AUDIO PROVIDING APPARATUS AND AUDIO PROVIDING METHOD

TECHNICAL FIELD

[0001] The inventive concept relates to an audio providing apparatus and method, and more particularly, to an audio providing apparatus and method which render and output audio signals having various formats to be optimal for an audio reproduction system.

BACKGROUND ART

[0002] At present, various audio formats are being used in the multimedia market. For example, an audio providing apparatus provides various audio formats from a 2-channel audio format to a 22.2-channel audio format. Particularly, an audio system, which uses channels such as 7.1 channel, 11.1 channel, and 22.2 channel for expressing a sound source in a three-dimensional space, is being provided.

[0003] However, most of currently provided audio signals have a 2.1-channel format or a 5.1-channel format and have a limitation in expressing a sound source in a three-dimensional space. Also, it is really difficult to setup, in homes, an audio system for reproducing 7.1-channel, 11.1-channel, and 22.2-channel audio signals.

[0004] Therefore, it is required to develop a method of actively rendering an audio signal according to a format of an input signal and an audio reproducing system.

DETAILED DESCRIPTION OF THE INVENTIVE CONCEPT

TECHNICAL PROBLEM

[0005] The inventive concept provides an audio providing method and an audio providing apparatus using the method, which optimize a channel audio signal for a listening environment by up-mixing or down-mixing the channel audio signal and render an object audio signal according to geometric information to provide a sound image optimized for the listening environment.

TECHNICAL SOLUTION

[0006] According to an aspect of the inventive concept, there is provided an audio providing apparatus including: an object rendering unit that renders an object audio signal

based on geometric information regarding the object audio signal; a channel rendering unit that renders an audio signal having a first channel number into an audio signal having a second channel number; and a mixing unit that mixes the rendered object audio signal with the audio signal having the second channel number.

[0007] The object rendering unit may include: a geometric information analyzer that converts the geometric information regarding the object audio signal into three-dimensional (3D) coordinate information; a distance controller that generates distance control information, based on the 3D coordinate information; a depth controller that generates depth control information, based on the 3D coordinate information; a localizer that generates localization information for localizing the object audio signal, based on the 3D coordinate information; and a renderer that renders the object audio signal, based on the distance control information, the depth control information, and the localization information.

[0008] The distance controller may acquire a distance gain of the object audio signal. As a distance of the object audio signal increases, the distance controller may decrease the distance gain of the object audio signal, and as the distance of the object audio signal decreases, the distance controller may increase the distance gain of the object audio signal.

[0009] The depth controller may acquire a depth gain, based on a horizontal projection distance of the object audio signal, and the depth gain may be expressed as a sum of a negative vector and a positive vector or may be expressed as a sum of the negative vector and a null vector.

[0010] The localizer acquires a panning gain for localizing the object audio signal according to a speaker layout of the audio providing apparatus.

[0011] The renderer may render the object audio signal into a multi-channel, based on the depth gain, the panning gain, and the distance gain of the object audio signal.

[0012] When the object audio signal is plurality, the object rendering unit may acquire a phase difference between a plurality of object audio signals having a correlation among the plurality of object audio signals and move one of the plurality of object audio signals by the acquired phase difference to combine the plurality of object audio signals.

[0013] When the audio providing apparatus reproduces audio by using a plurality of speakers having the same elevation, the object rendering unit may include: a virtual filter that corrects spectral characteristics of the object audio signal and adds virtual elevation information to the object audio signal; and a virtual renderer that renders the object audio signal, based on the virtual elevation information supplied by the virtual filter.

[0014] The virtual filter may have a tree structure consisting of a plurality of stages.

[0015] When a layout of the audio signal having the first channel number is a two-dimensional (2D) layout, the channel rendering unit may up-mix the audio signal having the first channel number to the audio signal having the second channel number which is greater than the first channel number, and a layout of the audio signal having the second channel number may be a three-dimensional (3D) layout having elevation information which differs from elevation information regarding the audio signal having the first channel number.

[0016] When a layout of the audio signal having the first channel number is a three-dimensional (3D) layout, the channel rendering unit may down-mix the audio signal having the first channel number to the audio signal having the second channel number which is less than the first channel number, and a layout of the audio signal having the second channel number may be a two-dimensional (2D) layout where a plurality of channels have the same elevation component.

[0017] At least one selected from the object audio signal and the audio signal having the first channel number may include information for determining whether to perform virtual three-dimensional (3D) rendering on a specific frame.

[0018] The channel rendering unit may acquire a phase difference between a plurality of audio signals having a correlation in an operation of rendering the audio signal having the first channel number into the audio signal having the second channel number, and move one of the plurality of audio signals by the acquired phase difference to combine the plurality of audio signals.

[0019] The mixing unit may acquire a phase difference between a plurality of audio signals having a correlation while mixing the rendered object audio signal with the audio signal having the second channel number, and move one of the plurality of audio signals by the acquired phase difference to combine the plurality of audio signals.

[0020] The object audio signal may include at least one of an identification (ID) and type information regarding the object audio signal for enabling a user to select the object audio signal.

[0021] According to another aspect of the inventive concept, there is provided an audio providing method including: rendering an object audio signal based on geometric information regarding the object audio signal; rendering an audio signal having a first channel number into an audio signal having a second channel number; and mixing the rendered object audio signal with the audio signal having the second channel number.

[0022] The rendering of the object audio signal may include: converting the geometric information regarding the object audio signal into three-dimensional (3D) coordinate

information; generating distance control information, based on the 3D coordinate information; generating depth control information, based on the 3D coordinate information; generating localization information for localizing the object audio signal, based on the 3D coordinate information; and rendering the object audio signal, based on the distance control information, the depth control information, and the localization information.

[0023] The generating of the distance control information may include: acquiring a distance gain of the object audio signal; decreasing the distance gain of the object audio signal as a distance of the object audio signal increases; and increasing the distance gain of the object audio signal as the distance of the object audio signal decreases.

[0024] The generating of the depth control information may include acquiring a depth gain, based on a horizontal projection distance of the object audio signal, and the depth gain may be expressed as a sum of a negative vector and a positive vector or may be expressed as a sum of the negative vector and a null vector.

[0025] The generating of the localization information may include acquiring a panning gain for localizing the object audio signal according to a speaker layout of an audio providing apparatus.

[0026] The rendering may include rendering the object audio signal to a multi-channel, based on the depth gain, the panning gain, and the distance gain of the object audio signal.

[0027] The rendering of the object audio signal may include: when the object audio signal is plurality, acquiring a phase difference between plurality of object audio signals having a correlation among the plurality of object audio signals; and moving one of the plurality of object audio signals by the acquired phase difference to combine the plurality of object audio signals.

[0028] When an audio providing apparatus reproduces audio by using a plurality of speakers having the same elevation, the rendering of the object audio signal may include: correcting spectral characteristics of the object audio signal and adding virtual elevation information to the object audio signal; and rendering the object audio signal, based on the virtual elevation information supplied by the virtual filter.

[0029] The acquiring may include virtual elevation information regarding the object audio signal by using a virtual filter which has a tree structure consisting of a plurality of stages.

[0030] The rendering of the audio signal having the first channel number into the audio signal having the second channel number may include, when a layout of the audio signal having the first channel number is a two-dimensional (2D) layout, up-mixing the audio signal

having the first channel number to the audio signal having the second channel number which is greater than the first channel number, and a layout of the audio signal having the second channel number may be a three-dimensional (3D) layout having elevation information which differs from elevation information regarding the audio signal having the first channel number.

[0031] The rendering of the audio signal having the first channel number to the audio signal having the second channel number may include, when a layout of the audio signal having the first channel number is a three-dimensional (3D) layout, down-mixing the audio signal having the first channel number to the audio signal having the second channel number which is less than the first channel number, and a layout of the audio signal having the second channel number may be a two-dimensional (2D) layout where a plurality of channels have the same elevation component.

[0032] At least one selected from the object audio signal and the audio signal having the first channel number may include information for determining whether to perform virtual three-dimensional (3D) rendering on a specific frame.

ADVANTAGEOUS EFFECTS

[0033] According to various embodiments of the present invention, an audio providing apparatus reproduce audio signals having various formats to be optimal for an output audio system.

DESCRIPTION OF THE DRAWINGS

[0034] FIG. 1 is a block diagram illustrating a configuration of an audio providing apparatus according to an exemplary embodiment of the present invention.

[0035] FIG. 2 is a block diagram illustrating a configuration of an object rendering unit according to an exemplary embodiment of the present invention.

[0036] FIG. 3 is a diagram for describing geometric information of an object audio signal according to an exemplary embodiment of the present invention.

[0037] FIG. 4 is a graph for describing a distance gain based on distance information of an object audio signal according to an exemplary embodiment of the present invention.

[0038] FIGS. 5A and 5B are graphs for describing a depth gain based on depth information of an object audio signal according to an exemplary embodiment of the present invention.

[0039] FIG. 6 is a block diagram illustrating a configuration of an object rendering unit for providing a virtual three-dimensional (3D) object audio signal, according to another exemplary embodiment of the present invention.

[0040] FIGS. 7A and 7B are diagrams for describing a virtual filter according to an exemplary embodiment of the present invention.

[0041] FIGS. 8A and 8B are diagrams for describing channel rendering of an audio signal according to various exemplary embodiments of the present invention.

[0042] FIG. 9 is a flowchart for describing an audio signal providing method according to an exemplary embodiment of the present invention.

[0043] FIG. 10 is a block diagram illustrating a configuration of an audio providing apparatus according to another exemplary embodiment of the present invention.

BEST MODE

[0044] Hereinafter, the present invention will be described in detail with reference to the accompanying drawings. FIG. 1 is a block diagram illustrating a configuration of an audio providing apparatus 100 according to an exemplary embodiment of the present invention. As illustrated in FIG. 1, the audio providing apparatus 100 includes an input unit 110, a de-multiplexer 120, an object rendering unit 130, a channel rendering unit 140, a mixing unit 150, and an output unit 160.

[0045] The input unit 110 may receive an audio signal from various sources. In this case, an audio source may include a channel audio signal and an object audio signal. Here, the channel audio signal is an audio signal including a background sound of a corresponding frame and may have a first channel number (for example, 5.1 channel, 7.1 channel, etc.). Also, the object audio signal may be an object having a motion or an audio signal of an important object in a corresponding frame. Examples of the object audio signal may include voice, gunfire, etc. The object audio signal may include geometric information of the object audio signal.

[0046] The de-multiplexer 120 may de-multiplex the channel audio signal and the object audio signal from the received audio signal. Also, the de-multiplexer 120 may respectively output the de-multiplexed object audio signal and channel audio signal to the object rendering unit 130 and the channel rendering unit 140.

[0047] The object rendering unit 130 may render the received object audio signal, based on geometric information regarding the received object audio signal. In this case, the object audio rendering unit 130 may render the received object audio signal according to a speaker layout of the audio providing apparatus 100. For example, when the speaker layout of the audio providing apparatus 100 is a two-dimensional (2D) layout having the same elevation, the object rendering unit 130 may two-dimensionally render the received object audio signal. Also, when the speaker layout of the audio providing apparatus 100 is a 3D layout having a plurality of elevations, the object rendering unit 130 may three-dimensionally render the received object audio signal. Also, although the speaker layout of the audio providing apparatus 100 is the 2D layout having the same elevation, the object rendering unit 130 may add virtual elevation information to the received object audio signal and three-dimensionally render the object audio signal. The object rendering unit 130 will be described in detail with reference to FIGS. 2 to 7B.

[0048] FIG. 2 is a block diagram illustrating a configuration of the object rendering unit 130 according to an exemplary embodiment of the present invention. As illustrated in FIG. 2, the object rendering unit 130 may include a geometric information analyzer 131, a distance controller 132, a depth controller 133, a localizer 134, and a renderer 135.

[0049] The geometric information analyzer 131 may receive and analyze geometric information regarding an object audio signal. In detail, the geometric information analyzer 131 may convert the geometric information regarding the object audio signal into 3D coordinate information necessary for rendering. For example, the geometric information analyzer 131, as illustrated in FIG. 3, may analyze the received object audio signal "O" into coordinate information (r, Θ, ϕ). Here, r denotes a distance between a position of a listener and the object audio signal, Θ denotes a azimuth angle of a sound image, and ϕ denotes an elevation angle of the sound image.

[0050] The distance controller 132 may generate distance control information, based on the 3D coordinate information. In detail, the distance controller 132 may calculate a distance gain of the object audio signal, based on a 3D distance "r" obtained through analysis by the geometric information analyzer 131. In this case, the distance controller 132 may calculate the distance gain in inverse proportion to the 3D distance "r". That is, as a distance of the object audio signal increases, the distance controller 132 may decrease the distance gain of the object audio signal, and as the distance of the object audio signal decreases, the distance controller 132 may increase the distance gain of the object audio signal. Also, when a position is closer to the origin point, the distance controller 132 may set an upper limit gain

value which is not purely inverse proportion, in order for the distance gain not to diverge. For example, the distance controller 132 may calculate the distance gain "d_g" as expressed in the following Equation (1):

$$d_g = \frac{1}{(0.3 + 0.7r)} \qquad \ldots(1)$$

[0051] That is, as illustrated in FIG. 4, the distance controller 132 may set the distance gain value "d_g" to 1 to 3.3, based on Equation (1).

[0052] The depth controller 133 may generate depth control information, based on the 3D coordinate information. In this case, the depth controller 133 may acquire a depth gain, based on a horizontal projection distance "d" of the object audio signal and the position of the listener.

[0053] In this case, the depth controller 133 may express the depth gain as a sum of a negative vector and a positive vector. In detail, when r<1 in 3D coordinates of the object audio signal, namely, when the object audio signal is located in a sphere consisting of a speaker included in the audio providing apparatus 100, the positive vector is defined as (r, Θ, ϕ), and the negative vector is defined as (r, Θ+180, ϕ). In order to define the object audio signal, the depth controller 133 may calculate a depth gain "v_p" of the positive vector and a depth gain "v_n" of the negative vector for expressing a geometric vector of the object audio signal as a sum of the positive vector and the negative vector. In this case, the depth gain "v_p" of the positive vector and the depth gain "v_n" of the negative vector may be calculated as expressed in the following Equation (2):

$$\begin{aligned} v_p &= \sin(dS\pi/2 + \pi/4) \\ v_n &= \cos(dS\pi/2 + \pi/4) \end{aligned} \qquad \ldots(2)$$

[0054] That is, as illustrated in FIG. 5A, the depth controller 133 may calculate the depth gain of the positive vector and the depth gain of the negative vector where the horizontal projection distance "d" is 0 to 1.

[0055] Moreover, the depth controller 133 may express the depth gain as a sum of the positive vector and the negative vector. In detail, a panning gain when there is no direction where a sum of multiplications of panning gains and positions of all channels converges to 0 may be defined as a null vector. Particularly, the depth controller 133 may calculate the depth gain "v_p" of the positive vector and a depth gain "v_{nll}" of the null vector so that when the horizontal projection distance "d" is close to 0, the depth gain of the null vector is mapped to 1, and when the horizontal projection distance "d" is close to 1, the depth gain of the positive

vector is mapped to 1. In this case, the depth gain "v_p" of the positive vector and the depth gain "v_{nll}" of the null vector may be calculated as expressed in the following Equation (3):

$$v_p = \sin(dS\pi/2)$$
$$v_{nll} = \cos(dS\pi/2) \qquad \ldots(3)$$

[0056] That is, as illustrated in FIG. 5B, the depth controller 133 may calculate the depth gain of the positive vector and the depth gain of the null vector where the horizontal projection distance "d" is 0 to 1.

[0057] Depth control is performed by the depth controller 133, and when the horizontal projection distance is close to 0, a sound may be output through all speakers. Therefore, a discontinuity, which occurs in a panning boundary reduces.

[0058] The localizer 134 may generate localization information for localizing the object audio signal, based on the 3D coordinate information. Particularly, the localizer 134 may calculate a panning gain for localizing the object audio signal according to the speaker layout of the audio providing apparatus 100. In detail, the localizer 134 may select a triplet speaker for localizing the positive vector having the same direction as that of a geometry of the object audio signal and calculate a 3D panning coefficient "g_p" for the triplet speaker of the positive vector. Also, when the depth controller 133 expresses a depth gain with the positive vector and the negative vector, the localizer 134 may select a triplet speaker for localizing the negative vector having a direction opposite to a direction of the trajectory of the object audio signal and calculate a 3D panning coefficient "g_n" for the triplet speaker of the negative vector.

[0059] The renderer 135 may render the object audio signal, based on the distance control information, the depth control information, and the localization information. Particularly, the renderer 135 may receive the distance gain "d_g" from the distance controller 132, receive a depth gain "v" from the depth controller 133, receive a panning gain "g" from the localizer 134, and apply the distance gain "d_g", the depth gain "v", and the panning gain "g" to the object audio signal to generate a multi-channel object audio signal. Particularly, when the depth gain of the object audio signal is expressed as a sum of the positive vector and the negative vector, the renderer 135 may calculate an mth-channel final gain "Gm" as expressed in the following Equation (4):

$$G_m = d_g S(g_{p,m} S v_p + g_{n,m} S v_n) \qquad \ldots(4)$$

where $g_{p,m}$ denotes a panning coefficient applied to an m channel when the positive vector is localized, and $g_{n,m}$ denotes a panning coefficient applied to the m channel when the negative vector is localized.

[0060] Moreover, when the depth gain of the object audio signal is expressed as a sum of the positive vector and the null vector, the renderer 135 may calculate the mth-channel final gain "Gm" as expressed in the following Equation (5):

$$G_m = d_g S(g_{p,m} S v_p + g_{nll,m} S v_{nll}) \qquad \ldots(5)$$

where $g_{p,m}$ denotes a panning coefficient applied to an m channel when the positive vector is localized, and $g_{n,m}$ denotes a panning coefficient applied to the m channel when the negative vector is localized. Furthermore, $\sum g_{nll,m}$ may become 0.

[0061] Moreover, the renderer 135 may apply the final gain to the object audio signal "x" to calculate a final output "Y_m" of an mth-channel object audio signal as expressed in the following Equation (6):

$$Y_m = XsG_m \qquad \ldots(6)$$

[0062] The final output "Y_m" of the object audio signal calculated as described above may be output to the mixing unit 150.

[0063] Moreover, when there are a plurality of object audio signals, the object rendering unit 130 may calculate a phase difference between the plurality of object audio signals and move one of the plurality of object audio signals by the calculated phase difference to combine the plurality of object audio signals.

[0064] In detail, in a case where a plurality of object audio signals are the same signals but have opposite phases while the plurality of object audio signals are being input, when the plurality of object audio signals are combined as-is, an audio signal is distorted due to overlapping of the plurality of object audio signals. Therefore, the object rendering unit 130 may calculate a correlation between the plurality of object audio signals, and when the correlation is equal to or greater than a predetermined value, the object rendering unit 130 may calculate a phase difference between the plurality of object audio signals and move one of the plurality of object audio signals by the calculated phase difference to combine the plurality of object audio signals. Accordingly, when a plurality of object audio signals similar thereto are input, distortion caused by combination of the plurality of object audio signals is prevented.

[0065] In the above-described exemplary embodiment, the speaker layout of the audio providing apparatus 100 is the 3D layout having different senses of elevation, but this is merely an exemplary embodiment. The speaker layout of the audio providing apparatus 100 may be a 2D layout having the same value of elevation. Particularly, when the speaker layout of the audio providing apparatus 100 is the 2D layout having the same sense of elevation, the

object rendering unit 130 may set a value of ϕ, included in the above-described geometric information regarding the object audio signal, to 0.

[0066] Moreover, the speaker layout of the audio providing apparatus 100 may be the 2D layout having the same sense of elevation, but the audio providing apparatus 100 may virtually provide a 3D object audio signal by using a 2D speaker layout.

[0067] Hereinafter, an exemplary embodiment for providing a virtual 3D object audio signal will be described with reference to FIGS. 6 and 7.

[0068] FIG. 6 is a block diagram illustrating a configuration of an object rendering unit 130' for providing a virtual 3D object audio signal, according to another exemplary embodiment of the present invention. As illustrated in FIG. 6, the object rendering unit 130' includes a virtual filter 136, a 3D renderer 137, a virtual renderer 138, and a mixer 139.

[0069] The 3D renderer 137 may render an object audio signal by using the method described above with reference to FIGS. 2 to 5B. In this case, the 3D renderer 137 may output the object audio signal, which is capable of being output through a physical speaker of the audio providing apparatus 100, to the mixer 139 and output a virtual panning gain "$g_{m,top}$" of a virtual speaker providing different senses of elevation.

[0070] The virtual filter 136 is a block that compensates a tone color of an object audio signal. The virtual filter 136 may compensate spectral characteristics of an input object audio signal, based on psychoacoustics and provide a sound image to a position of the virtual speaker. In this case, the virtual filter 136 may be implemented as filters of various types such as a head-related transfer function (HRTF) filter, a binaural room impulse response (BRIR) filter, etc.

[0071] Moreover, when the length of the virtual filter 136 is less than that of a frame, the virtual filter 136 may be applied through block convolution.

[0072] Moreover, when rendering is performed in a frequency domain such as a fast Fourier transform (FFT), a modified discrete cosine transform (MDCT), and a quadrature mirror filter (QMF), the virtual filter 136 may be applied as multiplication.

[0073] When a plurality of virtual top layer speakers are provided, the virtual filter 136 may generate the plurality of virtual top layer speakers by using a distribution formula of physical speakers and one elevation filter.

[0074] Moreover, when a plurality of virtual top layer speakers and a virtual back speaker are provided, the virtual filter 136 may generate the plurality of virtual top layer speakers and the virtual back speaker by using a distribution formula of physical speakers and a plurality of virtual filters, for applying a spectral coloration at different positions.

[0075] Moreover, in if N number of spectral coloration such as H1, H2, ..., HN are used, the virtual filter 136 may be designed in a tree structure so as to reduce the number of arithmetic operations. In detail, as illustrated in FIG. 7A, the virtual filter 136 may design a notch/peak, which is used to recognize a height in common, to H0 and connect K1 to KN, which are components obtained by subtracting a characteristic of H0 from H1 to HN, to HO in a cascade type. Also, the virtual filter 136 may have a tree structure consisting of a plurality of stages illustrated in FIG. 7B, based on a common component and spectral coloration.

[0076] The virtual renderer 138 is a rendering block for expressing a virtual channel as a physical channel. Particularly, the virtual renderer 138 may generate an object audio signal that is output to the virtual speaker according to a virtual channel distribution formula output from the virtual filter 136 and multiply the generated object audio signal of the virtual speaker by the virtual panning gain "$g_{m,top}$" to combine output signals. In this case, a position of the virtual speaker may be changed according to a degree of distribution to a plurality of physical flat cone speakers, and the degree of distribution may be defined as the virtual channel distribution formula.

[0077] The mixer 139 may mix a physical-channel object audio signal with a virtual-channel object audio signal.

[0078] Therefore, an object audio signal may be expressed as being located on a 3D layout by using the audio providing apparatus 100 having a 2D speaker layout.

[0079] Referring again to FIG. 1, the channel rendering unit 140 may render a channel audio signal having a first channel number into an audio signal having a second channel number. In this case, the channel rendering unit 140 may change the channel audio signal having the first channel number to the audio signal having the second channel number, based on a speaker layout.

[0080] In detail, when a layout of a channel audio signal is the same as a speaker layout of the audio providing apparatus 100, the channel rendering unit 140 may render the channel audio signal without changing a channel.

[0081] Moreover, when the number of channels of the channel audio signal is more than the number of channels of the speaker layout of the audio providing apparatus 100, the channel rendering unit 140 may down-mix the channel audio signal to perform rendering. For example, when a channel of the channel audio signal is 7.1 channel and the speaker layout of the audio providing apparatus 100 is 5.1 channel, the channel rendering unit 140 may down-mix the channel audio signal having 7.1 channel to 5.1 channel.

[0082] Particularly, when down-mixing the channel audio signal, the channel rendering unit 140 may determine an object where a geometry of the channel audio signal is stopped without any change, and perform down-mixing. Also, when down-mixing a 3D channel audio signal to a 2D signal, the channel rendering unit 140 may remove an elevation component of the channel audio signal to two-dimensionally down-mix the channel audio signal or to three-dimensionally down-mix the channel audio signal so as to have a sense of virtual elevation, as described above with reference to FIG. 6. Also, the channel rendering unit 140 may down-mix all signals except a front left channel, a front right channel, and a center channel which constitute a front audio signal, thereby implementing a signal with a right surround channel and a left surround channel. Also, the channel rendering unit 140 may perform down-mixing by using a multi-channel down-mix equation.

[0083] Moreover, when the number of channels of the channel audio signal is less than the number of channels of the speaker layout of the audio providing apparatus 100, the channel rendering unit 140 may up-mix the channel audio signal to perform rendering. For example, when a channel of the channel audio signal is 7.1 channel and the speaker layout of the audio providing apparatus 100 is 9.1 channel, the channel rendering unit 140 may up-mix the channel audio signal having 7.1 channel to 9.1 channel.

[0084] Particularly, when up-mixing a 2D channel audio signal to a 3D signal, the channel rendering unit 140 may generate a top layer having an elevation component, based on a correlation between a front channel and a surround channel to perform up-mixing, or divide channels into a center channel and an ambience channel through analysis of the channels to perform up-mixing.

[0085] Moreover, the channel rendering unit 140 may calculate a phase difference between a plurality of audio signals having a correlation in an operation of rendering the channel audio signal having the first channel number to the channel audio signal having the second channel number, and move one of the plurality of audio signals by the calculated phase difference to combine the plurality of audio signals.

[0086] At least one of the object audio signal and the channel audio signal having the first channel number may include guide information for determining whether to perform virtual 3D rendering or 2D rendering on a specific frame. Therefore, each of the object rendering unit 130 and the channel rendering unit 140 may perform rendering based on the guide information included in the object audio signal and the channel audio signal. For example, when guide information which allows virtual 3D rendering to be performed on an object audio signal in a first frame is included in the object audio signal, the object rendering unit

130 and the channel rendering unit 140 may perform virtual 3D rendering on the object audio signal and a channel audio signal in the first frame. Also, when guide information which allows 2D rendering to be performed on an object audio signal in a second frame is included in the object audio signal, the object rendering unit 130 and the channel rendering unit 140 may perform 2D rendering on the object audio signal and a channel audio signal in the second frame.

[0087] The mixing unit 150 may mix the object audio signal, which is output from the object rendering unit 130, with the channel audio signal having the second channel number which is output from the channel rendering unit 140.

[0088] Moreover, the mixing unit 150 may calculate a phase difference between a plurality of audio signals having a correlation while mixing the rendered object audio signal with the channel audio signal having the second channel number, and move one of the plurality of audio signals by the calculated phase difference to combine the plurality of audio signals.

[0089] The output unit 160 may output an audio signal which is output from the mixing unit 150. In this case, the output unit 160 may include a plurality of speakers. For example, the output unit 160 may be implemented with speakers such as 5.1 channel, 7.1 channel, 9.1 channel, 22.2 channel, etc.

[0090] Hereinafter, various exemplary embodiments of the present invention will be described with reference to FIGS. 8A to 8G.

[0091] FIG. 8A is a diagram for describing rendering of an object audio signal and a channel audio signal, according to a first exemplary embodiment of the present invention.

[0092] First, the audio providing apparatus 100 may receive a 9.1-channel channel audio signal and two object audio signals O1 and O2. In this case, the 9.1-channel channel audio signal may include a front left channel (FL), a front right channel (FR), a front center channel (FC), a subwoofer channel (Lfe), a surround left channel (SL), a surround right channel (SR), a top front left channel (TL), a top front right channel (TR), a back left channel (BL), and a back right channel (BR).

[0093] The audio providing apparatus 100 may be configured with a 5.1-channel speaker layout. That is, the audio providing apparatus 100 may include a plurality of speakers respectively corresponding to the front right channel, the front left channel, the front center channel, the subwoofer channel, the surround left channel, and the surround right channel.

[0094] The audio providing apparatus 100 may perform virtual filtering on signals respectively corresponding to the top front left channel, the top front right channel, the back

left channel, and the back right channel among a plurality of input channel audio signals to perform rendering.

[0095] Moreover, the audio providing apparatus 100 may perform virtual 3D rendering on a first object audio signal O1 and a second object audio signal O2.

[0096] The audio providing apparatus 100 may mix a channel audio signal having the front left channel, a channel audio signal having the virtually-rendered top front left channel and top front right channel, a channel audio signal having the virtually-rendered back left channel and back right channel, and the virtually-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the front left channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the front right channel, a channel audio signal having the virtually-rendered top front left channel and top front right channel, a channel audio signal having the virtually-rendered back left channel and back right channel, and the virtually-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the front right channel. Also, the audio providing apparatus 100 may output a channel audio signal having the front center channel to a speaker corresponding to the front center channel and output a channel audio signal having the subwoofer channel to a speaker corresponding to the subwoofer channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the surround left channel, a channel audio signal having the virtually-rendered top front left channel and top front right channel, a channel audio signal having the virtually-rendered back left channel and back right channel, and the virtually-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the surround left channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the surround right channel, a channel audio signal having the virtually-rendered top front left channel and top front right channel, a channel audio signal having the virtually-rendered back left channel and back right channel, and the virtually-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the surround right channel.

[0097] By performing the above-described channel rendering and object rendering, the audio providing apparatus 100 may establish a 9.1-channel virtual 3D audio environment by using a 5.1-channel speaker.

[0098] FIG. 8B is a diagram for describing rendering of an object audio signal and a channel audio signal, according to a second exemplary embodiment of the present invention.

[0099] First, the audio providing apparatus 100 may receive a 9.1-channel channel audio signal and two object audio signals O1 and O2.

[00100] The audio providing apparatus 100 may be configured with a 7.1-channel speaker layout. That is, the audio providing apparatus 100 may include a plurality of speakers respectively corresponding to the front right channel, the front left channel, the front center channel, the subwoofer channel, the surround left channel, the surround right channel, the back left channel, and the back right channel.

[00101] The audio providing apparatus 100 may perform virtual filtering on signals respectively corresponding to the top front left channel and the top front right channel among a plurality of input channel audio signals to perform rendering.

[00102] Moreover, the audio providing apparatus 100 may perform virtual 3D rendering on a first object audio signal O1 and a second object audio signal O2.

[00103] The audio providing apparatus 100 may mix a channel audio signal having the front left channel, a channel audio signal having the virtually-rendered top front left channel and top front right channel, and the virtually-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the front left channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the front right channel, a channel audio signal having the virtually-rendered back left channel and back right channel, and the virtually-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the front right channel. Also, the audio providing apparatus 100 may output a channel audio signal having the front center channel to a speaker corresponding to the front center channel and output a channel audio signal having the subwoofer channel to a speaker corresponding to the subwoofer channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the surround left channel, a channel audio signal having the virtually-rendered top front left channel and top front right channel, and the virtually-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the surround left channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the surround right channel, a channel audio signal having the virtually-rendered top front left channel and top front right channel, and the virtually-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the surround right channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the back left channel and the virtually-rendered first object audio signal O1 and second object audio signal O2 and output a

mixed signal to a speaker corresponding to the back left channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the back right channel and the virtually-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the back right channel.

[00104] By performing the above-described channel rendering and object rendering, the audio providing apparatus 100 may establish a 9.1-channel virtual 3D audio environment by using a 7.1-channel speaker.

[00105] FIG. 8C is a diagram for describing rendering of an object audio signal and a channel audio signal, according to a third exemplary embodiment of the present invention.

[00106] First, the audio providing apparatus 100 may receive a 9.1-channel channel audio signal and two object audio signals O1 and O2.

[00107] The audio providing apparatus 100 may be configured with a 9.1-channel speaker layout. That is, the audio providing apparatus 100 may include a plurality of speakers respectively corresponding to the front right channel, the front left channel, the front center channel, the subwoofer channel, the surround left channel, the surround right channel, the back left channel, the back right channel, the top front left channel, and the top front right channel.

[00108] Moreover, the audio providing apparatus 100 may perform 3D rendering on a first object audio signal O1 and a second object audio signal O2.

[00109] The audio providing apparatus 100 may mix the 3D-rendered first object audio signal O1 and second object audio signal O2 with audio signals respectively having the front right channel, the front left channel, the front center channel, the subwoofer channel, the surround left channel, the surround right channel, the back left channel, the back right channel, the top front left channel, and the top front right channel, and output a mixed signal to a corresponding speaker.

[00110] By performing the above-described channel rendering and object rendering, the audio providing apparatus 100 may output a 9.1-channel channel audio signal and a 9.1-channel object audio signal by using a 9.1-channel speaker.

[00111] FIG. 8D is a diagram for describing rendering of an object audio signal and a channel audio signal, according to a fourth exemplary embodiment of the present invention.

[00112] First, the audio providing apparatus 100 may receive a 9.1-channel channel audio signal and two object audio signals O1 and O2.

[00113] The audio providing apparatus 100 may be configured with an 11.1-channel speaker layout. That is, the audio providing apparatus 100 may include a plurality of speakers

respectively corresponding to the front right channel, the front left channel, the front center channel, the subwoofer channel, the surround left channel, the surround right channel, the back left channel, the back right channel, the top front left channel, the top front right channel, a top surround left channel, a top surround right channel, a top back left channel, and a top back right channel.

[00114] Moreover, the audio providing apparatus 100 may perform 3D rendering on a first object audio signal O1 and a second object audio signal O2.

[00115] The audio providing apparatus 100 may mix the 3D-rendered first object audio signal O1 and second object audio signal O2 with audio signals respectively having the front right channel, the front left channel, the front center channel, the subwoofer channel, the surround left channel, the surround right channel, the back left channel, the back right channel, the top front left channel, and the top front right channel, and output a mixed signal to a corresponding speaker.

[00116] Moreover, the audio providing apparatus 100 may output the 3D-rendered first object audio signal O1 and second object audio signal O2 to a speaker corresponding to each of the top surround left channel, the top surround right channel, the top back left channel, and the top back right channel

[00117] By performing the above-described channel rendering and object rendering, the audio providing apparatus 100 may output a 9.1-channel channel audio signal and a 9.1-channel object audio signal by using an 11.1-channel speaker.

[00118] FIG. 8E is a diagram for describing rendering of an object audio signal and a channel audio signal, according to a fifth exemplary embodiment of the present invention.

[00119] First, the audio providing apparatus 100 may receive a 9.1-channel channel audio signal and two object audio signals O1 and O2.

[00120] The audio providing apparatus 100 may be configured with a 5.1-channel speaker layout. That is, the audio providing apparatus 100 may include a plurality of speakers respectively corresponding to the front right channel, the front left channel, the front center channel, the subwoofer channel, the surround left channel, and the surround right channel.

[00121] The audio providing apparatus 100 may perform 2D rendering on signals respectively corresponding to the top front left channel, the top front right channel, the back left channel, and the back right channel among a plurality of input channel audio signals.

[00122] Moreover, the audio providing apparatus 100 may perform 2D rendering on a first object audio signal O1 and a second object audio signal O2.

[00123] The audio providing apparatus 100 may mix a channel audio signal having the front left channel, a channel audio signal having the 2D-rendered top front left channel and top front right channel, a channel audio signal having the 2D-rendered back left channel and back right channel, and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the front left channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the front right channel, a channel audio signal having the 2D-rendered top front left channel and top front right channel, a channel audio signal having the 2D-rendered back left channel and back right channel, and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the front right channel. Also, the audio providing apparatus 100 may output a channel audio signal having the front center channel to a speaker corresponding to the front center channel and output a channel audio signal having the subwoofer channel to a speaker corresponding to the subwoofer channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the surround left channel, a channel audio signal having the 2D-rendered top front left channel and top front right channel, a channel audio signal having the 2D-rendered back left channel and back right channel, and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the surround left channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the surround right channel, a channel audio signal having the 2D-rendered top front left channel and top front right channel, a channel audio signal having the 2D-rendered back left channel and back right channel, and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the surround right channel.

[00124] By performing the above-described channel rendering and object rendering, the audio providing apparatus 100 may output a 9.1-channel channel audio signal and a 9.1-channel object audio signal by using a 5.1-channel speaker. In comparison with FIG. 8A, the audio providing apparatus 100 according to the present embodiment may render a signal not into a virtual 3D audio signal but into a 2D audio signal.

[00125] FIG. 8F is a diagram for describing rendering of an object audio signal and a channel audio signal, according to a sixth exemplary embodiment of the present invention.

[00126] First, the audio providing apparatus 100 may receive a 9.1-channel channel audio signal and two object audio signals O1 and O2.

[00127] The audio providing apparatus 100 may be configured with a 7.1-channel speaker layout. That is, the audio providing apparatus 100 may include a plurality of speakers

respectively corresponding to the front right channel, the front left channel, the front center channel, the subwoofer channel, the surround left channel, the surround right channel, the back left channel, and the back right channel.

[00128] The audio providing apparatus 100 may perform 2D rendering on signals respectively corresponding to the top front left channel and the top front right channel among a plurality of input channel audio signals.

[00129] Moreover, the audio providing apparatus 100 may perform 2D rendering on a first object audio signal O1 and a second object audio signal O2.

[00130] The audio providing apparatus 100 may mix a channel audio signal having the front left channel, a channel audio signal having the 2D-rendered top front left channel and top front right channel, and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the front left channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the front right channel, a channel audio signal having the 2D-rendered back left channel and back right channel, and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the front right channel. Also, the audio providing apparatus 100 may output a channel audio signal having the front center channel to a speaker corresponding to the front center channel and output a channel audio signal having the subwoofer channel to a speaker corresponding to the subwoofer channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the surround left channel, a channel audio signal having the 2D-rendered top front left channel and top front right channel, and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the surround left channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the surround right channel, a channel audio signal having the 2D-rendered top front left channel and top front right channel, and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the surround right channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the back left channel and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the back left channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the back right channel and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the back right channel.

[00131] By performing the above-described channel rendering and object rendering, the audio providing apparatus 100 may output a 9.1-channel channel audio signal and a 9.1-channel object audio signal by using a 7.1-channel speaker. In comparison with FIG. 8B, the audio providing apparatus 100 according to the present embodiment may render a signal not into a virtual 3D audio signal but into a 2D audio signal.

[00132] FIG. 8G is a diagram for describing rendering of an object audio signal and a channel audio signal, according to a seventh exemplary embodiment of the present invention.

[00133] First, the audio providing apparatus 100 may receive a 9.1-channel channel audio signal and two object audio signals O1 and O2.

[00134] The audio providing apparatus 100 may be configured with a 5.1-channel speaker layout. That is, the audio providing apparatus 100 may include a plurality of speakers respectively corresponding to the front right channel, the front left channel, the front center channel, the subwoofer channel, the surround left channel, and the surround right channel.

[00135] The audio providing apparatus 100 may two-dimensionally down-mix signals respectively corresponding to the top front left channel, the top front right channel, the back left channel, and the back right channel among a plurality of input channel audio signals to perform rendering.

[00136] Moreover, the audio providing apparatus 100 may perform virtual 3D rendering on a first object audio signal O1 and a second object audio signal O2.

[00137] The audio providing apparatus 100 may mix a channel audio signal having the front left channel, a channel audio signal having the 2D-rendered top front left channel and top front right channel, a channel audio signal having the 2D-rendered back left channel and back right channel, and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the front left channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the front right channel, a channel audio signal having the 2D-rendered top front left channel and top front right channel, a channel audio signal having the 2D-rendered back left channel and back right channel, and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the front right channel. Also, the audio providing apparatus 100 may output a channel audio signal having the front center channel to a speaker corresponding to the front center channel and output a channel audio signal having the subwoofer channel to a speaker corresponding to the subwoofer channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the surround left channel, a channel audio signal having the 2D-rendered top front left channel and top

front right channel, a channel audio signal having the 2D-rendered back left channel and back right channel, and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the surround left channel. Also, the audio providing apparatus 100 may mix a channel audio signal having the surround right channel, a channel audio signal having the 2D-rendered top front left channel and top front right channel, a channel audio signal having the 2D-rendered back left channel and back right channel, and the 2D-rendered first object audio signal O1 and second object audio signal O2 and output a mixed signal to a speaker corresponding to the surround right channel.

[00138] By performing the above-described channel rendering and object rendering, the audio providing apparatus 100 may output a 9.1-channel channel audio signal and a 9.1-channel object audio signal by using a 5.1-channel speaker. In comparison with FIG. 8A, when it is determined that sound quality is more important than a sound image of a channel audio signal, the audio providing apparatus 100 according to the present embodiment may down-mix only a channel audio signal to a 2D signal and render an object audio signal into a virtual 3D signal.

[00139] FIG. 9 is a flowchart for describing an audio signal providing method according to an exemplary embodiment of the present invention.

[00140] First, the audio providing apparatus 100 receives an audio signal in operation S910. In this case, the audio signal may include a channel audio signal having a first channel number and an object audio signal.

[00141] In operation S920, the audio providing apparatus 100 separates the received audio signal. In detail, the audio providing apparatus 100 may de-multiplex the received audio signal into the channel audio signal and the object audio signal.

[00142] In operation S930, the audio providing apparatus 100 renders the object audio signal. In detail, as described above with reference to FIGS. 2 to 5B, the audio providing apparatus 100 may two-dimensionally or three-dimensionally render the object audio signal. Also, as described above with reference to FIGS. 6 to 7B, the audio providing apparatus 100 may render the object audio signal into a virtual 3D audio signal.

[00143] In operation S940, the audio providing apparatus 100 renders the channel audio signal having the first channel number into a second channel number. In this case, the audio providing apparatus 100 may down-mix or up-mix the received channel audio signal to perform rendering. Also, the audio providing apparatus 100 may perform rendering while maintaining the number of channels of the received channel audio signal.

[00144] In operation S950, the audio providing apparatus 100 mixes the rendered object audio signal with a channel audio signal having the second channel number. In detail, as illustrated in FIGS. 8A to 8G, the audio providing apparatus 100 may mix the rendered object audio signal with the channel audio signal.

[00145] In operation S960, the audio providing apparatus 100 outputs a mixed audio signal.

[00146] According to the above-described audio providing method, the audio providing apparatus 100 reproduces audio signals having various formats to be optimal for an audio system space.

[00147] Hereinafter, another exemplary embodiment of the present invention will be described with reference to FIG. 10. FIG. 10 is a block diagram illustrating a configuration of an audio providing apparatus 1000 according to another exemplary embodiment of the present invention. As illustrated in FIG. 10, the audio providing apparatus 1000 includes an input unit 1010, a de-multiplexer 1020, an audio signal decoding unit 1030, an additional information decoding unit 1040, a rendering unit 1050, a user input unit 1060, an interface 1070, and an output unit 1080.

[00148] The input unit 1010 receives a compressed audio signal. In this case, the compressed audio signal may include additional information as well as a compressed-type audio signal which includes a channel audio signal and an object audio signal.

[00149] The de-multiplexer 1020 may separate the compressed audio signal into the audio signal and the additional information, output the audio signal to the audio signal decoding unit 1030, and output the additional information to the additional information decoding unit 1040.

[00150] The audio signal decoding unit 1030 decompresses the compressed-type audio signal and outputs the decompressed audio signal to the rendering unit 1050. The audio signal includes a multi-channel channel audio signal and an object audio signal. In this case, the multi-channel channel audio signal may be an audio signal such as background sound and background music, and the object audio signal may be an audio signal, such as voice, gunfire, etc., for a specific object.

[00151] The additional information decoding unit 1040 decodes additional information regarding the received audio signal. In this case, the additional information regarding the received audio signal may include various pieces of information such as the number of channels, a length, a gain value, a panning gain, a position, and an angle of the received audio signal.

[00152] The rendering unit 1050 may perform rendering based on the received additional information and audio signal. In this case, the rendering unit 1050 may perform rendering according to a user command input to the user input unit 1060 by using various methods described above with reference to FIGS. 2 to 8G. For example, when the received audio signal is a 7.1-channel audio signal and a speaker layout of the audio providing apparatus 1000 is 5.1 channel, the rendering unit 1050 may down-mix the 7.1-channel audio signal to a 2D 5.1-channel audio signal and down-mix the 7.1-channel audio signal to a 3D 5.1-channel audio signal according to the user command which is input through the user input unit 1060. Also, the rendering unit 1050 may render the channel audio signal into a 2D signal and render the object audio signal into a virtual 3D signal according to the user command which is input through the user input unit 1060.

[00153] Moreover, the rendering unit 1050 may directly output the rendered audio signal through the output unit 1080 according to the user command and the speaker layout, but may transmit the audio signal and the additional information to an external device through the interface 1070. Particularly, when the audio providing apparatus 1000 has a speaker layout exceeding 7.1 channel, the rendering unit 1050 may transmit at least one of the audio signal and the additional information to the external device through the interface 1070. In this case, the interface 1070 may be implemented as a digital interface such as an HDMI interface or the like. The external device may perform rendering by using the received audio signal and additional information and output a rendered audio signal.

[00154] However, as described above, the rendering unit 1050 transmitting the audio signal and the additional information to the external device is merely an exemplary embodiment. The rendering unit 1050 may render the audio signal by using the audio signal and the additional information and output the rendered audio signal.

[00155] The object audio signal according to an exemplary embodiment of the present invention may include metadata including an identification (ID), type information, or priority information. For example, the object audio signal may include information indicating whether a type of the object audio signal is dialogue or commentary. Also, when the audio signal is a broadcast audio signal, the object audio signal may include information indicating whether a type of the object audio signal is a first anchor, a second anchor, a first caster, a second caster, or background sound. Also, when the audio signal is a music audio signal, the object audio signal may include information indicating whether a type of the object audio signal is a first vocalist, a second vocalist, a first instrument sound, or a second instrument sound. Also, when the audio signal is a game audio signal, the object audio signal may

include information indicating whether a type of the object audio signal is a first sound effect or a second sound effect.

[00156] The rendering unit 1050 may analyze the metadata included in the above-described object audio signal and render the object audio signal according to a priority of the object audio signal.

[00157] Moreover, the rendering unit 1050 may remove a specific object audio signal according to a user's selection. For example, when the audio signal is an audio signal for sports, the audio providing apparatus 1000 may display a user interface (UI) that shows a type of a currently input object audio signal to the user. In this case, the object audio signal may include a caster's voice, voiceover, shouting voice, etc. When a user command for removing a caster's voice from among a plurality of object audio signals is input through the user input unit 1060, the rendering unit 1050 may remove the caster's voice from among the plurality of object audio signals and perform rendering by using the other object audio signals.

[00158] Moreover, the rendering unit 1050 may raise or lower volume for a specific object audio signal according to a user's selection. For example, when the audio signal is an audio signal included in movie content, the audio providing apparatus 1000 may display a UI that shows a type of a currently input object audio signal to the user. In this case, the object audio signal may include a first protagonist's voice, a second protagonist's voice, bomb sound, airplane sound, etc. When a user command for raising the volume of the first protagonist's voice and the second protagonist's voice and lowering the volume of the bomb sound and the airplane sound among a plurality of object audio signals is input through the user input unit 1060, the rendering unit 1050 may raise the volume of the first protagonist's voice and the second protagonist's voice and lower the volume of the bomb sound and the airplane sound.

[00159] According to the above-described exemplary embodiments, a user manipulates a desired audio signal, and thus, an audio environment that is suitable for the user is established.

[00160] The audio providing method according to various exemplary embodiments may be implemented as a program and may be provided to a display apparatus or an input apparatus. Particularly, a program including a method of controlling a display apparatus may be stored in a non-transitory computer-readable recording medium and provided.

[00161] The non-transitory computer-readable recording medium denotes a medium that semi-permanently stores data and is readable by a device, instead of a medium that stores data for a short time like registers, caches, and a memories. In detail, various applications or

programs may be stored in a non-transitory computer-readable recording medium such as a CD, a DVD, a hard disk, a blue-ray disk, a USB memory, a memory card, or ROM.

[00162] While the inventive concept has been particularly shown and described with reference to exemplary embodiments thereof, it will be understood that various changes in form and details may be made therein without departing from the spirit and scope of the following claims.

CLAIMS

1. An audio providing apparatus comprising:

an object rendering unit that renders an object audio signal based on geometric information regarding the object audio signal;

a channel rendering unit that renders an audio signal having a first channel number into an audio signal having a second channel number; and

a mixing unit that mixes the rendered object audio signal with the audio signal having the second channel number.

2. The audio providing apparatus of claim 1, wherein the object rendering unit comprises:

a geometric information analyzer that converts the geometric information regarding the object audio signal into three-dimensional (3D) coordinate information;

a distance controller that generates distance control information, based on the 3D coordinate information;

a depth controller that generates depth control information, based on the 3D coordinate information;

a localizer that generates localization information for localizing the object audio signal, based on the 3D coordinate information; and

a renderer that renders the object audio signal, based on the distance control information, the depth control information, and the localization information.

3. The audio providing apparatus of claim 2, wherein

the distance controller acquires a distance gain of the object audio signal,

as a distance of the object audio signal increases, the distance controller decreases the distance gain of the object audio signal, and

as the distance of the object audio signal decreases, the distance controller increases the distance gain of the object audio signal.

4. The audio providing apparatus of claim 3, wherein

the depth controller acquires a depth gain, based on a horizontal projection distance of the object audio signal, and

the depth gain is expressed as a sum of a negative vector and a positive vector or is expressed as a sum of the negative vector and a null vector.

5. The audio providing apparatus of claim 4, wherein the localizer acquires a panning gain for localizing the object audio signal according to a speaker layout of the audio providing apparatus.

6. The audio providing apparatus of claim 5, wherein the renderer renders the object audio signal into a multi-channel, based on the depth gain, the panning gain, and the distance gain of the object audio signal.

7. The audio providing apparatus of claim 2, wherein when the object audio signal is plurality, the object rendering unit acquires a phase difference between a plurality of object audio signals having a correlation among the plurality of object audio signals and moves one of the plurality of object audio signals by the acquired phase difference to combine the plurality of object audio signals.

8. The audio providing apparatus of claim 1, wherein when the audio providing apparatus reproduces audio by using a plurality of speakers having the same elevation,

the object rendering unit comprises:

a virtual filter that corrects spectral characteristics of the object audio signal and adds virtual elevation information to the object audio signal; and

a virtual renderer that renders the object audio signal, based on the virtual elevation information supplied by the virtual filter.

9. The audio providing apparatus of claim 8, wherein the virtual filter has a tree structure consisting of a plurality of stages.

10. The audio providing apparatus of claim 1, wherein

when a layout of the audio signal having the first channel number is a two-dimensional (2D) layout, the channel rendering unit up-mixes the audio signal having the first channel number to the audio signal having the second channel number which is greater than the first channel number, and

a layout of the audio signal having the second channel number is a three-dimensional (3D) layout having elevation information which differs from elevation information regarding the audio signal having the first channel number.

11. The audio providing apparatus of claim 1, wherein

when a layout of the audio signal having the first channel number is a three-dimensional (3D) layout, the channel rendering unit down-mixes the audio signal having the first channel number to the audio signal having the second channel number which is less than the first channel number, and

a layout of the audio signal having the second channel number is a two-dimensional (2D) layout where a plurality of channels have the same elevation component.

12. The audio providing apparatus of claim 1, wherein at least one selected from the object audio signal and the audio signal having the first channel number comprises information for determining whether to perform virtual three-dimensional (3D) rendering on a specific frame.

13. The audio providing apparatus of claim 1, wherein the channel rendering unit acquires a phase difference between a plurality of audio signals having a correlation in an operation of rendering the audio signal having the first channel number into the audio signal having the second channel number, and moves one of the plurality of audio signals by the acquired phase difference to combine the plurality of audio signals.

14. The audio providing apparatus of claim 1, wherein the mixing unit acquires a phase difference between a plurality of audio signals having a correlation while mixing the rendered object audio signal with the audio signal having the second channel number, and moves one of the plurality of audio signals by the acquired phase difference to combine the plurality of audio signals.

15. The audio providing apparatus of claim 1, wherein the object audio signal comprises at least one of an identification (ID) and type information regarding the object audio signal for enabling a user to select the object audio signal.

16. An audio providing method comprising:

rendering an object audio signal based on geometric information regarding the object audio signal;

rendering an audio signal having a first channel number into an audio signal having a second channel number; and

mixing the rendered object audio signal with the audio signal having the second channel number.

17. The audio providing method of claim 16, wherein the rendering of the object audio signal comprises:

converting the geometric information regarding the object audio signal into three-dimensional (3D) coordinate information;

generating distance control information, based on the 3D coordinate information;

generating depth control information, based on the 3D coordinate information;

generating localization information for localizing the object audio signal, based on the 3D coordinate information; and

rendering the object audio signal, based on the distance control information, the depth control information, and the localization information.

18. The audio providing method of claim 17, wherein the generating of the distance control information comprises:

acquiring a distance gain of the object audio signal;

decreasing the distance gain of the object audio signal as a distance of the object audio signal increases; and

increasing the distance gain of the object audio signal as the distance of the object audio signal decreases.

19. The audio providing method of claim 18, wherein

the generating of the depth control information comprises acquiring a depth gain, based on a horizontal projection distance of the object audio signal, and

the depth gain is expressed as a sum of a negative vector and a positive vector or is expressed as a sum of the negative vector and a null vector.

20. The audio providing method of claim 19, wherein the generating of the localization information comprises acquiring a panning gain for localizing the object audio signal according to a speaker layout of an audio providing apparatus.

21. The audio providing method of claim 20, wherein the rendering comprises rendering the object audio signal to a multi-channel, based on the depth gain, the panning gain, and the distance gain of the object audio signal.

22. The audio providing method of claim 17, wherein the rendering of the object audio signal comprises:

when the object audio signal is plurality,

acquiring a phase difference between plurality of object audio signals having a correlation among the plurality of object audio signals; and

moving one of the plurality of object audio signals by the acquired phase difference to combine the plurality of object audio signals.

23. The audio providing method of claim 16, wherein when an audio providing apparatus reproduces audio by using a plurality of speakers having the same elevation,

the rendering of the object audio signal comprises:

correcting spectral characteristics of the object audio signal and adding virtual elevation information to the object audio signal; and

rendering the object audio signal, based on the virtual elevation information supplied by the virtual filter.

24. The audio providing method of claim 23, wherein the acquiring comprises virtual elevation information regarding the object audio signal by using a virtual filter which has a tree structure consisting of a plurality of stages.

25. The audio providing method of claim 16, wherein

the rendering of the audio signal having the first channel number into the audio signal having the second channel number comprises, when a layout of the audio signal having the first channel number is a two-dimensional (2D) layout, up-mixing the audio signal having the first channel number to the audio signal having the second channel number which is greater than the first channel number, and

a layout of the audio signal having the second channel number is a three-dimensional (3D) layout having elevation information which differs from elevation information regarding the audio signal having the first channel number.

26. The audio providing method of claim 16, wherein

the rendering of the audio signal having the first channel number to the audio signal having the second channel number comprises, when a layout of the audio signal having the first channel number is a three-dimensional (3D) layout, down-mixing the audio signal having the first channel number to the audio signal having the second channel number which is less than the first channel number, and

a layout of the audio signal having the second channel number is a two-dimensional (2D) layout where a plurality of channels have the same elevation component.

27. The audio providing method of claim 16, wherein at least one selected from the object audio signal and the audio signal having the first channel number comprises information for determining whether to perform virtual three-dimensional (3D) rendering on a specific frame.