Claims

1. An audio encoder for generating a downmix signal from an original signal, the original signal having a plurality of original channels, the downmix signal having a plural number of downmix channels, the audio encoder comprising:

a parameter estimator configured for estimating channel level and correlation information of the original signal, and

a bitstream writer for encoding the downmix signal into a bitstream, so that the downmix signal is encoded in the bitstream so as to have side information including the channel level and correlation information of the original signal,

wherein the channel level and correlation information of the original signal includes at least one interchannel level difference, ICLD,

wherein the channel level and correlation information of the original signal encoded in the side information includes at least correlation information describing energy relationships between at least one couple of different original channels, but less than the totality of the original channels,

wherein the ICLD is

$$\chi_i = 10 \cdot log_{10}\left(\frac{P_i}{P_{\mathrm{dmx,i}}}\right)$$

where

- χ_i is the ICLD for channel i
- P_i is the power of the channel i
- $P_{\mathrm{dmx,i}}$ is a linear combination of the values of the covariance information of the downmix signal, and

wherein the channel level and correlation information of the original signal includes at least one coherence value describing the coherence between two channels of a couple of original channels, wherein the coherence value is

$$\xi_{i,j} = \frac{C_{y_{i,j}}}{\sqrt{C_{y_{i,i}} \cdot C_{y_{j,j}}}}$$

where $C_{y_{i,j}}$ is a covariance between the channels i and j, $C_{y_{i,i}}$ and $C_{y_{j,j}}$ being respectively levels associated to the channels i and j.

2. The audio encoder of claim 1, configured to provide the channel level and correlation information as including interchannel coherences between at least one of the following channel combinations:

- a center channel and a right channel
- the center channel and a left channel
- the left channel and a left surround channel
- the right channel and a right surround channel
- the center channel and a low frequency effect (LFE) channel.

3. The audio encoder of claims 1-2, configured to choose whether to encode or not to encode at least part of the channel level and correlation information of the original signal on the basis of status information, so as to include, in the side information, an increased quantity of channel level and correlation information in case of comparatively lower payload.

4. The audio encoder of any of claims 1-3, configured to choose which part of the channel level and correlation information of the original signal is to be encoded in the side information on the basis of metrics on the channels, so as to include, in the side information, channel level and correlation information associated to more sensitive metrics.

5. The audio encoder of any of claims 1-4, wherein the channel level and correlation information of the original signal is in the form of entries of a matrix.

6. The audio encoder of claim 5, wherein the matrix is symmetrical or Hermitian, wherein the entries of the channel level and correlation information are provided for all or less than the totality of the entries in the diagonal of the matrix and/or for less than the half of the non-diagonal elements of the matrix.

7. The audio encoder of any of claims 1-6, wherein the bitstream writer is configured to encode identification of at least one channel.

8. The audio encoder of any of claims 1-7, wherein the original signal, or a processed version thereof, is divided into a plurality of subsequent frames of equal time length.

9. The audio encoder of claim 8, configured to encode in the side information channel level and correlation information of the original signal specific for each frame.

10. The audio encoder of claim 9, configured to encode, in the side information, the same channel level and correlation information of the original signal collectively associated to a plurality of consecutive frames.

11. The audio encoder of any of claims 9-10, configured to choose the number of consecutive frames to which the same channel level and correlation information of the original signal has been encoded so that:

a comparatively higher bitrate or higher payload implies an increase of the number of consecutive frames to which the same channel level and correlation information of the original signal is associated, and vice versa.

12. The audio encoder of any of claims 10-11, configured to reduce the number of consecutive frames to which the same channel level and correlation information of the original signal is associated at the detection of a transient.

13. The audio encoder of any of claims 8-12, wherein each frame is subdivided into an integer number of consecutive slots.

14. The audio encoder of claim 13, configured to estimate the channel level and correlation information for each slot and to encode in the side information the sum or average or another predetermined linear combination of the channel level and correlation information estimated for different slots,

wherein the audio encoder is configured to perform a transient analysis onto the time domain version of the frame to determine the occurrence of a transient within the frame.

15. The audio decoder of claim 14, configured to determine in which slot of the frame the transient has occurred, and:

to encode the channel level and correlation information of the original signal associated to the slot in which the transient has occurred and/or to the subsequent slots in the frame,

without encoding channel level and correlation information of the original signal associated to the slots preceding the transient.

16. The audio encoder of claim 14 or 15, configured to signal, in the side information, the occurrence of the transient being occurred in one slot of the frame.

17. The audio encoder of claim 16, configured to signal, in the side information, in which slot of the frame the transient has occurred.

18. The audio encoder of any of claims 14-17, configured to estimate channel level and correlation information of the original signal associated to multiple slots of the frame, and to sum them or average them or linearly combine them to obtain channel level and correlation information associated to the frame.

19. The audio encoder of any of claims 1-18, wherein the original signal is converted into a frequency domain signal, wherein the audio encoder is configured to encode, in the side information, the channel level and correlation information of the original signal in a band-by-band fashion,

wherein the audio encoder is configured to aggregate a number of bands of the original signal into a more reduced number of bands, so as to encode, in the side information, the channel level and correlation information of the original signal in an aggregated-band-by-aggregated-band fashion.

20. The audio encoder of claim 19, configured, in case of detection of a transient in the frame, to further aggregate the bands so that:

the number of the bands is reduced; and/or

the width of at least one band is increased by aggregation with another band.

21. The audio encoder of any of claims 19-20, further configured to encode, in the bitstream, at least one channel level and correlation information of one band as an increment in respect to a previously encoded channel level and correlation information.

22. The audio encoder of any of claims 1-21, configured to encode, in the side information of the bitstream, an incomplete version of the channel level and correlation information with respect to the channel level and correlation information estimated by the estimator.

23. The audio encoder of claim 22, configured to adaptively select, among the whole channel level and correlation information estimated by the estimator, selected information to be encoded

in the side information of the bitstream, so that remaining non-selected information channel level and/or correlation information estimated by the estimator is not encoded.

24. The audio encoder of claim 22, configured to reconstruct channel level and correlation information from the selected channel level and correlation information, thereby simulating the estimation, at the decoder, of non-selected channel level and correlation information, and to calculate error information between:

the non-selected channel level and correlation information as estimated by the encoder; and

the non-selected channel level and correlation information as reconstructed by simulating the estimation, at the decoder, of non-encoded channel level and correlation information; and

so as to distinguish, on the basis of the calculated error information:

properly-reconstructible channel level and correlation information; from

non-properly-reconstructible channel level and correlation information,

so as to decide for:

the selection of the non-properly-reconstructible channel level and correlation information to be encoded in the side information of the bitstream; and

the non-selection of the properly-reconstructible channel level and correlation information, thereby refraining from encoding in the side information of the bitstream the properly-reconstructible channel level and correlation information.

25. The audio encoder of any of claims 23-24, wherein the channel level and correlation information is indexed according to a predetermined ordering, wherein the encoder is configured to signal, in the side information of the bitstream, indexes associated to the predetermined ordering, the indexes indicating which of the channel level and correlation information is encoded.

26. The audio encoder of claim 25, wherein the indexes are provided through a bitmap.

27. The audio encoder of any of claims 25-26, wherein the indexes are defined according to a combinatorial number system associating a one-dimensional index to entries of a matrix.

28. The audio encoder of any of claims 25-27, configured to perform a selection among:

an adaptive provision of the channel level and correlation information, in which indexes associated to the predetermined ordering are encoded in the side information of the bitstream; and

a fixed provision of the channel level and correlation information, so that the channel level and correlation information which is encoded is predetermined, and ordered according to a predetermined fixed ordering, without the provision of indexes.

29. The audio encoder of claim 28, configured to signal, in the side information of the bitstream, whether channel level and correlation information is provided according to an adaptive provision or according to the fixed provision.

30. The audio encoder of any of claims 1-29, further configured to encode, in the bitstream, current channel level and correlation information as increment in respect to previous channel level and correlation information.

31. The audio encoder of any of claims 1-30, further configured to generate the downmix signal according to a static downmixing.

32. The audio encoder of any of claims 1-31, wherein the audio encoder is agnostic to the audio synthesizer.

33. A method for generating a downmix signal from an original signal, the original signal having a number of original channels, the downmix signal having a number of downmix channels, the method comprising:

estimating channel level and correlation information of the original signal, wherein the channel level and correlation information of the original signal includes at least one interchannel level difference, ICLD, wherein the channel level and correlation information of the original signal encoded in the side information further includes at least correlation information describing energy relationships between at least one couple of different original channels, but less than the totality of the original channels, and

encoding the downmix signal into a bitstream, so that the downmix signal is encoded in the bitstream so as to have side information including channel level and correlation information of the original signal,

wherein the ICLD is

$$\chi_i = 10 \cdot log_{10}\left(\frac{P_i}{P_{\mathrm{dmx,i}}}\right)$$

where

- χ_i is the ICLD for channel i
- P_i is the power of the channel i
- $P_{\mathrm{dmx,i}}$ is a linear combination of the values of the covariance information of the downmix signal,

wherein the channel level and correlation information of the original signal includes at least one coherence value describing the coherence between two channels of a couple of original channels, and

wherein the coherence value is

$$\xi_{i,j} = \frac{C_{y_{i,j}}}{\sqrt{C_{y_{i,i}} \cdot C_{y_{j,j}}}}$$

where $C_{y_{i,j}}$ is a covariance between the channels i and j, $C_{y_{i,i}}$ and $C_{y_{j,j}}$ being respectively levels associated to the channels i and j.

34. A non-transitory storage unit storing instructions which, when executed by a processor, cause the processor to perform a method according to claim 33.