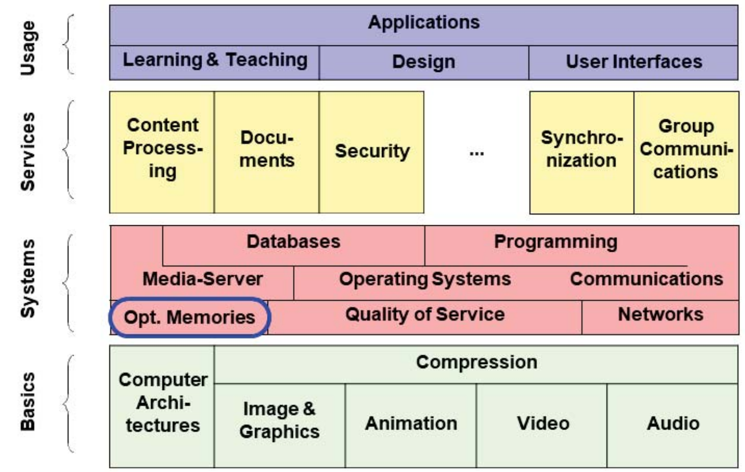
OPTICAL STORAGE MEDIA

UNIT 6

6.1. Basic Technology

Scope



6.1. Basic Technology (cont..)

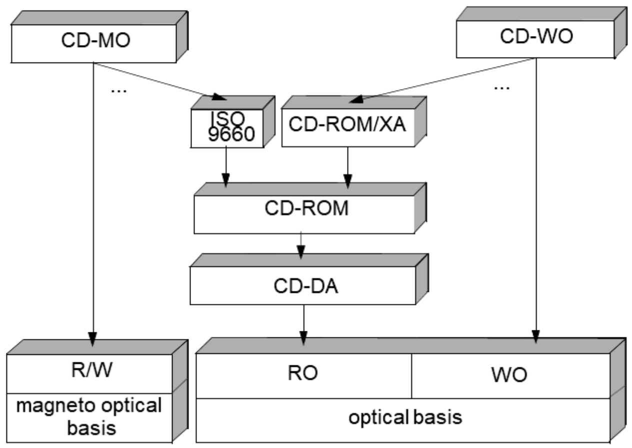
- **Optical storage** is the **storage** of data on an optically readable **medium**.
- Data is recorded by making marks in a pattern that can be read back with the aid of light, usually a beam of laser light precisely focused on a spinning **optical disc**.
- **Optical storage**, electronic storage medium that uses low-power laser beams to record and retrieve digital (binary) data.
- In optical-storage technology, a laser beam encodes digital data onto an optical, or laser, disk in the form of tiny pits arranged in concentric tracks on the disk‟s surface

6.1. Basic Technology (cont..)

- Optical storage media offer a higher storage density at a lower cost.
- Current magnetic data storage carriers take the form of floppy disks or hard disks and are used as secondary storage media.
- An optical disc drive is a device in a computer that can read CD-ROMs or other optical discs, such as DVDs and Blu-ray discs.
- Optical storage differs from other data storage techniques that make use of other technologies such as magnetism, such as floppy disks and hard disks, or semiconductors, such as flash memory and RAM.

6.1. Basic Technology (cont..)

Compact Disc and in addition Digital Versatile/Video Disk



6.1. Basic Technology (cont..)

History

1973	Video Long Play (VLP) - published
1983	Compact Disc Digital Audio (CD-DA) - available: • Red Book Standard
1985	Compact Disc Read Only Memory (CD-ROM): • Yellow Book Standard for physical format • High Sierra Proposal • ISO 9660 Standard for logical file format
1986	Compact Disc Interactive (CD-I) - announcement: • Green Book
1987	Digital Video Interactive (DVI) - first presentation
1988	CD-ROM Extended Architecture (CD-ROM-XA) announcement
1990	CD Write Once (CD-WO), CD Magneto Optical (CD-MO): • Orange Book
1996	Digital Video Disk DVD

6.2. Fundamentals

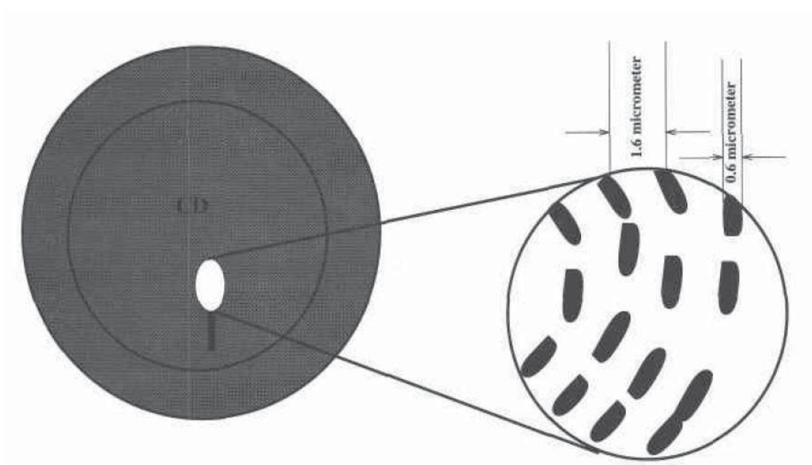


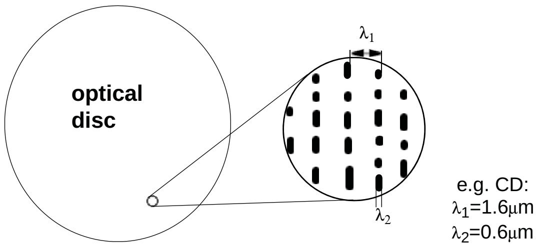
Fig. :Data on a CD as an example of an optical disk (track with "lands' and "pits').

6.2. Fundamentals (cont..)

- Magnetic disks is that on the former 1.66 data bits per µm can be stored. This results in a data density of 1,000,000 bits per mm^2, which implies 16,000 tracks per inch. In comparison, a floppy disk has 96 tracks per inch.

6.2. Fundamentals (cont..)

Pits and Lands



Information is stored in a spiral-shaped track:

- Series of pits and lands in substrate layer
- Transition from pit to land and from land to pit: „1"
- Between transitions: sequence of „0" s
- 16000 turns/inch (tpi)

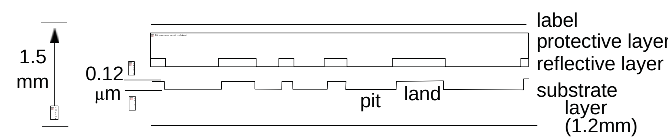
Laser focused onto reflective layer

- Lands - almost totally reflecting
- Pits - scattering

6.2. Fundamentals (cont..)

Fundamentals: Physical Structure

Cross-section through disc in direction of spiral track:

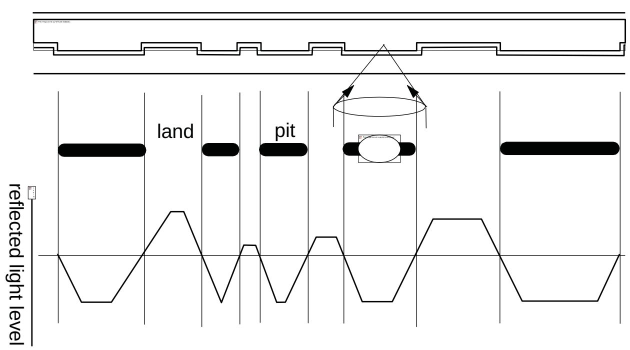


3 layers:

- Transparent substrate layer
- Reflective layer
- Protective layer

6.2. Fundamentals (cont..)

Fundamentals: Read Data



Laser focused onto reflective layer:

- Lands - almost totally reflecting
- Pits - scattering

Advantages of Optical Storage Media

High data density:

- 1.66 data bits / µm of track
- 16000 tpi ⇔ floppy disk: 96 tpi

Long term storage:

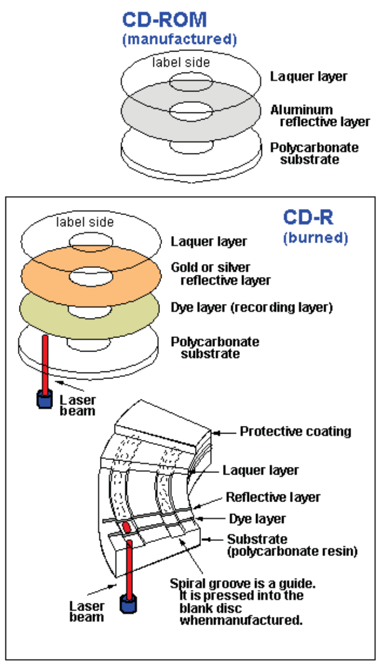
- Protection of data
- Surface out of focus ➔ insensitivity to dust, scratches

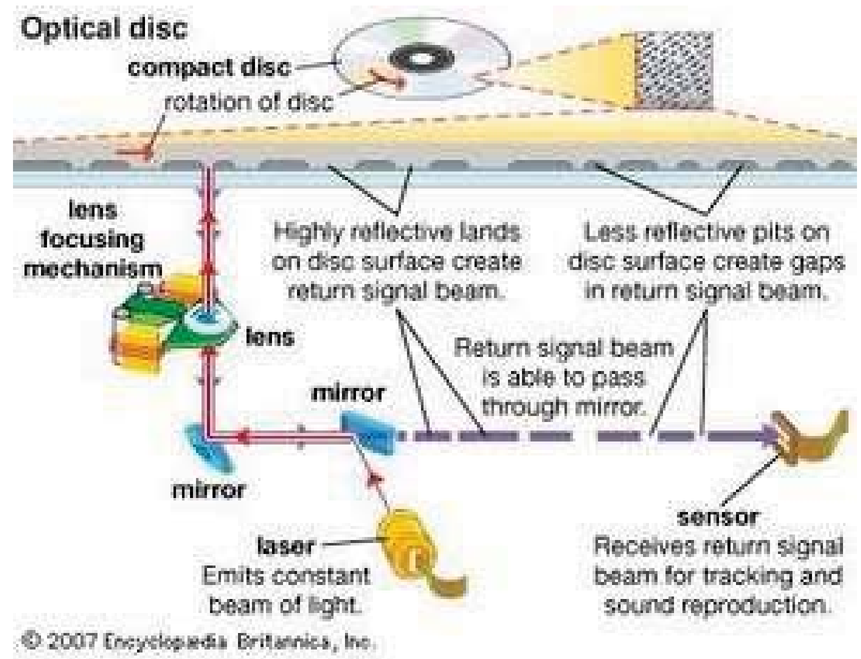
Low probability of head crashes:

- Distance between head and substrate surface > 1mm

Adequate error correction

Each digital disc is equivalent to the master





Laser Vision

Characteristics:

- Storage of video and audio
- Analogue encoding
- High quality of reproduced data
- Diameter: ~ 30 cm
- Storage capacity: ~ 2.6 GByte

History:

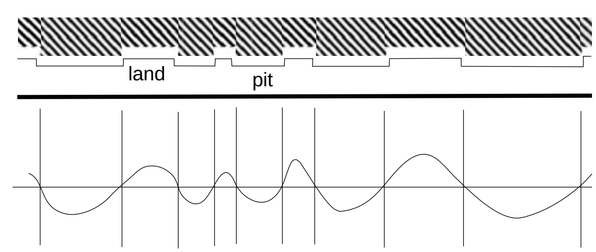
- Originally Video Long Play (VLP)
- 1973 first description in Philips Technical Review

Laser Vision: Fundamentals

Principles:

- Mix of audio and video
- Frequency modulation
- No quantization of pit length

Cross section through a Laser Vision disc:



6.3. Compact Disc Digital Audio (CD-DA)

- An Audio CD is a music CD like that you buy in a music store.
- It can be played on any standard CD player (such as a CD deck, or your car CD player, or a portable CD player).
- Music is stored on Audio CDs as uncompressed digital data, no data is lost and quality is very high, exactly as in WAV(Waveform Audio File Format, WAVE) digitally encoded files.
- When you put an Audio CD into your personal computer CD player and play it, audio is extracted on the fly and played by your PC sound hardware.
- **Compact Disc Digital Audio**(**CDDA** or **CD-DA**) is the standard format for audio compact discs.

6.3. Compact Disc Digital Audio (CD-DA)

Storage of audio data

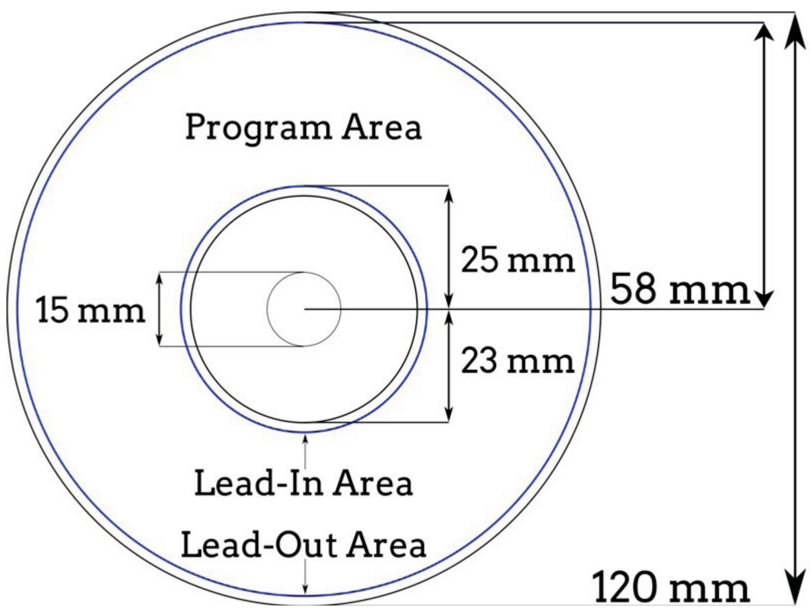
History:

- Development of basic technology by N. V. Philips
- Cooperation of N. V. Philips and Sony Corporation
- 1983 CD-DA available

Physical characteristics:

- Diameter: 120 mm
- Constant linear velocity (CLV), i.e. number of rotations/s depends on position of head relative to disc center
- Track shape: One spiral with appr. 20000 turns (LP: 850 turns)

Physical characteristics:



Data structure

- The audio data stream in an audio CD is continuous, but has three parts.
- The main portion, which is further divided into playable audio tracks, is the *program area*. This section is preceded by a *lead-in*track and followed by a *lead-out*track.
- The lead-in and lead-out tracks encode only silent audio, but all three sections contain subcodedata streams.
- The lead-in's subcode contains repeated copies of the disc's Table Of Contents (TOC), which provides an index of the start positions of the tracks in the program area and lead-out.

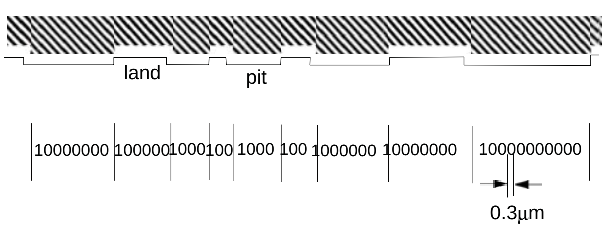
Data structure(cont..)

- The track positions are referenced by absolute time code, relative to the start of the program area, in MSF format: minutes, seconds, and fractional seconds called *frames*. Each time code frame is one seventy-fifth of a second, and corresponds to a block of 98 channel-data frames—ultimately, a block of 588 pairs of left and right audio samples.
- Time code contained in the sub channel data allows the reading device to locate the region of the disc that corresponds to the time code in the TOC. The TOC on discs is analogous to the partition table on hard drives. Nonstandard or corrupted TOC records are abused as a form of CD/DVD copy protection, in e.g. the key2Audio scheme.

CD-DA: Characteristics

- Audio data rate:
  - Sampling frequency: 44100Hz
  - 16 bit quantization (Each audio sample is a signed 16-bit two's complement integer, with sample values ranging from −32768 to +32767.)
  - Pulse code modulation (PCM)
  - Audio data rate = 1411200 bit/s = 176,4 Kbyte/s
- Quality:
  - Signal to noise ratio (S/N):
  - ~ 6dB/bit, 16 bit quantization ➔ S/N exactly 98 dB
  - LP, tape: S/N 50-60 dB
- Capacity: (without error correction data)
  - Playback time: maximal 74 min
  - Capacity =74 min ∗ 1411200 bit/s = 6265728000 bit ~ 747 MByte

CD-DA: Pits and Lands



Length of pits: multiples of 0.3µm

Coding:

- Transition from pit to land / from land to pit: „1‟
- Between transitions: sequence of „0‟s

CD-DA: Eight-to-Fourteen Modulation

- Eight-to-Fourteen Modulation, or EFM as it is abbreviated, is an encoding technique used by CDs and provides a way of countering errors by encoding a byte into 2 bytes.
- Using EFM the data is broken into 8-bit blocks (bytes). Each 8-bit block is translated into a corresponding 14-bit codeword using a predefined lookup **table**.
- The 14-bit codeword are chosen so that binary ones are always separated by a minimum of two and a maximum of ten binary zeroes.
- EFM maximizes the number of transitions possible with an arbitrary pit and land length which is determined by the wavelength of the laser light used to read the data.

CD-DA: Eight-to-Fourteen Modulation

Restricted laser resolution:

- Minimal distance between transitions (pit to land, land to pit)
- At least two “0” between two “1”

Generation of clock signal:

- Maximal distance between transitions (pit to land, land to pit)
- Not more than 10 consecutive “0”

Eight-to-Fourteen Modulation:

- 8 bit value is encoded using 14 bits
- 267 combinations possible
- 256 are used
  (criterion: efficient implementation with small number of gates)

CD-DA: Eight-to-Fourteen Modulation

Example for a code conversion table:

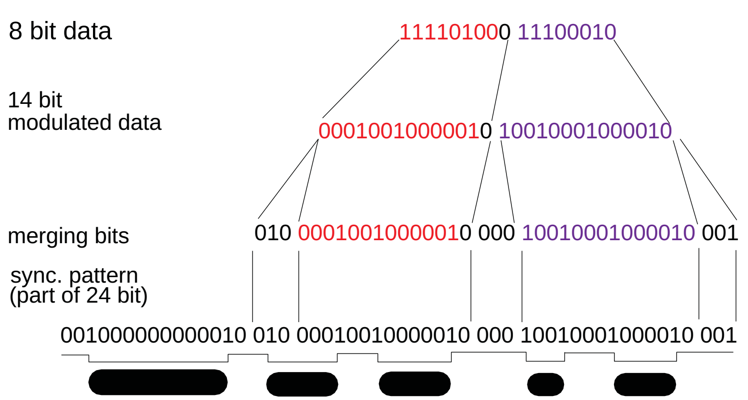
data bits	channel bits
00000000	01001000100000
00000001	10000100000000
...	...

Concatenation of independent 14 bit values
Í potential violation of:

- Min. distance of 2 bits
- Max. distance of 10 bits

Three additional *merging bits*

CD-DA: Eight-to-Fourteen Modulation Example



CD-DA: Error Handling

Typical Errors:

- Scratches, dust, fingerprints
- "Burst errors"
- To be detected and corrected

Two-level Reed-Solomon Code with frame interleaving:

- First level: byte level, EDC and ECC
  - two groups: each with 4 correction bytes for 24 data bytes
  - first group: correction of single byte errors
  - second group: correction of double byte errors, detection of further errors
- Second level: Frame interleaving
  - frame: 588 channel bits = 24 audio data bytes
  - distribution of consecutive data bytes and corresponding ECC bytes over adjacent frames

CD-DA: CIRC

Cross Interleaved Reed Solomon Code

Error rate: 10^{-8} (~ 1bit/100 millions of bits) Exact correction of 4000 data bits possible:

- 4000 data bits * 0.3 µm/channel bit
- ~ 2.5 mm
- Hence: burst errors within 2.5 mm can be corrected

Interpolation:

- Up to 12,300 data bits (~ 7 mm)
- Hence: error within 7mm can be repaired

CD-DA: Frames

Frame **consists of:**

- Data:
  - 2 groups of 12 audio data bytes each (actual data)
- Error detection and correction code:
  - 2 groups of 4 parity bytes
  - According to Reed-Solomon
- Control & display byte:
  - Together with c&d bytes of other frames it forms subchannel stream
  - E.g., subchannel byte for track start identification
- Synchronization pattern:
  - Start of a frame
  - 12 x “1” + 12 x “0” + 3 merging bits = 27 bits

CD-DA: Data Streams

Audio bit stream ~ 1.41 x 10^6 bit/s:

- 44,1 kHz sampling frequency~ 1411200 bit/s
- 16-bit stereo PCM
- Uniform quantization

Data bit stream ~ 1.94 x 10^6 bit/s:

- Audio bit stream
  + parity bytes
  + control & display byte

Channel bit stream ~ 4.32 x 10^6 bit/s:

- Data bit stream
  + EFM
  + merging bits
  + synchronization pattern

CD-DA: Areas

Areas:

- Lead-in area:
  - List of contents
  - Indication to start of each track
- Program area:
  - Up to 99 tracks of different lengths
  - Typically one track relates to one song
- Lead-out area

Random Access:

- Tracks
- Index points:
  - IP_0: start of track
  - IP_1: start of audio data
  - Track pregap: part between IP_0 and IP_1

CD-DA: Summary

Provides:

- Suitable means for typical errors caused by damage, dust in audio data
- CD-DA specification is base for family of optical storage media

But:

- Not conceived for:
  - video (different ECC, EDC scheme required)
  - discrete data (error rate too high)
- Desired: simultaneous play back of various media

6.4. Compact Disc Read Only Memory (CD-ROM) and Extended Architecture

- A **CD-ROM** is a pre-pressed optical compact disc which contains data.
- Computers can read CD-ROMs, but cannot write to CD-ROMs, which are not writable or erasable.

6.4. Compact Disc Read Only Memory (CD-ROM) and Extended Architecture

Storage of:

- Data, audio, compressed audio and video

Yellow Book CD-ROM Standard:

- CD-ROM mode 1: for any data
- CD-ROM mode 2: for compressed audio and video data
- But can not be combined on single track

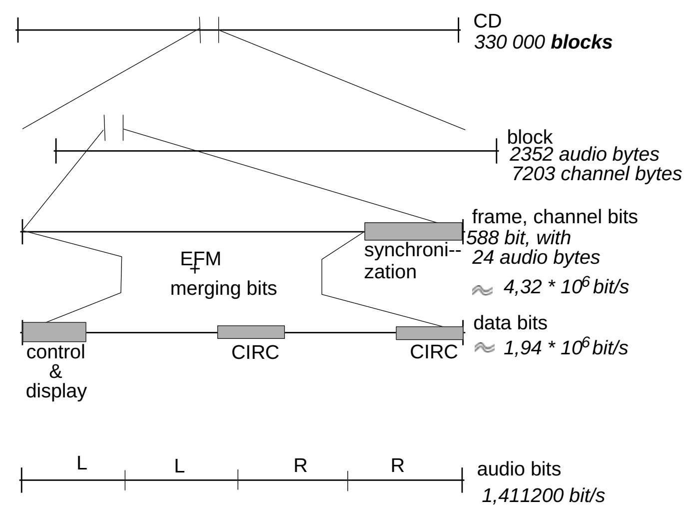
Within single track:

- Only CD-DA audio or only CD-ROM specific data

Mixed Mode Disc:

- Data tracks at the beginning
- Subsequent tracks for audio data

CD-ROM: Structure



CD-ROM: Structure

Fine granularity for random access:

- Tracks, IP not sufficient
- Structure with a higher resolution: block
- Blocks with fixed number of frames

Some numbers:

- 1 block = 32 frames
- 75 blocks/s (for single-speed CD-ROM)
- 1411200 bit/s / 75 blocks/s / 8bit/byte = 2352 byte/block

Allows for:

- Random access
- Better EDC, ECC

CD-ROM Mode 1

sync 12	header 4	user data 2048	EDC 4	blanks 8	ECC 276

2352 byte

- 4 bytes for the header, which carries an unambiguous (definite) specification of the block.
  - The first byte stores minutes, the second byte stores seconds and the third byte contains the block number. The fourth byte includes the mode specification.
- 2,048 bytes for the user data.
- 4 bytes for error detection.
- 8 unused bytes.
- 276 bytes for error correction. Hence, an error rate of 10^{-12} can be achieved

CD-ROM Mode 1 (cont..)

- A CD-ROM contains 333,000 blocks to be played in 74 minutes. The capacity of a CD-ROM with all blocks in mode 1 can be computed as follows

$$
\begin{aligned}
Capacity_{CD-ROM_{mode1}} &= \\
&= 333,000 blocks \times 2048 \frac{bytes}{block} = 681,984,000 bytes \\
&= 681,984,000 \times \frac{1}{1024 \frac{bytes}{Kbyte}} \times \frac{1}{1024 \frac{Kbytes}{Mbyte}} \approx 660 Mbytes
\end{aligned}
$$

The *data rate* in mode 1 is:

$$
\begin{aligned}
Rate_{CD-ROM_{mode1}} &= \\
&= 2,048 \frac{bytes}{Block} \times 75 \frac{Blocks}{s} = 153.6 \frac{Kbytes}{s} \equiv 150 \frac{Kbytes}{s}
\end{aligned}
$$

CD-ROM Mode 1

sync 12	header 4	user data 2048	EDC 4	blanks 8	ECC 276

2352 byte

1 block = 2352 byte:

- Header bytes include minutes, seconds, block number, mode
- Error rate = 10^{-12}

Capacity:

- Max. 74 min x 60 s/min x 75 block/s = 333000 blocks
- 333000 blocks/CD ~ 650 MByte (user data)

Data rate:

- 2048 byte/block x 75 block/ s ~150 KByte/s (single-speed)

Used by most CD-ROM applications

CD-ROM Mode 2

sync	header	user data
12	4	2336

2352 byte

Capacity:

- 333000 blocks x 2336 byte/block
  = 777888000 byte ~ 741.85 MByte

Data rate:

- 2336 byte/block x 75 block/s = 171 KByte/s (single-speed)

Problem: concatenation of mode 1 and mode 2 blocks

CD-ROM Mode 2 (cont..)

1. Additional error correction does not exist. The capacity and data rate of a CD-ROM with all blocks in mode 2 can be computed as follows:

$$Capacity_{CD-ROM_{mode2}} = 333,000\ blocks \times 2336 \frac{bytes}{block} \approx 777,888,000\ bytes$$

$$Rate_{CD-ROM_{mode2}} = 2336 \frac{bytes}{block} \times 75\ blocks/s \approx 175.2\ Kbytes/s$$

CD-ROM: Average Access Time

Time to position a block/sector:

- Synchronization time:
  - Adapt internal clock to disc signal
  - Range of milliseconds
- Seek time:
  - Adaptation of laser to radius: max. 1s
- Rotation delay (for constant velocity time):
  - Find sector within 1 rotation
  - Adapt disk speed
  - ~ 300 ms

Maximum access time > 1s

Average access time > 300ms (with data caching)

- Simultaneous reading of audio and other data in CD-ROM mode 1 not possible

CD-ROM: File System

CD-ROM:

- No logical file format
- No directory specification

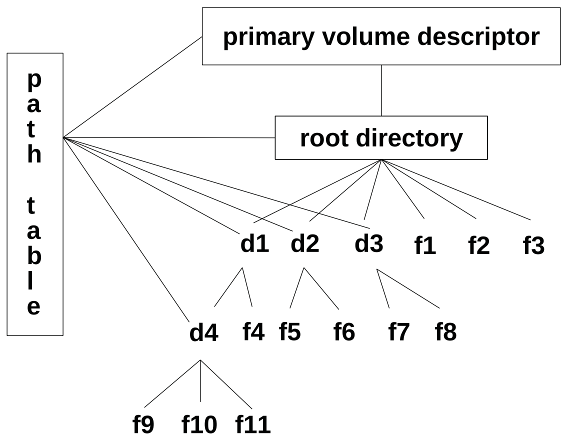
High Sierra Proposal:

- Developed by group of industry representatives
- Initial file system leading to ISO 9660

ISO 9660 file standard:

- Directory tree: information about files
- Path table:
  - List of all directories
  - Direct access to files of any level
- File interleaving

CD-ROM: File System - Path Table



ISO 9660 File System

First track:

- 16 blocks (sectors 0 to 15): system area
- Volume descriptors in subsequent blocks:
  - Primary descriptor:
    - Length of file system
    - Length and address of path table
  - Supplementary descriptors
- Volume descriptor terminator

Logical block size:

- Between 512 byte and 2048 byte (in steps of 2^i)
- Blocks of 512 byte, 1024 byte, and 2048 byte are used
- Files begin at logical block start

CD-ROM Extended Architecture (CD-ROM / XA)

History:

- N.V. Philips, Sony and Microsoft (announcement in 1988)

Goal:

- Simultaneous transfer of various media data

Characteristics:

- Based on CD-ROM mode 2, ISO 9660, CD-I
- Extension of Yellow Book standard
- Interleaving of blocks of different media within the same track
- Definition of a new type of track used for:
  - compressed audio (ADPCM) and video data
  - images
  - text, programs
- Distinction between two block formats: „Form 1“, „Form 2“

CD-ROM / XA (Mode 2) Form 1

sync	header	sub-header	user data	EDC	ECC
12	4	8	2048	4	276

2352 bytes

Subheader:

- Specification of CD-ROM Mode 2 XA-Format type
- 8 bytes

Improved error handling for:

- Text
- Program data

By:

- 4 byte for error detection
- 276 byte for error correction

CD-ROM / XA (Mode 2) Form 2

sync	header	sub-header	user data	EDC
12	4	8	2324	4

2352 byte

- Storage of compressed data (incl. audio, video)
- Only 4 bytes for error detection
- 13% more data bytes

CD-ROM / XA: Audio

	CD audio	level B stereo	level B mono	level C stereo	level C mono
compr. ratio	1	4:1	8:1	8:1	16:1
coding techn.	PCM	ADPCM	ADPCM	ADPCM	ADPCM
sampling freq.	44100 Hz	37800 Hz	37800 Hz	18900 Hz	18900 Hz
capacity	74 min	4 h 48 min	9 h 36 min	9 h 36 min	19 h 12 min
data rate	176 kByte/s	43 Kbyte/s	22 Kbyte/s	22 Kbyte/s	11 Kbyte/s

CD-ROM / XA: Drawbacks

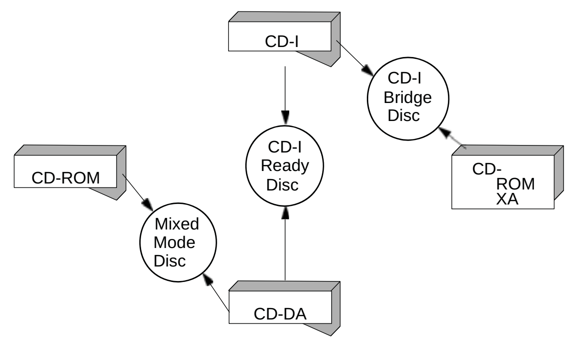
Compatibility to audio and video compression

- For some media only reference to standard
  - coding / decoding not part of CD technology
- MPEG audio not compatible (MPEG does not use ADPCM)

Interleaved storage of data of different types in the same track:

- Requires special disc layout
- Requires effective interleaving with choice of suitable audio level
- Complex application development

Further CD-ROM Based Developments



CD-Interactive (CD-I)

History:

- Developed by Philips and Sony
- 1986 CD-I announcement
- 1988 Green Book: CD-I extension based on Red and Yellow Book
- Originally for consumer market

CD-I system:

- CD-ROM based format with interleaving of different media
- Compression algorithms
- Software (operating system)
- Hardware (decoder)

Different quality levels Low

data rates

CD-I: Software and Hardware

CD-I Software with CD-RTOS operating system:

- CD-RTOS operating system is an extension of OS/9
- Real-time capabilities

CD-I Hardware (decoder):

- System board:
  - 680xx processor
  - Video-, audio -chips
- CD player with CD-DA components
- Mouse or joystick interface
- CD controller
- Connection to RGB monitor or TV
- Replacement of CD-DA
- Size of a video recorder

CD-I: Audio

	CD digital audio	CD-I „A" hi-fi music	CD-I „B" mid-fi music	CD-I „C" speech
sampling rate	44.1 kHz	37.8 kHz	37.8 kHz	18.9 kHz
freq. range	20 kHz	17 kHz	17 kHz	8.5 kHz
encod.	16 bit PCM	8 bit ADPCM	4 bit ADPCM	4 bit ADPCM
s/n ratio	98 dB	90 dB	60 dB	60 dB
max. playing time	74 min stereo	2.4 h stereo 4.8 h mono	4.8 h stereo 9.6 h mono	9.6 h stereo 19.2 h mono
appr. fidelity equival.	CD	mono LP	mono FM	mono AM

CD-I: Video

Coding of still images at different qualities and resolutions:

- Different amount of data
- Different data rates

YUV mode:

- Reproduction of natural images with many colors
- Encoding of changes of luminance and chominance values
- 360 x 240 pixel, 18 bit/pixel
- 262144 colors
- 360 * 240 * 18 bit/image = 194400 byte/image

Animations with few colors:

- Run-length encoding, about 10000 to 20000 Byte/image

MPEG for video encoding

CD-I: Video

Color Look-Up Table (CLUT):

- 4 bit/pixel (3.7 or 8 bit/pixel)
- For simple graphics fast data read-out
- Predefined color table
- 720 x 240 pixel, 16 colors
- 720 * 240 * 4 bit = 86400 Byte/image

RGB mode:

- For very good graphics
- 5 bit/pixel for each component
- 15 bit/pixel + 1 additional bit/pixel = 16 bit/pixel
- 360 x 240 pixel/image, 65538 colors
- 360 * 240 pixel/image * 16 bit/pixel = 172800 Byte/image

CD-I Ready Format

CD that can be played in CD-DA and CD-I players:

- Track pregap:
  - Contains CD-I specific information
  - Increased from 2-3 s to at least 182 s
- Audio players ignore the track pregap information and play only audio data part

Three different modes of play:

- Standard audio playback:
  - Track pregap information is ignored and only audio data played
- Reading, display and interpretation of the pregap data:
  - Audio data part is ignored
- Displaying pregap data as audio is played:
  - First step: loading CD-I information into player‟s RAM memory
  - Second step: start playing audio and information display

6.5. Principle of CD-Write and CD-Magneto optical

<u>CD-Write Once</u>

- So far, all of the CD technologies considered do not allow the user to write to the disk. Thus, the application scope is limited.
- This has led research laboratories to develop, besides the Read Only Storage Media, compact disks that can be recorded once or several times.
- The compact Disk write once (CD-WO), like WORM (write once Read Many),allows the user to write once to a CD and afterwards to read it many times.

Principle of the CD-WO

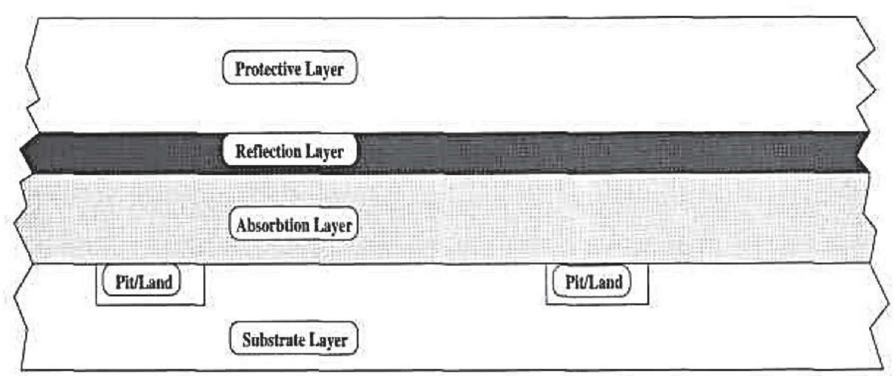


Figure 7.12: *Cross-section of a CD-WO disk.*

Principle of the CD-WO (cont..)

- Figure shows a cross-section of a CD-WO, vertical to the disk surface and data track.
- In the case of read-only CDs, the substrate (a polycarbonate) lies directly next to the reflection layer.
- In the case of a CD-WO, an absorption layer exists between the substrate and the reflection layer. This layer can be irreversibly (permanently) modified through strong thermal influence, which changes the reflection properties of the laser beams.

Principle of the CD-WO (cont..)

- In its original state, a CD-WO player recognizes a track consisting of lands.
- The **absorption layer** in the pre-grooved track is heated to above 250^0 C with a laser three to four times the intensity of a reading player. Hence, the absorption layer changes such that the reflection of the laser light now corresponds to a **pit**.
- This method determines the most remarkable property of the CD-WO: its data can be played by any devices which are meant only for read-only CDs.

6.5. Principle of CD-Write and CD-Magneto optical

<u>Compact Disk Magneto Optical</u>

- The Compact Disk Magneto Optical (CD-MO)has a high storage capacity and allows one to write multiple times to the CD.

Principle of the Magnetic-Optical Method

- The magnetic-optical method is based on the polarization of the magnetic field where the polarization is caused by a heat.
- To be written, the block (sector) is heated to above $150\,^0$C. Simultaneously, a magnetic field approximately 10 times the strength of the earth's magnetic field is created.
- The individual dipoles in the material are then polarized according to this magnetic field.
- Hereby, a pit corresponds to a low value of the magnetic field. A land is coded through a high value of the magnetic field.

Principle of the Magnetic-Optical Method (cont..)

- After the CD is irradiated (illuminated) with a laser beam, the polarization of the light changes corresponding to the existing magnetization. Using this process, the read operation is executed.
- For a delete activity, a constant magnetic field is created in the area of a block and the sector is simultaneously heated.

Areas of the CD-MO

- A CD-MO consists of an optional *read-only* area and a *write-many* (recordable) area.
- The read-only area (the premastered area) includes data which were written in a specified format onto the disk.
- CD-MO read-only area can be read by available playback devices.

6.6. Other Storage Media; DVD, Flash Drive, HD Cards, USB

<u>DVD</u>

Also known as: "Digital Versatile Disk"

Goal:

- **Create a new optical media to store an entire high-quality digital movie on a single side of a disk**

Technical overview of DVD:

- CD-like optical storage media
  - same size as CD $\Rightarrow$ allows for backward compatibility of reading devices
- Capacity considerably higher than CD
  - shorter pit/lands
  - tighter tracks
- EFM PLUS error correction scheme: more robust than CD scheme

Formats:

- single-sided single-layer
- single-sided double-layer: laser must switch focus to read both layers
- double-sided: disk must be flipped over to read both sides

DVD

- The 12 cm type is a standard DVD, and the 8 cm variety is known as a MiniDVD.
- These are the same sizes as a standard CD and a mini-CD, respectively.
- The capacity by surface area (MiB/cm^2) varies from 6.92 MiB/cm^2 in the DVD-1 to 18.0 MB/cm^2 in the DVD-18.

DVD

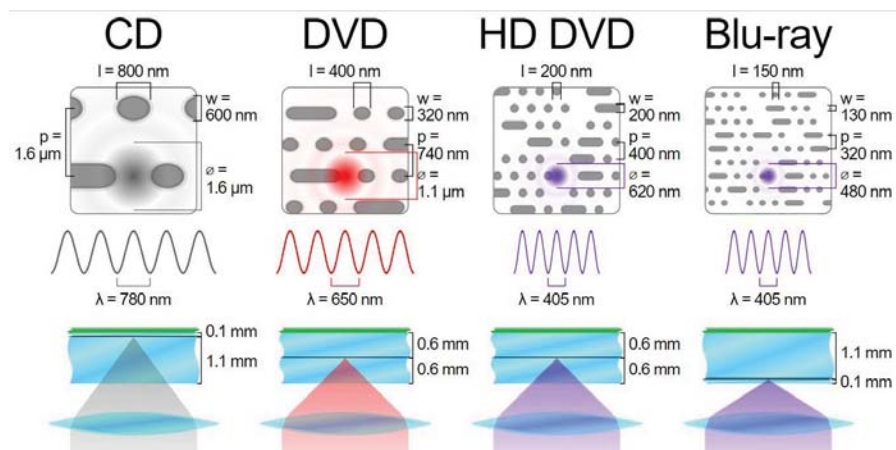
- Each DVD sector contains 2,418 bytes of data, 2,048 bytes of which are user data.
- There is a small difference in storage space between + and -(hyphen) formats:

- Scan of a DVD-R; the "a" portion has been recorded on while the "b" portion has not

DVD

- A dual-layer disc differs from its single layered counterpart by employing a second physical layer within the disc itself.
- The drive with dual-layer capability accesses the second layer by shining the laser through the first semitransparent layer.
- In some DVD players, the layer change can exhibit a noticeable pause, up to several seconds.
- This caused some viewers to worry that their dual-layer discs were damaged or defective, with the end result that studios began listing a standard message explaining the dual-layer pausing effect on all dual-layer disc packaging.

DVD



CD versus DVD

	CD	DVD
Disc diameter	120 mm	120 mm
Disc thickness	1.2 mm	1.2 mm
Laser wavelength	780 nm (infrared)	650 and 635 mm (red)
Track pitch	1.6 µm	0.74 µm
Min. pit/land length	0.83 µm	0.4 µm
Data layers	1	1 or 2
Sides	1	1 or 2
Data capacity	~ 650 MB	Single-Layer: 4.7 GB Dual-Layer: 8.5 GB Double-Sided: 9.4 GB

USB Flash Drive

- A **USB flash drive**, also variously known as a **thumb drive**, **pen drive**, **gig stick**, **flash stick**, **jump drive**, **disk key**, **disk on key** (after the original M-Systems DiskOnKey drive from 2000), **flash-drive**, **memory stick** (not to be confused with the Sony Memory Stick), **USB stick** or **USB memory**, is a data storage device that includes flash memory with an integrated USB interface.
- Most weigh less than 30 g (1 ounce).
- flash drives with anywhere from 8 to 256 GB are frequently sold; less frequent are 512 GB and 1 TB units.
- USB flash drives are often used for the same purposes for which floppy disks or CDs were once used; i.e. for storage, data back-up and transfer of computer files.

USB Flash Drive

- They are smaller, faster, have thousands of times more capacity, and are more durable and reliable because they have no moving parts.
- USB flash drives use the USB mass storage device class standard, supported natively by modern operating systems such as Windows, Linux, macOS and other Unix-like systems, as well as many BIOS boot ROMs.

Memory card

- A **memory card**, **flash card** or **memory cartridge** is an electronic flash memory data storage device used for storing digital information.
- These are commonly used in portable electronic devices, such as digital cameras, mobile phones, laptop computers, tablets, etc

MULTIMEDIA

Unit 1

1.1 Introduction

- Any kind of system that supports more than **one kind of media**
- Multimedia means integration of **continuous media** (e.g. audio, video) and **discrete media** (e.g. text graphics , images) through which digital information can be conveyed to the user in appropriate way.
  - ***Multi*** - many; much; multiple
  - ***Medium*** - a interleaving substance through which something is transmitted or carried on
- Classification based on perception (text, audio, video) is appropriate for defining multimedia

1.1 Introduction..

- **Multimedia** is defined as the **computer based** interactive environment that incorporates text, images, graphics, sound, audio, animation, video and virtual reality (VR).

1.1 Introduction..

Time always takes separate dimension in the media representation

- Based on **time-dimension** in the representation space, media can be
  - Time-independent (***Discrete***)
    - Text, Graphics
  - Time dependent (***Continuous***)
    - Audio, Video
    - Video, sequence of frames (images) presented to the user *periodically.*
    - Time dependent a periodic media is not continuous!!
- Discrete & Continuous here has no connection with internal representation!! (Relates to the viewers impression…)

1.1 Introduction..

- Multimedia is ***any*** combination of ***digitally manipulated*** text, art, sound, animation and video.
- A more strict version of the definition of multimedia do not allow just any combination of media.
  - It requires
    - Both continuous & discrete media to be utilized
    - Significant level of independence between media being used

1.1 Introduction..

- **Multimedia** is interactive when the end-user is allowed to control what and when the elements are delivered.
- Interactive Multimedia is Hypermedia, when the end-user is provided with the structure of linked elements through which he/she can navigate.

1.1 Introduction..

- **Multimedia** is ***linear***, when it is not interactive and the users just sit and watch as if it is a movie.
- **Multimedia** is ***nonlinear***, when the users are given the navigational control and can browse the contents at will.

1.2. The Medium Aspect
1.5 Media Combination and Independence

- **Multimedia System:** one way of defining multimedia can be found in the meaning of composed world multi-many, much multiple.
- **Medium:** An intervening substance through which some thing is transmitted or carried on.
- **Computer System medium:**
  1. Text
  2. Image
  3. Sound
  4. Video

1.2. The Medium Aspect
1.5 Media Combination and Independence….

Medium is defined as means for distribution and presentation of information.

Examples of a medium are text, graphics, speech, and music. Media can be classified with respect to different criteria.

We classify media according to perception, representation, presentation, storage, transmission, and information exchange.

1.2. The Medium Aspect
1.5 Media Combination and Independence….

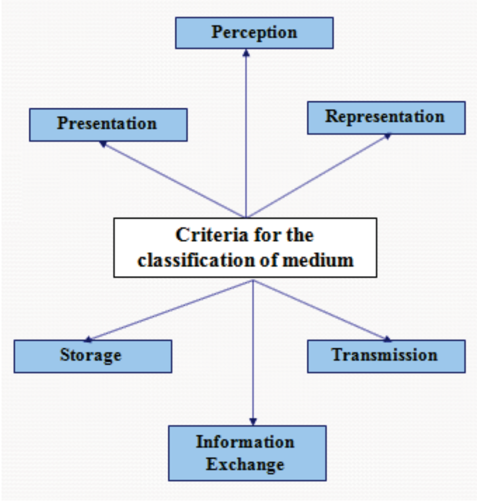


fig : Classification of Medium/Media

1.2. The Medium Aspect
1.5 Media Combination and Independence….

Classification of Media :

1. The perception media
2. The representation Media
3. The presentation Media
4. The storage media
5. The transmission media
6. The information Exchange media

1.2. The Medium Aspect
1.5 Media Combination and Independence….

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.

1.2. The Medium Aspect
1.5 Media Combination and Independence….

Classification of Media :

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 code
- Graphics are coded according to CEPT (Conference of European Postal and Telecommunications Administration) or CAPTAIN (Character and Pattern Telephone Access Information Networks) video text standard.
- Image can be coded as JPEG format
- Audio video sequence can be coded in different TV standard format (PAL, NTSC, SECAM and stored in the computer in MPEG format)

1.2. The Medium Aspect
1.5 Media Combination and Independence….

Classification of Media :

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.
- **Digital Media**: Soft copy presentation.
- **Paper Media**: Hard copy presentation.

1.2. The Medium Aspect
1.5 Media Combination and Independence….

Classification of 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 hard disk, CD-ROM, etc.

1.2. The Medium Aspect
1.5 Media Combination and Independence….

Classification of Media :

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.***

1.2. The Medium Aspect
1.5 Media Combination and Independence….

Classification of Media :

<u>6. The information Exchange media</u>

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.

1.3. Main properties of Multimedia

The uses of term multimedia are not every arbitrary combination of media. The main properties of multimedia system are:

1. *Combination of media*
2. *Computer support integration*
3. *Communication systems*
4. *Independence*

1.3. Main properties of Multimedia..

<u>1. Combination of media</u>

simple text processing program with in corporate image is often called a multimedia application.

Because ***two media are processed through one program***. But one should talk multimedia only when both continuous and discrete media are utilized. So text processing program with incorporated images is not a multimedia application.

1.3. Main properties of Multimedia..

<u>2. Computer support integration</u>

Computer is idle tools for media combinations. The system should be capable of computer-controlled media processing. The system should be programmable by a system programmer or even a user.

<u>3.Communication System</u>

Communication-capable multimedia systems must be approached. A reason for this is that most of today‟s computers are interconnected; considering multimedia functions from only the local processing viewpoint would be a restriction, if not a step back.

1.3. Main properties of Multimedia..

4. Independence:

An important aspect of different media is their level of independence from each other. In general there is a request for independence of different media but multimedia may requires several level of independence.

E.g. A computer controlled video recorder stores audio and video information's. There is inherently tight connection between two types of media. Both media are coupled together through common storage medium of tape. On the other hand for the purpose of presentation the combination of DAT (digital audio tape recorder) signals and computer available text satisfies the request for media independence.

1.4. Definition of Multimedia

Multimedia- Applications

Multimedia plays major role in following areas

- Instruction (Teaching)
- Business
  - Advertisements
  - Training materials
  - Presentations
  - Customer support services
- Entertainment
  - Interactive Games
- Enabling Technology
  - Accessibility to web based materials
  - Teaching-learning disabled children & adults
- Fine Arts & Humanities
  - Museum tours
  - Art exhibitions
  - Presentations of literature

1.4. Definition of Multimedia

Usages of Multimedia Application:

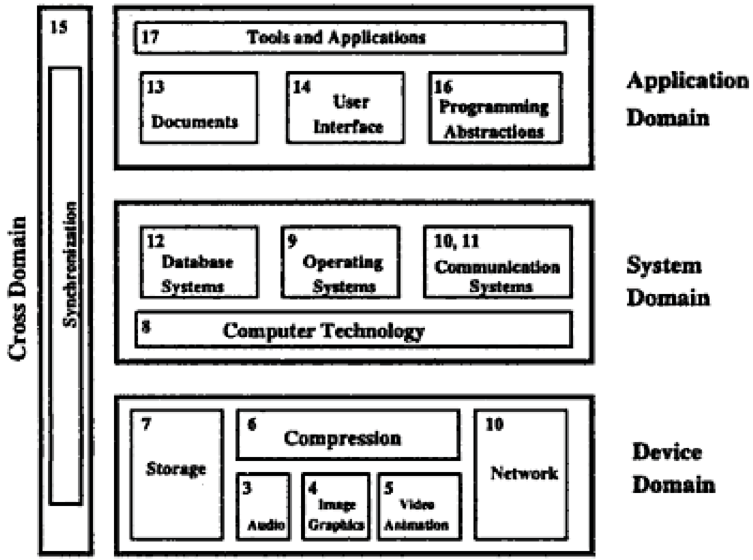
1. Education
2. Training
3. Entertainment
4. Advertisement
5. Presentation
6. Business Communication
7. Web page Design

1.4. Definition of Multimedia

Multimedia Tools:

- Text Editing Tools
- Drawing and Painting Tools
- Image Editing Tools
- Audio Editing Tools
- Video Editing Tools
- Animation and 3D Modeling Tools
- OCR (optical character recognition)Tools
- Voice Recognition Tools

Global Structure of Multimedia System



Global Structure of Multimedia System

1. Device domain
2. System domain
3. Application domain
4. Cross domain

Global Structure of Multimedia System

1. Device domain

It deals with interaction between multimedia application and multimedia devices

such as AGP(**Accelerated Graphics Port**) Card, Sound Card etc.

Basic concepts for the processing of digital audio and video data are based on digital signal processing. Different methods for the processing of image, graphics and animation are described.

The audio techniques section includes music MIDI (Musical Instrument Digital Interface) and speech processing.

Global Structure of Multimedia System

2. System domain

The interface between the device domain and the system domain is specified by the computer technology.

To utilize the device domain, several system services are needed. Basically, three services exits.

These services are mostly implemented in software. The operating system, serves as an interface between computer hardware/system and all other software components.

It provides the user with a programming and computational environment, which should be easy to operate.

The database system allows a structured access to data and a management of large databases. The communication system is responsible for data transmission according to the timing and reliability requirements of the networked multimedia.

Global Structure of Multimedia System

3. Application domain

Provides functions to the user to develop and present multimedia projects.

This includes software tools, and multimedia projects development methodology.

The services of the system domain are offered to the application domain through proper programming abstractions. Another topic embedded in the application domain is document handling.

Global Structure of Multimedia System

4. Cross domain

It turns out that, some aspects such as ***synchronization aspects***, are difficult to locate in one or two components or domains. The reason is that synchronization, being the temporal relationship among various media, relates to many components across all domains.

1.6. Traditional Data Stream Characteristics

In multimedia communication system data of discrete and continuous media are transmitted and information access takes place.

This transmitted information is divided into small individual unit known as ***packets***.

Packets can carry information of either continuous or discrete media

A sequence of individual packets transmitted in a time dependent fashion is called a *data stream* . The data stream will be used as a synonym ***data flow.***

1.6. Traditional Data Stream Characteristics

Transmission of information carrying different media leads to data stream with very different features.

The attributes of **synchronous**, **asynchronous** and **isochronous** data transmission conforms the field of computer communication and switching. Transmission of the data packets takes place in any of the following transmission modes:

Asynchronous Transmission mode:

Synchronous Transmission mode:

Isochronous Transmission mode:

1.6. Traditional Data Stream Characteristics

Asynchronous Transmission mode:

- The asynchronous transmission mode provides for **communication with no timely restriction (limit).**
- Packets reach **the receivers as fast as possible**.
- All information of discrete media can be transmitted as asynchronous data stream.
- If an asynchronous mode is chosen for **transmission of continuous media, additional technique most be applied to provide the time restriction.**
- E.g.: Ethernet, protocol of worldwide internet for e-mail transmission.
- **No time bound** (E.g. postal service)
- E.g. mail system.

1.6. Traditional Data Stream Characteristics

Synchronous Transmission mode:

- The synchronous transmission mode defines the ***maximum end to end delay for each packet of the data stream.***
- This upper bound will never be violated. (E.g. Ping operation)
- Moreover, a **packet can reach the receiver at any arbitrary(random) earlier time.** So most of the time the **receiver has to hold the packet temporarily**
- **A packet has a start frame and the end frame.**
- Start frame is used to tell the receiving station that a new packet of characters is arriving and used to synchronize the receiving station's internal clock.
- The end frame is used to indicate the end of packet.
- For retrieving uncompressed video at data rate 140Mbits/s & maximal end-to-end delay 1 second the receiver should have temporary storage 17.5 Mbytes

1.6. Traditional Data Stream Characteristics

Isochronous Transmission mode:

- **Defines maximum & minimum end-to-end delay**
- This means the delay jitter for individual packet is bounded.
- Isochronous transmission mode **minimizes the overhead of the receiver.**
- Upper time bound + lower time bound **(E.g. TV systems)**
- Data will reach destination in between these upper and lower bound time.
- Less storage buffer at receiver is needed than the synchronous transmission mode.
- Storage requirements at the receiver reduces

1.6. Traditional Data Stream Characteristics

Data stream characteristics for continuous media

Characteristics for continuous media can be classified according to:

Time interval between a complete transmission of consecutive packets

Periodicity

(a) Strongly Periodic
(b) Weakly Periodic
(c) Aperiodic

Variation in amount of consecutive packet amount

(a) Regular
(b) Weakly regular
(c) Irregular

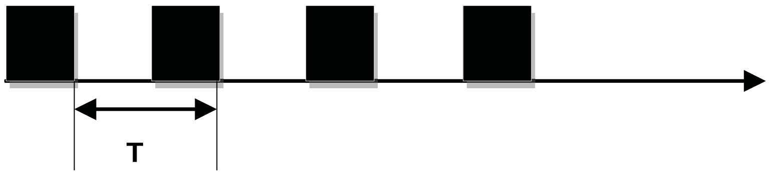
Contiguous packets

(a) Continuous
(b) Discrete

1.6. Traditional Data Stream Characteristics

Time interval based characteristics:

(1) Strongly periodic data stream transmission:

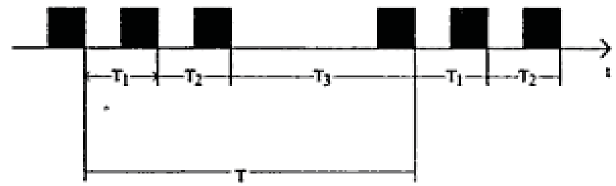


Time interval between two consecutive packets is constant. E.g. PCM-coded speech used in traditional telephone switching systems. It is also called strongly periodic.

1.6. Traditional Data Stream Characteristics

Time interval based characteristics:

(2) *Weakly periodic data stream transmission*:



Time interval between the consecutive packets is of periodic nature. Duration of time interval between two consecutive packets can be described by using a periodic function with finite period. But, time interval between two consecutive packets is not constant.

1.6. Traditional Data Stream Characteristics

Time interval based characteristics:

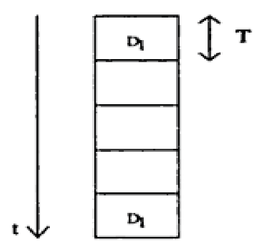
(3) Aperiodic data stream transmission:

Sequence of the time interval between packets is neither strongly nor weakly periodic. e.g. Cooperative application with shared Window.

1.6. Traditional Data Stream Characteristics

Packet Amount based characteristics:

(1) Strongly regular data stream

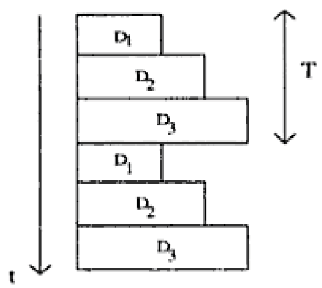


Data size of all the packets is constant. Amount of the data stays constant during the life time of a data stream. E.g. Uncompressed digital data transmission, video stream taken from a camera in uncompressed form, and the audio stream from an audio CD.

1.6. Traditional Data Stream Characteristics

Packet Amount based characteristics:

(2) Weakly regular data stream:

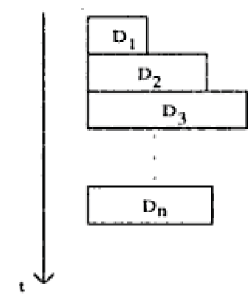


Data size of the packets changes periodically (with time). E.g. compressed video stream

1.6. Traditional Data Stream Characteristics

Packet Amount based characteristics:

(3) Irregular data stream



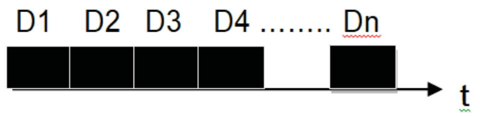
Data size of the packets is neither constant nor changing periodically according to a periodic function. Transmission and processing is more complicated.

1.6. Traditional Data Stream Characteristics

Contiguous packets based characteristics:

It characterizes continuity, or connection between consecutive packets. Is consecutive packets transmitted directly one after another, or is there a gap between the packets.

(1) *Continuous data stream*

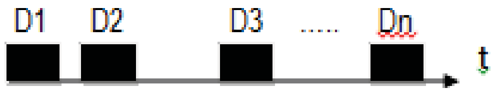


Packets are transmitted without intermediate gaps. Also called connected information transfer. It allows maximum data throughput. E.g. B-channel of ISDN with 64 Kb audio data transmission

1.6. Traditional Data Stream Characteristics

Contiguous packets based characteristics:

(2) Discrete data stream:



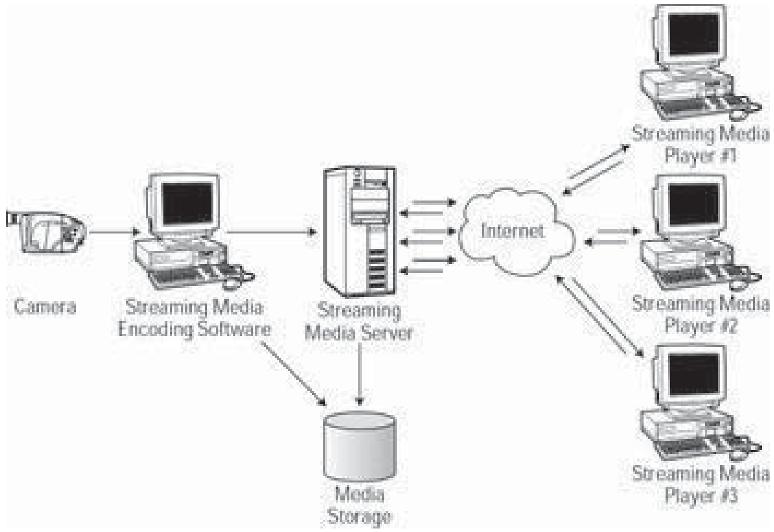
Gap exists among the packets. Also called unconnected data stream. Duration of the gap may vary.

1.6. Traditional Data Stream Characteristics

Streaming Media

- Popular approach to continuous media over the internet
- Playback at users computer is done while the media is being transferred (no waiting till complete download!!!)
- You can find streaming in
  - Internet radio stations
  - Distance learning
  - Cricket Live!!!
  - Movie Trailers

Streaming Media



1.7. Information Units

Continuous media consists of a time dependent sequence of individual information units. Such an information unit is called a LDU (logical data unit). With respect to time there can be Closed LDU and Open LDU.

- *Closed LDU*
  - Predefined duration is present.
  - E.g. data stream characteristics of audio samples in the computer.
- *Open LDU*
  - Duration is not known in advance.
  - E.g. Data stream sent from camera/ microphone to computer.

1.7. Information Units..

The most general names and best-known information units are the symphony and the movie. LDU is the whole symphony, individual sentences, individual notes, grouped samples of 1/75 second duration or just individual samples.

An example is an uncompressed video sequence consisting of individual video clips which presents a specific scene. Such a scene is comprised of a sequence of image. An image can be divided into group of pixels. Each pixel consist luminance and chrominance values. The image is therefore not the only possible LDU of a video sequence. A scene or a pixel also can be an LDU.

References:

- Multimedia: Computing, Communications and Applications", Ralf Steinmetz and Klara Nahrstedt, Pearson Education Asia
- "Multimedia Communications, Applications, Networks, protocols ad Standards", Fred Halsall, Pearson Education Asia
- "Multimedia Systems", John F. Koegel Buford, Pearson Education Asia

Assignments:

1. Define multimedia. Explain the application areas of multimedia.[8 marks/ 2015 fall]
2. What is Multimedia? Explain global structure of multimedia system. [8 marks/ 2014 fall]
3. Define application domain? Explain the different criteria that are used to classify media in the multimedia system. [8 marks/ 2013 fall]
4. Define Multimedia System. Explain properties of Multimedia System. [8 marks/ 2012 fall]
5. Write short notes
   1. Data stream with Transmission mode
   2. Data stream characteristics for continuous media
   3. Information unit

CHAPTER 2

Sound and Audio System

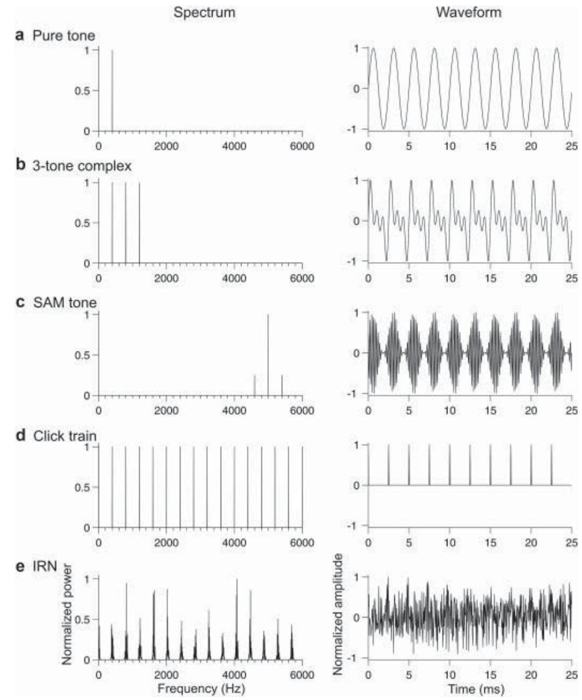
2.1. Basic Sound Concept: Representation and Format

- ***Sound is the physical phenomenon produced by the vibration of matter*** such as violin string or block of wood.
- When a matter vibrates, pressure variations are created in the air surrounding it.
- This alteration of high and low pressure is propagated through the air in a wave like motion. When the wave reaches the human ear, a sound is heard.

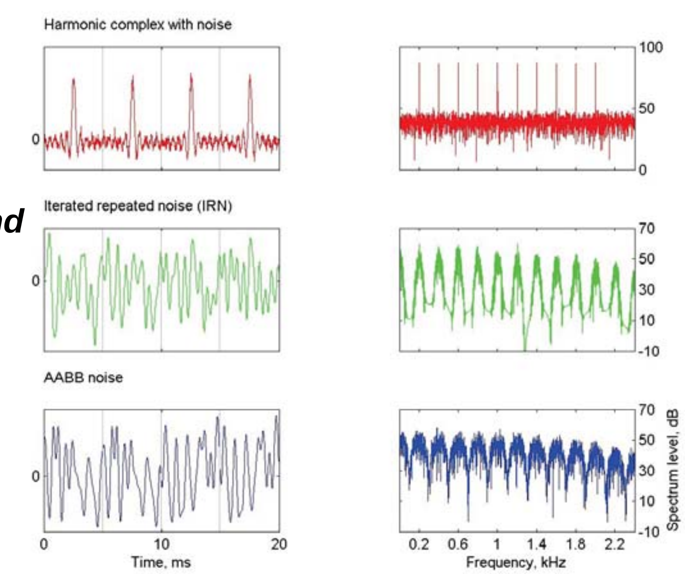
2.1. Basic Sound Concept: Representation and Format…

- The pattern of the oscillation is called a ***waveform***.
- The waveform repeats the same shape at regular intervals and this point is called a ***period***.
- ***Since sound wave forms occur naturally, Sound waves are never perfectly smooth or uniformly periodic.***
- ***Periodic sound****:* E.g. Musical instruments, vowel sounds, whistling wind, bird songs etc.
- ***Non-periodic sound****: E.g.* Unpitched percussion Instruments, coughs, sneezes, rushing water, Consonants,.

Periodic sound



Non Periodic sound



2.1. Basic Sound Concept: Representation and Format…

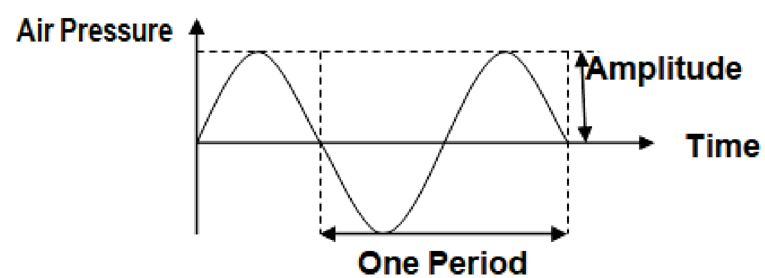


Figure: Oscillation of an air pressure wave

2.1. Basic Sound Concept: Representation and Format…

Frequency

- The frequency of a sound is the reciprocal value of the period. ***It represents the number of periods in a second***. It is measured in hertz (Hz) or cycles per second (cps).

1 KHz = 1000 Hz

Some of the frequency ranges are:

- Infra sound – 0 - 20 Hz
- Human audible sound – 20 Hz - 20KHz
- Ultra sound – 20KHz - 1GHz
- Hyper sound – 1GHz - 10THz

Human audible sound is also called audio or acoustic signals (waves). Speech is an acoustic signal produced by the humans.

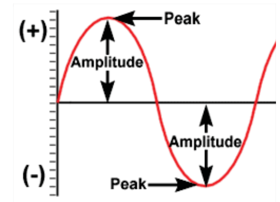
2.1. Basic Sound Concept: Representation and Format…

Amplitude

The amplitude of the sound is the measure of the displacement of the air pressure wave from its mean or quiescent state.

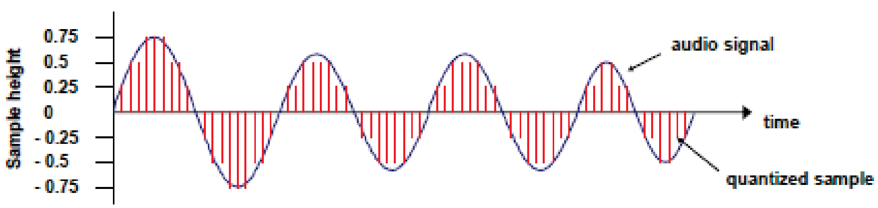
i.e. Distance moved by a point on a vibrating body or wave measured from its equilibrium position.

A sound has an amplitude property subjectively heard as loudness



2.1. Basic Sound Concept: Representation and Format…

Computer representation of sound



- A transducer converts pressure to voltage levels
- The analog signal is converted into a digital stream by discrete sampling:
- The analogous signal is sampled in regular time intervals, i.e. the amplitude of the wave is measured
- Discretization both in time and amplitude (quantization) to get representative values in a limited range (e.g. quantization with 8 bit: 256 possible values)
- Result: series of values:

0.25	0.5	0.5	0.75	0.75	0.75	0.5	0.5	0.25	0	- 0.25	- 0.5	…

2.1. Basic Sound Concept: Representation and Format…

Computer representation of sound

The smooth, continuous curve of a wave form is not directly represented in a computer. ***A computer measures the amplitude of the waveform at regular time intervals to produce a series of numbers called samples.***

Audio signals are converted into digital samples through Analog-to-Digital Converter (ADC). The reverse mechanism is performed by a Digital-to Analog Converter (DAC). E.g. of ADC is AM79C30A digital subscriber controller chip.

2.1. Basic Sound Concept: Representation and Format…

Sampling

Sound wave form the smooth, continuous is not directly represented in the computer.

The computer measures the amplitude of the wave form in the regular time interval to produce the series the numbers. Each of this measurement is called sample. This process is called sampling.

Sampling rate: the rate at which a continuous wave form is sampled is called sampling rate. Like frequency, sampling rate are measured in Hz. For loss less digitization the sampling rate should be at least twice of the maximum frequency response.

2.1. Basic Sound Concept: Representation and Format…

Sampling Rate..

Nyquist Sampling Theorem: “For lossless digitization, the sampling rate should be at least twice the maximum frequency responses”.

E.g. CD standard sampling rate of 44100Hz means that the waveform is sampled 44100 times per second. A sampling rate of 44100 Hz can only represent frequencies up to 22050 Hz.

2.1. Basic Sound Concept: Representation and Format…

Sample Height

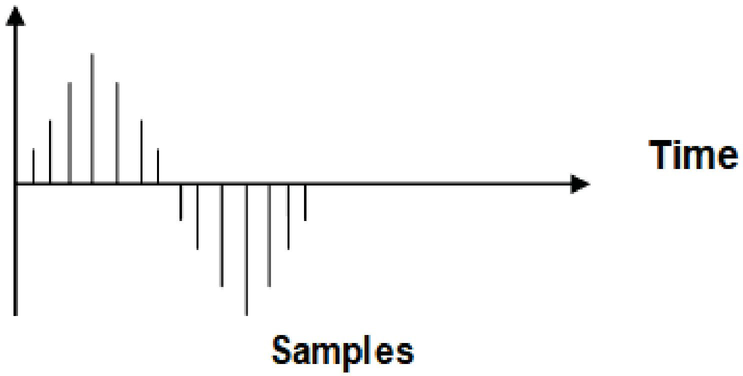


Figure: Sampled Waveform

2.1. Basic Sound Concept: Representation and Format…

Quantization:

The value of sample is discrete. Resolution/Quantization of a sample value depends on the no. of bits used in measuring the height of a waveform.

For E.g. An 8-bit quantization yields 256 values. 16-bit CD quality quantization results over 65536 values.

2.1. Basic Sound Concept: Representation and Format…

- **Quantization:**

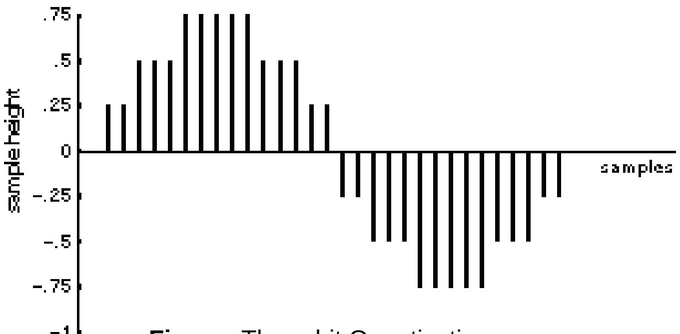
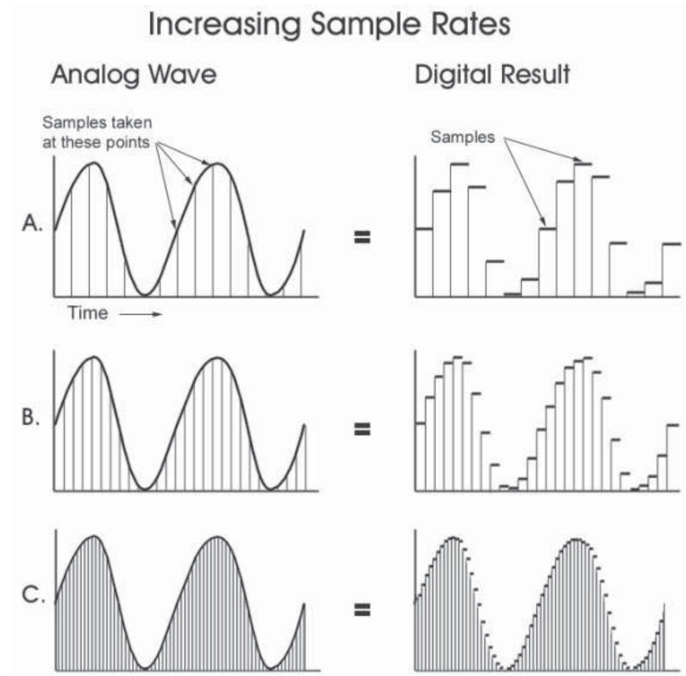
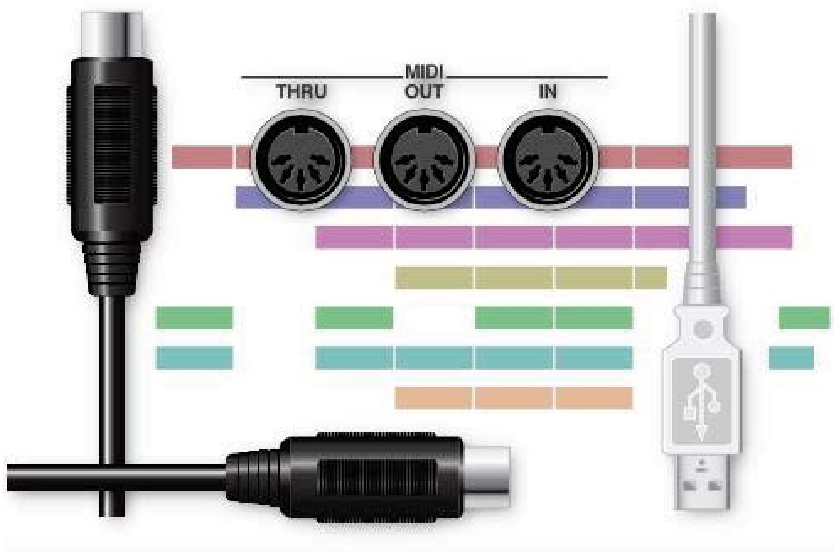


Figure: Three-bit Quantization

Lower the quantization; lower the quality of the sound. For 3-bit quantization, values are 8. i.e. 0.75, 0.5, 0.25, 0, -0.25, -0.5, -0.75 & 1



2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software



2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

- **Music:**

The relationship between music and computer has become more and more important, especially considering the development of **MIDI (Musical Instrument and Digital Interface)** and its important contribution in the music industry.

MIDI interface between musical instrument and computer.

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

- **MIDI Basic Concept:**

MIDI is the standard that manufacturer of electronic musical instrument have agreed upon.

It is set of specifications used for building the instrument so that the instrument of one manufacturer can without difficulty communicate musical information between one another.

So that **the instrument of different manufacturers can communicate without difficulty**. MIDI is defined in 1983.

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

MIDI Interface Components:

MIDI interface has two different components:

i. Hardware:

ii. A data format:

i. Hardware:

Hardware connects the equipment. It specifies the physical connection between musical instruments.

A MIDI Port is built in to the instrument and it specifies MIDI cable and deals with electronic signals that are sent over the cable.

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

ii. A data format:

Data format encodes the information travelling through the hardware. MIDI doesn't transmit an audio signal.

For this propose MIDI data format is used. The MIDI encoding includes, besides the instrument specification, the notion of the beginning and ending node, basic frequency and sound volume.

MIDI data format is digital. The data are grouped into MIDI message. Each MIDI message can communicate one musical event between instruments.

MIDI: Data Format

- Information traveling through the hardware is encoded in MIDI data format.
- The encoding includes *note* information like beginning of note, frequency and sound volume; upto 128 notes
- The MIDI data format is digital
- The data are grouped into MIDI messages
- Each MIDI message communicates one *musical event* between machines. An event might be pressing keys, moving slider controls,
- 10 mins of music encoded in MIDI data format is about 200 Kbytes of data. (compare against CD-audio!)

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

MIDI Reception Mode:

- Mode 1: Omni on / poly
- Mode 2: Omni on / mono
- Mode 3: Omni off / poly
- Mode 4: Omni off / mono

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

Omni On/Off:

If Omni if turned on, the MIDI device monitors all the MIDI channels and responds to all channels messages. If it is turned off, the MIDI device responds only to channel messages sent on the channels the device is set to receive.

Omni Poly/Mono:

Of Poly is set, the device can play several notes at a time. If the mono is set, the device plays notes like monophonic synthesizer-one note at a time.

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

The first half of the mode name specifies how MIDI device monitors the incoming MIDI channels. If Omni is turned on, the MIDI device monitors all the MIDI channels and responds to all channel messages, no matter which channels they are transmitted on. If the Omni is turned off, the MIDI device respond only to channel message, the device is set to receive.

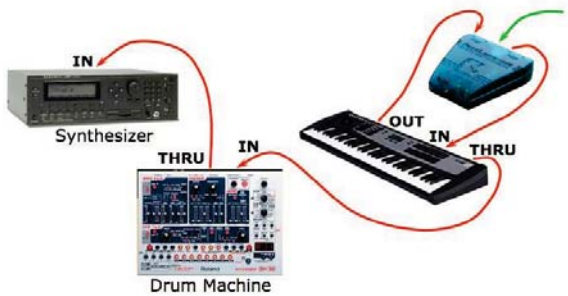
The second half of the mode name tells the MIDI device how to play nodes coming in over the MIDI cable. If the option Poly is set the device can play several nodes at a time. If the node is set to mono the device play only one node at a time like a monophonic synthesizer.

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

MIDI Devices

MIDI synthesizer device is the heart of MIDI system. Most **synthesizers** have following components.

1. ***Sound generators***
2. ***Microprocessor***
3. ***Keyboard***
4. ***Control panel***
5. ***Auxiliary controllers***
6. ***Memory***
7. ***Drum machine***
8. ***Master keyboard***
9. ***Guitar Synthesizer***



2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

- ***Sound generators***:
  It synthesizes the sound. It produces an audio signal that becomes sound when fed into a loud speaker. It can change quality of sound by varying the voltage oscillation of the audio. Sound generation is done in 2-ways:
  - Storing acoustic signals as MIDI data in advance
  - Creating acoustic signals synthetically
- ***Microprocessor***:
  Microprocessor communicates with the keyboard to know which notes the musician is playing. Microprocessor communicates with the control panel to know what commands the musician wants to sent to the microprocessor. The microprocessor then specifies note and sound commands to the sound generators (i.e. microprocessor sends and receives the MIDI message).
- ***Keyboard***:
  It affords the musician"s direct control of the synthesizer. Pressing keys means signaling microprocessor what notes to play and how long to play them. Keyboard should have at least 5 octaves and 61 keys.

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

- ***Control panel***:
  Controls those function that are not directly concerned with notes and duration. Control panel includes a slider, a button and a menu.
- ***Auxiliary controllers***:
  Gives more control over the notes played on keyboard. Pitch bend and modulation are the 2 common variables on the synthesizer
- ***Memory***:
  Stores patches for the sound generation and settings on the control panel.
- ***Drum machine***:
  Specialize in percussion sounds and rhythms.
- ***Master keyboard***:
  Increases the quality of the synthesizer
- ***Guitar Synthesizer***:
  Drum pad controllers, Guitar controllers and many more

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

Sequencer:

It is an electronic device in cooperating with both hardware and software, which is used as storage server for generated MIDI data. A sequencer may be computer. Sequencer transforms the note into MIDI message.

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

MIDI Message:

MIDI messages transmit information between MIDI devices and determine what kinds of musical events can be passed from device to device.

Formats of MIDI messages

- ***Status byte***: First byte of any MIDI message. It describes the kind of message
- ***Data byte***: The following bytes.

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

There are two types of MIDI messages

1. **Channel Message** (: *It go only to specified devices*)
   1. Channel Voice Message
   2. Channel mode Message
2. **System Message** (: *It go to all devices in a MIDI system*)
   1. System Real-time message
   2. System Common Message
   3. System exclusive Message

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

1. Channel Message: It go only to specified devices. There are two types of channel message.

i. **Channel Voice Message**: (*send actual performance between devices*)

Channel Voice message send actual performance data between MIDI devices describing keyboard action, controller action and control panel changes. E.g. note on, Note off, channel pressure, control change etc.

ii. **Channel Mode Message**: (*determine how to play notes*)

Channel Mode message determine the way that a receiving MIDI device respond to channel voice message. It deals with how to play notes coming in over MIDI cables. Channel mode message includes Omni On, Omni Off, note off, note on, etc. E.g. local control, All note off, Omni mode off etc.

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

2. System message: System message go to all devices in a MIDI system because no channel number are specified. There are three types of system message.

i. **System Real-time Message: (***avoid delay time***)**

These are very short and simple, consisting of only one byte. They carry extra data with them. ***These messages synchronize the timing MIDI devices in performance. To avoid delay these message are sent in the middle of other message if necessary***. E.g. System reset, Timing clock i.e. MIDI clock etc.

ii. **System Common Message: (***prepare sequence***)**

System common message are commands that ***prepares sequencers and synthesizer to play song***. E.g. Song selected, find the common starting place in the song.

iii. **System Exclusive message:** (*create customize message*)

System exclusive message allow MIDI manufacturers to ***create customized MIDI message to send between the MIDI devices.***

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

MIDI Software

The software application generally falls into four major categories.

1. Music recording and performance application
2. Musical notation and printing application
3. Synthesizer Patch editor and library patch
4. Music education application

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

1. ***Music recording and performance applications:*** Provides function as recording of MIDI messages. Editing and playing the messages in performance.
2. ***Musical notations and printing applications:*** Allows writing music using traditional musical notation. User can play and print music on paper for live performance or publication.
3. ***Synthesizer path editor and librarians:*** Allows information storage of different synthesizer patches in the computer's memory and disk drives. Editing of patches in computer.
4. ***Music education applications:*** Teaches different aspects of music using the computer monitor, keyboard and other controllers of attached MIDI instruments.

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

MIDI and SMPTE Timing Standard

MIDI reproduces traditional note length using MIDI clock, which are represented through timing clock message. Using MIDI clock a receiver can synchronize with the clock cycle of sender. To keep the standard timing reference the MIDI specification state **24 MIDI clock = 1 quarter note**.

As an alternative the SMPTE timing standard (***Society Of Motion Picture and Television Engineer***) can be used. The SMPTE timing standard was originally developed by NASA as a way to make incoming data from different tracking stations so that the receiving computer could fill what time each piece of data was created. SMPTE format consists of hour: minutes: second: frames: bits

-30 frames per second, SMPTE uses a 24 hour clock from 0 to 23 before recycling.

2.2.Basic Music (MIDI) Concept: Devices, Messages, Standards and Software

Processing chain of interactive computer music systems

- ***Sensing stage:*** Data is collected from controllers reading the gesture information from human performers on stage.
- ***Processing stage****:* Computer reads and interprets information coming from the sensors and prepares data for the response stage.
- ***Response stage:*** Computer ad some collection of sound producing devices share in realizing a musical output

Some of the interactive music systems:

- Max (OOP language)
- Cypher (It has a listener and a player)
- NeXT Computer (It has a music kit)
- M & Jam Factory (It has graphics control panel)

2.3. Speech : Generation, Analysis and Transmission

Speech

Speech can be „perceived", „understood", and „generated" by humans and by machines too. Human speech signal comprises a subjective lowest spectral component known as the pitch. Pitch is not proportional to the frequency. Human ear is sensitive in range from 600Hz-6000Hz.

So, a human adjust himself/ herself very efficiently to different speakers and their speech habit. The human brain can recognize the very fine line between speech and noise.

2.3. Speech : Generation, Analysis and Transmission

Properties of Speech Signals:

1. **Voice speech signals** show during certain time intervals almost periodic behavior. Therefore we can consider these signals as ***quasi-stationary*** signals for around 30ms
2. **The spectrum of the audio signals** shows characteristics maxima, which are mostly 3-5 frequency bands. These maxima called format.

2.3. Speech : Generation, Analysis and Transmission

Speech Generation:

Mid 19th century, Helmholtz built a mechanical vocal tract coupling together several mechanical resonators with which sound could be generated. In 1940, Dudley produced the 1st speech synthesizer through imitation of mechanical vibration using electrical oscillation.

Requirement for speech generation is real time signal generation.

1. Speech output system could transfer text into speech automatically without any lengthy processing.
2. Generated speech must be understandable and must sound natural.

2.3. Speech : Generation, Analysis and Transmission

Basic Notions:

- A ***phone*** is the smallest speech unit, such as the m of mat and the b of bat in English that distinguishes one utterance or word from another in a given language.
- ***Allophones*** mark the variants of a phone. For e.g. the aspirated p of pit and the un-aspirated p of spit are allophones of the English phoneme p.
- The ***morph*** marks the smallest speech unit which carries a meaning itself. Therefore, consider is a morph, but reconsideration is not.
- The ***voiced sound*** is generated through the vocal cords, m, v, l are the examples of voiced sounds.
- During the generation of an ***unvoiced sound*** the vocal cords are opened. F and s are unvoiced sounds.

2.3. Speech : Generation, Analysis and Transmission

Types of Speech generation:

The important requirement of speech generation is the generation of the real time signals. The easiest method for speech generation is to use pre-coded speech and play it back in the timely fashion.

2.3. Speech : Generation, Analysis and Transmission

Time-dependent sound concatenation:

CRUM

Individual speech units are composed like building blocks, where the composition can occur at different levels. Individual phones are understood as speech units. Example, **CRUM** phones word phone are shown individually as -

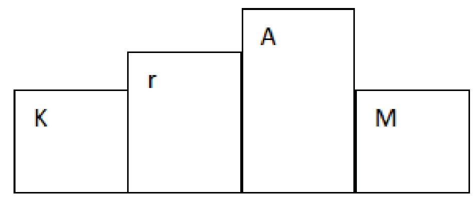


Fig: Phone sound concatenation

2.3. Speech : Generation, Analysis and Transmission

Time-dependent sound concatenation:

CRUM

It is possible with just a few phones to create from an unlimited vocabulary. The phone sound concatenation shows the problem during the transition between individual phones. This problem is called co articulation which is ***mutual sound effect***. To solve these problems, Di-phone sound concatenation is used. Two phones can constitute a Di-phone. In the above figure, Di-phone of word CRUM is shown.

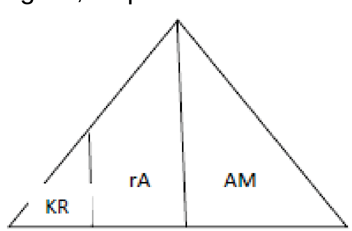


Fig: Di-phone concatenation

2.3. Speech : Generation, Analysis and Transmission

Time-dependent sound concatenation:

CRUM

The transition problem is not solved sufficiently on this level. To make the transition problem easier, syllabus can be created. The speech is generated through set of syllabus. The above given figure word sound concatenation and syllabus sound shows the syllabus sound of word CRUM. The best pronunciation of word is achieved is storage of whole word. This leads towards synthesize the speech sequence.

2.3. Speech : Generation, Analysis and Transmission

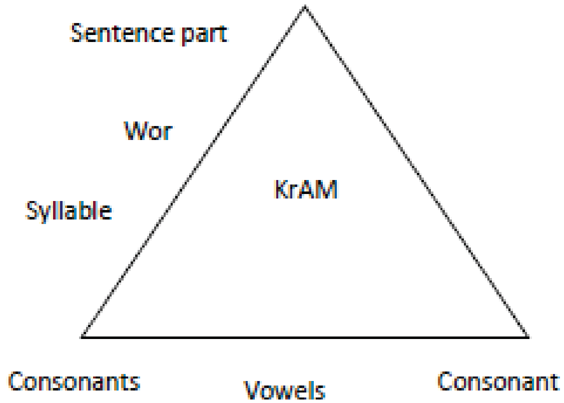


Fig: Word sound concatenation

2.3. Speech : Generation, Analysis and Transmission

- **co articulation**

The transition between individual sound units create an essential problem called co articulation which is the mutual sound influence throughout the several sound.

2.3. Speech : Generation, Analysis and Transmission

Frequency dependent sound concatenation:

Speech generation/output can also be based on a frequency dependent sound concatenation.

This can be done through a ***formant synthesis***. Formants are frequency maxima in the spectrum of the speech signal. Formants synthesis simulates the vocal tract through a filter.

A pulse signal with a frequency, corresponding to the fundamental speech frequency, is chosen as a simulation for voiced sounds. Unvoiced sounds are created through a noise generator.

The method used for the sound synthesis in order to simulate human speech is called the linear predictive coding (LPC) method. Using speech synthesis, an existent text can be transformed into an acoustic signal.

2.3. Speech : Generation, Analysis and Transmission

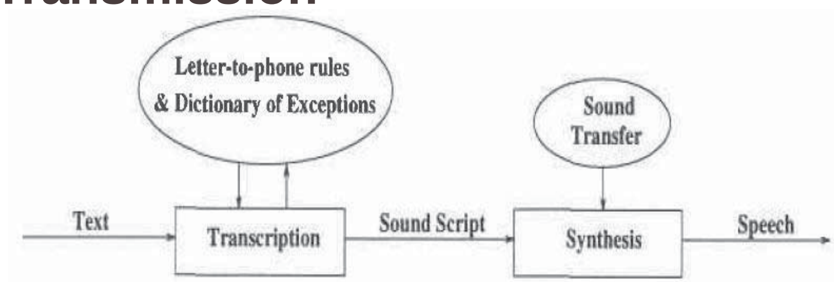


Fig: Components of a speech synthesis system with time-dependent sound concatenation.

In the first step the transcription is performed where text is translated into sound script. Most transcription methods work with later to phone rules. A dictionary of exception stored in the library.

In the second step the sound script is translated into a speech signals. Besides the problem of co-articulation ambiguous pronunciation most be considered.

2.3. Speech : Generation, Analysis and Transmission

Step 1:

- Performs transcription
- Text is translated into sound script
- This process is done using letter-to-phone rules and dictionary of exceptions
- User recognizes the formula deficiency in the transcription and improves the pronunciation manual

Step 2:

- Sound script is translated into a speech signal.
- Time or frequency dependent concatenation can follow

2.3. Speech : Generation, Analysis and Transmission

Speech Analysis

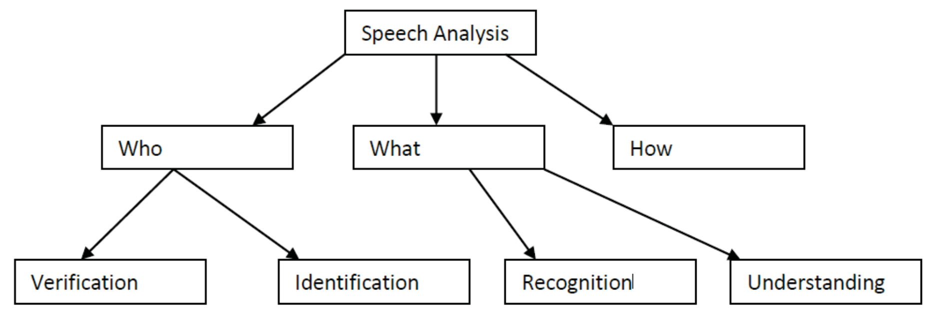


Fig. Speech analysis deals with research area shown in the above figure.

2.3. Speech : Generation, Analysis and Transmission

Speech Analysis

- Human speeches have certain characteristics determined by speaker. So speech analysis can server to analyze who is speaking. i.e. to recognize a speaker, for his identification and verification. The computer identifies and verifies the speaker using an acoustic finger print is digitally stored. It is digitally stored speech probe (certain statement of the speaker).
- Another main part of speech analysis is to analyze what has been said. To recognize and understand the speech itself.
- Another area of speech analysis tries to research speech pattern with respect to how the statement was said. E.g. a speaker sentence sounds differently if a person is angry or happy.

2.3. Speech : Generation, Analysis and Transmission

Speech Recognition System and Understanding:

The system which provides the recognition and understanding of speech signal applies this principle several times as follows:

1. In the first step, the principle is applied to a sound pattern and/or word model. An acoustic and **phonetic analysis is performed.**
2. In the second step, certain speech units go through **syntactical analysis**. In this step the errors in the previous step can be recognized.
3. The third step deals with **semantics(the study of meanings**) of the previously

2.3. Speech : Generation, Analysis and Transmission

Speech Recognition System and Understanding:

There are still many problems into speech recognition and understanding research.

1. Specific problem is presented by Room Acoustic with existed environment noise.
2. Word boundaries must be determined. Very often neighboring words follows into one another.
3. For the comparison of the speech elements to the existing pattern, time normalization is necessary. The same word can be spoken quickly or slowly.

2.3. Speech : Generation, Analysis and Transmission

Speech Transmission

The area of speech transmission deals with efficient coding of the speech signal to allow speech/sound transmission at low transmission rates over networks.

The goal is to provide the receiver with the same speech/sound quality as was generated at the sender side. This section includes some principles that are connected to speech generation and recognition'

2.3. Speech : Generation, Analysis and Transmission

Speech Transmission

Signal Form Coding

This kind of coding considers no speech-specific properties and parameters. Here, the goal is to achieve the most efficient coding of the audio signal. The data rate of a PCM-coded stereo-audio signal with CD-quality requirements is:

Rate = 2 * 44100/s * 16 bit / 8 bit/byte

= 176400b bytes/s = 1411200 bits / s

Telephone quality, in comparison to CD-quality, needs only 64 Kbit/s. Using Difference Pulse Code Modulation (DPCM), the data rate can be lowered to 56 Kbits/s without loss of quality. Adaptive Pulse Code Modulation (ADPCM) allows a further rate reduction to 32 Kbits/s.

2.3. Speech : Generation, Analysis and Transmission

Speech Transmission

Source Coding

Parameterized systems work with source coding algorithms. Here, the specific speech characteristics are used for data rate reduction. Channel vo-coder is an example of such a parameterized system.

The channel vo-coder is an extension of a sub-channel coding. The signal is divided into a set of frequency channels during speech analysis because only certain frequency maxima are relevant to speech.

A **vocoder** is a category of voice codec that analyzes and synthesizes the human voice signal for audio data compression, multiplexing, voice encryption, voice transformation, etc.

vocoder

2.3. Speech : Generation, Analysis and Transmission

Speech Transmission

Source Coding

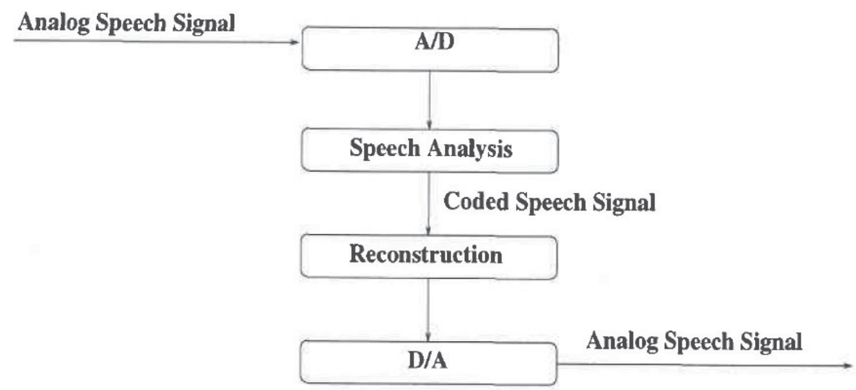


Fig: Source coding in parameterized systems: components of a speech transmission system.

2.3. Speech : Generation, Analysis and Transmission

Speech Transmission

Recognition / Synthesis (Mixture) Methods

There have been attempts to reduce the transmission rate using pure recognition/ synthesis methods. Speech analysis (recognition) follows on the sender side of a speech transmission system and speech synthesis (generation) follows

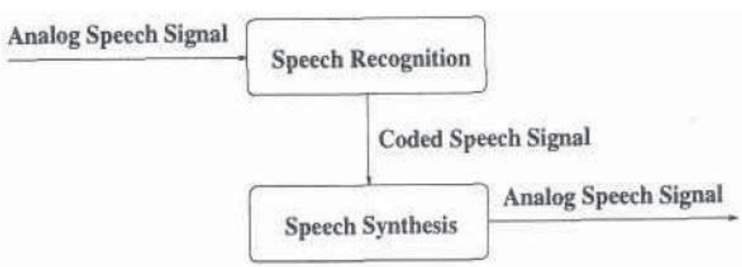


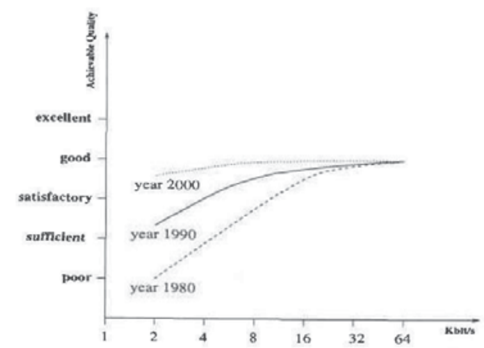
Fig: Recognition/synthesis systems: components of a speech transmission system.

2.3. Speech : Generation, Analysis and Transmission

Speech Transmission

Achieved Quality

The essential question regarding speech and audio transmission with respect to multimedia systems is how to achieve the minimal data rate for a given quality.



References:

- Multimedia: Computing, Communications and Applications", Ralf Steinmetz and Klara Nahrstedt, Pearson Education Asia
- "Multimedia Communications, Applications, Networks, protocols ad Standards", Fred Halsall, Pearson Education Asia
- "Multimedia Systems", John F. Koegel Buford, Pearson Education Asia

CHAPTER 3

IMAGES AND GRAPHICS

3.1. Basic Image Concept: Representation and Format

Introduction to Images:

Image is the spatial representation of an object. It may be 2D or 3D scene or another image. Images may be real or virtual. It can be abstractly thought of as continuous function defining usually a rectangular region of plane. **Example**:

- **Recorded image-** photographic, analog video signal or in digital format
- **Computer vision-** video image, digital image or picture
- **Computer graphics-** digital image
- **Multimedia-** deals about all above formats

3.1. Basic Image Concept: Representation and Format

Image:

An image may be defined as two dimensional function f(x, y) where, x y are the spatial co-ordinate and the amplitude of 'f' at any pair of co-ordinates x, y is called the intensity or gray level of image at that point x, y and amplitude values of f are all finite, discrete quantities we all the image is digital image.

3.1. Basic Image Concept: Representation and Format

Digital Image Representation:

A digital image is represented by a matrix of numeric values each representing a quantized intensity value. When *I* is a two-dimensional matrix, then *I(r,c)* is the intensity value at the position corresponding to row *r* and column *c* of the matrix.

Or

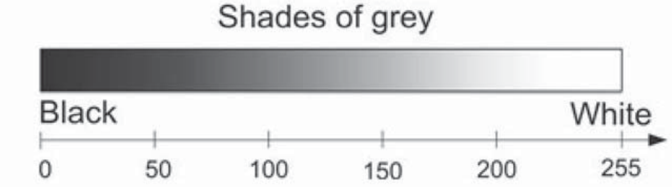
Digital image is represented by two dimensional matrix e.g. **I[x] [y]** .When I is the two dimensional matrix then, **I[x][y]** is the intensity value at the position corresponding to the row x and column y of the matrix.

3.1. Basic Image Concept: Representation and Format

Digital Image Representation:

The points at which an image is sampled are known as picture elements, or pixels.

The pixel values of intensity images are called gray scale levels. If there are just two intensity values (e.g., black and white), they are represented by the numbers 0 and 1. Such images are called binary-valued images. When 8-bit integers are used to store the intensity values, the gray levels range from 0 to 255.

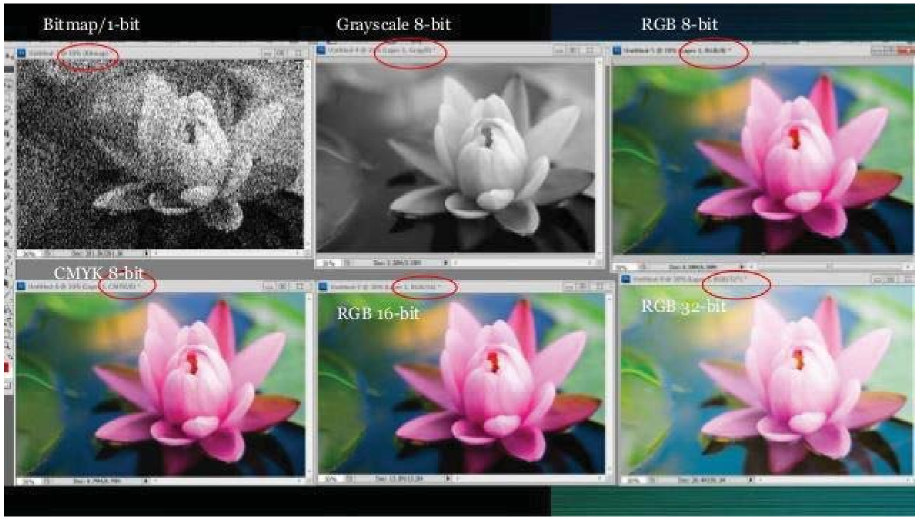


3.1. Basic Image Concept: Representation and Format

Digital Image Representation:

Digital images are often very large, because a number of bits (e.g., 8 bits for 256 discrete gray levels) are required for each pixel, and a large number of pixels (e.g., 640 X 480) are required for representing the images.

a) Binary Image b) Gray-Scale Image c) Color Image



3.1. Basic Image Concept: Representation and Format

Intensity

If n bit integers are used to store each pixel‟s intensity value, 2^n different intensities (colors) are possible numbered from 0 to 2^n -1.

Intensity is used to describe the brightness and purity of a **color**. When a hue is strong and bright, it is said to be high in **intensity**. When a **color** is faint, dull and gray, it is said to be low in **intensity**. When describing a hue, value refers to its lightness or darkness.

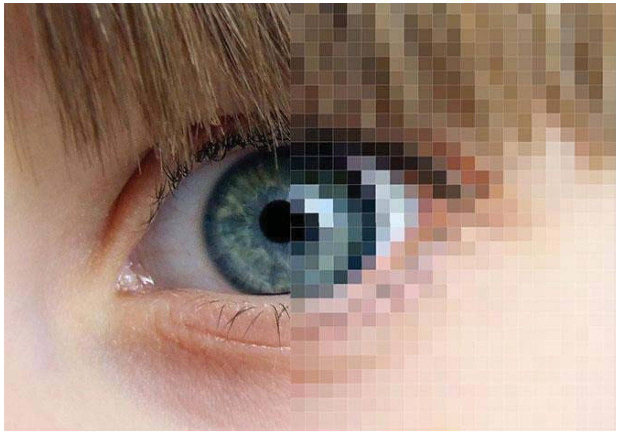
3.1. Basic Image Concept: Representation and Format

Intensity

3.1. Basic Image Concept: Representation and Format

Image resolution

It depends on the distance between pixel grids and imaging system also (persistence and spot profile).



3.1. Basic Image Concept: Representation and Format

Image file-size

It depends on total number of pixels and no. of bits per pixel (quantization).

Image file size = W*L*n bits

Where,

W = width (pixels)

L = length or height (pixels)

n = number of bits per pixel

Numerical

Q.N : Consider three different raster system with resolution 640 by 400, 1280 by 1024 and 2560 by 2048. what size frame buffer (in byte) is needed for each of the system to store 12 bits per pixel ? How much storage is required for each system if 24 bit per pixel are to be stored.

Ans :

1) for 12 bit per pixel system

the frame buffer(in byte) size needed

a. for 640 by 400 resolution = $\frac{(640 \text{ X } 400 \text{ X } 12)}{8}$

=384000 byte

b. for 1280 by 1024 resolution = $\frac{(1280 \text{ X } 1024 \text{ X } 12)}{8}$

=1966080 byte

c. for 2560 by 2048 resolution = $\frac{(2560 \text{ X } 2048 \text{ X } 12)}{8}$

=7864320 byte

Cont…..

2) for 24 bit per pixel system

the frame buffer(in byte) size needed

a. for 640 by 400 resolution = $\frac{(640 \text{ X } 400 \text{ X } 24)}{8}$

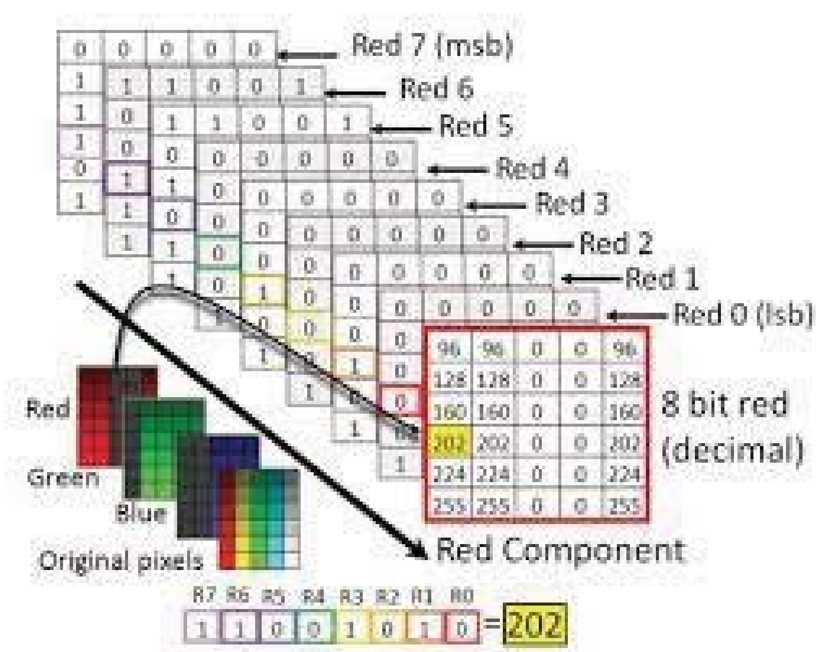
=768000 byte

b. for 1280 by 1024 resolution = $\frac{(1280 \text{ X } 1024 \text{ X } 24)}{8}$

=3932160 byte

c. for 2560 by 2048 resolution = $\frac{(2560 \text{ X } 2048 \text{ X } 24)}{8}$

=15728600byte



3.1. Basic Image Concept: Representation and Format

Image Formats

Image formats are basically of two kinds:

(1) **Captured Image Format**

(2) **Stored Image Format**

3.1. Basic Image Concept: Representation and Format

Image Format

1) Capture Image format

This is the format that comes out from an image frame grabber, such as VideoPix card, Parallax XVideo, etc. It is specified by mainly two parameters:

- Spatial Resolution (specified by pixel × pixel)
- Color encoding (specified by bits per pixel)

Both these parameter values depend on hardware and software for the input/output of images.

For example, for image capturing on a SPARCstation, the VideoPix card and its software are used. The spatial resolution is 320x240 pixels and the color can be encoded with 1-bit (a binary image format), 8-bit (color or grayscale) or 24-bit (color-RGB).

3.1. Basic Image Concept: Representation and Format

Image Format

1) Capture Image format

Captured Image Format is specified by:

spatial resolution (pixels x pixels) and *color encoding* (bits per pixel)

Examples:

VideoPix™ / SunVideo™ card:

spatial resolution: 320 x 240 pixels

color encoding: 1-bit (binary image), 8-bit (color or greyscale) 24-bit (color-RGB)

SunVideo™ card (capture and compression card with 30* fps (frames per second)):

CellB	30 fps	MPEG1 IP frames	17 fps
JPEG	30 fps	Capture YUV	30 fps
MPEG1 I frames	30 fps	Capture RGB-8/24	30 / 12 fps

* At least 25 frames per second are necessary if continuous motion is to be presented. However, it is possible to "cheat" a little and to work with smaller frame rates.

3.1. Basic Image Concept: Representation and Format

Image Format

2) Stored Image Format

While storing an image, we store a two-dimensional array of values, in which each value represents the data associated with a pixel in the image. For a bitmap, this value is a binary digit.

For a color image (pixmap), the value may be a collection of:

- Three numbers representing the intensities of the red, green and blue components of the color at that pixel.
- Three numbers that are indices to tables of the red, green and blue intensities.
- A single number that is an index to a table of color triples.
- An index to any number of other data structures that can represent a color.
- Four or five spectral samples for each other.

3.1. Basic Image Concept: Representation and Format

Image Format

2) Stored Image Format…

In addition, each pixel may have other information associated with it; e.g., three numbers indicating the normal to the surface drawn at that pixel.

Information associated with the image as a whole, e.g., width, height, depth, name of the creator, etc. may also have to be stored.

The image may be compressed before storage for saving storage space.

Some current image file formats for storing images include GIF, X11 Bitmap, Sun Rasterfile, PostScript, IRIS, JPEG, TIFF, etc.

3.1. Basic Image Concept: Representation and Format

Graphics Format

Graphics image formats are specified through graphics primitives and their attributes. Graphics primitives include lines, rectangles, etc. specifying 2D objects or polyhedron, etc. specifying 3D objects. A graphics package determines which primitives are supported. Attributes of the graphics primitives include line style, line width, color effect, etc., that affect the outcome of the graphical image.

3.1. Basic Image Concept: Representation and Format

Graphics Format

Graphics primitives and their attributes represent a higher level of an image representation where the graphical images are not represented by a pixel matrix. As an advantage, data to be stored per graphical image is reduced, and manipulation of the graphical image becomes easier. However, this higher level of representation needs to be converted at some point of the image processing into the lower level of the image representation, e.g., when the image is to be displayed. Graphics packages such as SRGP can be used for this conversion (from graphics primitives to bitmap or pixmap). A bitmap is an array of pixel values with one bit for each pixel. A pixmap is an array of pixel values with multiple bits (e.g., 8 bits for 256 colors) for each pixel.

3.1. Basic Image Concept: Representation and Format

Graphics image formats **are specified through:**

- *graphics primitives*: lines, rectangles, circles, ellipses, text strings (2D), polyhedron (3D)
- *attributes*: line style, line width, color affect.

Graphics primitives and their attributes represent a higher level of an image representation. The graphics package determines which primitives are supported.

Advantages:

+ Reduction of the graphical image data
+ Easier manipulation of graphical images.

Disadvantage:

- Additional conversion step from graphical primitives and attributes to its pixel representation

Formats:

- SRGP (Simple Raster Graphics Package), one way conversion to bit-/pixmap
- PHICS (Programmer's Hierarchical Interactive Graphics Systems) and
- GKS (Graphical Kernel System) only image representation is in pixmap

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Computer Image Processing

Computer graphics concern the pictorial synthesis of real or imaginary objects from their computer-based models. But Image Processing treats in converse process i.e. analysis of scenes or reconstruction of models from pictures of 2D and 3D objects. Computer image processing comprises of image synthesis (generation) and image analysis (recognition).

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Image Synthesis

Image synthesis deals with the generation of images of real or imaginary objects. It is an integral part of all computer user interfaces and is essential for visualizing 2D, 3D and higher dimensional objects. Some of the applications of image synthesis (areas which use image synthesis) are:

(1) *User Interface*

(2) *Office automation and electronic publishing*

(3) *Simulation and Animation for Scientific Visualization and Entertainment*

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

(1) User Interface

- Point-and-click facility
- Menu-driven

(2) Office automation and electronic publishing

- Desktop publishing
- Electronic publishing
- Hypermedia Systems

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

(3) Simulation and Animation for Scientific Visualization and Entertainment

- Simulation of real time systems
- Visualization of time-varying behavior of systems
- Abstract representation of complex mathematical expressions
- Models for fluid flow, chemical reaction etc.
- Cartoons
- Flying logos and more exciting visual for movies

Image Synthesis can be dynamic. Similarly, interactive graphics systems are used for image synthesis.

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Image synthesis **(generation) is an integral part of all computer graphical user interfaces and indispensable for visualising 2D, 3D and higher dimensional objects. E.g.:**

***Graphical User Interface (GUI):* desktop windows system with icons and menu items**

- Office Automation and Electronic Publishing: desktop publishing, Hypermedia systems
- Simulation and Animation for Scientific Visualisation and Entertainment

Pictures can be dynamically varied by adjusting the animation speed, portion of the total scene in view, the amount of details shown etc.

- *Motion Dynamics:* Objects are moved and enabled with respect to a stationary or also dynamic observer, e.g. flight simulator.
- *Update Dynamics:* Objects being viewed are changed in shape, color, or other properties, e.g. deformation of an in-flight aeroplane structure.

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Dynamics in Graphics

Graphics are not confined to static pictures. Pictures can be dynamically varied, e.g., the portion of the total scene in view, amount of detail shown, etc. of an animation can be controlled by the user. Dynamics can be of two kinds:

Motion Dynamics:

With motion dynamics

- Objects can be moved and enabled with respect to a stationary observer.
- The object can remain stationary and the view around it can move.
- Both the objects and the camera can move.

Update Dynamics:

Update Dynamics is the actual change of the shape, color, or other properties of the objects being viewed.

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Image Analysis

Image analysis is concerned with techniques for extracting descriptions of objects from images. Image analysis techniques include computation of perceived brightness and color, partial or complete recovery of three-dimensional data in the scene, location of discontinuities corresponding to objects in the scene and characterization of the properties of uniform regions in the image.

Image analysis involves activities where the different properties of the image, its orientation and other data related to it is extracted, evaluated, compared in order to get some result. It may deal with calculating the intensity, hue, saturation of the image, the centroid, and identification of noise, detection of patterns or recognition of the image itself. It might result in a complete recovery of 3D data in the scene.

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Image analysis has profuse importance in several areas and they are

- Criminology
- Biometrics
- Analysis of aerial surveillance photographs
- Medicine
- Analysis of slow scan television images of the moon or of plates gathered from space probes
- Machine vision

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Image analysis is used in areas such as

- Aerial surveillance photographs
- Slow scan images of the moon or of planets gathered from space probes
- Television images taken from an industrial robot‟s visual sensor
- X-ray images and computerized axial tomography (CAT) scans, etc.

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Sub-areas of image analysis includes

- Image Enhancement

  It deals with improving image quality by eliminating noise (extraneous or missing pixels) or by enhancing contrast.

- Pattern Detection and Recognition

  It deals with detecting and clarifying standard patterns and finding distortions from these patterns.

- Scene Analysis and Computer Vision

  It deals with recognizing and reconstructing 3D models of a scene from several 2D images. An example is an industrial root sensing the relative sizes, shapes, positions and colors of objects.

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Image Recognition

To fully recognize an object in an image means knowing that there is an agreement between the sensory projections and the observed image. How the object appears in the image is specified by the spatial configuration of the pixel values. Thus, agreement between the observed spatial configuration and the expected sensory projections is required.

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

This, in turn, requires the following capabilities.

- Infer explicitly or implicitly an object"s position and orientation from the spatial configuration. This requires the capability to infer which pixels are parts of the object. Also, object features such as special markings, curves, lines, etc. have to be distinguished.

- Confirm that the inference is correct. This depends on matching the distinguishing image features with corresponding object features.

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Image recognition steps

Image recognition involves the following six steps:

1. Image formatting
2. Conditioning
3. Labeling
4. Grouping
5. Extracting
6. Matching

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

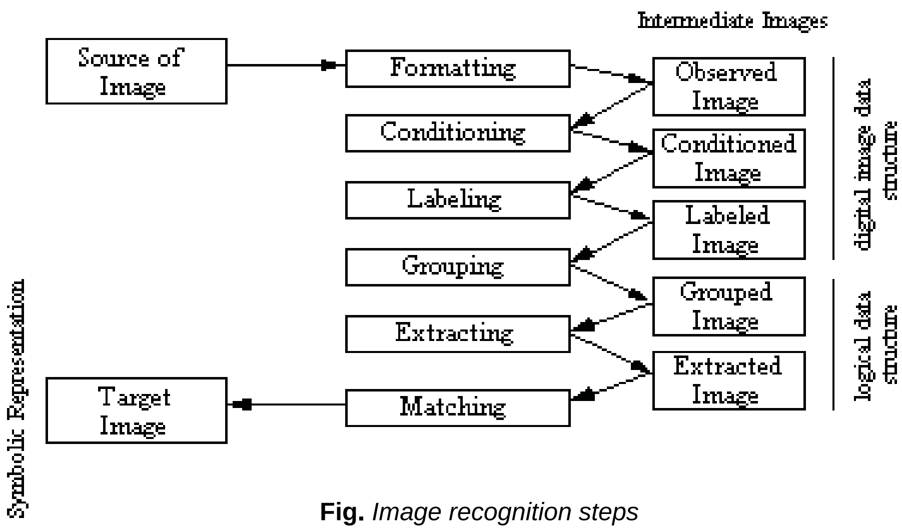


Fig. *Image recognition steps*

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

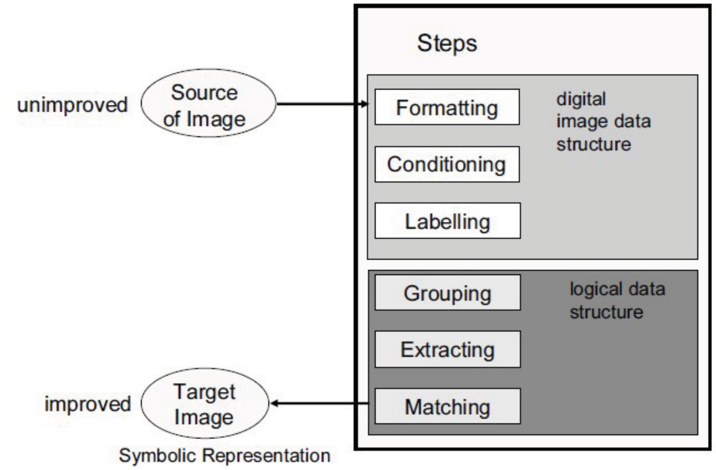


Fig. *Image recognition steps*

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

(1) Image Formatting

It refers to **capturing an image** from a camera and **bringing it into a digital form**, i.e., generating a digital representation of an image in the form of pixels.

(2) Conditioning

The image usually contains some **unwanted variations** or **noise** that makes the recognition process difficult and complex. Conditioning is a process in which the image are **eliminated** or suppressed so that they do not have influence over the recognition process.

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

(3) Labeling

The object of interest usually exhibit some pattern or structure dependent on the spatial position and the connectivity of the individual pixels. **Edge Detection** is an example of labeling in which the boundary of the object of interest is determined. An edge is said to occur at a point in the image if some image attribute changes in value discontinuously at that point. It may result in locating many edges of which some may not be the part of object of interest so they need to be filtered or ignored.

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

(4) Grouping

Grouping involves arranging the pixels or the events (e.g. edge) so that they **produce some meaningful shape and structure**. The grouping operation may involve calculation of neighborhood, connectivity and adjacency of the pixels. The edges may be grouped into lines (line fitting operation), curves may be grouped to form circle or any other structure. It may result in identifying the pixels of interest or creation of data structures of spatial event.

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

(5) Extracting

Extraction is a process in which the data and attributes related to **the images are computed or calculated**. It ends up **creating a list of pixel properties** like centroid, angular orientation, intensity, spatial position etc. The groups are evaluated for regions. For e.g. if the group is an arc, average curvature is a candidate property.

(6) Matching

Matching is the phase in which the image under consideration is related with some pre-defined object, properties, shape or structure. Matching may involve template matching in which the templates are already stored in the database and the image under consideration is compared with it or it may involve complex neural process to identify and classify the image.

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Image Transmission

Image transmission takes into account transmission of digital images through computer networks.

There are several requirements on the networks when images are transmitted

- The network must accommodate bursty transport data, because image transmission is bursty (due to large image sizes).
- Transport should be reliable.
- Time-dependence is not dominant as in audio/video transmission.

Image size depends on the image representation format used for transmission. There are several possibilities:

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Image Transmission …..

Raw image data transmission

The image is generated through as video digitizer and transmitted in its digital format. The image size is calculated as:

Size = Spatial_Resolution X Pixel_Quantization

For example, the transmission of an image with a resolution of 640 X 480 pixels and pixel quantization of 8 bits per pixel requires

640 X 480 X 8 = 2457600 bits (=307,200 bytes).

3.2. Image Processing Fundamentals: Synthesis, Analysis and Transmission.

Image Transmission …..

Compressed image data transmission

The image is generated through a video digitizer and compressed before transmission. Methods such as JPEG or MPEG are used for compression. The reduction of image size depends on the compression method and compression rate.

Symbolic image data transmission

The image is represented through symbolic data representation as image primitives (e.g., 2D or 3D geometric representation), attributes and other control information. This image representation method is used in computer graphics. Image size is equal to the structure size, which carries the transmitted symbolic information of the image.

3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

<u>Image enhancement :</u>

The process of improving the quality of a digitally stored image by manipulating the image with software.

It is quite easy, for example, to make an image lighter or darker, or to increase or decrease contrast. Advanced image enhancement software also supports many filters for altering images in various ways.

Programs specialized for image enhancement are sometimes called *image editors*.

3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

What Is Image Enhancement?

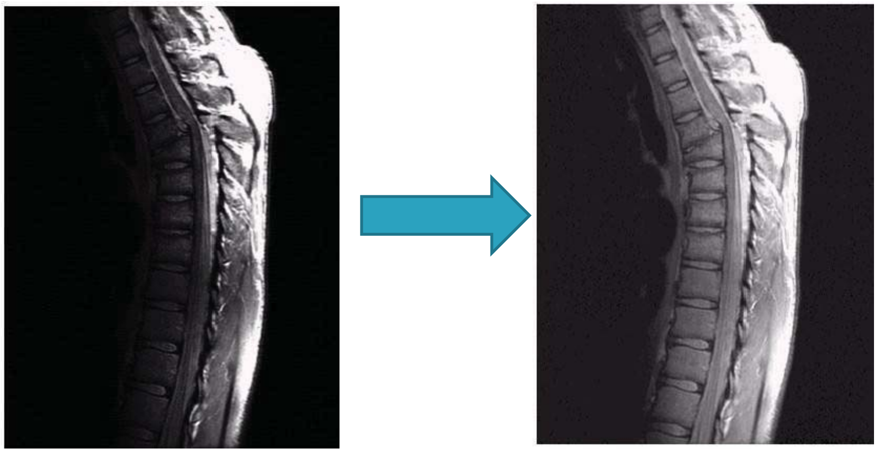
Image enhancement is the process of making images more useful

<u>The reasons for doing this include</u>

- – Highlighting interesting detail in images
- – Removing noise from images
- – Making images more visually appealing

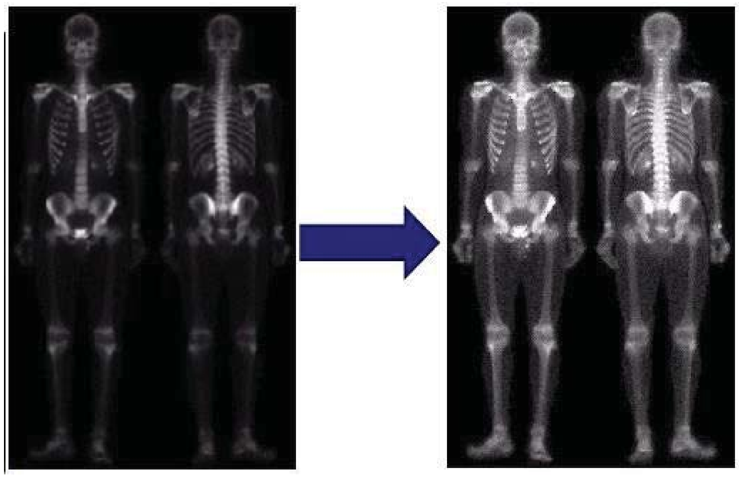
3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Image Enhancement Examples



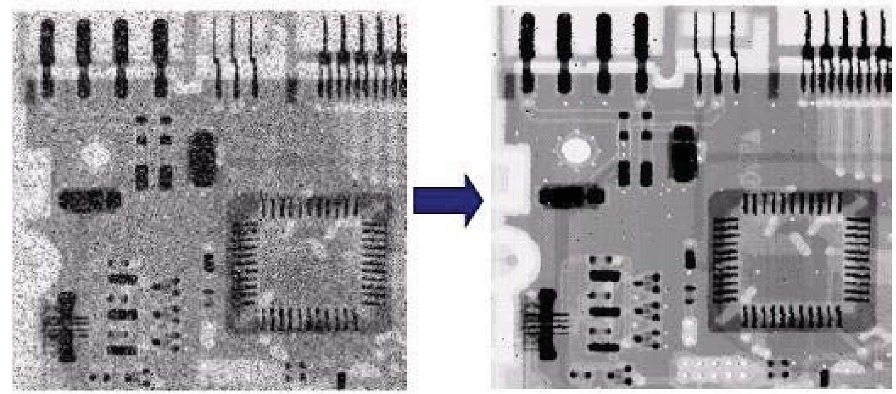
3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Image Enhancement Examples (cont…)



3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Image Enhancement Examples (cont…)



3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Enhancement by point processing

The simplest spatial domain operations occur when the neighborhood is simply the pixel itself

In this case T is referred to as a *grey level transformation function* or a *point processing operation*

Point processing operations take the form

$$s = T(r)$$

where s refers to the processed image pixel value and r refers to the original image pixel value

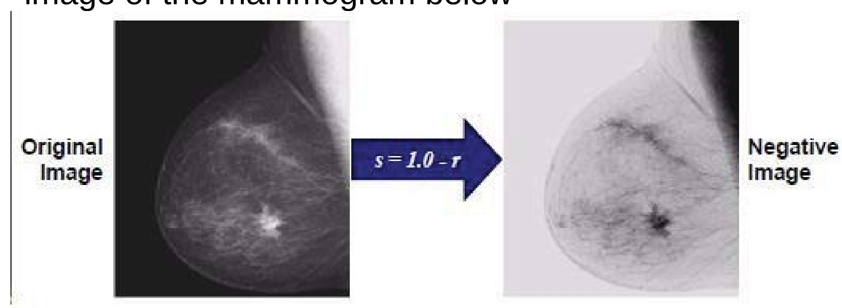
3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Point Processing Example:

Negative Images

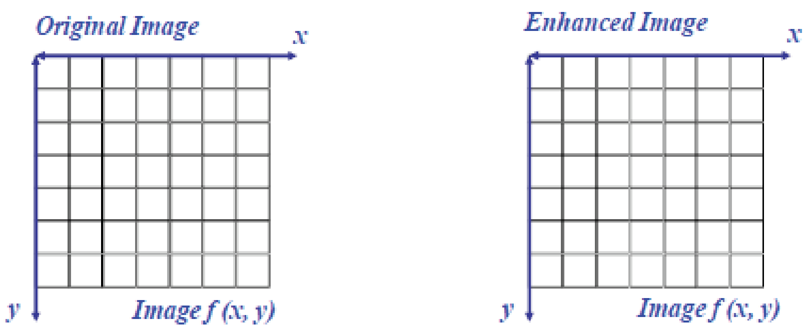
Negative images are useful for enhancing white or grey detail embedded in dark regions of an image

— Note how much clearer the tissue is in the negative image of the mammogram below



3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Enhancement by point processing

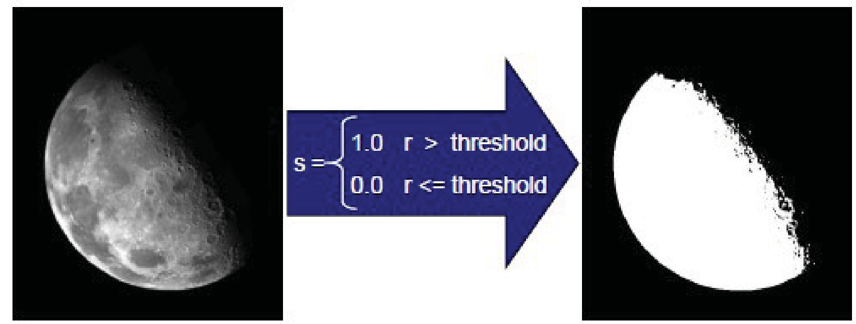


$$s = intensity_{max} - r$$

3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

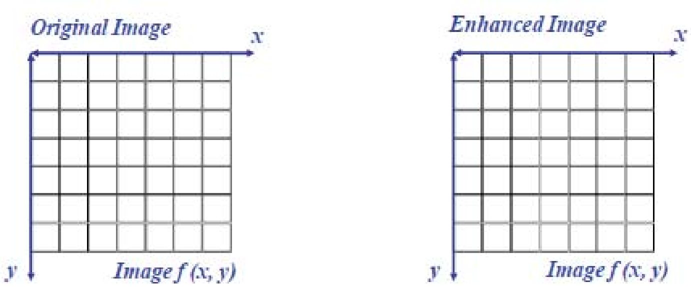
Enhancement by point processing

Thresholding transformations are particularly useful for segmentation in which we want to isolate an object of interest from a background



3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Enhancement by point processing



$$s = \begin{cases} 1.0 & r > threshold \\ 0.0 & r <= threshold \end{cases}$$

3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Spatial Filtering

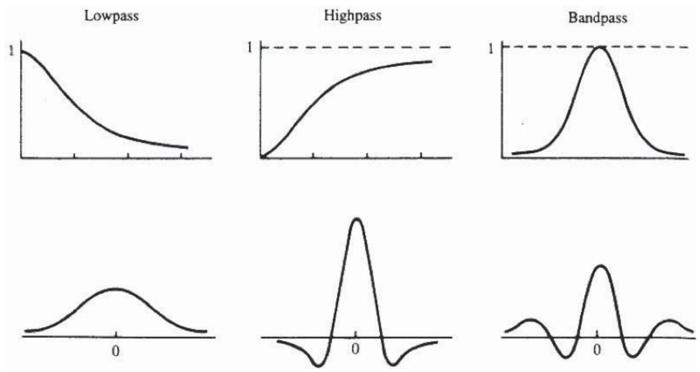
Use of spatial masks for filtering is called spatial filtering

May be linear or nonlinear

- Linear filters
  - Lowpass: attenuate (or eliminate) high frequency components such as characterized by edges and sharp details in an image
- Net effect is image blurring
  - Highpass: attenuate (or eliminate) low frequency components such as slowly varying characteristics
- Net effect is a sharpening of edges and other details
  - Bandpass: attenuate (or eliminate) a given frequency range
- Used primarily for image restoration(are of little interest for image enhancement)

3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

- Filters in the frequency domain and corresponding spatial filters
- Basic approach is to sum products between mask coefficients and pixel values



3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

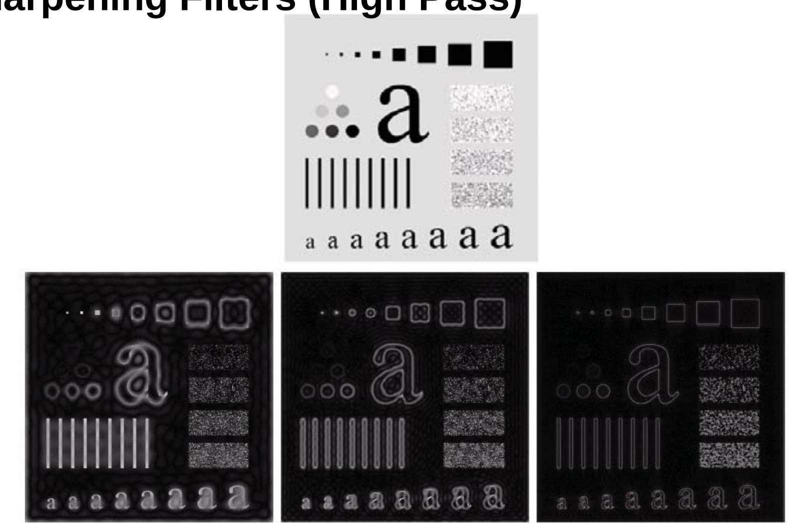
Sharpening Filters (High Pass)

Original

Filtered using ideal HPF

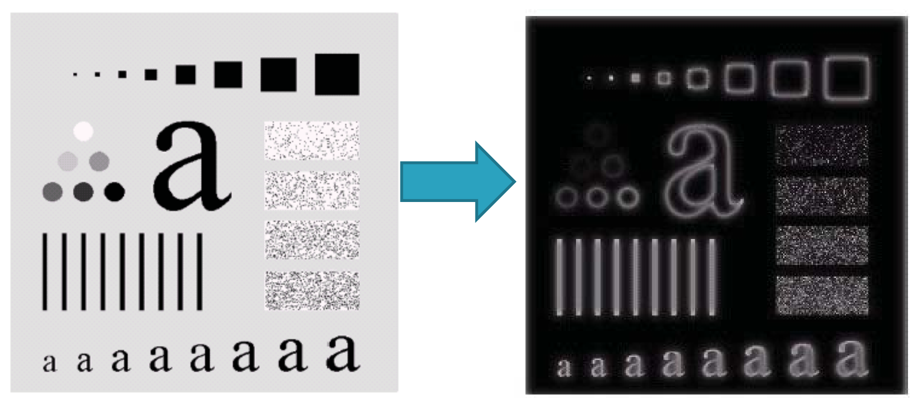
3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Sharpening Filters (High Pass)



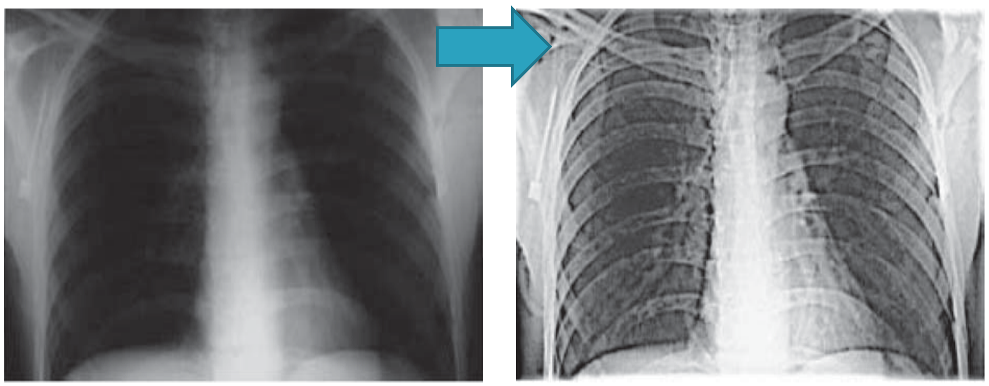
3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Sharpening Filters (High Pass)



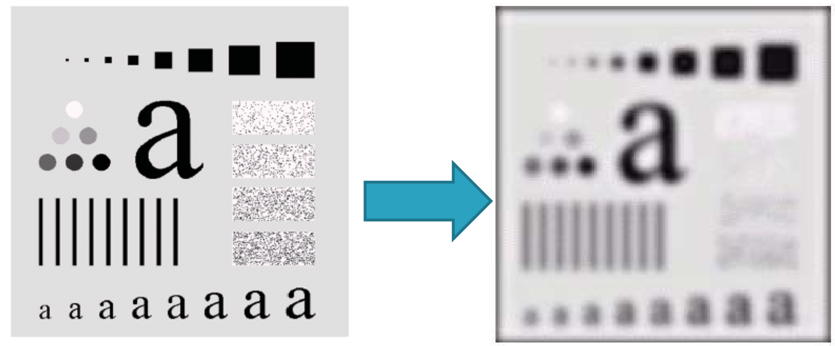
3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Sharpening Filters (High Pass)



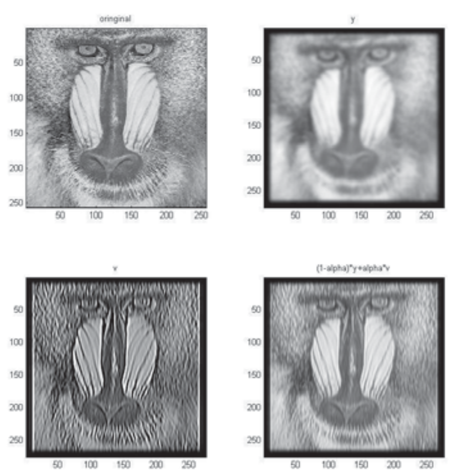
3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Ideal Low pass Filter



3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Band pass Filter



3.3. Image Enhancement : Enhancement by point processing, spatial filtering, Color image processing

Color image processing

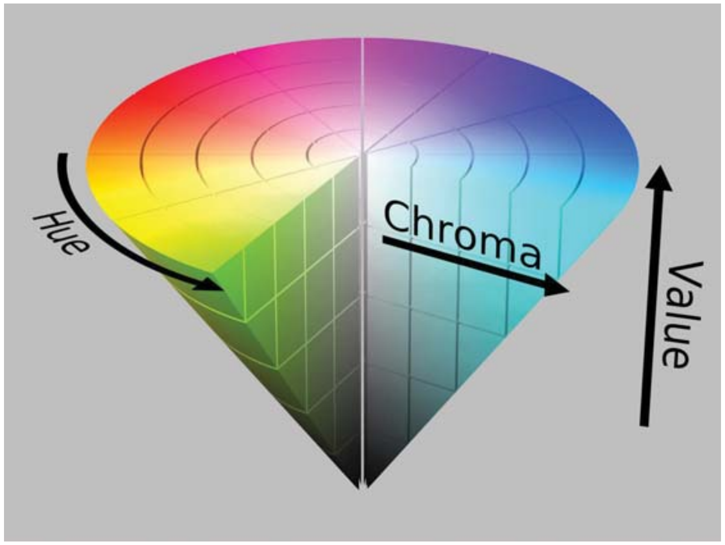
A **(digital) color image** is a digital image that includes color information for each pixel.

The RGB color space is commonly used in computer displays, but other spaces such as HSV, and are often used in other contexts.

A color image has three values (or channels) per pixel and they measure the intensity and chrominance of light.

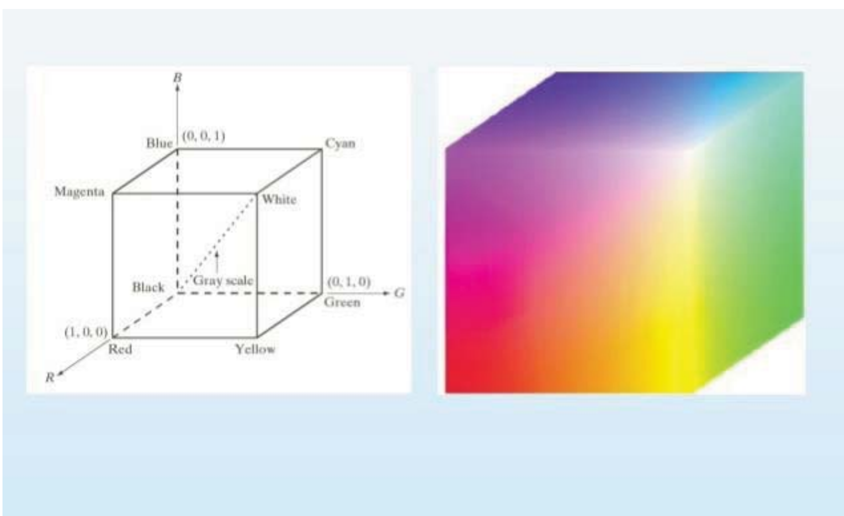
The actual information stored in the digital image data is the brightness information in each spectral band.

Color image processing (HSV)

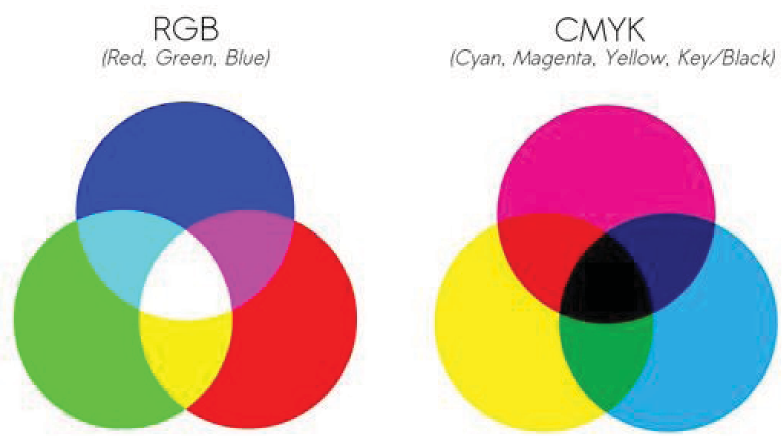


Color image processing (HSV)

Color image processing (RGB)



Color image processing (RGB / CMYK)



VIDEO AND ANIMATION

Chapter 4

4.1. Basic Video Concepts: Representation and Format

Video Signal Representation

In conventional TV sets, the video signal is displayed using a cathode ray tube. An electron beam carries corresponding pattern information such as intensity in a viewed scene.

Video signal representation includes three aspects:

1. the visual representation
2. transmission
3. digitization

4.1. Basic Video Concepts: Representation and Format

1. Visual Representation

The objective is to offer the viewer a sense of presence in the scene and of participation in the events portrayed. To meet the objective, the televised images should convey spatial and temporal content of the scene.

Important measures taken for the purpose include

1. ***Vertical Detail and the Viewing Distance***
2. ***Horizontal Detail and Picture Width***
3. ***Total Detail Content of the Image***
4. ***Perception of Depth***
5. ***Luminance and Chrominance***
6. ***Temporal Aspects of Illumination***
7. ***Continuity of Motion***
8. ***Flicker***
9. ***Temporal Aspect of Video Bandwidth***

4.1. Basic Video Concepts: Representation and Format

1)Visual Representation

1. Vertical Detail and the Viewing Distance

Geometry of the field occupied by the television image is based on the ratio of picture width and height, called as aspect ratio (W/L=4/3=1.33)

Viewing distance **D determines the angle H subtended by the picture height.** The angle is usually measured by the ration of the viewing distance to the picture height (D/H). The smallest detail that can be reproduced in the image is pixel

Both the number of pixels per scanned line and the number of lines per frame vary; the actual numbers used being determined by the *aspect ratio.* This is the ratio of the screen width to the screen height. The aspect ratio of current television tubes is 4/3 with older tubes – on which PC monitors are based – and 16/9 with widescreen television tubes.

4.1. Basic Video Concepts: Representation and Format

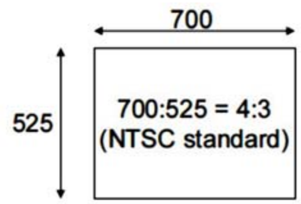
1)Visual Representation

1. Vertical Detail and the Viewing Distance (cont..)

Geometry of the television image is based on the *aspect ratio*, which is the ratio of the picture width W to the height H (W:H).

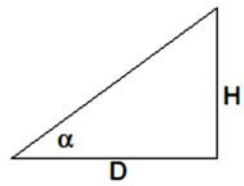
The conventional aspect ratio is 4/3 = 1.33 (W/H = 4/3)

Modern systems use 16/9 = 1.77



The *viewing distance* D determines the angle α subtended by the picture height H.

This angle is measured by the ratio of the picture height to viewing distance: $\tan(\alpha) = H/D$.



4.1. Basic Video Concepts: Representation and Format

1)Visual Representation(cont..)

2.*Horizontal Detail and Picture Width*

Picture width chosen for the conventional television services is **4/3 * picture height** Using the aspect ratio, we can determine the horizontal field of view from the horizontal angle.

The horizontal detail or the picture width is again dependent upon the aspect ratio and is given by **aspect ratio*height of picture.**

4.1. Basic Video Concepts: Representation and Format

1)Visual Representation(cont..)

3. Total Detail Content of the Image

The **vertical resolution** is equal to the number of picture elements separately presented in the **picture height**, while the number of elements in the **picture width** is equal to the **horizontal resolution** times the aspect ratio.

The product of horizontal and vertical elements gives the **total picture elements in the image**.

The total detail content of the image is given by the number of pixels used to represent the image and is given by the resolution.

Thus total detail content of the image is the product of **Vertical Detail * Horizontal Detail.**

4.1. Basic Video Concepts: Representation and Format

1)Visual Representation(cont..)

4.Perception of Depth

In natural vision, perception of the third **spatial(3D/4D) dimension, depth,** depends primarily on the angular separation of the images received by the **two eyes if the viewer.**

In the flat image television, a considerable degree of depth perception is inferred from the perspective appearance of the subject matter.

As the screen is two dimensional a special measure is taken to give the sense of depth. This involves the perspective representation of the image content. This is again governed by the focal length of the lenses and changes in depth of focus in camera.

4.1. Basic Video Concepts: Representation and Format

1)Visual Representation(cont..)

5.Luminance and Chrominance

Color vision is achieved through three signals, proportional to the relative **intensities of Red, green and blue lights (RGB) in each portion** of the scene.

During the transmission of the signals from the camera to the television (display), a different division of signals in comparison to the RGB division is often used.

This color encoding during transmission uses luminance and chrominance.

The term **luminance** is used to refer to the **brightness** of a source, and the **color information** is referred by the **chrominance**.

Human eye however is more sensitive to luminance than to chrominance. The color encoding during transmission uses luminance and two chrominance signals (for hue and saturation).

4.1. Basic Video Concepts: Representation and Format

1)Visual Representation(cont..)

5.Luminance and Chrominance(cont..)

Thus luminance (Y_s) and two chrominance blue chrominance(C_b), and the red chrominance(C_r) are then used to represent the video content where

$$Y_s = 0.299\,R_s + 0.587\,G_s + 0.144\,B_s$$

$$C_b = B_s - Y_s \text{ and}$$

$$C_r = R_s - Y_s$$

4.1. Basic Video Concepts: Representation and Format

1)Visual Representation(cont..)

6.Temporal Aspects of Illumination

In contrast to continuous pressure waves of an acoustic signal, a discrete sequence of individual pictures can be perceived as a continuous sequence.

This property is used in television and **motion pictures**. The rate of repetition of the images must be high enough to **guarantee smooth motion** from **frame to frame** and the persistence of vision extends over the interval between flashes.

One of the major characteristics of human eye is that if the still images are shown in rapid succession, **human eye perceive the motion** i.e. it does not notice the brief cut off of the light.

Thus if the rate of succession is sufficiently high still images can be used to represent motion as the persistence of vision extends over the interval between flashes.

4.1. Basic Video Concepts: Representation and Format

1)Visual Representation(cont..)

7.Continuity of Motion

NTSC (**N**ational **T**elevision **S**ystems **C**ommittee) specified the frame rate for maintaining the visual-aural carrier separation at **4.5 MHz was 30 frames/sec** but has been changed to 29.97 HZ.

PAL (**P**hase **Al**ternating **L**ine) adopted the frame repetition rate of **25 Hz**, and the frame rate is **25frames/sec**. Digital motion picture uses frame rate of **50 frames/sec**.

As stated above, motion can be represented by showing the still images in rapid succession. This can be done by showing the still images 15 frames per second. Video motion seems smooth and is achieved at only 30 frames per second.

4.1. Basic Video Concepts: Representation and Format

1)Visual Representation(cont..)

8.Flicker

Through a slow motion, a periodic fluctuation of brightness perception, a flicker effect arises. The marginal value to avoid flicker is at least 50 refresh cycles per second.

To achieve continuous flicker-free motion, we need relatively high refresh frequency.

Movies and television apply some technical measures to work with lower motion frequencies. When the **refresh rate of frames is smaller than the eye will notice the cutoff of light between the frames and there arises the flickering rate**. The marginal value to avoid flicker is at least 50 refresh cycles per second.

4.1. Basic Video Concepts: Representation and Format

1)Visual Representation(cont..)

9.Temporal Aspect of Video Bandwidth

The most important factor to determine at which bandwidth the video can be transmitted is its temporal specification. Temporal specification depends on the rate of the visual system to scan pixels, as well as on the human eye"s scanning capabilities.

In a HDTV (High Definition TV) device, a pixel can be scanned in less than a tenth of a millionth of a second. From the human visual perspective, the eye requires that a video frame to be scanned every 1/25 second, this time is equivalent to the time during which human eye does not see the flicker effect.

The choice of bandwidth for transmission of the video signal depends on the rate of the visual system to scan pixels, as well as on the human eye"s scanning capabilities. From the human visual perspective, the eye requires that a video frame be scanned every 1/25 second.

4.1. Basic Video Concepts: Representation and Format

2) Transmission

Standards for transmission of video are: NTSC (***N**ational **T**elevision **S**ystem **C**ommittee) and* PAL *(**P**hase **A**lternating **L**ine).* Approaches of color encoding are: RGB, YUV (luminance Y, two chrominance channels U and V), YIQ, Composite, etc.

Different color encoding techniques can be employed while transmitting the video signal and they are:

- *i.* ***RGB signal***
- *ii.* ***YUV signal***
- *iii.* ***YIQ signal***
- *iv.* ***Composite signal***

4.1. Basic Video Concepts: Representation and Format

2) Transmission (cont..)

i. RGB signal

In this case separate signal coding is done for the individual R, G and B components of the image and these components are transmitted separately.

ii.YUV signal

This model exploits the fact that the human eye is more sensitive to brightness i.e. luminance than to color information i.e. chrominance. Thus instead of separating the color component the brightness and the coloration of the image is separated.

Thus luminance(Y_s) and two chrominance blue chrominance (C_b), and the red chrominance (C_r) are then used to represent the video content where

$Y_s = 0.30\, R_s + 0.59\, G_s + 0.11\, B_s$

$U = (B_s - Y_s) \times 0.493$

$V = (R_s - Y_s) \times 0.877$

4.1. Basic Video Concepts: Representation and Format

2) Transmission (cont..)

iii.YIQ signal

This model is very much similar to YUV model and is used in NTSC system. The image information is broken down into three components a luminance component and two chrominance component given by following relation.

$$Y = 0.30\, R_s + 0.59\, G_s + 0.11\, B_s$$

$$I = 0.60\, R_s - 0.28\, G_s + 0.32\, B_s$$

$$Q = 0.21\, R_s + 0.52\, G_s + 0.31\, B_s$$

4.1. Basic Video Concepts: Representation and Format

2) Transmission(cont..)

iv. Composite signal

The alternative to component coding composes all information into one signal; consequently, the individual components (RGB, YUV, and YIQ) must be combined into one signal.

The basic information consists of luminance information and chrominance difference signals.

4.1. Basic Video Concepts: Representation and Format

3) Digitalization

Digitalization is a **process of changing the continuously varying signal into discrete components.**

This basically uses mathematical process like ***Fourier analysis*** or a series of step consisting of sampling and quantizing.

Sampling is the process that actually digitizes the spatial position of the image while quantizing digitizes its color information.

Sampling involves dividing the picture at **M*N array** of points while quantizing involves dividing the signal into a range of gray level values.

Finally the digital motion video is created by digitizing the pictures temporally.

4.1. Basic Video Concepts: Representation and Format

Computer Video Format

Computer video format depends on the input and output devices for the motion video medium

- Current video digitizers differ in digital image resolution, quantization and frame rate
- Most often used display is raster display
- The raster display architecture (as shown below)

4.1. Basic Video Concepts: Representation and Format

Some computer video controller standards **(Computer Video Formats)** are given below:

Color Graphics Adapter (CGA)

Resolution: 320*200
Color depth: 2 bits/pixel
Image size: 16 KB
No. of colors: 4
Storage Capacity: 320*200*4/8=16,000 bytes

Enhanced Graphics Adapter (EGA)

Resolution: 640*350
Color depth: 4 bits/pixel
Image size: 112 KB
No. of colors: 16
Storage Capacity: 640*480*4/8=112,000 bytes

4.1. Basic Video Concepts: Representation and Format

Video Graphics Array (VGA)

Resolution: 649*480
Color depth: 8 bits/pixel
Image size: 307.2 KB
No. of colors: 256
Storage Capacity: 640*480*8/8=307,200 bytes

8514/A Display Adapter Mode

Resolution: 1024*768
No. of colors: 256
Storage Capacity: 1024*768*8/8=786,432 bytes

4.1. Basic Video Concepts: Representation and Format

Extended Graphics Array (XGA)

Resolution: 640*480 / 1024*768
Color depth: 65,000 colors / 256
colors No. of colors: 256
Storage Capacity: 1024*768*8/8=786,432 bytes

Super VGA (SVGA)

Resolution: 1024*768
Color depth: 24 bits/pixel
Image size: 2.35 MB
No. of colors: 2^{24}
Storage Capacity: 1024*768*24/8=2,359,296 bytes

4.2. Television

Television

Conventional Systems

Conventional television systems employ the following standards:

NTSC
SECAM
PAL

4.2. Television

NTSC

- National Television Systems Committee
- Developed in the U.S.
- Oldest and most widely used television standard
- Color carrier is used with approximately 4.429 MHz or with approximately 3.57 MHz.
- Uses a quadrature amplitude modulation with a suppressed color carrier
- Works with a motion frequency of approximately 30 Hz
- A picture consists of 525 lines
- 4.2 MHz is used for the luminance and 1.5 MHz is used for each of the two chrominance channels.

4.2. Television

NTSC

N T S C **N**ational **T**elevision **S**ystem **C**ommittee	
Lines/Field	525/60
Horizontal Frequency	15.734 kHz
Vertical Frequency	60 Hz
Color Subcarrier Frequency	3.579545 MHz
Video Bandwidth	4.2 MHz
Sound Carrier	4.5 MHz

4.2. Television

SECAM

- **SE**quential **C**ouleur **A**vec **M**emoire
- Used in France and Eastern Europe
- Unlike NTSC and PAL, it is based on frequency modulation
- Uses a motion frequency of 25 Hz
- Each picture has 625 lines

SECAM **S**equential **C**ouleur **A**vec **M**emoire *or* **S**equential **C**olor with **M**emory		
SYSTEM	**SECAM B,G,H**	**SECAM D,K,K1,L**
Line/Field	625/50	625/50
Horizontal Frequency	15.625 kHz	15.625 kHz
Vertical Frequency	50 Hz	50 Hz
Video Bandwidth	5.0 MHz	6.0 MHz
Sound Carrier	5.5 MHz	6.5 MHz

4.2. Television

PAL

- Phase Alternating Line
- Invented in 1963 by W. Bruch
- Used in parts of Western Europe
- Uses a quadrature amplitude modulation similar to NTSC, but the color carrier is not suppressed

P A L **P**hase **A**lternating **L**ine			
SYSTEM	**PAL**	**PAL N**	**PAL M**
Line/Field	625/50	625/50	525/60
Horizontal Freq.	15.625 kHz	15.625 kHz	15.750 kHz
Vertical Freq.	50 Hz	50 Hz	60 Hz
Color Sub Carrier	4.433618 MHz	3.582056 MHz	3.575611 MHz
Video Bandwidth	5.0 MHz	4.2 MHz	4.2 MHz
Sound Carrier	5.5 MHz	4.5 MHz	4.5 MHz

4.2. Television

High Definition Television

The formats used in HDTV are:

- **720p** - 1280x720 pixels progressive
- **1080i** - 1920x1080 pixels interlaced
- **1080p** - 1920x1080 pixels progressive

HDTV is high-resolution digital television (DTV) combined with Dolby Digital surround sound (AC- 3).
HDTV is the highest DTV resolution in the new set of standards. This combination creates a stunning image with stunning sound. **HDTV** requires new production and transmission equipment at the **HDTV stations**, as well as new equipment for reception by the consumer. The higher resolution picture is the main selling point for HDTV.

4.2. Television

High Definition Television

High-Definition Television (HDTV) is defined by the image it presents to its viewer, which has the following characteristics:

Resolution

The HDTV image has approximately twice as many horizontal and vertical pixels as conventional systems. Luminance detail is also increased by employing a video bandwidth approximately five times that used in conventional systems.

Aspect Ratio

The aspect ratio of the HDTV images is 16/9 = 1.777.

Viewing Distance

Since the eye‟s ability to distinguish details is limited, the more detailed HDTV image should be viewed closer than conventional systems.

INTRODUCTION TO VIRTUAL REALITY(VR) AND ANIMATION

Virtual Reality(VR)

Virtual Reality(VR)

- **Virtual reality** or **virtual realities** (**VR**), also known as **immersive multimedia** or **computer-simulated reality**, is a computer technology that replicates an environment, real or imagined, and simulates a user's physical presence and environment to allow for user interaction. Virtual realities artificially create sensory experience, which can include sight, touch, hearing, and smell.
- Most up-to-date virtual realities are displayed either on a computer monitor or with a virtual reality headset (also called **head mounted display**)
- Some simulations include additional sensory information and focus on real sound through speakers or headphones targeted towards VR users

Why Virtual Reality(VR) ?

- VR is able to immerse you in a computer-generated world of your own making: a room, a city, the interior of human body. With VR, you can explore any uncharted territory of the human imagination.

Types of VR(Virtual Reality)

1. Windows on World(WoW) (also called Desktop VR)
2. Immersive VR
3. Telepresence
4. Mixed Reality(Augmented Reality)
5. Distributed VR

Types of VR(Virtual Reality)..

Windows on World(WoW)

- Also called Desktop VR.
- Using a conventional computer monitor to display the 3D virtual world.

Types of VR(Virtual Reality)..

Immersive VR

- Completely immerse the user's personal viewpoint inside the virtual 3D world.
- The user has no visual contact with the physical word.
- Often equipped with a Head Mounted Display (HMD).

Types of VR(Virtual Reality)..

Telepresence

- A variation of visualizing complete computer generated worlds.
- Links remote sensors in the real world with the senses of a human operator. The remote sensors might be located on a robot. Useful for performing operations in dangerous environments.

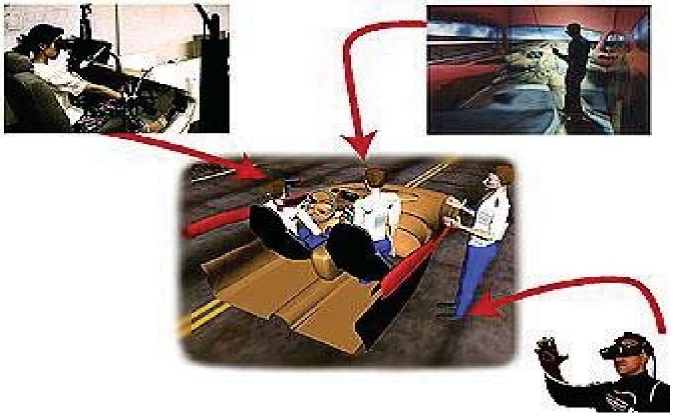
Types of VR(Virtual Reality)…

- Mixed Reality(Augmented Reality)
  - The seamless merging of real space and virtual space.
  - Integrate the computer-generated virtual objects into the physical world which become in a sense an equal part of our natural environment.

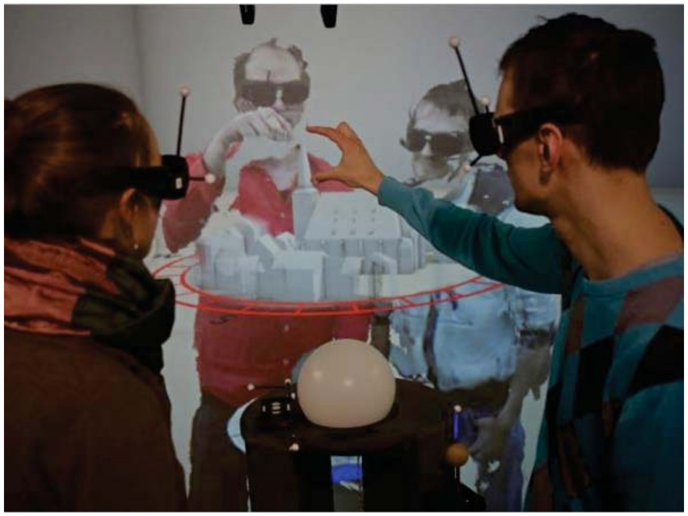
Augmented VR

Types of VR(Virtual Reality)…

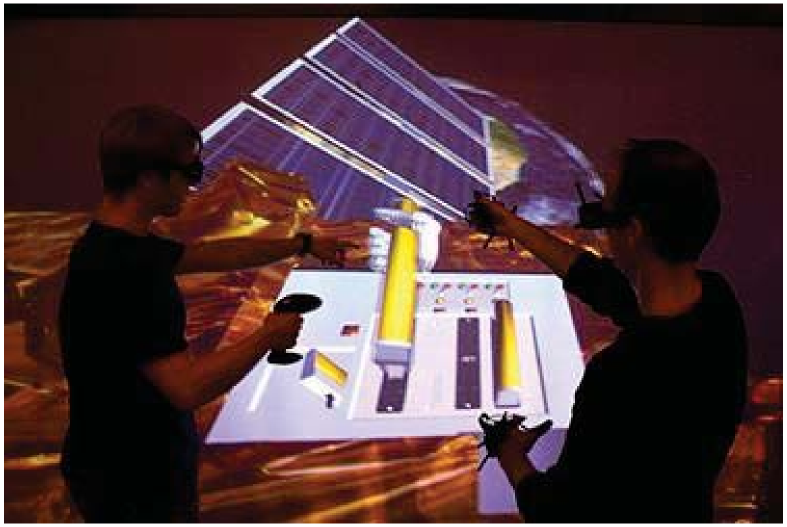
- Distributed VR
  - A simulated world runs on several computers which are connected over network and the people are able to interact in real time, sharing the same virtual world.



Distributed VR



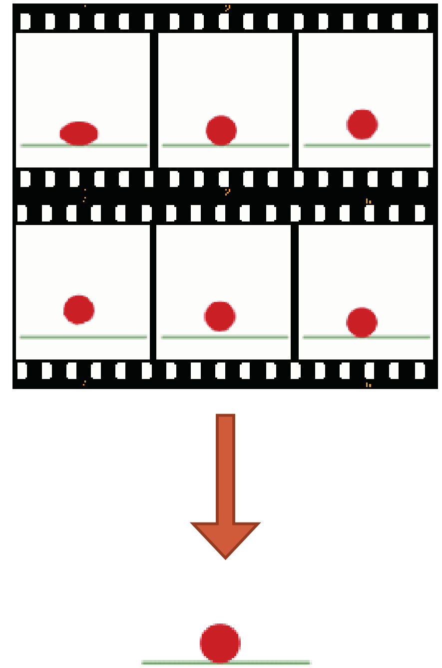
Distributed VR



4.3.ANIMATION

Introduction

- **Animation** is the process of making the illusion of motion and change by means of the rapid display of a sequence of static images that minimally differ from each other. The illusion—as in motion pictures in general—is thought to rely on the phenomenon. Animators are artists who specialize in the creation of animation.
- Animation is the creation of the —illusion of movement‖ using a series of still images.
- A collection of static image joined together and shown consecutively so that appear to move.



Introduction ….

- Animation creation methods include the traditional animation creation method and those involving stop motion animation of two and three-dimensional objects, paper cutouts, puppets and clay figures. Images are displayed in a rapid succession, usually 24, 25, 30, or 60 frames per second.

Introduction ….

to animate = "to bring to life"

Animation **covers changes in:**

time-varying positions (motion dynamics),

shape, color, transparency, structure and texture of an object (update dynamics) as well as lightning, camera position, camera orientation and focus.

Basic Concepts **of animation are:**

Input Process

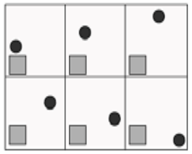
- Key frames, where animated objects are at extreme or characteristic positions must be digitized from drawings.
- Often a post-processing by a computer is required.

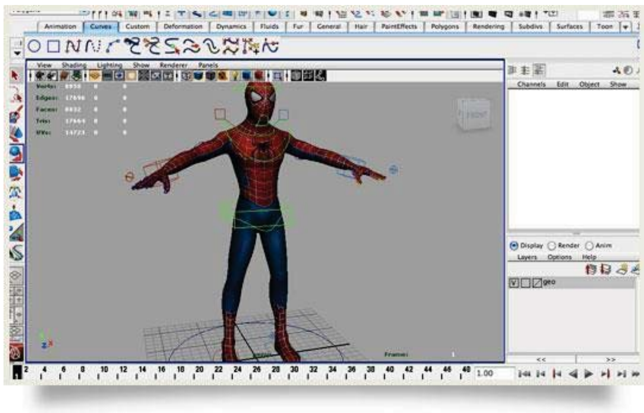
Composition Stage

Foreground and background figures are combined to generate an individual frame.

Placing of several low-resolution frames of an animation in an array leads to a *trail film* (pencil test), by the use of the *pan-zoom* feature (This feature is available for some frame buffers).

The frame buffer can take a part of an image (pan) and enlarge it to full screen (zoom).
Continuity is achieved by repeating the pan-zoom process fast enough.



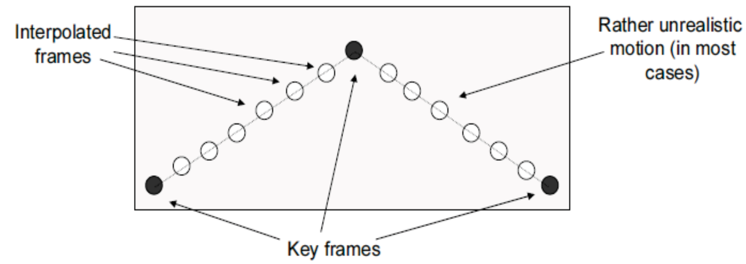


In-between Process

Composition of intermediate frames between key frames.

Performed by linear interpolation (lerping) between start- and end-positions.

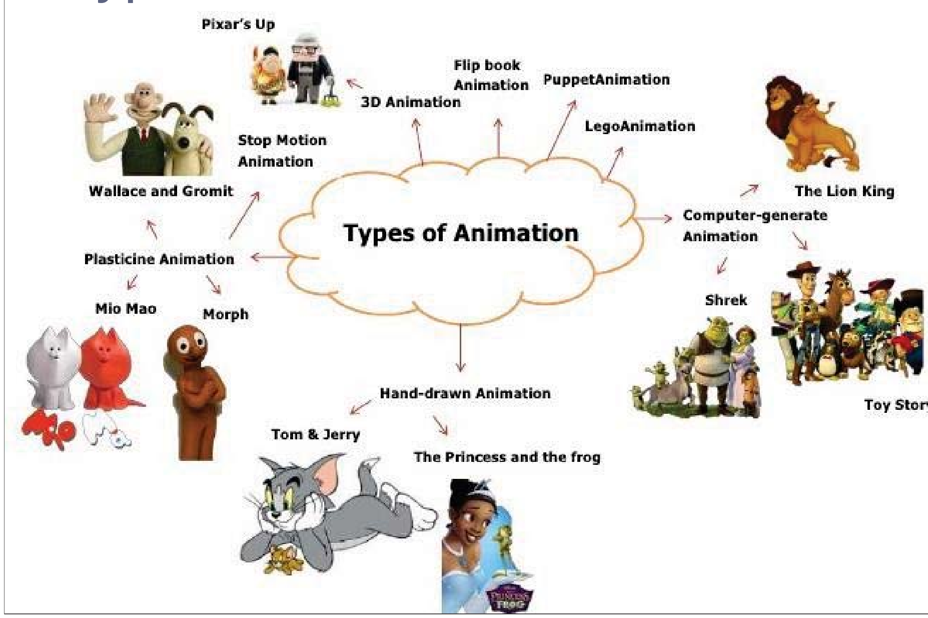
To achieve more realistic results, cubic spline-interpolation can be used.



Types of Animation

- Hand Drawn Animation
- Stop Motion Animation
- Computer animation

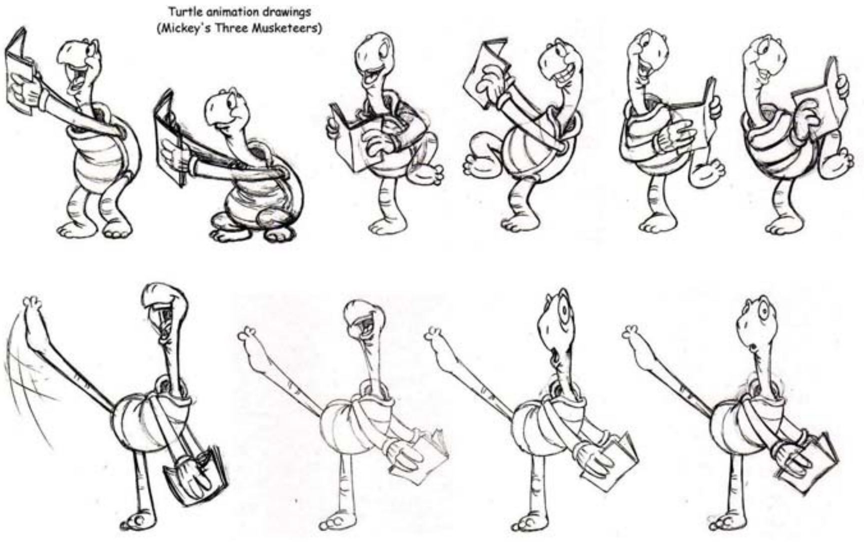
Types



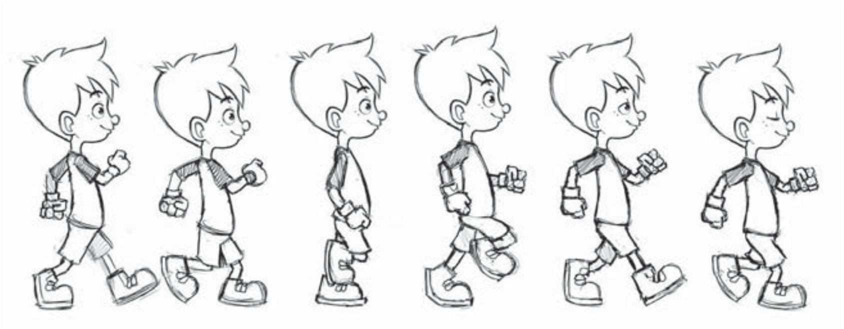
Hand Drawn Animation

- Done by an artist who draws each character and movement individually
- Very time consuming to have to draw, then color, then photograph each picture
- Draw pictures first, then color them on celluloid, then they pictures and animate them
- Very expensive due to hours of labor involved
- Examples: older Disney Movies i.e. Bambi, Fox and Hound, Cinderella etc.

Hand Drawn Animation



Hand Drawn Animation



Stop Motion Animation

- Can be done by virtually anyone, with no extensive training
- Does not take that much time relative to the other 2 methods
- Uses jointed figures or clay figures that can be moved to make motions
- Take still pictures of the individual movements, then use relatively inexpensive computer software to animate
- We use Movie Maker Software to complete our animations
- Not very expensive because all you need is a digital camera and the software comes with windows XP operating system
- Examples: star wars, robot chicken, old Rudolph the red Nosed Reindeer.

Stop Motion Animation

Stop Motion Animation

Computer Animation

- All characters and movements are generated using computer animation software
- Can also be very time consuming as they can get very complicated in movements and effects
- All characters are fully animated with no still pictures
- Can be very expensive because of the complexity of the stunts and animation being done
- Huge budgets because the animation sequences more complicated these days eg. The war scenes in Lord of the Rings etc.
- Examples: Toy story , finding Nemo, Matrix, Lord of the Rings

Basic Step Of Animation

1. Shooting reference video
2. Key posing
3. Blocking
4. Splining
5. Smoothing
6. Adding life

1. Shooting Reference Video

- This is a very important and **overlooked** step. It‘s weird how people really think they know what certain actions look like and how long they take, but in reality they are often wrong.
- Physical actions is something you **need to analyze** before animating, especially if you‘re a beginner.
- You have a shot of a guy throwing a baseball? Better YouTube some reference video of pitchers throwing balls. Don‘t assume you know what an action looks like just because you‘ve seen it before. Looking at an action as an **animator** is completely different than looking at it as a regular viewer.

2. Posing

- After shooting a reference, it‘s time to create the key poses of the shot.
- These poses are called key poses because they are the most important poses of the shot. These are the poses that convey the story of the shot. We better make sure we **get those poses right**, because we‘re going to build on those for the rest of the process.

3. Blocking

- Once we‘re happy with our key poses, we start **breaking down** the movement from each pose to the next by adding ‗in between‘ (also known as breakdown poses or passing poses). These are the poses that connect the key poses.
- We keep adding more poses until the movement looks as good as it could, while still staying in **stepped mode** (stepped mode is when you don‘t allow interpolation between poses, which results in a very choppy/blocky motion).

4. Splining

- Splining is a 3D animation term. It‘s the process in which you convert the interpolation of the keys from stepped to spline. In other words – you make the computer connect the movement between each of your poses, and that makes the movement **look smoother.**
- The problem is that the computer doesn‘t do a very good job at interpolating. It only works with what it has. That‘s why the better the blocking is – the better the splined version is going to look.

5. Smoothing and offset

- Now that all of our keys are on spline mode, we have to work on them. We need to clean up all the curves and make sure the movement looks smooth.
- It's also a good idea to **offset** some of the actions so it doesn't look so **'stop and start'**, as if the character is doing all the motion at once. By the end of this step your shot should look pretty solid and almost finished.

6. Adding life

- This step is the a lot of fun. We've already finished with the grunt work of animating and it's time to add the fun stuff. In this step we add **small imperfections** that bring life to the character. Maybe an extra blink or a mouth twitch here and there. The difference between the last 2 steps is small but very noticeable.

Transmission of Animation:

One of the following two approaches may be used for the transmission of animation over computer networks:

- ***Symbolic Representation***
- ***Pixmap Representation***

Symbolic Representation

The symbolic representation (e.g. circle) of animation objects (e.g. ball) is transmitted together with the operation commands (e.g. roll the ball) performed on the object, and at the receiver side the animation is displayed.

The transmission time is short because the symbolic representation of an animated object is smaller in byte size than its pixmap representation. However, the display time at the receiver takes longer because the scan-converting operation has to be performed at the receiver side.

The transmission rate (bits/sec or bytes/sec) of animated objects depends on:

- The size of the symbolic representation structure, where the animated object is encoded,
- The size of the structure, where the operation command is encoded, and
- The number of animated objects and operation commands sent per second.

Pixmap Representation

The pixmap representation of the animated objects is transmitted and displayed on the receiver side.

The transmission time is longer in comparison to the previous because of the size of the pixmap representation. However, the display time is shorter because the scan-conversion of the animated objects is avoided at the receiver side.

It is performed at the sender side where animation objects and operation commands are generated. The transmission rate of the animation is equal to the size of the pixmap representation of an animated object (graphical image) multiplied by the number of graphical images per second.

Animation Uses

- Animated Movies: Million dollar industry;
- 1-20 millions spent on each movie
- Animation and computer graphics effects are used in movies frequently
- Video games
- TV programs(e.g. Weather, News)
- Used online (images, ads, chatting)
- Simulations (science and Engineering)
- Virtual Reality (e.g. second life)

Morphing

It is derived from the word metamorphosis, which means the transformation shape, appearance of one thing into another. The transformation of object shapes from one form to another is called morphing.

Also it can be defined as:

- Transition from one object to another.
- Process of transforming one image into another.

Morphing

DATA COMPRESSION

Unit 5

5.1 Data Compression and Coding Fundamentals

Why we need image compression?

- One 90 minutes color movie, each second plays 24 frames. When we digitize it, each frame has 512×512 pixels, each pixel has three components R, G, B each one occupies 8 bits respectively, the total byte number is:

$$90 \times 60 \times 24 \times 3 \times 512 \times 512 = 97{,}200\text{MB}$$

- A CD may save 600 megabytes data, the movie needs 160 CDs to save.

5.1 Data Compression and Coding Fundamentals

Objective: To reduce the amount of data required to represent an image.

Important in data storage and transmission

- Progressive transmission of images (internet, www)
- Video coding (HDTV, Teleconferencing)
- Digital libraries and image databases
  - Medical imaging
  - Satellite images

5.1 Data Compression and Coding Fundamentals

Storage Space and Coding Requirements

Storage Requirement

Uncompressed graphics, audio and video data require considerable storage capacity.

For e.g. an image represented in 640 by 480 resolutions with 8 bits per pixel require 300 kilobytes of storage space.

So, in order to reduce the storage requirement of the multimedia data it should be compressed.

The compression should be such that there should not be a loss of quality of the data and hence compression should not reduce the information content of the data. For e.g. 90% of the raw audio can be deleted without affecting the quality of the audio.

5.1 Data Compression and Coding Fundamentals

Bandwidth Requirement

Uncompressed data transfer requires greater bandwidth or data rate for its communication.

If the data is compressed there can be a considerable reduction in the bandwidth requirement for the transmission of the data.

5.1 Data Compression and Coding Fundamentals

Bandwidth Requirement (cont..)

The following examples specify continuous media and derive the amount of storage required for one second of playback:

- An uncompressed audio signal of telephone quality sampled at 8 kHz and quantized with 8 bits per sample requires 64 kbits to store one second of playback and a bandwidth of 64 kbits/sec.
- An uncompressed stereo audio signal of CD quality is sampled at a rate of 44.1 kHz quantized at 3 16 bits per sample require 705.6 * 10 3 bits to store one second of playback and the bandwidth requirement is 705.6 * 10 bits/second.

5.1 Data Compression and Coding Fundamentals

Bandwidth Requirement (cont..)

As mentioned above compression in multimedia systems is subject to certain constraints.

- The **quality** of the compressed data should be as good as possible.
- To make **cost-effective** implementation possible, the complexity of the technique used should be minimal.
- The processing of the **algorithm** must not exceed certain **time span**.

5.1 Data Compression and Coding Fundamentals

Bandwidth Requirement (cont..)

In retrieval mode application, the following demands arise:

- ***Fast forward*** *and* ***backward data*** *retrieval* should be possible with simultaneous display.
- ***Random access*** to single images and audio frames of a data stream should be possible without extending the access time more than 0.5 second.
- **Decompression** of data should be independent of other data units.

5.1 Data Compression and Coding Fundamentals

Bandwidth Requirement (cont..)

For both dialogue and retrieval mode, the following requirements apply

- The format should be **independent** of frame size and video frame rate.
- The format should support **various** data rate.
- There should be **synchronization** between the ***audio*** and ***video***.
- Compression and decompression should **not require additional hardware.**
- The compression of data in **one system of multimedia** should ensure the **decompression** in the **other system**.

5.1 Data Compression and Coding Fundamentals

Source, Entropy and Hybrid Coding

Entropy encoding

- Data stream is considered to be a simple digital sequence without semantics
- Lossless coding, decompression process regenerates the data completely
- Used regardless of the media´s specific characteristics
- Examples: **Run-length encoding**, **Huffman encoding**, **Arithmetic encoding**

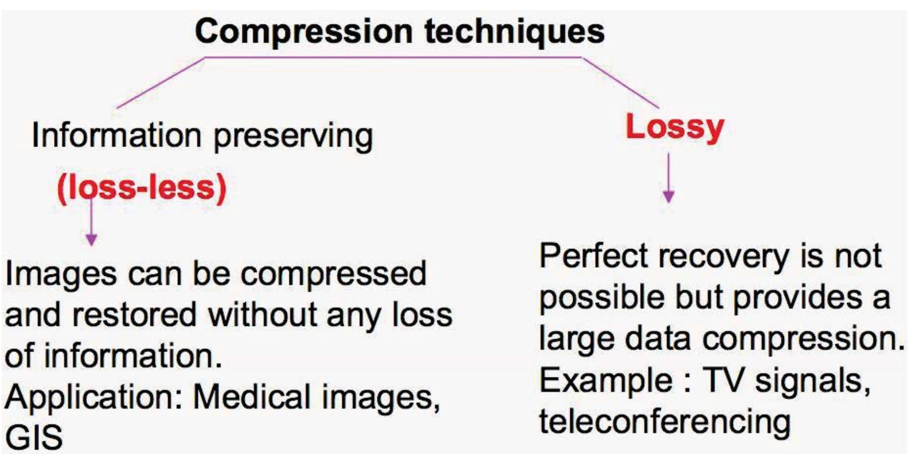
Source encoding

- Semantics of the data are taken into account
- Lossy coding (encoded data are not identical with original data)
- Degree of compression depends on the data contents
- Examples: Content prediction techniques; e.g. use of spatial redundancies between still images for data compression.
  Discrete Cosine Transformation (DCT) as transformation technique of the spatial domain into the two-dimensional frequency domain

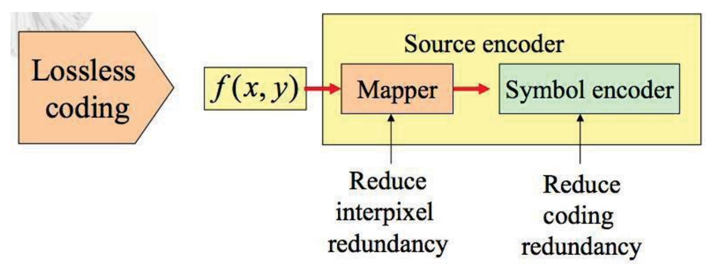
Hybrid encoding

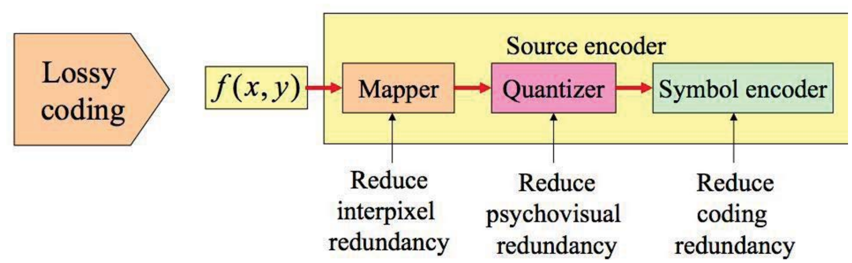
- Used by most multimedia systems
- Combination of entropy and source encoding
- Examples: JPEG, MPEG, H.261, DVI

Lossy Vs. Lossless Compression



Lossless Vs. Lossy Coding





5.1 Data Compression and Coding Fundamentals

Source, Entropy and Hybrid Coding (cont..)

Source Coding

Source coding takes into account the semantics and the characteristics of the data. Thus the degree of compression that can be achieved depends on the data contents.

Source coding is a lossy coding process in which there is some loss of information content. For e.g. in case of speech the speech is transformed from the time domain to frequency domain.

5.1 Data Compression and Coding Fundamentals

Source, Entropy and Hybrid Coding (cont..)

Source Coding(cont..)

In the psychoacoustic the encoder analyzes the incoming audio signals to identify perceptually important information by incorporating several psychoacoustic principles of the human ear.

One is the critical-band spectral analysis, which accounts for the ear‟s poorer discrimination in higher frequency regions than in lower-frequency regions. The encoder performs the psychoacoustic analysis based on either a side-chain FFT analysis or the output of the filter bank.

E.g. Differential Pulse Code Modulation, Delta Modulation, Fast Fourier Transform, Discrete Fourier Transform, Sub-band coding etc

5.1 Data Compression and Coding Fundamentals

Source, Entropy and Hybrid Coding (cont..)

Entropy Coding

Entropy coding is used regardless of the media‟s specific characteristics.

The data stream to be compressed is considered to be a simple digital sequence and the semantics of the data is ignored. It is concerned solely with how the information is represented.

5.1 Data Compression and Coding Fundamentals

Source, Entropy and Hybrid Coding (cont..)

Entropy Coding (cont..)

Run Length Coding

Typical applications of this type of encoding are when the source information comprises long substrings of the same character or binary digit. Instead of transmitting the source string in the form of independent code-words or bits, it is transmitted in the form of a different set of code-words which indicate not only the particular character or bit being transmitted but also indication of the number of characters/bits in the substring.

5.1 Data Compression and Coding Fundamentals

Source, Entropy and Hybrid Coding (cont..)

Entropy Coding (cont..)

Run Length Coding (cont..)

For e.g. if the string is AAAAABBBTTTTTTMMMMMMMM, it is encoded as A!5BBBT!6M!8 (In this case there is no point in encoding characters that repeats itself less than 4 times).

Diatomic encoding is a variation of run-length encoding based on a combination of two data bytes. This technique determines the most frequently occurring pairs of bytes. For e.g. in English language "E","T","TH","A","S","RE","IN" and "HE" occurs most frequently.

5.1 Data Compression and Coding Fundamentals

Source, Entropy and Hybrid Coding (cont..)

Entropy Coding (cont..)

Huffman Coding

Huffman coding is an example of variable length coding. It is based in the concept that the probability of occurrence of the characters is not same so different number of bits is assigned for different character.

Basically in variable length coding the characters that occur most frequently are assigned fewer numbers of bits. However in order to used variable length coding the destination must know the set of code-words being used by the source. In Huffman Coding the probability of occurrence of the characters are estimated and based on this estimation code-words are assigned to the characters.

Q.N.1:-Design Huffman code for probabilities: 0.1, 0.4, 0.06, 0.1, 0.04, and 0.3.

- Huffman coding: give the smallest possible number of code symbols per source symbols.

Step 1: Source reduction

Original source		Source reduction			
Symbol	Probability	1	2	3	4
a_2	0.4	0.4	0.4	0.4	0.6
a_6	0.3	0.3	0.3	0.3	0.4
a_1	0.1	0.1	0.2	0.3	
a_4	0.1	0.1	0.1		
a_3	0.06	0.1			
a_5	0.04				

Step 2: Code assignment procedure

Original source			Source reduction							
Sym.	Prob.	Code	1		2		3		4	
a_2	0.4	1	0.4	1	0.4	1	0.4	1	0.6	0
a_6	0.3	00	0.3	00	0.3	00	0.3	00	0.4	1
a_1	0.1	011	0.1	011	0.2	010	0.3	01		
a_4	0.1	0100	0.1	0100	0.1	011				
a_3	0.06	01010	0.1	0101						
a_5	0.04	01011								

The code is instantaneous uniquely decodable without referencing succeeding symbols.

Average length:
(0.4) (1) + 0.3 (2) + 0.1 X 3 + 0.1 X 4 + (0.06 + 0.04) 5 = 2.2 bits/symbol

Q.N.2:-Design Huffman code for probabilities: 0.1, 0.4, 0.15, 0.05, 0.1, and 0.2.

Q.N.3: let the word to be encoded be AABABBEEEEFEFC

Ans:- Length of string=14

No. of A =3 Probability of occurrence of A=3/14=**0.214**;

No. of B=3 Probability of occurrence of B=3/14=**0.214**

No. of E=5 Probability of occurrence of E=5/14=**0.357**

No. of F=2 Probability of occurrence of F=2/14=**0.142**

No. of C=1 Probability of occurrence of C=1/14=**0.01**

Here , Huffman code for probabilities are :
0.214, 0.214, 0.357, 0.142 and 0.01.

And solve ………..

Q.N.4:- Construct the Huffman code for (PoU 2014 / Fall)

Gray Level	0	1	2	3	4	5	6	7
No.of Pixel	5000	1000	500	530	1250	950	860	130

Q.N.5:- Construct the Huffman code for (PoU 2012 / Fall)

Gray Level	0	1	2	3	4	5	6	7
No. of Pixel	5320	1000	500	525	1236	956	856	128

5.1 Data Compression and Coding Fundamentals

Source, Entropy and Hybrid Coding (cont..)

Entropy Coding (cont..)

Arithmetic Coding

Unlike Huffman coding which used a separate codeword for each character, arithmetic coding yields a single codeword for each encoded string of characters.

The first step is to divide the numeric range from 0 to 1 into a number of different characters present in the message to be sent – including the termination character – and the size of each segment by the probability of the related character.

5.1 Data Compression and Coding Fundamentals

Source, Entropy and Hybrid Coding (cont..)

Hybrid Coding

This type of coding mechanism involve the combine use of both the source coding and the entropy coding for enhancing the compression ratio still preserving the quality of information content. The example of Hybrid Coding includes MPEG, JPEG, H.261, DVI techniques.

5.2. Basic Data Compression Techniques

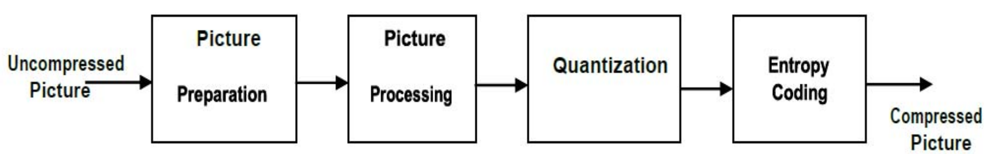
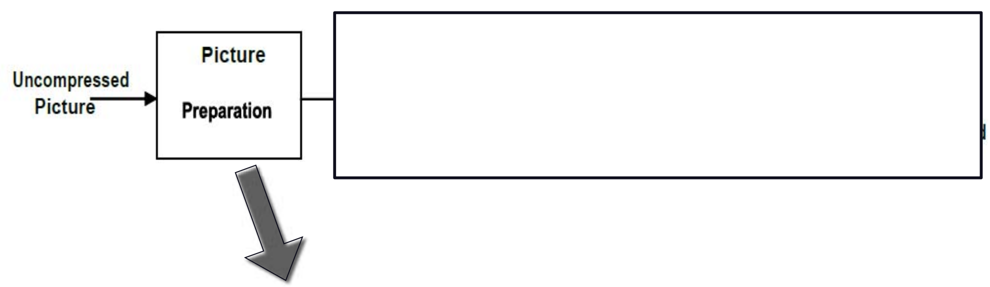


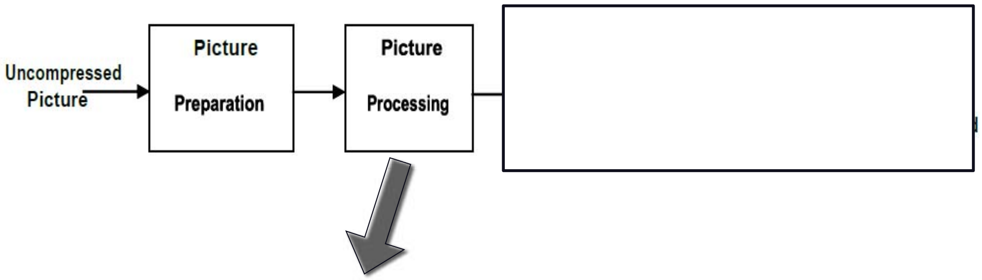
Figure: Major steps of data compression

5.2. Basic Data Compression Techniques



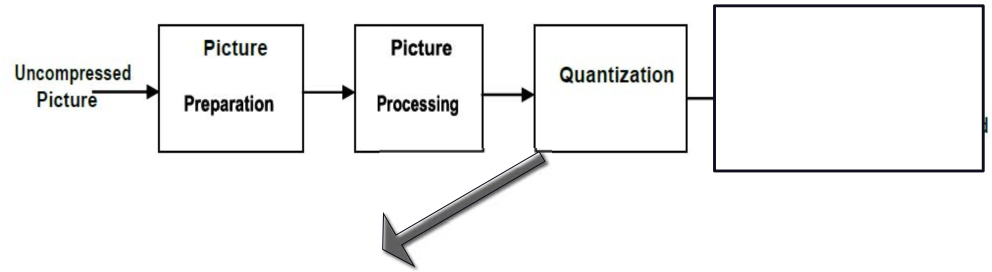
- **Picture preparation:** this includes analog to digital conversion and generating a appropriate digital representation of the information. An image is divided into block of 8 X 8 pixels and represented by fixed number of bits/pixel

5.2. Basic Data Compression Techniques



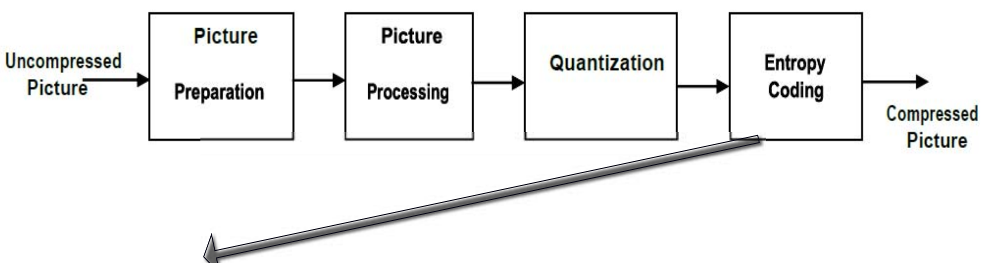
- **Picture processing:** Processing is the actually the first step of compression process which makes use of sophisticated algorithm of transformation from time to frequency domain. It can be performed using DCD (Discrete Cosine transform).

5.2. Basic Data Compression Techniques



- **Quantization Process:** Quantization process the result of the previous step. It specifies the mapping of the real number into integer. This can be considered as the equivalent of μ-law and A- law which apply the audio data. Quantization is performed using the different number of bits per coefficient.

5.2. Basic Data Compression Techniques



- **Entropy Coding:** Entropy coding is usually last step of data compression. It compresses a sequence of digital stream without loss. For e.g. a sequence of zero's in the data stream can be compressed by specifying the number of occurrence followed by zero it.

5.2. Basic Data Compression Techniques

- The **processing** and **quantization** cab be repeated iteratively several times in feedback loop.
- The term spatial domain (time domain) refer to the image plane itself and approaches in this category are based on discrete manipulation of pixel in an image.
- After compression, the compressed video stream contains the specification of the image starting point and an identification of the compression technique may be the part of the data stream. The error correction code may also be added to the stream.
- Decompression is the inverse process of compression.

5.3.Data Compression and Coding Standards:

JPEG (Joint Photographic Experts Group)

The JPEG standard for compressing continuous –tone still pictures (e.g. photographs) was developed by photographic experts working under the joint auspices of ITU, ISO and IEC.

JPEG is significant in compression because MPEG or the standard for motion picture compression is just the JPEG encoding applied to each frame separately.

5.3.Data Compression and Coding Standards:

JPEG (Joint Photographic Experts Group) (cont..)

There are some requirements of JPEG standard and they are:

- The JPEG implementation should be independent of image size.
- The JPEG implementation should be applicable to any image and pixel aspect ratio.
- Color representation itself should be independent of the special implementation.
- Image content may be of any complexity, with any statistical characteristics.
- The JPEG standard specification should be state of art (or near) regarding the compression factor and achieved image quality.
- Processing complexity must permit a software solution to run on as many available standard processors as possible. Additionally, the use of specialization hardware should substantially enhance image quality.

5.3.Data Compression and Coding Standards:

Steps in JPEG Compression

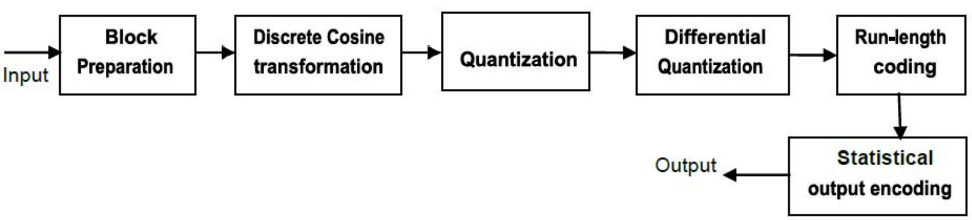


Figure: The operation of JPEG in lossy-sequential mode.

5.3.Data Compression and Coding Standards:

Steps in JPEG Compression (cont..)

Step 1: (Block Preparation)

This step involves the block preparation. For e.g. let us assume the input to be 640*480 RGB image with 24 bits/pixel. The luminance and chrominance component of the image is calculated using the YIQ model for NTSC system.

$$Y=0.30R + 0.59G + 0.11B$$

$$I= 0.60R – 0.28G – 0.32B$$

$$Q= 0.21R – 0.52G + 0.31B$$

5.3.Data Compression and Coding Standards:

Steps in JPEG Compression (cont..)

Step 1: (Block Preparation) (cont..)

For PAL system YUV model is used. Separate matrices are constructed for Y, I and Q each elements in the range of 0 and 255.

The square blocks of four pixels are averaged in the I and Q matrices to reduce them to **320*240**. Thus the data is compressed by a factor of two.

Now, 128 is subtracted from each element of all three matrices to put 0 in the middle of the range. Each image is divided up into 8*8 blocks. The Y matrix has 4800 blocks; the other two have 1200 blocks.

5.3.Data Compression and Coding Standards:

Steps in JPEG Compression (cont..)

Step 2: (Discrete Cosine Transformation)

Discrete Cosine Transformation is applied to each 7200 blocks separately.

The output of each DCT is an 8*8 matrix of DCT coefficients.

DCT element (0,0)is the average value of the block. The other element tells how much spectral power is present at each spatial frequency.

5.3.Data Compression and Coding Standards:

Steps in JPEG Compression (cont..)

Step 3: (Quantization)

In this step the less important DCT coefficients are wiped out.

This transformation is done by dividing each of the coefficients in the 8*8 DCT matrix by a weight taken from a table.

If all the weights are 1 the transformation does nothing however, if the weights increase sharply from the origin, higher spatial frequencies are dropped quickly.

5.3.Data Compression and Coding Standards:

Steps in JPEG Compression (cont..)

Step 4: (Differential Quantization)

This step reduces the (0,0) value of each block by replacing it with the amount it differs from the corresponding element in the previous block.

Since these elements are the averages of their respective blocks, they should change slowly, so taking the differential values should reduce most of them to small values.

The (0,0) values are referred to as the DC components; the other values are the AC components.

5.3.Data Compression and Coding Standards:

Steps in JPEG Compression (cont..)

Step 5: (Run length Encoding)

This step linearizes the 64 elements and applies run-length encoding to the list.

In order to concentrate zeros together, a zigzag scanning pattern is used.

Finally run length coding is used to compress the elements.

5.3.Data Compression and Coding Standards:

Steps in JPEG Compression (cont..)

Step 6: (Statistical Encoding)

Huffman encodes the numbers for storage or transmission, assigning common numbers shorter codes than uncommon ones.

JPEG produces a 20:1 or even better compression ratio. Decoding a JPEG image requires running the algorithm backward and thus it is roughly symmetric: decoding takes as long as encoding.

5.3.Data Compression and Coding Standards:

Lossy Sequential DCT-based Mode of JPEG

Image Processing

It basically involves the block preparation where the image samples are grouped into 8*8 pixels and passed to the encoder.

Then Discrete Cosine Transformation is applied to the blocks where the pixel values are shifted into the range [-128,127] with zero as the center.

Each of these values is then transformed using *Forward DCT (FDCT)*. DCT is similar to *Discrete Fourier Transformation* as it maps the values from the time to the frequency domain.

5.3.Data Compression and Coding Standards:

Lossy Sequential DCT-based Mode of JPEG

Quantization

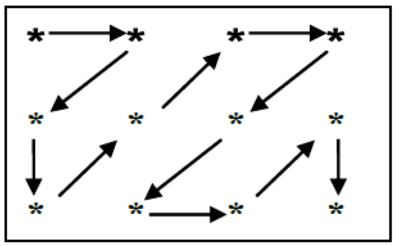
The JPEG application provides a table of 64 entries. Each entry will be used for the quantization of one of the 64 DCT-coefficients

5.3.Data Compression and Coding Standards:

Lossy Sequential DCT-based Mode of JPEG

Entropy Encoding

During the initial step of entropy encoding, the quantized DC-coefficients are treated separately from the quantized AC-coefficients.



5.3.Data Compression and Coding Standards:

Lossy Sequential DCT-based Mode of JPEG

Entropy Encoding

- The DC-coefficient determines the basic color of the data units.
- The DCT processing order of the AC coefficients involves the zigzag sequence to concentrate the number of zeros.

JPEG specifies Huffman and arithmetic encoding as entropy encoding methods. However, as this is lossy sequential DCT-based mode, only Huffman encoding is allowed. In lossy sequential mode the framework of the whole picture is not formed but parts of it are drawn i.e. sequentially done.

5.3.Data Compression and Coding Standards:

Expanded Lossy DCT-based Mode

It differs from the sequential mode in terms of number of bits per sample. Here 12 bits along with 8 bits per sample can be used.

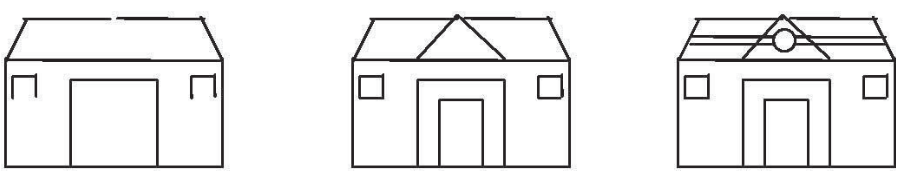


Figure: Progressive Picture Presentation

5.3.Data Compression and Coding Standards:

Expanded Lossy DCT-based Mode (cont..)

For the expanded lossy DCT-based mode, JPEG specifies *progressive encoding* in addition to sequential encoding.

At first, a very rough representation of the image appears which is progressively refined until the whole image is formed. This progressive coding is achieved by layered coding.

5.3.Data Compression and Coding Standards:

Expanded Lossy DCT-based Mode (cont..)

Progressiveness is achieved in two different ways:

- By using a *spectral selection* in the first run only, the quantized DCT coefficients of low frequencies of each data unit are passed in the entropy encoding. In successive runs, the coefficients of higher frequencies are processed.
- *Successive approximation* transfers all of the quantized coefficients in each run, but single bits are differentiated according to their significance. The most-significant bits are encoded first, then the less-significant bits.

5.3.Data Compression and Coding Standards:

Hierarchical Mode

This mode uses either the lossy DCT-based algorithms or the lossless compression technique. The main feature of this mode is the encoding of an image at different resolutions, i.e., the encoded data contains images at several resolutions. This involves the sampling of images at higher resolutions at first and then subsequently reducing the horizontal and vertical resolution. The compressed image is subtracted from the previous result. The process is repeated until the full resolution is reduced considerably and no further compression is possible.

This requires more storage but as images at different resolution are available the systems using low resolution need not decode the image as the low resolution image is easily available.

5.3.Data Compression and Coding Standards:

MPEG (Motion Picture Expert Group)

The Motion Pictures Expert Group was formed by the ISO to formulate a set of standards relating to a range of multimedia applications that involve the use of video with sound.

The coders associated with the audio compression part of these standards are known MPEG audio coders and a number of these use perceptual coding.

MPEG can deliver a data rate of at most 1856000 bits/second, which should not be exceeded. Data rates for audio are between 32 and 448 Kbits/second; this data rate enables video and audio compression of acceptable quality.

5.3.Data Compression and Coding Standards:

MPEG (Motion Picture Expert Group) (cont..)

The MPEG standard exploits the other standards and they are

- JPEG:

  JPEG is used as the motion picture is a continuous sequence of still image.
- H.261:

  H.261 video compression standard has been defined by the ITU-T for the provision of video telephony and videoconferencing services over an Integrated Services Digital Network (ISDN). It specifies two resolution formats with an aspect ratio of 4:3 are specified. *Common Intermediate Format* (CIF) defines a luminance component of 288 lines, each with 352 pixels. The chrominance components have a resolution with a rate of 144 lines and 176 pixels per line. *Quarter CIF* (QCIF) has exactly half of the CIF resolution i.e., 176*144 pixels for the luminance and 88*72 pixels for the other components.

5.3.Data Compression and Coding Standards:

MPEG (Motion Picture Expert Group) (cont..)

MPEG video uses video compression algorithms called ***Motion-Compensated Discrete Cosine Transform*** algorithms. The algorithms use the following basic algorithm

- ***Temporal Prediction****:* It exploits the temporal redundancy between video pictures.
- ***Frequency Domain Decomposition****:* It uses DCT to decompose spatial blocks of image data to exploit statistical and perceptual spatial redundancy.
- ***Quantization****:* It reduces bit rate while minimizing loss of perceptual quality.
- ***Variable-length Coding****:* It exploits the statistical redundancy in the symbol sequence resulting from quantization as well as in various types of side information.

5.3.Data Compression and Coding Standards:

MPEG (Motion Picture Expert Group) (cont..)

As far as audio compression is concerned the time-varying audio input signal is first sampled and quantized using PCM, the sampling rate and number of bits per sample being determined by the specific application.

The bandwidth that is available for transmission is divided into a number of *frequency sub bands* using a bank of *analysis function* which because of their role, are also known as *critical-band filters.*

5.3.Data Compression and Coding Standards:

Video Encoding

Video is nothing but simply a sequence of digitized pictures. The video that MPEG expects to process is composed of a sequence of frames or fields of luma and chroma.

Frame-Based Representation**:**

MPEG-1 is restricted to representing video as a sequence of frames. Each frame consists of three rectangular arrays of pixels, one for the luma (Y, black and white) component, and one each for the chroma (Cr and Cb, color difference) components. The chroma arrays in MPEG-1 are sub-sampled by a factor of two both vertically and horizontally relative to the luma array.

Field-Based Representation:

MPEG-2 is optimized for a wider class of video representations, including, most importantly, field-based sequences. *Fields* are created by dividing each frame into a set of two interlaced fields, with odd lines from the frame belonging to one field and even lines to the other. The fields are transmitted in interlaced video one after the other, separated by half a frame time.

5.3.Data Compression and Coding Standards:

Video Encoding (cont..)

MPEG provides 14 different image aspect ratios per pixel which are coded in the data stream. The image refresh frequency is also encoded in the data stream. Eight frequencies are defined: 23.976Hz, 24 Hz, 25 Hz, 29.97 Hz, 30 Hz, 50 Hz, 59.94 Hz and 60 Hz.

5.3.Data Compression and Coding Standards:

Types of Image Frames in MPEG

There are two basic types of compressed frame: those that are encoded independently and those that are predicted.

The first are as *intracoded frames* or *I frames.* In practice, there are two types of predicted frames: *predictive* or *P-frames* and *bidirectional* or *B-frames* and because of the way they are derived, the latter are also known as *intercoded* or *interpolation frames.*

5.3.Data Compression and Coding Standards:

Types of Image Frames in MPEG (cont..)

I-frames (Intra coded Images):

I- frames are encoded without reference to any other frames.

Each frame is treated as a separate (digitized) **picture** and the Y, C , and C matrices are encoded independently using the JPEG algorithm.

b r The level of compression obtained with I-frames is relatively small. I-frames must be present in the output stream at regular intervals in order to allow for the possibility of the contents of an encoded I-frame being corrupted during transmission. The number of frames/pictures between successive I-frames is known as a *group of pictures* or *GOP.*

5.3.Data Compression and Coding Standards:

Types of Image Frames in MPEG (cont..)

P-frames (Predictive-Coded Frames):

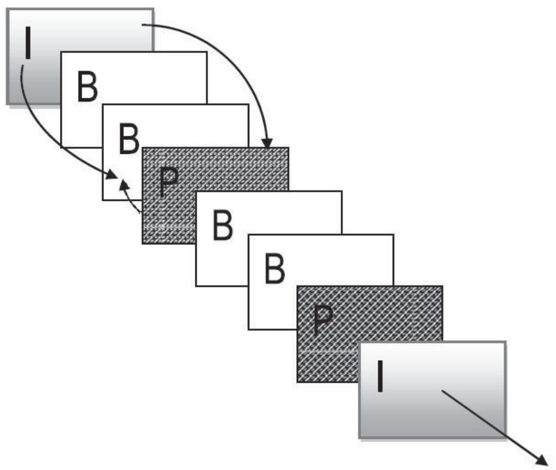
The encoding of a P-frame is relative to the contents of either a preceding I-frame or a preceding P frame.

As indicated, P-frames are encoded using a combination of motion estimation and motion compensation and hence significantly higher levels of compression can be obtained.

In practice, the number of P-frames between each successive pair of I-frames is limited since any errors present in the first P-frame will be propagated to the next. The number of frames between a P-frame and the immediately preceding I- or P-frame is called the *prediction span.* It is given the symbol M and typical values range from 1 through 3.

5.3.Data Compression and Coding Standards:

Types of Image Frames in MPEG (cont..)



5.3.Data Compression and Coding Standards:

Types of Image Frames in MPEG (cont..)

***B-frames (Bi-directionally predictive-coded frames)*:**

Motion estimation involves comparing small segments of two consecutive frames for differences and, should a difference be detected, a search is carried out to determine to which neighboring segment the original segments has moved. In order to minimize the time for each search, the search region is limited to just a few neighboring segments. Some applications may involve very fast moving objects; however, it is possible for a segment to have moved outside the search region. To allow for this possibility, in applications such as movies, in addition to P-frames, second types of frames are used called B-frames.

5.3.Data Compression and Coding Standards:

Types of Image Frames in MPEG (cont..)

D-frames (DC-Coded Frames):

D-frame has been defined for use in movie/video-on-demand applications.

D-frames are used for display in fast-forward or fast-rewind modes. D-frames are inserted at regular intervals throughout the stream. These are highly compressed frames and are ignored during the decoding of P- and Bcompression algorithm, the DC coefficient associated with each 8 * 8 block pixels- both for the luminance and the two chrominance signals- is the mean of all the values in the related block. Hence by using only encoded DC coefficients of each block of pixels in the periodically inserted D-frames, a lowresolution sequence of frames is provided each of which can be decoded at the higher speeds that are expected with the rewind and fast-forward operations.

5.3.Data Compression and Coding Standards:

Audio Encoding:

The time-varying audio input signal is first sampled and quantized using PCM, the sampling rate and number of bits per sample being determined by the specific application. The bandwidth that is available for transmission is divided into a number of *frequency sub bands* using a bank of *analysis filters* which, because of their role, are also known as *critical-band filters.*

Each frequency subbands is of equal width and, essentially, the bank of filters maps each set of 32 PCM samples into an equivalent frequency samples, one per subband, Hence each is known as a *sub band sample* and indicates the magnitude of each of the 32 frequency components that are present in a segment of the audio input signals of a time duration equal to 32 PCM samples.

5.3.Data Compression and Coding Standards:

Audio Encoding:

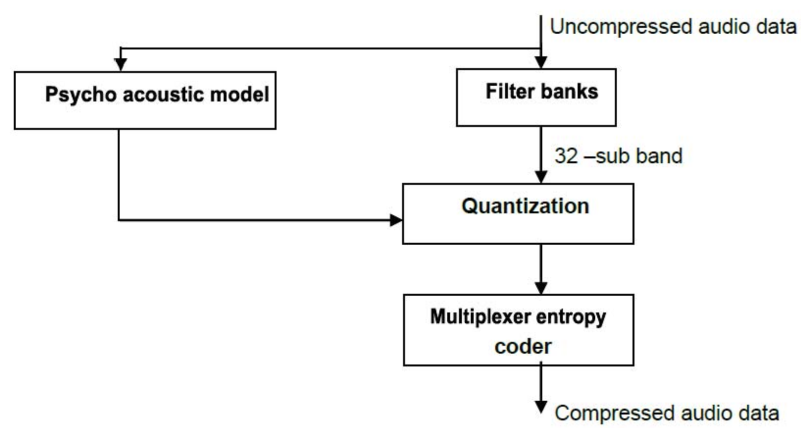


Figure: MPEG basic steps for audio encoding

5.3.Data Compression and Coding Standards:

DVI (Digital Video Interactive)

DVI is a technology that includes coding algorithms. The fundamental components are a VLSI chip set for the video subsystem, a well specified data format for audio and video files, an application user interface to the audio-visual kernel and compression, as well as decompression, algorithms. For encoding audio standard signal processor is used. Processing of images and video is performed by a video processor.

5.3.Data Compression and Coding Standards:

DVI (Digital Video Interactive) (cont..)

Audio and Still Image encoding:

Audio signals are digitized using 16-bits per sample. Audio signals may be PCM-encoded or compressed using the adaptive differential pulse coded modulation (ADPCM) technique. Supported sampling frequencies are: 11025Hz, 22050Hz and 44100 Hz for one or two PCM-coded channels. And 8268Hz, 31129Hz, 33075Hz for ADPCM.

5.3.Data Compression and Coding Standards:

DVI (Digital Video Interactive) (cont..)

Audio and Still Image encoding:

For Still Images, DVI assumes an internal digital YUV format for image preparation. Any video input signal must first be transformed into this format. The color of each pixel is split into luminance component and the two chrominance components (U and V). The luminance represents the gray scale image. With RGB, DVI computes the YUV signal using the following relationship.

$$Y=0.30R+0.59G+0.11B$$

$$U=B-Y$$

$$V=R-Y$$

5.3.Data Compression and Coding Standards:

DVI (Digital Video Interactive) (cont..)

Audio and Still Image encoding: (cont..)

It leads to:

$$U=-30R-0.59G+0.89B$$

$$V=0.70R-059G-0.11B$$

DVI Determines the components YUV according to the following:

$$Y=0.299R+0.587G+0.144B+16$$

$$U=0.577B-0.577Y+137.23$$

$$V=0.730R-0.730Y+139.67$$

5.3.Data Compression and Coding Standards:

DVI (Digital Video Interactive) (cont..)

Audio and Still Image encoding: (cont..)

DVI is able to process image in the 16-bit YUV format and the 24-bit YUV format. The 24-bit YUV format uses 8 bits for each component. The 16-bit YUV format coded the Y components of each pixel with 6 bits and the color difference components with 5 bits each.

There are 2 bitmap formats: Planer and Packed.

Planer:

All data of the Y component are stored first, followed by the U component values and then all V values.

Packed:

For the packed bitmap format, the Y, U, and V information of each pixel is stored together by the data of the next pixel.

Tutorial I

1. Define multimedia. Explain the application areas of multimedia.
2. What is Multimedia? Explain global structure of multimedia system.
3. Define application domain? Explain the different criteria that are used to classify media in the multimedia system.
4. Define Multimedia System. Explain properties of Multimedia System.
5. Write short notes
   a) Data stream with Transmission mode
   b) Data stream characteristics for continuous media
   c) Information unit

Tutorial II

1. Explain the different components of a MIDI device.
2. Illustrate the importance of MIDI. Explain the significance of MIDI messages.
3. What is MIDI? What features of MIDI make it suitable for multimedia applications?
4. How can speech be generated from a digital device? Explain in detail.
5. With necessary diagrams explain how the sound is digitized and the sound stored in a multimedia system.

Tutorial III

1. How does a digital computer represent a color image? Explain the digital image representation with conceptual diagram.
2. How do you mean by digital image? Explain the different types of image format.
3. What are the steps involve in Image Recognition? Explain with necessary block diagrams.
4. What are the application ares of image processing? Explain.
5. Briefly explain the spatial filtering technique for image enhancement.

Tutorial IV

1. What do you mean by computer based animation? List the different types of animation languages.
2. List three distinct models of color used in Multimedia. Explain why there are a number of different color models exploited in multimedia data formats.
3. Explain Tele-services and the implementation of Conversation services in Multimedia communication.
4. Explain the methods that are used to control animation. Discuss the YUV model for video transmission.
5. How long will it take to transmit a minute long video of spatial resolution 640×480, 32 bits per pixel and 12 frames per second through a communication link at a constant rate of 56 K bits PS.
6. Describe the television standards.

Tutorial V

1. How is source coding different from entropy encoding? Describe about the MPEG video compression.
2. What are the different types of compression technique used? Explain in detail any one Source encoding technique used for data compression.
3. How is source coding different from entropy encoding? Describe about the JPEG compression.
4. What is data compression? Why multimedia data should be compressed? Describe the JPEG compression with its different modes.
5. Define Run length encoding. Construct the Huffman code for:

Gray Level	0	1	2	3	4	5	6	7
No. of Pixel	4500	1500	900	750	1200	1300	550	100

COMPUTER TECHNOLOGY AND MULTIMEDIA OPERATING SYSTEM (MOS)

Unit 7

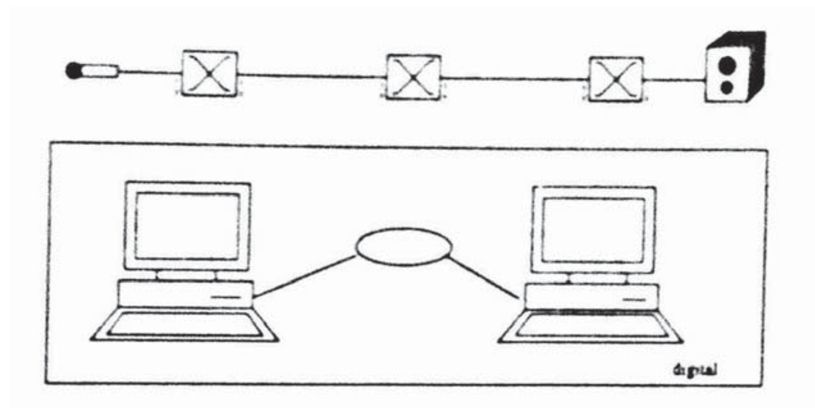
7.1. Communication Architecture

Local multimedia systems (i.e., multimedia workstations) frequently include a network interface (e.g., Ethernet card) through which they can communicate with each other. However, the transmission of audio and video cannot be carried out with only the conventional communication infrastructure and network adapters.

Until now, the solution was that continuous and discrete media have been considered in different environments, independently of each other. It means that fully different systems were built. For example, on the one hand, the *analog telephone system* provides audio transmission services using its original dial devices connected by copper wires to the telephone company's nearest *end office*. The end offices are connected to switching centers, called *toll offices*, and these centers are connected through high bandwidth intertoll trunks to *intermediate switching offices*. This hierarchical structure allows for reliable audio communication. On the other hand, *digital computer networks* provide data transmission services at lower data rates using network adapters connected by copper wires to switches and routers.

Even today, professional radio and television studios transmit audio and video

7.1. Communication Architecture



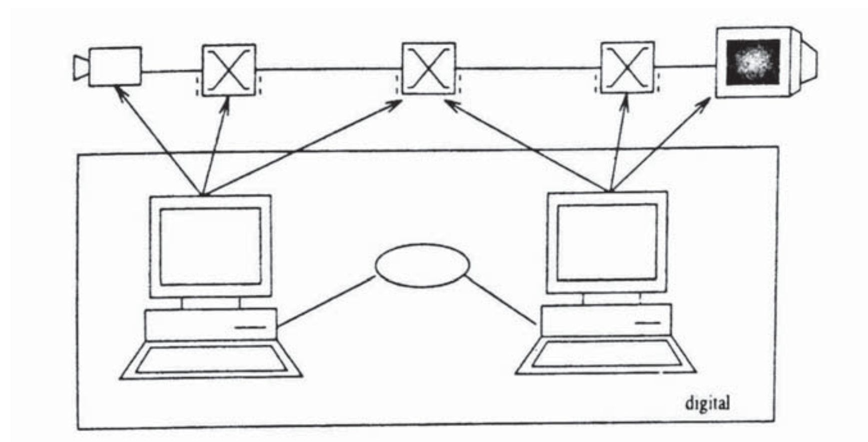
Analog and digital environments without interaction

7.1. Communication Architecture: Hybrid System

By using existing technologies, integration and interaction between analog and digital environments can be implemented. This integration approach is called the *hybrid approach*.

The main advantage of this approach is the high quality of audio and video and all the necessary devices for input, output, storage and transfer that are available. The hybrid approach is used for studying application user interfaces, application programming interfaces or application scenarios. The transmission techniques used in these cases are less important, although to meet the goal of full digital integration, this approach is not satisfactory.

7.1. Communication Architecture: Hybrid System



Computer control of all audio-video components.

7.1. Communication Architecture: Hybrid System

- This trend has emerged because with cooperative peers it is possible to asymptotically enhance the use of resources in sharing of data compared to the basic client-server architecture.
- The need for distribution of data is wide and one could argue that it is as fundamental a building block as the message passing of the Internet.
- As an answer to this need a new scalable architecture is introduced: **Hybrid Communication Architecture** (HCA), which provides both data sharing and message passing as communication primitives for applications.

7.1. Communication Architecture: Hybrid System

- HCA can be regarded as an abstraction layer for communication which is further encapsulated by a higher-level middleware.
- HCA is aimed at general use, and it is not designed for any particular application.
- One key idea is to combine data sharing with streaming since together they enable many applications not easily implementable with only one of these features.
- For example, a game application could share the game world state between clients and modify it by using streaming. The other distinctive feature of the system is the use of knowledge of the physical network topology in the optimization of the communication. With a feasible business model, fault-tolerance, and security features, HCA is aimed eventually for real-life adoption.

7.1. Communication Architecture: Hybrid System

The following interfaces and protocols form the skeleton of the HCA framework:

- Client interface is used by applications on top of HCA,
- DLL interface divides the client side implementation of HCA into dynamically and statically linked parts,
- Context interface passes information from the execution environment to the application,
- Security interface provides an access to cryptographic functions,
- OSLib interface provides an abstraction layer for different operating systems,
- Inter-domain protocol is used by communication nodes of the HCA network,
- Intra-domain protocol is used by communication nodes of the same domain as explained later.

7.1. Communication Architecture: Hybrid System

In a normal configuration, multiple applications with their HCA client interface static implementations and a single local node implementation are located on the same **client machine** at the edge of the HCA overlay network.

The **client machine** can be, for example, the user's home computer or office workstation which is shut down every now and then. The local node communicates with the nearest communication node using the inter-domain protocol.

This node probably resides on a different machine typically administered by a service provider organization or it could be located in a company intranet server.

7.1. Communication Architecture: Hybrid System

The HCA overlay network is formed by many communication nodes structured in the form of hierarchical domain tree.

The nodes are connected by the inter-domain protocol and typically reside in servers that are continuously online. Clients and persistence servers can be thought of residing at the edge of the overlay network.

Client interface, OSLib, Context interface and Security interface form together an operating-system- and hardware-independent platform for components on top of HCA.

7.1. Communication Architecture: Hybrid System

Client Interface

The HCA Client interface is the only part of the communication architecture visible to applications and higher-level layers of the middleware.

Because HCA is intended to be programming language neutral, there are language mappings of the client interface to each programming language supported.

The implementation of the client interface is normally divided to statically and dynamically linked parts so that it is possible to change the implementation without recompiling applications.

7.1. Communication Architecture: Hybrid System

DLL Interface and Local Node Implementation

HCA DLL (Dynamic link library) interface is an operating-system dependent application programming interface (API) for wrapping the local communication node to a separate process running on the client machine.

The DLL implementation shares the local communication node between multiple client processes and implements the communication between client processes and the local communication node process using fast inter-process communication, shared memory or other facilities provided by the operating system.

7.1. Communication Architecture: Hybrid System

OSLib

OSLib programming interface provides a system-wide hardware abstraction layer (HAL) for preventing the components of HCA of having unnecessary dependencies to particular operating system or hardware. With this modularization in implementations we gain the easy portability of HCA implementations over different platforms.

7.1. Communication Architecture: Hybrid System

Context Interface

Context interface is intended to pass information from the execution platform to the program. For example, communication node implementations can use it to read the configuration information of the node

Security Interface

Security interface contains a uniform interface for typical cryptographic algorithms. An abstract interface for both public key and secret key methods are specified.

Inter-domain Protocol

Inter-domain protocol is used between all communication nodes to form the overlay network of HCA

7.1. Communication Architecture: Hybrid System

Intra-domain Protocol

Each domain can use their own intra-domain protocol for coordination of the nodes inside the domain.

Nodes must implement the intra-domain protocol used by their enclosing domain. Small domains do not necessarily need an intra-domain protocol for managing nodes because it can be achieved by configuring each node manually.

7.1. Communication Architecture: Digital System

Digital systems are designed to store, process, and communicate information in digital form.

They are found in a wide range of applications, including process control, communication systems, digital instruments, and consumer products.

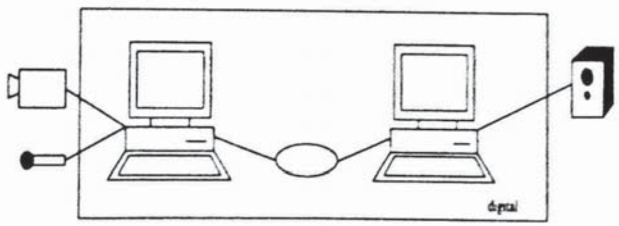
The digital computer, more commonly called the *computer*, is an example of a typical digital system.

7.1. Communication Architecture: Digital System

Connection to Workstations

In digital systems, audio-video devices can be connected directly to the computers (workstations) and digitized audio-video data are transmitted over shared data networks, Audio-video devices in these systems can be either analog or digital. Figure 8.5 shows an integrated system structure with analog devices and A/D and D/A interfaces. Figure 8.6 shows an integrated system structure with digital end-system devices and interfaces.

7.1. Communication Architecture: Digital System



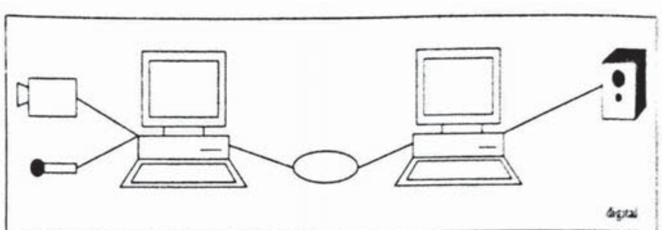
Integrated (with respect to hardware) system structure with analog end-system devices and A/D and D/A interfaces.

7.1. Communication Architecture: Digital System

An example of a digital system is the *Etherphone* system from Xerox PARC [Swi87]. A digital audio communication was demonstrated over an Ethernet, although not in a fully integrated form, i.e., the audio was not processed in the main memory.

Another example is an early project by AT&T in Naporville, which considered a similar system architecture to a Etherphone [LL89, LBH+90]. Here, a computer was directly connected to a *Fast Packet Switching* network. The processing of continuous media in the computer was allowed through extensions of the UNIX operating

7.1. Communication Architecture: Digital System



Integrated (with respect to hardware) system structure with digital end-system devices and interfaces.

7.1. Communication Architecture: Digital System

A computer manipulates information in digital, or more precisely, binary form. A binary number has only two discrete values — zero or one.

7.2. Multimedia Workstation

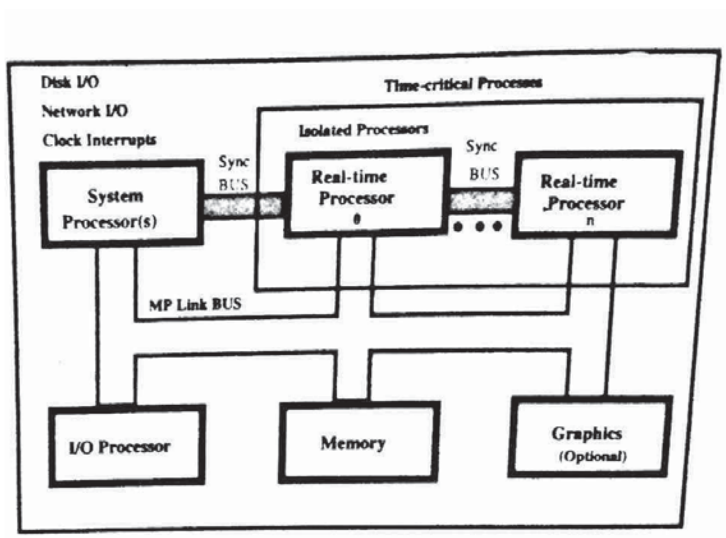
Current workstations are designed for the manipulation of discrete media information. The data should be exchanged as quickly as possible between the involved components, often interconnected by a common bus. Computationally intensive and dedicated processing requirements lead to dedicated hardware, firmware and additional boards. Examples of these components are hard disk controllers and FDDI-adapters.

7.2. Multimedia Workstation

A *multimedia workstation* is designed for the simultaneous manipulation of discrete and continuous media information. The main components of a multimedia workstation are:

- *Standard Processor(s)* for the processing of discrete media information.
- *Main Memory* and *Secondary Storage* with corresponding autonomous controllers.
- *Universal Processor(s)* for processing of data in real-time (signal processors).
- *Special-Purpose Processors* designed for graphics, audio and video media (containing, for example, a micro code decompression method for DVI processors) [Rip89, Tin89, Lut91].
- *Graphics and Video Adapters.*
- *Communications Adapters* (for example, the Asynchronous Transfer Mode Host Interface [TS93].
- Further *special-purpose adapters.*

7.2. Multimedia Workstation



7.3. Introduction to MOS / Function of MOS

The operating system provides a comfortable environment for the execution of programs, and it ensures effective utilization of the computer hardware.

The OS offers various services related to the essential resources of a computer: CPU, main memory, storage and all input and output devices.

In multimedia applications, a lot of data manipulation (e.g. A/D, D/A and format conversion) is required and this involves a lot of data transfer, which consumes many resources.

7.3. Introduction to MOS / Function of MOS

The integration of discrete and continuous multimedia data demands additional services from many operating system components.

The major aspect in this context is *real-time processing* of continuous media data.

7.3. Introduction to MOS / Function of MOS

Issues concerned:

- Process management: a brief presentation of traditional real-time scheduling algorithms.
- File systems: outlines disk access algorithms, data placement and structuring
- Interprocess communication and synchronization
- Memory management
- Database management
- Device management

7.3. Introduction to MOS / Function of MOS

Process management must take into account the timing requirement imposed by the handling of multimedia data.

• Concerns in process management (Scheduling):

	Traditional OS	MM OS
Timing requirement	No	Yes
Fairness	Yes	Yes

7.3. Introduction to MOS / Function of MOS

Single components are conceived as resources that are reserved prior to execution to obey timing requirements and this *resource reservation* has to cover all resources on a data path.

The *communication & synchronization between single processes* must meet the restrictions of real-time requirements and timing relations among different media.

7.3. Introduction to MOS / Function of MOS

Memory management has to provide access to data with a guaranteed timing delay and efficient data manipulation functions. (e.g. should minimize physical data copy operations.)

Database management should rely on file management services

7.5. Multimedia Real Time System

A real-time process is a process which delivers the results of the processing in a given time-span.

The main characteristic of real-time systems is the correctness of the computation.

- Errorless computation
- The time in which the result is presented

7.5. Multimedia Real Time System

Speed and efficiency are not the main characteristic of real-time systems. (e.g. the video data should be presented at the right time, neither too quickly nor too slowly)

Timing and logical dependencies among different related tasks, processed at the same time, must also be considered.

7.5. Multimedia Real Time System

Deadlines:

A deadline represents the latest acceptable time for the presentation of a processing result.

Soft deadline:

- a deadline which cannot be exactly determined and which failing to meet does not produce an unacceptable result.
- Its miss may be tolerated as long as (1) not too many deadlines are missed and/or (2) the deadlines are not missed by much.

7.5. Multimedia Real Time System

Hard deadline:

- a deadline which should never be violated.
- Its violation causes a system failure.
- Determined by the physical characteristics of real-time processes.

7.5. Multimedia Real Time System

Characteristics of real time systems

The necessity of deterministic and predictable behavior of real-time systems requires processing guarantees for time-critical tasks.

A real-time system is distinguished by the following features:

- Predictably fast response to time-critical events and accurate timing information:
- A high degree of schedulability: to meet the deadlines.
- Stability under transient overload: critical task first.

7.5. Multimedia Real Time System

Real time and multimedia

The real-time requirements of traditional real-time scheduling techniques usually have a high demand for security and fault-tolerance. (Most of them involve system control.)

7.5. Multimedia Real Time System

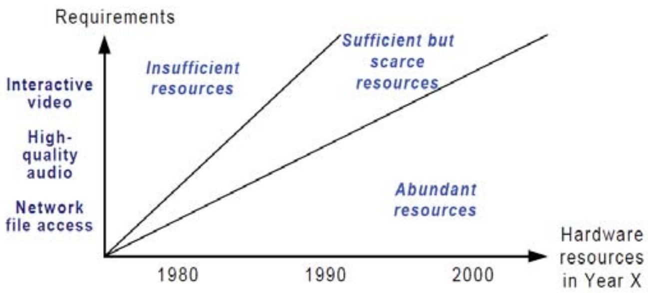
Real-time requirements of multimedia systems:

- The fault-tolerance requirements of multimedia systems are usually less strict than those of real-time systems that have a direct physical impact.
- For many multimedia system applications, missing a deadline is not a severe failure, although it should be avoided. (e.g. playing a video sequence)
- In general, all time-critical operations are periodic and schedulability considerations for periodic tasks are much easier.
- The bandwidth demand of continuous media is usually negotiable and the media is usually scalable.

7.5. Multimedia Real Time System
(Resource management)

Resource management

Multimedia systems with integrated audio and video processing are at the limit of their capacity even with data compression and utilization of new technology. (Demand increases drastically.)



7.5. Multimedia Real Time System
(Resource management)

No redundancy of resource capacity can be expected in the near future.

In a multimedia system, the given timing guarantees for the processing of continuous media must be adhered to along the data path.

The actual requirements depend on (1) the type of media and (2) the nature of the applications supported.

The shortage of resources requires careful allocation.

The resource is first allocated and then managed.

7.5. Multimedia Real Time System
(Resource management)

At the connection establishment phase, the resource management ensures that the new 'connection;' does not violate performance guarantees already provided to existing connections.

Applied to OS, resource management covers the CPU (including process management), memory management, the file system and the device management.

The resource reservation is identical for all resources, whereas the management is different for each.

7.5. Multimedia Real Time System
(Resource management)

Resources

A resource is a system entity required by tasks for manipulating data.

A resource can be active or passive.

- Active resource:
  - e.g. the CPU or a network adapter for protocol processing;
  - it provides a service.
- Passive resource:
  - e.g. main memory, communication bandwidth or file systems;
  - It denotes some system capability required by active resources.

7.5. Multimedia Real Time System
(Resource management)

Resources

A resource can be either used exclusively by one process at a time or shared between various processes.

Active ones are often exclusive while passive ones can usually be shared.

Each resource has a capacity in a given time-span. (e.g. processing time for CPU, the amount of storage for memory and etc.)

For real-time scheduling, only the temporal division of resource capacity among real-time processes is of interest.

7.5. Multimedia Real Time System
(Resource management)

Requirements

The requirements of multimedia applications and data streams must be served.

The transmission/processing requirements of local and distributed multimedia applications can be specified according to the following characteristics:

- *Throughput*: Determined by the needed data rate of a connection to satisfy the application requirements.
- *Delay "at the resource"* (local): The maximum time span for the completion of a certain task at this resource.
- *End-to-end delay* (global): The total delay for a data unit to be transmitted from the source to its destination.
- *Jitter*: Determines the maximum allowed variance in the arrival of data at the destination.
- *Reliability*: Defines error detection and error correction mechanisms used for the transmission and processing of multimedia tasks.
  - How to handle errors: Ignored, indicated and / or corrected.
  - Retransmission may not be acceptable for time critical data.
- These requirements are known as *Quality of Service* (QoS) parameters.

7.5. Multimedia Real Time System
(Resource management)

Components and phases

Resource allocation and management can be based on the interaction between clients and their respective resource managers.

The client selects the resource and requests a resource allocation by specifying its QoS specification.

The resource manager checks its own resource utilization and decides if the reservation request can be served or not.

Performance can be guaranteed once it is accepted.

7.5. Multimedia Real Time System
(Resource management)

Components and phases (cont…)

Phases of the resource reservation and management process

1.Schedulability

- The resource manager checks with the given QoS parameters (e.g. throughput and reliability).

2.QoS calculation

- The resource manager calculates the best possible performance (e.g. delay) the resource can guarantee for the new request.

3.Resource reservation

- Allocates the required capacity to meet the QoS guarantees for each request.

4. Resource scheduling

- Incoming messages (i.e. LDUs) from connections are scheduled according to the given QoS guarantees.

7.5. Multimedia Real Time System

(**Resource management)**

Allocation Scheme

Reservation of resources can be made either in a pessimistic or optimistic way:

The pessimistic approach avoids resource conflicts by making reservations for the worst case. (It's very conservative.)

The optimistic approach reserves resources according to an average workload only.

7.5. Multimedia Real Time System

(**Resource management)**

Allocation Scheme (cont..)

	Pessimistic approach	Optimistic approach
Account for	Worst case	Average case
QoS	Guaranteed	Best effort
Utilization	Low	High
Remarks		May need a monitor to detect overload situation and act

7.5. Multimedia Real Time System

(**Resource management)**

Continuous media resource model

A model is frequently adopted to define QoS parameters and the characteristics of the data stream.

It is based on the model of linear bounded arrival process (LBAP).

A distributed system is decomposed into a chain of resources traversed by the messages on their end-to-end path.

The data stream consists of LDUs (messages). Various data streams are independent of each other.

The model considers a burst of messages consists of messages that arrived ahead of schedule.

DOCUMENTATION, HYPERTEXT AND MHEG

Unit 8

8.1. Document Architecture and Multimedia Integration

A document consists of a set of structural information that can be different forms of media, and during presentation can be generated or recorded.

Documents

A multimedia document is a document which is comprised of information coded in at least one continuous (time-dependent) medium and in one discrete (time independent) medium. The media are synchronized. A multimedia document is closely related to its environment of tools, data abstractions, basic concepts and document architecture.

8.1. Document Architecture and Multimedia Integration

Documents

Continuous and discrete data are processed differently: text is processed within an editor program as a type of a programming language; a motion picture can be manipulated with the same editor program only through library calls.

Basic *system* concepts for document processing use multimedia abstractions and also serve as concepts for the information architecture in a document.

8.1. Document Architecture and Multimedia Integration

Document Architecture

In order to exchange documents the content of the document as well as its structure needs to be communicated. This requires the use of standard format for the documents so that it can be communicated to wide range of users.

The current formats on the process of standardization are ***Standard Generalized Markup Language (SGML)*** and the ***Open Document Architecture (ODA).*** It means a standard architecture of document is required.

Document architecture describes the connections among the individual elements represented as models.

8.1. Document Architecture and Multimedia Integration

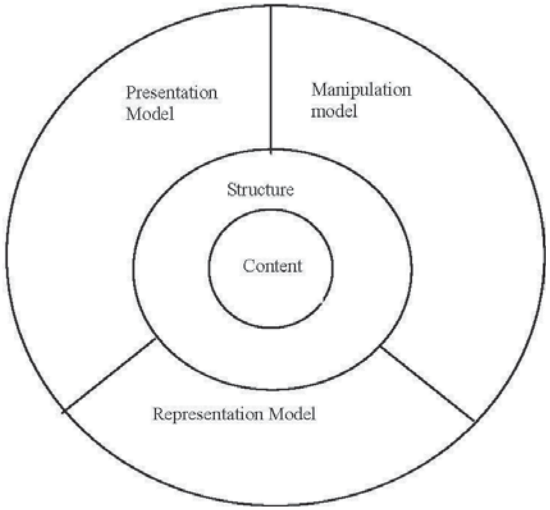


Figure: Document architecture and its elements

8.1. Document Architecture and Multimedia Integration

The ***manipulation model*** describes all the operations allowed for creation, change and deletion of multimedia information.

The ***representation model*** includes the relations between the individual information elements which need to be considered during presentation. It defines:

- *protocols* for exchanging information among different computers
- *formats* for storing the data

The ***presentation model*** describes how the content of the document is displayed or presented before the users. The Structure is described by the links and the synchronization parameters of the document. The content is the actual information that the document carries.

8.1. Document Architecture and Multimedia Integration

Manipulation of Multimedia Data

The manipulation tools for the multimedia document or data are the ***editors***, desktop publishing programs, text, image and audio processing programs.

The information included in the document classifies it into different document types like letters, articles, invitation etc. However the transformation from the actual information to its final representation behaves according to rules specific to the document architecture.

While processing the document exists in a process able representation. The subsequent formatting process determines the layout of the document. The result is a final representation of the document.

8.1. Document Architecture and Multimedia Integration

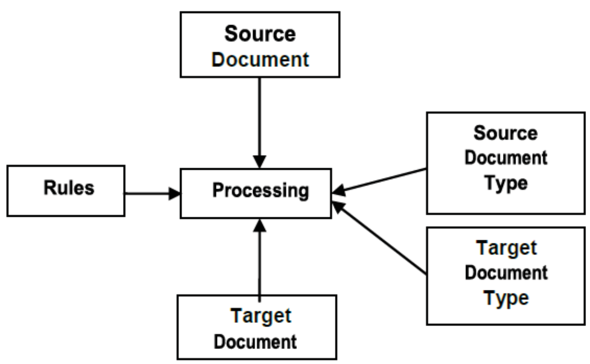


Figure: Processing of a document – from the information to the presentation

8.2. Hypertext, Hypermedia and Multimedia

Hypertext, Hypermedia and Multimedia

Ordinary documents are transformed into a linear document and the structural information is integrated into the actual content. In the case of hypertext and hypermedia, a graphical structure is possible in a document which may simplify the writing and reading processes.

A software system allowing extensive cross-referencing between related sections of text and associated graphic material called **Hypertext**.

An extension to hypertext providing multimedia facilities, such as those handling sound and video called **Hypermedia**.

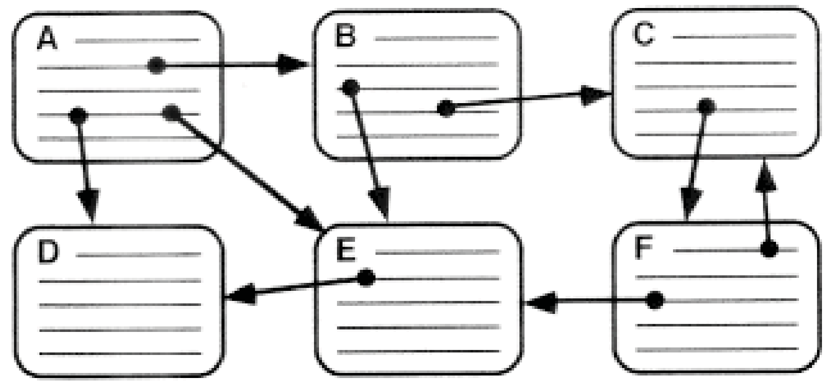
8.2. Hypertext, Hypermedia and Multimedia

Hypertext, Hypermedia and Multimedia

The reading of the document may not always be sequential. For e.g. when a reader is reading about how internet works, he might want to learn about the protocols which the sequential document does not contain in such case it is possible to provide the link to the key terms from the document to some other related documents.

Similarly when reading a biography the user may want to view the pictures of the person, or the videos related to him. It may not be possible to include that additional information in the same document but it is still possible to provide a link to the video or the audio.

Hypertext



Hypermedia

8.2. Hypertext, Hypermedia and Multimedia

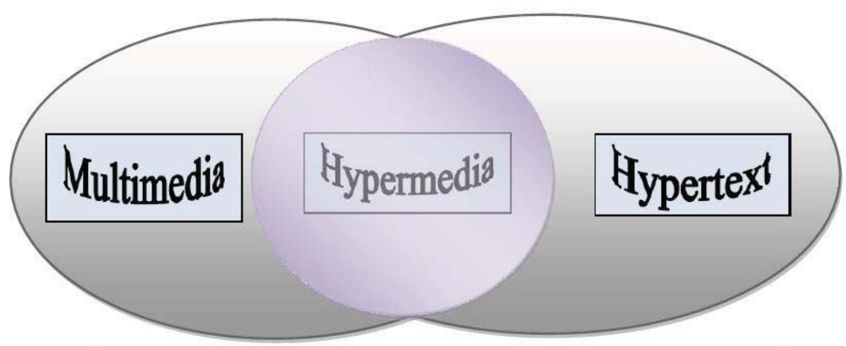


Figure: The hypertext, hypermedia and multimedia relationship

8.2. Hypertext, Hypermedia and Multimedia

- ***Hypertext*** can be a word, phrase, or a sentence that provides a non-linear link from a document to some other document. A *hypertext system* is mainly determined through non-linear links of information.
- ***Hypermedia*** on the other hand is nonlinear way of presenting information that allows users to access related works or images from a single computer screen. For example, a user reading an encyclopedia entry on jazz could also hear excerpts from recordings, read biographies of jazz artists, and view photos of them. Apple Computer Inc.'s
- **Hyper-card** is the best-known example of hypermedia. Presumably, this type of interface is similar to normal human cognitive processes. It is in fact the hypertext extended so that the link could be provided with the help of image or other data structure different than the text.
- A ***hypermedia*** *system* includes the non-linear information links of hypertext systems and the continuous and discrete media of multimedia systems.
- A ***multimedia*** *system* consists of information which is coded at least in a continuous and discrete medium.

8.2. Hypertext, Hypermedia and Multimedia

Non-Linear Information Chain

Hypertext and hypermedia have a major property a *non-linear information link.* There exists not only a reading sequence, but also the reader decides on his/her reading path. The reader can start from a page with a notion hypertext, then go through a cross reference to systems and finish with audio-visual information.

A hypertext structure is a graph, consisting of nodes and edges.

- The ***nodes*** are the actual information units. They are for example the text elements, individual graphics, audio or video LDUs. The information units are shown at the user interface mostly in their own windows.
- The ***edges*** provide links to other information units. They are usually called pointers or links. A pointer is mostly a directed edge and includes its own information too.

8.2. Hypertext, Hypermedia and Multimedia

Anchor

Exploring a non-linear document is *navigation.* The document is linked with other related document with the help of **pointers**.

This origin of a pointer is called an anchor. Representation of the anchor can be

- A ***media-independent*** representation can happen through the selection of general graphical elements, such as buttons.
- In a ***text***, individual words, paragraphs or sections of text can be used as pointer. These texts are visually different from rest of the document. The user is flown to the other document with a click to these texts.
- In ***images***, specific graphical objects or simply areas are defined as selection objects. A specific marking can occur through a color or stripe.
- In a ***motion*** *video*, media-independent representations of the anchor are preferred. A timely selection is supported.
- With respect to ***audio***, a media-independent solution is used for e.g. by using descriptive text or an image of the size of an icon.

8.2. Hypertext, Hypermedia and Multimedia

Application Areas of Hypertext, Hypermedia and Multimedia

- While giving lecture on hypermedia documents.
- In some classical computer applications like the “help” function
- In the area of commercial applications, repair and operational manuals can be found.
- The organization of ideas, brainstorming and the generation of scientific documents count, for example, as intellectual applications.
- Tourist information systems and interactive science-fiction movies count on the areas of entertainment and free-time activities.

8.3.Hypermedia system: Architecture, Nodes and Pointers

Architecture

- ***Presentation Layer*** At the upper layer, the presentation layer, all functions connected to the user interface are embedded. Here, nodes and pointers are mapped to the user interface.
- ***Hypertext Abstract Machine*** (**HAM**) is placed between the presentation and storage layers. It knows the structure of the document and it expects database related functions from the lower layers.
- ***Storage Layer*** It is the lowest layer. All functions connected with the storage of data, i.e., secondary storage management, belong to this layer.

8.3.Hypermedia system: Architecture, Nodes and Pointers

Nodes

A node is an information unit (LDU) in a hypertext document.

- The maximal stored *data amount* can be *limited* and mapped onto the screen size.
- Window-based systems with an *unlimited data amount* per node are the alternative. Forward and backward scrolling of pages is offered analogous to other windows at the user interface.

Pointers

Pointers are the **edges** of a hypertext graph. Pointers can be

- ***Simple pointers*** link two nodes of the graph without containing any further information.
- ***Typed pointers*** contain further information in addition to the link. Usually a tool tip text or a label is used with typed pointers.
- ***Implicit Pointers*** determine the relation between nodes automatically.
- ***Explicit Pointers*** is created by the author himself

8.3.Hypermedia system: Architecture, Nodes and Pointers

When is the destination of a pointer specified?

- In the classical case, the pointer is created during the generation of hypertext document, and hereby the origin of the destination node is determined. The author determines explicitly the links of the information units during document processing.
- A destination node can be determined first by using the pointer, i.e., during reading. The author specifies an algorithm for the creation of the pointers, but they are determined first during the reading depending on the context.

8.3.Hypermedia system: Architecture, Nodes and Pointers

Tools

The tools can be ***editors*** for processing the information.

Search tools allow the search of desired information.

Browser allows a shortened but clear representation of the nodes and edges. The nodes are described media-dependently.

Backtracking prevents the users from getting lost in the hyperspace.

8.4.Document Architecture: SGML and ODA

Document Architecture SGML

*Standard Generalized Markup Language (**SGML**)* is a standard for the uniformity in the content and their representation in the document.

The Content of the document is described within the tags.

SGML determines the form of **tags**, but it does not specify their location or meaning.

It is basically a set of rules that break a document into parts and identify the different parts of the document. It basically defines the syntax or the structure of the document's content. However SGML does not provide semantics.

8.4.Document Architecture: SGML and ODA

Example:

<BOOK>

<TITLE> Multimedia Communication </TITLE>

<AUTHOR> Jerry D. Gibson </AUTHOR>

<PRICE> NRs. 1000 </PRICE>

<PUBLICATION> PHI </PUBLICATION>

</BOOK>

8.4.Document Architecture: SGML and ODA

Processing of SGML document

Processing of SGML document is divided into two processes. Only the formatter knows the meaning of the tag and it transforms the document into a formatted document. The parser uses the tags, occurring in the document, in combination with the corresponding document type.

8.4.Document Architecture: SGML and ODA

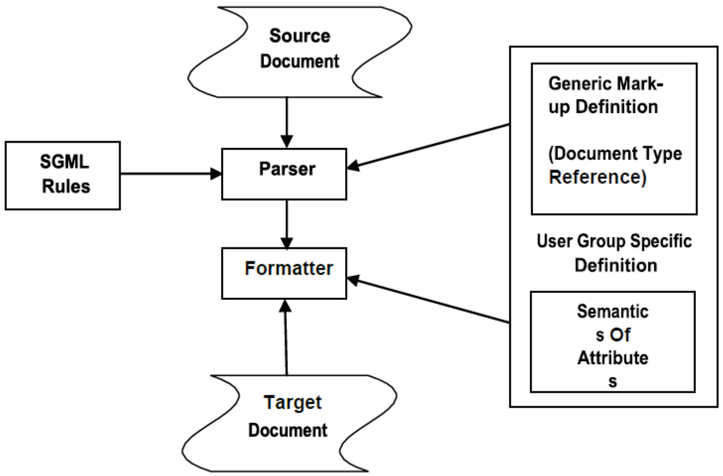


Figure: SGML Document processing - from the information to the presentation

8.4.Document Architecture: SGML and ODA

Tags are divided into different categories

- The ***descriptive markup*** *(tags)* describes the actual structure always in the form: <start-tag> information </end-tag>
- The ***entity reference*** provides connection to another element. This element replaces the entity reference. The following example shows entity reference in a mathematical context: &square x … should be x^2
- The ***markup declarations*** define the elements to which an entity reference refers. Squaring a variable x, square is defined as <!ELEMENT square (…)>
- Instructions for other programs in a text are entered through ***processing instructions***. Using processing instructions, different media can be inserted.

8.4.Document Architecture: SGML and ODA

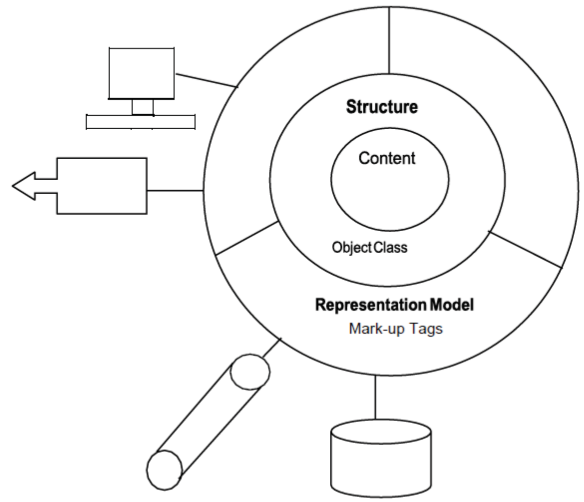


Figure: SGML: Document architecture - emphasis on the representation model

8.4.Document Architecture: SGML and ODA

Document Architecture ODA

The *Open Document Architecture (ODA)* is a standard document file format created by the ITU-T to replace all proprietary document file formats.

ODA defines a compound document format that can contain ***raw text, raster images and vector graphics***.

The documents have both logical and layout structures. Logically the text can be partitioned into chapters, footnotes and other sub elements, and the layout fills a function similar to Cascading Style Sheets in the web world.

8.4.Document Architecture: SGML and ODA

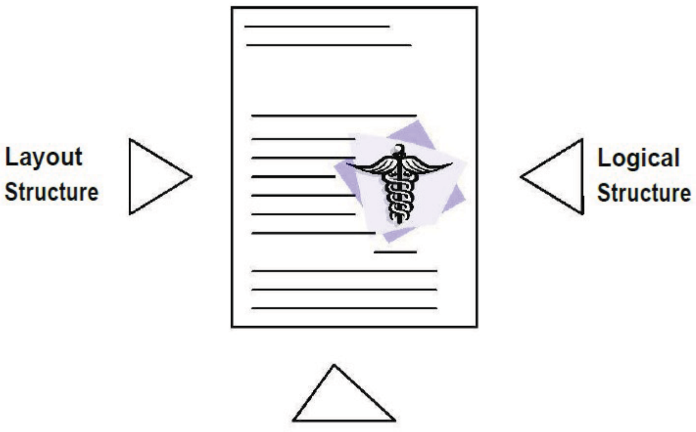


Figure: ODA: Content, layout and logical view

8.4.Document Architecture: SGML and ODA

The binary transport format for an ODA-conformant file is called Open Document Interchange Format and is based on abstract syntax notation one (ASN.1).

Thus the main property of ODA is the distinction among content, logical structure and layout structure. This is in contrast to SGML where only a logical structure and the contents are defined.

ODA also defines semantics. The following figure shows the layout, logical and content aspect of the linked documents. Each of these views represent one aspect, together we get the actual document.

8.4.Document Architecture: SGML and ODA

ODA and Multimedia

Multimedia requires, besides spatial representational dimensions, the *time* as the main part of a document. For the inclusion of continuous change is the standard is necessary which basically calls for an extension. The following changes will occur:

Contents

The *content portions* will change to *timed content portions.* The duration does not have to be specified earlier. These types of content portions are called *Open Timed Content Portions.* In case of a *Closed Timed Content Portion,* the duration is fixed. E.g. song

Structure

Operations between objects must be extended with a time dimension where the time relation is specified in the root node r in proportion to the child nodes c1, c2.

8.4.Document Architecture: SGML and ODA

ODA and Multimedia (cont..)

Content Architecture

Many functions are very often device-dependent. One of the most important aspects is a compatibility provision among different systems implementing ODA.

Logical Structures

Extensions for multimedia of the logical structure also need to be considered. For example, a film can include a logical structure.

Layout Structure

The layout structure needs extensions for multimedia. The time relation by a picture and audio must be included. Further, question such as *when will something be played?, From which point?,* And *With which attributes and dependencies?* must be answered.

8.5.MHEG

MHEG (Multimedia and Hypermedia Information Coding Expert Group)

Several cross platform video and audio standards have been established including still and motion JPEG, and a number of different MPEG standards.

So far, there has been no standard method of bringing all these formats together to produce multimedia presentations.

The MHEG model aims to solve this by providing a system independent presentation standard for hardware and software engineers and presentation authors to conform to.

In this way, a presentation created on one hardware platform should be viewable on others.

8.5.MHEG

What is MHEG?

MHEG is an abbreviation(contraction) for the Multimedia and Hypermedia Experts Group. This is another group of specialists, eminent in their field which has been set up by ISO, the International Standards Organization. This group was set the task of creating a standard method of storage, exchange and display of multimedia presentations.

8.5.MHEG

Its basic goals are:

- To provide a simple but useful, easy to implement framework for multimedia applications using the minimum system resources.
- To define a digital final form for presentations, this may be used for exchange of the presentations between different machines no matter what structure or platform.
- To provide extensibility i.e. the system should be expandable and customizable with additional application specific code, though this may make the presentation platform dependent.

8.5.MHEG

The Reasoning behind MHEG

In order for products to have maximum commercial success, they must appeal to the largest number of consumers possible.

There are various multimedia presentation packages available, but they are proprietary and do not work across different hardware platforms.

This means that an author, who wants to publish an interactive multimedia book for example, will have to produce several versions that comply with different standards, or risk losing potential customers.

The MHEG standard would do away with all this, allowing an author to produce his/her work in one universally acceptable format. It also has advantages for the hardware and software suppliers.

They are able to concentrate on producing the one standard MHEG engine rather than an engine for every presentation standard that is available.

8.5.MHEG

Applications of MHEG

The applications to which MHEG may be put are growing all the time, as people dream up more and more applications of multimedia. Here are a few examples:

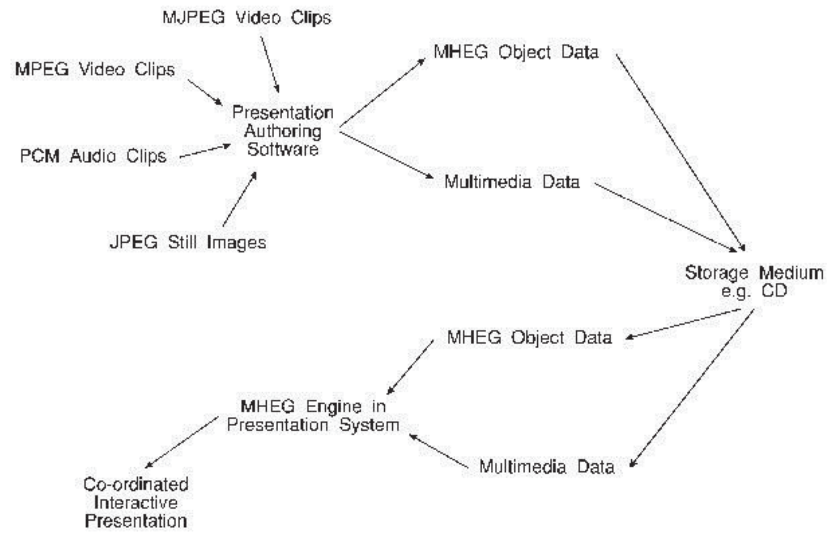
- CD-ROM based encyclopedias
- Interactive books for learning
- Video and news on demand systems
- Interactive home shopping

8.5.MHEG

Structure of MHEG

MHEG defines the abstract syntax through which presentations can be structured. This is the definition of data structures and the fields in those data structures, through which two computers may communicate. In the case of a multimedia book, these two computers would be that of the author and that of the user.

8.5.MHEG (Structure of MHEG)



8.5.MHEG (Structure of MHEG)

The MHEG model is object orientated, and defines a number of classes from which object instances are created when a presentation is designed. There are several classes, and these are used to describe the way video is displayed, audio is reproduced and how the user can interact with the ongoing presentation. The relationship that is created between instances of these classes‟ forms the structure of the presentation. In addition to just replaying existing multimedia data MHEG also defines some types of its own. e.g. an MHEG compliant system is able to overlay titles onto video scenes. It is also able to display menus and buttons to allow the user to make choices

8.5.MHEG

There are several different types of class defined by MHEG. Some are concerned with the structure of the presentation and grouping of objects, whilst others are involved in the interchange of information between machines.

8.5.MHEG (Class Hierarchy)

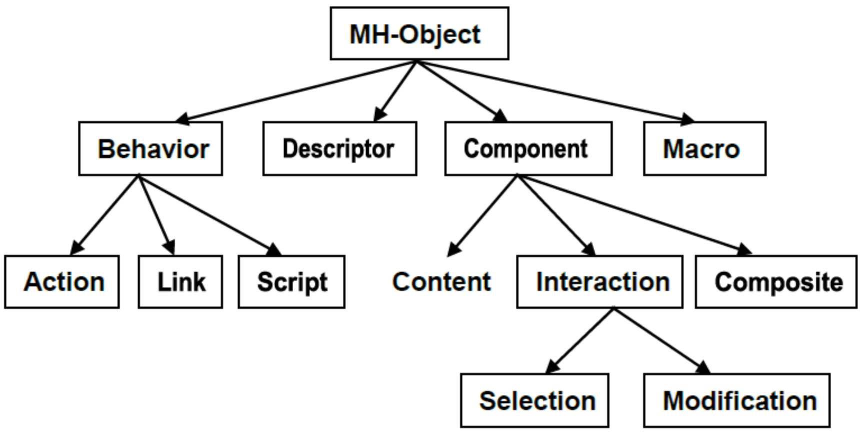


Figure: Class hierarchy of MHEG objects

8.5.MHEG (Class Hierarchy)

Content Class:

Content classes are each piece of multimedia data e.g. video or audio clip has its own MHEG object. If the amount of data involved is small (a textual title), then it may be contained in the MHEG object itself. If not, then the MHEG object will give a reference to the data e.g. a disk filename. Depending on the form of the data, extraction of smaller sections of the data may be possible e.g. one audio channel from an MPEG audio/video clip.

Behavior Class:

Classes under this heading control how and when data is presented to the user. They allow synchronizing of events and user interaction.

Action Class:

The action class allows events to be triggered sequentially or in parallel e.g. the replay of several clips of video one after the other.

8.5.MHEG (Class Hierarchy)

Link Class:

The link class establishes relationships between events and objects i.e. what actions to take on what objects in response to a particular event. A link object consists of a set of links.

Script Class:

It determines the behavior of the objects. It supports MHEG presentation in run-time environments.

Selection Class:

The selection class provides the possibility to model an interaction as a selection of a value from a predefined value set.

Modification Class:

It serves as the input and manipulation of data. No value set is predefined by a modification object.

Composite Class:

It has the task of composing all the necessary objects from the previously described classes into a presentation. If a single object is included or referenced, each composite object behaves as a *container*.

MULTIMEDIA COMMUNICATION SYSTEM

Unit 9

9.1. Definition of multimedia communication

- Multimedia applications have several requirements – as well for the data transmission as for controlling interactivity. For structuring control and transmission functionality and implement common protocols
- From the communication perspective, we can divide the higher layers of the Multimedia Communication System (MCS) into two architectural subsystems:
  - *Application Subsystem*
  - *Transport Subsystem*

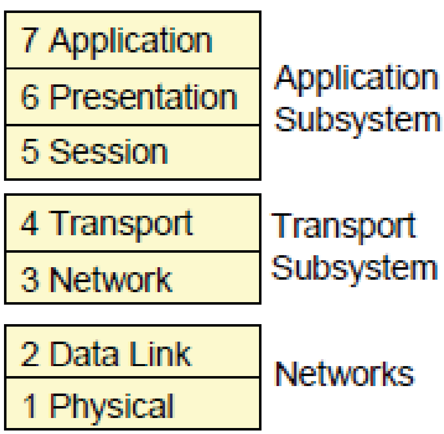
9.1. Definition of multimedia communication

Application Subsystem

This subsystem includes the software and the tools with which the end user directly interacts. E.g. of application can be email, video conferencing software etc.

- Responsible for the management and service issues for group cooperation and session orchestration
- Supporting a large scale of multimedia applications, e.g.
  - Multimedia Mail
  - Virtual Reality Applications
  - Video Conferencing
  - CSCW (Computer Supported Cooperative Work)

9.1. Definition of multimedia communication



9.1. Definition of multimedia communication

Transport Subsystem

- Transport and network layer protocols for multimedia applications (streaming)

9.2. Application Subsystem

- Streaming multimedia data often is used in *cooperative computing*
  - Beneath streaming services (transport subsystem), some control functionality for cooperation support is needed
  - Cooperative Computing is generally known as *Computer Supported Cooperative Work* (CSCW).
- Tools for cooperative computing:
  - Electronic mail
  - Shared whiteboards
  - Screen sharing tools
  - Application sharing
  - Text-based conferencing systems
  - Video conference systems (e.g. MBone Tools, ProShare from Intel, PictureTel,Teles Online, NetMeeting from Microsoft)

9.2. Application Subsystem

Cooperation Dimensions

Computer-Supported cooperation may be categorized according to the following parameters:

- *Time*
  - Asynchronous cooperative work (not at the same time)
  - Synchronous cooperative work (at the same time)
- *User Scale*
  - Single user, two users (“dialogue, point to point“, direct cooperation) or groups with more than two users
  - Static or dynamic groups, depending on if the members are pre-determined or not
- *Control*
  - Centralized, i.e. controlled by a “chairman”
  - Distributed, i.e. control protocols provide consistent cooperation
- *Locality*
  - Cooperation at the same place
  - Tele-cooperation of users at different places

9.2. Application Subsystem

- **Application Subsystem**

  This subsystem includes the software and the tools with which the end user directly interacts. E.g. of application can be email, video conferencing software etc.
- **Collaborative Computing**

  Collaborative computing is the computer supported cooperative work supported by the networks, PCs and the software that facilitates the cooperation. The examples of collaborative computing tools are *electronic mail, bulletin boards, screen sharing tools, text-based conferencing systems, telephone conference systems, conference rooms* and *video conference systems.*
- **Collaborative Dimensions** The collaboration dimensions are
  - *Time*
  - *User Scale*
  - *Control*

9.2. Application Subsystem

- ***Time:***

  According to time there can be two types of collaborative work and they are *asynchronous* where the cooperative works do not happen at the same time; while the other is the *synchronous* where the cooperative works happen at the same time.
- ***User Scale:***

  The user of the application can be a single user interacting to the other user or to a group of users. Email between two individuals is a user to user interaction where as Email to a group is a user to group interaction. Video conferencing is also a user to group interaction application.

9.2. Application Subsystem

There can be following types of groups.

- Dynamic or Static Group: When the new users can join the group for cooperative work in the real time the group is said to be dynamic where as when the number of members and their membership is predefined it is static group.
- The members of the group may simply be a *participant* of the collaborative work or he/she may be the *co-coordinator, conference initiator, conference chairman, a token holder* or just an *observer.*
- The members of the group may have *homogenous* or *heterogeneous* characteristics. For e.g. they may belong to different ethnic group or they may differ in the level of intelligence.

9.2. Application Subsystem

- Control:

  Control during the collaboration can be *centralized* or *distributed*. Centralized control means there is a chairman who controls the collaborative work and every group member reports to him. Distributed control means every group member has control over his/her own tasks in the collaborative works.

9.2. Application Subsystem

Group Communication Architecture

Group communication is a cooperative activity which may be synchronous or asynchronous which may be central control or a distributed control. A group communication architecture consists of:

- Support model
- System model
- Interface model

Group Communication Architecture

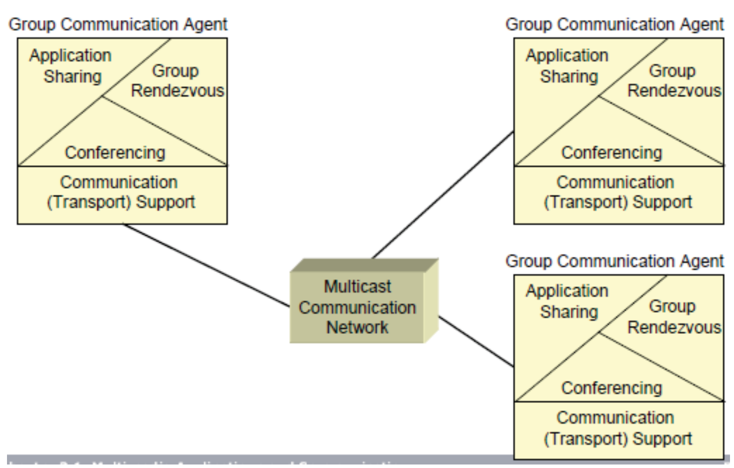


Figure: Communication Support Model

Group Communication Architecture

- Group communication:
  - Synchronous communication
  - Asynchronous communication
- Support Model: *Group communication agents* (cooperating via a multicast network)
  - *Group rendezvous* (organization of meetings and delivering information)
  - *Shared applications* (simultaneous replication and modification of information to multiple users, e.g. telepointing, joint editing)
  - *Conferencing* (audio/video*)*
- System Model:
  - Client/server model
- Interface Model:
  - Exchanging information within the support model (object oriented)

Group Communication Architecture

The Support Model:

It includes *group communication agents* that communicate via a multi-point multicast communication network as shown in the figure above. Group communication agents may use the following for their collaboration:

Group Rendezvous:

It represents methods which allows one to organize meetings, and to get information about the group, ongoing meetings and other static and dynamic information. The rendezvous can be Synchronous where the Directory services allows the access to information stored in a knowledge base about the conference, registered participants, authorized users and name and role of the participants.

Shared Applications:

It refers to the techniques which allow the replication of the information which can be delivered to all participants of the collaborative work. It may use the Centralized Architecture or Replicated Architecture. In the former a single copy of the application runs at one site say server.

The Support Model:

- **Conferencing**:

  It represents the service which manages the multiple users to communicate and interact with each other by the use of multimedia data. Thus conferencing is basically a management service that controls the communication among multiple users via multiple media, such as video and audio, to achieve simultaneous face-to-face communication.

Group Communication Architecture

- **The System Model:** It is based on a client-server model where the clients are applications that provides interface to the users who interact with the system while servers refers to function which makes it possible for the clients to communicate with each other and manage the communication.
- **The Interface Model:** It includes the *user presentation protocols* and *group work management protocols.* User presentation protocol is the interface available to the end users from which they can initiate, join, manage, communicate and terminate the conference. Group work management protocols specify the communication between the client and the servers for services like registration and querying the status of the conference.

Session Management

Session Management

A session is the total logged in time of a user or it can be the entire conference from its commencement to its termination. The management of session is a very important task and it should consider several issues like allowing users to join and leave the conferencing, selection of the coordinator, distributing information between users.

Architecture of Session Management It consists of following components:

- Session manager
- Media agent
- Shared workspace agent

Session Management

- *Session Manager:* It includes local and remote functionalities. The local functionalities includes the
  - Membership Control management: Authenticating the users or allowing members to join the session.
  - Control Management: It may involve floor management which involves the distribution and sharing of the information and resources available to the conference.
  - Media Control Management: It is required for the synchronization between the different media.
  - Configuration Management: It refers to the exchange and optimization of the QoS parameters.
  - Conference Control Management: It consists of functionalities for starting, changing and closing the conference.

Session Management

- ***Media Agents**:*

  These are responsible for decisions specific to each type of media. Each agent performs its own control mechanism over the particular medium, such as mute, unmute, change video quality, start sending, stop sending etc.

- ***Shared Workspace Agent**:*

  The shared workspace agent transmits shared objects (e.g., telepointer coordinate, graphical or textual object) among the shared applications.

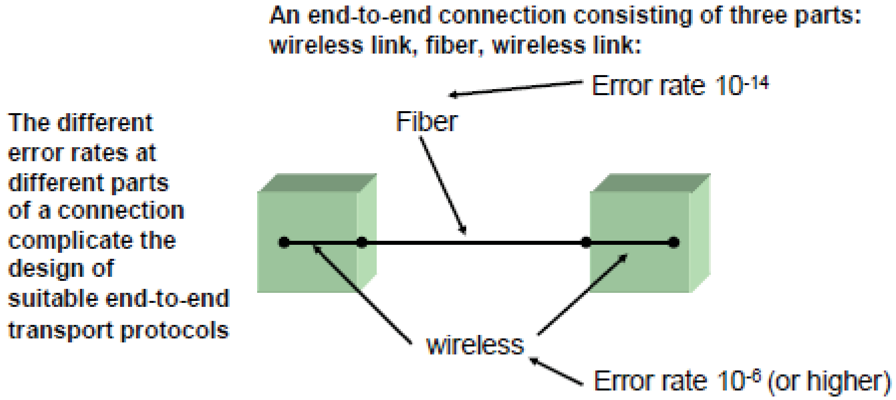
9.3. Transport Subsystem : Requirements, Transport Layer, Network Layer

Transport Subsystem

Multimedia applications have high requirements on network protocols:

1. High data throughput
   - Deliver as much data as possible in short time
2. Fast data forwarding
   - Deliver data as fast as possible
3. Service guarantees
   - Deliver data with regard to negotiated policies (throughput, delay)
4. Multicasting
   - 1:*n* and *m*:*n* point communication [*a*:*b* denotes: *a* senders and *b* receivers] *Needed: special transport resp. network protocols*

Most transport protocols are designed for unreliable and “relatively slow“ networks Modern high speed networks need ne transport protocols that take into account the properties of new network technology



Data Throughput (fast & effective)

- Audio and video data typically have a stream-like behavior
- Even in compressed mode, they demand high data throughput (16 kbit/s for compressed audio, 64 kbit/s for original PCM-audio, 2 Mbit/s for MPEG-coded video)
- In a workstation or in a network several of those streams may coexist
- Telephone services or video conferencing demand real-time computing of the data streams

This requires not only suitable transport protocols, but also high performance workstations which are able to compute several multimedia streams simultaneously and can transmit the packets in appropriate speed to the network interface

Service Guarantees / Multicasting

Service guarantees

- Service guarantees are important for the acceptance of multimedia applications
- Multimedia applications need guarantees, such as: throughput ≥ minvalue, delay ≤ maxvalue1, jitter ≤ maxvalue2
- To give service guarantees, resource management must be used – without this, in end-systems and switches/routers multimedia systems cannot provide reliable QoS to their users because transmission over unreserved resources may lead to dropping or delaying of packets

Service Guarantees / Multicasting

Multicasting

- Multicasting is important for multimedia applications in terms of sharing resources like the network bandwidth
- Many multimedia applications, such as video conferencing, have multicast characteristics

Transport Subsystem

Multimedia applications generate and consume a huge amount of data. They produce high demands for the underlying infrastructure and communication architecture:

- Data should be copied directly from adapter to adapter to reduce copy overhead, e.g. from the video board to the network interface by using **direct memory access** (**DMA**). With DMA technique, the application itself never really touches the data, it only sets the correct switches for the data flow by connecting sources to sinks
- The data transfer involved by the layered structure of the protocols form a bottleneck, hence other mechanisms must be found
- For error-recovery some protocols use retransmission techniques which impose requirements on buffer space for queues at the expense of larger end-to-end delays
- The synchronous behavior of most multimedia data streams must be mapped onto the asynchronous transfer mode of the underlying networks

Are the “old“ transport protocols like TCP and UDP suitable for multimedia transmission and, if not, which new protocols exist or have to be developed?

Transport Subsystem

Requirements

- **User and Application Requirements**

  Networked multimedia applications by themselves impose new requirements onto data handling in computing and communication because they need: *Sustainable data throughput, fast data forwarding, service guarantees, and multicasting.*
- *Data Throughput*

  This requirement wants the processing of the system to be **fast** and **effective**.
- *Fast Data Forwarding*

  The users or application wants **very low end to end delay** and jitter when communicating multimedia data. The holding time should be very less due to the real time requirement.

Transport Subsystem

- ***Service Guarantees****:* The loss of the information is undesired and the system or protocol used must ensure that the information is delivered to the intended destination.
- ***Multicasting*** *:* It is important for sharing the bandwidth and the communication protocol processing at end systems.
- **Processing and Protocol Constraints** Processing system and protocols have constraints which need to be considered while processing and transmitting multimedia information.
  - Following the “**shortest possible path**” for quicker delivery
  - Buffer management
  - Segmentation and reassembly
  - Re-transmission on error
  - Error-recovery
  - Asynchronous transfer

Transport Subsystem

Transport Layer

Transport protocols, to support multimedia transmission, need to have new features and provide the following functions:

- Timing information
- Semi-reliability
- Multicasting
- NAK (none-acknowledgement)-based error recovery mechanism
- Rate control

Internet Transport Protocols

- *TCP (Transmission Control Protocol)* It was designed to provide a **reliable end-to-end** byte stream over an unreliable inter-network.
- Each machine supporting TCP has a TCP transport entity, either a library procedure, a user process, or part of the kernel. In all cases, it manages TCP streams and interfaces to the IP layer.
- A TCP entity accepts user data streams from local processes, breaks them up into pieces not exceeding 64KB, and sends each piece as a separate IP datagram.

Internet Transport Protocols

- UDP (User Datagram Protocol) The Internet protocol suite supports a connectionless transport protocol, *UDP (User Datagram Protocol).*
- UDP provides a way for applications to send encapsulated **(summarized) IP datagrams and send them without having to establish a connection**.
- UDP transmits *segments* consisting of an 8-byte header followed by the payload

Real-time Transport Protocol (RTP)

- RTP is a **UDP protocol** used in the **client server** environment and in the real-time multimedia applications.
- The multimedia application consists of multiple audio, video, text, and possibly other streams.
- These are fed into the RTP library, which is in the user space along with the application. This library then multiplexes the streams and encodes them in RTP packets, which it then stuffs into a socket. At the other end of the socket, UDP packets are generated and embedded in IP packets.

Xpress Transport Protocol (XTP)

- XTP integrates transport and network protocol functionalities to have **more control over the environment** in which it operates. XTP is intended to be useful in a wide variety of environments, from real-time control systems to remote procedure calls in distributed operating systems and distributed databases to bulk data transfer.
- It defines for this purpose six service types: *connection, transaction, unacknowledged data gram, acknowledged datagram, isochronous stream* and *bulk data.*

Network Layer

- **Internet Protocol** In the TCP/IP protocol stack the network layer protocol is the *Internet Protocol (IP)* and, in order to transfer packets of information from one host to another, it is the IP in the two hosts, together with the IP in each Internet gateway and router involved that perform the routing and other harmonization functions necessary. The IP in each host has a unique Internet-wide address assigned to it.
- This is known as the host"s *Internet address* or, more usually, its *IP address.* Each IP address has two parts: a *network identifier* and a *host identifier.*

Routing

- Routing is used for guiding the packets from its source to the destination.
- Routers are dedicated for this purpose.
- Routing is based upon the congestion information, shortest path method.
- These routers may be administered by a common authority and are called *Autonomous Systems (AS)* for which protocol- *Interior Gateway Protocol (IGP).* ASs of gateways exchange reachability information by means of *Exterior Gateway Protocol (EGP).*

Internet Group Management Protocol (IGMP)

- Multicasting in Internet is done with the help of multicast routers.
- About once a minute, each multicast router sends a hardware multicast to the hosts on its LAN asking them to report back on the groups their processes currently belongs to.
- Each host sends back responses for all the class D addresses it is interested in.
- These query and response packets use a protocol called *IGMP (Internet Group Management Protocol)*.
- **It has only two kinds of packets: query and response**, each with a simple, fixed format containing some control information in the first word of the payload field and a class D address in the second word.

Resource Reservation Protocol (RSVP)

- In order to ensure that the real-time traffic flows does not exceed that which is allocated for it, the resources required for each flow are reserved in advance of each packet flow starting.
- **The resources can be bandwidth and buffer capacity.**
- The protocol used to do this is called *Resource Reservation Protocol.*
- Because many of the new real-time applications involve multiple participants, RSVP is used to reserve resources in each router along either a unicast or a multicast path.
- When making a reservation, a receiver can specify one or more sources that it wants to receive from. It can also specify whether these choices are fixed for the duration of the reservation or whether the receiver wants to keep open the option of changing sources later.
- The routers use this information to optimize bandwidth planning. Once a receiver has reserved bandwidth, it can switch to another source and keep that portion of the existing path that is valid for the new source.

9.4. Quality of Service(QoS) and Resource Management

- During a multimedia communication, the services in the multimedia systems need to be parameterized.
- Parameterization of the services is defined in ISO standards through the notion of Quality of Service (QoS).
- Each service can be characterized by a quality of service. As a simple example some services are reliable i.e. they do not loose data while some are unreliable as they may loose data.
- The parameters can be bandwidth, maximum and minimum end to end delay, jitter, buffer allocation etc. There are several issues that need to be addressed and they are:

9.4. Quality of Service and Resource Management

QoS Layering

The QoS requirement is associated with each layer of the OSI model, or the TCP/IP model. However the QoS for multimedia communication system (MCS) consists of three layers: application, system and devices.

Application means the software and the program parameters, where as system refers to the overall system of communication and then in the network.

QoS-layered model for the MCS

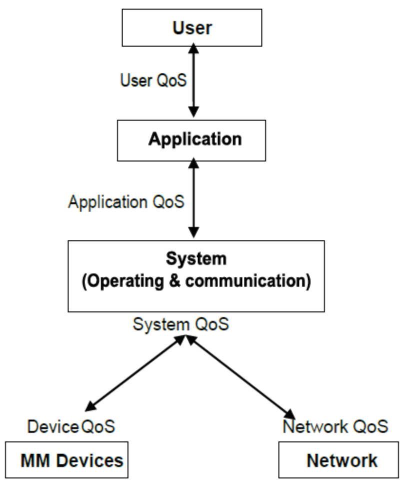


Figure: QoS-layered model for the MCS

QoS-layered model for the MCS

- **Service Objects**

  Services are performed on different objects, for example, media sources, media sinks, connections and Virtual Circuits (VCs), hence QoS parameterization species these *service objects.*

- **QoS Description**

  The QoS is described in terms of the required parameters by the end systems.

  The *application QoS* parameters may include media quality, transmission delay, jitter, synchronization.

  The system QoS parameters describe requirements on the communication services and OS services resulting from the application QoS. They may be specified in terms of both *quantitative* (bandwidth, PDU size, buffer size) and *qualitative criteria* (level of synchronization, order of data delivery, recovery).

QoS-layered model for the MCS

- *The network* QoS parameters describe requirements on network services. They may be specified in terms of network load (packet size, service time) and network performance (congestion, delay).
- *The device QoS parameters* typically specify timing and throughput demands for media data units.

QoS Parameter Values and Types of Service

There are three major type of service and they are:

- *Guaranteed Services*
- *Predictive Services*
- *Best-effort services*

In the **Guaranteed services** the QoS parameter values are deterministic or statistical in nature. It may ensure the lossless transmission of data i.e. reliable transmission.

In the **Predictive Services** the QoS parameter are predictable with the help of the past parameters. Though the exact value of the parameters may not be known a rough estimate can be made.

In the **Best-effort services**, the QoS parameters depend on the load of the network. It ensures that the best possible service is provided to the multimedia data.

9.4. Quality of Service and Resource Management

Resource Resource is a system used for processing, storing, manipulating data.

- The resource can be ***active* and *passive***. The active resource can be a CPU which processes data or manipulates data where as the passive resource is a bandwidth which only serves a particular purpose.
- A resource can be *exclusive* i.e. used by a single process or it may be *shared* where it is shared between various processes.
- The resource may be *single* or it may be *multiple.*

Resource Management Architecture

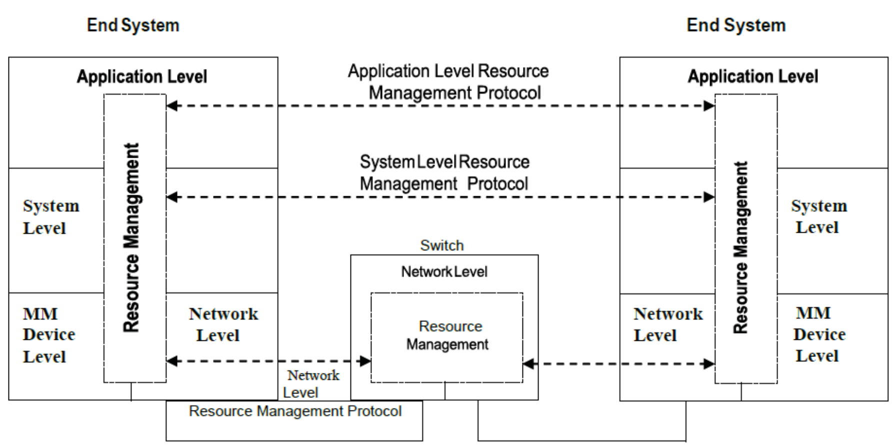


Figure: Resource Management in MCSs

Resource Management Architecture

The main goal of resource management is to offer guaranteed quality of service. It addresses three main actions

- Reserve and allocate resource
- Provide resources according to QoS specification
- Adapt to resource changes during on-going multimedia data processing.

Relation between QoS and Resources

The QoS parameters and their corresponding resources are mapped for the management of resources. For e.g. the end to end delay QoS parameter determines the behavior of transmission services along the path between source and sink with respect to packet scheduling, queuing and task scheduling. Description of a possible realization of resource allocation and management shows the QoS and resource relation.

QoS Negotiation (Cooperation)

- *Bilateral Peer-to-Peer Negotiation:* Negotiation occurs between the two service users and the service provider is not involved.
- *Bilateral Layer-to-Layer Negotiation:* It occurs between the service user and the service provider.
- *Unilateral Negotiation:* The user and the provider cannot modify the QoS parameter. It is based on “take it or leave it” model.
- *Hybrid Negotiation:* The negotiation between host and sender is bilateral layer-to-layer negotiation and negotiation between network and host-receiver is unilateral.

QoS Negotiation

- *Triangular Negotiation for Information: Exchange* The user specifies its required QoS parameters while the provider may change it according to the possibility and availability before confirming that the caller agrees upon it.
- *Triangular Negotiation for a Bounded Target:* It is similar to the above method but in this type of negotiation the caller specifies both the target value and the minimum required value. If the parameter that the provider does not meet the minimum required value the caller rejects the provider.
- *Triangular Negotiation for a Contractual Value:* In this case, the QoS parameters are specified through a minimal requested value and bound of strengthening. The goal of this negotiation is to agree on a contractual value, which in this case is the minimal request QoS parameter value.