

TV Loudness: Time for a New Approach?

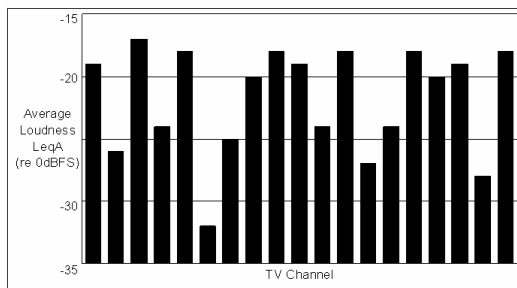
John Couling
Dolby Laboratories, Inc., England

Differences in subjective loudness between television channels and programmes are a significant cause of public annoyance. This problem cannot be solved solely through use of PPM or VU meters, or use of compression technology to reduce audio dynamic range. A method of measuring loudness as perceived by the home listener is proposed; an algorithm based on the LeqA measurement along with Dialogue Intelligence™. LeqA is a standardised technique that has been used for long-term acoustical measurement for some time. Dialogue Intelligence controls this measurement to select the same portions of an audio signal a listener uses to normalise a signal at home. Dolby Laboratories has realised this measurement in the Dolby® LM100 Broadcast Loudness Meter, which provides this measurement in a manner designed for use in broadcast and post-production.

Introduction

In analogue broadcasting, differences in loudness levels between TV programmes and channels have often been a cause of public annoyance. With the move to digital broadcasting, one might expect these problems to be solved, but this doesn't seem to be happening. In fact, with the larger dynamic range available in digital broadcast, the problem is getting worse.

Dolby recently performed a study of the loudness of material as received off-air from major European satellites and found 16dB of variation in the loudness of programming between channels and services. In a separate study looking at variations between material supplied for ingest at a satellite and cable play-out facility, a 17dB variation in the loudness of just one programming genre was identified.



Loudness of channels on a typical European satellite vary by up to 16dB

One contributory factor to these differences is the measurement technique currently employed. Broadcasters usually rely on peak programme meter (PPM) or volume unit (VU) measurements to normalize material, but neither of these account for how loud the material will be perceived by the listener at home. Rather, they indicate absolute signal levels, whereas humans use a very different long-term method to assess loudness. The listener's perception is based on factors such as the frequency response of the ear and the average level of the signal, not the peaks. When assessing loudness, the dialogue portions of the signal are particularly important - listeners will adjust their volume control so that they can understand the dialogue clearly.

Absolute measurements are, of course, still essential in order to calibrate reference level within a station to ensure correct gain structure and to measure peaks to avoid clipping and distortion. Any measurement of loudness should therefore be used in conjunction with a measurement of the true peak level, (both PPMs and VU meters under-read true peak value due to meter ballistics).

Current Practice

Current broadcasting practice is to use a combination of 'dynamic range compression' and 'peak alignment'.

In analogue television broadcasting, standard practice has been to drive programme peaks close to maximum modulation because of the limited signal-to-noise performance of the analogue transmission medium. This has led to a practice where programme levels are aligned based on the peak level, as read by the PPM or VU meter chosen by that facility.

As it is essential that you do not exceed the peak (leading to overmodulation), heavy audio compression processing is often used. The result is to give all programming a similar, restricted dynamic range, with very limited headroom above the speech peaks. These two techniques, reducing the dynamics and aligning peaks, do roughly align loudness levels, as when the dynamic range is small, and controlled, aligning the peaks will give an approximate alignment of the average programme loudness.

However, as studies have shown, there are still large discrepancies between the subjective loudness levels of different TV channels and programmes. This can be confirmed by simply flicking between TV channels or watching an advertising break.

These discrepancies are often attributed to recent developments in compressor technology, which have enabled programme producers to drive the average level of some material (often commercials) closer and closer to its peak level. As a result, some programme makers are now able to reduce programme dynamics by a much greater amount than is possible or desirable for other programming. For example a promo or commercial can be compressed so that the peak and average loudness are only a few dB apart; doing the same to a drama or film programme can result in unpleasant compressor artefacts or a loss of artistic intent. Consequently, these two signal types are no longer compressed to the same dynamic range and so when aligned by the peak level the average loudness is no longer consistent.

Variations in loudness levels are compounded by numerous other factors. These include differences in the line-up levels used over a wide selection of content, the reliance on less operators and more automation, the variety of metering standards in use, and the variety of transmission media in use.

It should always be remembered that when compression is applied to reduce the dynamics of a programme, this will lead to reduced presentation quality compared to other media. This would be of particular concern to movie services, such as pay-per-view services, which see DVD rental as a direct competitor.

A Case Study: Loudness in the Cinema

In order to understand how we might start to measure the loudness of broadcast programming, let us first look at how the film industry is solving its own loudness issues. Several years ago, people started complaining about loud replay levels in cinemas, particularly during the advertising shown before the main feature. As a result, many cinema operators turned down their systems, sometimes by as much as 10 dB. Of course, knowing that the cinema sound system was often turned down, some mixers tended to compensate and mix their ads a little louder. The situation began to spiral out of control, with cinema owners reducing levels yet further, and mixers compensating by mixing louder and louder ads. An undesirable side effect was that, once the volume of the replay system had been turned down for the pre-show content, it didn't always get turned back up for the feature itself.

A typical early solution was to use peak level alignment when mastering the content. However, differences in the meters used often caused costly rejections and confusion. Furthermore, simple peak reading does little to quantify perceived loudness or annoyance to the listener. Note that peak reading is an

even less appropriate measurement for use with multichannel sound. For these reasons, Dolby Laboratories worked with the film industry, advertising associations, and standards organisations to set up a loudness measurement standard for cinema advertising. The measure chosen was $Leq(m)$, a long-term average that correlates more closely with subjective impressions.

Dolby realised that the key to making this new standard practical and successful was to introduce a simple but effective meter specifically designed for use in film mixing studios. The Dolby meter has been joined by models from other manufacturers, and the $Leq(m)$ scale has quickly become a worldwide standard for sound-level measurement of cinema commercials and film trailers. The maximum permitted loudness of pre-feature content is progressively being reduced, and advertisers can be more confident that their mixes will not be turned down in the cinema. This is also benefitting producers of feature films, who can also now mix without being concerned the volume and impact of their movies will be reduced on playback.

Measuring Loudness in Broadcast

So how can we start to address these problems in broadcast applications? The answer depends on finding a practical method to assess the loudness of material in a broadcast environment and using this measurement to align programme levels.

Many different methods have been tried over the years to calculate how loud a human will assess a signal to be. Early analogue solutions used filters to mimic the frequency response of human ears. The most proven of these being the A-weighting curve, as used in acoustic measurement, which models the frequency response of the ear at conversational sound pressure levels. But these were not adopted in broadcast due to their expense at the time, and PPMs and VUs still do not utilize any frequency

response curve.

More recent techniques of varying complexity have been proposed, based on models of the human ear. These techniques tend to be computationally expensive and unrealistic for real-time use in a broadcast environment. The key to finding a workable solution for loudness measurement for broadcast applications is finding a simple, easy-to-repeat measurement, which can be calculated without overwhelming complexity.

In experiments Dolby found that the variation in measurement accuracy between the simplest and the most complex loudness measurement system is actually quite small, especially when compared with the variation of each listener. The $LeqA$ measurement method algorithm has been standardised and used for some time as an acoustical measurement tool. It performs a long-term A-weighted summation which produces results, in real-time, based on achievable, cost-effective processing. It therefore provides a solid basis for a broadcast loudness measurement algorithm. Importantly, the basic measurement accuracy is within the tolerance of the average listener. More complex techniques do not offer increased levels of performance to justify the expense they would incur to realise in a broadcast environment.

There is an important difference between the $LeqA$ and $Leq(m)$ weighting curves that should be noted. The A-weighting curve corresponds to the frequency response of the human ear at normal conversational levels, like those experienced at home. This makes it ideal for assessing perceived loudness and intelligibility. The m-weighting curve, applied by the film industry, is designed specifically to assess film programming at cinema playback levels focussing on how annoying a signal is⁽¹⁾.

An $LeqA$ -based measurement therefore meets all of the requirements for broadcast

loudness measurement. Results are repeatable, the measurement output is a single figure which cannot be interpreted differently by different operators, and the measurement can be performed in real-time without excessive expense. By combining a machine that can measure this with one that measures true peak, that is, with no lag time, the problem of broadcast loudness variation could be solved for analogue and digital broadcasts alike.

In the US, LeqA has already been standardised as a broadcast measurement method in ATSC standard A54⁽²⁾. The measurement is already being used by broadcasters preparing content for Dolby Digital transmission. (see box)

The Dolby Digital audio delivery system used by a growing number of broadcasters already employs a method for addressing programme loudness. A single audio data stream is transmitted to the home together with metadata – information about the audio created by the original production team. This metadata contains a value representing the loudness of a programme, and the home Dolby Digital decoder uses this metadata to automatically normalise all programmes to the same loudness level.

measure would dramatically improve measurement quality.

In order to recreate the home listeners' behaviour, Dolby introduced the concept of Dialogue Intelligence. This is a function that

allows a loudness meter to control its measurement based on dialogue, to give even more accurate and reliable results. It allows a meter to recognise the same dialogue portions of a signal that a human listener would use when setting playback volume, and uses

that knowledge to choose when to take a measurement. This additional intelligence helps every operator to make the correct measurement time after time, and makes for a very powerful algorithm.

Dialogue Intelligence

In order to measure loudness accurately, it is necessary to always measure the correct portion of a signal. For example, when watching television, a viewer will adjust the volume so that spoken dialogue is clearly audible and at a realistic, conversational level. They will not adjust the volume upwards during a quiet section of a programme where there is no dialogue, and they are unlikely to adjust it downwards following a short impulsive noise, for example a gunshot.

Therefore, once the measurement algorithm has been selected, it is essential to develop a method to decide when to apply that measurement. Previously, the choice of when to measure (or when to read the measurement value) has been a decision for the operator. Therefore, the accuracy and repeatability of the measure has often relied on the operator's interpretation. A method which automatically identifies when to

Dolby LM100 Broadcast Loudness Meter

The Dolby LM100 Broadcast Loudness Meter is the first stand-alone product capable of measuring the subjective loudness of programming and just as importantly, presenting the results in an easy-to-understand numerical format. The meter combines an algorithm based on the LeqA measurement with key improvements such as modes for both pre-recorded and live broadcast use as well as Dialogue Intelligence. The outcome is a simple measurement that can be used to adjust different content or channels to similar levels, on either analogue or digital TV services.

The meter has already been used by several major broadcasters, including ORF (Austria), SVT (Sweden), BBC (UK), Bell Express Vu (Canada), and HBO (US). It has also been used by postproduction houses, like SZM Studios (Germany) to check the

loudness levels of the material they send out to broadcasters.

If used by broadcasters during ingest and quality control, the LM100 would make it

development of advanced signal processing devices has exaggerated this problem.

The realisation of a loudness meter based on the LeqA algorithm coupled with Dialogue



possible to conform all material for storage at the same perceived loudness level. When played-out, all content would therefore be at the same, station-approved, loudness level. Outsourced material could easily be measured and conformed to a specified value before being sent to the station. At the station, a quality control process could easily check the loudness without any uncertainty due to different operators.

Multiple service satellite or cable operators could use the meter to measure the different loudness levels of individual channels within their services. Studies suggest there is often a net difference between services that can be removed through simple gain adjustment.

The meter can provide further user benefits through the addition of a range of user configurable alarms and extensive logging functions. These can provide information on a loudness history which can be used to analyse the source of loudness complaints or to understand the results of current practices.

Conclusion

Loudness problems compromise how entertaining broadcast sound can be in the home. It is natural to assume that digital technologies offer some improvement; however, this has been shown not to be the case as the variations in loudness are often even greater. The reliance on heavy compression, peak alignment, and the

Intelligence allows broadcasters to adopt a different approach. By measuring the subjective loudness of programming in a manner consistent with the home listener and using this measurement to adjust programme levels, loudness variations can be removed without compromising programme dynamics.

Dolby is working with standards bodies, including the ITU and EBU, who are currently examining loudness issues in broadcast. If a single broadcaster were to use this meter to ensure all of the material they broadcast was the same perceived loudness level, then the experience would improve for the listener at home. With the help of worldwide standards, it should be possible to address loudness issues between broadcasters too.

References

- (1) Allen, I., 1997. Are Movies Too Loud? Presented at SMPTE Film Conference, March 22, 1997.
- (2) United States Advanced Television Systems Committee. Standards Document A/54 pp. 49–52.

W03/132