Claims

1. An audio encoder, comprising:

a transport signal generator for generating two or more transport channels of a transport signal from audio input comprising at least one of a plurality of audio input objects and a plurality of audio input channels,

a voice activity determiner for determining a voice activity decision for the transport signal, which indicates whether or not the audio input within the transport signal exhibits voice activity, and

a bitstream generator for generating a bitstream depending on the audio input,

wherein, if the voice activity determiner has determined that the transport signal exhibits voice activity, the bitstream generator is adapted to encode the two or more transport channels within the bitstream,

wherein, if the voice activity determiner has determined that the transport signal does not exhibit voice activity, the bitstream generator is suitable to encode, instead of the two or more transport channels, information on a background noise, wherein the information on the background noise comprises information on a background noise of at least one of the two or more transport channels or information on a background noise of a derived signal which depends on at least one of the two or more transport channels.

2. An audio encoder according to claim 1,

wherein the voice activity determiner is configured to determine an individual voice activity decision for each transport channel of one or more transport channels of the transport signal, which indicates whether or not the audio input within the transport channel exhibits voice activity, and

wherein the voice activity determiner is configured to determine the voice activity decision for the transport signal depending on the individual voice activity decision of each transport channel of the one or more transport channels.

3. An audio encoder according to claim 2,

wherein the voice activity determiner is configured to determine an individual voice activity decision for each transport channel of the two or more transport channels of the transport signal, which indicates whether or not the audio input within said transport channel exhibits voice activity, and

wherein the voice activity determiner is configured to determine the voice activity decision for the transport signal depending on the individual voice activity decision of each transport channel of the two or more one transport channels of the transport signal.

4. An audio encoder according to claim 3,

wherein the voice activity determiner is configured to determine that the transport signal exhibits voice activity, if at least one of the two or more transport channels of the transport signal exhibits voice activity, and

wherein the voice activity determiner is configured to determine that the transport signal does not exhibit voice activity, if none of the two or more transport channels of the transport signal exhibits voice activity.

5. An audio encoder according to one of the preceding claims,

wherein the audio encoder is configured to determine, if the voice activity determiner has determined that the transport signal does not exhibit voice activity, whether to transmit the bitstream having encoded therein the information on the background noise, or whether to not generate and to not transmit the bitstream.

6. An audio encoder according to one of the preceding claims,

wherein the audio encoder comprises a mono signal generator for generating, if the voice activity determiner has determined that the transport signal does not exhibit voice activity, the derived signal as a mono signal from at least one of the two or more transport channels, and

wherein the audio encoder comprises an information generator for generating the information on the background noise as information on the background noise of the mono signal.

7. An audio encoder according to claim 6,

wherein the mono signal generator is configured to generate the mono signal by adding the two or more transport channels or by adding two or more channels derived from the two or more transport channels, or

wherein the mono signal generator is configured to generate the mono signal by choosing that transport channel of the two or more transport channels which exhibits a higher energy.

8. An audio encoder according to claim 6 or 7,

wherein the information generator is configured to generate the information on a background noise of the mono signal as the information on the mono signal.

9. An audio encoder according to claim 8,

wherein the information generator is configured to generate a silence insertion description of the background noise of the mono signal as the information on the background noise of the mono signal.

10. An audio encoder according to one of the preceding claims,

wherein the audio encoder comprises a direction information determiner for determining direction information depending on the audio input,

wherein the audio encoder comprises a direction information quantizer for quantizing the direction information to obtain quantized direction information, and

wherein the bitstream generator is configured to encode the quantized direction information within the bitstream.

11. An audio encoder according to claim 10,

wherein the transport signal generator is configured to generate the two or more transport channels of the transport signal from the audio input using the direction information.

12. An audio encoder according to claim 10 or 11,

wherein the audio input comprises the plurality of audio input objects,

wherein the direction information comprises information on an azimuth angle and on an elevation angle of an audio input object of the plurality of audio input objects of the audio input.

13. An audio encoder according to one of the preceding claims,

wherein the audio encoder comprises an active metadata generator for generating metadata comprising at least one of quantized direction information, object indices and power ratios of the plurality of audio input objects and or of the plurality of audio input channels of the audio input, if the voice activity determiner has determined that the transport signal exhibits voice activity.

14. An audio encoder according to one of the preceding claims,

wherein the audio input comprises the plurality of audio input objects, and

wherein the audio encoder comprises an inactive metadata generator for generating, if the voice activity determiner has determined that the transport signal does not exhibit voice activity, metadata comprising quantized direction information and control parameters, for example comprising a scaling factor and/or either a coherence or a correlation.

15. An audio encoder according to claim 13 and according to claim 14,

wherein the direction information that is generated by the inactive metadata generator differs in a quantization resolution from the metadata that is generated by the active metadata generator.

16. An audio encoder according to claim 14 or 15, further depending on claim 13,

wherein the inactive metadata generator is configured to generate the control parameters such that the control parameters differ in characteristics from a characteristics of power ratios and object indices that are generated by the active

metadata generator, for example wherein the control parameters comprise, e.g., the scaling factor and/or, e.g., either the coherence or the correlation.

17. An audio encoder according to one of the preceding claims,

wherein the audio input comprises a plurality of audio input objects and metadata being associated with the audio input objects.

18. An audio encoder according to one of the preceding claims,

wherein the transport signal generator is configured to generate the two or more transport channels of the transport signal from the audio input comprising by downmixing at least one of a plurality of audio input objects and a plurality of audio input channels to obtain a downmix as the transport signal, which comprises two or more downmix channels as the two or more transport channels.

19. An audio encoder according to one of claims 10 to 13 and according to claim 18,

wherein, if the audio input within the transport signal does not exhibit voice activity, the direction information quantizer is configured to determine the quantized direction information such that a quantization resolution of the quantized direction information is different from a quantization resolution used for computing the downmix.

20. An audio encoder according to one of the preceding claims, further depending on claim 14,

wherein the bitstream generator is configured to encode the control parameters within the bitstream, if the voice activity determiner has determined that the transport signal does not exhibit voice activity,

wherein the control parameters are suitable for steering a generation of an intermediate signal from random noise,

wherein the control parameters either comprises a plurality of parameter values for a plurality of subbands, or wherein the control parameters are single broadband control parameters.

21. An audio encoder according to claim 20,

wherein the audio encoder is configured generate the control parameters, by selecting, whether the control parameters either comprises the plurality of parameter values for the plurality of subbands, or whether the control parameters are the single broadband control parameters, depending on an available bitrate.

22. An audio encoder according to one of the preceding claims,

wherein the transport signal generator is configured to encode the audio input by applying Code-Excited Linear Prediction or by applying a Modified Discrete Cosine Transform or by applying a combination of the Code-Excited Linear Prediction and of the Modified Discrete Cosine Transform.

23. An audio encoder according to one of the preceding claims,

wherein, if the audio input comprises the plurality of audio input channels, but not the plurality of audio input objects, a number of the two or more transport channels is smaller than a number of the plurality of audio input channels,

wherein, if the audio input comprises the plurality of audio input objects, but not the plurality of audio input channels, the number of the two or more transport channels is smaller than a number of the plurality of audio input objects,

wherein, if the audio input comprises both the plurality of audio input objects and the plurality of audio input channels, the number of the two or more transport channels is smaller than a sum of the number of the plurality of audio input channels and the number of the plurality of the audio input objects;

or

wherein, if the audio input comprises the plurality of audio input channels, but not the plurality of audio input objects, a number of the two or more transport channels is smaller than or equal to a number of the plurality of audio input channels,

wherein, if the audio input comprises the plurality of audio input objects, but not the plurality of audio input channels, the number of the two or more transport channels is smaller than or equal to a number of the plurality of audio input objects,

wherein, if the audio input comprises both the plurality of audio input objects and the plurality of audio input channels, the number of the two or more transport channels is smaller than or equal to a sum of the number of the plurality of audio input channels and the number of the plurality of the audio input objects.

24. A system, comprising:

an audio encoder according to one of the preceding claims, and

an audio decoder,

wherein the audio decoder comprises:

an input interface for receiving a bitstream which depends on audio content comprising at least one of a plurality of audio objects and a plurality of audio channels; wherein a transport signal comprising two or more transport channels is encoded within the bitstream, and the audio content is encoded within the transport signal; or wherein information on a background noise is encoded within the bitstream instead of the transport signal, wherein the information on the background noise comprises information on a background noise of at least one of the two or more transport channels or information on a background noise of a derived signal which depends on at least one of the two or more transport channels; and

a renderer for generating one or more audio output signals depending on the audio content being encoded with the bitstream;

wherein, if the transport signal comprising the two or more transport channels is encoded within the bitstream, the renderer is configured to generate the one or more audio output signals depending on the two or more transport channels, and

wherein, if the information on the background noise is encoded within the bitstream instead of the transport signal, the renderer is configured to generate the one or more audio output signals depending on the information on the background noise,

wherein the audio encoder is configured to generate a bitstream from audio input, and

wherein the audio decoder is configured to generate one or more audio output signals from the bitstream.

25. A method for audio encoding, wherein the method comprises:

generating two or more transport channels of a transport signal from audio input comprising at least one of a plurality of audio input objects and a plurality of audio input channels,

determining a voice activity decision for the transport signal, which indicates whether or not the audio input within the transport signal exhibits voice activity, and

determining a bitstream depending on the audio input,

wherein, if it has been determined that the transport signal exhibits voice activity, the method comprises encoding the two or more transport channels within the bitstream,

wherein, if it has been determined that the transport signal does not exhibit voice activity, the method comprises encoding, instead of the two or more transport channels, information on a background noise of at least one of the two or more transport channels or information on a background noise of a derived signal which depends on at least one of the two or more transport channels.

26. A computer program for implementing the method of claim 25 when being executed on a computer or signal processor.