Chapter 2

Digital Audio Compression Performance Considerations and Line Formats

Different audio compression algorithms have different characteristics, and there are several different, but incompatible ways to get the compressed audio onto the digital facility

2. Compression Algorithms and Line Formats

2.1 CCS MUSICAM® Digital Audio Compression

There is quite a bit of confusion among the names used for compression algorithms. ISO/MPEG Layer II denotes the basic compression algorithm invented and tested in 1990 by the IRT, Philips and the CCITT. ISO/MPEG Layer II was given the name MUSICAM (Masking patterns adapted Universal Sub-band Integrated Coding And Multiplexing) in conjunction with its use in the Eureka DAB (Digital Audio Broadcasting) System proposed for Europe.

In the US, however, MUSICAM® is a registered trademark of MUSICAM USA, the d/b/a of Corporate Computer Systems, Inc. (CCS), and its meaning is different from MUSICAM as it is used in Europe.

To understand the difference, it is important to understand the basic design of the ISO / MPEG compression system. The system, in essence, contains an encoder and a decoder. Audio, data, and additional signaling inputs are compressed and encoded at the encoder for transmission through a digital transmission facility. When the encoded signal reaches the decoder, the process is reversed and audio and data are restored to their original form.

The inventors of the algorithm realized that improvements might be made to the encode process as psychoacoustic research continued, and for this reason they designed the decoder to be a slave to the encoder. This allows for improvements to be made to the encode process without making obsolete perhaps millions of decoders deployed in the field.

MUSICAM USA has been improving the basic ISO / MPEG encoder since 1990, when we first introduced the CDQ2000 codec series and became involved in the pioneering work of In-Band On-Channel (IBOC) Digital Audio Broadcasting in the US. Today, MUSICAM USA encoding technology has evolved to a level of quality that is far superior to baseline 1990 ISO / MPEG compression, all the while *remaining compatible* with all standard ISO / MPEG Layer II decoders.

The MUSICAM USA improvements have been judged superior by the world's largest and most respected broadcast organizations, including the BBC, Swedish Broadcasting, CBS, and Infinity Broadcasting. In addition, the Direct—TV Satellite Broadcast System uses the MUSICAM USA technology for its audio compression coding. As of this writing, over one million decoders are receiving MUSICAM crystal-clear digital audio signals.

2.1.1 MUSICAM Compression Concepts

The main principle of MUSICAM is the reduction of redundancy and irrelevance in the audio signal. Every audio signal contains irrelevant signal components that have nothing to do with the identification of the audio signal (i.e., determination of timbre and localization). These irrelevant signals are not significant to the human ear and are not required by the information processing centers in the brain. The reduction of irrelevance means that these signal components are not transmitted. This results in a lower bit rate without any perceived degradation of the audio signal. Furthermore, it is possible to allow a certain degree of quantizing noise that is inaudible to the human ear due to the masking effects of the audio itself.

Every audio signal produces a masking threshold in the ear depending on a time varying function of the signal. An example that could help us to understand masking effects is that of an automobile. There is always wind, road and engine noise present while driving a car. Turning on the radio will mask most, or even all of this noise, depending on the radio volume. The noise is still there, but we cannot hear it because the radio masks it. Since we cannot hear it, the MUSICAM algorithm will not transmit it, resulting in a large reduction in the number of bits required.

Refer to Chapter 8 for a detailed explanation of MUSICAM psychoacoustic parameter adjustment concepts.

2.2 ISO/MPEG Layer III

Rev. 6

To make your *CDQPrima* compatible with more users, MUSICAM USA has added MPEG Layer III compatibility to all models. Layer III builds on the psychoacoustic modeling of Layer II but incorporates an additional layer of redundancy reduction using a Huffman (entropy) coding process that takes advantage of the statistical properties of the signal. Thus Layer III achieves higher compression levels and is ideally suited for low bit rate applications, providing that the added delay can be tolerated.

At bit rates of 56 or 64 kb/s, using Layer III, *CDQPrima* can deliver Commentary grade monaural audio with up to 15 kHz bandwidth. At 112 or 128 kb/s the advantages of Layer III disappear. Although at these bit rates, Layer III does sound better than *standard* ISO/MPEG Layer II, and can reproduce "CD-like"[1] independent channel stereo, listening tests have shown that MUSICAM USA's enhanced Layer II, as used in all MUSICAM USA Layer II products, sounds better than Layer III. In addition, due to the errors created by the entropy coding process, Layer III does not cascade well at any bit rate, and digital artifacts quickly multiply under multiple encoding cycles. At higher bit rates, or with multiple cascadings, Layer II, whether standard or enhanced, is the clear winner.

2.3 G.722

Because of its near-instantaneous coding delay of approximately 20 milliseconds, G.722 is useful where off air monitoring is a requirement;

---

[1]Fraunhofer Institut für integrierte Schaltungen. "Information about MPEG Audio Layer-3", Version 2.00, 9/5/95

however, bandwidth is limited to 7.5 kHz monaural. Recent schemes to "turbocharge" G.722 for higher frequency response should be evaluated carefully for compatibility, audible artifacts and phase distortions. In addition, G.722 does not support ancillary data nor does it cascade well.

To accomplish G.722 coding, the ***CDQPrima*** uses Adaptive Differential Pulse Code Modulation (ADPCM) to reduce the digital bit rate needed to transmit the digital representation of an analog signal. The ***CDQPrima*** digitizes the incoming analog signal with a 20 bit linear Analog-to-Digital converter 16,000 times per second. The Nyquist theorem states that at this sampling rate, an analog signal of up to 8,000 Hz can be reconstructed from the sampled signal. Using this sampling rate and A/D converter resolution, the following uncompressed bit rate is derived:

PCM bit rate = 16,000 * 16
PCM bit rate = 256,000 bits per second

The ***CDQPrima*** then compresses this bit rate to 64,000 or 56,000 bits per second using ADPCM. To accomplish this compression, ADPCM uses the fact that the next sample of speech can be predicted by previous speech samples. The ***CDQPrima*** transmits only the difference between the predicted and actual samples. If the prediction process is effective, then the information to transmit consists of significantly fewer bits than the digital representation of the actual sample. The prediction accuracy is greatly enhanced by splitting the 8 kHz band into two 4 kHz bands. The signal in each band is predicted separately. This allows a more faithful representation of the analog signal than is possible by considering the whole 8 kHz band at once.

In conventional PCM, the binary representation of each sampled analog point is used. Differential PCM (DPCM) transmits the difference between the previous point and the current point. In this scheme, the prediction process involves only the previous point. In fact, the predicted value of the current point is exactly the last point. In the CCITT G.722 implementation of ADPCM, the predictor is very sophisticated and uses the previous six points to predict the current point. This results in a very accurate prediction and hence requires a very low bit rate.

It must be noted that some early implementations of codecs and ISDN telephones manufactured by PKI (Philips) do not follow the H.221 framing structure standard of the G.722 coding algorithm. All

CDQPrima codecs are compatible with these early PKI products as well as the standard compliant PKI products.

2.4 Performance Considerations

Before discussing the various quality aspects of MUSICAM, it is necessary to define the terms used to represent the field of use of the audio. The four commonly discussed fields of use are:

- Contribution
- Distribution
- Emission
- Commentary

The term *contribution* is used to describe audio quality suitable for digital mastering. Its use would be in the transmission of a digital master from one archive to another. It is assumed that the original copy is in a 16 bit linear PCM format and it is to be compressed, transmitted, decompressed and stored in a 16 bit linear PCM format at the far end. Because the audio may be the source of future compression/decompression cycles (cascading), any contribution grade compression system must be able to withstand many encode-decode cycles and the effects of post production without any apparent degradation.

Distribution grade systems are used to transmit audio between two storage devices. However, the number of encode-decode cycles is limited to only a few. Distribution grade systems are used when the numbers of audio compression-decompression cycles are limited.

Emission grade systems are used when there is only one compression-expansion cycle anticipated. This is the case when audio is compressed and transmitted from one place to another, decompressed and stored on an analog tape and the only future manipulations done are in the analog domain.

Commentary grade systems are used for transmitting voice grade audio.

These definitions do not mention the analog bandwidth or the exact quality of the audio. They are vague terms used to describe the ability of the audio to withstand multiple encode-decode cycles. In all cases, the compressed audio is assumed to be indistinguishable from the original.

The MUSICAM design allows the digital bit rate, analog bandwidth and quality to be generally related by the formula

$$\text{Audio Quality} = \frac{\text{Digital Bit Rate}}{\text{Analog Bandwidth}}$$

As indicated above, audio quality increases as the bit rate increases and the analog bandwidth is kept constant. Similarly, if the digital bit rate is kept constant while the analog bandwidth is decreased, then the quality improves.

2.5 Tolerance to Multiple Processing

To understand the effects of multiple encode and decode cycles, it is important to review the predominant effect that allows MPEG audio to achieve its compression. This is the hiding of quantization noise under a loud signal. MPEG audio adjusts the degree of quantization-induced noise in each sub-band and thus hides more noise (uses fewer bits) in the sub-bands that contain large amounts of audio energy.

The quantizing noise rises with each encode and decode cycle and after enough cycles the noise level becomes perceptible. The degradation process is gradual and depends upon the level of the quantizing noise on the original. For example, the following table lists the approximate numbers of total MPEG Layer II encode and decode cycles before the noise of the stereo signal becomes significant. The number of encode/decode cycles using other algorithms is less.

Bit Rate	Number of Transcodings
384 kb/s	15
256 kb/s	5
192 kb/s	2
128 kb/s	1

Table 2-1 Bit rate vs. number of transcodings for MUSCAM enhanced Layer II

It is important to understand that these figures are for stereo audio, are approximate, worst-case and the exact number depends highly on the source material.

2.6 Post Production Processing Effects

Post production processing of compressed audio is a complicated effect to model. For example, an equalizer changes the level of a range of frequencies, while limiting and compression are non-linear processes. Very little test data is available to ascertain the effects of post processing. Private communications with the IRT suggest that MPEG Layer II is robust against the effects of post processing and the degree of robustness depends on the compression rate. In particular, 384 kb/s audio is unaffected by post processing while 128 kb/s audio is somewhat sensitive to post processing. ISO/MPEG Layer III is not as robust against the effects of post processing, and indeed, even moderate equalization may make artifacts noticeable at lower bit rates. Digital post processing should be avoided with MPEG Layer III.

It is not easy to define tests to measure the effects of post processing, but an international standards body (CCIR) is specifically designing tests to determine the effects of both transcoding and post processing. These tests were conducted in November of 1991 and represented the first time such tests were performed by an independent organization.

MPEG Layer II represents the most tested, documented and reviewed audio compression algorithm in the world. It is significant to note that no other compression technique has survived this crucial review process as well as the MPEG Layer II algorithm, and many other algorithms have elected not to participate in this review process. It is precisely these untested algorithms that make the boldest claims. MPEG Layer II audio provides the security of the international review process to insure the highest quality audio possible with today's technology.

2.7 MUSICAM, ISO/MPEG Layer II and Layer III Sample Rates and Bit Rates

These algorithms support four different sampling rates[2]. Each sampling rate generates a frame of compressed audio for the defined time interval:

[2]The sampling rates shown here are for MUSICAM encoding. AES/EBU or S/PDIF digital input sampling rates can be any valid digital rate, including 44.1 kHz, and will be automatically rate-adapted to a valid MUSICAM rate.

Sampling Rate	Sampling Rate Level	Frame Period (milliseconds)	
		Layer II	Layer III
48,000	High	.024	.024
32,000	High	.036	.036
24,000	Low	.048	.048
16,000	Low	.072	.072

Table 6-2 Supported Sampling Rates

Each **sampling rate level** supports fourteen different **bit rates** (Layer II only):

High Level	Low Level
32 K Bits	8 K Bits
48 K Bits	16 K Bits
56 K Bits	24 K Bits
64 K Bits	32 K Bits
80 K Bits	40 K Bits
96 K Bits	48 K Bits
112 K Bits	56 K Bits
128 K Bits	64 K Bits
160 K Bits	80 K Bits
192 K Bits	96 K Bits
224 K Bits	112 K Bits
256 K Bits	128 K Bits
320 K Bits	144 K Bits
384 K Bits	160 K Bits

Table 6-3 Support Bit Rates at Selected Sampling Rate

When making the selection of sampling rate and bit rate combination, the pairing must be valid according to the above information. If an invalid combination is selected, the ***CDQPrima*** will not frame, but will not report an error condition. For example, 24 kHz sampling is not valid when using 256 kb/s bit rates.

Rev. 6

For ISO/MPEG Layer III operation, *CDQPrima* supports bit rates of 56, 64, 112, 128, 192, 256, and 320 kb/s only.

2.8 Ancillary Data

The ***CDQPrima*** provides for transmission of up to two independent asynchronous data channels by way of ancillary RS232 interfaces. These interfaces provide a transparent channel for the transmission of eight data bits. The data format is one start bit, eight data bits, one stop bit and no parity bits. These interfaces are capable of transmitting at the maximum data rate selected by the encoder and decoder and thus no data pacing such as XON/XOFF or CTS/RTS is provided.

The encoder RS232 data rate can be set from 300 to 38,400 b/s. Please note that the use of the ancillary data channel reduces the number of bits available to the audio channel; however, a reduction of audio bits is distributed evenly over both stereo channels and occurs only if ancillary data is actually present. The data rate can be thought of as a maximum data rate and if there is no ancillary data present, then no data bits are transmitted. A typical example of this situation occurs when the ***CDQPrima*** encoder is connected to a terminal. When the user types a character, the character is sent to the decoder at the data bit rate specified, but if there is no key pressed, there is no data to transmit and no audio bits are lost.

2.9 Compatibility with older CCS codecs

The ***CDQPrima*** is fully compatible with all older CCS audio codecs. The decoder automatically senses any of the MPEG types of bit-streams (CCSO, CCSN and MPEGL2) and adapts the bitstream. The encoder can be set to generate any of the MPEG types of bitstreams. This allows full compatibility with older CDQ1000, 2000, 2001, and 2012 codecs.

2.10 Digital Transmission Line Formats

2.10.1 Inverse Multiplexing Functionality: CCSIMUX and J.52

The inverse multiplexing functionality of ***CDQPrima*** can combine individual 56 or 64 kb/s switched and dedicated channels to create virtual double speed audio calls at rates that are double of 56 or 64 kb/s (2x56/64) applying CCSIMUX and even up to 6 x 64 kb/s using ITU-T J.52 (H.221) BONDING.

2.10.1.1 CCSIMUX

CDQPrima supports dual-port calls for backward compatibility with existing MISICAM USA audio codecs, such as CDQ2000, 2001 and 2012 as well as other manufacturers codecs, for high quality audio applications. In these applications, a bandwidth of 112 or 128 kb/s over the PSDN is achieved by using two ISDN B-channels accessed through the integrated terminal adapter or data interface cards, switched-56 units, or external terminal adapters (TAs), each connected to a 56 or a 64 kb/s channel. The audio codec provides two data ports, one for each channel and aggregates their bandwidth. ***CDQPrima*** can communicate at 112 or 128 kb/s with existing equipment in such configurations.

The ***CDQPrima***'s CCS2Line mode can be used to connect to other manufacturers codecs and terminal adapters that also support inverse multiplexing. Both codecs must be in two-line mode and set to the same data rate for proper operation. If only one 56 or 64 kb/s connection is established when using MPEG Layer II, monaural audio will still be passed between MUSICAM USA codecs. Therefore, when connected to a MUSICAM USA or CCS Audio Products codec using Layer II with two lines, if one line is lost, the audio quality is reduced to 8.5 kHz monaural, but the audio is *not* muted.

In a ***CDQPrima,*** multiple 56 or 64 kb/s channels can be accessed by one speed dial entry. Two single-channel calls are made over separate host ports, called the primary port and the secondary port.

2.10.1.2 ITU-T J.52

Within the ITU J.52 standard for the transmission of encoded audio signals according to ISO MPEG 11172-3 or 11318-3 the following items are addressed:

Time delay compensation:	The routing of B-channels within the ISDN network might be different and thus the mutual delay between such B-channels has to be compensated.
Organization of B-channels:	The receiver must know which B-channels are combined to a virtual channel.
Interfaces:	There are two main interfaces within ISDN, BRI and PRI. The adaptation of the signal can be done within the transmission device or within separate terminal adapters.

Error protection: The ISO bitstream provides the possibility to insert a CRC word for error recognition, such errors can not be easily corrected or concealed. Thus an additional possibility is required to correct transmission errors.

2.10.1.3 Transmission with H.221 framing

The ITU J.52 recommendation is based on the ITU recommendation H.221 frame structure for a 64 to 1920 kb/s channel in audiovisual teleservices. Within the H.221 structure, delay compensation in time intervals of 20 ms. is possible up to a total delay of 1.28 seconds. A disadvantage is that within a 64 kb/s H.221 frame, only 62.4 kb/s net data is available. The remaining part is for FAS (Frame Alignment Signal) and BAS (Bitrate Allocation Signal). When using this mode with internal terminal adapters, the ***CDQPrima*** will automatically adapt to the number of lines that are connected, up to a maximum of the bit rate selected.

!note:

The ***CDQPrima*** supports multiple line H.221 BONDING, and can operate over 6 ISDN B channels yielding a total connection rate of 384 kb/s. However, you cannot mix terminal adapter types. You must use internal terminal adapters at both ends or external terminal adapters at both ends. You cannot use BONDING with internal TA’s at one end and external BONDIONG TA’s at the other end, except if the external TAs are a 2X64 type.

2.10.2 Single Line Mode

H.221 BONDING, whether using internal or external terminal adapters can only support multiples of the basic rate 64 kb/s ISDN channel. CCS2Line mode can only support 112 or 128 kb/s over 2 ISDN B channels. The single line mode can support any allowable data rate, from 24 to 384 kb/s and allows connection to any digital transmission facility using the appropriate data interface. MUSICAM USA currently supports X.21, RS530, RS422 and V.35 interface modules, each supports all allowable data rates. In addition, the three available terminal adapters also support the single line format at 56 or 64 kb/s.

2.10.3 Independent Mono

Independent mono allows you to send two independent (different) audio programs on two (or more) different lines to multiple locations. Independent mono supports only 56 and 64 kb/s bit rates. MUSICAM Layer II or G.722 can be used for send and G.722 or MPEG Layer III can be used for return audio. If MUSICAM Layer II is used for return audio, only one location can be monitored. With G.722 or MPEG Layer III for return audio, two locations can be monitored.