



(51) International Patent Classification:
G10L 19/00 (2006.01) *H04S 3/00* (2006.01)

(21) International Application Number:
PCT/IB2010/050763

(22) International Filing Date:
22 February 2010 (22.02.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09153900.7 27 February 2009 (27.02.2009) EP

(71) Applicant (for all designated States except US): **KONINKLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **OOMEN, Arnoldus, W., J.** [NL/NL]; c/o High Tech Campus Building 44, NL-5656 AE Eindhoven (NL). **SCHUIJERS, Erik, G., P.** [NL/NL]; c/o High Tech Campus Building 44, NL-5656 AE Eindhoven (NL). **BREEBAART, Dirk, J.** [NL/NL]; c/o High Tech Campus Building 44, NL-5656 AE Eindhoven (NL).

(74) Agents: **UITTENBOGAARD, Frank** et al.; High Tech Campus Building 44, NL-5656 AE Eindhoven (NL).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

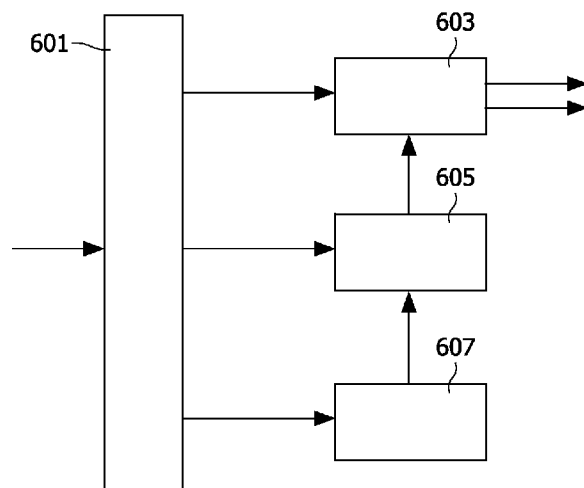
Declarations under Rule 4.17:

— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))

Published:

— with international search report (Art. 21(3))

(54) Title: PARAMETRIC STEREO ENCODING AND DECODING



215

FIG. 6

(57) Abstract: A decoder comprises a receiver (601) which receives a parametrically encoded signal comprising a downmix of multiple channels, a set of upmix parameters including a phase parameter, and a phase correction parameter. The phase correction parameter is provided for second parts of the downmix but not for first parts of the downmix. An upmixer (603, 605) generates the multi-channel signal by upmixing the downmix based on the set of upmix parameters and a modifying unit (607) modify the upmixing in response to the phase correction parameter for the second parts of the downmix. The second parts may be associated with scenarios wherein at least two of the encoded multi-channel signals are out of phase. The invention may provide improved audio quality for a reduced data rate.

Parametric stereo encoding and decoding

FIELD OF THE INVENTION

The invention relates to parametric stereo encoding and decoding.

BACKGROUND OF THE INVENTION

Digital encoding of various source signals has become increasingly important over the last decades as digital signal representation and communication increasingly has replaced analogue representation and communication. For example, mobile telephone systems, such as the Global System for Mobile communication, are based on digital speech encoding. Also distribution of media content, such as video and music, is increasingly based on digital content encoding.

One encoding technique is known as Parametric Stereo (PS) and includes downmixing a stereo signal to a mono signal which is then encoded. In addition, parameters are generated that allows a stereo signal resembling the original stereo signal to be recreated by upmixing of the mono signal.

Fig. 1 illustrates an example of a PS encoding/ decoding scheme. The scheme is based on the generation of an appropriate mono down-mix. Along with the calculation of the mono down-mix, parameters are calculated that enable a PS decoder to regenerate the stereo signal. PS schemes generally rely on a time-frequency representation, which can be e.g. based on a Discrete Fourier Transform (DFT) for parameter analysis and synthesis or a Quadrature Mirror Filterbank (QMF) for a lower-complexity alternative.

Examples of parameters that are used in different PS approaches include:

- *IID*: the Inter-channel Intensity Difference (typically given in dB).
- *IPD*: the Inter-channel Phase Difference (typically given in radians or degrees).
- *OPD*: the Overall Phase Difference (typically given in radians or degrees).
- *IC or sometimes the ICC*: the Inter-channel Coherence (typically calculated to be independent of inter-channel phase differences).

The encoder typically estimates such parameters for each sub-band in each time frame based on the mono downmix and the original stereo signals.

At the decoder, it is possible to create a stereo synthesis using various such parameters. For example, some schemes (such as Intensity Stereo) use the IID value to recreate stereo signals that have appropriate signal levels in the two channels. Other systems use both IID and IC to add a certain sense of ambience (captured by IC) and sound source positions (IID). It has also been proposed to use phase differences (IPD and OPD) as these contain important sound source localization properties. Thus, PS can be used to provide signals of different qualities and has the advantage of being scalable.

PS has been standardized within the MPEG-4 audio standard (ISO/IEC 14496-3:2005 Part 3: Audio). It has successfully been adopted into the High-Efficiency Advanced Audio Coding - HE-AAC v2 profile and also by the 3rd Generation Partnership Project 3GPP Release 6 as Enhanced aacPlus for use in cellular communication systems. However, these latter standards use PS approaches with a profile/level that does not use the phase parameters IPD and OPD as described in ISO/IEC 14496-3:2005 Part 3: Audio. Nevertheless, for higher bitrates it is advantageous to exploit phase parameters as these provide significant improvements in perceived quality.

Furthermore, the PS technology has been employed as a building block in constructing a multi-channel audio coding tool, MPEG Surround (ISO/IEC 23003-1:2007 Part 1: MPEG Surround). There it is referred to as a One-To-Two (OTT) module, since it effectively expands a single channel into two channels.

Within a recent MPEG Unified Speech and Audio coding work item (ISO/IEC 23003-3), it has been proposed to use phase parameters with PS encoding. In particular, the standards have introduced the use of IPD parameters in a very limited fashion. Namely, it allows that a broadband +90 degrees phase shift between left and right can be signaled on a frame-by-frame basis.

However, in Jimmy Lapierre and Roch Lefebvre, "On Improving Parametric Stereo Audio Coding", Presented at the 120th Convention, 2006 May 20–23 Paris, Audio engineering Society, France, Preprint 6804a more elaborate scheme is introduced. This scheme effectively employs the same upmix as the traditional PS scheme when phase information is included. However, in addition it is proposed to derive the OPD parameters from the IPD, ICC and IID parameters.

However, such approaches tend to provide suboptimal performance in many scenarios. Hence, an improved parametric stereo approach would be advantageous and in particular an approach allowing increased flexibility, reduced data rate, increased quality,

reduced complexity, facilitated implementation and/or improved performance would be advantageous.

SUMMARY OF THE INVENTION

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

According to an aspect of the invention there is provided a decoder for generating a multi-channel signal comprising: means for receiving a parametrically encoded signal comprising a downmix of multiple channels, a set of upmix parameters including a phase parameter, and a phase correction parameter, the phase correction parameter being provided for second parts of the downmix but not for first parts of the downmix; upmixing means for generating the multi-channel signal by upmixing the downmix based on the set of upmix parameters;

modifying means for modifying the upmixing in response to the phase correction parameter for the second parts of the downmix.

The inventors have in particular realized that whereas the use of phase information with parametric multichannel (and specifically stereo) encoding may provide improved performance for some parts of a parametrically encoded multichannel signal, it may also provide suboptimal performance for other parts. In particular, the inventors have realized that whereas an upmixing considering phase parameters may provide improved performance when the multichannel signal has certain characteristics, it may also provide reduced performance when the multichannel signal has other characteristics. Furthermore, the inventors have realized that improved operation can be achieved by only including additional phase information, in the form of a phase correction parameter, for some parts of the downmix and not for other parts of the downmix. This may specifically allow reduced data rate while maintaining a suitably high quality level. In particular, the inventors have realized that when estimating upmix phase operations, these can be derived appropriately from parameters such as IID and IPD for some multichannel signal characteristics but not for other multichannel signal characteristics. Typically, such estimation is unsuitable for signals of the multiple channels that are sufficiently out of phase with each other. This may lead to significant and noticeable audio quality degradation. The inventors have realized that such disadvantages may be mitigated while maintaining a suitably low data rate by a decoder being adapted to modify the upmix operation based on the presence of a phase correction parameter for only some parts of the signal.

The upmixing is specifically in response to the phase parameter which e.g. may be an IPD parameter. The upmixing accordingly reflects a phase characteristic of the multiple channels. Each of the parameters (including the phase correction parameter) may be provided as a set of parameter values for frequency and time intervals (e.g. time frequency tiles or blocks). E.g. a set of parameter values may be provided for each of a set of frequency and time blocks. The second parts may correspond to a subset of frequency time blocks. The set of parameters may specifically include IID, IC and/or IPD parameters.

The phase correction parameter may be an absolute phase correction parameter and may specifically be a replacement phase parameter. In other scenarios, the phase correction parameter may provide relative correction values, such as a relative offset or correction to a parameter value that is calculated by or employed in the decoder.

In accordance with an optional feature of the invention, the decoder further comprises: first determining means arranged to determine an overall phase offset in response to the set of upmix parameters for the first parts of the downmix; second determining means arranged to determine the overall phase offset in response to the phase correction parameter for the second parts of the downmix; and wherein the upmixing means is arranged to upmix the downmix based on the overall phase offset.

The invention may provide an improved audio quality to data rate ratio in many scenarios. Specifically, the approach may allow the upmixed signal to reflect overall phase offsets without artifacts and degradations often associated with such approaches. In particular, the invention may allow the upmix to take the overall phase offset into account with substantially improved performance for out of phase signals. For parts of the signal, the overall phase offset can be calculated from other PS parameters.

The overall phase offset may specifically be an OPD parameter.

In accordance with an optional feature of the invention, the overall phase offset is indicative of a phase difference between the downmix and at least one of the multiple channels.

This may allow improved performance and/or facilitated operation and/or facilitated implementation.

In accordance with an optional feature of the invention, the second determining means is arranged to further determine the overall phase offset in response to the set of upmix parameters.

This may in many embodiments provide improved quality and/or reduced complexity.

In accordance with an optional feature of the invention, the downmix is a phase compensated downmix for the second parts relative to the first parts, and the phase correction parameter is indicative of the phase compensation.

This may provide improved performance in many scenarios. Specifically, the system may allow a downmix to be generated wherein the contributions from two stereo channels do not cancel each other even if the input signals are out of phase. This may allow improved performance and an improved use of phase information in the upmix. In addition, it may allow the energy variation of the downmix signal to be reduced. In many embodiments, the phase compensation may ensure that the downmix and set of parameters is such that the decoder can calculate suitable phase parameters to use in the upmix from the set of parameters. Furthermore, as the compensation is different for the first and second parts, the compensation can be optimized for different signal characteristics in the two parts. This may e.g. allow a reduced data rate.

The phase compensation for the second parts may for example be represented by:

$$S = \alpha \cdot \exp(j\varphi_1)L + \beta \cdot \exp(j\varphi_2)R$$

where S is the phase compensated downmix, L and R are the left and right signal values of a stereo input signal (typically time frequency block values) respectively, α and β are design parameters (often set to one or chosen such that the power of the downmix signal corresponds to the sum of the powers of the left and right signals) and φ_1 and φ_2 are compensating phase values (which may e.g. be set as $\varphi_1 = \text{IPD}/2$, $\varphi_2 = -\text{IPD}/2$)

The phase compensation for the first parts may for example be represented by:

$$S = \gamma \cdot \exp(j\varphi)L + \delta \cdot \exp(j\varphi)R$$

where γ and δ are design parameters (typically set to one or chosen such that the power of the downmix signal corresponds to the sum of the powers of the left and right signals). The angle φ represents a constant phase compensation.

In accordance with an optional feature of the invention, a phase of the downmix is constant in the first parts and the phase compensation varies dynamically during the second parts.

This may provide improved performance in many embodiments. Specifically, it may in many embodiments allow the data rate to be reduced as the phase compensation for the first parts may be known and therefore need not be dynamically modified. Thus, for the first parts, the upmixing can be based on an assumption of a known constant phase compensation being performed at the encoder (including the encoder using a phase compensation with zero values, i.e. with no phase compensation). The phase of the downmix may be constant in the sense that the generation of the downmix applies a constant phase shift to the left and right signals. The phase shift may specifically be zero (such as e.g. for a simple summation of the right and left signals).

In accordance with an optional feature of the invention, the decoder further comprises: first determining means arranged to determine an overall phase offset for the downmix as a predetermined function of the set of stereo upmix parameters; and wherein the upmixing means is arranged to upmix the downmix based on the overall phase offset.

This may provide improved performance, and in particular improved perceived audio quality, in many scenarios.

In accordance with an optional feature of the invention, the predetermined function is arranged to estimate the overall phase offset of at least one of the multiple channels relative to a constant phase of the downmix during the first parts.

This may provide improved performance, and in particular improved perceived audio quality, in many scenarios.

In accordance with an optional feature of the invention, the upmixing means is arranged to estimate a difference signal for a difference between channels of the multichannel signal based on the downmix scaled with a prediction coefficient derived from the set of upmix parameters, and to generate the multichannel signal based on a sum and a difference of the downmix and said difference signal; and wherein the modifying means is arranged to modify the prediction coefficient for the second parts in response to the phase correction parameter.

This may provide improved performance, and in particular improved perceived audio quality, in many scenarios. In many embodiments, it may provide a reduced complexity, reduced data rate and/or improved audio quality.

In accordance with an optional feature of the invention, the upmixing means is arranged to enhance the difference signal by adding a scaled decorrelated mono downmix.

This may provide improved performance in many scenarios.

In accordance with an optional feature of the invention, the phase correction parameter is indicative of a correction for an interchannel phase difference of the set of upmix parameters.

This may provide improved performance in many scenarios.

In accordance with an optional feature of the invention, the second parts correspond to parts of the parametric signal associated with signals of the multiple channels meeting an out of phase criterion.

This may allow improved performance. In particular, the inventors have realized that using phase information in upmixing is particularly sensitive to out of phase signals and that this can be addressed while maintaining a relatively low data rate by using phase correction parameters for (only) the parts of the signal associated with out of phase characteristics.

The parts of the parametric signal that are associated with the channels meeting the out of phase criterion may correspond to a time interval and/or frequency interval around a time-frequency for which the phase criterion is met. Specifically, the phase correction parameter may gradually change around such a detection to ensure a smooth transition from the first parts (i.e. from parts of the downmix for which no phase correction parameter is provided).

In accordance with an optional feature of the invention, the out of phase criterion comprising a requirement that a phase difference between signals of the multiple channels fall within the intervals of $[\pi-a;\pi]$ and $[-\pi;-\pi+b]$ where a and b are each less or equal to $\pi/8$. This may provide a particularly advantageous criterion for many scenarios and embodiments.

In accordance with an optional feature of the invention, the parametrically encoded signal comprises a phase correction parameter presence indication indicative of the second parts.

This may allow improved performance and may in particular allow a reduced data rate in many scenarios.

In accordance with an optional feature of the invention, the phase correction parameter presence indication comprises a common indication for all time frequency blocks of each encoding frame of the parametrically encoded signal.

This may allow improved performance and may in particular allow a reduced data rate while still providing a suitably high audio quality.

In accordance with an optional feature of the invention, the phase correction parameter presence indication comprises individual presence indications for a plurality of sets of time frequency blocks of the down-mix.

This may allow improved performance and may in particular allow an improved data rate to audio quality ratio. In particular, it may allow a more accurate and flexible use of phase correction parameter values while maintaining a relatively low data rate. The individual presence indications may cover all time frequency blocks of the downmix signal.

In accordance with an optional feature of the invention, the phase correction parameter comprises individual parametric values for a plurality of sets of time frequency blocks of the second parts.

This may allow improved performance and may in particular allow an improved data rate to audio quality ratio. In particular, it may allow a more accurate and flexible use of phase correction parameter values while maintaining a relatively low data rate. The individual parametric values may cover all time frequency blocks of the downmix signal.

According to an aspect of the invention there is provided an encoder for encoding a multichannel signal comprising: downmix means for generating an encoded downmix of multiple channels of the multichannel signal; parameter means for generating a set of upmix parameters relating the downmix to the multiple channels; means for determining at least one section of the downmix in which a deviation between a value of phase parameter derived from the set of upmix parameters and a target value for the phase parameter meets a criterion;

means for generating a phase correction parameter for a part of the encoded downmix associated with the at least one section; and means for generating an encoded signal comprising the encoded downmix, the set of upmix parameters and the phase correction parameter for the part of the encoded downmix.

This may provide improved performance and/or facilitated implementation in many embodiments. The means for determining at least one section may specifically determine the at least one section as a section for which two signals of the multiple channels meet an out of phase criterion.

It will be appreciated that most of the described advantages, comments and features apply to the encoder as well as the decoder as appropriate.

According to an aspect of the invention there is provided a method of generating a multi-channel signal, the method comprising: receiving a parametrically

encoded signal comprising a downmix of multiple channels, a set of upmix parameters including a phase parameter, and a phase correction parameter, the phase correction parameter being provided for second parts of the downmix but not for first parts of the downmix; generating the multi-channel signal by upmixing the downmix based on the set of upmix parameters; and modifying the upmixing in response to the phase correction parameter for the second parts of the downmix.

According to an aspect of the invention there is provided a method of encoding a multichannel signal comprising: generating an encoded downmix of multiple channels of the multichannel signal; generating a set of upmix parameters relating the downmix to the multiple channels; determining at least one section of the downmix in which a deviation between a value of phase parameter derived from the set of upmix parameters and a target value for the phase parameter meets a criterion; generating a phase correction parameter for a part of the encoded downmix associated with the at least one section; and generating an encoded signal comprising the encoded downmix, the set of upmix parameters and the phase correction parameter for the part of the encoded downmix.

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

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

Fig. 1 is an illustration of an example of a parametric encoding system in accordance with the prior art;

Fig. 2 illustrates an example of a transmission system for communication of a stereo audio signal in accordance with some embodiments of the invention;

Fig. 3 illustrates an example of a relationship between an OPD and IPD phase parameter of a parametric encoding system;

Fig. 4 illustrates an example of a relationship between an OPD-IPD and IPD phase parameter of a parametric encoding system;

Fig. 5 illustrates an example of a parametric encoder in accordance with some embodiments of the invention; and

Fig. 6 illustrates an example of a parametric decoder in accordance with some embodiments of the invention.

DETAILED DESCRIPTION OF SOME EMBODIMENTS OF THE INVENTION

The following description focuses on embodiments of the invention applicable to a parametric stereo encoding scheme. However, it will be appreciated that the invention is not limited to this application but may be applied to other multichannel encoding schemes.

Fig. 2 illustrates a transmission system for communication of a stereo audio signal in accordance with some embodiments of the invention. The transmission system comprises a transmitter 201 which is coupled to a receiver 203 through a network 205 which specifically may be the Internet.

In the specific example, the transmitter 201 is a signal recording device and the receiver 203 is a signal player device but it will be appreciated that in other embodiments a transmitter and receiver may be used in other applications and for other purposes. For example, the transmitter 201 and/or the receiver 203 may be part of a transcoding functionality and may e.g. provide interfacing to other signal sources or destinations.

In the specific example where a signal recording function is supported, the transmitter 201 comprises a digitizer 207 which receives an analog stereo signal that is converted to a digital PCM (Pulse Code Modulated) stereo signal by sampling and analog-to-digital conversion.

The digitizer 207 is coupled to the encoder 209 of Fig. 2 which encodes the PCM signal in accordance with a Parametric Stereo (PS) encoding algorithm. The encoder 209 is coupled to a network transmitter 211 which receives the encoded signal and interfaces to the Internet 205. The network transmitter may transmit the parametrically encoded signal to the receiver 203 through the Internet 205.

The receiver 203 comprises a network receiver 213 which interfaces to the Internet 205 and which is arranged to receive the parametrically encoded signal from the transmitter 201.

The network receiver 213 is coupled to a decoder 215. The decoder 215 receives the parametrically encoded signal and decodes it in accordance with a PS decoding algorithm.

In the specific example where a signal playing function is supported, the receiver 203 further comprises a signal player 217 which receives the decoded audio signal from the decoder 215 and presents this to the user. Specifically, the signal player 217 may comprise a digital-to-analog converter, amplifiers and speakers as required for outputting the decoded audio signal.

In the system, the parametric stereo coding scheme utilizes phase information for the stereo channels. Specifically, a downmix is generated at the encoder 209 together with parametric stereo parameters that can be used by the decoder to upmix the downmixed mono signal. These parameters include at least one phase parameter which specifically may be an IPD. The decoder 215 recreates the original signal by upmixing the downmixed mono signal to a stereo channel using the set of stereo upmix parameters. Thus, the upmixing specifically considers a phase characteristics for at least one of the stereo channels of the downmix.

However, the inventors have realized that improved performance can be achieved by identifying scenarios where such an upmix may result in problems and using a phase correction parameter for corresponding parts of the downmix to overcome or mitigate such problems. In particular, the inventors have realized that including a phase correction parameter for parts of the downmix associated with a detection of the stereo channels to be encoded being out of phase may address a number of problems. Furthermore, by limiting the phase correction parameter to specifically identified parts of the downmix, the data rate for other parts of the downmix may be kept unchanged thereby resulting in a reduced average data rate.

In particular, using phase parameters such as IPD and OPD in the upmix may contribute significantly to the perceived quality in PS based audio codecs and may in particular substantially improve sound source localization. However, the usage of such parameters introduces several challenges and especially for signals that are (nearly) out-of-phase (either in single time/frequency tiles or in more continuous regions of a signal), significant degradation in perceived quality can occur if phase relations are not handled correctly. It is worth noting that such degradations do not result from quantization errors, but tend to be related to phase continuity in downmix and decoder output channels.

There are several reasons why out-of-phase signals will occur in stereo signals, and why it is important to preserve this property. Examples include:

- Stereo microphone techniques will often cause time delays between the recorded signals. For example in classical music recordings, this will occur in many cases especially when recordings are made from a limited set of microphones. The resulting time delays will cause out-of-phase signals in some frequency regions (depending on the inter-channel delay).
- Advanced panning techniques may employ temporal panning instead of the conventional amplitude panning. This technique will also create signals that are out-of-phase in some frequency bands.

- Cross-talk cancellation techniques deliberately introduce negative correlations to widen the perceived sound stage.
- Detuning (the effect of creating small frequency differences in the two channels, as is often employed in modern music based on synthesizers) will result in out-of-phase signals.

Thus, it is important for a stereo codec that it can effectively handle out of phase signals. However, stereo signals that are out of phase may cancel each other out and may furthermore result in phase parameters being calculated which vary very significantly for even small variations in the individual stereo signals. For example, small signal variations resulting in a relative phase change of, say 2° , may result in a calculated IPD value changing from $+179^\circ$ to -179° . Also, the relationship between IPD and OPD values may change very substantially for small changes of signals that are substantially out of phase. This may introduce phase discontinuities (in time or frequency) which may result in very noticeable artifacts.

Thus may be particularly critical for systems that use a decoder based calculation of OPD values from other PS parameters, such as e.g. proposed in the article Jimmy Lapierre and Roch Lefebvre, "On Improving Parametric Stereo Audio Coding", Presented at the 120th Convention, 2006 May 20–23 Paris, Audio engineering Society, France, Preprint 6804.

In more detail, a PS decoder may use the IID, IC and IPD parameters to determine a stereo signal from a downmix signal. This process is typically performed by upmixing the downmix and a decorrelated signal using a mixing matrix that depends on the IID, IC and IPD values. However, as IPD only describes the relative phase modification between the two output signals and not the phase difference between the downmix and the individual stereo channels, it cannot provide any information of how the phase modification should be distributed across the output stereo channels. The OPD parameter is indicative of the phase offset between the downmix and at least one of the stereo channels and it thus reflects how the phase should be distributed between the channels. The OPD may accordingly be included in the encoded signal by an encoder. However, although the OPD can be transmitted from encoder to decoder using a relatively limited bit budget, this approach does increase the overall data rate for the signal. Therefore, the mentioned document proposes that an OPD estimation is performed at the decoder side such that the OPD value is not included in the encoded signal but is instead calculated by the decoder from the other parameter values.

When the OPD parameter is available, the output stereo signals L' , R' can be constructed from the down-mix S and the decorrelated signal D by:

$$L' = \exp(j(OPD))(m_{11}S + m_{12}D),$$

$$R' = \exp(j(OPD - IPD))(m_{21}S + m_{22}D),$$

where the matrix elements m_{ij} are only dependent on IID and IC (as will be well known to the skilled person and as described in e.g. Breebaart, J., van de Par, S., Kohlrausch, A., and Schuijers, E. (2005). Parametric coding of stereo audio. *Eurasip J. Applied Signal Proc.*, issue 9: special issue on anthropomorphic processing of audio and speech, 1305-1322).

The OPD may e.g. be calculated from (ref. e.g. Jimmy Lapierre and Roch Lefebvre, “On Improving Parametric Stereo Audio Coding”, Presented at the 120th Convention, 2006 May 20–23 Paris, Audio engineering Society, France, Preprint 6804):

$$OPD = \angle(10^{IID/20} + IC \exp(jIPD))$$

However, such approaches result in a discontinuity for the OPD value as a function of the IPD value. For example, for an IID = 0 dB, and IC = +1, Fig. 3 illustrates the relationship between the OPD and the IPD and Fig. 4 illustrates the relationship between OPD-IPD (the phase modification applied to the right output channel) and the IPD parameter. As is clearly demonstrated, a discontinuity occurs for IPD = π corresponding to the input stereo signals being out of phase. Furthermore, the phase discontinuity amounts to π radians resulting in the effect that small variations of IPD around π results in the stereo output signals generated by the upmix completely changing signs. Thus, the signs of the output signals (e.g. in each time frequency block for which the IPD is close to π) may flip between signs.

These variations in IPD (and thus the sign inversions) may occur in both the temporal and frequency dimensions. For example, discontinuities can occur in time (e.g. if IPD changes in time from just below to just above π). However, as the processing of segments include time averaging (e.g. as part of FFT or QMF windowing) this may lead to cancellation of the output signal. In particular, such an approach typically results in audible artifacts perceived as ‘clicks’ or ‘warbling’ sounds. Similarly, the IPD changes (and thus sign inversions) may occur in the frequency domain between one subband and the next (e.g. if

IPD in one band is just below π and just above π in a neighboring band). This may similarly result in noticeable artifacts.

In the system of Fig. 2, such the encoder 209 divides the signal into first and second parts where the first parts correspond to time and/or frequency sections that are not associated with an out of phase scenario whereas the second parts correspond to time and/or frequency sections that are associated with an out of phase scenario. The encoder then proceeds to include a phase correction parameter for the second parts but not for the first parts. This phase correction parameter is specifically calculated to compensate for phase discontinuities and is included in the encoded signal. The decoder 215 then proceeds to use the phase correction parameter when upmixing the downmix of the second parts and specifically uses this to compensate for phase discontinuities. The inclusion of the phase correction parameter in the encoded signal may increase the data rate relative to an approach where e.g. OPD values are always calculated at the decoder. However, the approach may provide substantially improved audio quality and support for out of phase signals. Furthermore, as out of phase signals typically only occur for a small proportion of the frequencies and/or time, the increase in the data rate may be kept relatively small.

Fig. 5 illustrates elements of the encoder 209 in more detail. The encoder 209 comprises a downmix unit 501 which receives an input stereo signal. The downmix unit 501 proceeds to generate a mono downmix for the two stereo channels of the input stereo signal. The downmix may e.g. be generated as a simple summation of the signals of the two stereo channels, i.e. as:

$$S = L + R$$

where L and R are the signal values of left and right input stereo channels. The downmix unit 501 then proceeds to encode the downmix signal to generate an encoded downmix. It will be appreciated that any suitable encoding may be used. Typically, the encoding will include the generation of a number of time frequency tiles representing the downmix signal in the frequency domain for each of a plurality of time segments as will be well known to the skilled person.

The downmix unit 501 is coupled to a parameter generation unit 503 which proceeds to generate parametric stereo parameters based on the downmix and the original input signals. Specifically, the parameter unit 503 may generate a set of stereo upmix parameters comprising IID and IPD values, as well as in some embodiments IC values. It will

be appreciated that any suitable approach for generating the stereo upmix parameters may be used without detracting from the invention. Typically, the parameter unit 503 generates a set of IID and IPD values for each of a plurality of time frequency blocks. Typically, each time frequency block corresponds to one time segment or (encoding) frame and a given frequency band (such as an ERB band) which is wider than the frequency span of the time frequency tiles used for the encoding of the downmix signal.

It will be appreciated that various approaches and algorithms for calculating such PS parameters will be well known to the skilled person and consequently it will not be described in further detail.

The parameter unit 503 is fed to an out-of-phase detection unit 505 which is arranged to detect at least one section of the downmix in which the stereo channels meet an out of phase criterion. Each section for which an out-of-phase condition is detected may correspond to one or more time frequency blocks.

As an example, the criterion may be a requirement that the IPD value falls within one of the intervals $[\pi-a;\pi]$ and $[-\pi;-\pi+b]$ where a and b are suitable design parameter (which typically may be identical). Particularly advantageous performance trade-offs have been found when a and b are each less or equal to $\pi/8$. Thus, the out-of-phase detection unit 505 may identify all time frequency blocks for which the IPD value falls within these intervals and thus all the time frequency blocks for which the input stereo channels are considered to be (sufficiently) out of phase. It will be appreciated that in other embodiments, other out of phase criteria may be used, such as e.g. a criterion including a requirement that an interchannel correlation is below a given value.

The out-of-phase detection unit 505 is coupled to a correction unit 507 which is further coupled to the parameter unit 503. The correction unit 507 receives an indication of all the sections (e.g. all the time frequency blocks) that are considered to be out of phase. It then proceeds to select which parts of the downmix signal a phase correction parameter should be included. Generally, the parts are selected to include the time frequency blocks identified by the out-of-phase detection unit 505 but are typically not restricted to these. Indeed, typically surrounding time frequency blocks are also included in order to ensure a smooth and gradual transition from the parts for which there is no phase correction parameter to the time frequency blocks which are detected as being sufficiently out of phase.

The correction unit 507 then proceeds to calculate a phase correction parameter. Specifically, phase correction parameter values for all the time frequency blocks identified by the correction unit 507 are calculated. Specifically, the phase correction

parameter may be a value which can offset a phase parameter value calculated by the decoder such that phase discontinuities are avoided. For example, the phase correction value may specify that e.g. π should be added or subtracted to the phase parameter value calculated at the decoder. Thus, the phase correction parameter may be a relative parameter which is relative to a phase parameter calculated by the decoder or to parameters or phase values for other time frequency blocks. In other, embodiments or scenarios, the phase correction value may be a replacement value that directly provides the parameter value to be used by the decoder 215.

The downmix unit 501 and the correction unit 507 are coupled to a multiplexer 509 which combines the encoded downmix, the set of upmix stereo parameters and the phase correction parameter into a single encoded signal.

Thus, the encoder 209 generates a parametrically encoded stereo signal comprising a mono downmix as well as parametric stereo upmix parameters that can be used to upmix this downmix signal. These stereo parameters include a phase parameter and particularly include an IPD parameter. In addition, the encoded signal comprises a phase correction parameter, but only for parts of the encoded signal for which the input stereo signal are considered to be associated with input stereo channels being sufficiently out of phase. Thus, the signal is divided into first parts wherein no phase correction parameter is provided and second parts in which a phase correction parameter is provided. The second parts represent parts of the signal for which the encoder estimates that the phase based upmixing is likely to introduce mistakes or artifacts unless the phase correction parameter is used.

Fig. 6 illustrates elements of the decoder 215 in more detail.

The decoder 215 comprises a demultiplexer 601 which receives the parametrically encoded signal from the encoder. It furthermore comprises an upmix unit 603 which is arranged to generate a stereo signal from the downmix. The upmixing may specifically use the downmix signal as well as a second signal derived therefrom (e.g. a decorrelated signal derived from the downmix). For example, a matrix multiplication may be applied to each time frequency tile of the downmix and the decorrelated signal.

The decoder 215 further comprises a parameter derivation unit 605 which is coupled to the demultiplexer 601 and the upmix unit 603. The parameter derivation unit 605 generates suitable upmix values based on the set of stereo upmix parameters. Specifically, for the first parts of the downmix (for which no phase correction parameter is provided), the parameter derivation unit 605 calculates suitable upmix values based on a nominal or default

algorithm. This default algorithm may e.g. use a predetermined function for calculating phase values based on the provided set of parameters.

The decoder 215 also comprises a correction unit 607 which is coupled to the demultiplexer 601 and the parameter derivation unit 605 and which is arranged to modify the upmixing operation based on the phase correction parameter. Thus, for the second parts of the downmix, the correction unit 607 receives the phase correction parameter and controls the parameter derivation unit 605 to modify the generated upmix values to reflect the correction indicated by the phase correction parameter.

Thus, the decoder 215 generates a stereo signal based on a default phase based upmixing during (most) parts of the received signal while at the same time modifying these approaches for parts of the signal wherein the stereo channels are considered sufficiently out of phase to be likely to result in artifacts and degradation if the default upmixing is used.

As a specific example, the decoder 215 may be arranged to receive IPD, IID and ICC parameters in the encoded signal. It may further be arranged to calculate an OPD parameter from these parameters using a predetermined function, such as:

$$OPD = \angle(10^{IID/20} + IC \exp(jIPD))$$

The decoder may then proceed to generate the upmixed values as:

$$l = \exp(j(OPD))(m_{11}s + m_{12}s_d),$$

$$r = \exp(j(OPD - IPD))(m_{21}s + m_{22}s_d),$$

or

$$\begin{bmatrix} l \\ r \end{bmatrix} = \begin{bmatrix} c_1 \cdot \cos(\alpha + \beta) \cdot e^{j\text{opd}} & c_1 \cdot \sin(\alpha + \beta) \cdot e^{j\text{opd}} \\ c_2 \cdot \cos(-\alpha + \beta) \cdot e^{j(\text{opd}-\text{ipd})} & c_2 \cdot \sin(-\alpha + \beta) \cdot e^{j(\text{opd}-\text{ipd})} \end{bmatrix} \begin{bmatrix} s \\ s_d \end{bmatrix}$$

where c_1 , c_2 , α and β are a function of the IID and ICC parameters:

$$c_1 = \sqrt{\frac{iid}{1+iid}},$$

$$c_2 = \sqrt{\frac{1}{1+iid}},$$

$$\alpha = \frac{1}{2} \arccos(icc),$$

$$\beta = \arctan\left(\tan(\alpha) \cdot \frac{c_2 - c_1}{c_2 + c_1}\right).$$

and s_d is a decorrelated signal generated from the downmix s e.g. by allpass filtering (as will be known to the skilled person).

Thus, the upmixing may use a conventional upmix operation based on the phase parameters IPD and OPD. However, rather than using communicated OPD values, the decoder itself estimates/calculates suitable OPD values from the received PS parameters.

The encoder may be arranged to perform the same calculation thereby deriving the OPD value that is calculated by the decoder. However, for the second parts typically associated with sections of the signal for which the stereo signals are sufficiently out of phase to possibly result in phase discontinuities, the encoder proceeds to generate an offset phase correction. For example, for a time frequency block where the OPD varies by close to π compared to a neighbor time frequency block (where the neighbor may be in time and/or frequency), the encoder may include a phase correction parameter that indicates that a value of π should be added (or subtracted) from the calculated OPD value. Alternatively, the phase correction parameter may directly provide an OPD value that should be used by the decoder instead of the decoder calculated value. Thus, the encoder includes overall phase offset correction parameter values that may remove any discontinuities which will result from applying the predetermined function to calculate OPD.

Accordingly, the decoder 215 operates in two different modes depending on whether a phase correction parameter is provided or not. For the first parts of the downmix, an overall phase offset in the form of an OPD is generated based on a predetermined function, and for the second parts the overall phase offset is determined from the phase correction parameter. The overall phase offset is then used as the OPD value of the upmix operation. The overall phase offset is indicative of a phase difference between the downmix and at least one of the stereo channels and thus provides information of how the downmix phase should be distributed across the different output signal. Accordingly, an improved perceived audio quality and especially sound source position perception is achieved.

As mentioned, the OPD used by the upmixing may for the second parts be determined based on the overall phase correction value provided in the encoded signal, e.g. when this comprises a replacement value for the decoder generated value. In other scenarios,

the OPD may in addition to the overall phase correction parameter also be determined based on the set of stereo upmix parameters. Specifically, the IID, IC and IPD values may be used to generate an OPD estimate using the default function and the resulting value may be offset by the value given by the overall phase offset parameter.

Thus, in this example, the encoder 209 detects whether or not the OPD that can be estimated by the decoder 215 provides a reliable estimate which does not invoke any phase modifications. This information is then signaled in the PS bit-stream e.g. by including a phase correction parameter value for all time frequency blocks that are considered to not be estimated reliably. In other examples, a correction value may be provided for each segment or for groups of time frequency blocks. The correction values can be transmitted as absolute values, but also differentially with respect to the decoder estimate, differentially over frequency- and/or time (e.g. the correction value is transmitted as an offset to a (time and/or frequency) neighboring time frequency block).

At the decoder side (depending on the signaling information), the OPD information is estimated from the other PS parameters or is decoded from the bit-stream. The latter case may still employ the estimated data, depending on the coding scheme as outlined above, e.g. by transmitting the difference between the estimated OPD and the OPD derived in the decoder. Furthermore, although the parameters for the upmixing are changed for the second parts based on the phase correction parameter, the actual operations and functions may be the same as in traditional parametric stereo upmixing which uses IPD and OPD values.

In some embodiments, the encoder 209 is arranged to introduce a phase compensation when generating the downmix for the second parts relative to any phase compensation that is applied for the first parts. Specifically, no phase compensation – or a constant phase compensation - may be used for the first parts, whereas the phase is dynamically adjusted for the second parts (or a different constant phase compensation is used). Such a phase compensation may specifically ensure that the component from the two stereo channels in the downmix are not out of phase with each other, and thus may prevent or reduce the two stereo signals canceling each other. The encoder may then generate the phase correction parameter to be indicative of the phase compensation that is performed for the second parts. However, as there is no phase compensation (or a constant known phase compensation) applied for the first parts, the decoder does not require any phase compensation information for these parts.

Such an approach may provide advantageous operation in many embodiments and may in particular provide improved audio quality without substantially increasing the data rate.

In more detail, when signals are down-mixed, a preferred property of the down-mix signal is that it is energy preserving on a frequency scale that roughly corresponds to the human auditory system, such as the ERB scale. Said differently, in each time/frequency tile, the energy of the downmix should preferably be equal to the sum of the energies of the individual input signals. This is advantageous to minimize perceptual influences (e.g. coloring) of the down-mix process. For the first parts of the signal, using a simple downmix, such as simply summing the two stereo signals (and e.g. applying a simple scaling), provides a typically sufficient degree of energy preservation. Accordingly, a constant (such as 0°) phase compensation can be used thereby obviating the need for substantially increasing the data rate. However, for the second parts wherein the signals are (close to being) out of phase the signals fully or partially cancel each other out resulting in a very high variation in the energy. Furthermore, this cannot typically be compensated by simply scaling the downmix signal since the post-scale factor that would need to be employed would be extremely large, resulting in audible quantization artifacts. On the other hand, if no post-scaling is employed, the signal will be (almost) zero.

In some embodiments, the encoder may accordingly apply phase compensation for these second parts. For example, the downmix unit may calculate the downmix as:

$$S = \alpha \cdot \exp(j\phi_1)L + \beta \cdot \exp(j\phi_2)R$$

where S is the phase compensated downmix, L and R are the left and right signal values (typically time frequency block values) respectively α and β are design parameters (often set to one or such that the power of the downmix corresponds to the sum of the powers of the left and right signals) and ϕ_1 and ϕ_2 are compensating phase values.

Using unity scaling of the signals and setting the phase values to $\pm IPD/2$ yields:

$$S = \exp(j \cdot IPD / 2)L + \exp(-j \cdot IPD / 2)R$$

The phase compensation can be used to ensure that the energy of the downmix is maintained sufficiently constant. Furthermore, the phase values may be selected such that phase discontinuities (e.g. across time or frequency) are avoided. Specifically, the phase compensation may be selected such that the OPD of the downmix relative to the phase compensated signal will always correspond to that calculated at the decoder 215. However, in order to generate the original stereo signals instead of the phase compensated stereo signals, the decoder 215 must reverse the operation of the phase compensation. Accordingly, the phase correction parameter may indicate the phase compensation that has been applied thereby allowing the decoder 215 to reverse this phase compensation. Furthermore, as the correction is only applied to the out of phase parts of the encoded signal, the data rate increase is typically relatively low.

As a specific example, the encoder may be arranged to use the following downmix for the first parts:

$$S = L + R$$

and may generate IID, IC and IPD values for this downmix. No phase correction parameter values are included for these parts. The decoder may then for these first parts estimate an OPD value as previously described and perform the described OPD and IPD based upmixing resulting in the desired output stereo signals.

However, for the second parts, the encoder 209 may use the following downmix:

$$S = \exp(j\phi_1)L + \exp(j\phi_2)R$$

where the phase values are selected to gradually and smoothly change with time and frequency. Furthermore, from a conceptual point of view, the encoder can determine the IID, IC and IPD values that relate the downmix to the phase compensated stereo signals given by:

$$L' = \exp(j\phi_1)L \quad R' = \exp(j\phi_2)R$$

Furthermore, the phase values are selected such that the OPD values estimated from the IID, IC and IPD values will always be appropriate and will not have any phase

discontinuities (basically the phase values ϕ_1 and ϕ_2 are selected such that the phase compensated stereo signals L' and R' are not close to being out of phase with each other).

Accordingly, the decoder 215 can generate the phase compensated stereo signals L' and R' based on reliable parameter values including reliable IPD and OPD values. However, in order to recreate the original stereo signal R, L it is necessary for the decoder 215 to reverse the phase compensation. Accordingly, the encoder may include a phase correction parameter which specifies the phase compensation that has been performed for the second parts. Specifically, the applied phase values ϕ_1 and ϕ_2 may be included for each frequency time block of the second parts.

The decoder 215 may then modify the upmixing to include the operation:

$$\begin{bmatrix} l \\ r \end{bmatrix} = \begin{bmatrix} e^{-j\phi_1} & 0 \\ 0 & e^{-j\phi_2} \end{bmatrix} \begin{bmatrix} l' \\ r' \end{bmatrix}$$

where l' and r' correspond to the phase compensated stereo signals generated from the stereo upmix parameters. Thus the correction unit 607 modifies the upmixing based on the phase correction parameter for the second parts. Typically, the matrix operation to reverse the phase compensation will be included in the upmix matrix. Thus, the same operations may be performed but with modified values.

In other embodiments, the stereo signals may be generated without an explicit generation of OPD values. For example, stereo signals can be generated based on the downmix and the generation of a difference signal generated from the downmix.

More specifically, in accordance with such an embodiment a difference signal comprising a difference between the left stereo signal and the right stereo signal may be predicted based on the mono downmix scaled with a prediction coefficient, where the prediction coefficient is derived from the spatial parameters. The upmixing unit may then generate the left output signal and the right output signal based on a sum and a difference of the mono downmix signal and said difference signal.

The left right output signals can be reconstructed as follows:

$$\begin{aligned} l &= s + d, \\ r &= s - d, \end{aligned}$$

where s is the mono downmix, and d is the predicted difference signal. This is under the assumption that the encoder sum signal is calculated as:

$$s = \frac{l + r}{2}.$$

In practice gain normalization is often applied when constructing the left signal and the right signal:

$$l = \frac{1}{2c} \cdot (s + d),$$

$$r = \frac{1}{2c} \cdot (s - d),$$

where c is a gain normalization constant and is a function of the stereo parameters. Gain normalization ensures that a power of the mono downmix signal is (approximately) equal to a sum of powers of the left signal and the right signal. In this case the encoder sum signal may be calculated as:

$$s = c \cdot (l + r).$$

The spatial parameters are determined in the encoder beforehand and transmitted to the decoder. The spatial parameters are determined on a frame-by-frame basis for each time/frequency tile as:

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

$$icc = \frac{|\langle l, r \rangle|}{\sqrt{\langle l, l \rangle \cdot \langle r, r \rangle}},$$

$$ipd = \angle \langle l, r \rangle,$$

where iid is an interchannel intensity difference, icc is an interchannel coherence, ipd is an interchannel phase difference, and $\langle l, l \rangle$ and $\langle r, r \rangle$ are the left and right signal powers

respectively and $\langle l, r \rangle$ represents the non-normalized complex-valued covariance coefficient between the left and right signals.

For a typical complex-valued frequency domain such as the DFT (FFT), these powers are measured as:

$$\langle l, l \rangle = \sum_{k \in k_{tile}} l[k] \cdot l^*[k],$$

$$\langle r, r \rangle = \sum_{k \in k_{tile}} r[k] \cdot r^*[k],$$

$$\langle l, r \rangle = \sum_{k \in k_{tile}} l[k] \cdot r^*[k],$$

where k_{tile} represents the DFT bins corresponding to a parameter band. It is to be noted that also other complex domain representation could be used, such as e.g. a complex exponentially modulated QMF bank as described in P. Ekstrand, “Bandwidth extension of audio signals by spectral band replication”, in Proc. 1st IEEE Benelux Workshop on Model based Processing and Coding of Audio (MPCA-2002), Leuven, Belgium, Nov. 2002, pp. 73 – 79.

For low frequencies up to 1.5-2 kHz the above equations hold. However, for higher frequencies the ipd parameters are not relevant for perception and therefore they are set to a zero value resulting in:

$$\begin{aligned} iid &= \frac{\langle l, l \rangle}{\langle r, r \rangle}, \\ icc &= \frac{\Re\{\langle l, r \rangle\}}{\sqrt{\langle l, l \rangle \cdot \langle r, r \rangle}}, \\ ipd &= 0. \end{aligned}$$

Alternatively, since at higher frequencies, rather the broadband envelope than the phase differences are important for perception, the icc is calculated as:

$$icc = \frac{|\langle l, r \rangle|}{\sqrt{\langle l, l \rangle \cdot \langle r, r \rangle}}.$$

The gain normalization constant c is expressed as:

$$c = \sqrt{\frac{iid + 1}{iid + 1 + 2 \cdot icc \cdot \cos(ipd) \cdot \sqrt{iid}}}.$$

Since c may approach infinity due to left and right signals being out of phase, the value of the gain normalization constant c is typically limited as:

$$c = \min \left(\sqrt{\frac{iid + 1}{iid + 1 + 2 \cdot icc \cdot \cos(ipd) \cdot \sqrt{iid}}}, c_{\max} \right),$$

with $c_{\max}$ being the maximum amplification factor, e.g. $c_{\max} = 2$.

The prediction coefficient used to calculate the difference signal is based on estimating the difference signal from the mono downmix using waveform matching. Said waveform matching comprises e.g. a least-squares match of the mono downmix signal onto the difference signal, resulting in the difference signal provided as:

$$d = \alpha \cdot s,$$

where s is the mono downmix and α is the prediction coefficient.

Beside the least-squares matching a waveform matching using a different norm from L_2 -norm can be used. Alternatively, the p-norm error $\|d - \alpha \cdot s\|^p$ could be e.g. perceptually weighted. However, the least-squares matching is advantageous as it results in relatively simple calculations for deriving the prediction coefficient from the transmitted spatial image parameters.

It is well known that the least-squares prediction solution for the prediction coefficient α is given by:

$$\alpha = \frac{\langle s, d \rangle^*}{\langle s, s \rangle},$$

where $\langle s, d \rangle^*$ represents the complex conjugate of the cross correlation of the mono downmix and the difference signal and $\langle s, s \rangle$ represents the power of the mono downmix signal.

The prediction coefficient may specifically be given as a function of the stereo parameters:

$$\alpha = \frac{iid - 1 - j \cdot 2 \cdot \sin(ipd) \cdot icc \cdot \sqrt{iid}}{iid + 1 + 2 \cdot \cos(ipd) \cdot icc \cdot \sqrt{iid}}.$$

The upmixing unit may specifically enhance the difference signal by adding a scaled decorrelated mono downmix signal. The mono downmix is decorrelated, e.g. using all-pass filters to generate a decorrelated mono downmix. A first part of the difference signal is calculated by scaling the mono downmix with the prediction coefficient. Additionally the decorrelated mono downmix is also scaled by a scale factor. A resulting second part of the difference signal is consequently added to the first part of the difference signal resulting in an enhanced difference signal. The mono downmix and the enhanced difference signal are then used to calculate the left signal and the right signal.

In general it is not possible to accurately predict the difference signal from the mono downmix signal by just scaling with the prediction coefficient. This gives rise to a residual signal $d_{res} = d - \alpha \cdot s$. This residual signal has no correlation with the downmix signal as otherwise it would have been taken into account by means of the prediction coefficient. In many cases the residual signal comprises a reverberant sound field of a recording. The residual signal is effectively synthesized using the decorrelated mono downmix derived from the mono downmix.

The decorrelated mono downmix can be obtained by means of filtering the mono downmix. This filtering generates a signal with a similar spectral and temporal envelope as the mono downmix, but with a correlation substantially close to zero such that it corresponds to a synthetic variant of the residual component derived in the encoder. This effect is achieved by means of e.g. allpass filtering, delays, lattice reverberation filters, feedback delay networks or a combination thereof.

The scaling factor applied to the decorrelated mono downmix may be compensated for a prediction energy loss. The scaling factor applied to the decorrelated mono downmix ensures that the overall signal power of the output left and right signals matches the signal power of the left and right signal power at the encoder side, respectively. As such the scaling factor, henceforth referred to as β is interpreted as a prediction energy loss compensation factor. The difference signal d is then expressed as:

$$d = \alpha \cdot s + \beta \cdot s_d,$$

where s_d is the decorrelated mono downmix.

It can be shown that said scaling factor β can be expressed as:

$$\beta = \sqrt{\frac{\langle d, d \rangle}{\langle s, s \rangle} - |\alpha|^2}$$

in terms of signal powers corresponding to the difference signal d and the mono downmix s .

The scaling factor β applied to the decorrelated mono downmix can be derived as a function of the spatial parameters:

$$\beta = \sqrt{\frac{iid + 1 - 2 \cdot \cos(ipd) \cdot icc \cdot \sqrt{iid}}{iid + 1 + 2 \cdot \cos(ipd) \cdot icc \cdot \sqrt{iid}} - |\alpha|^2} = \frac{2 \cdot \sqrt{iid} \cdot (1 - icc^2)}{iid + 1 + 2 \cdot \cos(ipd) \cdot icc \cdot \sqrt{iid}}.$$

In case, no downmix normalization was applied in the encoder, i.e., the downmix signal was calculated as $s = \frac{1}{2}(l + r)$, the left and right output signals may then be expressed as:

$$\begin{bmatrix} l \\ r \end{bmatrix} = \begin{bmatrix} 1 + \alpha & \beta \\ 1 - \alpha & -\beta \end{bmatrix} \begin{bmatrix} s \\ s_d \end{bmatrix}.$$

Thus, the difference signal $d = \alpha \cdot s + \beta \cdot s_d$ is respectively added to the downmix s for the left channel and subtracted from the downmix s for the right channel.

In case a downmix normalization was applied, (i.e., the downmix signal was calculated as $s = c(l + r)$), the left and right output signals can be expressed as:

$$\begin{bmatrix} l \\ r \end{bmatrix} = \frac{1}{2c} \begin{bmatrix} 1+\alpha & \beta \\ 1-\alpha & -\beta \end{bmatrix} \begin{bmatrix} s \\ s_d \end{bmatrix}.$$

In the system of Fig. 2, the approach described above may be used to generate the output stereo signal for the first parts of the downmix where no phase correction parameter is provided (and the downmix is generated as $s = l + r$ or $s = c(l + r)$).

However, for the second parts, the downmix includes a phase compensation and the downmix may accordingly be given as

$$s = c'(\exp(j\varphi_1)L + \exp(j\varphi_2)R)$$

In this case it is necessary to consider the phase compensation in the decoder 215. Accordingly, the phase correction parameter provides the information of the phase compensation for these parts.

The left and right output signals may then be calculated as:

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

where

$$\begin{aligned} \alpha' &= \frac{\langle s, d \rangle^*}{\langle s, s \rangle} \\ &= \frac{iid - 1 + 2j \sin(ipd + \varphi_1 - \varphi_2) \cdot icc \cdot \sqrt{iid}}{iid + 1 + 2 \cos(ipd + \varphi_1 - \varphi_2) \cdot icc \cdot \sqrt{iid}} \\ \beta' &= \sqrt{\frac{\langle d, d \rangle}{\langle s, s \rangle} - |\alpha'|^2} \end{aligned}$$

with

$$\frac{\langle d, d \rangle}{\langle s, s \rangle} = \frac{iid + 1 - 2 \cos(ipd + \varphi_1 - \varphi_2) \cdot icc \cdot \sqrt{iid}}{iid + 1 + 2 \cos(ipd + \varphi_1 - \varphi_2) \cdot icc \cdot \sqrt{iid}}$$

For the downmix gain adjustment parameter c' the following assumption is made:

$$\begin{aligned}\langle l, l \rangle + \langle r, r \rangle &= \langle s, s \rangle \\ &= c'^2 \left(\langle l, l \rangle + \langle r, r \rangle + 2\Re \left\{ e^{j(\varphi_1 - \varphi_2)} \langle l, r \rangle \right\} \right)\end{aligned}$$

which leads to:

$$c' = \sqrt{\frac{iid + 1}{iicd + 1 + 2 \cos(ipd + \varphi_1 - \varphi_2) \cdot icc \cdot \sqrt{iid}}}$$

Thus, in this case the phase correction parameter may provide the phase compensation difference $\varphi_1 - \varphi_2$ for the time frequency blocks of the second parts of the downmix. Thus, rather than providing an overall phase offset, the phase correction parameter in this example reflects an interchannel phase difference correction value. The modifying unit 607 thus changes the generation of the upmixing coefficients to be based on the equations which include the correction value $\varphi_1 - \varphi_2$ for the time frequency blocks of the second parts but not for the first parts.

In some embodiments, the parametrically encoded signal may also comprise a phase correction parameter presence indication which is indicative of the second parts. Thus, this presence indication, which may be distinct from the phase correction parameter itself, may provide an efficient and low data rate indication of which parts of the encoded signal comprises phase correction parameter values that should be included when decoding the encoded stereo signal.

The presence indication may specifically be single bit values indicating whether a phase correction parameter should purely be estimated by the encoder or should be replaced (or compensated) by a phase correction parameter included in the encoded stereo signal.

In some embodiments, a common presence indication may be provided for each segment of the parametrically encoded signal. Thus, the encoded signal may be segmented in typically the time domain when being encoded. The presence indication may simply indicate whether there are any phase correction parameter values for the current segment. For example, the presence indication can be a single bit denoting that for the current

frame all time frequency blocks can be estimated reliably by the decoder. This may provide a very low data rate overhead (possibly a single bit per segment) and may reduce the complexity and/or resource usage of the decoder.

In other embodiments, a more detailed presence indication may be used. Specifically, the presence indication may comprise individual presentation indications for a plurality of sets of time frequency blocks of the down-mix. Specifically, each set may correspond to one time frequency block for which individual PS parameters are provided. Further the sets may cover all time frequency blocks of the signal. Thus, specifically a single presence indication bit may be included for each parameter time frequency block indicating whether e.g. the OPD for the block can be purely estimated by the decoder 215 or whether it must take into account a phase correction parameter provided for the block. It will be appreciated that in many embodiments, the phase correction parameter may indeed be provided for each time frequency block that belongs to the second parts of the downmix.

The previous description focuses on an example wherein the second parts are determined to correspond to parts associated with an out-of-phase condition for the two input channels. However, it will be appreciated that the invention is not limited to this specific example. Rather, the out-of-phase detection unit 505 may be replaced by other detection units arranged to detect one or more sections of the downmix for which the deviation between a value of phase parameter derived from the upmix parameters and a target value for the phase parameter meets a criterion. The target value may specifically be a value for the phase parameter which is calculated directly from the input signals or which has been compensated to reduce or remove undesired characteristics. For example, the target value for an OPD parameter may be calculated from the original input signals and compensated such that any phase discontinuities are removed. The value calculated from the set of upmix parameters may be calculated using a predetermined function which is also used by the decoder 215. Thus the detection unit 505 may be arranged to compare the value that will be calculated by the decoder 215 to the value it should preferably be. Any sections where the deviation (e.g. calculated on the basis of a psychoacoustic model) between these exceed a threshold may thus be identified and corrected by the inclusion of the phase correction parameter. Thus, improved audio quality is generated for the specific parts of the signal where the decoder 215 derived parameter value is not sufficiently accurate without affecting the data rate for other parts of the encoded signal.

Also, it will be appreciated that whereas the above description has focused on an application to a parametric stereo coding system, the described approaches and principles

may also be used in multichannel systems. For example, a multichannel system may use a plurality of downmix and upmix blocks (often referred to as One-To-Two, Two-To-Three, Two-To-One and Three-To-Two modules) to encode a multichannel signal. The encoder described above may in such a system be used as a Two-To-One block and the decoder may be used as a One-To-Two block. In other embodiments, the principles may e.g. be used in a coding system wherein three channels are encoded as a mono or stereo downmix with parameters allowing the upmix back to three channels.

It will be appreciated that the above description for clarity has described embodiments of the invention with reference to different functional units and processors. However, it will be apparent that any suitable distribution of functionality between different functional units or processors may be used without detracting from the invention. For example, functionality illustrated to be performed by separate processors or controllers may be performed by the same processor or controllers. Hence, references to specific functional units are only to be seen as references to suitable means for providing the described functionality rather than indicative of a strict logical or physical structure or organization.

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 and processors.

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.

Furthermore, although individually listed, a plurality of means, elements or method steps may be implemented by e.g. a single 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. A decoder for generating a multi-channel signal comprising:
means (601) for receiving a parametrically encoded signal comprising a downmix of multiple channels, a set of upmix parameters including a phase parameter, and a phase correction parameter, the phase correction parameter being provided for second parts of the downmix but not for first parts of the downmix;
upmixing means (603, 605) for generating the multi-channel signal by upmixing the downmix based on the set of upmix parameters;
modifying means (607) for modifying the upmixing in response to the phase correction parameter for the second parts of the downmix.
2. The decoder of claim 1 further comprising:
first determining means arranged to determine an overall phase offset in response to the set of stereo upmix parameters for the first parts of the downmix;
second determining means arranged to determine the overall phase offset in response to the phase correction parameter for the second parts of the downmix; and wherein
the upmixing means (603, 605) is arranged to upmix the downmix based on the overall phase offset.
3. The decoder of claim 2 wherein the overall phase offset is indicative of a phase difference between the downmix and at least one of the multiple channels.
4. The decoder of claim 2 wherein the second determining means is arranged to further determine the overall phase offset in response to the set of upmix parameters.
5. The decoder of claim 1 wherein the downmix is a phase compensated downmix for the second parts relative to the first parts, and the phase correction parameter is indicative of the phase compensation.

6. The decoder of claim 5 wherein a phase of the downmix is constant in the first parts and the phase compensation varies dynamically during the second parts.
7. The decoder of claim 5 further comprising:
first determining means arranged to determine an overall phase offset for the downmix as a predetermined function of the set of stereo upmix parameters; and
wherein
the upmixing means (603, 605) is arranged to upmix the downmix based on the overall phase offset.
8. The decoder of claim 7 wherein the predetermined function is arranged to estimate the overall phase offset of at least one of the multiple channels relative to a constant phase of the downmix during the first parts.
9. The decoder of claim 5 wherein
the upmixing means (603, 605) is arranged to estimate a difference signal for a difference between channels of the multichannel signal based on the downmix scaled with a prediction coefficient derived from the set of upmix parameters, and to generate the multichannel signal based on a sum and a difference of the downmix and said difference signal; and wherein the modifying means (607) is arranged to modify the prediction coefficient for the second parts in response to the phase correction parameter.
10. The decoder claimed in claim 9 wherein the upmixing means (603, 605) is arranged to enhance the difference signal by adding a scaled decorrelated mono downmix.
11. The decoder of claim 1 wherein the phase correction parameter is indicative of a correction for an interchannel phase difference of the set of upmix parameters.
12. The decoder of claim 1 wherein the second parts correspond to parts of the parametric signal associated with signals of the multiple channels meeting an out of phase criterion.
13. The decoder of claim 1 wherein the parametrically encoded signal comprises a phase correction parameter presence indication indicative of the second parts.

14. The decoder of claim 13 wherein the phase correction parameter presence indication comprises a common indication for all time frequency blocks of each encoding frame of the parametrically encoded signal.

15. The decoder of claim 13 wherein the phase correction parameter presence indication comprises individual presence indications for a plurality of sets of time frequency blocks of the down-mix.

16. The decoder of claim 1 wherein the phase correction parameter comprises individual parametric values for a plurality of sets of time frequency blocks of the second parts.

17. An encoder for encoding a multichannel signal comprising:
downmix means (501) for generating an encoded downmix of multiple channels of the multichannel signal;
parameter means (503) for generating a set of upmix parameters relating the downmix to the multiple channels;
means (505) for determining at least one section of the downmix in which a deviation between a value of phase parameter derived from the set of upmix parameters and a target value for the phase parameter meets a criterion;
means (507) for generating a phase correction parameter for a part of the encoded downmix associated with the at least one section; and
means (509) for generating an encoded signal comprising the encoded downmix, the set of upmix parameters and the phase correction parameter for the part of the encoded downmix.

18. The encoder of claim 17 wherein the downmix means is arranged to include a phase compensation to the downmix for the part of the downmix, and the phase correction parameter is indicative of the phase compensation.

19. A method of generating a multi-channel signal, the method comprising:
receiving a parametrically encoded signal comprising a downmix of multiple channels, a set of upmix parameters including a phase parameter, and a phase correction

parameter, the phase correction parameter being provided for second parts of the downmix but not for first parts of the downmix;

generating the multi-channel signal by upmixing the downmix based on the set of upmix parameters; and

modifying the upmixing in response to the phase correction parameter for the second parts of the downmix.

20. A method of encoding a multichannel signal comprising:

generating an encoded downmix of multiple channels of the multichannel signal;

generating a set of upmix parameters relating the downmix to the multiple channels;

determining at least one section of the downmix in which a deviation between a value of phase parameter derived from the set of upmix parameters and a target value for the phase parameter meets a criterion;

generating a phase correction parameter for a part of the encoded downmix associated with the at least one section; and

generating an encoded signal comprising the encoded downmix, the set of upmix parameters and the phase correction parameter for the part of the encoded downmix.

21. A computer program product for executing the method of any of the claims 19 or 20.

22. A parametrically encoded signal comprising a downmix of multiple channels, a set of upmix parameters including a phase parameter, and a phase correction parameter, the phase correction parameter being provided for second parts of the downmix but not for first parts of the downmix.

1/4

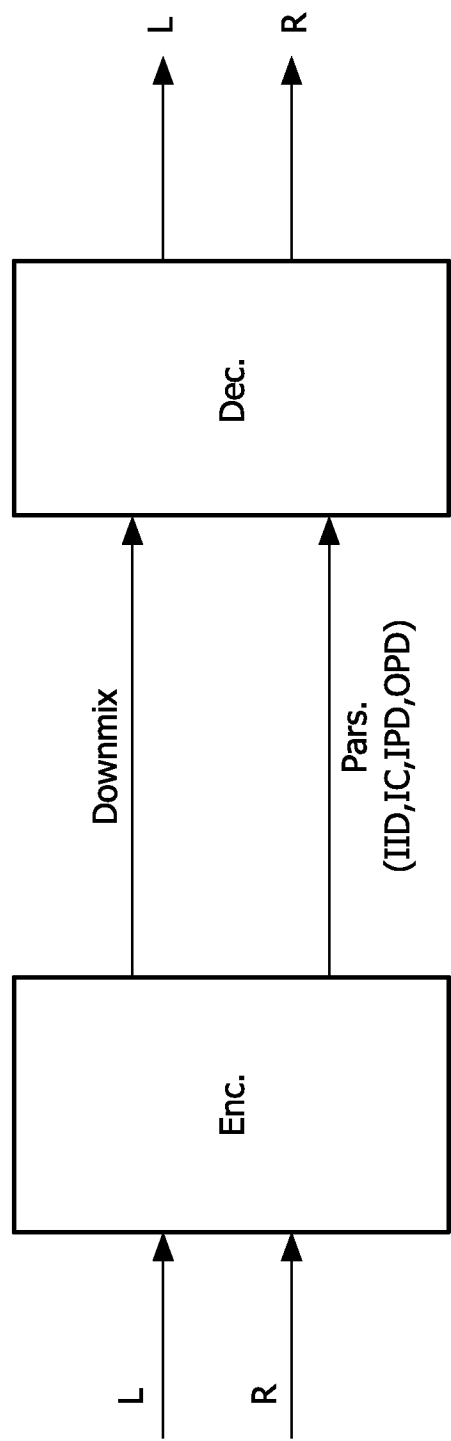


FIG. 1

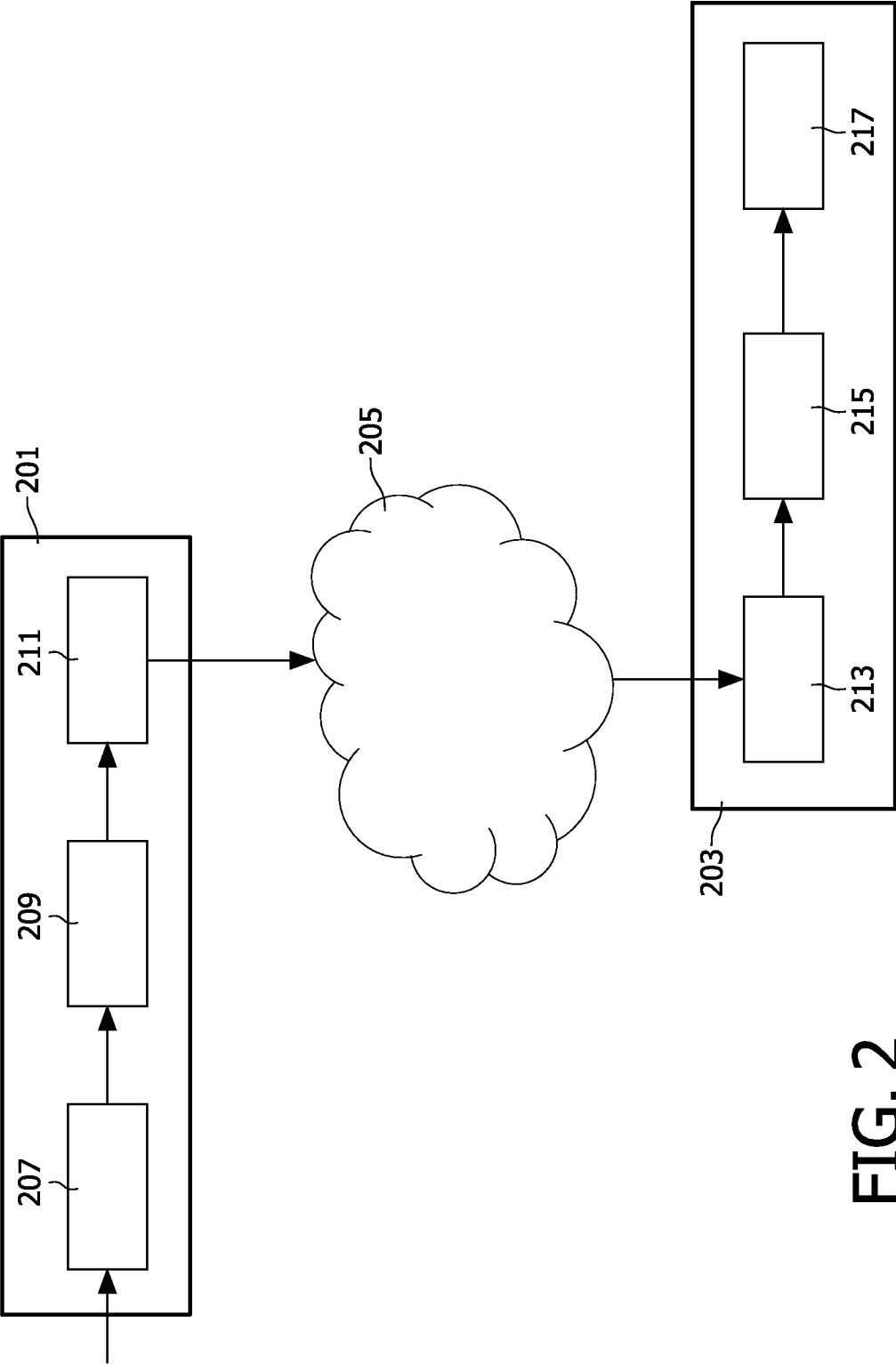
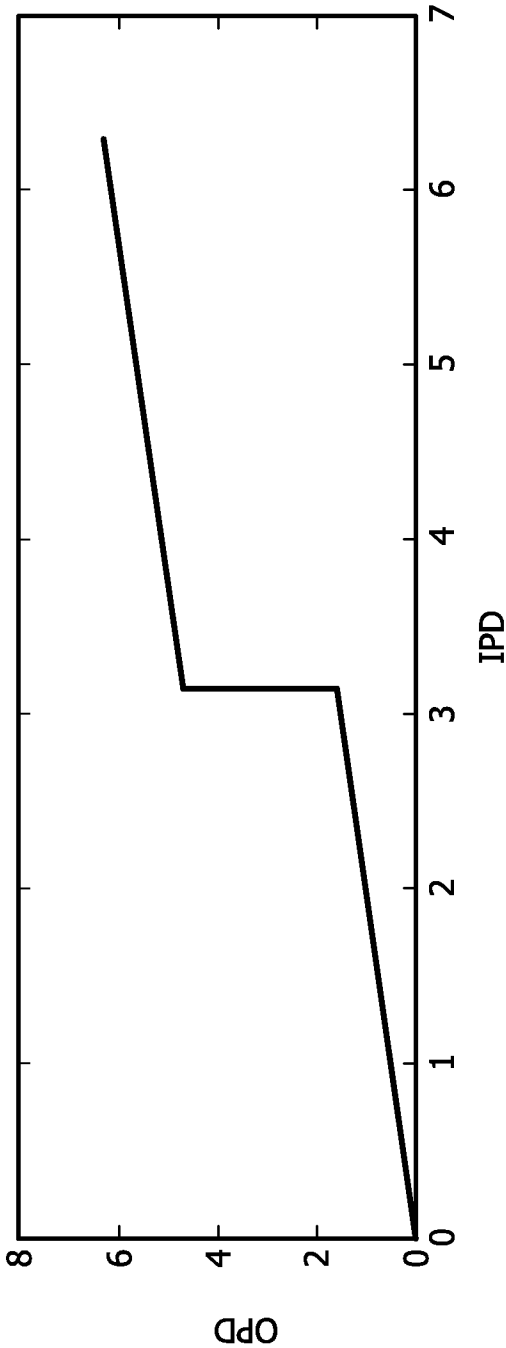
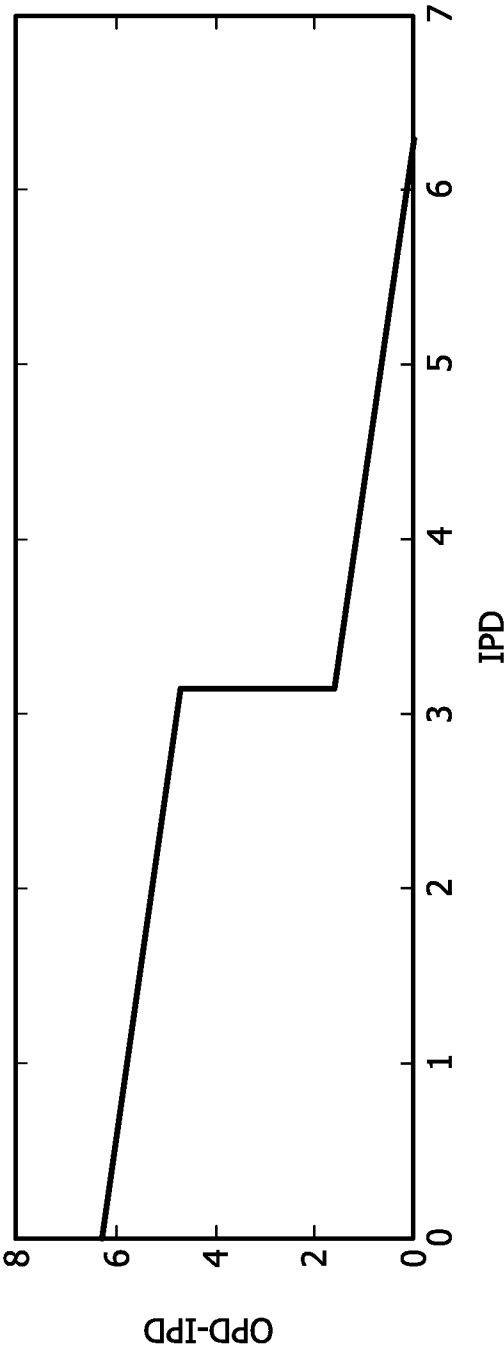


FIG. 2



4/4

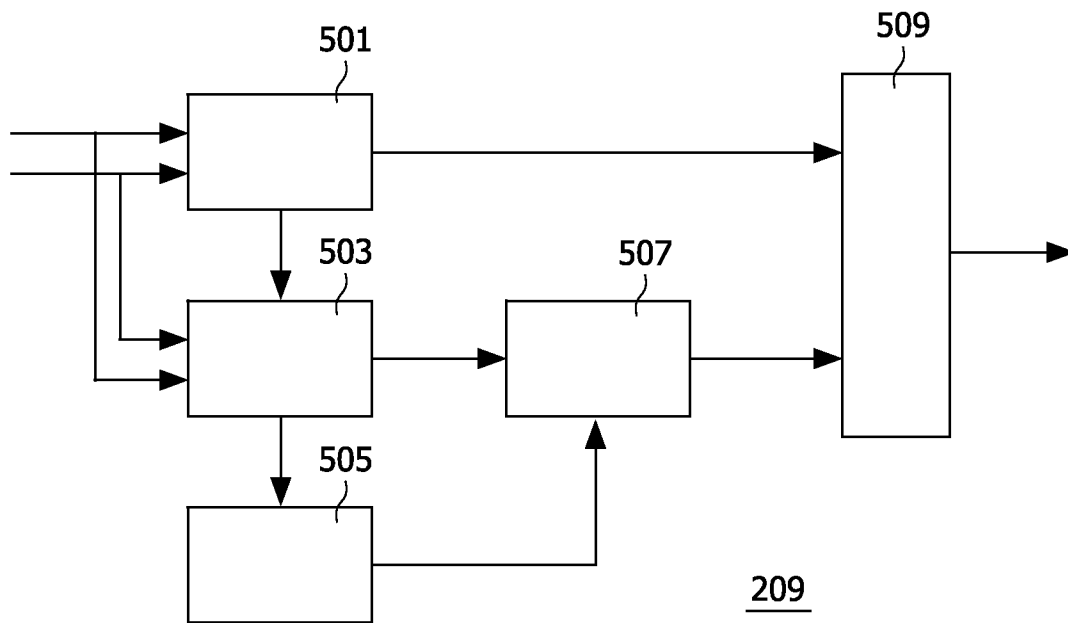


FIG. 5

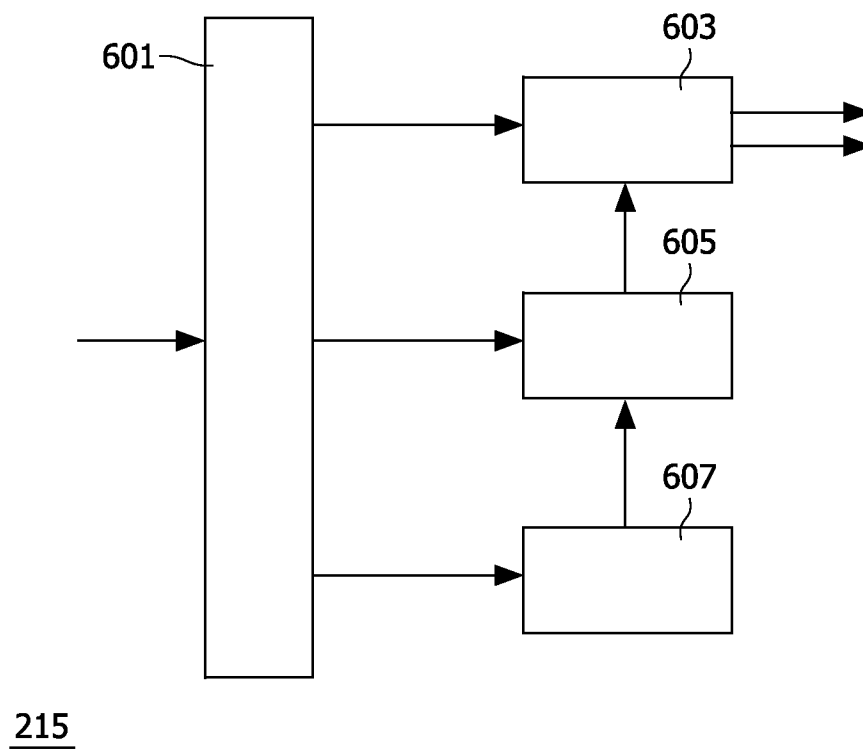


FIG. 6

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2010/050763

A. CLASSIFICATION OF SUBJECT MATTER

INV. G10L19/00

ADD. H04S3/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BREEBAART J ET AL: "Parametric Coding of Stereo Audio" INTERNET CITATION 1 June 2005 (2005-06-01), pages 1305-1322, XP002514252 ISSN: 1110-8657 Retrieved from the Internet: URL: http://www.jeroenbreebaart.com/papers/jasp/jasp2005.pdf [retrieved on 2009-02-10] cited in the application paragraph [05.4]	1,17, 19-22
A	US 5 632 005 A (DAVIS MARK F [US] ET AL) 20 May 1997 (1997-05-20) column 20, line 8 - column 22, line 30 column 10, line 11 - line 20 column 27, line 35 - column 30, line 1 ----- -/--	1,17, 19-22

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

16 April 2010

Date of mailing of the international search report

27/04/2010

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Krembel, Luc

INTERNATIONAL SEARCH REPORT

International application No

PCT/IB2010/050763

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JIMMY LAPIERRE AND ROCH LEFEBVRE: "On Improving Parametric Stereo Audio Coding" AES CONVENTION PAPER 6804,, 1 May 2006 (2006-05-01), pages 1-9, XP009131876 cited in the application paragraphs [02.3], [02.4] -----	1,17, 19-22
A,P	EP 2 124 224 A1 (LG ELECTRONICS INC [KR]) 25 November 2009 (2009-11-25) paragraph [0023] - paragraph [0031]; figure 1 -----	1,17, 19-22

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/IB2010/050763

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 5632005	A	20-05-1997	US 5633981 A	27-05-1997
			US 5583962 A	10-12-1996
EP 2124224	A1	25-11-2009	WO 2009142465 A2	26-11-2009
			US 2009325524 A1	31-12-2009