SUBJECT NAME:MULTIMEDIA AND ANIMATION

SUBJECT CODE: DECP 35A

UNIT 4

COMPRESSION AND DECOMPRESSION

Compression is the way of making files to take up less space. In multimedia systems, in order to manage large multimedia data objects efficiently, these data objects need to be compressed to reduce the file size for storage of these objects.

Compression tries to eliminate redundancies in the pattern of data.

For example, if a black pixel is followed by 20 white pixels, there is no need to store all 20 white pixels. A coding mechanism can be used so that only the count of the white pixels is stored. Once such redundancies are removed, the data object requires less time for transmission over a network. This in turn significantly reduces storage and transmission costs.

TYPES OF COMPRESSION

Compression and decompression techniques are utilized for a number of applications, such as facsimile system, printer systems, document storage and retrieval systems, video teleconferencing systems, and electronic multimedia messaging systems. An important standardization of compression algorithm was achieved by the CCITT when it specified Group 2 compression for facsimile system. .

When information is compressed, the redundancies are removed.

Sometimes removing redundancies is not sufficient to reduce the size of the data object to manageable levels. In such cases, some real information is also removed. The primary criterion is that removal of the real information should not perceptly affect the quality of the result. In the case of video, compression causes some information to be lost; some information at a delete level is considered not essential for a reasonable reproduction of the scene. This type of compression is called lossy compression. Audio compression, on the other hand, is not lossy. It is called lossless compression.

Lossless Compression.

In lossless compression, data is not altered or lost in the process of compression or decompression. Decompression generates an exact replica ofthe original object. Text compression is a good example of lossless compression. The repetitive nature of text, sound and graphic images allows replacement of repeated strings of characters or bits by codes. Lossless compression techniques are good for text data and for repetitive data in images all like binary images and gray-scale images.

Some of the commonly accepted lossless standards are given below:

- **Packpits encoding (Run-length encoding)**
- **CCITT Group 3 I D**
- **CCITT Group 3 2D**
- **CCITT Group 4**
- **Lempe l-Ziv and Welch algorithm LZW.**

Lossy compression is that some loss would occur while compressing information objects.

Lossy compression is used for compressing audio, gray-scale or color images, and video objects in which absolute data accuracy is not necessary.

The idea behind the lossy compression is that, the human eye fills in the missing information in the case of video.

But, an important consideration is how much information can be lost so that the result should not affect. For example, in a grayscale image, if several bits are missing, the information is still perceived in an acceptable manner as the eye fills in the gaps in the shading gradient.

Lossy compression is applicable in medical screening systems, video tele-conferencing, and multimedia electronic messaging systems.

Lossy compressions techniques can be used alone o~' in combination with other compression methods in a multimedia object consisting of audio, color images, and video as well as other specialized data types.

The following lists some of the lossy compression mechanisms:

- ü Joint Photographic Experts Group (JPEG)
- ü Moving Picture Experts Group (MPEG)
- ü Intel DVI
- ü CCITT H.261 (P * 24) Video Coding Algorithm
- ü Fractals.

Compression schemes are

Binary Image compression schemes

Binary Image Compression Scheme is a scheme by which a binary image containing black and white pixel is generated when a document is scanned in a binary mode.

The schemes are used primarily for documents that do not contain any continuous-tone information or where the continuous-tone information can be captured in a black and white mode to serve the desired purpose.

The schemes are applicable in office/business documents, handwritten text, line graphics, engineering drawings, and so on. Let us view the scanning process.A scanner scans a document as sequential scan lines, starting from the top of the page.

A scan line is complete line of pixels, of height equal to one pixel, running across the page. It scans the first line of pixels (Scan Line), then scans second "line, and works its way up to the last scan line of the page. Each scan line is scanned from left to right of the page generating black and white pixels for that scan line.

This uncompressed image consists of a single bit per pixel containing black and white pixels. Binary 1 represents a black pixel, binary 0 a white pixel. Several schemes have been standardized and used to achieve various levels of compressions. Let us review the more commonly used schemes.

Most people have downloaded large files, such as music or video, from the Internet. Because of the large size of these files, downloading them can take hours. To solve this problem, and make better use of disk space, large files are compressed, using various software. Once downloaded, they can then be decompressed, and viewed, using a decompression program.

How Compression Works

Compression software works by using mathematical equations to scan file data and look for repeating patterns. The software then replaces these repeating patterns with smaller pieces of data, or code, that take up less room. Once the compression software has identified a repeating pattern, it replaces that pattern with a smaller code that also shows the locations of the pattern. For example, in a picture, compression software replaces every instance of the color red with a code for red that also indicates everywhere in the picture red occurs.

Types of Compression

Compressed files usually end with .zip, .sit and .tar. These are called extensions, and they indicate different compression formats--different types of software used to compress files. For PCs, .zip is most common, .sit is used often with Macs and .tar used with Linux. When you see a file with one of these extensions, it may be either a single large file or a group of files bundled together.

Lossless Compression

Lossless compression is a way to compress files without losing any data. This method shoves the data closer together by replacing it with a type of shorthand. It can reduce file sizes by around half. The .zip format uses lossless compression. With this form, the file decompresss to provide an exact duplicate of the compressed file, with the same quality. However, it cannot compress files to a really small size, making it less useful for very large files.

Lossy Compression

To make files up to 80 percent smaller, lossy compression is used. Lossy compression software removes some redundant data from a file. Because data is removed, the quality of the decompressed file is less than the original. This method compresses graphic, audio and video files, and the slight damage to quality may not very noticeable. JPEG uses lossy compression, which is why files converted to JPEG lose some quality. MP3 also uses lossy compression to fit a great deal of music files in a small space, although the sound quality is lower than with WAV, which uses lossless compression.

Decompression

In order to use a compressed file, you must first decompress it. The software used to decompress depends on how the file was compressed in the first place. To decompress a .zip file you need software, such as WinZip. To decompress a .sit file, you need the Stuffit Expander program. WinZip does not decompress .sit files, but one version of StuffIt Expander can decompress both .zip and .sit files. Files ending in .sea or .exe are called self-extracting files. These are compressed files that do not require any special software to decompress. Just click on the file and it will automatically decompress and open.

What Does Decompression Mean?

Decompression is the process of restoring compressed data to its original form. Data decompression is required in almost all cases of compressed data, including lossy and lossless compression. Similar to compression of data, decompression of data is also based on different algorithms. Decompression is considered important, as compressed data needs to be restored back to standard state for usage. Decompression is widely used in data communications, multimedia, audio, video and file transmissions.

Decompression is also known as uncompression.

Compression of data is advantageous as it helps in the reduction of storage space, transmission capacity or usage of resources. Decompression is thus important for compressed data as all compressed data needs to be decompressed. The application needed for decompression largely depends on how the data was compressed in the first place. There are different techniques and algorithms available for decompression of data. In most cases, the software or application needed for data decompression also comes with the application or software used for data compression. For each compression technique, there exists a corresponding decompression technique as well. Certain compressed files, such as those ending in ".exe" and ".sea" are classified as self-extracting files. These files do not need any special application for decompression, as it is initiated automatically once the file is clicked or run. Most decompression software makes use of a decoder, which aids in the data decompression process.

In the case of lossless compression, the original data is obtained without any loss on decompression. This is not the case with decompression for lossy methods, as there could be loss in the original data, although it may not impact the receiver. Data decompression helps in removing the complications added by data compression. Like compression, decompression can also be slow and time consuming at times and errors in transmission are not uncommon during data decompression.

File Formats for multimedia

The following is an outline of current file formats used for the production and delivery of multimedia data.

Text Formats

RTF

Rich Text Format is the primary file format introduced in 1987 by Microsoft with the specification of their published products and for cross-platform documents interchange.

Plain text

Plain text files can be opened, read, and edited with most text editors. commonly used are Notepad (Windows), Gedit or nano (Unix, Linux), TextEdit (Mac OS and so on. Other computer programs are also capable of reading and importing plain text. Plain text is the original and popular way of conveying an e-mail.

Image Formats

TIFF (Tagged Image File Format)

This format is common in desktop publishing world (high quality output), and is supported by almost all software packages. Recent versions of TIFF allows image compression, and the format is comfortable for moving large files between computers.

BMP (Bitmap)

Initially this format is in use with Windows 3.1. It is quite large and uncompressed and hence BMP is used for the high-resolution or large images.

DIB (Device Independent Bitmap)

This format which is similar to BMP, allows the files to be displayed on a variety of devices.

GIF (Graphics Interchange Format)

GIF is a compressed image format. Most of the computer color images and backgrounds are GIF files. This file format is best suitable for graphics that uses only limited colors, and it is the most popular format used for online color photos. 13-bit Color look up table is used by the GIF format to identify its color values. This format is supported widely.

JPEG (Joint Photographic Experts Group)

JPEG was designed to attain maximum image compression. It uses lossy compression technique, where a compression method is referred that loses some of the data required for the image reconstruction. It works good with photographs, naturalistic artwork, and similar material but functions less on lettering, live drawings or simple cartoons.

TGA (Tagra)

It is the first popular format for high-resolution images. TGA is supported by Most of the video-capture boards.

PNG (Portable Network Graphics)

An extensible file format for the less loss, portable and well compressed storage of raster images. PNG acts as replacement for GIF and also replaces multiple common uses of TIFF. PNG works good with online viewing applications like worldwide web. so it is fully streameable with a best display option.

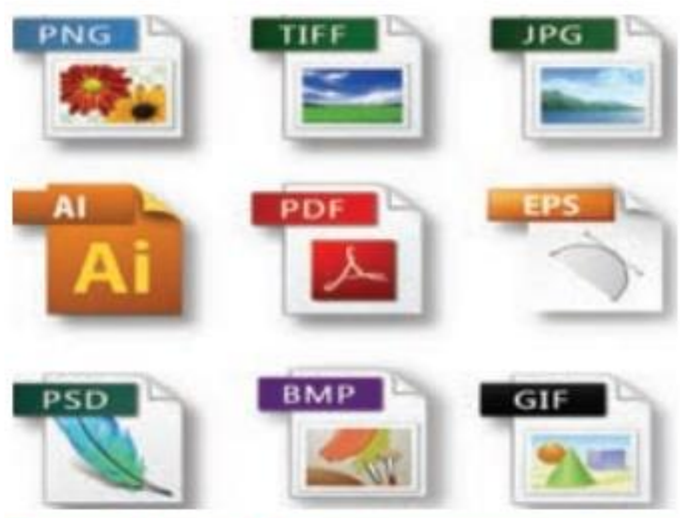


Figure 4.9 Image file Formats

Digital Audio File Formats

WAV (Waveform Audio File Format)

It is the most popular audio file format in windows for storing uncompressed sound files. In order to attain the reduced file size it can also be converted to other file formats like MP3.

MP3 (MPEG Layer-3 Format)

MPEG Layer-3 format is the most popular format for storing and downloading music. The MP3 files are roughly compressed to one-tenth the size of an equivalent WAV file.

OGG

A free, open source container format that is designed for obtaining better streaming and evolving at high end quality digital multimedia. It can be compared to MP3 files in terms of quality.

AIFF (Audio Interchange File Format)

A standard audio file format used by Apple which is like a WAV file for the Mac.

WMA (Windows Media Audio)

It is a popular windows media audio format owned by Microsoft and designed with Digital Right Management (DRM) abilities for copyright protection.

RA (Real Audio Format)

Real Audio format is designed for streaming audio over the Internet. The digital audio resources are usually stored as a computer file in computer's hard drive or CD/DVD. Besides the variety of audio file formats available, the most common formats are wave files (.WAV) and MPEG Layer-3 files (.MP3), WMA and RA.

Figure 4.10 Digital Audio File Formats

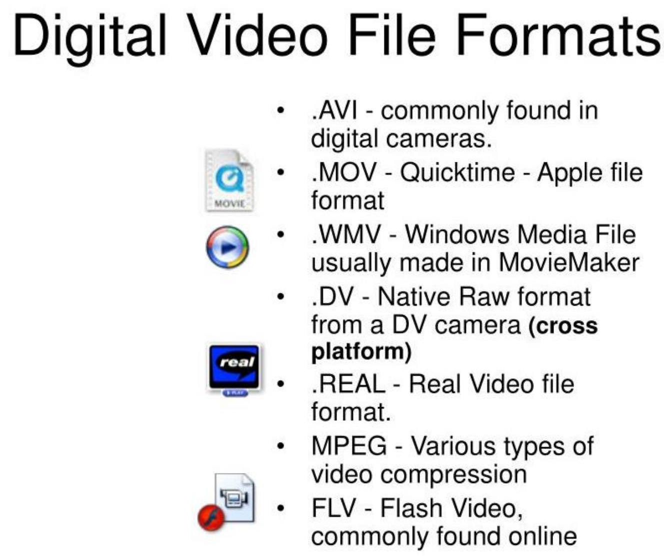
Digital Video File Formats

AVI (Audio/Video Interleave)

AVI is the video file format for Windows. Here sound and picture elements are stored in alternate interleaved chunks in the file.

MPEG (Moving Picture Experts Group)

MPEG is a standard for generating digital video and audio compression under the International Standards Organization (ISO) by the group of people. The group has developed MPEG-1, the standard on which Video CD and MP3 are based, MPEG-2, the standard that supports products as Digital Television set top boxes and DVD, MPEG-4, the standard for multimedia and mobile web.MPEG-7, the standard for search of audio and visual content. Research on MPEG-21 “Multimedia Framework” has started in 2000. Simply MPEG is the standards for digital video and audio compression.



MULTIMEDIA INPUT/OUTPUT TECHNOLOGIES

Multimedia Input and Output Devices

Wide ranges of Input and output devices are available for multimedia.

Image Scanners: Image scanners are the scanners by which documents or a manufactured part are scanned. The scanner acts as the camera eye and take a photograph of the document, creating an unaltered electronic pixel representation of the original.

Sound and Voice: When voice or music is captured by a microphone, it generates an electrical signal. This electrical signal has analog sinusoidal waveforms. To digitize, this signal is converted into digital voice using an analog-to-digital converter.

Full-Motion Video: It is the most important and most complex component of Multimedia System. Video Cameras are the primary source of input for full-motion video.

. **Pen Driver**: It is a pen device driver that interacts with the digitizer to receive all digitized information about the pen location and builds pen packets for the recognition context manager. Recognition context manager: It is the main part of the pen system. It is responsible for co-ordinating windows pen applications with the pen. It works with Recognizer, dictionary, and display driver to recognize and display pen drawn objects.

Recognizor: It recognizes hand written characters and converts them to ASCII.

Dictionary: A dictionary is a dynamic link library (DLL); The windows form pen computing system uses this dictionary to validate the recognition results.

Display Driver: It interacts with the graphics device interface' and display hardware. When a user starts writing or drawing, the display driver paints the ink trace on the screen.

Video and Image Display Systems Display System Technologies

There are variety of display system technologies employed for decoding compressed data for displaying. Mixing and scaling technology: For VGA screen, these technologies are used.

VGA mixing: Images from multiple sources are mixed in the image acquisition memory.

VGA mixing with scaling: Scalar ICs are used to sizing and positioning of images in predefined windows.

Dual buffered VGA mixing/Scaling: If we provide dual buffering, the original image is prevented from loss. In this technology, a separate buffer is used to maintain the original image.

Visual Display Technology Standards

MDA: Monochrome Display Adapter.

It was introduced by IBM .

:. It displays 80 x 25 rows and columns .
:. It could not display bitmap graphics .

:. It was introduced in 1981.

CGA: Color Graphics Adapter .

:. It was introduced in 1981.

.:. It was designed to display both text and bitmap graphicsi

it supported RGB color display,

.:. It could display text at a resolution of 640 x 200 pixels .

:. It displays both 40 x 25 and 80 x 25 row!' and columns of text characters.

MGA: Monochrome Gr.aphics Adapter .

:. It was introduced in 1982 .

:. It could display both text and graphics .

:. It could display at a resolution 720 x 350 for text and

720 x 338 for Graphics . MDA is compatible mode for this standard.

EGA: Enhanced Graphics Adapter .

:. It was introduced in 1984 .
:. It emulated both *MDt.* and CGA standards .

:. It allowed the display of both text and graphics in 16 colors at a resolution of 640 x· 350 pixels.

PGA: Professional Graphics Adapter.

.:. It was introduced in 1985 .

:. It could display bit map graphics at 640 x 480 resolution and 256 colors .
:.Compatible mode of this standard is CGA.
VGA: Video Graphics Array . :. It was introduced by IBM in 1988 .

:. It offers CGA and EGA compatibility .
:. It display both text and graphics .

:. It generates analog RGB signals to display 256 colors .

:. It remains the basic standard for most video display systems.

SVGA: Super Video Graphics Adapter. It is developed by VESA (Video Electronics Standard Association) . It's goal is to display with higher resolution than the VGA

with higher refresh rates with minimize flicker.

XGA: Extended Graphics Array

It is developed by IBM . It offers VGA compatible mode . Resolution of 1024 x 768 pixels in 256 colors is offered by it. XGA utilizes an interlace scheme for refresh rates.

Flat Panel Display system

Flat panel displays use a fluorescent tube for backlighting to give the display a sufficient level of brightness. The four basic technologies used for flat panel display are:

1. Passive-matrix monochrome
2. Active-matrix monochrome
3. Passive-matrix color
4. Active-matrix color.

LCD (Liquid Crystal Display)

Construction: Two glass plates each containing a light polarizer at right angles to the other plate, sandwich the nematic (thread like) liquid crystal material.

Liquid crystal is the compounds having a crystalline arrangement of molecules. But it flow like a liquId. Nematic liquid crystal compounds are tend to keep the long axes of rod-shaped molecules aligned. Rows of horizontal transparent conductors are built into one glass plate, and columns of vertical conductors are put into the other plate. The intersection of two conductors defines a pixel position.

Passive Matrix LCD

Working: Normally, the molecules are aligned in the 'ON' state.

Polarized light passing through the materials is twisted so that it will pass through the opposite polarizer. The light is then reflected back to the viewer. To turn off the pixel, we have to apply a voltage to the two intersecting conductors to align molecules so that the light is not twisted.

ACTIVE Matrix LCD

In this device, a transistor is placed at each pixel position, using thin-film transisor technology.

The transistors are used to control the voltage at pixel locations and to prevent charge from gradually leaking out of the liquid crystal cells.

PRINT OUTPUT TECHNOLOGIES

There are various printing technologies available namely Dot matrix, inkjet, laser print server and ink jet color. But, laser printing technology is the most common for multimedia systems.

To explain this technology, let us take Hewlett Packard Laser jet-III laser printer as an example. The basic components of the laser printer are

.:. Paper feed mechanism **.:.** Paper guide **.:.** Laser assembly **.:.** Fuser **.:.** Toner cartridge.

Working: The paper feed mechanism moves the paper from a paper tray through the paper path in the printer. The paper passes over a set of corona wires that induce a change in the paper .

The charged paper passes over a drum coated with fine-grain carbon (toner), and the toner attaches itself to the paper as a thin film of carbon .The paper is then struck by a scanning laser beam that follows the pattern of the text on graphics to be printed . The carbon particles attach themselves to the pixels traced by the laser beam . The fuser assembly then binds the carbon particles to the paper.

Role of Software in the printing mechanism:

The software package sends information to the printer to select and control printing features . Printer drivers (files) are controlling the actual operation of the printer and allow the application

software to access the features ofthe printer.

IMAGE SCANNERS

In a document imaging system, documents are scanned using a scanner. \The document being scanned is placed on the scanner bed or fed into the sheet feeder of the scanner .The scanner acts as the camera eye and takes a photograph of the document, creating an image of the original. The pixel representation (image) is recreated by the display software to render the image of the original document on screen or to print a copy of it.

Types of Scanners

A and B size Scanners, large form factor scanners, flat bed scanners, Rotory drum scanners and hand held scanners are the examples of scanners.

Charge-Coupled Devices All scanners use charge-coupled devices as their photosensors. CCDs consists of cells arranged in a fixed array on a small square or rectangular solid state surface. Light source moves across a document. The intensity of the light reflected by the mirror charges those cells. The amount of charge is depending upon intensity of the reflected light, which depends on the pixel shade in the document.

Image Enhancement Techniques

HalfTones In a half-tone process, patterns of dots used to build .scanned or printed image create the illusion of continuous shades of gray or continuous shades of color. Hence only limited number of shades are created. This process is implemented in news paper printers.

But in black and white photograph or color photograph, almost infinite levels of tones are used.

Dithering

Dithering is a process in which group of pixels in different patterns are used to approximate halftone patterns by the scanners. It is used in scanning original black and white photographs.

Image enhancement techniques includes controls of brightness, deskew (Automatically corrects page alignment), contrast, sharpening, emphasis and cleaning up blacknoise dots by software.

Image Manipulation

It includes scaling, cropping and rotation.

Scaling: Scaling can be up or down, the scaling software is available to reduce or enlarge. This software uses algorithms.

Cropping: To remove some parts of the image and to put the rest of the image as the subset of the old image.

Rotation: Image could be rotated at any degree for displaying it in different angles.

DIGITAL VOICE AND AUDIO

Digital Audio

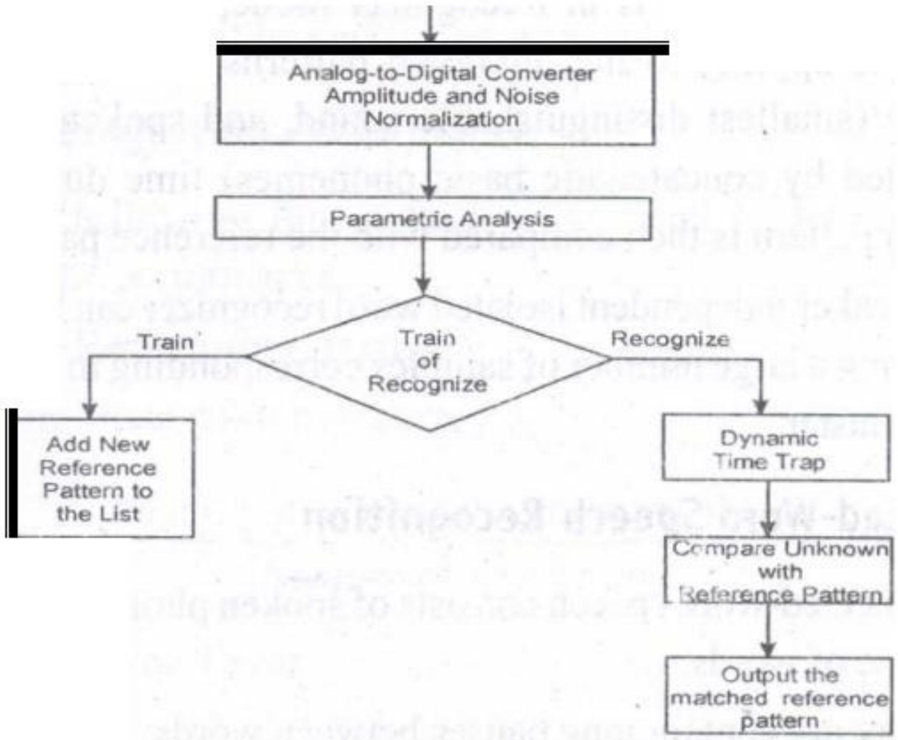
Sound is made up of continuous analog sine waves that tend to repeat depending on the music or voice. The analog waveforms are converted into digital fornlat by analog-to-digital converter (ADC) using sampling process.

Sampling process

Sampling is a process where the analog signal is sampled over time at regular intervals to obtain the amplitude of the analog signal at the sampling time.

Sampling rate

The regular interval at which the sampling occurs is called the sampling rate.



Digital Voice

Speech is analog in nature and is cOl1veli to digital form by an analog-to-digital converter (ADC). An ADC takes an input signal from a microphone and converts the amplitude of the sampled analog signal to an 8, 16 or 32 bit digital value.

The four important factors governing the

ADC process are **sampling rate resolution linearity and conversion speed.**

Sampling Rate: The rate at which the ADC takes a sample of an analog signal. **Resolution:** The number of bits utilized for conversion determines the resolution of ADC.

Linearity: Linearity implies that the sampling is linear at all frequencies and that the amplitude tmly represents the signal.

Conversion Speed: It is a speed of ADC to convert the analog signal into Digital signals. It must be fast enough.

VOICE Recognition System

Voice Recognition Systems can be classified into three types. 1.Isolated-word Speech Recognition.

2.Connected-word Speech Recognition.

3.Continuous Speech Recognition.

1. Isolated-word Speech Recognition.

It provides recognition of a single word at a time. The user must separate every word by a pause. The pause marks the end of one word and the beginning of the next word.

Stage 1: Normalization

The recognizer's first task is to carry out amplitude and noise normalization to minimize the variation in speech due to ambient noise, the speaker's voice, the speaker's distance from and position relative to the microphone, and the speaker's breath noise.

Stage2: Parametric Analysis

It is a preprocessing stage that extracts relevent time-varying sequences of speech parameters. This stage serves two purposes: (i) It extracts time-varying speech parameters. (ii) It reduces the amount of data of extracting the relevant speech parameters.

Training mode In training mode of the recognizer, the new frames are added to the reference list. **Recognizer mode**If the recognizer is in Recognizer mode, then dynamic time warping is applied to the unknown patterns to average out the phoneme (smallest distinguishable sound, and spoken words are constructed by concatenatic basic phonemes) time duration. The unknown pattern is then compared with the reference patterns.

A speaker independent isolated word recognizer can be achieved by groupi.ng a large number of samples corresponding to a word into a single cluster.

2Connected-Word Speech Recognition Connected-word speech consists of spoken phrase consisting of a sequence of words. It may not contain long pauses between words.

The method using Word Spotting technique

It Recognizes words in a connected-word phrase. In this technique, Recognition is carried out by compensating for rate of speech variations by the process called dynamic time warping (this process is used to expand or compress the time duration of the word), and sliding the adjusted connected-word phrase representation in time past a stored word template for a likely match.

Continuous Speech Recognition

This sytem can be divided into three sections:

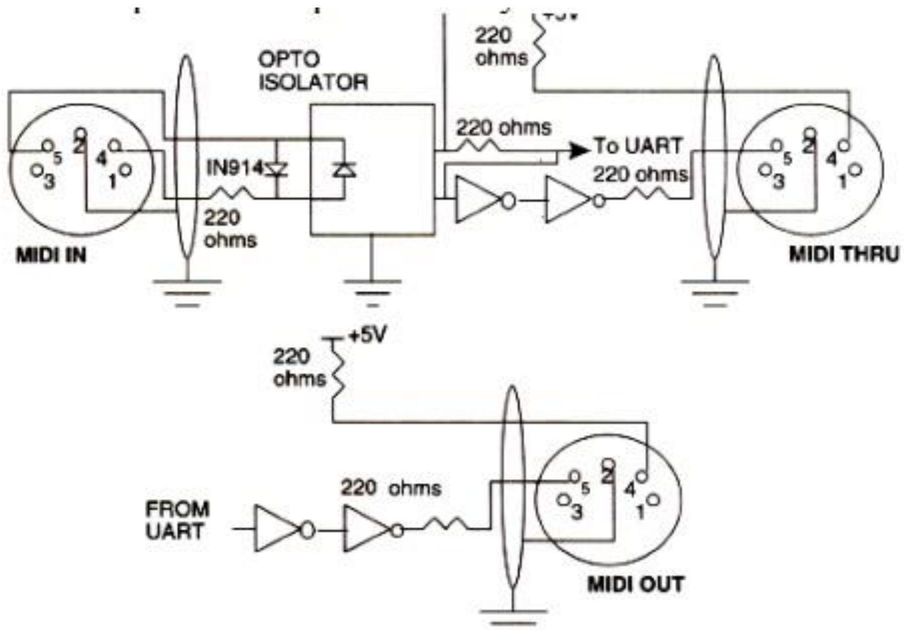
(i) A section consisting of digitization, amplitude normalization, time nonnalization and parametric representation.

(ii) Second section consisting of segmentation and labeling of the speech segment into a symbolic string based on a knowledgebased or rule-based systems.

(iii) The final section is to match speech segments to recognize word sequences.

Voice Recognition performance

It is categorized into two measures: Voice recognition performance and system performance. The following four measures are used to determine voice recognition performance.



Voice Recognition Applications

Voice mail integration: The voice-mail message can be integrated with e-mail messages to create an integrated message.

DataBase Input and Query Applications

A number of applications are developed around the voice recognition and voice synthesis function. The following lists a few applications which use Voice recognition.

- Application such as order entry and tracking

It is a server function; It is centralized; Remote users can dial into the system to enter an order or to track the order by making a Voice query.

- Voice-activated rolodex or address book

When a user speaks the name of the person, the rolodex application searches the name and address and voice-synthesizes the name, address, telephone numbers and fax numbers of a selected person. In medical emergency, ambulance technicians can dial in and register patients by speaking into the hospital's centralized system.

Police can make a voice query through central data base to take follow-up action ifhe catch any suspect.

Language-teaching systems are an obvious use for this technology. The system can ask the student to spell or speak a word. When the student speaks or spells the word, the systems performs voice recognition and measures the student's ability to spell. Based on the student's ability, the system can adjust the level of the course. This creates a self-adjustable learning system to follow the individual's pace.

Foreign language learning is another good application where"' an individual student can input words and sentences in the system. The system can then correct for pronunciation or grammar.

Musical Instrument Digital Interface (MIDI)

MIDI interface is developed by Daver Smith of sequential circuits, inc in 1982. It is an universal synthesizer interface

MIDI Specification 1.0

MIDI is a system specification consisting of both hardware and software ~omponents which define inter-coimectivity and a communication protocol for electronic sysnthesizers, sequences, rythm machines, personal computers, and other electronic musical instruments. The inter-connectivity defines the standard cabling scheme, connector type and input/output circuitry which enable these different MIDI instruments to be interconnected. The communication protocol defines standard multibyte messages that allow controlling the instrument"s voice and messages including to send response, to send status and to send exclusive.

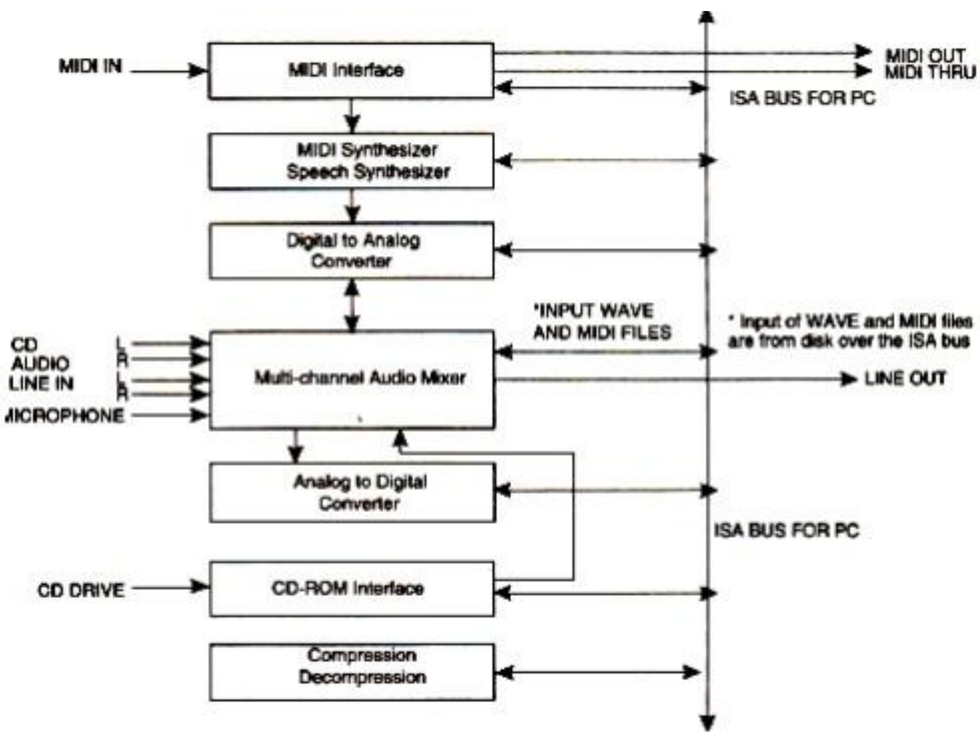
MIDI Hardware Specification

The MIDI. hardware specification require five pin panel mount requires five pin panel mount receptacle DIN connectors for MIDI IN, MIDI OUT and MIDI THRU signals. The MIDI IN connector is for input signals The MIDI OUT is for output signals MIDI THRU connector is for daisy-chaining multiple MIDI instruments.

MIDI Interconnections

The MIDI IN port of an instrument receives MIDI ncssages to play the instrument's internal synthesizer. The MIDI OUT port sends MIDI messages to play these messages to an external synthesizer. The MIDI THRU port outputs MIDI messages received by the MIDI IN port for daisy-chaining external synthesizers.

MIDI Input and output circuitry:



Communication Protocol

The MIDI communication protocol uses multibyte messages; There are two types of messages:

(i) **Channel messages**

(ii) System messages.

The channel message have three bytes. The first byte is called a status byte, and the other two bytes are called data bytes.

The two types of channel messages: (**i) Voice messages**

(ii) Mode messages.

System messages: The three types of system messages.

Common message: These messages are common to the complete system. These messages provide for functions.

System real.time messages: These messages are used for setting the system's real-time parameters. These parameters include the timing clock, starting and stopping the sequencer, resuming the sequencer from a stopped position and restarting the system.

System exclusive message: These messages contain manufacturer specific data such as identification, serial number, model number and other information.

SOUND BOARD ARCHITECTURE

A sound card consist of the following components:

MIDI Input/Output Circuitry, MIDI Synthesizer Chip, input mixture circuitry to mix CD audio input with LINE IN input and microphone input, analog-to-digital converter with a pulse code modulation circuit to convert analog signals to digital to create WAVfiles, a decompression and compression chip to compress and decompress audio files, a speech synthesizer to synthesize speech output, a speech recognition circuitry to recognize speech input and output circuitry to output stereo audio OUT or LINEOUT.

AUDIO MIXER

The audio mixer c:omponent of the sound card typically has external inputs for stereo CD audio, stereo LINE IN, and stereo microphone MICIN.

These are analog inputs, and they go through analog-to-digitaf conversion in conjunction with PCM or ADPCM to generate digitized samples.

Analog-to-Digital Converters: The ADC gets its input from the audio mixer and converts the amplitude of a sampled analog signal to either an 8-bit or 16-bit digital value.

Digital-to-Analog Converter (DAC): A DAC converts digital input in the 'foml of W AVE files, MIDI output and CD audio to analog output signals.

Sound Compression and Decompression: Most sound boards include a codec for sound compression and decompression.

ADPCM for windows provides algorithms for sound compression.

CD-ROM Interface: The CD-ROM interface allows connecting u CD ROM drive.to the sound board.

VIDEO IMAGES AND ANIMATION

VIDEO FRAME GRABBER ARCHITECTURE

A video frame grabber is used to capture, manipulate and enhance video images.

A video frame grabber card consists of video channel multiplexer, Video ADC, Input look-up table with arithmetic logic unit, image frame buffer, compression-decompression circuitry, output color look-up table, video DAC and synchronizing circuitry.

Video Channel Multiplexer:

A video channel multiplexer has multiple inputs for different video inputs. The video channel multiplexer allows the video channel to be selected under program control and switches to the control circuitry appropriate for the selected channel in aTV with multi – system inputs.

Analog to Digital Converter: The ADC takes inputs from video multiplexer and converts the amplitude of a sampled analog signal to either an 8-bit digital value for monochrome or a 24 bit digital value for colour.

Input lookup table: The input lookup table along with the arithmetic logic unit (ALU) allows performing image processing functions on a pixel basis and an image frame basis. The pixel image-processing functions ate histogram stretching or histogram shrinking for image brightness and contrast, and histogram sliding to brighten or darken the image. The frame-basis image-processing functions perform logical and arithmetic operations.

Image Frame Buffer Memory: The image frame buffer is organized as a l024 x 1024 x 24 storage buffer to store image for image processing and display.

Video Compression-Decompression: The video compressiondecompression processor is used to compress and decompress still image data and video data.

Frame Buffer Output Lookup Table: The frame buffer data represents the pixel data and is used to index into the output look uptable. The output lookup table generates either an 8 bit pixel value for monochrome or a 24 bit pixel value for color.

SVGA Interface: This is an optional interface for the frame grabber. The frame grabber can be designed to include an SVGA frame buffer with its own output lookup table and digital-to-analog converter.

Analog Output Mixer: The output from the SVGA DAC and the output from image frame buffer DAC is mixed to generate overlay output signals. The primary components involved include the display image frame buffer and the display SVGA buffer. The display SVGA frame buffer is overlaid on the image frame buffer or live video, This allows SVGA to display live video.

Video and Still Image Processing

Video image processing is defined as the process of manipulating a bit map image so that the image can be enhanced, restored, distorted, or analyzed.

Let us discuss about some of the terms using in video and still image processing.

Pixel point to point processing: In pixel point-to-point processing, operations are carried out on individual pixels one at a time.

Histogram Sliding: It is used to change the overall visible effect of brightening or darkening of the image. Histogram sliding is implemented by modifying the input look-up table values and using the input lookup table in conjunction with arithmetic logic unit.

Histogram Stretching and Shrinking: It is to increase or decrease the contrast.

In histogram shrinking, the brighter pixels are made less bright and the darker pixels are made less dark. **Pixel Threshold:** Setting pixel threshold levels set a limit on the bright or dark areas of a picture. Pixel threshold setting is also achieved through the input lookup table.

Inter- frame image processing

Inter- frame image processing is the same as point-to-point image processing, except that the image processor operates on two images at the same time. The equation of the image operations is as follows: Pixel output (x, y) = (Image l(x, y)

Operator (Image 2(x, y)

Image Averaging: Image averaging minimizes or cancels the effects of random noise.

Image Subtraction: Image subtraction is used to determine the change from one frame to the next .for image comparisons for key frame detection or motion detection.

Logical Image Operation: Logical image processing operations are useful for comparing image frames and masking a block in an image frame.

Spatial Filter Processing The rate of change of shades of gray or colors is called spatial frequency. The process of generating images with either low-spatial frequency-components or high frequency components is called spatial filter processing.

Low Pass Filter: A low pass filter causes blurring of the image and appears to cause a reduction in noise.

High Pass Filter: The high-pass filter causes edges to be emphasized. The high-pass filter attenuates low-spatial frequency components, thereby enhancing edges and sharpening the image.

Laplacian Filter: This filter sharply attenuates low-spatial-frequency components without affecting and high-spatial frequency components, thereby enhancing edges sharply.

Frame Processing Frame processing operations are most commonly for geometric operations, image transformation, and image data compression and decompression Frame processing operations are very compute intensive many multiply and add operations, similar to spatial filter convolution operations.

Image scaling: Image scaling allows enlarging or shrinking the whole or part of an image.

Image rotation: Image rotation allows the image to be rotated about a center point. The operation can be used to rotate the image orthogonally to reorient the image if it was scanned incorrectly. The operation can also be used for animation. The rotation formula is:

pixel output-(x, *y)* = pixel input *(x,* cos Q + *y* sin Q, - *x* sin Q + Y cos Q) where, Q is the orientation angle

x, yare the spatial co-ordinates of the original pixel.

Image translation: Image translation allows the image to be moved up and down or side to side. Again, this function can be used for animation.

The translation formula is:

Pixel output (x, y) =Pixel Input $(x + Tx, y + Ty)$ where

Tx and Ty are the horizontal and vertical coordinates. x, yare the spatial coordinates of the original pixel.

Image transformation: An image contains varying degrees of brightness or colors defined by the spatial frequency. The image can be transformed from spatial domain to the frequency domain by using frequency transform.

Image Animation Techniques

Animation: Animation is an illusion of movement created by sequentially playing still image frames at the rate of 15-20 frames per second.

Toggling between image frames: We can create simple animation by changing images at display time. The simplest way is to toggle between two different images. This approach is good to indicate a "Yes" or "No" type situation.

Rotating through several image frames: The animation contains several frames displayed in a loop. Since the animation consists of individual frames, the playback can be paused and resumed at any time.

FULL MOTION VIDEO

Most modem cameras use a CCD for capturing the image. HDTV video cameras will be all-digital, and the capture method will be significantly different based on the new NTSC HDTV Standard.

Full-Motion Video Controller Requirements

Video Capture Board Architecture: A full-motion video capture board is a circuit card in the computer that consists of the following components:

(i) Video input to accept video input signals.

(ii) S- Video input to accept RS 170 input.

(iii) Video compression-decompression processor to handle different video compression-decompression algorithms for video data.

(iv) Audio compression-decompression processor to compress and decompress audio data.

(v) Analog to digital converter.

(vi) Digital to analog converter.

(vii)Audio input for stereo audio LINE IN, CD IN. (viii) Microphone.

A video capture board can handle a variety of different audio and video input signals and convert them from analog to digital or digital to analog.

Video Channel Multiplexer: It is similar to the video grabber's video channel multiplexer.

Video Compression and Decompression: A video compression and decompression processor is used to compress and decompress video data.

The video compression and decompression processor contains multiple stages for compression and decompression. The stages include forward discrete cosine transformation and inverse discrete cosine transformation, quantization and inverse quantization, ZigZag and Zero run-length encoding and decoding, and motion estimation and compensation.

Audio Compression: MPEG-2 uses adaptive pulse code modulation (ADPCM) to sample the audio signal. The method takes a difference between the actual sample value and predicted sample value. The difference is then encoded by a 4-bit value or 8-bit value depending upon the sample rate

Analog to Digital Converter: The ADC takes inputs from the video switch and converts the amplitude of a sampled analog signal to either an 8-bit or 16-bit digital value.

STORAGE AND RETRIVAL TECHNOLOGY

Multimedia systems require storage for large capacity objects such as video, audio and images.

Another requirement is delivery of audio and video objects. Storage technologies include battery powered RAM, Nonvolatile flash, rotating magnetic disk drives, and rotating optical disk drives: Let us discuss these technologies in detail.

MAGNETIC MEDIA TECHNOLOGY

Magnetic hard disk drive storage is a mass storage medium.

It has advantages of it continual reduction in the price per mega byte of high-capacity storage. It has high capacity and available in low cost.

In this section let us concentrate on magnetic disk I/O subsystems most applicable to multimedia uses.

HARD DISK TECHNOLOGY

Magnetic hard disk storage remains a much faster mass storage to play an important role in multimedia systems.

It remains a much faster mass storage medium than any other mass storage medium.

ST506 and MFM Hard drives: ST506 is an interface that defines the signals and the operation of signals between a hard disk controller and the hard disk. It is developed by seagate. It is used to control platter speed and the movement of heads for a drive. Parallel data is converted to a series of encoded pulses by using a scheme called MFM (modified frequency modulation). The MFM encoding scheme offers greater packing of bits and accuracy than the FM encoding scheme. Other encoding scheme is Run-Length-Limited. Its drive capacity varies from 20 M Bytes to 200 M Bytes.

ESDI Hard Drive: ESDI (Enhanced Small Device Interface) was developed by a consortium of several manufacturers. It converts the data into serial bit streams.

It uses the Run-Length-Limited Scheme for encoding. The drive has data separator circuitry Drive capacity varies from 80 M Bytes to 2 GB. ESDI interface has two ribbon cables: (i) 36 pin cable for control signals. (ii) 20 pin cable for data signals.

IDE: Integrated Device Electronics (IDE) contains a,n integrated controller with drive.

The interface is 16 bit parallel data interface. The IDE interface supports two IDE drives. One is master drive and other is slave drive. Here, Jumper setting is required. The transfer rate is 8 MHz at bus speed.

New Enhanced IDE Interface

This new interface has a transfer rate of 9-13 M Bytes/See with maximum capacity around 8 GB. It supports upto four drives CD ROM and tape drives.

SCSI (Small Computer System Interface)

It is an ANSI X3T9.2 standard which supports SCSI and SCSI2 Standards. The Standard defines both software and hardware.

SCSI-I:It defines an 8-bit parallel data path between host adapter and device.

Here, host adapter is known as initiator and the device is known as target. There are one initiator and seven targets.

Nine control signals define the activity phases of the SCSI bus during a transaction between an initiator and a target. The phases are:

(i) arbitration phase (ii) selection phase (iii) command phase (iv) data phase (v) status phase (vi) message phase (vii) bus free phase.

Arbitrary Phase: In this phase an initiator starts arbitration and tries to acquire the bus.

Selection Phase: In this phase, an initiator has acquired the bus and selects the target to which it needs to communicate.

Command Phase: The target now enters into this phase. It requests a command from the initiator. Initiator places a command on the bus. It is accepted by the target.

Data Phase: The target now enters in this phase. It requests data transfer with the initiator. The data is placed on the bus by the target and is then accepted by the initiator.

Status Phase: Now, the target enters in status phase. It indicates the end of data transfer to the initiator. **Message Phase**: This is the last phase. It is to interrupt the initiator signaling completion of the read message. The bus free phase is a phase without any activity on the bus so that the bus can settle down before the next transaction. SCSI-l transfers data in 8-bit parallel form, and the transfer rate vades rom I M Bytes/See to 5 M Bytes/Sec. SCSI-I'drive capacity varies from 20 M bytes to 2 GB. SCSI-1 has over 64 commands specified to carry out transactions.

Commands include read, write, seek, enquiry, copy, verify, copy and verify, compare and so on.

SCSI-2

It has the same aspects of SCSI -1, But with faster data transfer . rates, and wider data width.

It includes few more new commands, and vender-unique command sets for optical drives, tape drives, scanners and so on. To make the bus wider, a system designer uses a second 68-pin connector in addition to the standard 50 pin connector.

Magnetic Storage Densities and Latencies

The Latency is divided into two categories: seek latency and rotational latency. Data management provides the command queuing mechanism to minimize latencies and also set-up the scatter-gather process to gather scattered data in CPU main memory.

Seek Latencies: There are three seek latencies available. They are· overlapped seek latency, Mid-transfer seek and Elevator seek.

Rotational Latencies: To reduce latency, we use two methods. They are:

(i) Zero latency read/write: Zero latency reads allow transferring data immediately after the head settles. It does not wait for disk revolution to sector property.

(ii) Interleaving factor: It keeps up with the data stream without skipping seccors. It determines the organization of sectors.

Transfer Rate and I/O per Second: I/O transfer nite varies from 1.2 M bytes/Sec. to 40 M bytes/Sec. Transfer rate is defined as the rate at which data is transferred from the drive buffer to the host adapter memory.

Data Management: It includes Command queueing and Scattergather. Command queueing allows execution of multiple sequential commands with system CPU intervention.Scatter is a process of setting the data for best fit in available block of memory or disk. Gather is a process which reassembles data into contiguous blocks on memory or disk ..

Figure below shows the relationship between seek latency, Rotational latency and Data transfer

It is a method of attaching multiple drives to a single host adapter. The data is written to the first drive first, then after filling it, the controller, allow the data to write in second drive, and so on. Meantime Between Failure (MTBF) = MTBF of single/drivel Total no. of dr

RAID (Redundant Array of Inexpensive Disks)

It is an alternative to mass storage for multimedia systems that combines throughput speed and reliability improvements.

RAID is an array of multiple disks. In RAID the data is spread across the drives. It achieves fault tolerance, large storage capacity and performance improvement.

If we use RAID as our hot backups, it will be economy. A number of RAID schemes havebeen developed:

1. Hot backup of disk systems
2. Large volume storage at lowercost
3. Higher performance at lower cost
4. Ease of data recovery
5. High MTBF.

There are six levels of RAID available.

(i) RAID Level 0 Disk Striping

It spreads data across drives. Data is striped to spread segments of data across multiple drives. Data striping provides high transfer rate. Mainly, it is used for database applications.

RAID level 0 provides performance improvement. It is achieved by overlapping disk reads and writes. Overlapping here means, while segment I is being written to drive 1, segment 2 writes can be initiated for drive 2.

RAID Level 1 Disk Mirroring

The Disk mirroring causes two copies of every file to be written on two separate drives. (Data redundancy is achieved).

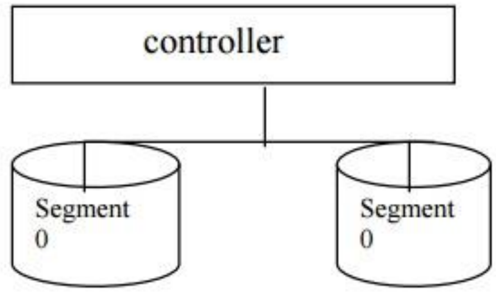
These drives are connected to a single disk controller. It is useful in mainframe and networking systems. Apart from that, if one drive fails, the other drive which has its copy can be used.

Performance: Writing is slow.

Reading can be speeded up by overlapping seeks.

Read transfer rate and number ofI/O per second is better than a single drive. I/O transfer rate (Bandwidth) = No. of drives x drive I/O transfer rate

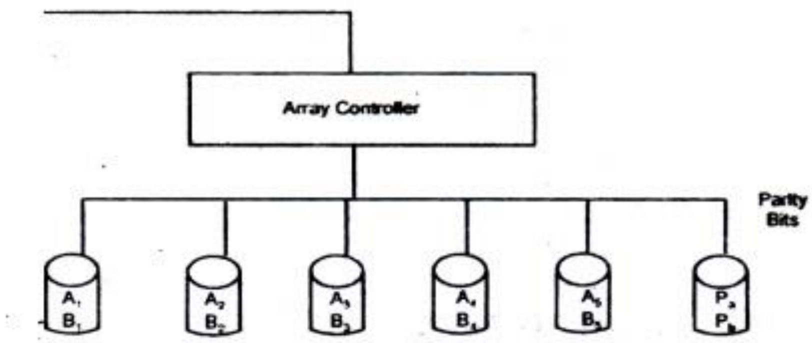
$$\text{No of I/O's Per second} = \frac{I/O\,transfer\,rate}{Average\,size\,of\,transfer}$$



Disk controller arrangement for RAID Level1

RAID Level 2, - Bit interleaving of Data: It contains arrays of multiple drives connected-to a disk array controller.

Data (written one bit at a time) is bit interleaved across multiple drives. Multiple check disks are used to detect and correct errors.



Host adapter organization of bit interleaving for RAID level2

It provides the ability to handle very large files, and a high level of integrity and reliability. It is good for multimedia system. RAID Level 2 utilizes a hamming error correcting code to correct single-bit errors and doublebit errors.

Drawbacks:

(i) It requires multiple drives for error correction (ii) It is an expensive approach to data redundancy. (iii) It is slow.

Uses: It is used in multimedia system. Because we can store bulk of video and audio data.

RAID Level-3 Parallel Disk Array: RAID 3 subsystem contains an array of multiple data drives and one parity drive, connected to a disk array controller.

The difference between RAID 2 and RAID 3 is that RAID 3 employs only parity checking instead of the full hamming code error detection and correction. It has the advantages of high transfer rate, cost effective than RAID 2, and data integrity.

RAID Level-4 Sector Interleaving: Sector interleaving means writing successive sectors of data on different drives.

As in RAID 3, RAID 4 employs multiple data drives and typically a single dedicated parity drive. Unlike RAID 3, where bits of data are Written to successive disk drives, an Ri\ID 4, the first sector of a block of data is written to the first drive, the second sector of data is written to the secohd drive, and so on. The data is interleaved at the data level.

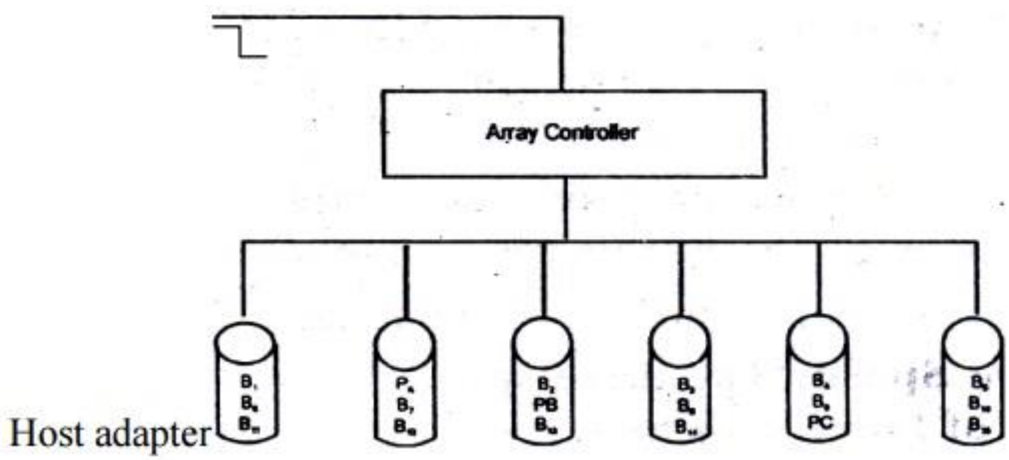
RAID *Leve1-4* offers *cost-effective* improvement in performance with data.

RAID Level-5 Block Interleaving: In RAID LevelS, as in all the other RAID systems, multiple drives are connected to a disk array controller.

The disk array controller contains multiple SCSI channels.

A RAID 5 system can be designed with a single SCSI host adapter with multiple drives connected to the single SCSI channel.

Unlike RAID Level-4, where the data is sector-interleaved, in RAID Level-5 the data is block-interleaved.



RAID LEVEL 5 DISK ARRAYS

Optical Media

CD ROM, WORM (Write once, Read many) and rewritable optical systems are optical drives. ·CD-ROMs have become the primary media of choice for music due to the quality of ,sound. WORMs and erasable opticel drives both use lasers to pack

information densely on a removable disk.

Optical Media can be classified by technology as follows:

- CD-ROM - Compact Disc Read Only Memory
- WORM - Write Once Read Many
- Rewritable - Erasable
- Multifunction - WORM and Erasable.

CD-ROM

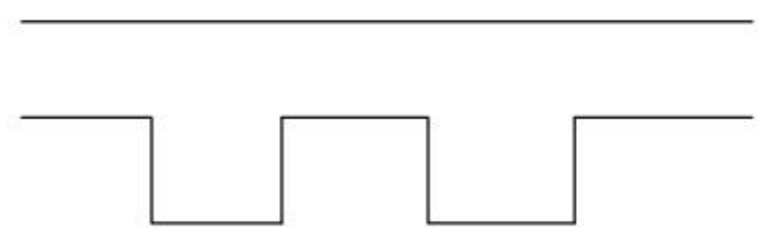
Physical Construction of CD ROMs:

It consists of a polycarbonate disk. It has 15 mm spindle hole in the center. The polycarbonate substrate contains lands and pits.

The space between two adjacent pits is called a land. Pits, represent binary zero, and the transition from land to pits and from pits to land is represented by binary one.

The polycarbonate substrate is covered by reflective aluminium or aluminium alloy or gold to increase the reflectivity of the recorded surface. The reflective surface is protected by a coat oflacquer to prevent oxidation. A CD-ROM consists of a single track which starts at the center from inside and spirals outwards. The data is encoded on this track in the form of lands and pits. A single track is divided into equal length sectors and blocks**.**

CD-ROM Physical Layers



Each sector or block consists of2352 bytes, also called a frame. For Audio CD, the data is indexed on addressed by hours, rninutes, seconds and frames. There are 75 frames in a second.

Magnetic Disk Organization: Magnetic disks are organized by CYlinder, track and sector. Magnetic hard disks contain concentric circular tracks. They are divided into sector.

Component of rewritable phase change cd-rom

Organization of magnetic media

CD-ROM Standards : A number of recording standards have emerged for CD-ROMs. They are:

CD-DA (DD-Digital Audio) Red Book: CD-ROM is developed by philips and sony to store audio information. CD-DA is the basic medium for the music industry.

The standard specifies multiple tracks, with one song per track. One track contains one frame worth of data: 2352 bytes. There are 75 frames in a second. Bandwidth = 176 *KB/s.*

CD-ROM Mode 1 Yellow Book: The Mode 1 Yellow Book Stnadard was developed for error correction. The Yellow Book Standard dedicates 288 bytes for error detection codes (EDCs) and error correction codes (ECCs).

CD-ROM Mode 2 Yellow Book

The Mode 2 Yellow Book standard was developed for compressed audio and video applications where, due to lossy compression, data integrity is not quite as important. This standard maintains the frame stmcture but it does not contain the *ECC/EDC* bytes. Removing the *ECC/EDC* bytes allows a frame to contain an additional 288 bytes of data, resulting in an increase of 14% more data. The frame stmcture is shown in the Table below:

Synchronization	Header	Data
12 Bytes	4 Bytes	2336 Bytes
0-11	13-15	16-2351

CD-ROMXA

XA stands for Extended Architecture. The standard was created for extending the present CD-ROM format.

CD-ROM XA contains multiple tracks. Each track's content is desclibed by mo~e. CD-ROM XA also allows interleaving audio and video objects with data for synchroni~ed playback. It does not support video compression. It supports audio compression. It uses Adaptive differential pulse Code Modulation algorithms.

CD-MO Orange Book Part 1

This standard defines an optional pre-mastered area conforming to the Red, Yellow or Green book standards for read-only, and a recordable area. It utilizes a read/write head similar to that found in magnetooptical drives. We can combine the pre-master multimedia objects as the base and develop their own versions.

CD-R Orange Book Part 2

This standard allows writing data once to a writeable disk. Here, the CD contains a polycarbonate substrate with pits and lands.

The polycarbonate layer is covered with an organic dye recording layer.

As in CD-ROM construction, the track starts from the center and spirals outwards. CD-R uses a high powered laser beam. The laser beam alters the state of the organic dye such that when the data is read, the altered state of dye disperses light instead of reflecting it. The reflected beam is measured for reading the state of each bit on the disk.

Mini-Disk

Mini-Disk for Data is known as MD-Data. It was developed by Sony Corporation. It is the data version of the new rewritable storage format. It can be used in three formats to support all users.

A premastered optical disk.

A recordable magneto-optical disk. A hybrid of mastered and recorded.

Its size is 2.5 inch. It provides large capacity. It is low cost. It is used in multimedia applications.

WORM Optical Drives

It records data using a high power laser to create a permanent burnt-in record of data. The laser beam makes permanent impressions on the surface of the disk.

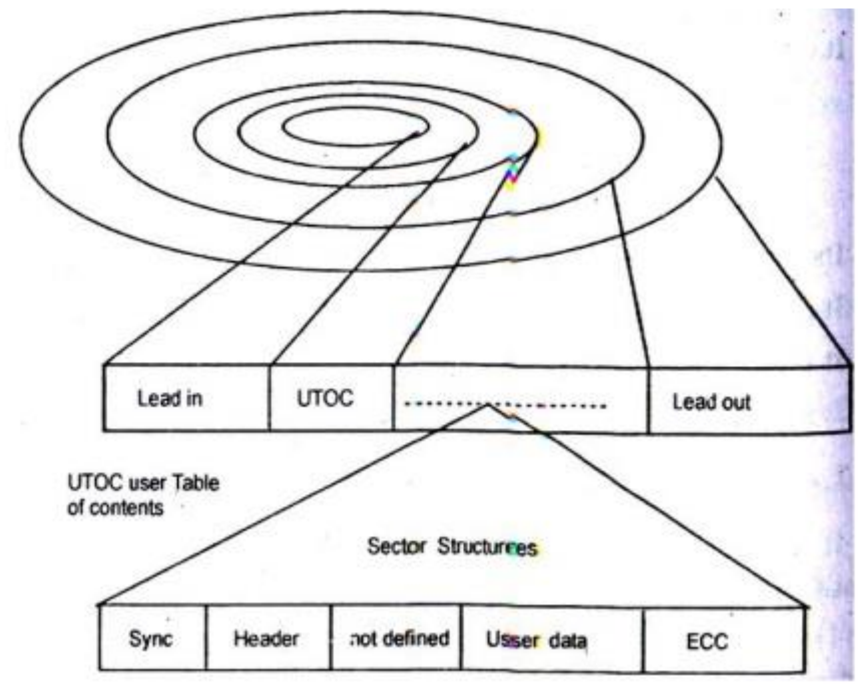
It creates pits. Information is written once. It cannot be written over and cannot be erased. i.e., Here data cannot be edited.

Recording of information: During recording, the input Signal is fed to a laser diode. The laser beam from the laser diode is modulated by the inpUt signal. It switches the laser beam on and off. if the beam is on, it strikes the three recording layers.

The beam is absorbed by the bismuth-tellurium layer. Heat is generated within the layer. This heat diffuses the atoms in the three recording layers. It forms four-element alloy layer. Now, the layer becomes recorded layers.

Reading Information from disk:

During disk read, a weaker, laser beam is focused on to the disk. It is reflectted back. The beam splitter mirror and lens arrangement sends the reflected beam to the photo detector. The photo sensor detects the beam and converts it into an electrical signal.



WORM DRIVE Applications

On-line catalogs Large-volume distribution Transaction logging Multimedia archival.

Rewritable Optical Disk Technologies

This technology allows erasing old data and rewriting new data over old data. There are two types of rewritable technology: (i) Magneto-optical ii)Phase change.

Magneto-Optical Technology

It uses a combination of magnetic and laser technology to achieve read/write capability. The disk recording layer uses a weak magnetic field to record data under high temperature. High temperature is achieved by laser beam.

When the beam is on, it heats the spot on the magneto optical disk to its curie temperature. The rise in temperature makes the spot extra sensitive to the magnetic field of bias field.

Magneto-optical drives require two passes to write data; in the first pass, the magneto optical head goes through an erase cycle, and in the second pass, it writes the data.

During the erase cycle, the laser beam is turned on and the bias field is modulated to change the polarity of spots to be erased. During the write cycle, the bias field is turned on and the laser beam is modulated to change the polarity of some spots to 1 according to the bit value.

Phase change Rewritable optical Disk

In phase change technology the recording layer changes the physical characteristics from crystalline to amorphous and back under the influence of heat from a laser beam.

To read the data, a low power laser beam is transmitted to the disk. The reflected beam is different for a crystalline state than for an amorphous state. The difference in reflectivity determines the polarity of the spot.

Benefits: it requires only one pass to write.

Dye Polymer Rewritable Disk

There is no need of magnetic technology here.

This technology consists of giant molecules formed from smaller molecules of the same kind with light-sensitive dye. This technology is also used in WORM drives.