chapter 1

Multimedia—An Overview

1.1 Introduction

The word "multimedia" comes from the Latin words "*multus*" which means numerous and "*media*" which means middle or center. More recently, the word media (singular medium) started to convey the sense "intermediary". Multimedia therefore means "multiple intermediaries" or "multiple means". The word "multimedia" may be used as a noun (e.g., multimedia is a new technological field) and as an adjective (e.g., a multimedia document). Multimedia, in the general sense, therefore, means "multiple intermediaries" between the source and sink of information or "multiple means" by which information is stored, transmitted, presented or perceived. The **multiple means** by which we can perceive information are the following:

1. text (e.g., book chapter, newspaper article)
2. image (e.g., photograph)
3. graphics (e.g., sketch, diagram)
4. sound (e.g., speech, music)
5. video (e.g., TV program, movie clip)
6. animation (e.g., animation clip)

1.2 Multimedia Presentation and Production • •

A **multimedia presentation** is basically a "show" whose content is expressed through various media types like text, image, sound, video, etc. There can be various objectives of the presentation, for example, to deliver some information about a company's performance (corporate presentation), to enhance the knowledge of students (computer based training), to present the facilities offered by a travel company to the tourists (tourist kiosk), to produce documentaries about famous personalities and their works (e.g., Satyajit Ray and his films), and so on. In fact, any subject matter, where information may be expressed through various visual and audio information, may be a potential application area for a multimedia presentation. The end users who execute and watch the presentations are called the **target audience**. Different types of presentations may have different categories of audience like company employees, students, professionals, factory workers, tourists, etc. The presentation is usually played back on a personal computer, either from the hard disk or a CD/DVD. Sometimes when the audience consists of a large number of people, the presentation may be projected on a big screen using a projection system. A presentation has to be created before it can be viewed. This process is known as **multimedia production**. The production work is carried out by a team of professionals equipped with the required skill and knowledge. These professionals are called **developers** or **authors** and the development work is called **authoring**. Authoring involves a number of steps which shall be discussed later in detail.

1.3 Characteristics of a Multimedia Presentation

There is no general consensus on the exact definition of multimedia. Several definitions with varying scopes have been proposed by authors. According to Fetterman and Gupta [Fetterman, 1993], digital multimedia is defined as the integration of up to six media types in an interactive, color computing environment. In contrast, Vaughan [Vaughan, 1998] proposes a more generic definition: Multimedia is any combination of text, graphic, art, sound, animation, and video delivered by any electronic means. Thus non-interactive and non-digital devices are also included within the purview of multimedia. As per Minoli and Keinath [Minoli, 1994], multimedia is an interdisciplinary application-oriented technology that capitalizes on the multi-sensory nature of humans. (Humans are multi-sensory as they use the senses of sight, hearing, touch, smell, and taste to communicate with each other). In this book we shall regard the following as important characteristics of a multimedia presentation:

• Multiple media • Non-linearity • Scope of interactivity • Integrity • Digital Representation

These points are discussed in detail below.

1.3.1 Multiple Media

The first characteristic of a multimedia presentation is of course the use of **multiple media**. Text has been the main mode of communication for many years during the pre-multimedia era, and still continues to be one, but it is now more and more supplemented by other media, which often prove more effective. Everybody has heard the phrase: "A picture is worth a thousand words". As computer technology progressed, in addition to text, pictures were also started being used to communicate ideas. Displaying pictures required improved display devices and powerful processing capabilities. Pictures were sub-divided into two types: the real-world pictures captured by a camera, called **images**, and the hand-drawn pictures like sketches, diagrams and portraits, called **graphics**. Text, images, and graphics are together referred to as **static** elements, because they do not change over time. The usage of **dynamic** or time-varying elements like sound and movies began with further improvement in technology. Unlike text and pictures, **sound** requires audio speakers to be played back and can comprise of music, human speech or other sound effects like breaking glass, gunshots, falling water, etc. **Movies** are essentially combinations of images and sound. Instead of a single picture, movies are composed of a series of pictures which are showed one after another at a certain speed, while the sound is played together in the background. Movies are again divided into two classes: those which depict real-world incidents are called **motion pictures** (recorded on film) or **motion video** (recorded on magnetic media), and those which depict artificial or imaginary scenarios, are called **animation**. In general, sound and movies require large storage space and powerful processors to store and play back and hence their use on the PC platform became widespread only after processors like Pentium were available.

The question that can arise in one's mind is: Can any combination of media types form a multimedia presentation? The general convention followed is that a legitimate multimedia presentation should contain at least one static media, like text, image or graphics, and at least one time-varying media, like audio, video, or animation. Different types of information might warrant the use of different types of media. For example, to show the exterior shape, size, and color of a car, one would use photographic **images**; to show the internal parts of its combustion engine, one would prefer **graphics** with specific parts highlighted or detailed out if required; to depict how the pistons of its engine move or how the gearing system works, one would prefer **animation**; to exhibit the quality of its music system, one would use **audio**; and to give a sense of its pickup from 0 to 100 kmph or to demonstrate the efficiency of its braking system, to bring the car to a halt within seconds, one would use **video**. In general, one should use static media like images and graphics if they are sufficient for explaining relevant concepts, (e.g., for depicting size, shape, color of objects) and should use video only when images are not sufficient for the purpose, because video takes up a lot of space and their use should be justified (e.g., when explaining dynamic processes or moving parts). Animation is used in situations where capturing video is not practicable, e.g., when the objects are too small to be photographed (e.g., motions of subatomic particles), or too large (e.g., planetary bodies), or an entity which is difficult to photograph (e.g., electromagnetic waves). Text is usually used as bulleted points to highlight main concepts, and needs to be kept to a minimum, as people do not prefer to read a lot of text from monitor screens, while audio can be used to augment the textual points and fill in the rest of the explanation.

1.3.2 Non-linearity

Non-linearity is the capability of "jumping" or navigating from one point within a presentation to another point without appreciable delay. TV shows and motion pictures are considered **linear** presentations because the user has to watch the information being displayed, in a pre-defined sequence of frames as determined by the producer or creator of the show. He/She cannot change the sequence of frames or the timing between them if desired. Devices like the video cassette player does give us some flexibility during playback of a recorded video cassette by allowing us to go forward or backward, start, stop, or pause the show if required. However, we cannot do these instantly because the tape has to rotate forward or backward to skip irrelevant portions, which causes delays in displaying non-contiguous portions. In a multimedia presentation, the user should be able to instantly navigate to different parts of the presentation and display the frames in any way he/she chooses, without appreciable delays, due to which it is called a **non-linear** presentation. The author needs to ensure that the presentation behaves in a consistent manner irrespective of the sequence or order in which it is viewed. In reality, however, the user is restricted to a certain extent because he/she cannot arbitrarily navigate within the presentation but only along those paths already defined and allowed by the author or developer.

1.3.3 Interactivity

To make non-linearity a possibility, a user needs to interact with a presentation. A linear presentation is usually watched passively without changing the order or sequence in which it is displayed. For a non-linear presentation, a user will have to specify how he/she would like to watch the presentation, e.g., directly navigate to an area of interest instead of watching it from the beginning. Such interaction is made possible through a set of **interactivity elements** embedded within the presentation like buttons, menu items or hyperlinks. To enable interactivity, a presentation should be able to accept user inputs and allow him/her to change the sequence of screens to be displayed. Interaction is not only limited to changing the order of display, a user can also interact to start/stop individual media elements within a presentation, e.g., click a button to start a video sequence within the presentation or pause the background music. This was never possible in the other modes like a TV show or a video cassette player. Interactivity is a powerful tool and hopes to achieve much more than simply switching on/off media elements: It allows one to get involved with the presentation content. It enables the author of a learning package to create an environment within the presentation where the learner can give inputs and ask the system to provide certain outputs or feedback, simulating a tutor. This is considered very useful for applications like training and learning. It has been observed that we, human beings, tend to remember only 20% of what we read, 40% of what we see and hear over an extended period of time, but almost 80% of what we actually do. Multimedia-based education programs like CBTs (computer-based training) and CAIs (computer-aided instructions) try to exploit this property to create better training and learning modules. On one hand, they use *several media types* to explain a subject matter, rather than simply relying on text, to create a more profound influence on the human mind; on the other hand, the modules are designed so as to provide ample *scopes for interactivity* like practice sessions, drills, and tutorials. This is to get the learner more involved which, in turn, creates higher retentivity in the mind. Interactivity is considered to be one of the salient features on which the next generation e-learning tools are expected to rely for greater effectiveness.

1.3.4 Digital Representation

Audio and video cassette players use media recorded on magnetic tapes in analog format. Magnetic tapes are called **sequential access** storage devices, wherein, data is recorded sequentially along the length of the tape. When a specific portion of the data is required to be played back, the portion before that needs to be skipped. This creates a delay in accessing the arbitrary portions of the data which is unsuitable for requirements like teaching and learning. Multimedia productions provide instant access to different portions of the presentation. This is best done using a digital computer which stores data on **random access** devices like hard disks, floppy disks, and compact discs. The rotation of the disk and the radial movement of the read/write head enables instant access to different portions of data recorded on the disk surface. Hence, multimedia presentations are produced and played back on the digital platform. Each of the media types like text, image, audio, and video needs to be expressed in the digital form before it can be utilized within a presentation. Digital representation has other advantages. Software-based programs can be used to edit the digitized media in various ways to improve their appearances and compress file sizes to increase performance efficiency.

1.3.5 Integrity

One would tend to believe that any set of media files played on a computer would comprise a multimedia presentation. For example, we can edit a text document in a word processor, play a video clip in the Media Player and listen to an Audio-CD using the CD-Player. All these applications collectively, however, would not constitute a multimedia presentation. To clarify that, we introduce another important characteristic of a multimedia presentation called **integrity**. It means that although there may be several media types present and playing simultaneously, they need to be integrated or be part of a single entity, which is the presentation. We should not be able to separate out the various media and control them independently; rather they should be controlled from within the framework of the presentation. Moreover, the presentation should decide how the individual elements can be controlled. For example, it might or might not be possible to turn off the background music depending on whether the author has allowed it within the presentation.

Further Readings

- Roger L. Fetterman, Satish. K. Gupta, "Mainstream Multimedia: Applying Multimedia in Business", Van Nostrand Reinhold, 1993.
- Tay Vaughan, "Multimedia: Making it Work", McGraw Hill Osborne-Media (7th edition), 2006.
- Daniel Minoli, Robert Keinath, "Distributed Multimedia through Broadband Communication Services", Artech House Publishers, 1993.
- Multimedia @ Wikipedia [*http http://en.wikipedia.org/wiki/Multimedia*]
- New media @ Wikipedia [*http://en.wikipedia.org/wiki/New_media*]
- Digital media @ Wikipedia [*http://en.wikipedia.org/wiki/Digital_media*]
- Interactivity @ Wikipedia [*http://en.wikipedia.org/wiki/Interactivity*]
- D. Svanaes, "Understanding Interactivity", PhD thesis, 2000 [*http://dag.idi.ntnu.no/interactivity.pdf*]
- Exploring models of interactivity [*http://web.utk.edu/~sjmcmill/Research/interactivity2.doc*]
- What is interactivity [*http://www.yupingliu.com/files/papers/liu_shrum_interactivity.pdf*]
- Interactivity: a forgotten art? [*http://www2.gsu.edu/~wwwitr/docs/interact/*]
- Interactivity in teaching and learning [*http://www.elearningartdesign.org/moodle/*]
- Multisensory Research Lab [*http://www.kc.vanderbilt.edu/multisensory/*]

1.4 Hardware and Software Requirements • •

Most PCs are IBM compatible but what makes them multimedia enabled? To answer this question the **Multimedia Marketing Council** (MMC) has come out with specifications for multimedia PCs which can assure a quality playback of multimedia productions. These are known as MPC (Multimedia PC) specifications. The MMC includes a number of companies like Creative Labs Inc, Fujitsu, Media Vision, Microsoft Corp, NEC Technologies, Olivetti, Philips Consumer Electronics Co., and Zenith Data Systems. Though the detailed specifications are included in the Appendix-A, here we will discuss the general hardware and software requirements of a multimedia computer and also mention some of the accessories used for processing media components. The **MPC specifications** help to standardize multimedia computers and provide guidelines to consumers for purchasing such multimedia PCs. The MPC-1 specification was published by the MMC in 1990. In 1993, a new specification MPC-2 was published as an enhanced standard, which continues alongside MPC-1 as the basic standard (instead of replacing MPC-1). MPC-3 specifications further extend the requirements for multimedia PCs. However, it must be noted that since these specifications were formulated in the 1990s, they might be outdated nowadays.

Hardware and software requirements of a multimedia PC can be categorized into two classes. Multimedia **playback** usually requires smaller amount of resources, those which are sufficient for viewing an existing multimedia presentation. Multimedia **production** generally requires greater and more powerful resources and should fulfill all requirements for designing and developing a multimedia presentation. Recommendations and minimum requirements for multimedia playback and production are briefly listed below:

Multimedia Playback

- Processor: Pentium-IV class
- Hard disk: 80 GB
- CD drive (32X), DVD drive (4X)
- OS: Windows XP / Vista / 7
- Document readers: Adobe PDF reader 9
- RAM: 1 GB
- Monitor: SVGA with 1024 × 768 display and 24-bit color depth
- Sound card with attached speakers
- Web browser: IE 7, Firefox 4, Chrome 5, Opera 11, Safari 5
- Media player: Windows Media Player 9, Apple QuickTime Player 7, Real Player 10, VLC player 1

Multimedia Production

- Processor: Pentium-IV class
- Hard disk: 160 GB
- CD drive (32X), DVD drive (4X)
- OS: Windows XP / Vista / 7
- Document readers: Adobe PDF reader 9
- Image scanners with 600 dpi resolution
- Document editors: MS Word 2007
- Image-processing software: Adobe Photoshop 7, CorelDraw 10
- Video processing software: Adobe Premiere 6, Avid Media Composer 4
- 3D animation software: Autodesk 3D Studio Max 7
- RAM: 2 GB
- Monitor: SVGA with 1024 × 768 display and 24-bit color depth
- Sound card with attached speakers and microphones
- Web browser: IE 7, Firefox 4, Chrome 5, Opera 11, Safari 5
- Media player: Windows Media Player 9, Apple QuickTime Player 7, Real Player 10, VLC player 1
- Still digital camera with 5 megapixel resolution
- Video capture card with composite video/S-video interfaces
- Sound-processing software: Sony SoundForge 7, Adobe Audition 3
- 2-D animation software: Adobe (Macromedia) Flash 8
- Authoring software: Adobe (Macromedia) Director 8

Further Readings

- Multimedia PC @ Wikipedia [http://en.wikipedia.org/wiki/Multimedia_PC]
- Multimedia PC Level 3 specifications [http://support.gateway.com/s/SOUND/U0064 7/U0064725.shtml]
- MPC level 1 and 2 specifications [http://support.microsoft.com/kb/106055]

1.5 Uses of Multimedia

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

Home Entertainment Application of multimedia technology related to home entertainment includes computer-based games for children, interactive encyclopedias, story-telling, cartoons, etc. Games are one of the best applications of multimedia because of the high amount of interactivity involved. Audio/Video-on-demand systems employing multimedia technologies would enable users to specify the movies they would like to watch, not simply by their names, but by a host of other criteria like actors/actresses, director, music director, genre (like action movie, detective thriller or science fiction), year of release, awards won, etc.

Education Education-based multimedia applications include learning packages and simulation of lab experiments (especially those which cannot be easily performed in practice). Different aspects of the course curriculum which cannot be explained or grasped easily through simple text and images (as offered by the books) could be presented through video clips, animation, 3D modeling, audio annotations, etc., for making them easily understandable. Moreover, expertise of the best instructors could be used to create content and distribute them to thousands of students all over the globe as part of "distance learning" or "e-learning" programs. The multi-sensory perceptions of such study material is expected to provide a good grasp of the subject matter and interactivity elements to provide for better retention.

Industry Computer based training (CBT) packages for employees, both technical and marketing. Successful organizations are required to maintain a high level of staff training and development. Although multimedia presentations could be expensive, reduction in costs associated with holding conventional courses could make these feasible. Some of the advantages of these courses include: (a) Many people can use each of these courses. (b) They do not need to spend time away from office. (c) People can learn at their own paces. (d) Full time instructors are not required. (e) Because the best instructors could be used to make these CBTs, they could be of a high quality. (f) Experimental setups could be reduced as these can be simulated. (g) Different groups of employees could be evaluated based on standardized evaluation procedures.

Information Kiosk In an information kiosk, information is accessed through a touch screen and viewed on a monitor. It can be a multi-lingual product catalog with provisions for placing orders. It can also be used for dispensing important information, e.g., train timings at a railway station. Shopping kiosks would provide consumers with access to an electronic shopping center offering a wide range of information, products, and services. Links to payment systems and distribution channels would enable the kiosks to accept orders and lead to the development of e-commerce practices. Kiosks can also be used to capture statistical data for an in-depth market research to be carried out on consumer trends.

Corporate Presentations Corporate presentations for emphasizing the salient features and activities of a company, its products, its business partners like suppliers and retailers, can be built by incorporating multimedia elements along with textual descriptions. This could be as simple as charts, logos, graphs of something like an audio/video presentation of annual general meeting activities, speeches from the general manager, or features of new product ranges.

Business Items difficult to stock like glassware, industrial equipment, etc., can be displayed to prospects by company sales people through multimedia presentation. Real-estate agents can display interior and exterior of buildings along with necessary information like dimension and price. Architects and engineers can transform blueprints into buildings and products. The benefits include saving on space, inventory, and distribution.

Travel and Tourism Travel companies can market packaged tours by showing prospective customers glimpses of the places they would like to visit, details of lodging, food, special attractions, etc. Hotel owners can utilize the technology to display details of facilities offered at various hotels at different locations. Museums and art galleries may provide digital versions of their specimens for public viewing, exhibition, auction, etc. A multimedia system implementing an intelligent travel agent software would enable the user to specify his/her travel needs and budget, and then ask the system to find places that satisfy these requirements. The system would then list out possible destinations that meet these criteria and create multimedia presentations about one or more of these.

Electronic Commerce Customized presentations for home and industrial products could be created and distributed to prospective customers. Customers can compare different products in relation to their quality, price, and appearance without leaving their homes and offices. Once they decide on purchasing an item, hyperlinks could connect them directly to an on-line store where they could place their orders. Additionally, such catalogs can also server as expert systems, which could take a list of criteria from customers and provide a list of products which satisfy them. This would serve to guide the customers into choosing the appropriate items they want. Such systems can also be useful in providing after-sales services by taking in a list of symptoms from a user and suggesting to them the possible reasons for the problems in their machines like automobiles, washing machines, or music systems. It could also suggest simple steps for troubleshooting.

Communication and Networks These applications allow real-time interactions between people who need to work together but cannot be in the same place at the same time. In order to cater to these needs, many public and private networks support services like voice-mail and tele-conferencing. Voice-mail enhances the effectiveness of normal e-mail by allowing messages to be composed and sent using audio instead of text, while tele-conferencing is used where the parties need to communicate in real time. Another real time application coming into prominence is Internet Telephony. It involves establishing long distance telephone calls by logging into a local server and using the Internet to connect to a remote server. Other similar uses include groupware/whiteboard applications in intranets, e-learning applications in virtual classrooms, audio/video-on-demand applications for entertainment.

Medicine and Healthcare Multimedia technologies can be used to prepare high-quality magnetic resonance 3D images of human bodies and practice complicated surgical procedures. Archives of X-ray images, CT scans, ultra-sonography images, etc., will enable doctors to provide better consultations, and could serve as an expert system. Transmission of these images over geographical boundaries can help in the growth of “tele-medicine” where doctors separated by distance can consult with one another and find solutions to complicated problems.

Engineering Applications ultimedia is used extensively in designing mechanical, electrical, electronic, and architectural parts by the use of Computer Aided Design (CAD) and Computer Aided Manufacturing (CAM) applications. They enable engineers to develop software representations of products from various viewpoints, rotate scale and move parts and portions, zoom on to critical parts, and try out various combinations before deciding on the final product implementation.

Content Based Storage and Retrieval (CBSR) Systems In recent times content-based storage and retrieval systems have assumed importance. Traditionally, data-searching activities have been performed on textual databases by string matching. As large repositories of media elements like images, audio, and video are growing up all over the world, efficient methods of searching non-textual media are being developed. An example is the matching of a fingerprint from police records to identify a criminal (fingerprint-recognition). Other similar applications include identifying a person from a photograph (face-recognition), or a recorded speech (speaker-recognition), or a video clip from surveillance cameras inside a store or bank. These techniques are proving to be immensely helpful especially in cases where there are no additional or prior information about the person in question.

Further Readings

- Multimedia home entertainment system [*http://news.ecoustics.com/bbs/messages/10381/15048.html*]
- Interactivity in Teaching and Learning [*http://www.elearningartdesign.org/moodle/*]
- Edutainment @ Wikipedia [*http://en.wikipedia.org/wiki/Edutainment*]
- Computer based training [*http://en.wikipedia.org/wiki/Computer-based_training*]
- Touchscreen [*http://en.wikipedia.org/wiki/Touchscreen*]
- Information kiosk systems [*http://www.touchscreens.com/products-kiosk.html*]
- Dublin tourism multimedia [*http://www.visitdublin.com/multimedia/dublin.aspx*]
- Thailand tourism multimedia [*http://www.tourismthailand.org/multimedia/*]
- E-commerce revolution [*http://www.shopmotion.com/e-commerce/en/factsheet.pdf*]
- Health informatics @ Wikipedia [*http://en.wikipedia.org/wiki/Health_informatics*]
- Bioinformatics @ Wikipedia [*http://en.wikipedia.org/wiki/Bioinformatics*]
- Electronic health records @ Wikipedia [*http://en.wikipedia.org/wiki/Electronic_Health_Records*]
- CAD @ Wikipedia [*http://en.wikipedia.org/wiki/Computer-aided_design*]
- CAM @ Wikipedia [*http://en.wikipedia.org/wiki/Computer-aided_manufacture*]
- List of computer aided technologies [*http://en.wikipedia.org/wiki/Computer-aided_technologies*]
- CBIR @ Wikipedia [*http://en.wikipedia.org/wiki/Content-based_image_retrieval*]
- Query by Image and Video Content [*http://www.computer.org/portal/web/csdl/doi/10.1109/2.410146*]
- Content based image retrieval [*http://www.jisc.ac.uk/uploaded_documents/jtap-039.doc*]
- Similarity functions for image retrieval [*http://www.svcl.ucsd.edu/publications/journal/2004/it04/it04.pdf*]
- Perspectives on visual information retrieval [*http://www.ims.tuwien.ac.at/~hme/papers/ei2004-vir.pdf*]

1.6 Analog and Digital Representations

The real world that we perceive with our senses is essentially an **analog** world, which roughly means "continuous". The various phenomena like light, heat, magnetism, sound, and electricity all appear as continuous entities. However, when we try to represent some of these within a **digital** computer, we need a radically different representation, which is essentially non-continuous or "discrete". Digitization parameters like **sampling rate** and **bit depth** decide the final quality of the digital media produced and hence knowledge of these are essential for multimedia production. Theorems like the Nyquist's Sampling Theorem provide guidelines for choosing these parameters efficiently.

If n be the bit-depth then number of steps is 2^n. If S be peak-to-peak height of the wave then height of each step is $S/2^n$ which is equal to the quantization error e. Hence $A_S = S$ and $A_N = e = S/2^n$.

$$SNR = \frac{A_S}{A_N} = \frac{S}{e} = 2^n \tag{1.3}$$

$$SNR_{dB} = 20\log_{10}\left(\frac{A_S}{A_N}\right) = 20\log_{10}(2^n) = 20n\log_{10}(2) = 20n(0.3) = 6n$$

Equation (1.3) implies that if bit-depth is increased by 1 during digitization, the signal-to-noise ratio increases by 6 dB.

Further Readings

- Quantization @ Wikipedia [*http://en.wikipedia.org/wiki/Quantization_(signal_processing)*]
- Quantization error @ Wikipedia [*http://en.wikipedia.org/wiki/Quantization_error*]
- Quantization noise [*http://www.dsplog.com/2007/03/19/signal-to-quantization-noise-in-quantized-sinusoidal/*]
- SNR @ Wikipedia [*http://en.wikipedia.org/wiki/Signal-to-noise_ratio*]
- Decibel @ Wikipedia [*http://en.wikipedia.org/wiki/Decibel*]
- Dynamic range @ Wikipedia [*http://en.wikipedia.org/wiki/Dynamic_range*]
- Dynamic range [*http://www.seis.com.au/TechNotes/TN200410A_SNR.html*]

1.10 Visual Display Systems

Multimedia consists of a large number of visual media like text, image, graphics, video, animation, etc. The primary user interface hardware for displaying these visual media is the Visual Display System (VDS). The VDS enables us to generate, process, and present these visual media. It also plays a role in deciding the quality of the final presentation. To plan a multimedia project, it is essential for us to be aware of the basic features and limitations of the VDS. The VDS consist of the following components: a monitor, an adapter card, and an adapter cable.

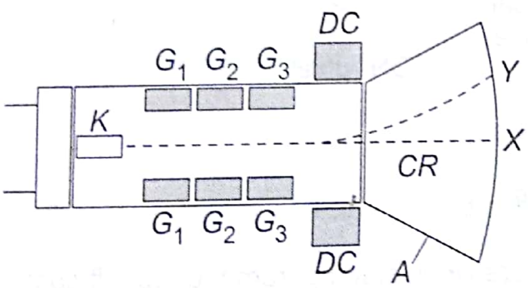


Fig. 1.13 *Components of a monochrome CRT*

1.10.1 Cathode Ray Tube (CRT)

By far, the most common type of monitors used with multimedia desktop systems are based on the picture tube called the **Cathode Ray Tube** (CRT). It is essentially a vacuum-sealed glass tube containing two electrodes inside it, the cathode (K) or negative electrode and the anode (A) or positive electrode. See Fig. 1.13. The front face of the tube is coated with a chemical compound called **phosphor** arranged in the form of a rectangular grid of a large number of dots. Phosphor has the property of emitting a glow of light when it is hit by charged particles like electrons. Each phosphor dot is called a **pixel**, short for "picture element", which is responsible for producing a dot of light on the monitor screen.

A CRT capable of producing a black-and-white image on the screen is called a **monochrome CRT**. A high voltage of the order of 18 kV is maintained between the cathode and anode. This produces a beam of electrons, known as **cathode rays**, from the cathode towards the anode. The beam of electrons is controlled by three other positive terminals. The **control grid** (G_1) helps to control the amount of electrons in the beam and thereby determine its strength, the **accelerating grid** (G_2) provides acceleration to the electrons in the forward direction, and the **focusing grid** (G_3) focuses the beam to a single point X on the screen ahead, so that the diameter of the beam is equal to the diameter of a single dot of phosphor. As the beam hits the phosphor dot, a single glowing spot of light is created at the center of the screen. This spot of light is known as a **glowing pixel**. When current flows through a pair of deflection coils (DC), the electrical field produced interacts with the electron beam, thereby deflecting it from its original path to another point Y. One of the coils, called the **horizontal deflection coil**, moves the beam horizontally across the screen and the other coil, called the **vertical deflection coil**, moves the beam vertically along the height of the screen. When both these coils are energized in required proportions, the electron beam can be moved in any direction, thus generating a single spot of light at any point on the CRT screen.

To draw an image on the screen, the electron beam starts from the upper left corner of the screen and sequentially moves over each pixel row from left to right. This is referred to as **forward trace**. During this phase, the electron beam is sometimes switched ON and sometimes remains OFF. When the beam is ON, the pixels over which it is moving get hit by the electrons and start glowing (active pixels). The pixels over which the beam remains OFF do not get a dose of the electrons and therefore remain dark. At the end of each horizontal line, the beam gets switched OFF and retraces diagonally to the beginning of the next row. This phase is referred to as **horizontal retrace**. The switching OFF conserves power and avoids activation of unwanted pixels. At the beginning of the next line, it is again switched ON and begins the next trace. The process continues until the beam reaches the lower right corner of the screen, after which it is again switched OFF and moves diagonally back to the starting point. This is referred to as **vertical refresh**. The entire process from beginning to end is called **raster scanning**. See Fig. 1.14. To display an image on the screen, as the electron beam moves over the pixel rows, some of the pixels are turned "ON" or activated while the other pixels remain "OFF" or deactivated. Activated pixels emit a glow of light. The image is recognized on screen by the glowing pixels over a dark background of deactivated pixels.

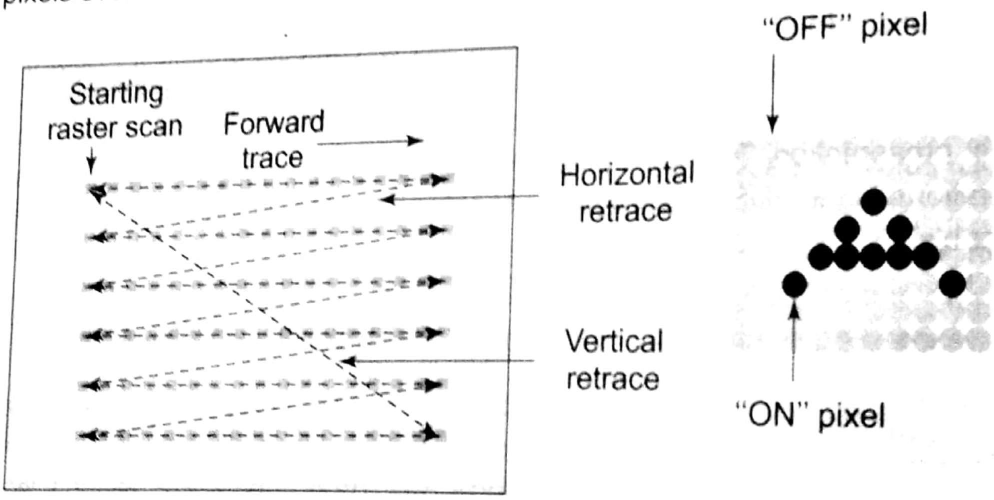


Fig. 1.14 *Raster scanning*

The electron beam is said to produce a complete **frame** of picture when starting from the top-left corner it moves over all the pixels and returns back to the starting point. The human brain has the capability to hold on to the image of an object before our eyes for a fraction of a second even after the object has been removed from before our eyes. This phenomenon is called **persistence of vision**. As the beam moves over each pixel, the glow of the pixel dies down although its image persists in our eyes for some time after that. So if the beam can come back to the pixel before its glow has completely disappeared, to us it will seem that the pixel is glowing continuously. It has been observed that we see a steady image on the screen only if the electron beam returns to its starting point within 1/50 th or 1/60 th of a second, i.e., about 50–60 frames are generated per second. The monitor is then said to have a **refresh rate** of 50 or 60 Hz. A monitor with a refresh rate of less than 50 Hz produces a perceptible **flicker** on the screen, which is caused by the previous image fading from the eye retina before the next image could be generated. Such flickering monitors should be avoided as it creates a strain on the eye. In practice, the frame refresh rate is usually determined by the frequency of the mains electric supply, which is either 50 Hz in Europe and 60 Hz in America.

The working principle of a **color CRT** is similar to that of a monochrome CRT, except that here each pixel consists of three colored dots instead of one and is called a **triad**. These processed phosphors produce lights of colors *red*, *green*, and *blue* (in short, *RGB*) and are called **primary colors**. These are so called because it has been experimentally observed that these three colored lights can combine in various proportions to produce all other colors that we see. Corresponding to the three dots, there are also three electron beams from the electrode (also called **electron gun**), each of which falls on the corresponding dot. As each of the three beams hits the corresponding dot in various intensities, they produce different proportions of the three primary colored lights, which mix together to create the sensation of a specific color in our eyes. See Fig. 1.15. Our eyes cannot distinguish the individual dots but see their net effect as a whole. Also, *red* and *blue* in equal proportions produce color *magenta*; blue and green likewise produce *cyan*, and green and red produce *yellow*. These colors, *cyan, magenta, yellow* (*CMY*) produced by mixing equal proportions of primary colors, are called **secondary colors.** All three primary colors in equal proportion produce *white*, while their absence leads to the color sensation *black*. A perforated screen called a **shadow mask** prevents the beams from falling in the gap between the dots. As the electron beam sweeps across the shadow mask, it gets heated up due to the electron bombardment on it. Expansion due to heating may disturb the alignment of the perforated holes with the phosphor dots, thereby causing the mask to lose its utility. To prevent this, the mask is made up of a special alloy of iron and nickel called **invar** which does not expand on heating.

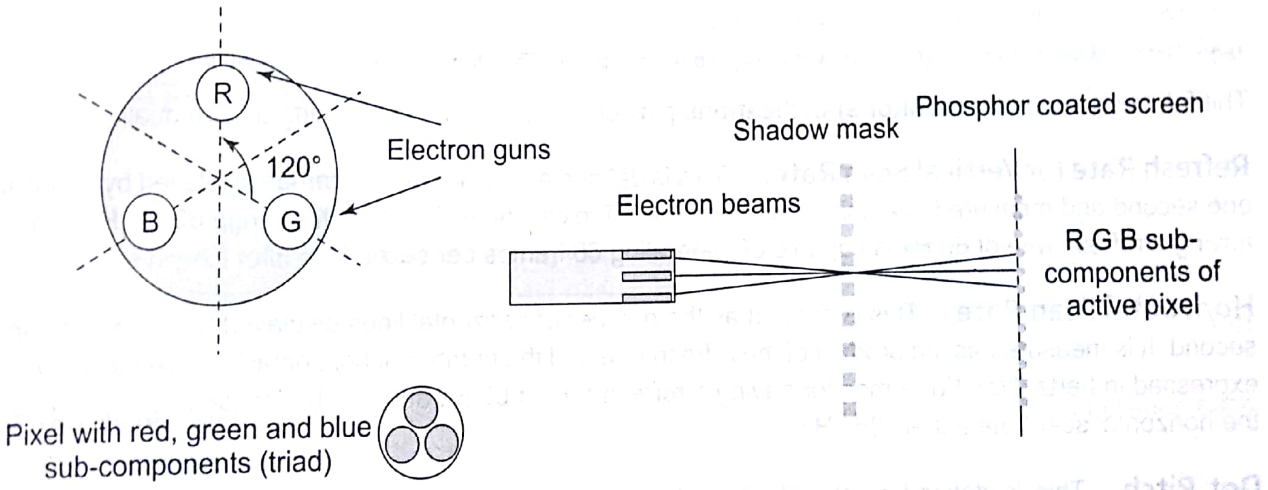


Fig. 1.15 *Components of a color CRT*

Ordinarily, a monitor of a refresh rate of less than 60 Hz would produce a perceptible flicker on the screen. **Interlacing** is a technique by which monitors of lower refresh rates can be made to produce images comparable in quality to that produced by monitors of higher refresh rates. To reduce the flicker, each frame produced by a raster scan is split into two halves, each called a **field**. The first field is made up of only odd-numbered pixel rows, i.e., 1, 3, 5,... and is called the **odd-field**. The second field is made up of even-numbered pixel rows, i.e., 2, 4, 6,... and is called the **even-field**. See Fig. 1.16. Since each field is made up of only half the total number of pixel rows, generating a field takes only half the duration as that for generating a frame. Due to persistence of vision, this kind of arrangement leads to a smooth blending of the rows of each field and helps reduce the flickering effect prominent is monitors of low refresh rates.

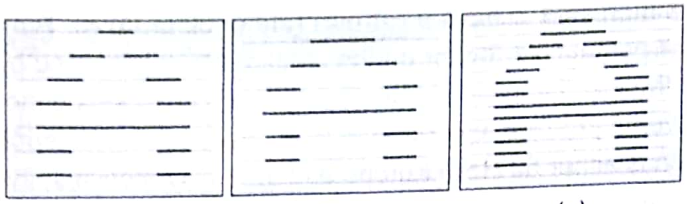


Fig. 1.16 *Interlacing of (a) Odd-field, (b) Even-field, and (c) Combined-fields*

A monitor which uses the interlacing principle to produce images is called an **interlaced monitor**. Modern monitors are mostly non-interlaced and should be preferred over interlaced monitors for purposes related to multimedia. Such type of monitors are referred to as **progressive**, implying that the entire frame is progressively scanned from start to end instead of being split into fields. To distinguish between progressive and interlaced scanning, the notations *i* and *p* are used. For example, 480*i* and 576*i* are standard television formats for NTSC and PAL, while 720*p* and 1080*p* are HDTV formats indicating that there are 720 and 1080 active lines on the screen that are scanned in a progressive mode.

Further Readings

- CRT @ About.com [*http://inventors.about.com/od/cstartinventions/a/CathodeRayTube.htm*]
- How computer monitors work [*http://computer.howstuffworks.com/monitor.htm*]
- How CRT and LCD monitors work [*http://www.bit-tech.net/hardware/2006/03/20/how_crt_and_lcd_ monitors_work/1*]
- CRT @ PCGuide [*http://www.pcguide.com/ref/crt/const.htm*]
- Temporal rate conversion [*http://www.microsoft.com/whdc/archive/TempRate.mspx*]

The following important **monitor specifications** play decisive roles in determining picture quality.

Refresh Rate (or Vertical Scan Rate) This is defined as the number of frames displayed by a monitor in one second and measured in a unit called hertz (Hz). Typical values lie within the range 60–70 Hz. A monitor having a refresh rate of 60 Hz is capable of generating 60 frames per second.

Horizontal Scan Rate This is defined as the number of horizontal lines displayed by the monitor in one second. It is measured as the product of the refresh rate and the number of horizontal lines on the screen, and expressed in hertz (Hz). For a monitor having a refresh rate of 60 Hz and 600 horizontal lines on the screen, the horizontal scan rate equals 36 kHz.

Dot Pitch This is defined as the shortest distance between two neighboring pixels (for monochrome monitors) or triads (for color monitors). It is usually of the order of 0.25 mm to 0.4 mm.

Pixel Addressability It is defined as the total number of pixels that can be addressed on the screen. It is measured by the product of number of horizontal pixel rows and the number of pixels per row. Modern monitors have pixel addressabilities of 640 × 480 or 800 × 600 or higher.

Aspect Ratio This is defined as the ratio of the horizontal number of pixels (number of pixels along a horizontal row) to the vertical number of pixels (number of horizontal rows of pixels). For a computer monitor or a TV screen, the ratio is fixed and equals 4:3. For certain display media like movie theatres and High Definition TVs, the ratio is 16:9.

Monitor Size The size of the monitor is defined as the longest diagonal length of the monitor. Standard computer monitors are usually between 15 inches to 20 inches in size.

Resolution This is defined as the total number of pixels per unit length of the monitor in the horizontal direction. It is expressed in a unit called **dots per inch (dpi)**. Standard monitors usually have a resolutions of 72 dpi–96 dpi.

Color Depth This is a measure of the total number of colors that can be displayed on a monitor and depends on the total number of varying intensities of the electron beam of a CRT. Modern color monitors usually have a color depth of 24 bits implying that they can display up to 2^{24} or 16.7 million colors.

Some **mathematical formulations** associated with monitor specifications are discussed below:

If w be the width and h the height of the monitor in pixels then pixel addressability P is given by

$$P = w.h \tag{1.4}$$

Let b be the bit-depth or color-depth supported by the monitor, in bits. Then total number of colors C that can be displayed is given by

$$C = 2^b \quad \rightarrow \quad b = \log_2(C) \tag{1.5}$$

The memory M in bits, required to store one screen of information of the monitor, is given by

$$M = w.h.b = P.b \tag{1.6}$$

Let W and H be the width and height of the monitor in inches. Then its size D, i.e., length of its diagonal, is given according to properties of right-angled traingles as

$$D = \sqrt{W^2 + H^2} \tag{1.7}$$

The resolution R in dots per inch (dpi) is the number of pixels per unit width or height of the monitor. In most cases, these are assumed to be equal unless specified otherwise.

$$R = \frac{w}{W} = \frac{h}{H} \tag{1.8}$$

If $m{:}n$ be the aspect ratio of the monitor ($m > n$) then we have

$$\frac{w}{h} = \frac{m}{n} \tag{1.9}$$

Let d be the dot-pitch of the pixels. Since R is the number of pixels in one inch, the distance between each pair of pixels is given by

$$d = \frac{1}{R} \tag{1.10}$$

1.10.2 Video Adapter Card

The video adapter is an expansion card, which usually sits on a slot on the motherboard. Nowadays, it can also be found integrated with the motherboard. It acts as an interface between the processor of the computer and the monitor. It accepts data from the processor through a communication channel called the bus, which connects to the slot on which the card sits. Earlier, machines used a bus called ISA (Industry Standard Architecture), but modern machines use an improved bus called PCI (Peripheral Components Interconnect) and in recent times, AGP (Accelerated Graphics Port). The card communicates with the monitor via the adapter cable, which remains connected to the card through a 15-pin D-type connector. See Fig. 1.17. The essential components of an adapter card are video RAM, graphics controller, and DAC.

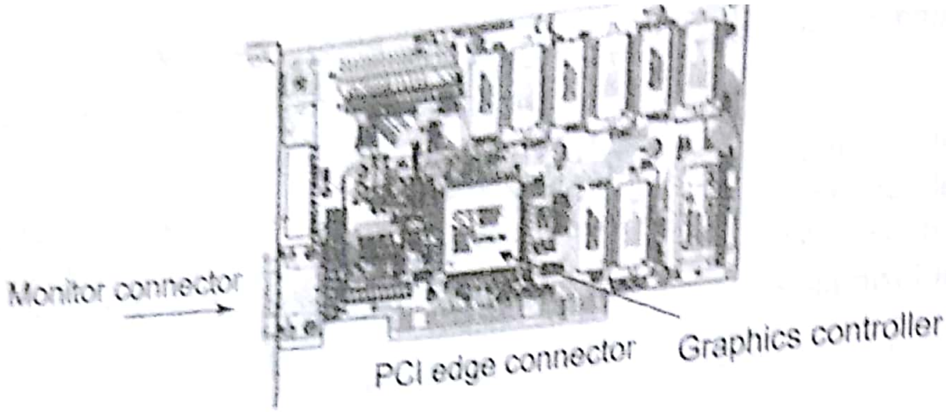


Fig. 1.17 *Video adapter card*

The **Video RAM** (VRAM) is a bank of memory (dynamic RAM) within the adapter card used for storing pixel attributes generated from the CPU. The amount of memory should be sufficient to hold the attributes of all the pixels on the screen and depends on the pixel addressability as well as the color depth. Thus, for an 8-bit image displayed at 640 x 480 mode, the minimum amount of display memory required is approximately 1 MB. The **graphics controller** is the main processor within the adapter card, responsible for coordinating the activities of all other components of the card. For the earlier generation video cards, the controller simply passed on the data from the processor to the monitor after conversion. For modern accelerated video cards, the controller also has the capability of manipulating the image data independently of the central processor. The controller also generates the synchronization (or sync) signals to control the movement of the electron guns of the CRT monitor. The **Digital-to-Analog Converter** (DAC) is one of the main reasons why the adapter is there. It takes the digital data stored in the VRAM and converts it to analog signals before sending them outwards to the monitor. Since the electron guns of the monitor need voltage signals to generate electron beams, they would not be able to function if digital data is directly fed to them.

The first PCs used a display adapter called **Mono Display Adapter** (MDA) that provided only 80 columns by 25 rows of monochrome text. It had 4 KB of video memory. Soon IBM introduced the **Color Graphics Adapter** (CGA) which allowed low-resolution color graphics with low-quality text. CGA supported a text mode of 80 columns by 25 rows of characters, and a graphics mode of 640 pixels by 200 pixels—usually of two or four colors. To compete against the CGA, the **Hercules Graphics Adapter** (HGA) was developed and provided MDA compatible text modes and a single 720 pixels by 348 pixels monographics mode. The next card to be introduced was the **Enhanced Graphics Adapter** (EGA). The EGA standard included all of the MDA and CGA modes and increased pixel addressability to 640 × 350 in 16 colors. All these early standards are obsolete nowadays, being subsequently superseded by the VGA and SVGA. See Table 1.1.

Table 1.1 *Early adapter specifications*

Year	Adapter	Text mode (columns × lines)	Graphics resolution	Colors	Memory in KB
1981	MDA	80 × 25	-	-	4
1981	CGA	80 × 25	640 × 200	16	16
1982	HGA	80 × 25	720 × 348	2	64
1984	EGA	80 × 25	640 × 350	16	256
1987	VGA	80 × 25	640 × 480	16	256
1989	SVGA	80 × 25	800 × 600	256	512

Video Graphics Array (VGA) refers to the display hardware introduced by IBM in 1987, and has become synonymous with 640 by 480 resolution. VGA was the last standard for display resolution introduced by IBM and has today become the lowest resolution that a PC is expected to support. VGA is referred to as an "array" instead of an "adapter" like its predecessors MGA, CGA, and EGA because it was implemented on a single chip instead of a separate adapter card. This allowed the VGA to be placed directly on the PC motherboard instead of requiring a adapter card to be placed on a slot on the motherboard. The VGA specifications include (a) 256 KB of video RAM, (b) 640 × 480 in 16 colors, 640 × 350 in 16 colors, 320 × 200 in 16 colors, 320 × 200 in 256 colors, (c) Clock frequency of 25.175 MHz, (d) Line frequency of 30 KHz, (e) Frame rate of 60 Hz, and (f) Signal levels of 0.7 V peak to peak.

Super VGA (SVGA) is an extension to the VGA standard defined by VESA (Video Electronics Standards Association), an open consortium set up to define video standards. Nowadays SVGA has become synonymous with a video resolution of 800 by 600 in 16 colors. SVGA uses the same 15-pin D-subminiature connector as the VGA. In the original DE-15 connector, the monitor bits 0, 1 and 2 were used to identify mono and color monitors and those which did and did not support the 1024 by 768 display mode. Later VESA modified the pin assignments according to VESA DDC standard. VESA DDC (Display Data Channel) is a method that allows the monitor to communicate its supported display mode to the adapter, and hence automatically change display parameters to provide a plug-and-play interface.

Extended Graphics Array (XGA) is an IBM standard introduced in 1990 and is synonymous with a display resolution of 1024 × 768. The initial XGA was expanded with two display modes: 800 × 600 with high color, i.e., 16 bits/pixel or 65536 colors; and 1024 × 768 with 256 colors. Subsequently, a number of XGA successors were developed supporting various screen dimensions, aspect ratios and megapixel counts. A summary of the various resolutions along with their megapixel counts is given in Table 1.2.

Table 1.2 *XGA Video resolutions*

Code	Name	Aspect ratio	Width	Height	Megapixels
VGA	Video Graphics Array	4:3	640	480	0.3
SVGA	Super VGA	4:3	800	600	0.5
XGA	eXtended Graphics Array	4:3	1024	768	0.8
W-XGA	Wide XGA (720p)	8:5	1280	800	1.0
S-XGA	Super XGA	5:4	1280	1024	1.3
U-XGA	Ultra XGA	4:3	1600	1200	1.9
WU-XGA	Wide Ultra XGA (1080p)	16:9	1920	1200	2.3
QXGA	Quad XGA (4 * XGA)	4:3	2048	1536	3.1
W-QXGA	Wide Quad XGA	8:5	2560	1600	4.1
S-QXGA	Super Quad XGA	5:4	2560	2048	5.2
U-QXGA	Ultra Quad XGA	4:3	3200	2400	7.7
WU-QXGA	Wide Ultra Quad XGA	8:5	3840	2400	9.2
HXGA	Hex XGA (16 * XGA)	4:3	4096	3072	12.6
W-HXGA	Wide Hex XGA	8:5	5120	3200	16.4
S-HXGA	Super Hex XGA	5:4	5120	4096	21.0
U-HXGA	Ultra Hex XGA	4:3	6400	4800	31.0
WU-HXGA	Wide Ultra Hex XGA	8:5	7680	4800	37.0

Further Readings

- Video card @ Wikipedia [*http://en.wikipedia.org/wiki/Video_card*]
- How graphics cards work @ HowStuffWorks [*http://www.howstuffworks.com/graphics-card.htm*]
- MDA Notes [*http://www.seasip.info/VintagePC/mda.html*]
- CGA Notes [*http://www.seasip.info/VintagePC/cga.html*]
- Technical details of CGA [*http://nemesis.lonestar.org/reference/video/cga.html*]
- EGA @ Wikipedia [*http://en.wikipedia.org/wiki/Enhanced_Graphics_Adapter*]
- HGC @ Wikipedia [*http://en.wikipedia.org/wiki/Hercules_Graphics_Card*]
- VESA official site [*http://www.vesa.org/*]
- VESA standards FAQ [*http://vesa.org/FAQ/standardsFAQ.htm*]
- VESA standards summaries [*http://www.vesa.org/Standards/summaries.htm*]
- VESA DDC pinout [*http://pinouts.ru/Video/VGAVesaDdc_pinout.shtml*]
- Video display standards [*http://www.cknow.com/cms/ref/video-display-standards.html*]
- Comparison of next-gen display devices [*http://www.tgdaily.com/hardware-features/33161-a-comparison-of-next-gen-display-interfaces*]

1.10.3 Video Connector and Cable

The video adapter card is connected to the monitor with the video adapter cable. The standard cable and connector used today is based on VGA signals. VGA uses a 15-pin D-subminiature connector (called VGA-15 or DE-15) to convey signal between the graphics controller and the monitor, although there are other variants like the 9-pin DE-9 connector and the mini-VGA connectors used for laptops. It carries the video signals (one for monochrome and three for color) from the adapter card to the monitor where these are used to activate the electron beams of the electron gun. Additionally, the cable also carries two synchronization signals (horizontal and vertical) which are fed to the deflection coils on the neck of the CRT to control the movement of the electron beams. All these signals are generated by the adapter card in such a way that the monitor can reproduce the image described by the pixel attributes stored in the VRAM. See Fig. 1.18.

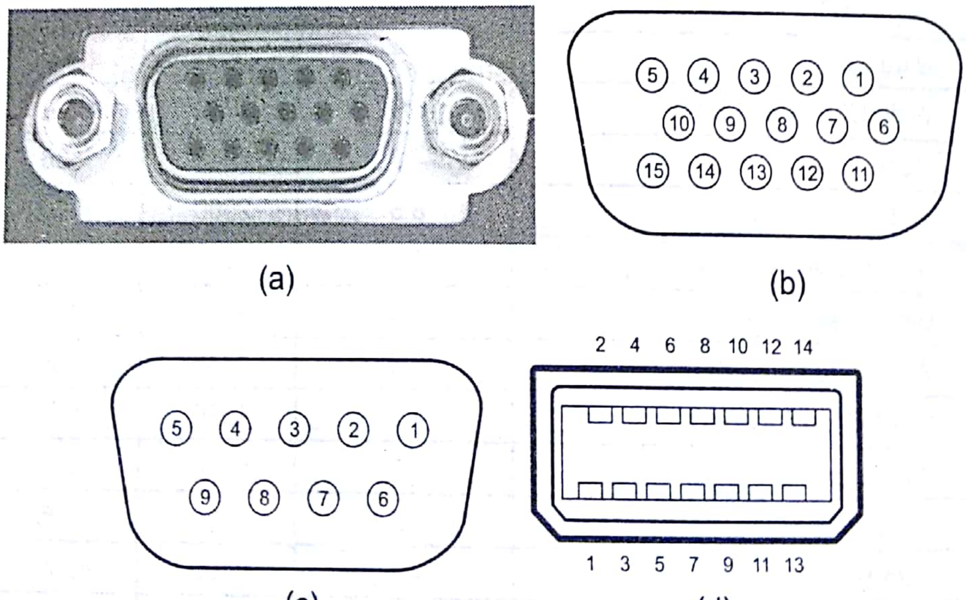


Fig. 1.18 *(a) 15-pin VGA connector (b) Pins of VGA-15 (c) Pins of VGA-9 (d) Mini-VGA connector*

Table 1.3 *Pin signal assignments of DE-15*

1	2	3	4	5
Red	Green	Blue	Monitor ID bit 2	Ground (Hsync)
6	**7**	**8**	**9**	**10**
Red_return	Green_return	Blue_return	+5 V DC	Ground (Vsync)
11	**12**	**13**	**14**	**15**
Monitor ID bit 0	Monitor ID bit 1	HSync	VSync	Monitor ID bit 3

Further Readings

- VGA 15 pinout [*http://pinouts.ru/Video/VGA15_pinout.shtml*]
- VGA 9 pinout [*http://pinouts.ru/Video/VGA9_pinout.shtml*]
- VGA chipset reference [*http://osdever.net./FreeVGA/vga/vga.htm*]
- VGA interface and video signals [*http://martin.hinner.info/vga/*]
- VGA electrical FAQ [*http://www.microvga.com/faq/electrical*]
- D-subminiature @ Wikipedia [*http://en.wikipedia.org/wiki/D-subminiature*]

1.10.4 Accelerated Graphics Port (AGP)

To combat the eventual saturation of the PCI bus with video information a new interface has been pioneered by Intel, designed specifically for the video subsystem. AGP was developed in response to the trend towards greater and greater performance requirements for video. AGP gets around these problems by two approaches. First, it removed the pressure from the existing PCI bus by providing an additional high-speed dedicated bus exclusively for transferring video data called the **AGB bus**. The PCI bus could transfer data at a maximum speed of 133 MB/second but the AGP improves on that by providing a dedicated bus operating at over 500 MB/second. To use the new bus, the AGP provides a new slot on the motherboard, called the **AGP slot**, that looks similar to the PCI slot but slightly offset further from the edge of the motherboard. The AGP slot is not pin compatible with the PCI slot, and thus existing PCI cards cannot be inserted into the AGP slot. Secondly, AGP removes the limitations due to the video memory by **extending memory capabilities**, allowing a part of the main memory to be used for storing video data. A part of the main system memory is reserved and works as a cache for the video memory on the adapter. This together provided a large storage area for storing pixel attributes and allowed for much faster data transfers than possible by the PCI bus, which had to fetch additional data required by the video memory from the hard disk. See Fig. 1.19.

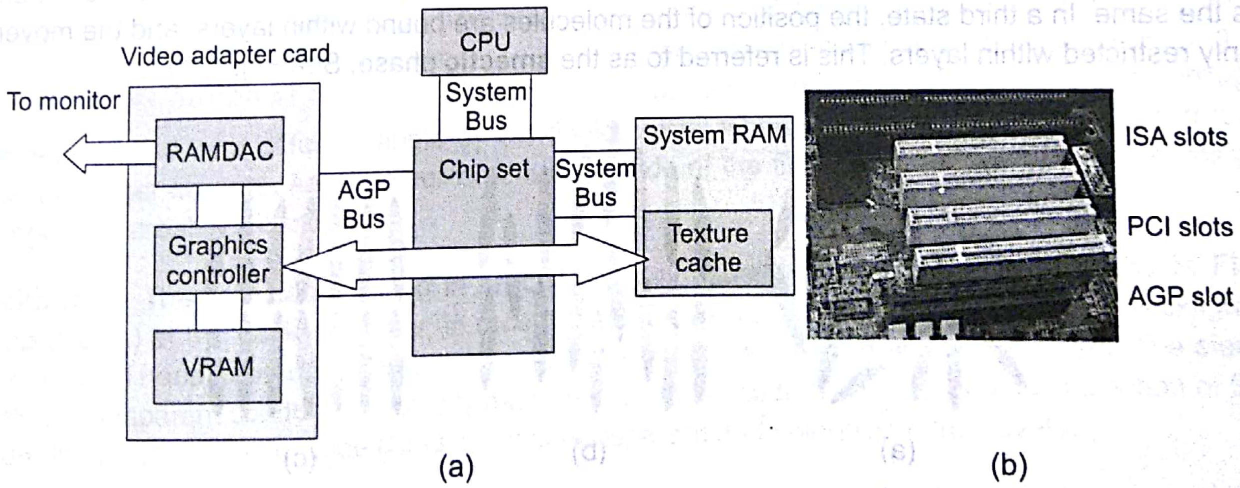


Fig. 1.19 *AGP (a) Data flow (b) Motherboard slot*

AGP is ideal for transferring the huge amount of data required for displaying 3D graphics and animatio[n] has helped remove bandwidth overheads from the PCI bus. AGP is actually considered a port and not a as it involves only two devices, the processor and the video card and are not expandable. AGP specifica are summarized in Table 1.4.

Table 1.4 *AGP specifications*

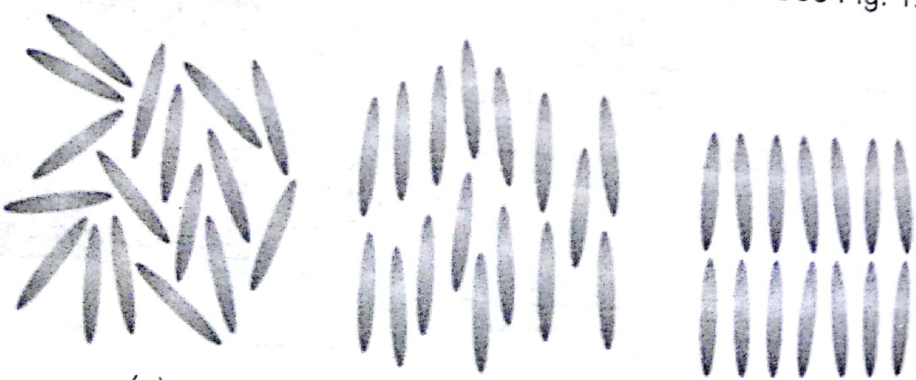
Bus	Speed	MB/second	Frequency (MHz)	Voltage (V)
PCI	-	133	33	3.3
AGP 1.0	1x	266	66	3.3
AGP 1.0	2x	533	66	3.3
AGP 2.0	4x	1066	66	1.5
AGP 3.0	8x	2133	66	0.8

Further Readings

- AGP 3.0 specifications [*http://download.intel.com/support/motherboards/desktop/sb/agp30.pdf*]
- What is AGP [*http://www.sysopt.com/features/mboard/article.php/3549951*]
- AGP pinout [*http://pinouts.ru/Slots/agp_pinout.shtml*]
- Universal AGP [*http://www.microsoft.com/whdc/archive/uagp.mspx*]
- AGP expansion slots [*http://www.motherboards.org/articles/tech-planations/920_4.html*]
- How AGP works [*http://computer.howstuffworks.com/agp.htm*]

1.10.5 Liquid Crystal Display (LCD)

Liquid crystals were first discovered in 1888 by the Austrian botanist Freidrich Reinitzer and the te[rm] **liquid crystal** was coined by the German physicist Otto Lehmann. The liquid crystal is a transparent orga[nic] substance consisting of long rod-like molecules with a cigar-like appearance. It was found that the substan[ce] has properties for manipulating direction of light rays flowing through it. This property is utilized for us[ing] it in display devices. In their natural state, the molecules are oriented arbitrarily, a condition known as t[he] **isotropic** phase. Under certain conditions, it is possible to orient the molecules such that their long axes li[ne] up. This is referred to as the **nematic** phase. The molecules are still able to move around, but their orientatio[n] remains the same. In a third state, the position of the molecules are bound within layers, and the moveme[nt] are mainly restricted within layers. This is referred to as the **smectic** phase. See Fig. 1.20.



(a) (b) (c)

Fig. 1.20 *Liquid-crystal molecule...*

An **LCD monitor** is a collection of LCD elements, each element generating a pixel on the screen, analogous to the phosphor dots on a CRT screen. An LCD element consists of the following major components: polarizing filters, liquid-crystal block, battery, and backlight.

Polarizing Filters Natural light waves are oriented at random angles and flow along various planes from the light source. An optical polarizing filter, or **polarizer,** can isolate a single plane of light from the collection. The filter acts as a net of finely parallel lines blocking all light except those flowing in a plane parallel to the lines. The light in this condition is said to be **polarized**. Depending on the plane of the polarizer, the polarized light can be oriented along the vertical, horizontal or any direction. A second polarizer whose lines are perpendicular to the first would block all the polarized light. See Fig. 1.21.

Liquid-crystal Block The nematic **liquid crystal** phase is the most important for application like LCD displays. By flowing the liquid crystal in a nematic state over a finely grooved surface, it is possible to control the alignment of the molecules as they follow the alignment of the grooves. A layer of liquid crystal material is placed in a container with two finely grooved surfaces whose grooves are perpendicular to each other. Thus, the molecules at the two surfaces are aligned perpendicular to each other and those at the intermediate layers are twisted by intermediate angles. This is referred to as **twisted nematic** state. See Fig. 1.22(a). The container with grooved surfaces is placed in between two perpendicular polarizing filters. Light passing through the first filter would be blocked by the second filter. However, in this case the liquid-crystal material placed in between twists the plane of light by 90 degrees as it passes through the material due to its molecular alignments. The light is now parallel to the second filter and comes out wholly through it to the eye of the observer. This constitutes a lighted pixel on the screen. See Fig. 1.22(b).

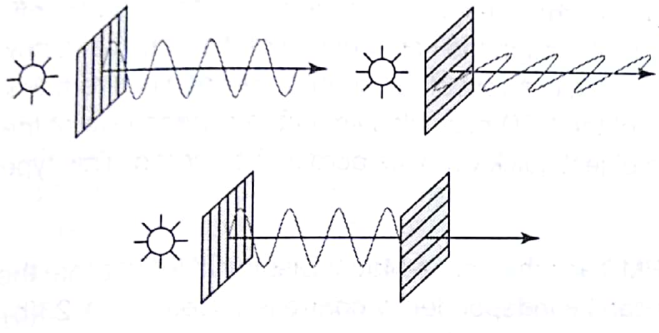


Fig. 1.21 *Polarizing filters*

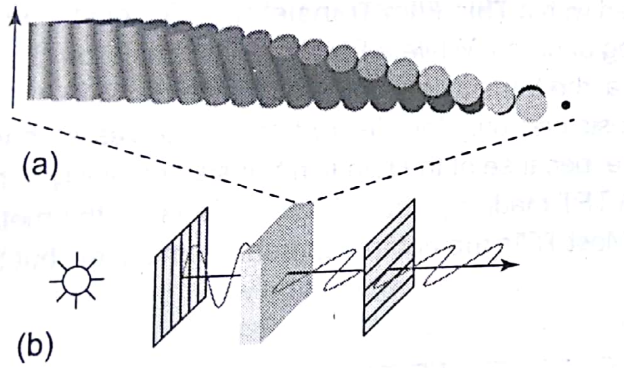


Fig. 1.22 *(a) LC molecules within two grooved surfaces*
(b) Twists light rays by 90 degrees

Battery A battery connected across the liquid-crystal container generates a current through the liquid crystal, and re-orients its molecules according to the direction of the current flow. This disturbs the orderly pattern of the liquid crystal molecules, so that now the molecules at the grooved surfaces are no longer twisted by 90 degrees. Polarized light through the first filter gets twisted by a different angle by the liquid crystal when it reaches the second filter, through which it can no longer emerge. An observer on the other side of the filter does not see any light coming out. This arrangement creates a dark pixel.

Backlight The light source used in an LCD display usually comes from a **Cold Cathode Fluorescent Lamp** (CCFL) at the back; however, in recent times LEDs are also used as a source of backlight. The light sources used need to produce white light, should be cheap, durable, and energy efficient. The electrodes are made of transparent conductor, typically made of indium tin oxide (ITO). ITO is a solid solution of 90% indium oxide (In_2O_3) and 10% tin oxide (SnO_2) and is transparent and colorless in thin layers.

Creating Light and Dark Pixels Images on the LCD screen can be displayed by using a large number of such cells, each functioning as a lighted or dark pixel, depending on whether the corresponding batteries

LCD TV An LCD TV is used for displaying television programs using LCD screens. They have a number of features which make them universal. They have tuners that can display either NTSC, PAL or SECAM signals. Their power systems allow them to be connected to either 110 V or 220 V AC. Other than TV broadcast signals they can also display VGA signals, which allows them to be used as computer monitors. Some LCD TVs also have a wireless port (e.g., Bluetooth) for connecting them to a computer. Companies such as Sharp Corporation, Samsung and Philips have developed large screen sizes for LCD TVs. In October 2004, 40" to 45" televisions were widely available and Sharp Corporation had announced the successful manufacture of a 65" LCD panel. In 2004, Samsung and Sony joined forces to build a factory in South Korea, with a capacity of producing 60,000 LCD panels a month, and in March 2005, Samsung announced an 82" HDTV TFT Panel. Key industry players are seen to be investing a large amount of money in the LCD sector.

Comparisons between CRT and LCD The CRT based devices are quite **bulky** and heavy because of the picture tube, and appropriate for table-top systems but difficult to be carried around. Since the electron gun is held within a vacuum chamber, increasing the screen size required thicker glass to support the atmospheric air force, which in turn increased the weight. The practical limits of CRT sizes are around 30 to 40 inches, beyond which the weight becomes too large. Since the CRT can bend the electron beam only through a specific angle while maintaining focus, the electron gun needs to be located at a distance from the screen which increases the **depth** of the monitors. For example, a 40-inch monitor needs to be around 25 inch in depth. In comparison, the LCD devices use a matrix of LCD cells that are quite small, take approximately a third of the volume of a CRT and weigh about one-fifth of a CRT with comparable screen sizes. LCD screens are therefore appropriate for **portable** equipment like laptops, calculators, palmtops, digital diaries, mobile phones, etc. LCD monitors can also, in theory, be built to any size subject to production complexities and constraints. Screen sizes of 60 to 70 inches are now common. A 40-inch LCD television weights around 25 kg compared to 130 kg for a CRT, and is less than 4 inches thick. Due to the high voltage needed for activating a CRT, **power consumption** in CRT devices are quite high. Adequate heat emission and cooling arrangements are also to be taken into consideration. By contrast, the LCD devices use a small amount of current in each cell and the total power consumption of LCD devices are just a fraction of that in CRT devices. The LCD cells are minute structures, which are difficult to fabricate. A large percentage of the manufactured cells have some defects in them and are discarded. This makes the LCD fabrication process **expensive** and results in the LCD screens being more costly than their CRT counterparts. The longevity of the LCD screens is also thought to be less than the CRT screens. In a CRT screen, the light produced is emitted directly by the phosphor material and, therefore, appreciable amount of **brightness** is observed. In contrast LCD screens do not emit light, they can only block light flowing through them. The light comes from a source of white light within the LCD device. Hence, LCD screens usually lack the brightness and contrast, of CRT screens. In CRT screens all the light emitted from the phosphor dots are visible by the observer, and there are usually no viewing problems. In LCD screens however, only single planes of light are isolated as part of the polarization process, which reduces the **viewing angle** to the observer. It is therefore often seen that the images on a LCD screen are only visible when looked straight on, but become hazy or blurred when viewed at an angle to the perpendicular line.

Further Readings

- Liquid Crystal Display @ Wikipedia [*http://en.wikipedia.org/wiki/Lcd*]
- History of LCD [*http://ieee.org/portal/cms_docs_iportals/iportals/aboutus/history_ center/LCD-History.pdf*]
- Polarization by Wikipedia [*http://en.wikipedia.org/wiki/Polarization*]
- How LCDs work [*http://electronics.howstuffworks.com/lcd.htm*]
- Liquid Crystal Institute (Kent State University) [*http://www.lci.kent.edu/*]
- How LCDs work [*http://solutions.3m.com/wps/portal/3M/en_US/Vikuiti1/BrandProducts/ secondary/optics101/*]

- Liquid crystals [http://nobelprize.org/educational_games/physics/liquid_crystals/history/]
- Making of LCD panels [http://www.auo.com/auoDEV/content/technology/technology_tftprocess_popup_en.htm]
- Links to liquid crystals and LCDs [http://courseware.ee.calpoly.edu/~dbraun/courses/lcd.html]
- Liquid Crystal Technology Group homepage [http://www.eng.ox.ac.uk/lc/index.html]
  LCD Monitor Parameters and Characteristics [http://www.xbitlabs.com/articles/other/display/lcd-guide.html]
- How TFTs work @ TFT Central [http://www.tftcentral.co.uk/working.htm]
- TFT LCD display [http://www.maxi-pedia.com/TFT+LCD+display+monitor+panel]
- LCD TV FAQ [http://www.lcd-tv-reviews.com/]

1.10.6 Plasma Display Panel (PDP)

In physics and chemistry, **plasma** (also called an ionized gas) is an energetic gas-phase state of matter, often referred to as "the fourth state of matter", in which some or all of the electrons in the outer atomic orbital have separated from the atom. The result is a collection of ions and electrons that are no longer bound to each other. Because these particles are ionized (charged), the gas behaves in a different fashion than neutral gas, e.g., it can be affected by electromagnetic fields. This state of matter was first identified by Sir William Crookes in 1879, and dubbed "plasma" by Irving Langmuir in 1928. Commonly encountered forms of plasma include the sun and other stars (which are plasmas heated by nuclear fusion), lit fluorescent lamps, lightning, the Aurora borealis, the solar wind, and interstellar nebulae. A plasma is also generated in front of a spacecraft's heat shield on reentering the atmosphere. There is still some debate as to whether plasma is an individual state of matter or simply a type of gas, but most physicists have accepted plasma as a state of matter.

Other than CRT and LCD, a third type of display screens often used is known as the **Plasma Display Panel** (PDP). A PDP is an emissive (i.e., discharge of electromagnetic radiation or particles) flat panel display where light is created by phosphors excited by a plasma discharge between two flat panels of glass. The gases used are usually a mixture of inert gases like neon and xenon. The PDP was invented at the University of Illinois by Donald L. Bitzer and H. Gene Slottow in 1964. One advantage of the PDP is that the light produced is extremely bright and thus they are mostly used for glow signs. A disadvantage is that they consume a large amount of power to ionize the gas. A PDP usually produces an excess of 1000 lux (a **lux** is an unit of illumination equal to one lumen per square metre) of illumination. By comparison, sunlight on an average day ranges from 32 kilo-lux to 100 kilo-lux, TV studios are lit at 1000 lux, a bright office has 400 lux, moonlight represents 1 lux, starlight measures 50 μlux. Moreover, PDPs can have large sizes, about 80" diagonally.

Plasma displays, like cathode ray tube (CRT) televisions, are phosphor-based, meaning they rely on glowing phosphor to create their images. A plasma panel starts with two sheets of 1/16" glass. Between the sheets of glass are spacers that create individual cells within the two glass sheets. Transparent electrodes are then coated inside the front and back pieces of glass, so that there are two electrodes per cell on the front piece of glass and one electrode per cell on the rear piece of glass. Next, each cell is coated with phosphor and filled with a mixture of neon and xenon gases. When an electrical charge is added to the front and back electrodes, a voltage differential is created, adding energy to the gas mixture and changing the gas to a plasma state. Once the gas changes into a plasma state, it releases ultraviolet energy which excites the phosphor coated inside each cell, causing it to emit visible light. For color reproduction, the ions in the plasma are used to excite three phosphor sub-components (RGB) to produce a desired color. Due to the bright illumination, PDPs have a large contrast ratio (ratio between the brightest and darkest pixels) which indicates more picture detail. Ratios are advertised having a value around 1000:1; by comparison an ordinary CRT monitor has a value around 50:1.

Advantages of PDP include the following: (a) A PDP like an LCD is very thin and can be mounted in places where traditional CRTs cannot be used. (b) Color reproduction and contrast are good. (c) Large screen sizes of 70" are possible. (d) PDP HDTV resolution displays are much less expensive than LCD displays of the same size. Main disadvantages are the following: (a) The PDP is quite fragile and may create shipping problems. (b) Detail in dark scenes are not as good as CRTs. (c) They are usually much heavier than LCD displays and cannot be used in portable equipment.

Further Readings

- Plasma the fourth state of matter [*http://fusedweb.pppl.gov/CPEP/Chart_Pages/5.Plasma4StateMatter.html*]
- Plasma science and technology [*http://www.plasmas.org/*]
- Introduction to Plasma physics [*http://farside.ph.utexas.edu/teaching/plasma/lectures/lectures.html*]
- Plasma display @ Wikipedia [*http://en.wikipedia.org/wiki/Plasma_display*]
- LCD TV vs. Plasma TV [*http://www.crutchfield.com/S-K09Efe8ktJz/learn/learningcenter/home/tv_flatpanel.html*]
- How plasma displays work [*http://www.howstuffworks.com/plasma-display.htm*]
- Plasma TV technology [*http://www.plasmatvscience.org/technology.html*]
- Plasma vs. LCD @ CNET Australia [*http://www.cnet.com.au/plasma-vs-lcd-which-is-right-for-you-240036500.htm*]
- Comparison of LCD & Plasma TV @ NYTimes [*http://www.nytimes.com/2006/12/25/technology/25flat.html?_r=1*]

1.10.7 Other Flat Panel Displays

Light Emitting Diodes (LED) These displays are typically used in outdoor glow signs and billboards, and in recent times in display signs on public transports. Most of these are based on a collection of R, G, B discrete diodes, spaced evenly apart, which act as individual pixel elements. The largest LED display in the world is 1500 feet long and located in Las Vegas, Nevada, USA. Most indoor LED display screens are based on Surface Mounted Device (SMD) panels, which consist of R, G, B panels mounted on a chipset board. The main difference with the outdoor type is the difference in brightness—while the outdoor LED displays have a brightness of around 2000 candela/m^2, the indoor type is limited to around 600 candela/m^2. An alternate way to use LEDs in display panels is an **LED-backlight** LCD panel. Instead of the CCFL lamps used in traditional LCDs to provide backlighting, LED lamps are used for the same purpose. Such uses of LED have been reported to produce a thinner panel and less power consumption. LED backlighting can be of two types: (1) Dynamic RGB LEDs placed behind the panel which result in a wider color gamut and better brightness-contrast, and (2) Edge LEDs placed at the rim of the screen with a diffusion mechanism to spread the light evenly behind the screen and which can result in a thinner display and lower power consumption.

Electroluminescent Displays (ELD) These are created by sandwiching a layer of electroluminescent material such as gallium arsenide between two layers of conductors. When current flows, the layer of material emits radiation in the form of visible light. The electric current excites the atoms of the EL material causing them to emit photons.

Surface-conduction Electron-emitted Display (SED) These use a technology similar to CRT but on a microscopic scale by using a grid of electron emitters for each sub-pixel of the display. Each emitter produces electrons which are caused to strike a colored phosphor dot to produce light. SEDs combine the high contrast ratios, wide viewing angles and fast response times of CRTs with the compact size and low power consumptions of LCDs. Canon's 55-inch prototype SED offered brightness of around 450 cd/m^2 a contrast ratio of 50000:1, and a response time of 1 ms. By contrast, LCD TVs offer a contrast ratio of 3000:1 and a response time of around 4 ms. SED is closely related to another display technology, the Field Emission Display (FED). The two differ in the details of electron emitters.

Further Readings

- EL displays [http://www.indiana.edu/~hightech/fpd/papers/ELDs.html]
- SED technologies [http://www.eetasia.com/ARTICLES/2007AUG/PDF/EEOL_2007AUG16_OPT_TA.pdf]
- Comparison between SED and FED [http://www.appliednanotech.net/tech/pdfs/fed_sed.pdf]
- FED TV reviews [http://www.fed-tv-reviews.com/]
- Field emission technologies [http://www.fe-tech.co.jp/en/index.html]
- Comparison of display techologies @ Wikipedia [http://en.wikipedia.org/wiki/Comparison_of_display_technology]
- Large screen television technology @ Wikipedia [http://en.wikipedia.org/wiki/Large-screen_television_technology]

1.10.8 Multiple Monitors

Multiple monitors imply the use of more than one physical monitor to a single computer in order to increase the available visual display area for computer programs. There are two popular modes in which such monitors might be connected: the clone mode and the span mode. The **clone mode** duplicates the same video frame information to each of the monitors usually using an external video signal splitter. The multiple monitors, therefore, display the same video image by sharing the same frame buffer, usually to cater to multiple viewers physically separated by some distance, but also sometimes to display the same information on a low-resolution color monitor and a high-resolution monochrome monitor. Different types of monitors, i.e., CRT and LCD, can be connected together, e.g., connecting a laptop to an external CRT monitor. The **span mode** splits the same video buffer across two or more monitors, thereby effectively increasing the total information that can be displayed, e.g., as the mouse is moved to the right side of the left screen, it will appear on the left side of the right screen for a dual-monitor setup. Thus, using two monitors, each with pixel addressability m by n, can generate a virtual addressability to $2m \times n$ or $m \times 2n$. The monitors use different frame buffer information provided this is supported by the operating system, as well as the graphics hardware. The video adapter cards in this case need to have two or more outputs ports usually VGA and/or DVI. Alternatively, two separate graphic cards each with one monitor can also be used. Spanning mode can be combined with multi-tasking applications to generate two different views on two different monitors, e.g., viewing a 3D graphical objects from two different angles or a normal view combined with a zoomed-in view, or even two different applications like Web browsing and e-mail on separate monitors.

Multiple monitors are frequently used in applications like graphic design, video editing, computer gaming, communications as well as banks, business accounting and share markets. Multi-monitor graphic cards are available from hardware vendors like NVIDIA, Matrox and ATI Technologies. From the programming viewpoint, multi-monitor rendering uses OpenGL or DirectX routines to create a large virtual frame buffer in which the graphics driver and OS writes data by demarcating areas for individual monitors.

Although multiple monitors helps to increase the viewing space and consequently the amount of information that can be projected at once, the main disadvantage is that the resources of the video adapter card including the VRAM, is divided among all the monitors which may lead to degraded performances. In recent times, this problem has been addressed by graphic cards supporting 1 GB or more of VRAM and powerful graphic processors to handle data transfers. Another disadvantage is the additional cost required for supporting the additional monitors often with added costs of additional adapters. Users might find a single wide screen LCD display more cost effective than multiple monitors. A third problem relates to the use of edge detection of monitor screens for scrolling and gaming purposes which may not work properly when multiple monitors are installed.

Further Readings

- Multi-monitor computing [http://9xmedia.com/PDFs/9X_Media_multiple_monitor_whitepaper.pdf]
- Multi-monitor rendering in OpenGL [http://www.rchoetzlein.com/theory/?p=103]
- Setting up multiple monitors [http://www.pcworld.com/article/157672/how_to_set_up_multiple_monitors.html]

1.10.9 3D Monitors

A 3D monitor is a display device capable of imparting a perception of depth to the viewer. The basic principle is to display two images of the same object or scene, slightly offset from each other, in such a way that the left eye sees only one of the images, while the right eye sees only the other image. These two offset images are then combined in the brain to provide the perception of three-dimensional depth. The term **3D image** is actually considered to be a misnomer, since essentially it is displaying dual 2D images, and the accurate term to describe this is **stereoscopic** or **stereo-vision**. A real 3D image should enable the viewer to view it from any direction, as is the case in a 3D modeling software like 3D Studio MAX, while in a 3D monitor we are obviously unable to view from behind the object or scene. The basic method of viewing stereo images involves two cameras capturing an image of an object from two slightly different angles such that the offset produced is similar to the one that we see with our left eye and right eye. These two images are then projected simultaneously on a screen, which is viewed by the viewer wearing polarized glasses, as in a 3D movie. The glasses ensure that image from one camera goes only to the left eye and blocked to the right eye, while image from the other camera goes only to the right eye and blocked to the left eye. These two images are then combined by the brain to give a depth perception. A 3D monitor, however, attempts to make the viewer see 3D images without glasses, described by a term called **auto-stereoscopic**. The simplest way is to use an interlaced arrangement where one field shows image from one camera and the other field shows image from the other camera. This principle relies on persistence of vision to combine the images together and provide a depth perception. Modern flat-panel displays providing auto-stereoscopy work on the principle of **parallax barrier,** whereby a series of precision slits are placed in front of an LCD display panel allowing the eye to see a different set of pixels and creating a sense of depth through parallax. **Parallex** is the apparent shift in the position of an object when viewed along different lines of sight, as it gets superimposed on different sets of distant objects. An alternative arrangement is to replace the parallex barrier by an array of tiny cylindrical lenses called **lenticular lenses,** which focus light from different sets of pixels to the left eye and right eye.

Further Readings

- N. Holliman, "3D display systems" [http://www.dur.ac.uk/n.s.holliman/Presentations/3dv3-0.pdf]

Conclusions

- The term "multimedia" essentially implies "multiple means" of conveying and perceiving information.
- A multimedia production involves a subject matter whose content is represented through various audiovisual media.
- Such content is created by a team of developers keeping in mind a specific target audience.

- LCD monitors are much less bulky and use low power consumption compared to the CRTs, making them ideal for mobile device screens.
- Other flat-panel monitors use ionized gases within sealed panels (plasma display panels) and Light Emitting Diodes (LED) to form images on screens.
- Multiple monitors can either be connected in clone mode for duplicating the same video information in each, or in span mode for splitting the information across them.

Solved Examples

Example 1.1 **A 15-inch monitor with an aspect ratio of 4:3 has a pixel addressability of 800 × 600. Calculate its resolution and dot-pitch.**

Let width W and height H of the monitor be $4x$ and $3x$ respectively. Size (diagonal) is $D = 15$

From equation (1.7),

$$15^2 = (4x)^2 + (3x)^2 \rightarrow x = 3$$

Hence $W = 4x = 12$ inch

From equation (1.8), resolution $R = w$ (pixels) / W (inch) = 800/12 = **66.67 dpi**

From equation (1.10), dot-pitch $d = 1/R = 1/66.67 = 0.015$ inch = (0.015×25.4) mm = **0.381 mm**

Example 1.2 **A monitor can display 4 shades of red, 8 shades of blue and 16 shades of green. Find out the color depth supported by the monitor.**

Each pixel is capable of displaying $4 \times 8 \times 16 = 512$ colors.

Let b be the bit-depth or color-depth supported by the monitor.

From equation (1.5), $512 = 2^b \rightarrow b =$ **9 bits**

Example 1.3 **A monitor has a pixel addressability of 800 × 600 and a color-depth of 24 bits. Calculate the minimum amount of display memory required on its adapter card to display an image on the screen.**

From equation (1.6),

Memory required $M = w.h.b = 800 \times 600 \times 24$ bits $= (800 \times 600 \times 24)/(1024 \times 1024 \times 8)$ MB = **2 MB**

Example 1.4 **A monitor of aspect ratio 4:3 with 900 dots along a vertical column, needs to display images with 32768 colors. Calculate the minimum amount of adapter display memory required in MB. Also calculate the monitor resolution in dpi for a 20-inch monitor.**

Height of monitor in pixels $h = 900$. Aspect ratio $m : n = 4:3$

Let width in pixels be w

From equation (1.9), $w = (m/n).h = (4/3)(900) = 1200$ pixels

Total number of colors $C = 32768$. Let bit-depth be b.

From equation (1.5), $32768 = 2^b \rightarrow b = 15$

From equation (1.6), memory required to display an image on screen $M = 1200 \times 900 \times 15$ bits = **2 MB**

Let the width and height of the monitor in inches be $W = 4x$ and $H = 3x$ respectively. Size $D = 20$ inch.

From equation (1.7), $20^2 = (4x)^2 + (3x)^2 \rightarrow x = 4$

So width $W = 4x = 16$ inches

From equation (1.8), resolution $R = w$ (pixels) / W (inch) = 1200/16 = **75 dpi**

Example 1.5 **A monitor can display R shades of red, G shades of green and B shades of blue colors. Show that the color depth supported by the monitor is ($\log_2 R + \log_2 G + \log_2 B$).**

Let number of bits allocated to red be r, to green be g and to blue be b.

From equation (1.5), $R = 2^r$, $G = 2^g$, $B = 2^b$

Hence $r = \log_2 R, \quad g = \log_2 G, \quad b = \log_2 B$

So total number of bits allocated to each pixel = bit-depth of monitor = $r + g + b = \log_2 R + \log_2 G + \log_2 B$

Example 1.6 **An analog signal containing components with frequency values ranging from 50 Hz to 5 kHz, is to be sampled. Determine the sampling frequency and the bandwidth of the band-limiting filter.**

From equation (1.1), sampling rate $F = 2 \times 5$ kHz = **10 kHz**

The bandwidth of the band-limiting filter is : **0 Hz to 5 kHz**

Example 1.7 **Repeat previous problem if the signal is transmitted over a communication channel with a bandwidth from 100 Hz to 4.5 kHz.**

In this case, the bandwidth of the transmission channel is smaller than that of the source signal.

From equation (1.1), sampling rate $F = 2 \times 4.5$ kHz = **9 kHz**

The bandwidth of the band-limiting filter is : **0 Hz to 4.5 kHz.**

Text

2.1 Introduction

Ever since the inception of human civilization, ideas have been largely articulated using the written mode. The flexibility and ease of use of the textual medium makes it ideal for learning. Word-processing programs emerged as one the earliest application programs. In multimedia presentations, text can be combined with other media in a powerful way to present information and express moods. Text can be of various types: plain text, consisting of fixed-size characters having essentially the same type of appearance; formatted text, where appearance can be changed using font parameters; and hypertext, which can serve to link different electronic documents and enable the user to jump from one to the other in a non-linear way. Internally, text is represented via binary codes as per the ASCII table. The ASCII table is however quite limited in its scope and a new standard has been developed to eventually replace the ASCII standard. This standard is called the **Unicode** standard and is capable of representing international characters from various languages throughout the world. Text can be inserted into an application using various means. The simplest way is directly typing text into the application by using the keyboard; alternatively, text can be copied from another pre-existing file or application and pasted into the application. Nowadays, we also generate text automatically from a scanned version of a paper document or image using OCR software. When text is saved onto the hard disk, it can be compressed using various algorithms so as to reduce the file size. All algorithms however work in a lossless mode, i.e., all information in the original file is maintained intact without loss of data in any way. Depending on how various visual properties of text are stored and the compression scheme followed, text can be stored into a number of file formats, each requiring its own specific application to open and modify the contents.

2.2 Types of Text

Essentially, there are three types of text that can be used to produce pages of a document : unformatted text, formatted text, and hypertext.

2.2.1 Unformatted Text

Also known as plaintext, this comprise of fixed-size characters from a limited character set. The character set is called **ASCII table** which is short for American Standard Code for Information Interchange and is one of the most widely used character sets. It basically consists of a table where each character is represented by a unique 7-bit binary code. This means there are 2^7 or 128 code words which can be used to identify the characters. The characters include *a* to *z*, *A* to *Z*, 0 to 9, and other punctuation characters like parenthesis, ampersand, single and double quotes, mathematical operators, etc. All the characters are of the same height. In addition to normal alphabetic, numeric, and punctuation characters, collectively called **printable characters**, the ASCII character set also includes a number of **control** characters. These include *BS* (backspace), *LF* (linefeed), *CR* (carriage return), *SP* (space), *DEL* (delete), *ESC* (escape), *FF* (form feed) and others. See Fig. 2.1. Whenever characters included in the table are required to be stored or processed by a computer, the corresponding numeric code is retrieved from the ASCII table in binary form and substituted for the actual text for internal processing. This includes both the printable and control characters, e.g., each line of text in a document is terminated by a Line-Feed character.

Later on as requirements increased an extended version of ASCII table was introduced known as the **extended ASCII** character set, while the original table came to be known as **standard ASCII** set. The extended set used an 8-bit representation and, therefore, had a provision of 256 characters. The first 128 characters were the same as the original character set, now known as the Standard ASCII table while the remaining codes were used to represent small simple graphical symbols. See Fig. 2.2. The symbols are all of the same width and height. Applications can use these symbols to create more complex graphical images and larger sizes of text.

2.2.2 Formatted Text

Formatted text are those where apart from the actual alphanumeric characters, other **control characters** are used to change the appearance of the characters, e.g., bold, underline, italics, varying shapes, sizes and colors, etc. Most text-processing software use such formatting options to change text appearance. It is also extensively used in the publishing sector for the preparation of papers, books, magazines, journals, and so on. In addition, a variety of document-formatting options are supported to enable an author to structure a document into chapters, sections and paragraphs, and with tables and graphics inserted at appropriate points. The control characters used to format the text is application dependent and may vary from one package to another, e.g., bold appearance in an MS-Word document may be produced by a different control code than that in an HTML document. To print such a document, the printer should also be capable of interpreting these control codes so that the appropriate appearance may be reproduced.

Further Readings

- ASCII @ Wikipedia [*http://en.wikipedia.org/wiki/Ascii*]
- Annotated history of ASCII [*http://wps.com/projects/codes/index.html*]
- Evolution of character codes [*http://www.transbay.net/~enf/ascii/ascii.pdf*]
- Standard code for ASCII [*http://wps.com/projects/codes/X3.4-1963/page5.JPG*]
- ASCII table [*http://www.asciitable.com/*]
- ASCII chart and other resources [*http://www.jimprice.com/jim-asc.shtml*]
- Plain text @ Wikipedia [*http://en.wikipedia.org/wiki/Plain_text*]
- Formatted text @ Wikipedia [*http://en.wikipedia.org/wiki/Formatted_text*]

2.2.3 Hypertext

Documents provide a method of transmitting information. Reading a document is an act of reconstructing knowledge. A book or an article on paper has a given structure and is represented in a sequential form. Although it is possible to read individual paragraphs without reading previous paragraphs, authors mostly assume sequential reading. Novels, as well as movies, always assume a pure sequential reception. Technical documentation (e.g., manuals) consists often of a collection of relatively independent information units. There also exits many cross-references in such documentation which lead to multiple searches at different places for the reader. A hypertext document can be used to handle such situations. The term **hyper** is usually used to mean something excessive (beyond normal), e.g., hyperactive, hypertension, etc. Here the term is used to mean certain extra capabilities imparted to normal or standard text. Like normal text, a hypertext document can be used to reconstruct knowledge through sequential reading but additionally, it can be used to link multiple documents in such a way that the user can navigate non-sequentially from one document to the other for cross-references. These links are called hyperlinks. **Hyperlinks** form one of the core structures in multimedia presentations, because multimedia also emphasizes a non-linear mode of presentation. For example, a multimedia tourist brochure can have text information about various places of interest, with photographic images, can have voice annotations about how to get to those places and the modes of travel available, video clips of how tourists are traveling and the facilities provided to them. All these information can be hyperlinked to each other, along with other relevant information like a list of hotels at each place along with their charges. A tourist can make use of a search facility to navigate to the information regarding a specific place of interest and then use the hyperlinks provided to view each category of information. Hypertext is mostly used on the World Wide Web for linking different Web pages together and allowing the user to navigate from one page to another. To create such documents, the user uses commands of a hypertext language like HTML or SGML to specify the links. Typically, hyperlinks take the form of an underlined text string and the user initiates the access and display of a particular document by pointing and clicking the mouse on the appropriate link. The underlined text string on which the user clicks the mouse is called an **anchor** and the document which opens as a result of clicking is called the **target** document. On the Web, target documents are specified by a specific nomenclature called web site address technically known as **uniform resource locators** or URL. An example of hypertext is shown below. The text string "Microsoft Home Page" forms the anchor which when mouse clicked opens the Microsoft Web site defined by the URL "www.microsoft.com". Whenever the mouse pointer is placed over a hypertext, the cursor changes to a hand icon.

<u>Microsoft Home Page</u>

Other than the Internet, hypertext can also be used in other applications like MS-Word, MS-PowerPoint Adobe Acrobat and multimedia presentations. Here, hypertext is used to jump to one page of the document to another page, or even from one point to another within the same page. Using programming languages like JavaScript or Lingo, hypertexts can be used to initiate playback of media files like a video clip. A hypertext consists of two different parts : the anchor and the link.

The **anchor** is the actual visual element (text) which not only is associated with some meaning by itself but also provides an entry point to another document. An important factor in designing a user interface is the concept of how the anchor can be represented properly. In most cases, the appearance of the text is changed from the surrounding text to designate a hypertext, e.g., by default it is colored blue with an underline. Moreover, the mouse pointer changes to a finger icon when placed over a hypertext. The user usually clicks over the hypertext in order to activate it and open a new document in the document viewer. In some cases instead of text an anchor can be an image, a video or some other non-textual element. In such cases the term **hypermedia** is more appropriate.

Pointers provide connection to other information units known as target documents. A link has to be defined at the time of creating the hyperlink, so that when the user clicks on an anchor, the appropriate target document can be fetched and displayed. Usually, some information about the target document should be available to the user before clicking on the anchor. If the destination is a text document, a short description of the content can be represented. In the case of an image, the image content can appear in thumbnail form on the screen. A visual representation of the video content can follow in the form of a moving icon. If the content of the destination node consists of audio information a visual representation of the audio content must follow. For example, in the case of a music passage a picture of the composer could be displayed.

The architecture of a hypertext system can be divided into three layers : Presentation layer, Hypertext Abstraction layer, and Storage layer.

At the uppermost layer, the **presentation layer**, all functions connected to the user interface are embedded. Here nodes and pointers are mapped to the user interface. This layer determines which data are presented and how they are presented. This layer takes over control of all other inputs. For example, an HTML document displayed in a Web browser shows text and images in various formats and locations on the screen and able to handle events like the user clicking the mouse on a hyperlink and take appropriate action, e.g., open the target document.

The **hypertext abstraction layer,** also called the Hypertext Abstract Machine (HAM), is placed between the presentation and storage layers. HAM knows the structure of the document and has the knowledge about the pointers and its attributes. This layer has the least system dependency compared to the other two layers and hence the most suitable for standardization. For example, the HTML code itself will be able to generate the display of the target document of a hyperlink by interpreting the element.

The **storage layer**, also called the database layer, is the lowest layer. All functions connected with the storage of data belong to this layer. The specific properties of the different discrete and continuous media need to be considered. Functionalities from traditional database systems are expected like transactions, locks, etc. For example, the storage of an HTML file along with graphic and audio files can be used to generate a display of a page containing the graphic and background music.

Further Readings

- Hypertext @ Wikipedia [*http://en.wikipedia.org/wiki/Hypertext*]
- Historical overview of hypertext [*http://www.mprove.de/diplom/text/2_hypertext.html*]
- Hypertext – behind the hype [*http://www.ericdigests.org/pre-9212/hype.htm*]
- History of hypertext [*http://www.omninerd.com/articles/The_History_of_Hypertext*]
- Hyperwords [*http://www.hyperwords.net/*]
- Glossary of hypertext terms @ W3C [*http://www.w3.org/2003/glossary/subglossary/hypertext-terms.rdf/*]
- HTML @ Wikipedia [*http://en.wikipedia.org/wiki/HTML*]
- HTML 4.01 specifications @ W3C [*http://www.w3.org/TR/html401/*]
- HTML 5 specifications @ W3C [*http://dev.w3.org/html5/spec/*]
- Getting started with HTML [*http://www.w3.org/MarkUp/Guide/*]
- HTML tutorial [*http://w3schools.com/html/default.asp*]

2.3 Unicode Standard • •

The **Unicode standard** is a new universal character coding scheme for written characters and text. It defines a consistent way of encoding multilingual text which enables textual data to be exchanged universally. The Unicode standard goes far beyond ASCII's limited capability by providing the capacity of encoding more than 1 million characters. Multilingual support is provided for European, Middle Eastern, and Asian languages. The Unicode Consortium was incorporated in 1991 to promote the Unicode standard. The Unicode Technical Committee (UTC) is the working group within the Consortium responsible for the creation, maintenance, and quality of the Unicode Standard. The Unicode standard draws a distinction between **characters**, which are the smallest component of written language, and **glyphs**, which represent the shapes, the characters can have when displayed. The Unicode standard deals only with character codes, representation of the glyphs is not part of the standard. These are to be dealt with by the font vendors and hardware vendors. A font and its rendering process define a mapping from the Unicode values to glyphs. For example, the Hindi character "Pa" is represented by the Unicode sequence 0000 1001 0010 1010 (U+092A), how it will be rendered on the screen will be decided by the font vendor. The first byte represents the language area while the next byte represents the actual character. Some of the languages and their corresponding codes are Latin (00), Greek (03), Arabic (06), Devanagari/Bengali (09), Gurumukhi/Gujrati (0A), Oriya/Tamil (0B), Tibetan (0F), Mongolian (18), etc.

The **Unicode Consortium** based in California overlooks the development of the standard. Members include major software and hardware vendors like Apple, Adobe, IBM, Microsoft, HP, etc. The Consortium first published the Unicode Standard 1.0 in 1991. The latest version is 4.1 released in 2005. The Unicode version 3.0 is identical to the ISO standard 10646. Several methods have been suggested to implement Unicode based on variations in storage space and compatibility. The mapping methods are called **Unicode Transformation Formats** (UTF) and **Universal Character Set** (UCS). Some of the major mapping methods are the following:

Mapping methods

UCS-4, UTF-32 The UTF-32 scheme uses 32 bits for each character. The simplest scheme, as it consists of fixed length encoding, is not efficient with regard to storage space and memory usage, and therefore rarely used. The 4 here means 4 bytes, and 32 means 32 bits. Initially the UCS-4 was proposed with a possible address range of 0 to FFFFFFFF(H), but Unicode requires only upto 10FFFF(H). Hence UTF-32 was proposed as a subset of UCS-4 that uses 32-bit code within the range 0 to 10FFFF(H) only.

UTF-16 The UTF-16 scheme uses 16 bits for each character. In its native format it can encode numbers up to FFFF(H), i.e., as *xxxxxxxx xxxxxxxx*. For coding beyond this, the original number is expressed as a combination of two 16-bit numbers. Let us consider the code 1D11E(H) which represents the musical G clef note. Since this is larger than FFFF(H), it cannot be represented using 16 bits. The following technique is used:

(a) an offset of 10000(H) is subtracted from it, i.e., 1D11E(H) – 10000(H) = D11E(H).

(b) This number is then expressed in binary as combination of two 10-bit binary numbers, i.e., 0000110100 0100011110. Let this be represented as xxxxxxxxxx xxxxxxxxxx.

(c) The actual Unicode is derived by prefixing it by 36(H) and 37(H), i.e., 110110xxxxxxxxxx 110111xxxxxxxxxx. In this case, the numbers become : 1101100000110100 1101110100011110, i.e., D834 DD1E.

UTF-8 In the UTF-8 scheme the bits of a Unicode character are divided into a series of 8-bit numbers. The output codes against various ranges of input codes are given below :

Table 2.1 *UTF-8 coding scheme*

Code range	Input code	Output code
000000–00007F	xxxxxxx	0xxxxxxx
000080–0007FF	xxx xxxxxxxx	110xxxxx 10xxxxxx
000800–00FFFF	xxxxxxxx xxxxxxxx	1110xxxx 10xxxxxx 10xxxxxx
010000–10FFFF	xxx xxxxxx xxxxxx xxxxxx	11110xx 10xxxxxx 10xxxxxx 10xxxxxx

Further Readings

- The Unicode Consortium [*http://www.unicode.org/*]
- The Unicode standard—a technical introduction [*http://www.unicode.org/standard/principles.html*]
- Unicode FAQ [*http://www.unicode.org/faq/basic_q.html*]

- Unicode 5.2.0 [http://www.unicode.org/versions/Unicode5.2.0/]
- Table of Unicode characters [http://unicode.coeurlumiere.com/]
- Unicode character search [http://www.fileformat.info/info/unicode/char/search.htm]

2.4 Font

The appearance of each character in case of formatted text is determined by specifying a **font** name. Font names refer to font files which contain the actual description of the character appearance. On the Windows platform, font files are stored in a specific folder called Fonts under the Windows folder. These files are usually in **vector** format meaning that character descriptions are stored mathematically. This is useful because characters may need to be scaled to various heights and mathematical descriptions can easily handle such variations without degrading the appearance of characters. Windows calls these fonts as **true type fonts** because their appearance stays the same on different devices like monitors, printers, and plotters, and they have a file extension of TTF. An alternative form of font files is the **bitmap** format where each character is described as a collection of pixels. The disadvantage of this is that it can lead to distortion when the characters are scaled to various sizes. Some of the standard font types included with the Windows OS package are shown below. See Fig. 2.3, upper row. Other than these there are thousands of font types made by various third-party organizations and many of them are freely downloadable over the Internet. Each application has a default font name associated with it. When a specific font requested by the user is not available in the system then the default font is used which is assumed to be always available. Some software application packages (e.g., Adobe Acrobat) allow font files to be embedded within the document file (e.g., PDF) so that on a target system that does not have the requisite font files, the embedded fonts are used instead. Some examples of third-party downloaded fonts are shown below. See Fig. 2.3, lower row.

Times Roman Arial Garamond Verdana Courier New

Aero Bauhaus BREMEN Harlow Papyrus

Fig. 2.3 *Examples of Windows default fonts (upper row) and third-party downloaded fonts (lower row)*

Font Size and Style Font characters may have a number of sizes. Size is usually specified in a unit called **point** (pt) where 1 point equals 1/72 of an inch. Sometimes the size may also be specified in pixels. Standard characters in textual documents usually range from 10 to 12 points. in size, while the upper limit may go well beyond 100. Specific font types can be displayed in a variety of styles. Some of the common styles used are **bold**, *italics*, underline, super^script and sub_script. Some application packages allow changing the horizontal gap between the characters, called **kerning**, and the vertical gap between two lines of text, called **leading**. Some packages allow a number of special effects on text to make it more dramatic, interesting and fun. This includes changing the appearance in a variety of ways like bending, slanting, warping, rotating, adding shadows and 3D effects, etc. Other ways of manipulating text include animations like scrolling, fading, changing colors, adding sound effects, etc.

TrueType Font TrueType font is an outline font standard developed by Apple Computer in the late 1980s. Outline fonts or vector fonts are collection of vector images, i.e., mathematically defined straight lines and Bezier curves, to define the border of glyphs. A glyph represents the actual visual representation of a character. A TrueType font file contains vector instructions to create specific glyphs and has extension TTF. A true type system has a rendering machine which reads in the vector instructions and produces an appropriately scaled rasterized output to be displayed on screen or printed on printers. This enables the system to adjust the resolution of the output glyphs as per requirements and maintain their smooth appearances irrespective of their sizes, unlike bitmap fonts which degrade in appearance as they are scaled.

PostScript Font PostScript fonts are outline font specifications developed by Adobe Systems which uses the PostScript (PS) language to encode font information. These files usually have the extension PFM, and are usually of the following types:

Type 1 Single-byte digital fonts which can support font hinting. It is essentially a simplification of the PS system for storing outline information only. Glyphs are described using cubic Bezier curves, as opposed to quadratic Bezier curves for TrueType fonts. In order to preview Type-1 fonts in typesetting applications, Adobe Type Manager was used.

Type 2 Provides a method for compact encoding of glyph procedures in an outline font program, and stored in a CFF (Compact Font Format) file. Type-2 charstring operators are a superset of Type-1 operators and provides smaller size and better rendering quality and performance. Type-2 charstrings are essentially designed to use less storage space than Type-1 by using operators with multiple arguments, various pre-defined default values, more efficient allotment of encoding values and shared sub-routines. A CFF file therefore provides a lossless compaction of Type-1 format using Type-2 charstrings.

Type 3 Contains glyphs defined using the full PS language rather than a subset. Hence it has capabilities not available in Type-1 fonts like shading, color, fill patterns.

CID fonts Adobe CID fonts are a new family of composite (multibyte) fonts designed to support the East Asian character sets like Chinese, Japanese, Korean, etc. CID-keyed refers to the character identifier (CID) numbers that are used to index and access the characters in the font. A CID (character identifier) font consists of a large font file that contains all the character outlines and a small character map file that contains a list of characters, encodings, and character identifiers. The combination of the font file and the character map file yields a font that is a specific character set and encoding information. Each CID font can support many character set and encoding combinations.

OpenType Font OpenType is a format for scalable digital fonts developed jointly by Microsoft and Adobe Systems, intended to supersede both the TrueType and the PostScript Type-1 fonts. In 2007 ISO incorporated OpenType as an ISO standard as part of the MPEG-4 specifications (MPEG-4 Part-22) called Open Font Format (OFF). OpenType is based on the basic spline based structure of TrueType but adds several options to enhance the font's typographic and language support capabilities, viz. the font character encoding is based on Unicode and can support any script, can have up to 65536 glyphs, can be used without modifications across several platforms like Mac, Windows, Unix, is usually smaller than Type-1 fonts. The font files can have extensions either OTF or TTF, however they have a different file icon of 'O' compared to 'TT' for TrueType.

Further Readings

- Font @ Wikipedia [*http://en.wikipedia.org/wiki/*
- Foreign language fonts [*http://www.kwintesser* … */foreign-language.html*]
- TrueType font @ Wi… [*http://en.wikipedi* … *ype*]
- TrueType refere… Apple [*http://d* … *om/fonts/TTRefMan/index.html*]
- Open font librar… *penfontlibrar*
- Unicode fonts fo… *://www.alan* … *de/fonts.html*]
- Unicode typefac… [*http://en.w* … */Unicode_typefaces*]
- Font resources [… *ype.com/*]
- Adobe type-1 for… *//partners.a* … */developer/en/font/T*

- CFF specifications [http://partners.adobe.com/public/developer/en/font/5176.CFF.pdf]
- Adobe CID fonts [http://www.adobe.com/products/postscript/pdfs/cid.pdf]
- Open type fonts @ Adobe [http://www.adobe.com/type/opentype/]
- Open type specifications @ Microsoft [http://www.microsoft.com/typography/otspec/default.htm]
- Free font downloads [http://www.fontyukle.net/en/]

2.5 Insertion of Text

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

Using a Keyboard The most common process of inserting text into a digital document is by typing the text using an input device like the keyboard. Usually, a text-editing software, like Microsoft Word, is used to control the appearance of text which allows the user to manipulate variables like the font, size, style, color, etc. Some image-processing and multimedia authoring software provide a separate text editing window where the user can type text and integrate it with the rest of the media like background images.

Copying and Pasting Another way of inserting text into a document is by copying text from a pre-existing digital document. The existing document is opened using the corresponding text processing program and portions of the text may be selected by using the keyboard or mouse. Using the "Copy" command, the selected text is copied to the clipboard. This text can then be inserted into another existing document or a new document or even in another place of the same document by choosing the "Paste" command, whereupon the text is copied from the clipboard into the target document. This text can then be edited as per the user's requirements.

Using Optical Character Recognition (OCR) A third way of inserting text into a digital document is by scanning it from a paper document. Text in a paper document including books, newspapers, magazines, letterheads, etc., can be converted into the electronic form using a device called the **scanner** (see Chapter 3 for more about scanners). The electronic representation of the paper document can then be saved as a file on the hard disk of the computer. The scanned file will however be an image file in which the text will be present as part of an image and will not be editable in a text processor. To be able to edit the text, it needs to be converted from the image format into the editable text format using a software called an **Optical Character Recognition** (OCR) software. The OCR software traditionally works by a method called **pattern matching**. Here, the software tries to match each and every scanned character to a already stored definition of characters within the software. For every match found, the software represents the character as an editable text object instead of an image object. This process is however largely dependant on the appearance of the characters for finding an adequate match and thus on the specific fonts. Some of the standard fonts may be converted with appreciable accuracy while other fonts may not be recognized. Recent research on OCR is based on another technology called **feature extraction**. Using this method, the software attempts to extract the core features of the characters and compare them to a table stored within itself for recognition. It is believed that if it can be done accurately enough, it will make the recognition process independent of the appearance or font of characters because the core set of features will remain same in any font representation. For example, if the software algorithm can extract the line, the circle and the arc in the positions shown below, it will recognize the character to be "a", whatever be the actual shape. See Fig. 2.4.

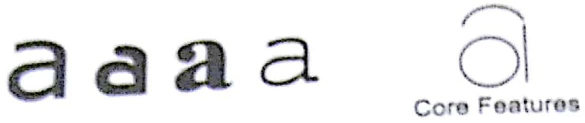


Fig. 2.4 *Core features can be isolated irrespective of font appearances*

Further Readings

- Text file @ Wikipedia [http://en.wikipedia.org/wiki/Text_file]
- On-line handwriting recognition [http://csis.pace.edu/~ctappert/dps/pdf/pen-tappert.pdf]
- OCR @ Wikipedia [http://en.wikipedia.org/wiki/Optical_character_recognition]
- OCR-A @ Wikipedia [http://en.wikipedia.org/wiki/OCR-A]

2.6 Text Compression

Large text documents covering a number of pages may take a lot of disk space. We can apply compression algorithms to reduce the size of the text file during storage. A reverse algorithm must be applied to decompress the file before its contents can be displayed on screen. However, to be meaningful, the compression-decompression process must not change the textual content in any way, not even a single character. The compression methods are explained below.

Run Length Encoding (RLE) The most basic type of text compression consist of replacing a string of repeated characters by a single instance of the character together with a count value which signifies how many times the character needs to be repeated. For example, if any text string is of the following type: ABCDDDDEEEEEEFFGH then it can be represented in a more compact way as follows: ABC!4D!6EFFGH. The '4' before D and 6 before E tells the decoder to repeat D four times and E six times while decoding. A special symbol '!' signifies that the number following is a count value and not part of the text content. It is clear that such a scheme can only be used if the number of repetitions is 3 or more. The scheme can also be extended for use with repeated words instead of characters.

Huffman Coding This type of coding is intended for applications in which the text to be compressed has known characteristics in terms of the characters used and their relative frequencies of occurrences. Using this information, instead of using fixed-length code words, an optimum set of variable-length code words is derived such that the shortest code word is used to represent the most frequently occurring characters. This approach is called the Huffman coding method. For further details, see Chapter 8.

Lempel–Ziv (LZ) Coding In the second approach followed by the Lempel–Ziv (LZ) method, instead of using a single character as a basis of the coding operation, a string of characters in used. For example, a table containing all the possible words that occur in a text document, is held by both the encoder and decoder. As each word occurs in the text, instead of representing the text as ASCII characters, the encoders stores only the index of where the word in stored in the table. The decoder converts the index into the appropriate words from the table. Thus the table is used as a dictionary and the LZ algorithm is also known as **dictionary-based algorithm**. The above method may not however produce efficient results for documents with a small subset of words in the dictionary. Hence, a variation of the above algorithm called **Lempel–Ziv–Welsh** (LZW) method allows the dictionary to be built up dynamically by the encoder and decoder for the document under processing. This dictionary becomes a better match for a specific document than a standard dictionary. For further details, see Chapter 8.

2.7 Text File Formats • •

2.7.1 TXT

TXT (Text) is an unformatted text document created by an editor like Notepad on Windows platform. Unformatted text documents can be used to transfer textual information between different platforms like Windows, DOS and UNIX. The data is encoded using ASCII codes but sometimes Unicode encodings like UTF-8 or UTF-16 may be used.

2.7.2 DOC, DOCX

DOC (Document) is a proprietary document file format developed by Microsoft as a native format for storing documents created by the MS-Word package in 1989. The DOC file format is used in various versions of MS Word over the years as shown below (Table 2.2). The DOC file format encompasses a rich set of formatting capabilities. It uses a master template which determines the margins, layouts and default fonts. Major formatting capabilities include support for different fonts, sizes, colors, styles (bold, italics, underline, superscript, subscript), lists (bulleted and numbered), Wordart (drawing text with graphical effects), justification (left, right, center), tables, borders and shading, paragraph indent, page feed, headers and footers, page numbering, sorting, watermark, and multiple columns. Additional facilities include inserting equations, tracking changes, support for thesaurus and spell check, aligning text in multiple columns, insertion of shapes and clip arts, symbols, and charts. From Word 2007, a new file format DOCX is used which emphasizes an XML based format. This is part of the Office Open XML format which has been standardized by European Computer Manufacturers Association (ECMA-376) and International Standards Organization/International Electrotechnical Commission (ISO/IEC 29500). The format defines a set of XML markup vocabularies for word-processing operations to facilitate extensibility and interoperability.

Table 2.2 *MS Word versions*

Year	MS Word ver
1989	Word for Windows 1.0
1990	Word for Windows 1.1
1991	Word for Windows 2.0
1993	Word for Windows 6.0
1995	Word 95
1997	Word 97
1998	Word 98
2000	Word 2000
2002	Word 2002
2003	Word 2003
2007	Word 2007
2010	Word 2010

2.7.3 RTF

RTF (Rich Text Format) is a proprietary document file format developed by Microsoft in 1987 for cross-platform document exchanges. It is the default format for Mac OS X's editor TextEdit. The WordPad editor earlier created RTF files by default although now it has switched to the DOC format. RTF control codes are human readable, similar to HTML code, e.g., \b for bold, \par for paragraph. RTF specifications are changed and published with major versions of MS Office. The different versions of RTF along with their MS Word versions are tabulated below (Table 2.3):

Table 2.3 *RTF versions*

Year	RTF ver	MS Word ver
1987	1/1.1	Word 3
1993	1.2	Word 6
1994	1.3	Word 6
1995	1.4	Word 95
1997	1.5	Word 97
1999	1.6	Word 2000
2001	1.7	Word 2002
2004	1.8	Word 2003
2008	1.9	Word 2007

2.7.4 PS, EPS

PS (PostScript) is a page-description language used mainly for desktop publishing. A page-description language is a high-level language that can describe the contents of a page such that it can be accurately displayed on output devices, usually printers. PostScript was developed in 1985 and soon became the ideal choice for graphical output for printing applications. In the same year, Apple Laserwriter was the first printer to ship with PostScript. Prior to PostScript, when printing graphics on a dot-matrix printer, the graphics had to be interpreted by the computer and then sent as a series of instructions to the printer to reproduce it. These printer control languages varied from printer to printer. PostScript offered a universal language that could be used for any brand of printer. PostScript represents all graphics and even text as vectors, i.e., as combinations of lines and curves. A PostScript compatible program converts an input document into the PS format, which is sent to the printer. A PostScript interpreter inside the printer converts the vectors back into the raster dots to be printed. See Fig. 2.5. This allows arbitrary scaling, rotating and other transformations. Thus to summarize, PostScript is a programming language similar to BASIC or FORTRAN or C, but designed to do only one thing: describe extremely accurately what a page looks like. The PS interpreter generates from this code the actual graphical output, which is then sent to the printer to be printed.

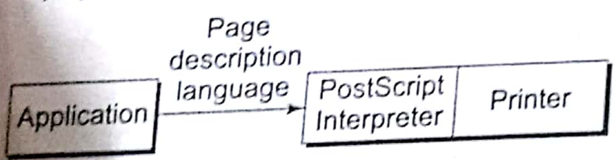


Fig. 2.5 *Interaction between application and postscript interpreter*

Another way to use PostScript is to generate the graphical output and then save this output as a file. Later this output file can be directly sent to the printer to be printed. Such an output file is called an **Encapsulated PostScript** (EPS) file. An EPS file can contain a low-resolution preview of the graphical content and it can be embedded or encapsulated within another document file as shown in Fig. 2.6. Hence, an EPS is essentially a saved ouput file generated from PS code which can be embedded within another document.

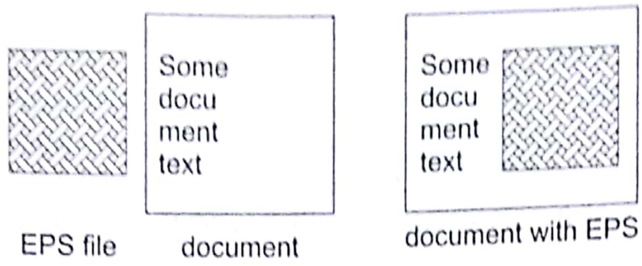


Fig. 2.6 *Encapsulating an EPS file*

2.7.5 PDF

PDF (Portable Document Format) is a file format developed by Adobe Systems in 1993 for cross-platform exchange of documents. In addition to text the format also supports images, graphics, fonts and recently 3D drawings. Initially a proprietary format, PDF was officially released as an open standard in 2008 published as ISO standard ISO/IEC 32000-1:2008. This implies anyone may write programs that can read and write PDFs without any associated royalty charges. Adobe started distributing its Acrobat Reader (now Adobe Reader) freely for reading PDF files. PDF readers can be downloaded for free from Adobe's site and there are several free open-source readers available as well. The PDF file format embodies three technologies: (1) as subset of the PostScript page description language for generating layout and graphical content, (2) a system to allow fonts to embed along with the document, and (3) a system for packaging all content with compression wherever appropriate. A PDF file can contain primarily 8 object types: Boolean, numbers, strings, names, arrays, dictionaries, streams, null. Objects may either be embedded within the document or linked to external files. PDF files support two types of layouts : non-optimized which consume less space but slower to access, and optimized which are structured so that a Web browser can read and display them without the entire file being loaded. Graphics support in PDF is very similar to PS, except the support for transparency which was added in PDF 1.4. Graphics in a PDF is described in terms of key graphical elements like color space, alpha (transparency) value, transformation matrix, Bezier curves, clipping paths, etc. Compression in PDF files can be based on LZW and RLE for text, and JPEG/JPEG2000 and CCITT fax for images. PDF currently also supports interactive elements like forms, through Adobe XML Forms Architecture (XFA) introduced in PDF 1.5. A brief summary of the various versions of PDF is given in the table below (Table 2.4):

Table 2.4 *PDF versions*

Year	PDF version	Salient features	Adobe Reader version
1993	1.0		1.0
1994	1.1	Password, encryption	2.0
1996	1.2	Radio buttons, checkboxes, mouse events, audio/video, Unicode	3.0
1999	1.3	Digital signature, ICC profiles, JavaScript, Web capture	4.0
2001	1.4	Transparency, interactive forms, XML tags, OCR text layer	5.0
2003	1.5	JPEG2000, XML Forms Architecture (XFA) ver 2	6.0
2005	1.6	3D artwork, open type font, AES encryption, XFA 2.2	7.0
2006	1.7	XFA 2.4, new string types, Read Out Loud	8.0
2008	1.7 ext. 3	XFA 2.6, Flash video, H.264	9.0
2009	1.7 ext. 5	XFA 3.0	9.1

PDF vs. PS While PS is a programming language to generate graphical and textual output, PDF is the generated output saved from a PS file, and hence more similar to an EPS file. However, the advancement of the EPS format has led to the invention of PDF, because it contains information not only on how a page looks, but also on how it behaves and what kind of information it contains. Moreover a PDF can also support embedded fonts, keywords for indexing and searching, interactive hyperlinks, transparency, etc. Hence PDF is sometimes referred to as a replacement of the EPS (output), not a replacement of PS (programming language).

2.7.6 DjVu

DjVu (pronounced de-ja-vu) is a file format developed by AT&T Laboratories in 1996 for storing text, graphics and images. It produces smaller file sizes than PDF, and is often considered a superior alternative to it. In 2002, DjVu was chosen by the Internet archive as the preferred format for e-books. DjVu is an open standard: the file format specifications as well as the source code of encoder and decoder are available. It uses image compression technique based on Wavelets which allow distribution of very high-resolution images over the Internet. It has been reported to achieve compression ratios 5 to 10 times better than existing compression techniques like JPEG and GIF for color documents, and 3 to 8 times better than TIFF for black-and-white documents. The DjVu plug-in is available for standard Web browsers on various platforms. The DjVu format is progressive. Users get an initial version of the page very quickly, and the visual quality of the page progressively improves as more bits arrive. One of the main techniques behind DjVu compression is the ability to separate an image into foreground and background components, which are then compressed separately at different resolutions. A list of stand-alone viewer programs exists to view DjVu files, e.g., WinDjView.

2.7.7 OXPS

OXPS (Open XML Paper Specification) is an open specification for a page description language and document format developed by Microsoft and later standardized by ECMA as standard ECMA-388. It is an XML based specification to define document layout with support for features like color gradients, transparencies, CMYK color space, color management schemes. It is similar in many respects to PDF but while PDF is based on PostScript, OXPS is based on XML. Windows Vista has native support for OXPS, while for other platforms like Windows XP, the XPS Essentials Pack includes an XPS viewer and document writer. XPS documents can also be viewed with the IE browser.

Further Readings

- MS Office file format specifications [*http://www.microsoft.com/interop/docs/OfficeBinaryFormats.mspx*]
- MS Word home page [*http://office.microsoft.com/word/*]
- Adobe PostScript 3 [*http://www.adobe.com/products/postscript/index.html*]
- PostScript FAQ [*http://en.wikibooks.org/wiki/PostScript_FAQ*]
- PostScript Language Reference [*http://partners.adobe.com/public/developer/en/ps/PLRM.pdf*]
- PostScript vs. PDF [*http://www.adobe.com/print/features/psvspdf/index.html*]
- Encapsulated PostScript spec [*http://partners.adobe.com/public/developer/en/ps/5002.EPSF_Spec.pdf*]
- PDF specification [*http://www.adobe.com/devnet/pdf/pdf_reference.html*]
- Adobe PDF reference version 1.7 [*http://www.adobe.com/devnet/acrobat/pdfs/pdf_reference_1-7.pdf*]
- ISO PDF 32000-1:2008 specification [*http://www.adobe.com/devnet/acrobat/pdfs/PDF32000_2008.pdf*]
- DjVu official site [*http://www.djvu.org*]
- OXPS official site [*http://www.ecma-international.org/publications/standards/Ecma-388.htm*]

Note 2.1: Open standard or open source are hardware and software whose specifications are publicly available. Anyone can use these specifications to develop hardware and software. Use of open standards increases compatibility between systems and platforms. However, open-source items may or may not be free. For example, most of the international specifications by ITU, ISO and IEC are considered open-source but requires licensing fees for implementation. Examples of open-source hardware include PCI and AGP, while those of open-source software include HTML, TCP/IP, and PDF.

Conclusions

- Unformatted text comprises of fixed-size characters from the ASCII character sets.
- Formatted text have associated appearance parameters like shape, size, color determined by font files.
- Hypertext consists of hyperlinks which enables the user to navigate between different points within documents.
- The anchor is the visual element associated with hypertext while the pointer defines the navigation pathway.
- The architecture of a hypertext document consists of a storage layer, an abstraction layer and a presentation layer.
- The ASCII character set have recently been superseded by the Unicode set capable of 32-bit encoding.
- To utilize storage space efficiently, the Unicode characters are encoded using 16-bit and 8-bit representations.
- True-Type fonts have vector representations and can be scaled to various sizes without quality degradations.
- An OCR algorithm converts a page of scanned text into editable alpha-numeric characters.
- The Huffman coding algorithm compresses text documents by representing more frequently occurring strings by shorter codes .
- Lempel–Ziv coding uses a dictionary to represent text strings by their code words in the dictionary.
- Text file formats like RTF and DOC can be used to store editable text while formats like PDF are used to display non-editable content.
- A page-description language like PostScript is a high-level language that can describe the contents of a page such that it can be accurately displayed on output devices, usually printers.

Review Questions

1. Explain how unformatted text is represented internally via the ASCII table.
2. Distinguish between standard ASCII table and extended ASCII table.
3. Differentiate between formatted and unformatted text.
4. What is the advantage of hypertext? Distinguish between anchor and target document.
5. What is the advantage of Unicode standard over ASCII standard?
6. Differentiate between Unicode Transformation Formats UTF-32, UTF-16, and UTF-8.
7. What is meant by font? Differentiate between bitmap fonts and vector fonts.
8. Distinguish between font size and font style.
9. Distinguish between TrueType, PostScript, and OpenType fonts.
10. How can text be inserted within an application?
11. In relation to OCR software, distinguish between pattern-matching and feature-extraction.