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

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

(21) **International Application Number:**
PCT/KR2007/006316

(22) **International Filing Date:**
6 December 2007 (06.12.2007)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**

60/869,077	7 December 2006 (07.12.2006)	US
60/877,134	27 December 2006 (27.12.2006)	US
60/883,569	5 January 2007 (05.01.2007)	US
60/884,043	9 January 2007 (09.01.2007)	US
60/884,347	10 January 2007 (10.01.2007)	US
60/884,585	11 January 2007 (11.01.2007)	US
60/885,347	17 January 2007 (17.01.2007)	US
60/885,343	17 January 2007 (17.01.2007)	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]; 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, 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,

(54) **Title:** A METHOD AND AN APPARATUS FOR PROCESSING AN AUDIO SIGNAL
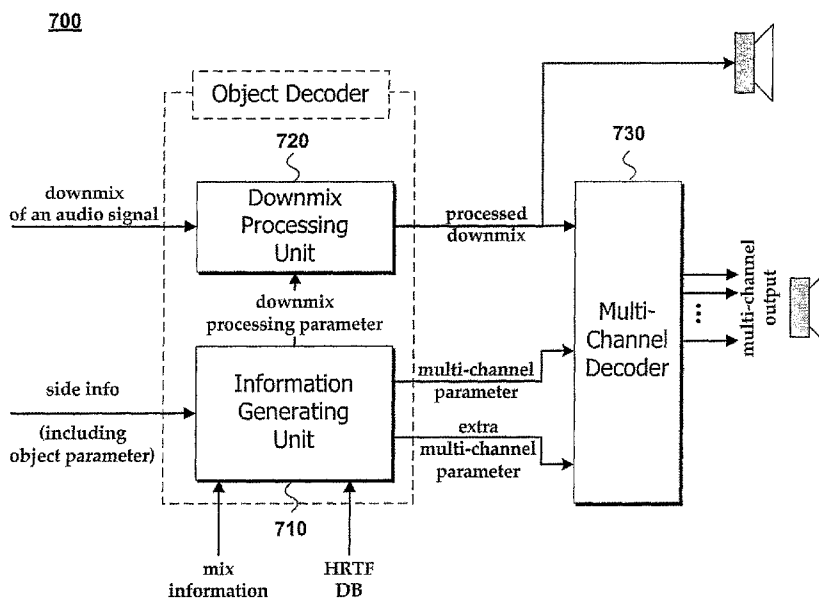
(57) **Abstract:** A method for processing an audio signal, comprising: receiving a downmix signal in time domain; if the downmix signal corresponds to a mono signal, bypassing the downmix signal; if the number of channel of the downmix signal corresponds to at least two, decomposing the downmix signal into a subband signal, and processing the subband signal using a downmix processing information, wherein the downmix processing information is estimated based on an object information and a mix information is disclosed.

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, 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】

【Invention Title】

A METHOD AND AN APPARATUS FOR PROCESSING AN AUDIO

ЗIGNAL

【Technical Field】

The present invention relates to a method and an apparatus for processing

an audio signal, and more particularly, to a method and an apparatus for decoding

an audio signal received on a digital medium, as a broadcast signal, and so on.

【Background Art】

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 repositioning/panning of the

individual sources can be controlled by user' selection.

【Disclosure】

【Technical Problem】

However, in order to control the individual object signals,

repositioning/panning of the individual sources included in a downmix signal

must be performed suitably.

However, for backward compatibility with respect to the channel-oriented decoding method (as a MPEG Surround), an object parameter must be converted flexibly to a multi-channel parameter required in upmixing process.

【Technical Solution】

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

An object of the present invention is to provide a method and an apparatus for processing an audio signal to control object gain and panning unrestrictedly.

Another object of the present invention is to provide a method and an apparatus for processing an audio signal to control object gain and panning based on user selection.

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】

The present invention provides the following effects or advantages.

First of all, the present invention is able to provide a method and an pparatus for processing an audio signal to control object gain and panning nrestrictedly.

Secondly, the present invention is able to provide a method and an pparatus for processing an audio signal to control object gain and panning based n user selection.

[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 pplication, illustrate embodiments of the invention and together with the lescription serve to explain the principle of the invention. In the drawings;

FIG. 1 is an exemplary block diagram to explain to basic concept of rendering a downmix signal based on playback configuration and user control.

FIG. 2 is an exemplary block diagram of an apparatus for processing an audio signal according to one embodiment of the present invention corresponding to the first scheme.

FIG. 3 is an exemplary block diagram of an apparatus for processing an audio signal according to another embodiment of the present invention corresponding to the first scheme.

FIG. 4 is an exemplary block diagram of an apparatus for processing an udio signal according to one embodiment of present invention corresponding to he second scheme.

FIG. 5 is an exemplary block diagram of an apparatus for processing an udio signal according to another embodiment of present invention corresponding o the second scheme.

FIG. 6 is an exemplary block diagram of an apparatus for processing an udio signal according to the other embodiment of present invention orresponding to the second scheme.

FIG. 7 is an exemplary block diagram of an apparatus for processing an udio signal according to one embodiment of the present invention corresponding o the third scheme.

FIG. 8 is an exemplary block diagram of an apparatus for processing an udio signal according to another embodiment of the present invention orresponding to the third scheme.

FIG. 9 is an exemplary block diagram to explain to basic concept of endering unit.

FIGS. 10A to 10C are exemplary block diagrams of a first embodiment of a lownmix processing unit illustrated in FIG. 7.

FIG. 11 is an exemplary block diagram of a second embodiment of a downmix processing unit illustrated in FIG. 7.

FIG. 12 is an exemplary block diagram of a third embodiment of a downmix processing unit illustrated in FIG. 7.

FIG. 13 is an exemplary block diagram of a fourth embodiment of a downmix processing unit illustrated in FIG. 7.

FIG. 14 is an exemplary block diagram of a bitstream structure of a compressed audio signal according to a second embodiment of present invention.

FIG. 15 is an exemplary block diagram of an apparatus for processing an audio signal according to a second embodiment of present invention.

FIG. 16 is an exemplary block diagram of a bitstream structure of a compressed audio signal according to a third embodiment of present invention.

FIG. 17 is an exemplary block diagram of an apparatus for processing an audio signal according to a fourth embodiment of present invention.

FIG. 18 is an exemplary block diagram to explain transmitting scheme for variable type of object.

FIG. 19 is an exemplary block diagram to an apparatus for processing an audio signal according to a fifth embodiment of present invention.

【Best Mode】

To achieve these objects and other advantages and in accordance with the purpose of the invention, as embodied and broadly described herein, a method for processing an audio signal, comprising: receiving a downmix signal in time domain; if the downmix signal corresponds to a mono signal, bypassing the downmix signal; if the number of channel of the downmix signal corresponds to at least two, decomposing the downmix signal into a subband signal, and processing the subband signal using a downmix processing information, wherein the downmix processing information is estimated based on an object information and a mix information.

According to the present invention, wherein the number of channel of the downmix signal is equal to the number of channel of the processed downmix signal.

According to the present invention, wherein the object information is included in a side information, and the side information includes a correlation flag information indicating whether an object is part of at least two channel object.

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

According to the present invention, wherein the downmix processing information corresponds to an information for controlling object panning if the number of channel the downmix signal corresponds to at least two.

According to the present invention, wherein the downmix processing information corresponds to an information for controlling object gain.

According to the present invention, further comprising, generating a multi-hannel signal using the processed subband signal.

According to the present invention, further comprising, generating a multi-hannel information using the object information and the mix information, wherein he multi-channel signal is generated based on the multi-channel information.

According to the present invention, further comprising, downmixing the lownmix signal to be a mono signal if the downmix signal corresponds to a stereo ignal.

According to the present invention, wherein the mix information is generated using at least one of an object position information and a playback configuration information.

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

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

In another aspect of the present invention, a computer-readable medium having instructions stored thereon, which, when executed by a processor, causes the processor to perform operations, comprising: receiving a downmix signal in

ime domain; if the downmix signal corresponds to a mono signal, bypassing the

lownmix signal; if the number of channel of the downmix signal corresponds to at

east two, decomposing the downmix signal into a subband signal, and processing

he subband signal using a downmix processing information, wherein the

lownmix processing information is estimated based on an object information and a

nix information.

In another aspect of the present invention, an apparatus for processing an

tudio signal, comprising: a receiving unit receiving a downmix signal in time

lomain; and, a downmix processing unit bypassing the downmix signal if the

lownmix signal corresponds to a mono signal, and decomposing the downmix

ignal into a subband signal and processing the subband signal using a downmix

processing information if the number of channel of the downmix signal

:orresponds to at least two, wherein the downmix processing information is

:stimated based on an object information and a mix information.

It is to be understood that both the foregoing general description and the

:ollowing detailed description of the present invention are exemplary and

:xplanatory and are intended to provide further explanation of the invention as

:laimed.

【Mode for Invention】

Reference will now be made in detail to the preferred embodiments of the resent invention, examples of which are illustrated in the accompanying drawings. /herever possible, the same reference numbers will be used throughout the rawings to refer to the same or like parts.

Prior to describing the present invention, it should be noted that most terms isclosed in the present invention correspond to general terms well known in the rt, but some terms have been selected by the applicant as necessary and will ereinafter be disclosed in the following description of the present invention. herefore, it is preferable that the terms defined by the applicant be understood on ie basis of their meanings in the present invention.

In particular, 'parameter' in the following description means information ncluding values, parameters of narrow sense, coefficients, elements, and so on. Iereinafter 'parameter' term will be used instead of 'information' term like an bject parameter, a mix parameter, a downmix processing parameter, and so on, /hich does not put limitation on the present invention.

In downmixing several channel signals or object signals, an object parameter nd a spatial parameter can be extracted. A decoder can generate output signal ising a downmix signal and the object parameter (or the spatial parameter). The iutput signal may be rendered based on playback configuration and user control

y the decoder. The rendering process shall be explained in details with reference

) the FIG. 1 as follow.

FIG. 1 is an exemplary diagram to explain to basic concept of rendering

ownmix based on playback configuration and user control. Referring to FIG. 1, a

ecoder 100 may include a rendering information generating unit 110 and a

endering unit 120, and also may include a renderer 110a and a synthesis 120a

astead of the rendering information generating unit 110 and the rendering unit 120.

A rendering information generating unit 110 can be configured to receive a

ide information including an object parameter or a spatial parameter from an

ncoder, and also to receive a playback configuration or a user control from a

levice setting or a user interface. The object parameter may correspond to a

)arameter extracted in downmixing at least one object signal, and the spatial

)arameter may correspond to a parameter extracted in downmixing at least one

:hannel signal. Furthermore, type information and characteristic information for

:ach object may be included in the side information. Type information and

:haracteristic information may describe instrument name, player name, and so on.

The playback configuration may include speaker position and ambient information

'speaker's virtual position), and the user control may correspond to a control

information inputted by a user in order to control object positions and object gains,

and also may correspond to a control information in order to the playback

onfiguration. Meanwhile the payback configuration and user control can be

presented as a mix information, which does not put limitation on the present

vention.

A rendering information generating unit 110 can be configured to generate a

ndering information using a mix information (the playback configuration and

ser control) and the received side information. A rendering unit 120 can

onfigured to generate a multi-channel parameter using the rendering information

case that the downmix of an audio signal (abbreviated 'downmix signal') is not

ansmitted, and generate multi-channel signals using the rendering information

nd downmix in case that the downmix of an audio signal is transmitted.

A renderer 110a can be configured to generate multi-channel signals using a

ix information (the playback configuration and the user control) and the received

ide information. A synthesis 120a can be configured to synthesis the multi-channel

ignals using the multi-channel signals generated by the renderer 110a.

As previously stated, the decoder may render the downmix signal based on

layback configuration and user control. Meanwhile, in order to control the

ndividual object signals, a decoder can receive an object parameter as a side

nformation and control object panning and object gain based on the transmitted

bject parameter.

1. Controlling gain and panning of object signals

Variable methods for controlling the individual object signals may be provided. First of all, in case that a decoder receives an object parameter and generates the individual object signals using the object parameter, then, can control the individual object signals base on a mix information (the playback configuration, the object level, etc.)

Secondly, in case that a decoder generates the multi-channel parameter to be inputted to a multi-channel decoder, the multi-channel decoder can upmix a downmix signal received from an encoder using the multi-channel parameter. The above-mention second method may be classified into three types of scheme. In particular, 1) using a conventional multi-channel decoder, 2) modifying a multi-channel decoder, 3) processing downmix of audio signals before being inputted to a multi-channel decoder may be provided. The conventional multi-channel decoder may correspond to a channel-oriented spatial audio coding (ex: MPEG Surround decoder), which does not put limitation on the present invention. Details of three types of scheme shall be explained as follow.

1.1 Using a multi-channel decoder

First scheme may use a conventional multi-channel decoder as it is without modifying a multi-channel decoder. At first, a case of using the ADG (arbitrary downmix gain) for controlling object gains and a case of using the 5-2-5 configuration for controlling object panning shall be explained with reference to

IG. 2 as follow. Subsequently, a case of being linked with a scene remixing unit

/ill be explained with reference to FIG. 3.

FIG. 2 is an exemplary block diagram of an apparatus for processing an

udio signal according to one embodiment of the present invention corresponding

ɔ first scheme. Referring to FIG. 2, an apparatus for processing an audio signal 200

hereinafter simply 'a decoder 200') may include an information generating unit

.10 and a multi-channel decoder 230. The information generating unit 210 may

eceive a side information including an object parameter from an encoder and a

nix information from a user interface, and may generate a multi-channel

)arameter including a arbitrary downmix gain or a gain modification

;ain(hereinafter simple 'ADG'). The ADG may describe a ratio of a first gain

:stimated based on the mix information and the obejct information over a second

;ain extimated based on the object information. In particular, the information

;enerating unit 210 may generate the ADG only if the downmix signal corresponds

:o a mono signal. The multi-channel decoder 230 may receive a downmix of an

1udio signal from an encoder and a multi-channel parameter from the information

;enerating unit 210, and may generate a multi-channel output using the downmix

;ignal and the multi-channel parameter.

14

The multi-channel parameter may include a channel level difference hereinafter abbreviated 'CLD'), an inter channel correlation (hereinafter abbreviated 'ICC'), a channel prediction coefficient (hereinafter abbreviated 'CPC').

Since CLD, ICC, and CPC describe intensity difference or correlation between two channels, and is to control object panning and correlation. It is able to control object positions and object diffuseness (sonority) using the CLD, the ICC, etc. Meanwhile, the CLD describe the relative level difference instead of the absolute level, and energy of the splitted two channels is conserved. Therefore it is unable to control object gains by handling CLD, etc. In other words, specific object cannot be mute or volume up by using the CLD, etc.

Furthermore, the ADG describes time and frequency dependent gain for controlling correction factor by a user. If this correction factor be applied, it is able to handle modification of down-mix signal prior to a multi-channel upmixing. Therefore, in case that ADG parameter is received from the information generating unit 210, the multi-channel decoder 230 can control object gains of specific time and frequency using the ADG parameter.

Meanwhile, a case that the received stereo downmix signal outputs as a stereo channel can be defined the following formula 1.

[formula 1]

$$y[0] = w_{11} \cdot g_0 \cdot x[0] + w_{12} \cdot g_1 \cdot x[1]$$

$$y[1] = w_{21} \cdot g_0 \cdot x[0] + w_{22} \cdot g_1 \cdot x[1]$$

where x[] is input channels, y[] is output channels, g_x is gains, and w_{xx} is weight.

It is necessary to control cross-talk between left channel and right channel in order to object panning. In particular, a part of left channel of downmix signal may output as a right channel of output signal, and a part of right channel of downmix signal may output as left channel of output signal. In the formula 1, w_{12} and w_{21} may be a cross-talk component (in other words, cross-term).

The above-mentioned case corresponds to 2-2-2 configuration, which means 2-channel input, 2-channel transmission, and 2-channel output. In order to perform the 2-2-2 configuration, 5-2-5 configuration (2-channel input, 5-channel transmission, and 2 channel output) of conventional channel-oriented spatial audio coding (ex: MPEG surround) can be used. At first, in order to output 2 channels for 2-2-2 configuration, certain channel among 5 output channels of 5-2-5 configuration can be set to a disable channel (a fake channel). In order to give cross-talk between 2-transmitted channels and 2-output channels, the above-mentioned CLD and CPC may be adjusted. In brief, gain factor g_x in the formula 1 is obtained using the above mentioned ADG, and weighting factor $w_{11} \sim w_{22}$ in the formula 1 is obtained using CLD and CPC.

In implementing the 2-2-2 configuration using 5-2-5 configuration, in order

to reduce complexity, default mode of conventional spatial audio coding may be

applied. Since characteristic of default CLD is supposed to output 2-channel, it is

able to reduce computing amount if the default CLD is applied. Particularly, since

there is no need to synthesis a fake channel, it is able to reduce computing amount

largely. Therefore, applying the default mode is proper. In particular, only default

CLD of 3 CLDs (corresponding to 0, 1, and 2 in MPEG surround standard) is used

for decoding. On the other hand, 4 CLDs among left channel, right channel, and

center channel (corresponding to 3, 4, 5, and 6 in MPEG surround standard) and 2

ADGs (corresponding to 7 and 8 in MPEG surround standard) is generated for

controlling object. In this case, CLDs corresponding 3 and 5 describe channel level

difference between left channel plus right channel and center channel ((l+r)/c) is

proper to set to 150dB (approximately infinite) in order to mute center channel.

And, in order to implement cross-talk, energy based up-mix or prediction based

up-mix may be performed, which is invoked in case that TTT mode

'bsTttModeLow' in the MPEG surround standard) corresponds to energy-based

mode (with subtraction, matrix compatibility enabled) (3rd mode), or prediction

mode (1st mode or 2nd mode).

FIG. 3 is an exemplary block diagram of an apparatus for processing an

audio signal according to another embodiment of the present invention

orresponding to first scheme. Referring to FIG. 3, an apparatus for processing an

udio signal according to another embodiment of the present invention 300

hereinafter simply a decoder 300) may include a information generating unit 310, a

cene rendering unit 320, a multi-channel decoder 330, and a scene remixing unit

50.

The information generating unit 310 can be configured to receive a side

nformation including an object parameter from an encoder if the downmix signal

orresponds to mono channel signal (i.e., the number of downmix channel is '1'),

nay receive a mix information from a user interface, and may generate a multi-

hannel parameter using the side information and the mix information. The

umber of downmix channel can be estimated based on a flag information

ncluded in the side information as well as the downmix signal itself and user

election. The information generating unit 310 may have the same configuration of

he former information generating unit 210. The multi-channel parameter is

nputted to the multi-channel decoder 330, the multi-channel decoder 330 may

ave the same configuration of the former multi-channel decoder 230.

The scene rendering unit 320 can be configured to receive a side information

ncluding an object parameter from and encoder if the downmix signal corresponds

o non-mono channel signal (i.e., the number of downmix channel is more than '2'),

nay receive a mix information from a user interface, and may generate a remixing

arameter using the side information and the mix information. The remixing

arameter corresponds to a parameter in order to remix a stereo channel and

generate more than 2-channel outputs. The remixing parameter is inputted to the

cene remixing unit 350. The scene remixing unit 350 can be configured to remix

he downmix signal using the remixing parameter if the downmix signal is more

han 2-channel signal.

In brief, two paths could be considered as separate implementations for

eparate applications in a decoder 300.

1.2 Modifying a multi-channel decoder

Second scheme may modify a conventional multi-channel decoder. At first, a

case of using virtual output for controlling object gains and a case of modifying a

device setting for controlling object panning shall be explained with reference to

FIG. 4 as follow. Subsequently, a case of Performing TBT(2x2) functionality in a

multi-channel decoder shall be explained with reference to FIG. 5.

FIG. 4 is an exemplary block diagram of an apparatus for processing an

audio signal according to one embodiment of present invention corresponding to

the second scheme. Referring to FIG. 4, an apparatus for processing an audio signal

according to one embodiment of present invention corresponding to the second

scheme 400 (hereinafter simply 'a decoder 400') may include an information

generating unit 410, an internal multi-channel synthesis 420, and an output

napping unit 430. The internal multi-channel synthesis 420 and the output

napping unit 430 may be included in a synthesis unit.

The information generating unit 410 can be configured to receive a side

information including an object parameter from an encoder, and a mix parameter

rom a user interface. And the information generating unit 410 can be configured to

generate a multi-channel parameter and a device setting information using the side

information and the mix information. The multi-channel parameter may have the

same configuration of the former multi-channel parameter. So, details of the multi-

channel parameter shall be omitted in the following description. The device setting

information may correspond to parameterized HRTF for binaural processing,

which shall be explained in the description of '1.2.2 Using a device setting

information'.

The internal multi-channel synthesis 420 can be configured to receive a

multi-channel parameter and a device setting information from the parameter

generation unit 410 and downmix signal from an encoder. The internal multi-

channel synthesis 420 can be configured to generate a temporal multi-channel

output including a virtual output, which shall be explained in the description of

'1.2.1 Using a virtual output'.

1.2.1 Using a virtual output

Since multi-channel parameter (ex: CLD) can control object panning, it is
ırd to control object gain as well as object panning by a conventional multi-
ıannel decoder.

Meanwhile, in order to object gain, the decoder 400 (especially the internal
ıulti-channel synthesis 420) may map relative energy of object to a virtual channel
:x: center channel). The relative energy of object corresponds to energy to be
:duced. For example, in order to mute certain object, the decoder 400 may map
ıore than 99.9% of object energy to a virtual channel. Then, the decoder 400
:specially, the output mapping unit 430) does not output the virtual channel to
/hich the rest energy of object is mapped. In conclusion, if more than 99.9% of
bject is mapped to a virtual channel which is not outputted, the desired object can
ıe almost mute.

1.2.2 Using a device setting information

The decoder 400 can adjust a device setting information in order to control
ıbject panning and object gain. For example, the decoder can be configured to
zenerate a parameterized HRTF for binaural processing in MPEG Surround
;tandard. The parameterized HRTF can be variable according to device setting. It is
ıble to assume that object signals can be controlled according to the following
formula 2.

[formula 2]

$L_{new} = a_1 * obj_1 + a_2 * obj_2 + a_3 * obj_3 + .. + a_n * obj_n,$

$R_{new} = b_1 * obj_1 + b_2 * obj_2 + b_3 * obj_3 + .. + b_n * obj_n,$

where obj_k is object signals, L_{new} and R_{new} is a desired stereo signal, and a_k and b_k are coefficients for object control.

An object information of the object signals obj_k may be estimated from an object parameter included in the transmitted side information. The coefficients a_k, b_k which are defined according to object gain and object panning may be estimated from the mix information. The desired object gain and object panning can be adjusted using the coefficients a_k, b_k.

The coefficients a_k, b_k can be set to correspond to HRTF parameter for binaural processing, which shall be explained in details as follow.

In MPEG Surround standard (5-1-5$_1$ configuration) (from ISO/IEC FDIS 23003-1:2006(E), Information Technology — MPEG Audio Technologies — Part1: MPEG Surround), binaural processing is as below.

[formula 3]

$$y_B^{n,k} = \begin{bmatrix} y_{L_B}^{n,k} \\ y_{R_B}^{n,k} \end{bmatrix} = H_2^{n,k} \begin{bmatrix} y_m^{n,k} \\ D(y_m^{n,k}) \end{bmatrix} = \begin{bmatrix} h_{11}^{n,k} & h_{12}^{n,k} \\ h_{21}^{n,k} & h_{22}^{n,k} \end{bmatrix} \begin{bmatrix} y_m^{n,k} \\ D(y_m^{n,k}) \end{bmatrix}, \quad 0 \le k < K$$

where y_B is output, the matrix H is conversion matrix for binaural processing.

[formula 4]

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

The elements of matrix H is defined as follows:

[formula 5]

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

[formula 6]

$$\left(\sigma_{\mathsf{X}}^{l,m}\right)^2 = \left(P_{\mathsf{X,C}}^m\right)^2\left(\sigma_{\mathsf{C}}^{l,m}\right)^2 + \left(P_{\mathsf{X,L}}^m\right)^2\left(\sigma_{\mathsf{L}}^{l,m}\right)^2 + \left(P_{\mathsf{X,Ls}}^m\right)^2\left(\sigma_{\mathsf{Ls}}^{l,m}\right)^2 + \left(P_{\mathsf{X,R}}^m\right)^2\left(\sigma_{\mathsf{R}}^{l,m}\right)^2 + \left(P_{\mathsf{X,Rs}}^m\right)^2\left(\sigma_{\mathsf{Rs}}^{l,m}\right)^2 + \dots$$
$$P_{\mathsf{X,L}}^m P_{\mathsf{X,R}}^m \rho_{\mathsf{L}}^m \sigma_{\mathsf{L}}^{l,m}\sigma_{\mathsf{R}}^{l,m} ICC_3^{l,m}\cos\left(\phi_{\mathsf{L}}^m\right) + \dots$$
$$P_{\mathsf{X,L}}^m P_{\mathsf{X,R}}^m \rho_{\mathsf{R}}^m \sigma_{\mathsf{L}}^{l,m}\sigma_{\mathsf{R}}^{l,m} ICC_3^{l,m}\cos\left(\phi_{\mathsf{R}}^m\right) + \dots$$
$$P_{\mathsf{X,Ls}}^m P_{\mathsf{X,Rs}}^m \rho_{\mathsf{Ls}}^m \sigma_{\mathsf{Ls}}^{l,m}\sigma_{\mathsf{Rs}}^{l,m} ICC_2^{l,m}\cos\left(\phi_{\mathsf{Ls}}^m\right) + \dots$$
$$P_{\mathsf{X,Ls}}^m P_{\mathsf{X,Rs}}^m \rho_{\mathsf{Rs}}^m \sigma_{\mathsf{Ls}}^{l,m}\sigma_{\mathsf{Rs}}^{l,m} ICC_2^{l,m}\cos\left(\phi_{\mathsf{Rs}}^m\right)$$

[formula 7]

$$\left(\sigma_{\mathsf{L}}^{l,m}\right)^2 = r_1\left(CLD_0^{l,m}\right)r_1\left(CLD_1^{l,m}\right)r_1\left(CLD_3^{l,m}\right)$$

$$\left(\sigma_{\mathsf{R}}^{l,m}\right)^2 = r_1\left(CLD_0^{l,m}\right)r_1\left(CLD_1^{l,m}\right)r_2\left(CLD_3^{l,m}\right)$$

$$\left(\sigma_{\mathsf{C}}^{l,m}\right)^2 = r_1\left(CLD_0^{l,m}\right)r_2\left(CLD_1^{l,m}\right)/g_{\mathsf{c}}^2$$

$$\left(\sigma_{\mathsf{Ls}}^{l,m}\right)^2 = r_2\left(CLD_0^{l,m}\right)r_1\left(CLD_2^{l,m}\right)/g_{\mathsf{s}}^2$$

$$\left(\sigma_{\mathsf{Rs}}^{l,m}\right)^2 = r_2\left(CLD_0^{l,m}\right)r_2\left(CLD_2^{l,m}\right)/g_{\mathsf{s}}^2$$

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

1.2.3 Performing TBT(2x2) functionality in a multi-channel decoder

FIG. 5 is an exemplary block diagram of an apparatus for processing an

audio signal according to another embodiment of present invention corresponding

to the second scheme. FIG. 5 is an exemplary block diagram of TBT functionality in

a multi-channel decoder. Referring to FIG. 5, a TBT module 510 can be configured

to receive input signals and a TBT control information, and generate output signals.

The TBT module 510 may be included in the decoder 200 of the FIG. 2 (or in

particular, the multi-channel decoder 230). The multi-channel decoder 230 may be

implemented according to the MPEG Surround standard, which does not put

limitation on the present invention.

[formula 9]

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{Wx}$$

where **x** is input channels, **y** is output channels, and **w** is weight.

The output y_1 may correspond to a combination input x_1 of the downmix

multiplied by a first gain w_{11} and input x_2 multiplied by a second gain w_{12}.

The TBT control information inputted in the TBT module 510 includes

elements which can compose the weight **w** (w_{11}, w_{12}, w_{21}, w_{22}).

In MEPG Surround standard, OTT(One-To-Two) module and TTT(Two-To-

Three) module is not proper to remix input signal although OTT module and TTT

module can upmix the input signal.

In order to remix the input signal, TBT (2x2) module 510 (hereinafter abbreviated 'TBT module 510') may be provided. The TBT module 510 may can be figured to receive a stereo signal and output the remixed stereo signal. The weight w may be composed using CLD(s) and ICC(s).

If the weight term $w_{11} \sim w_{22}$ is transmitted as a TBT control information, the decoder may control object gain as well as object panning using the received weight term. In transmitting the weight term **w**, variable scheme may be provided. At first, a TBT control information includes cross term like the w_{12} and w_{21}. Secondly, a TBT control information does not include the cross term like the w_{12} and w_{21}. Thirdly, the number of the term as a TBT control information varies adaptively.

At first, there is need to receive the cross term like the w_{12} and w_{21} in order to control object panning as left signal of input channel go to right of the output channel. In case of N input channels and M output channels, the terms which number is NxM may be transmitted as TBT control information. The terms can be quantized based on a CLD parameter quantization table introduced in a MPEG Surround, which does not put limitation on the present invention.

Secondly, unless left object is shifted to right position, (i.e. when left object is moved to more left position or left position adjacent to center position, or when only level of the object is adjusted), there is no need to use the cross term. In the

ase, it is proper that the term except for the cross term is transmitted. In case of N

input channels and M output channels, the terms which number is just N may be

transmitted.

Thirdly, the number of the TBT control information varies adaptively

according to need of cross term in order to reduce the bit rate of a TBT control

information. A flag information 'cross_flag' indicating whether the cross term is

present or not is set to be transmitted as a TBT control information. Meaning of the

flag information 'cross_flag' is shown in the following table 1.

[table 1] meaning of cross_flag

cross_flag	meaning
0	no cross term (includes only non-cross term) (only w_{11} and w_{22} are present)
1	includes cross term (w_{11}, w_{12}, w_{21}, and w_{22} are present)

In case that 'cross_flag' is equal to 0, the TBT control information does not

include the cross term, only the non-cross term like the w_{11} and w_{22} is present.

Otherwise ('cross_flag' is equal to 1), the TBT control information includes the cross

term.

Besides, a flag information 'reverse_flag' indicating whether cross term is

present or non-cross term is present is set to be transmitted as a TBT control

information. Meaning of flag information 'reverse_flag' is shown in the following

table 2.

[table 2] meaning of reverse_flag

reverse_flag	meaning
0	no cross term (includes only non-cross term) (only w_{11} and w_{22} are present)
1	only cross term (only w_{12} and w_{21} are present)

In case that 'reverse_flag' is equal to 0, the TBT control information does not

nclude the cross term, only the non-cross term like the w_{11} and w_{22} is present.

)therwise ('reverse_flag' is equal to 1), the TBT control information includes only

he cross term.

Futhermore, a flag information 'side_flag' indicating whether cross term is

resent and non-cross is present is set to be transmitted as a TBT control

nformation. Meaning of flag information 'side_flag' is shown in the following table

[table 3] meaning of side_config

side_config	meaning
0	no cross term (includes only non-cross term) (only w_{11} and w_{22} are present)
1	includes cross term (w_{11}, w_{12}, w_{21}, and w_{22} are present)
2	reverse (only w_{12} and w_{21} are present)

Since the table 3 corresponds to combination of the table 1 and the table 2,

details of the table 3 shall be omitted.

1.2.4 Performing TBT(2x2) functionality in a multi-channel decoder by

modifying a binaural decoder

The case of '1.2.2 Using a device setting information' can be performed

without modifying the binaural decoder. Hereinafter, performing TBT

functionality by modifying a binaural decoder employed in a MPEG Surround

decoder, with reference to FIG. 6.

FIG. 6 is an exemplary block diagram of an apparatus for processing an

audio signal according to the other embodiment of present invention

corresponding to the second scheme. In particular, an apparatus for processing an

audio signal 630 shown in the FIG. 6 may correspond to a binaural decoder

included in the multi-channel decoder 230 of FIG. 2 or the synthesis unit of FIG. 4,

which does not put limitation on the present invention.

An apparatus for processing an audio signal 630 (hereinafter 'a binaural

decoder 630') may include a QMF analysis 632, a parameter conversion 634, a

spatial synthesis 636, and a QMF synthesis 638. Elements of the binaural decoder

630 may have the same configuration of MPEG Surround binaural decoder in

MPEG Surround standard. For example, the spatial synthesis 636 can be configured

to consist of 1 2x2 (filter) matrix, according to the following formula 10:

[formula 10]

$$\mathbf{y}_B^{n,k} = \begin{bmatrix} y_{L_B}^{n,k} \\ y_{R_B}^{n,k} \end{bmatrix} = \sum_{i=0}^{N_q-1} \mathbf{H}_2^{n-i,k} \mathbf{y}_0^{n-i,k} = \sum_{i=0}^{N_q-1} \begin{bmatrix} h_{11}^{n-i,k} & h_{12}^{n-i,k} \\ h_{21}^{n-i,k} & h_{22}^{n-i,k} \end{bmatrix} \begin{bmatrix} y_{L_0}^{n-i,k} \\ y_{R_0}^{n-i,k} \end{bmatrix}, \quad 0 \le k < K$$

with y_0 being the QMF-domain input channels and y_B being the binaural

output channels, k represents the hybrid QMF channel index, and i is the HRTF

filter tap index, and n is the QMF slot index. The binaural decoder 630 can be

configured to perform the above-mentioned functionality described in subclause

1.2.2 Using a device setting information'. However, the elements h_{ij} may be

generated using a multi-channel parameter and a mix information instead of a

multi-channel parameter and HRTF parameter. In this case, the binaural decoder

600 can perform the functionality of the TBT module 510 in the FIG. 5. Details of the

elements of the binaural decoder 630 shall be omitted.

The binaural decoder 630 can be operated according to a flag information

binaural_flag'. In particular, the binaural decoder 630 can be skipped in case that a

flag information binaural_flag is '0', otherwise (the binaural_flag is '1'), the

binaural decoder 630 can be operated as below.

[table 4] meaning of binaural_flag

binaural_flag	Meaning
0	not binaural mode (a binaural decoder is deactivated)
1	binaural mode (a binaural decoder is activated)

1.3 Processing downmix of audio signals before being inputted to a multi-

hannel decoder

The first scheme of using a conventional multi-channel decoder have been

xplained in subclause in '1.1', the second scheme of modifying a multi-channel

ecoder have been explained in subclause in '1.2'. The third scheme of processing

ownmix of audio signals before being inputted to a multi-channel decoder shall

e explained as follow.

FIG. 7 is an exemplary block diagram of an apparatus for processing an

udio signal according to one embodiment of the present invention corresponding

o the third scheme. FIG. 8 is an exemplary block diagram of an apparatus for

rocessing an audio signal according to another embodiment of the present

nvention corresponding to the third scheme. At first, Referring to FIG. 7, an

pparatus for processing an audio signal 700 (hereinafter simply 'a decoder 700')

nay include an information generating unit 710, a downmix processing unit 720,

ind a multi-channel decoder 730. Referring to FIG. 8, an apparatus for processing

in audio signal 800 (hereinafter simply 'a decoder 800') may include an

nformation generating unit 810 and a multi-channel synthesis unit 840 having a

nulti-channel decoder 830. The decoder 800 may be another aspect of the decoder

700. In other words, the information generating unit 810 has the same configuration

of the information generating unit 710, the multi-channel decoder 830 has the same

)nfiguration of the multi-channel decoder 730, and, the multi-channel synthesis

ait 840 may has the same configuration of the downmix processing unit 720 and

uulti-channel unit 730. Therefore, elements of the decoder 700 shall be explained in

etails, but details of elements of the decoder 800 shall be omitted.

The information generating unit 710 can be configured to receive a side

iformation including an object parameter from an encoder and a mix information

·om an user-interface, and to generate a multi-channel parameter to be outputted

) the multi-channel decoder 730. From this point of view, the information

enerating unit 710 has the same configuration of the former information

enerating unit 210 of FIG. 2. The downmix processing parameter may correspond

) a parameter for controlling object gain and object panning. For example, it is able

) change either the object position or the object gain in case that the object signal is

ocated at both left channel and right channel. It is also able to render the object

ignal to be located at opposite position in case that the object signal is located at

·nly one of left channel and right channel. In order that these cases are performed,

he downmix processing unit 720 can be a TBT module (2x2 matrix operation). In

ase that the information generating unit 710 can be configured to generate ADG

lescribed with reference to FIG 2. in order to control object gain, the downmix

)rocessing parameter may include parameter for controlling object panning but

)bject gain.

Furthermore, the information generating unit 710 can be configured to

ceive HRTF information from HRTF database, and to generate an extra multi-

annel parameter including a HRTF parameter to be inputted to the multi-channel

coder 730. In this case, the information generating unit 710 may generate multi-

annel parameter and extra multi-channel parameter in the same subband domain

d transmit in syncronization with each other to the multi-channel decoder 730.

e extra multi-channel parameter including the HRTF parameter shall be

cplained in details in subclause '3. Processing Binaural Mode'.

The downmix processing unit 720 can be configured to receive downmix of

a audio signal from an encoder and the downmix processing parameter from the

iformation generating unit 710, and to decompose a subband domain signal using

abband analysis filter bank. The downmix processing unit 720 can be configured

) generate the processed downmix signal using the downmix signal and the

owmix processing parameter. In these processing, it is able to pre-process the

owmix signal in order to control object panning and object gain. The processed

owmix signal may be inputted to the multi-channel decoder 730 to be upmixed.

Furthermore, the processed downmix signal may be outputted and

laybacked via speaker as well. In order to directly output the processed signal via

peakers, the downmix processing unit 720 may perform synthesis filterbank using

he prepossed subband domain signal and output a time-domain PCM signal. It is

ble to select whether to directly output as PCM signal or input to the multi-

hannel decoder by user selection.

The multi-channel decoder 730 can be configured to generate multi-channel

utput signal using the processed downmix and the multi-channel parameter. The

nulti-channel decoder 730 may introduce a delay when the processed downmix

ignal and the multi-channel parameter are inputted in the multi-channel decoder

'30. The processed downmix signal can be synthesized in frequency domain (ex:

QMF domain, hybrid QMF domain, etc), and the multi-channel parameter can be

ynthesized in time domain. In MPEG surround standard, delay and

ynchronization for connecting HE-AAC is introduced. Therefore, the multi-

channel decoder 730 may introduce the delay according to MPEG Surround

standard.

The configuration of downmix processing unit 720 shall be explained in

detail with reference to FIG. 9 ~ FIG. 13.

1.3.1 A general case and special cases of downmix processing unit

FIG. 9 is an exemplary block diagram to explain to basic concept of

rendering unit. Referring to FIG. 9, a rendering module 900 can be configured to

generate M output signals using N input signals, a playback configuration, and a

user control. The N input signals may correspond to either object signals or channel

signals. Furthermore, the N input signals may correspond to either object

arameter or multi-channel parameter. Configuration of the rendering module 900

m be implemented in one of downmix processing unit 720 of FIG. 7, the former

:ndering unit 120 of FIG. 1, and the former renderer 110a of FIG. 1, which does not

ut limitation on the present invention.

If the rendering module 900 can be configured to directly generate M

hannel signals using N object signals without summing individual object signals

orresponding certain channel, the configuration of the rendering module 900 can

ie represented the following formula 11.

[formula 11]

$$\mathbf{C} = \mathbf{RO}$$

$$\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_M \end{bmatrix} = \begin{bmatrix} R_{11} & R_{21} & \cdots & R_{N1} \\ R_{12} & R_{22} & \cdots & R_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ R_{1M} & R_{2M} & \cdots & R_{NM} \end{bmatrix} \begin{bmatrix} O_1 \\ O_2 \\ \vdots \\ O_N \end{bmatrix}$$

C_i is a ith channel signal, O_j is j^{th} input signal, and R_{ji} is a matrix mapping j^{th}

nput signal to i^{th} channel.

If R matrix is separated into energy component E and de-correlation

:omponent, the formula 11 may be represented as follow.

[formula 12]

$$\mathbf{C} = \mathbf{RO} = \mathbf{EO} + \mathbf{DO}$$

$$
\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_M \end{bmatrix} = \begin{bmatrix} E_{11} & E_{21} & \cdots & E_{N1} \\ E_{12} & E_{22} & \cdots & E_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ E_{1M} & E_{2M} & \cdots & E_{NM} \end{bmatrix} \begin{bmatrix} O_1 \\ O_2 \\ \vdots \\ O_N \end{bmatrix} + \begin{bmatrix} D_{11} & D_{21} & \cdots & D_{N1} \\ D_{12} & D_{22} & \cdots & D_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ D_{1M} & D_{2M} & \cdots & D_{NM} \end{bmatrix} \begin{bmatrix} O_1 \\ O_2 \\ \vdots \\ O_N \end{bmatrix}
$$

It is able to control object positions using the energy component E, and it is

ble to control object diffuseness using the de-correlation component D.

Assuming that only i^{th} input signal is inputted to be outputted via j^{th} channel

nd k^{th} channel, the formula 12 may be represented as follow.

[formula 13]

$$
\mathbf{C}_{jk\_i} = \mathbf{R}_i \mathbf{O}_i
$$

$$
\begin{bmatrix} C_{j\_i} \\ C_{k\_i} \end{bmatrix} = \begin{bmatrix} \alpha_{j\_i} \cos(\theta_{j\_i}) & \alpha_{j\_i} \sin(\theta_{j\_i}) \\ \beta_{k\_i} \cos(\theta_{k\_i}) & \beta_{k\_i} \sin(\theta_{k\_i}) \end{bmatrix} \begin{bmatrix} o_i \\ D(o_i) \end{bmatrix}
$$

α_{j_i} is gain portion mapped to j^{th} channel, β_{k_i} is gain portion mapped to k^{th}

:hannel, θ is diffuseness level, and $D(o_i)$ is de-correlated output.

Assuming that de-correlation is omitted, the formula 13 may be simplified as

:ollow.

[formula 14]

$$
\mathbf{C}_{jk\_i} = \mathbf{R}_i \mathbf{O}_i
$$

$$
\begin{bmatrix} C_{j\_i} \\ C_{k\_i} \end{bmatrix} = \begin{bmatrix} \alpha_{j\_i} \cos(\theta_{j\_i}) \\ \beta_{k\_i} \cos(\theta_{k\_i}) \end{bmatrix} o_i
$$

If weight values for all inputs mapped to certain channel are estimated

:cording to the above-stated method, it is able to obtain weight values for each

ιannel by the following method.

1) Summing weight values for all inputs mapped to certain channel. For

example, in case that input 1 O_1 and input 2 O_2 is inputted and output

channel corresponds to left channel L, center channel C, and right channel R,

a total weight values $\alpha_{L(tot)}$, $\alpha_{C(tot)}$, $\alpha_{R(tot)}$ may be obtained as follows:

[formula 15]

$$\alpha_{L(tot)} = \alpha_{L1}$$

$$\alpha_{C(tot)} = \alpha_{C1} + \alpha_{C2}$$

$$\alpha_{R(tot)} = \alpha_{R2}$$

where α_{L1} is a weight value for input 1 mapped to left channel L, α_{C1} is a

veight value for input 1 mapped to center channel C, α_{C2} is a weight value for

nput2 mapped to center channel C, and α_{R2} is a weight value for input 2 mapped

o right channel R.

In this case, only input 1 is mapped to left channel, only input 2 is mapped

:o right channel, input 1 and input 2 is mapped to center channel together.

2) Summing weight values for all inputs mapped to certain channel, then

dividing the sum into the most dominant channel pair, and mapping de-

correlated signal to the other channel for surround effect. In this case, the

dominant channel pair may correspond to left channel and center channel in case that certain input is positioned at point between left and center.

3) Estimating weight value of the most dominant channel, giving attenuated correlated signal to the other channel, which value is a relative value of the estimated weight value.

4) Using weight values for each channel pair, combining the de-correlated signal properly, then setting to a side information for each channel.

1.3.2 A case that downmix processing unit includes a mixing part orresponding to 2x4 matrix

FIGS. 10A to 10C are exemplary block diagrams of a first embodiment of a lownmix processing unit illustrated in FIG. 7. As previously stated, a first mbodiment of a downmix processing unit 720a (hereinafter simply 'a downmix rocessing unit 720a') may be implementation of rendering module 900.

First of all, assuming that $D_{11} = D_{21} = aD$ and $D_{12} = D_{22} = bD$, the ormula 12 is simplified as follow.

[formula 15]

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} E_{11} & E_{21} \\ E_{12} & E_{22} \end{bmatrix} \begin{bmatrix} O_1 \\ O_2 \end{bmatrix} + \begin{bmatrix} aD & aD \\ bD & bD \end{bmatrix} \begin{bmatrix} O_1 \\ O_2 \end{bmatrix}$$

The downmix processing unit according to the formula 15 is illustrated FIG.

)A. Referring to FIG. 10A, a downmix processing unit 720a can be configured to

ypass input signal in case of mono input signal (m), and to process input signal in

ase of stereo input signal (L, R). The downmix processing unit 720a may include a

e-correlating part 722a and a mixing part 724a. The de-correlating part 722a has a

e-correlator aD and de-correlator bD which can be configured to de-correlate

iput signal. The de-correlating part 722a may correspond to a 2x2 matrix. The

nixing part 724a can be configured to map input signal and the de-correlated

ignal to each channel. The mixing part 724a may correspond to a 2x4 matrix.

Secondly, assuming that $D_{11} = aD_1$, $D_{21} = bD_1$, $D_{12} = cD_2$, and

$D_{22} = dD_2$, the formula 12 is simplified as follow.

[formula 15-2]

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} E_{11} & E_{21} \\ E_{12} & E_{22} \end{bmatrix} \begin{bmatrix} O_1 \\ O_2 \end{bmatrix} + \begin{bmatrix} aD_1 & bD_1 \\ cD_2 & dD_2 \end{bmatrix} \begin{bmatrix} O_1 \\ O_2 \end{bmatrix}$$

The downmix processing unit according to the formula 15 is illustrated FIG.

lOB. Referring to FIG. 10B, a de-correlating part 722' including two de-correlators

Ɔ₁, D₂ can be configured to generate de-correlated signals D₁(a*O₁+b*O₂),

Ɔ₂(c*O₁+d*O₂).

Thirdly, assuming that $D_{11} = D_1$, $D_{21} = 0$, $D_{12} = 0$, and $D_{22} = D_2$,

the formula 12 is simplified as follow.

[formula 15-3]

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} E_{11} & E_{21} \\ E_{12} & E_{22} \end{bmatrix} \begin{bmatrix} O_1 \\ O_2 \end{bmatrix} + \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} \begin{bmatrix} O_1 \\ O_2 \end{bmatrix}$$

The downmix processing unit according to the formula 15 is illustrated FIG.

0C. Referring to FIG. 10C, a de-correlating part 722″ including two de-correlators

D_1, D_2 can be configured to generate de-correlated signals $D_1(O_1)$, $D_2(O_2)$.

1.3.2 A case that downmix processing unit includes a mixing part

corresponding to 2x3 matrix

The foregoing formula 15 can be represented as follow:

[formula 16]

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} E_{11} & E_{21} \\ E_{12} & E_{22} \end{bmatrix} \begin{bmatrix} O_1 \\ O_2 \end{bmatrix} + \begin{bmatrix} aD(O_1 + O_2) \\ bD(O_1 + O_2) \end{bmatrix}$$

$$= \begin{bmatrix} E_{11} & E_{21} & \alpha \\ E_{12} & E_{22} & \beta \end{bmatrix} \begin{bmatrix} O_1 \\ O_2 \\ D(O_1 + O_2) \end{bmatrix}$$

The matrix **R** is a 2x3 matrix, the matrix **O** is a 3x1 matrix, and the C is a 2x1

matrix.

FIG. 11 is an exemplary block diagram of a second embodiment of a

downmix processing unit illustrated in FIG. 7. As previously stated, a second

embodiment of a downmix processing unit 720b (hereinafter simply 'a downmix

processing unit 720b') may be implementation of rendering module 900 like the

downmix processing unit 720a. Referring to FIG. 11, a downmix processing unit

!0b can be configured to skip input signal in case of mono input signal (m), and to

:ocess input signal in case of stereo input signal (L, R). The downmix processing

nit 720b may include a de-correlating part 722b and a mixing part 724b. The de-

)rrelating part 722b has a de-correlator D which can be configured to de-correlate

iput signal O_1, O_2 and output the de-correlated signal $D(O_1+O_2)$. The de-

)rrelating part 722b may correspond to a 1x2 matrix. The mixing part 724b can be

onfigured to map input signal and the de-correlated signal to each channel. The

iixing part 724b may correspond to a 2x3 matrix which can be shown as a matrix

l in the formula 16.

Furthermore, the de-correlating part 722b can be configured to de-correlate a

lifference signal O_1-O_2 as common signal of two input signal O_1, O_2. The mixing

)art 724b can be configured to map input signal and the de-correlated common

ignal to each channel.

1.3.3 A case that downmix processing unit includes a mixing part with

ieveral matrixes

Certain object signal can be audible as a similar impression anywhere

without being positioned at a specified position, which may be called as a 'spatial

iound signal'. For example, applause or noises of a concert hall can be an example

of the spatial sound signal. The spatial sound signal needs to be playback via all

ipeakers. If the spatial sound signal playbacks as the same signal via all speakers, it

hard to feel spatialness of the signal because of high inter-correlation (IC) of the

gnal. Hence, there's need to add correlated signal to the signal of each channel

gnal.

FIG. 12 is an exemplary block diagram of a third embodiment of a downmix

:ocessing unit illustrated in FIG. 7. Referring to FIG.12, a third embodiment of a

>wnmix processing unit 720c (hereinafter simply 'a downmix processing unit

!0c') can be configured to generate spatial sound signal using input signal O_i,

·hich may include a de-correlating part 722c with N de-correlators and a mixing

art 724c. The de-correlating part 722c may have N de-correlators D_1, D_2, $\cdots$, D_N

·hich can be configured to de-correlate the input signal O_i. The mixing part 724c

ιay have N matrix R_j, R_k, $\cdots$, R_l which can be configured to generate output signals

j, C_k, $\cdots$, C_l using the input signal O_i and the de-correlated signal $D_X(O_i)$. The R_j

ιatrix can be represented as the following formula.

[formula 17]

$$\mathbf{C}_{j\_i} = \mathbf{R}_j \mathbf{O}_i$$

$$C_{j\_i} = \begin{bmatrix} \alpha_{j\_i}\cos(\theta_{j\_i}) & \alpha_{j\_i}\sin(\theta_{j\_i}) \end{bmatrix} \begin{bmatrix} o_i \\ Dx(o_i) \end{bmatrix}$$

O_i is i^{th} input signal, R_j is a matrix mapping i^{th} input signal O_i to j^{th} channel,

ιnd C_{j_i} is j^{th} output signal. The θ_{j_i} value is de-correlation rate.

The θ_{j_i} value can be estimated base on ICC included in multi-channel

parameter. Furthermore, the mixing part 724c can generate output signals base on

patialness information composing de-correlation rate θ_{j_i} received from user-

nterface via the information generating unit 710, which does not put limitation on

present invention.

The number of de-correlators (N) can be equal to the number of output

hannels. On the other hand, the de-correlated signal can be added to output

hannels selected by user. For example, it is able to position certain spatial sound

ignal at left, right, and center and to output as a spatial sound signal via left

hannel speaker.

1.3.4 A case that downmix processing unit includes a further downmixing

part

FIG. 13 is an exemplary block diagram of a fourth embodiment of a

lownmix processing unit illustrated in FIG. 7. A fourth embodiment of a downmix

processing unit 720d (hereinafter simply 'a downmix processing unit 720d') can be

configured to bypass if the input signal corresponds to a mono signal (m). The

lownmix processing unit 720d includes a further downmixing part 722d which can

pe configured to downmix the stereo signal to be mono signal if the input signal

corresponds to a stereo signal. The further downmixed mono channel (m) is used

is input to the multi-channel decoder 730. The multi-channel decoder 730 can

ıntrol object panning (especially cross-talk) by using the mono input signal. In

ıis case, the information generating unit 710 may generate a multi-channel

arameter base on 5-1-5₁ configuration of MPEG Surround standard.

Furthermore, if gain for the mono downmix signal like the above-mentioned

rtistic downmix gain ADG of FIG. 2 is applied, it is able to control object panning

nd object gain more easily. The ADG may be generated by the information

enerating unit 710 based on mix information.

2. Upmixing channel signals and controlling object signals

FIG. 14 is an exemplary block diagram of a bitstream structure of a

ompressed audio signal according to a second embodiment of present invention.

ꓛIG. 15 is an exemplary block diagram of an apparatus for processing an audio

iignal according to a second embodiment of present invention. Referring to (a) of

ꓛIG. 14, downmix signal α, multi-channel parameter β, and object parameter γ are

ncluded in the bitstream structure. The multi-channel parameter β is a parameter

ꞓor upmixing the downmix signal. On the other hand, the object parameter γ is a

parameter for controlling object panning and object gain. Referring to (b) of FIG. 14,

downmix signal α, a default parameter β′, and object parameter γ are included in

the bitstream structure. The default parameter β′ may include preset information

for controlling object gain and object panning. The preset information may

orrespond to an example suggested by a producer of an encoder side. For example,

reset information may describes that guitar signal is located at a point between

eft and center, and guitar's level is set to a certain volume, and the number of

utput channel in this time is set to a certain channel. The default parameter for

ither each frame or specified frame may be present in the bitstream. Flag

nformation indicating whether default parameter for this frame is different from

lefault parameter of previous frame or not may be present in the bitstream. By

ncluding default parameter in the bitstream, it is able to take less bitrates than side

nformation with object parameter is included in the bitstream. Furthermore,

leader information of the bitstream is omitted in the FIG. 14. Sequence of the

itstream can be rearranged.

Referring to FIG. 15, an apparatus for processing an audio signal according

o a second embodiment of present invention 1000 (hereinafter simply 'a decoder

.000') may include a bitstream de-multiplexer 1005, an information generating unit

.010, a downmix processing unit 1020, and a multil-channel decoder 1030. The de-

nultiplexer 1005 can be configured to divide the multiplexed audio signal into a

lownmix α, a first multi-channel parameter β, and an object parameter γ. The

nformation generating unit 1010 can be configured to generate a second multi-

:hannel parameter using an object parameter γ and a mix parameter. The mix

)arameter comprises a mode information indicating whether the first multi-

annel information β is applied to the processed downmix. The mode information

lay corresponds to an information for selecting by a user. According to the mode

iformation, the information generating information 1020 decides whether to

:ansmit the first multi-channel parameter β or the second multi-channel parameter.

The downmix processing unit 1020 can be configured to determining a

rocessing scheme according to the mode information included in the mix

iformation. Furthermore, the downmix processing unit 1020 can be configured to

process the downmix α according to the determined processing scheme. Then the

lownmix processing unit 1020 transmits the processed downmix to multi-channel

lecoder 1030.

The multi-channel decoder 1030 can be configured to receive either the first

nulti-channel parameter β or the second multi-channel parameter. In case that

iefault parameter β′ is included in the bitstream, the multi-channel decoder 1030

:an use the default parameter β′ instead of multi-channel parameter β.

Then, the multi-channel decoder 1030 can be configured to generate multi-

:hannel output using the processed downmix signal and the received multi-

:hannel parameter. The multi-channel decoder 1030 may have the same

configuration of the former multi-channel decoder 730, which does not put

limitation on the present invention.

3. Binaural Processing

A multi-channel decoder can be operated in a binaural mode. This enables a multi-channel impression over headphones by means of Head Related Transfer Function (HRTF) filtering. For binaural decoding side, the downmix signal and multi-channel parameters are used in combination with HRTF filters supplied to the decoder.

FIG. 16 is an exemplary block diagram of an apparatus for processing an audio signal according to a third embodiment of present invention. Referring to FIG. 16, an apparatus for processing an audio signal according to a third embodiment (hereinafter simply 'a decoder 1100') may comprise an information generating unit 1110, a downmix processing unit 1120, and a multi-channel decoder 1130 with a sync matching part 1130a.

The information generating unit 1110 may have the same configuration of the information generating unit 710 of FIG. 7, with generating dynamic HRTF. The downmix processing unit 1120 may have the same configuration of the downmix processing unit 720 of FIG. 7. Like the preceding elements, multi-channel decoder 1130 except for the sync matching part 1130a is the same case of the former elements. Hence, details of the information generating unit 1110, the downmix processing unit 1120, and the multi-channel decoder 1130 shall be omitted.

The dynamic HRTF describes the relation between object signals and virtual peaker signals corresponding to the HRTF azimuth and elevation angles, which is ime-dependent information according to real-time user control.

The dynamic HRTF may correspond to one of HTRF filter coefficients itself, arameterized coefficient information, and index information in case that the nulti-channel decoder comprise all HRTF filter set.

There's need to match a dynamic HRTF information with frame of downmix ignal regardless of kind of the dynamic HRTF. In order to match HRTF nformation with downmix signal, it able to provide three type of scheme as ollows:

1) Inserting a tag information into each HRTF information and bitstream lownmix signal, then matching the HRTF with bitstream downmix signal based on he inserted tag information. In this scheme, it is proper that tag information may e included in ancillary field in MPEG Surround standard. The tag information nay be represented as a time information, a counter information, a index nformation, etc.

2) Inserting HRTF information into frame of bitstream. In this scheme, it is ossible to set to mode information indicating whether current frame corresponds o a default mode or not. If the default mode which describes HRTF information of

ırrent frame is equal to the HRTF information of previous frame is applied, it is

ɔle to reduce bitrates of HRTF information.

2-1) Furthermore, it is possible to define transmission information indicating

'hether HRTF information of current frame has already transmitted. If the

ʾansmission information which describes HRTF information of current frame is

ʄual to the transmitted HRTF information of frame is applied, it is also possible to

ɜduce bitrates of HRTF information.

3) Transmitting several HRTF informations in advance, then transmitting

lentifying information indicating which HRTF among the transmitted HRTF

ʾformations per each frame.

Furthermore, in case that HTRF coefficient varies suddenly, distortion may

ʾe generated. In order to reduce this distortion, it is proper to perform smoothing

ʾf coefficient or the rendered signal.

4. Rendering

FIG. 17 is an exemplary block diagram of an apparatus for processing an

ʾudio signal according to a fourth embodiment of present invention. The apparatus

ʾor processing an audio signal according to a fourth embodiment of present

ʾnvention 1200 (hereinafter simply 'a processor 1200') may comprise an encoder

l210 at encoder side 1200A, and a rendering unit 1220 and a synthesis unit 1230 at

ecoder side 1200B. The encoder 1210 can be configured to receive multi-channel

bject signal and generate a downmix of audio signal and a side information. The

endering unit 1220 can be configured to receive side information from the encoder

210, playback configuration and user control from a device setting or a user-

nterface, and generate rendering information using the side information, playback

onfiguration, and user control. The synthesis unit 1230 can be configured to

ynthesis multi-channel output signal using the rendering information and the

eceived downmix signal from an encoder 1210.

4.1 Applying effect-mode

The effect-mode is a mode for remixed or reconstructed signal. For example,

ive mode, club band mode, karaoke mode, etc may be present. The effect-mode

nformation may correspond to a mix parameter set generated by a producer, other

iser, etc. If the effect-mode information is applied, an end user don't have to

control object panning and object gain in full because user can select one of pre-

determined effect-mode informations.

Two methods of generating an effect-mode information can be distinguished.

First of all, it is possible that an effect-mode information is generated by encoder

1200A and transmitted to the decoder 1200B. Secondly, the effect-mode information

may be generated automatically at the decoder side. Details of two methods shall

be described as follow.

4.1.1 Transmitting effect-mode information to decoder side

The effect-mode information may be generated at an encoder 1200A by a producer. According to this method, the decoder 1200B can be configured to receive side information including the effect-mode information and output user-interface by which a user can select one of effect-mode informations. The decoder 200B can be configured to generate output channel base on the selected effect-node information.

Furthermore, it is inappropriate to hear downmix signal as it is for a listener n case that encoder 1200A downmix the signal in order to raise quality of object signals. However, if effect-mode information is applied in the decoder 1200B, it is possible to playback the downmix signal as the maximum quality.

4.1.2 Generating effect-mode information in decoder side

The effect-mode information may be generated at a decoder 1200B. The decoder 1200B can be configured to search appropriate effect-mode informations for the downmix signal. Then the decoder 1200B can be configured to select one of the searched effect-mode by itself (automatic adjustment mode) or enable a user to select one of them (user selection mode). Then the decoder 1200B can be configured to obtain object information (number of objects, instrument names, etc) included in side information, and control object based on the selected effect-mode information and the object information.

Furthermore, it is able to control similar objects in a lump. For example,

nstruments associated with a rhythm may be similar objects in case of 'rhythm

mpression mode'. Controlling in a lump means controlling each object

imultaneously rather than controlling objects using the same parameter.

Furthermore, it is able to control object based on the decoder setting and

levice environment (including whether headphones or speakers). For example,

)bject corresponding to main melody may be emphasized in case that volume

ietting of device is low, object corresponding to main melody may be repressed in

:ase that volume setting of device is high.

4.2 Object type of input signal at encoder side

The input signal inputted to an encoder 1200A may be classified into three

:ypes as follow.

1) Mono object (mono channel object)

Mono object is most general type of object. It is possible to synthesis internal

downmix signal by simply summing objects. It is also possible to synthesis internal

downmix signal using object gain and object panning which may be one of user

control and provided information. In generating internal downmix signal, it is also

possible to generate rendering information using at least one of object characteristic,

user input, and information provided with object.

In case that external downmix signal is present, it is possible to extract and

ansmit information indicating relation between external downmix and object.

2) Stereo object (stereo channel object)

It is possible to synthesis internal downmix signal by simply summing

ojects like the case of the former mono object. It is also possible to synthesis

iternal downmix signal using object gain and object panning which may be one of

ser control and provided information. In case that downmix signal corresponds to

mono signal, it is possible that encoder 1200A use object converted into mono

gnal for generating downmix signal. In this case, it is able to extract and transfer

iformation associated with object (ex: panning information in each time-frequency

omain) in converting into mono signal. Like the preceding mono object, in

enerating internal downmix signal, it is also possible to generate rendering

iformation using at least one of object characteristic, user input, and information

rovided with object. Like the preceding mono object, in case that external

lownmix signal is present, it is possible to extract and transmit information

ndicating relation between external downmix and object.

3) Multi-channel object

In case of multi-channel object, it is able to perform the above mentioned

nethod described with mono object and stereo object. Furthermore, it is able to

nput multi-channel object as a form of MPEG Surround. In this case, it is able to

enerate object-based downmix (ex: SAOC downmix) using object downmix

1annel, and use multi-channel information (ex: spatial information in MPEG

urround) for generating multi-channel information and rendering information.

[ence, it is possible to reduce computing amount because multi-channel object

resent in form of MPEG Surround don't have to decode and encode using object-

riented encoder (ex: SAOC encoder). If object downmix corresponds to stereo and

bject-based downmix (ex: SAOC downmix) corresponds to mono in this case, it is

ossible to apply the above-mentioned method described with stereo object.

4) Transmitting scheme for variable type of object

As stated previously, variable type of object (mono object, stereo object, and

multi-channel object) may be transmitted from the encoder 1200A to the decoder.

1200B. Transmitting scheme for variable type of object can be provided as follow:

Referring to FIG. 18, when the downmix includes a plural object, a side

nformation includes information for each object. For example, when a plural object

:onsists of Nth mono object (A), left channel of N+1th object (B), and right channel

)f N+1th object (C), a side information includes information for 3 objects (A, B, C).

The side information may comprise correlation flag information indicating

vhether an object is part of a stereo or multi-channel object, for example, mono

:bject, one channel (L or R) of stereo object, and so on. For example, correlation flag

nformation is '0' if mono object is present, correlation flag information is '1' if one

annel of stereo object is present. When one part of stereo object and the other part

stereo object is transmitted in succession, correlation flag information for other

art of stereo object may be any value (ex: '0', '1', or whatever). Furthermore,

orrelation flag information for other part of stereo object may be not transmitted.

Furthermore, in case of multi-channel object, correlation flag information for

ne part of multi-channel object may be value describing number of multi-channel

bject. For example, in case of 5.1 channel object, correlation flag information for

ft channel of 5.1 channel may be '5', correlation flag information for the other

hannel (R, Lr, Rr, C, LFE) of 5.1 channel may be either '0' or not transmitted.

4.3 Object attribute

Object may have the three kinds of attribute as follows:

a) Single object

Single object can be configured as a source. It is able to apply one parameter

o single object for controlling object panning and object gain in generating

lownmix signal and reproducing. The 'one parameter' may mean not only one

parameter for all time/frequency domain but also one parameter for each

ime/frequency slot.

b) Grouped object

Single object can be configured as more than two sources. It is able to apply

one parameter to grouped object for controlling object panning and object gain

although grouped object is inputted as at least two sources. Details of the grouped

object shall be explained with reference to FIG. 19 as follows: Referring to FIG. 19,

an encoder 1300 includes a grouping unit 1310 and a downmix unit 1320. The

grouping unit 1310 can be configured to group at least two objects among inputted

multi-object input, base on a grouping information. The grouping information may

be generated by producer at encoder side. The downmix unit 1320 can be

configured to generate downmix signal using the grouped object generated by the

grouping unit 1310. The downmix unit 1320 can be configured to generate a side

information for the grouped object.

c) Combination object

Combination object is an object combined with at least one source. It is

possible to control object panning and gain in a lump, but keep relation between

combined objects unchanged. For example, in case of drum, it is possible to control

drum, but keep relation between base drum, tam-tam, and symbol unchanged. For

example, when base drum is located at center point and symbol is located at left

point, it is possible to positioning base drum at right point and positioning symbol

at point between center and right in case that drum is moved to right direction.

Relation information between combined objects may be transmitted to a

decoder. On the other hand, decoder can extract the relation information using

combination object.

4.4 Controlling objects hierarchically

It is able to control objects hierarchically. For example, after controlling a

rum, it is able to control each sub-elements of drum. In order to control objects

ierarchically, three schemes is provided as follows:

a) UI (user interface)

Only representative element may be displayed without displaying all objects.

: the representative element is selected by a user, all objects display.

b) Object grouping

After grouping objects in order to represent representative element, it is

ossible to control representative element to control all objects grouped as

epresentative element. Information extracted in grouping process may be

ransmitted to a decoder. Also, the grouping information may be generated in a

lecoder. Applying control information in a lump can be performed based on pre-

letermined control information for each element.

c) Object configuration

It is possible to use the above-mentioned combination object. Information

:oncerning element of combination object can be generated in either an encoder or

a decoder. Information concerning elements from an encoder can be transmitted as

a different form from information concerning combination object.

It will be apparent to those skilled in the art that various modifications and

ariations can be made in the present invention without departing from the spirit

: scope of the inventions. Thus, it is intended that the present invention covers the

iodifications and variations of this invention provided they come within the scope

f the appended claims and their equivalents.

[Industrial Applicability]

Accordingly, the present invention is applicable to encode and decode an

udio signal.

[CLAIMS]

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

    receiving a downmix signal in time domain;

    if the downmix signal corresponds to a mono signal, bypassing the downmix signal;

    if the number of channel of the downmix signal corresponds to at least two, decomposing the downmix signal into a subband signal, and processing the subband signal using a downmix processing information,

    wherein the downmix processing information is estimated based on an object information and a mix information.
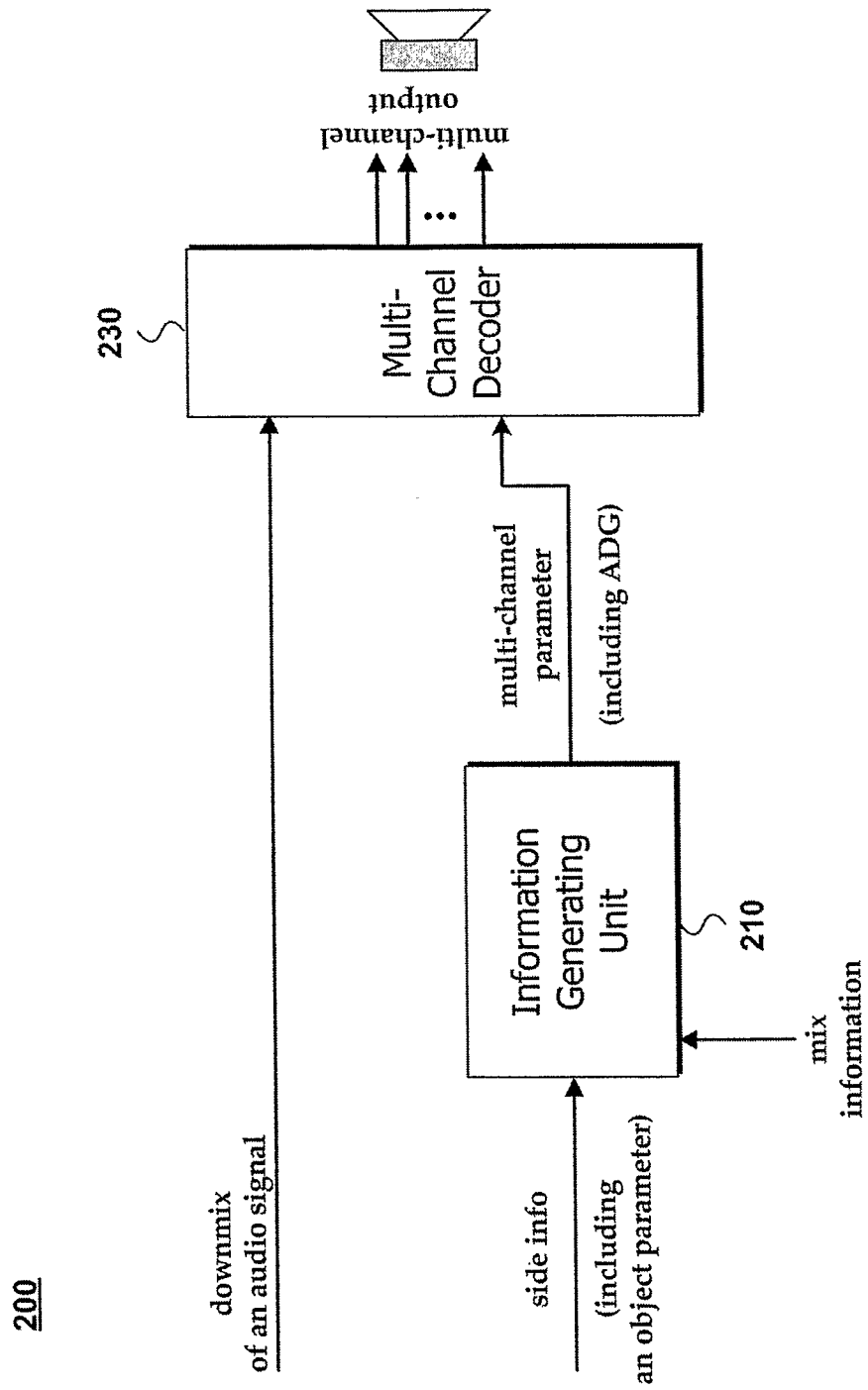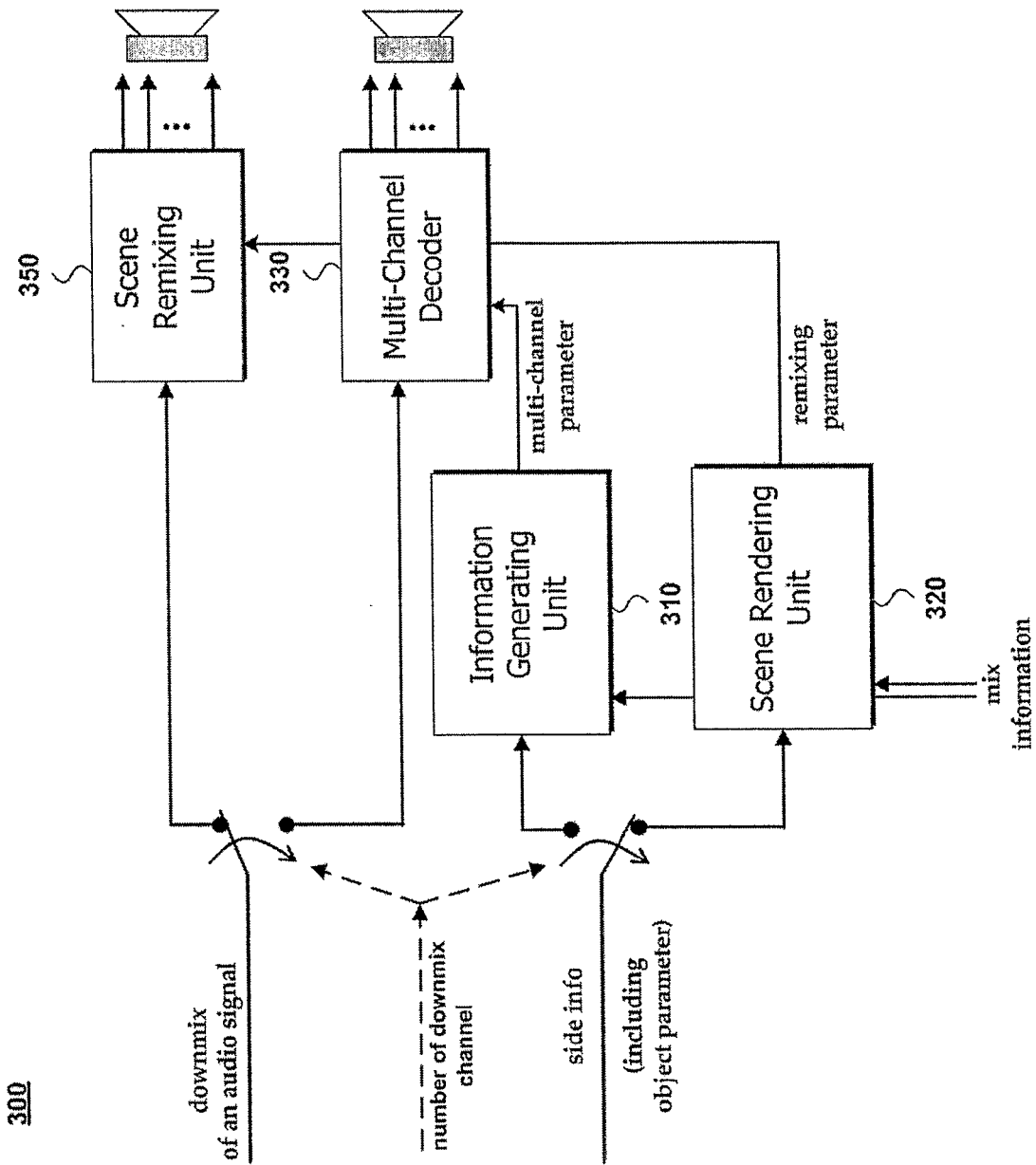
2. The method of claim 1, wherein the number of channel of the downmix signal is equal to the number of channel of the processed downmix signal.

3. The method of claim 1, wherein the object information is included in a side information, and the side information includes a correlation flag information indicating whether an object is part of at least two channel object.

4. The method of claim 1, wherein the object information includes at least one of an object level information and an object correlation information.

5. The method of claim 1, wherein the downmix processing information corresponds to an information for controlling object panning if the number of channel the downmix signal corresponds to at least two.

6. The method of claim 1, wherein the downmix processing information corresponds to an information for controlling object gain.

7. The method of claim 1, further comprising:

> generating a multi-channel signal using the processed subband signal.

8. The method of claim 7, further comprising

> generating a multi-channel information using the object information and the mix information,

> wherein the multi-channel signal is generated based on the multi-channel information.

9. The method of claim 1, further comprising:

> downmixing the downmix signal to be a mono signal if the downmix signal corresponds to a stereo signal.

10. The method of claim 1, wherein the mix information is generated using at least one of an object position information and a playback configuration information.

11. The method of claim 1, wherein the downmix signal is received as a broadcast signal.

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

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

    receiving a downmix signal in time domain;

    if the downmix signal corresponds to a mono signal, bypassing the downmix signal;

    if the number of channel of the downmix signal corresponds to at least two, decomposing the downmix signal into a subband signal, and processing the subband signal using a downmix processing information,

    wherein the downmix processing information is estimated based on an object information and a mix information.

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

    a receiving unit receiving a downmix signal in time domain; and,

    a downmix processing unit bypassing the downmix signal if the downmix signal corresponds to a mono signal, and decomposing the downmix signal into a subband signal and processing the subband signal using a downmix processing information if the number of channel of the downmix signal corresponds to at least two,

    wherein the downmix processing information is estimated based on an object information and a mix information.

FIG. 1
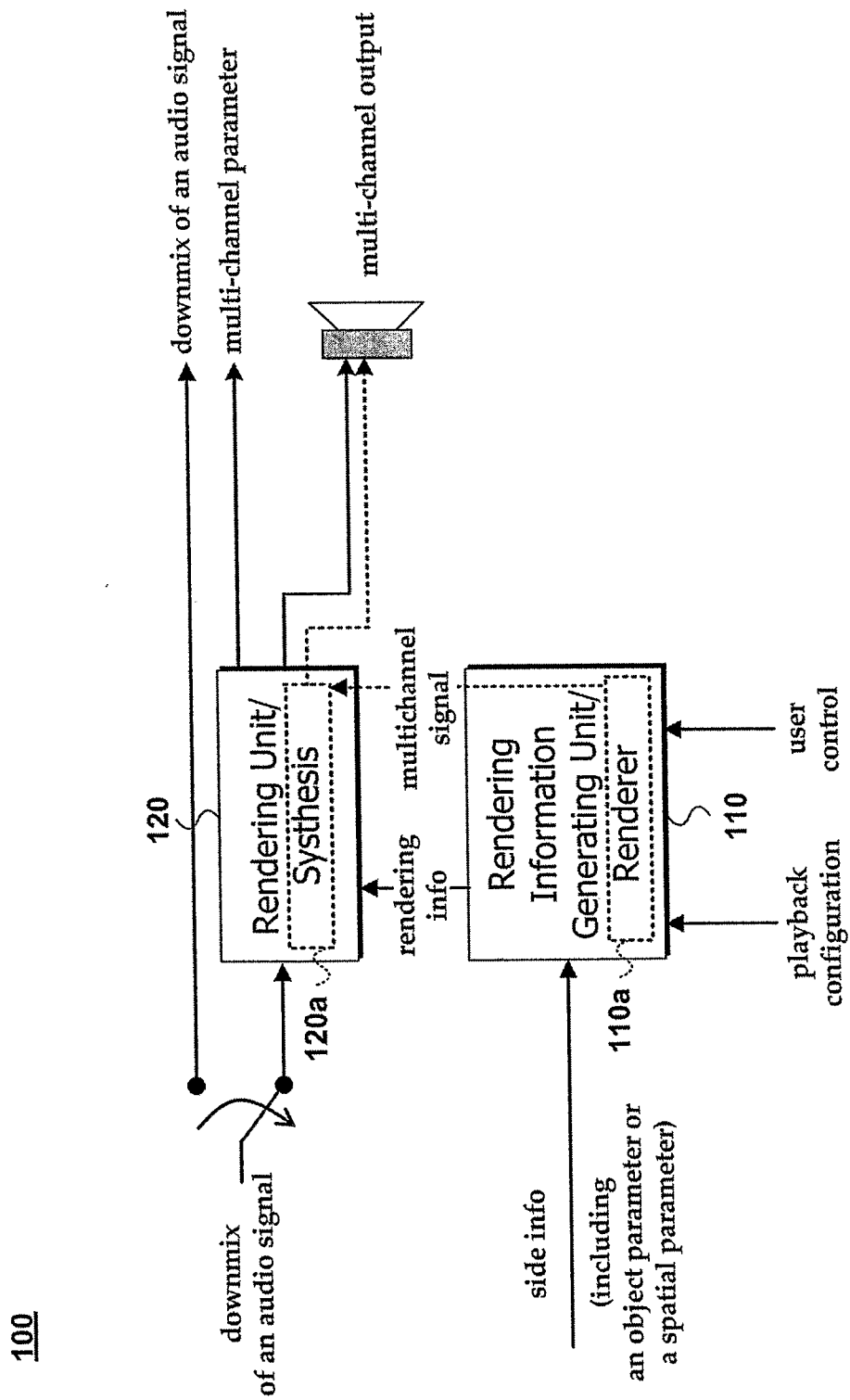
FIG. 2

FIG. 3

FIG. 4
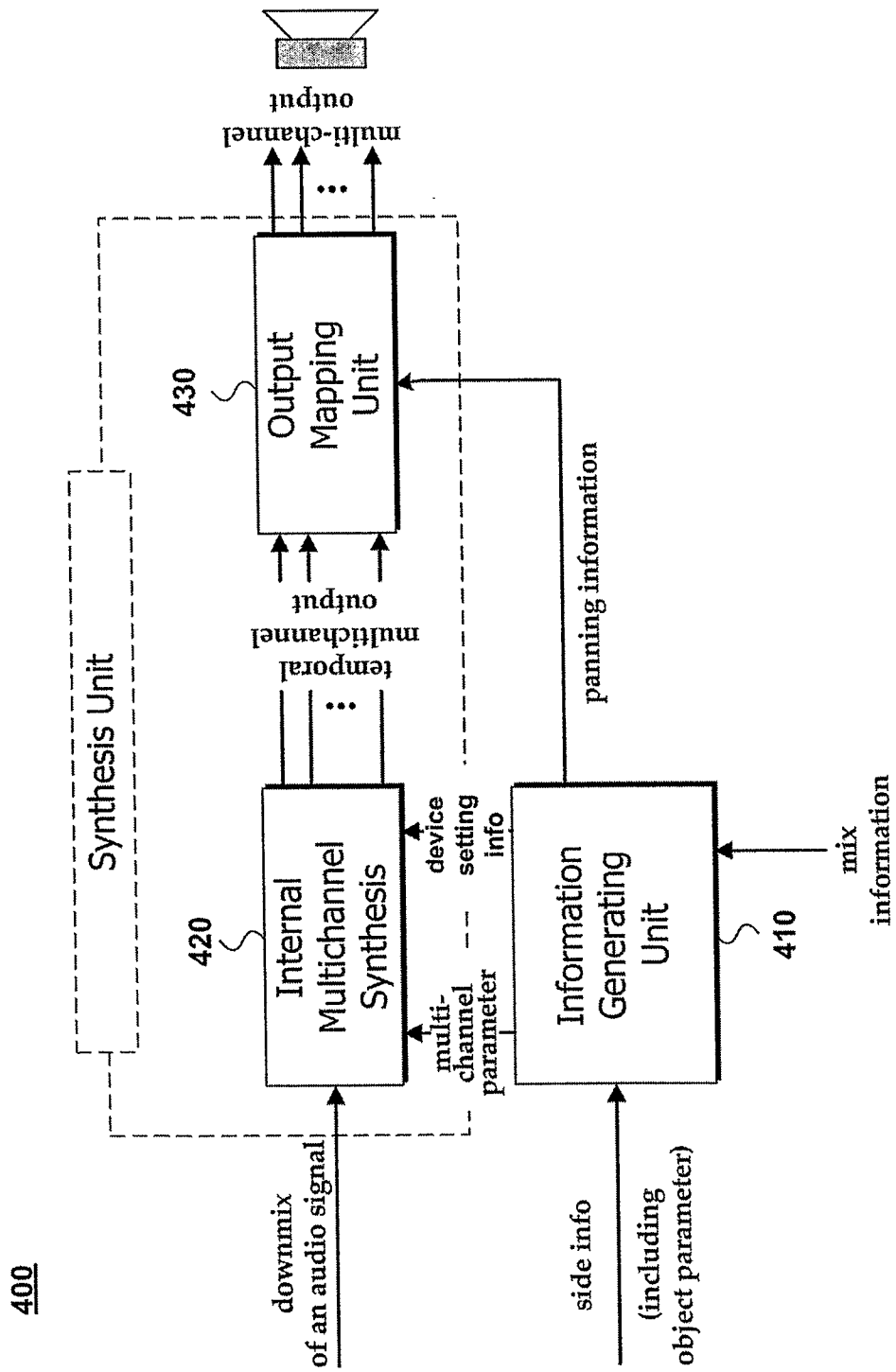
FIG. 5

FIG. 6

FIG. 7

FIG. 8

800
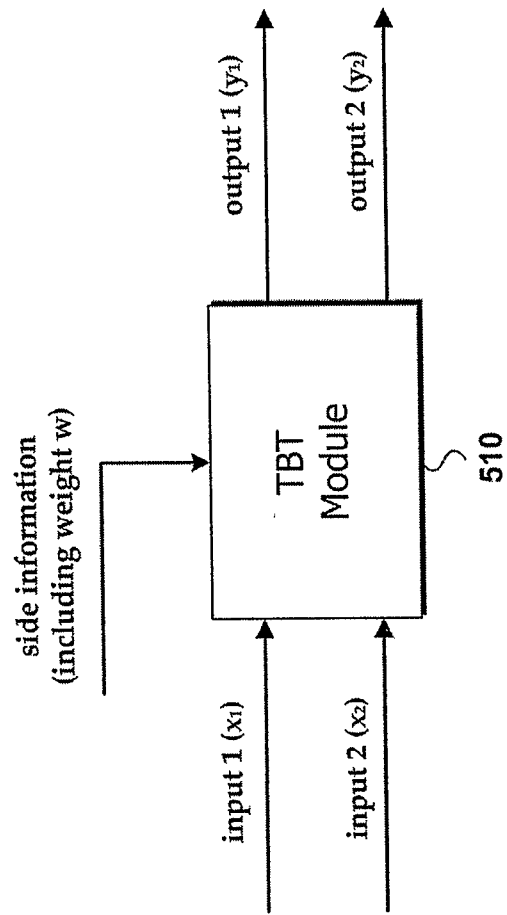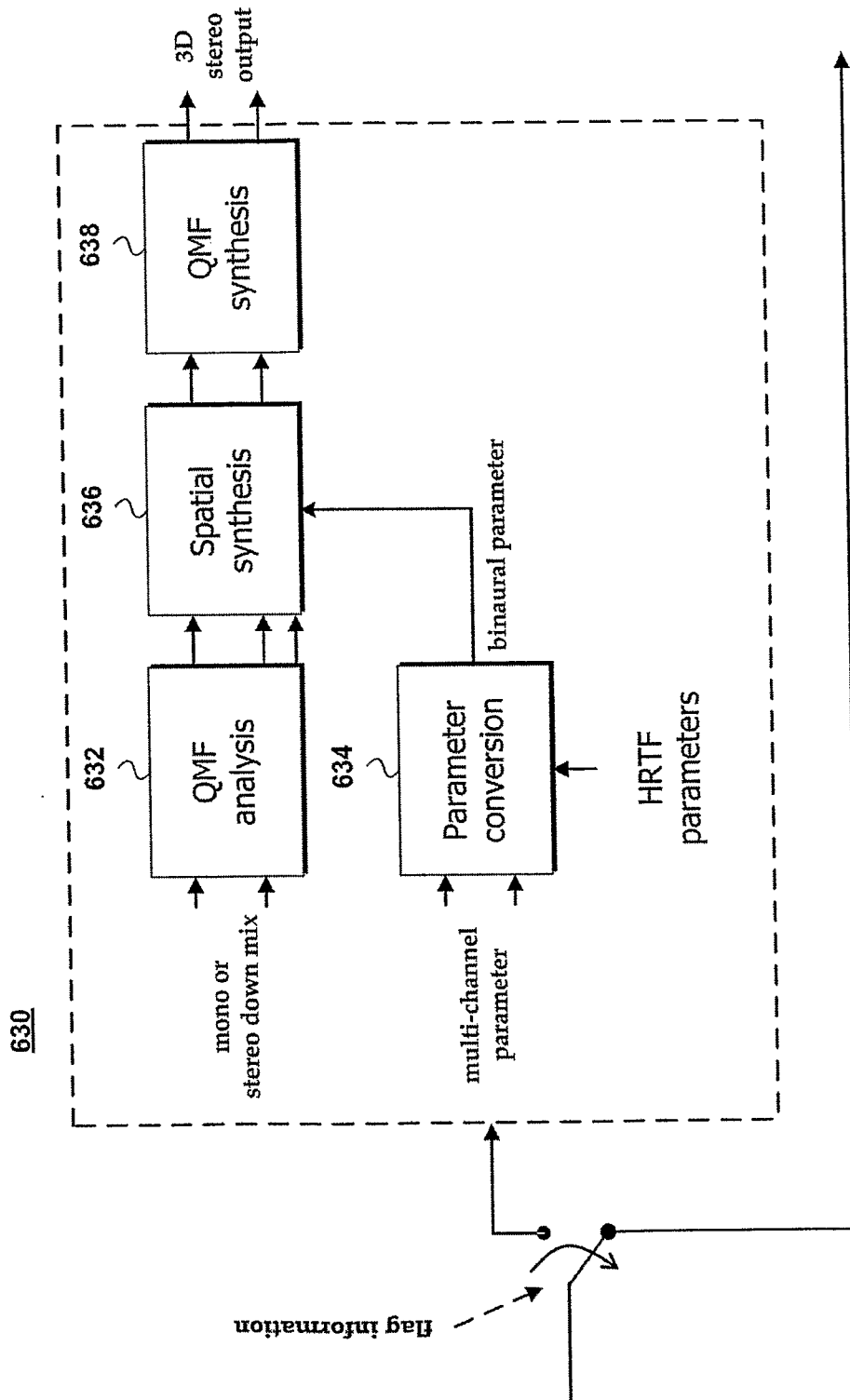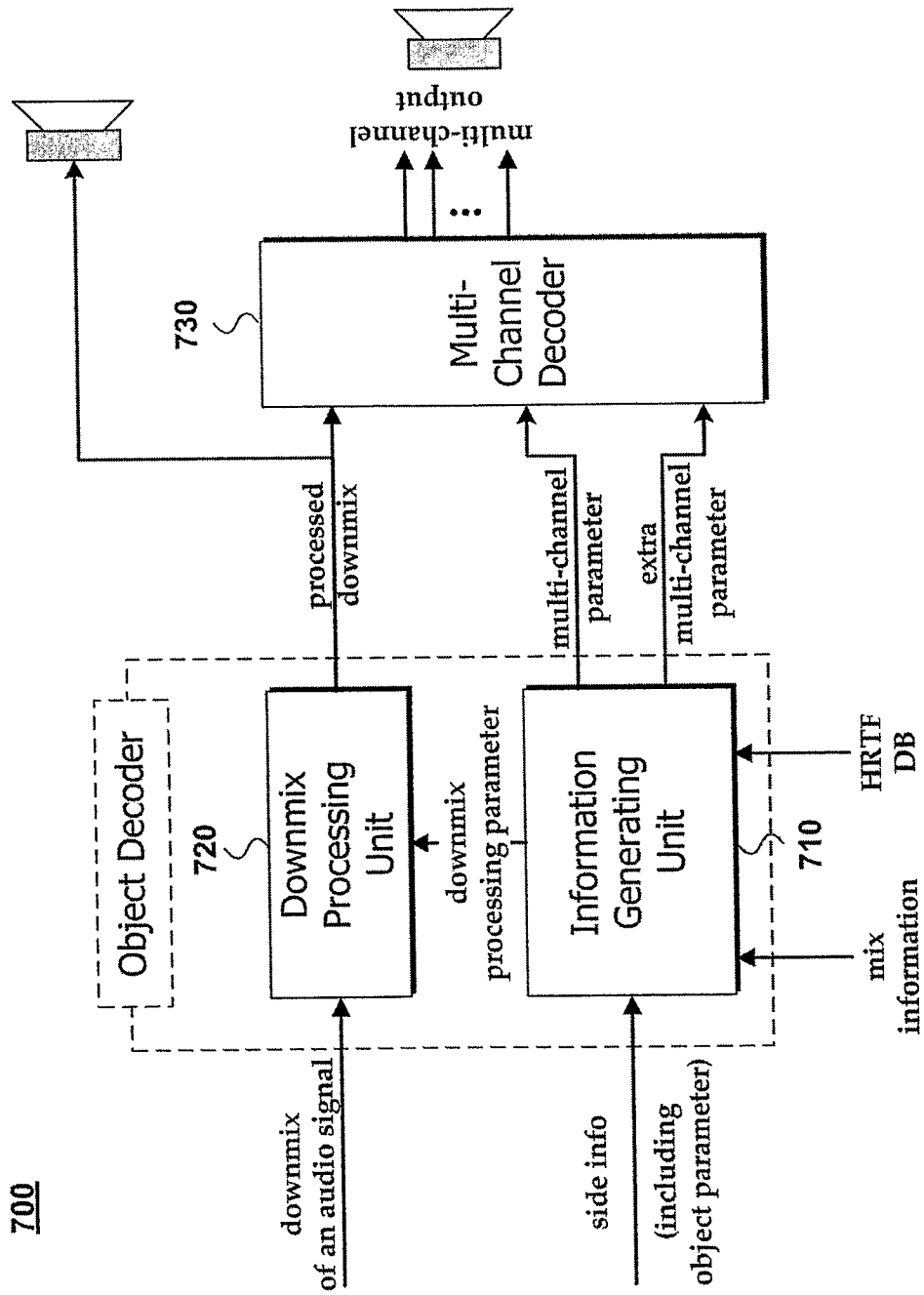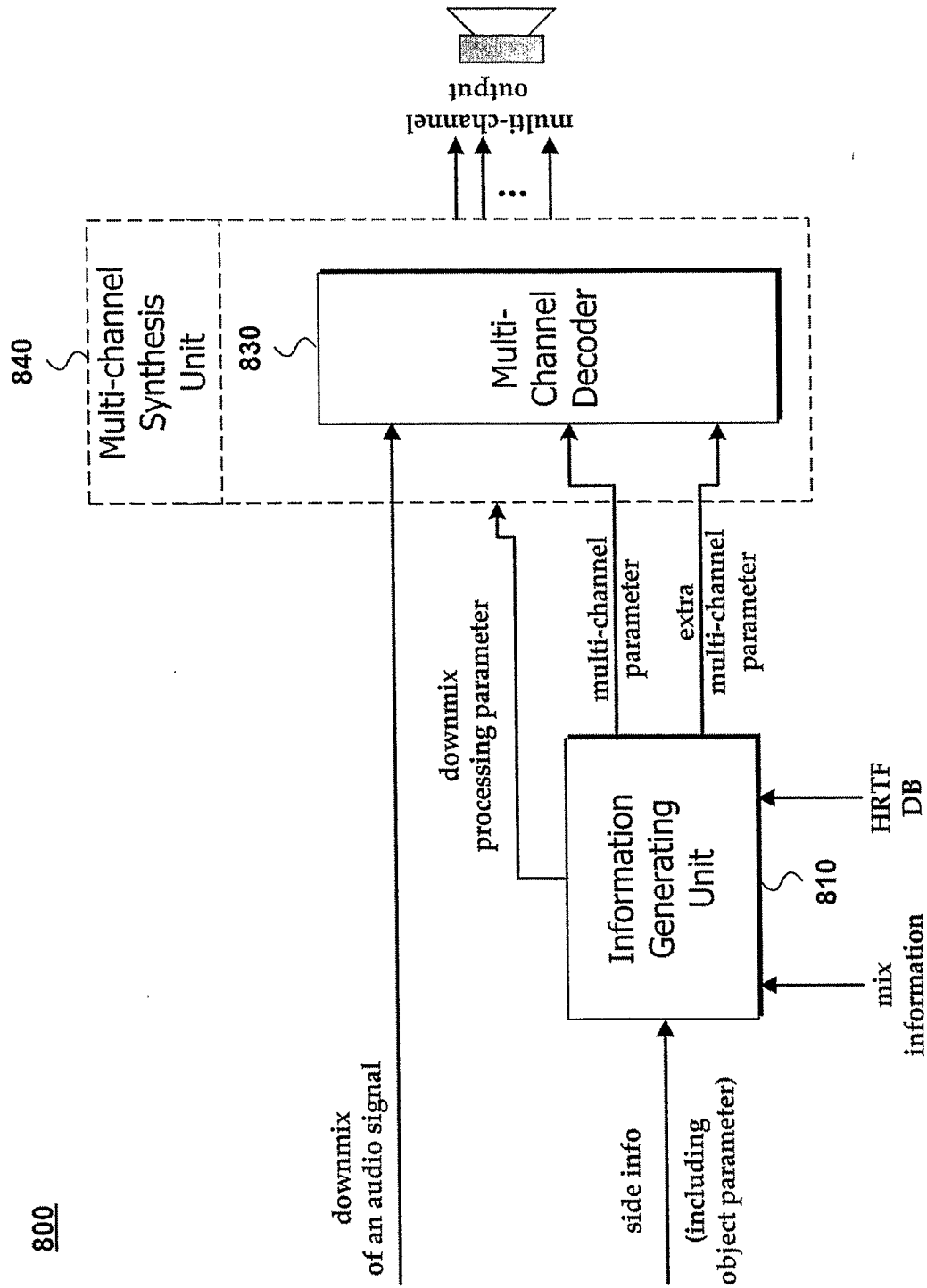
FIG. 9
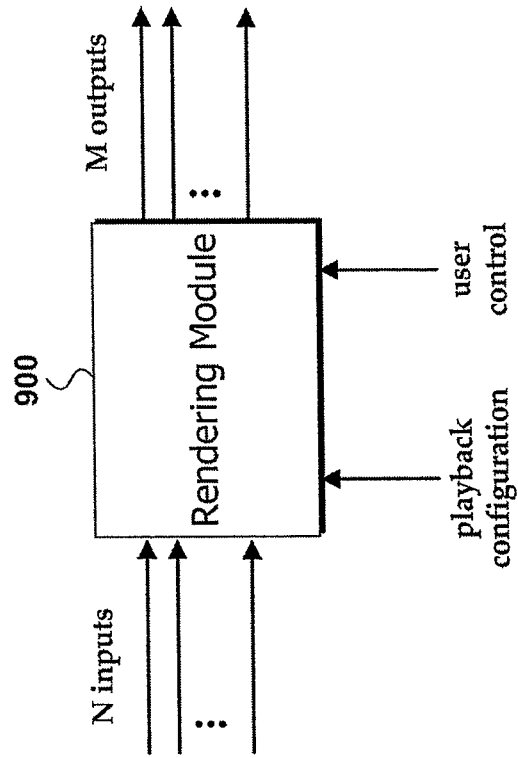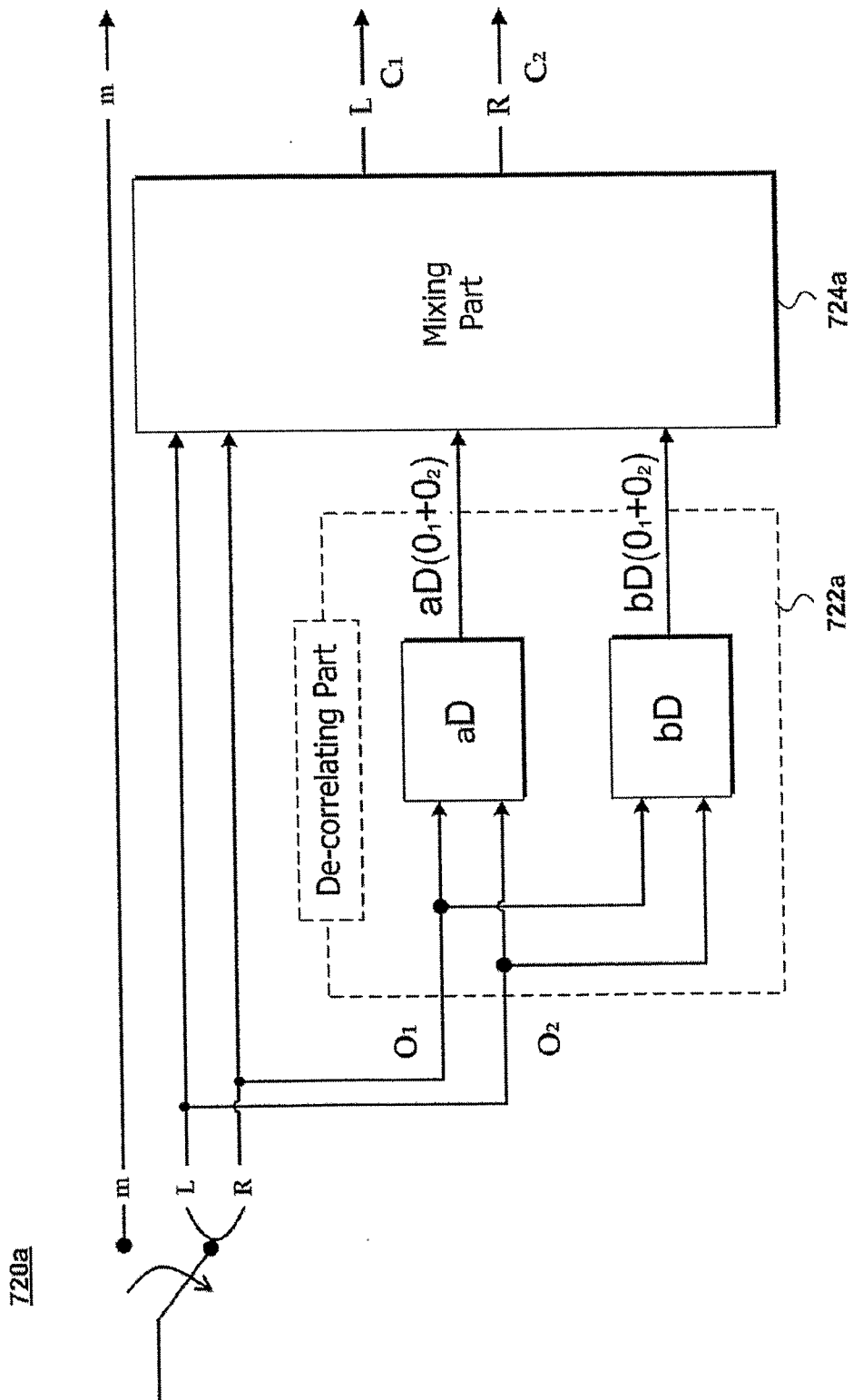
FIG. 10A

FIG. 10B
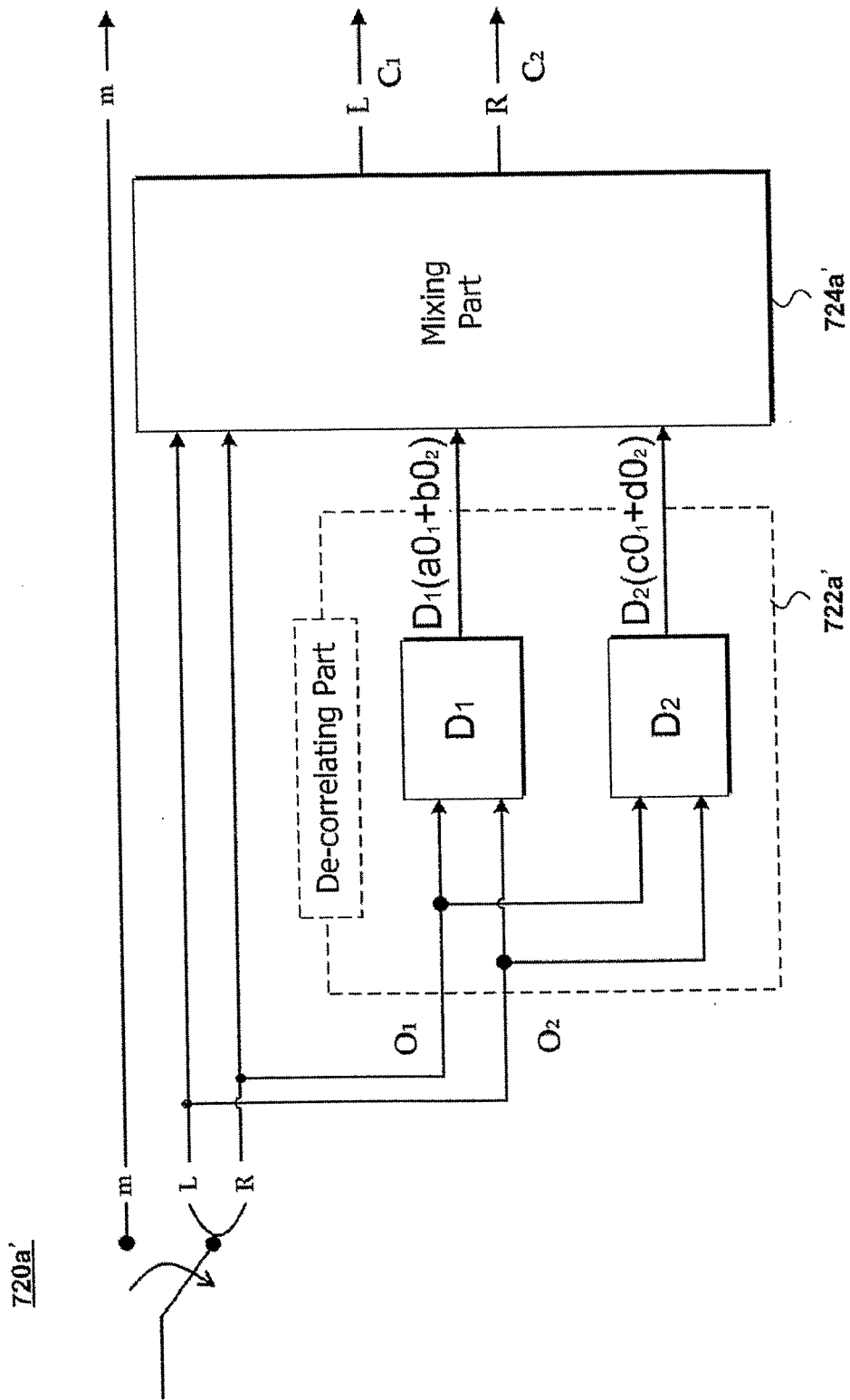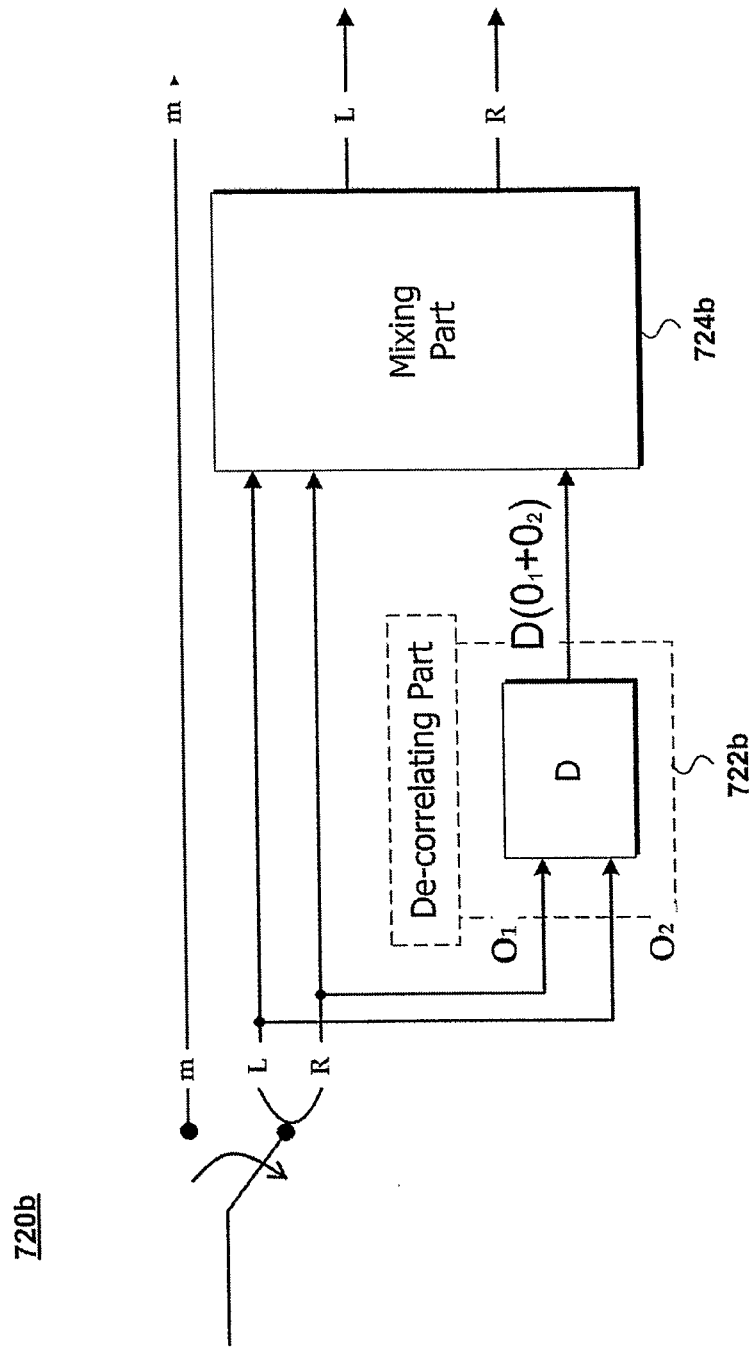
FIG. 10C

FIG. 11

FIG. 12

FIG. 13
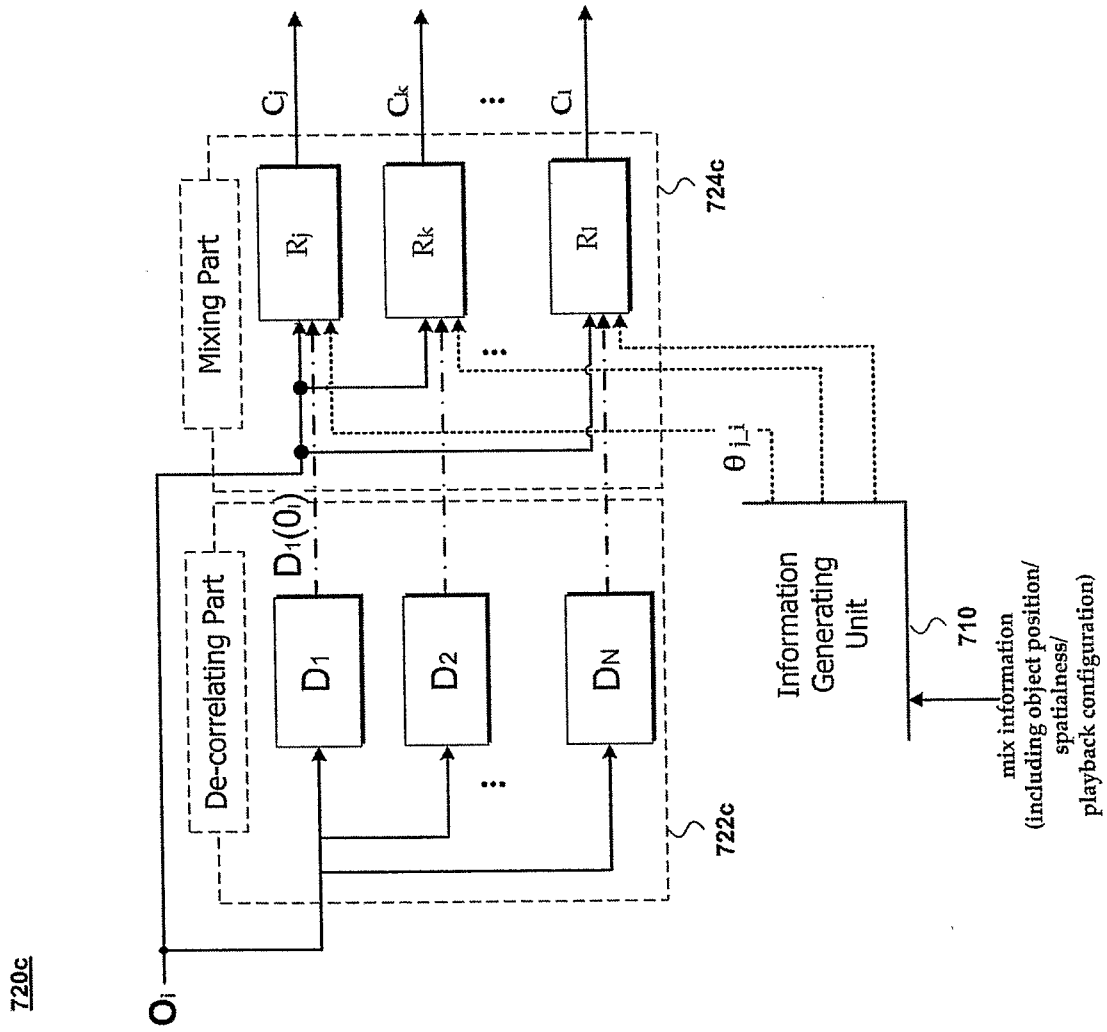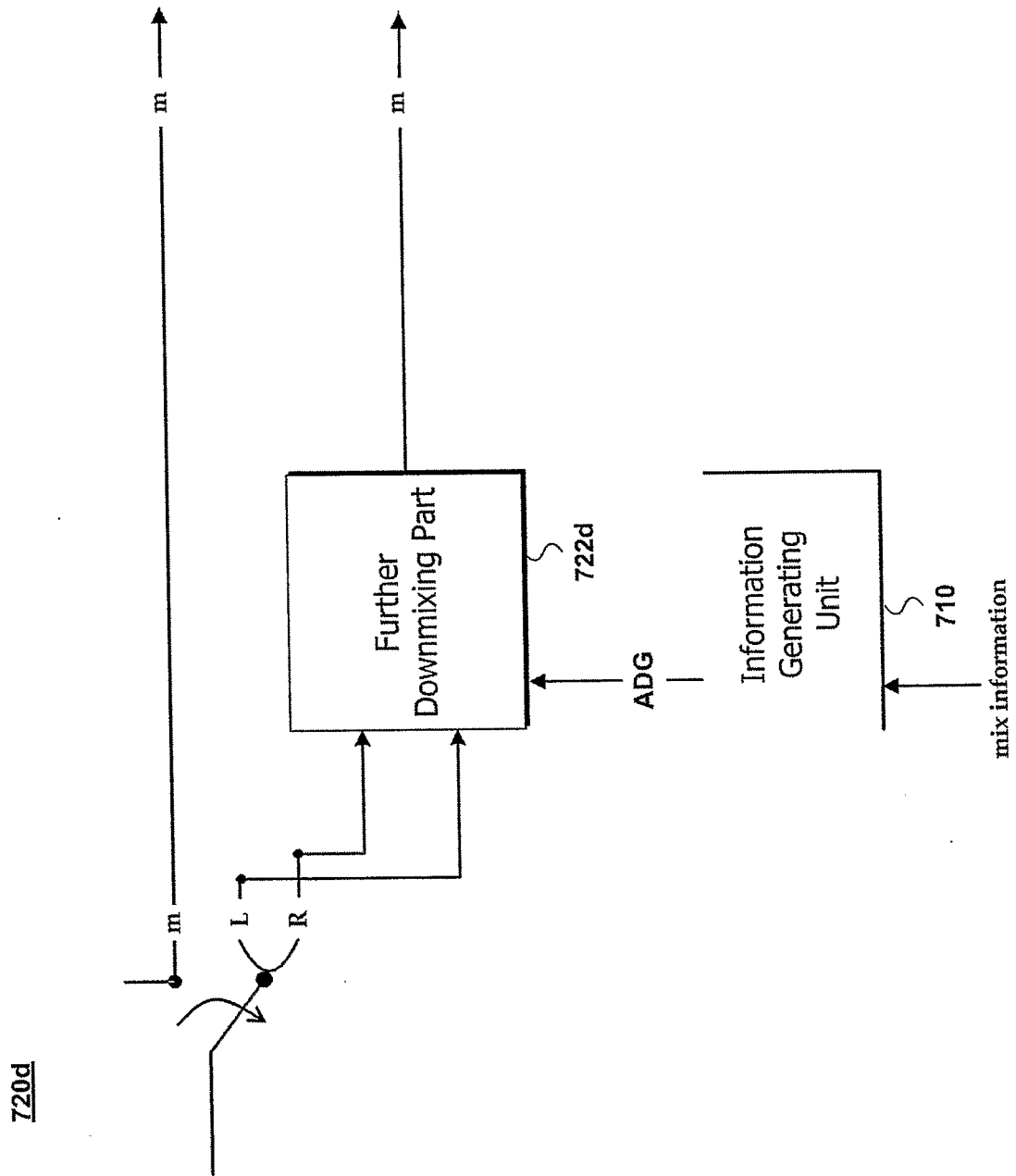
FIG. 14
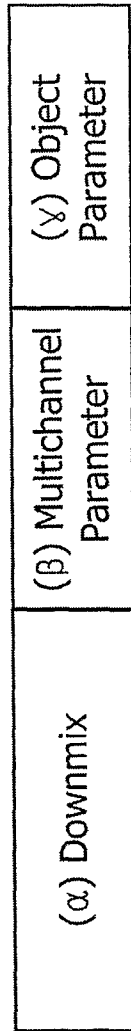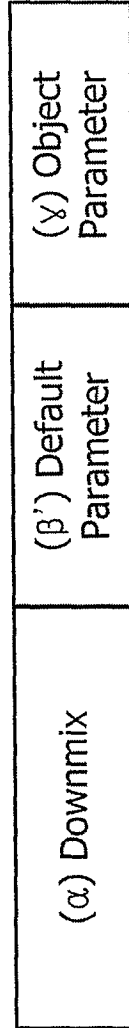
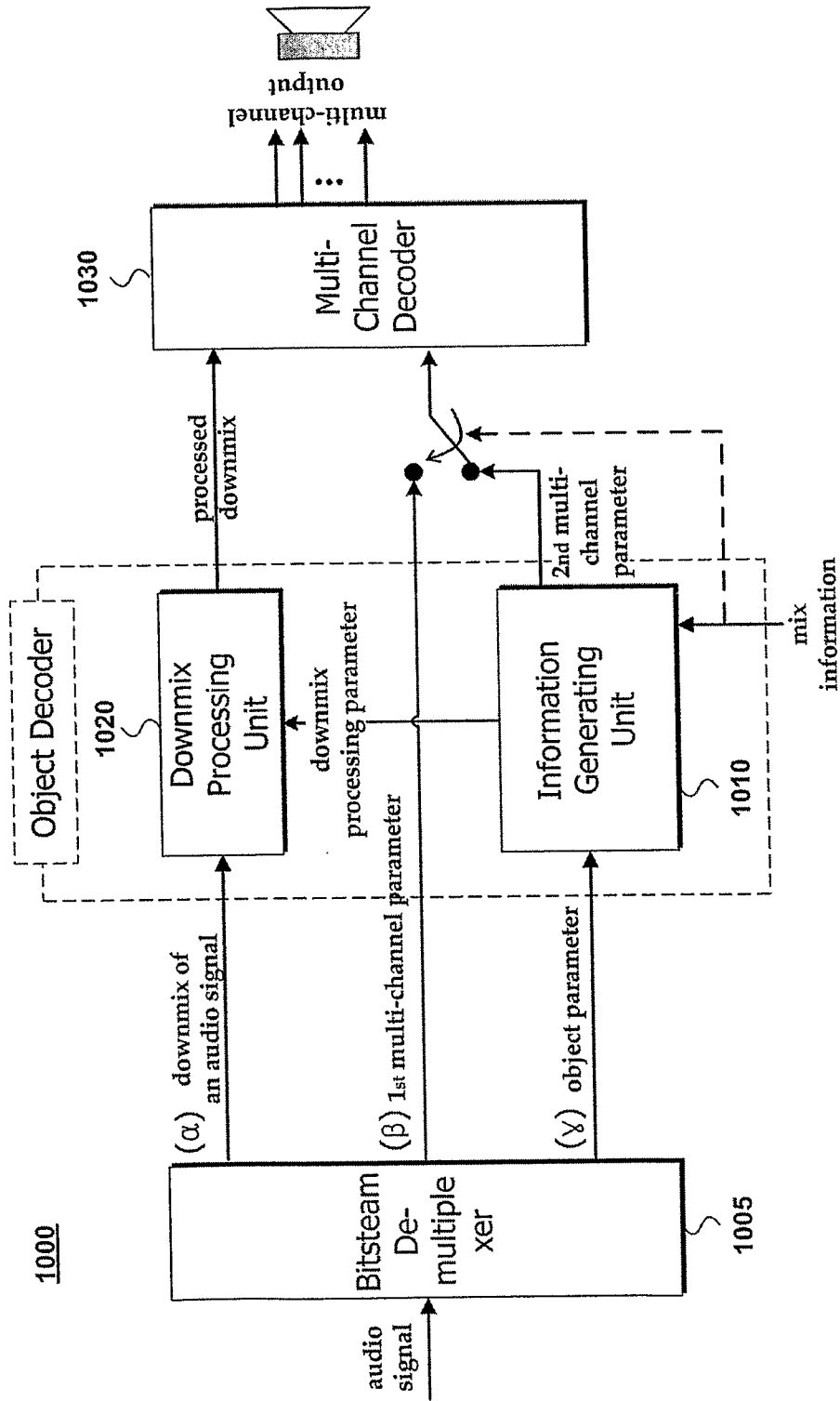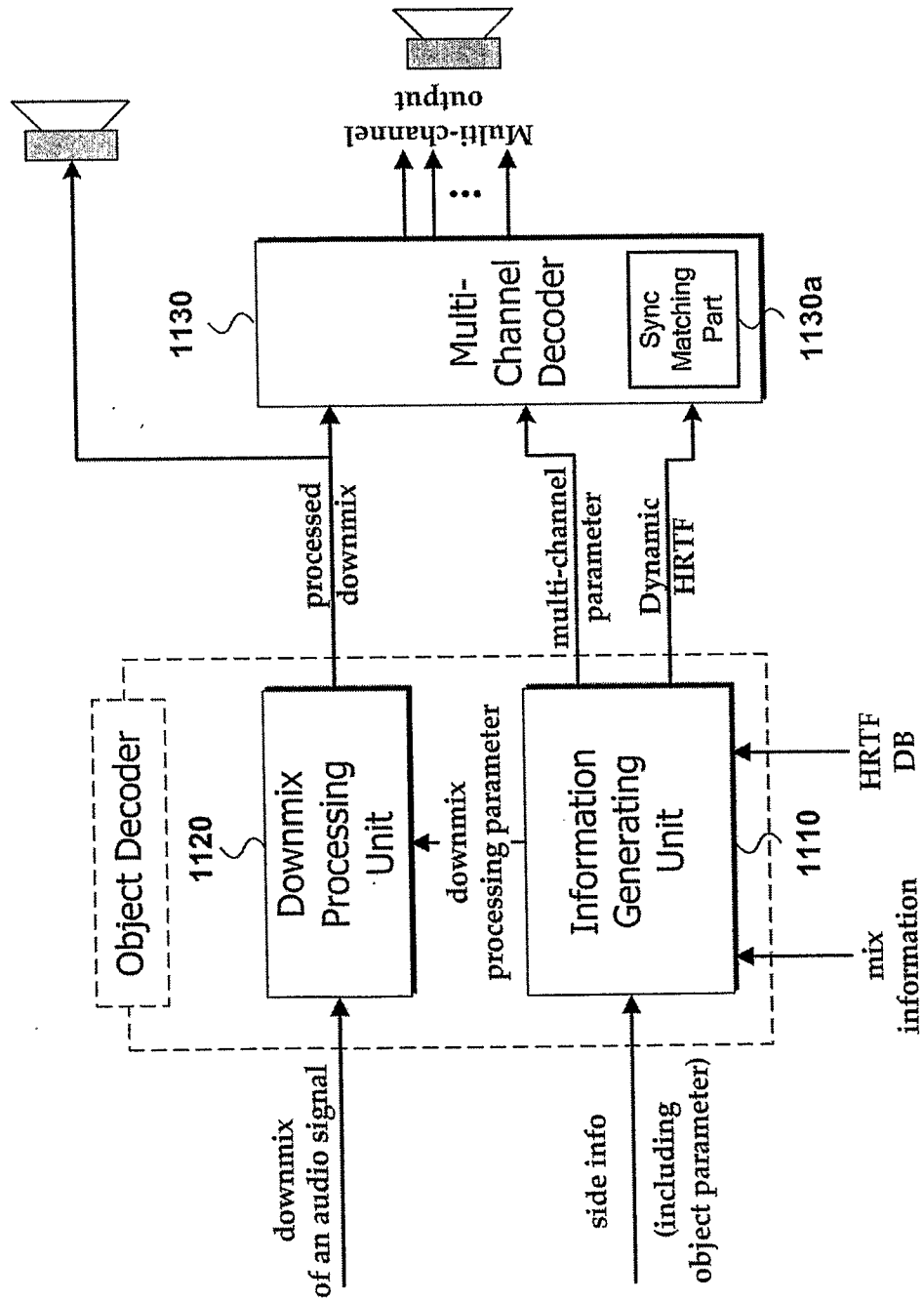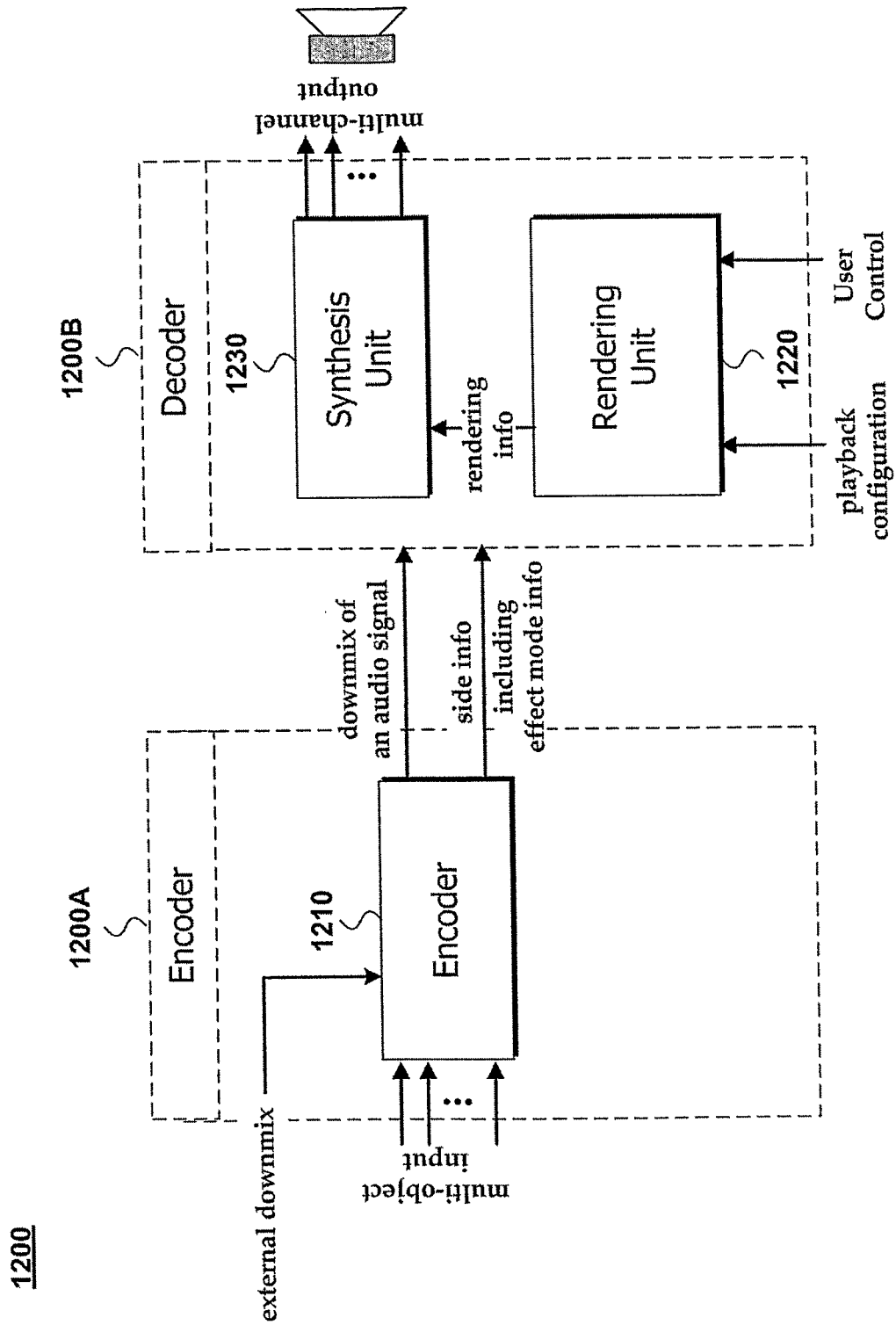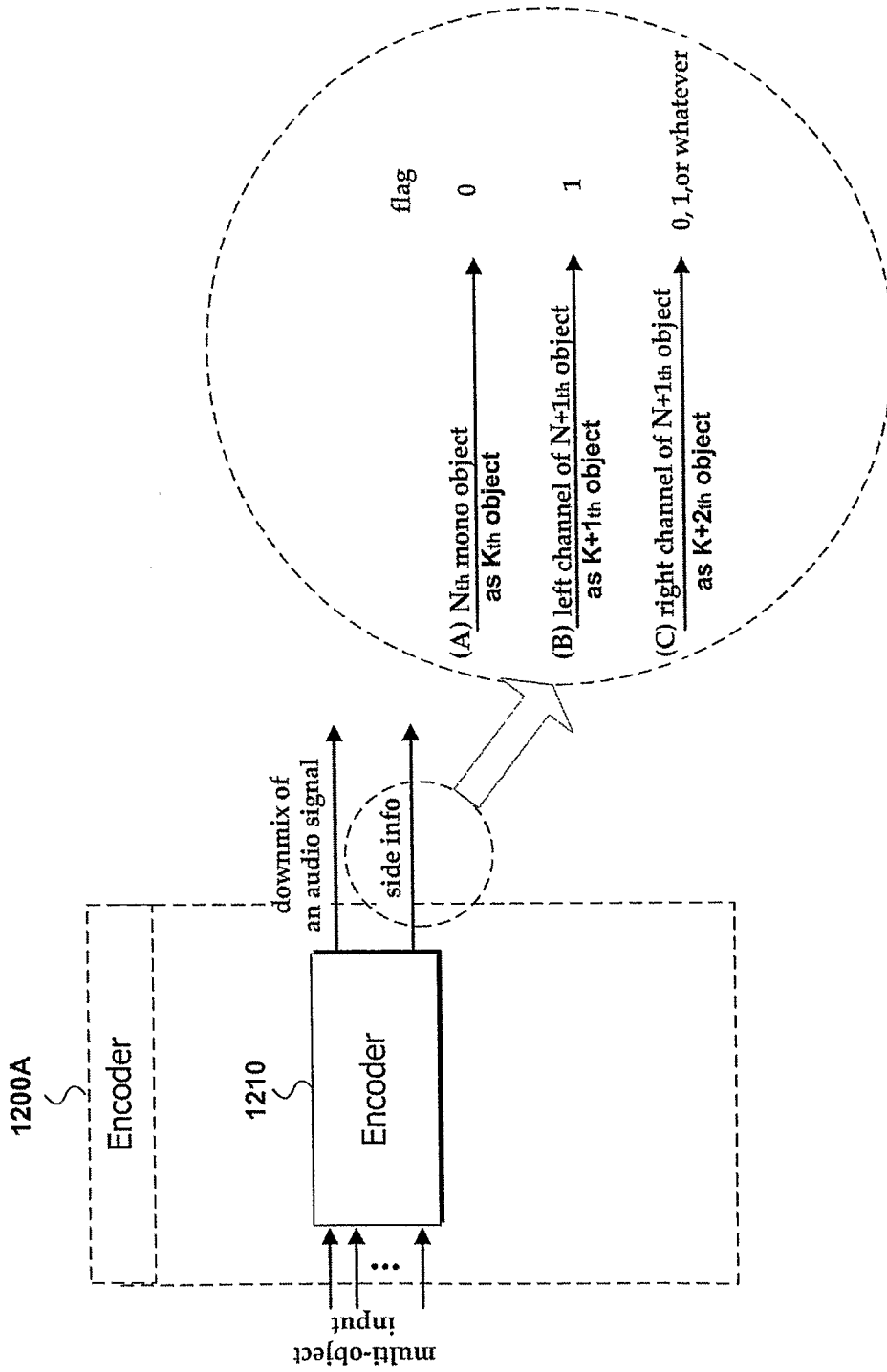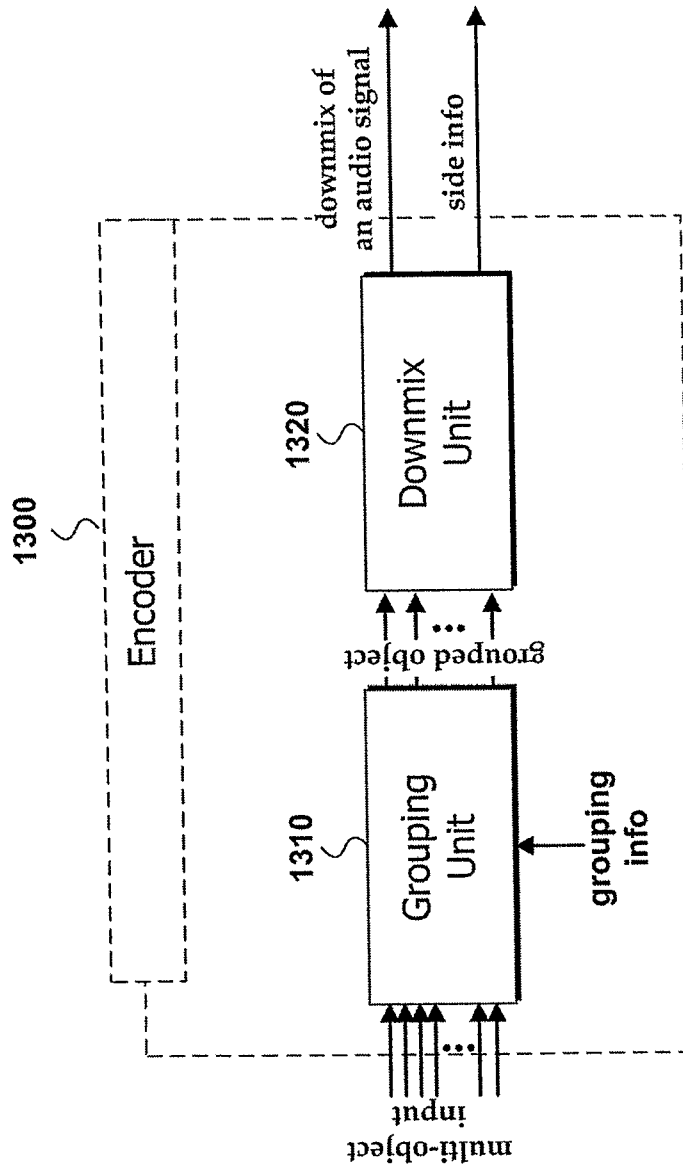| (α) Downmix	(β) Multichannel Parameter	(γ) Object Parameter

(a)

(α) Downmix	(β') Default Parameter	(γ) Object Parameter

(b)

FIG. 15

FIG. 16

FIG. 17

FIG. 18

FIG. 19

A. CLASSIFICATION OF SUBJECT MATTER

G10L 19/00(2006.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 8 H04S 3/02 H04S 5/00 H04R 5/00 H04B 1/00 H04B 1/20

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""DOWNMIX*""CHANNEL"

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2005/086139 A1 (DOLBY LABORATORIES LICENSING CO) 15 SEPTEMBER 2005 See Fig 1, Fig 2, Page 9 Line 1 ~ Page 16 Line 4	1, 13, 14
A	US 2006/0133618 A1 (VILEMOES et al.) 22 JUNE 2006 See Fig 3, Fig 4, Paragraph [0088] ~ [0124]	1, 13, 14
A	US 2006/0115100 A1 (FALLER) 01 JUNE 2006 See Fig 2, Abstract	14
A	JP 18-323408 A (JAPAN VICTOR CO) 30 NOVEMBER 2006 See Paragraph [0005] – [0007]	14

☐ Further documents are listed in the continuation of Box C.   ☒ See patent family annex.

* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of 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
17 MARCH 2008 (17.03.2008)	**17 MARCH 2008 (17.03.2008)**

Name and mailing address of the ISA/KR	Authorized officer
Korean Intellectual Property Office Government Complex-Daejeon, 139 Seonsa-ro, Seo-gu, Daejeon 302-701, Republic of Korea	SUH, Hawthorne
Facsimile No. 82-42-472-7140	Telephone No. 82-42-481-5670

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

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO2005086139A1	15.09.2005	AU2005219956A1	15.09.2005
		AU2005219956AA	15.09.2005
		BR200508343A	24.07.2007
		CA2556575AA	15.09.2005
		CA2556575A1	15.09.2005
		CN1926607A	07.03.2007
		EP01721312A1	15.11.2006
		EP1721312A1	15.11.2006
		JP19526522	13.09.2007
		JP2007526522T2	13.09.2007
		KR1020060132682	21.12.2006
		KR2006132682A	21.12.2006
		US20060212686A1	21.09.2006
		US20070140499A1	21.06.2007
		US20070148309A1	28.06.2007
		US2007140499A1	21.06.2007
		US2007140499AA	21.06.2007
		US2007148309A1	28.06.2007
		US2007148309AA	28.06.2007
US2006133618A1	22.06.2006	US2006133618AA	22.06.2006
US2006115100A1	01.06.2006	EP01817768A2	15.08.2007
		EP1817768A2	15.08.2007
		KR1020070094752	21.09.2007
		KR2007094752A	21.09.2007
		US2006115100AA	01.06.2006
		WO2006060280A2	08.06.2006
		WO2006060280C1	15.03.2007
		WO2006060280A3	14.09.2006
JP18323408A	30.11.2006	JP2006323408A2	30.11.2006