

STANDARDS AND INFORMATION DOCUMENTS

Call for comment on DRAFT AES71-xxxx AES Recommended Practice - Loudness Guidelines for Over the Top Television and Online Video Distribution

This document was developed by a writing group of the Audio Engineering Society Standards Committee (AESSC) and has been prepared for comment according to AES policies and procedures. It has been brought to the attention of International Electrotechnical Commission Technical Committee 100. Existing international standards relating to the subject of this document were used and referenced throughout its development.

Address comments by E-mail to standards@aes.org, or by mail to the AESSC Secretariat, Audio Engineering Society, PO Box 731, Lake Oswego, OR 97034, US.. **Only comments so addressed will be considered.** E-mail is preferred. **Comments that suggest changes must include proposed wording.** Comments shall be restricted to this document only. Send comments to other documents separately. Recipients of this document are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

This document will be approved by the AES after any adverse comment received within six weeks of the publication of this call on <http://www.aes.org/standards/comments>, **2018-03-28**, has been resolved. Any person receiving this call first through the *JAES* distribution may inform the Secretariat immediately of an intention to comment within a month of this distribution.

Because this document is a draft and is subject to change, no portion of it shall be quoted in any publication without the written permission of the AES, and all published references to it must include a prominent warning that the draft will be changed and must not be used as a standard.



The AES Standards Committee is the organization responsible for the standards program of the Audio Engineering Society. It publishes technical standards, information documents and technical reports. Working groups and task groups with a fully international membership are engaged in writing standards covering fields that include topics of specific relevance to professional audio. Membership of any AES standards working group is open to all individuals who are materially and directly affected by the documents that may be issued under the scope of that working group.

Complete information, including working group scopes and project status is available at <http://www.aes.org/standards>. Enquiries may be addressed to standards@aes.org

The AES Standards Committee is supported in part by those listed below who, as Standards Sustainers, make significant financial contribution to its operation.



This list is current as of 2018/3/5

Call for Comment on DRAFT

AES Recommended Practice - Loudness Guidelines for Over the Top Television and Online Video Distribution

Published by
Audio Engineering Society, Inc.
Copyright ©2018 by the Audio Engineering Society

Abstract:

This AES Recommended Practice (RP) describes the many issues related to online audio Loudness variations. It provides comprehensive recommendations documenting effective guidelines for managing audio Loudness of soundtracks of television and video content available to consumers by Over-The-Top (OTT) and by Online Video Distributors (OVD).

When followed, these guidelines will provide consistent Loudness, appropriate playback loudness range, reduce audio quality degradation from excessive limiting, preserve the original artistic intent, and improve the listening experience. This document does not provide specific recommendations about target playback loudness or dynamic range.

An AES standard implies a consensus of those directly and materially affected by its scope and provisions and is intended as a guide to aid the manufacturer, the consumer, and the general public. The existence of an AES standard does not in any respect preclude anyone, whether or not he or she has approved the document, from manufacturing, marketing, purchasing, or using products, processes, or procedures not in agreement with the standard. Prior to approval, all parties were provided opportunities to comment or object to any provision. Attention is drawn to the possibility that some of the elements of this AES standard or information document may be the subject of patent rights. AES shall not be held responsible for identifying any or all such patents. Approval does not assume any liability to any patent owner, nor does it assume any obligation whatever to parties adopting the standards document. Recipients of this document are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation. This document is subject to periodic review and users are cautioned to obtain the latest edition.

Audio Engineering Society Inc. 551 Fifth Avenue, New York, NY 10176, US.

www.aes.org/standards standards@aes.org

Contents

0 Introduction4
 0.1 General.....4
 0.2 Patents4
 0.3 Documentation conventions.....4
1 Scope4
2 Normative references5
3 Terms, definitions and abbreviations.....6
4 Guidelines9
 4.1 General.....9
 4.2 Delivery of content between providers and distributors9
 4.3 Distribution of content over systems with metadata capability9
 4.4 Distribution of content over systems without metadata capability9
 4.5 Distribution of content under conditions of limited dynamic range.....9
 4.6 Delivery or distribution of content where prior arrangements exist.....10
 4.7 Delivery or distribution of non-compliant content over systems with metadata capability10
 4.8 Delivery or distribution of non-compliant content over systems without metadata capability10
Annex A (normative): Content Considerations11
 A.1 Loudness11
 A.2 Peak Level11
 A.3 Dynamic Range11
 A.4 Content Type12
 A.5 Dynamic Range and Loudness Control.....12
Annex B (informative): OTT and OVD Device Considerations.....14
 B.1 Fixed Devices14
 B.2 Portable Devices.....15
Annex D (normative) Tables17
Annex E (informative): OTT and OVD Ecosystem Diagrams18
Bibliography19

Foreword

This foreword is not part of the *Loudness Guidelines for Over the Top Television and Online Video Distribution*.

The Audio Guidelines for Over the Top Television and Video Streaming (AGOTTVS) technical subcommittee was formed in early 2016 to study the many issues related to online audio Loudness variations. Its goal is to develop comprehensive recommendations, providing effective guidelines for managing audio Loudness of soundtracks of television and video content available to consumers by Over-The-Top (OTT) and by Online Video Distributors (OVD).

Comprehensive recommendations require a thorough process of user input, data collection, discussion and drafting on an ongoing basis.

In October of 2016 the group recognized an urgent need to publish preliminary loudness guidelines as AES Technical Council document AESTD1005.1.16-09 that addressed the fundamental concern of audio Loudness in the developing segment of on-line television and video content delivery, from creation through distribution and to the consumer experience.

With the release of the preliminary loudness guidelines the group raised awareness of more forthcoming, comprehensive and ongoing work and invited all interested parties to join the effort.

As a result, the AGOTTVS group increased in membership and met multiple times to continue their work drafting enhanced OTT and OVD loudness guidelines that were released as AES Technical Council document AESTD1006.1.17-10.

This group consists of volunteer members with expertise and/or interest in the creation, distribution and emission of professional audio. AGOTTVS membership is open to all stakeholders with a material interest in its work, regardless of AES membership status.

The members of the writing group that developed this document in draft are:

This draft standard was developed under project AES-X244 by writing group within task group SC-02-12-P led by Jim Starzynski, and with the following members: Frank Baumgarte, David Bialik, Tim Carroll, Roger Charlesworth, Jim DeFilippis, Chris Homer, Fabian Kuech, Thomas Lund, Scott Norcross, Kazuho Ono, Sean Richardson, Jeff Riedmiller, Steven A. Silva, Jean-Michel Trivi, and Adrian Wisbey.

Additional contributions were provided by Shujaat Ali, Leslie Gaston-Bird, Deborah Cornish, Terry Douds, Bob Katz, John Kean, Fadi Malak, Lon Neumann, Peter Poers, Jacalyn Schubring, Paul Tapper, and Will Wolcott.

Additional task group members were Eric Allamanche, Robert Bleidt, Manuel Briand, Andy Butler, Rob Byers, Richard Cabot, Dave Casey, Florian Camerer, Greg Coppa, Werner de Bruijn, Chris Fetner, Frank Foti, Ed Greene, Kimio Hamasaki, Clarence Hau, Alex Kosiorek, Andres A. Mayo, Scott Miller, Schuyler Quackenbush, Matthieu Parmentier, Skip Pizzi, Nils Peters, Patrick Waddell, Ben Waggoner, Robert Weigand, Dave Wilson, Yongjun Wu.

Morten Lave
Chair, Working Group SC-02-12

Note on normative language

In AES standards documents, sentences containing the word “shall” are requirements for compliance with the document. Sentences containing the verb “should” are strong suggestions (recommendations). Sentences giving permission use the verb “may”. Sentences expressing a possibility use the verb “can”.

Call for Comment on DRAFT

AES RP: Loudness Guidelines for Over the Top Television and Online Video Distribution

0 Introduction

0.1 General

Television content distributed as Over-the-Top-Television (OTT) or by Online Video Distribution is prone to the same Loudness management problems as early Digital Television (DTV). All provided a substantial increase in audio dynamic range capability compared to their analog predecessor. This created an opportunity for severe Loudness variation between Programs, channels and commercial advertising content. Where Loudness was not managed correctly, audiences became annoyed with the need to constantly adjust their listening volume.

Using the audio Loudness measurement recommendation, ITU-R BS.1770, organizations around the world independently developed guidelines for TV engineers to follow. These guidelines focused on maintaining and improving DTV's sonic integrity and listening experience by managing the Loudness and Loudness Range of Program and Interstitial Content. Recognizing that devices can receive multiple services or use multiple audio CODECs, guidelines were produced to align the playback Loudness of receiving devices by establishing suitable gain structures in the audio output paths.

0.2 Patents

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. AES shall not be held responsible for identifying any or all such patent rights.

0.3 Documentation conventions

Capitalization denotes the term is defined in the document glossary.

1 Scope

This AES document addresses OTT and OVD Loudness challenges by leveraging the established practices, noted above, providing new guidelines focused on the Loudness and Content Dynamic Range for connected set-top and mobile devices. When followed, these guidelines will:

- Provide consistent Loudness across different Programs, service providers and advertising content
- Provide appropriate ranges for playback Loudness for different devices and listening conditions
- Prevent excessive Peak Limiting or other processing from degrading the audio quality
- Preserve the artistic intent of wide Content Dynamic Range (movies, drama, live music)
- Improve the listening experience

Note that this document does not provide specific recommendations about device target playback loudness, Loudness Range or dynamic range.

For the purpose of this document OTT is defined as: The means to deliver video content via streaming, VOD, pay TV, IPTV and download via IP mechanisms. OVD is defined as any entity that offers video content by means of Internet Protocol (IP)-based transmission paths provided by a person or entity other than the OVD. OTT and OVD do not include delivery of content via means of traditional distribution e.g., broadcast TV, cable TV, satellite TV, and telco supplied TV, etc.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ATSC A/85-2013 *Techniques for Establishing and Maintaining Audio Loudness for Digital Television*, March 2013, Advanced Television Systems Committee (ATSC), Washington DC, US, <https://www.atsc.org/recommended-practice/a85-techniques-for-establishing-and-maintaining-audio-loudness-for-digital-television/>

EBU R128 *Loudness Normalisation and Permitted Maximum Level of Audio Signals*, June 2014, European Broadcast Union, EBU Technical Centre, Geneva, Switzerland, <https://tech.ebu.ch/docs/r/r128.pdf>

ARIB TR-B32 *Operational Guidelines for Loudness of Digital Television Programs*, September 2016, Association of Radio Industries and Businesses (ARIB), Tokyo, Japan, https://www.arib.or.jp/english/std_tr/broadcasting/tr-b32.html

OP-59 *Measurement and Management of Loudness in Soundtracks for Television Broadcasting*, February 2013, Free TV Australia, Mosman NSW Australia, http://www.freetv.com.au/media/Engineering/Free_TV_OP59_Measurement_and_Management_of_Loudness_for_TV_Broadcasting_Issue_3_July_2016.pdf

ITU-R BS.1770-4 *Algorithms to measure audio programme loudness and true-peak audio level*, October 2015, International Telecommunications Union (ITU), Geneva, Switzerland, http://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1770-4-201510-I!!PDF-E.pdf

ITU-R BS.1771-1 *Requirements for Loudness and True-peak Indicating Meters*, January 2012, International Telecommunications Union (ITU), Geneva, Switzerland, https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1771-1-201201-I!!PDF-E.pdf

3 Terms, definitions and abbreviations

For the purposes of this document, the following terms, definitions, and abbreviations apply.

3.1

Anchor Element

The perceptual Loudness reference point or element around which other elements are balanced in producing the final mix of the content, or that a reasonable viewer would focus on when setting the volume control. It is typically speech or dialog in broadcast or movie content. See ATSC A/85.

3.2

ITU-R BS.1770

Specifies an algorithm that provides a numerical value that estimates the perceived Loudness of the content that is measured. Loudness meters and measurement tools which have implemented the BS.1770 algorithm will report Loudness in units of “LKFS” [ATSC A/85] or “LUFS”. [EBU R128]

3.3

Casting

A process of:

Mirroring the playback of content from one device to another device.

A device indicating to a different target device what the target device should play

3.4

Content Dynamic Range

The difference between the quietest and loudest sound in an audio program or piece of content. It should be noted that a large difference between the Loudness of the dialog and the Loudness of the Full Program Mix Measurement may be an indication of wide Content Dynamic Range. Similarly, large values of LRA may be an indication of wide Content Dynamic Range.

3.5

Contouring

Process which modifies the audio program to better adapt the reproduced dynamic range of the content to the device and/or listening environment

3.6

dB FS

Amplitude expressed as a level in decibels relative to 100% full-scale. [AES17]

3.7

Dialog Level

The Loudness, in LKFS units, of the Anchor Element. See ATSC A/85.

3.8

Dynamic range control

DRC

The process of continually adjusting audio signal level to control the level difference between loud and soft passages according to some desired objective.

3.9

DRC Profile

A collection of parameters that describe how dynamic range control metadata is calculated.

3.10

File-Based-Scaling Device

A device used to apply an overall gain correction to audio content stored as files.

3.11

Full Program Mix Measurement

The Integrated Loudness measurement using the ITU-R BS.1770 algorithm, in units of LKFS or LUFS based on all channels and all elements of the audio content, over the duration of the program.

3.12

Integrated Loudness

An objective measurement of audio loudness of a program averaged over the entire length of the program.

3.13

Interstitial Content

Advertising, commercial, promotional or public service related material or essence. The typical duration is less than approximately two to three minutes. Synonymous with Short Form Content. See ATSC A/85.

3.14

LKFS

Unit of Loudness, K-weighted, relative to full scale, measured with equipment that implements the algorithm specified by ITU-R BS.1770. A unit of LKFS is equivalent to a decibel. See ATSC A/85.

3.15

Long Form Content

Show or Program related material or essence. The typical duration is greater than approximately two to three minutes. See ATSC A/85.

3.16

Loudness

A subjective measurement of the perceived audio level of a program or other piece of content

3.17

Loudness Normalization

The process of applying a global offset, based on the measured Integrated Loudness of an audio Program, such that after the offset is applied, the audio Program will be at the desired target Loudness.

3.18

Loudness Range

LRA

Quantifies the variation in a time-varying Loudness measurement. See EBU Tech 3342.

3.19

Loudness Unit

LU

Relative Loudness is expressed as a logarithmic ratio, 1 dB increase in level equals 1 LU increase in Loudness. See ITU-R BS.1771-1.

3.20

LUFS

Unit of Loudness, K-weighted, expressed as a level in LU relative to 100% full-scale. See EBU R128.

3.21

Maximum True Peak

The largest True Peak in the Program under consideration. See ATSC A/85.

3.22

Measured Loudness

The value reported, in units of LKFS or LUFS, when an audio signal is measured with equipment that implements the algorithm specified by ITU-R BS.1770. It is an approximation of perceived Loudness.

3.23

Over-the Top-Television

OTT

The means to deliver video content via streaming, VOD, pay TV, IPTV and download via IP mechanisms.

3.24

Online Video Distributor

OVD

Any entity that offers video content by means of the Internet or other Internet Protocol (IP)-based transmission path provided by a person or entity other than the OVD. See FCC 13-99.

3.25

Peak Limiting

Limiting

A process allowing signals below a specified input power or level to pass unaffected, while attenuating the peaks of stronger signals that exceed this threshold. Limiting is a type of dynamic range compression. Clipping is an extreme version of limiting.

3.26

Peak Program meter

PPM

A device which measures and displays or records the maximum excursion of an audio program.

3.27

Program

An individual, self-contained audio-visual or audio-only item to be presented in radio, television or other electronic media. A program segment or program act, an advertisement (commercial), trailer, promotional item ('promo'), interstitial or similar item shall be considered to be a Program in this context. See EBU R128..

3.28

Program Loudness

The Integrated Loudness over the duration of a Program. See EBU R128.

3.29

Short Form Content

Advertising, commercial, promotional or public service related material or essence. The typical duration is less than approximately two to three minutes. Synonymous with Interstitial Content. See ATSC A/85.

3.30

Target Program Loudness

A specified Loudness value for the Anchor Element, Dialog Level or Program Loudness, established to facilitate content exchange from a supplier to an operator. See ATSC A/85.

3.31

True Peak

The maximum absolute level of the signal waveform in the continuous time domain, measured per BS.1770. Its units are dBTP (decibels relative to nominal 100% true-peak).

4 Guidelines

4.1 General

To preserve the original sonic integrity of OTT and OVD content, to reduce annoying Loudness jumps when switching content, to prevent clipping and to provide an appropriate listening level across different devices, it is recommended that content providers and content distributors follow some basic guidelines. These vary based on the content being distributed and the application, as follows:

4.2 Delivery of content between providers and distributors

For delivery of content between providers and distributors it is recommended that:

- a) Where there is no prior arrangement by the parties regarding content delivery, the appropriate broadcast regional content delivery and exchange recommendations should be followed (see Annex D, Tables D.1 & D.2)
- b) An Anchor Element (for instance, dialog) should be used for the Integrated Loudness measurement, in lieu of a Full Program Mix Measurement, for conditions noted in the broadcast regional recommendations (see Annex D, Tables D.1 & D.2)
- c) Loudness and dynamic range control metadata that matches the content should be included

4.3 Distribution of content over systems with metadata capability

For distribution of content over systems with certain metadata capability it is recommended that:

- a) Where there is no prior arrangement by the parties regarding content distribution, the appropriate broadcast regional content delivery and exchange recommendations should be followed (see Annex D, Tables D.1 & D.2)
- b) An Anchor Element (for instance, dialog) should be used for the Integrated Loudness measurement, in lieu of a Full Program Mix Measurement, for conditions noted in the broadcast regional recommendations (see Annex D, Tables D.1 & D.2)
- c) Loudness and dynamic range control metadata that matches the content should be included

4.4 Distribution of content over systems without metadata capability

For distribution of content over systems with unknown or uncertain metadata capability it is recommended that:

- a) Where there is no prior arrangement by the parties regarding content distribution, the appropriate broadcast regional content delivery and exchange recommendations should be followed (see Annex D, Tables D.1 & D.2)
- b) An Anchor Element (for instance, dialog) should be used for the Integrated Loudness measurement, in lieu of a Full Program Mix Measurement, for conditions noted in the broadcast regional recommendations (see Annex D, Tables D.1 & D.2).

4.5 Distribution of content under conditions of limited dynamic range

For distribution of content for devices or listening conditions with limited dynamic range it is recommended that:

- i. If the distribution system supports metadata, ensure an appropriate DRC Profile is defined and metadata included for limited dynamic range playback
- ii. If the distribution system does not support metadata, application of Loudness and/or dynamic range control before encoding may be necessary to contour the audio

4.6 Delivery or distribution of content where prior arrangements exist

For delivery or distribution of content where prior arrangements exist it is recommended that:

- a) The maximum values for Loudness and True Peak are not exceeded (see Annex D, Table D.3).
- b) The Loudness and True Peak level of content should be measured for compliance either with values agreed to by prior arrangement or with the appropriate regional broadcast recommendations (see Annex D, Tables D.1 & D.2)

4.7 Delivery or distribution of non-compliant content over systems with metadata capability

For delivery or distribution of non-compliant content (content created under no recommendations or standards) over systems with certain metadata capability:

The Loudness and True Peak level of content should be measured for compliance either with values agreed to by prior arrangement or with the appropriate broadcast regional recommendations (see Annex D, Tables D.1 and D.2). Subject to contractual considerations, any non-compliant content should be:

- i. Normalized to a value agreed to by prior arrangement and authored with Loudness and DRC metadata matching the actual normalized content

Or

- ii. Normalized to the appropriate broadcast regional recommendations (see Annex D, Tables D.1 and D.2) and authored with Loudness and DRC metadata matching the actual normalized content

Or

- iii. Unaltered and authored with the Loudness and DRC metadata, matching the actual non-compliant content

4.8 Delivery or distribution of non-compliant content over systems without metadata capability

For delivery or distribution of non-compliant content (content created under no recommendations or standards) over systems with unknown or uncertain metadata capability where there is no prior arrangement by the parties regarding content distribution:

The Loudness and True Peak level of content should be measured for compliance with the appropriate broadcast regional recommendations (see Annex D, Tables D.1 and D.2). Subject to contractual considerations, any non-compliant content should be normalized to these recommendations.

Annex A (normative): Content Considerations

A.1 Loudness

Loudness is a subjective measurement of the perceived audio level of a Program or other piece of content. Program Loudness is the Integrated Loudness (as defined by ITU-R BS.1770) measured across the entire Program mix or the Anchor Element (such as dialog). Content having the same Program Loudness will have the same perceived Loudness to the listener.

Loudness Normalization achieves equal average Loudness of Programs while allowing the peak level to vary depending on the content and the artistic and technical needs. Generally, higher Target Program Loudness equates to lower Program dynamic range.

There are two ways to achieve Loudness Normalization for the listener, one is to actually normalize the audio signal, and the other is to use Loudness metadata to describe how loud a Program is. Not all consumer devices support the metadata solution and so the former method is recommended, with appropriate metadata, to provide the best compatibility across consumer devices.

A.2 Peak Level

Peaks in the audio signal generally do not affect the Loudness. They do, however, have a significant effect on perceived audio quality. A Program with a high peak to Loudness ratio is often perceived as sounding clearer and less fatiguing than one that has been excessively peak limited.

The ITU-R BS.1770 defines True Peaks as the peak level taking into account the fact that the actual audio peak may occur between digital samples and so be higher than any digital sample in the program and may even be above 0 dB FS. It is advisable to author content with a maximum True Peak level below 0 dB FS (see Annex D) to avoid clipping in decoding and playback systems.

The lower the Target Program Loudness the less likely peak control will be required. Programs that require the audio level to be increased to meet the required Target Program Loudness will need to have special attention paid to the maximum True Peak levels.

A.3 Dynamic Range

It is strongly encouraged that the original dynamic range of a Program be maintained to preserve the sonic integrity and artistic intent of the content. However it is recognized that there are situations where this may not be possible or desirable:

- a) The required distribution Loudness is higher than the delivered Program Loudness
- b) The delivered Content Dynamic Range exceeds the consumer's system dynamic range
- c) The consumer is in a noisy environment such as an aircraft

System dynamic range is the ratio between the minimum and maximum signal level a system can reproduce. In some countries (due to government regulations), personal listening devices may have a limited output level.

Some Programs may exceed the dynamic range of typical broadcast content. Examples may include cinema derived or premium episodic material created in cinema-style mixing environments where a lower target Loudness value is common. In these cases, an anchor-element (for instance, dialog) should be used for Integrated Loudness measurement, in lieu of a Full Program Mix Measurement for conditions noted in the regional recommendations.

A.4 Content Type

A.4.1 Short Form Content

Care should be taken when measuring Short Form Content such as advertisements, promotional items and interstitials. An infinite term measurement of all channels for the entirety of the content should be taken (refer to regional guidelines for measurement of Short Form Content).

In all cases, for all regions, see table D1.

A.4.2 Long Form Content

Care should be taken when measuring the Loudness of Long Form Content such as typical TV Programs (news, sports, variety shows, drama and movies).

For ATSC Regions, the Anchor Element (typically dialog) or a representative sample of the Anchor Element should be measured using BS.1770. If the Anchor Element cannot be isolated and measured, then the long term integrated or average Loudness of the contents full Program mix, over its entire duration, should be measured and reported (see ATSC A/85).

For EBU Regions, an infinite term measurement of all channels of the entirety of the content is the most general measurement method and appropriate for the vast majority of broadcast content. For audio Programs that have wide dynamic range where the LRA is greater than 20 LU (see EBU Tech 3343 sect 4.1) or where the voice Loudness is more than 3–5 LU different than the Loudness measurement of the Program Loudness Level then Anchor Element based Loudness Normalization may be used (see EBU R128).

In all cases, for all regions, see table D.2

A.4.3 System Generated Sounds:

These guidelines focus on television and video content from creation through distribution to the consumer. However, the Loudness of other audio sources on both fixed and mobile streaming devices are part of the full listening experience. Application Developers should strive to match the Loudness of system generated sounds to T.V. or video content to achieve a pleasing experience whenever appropriate.

Notwithstanding, it is understood that Loudness plays an important part in distinguishing incoming “rings” over content, localization of sound to picture, the delivery of content to the impaired and the alerting and delivery of important emergency information (e.g., Wireless Emergency Alerts-WEA), etc. when necessary. An increase in the Loudness of system generated sounds and related content above other sounds, in these and other similar situations when appropriate, is expected.

A.5 Dynamic Range and Loudness Control

A.5.1 Production & Distribution Environment

It is strongly encouraged to always include Loudness and dynamic range control metadata that matches the content when the distribution system supports it. This practice permits content to be contoured, in a non-destructive manner, for different devices and listening scenarios during playback. The use and support of Loudness and dynamic range control metadata, including support in audio CODECs is addressed and documented in ATSC A/85 and EBU Tech Doc 3344. A CODEC agnostic method of carrying Loudness and dynamic range control metadata is documented in MPEG-D DRC (ISO/IEC 23003-4).

The use of defined DRC characteristics or profiles is encouraged when authoring content metadata for distribution. This practice provides the opportunity to properly adjust the target Loudness and dynamic range DRC Profile on the playback device when necessary to account for limitations of the device, user preferences, Loudness or the listening environment.

A.5.2 Playback Environment

It is recognized that there currently is a large population of devices that use multiple CODECs. One or more of these CODECs may not support metadata. While many of these devices have successfully addressed this issue by adherence to the recommendations in CTA CEB11, EBU R128, and EBU Tech Doc 3344, where applicable, not all are universally applied, and can result in content being reproduced at differing target Loudness depending on the device and its method of connection.

Annex B (informative): OTT and OVD Device Considerations

Portable devices are used in many different locations that have many diverse environmental conditions. Most engaged listeners use headphones for every day content exhibition. In noisy environments DRC can be used to hear quiet passages. At home, consumers can also directly attach their devices to a TV or an AVR via HDMI or Cast via different wireless technologies. All of these scenarios create major challenges for quality reproduction of the audio.

Content distributed with a CODEC using correct metadata benefits from the communication of device capabilities over media interfaces. HDMI handshaking and the EDID data field can identify CODEC compatibility between devices that facilitates, for example, content and listener specific adjustment of target loudness and DRC. Content distributed with a system not using metadata will be dependent on fixed, non-content specific onboard device software, hardware, or combinations of both, to process the audio which leads to unpredictable presentation when connected to another device.

B.1 Fixed Devices

B.1.1 General

Home playback environments vary greatly. The noise level from external and internal sources like nearby traffic, appliances, other family members, distance from the sound source, room reverberation, quality of speakers, speaker placement and quality of mix greatly influence what the listeners hears. In all scenarios, mix intelligibility is an important factor for the listener.

B.1.2 Applications:

Fixed devices are generally intended for media playback in a home or other indoor environment e.g.: living rooms, hotel rooms, office spaces, classrooms etc.

B.1.3 Typical Capabilities:

Fixed Devices are intended to be paired with televisions, sound bars, or audio-video receivers.

Content is streamed via:

- Over the Top Television (OTT) application
- Online Video Distribution (OVD) or live streaming service
- Over the Air traditional broadcast
- MVPD (cable, satellite, telco) Transport Stream or IP services

Some fixed devices use internal hard drives to store (via download or side-load) or record programming for playback over extended periods of time. Many devices also support Blu-ray Disc or DVD playback.

These devices have the capability to reproduce the full audio dynamic range of content as supplied in the incoming audio stream. Devices also contain hardware and software digital signal processing (DSP) and decoding, capable of contouring the audio to the listener's needs. Examples of these are:

- Speaker/room equalization systems
- added room reverberation
- Metadata enabled decoder Loudness Normalization and dynamic range control
- independent dynamic range control compression

These features can be used independently or in combination and are designed to provide flexibility to the consumer to fine tune their listening experience.

B.2 Portable Devices

B.2.1 Loudness and DRC Applications in Typical Playback Environments

Similar to the features of fixed devices, portable devices may offer different audio Loudness and dynamic range processing configurations. This adaptability is a key to optimizing the sound of the personal streaming player for the constraints posed by challenging acoustic listening environments.

B.2.2 Applications

These devices support in-home and outdoor use. Most handhelds and tablets began as personal media audio players and developed to fully functioning personal communication devices with advanced video and audio playback capabilities. The smartphone is the dominant consumer device in this category.

B.2.3 Typical Capabilities

Handheld and tablet devices can be used as personal players accessing locally stored files or streaming content via Wi-Fi or cellular connection to the internet.

Content can be:

- Downloaded for later playback
- Side loaded for later playback
- Streamed using adaptive bitrate profiles

Multiple factors influence the quality of the audio from portable devices:

- Available bandwidth when the content is acquired
- Available local storage on the portable device
- DSP and/or CODEC support on the portable device
- Processing power of the portable device for app based decoding and DSP
- The maximum quality made available in the adaptive bitrate stream

Additional factors are the output configurations of the portable device:

- Internal mono or stereo micro speaker(s) with limited dynamic range and frequency response
- Connected devices (e.g.: headphones) with limited range (compared to connected devices in full range mode) and varied frequency response
- Bluetooth devices including headphones, speakers, and AVR's. Dynamic range and frequency response of content can be limited by the Bluetooth version, profile (CODEC) and available bandwidth.
- Wi-Fi and Bluetooth Casting (E.g.: AirPlay, Chromecast) can support quality playback depending on the available bandwidth
- HDMI (direct wired connection) can provide an unaltered, coded audio stream that is intended to be decoded downstream by a device with enhanced audio capability; or a decoded PCM version of the audio, subject to metadata enhanced audio processing or internal processing.
- The amplitude limitations inherent to some devices
- CENELEC regulations impact playback loudness of some devices in some parts of the world.

Annex C (informative): Mobile Loudness Pretesting Summary

The AGOTTVS group performed an informal Loudness survey that was carried out in August of 2017 with expert and non-expert listeners from various regions around the world and included several AGOTTVS members. The test was performed using seven selections of broadcast content, provided by AGOTTVS members, and normalized to three different Loudness targets, -24, -20 and -16 LKFS to determine if the volume on typical portable devices could be set to produce intelligible results.

Participants listened to the audio clips in quiet and noisy environments over headphones and the internal loudspeaker of the device they were using. The results indicated that in quiet environments, all clips could be set to an intelligible level, though some clips were deemed to have parts reported as “too loud.” In noisy environments, most clips could be set to an intelligible level, though it resulted in an even higher quantity of parts reported as “too loud.”

In summary, this preliminary work gives confidence that while the three target levels are practical, dynamic range control of some sort will be required, particularly for noisy environments. It also clarifies the need for more detailed testing.

Annex D (normative) Tables

Table D.1 - Recommendations for Short Form Content

Broadcast Region	Standard	Integrated Loudness	Maximum Short Term Loudness	Maximum True Peak	Anchor Element Measurement	Full Program Mix Measurement
North America	A/85	-24 ± 2 LKFS	N/A	-2 dB TP	Not permitted	Recommended
Europe	R128	-23 ± 0.5 LUFS	-18 LUFS and + 5 LU relative to IL	-1 dB TP	Not permitted	Recommended
Japan	TR-B32	-24 ± 1 LKFS	N/A	-1 dB TP	Not permitted	Recommended
Australia	OP-59	-24 ± 1 LKFS	N/A	-2 dB TP	Not permitted	Recommended

Table D.2 - Recommendations for Long Form Content

Broadcast Region	Standard	Integrated Loudness	Maximum Short Term Loudness	Maximum True Peak	Anchor Element Measurement	Full Program Mix Measurement
North America	A/85	-24 ± 2 LKFS	N/A	-2 dB TP	Recommended	Conditionally Permitted ¹
Europe	R128	-23 ± 0.5 LUFS	N/A	-1 dB TP	Conditionally Permitted ²	Recommended
Japan	TR-B32	-24 ± 1 LKFS	N/A	-1 dB TP	Not permitted	Recommended
Australia	OP-59	-24 ± 1 LKFS	N/A	-2 dB TP	Permitted	Permitted

1) If the Anchor Element cannot be isolated and measured (per ATSC A/85)

2) Permitted for wide dynamic range content (per EBU R128)

Table D.3 - Recommendations for maximum Loudness

	Integrated Loudness	Maximum True Peak
Short Form Content	-16 ± 1 LKFS/LUFS ³	-1 dB TP
Long Form Content	-16 ± 1 LKFS/LUFS ³	-1 dB TP

3) Not applicable for Japan

Annex E (informative): OTT and OVD Ecosystem Diagrams

OTT / OVD AUDIO WORKFLOW

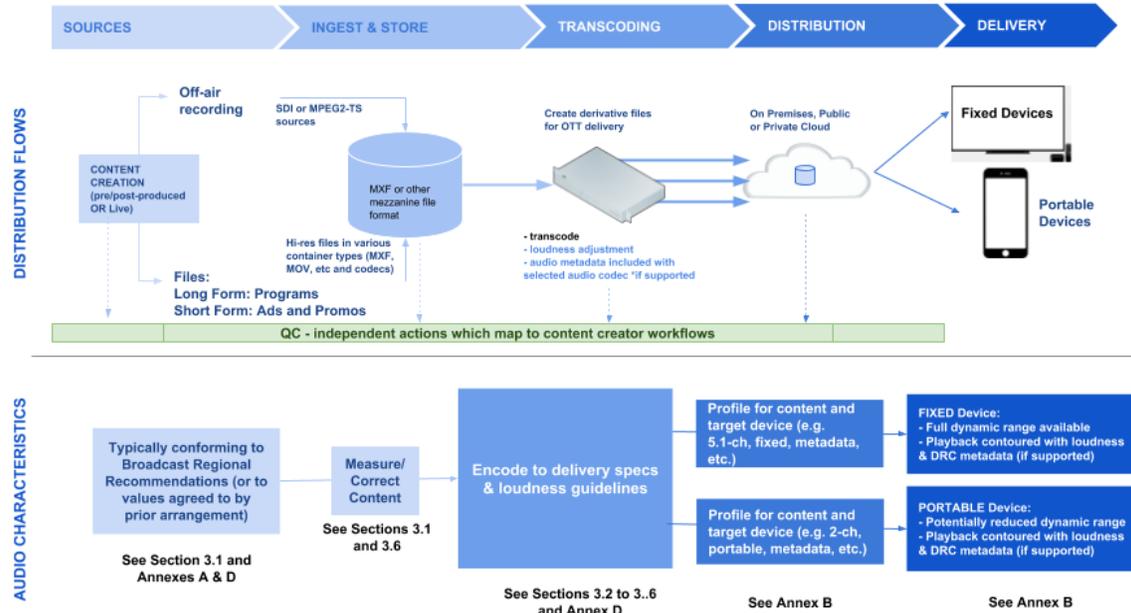


Figure E.1 – OTT/OVD Audio Workflow

LINEAR STREAMING AUDIO WORKFLOW

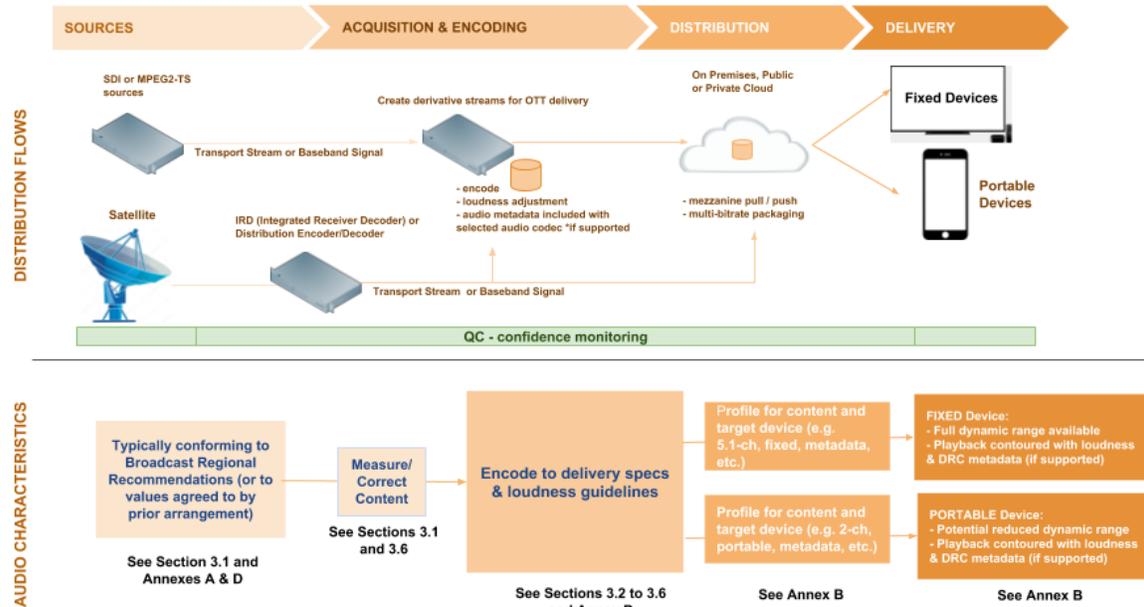


Figure E.2 – Linear Streaming Audio Workflow

Bibliography

FCC 13-99 (July 22, 2013) “Annual Assessment of the Status of Competition in the Market for the Delivery of Video Programming” US Federal Communications Commission, Washington DC
<https://www.fcc.gov/document/fcc-adopts-15th-report-video-competition>

N15071 (Feb 2015) “White Paper on MPEG-D Dynamic Range Control” ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Audio, International Organization for Standardization, Geneva, Switzerland
http://www.mpeg-audio.org/docs/w15071_MPEG-D_DRC.pdf

T. Lund, “Prevention of Hearing Loss from the Use of Personal Music Players”, AES 58th International Conference, Aalborg, Denmark, (June 28–30 2015), Audio Engineering Society, New York, NY, US
<http://www.aes.org/e-lib/browse.cfm?elib=17796>

IEC 60268-10 (1991) *Sound system equipment - Part 10: Peak programme level meters*, International Electrotechnical Commission, Geneva, Switzerland.

EN 50332-3 *Sound system equipment: headphones and earphones associated with personal music players – Maximum sound pressure measurement methodology – Part 3: Measurement method for sound dose management*, February 2017, CENELEC: European Committee for Electrical Standardization, Brussels, Belgium
https://www.cenelec.eu/dyn/www/f?p=104:110:542490721259701:::FSP_ORG_ID,FSP_PROJECT,FSP_LANG_ID:1257189,60972,25

EBU R68-1992 *Alignment level in digital audio production equipment and in digital audio recorders*, EBU Technical Centre, Geneva, Switzerland.

CTA-CEB11-C *NTSC/ATSC Loudness Matching*, Consumer Technology Association (CTA), January 2015, Washington DC, US, <https://www.cta.tech/Research-Standards/Standards-Listing.aspx>

EBU Tech Doc 3344 *Guidelines for Distribution and Reproduction in Accordance with EBU R 128*, July 2016, European Broadcast Union, EBU Technical Centre, Geneva, Switzerland,
<https://tech.ebu.ch/docs/tech/tech3344.pdf>

AES17-2015 *AES Standard Method for Digital Audio Engineering - Measurement of Digital Audio Equipment*, September 2015, Audio Engineering Society (AES), New York, NY, US,
<http://www.aes.org/standards/blog/2015/9/aes17-2015-measurement-digital-audio-150917>

EBU Tech Doc 3342 *Loudness Range: A Measure to Supplement EBU R 128 Loudness Normalization*, January 2016, European Broadcast Union, EBU Technical Centre, Geneva, Switzerland,
<https://tech.ebu.ch/docs/tech/tech3342.pdf>

ISO/IEC 23003-4:2015 *Information technology -- MPEG audio technologies -- Part 4: Dynamic Range Control*, International Organization for Standardization (ISO), Geneva, Switzerland,
<https://www.iso.org/standard/66482.html>