

CLAIMS

1. An audio decoding method comprising:

receiving a two-channel downmix signal, which is associated with metadata, the metadata comprising upmix parameters for parametric reconstruction of an M -channel audio signal based on the downmix signal, where $M \geq 4$;

receiving at least a portion of said metadata;

generating a decorrelated signal based on at least one channel of the downmix signal;

determining a set of mixing coefficients based on the received metadata; and

forming a K -channel output signal as a linear combination of the downmix signal and the decorrelated signal in accordance with the mixing coefficients, wherein $2 \leq K < M$,

wherein the mixing coefficients are determined such that a sum of a mixing coefficient controlling a contribution from the first channel of the downmix signal to a channel of the output signal, and a mixing coefficient controlling a contribution from the first channel of the downmix signal to another channel of the output signal, has the value 1,

wherein, if the downmix signal represents the M -channel audio signal according to a first coding format in which:

a first channel of the downmix signal corresponds to a certain linear combination of a first group of one or more channels of the M -channel audio signal;

a second channel of the downmix signal corresponds to a certain linear combination of a second group of one or more channels of the M -channel audio signal; and

the first and second groups constitute a certain partition of the M channels of the M -channel audio signal,

then the K -channel output signal represents the M -channel audio signal according to a second coding format in which:

each of the K channels of the output signal approximates a linear combination of a group of one or more channels of the M -channel audio signal;

the groups corresponding to the respective channels of the output signal constitute a partition of the M channels of the M -channel audio signal into K groups of one or more channels; and

at least two of the K groups comprise at least one channel from said first group.

2. The audio decoding method of claim 1, wherein $K = 2$, $K = 3$ or $K = 4$, and/or wherein $M = 5$ or $M = 6$.

3. The audio decoding method of claim 1, wherein the received metadata includes the upmix parameters and wherein the mixing coefficients are determined by processing the upmix parameters.

4. The audio decoding method of claim 1, wherein:

in the first coding format, each of the channels of the M -channel audio signal is associated with a non-zero gain controlling a contribution from this channel to one of the linear combinations to which the channels of the downmix signal correspond;

in the second coding format, each of the channels of the M -channel audio signal is associated with a non-zero gain controlling a contribution from this channel to one of the linear combinations approximated by the channels of the output signal; and

for each of the channels of the M -channel audio signal, the non-zero gain associated with the channel in the first coding format coincides with the non-zero gain associated with the channel in the second coding format.

5. The audio decoding method of claim 1, further comprising an initial step of receiving a bit-stream representing the downmix signal and the metadata,

wherein the downmix signal and said received metadata are extracted from the bit-stream.

6. The audio decoding method of claim 1, wherein the decorrelated signal is a two-channel signal, and wherein said output signal is formed by including no more than two decorrelated signal channels into said linear combination of the downmix signal and the decorrelated signal.

7. The audio decoding method of claim 6, wherein $K = 3$, and wherein forming the output signal amounts to a projection from four channels to three channels.

8. The audio decoding method of claim 1, wherein said first group consists of two or three channels.

9. The audio decoding method of claim 1, wherein the M -channel audio signal comprises either three or four channels representing different horizontal directions in a playback envi-

ronment for the M -channel audio signal, and two channels representing directions vertically separated from those of said three or four channels in said playback environment.

10. The audio decoding method of claim 9, wherein said first group consists of said three channels, and wherein said second group consists of the two channels representing directions vertically separated from those of said three channels in said playback environment.

11. The audio decoding method of claim 10, wherein the two channels representing directions vertically separated from those of said three channels in said playback environment are comprised in different groups of the K groups.

12. The audio decoding method of claim 9, wherein one of the K groups comprises both of the two channels representing directions vertically separated from those of said three or four channels in said playback environment.

13. The audio decoding method of claim 1, wherein the decorrelated signal comprises two channels, a first channel of the decorrelated signal being obtained based on the first channel of the downmix signal and a second channel of the decorrelated signal being obtained based on the second channel of the downmix signal.

14. The audio decoding method of claim 1, wherein said first group consists of N channels, where $N \geq 3$, wherein said first group is reconstructable as a linear combination of said first channel of the downmix signal and an $(N - 1)$ -channel decorrelated signal by applying dry upmix coefficients to said first channel of the downmix signal and wet upmix coefficients to channels of the $(N - 1)$ -channel decorrelated signal, wherein the received metadata includes wet upmix parameters and dry upmix parameters, and wherein determining the mixing coefficients comprises:

determining, based on the dry upmix parameters, the dry upmix coefficients;

populating an intermediate matrix having more elements than the number of received wet upmix parameters, based on the received wet upmix parameters and knowing that the intermediate matrix belongs to a predefined matrix class;

obtaining the wet upmix coefficients by multiplying the intermediate matrix by a predefined matrix, wherein the wet upmix coefficients corresponds to the matrix resulting from the

multiplication and includes more coefficients than the number of elements in the intermediate matrix; and

processing the wet and dry upmix coefficients.

15. The audio decoding method of claim 1, further comprising:

signaling indicating one of at least two coding formats of the M -channel audio signal, the coding formats corresponding to respective different partitions of the channels of the M -channel audio signal into respective first and second groups associated with the channels of the downmix signal,

wherein the K groups are predefined, and wherein the mixing coefficients are determined such that a single partition of the M -channel audio signal into the K groups of channels, approximated by the channels of the output signal, is maintained for said at least two coding formats.

16. The audio decoding method of claim 15, wherein:

in a first coding format of said at least two coding formats, said first group consists of three channels representing different horizontal directions in a playback environment for the M -channel audio signal, and said second group consists of two channels representing directions vertically separated from those of said three channels in said playback environment; and

in a second coding format of said at least two coding formats, each of said first and second groups comprises one of said two channels representing directions vertically separated from those of said three channels in said playback environment.

18. An audio decoding system comprising a decoding section configured to:

receive a two-channel downmix signal, which is associated with metadata, the metadata comprising upmix parameters for parametric reconstruction of an M -channel audio signal based on the downmix signal, where $M \geq 4$;

receive at least a portion of said metadata; and

provide a K -channel output signal based on the downmix signal and the received metadata, wherein $2 \leq K < M$,

the decoding section comprising:

a decorrelating section configured to receive at least one channel of the downmix signal and to output, based thereon, a decorrelated signal; and

a mixing section configured to
determine a set of mixing coefficients based on the received metadata, and
form the output signal as a linear combination of the downmix signal and the decorrelated signal in accordance with the mixing coefficients,

wherein the mixing section is configured to determine the mixing coefficients such that a sum of a mixing coefficient controlling a contribution from the first channel of the downmix signal to a channel of the output signal, and a mixing coefficient controlling a contribution from the first channel of the downmix signal to another channel of the output signal, has the value 1,

wherein, if the downmix signal represents the M -channel audio signal according to a first coding format in which:

a first channel of the downmix signal corresponds to a certain linear combination of a first group of one or more channels of the M -channel audio signal;

a second channel of the downmix signal corresponds to a certain linear combination of a second group of one or more channels of the M -channel audio signal; and

the first and second groups constitute a certain partition of the M channels of the M -channel audio signal,

then the K -channel output signal represents the M -channel audio signal according to a second coding format in which:

each of the K channels of the output signal approximates a linear combination of a group of one or more channels of the M -channel audio signal;

the groups corresponding to the respective channels of the output signal constitute a partition of the M channels of the M -channel audio signal into K groups of one or more channels; and

at least two of the K groups comprise at least one channel from said first group.

19. The audio decoding system of claim 18, further comprising an additional decoding section configured to:

receive an additional two-channel downmix signal which is associated with additional metadata, the additional metadata comprising additional upmix parameters for parametric reconstruction of an additional M -channel audio signal based on the additional downmix signal,

receive at least a portion of the additional metadata; and

provide an additional K -channel output signal based on the additional downmix signal and the additional received metadata,

the additional decoding section comprising:

an additional decorrelating section configured to receive at least one channel of the additional downmix signal and to output, based thereon, an additional decorrelated signal; and

an additional mixing section configured to:

determine a set of additional mixing coefficients based on the received additional metadata, and

form the additional output signal as a linear combination of the additional downmix signal and the additional decorrelated signal in accordance with the additional mixing coefficients,

wherein the additional mixing section is configured to determine the additional mixing coefficients such that a sum of a mixing coefficient controlling a contribution from the first channel of the additional downmix signal to a channel of the additional output signal, and a mixing coefficient controlling a contribution from the first channel of the additional downmix signal to another channel of the additional output signal, has the value 1,

wherein, if the additional downmix signal represents the additional M -channel audio signal according to a third coding format in which:

a first channel of the additional downmix signal corresponds to a linear combination of a first group of one or more channels of the additional M -channel audio signal; a second channel of the additional downmix signal corresponds to a linear combination of a second group of one or more channels of the additional M -channel audio signal; and

the first and second groups of channels of the additional M -channel audio signal constitute a partition of the M channels of the additional M -channel audio signal,

then the additional K -channel output signal represents the additional M -channel audio signal according to a fourth coding format in which:

each of the K channels of the additional output signal approximates a linear combination of a group of one or more channels of the M -channel audio signal;

the groups corresponding to the respective channels of the additional output signal constitute a partition of the M channels of the additional M -channel audio signal into K groups of one or more channels; and

at least two of the K groups of one or more channels of the additional M -channel audio signal comprise at least one channel from said first group of channels of the additional M -channel audio signal.

20. The decoding system of claim 18, further comprising:

a demultiplexer configured to extract, from a bitstream, the downmix signal, said received metadata, and a discretely coded audio channel; and

a single-channel decoding section operable to decode said discretely coded audio channel.