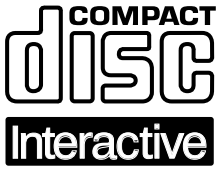


OptImage Interactive Services Company, L.P.

An Overview of CD-I Standards

OptImage Technical Bulletin 001

Jon W. Bäckström

Acknowledgments

I would like to gratefully acknowledge the important contributions made by OptImage and Philips personnel involved in the editing and review of this document. The high degree of technical accuracy in this document reflects the insightful feedback and patience offered by a number of dedicated CD-I professionals.

Disclaimer

The information contained herein is believed to be accurate as of the date of publication, however, OptImage will not be liable for any damages, including indirect or consequential, from reliance on the accuracy of this information. The information contained herein is subject to change without notice.

Trademarks

CD-I is a trademark of N.V. Philips and Sony Corporation

All other brand and product names are marks or registered marks of their respective companies.

Copyright

Copyright © 1993 OptImage Interactive Services Company, L.P. All Rights Reserved.

Reproduction of this document, in part or in whole, by any means, electrical, mechanical, magnetic, optical, chemical, manual or otherwise, is prohibited without written permission from OptImage Interactive Services Company, L.P.

Publication Editor:	Jon W. Bäckström
Publication Date:	January 8, 1993
Documentation Part Number:	OPT901
Revision:	A

OptImage Interactive Services Company, L.P.
1501 50th Street, Suite 100
West Des Moines, Iowa 50266

Phone Numbers:

In the USA: (515) 225-7000
(515) 225-0252 FAX
Sales Information:
(800) CDI-5484
Technical Support:
(515) 224-1234

In Europe: +31 (40) 73 6228
In Japan: (03) 3 665 9740

Online Information:

Internet: info@optimage.com
sales@optimage.com
support@optimage.com
Applelink: D6431
America Online, Keyword: OptImage
OptImage BBS: (515) 225-1933

Table of Contents

An Overview of CD-I Standards

CD-I Players

CD-I players are essentially small, integrated computer systems equipped with a sophisticated set of graphics and audio processors. CD-I players are capable of playing a variety of 5.25" optical media standards:

- CD-I: Program Discs
- CD-DA: Conventional Digital Audio Discs
- CD+G: Digital Audio Discs with Graphics
- Kodak Photo CD: Digital Photographic Images

CD-I players are designed as a world-wide standard, bridging differences between international broadcast systems. The data on any given CD-I disc should play the same way on another CD-I machine, anywhere else in the world.

Types of CD-I Players

There are three main types of CD-I players: consumer, industrial, and development. Consumer CD-I players are meant for use in the home or office and typically feature a minimum hardware configuration (in this case, they are referred to as “Base Case players”). Additional capabilities beyond the Base Case are referred to as “extended” features.

Industrial CD-I players typically offer additional features such as increased memory capacity, sturdier disc transports, floppy disk drives, and more input/output flexibility. Finally, development CD-I players are intended for use in the title authoring process and feature the ability to attach various peripheral devices needed in the development process, such as workstations, ethernet connections, hard disks, and emulators. (Emulators allow development CD-I players to test disc images by sending data to the player directly from a hard disk, rather than from optical media.)

CPU and Memory

Every CD-I player uses a Motorola 680x0 microprocessor and runs CD-RTOS, a real-time operating system which is stored in the player's ROM. Base Case CD-I players contain 1 Mb of RAM. This can be a limitation since video display data typically uses most of the available memory in a Base Case player, leaving less memory for program code and data. Industrial and development CD-I players, on the other hand, typically contain 4 or 8 Mb of RAM.

Full-Motion Video (FMV) display hardware is a new extended feature that adds an additional 512K of memory to the player. When an FMV data stream is not actively being decoded, this extra memory can be put to general use by the CD-I player.

User Control Devices

Users typically interact with a CD-I player through a pointing device, such as a thumbstick controller, a mouse, a trackball, or a touch-screen. It is also possible to use keyboard input for CD-I through the use of an RS-232 keyboard attached to a controller port.

Pointing devices may control the CD-I player using infa-red remote control or by attaching directly to the controller or serial ports.

Data Input/Output

Base Case consumer CD-I players may include one or two controller ports, but they do not generally have an extensive array of connectors for other types of external input/output (I/O). Industrial and development players, on the other hand, may offer a number of I/O connectors. For example, the Philips 605 development player features 5 Mb of memory, 2 controller ports, 2 serial ports, a parallel printer port, a 3.5" floppy disk drive, an ethernet connector, a SCSI port (for hard disks or other devices), an emulator interface for software testing, and an expansion slot for additional system extensions, such as Full-Motion Video.

Most CD-I players display video on a television screen or an RGB monitor. There are also a number of hand-held CD-I players that use an integrated color LCD screen to display video and communicate with the user. Often these hand-held players also feature an external video jack.

Audio and Video Output

CD-I players typically offer a single set of stereo outputs, which connect to an amplified speaker system. Most players are designed to integrate with a home entertainment system with the video outputs connecting to a TV or monitor using standard cables (S-video or coaxial) and the audio outputs connecting to a stereo sound system.

Video Planes

All CD-I players support four video planes: the Cursor plane, two general video planes for the display of images, and a background plane, which can be used for Full-Motion Video.

Cursor Plane

The cursor plane is a 16 x 16 pixel area that is used to display the pointer for user input. It is a single color and may be enabled or disabled. The cursor plane is always at the top level of the display, but can be temporarily disabled if the need arises.

Video Planes A and B

All CD-I video image formats can be loaded into either video plane except for CLUT8, which can only display on Plane A, and RGB555, a special high resolution format that requires both video planes. The two video planes can be mixed together using a built-in, programmable video synthesizer. The video planes can also be arbitrarily set in front or in back of each other and final display images can be made to dissolve between the two planes by using a variable image contribution factor.

In addition, one video plane can display images over another with a transparency effect. A palette entry can be selected to define a transparent color and any pixels of that color on the front video plane will display the corresponding pixel from the second video plane. For example, a cartoon animation might define its background color as being transparent, making the cartoon appear on top of another image in a second video plane.

Matte effects are essentially “holes” that are programmed into the front video plane, allowing an image from the back video plane to be seen through them. An example of an interesting matte effect might be to produce two video images of a control panel, one drawn with buttons up and the other with buttons down. Later, when a hotspot is defined and selected from the top image, a matte effect could display that specific part of the lower image under the area of the hotspot, giving the effect of an animated button control.

Background Plane

The background plane, or backdrop, can be used to display Full-Motion Video (FMV), an external video source, or a constant color. FMV uses specialized MPEG decompression hardware to display either full-screen or partial-screen video sequences on the background plane. The MPEG hardware can also process the FMV data stream to simultaneously provide audio data. By turning off the cursor and both video planes, a CD-I application can display FMV sequences without any other graphic element present on the display.

Alternatively, the video planes can also be used to display graphic images over an FMV sequence, providing graphic images superimposed over the video – also known as a genlock effect. This capability might be used to graphically label parts of a full-screen video frame. For example, while a person watches and listens to an FMV sequence of an orchestra playing, graphic displays from the video planes might display over the video, labeling the instruments in various parts of the orchestra or identifying a soloist.

Some CD-I players provide the capability to use the background plane as a display for video from an external source, such as a laserdisc player or a live video camera. The capability to display external video on the backdrop is part of the Base Case specification, but it is left to the CD-I player manufacturer to optionally provide the necessary synchronization hardware and a video input jack.

Video Standards

All CD-I players are designed to use the same data format and support the two basic video standards existing in the world today: NTSC and PAL/SECAM. This means that you can place a video sequence or animation on a CD-I disc and have it display properly on both NTSC and PAL television monitors the world over.

Video Standard	Normal Screen Resolution
NTSC Monitor	360 x 240
NTSC TV (525 lines / 60 Hz)	384 x 240
PAL/SECAM TV (625 lines / 50 Hz)	384 x 280

NTSC (525 Line Video)

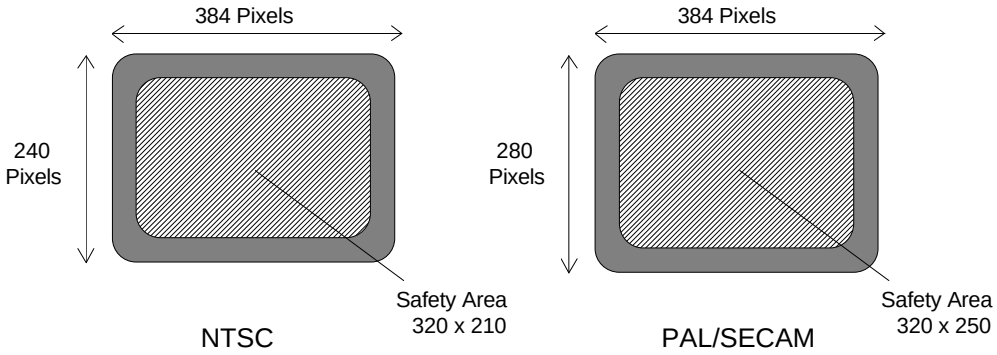
NTSC video is the standard in the United States, Canada, and Japan. NTSC televisions display CD-I images with a normal resolution of 384 pixels per line and 240 lines. The NTSC monitor mode, however, has a resolution of 360 x 240 pixels.

PAL/SECAM (625 Line Video)

The PAL format is the standard in Europe, Australia, and South America. PAL images for CD-I are displayed with a normal resolution of 384 pixels per line and 280 lines. The SECAM video standard is used in France, the former Soviet Union, and much of eastern Europe, but is considered the same as PAL for the purpose of explaining CD-I.

Safety areas

Television sets do not always display the entire video frame, and parts of a video image may actually scan beyond the visible edge of screen. For this reason, it is wise to keep important display elements within a "safety area" in the center of the display. This will insure that important information or controls are not hidden from the user on the edges of the display.



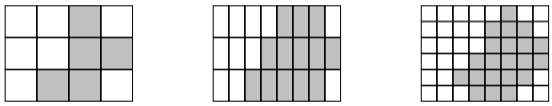
Universal compatibility

In order to insure that an image or animation will display properly on any CD-I player in the world, it is usually best to create all displays in PAL resolution, but observe an NTSC safety area. In other words, create a 384 x 280 display image but observe a 320 x 210 safety zone when designing the image content.

Resolutions

The resolution of a video image depends on the number of lines used to make up the television picture. To achieve world-wide compatibility, CD-I video images must be displayed using slightly different screen resolutions under the NTSC and PAL video standards.

Three resolutions exist for CD-I images. Most CD-I video modes use "normal resolution". "Double resolution" video modes double the number of pixels in the width, while "high resolution" video modes double the number of pixels in both the width and the height.



Video Standard	Normal Resolution	Double Resolution	High Resolution
NTSC Monitor	360 x 240	720 x 240	720 x 480
NTSC TV	384 x 240	768 x 240	768 x 480
PAL/SECAM TV	384 x 280	768 x 280	768 x 560

CD-I Coordinate Conventions

CD-I screen resolutions are commonly referenced by the 384 x 280 pixels normally supported in most video modes. Many authoring tools and software libraries, however, reference the CD-I screen in high-resolution, logical coordinates (768 x 560), a convention used by the CD-RTOS User Communications Manager (UCM). These "UCM coordinates" are then translated into the physical pixels supported by whatever mode the hardware happens to be set to.

In normal resolution formats, the 768 x 560 UCM coordinates can be addressed logically by software; but they are mapped into the 384 x 280 pixels that are actually physically addressable. Double resolution video formats, however, allow each horizontal pixel of the UCM coordinate system to be individually addressed.

CD-I Video Formats

The quality and complexity of an image is determined by the coding scheme used to produce it.

Images using a color look-up table (CLUT) assign a specific palette entry to each pixel on the display. Each CLUT entry in the palette specifies a particular mix of RGB components that determine the hue of all corresponding pixels. Since the size of the palette is finite, the number of palette entires determines the limits of multiple hue and image subtlety for a CLUT image.

DYUV images acheive a high degree of realism by specifying color as a change from the last pixel, rather than from a finite palette. This means that the coloring can be much more subtle but large shifts in color or luminance may require multiple pixels to acheive. For this reason, the DYUV video format is not always best for images that require crisp, detailed, one-pixel lines.

Standard Video Formats

The following table summarizes the standard video formats for CD-I players:

Video Format	Description	Possible Colors	Video Plane	Bits/ pixel	Resolution	Typical Use
RGB888	*Not CD-I*	16,777,216	- NA -	24	Normal	source encoding
RGB555	5-bit RGB	32,768	Both	16	Normal	detailed images
DYUV	Delta YUV	16,777,216†	A or B	16/2	Normal	"natural" images
CLUT8	8-bit CLUT	256	A	8	Normal	graphics
CLUT7	7-bit CLUT	128	A or B	7	Normal	graphics
CLUT4	4-bit CLUT	16	A or B	8/2	Double	graphics
RL7	Run Length	128	A or B	7	Normal	graphics / animation
RL3	Run Length	8	A or B	6/2	Double	graphics / animation

† Base Case CD-I players are only required to implement 6-bit RGB digital to analog converters in the video hardware. Because of this, DYUV images on Base Case players are usually displayed using a maximum of 262,144 colors [$(2^6)^3$]. The internal representation of DYUV images remains at a theoretical maximum of 16,777,216 colors, however, the same as a 24-bit RGB888 image.

QHY-B Images

High resolution DYUV images usually require five times the amount of memory than normal resolution DYUV images. The *Green Book* describes QHY (Quantized High Y) as a coding scheme that requires less data and, therefore, offers faster load times. At the same time, the QHY coding scheme offers a subjective quality that is close to that of full high resolution DYUV images. QHY requires extended hardware to display the images, however.

The QHY-B image format is a special simulated high-resolution DYUV mode for Base Case CD-I players. In QHY-B, the video hardware is set to interlace mode and one or two DYUV images are placed in the even and odd scan lines on one video plane. The second video plane uses RL3 or CLUT4 components on even and odd scan lines and the two video planes are combined with each other on the final display. This provides more detail and helps to reduce the flicker of the video interlacing.

CD-I Animation Formats

There are two basic types of animation in CD-I: run-length animation and DYUV movies. While both are flexible animation formats, the size, complexity, and speed of the animation is limited by the video data bandwidth of the real-time data stream.

Run-length Animations

Run-length animations are full-screen animations that are best suited for cartoon images or those involving a lower level of detail or image complexity. Two basic types of run-length animation formats exist:

Animation Type	Possible Colors	Maximum Resolution
RL7	128	384 x 280
RL3	8	768 x 280

Run-length data is encoded so that repeating horizontal pixels – a run – require only two bytes of data; one byte for the color look-up table entry (CLUT) and another to determine how many pixels to repeat (from 2 to 255). In RL3 formats, a pair of 3-bit CLUT entries are coded as one byte and a second byte specifies the length of the run for the pixel pair.

DYUV Movies

DYUV movies are small, partial-screen DYUV images displayed in rapid succession. DYUV movie animations are limited in size by how much video data, combined with an audio track, can be placed in the CD-I data stream. Since DYUV images can display over 16-million colors, they are often used to display recorded video clips and other "natural" images, DYUV movies are useful for displaying motion sequences with a high degree of detail and image complexity.

CD-I Audio Formats

CD-I audio uses an Adaptive Pulse-Code Moluation (ADPCM) coding scheme instead of the PCM coding used by CD-DA discs. ADPCM requires about half the amount of raw data, as compared with PCM coding, since only the differences from one sample to the next are recorded. This means that a CD-I disc can hold at least twice as much audio data as standard CD-DA discs. The remaining data path can be used for the video processors and program logic.

Audio Quality

There are three different CD-I audio levels (in both mono and stereo) and they provide tradeoffs in terms of audio quality and the amount of data they require.

Audio Level	Sample Rate	Encoding Method	Frequency Bandwidth	Data Ratio	Playing Time (Hours:Minutes)
CD-DA	44.1 kHz	PCM / 16 bits	20 kHz	1:1 - 100%	1:12
A Stereo	37.8 kHz	ADPCM / 8 bits	17 kHz	1:2 - 50%	2:24
A Mono	37.8 kHz	ADPCM / 8 bits	17 kHz	1:4 - 25%	4:48
B Stereo	37.8 kHz	ADPCM / 4 bits	17 kHz	1:4 - 25%	4:48
B Mono	37.8 kHz	ADPCM / 4 bits	17 kHz	1:8 - 12%	9:36
C Stereo	18.9 kHz	ADPCM / 4 bits	8.5 kHz	1:8 - 12%	9:36
C Mono	18.9 kHz	ADPCM / 4 bits	8.5 kHz	1:16 - 6%	19:12

Level A audio

Level A audio provides the highest quality audio source (37.8 kHz sample rate, 8 bit samples) While providing the best level of audio quality possible for CD-I titles, this format is not often used because of the large amounts of data needed to support it.

Level B audio

Level B audio provides medium quality audio (37.8 kHz sample rate, 4 bit samples) This format is most often used because it offers adequate fidelity along with a relatively moderate demand for data.

Level C audio

Level C audio provides narration quality audio (18.9 kHz sample rate, 4 bit samples). This is good for use with a complex animation sequence or large DYUV movies since more of the data stream is available for video data.

Audio Data Rates

Audio quality has a big effect on how much data is available for program logic or for loading images into video buffers. The following table summarizes the sector utilization for each type of CD-I audio format. The higher the audio quality, the more continuous data needed to support it.

Continuous Sector Utilization

Audio Level	*Relative Sector Number* 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Data Ratio
CD-DA	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	1:1
A Stereo	◆		◆		◆		◆		◆		◆		◆		◆		1:2
A Mono	◆				◆				◆				◆				1:4
B Stereo	◆				◆				◆				◆				1:4
B Mono	◆								◆								1:8
C Stereo	◆								◆								1:8
C Mono	◆																1:16

Having more sectors available for audio also means that simultaneous soundtracks can be implemented. For example, using the B Mono audio format, it would be possible use every fourth sector on the disc to implement two independent soundtracks in different languages. In this case, the two independent B Mono audio streams, each one using one out of every 8 sectors, would use the same about of raw data as a B Stereo or A Mono soundtrack.

Photo CD Formats

Photo-CD discs are compatible with CD-I and each disc actually contains a CD-I application that is used by a CD-I player or the Kodak Photo-CD machine. The images on the disc are accessed and displayed through the logic provided by the CD-I application.

For every image on a standard Photo-CD disc, data exists in five resolutions:

Image Format	Resolution	Pronounced
Base/16	192 x 128	"base over sixteen"
Base/4	384 x 256	"base over four"
Base	768 x 512	"base"
4Base	1536 x 1024	"four base"
16Base	3072 x 2048	"sixteen base"

The Kodak Photo-CD player uses special hardware decoding to display images while a CD-I player converts the Photo-CD data into a QHY-B or DYUV format in real-time.

A CD-I player uses the lesser three formats to display images on a television or monitor while the two higher resolution formats are used for digital image composition on computers or for transfer back to film.

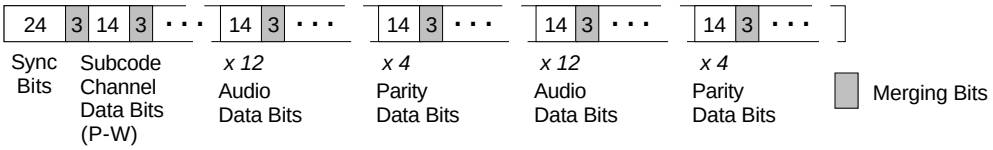
Optical Media Formats

The data formats used with optical media are described in the *Red Book* (digital audio), the *Yellow Book* (CD-ROM and CD-ROM XA), the *Orange Book* (magneto-optical and WORM – write-once, read-many – recordable discs), and the *Green Book* (CD-I). All of these optical media standards are defined and maintained by N.V. Philips and Sony Corporation.

CD-DA (*Red Book*)

It is important to know about Compact Disc-Digital Audio (CD-DA) discs, since the same data format – called "red tracks" – can exist on both CD-ROM and CD-I discs. The data on CD-DA discs (and in CD-ROM and CD-I red tracks) is essentially a continuous bit-stream, divided into frames. Each frame is 588 bits long and provides 24 bytes of audio data, plus 8 bits of channel data, one bit for each of eight subcode channels. The subcode channels are labeled P through W and are coded in EFM – eight to fourteen modulation – code (14 bits decode to 1 byte). Parity bits embedded in each frame of the bit-stream data are part of an efficient error correcting system known as Cross Interleave Reed-Solomon Code (CIRC). The parity bits allow hardware in the CD-I player to correct for errors in the data stream that might be caused by such things as dirt or scratches on the disc. A subcoding block consists of 98 frames and provides 2352 bytes of audio data along with 8 bytes of sub-channel data, 1 byte for each of the eight subcode channels. Data is delivered at the rate of 75 subcoding blocks per second.

CD-DA Bit-stream (one 588-bit frame)



Other optical specifications, such as CD-ROM and CD-I discs, use the same low-level data format but the 2352 bytes of audio data in a subcoding block (98 frames) is used to represent a physical sector on the disc.

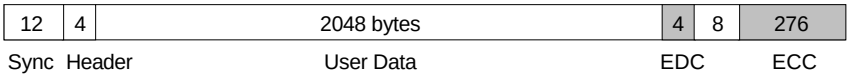
The subcode channels are used to provide information about the audio tracks being played, in addition to display data. Channel P is a simple track separator flag and is used to mark the end of individual tracks and to facilitate skipping through the music. Channel Q is used for control purposes and provides the track number and running time while an audio track is being played. Channel Q also carries a large amount of information about the data on a disc. It includes a table of contents, running times for each track, total disc running time, whether the disc contains any CD-ROM data, and a cyclical redundancy check (CRC) error detection routine.

The remaining six sub-code channels (R, S, T, U, V, W) are not given specific functions in the *Red Book*, but it is stated that they can be used for display purposes. CD+G digital audio discs use these channels to carry data for graphics and text, which can be displayed on a CD-I players (or other types of CD+G audio/video equipment) while listening to the audio tracks on a disc.

CD-ROM (*Yellow Book*) - Mode 1

The physical sector format for CD-ROM data is referred to as Mode 1. Each sector is 2352 bytes long and contains a 12-byte synchronization field, a 4-byte header, 2048 bytes of user data, and a 288-byte auxiliary field for EDC/ECC (error detection and error correction codes). The auxiliary field uses 4 bytes for EDC and 276 bytes for ECC. An 8-byte field between them is set to zero.

CD-ROM Sector Data - Mode 1



CD-I (*Green Book*) - Mode 2

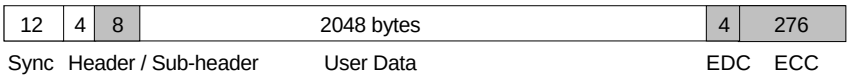
The physical sector format for CD-I data is referred to as Mode 2. Each sector is 2352 bytes long and contains a 12-byte synchronization field, a 4-byte header, and an 8-byte sub-header. Data is delivered at the rate of 75 sectors per second; however, two forms of sector data are used, one offering increased data rates by sacrificing EDC/ECC error correction.

The Mode 2 sub-header field is used to specify the file number for interleaved files, the channel number (e.g. parallel audio tracks in a real-time record of a bi-lingual title), coding information (e.g. Level B audio or CLUT4 image), and various sub-mode flags such as the sector data form, real-time sectors, end-of-file, end-of-record, sector data type, and time-coded trigger bits.

Form 1

Mode 2, Form 1 sectors are very similar to the Mode 1 sectors used for CD-ROM. Each Form 1 sector delivers 2048 bytes of user data and contains 280 bytes of error detection and error correcting codes. Form 1 sectors offer a data rate of 150 K bytes per second and are mainly used for loading program logic and video data that is not processed in real-time.

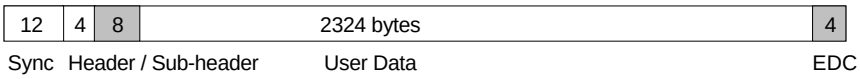
CD-I Sector Data - Mode 2, Form 1



Form 2

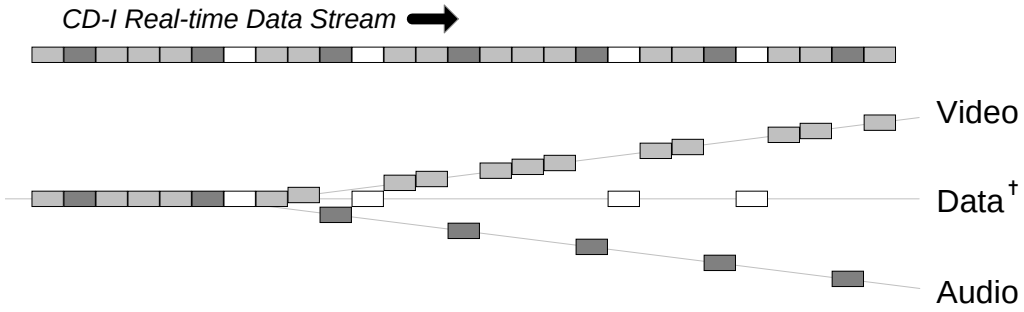
Mode 2, Form 2 sectors deliver 2324 bytes of user data per sector and do not contain error correction data. A reserved 4-byte field at the end of the sector can either be set to zero or used for quality control during disc production by loading it with the same error detection code used in Form 1 sectors. Form 2 sectors offer a data rate of 170 K bytes per second and are mainly used for loading audio and video data that is processed in real-time records.

CD-I Sector Data - Mode 2, Form 2



The Interleaved CD-I Data Stream

A unique feature of CD-I is the interleaved nature of real-time data in a disc image. As the disc spins in the CD-I player, program logic, audio data, and video data are simultaneously delivered to individual specialized processors as part of a real-time data stream.



† In this example, Form 2 sectors of real-time video and audio data are interleaved with Form 1 data sectors. Program data is *always* formatted as Form 1 sectors, but when they are embedded as part of a real-time data stream, they are *not* error corrected unless the CD-I player has built-in hardware support for ECC data correction that does not require any CPU intervention.

Different audio and video modes make varying demands on the real-time data stream for amounts of raw data and this affects the total amount of data that can be delivered to the each of the specialized processors in the CD-I machine. For example, the need for higher quality audio means that less data will be available in the real-time data stream for video data. Conversely, a fast animation sequence that requires a lot of video data will leave little room for higher quality audio in the real-time data stream.