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

(19) World Intellectual Property Organization
International Bureau

PCT

(43) International Publication Date
10 July 2008 (10.07.2008)

(10) International Publication Number
WO 2008/082276 A1

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

(21) International Application Number:
PCT/KR2008/000073

(22) International Filing Date: 7 January 2008 (07.01.2008)

(25) Filing Language: Korean

(26) Publication Language: English

(30) Priority Data:

60/883,569	5 January 2007 (05.01.2007)	US
60/884,043	9 January 2007 (09.01.2007)	US
60/885,347	17 January 2007 (17.01.2007)	US

(71) Applicant *(for all designated States except US):* LG ELECTRONICS INC. [KR/KR]; IP GROUP 16, Woomyeon-Dong, Seocho-Gu, Seoul 137-724 (KR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only):* OH, Hyen O [KR/KR]; LG ELECTRONICS INC. IP GROUP, 16 Woomyeon-dong, Seocho-gu, Seoul 137-724 (KR). JUNG, Yang Won [KR/KR]; LG ELECTRONICS INC. IP GROUP, 16 Woomyeon-dong, Seocho-gu, Seoul 137-724 (KR).

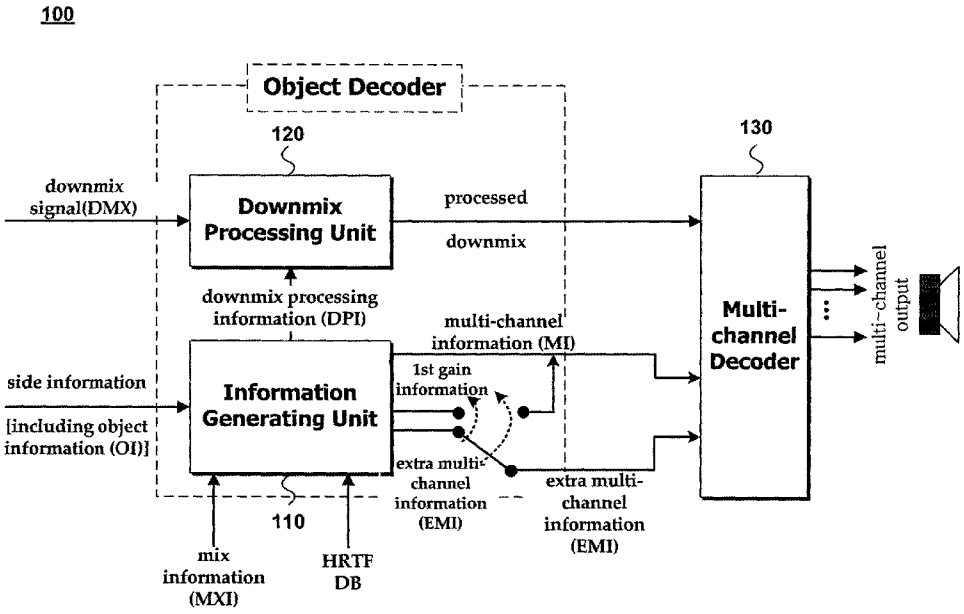
(74) Agents: KIM, Yong In et al; KBK & Associates, 15th Floor Yo Sam Building, 648-23, Yeoksam-dong,, Kangnam-gu,, Seoul 135-080 (KR).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available):* AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, 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, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, 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*

(54) Title: A METHOD AND AN APPARATUS FOR PROCESSING AN AUDIO SIGNAL



(57) Abstract: A method of processing an audio signal is disclosed. The present invention includes receiving downmix information, object information and mix information, generating and transferring multi-channel information using at least one of the downmix information, the object information and the mix information, and selectively generating and transferring either first gain information or extra multi- channel information including second gain information in accordance with a decoding mode using at least one of the object information and the mix information.

WO 2008/082276 A1

A METHOD AND AN APPARATUS FOR PROCESSING AN AUDIO SIGNAL

TECHNICAL FIELD

The present invention relates to an apparatus for processing an audio signal and method thereof. Although the present invention is suitable for a wide scope of applications, it is particularly suitable for processing an audio signal received *on* a digital medium, a broadcast signal or the like.

BACKGROUND ART

Generally, while downmixing several audio objects to be a mono or stereo signal, parameters from the individual object signals can be extracted. These parameters can be used in a decoder of an audio signal, and positioning/panning of the individual sources can be controlled by user' selection.

DISCLOSURE OF THE INVENTION

TECHNICAL PROBLEM

However, in order to control each object signal, sources included in downmix need to be appropriately positioned or panned.

Moreover, in order to provide backward compatibility

with a channel-oriented decoding scheme, an object parameter should be flexibly converted to a multi-channel parameter.

TECHNICAL SOLUTION

Accordingly, the present invention is directed to an apparatus for processing an audio 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 apparatus for processing an audio signal and method thereof, by which gain and panning of an object can be controlled without restriction.

Another object of the present invention is to provide an apparatus for processing an audio signal and method thereof, by which gain and panning of an object can be controlled based on a selection made by a user.

ADVANTAGEOUS EFFECTS

Accordingly, the present invention provides the following effects or advantages.

First of all, according to the present invention, gain and panning of an object can be controlled without restriction.

Secondly, according to the present invention, gain and panning of an object can be controlled based on a selection made by a user.

Thirdly, according to the present invention, gain and panning of an object can be controlled no matter what a downmix signal is a mono signal or a stereo signal.

DESCRIPTION OF DRAWINGS

The accompanying drawings, which are included to provide a further understanding of the invention and are incorporated in and constitute a part of this specification, illustrate embodiments of the invention and together with the description serve to explain the principles of the invention.

In the drawings:

FIG. 1 is a block diagram of an audio signal processing apparatus according to an embodiment of the present invention;

FIG. 2 is a detailed block diagram of an information generating unit of an audio signal processing apparatus according to an embodiment of the present invention; and

FIG. 3 and FIG. 4 are flowcharts for an audio signal processing method according to an embodiment of the present invention.

BEST MODE

Additional features and advantages of the invention will be set forth in the description which follows, and in part will be apparent from the description, or may be learned by practice of the invention. The objectives and other advantages of the invention will be realized and attained by the structure particularly pointed out in the written description and claims thereof as well as the appended drawings.

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 an audio signal according to the present invention includes receiving downmix information, object information and mix information, generating and transferring multi-channel information using at least one of the downmix information, the object information and the mix information, and selectively generating and transferring either first gain information or extra multi-channel information including second gain information in accordance with a decoding mode using at least one of the object information and the mix information.

According to the present invention, the method can

further include generating a multi-channel audio using either the first gain information or the extra multi¬channel information including the second gain information, the multi-channel information and the downmix information.

According to the present invention, the object information includes at least one of object level information and object correlation information.

According to the present invention, the multi-channel information corresponds to information for upmixing the downmix signal into the multi-channel signal and the multi¬channel information is generated using the object information and the mix information.

According to the present invention, the multi-channel information includes at least one of channel level information and channel correlation information.

According to the present invention, the first gain information is calculated per a time-subband variant.

According to the present invention, the first gain information indicates a ratio of a user gain calculated based on the object information and the mix information to an object level calculated from the object information.

According to the present invention, the multi-channel information and the first gain information are transferred together.

According to the present invention, the extra multi-channel information corresponds to HRTF information for binaural .

According to the present invention, generating either the first gain information or the extra multi-channel information includes if the decoding mode is not a binaural mode, generating the first gain information and if the decoding mode is the binaural mode, generating the extra multi-channel information.

According to the present invention, the HRTF information includes HRTF parameter and the object information.

According to the present invention, the HRTF parameter corresponds to a parameter extracted from an HRTF database.

According to the present invention, the second gain information corresponds to information for controlling a per-object level and the second gain information is generated based on the mix information.

According to the present invention, if the downmix signal corresponds to a mono signal, the method further includes bypassing the downmix signal, wherein in generating either the first gain information or the extra multi-channel information, if the decoding mode is not a

binaural mode, the first gain information is generated and wherein in generating either the first gain information or the extra multi-channel information, if the decoding mode is the binaural mode, the extra multi-channel information is generated.

According to the present invention, the method further includes if a channel number of the downmix signal is at least two, generating downmix processing information using at least one of the object information and the mix information and processing the downmix signal using the downmix processing information, wherein in generating either the first gain information or the extra multi-channel information, if the decoding mode is a binaural mode, the extra multi-channel information is generated.

According to the present invention, the mix information is generated based on at least one of object position information, object gain information and playback configuration information.

According to the present invention, the downmix signal is received via a broadcast signal.

According to the present invention, the downmix signal is received on a digital medium.

To further achieve these and other advantages and in accordance with the purpose of the present invention, a

computer-readable recording medium according to the present invention includes a program recorded therein, wherein the program is provided for executing receiving downmix information, object information and mix information, generating and transferring multi-channel information using at least one of the downmix information, the object information and the mix information, and selectively generating and transferring either first gain information or extra multi-channel information including second gain information in accordance with a decoding mode using at least one of the object information and the mix information.

To further achieve these and other advantages and in accordance with the purpose of the present invention, an apparatus for processing an audio signal according to the present invention includes an information receiving unit receiving downmix information, object information and mix information, an information generating unit generating multi-channel information using at least one of the downmix information, the object information and the mix information, the information generating unit selectively generating either first gain information or extra multi-channel information including second gain information in accordance with a decoding mode using at least one of the object information and the mix information, and an information

transferring unit transferring the multi-channel information, the information transferring unit transferring either the first gain information or the extra multi-channel information including the second gain information in accordance with the decoding mode.

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

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.

In this disclosure, information means a terminology that covers values, parameters, coefficients, elements and the like overall. So, its meaning can be construed different for each case. This does not put limitation on the present invention.

And, a multi-channel audio signal of the present invention is to be understood as a concept that includes a channel signal having a stereo effect (3D effect, binaural effect) applied thereto as well as a 3-channel or higher signal.

FIG. 1 is a block diagram of an audio signal processing apparatus according to an embodiment of the present invention.

Referring to FIG. 1, an audio signal processing apparatus 100 according to an embodiment of the present invention includes an information generating unit 110, a downmix processing unit 120, and a multi-channel decoder 130.

The information generating unit 110 receives side information including object information and mix information. The information generating unit 110 generates first gain information or extra multi-channel information (EMI) using the received information. In this case, an extra multi-channel parameter (EMI) includes HRTF (head-related transfer functions) information for a binaural mode and second gain information. Meanwhile, details for the object information (01), the mix information (MXI), the first gain information, the extra multi-channel information (EMI) and the like will be explained later with reference to FIG. 2. Moreover, in case of generating the first gain information, the information generating unit 110 transfers multi-channel information (MI) including the first gain information to the multi-channel decoder 130. In case of not generating the first gain information, the information

generating unit 110 transfers multi-channel information (MI) excluding the first gain information and the extra multi-channel information (EMI) to the multi-channel decoder 130. Its details will be explained later with reference to FIG. 2. In addition, the information generating unit 110 is capable of generating downmix processing information (DPI) using the object information (01) and the mix information (MXI).

The downmix processing unit 120 receives downmix information (hereinafter named 'downmix signal (DMX)') and then processes the downmix signal DMX using downmix processing information (DPI). In case that the downmix signal (DMX) corresponds to a mono signal, the downmix processing unit 120 bypasses the downmix signal (DMX) without processing it. In this case, in order to adjust a gain of the downmix signal (DMX), the information generating unit 110 is able to generate the first gain information. Meanwhile, in case that a channel number of the downmix signal (DMX) corresponds to at least two (i.e., the downmix signal is not a mono signal but a stereo or multi-channel signal), information for adjusting gain and panning of object may be included in the downmix processing information (DPI) or the extra multi-channel information (EMI) instead of being included in the first gain

information. This will be explained in detail later.

The multi-channel decoder 130 receives a processed downmix. The multi-channel decoder 130 generates a multi-channel signal by upmixing the processed downmix signal using the multi-channel information (MI). In case that the extra multi-channel information (EMI) is received, the multi-channel decoder 30 modifies the multi-channel signal using the received extra multi-channel information (EMI).

FIG. 2 is a detailed block diagram of an information generating unit of an audio signal processing apparatus according to an embodiment of the present invention.

Referring to FIG. 2, an information generating unit 110 includes an information receiving unit 112, a multi-channel information generating unit 114, a first gain information generating unit 114a, an extra multi-channel information generating unit 116, and an information transferring unit 118. Meanwhile, the information generating unit 110 may include the information receiving unit 112 and the information transferring unit 118. Alternatively, the information receiving unit 112 and the information transferring unit 118 may correspond to elements configured separate from the information generating unit 110. Moreover, the multi-channel information generating unit 114 may include the first gain

information generating unit 114a, which does not restrict various implementations of the present invention.

The information receiving unit 112 receives object information (01) via a broadcast signal, a digital medium or the like. In this case, the object information (01) may be the information extracted from the aforesaid side information. The object information (01) is information on objects included within a downmix signal and may include object level information, object correlation information and the like. Meanwhile, the information receiving unit 112 receives mix information (MXI) via a user interface or the like. In this case, the mix information (MXI) is the information generated based on object position information, object gain information, playback configuration information and the like. In particular, the object position information is the information inputted for a user to control position or panning of each object. The object gian information is the information inputted for a user to control gain fo each object. The playback configuration information is the information that includes the number of speakers, a position of each speaker, ambient information (virtual position of speaker) and the like. And, the playback configuration information can be inputted by a user, stored in advance or received from other devices.

The multi-channel information generating unit 114 generates multi-channel information (MI) using the object information (01) and the mix information (MXI). In this case, the multi-channel information (MI) is the information for upmixing a downmix signal (DMX) and may include channel level information, channel correlation information and the like.

The first gain information generating unit 114a generates first gain information using the object information (01) and the mix information (MXI). In this case, the first gain information is the information for modifying a gain of the downmix signal (DMX) and can be called a gain modifying factor or an arbitrary downmix gain (ADG). The first gain information can be represented as a ratio of a user gain estimated based on the object information (01) and the mix information (MXI) to an object level estimated from the object information (01). And, the first gain information can be calculated per a time-subband. If the first gain information is applied to the downmix signal (DMX), prior to upmixing the downmix signal (DMX), it is able to adjust a gain of the downmix signal per a specific time and per a specific frequency band. Hence, it is able to adjust a gain of each object accoding to user's control.

Meanwhile, in case that a downmix (DMX) is a mono signal, the first gain information generating unit 114a is able to generate first gain information. Furthermore, in case that a downmix signal (DMX) is a mono signal, when the extra multi- channel information generating unit 116 does not generate HRTF information for a binaural mode, the first gain information generating unit 114a is able to generate first gain information. In case that HTRF information for a binaural mode is generated, second gain information for adjusting an object gain can be included within the HRTF information. So, if the first gain information for adjusting a gain of object is generated, generation and transport of gain information may be overlapped. Details for the binaural mode and the like will be explained later together with the extra multi-channel generating unit 116 .

The extra multi-channel generating unit 116 generates extra multi-channel information (EMI) using object information (01) , mix information (MXI) and an HRTF database. The extra multi-channel information (EMI) may include HTRF information for binaural mode. In this case, the binaural mode is a processing mode for 3-dimensional stereo sound in a channel-oriented decoding scheme (e.g., MPEG Surround) .

Meanwhile, the HRTF information may include: 1) second gain information; 2) HRTF parameter; and 3) object information. In this case, the second gain information is the information for controlling a object gain and may be estimated based on mix information (MXI). And, the HRTF parameter may be the parameter extracted from the HTRF database. Since it is able to independently use the HRTF information for each decoder, an audio signal can be effectively decoded using the HRTF information. The object information may be object information (01) received via the information receiving unit 112.

Besides, it is able to assume that objects signals are controlled in a manner of Formula 1.

[Formula 1]

$$L_{new} = a, xobj_{]} + a_2 x\,obj_2 + a_3\, x\,obj_3 + \cdots + a_n \chi\, obj_n,$$
$$R_{new} = b, xobj, + b_2\, XObJ_2 + b_3\, x\,obj_3 + \text{--} + b_n\, XObJ_n$$

In this case, L_{new} and R_{new} indicate signals desired by a user. And, Obj_k indicate information representing characteristic (energy, correlation, etc.) of object and may be the information extracted from the aforesaid object information (OI). Moreover, a_k and b_k are coefficients for object control and may be the information extracted mix information (MXI) inputted by a user. To correspond to a_k

and b_k, the first gain information or the HRTF parameter can be set.

In particular, Formula 1 can be represented as Formula 2 as well.

[Formula 2]

$$L_{new} = \sum HRTF \times ch$$

In this case, 'HRTF' indicates an HRTF parameter and 'Ch' indicates a channel signal.

Besides, the following is possible.

[Formula 3]

$$L_{new} = \sum \tilde{HRTF} \times ch$$

In this case, is a factor to adjust a gain and may correspond to second gain information.

Meanwhile, in the MPEG Surround standard (5-l-5i configuration) (from ISO/IEC FDIS 23003-1:2006 (E), Information Technology - MPEG Audio Technologies - Parti: MPEG Surround), binaural processing can be represented as follows.

[Formula 4]

$$\mathbf{y}_{\mathrm{B}}^{n,k} = \begin{bmatrix} y_{\mathrm{L_B}}^{n,k} \\ y_{\mathrm{R_B}}^{n,k} \end{bmatrix} = \mathbf{H}_2^{n,k} \begin{bmatrix} y_{\mathrm{m}}^{n,k} \\ D\left(y_{\mathrm{m}}^{n,k}\right) \end{bmatrix} = \begin{bmatrix} h_{11}^{n,k} & h_{12}^{n,k} \\ h_{21}^{n,k} & h_{22}^{n,k} \end{bmatrix} \begin{bmatrix} y_{\mathrm{m}}^{n,k} \\ D\left(y_{\mathrm{m}}^{n,k}\right) \end{bmatrix}, \quad 0 \le k < K$$

In this case, 'y_B' is an output signal and a matrix H is a transform matrix for performing a binaural processing.

And, the matrix H can be expressed as follows .

[Formula 5]

$$\mathbf{H}_1^{l,m} = \begin{bmatrix} h_{11}^{l,m} & h_{12}^{l,m} \\ h_{21}^{l,m} & -\left(h_{12}^{l,m}\right)^* \end{bmatrix}, \quad 0 \le m < M_{\text{Proc}}, 0 \le l < L$$

Each component of the matrix H can be defined as follows .

[Formula 6]

$$h_{11}^{l,m} = \sigma_{\text{L}}^{l,m}\left(\cos\left(IPD_{\text{B}}^{l,m}/2\right) + j\sin\left(IPD_{\text{B}}^{l,m}/2\right)\right)\left(iid^{l,m} + ICC_{\text{B}}^{l,m}\right)d^{l,m},$$

$$h_{12}^{l,m} = \sigma_{\text{L}}^{l,m}\left(\cos\left(IPD_{\text{B}}^{l,m}/2\right) + j\sin\left(IPD_{\text{B}}^{l,m}/2\right)\right)\sqrt{1-\left(\left(iid^{l,m} + ICC_{\text{B}}^{l,m}\right)d^{l,m}\right)^2}$$

$$h_{21}^{l,m} = \sigma_{\text{R}}^{l,m}\left(\cos\left(IPD_{\text{B}}^{l,m}/2\right) - j\sin\left(IPD_{\text{B}}^{l,m}/2\right)\right)\left(1 + iid^{l,m}\,ICC_{\text{B}}^{l,m}\right)d^{l,m}$$

[Formula 7]

$$\begin{aligned}\left(\sigma_{\text{X}}^{l,m}\right)^2 = {} & \left(P_{\text{X,C}}^{m}\right)^2\left(\sigma_{\text{C}}^{l,m}\right)^2 + \left(P_{\text{X,L}}^{m}\right)^2\left(\sigma_{\text{L}}^{l,m}\right)^2 + \left(P_{\text{X,Ls}}^{m}\right)^2\left(\sigma_{\text{Ls}}^{l,m}\right)^2 + \left(P_{\text{X,R}}^{m}\right)^2\left(\sigma_{\text{R}}^{l,m}\right)^2 + \left(P_{\text{X,Rs}}^{m}\right)^2\left(\sigma_{\text{Rs}}^{l,m}\right)^2 + \ldots \\ & P_{\text{X,L}}^{m}P_{\text{X,R}}^{m}\rho_{\text{L}}^{m}\sigma_{\text{L}}^{l,m}\sigma_{\text{R}}^{l,m}ICC_3^{l,m}\cos\left(\phi_{\text{L}}^{m}\right) + \ldots \\ & P_{\text{X,L}}^{m}P_{\text{X,R}}^{m}\rho_{\text{R}}^{m}\sigma_{\text{L}}^{l,m}\sigma_{\text{R}}^{l,m}ICC_3^{l,m}\cos\left(\phi_{\text{R}}^{m}\right) + \ldots \\ & P_{\text{X,Ls}}^{m}P_{\text{X,Rs}}^{m}\rho_{\text{Ls}}^{m}\sigma_{\text{Ls}}^{l,m}\sigma_{\text{Rs}}^{l,m}ICC_2^{l,m}\cos\left(\phi_{\text{Ls}}^{m}\right) + \ldots \\ & P_{\text{X,Ls}}^{m}P_{\text{X,Rs}}^{m}\rho_{\text{Rs}}^{m}\sigma_{\text{Ls}}^{l,m}\sigma_{\text{Rs}}^{l,m}ICC_2^{l,m}\cos\left(\phi_{\text{Rs}}^{m}\right)\end{aligned}$$

[Formula 8]

$$\begin{aligned}\left(\sigma_{\text{L}}^{l,m}\right)^2 &= r_1\left(CLD_0^{l,m}\right)r_1\left(CLD_1^{l,m}\right)r_1\left(CLD_3^{l,m}\right) \\ \left(\sigma_{\text{R}}^{l,m}\right)^2 &= r_1\left(CLD_0^{l,m}\right)r_1\left(CLD_1^{l,m}\right)r_2\left(CLD_3^{l,m}\right) \\ \left(\sigma_{\text{C}}^{l,m}\right)^2 &= r_1\left(CLD_0^{l,m}\right)r_2\left(CLD_1^{l,m}\right)/g_c^2 \\ \left(\sigma_{\text{Ls}}^{l,m}\right)^2 &= r_2\left(CLD_0^{l,m}\right)r_1\left(CLD_2^{l,m}\right)/g_s^2 \\ \left(\sigma_{\text{Rs}}^{l,m}\right)^2 &= r_2\left(CLD_0^{l,m}\right)r_2\left(CLD_2^{l,m}\right)/g_s^2\end{aligned}$$

with $r_1(CLD) = \dfrac{10^{CLD/10}}{1+10^{CLD/10}}$ and $r_2(CLD) = \dfrac{1}{1+10^{CLD/10}}$.

In Formula 7, '$P_{X,C}$', '$P\chi_{,L}$' and the like are factors corresponding to HTRF parameters and can correspond to the second gain information in Formula 3. And, 'σ_C', 'σ_L' and the like in Formula 7 are factors indicating channel power and can correspond to the object power in Formula 1. Thus, since the correspondent relation is effected, it is able to generate a signal specified by a user using the HRTF parameters. In other words, it is able to generate output by applying HRTF parameter to value corresponding to each channel given by the Formulas.

The information transferring unit 118 transfers multi-channel information (MI) and also transfers either the first gain information or the extra multi-channel information (EMI). In particular, in case that the first gain information is generated by the first gain information generating unit 114a, the information transferring unit 118 transfers the multi-channel information including the first gain information. In case that the extra multi-channel information (EMI) is generated by the extra multi-channel information generating unit 116, the information transferring unit 118 transfers the multi-channel information (MI) excluding the first gain information and the extra multi-channel information (EMI). In this case, it is to be understood that it is able to transfer the first

gain information of default instead of excluding the first gain information from the multi- channel information (MI) .

Meanwhile, in case that the extra multi-channel information (EMI) including the HRTF information is transferred, the information transferring unit 118 transfers a specific HRTF parameter once and is then able to transfer information (e.g., index) capable of identifying the specific HRTF parameter.

After a bit stream matching a syntax of a channel-oriented standard (e.g., MPEG Surround) has been generated using the multi-channel information (MI) and the first gain information, the information transferring until 118 is able to transfer the generated bit stream. This does not put limitation on various implementations of the present invention.

FIG. 3 is a flowchart for an audio signal processing method according to an embodiment of the present invention.

Referring to FIG. 3, a downmix signal (DMX), object information (01) and mix information (MXI) are received [SIlO] . Multi-channel information is generated and then transferred using the object information (01) and the mix information (MXI) [S120] . If the downmix signal is not a mono signal ('no' in the step S130) (i.e., the downmix signal is a stereo signal) , steps S210 to S240 are executed.

This will be explained in detail later with reference to FIG. 4. In case that first gain information is generated regardless of whether the downmix signal is a mono signal or a stereo signal, it is a matter of course that the step S130 and the steps S210 to S240 can be omitted.

Meanwhile, in case that the downmix signal is the mono signal ('yes' in the step S130), it is decided whether information for a binaural mode will be generated or not [S140]. If the information for the binaural mode is not to be generated ('no' in the step S140), first gain information is generated for controlling an object gain [S150]. Subsequently, multi-channel information (MI) including the first gain information is transferred [S170]. In this case, the first gain information can be transferred together with the multi-channel information of the step S120. A multi-channel decoder receives the multi-channel information and is then able to control a gain of the downmix signal by applying the received multi-channel information.

In case that the information for the binaural mode is generated in the step S140 ('yes' in the step S140), HTRF information including second gain information, HRTF parameter and object parameter is generated using object information, mix information, HRTF database and the like

[S170] . Subsequently, extra multi-channel information (EMI) including the second gain information is transferred [S180] .

In case that the downmix signal is not the mono signal in the step S130, downmix processing information is preferentially generated using the object information (01) and the mix information (MXI) [S210] . A downmix is processed using the downmix processing information (DPI) generated in the step S210 [S220] . In case of the binaural mode ('yes' in the step S230) , the aforesaid steps S170 and S180 are executed. If it is not the binaural mode ('no' in the step S230) , all procedures are ended.

While the present invention has been described and illustrated herein with reference to the preferred embodiments thereof, it will be apparent to those skilled in the art that various modifications and variations can be made therein without departing from the spirit and scope of the invention. Thus, it is intended that the present invention covers the modifications and variations of this invention that come within the scope of the appended claims and their equivalents.

INDUSTRIAL APPLICABILITY

Accordingly, the present invention is applicable to a process for encoding/decoding an audio signal.

WHAT IS CLAIMED IS:

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

receiving downmix information, object information and mix information;

generating and transferring multi-channel information using at least one of the downmix information, the object information and the mix information; and

selectively generating and transferring either first gain information or extra multi-channel information including second gain information in accordance with a decoding mode using at least one of the object information and the mix information.

2. The method of claim 1, further comprising generating a multi-channel audio using either the first gain information or the extra multi-channel information including the second gain information, the multi-channel information and the downmix information.

3. The method of claim I_7 wherein the object information includes at least one of object level information and object correlation information.

4. The method of claim 1, wherein the multi-channel information corresponds to information for upmixing the downmix signal into the multi-channel signal and the multi-channel information is generated using the object information and the mix information.

5. The method of claim 4, wherein the multi-channel information includes at least one of channel level information and channel correlation information.

6. The method of claim 1, wherein the first gain information is calculated per a time-subband.

7. The method of claim 1, wherein the first gain information indicates a ratio of a user gain calculated based on the object information and the mix information to an object level calculated from the object information.

8. The method of claim 1, wherein the multi-channel information and the first gain information are transferred together.

9. The method of claim 1, wherein the extra multi-

channel information corresponds to HRTF information for binaural .

10. The method of claim 9, generating either the first gain information or the extra multi-channel information, comprising:

if the decoding mode is not a binaural mode, generating the first gain information; and

if the decoding mode is the binaural mode, generating the extra multi-channel information.

11. The method of claim 9, wherein the HRTF information includes HRTF parameter and the object information.

12. The method of claim 11, wherein the HRTF parameter corresponds to a parameter extracted from an HRTF database .

13. The method of claim 1, wherein the second gain information corresponds to information for controlling a object level, and the second gain information is generated based on the mix information.

14. The method of claim 1, if the downmix signal corresponds to a mono signal, the method further comprising bypassing the downmix signal,

wherein the generating either the first gain information or the extra multi-channel information, comprising:

if the decoding mode is not a binaural mode, generating the first gain information and

if the decoding mode is the binaural mode, generating the extra multi-channel information.

15. The method of claim 1, further comprising:

if a channel number of the downmix signal is at least two, generating downmix processing information using at least one of the object information and the mix information; and

processing the downmix signal using the downmix processing information,

wherein the generating either the first gain information or the extra multi-channel information, comprising:

if the decoding mode is a binaural mode, generating the extra multi-channel information.

16. The method of claim 1, wherein the mix information is generated based on at least one of object position information, object gain information and playback configuration information.

17. The method of claim 1, wherein the downmix information is received via a broadcast signal.

18. The method of claim 1, wherein the downmix information is received on a digital medium.

19. A computer-readable recording medium, comprising a program recorded therein, the program provided for executing:

receiving downmix information, object information and mix information;

generating and transferring multi-channel information using at least one of the downmix information, the object information and the mix information; and

selectively generating and transferring either first gain information or extra multi-channel information including second gain information in accordance with a decoding mode using at least one of the object information and the mix information.

20. An apparatus for processing an audio signal, comprising:

an information receiving unit receiving downmix information, object information and mix information;

an information generating unit generating multi¬channel information using at least one of the downmix information, the object information and the mix information, the information generating unit selectively generating either first gain information or extra multi-channel information including second gain information in accordance with a decoding mode using at least one of the object information and the mix information; and

an information transferring unit transferring the multi-channel information, the information transferring unit transferring either the first gain information or the extra multi-channel information including the second gain information in accordance with the decoding mode.

FIG. 1

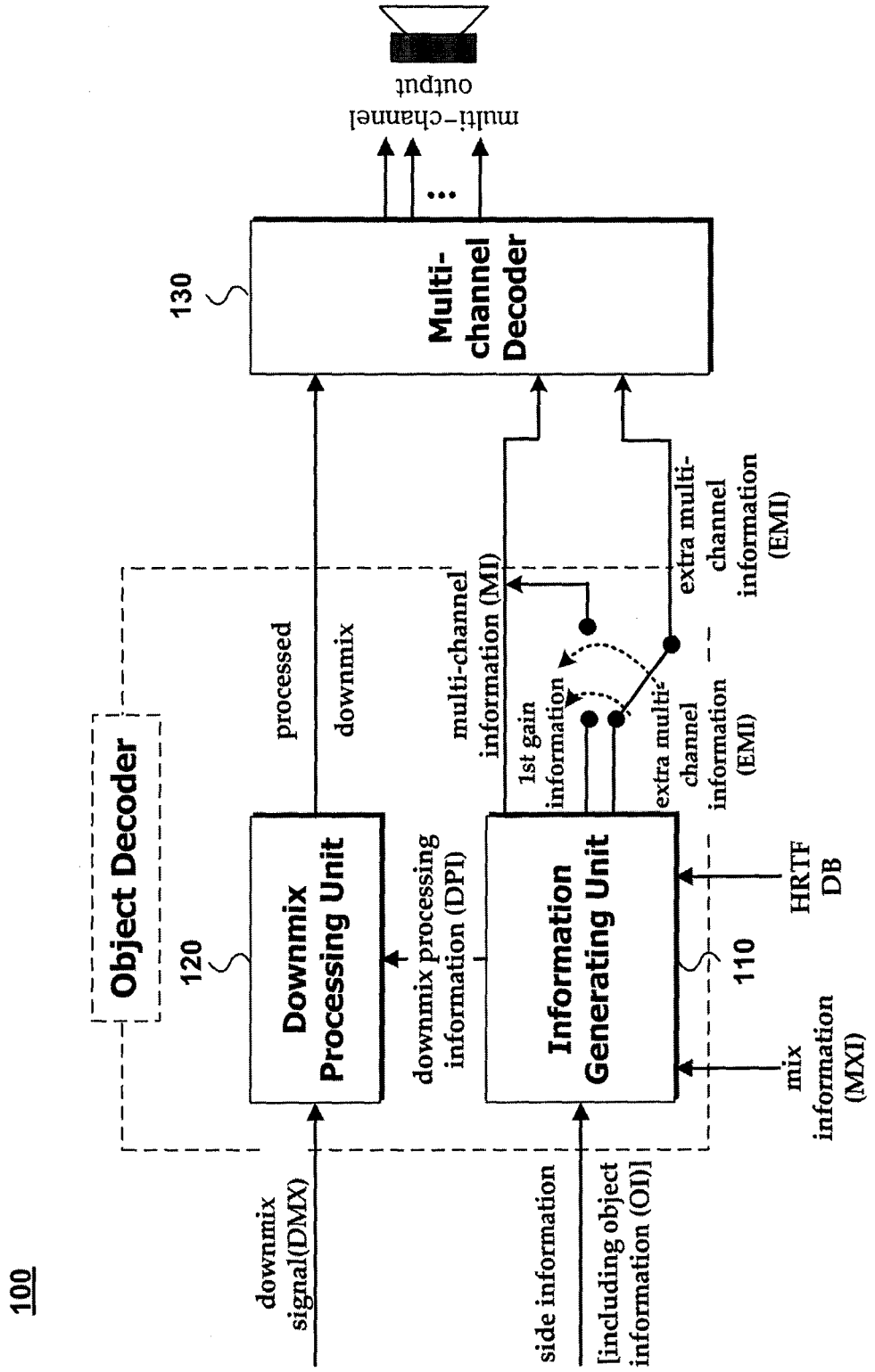

100
Object Decoder
120
Downmix Processing Unit
Information Generating Unit
110
Multi-channel Decoder
130
multi-channel output
downmix signal(DMX)
side information [including object information (OI)]
processed downmix
downmix processing information (DPI)
multi-channel information (MI)
1st gain information
extra multi-channel information (EMI)
extra multi-channel information (EMI)
HRTF DB
mix information (MXI)

FIG. 2

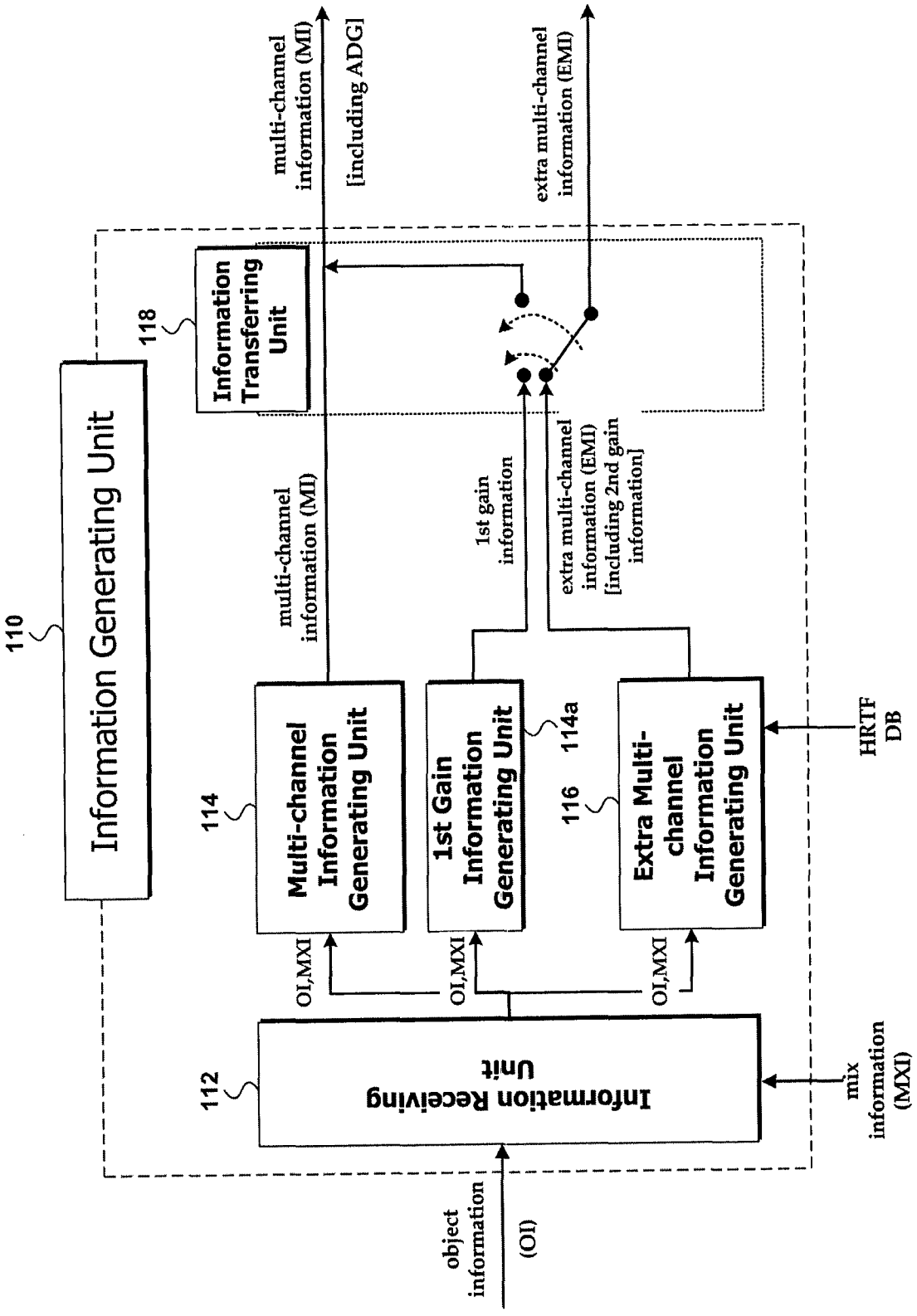

Information Generating Unit
110
Information Transferring Unit
118
Multi-channel Information Generating Unit
114
1st Gain Information Generating Unit
114a
Extra Multi-channel Information Generating Unit
116
Information Receiving Unit
112
multi-channel information (MI)
[including ADG]
extra multi-channel information (EMI)
multi-channel information (MI)
1st gain information
extra multi-channel information (EMI) [including 2nd gain information]
OI,MXI
OI,MXI
OI,MXI
HRTF DB
mix information (MXI)
object information (OI)

FIG. 3

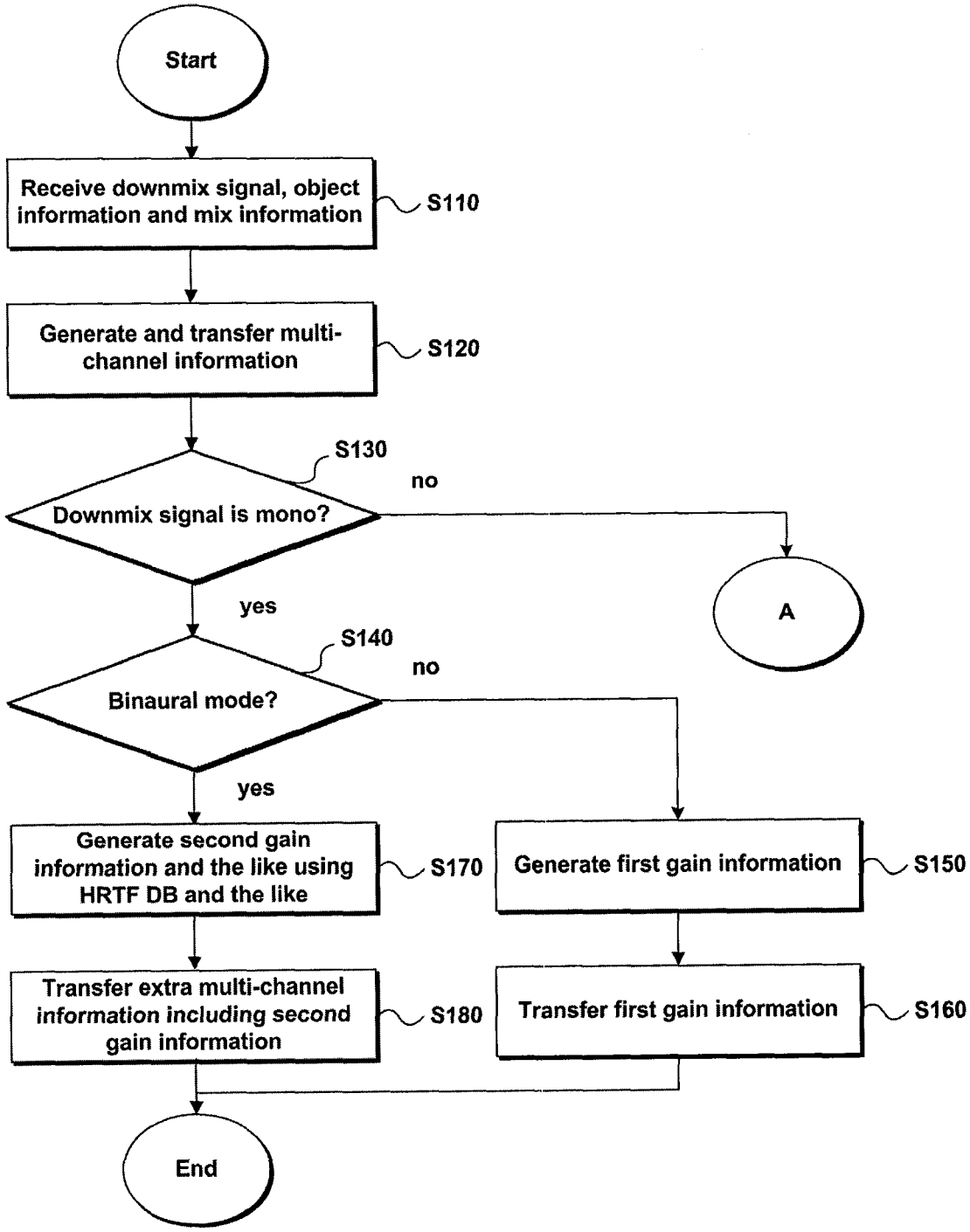
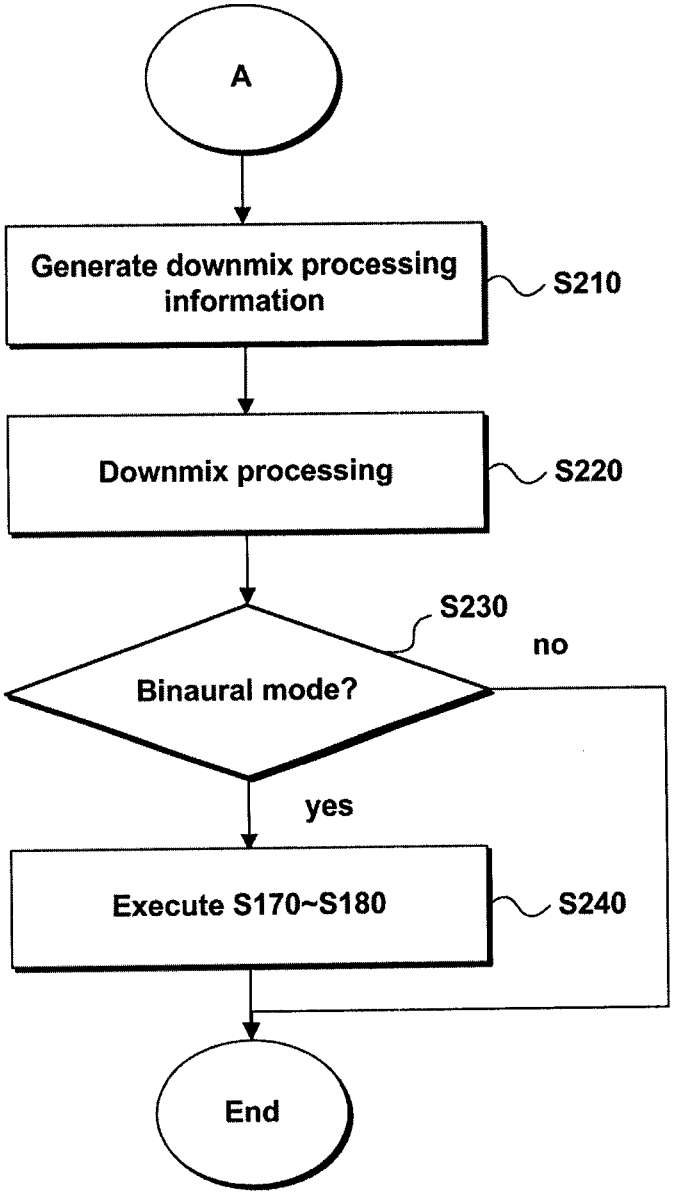

Start
Receive downmix signal, object information and mix information
S110
Generate and transfer multi-channel information
S120
S130
Downmix signal is mono?
no
yes
A
S140
Binaural mode?
no
yes
Generate second gain information and the like using HRTF DB and the like
S170
Generate first gain information
S150
Transfer extra multi-channel information including second gain information
S180
Transfer first gain information
S160
End

FIG. 4


A
Generate downmix processing information
S210
Downmix processing
S220
S230
Binaural mode?
no
yes
Execute S170~S180
S240
End

INTERNATIONAL SEARCH REPORT

International application No
PCT/KR2008/000073

A. CLASSIFICATION OF SUBJECT MATTER

GIOL 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 H04S 3/00 GlOL 21/02 H04R 1/10 H04R 3/00 H04R 25/00 H04S3/02 GlOL 21/00 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) "AUDIO"OBJECT""DOWNMFX""GAIN""PAN*"

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
A	WO 2006/008683A1 (KONINKLL JKE PIFILIPS ELECTRONICS N V et al) 26 JANUARY 2006 See Abstract, Fig 1 and Description	1, 19, 20
A	WO 2006/132857 A2 (DOLBY LABORA TPTIES LICENSING CO) 14 DECEMBER 2006 See the whole document	1, 19, 20
A	EP 1 640 972 A l (PHONAK AG) 29 MARCH 2006 See Column 2 Paragraph 11 ~ Column 3 Paragraph 12, Column 6 Paragraph 31 ~ Column 7 Paragraph 32	1
A	EP 1 691 348 A l (ECOLE POL YTECHNIQUE FEDERALE DE LAUSANNE) 16 AUGUST 2006 See the whole document	1
A	US 2006/0085200 A l (ALLAMANCHE et al) 20 APRIL 2006 See Page 3 Paragraph 52 ~ Page 8 Paragraph 99	1
A	US 2005/195981 A l (FALLER et al) 08 SEPTEMBER 2005 See the whole document	1

☐ 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
22 APRIL 2008 (22 04 2008)	22 APRIL 2008 (22.04.2008)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon, 139 Seonsa-ro, Seo-gu, Daejeon 302-701, Republic of Korea

Facsimile No 82-42-472-7140

Authorized officer

SUH, Hawthorne

Telephone No 82-42-481-5670

Form PCT/ISA/210 (second sheet) (April 2007)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/KR2008/000073

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
W02006008683A 1	26.01 .2006	CN 1985544A	20.06.2007
		EP01769655A 1	04.04.2007
		EP1769655A 1	04.04.2007
		KR 1020070039543	12.04.2007
		KR2007039543A	12.04.2007
		US200702307 10A 1	04. 10.2007
		US20072307 10A 1	04. 10.2007
		US20072307 10AA	04. 10.2007
W02006132857A2	14. 12.2006	AU2006255662A 1	14. 12.2006
		CA2610430A 1	14. 12.2006
		KR 10200800 15886	20.02.2008
		W02006132857A3	24.05.2007
EP01640972A 1	29.03.2006	EP1640972A 1	29.03.2006
		W020070738 18A 1	05.07.2007
EP01691348A 1	16.08.2006	AU200621219 1A 1	17.08.2006
		AU200621219 1AA	17.08.2006
		CA2597746AA	17.08.2006
		EP01844465A2	17. 10.2007
		EP1691348A 1	16.08.2006
		EP1844465A2	17. 10.2007
		W020060849 16A2	17.08.2006
		W020060849 16A3	08.03.2007
US2006085200A 1	20.04.2006	AU2005299070AA	04.05.2006
		CA2583146AA	04.05.2006
		EP01803325A 1	04.07.2007
		EP1803325A 1	04.07.2007
		KR2007061882A	14.06.2007
		N020071492A	19.07.2007
		US2006085200AA	20.04.2006
		W02006045373A 1	04.05.2006
US200519598 1A 1	08.09.2005	AT373402E	15.09.2007
		AU2005226536A 1	06. 10.2005
		AU2005226536AA	06. 10.2005
		BR200508146A	3 1.07.2007
		CA2557993AA	06. 10.2005
		CA2557993A 1	06. 10.2005
		EP01721489A 1	15. 11.2006
		EP01721489B 1	12.09.2007
		EP1721489A 1	15. 11.2006
		EP1721489B 1	12.09.2007

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/KR2008/000073

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		JP19526520	13.09.2007
		JP2007526520T2	13.09.2007
		KR1020060131866	20.12.2006
		KR2006131866A	20.12.2006
		N020064472A	03.10.2006
		US2005195981AA	08.09.2005
		W02005094125A1	06.10.2005