Multimedia Systems

By

Dr. Priyambada Subudhi
Assistant Professor
IIIT Sri City

Evaluation Criteria

Sl. No.	Mode of Assessment	% of Marks
1	Term Exam-1	20
2	Term Exam-2	20
3	Term Exam-3	20
4	Scheduled Quiz	10
5	Assignment	10
6	Project	20

Syllabus

- **Unit - 1** [4 Hours]: Introduction to Multimedia, Multimedia Data Representation, Components of Multimedia Systems.
- **Unit - 2** [6 Hours]: Text: Visual Representation, Digital Representation, File Formats: RTF, TIFF, Digital Image: Digital Image Representation (2D format, resolution), Color Models.
- **Unit - 3** [6 Hours]: Digitization of Video, Types of video signals (component, composite and Svideo), 3D Video and TV, Basic Sound Concepts: computer representation of sound, File Formats – WAV, MPEG Audio, Quantization and Transmission of Audio: PCM, DM, DPCM.
- **Unit - 4** [8 Hours]: Lossless Compression Algorithms: Run-Length Coding, Variable-Length Coding (Huffman Coding, Adaptive Huffman Coding), Arithmetic Coding, Adaptive Arithmetic Coding, Dictionary-Based Coding, Context-based Coding, Lossy Compression Algorithms: Standard Image Compression Techniques (JPEG, JPEG 2000), Video Compression Technique (MPEG).
- **Unit - 5** [8 Hours]: Fundamentals of data communication and networking, Bandwidth requirements of different media, Real time constraints: latency, video data rate, Multimedia over LAN and WAN, Multimedia conferencing, Video-on-demand broadcasting issues
- **Unit - 6** [4 Hours]: Basics of multimedia retrieval, Content Based Image Retrieval (CBIR), Key technologies and issues in current CBIR systems

Books

Text Books:

- Ze-Nian Li, and Mark S. Drew, "Fundamentals of Multimedia", PHI Learning.

Reference Books:

1. Fred Halsall, "Multimedia Communications: Applications, Networks, Protocols and Standards", Pearson.
2. Khalid Sayood, "Introduction to Data Compression", Elsevier Publication.
3. Asit Dan and Dinkar Sitaram, "Multimedia Servers", Elsevier, 2006.

What is Multimedia?

- The term multimedia is composed of two words
  - Multi : numerous or multiple
  - media : agent for something

- Multimedia could be defined as the usage of multiple media or agents for communication of information.

- These media or agents could be in the form of text, images, audio, video, graphics, animation etc.

- A good general working definition is

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

- Example: A music video and sound should be used together as one without another would lose its significance.

- The term media can be categorized based on few criteria
  - Perception media
  - Representation media
  - Presentation media
  - Storage media
  - Transmission media

Perception media

- "How do human perceive information"
- We perceive information from what we see and what we hear.
- Visual media
    - Text, graphics, images and video
- Auditory media
    - Music, sound, voice

Representation media

- “How information is encoded in the computer”
- Referring to how the information is represented internally in computer.
- Several options:
  - Text is encoded in ASCII
  - An audio data stream in PCM
  - Image in JPEG format
  - Video in MPEG format

Presentation media

- "Which medium is used to output information from the computer or input in the computer"
- Refers to physical means used by systems to reproduce information for humans, e.g: audio and visual devices.
- Input
  - Keyboards, cameras, microphone, Head mounted device
- Output
  - Paper, monitors, loud speakers

Storage media

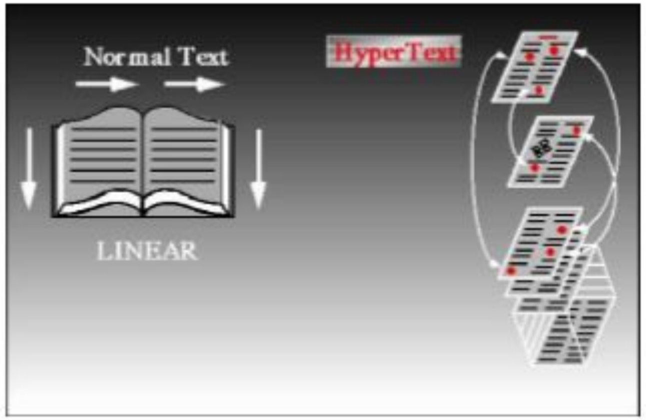
- "Where information is stored"
- Refers to various physical means for storing computer data
- E.g. magnetic tapes, magnetic disks or digital optical disks (CD-ROM, CD, DVD)

Transmission media

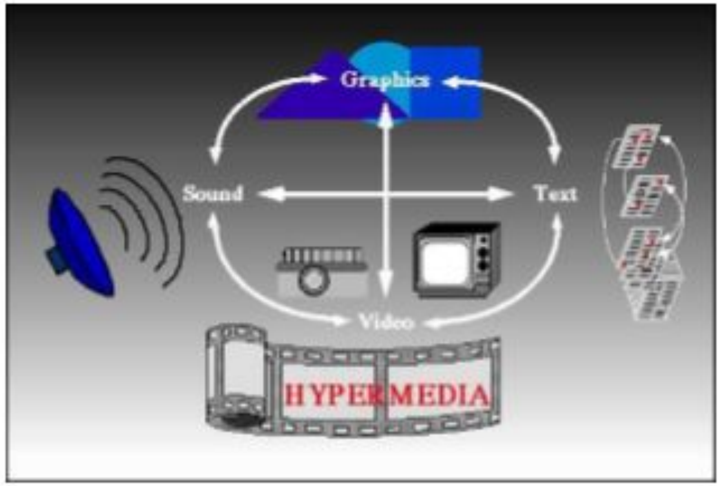
- "Which medium is used to transmit data"
- Refers to physical means – cable of various type (coaxial cable, twisted pair, fibre optics), radio tower, satellite – that allow the transmission of telecommunication signals.

Characteristics of multimedia

- It can increase the impact of the message or impact on the user.
- Digital and Voluminous
- Interactivity:
  - When the end-user is able to control the elements of media that are required, and subsequently obtains the required information in a non-linear way.
- Hypertext and Hypermedia support:
  - Hypertext is a text which contains links to other texts.
  - Traversal through pages of hypertext is therefore usually non-linear (as indicated below)



- HyperMedia is not constrained to be text-based. It can include other media, e.g., graphics, images, and especially continuous media – sound and video.
- The World Wide Web is the largest and most commonly used hypermedia application.



- It can involve more than one input device.
- It can repeated (over and over).
- It is generally dynamic, not static.

Multimedia Systems Lecture – 2

By

Dr. Priyambada Subudhi

Assistant Professor

IIIT Sri City

Categories of Multimedia

- Multimedia may be divided into following three categories based on their functions and how they are organized.

Static and Dynamic

- Static refers to cases when the multimedia data remains the same within a certain finite time, for example, one slide of a Microsoft PowerPoint presentation or one HTML Web page. Compare
- The dynamic case is when the data is changing, for example watching a video

Linear and non-linear:

- Information is read or viewed in a continuous sequence. E.g. A powerpoint presentation
- Non-linear multimedia information is not presented in sequential or chronological manner and usually interactive. E.g: Computer Game

Interactive and non-interactive:

- Interactive multimedia is the means to interface with different media through input and output devices allowing a user to make decisions as to what takes place next with multimedia

Real-time and orchestrated (recorded):

- Orchestrated refers to cases when there is no real-time requirement. E.g. compressing content on a DVD and distributing
- The most important constraint here is the quality of the compressed data.
- However, showing a game in a live broadcast over the Internet imposes a whole new set of engineering constraints in addition to compression quality, which relate to on-time data delivery and synchronization.

Person-to-machine versus person-to-person

- Based on whether the end user is interacting with a machine or with another person. For example, playing a CD-ROM game is a simple person-to-machine experience.
- videoconferencing is a person-to-person experience.

Single user, peer-to-peer, peer-to-multipeer, and broadcast

- Here, the manner of information distribution is used as a means to classify.
- a single-user scenario such as browsing the Web, or a peer-to-peer scenario when the information is exchanged from one person/computer to another, for example two friends instant messaging over the Internet.
- A peer-to-multipeer scenario extends the paradigm to sending messages to a limited set of intended viewers such as in a chat room.
- Broadcasting is the most general-purpose scenario, where information is sent not to any specific listener(s) but available to all those who want to listen, such as television and radio broadcasts

Multimedia Applications

- Definition: A Multimedia Application is an application which uses a collection of multiple media sources e.g. text, graphics, images, sound/audio, animation and/or video.
- Examples:
  - Video Teleconferencing: Transmission of synchronized video and audio in real-time through computer networks in between two or more multipoints (or participants) separated by locations.
  - Multimedia Store and Forward Mail: Allow users to generate, modify and receive documents that contain multimedia. Eg. Gmail, Hotmail, Yahoo etc.
  - Advertising and Purchasing: Most of the web sites visited have many advertisements with multimedia features with the objective of marketing merchandise or offering services online.

- ## For entertainment:
  - Multimedia is heavily used in the entertainment industry, especially to develop special effects in movies and animations. Multimedia games are a popular pastime and are software programs available either as CD-ROMs or online
- ## For Education:
  - Multimedia is used to produce computer-based training courses (popularly called CBTs) and reference books like encyclopedia and almanacs.
- ## For Healthcare:
  - Multimedia best use in healthcare is for real time monitoring of conditions of patients in critical illness or accident.
  - Multimedia makes it possible to consult a surgeon or an expert who can watch an ongoing surgery line on his PC monitor and give online advice at any crucial juncture.
- ## Other multimedia applications available to us at home are
  - Basic Television Services, Digital Audio, Video on demand, Home shopping, Digital multimedia libraries, E-Newspapers, e-magazines

History of Multimedia

- Multimedia to communicate ideas might begin with newspapers, which were perhaps the first mass communication medium, using text, graphics, and images.
- Thomas Alva Edison 'commissioned the invention of a motion picture camera in 1887.
- In 1895, Guglielmo Marconi sent his first wireless radio transmission at POlltecchio, Italy. Initially invented for telegraph, radio is a major medium for audio broadcasting.
- Television was the new medium for the twentieth century. It established video as a commonly available medium and has since changed the world of mass communication.

The connection between computers and digital media represented using the discrete binary format, emerged only over a short period:

1945 : As part of MIT's postwar deliberations on what to do with all those scientists employed on the war effort, Vannevar Bush (1890-1974) wrote a landmark article describing what amounts to a hypermedia system, called "Memex."

1960s Ted Nelson started the Xanadu project and coined the term "hypertext."

1967 Nicholas Negroponte formed the Architecture Machine Group at MIT.

1969 Nelson and van Dam at Brown University created an early hypertext editor called FRESS.

1976 The MIT Architecture Machine Group proposed a project entitled "Multiple Media." This resulted in the Aspen Movie Map, the first videodisk, in 1978.

1982 The Compact Disc (CD) was made commercially available by Philips and Sony, which was soon becoming the standard and popular medium for digital audio data.

1985 Negroponte and Wiesner co-founded the MIT Media Lab, a leading research institution investigating digital video and multimedia.

1990 Kristina Hooper Woolsey headed the Apple Multimedia Lab, with a staff of 100.

1991 MPEG1 was approved as an international standard for digital video. Its further development led to newer standards, MPEG-2, MPEG- 4, and further MPEGs, in the 1990s.

1991 The introduction of PDAs in 1991 began a new period in the use of computers in general and multimedia in particular.

1992 JPEG was accepted as the international standard for digital image compression, which remains widely used today.

1992 The first audio multicast on the multicast backbone (MBone) was made.

1995 The JAVA language was created for platform-independent application development, which was widely used for developing multimedia applications.

1996 DVD video was introduced; high-quality, full-length movies were distributed on a single disk.

1998 Handheld MP3 audio players were introduced to the consumer market, initially with 32 MB of flash memory.

2000 World Wide Web (WWW) size was estimated at over 1 billion pages.

2001 The first peer-to-peer file sharing (mostly MP3 music) system.

2003 Skype was released for free peer-to-peer voice over the Internet.

2004 Web 2.0 was recognized as a new way to utilize software developers and end-users use the Web.

2005 YouTube was created, providing an easy portal for video sharing, which was purchased by Google in late 2006.

2006 Twitter was created, and rapidly gained world wide popularity, with 500 million registered users in 2012.

2007 Apple launched the first generation of iPhone, running the iOS mobile operating system.

2008 The first Android-powered phone was sold and Google Play the Android's primary app store, was soon launched.

2009 The first LTE (Long Term Evolution) network was setup making an important step toward 4G wireless networking.

2010 Netflix, which used to be a DVD rental service provider, migrated its infrastructure to the Amazon AWS cloud computing platform, and became a major online streaming video provider.

Elements of Multimedia Data

- The common elements of multimedia includes
  - Text:

    All multimedia productions contain some amount of text. The text can have various types of fonts and sizes to suit the profession presentation of the multimedia software.
  - Graphics:
    - Graphics make the multimedia application attractive.
    - People do not like reading large amount of textual matter on the screen.
    - There are two types of Graphics:
      - Bitmap Images : Bitmap images are real images that can be captured from devices such as digital cameras or scanners. Generally bitmap images are not editable. Bitmap images require a large amount of memory.
      - Vector Graphics : Vector graphics are drawn on the computer and only require a small amount of memory. These graphics are editable.

- ## Audio:
  - A multimedia application may require the use of ***speech, music and sound*** effects. These are called audio or sound element of multimedia.
  - Speech is also a perfect way for teaching. Audio are of ***analog*** and ***digital*** types. Analog audio or sound refers to the original sound signal.
  - Computer stores the sound in ***digital form***. Therefore, the sound used in multimedia application is digital audio.
- ## Video:
  - The term video refers to the moving picture, accompanied by sound such as a picture in television.
  - Video element of multimedia application gives a lot of information in small duration of time.
  - Digital video is useful in multimedia application for showing real life objects.
  - Video have highest performance demand on the computer memory and on the bandwidth if placed on the internet.
  - Digital video files can be stored like any other files in the computer.

Animation:

- Animation is a process of making a static graphical elements look like it is moving.
- An animation is just a continuous series of still graphical elements that are displayed in a sequence.
- The animation can be used effectively for attracting attention.
- Animation also makes a presentation light and attractive.

Multimedia Data Representation: Text

- Source: Keyboard, speech input, optical character recognition, data stored on disk
- Stored and input character by character:
  - Storage of text is 1 byte per char / more bytes for Unicode.
  - For other forms of data (e.g. Spreadsheet les). May store format as text (with formatting) others may use binary encoding.
- Format: Raw text or formatted text e.g HTML, Rich Text Format (RTF), Word or a program language source (Java, Python, MATLAB etc.)
- Not temporal. But may have natural implied sequence e.g. HTML format sequence, Sequence of C program statements.
- Size Not significant w.r.t. other Multimedia data.

Images

- Input: digitally scanned photographs/pictures or directly from a digital camera.
- May also be generated by programs “similar” to graphics or animation programs.
- Still pictures which (uncompressed) are represented as a bitmap (a grid of pixels organized as a 2D array).
- The two dimensions specify the width and height of the images. Each pixel has also a bit depth which represents the number of bits assigned to each pixel.
- Stored at ***1 bit per pixel*** (Black and White), ***8 Bits per pixel*** (Gray Scale, Colour Map) or ***24 Bits per pixel*** (True Colour).
- Size: a **512x512** Gray scale image takes up ***1/4 MB***, a 512x512 24 bit image takes 3/4 MB with no compression.

Graphics

- Format: constructed by the composition of primitive objects such as ***lines***, ***polygons***, ***circles***, ***curves*** and ***arcs***.
- Input: Graphics are usually generated by a ***graphics editor*** program (e.g. Illustrator) or automatically by a program (e.g. Postscript).
- Graphics are usually editable or revisable (unlike Images).
- Graphics input devices: keyboard (for text and cursor control), mouse, trackball or graphics tablet.
- Graphics standards : OpenGL, PHIGS, GKS
- Graphics files usually store the primitive assembly
- Do not take up a very high storage overhead.

Audio

- Digital audio is characterized by a sampling rate in hertz, which gives the number of samples per second.

A sample can be defined as an ***individual unit*** of audio information.

- Each sample also has a size, the ***sample size***, which typically is anywhere from 8-bits to 16-bits depending on the application.

- CD Quality Audio requires 16-bit sampling at 44.1 KHz Even higher audiophile rates (e.g. 24-bit, 96 KHz)

- Audio signal is also described by dimensionality i.e. the dimensions of an audio signal signify the ***number of channels*** that are contained in the signal. These may be mono (one channel), stereo (two channels).
- Usually compressed (E.g. MP3, AAC, Flac, Ogg Vorbis).

Video

- Input: Analog Video is usually captured by a video camera and then digitized.
- Video is represented as a sequence of images. Each image in the sequence typically has the same properties of width, height, and pixel depth.
- There is one more temporal parameter known as frames per second or fps.
- Typically, videos have **25**, **30** or **50** frames per second.
- E.g. A ***512×512*** size monochrome video images take ***25×0.25 = 6.25MB*** for a second to store uncompressed.
- Digital video clearly needs to be compressed for most times.

Multimedia Systems
Lecture – 3

By

Dr. Priyambada Subudhi
Assistant Professor
IIIT Sri City

Multimedia Systems

- A Multimedia System is a system capable of processing multimedia data and applications.

- A Multimedia System is characterized by the ***processing***, ***storage***, ***generation***, ***manipulation*** and ***rendition*** of Multimedia information.

Characteristics of Multimedia Systems

- A Multimedia system has four basic characteristics:
  - Multimedia systems must be computer controlled.
  - Multimedia systems are integrated.
  - The information they handle must be represented digitally.
  - The interface to the final presentation of media is usually interactive.

Challenges for Multimedia Systems

- Distributed Networks
- Temporal relationship between data
  - Render different data at same time- continuously.
  - Sequencing within the media:

    playing frames in correct order/time frame in video
  - Synchronisation- inter-media scheduling

e.g. Video and Audio - Lip synchronization is clearly important for humans to watch playback of video and audio and even animation and audio.

Key Issues for Multimedia Systems

The key issues multimedia systems need to deal with here are:

- How to represent and store temporal information.
- How to strictly maintain the temporal relationships on play back/retrieval
- What process are involved in the above.
- Data has to represented digitally- Analog-Digital Conversion, Sampling etc.
- Large Data Requirements- bandwidth, storage

Data compression is usually mandatory

Desirable Features for a Multimedia System

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

- Very High Processing Power- needed to deal with large data processing and real time delivery of media. Special hardware commonplace.
- Multimedia Capable File System- needed to deliver real-time media -e.g. Video/Audio Streaming.
- Special Hardware/Software needed - e.g. RAID technology.
- Data Representations - File Formats that support multimedia should be easy to handle yet allow for compression/decompression in real-time.

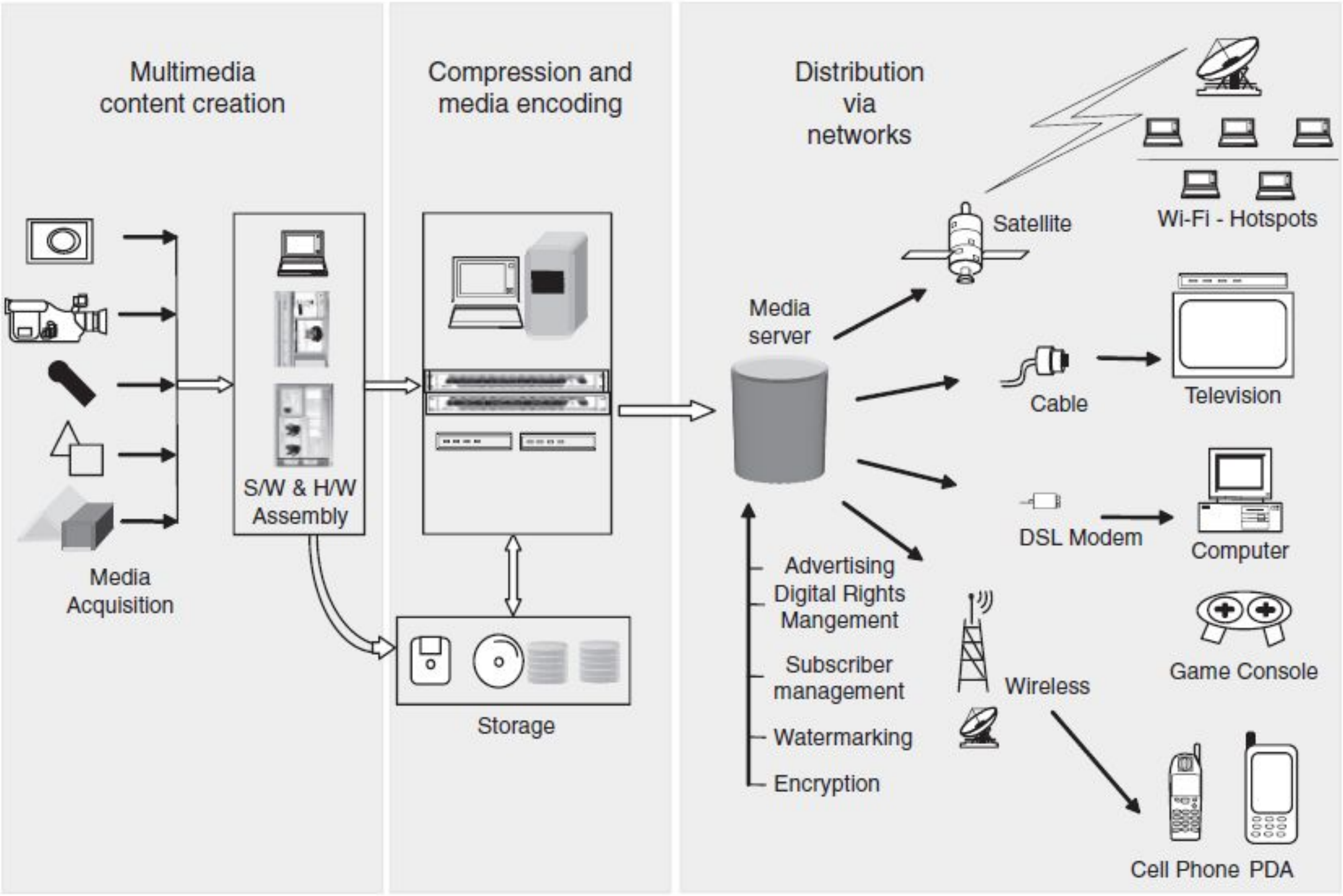
- Efficient and High I/O – input and output to the file subsystem needs to be efficient and fast. Needs to allow for real-time recording as well as playback of data. e.g. Direct to Disk recording systems.
- Special Operating System - Needs to support direct transfers to disk, real-time scheduling, fast interrupt processing, I/O streaming etc.
- Storage and Memory - large storage units (of the order of hundreds of Tb if not more) and large memory (several Gb or more). Large Caches also required and high speed buses for efficient management.
- Network Support - Client-server systems common as distributed systems common.
- Software Tools - user friendly tools needed to handle media, design and develop applications, deliver media.

Components of Multimedia Systems

Multimedia systems can be logically grouped into three parts whose primary functionalities are

- ***Content production***
- ***Compression and storage***
- ***Distribution to various end users and platforms***

Components of a multimedia system today



Content Production:

- It includes a variety of different instruments, which capture different media types in a digital format. These include digital ***cameras, camcorders or video cameras, sound recording devices, scanners to scan images***, *and* ***3D graphical objects***.
- Once the individual media elements are in their digital representations, they may be further combined to create coherent, interactive presentations using software (S/W) applications and hardware (H/W) elements.
- This content can be stored to disk, or in the case of real-time applications, the content can be sent directly to the end user via digital networks.

- **Compression and Storage:**
  - It deals with the compression of multimedia content
  - This entails the use of various compression technologies to compress video, audio, graphics, and so on.
- **Distribution**
  - It deals with media distribution across a variety of low-bandwidth and high-bandwidth networks.
  - This ranges from cellular to wireless networks, to cable, to digital subscriber line (DSL), to satellite networks.
  - Distribution normally follows standards protocols, which are responsible for collating and reliably sending information to end receivers.
  - The commonly used end receivers are computers, televisions, set-top boxes, cell phones, or even more application- or entertainment-specific items, such as video game consoles.

Multimedia Systems
Lecture – 4

By

Dr. Priyambada Subudhi
Assistant Professor
IIIT Sri City

Text: Visual Representation

- Text has become a part of our life. It consists of characters, punctuation symbols, etc. to convey a message.

- Text is one of the most imperative components of multimedia and an essential source of presenting information to a wide range of people.

- Proper use of text, keeping in mind elements such as ***font style, size*** and ***various design tools*** help the content creator to communicate the idea and message to the user.

Fonts and Faces

- Factors affecting legibility of text are as follows:
  - Size and style
  - Background and foreground colours
  - Leading
- A glyph is a graphic representation of a character's shape where a character may be represented by many glyphs.

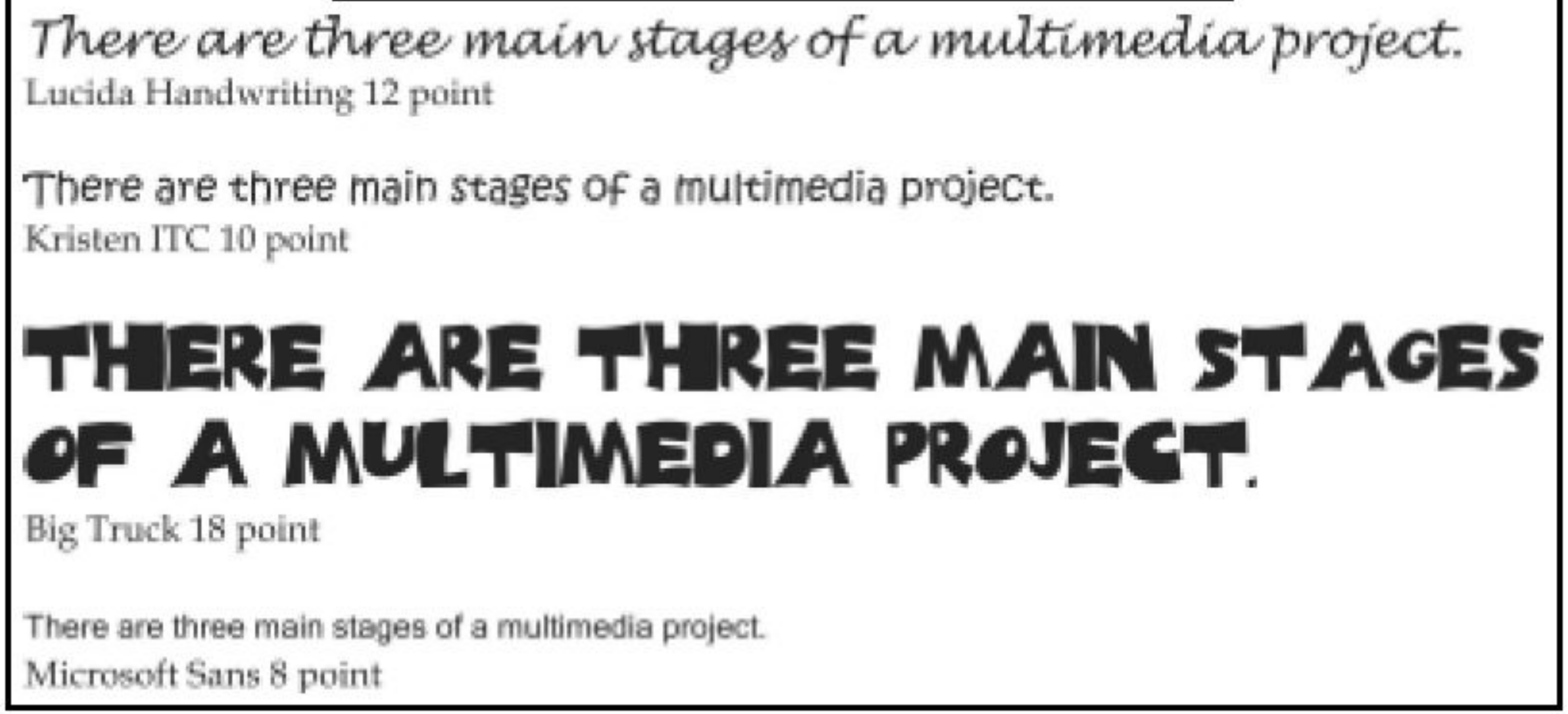
- A typeface is a particular set of glyphs or sorts (an alphabet and its corresponding accessories such as numerals and punctuation) that share a common design.
- E.g. ***Arial*** is a well known typeface.
- A font is a collection of character or glyphs of a single size and style belonging to a particular typeface family.
- Fonts are classified on the basis of ***spacing between characters***, ***words, presence or absence of serifs, their shape, stretch and weight*** such as ***bold or italics***.
- E.g. 12 point Arial is a font, and 10 point Arial is a separate font.

- The spacing between character pairs is called kerning and the space between lines is called leading.

Figure 2.1: Different Types of Fonts

- Font size is measured in points and it does not describe the height or width of its character. This happens because the height of two different fonts (in both upper and lower case) may differ.
- One point is approximately 1/72 of an inch i.e., 0.0138.

Figure 2.2: Examples of Different Fonts and Font Sizes



Serif and sans serif typeface: Times New Roman, Bodoni, Bookman are some fonts which come under serif category.

Arial, Avant Garde, Verdana are some examples of sans serif font.

Figure 2.3: Examples of Serif, Sans Serif and Decorative Fonts

Bodoni

Interactive multimedia is called hypermedia.

(This is a serif font. In this font, a line or curve extension from the end of a letter. Serif fonts are best used for body text.)

Avant Garde

Interactive multimedia is called hypermedia.

(This is a sans serif font. There are no extensions in this font. Sans Serif fonts are best used for titles.)

Matura MT Script

Interactive multimedia is called hypermedia.

(This is a decorative font. These fonts are stylish and formal and are best used for emphasis.)

Selecting Text Fonts

There are a few things that a user must keep in mind before selecting fonts for a multimedia presentation.

- Choose a font that is legible and easy to read.
- The different effects and colours of a font can be chosen to make the text look distinctive.
- Try to use few different colours within the same presentation.
- Try to use few typefaces within the same presentation. Play with the style and size to match up to the purpose and importance of the text.
- Drop caps and initial caps can be used to accent the words.

- To attract instant attention to the text, the words can be wrapped onto a sphere or bent like a wave.
- In case of text links (anchors) on web pages the messages can be highlighted.
- Meaningful words and phrases can be used for links and menu items.
- Overcrowding of text on a single page should be avoided.
- Do not use decorative passages for longer paragraphs.

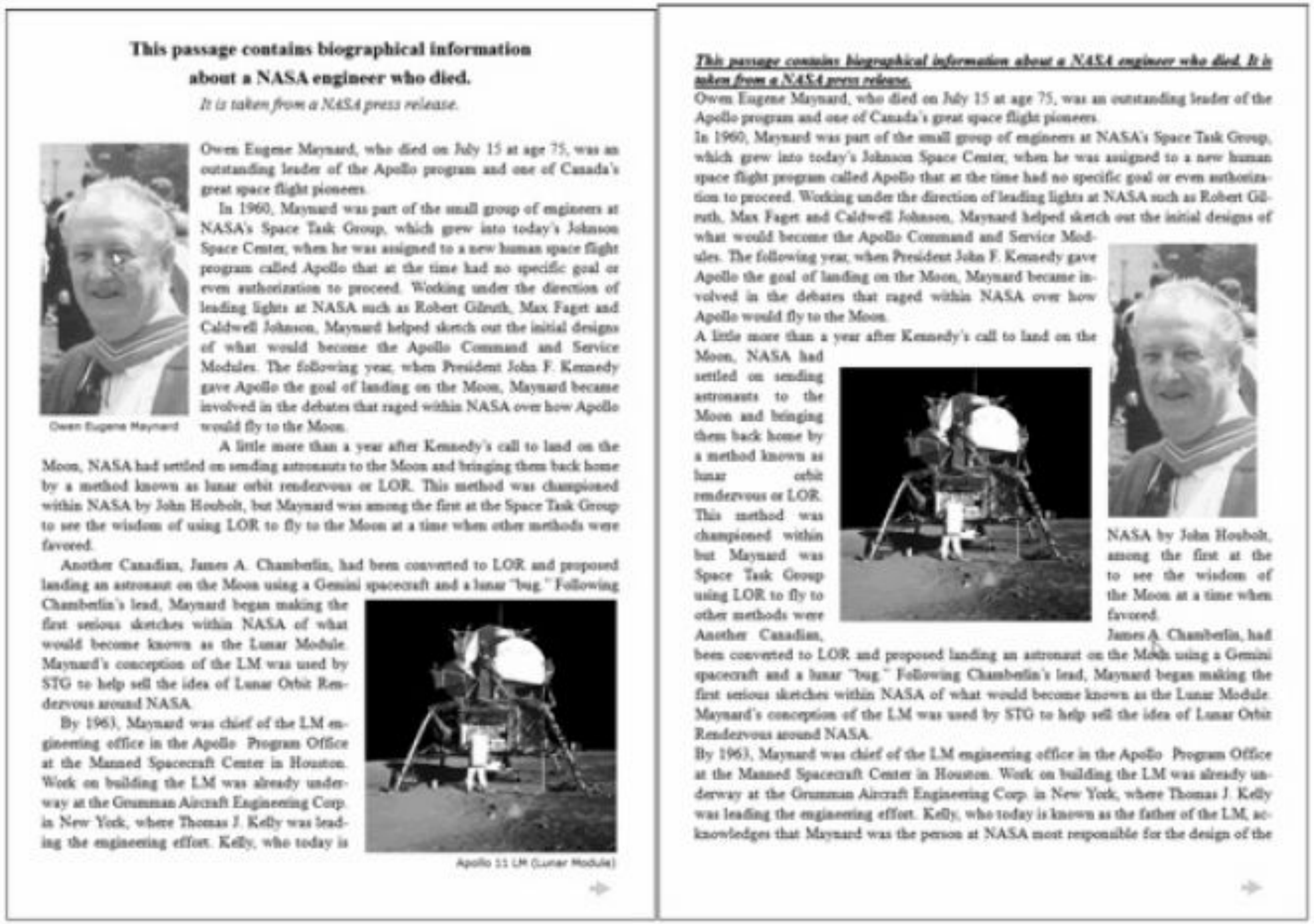
Using Text Elements in a Multimedia Presentation

The text elements used in multimedia are given below:

- **Menus for Navigation**
  - A user navigates through content using a menu.
  - A simple menu consists of a text list of topics.
- **Interactive Buttons**
  - A button is a clickable object that executes a command when activated.
  - Users can create their own buttons from bitmaps and graphics.
  - The design and labelling of the buttons should be treated as an industrial art project.
- **Fields for Reading**
  - Reading a hard copy is easier and faster than reading from the computer screen.
  - A document can be printed in one of two orientations - portrait or landscape.

- **HTML Documents**
  - HTML is the standard document format used for Web pages.
  - HTML documents are marked using tags.
  - Some of the commonly used tags are:
    - The <B> tag for making text bold faced.
    - The <OL> tag for creating an ordered list.
    - The <IMG> tag for inserting images.
- **Text Layout**
  - While creating a multimedia presentation, the presenter should plan the text layout to let a reader read it with ease.
  - The length of the text should neither too long nor too short.
  - The next point of concern is the consistency of pages.

Figure 2.4: Example of Good and Poor Page Layout



This passage contains biographical information about a NASA engineer who died.

It is taken from a NASA press release.

Owen Eugene Maynard

Owen Eugene Maynard, who died on July 15 at age 75, was an outstanding leader of the Apollo program and one of Canada's great space flight pioneers.

In 1960, Maynard was part of the small group of engineers at NASA's Space Task Group, which grew into today's Johnson Space Center, when he was assigned to a new human space flight program called Apollo that at the time had no specific goal or even authorization to proceed. Working under the direction of leading lights at NASA such as Robert Gilruth, Max Faget and Caldwell Johnson, Maynard helped sketch out the initial designs of what would become the Apollo Command and Service Modules. The following year, when President John F. Kennedy gave Apollo the goal of landing on the Moon, Maynard became involved in the debates that raged within NASA over how Apollo would fly to the Moon.

A little more than a year after Kennedy's call to land on the Moon, NASA had settled on sending astronauts to the Moon and bringing them back home by a method known as lunar orbit rendezvous or LOR. This method was championed within NASA by John Houbolt, but Maynard was among the first at the Space Task Group to see the wisdom of using LOR to fly to the Moon at a time when other methods were favored.

Another Canadian, James A. Chamberlin, had been converted to LOR and proposed landing an astronaut on the Moon using a Gemini spacecraft and a lunar "bug." Following Chamberlin's lead, Maynard began making the first serious sketches within NASA of what would become known as the Lunar Module. Maynard's conception of the LM was used by STG to help sell the idea of Lunar Orbit Rendezvous around NASA.

By 1963, Maynard was chief of the LM engineering office in the Apollo Program Office at the Manned Spacecraft Center in Houston. Work on building the LM was already underway at the Grumman Aircraft Engineering Corp. in New York, where Thomas J. Kelly was leading the engineering effort. Kelly, who today is

Apollo 11 LM (Lunar Module)

This passage contains biographical information about a NASA engineer who died. It is taken from a NASA press release.

Owen Eugene Maynard, who died on July 15 at age 75, was an outstanding leader of the Apollo program and one of Canada's great space flight pioneers.

In 1960, Maynard was part of the small group of engineers at NASA's Space Task Group, which grew into today's Johnson Space Center, when he was assigned to a new human space flight program called Apollo that at the time had no specific goal or even authorization to proceed. Working under the direction of leading lights at NASA such as Robert Gilruth, Max Faget and Caldwell Johnson, Maynard helped sketch out the initial designs of what would become the Apollo Command and Service Modules. The following year, when President John F. Kennedy gave Apollo the goal of landing on the Moon, Maynard became involved in the debates that raged within NASA over how Apollo would fly to the Moon.

A little more than a year after Kennedy's call to land on the Moon, NASA had settled on sending astronauts to the Moon and bringing them back home by a method known as lunar orbit rendezvous or LOR. This method was championed within NASA by John Houbolt, but Maynard was among the first at the Space Task Group to see the wisdom of using LOR to fly to the Moon at a time when other methods were favored.

Another Canadian, James A. Chamberlin, had been converted to LOR and proposed landing an astronaut on the Moon using a Gemini spacecraft and a lunar "bug." Following Chamberlin's lead, Maynard began making the first serious sketches within NASA of what would become known as the Lunar Module. Maynard's conception of the LM was used by STG to help sell the idea of Lunar Orbit Rendezvous around NASA.

By 1963, Maynard was chief of the LM engineering office in the Apollo Program Office at the Manned Spacecraft Center in Houston. Work on building the LM was already underway at the Grumman Aircraft Engineering Corp. in New York, where Thomas J. Kelly was leading the engineering effort. Kelly, who today is
known as the father of the LM, acknowledges that Maynard was the person at NASA most responsible for the design of the

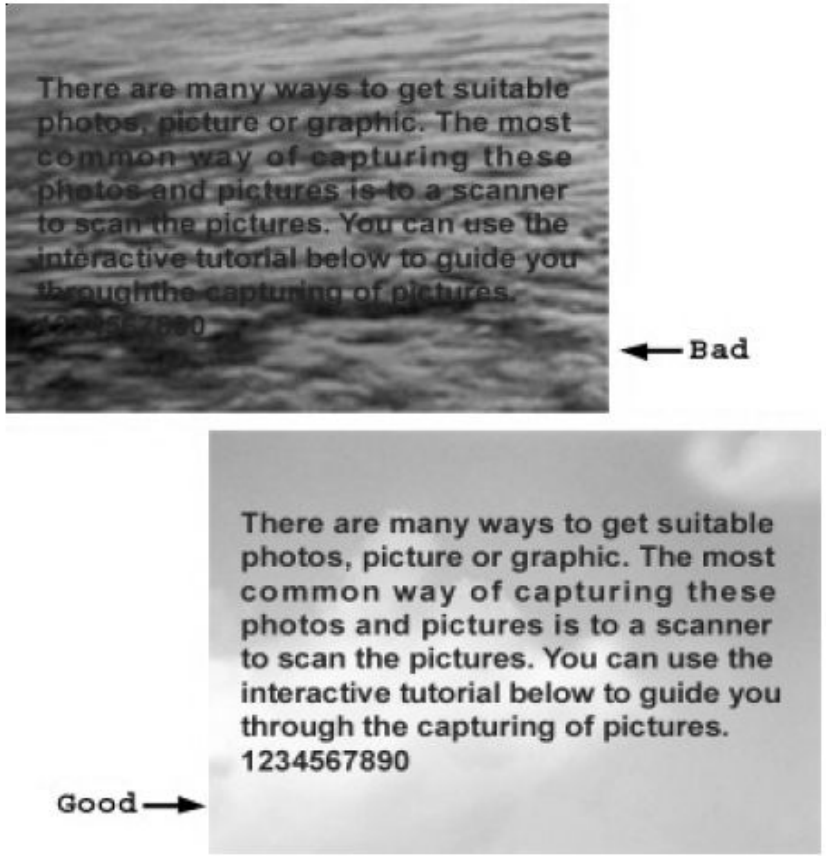
- The distance between two lines should be adjusted to a suitable value to increase readability.
- Ensure that the leading is not too small as then the text will be hard to read.

Figure 2.5: Effects of Different Leadings

Tight	If your software allows you to change the leading (the distance between two text lines), you should adjust it to a suitable value.
Normal	If your software allows you to change the leading (the distance between two text lines), you should adjust it to a suitable value.
Loose	If your software allows you to change the leading (the distance between two text lines), you should adjust it to a suitable value.

- The other one of the most common errors people make is while choosing the background colour. Using a background colour too close to the text or a background image highly in contrast to the text which makes the text difficult to read.

Figure 2.6: Use of Colour and Background



Use of Text in Webs

- Using text in websites attract a visitor's attention as well as help him in understanding the webpage better.

- It is far better than the use of meaningless graphics and images which do not contribute in understanding of the page.

- **Website Loading Speed:** A website which contains a lot of text loads faster than the websites that contains a lot of images and graphics, audio and video clips on the page.

Text in Films

- Most films start with titles and end with credits.
- Typography look different in different formats such as a in film subtitles, on websites, poster etc.
- While designing subtitles, a film maker will need to keep in mind that moving images interact with the top layer subtitles.
  - E.g. If subtitles are white and rest on top of a similar white tone in the image, the text will be difficult and impossible to read. To ensure this does not happen, a black outline around text should be used. Now the text will be viewable even against common black and white backgrounds.

Text in Advertisements

- Since the text ads are more of keyword oriented, they draw more attention than banner advertising.
- The text ads are inexpensive, thus making it affordable and effective for your business.
- There are a few websites which offers a flat free rental services to place your text based advertisements.
- The foremost benefit of having text based advertisements is that it helps in improving your search engine ranking.

Multimedia Systems
Lecture – 5

By

Dr. Priyambada Subudhi
Assistant Professor
IIIT Sri City

Digital Representation of Text

- Any type of data needs to be converted to its digital form (in the form 1s and 0s) to be represented in the Computer.

- Text (letters, symbols, and numerals) also needs to be encoded in digital form is what allows computers to manipulate and communicate text.

- Two character encoding standards define how characters are decoded from ones and zeros into the text you see on the screen right now, and into the different languages viewed every day on the World Wide Web.

- These two encoding standards are ASCII and Unicode.

ASCII

- The **American Standard Code for Information Interchange** (ASCII) was developed to create an international standard for encoding the Latin alphabet.
- In 1963, ASCII was adopted so information could be interpreted between computers; representing lower and upper letters, numbers, symbols, and some commands.
- Because ASCII is encoded using ones and zeros, the base 2 number system, it uses seven bits.
- Seven bits allow 2 to the power of 7 = 128 possible combinations of digits to encode a character.

ASCII control characters

00	NULL	(Null character)
01	SOH	(Start of Header)
02	STX	(Start of Text)
03	ETX	(End of Text)
04	EOT	(End of Trans.)
05	ENQ	(Enquiry)
06	ACK	(Acknowledgement)
07	BEL	(Bell)
08	BS	(Backspace)
09	HT	(Horizontal Tab)
10	LF	(Line feed)
11	VT	(Vertical Tab)
12	FF	(Form feed)
13	CR	(Carriage return)
14	SO	(Shift Out)
15	SI	(Shift In)
16	DLE	(Data link escape)
17	DC1	(Device control 1)
18	DC2	(Device control 2)
19	DC3	(Device control 3)
20	DC4	(Device control 4)
21	NAK	(Negative acknowl.)
22	SYN	(Synchronous idle)
23	ETB	(End of trans. block)
24	CAN	(Cancel)
25	EM	(End of medium)
26	SUB	(Substitute)
27	ESC	(Escape)
28	FS	(File separator)
29	GS	(Group separator)
30	RS	(Record separator)
31	US	(Unit separator)
127	DEL	(Delete)

ASCII printable characters

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

Extended ASCII characters

128	Ç	160	á	192	└	224	Ó
129	ü	161	í	193	┴	225	ß
130	é	162	ó	194	┬	226	Ô
131	â	163	ú	195	├	227	Ò
132	ä	164	ñ	196	─	228	õ
133	à	165	Ñ	197	┼	229	Õ
134	å	166	ª	198	ã	230	µ
135	ç	167	º	199	Ã	231	þ
136	ê	168	¿	200	╚	232	Þ
137	ë	169	®	201	╔	233	Ú
138	è	170	¬	202	╩	234	Û
139	ï	171	½	203	╦	235	Ù
140	î	172	¼	204	╠	236	ý
141	ì	173	¡	205	═	237	Ý
142	Ä	174	«	206	╬	238	¯
143	Å	175	»	207	¤	239	´
144	É	176	░	208	ð	240	≡
145	æ	177	▒	209	Ð	241	±
146	Æ	178	▓	210	Ê	242	‗
147	ô	179	│	211	Ë	243	¾
148	ö	180	┤	212	È	244	¶
149	ò	181	Á	213	ı	245	§
150	û	182	Â	214	Í	246	÷
151	ù	183	À	215	Î	247	¸
152	ÿ	184	©	216	Ï	248	°
153	Ö	185	╣	217	┘	249	¨
154	Ü	186	║	218	┌	250	·
155	ø	187	╗	219	█	251	¹
156	£	188	╝	220	▄	252	³
157	Ø	189	¢	221	¦	253	²
158	×	190	¥	222	Ì	254	■
159	ƒ	191	┐	223	▀	255	nbsp

How encoding ASCII works

- You already know how to convert between decimal and binary numbers
- You now need to turn letters into binary numbers
- Every character has a corresponding decimal number (for example, A → 65)
- ASCII uses 7 bits
- We use the first 7 columns of the conversion table to create 128 different numbers (from 0 to 127)

- Example: 1000001 gives us the number 65 (64+1), which corresponds to the letter A

64	32	16	8	4	2	1
1	0	0	0	0	0	1

- Here’s how ‘HELLO’ is encoded in ASCII in binary:

Latin character	ASCII
H	1001000
E	1000101
L	1001100
L	1001100
O	1001111

Extended ASCII

- Because ASCII encodes characters in 7 bits, moving to 8-bit computing technology meant there was one extra bit to be used.
- With this extra digit, **Extended ASCII** encoded up to 256 characters.
- However, the problem that developed was that countries that used different languages did different things with this extra capacity for encoding.
- Many countries added their own additional characters, and different numbers represented different characters in different languages.
- The problem of incompatible encoding systems became more urgent with the invention of the World Wide Web, as people shared digital documents all over the world, using multiple languages.
- To address the issue, the **Unicode Consortium** established a universal encoding system called **Unicode**.

Unicode

- ***Unicode (Unique, Universal, and Uniform character enCoding)*** is the new standard for representing characters of all the languages of the World.
- The latest version of Unicode contains a repertoire of more than 120,000 characters covering ***129 modern*** and ***historic scripts***, as well as ***multiple symbol sets***.
- ASCII character encoding is a subset of Unicode.
- Unicode can be implemented by different character encodings. The Unicode standard defines ***UTF-8***, ***UTF-16*** and ***UTF-32***.
- So, these use between 8 and 32 bits per character and has the advantage that it represents many more unique characters than ASCII because of the larger number of bits available to store a character code.
- It uses the same codes as ASCII up to 127.

- UTF-8 the dominant encoding on the World Wide Web (*used in over 92% of websites*), uses one byte for the first 128 code points, and up to 4 bytes for other characters. The first 128 Unicode code points are the ASCII characters, which means that any ASCII text is also a UTF-8 text.
- UTF-16 uses **16bits** to represent each character. This means that it is capable of representing *65,536* different characters.
- UTF-32 used **32bits** to represent each character, meaning it can represent a character set *4,294, 967,298* possible characters, enough for all known languages.

character	encoding	bits
A	UTF-8	01000001
A	UTF-16	00000000 01000001
A	UTF-32	00000000 00000000 00000000 01000001
あ	UTF-8	11100011 10000001 10000010
あ	UTF-16	00110000 01000010
あ	UTF-32	00000000 00000000 00110000 01000010

- UTF-16 and UTF-32 suffers from two major problems
  - Unicode uses ASCII for the English language, so A is still 65. However, encoded in 32 bits, the binary representation for the letter A would be 000000000000000000000000000000000001000001. This wastes a lot of valuable space!
  - Many older computers interpret eight zeros in a row (a null) as the end of a string of characters. So these computers wouldn’t send any characters that came after eight zeros in a row (they wouldn’t send an A if it was represented as 000000000000000000000000000000000001000001).

How UTF -8 encoding works

- The Unicode encoding method **UTF-8** solves these problems of UTF-16 and UTF-32. It works like as follows.
  - Up to character number 128, the regular ASCII value is used (so for example A is 01000001)
  - For any character beyond 128, UTF-8 separates the code into two bytes and adding '110' to the start of first byte to show that it is a beginning byte, and '10' to the start of second byte to show that it follows the first byte.
  - So, for each character beyond number 128, you have two bytes:

    [110xxxxx] [10xxxxxx]

  And you just fill in the binary for the number in between:

    [11000101][10000101] (that's the number 325 -> 00101000101)

- This works for the first 2048 characters.

- For characters beyond that, one more '1' is added at the beginning of the first byte and a third byte is also used:

  [1110xxxx] [10xxxxxx] [10xxxxxx]

- This gives you 16 spaces for binary code. In this manner, UTF-8 goes up to four bytes:

  [11110xxx] [10xxxxxx] [10xxxxxx] [10xxxxxx]

- How the Devanagari character क, with code point 2325 can be represented in UTF-8, UTF-16 and UTF-32.

Multimedia Systems
Lecture – 6

By

Dr. Priyambada Subudhi
Assistant Professor
IIIT Sri City

Text File Formats

- Text can be in the form of plain text or rich text.

Plain Text:

- Plain text is a text with no styles attached or embedded with it.
- It is just a plain text i.e. simple text.
- The files with plain text are saved with .txt or .TXT extension.

Let’s apply this theory in practice:

- Open Notepad, or whichever plain text editor you prefer
- Type a message and save it, e.g. ‘data is beautiful’
- Look at the size of the file — mine is 18 bytes
- Now, add another word, e.g. ‘data is SO beautiful’
- If you look at the file size again, you’ll see that it has changed — my file is now 3 bytes larger (SO[SPACE]: the ‘S’, the ‘O’, and the space)

Advantages of plain text –

- if you write something in plain text then whenever you open that file on any machine then the end result will be same that is it will show the text as it is on every system.
- Unlike rich text if you format the text and add the styles to it then it might create problem if the end person doesn't have suitable app for it then it will distort its styles as a result it will leave a bad impression on the user.
- It is fast and flexible as it doesn't uses styles and formatting therefore it is fast and flexible.
- file size of plain text format is less as compared to rich text format.

Disadvantages of plain text –

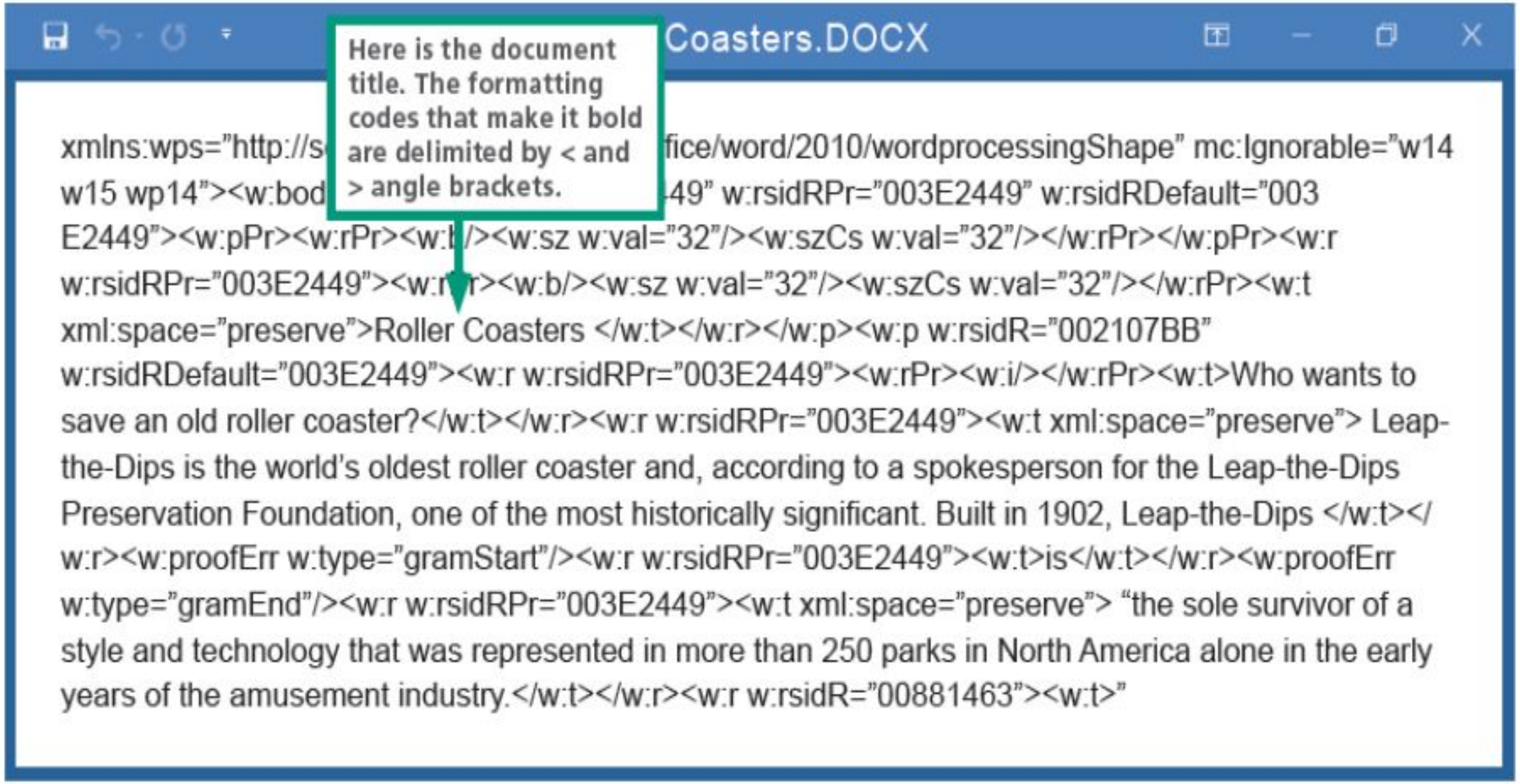
- In the todays era almost 90% of the word processor uses rich text format.
- It looks very basic i.e. without styles.
- To read the plain text it becomes boring sometimes.

Rich Text :

- Rich text is a text with styles attached or embedded with it so whenever you are copying the rich text and pasting to editor which support rich text then it will automatically include all the styles formatting spacing etc.

- Rich text format was created by Microsoft in 80s and they discontinued the development of rich text in 2008. Rich text files are usually save with .rtf or .RTF extension.

Formatted Text



Advantages of rich text –

- In almost 90% of word processors supports the rich text format.
- It looks interesting that a text with styles rather then simple text as it include styling and formatting as a result developer learn how to style and format the text.
- To keep the end user connected to the text without getting bored.

Disadvantages of rich text –

- If the rich text file is opened in word processor which do not support rich text then it doesn't looks good to the user as the formatting might not support the word processor.
- file size is large as compared to plain text as it includes styles and formatting information with it
- it is also bit complex to add the styles and formatting as it takes more time compared to plain text

Multimedia Systems
Lecture – 7

By

Dr. Priyambada Subudhi
Assistant Professor
IIIT Sri City

Digital Image Representation

- The most common form to represent natural images and other forms of graphics that are rich in detail is bitmap.

- The term bitmap refers to how a given pattern of bits maps to a specific color.

- They store this information in a grid of points (array of points) or pixels, which has a fixed width and height, and they can store various ranges of colours according to the image type.

A Bitmap image

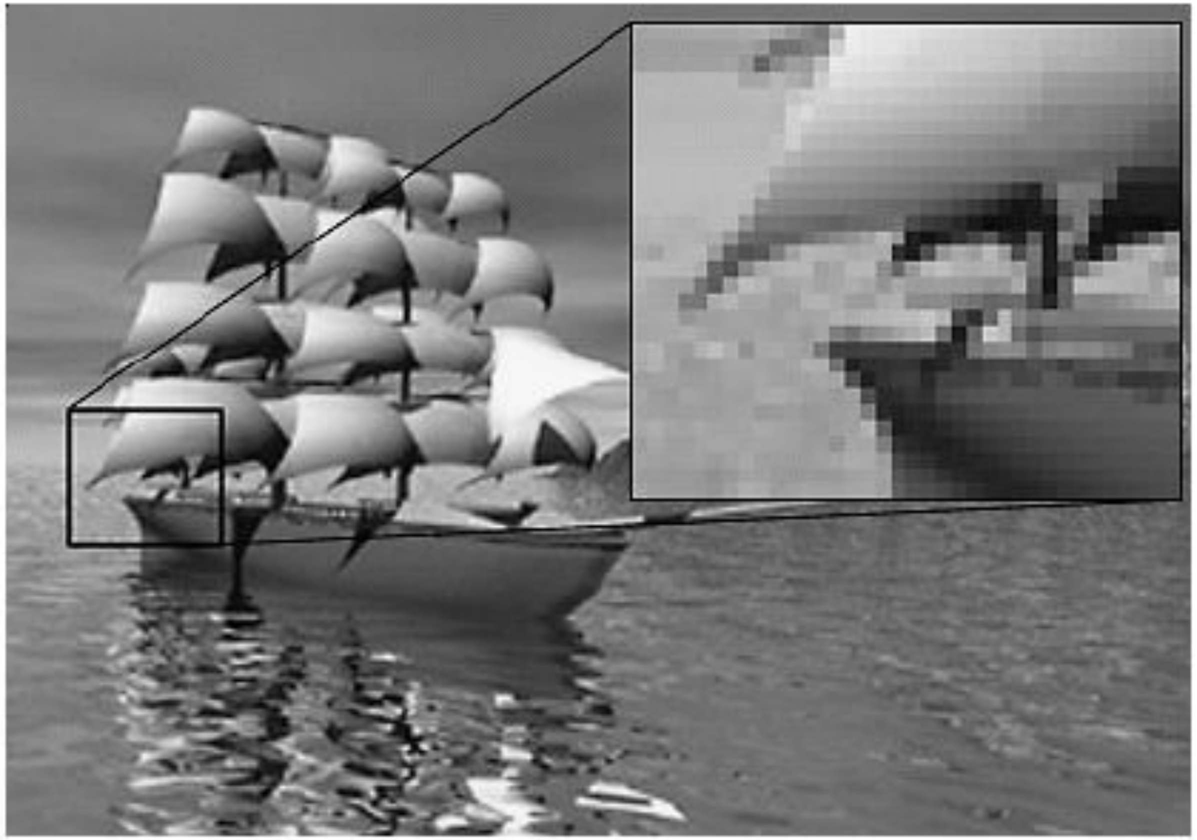


Image Data Type

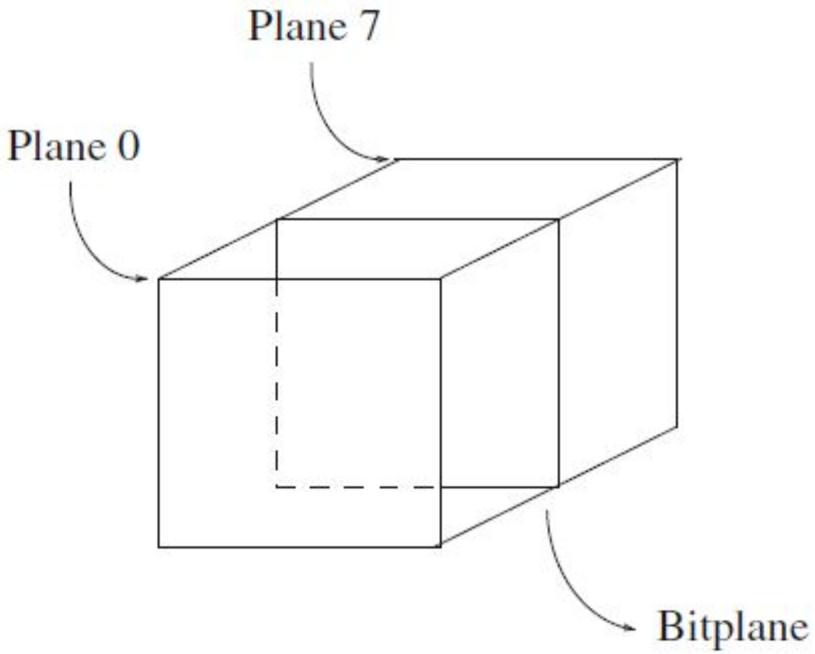
1-Bit Images

- A 1-bit image consists of on and off bits only and thus is the simplest type of image.
- Each pixel is stored as a single bit (0 or 1). Hence, such an image is also referred to as a *binary image*.
- Monochrome 1-bit images can be satisfactory for pictures containing only simple graphics and text.
- Moreover, fax machines use 1-bit data, so in fact 1-bit images are still important even though storage capacities have increased.

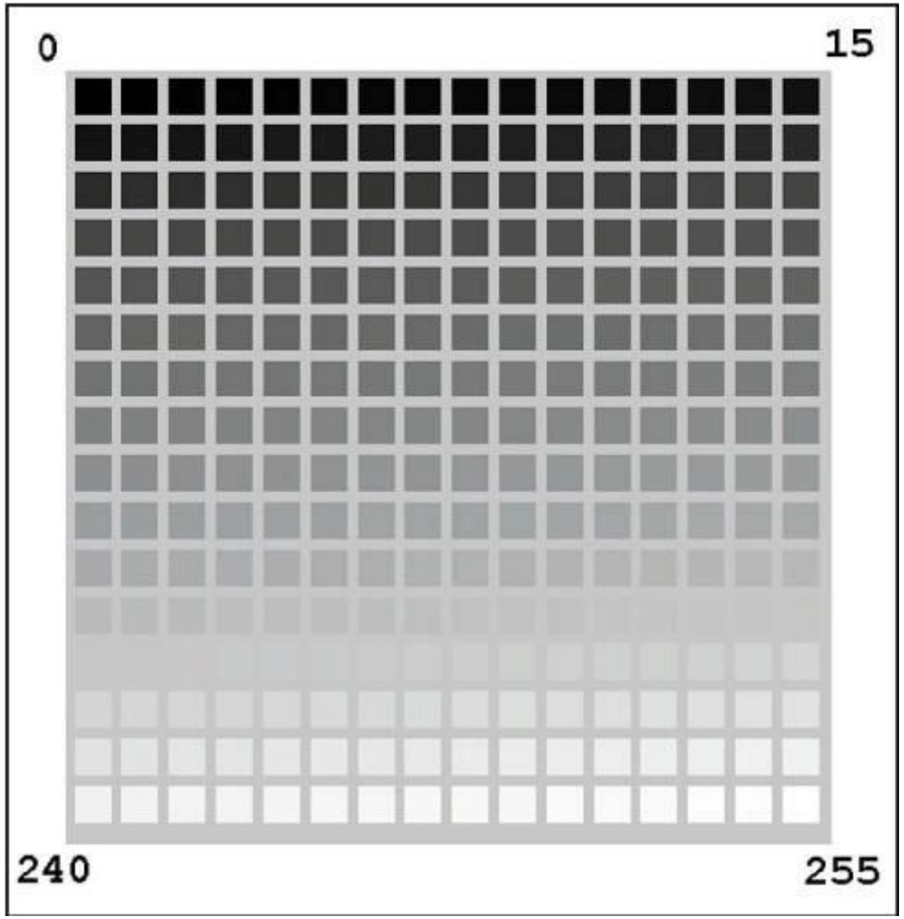
- This is the 1-bit monochrome Lena image of size 640×480.
- A 640×480 monochrome image requires 38.4 kilobytes (kB) of storage.

8-Bit Gray-Level Images

- In an 8-bit image, each pixel has a *gray value* between ***0 and 255***.
- Each pixel is represented by a single byte—for example, a dark pixel might have a value of 10, and a bright one might be 230.
- We can think of the 8-bit image as a set of 1-bit *bitplanes*, where each plane consists of a 1-bit representation of the image: a bit is turned on if the image pixel has a nonzero value at that bit level.



The 256 colors gray



- Each bitplane can have a value of 0 or 1 at each pixel but, together, all the bitplanes make up a single byte that stores values between 0 and 255
- For the least significant bit, the bit value translates to 0 or 1 in the final numeric sum of the binary number.
- Second bit makes a contribution of 0 or 2 to the final sum. The next bits stand for 0 or 4, 0 or 8, and so on, up to 0 or 128 for the most significant bit.
- Each pixel is usually stored as a byte (a value between 0 and 255), so a 640×480 grayscale image requires 300kB of storage (640×480 = 307,200).

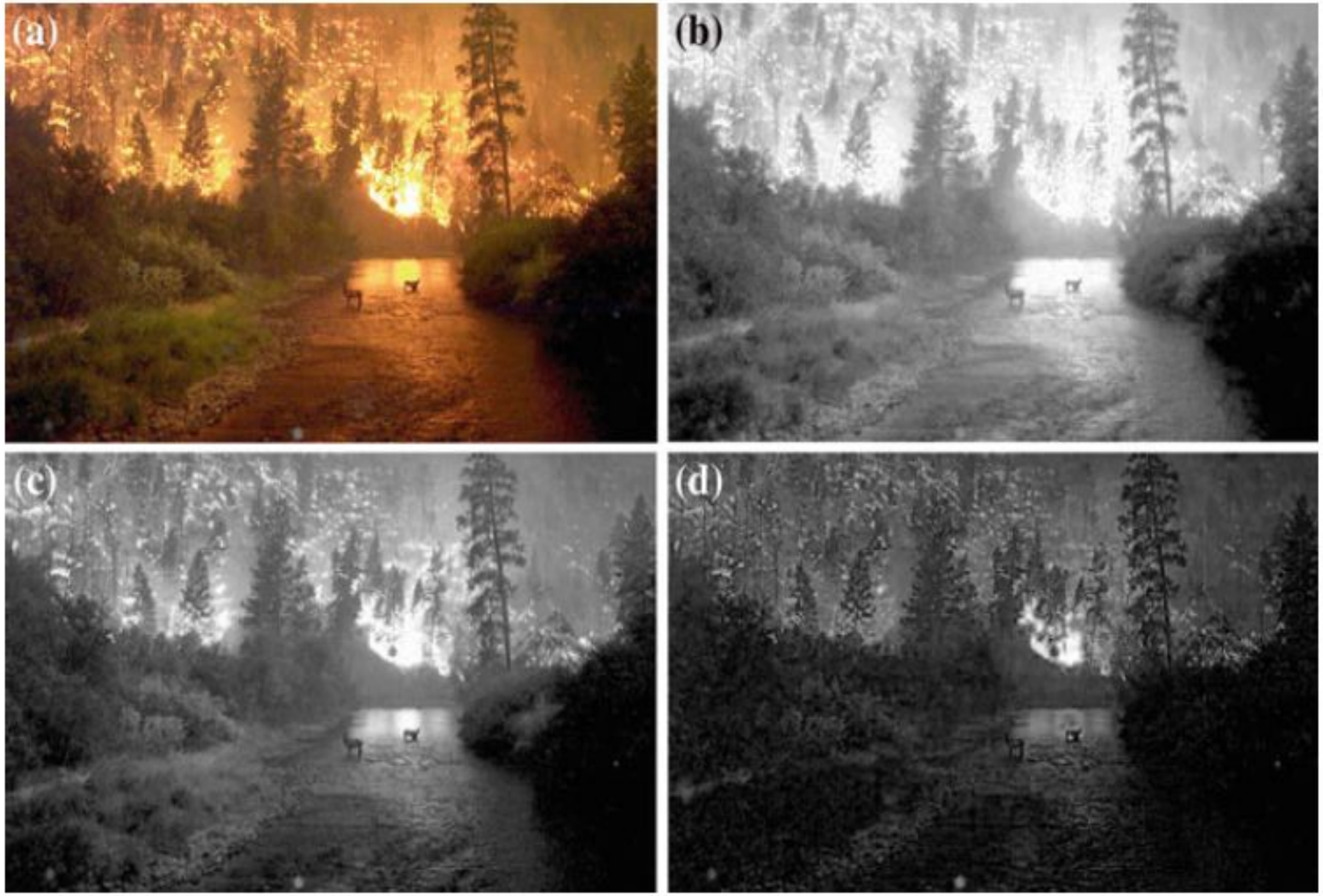
Lena gray-level image

24-Bit Color Images

- In a color 24-bit image, each pixel is represented by three bytes, usually representing RGB (Red, Blue, Green).
- Since each value is in the range 0–255, this format supports 256×256×256, or a total of 16,777,216 possible combined colors.
- However, such flexibility does result in a storage penalty:
  - a 640 × 480 24-bit color image would require 921.6kB of storage without any compression.
- Many 24-bit color images are actually stored as 32-bit images, with the extra byte of data for each pixel storing an α (*alpha*) value representing special-effect information.

Lena RGB color image

An RGB image and separate R, G and B channels



Higher Bit-Depth Image

- More information about the scene being imaged can be gained by using special cameras that view more than just three colors, i.e., RGB.

- E.g: To produce medical images of skin that can utilize the additional colors to better diagnose skin ailments such as carcinoma,   in satellite imaging, where extra information can give indication of types of crop growth.

- Such images are called *multispectral* (more than three colors) or hyperspectral (a great many image planes, say 224 colors for satellite imaging).

Resolution

- Image resolution refers to the number of pixels in a digital image (higher resolution always yields better quality). Fairly high resolution for such an image might be 1,600 × 1,200, whereas lower resolution might be 640 × 480.
- Resolution of screen or printing device is a measure of how finely a device displays graphics with pixels. It is used by printers, scanners, monitors (TV, computer), mobile devices and cameras.
- There are two types ways of measuring resolution:
  - The amount of ***dpi (dots per inch).*** dpi refers to the number of printed dots contained within one inch of an image printed by a printer.
  - ***ppi (Pixels Per Inch)*** is a term also used refers to the number of pixels contained within one inch of an image displayed on a computer monitor.

Example of a Bitmap Image at Different Levels of Magnification

- Bitmap images contain a fixed number of pixels, usually measured in pixels per inch (ppi).
- *Example: 1* A one inch by one inch image with a resolution of 72 ppi contains a total of 5184 pixels (72 pixels wide × 72 pixels high = 5184). The same one inch by one inch image with a resolution of 300 ppi would contain a total of 90,000 pixels.
- Example-2: The scan is pixels per inch. If you scan a 4x5 at 2000 pixels per inch, the resulting file will be 8000x10000 pixels.
- So you have 8000x10000 pixels available to print. If you make a 16x20 inch print, you have 500 pixels per inch. (8000/16) Once you have a given file and decide on a print size, the pixels per inch is determined. There is nothing you can do to change that.
- In the printer driver there is a choice of dots per inch to use. That is how many dots of ink to lay down on the paper per inch of paper. It is usual to have several dots for each pixel in the file being printed.
- For example, the Epson 3880 and P7000 printers have dpi choices from a low of 720x720 up to 2880x1440. The print quality should be better for the 2880x1440 dpi setting than for the 720x720 dpi setting.

Multimedia Systems
Lecture – 8

By

Dr. Priyambada Subudhi

Assistant Professor

IIIT Sri City

8-bit color images

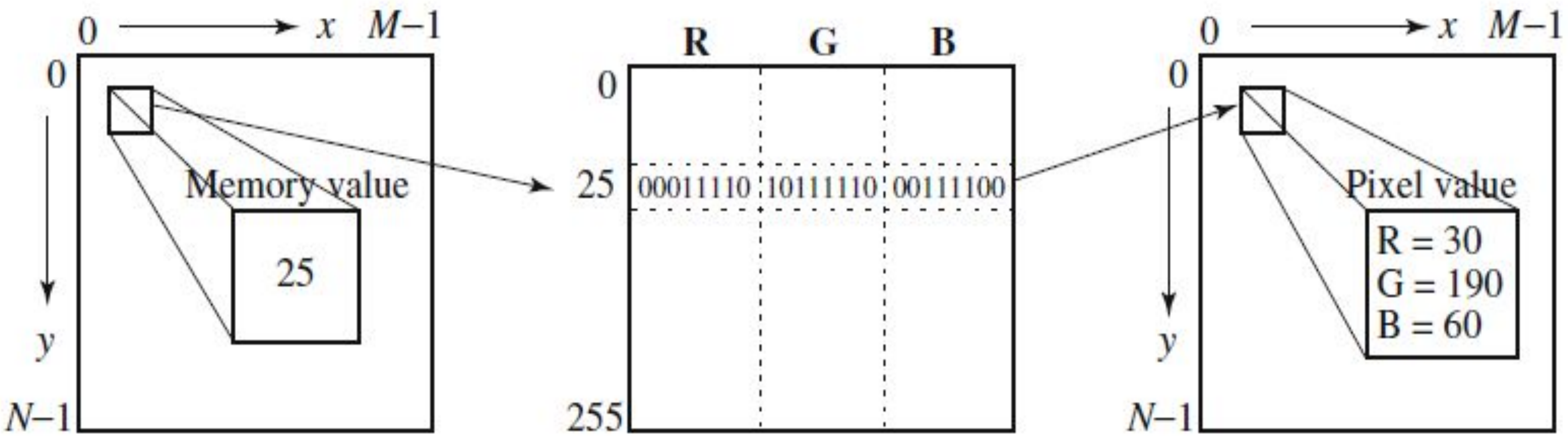
- Many systems can utilize color information stored with only 8 bits of information (the so-called 256 colors) in producing a screen image.

- Such image files use the concept of a lookup table to store color information.
  - Basically, the image stores not color, but instead just a set of bytes, each of which is actually an index into a table with 3-byte values that specify the color for a pixel with that lookup table index.

Color Lookup Tables

- The idea used in 8-bit color images is to store only the index, or code value, for each pixel.
- Then, if a pixel stores, say, the value 25, the meaning is to go to row 25 in a color lookup table (LUT).
- For an 8-bit image, the image file can store in the file header information just what 8-bit values for R, G, and B correspond to each index.
- The LUT is often called a *palette.*

Color LUT for 8-bit color images

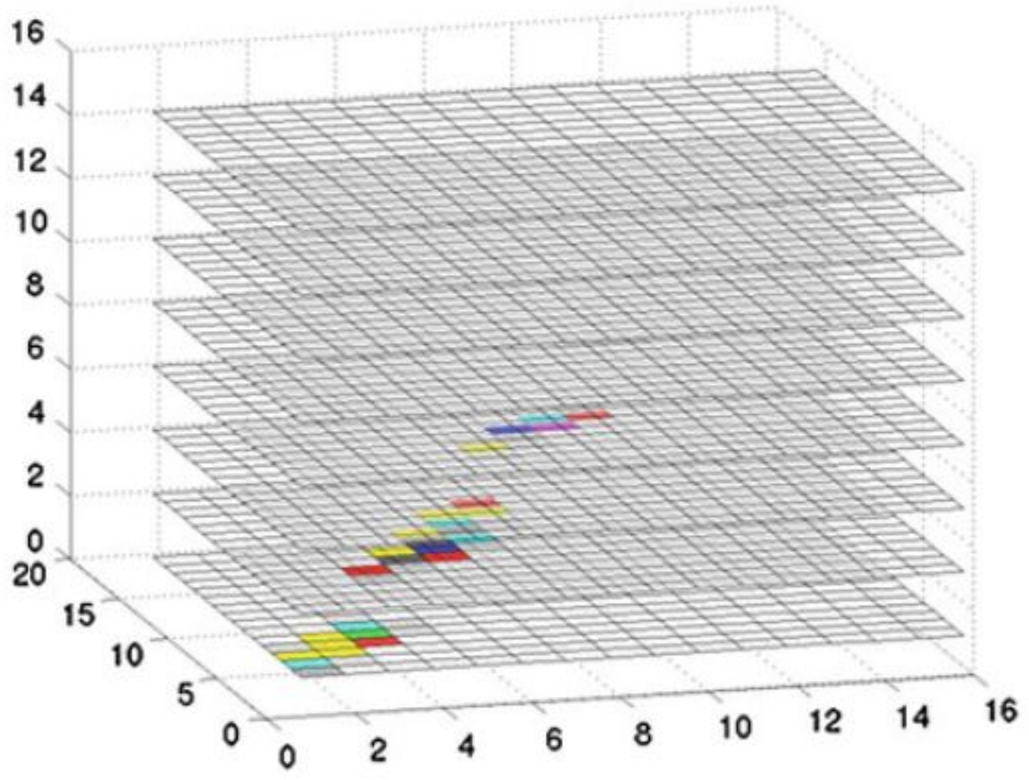


How to devise a color lookup table:
Color Histogram

- It makes sense to carefully choose just which colors to represent best in the image.
  - if an image is mostly red sunset, it is reasonable to represent red with precision and store only a few greens.
- Suppose all the colors in a 24-bit image were collected in a 256 × 256 × 256 set of cells, along with the count of how many pixels belong to each of these colors stored in that cell.
  - For example, if exactly 23 pixels have RGB values (45, 200, 91) then store the value 23 in a three-dimensional array, at the element indexed by the index values [45, 200, 91]
- This data structure is called a *color histogram*.

3D scatterplot of RGB colors in forestfire image

- The histogram has 16 × 16 × 16 bins and shows the count in each bin in terms of intensity and pseudocolor.
- We can see a few important clusters of color information, corresponding to the reds, yellows, greens, and so on, of the forestfire image.
- Large populations in 3D histogram bins can be subjected to a split-and merge algorithm to determine the "best" 256 colors.



Example of an 8-bit color image

Note the great savings in space for 8-bit images over 24-bit ones: a 640 × 480 8-bit color image requires only 300kB of storage, compared to 921.6kB for a color image.

24-bit color image

8-bit color image

How to Transform to 8 bit RGB

How to Select the Best 256 LUT RGB Entries Without Constructing a Color Histogram

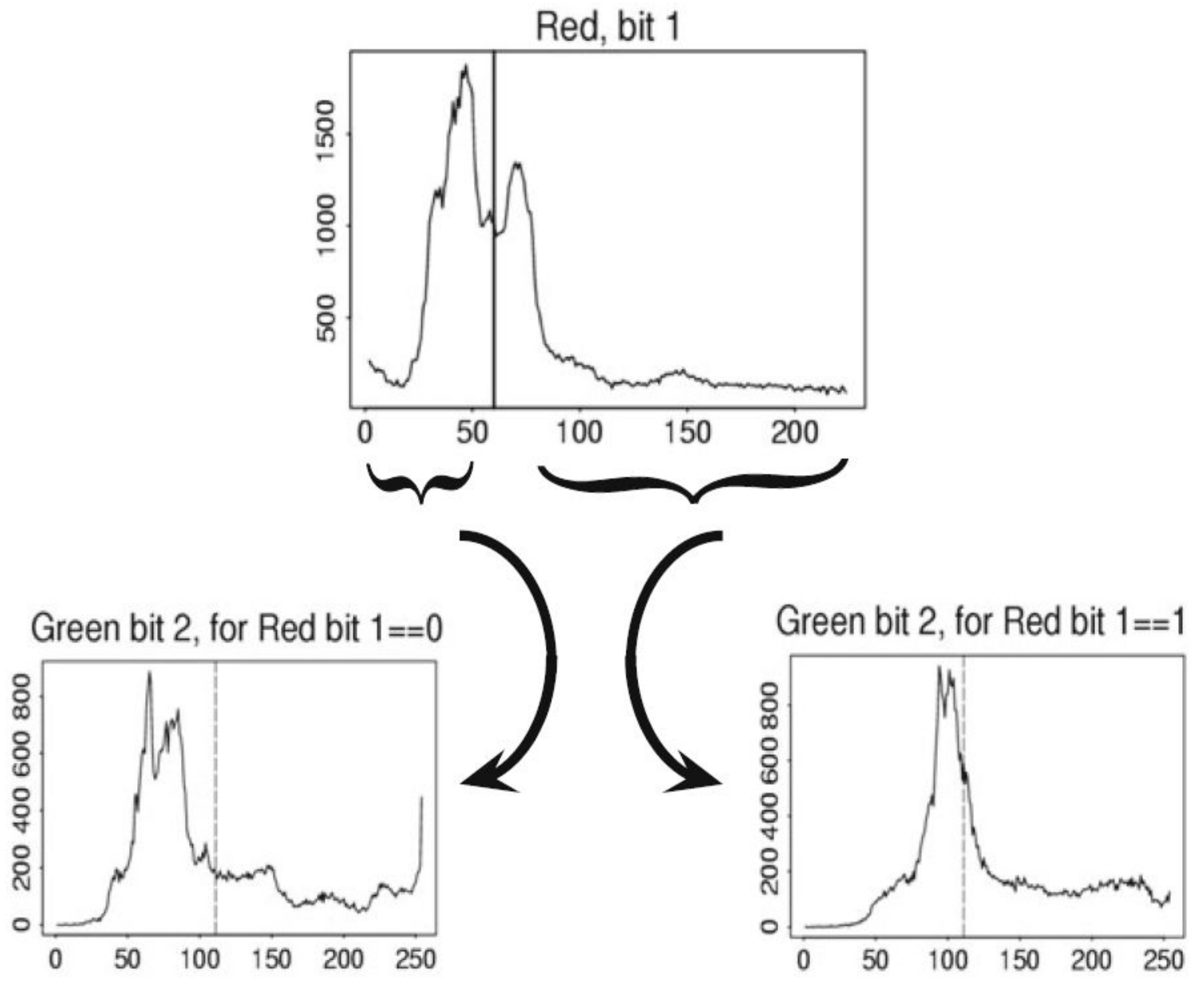
- Transform 3 bytes RGB to 8bit per pixel for 256 selective LUT table index by selecting RGB color separately.
- Sort R byte in the original image to select 8 (3bits) most popular R intensities.
  - Those 8 most popular R byte codes to LUT.
- Sort G byte to select 8 (3bits) most popular G intensities
  - Those 8 most popular G byte codes to LUT
- Sort B byte to select 4 (2bits) most popular B intensities.
  - Those 4 B byte codes to LUT
- All the combinations of 8R x 8G x 4B becomes 256 LUT RGB entries.
- To shrink R and G, we could simply divide the R or G byte value by (256/8 =) 32 and then truncate. Then each pixel in the image gets replaced by its 8-bit index, and the color LUT serves to generate 24-bit color.
- May result in edge artifacts.

Using median cut algorithm

Steps:

- Calculate the range of each color component (R, G, B).
- For the component with the largest range (let's say R), calculate the median value, M.
- Then values smaller than the median are labeled with a 0 bit and values larger than the median are labeled with a 1 bit.
- Next, we consider only pixels with a 0 label *from the first step* and sort their G values. Again, we label image pixels with another bit (and this is the second bit assigned), 0 for those less than the median in the greens and 1 for those greater.
- Now applying the same scheme to pixels that received a 1 bit for the red step, we have arrived at 2-bit labeling for all pixels.
- Carrying on to the blue channel, we have a 3-bit scheme.
- Repeating all steps, R, G, and B, results in a 6-bit scheme, and cycling through R and G once more results in 8 bits.
- These bits form our 8-bit color index value for pixels, and corresponding 24-bit colors can be the centers of the resulting small color cubes.

Histogram of R bytes for the 24-bit color image forestfire.bmp



Multimedia Systems Lecture – 9

By

Dr. Priyambada Subudhi
Assistant Professor
IIIT Sri City

Popular Image File Formats

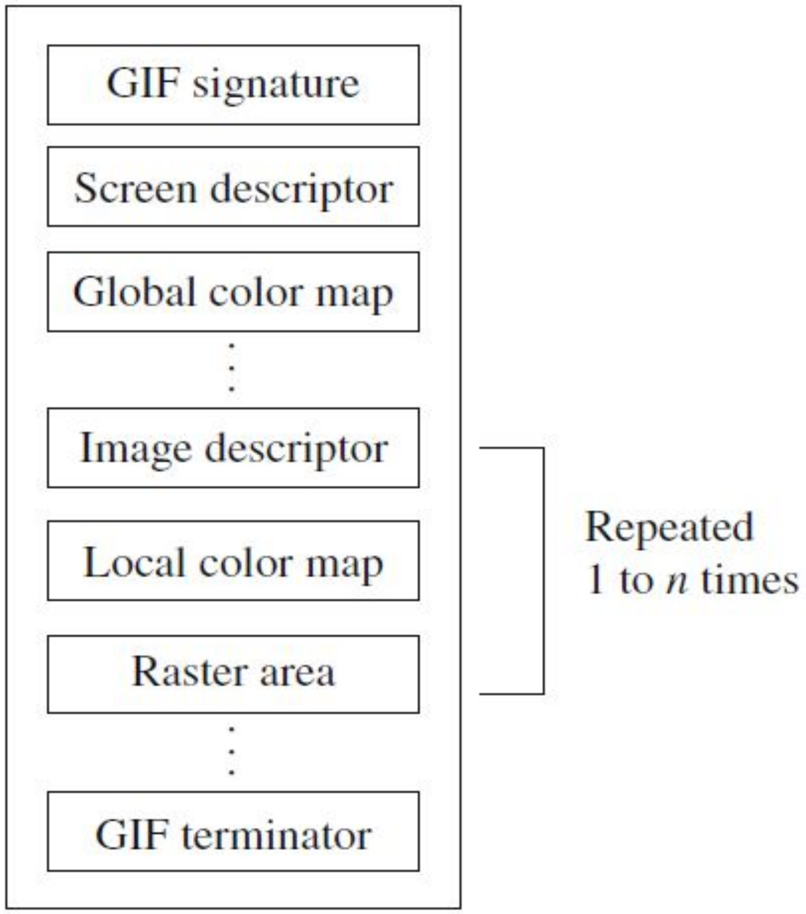
- 8-bit GIF: one of the most important formats because of its historical connection to the WWW and HTML markup language as the first image type recognized by net browsers.
- JPEG: currently the most important common file format.
- PNG: most popular lossless image format.
- TIFF: flexible file format due to the addition of tags.
- EXIF: allows the addition of image metadata.
- PS and PDF: vector based language, popular in publishing and academia

GIF (Graphic Interchange Format)

- GIF standard: We examine GIF standard because it is so simple! yet contains many common elements.
- Limited to 8-bit (256) color images only, which, while producing acceptable color images, is best suited for images with few distinctive colors (e.g., graphics or drawing).
- GIF standard supports interlacing — successive display of pixels in widely-spaced rows by a 4-pass display process.
- GIF actually comes in two flavors:
  1. GIF87a: The original specification.
  2. GIF89a: The later version. Supports simple animation via a Graphics Control Extension block in the data, provides simple control over delay time, a transparency index, etc.

GIF87

- Since many such formats bear a resemblance to it but have grown a good deal more complex than this “simple” standard, it is worth examining the file format for GIF87 in more detail.
- For the standard specification, the general file format is as shown in the figure.

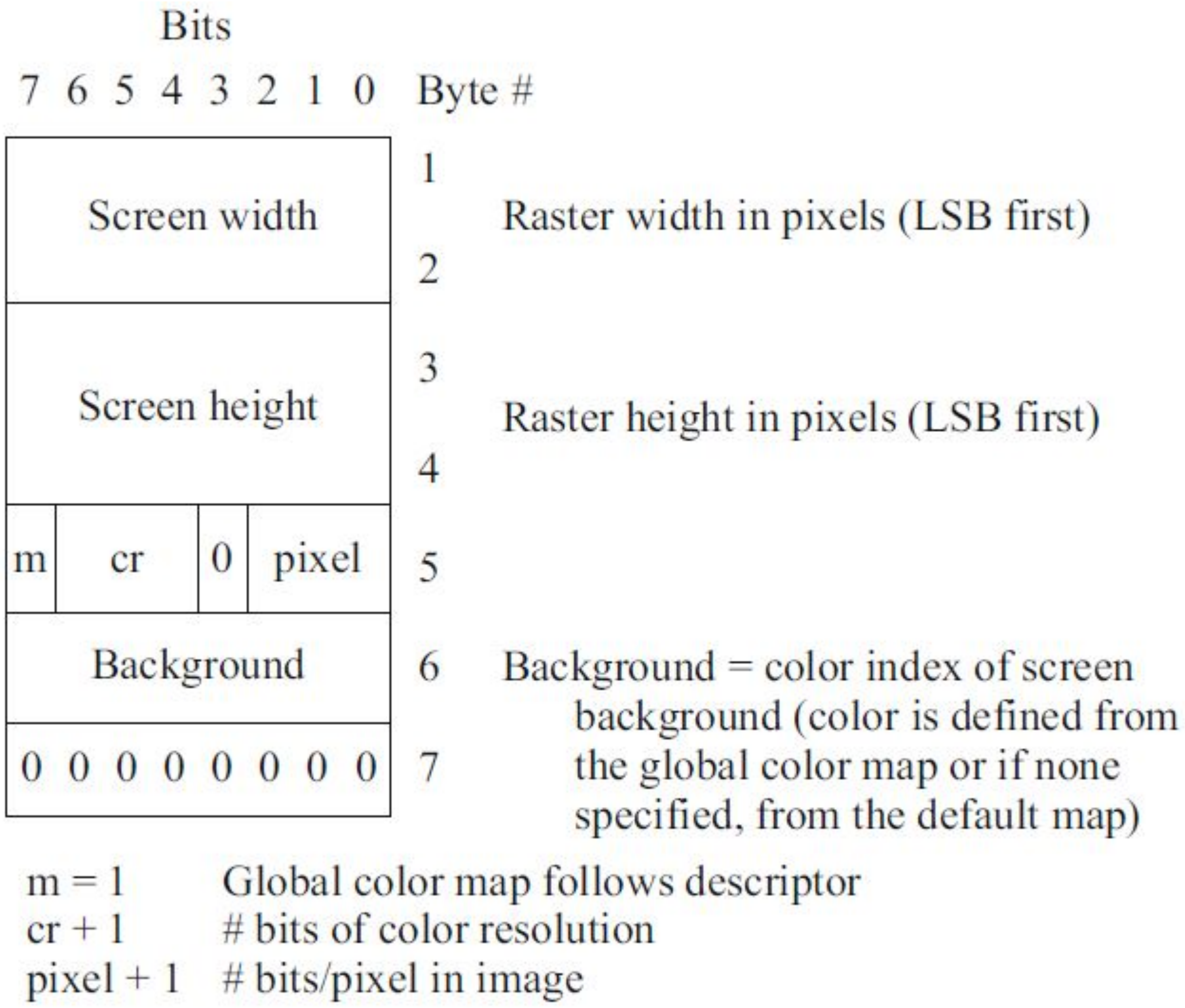


- The *Signature* is six bytes: GIF87a; the *Screen Descriptor* is a seven-byte set of flags.

- A GIF87 file can contain more than one image definition, usually to fit on several different parts of the screen.

- Therefore each image can contain its own color lookup table, a *Local Color Map*, for mapping 8 bits into 24-bit RGB values.

- The Screen Descriptor comprises a set of attributes that belong to every image in the file.



GIF Color Map

Bits 7 6 5 4 3 2 1 0	Byte #	
Red intensity	1	Red value for color index 0
Green intensity	2	Green value for color index 0
Blue intensity	3	Blue value for color index 0
Red intensity	4	Red value for color index 1
Green intensity	5	Green value for color index 1
Blue intensity	6	Blue value for color index 1
⋮		(continues for remaining colors)

GIF image descriptor

Bits 7 6 5 4 3 2 1 0	Byte #	
0 0 1 0 1 1 0 0	1	Image separator character (comma)
Image left	2 3	Start of image in pixels from the left side of the screen (LSB first)
Image top	4 5	Start of image in pixels from the top of the screen (LSB first)
Image width	6 7	Width of the image in pixels (LSB first)
Image height	8 9	Height of the image in pixels (LSB first)
m, i, 0, 0, 0, pixel	10	

m = 0	Use global color map, ignore 'pixel'
m = 1	Local color map follows, use 'pixel'
i = 0	Image formatted in Sequential order
i = 1	Image formatted in Interlaced order
pixel + 1	# bits per pixel for this image

Raster Area:

- The format of the actual image is defined as the series of pixel color index values that make up the image.
- The pixels are stored left to right sequentially for an image row.
- By default each image row is written sequentially, top to bottom.
- In the case that the Interlace or ‘i' bit is set in byte 10 of the Image Descriptor then the row order of the image display follows a four-pass process in which the image is filled in by widely spaced rows.
- The first pass writes every 8th row, starting with the top row of the image window. The second pass writes every 8th row starting at the fifth row from the top. The third pass writes every 4th row starting at the third row from the top. The fourth pass completes the image, writing every other row, starting at the second row from the top.

GIF four-pass interlace display row order

Image row	Pass 1	Pass 2	Pass 3	Pass 4	Result
0	*1a*				*1a*
1				*4a*	*4a*
2			*3a*		*3a*
3				*4b*	*4b*
4		*2a*			*2a*
5				*4c*	*4c*
6			*3b*		*3b*
7				*4d*	*4d*
8	*1b*				*1b*
9				*4e*	*4e*
10			*3c*		*3c*
11				*4f*	*4f*
12		*2b*			*2b*
⋮					

GIF Terminator

In order to provide a synchronization for the termination of a GIF image file, a GIF decoder will process the end of GIF mode when the character 0x3B hex or ';' is found after an image has been processed.

Multimedia Systems
Lecture – 10

By

Dr. Priyambada Subudhi

Assistant Professor

IIIT Sri City

JPEG

- The most important current standard for image compression is JPEG.
- This standard was created by a working group of the International Organization for Standardization (ISO) that was informally called the ***Joint Photographic Experts Group*** and is therefore so named.
- The human vision system has some specific limitations, which JPEG takes advantage of to achieve high rates of compression.
- The eye–brain system cannot see extremely fine detail.
- If many changes occur within a few pixels, we refer to that image segment as having *high **spatial frequency*** —that is, a great deal of change in (*x*, *y*) space.

- Therefore, color information in JPEG is decimated and then small blocks of an image are represented in the spatial frequency domain (u, v), rather than in (x, y).
- That is, the speed of changes in x and y is evaluated, from low to high, and a new "image" is formed by grouping the coefficients or weights of these speeds.
- Weights that correspond to slow changes are then favored, using a simple trick.
- Since we effectively throw away a lot of information by the division and truncation step, this compression scheme is "*lossy*"
- JPEG allows the user to set a desired *level of quality*, or *compression ratio* (input divided by output).

JPEG image with low quality specified by user.

This image is having a quality factor $Q = 10$. (The usual default quality factor is $Q = 75$). This image is a mere 1.5% of the original size. In comparison, a JPEG image with $Q = 75$ yields an image size 5.6% of the original, whereas a GIF version of this image compresses down to 23.0% of the uncompressed image size.

PNG

- **PNG format**: standing for ***Portable Network Graphics*** — system-independent image formats, meant to supersede the GIF standard, and extends it in important ways.
- Special features of PNG files include:
  - Support for up to ***16 bits*** per pixel in each color channel, i.e., 48-bit color.
  - Files may contain ***gamma-correction*** information for correct display of color images, as well as ***alpha-channel*** information for such uses as control of transparency.
  - The display progressively displays pixels in a 2-dimensional fashion by showing a few pixels at a time over ***seven*** passes through each ***8 x 8*** block of an image.
  - It supports both ***lossless*** and ***lossy*** compression with performance better than GIF.
  - PNG is widely supported by various web browsers and imaging software.

TIFF

- **TIFF**: stands for *Tagged Image File Format* is another popular image file format.
- The support for attachment of additional information (referred to as "tags") provides a great deal of flexibility.
- The most important tag is a ***format signifier***: what type of compression etc. is in use in the stored image.
- TIFF can store many different types of image: ***1-bit, grayscale, 8-bit color, 24-bit RGB***, etc.
- TIFF was originally a lossless format but now a JPEG tag allows one to opt for JPEG compression.

EXIF

- **EXIF** (*Exchange Image File*) is an image format for digital cameras.
- It enables the recording of image metadata (exposure, light source/flash, white balance, type of scene, etc.) for the standardization of image exchange.
- A variety of tags (many more than in TIFF) is available to facilitate higher quality printing, since information about the camera and picture-taking conditions can be stored and used, e.g., by printers for possible color-correction algorithms.

Windows BMP

- *BitMap* (BMP) is one major system standard image file format for Microsoft Windows.
- BMP supports many pixel formats, including indexed color (up to 8 bits per pixel), and 16, 24, and 32-bit color images.
- It makes use of *Run-Length Encoding (RLE)* compression.

Some Other Image Formats

- **Microsoft Windows WMF** (*Windows MetaFile*): the native vector file format for the Microsoft Windows operating environment:
- Consist of a collection of ***GDI (Graphics Device Interface)*** function calls, also native to the Windows environment.
- When a WMF file is "played" (typically using the Windows PlayMetaFile() function) the described graphics is rendered.
- WMF files are ***device-independent*** and are ***unlimited in size***.
- **Netpbm Format**: PPM (Portable PixMap), PGM (Portable GrayMap), and PBM (Portable BitMap) belong to a family of open source Netpbm formats. These formats are mostly common in the ***linux/unix environments***.

Multimedia Systems
Lecture – 11

By

Dr. Priyambada Subudhi

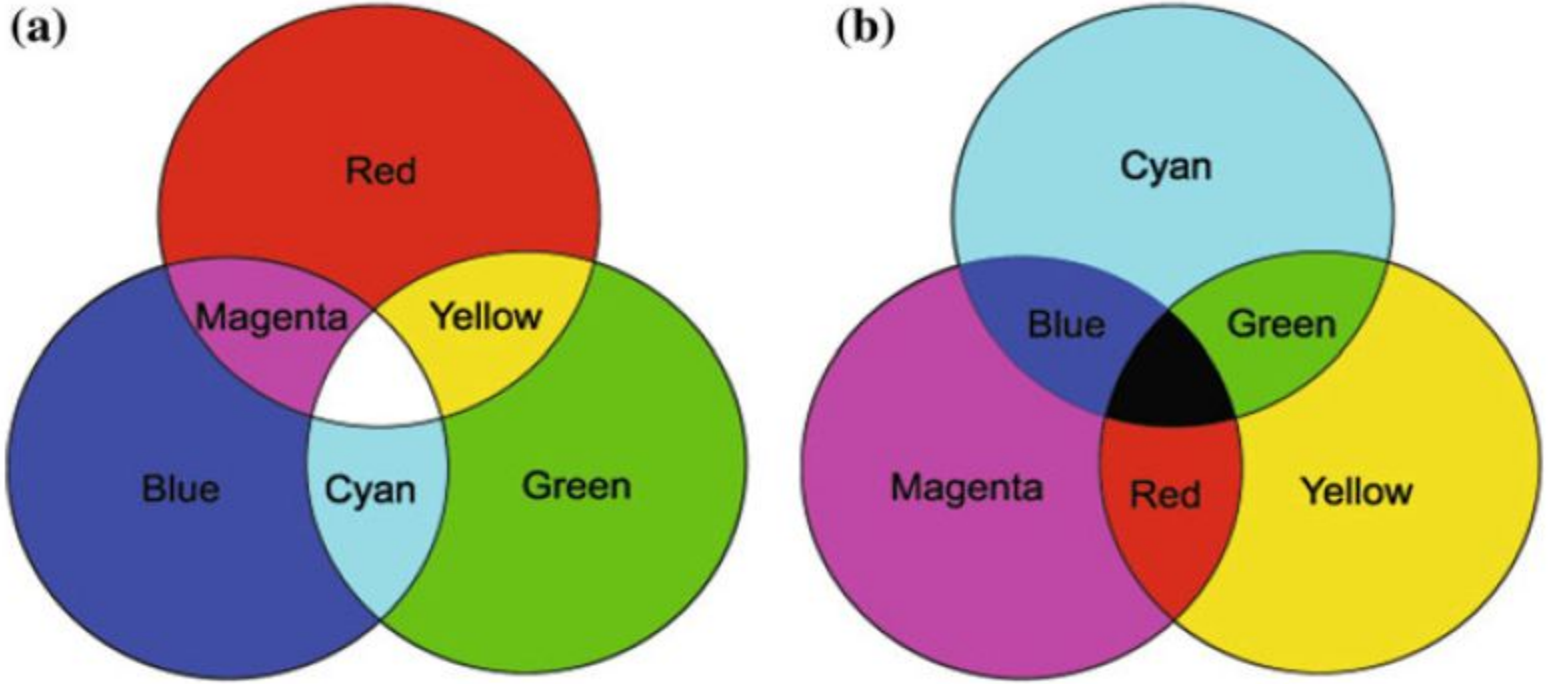
Assistant Professor

IIIT Sri City

Color Model

- A color model is an orderly system for creating a whole range of colors from a small set of ***primary colors*** *(are the set of colors that can be combined to make an useful range of colors)*.
- *Color Gamut*: Set of all colors that we can be produced from the primary colors.
- There are two types of color models.
  - Additive color model (e.g. RGB color model)
  - Subtractive color model (e.g. CMY color model)
- Additive color models use light to display color while subtractive models use printing inks.
- Colors perceived in additive models are the result of transmitted light while the colors perceived in subtractive models are the result of reflected light.

- There are several established color models used in computer graphics, but the two most common are the *RGB model* (***Red-Green-Blue***) for computer display and the *CMYK model* (***Cyan-Magenta-Yellow-Black***) for printing.

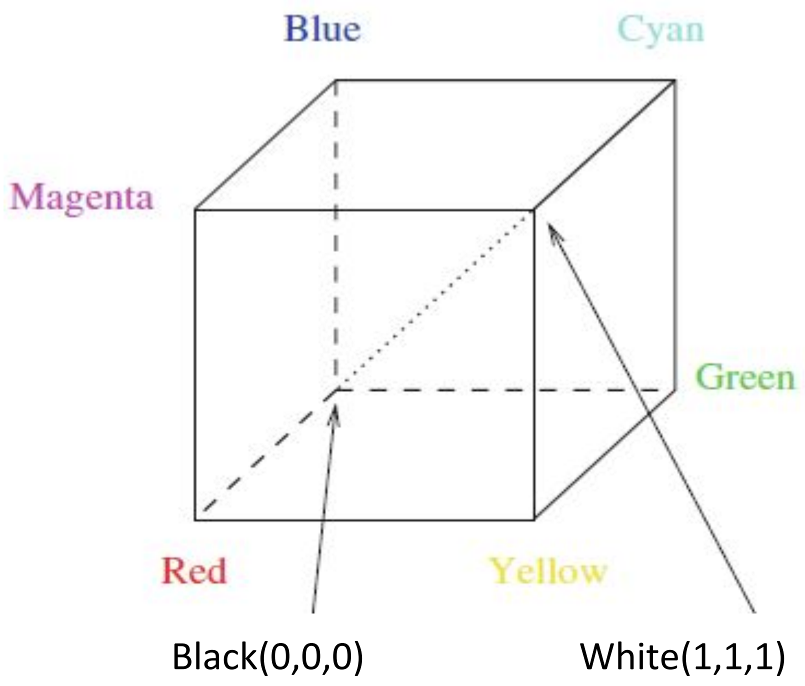


Additive and subtractive color. **a** RGB is used to specify additive color. **b** CMY is usedto specify subtractive color

RGB Model

- The red, green, and blue (RGB) color space is widely used throughout computer graphics.
- Unit Cube defined on R, G & B axes.
- The Origin **(0,0,0)** represents black and the diagonally opposite vertex **(1,1,1)** is White.
- Vertices of the cube on the axes represent primary colors, and the remaining vertices are the complementary color points for each of the primary colors.
- Shades of gray are represented along the main diagonal.

RGB color Cube



- Each color point within the unit cube can be represented as w weighted vector sum of the primary colors, using unit vectors **R**,**G** and **B**.

$$C(\lambda) = R\mathbf{R} + G\mathbf{G} + B\mathbf{B}$$

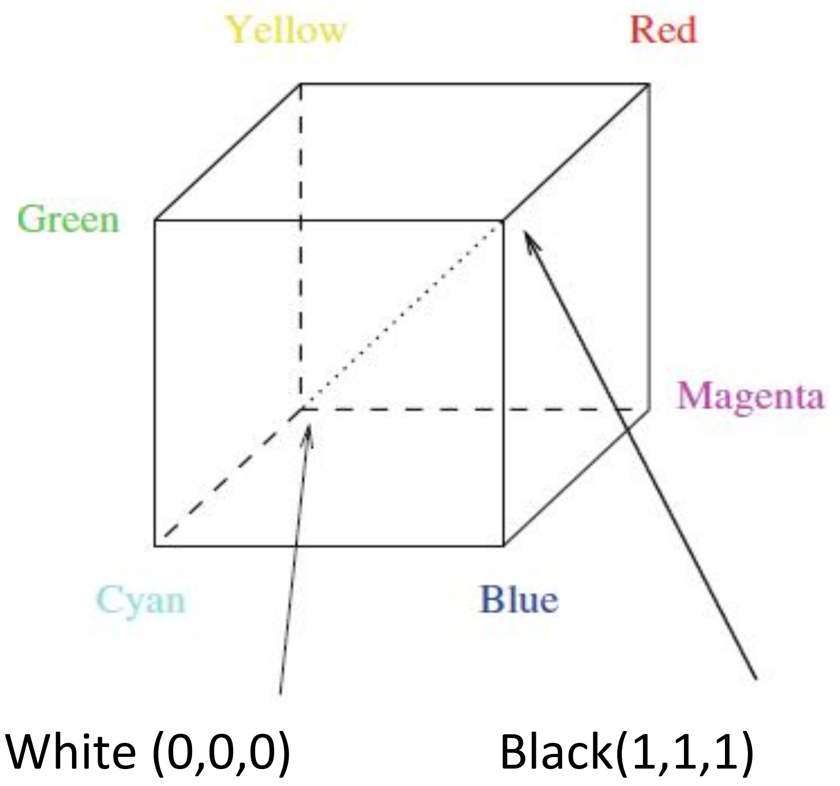
Where R, G and B are the assigned values in the range from 0 to 1.0

- For example, the magenta vertex is obtained by adding the maximum red and blue values to produce : (1,0,1)

CMY and CMYK model

- Subtractive Color Model.
- Stands for cyan-magenta-yellow.
- Used for hardcopy devices (ex. Printers).
- A printed color that looks red absorbs the other two components G and B and reflects R.
- Thus the C-M-Y coordinates are just the complements of the R-G-B coordinates.

CMY color Cube



- In additive color models such as RGB, white is the "additive" combination of all primary colored lights, while black is the absence of light.
- In the CMY model, it is the opposite: white is the natural color of the paper or other background, while black results from a full combination of colored inks.
- **Transformation from RGB to CMY and CMY to RGB**

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} C \\ M \\ Y \end{bmatrix}$$

- **[*Complementary colors*:** *Pairs of colors which, when combined in the right proportions, produce white. Example, in the RGB model: red & cyan , green & magenta , blue & yellow.]*

CMYK Model

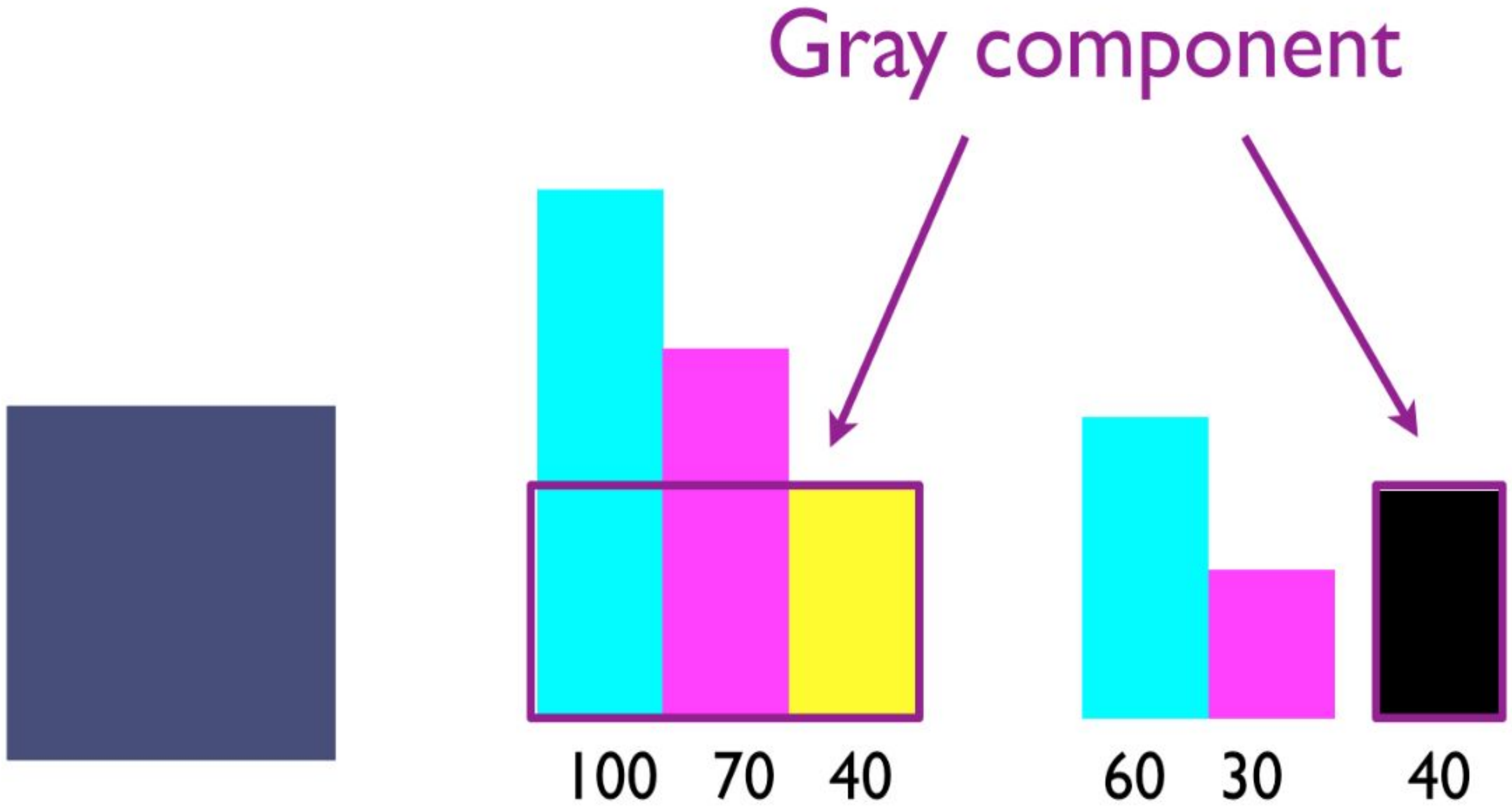
- Although cyan, magenta and yellow inks might be expected be sufficient for color printing, most actual color printing uses black ink in addition.
- This is partly because a mixture of the first three inks may not yield a black that is neutral enough, or dark enough.
- The use of black spares the use of the more expensive colored inks, and also reduces the total amount of ink used, thus speeding drying times.
- K used instead of equal amounts of CMY

- Truly “black” black ink is in fact cheaper than mixing colored inks to make black, so a simple approach to produce sharper printer colors is to calculate that part of the three-color mix that would be black, remove it from the color proportions, and add it back as real black. This is called “under color removal.”
- With *K* representing the amount of black, the new specification of inks is thus

$$K \equiv \min\{C, M, Y\}$$

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} \Rightarrow \begin{bmatrix} C - K \\ M - K \\ Y - K \end{bmatrix}$$

Under Color Removal



Multimedia Systems
Lecture – 12

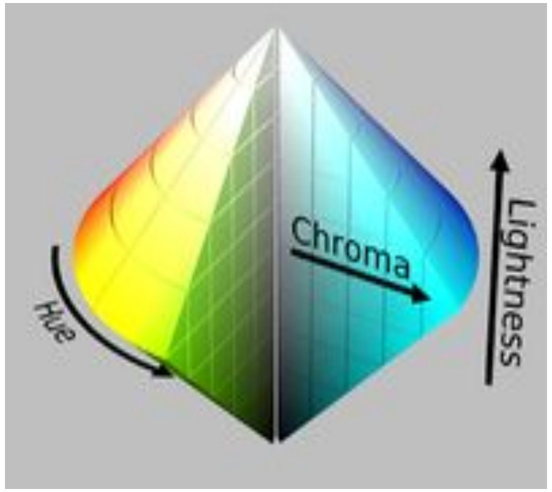
By

Dr. Priyambada Subudhi

Assistant Professor

IIIT Sri City

HSV model



- It generalizes how humans perceive color.
- How humans perceive colors is not like how RGB or CMYK make colors.
- The H stands for Hue, S stands for Saturation, and the V stand for value.
- Imagine a cone with a spectrum of red to blue from left to right, and from the centre to the edge, the color intensity increases.
- From bottom to up, the brightness increases. Hence resulting white at the center up layer.

- **Hue**: The hue represents the color. The hue value ranges from o to 360 degrees.

- **Saturation**: The saturation value tells us how much quantity of respective color must be added.
- A 100% saturation means that complete pure color is added, while a 0% saturation means no color is added, resulting in grayscale.

- **Value**: The value represents the brightness concerning the saturation of the color. The value 0 represents total black darkness, while the value 255 will mean a full brightness and depend on the saturation.

RGB to HSV

$$M = \max\{R, G, B\}$$

$$m = \min\{R, G, B\}$$

$$V = M$$

$$S = \begin{cases} 0 & \text{if } V = 0 \\ \big(V - m\big)/V & \text{if } V > 0 \end{cases}$$

$$H = \begin{cases} 0 & \text{if } S = 0 \\ 60(G - B)/(M - m) & \text{if } (M = R \text{ and } G \geq B) \\ 60(G - B)/(M - m) + 360 & \text{if } (M = R \text{ and } G < B) \\ 60(B - R)/(M - m) + 120 & \text{if } M = G \\ 60(R - G)/(M - m) + 240 & \text{if } M = B \end{cases}$$

HSV to RGB

- Given the values of H, S, and V, you can first compute m and M with the equations

  $M = V$

  $m = M(1-S)$.

- Now compute another number, z, defined by the equation

  $z = (M-m)[1 - |(H/60)mod\_2 - 1|]$,

where mod_2 means division modulo 2.

- Now you can compute R, G, and B according to the angle measure of H.
- There are six cases.

- When $0 \le H < 60$,
  $R = M$
  $G = z + m$
  $B = m$.

- If $60 \le H < 120$,
  $R = z + m$
  $G = M$
  $B = m$.

- If $120 \le H < 180$,
  $R = m$
  $G = M$
  $B = z + m$.

- When $180 \le H < 240$,
  $R = m$
  $G = z + m$
  $B = M$.

- When $240 \le H < 300$,
  $R = z + m$
  $G = m$
  $B = M$.

- if $300 \le H < 360$,
  $R = M$
  $G = m$
  $B = z + m$.

Multimedia Systems Lecture – 13

By

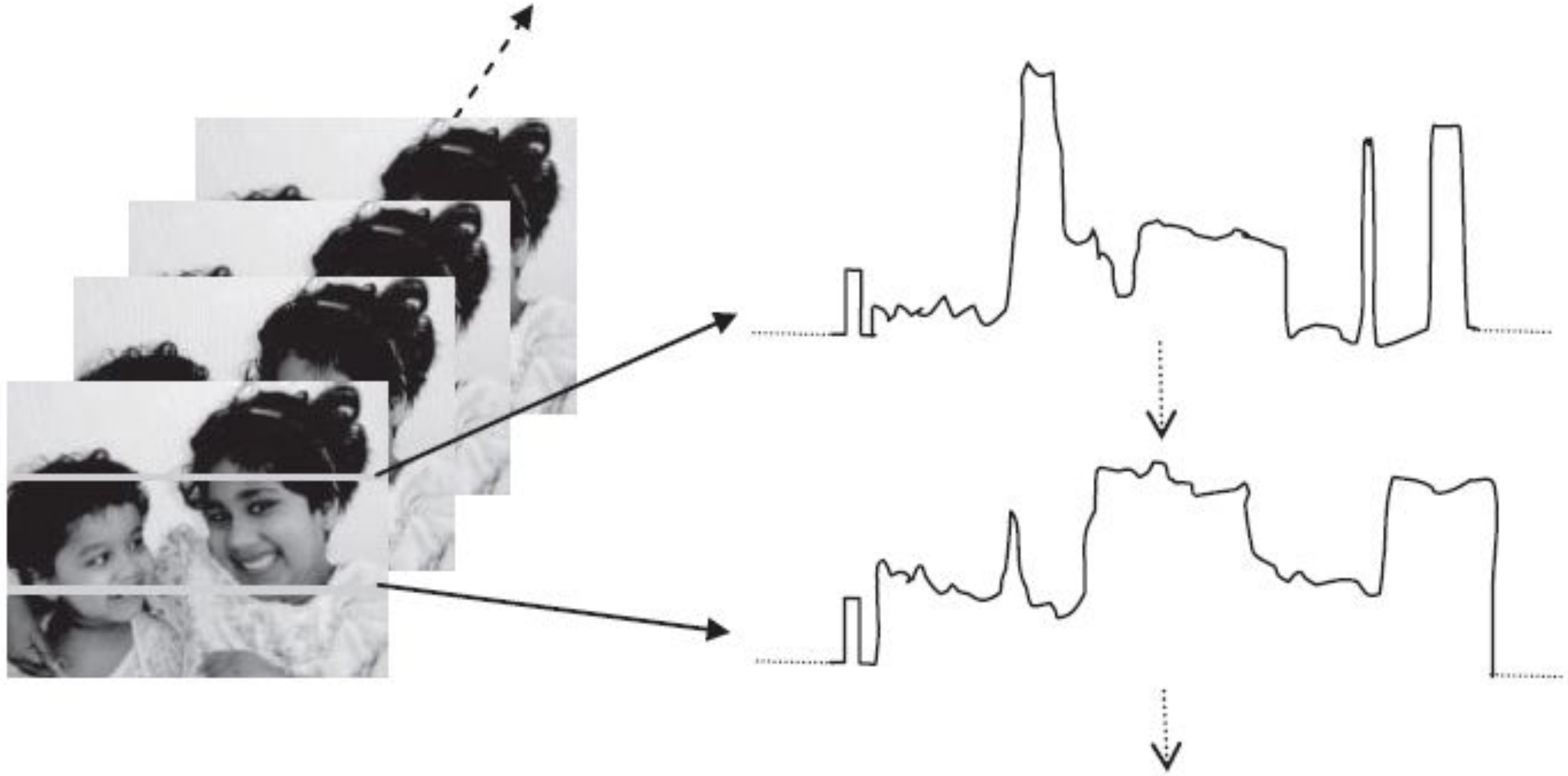
Dr. Priyambada Subudhi
Assistant Professor
IIIT Sri City

Video

- Video, whether analog or digital, is represented by a sequence of discrete images shown in quick succession. Each image in the video is called a frame, which is represented as a matrix of pixels defined by a width, height, and pixel depth.
- In addition, two important properties govern video representation: frame rate and scanning format.
- The rate at which the images are shown is the frame rate.
- If the frame rate is too slow, the human eye perceives an unevenness of motion called flicker.

- Although digital video can be considered a three-dimensional signal—a 2D image changing over time—analog video is converted to a 1D signal of scan lines.
- This scan line conversion was introduced to make analog television broadcast technology work, and is central to the manner in which televisions (and all other cathode-ray tubes) display images.
- The electron gun(s) in a television project electrons on the phosphor screen from left to right in a scan line manner and from top to bottom successively for each frame.
- The phosphor screen glows at each location on a scan line creating a color at all positions on the line.
- Scanning formats, which is an outcome of the analog technology, can be represented as interlaced or progressive.

- *Left: Video is represented as a sequence of images. Right: Analog video of one frame scanned as a 1D signal. Each scan line is scanned from left to right as an analog signal separated by horizontal syncs. Two scan lines are shown; each begins with a horizontal sync and traces through the intensity variation on that scan line.*

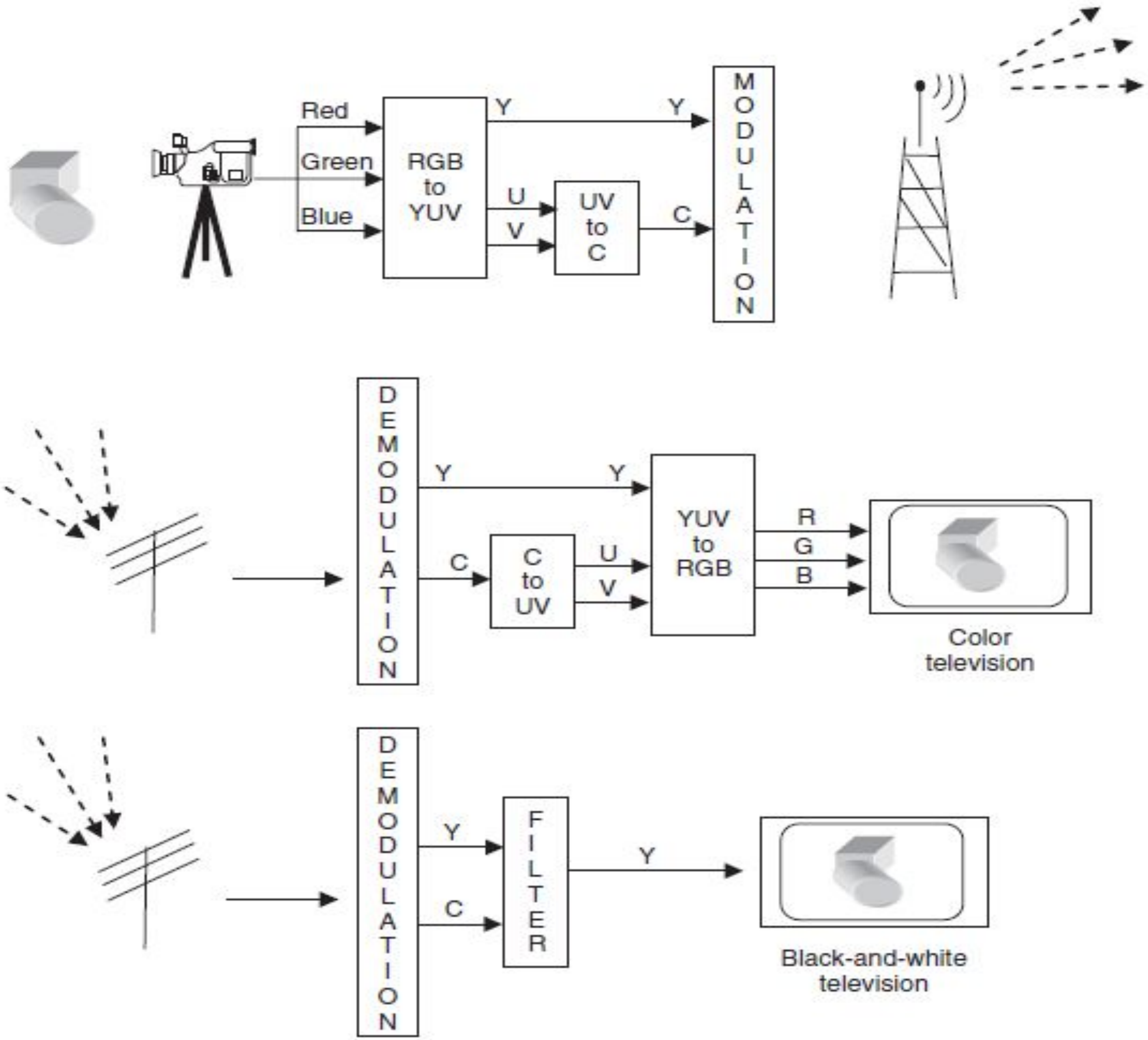


Digital display technologies display media in a digital format. Digital video display on these devices, such as LCD or plasma, does not require the scanning mechanism described previously. However, when the technology for digital video started to evolve, the television instruments were still rendering analog signals only. As a result, the digital video standards have their representations and formats closely tied to analog TV standards.

Analog Video and Television

- Analog video signal used in broadcast is scanned as a one-dimensional signal in time, where the spatiotemporal information is ordered as a function of time according to a predefined scanning convention.
- This 1D signal captures the time-varying image intensity information only along scanned lines.
- Television requires this analog scanned information to be broadcast from a broadcast station to all users.
- The standardization process implemented in the broadcast of analog video for television mandated a few requirements, which were necessary for making television transmission viable: YUV color space conversion and interlaced scanning.

Television works by sending scan line information in interlaced YUV format.



Conversion to YUV

- Video frames, like images, are represented using a color format, which is normally RGB. This RGB color space is used by cathode-ray tube–based display devices, such as the television, to display and render the video signal.

- For transmission purposes, however, the RGB signal is transformed into a YUV signal. The YUV color space aims to decouple the intensity information (Y or luminance) from the color information (UV or chrominance).

- The separation was intended to reduce the transmission bandwidth and is based on experiments with the human visual system, which suggests that humans are more tolerant to color distortions than to intensity distortions.

RGB to YUV

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Ranges:
R/G/B [0 ... 1]
Y [0 ... 1]
U [-0.436 ... +0.436]
V [-0.615 ... +0.615]

RGB to YUV color conversion for analog TV

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Ranges:
R/G/B [0 ... 255]
Y [16 ... 235]
Cb/Cr [16 ... 240]

RGB to YCbCr color conversion for SDTV

Analog Video Scanning

- Video is scanned as a 1D signal, where each raster line is interspaced with horizontal and vertical syncs.

- For horizontal synchronization in analog video, a small voltage offset from zero is used to indicate black and another value, such as zero, to indicate the start of a line.

- Vertical synchronization is carried out by the cycles in the power outlet (60 Hz for NTSC, 50 Hz for PAL). Every 1/60th of a second, the electron gun is reset by the vertical sync to draw the beginning of the next frame.

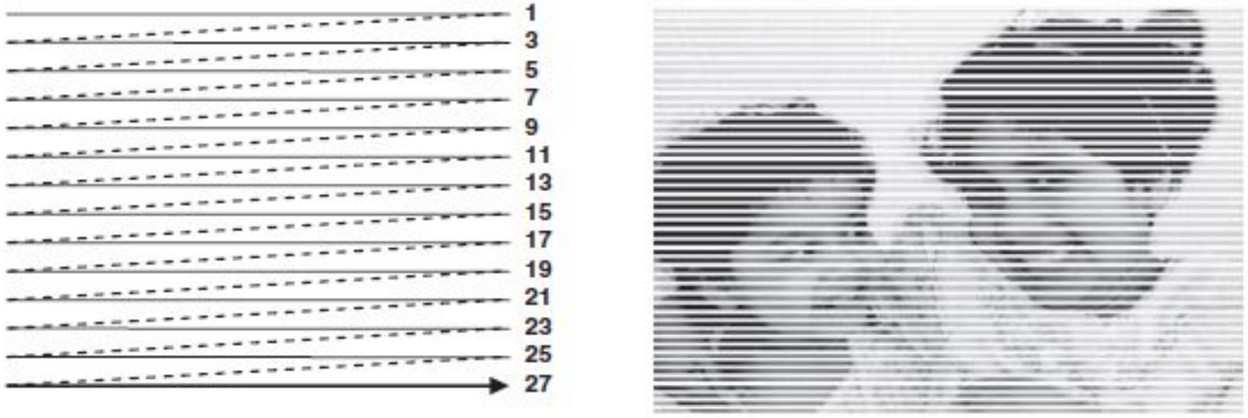
- In TV and in some monitors and multimedia standards, another system, *interlaced* scanning, is used.

- Here, the odd-numbered lines are traced first, then the even-numbered lines. This results in "odd" and "even" *fields*—two fields make up one frame.

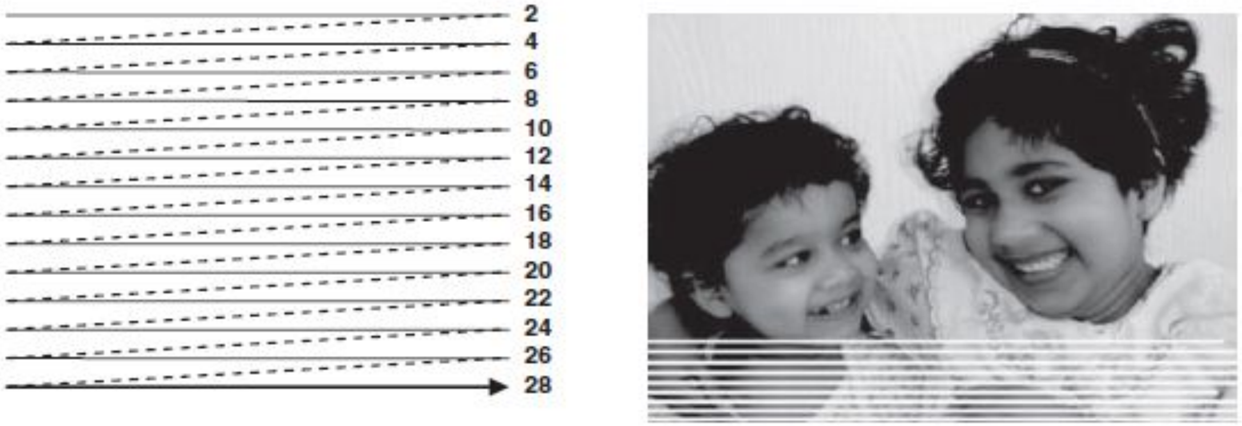
- But the resulting video drawn by interlaced scanning techniques might be unacceptable and has occasional flicker and artifacts. This is caused because the video is captured at different moments in time as two field and, hence, interlaced video frames exhibit motion artifacts when both fields are combined and displayed at the same moment.

- Video is of better quality when it is captured progressively and drawn progressively, which eliminates the occasional flicker.

- *Interlaced scanning. The top figure shows the upper “odd” field consisting of odd-numbered lines. The bottom shows a lower “even” field interspersed with the odd field. Both fields are shown in succession to meet the required frame rate.*

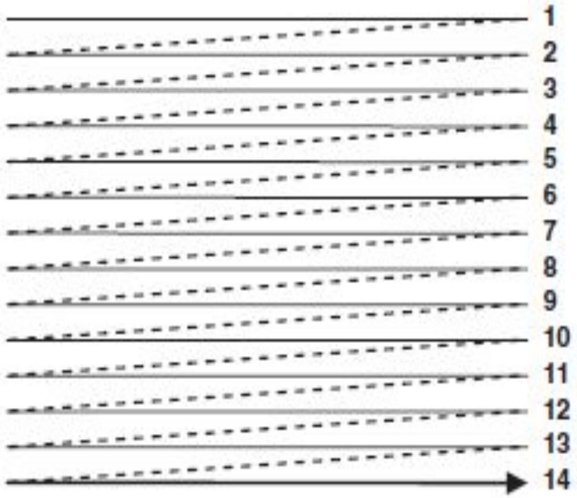


Upper field



Lower field

Progressive scanning. All the scan lines are drawn in succession, unlike in the interlaced case.



Analog Video Standards

NTSC Video

- ***NTSC***, named for the ***National Television System Committee***, is the analog television system that is mostly used in most of North America and Japan.
- It uses a familiar 4:3 *aspect ratio* (i.e., the ratio of picture width to height) and 525 scan lines per frame at 30 fps.
- NTSC video is an analog signal with no fixed horizontal resolution. Therefore, we must decide how many times to sample the signal for display. Each sample corresponds to one pixel output. A *pixel clock* divides each horizontal line of video into samples. The higher the frequency of the pixel clock, the more samples per line.

PAL Video

- *PAL (Phase Alternating Line)* is a TV standard originally invented by German scientists.
- It uses 625 scan lines per frame, at 25 frames per second (or 40 msec / frame), with a 4 : 3 aspect ratio and interlaced fields.
- This important standard is widely used in Western Europe, China, India and many other parts of the world.
- PAL uses the YUV color model with an 8 MHz channel, allocating a bandwidth of 5.5 MHz to Y and 1.8 MHz each to U and V.
- To improve picture quality, chroma signals have alternate signs (e.g., +U and — U) in successive scan lines; hence the name "Phase Alternating Line.
- This facilitates the use of a (line - rate) comb filter at the receiver — the signals in consecutive lines are averaged so as to cancel the chroma signals (which always carry opposite signs) for separating Y and C and obtain high - quality Y signals.

SECAM Video

- SECAM, which was invented by the French, is the third major broadcast TV standard.
- SECAM stands for Systeme Electronique Couleur Avec Memorie.
- SECAM also uses 625 scan lines per frame, at 25 frames per second, with a 4:3 aspect ratio and interlaced fields.
- SECAM and PAL are similar, differing slightly in their color coding scheme.
- In SECAM, U and V signals are modulated using separate color subcarriers at 4.25 MHz and 4.41 MHz, respectively. They are sent in alternate lines - that is, only one of the U or V signals will be sent on each scan line.

Multimedia Systems
Lecture – 14

By

Dr. Priyambada Subudhi
Assistant Professor
IIIT Sri City

Types of Video Signals

Composite Video

- Composite video is also called baseband video or RCA video. It is the analog waveform that conveys the image data in the conventional NTSC, PAL and SECAM television signal.
- Composite video contains both chrominance (color) and luminance (brightness) information, along with synchronization and blanking pulses, all together in a single signal.
- This was done to reduce bandwidth and achieve real-time transmission.
- However, in composite video, interference between the chrominance and luminance information is inevitable and tends to worsen when the signal is weak.
- This is why fluctuating colors, false colors, and intensity variations are seen when a distant NTSC television station sends signals that are weak and not properly captured at home with old-fashioned "rabbit ears," or outdoor "aerial" antennae.

S-Video

- S-Video (*Super-Video*, sometimes referred to as *Y/C Video*) is a video signal transmission in which the luminance signal and the chrominance signal are transmitted separately to achieve superior picture clarity.
- The luminance signal (*Y*) carries brightness information, and the chrominance signal (*C*) carries color information.
- Here, the chrominance signal (*C*) is formed by combining the two chrominance signals *U* and V into one signal along with their respective synchronization data, so at display time, the *C* signal can be separated into *U* and *V* signals.
- Separating the *Y* and *C* channels and sending them separately reduces problems caused by interference between the luminance and chrominance signals and yields a superior visual quality.

Component Video

- Component video strives to go a step further than S-Video by keeping all three *Y, U, V* (or equivalent) components separate.
- Consequently, the bandwidth required to broadcast component video is more than the composite or S-Video and, correspondingly, so is the visual quality.
- The separation of these components prevents artifacts due to intersignal interference.

Connectors for typical analog display interfaces. From left to right:, Composite video, S-video, and Component video

Digital Video

The advantages of digital representation for video are many. It permits

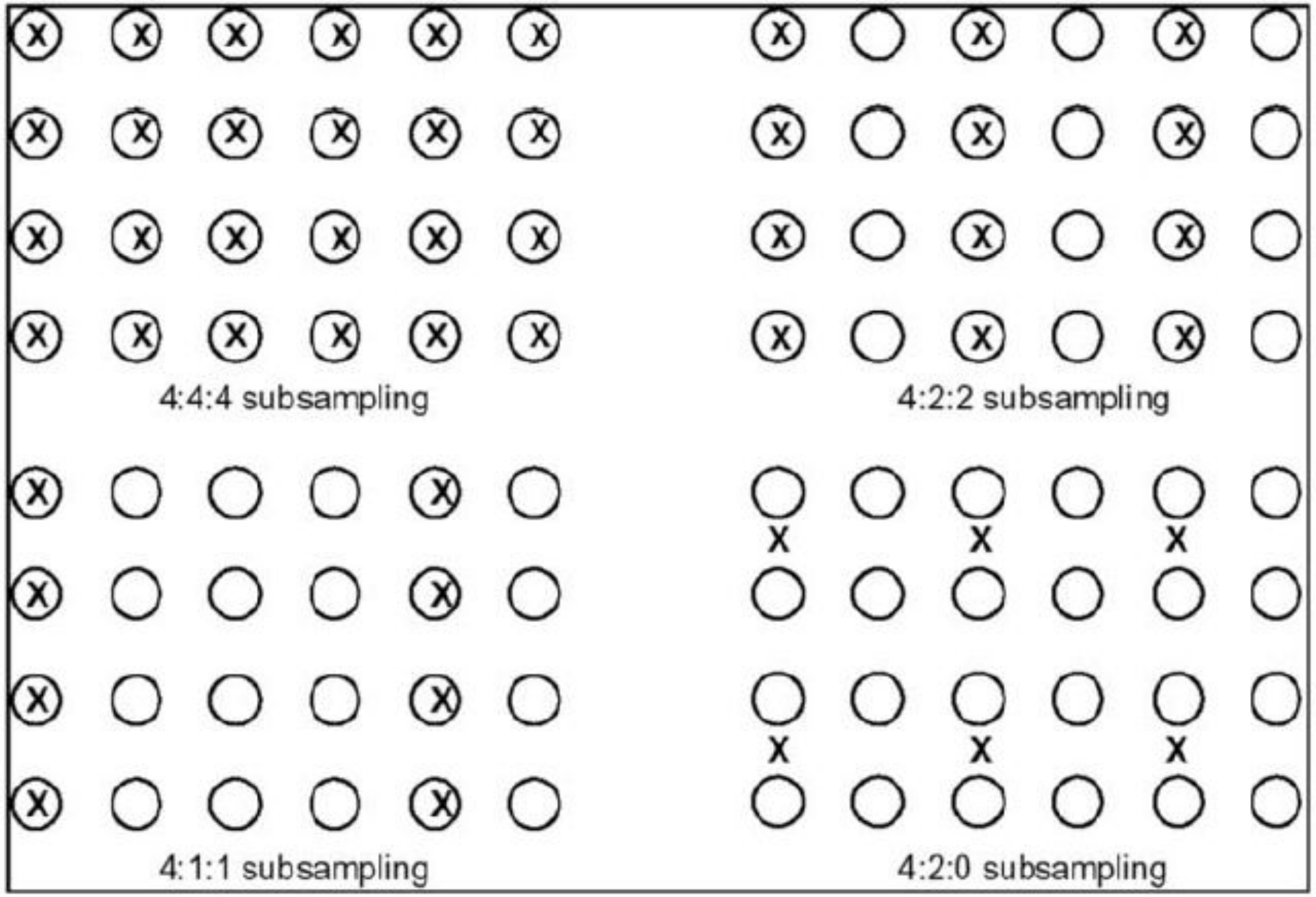
- Storing video on digital devices or in memory, ready to be processed (noise removal, cut and paste, and so on) and integrated into various multimedia applications.
- Direct access, which makes nonlinear video editing simple.
- Repeated recording without degradation of image quality.
- Ease of encryption and better tolerance to channel noise.

YUV Subsampling Schemes

- Video signals captured by digital cameras are represented in the RGB color space, which is also used to render video frames on a display device.
- However, for transmission and other intermediary processing, the YUV space is commonly used.
- The YUV space separates the color and luminance information.
- The color information (UV) is then further subsampled to gain more bandwidth.
- In analog video, subsampling is achieved by allocating half as much bandwidth to chrominance as to luminance.
- In digital video, subsampling can be done by reducing the number of bits used for the color channels on average.

- Depending on the way subsampling is done, a variety of subsampling ratios can be achieved.
- The circles represent pixel information.
- Potentially, we could store 1 byte each for Y, U, and V components, resulting in 24 bits per pixel.
- In subsampling, the luminance component Y is left untouched—that is, 1 byte is reserved for the luminance data per pixel.
- An X at a pixel position suggests that we also store the chrominance components for this position.

YUV subsampling schemes used in video



- In the 4:4:4 scheme, each pixel has luminance (8 bits) and chrominance (8 bits for U and 8 bits for V), resulting in 24 bits per pixel.
- In the 4:2:2 subsampling scheme, chrominance information is stored for every other pixel bringing the equivalent bits per pixel down to 16.
- In the 4:1:1 subsampling scheme, chrominance is stored every fourth pixel in a row.
- Whereas in the 4:2:0 scheme, the average of the U values for a 2×2 pixel area is stored, and similarly for the V values.
- Since there is only 1 U and 1 V sample for every four luminance samples, the equivalent bits per pixel is brought down to 12 bits per pixel.