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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 October 2007 (11.10.2007)

PCT

(10) International Publication Number
WO 2007/114594 A1

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

(21) International Application Number:
PCT/KR2007/001560

(22) International Filing Date: 30 March 2007 (30.03.2007)

(25) Filing Language: Korean

(26) Publication Language: English

(30) Priority Data:
60/787,172 30 March 2006 (30.03.2006) US
60/787,516 31 March 2006 (31.03.2006) 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*): **OH, Hyen O** [KR/KR]; 306-403, Hansin Apt., Gangseon-maeul 3-Danji, Juyeop 1-dong, Ilsan-gu, Goyang-si, Gyeonggi-do 411-744 (KR). **PANG, Hee Suck** [KR/KR]; 101, 4/7, #14-10, Yangjae-dong, Seocho-gu, Seoul 137-130 (KR). **KIM, Dong Soo** [KR/KR]; 502, Woolim Villa, #602-265, Namhyeon-dong, Gwanak-gu, Seoul 151-801 (KR). **LIM, Jae Hyun** [KR/KR]; 609, Parkville officetel, #1062-20, Namhyeon-dong, Gwanak-gu, Seoul 151-801 (KR). **JUNG, Yang Won** [KR/KR]; 2-803, Yeoksam Hanshin Apt., Dogok-dong, Gangnam-gu, Seoul 135-270 (KR).

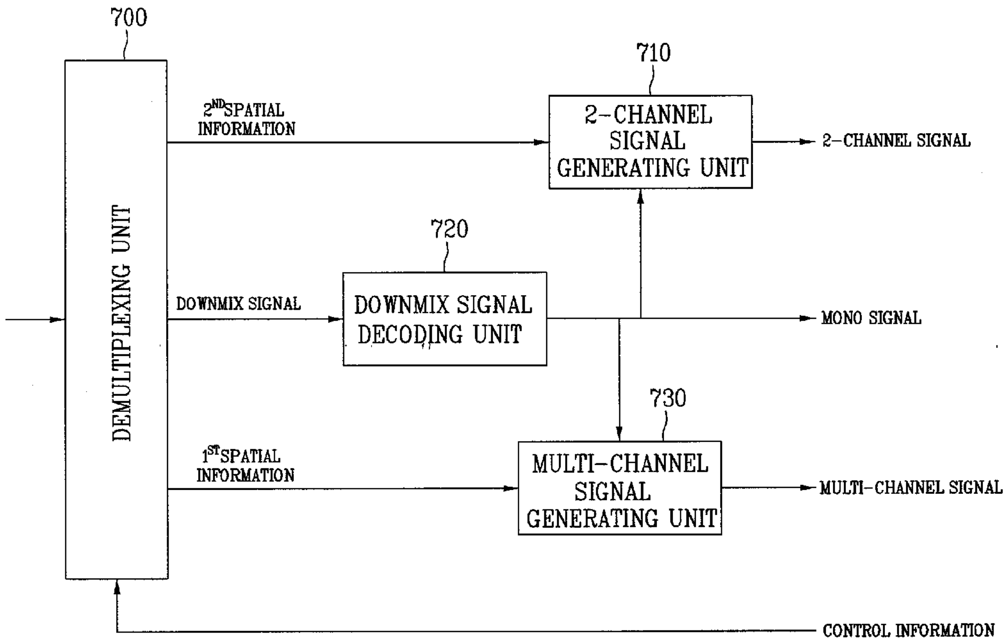
(74) Agents: **KIM, Yong In** et al.; 15th Floor, Yo Sam Building, 648-23, Yeoksam-dong, Kangnam-gu, Seoul 135-808 (KR).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, 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, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, 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.

(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, MT, NL, PL,

[Continued on next page]

(54) Title: APPARATUS FOR PROCESSING MEDIA SIGNAL AND METHOD THEREOF



(57) **Abstract:** The present invention relates to an apparatus for processing a media signal and method thereof. A method of processing a media signal according to the present invention includes extracting a downmix signal from a bitstream, extracting at least one of first spatial information and second spatial information from the bitstream, and generating multi-channels using the extracted spatial information and the downmix signal. And, the present invention provides a decoding method and apparatus for generating various kinds of multi-channels.

WO 2007/114594 A1

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.

Apparatus for Processing Media Signal and Method Thereof

TECHNICAL FIELD

The present invention relates to an apparatus for processing a media signal and method thereof.

BACKGROUND ART

In the present invention, media signals include an audio signal and a video signal. And, the audio signal is explained as an example in the following description.

Currently, 2-channel signal is most frequently generated and user. Yet, the use of multi-channel signals gradually increases. In the /following description, an audio signal including at least three channels is called a multi-channel signal to be discriminated from the 2-channel signal. In general, an encoder compresses a multi-channel signal into a mono- or stereo-type downmix signal instead of compressing channels of the multi-channel signal individually. A downmixing unit of the encoder extracts spatial information by downmixing multi-channels. The encoder transfers the compressed downmix signal and the spatial information to a decoder or stores them in a storage medium. The spatial information is used in reconstructing an original multi-channel signal from the compressed downmix signal. In case of using an encoder and decoder for 2-channel signal compression and reconstruction, the encoder generates a downmix signal and spatial information from a 2-channel signal and then transfers a bitstream including them to the decoder. The decoder upmixes the transferred bitstream to generate the original 2-channel signal. In case that the encoder and decoder are used for compression and reconstruction of a multi-channel signal, the encoder generates a downmix signal and spatial information from the multi-channel signal and then transfers a bitstream including the downmix signal and spatial information to the decoder. The decoder then upmixes the transferred bitstream to generate the original multi-channel signal.

DISCLOSURE OF THE INVENTION

TECHNICAL OBJECTS

Accordingly, the present invention is directed to an apparatus for processing a media signal and method thereof that substantially obviate one or more of the problems due to limitations and disadvantages of the related art.

An object of the present invention is to provide an encoding method and apparatus, by which spatial information for audio signal reconstruction having an audio quality close to an audio signal prior to downmixing can be generated.

Another object of the present invention is to provide an encoding method and apparatus, by which a bitstream including both spatial information used in generating a 2-channel signal and spatial information used in generating a multi-channel signal can be provided and generated.

Another object of the present invention is to provide a decoding method and apparatus, by which a 2-channel signal or a multi-channel signal can be selectively generated.

TECHNICAL SOLUTION

The present invention extracts a downmix signal from a bitstream and also extracts at least one of first spatial information and second spatial information from the bitstream. And, the present invention provides a method and apparatus for generating specific multi-channels using the extracted spatial information and the extracted downmix signal.

ADVANTAGEOUS EFFECTS

The present invention can provide an encoding method and apparatus for generating spatial information to reconstruct an audio signal having an audio quality close to a former audio signal prior to downmixing.

The present invention can provide a bitstream including both spatial information used in generating a 2-channel signal and spatial information used in generating a multi-channel signal. And, the present invention can provide an encoding method and apparatus for generating the bitstream.

And, the present invention can provide a decoding method and apparatus capable of generating a 2-channel signal or a multi-channel signal selectively.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram of a first encoding apparatus according to one embodiment of the present invention.

FIG. 2 is a block diagram of a second encoding apparatus according to another embodiment of the present invention.

FIG. 3 is a block diagram of a third encoding apparatus for generating spatial information using a decoded downmix signal according to one embodiment of the present invention.

FIG. 4 is a block diagram of a fourth encoding apparatus for generating spatial information using a decoded downmix signal according to another embodiment of the present invention.

FIG. 5 is a diagram of a bitstream of an audio signal according to one embodiment of the present invention.

FIG. 6 is a block diagram of a first decoding apparatus according to one embodiment of the present invention.

FIG. 7 is a block diagram of a second encoding apparatus according to another embodiment of the present invention.

BEST MODE FOR CARRYING OUT THE INVENTION

To achieve these and other advantages and in accordance with the purpose of the present invention, as embodied and broadly described, a method of processing a media signal includes extracting a downmix signal from a bitstream, extracting at least one of first spatial information and second spatial information from the bitstream, and generating multi-channels using the extracted spatial information and the downmix signal.

To further achieve these and other advantages and in accordance with the purpose of the present invention, a method of processing a media signal includes generating a first downmix signal from multi-channels, generating a second downmix signal from the first downmix signal, generating first spatial information using the multi-channels and the first downmix signal or the multi-channels and the second downmix signal, generating second spatial information using the first downmix signal and the second downmix signal, and generating a bitstream including the first spatial information and the second spatial information.

To further achieve these and other advantages and in accordance with the purpose of the present invention, a method of processing a media signal includes generating a first downmix signal from multi-channels, generating a second downmix signal from the first downmix signal, encoding the second downmix signal, decoding the encoded second downmix signal, generating second spatial information using the first downmix signal and the decoded second downmix signal, and generating first spatial information using the multi-channels and the decoded second downmix signal.

To further achieve these and other advantages and in accordance with the purpose of the present invention, a method of processing a media signal includes generating a first downmix signal from multi-channels, generating a second downmix signal from the first downmix signal, encoding the second downmix signal, decoding the encoded second downmix signal, generating second spatial information using the first downmix signal and the decoded second downmix signal, generating a modified first downmix signal using the decoded second downmix signal and the second spatial information, and generating first spatial information using the modified first downmix signal and the multi-channels.

To further achieve these and other advantages and in accordance with the purpose of the present invention, an apparatus for processing a signal includes a downmix signal extracting unit extracting a downmix signal from a bitstream, an information extracting unit extracting at least one of second spatial information for generating two channels from the downmix signal and first spatial information for generating at least three channels from the downmix signal from the bitstream, and a channel generating unit generating either the two channels or the at least three channels using the extracted information and the downmix signal.

To further achieve these and other advantages and in accordance with the purpose of the present invention, a bitstream structure includes first spatial information extracted in the course of generating a first downmix signal including at least two channels from multi-channels and second spatial information extracted in the course of generating a second downmix signal from the first downmix signal.

To further achieve these and other advantages and in accordance with the purpose of the present invention, a storage medium including the bitstream structure.

To further achieve these and other advantages and in accordance with the purpose of the present invention, a signal processing apparatus includes a first downmixing unit generating a first downmix signal from multi-channels, a second downmixing unit generating a second downmix signal from the first downmix signal, a first spatial information generating unit generating first spatial information using the multi-channels and the first downmix signal or the multi-channels and the second downmix signal, a second spatial information generating unit generating second spatial information using the first downmix signal and the second downmix signal, and a multiplexing unit generating a bitstream including the first spatial information and the second spatial information.

To further achieve these and other advantages and in accordance with the purpose of the present invention, a signal processing apparatus includes a downmixing unit generating a downmix signal from multi-channels, an encoding unit encoding the downmix signal, a decoding unit decoding the encoded downmix signal, and a spatial information generating unit generating spatial information using the multi-channels and the decoded downmix signal.

MODE FOR INVENTION

Reference will now be made in detail to the preferred embodiments of the present invention, examples of which are illustrated in the accompanying drawings. For facilitation in understanding the present invention, an audio signal encoding method and apparatus are explained prior to an audio signal decoding method and apparatus. Yet, the decoding method and apparatus according to the present invention are not limited by an encoding method and apparatus that will be explained in the following description. And, the present invention is applied to a coding scheme for generating two channels using spatial information and a coding scheme for generating multi-channels using spatial information as well as MP3 (MPEG ½-layer III) and AAC (advanced audio coding).

An encoding apparatus for compressing a 2-channel signal receives the 2-channel signal, downmixes the received signal into a moo signal, and extracts spatial information indicating a relation with the 2-channel signal. An encoding apparatus for compressing a multi-channel signal downmixes the multi-channel signal into one or two audio signals and extract information indicating a relation with the multi-channel signal. An encoding apparatus is cable to generates a 2-channel signal by downmixing a multi-channel signal or generates a mono signal by downmixing the 2-channel signal again. In this case, the encoding apparatus extracts spatial information from the relation between the multi-channel signal and the 2-channel signal in downmixing the multi-channel signal into the 2-channel signal or extracts spatial information from the relation between the 2-chanel signal and the mono signal in downmixing the 2-channel signal into the mono signal. An encoding apparatus is able to separately transfer spatial information for reconstructing 2-channel signal and spatial information for reconstructing multi-channel signal to a decoding apparatus. Alternatively, the encoding apparatus generates a bitstream including spatial information for reconstructing 2-channel signal and spatial information for reconstructing multi-channel signal and then transfer the bitstream to the decoding apparatus. In case that a signal the decoding apparatus is able to generate is either the 2-channel signal or the multi-channel signal, the decoding apparatus having received the bitstream including the spatial information for reconstructing the 2-channel signal and the spatial information for reconstructing the multi-channel signal extracts the spatial information for reconstruct the generatable channel signal from the bitstream only and is then able to reconstruct the channel signal using the extracted spatial information. In case that the decoding apparatus is capable of reconstruct both of the 2-channel signal and the multi-channel signal, the decoding apparatus extracts spatial information required for generating a channel signal selected by a user from the bitstream only and is then able to generate the channel signal selected by the user using the extracted spatial information.

An encoding method and apparatus for generating a bitstream including spatial information for reconstructing 2-channel signal and spatial information for reconstructing multi-channel signal are explained with reference to FIG. 1 and FIG. 2 as follows.

FIG. 1 is a block diagram of a first encoding apparatus according to one embodiment of the present invention.

Referring to FIG. 1, a first encoding apparatus includes a first downmixing unit 100, a second downmixing unit 110, a downmix signal encoding unit 120, a first spatial information generating unit 130, a second spatial information generating unit 140, and a multiplexing unit 150.

The first downmixing unit 100 receives a multi-channel signal and then downmixes the received signal into a first downmix signal having channels less than those of the multi-channel signal. And, the second downmixing unit 110 downmixes the first downmix signal into a second downmix signal having channels less than those of the first downmix signal.

Each of the downmixing units 100 and 110 can use an OTT (one-to-two) box or a TTT (two-to-three) box to transform two channels into one channel or transform three channels into two channels. The OTT or TTT box is a conceptional box included in an audio signal decoding apparatus to be used in generating multi-channels using a downmix signal and spatial information. The OTT box transforms one signal into two signals using spatial information. The TTT box transforms two signals into three signals using spatial information. In the following description, the OTT or TTT box is called a signal transforming unit. To correspond the OTT or TTT box used for the audio signal decoding apparatus, an OTT or TTT box is included in the downmixing unit 100 or 110 of the audio signal encoding apparatus to be used in outputting one or two down mix signals from inputted multi-channels.

The first/second downmix signal can be artificially generated instead of being generated by the downmixing unit 100/110. Since the second downmix signal is a signal including channels less than those of the first downmix signal, in case that the second downmix signal is a mono signal, the first downmix signal should include at least two channels. In case that the first downmix signal is a 2-channel signal, the multi-channel signal should include at least three channels.

The downmix signal encoding unit 120 compresses the second downmix signal and then sends the compressed downmix signal to the multiplexing unit 150. The first spatial information generating unit 130 generates first spatial information using the multi-channel signal and the second downmix signal and then sends the first spatial information to the multiplexing unit 150.

Spatial information is the information indicating a relation with a channel in downmixing a channel signal. And, the spatial information is used for a decoding apparatus to reconstruct an original channel signal from a downmix signal. First spatial information generated from downmixing a multi-channel signal includes CLD (channel level differences), ICC (interchannel correlations), CPC (channel prediction coefficients), or the like. The CLD indicates an energy difference between audio signals. The ICC indicates correlation or similarity between audio signals. And, the CPC indicates a coefficient for predicting an audio signal using another signal. The second spatial information generating unit 140 generates second spatial information using the first downmix signal and the second downmix signal and then sends the second spatial information to the multiplexing unit 150. In case that the first downmix signal is a 2-channel signal, the second spatial information can include IID (interchannel intensity difference) indicating an energy difference between two channels, IPD (interchannel phase difference) indicating a phase difference between two channels, ICC (interchannel correlation) indicating correlation between two channels, and the like.

Spatial information is the information extracted in the course of downmixing a channel signal according to a predetermined tree structure. In this case, the predetermined tree structure means the tree structure agreed between a decoding apparatus and an encoding apparatus. Spatial information is able to include tree structure information. In this case, the tree structure information is the information for a type of a tree structure. According to the type of the tree structure, the number of multi-channels, a per channel downmix sequence, and the like can be changed.

The multiplexing unit 150 generates a bitstream including the first spatial information and the second spatial information and then transfers the generated bitstream to the decoding apparatus together with or separately from a downmix signal.

The encoding apparatus is able to transfer the second downmix signal in a PCM signal format to the decoding apparatus. In this case, the multiplexing unit 150 generates a bitstream including the first spatial information and the second spatial information and then transfers the generated bitstream to the decoding apparatus together with or separately from a PCM signal. In case of transferring both of the PCM signal and the spatial information to the decoding apparatus, the multiplexing unit 150 generates one bitstream by embedding the first spatial information and the second spatial information in the PCM signal and then transfers the generated bitstream to the decoding apparatus.

And, the encoding apparatus is able to insert an identifier in the bitstream, In this case, the identifier indicates whether the transferred bitstream includes the second spatial information for the 2-channel signal generation, the first spatial information for the multi-channel signal generation, or both of the first spatial information and the second spatial information.

FIG. 2 is a block diagram of a second encoding apparatus according to another embodiment of the present invention.

Referring to FIG. 2, a second encoding apparatus includes a first downmixing unit 200, a second downmixing unit 210, a downmix signal encoding unit 220, a first spatial information generating unit 230, a second spatial information generating unit 240, and a multiplexing unit 250.

The first downmixing unit 200 receives a multi-channel signal and then downmixes the received signal into a first downmix signal having channels less than those of the multi-channel signal. And, the second downmixing unit 210 downmixes the first downmix signal into a second downmix signal having channels less than those of the first downmix signal.

The downmix signal encoding unit 220 compresses the second downmix signal and then sends the compressed signal to the multiplexing unit 250. The second downmix signal can be transferred in a PCM signal format to a decoding apparatus without passing through the downmix signal encoding unit 220.

The first spatial information generating unit 230 generates first spatial information using the multi-channel signal and the first downmix signal. The second spatial information generating unit generates second spatial information using the first downmix signal and the second downmix signal. And, the first spatial information generating unit 230 and the second spatial information generating unit 240 send the first spatial information and the second spatial information to the multiplexing unit 250, respectively.

The multiplexing unit 150 generates a bitstream by multiplexing the compressed downmix signal, the first spatial information, and the second spatial information together and then transfers the generated bitstream to the decoding apparatus.

The encoding apparatus separately generates a stream of the downmix signal, a stream for the first spatial information, and a stream for the second spatial information and then respectively transfers the separate streams to the decoding apparatus. Alternatively, the encoding apparatus generates a bitstream including the first spatial information and the second spatial information and then transfers the generated bitstream to the decoding apparatus together with the downmix signal.

The second encoding apparatus differs from the first encoding apparatus, which generates the first spatial information using the multi-channel signal and the second downmix signal, in generating the first spatial information using the multi-channel signal and the first downmix signal. So, the first spatial information generated by the first encoding apparatus differs from the first spatial information generated by the second encoding apparatus.

The decoding apparatus, which has received the downmix signal and the spatial information generated by the encoding apparatus explained in FIG. 1 or FIG. 2, reconstructs the 2-channel signal or the multi-channel signal using the spatial information and the downmix signal. The decoding apparatus decodes the downmix signal encoded and transferred by the encoding apparatus and then reconstructs the 2-channel signal or the multi-channel signal using the decoded downmix signal and the spatial information. So, an audio signal reconstructed by the decoding apparatus differs from an audio signal prior to downmixing in an audio quality. To prevent this, the encoding apparatus is able to generate spatial information using the downmix signal used for the decoding apparatus to reconstruct the audio signal.

An encoding method and apparatus for generating spatial information using a downmix signal user for a decoding apparatus to reconstruct an audio signal are explained with reference to FIG. 3 and FIG. 4 as follows.

FIG. 3 is a block diagram of a third encoding apparatus for generating spatial information using a decoded downmix signal according to one embodiment of the present invention.

Referring to FIG. 3, a third encoding apparatus includes a first downmixing unit 300, a second downmixing unit 310, a downmix signal encoding unit 320, a downmix signal decoding unit 330, a first spatial information generating unit 350, a second spatial information generating unit 340, and a multiplexing unit 360.

The third encoding apparatus differs from the first encoding apparatus in including the downmix signal decoding unit 330.

The first downmixing unit 300 downmixes a multi-channel signal into a first downmix signal and the second downmixing unit 310 downmixes the first downmix signal into a second downmix signal. The downmix signal encoding unit 320 encodes the second downmix signal. The downmix signal decoding unit 330 decodes the encoded second downmix signal. The second spatial information generating unit 340 generates second spatial information using the first downmix signal and the decoded second downmix signal.

The first encoding apparatus has a common feature with the third encoding apparatus in that the second spatial information is generated using the relation between the first downmix signal and the second downmix signal. Yet, the third encoding apparatus differs from the first encoding apparatus, which generates the second spatial information using the second downmix signal downmixed by the second downmixing unit 110, in encoding the second downmix signal, decoding the encoded second downmix signal, and then generating the second spatial information using the decoded second downmix signal. And, the second spatial information generated by the first encoding apparatus differs from the second spatial information generated by the third encoding apparatus.

The first spatial information generating unit 350 generates first spatial information using the multi-channel signal and the decoded second downmix signal. Unlike the first encoding apparatus generates the first spatial information using the second downmix signal, the third encoding apparatus encodes the second downmix signal, decodes the encoded signal again, and then generates the second spatial information using the decoded second downmix signal. Thus, the first encoding apparatus and the third encoding apparatus differ from each other. And, the first spatial information of the first encoding apparatus differs from that of the third encoding apparatus as well.

The multiplexing unit 360 multiplexes the encoded downmix signal, the first spatial information, and the second spatial information together and then transfers the multiplexed signal to the decoding apparatus.

The decoding apparatus decodes the second downmix signal encoded and transferred by the encoding apparatus and then reconstructs the 2-channel signal or the multi-channel signal by applying at least one of the first spatial information and the second spatial information to the decoded downmix signal. So, the channel signal reconstructed by the decoding apparatus has an audio quality closer to the audio signal prior to being downmixed by the encoding apparatus.

FIG. 4 is a block diagram of a fourth encoding apparatus for generating spatial information using a decoded downmix signal according to another embodiment of the present invention.

Referring to FIG. 4, a fourth encoding apparatus includes a first downmixing unit 400, a second downmixing unit 410, a downmix signal encoding unit 420, a downmix signal decoding unit 430, a first spatial information generating unit 460, a second spatial information generating unit 440, a first downmix signal generating unit 450, and a multiplexing unit 470.

The fourth encoding apparatus differs from the second encoding apparatus in including the downmix signal decoding unit 430 and the first downmix signal generating unit 450.

The first downmixing unit 400 downmixes a multi-channel signal into a first downmix signal and the second downmixing unit 410 downmixes the first downmix signal into a second downmix signal. The downmix signal encoding unit 420 encodes the second downmix signal and then sends it to the downmix signal decoding unit 430. The downmix signal decoding unit 430 decodes the encoded downmix signal and then sends it to the second spatial information generating unit 440. The second spatial information generating unit 440 generates second spatial information using the first downmix signal and the decoded second downmix signal.

The fourth encoding apparatus differs from the second encoding apparatus, which generates the second spatial information using the second downmix signal without being encoded and decoded, in generating the second spatial information using the downmix signal encoded by the downmix signal encoding unit 420 and then decoded by the downmix signal decoding unit 430 again.

The first downmix signal generating unit 450 generates a modified first downmix signal using the second downmix signal decoded by the downmix signal decoding unit 430 and the second spatial information. The modified first downmix signal differs from the first downmix signal downmixed by the first downmixing unit 400 in being generated from the encoded and re-decoded second downmix signal and the second spatial information generated using the encoded and re-decoded second downmix signal.

The first spatial information generating unit 460 generates first spatial information using the modified first downmix signal and the multi-channel signal. The first spatial information generating unit 460 differs from the second encoding apparatus, which generates the first spatial information using the first downmix signal intactly, in generating the first spatial information using the modified first downmix signal generated by the first downmix signal generating unit 450. And, the first spatial information generated by the first spatial information generating unit 460 differs from the first spatial information generated by the second encoding apparatus. The multiplexing unit 470 generates a bitstream including both of the first spatial information and the second spatial information.

And, the fourth encoding apparatus transfers the bitstream including the spatial information to the decoding apparatus together with or separately from the second downmix signal.

FIG. 5 is a diagram of a bitstream of an audio signal according to one embodiment of the present invention.

Referring to FIG. 5, an audio signal according to the present invention includes a downmix signal 500 and a spatial information signal 600. The audio signal exists in an ES elementary stream) form having frames arranged therein.

The downmix signal 500 and the spatial information signal 600 can be transferred in different ES forms to a decoding apparatus, respectively. Alternatively, they can be transferred in one ES form having the downmix and spatial information signals 500 and 600 combined together. In case of transferring the downmix signal 500 and the spatial information signal 600 in a combined form to the decoding apparatus, the spatial information signal 600 can be included in a location of ancillary data or extension data of the downmix signal 500.

The audio signal can include a codec identifier to enable a decoding apparatus to recognize basic information for audio codec without interpreting the audio signal. The codec identifier is the information indicating what kind of coding scheme is used in encoding the audio signal. The codec identifier can be included in a header 610 or spatial information 620 of the spatial information signal 600. And, the codec identifier can include a spatial information identifier. In this case, the spatial information identifier is the information indicating whether a bitstream includes second spatial information to generate 2-chanel signal from the audio signal, first spatial information to generate multi-channel signal from the audio signal, or both of the first spatial information and the second spatial information. so, the decoding apparatus is able to detect a type of the audio signal generatable from the downmix signal and the like and the like using the spatial information identifier.

The spatial information signal 600 can include the header 610 and the spatial information 620. Alternatively, the spatial information signal 600 can include the spatial information 620 only without including the header 610. Namely, the spatial information signal 600 is able to use a frame including the header 610 or a frame not including the header 610 together.

In case that the audio signal includes spatial information to generate multi-channel signal and spatial information to generate 2-channel signal, the header 610 can include a 2-channel signal header 611 and a multi-channel signal header 613.

In case that a signal reconstructible by the decoding apparatus is the 2-channel signal, the decoding apparatus decodes second spatial information 623 to generate the 2-channel signal using the 2-channel signal header 611 and then reconstructs the 2-channel signal using the decoded second spatial information 623.

In case that a signal reconstructible by the decoding apparatus is the multi-channel signal, the decoding apparatus decodes spatial information to generate the multi-channel signal using the multi-channel signal header 613. The spatial information for the multi-channel signal reconstruction can include the second spatial information 623 as well as the first spatial information 621. In case that the decoding apparatus reconstructs the 2-channel signal and then reconstructs the multi-channel signal from the reconstructed 2-channel signal, the multi-channel signal can be reconstructed using the second spatial information 623 for the 2-channel signal reconstruction and the first spatial information 621 for reconstructing the multi-channel signal from the 2-channel signal step by step. And, the spatial information signal can include the aforesaid tree structure information as well.

FIG. 6 is a block diagram of a first decoding apparatus according to one embodiment of the present invention.

Referring to FIG. 6, a first decoding apparatus includes a demultiplexing unit 700, a downmix signal decoding unit 720, a 2-channel signal generating unit 710, and a multi-channel signal generating unit 730.

The demultiplexing unit 700 parses a downmix signal and then sends the parsed signal to the downmix signal decoding unit 720. The downmix signal can be a mono signal. And, the downmix signal can be a signal on a frequency domain. The frequency domain can be a QMF domain.

The downmix signal decoding unit 720 decodes the downmix signal and then outputs the decoded downmix signal intactly. The downmix signal decoding unit 720 upmixes the downmix signal into a 2-channel signal or a multi-channel signal using spatial information and then outputs the upmixed signal. In case that the downmix signal is a PCM signal, the downmix can be outputted intact without passing through the downmix signal decoding unit 720.

A decoding apparatus is able to detect what kind of spatial information is included in a bitstream using a spatial information identifier included in the bitstream.

If a downmix signal is a mono signal and if a signal generatable by the first decoding apparatus is one of a 2-channel signal and a multi-channel signal, the decoding apparatus decides whether the downmix signal is a signal capable of generating the 2-channel signal or the multi-channel signal using a spatial information identifier. If the decoding apparatus decides that both spatial information for 2-channel signal generation and spatial information for multi-channel signal generation are included in a bitstream, the decoding apparatus extracts spatial information for specific signal generation from the spatial information for 2-channel signal generation and the spatial information for multi-channel signal generation only and is then able to generate a channel signal using the extracted information.

If a downmix signal is a PCM signal, the first spatial information 621 and the second spatial information 623 can be transmitted by being embedded in the downmix signal. In this case, the demultiplexing unit 700 is able to extract the first spatial information 621 and the second spatial information 623 from the downmix signal.

In case that the decoding apparatus is capable of generating 2-chanel signal only, the demultiplexing unit 700 of the decoding apparatus parses the second spatial information 623 for 2-channel signal generation in the transferred spatial information and then sends the parsed information to the 2-channel signal generating unit 710. In case that the decoding apparatus is capable of generating multi-channel signal only, the demultiplexing unit 700 of the decoding apparatus parses the first spatial information 621 for multi-channel signal generation in the transferred spatial information and then sends the parsed information to the multi-channel signal generating unit 730. Namely, if the decoding apparatus generates a multi-channel signal directly from a downmix signal and spatial information instead of generating multi-channel signal from 2-channel signal, the decoding apparatus need not use the second spatial information 623. So, the decoding apparatus extracts the first spatial information 621 only to use.

In case that the decoding apparatus is able to generate both 2-channel signal and multi-channel signal, the decoding apparatus is able to extract spatial information for user-selected channel signal generation by receiving control information from a user.

In case that a signal generatable by the decoding apparatus is 2-channel signal or a user selects 2-channel signal generation, the 2-chanel signal generating unit 710 generates 2-channel signal using the second spatial information 623 parsed and sent by the demultiplexing unit 700 and the decoded downmix signal and then outputs the generated signal. The 2-channel signal generating unit 710 generates the 2-channel signal by upmixing a mono downmix signal using a signal transforming unit (not shown in the drawing), and more particularly, an OTT box. In this case, the multi-channel signal generating unit 730 needs not to operate. The demultiplexing unit 700 can generate an identifier controlling an operation of the multi-channel signal generating unit 730 and send the generated identifier to the multi-channel signal generating unit 730. Hereinafter, the identifier controlling an operation of the 2-channel signal generating unit 710 or the multi-channel signal generating unit 730 is named an operation control identifier. The multi-channel signal generating unit 730 does not operate according to the operation control identifier received from the demultiplexing unit 700. And, it is unnecessary to consider the first spatial information

621.

In case that a signal generatable by the decoding apparatus is multi-channel signal or a user selects multi-channel signal generation, the multi-channel signal generating unit 730 generates multi-channel signal using the first spatial information 621 and then outputs the generated signal. The multi-channel signal generating unit 730 upmixes a downmix signal using a plurality of signal transforming units. As mentioned in the foregoing description, the signal transforming unit includes an OTT box or a TTT box. In this case, since the 2-channel signal generating unit 710 needs not to operate, the demultiplexing unit 700 generates an operation control identifier and then sends the generated operation control identifier to the 2-channel signal generating unit 710 to control an operation of the 2-channel signal generating unit 710. The 2-channel signal generating unit 710 does not operate according to the operation control identifier. And, it is unnecessary to consider the second spatial information 623.

The decoding apparatus can further include a modified spatial information generating unit (not shown in the drawing). The modified spatial information generating unit identifies a type of modified spatial information using spatial information and generates modified spatial information of the type identified based on the spatial information. In this case, the modified spatial information means the spatial information that is newly generated using spatial information. The modified spatial information can be generated by combining spatial information. The modified spatial information generating unit is able to generate modified spatial information using tree structure information, output channel information and the like included in the spatial information. The output channel information is the information for a speaker interconnecting with the decoding apparatus and can include the number of output channels, position information for each output channel, and the like. The output channel information is inputted to the decoding apparatus in advance by a manufacturer or can be inputted by a user.

The decoding apparatus decides whether the number of original multi-channels downmixed by the encoding apparatus is equal to the number of channels to be generated using the tree structure information and the output channel information. Hereinafter, the original multi-channels downmixed by the encoding apparatus are named first multi-channels. If the number of the first multi-channels downmixed by the encoding apparatus is different from the number of multi-channels to be generated, the decoding apparatus is able to modify spatial information using the modified spatial information generating unit. In this case, the modified spatial information can be generated by combining the aforesaid CLD, ICC, CPC, IPC, and the like. The decoding apparatus is able to generate multi-channels of which number differs from the number of the first multi-channels using the modified spatial information and the downmix signal.

FIG. 7 is a block diagram of a second encoding apparatus according to another embodiment of the present invention.

Referring to FIG. 7, a second decoding apparatus includes a demultiplexing unit 800, a downmix signal decoding unit 810, a 2-channel signal generating unit 820, and a multi-channel signal generating unit 830.

The demultiplexing unit 800 parses a downmix signal from a bitstream transferred from an encoding apparatus or a bitstream recorded in a storage medium and then sends the parsed signal to the downmix signal decoding unit 810.

The downmix signal decoding unit 810 decodes the downmix signal and outputs the decoded signal as a mono signal or generates 2-channel signal or multi-channel signal using spatial information.

In case that the decoding apparatus is able to generate 2-channel signal or that 2-channel signal generation is selected by a user despite that the decoding apparatus is able generate both 2-channel signal and multi-channel signal, the demultiplexing unit 800 extracts second spatial information 623 for 2-channel signal generation and then sends the extracted information to the 2-channel signal generating unit.

The 2-channel signal generating unit 820 generates 2-channel signal using the second spatial information 623 and the decoded downmix signal.

Since the second spatial information 623 is applied to the downmix signal on a frequency domain, the 2-chanel signal should be converted to a signal on a time domain in order for the decoding apparatus to output the 2-channel signal. The decoding apparatus is able to use FFT (fast Fourier transform), DFT (discrete Fourier transform), QMF or hybrid function, or the like in converting a time domain to a frequency domain, and vice versa. And, the decoding apparatus output a domain-converted 2-chanel signal.

In case that the decoding apparatus generates the 2-channel signal only, it is unnecessary to generate multi-channel signal. So, the demultiplexing unit 800 generates an operation control identifier in order for the multi-channel signal generating unit 830 not to operate and then sends the generated identifier to the multi-channel signal generating unit 830. The multi-channel signal generating unit 830 does not operate according to the operation control identifier. And, it is unnecessary to consider the first spatial information 621 for the multi-channel signal generation.

In case that the decoding apparatus is able to generate multi-channel signal or that multi-channel signal generation is selected by a user, the demultiplexing unit 800 extracts spatial information for the multi-channel signal generation. Since the second decoding apparatus generates multi-channel signal using 2-channel signal unlike the first decoding apparatus, the demultiplexing unit 800 extracts both second spatial information 623 for 2-channel signal generation and first spatial information 621 for generating multi-channel signal from the 2-channel signal. So, the first spatial information used by the first decoding apparatus is discriminated from the first spatial information used by the second decoding apparatus. In particular, the second spatial information used by the second decoding apparatus is the spatial information required for generating the multi-channel signal from the 2-channel signal, whereas the first spatial information used by the first decoding apparatus is the spatial information required for generating multi-channels from the downmix signal.

The 2-chanel signal generating unit 820 generates 2-channel signal using the second spatial information 623 and the decoded downmix signal and then sends the generated signal to the multi-channel signal generating unit 830.

The multi-channel signal generating unit 830 is able to generate multi-channel signal using the 2-channel signal sent by the 2-channel signal generating nit 820 and the first spatial information 621 extracted by the demultiplexing unit 800. In case that the 2-channel signal generation and the multi-channel signal generation are carried out on the same domain, i.e., a frequency domain, the multi-channel signal generating unit 830 is able to generate multi-channel signal using 2-channel signal on the frequency domain. In this case, the frequency domain includes a QMF domain, a hybrid domain, or the like. In particular, the multi-channel signal generating unit 830 is able to generate multi-channel signal by applying the first spatial information 621 to the 2-channel signal having a domain not converted to a time domain. In this case, it is unnecessary to convert the 2-channel signal to a signal on the time domain. And, a user is able to select and use the

2-channel signal or the multi-channel signal using the first decoding apparatus, the second decoding apparatus, or the like.

WHAT IS CLAIMED IS:

1. A method of processing a media signal, comprising:

extracting a downmix signal from a bitstream;

extracting at least one of first spatial information and second spatial information from the bitstream; and

generating multi-channels using the extracted spatial information and the downmix signal.

2. The method of claim 1, wherein the first spatial information is information for generating at least three channels and wherein the second spatial information is information for generating two channels.

3. The method of claim 2, wherein each of the first and second spatial informations comprises at least one selected from a group consisting of channel level differences, interchannel correlations, channel prediction coefficients, and interchannel phase differences.

4. The method of claim 1, wherein the downmix signal comprises a mono signal.

5. The method of claim 4, wherein the downmix signal comprises a signal on a frequency domain.

6. The method of claim 2, wherein the generating multi-channels is generating two channels by upmixing the downmix signal using a signal transforming unit if the extracted spatial information is the second spatial information.

7. The method of claim 1, further comprising modifying the spatial information, wherein the generating multi-channel is carried out using the modified spatial information and the downmix signal.

8. The method of claim 7, wherein the modified spatial information is generated by combining the spatial information.

9. The method of claim 8, wherein the downmix signal is a signal generated from downmixing first multi-channels and wherein a number of the multi-channels reconstructed using the modified spatial information and the downmix signal differs from a number of the first multi-channels.

10. The method of claim 4 or claim 7, wherein the generating multi-channel is performed by upmixing the downmix signal using a plurality of signal transforming units.

11. The method of claim 2, wherein the generating multi-channels comprises generating the two channels using the downmix signal and the extracted second spatial information or generating the at least three channels using the generated two channels and the extracted first spatial information.

12. The method of claim 11, wherein the two channels or the at least three channels are generated on a same domain and wherein the same domain is a frequency domain.

13. The method of claim 2, wherein the generating multi-channels comprises generating the two channels using the downmix signal and the extracted second spatial information or generating the at least three channels using the downmix signal and the extracted first spatial information.

14. The method of any one of claim 1, claim 11, and claim 13, wherein the extracting the spatial information and the generating the multi-channels are carried out according to a user's selection or a channel type generatable by an apparatus for performing the method.

15. The method of claim 11, wherein the two channels are generated on a frequency domain,

the method further comprising:

converting the generated two channels into a time domain; and

outputting the two channels on the time domain,

wherein the at least three channels are generated using the two channels on the frequency domain or the two channels on the time domain.

16. A method of processing a media signal, comprising:

generating a first downmix signal from multi-channels;

generating a second downmix signal from the first downmix signal;

generating first spatial information using the multi-channels and the first downmix signal or the multi-channels and the second downmix signal;

generating second spatial information using the first downmix signal and the second downmix signal; and

generating a bitstream including the first spatial information and the second spatial information.

17. The method of claim 16, wherein the first downmix signal is a 2-channel signal and wherein the second downmix signal is a mono signal.

18. A method of processing a media signal, comprising:

generating a first downmix signal from multi-channels;

generating a second downmix signal from the first downmix signal;

encoding the second downmix signal;

decoding the encoded second downmix signal;

generating second spatial information using the first downmix signal and the decoded second downmix signal; and

generating first spatial information using the multi-channels and the decoded second downmix signal.

19. A method of processing a media signal, comprising:

generating a first downmix signal from multi-channels;

generating a second downmix signal from the first downmix signal;

encoding the second downmix signal;

decoding the encoded second downmix signal;

generating second spatial information using the first downmix signal and the decoded second downmix signal;

generating a modified first downmix signal using the decoded second downmix signal and the second spatial information; and

generating first spatial information using the modified first downmix signal and the multi-channels.

20. The method of claim 18 or claim 19, further comprising generating a bitstream including the first spatial information and the second spatial information.

21. An apparatus for processing a signal, comprising:

a downmix signal extracting unit extracting a downmix signal from a bitstream;

an information extracting unit extracting at least one of second spatial information for generating two channels from the downmix signal and first spatial information for generating at least three channels from the downmix signal from the bitstream; and

a channel generating unit generating either the two channels or the at least three channels using the extracted information and the downmix signal.

22. A bitstream structure comprising:

first spatial information extracted in the course of generating a first downmix signal including at least two channels from multi-channels; and

second spatial information extracted in the course of generating a second downmix signal from the first downmix signal.

23. A storage medium including the bitstream structure of claim 22.

24. A signal processing apparatus comprising:

a first downmixing unit generating a first downmix signal from multi-channels;

a second downmixing unit generating a second downmix signal from the first downmix signal;

a first spatial information generating unit generating first spatial information using the multi-channels and the first downmix signal or the multi-channels and the second downmix signal;

a second spatial information generating unit generating second spatial information using the first downmix signal and the second downmix signal; and

a multiplexing unit generating a bitstream including the first spatial information and the second spatial information.

25. A signal processing apparatus comprising:

a downmixing unit generating a downmix signal from multi-channels;

an encoding unit encoding the downmix signal;

a decoding unit decoding the encoded downmix signal; and

a spatial information generating unit generating spatial information using the multi-channels and the decoded downmix signal.

FIG. 1

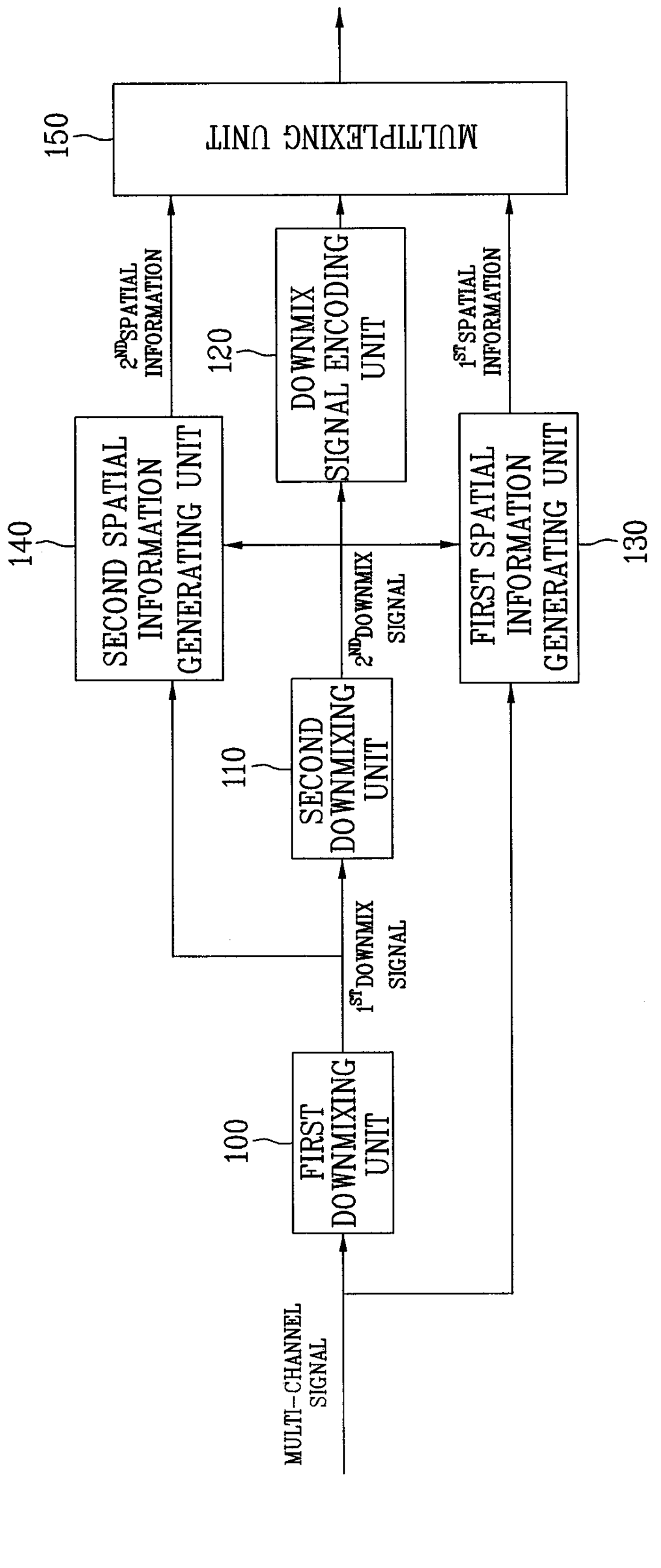

MULTI-CHANNEL SIGNAL
100
FIRST DOWNMIXING UNIT
1ST DOWNMIX SIGNAL
110
SECOND DOWNMIXING UNIT
2ND DOWNMIX SIGNAL
120
DOWNMIX SIGNAL ENCODING UNIT
130
FIRST SPATIAL INFORMATION GENERATING UNIT
1ST SPATIAL INFORMATION
140
SECOND SPATIAL INFORMATION GENERATING UNIT
2ND SPATIAL INFORMATION
150
MULTIPLEXING UNIT

FIG. 2

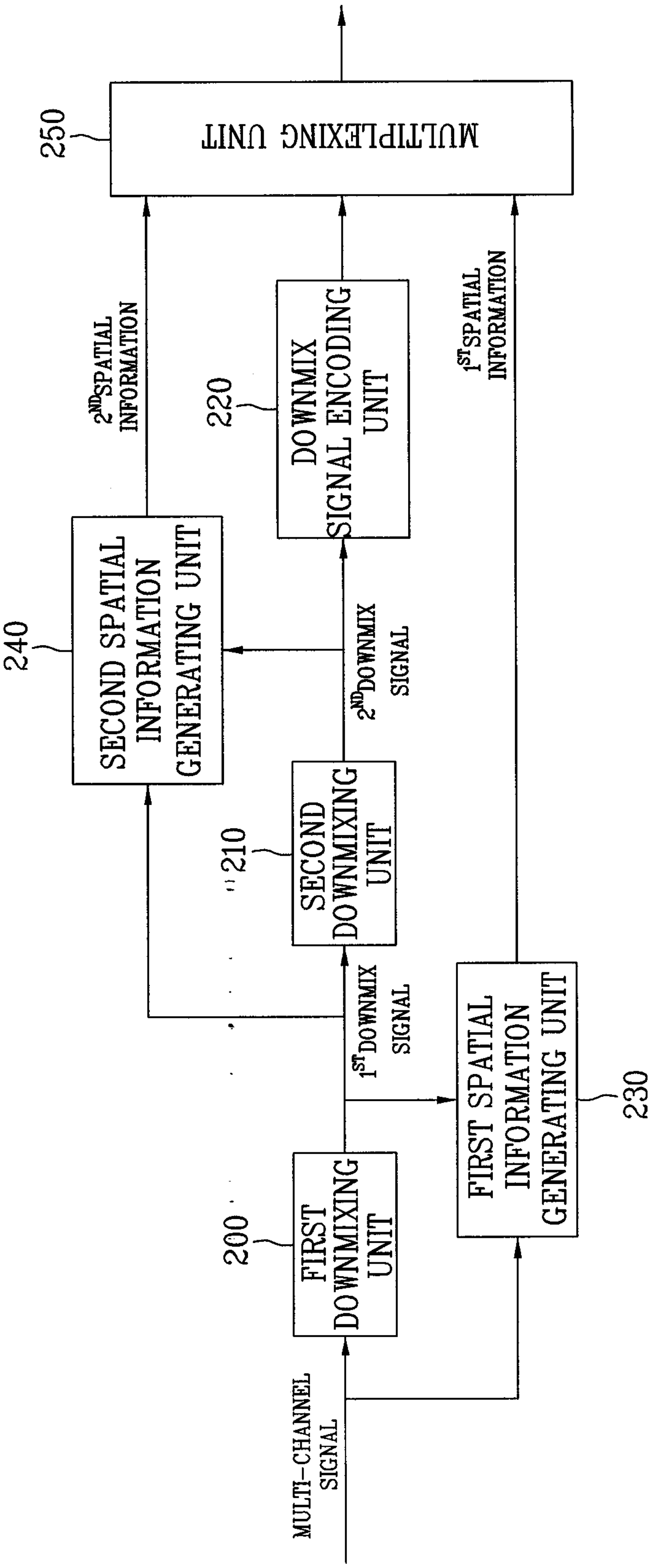

MULTI-CHANNEL SIGNAL
FIRST DOWNMIXING UNIT
200
1ST DOWNMIX SIGNAL
SECOND DOWNMIXING UNIT
210
2ND DOWNMIX SIGNAL
DOWNMIX SIGNAL ENCODING UNIT
220
FIRST SPATIAL INFORMATION GENERATING UNIT
230
1ST SPATIAL INFORMATION
SECOND SPATIAL INFORMATION GENERATING UNIT
240
2ND SPATIAL INFORMATION
MULTIPLEXING UNIT
250

FIG. 3

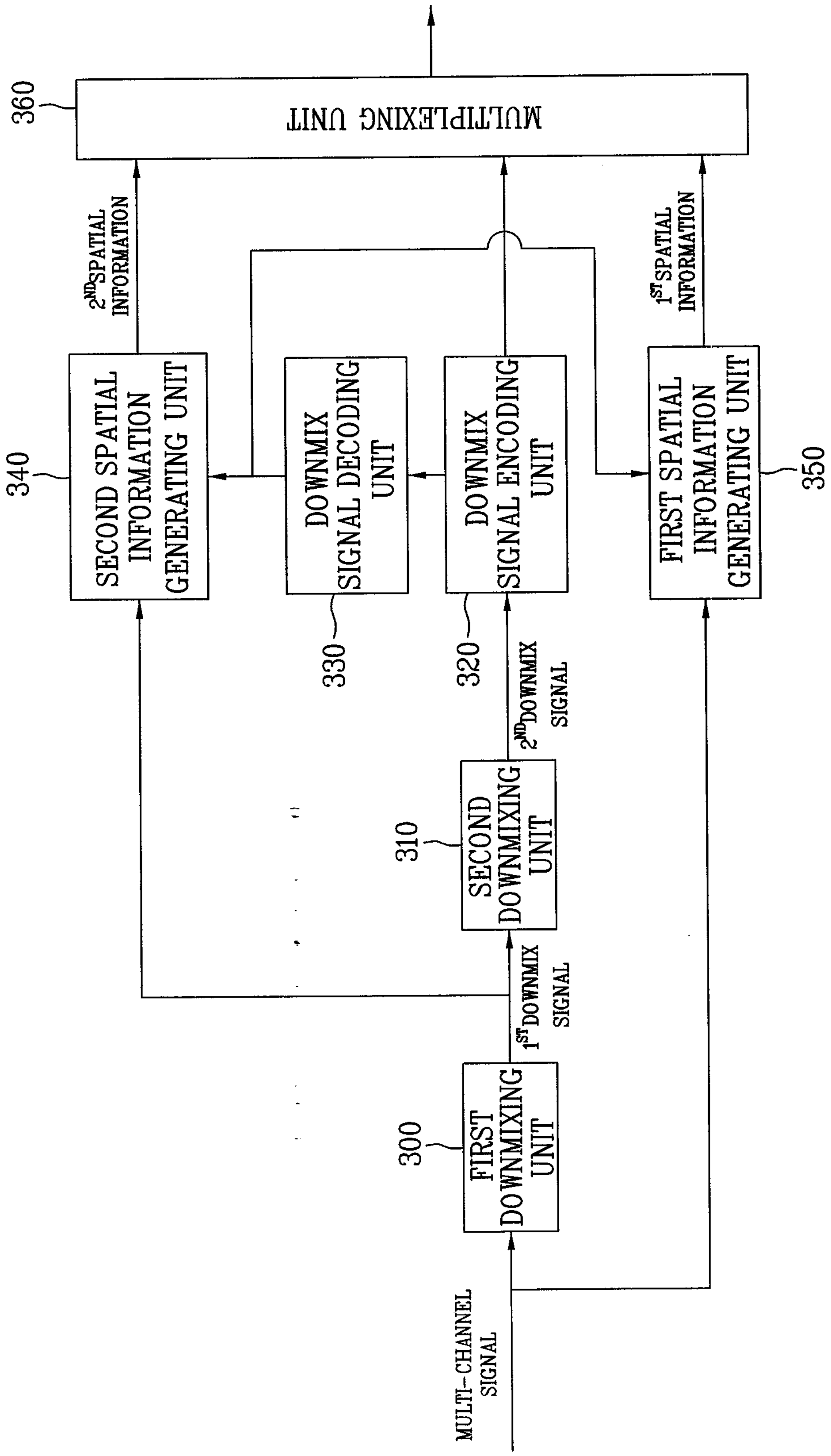

MULTI-CHANNEL SIGNAL
300
FIRST DOWNMIXING UNIT
1ST DOWNMIX SIGNAL
310
SECOND DOWNMIXING UNIT
2ND DOWNMIX SIGNAL
320
DOWNMIX SIGNAL ENCODING UNIT
330
DOWNMIX SIGNAL DECODING UNIT
340
SECOND SPATIAL INFORMATION GENERATING UNIT
350
FIRST SPATIAL INFORMATION GENERATING UNIT
2ND SPATIAL INFORMATION
1ST SPATIAL INFORMATION
360
MULTIPLEXING UNIT

FIG. 4

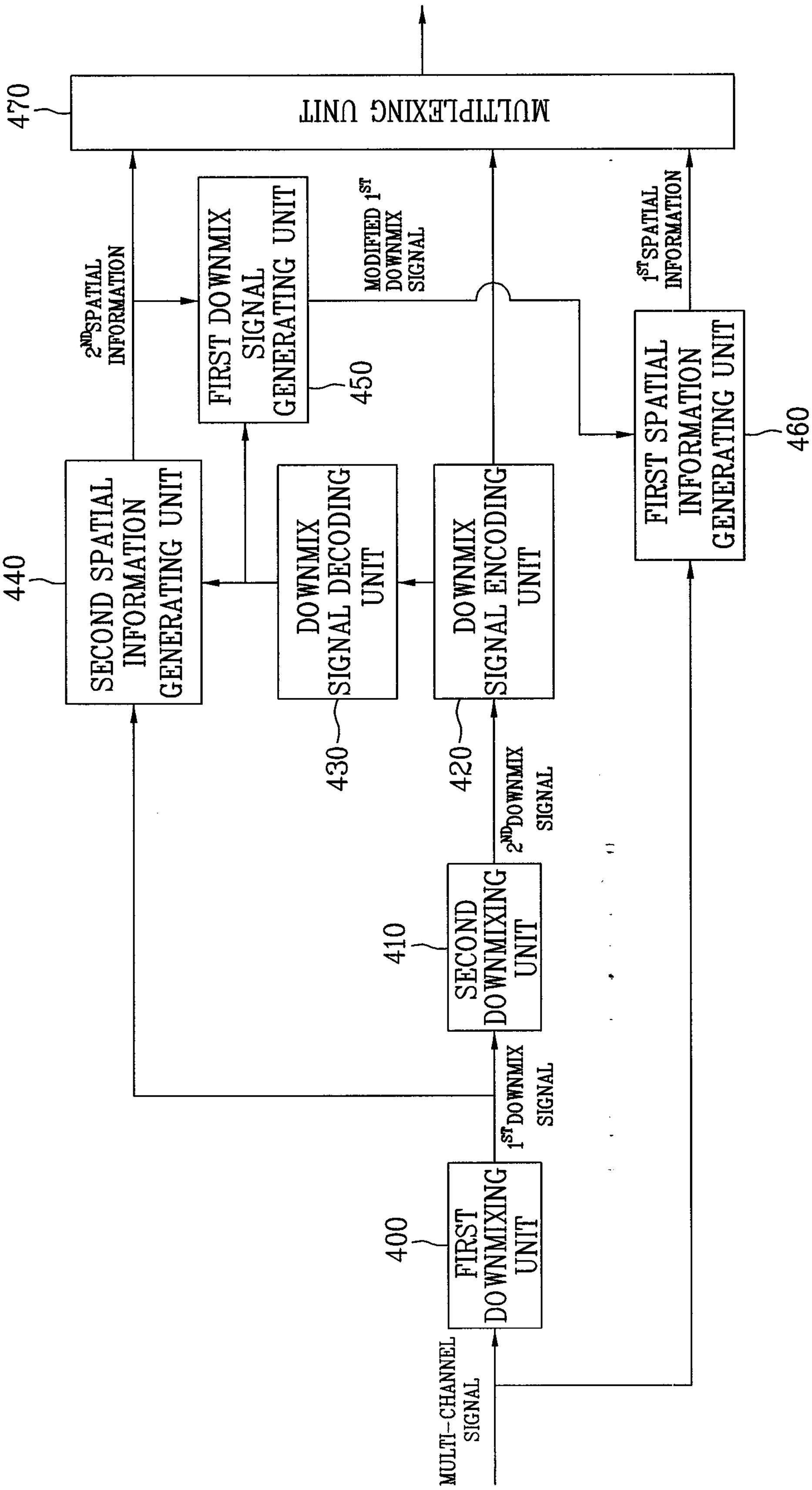

MULTI-CHANNEL SIGNAL
400
FIRST DOWNMIXING UNIT
1ST DOWNMIX SIGNAL
410
SECOND DOWNMIXING UNIT
2ND DOWNMIX SIGNAL
420
DOWNMIX SIGNAL ENCODING UNIT
430
DOWNMIX SIGNAL DECODING UNIT
440
SECOND SPATIAL INFORMATION GENERATING UNIT
2ND SPATIAL INFORMATION
450
FIRST DOWNMIX SIGNAL GENERATING UNIT
MODIFIED 1ST DOWNMIX SIGNAL
460
FIRST SPATIAL INFORMATION GENERATING UNIT
1ST SPATIAL INFORMATION
470
MULTIPLEXING UNIT

FIG. 5

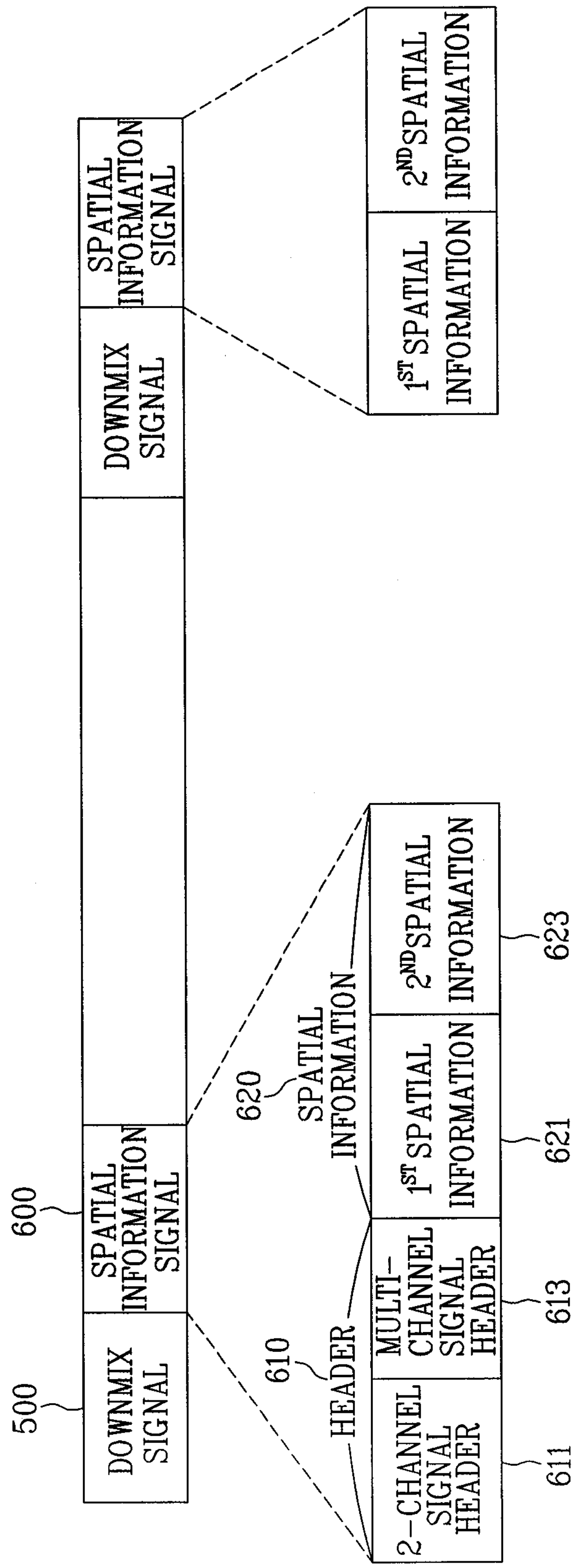

500
DOWNMIX SIGNAL
600
SPATIAL INFORMATION SIGNAL
DOWNMIX SIGNAL
SPATIAL INFORMATION SIGNAL
610
HEADER
2-CHANNEL SIGNAL HEADER
611
MULTI-CHANNEL SIGNAL HEADER
613
620
SPATIAL INFORMATION
1ST SPATIAL INFORMATION
621
2ND SPATIAL INFORMATION
623
1ST SPATIAL INFORMATION
2ND SPATIAL INFORMATION

FIG. 6

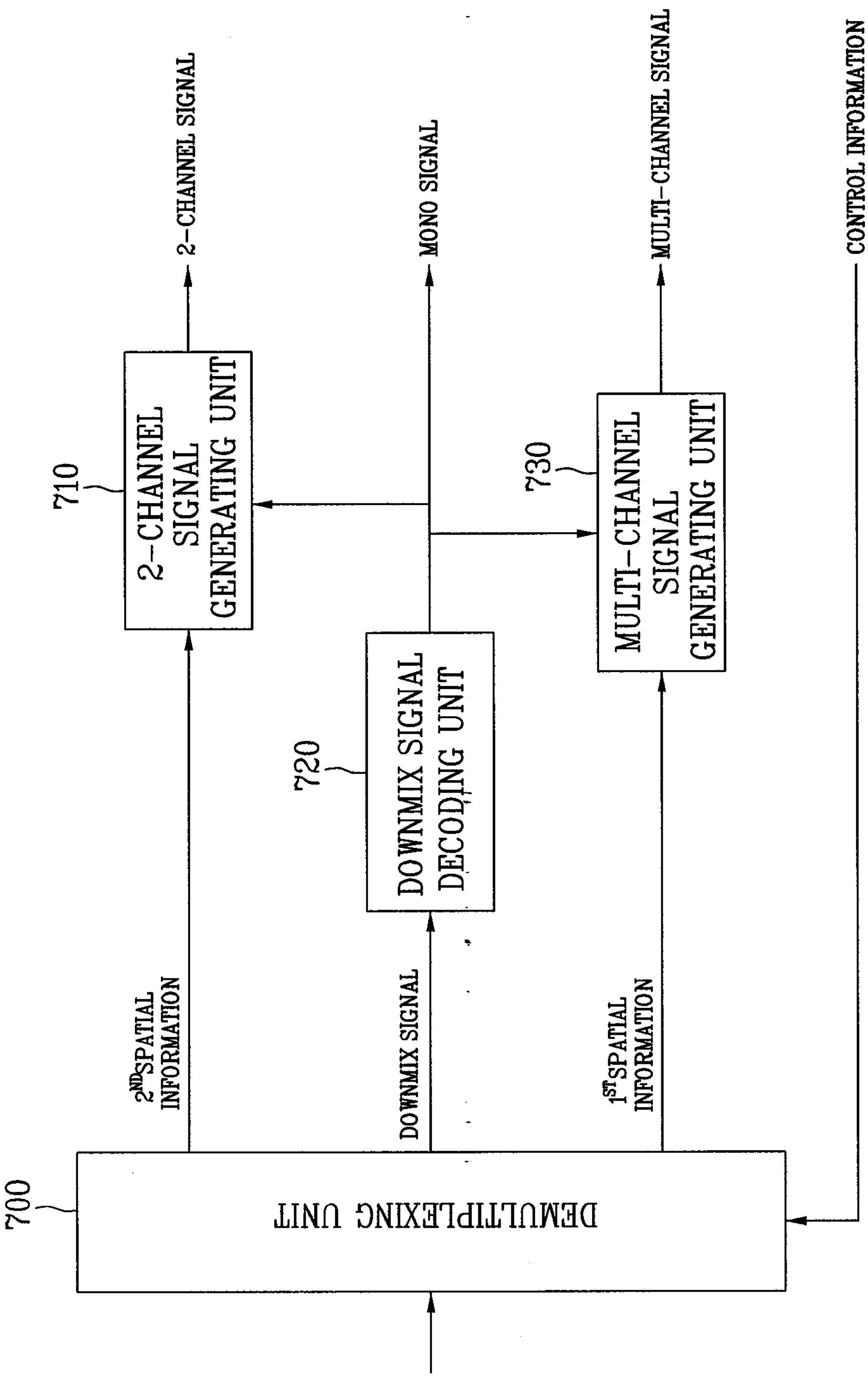
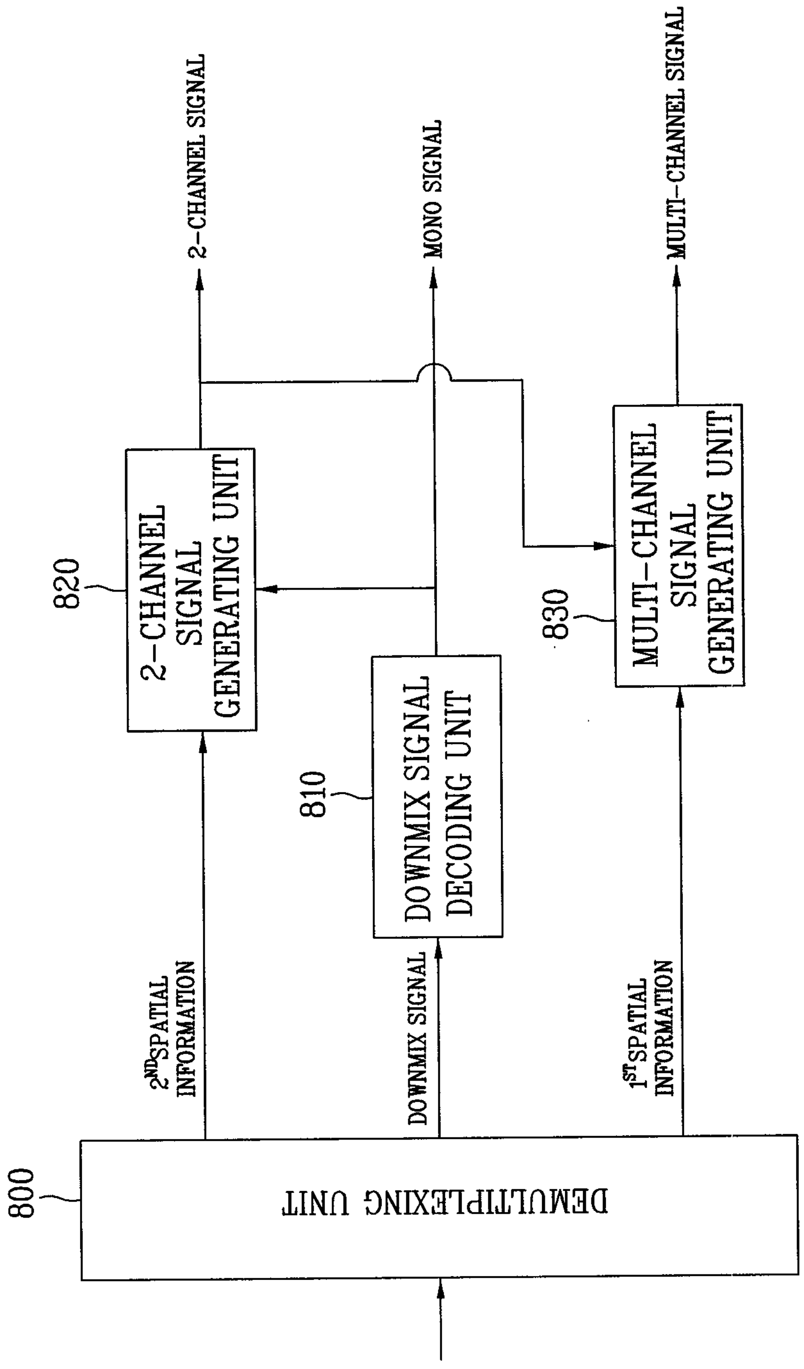

700
DEMULTIPLEXING UNIT
2ND SPATIAL INFORMATION
DOWNMIX SIGNAL
1ST SPATIAL INFORMATION
710
2-CHANNEL SIGNAL GENERATING UNIT
720
DOWNMIX SIGNAL DECODING UNIT
730
MULTI-CHANNEL SIGNAL GENERATING UNIT
2-CHANNEL SIGNAL
MONO SIGNAL
MULTI-CHANNEL SIGNAL
CONTROL INFORMATION

FIG. 7


800
DEMULTIPLEXING UNIT
2ND SPATIAL INFORMATION
DOWNMIX SIGNAL
1ST SPATIAL INFORMATION
810
DOWNMIX SIGNAL DECODING UNIT
820
2-CHANNEL SIGNAL GENERATING UNIT
830
MULTI-CHANNEL SIGNAL GENERATING UNIT
2-CHANNEL SIGNAL
MONO SIGNAL
MULTI-CHANNEL SIGNAL

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2007/001560

A. CLASSIFICATION OF SUBJECT MATTER

G10L 19/00(2006.01)i

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 19/00, G10L 19/02, G10L 19/04, H04S 5/00, H04S 3/00, H04R 5/00

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

Korean Utility models and applications for Utility Models since 1975
Japanese Utility Models and applications for Utility Models since 1975

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

eKIPASS(KIPO internal) "downmix, spatial, audio"

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	WO 2004008805 A1 (KONINKLIJKE PHILIPS ELECTRONICS N.V.) 22 January 2004 See claim 7; fig. 3.; page 12.	1 2-15,21
X	BREEBAART et al. 'MPEG Spatial Audio Coding / MPEG Surround: Overview and Current Status' In: Proc. 119th AES Convention. New York, October 2005. See fig. 5.	16,17,24
X Y	HERRE, J. et al. 'The Reference Model Architecture for MPEG Spatial Audio Coding' In: Proc. 118th AES Convention. May 2005. See page 12. Pages 3-4. Page 11. Page 6.	 22,23 2-5,15,21 7-9,14 6,10-13
A	US 20030236583 A1 (BAUMGARTE, F. et al.) 25 December 2003 See whole document.	1-25

☐ 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 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	Date of mailing of the international search report
19 JULY 2007 (19.07.2007)	**20 JULY 2007 (20.07.2007)**

Name and mailing address of the ISA/KR	Authorized officer
Korean Intellectual Property Office 920 Dunsan-dong, Seo-gu, Daejeon 302-701, Republic of Korea Facsimile No. 82-42-472-7140	JANG Ho Keun Telephone No. 82-42-481-8187

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2004008805 A1	22.01.2004	AU 2003244932 AA	02.02.2004
		BR 200305434 A	28.09.2004
		CN 1669359 A	14.09.2005
		EP 1523862 A1	20.04.2005
		JP 17533426 A	04.11.2005
		KR 1020050019851 A	03.03.2005
		RU 2005103637 A	10.07.2005
		US 20060206323 A1	14.09.2006
US 20030236583 A1	25.12.2003	EP 1376538 A1	02.01.2004
		JP 16078183 A2	11.03.2004
		US 2005058304 A1	17.03.2005