Unit_4
Multimedia Technology

Learning Objectives

- Multimedia
- Multimedia computer system
- Main components of multimedia and their associated technologies
- Common multimedia applications

Multimedia

- Media is something that can be used for presentation of information.
- Two basic ways to present some information are:
  - **Unimedia presentation:** Single media is used to present information
  - **Multimedia presentation:** More than one media is used to present information
- Multimedia presentation of any information greatly enhances the comprehension capability of the user as it involves use of more of our senses

Common Media

- Common media for storage, access, and transmission of information are:
  - Text (alphanumeric characters)
  - Graphics (line drawings and images)
  - Animation (moving images)
  - Audio (sound)
  - Video (Videographed real-life events)
- Multimedia in information technology refers to use of more than one of these media for information presentation to users

Multimedia Computer System

- Multimedia computer system is a computer having capability to integrate two or more types of media (text, graphics, animation, audio, and video)
- In general, size for multimedia information is much larger than plain text information
- Multimedia computer systems require:
  - Faster CPU
  - Larger storage devices (for storing large data files)
  - Larger main memory (for large data size)
  - Good graphics terminals
  - I/O devices to play any multimedia

Text Media

- Alphanumeric characters are used to present information in text form. Computers are widely used for text processing
- Keyboards, OCRs, computer screens, and printers are some commonly used hardware devices for processing text media
- Text editing, text searching, hypertext, and text importing/exporting are some highly desirable features of a multimedia computer system for better presentation and use of text information

Graphics Media

- *Computer graphics* deals with generation, representation, manipulation, and display of pictures (line drawings and images) with a computer
- Locating devices (such as a mouse, a joystick, or a stylus), digitizers, scanners, digital cameras, computer screens with graphics display capability, laser printers, and plotters are some common hardware devices for processing graphics media
- Some desirable features of a multimedia computer system are painting or drawing software, screen capture software, clip art, graphics importing, and software support for high resolution

Animation Media

- *Computer animation* deals with generation, sequencing, and display (at a specified rate) of a set of images (called frames) to create an effect of visual change or motion, similar to a movie film (video)
- Animation is commonly used in those instances where videography is not possible or animation can better illustrate the concept than video
- Animation deals with displaying a sequence of images at a reasonable speed to create an impression of movement. For a jerk-free full motion animation, 25 to 30 frames per second is required

Animation Media

Audio Media

- *Computer audio* deals with synthesizing, recording, and playback of audio or sound with a computer
- Sound board, microphone, speaker, MIDI devices, sound synthesizer, sound editor and audio mixer are some commonly used hardware devices for processing audio media
- Some desirable features of a multimedia computer system are audio clips, audio file importing, software support for high quality sound, recording and playback capabilities, text-to-speech conversion software, speech-to-text conversion software, and voice recognition software

Video Media

- *Computer video* deals with recording and display of a sequence of images at a reasonable speed to create an impression of movement. Each individual image of such a sequence is called a frame
- Video camera, video monitor, video board, and video editor are some of the commonly used hardware devices for processing video media
- Some desirable features of a multimedia computer system with video facility are video clips and recording and playback capabilities

Multimedia Applications

- Multimedia presentation
- Foreign language learning
- Video games
- Special effects in films
- Multimedia kiosks as help desks
- Animated advertisements
- Multimedia conferencing

Media Center Computer

- There is a growing trend of owning a personal computer (PC) at home like other electronic equipment
- New terminologies like "infotainment" and "edutainment" have evolved to refer to computers as versatile tools
- Media center PC provides following functionalities:
  - Server as PC, TV, radio, and music system
  - Serve as digital photo album and digital library
  - Server as Game station and DVD/CD Player
  - Allows play, pause, and record of TV programs
  - Provides Electronic Programming Guide (EPG)

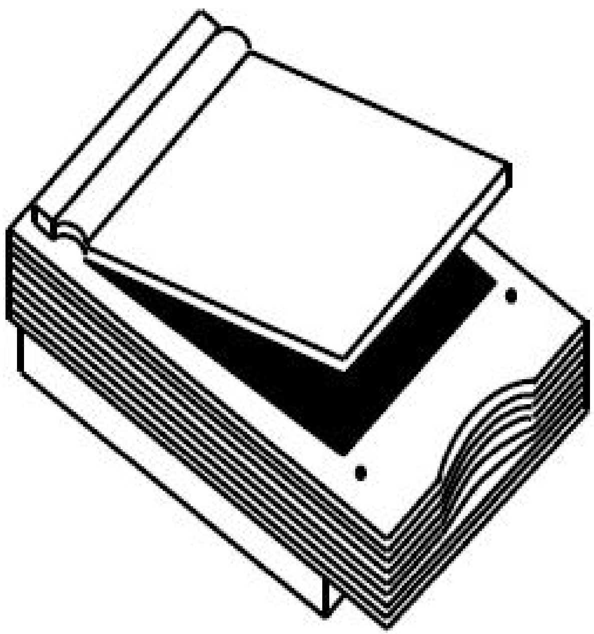
Data Scanning Devices

- Input devices that enable direct data entry into a computer system from source documents
- Eliminate the need to key in text data into the computer
- Due to reduced human effort in data entry, they improve data accuracy and also increase the timeliness of the information processed
- Demand high quality of input documents
- Some data scanning devices are also capable of recognizing marks or characters
- Form design and ink specification usually becomes more critical for accuracy

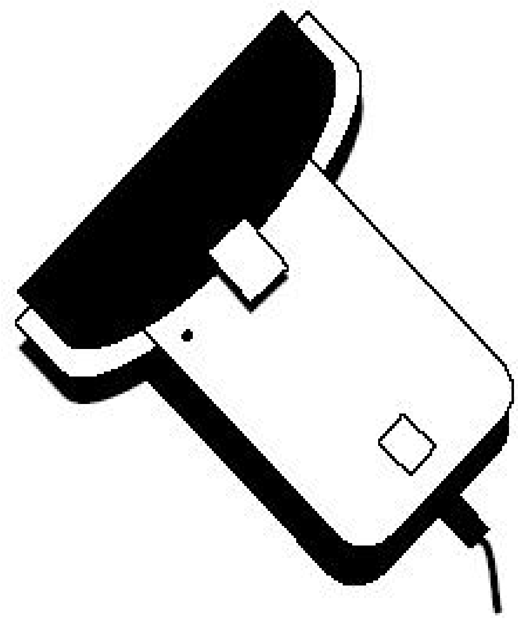
Image Scanner

- Input device that translates paper documents into an electronic format for storage in a computer
- Electronic format of a scanned image is its bit map representation
- Stored image can be altered or manipulated with an image-processing software

Two Common Types of Image Scanners



A flat-bed scanner



A hand-held scanner

Optical Character Recognition (OCR) Device

- Scanner equipped with a character recognition software (called OCR software) that converts the bit map images of characters to equivalent ASCII codes
- Enables word processing of input text and also requires less storage for storing the document as text rather than an image
- OCR software is extremely complex because it is difficult to make a computer recognize an unlimited number of typefaces and fonts
- Two standard OCR fonts are OCR-A (American standard) and OCR-B (European standard)

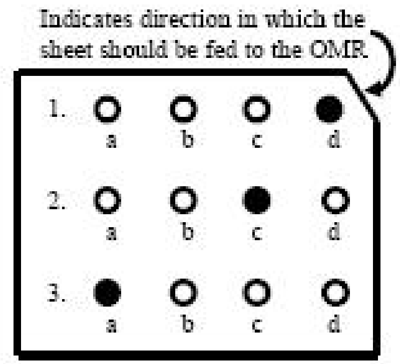
Optical Mark Reader (OMR)

- Scanner capable of recognizing a pre-specified type of mark by pencil or pen
- Very useful for grading tests with objective type questions, or for any input data that is of a choice or selection nature
- Technique used for recognition of marks involves focusing a light on the page being scanned and detecting the reflected light pattern from the marks

Sample Use of OMR

For each question, four options are given out of which only one is correct. Choose the correct option and mark your choice against the corresponding question number in the given answer sheet by darkening the corresponding circle with a lead pencil.

1. The binary equivalent of decimal 4 is:
   - a) 101
   - b) 111
   - c) 001
   - d) 100
2. The full form of CPU is:
   - a) Cursor Positioning Unit
   - b) Central Power Unit
   - c) Central Processing Unit
   - d) None of the above
3. Which is the largest unit of storage among the following:
   - a) Terabyte
   - b) Kilobyte
   - c) Megabyte
   - d) Gigabyte



(b) Pre-printed answer sheet

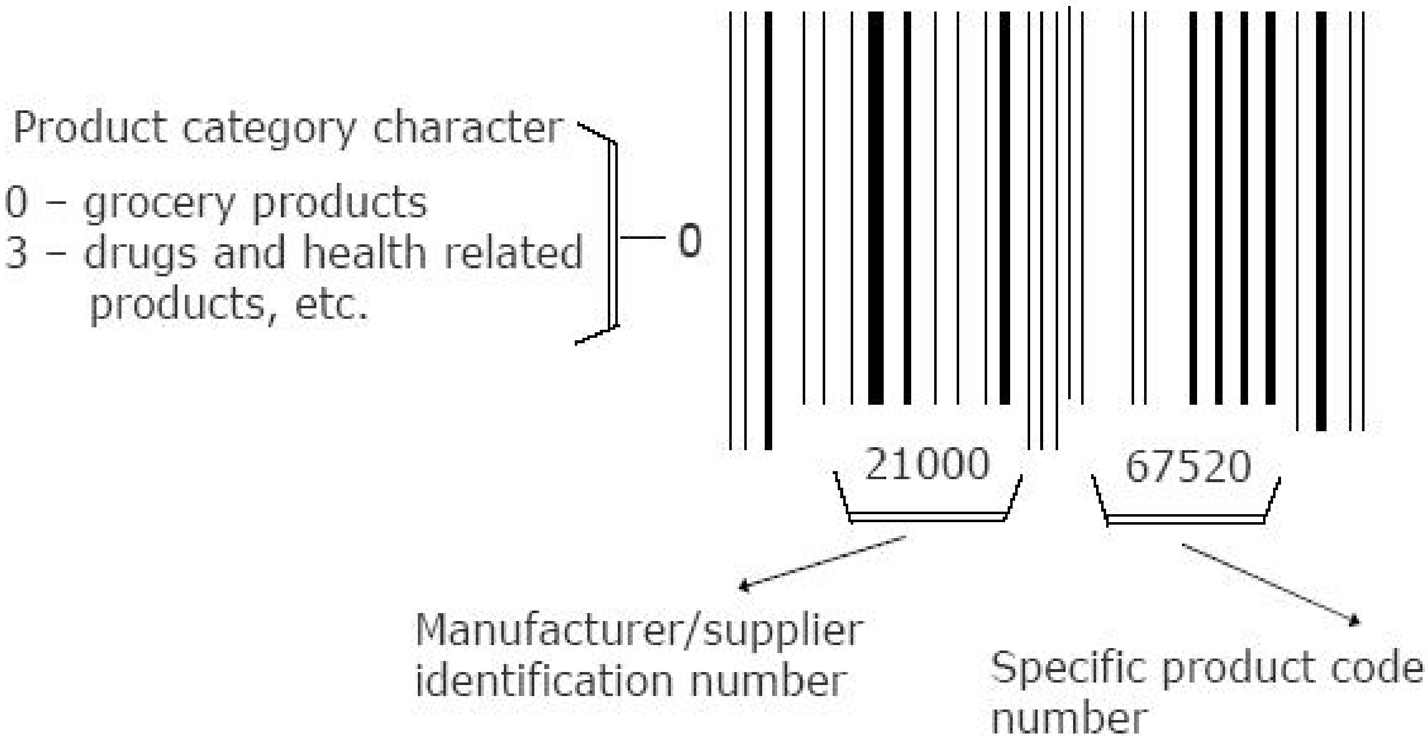
(a) Question sheet

A sample use of OMR for grading tests with objective type questions

Bar-code Reader

- Scanner used for reading (decoding) bar-coded data
- Bar codes represent alphanumeric data by a combination of adjacent vertical lines (bars) by varying their width and the spacing between them
- Scanner uses laser-beam to stroke across pattern of bar code. Different patterns of bars reflect the beam in different ways sensed by a light-sensitive detector
- Universal Product Code (UPC) is the most widely known bar coding system

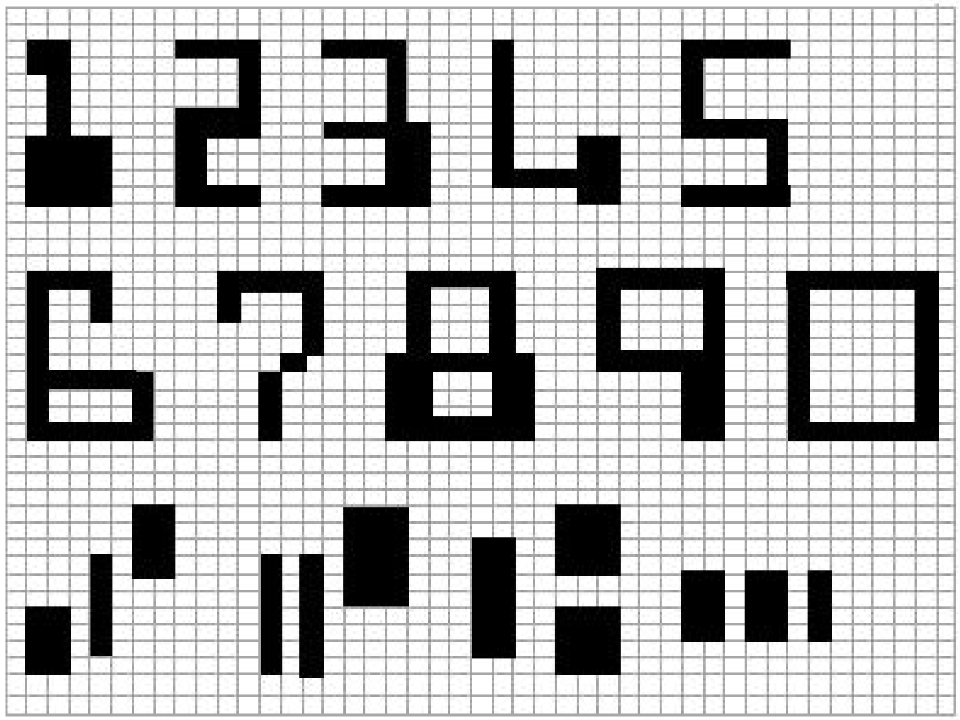
An Example of UPC Bar Code



Magnetic-Ink Character Recognition (MICR)

- MICR is used by banking industry for faster processing of large volume of cheques
- Bank's identification code (name, branch, etc.), account number and cheque number are pre-printed (encoded) using characters from a special character set on all cheques
- Special ink is used that contains magnetizable particles of iron oxide
- MICR reader-sorter reads data on cheques and sorts them for distribution to other banks or for further processing

MICR Character Set (E13B Font)

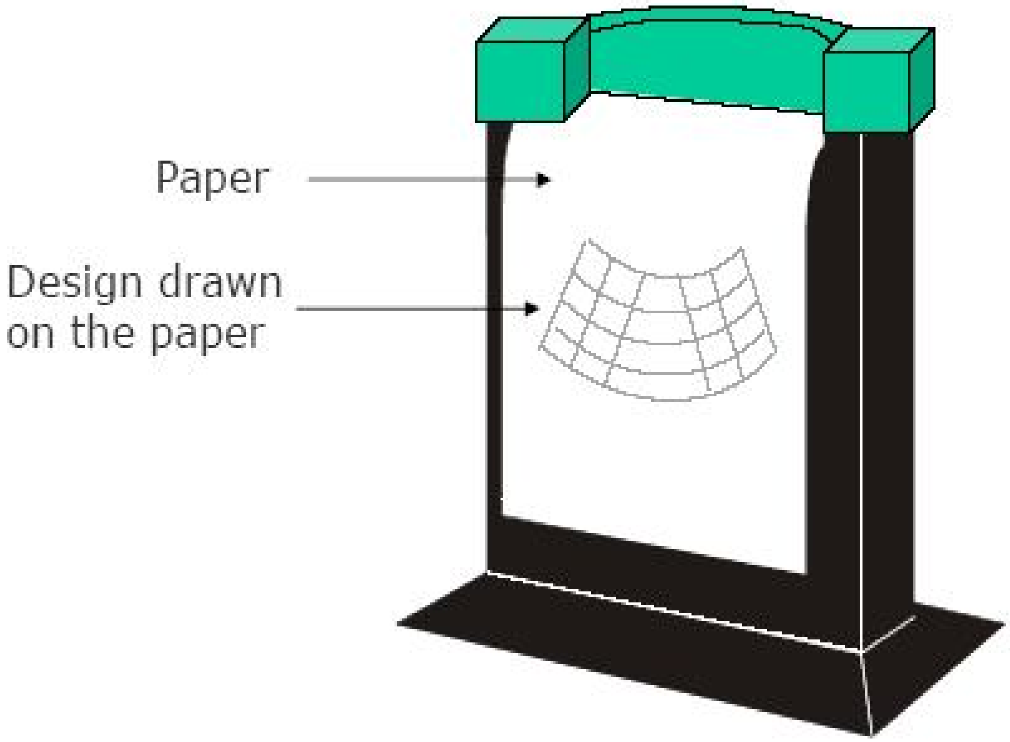


- It consists of numerals 0 to 9 and four special characters
- MICR is not adopted by other industries because it supports only 14 symbols

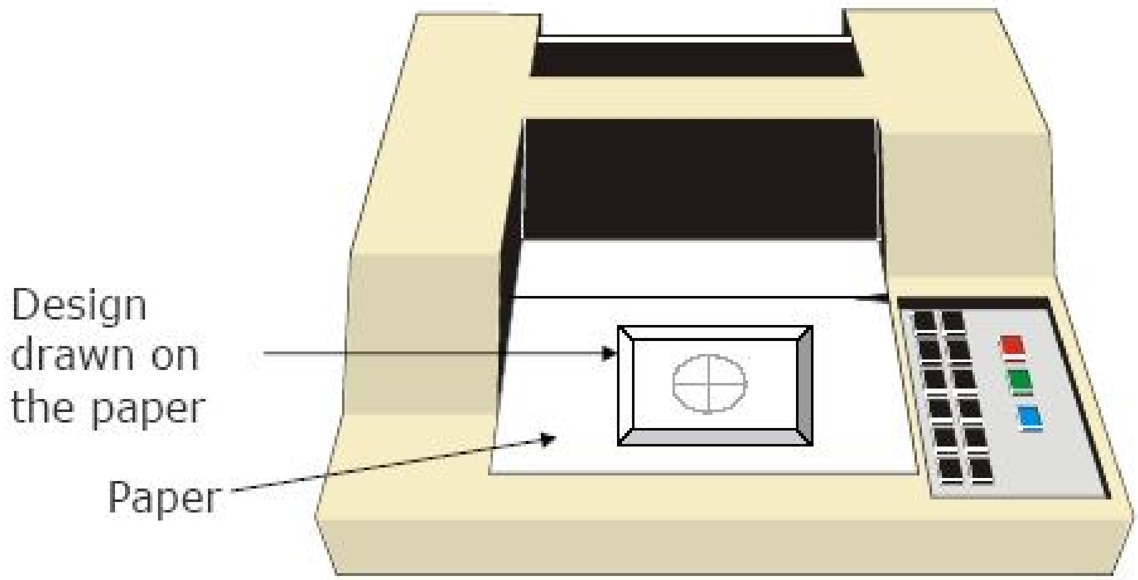
Plotters

- Plotters are an ideal output device for architects, engineers, city planners, and others who need to routinely generate high-precision, hard-copy graphic output of widely varying sizes
- Two commonly used types of plotters are:
  - *Drum plotter*, in which the paper on which the design has to be made is placed over a drum that can rotate in both clockwise and anti-clockwise directions
  - *Flatbed plotter*, in which the paper on which the design has to be made is spread and fixed over a rectangular flatbed table

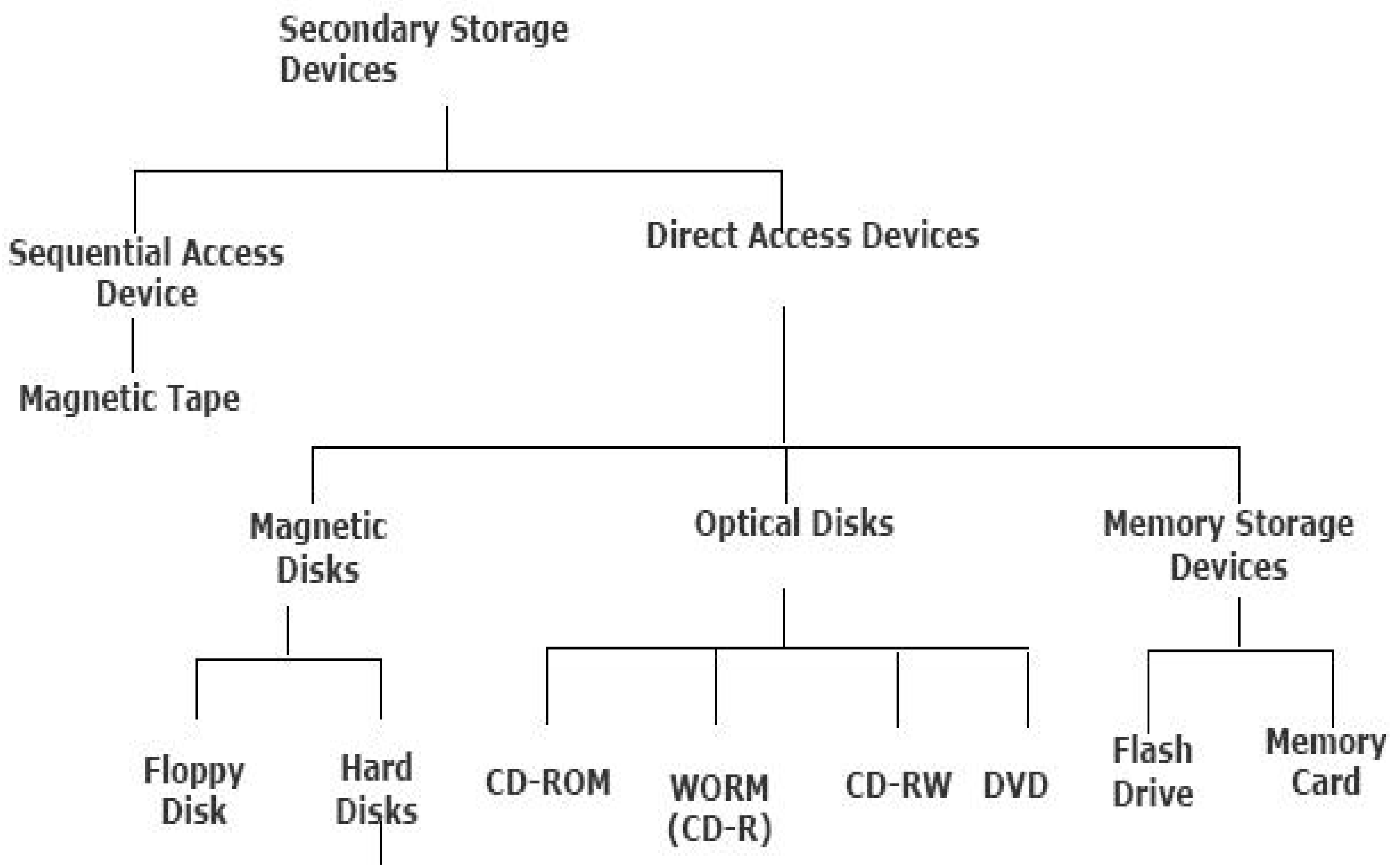
A Drum Plotter



A Flatbed Plotter



Classification of Commonly Used Secondary Storage Devices



Sequential-access Storage Devices

- Arrival at the desired storage location may be preceded by sequencing through other locations
- Data can only be retrieved in the same sequence in which it is stored
- Access time varies according to the storage location of the information being accessed
- Suitable for sequential processing applications where most, if not all, of the data records need to be processed one after another
- Magnetic tape is a typical example of such a storage device

Direct-access Storage Devices

- Devices where any storage location may be selected and accessed at random
- Permits access to individual information in a more direct or immediate manner
- Approximately equal access time is required for accessing information from any storage location
- Suitable for direct processing applications such as on-line ticket booking systems, on-line banking systems
- Magnetic, optical, and magneto-optical disks are typical examples of such a storage device

Optical Disk – Basics

- Consists of a circular disk, which is coated with a thin metal or some other material that is highly reflective
- Laser beam technology is used for recording/reading of data on the disk
- Also known as laser disk / optical laser disk, due to the use of laser beam technology
- Proved to be a promising random access medium for high capacity secondary storage because it can store extremely large amounts of data in a limited space

Optical Disk – Storage Organization

- Has one long spiral track, which starts at the outer edge and spirals inward to the center
- Track is divided into equal size sectors

(a) Track pattern on an optical disk (b) Track pattern on a magnetic disk

Difference in track patterns on optical and magnetic disks.

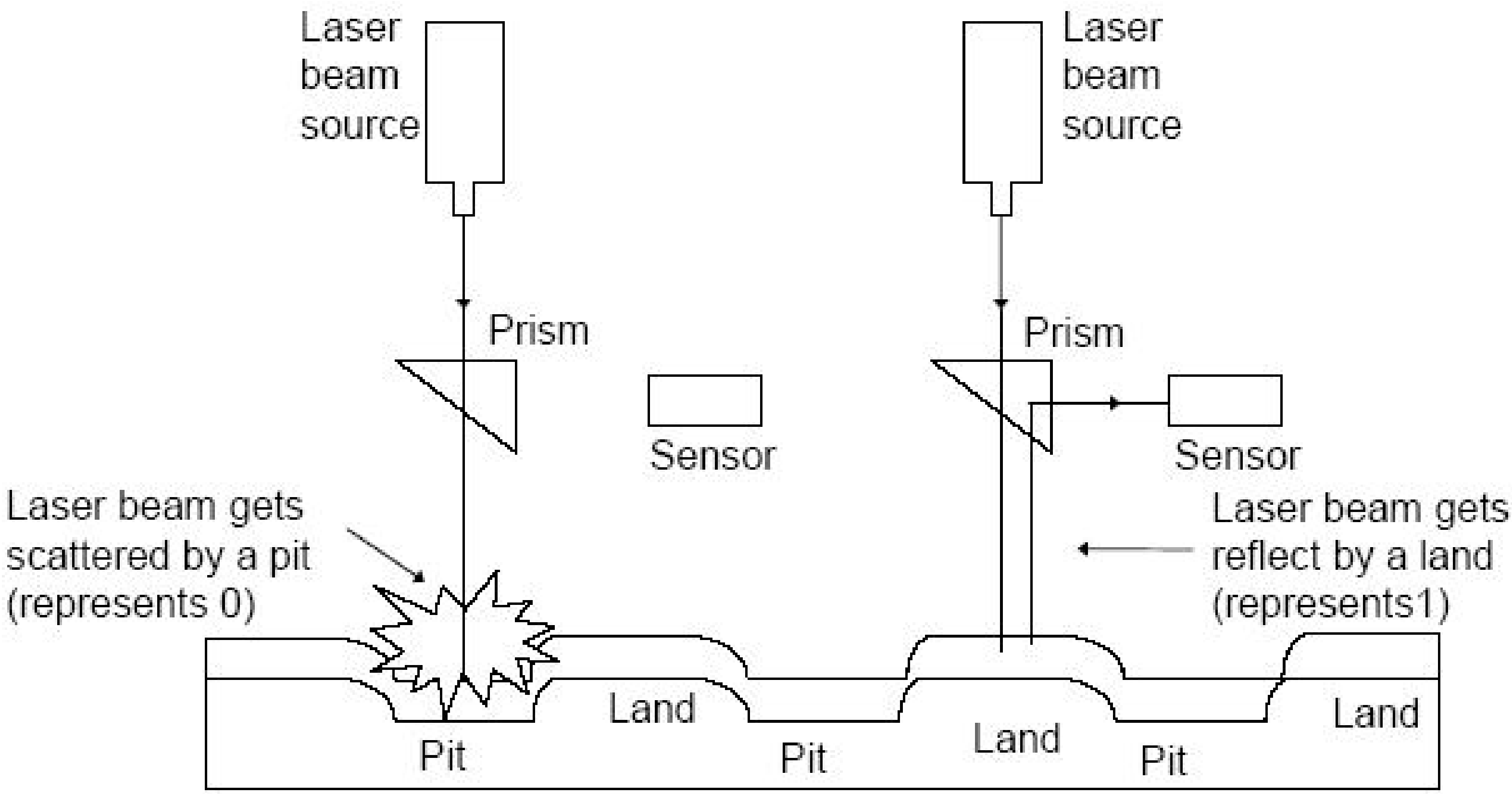
Optical Disk – Storage Capacity

Storage capacity of an optical disk

= Number of sectors
× Number of bytes per sector

The most popular optical disk uses a disk of 5.25 inch diameter with storage capacity of around 650 Megabytes

Optical Disk – Access Mechanism



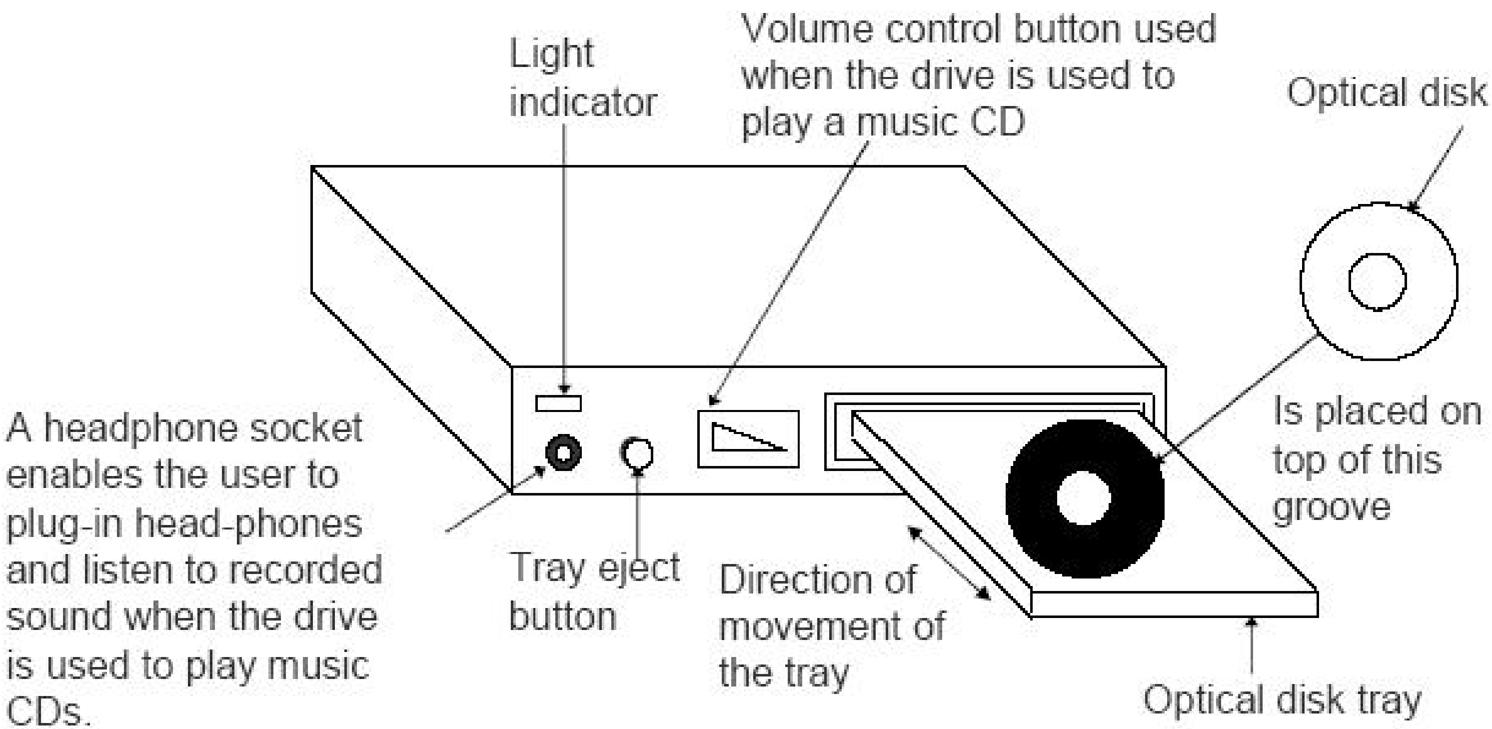
Optical Disk – Access Time

- With optical disks, each sector has the same length regardless of whether it is located near or away from the disk's center
- Rotation speed of the disk must vary inversely with the radius. Hence, optical disk drives use a constant linear velocity (CLV) encoding scheme
- Leads to slower data access time (larger access time) for optical disks than magnetic disks
- Access times for optical disks are typically in the range of 100 to 300 milliseconds and that of hard disks are in the range of 10 to 30 milliseconds

Optical Disk Drive

- Uses laser beam technology for reading/writing of data
- Has no mechanical read/write access arm
- Uses a constant linear velocity (CLV) encoding scheme, in which the rotational speed of the disk varies inversely with the radius

Optical Disk Drive



Types of Optical Disks

The types of optical disks in use today are:

CD-ROM

- Stands for Compact Disk-Read Only Memory
- Packaged as shiny, silver color metal disk of 5¼ inch (12cm) diameter, having a storage capacity of about 650 Megabytes
- Disks come pre-recorded and the information stored on them cannot be altered
- Pre-stamped (pre-recorded) by their suppliers, by a process called *mastering*

Types of Optical Disks

WORM Disk / CD-Recordable (CD-R)

- Stands for Write Once Read Many. Data can be written only once on them, but can be read many times
- Same as CD-ROM and has same storage capacity
- Allow users to create their own CD-ROM disks by using a CD-recordable (CD-R) drive that can be attached to a computer as a regular peripheral device
- Data to be recorded can be written on its surface in multiple recording sessions

Types of Optical Disks

CD-Read/Write (CD-RW)

- Same as CD-R and has same storage capacity
- Allow users to create their own CD-ROM disks by using a CD-recordable (CD-R) drive that can be attached to a computer as a regular peripheral device
- Data to be recorded can be written on its surface in multiple recording sessions
- Made of metallic alloy layer whose chemical properties are changed during burn and erase
- Can be erased and written afresh

Types of Optical Disks

Digital Video / Versatile Disk (DVD)

- Looks same as CD-ROM but has capacity of 4.7 GB or 8.5 GB
- Designed primarily to store and distribute movies
- Can be used for storage of large data
- Allows storage of video in 4:3 or 16:9 aspect-ratios in MPEG-2 video format using NTSC or PAL resolution
- Audio is usually Dolby® Digital (AC-3) or Digital Theater System (DTS) and can be either monaural or 5.1 Surround Sound

Advantages of Optical Disks

- The cost-per-bit of storage for optical disks is very low because of their low cost and enormous storage density.
- The use of a single spiral track makes optical disks an ideal storage medium for reading large blocks of sequential data, such as music.
- Optical disk drives do not have any mechanical read/write heads to rub against or crash into the disk surface. This makes optical disks a more reliable storage medium than magnetic tapes or magnetic disks.
- Optical disks have a data storage life in excess of 30 years. This makes them a better storage medium for data archiving as compared to magnetic tapes or magnetic disks.

Advantages of Optical Disks

- As data once stored on an optical disk becomes permanent, danger of stored data getting inadvertently erased/overwritten is removed
- Due to their compact size and light weight, optical disks are easy to handle, store, and port from one place to another
- Music CDs can be played on a computer having a CD-ROM drive along with a sound board and speakers. This allows computer systems to be also used as music systems

Limitations of Optical Disks

- It is largely read-only (permanent) storage medium. Data once recorded, cannot be erased and hence the optical disks cannot be reused
- The data access speed for optical disks is slower than magnetic disks
- Optical disks require a complicated drive mechanism

Categories of multimedia

- Non-Linear or Interactive multimedia:-non-liner contents offers user interactivity to control progress.
- Liner or Non-Interactive :- without any navigation control for the viewer
- Non-linear content is also known as hypermedia content.

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

Media types can be divided in two types depending on their behaviour with respect to time.

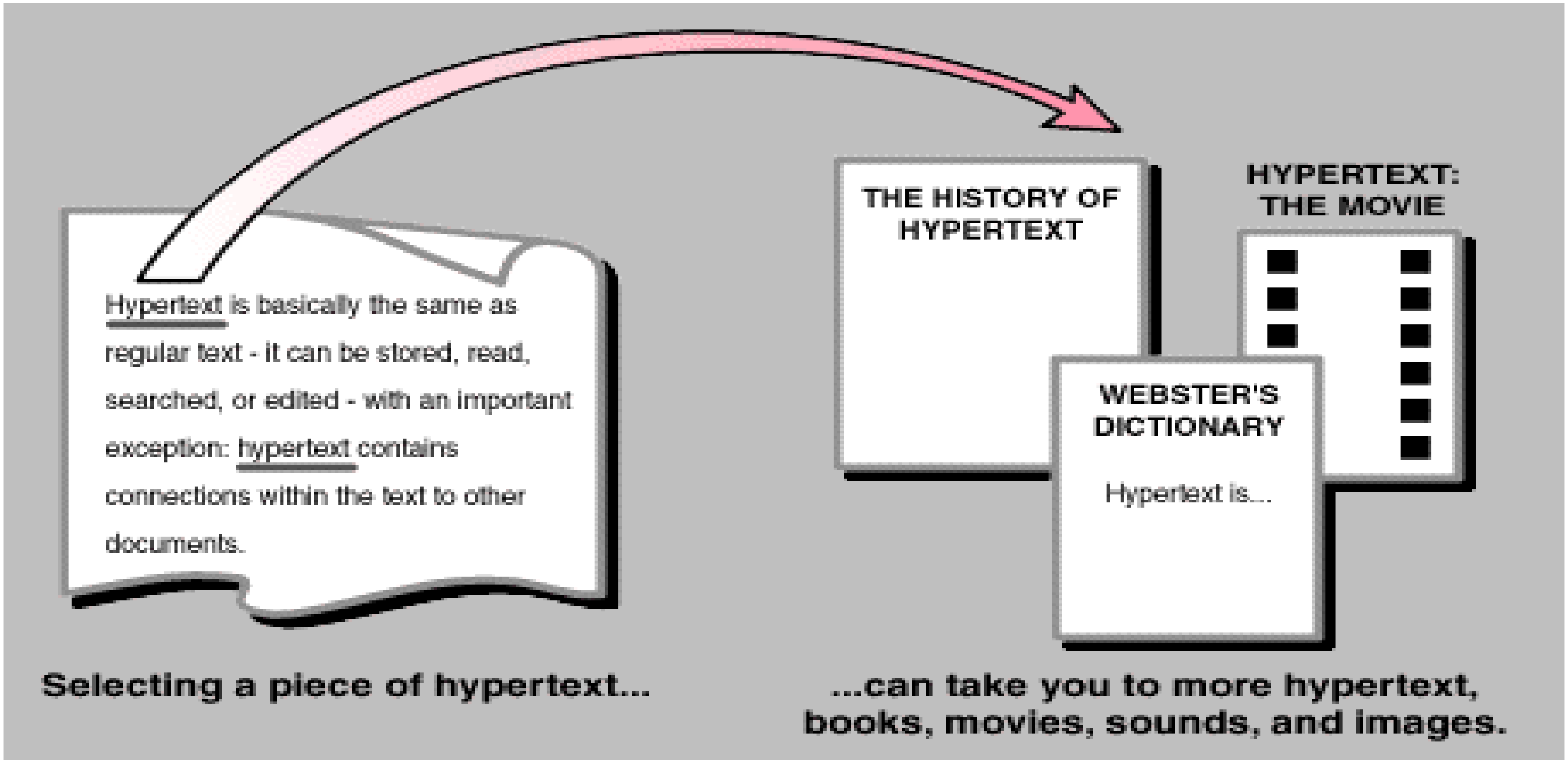
- Temporal/Dynamic/time-based :- these types change with time
  - audio
  - video
  - music
  - animation

- Non- temporal/static/non-time based :-
  remain the same with time
    - text
    - images
    - graphics

Hypertext

- **Hypertext** is text, displayed on a computer, with references to other text that the reader can immediately access, usually by a mouse click or keypress sequence.
- May contain tables, images and other presentational devices.

- Other means of interaction may also be present, such as a bubble with text appearing when the mouse over a particular area, a video clip starting, or a form to complete and submit.
- The most extensive example is the World Wide Web.


Hypertext is basically the same as regular text - it can be stored, read, searched, or edited - with an important exception: hypertext contains connections within the text to other documents.
THE HISTORY OF HYPERTEXT
HYPERTEXT: THE MOVIE
WEBSTER'S DICTIONARY
Hypertext is...
Selecting a piece of hypertext...
...can take you to more hypertext, books, movies, sounds, and images.

Types and uses of hypertext

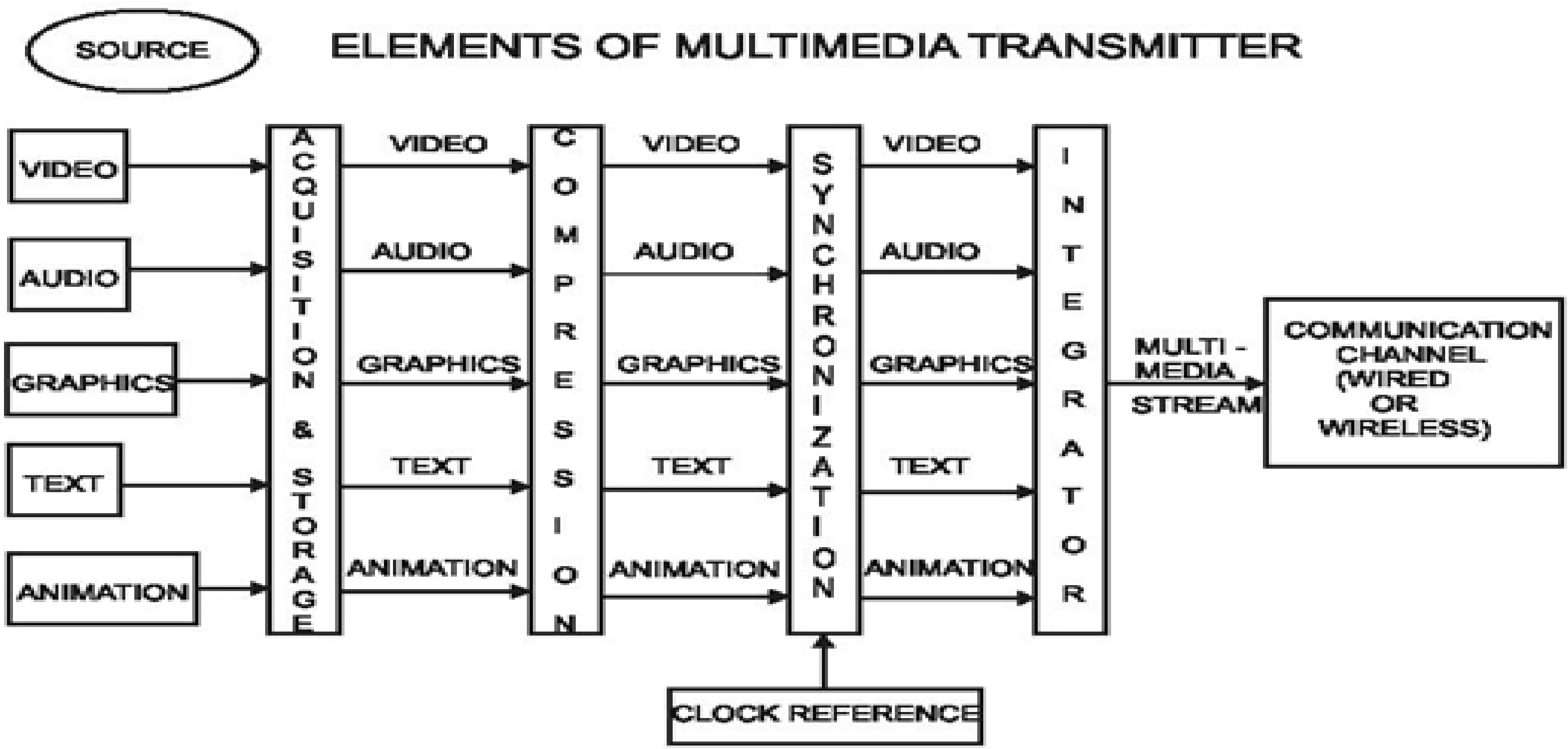
- Static (prepared and stored in advance) :- Static hypertext can be used to cross-reference collections of data in documents, software applications, or books on CDs.
- Dynamic (continually changing in response to user input). A well-constructed system can also incorporate other user-interface conventions, such as menus and command lines.

Hypermedia

- Hypermedia, a term derived from hypertext, extends the notion of the hypertext link to include links among any set of multimedia objects, including sound, motion video, and virtual reality.
- It can also indicate a higher level of user/network interactivity than the interactivity already implicit in hypertext.

Elements of multimedia communication systems

- We have audio, video, graphics, texts - all these media as the sources, but first and foremost, we need a system which can acquire the separate media streams, process them together to make it an integrated multimedia stream.


SOURCE
ELEMENTS OF MULTIMEDIA TRANSMITTER
VIDEO
AUDIO
GRAPHICS
TEXT
ANIMATION
ACQUISITION & STORAGE
COMPRESSION
SYNCHRONIZATION
INTEGRATOR
MULTI - MEDIA STREAM
COMMUNICATION CHANNEL (WIRED OR WIRELESS)
CLOCK REFERENCE

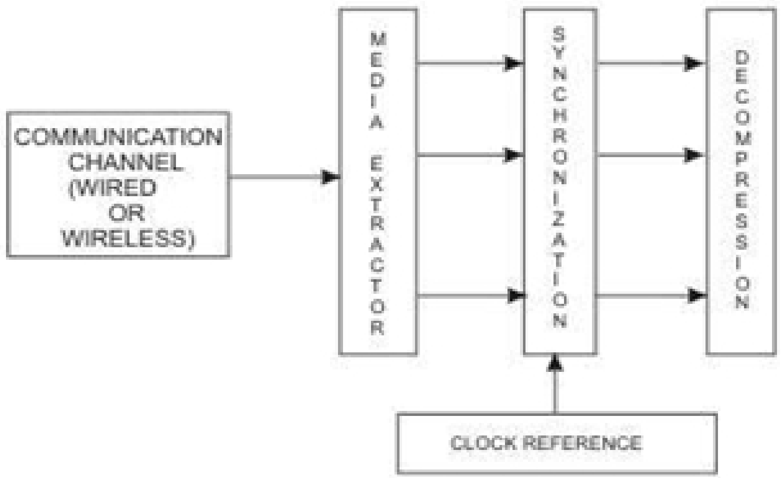


Fig 1.2 Elements of multimedia Receiver

- The media extractor separates the integrated media stream into individual media streams, which undergoes de-compression and then presented in a synchronized manner according to their time-stamps in different playback units, such as monitors, loudspeakers, printers/plotters, recording devices etc.

What is compression:--

- Compression is the process of reducing the size of a file by encoding its data information more efficiently.

How does Compression work?

- Compression is done by using compression algorithms that rearrange and reorganize data information so that it can be stored more economically.

Lossless Compression vs. Lossy Compression

- **Lossless Compression** is a type of compression that can reduce files without a loss of information in the process. The original file can be recreated exactly when uncompressed.
- Lossless compression is ideal for documents containing text and numerical data where any loss of textual information can't be tolerated.

- **Lossy Compression,** on the other hand, reduces the size of a file by eliminating bits of information.
- It permanently deletes any unnecessary data.
- This compression is usually used with images, audio and graphics where a loss of quality is affordable. However, the original file can’t be retained.

Lossless

- recover the exact original data after compression.
- mainly use for compressing database records, spreadsheets or word processing files, where exact replication of the original is essential.

lossy

- will result in a certain loss of accuracy in exchange for a substantial increase in compression.
- more effective when used to compress graphic images and digitised voice where losses outside visual or aural perception can be tolerated.
- Most lossy compression techniques can be adjusted to different quality levels, gaining higher accuracy in exchange for less effective compression.

Terminology

- Compressor-Software (or hardware) device that compresses data
- Decompressor-Software (or hardware) device that decompresses data
- Codec-Software (or hardware) device that compresses and decompresses data
- Algorithm-The logic that governs the compression/decompression process

CLASSIFICATION OF MMS

ENTROPY ENCODING	**RUN-LENGTH CODING**	
	HUFFMAN CODING	
	ARITHMETIC CODING	
SOURCE CODING	**PREDICTION**	**DPCM(differential pulse modulation)**
		DM
	TRANSFORMATION	**FFT**
		DCT
	LAYERED CODING	**BIT POSITION**
		SUBSAMPLING
		SUB-BAND CODING
	VECTOR QUANTIZATION	
HYBRID CODING	**JPEG**	
	MPEG	
	H.261	
	DVI RTV,DVI PLV	

What is Compression Ratio

Compression ratio:

$$\frac{\text{Number of bytes in the original image}}{\text{Number of bytes in the compressed image}}$$

e.g. input image = 10 MB

output image = 5 MB

compression ratio = 2:1

- Example :Data:
- ABBAAAACDEAAABBBDDEEAAA...
- Count symbols in stream:
- Symbol A B C D E
- COUNT 15 7 6 6 5

Lossless and lossy image compression

- **ENTROPY ENCODING→** Used regardless of the media's specific characteristics.
- Lossless Compression frequently involves some form of entropy encoding
- Based on information theoretic techniques. According to Shannon, the entropy of an information source S is defined as:

$$H(S) = n = Zi \; Pi \; log2 \; 1/Pi$$

where P_i is the probability that symbol S_i in S will occur.

- **SOURCE CODING→** Lossy process
- Relation exists between the encoded data and decoded data
- The data streams are similar but not identical.
- **HYBRID CODING→** combination of both

Run-length Encoding

- This encoding method is frequently applied to images.
- It is a small compression component used in JPEG compression.

In this instance:

- Sequences of image elements $X_1, X_2, \ldots, X_n$ (Row by Row)
- Mapped to pairs $(c_1, l_1), (c_2, l_2), \ldots, (c_n, l_n)$ where c_i represent image intensity or color and l_i the length of the pixels

Example:

Original Sequence:

- 111122233333311112222
  can be encoded as:

- (1,4),(2,3),(3,6),(1,4),(2,4)

- DATA :--

  ABCCCCCCCCDDD

Huffman Coding

- Based on the frequency of occurrence of a data item (pixels or small blocks of pixels in images).
- Use a lower number of bits to encode more frequent data
- Codes are stored in a Code table
- Code table constructed for each image or a set of images.
- Code book plus encoded data must be transmitted to enable decoding.

- A bottom-up approach

1. Put all nodes in an OPEN list, keep it sorted at all times (e.g., ABCDE).
2. Repeat until the OPEN list has only one node left:

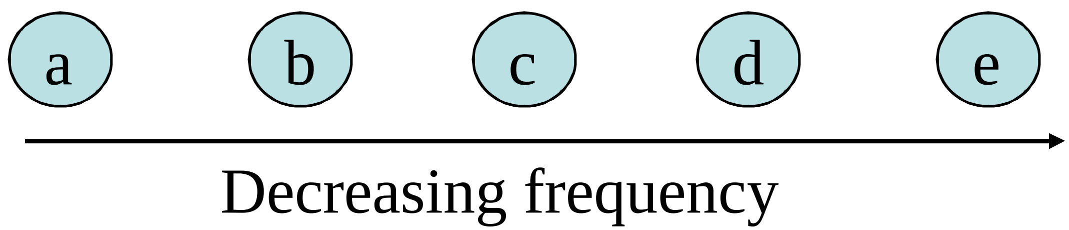
- From OPEN pick two nodes having the lowest frequencies/ probabilities, create a parent node of them.
- Assign the sum of the children's frequencies/probabilities to the parent node and insert it into OPEN
- Assign code 0, 1 to the two branches of the tree, and delete the children from OPEN.

- Although the number of bits required for less frequently increases rapidly, but there is a
- significant reduction in the number of bits required for the overall data due to the savings gained from the more-frequently-occurring symbols.
- Clearly this method requires at least one pass through the data to determine the symbol frequencies.

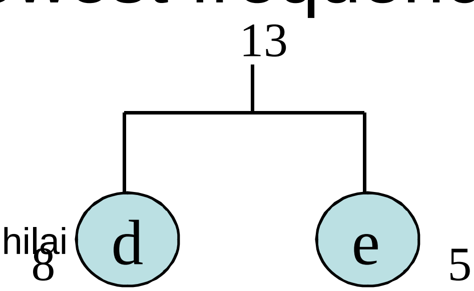
Huffman Coding

Symbol	a	b	c	d	e
Frequency	19	10	8	8	5

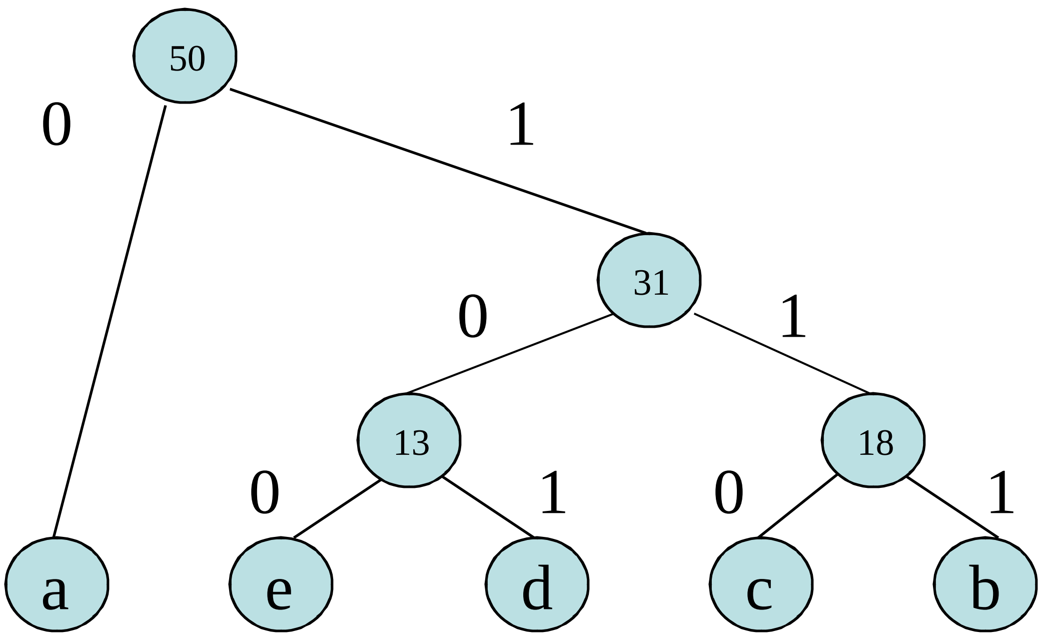
- 1. Show symbols as leaf nodes…



- 2. Combine nodes with lowest frequencies…



Huffman Coding II

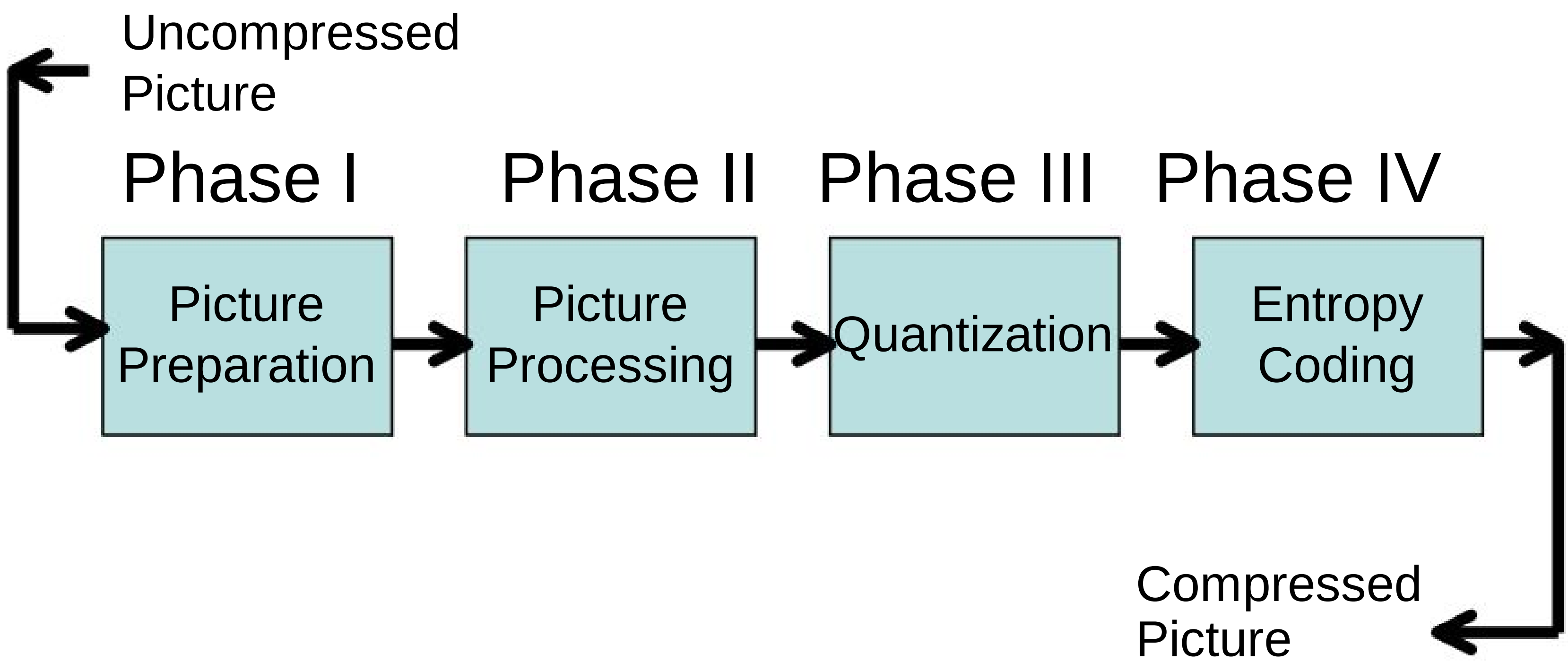


=

Symbol	Huffman Code
a	0
b	111
c	110
d	101
e	100

- Fewer bits per symbol – compression achieved

Basic Units of Encoder



- Picture preparation → includes analog-to digital conversion
- Generating an appropriate digital representation of the information
- Blocks of 8x8 pixels
- Represented by a fixed no. of bits per pixel.

- Processing→ first step
- Uses algorithms
- Transformation from time to frequency domain can be performed using DCT/FFT.

- Quantization process→ specify the mapping of real nos in to integers.
- Reduction of precision.
- Generates a limited number of symbols that can be used in the representation of the compressed image.
- Quantization is a many-to-one mapping which is irreversible.

- **ENTROPY ENCODING→** lossless process
- Used regardless of the media's specific characteristics.

Type of still image compression standards:

- (JPEG) Joint Photographic Experts Group

  a- Lossy copression of still images

  b- Lossless compression of still images
- (JBIG) Joint Bilevel Image Group
- (GIF) Graphics Interchange Format.
- (PNG) Portable Network Graphics.

JPEG

- **JPEG** is a commonly used for photographic images.
- JPEG typically achieves 10:1 compression.
- JPEG/Exif is the most common image format used by digital cameras and other photographic image capture devices;
- along with JPEG/JFIF, it is the most common format for storing and transmitting photographic images on the World Wide Web.

- The JPEG compression algorithm is best for photographs and paintings of realistic scenes with smooth variations of tone and color.
- For web usage, where the bandwidth used by an image is important, JPEG is very popular.
- JPEG/Exif is also the most common format saved by digital cameras.

- JPEG is *not* as well suited for line drawings and other textual or iconic graphics.
- JPEG is also not well suited to files that will undergo multiple edits,
- particularly if the image is cropped or shifted, or if encoding parameters are changed.

- It must not be used in astronomical or medical imaging or other purposes where the exact reproduction of the data is required.
- Lossless formats such as PNG must be used instead.

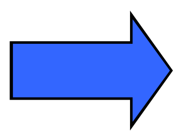
JPEG file extensions

- The most common filename extensions for files employing JPEG compression are
- **.jpg** and
- **.jpeg**, **.jpe**,
- **.jfif** and **.jif** are also used.

JPEG Compression: Basics

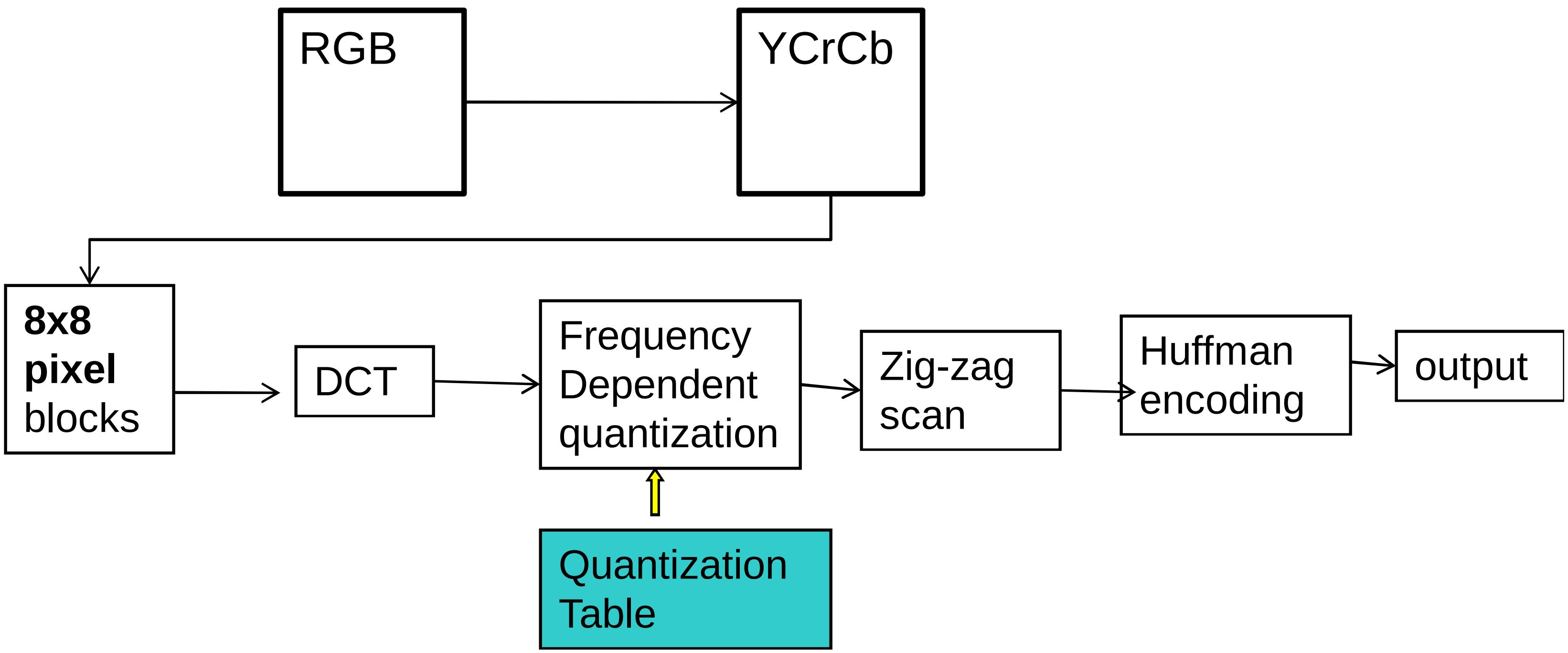
- JPEG can handle arbitrary color spaces RGB, CMYK, $YC_bC_{r,}$ or YUV.

Losslessly compressed image, ~150KB



JPEG compressed, ~14KB

Flow Chart of JPEG Compression Process



Color space transformation

- First, the image should be converted from RGB into a different color space called YCbCr.
- It has three components Y, Cb and Cr:
- the Y component represents the brightness or luminance of a pixel,
- the Cb and Cr components represent the chrominance.

Downsampling

- Human can see considerably more fine detail in the brightness of an image (the Y component) than in the color of an image (the Cb and Cr components).
- Using this knowledge, encoders can be designed to compress images more efficiently.
- The transformation into the YCbCr color model enables the next step, which is to reduce the spatial resolution of the Cb and Cr components
- called "downsampling" or "chroma subsampling".

Block splitting

- each channel must be split into 8×8 blocks of pixels.
- If the data for a channel does not represent an integer number of blocks then the encoder must fill the remaining area of the incomplete blocks with some form of dummy data.
- Filling the edge pixels with a fixed color (typically black)

Discrete cosine transform

- It helps to separate the images into parts of differing importance(w.r.t. the image quality)
- Tansforms a signal or image from the spatial domain to the frequency domain;
- Basic operation of DCT:
  - take input image N by M
  - g(i,j) is the intensity of the pixel in row i and column j
  - G(u,v) is the DCT cofficient in row u and column v.

$$G_{u,v} = \alpha(u)\alpha(v) \sum_{x=0}^{7} \sum_{y=0}^{7} g_{x,y} \cos\left[\frac{\pi}{8}\left(x + \frac{1}{2}\right)u\right] \cos\left[\frac{\pi}{8}\left(y + \frac{1}{2}\right)v\right]$$

As an example, one such 8×8 8-bit subimage might be:

128	128	128	128	128	128	128	128
118	111	112	117	120	123	123	122
125	121	115	111	119	119	118	117
120	121	113	113	125	124	115	108
120	120	116	119	124	120	115	110
117	113	111	122	120	110	116	119
109	113	111	122	120	110	116	119
111	121	124	118	115	121	117	113

DCT matrix

-80	4	-6	6	2	-2	-2	0
24	-8	8	12	0	0	0	2
10	-4	0	-12	-4	4	4	-2
8	0	-2	-6	10	4	-2	0
18	4	-4	6	-8	-4	0	0
-2	8	6	-4	0	-2	0	0
12	0	6	0	0	0	-2	-2
0	8	0	-4	-2	0	0	0

- Note the rather large value of the top-left corner. This is the DC coefficient.
- The remaining 63 coefficients are called the AC coefficients.

Quantization

- The human eye is good at seeing small differences in brightness over a relatively large area
- This is done by simply dividing each component in the frequency domain by a constant for that component, and then rounding to the nearest integer.
- This is the main lossy operation in the whole process.
- As a result of this, it is typically the case that many of the higher frequency components are rounded to zero, and many of the rest become small positive or negative numbers, which take many fewer bits to store.

A typical quantization matrix, as specified in the original JPEG Standard, is as follows:

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

- The quantized DCT coefficients are computed with

QuantizedCoefficient (i,j) = DCT(i,j)/Quantum(i,j)

- where *G* is the unquantized DCT coefficients;
- *Q* is the quantization matrix above;
- and *QC* is the quantized DCT coefficients.

Using this quantization matrix with the DCT coefficient matrix from above results in:

-5	0	0	0	0	0	0	0
2	-1	1	1	0	0	0	0
1	0	0	-1	0	0	0	0
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

- For example, using −80 (the DC coefficient) and rounding to the nearest integer

$$QC = -80/16 = -5$$

Zigzag

- It involves arranging the image components in a "zigzag" order employing run-length encoding (RLE) algorithm
- that groups similar frequencies together, inserting length coding zeros, and then using Huffman coding on what is left.

Zigzag ordering of JPEG image components

-5	0	0	0	0	0	0	0
2	-1	1	1	0	0	0	0
1	0	0	-1	0	0	0	0
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Zigzag

-5	0	0	0	0	0	0	0
2	-1	1	1	0	0	0	0
1	0	0	-1	0	0	0	0
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

The zigzag sequence for the above quantized coefficients

−5							
0	2						
1	−1	0					
1	0	1	0				
1	0	0	1	0			
0	0	−1	0	0	0		
0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	
0	0	0	0	0	0		
0	0	0	0	0			
0	0	0	0				
0	0	0					
0	0						
0							

Entropy coding

- Entropy coding is a special form of lossless data compression.

- The resulting data for all 8×8 blocks is further compressed with a loss-less algorithm, a variant of Huffman encoding.

JPEG encoding summary

128	128	128	128	128	128	128	128
118	111	112	117	120	123	123	122
125	121	115	111	119	119	118	117
120	121	113	113	125	124	115	108
120	120	116	119	124	120	115	110
117	113	111	122	120	110	116	119
109	113	111	122	120	110	116	119
111	121	124	118	115	121	117	113

Color space values

-80	4	-6	6	2	-2	-2	0
24	-8	8	12	0	0	0	2
10	-4	0	-12	-4	4	4	-2
8	0	-2	-6	10	4	-2	0
18	4	-4	6	-8	-4	0	0
-2	8	6	-4	0	-2	0	0
12	0	6	0	0	0	-2	-2
0	8	0	-4	-2	0	0	0

DCT matrix)

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Quantization Matrix

-5	0	0	0	[illegible]	0	0	0
2	-1	1	1	0	0	0	0
1	0	0	-1	0	0	0	0
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Final matrix

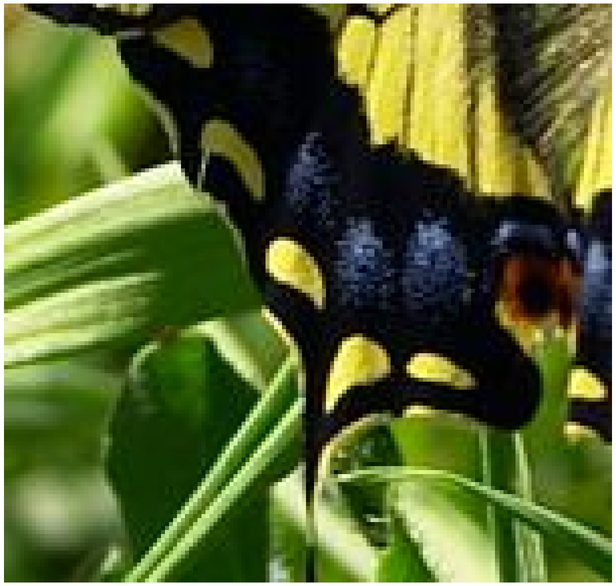
Examples of varying JPEG compression ratios

500KB image, minimum compression

40KB image, half compression

11KB image, max compres

Close-up details of different JPEG compression ratios



Uncompressed image (roughness between pixels still visible)

Half compression, blurring & halos around sharp edges

Max compression, 8-pixel blocks apparent, large distortion in high-frequency areas

JPEG Encoding modes

- Sequential mode
  - Image scanned in a raster scan with single pass, 8-bit resolution
- Sequential mode
  - Step-by-step buildup of image from low to high frequency, useful for applications with long loading times (internet, portable devices, etc)
- Hierarchical mode
  - Encoded using low spatial resolution image and encoding higher resolution images based on interpolated difference, for display on varying equipment

JPEG 2000

- **JPEG 2000** is a wavelet-based image compression standard.
- It was created by the Joint Photographic Experts Group committee in the year 2000 with the intention of superseding their original discrete cosine transform based JPEG standard.
- The standardized filename extension is **.jp2** .

Advantages of JPEG2000

1. Better image quality
2. 25% to 35% smaller file size at comprable image quality
3. Good image quality evan at very high compression ratio,over 80:1
4. Lossless compression mode(identical to original image.
5. Easy way to get ¼ ,1/8,1/16 of original image.

- JPEG 2000 requires far greater decompression time than JPEG
- allows more sophisticated progressive downloads, yet averages similar compression rates.
- JPEG 2000 is not widely supported in web browsers, and hence is not generally used on the World Wide Web.

FEATURES OF JPEG 2000

- *Superior compression performance*: At high bit rates, where artifacts become nearly imperceptible
- JPEG 2000 has a much more significant advantage over certain modes of JPEG: artifacts are less visible and there is almost no blocking.

- *Multiple resolution representation*: JPEG2000 decomposes the image into a multiple resolution representation in the course of its compression process. This representation can be put to use for other image presentation purposes beyond compression as such.
- *Progressive transmission* by pixel and resolution accuracy, commonly referred to as progressive decoding and signal-to-noise ratio (SNR)

- *Lossless and lossy compression*: the JPEG2000 standard provides both lossless and lossy compression in a single compression architecture.

- *Random codestream access and processing*, also referred as Region Of Interest (ROI): JPEG2000 codestreams offer several mechanisms to support spatial random access or region of interest access at varying degrees of granularity. This way it is possible to store different parts of the same picture using different quality.
- *Error resilience*: Like JPEG 1991, JPEG2000 is robust to bit errors introduced by noisy communication channels, due to the coding of data in relatively small independent blocks.

- *Flexible file format*: The JP2 and JPX file formats allow for handling of color-space information, metadata, and for interactivity in networked applications as developed in the JPEG Part 9 JPIP protocol.
- *Side channel spatial information*: it fully supports transparency and alpha planes.

Applications of JPEG 2000

- Consumer applications such as multimedia devices (e.g., digital cameras, personal digital assistants, 3G mobile phones, color facsimile, printers, scanners, etc.)
- Client/server communication (e.g., the Internet, Image database, Video streaming, video server, etc.)
- Military/surveillance (e.g., HD satellite images, Motion detection, network distribution and storage, etc.)

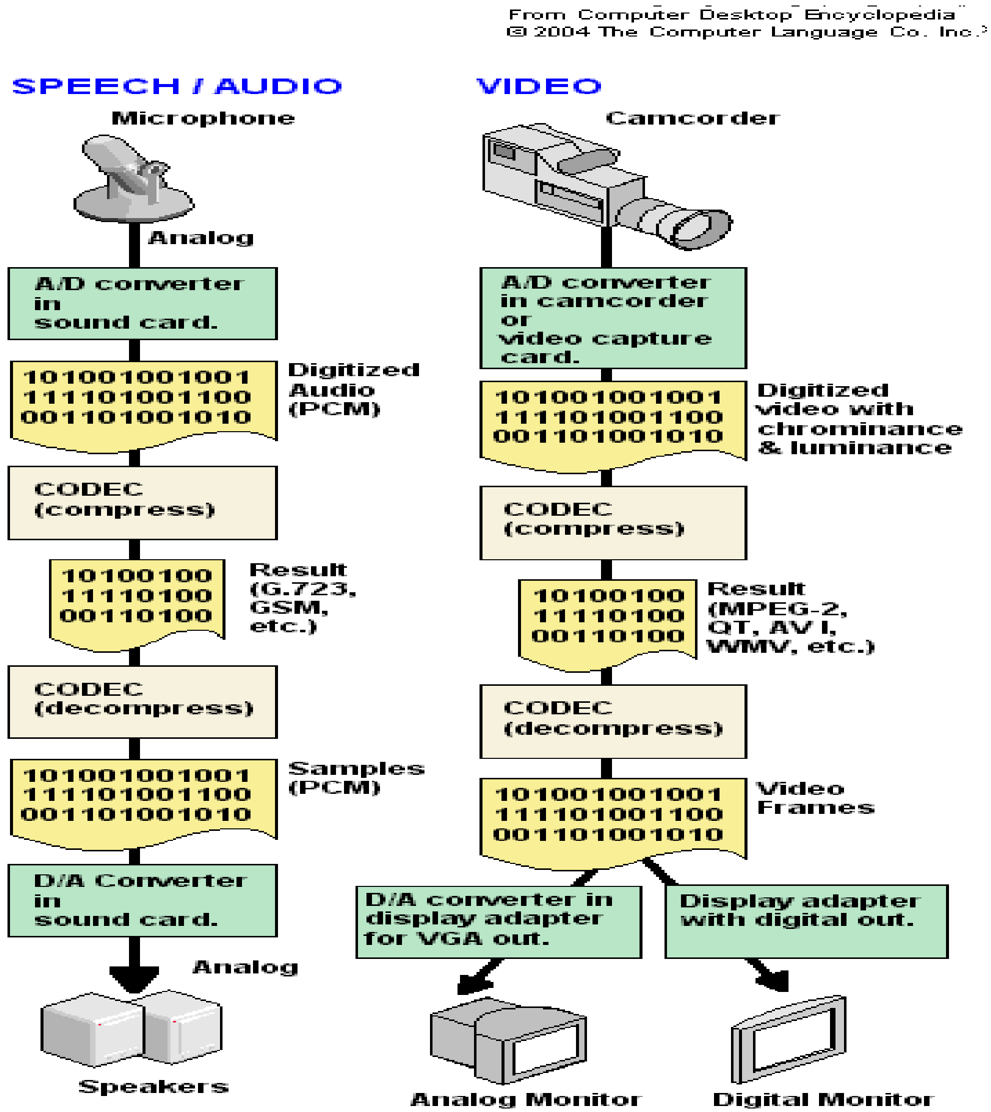
- Medical imagery, esp. the DICOM specifications for medical data interchange.
- Remote sensing
- High-quality frame-based video recording, editing and storage.
- JPEG 2000 has many design commonalities with the ICER image compression format that is used to send images back from the Mars rovers.

Video Compression

- **Video compression** refers to reducing the quantity of data used to represent digital video images.
- It is a combination of image compression and motion compensation.
- reduce the bandwidth required to transmit video via broadcast, via cable TV, or via satellite TV services.

- **Video** compression and decompression program, known as **Codecs**
- Algorithms that handle the **compression and decompression** of the digital video…
- The higher the compression ratio the worse the resulting image

Compression techniques



MPEG - (Motion Picture Experts Group)

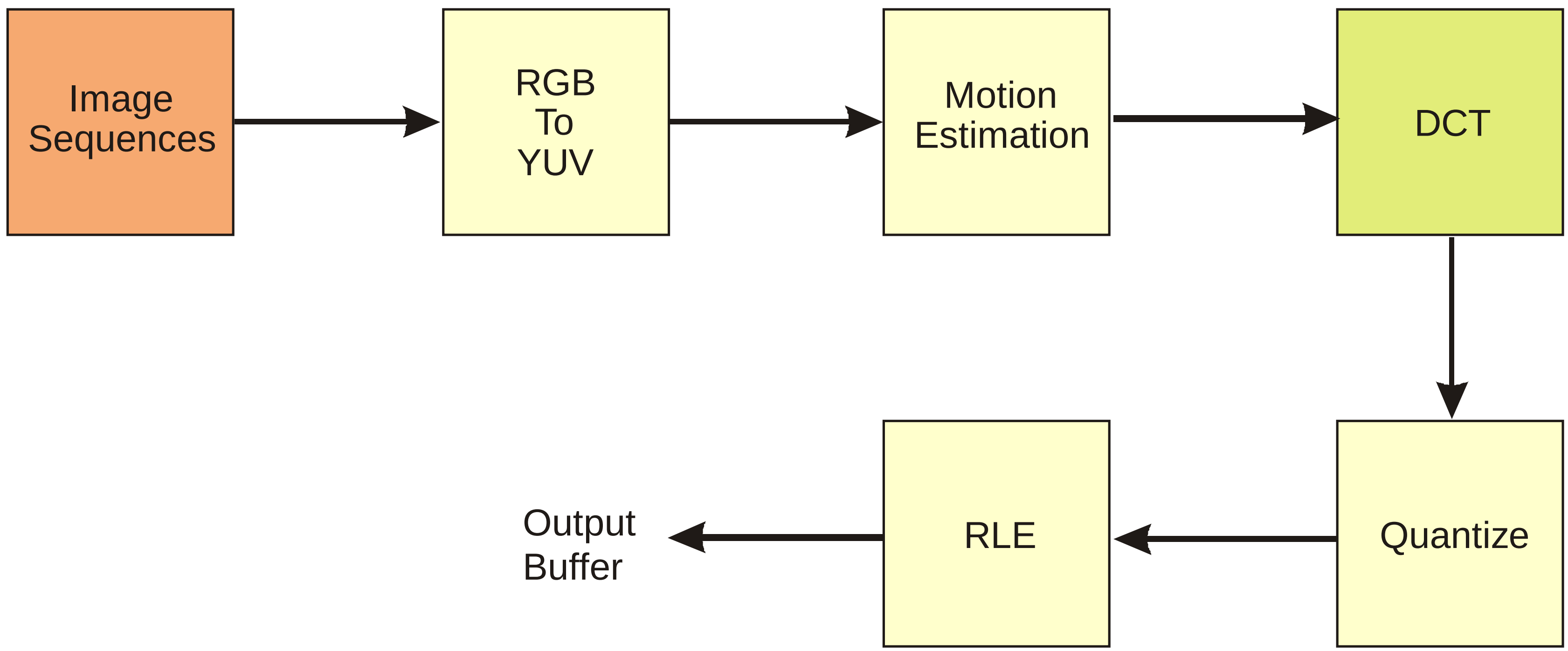
- Digital video and audio compression standards and file formats formulated by the MPEG.
- An international group of industry experts who's aim was to standardize compressed moving pictures and audio.
- They first met as a group in 1988.

- This methodology is considered *asymmetric* in that the encoder is more complex than the decoder.
- The encoder needs to be algorithmic or adaptive where as the decoder is 'dumb' and carries out fixed actions.

- **Moving Pictures Experts Group (MPEG)** standard for compressing video signals using DCT (Discrete Cosine Transform)…

- MPEG uses a type of **lossy compression**, since some data is removed.

MPEG Process



MPEG1

- Designed for up to 1.5 Mbit/sec
- based on CD-ROM video applications
- video on the Internet, transmitted as .mpg files.
- In addition, level 3 of MPEG-1 is the most popular standard for digital compression of audio--known as MP3.
- The quality however is not sufficient for TV broadcast.
- It may be used by TV content developers for TV stills (shots with no motion.)

MPEG-2

- Designed for between 1.5 and 15 Mbit/sec
- It is based on MPEG-1, but designed for the compression and transmission of digital broadcast television.
- The most significant enhancement from MPEG-1 is its ability to efficiently compress interlaced video.
- Used in DVD technology

- Very different from previous generations
- Aimed at low-bandwidth applications – at upper end, good enough for digital TV

MPEG-4

- MPEG-4 builds on MPEG-2 by allowing for greater use of multimedia/graphics within the video stream and for better compression.
- Its standards are for use in digital television, interactive graphics and interactive multimedia .
- MPEG-4 delivers video quality as good as MPEG-2 at a significantly lower bit rate.

- MPEG-4 is based on object-based compression, similar in nature to the Virtual Reality Modeling Language.
- Individual objects within a scene are tracked separately and compressed together to create an MPEG4 file.
- It also allows developers to control objects independently in a scene, and therefore introduce interactivity.

- Designed to transmit video and images over a narrow bandwidth
- **can mix video with text, graphics and 2-D and 3-D animation layers.**

- A major feature of MPEG-4 is its ability to **identify and deal with separate audio and video objects in the frame**, which **allows separate elements to be compressed more efficiently** and dealt with independently.

MPEG-7

- currently under development, is also called the Multimedia Content Description Interface.
- It's a way of attaching meta-data to multimedia.
- Contrary to the previous MPEG standards, which described actual content, MPEG-7 will represent information about the content.

MPEG-21

- work on this standard, also called the Multimedia Framework, has just begun.
- MPEG-21 will attempt to describe the elements needed to build an infrastructure for the delivery and consumption of multimedia content, and how they will relate to each other.
- More oriented to MPEG content delivery from the viewpoint of the consumer.
- It's envisioned to be an open framework for multimedia delivery and consumption.

- **MPEG-4 AVC -** *AVC* stands for Advanced Video Coding.
-

MPEG.org – A MPEG supersite :
- **MPEG Splicing -** The ability to cut into an MPEG bit stream for switching and editing.

MPEG-TS - MPEG Transport Stream.

Audio Compression Techniques

Audio Compression

- **Audio compression** is a form of data compression designed to reduce the transmission bandwidth requirement of digital audio streams and the storage size of audio files.
- Audio compression algorithms are implemented in computer software as audio codecs.
- specifically optimized audio lossless and lossy algorithms have been created.
- Lossy algorithms provide greater compression rates and are used in mainstream consumer audio devices.

Introduction

- Digital Audio Compression
  - Removal of redundant or otherwise irrelevant information from audio signal
  - Audio compression algorithms are often referred to as “audio encoders”
- Applications
  - Reduces required storage space
  - Reduces required transmission bandwidth

Audio Compression

- Audio signal – overview
  - Sampling rate (# of samples per second)
  - Bit rate (# of bits per second). Typically, uncompressed stereo 16-bit 44.1KHz signal has a 1.4MBps bit rate
  - Number of channels (mono / stereo / multichannel)
- Reduction by lowering those values or by data compression / encoding

Audio Data Compression

- Redundant information
  - Implicit in the remaining information
  - Ex. oversampled audio signal
- Irrelevant information
  - Perceptually insignificant
  - Cannot be recovered from remaining information

Audio Data Compression

- Lossless Audio **Compression**
  - Removes redundant data
  - Resulting signal is **same** as original – perfect reconstruction
- Lossy Audio **Encoding**
  - Removes irrelevant data
  - Resulting signal is **similar** to original

Audio Data Compression

- Audio vs. Speech Compression Techniques
  - Speech Compression uses a human vocal tract model to compress signals
  - Audio Compression does not use this technique due to larger variety of possible signal variations

Generic Audio Encoder

- Psychoacoustic Model
  - Psychoacoustics – study of how sounds are perceived by humans
  - Uses **perceptual coding**
    - eliminate information from audio signal that is inaudible to the ear
  - Detects conditions under which different audio signal components **mask** each other

Audio Compression

- Audio can also be compressed in a similar way to image data.
- For lossless audio compression, prediction is usually applied first.
  - Simple prediction.
  - Adaptive prediction.
- Stereo decorrelation.

Examples

- FLAC (free lossless audio codec)
  – Polynomial fitting prediction or linear prediction.
  – Rice coding.

- MPEG4-ALS
  – Lossless audio coding standard in MPEG4.
  – Adaptive Linear Prediction.
  – Supports up to 65535 channels.
  – Fast random access.

MPEG-4 ALS System Diagram

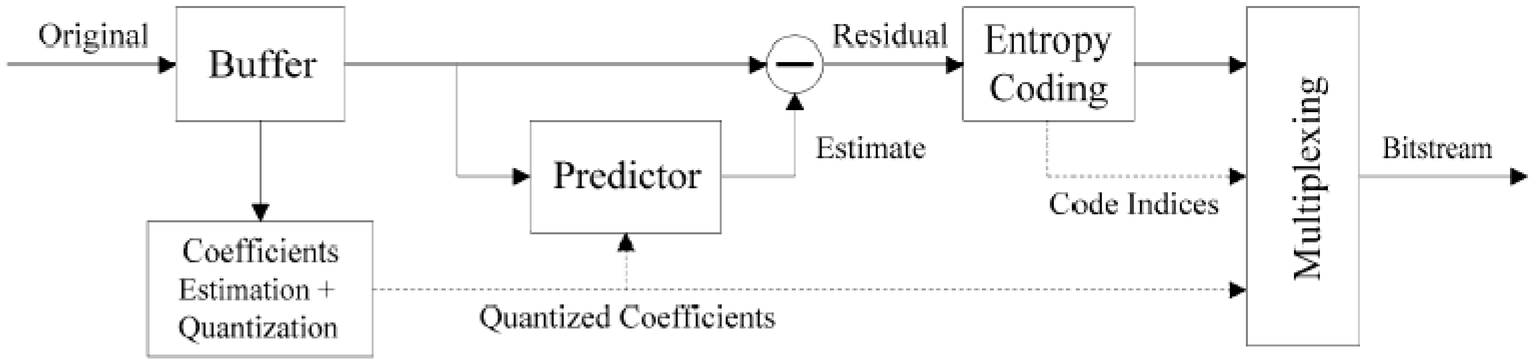


Fig. 1. MPEG-4 ALS encoder

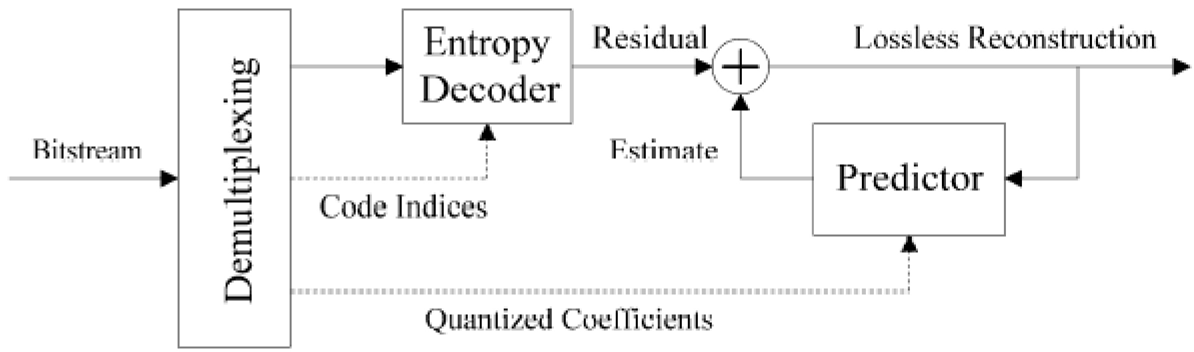
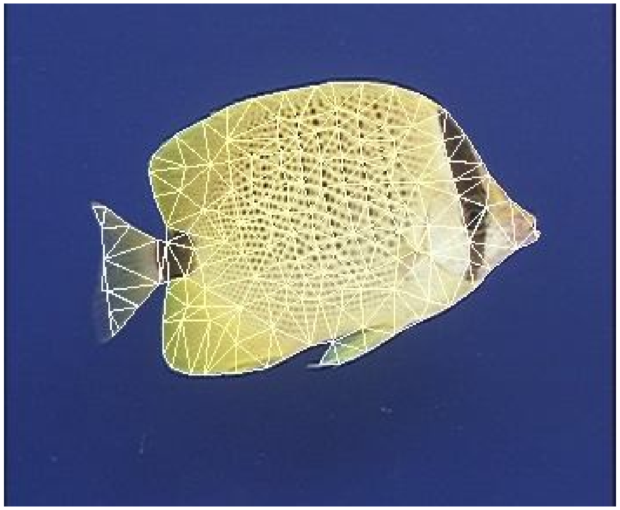


Fig. 2. MPEG-4 ALS decoder

- **MPEG-1 Audio Layer 3**, more commonly referred to as **MP3**, is a digital audio encoding format using a form of lossy data compression.
- It is a common audio format for consumer audio storage, as well as a de facto standard encoding for the transfer and playback of music on digital audio players.
- MP3's use of a lossy compression algorithm is designed to greatly reduce the amount of data required to represent the audio recording and still sound like a faithful reproduction of the original uncompressed audio for most listeners, but is not considered high fidelity audio by most audiophiles.
- An MP3 file that is created using the mid-range bitrate setting of 128 kbit/s will result in a file that is typically about 1/10th the size of the CD file created from the original audio source. An MP3 file can also be constructed at higher or lower bitrates, with higher or lower resulting quality.

MPEG-4 II - Meshes

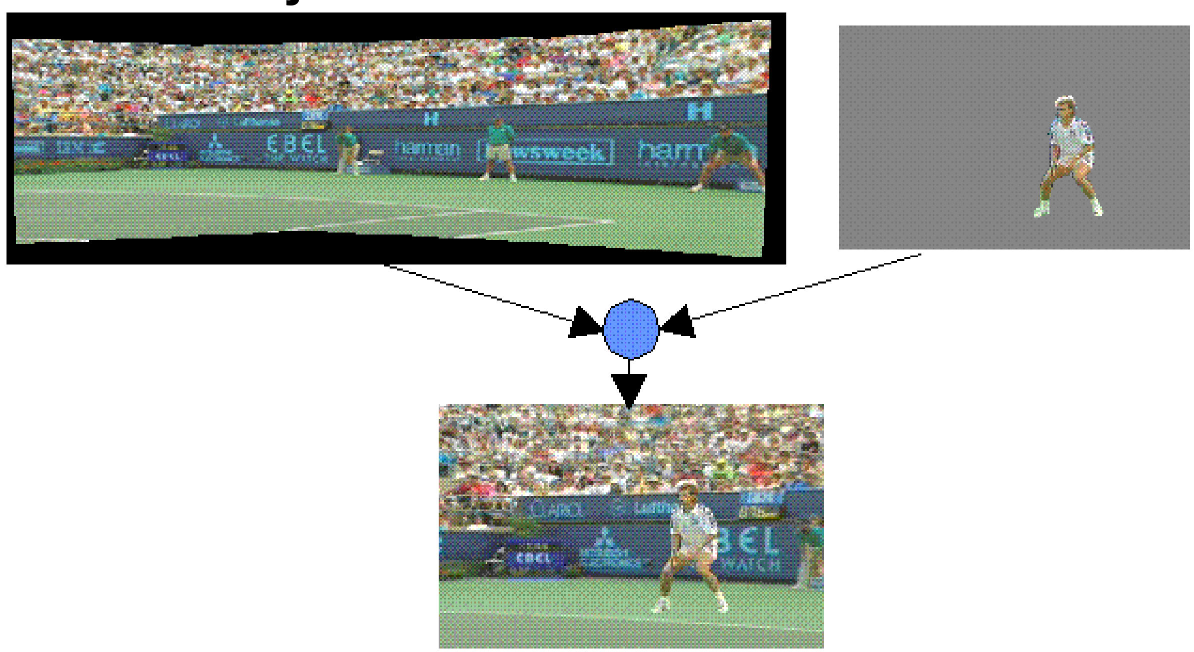
- 2-D animated meshes



- Textures mapped onto meshes
- Store vertices of mesh and movement

MPEG-4 III - Sprites

- MPEG-4 is object based – state of the art



- Panoramic images – massive compression

4. Audio Compression

- Techniques from image compression can be used
  - Huffman encodes output
  - DCT

- MP3 – huge!!!!
- How does it work…

MP3 – MPEG-1 Layer 3

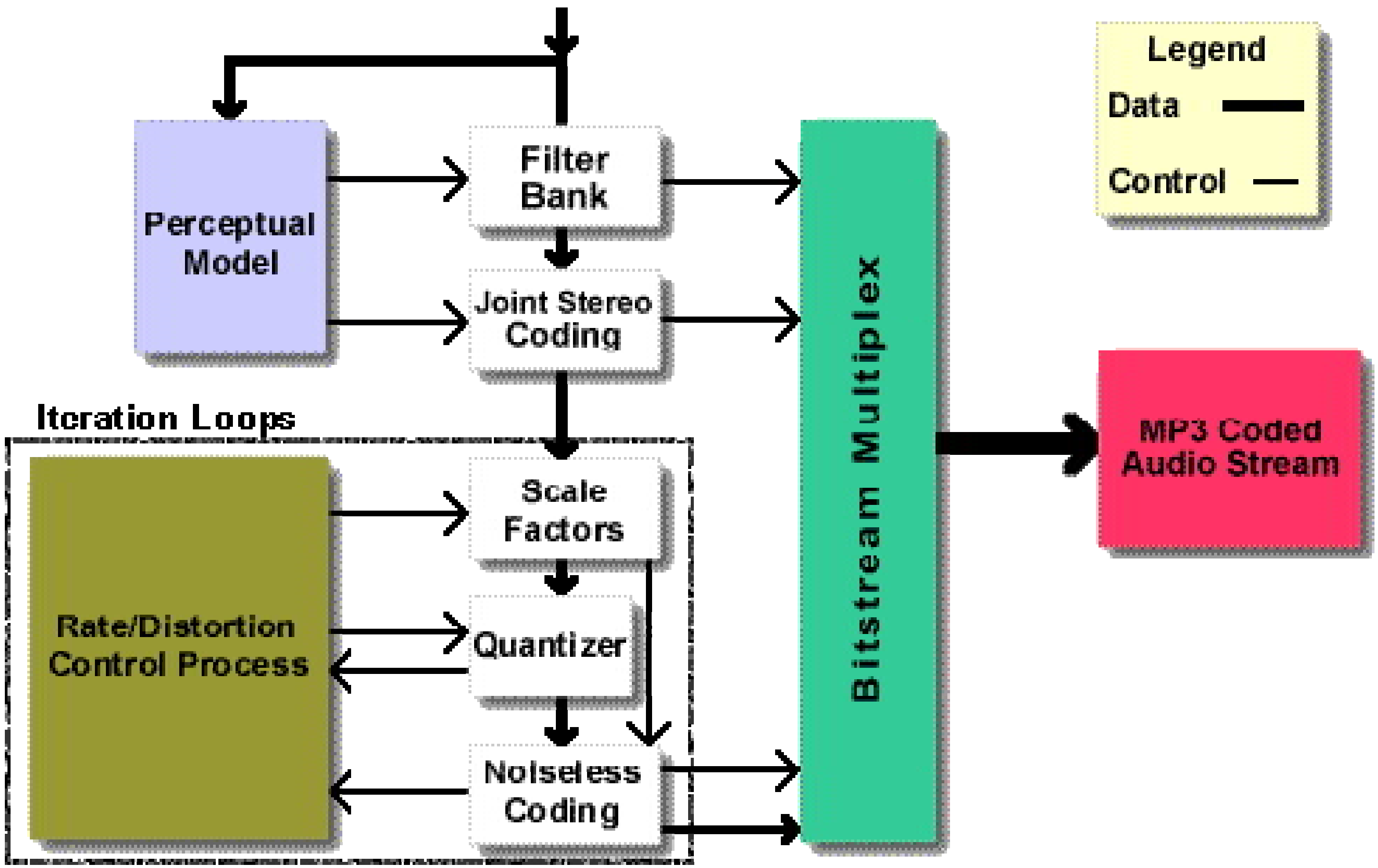
- 1. Minimal Audition Threshold
  - Don’t store anything under 5Khz
- 2. Masking Effect
  - Uses psychoacoustic model of the ear
  - Don’t store quiet and loud noises simultaneously

MP3 II

- 3. Joint Stereo (JS) coding
  - 1. Intensity Stereo (IS)
    - Ear unable to locate some frequencies – bass
    - Store signal in mono + minimum for spatialization
  - 2. Mid/Side (MS) Stereo
    - Used if left and right speakers are similar
    - Store middle (L+R) plus a side speaker (L or R) e.g.

	L		R	
Raw:	10		5	
Store:		7	5	Fewer bits
Decompress:	10		5	

MP3 III - schematic



Summary

Technique	Compression Ratio	When?
Huffman	1.5-2:1	1952
RLE	4-10:1	1966
LZW	2-10:1	1977&84
Quadtree	2:1	1980
VQ	10:1	1984
Directional Filtering	10-40:1	1985
Fractals	10-1000:1	1988
MPEG-1	10-100:1	1993
Surface Methods	10-50:1	1995
MPEG-2	10-200:1	1995
MPEG-4	10-500:1	1999

time