

METHOD AND APPARATUS FOR PROCESSING OF AUDIO DATA**TECHNOLOGY**

The present disclosure relates generally to a method of metadata-based dynamic processing of
5 audio data for playback and, in particular, to determining and applying one or more processing
parameters to the audio data for dynamic loudness adjustment and/or dynamic range
compression. The present disclosure further relates to a method of encoding audio data and
metadata for dynamic loudness adjustment and/or dynamic range compression into a bitstream.
The present disclosure yet further relates to a respective decoder and encoder as well as to a
10 respective system and computer program products. The present disclosure further relates to a
method of processing audio data for playback, a decoder for processing audio data for playback,
and respective computer program products.

While some embodiments will be described herein with particular reference to that disclosure, it
will be appreciated that the present disclosure is not limited to such a field of use and is
15 applicable in broader contexts.

BACKGROUND

Any discussion of the background art throughout the disclosure should in no way be considered
as an admission that such art is widely known or forms part of common general knowledge in the
field.

20 In playing back audio content, loudness is the individual experience of sound pressure. In
cinematic or television content, the loudness of dialogue in a program has been found to be the
most crucial parameter determining the perception of program loudness by a listener.

To determine the average loudness of a program, either of the full program or dialogue only,
analysis of the entire program has to be performed. The average loudness is typically required
25 for loudness compliance (for example, the CALM act in the US), and is also used for aligning
dynamic range control (DRC) parameters. The dynamic range of a program is the difference
between its quietest and loudest sounds. The dynamic range of a program depends on its content,
for example, an action movie may have a different and wider dynamic range than a documentary,
and reflects a creator's intent. However, capabilities of devices to play back audio content in the
30 original dynamic range vary strongly. Besides loudness management, dynamic range control is
thus a further key factor in providing optimal listening experience.

To perform loudness management and dynamic range control, the entire audio program or an audio program segment has to be analyzed and the resulting loudness and DRC parameters can be delivered along with audio data or encoded audio data to be applied in a decoder or playback device.

- 5 When analysis of an entire audio program or an audio program segment prior to encoding is not available, for example in real-time (dynamic) encoding, loudness processing or levelling is used to ensure loudness compliance and, if applicable, potential dynamic range constraints depending on playback requirements. This approach delivers processed audio that is “optimized” for a single playback environment.
- 10 There is thus an existing need for metadata-based processes that deliver “original” unprocessed audio with accompanying metadata allowing the playback device to use the metadata to modify the audio dynamically depending on device constraints or user requirements.

SUMMARY

- 15 In accordance with a first aspect of the present disclosure there is provided a method of metadata-based dynamic processing of audio data for playback. The method may include receiving, by a decoder, a bitstream including audio data and metadata for dynamic loudness. The method may further include decoding, by the decoder, the audio data and the metadata to obtain decoded audio data and the metadata. The method may further include determining, by
- 20 the decoder, from the metadata, one or more processing parameters for dynamic loudness adjustment based on a playback condition. The method may further include applying the determined one or more processing parameters to the decoded audio data to obtain processed audio data. And the method may include outputting the processed audio data for playback.

- The metadata for dynamic loudness adjustment may include a plurality of sets of metadata, with
- 25 each set corresponding to a respective (e.g., different) playback condition. Then, determining the one or more processing parameters for dynamic loudness adjustment from the metadata based on a (specific) playback condition may include selecting, in response to playback condition information provided to the decoder, a set of metadata corresponding to the (specific) playback condition, and extracting the one or more processing parameters for dynamic loudness
- 30 adjustment from the selected set of metadata. Therein, the playback condition information may be indicative of the (specific) playback condition or information derived therefrom.

In some embodiments, the metadata may be indicative of processing parameters for dynamic loudness adjustment for a plurality of playback conditions.

In some embodiments, said determining the one or more processing parameters may further include determining one or more processing parameters for dynamic range compression, DRC, based on the playback condition.

In some embodiments, the playback condition information may be indicative of a specific loudspeaker setup. In general, the playback condition may include one or more of a device type of the decoder, characteristics of a playback device, characteristics of a loudspeaker, a loudspeaker setup, characteristics of background noise, characteristics of ambient noise and characteristics of the acoustic environment.

In some embodiments, the selected set of metadata may include a set of DRC sequences, DRCSet. Further, each of the sets of metadata may include a respective set of DRC sequences, DRCSet. In general, said determining the one or more processing parameters may be said to further include selecting, by the decoder, at least one of a set of DRC sequences, DRCSet, a set of equalizer parameters, EQSet, and a downmix, corresponding to the playback condition.

In some embodiments, said determining the one or more processing parameters may further include identifying a metadata identifier indicative of the at least one selected DRCSet, EQSet and downmix to determine the one or more processing parameters from the metadata. Specifically, selecting the set of metadata may include identifying a set of metadata corresponding to a specific downmix. The specific downmix may be determined based on the loudspeaker setup.

In some embodiments, the metadata may include one or more processing parameters relating to average loudness values and optionally one or more processing parameters relating to dynamic range compression characteristics. Specifically, each set of metadata may include such one or more processing parameters relating to average loudness values and optionally one or more processing parameters relating to dynamic range compression characteristics.

In some embodiments, the bitstream may further include additional metadata for static loudness adjustment to be applied to the decoded audio data.

In some embodiments, the bitstream may be an MPEG-D DRC bitstream and the presence of metadata may be signaled based on MPEG-D DRC bitstream syntax.

In some embodiments, a loudnessInfoSetExtension()-element may be used to carry the metadata as a payload.

In some embodiments, the metadata may comprise one or more metadata payloads, wherein each metadata payload may include a plurality of sets of parameters and identifiers, with each set including at least one of a DRCSet identifier, drcSetId, an EQSet identifier, eqSetId, and a downmix identifier, downmixId, in combination with one or more processing parameters relating to the identifiers in the set.

In some embodiments, said determining the one or more processing parameters may involve selecting a set among the plurality of sets in the payload based on the at least one DRCSet, EQSet, and downmix selected by the decoder, wherein the one or more processing parameters determined by the decoder may be the one or more processing parameters relating to the identifiers in the selected set.

In accordance with a second aspect of the present disclosure there is provided a decoder for metadata-based dynamic processing of audio data for playback. The decoder may comprise one or more processors and non-transitory memory configured to perform a method including receiving, by the decoder, a bitstream including audio data and metadata for dynamic loudness adjustment; decoding, by the decoder, the audio data and the metadata to obtain decoded audio data and the metadata; determining, by the decoder, from the metadata, one or more processing parameters for dynamic loudness adjustment based on a playback condition; applying the determined one or more processing parameters to the decoded audio data to obtain processed audio data; and outputting the processed audio data for playback.

The metadata for dynamic loudness adjustment may include a plurality of sets of metadata, with each set corresponding to a respective (e.g., different) playback condition. Then, determining the one or more processing parameters for dynamic loudness adjustment from the metadata based on a (specific) playback condition may include selecting, in response to playback condition information provided to the decoder, a set of metadata corresponding to the (specific) playback condition, and extracting the one or more processing parameters for dynamic loudness adjustment from the selected set of metadata. Therein, the playback condition information may be indicative of the (specific) playback condition or information derived therefrom.

In accordance with a third aspect of the present disclosure there is provided a method of encoding audio data and metadata for dynamic loudness adjustment, into a bitstream. The method may include inputting original audio data into a loudness leveler for loudness processing to obtain, as an output from the loudness leveler, loudness processed audio data. The method may further include generating the metadata for dynamic loudness adjustment based on the

loudness processed audio data and the original audio data. And the method may include encoding the original audio data and the metadata into the bitstream.

In some embodiments, the metadata may include a plurality of sets of metadata. Each set of metadata may correspond to a respective (e.g., different) playback condition.

- 5 In some embodiments, the method may further include generating additional metadata for static loudness adjustment to be used by a decoder.

In some embodiments, said generating metadata may include comparison of the loudness processed audio data to the original audio data, wherein the metadata may be generated based on a result of said comparison.

- 10 In some embodiments, said generating metadata may further include measuring the loudness over one or more pre-defined time periods, wherein the metadata may be generated further based on the measured loudness.

In some embodiments, the measuring may comprise measuring overall loudness of the audio data.

- 15 In some embodiments, the measuring may comprise measuring loudness of dialogue in the audio data.

In some embodiments, the bitstream may be an MPEG-D DRC bitstream and the presence of the metadata may be signaled based on MPEG-D DRC bitstream syntax.

- 20 In some embodiments, a loudnessInfoSetExtension()-element may be used to carry the metadata as a payload.

- In some embodiments, the metadata may comprise one or more metadata payloads, wherein each metadata payload may include a plurality of sets of parameters and identifiers, with each set including at least one of a DRCSet identifier, drcSetId, an EQSet identifier, eqSetId, and a downmix identifier, downmixId, in combination with one or more processing parameters relating to the identifiers in the set, and wherein the one or more processing parameters may be parameters for dynamic loudness adjustment by a decoder.
- 25

In some embodiments, the at least one of the drcSetId, the eqSetId, and the downmixId may be related to at least one of a set of DRC sequences, DRCSet, a set of equalizer parameters, EQSet, and downmix, to be selected by the decoder.

In accordance with a fourth aspect of the present disclosure there is provided an encoder for encoding in a bitstream original audio data and metadata for dynamic loudness adjustment. The encoder may comprise one or more processors and non-transitory memory configured to perform a method including inputting original audio data into a loudness leveler for loudness processing
5 to obtain, as an output from the loudness leveler, loudness processed audio data; generating the metadata for dynamic loudness adjustment based on the loudness processed audio data and the original audio data; and encoding the original audio data and the metadata into the bitstream.

In accordance with a fifth aspect of the present disclosure there is provided a system of an encoder for encoding in a bitstream original audio data and metadata for dynamic loudness
10 adjustment, and a decoder for metadata-based dynamic processing of audio data for playback.

In accordance with a sixth aspect of the present disclosure there is provided a computer program product comprising a computer-readable storage medium with instructions adapted to cause the device to carry out a method of metadata-based dynamic processing of audio data for playback or a method of encoding audio data and metadata for dynamic loudness adjustment, into a
15 bitstream when executed by a device having processing capability.

In accordance with a seventh aspect of the present disclosure there is provided a computer-readable storage medium storing the computer program product described herein.

In accordance with an eighth aspect of the present disclosure there is provided a method of processing audio data for playback. The method may include receiving, by a decoder, a bitstream
20 including encoded audio data and metadata, wherein the metadata includes one or more dynamic range control (DRC) sets, and for each DRC set, an indication of whether the DRC set is configured for providing a dynamic loudness compensation effect. The method may further include parsing the metadata, by the decoder, to identify DRC sets that are configured for providing the dynamic loudness compensation effect. The method may further include decoding,
25 by the decoder, the encoded audio data to obtain decoded audio data. The method may further include selecting, by the decoder, one of the identified DRC sets configured for providing the dynamic loudness compensation effect. The method may further include applying to the decoded audio data, by the decoder, the one or more DRC gains corresponding to the selected DRC set to obtain dynamic loudness compensated audio data. And the method may include outputting the
30 dynamic loudness compensated audio data for playback.

In some embodiments, the metadata may include a plurality of DRC sets configured for providing the dynamic loudness adjustment, wherein each of the plurality of DRC sets may also

be associated with one or more playback conditions, and wherein the selecting may be performed in response to an indication of a playback condition provided to the decoder.

In some embodiments, in addition to providing a dynamic loudness compensation effect, the one or more DRC sets may also be configured to provide dynamic range control.

- 5 In some embodiments, the playback condition may include one or more of a device type of the decoder, characteristics of a playback device, characteristics of a loudspeaker, a loudspeaker setup, characteristics of background noise, characteristics of ambient noise and characteristics of the acoustic environment.

- 10 In some embodiments, the indication of whether the DRC set is configured for providing the dynamic loudness compensation effect may be provided in a parameter indicating one or more effects provided by the DRC set.

- In some embodiments, the parameter indicating one or more effects provided by the DRC set may be a drcSetEffect bitfield of an MPEG-D DRC bitstream, wherein individual bits of the drcSetEffect bitfield correspond to different effects, and one of the bits of the drcSetEffect
15 bitfield corresponds to the dynamic loudness compensation effect.

In some embodiments, the indication of whether the DRC set is configured for providing the dynamic loudness compensation effect may be whether the DRC set is specified in a dynamic loudness compensation bitstream payload.

- 20 In some embodiments, the dynamic loudness compensation bitstream payload may be included in an extension field of a previously defined bitstream syntax.

In some embodiments, the extension field may be a uniDrcConfigExtension field of an MPEG-D DRC bitstream, and the dynamic loudness compensation bitstream payload may be included only for specific values of a uniDrcConfigExtType parameter.

- 25 In some embodiments, a plurality of dynamic loudness compensation payloads specifying a plurality of DRC sets configured for providing the dynamic loudness compensation effect may be included in the extension field of the previously defined bitstream syntax.

In some embodiments, the indication of whether the DRC set is configured for providing the dynamic loudness compensation effect may be a field of a previously existing configuration element of a previously defined bitstream syntax.

- 30 In some embodiments, the field may be a dynamicLoudCompDRCSet parameter, and the previously existing configuration element may be a downmixInstructions element, a

drcInstructionsBasic element, or a drcInstructionsUniDRC element of an MPEG-D DRC bitstream.

In some embodiments, the field may be a previously existing field reserved for future use.

- 5 In some embodiments, the indication of whether the DRC set is configured for providing the dynamic loudness compensation effect may be a field of an updated version of a previously existing configuration element of a previously defined bitstream syntax.

In some embodiments, the field may be a dynamicLoudCompDRCSet parameter, and the updated version of the previously existing configuration element may be a downmixInstructionsV2 element or a drcInstructionsUniDrcV2 element.

- 10 In some embodiments, an indication that a dynamic loudness compensation effect is desired may be provided to the decoder through an interface, and the DRC set may be selected in response to the indication provided to the decoder through the interface.

- 15 In some embodiments, indications of additional desired effects may be provided to the decoder through the interface, the metadata may include a plurality of DRC sets configured to provide the dynamic loudness compensation effect, and the selection may depend on the additional desired effects.

In some embodiments, the indication that a dynamic loudness compensation effect is desired may be provided through a drcEffectTypeRequest parameter of a dynamicRangeControllerInterface payload.

- 20 In some embodiments, the metadata may include one or more static loudness values configured for providing static loudness adjustment to the decoded audio data.

In some embodiments, static loudness adjustment may be applied, in response to one or more of the static loudness values, to the decoded audio data or the dynamic loudness compensated audio data.

- 25 In some embodiments, a first one of the DRC sets may be configured for providing dynamic range control, and the first DRC set may comprise an indication that the selected DRC set configured for providing the dynamic loudness compensation effect may be configured for application in combination with the first DRC set.

- 30 In some embodiments, the selected DRC set may comprise an indication of whether the one or more DRC gains corresponding to the selected DRC set may only be applied in combination with DRC gains corresponding to the first DRC set.

In some embodiments, DRC gains corresponding to the first DRC set may be extracted from the bitstream and may be applied to the decoded audio data.

In accordance with a ninth aspect of the present disclosure there is provided a decoder for processing audio data for playback. The decoder may comprise one or more processors and non-
5 transitory memory configured to perform a method of processing audio data for playback as described above.

In accordance with a tenth aspect of the present disclosure there is provided a computer program product comprising a computer-readable storage medium with instructions adapted to cause the device to carry out a method of processing audio data for playback as described above.

10 In accordance with an eleventh aspect of the present disclosure there is provided a computer-readable storage medium storing the computer program product described herein.

BRIEF DESCRIPTION OF THE DRAWINGS

Example embodiments of the disclosure will now be described, by way of example only, with
15 reference to the accompanying drawings in which:

FIG. 1 illustrates an example of a decoder for metadata-based dynamic processing of audio data for playback.

FIG. 2 illustrates an example of a method of metadata-based dynamic processing of audio data for playback.

20 FIG. 3 illustrates an example of an encoder for encoding in a bitstream original audio data and metadata for dynamic loudness adjustment.

FIG. 4 illustrates an example of a method of encoding audio data and metadata for dynamic loudness adjustment, into a bitstream.

FIG. 5 illustrates an example of a device comprising one or more processors and non-transitory
25 memory configured to perform the methods described herein.

FIG. 6 illustrates an example of a method of processing audio data for playback.

DESCRIPTION OF EXAMPLE EMBODIMENTS

Overview

The average loudness of a program or dialogue is the main parameter or value used for loudness compliance of broadcast or streaming programs. The average loudness is typically set to -24 or -23 LKFS. With audio codecs that support loudness metadata, this single loudness value representing the loudness of the entire program, is carried in the bitstream. Using this value in the decoding process allows gain adjustments that result in predictable playback levels, so that programs play back at known consistent levels. Therefore, it is important that this loudness value is set properly and accurately. Since the average loudness is dependent on measuring the entire program prior to encoding, for real-time situations such as dynamic encoding with unknown loudness and dynamic range variation, this is, however, not possible.

When it is not possible to measure the loudness of the entire file prior to encoding, a dynamic loudness leveler is often used to modify or contour the audio data prior to encoding so that it meets the required loudness. This type of loudness management is often seen as an inferior method to meet compliance, as it often changes the dynamic range intercorrelation in the audio content and may thus change the creative intent. This is especially the case when it is desired to distribute one audio asset for all playback devices, which is one of the benefits of metadata driven codec and delivering systems.

In some approaches, audio content is mixed with the required target loudness, and the corresponding loudness metadata is set to that value. A loudness leveler might still be used in those situations as it will be used to help steer the audio content to the target loudness but it will not be that „active“ and is only used to when the audio content starts to deviate from the required target loudness.

In view of the above, methods and apparatus described herein aim at making real-time processing situations, also denoted as dynamic processing situations, also metadata driven. The metadata allow for dynamic loudness adjustment and dynamic range compression in real-time situations. The methods and apparatus as described advantageously enable:

- Use of real-time loudness adjustment and DRC in MPEG-D DRC syntax;
- Use of real-time loudness adjustment and DRC in combination with downmixId;
- Use of real-time loudness adjustment and DRC in combination with drcSetId;
- Use of real-time loudness adjustment and DRC in combination with eqSetId.

That is, depending on decoder settings (e.g., DRCSet, EQSet, and downmix), a decoder can search based on the syntax a given payload for an appropriate set of parameters and identifiers, by matching the aforementioned settings to the identifiers. The parameters included in the set whose identifiers best match the settings can then be selected as the processing parameters for dynamic
5 loudness adjustment to be applied to received original audio data for correction.

Further, multiple sets of parameters for dynamic processing (multiple instances of dynLoudCompValue) can be transmitted.

The metadata-driven dynamic loudness compensation, in addition to correcting the overall loudness, can also be used to “center” the DRC gain calculation and application. This centering
10 may be a result of correcting the loudness of the content, via the dynamic loudness compensation, and how DRC is typically calculated and applied. In this sense, metadata for dynamic loudness compensation can be said to be used for aligning DRC parameters.

Metadata-based dynamic processing of audio data

Referring to the example of Figure 1, a decoder 100 for metadata-based dynamic processing of
15 audio data for playback is described. The decoder 100 may comprise one or more processors and non-transitory memory configured to perform a method including the processes as illustrated in the example of Figure 2 by means of steps S101 to S105.

The decoder 100, may receive a bitstream including audio data and metadata and may be able to output the unprocessed (original) audio data, the processed audio data after application of
20 dynamic processing parameters determined from the metadata and/or the metadata itself depending on requirements.

Referring to the example of Figure 2, in step S101 the decoder 100 may receive a bitstream including audio data and metadata for dynamic loudness adjustment and optionally, dynamic range compression (DRC). The audio data may be encoded audio data, the audio data may
25 further be unprocessed. That is, the audio data may be said to be original audio data. The metadata may include a plurality of sets of metadata. For example, each payload of metadata may include such plurality of sets of metadata. These different sets of metadata may relate to respective playback conditions (e.g., to different playback conditions).

While the format of the bitstream is not limited, in an embodiment, the bitstream may be an
30 MPEG-D DRC bitstream. The presence of metadata for dynamic processing of audio data may then be signaled based on MPEG-D DRC bitstream syntax. In an embodiment, a

loudnessInfoSetExtension()-element may be used to carry the metadata as a payload as detailed further below.

In step S102, the audio data and the metadata may then be decoded, by the decoder, to obtain decoded audio data and the metadata. In an embodiment, the metadata may include one or more processing parameters relating to average loudness values and optionally one or more processing parameters relating to dynamic range compression characteristics. It is understood that each set of metadata may include respective processing parameters. The metadata allows to apply dynamic or real-time correction. For example, when encoding and decoding for live real-time playout, the application of the “real-time” or dynamic loudness metadata is desired to ensure that the live playout audio is properly loudness managed.

In step S103, the decoder then determines, from the metadata, one or more processing parameters for dynamic loudness adjustment based on a playback condition. This may be done by using the playback condition or information derived from the playback condition (e.g., playback condition information), to identify an appropriate set of metadata among the plurality of sets of metadata.

In an embodiment, a playback condition may include one or more of a device type of the decoder, characteristics of a playback device, characteristics of a loudspeaker, a loudspeaker setup, characteristics of background noise, characteristics of ambient noise and characteristics of the acoustic environment. Preferably, the playback condition information may be indicative of a specific loudspeaker setup. The consideration of a playback condition allows the decoder for a targeted selection of processing parameters for dynamic loudness adjustment with regard to device and environmental constraints.

In an embodiment, the process of determining the one or more processing parameters in step S103 may further include selecting, by the decoder, at least one of a set of DRC sequences, DRCSet, set of equalizer parameters, EQSet, and a downmix, corresponding to the playback condition. Thus, the at least one of a DRCSet, EQSet and downmix correlates with or is indicative of the individual device and environmental constraints due to the playback condition.

Preferably, step S103 involves selecting a set of DRC sequences, DRCSet. In other words, the selected set of metadata may include such set of DRC sequences.

In an embodiment, the process of determining in step S103 may further include identifying a metadata identifier indicative of the at least one selected DRCSet, EQSet and DownmixSet to determine the one or more processing parameters from the metadata. The metadata identifier

thus enables to connect the metadata with a corresponding selected DRCSet, EQSet, and/or downmix, and thus with a respective playback condition.

In an embodiment, the specific loudspeaker setup may be used to determine a downmix, which in turn may be used for identifying and selecting an appropriate one among the plurality of sets of metadata. In such case, the specific loudspeaker setup and/or the downmix may be indicated by the aforementioned playback condition information.

In an embodiment, the metadata may comprise one or more metadata payloads (e.g., dynLoudComp() payloads, such as shown in **Table 5** below), wherein each metadata payload may include a plurality of sets of parameters (e.g., parameters dynLoudCompValue) and identifiers, with each set including at least one of a DRCSet identifier, drcSetId, an EQSet identifier, eqSetId, and a downmix identifier, downmixId, in combination with one or more processing parameters relating to the identifiers in the set. That is, each payload may comprise an array of entries, each entry including processing parameters and identifiers (e.g., drcSetId, eqSetId, downmixId). The array of entries may correspond to the plurality of sets of metadata mentioned above. Preferably, each entry comprises the downmix identifier.

In a further embodiment, the determining in step S103 may thus involve selecting a set among the plurality of sets in the payload based on the downmix selected by the decoder (or alternatively, based on the at least one DRCSet, EQSet, and downmix), wherein the one or more processing parameters determined in step S103 may be the one or more processing parameters relating to the identifiers in the selected set. That is, depending on settings (e.g., DRCSet, EQSet, and downmix) present in the decoder, the decoder can search a given payload for an appropriate set of parameters and identifiers, by matching the aforementioned settings to the identifiers. The parameters included in the set whose identifiers best match the settings can then be selected as the processing parameters for dynamic loudness adjustment.

In step S104, the determined one or more processing parameters may then be applied, by the decoder, to the decoded audio data to obtain processed audio data. The processed audio data, for example live real-time audio data, are thus properly loudness managed.

In step S105, the processed audio data may then be output for playback.

In an embodiment, the bitstream may further include additional metadata for static loudness adjustment to be applied to the decoded audio data. Static loudness adjustment refers in contrast to dynamic processing for real-time situations to processing performed for general loudness normalization.

Carrying the metadata for dynamic processing separately from the additional metadata for general loudness normalization allows to not have the “real-time” correction applied.

For example, when encoding and decoding for live real-time playout, the application of dynamic processing is desired to ensure that the live playout audio is properly loudness managed. But for
5 a non-real-time playout, or a transcoding where the dynamic correction is not desired or required, the dynamic processing parameters determined from the metadata do not have to be applied.

By further keeping the (dynamic/real-time) metadata for dynamic processing separate from the additional metadata, the originally unprocessed content can be retained, if desired. The original
10 audio is encoded along with the metadata. This allows the playback device to selectively apply the dynamic processing and to further enable playback of original audio content on high-end devices capable of playing back original audio.

Keeping the dynamic loudness metadata distinct from the long-term loudness measurement/information, such as contentLoudness (in ISO/IEC 23003-4) as described above
15 has some advantages. If combined, the loudness of the content (or what it should be after the dynamic loudness metadata is applied) would not indicate the actual loudness of the content, as the metadata available would be a composite value. Besides removing this ambiguity of what the content loudness (or program or anchor loudness) is, there are some cases where this would be particularly beneficial:

20 Keeping the metadata for dynamic processing separate allows the decoder or playback device to turn off the application of dynamic processing and to apply an implemented real-time loudness leveler instead to avoid cascading leveling. This situation may occur, for example, if the device’s own real-time leveling solution is superior to the one used with the audio codec or, for example, if the device’s own real-time leveling solution cannot be disabled and therefore will always be
25 active, resolution in further processing leading to a compromised playback experience.

Keeping the metadata for dynamic processing separate further allows transcoding to a codec that does not support dynamic loudness processing and one wishes to apply their own loudness processing prior to re-encoding.

Live to broadcast, with a single encode for a live feed would be a further example. The dynamic
30 processing metadata may be used or stored for archive or on-demand services. Therefore, for the archive or on-demand services, a more accurate, or compliant loudness measurement, based on the entire program can be carried out, and the appropriate metadata reset.

For use-cases where a fixed target loudness is used throughout the workflow, for example in a R128 compliant situation where -23 LKFS is recommended, this is also beneficial. In this scenario, the addition of the dynamic processing metadata is a “safety” measure, where the content is assumed and close to the required target and the addition of the dynamic processing metadata is a secondary check. Thus, having the ability to turn it off is desirable.

Encoding audio data and metadata for dynamic loudness adjustment

Referring to the examples of Figures 3 and 4, an encoder for encoding in a bitstream original audio data and metadata for dynamic loudness adjustment and optionally, dynamic range compression, DRC, is described which may comprise one or more processors and non-transitory memory configured to perform a method including the processes as illustrated in the steps in the example of Figure 4.

In step S201, original audio data may be input into a loudness leveler, 201, for loudness processing to obtain, as an output from the loudness leveler, 201, loudness processed audio data.

In step S202, metadata for dynamic loudness adjustment may then be generated based on the loudness processed audio data and the original audio data. Appropriate smoothing and time frames may be used to reduce artifacts.

In an embodiment, step S202 may include comparison of the loudness processed audio data to the original audio data, by an analyzer, 202, wherein the metadata may be generated based on a result of said comparison. The metadata thus generated can emulate the effect of the leveler at the decoder site. The metadata may include:

- Gain (wideband and/or multiband) processing parameters such that when applied to the original audio will produced loudness compliant audio for playback;
- Processing parameters describing the dynamics of the audio such as
 - o Peak – sample and true peak
 - o Short-term loudness values
 - o Change of short-term loudness values.

In an embodiment, step S202 may further include measuring, by the analyzer, 202, the loudness over one or more pre-defined time periods, wherein the metadata may be generated further based on the measured loudness. In an embodiment, the measuring may comprise measuring overall

loudness of the audio data. Alternatively, or additionally, in an embodiment, the measuring may comprise measuring loudness of dialogue in the audio data.

In step S203, the original audio data and the metadata may then be encoded into the bitstream.

While the format of the bitstream is not limited, in an embodiment, the bitstream may be an

- 5 MPEG-D DRC bitstream and the presence of the metadata may be signaled based on MPEG-D DRC bitstream syntax. In this case, in an embodiment, a loudnessInfoSetExtension()-element may be used to carry the metadata as a payload as detailed further below.

In an embodiment, the metadata may comprise one or more metadata payloads, wherein each metadata payload may include a plurality of sets of parameters and identifiers, with each set

- 10 including at least one of a DRCSet identifier, drcSetId, an EQSet identifier, eqSetId, and a downmix identifier, downmixId, in combination with one or more processing parameters relating to the identifiers in the set, and wherein the one or more processing parameters may be parameters for dynamic loudness adjustment by a decoder. In this case, in an embodiment, the at least one of the drcSetId, the eqSetId, and the downmixId may be related to at least one of a set
- 15 of DRC sequences, DRCSet, a set of equalizer parameters, EQSet, and a downmix, to be selected by the decoder. In general, the metadata may be said to include a plurality of sets of metadata, with each set corresponding to a respective playback condition (e.g., to a different playback condition).

In an embodiment, the method may further include generating additional metadata for static

- 20 loudness adjustment to be used by a decoder. Keeping the metadata for dynamic loudness processing and the additional metadata separate in the bitstream and encoding further the original audio data into the bitstream has several advantages as detailed above.

The methods described herein may be implemented on a decoder or an encoder, respectively, wherein the decoder and the encoder may comprise one or more processors and non-transitory

- 25 memory configured to perform said methods. An example of a device having such processing capability is illustrated in the example of Figure 5 showing said device, 300, including two processors, 301, and non-transitory memory, 302.

It is noted that the methods described herein can further be implemented on a system of an encoder for encoding in a bitstream original audio data and metadata for dynamic loudness

- 30 adjustment and optionally, dynamic range compression, DRC, and a decoder for metadata-based dynamic processing of audio data for playback as described herein.

The methods may further be implemented as a computer program product comprising a computer-readable storage medium with instructions adapted to cause the device to carry out said methods when executed by a device having processing capability. The computer program product may be stored on a computer-readable storage medium.

5 *MPEG-D DRC modified bitstream syntax*

In the following, it will be described how the MPEG-D DRC bitstream syntax, as described in ISO/IEC 23003-4, may be modified in accordance with embodiments described herein.

The MPEG-D DRC syntax may be extended, e.g. the *loudnessInfoSetExtension()*-element shown in **Table 2** below, in order to also carry the dynamic processing metadata as a frame-
10 based *dynLoudComp* update.

For example, another switch-case UNIDRCLOUDEXT_DYNLOUDCOMP may be added in the *loudnessInfoSetExtension()*-element as shown in **Table 1**. The switch-case UNIDRCLOUDEXT_DYNLOUDCOMP may be used to identify a new element *dynLoudComp()* as shown in **Table 5**. The *loudnessInfoSetExtension()*-element may be an
15 extension of the **loudnessInfoSet()**-element as shown in **Table 2**. Further, the *loudnessInfoSet()*-element may be part of the *uniDRC()*-element as shown in **Table 3**.

Symbol	Value of loudnessInfoSetExtType	Purpose
UNIDRCLOUDEXT_TERM	0×0	Termination tag
UNIDRCLOUDEXT_EQ	0×1	Extension for equalization
UNIDRCLOUDEXT_DYNLOUDCOMP	0×2	Extension for dynamic processing
(reserved)	(All remaining values)	For future use

Table 4: loudnessInfoSet extension types

5 New dynLoudComp():

Syntax	No. of bits	Mnemonic
dynLoudComp() { dynLoudCompPresent = 1 drcSetId; eqSetId; downmixId; dynLoudCompValue; }	6 6 7 10	uimbsf uimbsf uimbsf uimbsf

Table 5: Syntax of dynLoudComp()-element

- The drcSetId enables dynLoudComp (relating to the metadata) to be applied per DRC-set.
- The eqSetId enables dynLoudComp to be applied in combination with different settings for the equalization tool.
- The downmixId enables dynLoudComp to be applied per DownmixID.

In some cases, in addition to the above parameters, it may be beneficial for the dynLoudComp() element to also include a **methodDefinition** parameter (specified by, e.g., 4 bits) specifying a loudness measurement method used for deriving the dynamic program loudness metadata (e.g.,

anchor loudness, program loudness, short-term loudness, momentary loudness, etc.) and/or a **measurementSystem** parameter (specified by, e.g., 4 bits) specifying a loudness measurement system used for measuring the dynamic program loudness metadata (e.g., EBU R.128, ITU-R BS-1770 with or without preprocessing, ITU-R BS-1771, etc.). Such parameters may, e.g., be

5 included in the **dynLoudComp()** element between the **downmixId** and **dynLoudCompValue** parameters.

Alternative Syntax 1

Syntax	No. of bits	Mnemonic
loudnessInfoSetExtension() { while (loudnessInfoSetExtType != UNIDRCLOUDEXT_TERM) { extSizeBits = bitSizeLen + 4; extBitSize = bitSize + 1; switch (loudnessInfoSetExtType) { UNIDRCLOUDEXT_EQ: loudnessInfoV1AlbumCount ; loudnessInfoV1Count ; for (i=0; i<loudnessInfoV1AlbumCount; i++) { loudnessInfoV1(); } for (i=0; i<loudnessInfoV1Count; i++) { loudnessInfoV1(); } break; UNIDRCLOUDEXT_DYNLOUDCOMP: loudnessInfoV2Count ; for (i=0; i<loudnessInfoV2Count; i++) { loudnessInfoV2(); } break; /* add future extensions here */ default: for (i=0; i<extBitSize; i++) { otherBit ; } } } }	4 4 extSizeBits 6 6 6 1	uimbsf uimbsf uimbsf uimbsf uimbsf uimbsf bslbf

Table 6: Syntax of loudnessInfoSetExtension()-element

Symbol	Value of loudnessInfoSetExtType	Purpose
UNIDRCLOUDEXT_TERM	0×0	Termination tag
UNIDRCLOUDEXT_EQ	0×1	Extension for equalization
UNIDRCLOUDEXT_DYNLOUDCOMP	0×2	Extension for dynamic processing
<i>(reserved)</i>	<i>(All remaining values)</i>	For future use

Table 7: loudnessInfoSet extension types

Syntax	No. of bits	Mnemonic
loudnessInfoV2() { drcSetId; eqSetId; downmixId; samplePeakLevelPresent; if (samplePeakLevelPresent==1) { bsSamplePeakLevel; } truePeakLevelPresent; if (truePeakLevelPresent==1) { bsTruePeakLevel; measurementSystem; reliability; } measurementCount; for (i=0; i<measurementCount; i++) { methodDefinition; methodValue; measurementSystem; reliability; } dynLoudCompPresent; if (dynLoudCompPresent==1) { dynLoudCompValue; } }	 6 6 7 1 12 1 12 4 2 4 4 2..8 4 2 1 10 	 uimsbf uimsbf uimsbf bslbf uimsbf bslbf uimsbf uimsbf uimsbf uimsbf uimsbf vclbf uimsbf uimsbf bslbf uimsbf

Table 8: Syntax of loudnessInfoV2() payload

- In some cases, it may be beneficial to modify the syntax shown above in Table 8 such that the **dynLoudCompPresent** and (if **dynLoudCompPresent==1**) **dynLoudCompValue** parameters follow the **reliability** parameter within the **measurementCount** loop of the **loudnessInfoV2()** payload, rather than being outside the **measurementCount** loop. Furthermore, it may also be beneficial to set **dynLoudCompValue** equal to 0 in cases where **dynLoudCompPresent** is 0.

Alternative Syntax 2

Alternatively, the dynLoudComp()-element could be placed into the uniDrcGainExtension()-element.

Syntax	No. of bits	Mnemonic
<pre> uniDrcGain() { nDrcGainSequences = gainSequenceCount; /* from drcCoefficientsUniDrc or drcCoefficientsUniDrc V1 */ for (s=0; s<nDrcGainSequences; s++) { if (gainCodingProfile[s]<3) { drcGainSequence(); } } uniDrcGainExtPresent; if (uniDrcGainExtPresent==1) { uniDrcGainExtension(); } } </pre>	1	bslbf

5

Table 9: Syntax of uniDrcGain()-element

Syntax	No. of bits	Mnemonic
<pre> uniDrcGainExtension() { while (uniDrcGainExtType != UNIDRCGAINEXT_TERM) { extSizeBits = bitSizeLen + 4; extBitSize = bitSize + 1; switch (uniDrcGainExtType) { UNIDRCLOUDEXT_DYNLOUDCOMP: dynLoudCompCount; for (i=0; i<dynLoudCompCount; i++) { dynLoudComp(); } break; /* add future extensions here */ default: for (i=0; i<extBitSize; i++) { otherBit; } } } } </pre>	4 3 extSizeBits 6 1	uimsbf uimsbf uimsbf uimsbf bslbf

Table 10: Syntax of uniDrcGainExtension()-element

Symbol	Value of uniDrcGainExtType	Purpose
UNIDRCGAINEXT_TERM	0x0	Termination tag
UNIDRCGAINEXT_DYNLOUDCOMP	0x1	Extension for dynamic processing
(reserved)	(All remaining values)	For future use

Table 11: UniDrc gain extension types

*Semantics***dynLoudCompValue**

This field contains the value for *dynLoudCompDb*. The values are encoded according to the table below. The default value is 0 dB.

Encoding	Size	Mnemonic	Value in [dB]	Approximate range
μ	10 bits	uimsbf	$\text{dynLoudCompDb} = -16 + \mu 2^{-5}$	-16 ... 16 dB, 0.0312 dB step size

Table 12: Coding of dynLoudCompValue field

5 *Updated MPEG-D DRC Loudness Normalization Processing*

```

if (targetLoudnessPresent) {
    if (dynLoudCompPresent) {
        loudnessNormalizationGainDb = targetLoudness - contentLoudness +
dynLoudCompDb;
    } else {
        loudnessNormalizationGainDb = targetLoudness - contentLoudness;
    }
} else {
    loudnessNormalizationGainDb = 0.0;
}
if (loudnessNormalizationGainDbMaxPresent) {
    loudnessNormalizationGainDb = min(loudnessNormalizationGainDb,
loudnessNormalizationGainDbMax);
}
if (loudnessNormalizationGainModificationDbPresent) {
    gainNorm = pow(2.0, (loudnessNormalizationGainDb +
loudnessNormalizationGainModificationDb) / 6);
} else {
    gainNorm = pow(2.0, loudnessNormalizationGainDb / 6);
}
for (t=0; t<drcFrameSize; t++) {
    for (c=0; c<nChannels; c++) {
        audioSample[c][t] = gainNorm * audioSample[c][t];
    }
}

```

Table 13: Loudness normalization processing

Pseudo-Code for selection and processing of dynLoudComp

```

/* Selection Process */

5  /* The following settings would be derived from user/decoder settings

    drcSetId = 1;

    eqSetID = 2;

    downmixId = 3;

    */

10 findMatchingDynLoudComp (drcSetId, eqSetID, downmixId) {

    dynLoudComp = UNDEFINED;

    /* Check if matching loudnessInfo set is present */

    if (targetLoudnessPresent(drcSetId, eqSetID, downmixId)) continue;

    else {

15     dynLoudCompPresent = false;

        exit(0);

    }

    /* If all values are defined */

    if (drcSetId != UNDEFINED && eqSetID != UNDEFINED && downmixId !=

20 UNDEFINED) {

        for (num=0; num<num_of_dynLoudCompValues; num++) {

            if (dynLoudCompArray[num].drcSetId == drcSetId &&

                dynLoudCompArray[num].eqSetID == eqSetID &&

                dynLoudCompArray[num].downmixId == downmixId) {

25             /* Correct entry found, assign to dynLoudComp */

                dynLoudComp = dynLoudCompArray[num].dynLoudCompValue;

```

```

        }

    }

}

else if (drcSetId == UNDEFINED) {

5     for (num=0; num<num_of_dynLoudCompValues; num++) {

        if (dynLoudCompArray[num].eqSetID == eqSetID &&

            dynLoudCompArray[num].downmixId == downmixId) {

                /* Correct entry found, assign to dynLoudComp */

                dynLoudComp = dynLoudCompArray[num].dynLoudCompValue;

10         }

    }

}

else if (eqSetID == UNDEFINED) {

    for (num=0; num<num_of_dynLoudCompValues; num++) {

15         if (dynLoudCompArray[num].drcSetId == drcSetId &&

            dynLoudCompArray[num].downmixId == downmixId) {

                /* Correct entry found, assign to dynLoudComp */

                dynLoudComp = dynLoudCompArray[num].dynLoudCompValue;

                }

20     }

}

else if (downmixId == UNDEFINED) {

    for (num=0; num<num_of_dynLoudCompValues; num++) {

        if (dynLoudCompArray[num].drcSetId == drcSetId &&

25         dynLoudCompArray[num].eqSetID == eqSetID) {

```

```
        /* Correct entry found, assign to dynLoudComp */  
        dynLoudComp = dynLoudCompArray[num].dynLoudCompValue;  
    }  
}  
5    }  
    else if (drcSetId == UNDEFINED && downmixId == UNDEFINED) {  
        for (num=0; num<num_of_dynLoudCompValues; num++) {  
            if (dynLoudCompArray[num].eqSetID == eqSetID) {  
                /* Correct entry found, assign to dynLoudComp */  
10         dynLoudComp = dynLoudCompArray[num].dynLoudCompValue;  
            }  
        }  
    }  
    else if (drcSetId == UNDEFINED && eqSetID == UNDEFINED) {  
15     for (num=0; num<num_of_dynLoudCompValues; num++) {  
        if (dynLoudCompArray[num].downmixId == downmixId) {  
            /* Correct entry found, assign to dynLoudComp */  
            dynLoudComp = dynLoudCompArray[num].dynLoudCompValue;  
        }  
20     }  
    }  
    else if (eqSetID == UNDEFINED && downmixId == UNDEFINED) {  
        for (num=0; num<num_of_dynLoudCompValues; num++) {  
            if (dynLoudCompArray[num].eqSetID == eqSetID) {  
25         /* Correct entry found, assign to dynLoudComp */
```

```
        dynLoudComp = dynLoudCompArray[num].dynLoudCompValue;
    }
}
}
5    if (dynLoudComp == UNDEFINED){
        dynLoudCompPresent = false;
        exit(1);
    }
    else {
10        dynLoudCompPresent = true;
        exit(0);
    }
}

15  /* Processing */
    if (targetLoudnessPresent) {
        if (dynLoudCompPresent) {
            loudnessNormalizationGainDb = targetLoudness – contentLoudness + dynLoudCompDb;
        } else {
20        loudnessNormalizationGainDb = targetLoudness – contentLoudness;
        }
    } else {
        loudnessNormalizationGainDb = 0.0;
    }

25    if (loudnessNormalizationGainDbMaxPresent) {
```

```

        loudnessNormalizationGainDb = min(loudnessNormalizationGainDb,
loudnessNormalizationGainDbMax);

    }

    if (loudnessNormalizationGainModificationDbPresent) {

5        gainNorm = pow(2.0, (loudnessNormalizationGainDb +
loudnessNormalizationGainModificationDb) / 6);

    } else {

        gainNorm = pow(2.0, loudnessNormalizationGainDb / 6);

    }

10    for (t=0; t<drcFrameSize; t++) {

        for (c=0; c<nChannels; c++) {

            audioSample[c][t] = gainNorm * audioSample[c][t];

        }

    }

15

    In some cases, in addition to the selection process shown in pseudo-code above (e.g., taking
    drcSetId, eqSetID and downmixId into consideration for selecting a dynLoudCompValue
    parameter), it may be beneficial for the selection process to also take a methodDefinition
    parameter and/or a measurementSystem parameter into consideration for selecting a
20    dynLoudCompValue parameter.

```

Alternative Updated MPEG-D DRC Loudness Normalization Processing

```

if (targetLoudnessPresent) {
    loudnessNormalizationGainDb = targetLoudness - contentLoudness + dynLoudCompDb;
}
else {
    loudnessNormalizationGainDb = 0.0;
}
if (loudnessNormalizationGainDbMaxPresent) {
    loudnessNormalizationGainDb = min(loudnessNormalizationGainDb,
    loudnessNormalizationGainDbMax);
}
if (loudnessNormalizationGainModificationDbPresent) {
    gainNorm = pow(2.0, (loudnessNormalizationGainDb +
    loudnessNormalizationGainModificationDb) / 6);
} else {
    gainNorm = pow(2.0, loudnessNormalizationGainDb / 6);
}
for (t=0; t<drcFrameSize; t++) {
    for (c=0; c<nChannels; c++) {
        audioSample[c][t] = gainNorm * audioSample[c][t];
    }
}

```

Table 14: Alternative Loudness normalization processing

In cases where the alternative loudness normalization processing of Table 14 above is used, the loudness normalization processing pseudo-code described above may be replaced by the

- 5 following alternative loudness normalization processing pseudo-code. Note that a default value of dynLoudCompDb, e.g., 0 dB, may be assumed to ensure that the value of dynLoudCompDb is defined, even for cases where dynamic loudness processing metadata is not present in the bitstream.

/* Alternative Processing */

- 10 **if** (targetLoudnessPresent) {
- loudnessNormalizationGainDb = targetLoudness - contentLoudness + dynLoudCompDb;
- }
- else** {
- loudnessNormalizationGainDb = 0.0;

```

    }

    if (loudnessNormalizationGainDbMaxPresent) {

        loudnessNormalizationGainDb = min(loudnessNormalizationGainDb,
        loudnessNormalizationGainDbMax);
5    }

    if (loudnessNormalizationGainModificationDbPresent) {

        gainNorm = pow(2.0, (loudnessNormalizationGainDb +
        loudnessNormalizationGainModificationDb) / 6);

    } else {
10        gainNorm = pow(2.0, loudnessNormalizationGainDb / 6);
    }

    for (t=0; t<drcFrameSize; t++) {

        for (c=0; c<nChannels; c++) {

            audioSample[c][t] = gainNorm * audioSample[c][t];
15        }

    }

```

Alternative Syntax 3

20 In some cases, it may be beneficial to combine the syntax described above in Table 1 – Table 5 with Alternative Syntax 1 described above in Table 6 – Table 8, as shown in the following Tables, to allow increased flexibility for transmission of the dynamic loudness processing values.

Syntax	No. of bits	Mnemonic
loudnessInfoSetExtension() { while (loudnessInfoSetExtType != UNIDRCLOUDEXT_TERM) { extSizeBits = bitSizeLen + 4; extBitSize = bitSize + 1; switch (loudnessInfoSetExtType) { UNIDRCLOUDEXT_EQ: loudnessInfoV1AlbumCount; loudnessInfoV1Count; for (i=0; i<loudnessInfoV1AlbumCount; i++) { loudnessInfoV1(); } for (i=0; i<loudnessInfoV1Count; i++) { loudnessInfoV1(); } break; UNIDRCLOUDEXT_DYNLOUDCOMP: loudnessInfoV2Count; for (i=0; i<loudnessInfoV2Count; i++) { loudnessInfoV2(); } break; UNIDRCLOUDEXT_DYNLOUDCOMP2: dynLoudCompCount; for (i=0; i<dynLoudCompCount; i++) { dynLoudComp(); } break; /* add future extensions here */ default: for (i=0; i<extBitSize; i++) { otherBit; } } } }	4 4 extSizeBits 6 6 6 1	uimbsf uimbsf uimbsf uimbsf uimbsf uimbsf bslbf

Table 15: Alternate Syntax 3 of loudnessInfoSetExtension()-element

Symbol	Value of loudnessInfoSetExtType	Purpose
UNIDRCLOUDEXTERM	0×0	Termination tag
UNIDRCLOUDEXTEQ	0×1	Extension for equalization
UNIDRCLOUDEXTDYNLOUDCOMP	0×2	Extension for dynamic processing
UNIDRCLOUDEXTDYNLOUDCOMP2	0×3	Extension for dynamic processing
(reserved)	(All remaining values)	For future use

Table 16: Alternate Syntax 3 of loudnessInfoSet extension types

Syntax	No. of bits	Mnemonic
loudnessInfoV2() { drcSetId; eqSetId; downmixId; samplePeakLevelPresent; if (samplePeakLevelPresent==1) { bsSamplePeakLevel; } truePeakLevelPresent; if (truePeakLevelPresent==1) { bsTruePeakLevel; measurementSystem; reliability; } measurementCount; for (i=0; i<measurementCount; i++) { methodDefinition; methodValue; measurementSystem; reliability; dynLoudCompPresent; if (dynLoudCompPresent==1) { dynLoudCompValue; } else { dynLoudCompValue=0; } } }	6 6 7 1 12 1 12 4 2 4 4 2..8 4 2 1 10	uimsbf uimsbf uimsbf bslbf uimsbf bslbf uimsbf uimsbf uimsbf uimsbf uimsbf vlclbf uimsbf uimsbf uimbsf uimsbf

Table 17: Alternate Syntax 3 of loudnessInfoV2() payload

Alternate Syntax 3 of dynLoudComp():

Syntax	No. of bits	Mnemonic
dynLoudComp() { dynLoudCompPresent = 1 drcSetId; eqSetId; downmixId; methodDefinition; measurementSystem; dynLoudCompValue; }	 6 6 7 4 4 10	 uimsbf uimsbf uimsbf uimsbf uimsbf uimsbf

Table 18: Alternate Syntax 3 of dynLoudComp()-element

Interface Extension Syntax

- 5 In some cases, it may be beneficial to allow control, e.g., by an end user, of whether or not dynamic loudness processing is performed, even when dynamic loudness processing information is present in a received bitstream. Such control may be provided by updating the MPEG-D DRC interface syntax to include an additional interface extension (e.g., UNIDRCINTERFACEEXT_DYNLOUD) which contains a revised loudness normalization control interface payload (e.g., loudnessNormalizationControlInterfaceV1()) as shown in the
- 10 following tables.

Syntax	No. of bits	Mnemonic
uniDrcInterfaceExtension() { while (uniDrcInterfaceExtType!=UNIDRCINTERFACEEXT_TERM) { extSizeBits = bitSizeLen + 4; extBitSize = bitSize + 1; switch (uniDrcInterfaceExtType) { UNIDRCINTERFACEEXT_EQ: loudnessEqParameterInterfacePresent; if (loudnessEqParameterInterfacePresent == 1) { loudnessEqParameterInterface(); } equalizationControlInterfacePresent; } }	 4 4 extSizeBits 1 1	 uimsbf uimsbf uimsbf bslbf bslbf

Syntax	No. of bits	Mnemonic
<pre> if (equalizationControlInterfacePresent == 1) { equalizationControlInterface(); } break; UNIDRCINTERFACEEXT_DYNLOUD: loudnessNormalizationControlInterfaceV1(); break; /* add future extensions here */ default: for (i=0; i<extBitSize; i++) { otherBit; } } } </pre>	1	bslbf

Table 19: Syntax of uniDRCInterfaceExtension() payload

Syntax	No. of bits	Mnemonic
<pre> loudnessNormalizationControlInterfaceV1() { loudnessNormalizationOn; if (loudnessNormalizationOn == 1) { targetLoudness; dynLoudnessNormalizationOn; } } </pre>	1 12 1	bslbf uimsbf bslbf

Table 20: Syntax of loudnessNormalizationControlInterfaceV1() payload

Symbol	Value of uniDRCInterfaceExtType	Purpose
UNIDRCINTERFACEEXT_TERM	0×0	Termination tag
UNIDRCINTERFACEEXT_EQ	0×1	Equalization Control
UNIDRCINTERFACEEXT_DYNLOUD	0×2	Dynamic Loudness Control
(reserved)	(All remaining values)	For future use

Table 21: UniDRC Interface extension types

Interface Extension Semantics

loudnessNormalizationOn This flag signals if loudness normalization processing should be switched on or off. The default value is 0. If *loudnessNormalizationOn* == 0, *loudnessNormalizationGainDb* shall be set to 0 dB.

targetLoudness This field contains the desired output loudness. The values are encoded according to the following Table.

5

Encoding	Size	Mnemonic	Value in [dB]	Approximate range
μ	12 bits	uimsbf	$L_{TL} = -\mu 2^{-5}$	-128 ... 0 dB, 0.0312 dB step size

Table 22: Coding of targetLoudness field

dynLoudnessNormalizationOn This flag signals if dynamic loudness normalization processing should be switched on or off. The default value is 0. If *dynLoudnessNormalizationOn* == 0, *dynloudnessNormalizationGainDb* shall be set to 0 dB.

Additional Methods of Enabling Dynamic Loudness Compensation

10 In addition to the above, additional methods for enabling dynamic loudness compensation are

possible. For example, instead of including a parameter which indicates a specific dynamic loudness compensation value to be applied to the signal, it is possible to define a specific set of DRC gains which already include the dynamic loudness compensation, such that, when those specific DRC gains are applied by a decoder, dynamic loudness compensation is applied. DRC gain sets suitable for applying such dynamic loudness compensation may be identified through specific bitstream elements, as explained further below. A benefit of such an approach is that it is not necessary to transmit an explicit dynamic loudness compensation gain in addition to other loudness information.

Additionally, a user can specify through an interface that the DRC gains including dynamic loudness compensation should be selected and applied, as explained further below.

Finally, additional bitstream elements may be provided to allow tighter creative control over whether and how dynamic loudness compensation is performed. For instance, a flag may indicate whether or not dynamic loudness compensation may be switched OFF by a user. In such cases, if the content creator allows for dynamic loudness compensation to be disabled, then dynamic loudness compensation will be applied or not as specified by the user.

Specifying DRC Sets with Dynamic Loudness Compensation through the DRC Set Effect Field

In the MPEG-D DRC standard, it is possible to indicate one or more DRC effects which are provided by a specific DRC set. For instance, a DRC set might be appropriate for “Late Night” viewing, for viewing in “Noisy Environments”, for viewing at “Low Playback Levels”, etc. A DRC set effect parameter may indicate which specific effects are provided by the DRC set.

Furthermore, the MPEG-D DRC standard allows a user to specify one or more desired DRC effects, as well as one or more optional fallback DRC effects. Such information from a user may be used to select a most appropriate DRC set from the available DRC sets. For example, if a DRC set exists which matches the desired DRC effects, then that set is selected. If such a set does not exist, but there is a set that matches a fallback DRC effect, that set may be selected.

The list of defined DRC effects may be extended to include a DRC effect which provides dynamic loudness compensation. Furthermore, the list of DRC effects which can be specified by a user through an interface may also be extended to include a DRC set effect which provides dynamic loudness compensation. As such, dynamic loudness compensation may be provided by including a DRC set indicated as providing a dynamic loudness compensation effect, and indicating, through an interface, a desire for a DRC set which provides dynamic loudness compensation, in which case, the decoder would select and apply the DRC set that corresponds

to the dynamic loudness compensation DRC effect.

As described above, a previous version of a table specifying DRC Set Effects that may be signaled in a bitstream (e.g., Table A.45 of ISO/IEC 23003-4:2020) may be updated to include an entry for a “Dynamic Loudness Compensation” effect as shown below. Specifically, a row with a particular bit position (e.g., 13) corresponding to a Dynamic Loudness Compensation effect may be added as shown in the table below.

bit position	drcSetEffect	Short name	Description
1 (LSB)	Late night	“Night”	For quiet environment, listening at low level, avoiding to disturb others.
2	Noisy environment	“Noisy”	Optimized to get the best experience in noisy environments, for instance by amplifying soft sections.
3	Limited playback range	“Limited”	Reduced dynamic range to improve quality on playback devices with limited dynamic range.
4	Low playback level	“LowLevel”	Listening at a low playback level.
5	Dialog enhancement	“Dialog”	The main effect is a more prominent dialogue within the content.
6	General compression	“General”	A DRC effect that reduces the dynamic range and is applicable to multiple playback scenarios.
7	Dynamic expansion	“Expand”	Dynamics enhancement.
8	Artistic effect	“Artistic”	To create an artistic sound effect.
9	Clipping prevention	“Clipping”	The main purpose is clipping prevention.
10	Fade-in/fade-out	“Fade”	Fade-in and fade-out envelope for gapless content (applicable when not playing in Album mode). It has no dynamic compression.
11	Ducking other	“Duck other”	An effect that attenuates all audio content except for the channelGroup it is associated with. It has no dynamic compression.
12	Ducking self	“Duck self”	An effect that attenuates all channelGroups that it is associated with. It is identical to “Ducking other”, however, it has the same channelGroup assignment as regular DRC gains.
13	Dynamic Loudness Compensation	“Dyn Loud Comp”	An effect that compensates for dynamic loudness changes in the decoder.
<i>Remaining values are reserved</i>			
NOTE Any information on DRC sets with “Ducking” effect in this document shall refer to both “Ducking other” and “Ducking self” if not stated otherwise.			

Table 23: Examples of DRC Effects

In the above example, a bit in a bitfield (e.g. a `drcSetEffect` bitfield of a `drcInstructionsBasic()` or a `drcInstructionsUniDrc()` or a `drcInstructionsUniDrcV1()` payload) associated with a particular DRC Set may be set (e.g., bit position 13) to indicate that the DRC Set provides the Dynamic Loudness Compensation effect.

- 5 Similarly, a previous version of a table specifying DRC Set Effects that may be specified to a decoder through an interface (e.g., Table 12 of ISO/IEC 23003-4:2020) may be updated to include an entry for a Dynamic Loudness Compensation effect as shown below. Specifically, a row with a particular Index Value (e.g., 9) corresponding to a Dynamic Loudness Compensation effect may be added as shown in the table below.

Index	drcSetEffect	Short name	Description
0	None	"None"	A DRC set that has no dynamic compression, for instance "Clipping", or applying no DRC set for compression.
1	Late night	"Night"	For quiet environment, listening at low level, avoiding to disturb others.
2	Noisy environment	"Noisy"	Optimized to get the best experience in noisy environments, for instance by amplifying soft sections.
3	Limited playback range	"Limited"	Reduced dynamic range to improve quality on playback devices with limited dynamic range.
4	Low playback level	"LowLevel"	Listening at a low playback level.
5	Dialog enhancement	"Dialog"	The main effect is a more prominent dialogue within the content.
6	General compression	"General"	A DRC effect that reduces the dynamic range and is applicable to multiple playback scenarios.
7	Dynamic expansion	"Expand"	Dynamics enhancement.
8	Artistic effect	"Artistic"	To create an artistic sound effect.
9	Dynamic Loudness Compensation	"DynLoudComp"	Enables dynamic loudness compensation in the decoder.

10 *Table 24: Examples of DRC Effects that may be specified through a interface*

- In the above example, an interface parameter (e.g., a `drcEffectTypeRequest` parameter of a `dynamicRangeControllerInterface()` payload) may be set to the specific value (e.g., 9) that corresponds to the Dynamic Loudness Compensation Effect to instruct the decoder that, if available, a DRC Set which provides the Dynamic Loudness Compensation Effect set should be selected and applied. However, because dynamic loudness compensation may be intended for use in combination with loudness normalization, in some cases a decoder may require that in order to select and apply a DRC set that provides a dynamic loudness compensation effect, a decoder must also perform loudness normalization. Such a requirement could be accomplished by only allowing the decoder to select a DRC set that provides a dynamic loudness compensation
- 15

effect when loudness normalization is also enabled in the decoder. Alternatively, or in addition, such a requirement could be satisfied by modifying the interface to require that when requesting a DRC set which provides a dynamic loudness compensation effect, loudness normalization must also be enabled (e.g., by setting both a loudness normalization on flag and a target loudness value).

Specifying DRC Sets with Dynamic Loudness Compensation using new extension payloads

Alternatively, a new payload (e.g., dynLoudInstructions()) including instructions for dynamic loudness compensation could be defined. One or more of such payloads could, e.g., be included in an extension field (e.g., UNIDRCCONFEXT_V2) of an existing bitstream. By including such payloads in an extension field, compatibility with existing decoders that were not designed to parse the new payload could be maintained, as such decoders would simply ignore the additional payload(s). Each of such payloads would be assigned a unique identifier (e.g., drcSetId) which corresponds to a DRC set that provides dynamic loudness compensation, and, to help easily identify such sets, a flag (e.g., a dynamicLoudCompDRCSet flag) could be set to 1 for each such DRC set signaled through the new payload.

Additionally, the user interface described above for selecting DRC sets that provide the dynamic loudness compensation effect could be used for DRC sets signaled through this new type of payload (e.g., dynLoudInstructions() payload) as well. Specifically, a user could indicate to the decoder through the interface (e.g., using the drcEffectTypeRequest field) that a DRC set that provides dynamic loudness compensation is desired. The decoder would then select, if present, a DRC set signaled through the new payloads (e.g., a DRC set having a DRC Set ID that is identified as corresponding to dynamic loudness compensation, and, for instance, having a dynamicLoudCompDRCSet flag set to 1). However, as noted above, because dynamic loudness compensation may be intended for use in combination with loudness normalization, in some cases a decoder may require that in order to select and apply a DRC set that provides a dynamic loudness compensation effect, a decoder must also perform loudness normalization. Such a requirement could be accomplished by only allowing the decoder to select a DRC set that provides a dynamic loudness compensation effect when loudness normalization is also enabled in the decoder. Alternatively, or in addition, such a requirement could be satisfied by modifying the interface to require that when requesting a DRC set which provides a dynamic loudness compensation effect, loudness normalization must also be enabled (e.g., by setting both a loudness normalization on flag and a target loudness value).

An advantage to using this type of signaling (e.g., instead of signaling DRC sets which provide

dynamic loudness compensation through an existing parameter, such as a drcSetEffect parameter), is that, as described above, additional parameters indicating whether and how a user can enable and/or disable dynamic loudness compensation can be included in such new payloads. For instance, the new payload may also include a parameter (e.g., a dynamicLoudCompSwOffAllowed parameter) which indicates whether or not switching dynamic loudness compensation off through the user interface is allowed for each DRC set that applies dynamic loudness compensation.

An example syntax for a new payload (e.g., a dynLoudInstructions() payload) for dynamic loudness compensation is shown in the table below.

Syntax	No. of bits	Mnemonic
dynLoudInstructions() { dynamicLoudCompDRCSet = 1; drcSetId; drcLocation; downmixId; additionalDownmixIdPresent; if (additionalDownmixIdPresent==1) { additionalDownmixIdCount; for (j=0; j<additionalDownmixIdCount; j++) { additionalDownmixId; } } else { additionalDownmixIdCount = 0; } dynamicLoudCompSwOffAllowed; }	 6 4 7 1 3 7 1	 uimbsbf uimbsbf uimbsbf bslbf uimbsbf uimbsbf bslbf

Table 25: Example Syntax of a Dynamic Loudness Instructions Payload

As described above, a unique DRC Set ID, as well as a flag (e.g., dynamicLoudCompDRCSet) which specifically indicates the DRC Set as a DRC set that provides dynamic loudness compensation, may be associated with the payload. Furthermore, the payload may be associated with one or more downmix identification parameters (e.g., downmixId parameters) which indicate that the DRC set is intended for use with one or more specific downmixes / downmix configurations of the audio program. Such DRC Set identifier (e.g., drcSetId parameter), as well as the flag (e.g., dynamicLoudCompDRCSet flag) allow for selection of the DRC Set associated with the payload in response to a user indication, provided through an interface, that a dynamic

loudness compensation DRC effect is desired by the user. Additionally, the inclusion of one or more downmix identifiers (e.g., downmixID parameters) enables the selection of a DRC set that provides dynamic loudness compensation which is intended for use with a particular downmix or downmixes of the audio program. Finally, as also described above, the inclusion of a parameter

5 (e.g., dynamicLoudComSwOffAllowed) indicating whether or not turning off dynamic loudness compensation is allowed, enables a content creator additional control over whether or not dynamic loudness compensation may be applied by decoders.

Finally, as indicated above, such additional payloads (e.g., dynLoudInstructions() payloads) may be contained in an extension payload of a configuration payload (e.g., a

10 uniDrcConfigExtension() payload). A specific extension type (e.g., a uniDrcConfigExtType of UNIDRCCONFEXT_V2 having a value of 0x2) may be used to indicate the presence of such additional payloads within the extension field of the configuration payload.

An example of a syntax for an extension field of a configuration payload (e.g., a uniDrcConfigExtension() payload) is shown in the table below.

Syntax	No. of bits	Mnemonic
uniDrcConfigExtension() { while (uniDrcConfigExtType != UNIDRCCONFEXT_TERM) { extSizeBits = bitSizeLen + 4; extBitSize = bitSize + 1; switch (uniDrcConfigExtType) { case UNIDRCCONFEXT_PARAM_DRC: drcCoefficientsParametricDrc(); parametricDrcInstructionsCount; for (i=0; i<parametricDrcInstructionsCount; i++) { parametricDrcInstructions (); } break; case UNIDRCCONFEXT_V1: downmixInstructionsV1Present; if (downmixInstructionsV1Present==1) { downmixInstructionsV1Count; for (i=0; i<downmixInstructionsV1Count; i++) { downmixInstructionsV1(); } } } } } }	4 4 extSizeBits 4 1 7	uimsbf uimsbf uimsbf uimsbf bslbf uimsbf

Syntax	No. of bits	Mnemonic
drcCoeffsAndInstructionsUniDrcV1Present;	1	bslbf
if (drcCoeffsAndInstructionsUniDrcV1Present==1) {		
drcCoefficientsUniDrcV1Count;	3	uimsbf
for (i=0; i<drcCoefficientsUniDrcV1Count; i++) {		
drcCoefficientsUniDrcV1();		
}		
drcInstructionsUniDrcV1Count;	6	uimsbf
for (i=0; i<drcInstructionsUniDrcV1Count; i++) {		
drcInstructionsUniDrcV1();		
}		
}		
loudEqInstructionsPresent;	1	bslbf
if (loudEqInstructionsPresent==1) {		
loudEqInstructionsCount;	4	uimsbf
for (i=0; i<loudEqInstructionsCount; i++) {		
loudEqInstructions();		
}		
}		
eqPresent;	1	bslbf
if (eqPresent==1) {		
eqCoefficients();		
eqInstructionsCount;	4	uimsbf
for (i=0; i<eqInstructionsCount; i++) {		
eqInstructions();		
}		
}		
break;		
case UNIDRCCONFEXT_V2:		
downmixInstructionsV1Present;	1	bslbf
if (downmixInstructionsV1Present==1) {		
downmixInstructionsV1Count;	7	uimsbf
for (i=0; i<downmixInstructionsV1Count; i++) {		
downmixInstructionsV1();		
}		
}		
drcCoeffsAndInstructionsUniDrcV1Present;	1	bslbf
if (drcCoeffsAndInstructionsUniDrcV1Present==1) {		
drcCoefficientsUniDrcV1Count;	3	uimsbf
for (i=0; i<drcCoefficientsUniDrcV1Count; i++) {		

Syntax	No. of bits	Mnemonic
<pre> drcCoefficientsUniDrcV1(); } drcInstructionsUniDrcV1Count; for (i=0; i<drcInstructionsUniDrcV1Count; i++) { drcInstructionsUniDrcV1(); } </pre>	6	uimsbf
<pre> } loudEqInstructionsPresent; if (loudEqInstructionsPresent==1) { loudEqInstructionsCount; for (i=0; i<loudEqInstructionsCount; i++) { loudEqInstructions(); } } eqPresent; if (eqPresent==1) { eqCoefficients(); eqInstructionsCount; for (i=0; i<eqInstructionsCount; i++) { eqInstructions(); } } dynLoudPresent; if (dynLoudPresent ==1) { dynLoudInstructionsCount; for (i=0; i<eqInstructionsCount; i++) { dynLoudInstructions(); } } break; /* add future extensions here */ default: for (i=0; i<extBitSize; i++) { otherBit; } } } } </pre>	1 4 1 4 1 4 1	bslbf uimsbf bslbf uimsbf bslbf uimsbf bslbf

Table 26: Example for Updated Syntax of a Configuration Payload including Extension Fields

The data associated with the specific extension type of the extension field of the configuration payload contains a parameter which indicates whether dynamic loudness compensation instructions are present (e.g., a `dynLoudPresent` parameter), and if so, a parameter which indicates the number of sets of dynamic loudness compensation instructions in the extension payload (e.g., a `dynLoudInstructionsCount` parameter), and each set of dynamic loudness compensation instructions (e.g., each `dynLoudInstructions()` payload), where the syntax for the dynamic loudness compensation instructions may be as shown in the table above.

10 *Signaling DRC Sets for Dynamic Loudness Compensation through existing configuration elements*

Still alternatively, it may be possible to indicate that a particular DRC set provides dynamic loudness compensation by adding a new field to an existing configuration element. For example, a new parameter (e.g., a `dynamicLoudCompDRCSet` parameter) that indicates whether or not a DRC set provides dynamic loudness compensation could be included in an existing configuration element (e.g., in a `downmixInstructions()` element, a `drcInstructionsBasic()` element, or a `drcInstructionsUniDrc()` element). Ideally, to avoid compatibility issues with legacy decoders, the new element could be included in a field which is already ignored by legacy decoders (e.g., in a reserved field). Alternatively, the new parameter could be included in an updated version of an existing configuration element (e.g., in a `downmixInstructionsV2()` element or a `drcInstructionsUniDrcV2()` element), which would be ignored by legacy decoders.

Finally, similarly to the selection process described above, DRC Sets that are signaled as providing dynamic loudness compensation in this manner could be selected through an interface to the decoder by indicating that a DRC Set is desired that provides a dynamic loudness compensation effect. When such indication is provided through the interface to the decoder, the decoder will identify and select a DRC Set which is indicated as providing dynamic loudness compensation (e.g., which has a `dynamicLoudCompDRCSet` parameter that equals 1). If multiple DRC sets are indicated as providing dynamic loudness compensation, a decoder may use other rules (e.g., predefined rules, such as those defined in ISO/IEC 23003-4:2020) for selecting the most appropriate DRC set of the multiple DRC sets that provide dynamic loudness compensation.

However, because dynamic loudness compensation may be intended for use in combination with loudness normalization, in some cases a decoder may require that in order to select and apply a

DRC set that provides a dynamic loudness compensation effect, a decoder must also perform loudness normalization. Such a requirement could be accomplished by only allowing the decoder to select a DRC set that provides a dynamic loudness compensation effect when loudness normalization is also enabled in the decoder. Alternatively, or in addition, such a requirement
5 could be satisfied by modifying the interface to require that when requesting a DRC set which provides a dynamic loudness compensation effect, loudness normalization must also be enabled (e.g., by setting both a loudness normalization on flag and a target loudness value).

Figure 6 illustrates an example method of processing audio data for playback as described above. Referring to the example of Figure 6, in step S301 the decoder 100 may receive a bitstream
10 including encoded audio data and metadata, wherein the metadata includes one or more dynamic range control (DRC) sets, and for each DRC set, an indication of whether the DRC set is configured for providing a dynamic loudness compensation effect, e.g., as described above.

While the format of the bitstream is not limited, in an embodiment, the bitstream may be an MPEG-D DRC bitstream. The presence of metadata for providing a dynamic loudness
15 compensation effect may then be signaled based on MPEG-D DRC bitstream syntax, e.g., as described above.

In step S302, the decoder may then parse the metadata to identify DRC sets that are configured for providing the dynamic loudness compensation effect, e.g., as described above.

In step S303, the decoder may then decode the audio data to obtain decoded audio data, e.g., as
20 described above.

In step S304, the decoder may then select one of the identified DRC sets configured for providing the dynamic loudness compensation effect, e.g., as described above.

In step S305, the decoder may then extract one or more DRC gains corresponding to the selected DRC set from the bitstream, e.g., as described above.

25 In step S306, the decoder may then apply the one or more DRC gains corresponding to the selected DRC set to the decoded audio data to obtain dynamic loudness compensated audio data, e.g., as described above.

Finally, in step S307, the dynamic loudness compensated audio data may then be output for playback, e.g., as described above.

Efficient Coding / Transmission of Dynamic Loudness Compensation Data

As described above, dynamic loudness compensation, in combination with dynamic range control, may be achieved by transmitting a DRC set that contains gains which are a combination of dynamic range control gains and dynamic loudness compensation gains. However, doing so may require redundant transmission of dynamic range control gain data. For instance, a first DRC set may contain gains for applying dynamic range control to a signal, and it may be important to allow for that DRC gain set to be applied independently of any other gains (e.g., dynamic loudness compensation gains). Therefore, if a DRC set is desired which contains gains for applying a combination of dynamic range control and dynamic loudness compensation, then a second DRC set needs to be specified which represents a combination of the dynamic range control gains of the first set and the desired dynamic loudness compensation gains. Unfortunately, doing so requires that both the first and the second DRC sets contain information about the dynamic range control gains of the first set, which is inefficient.

A more efficient way of transmitting data to accomplish the same goal is to provide only the dynamic range control gains in a first DRC set, and only the dynamic loudness compensation gains in a second DRC set, along with some additional metadata indicating the relationship between the two DRC sets. For instance, the first DRC set may include a parameter which indicates that there is another DRC set which depends on the first DRC set. Then, a decoder may understand that it is possible to apply the gains of the first set independently, e.g., in case only dynamic range control is desired. Additionally, the decoder will understand that it is also possible to combine the gains of the first and second DRC sets in order to obtain a combination of dynamic range control and dynamic loudness compensation. Finally, in cases where the gains of the second set are never intended to be applied alone, metadata can be included with the second DRC set to indicate whether the gains of that DRC set may only be used in conjunction with the first set, or whether they may be used independently of the gains of the first set.

For example, consider 2 DRC sets, with a first DRC set identified as DRC Set 1 (e.g., has a drcSetId equal to 1) providing gains for dynamic range control, and a second DRC Set identified as DRC Set 2 (e.g., has a drcSetId equal to 2) providing gains for dynamic loudness compensation. Since DRC Set 2 is intended for use in combination with DRC Set 1, DRC Set 1 could include a parameter indicating that DRC Set 2 depends on DRC Set 1 (e.g., a dependsOnDrcSet parameter of DRC Set 1 may be equal to 2). Furthermore, if dynamic loudness

- compensation is only intended for use in combination with dynamic range control, then DRC Set 2 could include a parameter indicating that the gains of DRC Set 2 may not be used independently (e.g., a noIndependentUse flag of DRC Set 2 may be set to a value 1). Conversely, if DRC Set 2 was not intended only for use in combination with dynamic range control, then
- 5 DRC Set 2 could include a parameter indicating that the gains of DRC Set 2 may be used independently (e.g., a noIndependentUse flag of DRC Set 2 may be set to a value of 0).

The above described Efficient Coding/Transmission of Dynamic Loudness Compensation data may be used with any of the above described Additional Methods of Enabling Dynamic Loudness Compensation.

- 10 For instance, in the first case, where the presence of a DRC Set which provides a dynamic loudness compensation effect is signaled through an existing effect flag (e.g., a bit of a drcSetEffect bitfield), DRC Set 1 could have a drcSetEffect parameter in which the bit corresponding to the dynamic loudness compensation effect is set to 0, while DRC Set 2 could have the bit corresponding to the dynamic loudness compensation effect set to 1, and all other
- 15 bits of the bitfield may be the same for both DRC Sets. Through such signaling, when a user requests a dynamic loudness compensation effect through the interface, a decoder will extract and apply the gains of DRC Set 1 and DRC Set 2 in order to apply a combination of dynamic range control and dynamic loudness compensation.

- The following table shows exemplary values of parameters accomplishing the above signaling of
- 20 efficient transmission of dynamic loudness compensation data.

<i>Parameter</i>	<i>DRC Set 1</i>	<i>DRC Set 2</i>
<i>drcSetID (Example)</i>	0x01	0x02
<i>dependsOnDrcSet</i>	0x02	Not set
<i>noIndependentUse</i>	0	1
<i>drcSetEffect</i>	"Dynamic compensation flag" set to 0, otherwise same.	"Dynamic compensation flag" set to 1, otherwise same.

Table 27: Exemplary Values of parameters for signaling efficient transmission of dynamic loudness compensation data

- In the second case, where the presence of a DRC Set which provides a dynamic loudness compensation effect is signaled using a new payload, DRC Set 2 would be signaled through the
- 25 new payload (e.g., the drcSetId parameter of the dynLoudInstructions payload would be set to 2). DRC Set 1 would be signaled as described above (e.g., a drcSetId parameter would be set to 1, and a dependsonDrcSet parameter would be set to 2, indicating that DRC Set 2 depends on DRC Set 1). Finally, it would not be possible to indicate through the new payload described above

(e.g., dynLoudInstructions) that DRC Set 2 should not be used independently (e.g., that it should always be used with DRC Set 1). As such, an additional parameter could be added to the new payload (e.g., a noIndependentUse parameter could be added to the dynLoudInstructions payload), such that, when the new parameter is set to 1, DRC Set 2 is only used with DRC Set 1, while when the new parameter is set to 0, DRC Set 2 may be used independently of DRC Set 1. Similarly to the first case, when a user requests a dynamic loudness compensation effect through the interface, a decoder will extract and apply the gains of both DRC Set 1 and DRC Set 2 in order to apply a combination of dynamic range control and dynamic loudness compensation.

Finally, in the third case, existing configuration elements are used which already include all of the necessary parameters for efficiently enabling dynamic loudness compensation as described above. Specifically, DRC Set 1 would include an indication that DRC Set 2 depends on DRC Set 1 as described above, and DRC Set 2 would include an indication of whether or not it may be used independently, or only in conjunction with DRC Set 2. Again, similarly to the first two cases, when a user requests a dynamic loudness compensation effect through the interface, a decoder will extract and apply the gains of DRC Set 1 and DRC Set 2 in order to apply a combination of dynamic range control and dynamic loudness compensation.

As indicated above, a benefit of using dependent DRC Sets to enable dynamic loudness compensation is that doing so allows for more efficient transmission of the data required for dynamic loudness compensation, because it eliminates the need to transmit redundant dynamic range control gains in the DRC Set which contains the dynamic loudness compensation gains.

Interpretation

Unless specifically stated otherwise, as apparent from the following discussions, it is appreciated that throughout the disclosure discussions utilizing terms such as “processing”, “computing”, “determining”, “analyzing” or the like, refer to the action and/or processes of a computer or computing system, or similar electronic computing devices, that manipulate and/or transform data represented as physical, such as electronic, quantities into other data similarly represented as physical quantities.

In a similar manner, the term “processor” may refer to any device or portion of a device that processes electronic data, e.g., from registers and/or memory to transform that electronic data into other electronic data that, e.g., may be stored in registers and/or memory. A “computer” or a “computing machine” or a “computing platform” may include one or more processors.

The methodologies described herein are, in one example embodiment, performable by one or more processors that accept computer-readable (also called machine-readable) code containing a set of instructions that when executed by one or more of the processors carry out at least one of the methods described herein. Any processor capable of executing a set of instructions

5 (sequential or otherwise) that specify actions to be taken are included. Thus, one example is a typical processing system that includes one or more processors. Each processor may include one or more of a CPU, a graphics processing unit, and a programmable DSP unit. The processing system further may include a memory subsystem including main RAM and/or a static RAM, and/or ROM. A bus subsystem may be included for communicating between the components.

10 The processing system further may be a distributed processing system with processors coupled by a network. If the processing system requires a display, such a display may be included, e.g., a liquid crystal display (LCD) or a cathode ray tube (CRT) display. If manual data entry is required, the processing system also includes an input device such as one or more of an alphanumeric input unit such as a keyboard, a pointing control device such as a mouse, and so

15 forth. The processing system may also encompass a storage system such as a disk drive unit. The processing system in some configurations may include a sound output device, and a network interface device. The memory subsystem thus includes a computer-readable carrier medium that carries computer-readable code (e.g., software) including a set of instructions to cause performing, when executed by one or more processors, one or more of the methods described

20 herein. Note that when the method includes several elements, e.g., several steps, no ordering of such elements is implied, unless specifically stated. The software may reside in the hard disk, or may also reside, completely or at least partially, within the RAM and/or within the processor during execution thereof by the computer system. Thus, the memory and the processor also constitute computer-readable carrier medium carrying computer-readable code. Furthermore, a

25 computer-readable carrier medium may form, or be included in a computer program product.

In alternative example embodiments, the one or more processors operate as a standalone device or may be connected, e.g., networked to other processor(s), in a networked deployment, the one or more processors may operate in the capacity of a server or a user machine in server-user network environment, or as a peer machine in a peer-to-peer or distributed network environment.

30 The one or more processors may form a personal computer (PC), a tablet PC, a Personal Digital Assistant (PDA), a cellular telephone, a web appliance, a network router, switch or bridge, or any machine capable of executing a set of instructions (sequential or otherwise) that specify actions to be taken by that machine.

Note that the term “machine” shall also be taken to include any collection of machines that individually or jointly execute a set (or multiple sets) of instructions to perform any one or more of the methodologies discussed herein.

Thus, one example embodiment of each of the methods described herein is in the form of a
5 computer-readable carrier medium carrying a set of instructions, e.g., a computer program that is for execution on one or more processors, e.g., one or more processors that are part of web server arrangement. Thus, as will be appreciated by those skilled in the art, example embodiments of the present disclosure may be embodied as a method, an apparatus such as a special purpose apparatus, an apparatus such as a data processing system, or a computer-readable carrier
10 medium, e.g., a computer program product. The computer-readable carrier medium carries computer readable code including a set of instructions that when executed on one or more processors cause the processor or processors to implement a method. Accordingly, aspects of the present disclosure may take the form of a method, an entirely hardware example embodiment, an entirely software example embodiment or an example embodiment combining software and
15 hardware aspects. Furthermore, the present disclosure may take the form of carrier medium (e.g., a computer program product on a computer-readable storage medium) carrying computer-readable program code embodied in the medium.

The software may further be transmitted or received over a network via a network interface device. While the carrier medium is in an example embodiment a single medium, the term
20 “carrier medium” should be taken to include a single medium or multiple media (e.g., a centralized or distributed database, and/or associated caches and servers) that store the one or more sets of instructions. The term “carrier medium” shall also be taken to include any medium that is capable of storing, encoding or carrying a set of instructions for execution by one or more of the processors and that cause the one or more processors to perform any one or more of the
25 methodologies of the present disclosure. A carrier medium may take many forms, including but not limited to, non-volatile media, volatile media, and transmission media. Non-volatile media includes, for example, optical, magnetic disks, and magneto-optical disks. Volatile media includes dynamic memory, such as main memory. Transmission media includes coaxial cables, copper wire and fiber optics, including the wires that comprise a bus subsystem. Transmission
30 media may also take the form of acoustic or light waves, such as those generated during radio wave and infrared data communications. For example, the term “carrier medium” shall accordingly be taken to include, but not be limited to, solid-state memories, a computer product embodied in optical and magnetic media; a medium bearing a propagated signal detectable by at least one processor or one or more processors and representing a set of instructions that, when

executed, implement a method; and a transmission medium in a network bearing a propagated signal detectable by at least one processor of the one or more processors and representing the set of instructions.

It will be understood that the steps of methods discussed are performed in one example
5 embodiment by an appropriate processor (or processors) of a processing (e.g., computer) system executing instructions (computer-readable code) stored in storage. It will also be understood that the disclosure is not limited to any particular implementation or programming technique and that the disclosure may be implemented using any appropriate techniques for implementing the functionality described herein. The disclosure is not limited to any particular programming
10 language or operating system.

Reference throughout this disclosure to “one embodiment”, “some embodiments” or “an example embodiment” means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment of the present disclosure. Thus, appearances of the phrases “in one embodiment”, “in some embodiments” or
15 “in an example embodiment” in various places throughout this disclosure are not necessarily all referring to the same example embodiment. Furthermore, the particular features, structures or characteristics may be combined in any suitable manner, as would be apparent to one of ordinary skill in the art from this disclosure, in one or more example embodiments.

As used herein, unless otherwise specified the use of the ordinal adjectives “first”, “second”,
20 “third”, etc., to describe a common object, merely indicate that different instances of like objects are being referred to and are not intended to imply that the objects so described must be in a given sequence, either temporally, spatially, in ranking, or in any other manner.

In the claims below and the description herein, any one of the terms comprising, comprised of or which comprises is an open term that means including at least the elements/features that follow,
25 but not excluding others. Thus, the term comprising, when used in the claims, should not be interpreted as being limitative to the means or elements or steps listed thereafter. For example, the scope of the expression a device comprising A and B should not be limited to devices consisting only of elements A and B. Any one of the terms including or which includes or that includes as used herein is also an open term that also means including at least the
30 elements/features that follow the term, but not excluding others. Thus, including is synonymous with and means comprising.

It should be appreciated that in the above description of example embodiments of the disclosure, various features of the disclosure are sometimes grouped together in a single example

embodiment, Fig., or description thereof for the purpose of streamlining the disclosure and aiding in the understanding of one or more of the various inventive aspects. This method of disclosure, however, is not to be interpreted as reflecting an intention that the claims require more features than are expressly recited in each claim. Rather, as the following claims reflect, inventive aspects lie in less than all features of a single foregoing disclosed example
5 embodiment. Thus, the claims following the Description are hereby expressly incorporated into this Description, with each claim standing on its own as a separate example embodiment of this disclosure.

Furthermore, while some example embodiments described herein include some but not other
10 features included in other example embodiments, combinations of features of different example embodiments are meant to be within the scope of the disclosure, and form different example embodiments, as would be understood by those skilled in the art. For example, in the following claims, any of the claimed example embodiments can be used in any combination.

In the description provided herein, numerous specific details are set forth. However, it is
15 understood that example embodiments of the disclosure may be practiced without these specific details. In other instances, well-known methods, structures and techniques have not been shown in detail in order not to obscure an understanding of this description.

Thus, while there has been described what are believed to be the best modes of the disclosure, those skilled in the art will recognize that other and further modifications may be made thereto without departing from the spirit of the disclosure, and it is intended to claim all such changes
20 and modifications as fall within the scope of the disclosure. For example, any formulas given above are merely representative of procedures that may be used. Functionality may be added or deleted from the block diagrams and operations may be interchanged among functional blocks. Steps may be added or deleted to methods described within the scope of the present disclosure.

25 In the following, two sets of enumerated example embodiments (EEE-As and EEE-Bs), which are not claims, describe some structures, features, and functionalities of some aspects of the example embodiments disclosed herein.

EEE-A1. A method of metadata-based dynamic processing of audio data for playback, the method including processes of:

- 30 (a) receiving, by a decoder, a bitstream including audio data and metadata for dynamic loudness adjustment;
- (b) decoding, by the decoder, the audio data and the metadata to obtain decoded audio data and the metadata;

(c) determining, by the decoder, from the metadata, one or more processing parameters for dynamic loudness adjustment based on a playback condition;

(d) applying the determined one or more processing parameters to the decoded audio data to obtain processed audio data; and

5 (e) outputting the processed audio data for playback.

EEE-A2. The method according to EEE-A1, wherein the metadata is indicative of processing parameters for dynamic loudness adjustment for a plurality of playback conditions.

EEE-A3. The method according to EEE-A1 or EEE-A2, wherein said determining the one or more processing parameters further includes determining one or more processing parameters for
10 dynamic range compression, DRC, based on the playback condition

EEE-A4. The method according to any one of EEE-A1 to EEE-A3, wherein the playback condition includes one or more of a device type of the decoder, characteristics of a playback device, characteristics of a loudspeaker, a loudspeaker setup, characteristics of background noise, characteristics of ambient noise and characteristics of the acoustic environment.

15 EEE-A5. The method according to any one of EEE-A1 to EEE-A4, wherein process (c) further includes selecting, by the decoder, at least one of a set of DRC sequences, DRCSet, a set of equalizer parameters, EQSet, and a downmix, corresponding to the playback condition.

EEE-A6. The method according to EEE-A5, wherein process (c) further includes identifying a metadata identifier indicative of the at least one selected DRCSet, EQSet, and downmix to
20 determine the one or more processing parameters from the metadata.

EEE-A7. The method according to any one of EEE-A1 to EEE-A6, wherein the metadata includes one or more processing parameters relating to average loudness values and optionally one or more processing parameters relating to dynamic range compression characteristics.

EEE-A8. The method according to any one of EEE-A1 to EEE-A7, wherein the bitstream further
25 includes additional metadata for static loudness adjustment to be applied to the decoded audio data.

EEE-A9. The method according to any one of EEE-A1 to EEE-A8, wherein the bitstream is an MPEG-D DRC bitstream and the presence of metadata is signaled based on MPEG-D DRC bitstream syntax.

30 EEE-A10. The method according to EEE-A9, wherein a loudnessInfoSetExtension()-element is used to carry the metadata as a payload.

EEE-A11. The method according to any one of EEE-A1 to EEE-A10, wherein the metadata comprises one or more metadata payloads, wherein each metadata payload includes a plurality of sets of parameters and identifiers, with each set including at least one of a DRCSet identifier, drcSetId, an EQSet identifier, eqSetId, and a downmix identifier, downmixId, in combination
5 with one or more processing parameters relating to the identifiers in the set.

EEE-A12. The method according to EEE-A11 when depending on EEE-A5, wherein process (c) involves selecting a set among the plurality of sets in the payload based on the at least one DRCSet, EQSet, and downmix selected by the decoder, and wherein the one or more processing parameters determined at process (c) are the one or more processing parameters relating to the
10 identifiers in the selected set.

EEE-A13. A decoder for metadata-based dynamic processing of audio data for playback, wherein the decoder comprises one or more processors and non-transitory memory configured to perform a method including processes of:

- 15 (a) receiving, by a decoder, a bitstream including audio data and metadata for dynamic loudness adjustment;
- (b) decoding, by the decoder, the audio data and the metadata to obtain decoded audio data and the metadata;
- (c) determining, by the decoder, from the metadata, one or more processing parameters for dynamic loudness adjustment based on a playback condition;
- 20 (d) applying the determined one or more processing parameters to the decoded audio data to obtain processed audio data; and
- (e) outputting the processed audio data for playback.

EEE-A14. A method of encoding audio data and metadata for dynamic loudness adjustment, into a bitstream, the method including processes of:

- 25 (a) inputting original audio data into a loudness leveler for loudness processing to obtain, as an output from the loudness leveler, loudness processed audio data;
- (b) generating metadata for dynamic loudness adjustment based on the loudness processed audio data and the original audio data; and
- (c) encoding the original audio data and the metadata into the bitstream.

30 EEE-A15. The method according to EEE-A14, wherein the method further includes generating additional metadata for static loudness adjustment to be used by a decoder.

EEE-A16. The method according to EEE-A14 or EEE-A15, wherein process (b) includes comparison of the loudness processed audio data to the original audio data, and wherein the metadata is generated based on a result of said comparison.

5 EEE-A17. The method according to EEE-A16, wherein process (b) further includes measuring the loudness over one or more pre-defined time periods, and wherein the metadata is generated further based on the measured loudness.

EEE-A18. The method according to EEE-A17, wherein the measuring comprises measuring overall loudness of the audio data.

10 EEE-A19. The method according to EEE-A17, wherein the measuring comprises measuring loudness of dialogue in the audio data.

EEE-A20. The method according to any one of EEE-A14 to EEE-A19, wherein the bitstream is an MPEG-D DRC bitstream and the presence of the metadata is signaled based on MPEG-D DRC bitstream syntax.

15 EEE-A21. The method according to EEE-A20, wherein a loudnessInfoSetExtension()-element is used to carry the metadata as a payload.

EEE-A22. The method according to any one of EEE-A14 to EEE-A21, wherein the metadata comprises one or more metadata payloads, wherein each metadata payload includes a plurality of sets of parameters and identifiers, with each set including at least one of a DRCSet identifier, drcSetId, an EQSet identifier, eqSetId, and a downmix identifier, downmixId, in combination
20 with one or more processing parameters relating to the identifiers in the set, and wherein the one or more processing parameters are parameters for dynamic loudness adjustment by a decoder.

EEE-A23. The method according to EEE-A22, wherein the at least one of the drcSetId, the eqSetId, and the downmixId is related to at least one of a set of DRC sequences, DRCSet, a set of equalizer parameters, EQSet, and a downmix, to be selected by the decoder.

25 EEE-A24. An encoder for encoding in a bitstream original audio data and metadata for dynamic loudness adjustment, wherein the encoder comprises one or more processors and non-transitory memory configured to perform a method including the processes of:

- (a) inputting original audio data into a loudness leveler for loudness processing to obtain, as an output from the loudness leveler, loudness processed audio data;
- 30 (b) generating the metadata for dynamic loudness adjustment based on the loudness processed audio data and the original audio data; and
- (c) encoding the original audio data and the metadata into the bitstream.

EEE-A25. A system of an encoder for encoding in a bitstream original audio data and metadata for dynamic loudness adjustment and/or dynamic range compression, DRC, according to EEE-A24 and a decoder for metadata-based dynamic processing of audio data for playback according to EEE-A13.

- 5 EEE-A26. A computer program product comprising a computer-readable storage medium with instructions adapted to cause the device to carry out the method according to any one of EEE-A1 to EEE-A12 or EEE-A14 to EEE-A23 when executed by a device having processing capability.

EEE-A27. A computer-readable storage medium storing the computer program product of EEE-A26.

- 10 EEE-A28. The method according to any one of EEE-A1 to EEE-A12 further comprising receiving, by the decoder, through an interface, an indication of whether or not to perform the metadata-based dynamic processing of audio data for playback, and when the decoder receives an indication not to perform the metadata-based dynamic processing of audio data for playback, bypassing at least the step of applying the determined one or more processing parameters to the
15 decoded audio data.

EEE-A29. The method according to EEE-A28, wherein until the decoder receives, through the interface, the indication of whether or not to perform the metadata-based dynamic processing of audio data for playback, the decoder bypasses at least the step of applying the determined one or more processing parameters to the decoded audio data.

- 20 EEE-A30. The method of any one of EEE-A1 to EEE-A12, EEE-A28, or EEE-A29, wherein the metadata is indicative of processing parameters for dynamic loudness adjustment for a plurality of playback conditions, and the metadata further includes a parameter specifying a loudness measurement method used for deriving a processing parameter of the plurality of processing parameters.

- 25 EEE-A31. The method of any one of EEE-A1 to EEE-A12, or EEE-A28 to EEE-A30, wherein the metadata is indicative of processing parameters for dynamic loudness adjustment for a plurality of playback conditions, and the metadata further includes a parameter specifying a loudness measurement system used for measuring a processing parameter of the plurality of processing parameters.

- 30 EEE-B1. A method of metadata-based dynamic processing of audio data for playback, the method including:

receiving, by a decoder, a bitstream including audio data and metadata for dynamic loudness adjustment, wherein the metadata for dynamic loudness adjustment comprises a plurality of sets of metadata, wherein each set of metadata corresponds to a respective playback condition; decoding, by the decoder, the audio data and the metadata to obtain decoded audio data and the metadata;

5 selecting, in response to playback condition information provided to the decoder, a set of metadata corresponding to a specific playback condition, and extracting, from the selected set of metadata, one or more processing parameters for dynamic loudness adjustment;

10 applying the extracted one or more processing parameters to the decoded audio data to obtain processed audio data; and

outputting the processed audio data for playback.

EEE-B2. The method according to EEE-B1, wherein said extracting the one or more processing parameters further includes extracting one or more processing parameters for dynamic range compression, DRC.

15 EEE-B3. The method according to EEE-B1 or EEE-B2, wherein the playback condition information is indicative of a specific loudspeaker setup.

EEE-B4. The method according to any one of EEE-B1 to EEE-B3, wherein the selected set of metadata includes a set of DRC sequences, DRCSet.

20 EEE-B5. The method according to any of EEE-B1 to EEE-B4, wherein selecting the set of metadata includes identifying a set of metadata corresponding to a specific downmix.

EEE-B6. The method according to any one of EEE-B1 to EEE-B5, wherein the sets of metadata each include one or more processing parameters relating to average loudness values and optionally one or more processing parameters relating to dynamic range compression characteristics.

25 EEE-B7. The method according to any one of EEE-B1 to EEE-B6, wherein the bitstream further includes additional metadata for static loudness adjustment to be applied to the decoded audio data.

EEE-B8. The method according to any one of EEE-B1 to EEE-B7, wherein the bitstream is an MPEG-D DRC bitstream and the presence of metadata is signaled based on MPEG-D DRC

30 bitstream syntax.

EEE-B9. The method according to EEE-B8, wherein a loudnessInfoSetExtension()-element is used to carry the metadata as a payload.

EEE-B10. The method according to any one of EEE-B1 to EEE-B9, wherein the metadata comprises one or more metadata payloads, wherein each metadata payload includes a plurality of sets of parameters and identifiers, with each set including a respective downmix identifier, downmixId, in combination with one or more processing parameters relating to the downmix identifier in the set.

EEE-B11. A decoder for metadata-based dynamic processing of audio data for playback, wherein the decoder comprises one or more processors and non-transitory memory configured to perform a method including:

receiving, by the decoder, a bitstream including audio data and metadata for dynamic loudness adjustment, wherein the metadata for dynamic loudness adjustment comprises a plurality of sets of metadata, wherein each set of metadata corresponds to a respective playback condition; decoding, by the decoder, the audio data and the metadata to obtain decoded audio data and the metadata;

selecting, in response to a playback condition provided to the decoder a set of metadata corresponding to a specific playback condition, and extracting, from the selected set of metadata, one or more processing parameters for dynamic loudness adjustment;

applying the extracted one or more processing parameters to the decoded audio data to obtain processed audio data; and

outputting the processed audio data for playback.

EEE-B12. A method of encoding audio data and metadata for dynamic loudness adjustment into a bitstream, the method including:

inputting original audio data into a loudness leveler for loudness processing to obtain, as an output from the loudness leveler, loudness processed audio data;

generating the metadata for dynamic loudness adjustment based on the loudness processed audio data and the original audio data; and

encoding the original audio data and the metadata into the bitstream.

EEE-B13. The method according to EEE-B12, wherein the method further includes generating additional metadata for static loudness adjustment to be used by a decoder.

EEE-B14. The method according to EEE-B12 or EEE-B13, wherein said generating metadata includes comparison of the loudness processed audio data to the original audio data, and wherein the metadata is generated based on a result of said comparison.

EEE-B15. The method according to EEE-B14, wherein said generating metadata further includes measuring the loudness over one or more pre-defined time periods, and wherein the metadata is generated further based on the measured loudness.

5 EEE-B16. The method according to EEE-B15, wherein the measuring comprises measuring overall loudness of the audio data.

EEE-B17. The method according to EEE-B15, wherein the measuring comprises measuring loudness of dialogue in the audio data.

10 EEE-B18. The method according to any one of EEE-B12 to EEE-B17, wherein the bitstream is an MPEG-D DRC bitstream and the presence of the metadata is signaled based on MPEG-D DRC bitstream syntax.

EEE-B19. The method according to EEE-B18, wherein a loudnessInfoSetExtension()-element is used to carry the metadata as a payload.

15 EEE-B20. The method according to any one of EEE-B12 to EEE-B19, wherein the metadata comprises a plurality of sets of metadata, wherein each set of metadata corresponds to a respective playback condition.

EEE-B21. The method according to any one of EEE-B12 to EEE-B20, wherein the metadata comprises one or more metadata payloads, wherein each metadata payload includes a plurality of sets of parameters and identifiers, with each set including a respective downmix identifier, downmixId, in combination with one or more processing parameters relating to the downmix
20 identifier in the set, and wherein the one or more processing parameters are parameters for dynamic loudness adjustment by a decoder.

EEE-B22. An encoder for encoding in a bitstream original audio data and metadata for dynamic loudness adjustment, wherein the encoder comprises one or more processors and non-transitory memory configured to perform a method including:
25 inputting original audio data into a loudness leveler for loudness processing to obtain, as an output from the loudness leveler, loudness processed audio data;
generating the metadata for dynamic loudness adjustment based on the loudness processed audio data and the original audio data; and
encoding the original audio data and the metadata into the bitstream.

30 EEE-B23. A system of an encoder for encoding in a bitstream original audio data and metadata for dynamic loudness adjustment, according to EEE-B22 and a decoder for metadata-based dynamic processing of audio data for playback according to EEE-B11.

EEE-B24. A computer program product comprising a computer-readable storage medium with instructions adapted to cause the device to carry out the method according to any one of EEE-B1 to EEE-B10 or EEE-B12 to EEE-B21 when executed by a device having processing capability.

EEE-B25. A computer-readable storage medium storing the computer program product of EEE-B24.

5

CLAIMS

1. A method of processing audio data for playback, the method including:
 - receiving, by a decoder, a bitstream including encoded audio data and metadata, wherein the metadata includes one or more dynamic range control (DRC) sets, and for each DRC set, an indication of whether the DRC set is configured for providing a dynamic loudness compensation effect;
 - 5 parsing the metadata, by the decoder, to identify DRC sets that are configured for providing the dynamic loudness compensation effect;
 - decoding, by the decoder, the encoded audio data to obtain decoded audio data;
 - 10 selecting, by the decoder, one of the identified DRC sets configured for providing the dynamic loudness compensation effect;
 - extracting from the bitstream, by the decoder, one or more DRC gains corresponding to the selected DRC set;
 - applying to the decoded audio data, by the decoder, the one or more DRC gains
 - 15 corresponding to the selected DRC set to obtain dynamic loudness compensated audio data; and
 - outputting the dynamic loudness compensated audio data for playback.
2. The method according to claim 1, wherein the metadata includes a plurality of DRC sets configured for providing the dynamic loudness adjustment, wherein each of the plurality of DRC sets is also associated with one or more playback conditions, and wherein the selecting is performed in response to an indication of a playback condition provided to the decoder.
- 20 3. The method according to claim 1 or claim 2, wherein, in addition to providing a dynamic loudness compensation effect, the one or more DRC sets are also configured to provide dynamic range control.
- 25 4. The method according to claim 2 or claim 3, wherein the playback condition includes one or more of a device type of the decoder, characteristics of a playback device, characteristics of a loudspeaker, a loudspeaker setup, characteristics of background noise, characteristics of ambient noise and characteristics of the acoustic environment.
- 30 5. The method according to any one of claims 1 to 4, wherein the indication of whether the DRC set is configured for providing the dynamic loudness compensation effect is provided in a parameter indicating one or more effects provided by the DRC set.

6. The method according to claim 5, wherein the parameter indicating one or more effects provided by the DRC set is a drcSetEffect bitfield of an MPEG-D DRC bitstream, wherein individual bits of the drcSetEffect bitfield correspond to different effects, and one of the bits of the drcSetEffect bitfield corresponds to the dynamic loudness compensation effect.

7. The method according to any one of claims 1 to 4, wherein the indication of whether the DRC set is configured for providing the dynamic loudness compensation effect is whether the DRC set is specified in a dynamic loudness compensation bitstream payload.

8. The method according to claim 7, wherein the dynamic loudness compensation bitstream payload is included in an extension field of a previously defined bitstream syntax.

9. The method according to claim 8, wherein the extension field is a uniDrcConfigExtension field of an MPEG-D DRC bitstream, and wherein the dynamic loudness compensation bitstream payload is included only for specific values of a uniDrcConfigExtType parameter.

10. The method according to claim 8 or claim 9, wherein a plurality of dynamic loudness compensation payloads specifying a plurality of DRC sets configured for providing the dynamic loudness compensation effect are included in the extension field of the previously defined bitstream syntax.

11. The method of any one of claims 1 to 4, wherein the indication of whether the DRC set is configured for providing the dynamic loudness compensation effect is a field of a previously existing configuration element of a previously defined bitstream syntax.

12. The method of claim 11, wherein the field is a dynamicLoudCompDRCSet parameter, and the previously existing configuration element is a downmixInstructions element, a drcInstructionsBasic element, or a drcInstructionsUniDRC element of an MPEG-D DRC bitstream.

13. The method of claim 12, wherein the field is a previously existing field reserved for future use.

14. The method of any one of claims 1 to 4, wherein the indication of whether the DRC set is configured for providing the dynamic loudness compensation effect is a field of an updated version of a previously existing configuration element of a previously defined bitstream syntax.

5 15. The method of claim 14, wherein the field is a dynamicLoudCompDRCSet parameter, and the updated version of the previously existing configuration element is a downmixInstructionsV2 element or a drcInstructionsUniDrcV2 element.

10 16. The method of any one of claims 1 to 15, wherein an indication that a dynamic loudness compensation effect is desired is provided to the decoder through an interface, and wherein the DRC set is selected in response to the indication provided to the decoder through the interface.

15 17. The method of claim 16, wherein indications of additional desired effects are provided to the decoder through the interface, wherein the metadata includes a plurality of DRC sets configured to provide the dynamic loudness compensation effect, and wherein the selection depends on the additional desired effects.

20 18. The method of claim 15 or 16, wherein the indication that a dynamic loudness compensation effect is desired is provided through a drcEffectTypeRequest parameter of a dynamicRangeControllerInterface payload.

25 19. The method of any one of claims 1 to 18, wherein the metadata includes one or more static loudness values configured for providing static loudness adjustment to the decoded audio data.

30 20. The method of claim 19, comprising applying static loudness adjustment, in response to one or more of the static loudness values, to the decoded audio data or the dynamic loudness compensated audio data.

35 21. The method of claim 1, wherein a first one of the DRC sets is configured for providing dynamic range control, wherein the first DRC set comprises an indication that the selected DRC set configured for providing the dynamic loudness compensation effect is configured for application in combination with the first DRC set.

22. The method of claim 21, wherein the selected DRC set comprises an indication of whether the one or more DRC gains corresponding to the selected DRC set may only be applied in combination with DRC gains corresponding to the first DRC set.

5 23. The method of claim 21 or claim 22, further comprising extracting from the bitstream the DRC gains corresponding to the first DRC set, and applying the DRC gains corresponding to the first DRC set to the decoded audio data.

 24. A decoder for processing audio data for playback, wherein the decoder comprises
10 one or more processors and non-transitory memory configured to perform the method of any one of claims 1 to 23.

 25. A computer program product comprising a computer-readable storage medium with instructions adapted to cause a device to carry out the method according to any one of claims 1
15 to 23.

 26. A computer-readable storage medium storing the computer program product of claim 25.