(12) STANDARD PATENT (11) Application No. **AU 2017200552 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
3D SOUND REPRODUCING METHOD AND APPARATUS

(51) International Patent Classification(s)
H04R 5/02 (2006.01) ***H04S 5/02*** (2006.01)

(21) Application No: **2017200552** (22) Date of Filing: **2017.01.27**

(43) Publication Date: **2017.02.23**
(43) Publication Journal Date: **2017.02.23**
(44) Accepted Journal Date: **2018.05.10**

(62) Divisional of:
2015207829

(71) Applicant(s)
Korea Advanced Institute of Science and Technology;Samsung Electronics Co., Ltd.

(72) Inventor(s)
Kim, Sun-Min;Park, Young-Jin;Jo, Hyun

(74) Agent / Attorney
Phillips Ormonde Fitzpatrick, PO Box 323, Collins Street West, VIC, 8007, AU

(56) Related Art
US 6839438 B1
EP 0959644 A2
US 2008/0159544 A1
US 2009/0067636 A1
US 6442277 B1
US 6766028 B1
YANG, D. T. ET AL: "High-Fidelity Multichannel Audio Coding", EURASIP Book Series on Signal Processing and Communications, 2006

2017200552 27 Jan 2017

Abstract

Provided are a three-dimensional (3D) sound reproducing method and apparatus. The method includes transmitting sound signals through a head related transfer filter (HRTF) corresponding to a first elevation, generating a plurality of sound signals by replicating the filtered sound signals, amplifying or attenuating each of the replicated sound signals based on a gain value corresponding to each of speakers, through which the replicated sound signals will be output, and outputting the amplified or attenuated sound signals through the corresponding speakers.

2017200552 27 Jan 2017

Description

Title of Invention: 3D SOUND REPRODUCING METHOD AND APPARATUS

[1] The present application is a divisional application from Australian Patent Application No. 2015207829 (which is a divisional application of Australian Patent Application No 2011274709), the entire disclosure of which is incorporated herein by reference.

Technical Field

[1a] Methods and apparatuses consistent with exemplary embodiments relate to re-producing three-dimensional (3D) sound, and more particularly, to localizing a virtual sound source to a predetermined elevation.

Background Art

[2] With developments in video and sound processing technologies, contents having high image and sound quality are being provided. Users demanding contents having high image and sound quality now require realistic images and sound, and accordingly, research into 3D image and sound is being actively conducted.

[3] 3D sound is generated by providing a plurality of speakers at different positions on a level surface and outputting sound signals that are equal to or different from each other according to the speakers so that a user may experience a spatial effect. However, sound may actually be generated from various elevations, as well as various points on the level surface. Therefore, a technology for effectively reproducing sound signals that are generated at different levels from each other is necessary.

[3a] A reference herein to a patent document or other matter which is given as prior art is not to be taken as an admission that that document or matter was known or that the information it contains was part of the common general knowledge as at the priority date of any of the claims.

Disclosure of Invention

Solution to Problem

[4] The present invention provides a 3D sound reproducing method and apparatus thereof for localizing a virtual sound source to a predetermined elevation.

Advantageous Effects of Invention

[5] According to the present embodiment, it is possible to provide 3D three-dimensional effect. And, according to the present embodiment, it is possible that the virtual sound source may be effectively localized to a predetermined elevation.

Brief Description of Drawings

2017200552 27 Jan 2017

[6] The above and other features and advantages of the present invention will become more apparent by describing in detail exemplary embodiments thereof with reference to the attached drawings in which:

[7] FIG. 1 is a block diagram of a 3D sound reproducing apparatus according to an exemplary embodiment;

[8] FIG. 2A is a block diagram of the 3D sound reproducing apparatus for localizing a virtual sound source to a predetermined elevation by using 5-channel signals;

[9] FIG. 2B is a block diagram of a 3D sound reproducing apparatus for localizing a virtual sound source to a predetermined elevation by using a sound signal according to another exemplary embodiment;

[10] FIG. 3 is a block diagram of a 3D sound reproducing apparatus for localizing a virtual sound source to a predetermined elevation by using a 5-channel signal according to another exemplary embodiment;

[11] FIG. 4 is a diagram showing an example of a 3D sound reproducing apparatus for localizing a virtual sound source to a predetermined elevation by outputting 7-channel signals through 7 speakers according to an exemplary embodiment;

[12] FIG. 5 is a diagram showing an example of a 3D sound reproducing apparatus for localizing a virtual sound source to a predetermined elevation by outputting 5-channel signals through 7 speakers according to an exemplary embodiment;

[13] FIG. 6 is a diagram showing an example of a 3D sound reproducing apparatus for localizing a virtual sound source to a predetermined elevation by outputting 7-channel signals through 5 speakers according to an exemplary embodiment;

[14] FIG. 7 is a diagram of a speaker system for localizing a virtual sound source to a predetermined elevation according to an exemplary embodiment; and

[15] FIG. 8 is a flowchart illustrating a 3D sound reproducing method according to an exemplary embodiment.

Best Mode for Carrying out the Invention

[16] Exemplary embodiments provide a method and apparatus for reproducing 3D sound, and in particular, a method and apparatus for localizing a virtual sound source to a predetermined elevation.

[17] According to a first aspect the present invention provides a method of rendering an audio signal, the method comprising: receiving multichannel audio signals and an input configuration; obtaining a first Head-Related Transfer Function (HRTF) based filter for a first height input channel signal among the multichannel audio signals, wherein the first height input channel signal is identified according to the input configuration; obtaining

2017200552 27 Jan 2017

first gains for the first height input channel signal; and performing elevation rendering on the multichannel signals including the first height input channel signal based on the first HRTF based filter and the first gains for elevating sound by a plurality of output channel signals comprising a horizontal configuration, wherein the input configuration comprises azimuth information and elevation information related to the multichannel audio signals.

[18] The predetermined filter may include head related transfer filter (HRTF).

[19] The transmitting the sound signals through the HRTF may include transmitting at least one of a left top channel signal representing a sound signal generated from a left side of a second elevation and a right top channel signal representing a sound signal generated from a right side of the second elevation through the HRTF.

[20] The method may further include generating the left top channel signal and the right top channel signal by np-mixing the sound signal, when the sound signal does not include the left top channel signal and the right top channel signal.

[21] The transmitting the sound signal through the HRTF a include transmitting at least one of a front left channel sign representing a sound signal generated from a front left side and a front right channel signal representing a sound signal generated from a front right side through the HRTF, when the sound signal does not include a left top channel signal representing a sound signal generated from a left side of a second elvation and a right top channel signal representing sound signal generated from a right side of the second elevation.

[22] The HRTF may be generated by dividing a first HRTF including information about a path from the first elevation to ears of a user by a second HRTF include information about a path from a location of a speaker, through which the sound signal will be output, to the ears of the user.

[23] The outputting the sound signal may include: generating a fist sound signal by mixing the sound signal that is obtained by amplifying the filtered left top channel signal according to a first gain value with the sound signal that is obtained by amplifying the filtered right top channel signal according a second gain value; generating a second sound according the sound signal that is obtained by amplifying the left top channel signal according to the second gain value with the sound signal that is obtained by amplifying the filtered right top channel signal according to the first gain value; and outputting the first sound signal through a speaker disposed on a left side and outputting the second sound signal though a speaker disposed on a right side.

[24] The outputting the sound signals may include: generating a third sound signal by mixing a sound signal that is obtained by amplifying a rear left signal representing a sound signal

2017200552 27 Jan 2017

generated from a rear left side a cording to a third gain value with the first sound signal; generating a fourth sound signal by mixing a sound signal that is obtained by amplifying a rear right signal representing a sound signal generated from a rear right side according to the third gain value with the second sound signal; and outputting the third sound signal through a left rear speaker and the fourth sound signal through a right rear speaker.

[25] The outputting the sound signals may further include muting at least one of the first sound signal and the second sound signal according to a location in the first elevation, where the virtual sound source is to be localized.

[26] The transmitting the sound signal though the HRTF may include: obtaining information about the location where the virtual sound source is to be localized; and determining the HRTF though which the sound signal is transmitted, based on the location information.

[27] The performing at least one of the amplifying, attenuating, and delaying processes may include determining at least one of the gain values and the delay values that will be applied to each of the replicated sound signals based on at least one of a location of the actual speaker, a location of a listener, and a location of the virtual sound source.

[28] The determining at least one of the gain value and the delay value may include determining at least one of the gain value and the delay value with respect to each of the replicated sound signals as a determined value, when information about the location of the listener is not obtained.

[29] The determining at least one of the gain value and the delay value may include determining at least one of the gain value and the delay value with respect to each of the replicated sound signals as an equal value, when information about the location of the listener is not obtained.

[30] According to a second aspect, the present invention provides a rendering an audio signal apparatus comprising: a receiver configured to receive multichannel audio signals and an input configuration; a controller configured to obtain a first Head-Related Transfer Function (HRTF) based filter for a first height input channel signal among the multichannel audio signals, wherein the first height input channel signal is identified according to the input configuration and configured to obtain first gains for the first height input channel signal; and a renderer configured to perform elevation rendering on the multichannel signals including the first height input channel signal based on the first HRTF based filter and the first gains for elevating sound by a plurality of output channel signals comprising a horizontal configuration, wherein the input configuration comprises azimuth information and elevation information related to the multichannel audio signals.

[31] The predetermined filter is head related transfer filter (HRTF).

2017200552 27 Jan 2017

[32] The filter unit may transmit at least one of a left top channel signal representing a sound signal generated from a left side of a second elevation and a right top channel signal representing a sound signal generated from a right side of the second elevation through the HRTF.

[33] The 3D sound reproducing apparatus may further comprising: an up-mixing unit which generates a left top channel signal and a right top channel signal, when the sound signal does not include the left top channel signal and the right top channel signal.

[34] The filter unit may transmit at least one of a front left channel signal representing a sound signal generated from a front left side and a front right channel signal rep-resenting a sound signal generated from a front right side through the HRTF, when the sound signal does not include a left top channel signal representing the sound signal generated from a left side of a second elevation and a right top channel signal rep-resenting the sound signal generated from a right side of the second elevation.

[35] The HRTF is generated by dividing a first HRTF including information about a path

2017200552 27 Jan 2017

from the first elevation to ears of a user by a second HRTF including information about a path from a location of a speaker, through which the sound signal will be output, to the ears of the user.

[36] The output unit comprises: a first mixing unit which generates a first sound signal by mixing a sound signal that is obtained by amplifying the filtered left top channel signal according to a first gain value with a sound signal that is obtained by amplifying the filtered right top channel signal according to a second gain value;

[37] a second mixing which generates a second sound signal by mixing a sound signal that is obtained by amplifying the filtered left top channel signal according to the second gain value with a sound signal that is obtained by amplifying the filtered right top channel signal according to the first gain value; and

[38] a rendering unit which outputs the first sound signal through a speaker disposed on a left side and outputting the second sound signal through a speaker disposed on a right side.

[39] The output unit comprises:

[40] a third mixing unit which generates a third sound signal by mixing a sound signal that is obtained by amplifying a rear left signal representing a sound signal generated from a rear left side according to a third gain value with the first sound signal; and

[41] a fourth mixing unit which generates a fourth sound signal by mixing a sound signal that is obtained by amplifying a rear right signal representing a sound signal generated from a rear right side according to the third gain value with the second sound signal;

[42] wherein the rendering unit outputs the third sound signal through a left rear speaker and the fourth sound signal through a right rear speaker.

[43] The rendering unit comprises a controller which mutes at least one of the first and second sound signals according to a location on the first elevation, where the virtual sound source is to be localized.

Mode for the Invention

[44] This application claims the benefit of U.S Provisional Application No. 61/362,014, filed on July 7, 2010 in the United States Patent and Trademark Office, Korean Patent Application No. 10-2010-0137232, filed on December 28, 2010, and Korean Patent Application No. 10-2011-0034415, filed on April 13, 2011, in the Korean Intellectual Property Office, the disclosures of which are incorporated herein in their entirety by reference.

[45] Hereinafter, exemplary embodiments will be described in detail with reference to accompanying drawings. In this description, the "term" unit means a hardware component and/or a software component that is executed by a hardware component such as a processor.

2017200552 27 Jan 2017

[46] FIG. 1 is a block diagram of a 3D sound reproducing apparatus 100 according to an exemplary embodiment.

[47] The 3D sound reproducing apparatus 100 includes a filter unit 110, a replication unit 120, an amplifier 130, and an output unit 140.

[48] The filter unit 110 transmits a sound signal through a predetermined filter generating 3D sound corresponding to a predetermined elevation. The filter unit 110 may transmit a sound signal through a head related transfer filter (HRTF) corresponding to a predetermined elevation. The HRTF includes information about a path from a spatial position of a sound source to both ears of a user, that is, a frequency transmission characteristic. The HRTF makes a user recognize 3D sound by a phenomenon whereby complex passage characteristics such as diffraction at skin of human head and reflection by pinnae, as well as simple passage differences such as an inter-aural level difference (ILD) and an inter-aural time difference (ITD), are changed according to sound arrival directions. Since only one HRTF exists in each direction in a space, the 3D sound may be generated due to the above characteristics.

[49] The filter unit 110 uses the HRTF filter for modeling a sound being generated from a position at an elevation higher than that of actual speakers that are arranged on a level surface. Equation 1 below is an example of HRTF used in the filter unit 110.

[50] $HRTF=HRTF_2/HRTF_1$ ---------- (1)

[51] $HRTF_2$ is HRTF representing passage information from a position of a virtual sound source to the ears of a user, and $HRTF_1$ is HRTF representing passage information from a position of an actual speaker to the ears of the user. Since a sound signal is output from the actual speaker, in order for the user to recognize that the sound signal is output from a virtual speaker, $HRTF_2$ corresponding to a predetermined elevation is divided by $HRTF_1$ corresponding to the level surface (or elevation of the actual speaker).

[52] An optimal HRTF corresponding to a predetermined elevation varies depending on each person, such as a fingerprint. However, it is impossible to calculate the HRTF for each user and to apply the calculated HRTF to each user. Thus, HRTF is calculated for some users of a user group, who have similar properties (for example, physical properties such as age and height, or propensities such as favorite frequency band and favorite music), and then, a representative value (for example, an average value) may be determined as the HRTF applied to all of the users included in the corresponding user group.

[53] Equation 2 below is a result of filtering the sound signal by using the HRTF defined in Equation 1 above.

[54] $Y_2(f)=Y_1(f)*HRTF$ ---------- (2)

[55] $Y_1(f)$ is a value converted into a frequency band from the sound signal output that a

2017200552 27 Jan 2017

user hears from the actual speaker, and $Y_2(f)$ is a value converted into a frequency band from the sound signal output that a user hear from the virtual speaker.

[56] The filter unit 110 may only filter some channel signals of a plurality of channel signals included in the sound signal.

[57] The sound signal may include sound signals corresponding to a plurality of channels. Hereinafter, a 7-channel signal is defined for convenience of description. However, the 7-channel signal is an example, and the sound signal may include a channel signal representing the sound signal generated from directions other than the seven directions that will now be described.

[58] A center channel signal is a sound signal generated from a front center portion, and is output through a center speaker.

[59] A front right channel signal is a sound signal generated from a right side of a front portion, and is output through a front right speaker.

[60] A front left channel signal is a sound signal generated from a left side of the front portion, and is output through a front left speaker.

[61] A rear right channel signal is a sound signal generated from a right side of a rear portion, and is output through a rear right speaker.

[62] A rear left channel signal is a sound signal generated from a left side of the rear portion, and is output through a rear left speaker.

[63] A right top channel signal is a sound signal generated from an upper right portion, and is output through a right top speaker.

[64] A left top channel signal is a sound signal generated from an upper left portion, and is output through a left top speaker.

[65] When the sound signal includes the right top channel signal and the left top channel signal, the filter unit 110 filters the right top channel signal and the left top channel signal. The right top signal and the left top signal that are filtered are then used to model a virtual sound source that is generated from a desired elevation.

[66] When the sound signal does not include the right top signal and the left top signal, the filter unit 110 filters the front right channel signal and the front left channel signal. The front right channel signal and the front left channel signal are then used to model the virtual sound source generated from a desired elevation.

[67] In some exemplary embodiments, the sound signal that does not include the right top channel signal and the left top channel signal (for example, 2.1 channel or 5.1 channel signal) is up-mixed to generate the right top channel signal and the left top channel signal. Then, the mixed right top channel signal and the left top channel signal may be filtered.

[68] The replication unit 120 replicates the filtered channel signal into a plurality of signals. The replication unit 120 replicates the filtered channel signal as many times as

2017200552 27 Jan 2017

the number of speakers through which the filtered channel signals will be output. For example, when the filtered sound signal is output as the right top channel signal, the left top channel signal, the rear right channel signal, and the rear left channel signal, the replication unit 120 makes four replicas of the filtered channel signal. The number of replicas made by the replication unit 120 may vary depending on the exemplary embodiments; however, it is desirable that two or more replicas are generated so that the filtered channel signal may be output at least as the rear right channel signal and the rear left channel signal.

[69] The speakers through which the right top channel signal and the left top channel signal will be reproduced are disposed on the level surface. As an example, the speakers may be attached right above the front speaker that reproduces the front right channel signal.

[70] The amplifier 130 amplifies (or attenuates) the filtered sound signal according to a predetermined gain value. The gain value may vary depending on the kind of the filtered sound signal.

[71] For example, the right top channel signal output through the right top speaker is amplified according to a first gain value, and the right top channel signal output through the left top speaker is amplified according to a second gain value. Here, the first gain value may be greater than the second gain value. In addition, the left top channel signal output through the right top speaker is amplified according to the second gain value and the left top channel signal output through the left top speaker is amplified according to the first gain value so that the channel signals corresponding to the left and right speakers may be output.

[72] In the related art, an ITD method has been mainly used in order to generate a virtual sound source at a desired position. The ITD method is a method of localizing the virtual sound source to a desired position by outputting the same sound signal from a plurality of speakers with time differences. The ITD method is suitable for localizing the virtual sound source at the same plane on which the actual speakers are located. However, the ITD method is not an appropriate way to localize the virtual sound source to a position that is located higher than an elevation of the actual speaker.

[73] In exemplary embodiments, the same sound signal is output from a plurality of speakers with different gain values. In this manner, according to an exemplary embodiment, the virtual sound source may be easily localized to an elevation that is higher than that of the actual speaker, or to a certain elevation regardless of the elevation of the actual speaker.

[74] The output unit 140 outputs one or more amplified channel signals through corresponding speakers. The output unit 140 may include a mixer (not shown) and a rendering unit (not shown).

2017200552 27 Jan 2017

[75] The mixer mixes one or more channel signals.

[76] The mixer mixes the left top channel signal that is amplified according to the first gain value with the right top channel signal that is amplified according to the second gain value to generate a first sound component, and mixes the left top channel signal that is amplified according to the second gain value and the right top channel signal that is amplified according to the first gain value to generate a second sound component.

[77] In addition, the mixer mixes the rear left channel signal that is amplified according to a third gain value with the first sound component to generate a third sound component, and mixes the rear right channel signal that is amplified according to the third gain value with the second sound component to generate a fourth sound component.

[78] The rendering unit renders the mixed or un-mixed sound components and outputs them to corresponding speakers.

[79] The rendering unit outputs the first sound component to the left top speaker, and outputs the second sound component to the right top speaker. If there is no left top speaker or no right top speaker, the rendering unit may output the first sound component to the front left speaker and may output the second sound component to the front right speaker.

[80] In addition, the rendering unit outputs the third sound component to the rear left speaker, and outputs the fourth sound component to the rear right speaker.

[81] Operations of the replication unit 120, the amplifier 130, and the output unit 140 may vary depending on the number of channel signals included in the sound signal and the number of speakers. Examples of operations of the 3D sound reproducing apparatus according to the number of channel signals and speakers will be described later with reference to FIGS. 4 through 6.

[82] FIG. 2A is a block diagram of a 3D sound reproducing apparatus 100 for localizing a virtual sound source to a predetermined elevation by using 5-channel signals according to an exemplary embodiment.

[83] An up-mixer 210 up-mixes 5-channel signals 201 to generate 7-channel signals including a left top channel signal 202 and a right top channel signal 203.

[84] The left top channel signal 202 is input into a first HRTF 111, and the right top channel signal 203 is input into a second HRTF 112.

[85] The first HRTF 111 includes information about a passage from a left virtual sound source to the ears of the user, and the second HRTF 112 includes information about a passage from a right virtual sound source to the ears of the user. The first HRTF 111 and the second HRTF 112 are filters for modeling the virtual sound sources at a predetermined elevation that is higher than that of actual speakers.

[86] The left top channel signal and the right top channel signal passing through the first

2017200552 27 Jan 2017

HRTF 111 and the second HRTF 112 are input into replication units 121 and 122.

[87] Each of the replication units 121 and 122 makes two replicas of each of the left top channel signal and the right top channel signal that are transmitted through the HRTFs 111 and 112. The replicated left top channel signal and right top channel signal are transferred to first to third amplifiers 131, 132, and 133.

[88] The first amplifier 131 and the second amplifier 132 amplify the replicated left top signal and right top signal according to the speaker outputting the signal and the kind of the channel signals. In addition, the third amplifier 133 amplifies at least one channel signal included in the 5-channel signals 201.

[89] In some exemplary embodiments, the 3D sound reproducing apparatus 100 may include a first delay unit (not shown) and a second delay unit (not shown) instead of the first and second amplifiers 131 and 132, or may include all of the first and second amplifiers 131 and 132, and the first and second delay units. This is because a same result as that of varying the gain value may be obtained when delayed values of the filtered sound signals vary depending on the speakers.

[90] The output unit 140 mixes the amplified left top channel signal, the right top channel signal, and the 5-channel signal 201 to output the mixed signals as 7-channel signals 205. The 7-channel signals 205 are output to each of the speakers.

[91] In another exemplary embodiment, when 7-channel signals are input, the up-mixer 210 may be omitted.

[92] In another exemplary embodiment, the 3D sound reproducing apparatus 100 may include a filter determining unit (not shown) and an amplification/delay coefficient determining unit (not shown).

[93] The filter determining unit selects an appropriate HRTF according to a position where the virtual sound source will be localized (that is, an elevation angle and a horizontal angle). The filter determining unit may select an HRTF corresponding to the virtual sound source by using mapping information between the location of the virtual sound source and the HRTF. The location information of the virtual sound source may be received through other modules such as applications (software or hardware), or may be input from the user. For example, in a game application, a location where the virtual sound source is localized may vary depending on time, and the filter determining unit may change the HRTF according to the variation of the virtual sound source location.

[94] The amplification/delay coefficient determining unit may determine at least one of an amplification (or attenuation) coefficient and a delay coefficient of the replicated sound signal based on at least one of a location of the actual speaker, a location of the virtual sound source, and a location of a listener. If the amplification/delay coefficient determining unit does not recognize the location information of the listener in advance, the amplification/delay coefficient determining unit may select at least one of a prede-

2017200552 27 Jan 2017

termined amplification coefficient and a delay coefficient.

[95] FIG. 2B is a block diagram of a 3D sound reproducing apparatus 100 for localizing a virtual sound source to a predetermined elevation by using a sound signal according to another exemplary embodiment.

[96] In FIG. 2B, a first channel signal that is included in a sound signal will be described for convenience of description. However, the present exemplary embodiment may be applied to other channels signals included in the sound signal.

[97] The 3D sound reproducing apparatus 100 may include a first HRTF 211, a replication unit 221, and an amplification/delay unit 231.

[98] A first HRTF 211 is selected based on the location information of the virtual sound source, and the first channel signal is transmitted through the first HRTF 211. The location information of the virtual sound source may include elevation angle information and horizontal angle information.

[99] The replication unit 221 replicates the first channel signal after being filtered into one or more sound signals. In FIG. 2B, it is assumed that the replication unit 221 replicates the first channel signal as many times as the number of actual speakers.

[100] The amplification/delay unit 231 determines amplification/delay coefficients of the replicated first channel signals respectively corresponding to the speakers, based on at least one of location information of the actual speaker, location information of a listener, and location information of the virtual sound source. The amplification/delay unit 231 amplifies/attenuates the replicated first channel signals based on the determined amplification (or attenuation) coefficients, or delays the replicated first channel signal based on the delay coefficient. In an exemplary embodiment, the amplification/delay unit 231 may simultaneously perform the amplification (or attenuation) and the delay of the replicated first channel signals based on the determined amplification (or attenuation) coefficients and the delay coefficients.

[101] The amplification/delay unit 231 generally determines the amplification/delay coefficient of the replicated first channel signal for each of the speakers; however, the amplification/delay unit 231may determine the amplification/delay coefficients of the speakers to be equal to each other when the location information of the listener is not obtained, and thus, the first channel signals that are equal to each other may be output respectively through the speakers. In particular, when the amplification/delay unit 231 does not obtain the location information of the listener, the amplification/delay unit 231 may determine the amplification/delay coefficient for each of the speakers as a predetermined value (or an arbitrary value).

[102] FIG. 3 is a block diagram of a 3D sound reproducing apparatus 100 for localizing a virtual sound source to a predetermined elevation by using 5-channel signals according to another exemplary embodiment. A signal distribution unit 310 extracts a front right

2017200552 27 Jan 2017

channel signal 302 and a front left channel signal 303 from the 5-channel signal, and transfers the extracted signals to the first HRTF 111 and the second HRTF 112.

[103] The 3D sound reproducing apparatus 100 of the present exemplary embodiment is the same as that described with reference to FIG. 2 except that the sound components applied to the filtering units 111 and 112, the replication units 121 and 122, and the amplifiers 131, 132, and 133 are the front right channel signal 302 and the front left channel signal 303. Therefore, detailed descriptions of the 3D sound reproducing apparatus 100 of the present exemplary embodiment will not be provided here.

[104] FIG. 4 is a diagram showing an example of a 3D sound reproducing apparatus 100 for localizing a virtual sound source to a predetermined elevation by outputting 7-channel signals through 7 speakers according to another exemplary embodiment.

[105] FIG. 4 will be described based on input sound signals, and then, described based on sound signals output through speakers.

[106] Sound signals including a front left channel signal, a left top channel signal, a rear left channel signal, a center channel signal, a rear right channel signal, a right top channel signal, and a front right channel signal are input in the 3D sound reproducing apparatus 100.

[107] The front left channel signal is mixed with the center channel signal that is attenuated by a factor B, and then, is transferred to a front left speaker.

[108] The left top channel signal passes through an HRTF corresponding to an elevation that is 30° higher than that of the left top speaker, and is replicated into four channel signals.

[109] Two left top channel signals are amplified by a factor A, and then, mixed with the right top channel signal. In some exemplary embodiments, after mixing the left top channel signal that is amplified by the factor A with the right top channel signal, the mixed signal may be replicated into two signals. One of the mixed signals is amplified by a factor D, and then, mixed with the rear left channel signal and output through the rear left speaker. The other of the mixed signals is amplified by a factor E, and then, output through the left top speaker.

[110] Two remaining left top channel signals are mixed with the right top channel signal that is amplified by the factor A. One of the mixed signals is amplified by the factor D, and then, is mixed with the rear right channel signal and output through the rear right speaker. The other of the mixed signals is amplified by the factor E, and is output through the right top speaker.

[111] The rear left channel signal is mixed with the right top channel signal that is amplified by the factor D and the left top channel signal that is amplified by a factor D/A, and is output through the rear left speaker.

[112] The center channel signal is replicated into three signals. One of the replicated center

2017200552 27 Jan 2017

channel signals is attenuated by the factor B, and then, is mixed with the front left channel signal and output through the front left speaker. Another replicated center channel signal is attenuated by the factor B, and after that, is mixed with the front right channel signal and output through the front right speaker. The other of the replicated center channel signals is attenuated by a factor C, and then, is output through the center speaker.

[113] The rear right channel signal is mixed with the left top channel signal that is amplified by the factor D and the right top channel signal that is amplified by the factor D(A, and then, is output through the rear right speaker.

[114] The right top signal passes through an HRTF corresponding to an elevation that is 30[higher than that of the right top speaker, and then, is replicated into four signals.

[115] Two right top channel signals are mixed with the left top channel signal that is amplified by the factor A. One of the mixed signals is amplified by the factor D, and is mixed with the rear left channel signal and output through the rear left speaker. The other of the mixed signals is amplified by the factor E, and is output through the left top speaker.

[116] Two replicated right top channel signals are amplified by the factor A, and are mixed with the left top channel signals. One of the mixed signals is amplified by the factor D, and is mixed with the rear right channel signal and output through the rear right speaker. The other of the mixed signals is amplified by the factor E, and is output through the right top speaker.

[117] The front right channel signal is mixed with the center channel signal that is attenuated by the factor B, and is output through the front right speaker.

[118] Next, sound signals that are finally output through the speakers after the above-described processes are as follows:

[119] (front left channel signal + center channel signal (B) is output through the front left speaker;

[120] (rear left channel signal + D((left top channel signal(A + right top channel signal)) is output through the rear left speaker;

[121] (E((left top channel signal(A + right top channel signal)) is output through the left top speaker;

[122] (C(center channel signal) is output through the center speaker;

[123] (E((right top channel signal(A + left top channel signal)) is output through the right top speaker;

[124] (rear right channel signal + D((right top channel signal(A + left top channel signal)) is output through the rear right speaker; and

[125] (front right channel signal + center channel signal(B) is output through the front right speaker.

2017200552 27 Jan 2017

[126] In FIG. 4, the gain values to amplify or attenuate the channel signals are merely examples, and various gain values that may make the left speaker and the right speaker output corresponding channel signals may be used. In addition, in some exemplary embodiments, gain values for outputting the channel signals that do not correspond to the speakers through the left and right speakers may be used.

[127] FIG. 5 is a diagram showing an example of a 3D sound reproducing apparatus 100 for localizing a virtual sound source to a predetermined elevation by outputting 5-channel signals through 7 speakers according to another exemplary embodiment.

[128] The 3D sound reproducing apparatus shown in FIG. 5 is the same as that shown in FIG. 4 except that sound components input into an HRTF are a front left channel signal and a front right channel signal. Therefore, sound signals output through the speakers are as follows:

[129] (front left channel signal + center channel signal (B) is output through the front left speaker;

[130] (rear left channel signal + D((front left channel signal(A + front right channel signal)) is output through the rear left speaker;

[131] (E((front left channel signal(A + front right channel signal)) is output through the left top speaker;

[132] (C(center channel signal) is output through the center speaker;

[133] (E((front right channel signal(A + front left channel signal)) is output through the right top speaker;

[134] (rear right channel signal + D((front right channel signal(A + front left channel signal)) is output through the rear right speaker; and

[135] (front right channel signal + center channel signal(B) is output through the front right speaker.

[136] FIG. 6 is a diagram showing an example of a 3D sound reproducing apparatus 100 for localizing a virtual sound source to a predetermined elevation by outputting 7-channel signals through 5 speakers, according to another exemplary embodiment.

[137] The 3D sound reproducing apparatus 100 of FIG. 6 is the same as that shown in FIG. 4 except for that the output signals that are supposed to output through the left top speaker (the speaker for the left top channel signal 413) and the right top speaker (the speaker for the right top channel signal 415) in Fig. 4, are output through the front left speaker (the speaker for the front left channel signal 611) and the front right speaker (the speaker for the front right channel signal 615) respectively. Therefore, sound signals output through the speakers are as follows:

[138] (front left channel signal+(center channel signal (B) + E((front left channel signal(A + front right signal)) is output through the front left speaker;

[139] (rear left channel signal + D((front left channel signal(A + front right channel

2017200552 27 Jan 2017

signal)) is output through the rear left speaker;

[140] (C(center channel signal) is output through the center speaker;

[141] (E((front right channel signal(A + front left channel signal)) is output through the right top speaker;

[142] (rear right channel signal + D((front right channel signal(A + front left channel signal)) is output through the rear right speaker; and

[143] (front right channel signal + (center channel signal(B) + E((front right channel signal(A + front left channel signal)) is output through the front right speaker.

[144] FIG. 7 is a diagram of a speaker system for localizing a virtual sound source to a predetermined elevation according to an exemplary embodiment.

[145] The speaker system of FIG. 7 includes a center speaker 710, a front left speaker 721, a front right speaker 722, a rear left speaker 731, and a rear right speaker 732.

[146] As described above with reference to FIGS. 4 through 6, for localizing a virtual sound source to a predetermined elevation, a left top channel signal and a right top channel signal that have passed through a filter are amplified or attenuated by gain values that are different according to the speakers, and then, are input into the front left speaker 721, the front right speaker 722, the rear left speaker 731, and the rear right speaker 732.

[147] Although not shown in FIG. 7, a left top speaker (not shown) and a right top speaker (not shown) may be disposed above the front left speaker 721 and the front right speaker 722. In this case, the left top channel signal and the right top channel signal passing through the filter are amplified by the gain values that are different according to the speakers and input into the left top speaker (not shown), the right top speaker (not shown), the rear left speaker 731, and the rear right speaker 732.

[148] A user recognizes that the virtual sound source is localized to a predetermined elevation when the left top channel signal and the right top channel signal that are filtered are output through one or more speakers in the speaker system. Here, when the filtered left top channel signal or the right top channel signal is muted in one or more speakers, a location of the virtual sound source in a left-and-right direction may be adjusted.

[149] When the virtual sound source is to be located at a center portion in a predetermined elevation, all of the front left speaker 721, the front right speaker 722, the rear left speaker 731, and the rear right speaker 732 output the filtered left top and right top channel signals, or only the rear left speaker 731 and the rear right speaker 732 may output the filtered left top and right top channel signals. In some exemplary embodiments, at least one of the filtered left top and right top channel signals may be output through the center speaker 710. However, the center speaker 710 does not contribute to the adjustment of the location of the virtual sound source in the left-

2017200552 27 Jan 2017

and-right direction.

[150] When it is desired that the virtual sound source be located at a right side in a predetermined elevation, the front right speaker 722, the rear left speaker 731, and the rear right speaker 732 may output the filtered left top and right top channel signals.

[151] When it is desired that the virtual sound source be located at a left side in a predetermined elevation, the front left speaker 721, the rear left speaker 731, and the rear right speaker 732 may output the filtered left top and right top channel signals.

[152] Even when it is desired that the virtual sound source be located at the right or left side in the predetermined elevation, the filtered left top and right top channel signals output through the rear left speaker 731 and the rear right speaker 732 may not be muted.

[153] In some exemplary embodiments, the location of the virtual sound source in the left-and-right direction may be adjusted by adjusting the gain value for amplifying or attenuating the left top and right top channel signals, without muting the filtered left and right top channel signals output through one or more speakers.

[154] FIG. 8 is a flowchart illustrating a 3D sound reproducing method according to an exemplary embodiment.

[155] In operation S810, a sound signal is transmitted through an HRTF corresponding to a predetermined elevation.

[156] In operation S820, the filtered sound signal is replicated to generate one or more replica sound signals.

[157] In operation S830, each of the one or more replica sound signals is amplified according to a gain value corresponding to a speaker, through which the sound signal will be output.

[158] In operation S840, the one or more amplified sound signals are output respectively through corresponding speakers.

[159] In the related art, a top speaker is installed at a desired elevation in order to output a sound signal being generated at the elevation; however, it is not easy to install the top speaker on the ceiling. Thus, the top speaker is generally placed above the front speaker, which may cause a desired elevation to not be reproduced.

[160] When the virtual sound source is localized to a desired location by using an HRTF, the localization of the virtual sound source may be performed effectively in the left-and-right direction on a horizontal plane. However, the localization using the HTRF is not suitable for localizing the virtual sound source to an elevation that is higher or lower than that of the actual speakers.

[161] In contrast, according to the exemplary embodiments, one or more channel signals passing through the HRTF are amplified by gain values that are different from each other according to the speakers, and are output through the speakers. In this manner,

2017200552 27 Jan 2017

the virtual sound source may be effectively localized to a predetermined elevation by using the speakers disposed on the horizontal plane.

[162] The exemplary embodiments can be written as computer programs and can be implemented in general-use digital computers that execute the programs which are stored in a computer readable recording medium.

[163] Examples of the computer readable recording medium include magnetic storage media (e.g., ROM, floppy disks, hard disks, etc.), and optical recording media (e.g., CD-ROMs, or DVDs).

[164] While exemplary embodiments been particularly shown and described, it will be understood by those of ordinary skill in the art that various changes in form and details may be made therein without departing from the spirit and scope of the inventive concept as defined by the following claims.

[165] Where the terms “comprise”, “comprises”, “comprised” or “comprising” are used in this specification (including the claims) they are to be interpreted as specifying the presence of the stated features, integers, steps or components, but not precluding the presence of one or more other features, integers, steps or components, or group thereto.

2017200552 27 Jan 2017

Claims

1. A method of rendering an audio signal, the method comprising:

receiving multichannel audio signals and an input configuration;

obtaining a first Head-Related Transfer Function (HRTF) based filter for a first height input channel signal among the multichannel audio signals, wherein the first height input channel signal is identified according to the input configuration;

obtaining first gains for the first height input channel signal; and

performing elevation rendering on the multichannel signals including the first height input channel signal based on the first HRTF based filter and the first gains for elevating sound by a plurality of output channel signals comprising a horizontal configuration,

wherein the input configuration comprises azimuth information and elevation information related to the multichannel audio signals.

2. The method of claim 1, wherein the horizontal configuration comprises 5.1 channel configuration or 5.0 channel configuration.

3. The method of claim 2, wherein the first HRTF based filter is used for each of the plurality of output channel signals.

4. The method of claim 1, wherein gains of surround output channel signals among the first gains are non-zero positive values.

5. The method of claim 1, wherein the first HRTF based filter is obtained based on each location of virtual output.

6. The method of claim 1, wherein the first gains are obtained based on each location of loudspeakers.

7. The method of claim 1, wherein the first height input channel signal is distributed to at least one of the plurality of output channel signals.

8. The method of claim 1, the method further comprising:

2017200552 27 Jan 2017

obtaining a second HRTF based filter for a second height input channel signal among the multichannel signals, wherein the second height input channel signal is identified by the input configuration; and

obtaining second gains for the second height input channel signal,

wherein the first HRTF based filter and the second HRTF based filter are independently obtained,

wherein the first gains and the second gains are independently obtained,

wherein the elevation rendering on the multichannel signals further including the second height input channel signal is performed further based on the second HRTF based filter and the second gains.

9. The method of claim 1, wherein a surround output channel signal is identified by at least one of 110 degree azimuth and -110 degree azimuth.

10. The method of claim 1, wherein a surround output channel signal is identified by 0 degree elevation.

11. The method of claim 1, wherein the first height input channel signal is located at top center.

12. A rendering an audio signal apparatus comprising:

a receiver configured to receive multichannel audio signals and an input configuration;

a controller configured to obtain a first Head-Related Transfer Function (HRTF) based filter for a first height input channel signal among the multichannel audio signals, wherein the first height input channel signal is identified according to the input configuration and configured to obtain first gains for the first height input channel signal; and

a renderer configured to perform elevation rendering on the multichannel signals including the first height input channel signal based on the first HRTF based filter and the first gains for elevating sound by a plurality of output channel signals comprising a horizontal configuration,

wherein the input configuration comprises azimuth information and elevation information related to the multichannel audio signals.

13. The apparatus of claim 12, wherein the horizontal configuration comprises 5.1 channel configuration or 5.0 channel configuration.

2017200552 27 Jan 2017

14. The apparatus of claim 13, wherein the first HRTF based filter is used for each of the plurality of output channel signals.

15. The apparatus of claim 12, wherein gains of surround output channel signals among the first gains are non-zero positive values.

16. The apparatus of claim 12, wherein the first HRTF based filter is obtained based on each location of virtual output.

17. The apparatus of claim 12, wherein the first gains are obtained based on each location of loudspeakers.

18. The apparatus of claim 12, wherein the first height input channel signal is distributed to at least one of the plurality of output channel signals.

19. The apparatus of claim 12,

the controller further obtain a second HRTF based filter for a second height input channel signal among the multichannel signals, wherein the second height input channel signal is identified according to the input configuration and obtaining second gains for the second height input channel signal,

wherein the first HRTF based filter and the second HRTF based filter are independently obtained,

wherein the first gains and the second gains are independently obtained,

wherein the elevation rendering on the multichannel signals further including the second height input channel signal is performed further based on the second HRTF based filter and the second gains.

20. The apparatus of claim 12, wherein a surround output channel signal is identified by at least one of 110 degree azimuth and -110 degree azimuth.

21. The apparatus of claim 12, wherein a surround output channel signal is identified by 0 degree elevation.

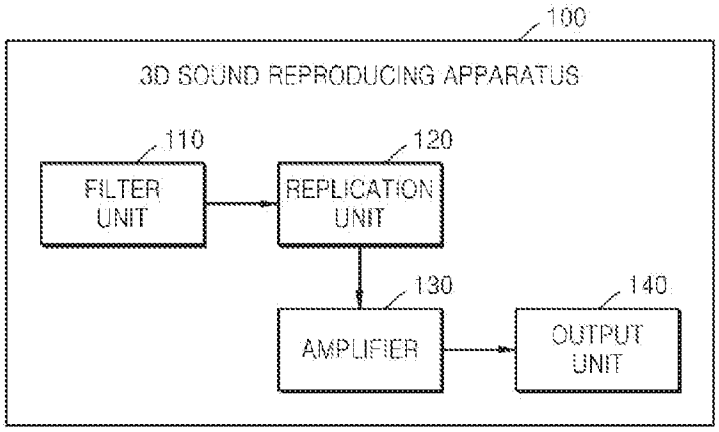
2017200552 27 Jan 2017

22. The apparatus of claim 12, wherein the first height input channel signal is located at top center.

23. A non-transitory computer readable recording medium having embodied thereon a computer program for executing the method of claim 1.

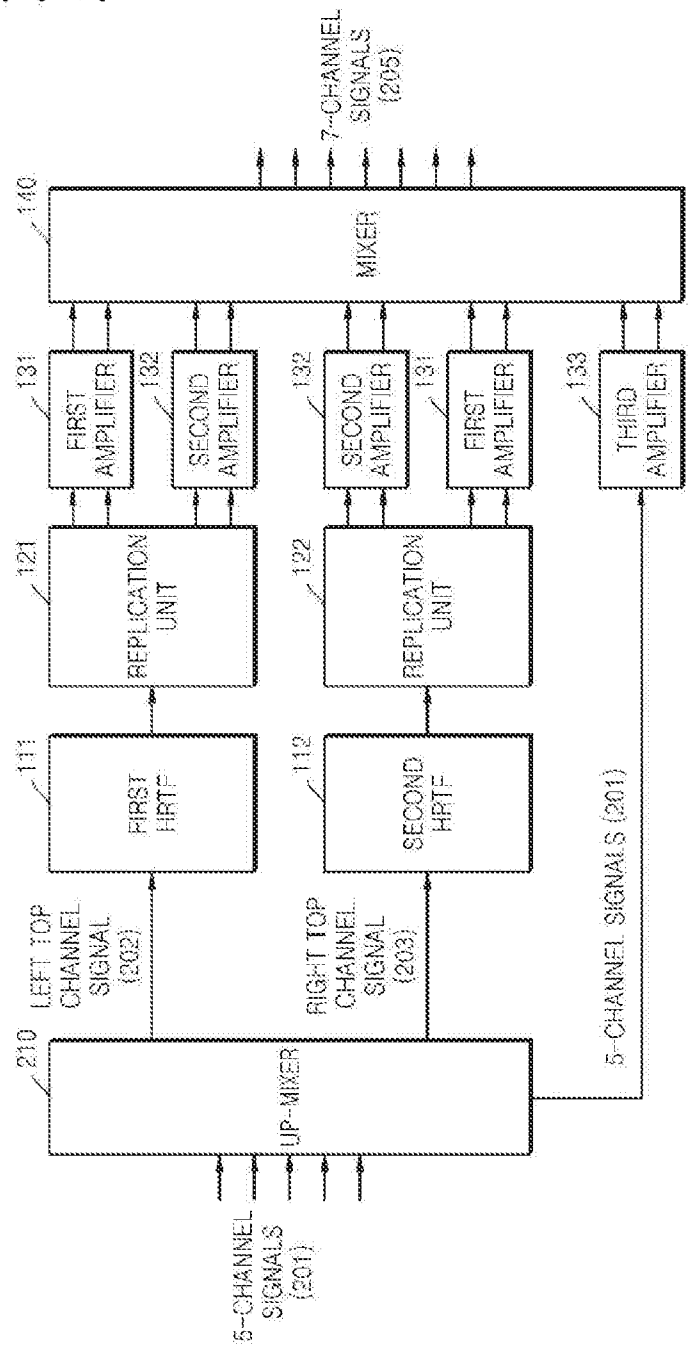
2017200552 27 Jan 2017

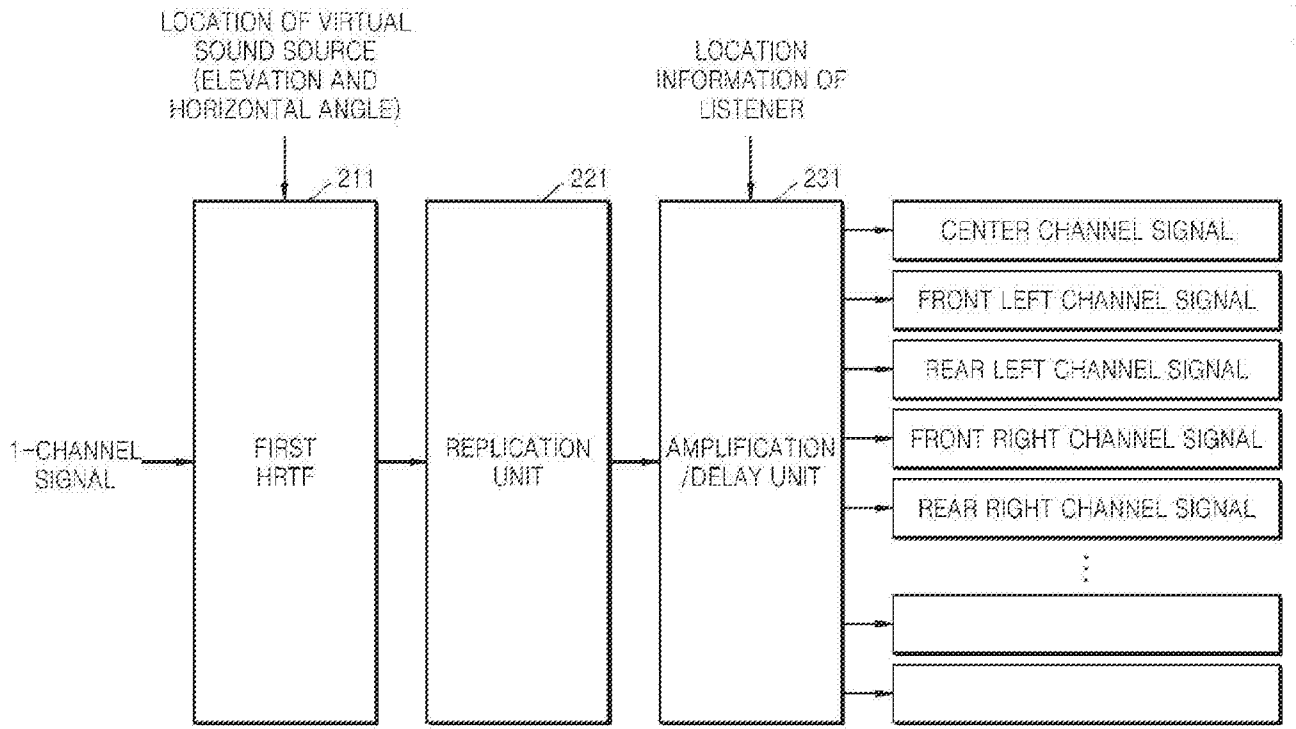
[Fig. 1]



2017200552 27 Jan 2017

[Fig. 2a]

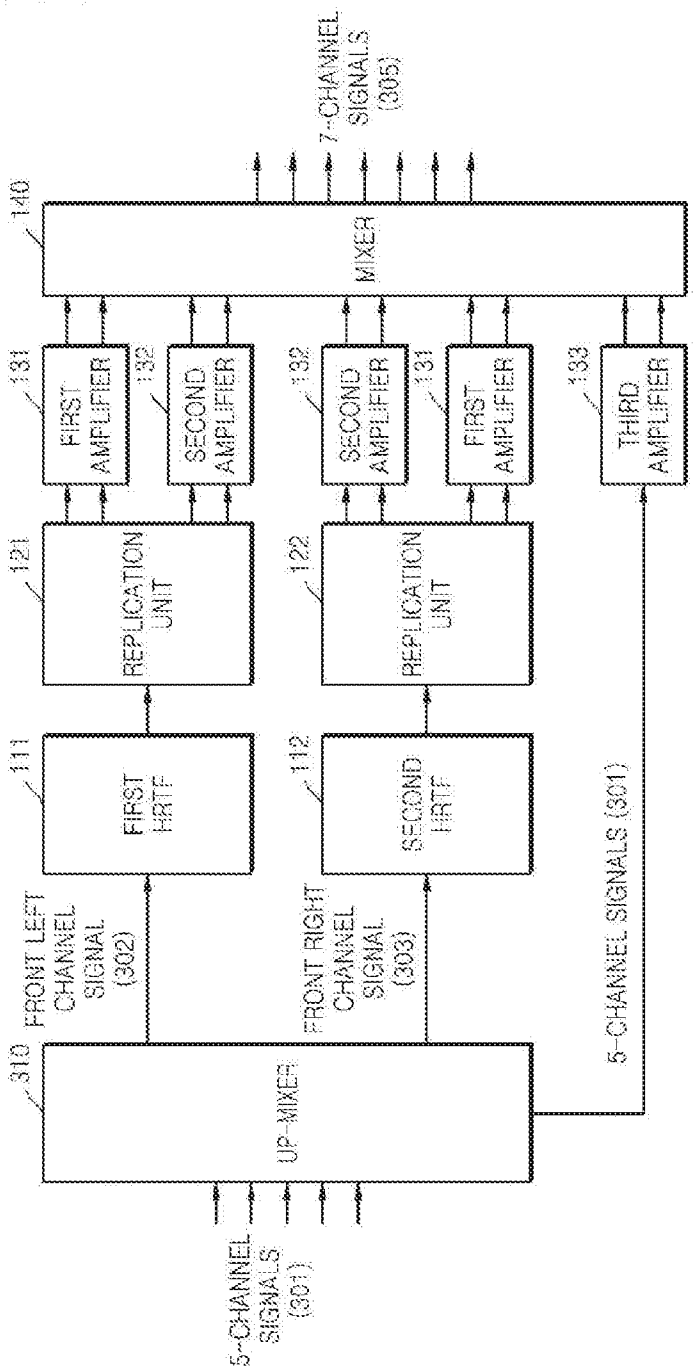



LOCATION OF VIRTUAL SOUND SOURCE (ELEVATION AND HORIZONTAL ANGLE)
LOCATION INFORMATION OF LISTENER
211
221
231
1-CHANNEL SIGNAL
FIRST HRTF
REPLICATION UNIT
AMPLIFICATION /DELAY UNIT
CENTER CHANNEL SIGNAL
FRONT LEFT CHANNEL SIGNAL
REAR LEFT CHANNEL SIGNAL
FRONT RIGHT CHANNEL SIGNAL
REAR RIGHT CHANNEL SIGNAL

[Fig. 2g]

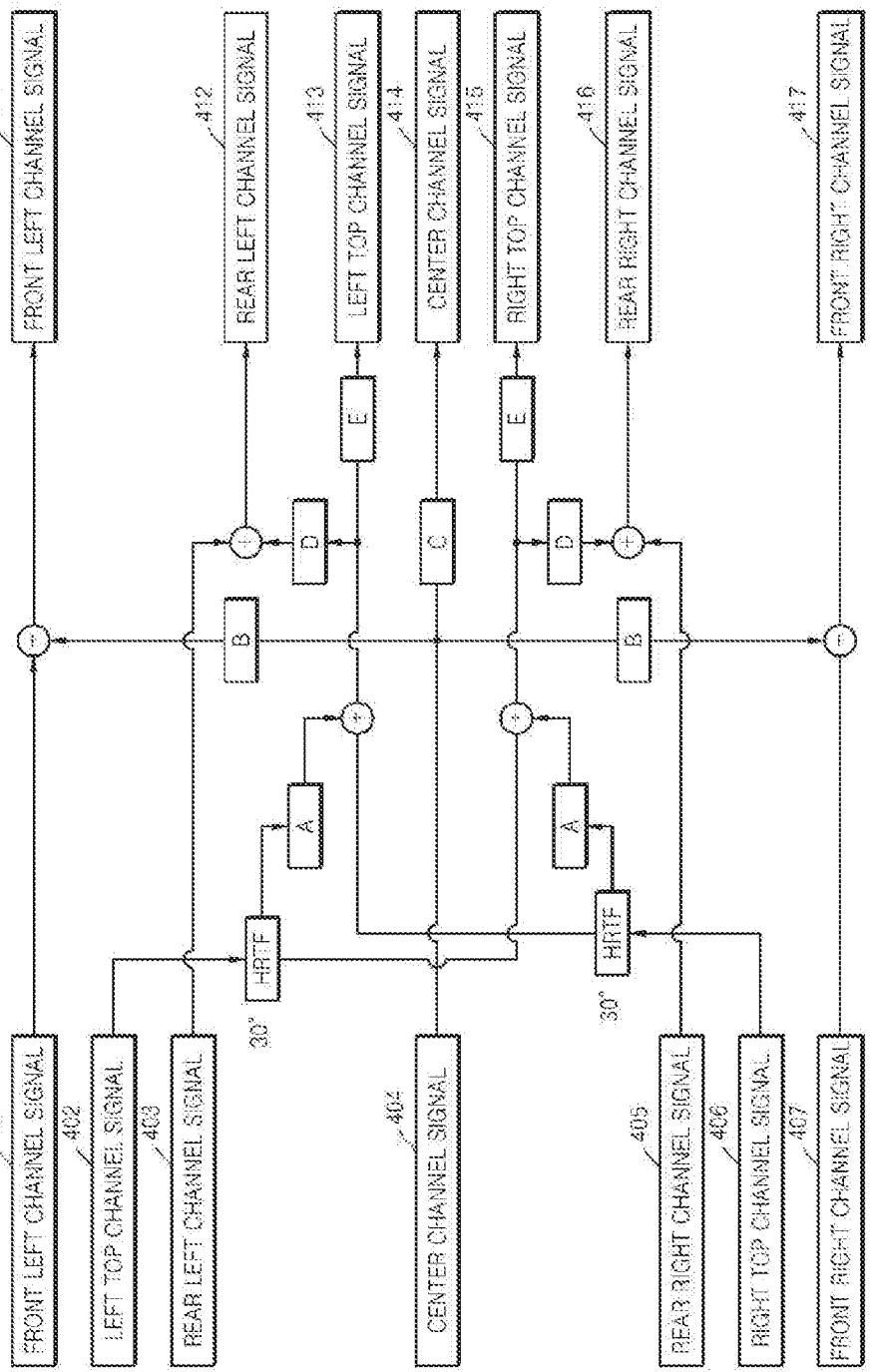
2017200552 27 Jan 2017

[Fig. 3]



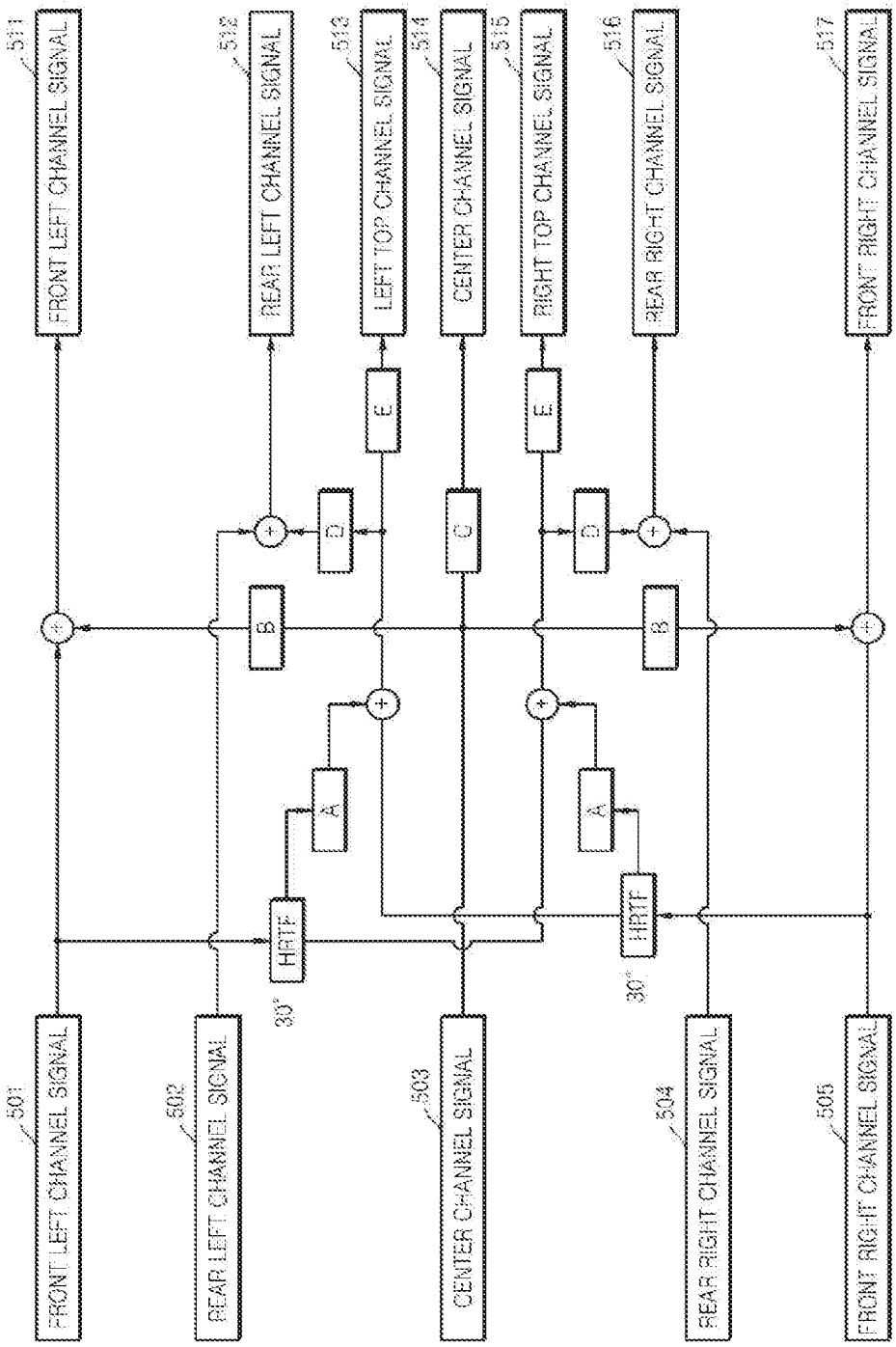
2017200552 27 Jan 2017

[Fig. 4]



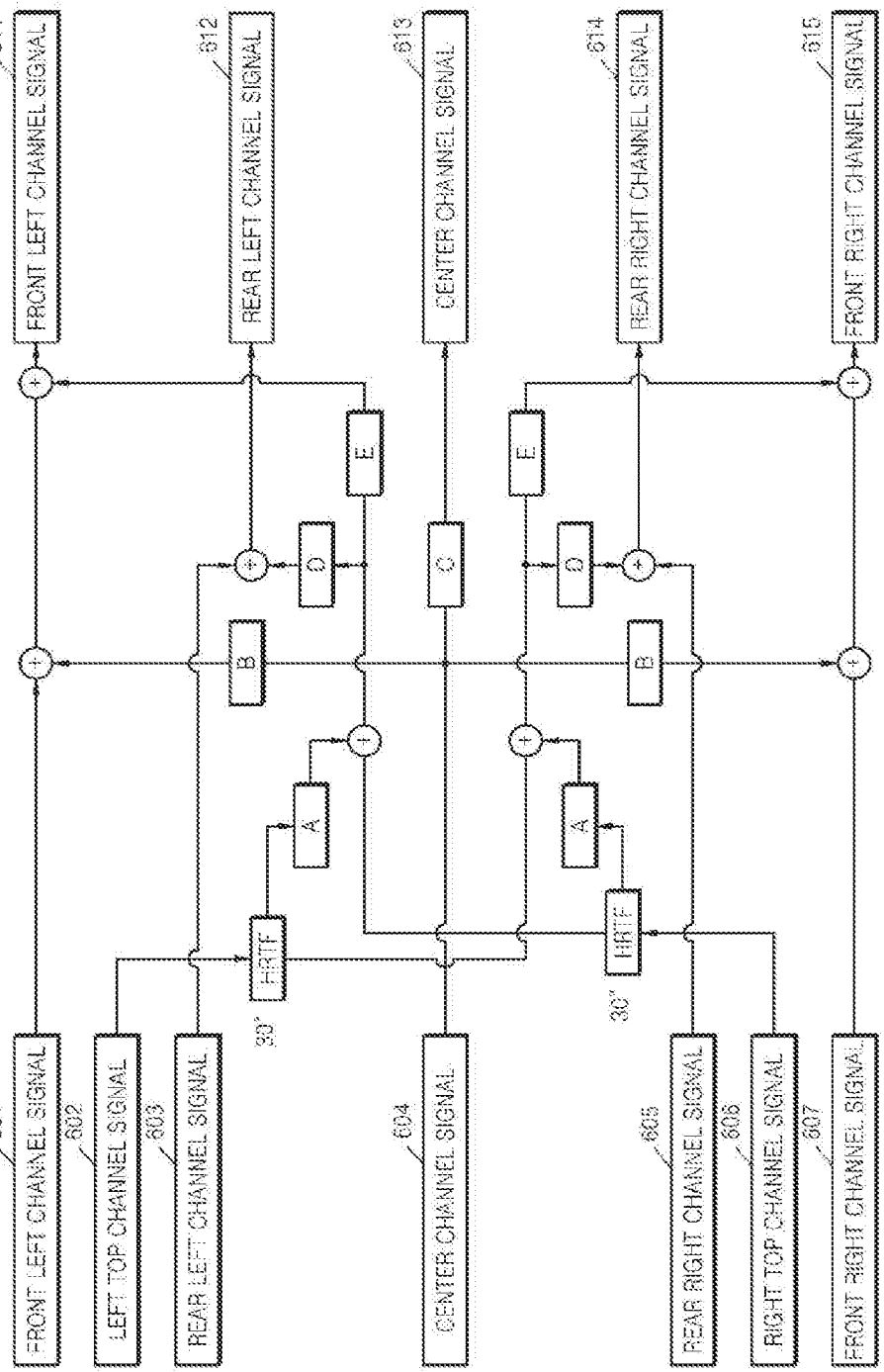
2017200552 27 Jan 2017

[Fig. 5]



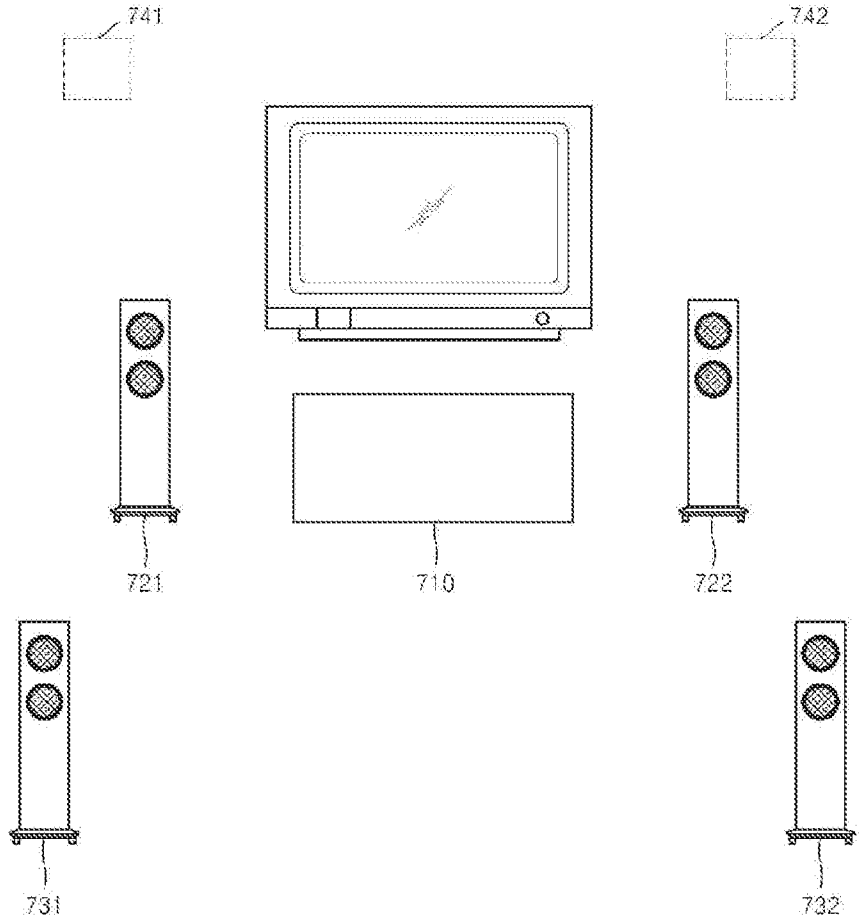
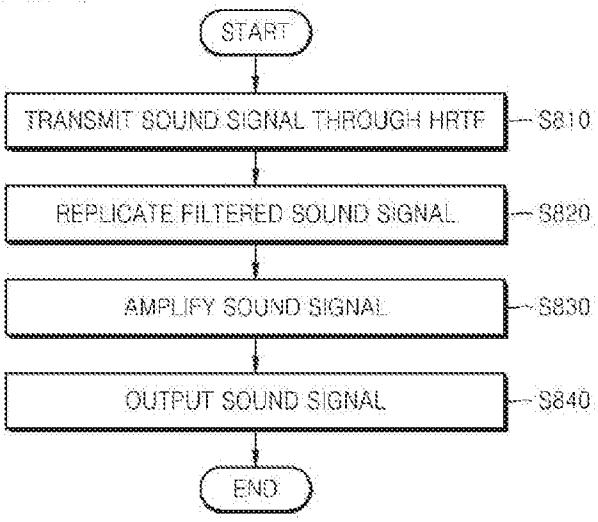
2017200552 27 Jan 2017

[Fig. 6]



20172005552 27 Jan 2017

[Fig. 7]



[Fig. 8]