Claims

1. Multisignal encoder for encoding at least three audio signals, comprising:

a signal preprocessor for individually preprocessing each audio signal to obtain at least three preprocessed audio signals, wherein the preprocessing is performed so that a preprocessed audio signal is whitened with respect to the signal before preprocessing;

an adaptive joint signal processor for performing a processing of the at least three preprocessed audio signals subsequent to an energy level normalization to obtain at least three jointly processed signals or at least two jointly processed signals and an unprocessed signal, wherein, for performing the energy level normalization, the adaptive joint signal processor is configured to:

calculate an information on a mean energy of the preprocessed audio signals;

calculate an information on an energy of each preprocessed audio signal,

to calculate the broadband energy normalization value for a specific preprocessed audio signal based on the information on the mean energy and the information on the energy of the specific preprocessed audio signal, and

to perform a broadband energy normalization of the at least three preprocessed audio signals using the broadband energy normalization values for the specific preprocessed audio signals, wherein each preprocessed audio signal has a normalized energy;

a signal encoder for encoding each signal to obtain one or more encoded signals; and

an output interface for transmitting or storing an encoded multisignal audio signal comprising the one or more encoded signals, side information relating to the preprocessing and side information relating to the processing comprising the broadband energy normalization values for the preprocessed audio signals.

2. Multisignal encoder of claim 1,

wherein the adaptive joint signal processor is configured to calculate a scaling ratio as the broadband energy normalization value for the specific preprocessed audio signal from the mean energy and the energy of the specific preprocessed audio signal, and

wherein the adaptive joint signal processor is configured for determining a flag indicating whether the scaling ratio is for an upscaling or a downscaling, and wherein the flag for each preprocessed audio signal is included in the side information relating to the processing.

3. Multisignal encoder of claim 2,

wherein the adaptive joint signal processor is configured to quantize the scaling ratio into the same quantization range irrespective of whether the scaling is an upscaling or a downscaling.

4. Multisignal encoder of one of the preceding claims, wherein the adaptive joint signal processor is configured:

to calculate cross-correlation values for each possible signal pair of the preprocessed signals having the normalized energies of the at least three preprocessed signals;

to select the signal pair having the highest cross-correlation value;

to determine a joint stereo processing mode for the selected signal pair; and

to joint stereo process the selected signal pair in accordance with the determined joint stereo processing mode to obtain a processed signal pair.

5. Multisignal encoder of claim 4, wherein the adaptive joint signal processor is configured to apply a cascaded signal pair preprocessing, or wherein the adaptive joint signal processor is configured to apply a non-cascaded signal pair processing,

wherein, in the cascaded signal pair preprocessing, the signals of a processed signal pair are selectable in a further iteration step consisting of a calculation of updated cross-correlation values, selecting the signal pair having the highest cross-correlation value,

the determination of a joint stereo processing mode for the selected signal pair and the joint stereo processing the selected signal pair in accordance with the determined joint stereo processing mode, or

wherein, in the non-cascaded signal pair processing, the signals of a processed signal pair are not selectable in an additional selecting the signal pair having the highest cross-correlation value, the determination of a joint stereo processing mode for the selected signal pair, and the joint stereo processing the selected signal pair in accordance with the determined joint stereo processing mode.

6. Multisignal encoder in accordance with one of the preceding claims,

wherein the adaptive joint signal processor is configured to determine the signal to be encoded individually as a signal remaining subsequent to a pairwise processing procedure, and

wherein the adaptive joint signal processor is configured to modify an energy normalization applied to the signal before performing the pairwise processing procedure such as reverting, or at least partly reverting the energy normalization applied to the signal before performing the pairwise processing procedure.

7. Multisignal encoder of one of the preceding claims,

wherein the adaptive joint signal processor is configured to determine, for each signal to be processed by the signal encoder, a bit distribution information, wherein the output interface is configured to introduce the bit distribution information, for each signal, into the side information relating to the processing.

8. Multisignal encoder of one of the preceding claims,

wherein the adaptive joint signal processor is configured for calculating a signal energy information of each signal to be processed by the signal encoder,

to calculate a total energy of the plurality of signals to be encoded by the signal encoder;

to calculate a bit distribution information for each signal based on the signal energy information and the total energy information, and

wherein the output interface is configured to introduce the bit distribution information, for each signal, into the side information relating to the processing.

9. Multisignal encoder of claim 8,

wherein the adaptive joint signal processor is configured to assign an initial number of bits to each signal, to assign a number of bits based on the bit distribution information, to perform a further refinement step, or to perform a final donation step, and

wherein the signal encoder is configured to perform the encoding of each signal using the assigned bits per signal.

10. Multisignal encoder of one of the preceding claims, wherein the signal preprocessor is configured to perform, for each audio signal:

a time-to-spectral conversion operation to obtain a signal spectrum for each audio signal;

a temporal noise shaping operation and/or a frequency domain noise shaping operation for each signal spectrum, and

wherein the signal preprocessor is configured to feed the signal spectra into the adaptive joint signal processor subsequent to the temporal noise shaping operation and/or the frequency domain noise shaping operation, and

wherein the adaptive joint signal processor is configured to perform the joint signal processing on the received signal spectra.

11. Multisignal encoder of one of the preceding claims, wherein the adaptive joint signal processor is configured

to determine, for each signal of a selected signal pair, a necessary bit rate for a full band separate encoding mode such as L/R or a necessary bit rate for a full band joint encoding mode such as M/S, or a bit rate for a bandwise joint encoding mode such as M/S plus necessary bits for bandwise signaling such as the M/S mask,

to determine the separate encoding mode or the joint encoding mode as the specific mode for all bands of a signal pair, when a majority of bands have been determined for the specific mode, and the minority of the bands being less than 10% of all bands have been determined to the other encoding mode, or to determine the coding mode requiring the lowest amount of bits, and

wherein the output interface is configured to include an indication into the side information relating to the processing, the indication indicating the specific mode for all bands of a frame instead of a coding mode mask for the frame.

12. Multisignal encoder of one of the preceding claims,

wherein the signal encoder comprises a rate loop processor for each individual signal or across two or more signals, the rate loop processor being configured for receiving and using a bit distribution information for the specific signal or for two or more signals.

13. Multisignal encoder of one of the preceding claims,

wherein the adaptive joint signal processor is configured to adaptively select signal pairs for joint coding, or wherein the adaptive joint signal processor is configured for determining, for each selected signal pair, a bandwise mid/side encoding mode, a full band mid/side encoding mode or a full band left/right encoding mode, and wherein the output interface is configured for including into the side information relating to the processing, an information indicating the selected coding mode in the encoded multisignal audio signal.

14. Multisignal encoder of one of the preceding claims,

wherein the adaptive joint signal processor is configured for forming a bandwise mid/side decision versus a left/right decision based on an estimated bitrate in each band when coded in the mid/side mode or in the left/right mode, and wherein a final joint coding mode is determined based on the results of the bandwise mid/side versus left/right decision.

15. Multisignal encoder of one of the preceding claims, wherein the adaptive joint signal processor is configured for performing a spectral band replication processing or an in-

telligent gap filling processing for determining parametric side information for the spectral band replication processing or the intelligent gap filling processing, and wherein the output interface is configured for including the spectral band replication or intelligent gap filling side information as additional side information into the side information relating to the processing.

16. Multisignal encoder of claim 15,

wherein the adaptive joint signal processor is configured for performing a stereo intelligent gap filling processing for an encoded signal pair and, additionally, for performing a single signal intelligent gap filling processing for the at least one signal to be encoded individually.

17. Multisignal encoder of one of the preceding claims,

wherein the at least three audio signals include a low frequency enhancement signal, and wherein the adaptive joint signal processor is configured to apply a signal mask, the signal mask indicating for which signals the adaptive joint signal processor is to be active, and wherein the signal mask indicates that the low frequency enhancement signal is not to be used in the pairwise processing of the at least three preprocessed audio signals.

18. Multisignal encoder of one of claims 1 to 3, wherein the adaptive joint signal processor is configured to calculate, as the information on the energy of each preprocessed signal, an energy of an MDCT spectrum of each preprocessed signal, or

to calculate, as the information on the mean energy of the at least three preprocessed audio signals, a mean energy of MDCT spectra of the at least three preprocessed audio signals.

19. Multisignal encoder of one of claims 1 to 3,

wherein the adaptive joint signal processor is configured for calculating a scaling factor for each signal based on the information on the energy for the specific signal and the information on the mean energy of the at least three preprocessed audio signals,

wherein the adaptive joint signal processor is configured for quantizing the scaling ratio to obtain a quantized scaling ratio value, the quantized scaling ratio value being used for deriving side information for the scaling ratio for each signal included into the encoded multisignal audio signal, and

wherein the adaptive joint signal processor is configured to derive, from the quantized scaling ratio value, a quantized scaling ratio, wherein the preprocessed audio signal is scaled using the quantized scaling ratio before being used for the pairwise processing of the scaled signal with another correspondingly scaled signal.

20. Multisignal encoder of one of the preceding claims,

wherein the adaptive joint signal processor is configured for calculating normalized inter-signal cross-correlation values for possible signal pairs in order to decide and select which signal pair has the highest degree of similarities and, therefore, is suitable to be selected as a pair for pairwise processing of the at least three preprocessed audio signals,

wherein the normalized cross-correlation values for each signal pair are stored in a cross-correlation vector, and

wherein the adaptive joint signal processor is configured for determining, whether a signal pair selection of one or more previous frames is to be retained or not by comparing a cross-correlation vector of the previous frame to the cross-correlation vector of the current frame, and wherein the signal pair selection of the previous frame is retained, when a difference between the cross-correlation vector of the current frame and the cross-correlation vector of the previous frame is lower than a predefined threshold.

21. Multisignal encoder of one of the preceding claims,

wherein the signal preprocessor is configured for performing a time-frequency conversion using a certain window length selected from a plurality of different window lengths,

wherein the adaptive joint signal processor is configured to determine, when comparing the preprocessed audio signals to determine a pair of signals to be pairwise processed, whether the pair of signals has the same associated window length, and

wherein the adaptive joint signal processor is configured to only allow a pairwise processing of two signals, when the two signals have associated therewith the same window length applied by the signal preprocessor.

22. Multisignal encoder of one of the preceding claims,

wherein the adaptive joint signal processor is configured to apply a non-cascaded signal-pair processing, in which the signals of the processed signal pair are not selectable in a further signal pair processing, wherein the adaptive joint signal processor is configured for selecting the signal pairs based on a cross-correlation between the signal pairs for the pairwise processing, and wherein the pairwise processing of several selected signal pairs is performed in parallel.

23. Multisignal encoder of claim 22,

wherein the adaptive joint signal processor is configured to determine, for a selected signal pair, a stereo encoding mode, and wherein, when the stereo encoding mode is determined to be a dual mono mode, the signals involved in this signal pair are at least partly re-scaled and indicated as signals to be encoded individually.

24. Multisignal encoder of one of claims 15 and 16,

wherein the adaptive joint signal processor is configured for performing, for a pairwise processed signal pair, a stereo intelligent gap filling (IGF) operation, if a stereo mode of the core region is different to a stereo mode of the IGF region, or if the stereo mode of the core is flagged as bandwise mid/side coding, or

wherein the adaptive joint signal processor is configured to apply a single signal IGF analysis for signals of a pairwise processed signal pair, if the stereo mode of the core region is not different to the stereo mode of the IGF region or if the stereo mode of the core is not flagged as bandwise mid/side encoding mode.

25. Multisignal encoder of one of the preceding claims,

wherein the adaptive joint signal processor is configured for performing an intelligent gap filling operation, before a result of the IGF operation is individually encoded by the signal encoder,

wherein a power spectrum is used for a tonality/noise determination in a quantization and intelligent gap filling (IGF), and wherein the signal preprocessor is configured for performing the same frequency domain noise shaping for an MDST spectrum as has been applied to the MDCT spectrum, and

wherein the adaptive joint signal processor is configured for performing the same mid/side processing to a preprocessed MDST spectrum, so that a result of a processed MDST spectrum is used within a quantization performed by the signal encoder or within an intelligent gap filling processing performed by the adaptive joint signal processor, or

wherein the adaptive joint signal processor is configured for applying the same normalization scaling based on the full band scaling vector for the MDST spectrum as it was done for the MDCT spectrum using the same quantized scaling vectors.

26. Multisignal encoder of one of the preceding claims, wherein the adaptive joint signal processor is configured for performing a pairwise processing of the at least three preprocessed audio signals to obtain the at least three jointly processed signals or at least two jointly processed signals and a signal to be encoded individually.

27. Multisignal encoder of one of the preceding claims, wherein the audio signals of the at least three audio signals are audio channels, or

wherein the audio signals of the at least three audio signals are audio component signals of a sound field description such as an Ambisonics sound field description, a B format description, an A format description or any other sound field description such as a sound field description describing a sound field with respect to a reference location.

28. Multisignal encoder of one of the preceding claims,

wherein the signal encoder is configured for individually encoding each signal to obtain at least three individually encoded signals, or for performing an (entropy) coding with more than one signals.

29. Multisignal decoder for decoding an encoded signal, comprising:

a signal decoder for decoding at least three encoded signals;

a joint signal processor configured

to extract, from the encoded signal, energy normalization values for the joint stereo decoded signals,

to perform a joint signal processing comprising a pairwise processing the decoded signals using a joint stereo mode as indicated by side information included in the encoded signal, and

to energy re-scale the joint stereo decoded signals using the energy normalization values to obtain at least three processed decoded signals; and

a post processor for post processing the at least three processed decoded signals in accordance with side information included in the encoded signal, wherein the post processing is performed so that the post processed signals are less white than the signals before post processing, and wherein the post processed signals represent a decoded audio signal.

30. Multisignal decoder of claim 29,

wherein the joint signal processor is configured to check, whether an energy normalization value extracted from the encoded signal for a specific signal has a predefined value, and

wherein the joint signal processor is configured to not perform an energy re-scaling or to only perform a reduced energy re-scaling to the specific signal, when the energy normalization value has the predefined value.

31. Multisignal decoder of one of claims 29 or 30, wherein the signal decoder is configured

to extract, from the encoded signal, a bit distribution value for each encoded signal,

to determine a used bit distribution for a signal using the bit distribution value for the signal, a number of remaining bits for all signals, and a further refinement step, or a final donation step; and

to perform the individual decoding based on the used bit distribution for each signal.

32. Multisignal decoder of one of claims 29 to 31, wherein the joint signal processor is configured

to perform a band replication or intelligent gap filling processing to the individually decoded signals using side information in the encoded signal to obtain spectrally enhanced individual signals, and

to perform a joint processing in accordance with a joint processing mode using the spectrally enhanced individual signals.

33. Multisignal decoder of claim 32,

wherein the joint signal processor is configured to transform a source range from one stereo representation to another stereo representation, when a destination range is indicated as having the other stereo representation.

34. Multisignal decoder of one of claims 29 to 33, wherein the joint signal processor is configured

for extracting, from the encoded signal, a flag indicating whether the energy normalization value is either an upscaling value or a downscaling value, and

to perform the energy rescaling using the energy normalization value as a downscaling, when the flag has a first value and as an upscaling, when the flag has the second value different from the first value.

35. Multisignal decoder of one of claims 29 to 34, wherein the joint signal processor is configured

to extract, from the encoded signal, a side information indicating signal pairs that result from a jointly coding operation,

to perform an inverse stereo or multichannel processing starting with a last signal pair to obtain the encoded signal, in order to convert back to original preprocessed spectra of each signal, and to perform the inverse stereo processing based on the stereo mode

and/or a bandwise mid/side decision as indicated in the side information for the encoded signal.

36. Multisignal decoder of one of claims 29 to 35,

wherein the joint signal processor is configured to energy-rescale all joint stereo decoded signals involved in signal pairs to a corresponding original energy level based on quantized energy normalization values for each joint stereo decoded signal, and wherein other signals that were not involved in signal pair processing are not energy-rescaled as the signals involved in signal pair processing.

37. Multisignal decoder of one of claims 29 to 36,

wherein the post processor is configured to perform, for each individual processed decoded signal, a temporal noise shaping operation or a frequency domain noise shaping operation and a conversion from a spectral domain into a time domain and a subsequent overlap/add operation between subsequent time frames of a post processed signal.

38. Multisignal decoder of one of claims 29 to 37,

wherein the joint signal processor is configured to extract, from the encoded signal, a flag indicating, whether several bands for a time frame of a signal pair are to be inversely processed using mid/side or left/right encoding, and wherein the joint signal processor is configured to use this flag to collectively subject the corresponding bands of the signal pair to either mid/side processing or left/right processing depending on the value of the flag, and

wherein, for a different time frame for the same signal pair or for a different signal pair at the same time frame, an encoding mode mask indicating, for each individual band, an individual coding mode, is extracted from the side information of the encoded signal, and wherein the joint signal processor is configured to whether apply inverse mid/side processing or left/right processing for the corresponding band as indicated for a bit associated with this band.

39. Multisignal decoder of one of the claims 29 to 38, wherein the encoded signal is an encoded multichannel signal, wherein the multisignal decoder is a multichannel decoder, wherein the encoded signal is an encoded multichannel signal, wherein the signal decoder is a channel decoder, wherein the encoded signals are encoded channels, wherein the joint signal processing is a joint channel processing, wherein the at least three processed decoded signals are at least three processed decoded signals, wherein the post processed signals are channels, or

wherein the encoded signal is an encoded multicomponent signal representing audio component signals of a sound field description such as an Ambisonics sound field description, a B format description, an A format description or any other sound field description such as a sound field description describing a sound field with respect to a reference location, wherein the multisignal decoder is a multicomponent decoder, wherein the encoded signal is an encoded multicomponent signal, wherein the signal decoder is a component decoder, wherein the encoded signals are encoded components, wherein the joint signal processing is a joint component processing, wherein the at least three processed decoded signals are at least three processed decoded components, and wherein the post processed signals are component audio signals.

40. Method for performing multisignal encoding at least three audio signals, comprising:

individually preprocessing each audio signal to obtain at least three preprocessed audio signals, wherein the preprocessing is performed so that a preprocessed audio signal is whitened with respect to the signal before preprocessing;

performing a processing of the at least three preprocessed audio signals subsequent to an energy level normalization to obtain at least three jointly processed signals or at least two jointly processed signals and an unprocessed signal, wherein the energy level normalization comprises:

calculating an information on a mean energy of the preprocessed audio signals;

calculating an information on an energy of each preprocessed audio signal,

calculating the broadband energy normalization value for a specific preprocessed audio signal based on the information on the mean energy and the information on the energy of the specific preprocessed audio signal, and

performing a broadband energy normalization of the at least three preprocessed audio signals using the broadband energy normalization values for the specific preprocessed audio signals, wherein each preprocessed audio signal has a normalized energy;

encoding each signal to obtain one or more encoded signals; and

transmitting or storing an encoded multisignal audio signal comprising the one or more encoded signals, side information relating to the preprocessing and side information relating to the processing comprising the broadband energy normalization values for the preprocessed audio signals.

41. Method for multisignal decoding an encoded signal, comprising:

individually decoding at least three encoded signals;

performing a joint signal processing comprising

extracting, from the encoded signal, energy normalization values for the joint stereo decoded signals,

performing a joint signal processing comprising a pairwise processing the decoded signals using a joint stereo mode as indicated by side information included in the encoded signal, and

to energy re-scaling the joint stereo decoded signals using the energy normalization values to obtain at least three processed decoded signals; and

post processing the at least three processed decoded signals in accordance with side information included in the encoded signal, wherein the post processing is performed so that the post processed signals are less white than the signals before post processing, and wherein the post processed signals represent a decoded audio signal.

42. Computer program for performing, when running on a computer or a processor, the method of claim 40 or the method of claim 41.