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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 May 2008 (22.05.2008)

PCT

(10) International Publication Number
WO 2008/060111 A1

(51) International Patent Classification:
G10L 19/00 (2006.01) *H03M 7/30* (2006.01)

(21) International Application Number:
PCT/KR2007/005740

(22) International Filing Date:
15 November 2007 (15.11.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/865,908	15 November 2006 (15.11.2006)	US
60/869,077	7 December 2006 (07.12.2006)	US
60/869,080	7 December 2006 (07.12.2006)	US
60/883,567	5 January 2007 (05.01.2007)	US
60/889,715	13 February 2007 (13.02.2007)	US
60/955,395	13 August 2007 (13.08.2007)	US

(71) Applicant *(for all designated States except US)*: **LG ELECTRONICS INC.** [KR/KR]; 20, Yeouido-dong, Yeongdeungpo-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). **JUNG, Yang Won** [KR/KR]; 202, #287-4 Yeonhui 3-dang, Seodaemun-gu, Seoul 120-830 (KR).

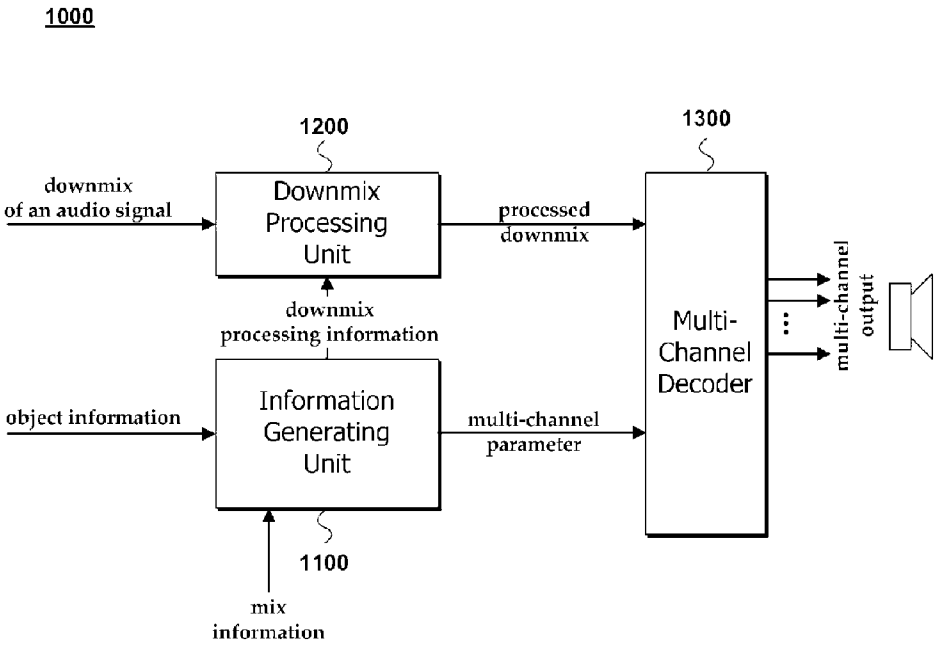
(74) Agents: **KIM, Yong In** et al.; KBK & Associates, 15th Floor, YoSam Building, 648-23 Yeoksam-dong, Kangnam-ku, Seoul 135-748 (KR).

(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, 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, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL,

[Continued on next page]

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



(57) Abstract: The present invention relates to a method and an apparatus for decoding an audio signal, is to provide a method for decoding an audio signal by using object information including an object level information and an object gain information to modify the downmix of an audio signal as changing the contribute of object to each downmix channel. The present invention is to provide an apparatus for decoding an audio signal by using object information including an object level information and an object gain information to modify the downmix of an audio signal as changing the contribute of each object to each downmix channel.

WO 2008/060111 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*

Description

A METHOD AND AN APPARATUS FOR DECODING AN AUDIO SIGNAL

Technical Field

[1] The present invention relates to a method and an apparatus for decoding an audio signal, and more particularly, to a method and an apparatus for decoding an audio signal received via various digital medium.

Background Art

[2] While downmixing several audio objects to a mono or a stereo signal, several informations(or parameters) from the individual object signals can be extracted. These informations can be used in a decoder of an audio signal. An output audio signal of multi-object control unit(MCU) can be generated using the informations corresponding to individual object signals.

[3] MCU(Mutipoint Control Unit) is a device that it can be used teleconference to articulate provided signals from remote place through conference call. Recently, experiments which use convergence techniques increase according to coming into the spotlight of the techniques.

[4] A conventional MCU combiner generally makes combined signal to be received multi-channel audio signals. But, when multi-channel audio signals having only multi-channel parameter are used in MCU, it only can control one of channels gain and panning, not control object gain and panning.

Disclosure of Invention

Technical Problem

[5] A decoder receives a downmix signal and a side information, and can generate an output signal using the side information. The output signal may be rendered based on other input information such as a user control or a playback configuration. In order to control the individual object signals, the decoder may receive multi-object signals and process to decode them.

[6] However, an apparatus and method for decoding whole multi-object signals needs a wide bandwidth. Accordingly, a new apparatus and method for decoding multi-object signals be needed to relieve the resource requirement like the wide bandwidth. More, for backward compatibility in the view of the channel-oriented decoding, an side information corresponding to object which can be converted flexibly to a multi-channel parameter is need.

Technical Solution

[7] Accordingly, the present invention has been made keeping in mind the above

problems, and is directed to a method and an apparatus for decoding an audio signal that substantially improves disadvantages of the related art and obviates one or more problems of related art.

[8] An object of the present invention is to provide a method for decoding an audio signal by using object information including an object level information and an object gain information to modify the downmix of an audio signal as changing the contribute of object to each downmix channel.

[9] The other object of the present invention is to provide an apparatus for decoding an audio signal by using object information including an object level information and an object gain information to modify the downmix of an audio signal as changing the contribute of each object to each downmix channel.

[10] Another object of the present invention is to provide a method and an apparatus for decoding an audio signal, comprising a downmix and a combined object parameter to be made in a MCU combiner, to control object gain and output in a remote conference and so on.

[11] Additional advantages, objects, and features of the invention will be set forth in part in the description which follows and in part will become apparent to those having ordinary skill in the art upon examination of the following or may be learned from practice of the invention. The objectives and other advantages of the invention may be realized and attained by the structure particularly pointed out in the written description and claims hereof as well as the appended drawings.

Advantageous Effects

[12] Various embodiments of the present invention provide a method and an apparatus for decoding multi-object audio signals fast and efficiently by reducing process time, computer resource, thereby relieving the resource requirement like the wide bandwidth. The object parameters according to the embodiments of the present invention can provide backward compatibility in the view of the channel-oriented decoding process.

Brief Description of the Drawings

[13] The accompanying drawings, which are included to provide a further understanding of the invention, illustrate the preferred embodiments of the invention, and together with the description, serve to explain the principles of the present invention. In the drawings;

[14] FIG. 1 is an exemplary block diagram of an apparatus for decoding an audio signal according to one embodiment of the present invention.

[15] FIG. 2 is a flow chart illustrating an audio signal decoding method in accordance with an embodiment of the present invention.

[16] FIG. 3 is an exemplary block diagram of an apparatus for decoding an audio signal according to other embodiment of the present invention.

[17] FIG. 4 is an exemplary block diagram of a parameter generating unit according to one embodiment of the present invention.

[18] FIG. 5 is an exemplary block diagram of a object gain information generating unit according to one embodiment of the present invention.

[19] FIG. 6 is an exemplary block diagram of a parameter generating unit according to other embodiment of the present invention.

[20] FIG. 7 is an exemplary block diagram of an apparatus for processing an audio signal according to other embodiment of the present invention.

[21] FIG. 8 is an exemplary block diagram of a MCU combining unit according to one embodiment of the present invention.

[22] FIG. 9 is an exemplary block diagram of a combined object parameter encoding unit according to one embodiment of the present invention.

Best Mode for Carrying Out the Invention

[23] To achieve there objects and other advantages in accordance with the purpose of the invention, as embodied and broadly described herein, the present invention of decoding method for an audio signal comprises receiving a downmix of an audio signal, an object information, and a mix information, the object information including an object level information, an object correlation information, and an object gain information, the object level information being generated by normalizing object level corresponding to object using one of the object level as a reference information, the object correlation information provided from combination of two selected objects, the object gain information comprising at least one of an object gain value information and an object gain ratio information; generating a downmix processing information using the object information and the mix information; and processing the downmix of an audio signal using the downmix processing information.

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

Mode for the Invention

[25] Reference will now be made in detail to the preferred embodiment of the present invention, examples of which are illustrated in the accompanying drawings. Wherever possible, the same reference numbers will be used throughout the drawings to refer to the same or like parts.

[26] Prior to describing the present invention, it should be noted that most terms disclosed in the present invention correspond to general terms well known in the art, but some

terms have been selected by the application as necessary and will hereinafter be disclosed in the following description of the present invention. Therefore, it is preferable that the terms defined by the applicant be understood on the basis of their m eanings in the present invention.

[27] FIG. 1 is an exemplary block diagram of an apparatus 1000 for decoding an audio signal according to one embodiment of the present invention. Fig. 3 is an exemplary block diagram of an apparatus 2000 for decoding an audio signal according to other embodiment of the present invention.

[28] The two embodiments of the apparatus 1000 and 2000 have a difference in that the apparatus 1000 has a multi-channel decoder 1300 while the apparatus 2000 doesn't have the multi-channel decoder 1300. Other elements, such as a parameter generating unit 1100 and 2000 and a downmix processing unit 1200 and 2200 are the same as that of FIGs. 1 and 3

[29] Referring FIG. 1, an apparatus 1000 for decoding an audio signal (hereinafter simply referred as 'a decoder 1000') include a parameter generating unit 1100, a downmix processing unit 1200, and a multi-channel decoder 1300. The parameter generating unit 1100 is configured to receive an object information and a mix information from an user control or a bitstream, and to generate a downmix processing information.

[30] The object information includes an object level information, an object correlation information, and an object gain information. The object level information can be generated by normalizing an object level corresponding to each object using one of the object levels as a reference information. The object correlation information can be provided from combination of two selected objects. The object gain information includes an object gain value information or an object gain ratio information. The downmix processing information includes a parameter for controlling object gain and object panning, which is inputted to the downmix processing unit 1200.

[31] The downmix processing unit 1200 is configured to receive a downmix of an audio signal with the downmix processing information from the parameter generating unit 1100. The downmix processing unit 1200 can process the downmix using the downmix processing information, thereby generating the processed downmix signal. For example, the downmix processing unit 1200 can apply the downmix processing information to the downmix of the audio signal in order to change one or all of object gain and object position of the downmix of the audio signal to generate the processed downmix.

[32] The processed downmix may be input to the multi-channel decoder 1300 to be upmixed and output.by an output device such as a speaker. A multi-channel parameter output from the parameter generating unit may be also input to the multi-channel decoder 1300 In some embodiments of the present invention, the multi-channel

decoder 1300 can be used as same as a decoder of MPEG Surround system.

[33] Alternatively, the processed downmix signal may be directly transmitted to and output by the output device as the device 2000 shown in FIG. 2. In order to directly output the processed signal via speakers, the downmix processing unit 2200 may perform synthesis filter bank and output PCM data. It is also able to select whether to directly output as PCM signal or input to the multi-channel decoder by user selection.

[34] FIG. 2 shows a flowchart of the present invention, and refer also to the FIG. 1. The method is a flow path of a decoding method for an audio signal. In step S110, a downmix of an audio signal, an object information, and a mix information is received. Step 120 generates a downmix processing information using the object information and the mix information. In step S130, a processed downmix is generated to process the downmix of the audio signal using the downmix processing information.

[35] The configuration of the parameter generating unit 1100 shall be explained in detail with reference to FIG. 4 to FIG.6.

[36] **1.Object information**

[37] **1.1 Reference information and object level information**

[38] FIG. 4 is an exemplary block diagram of an apparatus for processing an audio signal according to one embodiment of present invention, in particular, an exemplary block diagram of the parameter generating unit. Referring to FIG. 4, the parameter generating unit 1100 can be configured to receive an object information, and to generate a downmix processing information using the object parameter.

[39] The parameter generating unit 1100 can include an object level information decoding unit 1110a, an object gain information generating unit 1120a, and an object correlation information generating unit 1130a.

[40] The downmix of an audio signal includes many of object signals, and the object signals have object level each in an object signal.

[41] The object level information is generated by normalizing the object level using reference information, the reference information may be one of the object level, more particular, the reference information may be the largest object level among the all object levels.

[42] For example, it is assumed that the downmix of an audio signal includes object s_i, and the object level of each object s_i is Ps_i

[43] If the object level energies is transmitted as it is to encode the object parameter, the object parameter includes object information as follow:

[44] Ps_i can be obtained as various methods. For example, Ps_i may be "s_i(n)^2" or "E[s_i(n)^2]". Ps_i may be transmitted as the information corresponding to each object level information. Here, "s_i(n)" refers to a i^{th} object signal, and the s_i(n) can be either a time domain signal, or subband signal within a given band.

[45] However, if the object level information corresponding to each object signal is transmitted as the value itself the object level of an object signal may be difficult to be quantized due to an excessive increase in a variation of a dynamic range.

[46] Thus, the object level information may be normalized using the reference information, the largest object level energy of all object energies. If the reference information may be r_1, the object level information may be transmitted as in Math Figure below:

[47] [Math Figure 1]

[48] E[s_i(n)^2]/E[r_1(n)^2], r_1(n)=referenve information

[49] All of the object level information is comprised a range of equal or less than 1.

[50] Therefore, a dynamic range can be compressed enough to encode an audio signal.

[51] Additionally, the object level information may include reference information, default information, original object level energy to use other signal process. The object level information corresponds to each object, and the number of the object level information is same as the number of the objects in the downmix.

[52] **1.2 Object gain information**

[53] The object parameter comprises an object gain information including at least one of an object gain value information and an object gain ratio information. FIG. 5 is an exemplary block diagram of an apparatus for processing an audio signal according to one embodiment of present invention, in particular, an exemplary block diagram of the object gain information decoding unit of the parameter generating unit 1100.

[54] The object gain information generating unit 1120a includes an object gain value information generating unit 1121 and an object gain ratio information generating unit 1122. The object gain information relates to a downmix method one object signal to generate a downmix signal having more than one channel.

[55] **1.2.1 Object gain value information**

[56] The object gain value information comprises a gain value of an object. In some embodiments of the present invention, the object gain is applied to each object before generating the processed downmix.

[57] For example, when the downmix of an audio signal includes a plurality of objects, each object gain value information corresponding to each object is multiplied to the object level of each object to generate each gained object, and all of the gained objects are summed to generate the processed downmix.

[58] [Math Figure 2]

[59] X = sum{a_i * s_i}

[60] where X is processed downmix to be transtmitted to mono channel, s_i is an object level, and a_i is an object gain value information of an object contributing to each channel.

[61] **1.2.2 Object gain ratio information**

[62] The object gain information comprises further the object gain ratio information as well as the object gain value information. The object gain ratio information includes a ratio value between the gains of each object contributing to each channel of the processed downmix.

[63] The object gain ratio information can be used to process the down mix by the Downmix Processing Unit 1200, thereby obtaining the processed downmix to be transmitted through 2(i.e. stereo) and more channels. In the case of the stereo channel, a processed downmix to be transmitted trough each of the stereo channel is shown by Math Figure 3. The object gain ratio information can be obtained from Math Figure 4.

[64] [Math Figure 3]

[65] $x\_1 = sum\{a\_i * s\_i\}$

[66] $x\_2 = sum\{b\_i * s\_i\}$

[67] where x_1 and x_2 are processed downmix to be transtmitted through each channel, respectively, s_i is an object level, and a_i and b_i are an object gain value information of an object contributing to each channel.

[68] [Math Figure 4]

[69] $m\_i = a\_i / b\_i$

[70] where m_i is an object gain ratio information of each object.

[71] The object gain information, i.e.the object gain value information (a_i and b_i) and the object gain ration information (m_i) can be transmitted to parameter generating unit 1100 in various combination of the object gain information contained in a bitstream. The combinations include, for example, (a_i , b_i), (m_i, a_i) and (m_i, b_i). The parameter generating unit 1100 can decode the combinations to reconstruct the original object information. It can be understood that decoding of the combinations performed by the parameter generating unit 1100 can be adapted to the other decoder, for example multi-channel decoder 1300.

[72] Alternatively, when the object gain information is transmitted to the parameter generating unit 1100 in a combination of object gain value information (a_i , b_i), the object gain value informations can be scaled. If there is a convention that b_i be scaled to 1, though object level information and only a_i as an the object gain information is transmitted, the parameter generating unit 1100 can reconstruct the original object information according to the convention. By scaling the object gain value, the number of the parameters to be transmitted to the parameter gererating unit 1100, can be reduced.

[73] Alternatively, the object gain ration information (m_i) can be obtained fromwith a various value as Math Figure 5.

[74] [Math Figure 5]

[75] $m\_i = a\_i / b\_i,$

[76] $m_i = (a_i + \alpha) / (b_i + \beta)$,

[77] $m_i = (a_i * s_i) / (b_i * s_i)$

[78] (α,β is a very small number to prevent a numerator and a denominator to zero.)

[79] In case that the object gain ratio information comprises s_i, same m_i value may not be included same value of s_i. For example, in case of 1) a_i = 0.5, b_i=0.5, 2) a_i = 2, b_i = 2, all of case has same m_i (=1), but the cases have different values of a_i, b_i .

[80] To obtain the processed downmix to be transtmitted through each channel, new method can be used as Math Figure 6:

[81] [Math Figure 6]

[82] $x_1 = sum\{a_i'(n) * s_i'(n)\}$,

[83] $x_2 = sum\{b_i'(n) * s_i'(n)\}$

[84] (wherein a_i' and b_i' are values satisfied the following conditions,

[85] (a_i' + b_i' = C) or (a_i'^2 + b_i'^2 = C) or (a_i' = C or b_i' = C), wherein s_i' = g_i * s_i)

[86] Finally, the object gain ratio information can be transmitted m_i'(= a_i' / b_i'). The number of the parameters to be transmitted to the parameter gererating unit 1100 can be reduced. To prevent the distortion of an audio signal in the decoder 1000 or 2000, m_i can be transmitted.

[87] **1.3 Object correlation information**

[88] Referring to FIG. 4, the parameter decoding unit 1100 receives an object correlation information. The object correlation information is estimated between two objects and represents the correlation/coherence between two objects.

[89] In case that the two objects have same origin of channel and are transmitted through different channels, the object correlation information can be existed.

[90] First, if the object signal includes stereo objects, the stereo objects may generate a mono object downmixing stereo objects, and generate a descendant object parameter indicating relations between channels of the stereo objects(hereinafter, this method is 'mono method'). In this case, the object level information is generated using the object level energy of the mono object.

[91] Second, stereo objects recognizes two individual mono objects signal. In this case, the object level information is generated using the two individual mono objects level(hereinafter, this method is 'stereo method'). The amount of information to be transmitted using the second method has more than that of using the first method.

[92] To process a stereo object, for example, a first channel signal of stereo objects may be s_i, a second channel signal of stereo objects is s_j as each mono object signal.

[93] The object level of above channel signal may be Ps_i, Ps_j.

[94] In case of a stereo object, each object information representing L and R channels of given object is similar to each other. So, the object correlation information can be used

to represent similarity between the objects information.

[95] Therefore, to encode Ps_i and Ps_j, each mono object using stereo method is considered coupling constituted same object.

[96] The object correlation information includes one of channel power as representative, for example, left channel of stereo object, and normalized power value using the representative as follows.

[97] [Math Figure 7]

[98] Ps_j' = Ps_j / Ps_i or

[99] Ps_j' = 10log10(Ps_j) - 10log10(Ps_i) = 10log10(Ps_j/Ps_i)

[100] To reduce the transmitted bits of the object information, it is effective to use the object correlation information.

[101] And the object correlation information can be generated using the representative as follows.

[102] [Math Figure 8]

[103] Ps_i', Ps_j' = Ps_i, Ps_j / sqrt(Ps_i * Ps_j)

[104] The object correlation information represents relation between objects, whether or not the objects are both channels of the same stereo or multi-channel object, that is, each object is a different channel of same origin.

[105] Additionally, regarding to relation between two objects, a differential information can be used.

[106] The different information includes sum or substraction signal of the stereo object as follows.

[107] [Math Figure 9]

[108] M = (L + R)/2, S = (L - R)/2,

[109] Ps_M =(Ps_L + Ps_R)/2, Ps_S = (Ps_L - Ps_R)/2

[110] The object correlation information including above the M and Ps_M can improve transmission efficiency and be easy to perform the error balance.

[111] The number of the object correlation information varies adaptively according to constituted a same object in order to reduce the bit rate of a object parameter. A flag information 'correlation_flag' indicating whether an object is part of a stereo or multi-channel object, and can be received from the object information. The correlation_flag can be included the object information, and received the information generating unit 1100.

[112] Meaning of the flag information 'correlation_flag' is shown in the following Table 1.

[113] Table 1

[Table 1]

[Table]

Correlation_flag	Meaning
1	correlation
0	no correlation

[114] In case that 'correlation_flag'is equal to 0, the object correlation information is not transmitted to the object correlation information decoding unit 1130a. When the 'correlation_flag'is not received to the decoder 1000 or 2000, default value can be used to process the downmix of the audio signal. Otherwise('correlation_flag'is equal to 1), the object correlation information is transmitted to the object correlation information decoding unit 1130a similarity between the selected two objects.

[115] Besides, the object information further includes a reference information separately. When the reference information exists, the reference information can be a identifier for a MCU combiner.

[116] A method of encoding for an audio signal according to the present invention comprises the step of receiving a multi-object audio signal and the step of generating a downmix of an audio signal and an object information including an object level information, an object gain information, and an object correlation, the object level information and the object correlation information from the multi-object audio signal, characteristics of the object level information, the object gain information, and the object correlation is same as that of the decoding method. So, the method of encoding for an audio signal cording to the present invention may not be limited as above identified.

[117] Additionally, an apparatus of encoding for an audio signal according to the present invention comprises a downmixing unit generating a downmix of an audio signal from a multi-object audio signal, and an object information unit extracting an object information including an object level information, an object gain information, and an object correlation information from the multi-object audio signal. The apparatus of encoding for an audio signal may not be limited as above identified.

[118]

[119] **2. MCU Combiner**

[120] An audio signal comprising multi-object signals can be used MCU combiner to control object gain and output in a remote conference and so on. In case that uses the audio signal comprising multi-object signals, it is effective to control object gain and panning corresponding to characteristic of each object signal.

[121] For example, the multi-channel audio signal includes vocal sound, back ground

music(BGM), narration sound. As occasion demands, we can t delect or control a special kind of object signals when we only use or listen back ground music without vocal sound and narration sound or only make a communication with someone in a teleconference.

[122] Additionally, the method of decoding for the present invention using object information may be used an enhanced karaoke system.

[123] FIG. 6 is an exemplary block diagram of an apparatus for processing an audio signal according to an embodiment of present invention. Referring to FIG. 6, an apparatus for processing an audio signal according to embodiment may comprise an encoder 1 3100, an encoder 2 4100, a combining unit 5000 including a MCU combining unit 5100 and downmixer 5200. The encoder 1 3100 and the encoder 2 4100 can be configured to receive each an audio signal_1 or an audio signal_2, and to generate a downmix_1 and an object information_1 in the encoder 1 3100, and to generate a downmxi_2 and an object information_2 in the encoder 2 4100.

[124] The combining unit 5000 can be configured to receive the downmix_1 and the object information_1 from the encoder 1 3100, the downmix_2 and the object information_2 from the encoder 2 4100, and a control information from user control, and to generate a downmix and a combined object information.

[125] The downmix, output signal of the combining unit 5000, can be generated a conventional downmixing unit. Therefore, details of elements of the downmixer 5200 shall be omitted.

[126] 2.1 **combined object paramteter**

[127] FIG. 7 is an exemplary block diagram of an apparatus for processing an audio signal according to an embodiment of present invention, in particular, an exemplary block diagram of an MCU combining unit 8100. Referring to FIG. 7, the MCU combining unit 5100 can be configured to generated a combined object information using the object information_1, the object information_2, and the control information. The combined object information includes all information corresponding to the downmix_1 from the encoder 1 3100 and the downmix_2 from the encoder 2 4100.

[128] The MCU combining unit 5100 includes an object information decoding unit 5110 and a combined object information encoding unit 5120. The object information decoding unit 5110 can be configured to receive the object information_1 from the encoder 1 3100 and the object information_2 from the encoder 2 4100, and to generate a reference value_1, an object level information_1, and an object gain informaiton_1 from the object information_1, and a reference value_2, an object level information_2, and an object gain information_2. The reference values, the object level informations, and the object gain informations is same as that of FIG. 1 ~ FIG. 6. Therefore, details of generating method of those informations shall be omitted.

[129] And the MCU combining unit 5100 can be configured to receive at least two object informations from each multiple encoders without limitation of input signals, and to generate the combined object information comprising several information corresponding to the downmix.

[130] 2.2 **control information**

[131] FIG. 8 is an exemplary block diagram of an apparatus for processing an audio signal according to an embodiment of present invention, in particular, an exemplary block diagram of a combined object information encoding unit 5120. Referring to FIG. 8, the combined object information encoding unit 5120 can be configured to receive those informations and a control information from user control, and to generate a combined object information to be inputted in a decoder(not shown).

[132] The control information may process the object information_1 and the object information_2, and apply to combination of above the object information_1 and the object information_2 in the combined object information encoding unit 5120. The combined object information may be generated to be processed the control information, the control information indicating to combine some objects constituted the combined object information and to control object gain in the combination of the object informations.

[133] The control information includes an object control information, a gain control information, and a destination information. Each of the object control information, the gain control information, and the destination information may explain the followings.

[134] **2.2.1 object control information**

[135] The object control information may determine target objects to generate the combined object information. The object control information can determine a required subset of audio objects of object information_1 or object information_2.

[136] The object control information may be processed to the object level information in the object level information encoding unit 5112. The combined object information may include information corresponding to some objects determining by the object control information, and can be use according to several purposes.

[137] For example, the object information_1 comprises music including vocal, piano, guitar object signals, and the object information_2 comprises violin, vocal object signals. To generate an audio signal comprising piano, guitar, violin object signals, we can obtain the combined object information using the object control information from user control without vocal object signals.

[138] **2.2.2 gain control information**

[139] The object gain information encoding unit 5113 can be configured to receive a gain information_1 from the object information_1, a gain information_2 from the object information_2, a gain control information, and a destination information, and to generate

an object gain information of the object information.

[140] The gain control information may be used to control object gain for MCU combiner. Unlike the object control information, the gain control information may be processed object information in the object gain information encoding unit 5113, the object information is selected using the object control information in the object level information encoding unit 5112. The gain control information may be value within in the range of 0~1.

[141] **2.2.3 destination information**

[142] Among the range of the gain control information, If the gain control information corresponding to object information_i is 0, the object information does not included in the combined object information. When the gain control information is 0 or 1, the gain control information defines a destination information. The destination information may include the special gain control information having 0 or 1 value and the indicators which destinations are to be outputted the downmix.

[143] The destination information can be used for special function, for example, a whisper function, a secret meeting, and for controlling the destination of an object signal.

[144] Referring to the FIG. 8, the destination information may be inputted into the object gain information encoding unit 5123, and process the gain information_1 and the gain information_2 to control object gain of the combined object information. If a MCU combiner has 3-ports, the destination information may include each gain value(0, 1) corresponding to each output port.

[145] The gain control information and the destination information may be inputted at once or separately into the object gain information encoding unit 5113.

[146] **2.3 Process of generating a combined object information**

[147] FIG. 8 is an exemplary block diagram of the combined object information encoding unit 5120. Referring to FIG. 8, the combined object information encoding unit 5120 can be configured to receive a reference value_1, a reference value_2, an object level information_1, an object level information_2, an object gain information_1, an object gain information_2, an object control information, a gain control information, and a destination information, and to generate a combined object information using the object control information, the gain control information, and the destination information.

[148] **2.3.1 determination of reference information**

[149] Again referring to FIG. 8, the combined object information encoding unit 5120 includes a reference value generating unit 5121, an object level information encoding unit 5122, and an object gain information encoding unit 5123.

[150] To generate the combined object information, first, a reference information of the combined object information may be estimated. Each object information_i may include

reference information to normalize each object level, and to generate an object level information. But, in case of combining at least two object informations to generate a combined object information, the combined object information may determine to normalize the object level constituted to the object level information of the combined object information.

[151] The reference information of the combined object information may be determine by several methods. For example, the reference information of the combined object information may be the reference information_1 or the largest reference information of the object information_i.

[152] Instead of a change of the reference information, the combined object information may use the object level information of the object information_i as that of the combined object information.

[153] **2.3.2 object level information of the combined object information**

[154] The reference information generating unit 5121 may estimate the reference information of the combined object information as the above method. Before the change of the reference information of the combined object information, the object level information_i is normalized by the reference information_i.

[155] We assume that the object level information of the object information_1 is the [Math Figure 10], and the object level information of the combined object information is the [Math Figure 11].

[156] [Math Figure 10]

[157] OL_1n = EO_1n/(reference information of the object information_1)

[158] (OL_1n is a n^{th} object level information of the object information_1, EO_1n is a n^{th} object level energy of the object information_1)

[159] [Math Figure 11]

[160] OL_k = OL_1n * (reference information of the object information_1)/ (reference information of the object information)

[161] (OL_k is a k^{th} object level information of the combined object information)

[162] **2.3.2 object gain information**

[163] The object gain information encoding unit 5123 can be configured to receive an object gain_1, an object gain_2, a gain control information, and a destination information, and to generate an object gain information using the gain control information and the destination information. In case that the destination information from user control indicates on/off of the object information, that is, the destination information is 0 or 1, the object gain information of the object information_i is 0 or 1. In case that the gain control information may be inputted from user control, the object gain information_1 and the object gain information_2 can be changed using the gain control information.

[164] **2.3.3 object correlation information**

[165] The object correlation information indicates similarity/dissimilarity between the channels of a stereo object or a multi-channel object, so the object correlation information may be affected by combining object information in the MCU combining unit 5100.

[166] The object correlation information of the combined object information may be included the object correlation information of the object information_i as it is.

[167] It will be apparent to those skilled in the art that various modifications and variations can be made in the present invention without departing from the spirit or scope of the inventions. Thus, it is intended that the present invention covers the modifications and variations of this invention provided they come within the scope of the appended claims and their equivalents.

Industrial Applicability

[168] Accordingly, the present invention is applicable to encode and decode an audio signal.

[169]

Claims

[1] A method of decoding method for an audio signal, comprising:
receiving a downmix of an audio signal, an object information, and a mix information, the object information including an object level information, an object correlation information, and an object gain information, the object level information being generated by normalizing object level corresponding to object using one of the object level as a reference information, the object correlation information provided from combination of two selected objects, the object gain information comprising at least one of an object gain value information and an object gain ratio information;
generating a downmix processing information using the object information and the mix information; and
processing the downmix of the audio signal using the downmix processing information.

[2] The method of claim 1, wherein the reference information comprises the largest object level among the all object level.

[3] The method of claim 1, wherein the number of the object level information is same as the number of the objects in the downmix of the audio signal.

[4] The method of claim 1, wherein the object correlation information comprises a relation information representing a different object of same origin.

[5] The method of claim 1, wherein the object correlation information exists based on a correlation_flag.

[6] The method of claim 1, wherein the object correlation information comprises a default value based on a correlation_flag.

[7] The method of claim 1, wherein the object gain value information comprises gain value to be applied to object for generation of the downmix of the audio signal.

[8] The method of claim 1, wherein the object gain ratio information comprises gain ratio for relatively contributing to at least two channels of the downmix of the audio signal.

[9] The method of claim 1, wherein the object information further comprises a reference information.

[10] The method of claim 1, wherein the object information further comprises a correlation flag.

[11] The method of claim 1, further comprising:
obtaining the processed downmix of the audio signal as an output signal.

[12] The method of claim 1, further comprising:
upmixing the processed downmix using a multi-channel parameter;

[13] The method of claim 1, wherein the downmix of the audio signal is received as a broadcast signal.

[14] The method of claim 1, wherein the downmix of the audio signal is received on a digital medium.

[15] A computer-readable medium having instructions stored thereon, which, when executed by a decoder, causes the processor to perform operations, comprising:
receiving a downmix of an audio signal, an object information, and a mix information, the object information including an object level information, an object correlation information, and an object gain information, the object level information being generated by normalizing object level corresponding to object using one of the object level as a reference information, the object correlation information provided from combination of two selected objects, the object gain information including at least one of an object gain ratio information and an object gain value information;
generating a downmix processing information using the object information and the mix information;
processing the downmix of the audio signal using the downmix processing information.

[16] An apparatus for decoding an audio signal, comprising:
a information generating unit receiving an object information and a mix information, the object information including an object level information, an object correlation information, and an object gain information, the object level information being generated by normalizing object level corresponding to object using one of the object level as a reference information, the object correlation information provided from combination of two selected objects, the object gain information comprising at least one of an object gain value information and an object gain ratio information, and generating a downmix processing information using the object information and the mix information; and
a downmix processing unit receiving the downmix of the audio signal and the downmix processing information, and processing the downmix of the audio signal using the downmix processing information;

[17] A method of encoding for an audio signal, comprising:
receiving a multi-object audio signal; and
generating a downmix of an audio signal and an object information including an object level information, an object gain information, and an object correlation, the object level information and the object correlation information from the multi-object audio signal, the object level information being generated by normalizing object level corresponding to object using one of the object level as

a reference information, the object correlation information provided from combination of two selected objects, the object gain information comprising at least one of an object gain value information and an object gain ratio information ;

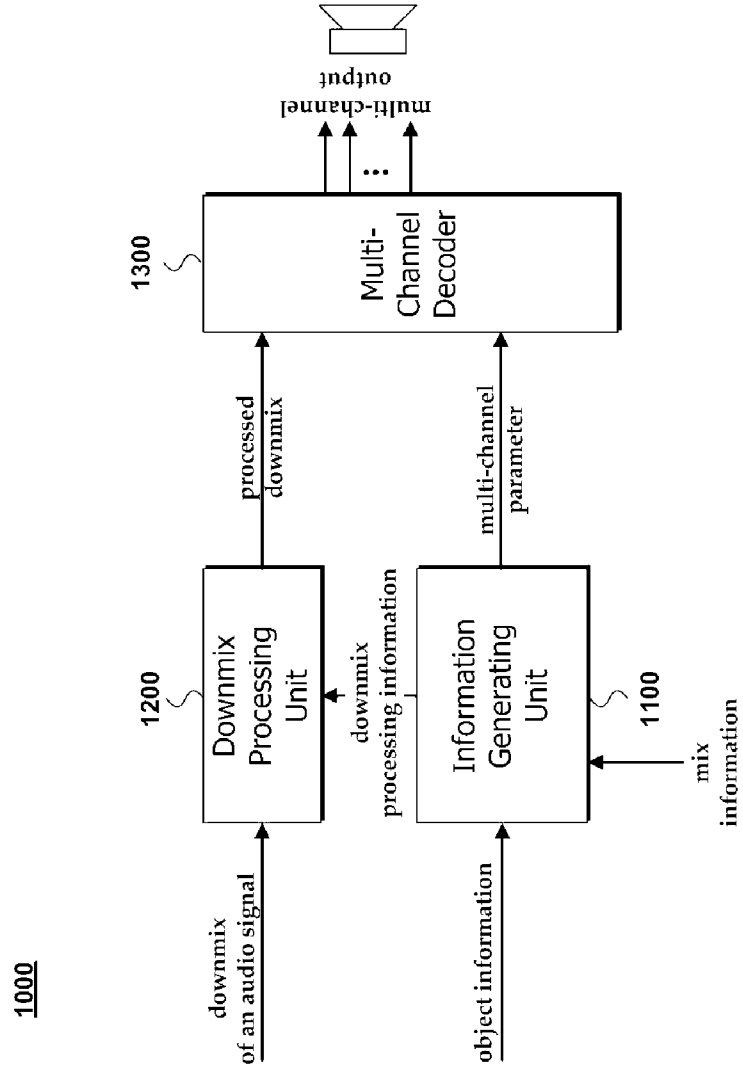
[18] The method of claim 17, wherein the reference information comprises the largest object level among the all object level.

[19] The method of claim 17, wherein the number of the object level information is same as the number of the objects in the downmix of the audio signal.

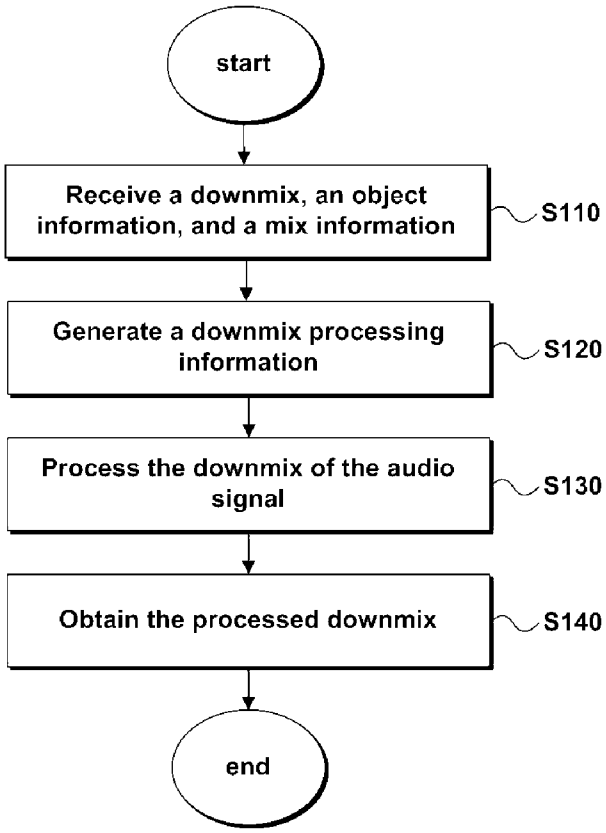
[20] The method of claim 17, wherein the object correlation information comprises a relation imformation representing a different object of same origin.

[21] An apparatus for encoding an audio signal, comprising:
a downmixing unit generating a downmix of an audio signal from a multi-object audio signal; and
an object information unit extracting an object information including an object level information, an object gain information, and an object correlation information from the multi-object audio signal, the object level information and the object correlation information from the multi-object audio signal, the object level information being generated by normalizing object level corresponding to object using one of the object level as a reference information, the object correlation information provided from combination of two selected objects, the object gain information comprising at least one of an object gain value information and an object gain ratio information.

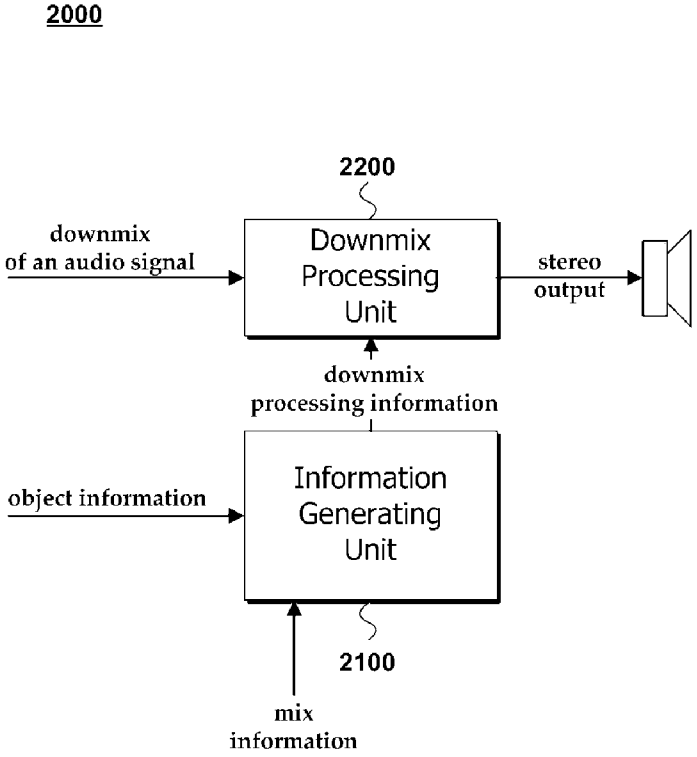
[Fig. 1]


multi-channel output
...
1300
Multi-Channel Decoder
processed downmix
multi-channel parameter
1200
Downmix Processing Unit
downmix processing information
Information Generating Unit
1100
mix information
downmix of an audio signal
object information
1000

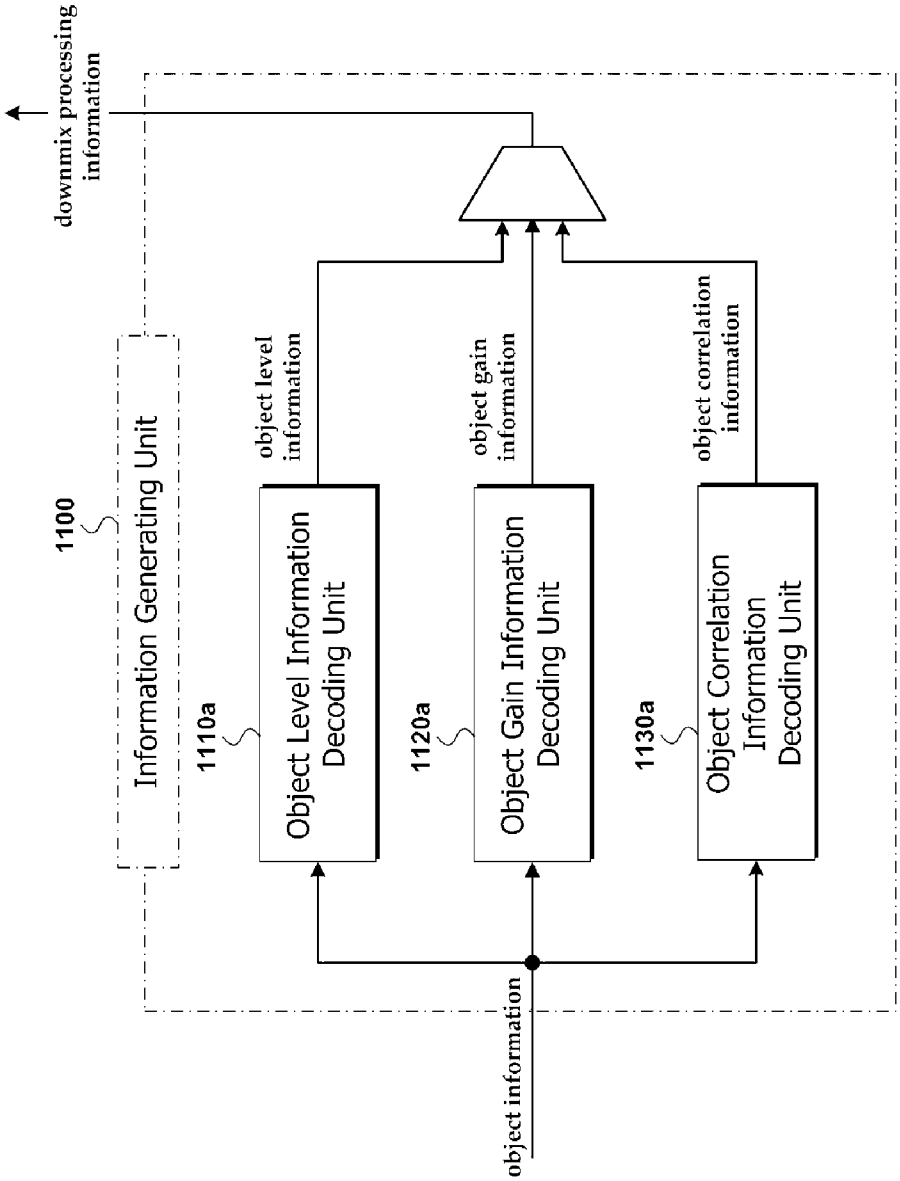
[Fig. 2]


start
Receive a downmix, an object information, and a mix information
S110
Generate a downmix processing information
S120
Process the downmix of the audio signal
S130
Obtain the processed downmix
S140
end

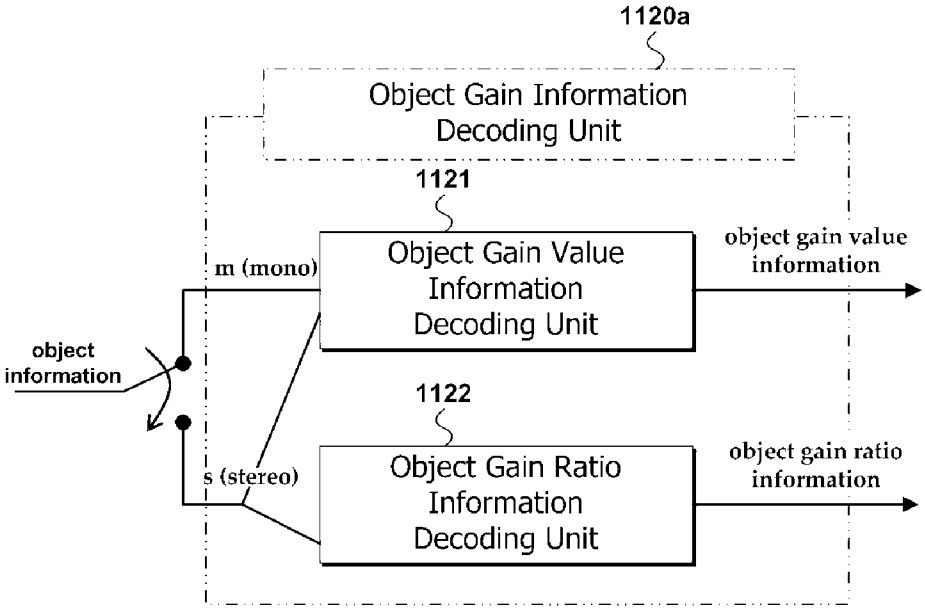
[Fig. 3]


2000
2200
downmix
of an audio signal
Downmix
Processing
Unit
stereo
output
downmix
processing information
object information
Information
Generating
Unit
2100
mix
information

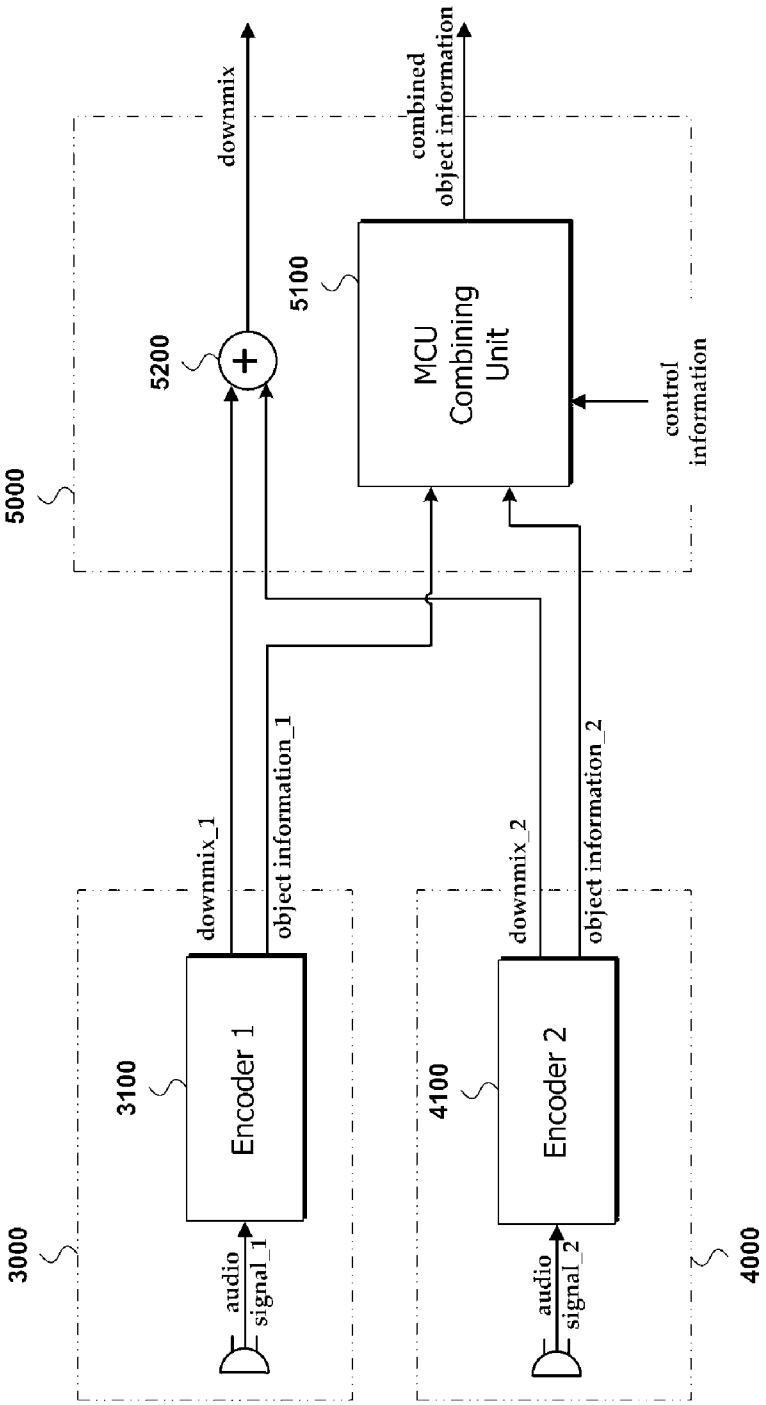
[Fig. 4]


downmix processing information
1100
Information Generating Unit
object level information
object gain information
object correlation information
1110a
Object Level Information Decoding Unit
1120a
Object Gain Information Decoding Unit
1130a
Object Correlation Information Decoding Unit
object information

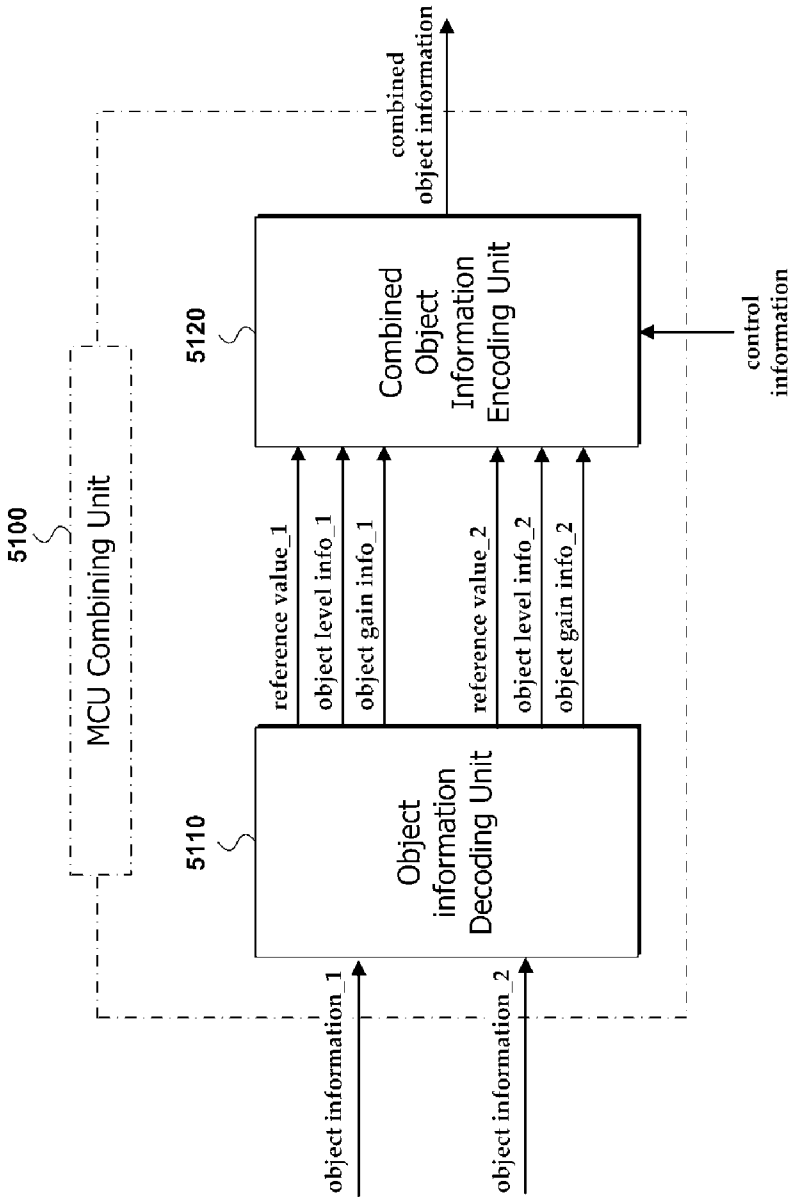
[Fig. 5]


1120a
Object Gain Information Decoding Unit
1121
Object Gain Value Information Decoding Unit
m (mono)
object gain value information
object information
1122
Object Gain Ratio Information Decoding Unit
s (stereo)
object gain ratio information

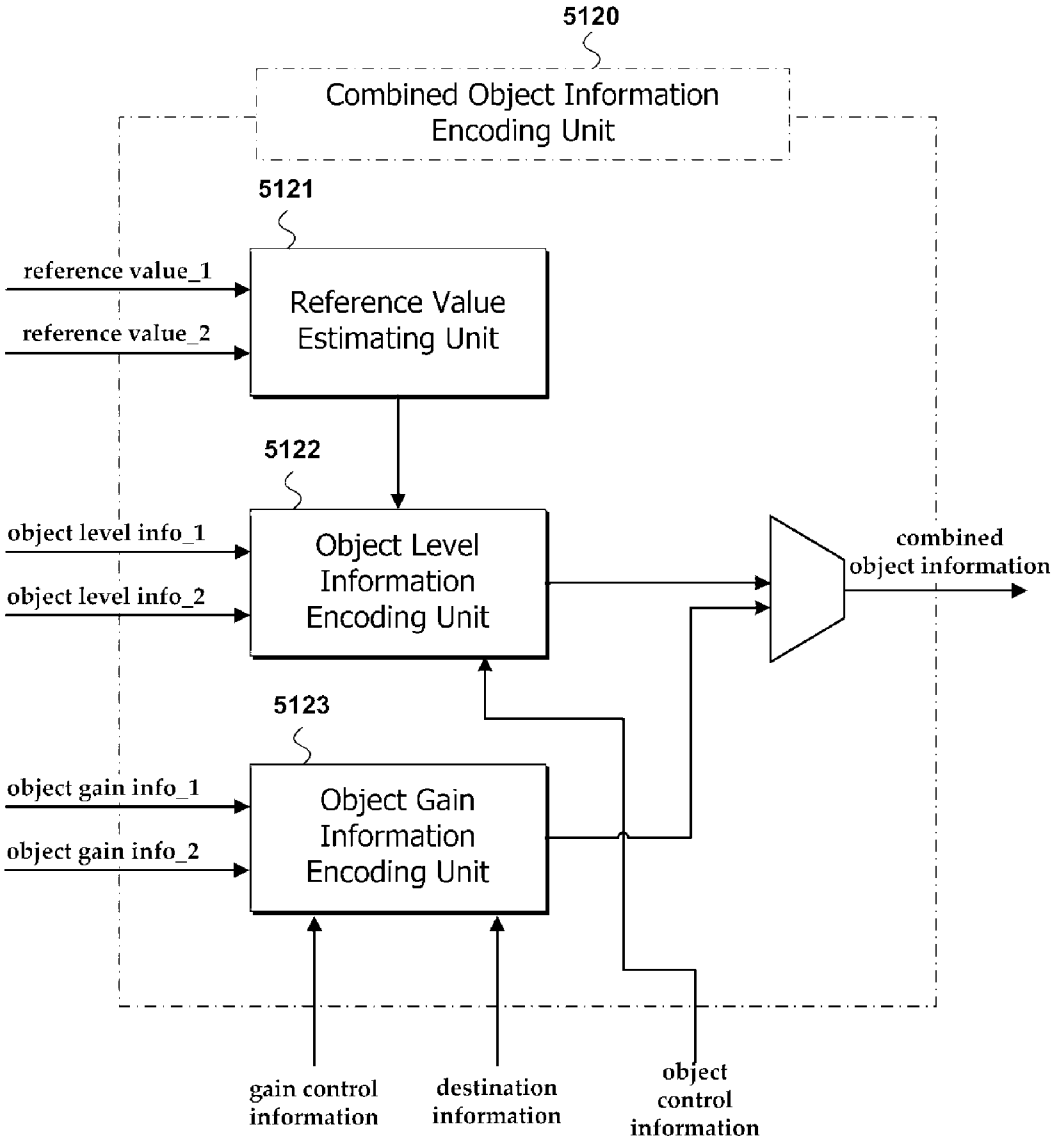
[Fig. 6]


downmix
combined object information
5100
MCU Combining Unit
5200
control information
5000
downmix_1
object information_1
downmix_2
object information_2
3100
Encoder 1
4100
Encoder 2
3000
4000
audio signal_1
audio signal_2

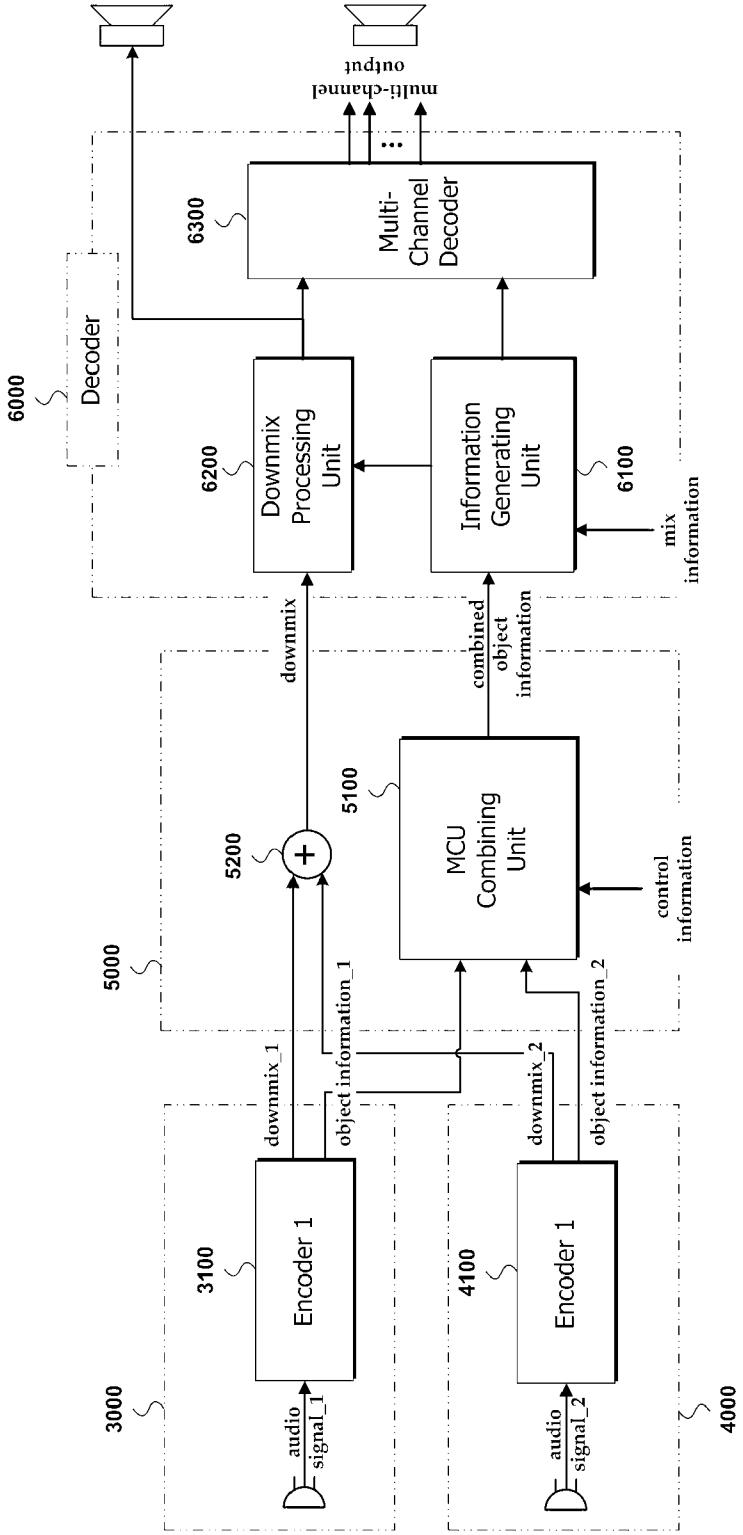
[Fig. 7]


5100
MCU Combining Unit
5110
Object information Decoding Unit
5120
Combined Object Information Encoding Unit
reference value_1
object level info_1
object gain info_1
reference value_2
object level info_2
object gain info_2
object information_1
object information_2
combined object information
control information

[Fig. 8]


5120
Combined Object Information Encoding Unit
5121
reference value_1
reference value_2
Reference Value Estimating Unit
5122
object level info_1
object level info_2
Object Level Information Encoding Unit
combined object information
5123
object gain info_1
object gain info_2
Object Gain Information Encoding Unit
gain control information
destination information
object control information

[Fig. 9]


multi-channel output
6300
Multi-Channel Decoder
6000
Decoder
6200
Downmix Processing Unit
Information Generating Unit
6100
mix information
downmix
combined object information
5100
MCU Combining Unit
5200
control information
5000
downmix_1
object information_1
downmix_2
object information_2
3100
Encoder 1
4100
Encoder 1
audio signal_1
audio signal_2
3000
4000

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2007/005740

A. CLASSIFICATION OF SUBJECT MATTER

G10L 19/00(2006.01)i, H03M 7/30(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 H04B 1/00 G01L 21/04 G01L 19/02 H04L 12/16

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""VOICE""DOWNMIX""MULTIPOINT""CONTROL""UNIT"

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2005/0117759 A1 (WU) 02 JUNE 2005 Read the whole document	1, 15, 16, 17, 21
A	US 6,122,619 A (KOLLURU et al.) 19 SEPTEMBER 2000 See Column 4, Line 24 - Column 10, Line 54	1, 15, 16, 17, 21
A	US 6,128,597 A (KOLLURU et al.) 03 OCTOBER 2000 See Column 7, Line 21 - Column 8, Line 25	1
A	WO 03/090207 A1 (OOMEN et al.) 30 OCTOBER 2003 Abstract	1
A	US 2003/0231600 A1 (POLOMSKI) 18 DECEMBER 2003 Fig 25, See Paragraph [0145] - Paragraph [0152]	1
A	US 6,584,077 B1 (POLOMSKI) 24 JUNE 2003 See Column 19, Line 47 - Column 20, Line 4	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
27 FEBRUARY 2008 (27.02.2008)	**27 FEBRUARY 2008 (27.02.2008)**

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

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/KR2007/005740

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2005/0117759 A1	02.06.2005	NONE	
US 6,122,619 A	19.09.2000	NONE	
US 6,128,597 A	03.10.2000	NONE	
WO 03/090207A1	30.10.2003	AT332003E	15.07.2006
		AT354161E	15.03.2007
		AU2003216682A1	03.11.2003
		AU2003216682AA	03.11.2003
		AU2003216686A1	03.11.2003
		AU2003216686AA	03.11.2003
		AU2003219426A1	03.11.2003
		AU2003219426AA	03.11.2003
		BR200304540A	20.07.2004
		BR200304541A	20.07.2004
		BR200304542A	20.07.2004
		CN1307612C	28.03.2007
		CN1312660C	25.04.2007
		CN1647155A	27.07.2005
		CN1647156A	27.07.2005
		CN1647157A	27.07.2005
		DE60306512C0	10.08.2006
		DE60306512T2	21.06.2007
		DE60311794C0	29.03.2007
		DE60311794T2	31.10.2007
		EP01500082A1	26.01.2005
		EP01500082B1	14.02.2007
		EP01500083A1	26.01.2005
		EP01500083B1	28.06.2006
		EP01500084A1	26.01.2005
		EP1500082A1	26.01.2005
		EP1500082B1	14.02.2007
		EP1500082A1	26.01.2005
		EP1500083A1	26.01.2005
		EP1500083B1	28.06.2006
		EP1500083A1	26.01.2005
		EP1500084A1	26.01.2005
		ES2268340T3	16.03.2007
		ES2280736T3	16.09.2007
		JP17523479	04.08.2005
		JP17523480	04.08.2005
		JP17523624	04.08.2005
		JP2005523479T2	04.08.2005
		JP2005523480T2	04.08.2005
		JP2005523624T2	04.08.2005
		KR1020040101552	02.12.2004
		KR1020040102163	03.12.2004
		KR1020040102164	03.12.2004
		US2005226426A1	13.10.2005
		US2005226426AA	13.10.2005

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/KR2007/005740

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		WO03090206A1	30.10.2003
		WO03090208A1	30.10.2003
		WO2003090206A1	30.10.2003
		WO2003090207A1	30.10.2003
		WO2003090208A1	30.10.2003
US 2003/0231600 A1	18.12.2003	AU4698796A1	14.08.1996
		DE19681223T	08.01.1998
		GB2312807A1	05.11.1997
		GB2312807B2	22.12.1999
		GB2312807A8	22.12.1997
		GB9715337A0	24.09.1997
		US2003231600A1	18.12.2003
		US2003231600AA	18.12.2003
		US6584077BA	24.06.2003
		WO9623388A1	01.08.1996
US 6,584,077 B1	24.06.2003	US2003231600A1	18.12.2003
		US2003231600AA	18.12.2003
		US6584077B1	24.06.2003
		US6584077BA	24.06.2003