

كلية التقنية الهندسية الكهربائية
قسم هندسة تقنيات الحاسوب

Multimedia Computing

حوسبة الوسائط المتعددة

مدرس المادة : علي نافع

Part one

1. Introduction to Multimedia

1.1 Introduction

The **media** in multimedia comes in different forms: **graphics, photography, text, audio, video, and animation**. The simplest explanation of multimedia is “**the combination of two or more media**”. However, multimedia is more complex than the term implies. It intertwines a myriad of media elements and, as a result, makes for a more comprehensive end product than when media components are experienced independently.

The kinds of media elements and how they are used play a significant role in multimedia environments. Consider a web site without any graphics, color, or images. How about a music video without sound? There is clearly something lacking. A combination of media adds richness and provides a complete sensory experience. As shown in Figure (1), multimedia is involved in media elements work together to make a stronger, more cohesive whole.

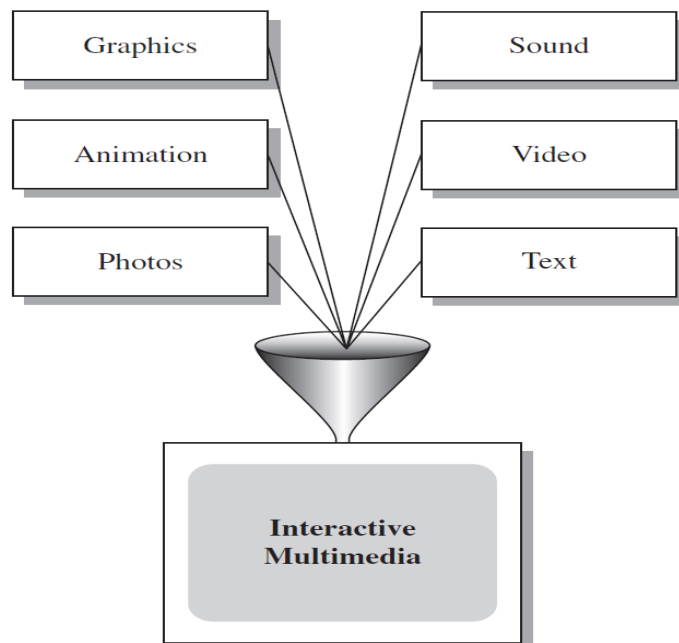


Figure 1: Multimedia Elements Funneling into One Unit

1.2 Multimedia Definition

Multimedia is the media that uses multiple forms of information content and information processing (e.g. **text, audio, graphics, animation, video, interactivity**) to inform or entertain the user. Multimedia also refers to the use of **electronic media** to store and experience multimedia content.

A Multimedia System is a system capable of processing multimedia data and applications. A Multimedia System is characterized by the processing, storage, generation and manipulation of Multimedia information.

A multimedia system consists of two types of forms of multimedia: **Non interactive and Interactive**

❖ **Non interactive:** user is an observer of information.

- User has no control over the flow of information.
- Developer establishes the sequence of media elements and how they are presented.

❖ **Interactive:** any computer-delivered electronic system that allows the user to control, combine, and manipulate different types of media, such as text, sound, video, computer graphics, and animation. **user is a participant in the flow of information.** The following are common types of interactive multimedia

1. **Basic interactivity**

Includes **menu** and **button** options to access content.

2. **Adaptive**

Adapt the information flow to the needs or interests of the users.

3. **Immersive**

Draws users into an alternate world.

Multimedia can be considered as an integrated if All multimedia components (audio, video, text, graphics) used in the system must be

somehow integrated. Every device, such as a microphone and a camera is connected to and controlled by a single computer. A single type of digital storage is used for all media types.

The terms **static and dynamic media** refer to the way in which a medium is presented. As shown in Figure (2).

- **Static media**, such as the newspaper is printed in black and white and cannot be altered once it has been written. The static data are those in which the temporal component is not present and do not need a synchronization among the data.
- **Dynamic media** is constantly updated and is interactive with player control. Dynamic media is developing in the modern world and the classic example is a website. In the dynamic data the temporal variable is present.



Static Media

1- Image

2- Text



Dynamic Media

1- Video

2- Animation

3- sound

Figure 2: Static and Dynamic Media

Media are divided into two types in respect to time in their representation space:

1. **Time independent (discrete):** Information is expressed only in its individual value. E.g.: text, image.
2. **Time dependent (continuous):** Information is expressed not only its individual value, but also by the time of its occurrences. E.g.: sound and video.

1.3 Classification of Media:

1. The perception media

Perception media help human to sense their environment. The central question is **how human perceive information in a computer environment**. The answer is through **seeing and hearing**.

Seeing: For the perception of information through seeing the usual such as text, image and video are used.

Hearing: For the perception of information through hearing media such as music noise and speech are used.

2. The representation Media

Representation media are defined by internal computer representation of information. The central question is how the computer information is coded? The answer is that various format are used to represent media information in computer.

- ❖ **Text, character** is coded in **ASCII** (American Standard Code for Information Interchange) code and **Unicode**.
- ❖ **Graphics** are coded according to **CEPT video** text standard.
- ❖ **Image** can be coded as **JPEG format**.
- ❖ **Audio video sequence** can be coded in different TV standard format in **MPEG**.

3. The Presentation Media

Presentation media refer to the tools and devices for the input and output of the information. The central question is, through which the

information is delivered by the computer and is introduced to the computer.

Output media: paper, screen and speaker are the output media.

Input Media: Keyboard, mouse, camera, microphone are the input media.

4. The **storage media**

Storage Media refer to the data carrier which enables storage of information. The central question is, how will information be stored? The answer is a **hard disk, CD-ROM**, etc.

5. The **transmission media**

Transmission Media are the different information carrier that enables continuous data transmission. The central question is, over which information will be transmitted? The answer is **co-axial cable, fiber optics as well as free air.**

6. The **information Exchange media**

Information exchange media includes all information carrier for transmission, i.e. all storage and transmission media. The central question is; which information carrier will be used for information exchange between different places? The answer is **combine uses of storage and transmission media.** E.g. **Electronic mailing system.**

1.4 Categories of Multimedia

Based on how multimedia programs are used, multimedia can be divided into two forms: **linear multimedia and non-linear multimedia.** As shown in Figure (3).

In **linear multimedia**, information is read or viewed in a continuous sequence. Usually, these presentations begin at a predetermined starting point and end at a predetermined end point. They can be automated so that each screen comes after a fixed time interval. Linear active content

progresses without any navigation control for the viewer such as a **cinema presentation, PowerPoint presentation.**

Non-linear multimedia information is not presented in sequential. Non-linear multimedia programs are usually interactive and require audience interaction. Non-linear content offers user interactivity to control progress as used with a computer game or used in self-paced computer based training. Non-linear content is also **known as hypermedia content.** Example: One of the most common examples of this form of multimedia is the **Web.**

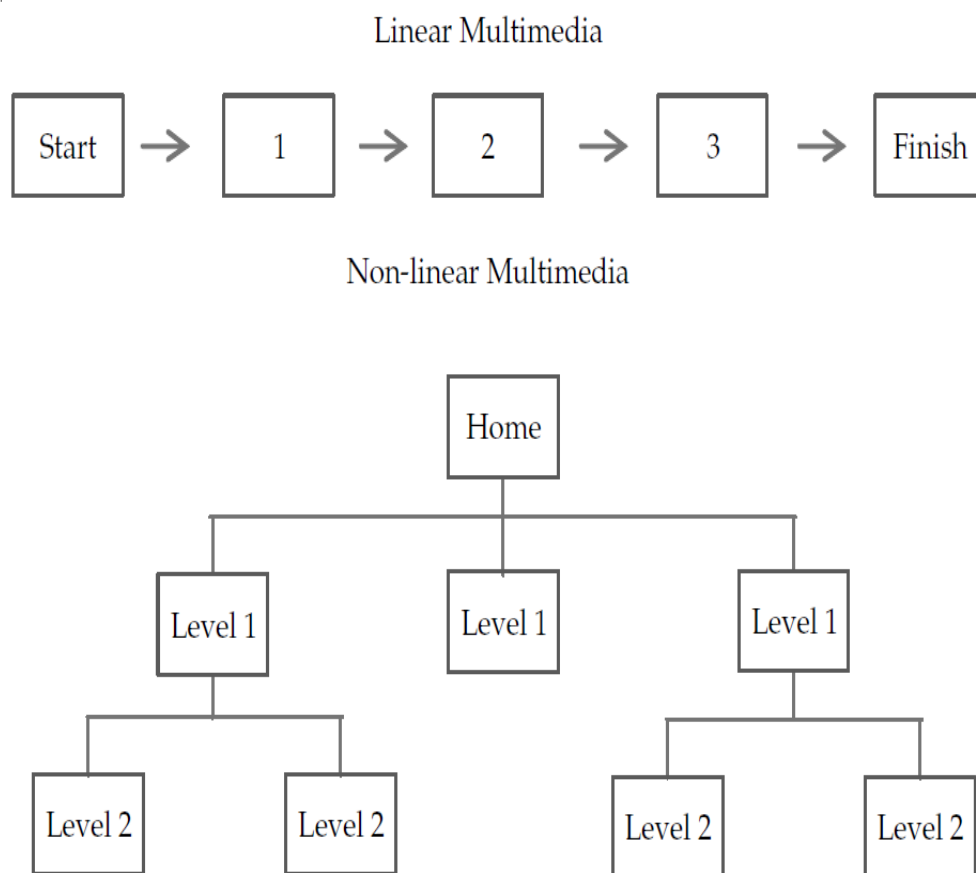


Figure 3: Linear and Non-linear Multimedia

2. Hypertext and Hypermedia

- A **hypertext** system: We may think of a book as a linear medium, basically meant to be read from beginning to end, a hypertext meant to be read **nonlinearly**, by following links that point to other parts of the document, or to other documents, as shown in Figure (4).

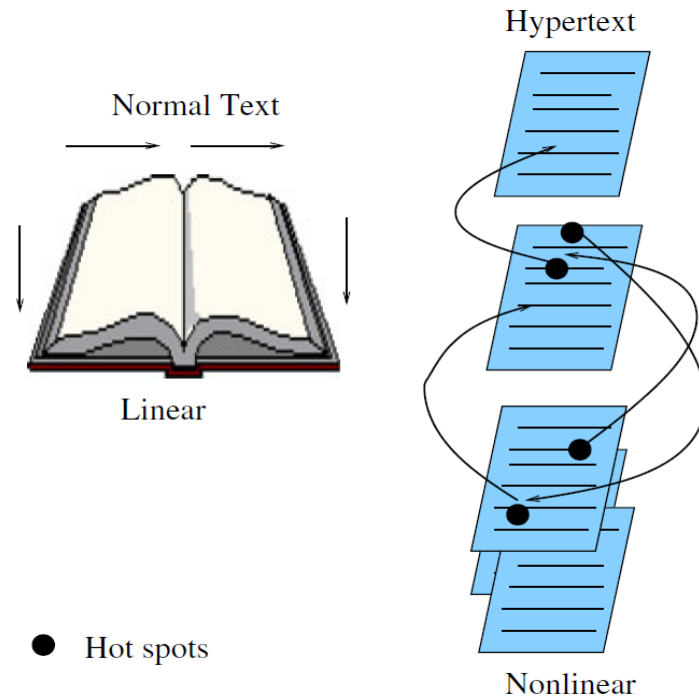


Figure 4: Hypertext is nonlinear

- **Hypermedia:** not constrained to be text-based, can include other media, e.g., graphics, images, and especially the continuous media — sound and video (e.g. home pages contain a link to video, images, and sound).

Both the terms hypertext and hypermedia follow a similar structure comprising of nodes that are interconnected by links except in hypermedia systems, the nodes can contain multiple forms of media such as text, images, audio, video, and graphics. The main difference lies in the way they are implemented. **Hypertext is used to represent multimedia content in electronic text format** whereas **hypermedia combines both hypertext and multimedia to provide access to a wealth of information usually in a non-linear sequence.**

Hypertext vs Hypermedia

Hypertext	Hypermedia
Text which links to other blocks of text within the same document or a different	Extension of hypertext but not limited to text elements
Represents text elements	Represents different media such as text, audio, image, video and still or moving graphics
Allows the users to navigate through text in a non-linear way	Includes multimedia elements to improve the multimedia experience

3. Components of Multimedia

The various components of multimedia are Text, Audio, Graphics, Video and Animation. All these components work together to represent information in an effective and easy manner.

1. **Text:** Text is the most common medium of representing the information. In multimedia, text is mostly use for titles, headlines, menu etc.
2. **Audio:** In multimedia audio means related with recording, playing etc. Audio is an important components of multimedia because this component increases the understandability and improves the clarity of the concept. audio includes speech, music etc.
3. **Graphics:** Every multimedia presentation is based on graphics. The used of graphics in multimedia makes the concept more effective and presentable.
4. **Video:** Video means moving pictures with sound. It is the best way to communicate with each other.
5. **Animation:** In computer animation is used to make changes to the images so that the sequence of the images appears to be moving pictures.

An animated sequence shows a number of frames per second to produce an effect of motion in the user's eye.

Multimedia involves multiple modalities of text, audio, images, drawings, animation, and video. Examples of how these modalities are put to use:

- **Geographically based**, real-time augmented-reality, massively **multiplayer online video games**, making use of any portable device such as smartphones, laptops, or tablets, which function as **GPS-aware mobile game consoles**. For example, a game in which players reinforce and link friendly “portals,” and attack enemy ones that are played on GPS-enabled devices where the players must physically move to the portals (which are overlaid on real sites such as interesting buildings, or parks) in order to interact with them.
- **Shapeshifting TV**, where **viewers vote** on the **plot path by phone text-messages**, which are parsed to direct plot changes in real-time.
- A **camera** that suggests what would be the best type of next shot so as to adhere to good technique guidelines for developing storyboards.
- A **Web-based video editor** that lets anyone create a new video by editing, and remixing professional videos on the cloud.
- **Cooperative education environments** that allow schoolchildren to share a single educational game using two mice at once that pass control back and forth.
- **Searching (very) large video and image databases** for target visual objects, using semantics of objects.
- **Compositing of artificial and natural video into hybrid scenes**, placing real appearing computer graphics and video objects into scenes so as to take the physics of objects and lights (e.g., shadows) into account.

- **Visual cues of video-conference participants**, taking into account gaze direction and attention of participants.
- **Making multimedia components editable**—allowing the user side to decide what components, video, graphics, and so on are actually viewed and allowing the client to move components around or delete them making components distributed.
- **Building “inverse-Hollywood” applications** that can recreate the process by which a video was made, allowing storyboard pruning and concise video summarization.

4. Multimedia Research Topics and Projects

In digital world, multimedia content is recorded or accessed by **digital information content processing devices**, ranging from smartphones, tablets, laptops, personal computers, smart TVs, and game consoles, to servers and data centers, over such distribution media as tapes, hard drives, and disks, or more popularly nowadays, wired and wireless networks. This leads to a wide variety of **research topics**:

- **Multimedia processing and coding**. This includes audio/image/video processing, compression algorithms, multimedia content analysis, content-based multimedia retrieval, multimedia security, and so on.
- **Multimedia system support and networking**. People look at such topics as network protocols, Internet and wireless networks, operating systems, servers and clients, and databases.
- **Multimedia tools, end systems, and applications**. These include hypermedia systems, user interfaces, authoring systems, **multimodal interaction** (multimodal interaction provides the user with multiple modes of interacting with a system), and integration, web-everywhere devices, multimedia education, including computer supported

collaborative learning and design, and applications of virtual environments.

Many exciting **research projects** are currently underway. Here are a few of them:

1. **Camera-based object tracking technology:** One aim is to develop control systems for industrial control, gaming, and so on that rely on moving scale models (toys) around a real environment (a board game, say). Tracking the control objects (toys) provides user control of the process.
2. **3D motion capture** can also be used for multiple actor capture, so that multiple real actors in a virtual studio can be used to automatically produce realistic animated models with natural movement.
3. **Multiple views** from several cameras or from a single camera under differing lighting can accurately acquire data that gives both the shape and surface properties of materials, thus automatically generating synthetic graphics models. This allows photo-realistic synthesis of virtual actors.
4. **3D capture technology:** allow acquiring dynamic characteristics of human facial expression during the speech, to synthesize highly realistic facial animation from speech.
5. **Specific multimedia applications:** aimed at handicapped persons with low vision capability and the elderly a rich field of endeavor in current research.
6. **Digital fashion:** aims to develop smart clothing that can communicate with other such enhanced clothing using wireless communication, so as to artificially enhance human interaction in a social setting. The vision here is to use technology to allow individuals to allow certain thoughts

and feelings to be broadcast automatically, for exchange with others equipped with similar technology.

7. **Electronic House call system:** an initiative for providing interactive health monitoring services to patients in their homes, relies on networks for delivery, challenging current capabilities.

5. Application of Multimedia

Multimedia finds its application in various areas including, but not limited to, advertisements, art, education, entertainment, engineering, medicine, mathematics, business, scientific research and spatial, temporal applications. A few application areas of multimedia are listed below:

- **Creative Industries**

Creative industries use multimedia for a variety of purposes ranging from fine arts, to entertainment, to commercial art, to journalism, to media and software services provided for any of the industries listed below. An individual multimedia designer may cover the spectrum throughout their career. Request for their skills range from technical, to analytical and to creative.

- **Commercial**

Much of the electronic old and new media utilized by commercial artists is multimedia. Exciting presentations are used to grab and keep attention in advertising. Industrial, business to business, and interoffice communications are often developed by creative services firms for advanced multimedia presentations beyond simple slide shows to sell ideas or liven-up training. Commercial multimedia developers may be hired to design for governmental services and nonprofit services applications as well.

- **Entertainment and Fine Arts**

In addition, 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. Some video games also use multimedia features. Multimedia applications that allow users to actively participate instead of just sitting by as passive recipients of information are called Interactive Multimedia.

- **Education**

In Education, multimedia is used to produce computer-based training courses (popularly called CBTs) and reference books like encyclopedia. A CBT lets the user go through a series of presentations, text about a particular topic, and associated illustrations in various information formats.

Eduainment is an informal term used to describe combining education with entertainment, especially multimedia entertainment.

- **Engineering**

Software engineers may use multimedia in Computer Simulations for anything from entertainment to training such as military or industrial training. Multimedia for software interfaces are often done as collaboration between creative professionals and software engineers.

- **Industry**

In the Industrial sector, multimedia is used as a way to help present information to shareholders, superiors and coworkers. Multimedia is also helpful for providing employee training, advertising and selling products all over the world via virtually unlimited web-based technologies.

- **Mathematical & Scientific Research**

In Mathematical and Scientific Research, multimedia is mainly used for modeling and simulation. For example, a scientist can look at a

molecular model of a particular substance and manipulate it to arrive at a new substance. Representative research can be found in journals such as the Journal of Multimedia.

- **Medicine**

In Medicine, doctors can get trained by looking at a virtual surgery or they can simulate how the human body is affected by diseases spread by viruses and bacteria and then develop techniques to prevent it.

- **Multimedia in Public Places**

In hotels, railway stations, shopping malls, museums, and grocery stores, multimedia will become available at stand-alone terminals or kiosks to provide information and help. Such installation reduce demand on traditional information booths and personnel, add value, and they can work around the clock, even in the middle of the night, when live help is off duty.

A menu screen from a supermarket kiosk that provide services ranging from meal planning to coupons. Hotel kiosk list nearby restaurant, maps of the city and airline schedules.

Example Multimedia Applications:

- World Wide Web
- Multimedia Authoring, e.g. Adobe/Macromedia Director
- Computer Games
- Multimedia Database systems

6. Multimedia on the Web

The World Wide Web (abbreviated as WWW or W3, commonly known as the **Web**), is a system of interlinked hypertext documents accessed via the Internet. With a web browser, one can view web pages that may contain text, images, videos, and other multimedia, and navigate between them via hyperlinks. WWW technology is maintained and developed by the World Wide Web Consortium (W3C). **The W3C has**

listed the **following three goals for the WWW**: universal access of web resources (by everyone everywhere), effectiveness of navigating available information, and responsible use of posted material.

6.1 HyperText Transfer Protocol (HTTP)

The **Hypertext Transfer Protocol (HTTP)** is an application protocol for distributed, collaborative, hypermedia information systems. HTTP is the foundation of data communication for the World Wide Web. HTTP was originally designed for transmitting hypermedia, but it also supports transmission of any file type. Hypertext is a multilinear set of objects, building a network by using logical links (the so - called hyperlinks) between the nodes (e.g. text or words).

6.2 HyperText Markup Language (HTML)

HyperText Markup Language (HTML) is the main markup language **for displaying web pages and other information that can be displayed in a web browser**. HTML is written in the form of HTML elements consisting of tags enclosed in angle brackets (like <html>), within the web page content. HTML tags most commonly come in pairs like <h1> and </h1>, for example. **A very simple HTML page is as follows**


```
<HTML>
<HEAD>
  <TITLE>
  A sample web page.
  </TITLE>
  <META NAME = "Author" CONTENT = "Cranky Professor">
</HEAD> <BODY>
  <P>
  We can put any text we like here, since this is
  a paragraph element.
  </P>
</BODY>
</HTML>
```

The first tag in a pair is the start tag, the second tag is the end tag (they are also called opening tags and closing tags). In between these tags web designers can add text, tags, comments and other types of text - based content. The purpose of a **web browser** is **to read HTML documents and compose them into visible or audible web pages**. The browser does not display the HTML tags, but uses the tags to interpret the content of the page. HTML elements form the building blocks of all websites. HTML allows images and objects to be embedded and can be used to create interactive forms. It provides a means to create structured documents by denoting structural semantics for text such as headings, paragraphs, lists, links, quotes and other items. It can embed scripts in languages such as JavaScript which affect the behavior of HTML webpages.

Web browsers can also refer to **Cascading Style Sheets** (CSS) to **define the appearance and layout of text and other material**. The W3C, maintainer of both the HTML and the CSS standards, encourages the use of CSS over explicitly presentational HTML markup.

6.3 Extensible Markup Language (XML)

Extensible Markup Language (XML) is a markup language that **defines a set of rules for encoding documents in a format that is both human-readable and machine-readable**. The design **goals of XML** focus on **simplicity, generality, and usability across the Internet**. It is a textual data format with strong support via Unicode for different human languages. Although the design of XML focuses on documents, the language is widely used for the representation of arbitrary data structures such as those used in web services. **XML is designed to store and transport data.**

Differences Between XML and HTML

1. XML is a text-based markup language which has the self-describing structure and can effectively define another markup language. On the other hand, HTML is a predefined markup language and has a limited capability.
2. XML provides logical structuring of the document while HTML structure is predefined where “head” and “body” tags are used.
3. When it comes to language type HTML is case insensitive. As against, XML is case sensitive.
4. HTML was designed with the emphasis on the presentational features of the data. In contrast, XML is data specific where the data storage and transfer was the prior concern.
5. XML does not permit any mistake if there are some errors in the code it could not be parsed. Inversely, in HTML small errors can be neglected.
6. Whitespaces in XML are used for a specific use as XML considers every single character. On the contrary, HTML can ignore the whitespaces.
7. The tags in XML are mandatory to be closed, whereas in HTML an open tag can also work completely fine.

8. Nesting in XML should be done correctly, it has a big importance in XML syntax. Conversely, HTML does not care much about nesting.

7. Multimedia Data Basics

Analog and Digital Signal Conversion

The world we sense is full of analog signals: Electrical sensors convert the medium they sense into electrical signals

- E.g. microphones: acoustic sensor, Cameras (Video): light sensor. (usually) continuous Analog signals (e.g. Sound and Light).

Analog: the continuous signal for which the time-varying feature (variable) of the signal is a representation of some other time-varying quantity.

Digital: a signal which is represented as a sequence of numbers (usually in binary numbers) discrete digital signals that computer can readily deal with ex.: digital image – matrix of pixels.

Special hardware devices: Analog-to-Digital Playback {a converse operation to Analog-to-Digital converters. E.g. Audio: Take analog signals from analog sensor (e.g. microphone) as shown in Figure (5).

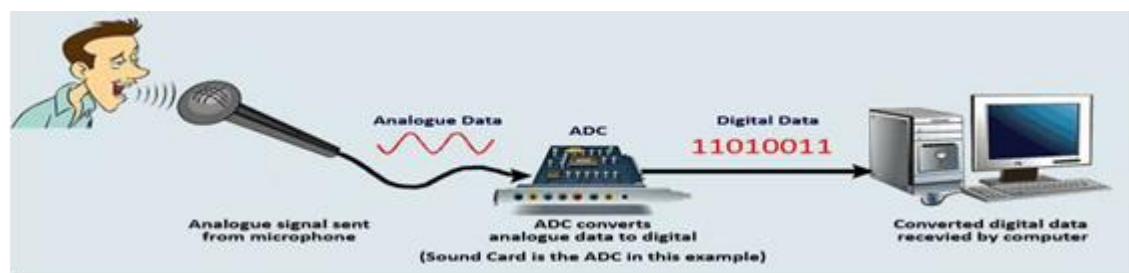


Figure 5: Analog to Digital Converter

Analog-to-Digital-to-Analog Pipeline

Begins at the conversion from the analog input and ends at the conversion from the output of the processing system to the analog output as shown in Figure (6):

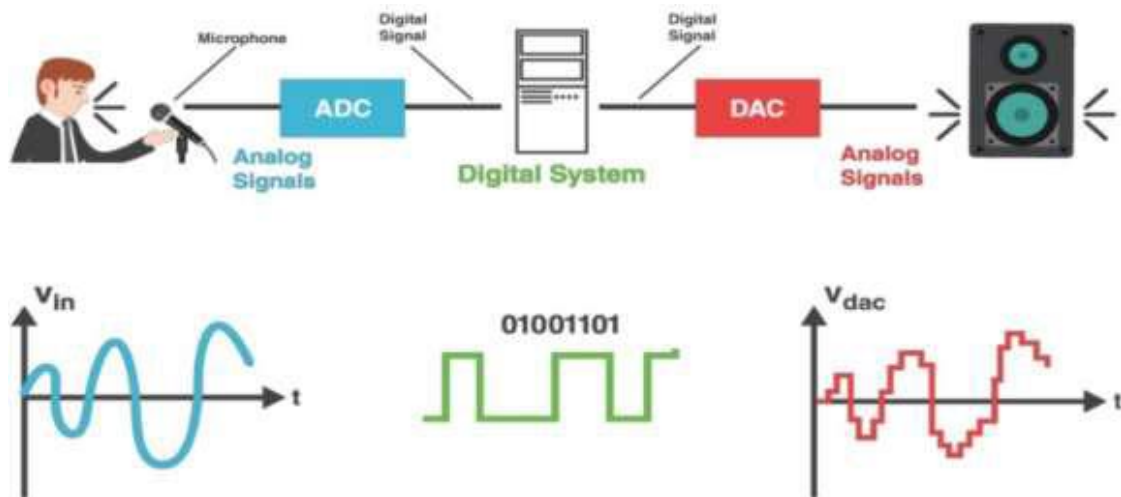


Figure 6: Analog to Digital to Analog Converter

Multimedia Data: Input and Format

How to capture and store each Media format?

- Note that text and graphics (and some images) are mainly generated directly by computer/device (e.g. drawing/painting programs) and do not require digitizing.
- Printed text and some handwritten text can be scanned via Optical Character Recognition.
- Handwritten text could also be digitized by electronic pen sensing. Printed imagery/graphics can be scanned directly to image formats.

All data must be in the form of digital information. The data may be in a variety of formats:

- 1- Text
- 2- Graphics
- 3- Images
- 4- Audio
- 5- Video

1) Text and Static Data

- characters that are used to create words, sentences, and paragraphs.
- Source: keyboard, optical character recognition, data stored on disk.
- Storage: 1 byte per character (text or format character), e.g. ASCII; more bytes for Unicode. For other forms of data (e.g. Spread sheet). May store as text (with formatting, e.g. CSV {Comma-Separated Values}) or may use binary encoding.
- Formatted Text: Raw text or formatted text e.g HTML, Rich Text Format (RTF), Word or a program language source (Java, Python, MATLAB).
- Compression: convenient to bundle for archiving and transmission of larger. E.g. Zip, RAR, 7-zip.

2) Graphics

- A digital representation of non-text information, such as a drawing, chart, or photograph.
- 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, Freehand).
- Graphics input devices: keyboard (for text and cursor control), mouse or graphics tablet. Graphics are usually selectable and editable or revisable (unlike images). Graphics usually store the primitive assembly. Do not take up a very high storage overhead.
- Graphics standards: Open Graphics Library, a standard specification defining a cross-language, cross-platform API for writing applications that produce 2D/3D graphics.
- Animation: can be generated via a sequence of slightly changed graphics. 2D animation: e.g. Flash | Key frame interpolation: **Tweening** (the process of generating intermediate frames between two images to give

the appearance that the first image evolves smoothly into the second image): motion & shape. 3D animation: e.g. Maya.

3) Images

- Still pictures which (uncompressed) are represented as a bitmap (a grid of pixels). Input: scanned for photographs or pictures using a digital scanner or from a digital camera.
- Input: May also be generated by programs similar to graphics or animation programs.
- Analog sources will require digitizing. Compression is commonly applied.
- Can usually only edit individual or groups of pixels in an image editing application, e.g. Photoshop.

4) Audio

- Audio signals are continuous analog signals, music, speech, or any other sound.
- Input: microphones and then digitized and stored.
- CD Quality Audio requires 16-bit sampling at 44.1 KHz.
- Usually compressed (E.g. MP3, AAC).

5) Video

- photographic images that are played back at speeds of 15 to 30 frames a second and they provide the appearance of full motion.
- Input: Analog Video is usually captured by a video camera and then digitized, although digital video cameras now essentially perform tasks.
- Raw video can be regarded as being a series of single images. There are typically 25, 30 or 50 frames per second.

Part two

8. Graphics and Image Data Representation

8.1 Digital Image Processing Basics

Digital Image Processing means processing digital images by means of a digital computer. We can also say that it is the use of computer algorithms, in order to get enhanced images either to extract some useful information. Image processing mainly includes the following steps:

1. Importing the image via image acquisition tools.
2. Analyzing and manipulating the image.
3. Output in which result can be altered image or a report which is based on analyzing that image.

8.2 What is an image

An **image** may be defined as a two-dimensional function, $f(x, y)$, where x and y are spatial (plane) coordinates, and the amplitude of f at any pair of coordinates (x, y) is called the **intensity** or **gray level** of the image at that point. When x , y , and the amplitude values of f are all finite, discrete quantities, we call the image a **digital image**. Note that a digital image is composed of a finite number of elements, each of which has a particular location and value.

In digital imaging, a **pixel** (or picture element) is a single point in a raster image. The **pixel** is the smallest addressable screen element; it is the smallest unit of picture that can be controlled. Each pixel has its own address. The address of a pixel corresponds to its coordinates.

Vector VS Raster

Vector	Raster
Vector images are mathematical calculations from one point to another that form lines and shapes. If you zoom into a vector graphic, it will always look the same.	Raster images are made of pixels. A pixel is a single point or the smallest single element in a display device. If you zoom in to a raster image, you may start to see many little tiny squares.
Vector programs best for creating logos, drawings and illustrations, technical drawings. For images that will be applied to physical products.	Raster programs best for editing photos and creating continuous tone images with soft color blends.

8.3 Image as a Matrix

An image can be defined by a two-dimensional array specifically arranged in rows and columns. Let the numbers of a row and column be M and N , then the size of an image is $M \times N$. The pixel values constitute a real matrix with the size of $M \times N$. Every element of this matrix is called image element, picture element, or pixel. Assume that an image $f(x, y)$ is sampled so that the resulting digital image has M rows and N columns. The values of the coordinates (x, y) now become discrete quantities. For notational clarity and convenience, we shall use integer values for these discrete coordinates. Thus, the values of the coordinates at the origin are $(x, y) = (0, 0)$. The next coordinate values along the first row of the image are represented as $(x, y) = (0, 1)$. It is important to keep in mind that the notation $(0, 1)$ is used to signify the second sample along the first row. It does not mean that these are the actual values of physical coordinates when the image was sampled. Figure (7) shows the coordinate convention used.

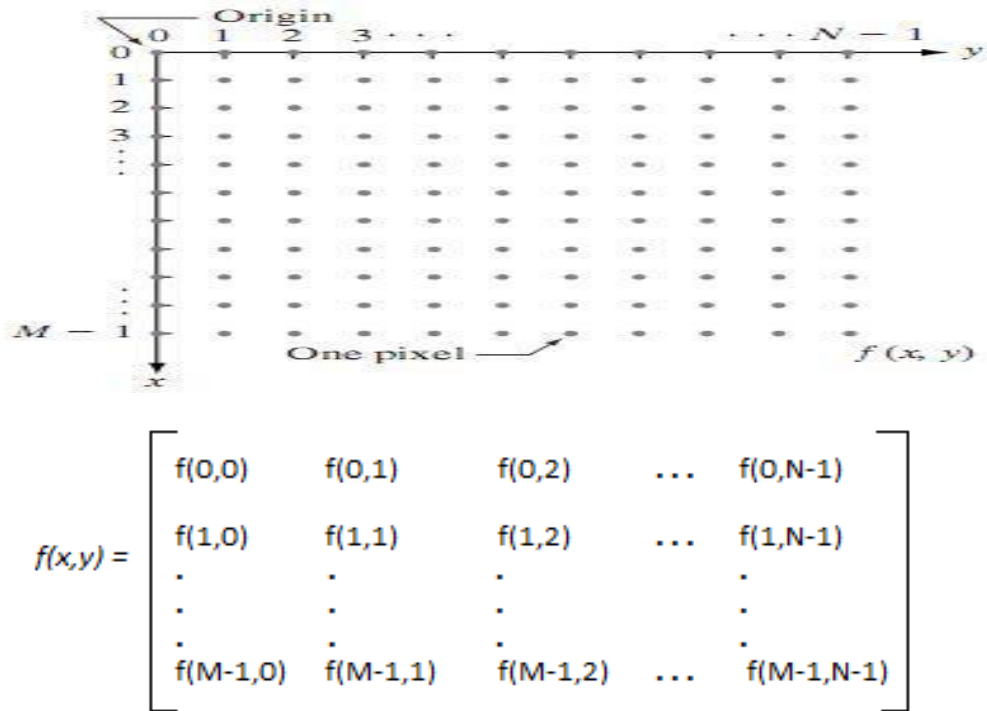


Figure 7: Coordinate convention used to represent digital images

DIGITAL IMAGE REPRESENTATION IN MATLAB: In MATLAB the start index is from 1 instead of 0. Therefore, $f(1,1) = f(0,0)$ as shown in Figure (8).

$$f = \begin{bmatrix} f(1, 1) & f(1, 2) & \dots & f(1, N) \\ f(2, 1) & f(2, 2) & \dots & f(2, N) \\ \vdots & \vdots & & \vdots \\ f(M, 1) & f(M, 2) & \dots & f(M, N) \end{bmatrix}$$

Figure 8: Represent Digital Image in Matlab

This **digitization** process requires decisions about values for M, N, and for the number L (max gray levels) of discrete gray levels allowed for each pixel. There are no requirements on M and N, **other than that they have to be positive integers. However, due to processing, storage, and sampling hardware considerations**, the number of gray levels typically is an integer power of 2:

$$L = 2^k.$$

We assume that the discrete levels are equally spaced and that they are integers in the interval $[0, L-1]$. **Sometimes the range of values spanned by the gray scale is called the dynamic range of an image, and we refer to images whose gray levels span a significant portion of the gray scale as having a high dynamic range.** Thus, for an 8 bit image (class: uint8) the number of discrete grayscale intensity levels are $[0, 255]$. **The number b of bits required to store a digitized image is**

$$B = M*N*k.$$

Assuming it has 1024 rows and it has 1024 columns. And since it is a gray scale image, it has 256 different shades of gray or it has bits per pixel. Then putting these values in the formula, we get

$$\text{Size of an image} = \text{rows} * \text{cols} * k$$

$$= 1024 * 1024 * 8 = 8388608 \text{ bits.}$$

But since its not a standard answer that we recognize, so will convert it into our format.

$$\text{Converting it into bytes} = 8388608 / 8 = 1048576 \text{ bytes.}$$

$$\text{Converting into kilo bytes} = 1048576 / 1024 = 1024\text{kB.}$$

$$\text{Converting into Mega bytes} = 1024 / 1024 = 1 \text{ MB.}$$

Example: In digital image of M rows and N columns and L discrete gray levels, calculate the bits required to store a digitized image for $M=N=32$ and $L=16$.

Sol: The number, b, of bits required to store a digitized image is:
 $b=M*N*k.$

For $L=16$, $L=2^k$, $16 = 2^k$, then $k=4$.

$$b = 32*32*4 = 4096 \text{ bits}$$

9. Image Digitization

The output of most sensors is a continuous voltage waveform whose amplitude and spatial behavior are related to the physical phenomenon being sensed. To create a digital image, we need to **convert the continuous sensed data into digital form. This involves two processes: sampling and quantization.**

9.1 Basic Concepts in Sampling and Quantization:

The basic idea behind sampling and quantization is illustrated in Figure (9). Figure 9(a) shows a continuous image, $f(x, y)$, that we want to convert to digital form. An image may be continuous with respect to the x and y coordinates, and also in amplitude. To convert it to digital form, we have to sample the function in both coordinates and in amplitude. **Digitizing the coordinate values is called sampling. Digitizing the amplitude values is called quantization.**

The one-dimensional function shown in Fig.9 (b) is a plot of amplitude (gray level) values of the continuous image along the line segment AB in Fig.9 (a). The random variations are due to image noise. To sample this function, we take equally spaced samples along line AB, as shown in Fig.9 (c). The location of each sample is given by a vertical tick mark in the bottom part of the figure. The samples are shown as small white squares superimposed on the function. The set of these discrete locations gives the sampled function. However, the values of the samples still span (vertically) a continuous range of gray-level values. In order to form a digital function, the gray-level values also must be converted (quantized) into discrete quantities. The right side of Fig. 9(c) shows the gray-level scale divided into eight discrete levels, ranging from black to white. The vertical tick marks indicate the specific value assigned to each of the eight gray levels. The continuous gray levels are quantized simply by assigning one of

the eight discrete gray levels to each sample. The assignment is made depending on the vertical proximity of a sample to a vertical tick mark. The digital samples resulting from both sampling and quantization are shown in Fig.9 (d). Starting at the top of the image and carrying out this procedure line by line produces a two- dimensional digital image.

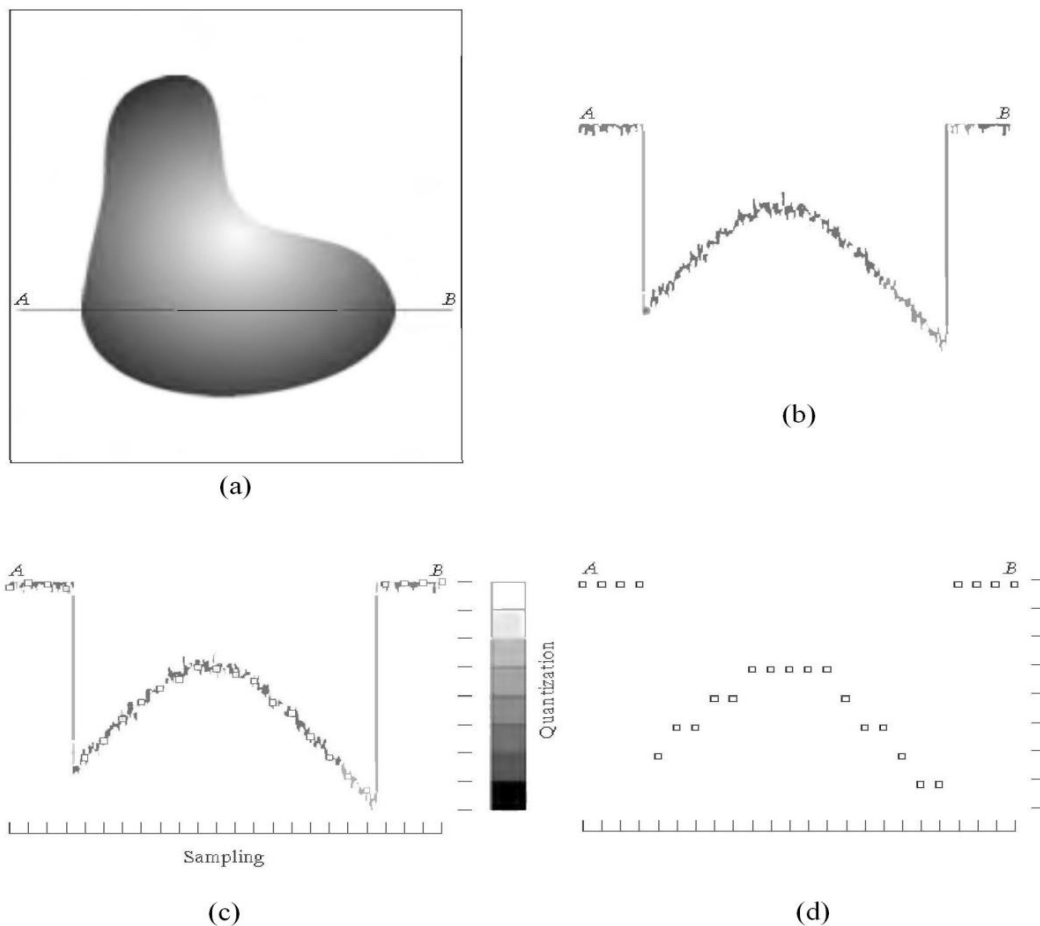


Fig.9. Generating a digital image (a) Continuous image (b) A scan line from A to B in the continuous image (c) Sampling and quantization. (d) Digital scan line.

Sampling in the manner just described assumes that we have a continuous image in both coordinate directions as well as in amplitude. In practice, the method of sampling is determined by the sensor arrangement used to generate the image. When an image is generated by a single sensing element combined with mechanical motion. However, sampling is

accomplished by selecting the number of individual mechanical increments at which we activate the sensor to collect data. Mechanical motion can be made very exact so, in principle; there is almost no limit as to how fine we can sample an image. However, practical limits are established by imperfections in the optics used to focus on the sensor an illumination spot that is inconsistent with the fine resolution achievable with mechanical displacements. When a sensing strip is used for image acquisition, the number of sensors in the strip establishes the sampling limitations in one image direction. Mechanical motion in the other direction can be controlled more accurately, but it makes little sense to try to achieve sampling density in one direction that exceeds the sampling limits established by the number of sensors in the other. Quantization of the sensor outputs completes the process of generating a digital image.

When a sensing array is used for image acquisition, there is no motion and the number of sensors in the array establishes the limits of sampling in both directions. Figure 10 illustrates this concept. Figure 10 (a) shows a continuous image projected onto the plane of an array sensor. Figure 10 (b) shows the image after sampling and quantization. Clearly, the quality of a digital image is determined to a large degree by the number of samples and discrete gray levels used in sampling and quantization.

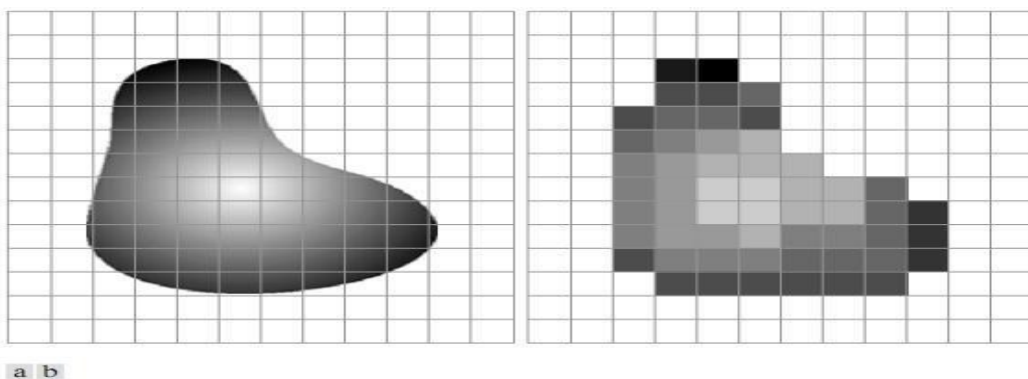


Fig.10 (a) Continuous image projected onto a sensor array (b) Result of image sampling and quantization.

10. Spatial resolution and quantization

Spatial resolution is the smallest discernible detail in an image. It is determined by the sampling process. The spatial resolution of a digital image reflects the amount of details that one can see in the image (i.e. the ratio of pixel “area” to the area of the image display). If an image is spatially sampled at $M \times N$ pixels, then the larger $M \times N$ the finer the observed details.

Gray-level resolution refers to the smallest discernible change in gray level. It is determined by the quantization process. As mentioned earlier, the number of gray levels is usually an integer power of 2. The most common number is 8 bits, however, 16 bits is used in some applications where enhancement of specific gray-level ranges is necessary.

10.1 Effect of reducing the spatial resolution

Decreasing spatial resolution of a digital image, within the same area, may result in what is known as *checkerboard pattern*. Also image details are lost when the spatial resolution is reduced.

To demonstrate the checkerboard pattern effect, we subsample the 1024×1024 image shown in the figure below to obtain the image of size 512×512 pixels. The 512×512 is then subsampled to 256×256 image, and so on until 32×32 image. The **subsampling process** means deleting the appropriate number of rows and columns from the original image. The number of allowed gray levels was kept at 256 in all the images.

To see the effects resulting from the reduction in the number of samples, we bring all the subsampled images up to size 1024×1024 by row and column pixel replication. The resulted images are shown in the figure (11) below.



Figure 11: A 1024×1024, 8-bit image subsampled down to size 32×32 pixels.

Compare Figure 12(a) with the 512×512 image in Figure 12(b), we find that the level of detail lost is simply too fine to be seen on the printed page at the scale in which these images are shown. Next, the 256×256 image in Figure 12(c) shows a very slight fine checkerboard pattern in the borders between flower petals and the black background. A slightly more pronounced graininess throughout the image also is beginning to appear. These effects are much more visible in the 128×128 image in Figure 13(d), and they become pronounced in the 64×64 and 32×32 images in Figures 12(e) and (f), respectively.

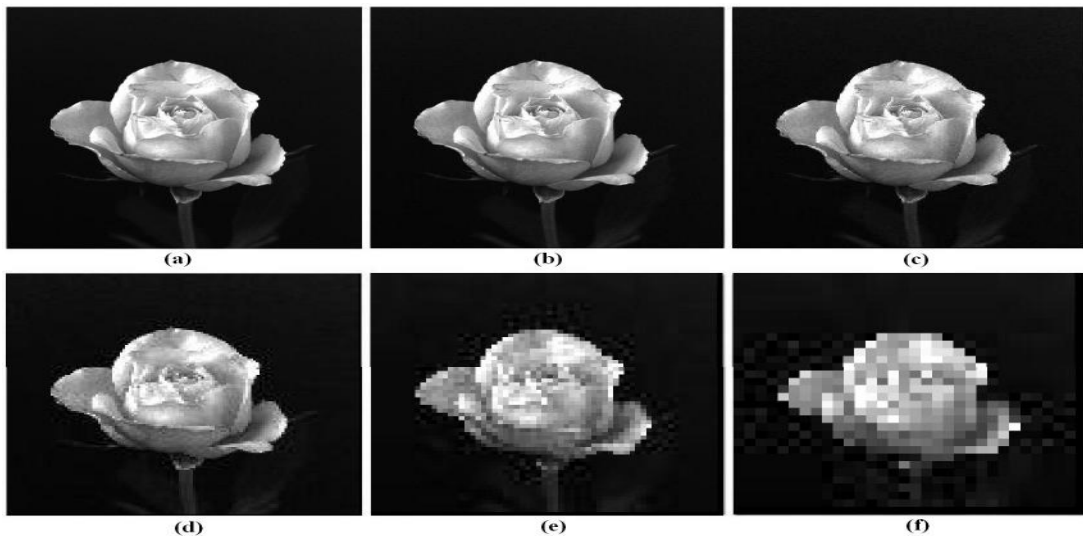


Figure 12: (a) 1024×1024, 8-bit image. (b) through (f) 512×512, 256×256, 128×128, 64×64, and 32×32 images resampled into 1024×1024 pixels by row and column duplication

10.2 Effect of reducing the gray-level resolution

Decreasing the gray-level resolution of a digital image may result in what is known as *false contouring*. This effect is caused by the use of an insufficient number of gray levels in smooth areas of a digital image.

To illustrate the false contouring effect, we reduce the number of gray levels of the 256-level image shown in Figure 13(a) from 256 to 2. The resulted images are shown in the figures 13(b) through (h). This can be achieved by reducing the number of bits from $k = 7$ to $k = 1$ while keeping the spatial resolution constant at 452×374 pixels.

We can clearly see that the 256-, 128-, and 64-level images are visually identical. However, the 32-level image shown in Figure 13(d) has an almost imperceptible set of very fine structures in areas of smooth gray levels (particularly in the skull). False contouring generally is quite visible in images displayed using 16 or less uniformly spaced gray levels, as the images in Figures 13(e) through (h) show.

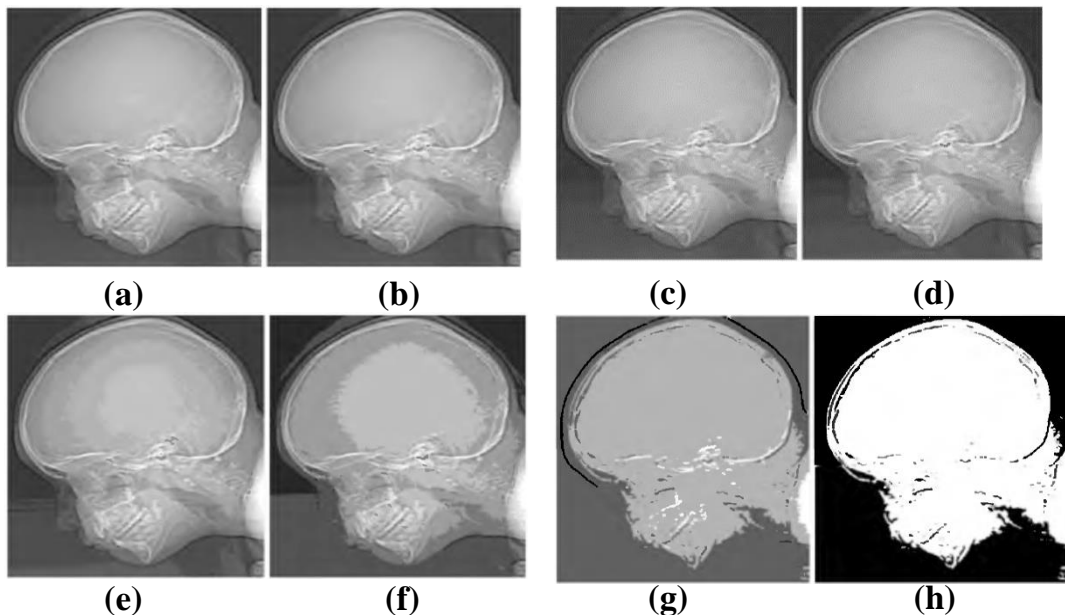


Figure 13 (a) 452×374 , 256-level image. (b)-(h) Image displayed in 128, 64, 32, 16, 8, 4, and 2 gray levels, while keeping the spatial resolution constant.

Resolution (Pixel density): The number of pixels in each dimension that can be displayed the density of pixels in the image.

Bit depth: the color information stored in an image. (The higher the bit depth, the more colors an image can store).

What is size of the image? In addition, what is a resolution?

Resolution = width x height.

Image Size = width x height x No. of bit per pixel.

Quantization = Number of bits per pixel (Quantization).

11. Type of image

A pixel is the smallest element of a digital image that you can access, and its address is the physical coordinates in which it is located. Based on pixels and how they are stored, there are the following different types of images (Binary Image, Gray Scale Image, RGB Image, Indexed Image):

1. Binary Image (1 - Bit Images)

Images consist of pixels in digital 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). Usually, **0 usually means “black” and a 1 means “white”**, though it is possible to use any two colors. Hence, such an image is also referred to as **a binary image**.

It is also called a 1 - bit **monochrome** image, since it contains no color. The figure (14) shows a 1 - bit monochrome image (called "Lena" by multimedia scientists this is a standard image used to illustrate many algorithms). A 640 x 480 monochrome image requires 37.5 kilobytes of storage $((640 \times 480 / 8) / 1024)$. **Monochrome 1 - bit images can be satisfactory for pictures containing only simple graphics and text.**



Fig. 14: Monochrome 1-bit Lena image.

2. 8-bit Gray-Scale Images

Each pixel has a gray-value between 0 and 255. Each pixel is represented by a single byte; **0** corresponds to “black,” and **255** to “white” e.g., a dark pixel might have a value of 10, and a bright one might be 230. Figure (15) below shows an example of a greyscale image. Here, we do not have colors, but just shades of grey. That is the shades between Black and White. When we read a $M \times N$ grayscale image, we get a 2D array of size $M \times N$. See the image below to understand how the shades of grey changes as the value changes from 0 to 255.

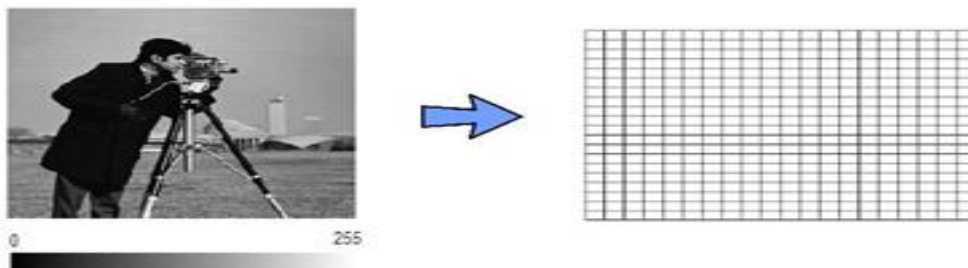


Fig. 15: Gray-Scale Image.

Bitmap the two-dimensional array of pixel values that represents the graphics/image data. **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 1600×1200 , whereas lower resolution might be 640×480 . Notice that here we are using an **aspect ratio** of 4:3. We don't have to adopt this ratio, but it has been found to look natural.

We can think of the 8 - bit image as a set of 1 bit-planes. The following figure (16) displays the concept of bit-planes graphically. Each

bit - plane can have a value of 0 or 1 at each pixel but, together, all the bit-planes make up a single byte that stores values between 0 and 255.

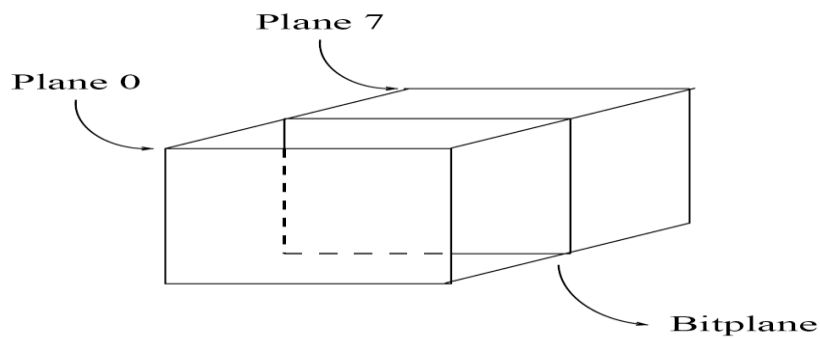


Fig. 16: Bit-planes for 8-bit grayscale image.

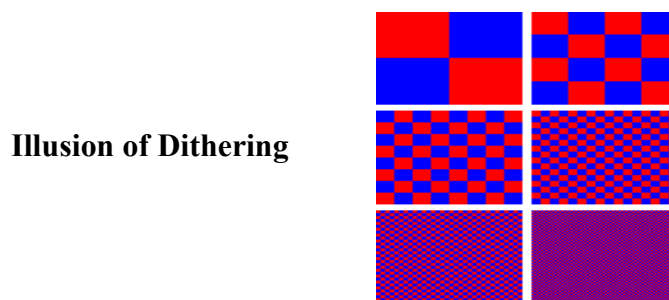
Each pixel is usually stored as a byte (a value between 0 and 255), so a 640 x 480 grayscale image requires 300 kilobytes of storage $(640 \times 480) \times 8/8 = 307,200$ Byte.

Dithering in Computer Graphics

Dithering is the attempt by a computer program to approximate a color from a mixture of other colors when the required color is not available.

For example, dithering occurs when a color is specified for a Web page that a browser on a particular operating system can't support. The browser will then attempt to replace the requested color with an approximation composed of two or more other colors it can produce.

The result may or may not be acceptable to the graphic designer. It may also appear somewhat **grainy** since it's composed of different pixel intensities rather than a single intensity over the colored space.



Illusion of Dithering

If we wish to print an image, things become more complex. Suppose we have available a 600 dot-per-inch (dpi) laser printer. Such a device can usually only print a dot or not print it. However, a 600×600 image will be printed in a 1-inch space and will thus not be very pleasing. Instead, dithering is used.

For printing all a 1-bit printer, dithering is used to calculate larger patterns of dots, such that values from 0 to 255 correspond to pleasing patterns that correctly represent darker and brighter pixel values.

The main strategy is to replace a pixel value by a larger pattern, say 2×2 or 4×4 , such that the number of printed dots approximates the varying-sized disks of ink used in analog, in **halftone printing** (e.g., for newspaper photos).

1. Halftone printing is an analog process that uses smaller or larger filled circles of black ink to represent shading, for newspaper printing.
2. For example, if we use a 2×2 **dither matrix** we can first re-map image values in $0..255$ into the new range $0..4$ by (integer) dividing by $256/5$. Then, e.g., if the pixel value is 0 we print nothing, in a 2×2 area of printer output. But if the pixel value is 4 we print all four dots.

$$\begin{pmatrix} 0 & 2 \\ 3 & 1 \end{pmatrix}$$

The rule is:

If the intensity is $>$ the dither matrix entry then print an on dot at that entry location: replace each pixel by an $n \times n$ matrix of dots.

Note that the image size may be much larger, for a dithered image, since replacing each pixel by a 4×4 array of dots, makes an image 16 times as large.

$$\begin{pmatrix} 0 & 8 & 2 & 10 \\ 12 & 4 & 14 & 6 \\ 3 & 11 & 1 & 9 \\ 15 & 7 & 13 & 5 \end{pmatrix}$$

3. 24-Bit Color Images

In a color 24 - bit image, each pixel is represented by three bytes, usually representing RGB (Red, Green, Blue). They are also called RGB images. When you read a color image with a size M x N, you will get a matrix with a dimension M x N x 3, ie a we get a 3 dimensional matrix. Each ‘plane’ in the matrix corresponds to one of the colors Red, Green and Blue (hence the name RGB images). 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 640x480 24-bit color image (((640 * 480 *24)/8)/1024) would require 900 kilobytes of storage without any compression.

In Figure (17 a) we see a representation of a typical RGB color image, and Figure (17 b) illustrate that in addition to referring to arrow or column as a vector, we can refer to a single-pixel red, green, and blue values as a color pixel vector (R,G,B).



Figure (17) :A color pixel vector consists of the red, green and blue pixel values (R, G, B) at one given row/column pixel coordinate(r , c)

4. The Indexed image

An indexed image consists of an array and a colormap matrix. An indexed image uses direct mapping of pixel values in the array to colormap values. By convention, this documentation uses the variable name `X` to refer to the array and `map` to refer to the colormap.

The colormap matrix is an m -by-3 array of class `double` containing floating-point values in the range $[0,1]$. Each row of `map` specifies the red, green, and blue components of a single color.

The pixel values in the array are direct indices into a colormap. The color of each image pixel is determined by using the corresponding value of `X` as an index into `map`. The relationship between the values in the image matrix and the colormap depends on the class of the image matrix:

- If the image matrix is of class `single` or `double`, the colormap normally contains integer values in the range $[1, p]$, where p is the length of the colormap. The value 1 points to the first row in the colormap, the value 2 points to the second row, and so on.
- If the image matrix is of class `logical`, `uint8` or `uint16`, the colormap normally contains integer values in the range $[0, p-1]$. The value 0 points to the first row in the colormap, the value 1 points to the second row, and so on.

A colormap is often stored with an indexed image and is automatically loaded with the image when you use the `imread` function. After you read the image and the colormap into the workspace as separate variables, you must keep track of the association between the image and colormap. However, you are not limited to using the default colormap—you can use any colormap that you choose.

The following figure (18) illustrates the structure of an indexed image. In the figure, the image matrix is of class double, so the value 5 points to the fifth row of the colormap.

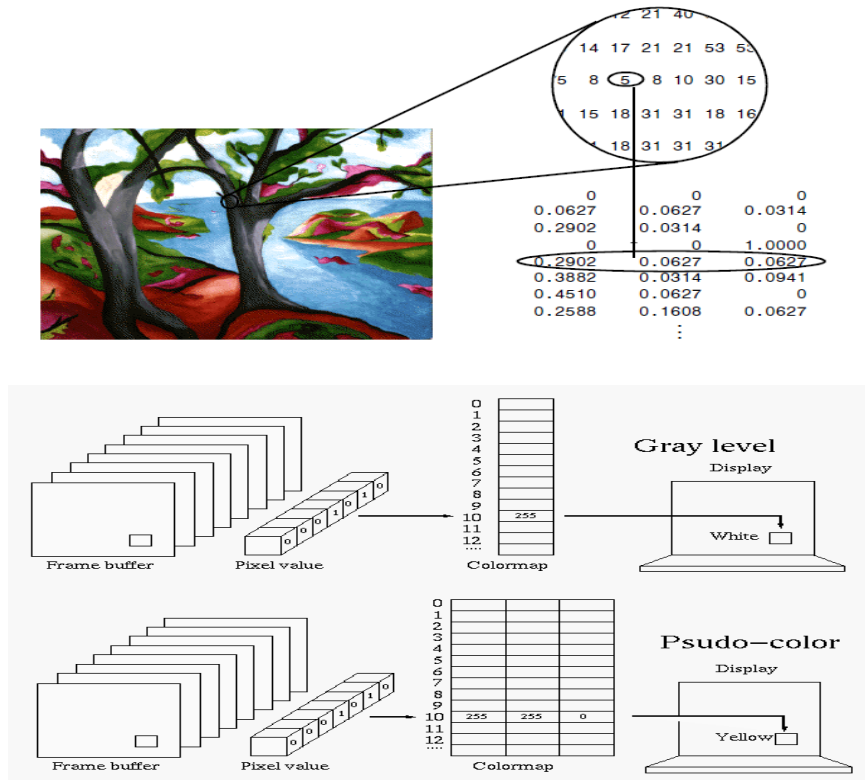


Figure (18) :Pixel Values Index to Colormap Entries in an Indexed Image

Example 1:

A binary image of (512× 512) dimensions. The number of bits needs to store the file size of this image is as follows:

- File size of binary image = $512 * 512 = 262, 144$ bits or
 = 32,768 Byte
 = 32 KB
 = 0.03125 MB
- File size of gray-scale image = $512 * 512 * 8 = 2,097,152$ bits or
 = 262,144 Byte
 = 256 KB

= 0.25 MB

- File size of RGB image = $512 * 512 * 8 * 3 = 6,291,456$ bits or
= 786,432 Byte
= 786 KB
= 0.75 MB

12. Image File Types

Some popular file formats for information exchange are described below.

Format Name	Description	Recognized Extensions
GIF	Graphics Interchange Format	.gif
JPEG	Joint Photographic Experts Group	.jpg, .jpeg
TIFF	Tagged Image File Format	.tif, .tiff
BMP	Windows Bitmap	.bmp
PNG	Portable Network Graphics	.png

Graphics Interchange Format (GIF):

GIF is a **bitmap lossless image file format** developed by Steve Wilhite in 1987 and commonly used in the World Wide Web due to its portability. **GIF format is useful for black and white, grayscale images and color images having less than 256 colors.** This format also **supports animations** and allows a separate palette of upto 256 colors for each frame. This format is well suited to simpler images such as graphics or logos with solid areas of color but, due to palette limitations, **not very suitable for color photographs and other images with color gradients.** Most color images are 24 bits per pixel and hence can't be stored as GIF. To store such images in GIF format, the image must first be converted into an 8-bit image. Although GIF format is a **lossless format** for color images with less than 256 colors, a rich truecolor image may "lose" 99.998% of the colors.

JPEG

JPEG is short for **Joint Photographic Experts Group**, and is the most **popular among the image formats used on the web**. JPEG files are very **'lossy' compression**, meaning so much information is lost from the original image when you save it in a JPEG file.

This is because **JPEG discards most of the information to keep the image file size small; which means some degree of quality is also lost**.

Almost every digital camera can shoot and save in the JPEG format. JPEG is very web friendly because the file is smaller, which means it takes up less room, and requires less time to transfer to a site. Moreover, it is less grainy than GIF, the old king of the internet roost. Since 1994, JPEG has been considered the standard.

Pros of JPEG:

- 24-bit color, with up to 16 million colors.
- Rich colors, great for photographs that need fine attention to color detail.
- Most used and most widely accepted image format.
- Compatible in most OS (Mac, PC, Linux).

Cons of JPEG:

- They tend to discard a lot of data.
- After compression, JPEG tends to create artifacts.
- Cannot be animated.
- Does not support transparency.

TIFF

TIFF is large raster file that doesn't lose quality **that stands for Tagged Image File Format**. Because of its extremely high quality, the format is primarily used in photography and desktop publishing. This file

type is known for using "**lossless compression**" meaning the original image data is maintained regardless of how often you might copy, re-save, or compress the original file.

Despite TIFF images' ability to recover their quality after manipulation, you should avoid using this file type on the web, it can take forever to load. TIFF files are also commonly used when saving photographs for print.

BMP

The **BMP** file format (Windows bitmap) handles graphic files within the Microsoft Windows OS. These files are large and **uncompressed**, but the images are rich in color, high in quality, simple and compatible in all Windows OS and programs. *BMP files are also called raster or paint images.*

BMP files are made of millions and millions of dots called 'pixels,' with different colors and arrangements to come up with an image or pattern. It **might be an 8-bit, 16-bit or 24-bit image**. BMP files are not great and not very popular. Being oversized, **bitmap files are not what you call 'web friendly', nor are they compatible in all platforms** and they do not scale well.

Pros of BMP:

- Works well with most Windows programs and OS, you can use it as a Windows wallpaper

Cons of BMP:

- Does not compress or scale well
- Again, very huge image files making it not web friendly

PNG

PNG is a **lossless raster format that stands for Portable Network Graphics**. Think of PNGs as the **next-generation GIF**. This format has

built-in transparency, but **can also display higher color depths**, which translates into millions of colors. PNGs are a **web standard** and are quickly becoming one of the most common image formats used online.