

DESCRIPTION

ENCODING DEVICE AND METHOD, DECODING DEVICE AND METHOD,
AND PROGRAM

5 TECHNICAL FIELD

[0001]

The present technology relates to an encoding device and method, a decoding device and method, and a program, and more particularly, to an encoding device and method, a decoding device and method, and a program
10 capable of realizing sense-of-distance control based on intention of a content creator.

BACKGROUND ART

15 [0002]

In recent years, object-based audio technology has attracted attention.

[0003]

In object-based audio, data of an object audio is
20 configured by a waveform signal with respect to an audio object and metadata indicating localization information of the audio object represented by a relative position from a listening position serving as a predetermined reference.

25 [0004]

Then, the waveform signal of the audio object is rendered into signals of a desired number of channels by, for example, vector based amplitude panning (VBAP) on the basis of the metadata and reproduced (see, for example,
30 Non Patent Document 1 and Non Patent Document 2).

[0005]

Furthermore, as a technology related to the object-based audio, for example, a technology for realizing audio reproduction with a higher degree of freedom in which a user can designate an arbitrary listening position has also been proposed (see, for example, Patent Document 1).

[0006]

In this technology, the position information of the audio object is corrected according to the listening position, and gain control or filter processing is performed according to a change in a distance from the listening position to the audio object, so that a change in frequency characteristics or volume accompanying a change in the listening position of the user, that is, a sense of distance to the audio object is reproduced.

CITATION LIST

NON PATENT DOCUMENT

[0007]

Non Patent Document 1: ISO/IEC 23008-3 Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 3:3D audio

Non Patent Document 2: Ville Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning", Journal of AES, vol. 45, no. 6, pp. 456-466, 1997

PATENT DOCUMENT

[0008]

Patent Document 1: WO 2015/107926 A

SUMMARY OF THE INVENTION

PROBLEMS TO BE SOLVED BY THE INVENTION

[0009]

However, in the above-described technology, the gain control and the filter processing for reproducing the change in frequency characteristics and volume
5 corresponding to the distance from the listening position to the audio object are predetermined.

[0010]

Therefore, when a content creator desires to reproduce a sense of distance based on the change in
10 frequency characteristics and volume in a different way therefrom, such a sense of distance cannot be reproduced. That is, it is not possible to realize sense-of-distance control based on the intention of the content creator.

[0011]

15 The present technology has been made in view of such a situation, and an object thereof is to realize the sense-of-distance control based on the intention of the content creator.

20 SOLUTIONS TO PROBLEMS

[0012]

An encoding device according to a first aspect of the present technology includes: an object encoding unit that encodes audio data of an object; a metadata encoding
25 unit that encodes metadata including position information of the object; a sense-of-distance control information determination unit that determines sense-of-distance control information for sense-of-distance control processing to be performed on the audio data; a sense-of-
30 distance control information encoding unit that encodes the sense-of-distance control information; and a

multiplexer that multiplexes the coded audio data, the coded metadata, and the coded sense-of-distance control information to generate coded data.

[0013]

5 An encoding method or a program according to the first aspect of the present technology includes the steps of: encoding audio data of an object; encoding metadata including position information of the object; determining sense-of-distance control information for sense-of-
10 distance control processing to be performed on the audio data; encoding the sense-of-distance control information; and

 multiplexing the coded audio data, the coded metadata, and the coded sense-of-distance control
15 information to generate coded data.

[0014]

 In the first aspect of the present technology, the audio data of the object is encoded, the metadata including the position information of the object is
20 encoded, the sense-of-distance control information for the sense-of-distance control processing to be performed on the audio data is determined, the sense-of-distance control information is encoded, and the coded audio data, the coded metadata, and the coded sense-of-distance
25 control information are multiplexed to generate the coded data.

[0015]

 A decoding device according to a second aspect of the present technology includes: a demultiplexer that
30 demultiplexes coded data to extract coded audio data of an object, coded metadata including position information

of the object, and coded sense-of-distance control
information for sense-of-distance control processing to
be performed on the audio data; an object decoding unit
that decodes the coded audio data; a metadata decoding
5 unit that decodes the coded metadata; a sense-of-distance
control information decoding unit that decodes the coded
sense-of-distance control information; a sense-of-
distance control processing unit that performs the sense-
of-distance control processing on the audio data of the
10 object on the basis of the sense-of-distance control
information; and a rendering processing unit that
performs rendering processing on the basis of the audio
data obtained by the sense-of-distance control processing
and the metadata to generate reproduction audio data for
15 reproducing a sound of the object.

[0016]

A decoding method or a program according to the
second aspect of the present technology includes the
steps of: demultiplexing coded data to extract coded
20 audio data of an object, coded metadata including
position information of the object, and coded sense-of-
distance control information for sense-of-distance
control processing to be performed on the audio data;
decoding the coded audio data; decoding the coded
25 metadata; decoding the coded sense-of-distance control
information; performing the sense-of-distance control
processing on the audio data of the object on the basis
of the sense-of-distance control information; and
performing rendering processing on the basis of the audio
30 data obtained by the sense-of-distance control processing
and the metadata to generate reproduction audio data for

reproducing a sound of the object.

[0017]

In the second aspect of the present technology, the coded data is demultiplexed to extract the coded audio data of the object, the coded metadata including the position information of the object, and the coded sense-of-distance control information for the sense-of-distance control processing to be performed on the audio data, the coded audio data is decoded, the coded metadata is decoded, the coded sense-of-distance control information is decoded, the sense-of-distance control processing is performed on the audio data of the object on the basis of the sense-of-distance control information, and the rendering processing is performed on the basis of the audio data obtained by the sense-of-distance control processing and the metadata to generate the reproduction audio data for reproducing the sound of the object.

BRIEF DESCRIPTION OF DRAWINGS

[0018]

Fig. 1 is a diagram illustrating a configuration example of an encoding device.

Fig. 2 is a diagram illustrating a configuration example of a decoding device.

Fig. 3 is a diagram illustrating a configuration example of a sense-of-distance control processing unit.

Fig. 4 is a diagram illustrating a configuration example of a reverb processing unit.

Fig. 5 is a diagram for describing an example of a control rule of gain control processing.

Fig. 6 is a diagram for describing an example of a

control rule of filter processing by a high-shelf filter.

Fig. 7 is a diagram for describing an example of a control rule of filter processing by a low-shelf filter.

Fig. 8 is a diagram for describing an example of a
5 control rule of reverb processing.

Fig. 9 is a diagram for describing generation of a wet component.

Fig. 10 is a diagram for describing the generation of the wet component.

10 Fig. 11 is a diagram illustrating an example of sense-of-distance control information.

Fig. 12 is a diagram illustrating an example of parameter configuration information of the gain control.

Fig. 13 is a diagram illustrating an example of
15 parameter configuration information of the filter processing.

Fig. 14 is a diagram illustrating an example of parameter configuration information of the reverb processing.

20 Fig. 15 is a flowchart for describing an encoding process.

Fig. 16 is a flowchart for describing a decoding process.

Fig. 17 is a diagram illustrating an example of a
25 table and a function for obtaining a gain value.

Fig. 18 is a diagram illustrating an example of the parameter configuration information of the gain control.

Fig. 19 is a diagram illustrating an example of the sense-of-distance control information.

30 Fig. 20 is a diagram illustrating an example of the sense-of-distance control information.

Fig. 21 is a diagram illustrating a configuration example of the sense-of-distance control processing unit.

Fig. 22 is a diagram illustrating an example of the sense-of-distance control information.

5 Fig. 23 is a diagram illustrating a configuration example of a computer.

MODE FOR CARRYING OUT THE INVENTION

[0019]

10 Hereinafter, embodiments to which the present technology is applied will be described with reference to the drawings.

[0020]

<First embodiment>

15 <Configuration example of encoding device>

The present technology relates to reproduction of audio content of object-based audio including sounds of one or more audio objects.

[0021]

20 Hereinafter, the audio object is also simply referred to as an object, and the audio content is also simply referred to as content.

[0022]

In the present technology, sense-of-distance control information for sense-of-distance control processing which is set by a content creator and reproduces a sense of distance from a listening position to the object is transmitted to a decoding side together with the audio data of the object. Therefore, it is possible to realize sense-of-distance control based on an intention of the content creator.

25

30

[0023]

Here, the sense-of-distance control processing is processing for reproducing a sense of distance from a listening position to an object when reproducing a sound of the object, that is, processing for adding the sense of distance to the sound of the object, and is signal processing realized by executing arbitrary one or more processing steps in combination.

[0024]

Specifically, for example, in the sense-of-distance control processing, gain control processing for audio data, filter processing for adding frequency characteristics and various acoustic effects, reverb processing, and the like are performed.

[0025]

Information for enabling the decoding side to reconfigure such sense-of-distance control processing is sense-of-distance control information, and the sense-of-distance control information includes configuration information and control rule information. In other words, the sense-of-distance control information includes the configuration information and the control rule information.

[0026]

For example, the configuration information configuring the sense-of-distance control information is information which is obtained by parameterizing the configuration of the sense-of-distance control processing set by the content creator and indicates one or more signal processing steps to be performed in combination to realize the sense-of-distance control processing.

[0027]

More specifically, the configuration information indicates the number of signal processing steps included in the sense-of-distance control processing, processing
5 executed in such signal processing, and the order of the processing.

[0028]

Note that, in a case where one or more signal processing steps configuring the sense-of-distance
10 control processing and the order of performing these signal processing steps are determined in advance, the sense-of-distance control information does not necessarily need to include the configuration information.

15 [0029]

Furthermore, the control rule information is information for obtaining a parameter which is obtained by parameterizing a control rule, which is set by the content creator, in each of the signal processing steps
20 configuring the sense-of-distance control processing and is used in each of the signal processing steps configuring the sense-of-distance control processing.

[0030]

More specifically, the control rule information
25 indicates the parameter which is used for each of the signal processing steps configuring the sense-of-distance control processing and the control rule in which the parameter changes according to the distance from the listening position to the object.

30 [0031]

On the encoding side, such sense-of-distance

control information and the audio data of each object are encoded and transmitted to the decoding side.

[0032]

Furthermore, on the decoding side, the sense-of-
5 distance control processing is reconfigured on the basis of the sense-of-distance control information, and the sense-of-distance control processing is performed on the audio data of each object.

[0033]

10 At this time, the parameter corresponding to the distance from the listening position to the object is determined on the basis of the control rule information included in the sense-of-distance control information, and the signal processing configuring the sense-of-
15 distance control processing is performed on the basis of the parameter.

[0034]

Then, 3D audio rendering processing is performed on the basis of the audio data obtained by the sense-of-
20 distance control processing, and reproduction audio data for reproducing the sound of the content, that is, the sound of the object is generated.

[0035]

Hereinafter, a more specific embodiment to which
25 the present technology is applied will be described.

[0036]

For example, a content reproduction system to which the present technology is applied includes an encoding device that encodes the audio data of each of one or more
30 objects included in content and the sense-of-distance control information to generate coded data, and a

decoding device that receives supply of the coded data to generate reproduction audio data.

[0037]

An encoding device configuring such a content reproduction system is configured as illustrated in Fig. 1, for example.

[0038]

An encoding device 11 illustrated in Fig. 1 includes an object encoding unit 21, a metadata encoding unit 22, a sense-of-distance control information determination unit 23, a sense-of-distance control information encoding unit 24, and a multiplexer 25.

[0039]

The audio data of each of one or more objects included in the content is supplied to the object encoding unit 21. The audio data is a waveform signal (audio signal) for reproducing the sound of the object.

[0040]

The object encoding unit 21 encodes the supplied audio data of each object, and supplies the resultant coded audio data to the multiplexer 25.

[0041]

The metadata of the audio data of each object is supplied to the metadata encoding unit 22.

25 [0042]

The metadata includes at least position information indicating an absolute position of the object in a space. The position information is coordinates indicating the position of the object in an absolute coordinate system, that is, for example, a three-dimensional orthogonal coordinate system based on a predetermined position in

the space. Furthermore, the metadata may include gain information or the like for performing gain control (gain correction) on the audio data of the object.

[0043]

5 The metadata encoding unit 22 encodes the supplied metadata of each object, and supplies the resultant coded metadata to the multiplexer 25.

[0044]

 The sense-of-distance control information
10 determination unit 23 determines the sense-of-distance control information according to a designation operation or the like by the user, and supplies the determined sense-of-distance control information to the sense-of-distance control information encoding unit 24.

15 [0045]

 For example, the sense-of-distance control information determination unit 23 acquires the configuration information and the control rule information designated by the user according to the
20 designation operation by the user, thereby determining the sense-of-distance control information including the configuration information and the control rule information.

[0046]

25 Furthermore, for example, the sense-of-distance control information determination unit 23 may determine the sense-of-distance control information on the basis of the audio data of each object of the content, information regarding the content such as a genre of the content,
30 information regarding a reproduction space of the content, and the like.

[0047]

Note that, in a case where each of the signal processing steps configuring the sense-of-distance control processing and the processing order of the signal processing steps are known on the decoding side, the configuration information may not be included in the sense-of-distance control information.

[0048]

The sense-of-distance control information encoding unit 24 encodes the sense-of-distance control information supplied from the sense-of-distance control information determination unit 23, and supplies the resultant coded sense-of-distance control information to the multiplexer 25.

[0049]

The multiplexer 25 multiplexes the coded audio data supplied from the object encoding unit 21, the coded metadata supplied from the metadata encoding unit 22, and the coded sense-of-distance control information supplied from the sense-of-distance control information encoding unit 24 to generate coded data (code string). The multiplexer 25 sends (transmits) the coded data obtained by the multiplexing to the decoding device via a communication network or the like.

[0050]

<Configuration example of decoding device>

Furthermore, the decoding device included in the content reproduction system is configured as illustrated in Fig. 2, for example.

[0051]

A decoding device 51 illustrated in Fig. 2 includes

a demultiplexer 61, an object decoding unit 62, a metadata decoding unit 63, a sense-of-distance control information decoding unit 64, a user interface 65, a distance calculation unit 66, a sense-of-distance control processing unit 67, and a 3D audio rendering processing unit 68.

[0052]

The demultiplexer 61 receives the coded data sent from the encoding device 11, and demultiplexes the received coded data to extract the coded audio data, the coded metadata, and the coded sense-of-distance control information from the coded data.

[0053]

The demultiplexer 61 supplies the coded audio data to the object decoding unit 62, supplies the coded metadata to the metadata decoding unit 63, and supplies the coded sense-of-distance control information to the sense-of-distance control information decoding unit 64.

[0054]

The object decoding unit 62 decodes the coded audio data supplied from the demultiplexer 61, and supplies the resultant audio data to the sense-of-distance control processing unit 67.

[0055]

The metadata decoding unit 63 decodes the coded metadata supplied from the demultiplexer 61, and supplies the resultant metadata to the sense-of-distance control processing unit 67 and the distance calculation unit 66.

[0056]

The sense-of-distance control information decoding unit 64 decodes the coded sense-of-distance control

information supplied from the demultiplexer 61, and supplies the resultant sense-of-distance control information to the sense-of-distance control processing unit 67.

5 [0057]

The user interface 65 supplies listening position information indicating the listening position designated by the user to the distance calculation unit 66, the sense-of-distance control processing unit 67, and the 3D
10 audio rendering processing unit 68, for example, according to an operation of the user or the like.

[0058]

Here, the listening position indicated by the listening position information is the absolute position
15 of a listener who listens to the sound of the content in the reproduction space. For example, the listening position information is coordinates indicating a listening position in the same absolute coordinate system as that of the position information of the object
20 included in the metadata.

[0059]

The distance calculation unit 66 calculates the distance from the listening position to the object for every object on the basis of the metadata supplied from
25 the metadata decoding unit 63 and the listening position information supplied from the user interface 65, and supplies distance information indicating the calculation result to the sense-of-distance control processing unit 67.

30 [0060]

On the basis of the metadata supplied from the

metadata decoding unit 63, the sense-of-distance control information supplied from the sense-of-distance control information decoding unit 64, the listening position information supplied from the user interface 65, and the distance information supplied from the distance calculation unit 66, the sense-of-distance control processing unit 67 performs the sense-of-distance control processing on the audio data supplied from the object decoding unit 62.

10 [0061]

At this time, the sense-of-distance control processing unit 67 obtains a parameter on the basis of the control rule information and the distance information, and performs the sense-of-distance control processing on the audio data on the basis of the obtained parameter.

[0062]

By such sense-of-distance control processing, the audio data of a dry component and the audio data of a wet component of the object are generated.

[0063]

Here, the audio data of the dry component is audio data, which is obtained by performing one or more processing steps on the audio data of the original object, such as a direct sound component of the object.

[0064]

The metadata of the original object, that is, the metadata output from the metadata decoding unit 63 is used as the metadata of the audio data of the dry component.

[0065]

Furthermore, the audio data of the wet component is audio data, which is obtained by performing one or more processing steps on the audio data of the original object, such as a reverberation component of the sound of the object.

[0066]

Therefore, it can be said that generating the audio data of the wet component is generating the audio data of a new object related to the original object.

[0067]

In the sense-of-distance control processing unit 67, necessary data of the metadata of the original object, the control rule information, the distance information, and the listening position information is appropriately used to generate the metadata of the audio data of the wet component.

[0068]

This metadata includes position information indicating at least the position of the object of the wet component.

[0069]

For example, the position information of the object of the wet component is polar coordinates expressed by an angle in a horizontal direction (horizontal angle) indicating the position of the object as viewed from the listener in the reproduction space, an angle in a height direction (vertical angle), and a radius indicating a distance from the listening position to the object.

[0070]

The sense-of-distance control processing unit 67 supplies the audio data and the metadata of the dry

component and the audio data and the metadata of the wet component to the 3D audio rendering processing unit 68.

[0071]

The 3D audio rendering processing unit 68 performs
5 the 3D audio rendering processing on the basis of the audio data and the metadata supplied from the sense-of-distance control processing unit 67 and the listening position information supplied from the user interface 65, and generates reproduction audio data.

10 [0072]

For example, the 3D audio rendering processing unit 68 performs VBAP, which is rendering processing in a polar coordinate system, or the like as the 3D audio rendering process.

15 [0073]

In this case, for the audio data of the dry component, the 3D audio rendering processing unit 68 generates position information expressed by polar coordinates on the basis of the position information
20 included in the metadata of the object of the dry component and the listening position information, and uses the obtained position information for the rendering process. This position information is polar coordinates expressed by a horizontal angle indicating the relative
25 position of the object as viewed from the listener, a vertical angle, and a radius indicating the distance from the listening position to the object.

[0074]

By such rendering processing, for example,
30 multichannel reproduction audio data including audio data of channels corresponding to a plurality of speakers

configuring a speaker system serving as an output destination is generated.

[0075]

The 3D audio rendering processing unit 68 outputs
5 the reproduction audio data obtained by the rendering processing to the subsequent stage.

[0076]

<Configuration example of sense-of-distance control processing unit>

10 Next, a specific configuration example of the sense-of-distance control processing unit 67 of the decoding device 51 will be described.

[0077]

Note that, here, an example will be described in
15 which the configuration of the sense-of-distance control processing unit 67, that is, one or more processing steps configuring the sense-of-distance control processing and the order of the processing are determined in advance.

[0078]

20 In such a case, the sense-of-distance control processing unit 67 is configured as illustrated in Fig. 3, for example.

[0079]

The sense-of-distance control processing unit 67
25 illustrated in Fig. 3 includes a gain control unit 101, a high-shelf filter processing unit 102, a low-shelf filter processing unit 103, and a reverb processing unit 104.

[0080]

In this example, gain control processing, filter
30 processing by a high-shelf filter, filter processing by a low-shelf filter, and reverb processing are sequentially

executed as the sense-of-distance control processing.

[0081]

The gain control unit 101 performs gain control on the audio data of the object supplied from the object
5 decoding unit 62 with the parameter (gain value) corresponding to the control rule information and the distance information, and supplies the resultant audio data to the high-shelf filter processing unit 102.

[0082]

10 The high-shelf filter processing unit 102 performs filter processing on the audio data supplied from the gain control unit 101 by the high-shelf filter determined by the parameter corresponding to the control rule information and the distance information, and supplies
15 the resultant audio data to the low-shelf filter processing unit 103.

[0083]

In the filter processing by the high-shelf filter, the high-frequency gain of the audio data is suppressed
20 according to the distance from the listening position to the object.

[0084]

The low-shelf filter processing unit 103 performs filter processing on the audio data supplied from the
25 high-shelf filter processing unit 102 by the low-shelf filter determined by the parameter corresponding to the control rule information and the distance information.

[0085]

In the filter processing by the low-shelf filter,
30 the low frequency of the audio data is boosted (emphasized) according to the distance from the listening

position to the object.

[0086]

The low-shelf filter processing unit 103 supplies the audio data obtained by the filter processing to the
5 3D audio rendering processing unit 68 and the reverb processing unit 104.

[0087]

Here, the audio data output from the low-shelf filter processing unit 103 is the audio data of the
10 original object described above, that is, the audio data of the dry component of the object.

[0088]

The reverb processing unit 104 performs reverb processing on the audio data supplied from the low-shelf
15 filter processing unit 103 with the parameter (gain) corresponding to the control rule information and the distance information, and supplies the resultant audio data to the 3D audio rendering processing unit 68.

[0089]

20 Here, the audio data output from the reverb processing unit 104 is the audio data of the wet component which is the reverberation component or the like of the original object described above. In other words, the audio data is the audio data of the object of
25 the wet component.

[0090]

<Configuration example of reverb processing unit>

Furthermore, more specifically, the reverb processing unit 104 is configured, for example, as
30 illustrated in Fig. 4.

[0091]

In the example illustrated in Fig. 4, the reverb processing unit 104 includes a gain control unit 141, a delay generation unit 142, a comb filter group 143, an all-pass filter group 144, an addition unit 145, an addition unit 146, a delay generation unit 147, a comb filter group 148, an all-pass filter group 149, an addition unit 150, and an addition unit 151.

[0092]

In this example, audio data of stereo reverberation components, that is, two wet components positioned on the left and right of the original object is generated for the mono audio data by the reverb processing.

[0093]

The gain control unit 141 performs gain control processing (gain correction processing) based on the wet gain value obtained from the control rule information and the distance information on the dry component audio data supplied from the low-shelf filter processing unit 103, and supplies the resultant audio data to the delay generation unit 142 and the delay generation unit 147.

[0094]

The delay generation unit 142 delays the audio data supplied from the gain control unit 141 by holding the audio data for a certain period of time, and supplies the delayed audio data to the comb filter group 143.

[0095]

Furthermore, the delay generation unit 142 supplies, to the addition unit 145, two pieces of audio data which are obtained by delaying the audio data supplied from the gain control unit 141, have different delay amounts from the audio data supplied to the comb

filter group 143, and have different delay amounts from each other.

[0096]

5 The comb filter group 143 includes a plurality of comb filters, performs filter processing by the plurality of comb filters on the audio data supplied from the delay generation unit 142, and supplies the resultant audio data to the all-pass filter group 144.

[0097]

10 The all-pass filter group 144 includes a plurality of all-pass filters, performs filter processing by the plurality of all-pass filters on the audio data supplied from the comb filter group 143, and supplies the resultant audio data to the addition unit 146.

15 [0098]

The addition unit 145 adds the two pieces of audio data supplied from the delay generation unit 142 and supplies the resultant audio data to the addition unit 146.

20 [0099]

The addition unit 146 adds the audio data supplied from the all-pass filter group 144 and the audio data supplied from the addition unit 145, and supplies the resultant audio data of the wet component to the 3D audio rendering processing unit 68.

[0100]

30 The delay generation unit 147 delays the audio data supplied from the gain control unit 141 by holding the audio data for a certain period of time, and supplies the delayed audio data to the comb filter group 148.

[0101]

Furthermore, the delay generation unit 147 supplies, to the addition unit 150, two pieces of audio data which are obtained by delaying the audio data supplied from the gain control unit 141, have different
5 delay amounts from the audio data supplied to the comb filter group 148, and have different delay amounts from each other.

[0102]

The comb filter group 148 includes a plurality of
10 comb filters, performs filter processing by the plurality of comb filters on the audio data supplied from the delay generation unit 147, and supplies the resultant audio data to the all-pass filter group 149.

[0103]

15 The all-pass filter group 149 includes a plurality of all-pass filters, performs filter processing by the plurality of all-pass filters on the audio data supplied from the comb filter group 148, and supplies the resultant audio data to the addition unit 151.

20 [0104]

The addition unit 150 adds the two pieces of audio data supplied from the delay generation unit 147 and supplies the resultant audio data to the addition unit 151.

25 [0105]

The addition unit 151 adds the audio data supplied from the all-pass filter group 149 and the audio data supplied from the addition unit 150, and supplies the resultant audio data of the wet component to the 3D audio
30 rendering processing unit 68.

[0106]

Note that, although the example in which the stereo (two) wet components are generated for one object has been described here, one wet component may be generated for one object, or three or more wet components may be generated. Furthermore, the configuration of the reverb processing unit 104 is not limited to the configuration illustrated in Fig. 4, and may be any other configuration.

[0107]

10 <Regarding control rule of parameter>

As described above, in each processing block configuring the sense-of-distance control processing unit 67, the parameters used for the processing in the processing blocks, that is, the characteristics of the processing change according to the distance from the listening position to the object.

[0108]

Here, an example of the parameter corresponding to the distance from the listening position to the object, that is, an example of a control rule of the parameter will be described.

[0109]

For example, the gain control unit 101 determines the gain value used for the gain control processing as the parameter corresponding to the distance from the listening position to the object.

[0110]

In this case, the gain value changes according to the distance from the listening position to the object as illustrated in Fig. 5, for example.

[0111]

For example, a portion indicated by an arrow Q11 indicates a change in the gain value corresponding to the distance. That is, a vertical axis represents the gain value as a parameter, and a horizontal axis represents the distance from the listening position to the object.

5 [0112]

As indicated by a polygonal line L11, the gain value is 0.0 dB when a distance d from the listening position to the object is between a predetermined minimum value Min and D_0 , and when the distance d is between D_0 and D_1 , the gain value linearly decreases as the distance d increases. Furthermore, the gain value is -40.0 dB when the distance d is between D_1 and the predetermined maximum value Max.

10

15 [0113]

From this, in the example illustrated in Fig. 5, it can be seen that control is performed in which the gain of the audio data is suppressed as the distance d increases.

20 [0114]

As a specific example, for example, in a case where the distance d is 1 m ($= D_0$) or less, the gain value is set to 0.0 dB, and when the distance d is between 1 m and 100 m ($= D_1$), the gain value can be linearly changed to -40.0 dB as the distance d increases.

25

[0115]

Here, when a point at which the parameter changes is referred to as a control change point, in the example of Fig. 5, a point (position) at which the distance $d = D_0$ and a point at which the distance $d = D_1$ in the polygonal line L11 are control change points.

30

[0116]

In this case, for example, as indicated by an arrow Q12, when the gain value "0.0" at the distance $d = D_0$ and the gain value "-40.0" at the distance $d = D_1$

5 corresponding to the control change point are transmitted to the decoding device 51, the decoding device 51 can obtain the gain value at an arbitrary distance d .

[0117]

10 Furthermore, in the high-shelf filter processing unit 102, for example, as indicated by an arrow Q21 in Fig. 6, the filter processing is performed in which the gain in the high frequency band is suppressed as the distance d from the listening position to the object increases.

15 [0118]

Note that, in the portion indicated by the arrow Q21, the vertical axis represents the gain value as a parameter, and the horizontal axis represents the distance d from the listening position to the object.

20 [0119]

In particular, in this example, the high-shelf filter realized by the high-shelf filter processing unit 102 is determined by a cutoff frequency F_c , a Q value indicating a sharpness, and a gain value at the cutoff frequency F_c .

25 frequency F_c .

[0120]

In other words, in the high-shelf filter processing unit 102, the filter processing is performed by the high-shelf filter determined by the cutoff frequency F_c , the Q value, and the gain value which are parameters.

30 value, and the gain value which are parameters.

[0121]

A polygonal line L21 in the portion indicated by the arrow Q21 indicates the gain value at the cutoff frequency F_c determined with respect to the distance d .
[0122]

5 In this example, the gain value is 0.0 dB when the distance d is between the minimum value Min and D_0 , and when the distance d is between D_0 and D_1 , the gain value linearly decreases as the distance d increases.
[0123]

10 Furthermore, when the distance d is between D_1 and D_2 , the gain value linearly decreases as the distance d increases, and similarly, when the distance d is between D_2 and D_3 and the distance d is between D_3 and D_4 , the gain value linearly decreases as the distance d
15 increases. Moreover, the gain value is -12.0 dB when the distance d is between D_4 and the maximum value Max .
[0124]

From this, in the example illustrated in Fig. 6, it can be seen that control is performed in which the gain
20 of the frequency component near the cutoff frequency F_c in the audio data is suppressed as the distance d increases.
[0125]

As a specific example, for example, in a case where
25 the distance d is 1 m ($= D_0$) or less, a frequency component of 6 kHz, which is the cutoff frequency F_c , or more can be set to pass through, and in a case where the distance d is between the distance d of 1 m and 100 m ($= D_4$), the frequency component of 6 kHz or more can be
30 changed to -12.0 dB as the distance d increases.
[0126]

Furthermore, in order to realize such a high-shelf filter in the decoding device 51, for example, as indicated by an arrow Q22, the cutoff frequency F_c , the Q value, and the gain value which are parameters are only required to be transmitted only for five control change points of the distances $d = D_0, D_1, D_2, D_3$, and D_4 .

[0127]

Note that, here, an example is described in which the cutoff frequency F_c is 6 kHz and the Q value is 2.0 regardless of the distance d , but these cutoff frequency F_c and Q value may also change according to the distance d .

[0128]

Moreover, in the low-shelf filter processing unit 103, for example, as indicated by an arrow Q31 in Fig. 7, the filter processing is performed in which the low-frequency gain is amplified as the distance d from the listening position to the object decreases.

[0129]

Note that, in the portion indicated by the arrow Q31, the vertical axis represents the gain value as a parameter, and the horizontal axis represents the distance d from the listening position to the object.

[0130]

In particular, in this example, the low-shelf filter realized by the low-shelf filter processing unit 103 is determined by the cutoff frequency F_c , the Q value indicating the sharpness, and the gain value at the cutoff frequency F_c .

[0131]

In other words, in the low-shelf filter processing

unit 103, the filter processing is performed by the low-shelf filter determined by the cutoff frequency F_c , the Q value, and the gain value which are parameters.

[0132]

5 A polygonal line L31 in the portion indicated by the arrow Q31 indicates the gain value at the cutoff frequency F_c determined with respect to the distance d .

[0133]

10 In this example, the gain value is 3.0 dB when the distance d is between the minimum value Min and D_0 , and when the distance d is between D_0 and D_1 , the gain value linearly decreases as the distance d increases. Furthermore, the gain value is 0.0 dB when the distance d is between D_1 and the maximum value Max .

15 [0134]

20 From this, in the example illustrated in Fig. 7, it can be seen that control is performed in which the gain of the frequency component near the cutoff frequency F_c in the audio data is amplified as the distance d decreases.

[0135]

25 As a specific example, for example, in a case where the distance d is 3 m ($= D_1$) or more, a frequency component of 200 Hz, which is the cutoff frequency F_c , or less can be set to pass through, and in a case where the distance d is between 3 m and 10 cm ($= D_0$), the frequency component of 200 Hz or less can be changed to +3.0 dB as the distance d decreases.

[0136]

30 Furthermore, in order to realize such a low-shelf filter in the decoding device 51, for example, as

indicated by an arrow Q32, the cutoff frequency F_c , the Q value, and the gain value which are parameters are only required to be transmitted only for two control change points of the distances $d = D_0$ and D_1 .

5 [0137]

Note that, here, an example is described in which the cutoff frequency F_c is 200 Hz and the Q value is 2.0 regardless of the distance d , but these cutoff frequency F_c and Q value may also change according to the distance

10 d .

[0138]

Moreover, in the reverb processing unit 104, for example, as indicated by an arrow Q41 in Fig. 8, the reverb processing is performed in which the gain (wet gain value) of the wet component increases as the distance d from the listening position to the object increases.

15

[0139]

In other words, control is performed in which the proportion of the wet component (reverberation component) generated by the reverb processing to the dry component increases as the distance d increases. Note that the wet gain value here is, for example, a gain value used in gain control in the gain control unit 141 illustrated in

20

25

Fig. 4.

[0140]

In the portion indicated by the arrow Q41, the vertical axis represents the wet gain value as a parameter, and the horizontal axis represents the distance d from the listening position to the object. Furthermore, a polygonal line L41 indicates the wet gain

30

value determined for the distance d .

[0141]

As indicated by the polygonal line L41, the wet gain value is negative infinity ($-\text{Inf dB}$) when the distance d from the listening position to the object is between the minimum value Min and D_0 , and when the distance d is between D_0 and D_1 , the wet gain value linearly increases as the distance d increases.

Furthermore, the wet gain value is -3.0 dB when the distance d is between D_1 and the maximum value Max .

[0142]

From this, in the example shown in Fig. 8, it can be seen that control is performed in which the wet component increases as the distance d increases.

[0143]

As a specific example, for example, in a case where the distance d is 1 m ($= D_0$) or less, the gain (wet gain value) of the wet component is set to $-\text{Inf dB}$, and in a case where the distance d is between the distance d of 1 m and 50 m ($= D_1$), the gain can be linearly changed to -3.0 dB as the distance d increases.

[0144]

Moreover, in order to realize such reverb processing in the decoding device 51, for example, as indicated by an arrow Q42, the wet gain value as a parameter is only required to be transmitted only for two control change points of the distances $d = D_0$ and D_1 .

[0145]

Furthermore, in the reverb processing, audio data of an arbitrary number of wet components (reverberation components) can be generated.

[0146]

Specifically, for example, as illustrated in Fig. 9, audio data of a stereo reverberation component can be generated for audio data of one object, that is, mono
5 audio data.

[0147]

In this example, an origin O of the XYZ coordinate system, which is a three-dimensional orthogonal coordinate system in the reproduction space, is the
10 listening position, and one object OB11 is arranged in the reproduction space.

[0148]

Now, the position of an arbitrary object in the reproduction space is represented by a horizontal angle
15 indicating the position in the horizontal direction viewed from the origin O and a vertical angle indicating the position in the vertical direction viewed from the origin O, and the position of the object OB11 is represented as (az,el) from a horizontal angle az and a
20 vertical angle el.

[0149]

Note that when a straight line connecting the origin O and the object OB11 is LN and a straight line obtained by projecting the straight line LN on the XZ
25 plane is LN', the horizontal angle az is an angle formed by the straight line LN 'and the Z axis. Furthermore, the vertical angle el is an angle formed by the straight line LN and the XZ plane.

[0150]

30 In the example of Fig. 9, for the object OB11, two objects OB12 and object OB13 are generated as wet

component objects.

[0151]

In particular, here, the object OB12 and the object OB13 are arranged at bilaterally symmetrical positions with respect to the object OB11 when viewed from the origin O.

[0152]

That is, the object OB12 and the object OB13 are arranged at positions shifted by 60 degrees to the left and right relatively from the object OB11, respectively.

[0153]

Therefore, the position of the object OB12 is a position (az+60,el) represented by the horizontal angle (az+60) and the vertical angle el, and the position of the object OB13 is a position (az-60,el) represented by the horizontal angle (az-60) and the vertical angle el.

[0154]

As described above, in a case where the wet components at bilaterally symmetrical positions with respect to the object OB11 are generated, the positions of the wet components can be designated by an offset angle with respect to the position of the object OB11. For example, in this example, an offset angle of ± 60 degrees of the horizontal angle is only required to be designated.

[0155]

Note that, although an example of generating two right and left wet components positioned on the right side and the left side with respect to one object has been described here, the number of wet components generated for one object may be any number, and for

example, wet components at upper, lower, left, and right positions may be generated.

[0156]

Furthermore, for example, in a case where
5 bilaterally symmetrical wet components are generated as illustrated in Fig. 9, the offset angle for designating the positions of the wet components may change according to the distance from the listening position to the object as illustrated in Fig. 10.

10 [0157]

In a portion indicated by an arrow Q51 in Fig. 10, the offset angle of the horizontal angle between the object OB12 and the object OB13 which are the wet components illustrated in Fig. 9 is illustrated.

15 [0158]

That is, in the portion indicated by the arrow Q51, the vertical axis represents the offset angle of the horizontal angle, and the horizontal axis represents the distance d from the listening position to the object
20 OB11.

[0159]

Furthermore, a polygonal line L51 indicates the offset angle of the object OB12 which is the left wet component determined for each distance d . In this
25 example, as the distance d decreases, the offset angle increases, and the object OB12 is arranged at a position farther away from the original object OB11.

[0160]

On the other hand, a polygonal line L52 indicates
30 the offset angle of the object OB13 which is the right wet component determined for each distance d . In this

example, as the distance d decreases, the offset angle decreases, and the object OB13 is arranged at a position farther away from the original object OB11.

[0161]

5 In a case where the offset angle changes according to the distance d in this manner, for example, as indicated by an arrow Q52, when the offset angle is transmitted to the decoding device 51 only for the control change point of the distance $d = D_0$, the wet
10 component can be generated at the position intended by the content creator.

[0162]

 As described above, when the sense-of-distance control processing is performed with the configuration
15 and the parameter corresponding to the distance d from the listening position to the object, the sense of distance can be appropriately reproduced. That is, it is possible to cause the listener to feel a sense of distance to the object.

20 [0163]

 At this time, when the content creator freely determines the parameter at each distance d , the sense-of-distance control based on the intention of the content creator can be realized.

25 [0164]

 Note that the control rule of the parameter corresponding to the distance d described above is merely an example, and by allowing the content creator to freely designate the control rule, it is possible to change how
30 to feel the sense of distance to the object.

[0165]

For example, since the change in sound with respect to the distance is different between outdoor and indoor, it is necessary to change the control rule depending on whether the space to be reproduced is outdoor or indoor.

5 [0166]

Therefore, for example, by determining (designating) the control rule according to the space where the content creator desires to reproduce with the content, the sense-of-distance control based on the intention of the content creator can be realized, and content reproduction with higher realistic feeling can be performed.

[0167]

Furthermore, in the sense-of-distance control processing unit 67, the parameter used for the sense-of-distance control processing can be further adjusted according to the reproduction environment of the content (reproduction audio data).

[0168]

20 Specifically, for example, the gain of the wet component used in the reverb processing, that is, the above-described wet gain value can be adjusted according to the reproduction environment of the content.

[0169]

25 When content is actually reproduced by a speaker or the like in the real space, reverberation of sound output from the speaker or the like occurs in the real space. At this time, how much reverberation occurs depends on the real space where the content is reproduced, that is, the reproduction environment.

30

[0170]

For example, when the content is reproduced in a highly reverberant environment, reverberation is further added to the sound of the reproduced content. Therefore, in a case where the content is actually reproduced, there is a case where the listener feels the sense of distance realized by the sense-of-distance control processing, that is, the sense of distance farther than the sense of distance intended by the content creator.

[0171]

Therefore, in a case where the reverberation in the reproduction environment is small, the sense-of-distance control processing is performed according to a preset control rule, that is, the control rule information, but in a case where the reverberation in the reproduction environment is relatively large, fine adjustment of the wet gain value determined according to the control rule may be performed.

[0172]

Specifically, for example, it is assumed that the user or the like operates the user interface 65 and inputs information regarding the reverberation of the reproduction environment such as type information, such as outdoors or indoors, of the reproduction environment and information indicating whether or not the reproduction environment is highly reverberant. In such a case, the user interface 65 supplies the information regarding reverberation of the reproduction environment input by the user or the like to the sense-of-distance control processing unit 67.

[0173]

Then, the sense-of-distance control processing unit

67 calculates the wet gain value on the basis of the control rule information, the distance information, and the information regarding the reverberation of the reproduction environment supplied from the user interface

5 65.

[0174]

Specifically, the sense-of-distance control processing unit 67 calculates the wet gain value on the basis of the control rule information and the distance
10 information, and performs determination processing on whether or not the reproduction environment is highly reverberant on the basis of the information regarding the reverberation of the reproduction environment.

[0175]

15 Here, for example, in a case where the information indicating that the reproduction environment is highly reverberant or the type information indicating a highly reverberant reproduction environment is supplied as the information regarding the reverberation of the
20 reproduction environment, it is determined that the reproduction environment is highly reverberant.

[0176]

Then, in a case where it is determined that the reproduction environment is not highly reverberant, that
25 is, the reproduction environment is less reverberant, the sense-of-distance control processing unit 67 supplies the calculated wet gain value to the reverb processing unit 104 as a final wet gain value.

[0177]

30 On the other hand, in a case where it is determined that the reproduction environment is highly reverberant,

the sense-of-distance control processing unit 67 corrects (adjusts) the calculated wet gain value with a predetermined correction value such as -6 dB, and supplies the corrected wet gain value to the reverb processing unit 104 as the final wet gain value.

5 [0178]

Note that the wet gain value correction value may be a predetermined value, or may be calculated by the sense-of-distance control processing unit 67 on the basis of the information regarding the reverberation of the reproduction environment, that is, the degree of reverberation in the reproduction environment.

10 [0179]

By adjusting the wet gain value according to the reproduction environment in this manner, it is possible to improve a deviation from the sense of distance intended by the content creator, the deviation being caused by the reproduction environment of the content.

15 [0180]

20 <Transmission of sense-of-distance control information>

Next, a transmission method of the sense-of-distance control information described above will be described.

25 [0181]

The sense-of-distance control information encoded by the sense-of-distance control information encoding unit 24 can have a configuration illustrated in Fig. 11, for example.

30 [0182]

In Fig. 11, "DistanceRender_Attn()" indicates

parameter configuration information indicating the control rule of the parameters used in the gain control unit 101.

[0183]

5 Furthermore, "DistanceRender_Filt()" indicates parameter configuration information indicating the control rule of the parameters used in the high-shelf filter processing unit 102 or the low-shelf filter processing unit 103.

10 [0184]

 Here, since the high-shelf filter and the low-shelf filter can be expressed by the same parameter configuration, the high-shelf filter and the low-shelf filter are described by the same syntax of the parameter configuration information DistanceRender_Filt().
15 Therefore, the sense-of-distance control information includes the parameter configuration information DistanceRender_Filt() of the high-shelf filter processing unit 102 and the parameter configuration information
20 DistanceRender_Filt() of the low-shelf filter processing unit 103.

[0185]

 Moreover, "DistanceRender_Revb()" indicates parameter configuration information indicating the
25 control rule of the parameter used in the reverb processing unit 104.

[0186]

 The parameter configuration information DistanceRender_Attn(), the parameter configuration
30 information DistanceRender_Filt(), and the parameter configuration information DistanceRender_Revb() included

in the sense-of-distance control information correspond to the control rule information.

[0187]

Furthermore, in the sense-of-distance control information illustrated in Fig. 11, parameter configuration information of four processing steps configuring the sense-of-distance control processing is arranged and stored in the order in which the processing steps are performed.

10 [0188]

Therefore, in the decoding device 51, the configuration of the sense-of-distance control processing unit 67 illustrated in Fig. 3 can be specified on the basis of the sense-of-distance control information. In other words, from the sense-of-distance control information illustrated in Fig. 11, it is possible to specify how many processing steps are included in the sense-of-distance control processing, what processing are performed in those processing steps, and in what order the processing is performed. Therefore, in this example, it can be said that the sense-of-distance control information substantially includes the configuration information.

[0189]

Moreover, the parameter configuration information DistanceRender_Attn(), the parameter configuration information DistanceRender_Filt(), and the parameter configuration information DistanceRender_Revb() illustrated in Fig. 11 are configured as illustrated in Figs. 12 to 14, for example.

30

[0190]

Fig. 12 is a diagram illustrating a configuration example, that is, a syntax example, of the parameter configuration information DistanceRender_Attn() of the gain control processing.

5 [0191]

In Fig. 12, "num_points" indicates the number of control change points of the parameter of the gain control processing. For example, in the example illustrated in Fig. 5, a point (position) at which the distance $d = D_0$ and a point at which the distance $d = D_1$ are control change points.

[0192]

In the example of Fig. 12, "distance[i]" indicating the distances d corresponding to the control change points and gain values "gain[i]" as a parameter at the distances d are included as many as the number of the control change points. When the distance distance[i] and the gain value gain[i] of each control change point is transmitted in this manner, the gain control illustrated in Fig. 5 can be realized in the decoding device 51.

[0193]

Fig. 13 is a diagram illustrating a configuration example, that is, a syntax example, of the parameter configuration information DistanceRender_Filt() of the filter processing.

[0194]

In Fig. 13, "filt_type" indicates an index indicating a filter type.

[0195]

30 For example, an index filt_type "0" indicates a low-shelf filter, an index filt_type "1" indicates a

high-shelf filter, and an index `filt_type` "2" indicates a peak filter.

[0196]

Furthermore, an index `filt_type` "3" indicates a
5 low-pass filter, and an index `filt_type` "4" indicates a high-pass filter.

[0197]

Therefore, for example, when the value of the index
`filt_type` is "0", it can be seen that the parameter
10 configuration information `DistanceRender_Filt()` includes information regarding a parameter for specifying the configuration of the low-shelf filter.

[0198]

Note that, in the example illustrated in Fig. 3,
15 the high-shelf filter and the low-shelf filter have been described as filter examples of the filter processing configuring the sense-of-distance control processing.

[0199]

On the other hand, in the example illustrated in
20 Fig. 13, the peak filter, the low-pass filter, the high-pass filter, and the like can also be used.

[0200]

Note that, as the filter for the filter processing configuring the sense-of-distance control processing,
25 only some of the low-shelf filter and the high-shelf filter, the peak filter, the low-pass filter, and the high-pass filter may be used, or other filters may be used.

[0201]

30 In the parameter configuration information `DistanceRender_Filt()` illustrated in Fig. 13, a region

after the index `filt_type` includes a parameter or the like for specifying the configuration of the filter indicated by the index `filt_type`.

[0202]

5 That is, "num_points" indicates the number of the control change points of the parameter of the filter processing.

[0203]

10 Furthermore, "distance[i]" indicating the distances `d` corresponding to the control change points, frequencies "freq[i]", Q values "Q[i]", and gain values "gain[i]" as parameters at the distances `d` are included as many as the number of the control change points indicated by the "num_points".

15 [0204]

 For example, when the index `filt_type` is "0" indicating a low-shelf filter, the frequency "freq[i]", the Q value "Q[i]", and the gain value "gain[i]", which are parameters, correspond to the cutoff frequency F_c , the Q value, and the gain value illustrated in Fig. 7.

[0205]

25 Note that the frequency `freq[i]` is a cutoff frequency when the filter type is the low-shelf filter and the high-shelf filter, the low-pass filter, or the high-pass filter, but is a center frequency when the filter type is the peak filter.

[0206]

30 As described above, when the distance `distance[i]`, the frequency "freq[i]", the Q value "Q[i]", and the gain value "gain[i]" of each control change point are transmitted, the high-shelf filter illustrated in Fig. 6

and the low-shelf filter illustrated in Fig. 7 can be realized in the decoding device 51.

[0207]

Fig. 14 is a diagram illustrating a configuration example, that is, a syntax example, of the parameter configuration information `DistanceRender_Revb()` of the reverb processing.

[0208]

In Fig. 14, "num_points" indicates the number of the control change points of the parameter of the reverb processing, and in this example, "distance[i]" indicating the distances *d* corresponding to those control change points and the wet gain values "wet_gain[i]" as the parameter at the distances *d* are included as many as the number of the control change points. The wet gain value wet_gain[i] corresponds to, for example, the wet gain value illustrated in Fig. 8.

[0209]

Furthermore, in Fig. 14, "num_wetobjs" indicates the number of generated wet components, that is, the number of objects of the wet components, and the offset angles indicating the positions of the wet components is stored as many as the number of the wet components.

[0210]

That is, "wet_azimuth_offset[i][j]" indicates the offset angle of the horizontal angle of a *j*-th wet component (object) at the distance distance[i] corresponding to an *i*-th control change point. The offset angle wet_azimuth_offset[i][j] corresponds to, for example, the offset angle of the horizontal angle illustrated in Fig. 10.

[0211]

Similarly, "wet_elevation_offset[i][j]" indicates the offset angle of the vertical angle of the j-th wet component at the distance distance[i] corresponding to the i-th control change point.

[0212]

Note that the number num_wetobjs of the generated wet components is determined by the reverb processing to be performed by the decoding device 51, and for example, the number num_wetobjs of the wet components is given from the outside.

[0213]

As described above, in the example of Fig. 14, the distance distance[i] and the wet gain value wet_gain[i] at each control change point, and the offset angle wet_azimuth_offset[i][j] and the offset angle wet_elevation_offset[i][j] of each wet component are transmitted to the decoding device 51.

[0214]

Therefore, in the decoding device 51, for example, the reverb processing unit 104 illustrated in Fig. 4 can be realized, and the audio data of the dry component and the audio data and the metadata of each wet component can be obtained.

[0215]

<Description of encoding process>

Next, an operation of the content reproduction system will be described.

[0216]

First, an encoding process performed by the encoding device 11 will be described with reference to a

flowchart in Fig. 15.

[0217]

In step S11, the object encoding unit 21 encodes the supplied audio data of each object, and supplies the
5 obtained coded audio data to the multiplexer 25.

[0218]

In step S12, the metadata encoding unit 22 encodes the supplied metadata of each object, and supplies the obtained coded metadata to the multiplexer 25.

10 [0219]

In step S13, the sense-of-distance control information determination unit 23 determines the sense-of-distance control information according to a designation operation or the like by the user, and
15 supplies the determined sense-of-distance control information to the sense-of-distance control information encoding unit 24.

[0220]

In step S14, the sense-of-distance control
20 information encoding unit 24 encodes the sense-of-distance control information supplied from the sense-of-distance control information determination unit 23, and supplies the obtained coded sense-of-distance control information to the multiplexer 25. Therefore, for
25 example, the sense-of-distance control information (coded sense-of-distance control information) illustrated in Fig. 11 is obtained and supplied to the multiplexer 25.

[0221]

In step S15, the multiplexer 25 multiplexes the
30 coded audio data from the object encoding unit 21, the coded metadata from the metadata encoding unit 22, and

the coded sense-of-distance control information from the sense-of-distance control information encoding unit 24 to generate coded data.

[0222]

5 In step S16, the multiplexer 25 sends the coded data obtained by the multiplexing to the decoding device 51 via a communication network or the like, and the encoding process ends.

[0223]

10 As described above, the encoding device 11 generates coded data including the sense-of-distance control information, and sends the coded data to the decoding device 51.

[0224]

15 As described above, by transmitting the sense-of-distance control information in addition to the audio data and the metadata of each object to the decoding device 51, it is possible to realize the sense-of-distance control based on the intention of the content
20 creator on the decoding device 51 side.

[0225]

<Description of decoding process>

Furthermore, when the encoding process described with reference to Fig. 15 is performed in the encoding
25 device 11, a decoding process is performed in the decoding device 51. Hereinafter, the decoding process by the decoding device 51 will be described with reference to a flowchart in Fig. 16.

[0226]

30 In step S41, the demultiplexer 61 receives the coded data sent from the encoding device 11.

[0227]

In step S42, the demultiplexer 61 demultiplexes the received coded data, and extracts the coded audio data, the coded metadata, and the coded sense-of-distance control information from the coded data.

[0228]

The demultiplexer 61 supplies the coded audio data to the object decoding unit 62, supplies the coded metadata to the metadata decoding unit 63, and supplies the coded sense-of-distance control information to the sense-of-distance control information decoding unit 64.

[0229]

In step S43, the object decoding unit 62 decodes the coded audio data supplied from the demultiplexer 61, and supplies the obtained audio data to the sense-of-distance control processing unit 67.

[0230]

In step S44, the metadata decoding unit 63 decodes the coded metadata supplied from the demultiplexer 61, and supplies the obtained metadata to the sense-of-distance control processing unit 67 and the distance calculation unit 66.

[0231]

In step S45, the sense-of-distance control information decoding unit 64 decodes the coded sense-of-distance control information supplied from the demultiplexer 61, and supplies the obtained sense-of-distance control information to the sense-of-distance control processing unit 67.

[0232]

In step S46, the distance calculation unit 66

calculates the distance from the listening position to the object on the basis of the metadata supplied from the metadata decoding unit 63 and the listening position information supplied from the user interface 65, and
5 supplies distance information indicating the calculation result to the sense-of-distance control processing unit 67. In step S46, the distance information is obtained for every object.

[0233]

10 In step S47, the sense-of-distance control processing unit 67 performs the sense-of-distance control processing on the basis of the audio data supplied from the object decoding unit 62, the metadata supplied from the metadata decoding unit 63, the sense-of-distance
15 control information supplied from the sense-of-distance control information decoding unit 64, the listening position information supplied from the user interface 65, and the distance information supplied from the distance calculation unit 66.

20 [0234]

For example, in a case where the sense-of-distance control processing unit 67 has the configuration illustrated in Fig. 3 and the sense-of-distance control information illustrated in Fig. 11 is supplied, the
25 sense-of-distance control processing unit 67 calculates the parameters used in each processing step on the basis of the sense-of-distance control information and the distance information.

[0235]

30 Specifically, for example, the sense-of-distance control processing unit 67 obtains a gain value at the

distance d indicated by the distance information on the basis of the distance $\text{distance}[i]$ and the gain value $\text{gain}[i]$ of each control change point, and supplies the gain value to the gain control unit 101.

5 [0236]

Furthermore, on the basis of the distance $\text{distance}[i]$, the frequency $\text{freq}[i]$, the Q value $Q[i]$, and the gain value $\text{gain}[i]$ of each control change point of the high-shelf filter, the sense-of-distance control processing unit 67 obtains the cutoff frequency, the Q value, and the gain value at the distance d indicated by the distance information, and supplies them to the high-shelf filter processing unit 102.

[0237]

15 Therefore, the high-shelf filter processing unit 102 can construct the high-shelf filter corresponding to the distance d indicated by the distance information.

[0238]

Similarly to the case of the high-shelf filter, the sense-of-distance control processing unit 67 obtains the cutoff frequency, the Q value, and the gain value of the low-shelf filter at the distance d indicated by the distance information, and supplies them to the low-shelf filter processing unit 103. Therefore, the low-shelf filter processing unit 103 can construct the low-shelf filter corresponding to the distance d indicated by the distance information.

25

[0239]

Moreover, the sense-of-distance control processing unit 67 obtains a wet gain value at the distance d indicated by the distance information on the basis of the

30

distance distance[i] and the wet gain value wet_gain[i] of each control change point, and supplies the wet gain value to the reverb processing unit 104.

[0240]

5 Therefore, the sense-of-distance control processing unit 67 illustrated in Fig. 3 is constructed from the sense-of-distance control information.

[0241]

 Furthermore, the sense-of-distance control
10 processing unit 67 supplies the offset angle wet_azimuth_offset[i][j] of the horizontal angle and the offset angle wet_elevation_offset[i][j] of the vertical angle, the metadata of the object, and the listening position information to the reverb processing unit 104.

15 [0242]

 The gain control unit 101 performs gain control processing on the audio data of the object on the basis of the gain value supplied from the sense-of-distance control processing unit 67, and supplies the resultant
20 audio data to the high-shelf filter processing unit 102.

[0243]

 The high-shelf filter processing unit 102 performs filter processing on the audio data supplied from the gain control unit 101 by the high-shelf filter determined
25 by the cutoff frequency, the Q value, and the gain value supplied from the sense-of-distance control processing unit 67, and supplies the resultant audio data to the low-shelf filter processing unit 103.

[0244]

30 The low-shelf filter processing unit 103 performs filter processing on the audio data supplied from the

high-shelf filter processing unit 102 by the low-shelf filter determined by the cutoff frequency, the Q value, and the gain value supplied from the sense-of-distance control processing unit 67.

5 [0245]

The sense-of-distance control processing unit 67 supplies, to the 3D audio rendering processing unit 68, the audio data obtained by the filter processing in the low-shelf filter processing unit 103 as the audio data of
10 the dry component together with the metadata of the object of the dry component. The metadata of the dry component is the metadata supplied from the metadata decoding unit 63.

[0246]

15 Furthermore, the low-shelf filter processing unit 103 supplies the audio data obtained by the filter processing to the reverb processing unit 104.

[0247]

Then, for example, as described with reference to
20 Fig. 4, the reverb processing unit 104 performs gain control based on the wet gain value for the audio data of the dry component, delay processing on the audio data, filter processing using a comb filter and an all-pass filter, and the like, and generates the audio data of the
25 wet component.

[0248]

Furthermore, the reverb processing unit 104 calculates the position information of the wet component on the basis of the offset angle `wet_azimuth_offset[i][j]`
30 and the offset angle `wet_elevation_offset[i][j]`, the metadata of the object (dry component), and the listening

position information, and generates the metadata of the wet component including the position information.

[0249]

The reverb processing unit 104 supplies the audio
5 data and metadata of each wet component generated in this manner to the 3D audio rendering processing unit 68.

[0250]

In step S48, the 3D audio rendering processing unit
68 performs rendering processing on the basis of the
10 audio data and the metadata supplied from the sense-of-distance control processing unit 67 and the listening position information supplied from the user interface 65, and generates reproduction audio data. For example, in step S48, VBAP or the like is performed as the rendering
15 processing.

[0251]

When the reproduction audio data is generated, the 3D audio rendering processing unit 68 outputs the generated reproduction audio data to the subsequent
20 stage, and the decoding process ends.

[0252]

As described above, the decoding device 51 performs the sense-of-distance control processing on the basis of the sense-of-distance control information included in the
25 coded data, and generates the reproduction audio data. In this way, it is possible to realize the sense-of-distance control based on the intention of the content creator.

[0253]

30 <First modification of first embodiment>

<Another example of parameter configuration>

information>

Note that, although the examples illustrated in Figs. 12, 13, and 14 have been described above as the parameter configuration information, the parameter

5 configuration information is not limited thereto, and any parameter configuration information may be used as long as the parameter of the sense-of-distance control processing can be obtained.

[0254]

10 For example, it is also conceivable to prepare in advance a table, a function (mathematical expression), or the like for obtaining a parameter for the distance d from the listening position to the object for each of one or more processing steps configuring the sense-of-
15 distance control processing, and include an index indicating the table or the function in the parameter configuration information. In this case, the index indicating the table or the function is the control rule information indicating the control rule of the parameter.

20 [0255]

In a case where the index indicating the table or the function for obtaining the parameter is set as the control rule information in this manner, for example, as illustrated in Fig. 17, a plurality of tables and
25 functions for obtaining the gain value of the gain control processing as the parameter can be prepared.

[0256]

In this example, for example, a function " $20\log_{10}(1/d)^2$ " for obtaining the gain value of the gain
30 control processing is prepared for the index value "1", and the gain value of the gain control processing

corresponding to the distance d can be obtained by substituting the distance d into this function.

[0257]

Furthermore, for example, a table for obtaining the gain value of the gain control processing is prepared for the index value "2", and when this table is used, the gain value as the parameter decreases as the distance d increases.

[0258]

10 The sense-of-distance control processing unit 67 of the decoding device 51 holds the table or the function in advance in association with such each index.

[0259]

In such a case, for example, the parameter configuration information `DistanceRender_Attn()` illustrated in Fig. 11 has the configuration illustrated in Fig. 18.

[0260]

20 In the example of Fig. 18, the parameter configuration information `DistanceRender_Attn()` includes the index "index" indicating the function or table designated by the content creator.

[0261]

25 Therefore, the sense-of-distance control processing unit 67 reads the table or the function held in association with the index "index", and obtains a gain value as the parameter on the basis of the read table or function and the distance d from the listening position to the object.

30 [0262]

In this way, when a plurality of patterns, that is,

a plurality of tables or functions for obtaining the parameter corresponding to the distance d is defined in advance, the content creator can designate (select) a desired pattern from among these patterns, thereby
5 performing the sense-of-distance control processing according to his/her intention.

[0263]

Note that, here, an example has been described in which the table or the function for obtaining the
10 parameter of the gain control processing is designated by the index. However, the present invention is not limited thereto, and also in the case of the filter processing of the high-shelf filter and the like or the reverb processing, the control rule of the parameter can be
15 designated by the index in the similar manner.

[0264]

<Second modification of first embodiment>

<Another example of sense-of-distance control information>

20 Furthermore, in the above description, an example has been described in which the parameter corresponding to the distance d is determined with the same control rule for all objects. However, the control rule of the parameter may be set (designated) for every object.

25 [0265]

In such a case, the sense-of-distance control information is configured as illustrated in Fig. 19, for example.

[0266]

30 In the example illustrated in Fig. 19, "num_objs" indicates the number of objects included in the content,

and for example, the number num_objs of objects is given to the sense-of-distance control information determination unit 23 from the outside.

[0267]

5 In the sense-of-distance control information, flags "isDistanceRenderFlg" indicating whether or not an object is the target of the sense-of-distance control are included as many as the number num_objs of the objects.

[0268]

10 For example, in a case where the value of the flag isDistanceRenderFlg of the i-th object is "1", the object is determined to be the target of the sense-of-distance control, and the sense-of-distance control processing is performed on the audio data of the object.

15 [0269]

 In a case where the value of the flag isDistanceRenderFlg of the i-th object is "1", the sense-of-distance control information includes the parameter configuration information DistanceRender_Attn(), two
20 pieces of parameter configuration information DistanceRender_Filt(), and the parameter configuration information DistanceRender_Revb() of the object.

[0270]

 Therefore, in this case, as described above, the
25 sense-of-distance control processing unit 67 performs the sense-of-distance control processing on the audio data of the target object, and outputs the obtained audio data and metadata of the dry component and the wet component.

[0271]

30 On the other hand, in a case where the value of the flag isDistanceRenderFlg of the i-th object is "0", it is

determined that the object is not the target of the sense-of-distance control, that is, is nontarget, and the sense-of-distance control processing is not performed on the audio data of the object.

5 [0272]

Therefore, for such an object, the audio data and metadata of the object are supplied without change from the sense-of-distance control processing unit 67 to the 3D audio rendering processing unit 68.

10 [0273]

In a case where the value of the flag isDistanceRenderFlg of the i-th object is "0", the sense-of-distance control information does not include the parameter configuration information

15 DistanceRender_Attn(), the parameter configuration information DistanceRender_Filt(), and the parameter configuration information DistanceRender_Revb() of the object.

[0274]

20 As described above, in the example illustrated in Fig. 19, the sense-of-distance control information encoding unit 24 encodes the parameter configuration information for every object. In other words, the sense-of-distance control information is encoded for every
25 object. Therefore, the sense-of-distance control based on the intention of the content creator can be realized for every object, and content reproduction with higher realistic feeling can be performed.

[0275]

30 In particular, in this example, when the flag isDistanceRenderFlg is stored in the sense-of-distance

control information, it is possible to set whether or not to perform the sense-of-distance control for every object and then perform different sense-of-distance control for every object.

5 [0276]

For example, with respect to an object of human voice, by setting a control rule different from that of other objects other than the object or by not performing the sense-of-distance control itself, it is possible to
10 cause the listener to feel less sense of distance, that is, to reproduce a sound that is always easy for the listener to hear (a sound that is easy to hear).

[0277]

<Third modification of first embodiment>

15 <Another example of sense-of-distance control information>

Furthermore, the control rule of the parameter may be set (designated) not for every object but for every object group including one or more objects.

20 [0278]

In such a case, the sense-of-distance control information is configured as illustrated in Fig. 20, for example.

[0279]

25 In the example illustrated in Fig. 20, "num_obj_groups" indicates the number of object groups included in the content, and for example, the number num_obj_groups of object groups is given to the sense-of-distance control information determination unit 23 from
30 the outside.

[0280]

In the sense-of-distance control information, flags "isDistanceRenderFlg" indicating whether or not an object group, more specifically, an object belonging to the object group is the target of the distance sense control
5 are included as many as the number num_obj_groups of the object group.

[0281]

For example, in a case where the value of the flag isDistanceRenderFlg of the i-th object group is "1", the
10 object group is determined to be the target of the sense-of-distance control, and the sense-of-distance control processing is performed on the audio data of the object belonging to the object group.

[0282]

15 In a case where the value of the flag isDistanceRenderFlg of the i-th object group is "1", the sense-of-distance control information includes the parameter configuration information DistanceRender_Attn(), two pieces of parameter
20 configuration information DistanceRender_Filt(), and the parameter configuration information DistanceRender_Revb() of the object group.

[0283]

Therefore, in this case, as described above, the
25 sense-of-distance control processing unit 67 performs the sense-of-distance control processing on the audio data of the object belonging to the target object group.

[0284]

On the other hand, in a case where the value of the
30 flag isDistanceRenderFlg of the i-th object group is "0", the object group is determined not to be the target of

the sense-of-distance control, and the sense-of-distance control processing is not performed on the audio data of the object of the object group.

[0285]

5 Therefore, for the object of such an object group, the audio data and metadata of the object are without change supplied from the sense-of-distance control processing unit 67 to the 3D audio rendering processing unit 68.

10 [0286]

 In a case where the value of the flag isDistanceRenderFlg of the i-th object group is "0", the sense-of-distance control information does not include the parameter configuration information

15 DistanceRender_Attn(), the parameter configuration information DistanceRender_Filt(), and the parameter configuration information DistanceRender_Revb() of the object group.

[0287]

20 As described above, in the example illustrated in Fig. 20, the sense-of-distance control information encoding unit 24 encodes the parameter configuration information for every object group. In other words, the sense-of-distance control information is encoded for every object group. Therefore, the sense-of-distance control based on the intention of the content creator can be realized for every object group, and content reproduction with higher realistic feeling can be performed.

30 [0288]

 In particular, in this example, when the flag

isDistanceRenderFlg is stored in the sense-of-distance control information, it is possible to set whether or not to perform the sense-of-distance control for every object group and then perform different sense-of-distance

5 control for every object group.

[0289]

For example, in a case where the same control rule is set for a plurality of percussive instruments such as a snare drum, a bass drum, a tom-tom, and a cymbal which
10 configure a drum set, the content creator can group the objects of the plurality of percussive instruments together into one object group.

[0290]

In this way, the same control rule can be set for
15 each object corresponding to each of the plurality of percussive instruments belonging to the same object group and configuring the drum set. That is, the same control rule information can be assigned to each of a plurality of objects. Moreover, as in the example illustrated in
20 Fig. 20, by transmitting the parameter configuration information for every object group, the information amount of the information such as the parameter transmitted to the decoding side, that is, the sense-of-distance control information can be further reduced.

25 [0291]

<Second embodiment>

<Configuration example of sense-of-distance control processing unit>

Furthermore, in the above description, an example
30 has been described in which the configuration of the sense-of-distance control processing unit 67 provided in

the decoding device 51 is determined in advance. That is, an example has been described in which one or more processing steps configuring the sense-of-distance control processing and the order of the processing which are indicated by the configuration information of the sense-of-distance control information are determined in advance.

[0292]

However, the present invention is not limited thereto, and the configuration of the sense-of-distance control processing unit 67 may be freely changed by the configuration information of the sense-of-distance control information.

[0293]

In such a case, the sense-of-distance control processing unit 67 is configured as illustrated in Fig. 21, for example.

[0294]

In the example illustrated in Fig. 21, the sense-of-distance control processing unit 67 executes a program according to the sense-of-distance control information, and realizes some processing blocks among a signal processing unit 201-1 to a signal processing unit 201-3, and a reverb processing unit 202-1 to a reverb processing unit 202-4.

[0295]

The signal processing unit 201-1 performs signal processing on the audio data of the object supplied from the object decoding unit 62 on the basis of the distance information supplied from the distance calculation unit 66 and the sense-of-distance control information supplied

from the sense-of-distance control information decoding unit 64, and supplies the resultant audio data to the signal processing unit 201-2.

[0296]

5 At this time, in a case where the reverb processing unit 202-2 functions, that is, in a case where the reverb processing unit 202-2 is realized, the signal processing unit 201-1 also supplies the audio data obtained by the signal processing to the reverb processing unit 202-2.

10 [0297]

 The signal processing unit 201-2 performs signal processing on the audio data supplied from the signal processing unit 201-1 on the basis of the distance information supplied from the distance calculation unit 66 and the sense-of-distance control information supplied from the sense-of-distance control information decoding unit 64, and supplies the resultant audio data to the signal processing unit 201-3. At this time, in a case where the reverb processing unit 202-3 functions, the signal processing unit 201-2 also supplies the audio data obtained by the signal processing to the reverb processing unit 202-3.

[0298]

 The signal processing unit 201-3 performs signal processing on the audio data supplied from the signal processing unit 201-2 on the basis of the distance information supplied from the distance calculation unit 66 and the sense-of-distance control information supplied from the sense-of-distance control information decoding unit 64, and supplies the resultant audio data to the 3D audio rendering processing unit 68. At this time, in a

case where the reverb processing unit 202-4 functions, the signal processing unit 201-3 also supplies the audio data obtained by the signal processing to the reverb processing unit 202-4.

5 [0299]

Note that, hereinafter, the signal processing units 201-1 to 201-3 will also be simply referred to as signal processing units 201 in a case where it is not particularly necessary to distinguish the signal processing units.

10 [0300]

The signal processing performed by the signal processing unit 201-1, the signal processing unit 201-2, and the signal processing unit 201-3 is the processing indicated by the configuration information of the sense-of-distance control information.

15 [0301]

Specifically, the signal processing performed by the signal processing unit 201 is, for example, gain control processing and filter processing by the high-shelf filter, the low-shelf filter, and the like.

20 [0302]

The reverb processing unit 202-1 performs reverb processing on the audio data of the object supplied from the object decoding unit 62 on the basis of the distance information supplied from the distance calculation unit 66 and the sense-of-distance control information supplied from the sense-of-distance control information decoding unit 64, and generates audio data of a wet component.

25 30 [0303]

Furthermore, the reverb processing unit 202-1

generates the metadata including the position information of the wet component on the basis of the sense-of-distance control information supplied from the sense-of-distance control information decoding unit 64, the
5 metadata supplied from the metadata decoding unit 63, and the listening position information supplied from the user interface 65. Note that, in the reverb processing unit 202-1, the metadata of the wet component is generated using the distance information as necessary.

10 [0304]

The reverb processing unit 202-1 supplies the metadata and the audio data of the wet component generated in this manner to the 3D audio rendering processing unit 68.

15 [0305]

The reverb processing unit 202-2 generates metadata and audio data of a wet component on the basis of the distance information from the distance calculation unit 66, the sense-of-distance control information from the
20 sense-of-distance control information decoding unit 64, the audio data from the signal processing unit 201-1, the metadata from the metadata decoding unit 63, and the listening position information from the user interface 65, and supplies the generated metadata and audio data to
25 the 3D audio rendering processing unit 68.

[0306]

The reverb processing unit 202-3 generates metadata and audio data of a wet component on the basis of the distance information from the distance calculation unit
30 66, the sense-of-distance control information from the sense-of-distance control information decoding unit 64,

the audio data from the signal processing unit 201-2, the metadata from the metadata decoding unit 63, and the listening position information from the user interface 65, and supplies the generated metadata and audio data to the 3D audio rendering processing unit 68.
[0307]

The reverb processing unit 202-4 generates metadata and audio data of a wet component on the basis of the distance information from the distance calculation unit 66, the sense-of-distance control information from the sense-of-distance control information decoding unit 64, the audio data from the signal processing unit 201-3, the metadata from the metadata decoding unit 63, and the listening position information from the user interface 65, and supplies the generated metadata and audio data to the 3D audio rendering processing unit 68.
[0308]

In the reverb processing unit 202-2, the reverb processing unit 202-3, and the reverb processing unit 202-4, processing similar to the case of the reverb processing unit 202-1 is performed, and the metadata and audio data of the wet component are generated.
[0309]

Note that, hereinafter, the reverb processing unit 202-1 to the reverb processing unit 202-4 will also be simply referred to as a reverb processing unit 202 in a case where it is not particularly necessary to distinguish the reverb processing units.
[0310]

In the sense-of-distance control processing unit 67, no reverb processing unit 202 may function, or one or

more reverb processing units 202 may function.

[0311]

Therefore, for example, the sense-of-distance control processing unit 67 may include the reverb processing unit 202 that generates a wet component positioned on the right and left with respect to the object (dry component) and a reverb processing unit 202 that generates a wet component positioned on the upper and lower sides with respect to the object.

10 [0312]

As described above, the content creator can freely designate each of the signal processing steps configuring the sense-of-distance control processing and the order in which the signal processing steps are performed.

15 Therefore, it is possible to realize the sense-of-distance control based on the intention of the content creator.

[0313]

20 <Another example of sense-of-distance control information>

Furthermore, in a case where the configuration of the sense-of-distance control processing unit 67 can be freely changed (designated) as illustrated in Fig. 21, the sense-of-distance control information has the configuration illustrated in Fig. 22, for example.

25 configuration illustrated in Fig. 22, for example.

[0314]

In the example illustrated in Fig. 22, "num_objs" indicates the number of objects included in the content, and in the sense-of-distance control information, flags "isDistanceRenderFlg" indicating whether or not the object is the target of the sense-of-distance control are

30 "isDistanceRenderFlg" indicating whether or not the

included as many as the number num_objs of the objects.
[0315]

Note that the number num_objs of these objects and the flag isDistanceRenderFlg are similar to those in the
5 example illustrated in Fig. 19, and thus the description thereof will be omitted.

[0316]

In a case where the value of the flag
isDistanceRenderFlg of the i-th object is "1", the sense-
10 of-distance control information includes id information
"proc_id" indicating signal processing and parameter
configuration information for each of the signal
processing steps configuring the sense-of-distance
control processing to be performed on the object.

15 [0317]

That is, for example, in accordance with the id
information "proc_id " indicating j-th (where $0 \leq j < 4$)
signal processing, the parameter configuration
information "DistanceRender_Attn()" of the gain control
20 processing, the parameter configuration information
"DistanceRender_Filt()" of the filter processing, the
parameter configuration information
"DistanceRender_Revb()" of the reverb processing, or
parameter configuration information
25 "DistanceRender_UserDefine()" of user definition
processing is included in the sense-of-distance control
information.

[0318]

Specifically, for example, in a case where the id
30 information "proc_id" is "ATTN" indicating the gain
control processing, the parameter configuration

information "DistanceRender_Attn()" of the gain control processing is included in the sense-of-distance control information.

[0319]

5 Note that the parameter configuration information "DistanceRender_Attn()", "DistanceRender_Filt()", and "DistanceRender_Revb()" is similar to the case in Fig. 11, and thus description thereof is omitted.

[0320]

10 Furthermore, the parameter configuration information "DistanceRender_UserDefine()" indicates parameter configuration information indicating the control rule of the parameter used in the user definition processing which is signal processing arbitrarily defined
15 by the user.

[0321]

 Therefore, in this example, in addition to the gain control processing, the filter processing, and the reverb processing, the user definition processing separately
20 defined by the user can be added as the signal processing configuring the sense-of-distance control processing.

[0322]

 Note that, here, a case where the number of the signal processing steps configuring the sense-of-distance
25 control processing is four has been described as an example, but the number of the signal processing steps configuring the sense-of-distance control processing may be any number.

[0323]

30 In the sense-of-distance control information illustrated in Fig. 22, for example, when 0-th signal

processing configuring the sense-of-distance control processing is set to the gain control processing, first signal processing is set to the filter processing by the high-shelf filter, second signal processing is set to the filter processing by the low-shelf filter, and third signal processing is set to the reverb processing, the sense-of-distance control processing unit 67 having the same configuration as that illustrated in Fig. 3 is realized.

10 [0324]

In such a case, in the sense-of-distance control processing unit 67 illustrated in Fig. 21, the signal processing unit 201-1 to the signal processing unit 201-3 and the reverb processing unit 202-4 are realized, and the reverb processing unit 202-1 to the reverb processing unit 202-3 are not realized (do not function).

[0325]

Then, the signal processing unit 201-1 to the signal processing unit 201-3, and the reverb processing unit 202-4 function as the gain control unit 101, the high-shelf filter processing unit 102, the low-shelf filter processing unit 103, and the reverb processing unit 104 illustrated in Fig. 3.

[0326]

25 As described above, even in a case where the sense-of-distance control information has the configuration illustrated in Fig. 22, basically, the encoding device 11 performs the encoding process described with reference to Fig. 15, and the decoding device 51 performs the decoding process described with reference to Fig. 16.

30 [0327]

However, in the encoding process, for example, in step S13, for every object, whether or not the object is to be subjected to the sense-of-distance control processing, the configuration of the sense-of-distance control processing, and the like are determined, and in
5 step S14, the sense-of-distance control information having the configuration illustrated in Fig. 22 is encoded.

[0328]

10 On the other hand, in the decoding process, in step S47, the configuration of the sense-of-distance control processing unit 67 is determined for every object on the basis of the sense-of-distance control information having the configuration illustrated in Fig. 22, and the sense-
15 of-distance control processing is appropriately performed.

[0329]

As described above, according to the present technology, the sense-of-distance control information is
20 transmitted to the decoding side together with the audio data of the object according to the setting of the content creator or the like, whereby the sense-of-distance control based on the intention of the content creator can be realized in the object-based audio.

25 [0330]

<Configuration example of computer>

By the way, the series of processes described above can be executed by hardware but can also be executed by software. In a case where the series of processing is
30 executed by software, a program configuring the software is installed in a computer. Here, the computer includes

a computer incorporated in dedicated hardware, a general-purpose personal computer capable of executing various functions by installing various programs, and the like, for example.

5 [0331]

Fig. 23 is a block diagram illustrating a configuration example of the hardware of the computer that executes the above-described series of processing by the program.

10 [0332]

In the computer, a central processing unit (CPU) 501, a read only memory (ROM) 502, and a random access memory (RAM) 503 are mutually connected by a bus 504.

[0333]

15 An input/output interface 505 is further connected to the bus 504. An input unit 506, an output unit 507, a recording unit 508, a communication unit 509, and a drive 510 are connected to the input/output interface 505.

[0334]

20 The input unit 506 includes a keyboard, a mouse, a microphone, an imaging element, and the like. The output unit 507 includes a display, a speaker, and the like. The recording unit 508 includes a hard disk, a nonvolatile memory, and the like. The communication unit
25 509 includes a network interface and the like. The drive 510 drives a removable recording medium 511 such as a magnetic disk, an optical disk, a magneto-optical disk, or a semiconductor memory.

[0335]

30 In the computer configured as described above, the above-described series of processing is performed, for

example, in such a manner that the CPU 501 loads the program recorded in the recording unit 508 into the RAM 503 via the input/output interface 505 and the bus 504 and executes the program.

5 [0336]

For example, the program executed by the computer (CPU 501) can be recorded and provided on the removable recording medium 511 as a package medium and the like. Furthermore, the program can be provided via a wired or
10 wireless transmission medium such as a local area network, the Internet, or digital satellite broadcasting.
[0337]

In the computer, the program can be installed in the recording unit 508 via the input/output interface 505
15 by mounting the removable recording medium 511 to the drive 510. Furthermore, the program can be received by the communication unit 509 and installed in the recording unit 508 via a wired or wireless transmission medium. In addition, the program can be installed in advance in the
20 ROM 502 or the recording unit 508.
[0338]

Note that the program executed by the computer may be a program in which processing is performed in time series in the order described in this description or a
25 program in which processing is performed in parallel or at a necessary timing such as when a call is made.
[0339]

Furthermore, the embodiments of the present technology are not limited to the above-described
30 embodiments, and various modifications can be made without departing from the gist of the present

technology.

[0340]

For example, the present technology can be configured as cloud computing in which one function is
5 shared by a plurality of devices via a network and jointly processed.

[0341]

Furthermore, each step described in the above-described flowcharts can be executed by one device or
10 shared by a plurality of devices.

[0342]

Moreover, in a case where one step includes a plurality of processes, the plurality of processes included in the one step can be executed by one device or
15 shared by a plurality of devices.

[0343]

Moreover, the present technology can have the following configurations.

[0344]

20 (1)

An encoding device including:

an object encoding unit that encodes audio data of an object;

a metadata encoding unit that encodes metadata
25 including position information of the object;

a sense-of-distance control information determination unit that determines sense-of-distance control information for sense-of-distance control processing to be performed on the audio data;

30 a sense-of-distance control information encoding unit that encodes the sense-of-distance control

information; and

a multiplexer that multiplexes the coded audio data, the coded metadata, and the coded sense-of-distance control information to generate coded data.

5 (2)

The encoding device according to (1),
in which the sense-of-distance control information includes control rule information for obtaining a parameter used in the sense-of-distance control
10 processing.

(3)

The encoding device according to (2),
in which the parameter changes according to a distance from a listening position to the object.

15 (4)

The encoding device according to (2) or (3),
in which the control rule information is an index indicating a function or a table for obtaining the parameter.

20 (5)

The encoding device according to any one of (2) to (4),

in which the sense-of-distance control information includes configuration information indicating one or more
25 processing steps which are performed in combination to realize the sense-of-distance control processing.

(6)

The encoding device according to (5),
in which the configuration information is
30 information indicating the one or more processing steps and an order of performing the one or more processing

steps.

(7)

The encoding device according to (5) to (6),
in which the processing is gain control processing,
5 filter processing, or reverb processing.

(8)

The encoding device according to any one of (1) to
(7),

in which the sense-of-distance control information
10 encoding unit encodes the sense-of-distance control
information for each of a plurality of the objects.

(9)

The encoding device according to any one of (1) to
(7),

15 in which the sense-of-distance control information
encoding unit encodes the sense-of-distance control
information for every object group including one or a
plurality of the objects.

(10)

20 An encoding method performed by an encoding device,
the method including:

encoding audio data of an object;

encoding metadata including position information of
the object;

25 determining sense-of-distance control information
for sense-of-distance control processing to be performed
on the audio data;

encoding the sense-of-distance control information;
and

30 multiplexing the coded audio data, the coded
metadata, and the coded sense-of-distance control

information to generate coded data.

(11)

A program for causing a computer to execute processing including the steps of:

- 5 encoding audio data of an object;
- encoding metadata including position information of the object;
- determining sense-of-distance control information for sense-of-distance control processing to be performed
- 10 on the audio data;
- encoding the sense-of-distance control information;
- and
- multiplexing the coded audio data, the coded metadata, and the coded sense-of-distance control
- 15 information to generate coded data.

(12)

A decoding device including:

- a demultiplexer that demultiplexes coded data to extract coded audio data of an object, coded metadata
- 20 including position information of the object, and coded sense-of-distance control information for sense-of-distance control processing to be performed on the audio data;
- an object decoding unit that decodes the coded
- 25 audio data;
- a metadata decoding unit that decodes the coded metadata;
- a sense-of-distance control information decoding unit that decodes the coded sense-of-distance control
- 30 information;
- a sense-of-distance control processing unit that

performs the sense-of-distance control processing on the audio data of the object on the basis of the sense-of-distance control information; and

a rendering processing unit that performs rendering processing on the basis of the audio data obtained by the sense-of-distance control processing and the metadata to generate reproduction audio data for reproducing a sound of the object.

(13)

The decoding device according to (12),
in which the sense-of-distance control processing unit performs the sense-of-distance control processing on the basis of a parameter obtained from control rule information included in the sense-of-distance control information and a listening position.

(14)

The decoding device according to (13),
in which the parameter changes according to a distance from the listening position to the object.

(15)

The decoding device according to (13) or (14),
in which the sense-of-distance control processing unit adjusts the parameter according to a reproduction environment of the reproduction audio data.

(16)

The decoding device according to any one of (13) to (15),

in which the sense-of-distance control processing unit performs, on the basis of the parameter, the sense-of-distance control processing in which one or more processing steps indicated by the sense-of-distance

control information is combined.

(17)

The decoding device according to (16),
in which the processing is gain control processing,
5 filter processing, or reverb processing.

(18)

The decoding device according to any one of (12) to
(17),

in which the sense-of-distance control processing
10 unit generates audio data of a wet component of the
object by the sense-of-distance control processing.

(19)

A decoding method performed by a decoding device,
the method including:

15 demultiplexing coded data to extract coded audio
data of an object, coded metadata including position
information of the object, and coded sense-of-distance
control information for sense-of-distance control
processing to be performed on the audio data;

20 decoding the coded audio data;
decoding the coded metadata;
decoding the coded sense-of-distance control
information;

performing the sense-of-distance control processing
25 on the audio data of the object on the basis of the
sense-of-distance control information; and

performing rendering processing on the basis of the
audio data obtained by the sense-of-distance control
processing and the metadata to generate reproduction
30 audio data for reproducing a sound of the object.

(20)

A program for causing a computer to execute processing including the steps of:

- demultiplexing coded data to extract coded audio data of an object, coded metadata including position
- 5 information of the object, and coded sense-of-distance control information for sense-of-distance control processing to be performed on the audio data;
- decoding the coded audio data;
- decoding the coded metadata;
- 10 decoding the coded sense-of-distance control information;
- performing the sense-of-distance control processing on the audio data of the object on the basis of the sense-of-distance control information; and
- 15 performing rendering processing on the basis of the audio data obtained by the sense-of-distance control processing and the metadata to generate reproduction audio data for reproducing a sound of the object.

20 REFERENCE SIGNS LIST

[0345]

- 11 Encoding device
- 21 Object encoding unit
- 22 Metadata encoding unit
- 25 23 Sense-of-distance control information determination unit
- 24 Sense-of-distance control information encoding unit
- 25 Multiplexer
- 51 Decoding device
- 30 61 Demultiplexer
- 62 Object decoding unit

	63	Metadata decoding unit
	64	Sense-of-distance control information decoding unit
	66	Distance calculation unit
	67	Sense-of-distance control processing unit
5	68	3D audio rendering processing unit
	101	Gain control unit
	102	High-shelf filter processing unit
	103	Low-shelf filter processing unit
	104	Reverb processing unit

CLAIMS

1. An encoding device comprising:
 - an object encoding unit that encodes audio data of
 - 5 an object;
 - a metadata encoding unit that encodes metadata including position information of the object;
 - a sense-of-distance control information determination unit that determines sense-of-distance
 - 10 control information for sense-of-distance control processing to be performed on the audio data;
 - a sense-of-distance control information encoding unit that encodes the sense-of-distance control information; and
 - 15 a multiplexer that multiplexes the coded audio data, the coded metadata, and the coded sense-of-distance control information to generate coded data.
2. The encoding device according to claim 1,
- 20 wherein the sense-of-distance control information includes control rule information for obtaining a parameter used in the sense-of-distance control processing.
- 25 3. The encoding device according to claim 2,
- wherein the parameter changes according to a distance from a listening position to the object.
4. The encoding device according to claim 2,
- 30 wherein the control rule information is an index indicating a function or a table for obtaining the

parameter.

5. The encoding device according to claim 2,
 wherein the sense-of-distance control information
5 includes configuration information indicating one or more
processing steps which are performed in combination to
realize the sense-of-distance control processing.
6. The encoding device according to claim 5,
10 wherein the configuration information is
information indicating the one or more processing steps
and an order of performing the one or more processing
steps.
- 15 7. The encoding device according to claim 5,
 wherein the processing is gain control processing,
filter processing, or reverb processing.
8. The encoding device according to claim 1,
20 wherein the sense-of-distance control information
encoding unit encodes the sense-of-distance control
information for each of a plurality of the objects.
9. The encoding device according to claim 1,
25 wherein the sense-of-distance control information
encoding unit encodes the sense-of-distance control
information for every object group including one or a
plurality of the objects.
- 30 10. An encoding method performed by an encoding device,
the method comprising:

encoding audio data of an object;
encoding metadata including position information of
the object;

determining sense-of-distance control information
5 for sense-of-distance control processing to be performed
on the audio data;

encoding the sense-of-distance control information;
and

multiplexing the coded audio data, the coded
10 metadata, and the coded sense-of-distance control
information to generate coded data.

11. A program for causing a computer to execute
processing including the steps of:

15 encoding audio data of an object;
encoding metadata including position information of
the object;

determining sense-of-distance control information
for sense-of-distance control processing to be performed
20 on the audio data;

encoding the sense-of-distance control information;
and

multiplexing the coded audio data, the coded
metadata, and the coded sense-of-distance control
25 information to generate coded data.

12. A decoding device comprising:

a demultiplexer that demultiplexes coded data to
extract coded audio data of an object, coded metadata
30 including position information of the object, and coded
sense-of-distance control information for sense-of-

distance control processing to be performed on the audio data;

an object decoding unit that decodes the coded audio data;

5 a metadata decoding unit that decodes the coded metadata;

a sense-of-distance control information decoding unit that decodes the coded sense-of-distance control information;

10 a sense-of-distance control processing unit that performs the sense-of-distance control processing on the audio data of the object on a basis of the sense-of-distance control information; and

a rendering processing unit that performs rendering
15 processing on a basis of the audio data obtained by the sense-of-distance control processing and the metadata to generate reproduction audio data for reproducing a sound of the object.

20 13. The decoding device according to claim 12,
wherein the sense-of-distance control processing unit performs the sense-of-distance control processing on a basis of a parameter obtained from control rule information included in the sense-of-distance control
25 information and a listening position.

14. The decoding device according to claim 13,
wherein the parameter changes according to a distance from the listening position to the object.

30

15. The decoding device according to claim 13,

wherein the sense-of-distance control processing unit adjusts the parameter according to a reproduction environment of the reproduction audio data.

5 16. The decoding device according to claim 13,
wherein the sense-of-distance control processing unit performs, on a basis of the parameter, the sense-of-distance control processing in which one or more processing steps indicated by the sense-of-distance
10 control information is combined.

17. The decoding device according to claim 16,
wherein the processing is gain control processing, filter processing, or reverb processing.

15

18. The decoding device according to claim 12,
wherein the sense-of-distance control processing unit generates audio data of a wet component of the object by the sense-of-distance control processing.

20

19. A decoding method performed by a decoding device, the method comprising:

demultiplexing coded data to extract coded audio data of an object, coded metadata including position
25 information of the object, and coded sense-of-distance control information for sense-of-distance control processing to be performed on the audio data;

decoding the coded audio data;

decoding the coded metadata;

30 decoding the coded sense-of-distance control information;

performing the sense-of-distance control processing
on the audio data of the object on a basis of the sense-
of-distance control information; and

performing rendering processing on a basis of the
5 audio data obtained by the sense-of-distance control
processing and the metadata to generate reproduction
audio data for reproducing a sound of the object.

20. A program for causing a computer to execute
10 processing including the steps of:

demultiplexing coded data to extract coded audio
data of an object, coded metadata including position
information of the object, and coded sense-of-distance
control information for sense-of-distance control

15 processing to be performed on the audio data;

decoding the coded audio data;

decoding the coded metadata;

decoding the coded sense-of-distance control
information;

20 performing the sense-of-distance control processing
on the audio data of the object on a basis of the sense-
of-distance control information; and

performing rendering processing on a basis of the
audio data obtained by the sense-of-distance control
25 processing and the metadata to generate reproduction
audio data for reproducing a sound of the object.