(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 May 2007 (18.05.2007)

PCT

(10) International Publication Number
WO 2007/055463 A1

(51) International Patent Classification:
H04S 3/00 (2006.01) *G10L 19/02* (2006.01)
G10L 19/00 (2006.01)

(21) International Application Number: PCT/KR2006/003424

(22) International Filing Date: 30 August 2006 (30.08.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/712,119	30 August 2005 (30.08.2005)	US
60/719,202	22 September 2005 (22.09.2005)	US
60/723,007	4 October 2005 (04.10.2005)	US
60/726,228	14 October 2005 (14.10.2005)	US
60/729,225	24 October 2005 (24.10.2005)	US
10-2006-0004051	13 January 2006 (13.01.2006)	KR
10-2006-0004057	13 January 2006 (13.01.2006)	KR
10-2006-0004062	13 January 2006 (13.01.2006)	KR
10-2006-0004063	13 January 2006 (13.01.2006)	KR
10-2006-0004055	13 January 2006 (13.01.2006)	KR
10-2006-0004065	13 January 2006 (13.01.2006)	KR
60/762,536	27 January 2006 (27.01.2006)	US

(71) **Applicant** *(for all designated States except US)*: **LG ELECTRONICS INC.** [KR/KR]; 20, Yoido-dong, Youngdungpo-gu, Seoul 150-721 (KR).

(72) **Inventors; and**

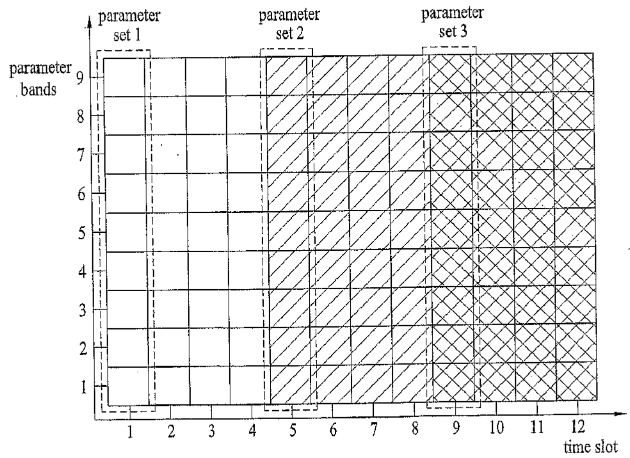
(75) **Inventors/Applicants** *(for US only)*: **PANG, Hee Suk** [KR/KR]; 101, #14-10, Yangjae-dong, Seocho-gu, Seoul 137-130 (KR). **OH, Hyeon O** [KR/KR]; 306-403, Gangseon Maeul 3-danji Apt., Juyeop 1(il)-dong, Ilsan-gu, Goyang-si, Gyeonggi-do 151-057 (KR). **KIM, Dong Soo** [KR/KR]; 1502, Woorim Villa, 602-265, Namhyeon-dong, Gwanak-gu, Seoul 151-801 (KR). **LIM, Jae Hyun** [KR/KR]; 609, Parkvill Officetel, 1062-20, Namhyeon-dong, Gwanak-gu, Seoul 151-801 (KR). **JUNG, Yang Won** [KR/KR]; 2-803, Yeoksam Hanshin Apt., Dogok-dong, Kangnam-gu, Seoul 135-270 (KR).

(74) **Agents: KIM, Yong In** et al.; 15th Floor Yo Sam Building, 648-23, Yeoksam-dong, Kangnam-gu, Seoul 135-808 (KR).

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

[Continued on next page]

(54) **Title:** APPARATUS FOR ENCODING AND DECODING AUDIO SIGNAL AND METHOD THEREOF



(57) **Abstract:** Spatial information associated with an audio signal is encoded into a bitstream, which can be transmitted to a decoder or recorded to a storage media. The bitstream can include different syntax related to time, frequency and spatial domains. In some embodiments, the bitstream includes one or more data structures (e.g., frames) that contain ordered sets of slots for which parameters can be applied. The data structures can be fixed or variable. A data structure type indicator can be inserted in the bitstream to enable a decoder to determine the data structure type and to invoke an appropriate decoding process. The data structure can include position information that can be used by a decoder to identify the correct slot for which a given parameter set is applied. The slot position information can be encoded with either a fixed number of bits or a variable number of bits based on the data structure type as indicated by the data structure type indicator. For variable data structure types, the slot position information can be encoded with a variable number of bits based on the position of the slot in the ordered set of slots.

WO 2007/055463 A1

(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, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

[TITLE OF THE INVENTION]

APPARATUS FOR ENCODING AND DECODING AUDIO SIGNAL AND METHOD THEREOF

Technical Field

The subject matter of this application is generally related audio signal processing.

Background Art

Efforts are underway to research and develop new approaches to perceptual coding of multi-channel audio, commonly referred to as Spatial Audio Coding (SAC). SAC allows transmission of multi-channel audio at low bit rates, making SAC suitable for many popular audio applications (e.g., Internet streaming, music downloads).

Rather than performing a discrete coding of individual audio input channels, SAC captures the spatial image of a multi-channel audio signal in a compact set of parameters. The parameters can be transmitted to a decoder where the parameters are used to synthesis or reconstruct the spatial properties of the audio signal.

In some SAC applications, the spatial parameters are transmitted to a decoder as part of a bitstream. The bitstream includes spatial frames that contain ordered sets of time slots for which spatial parameter sets can be applied. The bitstream also includes position information that can be used by a decoder to identify the correct time slot for which a given parameter set is applied.

Some SAC applications make use of conceptual elements in the encoding/decoding paths. One element is commonly referred to as One-To-Two (OTT) and another element is commonly referred to as Two-To-Three (TTT), where the names imply the number of input and output channels of a corresponding decoder element, respectively. The OTT encoder element extracts two spatial parameters and creates a downmix signal and residual signal. The TTT element mixes down three audio signals into a stereo downmix signal plus a residual signal. These elements can be combined to provide a variety of configurations of a spatial audio environment (e.g., surround sound).

Some SAC applications can operate in a non-guided operation mode, where only a stereo downmix signal is transmitted from an encoder to a decoder without a need for spatial parameter transmission. The decoder synthesizes spatial parameters from the downmix signal and uses those parameters to produce a multi-channel audio signal.

Disclosure of Invention

Spatial information associated with an audio signal is encoded into a bitstream, which can be transmitted to a decoder or recorded to a storage media. The bitstream can include different syntax related to time, frequency and spatial domains. In some embodiments, the bitstream includes one or more data structures (e.g., frames) that contain ordered sets of slots for which parameters can be applied. The data structures can be fixed or variable. A data structure type indicator can be inserted in the bitstream to enable a decoder to determine the data structure type and to invoke an appropriate decoding process. The data structure can include position information that can be used by a decoder to identify the correct slot for which a given parameter set is applied. The slot position information can be encoded with either a fixed number of bits or a variable number of bits based on the data structure type as indicated by the data structure type indicator. For variable data structure types, the slot position information can be encoded with a variable number of bits based on the position of the slot in the ordered set of slots.

In some embodiments, a method of encoding an audio signal includes: generating a parameter set corresponding to first or second information of an audio signal; and inserting the parameter set and corresponding first or second information in a bitstream representing the audio signal, wherein the first or second information is represented by a variable number of bits.

In some embodiments, a method of decoding an audio signal includes: determining a parameter set corresponding to first information or second information of an audio signal, where the parameter set and the corresponding first or second information is included in a bitstream representing an audio signal, and wherein the first or second information is represented in the bitstream by a variable number of bits; and decoding the audio signal based on the parameter set and the corresponding first or second information.

Other embodiments of time slot position coding of multiple frame types are disclosed that are directed to systems, methods, apparatuses, data structures and computer-readable mediums.

It is to be understood that both the foregoing general description and the following detailed description of the embodiments are exemplary and explanatory and are intended to provide further explanation of the invention as claimed.

Brief Description of Drawings

The accompanying drawings, which are included to provide a further understanding of the invention and are incorporated in and constitute part of this application, illustrate embodiment(s) of the invention, and together with the description, serve to explain the principle of the invention. In the drawings:

FIG. 1 is a diagram illustrating a principle of generating spatial information according to one embodiment of the present invention;

FIG. 2 is a block diagram of an encoder for encoding an audio signal according to one embodiment of the present invention;

FIG. 3 is a block diagram of a decoder for decoding an audio signal according to one embodiment of the present invention;

FIG. 4 is a block diagram of a channel converting module included in an upmixing unit of a decoder according to one embodiment of the present invention;

FIG. 5 is a diagram for explaining a method of configuring a bitstream of an audio signal according to one embodiment of the present invention;

FIGS. 6A and 6B are a diagram and a time/frequency graph, respectively, for explaining relationships between a parameter set, time slot and parameter bands according to one embodiment of the present invention;

FIG. 7A illustrates a syntax for representing configuration information of a spatial information signal according to one embodiment of the present invention;

FIG. 7B is a table for a number of parameter bands of a spatial information signal according to one embodiment of the present invention;

FIG. 8A illustrates a syntax for representing a number of parameter bands applied to an OTT box as a fixed number of bits according to one embodiment of the present invention;

FIG. 8B illustrates a syntax for representing a number of parameter bands applied to an OTT box by a variable number of bits according to one embodiment of the present invention;

FIG. 9A illustrates a syntax for representing a number of parameter bands applied to a TTT box by a fixed number of bits according to one embodiment of the present invention;

FIG. 9B illustrates a syntax for representing a number of parameter bands applied to a TTT box by a variable number of bits according to one embodiment of the present invention;

FIG. 10A illustrates a syntax of spatial extension configuration information for a spatial extension frame according to one embodiment of the present invention;

FIGS. 10B and 10C illustrate syntaxes of spatial extension configuration information for a residual signal in case that the residual signal is included in a spatial extension frame according to one embodiment of the present invention;

FIG. 10D illustrates a syntax for a method of representing a number of parameter bands for a residual signal according to one embodiment of the present invention;

FIG. 11A is a block diagram of a decoding apparatus in using non-guided coding according to one embodiment of the present invention;

FIG. 11B is a diagram for a method of representing a number of parameter bands as a group according to one embodiment of the present invention;

FIG. 12 illustrates a syntax of configuration information of a spatial frame according to one embodiment of the present invention;

FIG. 13A illustrates a syntax of position information of a time slot to which a parameter set is applied according to one embodiment of the present invention;

FIG. 13B illustrates a syntax for representing position information of a time slot to which a parameter set is applied as an absolute value and a difference value according to one embodiment of the present invention;

FIG. 13C is a diagram for representing a plurality of position information of time slots to which parameter sets are applied as a group according to one embodiment of the present invention;

FIG. 14 is a flowchart of an encoding method according to one embodiment of the present invention; and

FIG. 15 is a flowchart of a decoding method according to one embodiment of the present invention.

FIG. 16 is a block diagram of a device architecture for implementing the encoding and decoding processes described in reference to FIGS. 1-15.

Best Mode for Carrying Out the Invention

FIG. 1 is a diagram illustrating a principle of generating spatial information according to one embodiment of the present invention. Perceptual coding schemes for multi-channel audio signals are based on a fact that humans can perceive audio signals through three dimensional space. The three dimensional space of an audio signal can be represented using spatial information, including but not limited to the following known spatial parameters: Channel Level Differences (CLD), Inter-channel Correlation/Coherence (ICC), Channel Time Difference (CTD), Channel Prediction Coefficients (CPC), etc. The CLD parameter describes the energy (level) differences between two audio channels, the ICC parameter describes the amount of correlation or coherence between two audio channels and the CTD parameter describes the time difference between two audio channels.

The generation of CTD and CLD parameters is illustrated in FIG. 1. A first direct sound wave 103 from a remote sound source 101 arrives at a left human ear 107 and a second direct sound wave 102 is diffracted around a human head to reach a right human ear 106. The direct sound waves 102 and 103 differ from each other in arrival time and energy level. CTD and CLD parameters can be generated based on the arrival time and energy level differences of the sound waves 102 and 103, respectively. In addition, reflected sound waves 104 and 105 arrive at ears 106 and 107, respectively, and have no mutual correlations. An ICC parameter can be generated based on the correlation between the sound waves 104 and 105.

At the encoder, spatial information (e.g., spatial parameters) are extracted from a multi-channel audio input signal and a downmix signal is generated. The downmix signal and spatial parameters are transferred to a decoder. Any number of audio channels can be used for the downmix signal, including but not limited to: a mono signal, a stereo signal or a multi-channel audio signal. At the decoder, a multi-channel up-mix signal is created from the downmix signal and the spatial parameters.

FIG. 2 is a block diagram of an encoder for encoding an audio signal according to one embodiment of the present invention. The encoder includes a downmixing unit 202, a spatial information generating unit 203, a downmix signal encoding unit 207 and a multiplexing unit 209. Other configurations of an encoder are possible. Encoders can be implemented in hardware, software or a combination of both hardware and software. Encoders can be implemented in integrated circuit chips, chip sets, system on a chip (SoC), digital signal processors, general purpose processors and various digital and analog devices.

The downmixing unit 202 generates a downmix signal 204 from the multi-channel audio signal 201. In FIG. 2, $x_1,\ldots,x_n$ indicate input audio channels. As mentioned previously, the downmix signal 204 can be a mono signal, a stereo signal or a multi-channel audio signal. In the example shown, $x'_1,\ldots,x'_m$ indicate channel numbers of the downmix signal 204. In some embodiments, the encoder processes an externally provided downmix signal 205 (e.g., an artistic downmix) instead of the downmix signal 204.

The spatial information generating unit 203 extracts spatial information from the multi-channel audio signal 201. In this case, "spatial information" means information relating to the audio signal channels used in upmixing the downmix signal 204 to a multi-channel audio signal in the decoder. The downmix signal 204 is generated by downmixing the multi-channel audio signal. The spatial information is encoded to provide an encoded spatial information signal 206.

The downmix signal encoding unit 207 generates an encoded downmix signal 208 by encoding the downmix signal 204 generated from the downmixing unit 202.

The multiplexing unit 209 generates a bitstream 210 including the encoded downmix signal 208 and the encoded spatial information signal 206. The bitstream 210 can be transferred to a downstream decoder and/or recorded on a storage media.

FIG. 3 is a block diagram of a decoder for decoding an encoded audio signal according to one embodiment of the present invention. The decoder includes a demultiplexing unit 302, a downmix signal decoding unit 305, a spatial information decoding unit 307 and an upmixing unit 309. Decoders can be implemented in hardware, software or a combination of both hardware and software. Decoders can be implemented in integrated circuit chips, chip sets, system on a chip (SoC), digital signal processors, general purpose processors and various digital and analog devices.

In some embodiments, the demultiplexing unit 302 receives a bitstream 301 representing an audio signal and then separates an encoded downmix signal 303 and an encoded spatial information signal 304 from the bitstream 301. In FIG. 3, $x'_1,\ldots,x'_m$ indicate channels of the downmix signal 303. The downmix signal decoding unit 305 outputs a decoded downmix signal 306 by decoding the encoded downmix signal 303. If the decoder is unable to output a multi-channel audio signal, the downmix signal decoding unit 305 can directly output the downmix signal 306. In FIG. 3, $y'_1, \ldots, y'_m$ indicate direct output channels of the downmix signal decoding unit 305.

The spatial information signal decoding unit 307 extracts configuration information of the spatial information signal from the encoded spatial information signal 304 and then decodes the spatial information signal 304 using the extracted configuration information.

The upmixing unit 309 can up mix the downmix signal 306 into a multi-channel audio signal 310 using the extracted spatial information 308. In FIG. 3, $y_1, \ldots, y_n$ indicate a number of output channels of the upmixing unit 309.

FIG. 4 is a block diagram of a channel converting module which can be included in the upmixing unit 309 of the decoder shown in FIG. 3. In some embodiments, the upmixing unit 309 can include a plurality of channel converting modules. The channel converting module is a conceptual device that can differentiate a number of input channels and a number of output channels from each other using specific information.

In some embodiments, the channel converting module can include an OTT (one-to-two) box for converting one channel to two channels and vice versa, and a TTT (two-to-three) box for converting two channels to three channels and vice versa. The OTT and/or TTT boxes can be arranged in a variety of useful configurations. For example, the upmixing unit 309 shown in FIG. 3 can include a 5-1-5 configuration, a 5-2-5 configuration, a 7-2-7 configuration, a 7-5-7 configuration, etc. In a 5-1-5 configuration, a downmix signal having one channel is generated by downmixing five channels to a one channel, which can then be upmixed to five channels. Other configurations can be created in the same manner using various combinations of OTT and TTT boxes.

Referring to FIG. 4, an exemplary 5-2-5 configuration for an upmixing unit 400 is shown. In a 5-2-5 configuration, a downmix signal 401 having two channels is input to the upmixing unit 400. In the example shown, a left channel (L) and a right channel (R) are provided as input into the upmixing unit 400. In this embodiment, the upmixing unit 400 includes one TTT box 402 and three OTT boxes 406, 407 and 408. The downmix signal 401 having two channels is provided as input to the TTT box (TTTo) 402, which processes the downmix signal 401 and provides as output three channels 403, 404 and 405. One or more spatial parameters (e.g., CPC, CLD, ICC) can be provided as input to the TTT box 402, and are used to process the downmix signal 401, as described below. In some embodiments, a residual signal can be selectively provided as input to the TTT box 402. In such a case, the CPC can be described as a prediction coefficient for generating three channels from two channels.

The channel 403 that is provided as output from TTT box 402 is provided as input to OTT box 406 which generates two output channels using one or more spatial parameters. In the example shown, the two output channels represent front left (FL) and backward left (BL) speaker positions in, for example, a surround sound environment. The channel 404 is provided as input to OTT box 407, which generates two output channels using one or more spatial parameters. In the example shown, the two output channels represent front right (FR) and back right (BR) speaker positions. The channel 405 is provided as input to OTT box 408, which generates two output channels. In the example shown, the two output channels represent a center (C) speaker position and low frequency enhancement (LFE) channel. In this case, spatial information (e.g., CLD, ICC) can be provided as input to each of the OTT boxes. In some embodiments, residual signals (Res1, Res2) can be provided as inputs to the OTT boxes 406 and 407. In such an embodiment, a residual signal may not be provided as input to the OTT box 408 that outputs a center channel and an LFE channel.

The configuration shown in FIG. 4 is an example of a configuration for a channel converting module. Other configurations for a channel converting module are possible, including various combinations of OTT and TTT boxes. Since each of the channel converting modules can operate in a frequency domain, a number of parameter bands applied to each of the channel converting modules can be defined. A parameter band means at least one frequency band applicable to one parameter. The number of parameter bands is described in reference to FIG. 6B.

FIG. 5 is a diagram illustrating a method of configuring a bitstream of an audio signal according to one embodiment of the present invention. FIG. 5(a) illustrates a bitstream of an audio signal including a spatial information signal only, and FIGS. 5(b) and 5(c) illustrate a bitstream of an audio signal including a downmix signal and a spatial information signal.

Referring to FIG. 5(a), a bitstream of an audio signal can include configuration information 501 and a frame 503. The frame 503 can be repeated in the bitstream and in some embodiments includes a single spatial frame 502 containing spatial audio information.

In some embodiments, the configuration information 501 includes information describing a total number of time slots within one spatial frame 502, a total number of parameter bands spanning a frequency range of the audio signal, a number of parameter bands in an OTT box, a number of parameter bands in a TTT box and a number of parameter bands in a residual signal.

Other information can be included in the configuration information 501 as desired.

In some embodiments, the spatial frame 502 includes one or more spatial parameters (e.g., CLD, ICC), a frame type, a number of parameter sets within one frame and time slots to which parameter sets can be applied. Other information can be included in the spatial frame 502 as desired. The meaning and usage of the configuration information 501 and the information contained in the spatial frame 502 will be explained in reference to FIGS. 6 to 10.

Referring to FIG. 5(b), a bitstream of an audio signal may include configuration information 504, a downmix signal 505 and a spatial frame 506. In this case, one frame 507 can include the downmix signal 505 and the spatial frame 506, and the frame 507 may be repeated in the bitstream.

Referring to FIG. 5(c), a bitstream of an audio signal may include a downmix signal 508, configuration information 509 and a spatial frame 510. In this case, one frame 511 can include the configuration information 509 and the spatial frame 510, and the frame 511 may be repeated in the bitstream. If the configuration information 509 is inserted in each frame 511, the audio signal can be played back by a playback device at an arbitrary position.

Although FIG. 5(c) illustrates that the configuration information 509 is inserted in the bitstream by frame 511, it should be apparent that the configuration information 509 can be inserted in the bitstream by a plurality of frames which repeat periodically or non-periodically.

FIGS. 6A and 6B are diagrams illustrating relations between a parameter set, time slot and parameter bands according to one embodiment of the present invention. A parameter set means a one or more spatial parameters applied to one time slot. The spatial parameters can include spatial information, such as CDL, ICC, CPC, etc. A time slot means a time interval of an audio signal to which spatial parameters can be applied. One spatial frame can include one or more time slots.

Referring to FIG. 6A, a number of parameter sets 1,…,P can be used in a spatial frame, and each parameter set can include one or more data fields 1,…,Q-1. A parameter set can be applied to an entire frequency range of an audio signal, and each spatial parameter in the parameter set can be applied to one or more portions of the frequency band. For example, if a parameter set includes 20 spatial parameters, the entire frequency band of an audio signal can be divided into 20 zones (hereinafter referred to as "parameter bands") and the 20 spatial parameters of the parameter set can be applied to the 20 parameter bands. The parameters can be applied to the parameter bands as desired. For example, the spatial parameters can be densely applied to low frequency parameter bands and sparsely applied to high frequency parameter bands.

Referring to FIG. 6B, a time/frequency graph shows the relationship between parameter sets and time slots. In the example shown, three parameter sets (parameter set 1, parameter set 2, parameter set 3) are applied to an ordered set of 12 time slots in a single spatial frame. In this case, an entire frequency range of an audio signal is divided into 9 parameter bands. Thus, the horizontal axis indicates the number of time slots and the vertical axis indicates the number of parameter bands. Each of the three parameter sets is applied to a specific time slot. For example, a first parameter set (parameter set 1) is applied to a time slot #1, a second parameter set (parameter set 2) is applied to a time slot #5, and a third parameter set (parameter set 3) is applied to a time slot #9. The parameter sets can be applied to the other time slots by interpolating and/or copying the parameter sets to those time slots. Generally, the number of parameter sets can be equal to or less than the number of time slots, and the number of parameter bands can be equal to or less than the number of frequency bands of the audio signal. By encoding spatial information for portions of the time-frequency domain of an audio signal instead of the entire time-frequency domain of the audio signal, it is possible to reduce the amount of spatial information sent from an encoder to a decoder. This data reduction is possible since sparse information in the time-frequency domain is often sufficient for human auditory perception in accordance with known principals of perceptual audio coding.

An important feature of the disclosed embodiments is the encoding and decoding of time slot positions to which parameter sets are applied using a fixed or variable number of bits. The number of parameter bands can also be represented with a fixed number of bits or a variable number of bits. The variable bit coding scheme can also be applied to other information used in spatial audio coding, including but not limited to information associated with time, spatial and/or frequency domains (e.g., applied to a number of frequency subbands output from a filter bank).

FIG. 7A illustrates a syntax for representing configuration information of a spatial information signal according to one embodiment of the present invention. The configuration information includes a plurality of fields 701 to 718 to which a number of bits can be assigned.

A "bsSamplingFrequencyIndex" field 701 indicates a sampling frequency obtained from a sampling process of an audio signal. To represent the sampling frequency, 4 bits are allocated to the "bsSamplingFrequencyIndex" field 701. If a value of the "bsSamplingFrequencyIndex" field 701 is 15, i.e., a binary number of 1111, a "bsSamplingFrequency" field 702 is added to represent the sampling frequency. In this case, 24 bits are allocated to the "bsSamplingFrequency" field 702.

A "bsFrameLength" field 703 indicates a total number of time slots (hereinafter named "numSlots") within one spatial frame, and a relation of numSlots = bsFrameLength + 1 can exist between "numSlots" and the "bsFrameLength" field 703.

A "bsFreqRes" field 704 indicates a total number of parameter bands spanning an entire frequency domain of an audio signal. The "bsFreqRes" field 704 will be explained in FIG. 7B.

A "bsTreeConfig" field 705 indicates information for a tree configuration including a plurality of channel converting modules, such as described in reference to FIG. 4. The information for the tree configuration includes such information as a type of a channel converting module, a number of channel converting modules, a type of spatial information used in the channel converting module, a number of input/output channels of an audio signal, etc.

The tree configuration can have one of a 5-1-5 configuration, a 5-2-5 configuration, a 7-2-7 configuration, a 7-5-7 configuration and the like, according to a type of a channel converting module or a number of channels. The 5-2-5 configuration of the tree configuration is shown in FIG. 4.

A "bsQuantMode" field 706 indicates quantization mode information of spatial information.

A "bsOneIcc" field 707 indicates whether one ICC parameter sub-set is used for all OTT boxes. In this case, the parameter sub-set means a parameter set applied to a specific time slot and a specific channel converting module.

A "bsArbitraryDownmix" field 708 indicates a presence or non-presence of an arbitrary downmix gain.

A "bsFixedGainSur" field 709 indicates a gain applied to a surround channel, e.g., LS (left surround) and RS (right surround).

A "bsFixedgainLF" field 710 indicates a gain applied to a LFE channel.

A "bsFixedGainDM" field 711 indicates a gain applied to a downmix signal.

A "bsMatrixMode" field 712 indicates whether a matrix compatible stereo downmix signal is generated from an encoder.

A "bsTempShapeConfig" field 713 indicates an operation mode of temporal shaping (e.g., TES (temporal envelope shaping) and/or TP (temporal shaping)) in a decoder.

"bsDecorrConfig" field 714 indicates an operation mode of a decorrelator of a decoder.

And, "bs3DaudioMode" field 715 indicates whether a downmix signal is encoded into a 3D signal and whether an inverse HRTF processing is used.

After information of each of the fields has been determined/extracted in an encoder/decoder, information for a number of parameter bands applied to a channel converting module is determined/extracted in the encoder/decoder. A number of parameter bands applied to an OTT box is first determined/extracted (716) and a number of parameter bands applied to a TTT box is then determined/extracted (717). The number of parameter bands to the OTT box and/or TTT box will be described in detail with reference to FIGS. 8A to 9B.

In case that an extension frame exists, a "spatialExtensionConfig" block 718 includes configuration information for the extension frame. Information included in the "spatialExtensionConfig" block 718 will be described in reference to FIGS. 10A to 10D.

FIG. 7B is a table for a number of parameter bands of a spatial information signal according to one embodiment of the present invention. A "numBands" indicates a number of parameter bands for an entire frequency domain of an audio signal and "bsFreqRes" indicates index information for the number of parameter bands. For example, the entire frequency domain of an audio signal can be divided by a number of parameter bands as desired (e.g., 4, 5, 7, 10, 14, 20, 28, etc.).

In some embodiments, one parameter can be applied to each parameter band. For example, if the "numBands" is 28, then the entire frequency domain of an audio signal is divided into 28 parameter bands and each of the 28 parameters can be applied to each of the 28 parameter bands. In another example, if the "numBands" is 4, then the entire frequency domain of a given audio signal is divided into 4 parameter bands and each of the 4 parameters can be applied to each of the 4 parameter bands. In FIG. 7B, the term "Reserved" means that a number of parameter bands for the entire frequency domain of a given audio signal is not determined.

It should be noted a human auditory organ is not sensitive to the number of parameter bands used in the coding scheme. Thus, using a small number of parameter bands can provide a similar spatial audio effect to a listener than if a larger number of parameter bands were used.

Unlike the "numBands", the "numSlots" represented by the "bsFramelength" field 703 shown in FIG. 7A can represent all values. The values of "numSlots" may be limited, however, if the number of samples within one spatial frame is exactly divisible by the "numSlots." Thus, if a maximum value of the "numSlots" to be substantially represented is 'b', every value of the "bsFramelength" field 703 can be represented by ceil{$\log_2(b)$} bit(s). In this case, 'ceil(x)' means a minimum integer larger than or equal to the value 'x'. For example, if one spatial frame includes 72 time slots, then ceil{log2(72)} = 7 bits can be allocated to the "bsFrameLength" field 703, and the number of parameter bands applied to a channel converting module can be decided within the "numBands".

FIG. 8A illustrates a syntax for representing a number of parameter bands applied to an OTT box by a fixed number of bits according to one embodiment of the present invention. Referring to FIGS. 7A and 8A, a value of 'i' has a value of zero to numOttBoxes-1, where 'numOttBoxes' is the total number of OTT boxes. Namely, the value of 'i' indicates each OTT box, and a number of parameter bands applied to each OTT box is represented according to the value of 'i'. If an OTT box has an LFE channel mode, the number of parameter bands (hereinafter named "bsOttBands") applied to the LFE channel of the OTT box can be represented using a fixed number of bits. In the example shown in FIG. 8A, 5 bits are allocated to the "bsOttBands" field 801. If an OTT box does not have a LFE channel mode, the total number of parameter bands (numBands) can be applied to a channel of the OTT box.

FIG. 8B illustrates a syntax for representing a number of parameter bands applied to an OTT box by a variable number of bits according to one embodiment of the present invention. FIG.

8B, which is similar to FIG. 8A, differs from FIG. 8A in that "bsOttBands" field 802 shown in FIG. 8B is represented by a variable number of bits. In particular, the "bsOttBands" field 802, which has a value equal to or less than "numBands", can be represented by a variable number of bits using "numBands".

If the "numBands" lies within a range equal to or greater than 2^(n-1) and less than 2^(n), the "bsOttBands" field 802 can be represented by variable n bits.

For example: (a) if the "numBands" is 40, the "bsOttBands" field 802 is represented by 6 bits; (b) if the "numBands" is 28 or 20, the "bsOttBands" field 802 is represented by 5 bits; (c) if the "numBands" is 14 or 10, the "bsOttBands" field 802 is represented by 4 bits; and (d) if the "numBands" is 7, 5 or 4, the "bsOttBands" field 802 is represented by 3 bits.

If the "numBands" lies within a range greater than 2^(n-1) and equal to or less than 2^(n), the "bsOttBands" field 802 can be represented by variable n bits.

For example: (a) if the "numBands" is 40, the "bsOttBands" field 802 is represented by 6 bits; (b) if the "numBands" is 28 or 20, the "bsOttBands" field 802 is represented by 5 bits; (c) if the "numBands" is 14 or 10, the "bsOttBands" field 802 is represented by 4 bits; (d) if the "numBands" is 7 or 5, the "bsOttBands" field 802 is represented by 3 bits; and (e) if the "numBands" is 4, the "bsOttBands" field 802 is represented by 2 bits.

The "bsOttBands" field 802 can be represented by a variable number of bits through a function (hereinafter named "ceil function") of rounding up to a nearest integer by taking the "numBands" as a variable.

In particular, i) in case of 0<bsOttBands≤numBands or 0≤bsOttBands<numBands, the "bsOttBands" field 802 is represented by a number of bits corresponding to a value of ceil($\log_2$(numBands)) or ii) in case of 0≤bsOttBands≤numBands, the "bsOttBands" field 802 can be represented by ceil($\log_2$(numBands+1) bits.

If a value equal to or less than the "numBands" (hereinafter named "numberBands") is arbitrarily determined, the "bsOttBands" field 802 can be represented by a variable number of bits through the ceil function by taking the "numberBands" as a variable.

In particular, i) in case of 0<bsOttBands≤numberBands or 0≤bsOttBands<numberBands, the "bsOttBands" field 802 is represented by ceil($\log_2$(numberBands)) bits or ii) in case of 0≤bsOttBands≤numberBands, the "bsOttBands" field 802 can be represented by ceil($\log_2$(numberBands+1) bits.

If more than one OTT box is used, a combination of the "bsOttBands" can be expressed by Formula 1 below

$$\sum_{i=1}^{N} numBands^{i-1} \cdot bsOttBands_i, \qquad 0 \le bsOttBands_i < numBands$$

where, $bsOttBands_i$ indicates an i^{th} "bsOttBands". For example, assume there are three OTT boxes and three values (N=3) for the "bsOttBands" field 802. In this example, the three values of the "bsOttBands" field 802 (hereinafter named a1, a2 and a3, respectively) applied to the three OTT boxes, respectively, can be represented by 2 bits each. Hence, a total of 6 bits are needed to express the values a1, a2 and a3. Yet, if the values a1, a2 and a3 are represented as a group, then 27 (= 3*3*3) cases can occur, which can be represented by 5 bits, saving one bit. If the "numBands" is 3 and a group value represented by 5 bits is 15, the group value can be represented as 15=1x(3^2)+2*(3^1)+0*(3^0). Hence, a decoder can determine from the group value 15 that the three values a1, a2 and a3 of the "bsOttBands" field 802 are 1, 2 and 0, respectively, by applying the inverse of Formula 1.

In the case of multiple OTT boxes, the combination of "bsOttBands" can be represented as one of Formulas 2 to 4 (defined below) using the "numberBands". Since representation of "bsOttBands" using the "numberbands" is similar to the representation using the "numBands" in Formula 1, a detailed explanation shall be omitted and only the formulas are presented below.

[Formula 2]

$$\sum_{i=1}^{N}(numberBands+1)^{i-1}\cdot bsOttBands_i,\quad 0\le bsOttBands_i\le numberBands,$$

[Formula 3]

$$\sum_{i=1}^{N}numberBands^{i-1}\cdot bsOttBands_i,\quad 0\le bsOttBands_i< numberBands,$$

[Formula 4]

$$\sum_{i=1}^{N}numberBands^{i-1}\cdot bsOttBands_i,\quad 0< bsOttBands_i\le numberBands,$$

FIG. 9A illustrates a syntax for representing a number of parameter bands applied to a TTT box by a fixed number of bits according to one embodiment of the present invention. Referring to FIGS. 7A and 9A, a value of 'i' has a value of zero to numTttBoxes-1, where 'numTttBoxes' is a number of all TTT boxes. Namely, the value of 'i' indicates each TTT box. A number of parameter bands applied to each TTT box is represented according to the value of 'i'. In some embodiments, the TTT box can be divided into a low frequency band range and a high frequency band range, and different processes can be applied to the low and high frequency band ranges. Other divisions are possible.

A "bsTttDualMode" field 901 indicates whether a given TTT box operates in different modes (hereinafter called "dual mode") for a low band range and a high band range, respectively.

For example, if a value of the "bsTttDualMode" field 901 is zero, then one mode is used for the entire band range without discriminating between a low band range and a high band range. If a value of the "bsTttDualMode" field 901 is 1, then different modes can be used for the low band range and the high band range, respectively.

A "bsTttModeLow" field 902 indicates an operation mode of a given TTT box, which can have various operation modes. For example, the TTT box can have a prediction mode which uses, for example, CPC and ICC parameters, an energy-based mode which uses, for example, CLD parameters, etc. If a TTT box has a dual mode, additional information for a high band range may be needed.

A "bsTttModeHigh" field 903 indicates an operation mode of the high band range, in the case that the TTT box has a dual mode.

A "bsTttBandsLow" field 904 indicates a number of parameter bands applied to the TTT box.

A "bsTttBandsHigh" field 905 has "numBands".

If a TTT box has a dual mode, a low band range may be equal to or greater than zero and less than "bsTttBandsLow", while a high band range may be equal to or greater than "bsTttBandsLow" and less than "bsTttBandsHigh".

If a TTT box does not have a dual mode, a number of parameter bands applied to the TTT box may be equal to or greater than zero and less than "numBands" (907).

The "bsTttBandsLow" field 904 can be represented by a fixed number of bits. For instance, as shown in FIG. 9A, 5 bits can be allocated to represent the "bsTttBandsLow" field 904.

FIG. 9B illustrates a syntax for representing a number of parameter bands applied to a TTT box by a variable number of bits according to one embodiment of the present invention. FIG. 9B is similar to FIG. 9A but differs from FIG. 9A in representing a "bsTttBandsLow" field 907 of FIG. 9B by a variable number of bits while representing a "bsTttBandsLow" field 904 of FIG. 9A by a fixed number of bits. In particular, since the "bsTttBandsLow" field 907 has a value equal to or less than "numBands", the "bsTttBands" field 907 can be represented by a variable number of bits using "numBands".

In particular, in the case that the "numBands" is equal to or greater than 2^(n-1) and less than 2^(n), the "bsTttBandsLow" field 907 can be represented by n bits.

For example: (i) if the "numBands" is 40, the "bsTttBandsLow" field 907 is represented by 6 bits; (ii) if the "numBands" is 28 or 20, the "bsTttBandsLow" field 907 is represented by 5 bits; (iii) if the "numBands" is 14 or 10, the "bsTttBandsLow" field 907 is represented by 4 bits; and (iv) if the "numBands" is 7, 5 or 4, the "bsTttBandsLow" field 907 is represented by 3 bits.

If the "numBands" lies within a range greater than 2^(n-1) and equal to or less than 2^(n), then the "bsTttBandsLow" field 907 can be represented by n bits.

For example: (i) if the "numBands" is 40, the "bsTttBandsLow" field 907 is represented by 6 bits; (ii) if the "numBands" is 28 or 20, the "bsTttBandsLow" field 907 is represented by 5 bits; (iii) if the "numBands" is 14 or 10, the "bsTttBandsLow" field 907 is represented by 4 bits; (iv) if the "numBands" is 7 or 5, the "bsTttBandsLow" field 907 is represented by 3 bits; and (v) if the "numBands" is 4, the "bsTttBandsLow" field 907 is represented by 2 bits.

The "bsTttBandsLow" field 907 can be represented by a number of bits decided by a ceil function by taking the "numBands" as a variable.

For example: i) in case of $0<bsTttBandsLow\leq numBands$ or $0\leq bsTttBandsLow<numBands$, the "bsTttBandsLow" field 907 is represented by a number of bits corresponding to a value of $ceil(\log_2(numBands))$ or ii) in case of $0\leq bsTttBandsLow\leq numBands$, the "bsTttBandsLow" field 907 can be represented by $ceil(\log_2(numBands+1)$ bits.

If a value equal to or less than the "numBands", i.e., "numberBands" is arbitrarily determined, the "bsTttBandsLow" field 907 can be represented by a variable number of bits using the "numberBands".

In particular, i) in case of 0<bsTttBandsLow≤numberBands or 0≤bsTttBandsLow<numberBands, the "bsTttBandsLow" field 907 is represented by a number of bits corresponding to a value of ceil($\log_2$(numberBands)) or ii) in case of 0≤bsTttBandsLow≤numberBands, the "bsTttBandsLow" field 907 can be represented by a number of bits corresponding to a value of ceil($\log_2$(numberBands+1).

If the case of multiple TTT boxes, a combination of the "bsTttBandsLow" can be expressed as Formula 5 defined below.

[Formula 5]

$$\sum_{i=1}^{N} numBands^{i-1} \cdot bsTttBandsLow_i, \qquad 0 \le bsTttBandsLow_i < numBands$$

In this case, $bsTttBandsLow_i$ indicates an i^{th} "bsTttBandsLow". Since the meaning of Formula 5 is identical to that of Formula 1, a detailed explanation of Formula 5 is omitted in the following description.

In the case of multiple TTT boxes, the combination of "bsTttBandsLow" can be represented as one of Formulas 6 to 8 using the "numberBands". Since the meaning of Formulas 6 to 8 is identical to those of Formulas 2 to 4, a detailed explanation of Formulas 6 to 8 will be omitted in the following description.

[Formula 6]

$$\sum_{i=1}^{N}(numberBands+1)^{i-1}\cdot bsTttBandsLow_i,\quad 0\le bsTttBandsLow_i\le numberBands,$$

[Formula 7]

$$\sum_{i=1}^{N}numberBands^{i-1}\cdot bsTttBandsLow_i,\quad 0\le bsTttBandsLow_i<numberBands,$$

[Formula 8]

$$\sum_{i=1}^{N}numberBands^{i-1}\cdot bsTttBandsLow_i,\quad 0<bsTttBandsLow_i\le numberBands,$$

A number of parameter bands applied to the channel converting module (e.g., OTT box and/or TTT box) can be represented as a division value of the "numBands". In this case, the division value uses a half value of the "numBands" or a value resulting from dividing the "numBands" by a specific value.

Once a number of parameter bands applied to the OTT and/or TTT box is determined, parameter sets can be determined which can be applied to each OTT box and/or each TTT box within a range of the number of parameter bands. Each of the parameter sets can be applied to each OTT box and/or each TTT box by time slot unit. Namely, one parameter set can be applied to one time slot.

As mentioned in the foregoing description, one spatial frame can include a plurality of time slots. If the spatial frame is a fixed frame type, then a parameter set can be applied to a plurality of the time slots with an equal interval.

If the frame is a variable frame type, position information of the time slot to which the parameter set is applied is needed. This will be explained in detail later with reference to FIGS. 13A to 13C.

FIG. 10A illustrates a syntax for spatial extension configuration information for a spatial extension frame according to one embodiment of the present invention. Spatial extension configuration information can include a "bsSacExtType" field 1001, a "bsSacExtLen" field 1002, a "bsSacExtLenAdd" field 1003, a "bsSacExtLenAddAdd" field 1004 and a "bsFillBits" field 1007. Other fields are possible.

The "bsSacExtType" field 1001 indicates a data type of a spatial extension frame. For example, the spatial extension frame can be filled up with zeros, residual signal data, arbitrary downmix residual signal data or arbitrary tree data.

The "bsSacExtLen" field 1002 indicates a number of bytes of the spatial extension configuration information.

The "bsSacExtLenAdd" field 1003 indicates an additional number of bytes of spatial extension configuration information if a byte number of the spatial extension configuration information becomes equal to or greater than, for example, 15.

The "bsSacExtLenAddAdd" field 1004 indicates an additional number of bytes of spatial extension configuration information if a byte number of the spatial extension configuration information becomes equal to or greater than, for example, 270.

After the respective fields have been determined/extracted in an encoder/decoder, the configuration information for a data type included in the spatial extension frame is determined (1005).

As mentioned in the foregoing description, residual signal data, arbitrary downmix residual signal data, tree configuration data or the like can be included in the spatial extension frame.

Subsequently, a number of unused bits of a length of the spatial extension configuration information is calculated 1006.

The "bsFillBits" field 1007 indicates a number of bits of data that can be neglected to fill the unused bits.

FIGS. 10B and 10C illustrate syntaxes for spatial extension configuration information for a residual signal in case that the residual signal is included in a spatial extension frame according to one embodiment of the present invention.

Referring to FIG. 10B, a "bsResidualSamplingFrequencyIndex" field 1008 indicates a sampling frequency of a residual signal.

A "bsResidualFramesPerSpatialFrame" field 1009 indicates a number of residual frames per a spatial frame. For instance, 1, 2, 3 or 4 residual frames can be included in one spatial frame.

A "ResidualConfig" block 1010 indicates a number of parameter bands for a residual signal applied to each OTT and/or TTT box.

Referring to FIG. 10C, a "bsResidualPresent" field 1011 indicates whether a residual signal is applied to each OTT and/or TTT box.

A "bsResidualBands" field 1012 indicates a number of parameter bands of the residual signal existing in each OTT and/or TTT box if the residual signal exists in the each OTT and/or TTT box. A number of parameter bands of the residual signal can be represented by a fixed number of bits or a variable number of bits. In case that the number of parameter bands is represented by a fixed number of bits, the residual signal is able to have a value equal to or less than a total number of parameter bands of an audio signal. So, a bit number (e.g., 5 bits in FIG. 10C) necessary for representing a number of all parameter bands can be allocated.

FIG. 10D illustrates a syntax for representing a number of parameter bands of a residual signal by a variable number of bits according to one embodiment of the present invention. A "bsResidualBands" field 1014 can be represented by a variable number of bits using "numBands". If the numBands is equal to or greater than 2^(n-1) and less than 2^(n), the "bsResidualBands" field 1014 can be represented by n bits.

For instance: (i) if the "numBands" is 40, the "bsResidualBands" field 1014 is represented by 6 bits; (ii) if the "numBands" is 28 or 20, the "bsResidualBands" field 1014 is represented by 5 bits; (iii) if the "numBands" is 14 or 10, the "bsResidualBands" field 1014 is represented by 4 bits; and (iv) if the "numBands" is 7, 5 or 4, the "bsResidualBands" field 1014 is represented by 3 bits.

If the numBands is greater than 2^(n-1) and equal to or less than 2^(n), then the number of parameter bands of the residual signal can be represented by n bits.

For instance: (i) if the "numBands" is 40, the "bsResidualBands" field 1014 is represented by 6 bits; (ii) if the "numBands" is 28 or 20, the "bsResidualBands" field 1014 is represented by 5 bits; (iii) if the "numBands" is 14 or 10, the "bsResidualBands" field 1014 is represented by 4 bits; (iv) if the "numBands" is 7 or 5, the "bsResidualBands" field 1014 is represented by 3 bits; and (v) if the "numBands" is 4, the "bsResidualBands" field 1014 is represented by 2 bits.

Moreover, the "bsResidualBands" field 1014 can be represented by a bit number decided by a ceil function of rounding up to a nearest integer by taking the "numBands" as a variable.

In particular, i) in case of 0<bsResidualBands≤numBands or 0≤bsResidualBands<numBands, the "bsResidualBands" field 1014 is represented by ceil{$\log_2$(numBands)} bits or ii) in case of 0≤bsResidualBands≤numBands, the "bsResidualBands" field 1014 can be represented by ceil{$\log_2$(numBands+1)} bits.

In some embodiments, the "bsResidualBands" field 1014 can be represented using a value (numberBands) equal to or less than the numBands.

In particular, i) in case of 0<bsresidualBands≤numberBands or 0≤bsresidualBands<numberBands, the "bsResidualBands" field 1014 is represented by ceil{$\log_2$(numberBands)} bits or ii) in case of 0≤bsresidualBands≤numberBands, the "bsResidualBands" field 1014 can be represented by ceil{$\log_2$(numberBands+1)} bits.

If a plurality of residual signals (N) exist, a combination of the "bsResidualBands" can be expressed as shown in Formula 9 below.

[Formula 9]

$$\sum_{i=1}^{N} numBands^{i-1} \cdot bsResidualBands_i, \qquad 0 \le bsResidualBands_i < numBands,$$

In this case, $bsResidualBands_i$ indicates an i^{th} "bsresidualBands". Since a meaning of Formula 9 is identical to that of Formula 1, a detailed explanation of Formula 9 is omitted in the following description.

If there are multiple residual signals, a combination of the "bsresidualBands" can be represented as one of Formulas 10 to 12 using the "numberBands". Since representation of "bsresidualBands" using the "numberbands" is identical to the representation of Formulas 2 to 4, its detailed explanation shall be omitted in the following description.

[Formula 10]

$$\sum_{i=1}^{N}(numberBands+1)^{i-1}\cdot bsResidualBands_i,\quad 0\leq bsResidualBands_i\leq numberBands,$$

[Formula 11]

$$\sum_{i=1}^{N}numberBands^{i-1}\cdot bsResidualBands_i,\quad 0\leq bsResidualBands_i< numberBands,$$

[Formula 12]

$$\sum_{i=1}^{N}numberBands^{i-1}\cdot bsResidualBands_i,\quad 0< bsResidualBands_i\leq numberBands,$$

A number of parameter bands of the residual signal can be represented as a division value of the "numBands". In this case, the division value is able to use a half value of the "numBands" or a value resulting from dividing the "numBands" by a specific value.

The residual signal may be included in a bitstream of an audio signal together with a downmix signal and a spatial information signal, and the bitstream can be transferred to a decoder. The decoder can extract the downmix signal, the spatial information signal and the residual signal from the bitstream.

Subsequently, the downmix signal is upmixed using the spatial information. Meanwhile, the residual signal is applied to the downmix signal in the course of upmixing. In particular, the downmix signal is upmixed in a plurality of channel converting modules using the spatial information. In doing so, the residual signal is applied to the channel converting module. As mentioned in the foregoing description, the channel converting module has a number of parameter bands and a parameter set is applied to the channel converting module by a time slot unit. When the residual signal is applied to the channel converting module, the residual signal may be needed to update inter-channel correlation information of the audio signal to which the residual signal is applied. Then, the updated inter-channel correlation information is used in an up-mixing process.

FIG. 11A is a block diagram of a decoder for non-guided coding according to one embodiment of the present invention. Non-guided coding means that spatial information is not included in a bitstream of an audio signal.

In some embodiments, the decoder includes an analysis filterbank 1102, an analysis unit 1104, a spatial synthesis unit 1106 and a synthesis filterbank 1108. Although a downmix signal in a stereo signal type is shown in FIG. 11A, other types of downmix signals can be used.

In operation, the decoder receives a downmix signal 1101 and the analysis filterbank 1102 converts the received downmix signal 1101 to a frequency domain signal 1103. The analysis unit 1104 generates spatial information from the converted downmix signal 1103. The analysis unit 1104 performs a processing by a slot unit and the spatial information 1105 can be generated per a plurality of slots. In this case, the slot includes a time slot.

The spatial information can be generated in two steps. First, a downmix parameter is generated from the downmix signal. Second, the downmix parameter is converted to spatial information, such as a spatial parameter. In some embodiments, the downmix parameter can be generated through a matrix calculation of the downmix signal.

The spatial synthesis unit 1106 generates a multi-channel audio signal 1107 by synthesizing the generated spatial information 1105 with the downmix signal 1103. The generated multi-channel audio signal 1107 passes through the synthesis filterbank 1108 to be converted to a time domain audio signal 1109.

The spatial information may be generated at predetermined slot positions. The distance between the positions may be equal (i.e., equidistant). For example, the spatial information may be generated per 4 slots. The spatial information can also be generated at variable slot positions. In this case, the slot position information from which the spatial information is generated can be extracted from the bitstream. The position information can be represented by a variable number of bits. The position information can be represented as a absolute value and a difference value from a previous slot position information.

In case of using the non-guided coding, a number of parameter bands (hereinafter named "bsNumguidedBlindBands") for each channel of an audio signal can be represented by a fixed number of bits. The "bsNumguidedBlindBands" can be represented by a variable number of bits using "numBands". For example, if the "numBands" is equal to or greater than 2^(n-1) and less than 2^(n), the "bsNumguidedBlindBands" can be represented by variable n bits.

In particular, (a) if the "numBands" is 40, the "bsNumguidedBlindBands" is represented by 6 bits, (b) if the "numBands" is 28 or 20, the "bsNumguidedBlindBands" is represented by 5 bits, (c) if the "numBands" is 14 or 10, the "bsNumguidedBlindBands" is represented by 4 bits, and (d) if the "numBands" is 7, 5 or 4, the "bsNumguidedBlindBands" is represented by 3 bits.

If the "numBands" is greater than 2^(n-1) and equal to or less than 2^(n), then "bsNumguidedBlindBands" can be represented by variable n bits.

For instance: (a) if the "numBands" is 40, the "bsNumguidedBlindBands" is represented by 6 bits; (b) if the "numBands" is 28 or 20, the "bsNumguidedBlindBands" is represented by 5 bits; (c) if the "numBands" is 14 or 10, the "bsNumguidedBlindBands" is represented by 4 bits; (d) if the "numBands" is 7 or 5, the "bsNumguidedBlindBands" is represented by 3 bits; and (e) if the "numBands" is 4, the "bsNumguidedBlindBands" is represented by 2 bits.

Moreover, "bsNumguidedBlindBands" can be represented by a variable number of bits using the ceil function by taking the "numBands" as a variable.

For example, i) in case of $0<bsNumguidedBlindBands\leq numBands$ or $0\leq bsNumguidedBlindBands<numBands$, the "bsNumguidedBlindBands" is represented by $ceil\{log_2(numBands)\}$ bits or ii) in case of $0\leq bsNumguidedBlindBands\leq numBands$, the "bsNumguidedBlindBands" can be represented by $ceil\{log_2(numBands+1)\}$ bits.

If a value equal to or less than the "numBands", i.e., "numberBands" is arbitrarily determined, the "bsNumguidedBlindBands" can be represented as follows.

In particular, i) in case of

$0<bsNumguidedBlindBands \le numberBands$ or $0 \le bsNumguidedBlindBands<numberBands$, the "bsNumguidedBlindBands" is represented by $ceil\{\log_2(numberBands)\}$ bits or ii) in case of $0 \le bsNumguidedBlindBands \le numberBands$, the "bsNumguidedBlindBands" can be represented by $ceil\{\log_2(numberBands+1)\}$ bits.

If a number of channels (N) exist, a combination of the "bsNumguidedBlindBands" can be expressed as Formula 13.

[Formula 13]

$$\sum_{i=1}^{N} numBands^{i-1} \cdot bsNumGuidedBlindBands_i, \quad 0 \le bsNumGuidedBlindBands_i < numBands,$$

In this case, "$bsNumguidedBlindBands_i$" indicates an i^{th} "bsNumguidedBlindBands". Since the meaning of Formula 13 is identical to that of Formula 1, a detailed explanation of Formula 13 is omitted in the following description.

If there are multiple channels, the "bsNumguidedBlindBands" can be represented as one of Formulas 14 to 16 using the "numberBands". Since representation of "bsNumguidedBlindBands" using the "numberbands" is identical to the representations of Formulas 2 to 4, detailed explanation of Formulas 14 to 16 will be omitted in the following description.

[Formula 14]

$$\sum_{i=1}^{N} (numberBands+1)^{i-1} \cdot bsNumGuidedBlindBands_i, \quad 0 \le bsNumGuidedBlindBands_i \le numberBands,$$

[Formula 15]

$$\sum_{i=1}^{N} numberBands^{i-1} \cdot bsNumGuidedBlindBands_i, \quad 0 \le bsNumGuidedBlindBands_i < numberBands,$$

[Formula 16]

$$\sum_{i=1}^{N} numberBands^{i-1} \cdot bsNumGuidedBlindBands_i, \quad 0 < bsNumGuidedBlindBands_i \le numberBands,$$

FIG. 11B is a diagram for a method of representing a number of parameter bands as a group according to one embodiment of the present invention. A number of parameter bands includes number information of parameter bands applied to a channel converting module, number information of parameter bands applied to a residual signal and number information of parameter bands for each channel of an audio signal in case of using non-guided coding. In the case that there exists a plurality of number information of parameter bands, the plurality of the number information (e.g., "bsOttBands", "bsTttBands", "bsResidualBand" and/or "bsNumguidedBlindBands") can be represented as at least one or more groups.

Referring to FIG. 11B, if there are (kN+L) number information of parameter bands and if Q bits are needed to represent each number information of parameter bands, a plurality of number information of parameter bands can be represented as a following group. In this case, 'k' and 'N' are arbitrary integers not zero and 'L' is an arbitrary integer meeting 0≤L<N.

A grouping method includes the steps of generating k groups by binding N number information of parameter bands and generating a last group by binding last L number information of parameter bands. The k groups can be represented as M bits and the last group can be represented as p bits. In this case, the M bits are preferably less than N*Q bits used in the case of representing each number information of parameter bands without grouping them. The p bits are preferably equal to or less than L*Q bits used in case of representing each number information of the parameter bands without grouping them.

For instance, assume that two number information of parameter bands are b1 and b2, respectively. If each of the b1 and b2 is able to have five values, 3 bits are needed to represent each of the b1 and b2. In this case, even if the 3 bits are able to represent eight values, five values are substantially needed. So, each of the b1 and b2 has three redundancies. Yet, in case of representing the b1 and b2 as a group by binding the b1 and b2 together, 5 bits may be used instead of 6 bits (= 3 bits + 3 bits). In particular, since all combinations of the b1 and b2 include 25 (=5*5) types, a group of the b1 and b2 can be represented as 5 bits. Since the 5 bits are able to represent 32 values, seven redundancies are generated in case of the grouping representation. Yet, in case of a representation by grouping b1 and b2, redundancy is less than that of a case of representing each of the b1 and b2 as 3 bits. A method of representing a plurality of number information of parameter bands as groups can be implemented in various ways as follows.

If a plurality of number information of parameter bands have 40 kinds of values each, k groups are generated using 2, 3, 4, 5 or 6 as the N. The k groups can be represented as 11, 16, 22, 27 and 32 bits, respectively. Alternatively, the k groups are represented by combining the respective cases.

If a plurality of number information of parameter bands have 28 kinds of values each, k groups are generated using 6 as the N, and the k groups can be represented as 29 bits.

If a plurality of number information of parameter bands have 20 kinds of values each, k groups are generated using 2, 3, 4, 5, 6 or 7 as the N. The k groups can be represented as 9, 13, 18, 22, 26 and 31 bits, respectively. Alternatively, the k groups can be represented by combining the respective cases.

If a plurality of number information of parameter bands have 14 kinds of values each, k groups can be generated using 6 as the N. The k groups can be represented as 23 bits.

If a plurality of number information of parameter bands have 10 kinds of values each, k groups are generated using 2, 3, 4, 5, 6, 7, 8 or 9 as the N. The k groups can be represented as 7, 10, 14, 17, 20, 24, 27 and 30 bits, respectively.

Alternatively, the k groups can be represented by combining the respective cases.

If a plurality of number information of parameter bands have 7 kinds of values each, k groups are generated using 6, 7, 8, 9, 10 or 11 as the N. The k groups are represented as 17, 20, 23, 26, 29 and 31 bits, respectively. Alternatively, the k groups are represented by combining the respective cases.

If a plurality of number information of parameter bands have, for example, 5 kinds of values each, k groups can be generated using 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or 13 as the N. The k groups can be represented as 5, 7, 10, 12, 14, 17, 19, 21, 24, 26, 28 and 31 bits, respectively. Alternatively, the k groups are represented by combining the respective cases.

Moreover, a plurality of number information of parameter bands can be configured to be represented as the groups described above, or to be consecutively represented by making each number information of parameter bands into an independent bit sequence.

FIG. 12 illustrates syntax representing configuration information of a spatial frame according to one embodiment of the present invention. A spatial frame includes a "FramingInfo" block 1201, a "bsIndependencyfield 1202, a "OttData" block 1203, a "TttData" block 1204, a "SmgData" block 1205 and a "tempShapeData" block 1206.

The "FramingInfo" block 1201 includes information for a number of parameter sets and information for time slot to which each parameter set is applied. The "FramingInfo" block 1201 is explained in detail in FIG. 13A.

The "bsIndependencyFlag" field 1202 indicates whether a current frame can be decoded without knowledge for a previous frame.

The "OttData" block 1203 includes all spatial parameter information for all OTT boxes.

The "TttData" block 1204 includes all spatial parameter information for all TTT boxes.

The "SmgData" block 1205 includes information for temporal smoothing applied to a de-quantized spatial parameter.

The "TempShapeData" block 1206 includes information for temporal envelope shaping applied to a decorrelated signal.

FIG. 13A illustrates a syntax for representing time slot position information, to which a parameter set is applied, according to one embodiment of the present invention. A "bsFramingType" field 1301 indicates whether a spatial frame of an audio signal is a fixed frame type or a variable frame type. A fixed frame means a frame that a parameter set is applied to a preset time slot. For example, a parameter set is applied to a time slot preset with an equal interval. The variable frame means a frame that separately receives position information of a time slot to which a parameter set is applied.

A "bsNumParamSets" field 1302 indicates a number of parameter sets within one spatial frame (hereinafter named "numParamSets"), and a relation of "numParamSets = bsNumparamSets + 1" exists between the "numParamSets" and the "bsNumParamSets".

Since, e.g., 3 bits are allocated to the "bsNumParamSets" field 1302 in FIG. 13A, a maximum of eight parameter sets can be provided within one spatial frame. Since there is no limit on the number of allocated bits more parameter sets can be provided within a spatial frame.

If the spatial frame is a fixed frame type, position information of a time slot to which a parameter set is applied can be decided according to a preset rule, and additional position information of a time slot to which a parameter set is applied is unnecessary. However, if the spatial frame is a variable frame type, position information of a time slot to which a parameter set is applied is needed.

A "bsParamSlot" field 1303 indicates position information of a time slot to which a parameter set is applied. The "bsParamSlot" field 1303 can be represented by a variable number of bits using the number of time slots within one spatial frame, i.e., "numSlots". In particular, in case that the "numSlots" is equal to or greater than 2^(n-1) and less than 2^(n), the "bsParamSlot" field 1103 can be represented by n bits.

For instance: (i) if the "numSlots" lies within a range between 64 and 127, the "bsParamSlot" field 1303 can be represented by 7 bits; (ii) if the "numSlots" lies within a range between 32 and 63, the "bsParamSlot" field 1303 can be represented by 6 bits; (iii) if the "numSlots" lies within a range between 16 and 31, the "bsParamSlot" field 1303 can be represented by 5 bits; (iv) if the "numSlots" lies within a range between 8 and 15, the "bsParamSlot" field 1303 can be represented by 4 bits; (v) if the "numSlots" lies within a range between 4 and 7, the "bsParamSlot" field 1303 can be represented by 3 bits; (vi) if the "numSlots" lies within a range between 2 and 3, the "bsParamSlot" field 1303 can be represented by 2 bits; (vii) if the "numSlots" is 1, the "bsParamSlot" field 1303 can be represented by 1 bit; and (viii) if the "numSlots" is 0, the "bsParamSlot" field 1303 can be represented by 0 bit. Likewise, if the "numSlots" lies within a range between 64 and 127, the "bsParamSlot" field 1303 can be represented by 7 bits.

If there are multiple parameter sets (N), a combination of the "bsParamSlot" can be represented according to Formula 9.

[Formula 9]

$$\sum_{i=1}^{N} numSlots^{i-1} \cdot bsParamSlot_i, \qquad 0 \le bsParamSlot_i < numSlots,$$

In this case, "bsParamSlots$_i$" indicates a time slot to which an ith parameter set is applied. For instance, assume that the "numSlots" is 3 and that the "bsParamSlot" field 1303 can have ten values. In this case, three information (hereinafter named c1, c2 and c3, respectively) for the "bsParamSlot" field 1303 are needed. Since 4 bits are needed to represent each of the c1, c2 and c3, total 12 (= 4*3) bits are needed. In case of representing the c1, c2 and c3 as a group by binding them together, 1,000 (= 10*10*10) cases can occur, which can be represented as 10 bits, thus saving 2 bits. If the "numSlots" is 3 and if the value read as 5 bits is 31, the value can be represented as 31=1x(3^2)+5*(3^1)+7*(3^0). A decoder apparatus can determine that the c1, c2 and c3 are 1, 5 and 7, respectively, by applying the inverse of Formula 9.

FIG. 13B illustrates a syntax for representing position information of a time slot to which a parameter set is applied as an absolute value and a difference value according to one embodiment of the present invention. If a spatial frame is a variable frame type, the "bsParamSlot" field 1303 in FIG. 13A can be represented as an absolute value and a difference value using a fact that "bsParamSlot" information increases monotonously.

For instance: (i) a position of a time slot to which a first parameter set is applied can be generated into an absolute value, i.e., "bsParamSlot[0]"; and (ii) a position of a time slot to which a second or higher parameter set is applied can be generated as a difference value, i.e., "difference value" between "bsParamSlot[ps]" and "bsParamslot[ps-1]" or "difference value - 1" (hereinafter named "bsDiffParamSlot[ps]"). In this case, "ps" means a parameter set.

The "bsParamSlot[0]" field 1304 can be represented by a number of bits (hereinafter named "nBitsParamSlot(0)") calculated using the "numSlots" and the "numParamSets".

The "bsDiffParamSlot[ps]" field 1305 can be represented by a number of bits (hereinafter named "nBitParamSlot(ps)") calculated using the "numSlots", the "numParamSets" and a position of a time slot to which a previous parameter set is applied, i.e., "bsParamSlot[ps-1]".

In particular, to represent "bsParamSlot[ps]" by a minimum number of bits, a number of bits to represent the "bsParamSlot[ps]" can be decided based on the following rules: (i) a plurality of the "bsParamSlot[ps]" increase in an ascending series (bsParamSlot[ps]> bsParamSlot[ps-1]); (ii) a maximum value of the "bsParamSlot[0]" is "numSlots - NumParamSets"; and (iii) in case of 0 < ps < numParamSets, "bsParamSlot[ps]" can have a value between "bsParamSlot[ps-1] + 1" and "numSlots - numParamSets + ps" only.

For example, if the "numSlots" is 10 and if the "numParamSets" is 3, since the "bsParamSlot[ps]" increases in an ascending series, a maximum value of the "bsParamSlot[0]" becomes "10-3=7". Namely, the "bsParamSlot[0]" should be selected from values of 0 to 7. This is because a number of time slots for the rest of parameter sets (e.g., if ps is 1 or 2) is insufficient if the "bsParamSlot[0]" has a value greater than 7.

If "bsParamSlot[0]" is 5, a time slot position bsParamSlot[1] for a second parameter set should be selected from values between "5+1=6" and "10-3+1=8".

If "bsParamSlot[1]" is 7, "bsParamSlot[2]" can become 8 or 9. If "bsParamSlot[1]" is 8, "bsParamSlot[2]" can become 9.

Hence, the "bsParamSlot[ps]" can be represented as a variable bit number using the above features instead of being represented as fixed bits.

In configuring the "bsParamSlot[ps]" in a bitstream, if the "ps" is 0, the "bsParamSlot[0]" can be represented as an absolute value by a number of bits corresponding to "nBitsParamSlot(0)". If the "ps" is greater than 0, the "bsParamSlot[ps]" can be represented as a difference value by a number of bits corresponding to "nBitsParamSlot(ps)". In reading the above-configured "bsParamSlot[ps]" from a bitstream, a length of a bitstream for each data, i.e., "nBitsParamSlot[ps]" can be found using Formula 10.

[Formula 10]

$$f_b(x)=\begin{cases}0\,bit, & if\ x=1,\\ 1\,bit, & if\ x=2,\\ 2\,bits, & if\ 3\le x\le 4,\\ 3\,bits, & if\ 5\le x\le 8,\\ 4\,bits, & if\ 9\le x\le 16,\\ 5\,bits, & if\ 17\le x\le 32,\\ 6\,bits, & if\ 33\le x\le 64,\end{cases}$$

In particular, the "nBitsParamSlot[ps]" can be found as nBitsParamSlot[0]=f_b(numSlots - numParamSets + 1). If 0<ps<numParamSets, the "nBitsParamSlot[ps]" can be found as nBitsParamSlot[ps]=f_b(numSlots-numParamSets+ps-bsParamSlot[ps-1]). The "nBitsParamSlot[ps]" can be determined using Formula 11, which extends Formula 10 up to 7 bits.

[Formula 11]

$$f_b(x)=\begin{cases}0\,bit, & if\ x=1,\\ 1\,bit, & if\ x=2,\\ 2\,bits, & if\ 3\le x\le 4,\\ 3\,bits, & if\ 5\le x\le 8,\\ 4\,bits, & if\ 9\le x\le 16,\\ 5\,bits, & if\ 17\le x\le 32,\\ 6\,bits, & if\ 33\le x\le 64,\\ 7\,bits, & if\ 65\le x\le 128,\end{cases}$$

An example of the function $f_b(x)$ is explained as follows. If "numSlots" is 15 and if "numParamSets" is 3, the function can be evaluated as nBitsParamSlot[0] = f_b(15-3+1) = 4bits.

If the "bsParamSlot[0]" represented by 4 bits is 7, the function can be evaluated as nBitsParamSlot[1] = f_b(15-3+1-7) = 3bits. In this case, "bsDiffParamSlot[1]" field 1305 can be represented by 3 bits.

If the value represented by the 3 bits is 3, "bsParamSlot[1]" becomes 7+3 = 10. Hence, it becomes nBitsParamSlot[2] = fb(15-3+2-10) = 2bits. In this case, "bsDiffParamSlot[2]" field 1305 can be represented by 2 bits. If the number of remaining time slots is equal to a number of a remaining parameter sets, 0 bits may be allocated to the "bsDiffParamSlot[ps]" field. In other words, no additional information is needed to represent the position of the time slot to which the parameter set is applied.

Thus, a number of bits for "bsParamSlot[ps]" can be variably decided. The number of bits for "bsParamSlot[ps]" can be read from a bitstream using the function $f_b(x)$ in a decoder. In some embodiments, the function $f_b(x)$ can include the function ceil($\log_2(x)$).

In reading information for "bsParamSlot[ps]" represented as the absolute value and the difference value from a bitstream in a decoder, first the "bsParamSlot[0]" may be read from the bitstream and then the "bsDiffParamSlot[ps]" may be read for 0<ps<numParamSets. The "bsParamSlot[ps]" can then be found for

an interval 0≤ps<numParamSets using the "bsParamSlot[0]" and the "bsDiffParamSlot[ps]". For example, as shown in FIG. 13B, a "bsParamSlot[ps]" can be found by adding a "bsParamSlot[ps-1]" to a "bsDiffParamSlot[ps]+1".

FIG. 13C illustrates a syntax for representing position information of a time slot to which a parameter set is applied as a group according to one embodiment of the present invention. In case that a plurality of parameter sets exist, a plurality of "bsParamSlots" 1307 for a plurality of the parameter sets can be represented as at least one or more groups.

If a number of the "bsParamSlots" 1307 is (kN+L) and if Q bits are needed to represent each of the "bsParamSlots" 1307, the "bsParamSlots" 1307 can be represented as a following group. In this case, 'k' and 'N' are arbitrary integers not zero and 'L' is an arbitrary integer meeting 0≤L<N.

A grouping method can include the steps of generating k groups by binding N "bsParamSlots" 1307 each and generating a last group by binding last L "bsParamSlots" 1307. The k groups can be represented by M bits and the last group can be represented by p bits. In this case, the M bits are preferably less than N*Q bits used in the case of representing each of the "bsParamSlots" 1307 without grouping them. The p bits are preferably equal to or less than L*Q bits used in the case of representing each of the "bsParamSlots" 1307 without grouping them.

For example, assume that a pair of "bsParamSlots" 1307 for two parameter sets are d1 and d2, respectively. If each of the d1 and d2 is able to have five values, 3 bits are needed to represent each of the d1 and d2. In this case, even if the 3 bits are able to represent eight values, five values are substantially needed. So, each of the d1 and d2 has three redundancies. Yet, in case of representing the d1 and d2 as a group by binding the d1 and d2 together, 5 bits are used instead of using 6 bits (= 3 bits + 3 bits). In particular, since all combinations of the d1 and d2 include 25 (= 5*5) types, a group of the d1 and d2 can be represented as 5 bits only. Since the 5 bits are able to represent 32 values, seven redundancies are generated in case of the grouping representation. Yet, in case of a representation by grouping the d1 and d2, redundancy is smaller than that of a case of representing each of the d1 and d2 as 3 bits.

In configuring the group, data for the group can be configured using "bsParamSlot[0]" for an initial value and a difference value between pairs of the "bsParamSlot[ps]" for a second or higher value.

In configuring the group, bits can be directly allocated without grouping if a number of parameter set is 1 and bits can be allocated after completion of grouping if a number of parameter sets is equal to or greater than 2.

FIG. 14 is a flowchart of an encoding method according to one embodiment of the present invention. A method of encoding an audio signal and an operation of an encoder according to the present invention are explained as follows.

First, a total number of time slots (numSlots) in one spatial frame and a total number of parameter bands (numBands) of an audio signal are determined (S1401).

Then, a number of parameter bands applied to a channel converting module (OTT box and/or TTT box) and/or a residual signal are determined (S1402).

If the OTT box has a LFE channel mode, the number of parameter bands applied to the OTT box is separately determined.

If the OTT box does not have the LFE channel mode, "numBands" is used as a number of the parameters applied to the OTT box.

Subsequently, a type of a spatial frame is determined. In this case, the spatial frame may be classified into a fixed frame type and a variable frame type.

If the spatial frame is the variable frame type (S1403), a number of parameter sets used within one spatial frame is determined (S1406). In this case, the parameter set can be applied to the channel converting module by a time slot unit.

Subsequently, a position of time slot to which the parameter set is applied is determined (S1407). In this case, the position of time slot to which the parameter set is applied can be represented as an absolute value and a difference value. For example, a position of a time slot to which a first parameter set is applied can be represented as an absolute value, and a position of a time slot to which a second or higher parameter set is applied can be represented as a difference value from a position of a previous time slot. In this case, the position of a time slot to which the parameter set is applied can be represented by a variable number of bits.

In particular, a position of time slot to which a first parameter set is applied can be represented by a number of bits calculated using a total number of time slots and a total number of parameter sets. A position of a time slot to which a second or higher parameter set is applied can be represented by a number of bits calculated using a total number of time slots, a total number of parameter sets and a position of a time slot to which a previous parameter set is applied.

If the spatial frame is a fixed frame type, a number of parameter sets used in one spatial frame is determined (S1404). In this case, a position of a time slot to which the parameter set is applied is decided using a preset rule. For example, a position of a time slot to which a parameter set is applied can be decided to have an equal interval from a position of a time slot to which a previous parameter set is applied (S1405).

Subsequently, a downmixing unit and a spatial information generating unit generate a downmix signal and spatial information, respectively, using the above-determined total number of time slots, a total number of parameter bands, a number of parameter bands to be applied to the channel converting unit, a total number of parameter sets in one spatial frame and position information of the time slot to which a parameter set is applied (S1408).

Finally, a multiplexing unit generates a bitstream including the downmix signal and the spatial information (S1409) and then transfers the generated bitstream to a decoder (S1409).

FIG. 15 is a flowchart of a decoding method according to one embodiment of the present invention. A method of decoding an audio signal and an operation of a decoder according to the present invention are explained as follows.

First, a decoder receives a bitstream of an audio signal (S1501). A demultiplexing unit separates a downmix signal and a spatial information signal from the received bitstream (S1502). Subsequently, a spatial information signal decoding unit extracts information for a total number of time slots in one spatial frame, a total number of parameter bands and a number of parameter bands applied to a channel converting module from configuration information of the spatial information signal (S1503).

If the spatial frame is a variable frame type (S1504), a number of parameter sets in one spatial frame and position information of a time slot to which the parameter set is applied are extracted from the spatial frame (S1505). The position information of the time slot can be represented by a fixed or variable number of bits. In this case, position information of time slot to which a first parameter set is applied may be represented as an absolute value and position information of time slots to which a second or higher parameter sets are applied can be represented as a difference value. The actual position information of time slots to which the second or higher parameter sets are applied can be found by adding the difference value to the position information of the time slot to which a previous parameter set is applied.

Finally, the downmix signal is converted to a multi-channel audio signal using the extracted information (S1506).

The disclosed embodiments described above provide several advantages over conventional audio coding schemes.

First, in coding a multi-channel audio signal by representing a position of a time slot to which a parameter set is applied by a variable number of bits, the disclosed embodiments are able to reduce a transferred data quantity.

Second, by representing a position of a time slot to which a first parameter set is applied as an absolute value, and by representing positions of time slots to which a second or higher parameter sets are applied as a difference value, the disclosed embodiments can reduce a transferred data quantity.

Third, by representing a number of parameter bands applied to such a channel converting module as an OTT box and/or a TTT box by a fixed or variable number of bits, the disclosed embodiments can reduce a transferred data quantity. In this case, positions of time slots to which parameter sets are applied can be represented using the aforesaid principle, where the parameter sets may exist in range of a number of parameter bands.

FIG. 16 is a block diagram of an exemplary device architecture 1600 for implementing the audio encoder/decoder, as described in reference to FIGS. 1-15. The device architecture 1600 is applicable to a variety of devices, including but not limited to: personal computers, server computers, consumer electronic devices, mobile phones, personal digital assistants (PDAs), electronic tablets, television systems, television set-top boxes, game consoles, media players, music players, navigation systems, and any other device capable of decoding audio signals. Some of these devices may implement a modified architecture using a combination of hardware and software.

The architecture 1600 includes one or more processors 1602 (e.g., PowerPC®, Intel Pentium® 4, etc.), one or more display devices 1604 (e.g., CRT, LCD), an audio subsystem 1606 (e.g., audio hardware/software), one or more network interfaces 1608 (e.g., Ethernet, FireWire®, USB, etc.), input devices 1610 (e.g., keyboard, mouse, etc.), and one or more computer-readable mediums 1612 (e.g., RAM, ROM, SDRAM, hard disk, optical disk, flash memory, etc.). These components can exchange communications and data via one or more buses 1614 (e.g., EISA, PCI, PCI Express, etc.).

The term "computer-readable medium" refers to any medium that participates in providing instructions to a processor 1602 for execution, including without limitation, non-volatile media (e.g., optical or magnetic disks), volatile media (e.g., memory) and transmission media. Transmission media includes, without limitation, coaxial cables, copper wire and fiber optics. Transmission media can also take the form of acoustic, light or radio frequency waves.

The computer-readable medium 1612 further includes an operating system 1616 (e.g., Mac OS®, Windows®, Linux, etc.), a network communication module 1618, an audio codec 1620 and one or more applications 1622.

The operating system 1616 can be multi-user, multiprocessing, multitasking, multithreading, real-time and the like. The operating system 1616 performs basic tasks, including but not limited to: recognizing input from input devices 1610; sending output to display devices 1604 and the audio subsystem 1606; keeping track of files and directories on computer-readable mediums 1612 (e.g., memory or a storage device); controlling peripheral devices (e.g., disk drives, printers, etc.); and managing traffic on the one or more buses 1614.

The network communications module 1618 includes various components for establishing and maintaining network connections (e.g., software for implementing communication protocols, such as TCP/IP, HTTP, Ethernet, etc.). The network communications module 1618 can include a browser for enabling operators of the device architecture 1600 to search a network (e.g., Internet) for information (e.g., audio content).

The audio codec 1620 is responsible for implementing all or a portion of the encoding and/or decoding processes described in reference to FIGS. 1-15. In some embodiments, the audio codec works in conjunction with hardware (e.g., processor(s) 1602, audio subsystem 1606) to process audio signals, including encoding and/or decoding audio signals in accordance with the present invention described herein.

The applications 1622 can include any software application related to audio content and/or where audio content is encoded and/or decoded, including but not limited to media players, music players (e.g., MP3 players), mobile phone applications, PDAs, television systems, set-top boxes, etc. In one embodiment, the audio codec can be used by an application service provider to provide encoding/decoding services over a network (e.g., the Internet).

In the above description, for purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the invention. It will be apparent, however, to one skilled in the art that the invention can be practiced without these specific details. In other instances, structures and devices are shown in block diagram form in order to avoid obscuring the invention.

In particular, one skilled in the art will recognize that other architectures and graphics environments may be used, and that the present invention can be implemented using graphics tools and products other than those described above. In particular, the client/server approach is merely one example of an architecture for providing the dashboard functionality of the present invention; one skilled in the art will recognize that other, non-client/server approaches can also be used.

Some portions of the detailed description are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

Industrial Applicability

It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the discussion, it is appreciated that throughout the description, discussions utilizing terms such as "processing" or "computing" or "calculating" or "determining" or "displaying" or the like, refer to the action and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

The present invention also relates to an apparatus for performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general-purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but is not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, and magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions, and each coupled to a computer system bus.

The algorithms and modules presented herein are not inherently related to any particular computer or other apparatus. Various general-purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized apparatuses to perform the method steps. The required structure for a variety of these systems will appear from the description below. In addition, the present invention is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the invention as described herein. Furthermore, as will be apparent to one of ordinary skill in the relevant art, the modules, features, attributes, methodologies, and other aspects of the invention can be implemented as software, hardware, firmware or any combination of the three. Of course, wherever a component of the present invention is implemented as software, the component can be implemented as a standalone program, as part of a larger program, as a plurality of separate programs, as a statically or dynamically linked library, as a kernel loadable module, as a device driver, and/or in every and any other way known now or in the future to those of skill in the art of computer programming. Additionally, the present invention is in no way limited to implementation in any specific operating system or environment.

It will be apparent to those skilled in the art that various modifications and variations can be made to the disclosed embodiments without departing from the spirit or scope of the invention. Thus, it is intended that the present invention covers all such modifications to and variations of the disclosed embodiments, provided such modifications and variations are within the scope of the appended claims and their equivalents.

CLAIMS

WHAT IS CLAIMED IS:

1. A method of encoding an audio signal, comprising:

generating a parameter set corresponding to first or second information of an audio signal; and

inserting the parameter set and corresponding first or second information in a bitstream representing the audio signal, wherein the first or second information is represented by a variable number of bits.

2. A method of decoding an audio signal, comprising:

receiving a bitstream representing an audio signal, the bitstream including a parameter set and first or second information corresponding to the parameter set, where the first or second information is represented in the bitstream by a variable number of bits; and

decoding the audio signal based on the parameter set and the corresponding first or second information.

3. The method of claim 2, wherein the first information is position information of a time domain or a time slot to which the parameter set is applied, and the second information is position information of a frequency domain, a sub-band or a parameter band to which the parameter set is applied.

4. The method of claim 3, wherein the first or second information indicates a fixed position or a variable position.

5. The method of claim 4, wherein the first information is represented by a variable number of bits, the variable number of bits being determined by using number information of the time domain or the time slot, and the second information is represented by a variable number of bits, the variable number of bits being determined by using number information of the frequency domain, the sub-band or the parameter band.

6. The method of claim 5, wherein, when the number information is equal to or greater than 2^(n-1) and less than 2^(n), the variable number of bits is decided as n bits.

7. The method of claim 5, wherein, when the number information is greater than 2^(n-1) and equal to or less than 2^(n), the variable number of bits is decided as n bits.

8. The method of claim 2, wherein the first information is represented as a previous value and a difference value, wherein the previous value indicates position information to which a first parameter set is applied and the difference

value indicates position information to which a second parameter set is applied.

9. An apparatus for encoding an audio signal, comprising an encoder configured for:

generating a parameter set corresponding to first or second information of an audio signal; and

inserting the parameter set and corresponding first or second information in a bitstream representing the audio signal, wherein the first or second information is represented by a variable number of bits.

10. An apparatus for decoding an audio signal comprising a decoder configured for:

determining a parameter set corresponding to first information or second information of an audio signal, where the parameter set and the corresponding first or second information is included in a bitstream representing an audio signal, and wherein the first or second information is represented in the bitstream by a variable number of bits; and

decoding the audio signal based on the parameter set and the corresponding first or second information.

11. A data structure for inclusion in a bitstream representing an audio signal, the data structure comprising:

a first field including first or second information; and

a second field including a parameter set corresponding to the first or second information, wherein the first or second information is represented by a variable number of bits.

12. A computer-readable medium having instructions stored thereon, which, when executed by a processor, causes the processor to perform the operations of:

determining a parameter set corresponding to first information or second information of an audio signal, where the parameter set and the corresponding first or second information is included in a bitstream representing an audio signal, and wherein the first or second information is represented in the bitstream by a variable number of bits; and

decoding the audio signal based on the parameter set and the corresponding first or second information.

13. A system, comprising:

a processor; and

a computer-readable medium coupled to the processor and having instructions stored thereon, which, when executed by the processor, causes the processor to perform the operations of:

determining a parameter set corresponding to first information or second information of an audio signal, where the parameter set and the corresponding first or second information is included in a bitstream representing an audio signal, and wherein the first or second information is represented in the bitstream by a variable number of bits; and

decoding the audio signal based on the parameter set and the corresponding first or second information.

14. A system, comprising:

means for determining a parameter set corresponding to first information or second information of an audio signal, where the parameter set and the corresponding first or second information is included in a bitstream representing an audio signal, and wherein the first or second information is represented in the bitstream by a variable number of bits; and

means for decoding the audio signal based on the parameter set and the corresponding first or second information.

FIG. 1

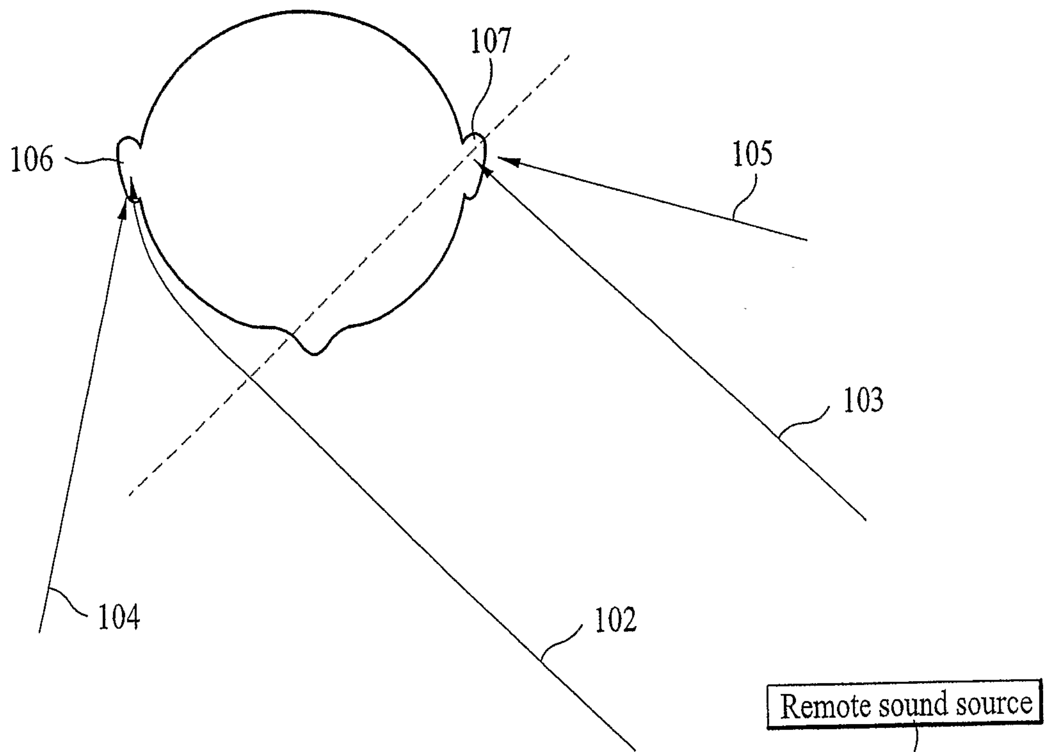


FIG. 2

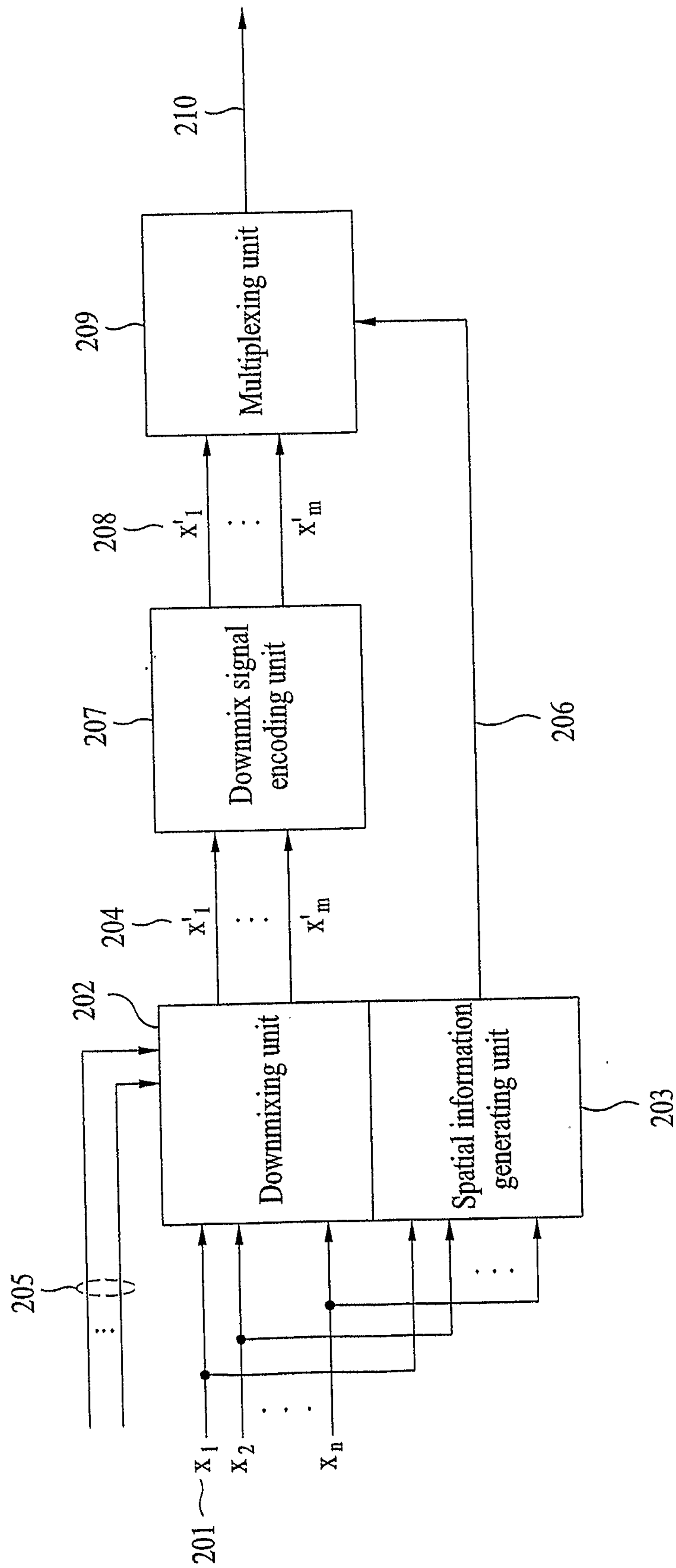

Multiplexing unit
Downmix signal encoding unit
Downmixing unit
Spatial information generating unit
201
202
203
204
205
206
207
208
209
210
x1
x2
xn
x'1
x'm

FIG. 3

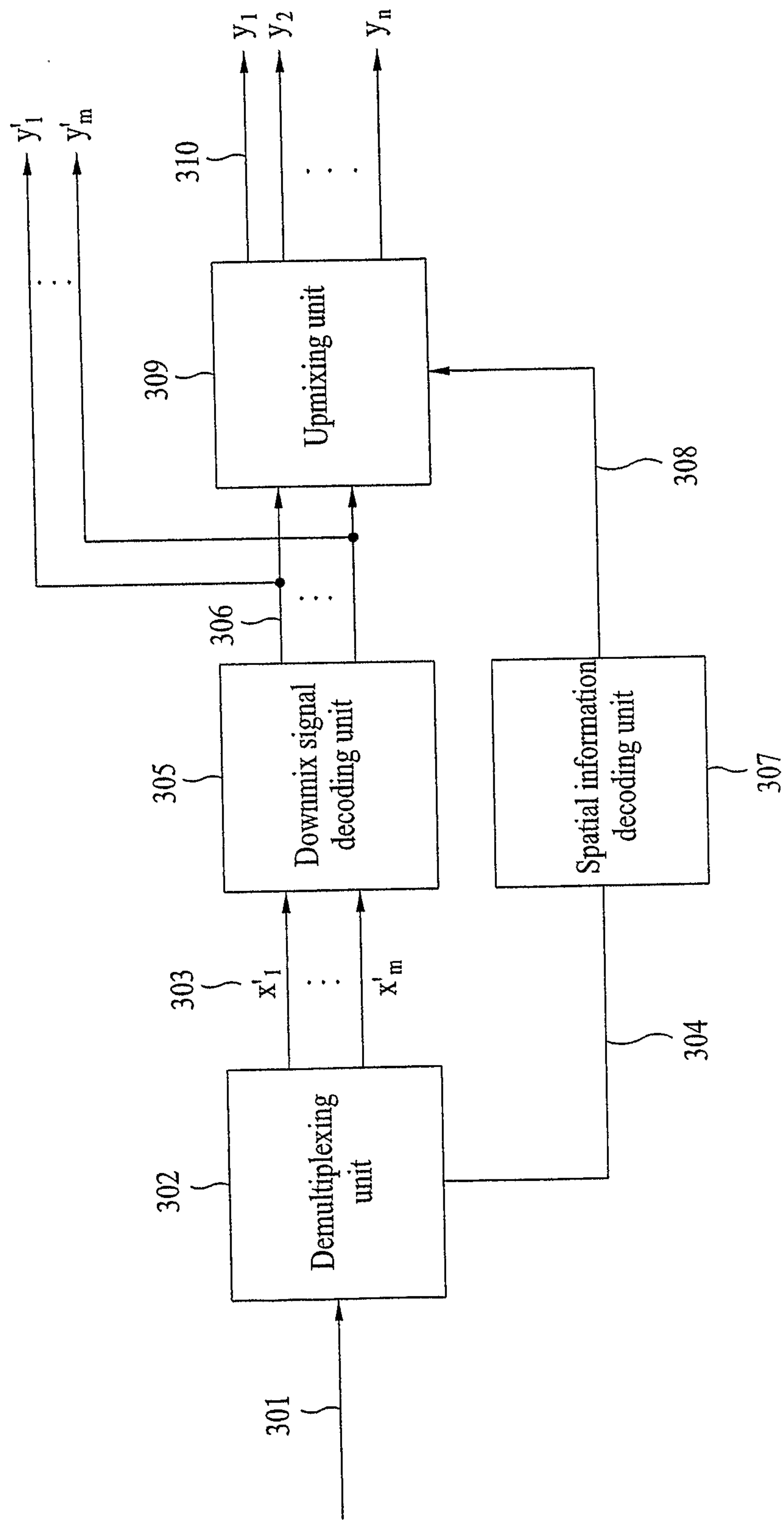

301
Demultiplexing unit
302
303
x'1
x'm
304
Downmix signal decoding unit
305
Spatial information decoding unit
307
306
308
Upmixing unit
309
310
y'1
y'm
y1
y2
yn

FIG. 4

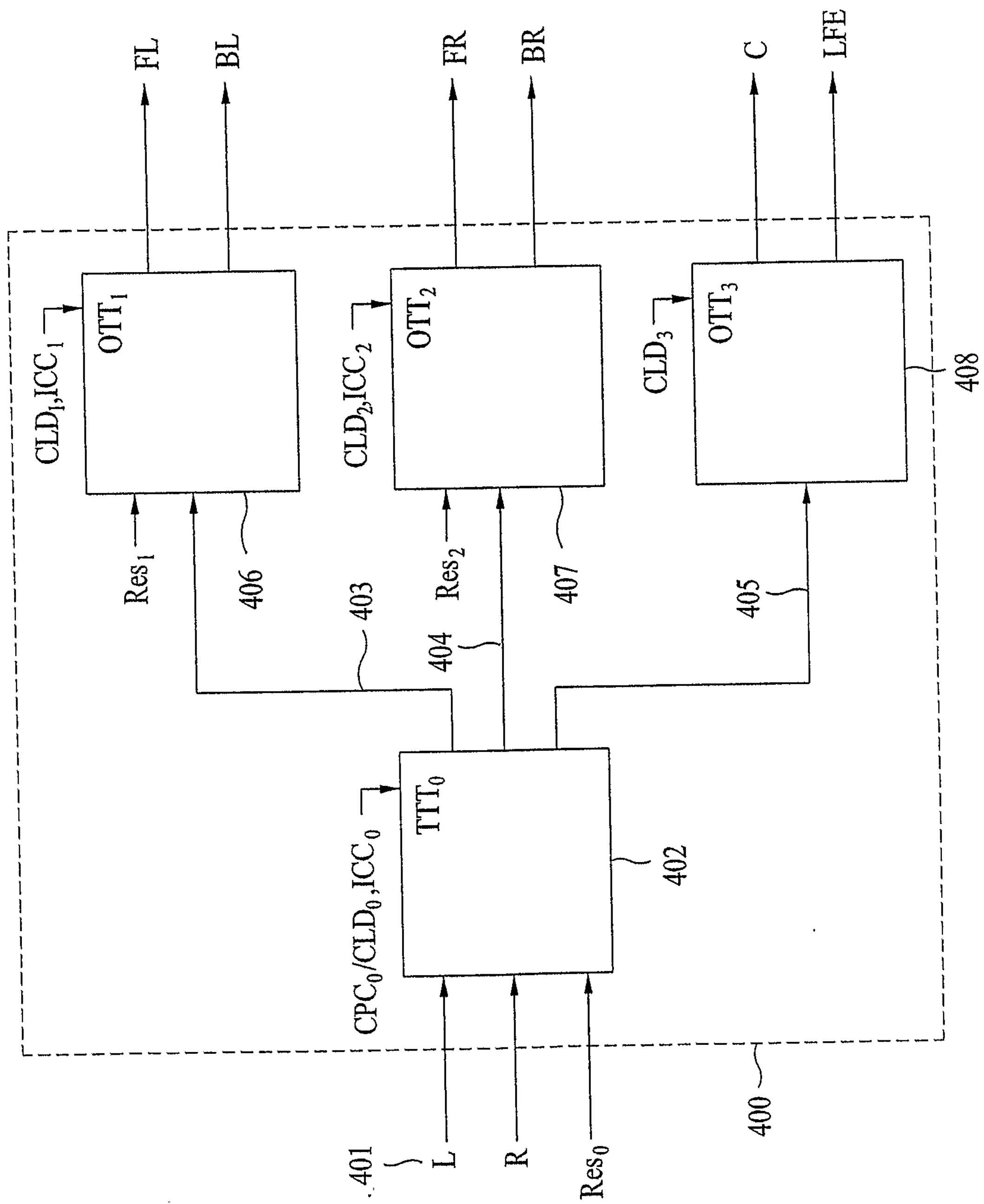

$CPC_0/CLD_0,ICC_0$
TTT_0
402
401
L
R
Res_0
400
403
404
405
CLD_1,ICC_1
OTT_1
Res_1
406
CLD_2,ICC_2
OTT_2
Res_2
407
CLD_3
OTT_3
408
FL
BL
FR
BR
C
LFE

FIG. 5

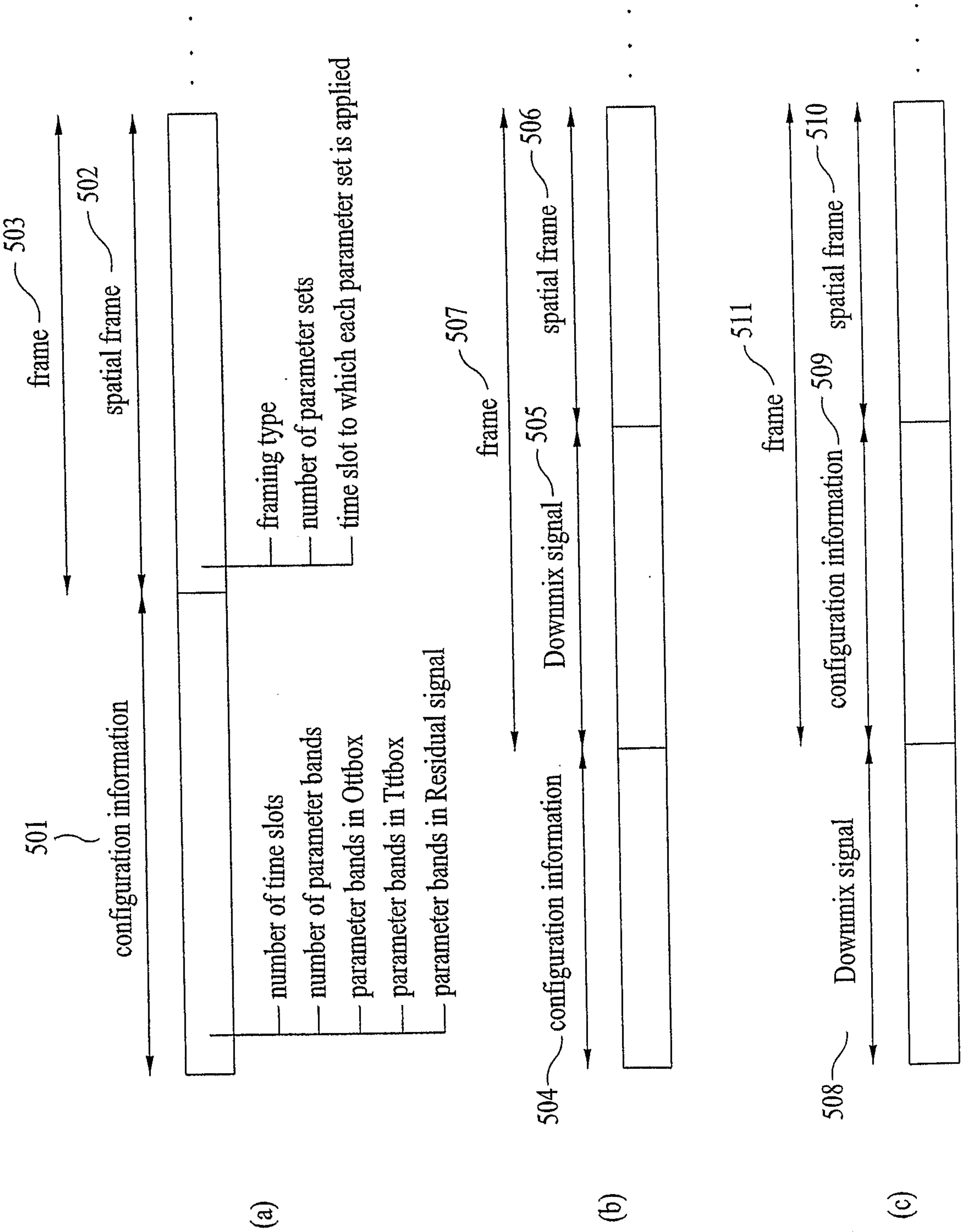

503
frame
502
spatial frame
501
configuration information
framing type
number of parameter sets
time slot to which each parameter set is applied
number of time slots
number of parameter bands
parameter bands in Ottbox
parameter bands in Tttbox
parameter bands in Residual signal
(a)
507
frame
506
spatial frame
505
Downmix signal
504
configuration information
(b)
511
frame
510
spatial frame
509
configuration information
508
Downmix signal
(c)

FIG. 6A

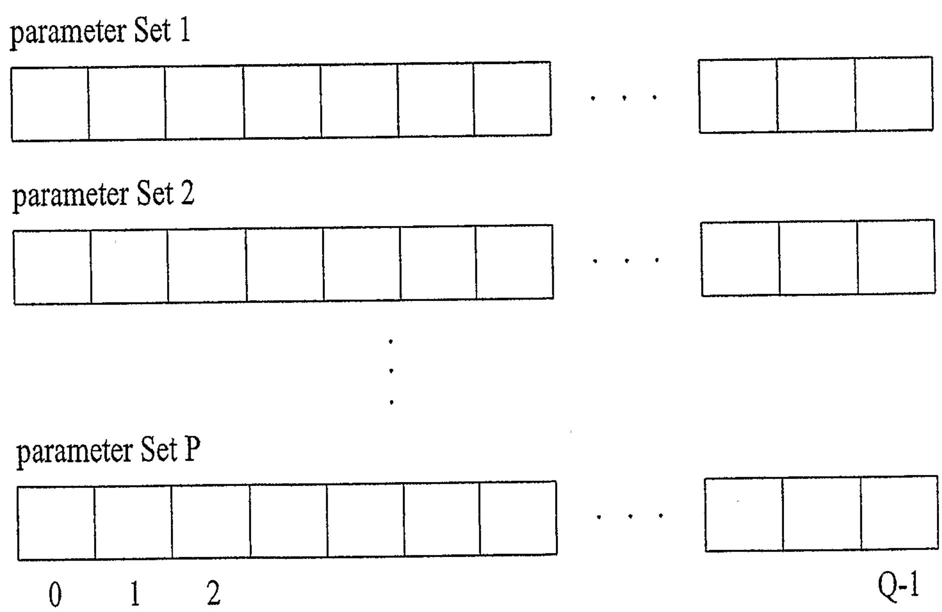


FIG. 6B

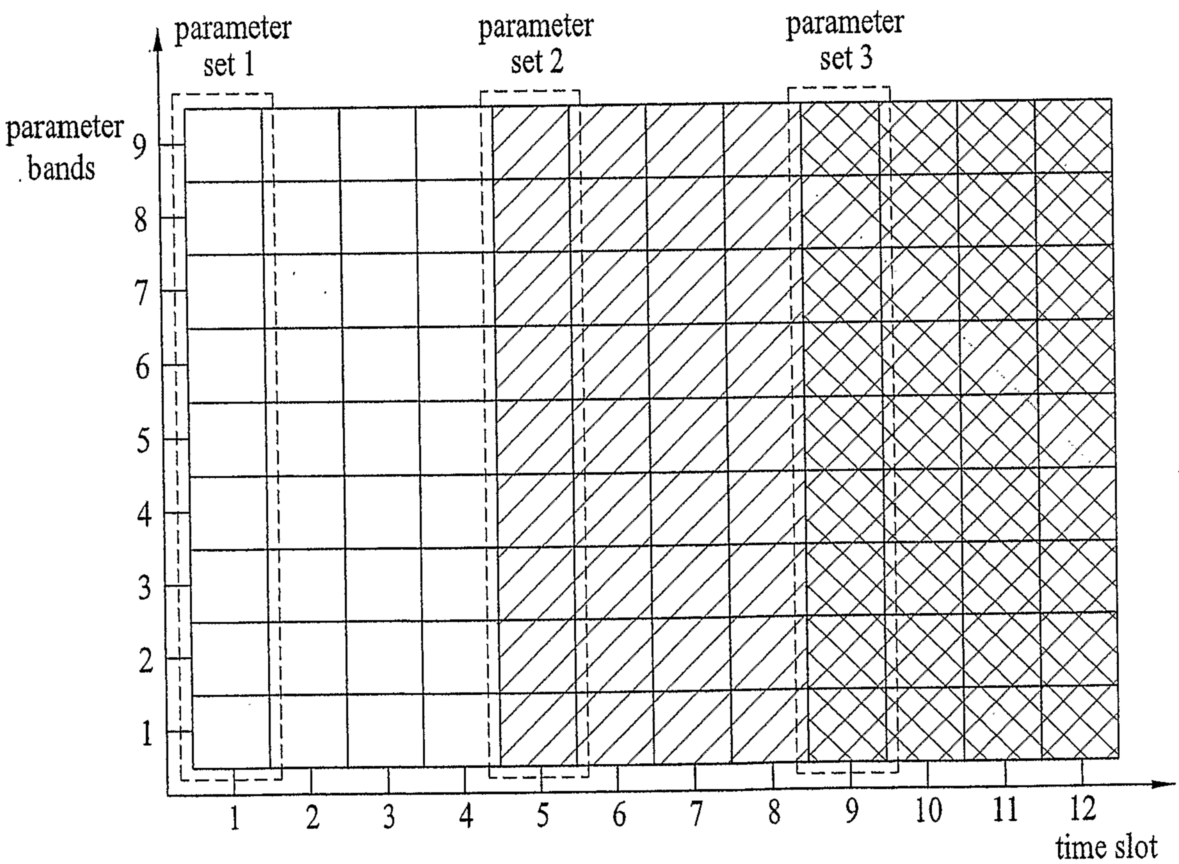

parameter set 1
parameter set 2
parameter set 3
parameter bands
9
8
7
6
5
4
3
2
1
1
2
3
4
5
6
7
8
9
10
11
12
time slot

FIG. 7A

	Syntax	No. of bits
	SpatialSpecificConfig()	
	{	
701	bsSamplingFrequencyIndex;	4
	if (bsSamplingFrequencyIndex == 0xf) {	
702	bsSamplingFrequency;	24
	}	
703	bsFrameLength;	7
704	bsFreqRes;	3
705	bsTreeConfig;	4
706	bsQuantMode;	3
707	bsOneIcc;	1
708	bsArbitraryDowmix;	1
709	bsFixedGainSur;	3
710	bsFixedGainLFE;	3
711	bsFixedGainDMX;	3
712	bsMatrixMode;	1
713	bsTempShapeConfig;	4
714	bsDecorrConfig;	4
715	bs3DaudioMode;	1
	⋮	
	for (i=0; i<numOttBoxes; i++) {	
716	OttConfig(i);	
	}	
	for (i=0; i<numTttBoxes; i++) {	
717	TttConfig(i);	
	}	
	⋮	
718	Spatial Extension Config()	
	}	

FIG. 7B

bsFreqRes	numBands
0	Reserved
1	28
2	20
3	14
4	10
5	7
6	5
7	4

FIG. 8A

Syntax	No. of bits
OttConfig(i)	
{	
if (ottModeLfe[i]) {	
801 bsOttBands[i];	5
}	
else {	
bsOttBands[i] = numBands;	
}	
}	

FIG. 8B

Syntax	No. of bits
OttConfig(i)	
{	
if (ottModeLfe[i]) {	
802 bsOttBands[i];	Bitsnumbands
}	
else {	
bsOttBands[i] = numBands;	
}	
}	

Minimum number of bits for representation of numBands

FIG. 9A

	Syntax	No. of bits
	TttConfig(i)	
	{	
901	bsTttDualMode[i];	1
902	bsTttModeLow[i];	3
	if (bsTttDualMode[i]) {	
903	bsTttModeHigh[i];	3
904	bsTttBandsLow[i];	5
905	bsTttBandsHigh[i] = numBands;	
	}	
	else {	
906	bsTttBandsLow[i] = numBands;	
	}	

FIG. 9B

Syntax	No. of bits
TttConfig(i)	
{	
bsTttDualMode[i];	1
bsTttModeLow[i];	3
if (bsTttDualMode[i]) {	
bsTttModeHigh[i];	3
907 bsTttBandsLow[i];	BitsnumBands
bsTttBandsHigh[i] = numBands;	
}	
else {	
bsTttBandsLow[i] = numBands;	
}	

Minimum number of bits for representation of numBands

FIG. 10A

	Syntax	No. of bits
	SpatialExtensionConfig()	
	{	
	while (BitsAvailable() >= 8) {	
1001	bsSacExtType;	
1002	cnt = bsSacExtLen;	4
	if (cnt==15) {	4
1003	cnt += bsSacExtLenAdd;	8
	}	
	if (cnt==15+255) {	
1004	cnt += bsSacExtLenAddAdd;	16
	}	
1005	bitsRead = SpatialExtensionConfigData(bsSacExtType)	
1006	nFillBits = 8*cnt-bitsRead;	
1007	bsFillBits;	
	}	
	}	
	}	

FIG. 10B

	Syntax	No. of bits
	SpatialExtensionConfigData(1)	
	{	
1008	bsResidualSamplingFrequencyIndex;	4
1009	bsResidualFramesPerSpatialFrame;	2
	for (i=0; i<numOttBoxes+numTttBoxes; i++) {	
1010	ResidualConfig(i);	
	}	
	}	

FIG. 10C

	Syntax	No. of bits
	ResidualConfig(i)	
	{	
1011	bsResidualPresent[i];	1
	if (bsResidualPresent[i]) {	
1012	bsResidualBands[i];	5
	}	
	}	

FIG. 10D

	Syntax	No. of bits
	ResidualConfig(i)	
	{	
1013	bsResidualPresent[i];	1
	if (bsResidualPresent[i]) {	
1014	bsResidualBands[i];	BitsnumBands
	}	
	}	

Minimum number of bits for representation of numBands

FIG. 11A

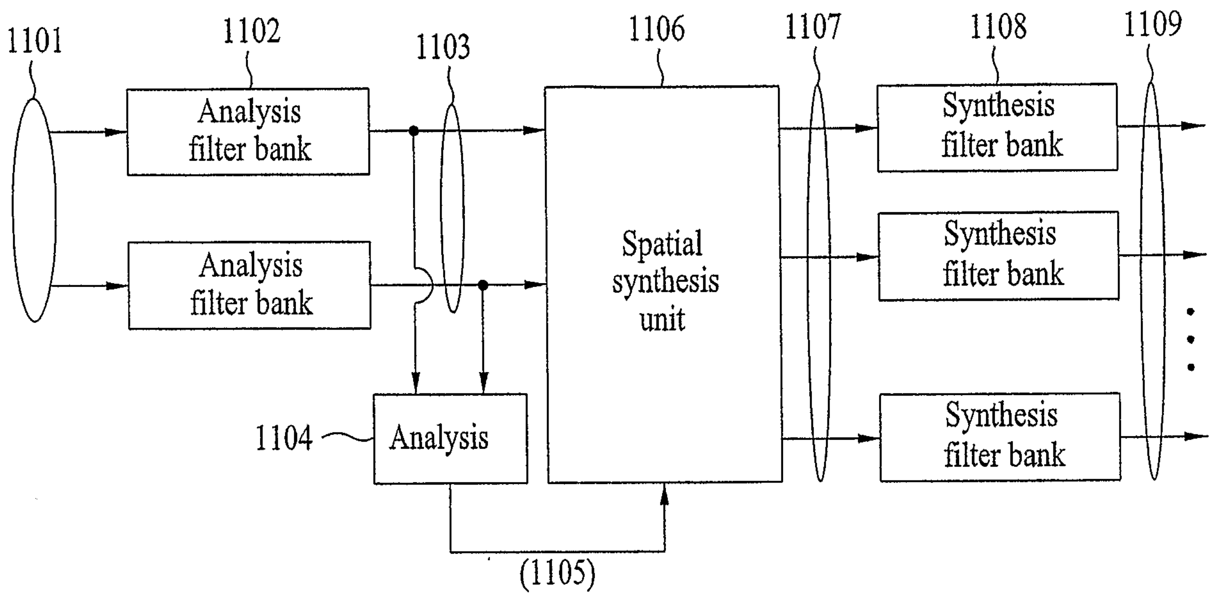


FIG. 11B

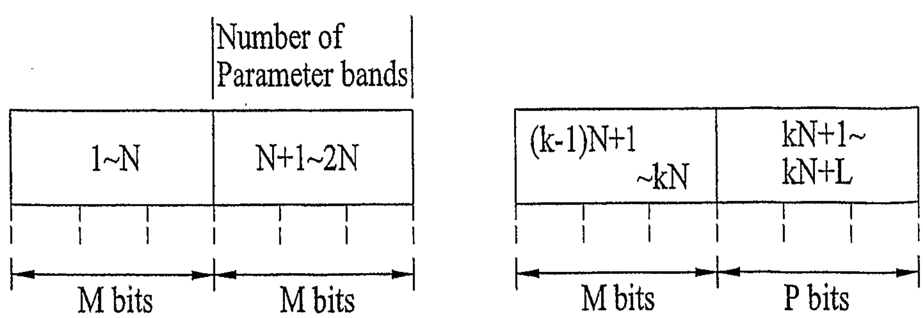


FIG. 12

	Syntax	No. of bits
	SpatialFrame()	
	{	
1201	FramingInfo();	
1202	bsIndependencyFlag;	1
1203	OttData();	
1204	TttData();	
1205	SmgData();	
1206	TempShapeData();	
	}	

FIG. 13A

	Syntax	No. of bits
	FramingInfo()	
	{	
1301	bsFramingType;	1
1302	bsNumParamSets;	3
	if (bsFramingType) {	
	for (ps=0; ps<numParamSets; ps++) {	
1303	bsDiffParamSlot[ps];	BitsnumSlots
	}	
	}	
	}	

BitsnumSlots: Minimum number of bits for representation of numSlots

FIG. 13B

	Syntax	No. of bits
	FramingInfo()	
	{	
	bsFramingType;	
	bsNumParamSets;	1
	if (bsFramingType) {	3
	for (ps=0; ps<numParamSets; ps++) {	
	if(ps==0){	
1304	bsParamSlot[0];	nBitsParamSlot(0)
	else{	
1305	bsDiffParamSlot[ps];	nBitsParamSlot(ps)
1306	bsParamSlot[ps] = bsParamSlot[ps-1] + bsDiffParamSlot[ps] + 1;	
	}	
	}	
	}	
	}	

FIG. 13C

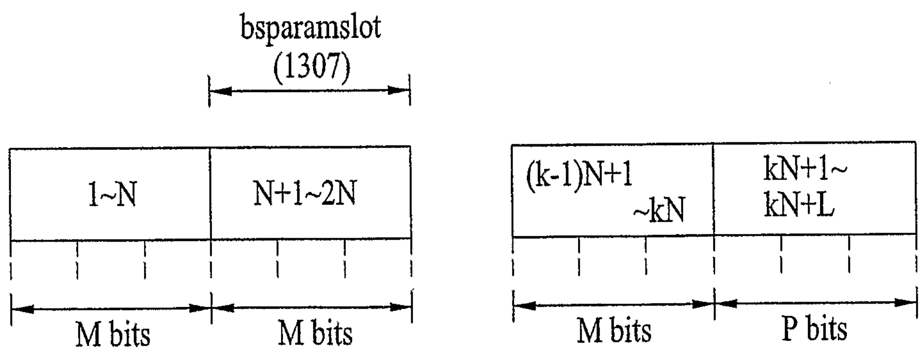

bsparamslot
(1307)
1~N
N+1~2N
(k-1)N+1
~kN
kN+1~
kN+L
M bits
M bits
M bits
P bits

FIG. 14

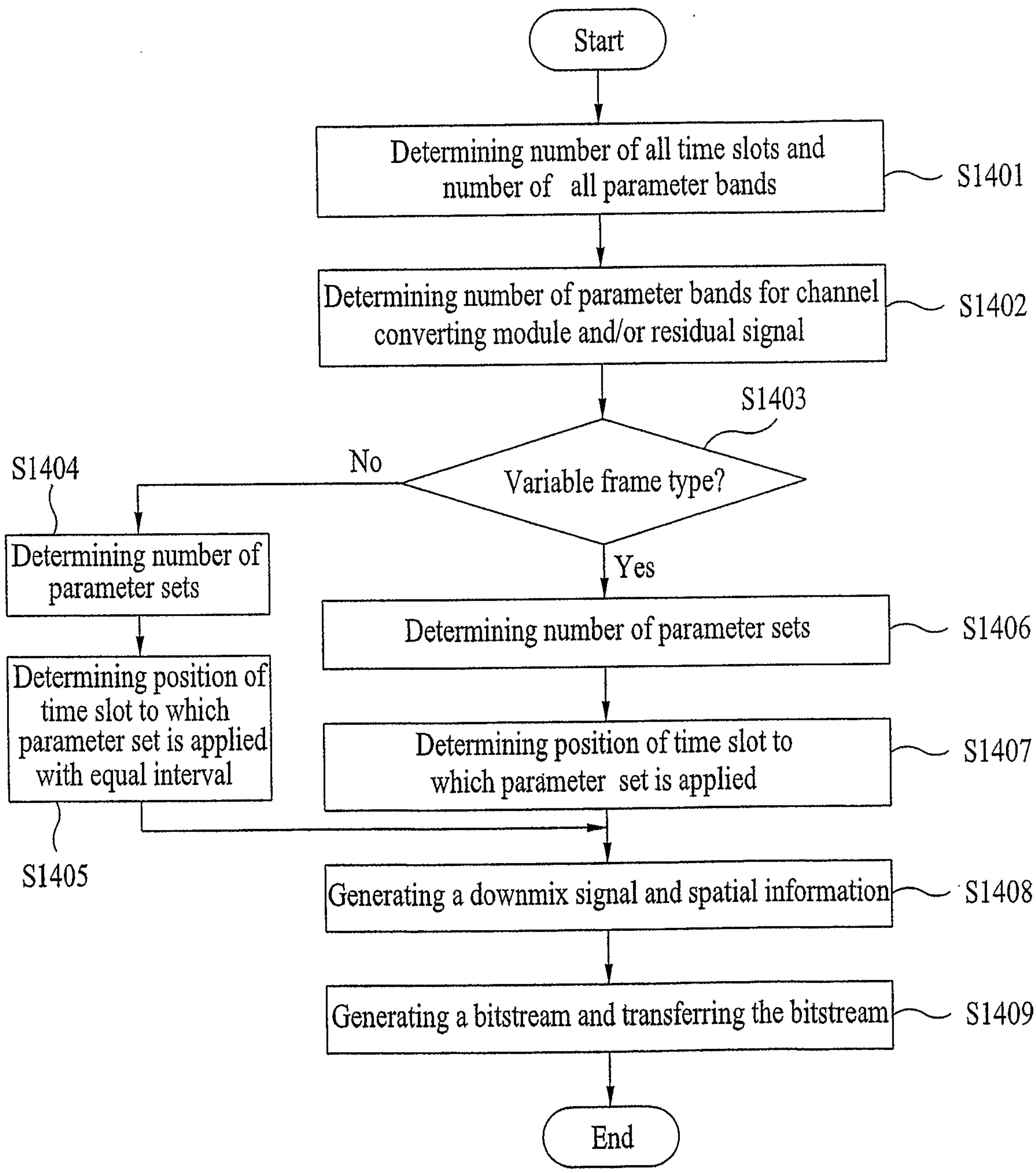

Start
Determining number of all time slots and number of all parameter bands
S1401
Determining number of parameter bands for channel converting module and/or residual signal
S1402
S1403
Variable frame type?
No
Yes
S1404
Determining number of parameter sets
Determining position of time slot to which parameter set is applied with equal interval
S1405
Determining number of parameter sets
S1406
Determining position of time slot to which parameter set is applied
S1407
Generating a downmix signal and spatial information
S1408
Generating a bitstream and transferring the bitstream
S1409
End

FIG. 15

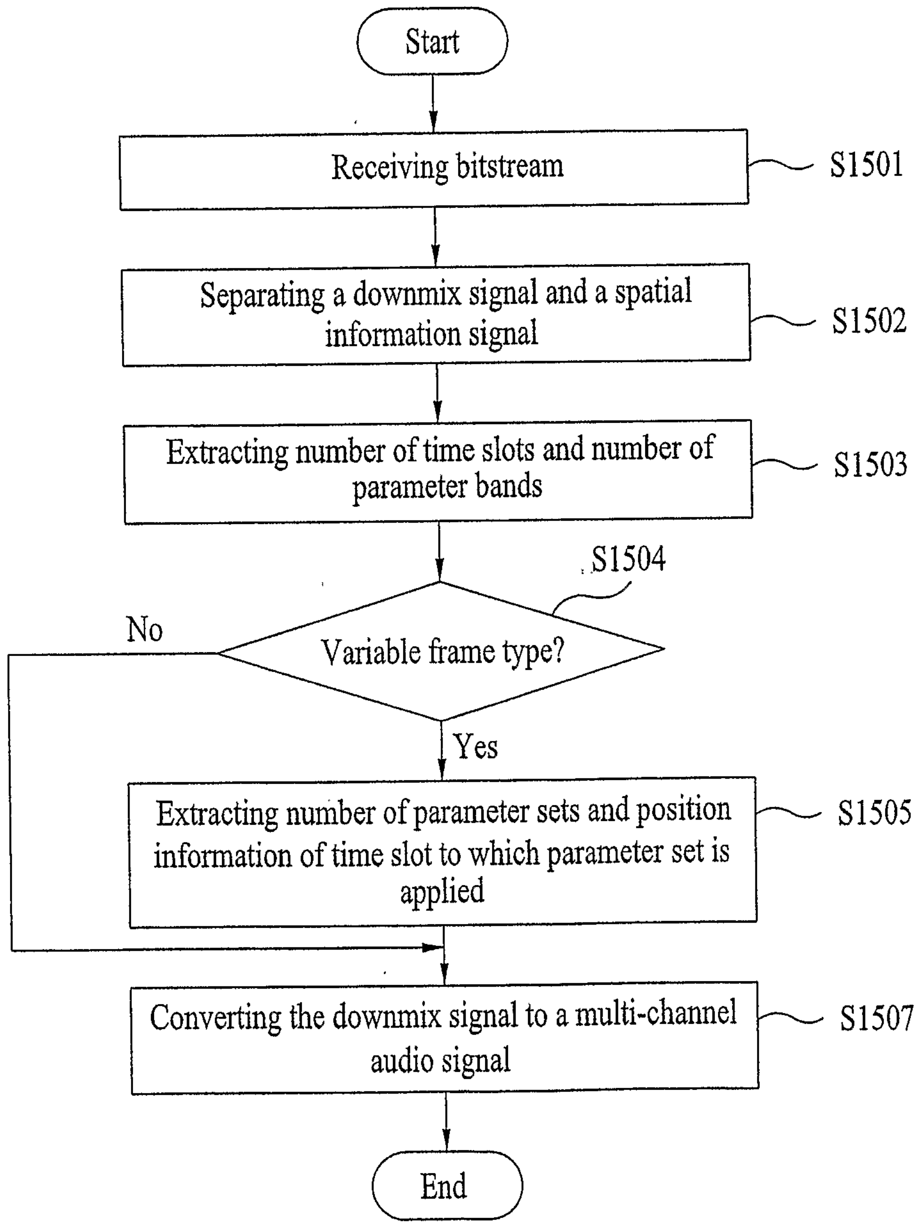
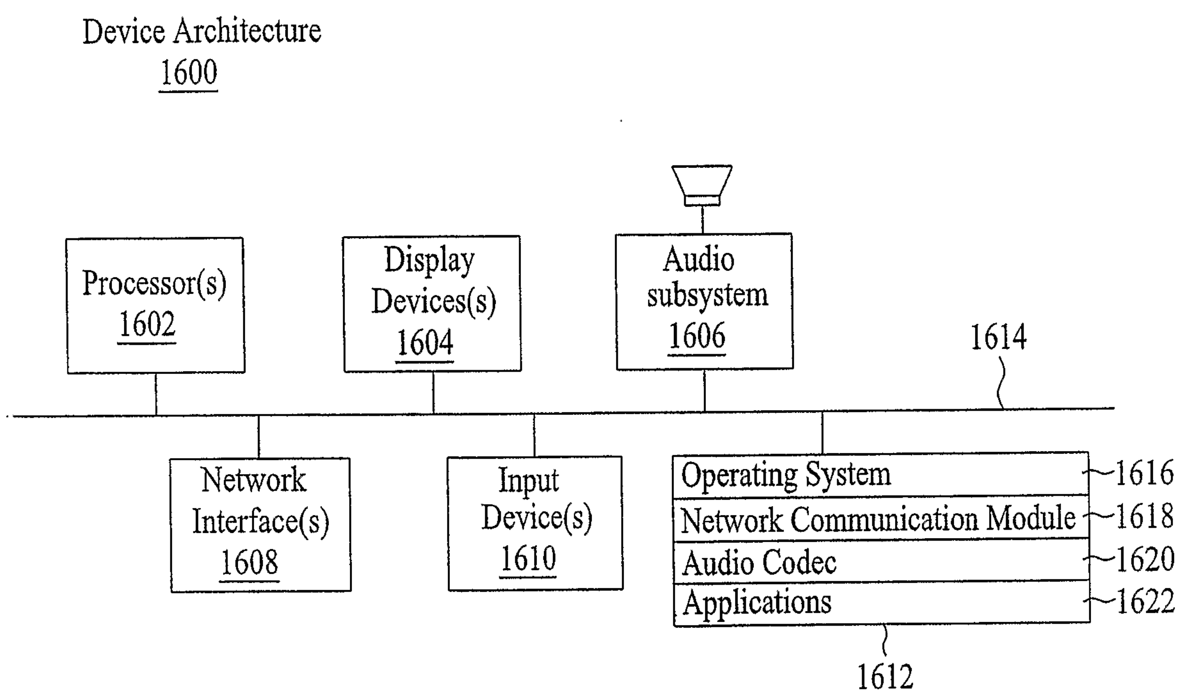

Start
Receiving bitstream
S1501
Separating a downmix signal and a spatial information signal
S1502
Extracting number of time slots and number of parameter bands
S1503
S1504
Variable frame type?
No
Yes
Extracting number of parameter sets and position information of time slot to which parameter set is applied
S1505
Converting the downmix signal to a multi-channel audio signal
S1507
End

FIG. 16


Device Architecture
1600
Processor(s)
1602
Display
Devices(s)
1604
Audio
subsystem
1606
1614
Network
Interface(s)
1608
Input
Device(s)
1610
Operating System
1616
Network Communication Module
1618
Audio Codec
1620
Applications
1622
1612

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR 2006/003424

A. CLASSIFICATION OF SUBJECT MATTER

IPC[8]: ***H04S 3/00*** (2007.01); ***G10L 19/00*** (2007.01); ***G10L 19/02*** (2007.01)

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)

IPC[8]: G10L, H04S

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 practicable, search terms used)

WPI, EPODOC, Elsevier, IEE, IEEEXplore, Research Disclosure, IBM TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5481643 A (TEN KATE et al.) 2 January 1996 (02.01.1996) *Figures; Abstract; Column 1, Line 14 - Column 5, Line 54; Claims* --	1-14
X	Hamdy K.N.; Ali M.; Tewfik A.H.; "Low bit rate high quality audio coding with combined harmonic and wavelet representations." In: IEEE International Conference on Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings. Volume 2, 7-10 May 1996. Pages: 1045-1048 --	1-14
X	US 5682461 A (SILZLE et al.) 28 October 1997 (28.10.1997) *Figures; Abstract; Column 1, Line 6 - Column 3, Line 59; Claims* --	1-14

☒ 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 application or patent 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 17 January 2007 (17.01.2007)	Date of mailing of the international search report 31 January 2007 (31.01.2007)
Name and mailing address of the ISA/ AT **Austrian Patent Office** Dresdner Straße 87, A-1200 Vienna Facsimile No. +43 / 1 / 534 24 / 535	Authorized officer MESA PASCASIO J. Telephone No. +43 / 1 / 534 24 / 327

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6021386 A (DAVIS et al.) 1 February 2000 (01.02.2000) *Figures; Abstract; Column 1, Line 33 - Column 11, Line 12* --	1-14
X	Webb J.; Lueck J.; "Video and audio coding for mobile applications." In: The application of programmable DSPs in mobile communications. Edited by A. Gatherer and E. Auslander. ISBN 0-471-48643-4. John Wiley & Sons Ltd., 2002. Pages: 190-199 --	1-14
X	US 2003/195742 A1 (TSUSHIMA et al.) 16 October 2003 (16.10.2003) *Figures; Abstract; Paragraphs 2-11; Claims* --	1-14
X	Boltze Th. et al.; "Audio services and applications." In: Digital Audio Broadcasting. Edited by Hoeg, W. and Lauterbach, Th. ISBN 0-470-85013-2. John Wiley & Sons Ltd., 2003. Pages: 75-83 --	1-14
X	Faller Ch.; "Parametric coding of spatial audio." In: Proceedings of the 7th International Conference on Digital Audio Effects (DAFx'04). Invited Paper. Naples, Italy. 5-8 October 2004. Pages: 151-156 ----	1-14

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2006/003424

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US A 5481643		US A 5481643	1996-01-02
US A 568261		US A 568261	0000-00-00
US A 6021386		JP A 2003099099	2003-04-04
		US A 6021386	2000-02-01
		US A 5909664	1999-06-01
		US A 5632005	1997-05-20
		US A 5633981	1997-05-27
		US A 5608805	1997-03-04
US A 2003195742		DE D1 60307252D	2006-09-14
		US A1 2003195742	2003-10-16
		WO A1 03085644	2003-10-16
		EP A1 1493146	2005-01-05
		CN A 1516865	2004-07-28
		JP A 2004004710	2004-01-08