Loudness Measurement and Control

Manik Gupta
Interra Systems
Cupertino, CA

Jozsef Pinter
Interra Systems
Cupertino, CA

Abstract - *Sudden changes in loudness because of loud commercials and variations in loudness levels between programs have been a major source of nuisance for the consumers. With the transition to Digital TV, loudness related issues have not only continued but are on the rise. Traditional techniques of loudness measurement are ineffective for digital broadcast. Techniques for digital broadcast need to be objective and accurate. More importantly, these new loudness measurement techniques need to be mandated to control loudness for a consistent consumer experience. Governments around the world are realizing the need to regulate loudness to protect consumer interests. International organizations such as ITU, ATSC, and EBU have formulated recommendations specifying concrete thresholds and restrictions for loudness measurement. This paper discusses some traditional loudness measurement techniques and shortcomings associated with these techniques. The paper provides an overview of ITU BS.1770 loudness algorithm followed by basic requirements of CALM Act. The paper highlights the need for automated solutions that can meter the loudness and accordingly normalize the audio to ensure conformance to regulations. The paper concludes with a proposal for loudness control method that can be adopted in a file-based workflow.*

INTRODUCTION

Content is an integral part of our lives spreading across multiple delivery platforms – cinema, television, internet and mobile. Media industry is always striving to improve the consumer experience. The Digital Television (DTV) system, with its expanded dynamic range, has resulted in improved consumer experience. Digital audio workflows now have greater flexibility and ability to deliver better audio quality.

With DTV systems, it was assumed that loudness issues would be a thing of past. But with the availability of wider dynamic range, issues with loudness have only increased. Today, customers complain more about loudness than in the analog days. There are several reasons for this.

The extended dynamic range of digital audio is much wider than what is available in consumer listening environments such as living room, bedroom, or while listening to audio using earplugs. Also, consumers prefer to watch programs with a more limited dynamic range during the night in order to avoid any disturbance [1]. Further, some channels increase the loudness of commercials deliberately to catch the audience's attention. Hence, audio content with high dynamic range coupled with loud commercials has led to 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.

Subsequent sections in this paper discuss the different loudness measurement techniques. Finally, the paper proposes an audio normalization approach or loudness control that can be effectively used in a file based workflow.

LOUDNESS MEASUREMENT: DIFFERENT TECHNIQUES

Audio loudness is subjective and very difficult to measure. Loudness is affected not only by sound pressure, but also by frequency, bandwidth, and duration. The first research on how the human ear hears different audio frequencies at different levels was conducted by Fletcher and Munson in 1933[2]. The curves as shown in FIG 1 depict how the human ear perceives loudness differently at different frequencies.

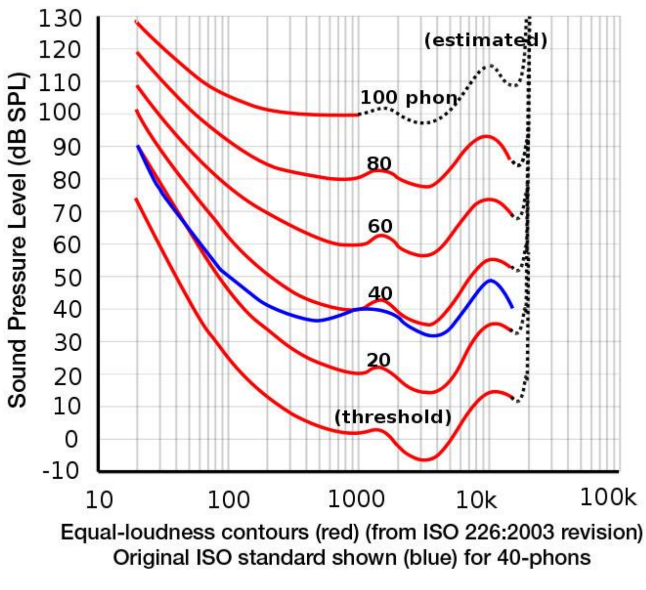

FIG 1

Approximations of above contours have been utilized by multiple audio meters over the years and are commonly referred as Frequency weighting networks. In such a network, intensity of each frequency is weighted according to the shape of equal-loudness contour. A-weighting is the most commonly used curve for measurement of sound pressure levels. A-weighting, which is based on the 40-phon Fletcher-Munson curves, has its limitation. It is applicable only for relatively quiet sounds and for pure tones. An example of a meter that uses Frequency Weighting networks is 'CBS Loudness Meter', which was developed in 1960s.

VU (Voltage Unit) meters and PPM (Peak Program Meters) are commonly used, but often misused as devices for audio loudness measurements. Both of them were designed to measure the audio levels rather than the 'perceived' audio loudness. VU meter is defined by IEC 60268-17. It has a flat frequency spectrum and low usable dynamic range. Hence, it doesn't take into account the non-linear nature of psycho-acoustic model. PPM was developed by BBC to overcome shortcomings of the VU meter. PPM is defined by IEC 60268-10. PPM was primarily designed to measure instantaneous peaks or transient in an audio signal. Neither VU or PPM accurately reflect the program loudness and are not very effective [3,4,5]. Most of the audio equipments and digital audio workstations that offer metering features include some kind of PPM or VU meter or a similar simulation. These meters cannot be trusted for loudness measurement. Sound engineers in virtual studio environments quickly learn ways to work around these meters and achieve maximum loudness out of a recording by destroying the dynamic range of the content and adding undesired distortion.

ITU-R BS.1770

In year 2006, ITU proposed an algorithm, which could objectively measure audio loudness. This algorithm was further modified in 2011 to include gating. It basically includes following steps.

- K Frequency Weighting that is composed of a first stage shelving filter simulating the head diffraction effects and a second stage 50 Hz high-pass filter
- Mean Square calculation for each channel
- Channel weighted summation followed by gating

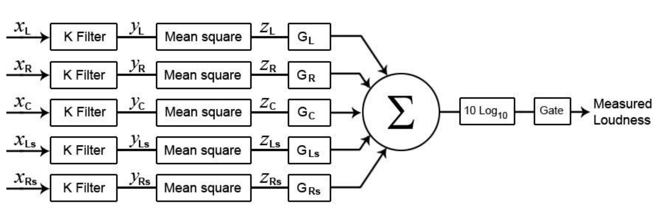

FIG 2 BLOCK DIAGRAM OF MULTI-CHANNEL LOUDNESS ALGORITHM

A key feature of this algorithm is its ability to measure loudness for multi-channel audio. All channels are processed in a similar way before channel specific scaling is done. Latest ITU recommendation (BS. 1770-2) includes gating. Gating ensures that long periods of silences or low levels do not influence the overall program loudness. Measured loudness is reported in LKFS units. A block diagram for multi-channel loudness algorithm is depicted in FIG 2[6]. 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 that has the benefit of a simple implementation. Different organizations such as ATSC, EBU, ARIB have developed their loudness specifications based on this measurement technique. The next section discusses basic details of ATSC A/85.

ATSC A-85

CALM Act calls for compliance with the ATSC (Advanced Television Systems Committee) technical recommendation, A/85[7]. ATSC recommends use of BS.1770[6] based measurement technique (without gating) as described above. Gating has been a recent addition to measurement. ATSC leaves inclusion of latest updates to user's discretion. ATSC A/85 broadly talks about three different aspects related to loudness.

- Program Loudness
- True Peak Level
- Dynamic Range

ATSC A/85 - Program Loudness

Program loudness is an indication about the average loudness of the audio content. Loudness is measured as per the ITU BS.1770 standard. The goal of ATSC A/85 is to present consistent audio loudness across commercials, programs, and channel changes. It, therefore, mandates presence of dialnorm metadata within an AC-3 Stream. Dialnorm or dialog normalization is metadata information that reflects the overall loudness of the content. It is specifically designed to control loudness of the broadcast content. Dialnorm has a finite range from –1 to –31 dB relative to 0 dBFS. It is almost similar to turning the volume up or down. Dialnorm value is determined on the basis of the loudness measurement of the anchor element. Dialogs are typically treated as anchor elements for most of the audio content. But for cases, such as music programs, anchor element could be the sound element on which viewer focuses the most. Dialnorm value is measured and then encoded within the AC-3 Stream and sent to television sets. TV sets can then control the loudness depending on the dialnorm value.

If dialnorm value is correctly set within the AC-3 Stream, then loudness remains within the comfort zone for all programs. Comfort zone is basically the range of loudness that is comfortable to a listener as depicted in FIG 3. It must be noted that program loudness may not be constant throughout the program. It can always vary within a zone according to the content requirements and artistic intent. If average loudness remains within the comfort zone during a program switch or start of commercial, a listener will not feel annoyed.

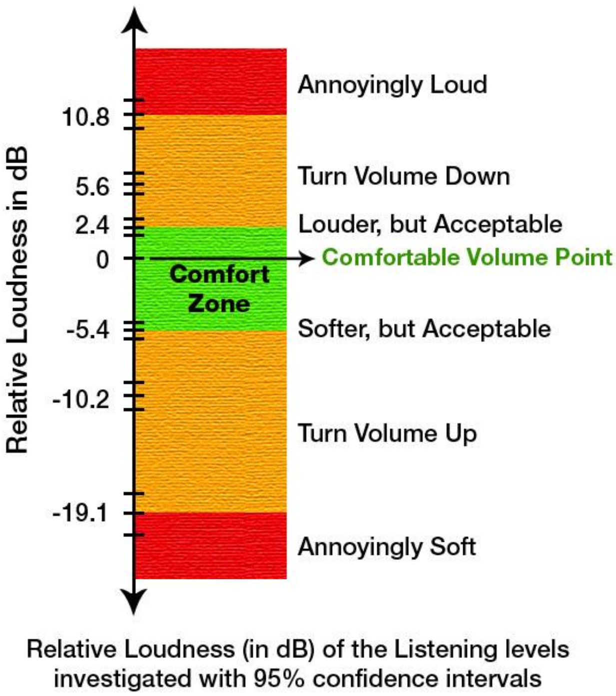

FIG 3 CRITICAL LOUDNESS LEVELS

ATSC Standard recommends program loudness of –24 LKFS with minor variation of ±2 dB.

ATSC A/85 - True Peak

Audio devices that are based on PPM are not capable of displaying peaks less than 10 msecs. 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 over sampling. 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.

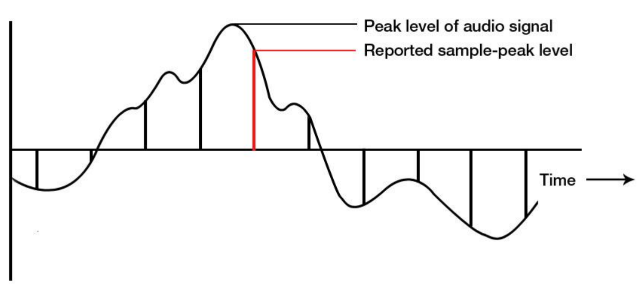

FIG 4 CONTINUOUS-SIGNAL PEAK LEVEL VERSUS SAMPLE-PEAK

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.

ATSC A/85 - Dynamic Range

DTV systems are capable of delivering very wide dynamic ranges. 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 smart phones. This can potentially lead to a conflicting situation. Such conflicts could arise either because of the equipment 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 within the AC-3 system. AC-3 system is capable of handling Dynamic Range Control (DRC) information within its bit stream. Dolby Digital system defines six compression 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. The viewer can then apply the gain control words depending on the desired dynamic range.

LOUDNESS CONTROL

With availability of objective loudness measurement techniques, it is now feasible to effectively control loudness. Hence, there is a need for a change in the leveling paradigm from peak normalization to loudness normalization. Peak normalization can still lead to variations in loudness across the content. The goal is to achieve consistent loudness across a program so that consumers do not feel the need to reach out to the remote control. Loudness control can be primarily achieved in two ways.

Metadata based Control

This approach suits formats such as AC-3, which have the provision for carrying metadata such as dialnorm, dynamic range control words, etc. AC-3 Streams should always carry a valid dialnorm value that matches with the program loudness. Audio streams that do not comply with the above requirement need correction. Software based correction tools can be deployed for such cases. The correction process for metadata is much simpler as decode and re-encode is not required. Hence, correction tools can correct the content at the metadata level. Dialnorm values can be re-authored and set to match that of the content.

Audio Signal Processing

This technique requires uncompressed audio signals to be processed to achieve target loudness. The metadata based strategy is more suitable for content after distribution stages where receivers can control the loudness on the basis of viewer’s settings. In production stages and for formats such as PCM, metadata information may not be available. Hence, a technique is needed which can process the uncompressed audio samples. Normalization algorithms can be devised which are able to control all three aspects of loudness: Program Loudness, Dynamic Range, and True Peak Level.

Metadata based control and audio signal processing can be used in conjunction with each other. Consider a feature film that has a wide dynamic range. In the first step, the audio signal can be processed in order to narrow the dynamics to be comfortable for a home theater system. In the second step, dynamic range control metadata with a Film Standard profile can be added. This will help the receiver to compress the dynamic range further if the viewer selects the 'Night Mode'.

The following points should be always considered while applying normalization algorithms to audio signals.

- Program Loudness correction should also consider dynamic range as a parameter during correction process. Correction shouldn't be a simple gain or attenuation process[8]. 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 low sound well, but would get back the original problem of loudness. Different approaches can be taken for short form and long form content to achieve better results. For long form content, anchor element should be considered for measuring loudness. For commercials or short form content, the entire content should be considered for loudness measurement.
- 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 rewrapped 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 audio distortion.

With good audio normalization algorithms, loudness can be effectively controlled at different stages in a workflow. Loudness measurement and control can be performed in two modes.

Offline Mode (For file based content)

When using file based content, there is a need to verify different loudness issues at each workflow stage and if possible correct them. Automated file-based QC or Content Verification systems are broadly deployed at multiple stages in a workflow. These systems can be used to control audio loudness in a non-real time environment. The next section in this paper describes loudness control in file-based workflows.

Real Time Mode (For News/Sports/Live Programming)

The principle of measuring the loudness applies to live productions as well. Normalization for live content is always very challenging, as it is very difficult to achieve a desired loudness. Intent of loudness measurements made during a live program is to produce content at loudness within the comfort zone.

LOUDNESS CONTROL IN FILE BASED WORKFLOWS

Due to flexibility of transforming digital media content, loudness should be measured and controlled at various stages of the workflow. Early detection of loudness issues closer to the content source prevents propagation of loudness issues and ensures efficient workflows. Automated content verification systems are currently used at different stages of the workflow for verification of media content. Verification typically includes checking the content as per the organizational requirements as well as audio and video quality assessment. The scope for such systems can be extended to include loudness normalization as an important feature. Content providers would benefit a lot from this integrated approach. Loudness control coupled with measurement offers multiple benefits to a broadcaster in terms of optimized resources and faster content delivery to the market.

In the following section, a Loudness Measurement and Control step (LMC) is described. This can be supported by content verification systems and other audio loudness measurement devices.

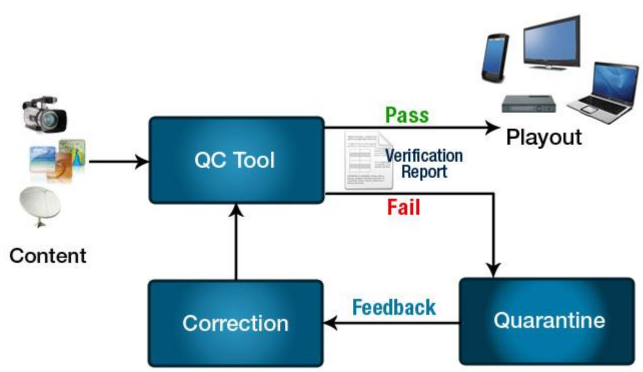

FIG 5 LMC STEPS

An LMC basically involves verification of content against the loudness specifications for that stage. If content meets all the requirements, it moves to the next stage. But if it fails, content is submitted to the Loudness Corrector along with the verification report. The Corrector can then try to correct the loudness on the basis of verification report. Corrected content can be re-fed into the QC system for re-analysis. The correction process should be able to perform both metadata based correction and also audio processing for loudness control.

The LMC step can be easily automated with use of watch folders coupled with quarantine functionality. It can then be used at various stages in the file based workflow as shown in the figure below.

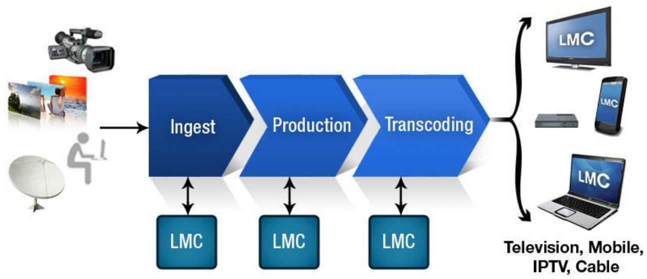

FIG 6 LMC IN A WORKFLOW

The figure above illustrates a sample workflow. The workflow begins with the ingest stage. After ingest, content is repurposed and transcoded to multiple formats to suit the needs of different delivery mediums. The LMC step after ingest is a must as it ensures that no bad content is further processed. Different delivery mediums have different loudness requirements. Loudness and dynamic range requirements differ for each delivery medium. Content meant for TV or cinema would have a wider dynamic range compared to content meant for mobiles. Hence, the LMC step can ensure that content meets the loudness requirements of each stage.

CONCLUSION

Availability of wide dynamic range in DTV has made loudness an important consumer-experience issue. Lack of objective loudness measurement in existing peak normalization techniques has led to a loudness war. With the availability of new standards, it is now feasible to control loudness objectively. International organizations such as ITU, ATSC, and EBU have formulated recommendations for loudness measurement. In the coming years, countries are expected to enforce loudness regulations, similar to the U.S. CALM Act, to protect consumer rights.

The availability of Loudness Measurement and Control (LMC) component in Automated Content Verification Systems can make loudness control easier to manage in a file-based workflow. They can be used at content exchange points between different organizations (broadcasters, content providers) working together in the content supply chain to maintain consistent loudness. The challenge for LMC components is accurate measurement of loudness and faithful normalization of audio. This would ensure faster time to delivery resulting in an efficient workflow and enhanced consumer experience.

REFERENCES

[1] Lund, T., *Control of Loudness in Digital TV*. Proceedings of NAB *BEC*. Las Vegas, NV, April *2006*.

[2] Fletcher-Munson curve, *Wikipedia* [http://en.wikipedia.org/wiki/Fletcher%E2%80%93Munson_curves]

[3] Katz, B., Mastering Audio, *The Art and the Science*, Focal Press

[4] Katz, B., *An Integrated Approach to Metering, Monitoring and Levelling Practices*, AES Journal No.9 2000

[5] Lund, T., "*Control of Loudness Across Broadcast Platforms*", *Proceedings of NAB* BEC. *Las Vegas, NV, April* 2009.

[6] *Algorithms to measure audio programme loudness and true-peak audio level*.[Document: ITU-R BS.1770-2]

[7] *ATSC Recommended Practice: Techniques for establishing and maintaining Audio Loudness for Digital Television*. [Document: ATSC A/85:2011]

[8] *Practical guidelines for Production and Implementation in accordance with EBU R 128* [Document: EBU Tech 3343]

AUTHOR INFORMATION

Manik Gupta, Interra Systems, Cupertino, CA

Jozsef Pinter, Interra Systems, Cupertino, CA