Chapter 7

Optical Storage Media

Current magnetic data storage carriers take the form of floppy disks or hard disks and are used as secondary storage media. Here, low average access time and adequate capacity can be offered for a reasonable price. However, since audio and video, either in compressed or uncompressed form, require higher storage capacity than other media, the storage cost for such continuous media data using traditional storage carriers is essentially higher.

Optical storage media offer a higher storage density at a lower cost. The *Audio Compact Disk*, the successor to *Long Play Disks* (LPs), is a commercially successful product in the entertainment industry. The computer industry has profited from this development, especially when audio and video should be stored digitally in the computer. This technology has been *the* main catalyst for the whole development of multimedia in computing because it is used in multimedia external devices. For example, external devices such as video recorders and DAT recorders (Digital Audio Tape) can be used for multimedia systems. The actual integration into the system is difficult, but not impossible [HS91, RSSS90, SHRS90]. For this reason we will discuss the optical storage medium in more detail in this chapter. Other data storage media will not be considered here because they do not have the special properties which should be taken into account with respect to integrated multimedia systems.

This chapter provides an overview of the fundamentals for optical storage media.

Moreover, some analog and WORM (Write Once Read Many) systems will be briefly discussed. The CD-ROM and CD-ROM/XA are explained as they are derived from the CD-DA. Further developments with respect to multimedia such as CD-I, Photo-CD and DVI are presented. In addition to the *read-only* CD developments, there already exist techniques for writing to CD-WO and CD-MO. At the end of this chapter, the relationships between current CD technologies are shown and further developments are mentioned.

7.1 History

The video disk in the form of the *Video Long Play* (VLP) has been available for quite some time and was described in detail by 1973. The read-only video disk has not become a commercial success, although there are several *Write-Once* (WO) optical disks of different sizes and formats available on the market. These initial developments were based on analog techniques with the highest quality requirements at a moderate cost.

Ten years later, towards the end of 1982, the *Compact Disk-Digital Audio* (CD-DA) was introduced. This optical disk allows the digital storage of stereo-audio information at a high level of quality. N.V. Philips, in cooperation with the Sony Corporation, specified the CD-DA and developed the basic technology [MGC82, DG82, HTV82, HS82]. The CD-DA specification was published as the *Red Book* [Phi82] on which all other CD formats are based. In the five years since its introduction, approximately 30 million CD-DA recording devices and over 450 million CD-DA disks were sold.

In 1983, the extension of compact disk technology to storage of computer data was announced by N.V. Phillips and the Sony Corporation, and it was presented for the first time in November 1985. This *Compact Disk-Read Only Memory* (CD-ROM) specification was described in the *Yellow Book* [Phi85] and later became the standard ECMA-119 [ECM88]. The standard ECMA-119 specifies the CD-ROM physical format. The CD-ROM logical format is specified in the ISO Standard 9660 and came from the *High Sierra Proposal* (a proposal of industrial companies). It allows data access over file names and directories.

The *Compact Disk-Interactive* (CD-I) was announced by N.V. Phillips and the Sony Corporation in 1986. The CD-I specification is described in the *Green Book* and includes, besides the standard CD technology, a complete system description [Phi88]. In 1987, the *Digital Video Interactive* (DVI) was publicly presented. Here, the algorithms for compression and decompression of audio and video data stored on a CD-ROM are of importance.

In 1988, the *Compact Disk-Read Only Memory / Extended Architecture* (CD-ROM/X. was announced. N.V. Phillips, the Sony Corporation and Microsoft specified the digital optical data carrier for several media and published the specification at the CD-ROM conference in Anaheim, California (USA) in 1989 [Phi89].

Since the beginning of 1990, developments in *Compact Disk-Write Once* (CD-WO) technology, as well as *Compact Disk-Magneto Optical* (CD-MO), have been well-known. The specifications of CD-WO and CD-MO are covered in the *Orange Book* [Phi91].

7.2 Basic Technology

In principle, optical storage media use the intensity of reflected laser light as an information source. A laser beam of approximately 780 nm wave length can be focused at approximately 1 μm. In a polycarbonate *substrate layer* we encounter holes, corresponding to the coded data, which are called *pits*. The areas between these pits are called *lands*. Figure 7.1 shows a cut through an optical disk along a track. In the middle of the figure, the lands and pits are schematically presented.

The *substrate layer* is covered with a thin *reflective layer*. The laser beam is focused on the reflective layer from the substrate layer. Therefore, the reflected beam has a strong intensity at the lands. The pits have a depth of 0.12 μm (from the substrate surface). The laser beam is lightly scattered at the pits, meaning it is reflected with a weak intensity. The signal, shown in Figure 7.1, denotes schematically the intensity of the reflected beam – a horizontal line is drawn as the threshold value. Hence, according to Figure 7.1, a compact disk consists of:

- The label.

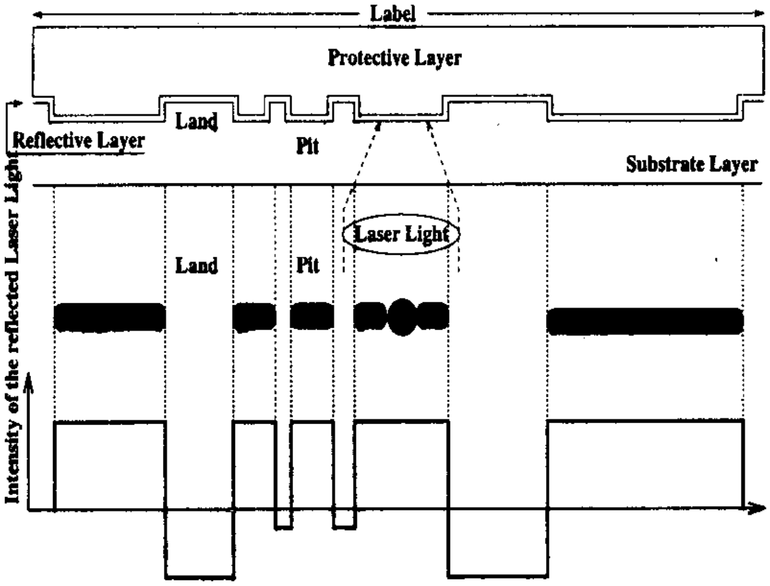

Figure 7.1: *Cut through an optical disk along the data trace. A schematic presentation with the layers (above), the "lands" and the "pits" (in the middle), and the signal (below).*

- The protective layer.
- The reflective layer.
- The substrate layer.

An optical disk consists of a sequential order of these pits and lands allocated in one track. Figure 7.2 shows an enlarged cut of such a structure.

In contrast to floppy disks and other conventional secondary storage media, the entire optical disk information is stored in one track. Thus, the stored information can be easily played back at a continuous data rate. This has advantages for audio and video data, as they are continuous data streams.

The track is a spiral. In the case of a CD, the distance between the tracks is 1.6 μm. The track width of each pit is 0.6 μm. The pits themselves have different lengths. Using these measurements, the main advantage of the optical disk in comparison

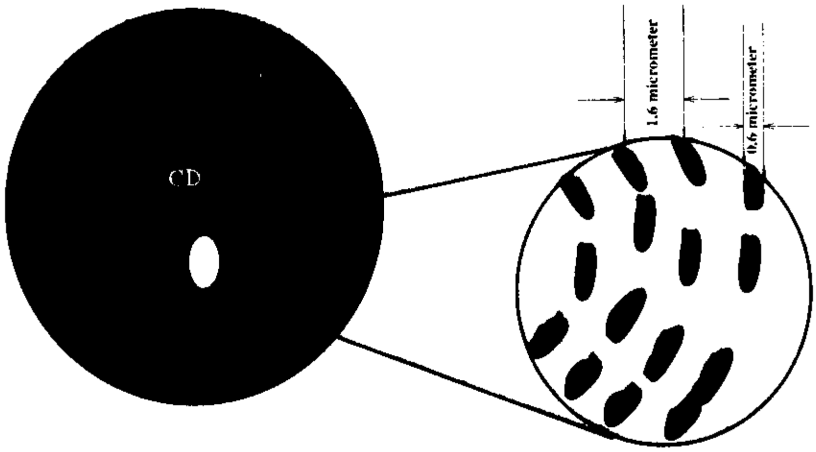

Figure 7.2: *Data on a CD as an example of an optical disk (track with "lands" and "pits").*

to 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.

While magnetization can decrease over time and in the case of tapes, for example, cross talk can occur, these effects are unknown in optical disks. Hence, this medium is very good for long-term storage. Only a decomposition or change of the material can cause irreparable damage. According to current knowledge, such effects will not occur.

The light source of the laser can be positioned at a distance of approximately one millimeter from the disk surface, and hence, it does not have to be positioned directly on the disk, respectively near the surface, as is the case with magnetic hard disks. This approach reduces friction and increases the life span of the involved components.

7.3 Video Disks and Other WORMs

The video disk, in the form of *Laser Vision*, serves as the output of motion pictures and audio. The data are stored in an analog-coded format on the disk; the reproduced data meet the highest quality requirements. The Laser Vision disk has a diameter of approximately 30 cm and stores approximately 2.6 Gigabytes.

Due to the similarities to LP records for audio information, the video disk was originally called the *Video Long Play* disk. It was described for the first time in Phillips' 1973 Technical Review [Phi73].

The motion picture on the video disk is encoded as frequency modulation, and the audio signal is mixed with the video signal. Figure 7.1 shows the principle of the recorded data. The main information of the mixed audio-video signal is the time at which the signal has the value zero. Hence, each zero cross-point corresponds to a change between a pit and a land on the disk. Such a change can occur at any time, and is written to the disk in a non-quantized form, i.e., the pit length is not quantized. Therefore, this method is time-continuous and *analog*.

Since the video disk was designed as *Read Only Memory* (ROM), many different write-once optical storage systems have come out, known as the *Write Once Read Many (WORM)* disk. An example is the *Interactive Video Disk*. This disk is played at a *Constant Angular Velocity (CAV)*. On each side, 36 minutes of audio and video at a rate of 30 frames per second can be stored and retrieved. One can also store around 54,000 studio quality images per side.

Write-once storage media have a capacity between 600 MBytes and approximately 8 Gigabytes. The diameter of the disks is between 3.5 and 14 inches. The main advantage of a WORM disk, compared to other mass storage media, is the ability to store large amounts of data which may not be changed later, i.e., an archive which is secure. To increase capacity, *juke-boxes* are available, which allow the stocking of several disks and lead to capacities of over 20 Gigabytes.

The following interesting peculiarities occur when using a WORM disk:

- *Media Overflow* can occur when a WORM disk is nearly full. Prior to writing data to a WORM disk, it must be determined if the recorded data can fit onto the disk and/or if the data should be stored on different disks. Further, it is required that if the data need to be stored on more than one physical disk, the time point at which data should be written to another physical disk must be determined. This approach is especially important for continuous media because these media streams can only be interrupted at certain points.

- *Packaging* refers to the problem of fixed-block structures in WORMs. Only data sentences of a given size can be written. For example, if the block size is 2,048 bytes, and only one byte is written, 2,047 bytes will be recorded without any information content.
- *Revision* refers to the problem of subsequently marking invalid areas. For example, during changes of a document invalid areas are be created on the disk(s). These areas have to be subsequently rewritten, which implies a document may be distributed over several WORMs after a while. Here, the distribution of the document over several disks should not disturb the output of the data stream of a continuous medium.

There are also other problems with WORM disks such as the following: besides the number of incompatible WORM disk formats, most multimedia systems lack adequate software support for the WORM disks, which results in a poor integration of WORM technology into the computer environment.

7.4 Compact Disk Digital Audio

7.4.1 Preliminary Technical Background

The CD has a diameter of 12 cm; the disk is played at a *Constant Linear Velocity* (CLV). Therefore, the number of rotations per time unit depends on the particular radius of the accessed data. The spiral-shaped CD track consists of approximately 20,000 windings. In comparison, an LP disk has only approximately 850 windings.

Information is stored according to the principle shown in Figures 7.1 and 7.3. The length of the pits is always a multiple of 0.3 μm. The transition from pit to land

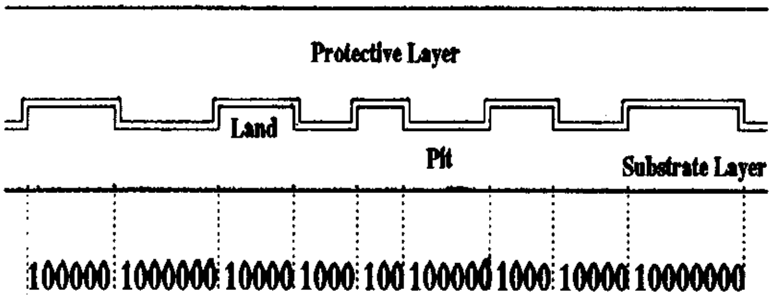

Figure 7.3: *Lands and pits with their related digital data stream.*

and from land to pit corresponds to the coding of a *1* in the data stream. A *0* is coded as no transition. Figure 7.3 shows a data stream as a sequence of lands and pits, and below it is the corresponding digital data stream.

Audio Data Rate

The audio data rate can be easily derived from the given sample frequency of 44.1 kHz and the 16-bit linear quantization. The stereo-audio signal obeys the pulse-code modulation rules and the following audio data rate is derived:

$$\begin{aligned}
Audio\ data\ rate_{CD-DA} &= 16\frac{bits}{sample} \times 2\ channels \times 44100\frac{samples}{s \times channel} \\
&= 1,411,200\frac{bits}{s} = 1,411,200\frac{bits/s}{8\ bits/byte} \\
&= 176.4\frac{kbytes}{s} \cong 172.3\frac{Kbytes}{s}
\end{aligned}$$

Analog LPs and cassette tapes have a signal-to-noise ratio between 50 dB and 60 dB. The quality of the CD-DA is substantially higher. As a first approximation, we can assume 6 dB per bit during the sampling process. Hence, with 16-bit linear sampling, we obtain the following:

$$S/N_{CD-DA} \cong 6\frac{dB}{bit} \times 16\ bits = 96\ dB$$

The signal-to-noise ratio is exactly 98 dB.

Capacity

A CD-DA play time is at least 74 minutes. With this value, the capacity of a CD-DA can be easily determined. The following example shows the computation of a capacity for pure audio data without taking into consideration additional information, such as error correction:

$$\begin{aligned} Capacity_{CD-DA} &= 74\ min \times\ 1,411,200\frac{bits}{s} = 6,265,728,000\ bits \\ &= 6,265,728,000\ bits \times \frac{1}{8\frac{bits}{byte}} \times \frac{1}{1024\frac{bytes}{Kbyte}} \times \frac{1}{\frac{Kbytes}{Mbyte}} \cong 747 Mbytes \end{aligned}$$

7.4.2 Eight-to-Fourteen Modulation

Each change from pit to land and from land to pit corresponds to the coding of a *1* which is sent across a communication channel as the *channel bit.*

Pits and lands may not follow too closely one after another on a CD-DA since the resolution of the laser is not sufficient to read such direct pit-land-pit-land-pit sequences (i.e., 11111 sequences) correctly. Therefore, an agreement was negotiated that at least two lands and two pits always occur consecutively in a sequence. Hence, between two *1s* there always exist at least two *0s*.

On the other hand, pit or land sequences are not allowed to be too long; they must keep a maximal distance. Otherwise, no phase-correct synchronization signal (clock) can be derived. Hence, the maximal length of the pits and lands is limited. At most, ten *0s* (as the channel bits) can follow one after another.

For these reasons, the bits written on a CD-DA, in the form of pits and lands, do not directly correspond to the actual information; before recording, *Eight-to-Fourteen Modulation* is applied [HS82]. Using this transformation, the regularity of the minimal and maximal distances is met.

Eight-bit words are coded as 14-bit values. Given the minimal and maximal allowable distances, there are 267 valid possibilities. 256 possibilities are used. For example, the code includes the following two entries shown in Table 7.1.

Audio Bits	Modulation Bits
00000000	01001000100000
00000001	10000100000000

Table 7.1: *Two values out of the code table.*

Audio Bits	00000000						00000001		
Modulation Bits			0	100	1000	100000		10000	100000000
Filling Bits	0	10					100		
Channel Bits .	0	10	0	100	1000	100000	100	10000	100000000
On the CD-DA	l	pp	p	lll	pppp	llllll	ppp	lllll	ppppppppp

Table 7.2: *Integration of the filling bits.*

Through the direct consecutive insert of the modulation bits (14-bit-values), the minimal allowed distance of two bits and the maximal distance of ten bits may still not be followed. Therefore, three additional bits are inserted between two consecutive modulation symbols so that the required regularity can be met. The filling bits are chosen depending on the neighboring bits. Table 7.2 shows an example (the audio bits are taken from Table 7.1) to clarify the integration of the filling bits and the whole transformation process from an 8-bit audio word to its CD-DA representation of lands and pits.

7.4.3 Error Handling

The goal of error handling on a CD-DA is the detection and correction of typical error patterns [HTV82]. A typical error, a consequence of a scratch and/or pollution, can be characterized as a *burst error.*

On the first level, a two-stage error correction is implemented according to the Reed-Solomon algorithm: for every 24 audio bytes, at the first stage, individual byte errors are recognized and corrected; at the second stage, double byte errors are recognized and corrected; also, other fault bytes in the sequence can be recognized, but they cannot be corrected using this approach.

On the second level, real consecutive data bytes are distributed over several frames. A *frame* consists of 588 channel bits, which correspond to 24 audio bytes. Therefore, the audio data is stored on the CD-DA in an interleaved form. This means that due to a burst error, only parts of data are modified.

An error rate of 10^{-8} is specified. More precisely, the burst errors, which are distributed over a maximal 7 frames, can be recognized and corrected. This corresponds to a track length of 7.7 mm. For example, a 2-mm-diameter hole in a CD-DA means that the audio data can still be played correctly. Nevertheless, experiments have shown that not all devices correct each error according to the given specification. The above described method for error correction is known as *Cross Interleaved Reed-Solomon Code* (CIRSC).

7.4.4 Frames, Tracks, Areas and Blocks of a CD-DA

Audio data, error correction, additional control and display-bytes, and a synchronization pattern all constitute *frames.*

- *Audio data* are divided into two groups, each consisting of 12 audio bytes. They contain the high and low bytes of the left and right channels.

- *Error-detection* and *error-correction* bytes (4 bytes per frame) are inserted, according to the above description, at both stages.

- Each frame has a *control* and a *display* byte. It consists of eight bits, which are marked with *P, Q, R, S, T, U, V* and *W* (*subchannel bits*). These subchannel bits are drawn together over 98 frames for each subchannel and used together. Hence, there are eight subchannels, each having 98 bits, of which 72 are used for the actual information. The 98 frames together build a *block*. Unfortunately, sometimes the blocks are also called frames.

  So the *P*-subchannel is used to differentiate between a CD-DA (with audio data) and a CD with other computer data. The *Q*-subchannel is used, for example, in the following:

  - The *lead-in area* for storage of the directory content.

  - The rest of the CD-DA for a specification of the relative time inside a track, and for the absolute time specification of the CD-DA.

- The *synchronization pattern* determines the beginning of a frame. The pattern consists of twelve *1s* and twelve *0s* as channel bits, and three filling bits.

Table 7.3 shows an overview of frame components with corresponding bytes and bits.

	Audio Bits	Modul.Bits	Fil.Bits	Ch.Bits	Together
Synchronization			3	+ 24	= 27 bits
Control & Display		i.e.	(14 + 3)		= 17 bits
12×Data	12 × 8	i.e. 12×	(14 + 3)		= 204 bits
4×Error Handling		i.e. 4×	(14 + 3)		= 68 bits
12×Data	12 × 8	i.e. 12×	(14 + 3)		= 204 bits
4×Error Connection		i.e. 4×	(14 + 3)		= 68 bits
Frames Together					= 588 bits

Table 7.3: *Components of a frame.*

Using these data, different data streams with corresponding rates can be recognized [MGC82]:

- The *audio bit stream* (also called the *audio data stream*) carries 1.4112×10^6 bits/s. Here, only the 16-bit quantized samples are taken.

- The *data bit stream* consists of the audio bit stream, including the control and display bytes, and necessary bytes for error handling. The number of bits can reach the value of 1.94×10^6 bits/s.

- The channel bit stream includes the data bit streams with the *Eight-to-Fourteen Modulation*, the *filling bits* and the *synchronization bits.* The data rate is approximately 4.32×10^6 bits/s.

A CD-DA consists of the following three *areas*:

- The *lead-in area* includes the directory of the CD-DA. Here, the beginning of the individual tracks are registered.

- The *program area* includes all tracks of the CD-DA. Here, the actual data are stored.

- At the end of each CD-DA there is a *lead-out area*. This area is used mainly to help the play-recorder when the reader head accidentally goes beyond the program area.

The program area of each CD-DA can consist of 99 *tracks* of different lengths. It includes at least one track and each track incorporates at most one song, or a sentence of a symphony. If the program area is randomly accessed (e.g., access 5th song), the reader head is positioned to the beginning of the particular track (e.g., position to the beginning of the 5th track).

Each track can have (according to the *Red Book*) several *index points*. Therefore, at certain places, direct positioning can occur. Mostly, only two a priori defined *Index Points (IPs)* are used, the IP_0 and IP_1. IP_0 marks the beginning of each track. IP_1 defines the beginning of the audio data inside each track. The area between IP_0 and IP_1 is called *track pregap*. CD-DA disks possess a track pregap of two to three seconds for each piece.

Another structure, called *block* (Figure 7.4), was introduced in addition to frames and tracks, but it does not have a special meaning in CD-DA technology. It is introduced at this point only because of comparability reasons to other CD technologies discussed in the next sections. A CD-DA block includes 32 frames and stores 2,352 bytes.

Figure 7.4: *CD-DA block layout according to the "Red Book."*

7.4.5 Advantages of Digital CD-DA Technology

Errors on a CD-DA can be caused by damage or pollution. The CD-DA is not sensitive, with respect to uncompressed audio, to the usual appearance of reading errors. The CD-DA specification, in the form of the *Red Book*, serves as the basis for all optical CD storage media. For example, *Eight-to-Fourteen Modulation* and the *Cross Interleaved Reed-Solomon Code* are always used. Hence, a fundamental specification has been developed which is used in many systems and means compatibility for many systems.

The disadvantage of the CD-DA technology is that its achieved reliability is too low for the storage of computer data. This lead to a further development of the CD technology.

7.5 Compact Disk Read Only Memory

The *Compact Disk Read Only Memory* (CD-ROM) was designed as the storage format for general computer data – in addition to uncompressed audio data [Che86, FE88, Hol88, LR86, OC89]. Further, CD-ROM technology has been planned to be the basis for storage of other media (e.g., video) [KSN+87, Wil89]. This was specified in the *Yellow Book* by N.V. Phillips and the Sony Corporation [Phi85]. The *Yellow Book* was later accepted as the ECMA standard [ECM88].

CD-ROM tracks are divided into *audio* (corresponding to the CD-DA) and *data* types. One track itself may either contain audio only or data only. A CD-ROM can contain both types of tracks, tracks with audio and other tracks with data. In such a mixed form, the data tracks are usually located at the beginning of the CD-ROM and then followed by the audio tracks. Such a CD is called a *Mixed Mode Disk* (see Figure 7.11).

7.5.1 Blocks

A CD-DA has an error rate of less than 10^{-8} and allows random access to individual tracks and index points. The use of a CD-ROM with its general-purpose computer data requires much better error correction and random access to a data unit with a higher resolution than the track.

This data unit is called a *block*, meaning the *physical block*. In the ISO 9660 standard, there also exists the notion of a *logical block*. The logical block has similar properties to the sectors of other media and file systems. A CD-ROM block consists of 2,352 bytes of a CD-DA block. Hence, the de facto CD-DA standard serves as the basis for the de facto CD-ROM standard.

Out of the 2,352 bytes of a block, 2,048 or 2,336 bytes (depending on whether computer data or audio data are stored on CD-ROM) can be used for user data. The remaining bytes are used for the identification of random access, as well as for another error correction layer, thereby lowering the error rate further.

75 blocks per second are played back, each consisting of 32 frames. Each frame is 73.5 bytes (588 bits).

$$Block = 1411200\frac{bits}{s} \times \frac{1}{75}s \times \frac{1}{8bits/byte} = 2352bytes$$

Figure 7.5 shows the data hierarchy of a CD-ROM with audio data.

7.5.2 Modes

The CD-ROM was specified with the following goal: it should serve to hold uncompressed CD-DA data and computer data. This goal is achieved by introducing two modes: *mode 1* and *mode 2*.

CD-ROM Mode 1

CD-ROM mode 1 serves as the actual storage of computer data (Figure 7.6). The block contains 2,048 bytes for information storage out of the available 2,352 bytes. The 2,352 bytes are split into the following groups:

- 12 bytes for *synchronization*, i.e., for the detection of the block beginning.

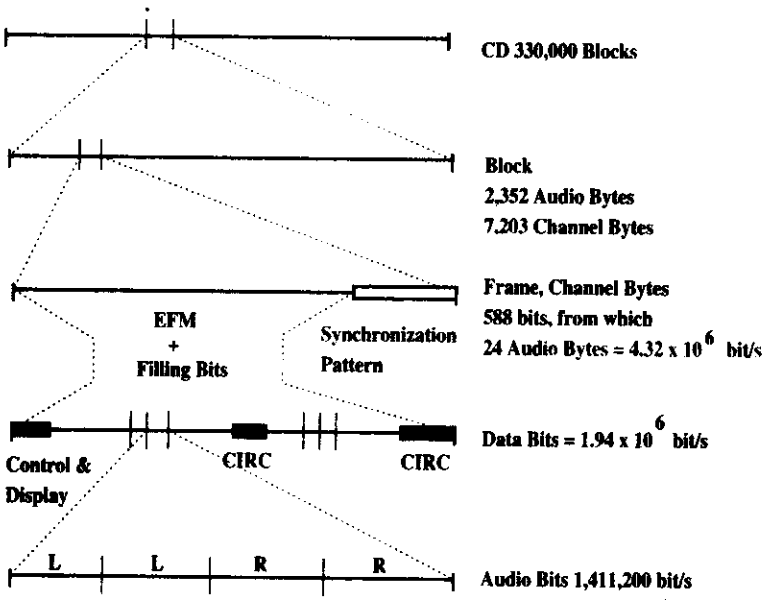

Figure 7.5: *CD-ROM data hierarchy with audio data.*

Sync	Header	User Data	EDC	Blanks	ECC
12	4	2048	4	8	276

|←—— 2,352 bytes ——→|

Figure 7.6: *CD-ROM mode 1 block layout according to the "Yellow Book."*

- 4 bytes for the *header*, which carries an unambiguous 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.

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 2

CD-ROM mode 2 holds data of any media. The data layout of a CD-ROM block in mode 2 is shown in Figure 7.7. Here, each block offers 2,336 bytes for information storage. The synchronization and header are processed in the same way as in mode

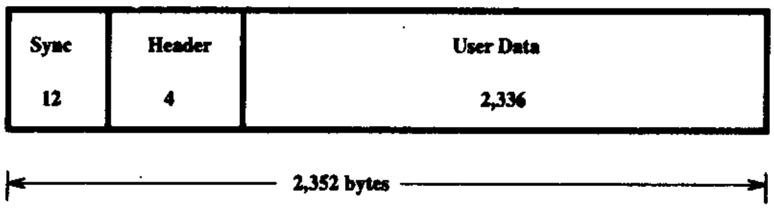

Figure 7.7: *CD-ROM mode 2 block layout according to the "Yellow Book."*

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:

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

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

7.5.3 Logical Data Format

At some point, it was recognized that the specification of the blocks in mode 1 was not a sufficient equivalent to the sectors. The logical data format with the directory was missing. Therefore, a group of industry representatives met in Del Webb's High Sierra Hotel & Casino in Nevada and worked out a proposal. It is called the *High Sierra Proposal* and it became the basis for the ISO 9660 standard, which exactly describes this format [KGTM90].

With this standard, a *directory tree* is defined, which includes all the information about the files. Additionally, there is a table in which all the directories are listed in compressed form. This path table – *Table of Contents* – allows direct access to files at any level. The table is loaded into the computer memory while mounting the CD. This static method is efficient because CD-ROM data cannot be changed (read only), yet most implementations use only the actual directory tree.

In the first track, ISO 9660 reserves the first 16 blocks (sectors 0 to 15) as the *system area*. This area can be used for indication of production-specific properties. Starting at sector 16, the *volume descriptor* (primary volume descriptor, supplementary volume descriptor, etc.) is stored. The supplementary volume descriptor can describe another file system, which also offers flexibility with respect to the allowed character string for file names. The most important descriptor is the primary volume descriptor. It includes, besides other information such as the logical block size, the length of its own defined file system, and the length and the addresses of the path tables.

Each volume descriptor is stored in a block with 2,048 bytes. A CD-ROM can include any number of volume descriptors. Mostly, individual volume descriptors are repeated as copies to provide increased access reliability. The volume descriptor area is closed with *volume descriptor terminators*, which are also implemented as a special block.

ISO 9660 uses the logical block size, which is equal to the power of two of a number beginning with 512 bytes. It is not allowed to have a logical block size greater than the size of the actual block (sector). The maximal logical block size is 2,048 bytes (de facto), although ISO 9660 does not strictly require this size.

If the underlying technology offers another physical block size, with ISO 9660, other logical block sizes are possible. Therefore, current block sizes of 512 bytes, 1,024 bytes and 2,048 bytes exist. The logical block size is the same for the entire file system. Files start at the beginning of logical blocks. While files can start and end inside of physical blocks (sector), directories always start at sector boundaries.

7.5.4 Limitations of the CD-ROM Technology

Any CD has a high storage capacity and constant data transfer rate. Random access to a CD track of approximately up to one second can be tolerated for audio playback. This access time, on one hand, still has important advantages in comparison to LPs, cassettes and DATs. On the other hand, these access times for a CD-ROM mean a significant disadvantage as a data carrier in comparison to magnetic disks (with a mean access time of under 10 milliseconds). The following effects contribute to the CD-ROM access time:

- *Synchronization time* occurs because the internal clock must be adjusted to be in phase with the signal reading the CD. Here, delays are in the range of milliseconds.
- Due to the Constant Linear Velocity (CLV) playback of a CD, reading the inner part of the CD requires around 200 rotations per second, and reading the outer part of the CD requires around 530 rotations per second. The *rotation delay* is specified as the maximal duration to position the laser above the desired sector within one rotation, and to correctly adjust the rotation speed. Hence, depending on the particular device, rotation delay can be up to 200 milliseconds.
- The *seek time* refers to the adjustment of the correct radius, in which the laser also must first locate the track and adjust itself. The seek time is at most one second.

These partially overlapping effects contribute to the highest maximal time of up to approximately one second for the worst case. However, the real values can be very different, depending on the actual and desired position of the information. Using cache-hierarchies, the access time for very good disk drives can be lowered to 200 milliseconds.

A constant audio stream on a CD-ROM block needs to be stored sequentially. We cannot, for example, play an audio track and the data of a CD-ROM mode 1 track simultaneously. Although very important for many multimedia systems, simultaneous playback of audio and other data is not possible.

7.6 CD-ROM Extended Architecture

The *Compact Disk Read Only Memory/Extended Architecture* (CD-ROM/XA) standard was established by N.V.Phillips and the Sony and Microsoft Corporations and is based on the CD-ROM specification [Fri92, GC89, Phi89]. The main motivation for this additional development was the concurrent output of several media, which was insufficiently considered in previous approaches. Before this standard specification, other definitions and systems could produce simultaneous output, such as CD-I (*Compact Disk Interactive*) and DVI (*Digital Video Interactive*). The experiences of CD-I were taken into account during the development of the CD-ROM/XA because of the tight connection of N.V. Phillips and the Sony Corporation to CD-I. Hence, many properties are identical.

The *Red Book* specifies a track for uncompressed audio data (Figure 7.4). The *Yellow Book* describes tracks for computer data with CD-ROM mode 1 (Figure 7.6) and tracks for compressed media with CD-ROM mode 2 (Figure 7.7). CD-ROM/XA uses CD-ROM mode 2 to define actual blocks. CD-ROM/XA defines, in addition to CD-ROM mode 2, a *subheader*, which describes the particular block (sector) as shown in Figures 7.8 and 7.9. This makes it possible to interleave different media using only mode 2 blocks. During playback, individual CD-ROM/XA data streams are separated.

7.6.1 Form 1 and Form 2

CD-ROM/XA differentiates blocks with form 1 and form 2 formats. This is similar to the CD-ROM modes:

- *Form 1*

  This CD-ROM mode 2 XA format provides improved error detection and correction. Analogous to the CD-ROM mode 1, four bytes are needed for detection and 276 bytes for correction. Contrary to CD-ROM mode 1, the unused eight bytes of mode 1 are used for subheaders. Figure 7.8 shows a block where 2,048 bytes are used as data.

Sync	Header	Sub-header	User Data	EDC	ECC
12	4	8	2048	4	276

|← 2,352 bytes →|

Figure 7.8: *CD-ROM/XA block layout according to the "Green Book" - layout of a CD-ROM block in mode 2, form 1.*

- *Form 2*

  This CD-ROM mode 2 XA format (Figure 7.9) allows 13% more storage capacity out of the entire block size (2,352 bytes) for user data, which means 2,324 bytes for user data. This is gained at the expense of worse error handling. In these form 2 blocks, compressed data of different media, including audio and video, can be stored.

Sync	Header	Sub-header	User Data	EDC
12	4	8	2324	4

|← 2,352 bytes →|

Figure 7.9: *CD-ROM/XA block layout according to the "Green Book" - layout of a CD-ROM block in mode 2, form 2.*

In the case of a CD-DA, CD-ROM or Mixed Mode Disk, a track always consists of homogeneous data, meaning either audio or computer data. The computer cannot, for example, concurrently retrieve uncompressed audio data and traditional computer data. The main advantage of the CD-ROM/XA is that within one track, blocks of different media can be stored, yet all are coded in CD-ROM mode 2. Therefore, it follows that CD-ROM/XA makes interleaved storage and retrieval possible.

7.6.2 Compressed Data of Different Media

CD-ROM/XA allows *interleaved* storage of different media. Audio can be compressed in different quality levels with *Adaptive Differential Pluse Code Modulation* (ADPCM). This compression improves the entire CD-DA duration from 74 uncompressed CD-DA minutes to over 19 hours of lower-quality audio by reducing the audio signal to four bits per sample. Further, this compression is necessary for simultaneous retrieval of other media. The following variants are possible:

- *Level B Stereo* has a compression factor of 4:1 in comparison to a CD-DA audio signal. The sample frequency of Level B is 37,800 Hz. This compression provides a CD-DA with a playback time capacity of 4 hours and 48 minutes, in comparison to 74 minutes when no compression at all is used. The data rate is around 43 Kbytes/s.

- *Level B Mono* has a compression factor of 8:1 in comparison to a CD-DA audio signal. This allows storage of 9 hours and 36 minutes of audio data. The data rate is approximately 22 Kbytes/s.

- *Level C Stereo* has a compression factor of 8:1 and results in the same storage capacity and data rate as Level B Mono. The sampling frequency is 18,900 Hz.

- *Level C Mono* works with a compression factor of 16:1. Hence, the storage capacity is a maximal 19 hours and 12 minutes with a data rate of approximately 11 Kbytes/s.

MPEG audio does not use ADPCM coding and therefore is not compatible with the CD-ROM/XA specification. But it can be assumed that in the future, compatibility will be designed into the specification. This media-specific encoding and decoding is not a part of the CD technology.

For implementations of the CD-ROM/XA format, the choice of medium and its corresponding quality can only be done by considering the maximal data rate. The same is applicable to other CD-based applications and formats, such as CD-I and DVI.

The logical format of a CD-ROM/XA uses the ISO 9660 specification. The ISO 9660 standard specifies *interleaved files*, i.e., several files are interleaved. *Channel interleaving* of audio, video and other data inside of one file is not specified in the ISO 9660 standard. Further, this standard does not prescribe any mandatory file content. Unfortunately, the notion of *file interleaving* is often used for interleaved files, as well as for channel interleaving.

7.7 Further CD-ROM-based Developments

The coexistence of different technologies is shown in Figure 7.10. It is important to point out that the CD-DA, CD-ROM and CD-ROM/XA specifications can be thought of as the layers in the communication systems. Basically, the CD-DA specification is valid for all layers. Although because of further developments, not all fundamental facts are summarized in the *Red Book*. For example, the *Mixed Mode Disk* is not defined there.

Based on these fundamental technologies, other CD-based developments which handle several media, or special media and application areas, have appeared or will soon appear. In the long run, we can see that further CD technologies will be based on CD-DA, CD-ROM and CD-ROM/XA.

7.7.1 Compact Disk Interactive

The *Compact Disk Interactive* (CD-I) was developed by N.V. Phillips and the Sony Corporation [vLZ89] prior to the specification of CD-ROM/XA. In 1986, CD-I was announced and in 1988 specified in the *Green Book* [Phi88] (based on the *Red Book*

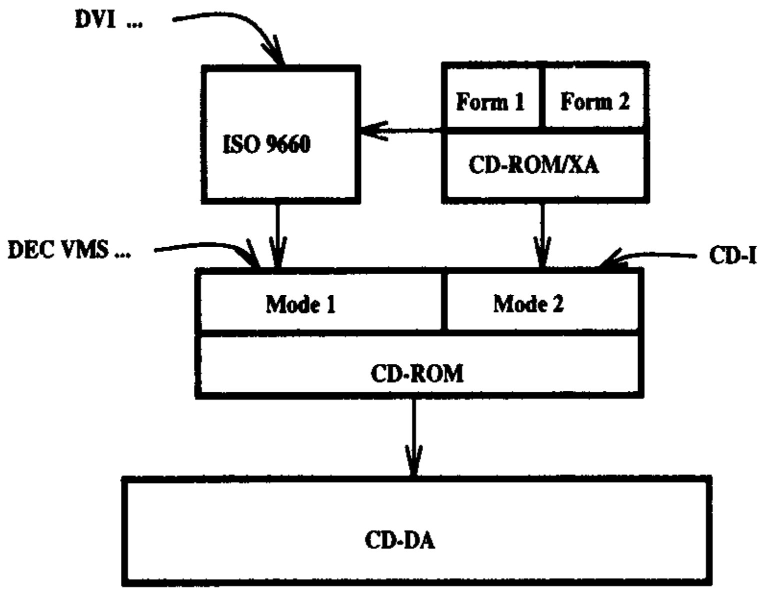

Figure 7.10: *CD Read Only technologies as layers.*

[Bas90, Int89, SvdM91]). CD-I was conceptually designed only for consumer electronics as an addition to a TV. Since October 1991, corresponding devices have been on the market.

CD-I represents an entire system. It contains a CD-ROM-based (not CD-ROM/XA) format with interleaving of different media and a definition of compression for different media. Further, CD-I defines a system software with CD-RTOS (Real-Time Operating System), which is an OS-9 derivation with extensions for real-time processing, and the output hardware for multimedia data.

The CD-I hardware is called the *Decoder*. Its size is comparable with the size of a current VCR. It consists of a main processor from the Motorola 68000 family and special video and audio elements. It also contains the CD player with a controller, joystick and mouse interface. Besides these components, an interface to an *RGB* monitor or TV is planned. CD-I devices are considered to be a replacement and/or extension of the CD-DA devices in the consumer environment.

	CD-DA	CD-I Level A	CD-I Level B	CD-I Level C
Sampling Frequency	44.1	37.8	37.8	18.9
Bandwidth (kHz)	20	17	17	8.5
Coding	16 bit PCM	8 bit ADPCM	4 bit ADPCM	4 bit ADPCM
Max. Recording Duration in Hours (Stereo/Mono)	74 min/-	4.8/2.4	9.6/4.8	19.2/9.6
Max. Number on Concurrent Channels (Stereo/Mono)	1/-	2/4	4/8	8/16
Portion (in %) of Entire Stream(Stereo/Mono)	100/	50/25	25/12.5	12.5/6.25
Signal-to-Noise Ratio (S/M) in dB	98	70	60	50
Quality	Audio CD	LP	FM Radio	AM Radio

Table 7.4: *CD-I audio coding (standards).*

Audio Encoding

Audio coding includes several quality levels differing in capacity and data rate, as shown in Table 7.4. The CD-I audio coding standard specifies quality levels, such as Level A, Level B and Level C, using ADPCM in the same way as CD-ROM/XA. The close relationship between CD-I and CD-ROM/XA is there because CD-I was the basis for the CD-ROM/XA specification. The low data rates can be used in combination with images or motion pictures. Several channels of lower quality can also be used for different languages.

Image Encoding

For coding of images using CD-I, different quality levels and resolutions can be used. The following overview shows that different size and data rates are also possible:

- *YUV* encoding serves for reproduction of natural pictures with many colors. The luminance component *Y* and the chrominance components *U*, *V* are coded with 360 × 240 pixels and 18 bits/pixel. There are 262,144 colors per image possible. These parameters lead to an image size of:

$$\frac{Data\ Size}{Image} = 360 \times 240 \times 18 \times \frac{1\ bit}{8\ bits/byte} = 194,400\ bytes$$

- CD-I can work with four bits per pixel using the *Color Look-Up Table* (CLUT). As an alternative, another 3.7 or 8 bits per pixel are available. This is suitable for simple graphics with fast access to a preloaded color table. Using the four bits per pixel, at most 16 colors can be presented. For example, with a resolution of 720 × 240 pixels and four bits/pixel, the image size is:

$$\frac{Data\ Size}{Image} = 720 \times 240 \times 4 \times \frac{1\ bit}{8\ bits/bytes} = 86,400\ bytes$$

- *RGB* encoding is used for high-quality image output. Here, each of the *R*, *G* and *B* components is coded with five bits/pixel. Using an additional bit, the colors are coded with 16 bits/pixel; hence, 65,538 colors per image can be presented. With the resolution of 36 × 240 pixels, the image size is:

$$\frac{Data\ Size}{Image} = 360 \times 240 \times 16 \times \frac{1\ bit}{8\ bits/byte} = 172,800\ bytes$$

Animation Encoding

Animation consists of motion pictures with few colors. The coding of animation is implemented with run-length coding and approximately 10,000 to 20,000 bytes per image. In the future, CD-I will use MPEG coding for video compression. The data format is similar to the ISO 9660 standard specification, but not quite compatible.

Although CD-I technology was actually designed by the consumer industry, today systems exist as parts of computers in commercial areas. CD-I is important in the context of CD technology because it provides the basis for CD-ROM/XA.

7.7.2 Compact Disk Interactive Ready Format

All CD formats are based on the same CD-DA standard, but it is not always possible to play a CD-I disk on a CD-DA device. It is also not correct to assume that all CD-DA devices will be replaced, for example, by CD-I devices. Therefore, there is a need for a format specification of an optical disk which can be played by CD-DA devices, as well as by CD-I devices – namely, the *Compact Disk Interactive Ready Format* [Fri92].

For this purpose, the *track pregap* area between the index points IP_0 and IP_1 at the beginning of a track is enlarged from 2-3 seconds to at least 182 seconds. The information specific to CD-I is stored in this area. This information can contain details about a particular piece of music, images or the biography of the composer and the conductor. A CD-I Ready Disk can be played in three different ways:

- Using the usual CD-DA media, the CD information in the track pregap area is ignored and only the audio is played.
- Using the second mode, only the CD-I data in the track pregap area is used. Data of all media are stored there and can be read, presented and interpreted.
- Using the third mode, the CD-I data from the track pregap area are presented simultaneously during audio output. This method is similar to the Mixed Mode Disk (see Section 7.5). For this purpose, the CD-I data are read and stored first. The next step is the output of the audio and corresponding CD-I data. Hence, a simultaneous presentation of the data is achieved.

7.7.3 Compact Disk Bridge Disk

The *Compact Disk Bridge Disk* (CD Bridge Disk) has a goal similar to the CD-I Ready Disk of being a format specification for an output device compatible with different CD formats. The CD-I Ready Disk has a fixed disk format for CD-DA and CD-I devices, whereas the CD Bridge Disk has a fixed disk format for CD-ROM/XA and CD-I devices [Fri92]. Figure 7.11 shows the placement of the CD Bridge Disk format among the different CD formats.

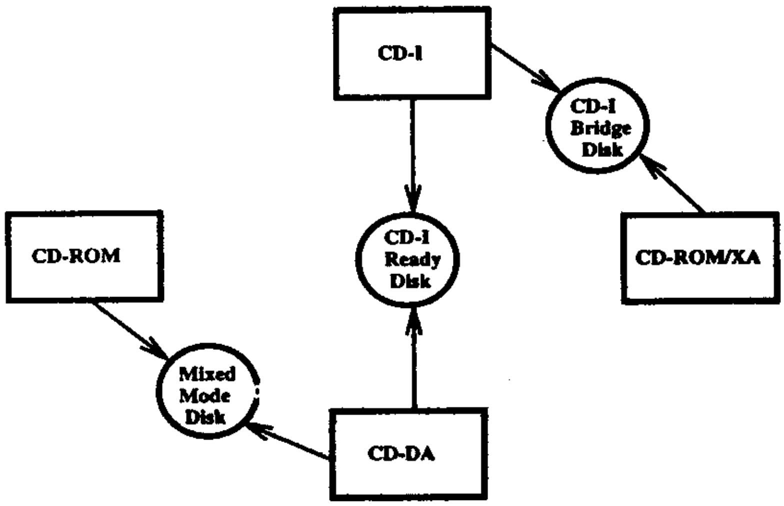

Figure 7.11: *CDs with several compatible formats: Mixed Mode Disk, CD-I Ready Disk and CD Bridge Disk.*

A CD Bridge Disk must satisfy the CD-I and CD-ROM/XA specifications. A common subset is defined which holds for both formats:

- All tracks with computer data (which are not uncompressed audio, CD-DA) must be written in CD-ROM mode 2. No CD-ROM mode 1 blocks are allowed to exist on the disk. All tracks with computer data can be followed by all audio tracks (CD-DA).
- Another example of compatibility with respect to both specifications is the track entry in the table of contents at the beginning of the CD. The reference to the CD-I tracks is not given in this area. All tracks with data are marked as CD-ROM/XA tracks.

7.7.4 Photo Compact Disk

The *Photo Compact Disk* (Photo-CD) from Eastman Kodak and N.V. Phillips Company is an example of a CD Bridge Disk [Fri92]. It was announced as the *Kodak Photo-CD System* in 1990 and will be, according to the press announcement, li-

Picture Name	compressed/uncompressed	line number	column number
Base/16	uncompressed	128	192
Base/4	uncompressed	256	384
Base	uncompressed	512	768
4Base	compressed	1024	1536
16Base	compressed	2048	3072

Table 7.5: *Image resolution of a Photo-CD [Klee92].*

censed from Agfa-Gevaert. Here, photographs of a high quality are stored. The Photo-CD is based on mechanisms described in Section 7.8, i.e., it has "write once" characteristics. The Photo-CD can be read as a CD Bridge Disk from CD-I and CD-ROM/XA devices. Additionally, it can be read and written as CD-WO (see Section 7.8) on special Photo-CD devices.

The Photo-CD is based on the following process: first, photographs are created conventionally using cameras and film; after film development, the pictures are digitalized with 8-bit luminance resolution and 2×8 bits chrominance resolution (each pixel is coded in 24 bits); subsequently, each picture is coded with one to five resolutions as shown in Table 7.5.

The integration of photos into digital computers and TVs brings many new applications to the home and business. Pictures can be retrieved by a computer or a TV extension. Using different resolutions, a digital zoom can be easily implemented. Using low resolutions, several pictures can be shown concurrently during an overview presentation. The pictures can be modified and inserted into documents.

7.7.5 Digital Video Interactive

Digital Video Interactive (DVI) describes – similar to CD-I – different components of a system. An overview of DVI, focusing on compression formats, is presented in Section 6.8. DVI consists of compression and decompression algorithms, highly integrated dedicated hardware components for compression and decompression of motion pictures in real-time, an application programming interface (to the Audio-Visual Kernel = AVK) and a fixed data format. In contrast to CD-I, the emphasis is not on CD-technology, but on compression algorithms [HKL+91, Lut91, Rip89].

DVI uses the CD-ROM mode 1 shown in Figure 7.6. In addition to this format, the ISO 9660 format for CD-ROM is used as the basis for the AVSS (Audio/Video Support System) interleaved file format. For example, Commodore uses the CD-ROM mode 1, as well as ISO 9660 format with the CDTV (Commodore Dynamic Total Vision). Yet, it is worth-while to mention that the ISO 9660 distinguishes among different *Interchange Levels*. DVI uses the basic mode (Interchange Level 1) with file names consisting of at most 8-Point-3-Characters (eight characters for filename prefix and three characters for filename suffix, prefix and suffix are separated by point) from a specific pre-defined character set. CDTV uses up to 30 characters from the Interchange Level 2, leading to larger file names.

7.8 Compact Disk Write Once

So far, all of the CD technologies considered (the sole exception being the Photo-CD described in Section 7.7.4) 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 [AFN90]. CD-WO is specified in the second part of the *Orange Book* [Phi91].

7.8.1 Principle of the CD-WO

The following section briefly explains the principle of CD-WO [Kle92]. Figure 7.12 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 modified through strong thermal influence, which changes the reflection properties of the laser beams.

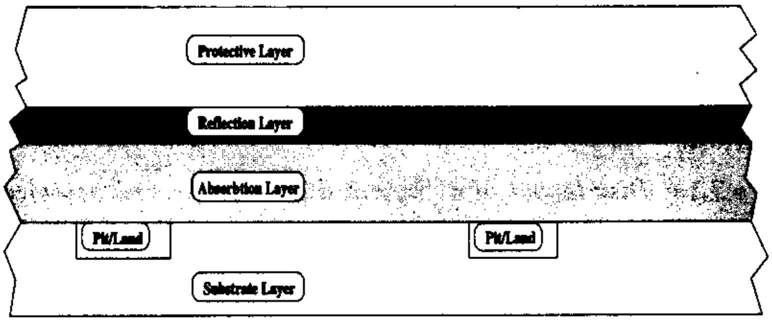

Figure 7.12: *Cross-section of a CD-WO disk.*

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 250^0C 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.

7.8.2 Sessions

All CD systems described so far assume that a *lead-in* area exists before the actual data area of a CD, and a *lead-out* area exists after the actual data of a CD (see Section 7.4.4). The content is written in a *table of contents* in the lead-in area. Each player needs this table of contents to position the player correctly. In the case of writing to a CD-WO, this lead-in area with the table of contents can be overwritten only after the entire write activity is complete. This means that the actual data of a CD-WO must be written before the table of contents is created. This further means that, in the meantime, the disk cannot be accessed and played by any other device.

Therefore, the principle of *several sessions* was introduced, as shown in Figure 7.13. Each session has its own lead-in area and lead-out area. During one write activity, all data for a session are written together with their table of contents, after which the disk can be played on other devices. Thereby the structure of a CD can be

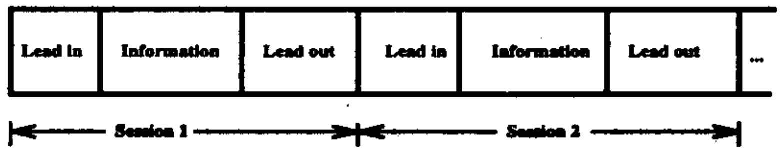

Figure 7.13: *Disk layout of a "hybrid disk" with several sessions.*

extended up to a maximal 99 sessions. However, because of the space requirement for lead-in and lead-out areas, at most 46 sessions can be stored. Each session consists again of its lead-in area, the data area and lead-out area. Until 1992, all available devices could read only one session. CD-WOs with only one session are called *regular CD-WOs*. A CD-WO with more than one session is called a *hybrid CD-WO*.

The *CD-WO recorder* typically works at a data rate double that of the player. This decreases the time for recording (write) by half the time for playing (read), but increases the requirements on the computer and its necessary software for installing and running a CD-WO. The reason is that this data rate must be sustained during the entire write activity. Simpler programs therefore produce an image (a one-to-one copy) of the CD-WO on a hard disk. In the second step, the data are transferred to the CD-WO. A storage-saving approach produces the data in the correct order and transfers them (without intermediate storage of the entire CD-WO information) at the necessary rate to the CD-WO.

The above description of CD-WO devices shows that a CD-WO disk could be a substitution, for example, for the CD-DA disk if it would have the same cost as CD-DA disk. However, the production process of the CD-WO is and will continue to be more expensive than the process for original CDs. Hence, CD-WO is only used in special application areas such as when huge data sets, because of technical or legal reasons, must be stored in an irreversible way. CD-WO also has application in the area of CD publishing because the production of an expensive and time-consuming master may be omitted. Fewer circulations with higher update (actual facts) can be produced.

7.9 Compact Disk Magneto Optical

The *Compact Disk Magneto Optical (CD-MO)* has a high storage capacity and allows one to write multiple times to the CD. CD-MO is specified in the first part of the *Orange Book* [Phi91].

7.9.1 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 150^0C. 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.

After the CD is irradiated 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.

7.9.2 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 in Figure 7.14) includes data which were written in a specified format onto the disk. Figure 7.14 shows the relationship between the premastered area of a CD-MO and read-only technologies. Therefore, only the CD-MO read-only area can be read by available playback devices.

The recordable area of a CD-MO cannot be played because the CD-MO read/write technology is fundamentally different from any CD-DA, CD-ROM, CD-ROM/XA

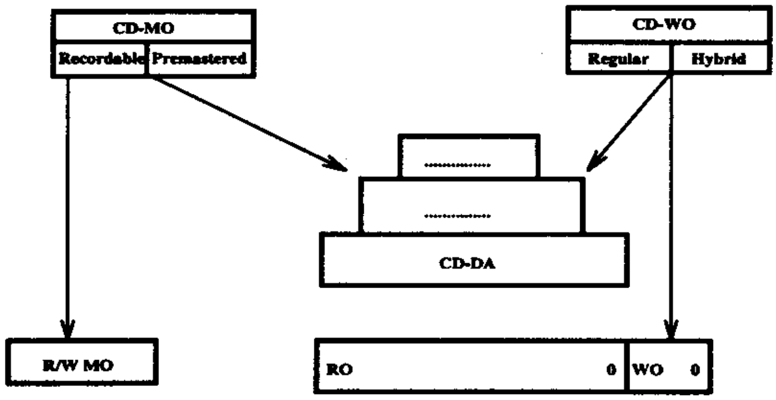

Figure 7.14: *CD-WO and CD-MO in relation to other CD technologies.*

or CD-WO device. Figure 7.14 shows the relationship between this recordable area and the fundamental magnetic-optical technology. Hence, this technology is unfortunately incompatible with all other CD systems, even though the same system parameters as in the other approaches were specified. For example, the dimensions and rotation speed are the same.

7.10 The Prospects of CD Technologies

Compact disk technology will remain the optical storage technology for all kinds of media. The relationship between different standards, shown in Figure 7.15, allows for a broad field of applications. Except for the CD-MO, the CD-DA with its optical technology still serves as the basis. A closer view and comparison of the formats show how the specification of CD technology has progressed. CD-ROM mode 1 defines improved error handling for computer data. In CD-ROM/XA form 1, based on CD-ROM mode 2, the same improved error handling is offered as in CD-ROM mode 1. This may mean that, for example, CD-ROM mode 1 could be omitted if so many applications did not already exist. The compression methods of CD-ROM/XA should use the JPEG, MPEG and CCITT ADPCM standards, and not limit themselves to the other coding methods for which cheaper integrated elements on the CD-ROM/XA controller boards already exist.

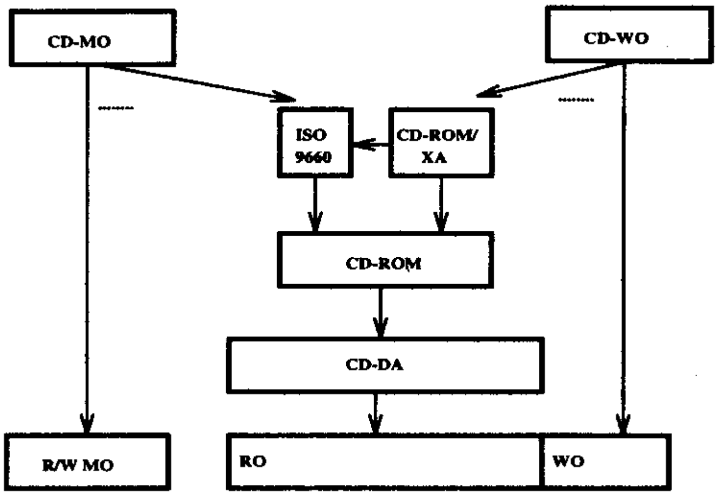

Figure 7.15: *The most important CD technologies and their mutual relations as a hierarchical structure.*

The disadvantage of all CDs is still the relatively high duration of the approximately 200 ms mean access. This property will not be improved substantially in the future. Unfortunately, there is also the incompatibility of the CD-MO to be considered.

The storage capacity of CDs is sufficient for many current systems. In research labs there are already optical storage media which work with a blue – instead of red – laser. One can envision, given the current pace of storage density increases, that in a few years there may exist a CD/2 which will have a capacity at least 16 times greater than today's CDs. Note that CD/2 is a fiction, standing for no ongoing research or development project.

Existing data rates require and determine efficient compression methods for audio, images and especially video. Analogous to the steps from TV to HDTV, from ISDN to B-ISDN and from MPEG to MPEG-2, one can also envision a CD/2 data transfer speed increase of up to factor 16. Indeed, today some CD-ROM built-in devices can continuously transfer above 600 Kbytes/s.

CD/2 devices should have the capability to also use current CDs as a storage medium. The difference would be that CD/2 disks would provide the aforementioned improved properties. This would allow a successive replacement of CD devices with new CD/2 devices and an essential acceleration of the CD/2 market introduction.

A harder requirement is to sustain the capability of a current CD device to play CD/2 disks. Even with a higher storage density, CD/2 could have, for example, several spirals running parallel as data tracks. A conventional track would be used by current devices. The other tracks should not influence – similar to the introduction of the color TV – the sampling of the existing track. It is important to note that this requirement could be fulfilled only at the expense of higher financial investments. Conversely, with the risk of user acceptance, a non-compatible system could make use of any new technology (e.g., the mini disk from Sony). Instead of increasing capacity to achieve the same goal, the size of the disk may be reduced to a Mini-CD with a storage capacity similar to today's CD.

Chapter 8

Computer Technology

A multimedia system is comprised of both hardware and software components, but the major driving force behind a multimedia development is research and development in hardware capabilities. For example, *Compact Disks* with their high storage capacity at a relatively low price provided the first step toward multimedia storage capabilities. Note that multimedia computer technology is a necessary but not sufficient pre-condition for *multimedia*. Unfortunately, computer hardware technology advances faster than the equivalent software. For example, many powerful computers with several signal-processors and RISC-processors have been built, but the capabilities of these processors are supported only very rarely by the current operating systems. Indeed, many operating systems cannot even fully use the capabilities of the processors in traditional workstations – they were designed for older generation of processors.

Besides the multimedia hardware capabilities of current personal computers (PCs) and workstations, computer networks with their increasing throughput and speed start to offer services which support multimedia communication systems. Also in this area, computer networking technology advances faster than the software. Today there are commercially-available computer networks with a throughput of over 100 Mbit/s, but full integration of high-speed networks and multimedia PCs or workstations are either missing or available only in research laboratories.

Many research projects are working toward solutions of the problems described above. Examples are European initiatives such as RACE, especially the German BERKOM project, and American high-speed network testbeds such as the AURORA, Blanca, Nectar, VISTAnet and CASA testbeds, and others [Par94].

We discuss in this chapter multimedia hardware components and their integration with the networks to support multimedia communication systems because without the necessary hardware, data storage capacity and continuous high data throughput, multimedia implementations would be impossible. The starting point for multimedia communication systems was the hybrid system where digital and analog components were integrated. The goal is a fully digital multimedia communication system [HS91].

Further, we give an overview of computer technology contributing to the development of multimedia workstations.

8.1 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 streams in the form of analog signals, although most network components (e.g., switches), over which these signals are transmitted, work internally in a digital mode. Figure 8.1 shows an analog and a digital environment without any interaction.

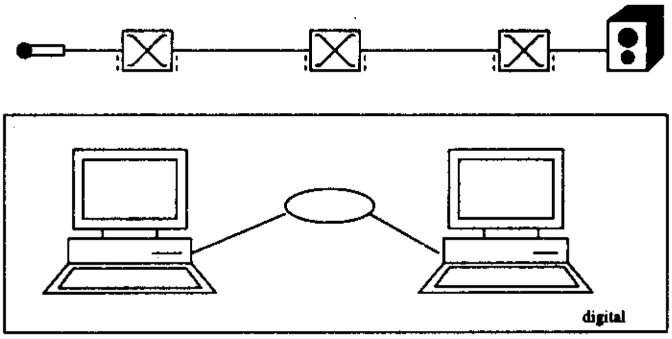

Figure 8.1: *Analog and digital environments without interaction.*

8.1.1 Hybrid Systems

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.

Integrated Device Control

One possible integration approach is to provide a control of analog input/output audio-video components in the digital environment. Moreover, the connection between the sources (e.g., CD player, camera, microphone) and destinations (e.g., video recorder, write-able CD), or the switching of audio-video signals can be controlled digitally.

Figure 8.2 shows a possible computer control and management of external analog components. The computer only controls and manages the streams, but it does not process the streams; hence, the existing quality and format of audio and/or video are not changed.

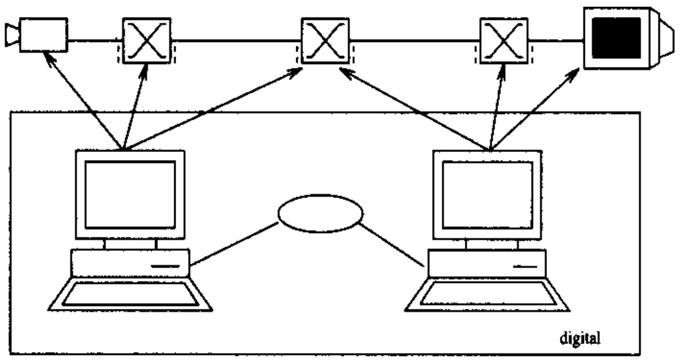

Figure 8.2: *Computer control of all audio-video components.*

Based on this hybrid approach, the Bell Communication Research Laboratory implemented the *Integrated Media Architecture Laboratory* (IMAL) in 1987. The IMAL project integrated a number of analog and digital networks [LD87]. The purpose of IMAL was to understand multimedia communication and to develop the necessary corresponding services.

Another example is the *Touring Machine System*™ from the Bellcore Information Networking Research Laboratory. A primary goal of the Touring Machine project is to provide a hybrid network infrastructure on which applications requiring complex multimedia communications can be developed, independent of the actual network fabric used to provide transport.

A third example is *Media Space*, created at the Xerox Palo Alto Research Center (PARC) in the mid-1980s. It provides an environment for collaborative work among the geographically separate research laboratories at Xerox PARC [BBI93].

Integrated Transmission Control

A second possibility to integrate digital and analog components is to provide a common transmission control. This approach implies that analog audio-video sources and destinations are connected to the computers for control purposes to transmit continuous data over digital networks, such as a cable network. Figure 8.3 show a possible system configuration.

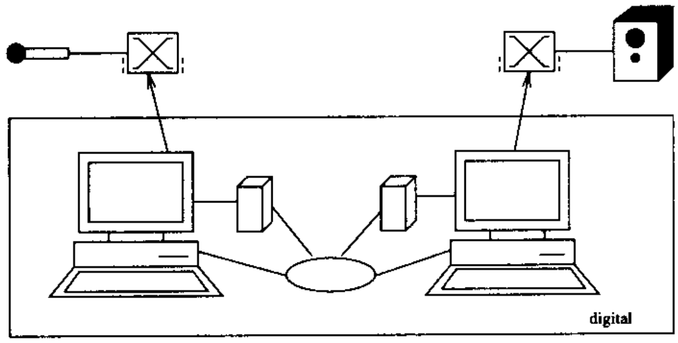

Figure 8.3: *Continuous data input into a digital network under computer control.*

For example, the video services of the *MIT Muse and Pygmalion Project* [CGR90, HS93, MTA+89] are based on such an approach: each computer is connected to the Ethernet for data communication, as well as to the cable network. The audio-video sources and destinations are connected to the computers through a Parallax family of adapters [Par87].

For both approaches above, the computer always controls the devices for processing and transmitting audio-video data. These data are not transferred through the computer but outside of it. Hence, in a very clever way, the problem of processing audio-video data under real-time conditions is avoided.

Integrated Transmission

The next possibility to integrated digital and analog components is to provide a common transmission network. This implies that external analog audio-video devices are connected to computers using A/D (D/A) converters outside of the computer, not only for control, but also for processing purposes. Continuous data are transmitted over shared data networks. Figure 8.4 shows a possible system configuration. This system structure implies several issues:

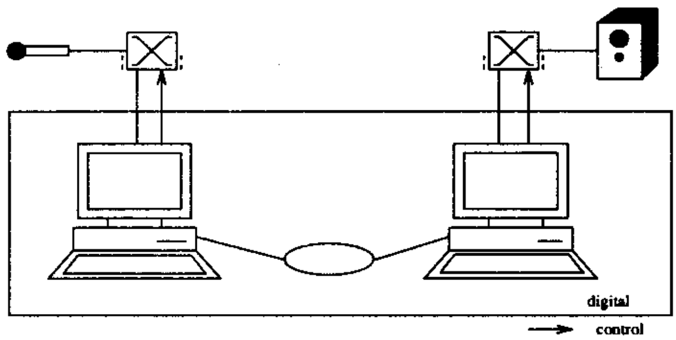

Figure 8.4: *Continuous data of local, external audio-video devices transmitted through the computer.*

- Computers control external devices.
- The continuous data are processed externally, as well as in the computer. This leads to diminished functionality and additional system software complexity because continuous data transferred between external audio-video-devices have a different processing environment than continuous data transferred, for example, from the hard disk to a video window.
- Before communication between different computers occurs, audio-video signals are digitized, coded and compressed (if possible).
- Synchronization of different media within a digital environment causes problems (see Chapter 15 on synchronization).

An example of this approach is the DIME project [RSSS90] of the IBM European Networking Center. Here, audio-video devices are connected to PS/2 computers through audio and video cards [Moo90] and transmitted over the data network.

Another example is the *Boulder Project* from US West, Colorado [CSA+89]. The research department is geographically split into two locations and between them a data connection of 45 Mbit/s is installed. Both places have a configuration of audio-video devices similar to IMAL. Between both laboratories, data of different media are transmitted.

8.1.2 Digital Systems

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.

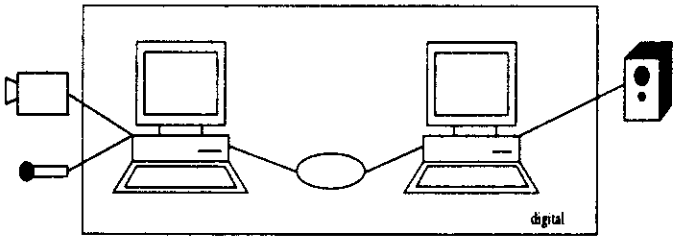

Figure 8.5: *Integrated (with respect to hardware) system structure with analog end-system devices and A/D and D/A interfaces.*

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

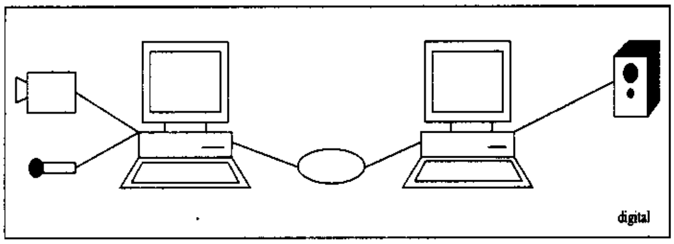

Figure 8.6: *Integrated (with respect to hardware) system structure with digital end-system devices and interfaces.*

system.

Connection to Switches

Another possibility to connect audio-video devices to a digital network is to connect them directly to the network switches. An example is the VuNet Asynchronous Transfer Mode (ATM) network, implemented by MIT's Telemedia, Networks and System Group [TAC+94]. Its configuration is shown in Figure 8.7.

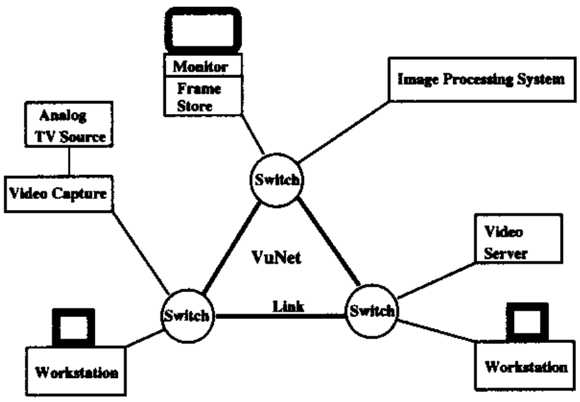

Figure 8.7: *The VuNet configuration.*

8.2 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.

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.*

The Silicon Graphics workstation, running IRIX with $REACT^{TM}$ [Ska94] (shown in Figure 8.8), is very close to a multimedia workstation with the above-mentioned components.

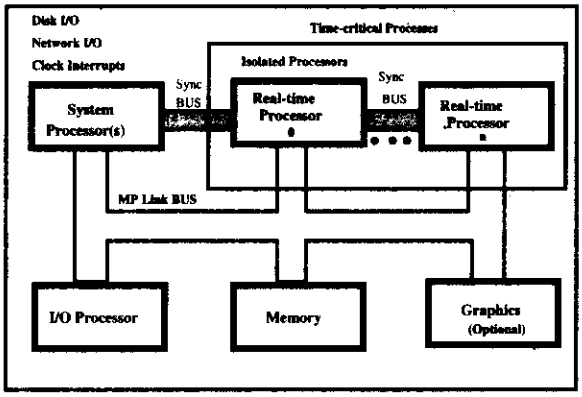

Figure 8.8: *Silicon Graphics workstation architecture (POWER Lock Processor Isolation).*

Bus

Within current workstations, data are transmitted over the traditional asynchronous *bus*, meaning that if audio-video devices are connected to a workstation, continuous data are processed in a workstation, and the data transfer is done over this bus, which provides low and unpredictable time guarantees. In multimedia workstations, in addition to this bus, the data will be transmitted over a second bus which can keep time guarantees. In later technical implementations, a bus may be developed which transmits two kinds of data according to their requirements (this is known as a *multi-bus system*).

The notion of a bus has to be divided into *system bus* and *periphery bus*. In their current versions, system busses such as ISA, EISA, Microchannel, Q-bus and VME-bus support only limited transfer of continuous data. The further-development of periphery busses, such as SCSI, is aimed at the development of data transfer for continuous media.

Multimedia Devices

The main peripheral components are the necessary input and output multimedia devices. Most of these devices were developed for or by consumer electronics, resulting in the relative low cost of the devices. Microphones, headphones, as well as passive and active speakers, are examples. For the most part, active speakers and headphones are connected to the computer because it, generally, does not contain an amplifier. The camera for video input is also taken from consumer electronics. Hence, a video interface in a computer must accommodate the most commonly used video techniques/standards, i.e., NTSC, PAL, SECAM with FBAS, RGB, YUV and YIQ modes. A monitor serves for video output. Besides Cathode Ray Tube (CRT) monitors (e.g., current workstation terminals), more and more terminals use the color-LCD technique (e.g., a projection TV monitor uses the LCD technique). Further, to display video, monitor characteristics, such as *color*, *high resolution*, and *flat and large shape*, are important.

Primary Storage

Audio and video data are copied among different system components in a digital system. An example of tasks, where copying of data is necessary, is a segmentation of the LDUs or the appending of a *Header* and *Trailer*. The copying operation uses system software-specific memory management designed for continuous media. This kind of memory management needs sufficient main memory (primary storage). Besides ROMs, PROMs, EPROMs and partially static memory elements, low-cost dynamic memory modules are especially needed. The steadily decreasing cost of these modules, together with steadily increasing storage capacities, profits the multimedia world.

The copying of data in a workstation can be performed using either the bus master method or the Central Processing Unit (CPU), although it is not always possible to use the former method. In most cases, the copying operation is performed by the CPU.

Secondary Storage

The main requirements put on secondary storage and the corresponding controller are a high storage density and low access time, respectively.

On the one hand, to achieve a high storage density, for example, a Constant Linear Velocity (CLV) technique was defined for the CD-DA (Compact Disc Digital Audio) (see Section 7.4). CLV guarantees that the data density is kept constant for the entire optical disk at the expense of a higher mean access time. On the other hand, to achieve time guarantees, i.e., lower mean access time, a Constant Angle Velocity (CAV) technique could be used. Because the time requirement is more important, the systems with a CAV are more suitable for multimedia than the systems with a CLV.

Further requirements put on secondary storage are a high capacity at a low price and a storage of discrete and continuous media together.

Processor

In a multimedia workstation, the necessary work is distributed among different processors. Although currently, and for the near future, this does not mean that all multimedia workstations must be multi-processor systems. The processors are designed for different tasks. For example, a *Dedicated Signal Processor (DSP)* allows compression and decompression of audio in real-time. Moreover, there can be special-purpose processors employed for video. Figure 8.9 shows an example of a multi-processor for multimedia workstations envisioned for the future (Intel '786 [Pre90]).

As a future development, a multimedia system could consist of multi-universal processors, which would have DSP properties unified with CPU properties. According to the application, these elements could be configured for exclusive processing of discrete or continuous media. Hence, this would provide the possibility of a multimedia workstation which supports customization of processors according to an application specification.

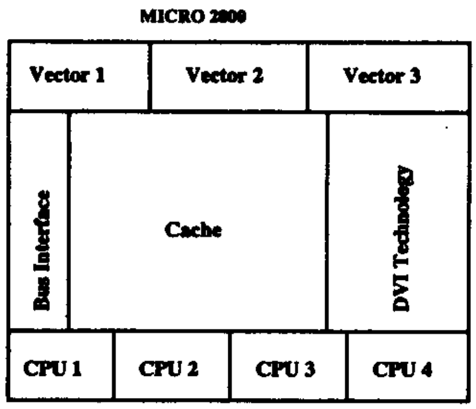

Figure 8.9: *Multi-processor system, Intel '786 [Pre90].*

Operating System

Another possible variant to provide computation of discrete and continuous data in a multimedia workstation could be distinguishing between *processes* for discrete data computation and for continuous data processing. These processes could run on separate processors. Given an adequate operating system, perhaps even one processor could be shared according to the requirements between processes for discrete and continuous data. The further development of multimedia operating systems is likely to go in this direction.

8.3 Comments

Digital systems consist, for most existing research and development projects, of traditional computers *enhanced* with hardware for multimedia functions. Examples of such enhancements are adapters connecting audio-video sources and destinations with the system bus in a workstation. These adapters do not place the data in the main memory during continuous media transfer, which would slow down the processing. Further improvements of traditional computers can be achieved by using compression methods. These methods considerably reduce the data rate of video, as well as that of audio (see Chapter 6).

Digital systems allow, in addition to the integration of audio-video devices, ***high flexibility*** because different multimedia devices can be programmed. For example, consider an audio mixer. With the help of signal processors, the mix functions are a set of filter functions implemented as programs. These programs could be modified easily in a digital environment and therefore customized according to any application requirements. The signal processor could also be used for other purposes. Another example represents DVI. DVI uses *Display Processor* VDP (Pixel Process), which is programmed in micro-code. The programs contain a part of the compression and decompression algorithm, which is loaded during the initialization phase. Therefore, some future extensions of DVI are possible without changing the underlying hardware.

Requirements, with respect to the performance capability of digital systems, do not always allow the use of software solutions in a digital environment. A hardware implementation mostly increases performance capability, but flexibility decreases. This boundary between dedicated hardware and general-purpose software components is dependent on technology, application and cost. Using a multimedia workstation architecture with a multimedia operating system, this boundary can be flexibly modified.