NAB TV TechCheck
The Weekly NAB Newsletter for Television Broadcast Engineers
May 4, 2009

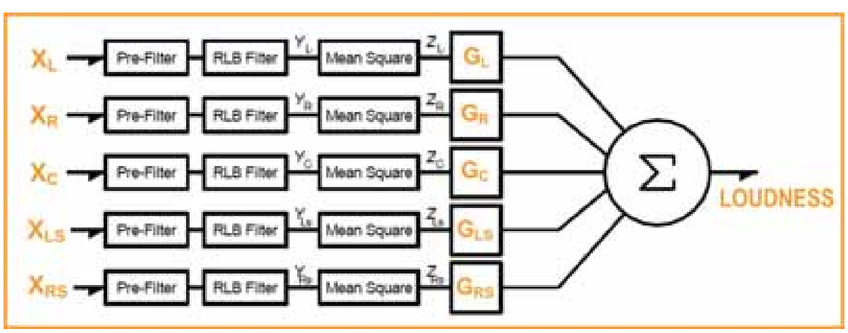
Loudness Monitoring in a Digital System

Audio loudness has been an issue with broadcasters almost since the inception of broadcasting, with research dating back to the 1930s. A session at this year's NAB Broadcast Engineering Conference (BEC, April 18-23, 2009, Las Vegas, NV) entitled "*Loudness, Lipsync and AFD for DTV*" includes a paper, excerpted here, which explores the use of ITU B.S. 1770 as a means to measure loudness, its effect on dialog normalization as specified in the ATSC A/53 standard, and common practices in delivering Dolby Digital audio.

LOUDNESS – Loudness is perceived by the listener, and while there is a relationship between the intensity of sound to loudness, it is not the same. The relationship of a listener's perceived loudness and intensity can be better understood by considering how the human auditory system works. As an audio wavefront enters the ear, it is sampled through a series of auditory filters which each have the characteristics of a bandpass filter. These filters are known as "critical bands," in which the intensity of the frequencies in each band is quantified and sent to the brain for processing.

Like a bandpass filter, each critical band also has an associated bandwidth, which at the lower end of the frequency spectrum is around 90 Hz, increasing to over 1 kHz at the upper end of human hearing. Within a critical band, there is a compression-like action that affects the way the intensity of the energy is quantified prior to being sent to the brain. This compression is nearly logarithmic, requiring ten times the intensity of energy in the critical band before there is a perceived doubling of loudness.

LOUDNESS AND PROCESSING – advancements in analog audio technology have allowed development of efficient multi-band processing, aiding in the delivery of sound to the end user. While such processing has helped maintain a high average percent of modulation, it has been at the expense of dynamic range, and has created a competition of stations wanting to be the loudest in a market. In today's digital environment, there is no need, or advantage, in processing audio to maintain a high average percent of modulation. In contrast to analog audio delivery, digital delivery allows for high signal to noise ratios, large dynamic ranges, and wideband responses that allow for delivery of sound as intended by the artist, not limited by the requirements of transmission.

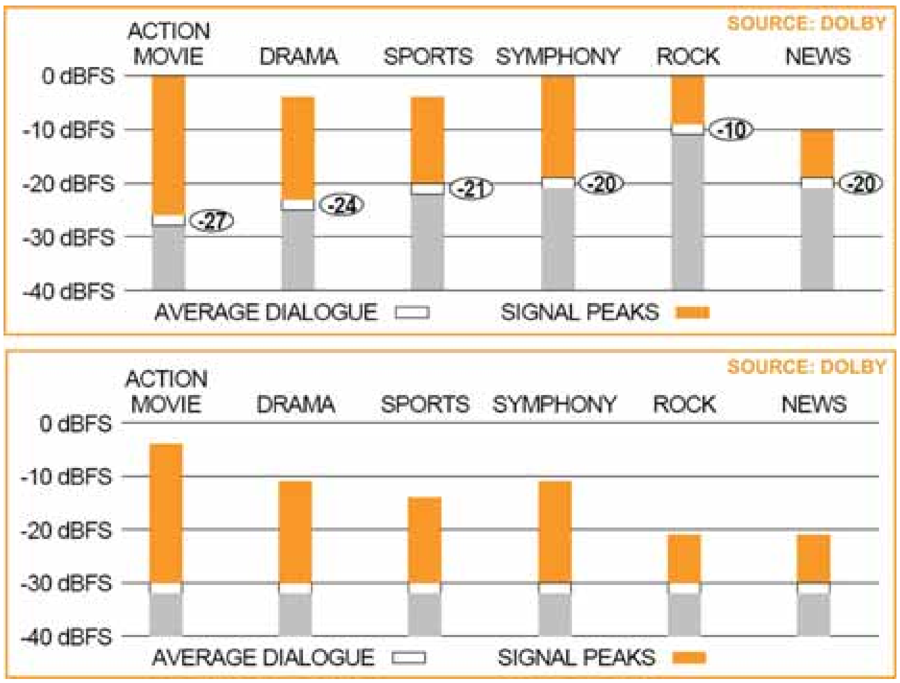


ITU-R BS.1770-1 – with the goal to deliver to the listener the highest quality audio at a consistent loudness, two tools now exist for systems using AC-3 audio coding: ITU-R BS.1770 and Dolby Dialog Normalization. BS-1770 Annex 1 specifies an algorithm for measuring multi-channel loudness. While the formal document only references a maximum listening environment of 5.1 channels, the algorithm is scalable up or down as required by the mix. This scalability is accomplished by processing each channel of sound, excluding the low frequency effects channel (LFE), through a series of components before being summed.

The block diagram shown illustrates the BS.1770 algorithm for a 5.1 channel listening environment, highlighting the four components applied to each channel:

- Pre-filter: accounts for the acoustical effects of the listener's head;
- RLB filter: Leq(RLB), Revised Low Frequency B response curve is used instead of the Leq(A) that was originally specified for use to determine Dialnorm. The major difference from the Leq(A) curve is that the low frequency range is wider, and the high frequencies do not roll off;
- Mean square: calculates the energy in the channel as integrated over an interval of time. For the purposes of BS-1770-1, the interval of time used for this calculation is the full length of the material used in testing;
- Gain: applies a weighting to each channel before the summing occurs. In essence, the primary channels (left, right, center) have a weighting of 1.00, where the surround channels have a weighting of 1.41.

DOLBY DIALOG NORMALIZATION – in the production process, sound engineers typically process the audio for the material. That is to say, the audio processing used for a symphony would be different than processing of a rock and roll band's performance. Dialog is then mixed into the material relative to the primary sound, while maintaining a lower than peak (0 dB relative to full scale, denoted dBFS) level. Consequently the dialog level for different types of program material will have different levels.



In the upper graph of the figure to the right, some typical types of program material and the associated dialog levels as found by Dolby Laboratories in their research are shown. By normalizing the dialog level to a fixed value, specified as -31 dBFS for AC-3 (Dolby Digital), loudness between differing materials is perceived by the listener as being equal. The lower graph of the figure shows the material identified in the upper graph after normalization. Normalization does not occur in the mix by the sound engineer, but instead is accomplished by setting the Dialnorm metadata value in the AC-3 bit stream. When the listener's decoder receives the Dialnorm value, the decoders gain is changed, acting as if the listener has changed the volume of their sound system.

This paper is authored by John Hartwell, Hartwell Consultants, and is included in its entirety in the *2009 NAB Broadcast Engineering Conference Proceedings*, available on-line from the NAB Store (www.nabstore.com). Audio recordings of the BEC sessions are also available for purchase – for more information, visit the NAB Show Online Learning Center at http://www.nabshow.com/2009/education/onlinelearningcenter.asp. For additional conference information visit the NAB Show Web page at www.nabshow.com.

ccolerid@nab.org for information about the NAB Satellite Uplink Operators Training Seminar. If you are interested in sponsorship opportunities for this event contact NAB Advertising at (800) 521-8624 or advertising@nab.org.