ELECTRONIC ENGINEERING

EETIMES®

www.eet.com The industry newspaper for engineers and technical management Monday December 6, 1999

FOCUS ENTERTAINMENT ELECTRONICS

Dolby Digital delivers 3-D sound

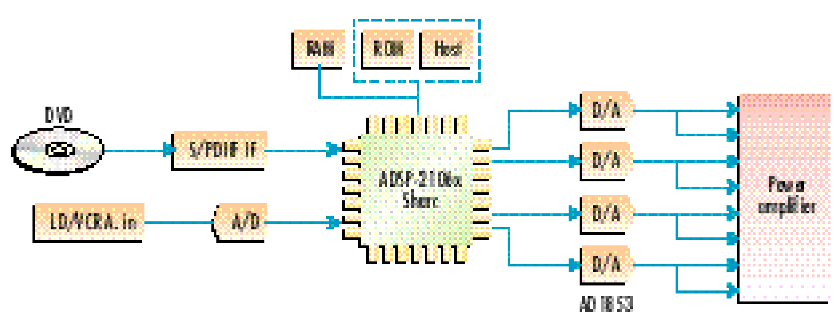
BY P V G MENON
Program Manager
Software & Systems Technology Division
Analog Devices Inc.
Norwood, Mass.

Most consumers agree that "surround sound," or audio that encircles all sides, significantly enhances the listening experience. Whether it is the soundtrack in a movie theater, music in the car or television at home, the demand is high for this kind of audio delivery.

Audio can be distributed in several ways. Each involves capturing additional content and coding the placement of sounds precisely—by specifying the exact channels from which they should emerge, how the sound should be "panned" or moved during playback, and so on. The most logical way to add this extra content to the existing transmission and distribution media is to encode the content and compress it. That has the added benefit of conserving expensive and scarce bandwidth. It also makes it possible to include different languages, mixes and video while fitting all of that content through all of the available digital pipelines (digital cable, digital television and digital broadcast satellite) without sacrificing perceptible sound quality.

Dolby Digital (formerly called AC-3) is one of the more popular coding schemes available. It is the mandatory coding scheme for DVD titles and is also the standard for DTV. Implementations (as Dolby calls them) range from custom silicon (ASICs) to general-purpose multifunction digital signal processor configurations, software decoders for PCs and so on.

In essence, the Dolby Digital coding scheme relies on a sophisticated "psycho-acoustic" model to compress the data. The basic principle is to code all of the sound that can be detected by the human ear with just enough resolution to ensure that no audible distortion or noise has been added. For example, loud sounds can totally mask the presence of other lower-level sounds



An audio/video processor can use a floating-point processor for Dolby Digital decoding. The benefit is the greater precision that can be obtained by using floating-point math for the series of complex calculations that most decoders use. The Dolby Digital encoded bitstream is decoded by the Dolby Digital decoder DSP and generates six channels of audio fed to D/A converters (DAC). This design uses Analog Devices' high-resolution 24-bit, 192-kHz AD1853 D/A converter.

and noise that are nearby in frequency. It is therefore possible to use fewer bits to quantize the louder sound without any apparent audible side effect. When done properly, the technique, known as perceptual coding, ensures that even though the delivered audio data is not exactly the same as the PCM original captured in the recording studio, the sound quality is perceptually unchanged.

The added effects of being surrounded by sound, combined with the audio effects made possible by six discrete channels, in fact provide a closer replication of the recording studio listening experience than is possible with two-channel CDs, in an uncompressed pulse-code modulation, 16-bit, 44.1-kHz format.

To enhance compression efficiency, the Dolby Digital encoder uses the concept of high-frequency coupling to share the bits used to describe the audio among the occupied channels, typically above 14 kHz. This is a more sophisticated process than the “joint stereo” mode found in other coders. Another technique, dynamic bit allocation, allocates bits to only those channels that contain audio data at a particular instant.

Given that Dolby Digital can be used for mono, stereo or any combination of channels up to 5.1, it is not surprising that the compression ratios range between 3.4:1 and 15:1. However, one would normally come across something like 10:1 compression in the case of 5.1-channel DVD material that is Dolby Digital-encoded.

The Dolby Digital bit stream relies on a block approach to the coding and transportation of data. Each block has a unique stream identifier (bit-stream information, or BSI,) to identify it as a Dolby Digital-coded stream while coming down an Audio Engineering Society/European Broadcaster Union (AES/EBU) carrier, and important information related to the sampling frequency, number of channels, data rate and the like. Error detection is built in, and most decoders that are implemented in silicon will use this information to help system designers prevent harsh noise from being sent to the speakers in the event that disk damage, bursts in data reception or similar mishaps occur. The BSI typically is read by the Dolby Digital decoder and is fed to a host controller for display of the relevant information on the front panel of a typical end product.

While decoding data, the approach is to first extract the cyclic redundancy check (CRC) checksums and verify that the incoming blocks of data are error-free. Once that is established, the decoder moves on to extracting the BSI, setting up the decoder moves on to extracting the BSI, setting up the memory for extracting individual channels and writing that data to the specified memory areas of the decoder. It also bears mentioning that several inputs are available to the decoder from a user perspective; they dictate how the decoder carries out its tasks.

For instance, a listener may choose to down-mix the six channels of information onto just two channels so that he or she can listen to it on a conventional stereo setup (a standard stereo television set, for example). Down-mixing the six channels of information onto just two channels—typically the front left and right speakers—ensures that no encoded information is lost. That is an important consideration, given that the five major channels—the front left, center and right; and the surround left and right—are treated as full-bandwidth channels with discrete data feeds. The low-frequency-effects (LFE) signal usually is ignored in most down-mix scenarios.

Decoders also must deal with other user inputs. Speaker configurations differ in most homes, for example. Those consumers with deep pockets might invest in full-bandwidth speaker arrangements, while the more common configuration is full-bandwidth speakers for the front left and right channels, with smaller, limited-bandwidth speakers for the rear satellite speakers. This setup is likely to be found in many homes that upgraded from a conventional hi-fi to home theater. A subwoofer may or may not be present, depending on individual preferences and budgets. Thus, there are many down-mix modes available: mono (1/0), plain stereo (2/0), three front-only (3/0), two front and two rear (2/2).

Then there’s the issue of bass management. Depending on the type of speakers (full bandwidth vs. limited-frequency smaller speakers), and depending on whether a subwoofer is present, the decoder can be instructed to divert the low-frequency (20-hertz to 100-hertz) content in each channel to a designated speaker or set of speakers. In a setup like a home theater, for example, where two full-bandwidth speakers serve the front left and right channels, and smaller, limited-bandwidth speakers serve the center and surround channels, it’s possible to have the bass manager divert the LFE content to the front left and right speakers without feeding any low-frequency content to the limited-bandwidth speakers. This prevents overloading the smaller speakers, as well as the loss of exciting bass content during playback.

Tight standards

Dolby Digital decoders have to perform certain mandatory functions and must adhere to the tight standard written by Dolby Laboratories. They must be tested with a comprehensive set of test vectors and be certified by Dolby before their companies can sell them to OEMs for inclusion in products. So how do they differ? Let's take the case of silicon implementations.

In the final analysis, a decoder is a piece of software, embedded or otherwise, that is implemented in a certain way on a particular chip. Software development and implementation is still an art form, and never more so than in the domain of audio processing. No doubt there are differences in the way two implementations of the same standard sound to a listener, even though both were designed to the same specification and have been subjected to the same tests. Audio domain knowledge and the degree to which programmers exploit architectural advantages account for these differences.

Greater precision can be obtained by using floating-point math for the series of complex calculations that most decoders use while doing their job.

Domain knowledge is invaluable in helping a decoder designer decide how to shape the signal going to the D/A converters. The maximum resolution currently available in D/A converters is 24 bits of precision, yet today's decoders can output 20 bits, 24 bits, even 32 bits or higher after the basic Dolby Digital decode process is completed. It is only from a thorough knowledge of the "art" of audio-signal processing that tricky decisions related to audio-data dithering, truncation and the like can be made.

In the area of architectural advantages, few would disagree that floating-point silicon architectures offer advantages over fixed-point architectures. That's because greater precision can be obtained by using floating-point math for the series of complex calculations that most decoders use while doing their job. In the course of these calculations, the decoder has to contend with a lot of mantissa processing, so it's essential that round-off errors be kept to the barest minimum to preserve the fidelity of the sound. Generally speaking, this translates to a simple rule: The higher the resolution of the decoder, the better the sound.

Greater resolution comes at a cost, however: With the push to make silicon perform more functions, decoder designers must budget their MIPS very carefully indeed. A longer word-length DSP architecture enables decoder designers to conserve MIPs while preserving fidelity, as opposed to tweaking implementations to run in double precision with a horrendous cost in terms of MIPS. Thus, a low-cost, 32-bit floating-point DSP, like Analog Devices' Sharc DSP, could support Dolby Digital decoding.

Several sources

In such setup, the source for the audio/video processor could be digital input from a DVD player, or an analog audio input from a laser disc player, a VCR or a TV. The analog input is converted using A/D converters. Then the input data comes to a DSP, in this case a 32-bit floating point DSP, where the Dolby Digital decoder is running. The analog inputs, which have been converted to PCM, will be processed by the built-in Dolby Surround Pro Logic decoder portion of the Dolby Digital decoder. The output will be processed as mono, stereo or 4-channel surround, depending on the needs of the consumer. The Dolby Digital encoded bit-stream is decoded by the Dolby Digital decoder DSP and generates the six channels of audio, which are then fed to D/A converters.

The resolution of the converter is extremely important. As with many things, you get the quality for which you pay. Since the DSP is generating 32 bits of decoded data, it is extremely important that the converters can process this information and output to the amplifier, and from there to the loudspeakers, with minimum distortion and truncation. In this example, the high-resolution 24-bit, 192-kHz AD1853 D/A converter from Analog Devices offers the best combination of THD+N (-107 dB), dynamic range (120 dB) and out-of-band noise (115 dB) available.

Thirty-two-bit de-coding becomes necessary when the quantization noise generated by the decoder algorithm exceeds the DAC's 24-bit noise floor, which introduces audible noise and distortion. Implementing double-precision routines and error-feedback schemes to work around the noise-floor problem adds DSP computation overhead by a factor ranging between 4 and 10. This would leave no room for the DSP to do any other task, which means that end-product or system designers might have to consider using multiple DSPs in their design, which would hike the cost.

Dolby Digital is already implemented in several high- and mid-range systems in the home-theater domain. Several Dolby-Digital-encoded PC games have also started to appear, enabling hard-core game enthusiasts to enjoy the sensation of being immersed in sound while they traverse fantastic and futuristic worlds. New applications are appearing, thanks to the development of technologies that enable the reproduction of virtual surround sound; that is, the effect of having five speakers around you when you actually have only two physical speakers installed.

What is important is that as DSP architectures evolve, and as Moore's Law ensures that consumers get more bang for their silicon buck with each year, decoder vendors will relentlessly add features. Consumers will soon have no more excuses for not sharing in the surround-sound experience. ■

WORLDWIDE HEADQUARTERS
One Technology Way, P.O. Box 9106
Norwood, MA 02062-9106
U.S.A.
Tel: +1 781 329 4700,
(1 800 262 5643, U.S.A. only)
Fax: +1 781 326 8703
World Wide Web Site: http://www.analog.com

EUROPE HEADQUARTERS
Am Westpark 1-3,
D-81373 München
Germany
Tel: +89 76903-0,
Fax: +89 76903-557

JAPAN HEADQUARTERS
New Pier Takeshiba
South Tower Building
1-16 1 Kaigan, Minato-ku
Tokyo 105-6891, Japan
Tel: +3 5402 8200
Fax: +3 5402 1063

SOUTHEAST ASIA HEADQUARTERS
4501 Nat West Tower
Times Square
Causeway Bay, Hong Kong, PRC
Tel: +2 2506 9336
Fax: +2 2506 4755