On the Control and Standards of Television Program Audio Loudness: A Postprint

Authors: Shang Kechen

Date: 2023-10-08T00:00:00+00:00

Abstract

With the accelerating pace of China's modernization drive, various radio and television production organizations and broadcasting stations are flourishing, and radio and television programs are blossoming in diverse forms with each passing day. Alongside the industry's developmental momentum and the international exchange, sharing, and cooperation of radio and television program resources, numerous issues concerning non-uniform technical standards for program production have become increasingly prominent. These include, for instance, loudness variations between different programs on the same channel, loudness discrepancies between advertisements and television programs, and loudness differences across different channels. This article takes existing international standards for loudness control and domestic standards as its entry point, offering a brief analysis from the perspectives of principles, implementation, and development.

Full Text

Abstract

With the accelerating pace of modernization in China, various radio and television production institutions and broadcasters are flourishing, leading to a diverse and rapidly evolving landscape of broadcast content. However, alongside this industry growth and international exchange of broadcast program resources, numerous inconsistencies in technical production standards have become apparent. These include loudness variations between different programs on the same channel, differences between advertisements and television programs, and disparities between different channels. This paper provides a brief analysis from the perspectives of principle, implementation, and development, using existing international and domestic standards for loudness control as a starting point.

Keywords: television program audio; level; loudness control

1.1 Sound Generation

As is well known, sound is produced by the vibration of objects. These vibrating objects excite the surrounding medium, enabling sound propagation. Sound cannot travel in a vacuum. The magnitude of vibration is called amplitude, which affects sound intensity and consequently loudness. When sound propagates through air, the air pressure around the sound source changes due to the vibration, creating what is known as sound pressure. Human perception of sound intensity, or the subjective perception of sound pressure magnitude, is called loudness. Human hearing is highly complex; the ear's perception of sound intensity varies with frequency and direction. To better understand human auditory perception, researchers have developed a series of methods to simulate the characteristics of human hearing. Furthermore, sound type and timbre also affect loudness perception. Therefore, loudness analysis of broadcast television audio—which is extremely complex—cannot simply apply existing methods for pure tones, complex tones, or noise. Instead, it requires comprehensive analysis of all factors for a holistic consideration.

1.2 Types of Sound

Various types of sound exist in daily life, such as traffic noise, speech, and musical instruments. Sounds that the human ear can subjectively perceive as having pitch are physically called complex tones, while those without perceptible pitch are called noise. Some well-defined electrical noise signals fall into this category, such as pink noise, white noise, and brown noise. Another category consists of sounds produced by single waveforms vibrating at specific periods with only one frequency component; these are called pure tones. Different waveforms produce different subjective perceptions: pure tones composed of sine waves sound clean; those from triangular or sawtooth waves sound brighter; and square waves produce a harsh, piercing quality. In reality, pure tones are almost exclusively generated by artificial audio signal generators, with only a few vibrating objects (such as tuning forks) producing sounds that can be approximated as pure tones.

1.3 Measurement Units

In physics, sound produced by object vibration is measured by the changes in atmospheric pressure caused by sound wave disturbances during air propagation, quantified as sound pressure. Pressure magnitude is typically measured in Pascals (Pa). The faintest sound detectable by the human ear is approximately 20 Pa, while the sound pressure causing pain is about 20 Pa—a difference of over one million times between these magnitudes, creating significant measurement challenges. Consequently, the decibel scale is widely adopted to simplify this process. The decibel concept involves taking the logarithm of the ratio between two quantities, allowing enormous values to be expressed with small numbers. In sound pressure measurement, the minimum audible sound pressure of 20 Pa is defined as the reference value. The sound pressure level is obtained by squaring the ratio of the measured value to the reference value, taking the logarithm, and multiplying by 10. Sound pressure level is generally expressed as dB SPL (Sound Pressure Level). 0 dB SPL indicates that the measured value equals the reference value.

2. Measurement Units for TV Program Signals

In broadcast television, audio cannot be considered as simple noise, pure tones, or complex tones. Given the diverse composition of program audio signals, while sound pressure magnitude is represented by dB SPL, how should electrical signal magnitude or digital audio signal magnitude be expressed and measured?

2.1 dB, dBu, dBV, dBFS and VU

For convenient measurement of both physical sound signals and electrical signals, the decibel scale is employed. The most obvious marker of the decibel system is “dB,” which follows any quantity measured in decibels. To prevent confusion, letters are often appended to distinguish different types.

VU (Volume Unit) indicates sound signal magnitude. In the analog era, electroacoustic equipment had 600Ω impedance. At 0 VU, the voltage was 1.228 V—the voltage required to produce +4 dBm power. 0 VU was established as the reference level between devices, as shown in Figure 1 [Figure 1: see original paper]. Typically, +3 VU above the reference value serves as headroom. Although VU meters can roughly reflect audio signal magnitude, they suffer from significant overall error, poor resolution, and cannot effectively display peak levels.

Subsequently, 0.775 V (the voltage required to produce 1 mW of power) was established as the reference value, expressed as 0 dBu. Since 0.775 is a decimal value, 1 V was later adopted as the reference, creating dBV. In the digital audio era, the maximum value of digital signals—that is, the maximum data processable by computers or audio equipment—is defined as 0 dB, called dBFS (decibels relative to full scale). Consequently, decibel values in digital audio are typically negative, as signals should not exceed full scale in principle (though with today’s floating-point technology, many digital audio workstations processing 32-bit or higher precision digital audio signals allow positive dBFS values).

Among these measurement methods, VU meters appeared earlier. However, VU meters only display average signal magnitude over time, with a slow response time of approximately 300 ms, making them unable to reflect instantaneous peak signals. To address this need, PPM (Peak Program Meter) was developed. PPM can reflect instantaneous level magnitude with an integration time of only 5 ms, allowing short-term levels to be read. In digital audio processing, this enables intuitive real-time monitoring of signal levels. However, neither VU nor PPM meters reflect actual program loudness, as loudness is affected by frequency, human hearing characteristics, duration, and other factors. Therefore, these level measurement methods are unsuitable for program loudness measurement. Figure 2 [Figure 2: see original paper] shows a typical PPM meter appearance.

2.2 Reference Level

Between different manufacturers and devices, the same digital audio signal produces inconsistent electrical output levels from digital equipment, resulting in inconsistent actual loudness. Two standards currently define this parameter: the maximum electrical signal level produced when digital audio equipment is at full scale. In the EBU R68 standard, 0 dBFS corresponds to an output of +18 dBu; at -18 dBFS, the equipment outputs 0 dBu. This value is defined as the alignment level. In SMPTE RP155, the standard specifies +24 dBu; at -20 dBFS, the equipment outputs 0 VU (+4 dBu), which serves as the alignment level. Reference levels are generally used for equipment calibration and interconnection. Users should clearly understand which standard their equipment employs to avoid 6 dB level jumps during program production, particularly when exchanging programs between China and Europe or America.

3. TV Program Loudness Measurement

3.1 TV Program Signal Characteristics

Television programs encompass diverse genres and styles, including speech, music, films, TV dramas, variety shows, and live broadcasts. Different programs have different dynamic range requirements for audio. For example, pop music signals often maintain high levels, while film broadcasts contain both very quiet dialogue passages and intense war scenes. For some live broadcasts (such as billiards matches) or wildlife documentaries, the source signal magnitude is inherently limited, and human subjective perception of loudness for such events in the real world is not particularly high. Consequently, audio production for television programs faces considerable challenges. Measuring loudness across these varied programs is one task; establishing different or unified standards for different program types is another.

3.2 Frequency Weighting

When listening to pure tone signals of the same sound pressure at different frequencies, human subjective perception varies. Based on this characteristic, the concept of “equal-loudness contours” was developed. Equal-loudness contours represent that at different sound pressure levels, points on the curve are perceived as having equal loudness. Since human perception of loudness varies with frequency, frequency characteristics must be considered in loudness analysis. Equal-loudness contours reveal that human perception of low-frequency sounds is relatively poor, while sensitivity is greater in the mid-high frequency range of 2-5 kHz. Therefore, typical weighting filter algorithms suppress low frequencies while slightly boosting high frequencies. This means low-frequency sounds in programs have minimal impact on measurement results, while high-frequency sounds significantly influence measurements.

Currently, common frequency weighting standards proposed by IEC include algorithms such as IEC A, B, and C weighting. A-weighting is primarily applied to low sound pressure noise measurement, C-weighting is used for high sound pressure level evaluation such as mechanical noise and cinema sound calibration, and B-weighting falls between them. In broadcast television program reproduction, television speakers are typically used to restore signals. Therefore, ITU modified the B-weighting curve to create a specialized weighting curve suitable for television program loudness measurement, called the ITU-R BS.1770 weighting curve, or K-weighting. The weighted loudness is expressed as LK, representing K-weighted loudness magnitude, with units equivalent to dB. When audiences watch television programs, the influence of the human head on sound signals must also be considered; consequently, the recommendation specifies a 4 dB boost for frequencies above 2 kHz. All these operations aim to approximate the audience's actual listening conditions, establishing standards based on frequency considerations for different program sources. Figure 3 [Figure 3: see original paper] shows equal-loudness contours, and Figure 4 [Figure 4: see original paper] illustrates several common weighting functions.

3.3 Time Integration

In any program, sound loudness magnitude is related not only to frequency but also to duration. For example, in a TV drama, even when using the frequency weighting algorithm described above to measure loudness, short, abrupt peak signals (such as gunshots or table slaps) may occur. If numerous instantaneous levels are included in overall program loudness calculation, the measured loudness will inevitably differ from the audience's subjective perception. ITU Recommendation 1770 does not provide explicit guidance on this issue, as program loudness measurement is based on statistics across the entire program duration. Therefore, EBU provides more detailed recommendations in EBU R 128. Specifically, EBU Technical Document 3342 defines the concept of "Loudness Range (LRA)," which addresses the impact of single, short loudness mutations on overall measurement. When analyzing with a 3-second window, mutations below the 10th percentile and above the 95th percentile are excluded from calculation, while the difference within this range represents the loudness range at that moment, providing a better and more intuitive understanding of loudness distribution throughout the work [5].

3.4 Gating for Different Program Types

ITU Recommendation 1770 only provides methods for long-term or continuous loudness monitoring. For programs with extended low-loudness segments—such as the billiards matches or wildlife documentaries mentioned earlier—where narration is interspersed with only faint ambient or animal sounds, standard loudness processing might achieve the target average loudness but result in excessively loud narration. During broadcast, quiet passages might subjectively sound too loud to the audience. After evaluating the ITU algorithm, EBU distinguished LKFS as LUFS (Loudness Units relative to Full Scale). Additionally, to address the aforementioned issues, EBU introduced gating methodology. When current program loudness falls below a defined threshold, it is excluded from overall loudness calculation, ensuring more accurate representation of perceived loudness across diverse program types.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.