Loudness Management Is Settled…What's Next? **04-09-14**

Author: Oliver A. Masciarotte
NAB 2014 Education Program: Broadcast Engineering Conference Location: S227
Part of: CALM Compliance Issues

After well over a decade of research, discussion and collaboration, the foundations are in place for loudness management across a broad spectrum of delivery platforms. The ITU's BS.1770-3 is a recommended standard for television and a basis for future discussion regarding loudness management of radio, IP streaming and file–based playback in the home and on mobile platforms [1].

The European Union began the process and recently, our neighbor to the north adopted measures similar to ours four months ahead of us [2]. Recently, additional EU member countries have adopted R 128–based loudness recommendations, and private broadcasters in the Republic of South Africa are leading the way on that continent [3, 4, 5, 6]. Australia, New Zealand and Japan have also adopted variations on 1770 [7, 8]. In South America, the economic powerhouse that is Brasil has settled on a Programme Loudness of -24 LU and a True–Peak of -3 dBTP. Their recent mandate reinforced and updated prior legislation dating from 2001, which banned a sudden rise in loudness during commercials relative to scheduled programming [9]. That initial law, as was the case in other parts of the globe including here in the U.S., was less effective than expected due, in part, to its reliance on metadata. The digital entertainment ecosystem is a complex blend of audio, video and metadata interdependencies, and early loudness management efforts relied too heavily on contributor coöperation in the analog realm, and presupposing the behavior of decoders in the consumer's home once digital took hold.

Truly compliant loudness control, in the modern sense, is file–based. This because the ITU's 1770 family of standards dictate that Programme Loudness, the principal metric of loudness measurement, is calculated over the entire duration of a program. This means that real–time processes simply cannot explicitly control Programme Loudness. Realtime hardware processors also introduce delays that complicate workflows downstream. Programme Loudness Adjustment is a linear gain change, based on the Programme Loudness integrated over the entire program. Being a simple gain change, it preserves the mix and is basically reversible; Normalize_dB = PLmeasured – PLtarget. Also because it is a simple gain change, a running average can be used to provide a not strictly compliant AGC or automatic gain control approach in real time. True–Peak, another aspect of loudness mandates, refers to a measurement method that captures inter–sample excursions. A True-Peak limiter is an oversampling dynamics processor that takes inter–

sample peaks into account, which can cause clipping, overloads and distortion in a listener's D/A converter.

There are four fundamental loudness management methods:

- Altering the loudness level of all audio files to be "on target"
- Modifying an asset only "on demand" prior to delivery
- Using the result of a loudness measurement to adjust the program level only on playout without altering the original file
- Transporting accurate loudness metadata to the consumer where normalization is performed in their home

It was the first and last approaches that were tried in previous decades with less than complete success. Each of these techniques have their merits though, as with most things broadcast–related, only yields significant benefits through automation. One of the key improvements delivered by current standards is precise definitions for loudness measurement. This has enabled international standardization across a wide range of workflows, along with the easing of file interchange of corrected assets.

Origination is, by its very nature, a manual process…a dance between the engineer, his or her gear, the talent and the producer. That said, everyone in our industry should, by now, have incorporated the new measurement standards into their workflows. Even if engineers are not manually mixing to target, they should begin to build a sense of where their work sits in the grand scheme of modern loudness management. If your budget accommodates and it fits your workflow, use DAW–centric automation where possible to save time while improving consistency. This allows engineers to "use their ears," be creative, and not be overly concerned about precisely hitting target loudness.

As televisions have gotten thinner, and content has gone mobile, audio quality has suffered. As a result, it is more important than ever to maintain intelligibility across all delivery platforms and media. Intelligibility is not currently part of any standard. Yet, next to loudness, poor dialog intelligibility is, anecdotally, the second most common complaint heard from viewers. Dialog Consistency is a new, non–mandated metric that gauges whether the dialog amplitude is consistent enough for automated processing.

In post–production, mastering, and archiving, loudness management really comes into its own. Writing in the EBU's PLOUD loudness working group mail list, Thomas Lund observed that "…balancing decisions should be made at the production stage rather than on a composite signal downstream [10]." Equally important is the necessity of managing all content prior to slotting it into a loudness–aligned broadcast environment. If possible, loudness adjust all content and, make sure that accurate dialnorm metadata is present. Again, automate where possible and, in these stages of a program's lifecycle, a more global approach to automation that addresses the entire facility will save money, allow for additional creativity, and result in less rework. When it comes to playout, it may be prudent to adjust on demand, and it often makes sense to move normalized content back into the archive to reduce utilization of your IT infrastructure.

As a pundit said long ago, “The problem with video is audio…” Loudness control is essentially an audio problem within a video realm, which has become as much about IT as it is about audio and video. Audio people generally know how to solve audio problems, but the “file” in “file-based” is often a video asset. Loudness control for broadcasters and networks is an enterprise–level problem requiring enterprise–class solutions. Automating any process dictates that you define the task before automating it. Allow your audio experts to become part of the overall discussion with all stakeholders.

With in–house standards in place, practitioners can begin to “test drive” methodologies to see how they work in the real world. This has exposed some weaknesses in the current recommendations. One area of particular interest is the difficulty of automatically controlling loudness for High Dynamic Range (HDR) audio content, such as Hollywood blockbusters, without sacrificing the perception of the wide dynamic range that the original mix contained. Content with high dynamic range can easily have a difference of Programme Loudness to True–Peak of well over 12 dB. Many automated loudness control techniques produce results that sound compressed and less than exciting when applied to HDR programs.

Another problem area that has merited extra attention is the handling of short form content, usually commercials. Because of their very brief duration, clever mixers are able build an unpleasantly loud or annoying mix and still have interstitials hit mandated targets. It is possible to hit a Program Loudness target, but still have transients that perceptually pop out and create the sensation of being louder. Typically, commercials are aggressively peak limited. If a loudness processor further limits any transients, the quality of the audio can be seriously impaired. This means that broadcasters need additional tools, in addition to Programme Loudness and True–Peak limiting, to control this special class of content so that it intercuts seamlessly with long form shows.

There currently are other standardized parameters of interest for loudness management. These include Maximum Short-Term Loudness, Maximum Momentary Loudness and Loudness Range. Maximum Short-Term Loudness or MSTL, meaures loudness integrated over a 3 second sliding window, with adjustable target maximum STL and window overlap. If engaged, MSTL reduces overall energy, but preserves short events. Maximum Momentary Loudness (MML) limits loudness integrated over a 400 millisecond sliding window, with an adjustable maximum Momentary Loudness (amplitude) target and window overlap. MML reduces energy of momentarily annoying sonic events.

Loudness Range Adjustment, or LRA, is single band or broadband compression that works around an adjustable rotation point, typically dictated by a prior Programme Loudness or dialog–anchored measurement. LRA reduces dynamic range by making louder sounds quieter and quiet sounds louder. Dialog anchoring refers to the familiar practice of using weighted and averaged speech amplitude as a basis for later adjustment. Dialog Level can be assessed using Dolby Laboratories's proprietary but open source Dialogue Intelligence™ algorithm to automatically detect the presence of "speech" within a track. The amount of speech detected is represented by a percentage of the entire program, while an average Dialog Loudness value (level) also results from this measurement. The amount can be used by automation to decide whether program or dialog anchoring is more appropriate for a particular show.

A/85, referenced in the CALM Act, includes several mentions of dialog anchoring as part of accurate dialnorm measurement and carriage. A/85 mandates that dialnorm, "…indicates the loudness of the encoded audio." Dolby Digital's Dialog Normalization is, as its name implies, a dialog–anchored metric. Dialogue Intelligence was designed for setting dialnorm, and should be the preferred method for obtaining and setting consistent metadata across all delivery platforms. This, in turn, guarantees consistent gain compensation behavior via that metadata value. It also provides consistent DRC behavior, the dynamic range reduction mechanism built into Dolby Digital that relies on dialnorm. Note that ITU-R BS.1770-1, 2 or 3 make no mention of "dialog." All those references to anchor elements, dialog, et cetera in A/85 have nothing to do with the actual measurement of Programme Loudness, Loudness Range, True-Peak and the other modern metrics that we also rely on for loudness management. Dialog Normalization is a metadata element; it is not the audio itself. Being metadata, it does not necessarily effect the audio in the listener's home or on the go. That said, it may also have a profound effect, given that most consumers are not privy to the settings or inner working of their audio playback chain.

As with "dialog," intelligibility is not currently mentioned in any standard. Neither is there any differentiation between long form versus short form content. Modern theatrical mixes are far too dynamic for broadcast without seriously compromising fidelity and what A/85 refers to as the "artistic intent" of a program.

To those of us in the business of loudness management, it became apparent several years ago that then current methods of managing HDR mixes were not up to the task. Taking 50 or 60 dB of dynamic range and squeezing it into a 20 dB deliverable can produced less than optimum results. Transients were rounded and lost their impact while intelligibility suffered. Pumping of lower amplitude content became apparent while image shifts and soundstage diminution often joined the list of impairments.

Film mixes *are* dynamic and rely on multiple loudspeakers to perceptually balance the mix. When a 5.1 mix is created, often the measured dialogue is far away from the average program loudness…and thus not suitable for broadcast. Users “turn it up” to understand the dialog and are then annoyed during loud explosions and action scenes. Loudness Range, when applied to HDR material, produces unacceptable changes to perceived imaging and placement, while also impacting intelligibility and spectral balance. This realization that something better was needed led to what is called “Movie Adaptation," an automated multifactorial approach to managing HDR content. Its mandate: to reduce dynamics without making the audio subjectively sound compressed. At the same time, preserve the intelligibility of dialogue without resorting to equalization or psychoacoustic manipulations.

Engineers can use complex workflows that combine standard gated measurement with dialog measurement to produce acceptable results. However, the issue is not that it can be done, it is that it must be done on an individual basis for each show. Successful workflows must take the special attributes of each essence into account with little or no human intervention. Also, linear PCM lives, hand in hand, with Dolby formats and various video containers. Multiple languages and myriad delivery platforms clutter the landscape.

The key to efficient adoption, from production to playout, is intelligent, file–based automated processing that supports multiple parallel workflows as either a stand-alone island or under the control of an existing DAM/MAM infrastructure. Quality control for PCM, Dolby E and lossy content is yet another aspect of many workflows that benefit from automation. Quality control should be applied throughout production and post.

For loudness control, broadcast television should not be the end point. Over the past decade, a dramatic increase in the number of delivery platforms, from DRM and HD radio to mobile, OTT, SVOD and telematics offerings, all point to a movement in the audio community to broaden the scope of loudness management. Current work by PLOUD indicates that there is an interest in the audio community to loudness manage radio, while the CEA has established the WG15 IP Streaming Media Working Group to develop standards, bulletins and other technical documents related to audio and video content streamed over the internet. Recent numbers indicate that up to 25% of “television” viewers watch their favorite shows on mobile, handheld or computer [11]. The OTT community knows that high quality audio, delivered to headphones or home theaters, will result in more cord cutting and increased revenues.

Standards are one thing, but implementation is key to the success of any regulation. In partnership with their PLOUD group members, the EBU has provided practical guidelines for content creation, distribution and broadcast that are critical to successful adoption. Even the AES, not known for their proactive stance, last year revised the AES41 standard

(AES41-1-2012) to now include “fragile” loudness metadata carriage in the ubiquitous mp2 and mp3 codecs. With more informed normalization on playout, this could have a major impact on mobile audio and OTT delivery. Metadata for Dolby Digital and Dolby Digital Plus is also critical, as accurate dialnorm values are required for proper loudness normalization by consumer electronics such as set–top boxes, OTT and mobile platforms.

Lest we forget, A/85:2013 states that “…the BS.1770 measurement methods by all involved in audio production will assist the industry to manage audio loudness of content consistent with artistic intent.” The intent of all our efforts is the ability to intercut any and all program types without jarring changes in perceived loudness. The tools and metrics to preserve that artistic intent are available in the marketplace.

Many local and some national broadcasting institutions in the Americas still rely on outmoded or non–compliant methods to manage loudness, instead counting on realtime stopgap measures on playout best employed as backup and for penalty prevention. There appears to be a “wait and see (who is fined)…” attitude about loudness regulations, though there is good progress in the ongoing replacement of QPPM metering with certified, loudness–oriented examples.

To capture the modern viewer, deliver a great customer experience. Audio is half of that deliverable, which means that loudness must be managed without losing the vividness of the audio experience. New tools and methodologies to deliver compliant audio are available that retain quality without sacrificing productivity. The very definition of entertainment is changing, as young audiences do not differentiate between platforms and delivery methods, be it RF or IP, broadcast or podcast. That said, it is your attention to staying informed about current standards and best practices which is our best hope for a high fidelity, loudness controlled future.

References:

[1] ITU, “Rec. ITU-R BS.1770-3, Algorithms to measure audio programme loudness and true–peak audio level,” International Telecommunications Union, 2012

[2] CRTC, "Broadcasting Regulatory Policy CRTC 2012-273,” Canadian Radio-television and Telecommunications Commission, 2012

[3] CNA, “RECOMANDARE privind nivelul tăriei sonore în programele audiovizuale,” Consiliul Naţional al Audiovizualului, 2013

[4] EBU, “Annual Report 2012-13,” European Broadcast Union, 2013

[5] Valter, T., “Loudness status in Europe,” TVBEurope, London UK, May 2013

[6] Telestream, “Case Study: Managing transcoding and workflows at SuperSport,” Telestream, Nevada City USA, 2012

[7] Free TV, “Measurement & Management of Loudness in Soundtracks for Television Broadcasting,” Commercials Advice, 2012

[8] ARIB, "Operational Guidelines for Loudness of Digital Television Programs," Association of Radio Industries and Businesses, 2013
[9] 111th Congress, "House Report 111-374 COMMERCIAL ADVERTISEMENT LOUDNESS MITIGATION ACT," U.S. Government Printing Office, 2009
[10] Lund, T., e–mail communication to PLOUD listserve subscribers 01-21-14
[11] Strauss, W., "The Great Audio Divide," Audio Media, London UK, February 2014