ALAGAPPA UNIVERSITY

[Accredited with 'A+' Grade by NAAC (CGPA:3.64) in the Third Cycle and Graded as Category–I University by MHRD-UGC]

(A State University Established by the Government of Tamil Nadu)

KARAIKUDI – 630 003

Directorate of Distance Education

B.Sc. (Information Technology)

VI - Semester

129 63

MULTIMEDIA AND ITS APPLICATION

Authors:

Anirban Mukhopadhyay, Associate Professor, RCC Institute of Information Technology, Kolkata

Arup Chattopadhyay, Systems Manager & Head, Scientific & Technical Group, DOEACC Centre, Jadavpur University Campus, Kolkata

Units (1.0-1.2, 1.4-1.9, 2.3, 3.3, 4, 10, 11)

Rajneesh Agrawal, Senior Scientist, Department of IT, Government of India

Unit (1.3)

Vikas®Publishing House, Units (2.0-2.2, 2.4-2.8, 3.0-3.2, 3.4-3.9, 5-9, 12-14)

VIKAS®

Vikas® is the registered trademark of Vikas® Publishing House Pvt. Ltd.

VIKAS® PUBLISHING HOUSE PVT. LTD.
E-28, Sector-8, Noida - 201301 (UP)
Phone: 0120-4078900 • Fax: 0120-4078999
Regd. Office: A-27, 2nd Floor, Mohan Co-operative Industrial Estate, New Delhi 1100 44
• Website: www.vikaspublishing.com • Email: helpline@vikaspublishing.com

Work Order No.AU/DDE/DE12-27/Preparation and Printing of Course Materials/2020 Dated 12.08.2020 Copies - 500

SYLLABI-BOOK MAPPING TABLE

Multimedia and its Application

Syllabi	Mapping in Book
BLOCK I: FUNDAMENTAL CONCEPTS **Unit 1:** Fundamental Concepts in Text and Image: Multimedia and Hypermedia, World Wide Web, Overview of Multimedia Software Tools. **Unit 2:** Graphics and Image Data Representation Graphics/Image Data Types, File Formats. **Unit 3:** Color in Image and Video: Color Science, Color Models in Images, Color Models in Video.	**Unit 1:** Fundamental Concepts in Text and Image **(Pages 1-20);** **Unit 2:** Graphics and Image **(Pages 21-34);** **Unit 3:** Color in Image and Video **(Pages 35-44)**
BLOCK II: FUNDAMENTAL CONCEPTS IN VIDEO AND DIGITAL AUDIO **Unit 4:** Fundamental Concepts in Video and Digital Audio: Types of Video Signals, Analog Video, Digital Video. **Unit 5:** Digitization of Sound, MIDI. **Unit 6:** Quantization and Transmission of Audio.	**Unit 4:** Fundamental Concepts in Video and Digital Audio **(Pages 45-56);** **Unit 5:** Digitization of Sound **(Pages 57-70);** **Unit 6:** Quantization and Transmission of Audio **(Pages 71-78)**
BLOCK III: MULTIMEDIA DATA COMPRESSION **Unit 7:** Lossless Compression Algorithm: Run-Length Coding, Variable Length Coding, Dictionary Based Coding, Arithmetic Coding. **Unit 8:** Lossless Image Compression, Lossy Compression Algorithm: Quantization.	**Unit 7:** Lossless Compression Algorithm **(Pages 79-92);** **Unit 8:** Lossless Image Compression **(Pages 93-102)**
BLOCK IV: BASIC VIDEO COMPRESSION TECHNIQUES **Unit 9:** Introduction to Video Compression, Video Compression based on Motion Compensation. **Unit 10:** Search for Motion Vectors, MPEG. **Unit 11:** Basic Audio Compression Techniques.	**Unit 9:** Basic Video Compression Techniques **(Pages 103-114);** **Unit 10:** Search for Motion Vectors **(Pages 115-120);** **Unit 11:** Basic Audio Compression **(Pages 121-132)**
BLOCK V: MULTIMEDIA NETWORKS **Unit 12:** Multimedia Networks: Basics of Multimedia Networks, Multimedia Network. **Unit 13:** Communications and Applications: Quality of Multimedia Data Transmission, Multimedia over IP. **Unit 14:** Multimedia over ATM Networks, Transport of MPEG-4, Media on Demand (MoD).	**Unit 12:** Multimedia Networks **(Pages 133-144);** **Unit 13:** Communications and Applications **(Pages 145-158);** **Unit 14:** Multimedia Over ATM Networks **(Pages 159-168)**

CONTENTS

INTRODUCTION

NOTES

Multimedia is a form of communication that combines different content forms, such as text, audio, images, animations, or video into a single presentation, in contrast to traditional mass media, such as printed material or audio recordings. Popular examples of multimedia include video podcasts, audio slideshows and animated videos.

Multimedia can be recorded for playback on computers, laptops, smartphones, and other electronic devices, either on demand or in real time (streaming). In the early years of multimedia, the term 'Rich Media' was synonymous with interactive multimedia. Over time, hypermedia extensions brought multimedia to the World Wide Web (WWW). Fundamentally, the term 'Multimedia' means that computer information can be represented through audio, video, and animation in addition to traditional media (i.e., text, graphics drawings, and images). Hypermedia can be considered as one of the multimedia applications.

Multimedia is the field concerned with the computer-controlled integration of text, graphics, drawings, still and moving images (Video), animation, audio, and any other media where every type of information can be represented, stored, transmitted and processed digitally.

A ***Multimedia Application*** is an application which uses a collection of multiple media sources, for example text, graphics, images, sound/audio, animation and/or video.

In common usage, multimedia refers to an electronically delivered combination of media including video, still images, audio, and text in such a way that can be accessed interactively. Multimedia presentations may be viewed by person on stage, projected, transmitted, or played locally with a media player. A broadcast may be a live or recorded multimedia presentation. Broadcasts and recordings can be either analog or digital electronic media technology.

Multimedia finds its application in various areas including, but not limited to, advertisements, art, education, entertainment, engineering, medicine, mathematics, business, scientific research and spatial temporal applications.

This book, *Multimedia and Its Applications*, is divided into five blocks, which are further subdivided into fourteen units. This book provides a basic understanding of the subject and helps to grasp its fundamentals. In a nutshell, it explains various aspects, such as fundamental concepts in text and image, multimedia and hypermedia, graphics and image data representation, file formats, color in image and video, fundamental concepts in video and digital audio, digitization of sound, MIDI, quantization and transmission of audio, lossless compression algorithm, lossless image compression, Lossy compression algorithm, quantization, introduction to video compression, video compression based on motion compensation, search for motion vectors, MPEG, basic audio compression techniques, multimedia networks, communications and applications, quality of multimedia data transmission, multimedia over IP, multimedia over ATM networks, transport of MPEG-4, Media-on-Demand (MoD).

The book follows the Self-Instructional Mode (SIM) wherein each unit begins with an 'Introduction' to the topic. The 'Objectives' are then outlined before going on to the presentation of the detailed content in a simple and structured format. 'Check Your Progress' questions are provided at regular intervals to test the student's understanding of the subject. 'Answers to Check Your Progress Questions', a 'Summary', a list of 'Key Words', and a set of 'Self-Assessment Questions and Exercises' are provided at the end of each unit for effective recapitulation.

BLOCK - I
FUNDAMENTAL CONCEPTS

NOTES

UNIT 1 FUNDAMENTAL CONCEPTS IN TEXT AND IMAGE

Structure

1.0 INTRODUCTION

Hypermedia is an enhancement of hypertext, the non-sequential access of text documents, using a multimedia environment and providing users the flexibility to select which document they want to view next based on their current interests. Multimedia is defined as the integration of sound, animation, and digitized video with more traditional types of data such as text. The hypermedia model is fundamental to the structure of the World Wide Web, which is often based on a relational database organization. In this model, documents are interconnected as in a network, which facilitates extensive cross-referencing of related items. Users can browse effectively through the data by following links connecting associated topics or keywords. Object-oriented and hypermedia models are becoming routine for managing large multimedia systems such as digital libraries.

The HyperText Transfer Protocol (HTTP) is an application layer protocol for distributed, collaborative, hypermedia information systems. HTTP is the foundation of data communication for the World Wide Web, where hypertext documents include hyperlinks to other resources that the user can easily access, for example by a mouse click or by tapping the screen in a Web browser.

In this unit, you will study about the multimedia and hypermedia, overview of WWW or World Wide Web and its history, prime feature of multimedia in the field of business, functions of marquee, podcasting and its use, constituents of a Website and various multimedia software.

NOTES

1.1 OBJECTIVES

After going through this unit, you will be able to:

- Analyse the significance of multimedia and hypermedia
- Interpret the characteristics of a multimedia
- Discuss the advantages and limitations of multimedia and World Wide Web
- Understand the World Wide Web and its history
- Explain the prime feature of multimedia in the field of business
- Define the functions of Marquee
- Know about the Podcasting and its use
- Comprehend about the different constituents of a Website
- Recognise the various multimedia software and their description

1.2 MULTIMEDIA AND HYPERMEDIA

The technology of multimedia did not happen overnight. Desktop computers are changed into sophisticated systems since the late 1980s and it helps us to get our jobs done, deliver information and provide us entertainment. The evolution of multimedia technology has changed the way we look at computers. The first computers were seen as single purpose machines that solved incredibly complex mathematical problems. Mainframe computers were used to manage large corporate databases and financial systems during 1960s and 1970s which saw computer terminals throughout an organization being used for publishing and information management. The 1980s brought the desktop computer so everyone could have a computer at his or her desk for word processing, spreadsheets, and even for playing games. Bringing the computer to the individual in the office, the home, and the classroom meant looking at the computer as more than just a fancy typewriter or automated bookkeeper. In the mid 1980s and in 1990's, computer developers started looking at how computers could be used as never before. At the same time, advances in technology brought about faster desktop computers, increased working memory capacity in computers, higher data storage capacity in disk drives and CD-ROMs, digital audio and video, local and wide area networks that connected users to the world. Computer developers started looking to multimedia – the delivery of information using text, pictures, audio and video—as a way to utilize computers in a uniquely personal way. Multimedia computers could be used to increase efficiency and productivity on the job, provide information at our fingertips in the home, and help students learn more effectively both in and out of the classroom. These personal gains meant that people would see computers as practical and useful tools in their everyday lives. Since, the late 1980s, multimedia technology and applications have found many places in our lives. We now use multimedia at home where a wide variety of games and reference products, such as encyclopaedias as encyclopedias and cookbooks are put to use. At the office also where marketing presentations and training are essential multimedia are used for getting a new job done. At school multimedia becomes necessary and the interactive software programs assist students in learning mathematics, science, and new languages. In shopping malls we need multimedia where interactive computer terminals, called kiosks, help us to design greeting cards or to find out where specific stores are located. Thus we can see that impact of multimedia technology plays a key role on our daily lives.

1.2.1 Multimedia and Hypermedia

Multimedia is the integration of multiple forms of media. It means that computer information can be represented through audio, video and animation in addition to traditional media (such as text, graphics drawings and images). Multimedia includes text, graphics, audio, video, etc. For example, a presentation involving audio and video clips would be considered as 'multimedia presentation.' Educational software that involves animations, sound, and text is called 'multimedia software.' CDs and DVDs are often considered to be 'multimedia formats' since they can store a lot of data and most forms of multimedia require a lot of disk space. Due to the advancements in computer speeds and storage space, multimedia is commonplace today. Some examples of multimedia are digital video editing and production systems, electronic newspapers/magazines, games, groupware, home shopping, interactive TV, multimedia courseware, video conferencing, video-on-demand and virtual reality.

In a Web browser most Web navigation is done by clicking text-based links that open new pages in the Web browser. These links, which are often blue and underlined, are referred to as hypertext, since they allow the user to jump from page to page. Hypermedia is an extension of hypertext that allows images, movies, and flash animations to be linked to other content. The most common type of hypermedia is an image link. Hypermedia can be considered as one of the multimedia applications. Photos or graphics on the Web are often linked to other pages. For example, clicking a small 'thumbnail' image may open a larger version of the picture in a new window. Clicking a promotional graphic may direct you to an advertiser's Website. Flash animations and videos can also be turned into hyperlinks by embedding one or more links that appear during playback. You can tell if an image or video is a hyperlink by moving the cursor over it. If the cursor changes into a small hand, which means the image or video is linked to another page. Clicking the text, image, or video will open up a new location in your Web browser. Therefore, you should only click a hypertext or hypermedia link when you are ready to leave the current page. If you want to open the link in a new window, you can usually right click the link and select 'Open Link in New Window.'

MHEG or Multimedia and Hypermedia information coding Expert Group is the latest standard related to multimedia presentation. Just as there is a group for

NOTES

multimedia presentation in audio, video and text in an interactive way, known as Motion Picture Expert Group (MPEG), there is another group that describes interactive television services.

MHEG defines standards of information coding and is defined in ISO/IEC 13522. Subsequently, various revisions have been done to keep up with developments in multimedia. The latest version is known as MHEG-5, which was created in November 1994.

MHEG model provides a set of standard method; covering other standards, such as still picture format, **Joint Photographic Experts Group (**JPEG) and different standards of MPEG together to produce multimedia presentation. This provides a system independent presentation standard. This group has created standard set of methods for storage, exchange and display of multimedia presentations.

Objectives of MHEG

The following are the objectives of MHEG:

- To offer simple, easily implemental framework, using minimum system resources for multimedia applications.
- To define standard format in digital form for presentations which is interactive, and hardware and platform independent.
- To add features, such as extensibility, expandability and customizability by adding code specific to the application. This creates some dependency on platform, but it is desirable to maintain some kind of individualized specialty.

Multimedia and World Wide Web

The World Wide Web and multimedia are perhaps the two most common words and are related to each other. The World Wide Web is a global hypermedia system. Multimedia is a combination of text, image, graphic, audio, video, animation, etc., which we can now share over the Internet. The proliferation of multimedia on the World Wide Web has led to the introduction of Web search engines for images, video and audio. On the Web, multimedia is typically embedded within documents that provide a wealth of indexing information. The primary sources of information for indexing multimedia documents are text extracted from HTML pages and multimedia document headers. Off-line analysis of the content of multimedia documents can be successfully employed in Web search engines when combined with these other information sources. Although multimedia design and evaluation includes a lot more than the World Wide Web, it is important to remember the size and importance of the Web. In terms of the speed with which technology and innovations are moving and the potential it has to expand and reach a global audience, the World Wide Web is one of the driving forces behind much multimedia development. For this reason, it has to be considered as a special case within multimedia design.

NOTES

1.3 WORLD WIDE WEB

World Wide Web (WWW) represents the networking of Internetworking resources and collection of the Internet sites. WWW is known as Web started in computer world in year 1989 at CERN (European Council for Nuclear Research). Web is consisted of Web pages that are basically linked documents. These pages can be accessed across the network. When one page is linked to other, it is considered as hypertext. The word 'hypertext' was invented by Vannevar Bush in the year 1945. Getting a Web page is possible through Web browser. Many Web page editors are used to publish the Web pages, for example, Dreamweaver, ColdFusion, FrontPage, GoLive, etc. These editors are popularly used to design the Web page. Browser is the software that enables Web page to be viewed by clients. WWW or Web is a global information medium which users can read and write via computers connected to the Internet. Table 1.1 summarizes the history of WWW.

Table 1.1 *History of WWW*

Range of Years	**Development**
1980-1991: Development of the World Wide Web	The NeXTcube used by Tim Berners-Lee at CERN became the first Web server. In 1980, Tim Berners Lee, an independent contractor at the European Organization for Nuclear Research (CERN) Switzerland built ENQUIRE as a collection of personal database of people and software models but which deals hypertext and each new page of information in ENQUIRE had to be linked to an existing page. Berners Lee built all the tools necessary for a working Web, such as the HyperText Transfer Protocol or HTTP 0.9, the HyperText Markup Language or HTML (the first Web browser), the first HTTP server software, and the first Web pages that described the project itself. The browser could access Usenet newsgroups and File Transfer Protocol or FTP files as well. Later, Berners Lee posted a short summary of the World Wide Web project on the newsgroup. This date also marked the debut of the Web as a publicly available service on the Internet. The WWW project aims to allow all links to be made to any information anywhere.
1992–1995: Growth of the WWW	Early Web sites link for both the HTTP Web protocol and the then popular Gopher protocol which provided access to content through hypertext menus presented in the form of a file system rather than through HTML files. Some of the Web sites were also indexed by Wide Area Information Server or WAIS enabling users to submit full-text searches similar to the capability later provided by search engines. There was no graphical browser available for computers. In April 1992, Erwise was released and an application has been created by Pei-Yuan Wei which included advanced features, such as embedded graphics, scripting, and animation.
1. Early Browsers	In 1993, the Mosaic Web browser, a graphical browser developed by a team at the National Center for Supercomputing Applications or NCSA. In 1992, Andreessen and Eric Bina released an X Window browser in February 1993. It gained popularity due to its strong support of integrated multimedia and recommendations for new features. The first Microsoft Windows browser Cello was written by Thomas R. Bruce.
2. Web Organization	In September 1994, Berners Lee found the World Wide Web Consortium or W3C at the Massachusetts Institute of Technology with support from the Defense Advanced Research Projects Agency or DARPA and the European Commission. It comprised various companies that were willing to create standards and recommendations to improve the quality of the Web. Berners Lee made the Web available freely, i.e., with no patent and no royalties due.
1996–1998: Commercialization of the WWW	By 1996, WWW became the possibilities of free publishing and instant world wide information increasing familiarity with two-way communication over the 'Web' led to the possibility of direct Web based commerce, i.e., e-commerce.

1999–2001: Emerging Dot-com	Traditional media outlets (newspaper publishers, broadcasters and cable casters) also found the Web to be a useful and profitable additional channel for content distribution. A dot-com company is a company through which most of its business on the Internet is linked with Web site that uses the popular top level domain '.com' (commercial).
2002–Present: The Web Becomes Everywhere	Telecommunications companies had a great deal of over capacity as many Internet business clients went grow. During this time, a handful of companies found success developing business models that helped in making the World Wide Web or WWW. These include online air ticket booking sites, Google's search engine and Amazon.com online department store. The popularity of YouTube and similar services, combined with the increasing availability and affordability of high speed connections has made video content far more common on all kinds of Web sites. Many video content hosting and creation sites provide an easy means for their videos to be embedded on third party Web sites without payment or permission.

Web pages are addressed by Uniform Resource Locator or URL. For example, the URL is responded with ***abc*** as follows:

http://www.abc.com/abcspace

Here, ***http://*** is protocol that uses server to fetch the Uniform Resource Locator or URL and ***www.abc.com*** is fully qualified domain name and ***/abcspace*** shows the specified Web page that is located on the Web server. If user clicks on hyperlink to get the information, browsers (Internet Explorer , Netscape Navigator, Mozilla Firefox , etc.) fetch the pages which have been requested to. In fact, browser works as HTML interpreter. HTML is a markup language that is used to write the Web pages. In 1990's, a solid core of expertise were reached in the field of Internet business that leverages the Internet as a tool to achieve strategic business success. Business consulting includes the various features, such as e-business consulting, e-marketing consulting, Website usability, Internet trends, network traffic congestion and Search Engine Optimization or SEO over WWW. SEO is often considered the more technical part of Web marketing because it helps in the promotion of sites and at the same time it requires some technical knowledge about basic HTML. SEO is sometimes also called SEO copyrighting because most of the techniques that are used to promote sites in search engines deal with text. Generally, SEO can be defined as the activity of optimizing Web pages or whole sites in order to make them more search engine-friendly, thus getting higher positions in search results. One of the basic truths in SEO is that even if you do all the things that are necessary to do, this does not automatically guarantee you top ratings but if you neglect basic rules, this certainly will not go unnoticed. Search engines perform several activities in order to deliver search results, such as crawling, indexing, processing, calculating relevancy and retrieving. First, search engines use crawling process the Web to see what is there. This task is performed by piece of software, called a crawler or a spider or Googlebot, if you use Google. Spider software, follows links from one Web page to another and indexes everything they find on their way. Crawlers are not humans and they do not see images, Flash movies, JavaScript, frames, password protected pages and directories so for this, you need to run Spider Simulator because goodies are viewable by the spider. If they are not viewable, they will not be spidered and not indexed and not even

NOTES

processed, etc. The main application areas where WWW is attracting businesses are as follows:

- **In the Field of Publicity, Marketing and Advertising:** The WWW is a proper media to promote business. For this, the very first step is taken as to set up a site on the WWW so that people across the world can get the proper ideas about products and advertising cost. The site is an online field of publicity and advertising scheme. Business promotion schemes are announced in the field of business Website at regular interval so that consumers can get aware of the promotion schemes about the products and get a good rebate. Users and consumers can a glance of virtual world of shopping, such as a Window e-shopping which really facilitates a rough idea of products and cost. They can visit all the products in virtual malls and virtual showrooms and also get a catalogue and types of the materials via courtesy of the Web. For example, online shopping designer bags for ladies and toys for children sites are developed as multimedia presentation which gives an overall idea about the products. The categories are applied to examine the use of multimedia on the company's Web pages, for example text only, text and graphics, text and graphics and photographs, text and graphics or photographs with sound or video clips which make the Websites more effective in the field of business growth.
- **In the Field of Research and Development:** Many companies work in the field of research and development of business fields. WWW facilitates the additional resource of collecting information of the regarding fields. The management levels post a query, join a discussion group and receive the solution how to overcome a business problem. Millions of Web pages contain databases of information of the specific topics so that they can research and know the trends of the business growth and possible solution. The e-mail facility is used to send and receive the message working as digital equipment which has a very vast computers link up and exchanges about millions of e-mail messages in each month with the outside world from business community.
- **In the Field of Business Collaboration:** Once link is formed between companies, it is easy to communicate via the Internet. For example, a fruitful collaboration has been done between two famous companies named as IBM or International Business Machines and Bellcore by using Internet links to share a workstation. That is why WWW is known as 'business tools' due to its performance and significance. A data has been published in the Internet (in 1995, Yahoo) that there are over 20,000 business corporations working on Websites. The business community is being expanded day-by-day by the side of the information superhighway and business activity. From 1997, business commerce has been growing in another concept as B2C or Business to Consumer concept. One survey has been done as follows by 'The Economist', 1996 at that time,

The Great profit is not in consumer shopping but in business-to-business commerce, since most business transactions were already done at a distance whether by fax, telephone, post or private electronic links.

NOTES

Many business communities believe that the major problems for the future of electronic commerce are payment security and access time**.** Therefore, it is said about the interrelation between business community and WWW that:

'Better marketing and promotion of the Website drive prospective customers to the site.'

Business on WWW survey is done by various famous sites involve the increasing diversity of business sectors. The design of Websites should be become more interesting and attractive. In this survey, it was also noticed that the virtual shopping centers adopt this technology more quickly than individual businesses on the Web. Electronic mail is widely used at most of the sites observed and will be remaining an important tool for communication particularly between businesses and users. This survey clearly indicates that electronic commerce is still in its infancy but has great potential. The results obtained from the e-mail questionnaire made a useful supplement to the observational findings. The majority of companies are using their Websites for publicity and advertising purposes. Most companies look to the potential of WWW as a marketing tool.

Survey Done by Yahoo

Yahoo did the survey that has two parts as **observation of Websites** and **electronic mail survey** of sample companies. Yahoo Website chose three hundred companies made a survey on all selected companies and listed in Yahoo! Directory. It observed in a detail and systematic business activities. The company's Website classified according to the typology of Cockburn and Wilson is as follows:

- A Web presence with basic information about the company but no further details on specific products or services.
- A Web presence with company information and some information about products or services.
- A Web presence with company information and products or services information together with some price details but with facilities for conventional purchasing only.
- A Web presence with company information and products or services information with price details and the ability to order products or services via e-mail but with billing occurring conventionally.
- A Web presence with company information and products or services information including price details with the ability to cope with online ordering and payment.
- A Web presence with company information and products or services information including price details with pre-registration of credit card

NOTES

details by conventional means to gain account number which may be used to order goods online.

- A Web presence with company information and providing free products or services.

The growth of e-business community depends on better payment transaction and the confidence of consumers in the Web security. The average level of educational and computer literacy is also important otherwise the huge potential markets of the developing world will be particularly difficult to exploit. The rapid growth indicates the increased awareness amongst the business community of the business potential of the WWW. Between August 1995 and August 1997, a number of companies were registered with the Yahoo! Directory and these companies are predominantly related to real estate, business services and retailers. E-mail ordering refers to companies that allow customers to send the order via e-mail but the payment is still carried out using conventional methods, such as credit cards. Screenshot below displays the design view of a Web page of FUN CITY.

In the above screenshot, the three hyperlinks are connected by clicking on **Interests, PhotoAlbum** and **Favorites.** Once design view is completed, the Web page gives a view of home page connected via localhost or if possible via remote host. Pages are accessed by clicking the link. If you click on any of the three hyperlinks the linked hyperlink brings you to the corresponding pages. A hyperlink points to a whole document or to a specific element within a document. Hypertext is text with hyperlinks. A hyperlink has an anchor which represents the location within a document from which the hyperlink can be followed and the document containing a hyperlink is known as its source document. It is arranged in terms of a word, group of words or image that you can click on to jump to a new document or a new section within the current document.

WWW and Online Business

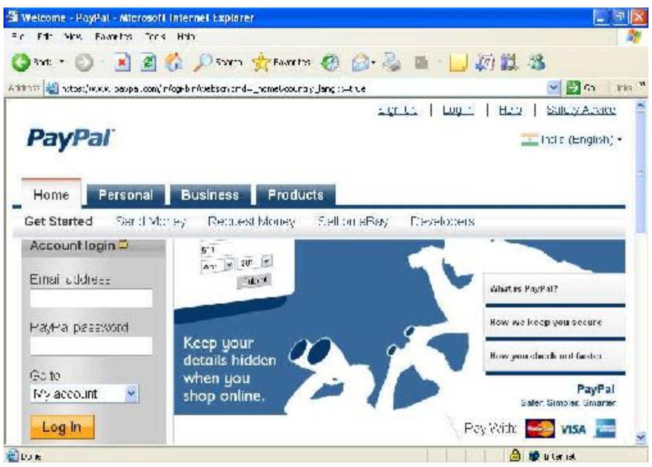
The business infrastructure community decides the platform, product, brand name, applications, data, and business rules to integrate with the cohesive system. It also supports business customization. Through program-to-program interpolation

NOTES

mechanisms, such as Web services, distributed object products, XML, etc. Business uses interaction kiosks for Internet business that lets you to focus on the core business in the field of private and public sector businesses. The private sector business refers to finance and investment banking, media, e-commerce and real estate whereas public business sector refers to e-commerce field. The prime feature of multimedia in the field of business is that it presents What You See Is What You Get or WYSIWYG environment for the produced material. It also includes the various supported image formats, animation effects, sound effects, transaction effects on the produced materials that are presented on the Web pages for e-business marketing. And, the base of interaction for produced materials, such item id, item name, item price, item background effect, Flash effects on produced images, sound effects, etc., on various pages is navigated through indexed hyperlink. Text effects for produced material create catching and effective marquee headlines. Marquee is used to make scrolling text or images to the Web page. The prime agenda behind business presentation is to include the multimedia authoring software. In fact, the software uses Compact Disks or CDs and videos on the Windows platform for business presentation that helps to decide the business infrastructure. Gaining profit in business does not impact on local basis but also on global business boundary. This technique is used for selling boundaries and also required for colossal step. For example, amazon.com and ebay.com were presented with entrepreneurs with a vital business tools and run across WWW and also still running in profit for online shopping and e-business domain. The amazon.com is enriched with shopping basket facility along with validation and consistency checking too. For example, if a customer is planning his vacation with an Internet travel agent service, the travel agent may be capable of checking on the arrival and departure times on hotel reservations and viability of flight connections. In fact, such operations can also be invoked each time when an object is added to the basket. Amazon.com, a pioneer hub of e-commerce, was launched by Jeff Bezos in 1994. And, eBay is an American Internet company that manages eBay.com. It is an online auction and shopping Website in which people and businesses buy and sell broad categories of produced items through online transaction. In these days, PaisaPay is also being used popularly and widely. PaisaPay is an eBay's online payment processing service where buyers can pay sellers through credit card, debit card and online bank transfer using PaisaPay. Basically, amazon.com and ebay.com optimize the search engine facilities for your Website that offer moderate list of options, searchable description and the third party data providers, such as Super Pages, Yellow Pages, City Searches, etc. They also offer a free service in which you can enter your Website domain area, tell the customer Who You Are (WAY) facility, get reviews, etc.

NOTES

The above screenshot shows the products which are provided online for the customers. It includes various products of books, toys, jewellery, sports, etc. In the field of business, multimedia plays a very important role to decide the business issues of podcasting, mobile marketing and other technologies. For growing business through multimedia technologies involve shopping cart software, online payment process that accepts credit card and PayPal account.



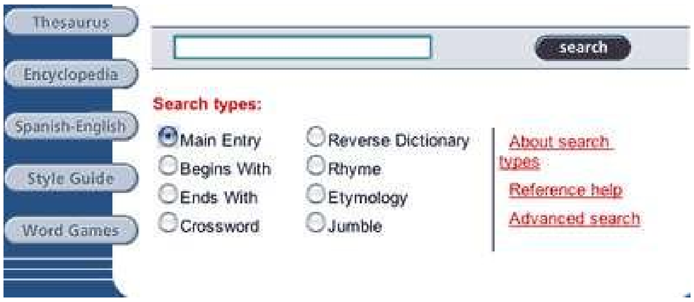
Podcasting is basically online audio content that is delivered via a Really Simple Syndication or RSS feed. It delivers audio content through portable media players. They are saved in Moving Picture Experts Group or MPEG Audio Layer III or MP3 format and listened through any system unit. The business field is enriched via podcasting in the network era. Podcasting is also considered through the new form of marketing and is able to:

- Target the audience and customers.
- Deliver the platform in low cost message.
- Send the multimedia messages to the customers in the form of audio and video means.

Basically, podcasting is tagged with the multimedia search engines to index the tagging technology correctly. Tagging is done in HTML coding and is used in RSS. Hence, tagging technology is considered as the heart of podcasting. RSS is an XML format that is used to define channels of information that contain elements which are typically stories or Web log entries. The video capable iPods support only two basic formats. They are certain types of H.264 video and certain types of MPEG4 video. Multimedia can often be an attractive investment for small businesses. Podcast client for the iPhone and iPod Touch is a unique iPhone app that lets you subscribe to your favorite podcasts, manage them, and listen to them in a unique and well designed format for the iPhone. Mobile channel gives access to a one-to-one, direct-to-consumer, personal contact with the customers. To take advantage of this, the brand managers are experimenting with mobile marketing by bombarding the customer with a lot of messages using all possible access channels. The customer receives numerous telesales call and unwanted service messages via out dialers, text or data.

World Wide Web (WWW) is considered a major reason behind the popularity of the Internet. WWW refers to a system of information and communications through which the users can access hypermedia information on servers. It is treasure trove of boundless information, in which all items have a reference through which they can be retrieved. In other words, any kind of information on any topic is readily available on the Internet. WWW comprises of a collection of Websites which are publicly accessible. A Website usually contains multiple pages which are replete with different types of information about different topics. Following are the constituents of a Website:

- **Home Page:** This page tells the visitors across the Internet about what is that. It also tells the visitors across net about what is that. The home page also locates the relevant sites on the Internet. It also provides detailed information about its service domain.



Screenshot above shows the Home Page of the site that provides various buttons to search the types. The Website depicted here is meant for online searching of thesaurus, encyclopedia, style guide, word games and Spanish-English dictionary.

The Web service provider helps establish a home page on the Internet. It is also known as electronic description of the organizations and its products and services. In real life, it is similar to a brochure or a catalog. An attractive home page grabs the attention of the visitors and that is why its cost per page is higher as compared to regular Web pages. A home page can be setup with the help of Integrated Services Digital Network (ISDN) line as it generates HTML which represents the graphical interface.

For example, if you type www.google.com, it brings you the Google home page which is illustrated in screenshot below.

- **About Page:** This page provides the visitors with a detailed information of the launched Website. For example, an online air ticket booking system gives a detailed description of its services for an air reservation domain.
- **Press Page:** This page on the site helps people interact with media. It also encourages the visitors to include the address, e-mail and phone/mobile numbers. It also publishes news on its Website on a daily basis. Before launching any site, it is essential to check whether the press page, contact page and other attractive features are available on the Website.
- **Contact Page:** The contact page provides the visitors with the city map, e-mail address, phone numbers and other relevant information, so that the visitors can personally contact the organization or the service providers.

1.4 OVERVIEW OF MULTIMEDIA SOFTWARE TOOLS

Table 1.2 lists software used in multimedia.

Table 1.2 *Various Multimedia Software and Their Description*

Package	Description
Audacity 1.2.3	A fast multitrack audio editor and recorder for Linux, MacOS, and Windows. Supports WAV, AIFF, Ogg, and MP3 formats. It is used for envelope editing, mixing, built-in effects and plug-ins, all with unlimited undo.
Ape2CD	Ape2CD can split .Ape file to CD track files based on .Cue file and then burn to CD. The operating systems for this software are Win95, Win98, WinME, Windows2000, WinXP, Windows2003, Wi 4.x.
Blender 2.37a	It is the software solution which is used for 3D modelling, animation, rendering and post-production, interactive creation and playback. Professionals and novices can easily and inexpensively publish stand-alone, secure, multi-platform content on the Web, CD-ROMs, and other media irrespective of whether they are users of Linux, Windows, Irix, FreeBSD, Sun Solaris or OSX.

NOTES

CD-DA X-Tractor 0.24	A CD ripper that features intelligent jitter correction, Internet CDDB support, and on-the-fly MP3 and Ogg Vorbis encoding.
CDex 1.51	A utility for extraction (ripping) of audio files from an audio CD.
Celestia 1.3.2	A real-time space simulation that allows you to experience our universe in three dimensions. Unlike most planetarium software, Celestia does not confine you to the surface of the Earth. You can travel throughout the solar system, to any of the over 100,000 stars, or even beyond the galaxy.
FLAC tools 1.1.2	Free Lossless Audio Codec (FLAC) is similar to MP3, but 'lossless', which means that audio is compressed in FLAC without throwing away any information. This is similar to the working of Zip, except that with FLAC you get much better compression. This is because it is designed specifically for audio, and you can play back compressed FLAC files in your player just like you would an MP3 file.
FlaskMPEG 0.78.39	A free, easy-to-use video conversion software which enables conversion of digital video in the MPEG format in to other formats.
Foobar2000 0.8.3	An advanced audio player for the Windows platform and includes some of the basic features such as replay gain support, low memory footprint and native support for several popular audio formats.
GNUMP3d 2.9.5	A small, stable, portable, and secure, self-contained, streaming server for MP3s, Ogg Vorbis files, movies and other media. Also, it can be easily installed and configured. It is portable across various types of Unix and Microsoft Windows platforms.
GOCR 0.39	An Optical Character Recognition (OCR) program, which was developed under the GNU Public License. It enables conversion of scanned images of text back to text files. GOCR can be used with different front-ends, which makes it very easy to port to different OSes and architectures. It can open many different image formats.
ImageMagick 6.2.4-5	A collection of tools and libraries for reading, writing and manipulating an image in numerous image formats (over 89 major formats). Some of the common formats include JPEG, TIFF, PNG, PhotoCD, PDF, and GIF. ImageMagick enables you to create images dynamically, thus making it suitable for Web applications. You can also resize, rotate, sharpen, reduce, colour or add special effects to an image or image sequence and save your completed work in the same or different image format. Image processing operations are available from the command line, or from the C, C++, Perl, Java, PHP, Python, or Ruby programming languages. A high-quality 2D providing a subset of SVG capabilities is also included. ImageMagick focusses on performance, minimizing bugs, and providing stable APIs and ABIs.
Inkscape 0.42.2	An open-source SVG editor which mimics Illustrator, CorelDraw, Visio and in its capabilities. Supported SVG features include basic shapes, paths, text, alpha blending, transforms, gradients, node editing, svg-to-png export, grouping, and more.
Jesterware DVD Audio Ripper	Jesterware DVD Audio Ripper is a powerful DVD-to-audio conversion software. Operating systems used for this software are Win98, WinME, Windows2000, WinXP, Windows2003, and Windows 4.x.
Media Player Classic 6.4.8.2	An enhanced version of Windows Media Player 6.
NotifyCD 1.60 beta 16	A very lightweight (but not in features) CD player for Windows (9x/ME/NT/2K/XP) which works completely from the system tray.
Oggenc 1.1.1	The standard Xiph oggenc with VBR quality setting and hard Min and Max bitrate limits.
OggdropXPd 1.8.6	A drag-and-drop Ogg Vorbis encoder/decoder/player for the experienced user. Features include compression from lossless files (Monkeys Audio, LPAC, FLAC and OptimFROG), auto-tagging, renaming of encoded files, setting of advanced encoder parameters, use of VorbisGain tags on decode, Play list (.m3u) creation, and others.
Project Mayo 0.5	The Playa was built for DivX users. Features such as progressive download, and the broken .avi file work when building 'The Playa'. Earlier, it took separate tools just for accomplishing this and there were no progressive download players which existed for .avi files on Windows.
Protected Music Converter	Protected Music Converter is a easy-to-use software for protected music conversion. This software works on the following operating systems: (a) WinNT 3.x, (b) WinNT 4.x, (c) WinXP, (d) Windows2000, (e) Windows2003, (f) Windows 4.x.
Sodipodi 0.34	A vector-based drawing program, such as CorelDraw or Adobe Illustrator from the proprietary software world, and Sketch or Karbon14 from the free software world. Sodipodi uses W3C SVG as its native file format. Hence, it is a very useful tool for web designers. It has a relatively modern display engine, giving you finely ant aliased display, alpha transparencies, vector fonts and so on.
Solfege 2.4.0	A computer program which helps to practise ear training. It can be useful when practising simple and mechanical exercises. The following are the exercises written so far: Recognize melodic and harmonic intervals, compare interval sizes, sing the intervals the computer asks for, identify chords, sing chords, dictation, scales, remembering rhythmic patterns.

Speex 1.0.5	A patent-free audio compression format which was designed for enabling speech. The Speex Project aims at lowering the barrier of entry for voice applications by providing a free alternative to expensive proprietary speech codec's. Also, it is well adjusted to the Internet applications and provides useful features that are not present in most other codec's. Speex is based on CELP and is designed to compress voice at bitrate ranging from 2 to 44 kbps.
VirtualDub 1.6.10	A video capturing/processing utility for 32-bit Windows platforms (95/98/ ME/NT4/ 2000/XP). It lacks the editing power of a general-purpose editor, such as Adobe Premiere, but is streamlined for fast linear operations over video. It has batch-processing capabilities thereby enabling processing of large numbers of files and can be extended with third-party video filters. VirtualDub is mainly geared toward processing AVI files, although it can read (not write) MPEG-1 and also handle sets of BMP images.
VLC 0.8.2	VLC (initially Video LAN Client) is a highly portable multimedia player for various audio and video formats (MPEG-1, MPEG-2, MPEG-4, DivX, mp3, etc.) as well as DVDs, VCDs, and various streaming protocols. It can also be used as a server to stream in unicast or multicast in IPv4 or IPv6 on a high-bandwidth network.
XviD 1.0.3	An ISO MPEG-4 compliant video codec. It is not a product; it is an open source project which is developed and maintained by many people from all over the world.
Zinf	A simple, but powerful audio player for Linux and Win32. It supports MP3, Ogg/Vorbis, WAV and Audio CD playback, with a powerful music browser, theme support and a download manager. It is based on the FreeA*p audio player which was developed by EMusic.com.

Check Your Progress

1. What was the purpose of main frame computers?
2. What is hypermedia?
3. List any two objectives of MHEG.
4. Define the term World Wide Web (WWW).
5. How are the Web pages addressed?
6. What is a hyperlink?
7. Define the term homepage.
8. What is the purpose of a contact page?

1.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Mainframe computers were used to manage large corporate databases and financial systems during 1960s and 1970s.
2. Hypermedia is an extension of hypertext that allows images, movies, and flash animations to be linked to other content.
3. The following are the objectives of MHEG:
   - To offer simple, easily implemental framework, using minimum system resources for multimedia applications.
   - To define standard format in digital form for presentations which is interactive and hardware and platform independent.
4. WWW or World Wide Web represents the networking of Internetworking resources and collection of Internet sites. WWW is known as Web started

NOTES

in computer world in year 1989 at CERN (European Council for Nuclear Research).

5. Web pages are addressed by Uniform Resource Locator or URL. For example, the URL is responded with abc as follows: http://www.abc.com/abcspace.
6. A hyperlink points to a whole document or to a specific element within a document.
7. The homepage locates the relevant sites on the Internet. It also provides detailed information about its service domain.
8. The contact page provides the visitors with the city map, email address, phone numbers and other relevant information, so that the visitors can personally contact the organization or the service providers.

1.6 SUMMARY

- The evolution of multimedia technology has changed the way we look at computers.
- Mainframe computers were used to manage large corporate databases and financial systems during 1960s and 1970s saw computer terminals throughout an organization being used for publishing and information management.
- Multimedia computers could be used to increase efficiency and productivity on the job, provide information at our fingertips in the home, and help students learn more effectively both in and out of the classroom.
- Multimedia is the integration of multiple forms of media. It means that computer information can be represented through audio, video and animation in addition to traditional media (such as text, graphics drawings and images).
- Multimedia includes text, graphics, audio, video, etc. For example, a presentation involving audio and
- video clips would be considered as multimedia presentation.
- Educational software that involves animations, sound, and text is called ‘multimedia software.
- In a Web browser most Web navigation is done by clicking text-based links that open new pages in the Web browser.
- Hypermedia is an extension of hypertext that allows images, movies, and flash animations to be linked to other content.
- Hypermedia can be considered as one of the multimedia applications. Photos or graphics on the Web are often linked to other pages.
- MHEG or Multimedia and Hypermedia information coding Expert Group is the latest standard related to multimedia presentation.

- MHEG model provides a set of standard method; covering other standards, such as still picture format, Joint Photographic Experts Group (JPEG) and different standards of MPEG together to produce multimedia presentation.
- The World Wide Web and multimedia are perhaps the two most common words and are related to each other. The World Wide Web is a global hypermedia system.
- Multimedia is a combination of text, image, graphic, audio, video, animation, etc., which we can now share over the Internet.
- The proliferation of multimedia on the
- World Wide Web has led to the introduction of Web search engines for images, video and audio. On the Web, multimedia is typically embedded within documents that provide a wealth of indexing information.
- World Wide Web (WWW) represents the networking of Internetworking resources and collection of Internet sites.
- WWW is known as Web started in computer world in year 1989 at CERN (European Council for Nuclear Research). Web is consisted of Web pages that are basically linked documents.
- The word 'hypertext' was invented by Vannevar Bush in the year 1945. Getting a Web page is possible through Web browser.
- The NeXT cube used by Tim Berners-Lee at CERN became the first Web server.
- E-mail ordering refers to companies that allow customers to send the order via e-mail, but the payment is still carried out using conventional methods, such as credit cards.
- A hyperlink has an anchor which represents the location within a document from which the hyperlink can be followed and the document containing a hyperlink is known as its source document.
- The prime feature of multimedia in the field of business is that it presents What You See Is What You Get or WYSIWYG environment for the produced material.
- Marquee is used to make scrolling text or images to the Web page. The prime agenda behind business presentation is to include the multimedia authoring software.
- Podcasting is basically online audio content that is delivered via a Really Simple Syndication or RSS feed.
- Podcast delivers audio content through portable media players. They are saved in Moving Picture Experts Group or MPEG Audio Layer III or MP3 format and listened through any system unit.
- Podcasting is tagged with the multimedia search engines to index the tagging technology correctly.

NOTES

- WWW refers to a system of information and communications through which the users can access hypermedia information on servers.
- The homepage locates the relevant sites on the Internet. It also provides detailed information about its service domain.
- The Web service provider helps establish a home page on the Internet. It is also known as electronic description of the organizations and its products and services.
- About page provides the visitors with a detailed information of the launched Website.
- Press page on the site helps people interact with media. It also encourages the visitors to include the address, e-mail and phone/mobile numbers. It also publishes news on its Website daily.
- The contact page provides the visitors with the city map, email address, phone numbers and other relevant information, so that the visitors can personally contact the organization or the service providers.

1.7 KEY WORDS

- **Multimedia computers:** They be used to increase efficiency and productivity on the job, provide information at our fingertips in the home, and help students learn more effectively both in and out of the classroom.
- **Multimedia:** It is the integration of multiple forms of media. It means that computer information can be represented through audio, video and animation in addition to traditional media.
- **Hypermedia:** It is an extension of hypertext that allows images, movies, and flash animations to be linked to other content.
- **MHEG:** Multimedia and Hypermedia information coding Expert Group is the latest standard related to multimedia presentation.
- **WWW:** WWW or World Wide Web represents the networking of Internetworking resources and collection of Internet sites.
- **HTML:** It is a markup language that is used to write the Web pages.
- **Podcasting:** It is basically online audio content that is delivered via a Really Simple Syndication or RSS feed. It delivers audio content through portable media players.
- **Homepage:** A home page (or homepage) is the main Web page of a Website. This page provides the visitors with a detailed information of the launched Website.
- **Press page:** This page on the site helps people interact with media. It also encourages the visitors to include the address, e-mail and phone/mobile numbers.

- **Contact page:** The contact page provides the visitors with the city map, email address, phone numbers and other relevant information, so that the visitors can personally contact the organization or the service providers.

NOTES

1.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What are multimedia computers?
2. Differentiate between the terms hyperlink and hypertext.
3. How can flash animations be turned into hyperlinks?
4. Define the term MHEG.
5. What is multimedia?
6. State the function of a spider software.
7. What is the use of electronic mail?
8. Define the term E-mail ordering.
9. What is the purpose of tagging?
10. How can a homepage be setup?
11. What is a about page?
12. State the significance of press page.

Long-Answer Questions

1. Briefly discuss why WWW is known as business tool with the help of an example.
2. Elaborate on the characteristic's features, advantages of multimedia giving suitable examples.
3. Discuss in detail about the Motion Picture Expert Group listing their objectives.
4. Explain the terms World Wide Web (WWW) and Web browser giving appropriate examples.
5. Give an overview on the history of WWW.
6. Explain the term Web page with the help of examples.
7. Elaborate the main application areas where WWW is attracting businesses with the help of appropriate examples.
8. Briefly explain the connection between WWW and online business.
9. Give an overview on the term podcasting along with its features.
10. Explain the term Website and the constituents of a Website.

NOTES

1.9 FURTHER READINGS

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

UNIT 2 GRAPHICS AND IMAGE

NOTES

Structure

2.0 INTRODUCTION

Multimedia is a form of communication that combines different content forms, such as text, audio, images, animations, or video into a single presentation, in contrast to traditional mass media, such as printed material or audio recordings. Popular examples of multimedia include video podcasts, audio slideshows and Animated videos. Multimedia can be recorded for playback on computers, laptops, smartphones, and other electronic devices, either on demand or in real time (streaming). In the early years of multimedia, the term rich media was synonymous with interactive multimedia. Over time, hypermedia extensions brought multimedia to the World Wide Web. Digital audio is a representation of sound recorded in, or converted into, digital form. In digital audio, the sound wave of the audio signal is typically encoded as numerical samples in a continuous sequence. A sample is a value or set of values at a point in time and/or space. In signal processing, sampling is the reduction of a continuous-time signal to a discrete-time signal. Data compression is subject to a space–time complexity trade-off.

The most common image file format used for windows is Device-Independent Bitmap (DIB) and is written as BMP files. This format can be used alone or may be masked with RIFF (Resource Interchange File Format). In fact, RIFF is the most preferred windows image format used since this format contains many file types including MIDI and formatted text.

In this unit, you will study about the graphics and images, data representation, image data types, image file formats, Web document images, file format specifications, Windows formats, cross platform formats.

2.1 OBJECTIVES

After going through this unit, you will be able to:

- Define data representation and image data types

- Understand the various Image file formats
- Analyse the significance of Web document images and their types
- Know about the file format specification
- Discuss about the Windows format
- Recognize the different types of cross platform formats

NOTES

2.2 DATA REPRESENTATION/IMAGE DATA TYPES

Multimedia is a form of communication that combines different content forms, such as text, audio, images, animations, or video into a single presentation, in contrast to traditional mass media, such as printed material or audio recordings. Popular examples of multimedia include video podcasts, audio slideshows and Animated videos. Multimedia can be recorded for playback on computers, laptops, smartphones, and other electronic devices, either on demand or in real time (streaming). In the early years of multimedia, the term rich media was synonymous with interactive multimedia. Over time, hypermedia extensions brought multimedia to the World Wide Web.

Digital Audio

Digital audio is a representation of sound recorded in, or converted into, digital form. In digital audio, the sound wave of the audio signal is typically encoded as numerical samples in a continuous sequence. For example, in CD audio, samples are taken 44,100 times per second, each with 16-bit sample depth. Digital audio is also the name for the entire technology of sound recording and reproduction using audio signals that have been encoded in digital form. Following significant advances in digital audio technology during the 1970s and 1980s, it gradually replaced analog audio technology in many areas of audio engineering and telecommunications in the 1990s and 2000s.

In a digital audio system, an analog electrical signal representing the sound is converted with an Analog-to-Digital Converter (ADC) into a digital signal, typically using Pulse-Code Modulation (PCM). This digital signal can then be recorded, edited, modified, and copied using computers, audio playback machines, and other digital tools. When the sound engineer wishes to listen to the recording on headphones or loudspeakers (or when a consumer wishes to listen to a digital sound file), a Digital-to-Analog Converter (DAC) performs the reverse process, converting a digital signal back into an analog signal, which is then sent through an audio power amplifier and ultimately to a loudspeaker. Digital audio systems may include compression, storage, processing, and transmission components. Conversion to a digital format allows convenient manipulation, storage, transmission, and retrieval of an audio signal. Unlike analog audio, in which making copies of a recording results in generation loss and degradation of signal quality, digital audio allows an infinite number of copies to be made without any degradation of signal quality.

Sampling

In signal processing, sampling is the reduction of a continuous-time signal to a discrete-time 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 is a value or set of values at a point in time and/or space. In signal processing, sampling is the reduction of a continuous-time signal to a discrete-time 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 is 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 original signal is retrievable from a sequence of samples, up to the Nyquist limit, by passing the sequence of samples through a type of low pass filter called a reconstruction filter.

Data Compression

In signal processing, data compression, source coding, or bit-rate reduction is the process of encoding information using fewer bits than the original representation. Any compression is either lossy or lossless. Lossless compression reduces bits by identifying and eliminating statistical redundancy. No information is lost in lossless compression. Lossy compression reduces bits by removing unnecessary or less important information. Typically, a device that performs data compression is referred to as an encoder, and one that performs the reversal of the process (decompression) as a decoder. The process of reducing the size of a data file is often referred to as data compression. In the context of data transmission, it is called source coding; encoding done at the source of the data before it is stored or transmitted. Source coding should not be confused with channel coding, for error detection and correction or line coding, the means for mapping data onto a signal.

Compression is useful because it reduces the resources required to store and transmit data. Computational resources are consumed in the compression and decompression processes. 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 (when using lossy data compression), and the computational resources required to compress and decompress the data.

NOTES

Lossless data compression algorithms usually exploit statistical redundancy to represent data without losing any information, so that the process is reversible. Lossless compression is possible because most real-world data exhibit statistical redundancy. For example, an image may have areas of color that do not change over several pixels; instead of coding "red pixel, red pixel, the data may be encoded as '279 red pixels'. This is a basic example of run-length encoding; there are many schemes to reduce file size by eliminating redundancy.

In the late 1980s, digital images became more common, and standards for lossless image compression emerged. In the early 1990s, lossy compression methods began to be widely used. In these schemes, some loss of information is accepted as dropping nonessential detail can save storage space. There is a corresponding trade-off between preserving information and reducing size. Lossy data compression schemes are designed by research on how people perceive the data in question. For example, the human eye is more sensitive to subtle variations in luminance than it is to the variations in color. JPEG image compression works in part by rounding off nonessential bits of information. Several popular compression formats exploit these perceptual differences, including psychoacoustics for sound, and psychovisual for images and video.

Image Data Types

Image file formats are standardized means of organizing and storing digital images. An image file format may store data in an uncompressed format, a compressed format (which may be lossless or lossy), or a vector format. Image files are composed of digital data in one of these formats so that the data can be rasterized for use on a computer display or printer. Rasterization converts the image data into a grid of pixels. Each pixel has several bits to designate its color (and in some formats, its transparency). Rasterizing an image file for a specific device considers the number of bits per pixel (the color depth) that the device is designed to handle.

The image data type is determined by the pixel depth in bits per channel, or bit depth. Bit depth for each channel can be 8, 16 or 32 and is included in the function name as one of these numbers (see Function Naming for details). The data may be signed (s), unsigned (u), or floating-point real (f). For some arithmetic and FFT/DFT functions, data in complex format (sc or fc) can be used, where each channel value is represented by two numbers: real and imaginary part. All channels in an image must have the same data type.

For example, in an absolute color 24-bit RGB image, three consecutive bytes (24 bits) per pixel represent the three channel intensities in the pixel mode. This data type is identified in function names as the 8u_C3 descriptor, where 8u represents 8-bit unsigned data for each channel and C3 represents three channels. Some functions operate with images in 16-bit packed RGB format (see RGB Image Formats in Image Color Conversion for more details). In this case data of all 3 channels are represented as 16u data type. For example, in an absolute color 16-bit packed RGB image, two consecutive bytes (16 bits) per pixel represent the three channel intensities in the pixel mode. This data type is identified in function names as the 16u_C3 descriptor, where 16u represents 16-bit unsigned data (not a bit depth) for all packed channels together and C3 stands for three channels.

For Macintosh platform PCs, PICT is a file format used for Macintosh graphics. It was developed by Apple Computers in 1984.

The PICT files also contain bitmaps of line-art, grayscale or RGB data.

2.3 IMAGE FILE FORMATS

Utilities, such as Windows Paintbrush or Paint Shop Pro, etc., generate *bitmapped* image files in BMP format or in the more efficient PCX format.

Static bitmapped images are often compressed to reduce the file sizes and thus, to save some disk space and shorten the time it takes to transfer those files over a communication link. The most common compressed file formats are:

- Graphics Image Format (*.GIF).
- Tagged Information File Format TIFF (*.TIF).
- Joint Photographic Experts Group JPEG (*.JPG).
- Windows Bitmap (*.BMP) and Windows Device Independent Bitmap (*.DIB)

GIF, TIF, DIB and PCX files are compressed in *lossless* fashion using either *RLE*[*] or *LZW*[†] compression algorithm. That is, only truly redundant bits are squeezed out, and they all can be returned exactly as they were when the file is decompressed. None of the original images data is deleted in the compression process.

JPEG, on the other hand, is an example of a lossy compression. Data from the original image which is deemed to be redundant is thrown away in the compression process. (And as you would expect lossy compression yields smaller compressed files than does lossless compression.) This means that the image resulting from the decompressed files will differ from the originals to some degree. The tricky part of these algorithms is their attempt to lose only 'unimportant' features of the images and people are least likely to notice those absent features while viewing the reconstructed image.

Researches are on for evolving more effective compression techniques. The objective is to reduce the compression as well as decompression cost and time without any significant degradation in image quality but at the same time enhancing the storage savings.

Graphics Image File Format

There are several graphic image file formats that are used by most of the graphics systems. The following are the most regularly used formats:

NOTES

Web Document Images

Web document images can be of two types as follows:

(a) ***Graphics Interchange Format (GIF):*** Images made using this format use a fixed colour palette which is limited to only 256 colours. This format downloads small, compressed files quickly from the Web. This format is most suitable for images with solid colours or uniform colour areas, such as illustrations and logos.

(b) ***Joint Photographic Experts Group (JPEG):*** These files are used for photographic (continuous colour tone) images, i.e., those images that have a continuous colour tone. Unlike the Graphics Interchange Format files, the Joint Photographic Experts Group format takes advantage of the full spectrum of colours available to the display unit. The JPEG format also uses compression for making smaller files and for obtaining faster downloads over the World Wide Web. However, unlike the compression method used in GIF files, the JPEG compression is also 'lossy compression' which means it discards some data in the decompression process. Once a file is saved in the JPEG format some data is lost permanently. But this does not affect the image.

Printed Documents

The following are the two types of printed documents.

(a) ***Encapsulated PostScript (EPS):*** It is an image file format used for both vector graphics and bitmaps. EPS files have a PostScript description of the graphic data within them. The EPS files are exclusive in that the graphics users use them for bitmap images, vector graphics, type or even entire pages.

(b) ***Tagged-Image File Format (TIFF):*** Such files are used for bitmap format only. The **TIFF** formats are the files that are supported by virtually all graphics applications.

Macintosh Formats

For Macintosh platform PCs, PICT is a file format used for Macintosh graphics. It was developed by Apple Computers in 1984. Both bitmap and vector-drawn objects can be drawn using PICT file format. It is a meta-format and can be used to make Revolution or Canvas. These file formats are used to exchange graphics between various Macintosh applications. Macintosh drawing programs, such as Illustrator or Freehand, also allow importing of bit-map, but there is no facility for editing a file.

File formats specifications

There are some specific PICT file formats:

- PICT images can be stored as a Macintosh resource of type PICT or as a data file of type 'PICT'.

NOTES

- All the QuickDraw commands such Font: size, style, type; Lines, Circles, Bitmaps, and so on, can be used to draw the image on the Macintosh screen in the Object Oriented file
- The PICT files also contain bitmaps of line-art, grayscale or RGB data
- These files are also supported by Windows platform PCs using QuickTime for Windows that can contain only one bitmap file

There are two different versions of the PICT formats:

- **PICT 1 Formats:** This is an old format which only allows 8 colours, but no modern colour primitive. This was the original version of Macs developed for black- and-white.
- **PICT 2 Formats:** This one is the new format which can store bitmap type objects. This format includes regions, lines, colour settings, Ovals and other primitives. The file format contains black-and white, 4-bit, 8-bit, 16-bit and 24-bit colour bitmap objects.

Windows Formats

The most common image file format used for windows is Device-Independent Bitmap (DIB) and is written as BMP files. This format can be used alone or may be masked with RIFF (Resource Interchange File Format). In fact, RIFF is the most preferred windows image format used since this format contains many file types including MIDI and formatted text.

Windows often support DIB, BMP, PCX and TIFF bitmap formats. The window bitmap file has a BMP format. PCX files are used in MS-DOS paint and desktop publishing software. Tagged Interchange File Format (TIFF) is used for archiving important images and is considered as a standard format in printing and publishing industry.

The bitmap formats often used are:

- Raster image file types and their formats

.bmp	Bitmap Image File
.gif	Graphical Interchange Format File
.jpg	JPEG Image File
.png	Portable Network Graphic
.psd	Adobe Photoshop Document
.pspimage	PaintShop Pro Image
.thm	Thumbnail Image File
.tif	Tagged Image File
.yuv	YUV Encoded Image File

Fig 2.1 *Raster Image File Types*

- Vector image file types and their formats

.ai	Adobe Illustrator File
.drw	Drawing File
.eps	Encapsulated PostScript File
.ps	PostScript File
.svg	Scalable Vector Graphics File

Fig 2.2 *Vector Image File Types*

- 3D Image

.3dm	Rhino 3D Model
.max	3ds Max Scene File

Fig 2.3 *3 D Image File Types*

Some of the commonly used windows image formats are discussed in the following paragraphs:

BMP is the window standard image format for DOS and other Windows compatible PCs. The modes of colour supported by BMP are RGB, indexed-colour, grayscale and bitmap, but it does not support alpha channels.

The JPEG format is the most common for displaying photographs and continuous-tone images in HTML for online services, such as WWW. The modes that JPEG supports are CMYK, RGB and grayscale. However, it does not support alpha channels. The colour information in an RGB image is retained by this format, but it compresses file size by discarding some selective data. The file format automatically decompresses an image when opened. A lower-level compression of an image displays a better quality of image in this file format, whereas a higher quality image displays a lower quality of image.

GIF is also used to display images and indexed-colour graphics formats used in HTML for WWW. It is basically used to reduce the size of the file and hence reduce the file transfer time electronically due to its LZW-compressed format design.

PNG file format is used for lossless compression to display images on WWW. The PNG format supports RGB and greyscale with single alpha channel and indexed-colour and Bitmap without alpha channels and supports 24-bit images and is useful to create transparent background with no jagged edges. However, older versions of web browsers do not support this file format.

Cross Platform Formats

The major concern for multimedia developers is file compatibility. It is important to ensure the compatibility of various file formats across various platforms. The compatible issue of file format is to use the formats intended for the platforms. For example, a developer needs to own a separate copy of the file for each platform even if the files used in Word or Photoshop are available for both the platforms, viz. Windows and MAC. The cross-platform compatibility challenge is whether a file is compatible with each platform. For example, a TIFF file format must be used by both the platforms, but a bitmap image used is developed for Microsoft whereas PICT image format is developed by Apple for MAC, and the two cannot be used interchangeably.

Another issue is to make sure that different application programs may be able to process the format on a given platform. Since, media developers use tools and media specific software, to create or edit the elements used in the multimedia applications, thus it is important to ensure that the file formats used will be compatible with the media specific and authoring tools used. For example, some latest file formats, such as PNG were not supported by most-editing programs and took some time to become compatible with the programs.

The two most commonly used bitmap formats, JPEG and GIF, are considered as cross-platforms formats used on Web and are displayed by almost all the browsers. The most common format used to deliver the product that contains various types of assets is Adobe's PDF (Portable Document File) which is capable of managing bitmap and drawn art including text and other elements used in multimedia. It is a good example of a cross-platform format application.

Check Your Progress

1. What is multimedia?
2. Define the term digital audio.
3. State the significance of sampling.
4. What does image file format means?
5. Name any one common compressed file formats.
6. Define the term Encapsulated PostScript (EPS).
7. Why is the PNG file format used?
8. List the two most used bitmap formats.

2.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Multimedia is a form of communication that combines different content forms, such as text, audio, images, animations, or video into a single presentation, in contrast to traditional mass media, such as printed material or audio recordings.
2. Digital audio is a representation of sound recorded in, or converted into, digital form. In digital audio, the sound wave of the audio signal is typically encoded as numerical samples in a continuous sequence.

NOTES

3. In signal processing, sampling is the reduction of a continuous-time signal to a discrete-time signal.
4. Image file formats are standardized means of organizing and storing digital images. An image file format may store data in an uncompressed format, a compressed format (which may be lossless or lossy), or a vector format.
5. The common compressed file formats are Graphics Image Format (*.GIF).
6. It is an image file format used for both vector graphics and bitmaps. EPS files have a PostScript description of the graphic data within them. The EPS files are exclusive in that the graphics users use them for bitmap images, vector graphics, type, or even entire pages.
7. PNG file format is used for lossless compression to display images on WWW. The PNG format supports RGB and greyscale with single alpha channel and indexed-colour and Bitmap without alpha channels and supports 24-bit images and is useful to create transparent background with no jagged edges.
8. The two most used bitmap formats, JPEG and GIF, are considered as cross-platforms formats used on Web and are displayed by almost all the browsers.

2.5 SUMMARY

- Multimedia is a form of communication that combines different content forms, such as text, audio, images, animations, or video into a single presentation, in contrast to traditional mass media, such as printed material or audio recordings.
- Multimedia can be recorded for playback on computers, laptops, smartphones, and other electronic devices, either on demand or in real time (streaming).
- Digital audio is a representation of sound recorded in, or converted into, digital form. In digital audio, the sound wave of the audio signal is typically encoded as numerical samples in a continuous sequence.
- Digital audio is a representation of sound recorded in, or converted into, digital form. In digital audio, the sound wave of the audio signal is typically encoded as numerical samples in a continuous sequence.
- Digital audio systems may include compression, storage, processing, and transmission components.
- Conversion to a digital format allows convenient manipulation, storage, transmission, and retrieval of an audio signal.
- In signal processing, sampling is the reduction of a continuous-time signal to a discrete-time signal.

NOTES

- A sampler is a subsystem or operation that extracts samples from a continuous signal.
- In signal processing, data compression, source coding, or bit-rate reduction is the process of encoding information using fewer bits than the original representation.
- Any compression is either lossy or lossless. Lossless compression reduces bits by identifying and eliminating statistical redundancy.
- Compression is useful because it reduces the resources required to store and transmit data. Computational resources are consumed in the compression and decompression processes.
- Data compression is subject to a space–time complexity trade-off.
- Lossless data compression algorithms usually exploit statistical redundancy to represent data without losing any information, so that the process is reversible.
- Lossless compression is possible because most real-world data exhibit statistical redundancy.
- Image file formats are standardized means of organizing and storing digital images.
- Rasterization converts the image data into a grid of pixels. Each pixel has several bits to designate its color (and in some formats, its transparency).
- Rasterizing an image file for a specific device considers the number of bits per pixel (the color depth) that the device is designed to handle.
- The image data type is determined by the pixel depth in bits per channel, or bit depth.
- Static bitmapped images are often compressed to reduce the file sizes and thus, to save some disk space and shorten the time it takes to transfer those files over a communication link.
- JPEG, on the other hand, is an example of a lossy compression. Data from the original image which is deemed to be redundant is thrown away in the compression process.
- The JPEG format also uses compression for making smaller files and for obtaining faster downloads over the World Wide Web.
- PICT images can be stored as a Macintosh resource of type PICT or as a data file of type 'PICT'.
- The most common image file format used for windows is Device-Independent Bitmap (DIB) and is written as BMP files.
- Windows often support DIB, BMP, PCX and TIFF bitmap formats. The window bitmap file has a BMP format.

NOTES

- Tagged Interchange File Format (TIFF) is used for archiving important images and is considered as a standard format in printing and publishing industry.
- The JPEG format is the most common for displaying photographs and continuous-tone images in HTML for online services such as WWW.
- GIF is also used to display images and indexed-colour graphics formats used in HTML for WWW.
- PNG file format is used for lossless compression to display images on WWW.
- The PNG format supports RGB and greyscale with single alpha channel and indexed-colour and Bitmap without alpha channels and supports 24-bit images and is useful to create transparent background with no jagged edges.
- The two most used bitmap formats, JPEG and GIF, are considered as cross-platforms formats used on Web and are displayed by almost all the browsers.

2.6 KEY WORDS

- **Multimedia:** It is a form of communication that combines different content forms, such as text, audio, images, animations, or video into a single presentation, in contrast to traditional mass media, such as printed material or audio recordings.
- **Digital audio:** It is a representation of sound recorded in, or converted into, digital form.
- **Graphics Interchange Format (GIF):** Images made using this format use a fixed colour palette which is limited to only 256 colours. This format downloads small, compressed files quickly from the Web.
- **Joint Photographic Experts Group (JPEG):** These files are used for photographic (continuous colour tone) images, i.e., those images that have a continuous colour tone.
- **Encapsulated PostScript (EPS):** It is an image file format used for both vector graphics and bitmaps. EPS files have a PostScript description of the graphic data within them.
- **Tagged-Image File Format (TIFF):** The TIFF formats are the files that are supported by virtually all graphics applications.
- **PICT 1 Formats:** This is an old format which only allows 8 colours, but no modern colour primitive.
- **PICT 2 Formats:** This one is the new format which can store bitmap type objects. This format includes regions, lines, colour settings, ovals, and other primitives.

- **Bitmap Image File (BMP):** It is the window standard image format for DOS and other Windows compatible PCs.
- **Joint Photographic Experts Group (JPEG):** This format is the most common for displaying photographs and continuous-tone images in HTML for online services such as WWW.
- **Graphics Interchange Format (GIF):** It is used to display images and indexed-colour graphics formats used in HTML for WWW.
- **Portable Network Graphics (PNG):** This file format is used for lossless compression to display images on WWW.

2.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is a data representation?
2. Define the characteristics of an image file format utilities.
3. What are the various types of Web document images?
4. Define the terms Macintosh.
5. What is the significance of Encapsulated PostScript??
6. What are various PICT file formats?
7. What is the PNG file format used for?
8. What are the two most used bitmap formats?
9. State the full form of BMP and PCX.
10. Classify the various types of printed documents.

Long-Answer Questions

1. Briefly discuss the significance of data representation giving their features and examples.
2. Discuss in detail about the various types of image data giving their features and examples.
3. Explain the significance and different types of image file format.
4. Describe the Web document images and their types.
5. What are printed documents? Why is it required? Discuss in brief.
6. Briefly discuss about cross platform formats and its compatibility challenge.

2.9 FURTHER READINGS

NOTES

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

UNIT 3 COLOR IN IMAGE AND VIDEO

NOTES

Structure

3.0 INTRODUCTION

Color is nothing more than how a person perceives the light waves that enters their eye. The scientific definition of color is the quality of an object or substance with respect to light reflected by the object, usually determined visually by measurement of hue, saturation, and brightness of the reflected light. Color categories and physical specifications of color are associated with objects through the wavelengths of the light that is reflected from them and their intensities. This reflection is governed by the object's physical properties such as light absorption, emission spectra, etc. The science of color is sometimes called chromatics, colorimetry, or simply color science. It includes the study of the perception of color by the human eye and brain, the origin of color in materials, color theory in art, and the physics of electromagnetic radiation in the visible range (that is, what is commonly referred to simply as light).

In this unit, you will study about the basic concepts of color science or chromatics, visible lights, color models in videos, additive color, vector graphs.

3.1 OBJECTIVES

After going through this unit, you will be able to:

- Analyse the significance color science
- Understand the color models in images
- Know about spectral colors and color reproduction
- Comprehend about the different types of color models in video
- Recognise the various types of coloring

3.2 COLOR SCIENCE

NOTES

The science of color is sometimes called chromatics, colorimetry, or simply color science. It includes the study of the perception of color by the human eye and brain, the origin of color in materials, color theory in art, and the physics of electromagnetic radiation in the visible range (that is, what is commonly referred to simply as light). Color categories and physical specifications of color are associated with objects through the wavelengths of the light that is reflected from them and their intensities. This reflection is governed by the object's physical properties, such as light absorption, emission spectra, etc.

Electromagnetic radiation is characterized 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 390 nm to 700 nm), it is known as "visible light". Most light sources emit light at many different wavelengths; a source's spectrum is a distribution giving its intensity at each wavelength. Although the spectrum of light arriving at the eye from a given direction determines the color sensation in that direction, there are many more possible spectral combinations than color sensations.

The color of an object depends on both the physics of the object in its environment and the characteristics of the perceiving eye and brain. Physically, objects can be said to have the color of the light leaving their surfaces, which normally depends on the spectrum of the incident illumination and the reflectance properties of the surface, as well as potentially on the angles of illumination and viewing. Some objects not only reflect light, but also transmit light or emit light themselves, which also contributes to the color. A viewer's perception of the object's color depends not only on the spectrum of the light leaving its surface, but also on a host of contextual cues, so that color differences between objects can be discerned mostly independent of the lighting spectrum, viewing angle, etc. This effect is known as color constancy.

The upper disk and the lower disk have the same objective color and are in identical gray surroundings; based on context differences, humans perceive the squares as having different reflectance's, and may interpret the colors as different color categories; see checker shadow illusion.

Some generalizations of the physics can be drawn, neglecting perceptual effects for now:

- Light arriving at an opaque surface is either reflected 'specularly', i.e., in the manner of a mirror, scattered, i.e., reflected with diffuse scattering, or absorbed—or some combination of these.
- Opaque objects that do not reflect specularly which tend to have rough surfaces have their color determined by which wavelengths of light they scatter strongly with the light that is not scattered being absorbed. If

NOTES

objects scatter all wavelengths with roughly equal strength, they appear white. If they absorb all wavelengths, they appear black.

- Opaque objects that specularly reflect light of different wavelengths with different efficiencies look like mirrors tinted with colors determined by those differences. An object that reflects some fraction of impinging light and absorbs the rest may look black but also be faintly reflective; examples are black objects coated with layers of enamel or lacquer.

3.3 COLOR MODELS IN IMAGES

Images can be generated and stored in a personal computer in two typically different ways. One is called vector graphics and the other is referred to as bitmapped. A piece of vector art is a file that contains descriptions of how to generate the image but not the actual image itself. A vector graphics program (generically called drawing program) creates a sequential list of graphic commands to draw lines, curves, text, etc., with associated parameters, such as screen location, size, color, rotation angle, width, style, etc. This type of list-file is often referred to as a display list/file. Such a file must be rasterized before it can be presented as an actual image on screen.

While a vector graphic is edited, the properties of the lines and curves, which explain its shape, are changed. Without altering the quality of its form, one can reshape, resize, move and modify the color of a vector graphic. Being resolution-independent, vector graphics may be displayed on output devices of varying resolutions without losing any quality.

A *bitmapped image*, in contrast, has actual pixel image data in the file. That is, it simply holds the color number for each dot or pixel in an image. The size of such files depends on the image size and how many colors are to be used per pixel, i.e., the color depth. For a 256-color range and standard VGA (640 × 480 pixels) full screen display the size of a bitmap file is 640 × 480 × 8 bits, i.e., 307200 × 8 bits or 307200 bytes or 300 KB. This is because for each pixel 8 bits are required for storing color value anything up to 256. However, true-color images (24-bit, more than 16 million color) provide the highest quality, and they are the best way to represent photographs on computer screen. Out of the 24 bits per pixel in high quality true-color images, 8 bits are used to describe intensities of each of the three basic color signals—Red, Green and Blue of *RGB color model*. White color is displayed when all RGB signals are at full intensity, and black occurs when there is no signal.

As soon as a bitmap graphic is edited, the pixels are altered but not the lines and curves. The former are resolution-dependent as the data unfolding the image is set to a grid of a particular size. However, editing a bitmap graphic alters the quality of its appearance. For instance, resizing a bitmap graphic make the edges of the image ragged because pixels are redistributed within the grid. Showing a bitmap graphic on an output device that has a lower-resolution than the image also degrades the quality of its appearance.

NOTES

***Left*:** On the left, the vector image of a leaf is shown by points through which lines and curves pass making the shape of the outline of the leaf. The color of the leaf is determined by the color of the outline and the area enclosed by the outline.

***Right*:** The bitmapped-image of a leaf is described by the specific location and color value of each pixel, creating a more realistic image.

Utilities, such as Windows Paintbrush or Paint Shop Pro generate *bitmapped* image files in BMP format or in the more efficient PCX format.

Static bitmapped images are often compressed to reduce the file sizes and thus to save some disk space and shorten the time it takes to transfer those files over a communication link. The most common compressed file formats are:

- Graphics Image Format (*.GIF).
- Tagged Information File Format or TIFF (*.TIF).
- Joint Photographic Experts Group or JPEG (*.JPG).
- Windows Bitmap (*.BMP) and Windows Device Independent Bitmap (*.DIB).

3.4 COLOR MODELS IN VIDEO

Two common color models in video are YUV and YIQ. YUV uses properties of the human eye to prioritize information. Y is the black and white (luminance) image, U and V are the color difference (chrominance) images.

YUV

YUV is a color encoding system typically used as part of a color image pipeline. It encodes a color image or video taking human perception into account, allowing reduced bandwidth for chrominance components, thereby typically enabling transmission errors or compression artifacts to be more efficiently masked by the human perception than using a 'direct' RGB-representation. Other color encodings have similar properties, and the main reason to implement or investigate properties

NOTES

of Y2 UV would be for interfacing with analog or digital television or photographic equipment that conforms to certain Y2 UV standards. The Y2 UV model defines a color space in terms of one luma component (Y2) and two chrominance components, called U (blue projection) and V (red projection) respectively. The Y2 UV color model is used in the PAL composite color video (excluding PAL-N) standard.

Y2 stands for the luma component (the brightness) and U and V are the chrominance (color) components; luminance is denoted by Y and luma by Y2 – the prime symbols (‘) denote gamma correction, with “luminance” meaning physical linear-space brightness, while “luma” is (nonlinear) perceptual brightness. Y2 UV was invented when engineers wanted color television in a black-and-white infrastructure. They needed a signal transmission method that was compatible with black-and-white (B&W) TV while being able to add color. The luma component already existed as the black and white signal; they added the UV signal to this as a solution.

YIQ

YIQ is the color space used by the NTSC color TV system, employed mainly in North and Central America, and Japan. I stand for in-phase, while Q stands for quadrature, referring to the components used in quadrature amplitude modulation. Some forms of NTSC now use the YUV color space, which is also used by other systems, such as PAL. The Y component represents the luma information, and is the only component used by black-and-white television receivers. I and Q represent the chrominance information. In YUV, the U and V components can be thought of as X and Y coordinates within the color space. I and Q can be thought of as a second pair of axes on the same graph, rotated 33°; therefore, IQ and UV represent different coordinate systems on the same plane.

The YIQ representation is sometimes employed in color image processing transformations. For example, applying a histogram equalization directly to the channels in an RGB image would alter the color balance of the image. Instead, the histogram equalization is applied to the Y channel of the YIQ or YUV representation of the image, which only normalizes the brightness levels of the image.

Check Your Progress

1. What is called chromatics or color science?
2. What is known as visible light?
3. List any two common color models in video.
4. Define the term additive color.
5. How does the light arriving at an opaque surface reflect?
6. What does a vector graphics program do?

NOTES

3.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. The science of color is sometimes called chromatics, colorimetry, or simply color science.
2. Electromagnetic radiation is characterized 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 390 nm to 700 nm), it is known as visible light.
3. Two common color models in video are YUV and YIQ. YUV uses properties of the human eye to prioritize information. Y is the black and white (luminance) image, U and V are the color difference (chrominance) images.
4. Additive color is light created by mixing light of two or more different colors. Red, green, and blue are the additive primary colors normally used in additive color systems such as projectors and computer terminals.
5. Light arriving at an opaque surface is either reflected specularly, i.e., in the manner of a mirror, scattered, i.e., reflected with diffuse scattering), or absorbed—or some combination of these.
6. A vector graphics program (generically called drawing program) creates a sequential list of graphic commands to draw lines, curves, text, etc., with associated parameters, such as screen location, size, color, rotation angle, width, style, etc.

3.6 SUMMARY

- The science of color is sometimes called chromatics, colorimetry, or simply color science.
- Color science includes the study of the perception of color by the human eye and brain, the origin of color in materials, color theory in art and the physics of electromagnetic radiation in the visible range i.e., what is commonly referred to simply as light.
- Color categories and physical specifications of color are associated with objects through the wavelengths of the light that is reflected from them and their intensities.
- Electromagnetic radiation is characterized by its wavelength or frequency and its intensity.

NOTES

- The color of an object depends on both the physics of the object in its environment and the characteristics of the perceiving eye and brain.
- The upper disk and the lower disk have the same objective color and are in identical gray surroundings.
- Light arriving at an opaque surface is either reflected specularly i.e., in the manner of a mirror, scattered i.e., reflected with diffuse scattering or absorbed or some combination of these.
- Opaque objects that do not reflect specularly which tend to have rough surfaces have their color determined by which wavelengths of light they scatter strongly with the light that is not scattered being absorbed.
- Opaque objects that specularly reflect light of different wavelengths with different efficiencies look like mirrors tinted with colors determined by those differences.
- Images can be generated and stored in a personal computer in two typically different ways.
- One is called vector graphics and the other is referred to as bitmapped.
- A vector graphics program generically called drawing program creates a sequential list of graphic commands to draw lines, curves, text, etc., with associated parameters, such as screen location, size, color, rotation angle, width, style, etc.
- A bitmapped image, in contrast, has in the file the actual pixel image data.
- Static bitmapped images are often compressed to reduce the file sizes and thus to save some disk space and shorten the time it takes to transfer those files over a communication link.
- Most light sources are mixtures of various wavelengths of light. Many such sources can still effectively produce a spectral color, as the eye cannot distinguish them from single-wavelength sources.
- Two different light spectra that have the same effect on the three-color receptors in the human eye will be perceived as the same color.
- The characteristics of the color sensors in the devices are often far from the characteristics of the receptors in the human eye.
- A color reproduction system “tuned” to a human with normal color vision may give very inaccurate results for other observers.
- Additive color is light created by mixing light of two or more different colors.

NOTES

- Red, green, and blue are the additive primary colors normally used in additive color systems, such as projectors and computer terminals.
- Subtractive coloring uses dyes, inks, pigments, or filters to absorb some wavelengths of light and not others.
- The color that a surface displays comes from the parts of the visible spectrum that are not absorbed and therefore remain visible.
- Structural colors are colors caused by interference effects rather than by pigments.
- Color effects are produced when a material is scored with fine parallel lines, formed of one or more parallel thin layers, or otherwise composed of microstructures on the scale of the color's wavelength.
- Two common color models in video are YUV and YIQ. YUV uses properties of the human eye to prioritize information.
- Y is the black and white or luminance image, U and V are the color difference or chrominance images.
- YUV is a color encoding system typically used as part of a color image pipeline.
- YUV encodes a color image or video taking human perception into account, allowing reduced bandwidth for chrominance components.
- YUV typically enabling transmission errors or compression artifacts to be more efficiently masked by the human perception than using a direct RGB-representation.
- YIQ is the color space used by the NTSC color TV system, employed mainly in North and Central America, and Japan. I stand for in-phase, while Q stands for quadrature, referring to the components used in quadrature amplitude modulation.
- The YIQ representation is sometimes employed in color image processing transformations.

3.7 KEY WORDS

- **Color science:** The science of color is sometimes called chromatics, colorimetry, or simply color science.
- **Visible light:** The wavelength is within the visible spectrum the range of wavelengths humans can perceive, approximately from 390 nm to 700 nm, it is known as visible light.

- **Static bitmapped images:** They are often compressed to reduce the file sizes and thus to save some disk space and shorten the time it takes to transfer those files over a communication link.
- **YUV:** It is a color encoding system typically used as part of a color image pipeline.
- **YIQ:** It is the color space used by the NTSC color TV system, employed mainly in North and Central America, and Japan. I stand for in-phase, while Q stands for quadrature, referring to the components used in quadrature amplitude modulation.

3.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define the term color constancy.
2. What are color categories and physical specifications of color associated?
3. How can images be generated and stored in a personal computer?
4. How are white images displayed in bitmapped images?
5. List the most common compressed file formats.

Long-Answer Questions

1. Briefly discuss what are color models with the help of an example.
2. Elaborate color models in videos with the help of appropriate examples.
3. Briefly explain the difference between YUV and YIQ.
4. Explain what does the color of an object depends on with the help of a suitable examples.

3.9 FURTHER READINGS

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

NOTES

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

BLOCK - II
FUNDAMENTAL CONCEPTS IN VIDEO AND DIGITAL AUDIO

NOTES

UNIT 4 FUNDAMENTAL CONCEPTS IN VIDEO AND DIGITAL AUDIO

Structure

4.0 INTRODUCTION

Video signals can be organized in three different ways: Component video, Composite video, and S - video. Component Video. Component video is a video signal that has been split into two or more component channels. The most common type of component signal separates a video signal into three components. For the analog output of many devices, the three components are red, green, and blue signals. There are three types of video signals as follows: Composite Video. Component Video. S-Video. Composite video is an analog video signal format that carries standard-definition video (typically at 480i or 576i resolution) as a single channel. Video information is encoded on one channel, unlike the higher-quality S-video (two channels) and the even higher-quality component video (three or more channels). In all these video formats, audio is carried on a separate connection. Component video is an analog video signal that has been split into two or more component channels. In popular use, it refers to a type of component analog video (CAV) information that is transmitted or stored as three separate signals. Component video can be contrasted with composite video in which all the video information is combined into a single signal that is used in analog television. S-Video (also known as separate video and Y/C) is a signalling standard for standard definition video, typically 480i or 576i. By separating the black-and-white and coloring signals, it achieves better image quality than composite video, but has lower color resolution than component video.

Digital video is perhaps the most prominent multimedia object to create impact on its audience. With the advancement of cheaper storage technology and better video compression techniques, digital video has become a very widely used multimedia object.

In this unit, you will study about the video signals and its type, analog video format, video camera and analog video camera, digital video camera, various video file formats and different digital video standards.

4.1 OBJECTIVES

After going through this unit, you will be able to:

- Analyse the significance of video signals
- Interpret the characteristics of type of video signals
- Know about component, s-video or composite form
- Discuss about the analog video formats
- Understand the difference between video camera and analog video camera
- Explain the prime feature of digital video camera
- Define the functions of transmission of video signals
- Know about the various video file formats
- Comprehend about the different digital video standards

4.2 TYPES OF VIDEO SIGNALS

The video signals can be divided into below types:

In *component video*, separate signals are sent for each part of the three luminance-chrominance components. Component video has three separate paths for the information and three connectors at the end. This is the most accurate format for representing and transmitting video, as crosstalk between the different components is eliminated. However, it is also very expensive format. Only a few high-end analog video cameras like Betacam use component video connections (Refer Figure 4.1).

NOTES

Fig. 4.1 *Typical RGB Component Video* Connection

The next way to transmit video is *s-video*, which utilizes two data paths, one for the luminance component, and the other for two chrominance components.

In *composite video* the video signal is sent on just one channel. The signal is sent on a single channel by compositing the signal. The main disadvantage to this technology is that the quality of the signal may deteriorate due to crosstalk between the color (chrominance) and luminance components. Thus, composite video is the lowest quality of all the three alternatives.

Some of the popular type of analog video camera formats during the 1980s was VHS, S-VHS and Hi-8. S-VHS had much higher resolution and higher color quality support than VHS. The Hi-8 video uses a smaller tape, and thus video cameras were smaller than for VHS. Then came the *Betacams*. The Betacams used small size tapes and gave extremely high quality video and became very popular in the news reporting circuit.

The resolution of any given video camera depends on whether it is NTSC-, PAL- or SECAM compliant. The number of horizontal lines in a frame is called the *vertical resolution* of an image. Now, for an analog video camera, the picture information is sent as a continuous waveform rather than as discrete pixels in case of the digital video camera. So the horizontal resolution for an analog video camera is only indicative. It is not fixed but lies within a range. That is why you will find that the preceding standards specify only the vertical resolution, but no horizontal resolution.

NOTES

The analog video formats are summarized in the Table 4.1.

Table 4.1 *Analog Video Formats*

Video Format	**Year Introduced**	**Colour Transmission Format**	**Horizontal Resolution**	**Tape Width**	**Quality**
VHS	1976	composite	~240	½" (12.5 mm)	consumer
Betamax	1976	composite	~240	½" (12.5 mm)	consumer
8mm (Video 8)	1984	composite	~240?300	8 mm	consumer
S-VHS	1987	s-video	~400?425	½" (12.5 mm)	high-end
consumer					
Hi-8	1998	s-video	~400?425	8 mm	high-end
consumer					
U-Matic	1971	composite	~250?340	¾" (18.75 mm)	professional
M-II	1986	component	~400?440	½" (12.5 mm)	professional
Betacam	1982	component	~300?320	½" (12.5 mm)	consumer
Betacam SP	1986	component	~340?360	½" (12.5 mm)	professional

Note: Vertical resolution depends on whether the camera is NTSC-, PAL- or SECAM-compliant.

Analog video has become outdated with the availability of digital video cameras at affordable price for personal and professional use. However, analog video cameras can still be connected to the PC with a video capture card and appropriate software—and video clippings can be captured in digital format directly on the computer. Also, using an interface device and a VHS player, an old VHS tape can be played and captured digitally on the computer. Video capturing is the processing of encoding and saving video in digital format on a computer.

4.3 ANALOG VIDEO, DIGITAL VIDEO

Digital video is perhaps the most prominent multimedia object to create impact on its audience. With the advancement of cheaper storage technology and better video compression techniques, digital video has become a very widely used multimedia object. The term '*video*' has been derived from Latin, meaning 'I see' or 'I apprehend'. The term videography refers to the process of capturing moving pictures.

Video technology has evolved from television technology, but it has now developed to a great extent to allow consumer digital video recording and playback. The standards that have evolved initially through analog television and film, then digital video, and finally to present day HDTV have jumbled up with lot of ambiguous terminologies and nomenclature. To get rid of this ambiguity, you have to know the basics of analog video, television technology and film. These are the three close kins of digital video and have many features in common. Moreover, to understand the mechanism of digital data communication you have to understand how data is communicated in an analog manner. You have to understand the concepts of frame, number of lines in a frame, frame rate, etc. These concepts evolved from analog television, knowing them makes it easier to understand the similar issues in digital video. Another similarity between television and digital video technology is their requirement for bandwidth. Just as the television in the analog domain requires bandwidth in the airwaves, the digital video requires bandwidth for transmission across the computer network. Bandwidth is a costly component, so its requirement is to be controlled and minimized to keep things economical. Digital technology is all around us, so you should have a fair idea about the frequently used terminologies, such as television screen sizes, resolution, etc.

Video Camera

To understand how digital video camera functions, you should first study some fundamental aspects of analog video camera.

Analog Video Camera

In an analog video camera, light comes through the lens and hits an imaging chip, which reacts to the light with continuously varying voltages. The stronger the light, the stronger the voltage is. These voltages, after magnification and signal processing, magnetize the tape particles in a continuously varying (analog) pattern that stores the signal.

In analog systems, the video signal from the camera is delivered to the video through the connector cables of a VCR, where it is recorded on magnetic videotape. The video signal is written to tape by a spinning recording head that changes the local magnetic properties of the tape's surface in a series of long diagonal stripes. The head is tilted at a slight angle compared with the path of the tape; it follows a helical (spiral) path which is called 'helical scan' recording. Each stripe represents information for one field of a video frame.

You know that color models divide color into three components. In case of digital still images, you can use the RGB color model—where a color is separated into its red, green, and blue components. In the analog video, luminance/chrominance models, such as the YUV and YIQ, it functions better. The luminance — chrominance color models also need three pieces of information — one luminance (value or brightness) and two chrominances to represent a color. In case of an analog video, the information can be passed on in either of three ways — in *component*, *s-video* or *composite* form.

NOTES

Digital Video Camera

Just like the analog video camera, digital video cameras light coming in through the lens and scans the image line-by-line. However, here the light intensity is digitally encoded, unlike the analog camera, where the light is recorded as continuously changing voltage. Just like the analog camera, the encoded pixel values can also be stored on videotape. However, the information is analog and thus different from the analog signal.

A unique sensor converts light into an electronic signal, which is called Charge-Coupled-Device (CCD), when the light reflected from a subject passes through a video camera lens. CCD was invented way back in 1969 at Bell Labs, by George Smith and Willard Boyle and is used in many digital optical devices including camcorders and digital still cameras to convert light energy to electronic signal. They are also used in astronomical telescopes, scanners and bar code readers. CCD is a light-sensitive integrated circuit. It is at times referred to as a *chip* or microchip that stores and shows the data for an image in such a way that each pixel (picture element) in the image is changed into an electrical charge, the intensity of which is related to a color in the color spectrum. The camera processes the output of the CCD into a signal having three channels of color information and synchronization pulses (sync). Different video standards are there for managing the CCD output, each deal with the amount of separation between the components of the signal.

Some high-quality cameras used for professional broadcasting have as many as three CCDs (one for each color of red, green and blue) to augment the resolution of the camera (Refer Figure 4.2).

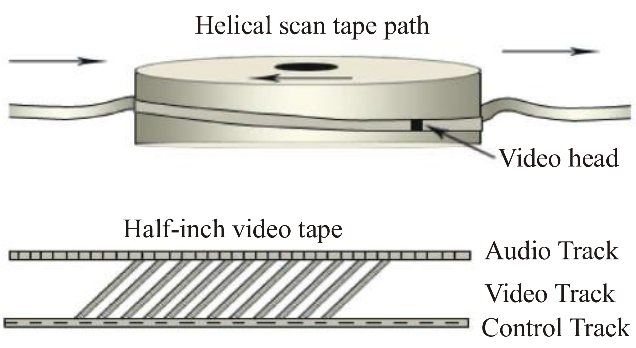


Fig. 4.2 *CCDs for High Quality Cameras*

Transmission of Video Signals

Basically, video or motion pictures are created by displaying images depicting progressive stages of motion at a rate fast enough so that the projection of individual images overlap on the eye. Persistence of vision of human eye, which allows any projected image to persist for 40-50 ms, requires a frame rate of 25-30 frames per second to ensure perception of smooth motion picture.

NOTES

In a video display:

- *Horizontal resolution* is the number of distinct vertical lines that can be produced in a frame.
- *Vertical resolution* is the number of horizontal scan lines in a frame.
- *Aspect ratio* is the width-to-height ratio of a frame.
- *Interlace*[†] *ratio* is the ratio of the frame rate to the field rate.

Constitution-wise there are three types of video signals—*component video*, *composite video* and s-*video*. Most computer systems and high-end video systems use component video whereby the three signals R, G and B are transmitted through three separate wires corresponding to red, green and blue image planes, respectively. However, because of the complexities of transmitting the three signals of component video in exact synchronism and relationship these signals are encoded using a frequency-interleaving scheme into a composite format that can be transmitted through a single cable. Such format known as composite video, used by most video systems and broadcast TV, uses one luminance and two chrominance signals. *Luminance* (Y) is a monochrome video signal that controls only the brightness of an image. *Chrominance* is actually two signals (I and Q or U and V), called color differences (B–Y, R–Y) and contains color information of an image. Each chrominance component is allocated half as much bandwidth as the luminance, a form of analog data compression[*], which is justified by the fact that human eyes are less sensitive to variations in color than to variations in brightness. Theoretically, there are infinity of possible combinations (additive) of R, G and B signals to produce Y, I and Q or Y, U and V signals. The common CCIR 601 standard defines:

$$\text{Luminance (Y)} = 0.299R + 0.587\,G + 0.114B$$

$$\text{Chrominance (U)} = 0.596R - 0.247\,G - 0.322B$$

$$\text{Chrominance (V)} = 0.211R - 0.523G + 0.312B$$

The inverse of the preceding transformation formula gives

$$\text{Red (R)} = 1.0\,Y + 0.956\,U + 0.621\,V$$

$$\text{Green (G)} = 1.0\,Y - 0.272\,U - 0.647\,V$$

$$\text{Blue (B)} = 1.0\,Y - 1.061\,U - 1.703\,V$$

Unlike composite video, s-video (separated video or super video as S–VHS) uses two wires, one for luminance and another for a composite chrominance signal. Component video gives the best output since there is no crosstalk or interference between the different channels unlike composite video or s-video.

Digital Video Standards

To improve picture quality and transmission efficiency, new generation televisions systems are designed based on international standards that exploit the advantage of digital signal processing. These standards include *High Definition Television* or *HDTV*, *Improved Definition Television* or *IDTV*, *Double Multiplexed Analog Components* or *D2-MAC*, *Advanced Compatible Television*, *First System* or *ACTV-I*. The HDTV standard that support progressive (non-interlaced) video scanning has much wider aspect ratio (16:9 instead of 4:3), greater field of view, higher horizontal and vertical resolution (9600 and 675 respectively in the USA) and more bandwidth (9 MHz in USA) as compared to conventional color TV systems.

Video File Formats

In a television transmission system, every part of every moving image is converted into analog electronic signals and transmitted. The VCR can store TV signals on magnetic tapes, which can be played to reproduce stored images. There are three main standards for analog video signals used in television transmission: NTSC, SECAM and PAL. Television standard in India is based on the Phase Alternate Lines (PAL) system. The PAL system is followed in the countries with Alternating Current (AC) frequency of 50Hz, such as in UK, Western Europe, Australia, China, India, South Africa and South America. In this system, the screen resolution is 625 lines but the scan rate is 25 frames per second (to suit the AC frequency). The National Television Standards Committee guided the development of America's television standard hence it is called NTSC. Another 50Hz Standard is Sequential Color Avec Memoire (SECAM). This standard is mainly followed in France, the Eastern Block and Middle Eastern Countries.

Check Your Progress

1. What are the three types of video signals?
2. Which is the prominent multimedia object to create impact on its audience?
3. How does analog video camera works?
4. Define the term Charge-Coupled-Device (CCD).
5. What is a horizontal resolution?
6. What is the purpose of a VCR?

4.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. The three types of video signals are: composite video, component video and s- video.
2. Digital video is perhaps the most prominent multimedia object to create impact on its audience.
3. In an analog video camera, light comes through the lens and hits an imaging chip, which reacts to the light with continuously varying voltages.

4. A unique sensor converts light into an electronic signal, which is called Charge-Coupled-Device (CCD).
5. Horizontal resolution is the number of distinct vertical lines that can be produced in a frame.
6. The VCR can store TV signals on magnetic tapes, which can be played to reproduce stored images

4.5 SUMMARY

- In component video, separate signals are sent for each part of the three luminance-chrominance components.
- Component video has three separate paths for the information and three connectors at the end.
- In composite video the video signal is sent on just one channel. The signal is sent on a single channel by compositing the signal.
- Some of the popular type of analog video camera formats during the 1980s was VHS, S-VHS and Hi-8. S-VHS.
- The Hi-8 video uses a smaller tape, and thus video cameras were smaller than for VHS.
- The Betacams used small size tapes and gave extremely high-quality video and became extremely popular in the news reporting circuit.
- The resolution of any given video camera depends on whether it is NTSC-, PAL- or SECAM compliant.
- The number of horizontal lines in a frame is called the vertical resolution of an image.
- Analog video has become outdated with the availability of digital video cameras at affordable price for personal and professional use.
- Video capturing is the processing of encoding and saving video in digital format on a computer.
- A unique sensor converts light into an electronic signal, which is called Charge-Coupled-Device (CCD), when the light reflected from a subject pass-through a video camera lens.
- CCD was invented way back in 1969 at Bell Labs, by George Smith and Willard Boyle and is used in many digital optical devices including camcorders and digital still cameras to convert light energy to electronic signal.
- Different video standards are there for managing the CCD output, each deal with the amount of separation between the components of the signal.

NOTES

- Persistence of vision of human eye, which allows any projected image to persist for 40-50 ms, requires a frame rate of 25-30 frames per second to ensure perception of smooth motion picture.
- Horizontal resolution is the number of distinct vertical lines that can be produced in a frame.
- Vertical resolution is the number of horizontal scan lines in a frame.
- Aspect ratio is the width-to-height ratio of a frame.
- Interlace† ratio is the ratio of the frame rate to the field rate.
- Constitution-wise there are three types of video signals—component video, composite video and s-video.
- Digital video is perhaps the most prominent multimedia object to create impact on its audience.
- The term 'video' has been derived from Latin, meaning 'I see' or 'I apprehend'.
- Video technology has evolved from television technology, but it has now developed to a great extent to allow consumer digital video recording and playback.
- In an analog video camera, light comes through the lens and hits an imaging chip, which reacts to the light with continuously varying voltages.
- In analog systems, the video signal from the camera is delivered to the video through the connector cables of a VCR, where it is recorded on magnetic videotape.
- The video signal is written to tape by a spinning recording head that changes the local magnetic properties of the tape's surface in a series of long diagonal stripes.
- The VCR can store TV signals on magnetic tapes, which can be played to reproduce stored images.

4.6 KEY WORDS

- **Digital video:** It is the most prominent multimedia object to create impact on its audience.
- **Videography:** The term videography refers to the process of capturing moving pictures.
- **Component video:** In component video, separate signals are sent for each part of the three luminance-chrominance components.
- **Composite video:** In composite video the video signal is sent on just one channel. The signal is sent on a single channel by compositing the signal.

- **Betacams:** The Betacams used small size tapes and gave extremely high-quality video and became extremely popular in the news reporting circuit.
- **Video capturing:** Video capturing is the processing of encoding and saving video in digital format on a computer.
- **Charge-Coupled-Device (CCD):** Charge- Coupled-Device (CCD) is a unique sensor converts light into an electronic signal.
- **Horizontal resolution:** It is the number of distinct vertical lines that can be produced in a frame.
- **Vertical resolution:** It is the number of horizontal scan lines in a frame.
- **Aspect ratio:** It is the width-to-height ratio of a frame.
- **Interlace† ratio:** It is the ratio of the frame rate to the field rate.
- **Luminance (Y):** It is a monochrome video signal that controls only the brightness of an image.
- **Chrominance:** It is two signals (I and Q or U and V), called color differences (B–Y, R–Y) and contains color information of an image.

4.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is component video?
2. Name some of the popular type of analog video camera formats during the 1980s.
3. What are Betacams?
4. Define the term vertical resolution of a frame.
5. State the function of a digital video.
6. What is the helical scan recording?
7. Name the three components of a color which a color models divide.
8. What is a vertical resolution?
9. State the difference between aspect ratio and interlace ratio.
10. Name the standards for analog video signals used in television transmission.

Long-Answer Questions

1. Briefly discuss the types of video signals with the help of an example.
2. Elaborate on the analog video formats giving suitable examples.
3. Discuss in detail about the analog digital camera listing their objectives.

4. Explain the term VCR giving appropriate examples.
5. Give an overview on the history of digital video standards.
6. Explain the terms vertical resolution and horizontal resolution with the help of examples.
7. Describe briefly about video file formats giving suitable examples.

4.8 FURTHER READINGS

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction.* New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

UNIT 5 DIGITIZATION OF SOUND

NOTES

Structure

5.0 INTRODUCTION

Digitization is a process of converting the analog signals to a digital signal. ... A sampling rate is the number of times the analog sound is taken per second. A higher sampling rate implies that more samples are taken during the given time interval and ultimately, the quality of reconstruction is better. MIDI stands for Musical Instrument Digital Interface. It is a musical interface to musical instruments. It is a technology that enables a user to create and play music easily. This can also make learning music much easier than that can be made possible by using conventional musical instruments. There are many such devices on which MIDI runs. MIDI runs on cell phones as well as the music keyboard of personal computers. All these devices see data in a format, defined for MIDI.

The Ethernet is a family of wired computer networking technologies commonly used in Local Area Networks (LAN), Metropolitan Area Networks (MAN) and Wide Area Networks (WAN). It was commercially introduced in 1980 and first standardized in 1983 as IEEE 802.3. Ethernet has since been refined to support higher bit rates, a greater number of nodes, and longer link distances, but retains much backward compatibility. Over time, Ethernet has largely replaced competing wired LAN technologies such as Token Ring, FDDI and ARCNET.

In this unit, you will study about the digitization of sound concept of MIDI, MIDI systems, hardware aspects of MIDI, types of MIDI, the concept of Ethernet, MIDI data format and channel voice messages.

5.1 OBJECTIVES

NOTES

After going through this unit, you will be able to:

- Analyse the significance of digitization of sound
- Interpret the type of video signals
- Discuss the advantages and limitations of MIDI
- Know about the components of MIDI system
- Discuss the hardware aspects of MIDI
- Comprehend the different types of MIDI
- Know about the concept of the Ethernet
- Analyse the MIDI data format
- Learn about channel voice messages

5.2 DIGITIZATION OF SOUND

Capability of computers was enhanced when sound was digitized. Once sound is put in the digital format, its manipulation becomes easy. Software programs are needed to read these files having graphic formats.

Sound may be made in a multimedia creation in the following ways:

- Creation of music that changes mood of a person.
- Narration for direct and effective expression for better communication.
- Creating sound effects for highlighting points. It may be a drum beat.

A narrative soundtrack can be created by connecting microphone through microphone jack. This jack is usually located at the back panel of a computer, in the sound card. Computers these days come with builtin microphones and in that case no sound card is required. One should check whether good quality sound will be produced or not.

There is another way to accomplish this task of acquiring sound in digital format, for manipulation by computer programs. One may record the voice and digitize it using the capability of the software. After recording it should be played to check for the quality. Music and sound effect can be transferred into computer in that way.

After opening audio editing software, settings should be adjusted 16 bit, 22.05 kHz that produces quality enough on home computers. To get sound quality, it should be set to 44 kHz. This setting consumes large amount of space on the disk and computer becomes slow. But if sound file is small, this setting may be kept for better quality of sound.

NOTES

Fig. 5.1 *Digitization of Sound*

Software programs that handle digital sound have on-screen player containing controls found in home stereo systems. By bringing the mouse pointer on these buttons their functions are displayed. The control buttons for actions are: play, pause, fast forward or review, record, stop. After setting suitable level, record button should be pressed to start recording spoken voice into the microphone. One can also play a pre-recorded music track. In one go one may not get the desired sound narration effect and few retakes may be needed. After recording is over, it should be saved in a format supported by the software. These files are normally saved in *.wav* or *.aiff* file (see Figure 5.2).

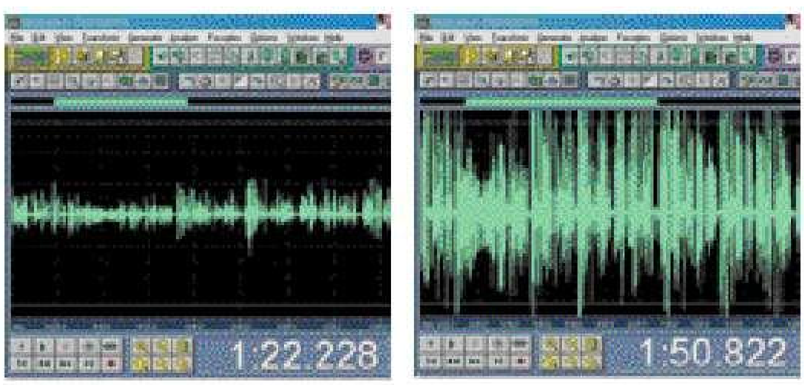


Fig. 5.2 *Waveform Representation of Sound*

While recording adjustment in input volume may be required. A recording done at low level has been shown above on the left side needs some boosting to make it more live. But with boosting of recorded sound noise in the background also gets boosted. When recording is done at very high level, which is shown above on the right side, will distort or cut out. If such problems do occur, re-record should be done.

Audio editors enable one for making selection certain parts of sound, and may be pasted in another part of a sound track by cutting or copying. These days, advanced audio editors are available that contain every function available in professional recording studios. They enable recording on multiple tracks, mixing these into one. Some special effects can also be added and tone and pitch of the sound can also be changed. After finishing this work, it should be given an easily recognizable name and saved in folder containing audio files.

5.3 MIDI

MIDI stands for Musical Instrument Digital Interface. It is a musical interface to musical instruments. It is a technology that enables a user to create and play music easily. This can also make learning music much easier than that can be made possible by using conventional musical instruments. This musical interface appeared as a protocol in 1982 as an industry standard. One need not play an actual musical instrument to create or play a musical note. One can do the same by using the keyboard of a computer connected to an audio output device, a speaker or a set of speakers.

This digital interface for music is software that will run on hardware devices. There are many such devices on which MIDI runs. MIDI runs on cell phones as well as the music keyboard of personal computers. All these devices see data in a format, defined for MIDI. This interface describes the method of playing music in a way similar to sheet music. MIDI messages describe which notes should be played, for how long and at what volume. MIDI message is a component of MIDI. MIDI permits communication between computer systems, synthesizers, controllers, samplers, sound cards and drum machines in proper coordination with the exchange of system data. Standard MIDI Files (SMF) have file extension writtens as *.mid*. The files have MIDI instructions telling about notes, volumes, sounds and even certain effects. To play music, these files are put in a 'player'. The player may be a hardware device or software package stored in the system. A sound-engine produces the final ***sound.*** The sound-engine is either a part of the player or connected to the player.

There are various components of a MIDI system. These are described as follows:

5.3.1 Components of MIDI System

A MIDI system comprises of the following essential parts:

1. Synthesizer

This part of MIDI system is made of:

- A sound generators, capable of producing sound with different pitches, intensity and tone
- A microprocessor, keyboard, control panels, memory, etc.

NOTES

2. Sequencer

- It is a separate unit for a personal computer that acts as a storage server for data in the MIDI. It may be a software program, such as music editor.
- It has ports for inputs and outputs. These are named MIDI INs and MIDI OUTs.

3. Track

- Track is contained inside the sequencer and it organizes recordings.
- One may turn these tracks on or off during playback or recording.

4. Channel

- MIDI channels create separation of information.
- In one cable, 16 channels are handled.
- Each MIDI message has coded channel numbers.

5. Timbre

- This tells about sound quality, e.g., flute sound or sound of other instruments.
- A sound may be called multi-timbre that means it can play various sounds simultaneously. These may combine sound emitted by musical instruments, such as piano, drums, violin, guitar, etc.

6. Pitch

- This is a name for musical note played by the instrument.

7. Voice

- Voice is a part of the 'synthesizer' that generates sound.
- A synthesizer may have a number of voice grades (36, 24, 20, 12, etc.)
- Every voice type works simultaneously and yet independently for producing sounds with different pitch and timbre.

8. Patch

- This quality of sound tells about control settings defining a particular timbre.

5.3.2 Hardware Aspects of MIDI

MIDI Connectors

A device that handles music from different instruments has three numbers of 5-pin ports. These ports are located at the back panel of a MIDI unit.

- **MIDI IN:** Via this connector, MIDI data is received.
- **MIDI OUT:** Via this connector MIDI data is transmitted by the device.

- **MIDI THROUGH:** Via this connector, data is echoed by the device received by MIDI IN.

Note: Only data received in MIDI IN is echoed by MIDI THROUGH. Data generated is routed through MIDI (as shown in Figure 5.3).

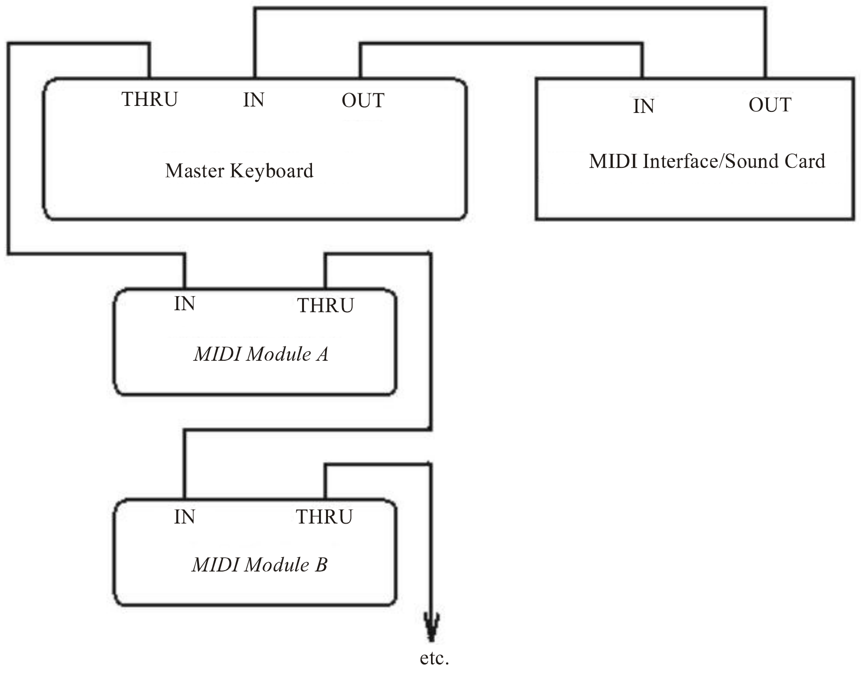


Fig. 5.3 *Data Generated*

MIDI Sequencer Setup

- The synthesizer's MIDI OUT gets attached to the sequencer's MIDI IN.
- The sequencer's MIDI OUT gets connected to the synthesizer's MIDI IN and to every additional module of sound.
- While recording, also known as the synthesizer, the keyboard-equipped sends MIDI message for sequencer to record these.

During playback, messages come via the sequencer and are received by audio modules, then by the synthesizer that plays music.

5.3.3 Types of MIDI

Original specification of MIDI defines a *connector* (a physical device) with a digital format in which the message *is contained*, to connect devices to have control over them in a 'real time' mode. Subsequently, standard MIDI files had been developed with a format for *storage so that it might be* recalled later.

MIDI messages specification is also known as 'MIDI Protocol'. This is its most important part. Originally, it was intended just to be used with MIDI DIN transport for connecting two keyboards. But now, MIDI messages find use inside cell phones and computers for generating music. These messages are transported to many interfaces—consumer as well as professional—that make use of FireWire, USB, etc. to various devices that are MIDI-equipped. Different message groups are used for different applications.

Different cables and connectors are used for transporting MIDI data between devices. MIDI DIN transport has a specific character which is not the character of the MIDI. Each transport is known by its performance characteristics but its importance need not be overemphasized as the purpose of these transports is to connect devices that are to be used for MIDI applications.

Standard MIDI Files with their variants are the final part of this interface. These are distributed music that can be played on MIDI players having a variety of software and hardware. Today, most of the computer systems are capable of playing MIDI files that have *.mid format. Also, there are numerous websites that offer such files free or at some nominal price. MIDI files can be created easily by using software available commercially. Some of these software packages are even free.

Many different types of 'transports' are available for use of MIDI messages. The speed of such transport finds the volume of MIDI data to be carried and the speed with which it is received.

5-Pin MIDI DIN

This transport was developed in 1983 using a 5-pin 'DIN' connector. It is not fast in comparison to today's common digital transports working at very high speed. Ethernet, FireWire and USB are such high-speed digital transport systems. But MIDI-DIN is still used. For connecting a MIDI device with another, MIDI cables are normally used and considered best.

To connect a MIDI device to a computer means insalling MIDI interface or sound card, for having a MIDI DIN connector on the computer. Such cards need driver-software for making MIDI work. There are few standards followed by companies including 'MPU-401' and 'SoundBlaster'. With gradual development that took few years, components of MIDI interface and sound card formed the standard on PC-motherboards, but configuring them is not so easy.

Ports: Serial, Parallel and Joystick

Prior to the advent of FireWire and USB, PCs were equipped with ports, namely serial, parallel and joystick. These were used to connect instruments, using some special adapters that were MIDI-equipped instruments. These connectors were available on computers which provided an economical alternative to add-on cards without special configuration.

USB and FireWire

Most of the computers today have connectors, such as FireWire or USB. Most common methods employ these connectors to connect computers with MIDI

NOTES

devices by use at one end format adapters. Such adapters may be a short cable having a USB at the other and MIDI DIN. But it may be very complex. This may contain a rack mountable CPU and many MIDIs with ports for Audio In and Audio Out. FireWire and USB are interfaces of 'plug-and-play' type that are self-configuring. One can just plug in a MIDI interface to FireWire or USB and boot up.

USB technology has support for communicating between a device and a host (PC). USB devices cannot be connected to each other the way it is done with MIDI devices having DIN ports. FireWire supports connections, of the type 'peer-to-peer'. So MMA (MIDI Manufactures Association) developed specification for MIDI data is able to work on standard IEEE-1394. This standard defines FireWire.

Ethernet

Many MIDI instruments may be connected to one or more computers using Ethernet. In the MIDI industry, there is no effort to use Ethernet, and hence, no efforts by MMA to develop standards that make MIDI work on Ethernet. But some standard setting organizations made efforts to develop a format so that MIDI can be connected using Ethernet.

In the past 20 years or more, original MIDI specification has been enhanced by defining few more performance-control messages with the creation of specifications as follows:

- MIDI machine control
- MIDI show control
- MIDI time code
- General MIDI
- Downloadable sounds
- Scalable Polyphony MIDI

Alternate Applications MIDI Machine Control and MIDI Show Control are extensions. They address recording equipment with theatrical control for studio instead of musical instruments. MIDI also controls devices in which standard messages have not been defined. Such devices are those that do audio mixing console automation.

Popularity of MIDI files is due to the fact that audio digital files *.aiff, .wav*, etc., are required to capture and then store actual sound, whereas a MIDI file need not do these. Instead, MIDI files could create an events' list describing steps taken by soundcard or other playback devices to generate sound. MIDI files are very small when compared to these digital audio files. In MIDI, events can also be edited, music can be rearranged, edited and even composed interactively.

MIDI Messages

MIDI messages enable MIDI devices to communicate mutually.

The structure of MIDI messages is as follows:

- MIDI message keeps one status byte and maximum two data bytes.
- Status byte
  - o Status byte has the setting '1' for the most significant bit.
  - o To identify the channel to which it belongs, four low-order bits are used that can create 16 channels.
  - o Message is identified by remaining 3 bits.
- Data byte has '0' as the most significant bit.

Digital Audio and MIDI

- Many applications of MIDI and Digital Audio are used together.
- Recording studios, that are modernized use the recording of MIDI files on hard disk.
- Analog sounds are digitized and recorded on hard disk. These analog sounds may come from speech or songs that are live vocals or a guitar or saxophone, a flute, etc.
- Digital audio uses keyboards, samples, drums and loop effects to play MIDI files.
- Sound generators use a mix of:
  - o Synthesis
  - o Samples
  - o Samplers that digitize sound
- Playback of the generated sound
- Loop (beats)
- Simulation of many traditional musical instruments

5.3.4 MIDI Data Format

Most of the communication in MIDI has packets containing multi-bytes that have status bit in the beginning followed by data bytes (one or two). As we know, a byte is a packet of 8 bits containing either 0s or 1s. Status bytes start with '1' for example, 1*** ****. This is known as a call 'set'. Data bytes have '0' in the beginning for example, 0*** **** which is known as 'reset.' Every byte has one start and one stop bit. This makes each packet of length 10 bits. Messages fall into the following five formats:

1. Channel voice

This type of message controls 16 voices (patches, timbres) of instruments, plays notes, transmits data for the controller, etc.

2. Channel mode

This defines response of the instrument to Voice messages. These are transmitted over basic channel of the instrument.

3. System common

These messages are intended for instruments in the networked environment and devices.

4. System real-time

This is applicable to all instruments under the network and devices. This has only status bytes and its use is made for synchronizing devices, which is essentially a timing clock.

5. System exclusive

This was **o**riginally for use in manufacturer-specific codes. But this was expanded for inclusion of MIDI Time Code, MIDI Sample Dump Standard and MIDI Machine Control

Channel Voice Messages

Most of the MIDI devices can receive MIDI messages on any selectable channel from 16 MIDI channels. It may select more than one such MIDI channel. A particular voice of a device that contains pitch, timbre etc., will give response to messages that are sent on the channel to which it is tuned. It ignores other channel messages. This is quite analogous to a television set that receives only the station to which it is tuned and rejects all others. The OMNI mode is an exception to this. An instrument set that receives in this mode will be accepting and responding to every channel message, regardless of the channel number.

Check Your Progress

1. What is hypermedia?
2. Define the term MIDI.
3. What is a MIDI protocol?
4. What are channel modes?
5. Define the term system real time.

5.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

NOTES

1. Sound may be made in a multimedia creation in the following three used ways:
   - Creation of music that changes mood of a person.
   - Narration for direct and effective expression for better communication.
   - Creating sound effects for highlighting points. It may be a drumbeat.
2. MIDI stands for Musical Instrument Digital Interface. It is a musical interface to musical instruments. It is a technology that enables a user to create and play music easily.
3. MIDI messages specification is also known as 'MIDI Protocol'. This is its most important part.
4. Cannel modes defines response of the instrument to Voice messages. These are transmitted.
5. System real time is applicable to all instruments under the network and devices. This has only status bytes and its use is made for synchronizing devices, which is essentially a timing clock.

5.5 SUMMARY

- Capability of computers was enhanced when sound was digitized. Once sound is put in the digital format, its manipulation becomes easy.
- A narrative soundtrack can be created by connecting microphone through microphone jack. This jack is usually located at the back panel of a computer, in the sound card.
- Software programs that handle digital sound has on-screen player containing controls found in home stereo systems.
- Audio editors enable one for making selection certain parts of sound and may be pasted in another part of a soundtrack by cutting or copying.
- MIDI stands for Musical Instrument Digital Interface. It is a musical interface to musical instruments. It is a technology that enables a user to create and play music easily.
- MIDI permits communication between computer systems, synthesizers, controllers, samplers, sound cards and drum machines in proper coordination with the exchange of system data.
- MIDI messages specification is also known as 'MIDI Protocol'.

NOTES

- Different cables and connectors are used for transporting MIDI data between devices. MIDI DIN transport has a specific character which is not the character of the MIDI.
- Many MIDI instruments may be connected to one or more computers using Ethernet.
- In the MIDI industry, there is no effort to use Ethernet, and hence, no efforts by MMA to develop standards that make MIDI work on Ethernet.
- Alternate Applications MIDI Machine Control and MIDI Show Control are extensions. They address recording equipment with theatrical control for studio instead of musical instruments.
- MIDI also controls devices in which standard messages have not been defined. Such devices are those that do audio mixing console automation.
- Recording studios, that are modernized use the recording of MIDI files on hard disk.
- Analog sounds are digitized and recorded on hard disk. These analog sounds may come from speech or songs that are live vocals or a guitar or saxophone, a flute, etc.
- Most of the communication in MIDI has packets containing multi-bytes that have status bit in the beginning followed by data bytes.
- Most of the MIDI devices can receive MIDI messages on any selectable channel from 16 MIDI channels. It may select more than one such MIDI channel.

5.6 KEY WORDS

- **MIDI:** MIDI stands for Musical Instrument Digital Interface. It is a musical interface to musical instruments.
- **MIDI IN:** Via this connector, MIDI data is received.
- **MIDI OUT:** Via this connector MIDI data is transmitted by the device.
- **MIDI THROUGH:** Via this connector, data is echoed by the device received of sound by MIDI IN.
- **Channel voice:** This type of message controls 16 voices (patches, timbres) of instruments, plays notes, transmits data for the controller, etc.
- **Channel mode:** This defines response of the instrument to Voice messages. These are transmitted over basic channel of the instrument.
- **System real-time:** This is applicable to all instruments under the network and devices. This has only status bytes and its use is made for synchronizing devices, which is essentially a timing clock.

5.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

NOTES

Short-Answer Questions

1. What was the purpose of digitization of sound?
2. Differentiate between the term's synthesizer and sequencer.
3. What is the role of track in MIDI system?
4. Define the term Standard MIDI Files (SMF).
5. State the function of a MIDI message.
6. State the significance of Pin MIDI DIN.
7. How digital audio and MIDI are related to each other?
8. How can a narrative soundtrack be created stating its significance?

Long-Answer Questions

1. Briefly discuss the significance of MIDI and list its various components.
2. Elaborate on the characteristics of MIDI sequencer setup giving suitable examples.
3. Discuss in detail about the Ethernet listing their objectives.
4. Give an overview on the MIDI messages with the help of examples.
5. Elaborate on hardware aspects of MIDI with suitable examples.

5.8 FURTHER READINGS

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

NOTES

UNIT 6 QUANTIZATION AND TRANSMISSION OF AUDIO

NOTES

Structure

6.0 INTRODUCTION

Quantization, in mathematics and digital signal processing, is the process of mapping input values from a large set to output values in a smaller set, often with a finite number of elements. Rounding and truncation are typical examples of quantization processes.

Audio transmission systems are how audio signals are routed, processed, and assigned to the desired monitor and recording output channels. In the early days of electrical recording, rarely more than one or two microphones were used, and the signal routing was to a single output channel. Audio transmission systems used in the motion picture industry were more complex from their inception because of the necessity of recording dialogue, sound effects, and music at different times. In a sense, this represented the beginning of multichannel recording even though the final product was a monophonic one.

In this unit, you will study about the quantization, and its types, mathematical properties of quantization, noise and error characteristics, noise model and audio transmission.

6.1 OBJECTIVES

After going through this unit, you will be able to:

- Analyse the significance of quantization
- Understand the significance of transmission of audio
- Explain the prime feature of quantization in the field of transmission

NOTES

- Define the functions of Analog-to-Digital Converter (ADC)
- Know about the noise and error characteristics
- Comprehend about the quantization noise model
- Recognise the various audio transmission and their description

6.2 QUANTIZATION AND TRANSMISSION OF AUDIO

Quantization, in mathematics and digital signal processing, is the process of mapping input values from a large set to output values in a countable smaller set, often with a finite number of elements. Rounding and truncation are typical examples of quantization processes. Quantization is involved to some degree in nearly all digital signal processing, as the process of representing a signal in digital form ordinarily involves rounding. Quantization also forms the core of essentially all lossy compression algorithms.

The difference between an input value and its quantized value, such as round-off error is referred to as quantization error. A device or algorithmic function that performs quantization is called a quantizer. An analog-to-digital converter is an example of a quantizer (Refer Figure 6.1).

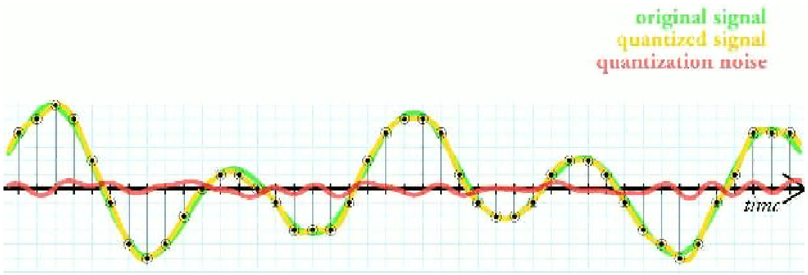


Fig. 6.1 *Conceptual Model for Quantization*

6.2.1 Types of Quantization

1. Analog-to-Digital Converter

An Analog-to-Digital Converter (ADC) can be modelled as two processes: sampling and quantization. Sampling converts a time-varying voltage signal into a discrete-time signal, a sequence of real numbers. Quantization replaces each real number with an approximation from a finite set of discrete values. Most commonly, these discrete values are represented as fixed-point words. Though any number of quantization levels is possible, common word-lengths are 8-bit (256 levels), 16-bit (65,536 levels) and 24-bit (16.8 million levels). Quantizing a sequence of numbers produces a sequence of quantization errors which is sometimes modelled as an additive random signal called quantization noise because of its stochastic behaviour. The more levels a quantizer uses, the lower is its quantization noise power.

NOTES

2. Rate Distortion Optimization

Rate distortion optimized quantization is encountered in source coding for lossy data compression algorithms, where the purpose is to manage distortion within the limits of the bit rate supported by a communication channel or storage medium. The analysis of quantization in this context involves studying the amount of data (typically measured in digits or bits or bit rate) that is used to represent the output of the quantizer and studying the loss of precision that is introduced by the quantization process (which is referred to as the distortion).

3. Mid-Riser and Mid-Tread Uniform Quantizers

Most uniform quantizers for signed input data can be classified as being of one of two types: mid-riser and mid-tread. The terminology is based on what happens in the region around the value 0 and uses the analogy of viewing the input-output function of the quantizer as a stairway. Mid-thread quantizers have a zero-valued reconstruction level (corresponding to a tread of a stairway), while mid-riser quantizers have a zero-valued classification threshold (corresponding to a riser of a stairway).

4. Dead Zone Quantizers

A dead-zone quantizer is a type of mid-tread quantizer with symmetric behaviour around 0. The region around the zero-output value of such a quantizer is referred to as the dead zone or dead band. The dead zone can sometimes serve the same purpose as a noise gate or squelch function. Especially for compression applications, the dead-zone may be given a different width than that for the other steps.

Mathematical Properties of Quantization

As quantization is a many-to-few mapping, it is an inherently non-linear and irreversible process, i.e., because the same output value is shared by multiple input values, it is impossible, in general, to recover the exact input value when given only the output value. The set of possible input values may be infinitely large, and may possibly be continuous and therefore uncountable, such as the set of all real numbers, or all real numbers within some limited range. The set of possible output values may be finite or countably infinite. The input and output sets involved in quantization can be defined in a rather general way. For example, vector quantization is the application of quantization to multi-dimensional (vector-valued) input data.

6.2.2 Noise and Error Characteristics

Additive Noise Model

A common assumption for the analysis of quantization error is that it affects a signal processing system in a similar manner to that of additive white noise – having negligible correlation with the signal and an approximately flat power spectral

NOTES

density. The additive noise model is commonly used for the analysis of quantization error effects in digital filtering systems, and it can be extremely useful in such analysis. It has been shown to be a valid model in cases of high-resolution quantization with smooth probability density functions. Additive noise behaviour is not always a valid assumption. Quantization error is deterministically related to the signal and not entirely independent of it. Thus, periodic signals can create periodic quantization noise. And in some cases, it can even cause limit cycles to appear in digital signal processing systems. One way to ensure effective independence of the quantization error from the source signal is to perform dithered quantization (sometimes with noise shaping), which involves adding random (or pseudo-random) noise to the signal prior to quantization.

Quantization Error Models

In the typical case, the original signal is much larger than one Least Significant Bit (LSB). When this is the case, the quantization error is not significantly correlated with the signal, and has an approximately uniform distribution. When rounding is used to quantize, the quantization error has a mean of zero and the root mean square (RMS) value is the standard deviation of this distribution, given by $\frac{1}{\sqrt{12}}\mathrm{LSB} \approx 0.289\ \mathrm{LSB}$. At lower amplitudes, the quantization error becomes dependent on the input signal, resulting in distortion. This distortion is created after the anti-aliasing filter, and if these distortions are above 1/2 the sample rate they will alias back into the band of interest. To make the quantization error independent of the input signal, the signal is dithered by adding noise to the signal. This slightly reduces signal to noise ratio but can eliminate the distortion.

Quantization Noise Model

Quantization noise is a model of quantization error introduced by quantization in the Analog-to-Digital Conversion (ADC). It is a rounding error between the analog input voltage to the ADC and the output digitized value. The noise is non-linear and signal dependent. It can be modelled in several different ways. In an ideal analog-to-digital converter, where the quantization error is uniformly distributed between –1/2 LSB and +1/2 LSB, and the signal has a uniform distribution covering all quantization levels, the Signal-to-quantization-noise ratio (SQNR) can be calculated from

$$\mathrm{SQNR} = 20\log_{10}(2^Q) \approx 6.02 \cdot Q\ \mathrm{dB}$$

Where, Q is the number of quantization bits.

6.3 AUDIO TRANSMISSION

Audio transmission systems are how audio signals are routed, processed, and assigned to the desired monitor and recording output channels. In the early days of electrical recording, rarely more than one or two microphones were used, and the signal routing was to a single output channel. Audio transmission systems used in the motion picture industry were more complex from their inception because of the necessity of recording dialogue, sound effects, and music at different times.

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. Linear Pulse-Code Modulation (LPCM) is a specific type of PCM in which the quantization levels are linearly uniform. This contrasts with PCM encodings in which quantization levels vary as a function of amplitude (as with the A-law algorithm or the ì-law algorithm). Though PCM is a more general term, it is often used to describe data encoded as LPCM. A PCM stream has two basic properties that determine the stream's fidelity to the original analog signal: the sampling rate, which is the number of times per second that samples are taken; and the bit depth, which determines the number of possible digital values that can be used to represent each sample.

Check Your Progress

1. What is known as quantization?
2. Define the term quantizer.
3. What is rate distortion optimized quantization?
4. What is dead-zone quantization?
5. What is quantization error?
6. State the role of Pulse-Code Modulation (PCM).

6.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Quantization, in mathematics and digital signal processing, is the process of mapping input values from a large set to output values in a countable smaller set, often with a finite number of elements.
2. A device or algorithmic function that performs quantization is called a quantizer.
3. Rate distortion optimized quantization is encountered in source coding for lossy data compression algorithms, where the purpose is to manage distortion within the limits of the bit rate supported by a communication channel or storage medium.
4. A dead-zone quantizer is a type of mid-tread quantizer with symmetric behaviour around 0. The region around the zero-output value of such a quantizer is referred to as the dead zone or dead band.

5. Quantization error is deterministically related to the signal and not entirely independent of it. Thus, periodic signals can create periodic quantization noise.
6. 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.

6.5 SUMMARY

- Quantization, in mathematics and digital signal processing, is the process of mapping input values from a large set to output values in a countable smaller set, often with a finite number of elements.
- Rounding and truncation are typical examples of quantization processes.
- Quantization is involved to some degree in nearly all digital signal processing, as the process of representing a signal in digital form ordinarily involves rounding.
- The difference between an input value and its quantized value, such as round-off error is referred to as quantization error. A device or algorithmic function that performs quantization is called a quantizer.
- An Analog-to-Digital Converter (ADC) can be modelled as two processes: sampling and quantization.
- Sampling converts a time-varying voltage signal into a discrete-time signal, a sequence of real numbers.
- Quantization replaces each real number with an approximation from a finite set of discrete values.
- Rate distortion optimized quantization is encountered in source coding for lossy data compression algorithms, where the purpose is to manage distortion within the limits of the bit rate supported by a communication channel or storage medium.
- The analysis of quantization in this context involves studying the amount of data (typically measured in digits or bits or bit rate) that is used to represent the output of the quantizer and studying the loss of precision that is introduced by the quantization process (which is referred to as the distortion).
- Mid-tread quantizers have a zero-valued reconstruction level (corresponding to a tread of a stairway), while mid-riser quantizers have a zero-valued classification threshold (corresponding to a riser of a stairway).
- A dead-zone quantizer is a type of mid-tread quantizer with symmetric behaviour around 0. The region around the zero-output value of such a quantizer is referred to as the dead zone or dead band.
- As quantization is a many-to-few mapping, it is an inherently non-linear and irreversible process i.e., because the same output value is shared by multiple

input values, it is impossible, in general, to recover the exact input value when given only the output value.

- A common assumption for the analysis of quantization error is that it affects a signal processing system in a similar manner to that of additive white noise – having negligible correlation with the signal and an approximately flat power spectral density.
- The additive noise model is commonly used for the analysis of quantization error effects in digital filtering systems, and it can be extremely useful in such analysis.
- Quantization error is deterministically related to the signal and not entirely independent of it.
- Quantization noise is a model of quantization error introduced by quantization in the Analog-to-Digital Conversion (ADC).
- Audio transmission systems are how audio signals are routed, processed, and assigned to the desired monitor and recording output channels.
- Audio transmission systems used in the motion picture industry were more complex from their inception because of the necessity of recording dialogue, sound effects, and music at different times.
- 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.
- Linear Pulse-Code Modulation (LPCM) is a specific type of PCM in which the quantization levels are linearly uniform. This contrasts with PCM encodings in which quantization levels vary as a function of amplitude (as with the A-law algorithm or the ì-law algorithm).

6.6 KEY WORDS

- **Quantization:** It is the process of mapping input values from a large set to output values in a countable smaller set, often with a finite number of elements.
- **Sampling:** It converts a time-varying voltage signal into a discrete-time signal, a sequence of real numbers.
- **Dead-zone quantizer:** It is a type of mid-tread quantizer with symmetric behaviour around 0.
- **Quantization error:** It is deterministically related to the signal and not entirely independent of it.
- **Audio transmission:** It is how audio signals are routed, processed, and assigned to the desired monitor and recording output channels.

- **Pulse-Code Modulation (PCM):** It is a method used to digitally represent sampled analog signals.
- **Linear Pulse-Code Modulation (LPCM):** It is a specific type of PCM in which the quantization levels are linearly uniform.

6.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define the term sampling.
2. What is the role of quantizing?
3. What is the purpose of rate distortion optimized quantization?
4. State the difference between mid-riser and mid-thread uniform quantizers.
5. What is a quantization error model?

Long-Answer Questions

1. Briefly discuss the significance of quantization with the help of an example.
2. Elaborate on the types of quantization giving suitable examples.
3. Discuss in detail about the mathematical properties of quantization listing their objectives.
4. Explain the terms audio transmission giving appropriate examples.

6.8 FURTHER READINGS

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

BLOCK - III
MULTIMEDIA DATA COMPRESSION

NOTES

UNIT 7 LOSSLESS COMPRESSION ALGORITHM

Structure

7.0 INTRODUCTION

Lossless compression is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data. By contrast, lossy compression permits reconstruction only of an approximation of the original data, though usually with greatly improved compression rates and therefore, reduced media sizes). Algorithms used in Lossless compression are: Run Length Encoding, Lempel-Ziv-Welch, Huffman Coding, Arithmetic encoding, etc. Lossy compression is used in Images, audio, video. Lossless Compression is used in Text, images, sound. Run Length Encoding (RLE) is a form of lossless data compression in which runs of data, sequences in which the same data value occurs in many consecutive data elements are stored as a single data value and count, rather than as the original run. In coding theory, a variable-length code is a code which maps source symbols to a variable number of bits. Variable-length codes can allow sources to be compressed and decompressed with zero error (lossless data compression) and still be read back symbol by symbol. With the right coding strategy an independent and identically distributed source may be compressed almost arbitrarily close to its entropy. Variable-length codes can be strictly nested in order of decreasing generality as non-singular codes, uniquely decodable codes, and prefix codes. Arithmetic Coding (AC) is a form of entropy encoding used in lossless data compression. Normally, a string of characters, such as the words 'hello there' is represented using a fixed number of bits per character, as in the ASCII code. Arithmetic coding differs from other forms of entropy encoding, such as Huffman coding.

NOTES

In this unit, you will study about lossless compression algorithm, run length encoding, variable length coding, non- singular code, dictionary coding, arithmetic coding, Huffman coding.

7.1 OBJECTIVES

After going through this unit, you will be able to:

- Analyse the significance of lossless compression
- Interpret the characteristics of a lossless compression algorithm
- Discuss the advantages and limitations of lossless compression
- Understand the Run Length Encoding (RLE)
- Explain the prime feature of variable length code
- Define the functions of non- singular code
- Know about the dictionary coder and its use
- Comprehend about the arithmetic coding
- Recognise the significance of Huffman coding

7.2 RUN LENGTH ENCODING

Lossless compression is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data. By contrast, lossy compression permits reconstruction only of an approximation of the original data, though usually with greatly improved compression rates (and therefore reduced media sizes).

Run Length Encoding (RLE) is a form of lossless data compression in which runs of data (sequences in which the same data value occurs in many consecutive data elements) are stored as a single data value and count, rather than as the original run. This is most efficient on data that contains many such runs, for example simple graphic images such as icons, line drawings, Conway's Game of Life, and animations. For files that do not have many runs, RLE could increase the file size.

RLE may also be used to refer to an early graphics file format supported by CompuServe for compressing black and white images but was widely supplanted by their later Graphics Interchange Format (GIF).

Run Length Encoding (RLE) schemes were employed in the transmission of analog television signals as far back as 1967. In 1983, run-length encoding was patented by Hitachi. RLE is particularly well suited to palette-based bitmap images, such as computer icons and was a popular image compression method on early online services, such as CompuServe before the advent of more sophisticated formats, such as GIF. It does not work well on continuous-tone images, such as photographs, although JPEG uses it 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. The International Telecommunication Union also describes a standard to encode run-length-colour for fax machines, known as T.45. The standard, which is combined with other techniques into Modified Huffman coding, is relatively efficient because most faxed documents are generally white space, with occasional interruptions of black.

Example of Run length Encoding

Consider a screen containing plain black text on a solid white background. There will be many long runs of white pixels in the blank space, and many short runs of black pixels within the text. A hypothetical scan line, with B representing a black pixel and W representing white, might read as follows:

WWWWWWWWWWWWBWWWWWWWWWWWWBBBWWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWBWWWWWWWWWWWWWW

With a run-length encoding data compression algorithm applied to the above hypothetical scan line, it can be rendered as follows:

12W1B12W3B24W1B14W

This can be interpreted as a sequence of twelve Ws, one B, twelve Ws, three Bs, etc., and represents the original 67 characters in only 18. While the actual format used for the storage of images is generally binary rather than ASCII characters like this, the principle remains the same. Even binary data files can be compressed with this method; file format specifications often dictate repeated bytes in files as padding space. However, newer compression methods, such as DEFLATE often use LZ77-based algorithms, a generalization of run-length encoding that can take advantage of runs of strings of characters (, such as BWWBWWBWWBWW).

Run-length encoding can be expressed in multiple ways to accommodate data properties as well as additional compression algorithms. For instance, one popular method encodes run lengths for runs of two or more characters only, using an 'escape' symbol to identify runs, or using the character itself as the escape, so that any time a character appears twice it denotes a run. On the previous example, this would give the following:

WW12BWW12BB3WW24BWW14

This would be interpreted as a run of twelve Ws, a B, a run of twelve Ws, a run of three Bs, etc. In data where runs are less frequent, this can significantly improve the compression rate.

NOTES

One other matter is the application of additional compression algorithms. Even with the runs extracted, the frequencies of different characters may be large, allowing for further compression; however, if the run lengths are written in the file in the locations where the runs occurred, the presence of these numbers interrupts the normal flow and makes it harder to compress. To overcome this, some run-length encoders separate the data and escape symbols from the run lengths, so that the two can be handled independently. For the example data, this would result in two outputs, the string 'WWBWWBBWWBWW' and the numbers (12,12,3,24,14).

7.3 VARIABLE - LENGTH CODING

In coding theory, a variable-length code is a code which maps source symbols to a variable number of bits. Variable-length codes can allow sources to be compressed and decompressed with zero error (lossless data compression) and still be read back symbol by symbol. With the right coding strategy an independent and identically distributed source may be compressed almost arbitrarily close to its entropy. This contrasts with fixed length coding methods, for which data compression is only possible for large blocks of data, and any compression beyond the logarithm of the total number of possibilities comes with a finite (though perhaps arbitrarily small) probability of failure. Some examples of well-known variable-length coding strategies are Huffman coding, Lempel–Ziv coding, arithmetic coding, and context-adaptive variable-length coding.

Codes and Their Extensions

The extension of a code is the mapping of finite length source sequences to finite length bit strings, that is obtained by concatenating for each symbol of the source sequence the corresponding codeword produced by the original code. Using terms from formal language theory, the precise mathematical definition is as follows:

Let S and T be two finite sets, called the source and target alphabets, respectively. A code $C:S \rightarrow T^*$ is a total function mapping each symbol from S to a sequence of symbols over T, and the extension of C to a homomorphism of S^* into T^*, which naturally maps each sequence of source symbols to a sequence of target symbols, is referred to as its extension.

7.3.1 Classes of Variable-Length Codes

Variable-length codes can be strictly nested in order of decreasing generality as non-singular codes, uniquely decodable codes, and prefix codes. Prefix codes are always uniquely decodable, and these in turn are always non-singular:

Non-Singular Codes

A code is non-singular if each source symbol is mapped to a different non-empty bit string, i.e., the mapping from source symbols to bit strings is injective.

- For example, the mapping M1 = { a $\leftrightarrow$ 0, b $\leftrightarrow$ 0, c $\leftrightarrow$ 0 } is not non-singular because both 'a' and 'b' map to the same bit string '0' ; any extension of this mapping will generate a lossy (non-lossless) coding. Such singular coding may still be useful when some loss of information is acceptable (for example when such code is used in audio or video compression, where a lossy coding becomes equivalent to source quantization).
- However, the mapping M2 = { a $\leftrightarrow$ 1, b $\leftrightarrow$ 011, c $\leftrightarrow$ 01110, d $\leftrightarrow$ 1110, e $\leftrightarrow$ 10011, f $\leftrightarrow$ 0 } is non-singular ; its extension will generate a lossless coding, which will be useful for general data transmission (but this feature is not always required). Note that it is not necessary for the non-singular code to be more compact than the source (and in many applications, a larger code is useful, for example to detect and/or recover from encoding or transmission errors, or in security applications to protect a source from undetectable tampering).

Uniquely Decodable Codes

A code is uniquely decodable if its extension is non-singular. Whether a given code is uniquely decodable can be decided with the Sardinas–Patterson algorithm.

- The mapping M3 = { a $\leftrightarrow$ 0, b $\leftrightarrow$ 01, c $\leftrightarrow$ 011 } is uniquely decodable (this can be demonstrated by looking at the follow-set after each target bit string in the map, because each bitstring is terminated as soon as we see a 0 bit which cannot follow any existing code to create a longer valid code in the map, but unambiguously starts a new code).
- Consider again the code M2 from the previous section. This code is not uniquely decodable, since the string 011101110011 can be interpreted as the sequence of codewords 01110 – 1110 – 011, but also as the sequence of codewords 011 – 1 – 011 – 10011. Two possible decoding of this encoded string are thus given by cdb and babe. However, such a code is useful when the set of all possible source symbols is completely known and finite, or when there are restrictions (for example a formal syntax) that determine if source elements of this extension are acceptable. Such restrictions permit the decoding of the original message by checking which of the possible source symbols mapped to the same symbol are valid under those restrictions.

Prefix Codes

A code is a prefix code if no target bit string in the mapping is a prefix of the target bit string of a different source symbol in the same mapping. This means that symbols can be decoded instantaneously after their entire codeword is received. Other commonly used names for this concept are prefix-free code, instantaneous code, or context-free code.

- The example mapping M3 in the previous paragraph is not a prefix code because we do not know after reading the bit string '0' if it encodes an "a" source symbol, or if it is the prefix of the encodings of the 'b' or 'c' symbols.

NOTES

- An example of a prefix code is shown below in Table 7.1:

Table 7.1 *Prefix Codes*

Symbol	Codeword
a	0
b	10
c	110
d	111

Example of encoding and decoding:

aabacdab → 00100110111010 → |0|0|10|0|110|111|0|10| → aabacdab

A special case of prefix codes is block codes. Here all codewords must have the same length. The latter are not very useful in the context of source coding, but often serve as error correcting codes in the context of channel coding.

Another special case of prefix code is variable-length quantity codes, which encode arbitrarily large integers as a sequence of octets—i.e., every codeword is a multiple of 8 bits.

7.4 DICTIONARY BASED CODING

A dictionary coder, also sometimes known as a substitution coder, is a class of lossless data compression algorithms which operate by searching for matches between the text to be compressed and a set of strings contained in a data structure (called the 'dictionary') maintained by the encoder. When the encoder finds such a match, it substitutes a reference to the string's position in the data structure.

Methods and Applications

Some dictionary coders use a 'static dictionary', one whose full set of strings is determined before coding begins and does not change during the coding process. This approach is most often used when the message or set of messages to be encoded is fixed and large; for instance, an application that stores the contents of a book in the limited storage space of a PDA generally builds a static dictionary from a concordance of the text and then uses that dictionary to compress the verses. This scheme of using Huffman coding to represent indices into a concordance has been called 'Huffword'. In a related and more general method, a dictionary is built from redundancy extracted from a data environment (various input streams) which dictionary is then used statically to compress a further input stream. For example, a dictionary is built from old English texts then is used to compress a book.

More common methods are where the dictionary starts in some predetermined state, but the contents change during the encoding process, based

NOTES

on the data that has already been encoded. Both the LZ77 and LZ78 algorithms work on this principle. In LZ77, a circular buffer called the "sliding window" holds the last *N* bytes of data processed. This window serves as the dictionary, effectively storing *every* substring that has appeared in the past *N* bytes as dictionary entries. Instead of a single index identifying a dictionary entry, two values are needed: the *length*, indicating the length of the matched text, and the *offset* (also called the *distance*), indicating that the match is found in the sliding window starting *offset* bytes before the current text. LZ78 uses a more explicit dictionary structure; at the beginning of the encoding process, the dictionary is empty. An index value of zero is used to represent the end of a string, so the first index of the dictionary is one. At each step of the encoding process, if there is no match, then the last matching index (or zero) and character are both added to the dictionary and output to the compressed stream. If there is a match, then the working index is updated to the matching index, and nothing is output.

LZW is like LZ78, but the dictionary is initialized to all possible symbols. The typical implementation works with 8-bit symbols, so the dictionary "codes" for hex 00 to hex FF (decimal 255) are pre-defined. Dictionary entries would be added starting with code value hex 100. Unlike LZ78, if a match is not found (or if the end of data), then only the dictionary code is output. This creates a potential issue since the decoder output is one step behind the dictionary. Refer to LZW for how this is handled. Enhancements to LZW include handing symbol sizes other than 8 bits and having reserved codes to reset the dictionary and to indicate end of data.

Algorithm LZW Compression

```
BEGIN
        s = next input character;
        While not EOF
                    {c = next input character;

                     If s + c exists in the dictionary
                                 s = s + c
                           else
                                {output the code for s
;
                                   add string s + c to
the dictionary with a new code;
                                    s = c;
                                }
                           }
              output the code for s;
       END
```

7.5 ARITHMETIC CODING

NOTES

Arithmetic Coding (AC) is a form of entropy encoding used in lossless data compression. Normally, a string of characters, such as the words "hello there" is represented using a fixed number of bits per character, as in the ASCII code. When a string is converted to arithmetic encoding, frequently used characters will be stored with fewer bits and not-so-frequently occurring characters will be stored with more bits, resulting in fewer bits used in total. Arithmetic coding differs from other forms of entropy encoding, such as Huffman coding, in that rather than separating the input into component symbols and replacing each with a code, arithmetic coding encodes the entire message into a single number, an arbitrary-precision fraction q, where 0.0 d" q < 1.0. It represents the current information as a range, defined by two numbers. A recent family of entropy coders called asymmetric numeral systems allows for faster implementations thanks to directly operating on a single natural number representing the current information.

Implementation Details and Examples

Equal Probabilities

In the simplest case, the probability of each symbol occurring is equal. For example, consider a set of three symbols, A, B, and C, each equally likely to occur. Simple block encoding would require 2 bits per symbol, which is wasteful: one of the bit variations is never used. Symbols A, B and C might be encoded respectively as 00, 01 and 10, with 11 unused. A more efficient solution is to represent a sequence of these three symbols as a rational number in base 3 where each digit represents a symbol. For example, the sequence "ABBCAB" could become 0.0112013, in arithmetic coding as a value in the interval [0, 1]. The next step is to encode this ternary number using a fixed-point binary number of sufficient precisions to recover it, such as 0.00101100102 – this is only 10 bits; 2 bits are saved in comparison with naïve block encoding. This is feasible for long sequences because there are efficient, in-place algorithms for converting the base of arbitrarily precise numbers.

To decode the value, knowing the original string had length 6, one can simply convert back to base 3, round to 6 digits, and recover the string.

Encoding and Decoding

The current interval (at the very start of the encoding process, the interval is set to [0,1], but that will change). The probabilities the model assigns to each of the various symbols that are possible at this stage. The encoder divides the current interval into sub-intervals, each representing a fraction of the current interval proportional to the probability of that symbol in the current context. Whichever interval corresponds to the actual symbol that is next to be encoded becomes the interval used in the next step.

Example: for the four-symbol model above:

the interval for NEUTRAL would be [0, 0.6)

the interval for POSITIVE would be [0.6, 0.8)

the interval for NEGATIVE would be [0.8, 0.9)

the interval for END-OF-DATA would be [0.9, 1).

When all symbols have been encoded, the resulting interval unambiguously identifies the sequence of symbols that produced it. Anyone who has the same final interval and model that is being used can reconstruct the symbol sequence that must have entered the encoder to result in that final interval.

It is not necessary to transmit the final interval, however; it is only necessary to transmit one fraction that lies within that interval. It is only necessary to transmit enough digits (in whatever base) of the fraction so that all fractions that begin with those digits fall into the final interval; this will guarantee that the resulting code is a prefix code.

One advantage of arithmetic coding over other similar methods of data compression is the convenience of adaptation. Adaptation is the changing of the frequency (or probability) tables while processing the data. The decoded data matches the original data if the frequency table in decoding is replaced in the same way and in the same step as in encoding. The synchronization is, usually, based on a combination of symbols occurring during the encoding and decoding process.

Connections with Other Compression Methods

Huffman Coding

Because arithmetic coding does not compress one datum at a time, it can get arbitrarily close to entropy when compressing iid strings. By contrast, using the extension of Huffman coding (to strings) does not reach entropy unless all probabilities of alphabet symbols are powers of two, in which case both Huffman and arithmetic coding achieve entropy. When naively Huffman coding binary strings, no compression is possible, even if entropy is low (e.g. ({0, 1}) has probabilities {0.95, 0.05}). Huffman encoding assigns 1 bit to each value, resulting in a code of the same length as the input. By contrast, arithmetic coding compresses bits well, approaching the optimal compression ratio of

$$1 - [-0.95 \log_2 (0.95) + - 0.05 \log_2 (0.05)] \sim 71.4\%$$

One simple way to address Huffman coding's suboptimality is to concatenate symbols ("blocking") to form a new alphabet in which each new symbol represents a sequence of original symbols – in this case bits – from the original alphabet. In the above example, grouping sequences of three symbols before encoding would produce new "super-symbols" with the following frequencies:

- 000: 85.7%
- 001, 010, 100: 4.5% each

NOTES

- 011, 101, 110: 0.24% each
- 111: 0.0125%

With this grouping, Huffman coding averages 1.3 bits for every three symbols, or 0.433 bits per symbol, compared with one bit per symbol in the original encoding, i.e., 56.7% compression. Allowing arbitrarily large sequences gets arbitrarily close to entropy – just like arithmetic coding – but requires huge codes to do so, so is not as practical as arithmetic coding for this purpose.

An alternative is encoding run lengths via Huffman-based Golomb-Rice codes. Such an approach allows simpler and faster encoding/decoding than arithmetic coding or even Huffman coding, since the latter requires a table lookup. In the {0.95, 0.05} example, a Golomb-Rice code with a four-bit remainder achieves a compression ratio of 71.1%, far closer to optimum than using three-bit blocks. Golomb-Rice codes only apply to Bernoulli inputs, such as the one in this example, however, so it is not a substitute for blocking in all cases.

Check Your Progress

1. What is lossless compression?
2. Define the term Run Length Encoding (RLE).
3. When were the RLE schemes employed?
4. State the significance of variable length code.
5. When is any code called non-singular?
6. What is a dictionary coder?
7. Where is the arithmetic coding used?
8. State the significance of Huffman coding.

7.6 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Lossless compression is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data.
2. Run-Length Encoding (RLE) is a form of lossless data compression in which runs of data (sequences in which the same data value occurs in many consecutive data elements) are stored as a single data value and count, rather than as the original run.
3. Run-Length Encoding (RLE) schemes were employed in the transmission of analog television signals as far back as 1967.
4. In coding theory, a variable-length code is a code which maps source symbols to a variable number of bits. Variable-length codes can allow sources to be

compressed and decompressed with zero error (lossless data compression) and still be read back symbol by symbol.

5. A code is non-singular if each source symbol is mapped to a different non-empty bit string, i.e., the mapping from source symbols to bit strings is injective.
6. A dictionary coder, also sometimes known as a substitution coder, is a class of lossless data compression algorithms which operate by searching for matches between the text to be compressed and a set of strings contained in a data structure (called the 'dictionary') maintained by the encoder.
7. Arithmetic Coding (AC) is a form of entropy encoding used in lossless data compression.
8. Huffman encoding assigns 1 bit to each value, resulting in a code of the same length as the input.

7.7 SUMMARY

- Lossless compression is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data.
- Run Length Encoding (RLE) is a form of lossless data compression in which runs of data (sequences in which the same data value occurs in many consecutive data elements) are stored as a single data value and count, rather than as the original run.
- RLE may also be used to refer to an early graphics file format supported by CompuServe for compressing black and white images but was widely supplanted by their later Graphics Interchange Format (GIF).
- Run Length Encoding (RLE) schemes were employed in the transmission of analog television signals as far back as 1967.
- The standard, which is combined with other techniques into Modified Huffman coding, is relatively efficient because most faxed documents are generally white space, with occasional interruptions of black.
- In coding theory, a variable-length code is a code which maps source symbols to a variable number of bits.
- Variable-length codes can allow sources to be compressed and decompressed with zero error (lossless data compression) and still be read back symbol by symbol.
- The extension of a code is the mapping of finite length source sequences to finite length bit strings, that is obtained by concatenating for each symbol of the source sequence the corresponding codeword produced by the original code.

NOTES

- Variable-length codes can be strictly nested in order of decreasing generality as non-singular codes, uniquely decodable codes, and prefix codes.
- A code is non-singular if each source symbol is mapped to a different non-empty bit string, i.e., the mapping from source symbols to bit strings is injective.
- A code is uniquely decodable if its extension is non-singular. Whether a given code is uniquely decodable can be decided with the Sardinas–Patterson algorithm.
- A code is a prefix code if no target bit string in the mapping is a prefix of the target bit string of a different source symbol in the same mapping.
- A dictionary coder, also sometimes known as a substitution coder, is a class of lossless data compression algorithms which operate by searching for matches between the text to be compressed and a set of strings contained in a data structure (called the 'dictionary') maintained by the encoder.
- More common methods are where the dictionary starts in some predetermined state, but the contents change during the encoding process, based on the data that has already been encoded.
- Both the LZ77 and LZ78 algorithms work on this principle. In LZ77, a circular buffer called the "sliding window" holds the last *N* bytes of data processed.
- LZ78 uses a more explicit dictionary structure; at the beginning of the encoding process, the dictionary is empty.
- Arithmetic Coding (AC) is a form of entropy encoding used in lossless data compression.
- One advantage of arithmetic coding over other similar methods of data compression is the convenience of adaptation. Adaptation is the changing of the frequency (or probability) tables while processing the data.
- The decoded data matches the original data if the frequency table in decoding is replaced in the same way and in the same step as in encoding.

7.8 KEY WORDS

- **Lossless compression:** It is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data.
- **Run Length Encoding (RLE):** It is a form of lossless data compression in which runs of data, sequences in which the same data value occurs in many consecutive data elements are stored as a single data value and count, rather than as the original run.

- **Variable-length code:** It is a code which maps source symbols to a variable number of bits.
- **Non-singular code:** A code is non-singular if each source symbol is mapped to a different non-empty bit string.
- **Arithmetic Coding (AC):** It is a form of entropy encoding used in lossless data compression.

7.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. State the significance of lossless compression in run length coding.
2. What is known as T.45 and state its significance?
3. List any one example of run length coding.
4. What are the various ways in which run length coding can be expressed?
5. What is the extension of a code?
6. What are the uniquely decodable codes?
7. Name the principle on which both LZ77 and LZ78 algorithms work.
8. Write an algorithm for LZW compression.
9. State the significance of both encoding and decoding.
10. Define the term Huffman coding.

Long-Answer Questions

1. Discuss run length encoding with the help of a suitable examples.
2. Elaborate on variable length coding and its various codes and their extensions.
3. Describe the non-singular and uniquely decodable code with the help of suitable examples.
4. Discuss in brief about dictionary-based coding along with its methods and applications.
5. Describe the arithmetic coding. Discuss its implementation details with the help of suitable examples.

7.10 FURTHER READINGS

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

NOTES

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

UNIT 8 LOSSLESS IMAGE COMPRESSION

NOTES

Structure

8.0 INTRODUCTION

Lossless compression is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data. By contrast, lossy compression permits reconstruction only of an approximation of the original data, though usually with greatly improved compression rates and therefore reduced media sizes. By operation of the pigeonhole principle, no lossless compression algorithm can efficiently compress all possible data. In lossless compression, the original file is compressed so that it can be completely restored flaw when decompression occurs. In lossy compression, the original file is processed so that some information is permanently removed. Lossy compression techniques eliminate information that is not perceived when people see pictures or hear sounds.

Quantization, in mathematics and digital signal processing, is the process of mapping input values from a large set to output values in a countable smaller set, often with a finite number of elements. Rounding and truncation are typical examples of quantization processes. Quantization is involved to some degree in nearly all digital signal processing, as the process of representing a signal in digital form ordinarily involves rounding. Quantization also forms the core of essentially all lossy compression algorithms.

In this unit, you will study about the significance of lossless image compression, lossless compression algorithm, features of quantization, types of quantization and mathematical properties of quantization.

NOTES

8.1 OBJECTIVES

After going through this unit, you will be able to:

- Analyse the significance of lossless image compression
- Interpret the techniques used in lossless image compression
- Discuss the significance of lossless compression algorithm
- Explain the features of quantization
- Know about the types of quantization
- Comprehend about the mathematical properties of quantization

8.2 LOSSLESS IMAGE COMPRESSION

Lossless compression is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data. By contrast, lossy compression permits reconstruction only of an approximation of the original data, though usually with greatly improved compression rates (and therefore reduced media sizes). By operation of the pigeonhole principle, no lossless compression algorithm can efficiently compress all possible data. For this reason, many different algorithms exist that are designed either with a specific type of input data in mind or with specific assumptions about what kinds of redundancy the uncompressed data are likely to contain. Lossless data compression is used in many applications. For example, it is used in the ZIP file format and in the GNU tool gzip. It is also often used as a component within lossy data compression technologies (e.g., lossless mid/side joint stereo pre-processing by MP3 encoders and other lossy audio encoders). Lossless compression is used in cases where it is important that the original and the decompressed data be identical, or where deviations from the original data would be unfavourable. Typical examples are executable programs, text documents, and source code. Some image file formats, like PNG or GIF, use only lossless compression, while others like TIFF and MNG may use either lossless or lossy methods. Lossless audio formats are most often used for archiving or production purposes, while smaller lossy audio files are typically used on portable players and in other cases where storage space is limited, or exact replication of the audio is unnecessary.

Following methods are used in lossy compression technique:

- **Reduction of Color Space:** In lossy compression, color space is reduced to some most common colors of the image. Colors are specified in the header of compressed images color palette. Every pixel makes a reference to the index of a color in this color palette. To avoid posterization this method is combined with dithering.

- **Chroma Subsampling:** This method makes the use of the fact that human eye has perception for spatial changes more in case of brightness in comparison to that of color. This averages or drops some of the information on chrominance of the image.
- **Transform Coding:** This method is most commonly used and in it a transform that is Fourier related is applied. This is followed by quantization and entropy coding.
- **Fractal Compression:** This technique works on the principle of self similarity.

The main objective of image compression is to produce the best quality of image at a given bit rate known as compression rate.

8.3 LOSSY COMPRESSION ALGORITHM

Computers can work with art, photographs, videos and sounds but they can do so only when these multimedia resources are stored in digitized files. These files require huge amounts of storage space. For example, an average sized hard disk could store only a dozen audio Compact Disks (CDs) and there would not be much memory space left for system software or applications. To reduce the size of multimedia files, most software use compression or decompression algorithms called *codecs* which use two different approaches to compress the multimedia streaming files namely *lossless compression* and *lossy compression*. In lossless compression, the original file is compressed so that it can be completely restored without flaw when decompression occurs. In lossy compression, the original file is processed so that some information is permanently removed. Lossy compression techniques eliminate information that is not perceived when people see pictures or hear sounds. Joint Photographic Experts Group (JPEG) is designed for compressing full color, gray scale images or continuous tone artwork. Any smooth variation in color, such as occurring in highlighted or shaded areas will be represented more faithfully and in less space by JPEG than by GIF. Plain black and white images should never be converted to JPEG. The Graphics Interchange Format (GIF) does significantly better on images with only a few selected colors, such as line drawings and simple cartoons. There has to be at least 16 gray levels before JPEG is useful for gray scale images. JPEG is lossy. The decompressed image is not quite the same as the original. A lossless compression algorithm is one that guarantees its decompressed output to be bit for bit identical to the original input. This scheme does not discard any data during the encoding process while the lossy scheme throws useless data away during encoding. That is, in fact, how lossy schemes manage to obtain superior compression ratios over most lossless schemes. JPEG has been designed specifically to discard information that the human eye cannot easily see. Because the human eye is much more sensitive to brightness

NOTES

variations in gray scale than to color variations and JPEG can compress color data more heavily than brightness data. Gray scale images do not compress well by large factors. GIF is lossless for gray scale images up to 256 levels while JPEG is not. The more complex and subtly rendered image the more likely JPEG will do. Multimedia data can be compressed with the help of following algorithms:

- **Lossy Compression Algorithm:** It is the compression technique which will lose data in the original source while trying to keep the visible quality at the almost same amount. The compression ratio will be very high. Most probably the ratio will be a value near 10. It reduces non-sensitive information to the human eyes and the compressed media will not be the media that was available before compression. The advantage of lossy compression is that it can reduce the file size more than in the lossless compression. But, the original file cannot be taken after the decompression.
- **Lossless Compression Algorithm:** There will be no data loss in this type of compression as it is defined by the name. Both original data and the compressed data are the same in this compression. The algorithms for the compression and decompression are exact inverse of each other in the lossless compression. The main mechanism in this compression is to remove the redundant data in the compression and adding them in the decompression. The advantage of lossless compression is that the original format of the data remains even it is compressed. But, the reduction of the size of the data is small. Sometimes, the file size can be increased instead of decrease if compressed file is not required for application.

8.4 QUANTIZATION

Quantization, in mathematics and digital signal processing, is the process of mapping input values from a large set to output values in a countable smaller set, often with a finite number of elements. Rounding and truncation are typical examples of quantization processes. Quantization is involved to some degree in nearly all digital signal processing, as the process of representing a signal in digital form ordinarily involves rounding. Quantization also forms the core of essentially all lossy compression algorithms.

The difference between an input value and its quantized value, such as round-off error is referred to as quantization error. A device or algorithmic function that performs quantization is called a quantizer. An analog-to-digital converter is an example of a quantizer (Refer Figure 8.1).

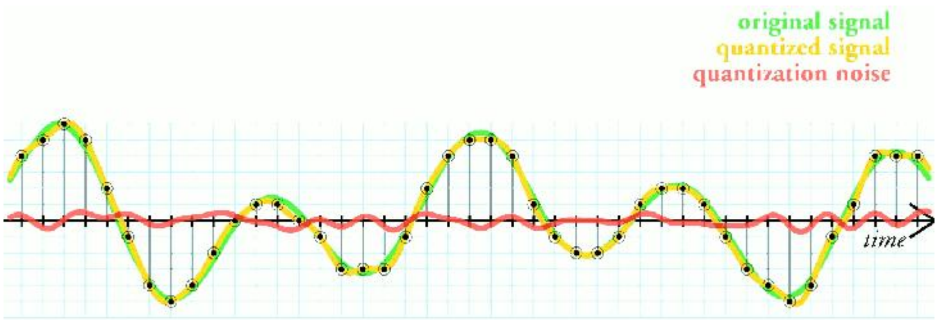


Fig. 8.1 *Conceptual Model for Quantization*

8.4.1 Types of Quantization

1. Analog-to-Digital Converter

An Analog-to-Digital Converter (ADC) can be modelled as two processes: sampling and quantization. Sampling converts a time-varying voltage signal into a discrete-time signal, a sequence of real numbers. Quantization replaces each real number with an approximation from a finite set of discrete values. Most commonly, these discrete values are represented as fixed-point words. Though any number of quantization levels is possible, common word-lengths are 8-bit (256 levels), 16-bit (65,536 levels) and 24-bit (16.8 million levels). Quantizing a sequence of numbers produces a sequence of quantization errors which is sometimes modelled as an additive random signal called quantization noise because of its stochastic behaviour. The more levels a quantizer uses, the lower is its quantization noise power.

2. Rate Distortion Optimization

Rate distortion optimized quantization is encountered in source coding for lossy data compression algorithms, where the purpose is to manage distortion within the limits of the bit rate supported by a communication channel or storage medium. The analysis of quantization in this context involves studying the amount of data (typically measured in digits or bits or bit rate) that is used to represent the output of the quantizer and studying the loss of precision that is introduced by the quantization process (which is referred to as the distortion).

3. Mid-Riser and Mid-Tread Uniform Quantizers

Most uniform quantizers for signed input data can be classified as being of one of two types: mid-riser and mid-tread. The terminology is based on what happens in the region around the value 0 and uses the analogy of viewing the input-output function of the quantizer as a stairway. Mid-tread quantizers have a zero-valued reconstruction level (corresponding to a tread of a stairway), while mid-riser quantizers have a zero-valued classification threshold (corresponding to a riser of a stairway).

NOTES

4. Dead Zone Quantizers

A dead-zone quantizer is a type of mid-tread quantizer with symmetric behaviour around 0. The region around the zero-output value of such a quantizer is referred to as the dead zone or dead band. The dead zone can sometimes serve the same purpose as a noise gate or squelch function. Especially for compression applications, the dead-zone may be given a different width than that for the other steps.

Mathematical Properties of Quantization

As quantization is a many-to-few mapping, it is an inherently non-linear and irreversible process i.e., because the same output value is shared by multiple input values, it is impossible, in general, to recover the exact input value when given only the output value. The set of possible input values may be infinitely large, and may possibly be continuous and therefore uncountable, such as the set of all real numbers, or all real numbers within some limited range. The set of possible output values may be finite or countably infinite. The input and output sets involved in quantization can be defined in a rather general way. For example, vector quantization is the application of quantization to multi-dimensional (vector-valued) input data.

Check Your Progress

1. What is lossless compression?
2. State the difference between lossless compression and lossy compression.
3. Define the term lossy compression algorithm.
4. What is quantization?
5. Define the term quantizer.

8.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Lossless compression is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data.
2. In lossless compression, the original file is compressed so that it can be completely restored without flaw when decompression occurs. In lossy compression, the original file is processed so that some information is permanently removed.
3. Lossy compression algorithm is the compression technique which will lose data in the original source while trying to keep the visible quality at the almost same amount.

4. Quantization, in mathematics and digital signal processing, is the process of mapping input values from a large set to output values in a countable smaller set, often with a finite number of elements.
5. A device or algorithmic function that performs quantization is called a quantizer.

8.6 SUMMARY

- Lossless compression is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data.
- Lossless compression is used in cases where it is important that the original and the decompressed data be identical, or where deviations from the original data would be unfavourable.
- Lossless audio formats are most often used for archiving or production purposes, while smaller lossy audio files are typically used on portable players and in other cases where storage space is limited, or exact replication of the audio is unnecessary.
- In lossy compression, color space is reduced to some most common colors of the image. Colors are specified in the header of compressed images' color palette.
- Computers can work with art, photographs, videos and sounds but they can do so only when these multimedia resources are stored in digitized files.
- To reduce the size of multimedia files, most software use compression or decompression algorithms called codecs.
- Codecs use two different approaches to compress the multimedia streaming files namely lossless compression and lossy compression.
- In lossless compression, the original file is compressed so that it can be completely restored without flaw when decompression occurs.
- In lossy compression, the original file is processed so that some information is permanently removed.
- Lossy compression techniques eliminate information that is not perceived when people see pictures or hear sounds.
- Joint Photographic Experts Group (JPEG) is designed for compressing full color, gray scale images or continuous tone artwork.
- A lossless compression algorithm is one that guarantees its decompressed output to be bit for bit identical to the original input.
- Gray scale images do not compress well by large factors. GIF is lossless for gray scale images up to 256 levels while JPEG is not.

NOTES

- Quantization, in mathematics and digital signal processing, is the process of mapping input values from a large set to output values in a countable smaller set, often with a finite number of elements.
- Quantization is involved to some degree in nearly all digital signal processing, as the process of representing a signal in digital form ordinarily involves rounding.
- Quantization also forms the core of essentially all lossy compression algorithms.
- A device or algorithmic function that performs quantization is called a quantizer.
- Sampling converts a time-varying voltage signal into a discrete-time signal, a sequence of real numbers.
- Quantization replaces each real number with an approximation from a finite set of discrete values.
- Quantizing a sequence of numbers produces a sequence of quantization errors which is sometimes modelled as an additive random signal called quantization noise because of its stochastic behaviour.
- Rate distortion optimized quantization is encountered in source coding for lossy data compression algorithms, where the purpose is to manage distortion within the limits of the bit rate supported by a communication channel or storage medium.
- Most uniform quantizers for signed input data can be classified as being of one of two types: mid-riser and mid-tread.
- A dead-zone quantizer is a type of mid-tread quantizer with symmetric behaviour around 0.
- The region around the zero-output value of such a quantizer is referred to as the dead zone or dead band.
- As quantization is a many-to-few mapping, it is an inherently non-linear and irreversible process i.e., because the same output value is shared by multiple input values, it is impossible, in general, to recover the exact input value when given only the output value.

8.7 KEY WORDS

- **Lossless compression:** It is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data.
- **Lossy compression algorithm:** It is the compression technique which will lose data in the original source while trying to keep the visible quality at the almost same amount.

- **Lossless compression algorithm:** In this algorithm will be no data loss in this type of compression as it is defined by the name.
- **Quantization:** In mathematics and digital signal processing, is the process of mapping input values from a large set to output values in a countable smaller set, often with a finite number of elements.
- **Sampling**: Sampling converts a time-varying voltage signal into a discrete-time signal, a sequence of real numbers.
- **Rate distortion optimized quantization:** It is encountered in source coding for lossy data compression algorithms, where the purpose is to manage distortion within the limits of the bit rate supported by a communication channel or storage medium.
- **Mid-tread quantizers:** They have a zero-valued reconstruction level corresponding to a tread of a stairway.
- **Mid-riser quantizers:** They have a zero-valued classification threshold corresponding to a riser of a stairway.
- **A dead-zone quantizer:** It is a type of mid-tread quantizer with symmetric behaviour around 0.

8.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is lossless audio format used for?
2. Define the term compression rate.
3. State the significance of lossless audio formats.
4. How lossy compression is different from lossless compression.
5. What is quantization error?
6. List the mathematical properties of quantization.

Long-Answer Questions

1. Briefly discuss the techniques used in lossless compression giving appropriate examples.
2. Elaborate on the characteristics on lossless image compression giving suitable examples.
3. Discuss in detail about codecs listing their uses.
4. Explain the term quantization and its types with the help of examples.
5. Briefly explain the connection between lossless compression algorithm and lossy compression algorithm.

8.9 FURTHER READINGS

NOTES

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

BLOCK - III
BASIC VIDEO COMPRESSION TECHNIQUES

NOTES

UNIT 9 BASIC VIDEO COMPRESSION TECHNIQUES

Structure

9.0 INTRODUCTION

Video compression is a practical implementation of source coding in information theory. In practice, most video codecs are used alongside audio compression techniques to store the separate but complementary data streams as one combined package using so-called container formats. Uncompressed video requires an extremely high data rate. Although lossless video compression codecs perform at a compression factor of 5 to 12, a typical H.264 lossy compression video has a compression factor between 20 and 200. The two key video compression techniques used in video coding standards are the Discrete Cosine Transform (DCT) and Motion Compensation (MC). Most video coding standards, such as the H.26x and MPEG formats, typically use motion compensated DCT video coding (block motion compensation).

MPEG-1 standard does not actually explain a compression algorithm, but it defines a datastream syntax and a decompressor. The datastream architecture is based on a sequence of frames, each of which contains the data that is needed to create a single displayed image. The Digital Display Working Group created the Digital Video Interface (DVI) technology that accommodates for interconnecting the single connector to both digital and analog interfaces. The DVI technology supports high-speed connection (digital) to display the videos and animations. It facilitates a common interface for all the display devices.

In this unit, you will study about the video compression and its characteristics, MPEG-1, prime features of I- frame, P-frame, B-frame, and D- frame, digital video interface technology, DVI technology and its features.

9.1 OBJECTIVES

After going through this unit, you will be able to:

- Analyse the significance of video compression
- Interpret the characteristics of spatial and intra frame compression
- Discuss about video compression based on motion compensation
- Know about MPEG-1 and its characteristics
- Explain the prime feature of I-frame
- Define the functions of B-frame
- Know about the P-frame and D-frame and its uses
- Comprehend about the digital video interface technology
- Recognise the various features of DVI technology

9.2 VIDEO COMPRESSION

Spatial redundancy is removed by compressing each individual image frame in isolation and the techniques used are generally called *spatial compression* or *intra frame compression*. Temporal redundancy is removed by storing only the differences of subsequent frames instead of compressing each frame independently and the technique is known as *temporal compression* or *inter frame compression*.

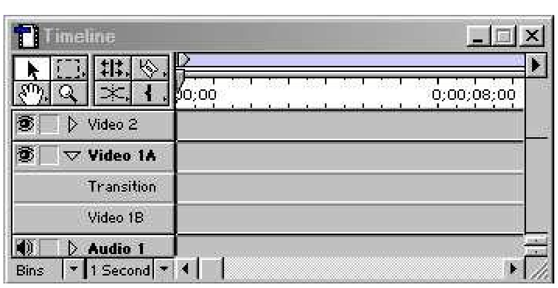
Spatial compression applies different lossless and lossy method same as those applied for still images. Some of these methods are as follows:

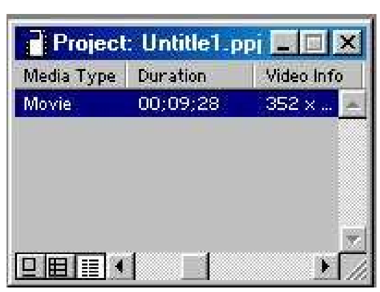
- Truncation of least significant image data
- Run Length Encoding (RLE)
- Interpolative techniques
- Predictive technique Differential Pulse Code Modulation (DPCM)**,** Adaptive DPCM (ADPCM)
- Transform coding techniques-Discrete Cosine Transform (DCT)
- Statistical or Entropy Coding–Huffman Coding, LZW coding, Arithmetic coding

The most simplistic approach for temporal compression is to perform a pixel by pixel comparison (subtraction) between two consecutive frames. The compare should produce zero for pixels which have not changed and non-zero for pixels which are somehow involved in motion. Then only the pixels with non-zero differences can be coded and stored, thus reducing the burden of storing all the pixel values of a frame. But there are certain problems with this approach. Firstly even if there is no object motion in a frame, slightest movement of camera would produce non-zero difference of all or most pixels. Secondly quantization noise would yield non-zero difference of stationary pixels.

In an alternative approach the motion generators camera and/or object can be compensated by detecting the displacements (*motion vectors*) of corresponding pixel blocks or regions in the frames and measuring the differences of their content (*prediction error*). Such approach of temporal compression is said to be based on *motion compensation*. For efficiency each image is divided into *macroblocks* of size N × N. The current image frame is referred to as *target frame*. Each macroblock of the target frame is examined with reference to most similar macroblocks in previous and/or next frame called *reference frame*. This examination is known as *forward prediction* or *backward prediction* depending on whether the reference frame is a previous frame or next frame. If the target macroblock is found to contain no motion, a code is sent to the decompressor to leave the block the way it was in the reference frame. If the block does have motion the motion vector and difference block need to be coded so that the decompressor can reproduce the target block from the code.

For building multimedia applications digitized video is often required to be edited using specialized softwares. *Adobe Premiere* is a popular mid range non-linear video editing application that provides some postproduction facilities on desktop platform. Three windows are used by Premiere namely *project*, *timeline* and *monitor*. Project window is used for importing and displaying raw video and audio clips and still images with all relevant information. The timeline window provides a visual display of the linear extent of the completed movie showing the order of its component clips. It uses multiple audio and video tracks for transitions and overlays. Audio and video clips can be dragged from the project window and dropped on the timeline for assembly.





NOTES

Screen above shows an Adobe Premiere 5.1: timeline, project and monitor window.

The monitor window is used for editing and viewing the video frames. Editing includes trimming, overlaying, applying effects, such as dissolve, wipes, spins, page turns, etc., for transition of one clip to another. Serious post production operations (changes) like color and contrast corrections, blurring or sharpening of images, element insertion and compositing, applying filter to a clip and vary it over time, sophisticated interpolation between key frames and so on can be done with more control and perfection using dedicated post production softwares, such as *Adobe After Effects*.

9.3 VIDEO COMPRESSION BASED ON MOTION COMPENSATION

MPEG (Moving Picture Experts Grop) is the international standard for audio and video digital compression and MPEG-1 is most relevant for video at low data rate (upto 1.5 M bit/s) to be incorporated in multimedia. MPEG-1 is standard with five parts, namely—systems, video and audio, conformance testing and software simulation (a full C-language implementation of the MPEG-1 encoder and decoder). Though higher standards like MPEG-2, MPEG-4, MPEG-7 and MPEG-21 have evolved in search of a higher compression ratio, better video quality, effective communication and technological upgradation, you will study MPEG-1 only for understanding of the basic MPEG scheme.

9.3.1 MPEG-1

MPEG-1 standard does not actually explain a compression algorithm but it defines a datastream syntax and a decompressor. The datastream architecture is based on a sequence of frames, each of which contains the data that is needed to create a single displayed image. There are four different kinds of frames (depending on how each image is to be decoded, which are as follows:

1. ***I-frames*** (intra-coded images): They are self-contained and are coded without any reference to other images. These frames are spatially compressed using a transform-coding method similar to JPEG. The compression ratio for I–frames is the lowest within MPEG. An I–frame must exist at the start of any video stream and also at any random access entry-point in the stream.

2. ***P-frames*** (predictive-coded images): They are compressed images resulting from the removal of temporal redundancy between successive frames. These frames are coded by a forward predictive-coding method in which the target macroblocks are predicted from most similar reference macroblocks in the preceding I or P–frame. Only the difference between the spatial location of the macroblocks, that is, the motion vector and the difference in the content of the macroblocks are coded. Instead of the difference, a macroblock itself is coded as non-motion compensated macroblock when a good match as reference macroblock is not found. P-frames usually achieve a large compression ratio (three times as much in I–frames).
3. ***B-frames*** (bi-directionally predictive-coded frames): They are coded by interpolation between two macroblocks—one from forward prediction (from previous I or P-frame) and the other from backward prediction (from future I or P-frame). Interpolative motion compensation is used here. If matching in both directions is successful, two motion vectors will be sent and the two corresponding matching macroblocks will be averaged (interpolated) for comparing to the target macroblock in order to generate the difference macroblock. If an acceptable match can be found in only one of the reference frames, then only one motion vector and its corresponding macroblocks and used for generating the difference macroblock. Maximum compression ratio (one and half times as much as in P-frame) is achieved in B-frames.
4. ***D-frames*** (DC-coded frames): They are intraframe-coded and are used for fast forward or fast-rewind modes.

9.3.2 Digital Video Interface Technology

The Digital Display Working Group created the Digital Video Interface (DVI) technology that accommodates for interconnecting the single connector to both digital and analog interfaces. The DVI technology supports high-speed connection (digital) to display the videos and animations. It facilitates a common interface for all the display devices. The high-speed digital signals in DVI technology support up to 350 Mpixels/sec as well as SVGA. The plug and play service is supported via hot-plug detection in this interface. It is frequently being used in LCD displays. The nature of operation of this technology is implemented in the pure digital form. Therefore, intermediate analog signals are not needed to use the DVI technology. It encounters analog-to-digital conversion to remove the synchronization of the operations and aliasing problems. Conceptually, DVI implies a collection of video interface through which LCD flat monitors get good quality modern video-graphics cards. For this, DVI cables are being used including both VGA and DVI output port (Refer Figure 9.1)

NOTES

Fig. 9.1 *Digital Video Interface*

The technology behind DVI uses Transmitting Minimized Differential Signaling (TMDS) to transmit the desired data over DVI connection. One TMDS is linked to transmit the data but sometimes dual links and two TMDS channels are also preferred to link if it is assembled to the system unit. The three data channels (RGB fiber optics) are maintained by a single link in which one channel (clock control channel) is used to access the demanded video even if the dual link is connected to the system unit (Figure 9.2).

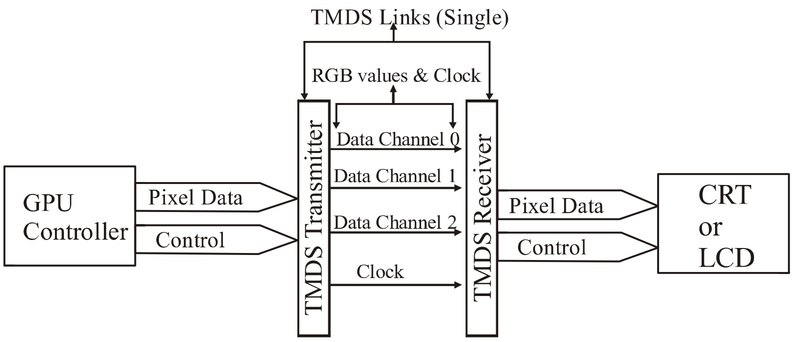


Fig. 9.2 *Data Channels Linked with DVI Connection*

In Figure 9.2, the Graphics Processing Unit (GPU) controller is used to send the pixel data that is passed to various data channels. The data channels are named data channel 0, data channel 1 and data channel 2 that are bound with time constraints. After controlling and checking, each pixel is used in graphics and the specific data are sent to the CRT or LCD monitors to be processed that are to be viewed by the customers/viewers. The TDMS link of 10-bit is generally operated till 165 MHz. It can also be linked up to 1.65 Gbps of bandwidth. A digital flat panel is displayed by 1920×1080 screen resolution and the electric power is generated at 60 Hz. Basically, this flat panel supports dual-link TMDS because dual link supports 2Gbps of bandwidth and is operated to match every second link to the previous one. The dual TMDS uses 2048 × 1536 screen resolution to achieve better graphics for the DVI technology.

9.3.3 Working of DVI

The system unit (PC/VGA/CRT) creates and transmits the video signal in the form of digital signals (0 and 1). The digital CRT monitor displays the analog signals but the video card of the system unit (VGA connection) converts the digital signals to analog ones to display the video data. The role of DVI technology is important at this step because the LCD monitor uses the graphics interpreter with the help of the DVI connection and changes analog signals back to digital ones.

In Figure 9.3, most of the video cards are used with DVI technology to convert the digital-to-analog signal and then convert analog-to-digital signal in the LCD display unit. The data in the PC is sent to the processing of electronic signals and then it appears on the LCD monitor.

NOTES

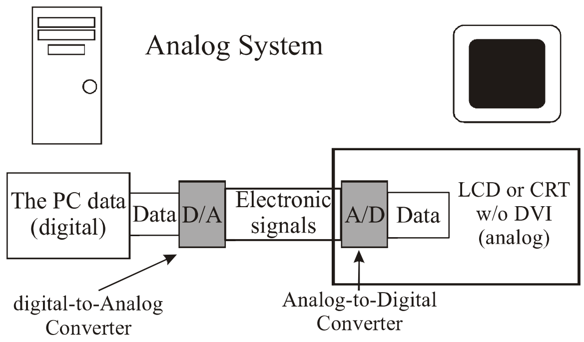


Fig. 9.3 *LCD Monitor with DVI Technology*

The image quality is not good and the display unit shows a lower resolution. The function of DVI is to remove the bad quality of graphics that appear in the video or animation.

9.3.4 Types of DVI

The DVI represents the mode of video interface technology, which is used to optimize Digital Flat Panel (DFP) standards. The types of DVI connections are as follows:

1. DVI-D (true digital video)

In the DVI-D system, the cables are directly connected to the source video, for example, video cards and digital LCD monitors. It offers a high quality image in comparison to the analog image. The analog signal is sent to the monitor and then changed to the digital format. It enhances the source connection and omits the process of analog conversion.

2. DVI-A (high-resolution analog)

In the DVI-A system, the cables are used to transmit the DVI signal to display in the analog format for the LCDs and CRT monitors.

3. DVI-I (high-resolution analog)

In this system, the cables are integrated and are used to transmit the digital source signal to a digital display or it can change the analog source to an analog display. It makes the DVI-I connection flexible and better than other DVI technologies.

NOTES

9.3.5 Features of DVI Technology

The features of DVI technology are as follows:

1. It uses proprietary chips and the data compression method to create a form of multimedia that is to be integrated into the desktop system unit.
2. It is helpful to play back the full motion videos, multiple stereo sound tracks, live television shows, colour graphics, etc.
3. It incorporates and stores DVI into the storage devices for desktop system unit and Winchester hard drives.
4. It plugs the interface boards to distribute the available expression slot on the motherboard and then installs the software.
5. It provides an enhanced technique manifesting the text-oriented e-mail and marks the sender of the message with the motion video communication.
6. It supervises the information tracking to transmit the videographic instructions for subordinating the messages sent, as per VoD technology.
7. It is used to send and receive the applications of videographic presentations, videographic voice mail, audio–visual databases, audio–visual references, sales messages, etc.

Check Your Progress

1. What is known as spatial compression?
2. Define the term reference frame.
3. Define the term MPEG.
4. State the purpose of I frames.
5. What is D-frames?
6. List the name of the technology used behind DVI.
7. What are DVI-A systems?
8. Define the term Adobe Premiere.

9.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Spatial redundancy is removed by compressing each individual image frame in isolation and the techniques used are generally called spatial compression or intra frame compression.
2. Each macroblock of the target frame is examined with reference to most similar macroblocks in previous and/or next frame called reference frame.
3. MPEG (Moving Picture Experts Group) is the international standard for audio and video digital compression and MPEG-1 is most relevant for video at low data rate (upto 1.5 M bit/s) to be incorporated in multimedia.

NOTES

4. I frame are self-contained and are coded without any reference to other images. These frames are spatially compressed using a transform-coding method like JPEG.
5. D- frames are intraframe-coded and are used for fast forward or fast-rewind modes.
6. The technology behind DVI uses Transmitting Minimized Differential Signalling (TMDS) to transmit the desired data over DVI connection. One TMDS is linked to transmit the data but sometimes dual links and two TMDS channels are also preferred to link if it is assembled to the system unit.
7. In the DVI-A system, the cables are used to transmit the DVI signal to display in the analog format for the LCDs and CRT monitors.
8. Adobe Premiere is a popular mid-range nonlinear video editing application that provides some postproduction facilities on desktop platform.

9.5 SUMMARY

- Spatial redundancy is removed by compressing each individual image frame in isolation and the techniques used are generally called spatial compression or intra frame compression.
- Temporal redundancy is removed by storing only the differences of subsequent frames instead of compressing each frame independently and the technique is known as temporal compression or inter frame compression.
- Spatial compression applies different lossless and lossy method same as those applied for still images.
- The most simplistic approach for temporal compression is to perform a pixel-by-pixel comparison (subtraction) between two consecutive frames.
- The current image frame is referred to as target frame. Each macroblock of the target frame is examined with reference to most similar macroblocks in previous and/or next frame called reference frame.
- Adobe Premiere is a popular mid-range nonlinear video editing application that provides some postproduction facilities on desktop platform.
- Audio and video clips can be dragged from the project window and dropped on the timeline for assembly.
- The monitor window is used for editing and viewing the video frames. Editing includes trimming, overlaying, applying effects, such as dissolve, wipes, spins, page turns, etc., for transition of one clip to another.
- MPEG (Moving Picture Experts Group) is the international standard for audio and video digital compression and MPEG-1 is most relevant for video at low data rate (up to 1.5 M bit/s) to be incorporated in multimedia.

NOTES

- MPEG-1 is standard with five parts, namely—systems, video and audio, conformance testing and software simulation (a full C-language implementation of the MPEG-1 encoder and decoder).
- MPEG-1 standard does not actually explain a compression algorithm, but it defines a datastream syntax and a decompressor.
- The datastream architecture is based on a sequence of frames, each of which contains the data that is needed to create a single displayed image.
- I-frames (intra-coded images) are self-contained and are coded without any reference to other images. These frames are spatially compressed using a transform-coding method like JPEG.
- P-frames (predictive-coded images) are compressed images resulting from the removal of temporal redundancy between successive frames.
- B-frames (bi-directionally predictive-coded frames) are coded by interpolation between two macroblocks—one from forward prediction (from previous I or P-frame) and the other from backward prediction (from future I or P-frame).
- D-frames (DC-coded frames) are intraframe-coded and are used for fast forward or fast-rewind modes.
- The Digital Display Working Group created the Digital Video Interface (DVI) technology that accommodates for interconnecting the single connector to both digital and analog interfaces.
- The DVI technology supports high-speed connection (digital) to display the videos and animations. It facilitates a common interface for all the display devices.
- The plug and play service is supported via hot-plug detection in this interface. It is frequently being used in LCD displays.
- The technology behind DVI uses Transmitting Minimized Differential Signalling (TMDS) to transmit the desired data over DVI connection. One TMDS is linked to transmit the data but sometimes dual links and two TMDS channels are also preferred to link if it is assembled to the system unit.
- The Graphics Processing Unit (GPU) controller is used to send the pixel data that is passed to various data channels.
- The data channels are named data channel 0, data channel 1 and data channel 2 that are bound with time constraints.
- The system unit (PC/VGA/CRT) creates and transmits the video signal in the form of digital signals (0 and 1).
- The DVI represents the mode of video interface technology, which is used to optimize Digital Flat Panel (DFP) standards.
- In the DVI-D system, the cables are directly connected to the source video, for example, video cards and digital LCD monitors. It offers a high-quality image in comparison to the analog image.

- In the DVI-A system, the cables are used to transmit the DVI signal to display in the analog format for the LCDs and CRT monitors.
- DVI technology uses proprietary chips and the data compression method to create a form of multimedia that is to be integrated into the desktop system unit.
- DVI technology incorporates and stores DVI into the storage devices for desktop system unit and Winchester hard drives.
- DVI technology provides an enhanced technique manifesting the text-oriented e-mail and marks the sender of the message with the motion video communication.
- DVI technology supervises the information tracking to transmit the video graphic instructions for subordinating the messages sent, as per VoD technology.

9.6 KEY WORDS

- **Spatial compression:** Spatial redundancy is removed by compressing each individual image frame in isolation and the techniques used are generally called spatial compression.
- **Adobe Premiere:** Adobe Premiere is a popular mid-range nonlinear video editing application that provides some postproduction facilities on desktop platform.
- **I-frames (intra-coded images):** They are self-contained and are coded without any reference to other images.
- **P-frames (predictive-coded images):** They are compressed images resulting from the removal of temporal redundancy between successive frames.
- **B-frames (bi-directionally predictive-coded frames):** They are coded by interpolation between two macroblocks—one from forward prediction (from previous I or P-frame) and the other from backward prediction (from future I or P-frame).
- **D-frames (DC-coded frames):** They are intraframe-coded and are used for fast forward or fast-rewind modes.
- **DVI-D (true digital video):** In the DVI-D system, the cables are directly connected to the source video, for example, video cards and digital LCD monitors.
- **DVI-A (high-resolution analog):** In the DVI-A system, the cables are used to transmit the DVI signal to display in the analog format for the LCDs and CRT monitors.
- **DVI-I (high-resolution analog):** In this system, the cables are integrated and are used to transmit the digital source signal to a digital display, or it can change the analog source to an analog display.

9.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

NOTES

Short-Answer Questions

1. Define the term temporal compression.
2. List the methods of spatial compression.
3. What is a target frame?
4. State the significance of MPEG-1.
5. How is I-frame different from P-frame?
6. What is a DVI technology?
7. State the significance of Transmitting Minimized Differential Signalling (TMDS).
8. List all the features of DVI technology.

Long-Answer Questions

1. Briefly discuss why video compression is important with the help of an example.
2. Elaborate on the characteristics of spatial redundancy and list it's all the methods.
3. Discuss in detail about the Adobe Premiere listing their objectives.
4. Explain the terms MPEG (Moving Picture Experts Group) giving all appropriate examples.
5. Give an overview on the MPEG-1 listing all its frames in brief.
6. Explain the term DVI along with the working of DVI.
7. Elaborate the types of all DVI with the help of appropriate examples.
8. Briefly explain the features of DVI technology.

9.8 FURTHER READINGS

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

UNIT 10 SEARCH FOR MOTION VECTORS

NOTES

Structure

10.0 INTRODUCTION

The Moving Picture Experts Group (MPEG) is an alliance of working groups of ISO and IEC that sets standards for media coding, including compression coding of audio, video, graphics and genomic data, and transmission and file formats for various applications. Together with the JPEG group, MPEG is organized under ISO/IEC JTC 1/SC 29. MPEG formats are used in various multimedia systems. The most well-known older MPEG media formats typically use MPEG-1, MPEG-2, and MPEG-4 AVC media coding and MPEG-2 systems transport streams and program streams. Newer systems typically use the MPEG base media file format and dynamic streaming.

Encapsulated PostScript (EPS) is a Document Structuring Conventions conforming (DSC) PostScript document format usable as a graphics file format. EPS files are more-or-less self-contained, reasonably predictable PostScript documents that describe an image or drawing and can be placed within another PostScript document.

Tag Image File Format, abbreviated TIFF or TIF, is a computer file format for storing raster graphics images, popular among graphic artists, the publishing industry, and photographers. TIFF is widely supported by scanning, faxing, word processing, optical character recognition, image manipulation, desktop publishing, and page-layout applications.

In this unit, you will study about the MPEG or Moving Picture Experts Group, different kinds of frame of MPEG, encapsulated PostScript and tagged image file format.

10.1 OBJECTIVES

After going through this unit, you will be able to:

NOTES

- Analyse the significance of MPEG
- Interpret the different kinds of frame of MPEG
- Understand the encapsulated PostScript
- Know about the tagged image file format

10.2 MPEG

MPEG or Moving Picture Experts Group is an international standard for audio/video digital compression; MPEG-1 is the most relevant for video compression at low data rate (upto 1.5 M bit/s). This standard has five parts, namely – systems, video and audio, conformance testing and software simulation. Though higher standards, for example, MPEG-2, MPEG-4, MPEG-7 and MPEG-21 have evolved in search of higher compression ratio, better video quality, effective communication and technological upgradation.

The datastream architecture is based on a sequence of frames, each of which contains the data needed to create a single displayed image. The four different kinds of frames, depending on how each image is to be decoded are as follows:

- **I-frames or Intra Coded Images:** These images are self-contained, i.e., coded without any reference to other images. These frames are spatially compressed using a transform coding method similar to JPEG. The compression ratio for I-frames is the lowest within MPEG. An I-frame must exist at the start of a video stream and also at any random access entry point in the stream.
- **P-Frames or Predictive Coded Images:** These compressed images result from removal of temporal redundancy between successive frames. These frames are coded by a forward predictive coding method in which the target macroblocks are predicted from the most similar reference macroblocks in the preceding I or P-frame. Only the difference between the spatial location of the macroblocks, i.e., the motion vector and the difference in content of the macroblocks are coded. A macroblock itself is coded as non-motion compensated macroblock when a good match as reference macroblock is not found. Usually, in P-frames a large compression ratio (three times as much in I-frames) is achieved.
- **B-Frames or Bi-Directionally Predictive Coded Frames:** These are coded by interpolating between two macroblocks – one from forward prediction (from previous I or P-frame) and the other from backward prediction (from future I or P-frame). Interpolative motion compensation is

used here. If matching in both directions is successful two motion vectors are sent and the two corresponding matching macroblocks are averaged (interpolated) and compared to the target macroblock for generating the difference macroblock. If an acceptable match is found in only one of the reference frames then only one motion vector and its corresponding macroblock is used for generating the difference macroblock. Maximum compression ratio (one and half times as much as in P-frame) is achieved in B-frames.

- **D-Frames or DC Coded Frames:** These are intraframe coded and are used for fast forward or fast rewind modes.

Encapsulated PostScript: Encapsulated PostScript (EPS) is an image file format used for both vector graphics and bitmaps. The EPS files have a PostScript description of the graphic data within them. The EPS files are exclusive in that the graphics users use them for bitmap images, vector graphics, type or even entire pages.

Tagged Image File Format: Tagged Image File Format (TIFF) files are used for bitmap format only. The TIFF formats are the files that are supported by virtually all graphics applications.

Check Your Progress

1. What is the full form of MPEG?
2. What is the data stream architecture based on?
3. Define the term Tagged Image File Format (TIFF).
4. State the significance of P-frames.
5. What is the B-frames also called?

10.3 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. MPEG full form is Moving Picture Experts Group.
2. The data stream architecture is based on a sequence of frames, each of which contains the data needed to create a single displayed image.
3. Tagged Image File Format (TIFF) files are used for bitmap format only. The TIFF formats are the files that are supported by virtually all graphics applications.
4. P-frames are coded by a forward predictive coding method in which the target macroblocks are predicted from the most similar reference macroblocks in the preceding I or P-frame.
5. B-frames are known as Bi-Directionally Predictive Coded Frame.

10.4 SUMMARY

NOTES

- MPEG or Moving Picture Experts Group is an international standard for audio/video digital compression; MPEG-1 is the most relevant for video compression at low data rate (upto 1.5 M bit/s).
- MPEG standard has five parts, namely – systems, video and audio, conformance testing and software simulation.
- The data stream architecture is based on a sequence of frames, each of which contains the data needed to create a single displayed image.
- I-frame images are self-contained, i.e., coded without any reference to other images. These frames are spatially compressed using a transform coding method like JPEG.
- The compression ratio for I-frames is the lowest within MPEG. An I-frame must exist at the start of a video stream and at any random-access entry point in the stream.
- P-frames compressed images result from removal of temporal redundancy between successive frames.
- A macroblock itself is coded as non-motion compensated macroblock when a good match as reference macroblock is not found.
- B-frames are coded by interpolating between two macroblocks – one from forward prediction (from previous I or P-frame) and the other from backward prediction (from future I or P-frame).
- D-frames are intraframe coded and are used for fast forward or fast rewind modes.
- Encapsulated PostScript (EPS) is an image file format used for both vector graphics and bitmaps.
- The EPS files have a PostScript description of the graphic data within them. The EPS files are exclusive in that the graphics users use them for bitmap images, vector graphics, type, or even entire pages.
- Tagged Image File Format (TIFF) files are used for bitmap format only.
- The TIFF formats are the files that are supported by virtually all graphics applications.

10.5 KEY WORDS

- **MPEG:** Moving Picture Experts Group is an international standard for audio/video digital compression; MPEG-1 is the most relevant for video compression at low data rate (upto 1.5 M bit/s).

- **I-frames or Intra Coded Images:** These images are self-contained, i.e., coded without any reference to other images. These frames are spatially compressed using a transform coding method like JPEG.
- **P-Frames or Predictive Coded Images:** These compressed images result from removal of temporal redundancy between successive frames.
- **B-Frames or Bi-Directionally Predictive Coded Frames:** These are coded by interpolating between two macroblocks one from forward prediction from previous I or P-frame and the other from backward prediction from future I or P-frame.
- **D-Frames or DC Coded Frames:** These are intraframe coded and are used for fast forward or fast rewind modes.
- **Encapsulated PostScript:** Encapsulated PostScript (EPS) is an image file format used for both vector graphics and bitmaps.
- **Tagged Image File Format:** Tagged Image File Format (TIFF) files are used for bitmap format only.

NOTES

10.6 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is known as MPEG?
2. What are I-frames or intra coded images?
3. Define the term D-frames.
4. State the use of Encapsulated PostScript (EPS).

Long-Answer Questions

1. Elaborate on the standards of MPEG giving suitable examples.
2. Discuss in detail about the MPEG or Motion Picture Expert Group listing their type of frames.

10.7 FURTHER READINGS

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

UNIT 11 BASIC AUDIO COMPRESSION

NOTES

Structure

11.0 INTRODUCTION

In signal processing, data compression, source coding, or bit-rate reduction is the process of encoding information using fewer bits than the original representation. Any compression is either lossy or lossless. Lossless compression reduces bits by identifying and eliminating statistical redundancy. No information is lost in lossless compression. Lossy compression reduces bits by removing unnecessary or less important information.

Audio compression is a type of lossy or lossless compression in which the amount of data in a recorded waveform is reduced to differing extents for transmission respectively with or without some loss of quality, used in CD and MP3 encoding, Internet radio, and the like. Lossy compression is most used to compress multimedia data (audio, video, and images), especially in applications such as streaming media and internet telephony. By contrast, lossless compression is typically required for text and data files, such as bank records and text articles. It can be advantageous to make a master lossless file which can then be used to produce additional copies from. This allows one to avoid basing new compressed copies off a lossy source file, which would yield additional artifacts and further unnecessary information loss. Lossless compression is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data. By contrast, lossy compression permits reconstruction only of an approximation of the original data, though usually with greatly improved compression rates (and therefore reduced media sizes). Lossless compression is used in cases where it is important that the original and the decompressed data be identical, or where deviations from the original data would be unfavourable. Typical examples are executable programs, text documents, and source code. Some image file formats, like PNG or GIF, use only lossless compression, while others like

NOTES

TIFF and MNG may use either lossless or lossy methods. Lossless audio formats are most often used for archiving or production purposes, while smaller lossy audio files are typically used on portable players and in other cases where storage space is limited, or exact replication of the audio is unnecessary.

In this unit, you will study about the multimedia compression technique, characteristics of audio compression, pulse code modulation, companding technique, lossy and lossless compression technique.

11.1 OBJECTIVES

After going through this unit, you will be able to:

- Learn about the multimedia compression technology
- Discuss the advantages and limitations of multimedia compression
- Understand the characteristic of audio compression
- Explain the prime feature of pulse code modulation
- Discuss the coding techniques of pulse code modulation
- Know about the companding technique and its uses
- Comprehend about the different between lossy and lossless compression
- Recognise the various compression techniques and methods used in compression

11.2 BASIC AUDIO COMPRESSION TECHNIQUES

Multimedia compression technology employs the tools and techniques in order to reduce the file size of various media formats. With the development of the Internet services the importance of compress algorithm is highlighted because it performs fast in networks due to its high file size. With the popularity of voice and video conferencing over the Internet, compression method for multimedia is preferred for unreliable network infrastructure. For example, the audio and video files are compressed to provide a quicker and a better download option. After finishing the operations of recording and editing the audio file, you need to create a data compressed audio file in order to deliver it over the Internet. Microsoft Video For Windows (MVFW) software is required to capture and compress the video which requires the following codec:

- **Cinepak:** This codec is also known as compact video codec. It produces high quality audio and video content and hence takes more time to compress the files.
- **Microsoft Video I:** Once video is captured, this software is used to correct the video streaming file.
- **Indeo:** This software is used for real time compression.

NOTES

- **Discrete Cosine Transforms:** Discrete Cosine Transforms (DCT) compress the selected image into 8×8 pixel blocks.
- **Discrete Wavelet Transform:** Discrete Wavelet Transform (DWT) is used to compress professional level quality.
- **Fractal:** It is used to compress the file in asymmetrical way, such as for seven hours clip this process takes 30 seconds to clip the file. But, resolution of image and picture is produced high by this technique.

A multimedia compression system is used to generate the frame rate scalable data. It includes all relevant characteristics of the data, for example frame rate, resolution and quality for video. The scalable data generated by the compression system includes multiple additive layers for each characteristic across which the data is scalable. For video, the frame rate layers are additive temporal layers, the resolution layers are additive base and enhancement layers, and the quality layers are additive index planes of embedded codes. Various techniques can be used for generating these layers, for example, Laplacian pyramid decomposition or wavelet decomposition is used for generating the resolution layers, and tree structured vector quantization or tree structured scalar quantization for generating the quality layers for selected graphics in the multimedia stream file. The system further provides for embedded inter frame compression in the context of frame rate scalability and non-redundant layered multicast network delivery of the scalable data.

11.2.1 Audio Compresssion

The uncompressed data rate increases as more bits are used for quantization. Stereo information as opposed to mono, doubles the amount of bandwidth needed to transmit a digital audio signal. Table 11.1 shows how audio quality is related to data rate and bandwidth.

Table 11.1 *Data Rate and Bandwidth of Audio*

Quality	Sample Rate (kHz)	Bits per Sample	Mono/Stereo (kB/sec)	Data Rate (uncompressed)	Frequency Band (Hz)
Telephone	8	8	Mono	8	200 – 3400
FM Radio	22.05	16	Stereo	88.2	20 – 11000
CD	44.1	16	Stereo	176.4	5 – 20000
DAT	48	16	Stereo	192.0	5 – 20000
DVD Audio	192 (max)	24 (max)	6 channels (max)	1200.0 (max)	0 – 96000 (max)

One minute of recorded sound (mono) on CD means that 60 × 44,100 samples are used. That is 26,46,000 samples and with each sample consisting of 16-bits that makes 52,92,000 bytes, i.e., around 5 MB. So, 1 minute of stereo sound would require around 10 MB of storage on CD. Using effective compression techniques storage requirement of audio files can be reduced to a great extent.

NOTES

PCM **or** ***Pulse Code Modulation*** is the formal term for audio digitization method that includes sampling and quantization. The resulting digitized samples can be thought of as infinitely narrow vertical 'pulses'.

PCM basically enables other coding techniques discussed below:

DPCM **or** ***Differential Pulse Code Modulation*** technique is based on storing the difference between consecutive samples instead of absolute value of the sample in fewer bits. However, unlike video, audio waveform can change rapidly and hence there is lesser temporal redundancy which means difference signal value is also higher.

DPCM actually employs a predictive coding scheme followed by quantization. It computes a predicted value for a sample based on preceding samples and stores the prediction error, i.e., the difference between the prediction and actual value.

ADPCM **or** ***Adaptive DPCM*** obtains further compression by dynamically varying the step size used to quantize signal differences. Large differences are quantized using large steps, small differences using small steps, so the amount of details stored is scaled or adapted to the size of the difference. Sometimes the predictor coefficients are adaptively modified in the process known as *Adaptive Predictive Coding* or *APC*.

DM **or** ***Delta Modulation*** is a simplified version of DPCM which uses only single quantized error value. It works well with more or less constant audio signals.

LPC **or** ***Linear Predictive Coding*** is a typical voice coding or vocoding technique that actually generates a set of parameters modelling the shape and excitation of the vocal tract, not actual signals or differences. It is like (MIDI) where the description of the signal is sent rather than the signal itself.

Companding is an effective speech compression technique mostly used in the telecommunications. It employs nonlinear quantization levels with higher levels spaced further apart than the lower ones, so that quiet sounds are represented in greater detail than the louder ones. This technique exploits the statistics that we generally hear more in the low volume range. Companding produces compression as fewer bits are needed to represent the full range of input signal values than a linear quantization scheme would require. When the signal is reconstructed it is passed through an expander circuit that reverses the compression algorithm. First compressing then expanding taken together is termed as companding. A common non-linear quantization function is the μ-law given by,

$$y = \frac{\log(1 + \mu x)}{\log(1 + \mu)} \text{ for } x \geq 0$$

where, x is the input value, y is the output and μ is the parameter that determines the amount of companding.

NOTES

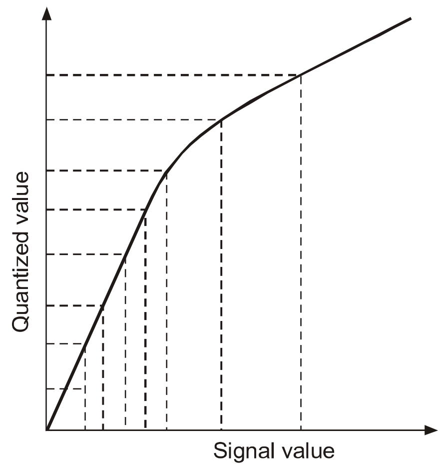


Fig. 11.1 *A Non-Linear Quantization Curve*

Audio compression is also achieved by employing variable length *Huffman Coding* technique whereby short codes are assigned to smaller size.

Silence or near silence can be coded in a speech compression technique applying the concept of *run length encoding*. Only the total length (time duration) is often recorded of samples falling below a threshold level as if they are all zero value instead of recording each such sample individually. This is, however, equivalent to applying noise gate and not strictly lossless in nature.

The World Wide Web (WWW) is capable of using most forms of media in combination. The addition of graphics, movies, sound and animation has transformed the Web.

11.2.2 Lossy and Lossless Compression

Computers can work with art, photographs, videos and sounds but they can do so only when these multimedia resources are stored in digitized files. These files require huge amounts of storage space. For example, an average sized hard disk could store only a dozen audio Compact Disks (CDs) and there would not be much memory space left for system software or applications. To reduce the size of multimedia files, most software use compression or decompression algorithms called *codecs* which use two different approaches to compress the multimedia streaming files namely *lossless compression* and *lossy compression*. In lossless compression, the original file is compressed so that it can be completely restored without flaw when decompression occurs. In lossy compression, the original file is processed so that some information is permanently removed. Lossy compression techniques eliminate information that is not perceived when people see pictures or hear sounds. Joint Photographic Experts Group (JPEG) is designed for compressing full color, gray scale images or continuous tone artwork. Any smooth variation in color, such as occurring in highlighted or shaded areas will be represented more faithfully and in less space by JPEG than by GIF. Plain black and white images

NOTES

should never be converted to JPEG. The Graphics Interchange Format (GIF) does significantly better on images with only a few selected colors, such as line drawings and simple cartoons. There has to be at least 16 gray levels before JPEG is useful for gray scale images. JPEG is lossy. The decompressed image is not quite the same as the original. A lossless compression algorithm is one that guarantees its decompressed output to be bit for bit identical to the original input. This scheme does not discard any data during the encoding process while the lossy scheme throws useless data away during encoding. That is, in fact, how lossy schemes manage to obtain superior compression ratios over most lossless schemes. JPEG has been designed specifically to discard information that the human eye cannot easily see. Because the human eye is much more sensitive to brightness variations in gray scale than to color variations and JPEG can compress color data more heavily than brightness data. Gray scale images do not compress well by large factors. GIF is lossless for gray scale images up to 256 levels while JPEG is not. The more complex and subtly rendered image the more likely JPEG will do. Multimedia data can be compressed with the help of following algorithms:

- **Lossy Compression Algorithm:** It is the compression technique which will lose data in the original source while trying to keep the visible quality at the almost same amount. The compression ratio will be very high. Most probably the ratio will be a value near 10. It reduces non-sensitive information to the human eyes and the compressed media will not be the media that was available before compression. The advantage of lossy compression is that it can reduce the file size more than in the lossless compression. But, the original file cannot be taken after the decompression.
- **Lossless Compression Algorithm:** There will be no data loss in this type of compression as it is defined by the name. Both original data and the compressed data are the same in this compression. The algorithms for the compression and decompression are exact inverse of each other in the lossless compression. The main mechanism in this compression is to remove the redundant data in the compression and adding them in the decompression. The advantage of lossless compression is that the original format of the data remains even it is compressed. But, the reduction of the size of the data is small. Sometimes, the file size can be increased instead of decrease if compressed file is not required for application.

Compression Techniques

Any analog element or object that human being can understand has to be digitized first to make it understandable by computer since it recognizes every object or element or a character or a point as a stream of 0s and 1s. Computer can accept a signal only in digital format, process the same and produce results in a digital format that is again reconstructed to produce an analog presentation. Digitization offers many advantages as it is possible to manipulate these digitized streams and it can be compressed to occupy lesser space on the storage media. Also, since these are recorded on hard disk in which random access is possible both sequential and random access methods can be utilized. By using digital representation, matter can be stored for a very long period.

With the advent of digital computers, video signals are digitized and this is known as digital video. The term 'digital video' refers to a type of recording system in which digitized video signals are recorded. A video as seen by human being is an analog presentation that is captured by a camera, camcorder or video camera. In a camera, an image is captured and digitized in a format for storage on the storage media inside the camera. With the improvement in technology cheaper systems appeared that used compressed data; digital betacam was introduced by Sony.

Compression of Digital Video

Properties of digital video apply to videos that are uncompressed. Need of compression arises due to the fact that uncompressed videos carrying many details need quite high bit rate. By compression each frame in a video needs smaller number of bits by a factor known as *compression factor*.

Since a video is formed by several image frames at some rate (mostly 30 fps); video compression is achieved by image compression.

Image Compression

Image compression denotes compression of data for digital images. Objective of such compression is the reduction of redundancy of data pertaining to the image data that enables storage and transmission of data in an efficient form. But such reduction of data may lead to loss of some data pertaining to the image and for this reason, an image compression may be lossy.

Thus, there are two types of image compression techniques known as lossy and lossless. A lossy compression, when applied at low bit rate, is subjected to compression artifacts. Lossless compression is preferred in technical drawing, medical imaging, comics and icons.

Methods used in Image Compression

Following are the methods used in image compression

Lossless Compression: In this method, data is compressed in such a way that on decompression the display would be an exact replica of original data.

- **Run Length Encoding:** This is default method used in Picture Comment eXtreme (PCX) and one of the methods used in Bitmap (BMP), TIFF and Truvision Graphics Adapter (TGA) formats.
- **Entropy Encoding:** Differential pulses code module (DPCM), predictive coding and chain codes.
- **Adaptive Dictionary Algorithms:** Lempel-Ziv-Welch (LZW) is an algorithm used in TIFF and GIF image formats.
- **Deflation:** This method is used in PNG, TIFF and Multiple-image Network Graphic (MNG) formats.

Lossy Compression

File compression technique that lose some of the original data is called lossy compression. Following methods are used in lossy compression technique:

NOTES

- **Reduction of Color Space:** In lossy compression, color space is reduced to some most common colors of the image. Colors are specified in the header of compressed images' color palette. Every pixel makes a reference to the index of a color in this color palette. To avoid posterization this method is combined with dithering.
- **Chroma Subsampling:** This method makes the use of the fact that human eye has perception for spatial changes more in case of brightness in comparison to that of color. This averages or drops some of the information on chrominance of the image.
- **Transform Coding:** This method is most commonly used and in it a transform that is Fourier related is applied. This is followed by quantization and entropy coding.
- **Fractal Compression:** This technique works on the principle of self similarity.

The main objective of image compression is to produce the best quality of image at a given bit rate known as *compression rate*.

Check Your Progress

1. What does the multimedia compression technology employ?
2. State the feature of Cinepak.
3. Why is multimedia compression system used?
4. Define the term PCM or Pulse Code Modulation.
5. What is the significance of Joint Photographic Experts Group (JPEG)?
6. State whether JPED is lossy or lossless.
7. Define the term compression factor.
8. List the methods used in image compression.

11.3 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Multimedia compression technology employs the tools and techniques to reduce the file size of various media formats.
2. Cinepak codec is also known as compact video codec. It produces high quality audio and video content and hence takes more time to compress the files.

3. A multimedia compression system is used to generate the frame rate scalable data. It includes all relevant characteristics of the data, for example frame rate, resolution, and quality for video.
4. PCM or Pulse Code Modulation is the formal term for audio digitization method that includes sampling and quantization. The resulting digitized samples can be thought of as infinitely narrow vertical 'pulses.
5. Joint Photographic Experts Group (JPEG) is designed for compressing full color, gray scale images or continuous tone artwork. Any smooth variation in color, such as occurring in highlighted or shaded areas will be represented more faithfully and in less space by JPEG than by GIF.
6. JPEG is lossy. The decompressed image is not quite the same as the original. A lossless compression algorithm is one that guarantees its decompressed output to be bit for bit identical to the original input.
7. Compression each frame in a video needs smaller number of bits by a factor known as compression factor.
8. The two methods used in image compression are lossy compression and lossless compression.

11.4 SUMMARY

- Multimedia compression technology employs the tools and techniques to reduce the file size of various media formats.
- A multimedia compression system is used to generate the frame rate scalable data. It includes all relevant characteristics of the data, for example frame rate, resolution, and quality for video.
- The scalable data generated by the compression system includes multiple additive layers for each characteristic across which the data is scalable.
- Laplacian pyramid decomposition or wavelet decomposition is used for generating the resolution layers, and tree structured vector quantization or tree structured scalar quantization for generating the quality layers for selected graphics in the multimedia stream file.
- PCM or Pulse Code Modulation is the formal term for audio digitization method that includes sampling and quantization. The resulting digitized samples can be thought of as infinitely narrow vertical 'pulses'.
- DPCM or Differential Pulse Code Modulation technique is based on storing the difference between consecutive samples instead of absolute value of the sample in fewer bits.
- DM or Delta Modulation is a simplified version of DPCM which uses only single quantized error value. It works well with constant audio signals.
- LPC or Linear Predictive Coding is a typical voice coding or vocoding technique that generates a set of parameters modelling the shape and excitation of the vocal tract, not actual signals, or differences.

NOTES

- Companding is an effective speech compression technique mostly used in the telecommunications. It employs nonlinear quantization levels with higher levels spaced further apart than the lower ones, so that quiet sounds are represented in greater detail than the louder ones.
- Audio compression is also achieved by employing variable length Huffman Coding technique whereby short codes are assigned to smaller size.
- Silence or near silence can be coded in a speech compression technique applying the concept of run length encoding.
- The World Wide Web (WWW) can use most forms of media in combination. The addition of graphics, movies, sound, and animation has transformed the Web.
- To reduce the size of multimedia files, most software use compression or decompression algorithms called codecs which use two different approaches to compress the multimedia streaming files namely lossless compression and lossy compression.
- In lossless compression, the original file is compressed so that it can be completely restored without flaw when decompression occurs.
- In lossy compression, the original file is processed so that some information is permanently removed. Lossy compression techniques eliminate information that is not perceived when people see pictures or hear sounds.
- Joint Photographic Experts Group (JPEG) is designed for compressing full color, gray scale images or continuous tone artwork.
- JPEG is lossy. The decompressed image is not quite the same as the original. A lossless compression algorithm is one that guarantees its decompressed output to be bit for bit identical to the original input.
- Lossy compression algorithm is the compression technique which will lose data in the original source while trying to keep the visible quality at the almost same amount.
- The advantage of lossy compression is that it can reduce the file size more than in the lossless compression. But the original file cannot be taken after the decompression.
- Any analog element or object that human being can understand has to be digitized first to make it understandable by computer since it recognizes every object or element or a character or a point as a stream of 0s and 1s.
- Digitization offers many advantages as it is possible to manipulate these digitized streams and it can be compressed to occupy lesser space on the storage media.
- Image compression denotes compression of data for digital images. Objective of such compression is the reduction of redundancy of data pertaining to the image data that enables storage and transmission of data in an efficient form.

- A lossy compression, when applied at low bit rate, is subjected to compression artifacts. Lossless compression is preferred in technical drawing, medical imaging, comics, and icons.
- In lossy compression, color space is reduced to some most common colors of the image. Colors are specified in the header of compressed images' color palette.

11.5 KEY WORDS

- **Cinepak:** It produces high quality audio and video content and hence takes more time to compress the files.
- **Microsoft Video I:** Once video is captured, this software is used to correct the video streaming file.
- **Indeo:** This software is used for real time compression.
- **Discrete Wavelet Transform (DWT):** Discrete Wavelet Transform (DWT) is used to compress professional level quality.
- **Fractal:** It is used to compress the file in asymmetrical way, such as for seven hours clip this process takes 30 seconds to clip the file.
- **PCM or Pulse Code Modulation:** It is the formal term for audio digitization method that includes sampling and quantization.
- **Lossy compression algorithm:** It is the compression technique which will lose data in the original source while trying to keep the visible quality at the almost same amount.
- **Lossless compression algorithm:** There will be no data loss in this type of compression and both original data and the compressed data are the same in this compression.
- **Lossless compression:** In this method, data is compressed in such a way that on decompression the display would be an exact replica of original data.
- **Chroma subsampling:** This method makes the use of the fact that human eye has perception for spatial changes more in case of brightness in comparison to that of color.

11.6 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define the term Microsoft Video for Windows (MVFW).
2. What is the role of fractal?
3. What is the purpose of ADPCM or Adaptive DPCM?

NOTES

4. What is the Huffman coding technique?
5. List the significance of a non-linear quantization curve.
6. State the difference between lossless compression and lossy compression.
7. Why should the plain black and white images never be converted into JPEG?
8. What is lossy compression algorithm?
9. What is run length encoding?
10. List any one method used in lossy compression technique.

Long-Answer Questions

1. Briefly discuss why multimedia compression technology is known as business tool with the help of an example.
2. Elaborate on the Microsoft Video for Windows (MVFW) and its codec required to compress the video.
3. Discuss in detail about the Laplacian pyramid decomposition used giving suitable examples.
4. Explain the term PCM or Pulse Code Modulation and its other coding techniques.
5. Give an overview on companding with its mathematical formulae.
6. Explain the terms lossy and lossless compression with the help of examples.
7. Elaborate on the algorithms which can help to compress the multimedia data.
8. Briefly explain the concept of image compression and list all the methods used in image compression.

11.7 FURTHER READINGS

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

BLOCK - V
MULTIMEDIA NETWORKS

NOTES

UNIT 12 MULTIMEDIA NETWORKS

Structure

12.0 INTRODUCTION

Multimedia is a form of communication that combines different content forms, such as text, audio, images, animations, or video into a single presentation, in contrast to traditional mass media, such as printed material or audio recordings. Popular examples of multimedia include video podcasts, audio slideshows and animated videos. Multimedia can be recorded for playback on computers, laptops, smartphones, and other electronic devices, either on demand or in real time (streaming). In the early years of multimedia, the term rich media was synonymous with interactive multimedia. Over time, hypermedia extensions brought multimedia to the World Wide Web.

Digital audio is a representation of sound recorded in, or converted into, digital form. In digital audio, the sound wave of the audio signal is typically encoded as numerical samples in a continuous sequence. Digital video is an electronic representation of moving visual images (video) in the form of encoded digital data. This contrasts with analog video, which represents moving visual images in the form of analog signals. Digital video comprises a series of digital images displayed in rapid succession. Networked multimedia is to build the multimedia on network and distributed systems, so different users on different machines can share image, sound, video, voice.

In this unit, you will study about the multimedia and its significance in networking, audio and video digitization, packet switching and circuit switching, LAN and its uses.

12.1 OBJECTIVES

NOTES

After going through this unit, you will be able to:

- Analyse the significance of multimedia
- Interpret the characteristics of audio digitization
- Discuss the advantages and limitations of video and its digitization
- Understand the multimedia networks
- Explain the prime feature of packet switching and circuit switching
- Know about the LAN and its use

12.2 BASICS OF MULTIMEDIA NETWORKS

Multimedia is a form of communication that combines different content forms such as text, audio, images, animations, or video into a single presentation, in contrast to traditional mass media, such as printed material or audio recordings. Popular examples of multimedia include video podcasts, audio slideshows and Animated videos. Multimedia can be recorded for playback on computers, laptops, smartphones, and other electronic devices, either on demand or in real time (streaming). In the early years of multimedia, the term "rich media" was synonymous with interactive multimedia. Over time, hypermedia extensions brought multimedia to the World Wide Web.

Multimedia may be broadly divided into linear and non-linear categories:

- Linear active content progresses often without any navigational control for the viewers, such as a cinema presentation.
- Non-linear uses interactivity to control progress as with a video game or self-paced computer-based training. Hypermedia is an example of non-linear content.

Multimedia presentations can be live or recorded:

- A recorded presentation may allow interactivity via a navigation system.
- A live multimedia presentation may allow interactivity via an interaction with the presenter or performer.

Audio and its Digitization

Digital audio is a representation of sound recorded in, or converted into, digital form. In digital audio, the sound wave of the audio signal is typically encoded as numerical samples in a continuous sequence. In a digital audio system, an analog electrical signal representing the sound is converted with an Analog-to-Digital Converter (ADC) into a digital signal, typically using Pulse-Code Modulation

NOTES

(PCM). This digital signal can then be recorded, edited, modified, and copied using computers, audio playback machines, and other digital tools. When the sound engineer wishes to listen to the recording on headphones or loudspeakers (or when a consumer wishes to listen to a digital sound file), a Digital-to-Analog Converter (DAC) performs the reverse process, converting a digital signal back into an analog signal, which is then sent through an audio power amplifier and ultimately to a loudspeaker. Digital audio systems may include compression, storage, processing, and transmission components. Conversion to a digital format allows convenient manipulation, storage, transmission, and retrieval of an audio signal. Unlike analog audio, in which making copies of a recording results in generation loss and degradation of signal quality, digital audio allows an infinite number of copies to be made without any degradation of signal quality. If an audio signal is analog, a digital audio system starts with an ADC that converts an analog signal to a digital signal. The ADC runs at a specified sampling rate and converts at a known bit resolution. CD audio, for example, has a sampling rate of 44.1 kHz (44,100 samples per second), and has 16-bit resolution for each stereo channel. Analog signals that have not already been bandlimited must be passed through an anti-aliasing filter before conversion, to prevent the aliasing distortion that is caused by audio signals with frequencies higher than the Nyquist frequency (half the sampling rate).

A digital audio signal may be stored or transmitted. Digital audio can be stored on a CD, a digital audio player, a hard drive, a USB flash drive, or any other digital data storage device. The digital signal may be altered through digital signal processing, where it may be filtered or have effects applied. Sample-rate conversion including up sampling and down sampling may be used to conform signals that have been encoded with a different sampling rate to a common sampling rate prior to processing. Audio data compression techniques, such as MP3, Advanced Audio Coding, Ogg Vorbis, or FLAC, are commonly employed to reduce the file size. Digital audio can be carried over digital audio interfaces, such as AES3 or MADI. Digital audio can be carried over a network using audio over Ethernet, audio over IP or other streaming media standards and systems. For playback, digital audio must be converted back to an analog signal with a DAC. According to the Nyquist–Shannon sampling theorem, with some practical and theoretical restrictions, a bandlimited version of the original analog signal can be accurately reconstructed from the digital signal.

Video and its Digitization

Digital video is an electronic representation of moving visual images (video) in the form of encoded digital data. This contrasts with analog video, which represents moving visual images in the form of analog signals. Digital video comprises a series of digital images displayed in rapid succession. Digital video can be copied and reproduced with no degradation in quality. In contrast, when analog sources are

NOTES

copied, they experience generation loss. Digital video can be stored on digital media such as Blu-ray Disc, on computer data storage, or streamed over the Internet to end users who watch content on a desktop computer screen or a digital smart TV. Today, digital video content, such as TV shows and movies also include a digital audio soundtrack.

- Each frame is divided in small grids, called pixels. For black and white TV, grey level of each pixel is represented by 8 bits.
- In case of color, each pixel is represented by 24 bits, 8-bit for each primary color R, G, B.

12.3 MULTIMEDIA NETWORK

Networked multimedia is to build the multimedia on network and distributed systems, so different users on different machines can share image, sound, video, voice, and many other features and to communicate with each under these tools.

Circuit Switching

Circuit switching is a method of implementing a telecommunications network in which two network nodes establish a dedicated communications channel (circuit) through the network before the nodes may communicate. The circuit guarantees the full bandwidth of the channel and remains connected for the duration of the communication session. The circuit functions as if the nodes were physically connected as with an electrical circuit. In circuit switching, the bit delay is constant during a connection (as opposed to packet switching, where packet queues may cause varying and potentially indefinitely long packet transfer delays). No circuit can be degraded by competing users because it is protected from use by other callers until the circuit is released and a new connection is set up. Even if no actual communication is taking place, the channel remains reserved and protected from competing users. While circuit switching is commonly used for connecting voice circuits, the concept of a dedicated path persisting between two communicating parties or nodes can be extended to signal content other than voice. The advantage of using circuit switching is that it provides for continuous transfer without the overhead associated with packets, making maximal use of available bandwidth for that communication. One disadvantage is that it can be relatively inefficient because unused capacity guaranteed to a connection cannot be used by other connections on the same network. In addition, calls cannot be established or will be dropped if the circuit is broken.

NOTES

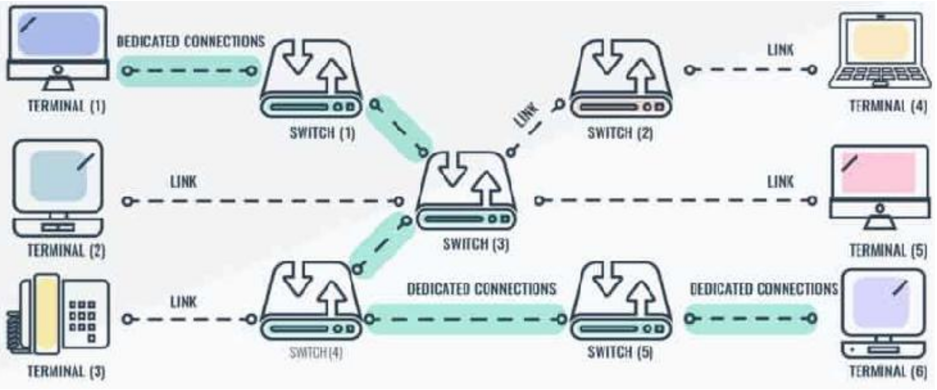


Fig. 12.1 *Circuit Switching*

Packet Switching

In telecommunications, packet switching is a method of grouping data that is transmitted over a digital network into packets. Packets are made of a header and a payload. Data in the header is used by networking hardware to direct the packet to its destination, where the payload is extracted and used by application software. Packet switching is the primary basis for data communications in computer networks worldwide. Packet switching is used to optimize the use of the channel capacity available in digital telecommunication networks, such as computer networks, and minimize the transmission latency (the time it takes for data to pass across the network), and to increase robustness of communication. Packet switching is used in the Internet and most local area networks. The Internet is implemented by the Internet Protocol Suite using a variety of Link Layer technologies. For example, Ethernet and Frame Relay are common. Newer mobile phone technologies (e.g., GSM, LTE) also use packet switching. Packet switching is associated with connectionless networking because, in these systems, no connection agreement needs to be established between communicating parties prior to exchanging data.

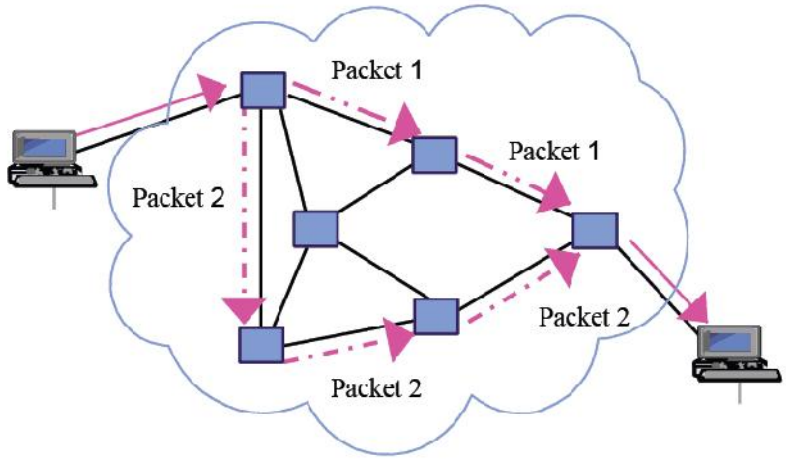


Fig. 12.2 *Packet Switching*

Local Area Network

NOTES

A Local Area Network (LAN) is a computer network that interconnects computers within a limited area such as a residence, school, laboratory, university campus or office building. By contrast, a Wide Area Network (WAN) not only covers a larger geographic distance, but also generally involves leased telecommunication circuits. Ethernet and Wi-Fi are the two most common technologies in use for local area networks. Historical network technologies include ARCNET, Token Ring, and AppleTalk. A router assigns IP addresses to each device on the network and facilitates a shared Internet connection between all the connected devices. A network switch connects to the router and facilitates communication between connected devices but does not handle Local Area Network IP configuration or sharing Internet connections. Switches are ideal tools for increasing the number of LAN ports available on the network. The Local Area Network layout, also known as Local Area Network topology, describes the physical and logical way devices and network segments are interconnected. LANs are categorized by the physical signal transmission medium or the logical way data travels through the network between devices, independent of the physical connection.

LANs generally consist of cables and switches, which can be connected to a router, cable modem, or ADSL modem for Internet access. LANs can also include such network devices as firewalls, load balancers, and network intrusion detection. Logical network topology examples include twisted pair Ethernet, which is categorized as a logical bus topology, and token ring, which is categorized as a logical ring topology. Physical network topology examples include star, mesh, tree, ring, point-to-point, circular, hybrid, and bus topology networks, each consisting of different configurations of nodes and links.

The function of Local Area Networks is to link computers together and provide shared access to printers, files, and other services. Local area network architecture is categorized as either peer-to-peer or client-server. On a client-server local area network, multiple client-devices are connected to a central server, in which application access, device access, file storage, and network traffic are managed. Applications running on the Local Area Network server provide services such as database access, document sharing, email, and printing. Devices on a peer-to-peer local area network share data directly to a switch or router without the use of a central server. LANs can interconnect with other LANs via leased lines and services, or across the Internet using virtual, private network technologies. This system of connected LANs is classified as a Wide Local Area Network or a metropolitan area network. Local Area and Wide Area Networks differ in their range. An Emulated Local Area Network enables routing and data bridging an Asynchronous Transfer Mode (ATM) network, which facilitates the exchange of Ethernet and token ring network data.

NOTES

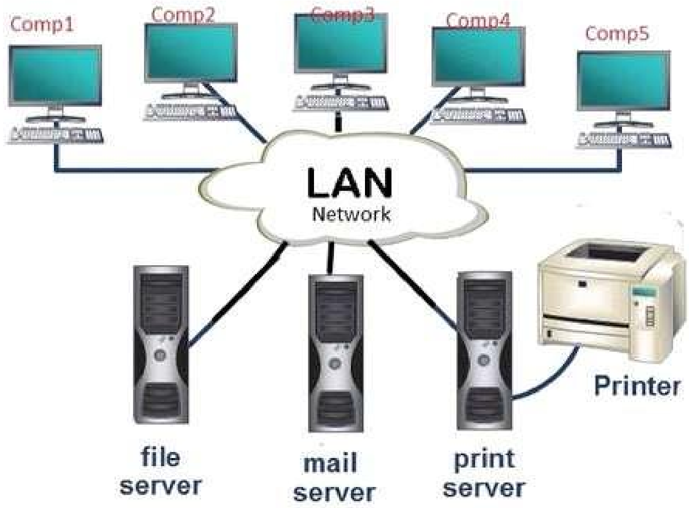


Fig. 12.3 *LAN (Local Area Network)*

There are several advantages of Local Area Networks in business:

- Reduced Costs: LANs present a significant reduction in Local Area Network hardware costs and efficient resource pooling.
- Increased Storage Capacity: By pooling all data into a central data storage server, the number of storage servers required is decreased and the efficiency of operations is increased.
- Optimized Flexibility: Data can be accessed by any device from anywhere via Internet connection.
- Streamlined Communication: Files and messages can be transferred in real time and accessed easily from anywhere on any device.

For users of network-based conferencing products, latency is the primary problem today. The IETF (Internet Engineering Task Force) and the IEEE are addressing these quality-of-service concerns with several efforts, including:

RSVP (Receiver-Initiated Resource Reservation Protocol)

RSVP is a transport layer protocol designed to reserve resources across a network using the integrated services model. RSVP operates over an IPv4 or IPv6 and provides receiver-initiated setup of resource reservations for multicast or unicast data flows. It does not transport application data but is like a control protocol, like Internet Control Message Protocol (ICMP) or Internet Group Management Protocol (IGMP). RSVP is described in RFC 2205. RSVP can be used by hosts and routers to request or deliver specific levels of Quality of Service (QoS) for application data streams. RSVP defines how applications place reservations and how they can relinquish the reserved resources once no longer required. RSVP operations will generally result in resources being reserved in each node along a path. RSVP is not a routing protocol but was designed to interoperate with current and future routing protocols.

RTP (Real-Time Transport Protocol)

NOTES

The Real-time Transport Protocol (RTP) is a network protocol for delivering audio and video over IP networks. RTP is used in communication and entertainment systems that involve streaming media, such as telephony, video teleconference applications including WebRTC, television services and web-based push-to-talk features. RTP typically runs over User Datagram Protocol (UDP). RTP is used in conjunction with the RTP Control Protocol (RTCP). While RTP carries the media streams (e.g., audio and video), RTCP is used to monitor transmission statistics and Quality of Service (QoS) and aids synchronization of multiple streams. RTP is one of the technical foundations of Voice over IP and in this context is often used in conjunction with a signalling protocol, such as the Session Initiation Protocol (SIP) which establishes connections across the network. RTP is designed for end-to-end, real-time transfer of streaming media. The protocol provides facilities for jitter compensation and detection of packet loss and out-of-order delivery, which are common especially during UDP transmissions on an IP network. RTP allows data transfer to multiple destinations through IP multicast. RTP is regarded as the primary standard for audio/video transport in IP networks and is used with an associated profile and payload format. The design of RTP is based on the architectural principle known as application-layer framing where protocol functions are implemented in the application as opposed to in the operating system’s protocol stack.

RTCP (Real-Time Transport Control Protocol)

The RTP Control Protocol (RTCP) is a sister protocol of the Real-time Transport Protocol (RTP). Its basic functionality and packet structure is defined in RFC 3550. RTCP provides out-of-band statistics and control information for an RTP session. It partners with RTP in the delivery and packaging of multimedia data but does not transport any media data itself. The primary function of RTCP is to provide feedback on the quality of service (QoS) in media distribution by periodically sending statistics information such as transmitted octet and packet counts, packet loss, packet delay variation, and round-trip delay time to participants in a streaming multimedia session. An application may use this information to control quality of service parameters, perhaps by limiting flow, or using a different codec.

Check Your Progress

1. Define the term multimedia.
2. List the categories multimedia can be divided into.
3. What is digital audio?
4. Define the term circuit switching.
5. What is LAN (Local Area Network)?

12.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

NOTES

1. Multimedia is a form of communication that combines different content forms, such as text, audio, images, animations, or video into a single presentation, in contrast to traditional mass media, such as printed material or audio recordings.
2. Multimedia can be divided into linear and nonlinear.
3. Digital audio is a representation of sound recorded in, or converted into, digital form. In digital audio, the sound wave of the audio signal is typically encoded as numerical samples in a continuous sequence.
4. Circuit switching is a method of implementing a telecommunications network in which two network nodes establish a dedicated communications channel (circuit) through the network before the nodes may communicate.
5. A Local Area Network (LAN) is a computer network that interconnects computers within a limited area such as a residence, school, laboratory, university campus or office building.

12.5 SUMMARY

- Multimedia is a form of communication that combines different content forms, such as text, audio, images, animations, or video into a single presentation, in contrast to traditional mass media, such as printed material or audio recordings.
- Multimedia can be recorded for playback on computers, laptops, smartphones, and other electronic devices, either on demand or in real time (streaming).
- Linear active content progresses often without any navigational control for the viewers, such as a cinema presentation.
- Non-linear uses interactivity to control progress as with a video game or self-paced computer-based training. Hypermedia is an example of non-linear content.
- Digital audio is a representation of sound recorded in, or converted into, digital form. In digital audio, the sound wave of the audio signal is typically encoded as numerical samples in a continuous sequence.
- In a digital audio system, an analog electrical signal representing the sound is converted with an Analog-to-Digital Converter (ADC) into a digital signal, typically using Pulse-Code Modulation (PCM).

NOTES

- Digital audio systems may include compression, storage, processing, and transmission components. Conversion to a digital format allows convenient manipulation, storage, transmission, and retrieval of an audio signal.
- A digital audio signal may be stored or transmitted. Digital audio can be stored on a CD, a digital audio player, a hard drive, a USB flash drive, or any other digital data storage device.
- The digital signal may be altered through digital signal processing, where it may be filtered or have effects applied.
- Digital video is an electronic representation of moving visual images (video) in the form of encoded digital data.
- Digital video can be stored on digital media, such as Blu-ray Disc, on computer data storage, or streamed over the Internet to end users who watch content on a desktop computer screen or a digital smart TV.
- Circuit switching is a method of implementing a telecommunications network in which two network nodes establish a dedicated communications channel (circuit) through the network before the nodes may communicate.
- The circuit guarantees the full bandwidth of the channel and remains connected for the duration of the communication session.
- In telecommunications, packet switching is a method of grouping data that is transmitted over a digital network into packets. Packets are made of a header and a payload.
- Packet switching is associated with connectionless networking because, in these systems, no connection agreement needs to be established between communicating parties prior to exchanging data.
- A Local Area Network (LAN) is a computer network that interconnects computers within a limited area such as a residence, school, laboratory, university campus or office building.
- LANs generally consist of cables and switches, which can be connected to a router, cable modem, or ADSL modem for Internet access.
- The function of Local Area Networks is to link computers together and provide shared access to printers, files, and other services. Local area network architecture is categorized as either peer-to-peer or client-server.
- LANs can interconnect with other LANs via leased lines and services, or across the Internet using virtual, private network technologies.
- RSVP is a transport layer protocol designed to reserve resources across a network using the integrated services model. RSVP operates over an IPv4 or IPv6 and provides receiver-initiated setup of resource reservations for multicast or unicast data flows.
- The Real-time Transport Protocol (RTP) is a network protocol for delivering audio and video over IP networks.

NOTES

- RTP is used in conjunction with the RTP Control Protocol (RTCP). While RTP carries the media streams (e.g., audio and video), RTCP is used to monitor transmission statistics and Quality of Service (QoS) and aids synchronization of multiple streams.
- The RTP Control Protocol (RTCP) is a sister protocol of the Real-time Transport Protocol (RTP). Its basic functionality and packet structure is defined in RFC 3550.
- RTCP provides out-of-band statistics and control information for an RTP session. It partners with RTP in the delivery and packaging of multimedia data but does not transport any media data itself.

12.6 KEY WORDS

- **Circuit switching:** It is a method of implementing a telecommunications network in which two network nodes establish a dedicated communications channel (circuit) through the network before the nodes may communicate.
- **Packet switching:** It is a method of grouping data that is transmitted over a digital network into packets.
- **A Local Area Network (LAN):** It is a computer network that interconnects computers within a limited area, such as a residence, school, laboratory, university campus or office building.
- **Wide Area Network (WAN):** It not only covers a larger geographic distance, but also generally involves leased telecommunication circuits. Ethernet and Wi-Fi are the two most common technologies in use for local area networks.
- **RSVP:** It is a transport layer protocol designed to reserve resources across a network using the integrated services model.
- **The Real-time Transport Protocol (RTP):** It is a network protocol for delivering audio and video over IP networks.
- **The RTP Control Protocol (RTCP):** Its basic functionality and packet structure is defined in RFC 3550. RTCP provides out-of-band statistics and control information for an RTP session.

12.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is multimedia and its categories?
2. State the significance of audio and its digitization.

3. List any advantage of video and its digitization.
4. List any one difference between circuit switching and packet switching.
5. What are the advantages of LAN?

NOTES

Long-Answer Questions

1. Briefly discuss multimedia with the help of an example.
2. Elaborate on the characteristics of digitization with the help of an example.
3. Discuss in detail about the circuit switching listing their objectives.
4. Elaborate on RSVP (Receiver-Initiated Resource Reservation Protocol) and RTP (Real-Time Transport Protocol) with the help of examples.

12.8 FURTHER READINGS

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

UNIT 13 COMMUNICATIONS AND APPLICATIONS

NOTES

Structure

13.0 INTRODUCTION

Quality of Service (QoS) is the description or measurement of the overall performance of a service, such as a telephony or computer network or a cloud computing service, particularly the performance seen by the users of the network. To quantitatively measure quality of service, several related aspects of the network service are often considered, such as packet loss, bit rate, throughput, transmission delay, availability, jitter, etc. Quality of service is particularly important for the transport of traffic with special requirements.

RSVP is a transport layer protocol designed to reserve resources across a network using the integrated services model. RSVP operates over an IPv4 or IPv6 and provides receiver-initiated setup of resource reservations for multicast or unicast data flows. The Real-time Transport Protocol (RTP) is a network protocol for delivering audio and video over IP networks. RTP is used in communication and entertainment systems that involve streaming media, such as telephony, video teleconference applications including WebRTC, television services and web-based push-to-talk features. The RTP Control Protocol (RTCP) is a sister protocol of the Real-time Transport Protocol (RTP). Its basic functionality and packet structure is defined in RFC 3550. RTCP provides out-of-band statistics and control information for an RTP session. The Real Time Streaming Protocol (RTSP) is a network control protocol designed for use in entertainment and communications systems to control streaming media servers. The protocol is used for establishing and controlling media sessions between endpoints.

In this unit, you will study about the significance of Quality of Service (QoS), QoS protocols, features of multimedia over IP, RSVP (Receiver-Initiated Resource Reservation Protocol), Real-time Transport Protocol (RTP), RTP Control Protocol

(RTCP) and Real Time Streaming Protocol (RTSP).

13.1 OBJECTIVES

NOTES

After going through this unit, you will be able to:

- Analyse the significance of Quality of Service (QoS)
- Interpret the qualities of QoS traffic
- Discuss the various QoS protocols
- Understand the features of multimedia over IP
- Explain the prime feature of RSVP (Receiver-Initiated Resource Reservation Protocol)
- Define the functions of Real-time Transport Protocol (RTP)
- Know about the RTP Control Protocol (RTCP)
- Comprehend about the Real Time Streaming Protocol (RTSP)

13.2 QUALITY OF MULTIMEDIA DATA TRANSMISSION

Quality of service of multimedia data transmission depends upon the different factor, such as transform method, quantization, block partitioning method, motion estimation and compensation strategy, optimization strategy.

Quality of Service (QoS)

Quality of Service (QoS) is the description or measurement of the overall performance of a service, such as a telephony or computer network or a cloud computing service, particularly the performance seen by the users of the network. To quantitatively measure quality of service, several related aspects of the network service are often considered, such as packet loss, bit rate, throughput, transmission delay, availability, jitter, etc. In the field of computer networking and other packet-switched telecommunication networks, quality of service refers to traffic prioritization and resource reservation control mechanisms rather than the achieved service quality. Quality of service is the ability to provide different priorities to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow.

Quality of service is particularly important for the transport of traffic with special requirements. Developers have introduced Voice over IP or VoIP technology to allow computer networks to become as useful as telephone networks for audio conversations, as well as supporting new applications with even stricter network performance requirements.

In the field of computer networking and other packet-switched telecommunication networks, tele traffic engineering refers to traffic prioritization and resource reservation control mechanisms rather than the achieved service quality. Quality of service is the ability to provide different priorities to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow. For example, a required bit rate, delay, delay variation, packet loss or bit error rates may be guaranteed. Quality of service is important for real-time streaming multimedia applications, such as voice over IP, multiplayer online games, and IPTV, since these often require fixed bit rate and are delay sensitive. Quality of service is especially important in networks where the capacity is a limited resource, for example in cellular data communication.

A network or protocol that supports QoS may agree on a traffic contract with the application software and reserve capacity in the network nodes, for example during a session establishment phase. During the session it may monitor the achieved level of performance, for example the data rate and delay, and dynamically control scheduling priorities in the network nodes. It may release the reserved capacity during a tear down phase. A best-effort network or service does not support quality of service. An alternative to complex QoS control mechanisms is to provide high quality communication over a best-effort network by over-provisioning the capacity so that it is sufficient for the expected peak traffic load. The resulting absence of network congestion reduces or eliminates the need for QoS mechanisms. QoS is sometimes used as a quality measure, with many alternative definitions, rather than referring to the ability to reserve resources. Quality of service sometimes refers to the level of quality of service, i.e., the guaranteed service quality. High QoS is often confused with a high level of performance, for example high bit rate, low latency, and low bit error rate. QoS is sometimes used in application layer services, such as telephony and streaming video to describe a metric that reflects or predicts the subjectively experienced quality. In this context, QoS is the acceptable cumulative effect on subscriber satisfaction of all imperfections affecting the service. Other terms with similar meaning are the Quality of Experience (QoE), Mean Opinion Score (MOS), Perceptual Speech Quality Measure (PSQM) and Perceptual Evaluation of Video Quality (PEVQ).

13.2.1 Qualities of Traffic

In packet-switched networks, quality of service is affected by various factors, which can be divided into human and technical factors. Human factors include stability of service quality, availability of service, waiting times and user information. Technical factors include reliability, scalability, effectiveness, maintainability, and network congestion. Many things can happen to packets as they travel from origin to destination, resulting in the following problems as seen from the point of view of the sender and receiver:

- **Goodput**

Due to varying load from disparate users sharing the same network resources, the maximum throughput that can be provided to a certain data stream may be too low for real-time multimedia services.

- **Packet Loss**

NOTES

The network may fail to deliver (drop) some packets due to network congestion. The receiving application may ask for this information to be retransmitted, possibly resulting in congestive collapse or unacceptable delays in the overall transmission.

- **Errors**

Sometimes packets are corrupted due to bit errors caused by noise and interference, especially in wireless communications and long copper wires. The receiver must detect this, and, just as if the packet were dropped, may ask for this information to be retransmitted.

- **Latency**

It might take a long time for each packet to reach its destination because it gets held up in long queues, or it takes a less direct route to avoid congestion. In some cases, excessive latency can render an application, such as VoIP or online gaming unusable.

- **Packet Delay Variation**

Packets from the source will reach the destination with different delays. A packet's delay varies with its position in the queues of the routers along the path between source and destination, and this position can vary unpredictably. Delay variation can be absorbed at the receiver, but in so doing increases the overall latency for the stream.

- **Out-of-Order Delivery**

When a collection of related packets is routed through a network, different packets may take different routes, each resulting in a different delay. The result is that the packets arrive in a different order than they were sent. This problem requires special additional protocols for rearranging out-of-order packets. The reordering process requires additional buffering at the receiver, and, as with packet delay variation, increases the overall latency for the stream.

NOTES

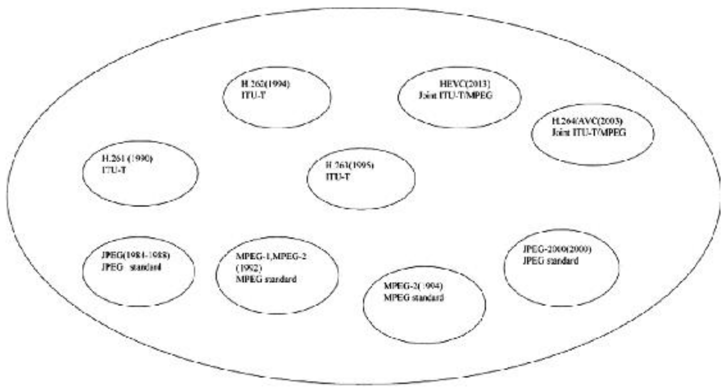


Fig. 13.1 *Quality of Multimedia Data Transmission in Multimedia*

Protocols

Several QoS mechanisms and schemes exist for IP networking.

- The Type of Service (ToS) field in the IPv4 header (now superseded by DiffServ)
- Differentiated Services (DiffServ)
- Integrated Services (IntServ)
- Resource Reservation Protocol (RSVP)
- RSVP-TE

QoS capabilities are available in the following network technologies.

- Multiprotocol Label Switching (MPLS) provides eight QoS classes
- Frame Relay
- X.25
- Some DSL modems
- Asynchronous Transfer Mode (ATM)
- Ethernet supporting IEEE 802.1Q with Audio Video Bridging and Time-Sensitive Networking
- Wi-Fi supporting IEEE 802.11e

The Internet carries all types of traffic, each type has different characteristics and requirements. For example, a file transfer application requires that some quantity of data is transferred in an acceptable amount of time, while Internet telephony requires that most packets get to the receiver in less than 0.3 seconds. If enough bandwidth is available, best-effort service fulfils all these requirements. When resources are scarce, however, real-time traffic will suffer from the congestion.

13.3 MULTIMEDIA OVER IP

NOTES

The solution for multimedia over IP is to classify all traffic, allocate priority for different applications and make reservations. The Integrated Services working group in the IETF (Internet Engineering Task Force) developed an enhanced Internet service model called Integrated Services that includes best-effort service and real-time service. The real-time service will enable IP networks to provide quality of service to multimedia applications. Resource Reservation Protocol (RSVP), together with Real-time Transport Protocol (RTP), Real-Time Control Protocol (RTCP), Real-Time Streaming Protocol (RTSP), provides a working foundation for real-time services. Integrated Services allows applications to configure and manage a single infrastructure for multimedia applications and traditional applications. It is a comprehensive approach to provide applications with the type of service they need and in the quality they choose.

RSVP (Receiver-Initiated Resource Reservation Protocol)

RSVP is a transport layer protocol designed to reserve resources across a network using the integrated services model. RSVP operates over an IPv4 or IPv6 and provides receiver-initiated setup of resource reservations for multicast or unicast data flows. It does not transport application data but is like a control protocol, like Internet Control Message Protocol (ICMP) or Internet Group Management Protocol (IGMP). RSVP is described in RFC 2205. RSVP can be used by hosts and routers to request or deliver specific levels of Quality of Service (QoS) for application data streams. RSVP defines how applications place reservations and how they can relinquish the reserved resources once no longer required. RSVP operations will generally result in resources being reserved in each node along a path. RSVP is not a routing protocol but was designed to interoperate with current and future routing protocols.

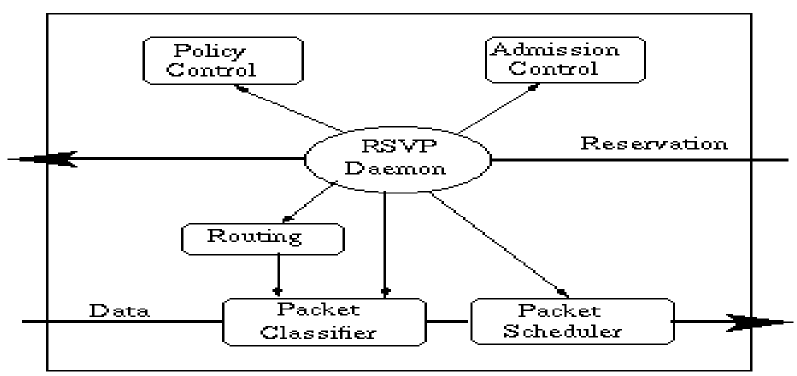


Fig 13.2 *Reservation at a Node on the Data Flow Path*

RSVP Features

- **RSVP flows are simplex.**

  RSVP distinguishes senders and receivers. Although in many cases, a host can act both as a sender and as a receiver, one RSVP reservation only reserves resources for data streams in one direction.

- **RSVP supports both multicast and unicast and adapts to changing memberships and routes.**

  RSVP is designed for both multicast and unicast. Since the reservations are initiated by the receivers and the reservation states are soft, RSVP can easily handle changing memberships and routes. A host can send IGMP (Internet Group Management Protocol) messages to join a multicast group. Reservation merging enables RSVP to scale to large multicast groups without causing heavy overhead for the sender.

- **RSVP is receiver-oriented and handles heterogeneous receivers.**

  In heterogeneous multicast groups, receivers have different capacities and levels of QoS. The receiver-oriented RSVP reservation requests facilitate the handling of heterogeneous multicast groups. Receivers are responsible for choosing its own level of QoS, initiating the reservation, and keeping it active if it wants. The senders divide traffic in several flows, each is a separate RSVP flow with different level of QoS. Each RSVP flow is homogeneous, and receivers can choose to join one or more flows. This approach makes it possible for heterogeneous receivers to request different QoS tailored to their capacities and requirements.

- **RSVP has good compatibility.**

  Efforts have been made to run RSVP over both IPv4 and IPv6. It provides opaque transport of traffic control and policy control messages to be more adaptive to new technologies. It also provides transparent operation through non-supporting regions.

RTP (Real-Time Transport Protocol)

The Real-time Transport Protocol (RTP) is a network protocol for delivering audio and video over IP networks. RTP is used in communication and entertainment systems that involve streaming media, such as telephony, video teleconference applications including WebRTC, television services and web-based push-to-talk features. RTP typically runs over User Datagram Protocol (UDP). RTP is used in conjunction with the RTP Control Protocol (RTCP). While RTP carries the media streams (e.g., audio and video), RTCP is used to monitor transmission statistics and Quality of Service (QoS) and aids synchronization of multiple streams. RTP is one of the technical foundations of Voice over IP and in this context is often used in conjunction with a signalling protocol, such as the Session Initiation Protocol (SIP) which establishes connections across the network. RTP is designed for end-to-end, real-time transfer of streaming media. The protocol provides facilities for jitter compensation and detection of packet loss and out-of-order delivery, which are common especially during UDP transmissions on an IP network. RTP allows data transfer to multiple destinations through IP multicast. RTP is regarded as the primary standard for audio/video transport in IP networks and is used with an

NOTES

associated profile and payload format. The design of RTP is based on the architectural principle known as application-layer framing where protocol functions are implemented in the application as opposed to in the operating system's protocol stack.

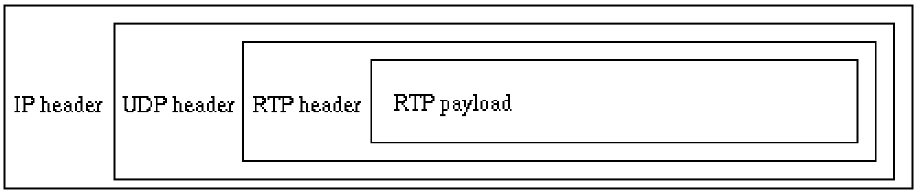


Fig 13.3 *RTP Data in an IP Packet*

RTP Features

- RTP provides end-to-end delivery services for data with real-time characteristics, such as interactive audio and video. But RTP itself does not provide any mechanism to ensure timely delivery. It needs support from lower layers that have control over resources in switches and routers. RTP depends on RSVP to reserve resources and to provide the requested quality of service.
- RTP does not assume anything about the underlying network, except that it provides framing. RTP is typically run on the top of UDP to make use of its multiplexing and checksum service, but efforts have been made to make RTP compatible with other transport protocols, such as ATM AAL5 and IPv6.
- Unlike usual data transmission, RTP does not offer any form of reliability or flow/congestion control. It provides timestamps, sequence numbers as hooks for adding reliability and flow/congestion control, but how to implement is totally left to the application.
- RTP is a protocol framework that is deliberately not complete. It is open to new payload formats and new multimedia software. By adding new profile and payload format specifications, one can tailor RTP to new data formats and new applications.
- RTP/RTCP provides functionality and control mechanisms necessary for carrying real-time content. But RTP/RTCP itself is not responsible for the higher-level tasks like assembly and synchronization. These must be done at application level.
- The flow and congestion control information of RTP is provided by RTCP sender and receiver reports.

RTCP (Real-Time Transport Control Protocol)

The RTP Control Protocol (RTCP) is a sister protocol of the Real-time Transport Protocol (RTP). Its basic functionality and packet structure is defined in RFC

NOTES

3550. RTCP provides out-of-band statistics and control information for an RTP session. It partners with RTP in the delivery and packaging of multimedia data but does not transport any media data itself. The primary function of RTCP is to provide feedback on the quality of service (QoS) in media distribution by periodically sending statistics information, such as transmitted octet and packet counts, packet loss, packet delay variation, and round-trip delay time to participants in a streaming multimedia session. An application may use this information to control quality of service parameters, perhaps by limiting flow, or using a different codec.

Real-Time Streaming Protocol (RTSP)

The Real Time Streaming Protocol (RTSP) is a network control protocol designed for use in entertainment and communications systems to control streaming media servers. The protocol is used for establishing and controlling media sessions between endpoints. Clients of media servers issue commands, such as play, record, and pause, to facilitate real-time control of the media streaming from the server to a client (Video on Demand) or from a client to the server (Voice Recording). The transmission of streaming data itself is not a task of RTSP. Most RTSP servers use the Real-time Transport Protocol (RTP) in conjunction with Real-time Control Protocol (RTCP) for media stream delivery. However, some vendors implement proprietary transport protocols. The RTSP server software from RealNetworks, for example, also used RealNetworks' proprietary Real Data Transport (RDT).

RTSP Features

- RTSP is an application-level protocol with syntax and operations like HTTP but works for audio and video. It uses URLs like those in HTTP.
- An RTSP server needs to maintain states, using SETUP, TEARDOWN, and other methods.
- RTSP messages are be carried out-of-band. The protocol for RTSP may be different from the data delivery protocol.
- Unlike HTTP, in RTSP both servers and clients can issue requests.
- RTSP is implemented on multiple operating system platforms, it allows interoperability between clients and servers from different manufacturers.

Check Your Progress

1. Define the term Quality of Service (QoS).
2. List any two protocols of QoS mechanisms.
3. What is RSVP (Receiver-Initiated Resource Reservation Protocol)?
4. State the functionality of The Real-time Transport Protocol (RTP).
5. Define the term The Real Time Streaming Protocol (RTSP).

13.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

NOTES

1. Quality of Service (QoS) is the description or measurement of the overall performance of a service, such as a telephony or computer network or a cloud computing service, particularly the performance seen by the users of the network.
2. The two protocols of QoS mechanisms are:
   - The Type of Service (ToS) field in the IPv4 header (now superseded by DiffServ)
   - Differentiated Services (DiffServ)
3. RSVP is a transport layer protocol designed to reserve resources across a network using the integrated services model. RSVP operates over an IPv4 or IPv6 and provides receiver-initiated setup of resource reservations for multicast or unicast data flows.
4. The Real-time Transport Protocol (RTP) is a network protocol for delivering audio and video over IP networks. RTP is used in communication and entertainment systems that involve streaming media, such as telephony, video teleconference applications including WebRTC, television services and web-based push-to-talk features.
5. The Real Time Streaming Protocol (RTSP) is a network control protocol designed for use in entertainment and communications systems to control streaming media servers.

13.5 SUMMARY

- Quality of service of multimedia data transmission depends upon the different factor, such as transform method, quantization, block partitioning method, motion estimation and compensation strategy, optimization strategy.
- Quality of Service (QoS) is the description or measurement of the overall performance of a service, such as a telephony or computer network or a cloud computing service, particularly the performance seen by the users of the network.
- To quantitatively measure quality of service, several related aspects of the network service are often considered, such as packet loss, bit rate, throughput, transmission delay, availability, jitter, etc.
- Quality of service is the ability to provide different priorities to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow.

NOTES

- Quality of service is particularly important for the transport of traffic with special requirements.
- In the field of computer networking and other packet-switched telecommunication networks, tele traffic engineering refers to traffic prioritization and resource reservation control mechanisms rather than the achieved service quality.
- Quality of service is the ability to provide different priorities to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow.
- Quality of service is important for real-time streaming multimedia applications, such as voice over IP, multiplayer online games, and IPTV, since these often require fixed bit rate and are delay sensitive.
- Quality of service is especially important in networks where the capacity is a limited resource, for example in cellular data communication.
- A network or protocol that supports QoS may agree on a traffic contract with the application software and reserve capacity in the network nodes, for example during a session establishment phase.
- QoS is sometimes used as a quality measure, with many alternative definitions, rather than referring to the ability to reserve resources.
- Quality of service sometimes refers to the level of quality of service, i.e., the guaranteed service quality.
- The Integrated Services working group in the IETF (Internet Engineering Task Force) developed an enhanced Internet service model called Integrated Services that includes best-effort service and real-time service.
- The real-time service will enable IP networks to provide quality of service to multimedia applications. Resource Reservation Protocol (RSVP), together with Real-time Transport Protocol (RTP), Real-Time Control Protocol (RTCP), Real-Time Streaming Protocol (RTSP), provides a working foundation for real-time services.
- RSVP is a transport layer protocol designed to reserve resources across a network using the integrated services model. RSVP operates over an IPv4 or IPv6 and provides receiver-initiated setup of resource reservations for multicast or unicast data flows.
- RSVP distinguishes senders and receivers. Although in many cases, a host can act both as a sender and as a receiver, one RSVP reservation only reserves resources for data streams in one direction.
- RSVP supports both multicast and unicast and adapts to changing memberships and routes.
- RSVP is designed for both multicast and unicast. Since the reservations are initiated by the receivers and the reservation states are soft, RSVP can

easily handle changing memberships and routes.

- RSVP is receiver-oriented and handles heterogeneous receivers.
- The Real-time Transport Protocol (RTP) is a network protocol for delivering audio and video over IP networks. RTP is used in communication and entertainment systems that involve streaming media, such as telephony, video teleconference applications including WebRTC, television services and web-based push-to-talk features.
- RTP typically runs over User Datagram Protocol (UDP). RTP is used in conjunction with the RTP Control Protocol (RTCP). While RTP carries the media streams (e.g., audio and video), RTCP is used to monitor transmission statistics and Quality of Service (QoS) and aids synchronization of multiple streams.
- RTP is one of the technical foundations of Voice over IP (VoIP) and in this context is often used in conjunction with a signalling protocol, such as the Session Initiation Protocol (SIP) which establishes connections across the network.
- RTP is designed for end-to-end, real-time transfer of streaming media. The protocol provides facilities for jitter compensation and detection of packet loss and out-of-order delivery, which are common especially during UDP transmissions on an IP network.
- RTP provides end-to-end delivery services for data with real-time characteristics, such as interactive audio and video.
- RTP does not assume anything about the underlying network, except that it provides framing.
- The RTP Control Protocol (RTCP) is a sister protocol of the Real-time Transport Protocol (RTP). Its basic functionality and packet structure is defined in RFC 3550.
- The Real Time Streaming Protocol (RTSP) is a network control protocol designed for use in entertainment and communications systems to control streaming media servers.

13.6 KEY WORDS

- **Quality of Service (QoS):** It is the description or measurement of the overall performance of a service, such as a telephony or computer network or a cloud computing service, particularly the performance seen by the users of the network.

- **Receiver-Initiated Resource Reservation Protocol (RSVP):** It is a transport layer protocol designed to reserve resources across a network using the integrated services model.
- **The Real-time Transport Protocol (RTP):** It is a network protocol for delivering audio and video over IP networks.
- **RTP Control Protocol (RTCP):** It is a sister protocol of the Real-time Transport Protocol. Its basic functionality and packet structure is defined in RFC 3550.
- **The Real Time Streaming Protocol (RTSP):** It is a network control protocol designed for use in entertainment and communications systems to control streaming media servers.

13.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Where is the Quality of Service (QoS) used?
2. Why has the Voice over IP (VoIP) technology been introduced?
3. Name the problems faced by the packet- switched networks.
4. List the various protocols of Quality of Service (QoS).
5. What is the significance of multimedia over IP?
6. List any two features of RSVP (Receiver-Initiated Resource Reservation Protocol).
7. What is UDP transmissions?
8. Define the term Real Data Transport (RDT).

Long-Answer Questions

1. Discuss briefly about Quality of Service (QoS) with the help of suitable examples.
2. Elaborate on quality of service which is affected by various factors giving appropriate examples.
3. Describe briefly about RSVP (Receiver-Initiated Resource Reservation Protocol) and list its various features.
4. Discuss the importance of Real-time Transport Protocol (RTP) and list its

various features with the help of suitable examples.

5. Differentiate between RSVP, RTP and RTCP giving suitable advantages and disadvantages of each.

NOTES

13.8 FURTHER READINGS

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.

NOTES

UNIT 14 MULTIMEDIA OVER ATM NETWORKS

Structure

14.0 INTRODUCTION

MPEG Transport Stream (MPEG-TS, MTS) or Simply Transport Stream (TS) is a standard digital container format for transmission and storage of audio, video, and Program and System Information Protocol (PSIP) data. It is used in broadcast systems, such as DVB, ATSC and IPTV. Transport stream specifies a container format encapsulating packetized elementary streams, with error correction and synchronization pattern features for maintaining transmission integrity when the communication channel carrying the stream is degraded.

Media on Demand (MoD) is a new generation of video on demand, which not only allows users to watch/listen video and audio content like movies and TV shows, but also provides functions including real-time information, interactive games, attractions guidance, route information, advertising system, shopping and ordering service. Users can select content whenever they want, rather than having to watch it at a specific broadcast time. In the transportation industry, media on demand technology was first applied by FUNTORO, which offer media on demand as in vehicle infotainment to bus and railway passengers through high-definition interactive monitors embedded in seatback or armrest.

Video on Demand (VoD) is a media distribution system that allows users to access videos without a traditional video playback device and the constraints of a typical static broadcasting schedule. In the 20th century, broadcasting in the form of over-the-air programming was the most common form of media distribution.

In this unit, you will study about the transport of MPEG-4, uses of the transport of MPEG-4, concept of MoD (Media on Demand) and the features of VoD (Video on Demand) services.

NOTES

14.1 OBJECTIVES

After going through this unit, you will be able to:

- Analyse the significance of transport of MPEG-4
- Interpret the use of transport of MPEG-4
- Understand the concept of MoD (Media on Demand) services
- Know about the various types of VoD (Video on Demand) services

14.2 TRANSPORT OF MPEG-4

MPEG Transport Stream (MPEG-TS, MTS) or Simply Transport Stream (TS) is a standard digital container format for transmission and storage of audio, video, and Program and System Information Protocol (PSIP) data. It is used in broadcast systems, such as DVB, ATSC and IPTV. Transport stream specifies a container format encapsulating packetized elementary streams, with error correction and synchronization pattern features for maintaining transmission integrity when the communication channel carrying the stream is degraded. Transport streams differ from the similarly named MPEG program stream in several important ways: program streams are designed for reasonably reliable media, such as discs (like DVDs), while transport streams are designed for less reliable transmission, namely terrestrial or satellite broadcast. Further, a transport stream may carry multiple programs.

A transport stream encapsulates several other sub streams, often Packetized Elementary Streams (PESs) which in turn wrap the main data stream using the MPEG codec or any number of non-MPEG codecs, such as AC3 or DTS audio, and MJPEG or JPEG 2000 video), text and pictures for subtitles, tables identifying the streams, and even broadcaster-specific information, such as an electronic program guide. Many streams are often mixed, such as several different television channels, or multiple angles of a movie. Each stream is chopped into (at most) 188-byte sections and interleaved together; because of the tiny packet size, streams can be interleaved with less latency and greater error resilience compared to program streams and other common containers, such as AVI, MOV/MP4, and MKV, which generally wrap each frame into one packet. This is particularly important for videoconferencing, where large frames may introduce unacceptable audio delay. Transport streams tend to be broadcast as Constant Bitrate (CBR) and filled with padding bytes when not enough data exists.

NOTES

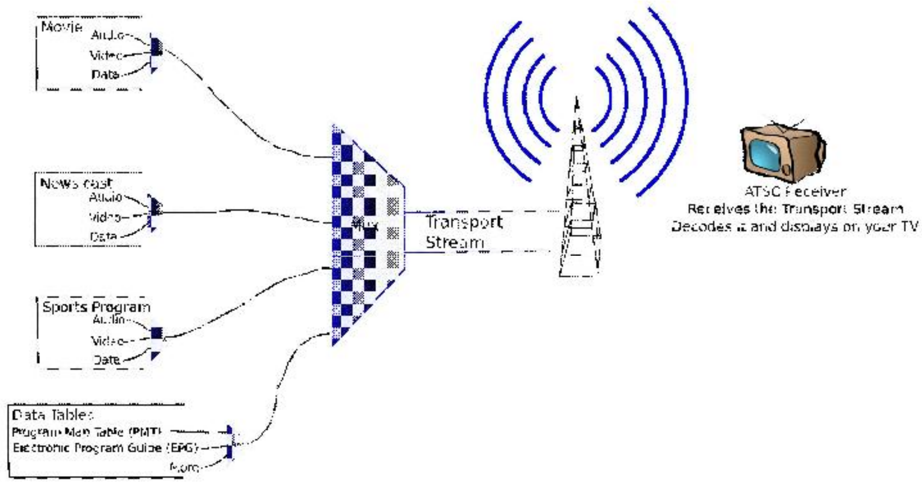


Fig 14.1 *Multiple MPEG Programs are Combined and sent to a Transmitting Antenna*

Uses of Transport in MPEG

1. Use in Digital Video Cameras

Transport Stream was originally designed for broadcast. Later it was adapted for use with digital video cameras, recorders, and players by adding a 4-byte Timecode (TC) field to the standard 188-byte packets, resulting in a 192-byte packet. This is what is informally called M2TS stream. The Blu-ray Disc Association calls it “BDAV MPEG-2 transport stream”. JVC called it TOD when used in HDD-based camcorders like GZ-HD7. The timecode allows quick access to any part of the stream either from a media player, or from a non-linear video editing system. It is also used to synchronize video streams from several cameras in a multiple-camera setup.

2. Uses in Blu-Ray

Blu-ray Disc video titles authored with menu support are in the Blu-ray Disc Movie (BDMV) format and contain audio, video, and other streams in a BDAV container, which is based on the MPEG-2 transport stream format. Blu-ray Disc video uses these modified MPEG-2 transport streams, compared to DVD’s program streams that do not have the extra transport overhead. There is also the BDAV (Blu-ray Disc Audio/Visual) format, the consumer-oriented alternative to the BDMV format used for movie releases. The BDAV format is used on Blu-ray Disc recordable for audio/video recording. Blu-ray Disc employs the MPEG-2 transport stream recording method. This enables transport streams of a BDAV converted digital broadcast to be recorded as they are with minimal alteration of the packets. It also enables simple stream cut style editing of a BDAV converted digital broadcast that is recorded as is and where the data can be edited just by discarding unwanted packets from the stream. Although it is quite natural, a function for high-speed and easy-to-use retrieval is built in.

NOTES

14.3 MEDIA-ON-DEMAND

Video on Demand (VoD) technology is used in many applications, for example entertainment, e-commerce based applications, distance learning, etc. It refers to a system in which users are allowed to choose and watch the audio/video content on demand. The users interact with DVD player, for example, pause, jump and stop. The video is delivered by the transmission of unicast stream as request sent by clients. The throughput and bandwidth factors make VoD scalable. Using VoD mechanism, it is always considered to check the files with VoD servers which are needed to deliver the data and open streams as per client requests. The client gets to listen to two streams at a time in which one stream is used to deliver the data and other stream is used to send the data. VoD provides a broadband interactive service in which users can avail the various services, such as selecting and choosing remote, delivering the ordered content, such as TV program, multimedia applications, movies, sport event, music, etc. It completely replaces the Video Cassette Recorders (VCRs) because it works the same as a video player, for example, it provides STOP, PAUSE, FORWARD and REWIND services. For this, it frequently requires Set Top Box (STB). The services collectively represent the virtual multimedia or video content shop. The implementation of VoD allows the operators to:

- Obtain more profit and revenue by registering more viewers.
- Obtain high broadband data network operator by offering the advanced multimedia services.
- Watch and order the selected movie or multimedia content in-house.
- The possibility of remote accessing of all the services in terms of 24×7 timing.

The VoD distribution requires VoD servers and private networks to establish frequently all the services. This technology prefers the client server computing to deliver the video metadata as per requested by the users. In client server computing, many applications are processed on a computer. The system unit is worked as client side which can be obtained in application services, for example, database services received from various computers. The client server computing describes the relationship between information and programs. VoD needs an integrated and high secured mechanism to generate the encrypted file. This file supports the automation of content ingestion process from the recipient side for the video files and metadata. Sometimes, optional watermark insertion is provisioned on VoD servers that provides robust operational front end added values. It encrypts the video metadata using client and server authentication. Sometimes, for large file it uses VoD trick play feature. It uses Digital Video Broadcast (DVB) standard and

Control Message (CM) format metadata within stream that includes watermark management and copy control for the various MPEG-2 transport.

Table 14.1 summarizes the various types of VoD services.

NOTES

Table 14.1 *Types of VoD Services*

VoD Services	Functions
Broadcast (No-VoD) service	This service provides broadcast service which is the same as TV broadcast.
Pay-per-view (PPV) service	This service provides user sign ups and pay for specific video facility. It almost works in the same way as CAble TV (CATV) Pay Per View (PPV) service works.
Quasi VoD (Q-VoD) service	This service provides its services for a group and is made on threshold services across net.
Near VoD(N-VoD) service	This service provides the functions of forwarding and reversing multimedia transitions in a specific time intervals, for example, 5 minutes. This service is provided by various channels with the skewed programming at a given time.
True (T-VoD) service	This service provides those services which are available on session presentation. The user has full control on Video Cassette Recorder (VCR) facilities along with virtual capability, forward and reverse playing services, and freezing and sleeping of video metadata.

VoD provides a smooth way of viewing the films and television programs which helps to eliminate the need of going to the video store to get a wide collection. You can watch the movie, rewind the movie and watch it again. VoD is an exclusive feature of digital cable that allows you to hire the programs and movies by paying a little fee. You can get the menu list too. VoD consists of server and local database which are vehemently used in transferring the video data on demand. It is mainly connected with high speed backbone network. The video server is used to control the transmission, guaranteed stream encryption and transmission, data retrieval, request handling, and support functions for the video metadata demanded by users. A video encoder sometimes referred to as the video server is used to send the digital images across an Internet Protocol (IP) network, for example, Intranet, Internet or Local Area Network (LAN) setup. It turns the analog video system into the network video system which makes users to access the live images across net. For this, either Web browser or video management system software is used that essentially authorizes the various viewers who want to get the video's metadata and information simultaneously that are added to the system. VoD is used in interactive multimedia services as listed in Table 14.2.

NOTES

Table 14.2 *Interactive Multimedia Services*

Applications of VoD	Function
Movies available on demand	The customers are able to select and watch the movie with full VCR capabilities.
Interactive video games	The customers are able to play with computer games without buying the physical copy of amusements and games.
Interactive news television	The customers are able to watch the tailored newscasts.
Catalogue browsing	The customers are able to purchase the commercial products by this service.
Distance learning	The customers who reside in the remote areas are able to subscribe the courses but their PCs are connected to the servers.
Interactive advertising	The customers are able to make a survey of advertisement and publicity with ad samples.
Video conferencing	The customers are able to negotiate and conversant with other users by interacting the multimedia and broadcast services.

Screen below shows VoD application in which MPEG-21 and time dependent metadata technology has been used to deliver the MPEG-4 videos demanded by the clients. The streaming server known as broker uses MPEG-21 if the client application starts. The digital item contains list of available movies and digital item identification on the multimedia streaming server. Once you start the selected movie the MPEG-4 audio and video stream is sent as per client request. The hardware setting supports Windows XP, Windows Vista operating system, Java Runtime Environment (JRE), QuickTime for Java and running application supports to extract the .zip file.

MPEG-4 supports VoD in which users adjust the CPU load in high, medium and low state and also check the battery capacity too while watching the video.

NOTES

Check Your Progress

1. Define the term MPEG-TS.
2. State any two uses of transport in MPEG.
3. Where is the VoD used?
4. State the significance of VoD.
5. Why is video encoder used to send the digital images across an Internet Protocol (IP) network?

14.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. MPEG Transport Stream (MPEG-TS, MTS) or Simply Transport Stream (TS) is a standard digital container format for transmission and storage of audio, video, and Program and System Information Protocol (PSIP) data.
2. The two uses of transport in MPEG are use in digital video camera and use in blue ray.
3. Video on Demand (VoD) technology is used in many applications, for example entertainment, e-commerce-based applications, distance learning, etc.
4. VoD consists of server and local database which are vehemently used in transferring the video data on demand. It is mainly connected with high-speed backbone network.
5. A video encoder sometimes referred to as the video server is used to send the digital images across an Internet Protocol (IP) network, for example, Intranet, Internet or Local Area Network (LAN) setup. It turns the analog video system into the network video system which makes users to access the live images across net.

14.5 SUMMARY

- MPEG Transport Stream (MPEG-TS, MTS) or Simply Transport Stream (TS) is a standard digital container format for transmission and storage of audio, video, and Program and System Information Protocol (PSIP) data.
- MPEG Transport Stream is used in broadcast systems, such as DVB, ATSC and IPTV.
- Transport stream specifies a container format encapsulating packetized elementary streams, with error correction and synchronization pattern

NOTES

features for maintaining transmission integrity when the communication channel carrying the stream is degraded.

- Transport streams differ from the similarly named MPEG program stream in several important ways program streams are designed for reasonably reliable media, such as discs (like DVDs), while transport streams are designed for less reliable transmission, namely terrestrial or satellite broadcast.
- A transport stream encapsulates several other sub streams, often Packetized Elementary Streams (PESs) which in turn wrap the main data stream using the MPEG codec or any number of non-MPEG codecs, such as AC3 or DTS audio, and MJPEG or JPEG 2000 video), text and pictures for subtitles, tables identifying the streams, and even broadcaster-specific information, such as an electronic program guide.
- Transport Stream was originally designed for broadcast. Later it was adapted for use with digital video cameras, recorders, and players by adding a 4-byte Timecode (TC) field to the standard 188-byte packets, resulting in a 192-byte packet.
- Blu-ray Disc video titles authored with menu support are in the Blu-ray Disc Movie (BDMV) format and contain audio, video, and other streams in a BDAV container, which is based on the MPEG-2 transport stream format.
- Blu-ray Disc video uses these modified MPEG-2 transport streams, compared to DVD's program streams that do not have the extra transport overhead.
- Video on Demand (VoD) technology is used in many applications, for example entertainment, e-commerce-based applications, distance learning, etc.
- Using VoD mechanism, it is always considered to check the files with VoD servers which are needed to deliver the data and open streams as per client requests.
- VoD provides a broadband interactive service in which users can avail the various services, such as selecting and choosing remote, delivering the ordered content, such as TV program, multimedia applications, movies, sport event, music, etc.
- The VoD distribution requires VoD servers and private networks to establish frequently all the services.
- VoD needs an integrated and high secured mechanism to generate the encrypted file. This file supports the automation of content ingestion process from the recipient side for the video files and metadata.

NOTES

- VoD provides a smooth way of viewing the films and television programs which helps to eliminate the need of going to the video store to get a wide collection.
- VoD consists of server and local database which are vehemently used in transferring the video data on demand. It is mainly connected with high-speed backbone network.
- MPEG-4 supports VoD in which users adjust the CPU load in high, medium and low state and also check the battery capacity too while watching the video.

14.6 KEY WORDS

- **MPEG transport stream:** It is a standard digital container format for transmission and storage of audio, video, and Program and System Information Protocol (PSIP) data.
- **Video on Demand (VoD):** It refers to a system in which users can choose and watch the audio/video content on demand.
- **Broadcast (No-VoD) service:** This service provides broadcast service which is the same as TV broadcast.
- **Pay-Per-View (PPV) service:** This service provides user sign ups and pay for specific video facility.
- **True (T-VoD) service:** This service provides those services which are available on session presentation.

14.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define the term VoD.
2. What does a VoD distribution requires?
3. State the various uses of transport in MPEG-4.

Long-Answer Questions

1. Discuss the significance of transport of MPEG-4 with the help of suitable examples.
2. Elaborate the types of VoD services and their functions in detail.
3. Describe the various applications of VoD with the help of appropriate examples.

NOTES

14.8 FURTHER READINGS

Li, Ze-Nian and Mark S. Drew. 2004. *Fundamentals of Multimedia*. New Jersey: Pearson Education International.

Chapman, Nigel and Jenny Chapman. 2009. *Digital Multimedia*, 3rd Edition. New Delhi: Wiley-Dreamtech India Pvt. Ltd.

Heath, Steve. 1999. *Multimedia and Communications Technology*, 2nd Edition. Massachusetts: Elsevier (Focal Press).

Mukhopadhyay, Anirban and Arup Chattopadhyay. 2010. *Introduction to Computer Graphics and Multimedia*. New Delhi: Vikas Publishing House Pvt. Ltd.

Mukherjee, D. P. 1998. *Fundamentals of Computer Graphics and Multimedia*. New Delhi: Prentice-Hall of India.

Villamil, John and Louis Molina. 1998. *Multimedia: An Introduction*. New Delhi: Prentice-Hall of India

Rosch, Winn L. 1995. *Multimedia Bible*. Lexington (US): Sam Publishing Company.

Buford, I. K. 1994. *Multimedia Systems*. New York: Addison-Wesley.