

**INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 N15071
February 2015, Geneva, Switzerland**

Title **White Paper on MPEG-D Dynamic Range Control**
Source **Communication Subgroup**

Contents

OVERVIEW OF MPEG-D PART 4 – DYNAMIC RANGE CONTROL	2
1 ABSTRACT	2
2 INTRODUCTION	2
3 FUNCTIONALITY	3
3.1 DYNAMIC RANGE CONTROL	3
3.2 LOUDNESS CONTROL	5
3.2.1 Loudness Normalization.....	5
3.3 PEAK AND CLIPPING CONTROL	6
3.3.1 Clipping Prevention Gains and Peak Limiting.....	8
3.4 FADING.....	8
3.5 DUCKING	9
3.6 DOWNMIX.....	9
4 SYSTEM CONFIGURATION AND CONTROL	10
4.1 PRODUCTION SIDE	10
4.2 PLAYBACK SIDE	11
4.2.1 Control of Downmix.....	12
4.2.2 Control of Loudness Normalization	12
4.2.3 Control of Dynamic Range Compression.....	12
5 SYSTEM INTEGRATION	14
5.1 PAYLOAD TRANSPORT	14
5.1.1 Transport within the Audio Payload.....	14
5.1.2 Transport using the ISO Base Media File Format.....	15
5.2 DRC PROCESSING OF AUDIO SIGNAL	16
5.2.1 Time Domain Processing	16
5.2.2 Sub-Band Domain Processing.....	17
5.2.3 Synchronization and Delay	17
6 MPEG-D DRC EXTENSIONS FOR MPEG-H 3D AUDIO	17
7 MPEG-D DRC AND LEGACY METADATA.....	17
8 FURTHER RESOURCES	17
9 SUMMARY	18
10 REFERENCES	18

Overview of MPEG-D Part 4 – Dynamic Range Control

1 Abstract

The MPEG audio standard for Dynamic Range Control (DRC) [1] defines a unified and flexible format to support comprehensive dynamic range and loudness control. It addresses a wide range of content delivery use cases including media streaming and broadcast applications. The associated metadata is attached to the audio content and can be applied during playback to enhance the user experience in scenarios such as ‘in the car’ or ‘late at night’. The integrated loudness control can be used, for instance, to meet regulatory requirements. An advanced peak and clipping control is fully integrated. This White Paper provides a brief overview of the main functionality, various configurations, and implementation aspects of MPEG-D DRC.

2 Introduction

People consume audio and multimedia content using various devices ranging from smart phones to home theaters. Mobile devices are used in acoustic environments that may severely impact the playback audio quality. Moreover, a multitude of different content is available and can be accessed instantaneously at random. Given this scenario, dynamic range and loudness control are important tools to enhance the perceived audio quality.

To be more specific, here are a few examples that illustrate some common problems that may result from the scenario described above:

- The user needs to adjust the volume because the loudness of different assets is not consistent.
- The intelligibility of movie dialog is adversely affected in soft parts due to a noisy listening environment.
- The level of loud parts of a movie is annoyingly high when soft parts are just loud enough; or soft parts are inaudible when loud parts are at a reasonable level.
- The dynamic range of an asset is too large for the employed playback device (e.g. low-quality loudspeakers) or for the desired playback level.
- The audio signal clips after a downmix.
- Existing dynamic range compression of the playback device is not aware of the content characteristics and may degrade the audio quality beyond expectation.

MPEG-D DRC supports solutions for these types of problems based on dynamic range and loudness control that additionally considers potential clipping. MPEG-D DRC defines a comprehensive metadata format – including time-varying gain values – and how it is applied. The metadata is typically generated by the content author and attached to the content. The audio content is delivered unmodified and the metadata can be applied

at the receiver if desired. The content author has full control of the whole process and can ensure that the DRC metadata produces a high-quality result in all scenarios. The standard supports virtually any dynamic range compressor, i.e. it is not limited to a specific compression algorithm.

3 Functionality

The MPEG-D DRC standard supports control of audio levels in the following main categories.

- Loudness control of an entire audio asset.
- Dynamic range control within an audio asset.
- Peak and clipping control.

Additionally, arbitrary fade-in and fade-out transitions can be applied. The main audio can be ‘ducked’ with arbitrary attenuation transitions for audio overlays such as video description or voice-overs.

MPEG-D DRC supports a flexible configuration that permits both, codec-agnostic application and audio format specific configurations to deliver the best audio quality.

In the following the functionality provided by MPEG-D DRC is described in more detail.

3.1 Dynamic Range Control

Traditionally, dynamic range control is associated with compression, i.e. reduction, of the dynamic range by means of a dynamic range compressor. This results in softer segments of an audio asset being louder and/or loud segments being softer. A simple dynamic range compressor generates time-varying gain values that are applied to the audio signal to achieve the desired compression effect.

Instead of applying the gain values immediately, MPEG-D DRC represents the gain values by metadata so that they can be applied at the receiver if desired. Hence the metadata supports virtually any compressor including multi-band compressors with flexible band configurations.

Compressor gains of multiple independent compressors can be included to achieve different compression effects where each compressor can be optimized for a specific playback scenario. The playback system can then request a dynamically compressed version of the audio asset based on a parametric description of the playback scenario. In response, MPEG-D DRC produces the best matching compressed version of the audio asset for that playback scenario based on the provided metadata.

Some or all of the following criteria can be used by the playback system to request a compressed audio version:

- Target loudness level
- DRC effect type
- Dynamic range
- DRC characteristic

It is common practice to use different compressor configurations for different playback target levels. While for low playback target levels only moderate or even no compression is desired, high target levels imply the need for a considerable amount of signal compression. It is possible to define a specific target level range for each set of gain sequences included in the MPEG-D DRC metadata to identify the most suitable one for a given target level defined by the playback system. Fig. 1 illustrates an example of three different sets of DRCs that are dedicated to a specific playback scenario (AVR, TV, mobile device) and which have corresponding target level ranges defined. MPEG-D DRC will select the best match based on the requested target loudness (i.e. set 2 in this example).

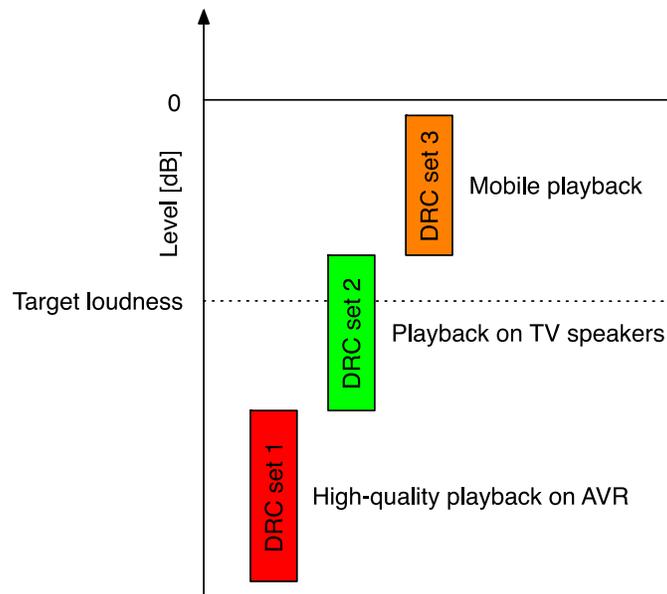


Fig. 1 Illustration of the target level range information included for different sets of DRC sequences.

The supported effect types listed in Table 1 serve a wide range of listening scenarios. The effect types that are provided with a specific asset can be accessed by the user or an application via a control interface.

Table 1 DRC effect types

Effect type	Description
Late night	For quiet environment, listening at low level, avoiding disturbing others while soft parts are still intelligible.
Noisy environment	Optimized to get the best experience in noisy environments, for instance by amplifying soft sections
Limited playback range	Reduced dynamic range to improve quality on playback devices with limited dynamic range such as the internal loudspeakers of mobile devices.
Low playback level	Listening at a low playback level

Dialog enhancement	The main effect is a more prominent dialog within the content
General compression	A general DRC effect that reduces the dynamic range and is applicable to multiple playback scenarios
Dynamic expansion	Enhancement of the dynamic range for instance by reverting the dynamic compression of the associated audio asset
Artistic effect	To create an artistic sound effect

3.2 Loudness Control

The loudness level of an entire audio asset can be adjusted to a desired target level by a constant gain using the loudness control. MPEG-D DRC can adjust the loudness based on metadata such that, for instance, subsequent assets are played back at a consistent loudness so that otherwise necessary manual volume adjustments by the user can be avoided or regulatory requirements can be met.

3.2.1 Loudness Normalization

Loudness normalization of content is based on a loudness measure of each asset item such as the Program Loudness [11] or the Dialnorm value [12]. After normalization, all assets have the same target loudness value. Optionally, loudness normalization can be based on loudness values generated by a user or expert/panel. While standards-based loudness measurement may be a regulatory requirement, subjective measurements can be used in other situations to avoid misadjustments due to occasional deviations between a loudness model and subjective results.

Fig. 2 illustrates an example for the loudness progression of three songs. The upper part shows the unmodified content. The second song is softer and the third song is louder than the first one. The middle part shows the result after loudness normalization. The overall loudness of each individual song is the same. The lower part illustrates the effect of both, loudness normalization and dynamic range compression. The three songs have the same loudness and similar dynamic range.

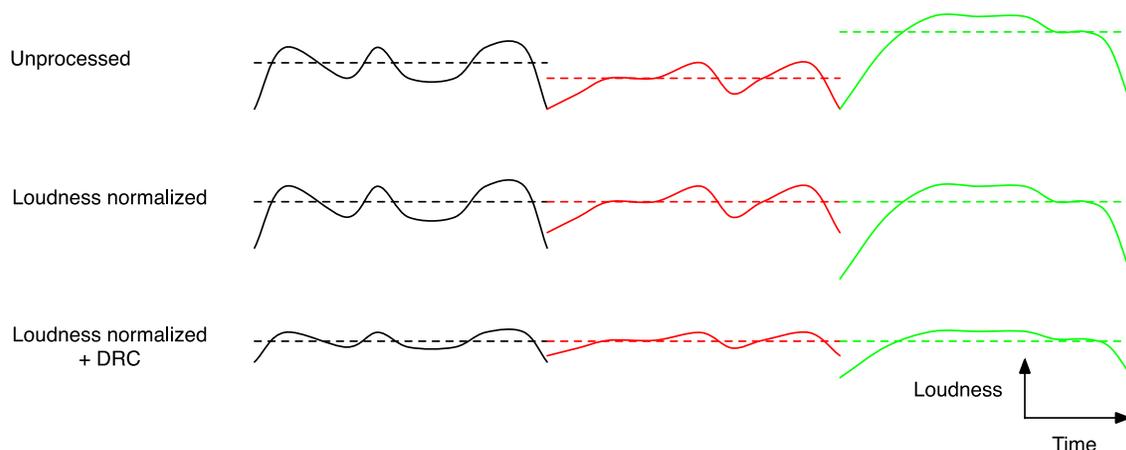


Fig. 2 Illustration of the time-varying loudness of three songs played back in succession. Program loudness of each song (dashed).

The loudness differences between the songs of an album are usually resulting from artistic intent. MPEG-D DRC will preserve the loudness differences between songs if the so-called album mode is active. Note that the loudness of the entire album can still be adjusted to match the preceding song or a desired target level. If using a custom playlist or shuffle playback, a song of the album can be normalized as usual on a per item basis. The result is illustrated in Fig. 3 where album mode was applied during loudness normalization in the lower graph.

Album mode is a control setting of MPEG-D DRC at the playback-side that should be used when songs are played back in album order.

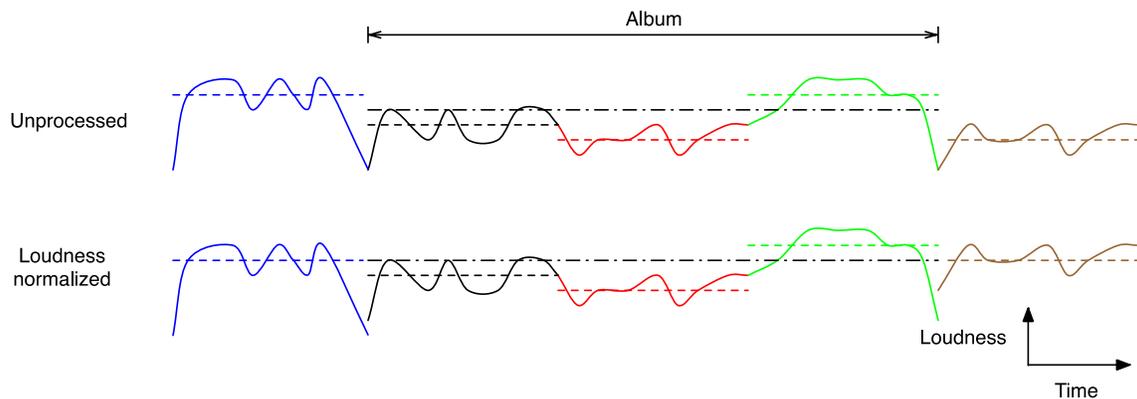


Fig. 3 Illustration of the time varying loudness of a progression of songs where three songs belong to an album. Program loudness of each song (dashed). Album loudness (dash-dotted).

Dynamic range processing or downmixing multichannel audio to a lower number of playback channels can change the loudness compared to the original content. An individual loudness metadata set can be provided for each dynamic range compression and/or downmix configuration given by MPEG-D DRC to enable accurate loudness control for each possible configuration.

3.3 Peak and Clipping Control

The largest peak magnitude of the audio signal is relevant for clipping prevention. Clipping can only be prevented if the gain applied to the audio signal does not cause the peak magnitude to exceed full scale. Hence, more aggressive compression and/or limiting must often be applied to achieve higher target loudness levels without clipping. MPEG-D DRC supports various tradeoffs between clipping, compression, and overall loudness based on the compressor metadata available with the audio asset and decoder-side controls. In addition to the peak magnitude (Sample Peak), True Peak [11] is supported.

Fig. 4 illustrates the dynamic range and peak value of three songs before and after loudness normalization. The third song is not fully normalized to avoid clipping. As a measure of dynamic range, the difference between the maximum short-term loudness [13] and the content loudness is used. Content loudness is the overall loudness of the asset based on one of the supported loudness measurements [1].

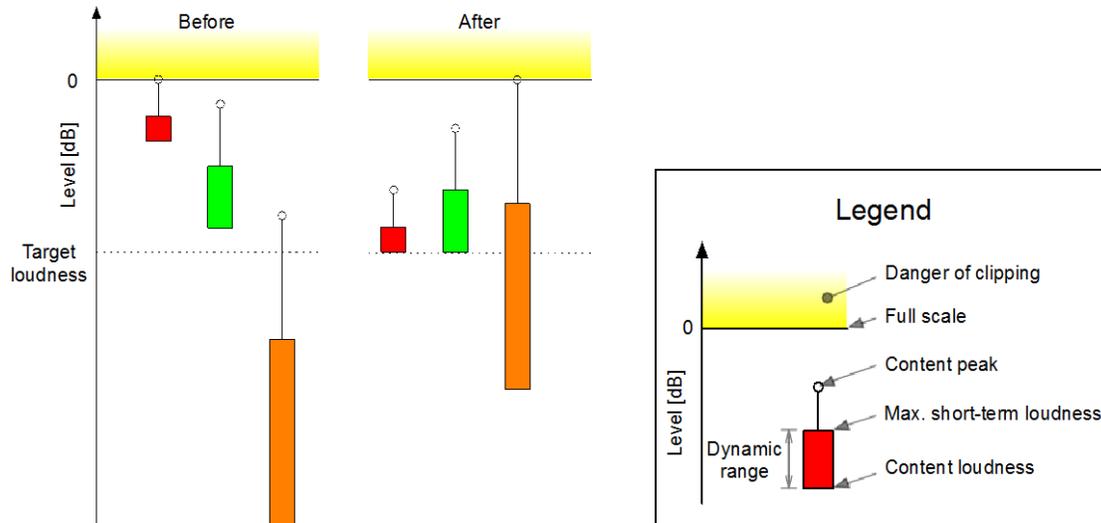


Fig. 4 Dynamic range and peak of three items before and after loudness normalization (left). Legend (right).

Fig. 5 illustrates the dynamic range and peak value of three songs before and after loudness normalization and dynamic range compression. Due to the compression, the third song is louder than indicated in Fig. 4 after normalization. If desired, MPEG-D DRC can limit the maximum normalization gain as illustrated in Fig. 5.

In the hypothetical case that the third asset additionally contains DRC gain metadata with stronger compression, MPEG-D DRC is able to recognize and apply the stronger compression to fully achieve loudness normalization.

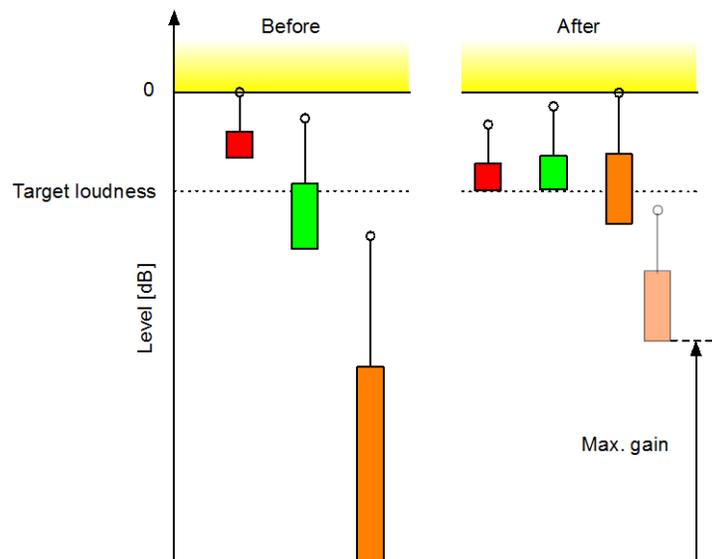


Fig. 5 Dynamic range and peak of three items before and after loudness normalization and dynamic range compression.

To achieve a very high target level it may be acceptable to allow some clipping. This is supported by specifying a maximum output peak level above full scale. As indicated in Fig. 6, most of the content is reproduced at a higher level compared to Fig. 5.

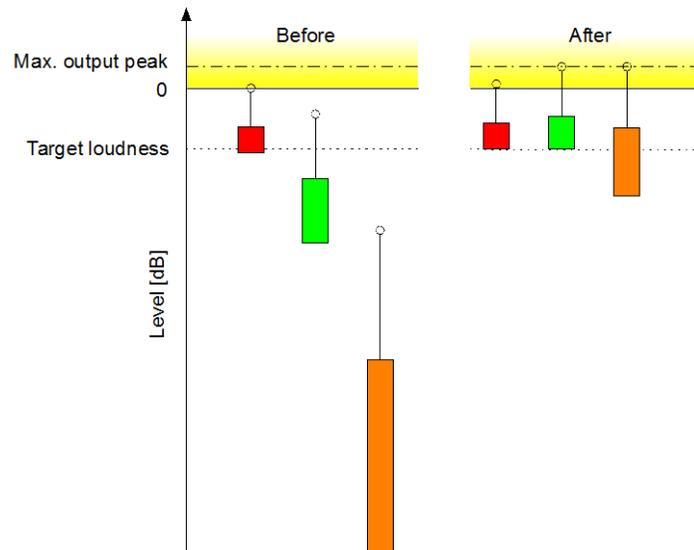


Fig. 6 Dynamic range and peak of three items before and after loudness normalization and dynamic range compression with some clipping.

3.3.1 Clipping Prevention Gains and Peak Limiting

Clipping prevention gains can be applied to the audio content in the same way as compressor gain values. Their purpose is the reduction of magnitude peaks to avoid clipping. The gain values are usually generated by a peak limiter at the encoder side. MPEG-D DRC automatically scales back and applies the clipping prevention gains according to the remaining headroom after adjusting the level to the target loudness if applicable. If such an adjustment is not required, the clipping prevention gains can be merged with the DRC gains to one combined gain sequence at the encoder side.

In addition to clipping prevention gains, MPEG-D DRC also specifies an optional peak limiter at the decoder side to account for potential clipping due to quantization errors, parametric coding tools, or arbitrary downmixing and rendering.

3.4 Fading

Gapless albums usually have no fade-out and fade-in between songs. If a song of a gapless album is played back in a different sequence than the one used in the album, there can be unintended hard transitions at the beginning and end of a song. Time-varying gains can be supplied to provide fade-in and fade-out transitions that are applied by MPEG-D DRC if the album mode is not active.

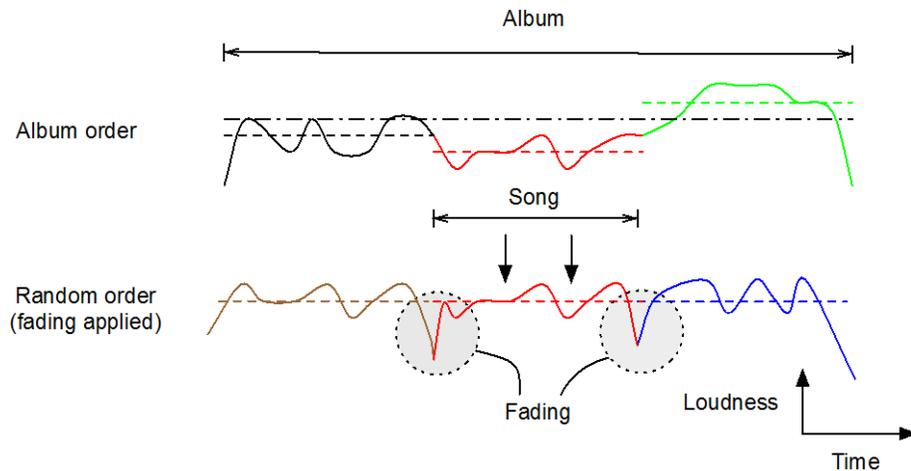


Fig. 7 Varying loudness of gapless album songs. Top: album order (no fading applied). Bottom: other order (fading applied).

3.5 Ducking

Ducking is useful if an audio signal is overlaid over the main audio signal, such as narration, director’s comments, video description, and other related use cases. In common approaches the main audio signal is attenuated when the narration is active. The attenuation and transitions can be precisely controlled and delivered by MPEG-D DRC. Ducking is activated when the corresponding overlaid signal channel is part of the selected audio mix. MPEG-D DRC eliminates the look-ahead needed in traditional playback systems.

Fig. 8 illustrates a narration channel signal and the corresponding ducking gain.

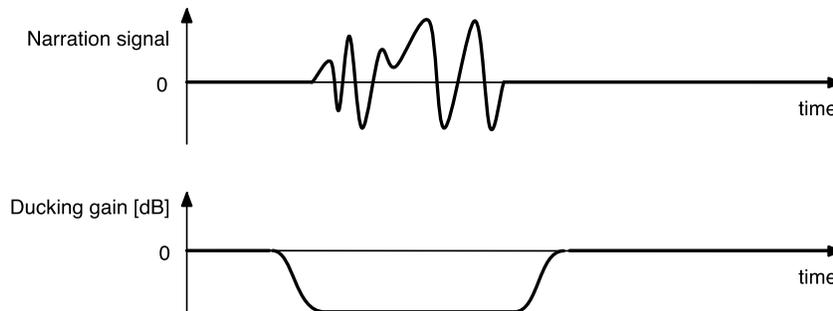


Fig. 8 Narration channel signal (top) and corresponding ducking gain (bottom).

3.6 Downmix

MPEG-D DRC fully supports downmixes in the same way as the base layout of the audio asset. This includes dynamic range compression, loudness control, peak and clipping control. More precisely, compression gains can be specified for application before and/or after the downmix. Fig. 9 shows an example system that applies five different gain

sequences before and after the downmix. Three independent DRC gain sequences are applied to the three channel groups before the downmix.

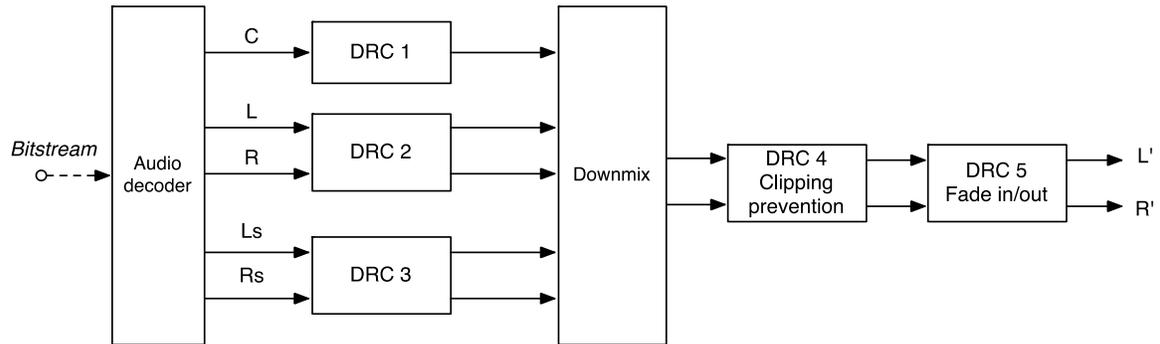


Fig. 9 Example configuration with DRC channel groups, downmix, clipping prevention, and fading.

4 System Configuration and Control

MPEG-D DRC can be configured in various ways to fit into a larger system. If an audio codec is used, the MPEG-D DRC metadata can be transported in an extension payload of the codec format or in a separate MPEG-4 track. The DRC tool will adapt to the frame size and sample rate of the audio codec to avoid unnecessary delay and complexity. Some basic configuration examples for the production and the playback side are discussed below.

4.1 Production Side

For offline production, the DRC, peak, and loudness metadata can be generated without any delay constraints. Measuring the program loudness of the whole audio asset in a first pass is recommended. During the second pass, the program loudness can be used to normalize the audio level before it enters the dynamic range compressor for the generation of DRC gains. The DRC gains are encoded and multiplexed into the output bitstream. By applying the DRC gains to the audio signal, the peak and loudness metadata for the dynamically compressed audio signal can be generated. The loudness and peak metadata should then be merged into the output bitstream.

If multiple dynamic range compressors are used to provide different DRC options at the receiver, they can be applied in parallel. Consequently, the peak and loudness measurement is then done independently for each dynamic range compressor. An example system is shown in Fig. 10.

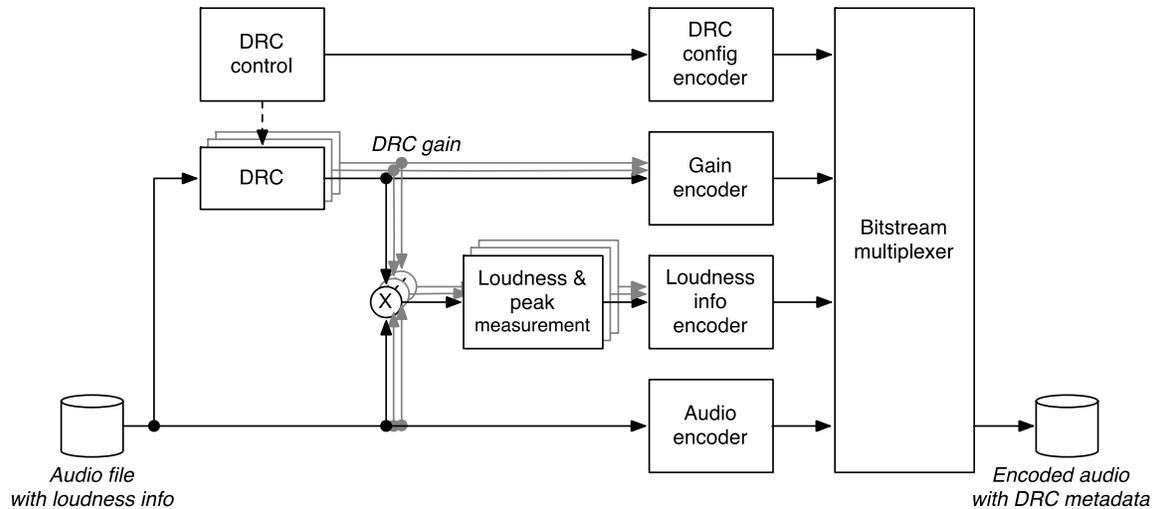


Fig. 10 Example for offline production.

A new DRC can be defined by parametric changes to the DRC gains of an existing DRC. This technique saves bitrate, because such a DRC does not require the transmission of any additional gain values. Typically, only the gain modification parameters, loudness information, and configuration are transmitted, but no dynamic payload.

In a live-streaming scenario, the overall loudness and peak value cannot be measured before the broadcast. But if the production system is carefully controlled and calibrated, the targeted loudness of the final broadcast can be well approximated. This approximation can be used to determine the loudness metadata for normalizing the DRC inputs. The loudness and peak information can be updated during the broadcast if necessary. This may occur for instance if an advertisement is inserted.

In a streaming scenario, the MPEG-D DRC configuration and loudness information must be repeatedly transmitted to permit a proper initialization of a decoder going online.

4.2 Playback Side

An example playback system with MPEG-D DRC is shown in Fig. 11. At the decoder side, MPEG-D DRC has a defined interface that exposes all relevant parameters to control the downmix, dynamic range compression, and loudness normalization. These controls are explained in more detail in the following sub-sections.

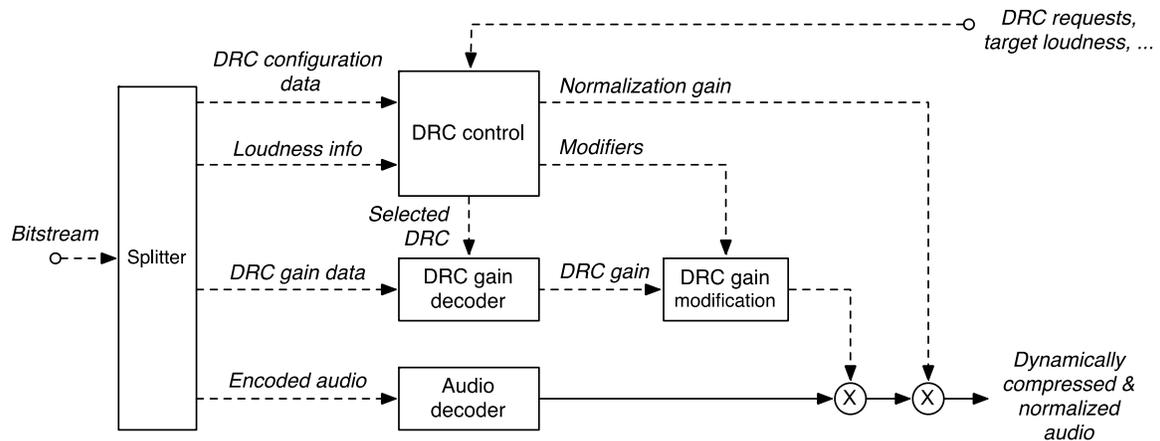


Fig. 11 Example playback system with MPEG-D DRC.

4.2.1 Control of Downmix

If applicable, a downmix is performed based on the requested downmix identifier, target channel layout, or the number of playback channels [1].

4.2.2 Control of Loudness Normalization

Loudness normalization can be activated and controlled by specifying the target loudness. Album Mode can be activated. A loudness value of a specific loudness measurement scheme can be requested for normalization, e.g. if a dialog-based loudness measurement is preferred over the loudness of the full program. Furthermore, the maximum normalization gain and the maximum output peak level can be specified.

4.2.3 Control of Dynamic Range Compression

The dynamic range compression can be activated and controlled by several parameters at the decoder side. As discussed in Section 3.1, a specific type of dynamic range compressor can be requested based on one or more items of

- Target loudness level
- DRC effect type (see Table 1),
- Dynamic range,
- DRC characteristic (see Fig. 12).

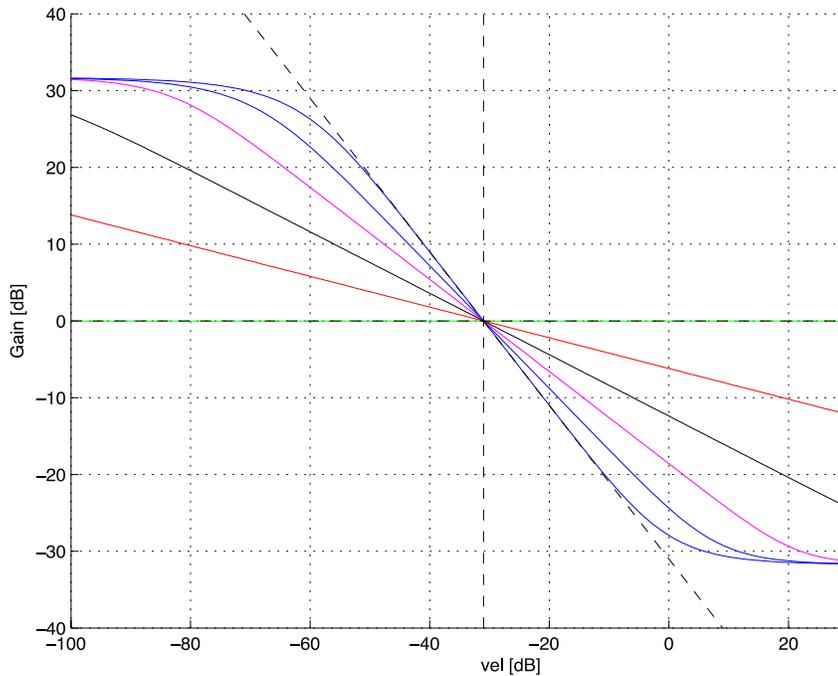


Fig. 12 Example DRC characteristics defined in MPEG-D DRC [6].

The DRC tool selects the best matching dynamic range compressor and the associated time-varying gains can optionally be modified before they are applied to the audio signal. This can be achieved by specifying independent scaling factors for positive and negative gain levels (compress and boost, see Fig. 13), or by specifying a target DRC characteristic. A target characteristic can only be achieved, if the DRC characteristic of the compressor used is known. Using a target characteristic permits more flexibility of gain modifications than with the compress and boost factors. Please note that modifications that result in a gain increase may result in a larger peak value than expected based on the metadata. In this case, the use of a peak limiter is recommended to avoid clipping.

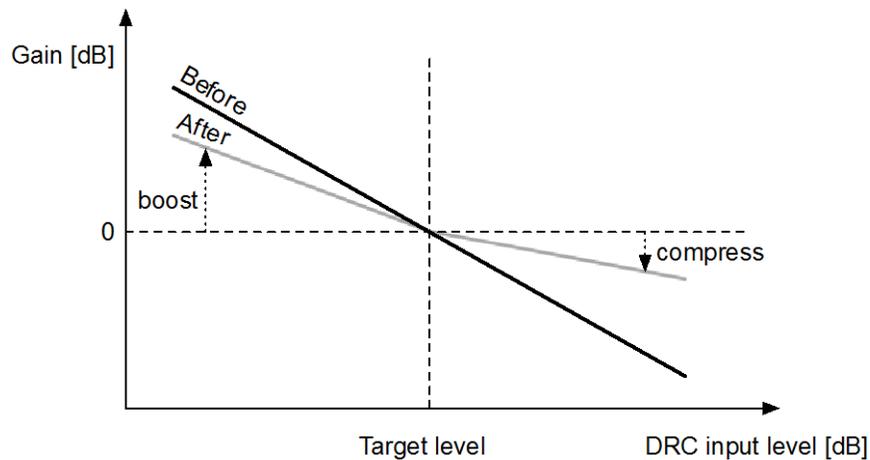


Fig. 13 Effect of boost and compress factor on the DRC characteristic (black line).

The dynamic range compressor can be changed in-process. If a new dynamic range compressor is requested during the processing of the audio signal, the new compressor will be applied immediately while avoiding any glitches in the output audio signal.

5 System Integration

MPEG-D DRC specifies different methods of metadata payload transport and how the metadata can be applied to the audio signal at the receiving side. The best choice often depends on the representation of the audio signal itself and whether an MPEG-4 Systems layer can be used. For instance, if the audio signal is encoded by an audio codec such as MPEG-AAC, the MPEG-D DRC metadata can either be included in the AAC bitstream or as a separate metadata track in the MPEG-4 file format. If the audio signal is encoded by MPEG-H 3D Audio, the MPEG-D DRC metadata is always embedded in the MPEG-H bitstream for optional coupling to other MPEG-H metadata types.

While in many scenarios the DRC gain is applied to the time-domain audio signal, the gain can alternatively be applied in the audio codec's sub-band domain, such as the QMF domain of an HE-AAC codec.

5.1 Payload Transport

Typically, an MPEG-D DRC encoder will generate a payload with variable bit rate for the dynamic part which contains the DRC gain sequences. The smallest bit rate for one gain sequence is approximately 0.4 kbit/s at a DRC frame size of 1024 if the extra bits needed for the codec's transport are ignored. Multiple DRCs with different compression strengths can be based on the same gain sequence to minimize the bit rate.

Two main categories of MPEG-D DRC payload transport are specified. These are the transport within the audio stream and the transport in a separate metadata track as defined in the ISO Base Media File Format [5].

5.1.1 Transport within the Audio Payload

Recent MPEG audio codecs specify extension payload fields that can carry the MPEG-D DRC payload in-stream. Legacy decoders that cannot decode MPEG-D DRC will ignore it.

5.1.1.1 Transport in AAC-based Formats

In order to carry the MPEG-D DRC payload, the Fill Element of the MPEG-4 AAC format [2] [3] is used as indicated in Fig. 14.

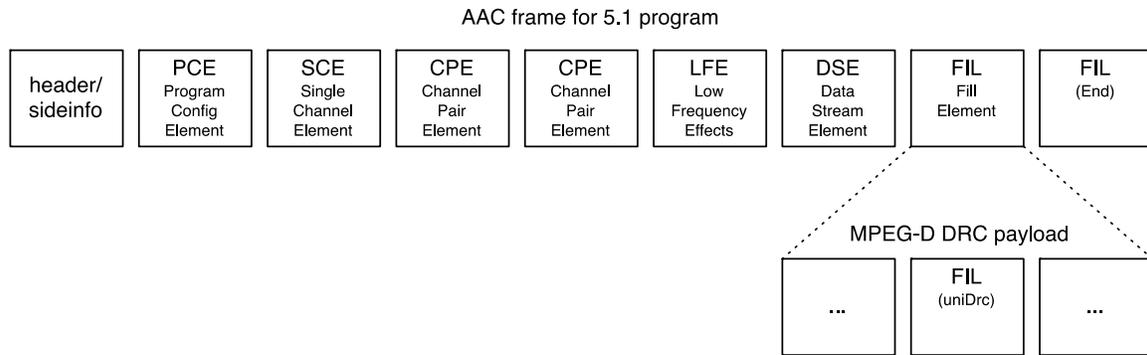


Fig. 14 Transport of MPEG-D DRC payload in the Fill Element of an AAC bitstream frame for a 5.1 program (Example).

Enhanced codecs that use AAC as the base layer can carry the MPEG-D DRC payload in the same location. Those codecs include HE-AAC [2], MPEG-D Surround [7], and MPEG-D SAOC [8].

5.1.1.2 Transport within USAC and MPEG-H 3D Audio

For USAC [9] and MPEG-H 3D Audio [10], the dynamic MPEG-D DRC payload can be carried in a general extension container to the audio codec. The static MPEG-D DRC configuration information is carried in a separate configuration container attached to the dynamic payload. The static loudness information can be transported in a separate container if desired.

5.1.2 Transport using the ISO Base Media File Format

The MPEG-D DRC payload transport in a separate metadata track of the ISO Base Media File Format (BMFF) [4][5] is codec agnostic and can support any audio format that can be carried in BMFF. An example of a valid ISO file box structure containing MPEG-D DRC metadata is shown in Fig. 15. The information is organized in separate boxes for the DRC track, the loudness information, and the DRC gains. Reference fields are used to link an audio track with the applicable DRC track and the DRC gains (see Fig. 15 left). Alternatively, the DRC gains can be carried in the audio frames if the audio codec supports it (see Fig. 15 right).

MPEG-D DRC supports multiple audio tracks, i.e. it can be applied across audio tracks. For instance, if a separate narration track is present, it can be overlaid on top of the main audio content using the ducking feature (see Section 3.5).

The location information of the DRC payload can be conveyed at the File Format level, even if the dynamic DRC payload is carried within one or more audio tracks and not in a separate metadata track.

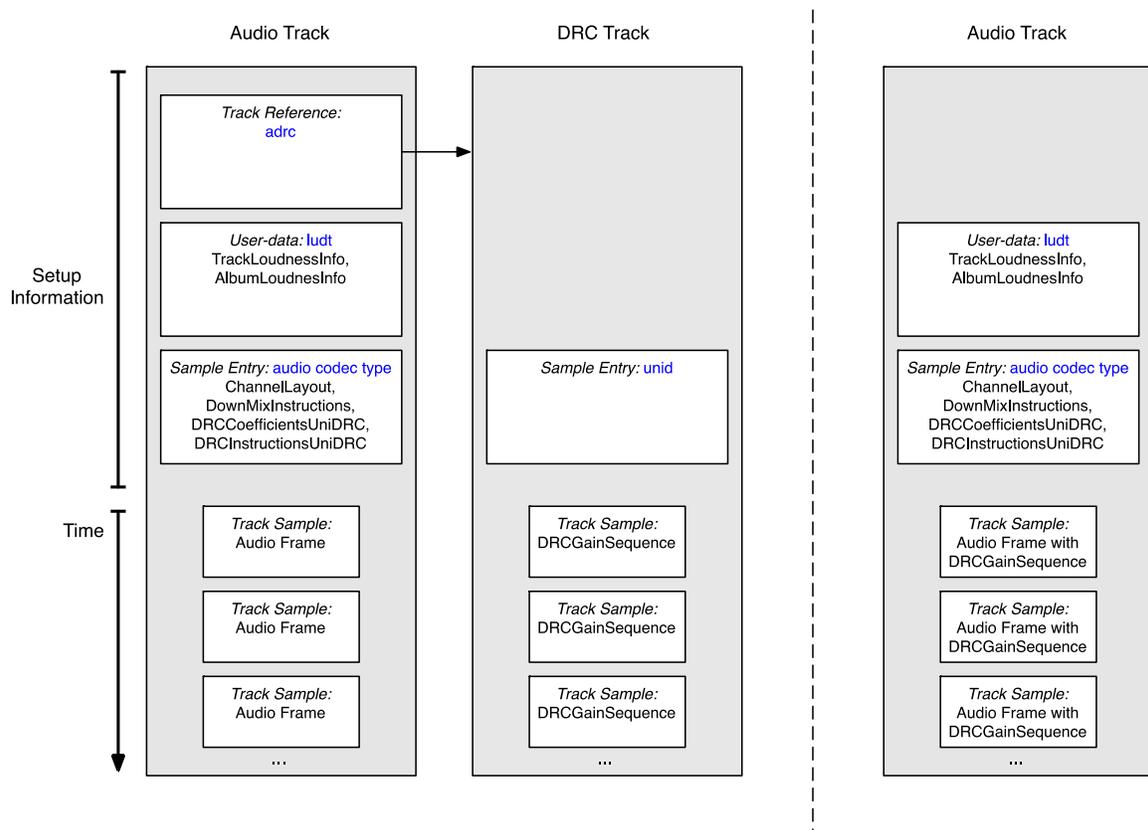


Fig. 15 Illustration of boxes associated with MPEG-D DRC in an ISO file. Left: audio track with referenced DRC track. Right: audio track including DRC gain sequences.

5.2 DRC Processing of Audio Signal

At the decoder side the desired DRC effects are achieved by applying time-varying gains to the audio signal. The DRC metadata controls whether the gains are applied in the time domain or sub-band domain. Furthermore, for multi-band DRC different gains can be applied to different bands in the time or sub-band domain.

If the configuration includes a downmix, the DRC gains can be applied before and/or after the downmix. At any location where DRC gains are applied they can be specified for individual channel groups.

5.2.1 Time Domain Processing

In case of time-domain processing, the DRC gains are applied to the final time-domain output of the audio decoder. The gains are applied directly to the audio signal for a full-band DRC. If a multi-band DRC is present, the audio signal is first split into up to four bands before the individual gains are applied to each band. All bands are summed to form the result. MPEG-D DRC defines a Linkwitz-Riley filterbank for time-domain multi-band DRC processing with adjustable crossover frequencies for this purpose.

5.2.2 Sub-Band Domain Processing

Sub-band processing takes advantage of the sub-band representation of the audio signal inside of an audio decoder. For instance, the QMF signal representation of an HE-AAC decoder is suitable for DRC gain application. The same holds for USAC and MPEG-H 3D Audio decoders, which also make use of the QMF signal representation. The MDCT domain such as in AAC-LC is not suitable for sub-band DRC gain processing due to potential aliasing distortions. For multi-band DRC, the decoder sub-bands are grouped into DRC bands for individual gain application. Therefore, no additional filterbank processing beyond the regular decoder operation is necessary.

5.2.3 Synchronization and Delay

DRC frames and audio frames are synchronized, i.e. the decoded DRC gains of the current DRC frame apply to the current audio frame in the domain specified. Usually the DRC frame size is identical to the audio codec frame size. Hence there is no additional audio delay caused by the DRC processing.

If a backwards-compatible audio decoder is used instead of a full-featured decoder, the sub-band domain signal may not be provided by the decoder. In that case, sub-band DRC gains must be applied in the time-domain with an appropriate delay.

6 MPEG-D DRC Extensions for MPEG-H 3D Audio

The MPEG-H 3D Audio codec [10] was developed to allow highly efficient immersive audio transmission and new capabilities like personalization and adaptation of the audio content to different use scenarios at the decoder side. To account for these new features, the static MPEG-D DRC configuration payload was extended by additional syntax elements within the MPEG-H specification that allow for coupling of individual DRC and loudness metadata sets to specific audio scenes if desired. Dependent on the chosen audio scene at the decoder side, optimized DRC and loudness metadata sets can be applied to improve the perceived audio quality in interactive MPEG-H scenarios. Apart from the described additions, the default feature set of MPEG-D DRC is still fully supported.

7 MPEG-D DRC and Legacy Metadata

The audio codecs defined by MPEG-AAC [2] and ATSC AC-3 [12] include metadata specifications for DRC and loudness as summarized in [14]. In contrast to MPEG-D DRC, the metadata format is codec specific and limited to a maximum of two different DRC gain sequences. Only one loudness value is supported assuming that the compressed and/or downmixed versions have the same loudness as the reference.

MPEG-D DRC can seamlessly integrate the legacy metadata and it supports metadata to describe the resulting DRC effect and loudness in a unified and backwards-compatible way.

8 Further Resources

MPEG reference software is currently (2015) being developed for MPEG-D DRC and MPEG-H 3D Audio including the enhanced MPEG-D DRC. Moreover, an extension of

the ISO Base Media File Format reference software to handle MPEG-D DRC metadata is underway. As with other MPEG standards, MPEG will specify how to test conformance of an implementation and supply test vectors to support adoption and to ensure compliance.

9 Summary

The MPEG-D Part 4 standard for Dynamic Range Control describes a universally applicable tool for dynamic range and loudness control. The associated format is flexible and codec agnostic. Content providers can use this format to assure that the dynamic range compression meets their quality targets for various playback scenarios and content types. Consumers can benefit from the interoperability of this format and from a better user experience due to improved playback quality in many scenarios and more consistent playback loudness of their content.

The bitrate of the MPEG-D DRC payload is typically very small in comparison to the audio bitrate. The transport of the MPEG-D DRC payload in various MPEG audio codec formats and the ISO Base Media File format is fully specified. In the framework of the MPEG-H 3D Audio codec, MPEG-D DRC was adopted as integral part of MPEG-H with some extensions to support the advanced features of the codec.

10 References

- [1] ISO/IEC 23003-4, “Information technology — MPEG audio technologies — Part 4: Dynamic range control,” 2015.
- [2] ISO/IEC 14496-3, “Information technology — Coding of audio-visual objects — Part 3: Audio,” 4th edition, 2009.
- [3] ISO/IEC 14496-3, “Information technology — Coding of audio-visual objects — Part 3: Audio. Amendment 5: Support for Dynamic Range Control”. (To be published in 2015)
- [4] ISO/IEC 14496-12, “Information Technology — Coding of audio-visual objects — Part 12: ISO Base Media File Format”, 4th edition, 2012.
- [5] ISO/IEC 14496-12 Amendment 4, “Improved Audio Support”. (To be published in 2015)
- [6] ISO/IEC 23001-8/AMD1, “Information technology — MPEG systems technologies — Part 8: Coding independent code points, Amendment 1: Additional audio code points”. (To be published in 2015)
- [7] ISO/IEC 23003-1, “Information technology — MPEG audio technologies — Part 1: MPEG Surround”, 2007.
- [8] ISO/IEC 23003-2, “Information technology — MPEG audio technologies — Part 2: Spatial Audio Object Coding (SAOC)”, 2010.
- [9] ISO/IEC 23003-3, “Information technology — MPEG audio technologies — Part 3: Unified speech and audio coding”, 2012.

- [10] ISO/IEC 23008-3, “Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 3: 3D Audio,” 2015
- [11] ITU-R, “Recommendation ITU-R BS. 1770-3. Algorithms to measure audio programme loudness and true-peak audio level,” 2012.
- [12] ATSC, “ATSC Standard: Digital Audio Compression (AC-3, E-AC-3),” Doc. A/52, 2012.
- [13] ITU-R, “Recommendation ITU-R BS.1771-1. Requirements for loudness and true-peak indicating meters,” 2012.
- [14] Fraunhofer-IIS, “White Paper: HE-AAC Metadata for Digital Broadcasting,” 2013.