Disc Manufacturing, Inc.

A QUIXOTE COMPANY

INTEGRATING MIXED-MODE

CD-ROM

Acknowledgment

I wish to thank Pamela A. Sansbury and Nancy A. Klocko for their editing contributions to this paper.

WHO IS DMI?

Disc Manufacturing, Inc. (DMI) manufactures all compact disc formats (i.e., CD-Audio, CD-ROM, CD-ROM XA, CDI, PHOTO CD, 3DO, KARAOKE, etc.) at two plant sites in the U.S.; Huntsville, AL, and Anaheim, CA. To help you, DMI has one of the largest Product Engineering/Technical Support staff and sales force dedicated solely to CD-ROM in the industry.

The company has had a long term commitment to optical disc technology and has performed developmental work and manufactured (laser) optical discs of various types since 1981. In 1983, DMI manufactured the first compact disc in the United States. DMI has developed extensive mastering expertise during this time and is frequently called upon by other companies to provide special mastering services for products in development.

In August 1991, DMI purchased the U.S. CD-ROM business from the Philips and Du Pont Optical Company (PDO). PDO employees in sales, marketing and technical services were retained.

DMI is a wholly-owned subsidiary of Quixote Corporation, a publicly owned corporation whose stock is traded on the NASDAQ exchange as QUIX. Quixote is a diversified technology company composed of Energy Absorption Systems, Inc. (manufactures highway crash cushions), Stenograph Corporation (manufactures shorthand machines and computer systems for court reporting) and Disc Manufacturing, Inc.

We would be pleased to help you with your CD project or answer any questions you may have. Please give us a call at 1-800-433-DISC for pricing or further information.

Nancy Klocko
302-479-2500
Fax: 302-479-2527

We have four additional technical papers available entitled

Compact Disc Terminology - 2nd Edition

An Overview to MultiMedia CD-ROM Production

Introduction to ISO 9660

A Glossary of CD and CD-ROM Terms

These are available upon request
800-433-DISC
302-479-2500
Fax: 302-479-2527

Integrating Mixed-Mode CD ROM

by
J. Philip Busk
Disc Manufacturing, Inc.
2/6/92

Imagine using an encyclopedia application and not only being able to look at a picture of a lion, but also being able to hear it roar. This is what mixed-mode CD ROM is all about; sound enhancing your application. Mixed-mode CD ROM integrates CD ROM data and CD Audio on the same disc. The thought of using sound to enhance your application is very appealing; but how do you go from the simple idea of adding music or voice-overs to your data, to really making it happen?

Background

To include audio with data in a mixed-mode application, it must first be in a format that can be used by computers, audio players or specialized players. Some of the audio formats you may be familiar with are digitized audio, Adaptive Pulse Code Modulation (ADPCM), and CD Audio (sometimes referred to as Red Book Audio, taking the name from the CD Audio specification).

All audio sounds can be represented by an analog signal, but analog signals cannot be stored on computer storage media. To store information on a computer, it must be in digital form. To capture an analog signal in digital form, the information must first be “sampled”. Sampling is a term used to describe how an analog signal is recorded in evenly spaced intervals. Each of these samples is assigned an integer value that reflects the sample's amplitude. The number of integer values or bits assigned to the amplitude is generally referred to as the sampling resolution.

The process of converting analog signals into digital form is known as digitizing. A digitized sound is an approximation of an analog sound. The quality of digitized audio is improved as more samples per second (higher sampling rate) are used to define the audio signal and more bits are used to define the sample (higher resolution). Figure 1 shows an analog signal and three digitized representations of that signal. As the sampling frequency and resolution are increased, the digitized representation more closely resembles the original analog signal.

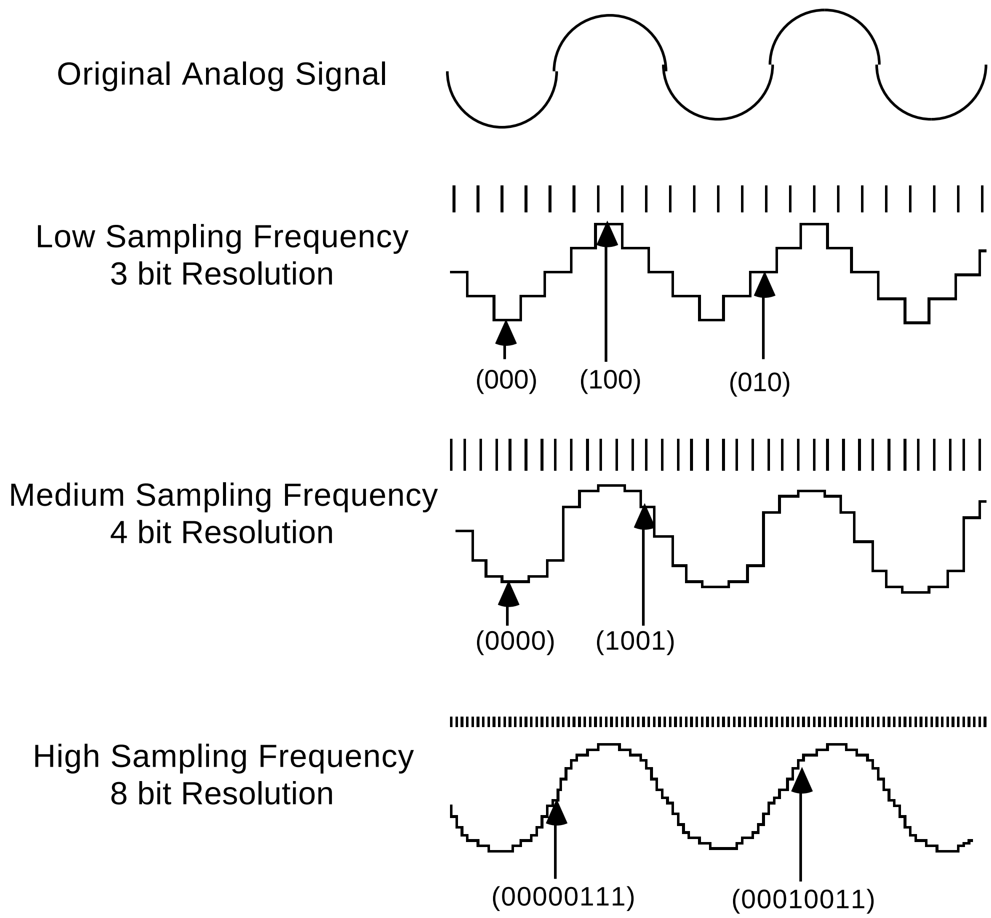

Figure 1: Digital Sampling

Digitized Audio

Digitized audio is treated by a computer as a data file. That is, as far as the storage media is concerned, digitized audio is just another file to be stored and accessed by the computer. There are several plug-in boards on the market which allow an IBM compatible PC user to digitize analog sound and store the information, in data form, on a magnetic disk. To play back digitized audio you need an audio playback card on your PC. The same applies to users of Apple computers, where some models have built-in audio recording and playback capabilities.

ADPCM

ADPCM audio is a compression scheme used by the Compact Disc Interactive (CD-I) and CD ROM Extended Architecture (CD ROM XA) formats. In the simplest terms, the audio file structure is interleaved with the data. CD-I offers three sound quality levels for compressed audio. A-Level is equivalent to a high quality recording. B-Level is equivalent to the very best FM radio broadcasts and C-Level is equivalent to the very best AM radio broadcasts. ADPCM playback requires special hardware, built into CD-I and CD ROM XA players, to separate the audio from the data.

CD Audio

CD Audio is very high quality digitized audio. Some inexpensive "digitizers" use a sampling rate of 44.1KHz (many use 22KHz and lower), but they usually only sample with 8 bits of resolution. The universal CD Audio standard uses a sampling rate of 44.1 KHz and a resolution of 16 bits. A comparison of the different quality levels and the amount of storage space needed for each level is shown in Figure 2.

Quality Level	Sampling Rate (KHz)	Resolution (bit)	Bytes Needed to Store 1 Minute of Stereo Audio (Mb)
CD Audio	44.1	16	10.09
ADPCM Level A	37.8	8	4.33
ADPCM Level B	37.8	4	2.16
ADPCM Level C	18.9	4	1.08
Digitized 22KHz	22	8	2.52
Digitized 11KHz	11	8	1.26

Figure 2: Audio Quality Level Comparison

CD Audio can be played back from any audio-capable CD ROM drive; therefore, no additional audio playback cards are required on the computer. In fact, a program capable of controlling the CD ROM player can tell the player where to start and stop playing audio segments giving the user interactive control. CD Audio is a good way to deliver high quality audio without requiring your customer to acquire additional hardware (playback card).

Because CD Audio is a format that will play on any CD ROM drive with audio capability, it offers virtually any computer high-quality, random-access audio without the need for additional hardware (as previously required for digitized audio) other than speakers or headphones. Unlike ADPCM, CD Audio is not interleaved with data on the disc. When planning a mixed-mode disc, remember the CD player can only access one track at a time.

The integration of CD Audio on CD ROM, concentrating on accessing CD Audio by Track, Absolute Time (A-Time), or by Track Relative Time will be discussed in the following paragraphs.

Using CD Audio in Mixed-Mode Applications

The easiest way to learn how to use CD Audio (also referred to as sound and/or audio in this article) in a mixed-mode application is with an example. In the example, we'll display pictures of animals and let the user hear what each animal sounds like. The animals in this mixed-mode application are a cow, cat, and a pig. We'll use a simple graphics program to display the animal pictures and an audio subroutine that addresses the animal sounds on the CD.

In a mixed-mode application, CD Audio is addressed in segments. Each of the segments represents narration, music, or a combination of narration and music that can be played in conjunction with displayed pictures or text. The audio segments for the example consists of; "moo", "oink", and "meow".

When a mixed-mode application is completed, the data or CD ROM portion of the disc resides on Track 1. CD Audio segments reside on tracks two through ninety nine.

A diagram of a mixed-mode disc is shown in Figure 3. Notice the data is on the first track and the audio follows on a contiguous spiral track that starts in the middle of the disc.

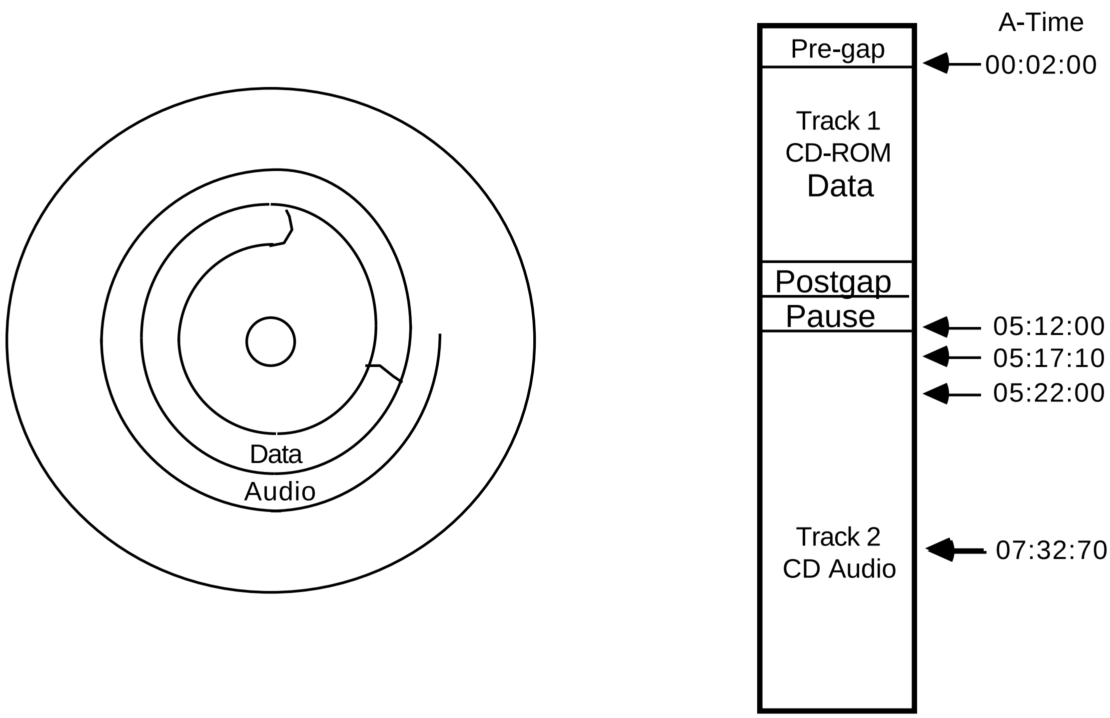

Figure 3: Mixed-mode disc spiral track

Accessing CD Audio using A-Time

A-Time is a timecode which registers the elapse time from the beginning of a CD. This timecode is based on the fact that in each second, 75 sectors of CD information (data or audio) are being played. Hence, the A-Time on a CD is expressed in minutes, seconds and CD sectors, the latter running from 0 through 74 in one second. Accessing audio on a CD or CD ROM by A-Time is not possible on most consumer audio players, but with the use of special software routines, the audio can be addressed on audio capable CD ROM players.

A-Time makes it possible for a computer application to access and play specific audio selections. CD Audio segments are normally addressed by the audio subroutine with a start and stop time. The start and stop times of each of the sounds on the CD ROM become the addresses of the sounds. In the example, "moo" has a start time of 05:12:00 (five minutes and twelve seconds) and a stop time of 05:14:30. "Oink"has a start time of 05:17:10 (Five minutes, seventeen seconds and 10 sectors) and a stop time of 05:20:05. If a picture of a cow was displayed and "moo" was to be played, the application's audio subroutine would address and play "moo" by telling the CD ROM player to play the CD from 05:12:00 to 05:14:30. Part B of Figure 3 shows the spiral track of data and audio stretched into a straight line indicating where each of the audio segments is addressed.

Integrating A-Time into your Application

Before a complete mixed-mode CD ROM can be developed, audio addresses need to be determined and programmed into the application. The best way to determine A-Time addresses is with a CD that already contains the audio portion of your program. This means the audio needs to be on an initial CD before the mixed-mode application can be completed.

This procedure has been made much simpler and less expensive in recent years because of CD write once (CD-WO) technology, better known as one-off discs. One-off discs are write once CDs that play just like regular CD ROMs but are very inexpensive compared to mastering and replicating discs. The audio portion of your application is sent to a mastering facility and put onto a one-off disc (data track size needs to be predetermined, see below). The one-off is returned and used to determine A-Time addresses. After the A-Time addresses are found, you can send your completed application (audio addresses programmed) to the mastering facility, where the application will be combined with the audio portion of the disc.

In the animal sounds example, we would first record "moo", "oink", and "meow" on either DAT or Sony 1630 U-matic tape. Although audio can be recorded on standard 1/4" analog tape, it is preferable to convert the tape to U-matic or DAT. The farther away you get from the preferred input (U-matic), the more it will cost (both in time and money) to get the audio into a digital form acceptable for CD mastering.

It is recommend that in the recording phase, you maintain at least two seconds of audio silence between individual audio segments that you plan to access separately within a single track. These two seconds of silence are used to facilitate CD ROM drive addressing accuracy. CD specifications dictate that each track must be at least four seconds long. Figure 4 shows a layout of the audio for the animal sounds example.

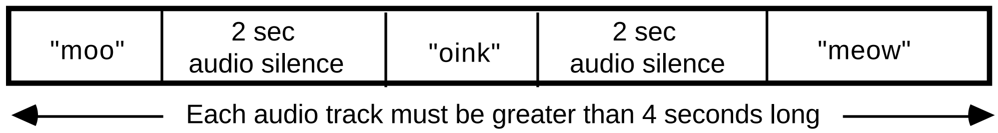

Figure 4: Animal sounds audio layout to be recorded on a single track

The recorded audio is sent to the mastering facility where it is recorded on a one-off disc. Preliminary data and applications can be placed on Track 1 at this time. Remember that the Track 1 size will increase when the audio access points are included in the data; this will change the addressing points unless the data track length is predetermined.

The length of the data track can be predetermined by using this calculation: 8 to 9 MB of data require about 1 minute of CD time. For example: if your data and graphics program is 50 MB in size, Track 1 will be approximately 6 minutes long. You can use this calculation to simply predetermine the start of your audio track to a set value.

A typical mixed-mode production flow chart is shown in Figure 5 on the next page.

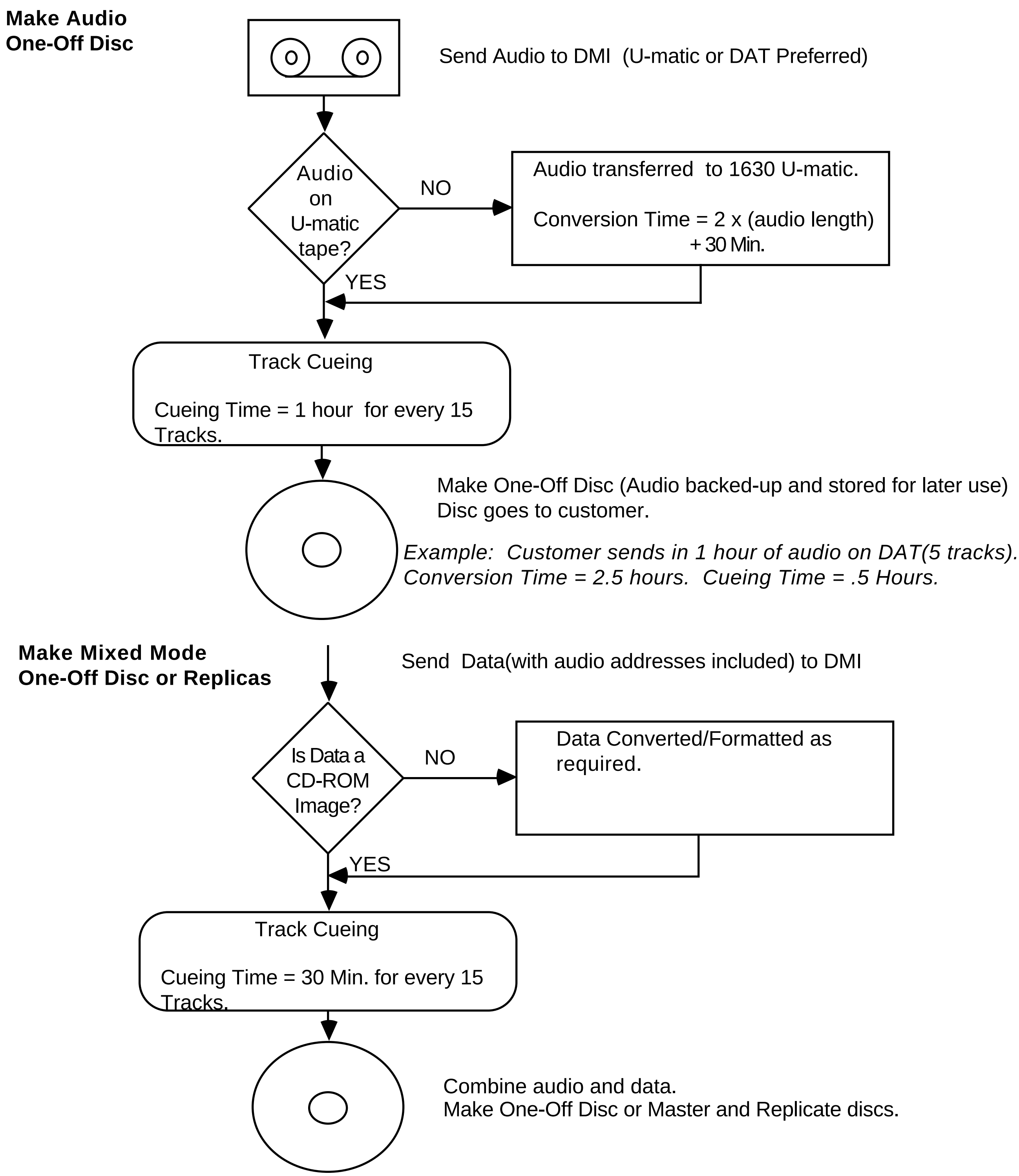

Figure 5: Mixed-mode production flow chart

Determining A-Time Addresses

The A-Time address of an audio segment is a combination of an offset factor added to the absolute time of the audio segment from the start of a specific track (track relative time). The offset factor is the elapsed time of your first track (the data track) plus 6 seconds for pre-gap, postgap, and pause periods. A pause period is a required specification that is added to the beginning of the audio track. This is illustrated in Figure 6.

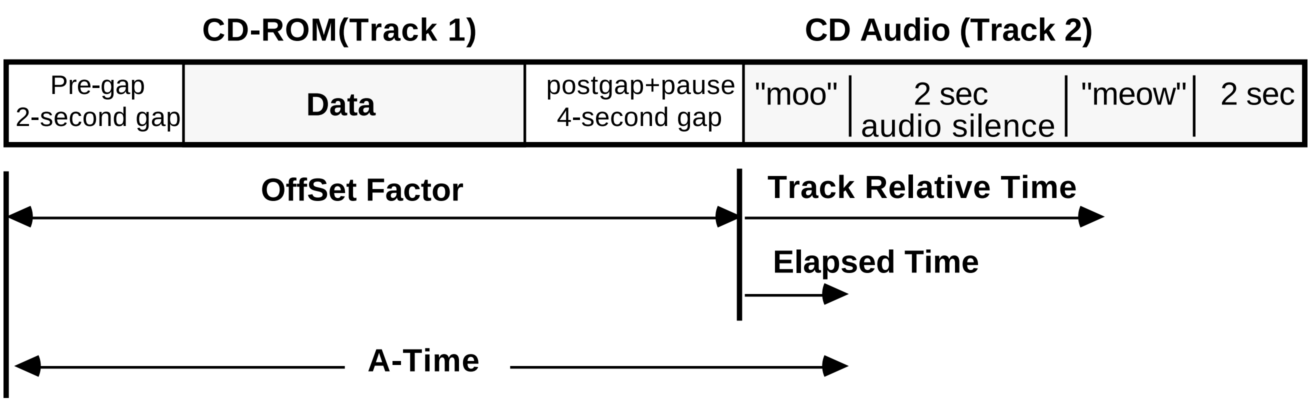

Figure 6:Diagram of Mixed-Mode CD ROM layout

Determining the Offset Factor

To determine the A-Time addresses of your audio segments, you can use a CD Audio player or audio capable CD ROM drive. With the disc loaded, switch your display so the disc's total elapsed time is displayed. To determine the offset factor, you must scan or play through the first track, which contains the data. Playing data can produce a very loud noise that may damage your speakers. Before playing a data track, TURN DOWN THE VOLUME OF YOUR AUDIO SYSTEM. At the end of Track 1, the counter will stop, and your system will switch to Track 2. You must determine the elapsed time of Track 1 right before the player switches to Track 2. It may be easier to immediately switch to Track 2, scan backwards a few seconds (switching to Track 1), then scan forward to determine the elapsed time of Track 1. Add 6 seconds to this time (for pre-gap, post-gap, and pause), and you have calculated the offset factor.

The elapsed time of each audio segment, combined with the offset factor, will determine the final A-time addresses for each segment.

Determining Elapsed Time

After scanning or playing to Track 2, which is where the audio program will start, press your PAUSE button. Verify that the TRACK number is “2” and that the TIME counter displays zero time (00:00:00). Keep your display set to TIME. From this point on, the elapsed time from the start of Track 2 is displayed. Press the PLAY button to start your segment.

As you move from one audio segment to the next, press the PAUSE button between every segment and record a start or end timecode. Later on, you can fine tune by playing the sector values on a CD ROM drive using the audio subroutine and see if this provides a satisfactory addressing point for your audio program. You can experiment with the sector values until the proper start and end codes have been determined. For example if "moo" ends at 5:14:30, using a CD Audio player, you may have determined the end time of "moo"to be 5:15:00. By fine tuning on a CD ROM drive (a CD ROM drive will allow addressing by the sector while most audio players only allow addressing by the second), you can change the stop time to 5:14:30 and eliminate a half second of delay.

The result of this exercise will be a table of all the audio segments start and end values. The final step in determining A-Time addresses is to add the offset factor to the elapsed times. These values can now be entered into your application. The audio access software will use these values to correctly access the audio segments from the mixed-mode CD. The offset factors and A-Time addresses of the animal sounds example are shown in Figure 7.

Sound	Offset	Start from Track 2	A-Time Address Start	Stop
"moo"	05:12:00	00:00:00	05:12:00	05:14:30
"oink"	05:12:00	00:05:10	05:17:10	05:20:05
"meow"	05:12:00	00:10:00	05:22:00	05:25:60

Figure 7: Animal sounds example A-Time Addresses

Accessing CD Audio using Tracks

Tracks on a CD mimic the visual track separations found on conventional record albums. Like A-Time, tracks also make it possible to address specific audio segments.

Integrating Track Access into your Application

Track addressing has some limitation (discussed in the trade-off section of this article), but if appropriate, it is probably the easiest way to access individual audio segments. Each addressable audio segment is simply accessed as a track. In other words, if twenty audio segments will need to be individually accessed, they would each go into a separate track starting from Track 2 (Track 2 - Track 21).

If your application will access the sounds "moo", "oink" and "meow"; "moo" would be located on Track 2, "oink" would be located on Track 3 and "meow" on Track 4. To play the sound "moo", the application would simply instruct the CD ROM drive to play Track 2, to play "oink" play Track 3, etc. The track layout for the example is shown in Figure 8.

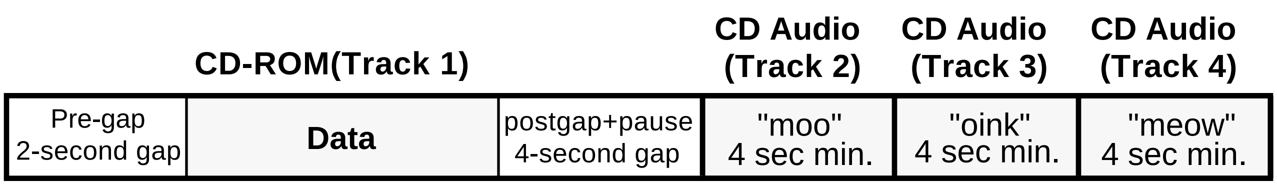

Figure 8: Track layout of animal sounds

Accessing CD Audio using Track Relative Time

Track Relative Time is a combination of track and A-Time. Instead of addressing the elapsed time from the beginning of the disc, like A-Time, you can address the elapsed time from the beginning of a specific track. Many mixed-mode CD ROM's rely on track relative time addressing so if the size of the data track changes, it does not effect the audio access times. Audio accessed by A-Time will require recalculating and reprogramming the audio access points when the size of the data track changes. A comparison of A-Time and Track Relative Time start and stop calculations is shown in Figure 9.

The start and stop times for the first selection in Part A of Figure 9 are:

Track Relative Times:	Start at 00:00:00 and End at 00:05:10
A-Time:	Start at 05:12:00 and End at 05:17:10

If the size of the data track increases by 4 minutes (shown in Part B of Figure 9), and the application needs to play the same audio section as the example above, then it would play the following times:

Track Relative Times:	Start at 00:00:00 and End at 00:05:10
A-Time:	Start at 09:12:00 and End at 09:17:10

Audio access points in an application that use track relative time addressing do not need to be changed when the size of the data track changes.

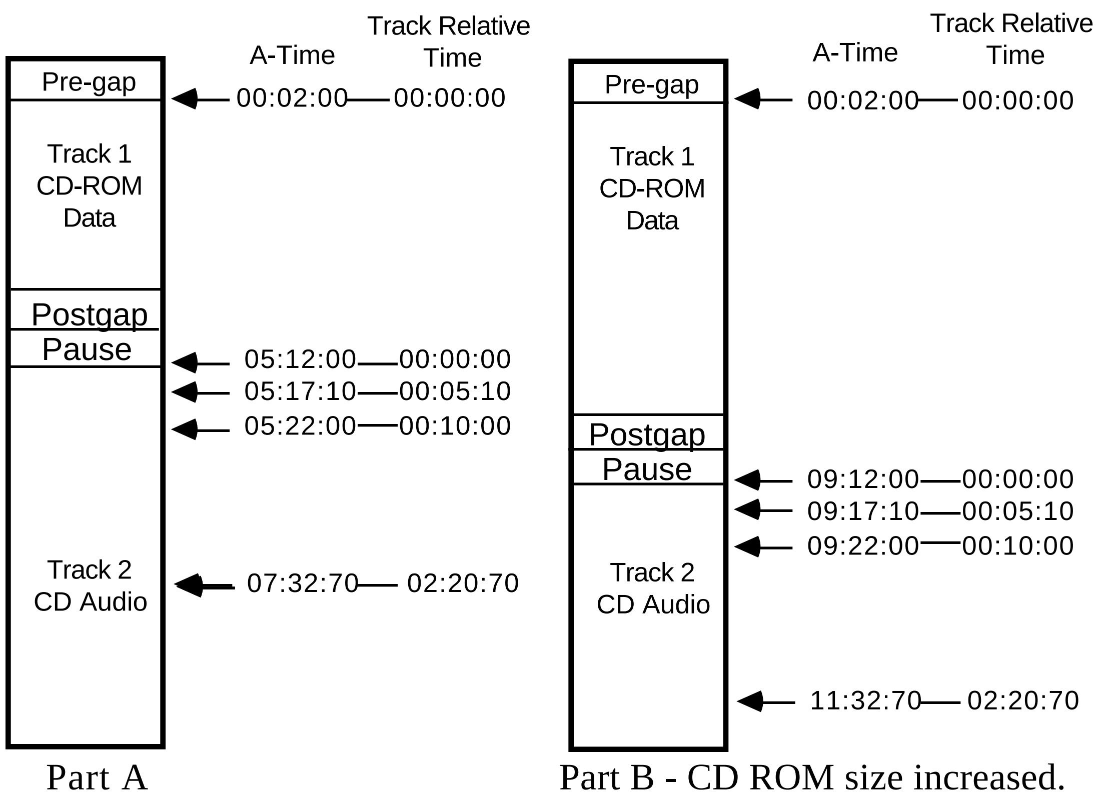

Figure 9: Track layout for a mixed-mode disc before and after the CD ROM track size is increased.

Integrating Track Relative Time

Integrating track relative time is similar to integrating A-Time except that the access points do not change when the data track size changes. Using the "determining elapsed time" procedure described previously, you can switch to Track 2 and determine your times from this point. The elapsed time becomes the track relative time. There is no need to determine the offset factor for track relative addressing. In Figure 6 you can see that the track relative time and the elapsed times are the same.

Trade-Offs

Before beginning any project, some key issues need to be addressed. Some of these issues include: Defining the delivery system, how will the data be captured, what type of computer will the customer base be using, will most of the customer base have an audio playback card (for digitized audio playback), or should CD Audio be used.

If CD Audio is used, the choice of which CD Audio access method to use is based on trade-offs. Some of the trade-offs to consider are:

1) A-Time access points will change if the data size changes. Track Relative Time access points will not change.

2) Unless you use track access you will need to make two discs; a preliminary disc used to determine A-Times or track relative times and a final disc with your completed application. The data size of the second disc will change when the application is either added or modified.

3) Track locations on a CD need to be cued. Cueing is the way the mastering facility adds a "marker" to the CD so that the CD player can determine the location of the tracks. It takes about 1 hour of work for every 15 tracks to be cued. A general rule of thumb is that any project using over 20-30 tracks should use track relative timing.

4) A-Time addressing requires more editing time by the mastering facility.

One important trade-off to keep in mind is that 1 minute of audio will take up about the same amount of space on a CD ROM as 9Mb of data. With 100Mb of data on the first track, there is still room for up to 1 hour of full CD quality, stereo sound. Figure 11 shows this trade-off for different sized data tracks. Data must be located on the first track of a CD ROM disc.

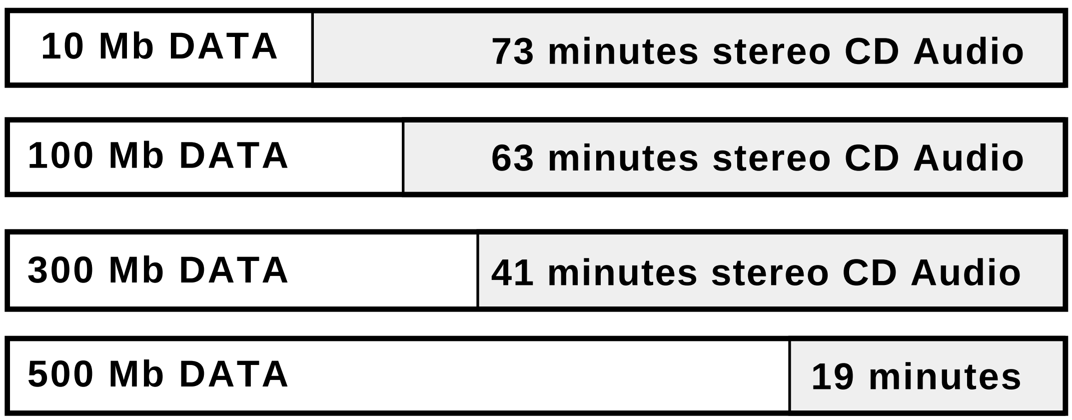

Figure 11: Trade-off between data and CD Audio capacity

Conclusion

CD ROM, with its combined storage potential for both data, graphics and audio in a random-access environment, can be an unparalleled carrier of multimedia information. Now that you have a little better understanding of how to use CD Audio on CD ROM, its time to get your applications working with audio.

We look forward to answering your questions and helping you with your mixed-mode CD ROM project.

Call 1-800-433-DISC or 1-302-479-2500
Disc Manufacturing, Inc.
1409 Foulk Road, Suite 102
Wilmington, DE 19803