White Paper

Loudness Measurement & Control

CALM; EBU Compliance

Interra Systems, Inc.
1601 S. De Anza Boulevard, Suite 212
Cupertino, CA 95014
USA

1. INTRODUCTION

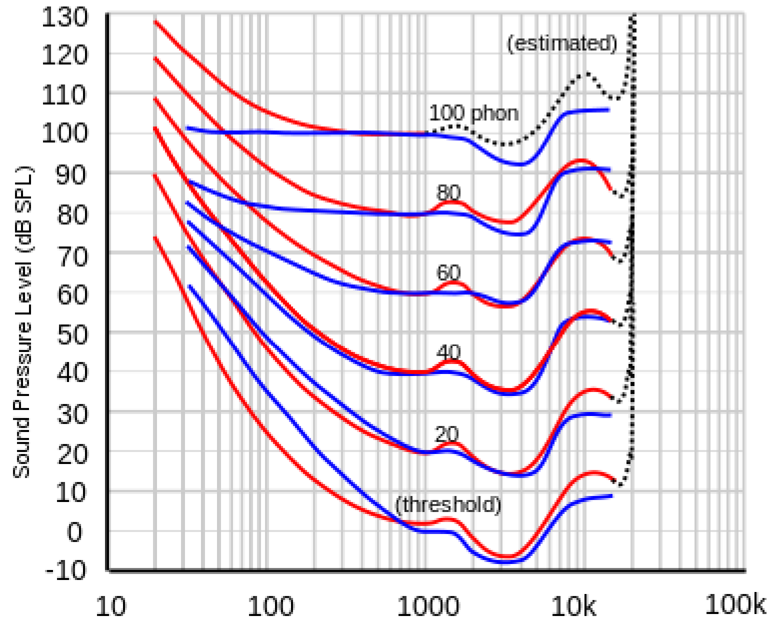
The most rudimentary issue faced by the broadcasting, satellite and cable television industry over the years has been the inconsistent loudness. The notion that people tend to pay more attention to loud records has resulted in encouraging the sound mastering engineer to produce louder records. Further, some broadcasters increase the loudness of commercials deliberately to catch the audience attention. There is a physical limitation in case of vinyl records as to how loud it can go, thus back then, even the louder records tend to retain some dynamics and never reached an extreme level. However, the advent of CDs and Digital Television (DTV) systems, removed these restrictions and resulted in hyper-compressed louder music with minimum crest factor. Hyper-compressed music, however, sound flat and may result in "listening fatigue", thus, it may discourage continued listening. This inconsistent loudness has resulted in frequent consumer complaints.

These complaints have prompted governments to come up with the regulations on loudness control. The U.S. has passed the legislation H.R. 1084 - also known as the CALM (Commercial Advertisement Loudness Mitigation) Act. CALM Act requires all broadcasters to broadcast advertisements at a loudness level no more than that of the accompanying program. Similarly, in Europe, EBU (European Broadcasting Union) has published Loudness Recommendation EBU R128. It tells how broadcasters can measure and normalize audio using Loudness meters.

This paper discusses the different loudness measurement techniques and proposes an audio normalization approach or loudness control that can be effectively used in a file-based workflow.

2. LOUDNESS MEASUREMENT: DIFFERENT TECHNIQUES AND SPECIFICATION

Loudness is the way in which we perceive sound level. Loudness is quite subjective and difficult to measure. Loudness is not only influenced by Sound Pressure Level (SPL) but also frequency, timbre, and duration of a sound.



Equal-loudness contours (red) (from ISO 226:2003 revision)
Fletcher–Munson curves shown (blue) for comparison

The first research on how the human ear hears different audio frequencies at different levels was conducted by Fletcher and Munson in 1933. The curves known as Fletcher-Munson curve(s) of equal loudness, show human ear perception against the different frequencies. Approximations of above contours have been utilized by multiple audio meters over the years and are commonly referred as Frequency Weighting Networks.

Voltage Unit (VU) and Peak Program Meters (PPM) are commonly used to measure audio level but often misused as devices for audio loudness measurement. Both of them were designed to measure the Sound Pressure Level (SPL) rather than the "perceived" audio loudness. Sound engineers circumvent these meters by employing multiband compressors achieving heavy compression, thus making a softer sound louder and bringing down the peaks, making content hyper-compressed.

3. RECOMMENDATION ITU-R BS.1770

Considering the problems in audio level based approaches, ITU (International Telecommunication Union) in the year 2006 proposed an algorithm, which could measure loudness objectively. Since then three more revisions have been released, the latest being 2015 (1770-4).

The loudness metering algorithm consists of a measurement of the energy equivalent average (Leq) with "K-Weighting" filter curve which forms the basis for matching an inherently subjective impression with an objective measurement.

"K-Weighting" accounts for the acoustic effects of the head, low frequencies are de-emphasized, boosting high frequencies. This weighing curve is applied to all the channels except the Low-Frequency Effects Channel (LFE), with different gain factors for front (~1.0 dB) and surrounds channels (~1.5 dB). The reason for this is humans' perceive sounds coming from behind louder than those from the front.

In its 2nd revision, ITU Algorithm also involves the concept of gating. The Algorithm talks about two kinds of gating - Absolute and Relative. The Absolute Gating, which is done at -70 LKFS, and it basically prevents inaudible segments from affecting the Loudness Measurement. The Relative Gating is set at –10 dB relative to the level measured after application of the first threshold. Its purpose is to ensure that any long segments with low level don't play a role in loudness measurement. Gating ensures that silence regions at the start and end of the program or pauses between the dialogues do not influence the overall program loudness.

A key feature of this algorithm is its ability to measure loudness for multi-channel audio. Measured loudness is reported in LKFS units, which is "Loudness, K-weighted, relative to Full Scale". The LKFS unit is equivalent to a decibel. A block diagram for multi-channel loudness algorithm is depicted in FIG 2.

Recently, ITU had released BS.1770-4 that specifies loudness measurements for advanced sound systems having a large number of channels. The algorithm is an extension of the above algorithm in which the third stage is modified to have weighting coefficient on the basis of azimuth and elevation angles of channel position.

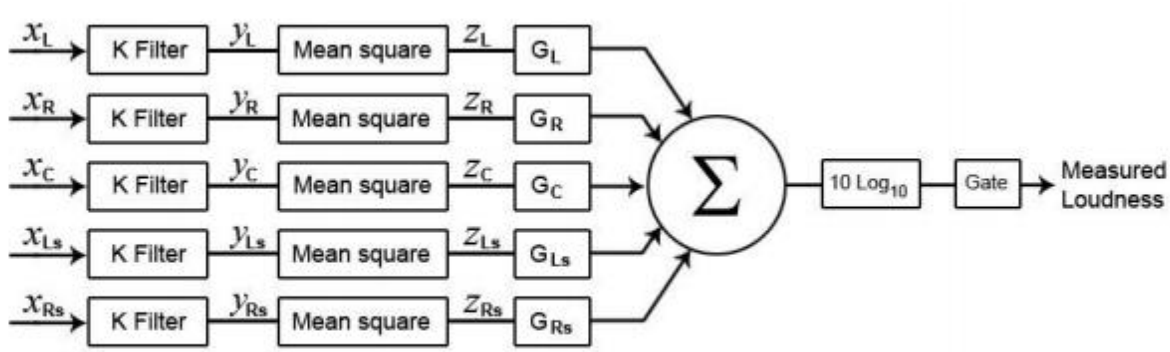


FIG 2 BLOCK DIAGRAM OF MULTI-CHANNEL LOUDNESS ALGORITHM

With this updated standard, the industry finally hopes to agree on a common algorithm for the measurement of loudness and the true peak levels of programs. ITU BS.1770 is a robust standard with simple implementation. Different organizations such as ATSC, EBU, ARIB have developed their loudness specifications based on this measurement technique. As an example, we will look into ATSC A/85 specs for loudness measurements based on the ITU standard. The next section discusses basic details of ATSC A/85.

4. CALM ACT AND ATSC A/85

CALM Act calls for compliance with the ATSC (Advanced Television Systems Committee) technical recommendation, A/85. ATSC recommends the use of BS.1770-3 based measurement technique as described above. ATSC/A85 proposes a new gating strategy based on Anchor element. The audience's impression of loudness of the program (i.e., TV show, ad, etc.) is formed by the Anchor element of the program and not by soft or loud portions of the program. Soft and loud portion create the dynamic range of the program, but the audience set the volume level of the program based on the loudness of the Anchor element. The dialog forms the Anchor element in programs where the dialog is predominant, e.g., television programs. But in cases, such as Music, the anchor element is the sound element on which the audience focuses the most.

Loudness calculation is done on the part that contains the Anchor element by applying BS.1770. Gating on the basis of the dialog is called speech or Dialog Gating, and it should be applied to long form content and speech dominated content. However, for short form content like commercials and promos, average loudness should be measured over the entire length of the program. ATSC Standard recommends program loudness of –24 LKFS with minor variation of +-2 dB

Apart from Average Loudness, ATSC A/85 talks about the following aspects related to loudness:

DIALOG NORMALIZATION

ATSC mandates the presence of Dialog normalization a metadata in the stream that reflects the average loudness of the content. It is designed to control the playback gain to maintain the loudness of the broadcast content. Dialnorm is an integer value in the range -1 to -31 corresponding to playback gain of -30 dB to 0dB. If Dialnorm value is properly encoded and used during playback, we will have consistent loudness throughout the playback of the program.

TRUE PEAK MEASUREMENTS

Audio devices that are based on PPM are not capable of displaying peaks less than 10 msec. Simply finding the absolute sample value within a given time period is also not enough. This can potentially lead to false peak detection if inter-sample peaks are ignored. With a small phase shift in audio, reported sample peak level may differ considerably. Such false or incorrect detections of true peaks can lead to audio clipping and other kinds of audio issues. True peak measurement is ideally defined as a measure of maximum absolute sample value of an audio signal in a continuous time domain. This measure can be easily achieved in the digital world by means of oversampling. ITU BS.1770 standard describes one such approach where audio signal can be over-sampled by a factor of 4 for estimating the true peak. ATSC recommends the same approach as well. ATSC recommends true peak levels should be kept below –2dBTP in order to provide sufficient headroom to avoid clipping at any of the workflow stages.

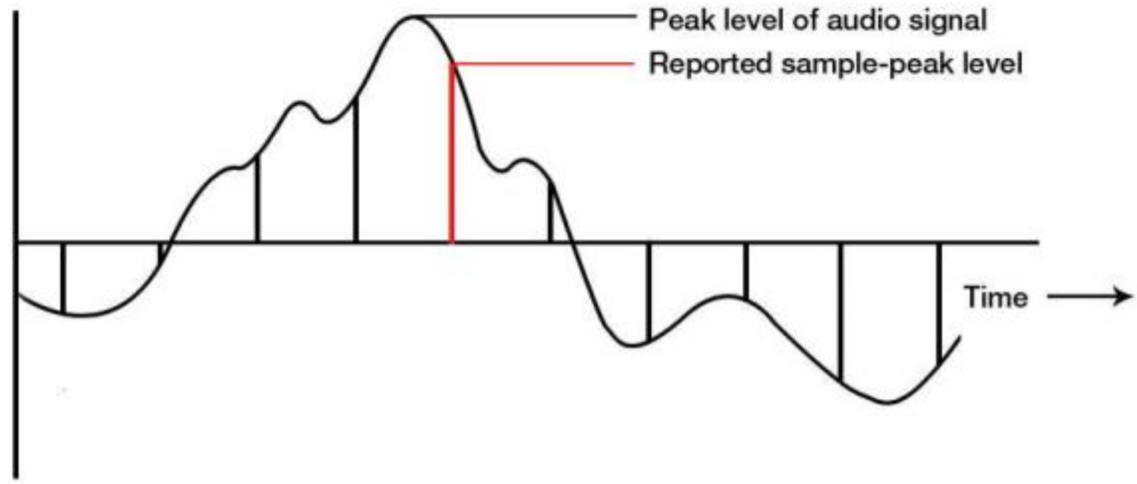


FIG3 CONTINUOUS-SIGNAL PEAK LEVEL VERSUS SAMPLE-PEAK

DYNAMIC RANGE MANAGEMENT

DTV systems are capable of delivering very wide dynamic ranges. Also, average loudness of -24 LKFS allow headroom of 24dB, thus allowing sound engineers to play with the dynamic range. Content producers can take advantage of this by adding artistic effects leading to wider dynamic ranges.

Some of these effects may be suitable for cinema viewing but may not be a good choice for platforms such as computers or smartphones. This can potentially lead to a conflicting situation. Such conflicts could arise either because of the equipment's limitations or consumer's wishes. Thus, the goal of preserving original dynamic range and satisfying the customer can be at odds. A possible solution lies in maintaining Dynamic Range Control (DRC) information within the bit stream. The Dolby AC-3 system is capable of handling DRC information within its bit stream. The system defines six compressions presets: Film Standard, Film Light, Music Standard, Music Light, Speech and None. These presets result in dynamic compression around the dialnorm value. AC-3 stream can convey the DRC options to the viewer, allowing a user to optionally choose how much dynamic range they desire.

5. EBU R128

EBU R128 is a loudness recommendation which is widely used in Europe. Apart from **Program Loudness**, EBU R128 also talks about two different types of loudness measures: **Momentary Loudness** and **Short-term Loudness**. The Momentary Loudness is basically used to measure sudden changes in audio levels, e.g., gun shots. It is measured using sliding time-window of length 400ms. The Short-term Loudness measures more general changes in levels using sliding window of 3 seconds.

In addition to loudness measurements, EBU R128 also talks about True peak level, already discussed above, as well as the Loudness Range. The **Loudness Range (LRA)** measures the variation of loudness on a macroscopic time-scale. LRA is based on statistical distribution of measured loudness, by measuring the difference between the softest and loudest part of the program. However, a lower percentile of 10%, e.g., fade out of a music track, as well as an upper percentile of 95 %, e.g., gunshot, are ignored to ensure that such events cannot manipulate the results in an undesirable way.

6. LOUDNESS CONTROL

Listening comfort of the consumer makes Loudness control an essential requirement. With the advent of various object loudness measurement techniques and compliance standards like CALM, loudness control has become even more important for content producers. The goal of loudness control should be to achieve consistent loudness across different programs as well within a program so that users don't feel the need to reach out to their remote controls. Traditionally, peak normalization has often been used as the technique to control the loudness of programs. But just peak normalization cannot properly normalize loudness based on modern metrics. Broadly, loudness control techniques can be divided into two categories.

METADATA BASED CONTROL

The metadata based control is currently being used only in Dolby formats such as AC-3, which carry the loudness specific metadata such as Dialnorm and Dynamic Range Control Profiles. Proper setting of Dialnorm and DRC Profiles is critical for automatic adjustments of audio level during playout time. If the audio material does not match these metadata, there is a need to either modify the audio signal or the metadata. An easier way is to correct the metadata because it does not require decode and re-encode in the process. So, this method aims to recode Dialnorm and dynamic range words to properly match that of a signal.

SIGNAL PROCESSING BASED CONTROL

The option of metadata based control is not available with non-Dolby formats such as PCM or AAC. This method aims to modify audio signal to achieve target loudness. Normalization algorithms need to be devised to control all three aspects of loudness: Program Loudness, True Peak and Dynamic Range.

Any normalization algorithm must follow these following requirements:

- Program Loudness correction should also consider dynamic range as a parameter during correction process. Correction should not be simple gain or attenuation process. Ignoring the dynamic range could lead to a situation where low audio levels may become so low that audio is inaudible. So at the end of the day, the viewer may increase the volume on the TV to hear the sound well but would get the original problem of loudness. Different approaches can be taken for short form and long form content to achieve better results.

- The Structural integrity of the media file should be maintained during the correction process. After normalization of audio samples, audio will need to be re-encoded and re-wrapped into the main container file. This step can potentially introduce encoding or wrapping errors in the content. If any of the video/audio/ancillary information is altered or lost during this step, it could adversely impact the content distribution chain.

- Good Processing algorithm should ensure that it doesn't lead to any kind of audio distortion. Processing of audio means altering the audio samples. If not done correctly, this can potentially lead to distortion.

7. WORKFLOWS FOR LOUDNESS CONTROL

There can be two kinds of workflows for loudness correction for any facility.

REAL TIME WORKFLOW

Real-time workflow for loudness control is required for live content such as News, Sports. Limitation of such systems is that processing must work in a single pass. So, the system cannot analyze the whole content before making correction to the audio signal. This makes the normalization of live content extremely challenging. The methodology the real-time systems adopt is around identifying local segments, calculation of average loudness and true peak level for that segment and then correcting them. These systems can only provide a quick fix for True Peak Level or short term loudness issues. It would be very difficult to achieve a desired program loudness level based on CALM or other loudness control regulations around the world.

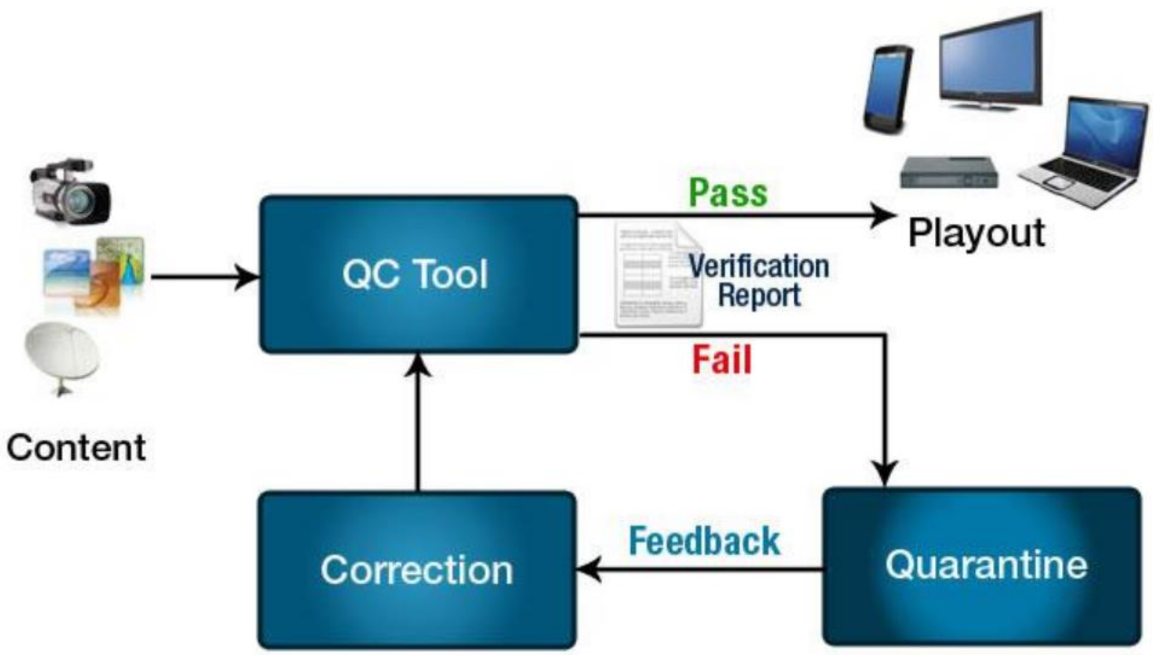
FILE-BASED WORKFLOW

File-based workflows enable much more flexibility to transform and transfer media content. Hence loudness correction and compliance to various control regulations is relatively easy. It would be preferable to identify and correct loudness issues at the point of transcode stage, to avoid propagation of loudness issues later in the workflow. Automated file-based QC or content verification systems are currently deployed at multiple stages of the workflow for verification of media content. Verification typically refers to ensuring that the content conforms to regulatory compliance while ensuring good audio and video quality. These systems can be extended to include loudness calculation and normalization as an integrated feature. Loudness control coupled with measurement offers multiple benefits in terms of optimized resources and faster content delivery to the market.

THE INTERRA SYSTEMS SOLUTION

Interra Systems provides a feature in Baton™, its enterprise-class content verification system. It is the complete solution for loudness management in file-based workflows.

Loudness Measurement and Control Step (LMC) in Baton can be explained through the following diagram.



LMC STEPS

LMC basically refers to a measurement of loudness of the media content and its verification against configurable loudness specifications. If the content meets all the requirements, it passes to the next stage. But if it fails, content is sent to the Loudness Corrector with its verification report.

The Loudness Corrector then corrects the file to conform to loudness specifications. Corrected content is passed again through the QC system for re-analysis. Both metadata correction and audio processing are performed in the correction process for Loudness Control. Click here for more: http://www.interrasystems.com/file-based-qc.php

8. ABOUT INTERRA SYSTEMS

Interra Systems provides software-based QC, monitoring, and analysis solutions to the digital media industry. The company's solutions include Baton, a market-leading, enterprise-class QC solution that automatically ensures media content readiness; Orion, a real-time content monitoring solution that simplifies the delivery of superior quality video; and Vega, a family of audio/video analyzers for standards compliance, debug, and interoperability of encoded streams.
For more information, please visit http://www.interrasystems.com .