



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
07.10.2026 Bulletin 2026/41

(51) International Patent Classification (IPC):
G10L 19/008 ^(2013.01)

(21) Application number: **25168242.3**

(52) Cooperative Patent Classification (CPC):
G10L 19/008

(22) Date of filing: **03.04.2025**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC ME MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA
Designated Validation States:
GE KH LA MA MD TN

(72) Inventors:
• **SCHUIJERS, Erik Gosuinus Petrus**
Eindhoven (NL)
• **KECHICHIAN, Patrick**
Eindhoven (NL)

(74) Representative: **Philips Intellectual Property & Standards**
High Tech Campus 34
5656 AE Eindhoven (NL)

(71) Applicant: **Koninklijke Philips N.V.**
5656 AE Eindhoven (NL)

(54) **GENERATION OF A STEREO CHANNEL AUDIO SIGNAL AND AUDIO DATA SIGNAL REPRESENTING A STEREO CHANNEL AUDIO SIGNAL**

(57) A decoder audio apparatus comprises a receiver (101) receiving an audio data signal comprising a downmix audio signal for a stereo signal as well as spatial parameter sets wherein a second set comprises parameters with a lower frequency resolution and a higher time resolution than parameters for a first set. A decorrelator (103) decorrelates the mono downmix to generate a decorrelated signal and a transformation circuit (105) generates a first stereo signal from the mono downmix signal and the decorrelated signal using the parameters

of the first set of spatial parameters. A signal generator (107) generates an output audio stereo signal from the first stereo signal with the generation depending on parameters of the second set of spatial parameters. The signal generator (107) may perform a second transformation/upmixing using coefficients determined from parameters of the second set. The approach may provide improved temporal accuracy for fast signal components with little increase in the data rate.

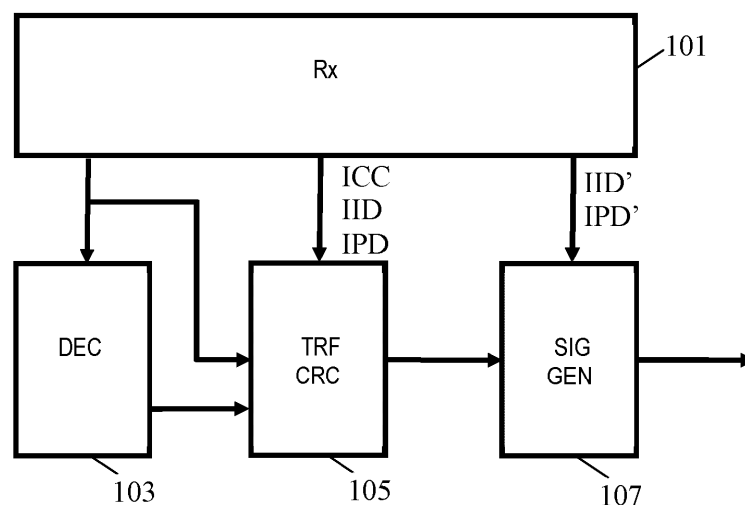


FIG. 1

Description

FIELD OF THE INVENTION

[0001] The invention relates to generation of a stereo channel audio signal and/or an audio data signal representing a stereo channel audio signal, and in particular, but not exclusively, to parametric encoding and/or decoding of stereo signals.

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 the 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 or an original one.

[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/ spatial 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 spatial parameters. Specifically, the two signals may be multiplied by a time- and frequency-dependent 2x2 matrix having coefficients determined from the spatial parameters to provide the output stereo signal.

[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) multichannel audio signal provided to the encoder and the multichannel audio signal recreated at the decoder. Typically, the audio quality may be degraded, and recreation of the stereo channel is typically suboptimal. Further, the data rate may still be higher than desired and/or the complexity/ resource usage of the processing may be higher than preferred.

[0009] A further issue is that it is often particularly desirable to reduce the complexity and computational load, especially at the decoder side.

[0010] Hence, an improved approach would be advantageous. In particular an approach allowing increased flexibility, improved adaptability, an improved performance, 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 audio experience would be advantageous.

SUMMARY OF THE INVENTION

[0011] Accordingly, the Invention seeks to preferably mitigate, alleviate or eliminate one or more of the above mentioned disadvantages singly or in any combination.

[0012] According to an aspect of the invention there is provided an audio apparatus for generating an output audio stereo signal, the audio apparatus comprising: a receiver arranged to receive an audio data signal comprising: a mono downmix signal of an audio stereo signal, a first set of spatial parameters, the first set of spatial parameters being indicative of relative properties of channel signals of the audio stereo signal, a second set of spatial parameters, the second set of spatial parameters being indicative of relative properties of channel signals of the audio stereo signal, the second set of spatial parameters having a lower frequency resolution and a higher time resolution than the first set of spatial parameters; a decorrelator arranged to decorrelate the mono downmix signal to generate a decorrelated signal; a first transformation circuit arranged to generate a first stereo signal from the mono downmix signal and the decorrelated signal using the first set of parameters; and a signal generator arranged to generate the output audio stereo signal from the first stereo signal in dependence on the second set of parameters.

[0013] The approach may provide an improved audio experience in many embodiments. For many signals and scenarios, the approach may provide improved generation/ reconstruction of an audio stereo signal with an improved perceived audio quality. In particular, the approach may provide improved representation and replication of audio signal components that have particular characteristics, such as transients and fast changing signal components. The approach may provide additional perceptual cues that are particularly relevant and important for the human perception, including the spatial perception of audio, and in particular cues may be provided to represent short term events that extend over a frequency range.

[0014] The approach may in many embodiments allow a facilitated and/or reduced complexity and/or reduced resource usage implementation. In particular, the cascaded processing of the first parameter set followed by processing of the second parameter set may provide an efficient implementation and may in many scenarios facilitate an enhancement or improvement of existing implementations to provide improved performance.

[0015] The approach may provide improved audio quality with only a very small relative data overhead. In many cases, the additional data required to represent the second set of spatial parameters may be very small in comparison to the data required to represent the mono downmix signal and/or the first set of spatial parameters. In many embodiments, the second set of spatial parameters may only include parameters for small parts of the time frequency grid for which the first parameter set includes parameters.

[0016] 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. Indeed, in many embodiments, substantially improved audio quality may be achieved with very little increase in the overall data rate.

[0017] The approach may in many embodiments allow reduced complexity and/or resource usage at the decoder/ generating/ reconstruction side.

[0018] The first transformation circuit may specifically be arranged to generate the first stereo signal by applying a matrix multiplication to (samples of) the downmix audio signal and the decorrelated signal with the matrix coefficients being determined from the first set of spatial parameters.

[0019] In some embodiments, the first set of parameters are provided with a quantization of a first time frequency (quantization) grid. In some embodiments, the second set of parameters are provided with a quantization of a second time frequency (quantization) grid. The second time frequency grid may have a higher time resolution and a lower frequency resolution than the first time frequency grid.

[0020] At least some of the parameters of the second set of spatial parameters apply to a larger frequency range than the parameters of the first set of spatial parameters covering a frequency interval within the larger frequency range. At least some of the parameters of the first set of spatial parameters apply to a smaller time interval than a parameter of the first set of spatial parameters that cover the smaller time interval. A lower frequency resolution may correspond to a larger frequency range, and a higher time resolution may correspond to a smaller time interval.

[0021] In many embodiments, the time and/or spatial resolution for the first set of spatial parameters may be a fixed or predetermined resolution. In many embodiments, the time and/or spatial resolution for the second set of spatial parameters may be a variable resolution. In many embodiments, the time and/or spatial resolution for the second set of spatial parameters may be a signal dependent resolution.

[0022] According to an optional feature of the invention, the signal generator comprises a second transformation circuit arranged to upmix the first stereo signal using coefficients determined from the second set of parameters; and wherein the first transformation circuit is arranged to upmix the mono downmix signal and the decorrelated signal on a first time frequency (quantization) grid, and the signal generator is arranged to upmix the first stereo signal on a second time frequency (quantization) grid, the second time frequency (quantization) grid having a higher time resolution and a lower frequency resolution than the first time frequency (quantization) grid.

[0023] This may provide an advantageous approach for many scenarios, including e.g. providing an advantageous trade-off between complexity, computational resources and/or the perceived audio quality of the generated audio stereo signal.

[0024] In many embodiments, the first time frequency (quantization) grid may be a fixed and/or predetermined time frequency (quantization) grid. In many embodiments, the second time frequency (quantization) grid may be a variable and/or signal dependent time frequency (quantization) grid.

[0025] According to an optional feature of the invention, the signal generator is arranged to generate the output stereo signal to include a first contribution from a first channel signal of the first stereo signal and a second contribution from a second channel signal of the first stereo signal, the first contribution being dependent on the second set of spatial parameters and the second contribution not being dependent on the second set of spatial parameters.

[0026] This may provide an advantageous approach in many scenarios and may in particular provide an advantageous trade-off between performance, data rate and computational resource/complexity. The output stereo signal may be generated by combining at least the first contribution and the second contribution (e.g. by a summation).

[0027] According to an optional feature of the invention, the first contribution is an intermediate stereo signal generated by upmixing the first channel signal dependent on the second set of parameters.

[0028] This may provide an advantageous approach in many scenarios and/or embodiments. The upmixing may include only the first channel signal and may not include the second channel signal.

[0029] According to an optional feature of the invention, the signal generator includes a synthesis filter for converting the first channel signal from a time frequency grid of the generation of the first stereo signal by the first transformation circuit to a time frequency grid of the upmixing of the first channel signal, the time frequency grid of the upmixing of the first channel signal having a higher time resolution and a lower frequency resolution than the time frequency grid of the generation of the first stereo signal by the first transformation circuit.

[0030] This may provide an advantageous approach in many scenarios and/or embodiments.

[0031] The time frequency grid of the first upmixing may be a time frequency grid of the first set of spatial parameters. The time frequency grid of the second upmixing may be a time frequency grid of the second set of spatial parameters.

[0032] According to an optional feature of the invention, the second contribution is of equal magnitude but opposite phase for the channel signals of the output stereo signal.

[0033] This may provide an advantageous approach in many scenarios and/or embodiments. It may provide a particularly advantageous implementation in many embodiments.

[0034] According to an optional feature of the invention, the first transformation circuit is arranged to generate the first stereo signal as a mid-side stereo signal.

[0035] This may provide an advantageous approach in many scenarios and/or embodiments. It may provide a particularly advantageous implementation in many embodiments. The first stereo signal may be generated as a sum-difference stereo signal.

[0036] According to an optional feature of the invention, the first transformation circuit is arranged to generate the first stereo signal with a first channel signal that includes a contribution from the decorrelated signal and a second channel signal that does not include a contribution from the decorrelated signal.

[0037] This may provide an advantageous approach in many scenarios and/or embodiments. The first channel signal may be a mid signal and the second channel signal may be a side signal.

[0038] According to an optional feature of the invention, the second set of parameters includes parameters for only a smaller set of time frequency intervals than the first set of parameters.

[0039] This may provide an advantageous approach in many scenarios and/or embodiments. It may in particular in many scenarios provide improved audio quality for a very small data overhead.

[0040] According to an optional feature of the invention, the second set of parameters include fewer parameters for each time frequency interval for which the second set of parameters include parameters than the first set of parameters include for each time frequency interval for which the first set of parameters include parameters.

[0041] This may provide an advantageous approach in many scenarios and/or embodiments. It may in particular in many scenarios provide improved audio quality for a very small data overhead.

[0042] In many embodiments, the first set of parameters may include a first parameter (category/type) that is not included in the second set of parameters. The first parameter (category/type) may be at least one of an interchannel correlation parameter, and interchannel phase difference parameter, or an interchannel time difference parameter.

[0043] According to an aspect of the invention, there is provided an audio apparatus comprising: a downmixer arranged to generate a mono downmix audio signal for an audio stereo signal; a first spatial parameter circuit arranged to determine a first set of spatial parameters for upmixing the mono downmix audio signal to the stereo signal, the first set of spatial parameters being indicative of relative signal properties of channels of the stereo signal; a determination circuit arranged to determine stereo signal components of the stereo signal for which a temporal variation meets a criterion; a second spatial parameter circuit arranged to determine a second set of spatial parameters, the second set of spatial parameters being indicative of relative signal properties of channels of the stereo signal components, and the second set of spatial

parameters having a lower frequency resolution and a higher time resolution than the first set of spatial parameters; and a data signal circuit arranged to generate an audio data signal to comprise the mono downmix audio signal, the first set of spatial parameters, and the second set of spatial parameters.

[0044] According to an aspect of the invention, there is provided an audio data signal comprising: a mono downmix signal of an audio stereo signal, a first set of spatial parameters, the first set of spatial parameters being indicative of relative properties of channel signals of the stereo signal, a second set of spatial parameters, the second set of spatial parameters being indicative of relative properties of channel signals of the stereo signal, the second set of spatial parameters having a lower frequency resolution and a higher time resolution than the first set of spatial parameters.

[0045] 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 downmix signal of an audio stereo signal, a first set of spatial parameters, the first set of spatial parameters being indicative of relative properties of channel signals of the stereo signal, a second set of spatial parameters, the second set of spatial parameters being indicative of relative properties of channel signals of the stereo signal, the second set of spatial parameters having a lower frequency resolution and a higher time resolution than the first set of spatial parameters; decorrelating the mono downmix signal to generate a decorrelated signal; generating a first stereo signal from the mono downmix signal and the decorrelated signal using the first set of spatial parameters; and generating the output audio stereo signal from the first stereo signal in dependence on the second set of spatial parameters.

[0046] According to an aspect of the invention, there is provided a method of generating an audio data signal, the method comprising: generating a mono downmix audio signal for an audio stereo signal; determining a first set of spatial parameters for upmixing the mono downmix audio signal to the stereo signal, the first set of spatial parameters being indicative of relative signal properties of channels of the audio stereo signal; determining stereo signal components of the stereo signal for which a temporal variation meets a criterion; determining a second set of spatial parameters, the second set of spatial parameters being indicative of relative signal properties of channels of the stereo signal components, and the second set of spatial parameters having a lower frequency resolution and a higher time resolution than the first set of spatial parameters; and generating an audio data signal to comprise the mono downmix audio signal, the first set of spatial parameters, and the second set of spatial parameters.

[0047] 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

[0048] Embodiments of the invention will be described, by way of example only, with reference to the drawings, in which

FIG. 1 illustrates an example of some elements of an example of an audio apparatus in accordance with some embodiments of the invention;

FIG. 2 illustrates an example of some elements of an example of an audio apparatus in accordance with some embodiments of the invention;

FIG. 3 illustrates an example of some elements of an example of an audio apparatus in accordance with some embodiments of the invention;

FIG. 4 illustrates an example of some elements of an example of an implementation of a filter bank suitable for an audio apparatus in accordance with some embodiments of the invention;

FIG. 5 illustrates an example of some elements of an example of an implementation of a synthesis filter bank suitable for an audio apparatus in accordance with some embodiments of the invention;

FIG. 6 illustrates an example of some elements of an example of an implementation of a synthesis filter bank suitable for an audio apparatus in accordance with some embodiments of the invention;

FIG. 7 illustrates an example of some elements of an example of an implementation of a synthesis filter bank suitable for an audio apparatus in accordance with some embodiments of the invention;

FIG. 8 illustrates an example of time frequency grids for an audio apparatus in accordance with some embodiments of the invention;

FIG. 9 illustrates an example of time frequency grids for an audio apparatus in accordance with some embodiments of the invention;

FIG. 10 illustrates an example of some elements of an example of an audio apparatus in accordance with some embodiments of the invention;

FIG. 11 illustrates an example of some elements of an example of an audio apparatus in accordance with some embodiments of the invention; and

FIG. 12 illustrates some elements of a possible arrangement of a processor for implementing elements of an audio apparatus in accordance with some embodiments of the invention.

DETAILED DESCRIPTION OF SOME EMBODIMENTS OF THE INVENTION

[0049] FIG. 1 illustrates some elements of an audio apparatus for generating an output audio stereo signal in accordance with some embodiments of the invention. FIG. 2 illustrates an example of an audio apparatus arranged to generate an audio data signal representing a stereo audio signal. The audio data signal generated by the audio apparatus of FIG. 2 may specifically be fed to the audio apparatus of FIG. 1 which may be arranged to generate the output stereo audio signal as a replica of the input stereo audio signal to the audio apparatus of FIG. 2. The audio apparatus of FIG. 1 will also be referred to as a decoder audio apparatus (or just as a decoder) and the audio apparatus of FIG. 2 will also be referred to as an encoder audio apparatus (or just as an encoder).

[0050] The decoder audio apparatus comprises a receiver 101 which is arranged to receive an audio data signal/bitstream comprising a mono downmix audio signal which is a downmix of a stereo audio signal. The audio data signal/bitstream may specifically be one generated by the encoder audio apparatus to represent the original input stereo audio signal.

[0051] In addition to the mono downmix audio signal, the received data signal includes parametric data which comprises spatial parameters that may be used to upmix the mono downmix audio signal. The spatial parameters may specifically be parameters that indicate relationships between the channel signals of input stereo audio signal and/or in some cases between the mono downmix signal and audio channels of the input stereo audio signal. Typically, the spatial parameters may be indicative of time differences, phase differences, level/intensity differences and/or a measure of similarity, such as correlation.

[0052] The spatial parameters may comprise one or more of the following:

- A level difference parameter being indicative of a level difference between channels (and specifically two channels) of the first multichannel audio signal. The level difference parameter may specifically be an Interaural Intensity Difference (IID) and/or an Interaural Level Difference (ILD) as e.g. known from ISO/IEC 23003-3:2020 Information technology - MPEG audio technologies - Part 3: Unified speech and audio coding
- A correlation parameter being indicative of a coherence between channels (and specifically two channels) of the first multichannel audio signal. The correlation parameter may specifically be an Inter-channel Cross Correlation (ICC) parameter as e.g. known from ISO/IEC 23003-3:2020 Information technology - MPEG audio technologies - Part 3: Unified speech and audio coding.
- A phase difference parameter being indicative of a phase difference between channels (and specifically two channels) of the first multichannel audio signal. The phase difference parameter may specifically be an Inter-channel Phase Difference (IPD), Overall Phase Difference (OPD), or Channel Phase Difference (CPD) parameter as e.g. known from ISO/IEC 23003-3:2020 Information technology - MPEG audio technologies - Part 3: Unified speech and audio coding.

[0053] A set of parameters may specifically include an IID, ICC, and IPD parameter determined in accordance with ISO/IEC 23003-3:2020 Information technology - MPEG audio technologies - Part 3: Unified speech and audio coding. In particular, the spatial parameters IID, ICC, and IPD may be determined as:

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

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

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

where l and r represent the signal values of two channels/signals of the first multichannel audio signal (specifically the left and right channel signal of a stereo signal) and $\langle a, b \rangle$ represents the complex-valued inner product between the vectors a and b .

[0054] The encoder audio apparatus is arranged to receive a stereo audio signal and to generate an audio data signal that represents the stereo audio signal with the representation including a mono downmix audio signal and the sets of spatial parameters. Specifically, the encoder audio apparatus may be a Parametric Stereo (PS) based encoder that receives a stereo signal and encodes it as a mono audio signal with associated spatial parametric data.

[0055] Typically, the downmix audio signal is encoded and the receiver 101 is arranged to decode the downmix audio signal to provide the downmix audio signal, i.e. the mono signal in the specific example as well as the sets of spatial parameters and any other required data.

[0056] Typically, the spatial parameters are provided on a per time and per frequency basis. In particular, each parameter value may be provided for a given time interval and for a given frequency band (also referred to as a time frequency tile). Thus, the spatial parameters are provided with a given time and frequency resolution, and specifically may be provided on a time frequency (quantization) grid.

[0057] The audio data signal of the described approach includes a set of spatial parameters (also referred to as the first set of spatial parameters) which are provided at a given time and frequency resolution, henceforth also referred to as the first time and frequency resolution, and specifically the values are provided in accordance with a given/first time frequency grid.

[0058] The first set of spatial parameters may typically include parameters for the full bandwidth and duration of the signal. In some embodiments, the first set of spatial parameters (also referred to as the first parameter set) may comprise a plurality of parameters that cover a (contiguous) frequency interval which includes a plurality (often more than 8 or 16) time frequency tiles, but which may not cover the full frequency interval of the stereo signal.

[0059] The first set of spatial parameters may specifically be a set of parameters corresponding to a PS encoded signal. In particular, for each time frequency tile, the first parameter set may include an interchannel intensity parameter, an interchannel phase difference, and an interchannel correlation parameter. Specifically, for each time frequency tile, the first parameter set may include an IID, IPD, and ICC value as defined above.

[0060] In the described approach, the audio data signal further comprises a second set of spatial parameters. The second set of spatial parameters (also referred to as the second parameter set) comprises additional spatial parameters that provide further information on interchannel properties of the stereo signal.

[0061] The parameters (or at least some of the parameters) of the second parameter set may be provided for time frequency tiles covering time and frequency intervals for which the first set of parameters also includes/provides parameters. The first and second parameter sets may thus provide parameters for overlapping time frequency intervals. Thus, in many embodiments, the second parameter set may provide overlapping and supplemental information to the information provided by the first parameter set.

[0062] In some embodiments, the second parameter set may provide parameter values for the entire bandwidth and duration of the stereo signal. In some embodiments, the second parameter set may include parameters of the frequency and time interval(s) for which the first parameter set includes parameters. However, in many embodiments, the second parameter set may comprise parameters for only one or more time frequency interval subsets of the time frequency interval(s) for which the first parameter set provides parameters. Indeed, in many embodiments, the amount of time frequency intervals covered by parameters of the second parameter set is no more than 5%, 10%, 20%, or 50% of the time frequency intervals covered by parameters of the first parameter set. Thus, in many embodiments, the second parameter set may provide complementary/additional spatial parameters for often small time frequency parts of the stereo signal. The second parameter set may in this way provide additional information of particular signal parts that in many cases may result in the possibility of improved quality of a rendered/generated stereo signal, the second parameter set may include parameters for only a smaller set of time frequency intervals than the first set of parameters.

[0063] The second set of spatial parameters has a lower frequency resolution and a higher time resolution than the first set of spatial parameters. The second parameter set may be provided on a time frequency grid which is finer in the time dimension (lower distance between grid points, smaller time intervals, more compact grid in the time dimension) and coarser in the frequency dimension (higher distance between grid points, larger frequency intervals, less compact grid in that dimension).

[0064] Thus, the second parameter set in particular provides additional parameters that have an increased time resolution/less coarse time quantization but a lower frequency resolution/more coarse frequency quantization. The second parameter set may accordingly provide additional and improved information on parts of the stereo signal with, in particular, improved time responsiveness/detail at the cost of reduced frequency responsiveness/detail. The inventors have realized that such a trade-off/ supplemental data may be particularly advantageous to support a first parameter set to provide overall improved representation of many signal components/events that may occur in many practical signals. For example, the approach may provide a much improved representation and characterization of e.g. transient type signal components.

[0065] In many embodiments, the second parameter set and the first parameter set may provide the same number/type of parameters for a given time frequency tile. For example, the first parameter set and the second parameter set may both provide an interchannel intensity parameter, an interchannel phase (or time) difference, and an interchannel correlation parameter for each time frequency tile. Specifically, for each time frequency tile, the first parameter set may include an IID, IPD, and ICC value as defined above.

[0066] However, in other embodiments, the second parameter set may include fewer parameters for at least some time frequency intervals for which both the first and the second parameter set provides parameters. A given time frequency

interval may be covered by both parameters of the first parameter set and the second parameter set (and thus the time frequency interval may be covered by one or more time frequency tiles of the first parameter set as well as by one or more time frequency tiles of the second parameter set). For such a time frequency interval, there may be fewer parameters in the second parameter set than in the first parameter set. For example, the first parameter set may include interchannel intensity parameter, an interchannel phase (or time) difference, and an interchannel correlation parameter whereas the second parameter set may include only an interchannel intensity parameter, or e.g. may only include an interchannel intensity parameter and an interchannel phase (or time) difference parameter.

[0067] The receiver 101 is coupled to a decorrelator 103 which receives the mono downmix audio signal and which applies a decorrelation thereto to generate a decorrelated (mono downmix audio) signal. It will be appreciated that a large number of different algorithms and functions for decorrelating an audio signal is known to the skilled person, and that any suitable approach or algorithm may be used without detracting from the invention.

[0068] The decorrelator 103 is coupled to a first transformation circuit 105 which is also coupled to the receiver 101. The first transformation circuit 105 receives the mono downmix signal and the decorrelated signal and proceeds to therefrom generate a first stereo signal as a function of the first parameter set. The first transformation circuit 105 may be arranged to perform an upmixing of the mono downmix signal and the decorrelated signal with the upmixing being based on the first parameter set.

[0069] In many embodiments, a first channel signal of the first stereo signal may be generated as a linear combination of the mono downmix signal and the decorrelated signal with the weights of the linear combination for each time frequency interval being determined as a function of at least one parameter of the first parameter set provided for the time frequency interval. Similarly, a second channel signal of the first stereo signal may be generated as a linear combination of the mono downmix signal and the decorrelated signal with the weights of the linear combination for each time frequency interval being determined as a function of at least one parameter of the first parameter set provided for the time frequency interval.

[0070] In many embodiments, time frequency interval/tile values of the first stereo signal may be generated as a result of a matrix multiplication/transformation of a 2 by 2 matrix and a two component vector comprising the time frequency value of the mono downmix signal and the decorrelated signal, and with the coefficients of the matrix being determined as a function of the first parameter set (with this processing being performed in time frequency intervals/tiles).

[0071] It will be appreciated that different approaches and algorithms for generating the first stereo signal based on the first parameter set may be used and that the particular implementation may depend on e.g. the specific spatial parameters, properties, and performance that is desired in the particular implementation. In particular, in some embodiments, the first stereo signal may be generated in accordance with an upmixing such as is used for upmixing when decoding a PS encoded stereo signal. In other embodiments, the first stereo signal may be generated in other ways to provide desired properties. For example, in many embodiments (as will be described in more detail later), the first stereo signal may be generated as a mid-side (sum-difference) stereo signal.

[0072] The first transformation circuit 105 is coupled to a signal generator 107 which is fed the first stereo signal, and which is arranged to generate an output audio stereo signal from the first stereo signal in dependence on the second set of parameters.

[0073] The signal generator 107 may in many embodiments perform a similar operation as that of the first transformation circuit 105 but with the operation being controlled by the second parameter set rather than the first parameter set. In particular, in many embodiments, a first channel signal of the output stereo signal may be generated as a linear combination of the channel signals of the first stereo signal with the weights of the linear combination for each time frequency interval being determined as a function of at least one parameter of the second parameter set provided for the time frequency interval. Similarly, a second channel signal of the output stereo signal may be generated as a linear combination of the channel signals of the first stereo signal with the weights of the linear combination for each time frequency interval being determined as a function of at least one parameter of the second parameter set provided for the time frequency interval.

[0074] In many embodiments, time frequency interval/tile values of the output stereo signal may be generated as a result of a matrix multiplication/transformation of a 2 by 2 matrix and a two component vector comprising the time frequency values of the first stereo signal and with the coefficients of the matrix being determined as a function of the second parameter set (with this processing being performed in time frequency intervals/tiles).

[0075] It will be appreciated that different approaches and algorithms for generating the output stereo signal based on the second parameter set may be used and that the particular implementation may depend on e.g. the specific spatial parameters, properties, and performance that is desired in the particular implementation.

[0076] As previously mentioned, the first parameter set and second parameter set have different time frequency resolutions, and in particular the second parameter set has a higher time resolution and lower frequency resolution than the first parameter set. Correspondingly, the operations of the first transformation circuit 105 and the signal generator 107 are performed at different resolutions. Specifically, the linear combinations/matrix operations may be performed on time frequency tiles/intervals that are different for the operation of the first transformation circuit 105 and the signal generator 107, and in particular with the operation of the signal generator 107 being performed on time frequency tiles that have a

lower resolution in the frequency domain and a higher resolution in the time domain.

[0077] In many embodiments, the signal generator 107 may include a converter which converts the first stereo signal from the time frequency grid of the operation of the first transformation circuit 105 to the time frequency grid of the signal generator 107, i.e. from the time frequency grid of the linear combination/matrix operation of the first transformation circuit 105 to the time frequency grid of the linear combination/matrix operation of the signal generator 107. The signal generator 107 may specifically include a synthesis filter which is arranged to receive the first stereo signal and convert it into a time frequency grid that is used for the upmixing/processing/linear combination to generate the output stereo signal.

[0078] The decoder apparatus of FIG. 1 may accordingly generate an output stereo signal from the received audio data signal using spatial parameters and processing on different time frequency grids. In particular, a first upmixing that may be used to generate a stereo signal based on spatial parameters on a given time frequency grid and this may subsequently be supplemented by additional processing/upmixing that is performed with a higher time resolution but lower frequency resolution. The additional processing may specifically be performed in a relatively small subset of the time and frequency extension of the stereo signal, and thus may in particular be used to adapt (in some cases fine-tune) particular signal components. The approach may in particular allow improved temporal characterization and generation of the output stereo signal thereby in many scenarios allowing improved representation and replication of signal components with high/fast temporal variation. The approach may allow improved audio quality while allowing a low data rate overhead as the second parameter set may typically have a relatively low data rate compared to the first parameter set (e.g. due to second parameter set only providing information on time frequency subintervals) and further allowing low complexity and/or low resource processing to generate the output stereo signal from the received audio data signal.

[0079] The approach reflects the inventor's research and insight that the cues described by the approaches such as Parametric Stereo (PS) system do not provide a complete representation of the spatial image as perceived by the human auditory system. In addition to the conventional approach used for to achieve and represent spatial perception, e.g. as based on the per-band IID, ICC and IPD (or ITD) parameters, the human auditory system is also sensitive to short-lived cues, such as e.g. short term variations in the interchannel intensity differences and/or interchannel time differences, but that this is highly dependent on these properties extending over many (e.g. critical) frequency bands, i.e., that they are consistent over frequency. Such cues can be referred to as "short (term)/fast consistent cues" with the term "short (term)/fast" indicating that the properties are present for a time interval typically significantly shorter than the update rate of regular spatial parameters (such as PS parameters) and the term "consistent" indicating that they are consistently present/manifesting over several, and often many, critical bands.

[0080] The encoder apparatus of FIG. 2 may generate the audio data signal that is fed to the decoder apparatus of FIG. 1. The encoder apparatus comprises a receiver 201 which receives a stereo signal to be encoded by the audio data signal. The stereo signal will also be referred to as the input or original stereo signal. In many embodiments, the receiver 201 may directly receive a stereo signal which is to be represented by the data signal. In other embodiments, the receiver 201 may additionally or alternatively receive a number of audio components such as audio objects, mono signals, single source signals, multichannel signals etc.

[0081] The input stereo signal is fed to a downmixer 203 which proceeds to generate a mono downmix signal for the input stereo signal. In many embodiments and scenarios, the downmixer 203 may generate a mono downmix audio signal from a received stereo signal, e.g. simply by summing the channels of the stereo signal in accordance with a standard PS downmix approach.

[0082] The encoder apparatus of FIG. 2 further comprises a first spatial parameter circuit 205 which is arranged to determine the first set of spatial parameters for the stereo signal represented by the mono downmix audio signal. Specifically, the received stereo signal may be fed to the parameter circuit 205 which may determine the spatial parameters from an analysis/processing of the stereo signal.

[0083] The first spatial parameter circuit 205 may provide a first set of frequency subband spatial parameters for the stereo signal where the set of frequency subband spatial parameters are indicative of relative signal properties of the channels of the stereo signal. The frequency subband spatial parameters are provided for individual subbands of the stereo signal.

[0084] The spatial parameters are indicative of/reflect relative properties of the channel signals of the stereo audio signal. In particular, the spatial parameters may include parameters that are indicative of at least one of relative intensities/levels of the stereo channels, relative (frequency domain) phases of the stereo channels, a relative time difference between the channels, and/or a correlation between the channels. Specifically, the spatial parameters may include one or more of an inter-channel intensity difference, inter-channel level difference, inter-channel time difference, inter-channel phase difference, and/or inter-channel correlation.

[0085] The first parameter set may specifically comprise spatial parameters as used for encoding a stereo signal using a Parametric Stereo (PS) encoding of the stereo signal, such as for example by a mono downmix audio signal given as:

$$m = c(l + r)$$

where the parameter c is chosen such that the power of the stereo signal is preserved in the downmix, the power being defined using the 2-norm:

$$\|m\|^2 = \|l\|^2 + \|r\|^2$$

and thus e.g.:

$$c = \sqrt{\frac{\|l\|^2 + \|r\|^2}{\|l + r\|^2}}$$

[0086] The PS parameters are specifically an Inter-channel Intensity Difference IID, an Inter-channel Correlation ICC, and in some cases an Inter-channel Phase Difference IPD parameter. These may specifically be defined/determined as:

$$\text{IID} = \frac{\|l\|^2}{\|r\|^2}$$

$$\text{ICC} = \frac{|\langle l, r \rangle|}{\sqrt{\|l\|^2 \|r\|^2}}$$

$$\text{IPD} = \arg \langle l, r \rangle$$

where the complex-valued inner product is defined as:

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

and:

$$\|x\|^2 = \langle x, x \rangle$$

[0087] The spatial parameters are typically provided for specific time frequency tiles, and thus specifically each parameter value is generated/provided for a given frequency subband and for a given time segment.

[0088] The first spatial parameter circuit 205 may determine spatial parameters that are indicative of relative properties of the channels (channel signals) of the stereo signal.

[0089] In many embodiments, the first spatial parameter circuit 205 may receive the input stereo signal and process/analyze this to generate the spatial parameters. Specifically, the first spatial parameter circuit 205 may calculate the IID, ICC, and IPD values in accordance with the formulas indicated above.

[0090] The first set of spatial parameters may comprise sets of spatial parameters for time frequency time intervals/tiles, each set of spatial parameters comprising at least one of: a level difference parameter indicative of a level difference between channels of the multichannel audio signal; a correlation parameter indicative of a coherence between channels of the multichannel audio signal; a timing difference parameter indicative of a timing difference between channels of the multichannel audio signal, and a phase difference parameter indicative of a phase difference between channels of the multichannel audio signal.

[0091] The encoder apparatus of FIG. 2 further comprises a second spatial parameter circuit 207 which is arranged to determine the second set of spatial parameters for the stereo signal represented by the mono downmix audio signal. Specifically, the received stereo signal may be fed to the second parameter circuit 207 which may determine the spatial parameters from an analysis/processing of the stereo signal.

[0092] The second spatial parameter circuit 207 may provide a second set of frequency subband spatial parameters for the stereo signal where the set of frequency subband spatial parameters are indicative of relative signal properties of the channels of the stereo signal. The frequency subband spatial parameters are provided for individual subbands of the stereo signal. Indeed, the comments provided above with reference to the first parameter set may apply equally to the second parameter set (*mutatis mutandis*).

[0093] However, as previously described, the second parameter set comprises parameters with a higher time resolution

and a lower frequency resolution. The second parameter set is typically provided on a time frequency grid which has a finer time granularity and a coarser frequency granularity.

[0094] In addition, the second parameter set typically includes parameters for only a subset of the time frequency region for which the first parameter set provides parameters. Indeed, typically, the second parameter set may only provide parameters for time and frequency regions which are determined to have particular properties that may particularly benefit from the second parameter set having a higher time resolution. In many embodiments, the first time frequency grid may be a predetermined fixed time frequency grid whereas the second time frequency grid may be a variable or signal dependent time frequency grid that specifically may be selected to be suitable for the specific signal components additionally represented by the second set of spatial parameters.

[0095] The encoder apparatus of FIG. 2 further comprises a determination circuit 209 which is arranged to determine stereo signal components of the stereo signal that have a temporal variation which meet a given criterion. The determination circuit 209 may specifically be arranged to determine signal components that have a sufficiently fast variation, e.g. it may identify signal components which exhibit a fast change. As a specific example, the determination circuit 209 may determine signal components corresponding to transients or step changes in the stereo signal.

[0096] In some embodiments, the determination may be both in the time and frequency domain, i.e. signal components may be determined for individual frequency bands if a temporal change in those frequency bands is sufficiently high in accordance with a given criterion. For example, the determination circuit 209 may include a filter bank dividing the stereo signal into frequency subbands with each subband then being analyzed to determine a suitable temporal change/variation measure. Signal components may then be determined as individual frequency subbands during time intervals where the temporal change/variation measure meet the criterion (e.g. exceeds a threshold).

[0097] The determination circuit 209 may typically be arranged to operate with a frequency division and time division that matches the time frequency grid of the second parameter set. Similarly, the second spatial parameter circuit 207 may operate at this resolution and specifically the receiver 201 may in some embodiments be arranged to transform the stereo signal into a time frequency representation with the granularity corresponding to the time frequency grid of the second parameter set. This representation may then be fed to the determination circuit 209 to identify (typically groups/regions) of time frequency tiles for which a temporal variation measure exceeds a threshold.

[0098] The second spatial parameter circuit 207 is typically arranged to determine the second parameter set for only the time frequency tiles (/regions) identified by the determination circuit 209 and thus a very compact second parameter set may typically be generated.

[0099] In many embodiments, the time frequency grid for the second set of spatial parameters may further be determined to be suitable for representing the properties of the detected signal components. For example, the time resolution may be set to be sufficient to represent the temporal variations of the detected signal components.

[0100] The downmixer 203, first spatial parameter circuit 205, and second spatial parameter circuit 207 are coupled to a data signal circuit 211 which generates the audio data signal to comprise the mono downmix audio signal, the first set of parameters, and the second set of parameters. It will be appreciated that any suitable approach, encoding, and format for the audio data signal may be used without depending on the requirements and preferences of the individual embodiment.

[0101] The encoder apparatus of FIG. 2 further comprises a data signal circuit 211 which generates the audio data signal and specifically it generates the audio data signal to include data representing the mono downmix audio signal and data representing the spatial parameters. The data signal circuit 211 may specifically include an audio signal encoder arranged to encode the mono downmix audio signal. It will be appreciated that any suitable audio signal encoding algorithm and approach may be used, such as an Advanced Audio Coding (AAC) mono encoding algorithm. Likewise, the spatial parameters may be encoded using a suitable encoding approach.

[0102] The data signal circuit 211 may generate an output data signal in accordance with any suitable format, and may specifically generate the output data signal to follow a suitable standard for an audio signal.

[0103] In the following, a particularly advantageous implementation of a decoder audio apparatus in accordance with some embodiments of the invention will be described. The decoder audio apparatus will be described with reference to FIG. 3 which may be considered an implementation of the decoder audio apparatus of FIG. 1.

[0104] In the example, the receiver 101 receives an audio data signal comprising mono downmix signal $M'[k, l]$ which is already in a time frequency representation. Specifically, a signal value is provided for each frequency bin/interval k in each time segment l . In many embodiments, the processing of the first transformation circuit 105 may be performed in the received time frequency grid. However, in the specific embodiment of FIG. 3, the receiver 101 may further increase the received frequency resolution by splitting some frequency bands into smaller frequency bands.

[0105] Specifically, the received mono downmix signal $M'[k, l]$ may be a quadrature mirror filter (QMF) domain mono signal $M'[k, l]$ and this may by the receiver 101 be converted into a hybrid QMF (hQMF) domain signal $M''[k, l]$. The receiver 101 may for example implement a structure as illustrated in FIG. 4 which shows a hybrid portion of a hybrid analysis filterbank where the lower frequency bands are split into a plurality of smaller frequency bands without impacting the time resolution of the hybrid subband domain samples. The increased frequency resolution may typically result in an improved perceived audio quality with the increased frequency resolution in particular being provided for frequency bands of high

perceptual relevance.

[0106] In the example, the resulting hQMF domain mono downmix signal $M''[k,l]$ is further processed and is specifically fed to the decorrelator 103 which may include a set of decorrelator filters to generate the decorrelated hQMF domain signal $D[k,l]$.

[0107] The hQMF domain mono downmix signal $M''[k,l]$ and the decorrelated hQMF domain signal $D[k,l]$ are fed to the first transformation circuit 105 which is arranged to generate a stereo signal, henceforth also referred to as the intermediate signal, based on the first parameter set.

[0108] In some embodiments, the first transformation circuit 105 may be arranged to generate the intermediate stereo signal to have properties as indicated by the first parameter set. Indeed, in some embodiments, the first parameter set may be generated in accordance with a PS encoding approach and indeed the intermediate stereo signal may be generated using PS decoding approach. Thus, in some embodiments, the intermediate stereo signal may be generated to resemble the original stereo signal as much as possible as indicated by the first parameter set.

[0109] The first transformation circuit 105 may specifically generate the intermediate stereo signal with the channel signals being linear combinations of the mono downmix signal and the decorrelated signal. For each time frequency tile, the first transformation circuit 105 may generate each channel signal sample as a linear combination of the signal samples of the time frequency tile of the mono downmix signal and the decorrelated signal respectively where the weights/coefficients are determined from the first parameter set. Specifically, the first transformation circuit 105 may multiply a 2x2 matrix of coefficients by a two value vector comprising the two signal samples for the time frequency tile (i.e. the signal samples of the mono downmix signal and the decorrelated signal respectively) to generate a two signal sample vector for the intermediate signal.

[0110] Specifically, the intermediate signal may for each time frequency tile be generated as:

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

where the parameters α (typically complex-valued), β and c are a function of the first parameter set, l, r are the left and right channel signals of the intermediate signal, and m, d represent respectively the samples of the mono downmix signal and the decorrelated signal for the time frequency tile.

[0111] The parameters α , β and c may e.g. be determined as:

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

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

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

[0112] However, in many embodiments, the first transformation circuit 105 is arranged to generate a modified intermediate signal and is specifically arranged to generate the intermediate signal as a mid-side (/sum-difference) stereo signal for the original stereo signal.

[0113] In particular, instead of directly generating the left and right (hQMF) signals of the desired stereo signal, a partial upmix is applied, resulting in a mid- and a side-signal $M'''[k,l]$ and $S[k,l]$ respectively. This can be achieved by decomposing the upmix matrix:

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

$$= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \frac{1}{2c} \begin{bmatrix} 1 & 0 \\ \alpha & \beta \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} m_{mid} \\ s \end{bmatrix} = \frac{1}{2c} \begin{bmatrix} 1 & 0 \\ \alpha & \beta \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix},$$

$$\begin{bmatrix} l \\ r \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} m_{mid} \\ s \end{bmatrix}$$

where m_{mid} is the mid signal and s is the side signal.

[0114] In the specific approach, the first transformation circuit 105 generates the intermediate stereo signal such that the decorrelated signal provides a contribution to only one of the channel signals. In particular, the coefficients for the matrix are determined such that one of the coefficients for the decorrelated signal are zero. Such an approach has been found to provide simplified operation/reduced complexity while still achieving high audio quality.

[0115] In the specific approach indicated above, the decorrelated signal provides only a contribution to the side signal and not to the mid signal, i.e. the coefficient of row 1, column 2 of the matrix is zero such that the mid signal, in contrast to the side signal, does not contain any component of the decorrelated signal, i.e. $M''(k,l) = M'(k,l)$.

[0116] The intermediate stereo signal is fed to the signal generator 107 which proceeds to perform a post-processing where the intermediate stereo signal is modified based on the second parameter set to generate an output stereo signal. In particular, the intermediate stereo signal is modified with a higher temporal resolution but a lower frequency resolution to provide an output stereo signal that may more closely resembles the original stereo signal. The signal generator 107 may in some embodiment fine-tune/adapt the output stereo signal over short time intervals to provide a temporally more accurate replication of the stereo signal for e.g. transients or other fast changing components.

[0117] For example, in situations where the intermediate stereo signal is directly generated to be a replica of the stereo signal as reflected by the second parameter set, the signal generator 107 may in the particular time frequency intervals for which the second parameter set comprise spatial parameters proceed to modify the generated intermediate stereo signal such that the generated output stereo signal in these time frequency intervals has properties that match the properties indicated by the parameters of the second parameter set.

[0118] The signal generator 107 accordingly apply processing of the signals with a higher time resolution but lower frequency resolution. In many embodiments, the signal generator 107 includes a conversion circuit that converts at least parts of the intermediate stereo signal from the time frequency grid used by the first transformation circuit 105 (and specifically the linear combination/upmixing/matrix multiplication) to a time frequency grid used by the signal generator 107 to process the intermediate stereo signal to generate the output stereo signal.

[0119] In the example of FIG. 3, the signal generator 107 includes a set of conversion circuits and specifically synthesis filter 301, 303 which convert the channel signals of the intermediate stereo signal to the time frequency grid of processing of the signal generator 107. In the specific example, a first synthesis filter 301 converts the mid-signal and a second synthesis filter 303 converts the side signal. Indeed, in the specific example where the intermediate stereo signal is a mid-side signal, the second synthesis filter 303 may directly convert the side signal to the time domain, i.e. the second synthesis filter 303 may generate a time domain side signal.

[0120] An example of elements of such a second synthesis filter 303 is illustrated in FIG. 5. In the specific example, the first (hybrid) QMF bands of $S[k,l]$ are simply summed to form the QMF domain signal $S'[k,l]$. The QMF domain signal $S'[k,l]$ is synthesized by upsampling and filtering the QMF domain signals and summing the output. Since the QMF domain samples are typically complex-valued, only the real part of the output is preserved to create the output signal $s[n]$.

[0121] Further, in the example of a mid-side intermediate stereo signal, the (time domain) side signal may not be processed or adapted depending on the second parameter set but rather may directly be respectively added/subtracted from the channel output stereo signal. The contribution of the side signal to the channel signals of the output stereo signal are specifically of equal magnitude but opposite phase for the channel signals of the output stereo signal.

[0122] The mid-signal, or indeed typically both channel signals for a non mid-side scenario, may be converted into a time frequency grid suitable for the further processing in dependence on the second parameter set.

[0123] The channel signals, and in the particular example the mid-signal, is further processed, but on a different time/frequency grid than the one used by the first transformation circuit 105 to generate the intermediate stereo signal. For this purpose, the first synthesis filter 301 may comprise a synthesis filterbank S_1 which combines groups of QMF bands (hence it is assumed that the hybrid part of the hQMF signal is already combined to form a QMF domain signal), resulting in a time frequency grid that has a higher time resolution, but a lower frequency resolution.

[0124] FIG. 6 illustrates an example of elements of a suitable synthesis filter, e.g. for implementing the first synthesis filter 301. FIG. 6 illustrates an example of a filterbank configuration for lowering the frequency resolution and increasing the time resolution with a factor of R (where $R=4$ in this case).

[0125] It will be appreciated that in some embodiments, the time frequency grid used for the processing by the first transformation circuit 105 may be the same as the time frequency grid on which the parameters of the first parameter set are provided and/or the time frequency grid used for the processing by the signal generator 107 may be the same as the

time frequency grid on which the parameters of the second parameter set are provided. However, in other embodiments, this may potentially not be the case, and a conversion of the spatial parameters from the time frequency grid in which they are provided to that of the processing of the corresponding signal may be performed. For example, increased resolution may be achieved using interpolation between the provided parameters, and reduced resolution may be achieved by combination/averaging of the provided parameters.

[0126] The signal generator 107 is arranged to generate a contribution to both channel signals of the output stereo signal from the first channel signal of the intermediate stereo signal, i.e. specifically from the mid signal.

[0127] In the example, the first channel signal/mid signal is fed to a second transformation circuit 305 which is arranged to upmix the first stereo signal using coefficients that are determined from the second parameter set. In the specific example of FIG. 3, the second transformation circuit 305 is fed only the mid signal and is arranged to generate a second intermediate stereo signal by upmixing the single mid signal based on the second parameter set. In particular, the signal samples of the (converted) time frequency tiles may be multiplied by coefficients/weights that are determined from the second parameter set.

[0128] In embodiments where the intermediate signal is generated to directly correspond to the original signal as represented by the first parameter set, the second transformation circuit 305 may be fed both channel signals and directly apply a conversion, e.g. by applying a matrix operation that multiplies a vector of the signal samples of the intermediate stereo signal by a 2x2 matrix, where the matrix coefficients are determined from the second parameter set. In such cases, a unity matrix may be used for all time frequency tiles for which no second parameter set is provided (corresponding to no change/multiplication) whereas for the time frequency tiles for which the second parameter set provides parameters, the signal samples are multiplied by a matrix with coefficients that will result in a stereo signal having interchannel properties as indicated by the second parameter set.

[0129] In the example of FIG. 3, the second intermediate stereo signal is fed to a third synthesis filter 307 and a fourth synthesis filter 309 which are arranged to convert the generated time frequency channel signals to the time domain. FIG. 7 illustrates a possible implementation of (either of) the third synthesis filter 307 and a fourth synthesis filter 309.

[0130] It is noted that the filter banks can be realized efficiently by employing a polyphase decomposition and efficient implementation of the resulting modulation matrix, e.g. by an FFT with pre- and post-twiddle factors.

[0131] The resulting time domain signals are fed to respectively a first combiner 311 and a second combiner 313 which are arranged to combine the contribution from the second intermediate stereo signal with the contribution from the side signal. In the present case, the time domain side signal is added to one channel signal of the output stereo signal and subtracted from the other channel signal, i.e. the side signal is combined with the second intermediate stereo signal by being added as a difference signal. In order to compensate for the processing delay of the processing generating the second intermediate stereo signal, the time domain side signal is delayed by a delay 315.

[0132] It will be appreciated that in embodiments where the intermediate stereo signal is not a mid-side signal, the second transformation circuit 305 may as described receive both channel signals and there may be no path corresponding to the side signal path of FIG. 3, i.e. the second synthesis filter 303, the delay 315, and the combiners 311, 313 may not be included in the signal generator 107.

[0133] The processing of the decoder audio apparatus may thus employ signals, parameters, and processing on a range of different time frequency grids/resolutions. As a specific example, FIG. 8 shows a simplified view of a time- and frequency resolution of the different (hybrid) QMF domain samples for the example of FIG. 3. The QMF domain signal $M'[k,l]$, is converted to the hQMF domain $M''[k,l]$ ($M'''[k,l]$) by splitting the lower frequency bands, without changing the time resolution. Hybrid synthesis follows the reverse process. Finally, synthesis filterbanks change both the time- and frequency resolution of $M'''[k,l]$, producing $N[q,r]$, which has higher time, but lower frequency resolution.

[0134] The synthesis of the quick consistent cues represented by the second parameter set may be realized by applying a left and right (typically complex-valued) gain profile to the altered resolution QMF domain samples $N[q,r]$. A simple example, having two quick consistent IID cue areas is shown in FIG. 9. The particular example shows two consistent IID cues that have been synthesized as gain values in a time frequency grid (the gray scale indicates gain values with the neutral predominant greyscale value indicating a gain of zero (corresponding to time frequency tiles for which the second parameter set provides no parameters)).

[0135] The example reflects a typical example where the amount of quick consistent cues is limited to specific time frequency regions. These may typically be encoded as areas on the time frequency plane, i.e., for a frame the following data may e.g. be provided:

- The number of consistent cues
- For each consistent cue:
 - Start position (lowest time instance, lowest band)
 - Size (nr of time slots, nr of bands)
 - Cue: IID, and possibly IPD/ITD value

[0136] For the regions for which the second parameter set includes parameters, the earlier mentioned upmix matrix may be used to reinstate cues. In case an ITD value is transmitted, it can be synthesized as an actual time delay, converted to a phase rotation, or a combination of both.

[0137] Different approaches and algorithms for determining suitable parameters for the upmixing by the second transformation circuit 305 may be used in different embodiments. In order to generate the output stereo signal based on the second parameter set from the intermediate stereo signal, which is generated by an upmixing/transform based on the first parameter set, the second transformation circuit 305 may take into account the processing based on the first parameter set. For example, it may be provided with information of the parameters used by the upmixing by the first transformation circuit 105.

[0138] For example, where the first transformation circuit 105 generates the intermediate stereo signal as a PS compatible mid-side signal, i.e. as:

$$\begin{bmatrix} m_{mid} \\ s \end{bmatrix} = \frac{1}{2c} \begin{bmatrix} 1 & 0 \\ \alpha & \beta \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix}$$

the parameters α and β may be provided to the second transformation circuit 305. When applying two complex-valued gains to the QMF domain signal $N[q,r]$, it can be shown that effectively the following signals are generated:

$$L[q, r] = g_1 \cdot N[q, r] + S[q, r]$$

$$R[q, r] = g_2 \cdot N[q, r] - S[q, r]$$

where $N[q,r]$ is the QMF domain mid-signal, $S[q,r]$ is the QMF domain representation of $s[n]$ and g_1 and g_2 are two complex-valued gains.

[0139] Rewriting this, leads to:

$$L[q, r] = \frac{1}{2c} ((g_1 + \alpha_{local}) \cdot M[q, r] + \beta_{local} \cdot D[q, r])$$

$$R[q, r] = \frac{1}{2c} ((g_2 - \alpha_{local}) \cdot M[q, r] - \beta_{local} \cdot D[q, r])$$

where α_{local} and β_{local} are approximations of the α and β parameters at time frequency grid $[q,r]$ from the processing by the first transformation circuit 105 at corresponding $[k, l]$ in the hybrid QMF domain of the second transformation circuit 305.

[0140] It can be shown that the complex-valued gain parameters are then restricted by:

1. Power preservation:

$$|g_1 + \alpha_{local}|^2 + |g_2 - \alpha_{local}|^2 = |1 + \alpha_{local}|^2 + |1 - \alpha_{local}|^2$$

2. Synthesis desired IID:

$$IID_{desired} = \frac{|g_1 + \alpha_{local}|^2 + \beta_{local}^2}{|g_2 - \alpha_{local}|^2 + \beta_{local}^2}$$

3. Synthesis desired IPD:

$$IPD_{desired} = \angle \{ (g_1 + \alpha_{local}) \cdot (g_2 - \alpha_{local})^* - \beta_{local}^2 \}$$

[0141] These considerations may then be used to determine the required parameters g_1 and g_2 to be used for the upmixing of the mid signal.

[0142] It will be appreciated that the encoder can be realized in different ways depending on the preferences and

requirements of the individual embodiment.

[0143] FIG. 10 illustrates a particular example using a parallel architecture. The top part corresponds to a legacy PS encoder. The left and right signals $l[n]$, $r[n]$ are converted to the (hybrid) QMF domain, where the IID, ICC and IPD parameters are estimated in pre-defined bands. The PS parameters are used to construct the mono downmix $M[k, l]$. In parallel, the left and right time domain signals are fed to two analysis filterbanks that produce a different time frequency grid, operating at a lower frequency resolution and a higher time resolution. In the corresponding left and right time frequency grids $L[q, r]$, $R[q, r]$, perceptually relevant cues are extracted (quick consistent cue estimation block). Several features can be extracted to indicate perceptual relevance:

- Cues should be short-lived; long-term cues will be captured by the regular PS parameterization.
- Cues should be consistent over a frequency range; this can e.g. be checked by measuring the standard deviation of a cue (e.g. IID) over the frequency range. Only if the standard deviation is below a certain threshold is it considered.
- Cues should be distinct enough; only if a (consistent) cue is strong enough compared to the background cues will it be picked up by the human auditory system.

[0144] In some embodiments of the cue estimation block, a log-IID (equiv. IPD/ITD) parameter map $P_{IID}[q, r]$ is constructed from the left and right time frequency grids such that,

$$P_{IID}[q, r] = 10 \log_{10} \frac{|L[q, r]|^2}{|R[q, r]|^2}$$

[0145] Next, the mean of the parameter map is removed,

$$P'_{IID}[q, r] = P_{IID}[q, r] - \frac{1}{N_q N_r} \sum_q \sum_r P_{IID}[q, r]$$

where N_q and N_r are the total number of frequency and time tiles, respectively.

[0146] Next, the tiles can be thresholded such that only perceptually relevant (sufficiently deviating from the mean) tiles are preserved with all other tiles set to zero. These non-zero regions can then be tested against the requirements listed above, i.e. given the time width of a region, and the consistency of IID values within a contiguous region of non-zero IIDs.

[0147] Finally, the two sets of parameters may be combined into a single parameter bit-stream.

[0148] An alternative approach is provided in FIG. 11. The left and right time-domain signals are first converted to the first time frequency grid, with the lower frequency resolution and the higher time resolution resulting in the QMF domain signals $L[q, r]$ and $R[q, r]$. These signals are used to estimate the "quick consistent cues". Then, the QMF domain signals $L[q, r]$ and $R[q, r]$ are further filtered by applying a small filterbank on each band q . Optionally, the hybrid analysis part of FIG. 4 can be added to realize a desired higher frequency resolution for the lower frequency region.

[0149] It is noted that, as an alternative to extracting the complex-valued gains in the decoder from the desired IID, IPD (ITD) and the local background stereo parameters, the gains may be directly extracted and transmitted to the decoder audio apparatus.

[0150] The audio apparatus(es) may specifically be implemented in one or more suitably programmed processors. In particular, the decoder audio apparatus and/or the encoder audio apparatus may be implemented in one or more such suitably programmed processors. The different functional blocks 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.

[0151] FIG. 12 is a block diagram illustrating an example processor 1200 according to embodiments of the disclosure. Processor 1200 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 1200 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.

[0152] The processor 1200 may include one or more cores 1202. The core 1202 may include one or more Arithmetic Logic Units (ALU) 1204. In some embodiments, the core 1202 may include a Floating Point Logic Unit (FPLU) 1206 and/or a Digital Signal Processing Unit (DSPU) 1208 in addition to or instead of the ALU 1204.

[0153] The processor 1200 may include one or more registers 1212 communicatively coupled to the core 1202. The registers 1212 may be implemented using dedicated logic gate circuits (e.g., flip-flops) and/or any memory technology. In some embodiments the registers 1212 may be implemented using static memory. The register may provide data,

instructions and addresses to the core 1202.

[0154] In some embodiments, processor 1200 may include one or more levels of cache memory 1210 communicatively coupled to the core 1202. The cache memory 1210 may provide computer-readable instructions to the core 1202 for execution. The cache memory 1210 may provide data for processing by the core 1202. In some embodiments, the computer-readable instructions may have been provided to the cache memory 1210 by a local memory, for example, local memory attached to the external bus 1216. The cache memory 1210 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.

[0155] The processor 1200 may include a controller 1214, which may control input to the processor 1200 from other processors and/or components included in a system and/or outputs from the processor 1200 to other processors and/or components included in the system. Controller 1214 may control the data paths in the ALU 1204, FPLU 1206 and/or DSPU 1208. Controller 1214 may be implemented as one or more state machines, data paths and/or dedicated control logic. The gates of controller 1214 may be implemented as standalone gates, FPGA, ASIC or any other suitable technology.

[0156] The registers 1212 and the cache 1210 may communicate with controller 1214 and core 1202 via internal connections 1220A, 1220B, 1220C and 1220D. Internal connections may be implemented as a bus, multiplexer, crossbar switch, and/or any other suitable connection technology.

[0157] Inputs and outputs for the processor 1600 may be provided via a bus 1616, which may include one or more conductive lines. The bus 1616 may be communicatively coupled to one or more components of processor 1600, for example the controller 1614, cache 1610, and/or register 1612. The bus 1616 may be coupled to one or more components of the system.

[0158] The bus 1616 may be coupled to one or more external memories. The external memories may include Read Only Memory (ROM) 1632. ROM 1632 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) 1633. RAM 1633 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) 1635. The external memory may include Flash memory 1634. The External memory may include a magnetic storage device such as disc 1636. In some embodiments, the external memories may be included in a system.

[0159] 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.

[0160] 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.

[0161] 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 audio apparatus for generating an output audio stereo signal, the audio apparatus comprising:

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

a mono downmix signal of an audio stereo signal,

a first set of spatial parameters, the first set of spatial parameters being indicative of relative properties of channel signals of the audio stereo signal,
a second set of spatial parameters, the second set of spatial parameters being indicative of relative properties of channel signals of the audio stereo signal, the second set of spatial parameters having a lower frequency resolution and a higher time resolution than the first set of spatial parameters;

a decorrelator (103) arranged to decorrelate the mono downmix signal to generate a decorrelated signal;
a first transformation circuit (105) arranged to generate a first stereo signal from the mono downmix signal and the decorrelated signal using the first set of parameters; and
a signal generator (107) arranged to generate the output audio stereo signal from the first stereo signal in dependence on the second set of parameters.

2. The audio apparatus of claim 1 wherein the signal generator (107) comprises a second transformation circuit (305) arranged to upmix the first stereo signal using coefficients determined from the second set of parameters; and wherein the first transformation circuit (105) is arranged to upmix the mono downmix signal and the decorrelated signal on a first time frequency grid, and the signal generator (107) is arranged to upmix the first stereo signal on a second time frequency grid, the second time frequency grid having a higher time resolution and a lower frequency resolution than the first time frequency grid.

3. The audio apparatus of any previous claim wherein the signal generator (107) is arranged to generate the output stereo signal to include a first contribution from a first channel signal of the first stereo signal and a second contribution from a second channel signal of the first stereo signal, the first contribution being dependent on the second set of spatial parameters and the second contribution not being dependent on the second set of spatial parameters.

4. The audio apparatus of claim 3 wherein the first contribution is an intermediate stereo signal generated by upmixing the first channel signal dependent on the second set of parameters.

5. The audio apparatus of claim 4 wherein the signal generator (107) includes a synthesis filter (301) for converting the first channel signal from a time frequency grid of the generation of the first stereo signal by the first transformation circuit (105) to a time frequency grid of the upmixing of the first channel signal, the time frequency grid of the upmixing of the first channel signal having a higher time resolution and a lower frequency resolution than the time frequency grid of the generation of the first stereo signal by the first transformation circuit (105).

6. The audio apparatus of any of claims 3-5 wherein the second contribution is of equal magnitude but opposite phase for the channel signals of the output stereo signal.

7. The audio apparatus of any previous claim wherein the first transformation circuit (105) is arranged to generate the first stereo signal as a mid-side stereo signal.

8. The audio apparatus of any previous claim wherein the first transformation circuit (105) is arranged to generate the first stereo signal with a first channel signal that includes a contribution from the decorrelated signal and a second channel signal that does not include a contribution from the decorrelated signal.

9. The audio apparatus of any previous claim wherein the second set of parameters includes parameters for only a smaller set of time frequency intervals than the first set of parameters.

10. The audio apparatus of any previous claim wherein the second set of parameters include fewer parameters for each time frequency interval for which the second set of parameters include parameters than the first set of parameters include for each time frequency interval for which the first set of parameters include parameters.

11. An audio apparatus comprising:

a downmixer (203) arranged to generate a mono downmix audio signal for an audio stereo signal;
a first spatial parameter circuit (205) arranged to determine a first set of spatial parameters for upmixing the mono downmix audio signal to the stereo signal, the first set of spatial parameters being indicative of relative signal properties of channels of the stereo signal;
a determination circuit (209) arranged to determine stereo signal components of the stereo signal for which a temporal variation meets a criterion;

a second spatial parameter circuit (207) arranged to determine a second set of spatial parameters, the second set of spatial parameters being indicative of relative signal properties of channels of the stereo signal components, and the second set of spatial parameters having a lower frequency resolution and a higher time resolution than the first set of spatial parameters; and

a data signal circuit (211) arranged to generate an audio data signal to comprise the mono downmix audio signal, the first set of spatial parameters, and the second set of spatial parameters.

12. An audio data signal comprising:

a mono downmix signal of an audio stereo signal,
 a first set of spatial parameters, the first set of spatial parameters being indicative of relative properties of channel signals of the stereo signal,
 a second set of spatial parameters, the second set of spatial parameters being indicative of relative properties of channel signals of the stereo signal, the second set of spatial parameters having a lower frequency resolution and a higher time resolution than the first set of spatial parameters.

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

receiving an audio data signal comprising:

a mono downmix signal of an audio stereo signal,
 a first set of spatial parameters, the first set of spatial parameters being indicative of relative properties of channel signals of the stereo signal,
 a second set of spatial parameters, the second set of spatial parameters being indicative of relative properties of channel signals of the stereo signal, the second set of spatial parameters having a lower frequency resolution and a higher time resolution than the first set of spatial parameters;

decorrelating the mono downmix signal to generate a decorrelated signal;
 generating a first stereo signal from the mono downmix signal and the decorrelated signal using the first set of spatial parameters; and
 generating the output audio stereo signal from the first stereo signal in dependence on the second set of spatial parameters.

14. A method of generating an audio data signal, the method comprising generating a mono downmix audio signal for an audio stereo signal;

determining a first set of spatial parameters for upmixing the mono downmix audio signal to the stereo signal, the first set of spatial parameters being indicative of relative signal properties of channels of the audio stereo signal;
 determining stereo signal components of the stereo signal for which a temporal variation meets a criterion;
 determining a second set of spatial parameters, the second set of spatial parameters being indicative of relative signal properties of channels of the stereo signal components, and the second set of spatial parameters having a lower frequency resolution and a higher time resolution than the first set of spatial parameters; and
 generating an audio data signal to comprise the mono downmix audio signal, the first set of spatial parameters, and the second set of spatial parameters.

15. A computer program product comprising computer program code means adapted to perform all the steps of claims 13 or 14 when said program is run on a computer.

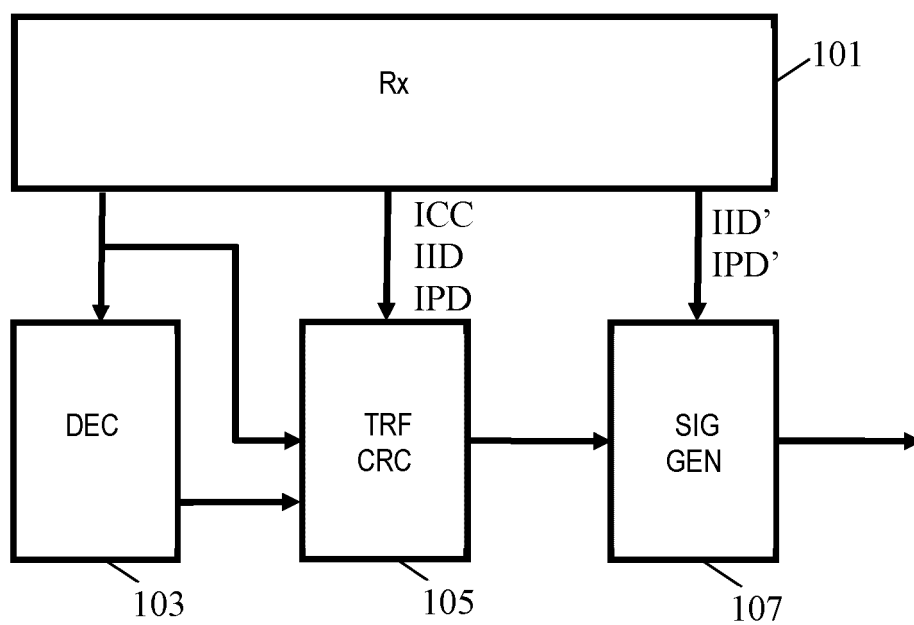


FIG. 1

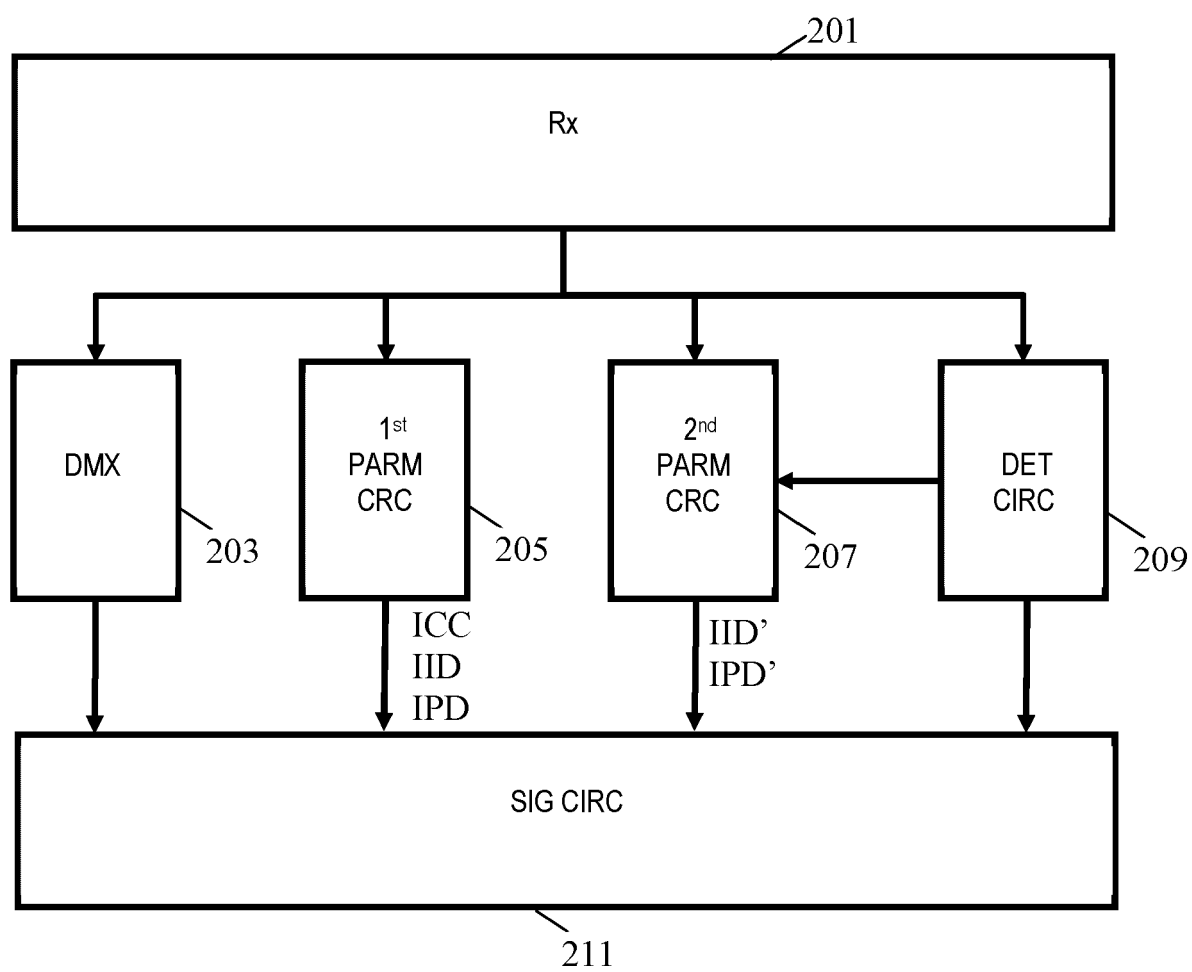


FIG. 2

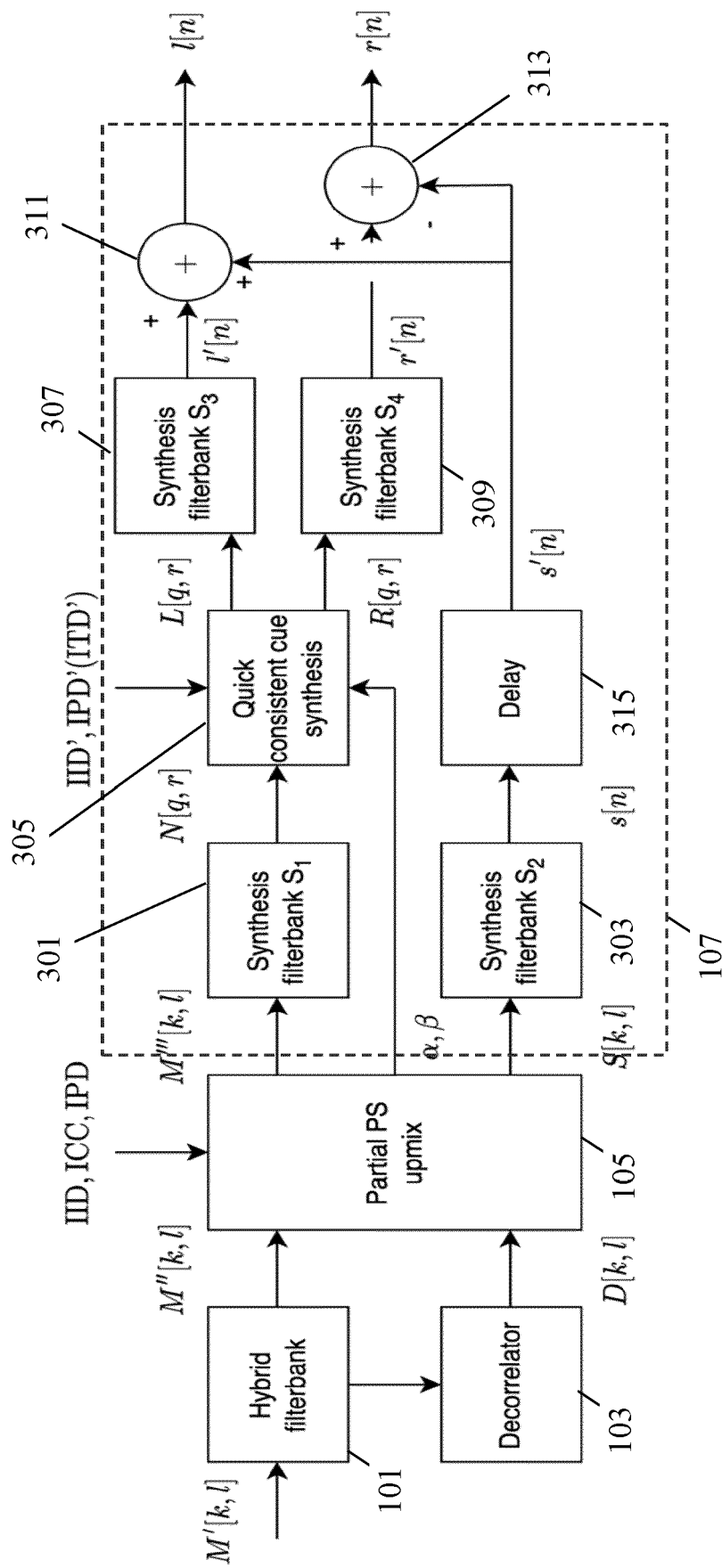
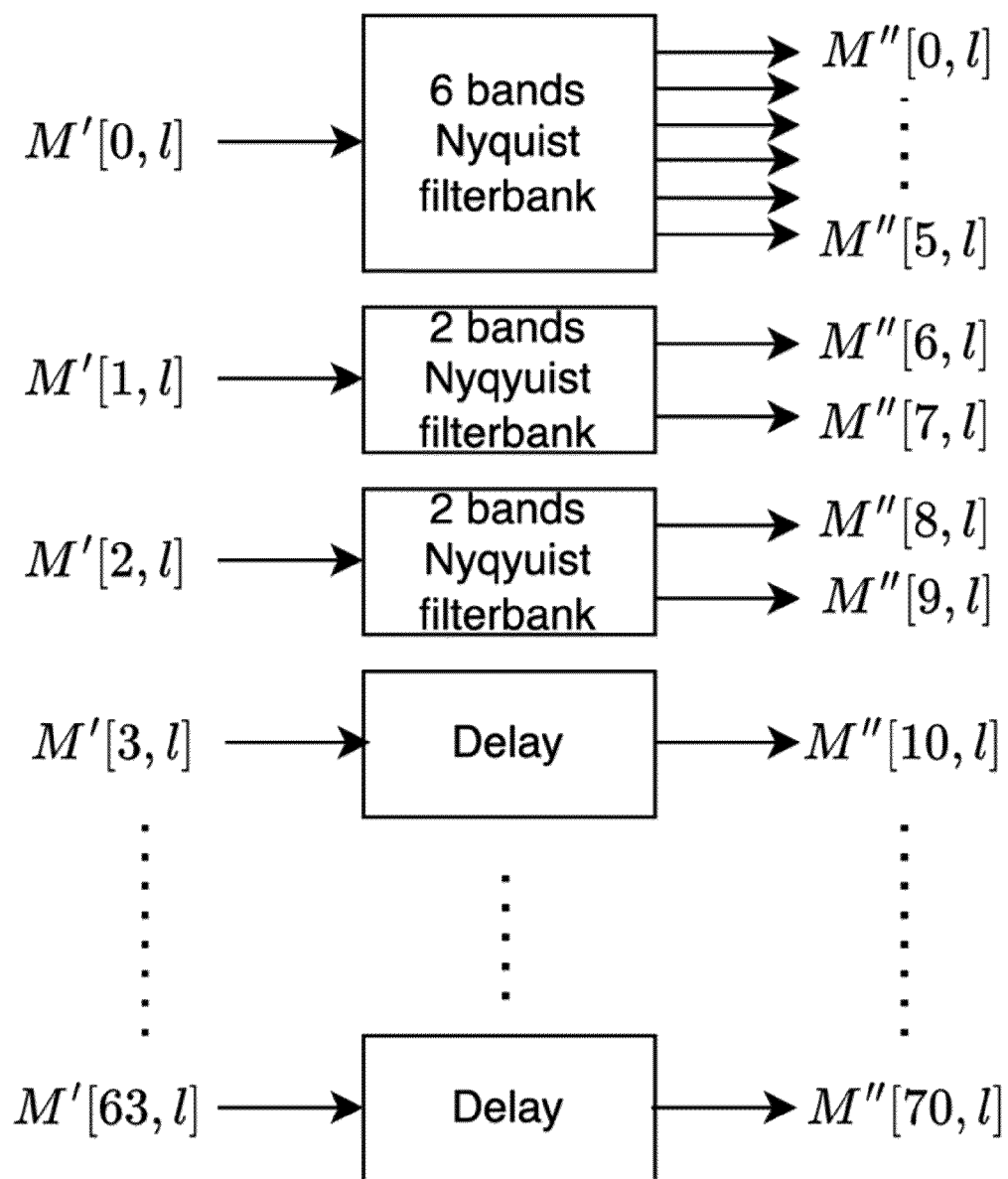


FIG. 3

**FIG. 4**

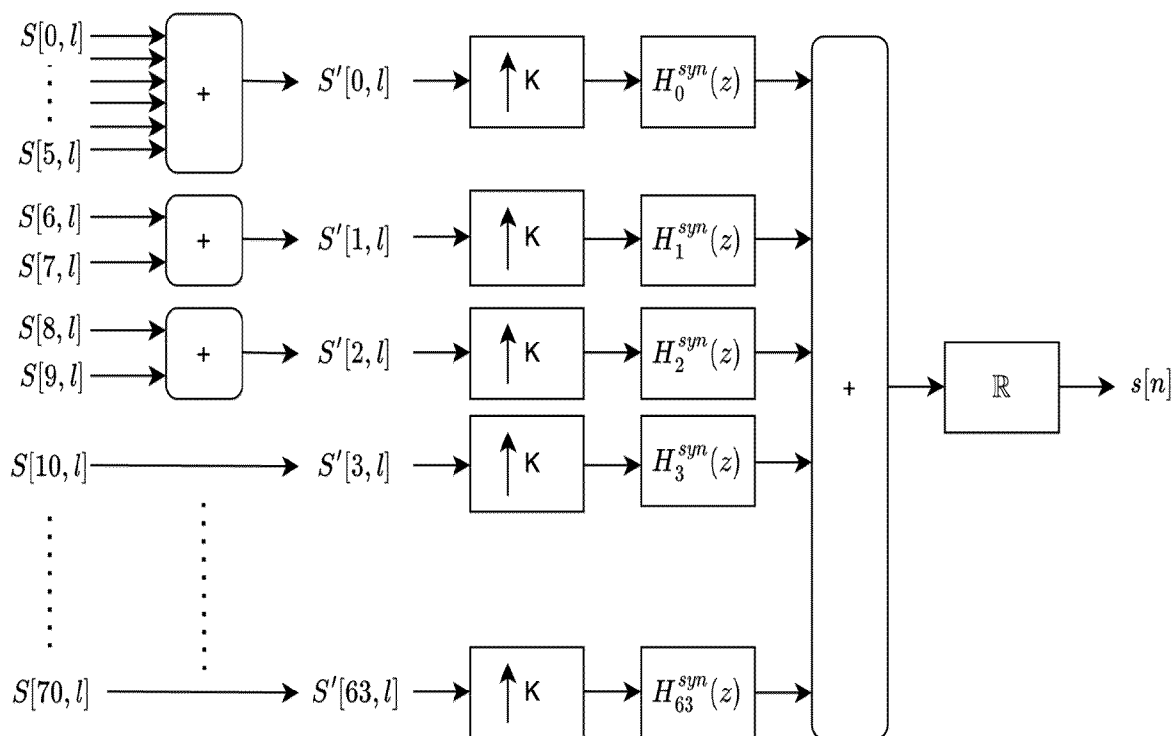


FIG. 5

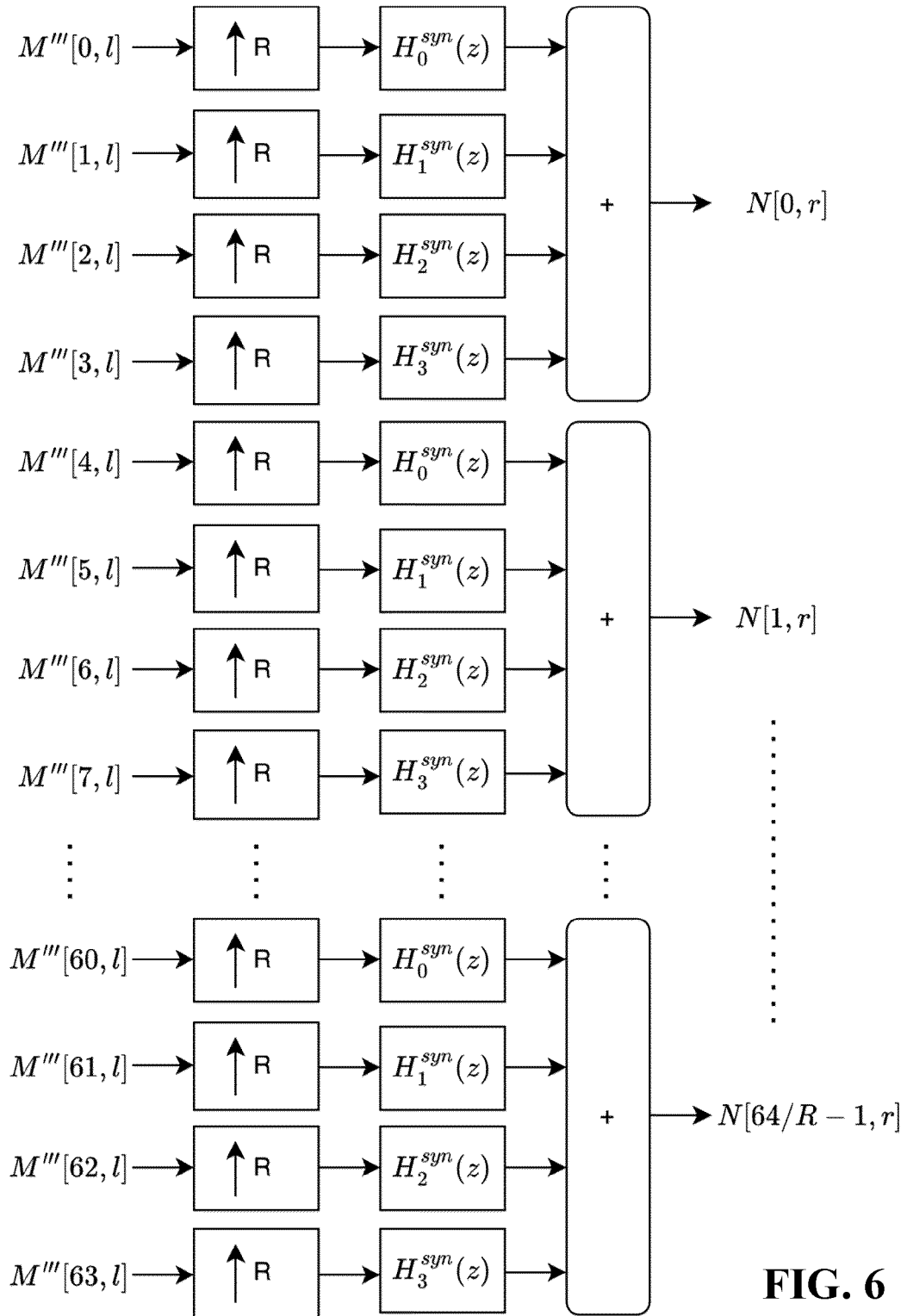


FIG. 6

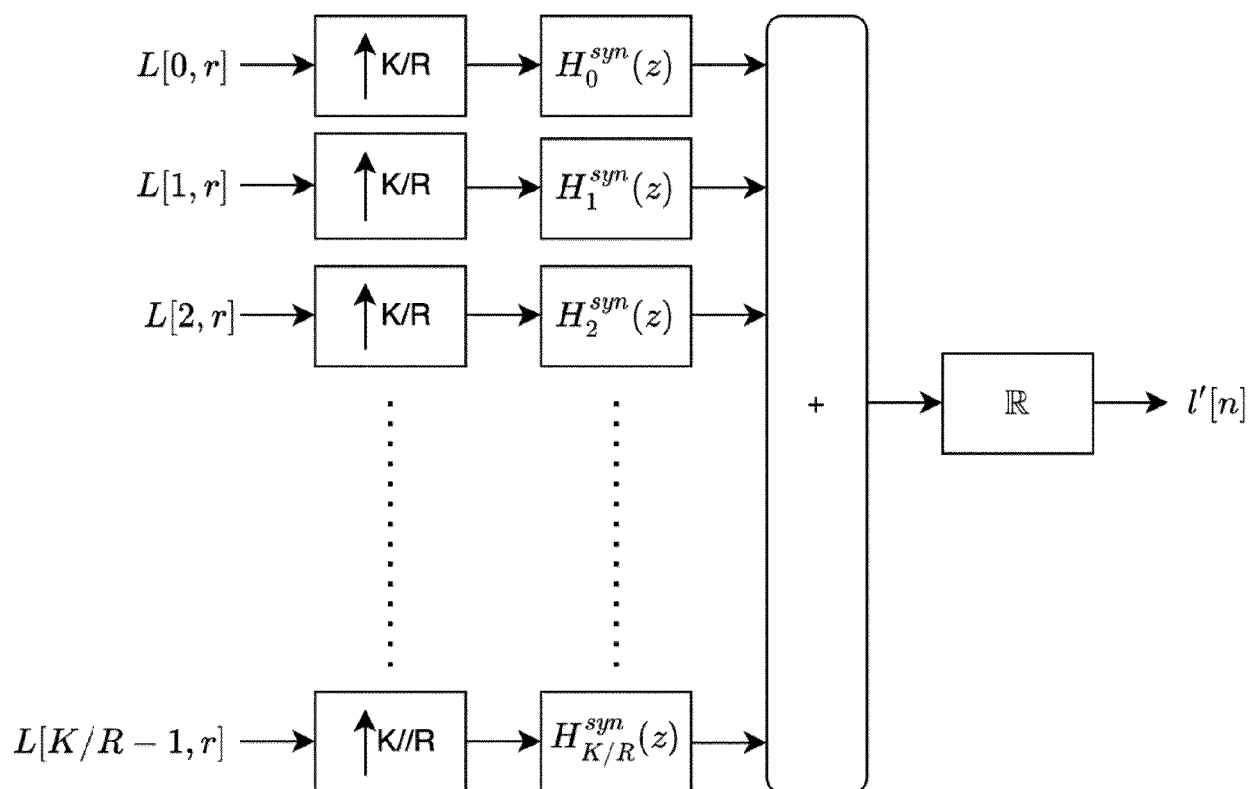


FIG. 7

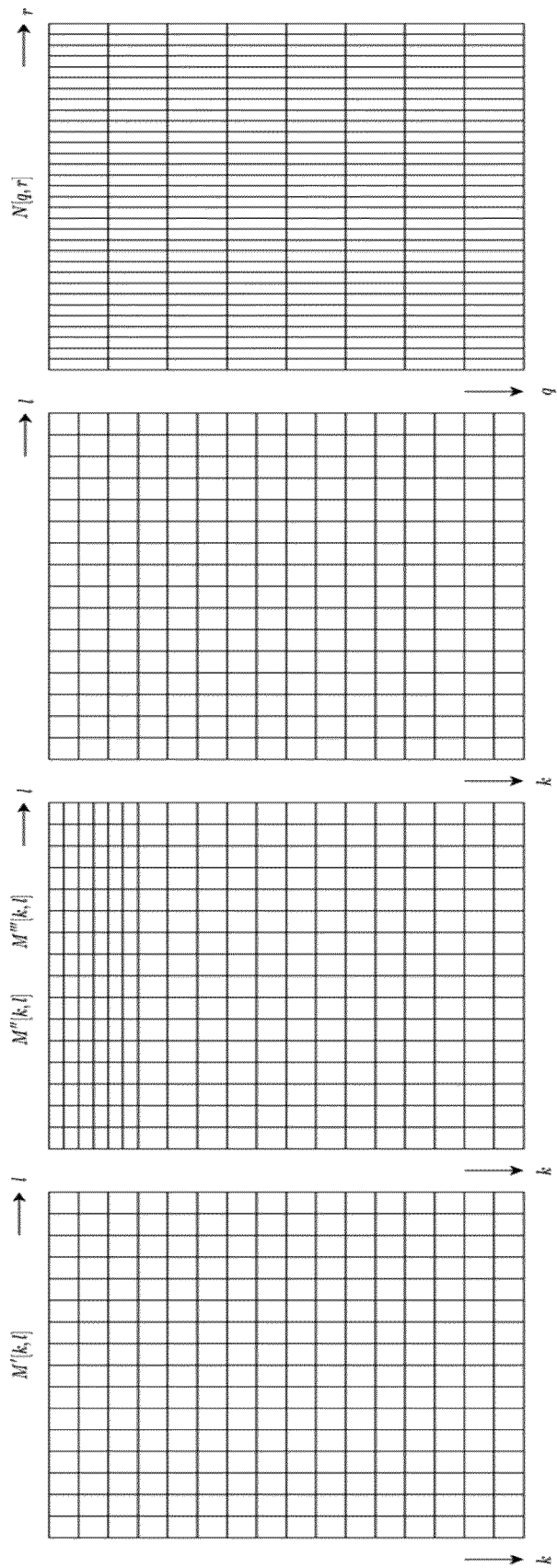


FIG. 8

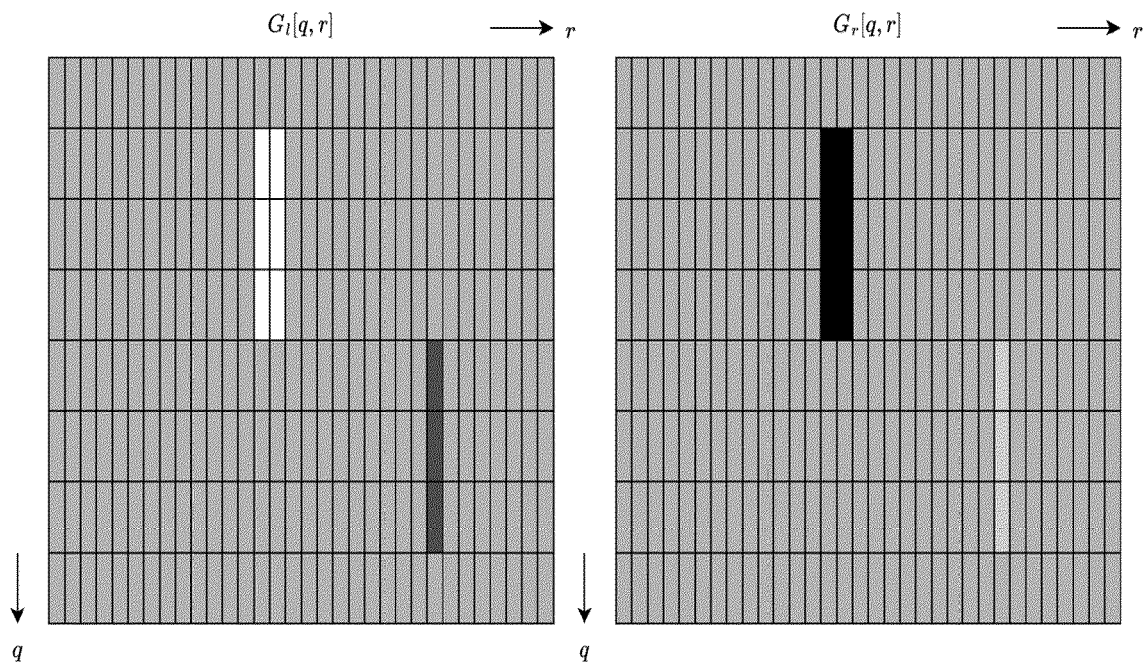


FIG. 9

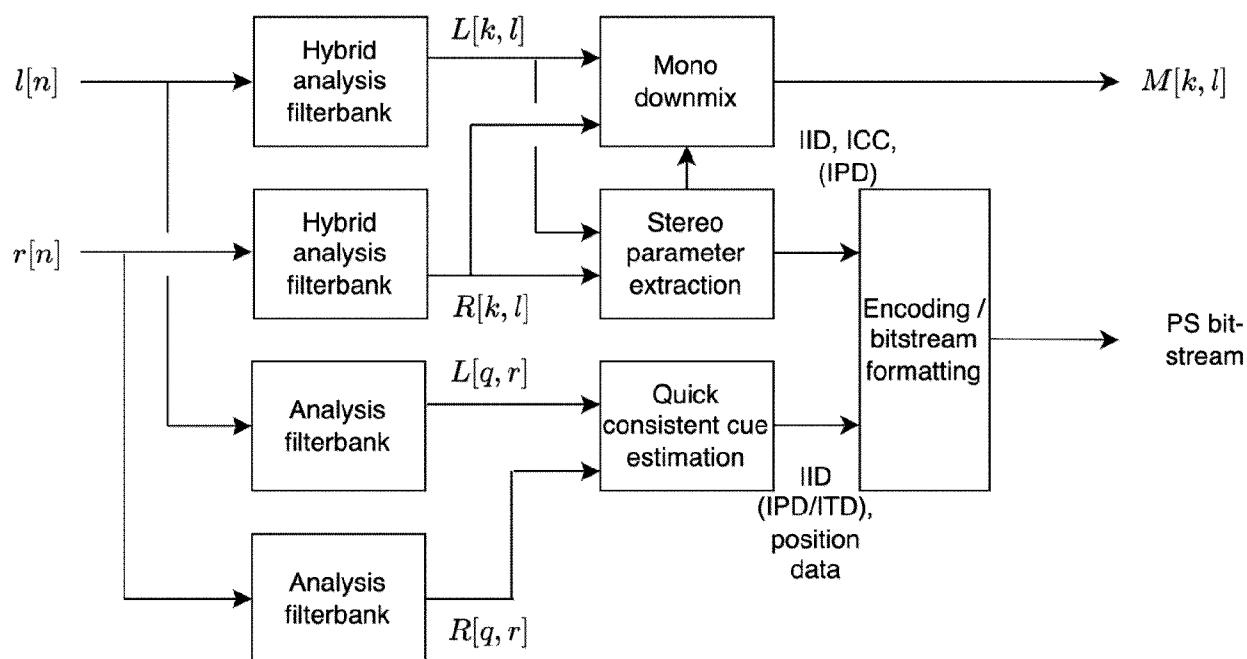


FIG. 10

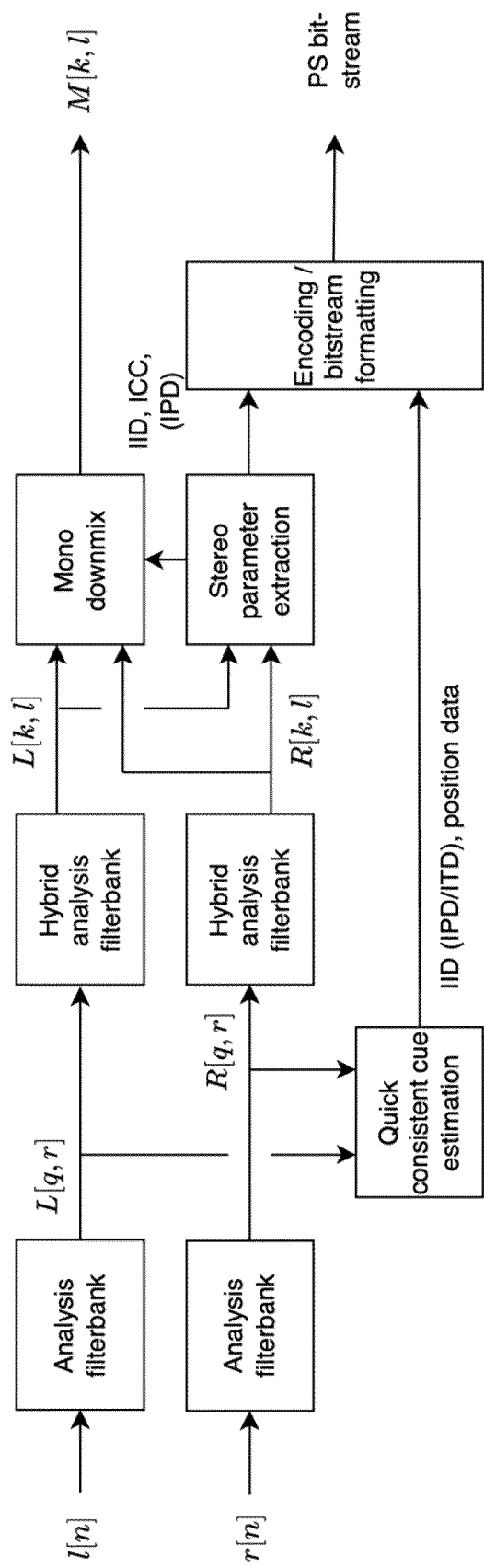


FIG. 11

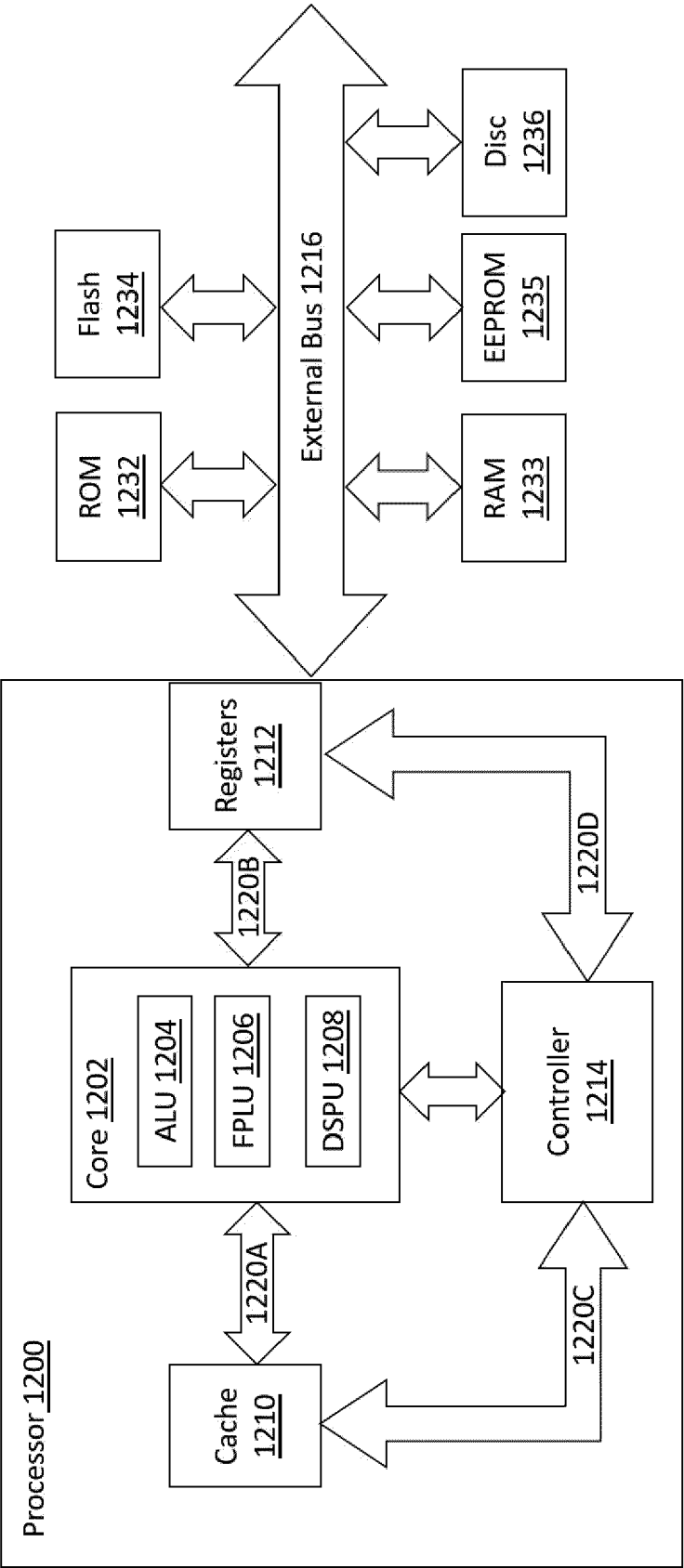


FIG. 12



EUROPEAN SEARCH REPORT

Application Number

EP 25 16 8242

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,D	SCHUIJERS E ET AL: "Advances in Parametric Coding for High-Quality Audio", AUDIO ENGINEERING SOCIETY CONVENTION PAPER, NEW YORK, NY, US, no. 5852, 22 March 2003 (2003-03-22), pages 1-11, XP002365269, * section 3.3, from page 6 to page 9 *	11,12, 14,15	INV. G10L19/008
A	* section 3.3, from page 6 to page 9 *	1-10,13	
A	US 2015/213806 A1 (DISCH SASCHA [DE] ET AL) 30 July 2015 (2015-07-30) * paragraph [0092] - paragraph [0102]; figures 1A,1B,2A,2B *	1-15	
A	US 2006/136229 A1 (KJOERLING KRISTOFER [SE] ET AL) 22 June 2006 (2006-06-22) * paragraph [0044] - paragraph [0054] *	1-15	
A	US 2016/232900 A1 (VILLEMOES LARS [SE] ET AL) 11 August 2016 (2016-08-11) * the whole document *	1-15	
A,D	SCHUIJERS ERIK ET AL: "Low Complexity Parametric Stereo Coding", AES CONVENTION 116; MAY 2004, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, 1 May 2004 (2004-05-01), XP040506843, * sections 2 and 3.6; figures 2,7 *	1-15	TECHNICAL FIELDS SEARCHED (IPC) G10L
The present search report has been drawn up for all claims			
Place of search The Hague		Date of completion of the search 3 July 2025	Examiner De Meuleneire, M
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 (P24C01)

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

EP 25 16 8242

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.

03-07-2025

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2015213806 A1	30-07-2015	AR 092927 A1	06-05-2015
		AU 2013326516 A1	23-04-2015
		BR 112015007532 A2	22-08-2017
		CA 2887228 A1	10-04-2014
		CN 104838442 A	12-08-2015
		EP 2717261 A1	09-04-2014
		EP 2904609 A1	12-08-2015
		ES 2703327 T3	08-03-2019
		HK 1213360 A1	30-06-2016
		JP 6285939 B2	28-02-2018
		JP 2015535958 A	17-12-2015
		KR 20150073180 A	30-06-2015
		MX 345497 B	02-02-2017
		MY 191498 A	28-06-2022
		PL 2904609 T3	30-04-2019
		PT 2904609 T	30-01-2019
		RU 2015116434 A	27-11-2016
		SG 11201502608X A	28-05-2015
		TR 201900436 T4	21-02-2019
		TW 201423728 A	16-06-2014
		US 2015213806 A1	30-07-2015
		WO 2014053537 A1	10-04-2014
		ZA 201503024 B	28-09-2016

US 2006136229 A1	22-06-2006	AT E387705 T1	15-03-2008
		DE 602005005083 T2	26-03-2009
		EP 1807824 A1	18-07-2007
		EP 1909265 A2	09-04-2008
		HK 1106862 A1	20-03-2008
		TW I330826 B	21-09-2010
		US 2006136229 A1	22-06-2006
		WO 2006058590 A1	08-06-2006

US 2016232900 A1	11-08-2016	CN 105531761 A	27-04-2016
		EP 3044783 A1	20-07-2016
		JP 6212645 B2	11-10-2017
		JP 2016536646 A	24-11-2016
		US 2016232900 A1	11-08-2016
		WO 2015036350 A1	19-03-2015

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*, 2003 [0006]
- **E. SCHUIJERS ; J. BREEBAART ; H. PURNHA-GEN ; J. ENGDEGÅRD**. Low Complexity Parametric Stereo Coding. *116th AES*, 2004 [0006]