WHITE PAPER

Loudness Monitoring and Control in Broadcast Systems

This paper examines loudness issues, measurement and correction methods, and some recent work by standards organizations. Some real-world examples and implementations are explained, and a new method of displaying audio loudness levels is described, consistent with the broadly implemented ITU-R BS.1770 recommendation.

What's the Issue?

Over recent years, the transition to digital audio production workflows has resulted in greater creative flexibility, higher productivity and enhanced quality. The consumer's experience has also been expanded beyond belief, particularly in relation to digital broadcasting — HD images and multichannel audio provide compelling experiences and have played a major part in paving the technology path to the home.

Often however, the rollout of digital television has been accompanied by consumer complaints regarding audio "loudness." Despite recommended industry practices, the general use of peak measurement in audio metering has led to over-compression of audio signals, which has been exacerbated by broadcasters' and advertisers' drive to make their output sound bigger and better than the competition.

Loudness is a perceptual quantity defined as the magnitude of the physiological effect produced when a sound stimulates the ear. Variations of 10+ dB are highly objectionable and can easily arise with digital delivery, leading to listener stress and complaints. Consequences for the broadcaster might include reduction of audience share and, potentially, a fine by the relevant regulator.

A key issue is that although a number of ways have been developed to characterize audio levels, including VU (volume unit) and peak programme meter, none of these serves as a good measurement to represent the perception of overall program volume by the human ear.

The Need for Loudness Standards and Recent Evolution

The original ITU recommendation (ITU-R BS.1770) published in 2006 was developed to provide guidance to broadcasters requiring a method to control the loudness value of their content. Generally, the loudness deviations are noticed in the distribution of the content where a listener can compare various programs on a channel, or compare the loudness of one channel to another when changing broadcast channels.

This and its subsequent descendents BS.1771 and BS.1864 — though providing a fair amount of information as to how to measure the loudness aspect of content — were still somewhat incomplete in respect to the details required to satisfy all issues surrounding loudness.

Adding to the pressure, growing numbers of consumers have voiced their concerns in the past few years about variations in loudness, resulting in national legislative activities that have begun to regulate broadcasters more closely.

In the United Kingdom, the BCAP (Broadcast Committee on Advertising Practice) went into effect January 1, 2009, limiting program and advertisement loudness.

In Italy, legislation was passed in 2009 requiring broadcasters to monitor loudness and keep an actual log that would be available for public inspection, proving the broadcaster's compliance.

In December 2009, the United States adopted the recommendations summarized in the document "ATSC A/85:2009 Recommended Practice: Techniques for Establishing and Maintaining Audio Loudness for Digital Television."

In August 2010, the EBU released their recommendation for loudness monitoring "EBU – Recommendation R 128 Loudness normalization and permitted maximum level of audio signals." The document contains extensions to and further definitions of the ITU-R BS.1770. Specifically, this recommendation considers program and channel loudness.

The growing objection by viewers in the U.S. to wide variations in audio loudness led to the clearly worded CALM bill (Commercial Advertisement Loudness Mitigation Act) passed in December 2010:

> "*In the new world of digital broadcasting, the CALM Act appropriately instructs the expert federal agency to adopt a national standard for commercials, which require their volume not be any louder than the average level of the programs they accompany…….. In particular, there are differences in the compressed audio levels of television shows and commercials. While the audio of a television show usually matches natural sound more closely, the audio of a commercial has less distinction between loud and soft sounds, resulting in everything seeming louder……… guidance is designed to assist broadcasters in complying with requirements that advertisements not be noisy or 'excessively strident.'*"

Some Practical Issues

As mentioned above, the key consumer issue is the perceived change of audio levels between different portions of audio content. This fundamental problem has tested audio content producers and broadcasters for some time — how to make the "average" audio level be perceived consistently throughout a program, regardless of program type and content, and beyond that, consistent between channels. That's a key operational challenge for multichannel broadcasters and playout providers.

The audio content itself has also become ever more complex. More and more organizations are opting to transport surround sound material as discrete, embedded audio signals, and the number of audio programs within an HD embedded infrastructure is also increasing, in particular for those who play out mixes for multilingual distribution.

There are always challenges in achieving an international consensus on a format or standard, yet the ITU-R BS.1770/1 has been gaining momentum as the reference point for broadcasters globally. The current loudness scale is measured in Loudness Units, as opposed to dB as in other meters. The ITU loudness algorithm takes into account perceived loudness levels and applies filters accordingly (PPM meters do not, as we'll describe later).

The practical implication of this move requires the loudness of audio content to be measured by the content provider and broadcaster using specific, and preferably automated, loudness meters. The method of measurement needs to become a standard practice as part of the content workflow. If a broadcaster is cited by regulatory bodies because it has exceeded a certain figure over the course of a two-hour movie, having the appropriate measurement equipment in place will enable the broadcaster to properly contest the issue.

Automated loudness corrections should be imperceptible to the listener. This means characteristics such as equalization and spectral density must remain untouched, and adjustments are only made when content strays beyond user-configured target loudness ranges, in-line with current international standards.

The ITU-R standard is being increasingly used by broadcast-governing bodies in different territories as the foundation on which to implement new regulations. The ATSC recommended practice A/85, for example, has provided a practical approach for the measurement of loudness by program providers and broadcasters using ITU-R BS.1770 as a basis.

Without a truly global standard, facilities that produce commercials for international markets could well have to mix them multiple times or have multiple meters, which is expensive and time consuming. The compromise may be to achieve a common goal that may not meet everybody's aspirations, but will "accomplish the job."

What's more, some programming doesn't fit the ITU-R template because it is based on dialogue levels. Some sports, for example, feature periods of hushed calm punctuated by sharp noise. This same issue applies to sounds captured by localized microphones that cause audio spikes outside of the norm.

Audio loudness needs to be monitored and managed throughout the entire workflow, along with the metadata that impacts how the audio content is delivered. Each facility should perform this using a combination of real-time and offline tools. Broadcasters can now achieve excellent loudness consistency that is virtually imperceptible to the consumer, via Perceptual Loudness Management (described later). This technique can also be used offline to manage loudness with natural sounding and consistent results.

Methods of Describing Audio Levels

Audio loudness measurement fundamentally requires a specific method of defining audio levels — a "meter ballistic."

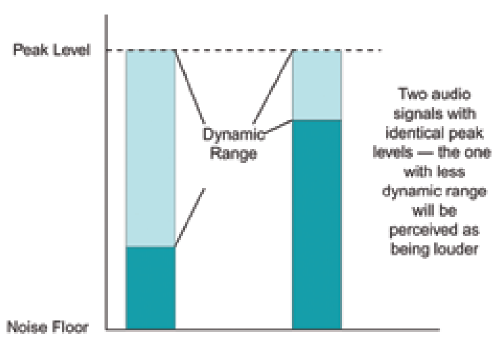
The most commonly used meter ballistics do a poor job of representing the average audio level perceived by a listener, since they were designed for another purpose. Often, the main application of existing ballistics is to prevent saturation and distortion on a recording medium or transmission path. While certainly important, this is a very different application than determining loudness.

One of the most common audio measurement ballistics is the Volume Indicator (VU meter), which traces its roots back to the 1930s. It is an early attempt to display an "average" audio level; this meter ballistic has an integration time of 300 ms, so it does not respond to fast audio peaks and troughs. This meter is defined by IEC 60268-17.

Another set of common audio ballistics attempts to show the peak excursions. BBC developed its peak program meter (PPM) in the 1930s for this purpose. Technically, the PPM is a quasi-peak meter, since it uses a 10 ms integration time. The PPM is defined by IEC 60268-10.

Peak meters are also commonly used with digital audio content. Unlike the PPM, this ballistic has zero integration time, responding to every digital audio sample.

Clearly, none of these ballistics accurately reflect the average level perceived by human ears. This loudness quantity is as important as other audio level measurements in terms of the listening experience for the consumer. These and other common meter ballistics can, however, produce widely varying loudness levels when used to monitor different types of audio content. The dynamic range of any audio content has a major impact on the perceived loudness, but is not reflected well by any commonly used meter ballistic. Peak reading meter types can give especially erroneous results, since there is (by definition) no attempt to average the audio level. The figure below shows an example.



The industry has made several attempts to quantify this perceived quantity of loudness, one of the earliest being the "CBS Loudness" measure, which was described in 1982. This was an attempt at making a perceptual rather than a purely electrical measurement, and for many years this was the closest to a reference loudness specification available.

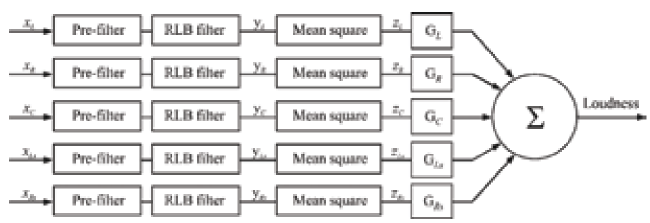
Dialog normalization (dialnorm) is a metadata parameter associated with compressed audio content, which is designed specifically to control dialog levels across channels at the consumer premises. This is a valid concept, but there have been implementation issues that, so far, have prevented it from being as effective as intended. There are multiple steps needed to properly employ dialog normalization. The first still requires a reading of the perceived loudness of the audio, regardless of content — with or without spoken dialog, for instance. Computation currently requires proprietary hardware and algorithms, firing ongoing discussions on the subject.

The dialnorm process outputs a dialog normalization level in dBFS, which is input to the encoder and carried in the compressed audio stream as metadata. This is, however, one common place where the process fails. Many encoders in use still have their default dialog normalization level being encoded — regardless of what the actual value may be. Needless to say, this does not yield correct results.

A simple mean-squared power measure has also been developed, commonly labeled Leq. This measurement can be applied unweighted (usually just referred to as Leq) or can use a weighting filter before the power measurement. The type of filter is typically indicated along with the measurement means, such as Leq (RLB).

The revised ITU-R BS.1770 document was based on the results of extensive perceptual trials, aiming to match an electronic loudness response to many different sets of listener results. The ITU has extended the Leq (RLB) ballistic to work with multichannel content. This required adding a high-frequency shelving filter to the RLB response.

A key feature of the ITU-R BS.1770 measurement technique is its multichannel capability, achieved by using the sum of five channels of a surround sound mix (the LFE channel is not used). All five channels pass through the same filtering and power summing process, with the individual results being scaled and then summed together for a final result. A block diagram of the measurement technique is shown below:



The pre-filter block is the shelving filter mentioned previously, which is then followed by an RLB filter. The Leq (mean square) level measurement is next, followed by gain blocks (Gx) that account for the relative angle of each surround channel. Finally, the scaled results of each channel are summed.

Extensive tests by the ITU have proven that this technique has a strong correlation to perceived loudness across different program materials and across different listener groups; hence, it has seen a growing industry acceptance.

Loudness Tools in the Broadcast Workflow

Depending on individual production and distribution environments, the following loudness measurement methods are most commonly used in various combinations:

- **Loudness Metering**
  Increasingly, the industry is recognizing that the ITU-R BS.1770 recommendation is the most practical to implement. Devices that implement it are deployed throughout the workflow to verify that audio content is within tolerances.
- **Offline Analysis and Correction**
  When using file-based material, there is an opportunity to employ non-real-time loudness measuring and correction. In such cases, an automated, file-based process is ideal as a complement to the above method.
- **Real-time Correction**
  When live material is being broadcast, there are particular considerations regarding loudness depending on the type of programming; sports coverage was quoted as an example earlier.

The ultimate application of real-time correction is at final playout, serving as the last line of defense if unchecked content were to get this far without correction. Current processing products can be set to "Protection Only" for this purpose.

New Operational Displays for Loudness Metering

Once industry practices and standards become established, the implementation challenges of effectively displaying and interpreting the measurement information for a range of users become key issues. This section describes how Harris products help solve these issues, including loudness measuring and processing in baseband systems, as well as loudness analysis and correction in a file-based environment.

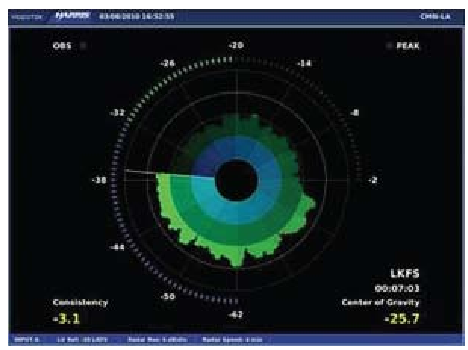
As described, loudness measurement techniques are based on meter ballistics, so displaying the individual channel loudness in a way similar to traditional audio meters conveys this information in the most readily understood manner. Audio levels are still labeled in dBFS, and the meters themselves should be labeled in "LU" or loudness units, differentiating them from "VU" or peak reading level meters.

However, looking at loudness channel-by-channel over a short time span does not convey much about the perceived loudness of the entire audio program. Therefore, a longer-term display method is required that will also show the ITU-R BS.1770-mandated sum of the channels described earlier.

A screen capture of a comprehensive ITU-R BS.1770 loudness analysis screen is shown below. It is available in a number of Harris products.

Interpretation is straightforward — each of the first six audio meters displays the True Peak of the surround channels as defined by ITU-R BS.1770. The rightmost meter displays the combined loudness of those channels (except for LFE, as defined by the spec). The chart below the meters offers a greater understanding of loudness levels over particular program segments. This replicates a trend plot or strip chart recorder, displaying a history of the total program loudness level over a time window that can range from 15 seconds to 24 hours, allowing loudness monitoring of anything from a promo segment up to a full broadcast day. To simplify loudness monitoring even further, the average value of the historical loudness is displayed as a single numeric value.

The radar display screen shown below is complementary to the loudness analysis screen and assists engineers in visualizing the loudness of multichannel surround sound streams. By combining the multichannel audio into one display, the operator can get a better image of the program loudness. When we perceive a program is loud, we are not discerning any particular section — it is the entire program that causes our brain to interpret the audio as loud. The radar display is designed to simulate that action; when the program summary is loud, the radar display extends to the outer perimeter of the display, showing clearly that it is TOO LOUD.



The radar display also includes a meter that wraps the radar's perimeter, providing a visual "envelope" to indicate short-term loudness. This display simulates our perception of loud program content by measuring and rasterizing, and providing us with alarms that there is a problem. This innovative and intuitive display is licensed by Harris from TC Electronic.

Another tool is the Harris CMN-LA loudness analyzer, which provides comprehensive audio monitoring that makes it easy to confirm compliance with the latest loudness requirements. Loudness and true peak measurements are made to the ITU-R BS.1770 standard with five times oversampling. Metering of up to 16 channels simultaneously makes for rapid alignment checks.

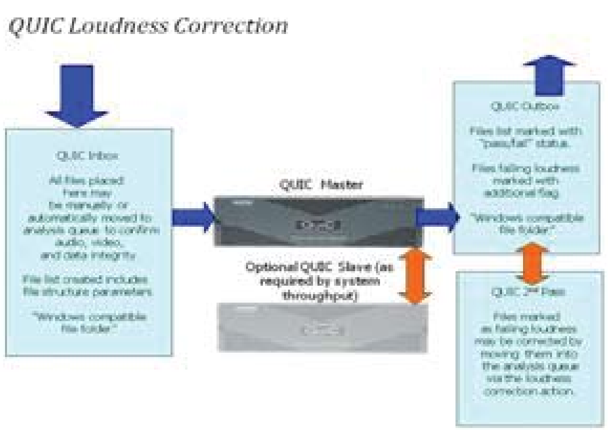
The CMN-LA provides full-screen, quad-screen or loudness display modes. In addition, it provides overlay display capabilities for picture-in-picture (PIP) functions. Quick setup and parameter changes are possible with direct access to display functions and screen location, 99 presets, context-sensitive shortcut menus and an intuitive navigation system. The CMN-LA features extensive audio loudness-related alarm capabilities. All real-time signal alarms have user-adjustable limits, time stamps from DVITC and an internal clock.

The CMN-LA is ideally suited for content owners and playout providers in quality control, troubleshooting or compliance-checking applications.

Offline, File-based Loudness Operations

Increasingly, video and audio content is stored, repurposed and played out as files, providing an offline QA opportunity at each of these stages. The Harris QuiC™ media analysis server is a fully automated, file-based test and measurement server platform that verifies the quality of compressed digital content residing on servers and storage networks before content is distributed.

The diagram below shows how the “drop folder” and pass/fail techniques are used for this purpose, thereby requiring minimal manual intervention. QuiC provides loudness evaluation based on the ITU-R BS.1770 standard and TC Electronic ® technology, and examines audio content for “center of gravity” and “consistency” in loudness parameters.



In the context of this white paper, the QuiC system’s capabilities include loudness-specific functions used to provide measurement and correction capabilities for clip loudness.

QuiC is valid for any type of normal broadcast content including dialog, music, effects and more, and conforms to the ITU-R BS.1770 standard. Center of Gravity (COG) identifies the average loudness over the entire run of a clip, and Short Term Loudness (STL) provides a measure of loudness on a frame-by-frame basis. STL alarms can also be set up to alarm on frames that exceed the desired STL threshold.

In addition to loudness monitoring capabilities, QuiC also can provide tools for the correction of loudness errors. With loudness correction, audio levels can now be adjusted to an acceptable value per the ITU-R BS.1770 standard.

Real-time Loudness Monitoring and Correction

Advanced audio applications can be implemented today with processing equipment in the Harris X85™ multiple application video and audio platform, and the NEO® and 6800+™ core processing platforms.

Consider the issue of loudness control when decoding Dolby® E and re-encoding as Dolby® Digital AC-3 — the building-block approach of the 6800+ modular products fits this application.

Another consumer issue is the switching between stereo and surround sound by the home receiver utilizing audio metadata, which can cause clicks and pops. This issue can be readily solved by utilizing the DTS Neural Loudness Control an industry-leading process for measuring and correcting loudness, and the unique DTS Neural Surround™ MultiMerge, which provides a constant 5.1 output when the input transitions between stereo and surround sound.

DTS Neural Loudness Control was designed as a real-time perceptual loudness control solution. It is based on a wideband loudness correction approach that does not modify any other aspects of the original audio signal. After measurement or Dialnorm capture, DTS Neural Loudness Control applies a user-based rule set to determine the difference between the desired target loudness that is allowed before correction is applied. After user and ballistic rules are applied, it applies attenuation or gain in the same way that an operator would. In the final stage, a soft-knee limiter subtly removes any content overshoots that exceed the capacity of the signal path.

With up to 32 channels of internal processing, the X85 multi-application processor offers unsurpassed audio capability. Audio sources can be selected from embedded audio inputs and AES inputs. For Dolby® applications, a Dolby® E/AC-3 decoder and Dolby® E or AC-3 encoder can be added with full audio metadata processing. Also available are add-on options for DTS Neural Surround™ UpMix, DownMix and MultiMerge, as well as DTS Neural Loudness Control.

Conclusion

Loudness monitoring, measurement and correction will continue to demand attention as viewers become ever more discerning.

Many useful audio metering ballistics exist, but until recently, none have accurately reflected how the human ear perceives the loudness of audio program content. Determined to rectify that situation, the ITU performed extensive research and subsequently created ITU-R BS.1770, a specification that defines a technique to monitor multichannel audio content for loudness.

Putting this specification into practice, the Harris® Videotek® CMN, TVM, and VTM Series test and measurement equipment, as well as the QuiC file-based solution, offer audio loudness measurements based on ITU-R BS.1770.

In cooperation with technology partners, Harris extends the utility of this specification by allowing short-term examination of loudness, as well as historical loudness trending and loudness logging for compliance verification. In this way, operational and technical personnel can, with minimal training, monitor content loudness by program segment or view an entire day of programming.

With standard practices in place for measuring the quality of the video signal, an equivalent standardization for the audio level of the signal is needed more than ever before. A switch to a new paradigm in audio leveling, from peak to loudness measurement, is long overdue, and the problem is being addressed via the introduction of loudness level metadata, an international standard for loudness measurement and innovative operational tools.

Harris' History in Loudness

Harris Corporation and, specifically, its Broadcast Communication Division is known worldwide as an innovator in providing relevant solutions for content production, infrastructure and delivery. Harris employs literally thousands of scientists, engineers and technicians in developing the best-in-class quality and intellectual properties among its peers in the broadcast industry.

Harris also partners with new, highly creative companies to attain the latest in user experience designs, positioning Harris products as not only of the utmost in customer cost of ownership, but also technologically fresh. Our broadcast division has developed a complete range of loudness control and measurement products and is offering the most comprehensive portfolio of interoperable products in the industry!

This work is a culmination of years of effort from our internal research and development, in concert with key industry partnerships. We recognized as early as 2005 that loudness had become a major issue in the (then) newly launched United States DTV transmissions. We soon set out to provide solutions to our customers and found that final standards were not yet in place, though working groups within the ATSC and EBU had already been given the task to find an amicable solution.

We knew that we needed to develop products offering loudness control and monitoring with the flexibility to adapt to future standards. These loudness control and monitoring products would have to accept changes as the standards were written and the likelihood that in more than one region, there may be differing standards.

In July 2006, the ITU produced a recommendation, ITU-R BS.1770/1771, which was the first effort from a standards organization to provide guidance to solving the broadcast loudness problems. Compliant Harris baseband monitoring products were beta-tested beginning in April 2007 and were released to customers, fully meeting the standards, in April 2008.

Harris began field tests of the loudness compliance control products in the summer of 2007. The first product order was shipped in September 2007. File-based loudness measurements were released in October 2007, following extensive testing, and file-based loudness correction products were released to the industry in December 2007.

HARRIS®
assuredcommunications®

Broadcast Communications Division
9800 South Meridian Boulevard | Suite 300 | Englewood, CO USA 80112 | Tel: (303) 476 5000
www.broadcast.harris.com

WP_LOUDNESS_MNTR_1210