(54) **PROCESSING OF AUDIO STEREO SIGNAL**

(57)   An audio decoder apparatus comprises a receiver (101) receiving a mono audio signal being a downmix of two channel signals of a first audio stereo signal and upmix parameters comprising a first parameter indicative of a level difference between the two channel signals, a second parameter indicative of a correlation between the two channel signals, and a third parameter indicative of a phase difference between the two channel signals. A coefficient generator (107) generates coefficients for an upmix matrix from the upmix parameters and a generator (103) generates an output audio stereo signal by applying the upmix matrix to the mono audio signal and an auxiliary signal. The coefficient generator (107) determine a signal cancellation measure for signal cancellation in a sum of the channel signals from the upmix parameters and determines the coefficients for the upmix matrix in dependence on the signal cancellation measure. The approach may e.g. mitigate numerical processing issues.
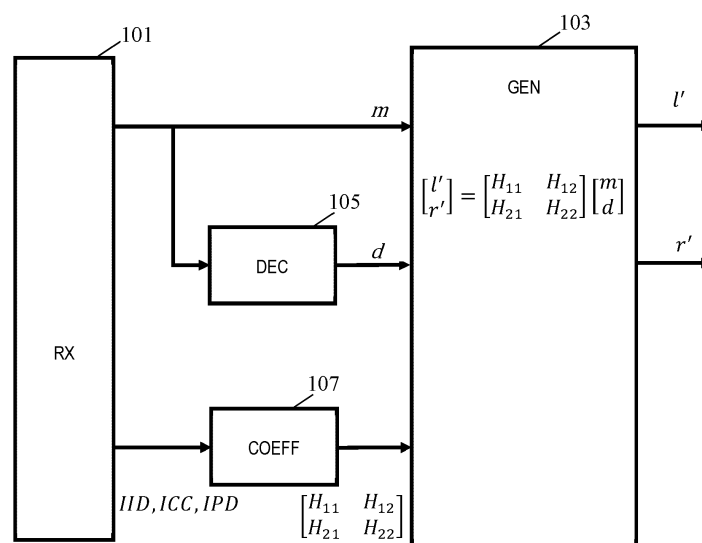
FIG. 1

Description

FIELD OF THE INVENTION

[0001]   The invention relates to processing, such as encoding/decoding/ downmixing/upmixing/generation of an audio stereo signal, and in particular, but not exclusively, to generation of an audio stereo signal from upmixing of a mono downmix signal using upmix parametric data.

BACKGROUND OF THE INVENTION

[0002]   Spatial audio applications have become numerous and widespread and increasingly form at least part of many audiovisual experiences. Indeed, new and improved spatial experiences and applications are continuously being developed which result in increased demands on audio processing and rendering.

[0003]   For example, in recent years, Virtual Reality (VR) and Augmented Reality (AR) have received increasing interest and a number of implementations and applications are reaching the consumer market. Indeed, equipment is being developed for both rendering the experience as well as for capturing or recording suitable data for such applications. For example, relatively low-cost equipment is being developed for allowing gaming consoles to provide a full VR experience. It is expected that this trend will continue and indeed will increase in speed with the market for VR and AR reaching a substantial size within a short time scale. In the audio domain, a prominent field explores the reproduction and synthesis of realistic and natural spatial audio. The ideal aim is to produce natural audio sources such that the user cannot recognize the difference between a synthetic and an original source.

[0004]   A lot of research and development effort has focused on providing efficient and high-quality audio encoding and audio decoding for spatial audio. A frequently used spatial audio representation is multichannel audio representations, including stereo representation, and efficient encoding of such multichannel audio based on downmixing multichannel audio signals to downmix channels with fewer channels have been developed. One of the main advances in low bit-rate audio coding has been the use of parametric multichannel coding where a downmix signal is generated together with parametric data that can be used to upmix the downmix signal to recreate the multichannel audio signal.

[0005]   In particular, instead of traditional mid-side or intensity coding, in parametric multichannel audio coding a multichannel input signal is downmixed to a lower number of channels (e.g. two to one) and multichannel image (stereo) parameters are extracted. Then the downmix signal is encoded using a more traditional audio coder (e.g. a mono audio encoder). The bitstream of the downmix is multiplexed with the encoded multichannel image parameter bitstream. This bitstream is then transmitted to the decoder, where the process is inverted. First the downmix audio signal is decoded, after which the multichannel audio signal is reconstructed guided by the encoded multichannel image upmix parameters.

[0006]   An example of stereo coding is described in E. Schuijers, W. Oomen, B. den Brinker, J. Breebaart, "Advances in Parametric Coding for High-Quality Audio", 114th AES Convention, Amsterdam, The Netherlands, 2003, Preprint 5852. In the described approach, the downmixed mono signal is parametrized by exploiting the natural separation of the signal into three components (objects): transients, sinusoids, and noise. In E. Schuijers, J. Breebaart, H. Purnhagen, J. Engdegård, "Low Complexity Parametric Stereo Coding", 116th AES, Berlin, Germany, 2004, Preprint 6073 more details are provided describing how parametric stereo was realized with a low (decoder) complexity when combining it with Spectral Band Replication (SBR).

[0007]   In the described approaches, the decoding is based on the use of the so-called de-correlation process. The de-correlation process generates a decorrelated helper signal from the monaural signal. In the stereo reconstruction process, both the monaural signal and the decorrelated helper signal are used to generate the upmixed stereo signal based on the upmix parameters. Specifically, the two signals may be multiplied by a time- and frequency-dependent 2x2 matrix having coefficients determined from the upmix parameters to provide the output stereo signal. The approach allows parametric stereo encoding/decoding to be realized with a low (decoder) complexity when combining it with Spectral Band Replication (SBR). The de-correlation process generates a synthetic helper signal d[n] from the monaural signal m[n]. In the stereo reconstruction process both signals m[n] and d[n] are mixed to form the stereo pair l[n], r[n]. To further reduce (decoder) complexity, it is also described how the de-correlation process can be moved into the sub-band domain. This is the form in which it has also been standardized for HE-AACv2 (see e.g. A. C. den Brinker, J. Breebaart, P. Ekstrand, J. Engdegård, F. Henn, K. Kjörling, W. Oomen, and H. Purnhagen, "An Overview of the Coding Standard MPEG-4 Audio Amendments 1 and 2: HE-AAC, SSC, and HE-AAC v2", EURASIP Journal on Audio Speech and Music Processing, January 2009 and ISO/IEC 14496-3:2005, Information technology - Coding of audio-visual objects - Part 3: Audio).

[0008]   However, although Parametric Stereo (PS) and similar downmix encoding/ decoding approaches were a leap forward from traditional stereo and multichannel coding, the approach is not optimal in all scenarios. In particular, known encoding and decoding approaches tend to introduce some distortion, changes, artefacts etc. that may introduce differences between the (original) stereo audio signal input to the encoder and the stereo audio signal recreated at the decoder. Typically, the audio quality may be degraded and imperfect recreation of the multichannel occurs. Further, the

data rate may still be higher than desired and/or the complexity/ resource usage of the processing may be higher than preferred. The encoding and decoding process is typically not ideal and in particular for some particular signals, the process may introduce undesired effects, degradations, inaccuracies, and/or artefacts.

[0009]   Hence, an improved approach would be advantageous. In particular, an approach allowing increased flexibility, improved adaptability, an improved performance, prevention or mitigation of numerical issues of audio processing including encoding and decoding, increased audio quality, improved audio quality to data rate trade-off, reduced complexity and/or resource usage, reduced computational load, facilitated implementation and/or an improved spatial audio experience would be advantageous.

SUMMARY OF THE INVENTION

[0010]   Accordingly, the Invention seeks to preferably mitigate, alleviate or eliminate one or more of the above mentioned disadvantages singly or in any combination.

[0011]   According to an aspect of the invention there is provided apparatus for generating an output audio stereo signal, the apparatus comprising: a receiver arranged to receive an audio data signal comprising: a mono audio signal being a downmix of two channel signals of a first audio stereo signal; upmix parameters for the mono audio signal, the set of upmix parameters comprising a first parameter indicative of a level difference between the two channel signals, a second parameter indicative of a correlation between the two channel signals, and a third parameter indicative of a phase difference between the two channel signals; a coefficient generator arranged to generate coefficients for an upmix matrix from the upmix parameters; a generator arranged to generate the output audio stereo signal by applying the upmix matrix to samples of the mono audio signal and an auxiliary mono audio signal; wherein the coefficient generator is arranged to: determine a signal cancellation measure from the upmix parameters, the signal cancellation measure being indicative of a signal cancellation in a summation of the two channel signals; and determine the coefficients for the upmix matrix in dependence on the signal cancellation measure.

[0012]   The approach may provide an improved audio experience in many embodiments and applications. For many signals and scenarios, the approach may provide improved generation/ reconstruction of a stereo audio signal with an improved perceived audio quality.

[0013]   The approach may provide an efficient implementation and may in many embodiments allow a reduced complexity and/or resource usage. The approach may in many scenarios allow a reduced data rate for data representing a multichannel audio signal using a downmix signal.

[0014]   The approach may in particular mitigate and compensate for numerical issues, and may typically provide more well-behaved parameter values and/or calculations. In particular, the processing and parameter determination may prevent that denominators for equations evaluated to determine the upmix coefficients approach zero. The approach may prevent or reduce the risk of parameter values (whether final or intermediate) exceeding suitable dynamic ranges, and may specifically prevent or reduce the risk of these parameter values approaching infinity. The approach may further achieve such effects while allowing optimum (or improved) determination of upmix parameters for most scenarios. For example, modifications that prevent or mitigate numerical issues may be focused on scenarios where these are likely to occur without having a significant impact on the operation in other scenarios.

[0015]   As a specific example, an approach for determining upmix parameters may closely follow the approach of ISO/IEC 14496-3:2005 for many scenarios while specifically preventing or mitigating numerical issues associated with this approach for some scenarios and signals.

[0016]   The approach may reduce distortions and artefacts resulting from numerical issues when determining upmix parameter values.

[0017]   Further, the approach may allow encoder and decoder side coordination. In particular, in many embodiments, the apparatus for generating an output audio stereo signal may determine the signal cancellation measure based only on the received upmix parameters, and these may specifically reflect properties of the channel signals of the input stereo signal at the encoder. Accordingly, the same information may be available at both the encoding and decoding side and the same signal cancellation measure may be determined at the encoding and decoding side. The upmix parameters may accordingly be determined to match the applied downmix parameters at the encoding side. In particular, the upmix parameters may be determined such that the upmix matrix closely complements the downmix matrix. Specifically, the upmix matrix may be determined as the inverse of the downmix matrix thereby resulting in the sequence of the downmix matrix multiplication and the upmix matrix multiplication resulting in the unity matrix.

[0018]   The samples of the mono audio signal may be frequency domain samples, or may span a particular time and frequency range (specifically subband domain samples). The samples of an auxiliary audio signal may be time domain samples, may be frequency domain samples, or may span a particular time and frequency range (specifically subband domain samples).

[0019]   The upmix parametric data may comprise data being indicative of relative properties between channel signals of the first stereo audio signal. The upmix parameters may comprise data being indicative of differences in properties

between channels of the stereo audio signal. The upmix parameters comprise data being perceptually relevant for the synthesis of the output stereo audio signal. The properties may for example be differences in phase and/or intensity and/or timing and/or correlation. The upmix parameters may in some embodiments and scenarios represent abstract properties not directly understandable by a human person/expert (but may typically facilitate a better reconstruction/lower data rate etc). The upmix parameters may comprise data including at least one of interchannel intensity differences, interchannel timing differences, interchannel correlations and/or interchannel phase differences for channel signals of the stereo audio signal.

[0020]   The upmix parameters may specifically include Interaural Intensity Differences (IIDs), Interaural Level Differences (ILD), Inter-channel Phase Differences (IPDs), Overall Phase Differences (OPDs), Inter-channel Cross Correlations (ICCs), Channel Phase Differences (CPDs) parameters

[0021]   The generator may be arranged to generate the output stereo audio signal by applying a matrix multiplication to the mono audio signal and the auxiliary audio signal with the coefficients of the upmix matrix being determined as a function of parameters of the upmix parameters. The upmix matrix be time- and frequency-dependent. Equivalently, the upmix matrix may be provided for a time and/or frequency segment, and different matrices may be provided for different time and/or frequency segments.

[0022]   The auxiliary signal may be a decorrelated signal generated from the mono audio signal. The decorrelated signal may be generated to have the same level and/or frequency distribution as the mono audio signal. The auxiliary signal may in some cases be a signal received with the mono audio signal, and may in particular be a side or residual signal for the first audio stereo signal.

[0023]   The signal cancellation measure may be indicative of a degree or level of signal cancellation in a summation of the two channel signals. The signal cancellation measure may be indicative of a signal level/power/amplitude of a sum signal being a summation of the two channel signals relative to the sum of the signal level/power/ amplitudes of the two channel signals.

[0024]   The signal cancellation measure may be indicative of a (degree/level) of signal cancellation in a summation of the two channel signals and/or equivalently may be indicative of a (degree/level) of signal cancellation in a difference/subtraction between the two channel signals (which may be considered a negative signal cancellation for the sum signal).

[0025]   The signal cancellation measure may in some embodiments be a normalized signal cancellation measure, and specifically normalized with respect to a level/power/energy of the first stereo signal. The signal cancellation measure may in some embodiments be in the range from -1 to +1. In some embodiments, the sign of the signal cancellation measure may indicate whether signal cancellation occurs in a summation of the channel signals and/or in a difference/ subtraction between the channel signals.

[0026]   According to an optional feature of the invention, the coefficient processor is arranged to adapt the upmix coefficients to deviate from coefficients for the mono audio signal being a sum signal of the channel signals for the signal cancellation measure meeting a first signal cancellation requirement.

[0027]   This may provide a particularly advantageous implementation and/or performance, and may in particular in many scenario prevent or mitigate numerical issues, artefacts, and/or signal distortions.

[0028]   According to an optional feature of the invention, the coefficient processor is arranged to increase a deviation of upmix coefficients from coefficients for the mono audio signal being a sum signal of the channel signals for the signal cancellation measure being indicative of an increasing signal cancellation in the sum signal of the channel signals.

[0029]   This may provide a particularly advantageous implementation and/or performance, and may in particular in many scenario prevent or mitigate numerical issues, artefacts, and/or signal distortions.

[0030]   According to an optional feature of the invention, the coefficient processor is arranged to increase a deviation of upmix coefficients from coefficients for the mono audio signal being a sum signal of the channel signals for the signal cancellation measure being indicative of an increasing signal cancellation in a difference signal of the channel signals.

[0031]   This may provide a particularly advantageous implementation and/or performance, and may in particular in many scenario prevent or mitigate numerical issues, artefacts, and/or signal distortions.

[0032]   According to an optional feature of the invention, the coefficients for the mono audio signal being a sum signal of the channel signals are given as

$$\frac{1}{2c}\begin{bmatrix} 1+\alpha & \beta \\ 1-\alpha & -\beta \end{bmatrix}$$

where

$$\alpha = \frac{IID - 1 + 2j \cdot \sin{(IPD)} \cdot ICC \cdot \sqrt{IID}}{IID + 1 + 2 \cdot \cos{(IPD)} \cdot ICC \cdot \sqrt{IID}}$$

$$\beta = \frac{2 \cdot \sqrt{IID \cdot (1 - ICC^2)}}{IID + 1 + 2 \cdot \cos{(IPD)} \cdot ICC \cdot \sqrt{IID}}$$

$$c = \min\left(\sqrt{\frac{IID + 1}{IID + 1 + 2 \cdot \cos{(IPD)} \cdot ICC \cdot \sqrt{IID}}}, c_{max}\right)$$

in which IID is an Interaural Intensity Differences upmix parameter, ICC is an Interchannel Cross Correlations upmix parameter, and IPD is an Inter-channel Phase Difference upmix parameter

[0033] This may provide a particularly advantageous implementation and/or performance, and may in particular in many scenario prevent or mitigate numerical issues, artefacts, and/or signal distortions.

[0034] According to an optional feature of the invention, the signal cancellation measure is determined substantially as:

$$R = \frac{2 \cdot ICC \cdot \cos{(IPD)}}{\sqrt{IID} + \frac{1}{\sqrt{IID}}}$$

where IID is an Interaural Intensity Difference upmix parameter, ICC is an Inter-channel Cross Correlations upmix parameter, and IPD is an Inter-channel Phase Difference upmix parameter.

[0035] This may provide a particularly advantageous implementation and/or performance, and may in particular in many scenario prevent or mitigate numerical issues, artefacts, and/or signal distortions.

[0036] According to an optional feature of the invention, the coefficient processor is arranged to generate a first intermediate parameter indicative of a prediction of a difference signal for the channel signals from the mono audio signal, and to generate the upmix coefficients in response to the first intermediate parameter, the first intermediate parameter being dependent on the signal cancellation measure.

[0037] This may provide a particularly advantageous implementation and/or performance, and may in particular in many scenario prevent or mitigate numerical issues, artefacts, and/or signal distortions.

[0038] According to an optional feature of the invention, the coefficient processor (107) is arranged to generate a second intermediate parameter indicative of a residual signal for the prediction, and to generate the upmix coefficients in response to the intermediate parameter, the second intermediate parameter being dependent on the signal cancellation measure.

[0039] This may provide a particularly advantageous implementation and/or performance, and may in particular in many scenario prevent or mitigate numerical issues, artefacts, and/or signal distortions.

[0040] According to an optional feature of the invention, the coefficient processor (107) is arranged to generate the upmix matrix as:

$$\frac{1}{2c} \cdot \frac{1}{g_{11} \cdot g_{22} - g_{12} \cdot g_{21}} \begin{bmatrix} g_{22} - g_{21} + \alpha(g_{11} - g_{12}) & \beta(g_{11} - g_{12}) \\ g_{22} + g_{21} - \alpha(g_{11} + g_{12}) & -\beta(g_{11} + g_{12}) \end{bmatrix}$$

where c is a gain parameter and α and β are parameters dependent on the upmix parameters and the signal cancellation measure, and the parameters $g_{1,1}$, $g_{1,2}$, $g_{2,1}$, and $g_{2,2}$, are dependent on the signal cancellation measure.

[0041] According to an optional feature of the invention, the coefficient processor (107) is arranged to generate the upmix matrix as:

$$\frac{1}{2c} \cdot \frac{1}{g_{11} \cdot g_{22} - g_{12} \cdot g_{21}} \begin{bmatrix} g_{22} - g_{21} + \alpha(g_{11} - g_{12}) & \beta(g_{11} - g_{12}) \\ g_{22} + g_{21} - \alpha(g_{11} + g_{12}) & -\beta(g_{11} + g_{12}) \end{bmatrix}$$

where c is a gain parameter and

$$\alpha = \frac{g_A \cdot g_B \cdot IID + g_C \cdot g_D + (g_C \cdot g_B \cdot e^{j \cdot IPD} + g_A \cdot g_D \cdot e^{-j \cdot IPD}) \cdot ICC \cdot \sqrt{IID}}{g_A{}^2 \cdot IID + g_C{}^2 + 2 \cdot g_A \cdot g_C \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}}$$

$$\beta = \sqrt{\frac{g_B{}^2 \cdot IID + g_D{}^2 + 2 \cdot g_B \cdot g_D \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}}{g_A{}^2 \cdot IID + g_C{}^2 + 2 \cdot g_A \cdot g_C \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}} - |\alpha|^2}$$

with:

$$g_A = g_{11} + g_{12}$$

$$g_B = g_{21} + g_{22}$$

$$g_C = g_{11} - g_{12}$$

$$g_D = g_{21} - g_{22}$$

[0042] In which IID is an Interaural Intensity Differences upmix parameter, ICC is an Interchannel Cross Correlations upmix parameter, and IPD is an Inter-channel Phase Difference upmix parameter; and wherein $\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$ is a dependent on the signal cancellation measure.

[0043] This may provide a particularly advantageous implementation and/or performance, and may in particular in many scenario prevent or mitigate numerical issues, artefacts, and/or signal distortions.

[0044] In some embodiments,

$$\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} \sqrt{1 - z^2} & z \\ -z & \sqrt{1 - z^2} \end{bmatrix}$$

where z is dependent on the signal cancellation measure.

[0045] This may provide a particularly advantageous implementation and/or performance, and may in particular in many scenario prevent or mitigate numerical issues, artefacts, and/or signal distortions.

[0046] According to an optional feature of the invention,

$$\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} 1 & \epsilon_1 \\ \epsilon_2 & 1 \end{bmatrix}$$

where ϵ_1 and ϵ_2 are dependent on the signal cancellation measure.

[0047] This may provide a particularly advantageous implementation and/or performance, and may in particular in many scenario prevent or mitigate numerical issues, artefacts, and/or signal distortions.

[0048] According to an aspect of the invention, there is provided an apparatus for generating an audio data signal, the apparatus comprising: a receiver arranged to receive an audio stereo signal comprising two channel signals; a downmixer arranged to generate a mono audio signal as a combination of the two channel signals in dependence on a set of downmix coefficients; a parameter generator arranged to generate a set of upmix parameters comprising a first parameter indicative of a level difference between the two channel signals, a second parameter indicative of a correlation between the two channel signals, and a third parameter indicative of a phase difference between the two channel signals; a downmix coefficient processor arranged to generate the set of downmix coefficients in dependence on the set of upmix parameters; a data signal generator arranged to generate the audio signal to include the mono audio signal and the set of upmix parameters; a signal cancellation estimator which is arranged to determine a signal cancellation measure from the set of upmix parameters, the signal cancellation measure being indicative of a signal cancellation in a summation of the two channel signals; and wherein the downmix coefficient processor is arranged to generate the set of downmix coefficients in

dependence on the signal cancellation measure

[0049] According to an aspect of the invention, there is provided a method of generating an output audio stereo signal, the method comprising: receiving an audio data signal comprising: a mono audio signal being a downmix of two channel signals of a first audio stereo signal; upmix parameters for the mono audio signal, the set of upmix parameters comprising a first parameter indicative of a level difference between the two channel signals, a second parameter indicative of a correlation between the two channel signals, and a third parameter indicative of a phase difference between the two channel signals; generating coefficients for an upmix matrix from the upmix parameters; generating the output audio stereo signal by applying the upmix matrix to samples of the mono audio signal and an auxiliary mono audio signal; wherein generating the coefficients comprises: determining a signal cancellation measure from the upmix parameters, the signal cancellation measure being indicative of a signal cancellation in a summation of the two channel signals; and determining the coefficients for the upmix matrix in dependence on the signal cancellation measure.

[0050] According to an aspect of the invention, there is provided a method of generating an audio data signal, the method comprising: receiving an audio stereo signal comprising two channel signals; generating a mono audio signal as a combination of the two channel signals in dependence on a set of downmix coefficients; generating a set of upmix parameters comprising a first parameter indicative of a level difference between the two channel signals, a second parameter indicative of a correlation between the two channel signals, and a third parameter indicative of a phase difference between the two channel signals; generating the set of downmix coefficients in dependence on the set of upmix parameters; generating the audio signal to include the mono audio signal and the set of upmix parameters; determining a signal cancellation measure from the set of upmix parameters, the signal cancellation measure being indicative of a signal cancellation in a summation of the two channel signals; and wherein generating the set of downmix coefficients is in dependence on the signal cancellation measure.

[0051] These and other aspects, features and advantages of the invention will be apparent from and elucidated with reference to the embodiment(s) described hereinafter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0052] Embodiments of the invention will be described, by way of example only, with reference to the drawings, in which

FIG. 1 illustrates some elements of an example of an audio apparatus in accordance with some embodiments of the invention;
FIG. 2 illustrates some elements of an example of an audio apparatus in accordance with some embodiments of the invention;
FIG. 3 illustrates an example of a parameter determination flow foreign object an audio apparatus of FIG. 1 or 2;
FIG. 4 illustrates in example of a signal cancellation measure in accordance with some embodiments of the invention;
FIG. 5 illustrates an example of an intermediate parameter as a function of upmix parameters in accordance with some embodiments of the invention;
FIG. 6 illustrates an example of an intermediate parameter as a function of upmix parameters in accordance with some embodiments of the invention;
FIG. 7 illustrates an example of an intermediate parameter as a function of upmix parameters in accordance with some embodiments of the invention;
FIG. 8 illustrates an example of an intermediate parameter as a function of upmix parameters in accordance with some embodiments of the invention; and
FIG. 9 illustrates some elements of a possible arrangement of a processor for implementing elements of an apparatus in accordance with some embodiments of the invention.

DETAILED DESCRIPTION OF SOME EMBODIMENTS OF THE INVENTION

[0053] FIG. 1 and 2 illustrate elements of audio apparatuses in accordance with some embodiments of the invention. The audio apparatus of FIG. 1 may typically be considered to perform a decoding and upmix function/ operation and will accordingly also for brevity be referred to as a decoder. The audio apparatus of FIG. 2 may typically be considered to perform an encoding and downmix function/ operation and will accordingly also for brevity be referred to as an encoder.

[0054] The audio apparatus of FIG. 1 comprises a receiver 101 which is arranged to receive a data signal/ bitstream comprising a downmix mono audio signal that is a downmix of a stereo audio signal which comprises two channel signals, typically corresponding to a left channel signal and a right channel signal. The stereo signal is in the specific example one that has been provided to the encoder of FIG. 2 and downmixed to the mono audio signal by this encoder.

[0055] In addition, the received data signal includes upmix parametric data for upmixing the downmix audio signal. The upmix parametric data may specifically be a set of parameters that indicate relationships between the signals of the two different audio channels of the stereo audio signal, i.e. of the channel signals that are combined into the downmix mono

audio signal. Typically, the upmix parameters may be indicative of time differences, phase differences, level/intensity differences and/or a measure of similarity, such as correlation, between the two channel signals (i.e. between the input left and right signal). Typically, the upmix parameters are provided on a per time and per frequency basis (time frequency tiles). For example, new parameters may periodically be provided for a set of subbands. Parameters may specifically include Interaural Intensity Differences (IIDs), Interaural Level Differences (ILD), Inter-channel Phase Differences (IPDs), Overall Phase Differences (OPDs), Inter-channel Cross Correlations (ICCs), Channel Phase Differences (CPDs) parameters as known from Parametric Stereo encoding (as well as from higher channel encodings).

[0056] Typically, the mono audio signal is an encoded audio signal that has been encoded in accordance with a suitable mono signal encoding standard or approach, and the receiver 101 may decode the received encoded mono audio signal using a decoding approach corresponding to the encoding approach of the encoder.

[0057] The receiver 101 is coupled to a generator 103 which generates an output stereo audio signal corresponding to the stereo audio signal from the downmix signal. The generator 103 is arranged to generate the output stereo audio signal from the mono audio signal and an auxiliary audio signal in dependence on the parametric upmix data. The generator may specifically generate the output stereo audio signal by applying a 2x2 matrix multiplication to the samples of the mono audio signal and the auxiliary audio signal. The coefficients of the 2x2 matrix, also known as an upmix matrix, are determined from the upmix parameters of the upmix parametric data, typically on a time and frequency band basis.

[0058] Typically, the upmixing includes generating an auxiliary audio signal in the form of a decorrelated signal of the mono audio signal. It has been found that by generating a decorrelated signal and mixing this with the mono audio signal, an improved quality of the upmix signal is perceived and therefore decoders have been developed to exploit this. The decorrelated signal is typically generated by a decorrelator 105, such as an all-phase filter that is applied to the mono audio signal. In some cases, the auxiliary signal may be a signal received together with the mono audio signal, in specifically may be a signal generated from a received residual or side signal generated at the encoder side and transmitted to the decoder side.

[0059] In the apparatus of FIG. 1, only a mono audio signal m is received and a decorrelator 105 is used to generate a decorrelated signal d as a decorrelated version of the mono audio signal (typically with the same energy/ level and spectral shape as the mono audio signal). The output stereo audio signal l',r' (where ' indicates that it is the decoder replica of the original input stereo audio signal provided to the encoder) is generated by multiplying the (samples of) the mono audio signal m and the decorrelated signal d by the upmix matrix H to generate the (samples of) the output stereo audio signal l',r':

$$\begin{bmatrix} l' \\ r' \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix}$$

[0060] The decoder of FIG. 1 further comprises a coefficient processor 107 which is arranged to generate the coefficients for the upmix matrix H from the received upmix parameters as will be described in more detail later. In particular, the coefficients for the upmix matrix H may be generated from received IID, ICC, IPD parameters.

[0061] The coefficients of the upmix matrix may in some examples be generated for each sample instant of the signals but is typically generated at a much lower update rate. In such cases, the same coefficients may for example be used for a group/block/segment of samples, or the coefficient processor 107 may for example be arranged to interpolate between determined values. For example, the upmix matrix H may be defined at discrete time points sampled at a lower rate than that which the samples are determined and temporal interpolation may be used to provide more appropriate time varying coefficients.

[0062] FIG. 2 illustrates an example of an apparatus, henceforth also referred to as an encoder, which may generate the audio data signal that may be received by the decoder of FIG. 1.

[0063] In the example, the encoder comprises a receiver 201 which receives an input stereo audio signal that is to be encoded and transmitted. The stereo audio signal includes two channel signals l,r that are fed to a downmixer 203 which is arranged to generate a mono audio signal comprising the majority of the signal energy of the channel signals l,r as well as typically a residual signal or side signal s.

[0064] The encoder further comprises an upmix parameter generator 205 which is arranged to determine upmix parameters characterizing properties of the input channel signals l,r. In particular, the upmix parameter generator 205 is arranged to generate IID, ICC, IPD parameters.

[0065] The encoder further comprises a downmix coefficient processor 207 which is arranged to determine the downmix coefficients for the downmix based on the upmix parameters (which accordingly may also be considered to be downmix parameters). The upmix/downmix parameters may specifically reflect how the downmix is performed in the encoder and how the upmix should be performed in the decoder.

[0066] The encoder further comprises a data signal generator 209 which is arranged to receive at least the downmix mono signal m and the upmix parameters and to generate the data signal to include these. The data signal generator 209 may specifically be arranged to generate suitable data representing these signals and parameters and may thus include

suitable encoder functions, bitstream formatting functions, etc. as will be well known to the skilled person. In many embodiments, the data signal generator 209 is arranged to generate the data signal to not include the residual/side signal s but in some embodiments this signal may also be encoded and included in the data signal. In such cases, the residual/side signal s is typically encoded at a much lower data rate than the mono audio signal m reflecting the reduced energy and reduced perceptual impact on the stereo signal generated at the decoder side.

[0067] An approach for Parametric Stereo is defined by Moving Pictures Expert Group (MPEG) of the International Organization for Standardization/International Electrotechnical Commission in ISO/IEC 23003-3:2020, Information technology - MPEG audio technologies - Part 3: Unified speech and audio coding.

[0068] In the standard, the upmix (for each parametric stereo band) is described as a generalized 2x2 mixing of a downmix mono signal m and a decorrelated signal d:

$$\begin{bmatrix} l' \\ r' \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix}$$

[0069] The decorrelated signal d is derived by applying a reverberant type processing on the signal m. In more detail, the upmix can be described as (for convenience the notation uses the opposite sign of α than used in the standard specification):

$$\begin{bmatrix} l' \\ r' \end{bmatrix} = \frac{1}{2c} \begin{bmatrix} 1+\alpha & \beta \\ 1-\alpha & -\beta \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix}$$

[0070] The approach uses a set of intermediate parameters α, β and c that are all a function of the IID, ICC and IPD parameters. FIG. 3 shows the parameter flowchart. From the IID, ICC and IPD parameters the intermediate parameters c, α and β are calculated. These are then used to calculate the entries of the H matrix.

[0071] The c, α and β parameters are specifically defined as:

$$\alpha = \frac{IID - 1 + 2j \cdot \sin{(IPD)} \cdot ICC \cdot \sqrt{IID}}{IID + 1 + 2 \cdot \cos{(IPD)} \cdot ICC \cdot \sqrt{IID}}$$

$$\beta = \frac{2 \cdot \sqrt{IID \cdot (1 - ICC^2)}}{IID + 1 + 2 \cdot \cos{(IPD)} \cdot ICC \cdot \sqrt{IID}}$$

$$c = \min{\left(\sqrt{\frac{IID + 1}{IID + 1 + 2 \cdot \cos{(IPD)} \cdot ICC \cdot \sqrt{IID}}}, c_{max} \right)}$$

[0072] The rationale behind this upmix can be seen by dissecting the inverse of this upmix matrix, i.e. the downmix matrix that is used by the encoder/ downmixer 203:

$$\begin{bmatrix} m \\ d' \end{bmatrix} = c \begin{bmatrix} \frac{1}{1-\alpha} & -\frac{1}{1-\alpha} \\ \frac{}{\beta} & \frac{}{\beta} \end{bmatrix} \begin{bmatrix} l \\ r \end{bmatrix}$$

[0073] The dissection shows the corresponding structure:

$$\begin{bmatrix} m \\ d' \end{bmatrix} = c \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\beta} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\alpha & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} l \\ r \end{bmatrix}$$

[0074] In the first step (the most right hand matrix multiplication), traditional mid and side signals are formed as

respectively a sum signal *l+r* and a difference signal *l-r*.

[0075] In the second step, the best possible (least squares) prediction of the side signal from the mono signal is realized. Thus, the parameter value α is a complex value that is determined to provide the optimal predication of the difference signal from the sum signal, and thus specifically of the difference signal from the mono audio signal m that is generated (noting that the remaining matrix multiplication retain the m signal as a direct sum signal *l+r*).

[0076] The residual signal of the difference signal is then scaled in the third step to ensure that both signals *m* and *d'* have equal signal power. Thus, the parameter value β is a gain parameter that adapts the level of the decorrelated signal d' to have a signal power corresponding to that of the mid/sum signal *m*. It should be noted that the residual signal is uncorrelated from the mid (mono) signal (due to the prediction using the parameter α).

[0077] The last parameter c is a coefficient which is used to maintain signal power in the downmix and specifically it is set to ensure that $c \cdot (l+r)$ has approximately the same power as the sum of the signal powers of left and right channel signals. The value is clipped/limited to a value *of c_{max}* in order to maintain a practical range of values.

[0078] The standardized approach for parametric stereo encoding and decoding provides for a very advantageous operation, and in particular provides a high audio quality to data rate ratio/trade-off. However, the Inventor has realized that in some situations, and in particular for some signals, the standardized approach leads to less than optimal encoding, and indeed may lead to significant degradation and distortion in some particular situations. As will be described in the following, the Inventor has furthermore realized that such effects and scenarios may be mitigated or reduced by performing specific modified operations.

[0079] The Inventor has in particular realized that issues may occur when the channel signals of the input stereo audio signal are identical or are identical except for being 180° out of phase. In such cases, the intermediate parameters may approach values that result in numerical problems and issues in the processing resulting in degradations and distortions to the resulting decoded stereo audio signal. In particular, in these scenarios, values of the encoding and/or decoding may approach infinite values that cannot be appropriately represented.

[0080] For example, in the case of the channel signals being identical but 180° out of phase, *l=-r*, the upmix parameters will have the following values:

$$IID=1,$$

$$ICC=1,$$

$$IPD=\pi$$

[0081] However, this results in parameters:

$$\alpha = \frac{0}{0}$$

$$\beta = \frac{0}{0}$$

$$c = \min\left(\sqrt{\frac{2}{0}}, c_{max}\right)$$

[0082] As a result, all parameters are troublesome to handle numerically resulting both in encoder and decoder problems.

[0083] Thus, the Inventor has realized that when the channel signals are substantially identical but 180° out of phase (l=-r), all parameters become numerically unstable. In addition, the downmix signal may start to include (time-frequency) gaps in which the sum signal may essentially have no energy (i.e. a zero signal), and this may make it extremely difficult to reconstruct a stereo signal.

[0084] In the encoder and decoder apparatuses of FIGs. 2 and 1, an approach is adopted which in many scenarios may mitigate and address such issues.

[0085] In the approach, the coefficient generator 107 is arranged to determine intermediate parameters from the upmix

parameters and then to determine the upmix matrix coefficients from the intermediate parameters. The intermediate parameters may correspond closely to those applied in ISO/IEC 23003-3:2020 but may be modified in particular for some specific inter-signal properties.

[0086] Further, in the decoder of FIG. 1, the coefficient processor 107 is arranged to determine a signal cancellation measure from the upmix parameters where the signal cancellation measure is indicative of a signal cancellation in a summation of the two channel signals of the original input stereo audio signal to the encoder. The properties of these channel signals of the original input stereo audio signal are represented by the upmix parameters. Indeed, the upmix parameters, such as specifically the IID, IPD, and ICC, are dependent on the input channel signals, and in particular on the relative differences between the input channel signals. Indeed, typically, the upmix parameters are dependent only on properties of the channel signals, and specifically the relative properties of the channel signals.

[0087] Accordingly, the coefficient processor 107 may on the basis of the information provided by the upmix parameters indicating relative properties of the channel signals proceed to determine how much signal cancellation would result when adding the channel signals together. The signal cancellation measure may be indicative of the energy/power/amplitude (square root of power)/ signal level for the sum signal 1+r of the channel signals relative to the sum/combination of the energy/power/amplitude (square root of power)/ signal level of the two individual channel signals 1 and r.

[0088] For example, in the case where 1=-r, i.e. the two signals are identical but have a phase offset of π, the sum signal 1+r=0, i.e. the two channel signals will cancel completely if added together/summed. Further, for this specific case IID=1, ICC=1, IPD= π, and thus the situation can be detected by evaluating these upmix parameters.

[0089] In the case where 1=r, i.e. the two signals are identical, the sum signal of 1+r=21=2r and accordingly there is a negative signal cancellation with indeed the signal energy of the sum signal being four times that of the left or right signal individually. Further, for this specific case IID=1, ICC=1, IPD= 0, and thus the situation can be detected by evaluating these upmix parameters.

[0090] These two examples may be considered to correspond to the extreme cases of signal cancellation.

[0091] As mentioned, the signal cancellation measure may be indicative of a sum signal energy measure determined from the upmix parameters where the sum signal energy measure may be indicative of an energy level of a sum signal that is a summation of the channel signals relative to combination/summation of an energy level of the individual channel signals.

[0092] As a low complexity example, a signal cancellation measure may be generated to reflect a difference between the received upmix parameters and the upmix parameters corresponding to the extreme scenarios of in-phase or out-of-phase signal cancellation. For example, the signal cancellation measure may be determined based on a comparison of the received upmix parameters relative to the upmix parameters that correspond to the maximum cancellation and/or the minimum (the inverse) cancellation (i.e. amplification) of the sum signal.

[0093] In many embodiments, the signal cancellation measure may be determined as a function of the upmix parameters. For example, the signal cancellation measure may be determined as:

$$R = \begin{cases} IID \cdot ICC \cdot \cos(IPD) & IID \leq 1 \\ \dfrac{ICC \cdot \cos(IPD)}{IID} & IID > 1 \end{cases}$$

[0094] This value will attain the value of 1 and -1 respectively in the extreme situations of complete signal cancellation and provide increasingly different values for other values of the upmix parameters. It may thus provide a suitable indication of how close the stereo signal is to respectively a scenario where the channel signals cancel out in the sum signal 1+r or in the difference signal 1-r (corresponding to a maximum negative signal cancellation for the sum signal).

[0095] For the specific signal cancellation measure, the closer the absolute value is to 1, the closer the input stereo audio signal is to a scenario in which the input channel signals cancel out in either the sum signal or the difference signal. Further, the sign of the signal cancellation measure indicates in which of these signals the cancellation occurs.

[0096] Another example of a signal cancellation measure is the following:

$$R = \frac{2 \cdot ICC \cdot \cos(IPD) \cdot \sqrt{IID}}{IID + 1} = \frac{2 \cdot ICC \cdot \cos(IPD)}{\sqrt{IID} + \dfrac{1}{\sqrt{IID}}}$$

[0097] This signal cancellation measure has some properties that are particular advantageous in many scenarios and embodiments. Since

$$0 \le ICC \le 1,$$

$$-1 \le \cos{(IPD)} \le 1,$$

and

$$\sqrt{IID} + \frac{1}{\sqrt{IID}} \ge 2,$$

it follows that

$$-1 \le R \le 1.$$

[0098]   Further, R approaches -1 only for highly correlated out-of-phase channel signals, i.e. when the two channel signals cancel each other in a sum signal. R=-1 accordingly indicates that if the two channel signals were added together, they would cancel, resulting in a zero signal. Also, R approaches 1 only for highly correlated in-phase channel signals, i.e. when the two channel signals cancel each other in a difference signal. R=1 thus indicates that if one of the two channel signals was subtracted from the other, then they would cancel resulting in a zero signal.

[0099]   The specific signal cancellation measure R is particularly advantageous in many scenarios. Considering the ratio of the power of the unadjusted mid- and side-signals:

$$\frac{\|m\|^2}{\|s\|^2} = \frac{IID + 1 + 2 \cdot ICC \cdot \cos{(IPD)} \cdot \sqrt{IID}}{IID + 1 - 2 \cdot ICC \cdot \cos{(IPD)} \cdot \sqrt{IID}}$$

there may be identified two terms that can potentially cancel out: $IID + 1$ and $2 \cdot ICC \cdot \cos{(IPD)} \cdot \sqrt{IID}$. The ratio of these two factors may provide a particularly advantageous signal cancellation measure for indicating how close the current scenario is to the problematic in-phase and out-of-phase scenarios, and in particular how close the current scenario is to a full signal cancellation in either the sum of the channel signals or the difference of the channel signals. The specific signal cancellation measure described above accordingly provides a particularly advantageous measure in many embodiments.

[0100]   FIG. 4 illustrates how the R value above varies with the upmix parameters. As can be seen, it provides a good indication of when signal cancellation may occur.

[0101]   The following description will focus on the use of this specific signal cancellation measure which has been found to provide particularly advantageous performance and allowing improved encoding, decoding, and rendering. However, it will be appreciated that other values and formulas for determining the signal cancellation measure from the upmix parameters may be used in other embodiments.

[0102]   The coefficient processor 107 is arranged to determine the coefficients for the upmix matrix in dependence on the signal cancellation measure. In particular, the coefficient processor 107 may be arranged to modify operation such that the operation is adapted to compensate/modify the operation in scenarios where signal cancellation may occur in a sum signal and/or a difference signal.

[0103]   The coefficient processor 107 may specifically be arranged to modify the determination of the coefficients for situations approaching signal cancellation such that the numerical problems are mitigated, and in particular such that the determination, and in particular intermediate parameters, do not approach problematic values, and specifically that they do not approach infinity. The coefficient processor 107 may specifically be arranged to adapt the operation/ equations for determining the coefficients such that the required dynamic ranges of intermediate calculations and intermediate parameters may be more constrained thereby allowing practical applications and reducing the numerical challenges and issues.

[0104]   In the example, the coefficient processor 107 is arranged to adapt the upmix coefficients (H_{11}, H_{21}) for the mono audio signal to deviate from coefficients for the mono audio signal being a sum signal of the channel signals for the signal cancellation measure meeting a first signal cancellation requirement. Indeed, in the case where the mono audio signal is a sum signal, the optimum coefficients for determining the channel signals of the output stereo signals will be given as a function of the upmix parameters. However, the coefficient processor 107 may be arranged to generate the coefficients such that they differ and deviate from such values in case of the signal cancellation measure meeting a requirement. Specifically, the coefficient processor 107 may be arranged to differ from these values if the signal cancellation measure is indicative of a signal cancellation in the sum and/or difference signal above a given threshold.

[0105] The cancellation requirement may require the signal cancellation measure to be indicative of a signal cancellation of the sum signal above a threshold. Thus, in contrast to existing systems where the calculation of the coefficients in the encoder is based on the received mono audio signal (the mid signal) being a sum signal/downmix of the original input channel signals, the coefficient processor 107 in the approach of FIG. 1 proceeds to, for at least some values, deviate from the values that are optimum for the mono audio signal being a sum signal.

[0106] In many embodiments, the degree of deviation from the coefficients for a mono signal being a sum signal of the channel signals may depend on the signal cancellation measure, and may specifically be a monotonically increasing function of a degree of signal cancellation in a sum of the channel signals. Thus, as the signal cancellation in the sum signal increases, the determination of the coefficients is modified to increasingly deviate from the coefficients that would be determined for the mono audio signal being a sum signal. In particular, for an increasing signal cancellation (and thus reduced signal level) of a sum signal, the upmix coefficients determined for a mono audio signal being a direct sum signal may become increasingly large and may indeed approach infinity or be non-defined. However, in the described approach, the signal cancellation measure is determined and used to control the coefficient determination such that this is mitigated and prevented, and thus coefficients are determined which deviate from the potentially ideal coefficients for a sum signal, but which have reduced numerical issues.

[0107] In many embodiments, the coefficient processor 107 is arranged to increase a deviation of the upmix coefficients for the mono audio signal from the coefficient values for the mono audio signal being a sum signal of the channel signals for the signal cancellation measure being indicative of an increasing signal cancellation in a sum signal for a sum signal of the channel signals. Thus, the deviation from reference coefficient values increases for the signal cancellation measure being indicative of an increasing signal cancellation in the sum of the channel signals where the reference coefficient values are (optimum) coefficients for the mono audio signal being a sum signal.

[0108] In many embodiments, the coefficient processor 107 is arranged to increase a deviation of upmix coefficients for the mono audio signal from the coefficient values for the mono audio signal being a difference/subtraction signal of the channel signals for the signal cancellation measure being indicative of an increasing signal cancellation in a difference/subtraction signal for a difference/subtraction signal of the channel signals. Thus, the deviation from reference coefficient values increases for the signal cancellation measure being indicative of an increasing signal cancellation in the difference between of the channel signals where the reference coefficient values are (optimum) coefficients for the mono audio signal being a sum signal.

[0109] In many embodiments, the coefficient processor 107 is arranged to determine the upmix coefficients for the mono audio signal to be coefficients for the mono audio signal being a sum signal of the channel signals for the signal cancellation measure meeting a second signal cancellation requirement. In particular, for the signal cancellation measure being indicative of the signal cancellation in the sum signal of the channel signals being below a threshold, the coefficient processor 107 may generate the coefficients based on the mono audio signal being a sum signal, and indeed in some embodiments, the coefficient processor 107 may in this scenario determine the coefficients substantially as defined in e.g. ISO/IEC 23003-3:2020.

[0110] In many embodiments, the coefficient processor 107 may be arranged to generate the upmix coefficients as (optimum) coefficients for the mono audio signal being a sum signal of the input channel signals when the signal cancellation measure is indictive of a low signal cancellation in the sum and/or difference signal and to generate the upmix coefficients to deviate from upmix coefficients that are optimum for the mono audio signal being a sum signal when the signal cancellation measure is indicative of high signal cancellation.

[0111] In most embodiments, the above described approach may be applied to all the coefficients of the upmix matrix, and indeed all of the coefficients may be determined to deviate from the coefficients that would apply to a sum signal for at least some values of the signal cancellation measure. However, it will be appreciated that in some embodiments, the approach may only be applied to a subset of one, two, or three of the coefficients. Specifically, in some embodiments, the approach may only be applied to the coefficients for the mono audio signal or for the coefficients for the auxiliary audio signal.

[0112] A similar approach may be used for the upmix coefficients (H_{12}, H_{22}) for the auxiliary signal but in this case with the deviation being introduced for signal cancellation in the difference signal of the channel signals, corresponding to a maximum negative cancellation in the sum signal (i.e. maximum level increase in the sum signal).

[0113] The coefficients for the mono audio signal being a sum signal of the channel signals, henceforth for brevity also referred to as reference coefficients, may specifically be optimum coefficients for generating the output stereo audio signal from a mono audio signal being a sum signal of the input channel signals and an auxiliary signal, which specifically may be a decorrelated version of the mono audio signal.

[0114] The reference coefficients (the coefficients for the mono audio signal being a sum signal 1+r) may specifically be the coefficients determined in accordance with the approach defined by ISO/IEC 23003-3:2020, i.e. the reference coefficients may specifically be determined as:

$$\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} = \frac{1}{2c} \begin{bmatrix} 1+\alpha & \beta \\ 1-\alpha & -\beta \end{bmatrix}$$

where

$$\alpha = \frac{IID - 1 + 2j \cdot \sin(IPD) \cdot ICC \cdot \sqrt{IID}}{IID + 1 + 2 \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}}$$

$$\beta = \frac{2 \cdot \sqrt{IID \cdot (1 - ICC^2)}}{IID + 1 + 2 \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}}$$

$$c = \min\left(\sqrt{\frac{IID + 1}{IID + 1 + 2 \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}}}, c_{max} \right)$$

in which IID is an Interaural Intensity Differences upmix parameter, ICC is an Interchannel Cross Correlations upmix parameter, and IPD is an Inter-channel Phase Difference upmix parameter.

[0115] The values IID, ICC, and IPD may specifically be determined in accordance with ISO/IEC 23003-3:2020. In particular, the upmix parameters IID, ICC, and IPD may be determined as:

$$IID = \frac{\langle l, l \rangle}{\langle r, r \rangle}$$

$$ICC = \left| \frac{\langle l, r \rangle}{\sqrt{\langle l, l \rangle \langle r, r \rangle}} \right|$$

$$IPD = \angle \left\{ \frac{\langle l, r \rangle}{\sqrt{\langle l, l \rangle \langle r, r \rangle}} \right\}$$

[0116] With the Hilbert inner product defined for complex-valued coefficients as:

$$\langle x, y \rangle = \sum_{\forall i} x_i \cdot y_i^*$$

[0117] The summation over i could refer to a group of frequency domain coefficients, or could refer to a summation both over a window in time and frequency in case of (complex-valued) subband representations.

[0118] It should be noted that the upmix parameters IID, ICC, and IPD depend only on the input signals and are not modified based on the downmix, signal cancellation, or indeed any part of the processing or encoding of the input stereo signal. This is highly beneficial in many scenarios and embodiments. Indeed, a particular advantage is that the perceptual sensitivities of the parameters are well known and understood.

[0119] Thus, the coefficient processor 107 may be arranged to determine coefficients that deviate from the optimum coefficients that would be applied to a mono audio signal being a sum signal of the channel signals. However, the deviation is dependent on the signal cancellation measure and thus may specifically be targeted at the specific situations where signal cancellation will occur in the sum signal. Thus, although the approach may accordingly deviate from the theoretical or optimum processing, it may in practice mitigate and often remove numerical issues and the difficulties and degradations associated therewith. This may provide improved audio quality, improved robustness, and/or reduced degradation/

artefacts in many scenarios.

[0120] In some embodiments, this approach may be used with an encoder that fixedly generates the mono audio signal as a sum signal, and thus a suboptimal upmixing may be performed. However, the benefit of mitigating and reducing numerical problems and issues may often substantially outweigh the effects of modifying the upmix coefficients, especially as this can be limited to specific scenarios in which the numerical issues would be highly detrimental and cause substantial distortion.

[0121] However, in many embodiments and systems, the encoder may be arranged to also determine the downmix coefficients to reflect the differences in the upmix coefficients, i.e. the deviation of the upmix coefficients may be complemented by a corresponding operation at the encoder such that the generated mono audio signal (and possibly a residual signal) is modified in scenarios where the signal cancellation may exceed a given level. Thus, in many embodiments, the downmixing at the encoder and the upmixing at the decoder may be complementary and both may be dependent on the signal cancellation in a sum/difference of the two input channel signals.

[0122] For example, in some embodiments, the mono audio signal may in the encoder be generated as a sum signal of the input channel signals, i.e. as *m=l+r.* However, in the specific scenario where l=-r, the signal may be modified to not be determined simply as the sum signal. In particular, if the signal cancellation in the sum signal increases towards complete cancellation, the mono audio signal may be generated to include a component of the difference signal s=1-r. As a result, a minimum level of the mono audio signal may be retained.

[0123] However, as the generation of the mono audio signal has changed, the decoder may be arranged to complement and compensate for this changed operation. In particular, for situations where the encoder modifies the operation to prevent signal cancellation, the upmixing at the decoder side may be modified correspondingly.

[0124] A similar approach may be used for signal cancellation in the difference signal. In this case, when the input channel signals are such that the difference signal approaches l=r such that 1-r will approach zero, the generation of the difference signal may be modified to include an element of the sum signal thereby preventing that the signal level falls below a given value.

[0125] Thus, in many embodiments, the encoder may modify the downmixing depending on the signal cancellation in a sum and/or difference signal for the two channel signals, and in particular may modify the downmixing coefficients of the downmixing matrix generating the mono audio signal, and optionally a side or auxiliary signal.

[0126] The encoder of FIG. 2 accordingly also includes a signal cancellation estimator 211 which receives the upmix parameters determined by the upmix parameter generator 205. The signal cancellation estimator 211 is then arranged to determine a signal cancellation measure from the set of upmix parameters where the signal cancellation measure is again indicative of a signal cancellation in a summation of the two channel signals of the input stereo signal. The signal cancellation estimator 211 may specifically be arranged to determine the signal cancellation measure using the same algorithm, formulas, and approach as the coefficient processor 107 of the decoder. Thus, the description provided on the generation of the signal cancellation measure by the coefficient processor 107 apply equally (mutatis mutandis) to the determination of the signal cancellation measure signal generated by the cancellation estimator 211.

[0127] The signal cancellation estimator 211 may accordingly generate a signal cancellation measure which is identical to that generated by the coefficient processor 107 of the decoder. The encoder and decoder may accordingly in many embodiments generate the same signal cancellation measure, and thus may be arranged to use coordinated and complementary approaches for generating the coefficients for respectively the downmix matrix of the encoder and the upmix matrix of the encoder. Indeed, in many embodiments, the coefficients may be generated such that the two matrices are the inverse of each other thereby resulting in an overall downmix and upmix operation that restores the original input stereo signal.

[0128] In the following, specific approaches will be described that may provide particularly advantageous implementations. The described approaches may typically provide a compatibility with existing Standards and Technical Specifications, such as the ISO/IEC 23003-3:2020 specification.

[0129] As previously mentioned, an encoder matrix corresponding to the decoder matrix of ISO/IEC 23003-3:2020 can be represented by:

$$\begin{bmatrix} m \\ d' \end{bmatrix} = c \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\beta} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\alpha & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} l \\ r \end{bmatrix}$$

[0130] The parameters α, β, and c are parameters determined from the upmix parameters to provide specific functions/compensate for specific properties of the channel signals. It should be noted that the signal d' is typically not explicitly calculated in the encoder but the parameters α, β, and c involved in generating this signal are determined.

[0131] Specifically, the parameter value α is determined to be generate a prediction of the difference signal 1-r from the sum signal l+r. It is thus a parameter that indicates a prediction of the difference signal from the sum signal.

[0132] The parameter β is a gain parameter which adapts the gain of the decorrelated signal d' to match that of the mono audio signal m. Thus, the parameter β is determined to indicate the relative difference (and specifically the ratio) between energies/levels/ amplitudes of the residual signal resulting from the prediction and the generated mono audio signal.

[0133] Finally, the parameter is determined to adjust the overall gain/energy of the mono audio signal.

[0134] In some embodiments of the approach of the apparatuses of FIG. 1 and 2, such a downmix matrix may be modified to include an additional matrix multiplication, such as for example by adding an additional gain matrix that is multiplied with the sum and difference signals resulting from the first matrix multiplication, e.g.:

$$\begin{bmatrix} m \\ d' \end{bmatrix} = c \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\beta} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\alpha & 1 \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} l \\ r \end{bmatrix}$$

[0135] The gains/coefficients g_{11}, g_{12}, g_{21} and g_{22} of the gain matrix may then be determined to compensate for signal cancellation in the sum signal and difference signal respectively. Accordingly, the gains/coefficients may be determined based on a signal cancellation measure that is determined in the encoder and which reflects the signal cancellation in the sum and/or difference signals for the input channel signals.

[0136] Further, the gain coefficients may be determined as a function of the upmix parameters/ parametric stereo parameters. These values are only dependent on the input signals and represent properties of the input stereo audio signal. In particular, the upmix parameters are not dependent on the output mono audio signal but can be determined directly from the input stereo audio signal without any consideration of any other signals.

[0137] Specifically, the encoder may be arranged to determine the upmix parameters ICC, IID, and IPD from the input stereo audio signal, i.e. from the input channel signals. It may then determine the gains/coefficients of the gain matrix from the upmix parameters. Specifically, the gains/coefficients may be determined such that for the input channel signals being substantially identical but out of phase, corresponding to a high signal cancellation for the sum signal, the gain matrix multiplication results in some of the difference signal being added to the mid signal, i.e. the gain matrix multiplication may results in the sum signal being modified to include some of the difference signal thereby preventing a full signal cancellation in the sum signal.

[0138] Similarly, the gains/coefficients may be determined such that for the input channel signals being substantially identical and in phase, corresponding to a high signal cancellation for the difference signal, the gain matrix multiplication results in some of the sum signal being added to the difference signal, i.e. the gain matrix multiplication may results in the difference signal being modified to include some of the sum signal thereby preventing a full signal cancellation for the difference signal.

[0139] The gains may (also) be determined as a function of the upmix/ parametric stereo parameters thereby allowing them to be determined equally at the encoder and decoder side.

[0140] The matrix can be condensed into a single matrix:

$$c \begin{bmatrix} g_{11} + g_{12} & g_{11} - g_{12} \\ \frac{-\alpha(g_{11}+g_{12})+g_{21}+g_{22}}{\beta} & \frac{-\alpha(g_{11}-g_{12})+g_{21}-g_{22}}{\beta} \end{bmatrix},$$

[0141] For the case of

$$G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} 1 & \epsilon_1 \\ \epsilon_2 & 1 \end{bmatrix}$$

this can be further simplified to:

$$c \begin{bmatrix} 1 + \epsilon_1 & 1 - \epsilon_1 \\ \frac{-\alpha(1+\epsilon_1)+\epsilon_2+1}{\beta} & \frac{-\alpha(1-\epsilon_1)+\epsilon_2-1}{\beta} \end{bmatrix},$$

[0142] In some embodiments, the downmix coefficient processor 207 may be arranged to determine the gains such that for scenarios in which signal cancellation does not occur to a significant degree in the sum or difference signals (as indicated by the signal cancellation measure), the matrix may be determined to closely resemble the unity matrix ($g_{11} = g_{22}$ = 1, $g_{12} = g_{21} = 0$).

[0143] The coefficient processor 107 may then for a scenario in which significant signal cancellation occurs in the difference signal 1-r (IID≈1, ICC≈1, IPD≈0), determine the gain matrix to have the following properties:

$$g_{11} \approx 1$$

$$g_{12} \approx 0$$

$$g_{21} \neq 0$$

$$g_{22} \approx 1$$

[0144] In this case, the downmix operation is modified to such that some of the sum signal is mixed into the different signal.

[0145] The coefficient processor 107 may then for a scenario in which significant signal cancellation occurs in the difference signal 1-r (IID≈1, ICC≈1, IPD≈π), determine the gain matrix to have the following properties:

$$g_{11} \approx 1$$

$$g_{12} \neq 0$$

$$g_{21} \approx 0$$

$$g_{22} \approx 1$$

[0146] In this case, the downmix operation is modified to such that some of the difference signal is mixed into the sum signal to generate the mono audio signal.

[0147] Alternatively or additionally, the encoder may for highly correlated out-of-phase signals modify the downmix coefficients such that some of the difference signal 'leaks' (is added) to the sum signal. This may ensure that the situation where the sum signal/ mono audio signal diminishes is prevented.

[0148] For other scenarios, the downmixing may remain close to the approach of e.g. the ISO/IEC 23003-3:2020 specification.

[0149] For the decoder, the upmix matrix can be realized by inverting each matrix of the above described downmix:

$$\begin{bmatrix} m \\ d' \end{bmatrix} = c \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\beta} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\alpha & 1 \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} l \\ r \end{bmatrix}$$

Inversion leads to:

$$\begin{bmatrix} l' \\ r' \end{bmatrix} = \frac{1}{2c} \cdot \frac{1}{g_{11} \cdot g_{22} - g_{12} \cdot g_{21}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} m \\ d' \end{bmatrix}$$

which can be written as a single matrix as:

$$\begin{bmatrix} l' \\ r' \end{bmatrix} = \frac{1}{2c} \cdot \frac{1}{g_{11} \cdot g_{22} - g_{12} \cdot g_{21}} \begin{bmatrix} g_{22} - g_{21} + \alpha(g_{11} - g_{12}) & \beta(g_{11} - g_{12}) \\ g_{22} + g_{21} - \alpha(g_{11} + g_{12}) & -\beta(g_{11} + g_{12}) \end{bmatrix} \begin{bmatrix} m \\ d' \end{bmatrix}$$

[0150] With the above definition of

$$G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} 1 & \epsilon_1 \\ \epsilon_2 & 1 \end{bmatrix},$$

[0151] This simplifies to:

$$\begin{bmatrix} l' \\ r' \end{bmatrix} = \frac{1}{2c} \cdot \frac{1}{1 - \epsilon_1 \cdot \epsilon_2} \begin{bmatrix} 1 - \epsilon_2 + (1 - \epsilon_1)\alpha & (1 - \epsilon_1)\beta \\ 1 + \epsilon_2 - (1 + \epsilon_1)\alpha & -(1 + \epsilon_1)\beta \end{bmatrix} \begin{bmatrix} m \\ d' \end{bmatrix}$$

[0152] For the unity matrix $G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ this upmix reduces to the traditional PS prediction upmix.

[0153] The generalized equations for α and β are given below. It is noted that these simplify for the different gain matrices G and G^{-1} as described above.

$$\alpha = \frac{g_A \cdot g_B \cdot IID + g_C \cdot g_D + (g_C \cdot g_B \cdot e^{j \cdot IPD} + g_A \cdot g_D \cdot e^{-j \cdot IPD}) \cdot ICC \cdot \sqrt{IID}}{g_A{}^2 \cdot IID + g_C{}^2 + 2 \cdot g_A \cdot g_C \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}}$$

$$\beta = \sqrt{\frac{g_B{}^2 \cdot IID + g_D{}^2 + 2 \cdot g_B \cdot g_D \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}}{g_A{}^2 \cdot IID + g_C{}^2 + 2 \cdot g_A \cdot g_C \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}} - |\alpha|^2}$$

and with:

$$g_A = g_{11} + g_{12}$$

$$g_B = g_{21} + g_{22}$$

$$g_C = g_{11} - g_{12}$$

$$g_D = g_{21} - g_{22}$$

[0154] In the approach, the parameter c is a gain parameter/ coefficient that in many embodiments may be set to a suitable value by the decoder, and specifically it may be a design parameter that can be set in accordance with any suitable algorithm or criterion.

[0155] For example, in some embodiments similar and detailed formulas may be used to determine the gain parameter c from received upmix parameters, and further the gain coefficient c may be dependent on the signal cancellation measure. However, in other cases (e.g. when backwards compatibility towards mono is not considered relevant), c may simply be set to a constant value, e.g. c=0.5.

[0156] The values of the gain matrix $G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$ are dependent on the signal cancellation measure but it will be appreciated that the exact dependency will depend on the specific preferences and requirements of the individual embodiment and application.

[0157] FIG. 5 and 6 illustrate an example of the absolute difference between the intermediate parameter α determined from the equations above and from the equations of ISO/IEC 14496-3:2005.

[0158] FIG. 7 and 8 illustrate an example of the absolute difference between the intermediate parameter β determined from the equations above and from the equations of ISO/IEC 14496-3:2005.

[0159] As can be seen, the approach allows for the deviation from the parameter values of ISO/IEC 14496-3:2005 to mainly be restricted to scenarios where the upmix parameters indicate that substantial signal cancellation occurs.

[0160] In some embodiments, the gain matrix may advantageously be given by:

$$G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} \cos(\gamma) & \sin(\gamma) \\ -\sin(\gamma) & \cos(\gamma) \end{bmatrix}$$

with the decoder inverse matrix:

$$G^{-1} = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) \\ \sin(\gamma) & \cos(\gamma) \end{bmatrix}$$

[0161] The parameter-dependent gain matrix G can then be established by defining some function $\gamma = f(R)$. For example:

$$\gamma = \gamma_{max} \cdot |R|^m$$

where γ_{max} is a value denoting the maximum allowed mixing, e.g. $\cdot\, \gamma_{max} = \frac{\pi}{8}$, and m is a suitable value, typically a high value (e.g. 4) that may ensure that modifications to a traditional approach only start to occur close to the critical scenarios.
[0162] In some embodiments, the gain matrix may advantageously be given by:

$$G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} \sqrt{1-z^2} & z \\ -z & \sqrt{1-z^2} \end{bmatrix}$$

with inverse matrix:

$$G^{-1} = \begin{bmatrix} \sqrt{1-z^2} & -z \\ z & \sqrt{1-z^2} \end{bmatrix}$$

[0163] Similarly, $z = f(R)$, e.g., $z = z_{max} \cdot |R|^m$, with e.g. $z_{max} = 1$ and $m=4$.
[0164] The above gain matrix definitions are symmetric, resulting in easily invertible matrices. However, that has a small disadvantage that there is unnecessary power loss for the in-phase and out-of-phase cases, e.g. if z is small in option A2, this also means that the mid signal is (unnecessarily) scaled with a factor smaller than 1. An alternative asymmetric matrix can be defined as:

$$G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} 1 & \epsilon_1 \\ \epsilon_2 & 1 \end{bmatrix}$$

with inverse matrix:

$$G^{-1} = \frac{1}{1 - \epsilon_1 \epsilon_2} \begin{bmatrix} 1 & -\epsilon_1 \\ -\epsilon_2 & 1 \end{bmatrix}$$

[0165] The weights ϵ_1 and ϵ_2 can for example be defined as a function of the upmix parameters as following:

$$\epsilon_1 = \begin{cases} |R|^m & \text{for} \quad R < 0 \\ 0 & \text{for} \quad R \geq 0 \end{cases}$$

$$\epsilon_2 = \begin{cases} 0 & \text{for} \quad R < 0 \\ R^m & \text{for} \quad R \geq 0 \end{cases}$$

[0166] It is noted that in this case the inverse matrix further simplifies as the determinant of the matrix G always equals 1:

$$G^{-1} = \begin{bmatrix} 1 & -\epsilon_1 \\ -\epsilon_2 & 1 \end{bmatrix}$$

[0167] It will be appreciated that other approaches and gain matrices may be used in other embodiments.

[0168] In the described approach, the coefficient processor 107 may accordingly proceed to generate an intermediate parameter α in dependence on the upmix parameters and the signal cancellation measure (as the gains g are dependent on the signal cancellation measure). The intermediate parameter α is indicative of the prediction of a difference signal of the channel signals from the mono audio signal where the difference signal may specifically be a subtraction signal 1-r (or r-1). The coefficient processor 107 may then generate the coefficients of the upmix matrix in dependence on the first intermediate parameter.

[0169] Further, the coefficient processor 107 may be arranged to generate a second intermediate parameter β which is indicative of a residual signal that results after the prediction based on the first intermediate parameter α. The second intermediate parameter β is determined in dependence on the upmix parameters and the signal cancellation measure. The coefficient processor 107 may then proceed to generate the upmix matrix coefficients in dependence on these intermediate parameters.

[0170] As another example, the default encoder matrix may be extended by adding two additional gain matrices, one prior to prediction and one after prediction:

$$\begin{bmatrix} m \\ d' \end{bmatrix} = c \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\beta} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \varepsilon_2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\alpha & 1 \end{bmatrix} \begin{bmatrix} 1 & \varepsilon_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} l \\ r \end{bmatrix}$$

[0171] This can be written as a single matrix operation as:

$$\begin{bmatrix} m \\ d' \end{bmatrix} = c \begin{bmatrix} 1 + \varepsilon_1 & 1 - \varepsilon_1 \\ \frac{\varepsilon_2(1 + \varepsilon_1) + 1 - \alpha(1 + \varepsilon_1)}{\beta} & \frac{\varepsilon_2(1 - \varepsilon_1) - 1 - \alpha(1 - \varepsilon_1)}{\beta} \end{bmatrix} \begin{bmatrix} l \\ r \end{bmatrix}$$

[0172] The decoder upmix matrix may be determined as the inverse of the encoder downmix matrix:

$$\begin{bmatrix} l' \\ r' \end{bmatrix} = \frac{1}{2c} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -\varepsilon_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\varepsilon_2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} m \\ d' \end{bmatrix}$$

[0173] This can be written as a single matrix operation as:

$$\begin{bmatrix} l' \\ r' \end{bmatrix} = \frac{1}{2c} \begin{bmatrix} 1 - \varepsilon_2 + \varepsilon_1\varepsilon_2 + (1 - \varepsilon_1)\alpha & (1 - \varepsilon_1)\beta \\ 1 + \varepsilon_2 + \varepsilon_1\varepsilon_2 - (1 + \varepsilon_1)\alpha & -(1 + \varepsilon_1)\beta \end{bmatrix} \begin{bmatrix} m \\ d' \end{bmatrix}$$

[0174] In this case the α and β parameters may again be determined using the equations above.

[0175] Again, the same signal cancellation measure R may be employed to control the first additional gain matrix (ε_1):

$$R_1 = \frac{2 \cdot ICC \cdot \cos(IPD) \cdot \sqrt{IID}}{IID + 1} = \frac{2 \cdot ICC \cdot \cos(IPD)}{\sqrt{IID} + \frac{1}{\sqrt{IID}}}$$

[0176] The weights ε_1 and ε_2 can for example be defined as a function of the signal cancellation measure, and thus the upmix parameters, as follows:

$$\epsilon_1 = \begin{cases} |R_1|^m & \text{for} \quad R_1 < 0 \\ 0 & \text{for} \quad R_1 \geq 0 \end{cases}$$

[0177] It should be noted that the combined encoder downmix matrices in the examples are given as:

$$\begin{bmatrix} l' \\ r' \end{bmatrix} = \frac{1}{2c} \cdot \frac{1}{g_{11} \cdot g_{22} - g_{12} \cdot g_{21}} \begin{bmatrix} g_{22} - g_{21} + \alpha(g_{11} - g_{12}) & \beta(g_{11} - g_{12}) \\ g_{22} + g_{21} - \alpha(g_{11} + g_{12}) & -\beta(g_{11} + g_{12}) \end{bmatrix} \begin{bmatrix} m \\ d' \end{bmatrix}$$

and

$$\begin{bmatrix} l' \\ r' \end{bmatrix} = \frac{1}{2c} \begin{bmatrix} 1 - \varepsilon_2 + \varepsilon_1 \varepsilon_2 + (1 - \varepsilon_1)\alpha & (1 - \varepsilon_1)\beta \\ 1 + \varepsilon_2 + \varepsilon_1 \varepsilon_2 - (1 + \varepsilon_1)\alpha & -(1 + \varepsilon_1)\beta \end{bmatrix} \begin{bmatrix} m \\ d' \end{bmatrix}$$

[0178] If the normalizing factor of the determinant (g11 g22 - g12 g21) is disregarded, it can be seen that:

$$g_{11} = 1$$

$$g_{12} = \epsilon_1$$

$$g_{21} = \epsilon_2$$

$$g_{22} = 1 + \epsilon_1 \epsilon_2$$

[0179] The normalizing factor of the determinant can be included in/ compensated for by the gain factor c and accordingly it can be seen that the approaches of the examples are equivalent.

[0180] The audio apparatus(es) may specifically be implemented in one or more suitably programmed processors. In particular, the artificial neural networks may be implemented in one more such suitably programmed processors. The different functional blocks, and in particular the artificial neural networks, may be implemented in separate processors and/or may e.g. be implemented in the same processor. An example of a suitable processor is provided in the following.

[0181] FIG. 9 is a block diagram illustrating an example processor 900 according to embodiments of the disclosure. Processor 900 may be used to implement one or more processors implementing an apparatus as previously described or elements thereof (including in particular one more artificial neural network). Processor 900 may be any suitable processor type including, but not limited to, a microprocessor, a microcontroller, a Digital Signal Processor (DSP), a Field ProGrammable Array (FPGA) where the FPGA has been programmed to form a processor, a Graphical Processing Unit (GPU), an Application Specific Integrated Circuit (ASIC) where the ASIC has been designed to form a processor, or a combination thereof.

[0182] The processor 900 may include one or more cores 902. The core 902 may include one or more Arithmetic Logic Units (ALU) 904. In some embodiments, the core 902 may include a Floating Point Logic Unit (FPLU) 906 and/or a Digital Signal Processing Unit (DSPU) 908 in addition to or instead of the ALU 904.

[0183] The processor 900 may include one or more registers 312 communicatively coupled to the core 902. The registers 912 may be implemented using dedicated logic gate circuits (e.g., flip-flops) and/or any memory technology. In some embodiments the registers 912 may be implemented using static memory. The register may provide data, instructions and addresses to the core 902.

[0184] In some embodiments, processor 900 may include one or more levels of cache memory 910 communicatively coupled to the core 902. The cache memory 910 may provide computer-readable instructions to the core 902 for execution. The cache memory 910 may provide data for processing by the core 902. In some embodiments, the computer-readable instructions may have been provided to the cache memory 910 by a local memory, for example, local memory attached to the external bus 916. The cache memory 910 may be implemented with any suitable cache memory type, for example, Metal-Oxide Semiconductor (MOS) memory such as Static Random Access Memory (SRAM), Dynamic Random Access Memory (DRAM), and/or any other suitable memory technology.

[0185] The processor 900 may include a controller 914, which may control input to the processor 900 from other processors and/or components included in a system and/or outputs from the processor 900 to other processors and/or components included in the system. Controller 914 may control the data paths in the ALU 904, FPLU 906 and/or DSPU 908. Controller 914 may be implemented as one or more state machines, data paths and/or dedicated control logic. The gates of controller 914 may be implemented as standalone gates, FPGA, ASIC or any other suitable technology.

[0186] The registers 912 and the cache 910 may communicate with controller 914 and core 902 via internal connections 920A, 920B, 920C and 920D. Internal connections may be implemented as a bus, multiplexer, crossbar switch, and/or any

other suitable connection technology.

[0187] Inputs and outputs for the processor 900 may be provided via a bus 916, which may include one or more conductive lines. The bus 916 may be communicatively coupled to one or more components of processor 900, for example the controller 914, cache 910, and/or register 912. The bus 916 may be coupled to one or more components of the system.

[0188] The bus 916 may be coupled to one or more external memories. The external memories may include Read Only Memory (ROM) 932. ROM 932 may be a masked ROM, Electronically Programmable Read Only Memory (EPROM) or any other suitable technology. The external memory may include Random Access Memory (RAM) 933. RAM 933 may be a static RAM, battery backed up static RAM, Dynamic RAM (DRAM) or any other suitable technology. The external memory may include Electrically Erasable Programmable Read Only Memory (EEPROM) 935. The external memory may include Flash memory 934. The External memory may include a magnetic storage device such as disc 936. In some embodiments, the external memories may be included in a system.

[0189] The invention can be implemented in any suitable form including hardware, software, firmware or any combination of these. The invention may optionally be implemented at least partly as computer software running on one or more data processors and/or digital signal processors. The elements and components of an embodiment of the invention may be physically, functionally and logically implemented in any suitable way. Indeed the functionality may be implemented in a single unit, in a plurality of units or as part of other functional units. As such, the invention may be implemented in a single unit or may be physically and functionally distributed between different units, circuits and processors.

[0190] Although the present invention has been described in connection with some embodiments, it is not intended to be limited to the specific form set forth herein. Rather, the scope of the present invention is limited only by the accompanying claims. Additionally, although a feature may appear to be described in connection with particular embodiments, one skilled in the art would recognize that various features of the described embodiments may be combined in accordance with the invention. In the claims, the term comprising does not exclude the presence of other elements or steps.

[0191] Furthermore, although individually listed, a plurality of means, elements, circuits or method steps may be implemented by e.g. a single circuit, unit or processor. Additionally, although individual features may be included in different claims, these may possibly be advantageously combined, and the inclusion in different claims does not imply that a combination of features is not feasible and/or advantageous. Also, the inclusion of a feature in one category of claims does not imply a limitation to this category but rather indicates that the feature is equally applicable to other claim categories as appropriate. Furthermore, the order of features in the claims do not imply any specific order in which the features must be worked and in particular the order of individual steps in a method claim does not imply that the steps must be performed in this order. Rather, the steps may be performed in any suitable order. In addition, singular references do not exclude a plurality. Thus, references to "a", "an", "first", "second" etc do not preclude a plurality. Reference signs in the claims are provided merely as a clarifying example shall not be construed as limiting the scope of the claims in any way.

Claims

1. An apparatus for generating an output audio stereo signal, the apparatus comprising:

   a receiver (101) arranged to receive an audio data signal comprising:

   a mono audio signal being a downmix of two channel signals of a first audio stereo signal;
   upmix parameters for the mono audio signal, the set of upmix parameters comprising a first parameter indicative of a level difference between the two channel signals, a second parameter indicative of a correlation between the two channel signals, and a third parameter indicative of a phase difference between the two channel signals;
   a coefficient generator (107) arranged to generate coefficients for an upmix matrix from the upmix parameters;
   a generator (103) arranged to generate the output audio stereo signal by applying the upmix matrix to samples of the mono audio signal and an auxiliary mono audio signal;
   wherein
   the coefficient generator (107) is arranged to:

   determine a signal cancellation measure from the upmix parameters, the signal cancellation measure being indicative of a signal cancellation in a summation of the two channel signals; and
   determine the coefficients for the upmix matrix in dependence on the signal cancellation measure.

2. The apparatus of claim 1 wherein the coefficient processor (107) is arranged to adapt the upmix coefficients to deviate from coefficients for the mono audio signal being a sum signal of the channel signals for the signal cancellation

measure meeting a first signal cancellation requirement.

3. The apparatus of claim 1 or 2 wherein the coefficient processor (107) is arranged to increase a deviation of upmix coefficients from coefficients for the mono audio signal being a sum signal of the channel signals for the signal cancellation measure being indicative of an increasing signal cancellation in the sum signal of the channel signals.

4. The apparatus of any previous claim wherein the coefficient processor (107) is arranged to increase a deviation of upmix coefficients from coefficients for the mono audio signal being a sum signal of the channel signals for the signal cancellation measure being indicative of an increasing signal cancellation in a difference signal of the channel signals.

5. The apparatus of any of claims 2 to 4 wherein the coefficients for the mono audio signal being a sum signal of the channel signals are given as

$$\frac{1}{2c}\begin{bmatrix} 1+\alpha & \beta \\ 1-\alpha & -\beta \end{bmatrix}$$

where

$$\alpha = \frac{IID - 1 + 2j \cdot \sin(IPD) \cdot ICC \cdot \sqrt{IID}}{IID + 1 + 2 \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}}$$

$$\beta = \frac{2 \cdot \sqrt{IID \cdot (1 - ICC^2)}}{IID + 1 + 2 \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}}$$

$$c = \min\left(\sqrt{\frac{IID + 1}{IID + 1 + 2 \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}}}, c_{max} \right)$$

in which IID is an Interaural Intensity Differences upmix parameter, ICC is an Interchannel Cross Correlations upmix parameter, and IPD is an Inter-channel Phase Difference upmix parameter.

6. The apparatus of any previous claim wherein the signal cancellation measure is determined substantially as:

$$R = \frac{2 \cdot ICC \cdot \cos(IPD)}{\sqrt{IID} + \frac{1}{\sqrt{IID}}}$$

where IID is an Interaural Intensity Difference upmix parameter, ICC is an Inter-channel Cross Correlations upmix parameter, and IPD is an Inter-channel Phase Difference upmix parameter.

7. The apparatus of any previous claims wherein the coefficient processor (107) is arranged to generate a first intermediate parameter in dependence on the upmix parameters and the signal cancellation measure, the first intermediate parameter indicative of a prediction of a difference signal for the channel signals from the mono audio signal, and to generate the coefficients in dependence on the first intermediate parameter.

8. The apparatus of claim 7 wherein the coefficient processor (107) is arranged to generate a second intermediate parameter indicative of a residual signal for the prediction in dependence on the upmix parameters and the signal cancellation measure, and to generate the coefficients in dependence on the intermediate parameter.

9. The apparatus of any previous claims wherein the coefficient processor (107) is arranged to generate the upmix matrix as:

$$\frac{1}{2c} \cdot \frac{1}{g_{11} \cdot g_{22} - g_{12} \cdot g_{21}} \begin{bmatrix} g_{22} - g_{21} + \alpha(g_{11} - g_{12}) & \beta(g_{11} - g_{12}) \\ g_{22} + g_{21} - \alpha(g_{11} + g_{12}) & -\beta(g_{11} + g_{12}) \end{bmatrix}$$

where c is a gain parameter and α and β are parameters dependent on the upmix parameters and the signal cancellation measure, and the parameters $g_{1,1}, g_{1,2}, g_{2,1},$ and $g_{2,2}$, are dependent on the signal cancellation measure.

10. The apparatus of any previous claims wherein the coefficient processor (107) is arranged to generate the upmix matrix as:

$$\frac{1}{2c} \cdot \frac{1}{g_{11} \cdot g_{22} - g_{12} \cdot g_{21}} \begin{bmatrix} g_{22} - g_{21} + \alpha(g_{11} - g_{12}) & \beta(g_{11} - g_{12}) \\ g_{22} + g_{21} - \alpha(g_{11} + g_{12}) & -\beta(g_{11} + g_{12}) \end{bmatrix}$$

where c is a gain parameter and

$$\alpha = \frac{g_A \cdot g_B \cdot IID + g_C \cdot g_D + (g_C \cdot g_B \cdot e^{j \cdot IPD} + g_A \cdot g_D \cdot e^{-j \cdot IPD}) \cdot ICC \cdot \sqrt{IID}}{g_A{}^2 \cdot IID + g_C{}^2 + 2 \cdot g_A \cdot g_C \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}}$$

$$\beta = \sqrt{\frac{g_B{}^2 \cdot IID + g_D{}^2 + 2 \cdot g_B \cdot g_D \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}}{g_A{}^2 \cdot IID + g_C{}^2 + 2 \cdot g_A \cdot g_C \cdot \cos(IPD) \cdot ICC \cdot \sqrt{IID}} - |\alpha|^2}$$

with:

$$g_A = g_{11} + g_{12}$$

$$g_B = g_{21} + g_{22}$$

$$g_C = g_{11} - g_{12}$$

$$g_D = g_{21} - g_{22}$$

In which IID is an Interaural Intensity Differences upmix parameter, ICC is an Interchannel Cross Correlations upmix parameter, and IPD is an Inter-channel Phase Difference upmix parameter; and wherein $\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$ is a dependent on the signal cancellation measure.

11. The apparatus of claim 9 wherein:

$$\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} 1 & \epsilon_1 \\ \epsilon_2 & 1 \end{bmatrix}$$

where ε_1 and ε_2 are dependent on the signal cancellation measure.

12. An apparatus for generating an audio data signal, the apparatus comprising:

a receiver (201) arranged to receive an audio stereo signal comprising two channel signals;
a downmixer (203) arranged to generate a mono audio signal as a combination of the two channel signals in dependence on a set of downmix coefficients;

a parameter generator (205) arranged to generate a set of upmix parameters comprising a first parameter indicative of a level difference between the two channel signals, a second parameter indicative of a correlation between the two channel signals, and a third parameter indicative of a phase difference between the two channel signals;

a downmix coefficient processor (207) arranged to generate the set of downmix coefficients in dependence on the set of upmix parameters;

a data signal generator (209) arranged to generate the audio signal to include the mono audio signal and the set of upmix parameters;

a signal cancellation estimator (211) which is arranged to determine a signal cancellation measure from the set of upmix parameters, the signal cancellation measure being indicative of a signal cancellation in a summation of the two channel signals; and

wherein the downmix coefficient processor (207) is arranged to generate the set of downmix coefficients in dependence on the signal cancellation measure.

13. A method of generating an output audio stereo signal, the method comprising:
receiving an audio data signal comprising:

a mono audio signal being a downmix of two channel signals of a first audio stereo signal;
upmix parameters for the mono audio signal, the set of upmix parameters comprising a first parameter indicative of a level difference between the two channel signals, a second parameter indicative of a correlation between the two channel signals, and a third parameter indicative of a phase difference between the two channel signals;
generating coefficients for an upmix matrix from the upmix parameters;
generating the output audio stereo signal by applying the upmix matrix to samples of the mono audio signal and an auxiliary mono audio signal;
wherein
generating the coefficients comprises:

determining a signal cancellation measure from the upmix parameters, the signal cancellation measure being indicative of a signal cancellation in a summation of the two channel signals;
and determining the coefficients for the upmix matrix in dependence on the signal cancellation measure.

14. A method of generating an audio data signal, the method comprising:

receiving an audio stereo signal comprising two channel signals;
generating a mono audio signal as a combination of the two channel signals in dependence on a set of downmix coefficients;
generating a set of upmix parameters comprising a first parameter indicative of a level difference between the two channel signals, a second parameter indicative of a correlation between the two channel signals, and a third parameter indicative of a phase difference between the two channel signals;
generating the set of downmix coefficients in dependence on the set of upmix parameters;
generating the audio signal to include the mono audio signal and the set of upmix parameters;
determining a signal cancellation measure from the set of upmix parameters, the signal cancellation measure being indicative of a signal cancellation in a summation of the two channel signals; and
wherein generating the set of downmix coefficients is in dependence on the signal cancellation measure.

15. A computer program product comprising computer program code means adapted to perform all the steps of claim 14 when said program is run on a computer.
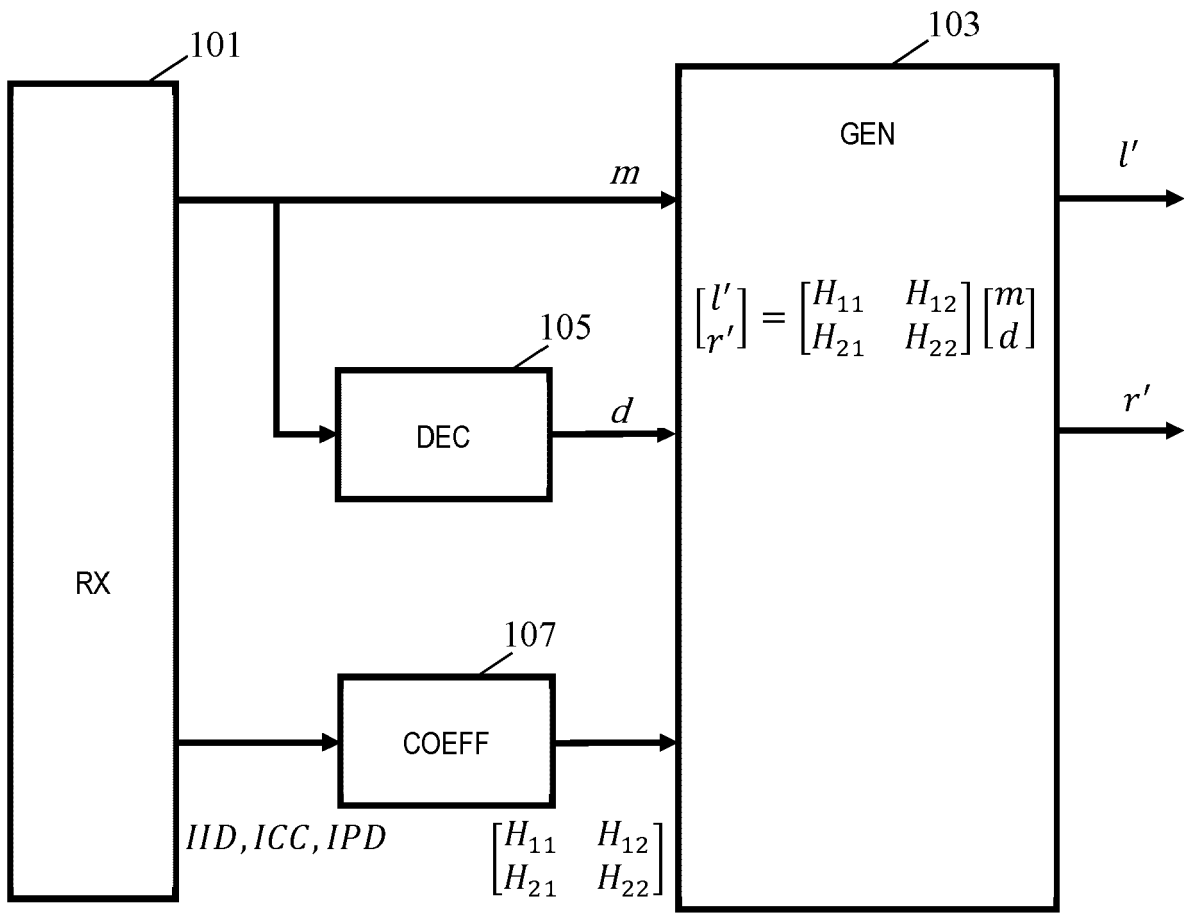
FIG. 1

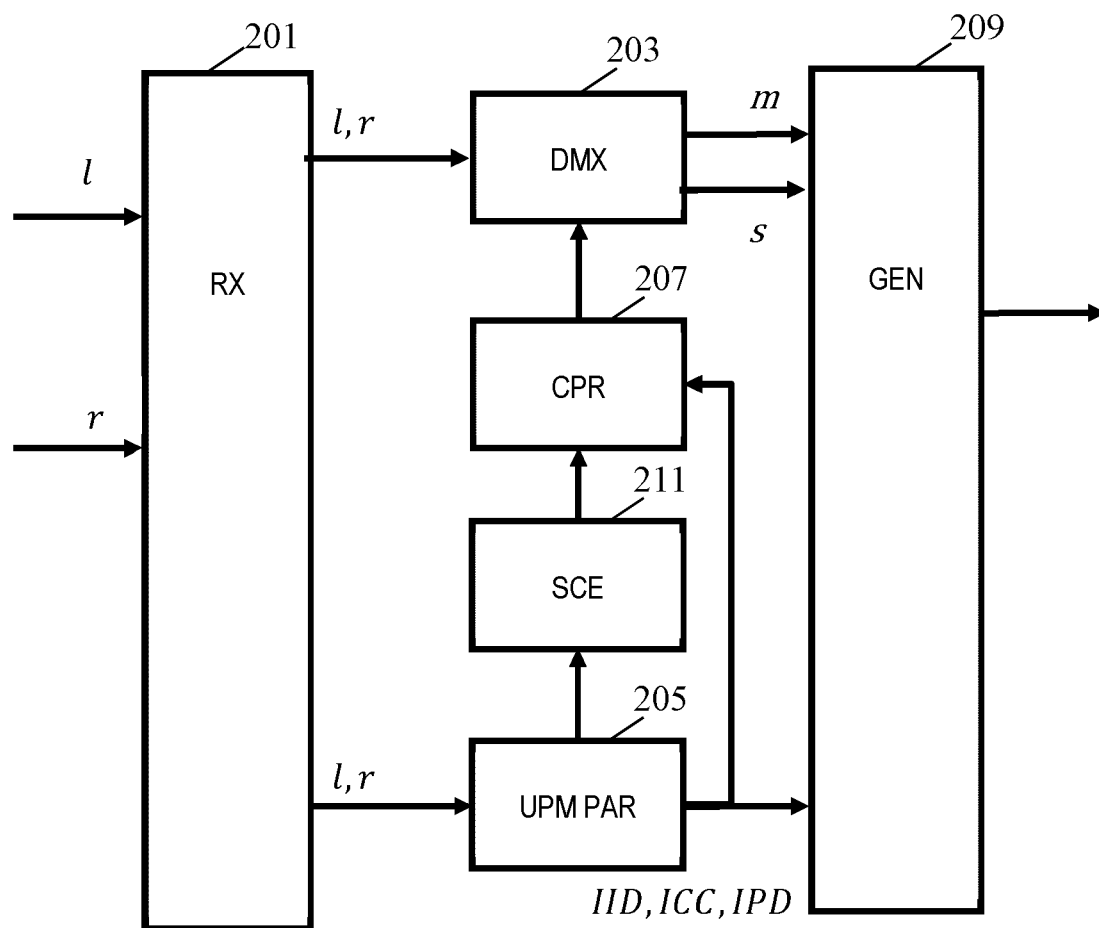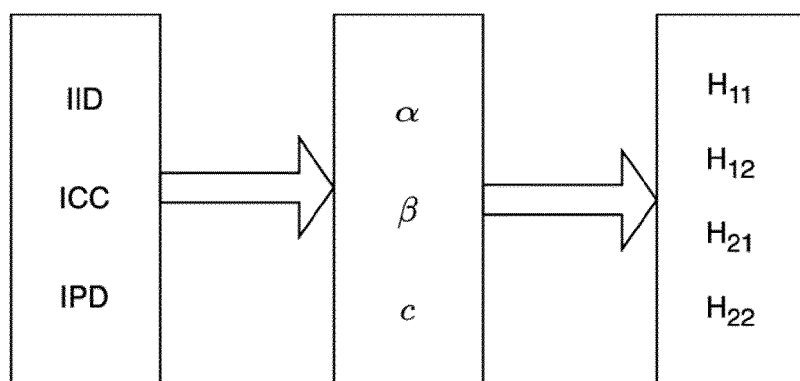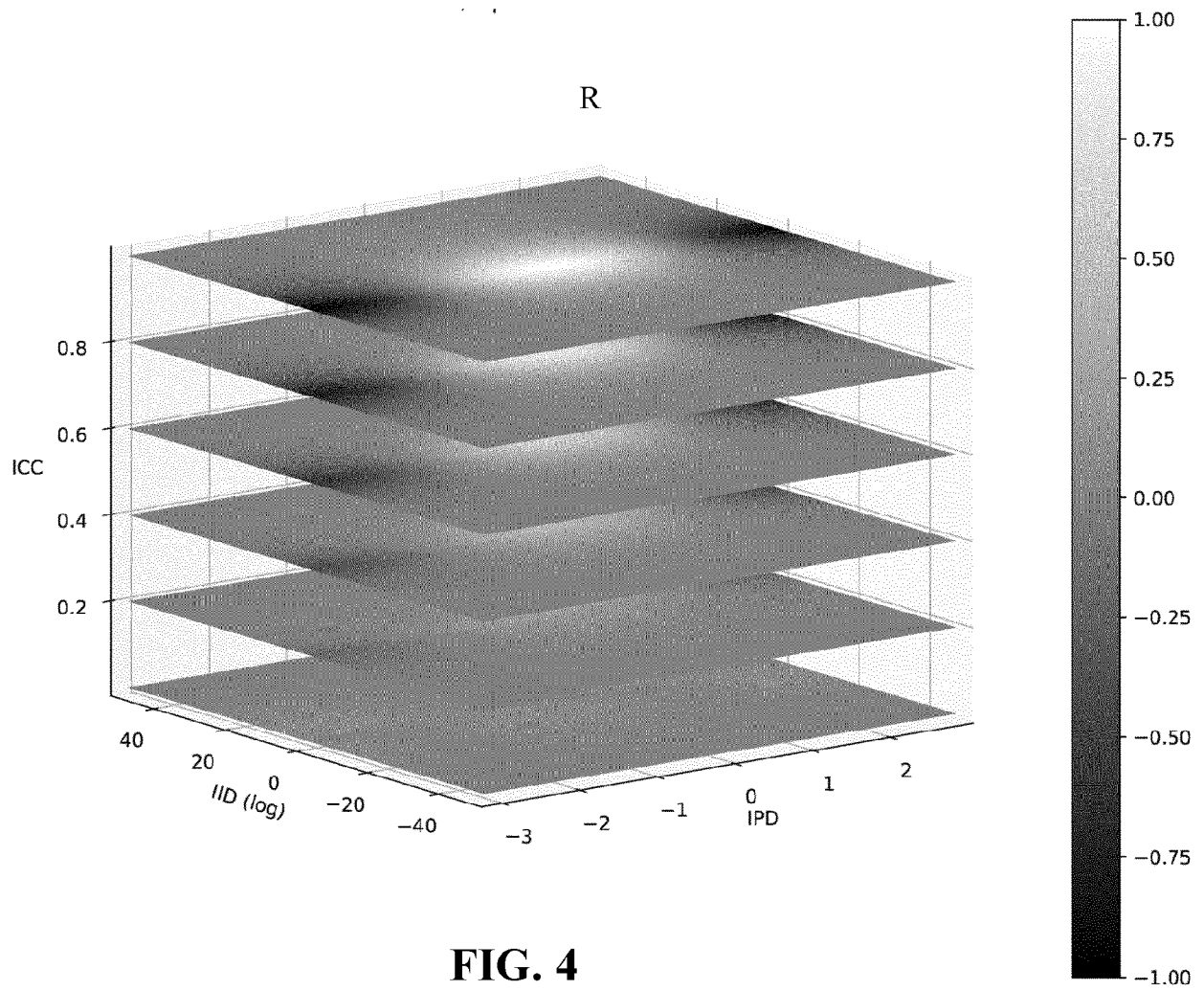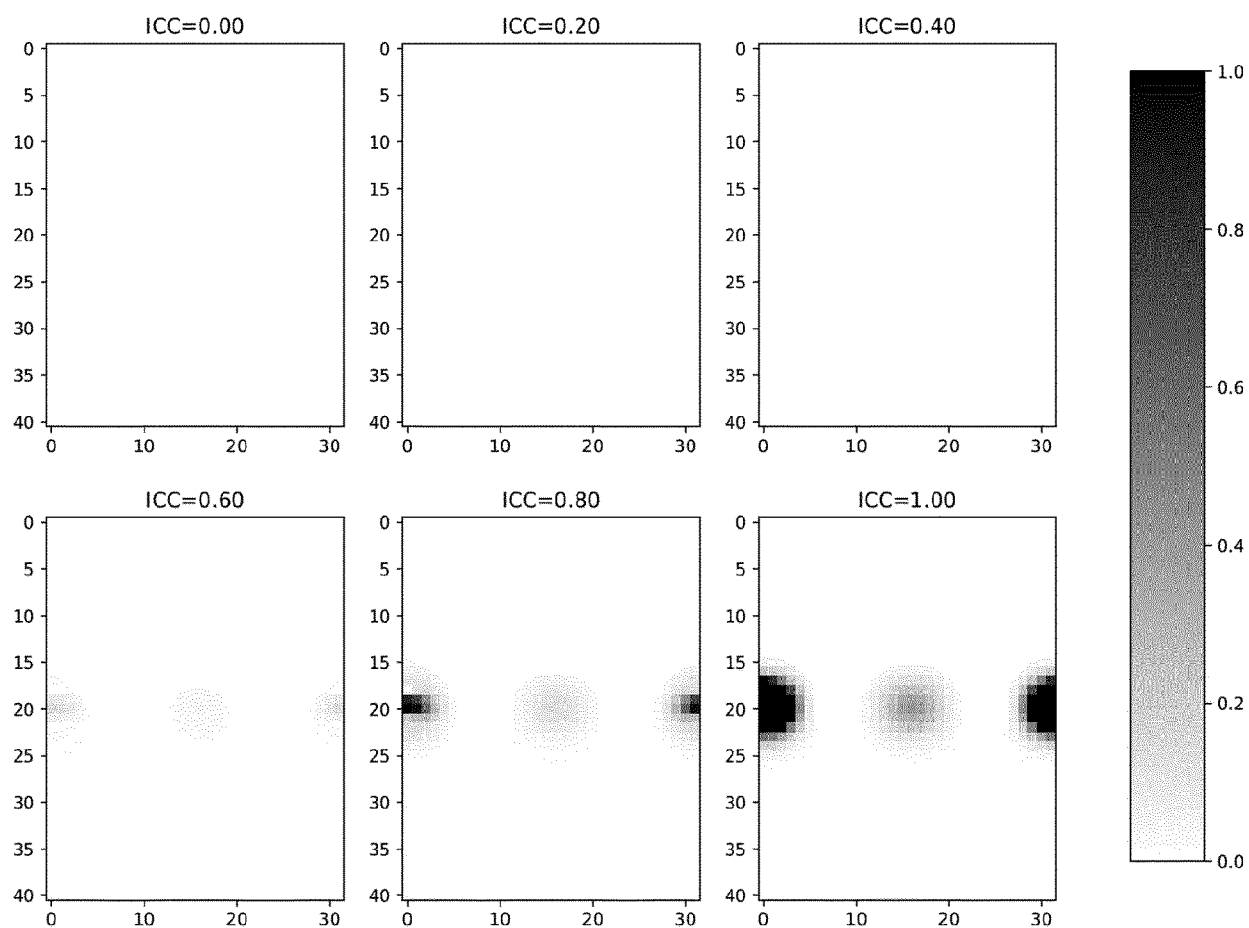**FIG. 2**

IID

ICC

IPD

α

β

c

H_{11}

H_{12}

H_{21}

H_{22}

FIG. 3

FIG. 4

FIG. 5

FIG. 6

FIG. 7

FIG. 8

Processor 900

Core 902

ALU 904

Cache
910

920A

FPLU 906

920B

Registers
912

DSPU 908

920C

Controller
914

920D

ROM
932

Flash
934

External Bus 916

RAM
933

EEPROM
935

Disc
936

FIG. 9

Europäisches Patentamt
European Patent Office
Office européen des brevets
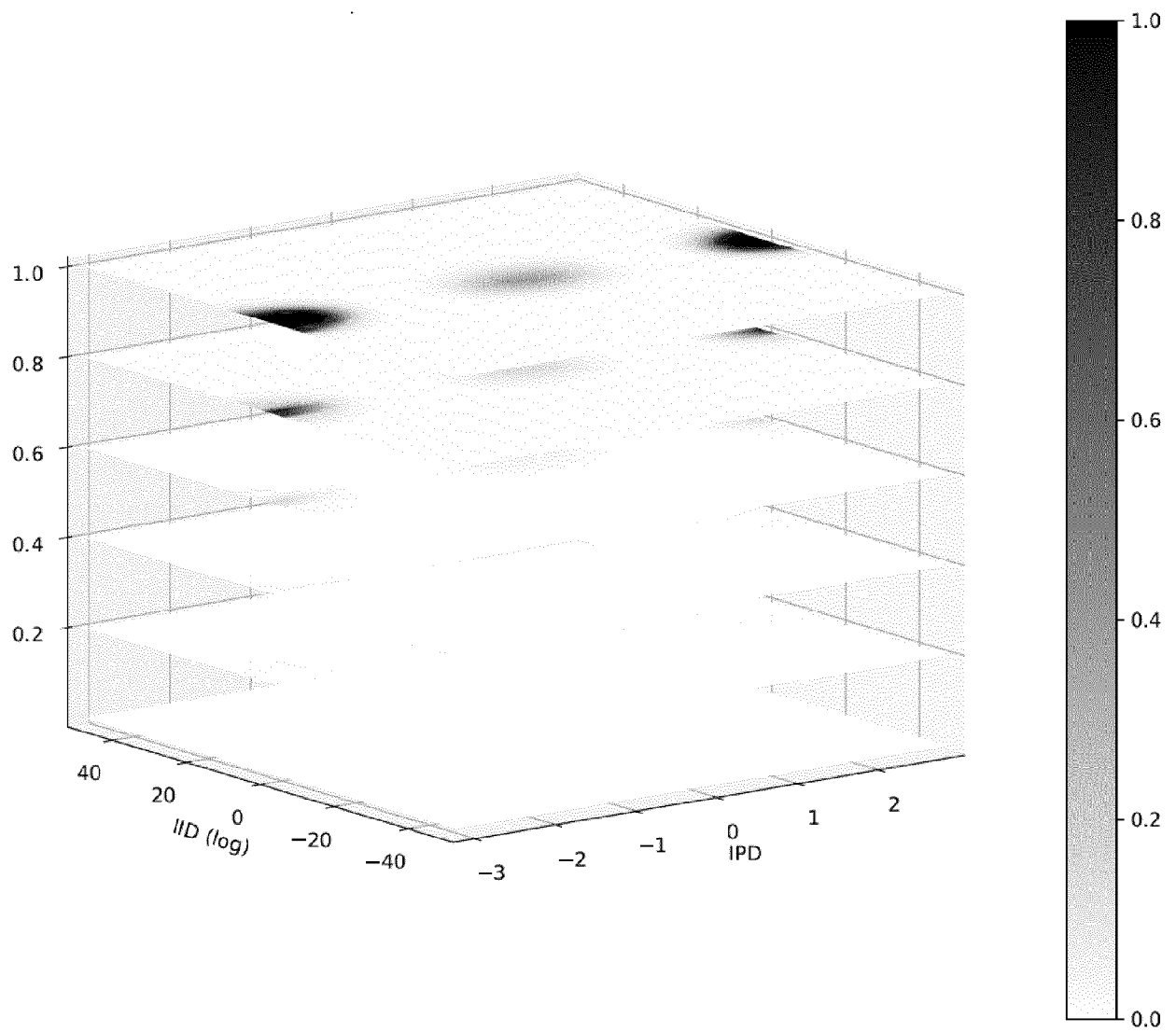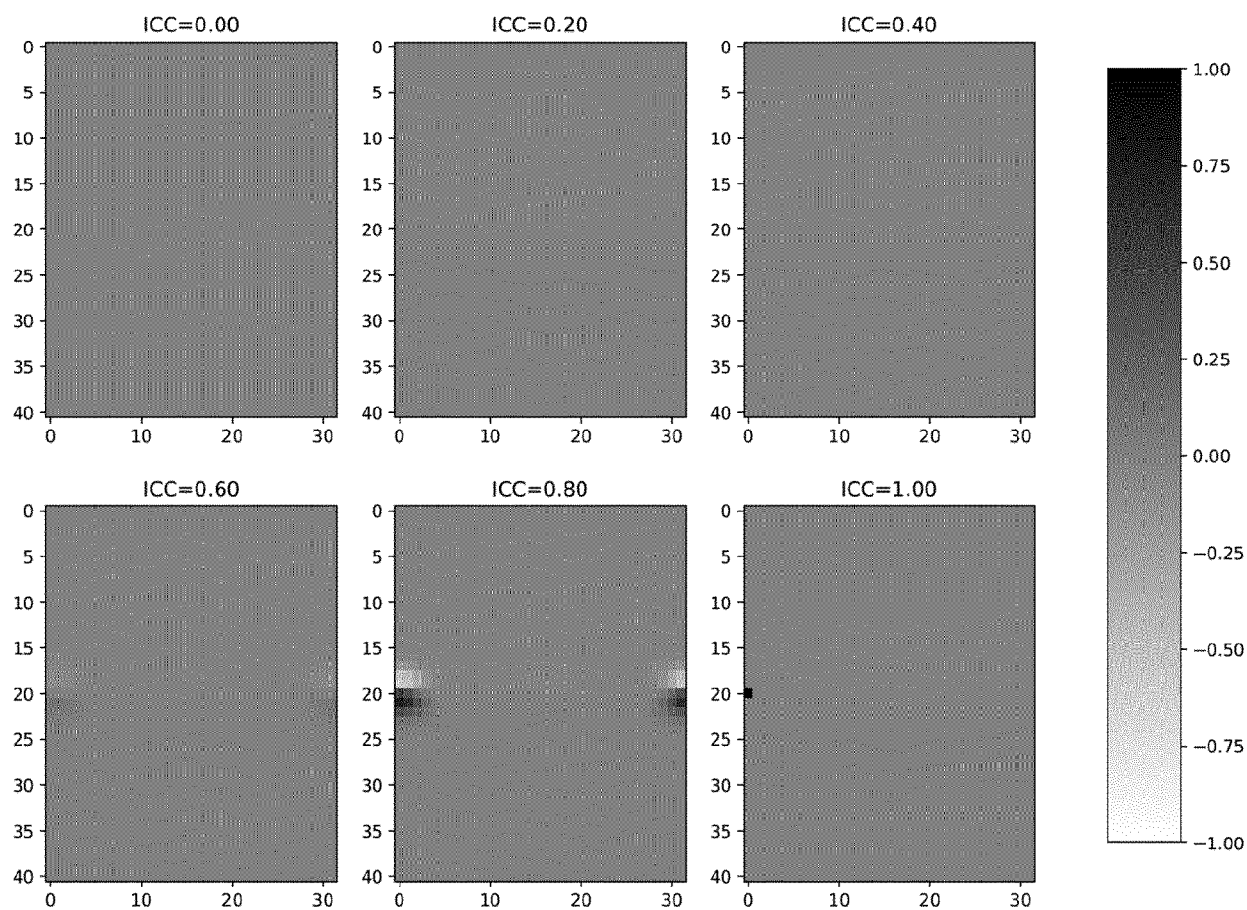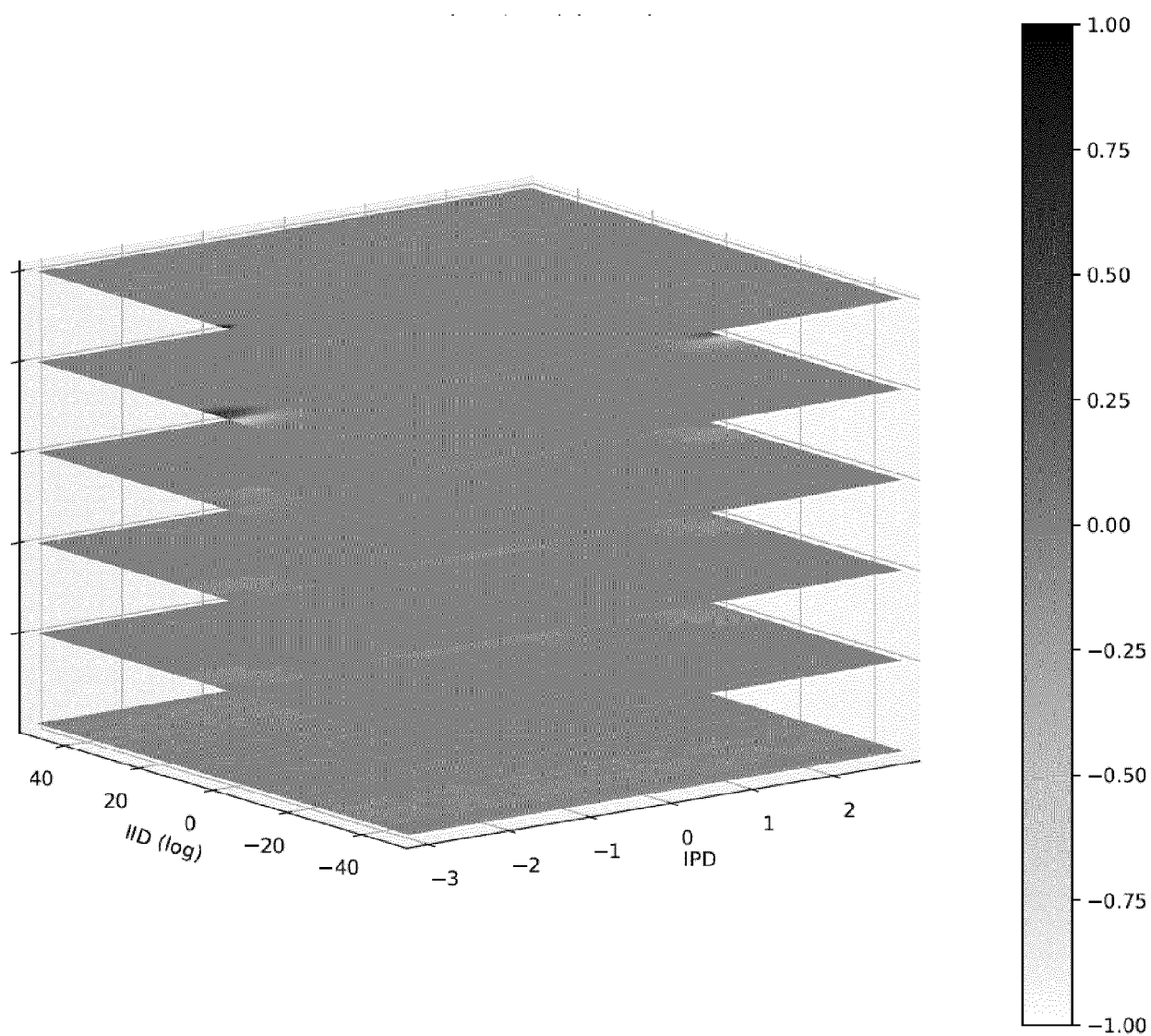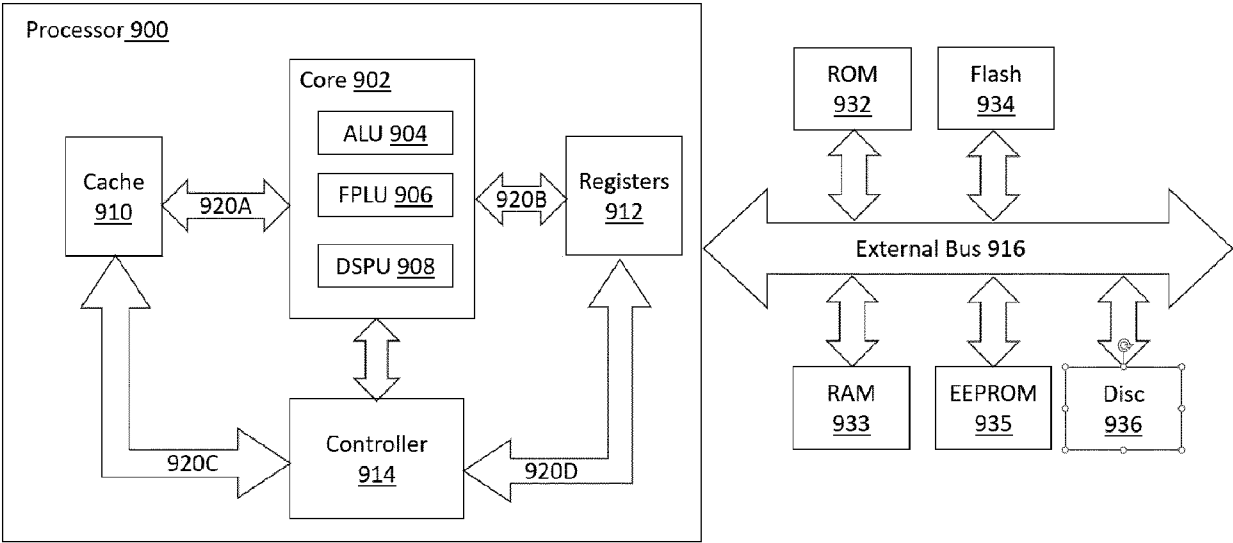
EUROPEAN SEARCH REPORT

Application Number

EP 23 18 7751

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	WO 2010/097748 A1 (KONINKL PHILIPS ELECTRONICS NV [NL]; OOMEN ARNOLDUS W J [NL] ET AL.) 2 September 2010 (2010-09-02)	1-5,7-15	INV. G10L19/008 H04S1/00
A	* figures 5,6 * * page 5, paragraph 2 * * page 14 - page 18 * * pages 26,27,28 *	6	H04S7/00
A	EP 1 869 668 A1 (CODING TECH AB [SE]; KONINKL PHILIPS ELECTRONICS NV [NL]) 26 December 2007 (2007-12-26) * figures 1,6 * * paragraphs [0037] - [0060], [0080] - [0082] *	1-15	
A	EP 3 571 695 A1 (QUALCOMM INC [US]) 27 November 2019 (2019-11-27) * figures 2,3 *	1-15	
A	EP 2 609 590 A1 (FRAUNHOFER GES FORSCHUNG [DE]) 3 July 2013 (2013-07-03) * figure 2 *	1-15	

TECHNICAL FIELDS SEARCHED (IPC)

G10L
H04S

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
Munich	24 November 2023	Chétry, Nicolas

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.

EP 23 18 7751

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

24-11-2023

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 2010097748	A1	02-09-2010	TW	201103008 A	16-01-2011
			WO	2010097748 A1	02-09-2010
EP 1869668	A1	26-12-2007	AT	E454693 T1	15-01-2010
			BR	PI0612218 A2	26-10-2010
			CN	101160619 A	09-04-2008
			EP	1869668 A1	26-12-2007
			ES	2338918 T3	13-05-2010
			HK	1110985 A1	25-07-2008
			JP	4685925 B2	18-05-2011
			JP	2008536184 A	04-09-2008
			KR	20070120527 A	24-12-2007
			MY	147609 A	31-12-2012
			PL	1869668 T3	30-06-2010
			RU	2380766 C2	27-01-2010
			TW	I303411 B	21-11-2008
			US	2006233379 A1	19-10-2006
			WO	2006108573 A1	19-10-2006
EP 3571695	A1	27-11-2019	AU	2017394681 A1	20-06-2019
			BR	112019014544 A2	27-02-2020
			CN	110100280 A	06-08-2019
			CN	116033328 A	28-04-2023
			EP	3571695 A1	27-11-2019
			KR	20190107025 A	18-09-2019
			KR	20230138046 A	05-10-2023
			SG	11201904753W A	27-08-2019
			TW	201832572 A	01-09-2018
			US	2018204579 A1	19-07-2018
			US	2019295559 A1	26-09-2019
			WO	2018136167 A1	26-07-2018
EP 2609590	A1	03-07-2013	AR	082542 A1	12-12-2012
			AR	082543 A1	12-12-2012
			AR	098078 A2	27-04-2016
			AU	2011295367 A1	11-04-2013
			AU	2011295368 A1	28-03-2013
			BR	112013004362 A2	19-09-2017
			BR	112013004365 A2	02-06-2020
			CA	2809404 A1	01-03-2012
			CA	2809437 A1	01-03-2012
			CA	2887939 A1	01-03-2012
			CN	103180898 A	26-06-2013
			CN	103460282 A	18-12-2013
			EP	2609590 A1	03-07-2013
			EP	2609591 A1	03-07-2013

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

EPO FORM P0459

page 1 of 2

ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.

EP 23 18 7751

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

24-11-2023

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
		EP	2924687 A1	30-09-2015
		EP	3144932 A1	22-03-2017
		EP	3471091 A1	17-04-2019
		ES	2544077 T3	27-08-2015
		ES	2585402 T3	05-10-2016
		ES	2706490 T3	29-03-2019
		HK	1186833 A1	21-03-2014
		HK	1187144 A1	28-03-2014
		JP	5775582 B2	09-09-2015
		JP	5775583 B2	09-09-2015
		JP	6196249 B2	13-09-2017
		JP	2013539553 A	24-10-2013
		JP	2013539554 A	24-10-2013
		JP	2015129953 A	16-07-2015
		KR	20130069770 A	26-06-2013
		KR	20130079507 A	10-07-2013
		MY	156770 A	31-03-2016
		MY	178197 A	06-10-2020
		MY	180970 A	14-12-2020
		PL	2609590 T3	30-10-2015
		PL	2609591 T3	30-11-2016
		PL	3144932 T3	30-04-2019
		PT	2609591 T	12-07-2016
		PT	3144932 T	04-02-2019
		RU	2013112853 A	27-09-2014
		RU	2013112903 A	27-09-2014
		RU	2015102326 A	10-06-2015
		SG	187950 A1	28-03-2013
		SG	188254 A1	30-04-2013
		SG	2014006738 A	28-03-2014
		TR	201900417 T4	21-02-2019
		TW	201214414 A	01-04-2012
		TW	201214417 A	01-04-2012
		US	2013173273 A1	04-07-2013
		US	2013173274 A1	04-07-2013
		US	2014222441 A1	07-08-2014
		WO	2012025282 A1	01-03-2012
		WO	2012025283 A1	01-03-2012
		ZA	201302050 B	23-12-2013

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

page 2 of 2

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Non-patent literature cited in the description

- **E. SCHUIJERS** ; **W. OOMEN** ; **B. DEN BRINKER** ; **J. BREEBAART**. Advances in Parametric Coding for High-Quality Audio. *114th AES Convention, Amsterdam, The Netherlands*, 2003 **[0006]**
- **E. SCHUIJERS** ; **J. BREEBAART** ; **H. PURNHAGEN** ; **J. ENGDEGÅRD**. Low Complexity Parametric Stereo Coding. *116th AES, Berlin, Germany*, 2004 **[0006]**
- **A. C. DEN BRINKER** ; **J. BREEBAART** ; **P. EKSTRAND** ; **J. ENGDEGÅRD** ; **F. HENN** ; **K. KJÖRLING** ; **W. OOMEN** ; **H. PURNHAGEN**. An Overview of the Coding Standard MPEG-4 Audio Amendments 1 and 2: HE-AAC, SSC, and HE-AAC v2. *EURASIP Journal on Audio Speech and Music Processing*, January 2009, vol. 14496 (3), 2005 **[0007]**