

DESCRIPTION

TITLE OF THE INVENTION

INFORMATION PROCESSING DEVICE AND METHOD, AND PROGRAM

5 TECHNICAL FIELD

[0001]

The present technology relates to an information processing device, an information processing method, and a program, and in particular, relates to an information
10 processing device, an information processing method, and a program capable of reducing an amount of transmission of directivity data.

BACKGROUND ART

15 [0002]

In the related art, it is known that audio reproduction with a higher realistic feeling can be realized by considering directivity of a sound source.

[0003]

20 For example, when directivity data representing directivity of sound from an object is prepared together with audio data of the object, audio reproduction based on directional characteristics of the object can be performed using the audio data and the directivity data.

25 [0004]

Furthermore, as a technology regarding directivity, for example, a technology has been proposed in which a user can perform recording by arbitrarily selecting a directivity direction at the time of recording, and the
30 user selects and reproduces a desired directivity direction separately from the directivity direction at

the time of recording (see, for example, Patent Document 1).

CITATION LIST

5 PATENT DOCUMENT

[0005]

Patent Document 1: Japanese Patent Application Laid-Open
No. 2021-100209

10 SUMMARY OF THE INVENTION

PROBLEMS TO BE SOLVED BY THE INVENTION

[0006]

Meanwhile, since directional characteristics
(directivity) are different for each sound source, in a
15 case where audio data of an object and directivity data
of the object are provided as content, it is necessary to
prepare directivity data for each type of sound source,
that is, for each type of object. In addition, when it
is attempted to provide information regarding directivity
20 for more directions and frequencies, the data amount of
the directivity data increases.

[0007]

Then, the amount of transmission of the directivity
data to the distribution destination of the content
25 increases, and there is a possibility that a transmission
delay occurs or a transmission rate increases.

[0008]

The present technology has been made in view of
such a situation, and an object thereof is to reduce an
30 amount of transmission of directivity data.

SOLUTIONS TO PROBLEMS

[0009]

An information processing device according to a first aspect of the present technology includes an acquisition unit configured to acquire model data obtained by modeling directivity data indicating directivity of a sound source, and a calculator configured to calculate the directivity data on the basis of the model data.

10 [0010]

An information processing method or a program according to the first aspect of the present technology includes the steps of acquiring model data obtained by modeling directivity data indicating directivity of a sound source, and calculating the directivity data on the basis of the model data.

[0011]

In the first aspect of the present technology, model data obtained by modeling directivity data indicating directivity of a sound source is acquired, and the directivity data is calculated on the basis of the model data.

[0012]

An information processing device according to a second aspect of the present technology includes a modeling unit configured to model directivity data representing directivity of a sound source with a mixture model including a plurality of distributions, and a model data generation unit configured to generate model data including a model parameter constituting the mixture model, the model parameter being obtained by the

modeling.

[0013]

An information processing method or a program
according to the second aspect of the present technology
5 includes the steps of modeling directivity data
representing directivity of a sound source with a mixture
model including a plurality of distributions, and
generating model data including a model parameter
constituting the mixture model, the model parameter being
10 obtained by the modeling.

[0014]

In the second aspect of the present technology,
directivity data representing directivity of a sound
source is modeled with a mixture model including a
15 plurality of distributions, and model data including a
model parameter constituting the mixture model, the model
parameter being obtained by the modeling, is generated.

BRIEF DESCRIPTION OF DRAWINGS

20 [0015]

Fig. 1 is a diagram for describing a mixed Gaussian
distribution.

Fig. 2 is a diagram for describing a vMF
distribution and a Kent distribution.

25 Fig. 3 is a diagram illustrating an example of
directivity.

Fig. 4 is a diagram for describing data points.

Fig. 5 is a diagram illustrating an example of
model data.

30 Fig. 6 is a diagram for describing a relationship
between a band and a bin.

Fig. 7 is a diagram illustrating a reduction example of the data amount of the directivity data.

Fig. 8 is a diagram for describing a residual of directivity data.

5 Fig. 9 is a diagram depicting an example of a configuration of a server.

Fig. 10 is a flowchart for describing the encoding process.

10 Fig. 11 is a diagram illustrating a configuration example of an information processing device.

Fig. 12 is a flowchart illustrating a directivity data generation process.

Fig. 13 is a flowchart for describing an output audio data generation process.

15 Fig. 14 is a diagram illustrating appearance probability of difference information.

Fig. 15 is a diagram illustrating an example of model data.

20 Fig. 16 is a diagram illustrating an example of model data.

Fig. 17 is a diagram for describing transmission of a Huffman encoding table.

Fig. 18 is a diagram illustrating an example of a Huffman encoding table.

25 Fig. 19 is a diagram illustrating a configuration example of a server.

Fig. 20 is a flowchart illustrating a directivity data generation process.

30 Fig. 21 is a diagram illustrating a configuration example of a directivity data encoding unit.

Fig. 22 is a diagram illustrating a configuration

example of a difference encoding unit.

Fig. 23 is a flowchart illustrating a model data generation process.

Fig. 24 is a diagram illustrating a configuration
5 example of a distribution model decoding unit.

Fig. 25 is a diagram illustrating an example of model data.

Fig. 26 is a diagram illustrating an arrangement example of data points.

10 Fig. 27 is a diagram illustrating a description example of a data point.

Fig. 28 is a diagram illustrating an example of a scale factor of each bin.

Fig. 29 is a diagram illustrating an example of a
15 minimum value of each bin.

Fig. 30 is a diagram illustrating an example of model data.

Fig. 31 is a diagram illustrating an example of Syntax of SymmetricDir().

20 Fig. 32 is a diagram for describing a rotation operation.

Fig. 33 is a diagram for describing a symmetry operation.

Fig. 34 is a diagram illustrating an example of
25 Syntax of NonSymmetricDir().

Fig. 35 is a diagram illustrating an example of model data.

Fig. 36 is a diagram illustrating an example of a weight used to calculate output values of a mixture model
30 of each bin.

Fig. 37 is a diagram illustrating an example of

model data.

Fig. 38 is a diagram illustrating an example of
Syntax of NonSymmetricDir().

Fig. 39 is a diagram illustrating an example of
5 Syntax of LeftRightLineSymmetricDir().

Fig. 40 is a diagram for describing a distribution
according to weights.

Fig. 41 is a diagram illustrating a configuration
example of a computer.

10

MODE FOR CARRYING OUT THE INVENTION

[0016]

Hereinafter, an embodiment to which the present
technology is applied will be described with reference to
15 the drawings.

[0017]

<First embodiment>

<About present technology>

The present technology is to reduce an amount of
20 transmission of directivity data by modeling the
directivity data.

[0018]

In the present technology, for example, audio data
and directivity data of a 3D sound source are provided as
25 content.

[0019]

Specifically, for example, sounds of one or more
audio objects (hereinafter, simply referred to as an
object) are collected (recorded) as a 3D sound source,
30 and audio data of each object is generated. Furthermore,
for each type of object, that is, each sound source type,

directivity data representing a directional characteristic, that is, directivity, of the object (sound source) is prepared.

[0020]

5 Furthermore, the audio data of each object and the directivity data for each sound source type are provided as content data. That is, the directivity data is transmitted to the device on the reproduction side together with the audio data of the object. Then, on the reproduction side, audio reproduction considering the directivity data is performed on the basis of the audio data and the directivity data constituting the content.

[0021]

15 The directivity data can be obtained, for example, by recording sound of an object with a plurality of microphones. Note that the recording of the directivity data may be performed simultaneously with the recording of the audio data of the object, or may be performed at a timing different from the recording of the audio data of the object.

[0022]

25 The directivity data is prepared for each sound source type such as a voice, a musical instrument, or a speaker. Further, the directivity data is, for example, data having information about the amplitude and the phase of the sound from the sound source for each target frequency in the entire frequency band from the frequency to be DC to the Nyquist frequency for a position in each direction viewed from the sound source.

30 [0023]

For example, the direction viewed from the sound

source is represented by an angle in the horizontal direction viewed from the sound source position, that is, an azimuth angle, and an angle in the vertical direction viewed from the sound source position, that is, an elevation angle. At this time, for example, the range of the azimuth angle is set to a range of 0 degrees to 360 degrees, and the range of the elevation angle is set to a range of -90 degrees to +90 degrees.

[0024]

10 In the present technology, in discretizing and compressing such directivity data, parametric compression by modeling is performed instead of directly compressing the data.

[0025]

15 Note that, in the present technology, the directivity data to be modeled is obtained by appropriately discretizing and normalizing the directivity data obtained by recording or the like.

[0026]

20 In the following description, it is assumed that the directivity data to be modeled includes a gain (hereinafter, referred to as a directivity gain) indicating a directional characteristic of each of a plurality of discrete frequencies of a sound source at a plurality of data points.

25 [0027]

 For example, the position to be the data point can be expressed by coordinates (polar coordinates) of a polar coordinate system with the sound source position as the origin, that is, an azimuth angle indicating a position in the horizontal direction viewed from the

30

sound source position and an elevation angle indicating a position in the vertical direction viewed from the sound source position. Note that the distance (radius) from the sound source position may be used to represent the position of the data point. Furthermore, the directivity gain can be obtained by normalizing the amplitude (sound pressure) of the sound from the sound source at the data point.

[0028]

10 (Method of recording directivity data of sound source)

A method of recording the directivity data for each sound source type will be described.

[0029]

In the present technology, a vMF (von Mises Fisher) distribution on a spherical face, a Kent distribution, or a mixture model including at least any one of the vMF distribution or the Kent distribution, corresponding to a multivariate/univariate Gaussian distribution defined on a plane, is used for modeling the directivity data.

20 [0030]

Note that the vMF distribution, the Kent distribution, and the mixture model are described in detail in, for example, "John T. Kent (1982). The Fisher-Bingham Distribution on the Sphere".

25 [0031]

First, a general mixed Gaussian distribution will be described.

[0032]

For example, a two-dimensional Gaussian distribution is illustrated in a portion indicated by the arrow Q11 in Fig. 1. In this example, there are two

Gaussian distributions on a straight line. That is, the curve L11 indicates one Gaussian distribution, and the curve L12 indicates another Gaussian distribution.

[0033]

5 In addition, the curve L13 indicates a mixed Gaussian distribution obtained by mixing the Gaussian distribution indicated by the curve L11 and the Gaussian distribution indicated by the curve L12.

[0034]

10 On the other hand, three distributions on a plane are illustrated in a portion indicated by the arrow Q12 in Fig. 1. It is also possible to mix a plurality of distributions on such a plane.

[0035]

15 Usually, the mixed Gaussian distribution is used to express a probability density function (pdf) on a plane. By expressing a desired pdf with a small number of model parameters and as few number of mixtures as possible, it is possible to reduce the amount of information.

20 [0036]

 In the present technology, the directivity data on the spherical surface, that is, the shape (distribution) of the directivity gain is modeled using the mixture model of the vMF distribution and the Kent distribution
25 corresponding to the Gaussian distribution defined on the spherical surface.

[0037]

 The mixture model may include one or more vMF distributions, one or more Kent distributions, or one or
30 more vMF distributions and one or more Kent distributions. That is, the mixture model includes one

or more distributions including at least any one of the
vMF distribution or the Kent distribution.

[0038]

When the position vector indicating the position of
5 the spherical surface, that is, the coordinates of the
orthogonal coordinate system (Cartesian coordinate
system) is $\mathbf{x}$, the value $f(\mathbf{x})$ of the Kent distribution
corresponding to the position vector $\mathbf{x}$, that is, the
value $f(\mathbf{x})$ of the Kent distribution at the position
10 indicated by the position vector $\mathbf{x}$ can be expressed by
the following Expression (1).

[0039]

[Math 1]

$$f(\mathbf{x}) = \frac{1}{c(\kappa, \beta)} \exp \left\{ \kappa \gamma_1^T \cdot \mathbf{x} + \beta \left[(\gamma_2^T \cdot \mathbf{x})^2 - (\gamma_3^T \cdot \mathbf{x})^2 \right] \right\} \quad \cdot \cdot \cdot (1)$$

15 [0040]

Note that, in Expression (1), κ represents a degree
of parameter concentration, and β represents an
ellipticity. Further, γ_1 represents a vector defining
the center of the mean direction distribution, γ_2
20 represents a major axis vector, and γ_3 represents a minor
axis vector.

[0041]

Further, $c(\kappa, \beta)$ is a normalization constant
expressed by the following Expression (2). Note that, in
25 Expression (2), Γ represents a gamma function, and I
represents a modified Bessel function of the first kind.

[0042]

[Math 2]

$$c(\kappa, \beta) = 2\pi \sum_{j=0}^{\infty} \frac{\Gamma\left(j + \frac{1}{2}\right)}{\Gamma(j+1)} \beta^{2j} \left(\frac{1}{2}\kappa\right)^{-2j - \frac{1}{2}} I_{2j + \frac{1}{2}}(\kappa) \quad \cdot \cdot \cdot (2)$$

[0043]

In addition, the value of the vMF distribution at the position indicated by the position vector x can also be expressed by the expression similar to Expression (1). In such a case, the value of the ellipticity β in Expression (1) is set to 0.

[0044]

Fig. 2 illustrates an example of the vMF distribution and the Kent distribution.

[0045]

In Fig. 2, an example of the vMF distribution is illustrated in a portion indicated by the arrow Q21. Specifically, the vector V11 represents the vector γ_1 shown in Expression (1).

[0046]

The vMF distribution does not have the ellipticity β , the major axis vector γ_2 , and the minor axis vector γ_3 as parameters, and is a circular distribution that spreads isotropically around the position indicated by the vector V11 (vector γ_1) on the spherical surface. That is, a circular distribution can be reproduced by using the vMF distribution (vMF distribution model).

[0047]

On the other hand, an example of the Kent distribution is illustrated in a portion indicated by the arrow Q22. Specifically, vectors V21 to V23 represent the vector γ_1 , the major axis vector γ_2 , and the minor axis vector γ_3 shown in Expression (1).

[0048]

The Kent distribution is an elliptical distribution centered on the position indicated by the vector V_{21} (vector γ_1) on the spherical surface and having the major axis vector γ_2 and the minor axis vector γ_3 on the spherical surface as the major axis and the minor axis, respectively. That is, by using the Kent distribution (Kent distribution model), it is possible to reproduce the distribution of the elliptical shape defined by the ellipticity β , the major axis vector γ_2 , and the minor axis vector γ_3 .

[0049]

The Kent distribution has a high degree of freedom because the shape of the ellipse can be changed by parameters such as the ellipticity β , but the number of parameters is larger than that of the vMF distribution.

[0050]

In the present technology, the directivity data is represented (modeled) by using a mixture model obtained by mixing the vMF distribution and the Kent distribution.

[0051]

For example, at the position indicated by the position vector x as in Expression (1), the output value $F(x; \Theta)$ of a mixture model using N Kent distributions $f(x; \theta_i)$ can be expressed by the following Expression (3). That is, the mixture model $F(x; \Theta)$ can be expressed by weighted addition of the N Kent distributions $f(x; \theta_i)$.

[0052]

30 [Math 3]

$$F(x; \Theta) = \sum_{i=1}^N \phi_i f(x; \theta_i) \quad \cdot \cdot \cdot (3)$$

[0053]

Note that, in Expression (3), the Kent distribution $f(x; \theta_i)$ is similar to that shown in the above Expression (1), and indicates the i -th Kent distribution among the N Kent distributions to be mixed.

[0054]

Further, θ_i is a parameter constituting the Kent distribution $f(x; \theta_i)$, more specifically, a set of parameters, and the parameter θ_i includes a degree of parameter concentration κ , an ellipticity β , a vector γ_1 , a major axis vector γ_2 , and a minor axis vector γ_3 in Expression (1). The parameter Θ of the Mixture model $F(x; \Theta)$ is a set of parameters θ_i of the N Kent distributions $f(x; \theta_i)$.

[0055]

Furthermore, in Expression (3), ϕ_i represents a weight (weight coefficient) of the i -th Kent distribution $f(x; \theta_i)$ when mixing N Kent distributions, and as shown in Expression (4), the sum of the weights ϕ_i of N Kent distributions $f(x; \theta_i)$ is 1.

[0056]

[Math 4]

$$\sum_{i=1}^N \phi_i = 1 \quad \cdot \cdot \cdot (4)$$

25 [0057]

The directivity data used in the present technology can be obtained by performing recording (sound collection) with a microphone array including a plurality of microphones disposed around an object.

[0058]

As an example, when the performance sound of the trumpet is recorded, the directivity illustrated in Fig. 3 is observed. Specifically, the directivity of each frequency in a horizontal plane, that is, a plane having an elevation angle of 0 degrees is illustrated on the left side in the figure, and the directivity of each frequency in a median plane is illustrated on the right side in the figure.

10 [0059]

In this example, it can be seen that the outline of the directivity changes depending on the frequency (pitch) in both the horizontal plane and the median plane, and the directivity is small at the frequency on the low frequency side, but the directivity is larger (sharper) as the frequency increases. For example, on a horizontal plane, a sound pressure difference of about 25 dB at maximum occurs at 8000 Hz depending on the direction.

20 [0060]

Meanwhile, in the directivity data to be modeled, for example, as illustrated in Fig. 4, a plurality of data points is provided on the spherical surface centered on the sound source position. In the example of Fig. 4, one point represents one data point, and it can be seen that there are a large number of data points on the entire spherical surface.

[0061]

Here, for example, when data points are provided at intervals of 2 degrees (in increments of 2 degrees) in the azimuth angle direction and at intervals of 2 degrees

in the elevation angle direction, 16022 data points are provided on the entire spherical surface. Furthermore, in such a case, when an attempt is made to transmit a directivity gain (sound pressure) in 512 bins (frequency
5 bins) for 19 Hz to 20 kHz for each data point, the directivity data of one sound source is about 31 MB.
[0062]

As described above, since the data size of the directivity data for each sound source type is large, the
10 amount of transmission increases.
[0063]

In addition, since there is a vowel sound, a consonant sound, and the like in the sound, and the directivity of the musical instrument varies depending on
15 the playing method, and the variety of sound source types is very large, a large number of pieces of directivity data are required when it is attempted to prepare the directivity data for each sound source type.
[0064]

20 For these reasons, the amount of transmission of directivity data increases, and an increase in the amount of directivity data transmission causes a transmission delay and an increase in the transmission rate.
Therefore, in some cases, it may not be possible to
25 reproduce the directivity according to the sound source type, the frequency, the direction of the object and the listener, and the like.
[0065]

Therefore, in the present technology, by modeling
30 the directivity data using the mixture model as described above, the amount of transmission of the directivity data

can be reduced.

[0066]

<Example of model data>

Here, a specific example of the model data obtained
5 by modeling the directivity data will be described.

[0067]

In the present technology, at the time of
transmitting the directivity data, the directivity data
based on the mixture model including the vMF distribution
10 and the Kent distribution is modeled, and model data
including model parameter and the like constituting the
mixture model obtained as a result thereof is generated.
Then, the model data is transmitted to the device on the
reproduction side of the content. As a result,
15 transmission of the original directivity data having a
large data size is unnecessary. In other words, the
amount of data (amount of transmission) at the time of
transmission of the directivity data can be reduced.

[0068]

20 Here, an example of model data of one sound source
type designated by num_sound_types_id is illustrated in
Fig. 5. In this example, model data of one sound source
type is described as directivityConfig.

[0069]

25 The model data includes the azimuth angle
"azimuth_table[i]", the elevation angle
"elevation_table[i]", and the radius "distance[i]"
indicating the position of the data point in the original
directivity data before modeling by the number indicated
30 by the number of data points "num_point_indices".

[0070]

The position of the data point is represented by coordinates of a polar coordinate system having the sound source position as an origin, the polar coordinate system including an azimuth angle "azimuth_table[i]" that is an angle of the data point in the horizontal direction viewed from the sound source position, an elevation angle "elevation_table[i]" that is an angle of the data point in the vertical direction viewed from the sound source position, and a radius "distance[i]" that is a distance from the sound source position to the data point.

[0071]

In addition, the model data includes the number of frequency points "bin_count" and the frequency "freq[i_bin]". In the original directivity data before modeling, the entire frequency band of interest is divided into frequency bins, that is, bins, which are frequency bands (frequencies) of the number indicated by the number of frequency points "bin_count", and the center frequency (Hz) of the i-th bin among these bins is the frequency "freq[i_bin]".

[0072]

Thus, the original directivity data before modeling includes a directivity gain for each of one or more bins (frequency bin) at each of the plurality of data points.

[0073]

Further, the model data includes, as parameters related to the Kent distribution and the vMF distribution, the number of bands to be modeled "band_count", the number of mixtures "mix_count[i_band]" in each band, and bin information "bin_range_per_band[i_band]" of the original directivity

data before modeling included in each band.

[0074]

For example, in modeling, the entire target frequency band of interest is divided into bands that are frequency bands whose number is indicated by the number of bands "band_count", and the distribution of the directivity gain is represented by the mixture model for each of the bands. In other words, the model parameter constituting the mixture model representing the distribution of the directivity gain in each band is estimated. Note that the frequency indicated by one or more bins, that is, the center frequency "freq[i_bin]" of the bin is always included in (belongs to) the frequency band indicated by each band.

15 [0075]

The number of mixtures "mix_count[i_band]" indicates the number of distributions constituting the mixture model representing the distribution of the directivity gain of the i-th band, that is, the number of Kent distributions and vMF distributions, and the number of mixtures corresponds to N in Expression (3).

[0076]

The bin information "bin_range_per_band[i_band]" of the directivity data is information indicating a bin of the original directivity data before modeling, the bin included in the i-th band. For example, the bin information is index information indicating a bin of the highest frequency belonging to the i-th band. By referring to the bin information "bin_range_per_band[i_band]", it is possible to identify in which band after modeling the bin (frequency bin) for

the original directivity data before modeling is included.

[0077]

In addition, the model data includes, as parameters
 5 related to the Kent distribution and the vMF
 distribution, the weight ϕ_i , the degree of parameter
 concentration κ , and the vector γ_1 described above for
 each distribution (Kent distribution or vMF distribution)
 constituting the mixture model for each band.

10 [0078]

In this example, "weight[i_band][i_mix]" and
 "kappa[i_band][i_mix]" indicate the weight ϕ_i and the
 degree of parameter concentration κ of the distribution
 indicated by "i_mix" for the i-th band indicated by
 15 "i_band".

[0079]

Further, "gamma1[i_band][i_mix][x]" and
 "gamma1[i_band][i_mix][y]" indicate an X component (X
 coordinate) and a Y component (Y coordinate) constituting
 20 a vector γ_1 of the distribution indicated by "i_mix" for
 the i-th band "i_band".

[0080]

The model data includes a selection flag
 "dist_flag" indicating which distribution of the Kent
 25 distribution or the vMF distribution the distribution
 indicated by "i_mix" for the i-th band "i_band"
 constituting the mixture model is.

[0081]

A value "1" of the selection flag "dist_flag"
 30 indicates that the distribution is the Kent distribution,
 and a value "0" of the selection flag "dist_flag"

indicates that the distribution is the vMF distribution.

[0082]

In a case where the value of the selection flag
 "dist_flag" is "1", the model data includes the
 5 ellipticity β , the major axis vector γ_2 , and the minor
 axis vector γ_3 described above.

[0083]

"beta[i_band][i_mix]" indicates the ellipticity β
 of the distribution (Kent distribution) indicated by
 10 "i_mix" for the i-th band indicated by "i_band". In
 addition, "gamma2[i_band][i_mix][x]" and
 "gamma2[i_band][i_mix][y]" indicate an X component (X
 coordinate) and a Y component (Y coordinate) constituting
 the major axis vector γ_2 of the distribution (Kent
 15 distribution) indicated by "i_mix" for the i-th band
 "i_band".

[0084]

Similarly, "gamma3[i_band][i_mix][x]" and
 "gamma3[i_band][i_mix][y]" indicate an X component (X
 20 coordinate) and a Y component (Y coordinate) constituting
 a minor axis vector γ_3 of the distribution (Kent
 distribution) indicated by "i_mix" for the i-th band
 "i_band".

[0085]

25 The model data also includes the directivity data
 in each bin, more specifically, the scale factor
 "scale_factor[i_bin]" indicating the dynamic range of the
 directivity gain, and the offset value of the directivity
 data (directivity gain) in each bin, that is, the minimum
 30 value "offset[i_bin]".

[0086]

Hereinafter, a parameter set including the ellipticity β , the degree of parameter concentration κ , the weight ϕ_i , the vector γ_1 , the major axis vector γ_2 , the minor axis vector γ_3 , the scale factor, and the
 5 minimum value (offset value) included in the model data is also referred to as a model parameter.

[0087]

The model data also includes difference information "diff_data[i_point]" indicating a difference between the
 10 value (directivity gain) of the original directivity data before modeling and the value (directivity gain) of the directivity data indicated by the mixture model obtained by the modeling at the data point. In other words, the difference information is information indicating a
 15 difference between the unmodeled directivity data and the modeled directivity data at the data point.

[0088]

Note that it may be possible to select whether or not the difference information is stored.
 20 "diff_data[i_point]" stored in the model data may be Huffman encoded difference information.

[0089]

In the device on the reproduction side (decoding side), for example, the output value $F(x; \Theta)$ of the
 25 mixture model at each data point, that is, the directivity gain is calculated on the basis of the model data of the configuration (format) illustrated in Fig. 5.

[0090]

Each bin of the original directivity data before
 30 modeling belongs to any band of the number of bands described by the number of bands "band_count" at the time

of the modeling determined in consideration of the similarity in shape of the directivity data.

[0091]

Furthermore, a relevant relationship between each
 5 bin and a band is described by bin information
 "bin_range_per_band[i_band]", and a maximum index that is
 index information indicating a bin of the highest
 frequency belonging to the band is written as the bin
 information.

10 [0092]

In this case, for example, as illustrated in Fig.
 6, the number of bins belonging to each band may be
 different for each band.

[0093]

15 In this example, two bin 0 (bin 0) and bin 1 belong
 to the first band 0 (band 0) with the lowest frequency,
 one bin 2 belongs to the next band 1, and two bin 3 and
 bin 4 belong to the next band 2.

[0094]

20 Therefore, the value of the bin information
 "bin_range_per_band[i_band]" of the band 0 is a value "1"
 indicating the bin 1, that is, "bin_range_per_band[0] =
 1". Similarly, the value of the bin information about
 the band 1 is "2", that is, "bin_range_per_band[1] = 2",
 25 and the value of the bin information about the band 2 is
 "4", that is, "bin_range_per_band[2] = 4".

[0095]

Since the model data includes a model parameter, a
 mixture model $F'(x; \theta)$ for each band can be obtained from
 30 the model parameters. Here, the mixture model $F'(x; \theta)$
 corresponds to the mixture model $F(x; \theta)$ for each bin

indicated by Expression (3).

[0096]

The directivity data before modeling has a directivity gain value for each bin of each data point.

5 Therefore, the mixture model $F'(x; \Theta)$ for each band obtained from the model parameter, more specifically the output value $F'(x; \Theta)$ of the mixture model, need to be converted to the original mixture model $F(x; \Theta)$ for each bin.

10 [0097]

Therefore, in the device on the reproduction side (decoding side), the output value $F(x; \Theta)$ of the mixture model for each bin at a data point is calculated on the basis of the mixture model $F'(x; \Theta)$ for each band, the scale factor "scale_factor[i_bin]" for each bin, and the minimum value "offset[i_bin]" for each bin.

[0098]

That is, $F(x; \Theta) = F'(x; \Theta) \times \text{scale_factor}[i_bin] + \text{offset}[i_bin]$ is calculated. In this calculation, the output value $F'(x; \Theta)$ of the mixture model for each band is corrected according to the dynamic range of each bin.

[0099]

Further, in a case where the differential compression is used together, that is, in a case where the model data includes the difference information "diff_data[i_point]" for each data point, the output value $F(x; \Theta)$, the difference information is added to the output value $F(x; \Theta)$ obtained by the calculation to make the final output value $F(x; \Theta)$.

[0100]

By the above calculation, the original directivity data before modeling is restored from the model data. Note that, on the reproduction side, the position to be each data point and the frequency of each bin can be identified from the azimuth angle "azimuth_table[i]", the elevation angle "elevation_table[i]", the radius "distance[i]", and the frequency "freq[i_bin]" stored in the model data.

[0101]

10 In practice, Fig. 7 illustrates the data amount of the model data when the directivity data is modeled so that the model data has the configuration illustrated in Fig. 5.

[0102]

15 In this example, the number of data points in the original directivity data before modeling is 2522, and the number of bins is 29. Furthermore, at the time of modeling, the number of bands "band_count" is set to "3", and modeling with a mixture model including the vMF distribution (without ellipticity β , major axis vector γ_2 , minor axis vector γ_3) is performed.

[0103]

25 In the modeling of Fig. 7, it can be seen that the original directivity data having the data amount of 306 KB before modeling is converted into the model parameter having the data amount of 0.85 KB, and the data amount is compressed to about 1/360.

[0104]

30 In addition, in the example of Fig. 5, the model data includes difference information as necessary, and the directivity data is restored using the difference

information as appropriate.

[0105]

That is, for example, in a case where the difference can be perceived from the viewpoint of auditory psychology, difference encoding is used together with modeling of the present technology, and the directivity data is restored to an unperceptible extent.

[0106]

For example, it is assumed that modeling is performed for the directivity data indicated by the arrow Q41 in Fig. 8. Note that, in Fig. 8, the shade of color on each spherical surface indicates the magnitude of the directivity gain.

[0107]

In this example, it is assumed that as a result of modeling the directivity data indicated by the arrow Q41, the mixture model indicated by the arrow Q42, more specifically, the directivity data represented by the mixture model is obtained.

[0108]

In the portion indicated by the arrow Q42, each of a plurality of straight lines drawn on the spherical surface represents the above-described vector γ_1 . For example, the vector V51 represents one vector γ_1 .

[0109]

In a case where the directivity data indicated by the arrow Q41 and the mixture model indicated by the arrow Q42 are obtained, the residual data indicated by the arrow Q43 is obtained as the difference information when the difference between the directivity data and the mixture model is obtained.

[0110]

In the example illustrated in Fig. 5, the value (residual) at each data point of the residual data indicated by the arrow Q43 is stored in the model data as the difference information "diff_data[i_point]".

[0111]

Note that there is a system called Higher Order Ambisonics (HOA) in a directivity expression system. The HOA has an advantage that not only amplitude information but also phase information can be recorded. However, as the shape of the directivity is more complicated, a higher-order term is required, and the amount of data increases. In addition, since the coefficient diverges in the HOA, there are prohibited frequencies that cannot be used.

[0112]

As for the directivity, in general, the shape is more complicated and the degree of protrusion is higher in the high frequency range. In addition, in the high frequency range, the use value of the phase information relatively decreases. Therefore, in the case of reducing the data amount of the directivity data, it is more advantageous to adopt the method of modeling by the mixture distribution model as in the present technology than to use the HOA. Note that, in the low frequency range, the shape of the directivity is relatively gentle, and physical phenomena such as diffraction and interference can be reproduced by recording the phase. Therefore, HOA may be used in the low frequency range, and a method of modeling by a mixture distribution model may be used in the high frequency range.

[0113]

In a case where the model data is transmitted to the reproduction side (decoding side), in the directivity data (amplitude data) generated (restored) on the basis of the model data, a directivity gain exists only at a specific discrete frequency point, that is, a specific bin. In other words, since there is a frequency at which the directivity gain does not exist, the rendering process may not be performed if the directivity data generated from the model data is used as it is.

[0114]

In addition, since the data points are also discretely disposed, when the viewpoint position (sound reception position) of the user or the object moves and the positional relationship between the user and the object changes, the data points of the directivity data used for the rendering process also change. In such a case, glitch (waveform discontinuity) occurs when the interval between the data points adjacent to each other is wide.

[0115]

Therefore, by performing the interpolation process in the frequency direction and the temporal direction on the directivity data, the directivity gain may be obtained for more frequencies (bins) and directions (data points).

[0116]

For example, as the interpolation process in the frequency direction, it is conceivable to perform the primary interpolation process, the secondary interpolation process, or the like using the directivity

gain of the bin indicating a plurality of frequencies in the vicinity of a specific frequency to be obtained.

[0117]

Furthermore, for example, as the interpolation
5 process in the temporal direction, it is conceivable to perform the bilinear interpolation process in the azimuth angle direction or the elevation angle direction using a directivity gain for each bin at a plurality of data points in the vicinity of the direction (position) to be
10 obtained.

[0118]

Note that the influence on the calculation amount and the sound quality at the time of modeling the directivity data varies (trade-off) depending on various
15 parameters such as the frame length (number of samples/frame) of the audio data, the number of mixtures and the model (distribution) to be selected in the mixture model, and the number of data points.

[0119]

20 That is, for example, in a case where the position and direction of the user (listener) or the object change for each frame of the audio data, the occurrence of the waveform discontinuity can be suppressed by performing the interpolation process in the temporal direction, and
25 the audio reproduction with higher quality can be realized.

[0120]

Furthermore, for example, on the reproduction side, it is possible to adjust the balance between the
30 calculation amount and the sound quality by appropriately determining whether to increase the number of mixtures of

the mixture models in order to obtain more accurate directivity data or to use the Kent distribution having a larger number of parameters than the vMF distribution but a higher expression capability.

5 [0121]

Furthermore, the content creator or the like can also determine, for example, whether to increase the number of data points of the directivity data or deal with the small number of data points by the interpolation process at the time of reproduction according to the shape of the directivity of the sound source (object).

[0122]

In addition, difference information indicating an error (difference) between the original directivity data to be modeled (encoded) and the mixture model, that is, the modeled directivity data may be encoded by any encoding method such as Huffman encoding and transmitted.

[0123]

Furthermore, for example, a method of using the directivity data (rendering method) such as the interpolation process in the frequency direction and the interpolation process in the temporal direction, and whether or not to use various types of information such as difference information may be switched by a flag or the like.

[0124]

For example, the flag may enable switching between a parameter for low accuracy for low resource reproduction device or the like and a parameter for high accuracy for high resource reproduction device or the like, that is, switching of parameter accuracy. In such

a case, for example, the parameter is switched according to the resource of the reproduction device, the network environment at the time of content distribution, and the like.

5 [0125]

Note that, although the example in which the present technology is applied to the directivity data is described above, the present technology can also be applied to color, transparency information, and the like for texture data in a video, for example, volumetric point cloud data.

[0126]

Furthermore, in the present technology, for example, in a case where reproduction of a microstructure is important, a content creator or the like may manually (manually) add the number of mixtures models or adjust various parameters such as model parameter.

[0127]

<Configuration example of server>

20 Fig. 9 is a diagram illustrating a configuration example of a server to which the present technology is applied.

[0128]

A server 11 illustrated in Fig. 9 is an information processing device including, for example, a computer and the like, and distributes content.

[0129]

For example, the content includes audio data (object audio data) of each of one or more objects and directivity data that is prepared for each sound source type and represents directivity of a sound source

30

(object), that is, a directional characteristic.

[0130]

Such content can be obtained, for example, by recording directivity data together with a sound of a 3D sound source with a microphone array or the like.
5 Further, the content may include video data corresponding to the audio data.

[0131]

The server 11 includes a modeling unit 21, a model data generation unit 22, an audio data encoding unit 23,
10 and an output unit 24.

[0132]

The modeling unit 21 models the input directivity data of each sound source type, and supplies model
15 parameter and difference information obtained as a result to the model data generation unit 22.

[0133]

The model data generation unit 22 generates model data on the basis of the model parameter and the
20 difference information supplied from the modeling unit 21, and supplies the model data to the output unit 24.

[0134]

The audio data encoding unit 23 encodes the input audio data of each object, and supplies encoded audio
25 data obtained as a result to the output unit 24.

[0135]

The output unit 24 generates and outputs an encoded bit stream by multiplexing the model data supplied from the model data generation unit 22 and the encoded audio
30 data supplied from the audio data encoding unit 23.

[0136]

Note that, in order to simplify the description, an example in which the model data and the encoded audio data are simultaneously output will be described, but the model data and the encoded audio data may be individually
5 generated to output at different timings. In addition, the model data and the encoded audio data may be generated by different devices.

[0137]

<Description of encoding process>

10 Next, an operation of the server 11 will be described. That is, hereinafter, the encoding process by the server 11 will be described with reference to the flowchart in Fig. 10.

[0138]

15 In step S11, the modeling unit 21 models the input directivity data of each sound source type, and supplies the model parameter and the difference information obtained as a result to the model data generation unit 22.

20 [0139]

For example, the modeling unit 21 models the directivity data by representing (expressing) the directivity data with the mixture model including a plurality of distributions shown in the above-described
25 Expression (3).

[0140]

As a result, the degree of parameter concentration κ , the ellipticity β , the weight ϕ_1 , the vector γ_1 , the major axis vector γ_2 , the minor axis vector γ_3 , the scale
30 factor, and the minimum value constituting the mixture model shown in Expression (3) are obtained as model

parameter.

[0141]

In addition, the modeling unit 21 generates information indicating the number of data points, the position of the data points, the number of frequency points, the center frequency of the bin, and the like as the information regarding the original directivity data before modeling.

[0142]

Furthermore, for example, the modeling unit 21 generates, as difference information, a residual (difference) between the modeled directivity data, that is, the directivity data represented by the mixture model, and the original directivity data before modeling.

[0143]

Note that the difference information may be generated, for example, in a case where a specific condition is satisfied, such as a case where a residual between the directivity data represented by the mixture model and the original directivity data is equal to or more than a predetermined value, or in a case where a creator or the like of the content instructs generation of the difference information.

[0144]

The modeling unit 21 supplies the model parameter obtained in this manner, information regarding the original directivity data before modeling, and difference information to the model data generation unit 22.

[0145]

In step S12, the model data generation unit 22 generates model data by packing the model parameter

supplied from the modeling unit 21, the information regarding the original directivity data before modeling, and the difference information, and supplies the model data to the output unit 24.

5 [0146]

At this time, the model data generation unit 22 generates model data in the format illustrated in Fig. 5 by Huffman encoding the difference information and packing the encoded difference information (hereinafter, referred to as difference code data), the model parameter, and the like obtained as a result. Note that the model parameter and the model data may be encoded.

[0147]

In step S13, the audio data encoding unit 23 encodes the input audio data of each object, and supplies the resultant encoded audio data to the output unit 24.

[0148]

Note that, when there is metadata for the audio data of each object, the audio data encoding unit 23 also encodes the metadata of each object (audio data), and supplies encoded metadata obtained as a result to the output unit 24.

[0149]

For example, the metadata includes object position information indicating an absolute position of the object in a three-dimensional space, object direction information indicating an orientation of the object in the three-dimensional space, sound source type information indicating a type of the object (sound source), and the like.

[0150]

In step S14, the output unit 24 multiplexes the model data supplied from the model data generation unit 22 and the encoded audio data supplied from the audio data encoding unit 23 to generate and output an encoded bit stream. When the object includes the metadata, the output unit 24 generates an encoded bit stream including the model data, the encoded audio data, and the encoded metadata.

[0151]

For example, the output unit 24 transmits an encoded bit stream to an information processing device functioning as a client (not illustrated). When the encoded bit stream is transmitted, the encoding process ends.

[0152]

As described above, the server 11 models the directivity data to output the encoded bit stream including the model parameter and the difference information obtained as a result. In this way, it is possible to reduce the data amount of directivity data to be transmitted to the client, that is, the amount of transmission of directivity data. As a result, occurrence of a transmission delay and an increase in the transmission rate can be suppressed.

[0153]

<Configuration example of information processing device>

The information processing device configured to acquire the encoded bit stream output from the server 11 and generate the output audio data for reproducing the sound of the content is configured as illustrated in Fig. 11, for example. An information processing device 51

illustrated in Fig. 11 includes, for example, a personal computer, a smartphone, a tablet, a game device, or the like.

[0154]

5 The information processing device 51 includes an acquisition unit 61, a distribution model decoding unit 62, an audio data decoding unit 63, and a rendering processing unit 64.

[0155]

10 The acquisition unit 61 acquires the encoded bit stream output from the server 11, and extracts the model data and the encoded audio data from the encoded bit stream. The acquisition unit 61 supplies the model data to the distribution model decoding unit 62 and supplies
15 the encoded audio data to the audio data decoding unit 63.

[0156]

 The distribution model decoding unit 62 calculates the directivity data from the model data. The
20 distribution model decoding unit 62 includes an unpacking unit 81, a directivity data calculator 82, a difference information decoding unit 83, an addition unit 84, and a frequency interpolation processing unit 85.

[0157]

25 The unpacking unit 81 performs unpacking on the model data supplied from the acquisition unit 61 to extract the model parameter, the information regarding the original directivity data before modeling, and the difference code data from the model data. In addition,
30 the unpacking unit 81 supplies information regarding the model parameter and the original directivity data before

modeling to the directivity data calculator 82, and supplies difference code data to the difference information decoding unit 83.

[0158]

5 The directivity data calculator 82 calculates (restores) the directivity data on the basis of the information regarding the model parameter supplied from the unpacking unit 81 and the original directivity data before modeling, and supplies the directivity data to the
10 addition unit 84. Note that, hereinafter, the directivity data calculated (restored) on the basis of the model parameter by the directivity data calculator 82 is also referred to as rough directivity data.

[0159]

15 The difference information decoding unit 83 decodes the difference code data supplied from the unpacking unit 81 by a method compatible with Huffman encoding, and supplies difference information obtained as a result to the addition unit 84 as a directivity data residual.

20 [0160]

 The addition unit 84 adds the rough directivity data supplied from the directivity data calculator 82 and the directivity data residual (difference information) supplied from the difference information decoding unit 83
25 to generate directivity data close to the original directivity data, and supplies the generated directivity data to the frequency interpolation processing unit 85.

[0161]

 The frequency interpolation processing unit 85
30 performs the interpolation process in the frequency direction on the directivity data supplied from the

addition unit 84, and supplies directivity data obtained as a result to the rendering processing unit 64.

[0162]

5 The audio data decoding unit 63 decodes the encoded audio data supplied from the acquisition unit 61, and supplies the resultant audio data of each object to the rendering processing unit 64.

[0163]

10 Furthermore, in a case where the encoded metadata is included in the encoded bit stream, the audio data decoding unit 63 decodes the encoded metadata supplied from the acquisition unit 61, and supplies the metadata obtained as a result to the rendering processing unit 64.

[0164]

15 The rendering processing unit 64 generates output audio data on the basis of the directivity data supplied from the frequency interpolation processing unit 85 and the audio data supplied from the audio data decoding unit 63.

20 [0165]

The rendering processing unit 64 includes a directivity data holding unit 86, a head related transfer function (HRTF) data holding unit 87, a temporal interpolation processing unit 88, a directivity convolution unit 89, and an HRTF convolution unit 90.

[0166]

30 The viewpoint position information, the listener direction information, the object position information, and the object direction information are supplied to the directivity data holding unit 86 and the HRTF data holding unit 87 according to designation by the user or

the like, measurement by a sensor or the like, or the like.

[0167]

For example, the viewpoint position information is
5 information indicating a viewpoint position (listening
position) in a three-dimensional space of a user
(listener) viewing the content, and the listener
direction information is information indicating a
direction of a face of the user viewing the content in
10 the three-dimensional space.

[0168]

Furthermore, in a case where the encoded metadata
is included in the encoded bit stream, the object
position information and the object direction information
15 are extracted from the metadata obtained by decoding the
encoded metadata and supplied to the directivity data
holding unit 86 and the HRTF data holding unit 87.

[0169]

In addition, the sound source type information
20 obtained by being extracted from the metadata or the like
is also supplied to the directivity data holding unit 86,
and the user ID indicating the user who views the content
is appropriately supplied to the HRTF data holding unit
87.

25 [0170]

The directivity data holding unit 86 holds the
directivity data supplied from the frequency
interpolation processing unit 85. Further, the
directivity data holding unit 86 reads directivity data
30 corresponding to the viewpoint position information, the
listener direction information, the object position

information, the object direction information, and the sound source type information supplied from the held directivity data, and supplies the directivity data to the temporal interpolation processing unit 88.

5 [0171]

The HRTF data holding unit 87 holds the HRTF for each of a plurality of directions viewed from the user (listener) for each user indicated by the user ID.

[0172]

10 The HRTF data holding unit 87 reads the HRTF corresponding to the viewpoint position information, the listener direction information, the object position information, the object direction information, and the user ID supplied from the held HRTF, and supplies the
15 HRTF to the HRTF convolution unit 90.

[0173]

The temporal interpolation processing unit 88 performs the interpolation process in the temporal direction on the directivity data supplied from the
20 directivity data holding unit 86, and supplies directivity data obtained as a result to the directivity convolution unit 89.

[0174]

The directivity convolution unit 89 convolves the
25 audio data supplied from the audio data decoding unit 63 and the directivity data supplied from the temporal interpolation processing unit 88, and supplies the resultant audio data to the HRTF convolution unit 90. By the convolution of the directivity data, the directional
30 characteristic of the object (sound source) is added to the audio data.

[0175]

The HRTF convolution unit 90 convolves the audio data supplied from the directivity convolution unit 89, that is, the audio data in which the directivity data is convolved, and the HRTF supplied from the HRTF data holding unit 87 to output the audio data obtained as a result as output audio data. By convolution of the HRTF, it is possible to obtain the output audio data in which the sound of the object is localized at the position of the object viewed from the user (listener).

[0176]

<Description of directivity data generation process>

Next, the operation of the information processing device 51 will be described.

15 [0177]

First, the directivity data generation process performed when the information processing device 51 generates directivity data of each sound source type will be described. That is, the directivity data generation process by the information processing device 51 will be described below with reference to the flowchart of Fig. 12.

[0178]

The directivity data generation process is started when the acquisition unit 61 receives the encoded bit stream transmitted from the server 11 and the acquisition unit 61 supplies the model data extracted from the encoded bit stream to the unpacking unit 81.

[0179]

30 In step S51, the unpacking unit 81 unpacks the model data supplied from the acquisition unit 61, and

supplies information regarding the model parameter extracted from the model data and the original directivity data before modeling to the directivity data calculator 82.

5 [0180]

In step S52, the directivity data calculator 82 calculates (generates) rough directivity data on the basis of the information regarding the model parameter supplied from the unpacking unit 81 and the original
10 directivity data before modeling, and supplies the rough directivity data to the addition unit 84.

[0181]

For example, the directivity data calculator 82 calculates the output value $F(x; \theta)$ of the mixture model
15 for each bin at the data point on the basis of the mixture model $F'(x; \theta)$ of each band obtained by the model parameter, the scale factor "scale_factor[i_bin]" for each bin, and the minimum value "offset[i_bin]" for each bin. As a result, rough directivity data including a
20 directivity gain (amplitude data) for each bin at each data point is obtained.

[0182]

In step S53, the unpacking unit 81 determines whether or not the difference code data is included in
25 the model data supplied from the acquisition unit 61, that is, whether or not the difference code data is present.

[0183]

In a case where it is determined in step S53 that
30 the difference code data is included, the unpacking unit 81 extracts the difference code data from the model data

and supplies the difference code data to the difference information decoding unit 83, and thereafter, the process proceeds to step S54.

[0184]

5 In step S54, the difference information decoding unit 83 decodes the difference code data supplied from the unpacking unit 81, and supplies a directivity data residual (difference information) obtained as a result to the addition unit 84.

10 [0185]

 In step S55, the addition unit 84 adds the directivity data residual supplied from the difference information decoding unit 83 to the rough directivity data supplied from the directivity data calculator 82.

15 [0186]

 The addition unit 84 supplies the directivity data obtained by the addition to the frequency interpolation processing unit 85, and thereafter, the process proceeds to step S56.

20 [0187]

 On the other hand, in a case where it is determined in step S53 that the difference code data is not included, the process in steps S54 and S55 is skipped, and then the process proceeds to step S56. In this case,
25 the addition unit 84 supplies the rough directivity data supplied from the directivity data calculator 82 to the frequency interpolation processing unit 85 as restored directivity data as it is.

[0188]

30 When it is determined in step S53 that the difference code data is not included or the process in

step S55 is performed, the process in step S56 is performed.

[0189]

In step S56, the frequency interpolation processing
5 unit 85 performs the interpolation process in the frequency direction on the directivity data supplied from the addition unit 84, and supplies the directivity data obtained by the interpolation process to the directivity data holding unit 86 to hold.

10 [0190]

For example, it is assumed that the audio data of the object is data in a frequency domain, and the audio data has a frequency component value for each of a plurality of frequency bins. In such a case, in the
15 interpolation process in the frequency direction, for example, the interpolation process of calculating a directivity gain of a necessary bin is performed such that the directivity data has a directivity gain for all frequency bins in which the audio data has a frequency
20 component value.

[0191]

Specifically, for example, the frequency interpolation processing unit 85 performs the interpolation process based on the directivity gains of a plurality of bins (frequencies) at a predetermined data
25 point in the directivity data, thereby calculating a directivity gain of a new frequency (bin) at the same data point that does not exist in the original directivity data. By such an interpolation process in
30 the frequency direction, it is possible to obtain directivity data including directivity gains at more

frequencies.

[0192]

When the interpolation process in the frequency direction is performed and the directivity data after the interpolation process is held in the directivity data holding unit 86, the directivity data generation process ends.

[0193]

As described above, the information processing device 51 calculates the directivity data on the basis of the model data. In this way, the data amount of directivity data to be transmitted, that is, the amount of transmission of directivity data can be reduced. As a result, occurrence of a transmission delay and an increase in the transmission rate can be suppressed.

[0194]

<Description of output audio data generation process>

Next, the output audio data generation process performed by the information processing device 51 will be described with reference to the flowchart in Fig. 13. This output audio data generation process is performed at any timing after the directivity data generation process described with reference to Fig. 12 is performed.

[0195]

In step S81, the audio data decoding unit 63 decodes the encoded audio data supplied from the acquisition unit 61, and supplies the resultant audio data to the directivity convolution unit 89. For example, audio data in a frequency domain is obtained by decoding.

[0196]

Note that, in a case where the encoded metadata is supplied from the acquisition unit 61, the audio data decoding unit 63 decodes the encoded metadata, and supplies the object position information, the object
5 direction information, and the sound source type information included in the metadata obtained as a result to the directivity data holding unit 86 and the HRTF data holding unit 87 as appropriate.

[0197]

10 Further, the directivity data holding unit 86 supplies the directivity data corresponding to the viewpoint position information, the listener direction information, the object position information, the object direction information, and the sound source type
15 information supplied to the temporal interpolation processing unit 88.

[0198]

For example, the directivity data holding unit 86 identifies the relationship between the object and the
20 viewpoint position (listening position) of the user in the three-dimensional space from the viewpoint position information, the listener direction information, the object position information, and the object direction information, and identifies the data point corresponding
25 to the identifying result.

[0199]

As an example, for example, when a direction from the object to the viewpoint position is a viewpoint position direction, a position on the spherical surface
30 of the mixture model in the viewpoint position direction when viewed from the center of the mixture model is

identified as the target data point position. Note that there may be no actual data point at the target data point position.

[0200]

5 The directivity data holding unit 86 extracts a directivity gain of each bin at a plurality of data points near the identified target data point position from the directivity data of the sound source type indicated by the sound source type information.

10 [0201]

 Then, the directivity data holding unit 86 supplies data including the directivity gains of the respective bins at the plurality of extracted data points to the temporal interpolation processing unit 88 as directivity data according to the relationship between the positions and directions of the object and the user (listener).

[0202]

 Further, the HRTF data holding unit 87 supplies the HRTF corresponding to the viewpoint position information, the listener direction information, the object position information, the object direction information, and the user ID supplied to the HRTF convolution unit 90.

[0203]

25 Specifically, for example, the HRTF data holding unit 87 identifies the relative direction of the object viewed from the listener (user) as the object direction on the basis of the viewpoint position information, the listener direction information, the object position information, and the object direction information. Then, 30 the HRTF data holding unit 87 supplies an HRTF in the direction corresponding to the object direction among the

HRTFs in the respective directions corresponding to the user ID to the HRTF convolution unit 90.

[0204]

In step S82, the temporal interpolation processing
5 unit 88 performs the interpolation process in the
temporal direction on the directivity data supplied from
the directivity data holding unit 86, and supplies
directivity data obtained as a result to the directivity
convolution unit 89.

10 [0205]

For example, the temporal interpolation processing
unit 88 calculates the directivity gain of each bin at
the target data point position by the interpolation
process on the basis of the directivity gain of each bin
15 at the plurality of data points included in the
directivity data. That is, the directivity gain at a new
data point (target data point position) different from
the original data point is calculated by the
interpolation process.

20 [0206]

The temporal interpolation processing unit 88
supplies data including the directivity gain of each bin
at the target data point position to the directivity
convolution unit 89 as directivity data obtained by the
25 interpolation process in the temporal direction.

[0207]

In step S83, the directivity convolution unit 89
convolves the audio data supplied from the audio data
decoding unit 63 and the directivity data supplied from
30 the temporal interpolation processing unit 88, and
supplies the resultant audio data to the HRTF convolution

unit 90.

[0208]

In step S84, the HRTF convolution unit 90 convolves the audio data supplied from the directivity convolution unit 89 and the HRTF supplied from the HRTF data holding unit 87 to output the output audio data obtained as a result.

[0209]

In step S85, the information processing device 51 determines whether or not to end the processing.

[0210]

For example, in a case where encoded audio data of a new frame is supplied from the acquisition unit 61 to the audio data decoding unit 63, it is determined in step S85 that the process is not ended. On the other hand, for example, in a case where the encoded audio data of the new frame is not supplied from the acquisition unit 61 to the audio data decoding unit 63 and the output audio data of all the frames of the content is generated, it is determined in step S85 that the process is ended.

[0211]

In a case where it is determined in step S85 that the process is not yet ended, thereafter, the process returns to step S81, and the above-described process is repeatedly performed.

[0212]

On the other hand, in a case where it is determined in step S85 that the process is ended, the information processing device 51 ends the operation of each unit and the output audio data generation process ends.

[0213]

As described above, the information processing device 51 selects appropriate directivity data and HRTF, and convolves the directivity data and the HRTF in the audio data to obtain output audio data. By doing so, it is possible to realize high-quality audio reproduction with a more realistic feeling in consideration of the directional characteristic of the object (sound source) and the relationship between the positions and orientations of the object and the listener.

10 [0214]

<Second embodiment>

<Encoding of difference information>

Meanwhile, the directivity data has different directivity shapes for each sound source type and each frequency band.

15 [0215]

In addition, in the server 11, difference information indicating a difference between the unmodeled directivity data and the modeled directivity data is appropriately generated. In the above example, it is described that the difference information is encoded by an encoding method such as Huffman encoding, and the difference code data is obtained.

[0216]

25 The method of encoding the difference information may be selected by the server 11, that is, the encoder, so that the appropriate encoding may be performed according to the sound source type and the frequency band for not only the modeling based on the Kent distribution and the vMF distribution but also the encoding of the difference information.

30

[0217]

Here, a case where the difference information is Huffman encoded will be described as an example.

[0218]

5 In Huffman encoding, for example, as illustrated in Fig. 14, a distribution of appearance probability (probability density function) is generated on the basis of difference information for each of a plurality of bins obtained from one piece of directivity data to be
10 encoded.

[0219]

Note that, in Fig. 14, the horizontal axis represents the value (dB value) of the difference information, and the vertical axis represents the
15 appearance probability of each value of the difference information.

[0220]

For example, all bins (frequencies) at all data points of the directivity data are targeted, and a
20 histogram is generated from the difference information about each bin, so that the appearance probability of each value of the difference information is obtained. Note that the distribution of the appearance probability (probability density function) may be obtained for each
25 bin, may be obtained for a bin included in a specific frequency band, may be obtained for all bins, or any of them may be selectable.

[0221]

In the server 11, one appropriate Huffman encoding
30 table is selected from a plurality of Huffman encoding tables prepared in advance or one new Huffman encoding

table is generated on the basis of the appearance probability of the difference information.

[0222]

5 All bins (frequencies) at all data points of the directivity data are targeted, and one Huffman encoding table may be selected or generated for all those bins, or one Huffman encoding table may be selected or generated for one or more bins.

[0223]

10 The Huffman encoding of the difference information is performed using the Huffman encoding table selected or generated in this manner.

[0224]

15 The Huffman encoding table is a table for converting data before encoding into a Huffman code, the table indicating a correspondence between data before encoding, that is, difference information and a Huffman encode (code data) obtained by encoding.

[0225]

20 In addition, when the difference code data obtained by Huffman encoding the difference information is decoded, the reverse table corresponding to the Huffman encoding table is used.

[0226]

25 The reverse table is a table for converting a Huffman code into decoded data, the table indicating a correspondence between the Huffman code (code data) and the decoded data. This reverse table can be generated from a Huffman encoding table.

30 [0227]

In a case of Huffman encoding the difference

information, both the server 11 (encoder) and the information processing device 51 (decoding unit) may hold the Huffman encoding table in advance. In such a case, the server 11 notifies the information processing device 51 of the ID information indicating the Huffman encoding table used for Huffman encoding the difference information.

[0228]

Furthermore, the server 11 may store the Huffman encoding table or the reverse table in the encoded bit stream to it to the information processing device 51.

[0229]

Specifically, since the size (data amount) of the reverse table is large, the Huffman encoding table may be transmitted from the server 11 to the information processing device 51, and the information processing device 51 may generate the reverse table on the basis of the Huffman encoding table at the time of decoding or the like.

[0230]

In addition, in the distribution of the appearance probability (probability density function), there are a value of difference information with a low appearance probability (appearance frequency) and a value of difference information with a high appearance probability. Therefore, for example, a range corresponding to data of a narrow dynamic range including a value of difference information having a high appearance probability, such as a range of ± 3 dB as a range of possible values of difference information, may be set as a target range, and a Huffman encoding table

for only the target range may be used.

[0231]

In such a case, for difference information about a value outside the target range, that is, difference
5 information about an irregular value having a low appearance probability, the difference information may be stored as it is in the model data. In other words, the difference information is treated as the difference code data as it is.

10 [0232]

As described above, the efficient Huffman encoding table is selected or generated according to the probability density function of the difference information, and the information regarding which Huffman
15 encoding table is used is described in the encoded bit stream, so that the difference information can efficiently be encoded and transmitted.

[0233]

Furthermore, in encoding the difference
20 information, by using one or more methods in combination, the dynamic range can be further reduced, and the encoding efficiency can be improved. Specifically, multi-stage difference encoding can be implemented by combining a plurality of methods.

25 [0234]

For example, in the multi-stage difference encoding, it is conceivable to perform encoding by combining at least two or more of the spatial adjacency difference method, the inter-frequency difference method,
30 or the complex difference method.

[0235]

In addition, for example, a mode indicating the presence or absence and the method of the multi-stage difference encoding is recorded as `enc_mode` or the like in the model data. At this time, for example, in a case

5 where the multi-stage difference encoding method is recorded in the lower 4 bits and which of a real number or a complex number the target is is recorded in the upper 4 bits, the following information is stored in the model data.

10 [0236]

(Target data is real number)

0x00: no multi-stage difference encoding

0x01: spatial adjacency difference method

0x02: inter-frequency difference method

15 0x03: spatial adjacency difference method + inter-frequency difference method

(Target data is complex number)

0x1*: lower bits are the same as those in the case where the target data is real number

20 [0237]

In the spatial adjacency difference method, when difference information at a data point to be processed is encoded, a difference between difference information at the data point to be processed and difference information

25 at another data point in the vicinity of the data point to be processed is obtained as spatial difference information. For example, a difference in difference information between data points adjacent to each other is obtained as the spatial difference information. Then,

30 the obtained spatial difference information is Huffman encoded to be difference code data.

[0238]

In the spatial adjacency difference method, a property that data at a spatially close position (data point) in the directivity data, that is, a directivity gain and difference information, easily take close values is used.

[0239]

In the inter-frequency difference method, when difference information about a bin (frequency) to be processed is encoded, a difference between difference information in a bin to be processed and difference information in another bin such as a bin adjacent to the bin to be processed, the another bin indicating a close frequency, is obtained as inter-frequency difference information. Then, the obtained inter-frequency difference information is Huffman encoded to be difference code data.

[0240]

In the inter-frequency difference method, a property that data of a close frequency (bin), that is, a directivity gain or difference information can easily take close values is used.

[0241]

For example, in a case where the spatial adjacency difference method and the inter-frequency difference method are used in combination, the difference in the spatial difference information between the adjacent bins is obtained as the inter-frequency difference information, and the inter-frequency difference information is Huffman encoded, or the difference in the inter-frequency difference information between the

adjacent data points is obtained as the spatial difference information, and the spatial difference information is Huffman encoded.

[0242]

5 The complex difference method is used in a case where the directivity data has not only the information regarding the amplitude described above but also the information regarding the phase.

[0243]

10 For example, in a case where the directivity data has information regarding the amplitude and the phase, the information regarding the amplitude and the phase, that is, the directivity gain is expressed by a complex number. That is, the directivity data has complex number
15 data (hereinafter, also referred to as complex directivity gain) indicating the amplitude and the phase for each bin for each data point, and the difference information is also complex number data.

[0244]

20 In the complex difference method, a real part and an imaginary part of difference information represented by a complex number are independently (individually) Huffman encoded, or Huffman encoding is performed on two-dimensional data (complex directivity gain) including the
25 real part and the imaginary part. Note that, in the complex difference method, it may be possible to select whether to individually perform Huffman encoding on each of the real part and the imaginary part or to perform Huffman encoding on two-dimensional data.

30 [0245]

Hereinafter, each method of performing encoding by

combining at least one or more methods of the spatial
adjacency difference method, the inter-frequency
difference method, or the complex difference method, and
a method of Huffman encoding the difference information
5 as it is are also referred to as one difference encoding
method or difference encoding mode. Specifically, it can
be said that the difference encoding method in which the
difference information is directly Huffman encoded is a
method in which encoding using a difference, that is,
10 difference encoding is not performed.

[0246]

For example, the server 11 selects the most
efficient method from the plurality of difference
encoding methods (difference encoding mode) on the basis
15 of the difference information and the like, and Huffman
encodes the difference information by the selected
difference encoding method.

[0247]

Specifically, for example, the code amount (data
20 amount) of the difference code data in each difference
encoding method may be obtained by an operation based on
the difference information, and the method having the
smallest code amount among the difference encoding
methods may be selected as the most efficient method.

25 [0248]

In addition, an appropriate difference encoding
method may be selected on the basis of, for example, a
sound source type of directivity data, an environment at
the time of recording directivity data such as an
30 anechoic chamber, or the like.

[0249]

<Representation method of directivity data>

The example in which the directivity data is modeled by the mixture model (mixture distribution model) including the Kent distribution and the vMF distribution has been mainly described above.

[0250]

In addition, for example, it is described that the directivity data may be modeled by the HOA and the information regarding the phase may also be recorded in the low frequency range, that is, the low frequency bin, and the directivity data may be modeled by the mixture model including the Kent distribution and the vMF distribution in the high frequency range in which the degree of importance of the phase is relatively low, that is, the high frequency bin. In this case, for example, it is conceivable to switch between modeling by the mixture model and modeling by the HOA at a predetermined frequency in the vicinity of 1.5 kHz to 2 kHz. For example, intensity stereo that does not use phase information in an audio codec or the like is used in the above band or higher. The method of combining the HOA and the mixture model as described above is considered to be effective in a case of having sharp front directivity, for example, a whistle or a trumpet.

[0251]

Note that the present invention is not limited thereto, and the model data may be generated by combining at least one or more methods of the HOA method, the mixing method, the complex mixing method, or the difference method for each frequency band, that is, for each bin or band, or in common for all frequency bands.

In such a case, for example, the directivity data is modeled by one or more methods different from each other such as the HOA method and the mixing method, and model data including model parameter and the like obtained as a
5 result thereof is generated.

[0252]

The HOA method is a method of modeling directivity data including a complex directivity gain for each bin at each data point using HOA. That is, the HOA method is a
10 method of modeling the directivity data by spherical harmonic function expansion.

[0253]

Specifically, in the HOA method, spherical harmonic function expansion is performed on the directivity data,
15 and as a result, a spherical harmonic coefficient that is a coefficient for the spherical harmonic function of each dimension is obtained as a model parameter. From the spherical harmonic coefficient of each dimension, it is possible to obtain the directivity data including the
20 complex directivity gain after modeling by the HOA.

[0254]

As described above, in the modeling by the HOA method, the expression including the phase is possible, but in order to perform the fine expression, it is
25 necessary to increase the order of the spherical harmonic function expansion, that is, to obtain the spherical harmonic coefficient up to the high-order term, and in such a case, the data amount of the model data increases. Specifically, in the modeling by the HOA method, it is
30 not possible to finely express the distribution of the amplitude and the phase only in a specific azimuth

(direction).

[0255]

Conversely, in a case where the spherical harmonic coefficient is obtained only for the low-order term, only
5 a relatively gentle change in amplitude or phase can be described.

[0256]

The mixing method is a method of performing modeling by a mixture model including the Kent
10 distribution and the vMF distribution described above. In the mixing method, it is possible to describe the shape of the directivity gain that changes drastically in a specific azimuth (direction) seen from the sound source, that is, the position of the data point.

15 [0257]

The complex mixing method is a method of modeling directivity data including a complex directivity gain, that is, amplitude and phase data, using a mixture distribution (mixture model) corresponding to a complex
20 number.

[0258]

As an example of the complex mixing method, for example, modeling by the following two methods is considered.

25 [0259]

First, as a first method, it is conceivable to perform modeling by describing each of a real part and an imaginary part of a complex directivity gain or each of an amplitude and a phase angle obtained from the complex
30 directivity gain independently with a mixture model of a probability density function for a real number.

[0260]

As a second method, a method of performing modeling by describing directivity data (distribution of complex directivity gains) using a complex Bingham distribution mixture model corresponding to a complex number, a
 5 complex watson distribution mixture model, or the like is considered.

[0261]

In this case, for example, the directivity data is
 10 modeled by a mixture model including one or more complex Bingham distributions or a mixture model including one or more complex watson distributions, and as a result, the model parameter similar to that in the case of the mixing method is obtained. Directivity data including the
 15 complex directivity gain after modeling by the complex mixing method can be obtained from the model parameter obtained in this manner.

[0262]

As an example, in a case where the distribution of
 20 the target complex number data is described as the complex Bingham distribution as it is, the description is made in a format illustrated in the following Expression (5). That is, the value $f(z)$ of the complex Bingham distribution is expressed by the following Expression
 25 (5).

[0263]

[Math 5]

$$f(z) = C(A)^{-1} \exp(z^* A z) \quad \cdot \cdot \cdot (5)$$

[0264]

30 The complex number vector z in Expression (5)

corresponds to the position vector x of the spherical surface in the Kent distribution or the VMF distribution, and z^* is its complex conjugate. The complex matrix A is a $k \times k$ dimensional matrix indicating a position, steepness, a direction, and a shape, and the normalization coefficient $C(A)$ is expressed by the following Expression (6).

[0265]

[Math 6]

$$C(A) = 2\pi^k \sum_{j=1}^k a_j \exp(\lambda_j) \quad \dots (6)$$

[0266]

where, the definition of a_j is as in the following Expression (7). λ_j is an eigenvalue of the complex matrix A , and $\lambda_1 < \lambda_2 < \lambda_3 < \dots \lambda_k$.

[0267]

[Math 7]

$$a_j = \frac{1}{\prod_{i \neq j} (\lambda_j - \lambda_i)} \quad \dots (7)$$

[0268]

The number of mixtures and weights in the mixture model including one or more complex Bingham distributions, that is, the complex Bingham mixture model, are common to the formulation of the mixture model including the Kent distribution and the VMF distribution described above. The value $F(x; \theta)$ of the mixture model including N complex Bingham distributions $f(z; \theta_i)$ can be described by weighting according to the following Expression (8). Note that, as illustrated in Expression (9), the sum of the weights is 1, θ represents a set of all parameters, θ_i represents a set of parameters of each

complex Bingham distribution (parameters constituting the complex Bingham distribution), and ϕ_i represents a weight for each complex Bingham distribution.

[0269]

5 [Math. 8]

$$F(x; \Theta) = \sum_{i=1}^N \phi_i f(z; \theta_i) \quad \cdot \cdot \cdot (8)$$

[Math. 9]

$$\sum_{i=1}^N \phi_i = 1 \quad \cdot \cdot \cdot (9)$$

[0270]

10 The difference method is a method of generating model data using the difference.

[0271]

For example, in a case where the model data is generated by combining one or more other methods such as the HOA method and the mixing method and the difference method, in the difference method, difference information indicating a difference between the directivity data before modeling and the directivity data after modeling by the one or more of other methods is encoded by the any difference encoding method, and the difference code data obtained as a result is stored in the model data. Note that the difference of the directivity data obtained by the difference method may be modeled by the HOA method or the like.

25 [0272]

In the difference method, for example, at least any one of a difference between spatial positions (between data points) or a difference between frequencies (between bins or bands) is obtained for difference information,

and the difference obtained as a result is Huffman encoded to be difference code data. At this time, in a case where the difference of the difference information to be Huffman encoded is a complex number, each of the
5 real part and the imaginary part of the difference may be individually Huffman encoded, the complex number may be Huffman encoded as it is, or each of the amplitude component and the phase component obtained from the difference may be individually Huffman encoded.

10 [0273]

On the other hand, in a case where only the difference method is used, that is, in a case where the model data is generated by the difference method alone, the encoding (Huffman encoding) on the directivity data
15 is performed as in the any difference encoding method described above.

[0274]

At this time, at least one or more of the spatial adjacency difference method, the inter-frequency
20 difference method, or the complex difference method including at least one method of the spatial adjacency difference method and the inter-frequency difference method are used. That is, the difference in the directivity gain between the positions (between the data
25 points) and between the frequencies (between the bins or between the bands) in the space is obtained, and the difference is Huffman encoded. When the difference is expressed by a complex number, the real part and the imaginary part of the difference may be individually
30 Huffman encoded, or the difference (complex number) may be Huffman encoded. Further, each of the amplitude

component and the phase component obtained from the difference may be individually Huffman encoded.

[0275]

Furthermore, model data including data
5 (hereinafter, also referred to as encoded directivity data) including a Huffman code obtained by Huffman encoding for the difference obtained by the difference method is generated. In this case, since there is no residual in the directivity data, the model data does not
10 include the difference code data.

[0276]

Note that, in the difference method, lossless compression is possible, but the compression rate varies depending on data. In addition, in a case where
15 processing of obtaining a multi-stage difference by combining a plurality of methods such as the spatial adjacency difference method and the inter-frequency difference method, that is, obtaining the differences is performed a plurality of times, unlike the case of one-
20 dimensional data, it is necessary to define the data order when the difference code data and the encoded directivity data are stored in the model data, and the compression rate varies depending on the data order.

[0277]

25 In generating the model data from the directivity data, it is also conceivable to model the average value of the directivity gains for each bin or each band of the directivity data, that is, the average directional characteristic. In such a case, the difference
30 information is calculated after applying the offset and the scale factor to the average directional

characteristic and matching the dynamic range.

[0278]

In a case where model data is generated by combining the HOA method, the mixing method, the complex
5 mixing method, and the difference method as described above, classification into the following five methods is mainly considered as a method of generating model data.

[0279]

The five methods mentioned herein are a band hybrid
10 method, an addition hybrid method, a multiplication hybrid method, a spherical harmonic coefficient modeling method, and a combination hybrid method. Each method will be described below.

[0280]

15 (Band hybrid method)

The band hybrid method is a method of switching whether to generate the model data using which of the HOA method, the mixing method, the complex mixing method, and the difference method for each frequency band, that is,
20 for each bin or each band. In this case, for example, recording with a complex directivity gain may be performed in the low frequency range, and recording with a real number directivity gain may be performed in the high frequency range.

25 [0281]

As a specific example, for example, it is possible to model the directivity data by a method different for each band (frequency band), such as performing modeling by the HOA method in the low frequency band and
30 performing modeling by the mixing method in the high frequency band.

[0282]

Furthermore, for example, modeling by a complex mixing method using a complex Bingham distribution or the like may be performed in a low frequency band, and
 5 modeling by a mixing method may be performed in a high frequency band.

[0283]

(Addition hybrid method)

In the addition hybrid method, difference
 10 information indicating a difference from the modeled directivity data is further modeled or encoded by the difference method.

[0284]

Specific examples of the addition hybrid method
 15 include the following methods (AH1) to (AH4). Specifically, in the following example, the process is executed in order from the method described on the left side.

[0285]

20 Method (AH1): Mixing method + difference method
 Method (AH2): HOA method (low order) + mixing method
 Manner (AH3): HOA Manner (low order) + difference method
 25 Method (AH4): HOA method (low order) + mixing method + difference method

[0286]

In method (AH1), the directivity data is first modeled in a mixing method. Next, difference information
 30 indicating a difference between the directivity data before modeling and the directivity data after modeling

by the mixing method is encoded by the difference method,
and difference code data is generated.

[0287]

Then, model data including the model parameter
5 obtained by modeling by the mixing method and the
difference code data is generated.

[0288]

In method (AH2), first, the directivity data is
modeled in the HOA method. Specifically, in the modeling
10 by the HOA method, spherical harmonic function expansion
up to a low-order term is performed. Next, difference
information indicating a difference between the
directivity data before modeling and the directivity data
after modeling by the HOA method is further modeled by
15 the mixing method.

[0289]

Then, model data including the model parameter
obtained by modeling by the HOA method and the model
parameter obtained by modeling the difference information
20 by the mixing method is generated.

[0290]

In the method (AH3), as in the method (AH2),
modeling up to a lower-order term is performed by the HOA
method, and then difference information obtained by
25 modeling by the HOA method is encoded by the difference
method, and difference code data is generated.

[0291]

Then, model data including the model parameter
obtained by modeling by the HOA method and the difference
30 code data is generated.

[0292]

In the method (AH4), as in the method (AH2), modeling up to the lower-order term is performed by the HOA method, and then modeling of the difference information is further performed by the mixing method.

5 [0293]

Next, difference information indicating a difference between difference information obtained by modeling by the HOA method and difference information after modeling by the mixing method is encoded by the difference method, and difference code data is generated. In other words, difference information indicating a difference between the directivity data after modeling modeled by the combination of the HOA method and the mixing method and the directivity data before modeling is encoded by the difference method, and difference code data is generated.

[0294]

Then, model data including a model parameter obtained by modeling by the HOA method, a model parameter obtained by modeling the difference information by the mixing method, and difference code data is generated.

[0295]

Hereinafter, the difference information to be modeled is also particularly referred to as intermediate difference information in order to further distinguish the difference information to be modeled from the difference information to be encoded in the difference method after modeling the directivity data in the predetermined method.

30 [0296]

For example, in the method (AH4), difference

information obtained by modeling by the HOA method is intermediate difference information, and the intermediate difference information is modeled by the mixing method. Then, difference information indicating a difference
5 between the original intermediate difference information and the intermediate difference information after modeling by the mixing method is encoded by the difference method.

[0297]

10 Among the above methods (AH1) to (AH4), data that completely matches the original directivity data cannot be obtained on the decoding side in the method (AH2), but data that completely matches the original directivity data is obtained in the method (AH1), the method (AH3),
15 and the method (AH4).

[0298]

 In addition, the directivity data may be modeled or encoded by a single method instead of the addition hybrid method. That is, for example, the directivity data may
20 be modeled or encoded by only any one of the HOA method, the mixing method, and the difference method, and the model data including the model parameter or the encoded directivity data obtained as a result thereof may be generated.

25 [0299]

(Multiplication hybrid method)

 In the multiplication hybrid method, the directivity data is modeled by a predetermined method, and the ratio (quotient) of the modeled directivity data
30 and the unmodeled directivity data is further modeled by another method different from the predetermined method.

[0300]

Specific examples of the multiplication hybrid method include the following methods (MH1) and (MH2).

[0301]

5 Method (MH1): HOA method (low order) × amplitude modulation (mixing method)

 Method (MH2): HOA method (low order) × amplitude phase modulation (mixing method)

[0302]

10 In the method (MH1), first, the directivity data is modeled by the HOA method. Specifically, in the modeling by the HOA method, spherical harmonic function expansion up to a low-order term is performed.

[0303]

15 Next, a value (hereinafter, also referred to as amplitude modulation information) obtained by dividing the directivity data before modeling by the directivity data after modeling by the HOA method is further modeled by the mixing method. At this time, for example, the
20 absolute value (amplitude component) of the complex number (complex directivity gain) constituting the amplitude modulation information may be a value for modeling by the mixing method, or the ratio of the amplitude components of the directivity data before and
25 after modeling may be the amplitude modulation information. Then, model data including the model parameter obtained by modeling by the HOA method and the model parameter obtained by modeling the amplitude modulation information by the mixing method is generated.

30 [0304]

 At the time of decoding, the directivity data

calculated from the model parameter for the HOA method is multiplied by the amplitude modulation information calculated from the model parameter for the mixing method, and the final directivity data is calculated.

5 [0305]

In such a method (MH1), amplitude modulation information indicating a small amplitude swing according to a high frequency azimuth (direction from a sound source), the information not being able to be expressed
10 by modeling up to a low-order term in the HOA method, is modeled by the mixing method and recorded (stored) in model data. At the time of decoding, the directivity data calculated from the model parameter for the HOA method is modulated by the amplitude modulation
15 information, and the directivity data with less error is obtained.

[0306]

In the method (MH2), as in the method (MH1), modeling up to the lower-order term in the HOA method is
20 performed on the directivity data.

[0307]

Next, a value (hereinafter, also referred to as amplitude phase modulation information) obtained by dividing the directivity data before modeling by the
25 directivity data after modeling by the HOA method is further modeled by the mixing method. At this time, for example, the real part and the imaginary part of the complex number (complex directivity gain) constituting the amplitude phase modulation information, or the
30 amplitude component and the phase component are to be modeled by the mixing method. Note that the amplitude

phase modulation information may be modeled by a complex mixing method. Then, model data including the model parameter obtained by modeling by the HOA method and the model parameter obtained by modeling the amplitude phase modulation information by the mixing method is generated.
5 [0308]

At the time of decoding, the directivity data calculated from the model parameter for the HOA method is multiplied by the amplitude phase modulation information
10 calculated from the model parameter for the mixing method, and the final directivity data is calculated.
[0309]

In such a method (MH2), the amplitude phase modulation information indicating the rotational change
15 in the high frequency phase according to the azimuth (direction from the sound source), the information not being able to be expressed by the modeling up to the low-order term in the HOA method, is modeled by the mixing method and recorded (stored) in the model data. At the
20 time of decoding, the directivity data calculated from the model parameter for the HOA method is modulated by the amplitude phase modulation information, and the directivity data with less error is obtained.
[0310]

25 In the multiplication hybrid method or another method, in a case where a directivity gain (complex directivity gain) or intermediate difference information expressed by a complex number is modeled, modeling may be performed independently (individually) by different or
30 the same method between a real part and an imaginary part of the complex number. For example, the real part may be

modeled by the mixing method, and the imaginary part may also be modeled by the mixing method.

[0311]

Similarly, the amplitude component and the phase
5 component may be independently (individually) modeled by any method, or data of a complex number may be modeled by a complex mixing method.

[0312]

(Spherical harmonic coefficient modeling method)

10 In the spherical harmonic coefficient modeling method, the directivity data is modeled by the HOA method, model parameter obtained as a result thereof, that is, the spherical harmonic coefficient, is further modeled by the mixing method, and the model parameter
15 obtained as a result thereof is stored in the model data.

[0313]

Therefore, in the spherical harmonic coefficient modeling method, it can be said that the directivity data is modeled in two stages of the HOA method and the mixing
20 method. At the time of decoding, first, the spherical harmonic coefficient is calculated on the basis of the model parameter for the mixing method, and further, the directivity data (rough directivity data) is calculated on the basis of the spherical harmonic coefficient.

25 [0314]

In addition, for example, each of the real part and the imaginary part of the spherical harmonic coefficient as the model parameter, or each of the amplitude component and the phase component obtained from the model
30 parameter may be individually (independently) modeled by any method such as a mixing method. In addition, the

spherical harmonic coefficient may be modeled by a complex mixing method, that is, one or more complex Bingham distributions or the like.

[0315]

5 (Combination hybrid method)

In the combination hybrid method, model data is generated using a combination of at least two or more of the band hybrid method, the addition hybrid method, the multiplication hybrid method, or the spherical harmonic
10 coefficient modeling method described above.

[0316]

In addition, for example, information indicating a combination of one or more methods used for generating the model data, such as the HOA method and the mixing
15 method, may be stored in the model data. In such a case, a combination of one or more methods used for generating model data can be appropriately selected and switched on the server 11 side.

[0317]

20 In a case where the directivity data is modeled as described above, the model data has a configuration illustrated in Figs. 15 and 16, for example. Note that Fig. 16 illustrates a portion following the portion illustrated in Fig. 15. Furthermore, in Figs. 15 and 16,
25 description of portions corresponding to those in the case illustrated in Fig. 5 will be appropriately omitted.

[0318]

The example illustrated in Figs. 15 and 16 is an example in which the directivity information (directivity
30 data) of one type of sound source designated by num_sound_types_id is described as directivityConfig.

Specifically, here, the vMF distribution, the Kent distribution, and Syntax in a case where difference data (difference information) exists are illustrated as examples of implementing the hybrid method, and the bit
5 depth of each piece of information is merely an example.
[0319]

The model data illustrated in Figs. 15 and 16 basically includes the same data as the model data illustrated in Fig. 5, but the examples of Figs. 15 and
10 16 are different from the example of Fig. 5 in the bit depth and the data configuration of some pieces of data.
[0320]

Specifically, in the examples illustrated in Figs. 15 and 16, the azimuth angle "azimuth_table[i]" and the
15 elevation angle "elevation_table[i]" are 16-bit unsigned shorts.
[0321]

In addition, the number of bands "band_count" and the number of mixtures "mix_count[i_band]" are 8-bit
20 unsigned char, and the selection flag "dist_flag" is set as 1-bit bool.
[0322]

Further, in this example, the model data includes the ID of the hybrid mode (difference encoding mode
25 (difference encoding method)) used for encoding the difference information, that is, "mode" indicating the difference encoding mode information. The model data also includes an index "table_index" indicating a Huffman encoding table used for encoding the difference
30 information.
[0323]

The model data further includes "int db_resolution" indicating a quantization step size such as quantization every 1.0 dB. For example, for "int db_resolution", the value "0" indicates no quantization, the value "1" indicates 0.01 dB, the value "2" indicates 0.2 dB, the value "3" indicates 0.4 dB, and the value "256" indicates 25.6 dB.

[0324]

In addition, the model data also stores a Huffman encode (Huffman encode) obtained by Huffman encoding the difference information for each data point for each bin, that is, "diff_data[i_bin][i_point]" which is difference code data.

[0325]

Furthermore, the information stored in the model data or about the configuration illustrated in Fig. 17 separately from the model data is transmitted from the server 11 to the information processing device 51. The information illustrated in Fig. 17 includes a Huffman encoding table or a reverse table.

[0326]

In the example illustrated in Fig. 17, "diff_mode_count" is information indicating the total number of difference encoding methods, and "int_nbits_res_data" is stored by the total number "diff_mode_count".

[0327]

This "int_nbits_res_data" is information indicating the maximum bit depth of the Huffman code, that is, the maximum word length of the Huffman code, and is, for example, 7 bits in the case of 1.0 dB increments, and can

express a range from 0 dB to 128 dB.

[0328]

"element_count" is information indicating the number of elements of the Huffman encoding table or the reverse table, and "Huff_dec_table[i_element]", which is
 5 an element corresponding to the number of elements, is stored. Specifically, in this example, "Huff_dec_table[i_element]" is an element of the reverse table.

10 [0329]

Furthermore, the Huffman encoding table is as illustrated in Fig. 18, for example. That is, Fig. 18 illustrates a specific example of the Huffman encoding table.

15 [0330]

For example, as a specific example, in a case where int db_resolution = 1 dB is set in Fig. 16, encoding is performed as follows.

[0331]

20 0:0dB
 10:+1dB
 11:+2dB

[0332]

 element_count=4;
 25 int_nbits_res_data=2; // Maximum word length of the huffman decode table (reverse table for obtaining data from index)

 Huff_dec_table[4]={0,0,1,2};

[0333]

30 Huff_dec_table is a reverse table in a case where the maximum word length is 2 bits.

0:0dB

1:0dB

2:1dB

3:2dB

5 [0334]

Furthermore, at the time of decoding, the process is performed in the following procedure.

(1) Acquire bit string from bitstream with maximum word length

10 (2) Refer to huff_dec_table by setting the bit string to i_element (equivalent to recording the huffman code with the maximum word length)

(3) Obtain data in which element of i_element is restored

15 (4) Restore the above data based on db_resolution to obtain dB value

Note that an offset value is required for restoration.

[0335]

20 In addition, the sound pressure (dB value) of the original data can be obtained by $Db = \text{Huff_dec_table}[\text{code}] * \text{db_resolution}$.

[0336]

<Configuration example of server>

25 In a case where the server 11 generates model data by combining one or more methods or encodes difference information in the difference encoding mode, for example, the server 11 is configured as illustrated in Fig. 19.

[0337]

30 Note that, in Fig. 19, portions corresponding to those in the case of Fig. 9 are denoted by the same

reference numerals, and the description thereof will be omitted as appropriate.

[0338]

5 A server 11 illustrated in Fig. 19 is an information processing device including a computer and the like, and functions as an encoding device as in the case in Fig. 9.

[0339]

10 The server 11 includes a directivity data encoding unit 201, the audio data encoding unit 23, and the output unit 24.

[0340]

The directivity data encoding unit 201 generates model data on the basis of the supplied directivity data.
15 The directivity data encoding unit 201 includes a model parameter estimation unit 211, a residual calculator 212, an encoding method selection unit 213, a Huffman encoding unit 214, and a model data generation unit 215.

[0341]

20 Specifically, in this example, the model parameter estimation unit 211 and the residual calculator 212 correspond to the modeling unit 21 in Fig. 9, and the encoding method selection unit 213 to the model data generation unit 215 correspond to the model data
25 generation unit 22 in Fig. 9.

[0342]

The model parameter estimation unit 211 models the supplied directivity data to be processed by at least one or more methods such as the HOA method or the mixing
30 method, and supplies the model parameter for each method obtained as a result to the residual calculator 212 and

the model data generation unit 215.

[0343]

5 The residual calculator 212 calculates difference information on the basis of the supplied directivity data to be processed and the model parameter supplied from the model parameter estimation unit 211, and supplies the difference information to the encoding method selection unit 213 and the Huffman encoding unit 214.

[0344]

10 On the basis of the supplied directivity data to be processed and the difference information supplied from the residual calculator 212, the encoding method selection unit 213 selects a difference encoding mode and a Huffman encoding table when Huffman encoding the difference information, and supplies encoding mode information indicating the selection result to the Huffman encoding unit 214 and the model data generation unit 215.

[0345]

20 The encoding mode information includes difference encoding mode information indicating the selected difference encoding mode (difference encoding method) and table index information indicating the selected Huffman encoding table. Note that only the difference information may be used in generating the encoding mode information in the encoding method selection unit 213.

[0346]

30 The Huffman encoding unit 214 Huffman encodes the difference information supplied from the residual calculator 212 on the basis of the encoding mode information supplied from the encoding method selection

unit 213, and supplies difference code data obtained as a result to the model data generation unit 215.

[0347]

The model data generation unit 215 generates model data including the model parameter for each method supplied from the model parameter estimation unit 211, the difference code data supplied from the Huffman encoding unit 214, and the encoding mode information supplied from the encoding method selection unit 213, and supplies the model data to the output unit 24. Note that, in the case where the difference information is not decoded, the difference code data is not included in the model data. In addition, more specifically, the model data also stores information regarding the above-described directivity data. In addition, information indicating a method used for modeling the directivity data may be stored in the model data.

[0348]

Even in a case where the server 11 has the configuration illustrated in Fig. 19, the server 11 performs the encoding process described with reference to Fig. 10. However, in steps S11 and S12, more specifically, processing described below is performed.

[0349]

That is, in step S11, the model parameter estimation unit 211 models the supplied directivity data to be processed by at least one or more methods, and the residual calculator 212 calculates difference information as necessary.

[0350]

In other words, for example, the HOA method, the

mixing method, the complex mixing method, the difference method, and the like are combined as necessary, whereby the model parameter and the difference information are calculated by the band hybrid method, the addition hybrid
 5 method, the multiplication hybrid method, the spherical harmonic coefficient modeling method, the combination hybrid method, and the like described above.

[0351]

In step S12, the difference encoding mode and the
 10 Huffman encoding table are selected by the encoding method selection unit 213, and the Huffman encoding is performed by the Huffman encoding unit 214 as necessary, and the model data is generated by the model data generation unit 215.

15 [0352]

Specifically, for example, in a case where the model parameter is calculated by the method of the addition hybrid method (AH4), the model parameter estimation unit 211 first models the directivity data by
 20 the HOA method, and as a result, obtains the spherical harmonic coefficient as the model parameter.

[0353]

In addition, the model parameter estimation unit 211 obtains a difference between the directivity data
 25 after modeling by the HOA method and the directivity data before modeling as intermediate difference information, and models the intermediate difference information by the mixing method. By modeling the intermediate difference information by the mixing method, the degree of parameter
 30 concentration κ , the ellipticity β , the weight φ_i , the vector γ_1 , the major axis vector γ_2 , the minor axis

vector γ_3 , the scale factor, and the minimum value are obtained as model parameters.

[0354]

5 The model parameter estimation unit 211 supplies the model parameter obtained by modeling the directivity data by the HOA method and the model parameter obtained by modeling the intermediate difference information by the mixing method to the residual calculator 212 and the model data generation unit 215.

10 [0355]

Then, the residual calculator 212 generates difference information on the basis of the model parameter supplied from the model parameter estimation unit 211 and the supplied directivity data. This
15 difference information is a residual between the directivity data after modeling modeled by the combination of the HOA method and the mixing method and the directivity data before modeling.

[0356]

20 Further, the Huffman encoding unit 214 Huffman encodes the difference information supplied from the residual calculator 212 according to the encoding mode information supplied from the encoding method selection unit 213.

25 [0357]

At this point, the process is performed by the method indicated by the difference encoding mode information. That is, the difference information is Huffman encoded by one or more methods of the spatial
30 adjacency difference method, the inter-frequency difference method, and the complex difference method, or

the difference information is not Huffman encoded.

[0358]

For example, in a case of performing Huffman encoding by the spatial adjacency difference method, the Huffman encoding unit 214 obtains a difference in difference information between data points adjacent to each other as spatial difference information, and generates difference code data by Huffman encoding the spatial difference information.

10 [0359]

The model data generation unit 215 generates model data including the HOA method model parameter and the mixing method model parameter supplied from the model parameter estimation unit 211, and the encoding mode information supplied from the encoding method selection unit 213. Specifically, in a case where the Huffman encoding of the difference information is performed, the model data generation unit 215 also stores the difference code data supplied from the Huffman encoding unit 214 in the model data.

[0360]

Note that, in a case where the model data is generated by the difference method alone, the model parameter estimation unit 211 obtains the difference (hereinafter, also referred to as difference directivity data) of the directivity data by at least one or more methods of the spatial adjacency difference method or the inter-frequency difference method on the basis of the supplied directivity data. The difference directivity data is directivity data between data points and bins, that is, a difference in directivity gain.

[0361]

In this case, the encoding method selection unit 213 generates the encoding mode information on the basis of the difference directivity data supplied from the model parameter estimation unit 211 via the residual calculator 212. In addition, on the basis of the encoding mode information supplied from the encoding method selection unit 213, the Huffman encoding unit 214 Huffman encodes the difference directivity data supplied from the model parameter estimation unit 211 via the residual calculator 212 by a designated difference encoding method to generate encoded directivity data.

[0362]

Then, the model data generation unit 215 generates model data including the encoded directivity data supplied from the Huffman encoding unit 214 and the encoding mode information supplied from the encoding method selection unit 213, and supplies the model data to the output unit 24.

[0363]

<Description of directivity data generation process>

The information processing device 51 that has received the supply of the encoded bit stream from the server 11 having the configuration illustrated in Fig. 19 performs, for example, the directivity data generation process illustrated in Fig. 20, and then performs the output audio data generation process described with reference to Fig. 13 at any timing.

[0364]

The directivity data generation process performed by the information processing device 51 functioning as

the decoding device will be described below with reference to the flowchart of Fig. 20.

[0365]

Note that, in step S111, the process similar to the process in step S51 in Fig. 12 is performed. That is, in step S111, the unpacking unit 81 unpacks the model data, and extracts the model parameter, the information regarding the original directivity data before modeling, the difference code data, and the like from the model data.

[0366]

In step S112, the unpacking unit 81 determines whether or not there is a model parameter that has not yet been supplied to the directivity data calculator 82 among the model parameters for each method extracted by the unpacking.

[0367]

In a case where it is determined in step S112 that there is a model parameter, the unpacking unit 81 supplies a model parameter that has not yet been supplied to the directivity data calculator 82, that is, has not yet been processed, to the directivity data calculator 82, and the process proceeds to step S113.

[0368]

In step S113, the directivity data calculator 82 calculates data on the basis of the model parameter on the basis of the model parameter of one method supplied from the unpacking unit 81.

[0369]

For example, in step S113, on the basis of the model parameter for each manner, such as the HOA method,

the mixing method, or the like, a directivity gain, intermediate difference information, amplitude modulation information, amplitude phase modulation information, and the like constituting the modeled directivity data are
5 calculated as data based on the model parameter for each method such as the HOA method, the mixing method, and the like.

[0370]

When the process in step S113 is executed,
10 thereafter, the process returns to step S112, and the above process is repeated.

[0371]

In addition, in a case where it is determined in step S112 that there is no model parameter that is not
15 supplied to the directivity data calculator 82, thereafter, the process proceeds to step S114.

[0372]

In step S114, the unpacking unit 81 determines whether or not the difference code data is included in
20 the model data supplied from the acquisition unit 61, that is, whether or not the difference code data is present.

[0373]

In a case where it is determined in step S114 that
25 the difference code data is included, the unpacking unit 81 supplies the encoding mode information and the difference code data extracted from the model data to the difference information decoding unit 83, and thereafter, the process proceeds to step S115.

30 [0374]

In step S115, the difference information decoding

unit 83 acquires the encoding mode information and the difference code data output from the unpacking unit 81.
[0375]

In step S116, the difference information decoding
5 unit 83 decodes the difference code data on the basis of the acquired encoding mode information, and supplies difference information (directivity data residual) obtained as a result to the addition unit 84.
[0376]

10 For example, it is assumed that it is identified by the difference encoding mode information included in the encoding mode information that the encoding by the spatial adjacency difference method is performed.
[0377]

15 In such a case, the difference information decoding unit 83 decodes the difference code data supplied from the unpacking unit 81 using the reverse table identified by the table index information included in the encoding mode information to obtain the spatial difference
20 information at each data point.
[0378]

Then, the difference information decoding unit 83 adds difference information at another decoded data point near the data point to the spatial difference information
25 at the data point to be processed to obtain difference information about the data point to be processed.
[0379]

In a case where it is determined that the process in step S116 is performed or that there is no difference
30 code data in step S114, the process in step S117 is then performed.

[0380]

In step S117, the directivity data calculator 82 and the addition unit 84 calculate the directivity data.

[0381]

5 That is, the directivity data calculator 82 calculates rough directivity data on the basis of the data obtained by the process in step S113 performed one or more times, and supplies the rough directivity data to the addition unit 84.

10 [0382]

As a specific example, for example, it is assumed that the model parameter is calculated by the addition hybrid method (AH4) on the server 11 side.

[0383]

15 In such a case, in the first process of step S113, the modeled directivity data (rough directivity data) is calculated on the basis of the model parameter of the HOA method. In addition, in the second process in step S113, the intermediate difference information after modeling is
20 calculated on the basis of the model parameter of the mixing method.

[0384]

Therefore, the directivity data calculator 82 obtains the final rough directivity data by adding the
25 intermediate difference information to the rough directivity data, that is, by adding the intermediate difference information for each bin at each data point to the directivity gain for each bin at each data point.

[0385]

30 The addition unit 84 calculates the directivity data by adding the difference information (directivity

data residual) supplied from the difference information decoding unit 83 to the final rough directivity data obtained by the directivity data calculator 82 in this manner, and supplies the directivity data to the
5 frequency interpolation processing unit 85. Note that, in a case where there is no difference information, the final rough directivity data is directly used as the directivity data.

[0386]

10 In addition, for example, it is assumed that the model parameter is calculated by the multiplication hybrid method (MH1) on the server 11 side.

[0387]

In such a case, in the first process of step S113,
15 the modeled directivity data (rough directivity data) is calculated on the basis of the model parameter of the HOA method. In addition, in the second process in step S113, the amplitude modulation information after modeling is calculated on the basis of the model parameter of the
20 mixing method.

[0388]

Therefore, the directivity data calculator 82 obtains the final directivity data by multiplying the rough directivity data by the amplitude modulation
25 information, that is, by multiplying the directivity gain for each bin at each data point by the amplitude modulation information for each bin at each data point. In this case, since the process of steps S115 and S116 is not performed and there is no difference information, the
30 directivity data obtained by the directivity data calculator 82 is directly supplied to the frequency

interpolation processing unit 85 via the addition unit 84.

[0389]

Note that, for example, the model data may be generated by the difference method alone on the server 11 side. In such a case, the process in step S113 is not performed, and the difference information decoding unit 83 decodes the encoded directivity data in steps S115 and S116.

10 [0390]

That is, the difference information decoding unit 83 decodes the encoded directivity data supplied from the unpacking unit 81 using the reverse table identified by the table index information included in the encoding mode information to obtain the difference directivity data.

[0391]

Then, in step S117, the difference information decoding unit 83 calculates the directivity data on the basis of the value (difference) for each bin at each data point constituting the difference directivity data.

[0392]

Specifically, for example, in a case where the difference directivity data is calculated by the spatial adjacency difference method on the server 11 side, the difference information decoding unit 83 adds a directivity gain of the same bin at another restored data point in the vicinity of the data point to a value (difference) for each bin at the data point to be processed, thereby obtaining the directivity gain for each bin at the data point to be processed.

[0393]

Furthermore, for example, in a case where the difference directivity data is calculated by the inter-frequency difference method on the server 11 side, the difference information decoding unit 83 adds a
5 directivity gain of another restored bin in the vicinity of the bin to be processed at the same data point to a value (difference) of the bin to be processed of the data point, thereby obtaining a directivity gain of the bin to be processed.

10 [0394]

As described above, even in a case where the encoded directivity data is stored in the model data, the amount of transmission of the directivity data can be reduced.

15 [0395]

When the process of step S117 is performed and the directivity data is calculated, then the process of step S118 is performed and the directivity data generation process is ended. Since the process of step S118 is
20 similar to the process of step S56 of Fig. 12, the description thereof is omitted.

[0396]

As described above, the information processing device 51 calculates the directivity data on the basis of
25 the model data. In this way, the amount of transmission of directivity data can be reduced. As a result, occurrence of a transmission delay and an increase in the transmission rate can be suppressed.

[0397]

30 <Specific example of addition hybrid method>

<Configuration example of directivity data encoding unit>

In a case where the model data is fixedly generated by the addition hybrid method in the server 11, the configuration of the directivity data encoding unit 201 in the server 11 illustrated in Fig. 19 can be, for example, the configuration illustrated in Fig. 21. Note that, in Fig. 21, portions corresponding to those in a case of Fig. 19 are denoted by the same reference numerals, and description thereof will be omitted as appropriate.

10 [0398]

In the example of Fig. 21, the directivity data encoding unit 201 includes a model parameter estimation unit 241, an arithmetic unit 242, a model parameter estimation unit 243, an arithmetic unit 244, a difference encoding unit 245, and a model data generation unit 215.

15 [0399]

The model parameter estimation unit 241 to the arithmetic unit 244 correspond to the model parameter estimation unit 211 in Fig. 19.

20 [0400]

The model parameter estimation unit 241 models the supplied directivity data to be processed by the mixing method, supplies the model parameter obtained as a result to the model data generation unit 215, and supplies the directivity data after modeling by the mixing method to the arithmetic unit 242.

[0401]

The arithmetic unit 242 calculates the intermediate difference information by subtracting (obtaining a difference) the modeled directivity data supplied from the model parameter estimation unit 241 from the supplied

30

directivity data to be processed, and supplies the intermediate difference information to the model parameter estimation unit 243 and the arithmetic unit 244.

5 [0402]

The model parameter estimation unit 243 models the intermediate difference information supplied from the arithmetic unit 242 by the HOA method, supplies the model parameter obtained as a result to the model data generation unit 215, and supplies the intermediate difference information after modeling by the HOA method to the arithmetic unit 244.

[0403]

The arithmetic unit 244 calculates difference information by subtracting (obtaining a difference) the intermediate difference information after modeling supplied from the model parameter estimation unit 243 from the intermediate difference information supplied from the arithmetic unit 242, and supplies the difference information to the difference encoding unit 245.

[0404]

The difference encoding unit 245 generates encoding mode information and difference code data on the basis of the difference information supplied from the arithmetic unit 244 and the supplied directivity data to be processed as appropriate, and supplies the encoding mode information and the difference code data to the model data generation unit 215.

[0405]

30 Note that, here, an example is described in which the model parameter estimation unit 241 performs modeling

by the mixing method, and the model parameter estimation unit 243 performs modeling by the HOA method.

[0406]

However, the present invention is not limited thereto, and modeling may be performed by any method in the model parameter estimation unit 241 and the model parameter estimation unit 243. For example, the model parameter estimation unit 241 may perform modeling by the HOA method, and the model parameter estimation unit 243 may perform modeling by the mixing method.

[0407]

<Configuration example of difference encoding unit>

Furthermore, the difference encoding unit 245 can have a configuration illustrated in Fig. 22, for example. Note that, in Fig. 22, portions corresponding to those in a case of Fig. 19 are denoted by the same reference numerals, and description thereof will be omitted as appropriate.

[0408]

In the example of Fig. 22, difference encoding unit 245 includes a residual calculator 212, an encoding method selection unit 213, a multi-stage difference processing unit 271, and a Huffman encoding unit 214.

[0409]

The residual calculator 212 calculates difference information on the basis of the supplied directivity data to be processed and the modeled directivity data and intermediate difference information supplied from the model parameter estimation unit 241 and the model parameter estimation unit 243, and supplies the difference information to the encoding method selection

unit 213 and the multi-stage difference processing unit 271.

[0410]

5 The multi-stage difference processing unit 271
generates the multi-stage difference information in the
difference encoding mode indicated by the encoding mode
information supplied from the encoding method selection
unit 213 on the basis of either the difference
information from the residual calculator 212 or the
10 difference information from the arithmetic unit 244.

[0411]

For example, the spatial difference information is
obtained as the multi-stage difference information in a
case where the Huffman encoding is performed by the
15 spatial adjacency difference method as the difference
encoding mode, and the inter-frequency difference
information is obtained as the multi-stage difference
information in a case where the Huffman encoding is
performed by the inter-frequency difference method as the
20 difference encoding mode. Similarly, in a case where the
Huffman encoding is performed by the spatial adjacency
difference method and the inter-frequency difference
method as the difference encoding mode, information to be
Huffman encoded obtained by obtaining the spatial
25 difference information and the inter-frequency difference
information is the multi-stage difference information.

[0412]

30 The multi-stage difference processing unit 271
supplies the obtained multi-stage difference information
to the encoding method selection unit 213 and the Huffman
encoding unit 214.

[0413]

The encoding method selection unit 213 generates the encoding mode information on the basis of the supplied directivity data to be processed, the difference information supplied from the residual calculator 212 or the arithmetic unit 244, and the multi-stage difference information supplied from the multi-stage difference processing unit 271, and supplies the encoding mode information to the multi-stage difference processing unit 271, the Huffman encoding unit 214, and the model data generation unit 215.

[0414]

The Huffman encoding unit 214 Huffman encodes the multi-stage difference information supplied from the multi-stage difference processing unit 271 on the basis of the encoding mode information supplied from the encoding method selection unit 213, and supplies difference code data obtained as a result to the model data generation unit 215.

20 [0415]

<Description of model data generation process>

In a case where the directivity data encoding unit 201 has the configuration illustrated in Fig. 21, the directivity data encoding unit 201 performs the model data generation process illustrated in Fig. 23 as processing corresponding to steps S11 and S12 in Fig. 10.

[0416]

That is, in step S151, the model parameter estimation unit 241 performs modeling by the mixing method on the supplied directivity data to be processed.

[0417]

The model parameter estimation unit 241 supplies model parameter obtained by modeling to the model data generation unit 215, and supplies directivity data after modeling by the mixing method to the arithmetic unit 242.

5 [0418]

In step S152, the arithmetic unit 242 calculates intermediate difference information on the basis of the supplied directivity data to be processed and the modeled directivity data supplied from the model parameter estimation unit 241, and supplies the intermediate difference information to the model parameter estimation unit 243 and the arithmetic unit 244.

[0419]

In step S153, the model parameter estimation unit 243 performs modeling by the HOA method on the intermediate difference information supplied from the arithmetic unit 242.

[0420]

The model parameter estimation unit 243 supplies model parameter obtained by modeling to the model data generation unit 215, and supplies intermediate difference information after modeling by the HOA method to the arithmetic unit 244.

[0421]

25 In step S154, the arithmetic unit 244 calculates difference information on the basis of the intermediate difference information supplied from the arithmetic unit 242 and the intermediate difference information after modeling supplied from the model parameter estimation unit 243, and supplies the difference information to the difference encoding unit 245.

[0422]

In step S155, the difference encoding unit 245 performs difference encoding on the basis of the difference information supplied from the arithmetic unit
5 244.

[0423]

That is, for example, the encoding method selection unit 213 of the difference encoding unit 245 generates encoding mode information on the basis of the supplied
10 directivity data to be processed, the difference information supplied from the arithmetic unit 244, and the multi-stage difference information supplied from the multi-stage difference processing unit 271 in the previous processing such as the previous frame, and
15 supplies the encoding mode information to the multi-stage difference processing unit 271, the Huffman encoding unit 214, and the model data generation unit 215. Note that the encoding method selection unit 213 may generate the encoding mode information using the difference
20 information supplied from the residual calculator 212.

[0424]

Furthermore, the multi-stage difference processing unit 271 generates multi-stage difference information on the basis of, for example, the difference information
25 supplied from the arithmetic unit 244 and the encoding mode information supplied from the encoding method selection unit 213, and supplies the multi-stage difference information to the encoding method selection unit 213 and the Huffman encoding unit 214.

30 [0425]

The Huffman encoding unit 214 Huffman encodes the

multi-stage difference information supplied from the multi-stage difference processing unit 271 on the basis of the encoding mode information supplied from the encoding method selection unit 213, and supplies
5 difference code data obtained as a result to the model data generation unit 215.

[0426]

In step S156, the model data generation unit 215 performs packing to generate model data, and supplies the
10 model data to the output unit 24.

[0427]

Specifically, the model data generation unit 215 generates model data including the model parameter of the mixing method from the model parameter estimation unit
15 241, the model parameter of the HOA method from the model parameter estimation unit 243, the encoding mode information from the encoding method selection unit 213, and the difference code data from the Huffman encoding unit 214. When the model data is generated in this
20 manner, the model data generation process ends.

[0428]

As described above, the directivity data encoding unit 201 generates the model data by the addition hybrid method. By doing so, the amount of transmission of
25 directivity data can be reduced, and occurrence of a transmission delay and an increase in a transmission rate can be suppressed.

[0429]

<Configuration example of distribution model decoding
30 unit>

Furthermore, in a case where the directivity data

encoding unit 201 has the configuration illustrated in Fig. 21, the distribution model decoding unit 62 of the information processing device 51 has the configuration illustrated in Fig. 24, for example. Note that, in Fig. 24, portions corresponding to those in a case of Fig. 11 are denoted by the same reference numerals, and description thereof will be omitted as appropriate.

[0430]

The distribution model decoding unit 62 illustrated in Fig. 24 includes an unpacking unit 81, a calculation unit 301, a calculation unit 302, a difference information decoding unit 83, an arithmetic unit 303, an arithmetic unit 304, and a frequency interpolation processing unit 85. In this example, the calculation unit 301 and the calculation unit 302 correspond to the directivity data calculator 82 illustrated in Fig. 11.

[0431]

The calculation unit 301 calculates directivity data (rough directivity data) after modeling by the mixing method on the basis of the model parameter of the mixing method supplied from the unpacking unit 81, and supplies the directivity data to the arithmetic unit 304. The calculation unit 302 calculates intermediate difference information after modeling by the HOA method on the basis of the model parameter of the HOA method supplied from the unpacking unit 81, and supplies the intermediate difference information to the arithmetic unit 303.

[0432]

The difference information decoding unit 83 calculates difference information (directivity data

residual) on the basis of the encoding mode information and the difference code data supplied from the unpacking unit 81, and supplies the difference information to the arithmetic unit 303. The arithmetic unit 303 adds
5 (combines) the difference information supplied from the difference information decoding unit 83 and the intermediate difference information supplied from the calculation unit 302, and supplies the addition result (difference information) to the arithmetic unit 304.

10 [0433]

The arithmetic unit 304 adds the directivity data (rough directivity data) supplied from the calculation unit 301 and the addition result (difference information) supplied from the arithmetic unit 303, and supplies
15 directivity data obtained as a result to the frequency interpolation processing unit 85.

[0434]

In a case where the distribution model decoding unit 62 has the configuration illustrated in Fig. 24, the
20 calculation unit 301 calculates the directivity data (rough directivity data) in the first step S113 in the directivity data generation process of Fig. 20 described above. In addition, in the second step S113, the calculation unit 302 calculates the intermediate
25 difference information.

[0435]

Then, the difference information decoding unit 83 performs the process of steps S115 and S116 to generate the difference information, and the arithmetic unit 303
30 and the arithmetic unit 304 perform the addition process in step S117 to generate the directivity data.

[0436]

<Arrangement of data points>

Incidentally, the configuration of the model data described above is not limited to the configuration illustrated in Fig. 5 and the configurations illustrated in Figs. 15 and 16, and may be the configuration illustrated in Fig. 25.

[0437]

Note that, in Fig. 25, description of portions corresponding to those in the case illustrated in Fig. 5 will be appropriately omitted. In Fig. 25, bslbf indicates a bit string, left bit first, that is, the left bit is the head. Furthermore, uimbsf indicates unsigned integer most significant bit first, that is, an unsigned integer in which the most significant bit is the head.

[0438]

The model data illustrated in Fig. 25 includes the number of frequency points "bin_count" indicating the number of frequency bins, and the frequency "bin_freq[i]" at the center of the frequency bins is stored by the number of frequency points "bin_count".

[0439]

In addition, the number of mixtures "mix_count[j]" indicating the number of distributions constituting the mixture model in each band and the bin information "bin_range_per_band[j]" indicating the bin included in the band are stored by the number of bands "band_count".

[0440]

Furthermore, for each band, a degree of parameter concentration κ , a weight ϕ_i , and a vector γ_1 as model parameters, and a selection flag "dist_flag" are stored

by the number of mixtures "mix_count[k]".

[0441]

In this example, "kappa[j][k]" indicates the degree of parameter concentration κ , and "weight[j][k]"

5 indicates the weight ϕ_i . Furthermore, "gamma_x[j][k]", "gamma_y[j][k]", and "gamma_z[j][k]" indicate an X component (X coordinate), a Y component (Y coordinate), and a Z component (Z coordinate) constituting the vector γ_1 .

10 [0442]

In a case where the selection flag "dist_flag" is "1", that is, in a case where the distribution is the Kent distribution, the ellipticity β , the major axis vector γ_2 , and the minor axis vector γ_3 are further

15 stored.

[0443]

Here, "beta[j][k]" indicates the ellipticity β , and "gamma2_x[j][k]", "gamma2_y[j][k]", and "gamma2_z[j][k]" indicate the X component, the Y component, and the Z

20 component constituting the major axis vector γ_2 .

"gamma3_x[j][k]", "gamma3_y[j][k]", and "gamma3_z[j][k]" indicate the X component, the Y component, and the Z component constituting minor axis vector γ_3 .

[0444]

25 The model data also includes a scale factor "scale_factor[i]" indicating the dynamic range of the directivity gain and an offset value of the directivity data in each bin, that is, a minimum value "offset[i]" by the number of frequency points "bin_count".

30 [0445]

In addition, the model data also includes

information for identifying the position of each data point.

[0446]

5 In the directivity recording method described above, it is assumed that the value of the directivity data at the data point defined by the original data (original directivity data), that is, the directivity gain is restored as accurately as possible.

[0447]

10 In the information processing device 51, the decoded directivity data is used when the rendering process is performed. However, what is required in this case is not limited to the value (directivity gain) at the data point described in the original directivity data, but is the directivity gain at the position
15 (orientation) used at the time of the rendering process.

[0448]

Therefore, for example, it is necessary to record the directivity data not only in the data arrangement
20 (hereinafter, referred to as grid pattern data arrangement) in which the data (directivity gain), that is, the data points are disposed at grid points obtained by dividing the latitude and longitude at equal intervals on the spherical surface, but also in various data
25 arrangements. In other words, Syntax for recording coordinate information about data points as efficiently as possible is necessary.

[0449]

30 As a method of disposing the data points in the directivity data, for example, the following method (arrangement) can be considered.

[0450]

- Grid pattern data arrangement
- Uniform data arrangement
- Non-uniform data arrangement

5 [0451]

Here, the uniform data arrangement is, for example, a data arrangement in which a plurality of data points is uniformly disposed on a spherical surface centered on a sound source position as illustrated in Fig. 26. In
10 other words, in a uniform data arrangement, the data points are disposed at a constant density in any region on the spherical surface.

[0452]

In the example of Fig. 26, it can be seen that each
15 point on the spherical surface represents a data point, and the data points are disposed at a constant density in any azimuth viewed from the sound source position, that is, a directivity gain (directivity data) is recorded at a constant density.

20 [0453]

The recording of the directivity data by such uniform data arrangement is particularly effective in a case where the direction of the listener (user) viewed from the sound source changes evenly with time.

25 [0454]

In addition, the non-uniform data arrangement is a data arrangement in which a plurality of data points is non-uniformly disposed on the spherical surface centered on the sound source position. In other words, in the
30 non-uniform data arrangement, the data points are disposed at different densities for each region on the

spherical surface. Therefore, it can be said that the grid pattern data arrangement is one arrangement example of the non-uniform data arrangement, but in the following description, the non-uniform data arrangement does not
5 include the grid pattern data arrangement.
[0455]

As a specific example of the non-uniform data arrangement, for example, it is conceivable to arrange the data points with high density in a region
10 corresponding to the front direction of the sound source which is important for audibility on the spherical surface centered on the sound source position, and in a region corresponding to the direction in which the user's viewpoint and the sound source are likely to approach as
15 a positional relationship. In the non-uniform data arrangement, it is also conceivable to arrange the data points with high density in a region where the directivity gain is large.
[0456]

20 As another example of the non-uniform data arrangement, it is conceivable to densely arrange the data points, that is, the directivity gains in a site (region) where the amount of change in the directivity gain is large as a whole or an important region on the
25 spherical surface centered on the sound source position, and coarsely arrange the data points in a region where the degree of importance is low.
[0457]

In any of the grid pattern data arrangement, the
30 uniform data arrangement, and the non-uniform data arrangement described above, it is conceivable to

cooperate with the priority of the object in determining the priority or the like of the directivity data. For example, the priority of the directivity data may be determined on the basis of the priority of the sound source type of the object in the content in which the

5 directivity data is utilized.

[0458]

As an example of cooperation with the priority of an object, for example, in a case where there is a

10 plurality of objects in content, in the case of content of music, it is conceivable to set the priority of an object corresponding to a vocal to be high.

[0459]

Furthermore, for example, in a case where there is

15 a sound source type with high priority, that is, an object sound source with high priority, such as a vocal in music content or a voice in movie content, it is conceivable to allocate more bits to the description of the directivity data of the sound source type. That is,

20 in the directivity data of the sound source type with higher priority, it is conceivable to provide more data points and record the directivity data with high definition.

[0460]

25 In a case where the arrangement positions and the like of the data points are recorded in the data arrangement as described above, for example, the information illustrated in Fig. 27 may be further described in the model data including the information

30 illustrated in Fig. 25. That is, Fig. 27 illustrates an example of a description format (Syntax) of information

or the like for identifying the position of each data point.

[0461]

Note that, here, it is assumed that the distance
5 from the sound source position (sound source center) to
each data point is constant. That is, an example in
which each data point is disposed on the surface of the
sphere centered on the sound source position will be
described. However, the present invention is not limited
10 thereto, and the distance from the sound source position
to the data point may be different for each data point.

[0462]

In the example of Fig. 27, "position_type" is
information indicating the arrangement format
15 (arrangement method) of the data points, that is, the
coordinate recording method.

[0463]

For example, in a case where the arrangement of the
data points is the grid pattern data arrangement, the
20 value of the coordinate recording method "position_type"
is "0x000".

[0464]

Further, for example, in a case where the
arrangement of the data points is uniform data
25 arrangement, the value of the coordinate recording method
"position_type" is "0x001", and in a case where the
arrangement of the data points is non-uniform data
arrangement, the value of the coordinate recording method
"position_type" is "0x010".

30 [0465]

"priority_index" is priority information indicating

the priority of the directivity data, more specifically,
the priority of the directivity data. For example, since
the directivity data is prepared for each type of object,
that is, for each sound source type, it can be said that
5 the priority information indicates the priority of the
directivity data for each type of sound source (object).
This priority may change over time.

[0466]

Specifically, for example, in a case where the
10 value of the priority "priority_index" is "0x000", that
is, in a case where the value indicating the priority is
the minimum, it is indicated that the priority of the
directivity data is the maximum. Here, the higher the
priority of the directivity data, the smaller the value
15 indicating the priority.

[0467]

Furthermore, in a case where the priority of the
directivity data is the maximum, for example, regarding
the directivity data, all the data points before modeling
20 (before encoding) may be restored (decoded) without
reducing the spatial resolution in the information
processing device 51 on the decoding side.

[0468]

That is, the information processing device 51, more
25 specifically, the distribution model decoding unit 62 may
calculate the directivity data having the same position
and the same number of data points as those before
modeling on the basis of the model data. In addition,
for example, the density (number) of data points
30 constituting the directivity data may be determined
according to the priority of the directivity data.

[0469]

Furthermore, in this example, information for identifying the arrangement position (coordinates) of the data point is described according to the value of the coordinate recording method "position_type".

[0470]

Specifically, in a case where the value of the coordinate recording method "position_type" is "0x000", that is, in the case of the grid pattern data arrangement, the azimuth angle direction interval "azimuth_interval" and the elevation angle direction interval "elevation_interval" are described (stored).

[0471]

The azimuth angle direction interval "azimuth_interval" indicates an angle (difference in azimuth angle) indicating an interval in the azimuth angle direction between data points adjacent to each other in the azimuth angle direction on the spherical surface.

[0472]

The elevation angle direction interval "elevation_interval" indicates an angle (difference in elevation angle) indicating the elevation angle direction interval between the data points adjacent to each other in the elevation angle direction on the spherical surface.

[0473]

Furthermore, in the grid pattern data arrangement, at least one position as a reference such as a position in the front direction viewed from the sound source position is known as the arrangement position of the data

points on the information processing device 51 side.
Therefore, the positions of all the data points can be
identified from the azimuth angle direction interval and
the elevation angle direction interval, and the
5 predetermined reference position.

[0474]

In a case where the value of the coordinate
recording method "position_type" is "0x001", that is, in
a case of uniform data arrangement, the number of data
10 points "uniform_dist_point_count" indicating the number
of data points uniformly distributed (disposed) on the
spherical surface is described (stored).

[0475]

In the uniform data arrangement, for example, on
15 the information processing device 51 side, the
arrangement position of each data point is known for each
number of data points, and the positions of all the data
points can be identified from the number of data points.

[0476]

20 In a case where the value of the coordinate
recording method "position_type" is "0x010", that is, in
the case of non-uniform data arrangement, together with
the number of the mandatory data points
"num_mandatory_point", the azimuth angle data
25 "azimuth_table[i]" and the elevation angle data
"elevation_table[i]" indicating the position of the
mandatory data point are described (stored) by the number
of the mandatory data points.

[0477]

30 Further, in a case where the value of the
coordinate recording method "position_type" is "0x010",

the data point arrangement resolution, in other words,
the data point arrangement resolution indicating the
arrangement density of the data points "gain_resolution",
is also described (stored). For example, the data point
5 arrangement resolution "gain_resolution" is a decibel
value indicating an amount of fluctuation of data
(directivity gain).

[0478]

In the non-uniform arrangement, the data point is
10 set for each amount of fluctuation of the directivity
gain indicated by the data point arrangement resolution
"gain_resolution". That is, the number of data points in
the directivity data obtained by decoding changes
according to the data point arrangement resolution.

15 [0479]

Specifically, in the non-uniform arrangement, data
points that always exist (are disposed), that is, data
points that are always restored at the time of decoding,
regardless of the data point arrangement resolution are
20 set as the mandatory data points. The number of
mandatory data points "num_mandatory_point" indicating
the number of mandatory data points is described.

[0480]

Furthermore, the azimuth angle data
25 "azimuth_table[i]" and the elevation angle data
"elevation_table[i]" are an azimuth angle and an
elevation angle indicating the positions (coordinates) in
the azimuth angle direction and the elevation angle
direction of the mandatory data points, respectively.

30 [0481]

Therefore, on the decoding side, the arrangement

position of each mandatory data point can be identified by the azimuth angle data "azimuth_table[i]" and the elevation angle data "elevation_table[i]". Note that the azimuth angle data and the elevation angle data are not limited to coordinates, that is, the azimuth angle and the elevation angle as long as it is the information for identifying the arrangement position of the mandatory data point, and may be any other information such as an index that can obtain the azimuth angle and the elevation angle.

[0482]

In the non-uniform arrangement, when the arrangement position of the mandatory data point is identified, the arrangement positions of the data points other than the mandatory data point in the directivity data are identified on the basis of the arrangement position of the mandatory data point and the data point arrangement resolution "gain_resolution".

[0483]

Specifically, first, on the basis of the model data, more specifically, the model parameter, a mixture model $F(x; \theta)$ is obtained. This mixture model $F(x; \theta)$ gives the value of the directivity gain at any position on the spherical surface surrounding the sound source position.

[0484]

Next, data points (hereinafter, also referred to as non-mandatory data points) that are not a mandatory data point are disposed on the spherical surface on the basis of the mixture model $F(x; \theta)$, the position of the mandatory data point, and the data point arrangement

resolution.

[0485]

The position of the non-mandatory data point is a position in which the value of the directivity gain indicated by the mixture model $F(x; \theta)$ is changed from the value of the directivity gain at the mandatory data point on the spherical surface by the amount of fluctuation indicated by the data point arrangement resolution, for example, 3 dB.

10 [0486]

Therefore, for example, in a case where the amount of fluctuation indicated by the data point arrangement resolution is +3 dB and the value of the directivity gain at any one mandatory data point is 48 dB, the non-mandatory data point is disposed at a position where the directivity gain is 51 dB on the spherical surface.

[0487]

At this time, another non-mandatory data point may be further set at a position where the value of the directivity gain on the spherical surface is a value changed from the value of the directivity gain at the non-mandatory data point that has already been set by the amount of fluctuation indicated by the data point arrangement resolution. That is, the non-mandatory data points may be disposed at intervals corresponding to the amount of fluctuation indicated by the data point arrangement resolution with respect to the mandatory data point.

[0488]

30 In addition, for example, the non-mandatory data points whose number corresponds to the data point

arrangement resolution may be disposed at equal intervals between the mandatory data points adjacent to each other in the azimuth angle direction and the elevation angle direction.

5 [0489]

As described above, the arrangement positions of all the data points constituting the directivity data in the non-uniform arrangement, that is, the arrangement positions of all the mandatory data points and the non-
10 mandatory data points are identified.

[0490]

As described above, in the non-uniform arrangement, the arrangement positions and the number of non-mandatory data points of the directivity data obtained on the
15 decoding side vary depending on the data point arrangement resolution "gain_resolution".

[0491]

In the above example, the spatial resolution of the directivity data, that is, the number of data points can
20 be adjusted according to the value of the priority "priority_index" on the decoding side (the information processing device 51) even in a case where the arrangement format of the data points (coordinate recording method) is any format of the grid pattern data
25 arrangement, the uniform data arrangement, and the non-uniform data arrangement. In this case, in each arrangement format, the number of data points changes according to the value of the priority "priority_index".

[0492]

30 Specifically, for example, in the grid pattern data arrangement, the spatial resolution of the directivity

data can be reduced by increasing the azimuth angle direction interval "azimuth_interval" and the elevation angle direction interval "elevation_interval".

[0493]

5 In addition, in the uniform data arrangement, the spatial resolution of the directivity data can be reduced by reducing the number of data points "uniform_dist_point_count".

[0494]

10 Similarly, in the non-uniform data arrangement, the spatial resolution of the directivity data can be reduced by increasing the data point arrangement resolution "gain_resolution".

[0495]

15 As a method of adjusting the spatial resolution of the directivity data, that is, the data amount of the directivity data obtained by decoding, for example, a method of multiplying the value of the priority "priority_index" by the azimuth angle direction interval "azimuth_interval" or the elevation angle direction interval "elevation_interval" is considered.

[0496]

25 Furthermore, as a method of adjusting the spatial resolution of the directivity data, for example, a method of multiplying the number of data points "uniform_dist_point_count" by the reciprocal of the value of the priority "priority_index", a method of multiplying the data point arrangement resolution "gain_resolution" by the value of the priority "priority_index", or the like can be considered.

[0497]

In this way, the information processing device 51 can obtain directivity data of an appropriate spatial resolution. That is, the spatial resolution (the number of data points) of the directivity data can be
5 appropriately adjusted.
[0498]

Note that, also in the model data illustrated in Figs. 5, 15, and 16, as information for identifying the position of each data point, information (hereinafter,
10 also referred to as data point position information) of the configuration illustrated in Fig. 27 may be stored instead of the azimuth angle, the elevation angle, and the like for each data point.
[0499]

15 In a case where the model data includes the data point position information having the configuration illustrated in Fig. 27, the model data generation unit 22 generates the model data including information illustrated in Figs. 25 and 27 in step S12 of the
20 encoding process described with reference to Fig. 10. That is, model data including the data point position information is generated.
[0500]

Note that even in a case where the server 11 has
25 the configuration illustrated in Fig. 19, model data including the data point position information may be generated by the model data generation unit 215.
[0501]

In addition, in a case where information for each
30 data point such as difference information is obtained at the time of generating the model data, each piece of

information such as difference information is calculated for each data point of the decoded directivity data, that is, each data point identified by the data point position information.

5 [0502]

Further, in a case where the data point position information is included in the model data, the directivity data calculator 82 generates the directivity data also using the data point position information in
10 step S52 of the directivity data generation process described with reference to Fig. 12.

[0503]

That is, on the basis of the data point position information included in the model data, the directivity
15 data calculator 82 identifies the arrangement format (coordinate recording method) of the data points and identifies the arrangement position of each data point in the directivity data. At this time, the directivity data calculator 82 identifies the arrangement position of the
20 data points using the priority information about the directivity data as necessary.

[0504]

Further, the directivity data calculator 82 calculates the output value $F(x; \theta)$ of the mixture model
25 for each bin at the data point on the basis of the mixture model $F'(x; \theta)$ of each band calculated from model parameter and the like, the result of identifying the arrangement position of each data point, the scale factor for each bin, and the minimum value for each bin. As a
30 result, rough directivity data including a directivity gain for each bin at each data point is obtained.

[0505]

Similarly, in a case where the data point position information is included in the model data, also in the directivity data generation process described with
5 reference to Fig. 20, the result of identifying the arrangement position of the data points is appropriately used in steps S113, S116, and S117.

[0506]

In the above description, the spatial adjacency
10 difference method and the inter-frequency difference method have been described as the difference encoding method.

[0507]

For example, in the inter-frequency difference
15 method, difference information and a difference in directivity gain between adjacent bins, that is, between adjacent frequencies are obtained.

[0508]

In such an inter-frequency difference method, a
20 property that the value of the directivity gain is close between adjacent frequencies (bins), that is, the shape of the directivity data is close in the directivity data is used.

[0509]

25 Similarly, in the spatial adjacency difference method, difference information and a difference in directivity gain between adjacent data points, that is, between adjacent positions are obtained.

[0510]

30 In such a spatial adjacency difference method, a property that a difference in directivity gain is small

between spatially close positions in the directivity data is used. That is, a property that the directivity gain on the spherical surface changes continuously in many cases in the directivity data, and the value of the
 5 directivity gain is close when the position (orientation) is close is used.

[0511]

In general, in a case where directivity or a head-related transfer function (HRTF) is recorded, for
 10 example, in a file in a Spatially Oriented Format for Acoustics (SOFA), data is defined on a spherical surface, and data points are often recorded in the following manner.

[0512]

```
15      for elev in elevation
          for azi in azimuth
              data_point(azi, elev)
          end
      end
```

20 [0513]

Specifically, for example, on the same latitude on the spherical surface, that is, on the circumference corresponding to the latitude, the data points are disposed at longitude positions adjacent to each other
 25 along the circumference. At this time, the data points are disposed at equal intervals, for example, so as to go around the circumference.

[0514]

Then, when the data point is provided for the
 30 latitude to be processed, then, the data point is provided on the spherical surface by disposing the data

point at each longitude position on the circumference corresponding to the latitude while sequentially changing the value of the latitude.

[0515]

5 In this way, directivity data of a method such as the grid pattern data arrangement can be obtained. In such a grid pattern data arrangement, the data density around the poles such as the south pole and the north pole, that is, the density of data points increases.

10 [0516]

 However, when actually recording the directivity data (directivity gain) as described above, it is desirable to record the directivity data in a data distribution in which the data (data points) is dense in an important azimuth in which it is necessary to record a change in the directivity gain with high definition or is uniform (uniform distribution) as a whole. The important azimuth mentioned here is, for example, a front direction or the like, a direction often used at the time of rendering, a direction of a position where the value of the directivity gain is large, or the like.

[0517]

 Furthermore, in a case where recording of the directivity data is actually considered, it is conceivable that the data on the horizontal plane is recorded densely and is recorded sparsely around the pole due to the convenience of recording.

[0518]

 Therefore, difference encoding may be performed by performing sorting (rearrangement) as follows.

[0519]

(Method DE1): Difference encoding in order of sorting data points on the basis of a predetermined criterion

(Method DE2): Difference encoding by sorting the
5 decibel values of the directivity gains in ascending or descending order

(Method DE3): Difference encoding by sorting in descending order of priority
[0520]

10 In the method DE1, the data points, that is, the difference information and the directivity gain at the data points are sorted (redisposed) in a predetermined order with respect to the data arrangement such as the grid pattern data arrangement, the uniform data
15 arrangement, and the non-uniform data arrangement. Then, the difference information and the difference in directivity gain are obtained between the data points adjacent to each other after the sorting. In this case, the order of sorting is known on the decoding side, that
20 is, on the information processing device 51 side.
[0521]

In the method DE2, the data points are sorted in ascending or descending order of values (decibel values (dB values)) to be calculated of differences such as
25 difference information and directivity gains at the data points. At this time, whether sorting is performed in ascending order or descending order is only required to be determined in advance.
[0522]

30 In addition, when sorting is performed in ascending order or descending order, difference information and a

difference in directivity gain are obtained between data points adjacent to each other after sorting. In this way, the difference information and the difference in directivity gain between the data points can be further
5 reduced.
[0523]

Note that, in the method DE2, information indicating the arrangement order of the sorted data points is stored in the model data so that the order of
10 sorting can be identified on the decoding side (information processing device 51 side). For example, information indicating the arrangement order of the sorted data points may be stored in the data point position information illustrated in Fig. 27.
15 [0524]

Furthermore, the information indicating the arrangement order of the sorted data points may be any information such as, for example, information obtained by disposing indexes indicating the data points in the
20 sorting order.
[0525]

In the method DE3, among the respective azimuths (directions) viewed from the sound source position, the data points are sorted in order from the data point in
25 the azimuth with high priority such as the front azimuth and the azimuth with a large directivity gain, and the difference information and the difference in directivity gain are obtained between the data points adjacent to each other after the sorting. As a result, the data
30 amount of the difference information or the like difference encoded can be kept within the predetermined

bit depth.

[0526]

Also in the method DE3, as in the case of the method DE2, information indicating the arrangement order of the sorted data points is stored in the model data.

[0527]

In the methods DE1 to DE3, the example of obtaining the difference between the data points is described, but the difference is only required to be calculated for at least any one of between the data points or between the bins.

[0528]

Therefore, for example, in each of the methods DE1 to DE3, the rearrangement may be performed in consideration of not only the position of the data point but also the frequency, that is, the bin.

[0529]

In such a case, for example, in the method DE1, the difference information and the directivity gains are sorted in order of predetermined data points and frequencies (bins), and the difference information and the difference in directivity gain adjacent to each other after sorting, that is, the difference between the data points and the bins are obtained. Note that, after sorting is performed in a predetermined order, a difference may be obtained between both data points and bins, or a difference may be obtained only between bins.

[0530]

In addition, for example, in the method DE2, regarding the difference information and the directivity gains sorted in ascending or descending order for the

same bin, the same data point, and the like, the difference information and the difference in directivity gain adjacent to each other, that is, the difference between the data points and the bins can be obtained.

5 [0531]

Similarly, in the method DE3, the difference information and the directivity data in the bins of the data points are sorted according to the priority of the data points and the frequencies (bins), and the
10 difference information and the difference in directivity gain adjacent to each other after sorting, that is, the difference between the data points and the bins can be obtained. In other words, in this example, the data points and bins are sorted in order of priority.

15 [0532]

Note that, in a case where sorting is performed by any of the above methods, sorting may be performed for each group including one or more bins or data points, for example, only bins of the same frequency, only a
20 plurality of bins belonging to a predetermined frequency band, or only bins in the same data point or a plurality of data points adjacent to each other are to be sorted.

[0533]

In addition, each variable (information) in the
25 encoded bit stream such as in the model data may be tabulated, and only an index indicating a value of the tabulated variable may be transmitted.

[0534]

That is, in the example described above, for
30 various variable values such as model parameter in the model data and the like, Syntax is described in the

following manner.

[0535]

· Record variable value in Syntax in floating point format

- 5 · Assign value according to dynamic range and necessary resolution in integer format such as 9 bits (value between 0 to 1 is expressed in 512 stages) and 11 bits

[0536]

- 10 Here, in the floating-point format in which the variable value is recorded, any value can be taken as the variable value in the format of float (32 bits).

[0537]

- On the other hand, in order to actually further
15 reduce the bit depth, Syntax may be described in the following manner.

[0538]

- That is, in a case where the variable value (parameter) to be described often takes a specific value
20 or can be represented by a specific value, a value to be actually used, that is, a variable value to be described is tabulated. Then, only the index obtained by the tabulation is described in the encoded bit stream such as the model data, that is, in Syntax.

- 25 [0539]

- In this case, the table itself is transmitted to the decoding side separately from the encoded bit stream. In this way, the variable value can be described with a small bit depth, and the data amount (amount of
30 transmission) of the encoded bit stream can be reduced.

[0540]

As a specific example, for example, it is conceivable to table only a partial range of possible values of the variable value, such as only the range of 0.0 to 0.1 or only the range of 0.9 to 1.0 of the variable values.

[0541]

In such a case, for example, for each discrete value (variable value) within a range to be tabulated, such as a range of 0.0 to 0.1, an index indicating the value is determined. Then, in a case where the actual variable value is a value within a range to be tabulated, an index corresponding to the actual variable value is stored in the model data or the like and transmitted.

[0542]

On the other hand, in a case where the actual variable value is out of the range to be tabulated, the actual variable value is stored in the model data and transmitted.

[0543]

Furthermore, it is also conceivable to perform parametric expression (compressed expression) on the scale factor "scale_factor[i]" and the offset value, that is, the minimum value "offset[i]" described above.

[0544]

In the above description, using the scale factor "scale_factor[i]" and the minimum value "offset[i]" of each bin, the mixture model $F(x; \Theta)$ is determined in the following manner.

[0545]

For i_bin in bin

$$F(x; \Theta) = F'(x; \Theta) \times scale_factor[i] + offset[i]$$

End

[0546]

where $F'(x; \Theta)$ is an output value of the mixture model for each band.

5 [0547]

In addition, the scale factor "scale_factor[i]" is a ratio between a sum of the vMF distribution and the Kent distribution (model data sum), that is, a sum of the values (directivity gain) at each data point of the mixture model $F'(x; \Theta)$, and a sum of values at the data points of the original (original) directivity data before modeling in the bin indicated by the index i, that is, the i-th bin. This scale factor is a float value representing a dynamic range.

15 [0548]

Note that the model data sum is a sum of values (directivity gains) defined on the spherical surface, and ideally is 1, but it does not become 1 because it is actually discretized. In addition, the original directivity data before modeling is dB-scale data, and is offset in the positive direction when the scale factor is calculated.

[0549]

The minimum value "offset[i]" is the original (original) directivity data before modeling in the i-th bin, that is, the minimum value (dB value) of the directivity gain, and is expressed by a float value.

[0550]

By the calculation using such a scale factor and a minimum value, the output value of the mixture model can be corrected and restored according to the dynamic range

of each bin.

[0551]

In this case, a scale factor and a minimum value corresponding to the number of bins are required, and
5 when the frequency resolution of the directivity data is made high definition, the amount of information required to record the scale factor and the minimum value, that is, the bit depth, increases in proportion to the number of bins.

10 [0552]

Therefore, the amount of information (bit depth) necessary for recording the scale factor and the minimum value may be reduced by parametrically expressing the scale factor and the minimum value.

15 [0553]

For example, as an example, the values illustrated in Figs. 28 and 29 are obtained as the scale factor and the minimum value (offset value) for the directivity data of each of the six sound source types.

20 [0554]

Fig. 28 illustrates scale factors of the six sound source types. Note that, in Fig. 28, the vertical axis represents the value of the scale factor which is a dimensionless ratio, and the horizontal axis represents
25 the index i of the bin.

[0555]

In this example, depending on the sound source type, the scale factor varies greatly between adjacent bins, or the scale factor varies less between adjacent
30 bins.

[0556]

Fig. 29 illustrates the minimum value (offset value) of each of the six sound source types. Note that, in Fig. 29, the vertical axis represents the minimum value (offset value) that is the dB value, and the horizontal axis represents the index i of the bin.

5 [0557]

Even in the minimum value, as in the case of the scale factor, it can be seen that the minimum value greatly fluctuates or the fluctuation is small between adjacent bins depending on the sound source type.

10 [0558]

As described above, the magnitude of the variation of the scale factor or the minimum value greatly differs between adjacent frequencies (between adjacent bins) depending on the sound source type.

15 [0559]

Therefore, when the scale factor and the minimum value are modeled, that is, parametrically expressed, there may be a case where modeling can be performed with a small number of parameters and a case where the number of parameters increases.

20 [0560]

Therefore, for example, in a case where the variation between the bins is large and the coding efficiency cannot be improved by the parametric expression of the scale factor or the minimum value, the model data generation unit 22 and the model data generation unit 215 stores (describes) the scale factor or the minimum value of each bin as it is in the model data.

25 [0561]

30

On the other hand, in a case where the variation between the bins is small and the coding efficiency can be improved, the model data generation unit 22 and the model data generation unit 215 parameterize the scale factor or the minimum value and store (describe) the scale factor or the minimum value in the model data.
[0562]

As an example of the parameterization (parametric representation), for example, curve fitting by function approximation or the like is exemplified.
[0563]

In such a case, the model data generation unit 22 and the model data generation unit 215 generate a function approximation parameter for obtaining an approximation function corresponding to a graph representing a scale factor or a minimum value of each bin by curve fitting or the like. Then, the model data generation unit 22 and the model data generation unit 215 store the function approximation parameter in the model data instead of the scale factor or the minimum value of each bin.
[0564]

On the decoding side, the directivity data calculator 82 and the calculation unit 301 obtain the scale factor or the minimum value in each bin from the approximation function on the basis of the function approximation parameter and the index i of the bin, and use the scale factor or the minimum value as the model parameter.
[0565]

In this way, the scale factors and the minimum

values of all the bins were required to be stored in the model data, but only the function approximation parameters need to be described, and the data amount can be compressed. Note that, as the function approximation,
5 any approximation such as approximation by a linear function, an n-th order function ($n \geq 2$), or polynomial approximation can be performed.

[0566]

Furthermore, in a case where the dynamic range of
10 the scale factor or the minimum value is large, the dynamic range may be compressed by performing, as preprocessing of function approximation, processing of taking the logarithm of the scale factor or the minimum value, processing of converting the scale factor or the
15 minimum value by a nonlinear function, or the like.

[0567]

In addition, the band hybrid method, the addition hybrid method, the multiplication hybrid method, the spherical harmonic coefficient modeling method, and the
20 combination hybrid method have been described above as examples of the method in the case of generating the model data by combining the HOA method, the mixing method, the complex mixing method, and the difference method.

25 [0568]

However, the present invention is not limited thereto, and it is of course possible to generate model data by other combinations.

[0569]

30 For example, the model data may be generated by switching any method such as the HOA method, the mixing

method, the complex mixing method, the difference method, the band hybrid method, or the addition hybrid method described above for each azimuth viewed from the sound source position, that is, for each data point or for each region including a plurality of data points.
[0570]

In the directivity data, there is a high possibility that the frequency of use of the data of the horizontal plane, that is, the data on the equator (directivity gain) is high, and conversely, the frequency of use of the data near the pole is low. Therefore, the bit depth of the model data can be appropriately reduced by switching the method for each region. Note that the horizontal plane here is a plane including a plurality of positions at which the latitude viewed from the sound source position, that is, the elevation angle (elevation), is 0 degrees.
[0571]

As a specific example, for example, it is conceivable to combine the HOA method with the mixing method, more specifically, the method of modeling by the vMF distribution. At this time, for example, the order of the spherical harmonic function expansion in the HOA method may be set to the first order, and whether the HOA method and the mixing method are used in combination or only the mixing method is used may be switched for each region (orientation).
[0572]

In addition, it is also conceivable to generate model data by changing the order of spherical harmonic function expansion in the HOA method for each region.

Furthermore, it is also conceivable to switch between the HOA method and a combination of the mixing method and the HOA method for each region, and to change the order of the spherical harmonic function expansion in the HOA method for each region.
[0573]

In addition, it is also conceivable that data points in the vicinity of the horizontal plane are recorded with high definition by using a method of modeling directivity data by circular harmonic function expansion instead of spherical harmonic function expansion, and for data points other than those in the vicinity of the horizontal plane, the directivity gain is recorded sparsely by another method.

[0574]
<Third embodiment>
<Utilization of symmetry of data>

Meanwhile, the directivity data may have symmetry depending on the shape of the original sound source.
[0575]

For example, a shape of a speaker as a sound source is bilaterally symmetric, and directivity data of the speaker is also symmetric. However, in a case where a tweeter and a woofer are present in the speaker, reproduction bands of the tweeter and the woofer are different, and thus, the directivity data is not symmetric in the up-and-down direction.
[0576]

In addition, a regular dodecahedron speaker and the like are also commercialized, and symmetry is established in 12 directions in the regular dodecahedron speaker.

Further, in the case of a full range speaker having a cubic shape, not only the left-and-right symmetry but also the up-and-down symmetry may be established. On the other hand, a human also has an outer shape that is bilaterally symmetrical, and bilateral symmetry is established to some extent, but has a shape that is not symmetrical in the up-and-down direction with the head, the torso, and the legs, and directivity is not symmetrical in the up-and-down direction.

10 [0577]

From these, in a case where there is symmetry in the directivity data, it is possible to reduce the amount of transmission data by utilizing the symmetry.

[0578]

15 In such a case, Syntax of the model data is, for example, as illustrated in Fig. 30.

[0579]

The model data illustrated in Fig. 30 includes the number of frequency points "bin_count" indicating the number of bins, and the frequency "bin_freq[i]" at the center of the bin is stored by the number of frequency points "bin_count".

[0580]

25 In addition, the number of bands "band_count" is also stored, and the symmetry information "use_symmetry" related to the use of the symmetry of the directivity data is stored for the number of bands "band_count", that is, for each band.

[0581]

30 For example, the values "4", "3", "2", "1", and "0" of the symmetry information "use_symmetry" indicate that

an up-and-down and left-and-right symmetry operation is performed, a left-and-right symmetry operation is performed, an up-and-down symmetry operation is performed, any symmetry and rotation are utilized, and
5 any symmetry and rotation operation is not performed, respectively.

[0582]

Specifically, in a case where the value of the symmetry information "use_symmetry" is "0", the
10 directivity data is described by a model in which directivity gains in all directions are composed of the above-described VMF distribution, Kent distribution, or the like, that is, by a mixture model or the like. Furthermore, values "5" to "7" of the symmetry
15 information "use_symmetry" are reserved.

[0583]

In the model data, operation-related information for a rotation operation or a symmetry operation is stored according to a value of the symmetry information
20 "use_symmetry".

[0584]

In a case where the value of the symmetry information "use_symmetry" is "4", the operation-related information "LeftRightVerticalLineSymmetricDir()" for the
25 up-and-down and left-and-right symmetry operation is described in the model data. In a case where the value of the symmetry information "use_symmetry" is "3", the operation-related information "LeftRightLineSymmetricDir()" for the left-and-right
30 symmetry operation is described in the model data.

[0585]

Furthermore, in a case where the value of the symmetry information "use_symmetry" is "2", the operation-related information "VerticalLineSymmetricDir()" for the up-and-down symmetry operation is described in the model data.

5 [0586]

In a case where the value of the symmetry information "use_symmetry" is "1", the operation-related information "SymmetricDir()" for any symmetry or rotation operation is described in the model data.

10 [0587]

In a case where the value of the symmetry information "use_symmetry" is "0", no operation is performed on the model data, and the information "NonSymmetricDir()" for obtaining the directivity data is described.

15 [0588]

Fig. 31 illustrates Syntax of "SymmetricDir()".

[0589]

In this example, as in the case in Fig. 25, "SymmetricDir()" of the model data stores the number of mixtures "mix_count[j]" and bin information "bin_range_per_band[j]", "kappa[j][k]", "weight[j][k]", "gamma_x[j][k]", "gamma_y[j][k]", and "gamma_z[j][k]" as model parameters, and a selection flag "dist_flag[j][k]".

20 [0590]

In addition, "beta[j][k]", "gamma2_x[j][k]", "gamma2_y[j][k]", "gamma2_z[j][k]", "gamma3_x[j][k]", "gamma3_y[j][k]", and "gamma3_z[j][k]" as model parameters are also stored according to the value of the selection flag "dist_flag[j][k]".

30

[0591]

Further, "SymmetricDir()" stores operation count information "sym_operation_count" and an operation flag "sym_operation_flag".

5 [0592]

The operation count information "sym_operation_count" is information indicating the number of times of performing a rotation operation, which is an operation of rotating and copying, or a symmetry
10 operation, which is an operation of copying to a symmetric position, on one distribution (distribution model) such as the vMF distribution or the Kent distribution.

[0593]

15 The operation flag "sym_operation_flag" is flag information indicating which of the rotation operation and the symmetry operation is performed. For example, in a case where the value of the operation flag "sym_operation_flag" is "1", it indicates that the
20 rotation operation is performed, and in a case where the value is "0", it indicates that the symmetry operation is performed.

[0594]

Specifically, here, the operation flag
25 "sym_operation_flag" is included by the number of times indicated by the operation count information "sym_operation_count", and information necessary for the operation is stored according to the value of the operation flag.

30 [0595]

That is, in a case where the value of the operation

flag "sym_operation_flag" is "1", the rotation axis azimuth angle "sym_azi", the rotation axis elevation angle "sym_elev", and the rotation angle "sym_rotation" required for the rotation operation are stored.

5 [0596]

Here, the rotation axis azimuth angle "sym_azi" and the rotation axis elevation angle "sym_elev" are an azimuth angle and an elevation angle indicating the direction of the rotation axis viewed from the sound
10 source position when the rotation operation is performed. That is, the rotation axis is determined by the rotation axis azimuth angle and the rotation axis elevation angle. Furthermore, the rotation angle "sym_rotation" is an angle at the time of rotation with the rotation axis as
15 the center (axis) in the rotation operation.

[0597]

In addition, in a case where the value of the operation flag "sym_operation_flag" is not "1", that is, in a case where the value of the operation flag is "0", a
20 yaw angle "sym_yaw", a pitch angle "sym_pitch", and a roll angle "sym_roll" indicating a direction of a spherical cross section, that is, a symmetry plane, necessary for the symmetry operation viewed from the sound source position are stored. That is, the symmetry
25 plane is determined by the yaw angle, the pitch angle, and the roll angle.

[0598]

Therefore, for example, in a case where the value of the operation count information "sym_operation_count"
30 is "2", the operation indicated by each of the two operation flags "sym_operation_flag" is performed. That

is, the rotation operation and the symmetry operation are performed twice.

[0599]

Further, as in the case in Fig. 25, the scale
 5 factor "scale_factor[i]" and the minimum value
 "offset[i]" are also stored in "SymmetricDir()" by the
 number of frequency points "bin_count".

[0600]

Here, the rotation operation and the symmetry
 10 operation will be described with reference to Figs. 32
 and 33. Note that the examples illustrated in Figs. 32
 and 33 are examples in which a rotation operation or a
 symmetry operation is performed on the Kent distribution.

[0601]

15 Fig. 32 illustrates an example in which a rotation
 operation is performed on the Kent distribution.

[0602]

In this example, the directivity gain on the sphere
 SP11 is represented by a Kent distribution, and vectors
 20 V81 to V83 represent a vector γ_1 , a major axis vector γ_2 ,
 and a minor axis vector γ_3 of the Kent distribution.

[0603]

These vectors V81 to V83 are obtained by model
 parameter stored in the model data, that is,
 25 "gamma_x[j][k]" to "gamma_z[j][k]" and "gamma2_x[j][k]"
 to "gamma2_z[j][k]".

[0604]

In a case where the rotation operation is
 performed, the directivity data calculator 82 of the
 30 information processing device 51 obtains the rotation
 axis RS11 on the basis of the rotation axis azimuth angle

"sym_azi" and the rotation axis elevation angle
 "sym_elev" read from the model data.

[0605]

The directivity data calculator 82 obtains the Kent
 5 distribution $f(x; \theta_i)$ using the vector V81 to the vector
 V83.

[0606]

Further the directivity data calculator 82 obtains
 the Kent distribution $f(x; \theta_i)$ using the vector V'81 to
 10 the vector V'83.

[0607]

Here, the vector V'81 to the vector V'83 are
 vectors after rotation obtained by rotating the vectors
 V81 to V83 by the rotation angle "sym_rotation" stored in
 15 the model data around the rotation axis RS11.

[0608]

In this case, the vector V'81 to the vector V'83
 are used as the vector γ_1 , the major axis vector γ_2 , and
 the minor axis vector γ_3 of the Kent distribution.

20 [0609]

Therefore, in this example, the directivity data
 calculator 82 calculates the rotated model parameter by
 performing a rotation operation on the model parameter
 such as the vector γ_1 of the Kent distribution on the
 25 basis of the rotation axis azimuth angle and the like.
 Then, the directivity data calculator 82 obtains the Kent
 distribution on the basis of each of the model parameter
 before rotation and the rotated (post-rotation) model
 parameter, and calculates the mixture model, that is, the
 30 directivity data (directivity gain) using the obtained
 Kent distribution. In other words, one distribution is

obtained by synthesis from the Kent distribution obtained from the model parameter before the rotation operation and the Kent distribution obtained from the model parameter after the rotation operation, and the mixture model is obtained using the distribution. Note that the two Kent distributions may be directly used for calculation of the mixture model, or only a partial region of each of the two Kent distributions, such as the right half and the left half, may be used for calculation of the mixture model. This applies not only to the case of the rotation operation but also to the case of the symmetry operation.

[0610]

Fig. 33 illustrates an example in which a symmetry operation is performed on the Kent distribution. Note that, in Fig. 33, portions corresponding to those in a case of Fig. 32 are denoted by the same reference numerals, and description thereof will be omitted as appropriate.

[0611]

In this example, the directivity data calculator 82 obtains the cross section SF11 of the sphere SP11 to be the symmetric plane on the basis of the yaw angle "sym_yaw", the pitch angle "sym_pitch", and the roll angle "sym_roll" read from the model data. The cross section SF11 is a plane including the center (sound source position) of the sphere SP11.

[0612]

The directivity data calculator 82 obtains the Kent distribution $f(x; \theta_i)$ using the vector V81 to the vector V83.

[0613]

Further, the directivity data calculator 82 obtains the Kent distribution $f(x; \theta_i)$ using the vector $V''81$ to the vector $V''83$.

5 [0614]

Here, the vector $V''81$ to the vector $V''83$ are vectors obtained by folding (symmetrically moving) the vectors $V81$ to $V83$ with the cross-section $SF11$ as a symmetry plane. That is, the vectors $V''81$ to $V''83$ and
 10 the vectors $V81$ to $V83$ are symmetric (plane-symmetric) with respect to the cross-section $SF11$.

[0615]

In the directivity data calculator 82, the vector $V''81$ to the vector $V''83$ are used as the vector γ_1 , the
 15 major axis vector γ_2 , and the minor axis vector γ_3 of the Kent distribution.

[0616]

Therefore, in this example, the directivity data calculator 82 performs a symmetry operation on the model
 20 parameter such as the vector γ_1 of the Kent distribution on the basis of the yaw angle and the like, thereby calculating the model parameter symmetrically moved (symmetry operation). Then, the directivity data calculator 82 obtains the Kent distribution on the basis
 25 of each of the model parameter before the symmetrical movement and the symmetrically moved (symmetrically moved) model parameter, and calculates the directivity data (directivity gain) from the obtained Kent distribution and the like.

30 [0617]

Fig. 34 illustrates an example of Syntax of the

information "NonSymmetricDir()" for obtaining the directivity data in the model data illustrated in Fig. 30.

[0618]

5 In the example illustrated in Fig. 34, as in the case in Fig. 25, the number of mixtures "mix_count[j]" and bin information "bin_range_per_band[j]", "kappa[j][k]", "weight[j][k]", "gamma_x[j][k]", "gamma_y[j][k]", and "gamma_z[j][k]" as model parameters, 10 and a selection flag "dist_flag[j][k]" are stored.

[0619]

 In addition, "beta[j][k]", "gamma2_x[j][k]", "gamma2_y[j][k]", "gamma2_z[j][k]", "gamma3_x[j][k]", "gamma3_y[j][k]", and "gamma3_z[j][k]" as model 15 parameters are also stored according to the value of the selection flag "dist_flag[j][k]".

[0620]

 Further, a scale factor "scale_factor[i]" and a minimum value "offset[i]" are also stored by the number 20 of frequency points "bin_count".

[0621]

 In this example, since the rotation operation and the symmetry operation are not performed, the model parameter constituting all the distributions are 25 described in the model data.

[0622]

 In addition, the data format (Syntax) of the operation-related information "LeftRightVerticalLineSymmetricDir()", 30 "LeftRightLineSymmetricDir()", and "VerticalLineSymmetricDir()" is the same as

"NonSymmetricDir()" illustrated in Fig. 34.

[0623]

However, in a case where these

"LeftRightVerticalLineSymmetricDir()",

5 "LeftRightLineSymmetricDir()", or

"VerticalLineSymmetricDir()" are stored, that is, in a

case where the value of the symmetry information

"use_symmetry" is "4", "3", or "2", the directivity data

calculator 82 performs the symmetry operation at the time

10 of decoding the directivity data.

[0624]

Specifically, in a case where the value of the

symmetry information "use_symmetry" is "3", the

directivity data calculator 82 performs a left-and-right

15 symmetry operation for the front median plane on the

distribution corresponding to the model parameter

described in the model data, and obtains a new vMF

distribution or Kent distribution.

[0625]

20 The left-and-right symmetry operation performed in

this case is a symmetry operation in which the front

median plane (median plane) viewed from the sound source

is a cross section SF11 illustrated in Fig. 33. In other

words, the left-and-right symmetry operation is realized

25 by performing the symmetry operation described with

reference to Fig. 33 with the median plane as the cross

section SF11. In this case, when the distribution

obtained from the model parameter before the left-and-

right symmetry operation and the distribution obtained

30 from the model parameter after the left-and-right

symmetry operation are combined, one distribution that is

left-and-right symmetric when viewed from the sound source is obtained.

[0626]

Furthermore, in a case where the value of the symmetry information "use_symmetry" is "2", the directivity data calculator 82 performs an up-and-down symmetry operation for the front horizontal plane on the distribution corresponding to the model parameter described in the model data with respect to the front horizontal plane, and obtains a new vMF distribution or Kent distribution.

[0627]

The up-and-down symmetry operation performed in this case is a symmetry operation in which the front horizontal plane (horizontal plane) viewed from the sound source is the cross-section SF11 illustrated in Fig. 33. In other words, the up-and-down symmetry operation is realized by performing the symmetry operation described with reference to Fig. 33 with the horizontal plane as the cross section SF11. In this case, when the distribution obtained from the model parameter before the up-and-down symmetry operation and the distribution obtained from the model parameter after the up-and-down symmetry operation are combined, one distribution that is symmetric in the up-and-down direction when viewed from the sound source is obtained.

[0628]

Furthermore, in a case where the value of the symmetry information "use_symmetry" is "4", the directivity data calculator 82 performs an up-and-down and left-and-right symmetry operation for the front face

on the distribution corresponding to the model parameter described in the model data, and obtains a new distribution. Here, the up-and-down and left-and-right symmetry operation is an operation of obtaining a symmetrical distribution in the up-and-down and left-and-right directions by performing an up-and-down symmetry operation and a left-and-right symmetry operation on the distribution to be operated. Note that the vMF distribution and the Kent distribution on which the symmetry operations including the left-and-right symmetry operation and the up-and-down symmetry operation are performed are effective over the entire spherical surface where the directivity data is defined at the time of decoding (at the time of restoration). In addition, a boundary may be defined in the distribution to be operated or the distribution obtained by the operation, and the directivity gain may be discontinuous at the boundary.

[0629]

20 <Fourth embodiment>

<Crossfade between bands>

In the above, the method of reducing the data amount by modeling the directivity data for each frequency band, that is, for each band is described.

25 [0630]

However, there is a case where the directivity data, that is, the directivity gain gradually fluctuates according to the frequency, and in such a case, the crossfade of the modeled data is considered to be effective.

[0631]

Fig. 35 illustrates an example of Syntax of the model data in a case where the crossfade is performed.

[0632]

In the example illustrated in Fig. 35, in addition to the information illustrated in Fig. 25, a crossfade flag "fade_flag" and an upper limit bin index "bin_range_per_band_fadein[j]" are further stored (included).

[0633]

Specifically, in this example, the crossfade flag "fade_flag" in each band is stored by the number of bands "band_count".

[0634]

The crossfade flag "fade_flag" is flag information indicating whether or not perform a crossfade between bands adjacent to each other, that is, perform weighted addition of the mixture model $F'(x; \theta)$ for each band, in calculating the mixture model $F(x; \theta)$ for each bin.

[0635]

For example, in a case where the value of the crossfade flag "fade_flag" is "1", the crossfade between the bands is performed, and in a case where the value is "0", the crossfade between the bands is not performed. Note that the crossfade between the bands is used in the second or subsequent bands.

[0636]

In addition, in a case where the value of the crossfade flag "fade_flag" is "1", the upper limit bin index "bin_range_per_band_fadein[j]" is stored.

[0637]

The upper limit bin index

"bin_range_per_band_fadein[j]" is an index indicating the upper limit bin in which the inter-band crossfade is performed, that is, the bin having the highest frequency among the bins in the band in which the inter-band crossfade is performed.

5 [0638]

In the crossfade between bands, the directivity data calculator 82 performs weighted addition on the output value $F'(x; \theta)$ of the mixture model obtained for the predetermined band and the output value $F'(x; \theta)$ of the mixture model obtaining for another band adjacent to the predetermined band.

10 [0639]

Then, the directivity data calculator 82 multiplies the output value obtained by the weighted addition by the scale factor, and sets a value obtained by adding the minimum value (offset value) to the multiplication result to the output value $F(x; \theta)$ of the mixture model in the target bin in another band.

15 [0640]

In this case, the target of the crossfade is each bin from the bin with the lowest frequency in the another band to the upper limit bin indicated by the upper limit bin index "bin_range_per_band_fadein[j]" in the another band, and the crossfade is not performed in the another bin. For a bin in which no crossfade is performed, the output value $F(x; \theta)$ of the mixture model is obtained from the output value $F'(x; \theta)$ of the mixture model in the band to which the bin belongs, the scale factor, and the minimum value.

20 [0641]

Therefore, in a case where the crossfade between the bands is performed, in the calculation of the directivity data (directivity gain), a procedure of setting the weighted sum (weighted addition value) of the output values of the restored mixture models between the adjacent bands for the output value of the mixture model of the final band is added before the scale factor and the minimum value are applied.

[0642]

10 Fig. 36 illustrates a conceptual diagram of the crossfade between bands.

[0643]

In Fig. 36, the vertical axis represents the weight used at the time of crossfading, and the horizontal axis represents the frequency. In addition, here, a case where the number of bands is three is illustrated as an example.

[0644]

In the drawing, the weight at the time of the weighted addition in a case where the crossfade between the bands is not performed is illustrated on the left side.

[0645]

25 Straight lines L51 to L53 shows a weight of the output value $F'(x; \theta)$ of the mixture model for each band from band "bin_range_per_band[0]" to band "bin_range_per_band[2]", the weight being used to calculate the output value $F(x; \theta)$ of the mixture model for each bin.

30 [0646]

In particular, in this example: the ranges in the

frequency direction of straight lines L51 to L53 do not overlap with each other, and the weight of the output value $F'(x; \theta)$ of the mixture model for each band for each bin (frequency) is 1. Therefore, it can be seen
 5 that crossfade between bands is not substantially performed.

[0647]

On the other hand, the weight at the time of the weighted addition in a case where the crossfade between
 10 the bands is performed is illustrated on the right side in the figure.

[0648]

Broken lines L61 to L63 shows the weight of the output value $F'(x; \theta)$ of the mixture model for each band
 15 from band "bin_range_per_band[0]" to band "bin_range_per_band[2]", which are used to calculate the output value $F(x; \theta)$ of the mixture model for each bin.

[0649]

In this example, the right end of the broken line
 20 L61 indicating the weight of the output value $F'(x; \theta)$ of for example, a mixture model of the band "bin_range_per_band[0]" is located at a frequency position outside the range of the band "bin_range_per_band[0]".

25 [0650]

Specifically, the frequency (bin) of the end portion on the right side of the broken line L61 is a bin in the band "bin_range_per_band[1]" adjacent to the band "bin_range_per_band[0]", and this bin is the upper limit
 30 bin "bin_range_per_band_fadein[1]".

[0651]

Therefore, for example, for each bin between the lowest frequency bin and the upper limit bin "bin_range_per_band_fadein[1]" among the bins in the band "bin_range_per_band[1]", it can be seen that the output value $F(x; \Theta)$ of the mixture model for each bin is obtained by performing the crossfading between bands. In this case, in each bin, the weights are calculated so that the sum of the weights used to calculate the output value $F(x; \Theta)$ of the mixture model is 1.

10 [0652]

On the other hand, for each bin having a frequency higher than the upper limit bin among the bins in the band "bin_range_per_band[1]", the value of the weight indicated by the broken line L62 is 1, and it can be seen that the crossfade between the bands is not performed in the bin.

[0653]

In a case where the crossfade between the bands is performed, the weight $\text{model_weight}_{i_band-1}[i_bin]$ of the output value of the mixture model of the lower frequency band "i_band-1" for a predetermined bin "i_bin" can be obtained by the following Expression (10).

[0654]

[Math. 10]

$\text{model_weight}_{i_band-1}[i_bin] =$

$$\begin{cases} (i_bin - \text{bin_range_per_band}[i_band-1]) / (\text{bin_range_per_band_fadein}[i_band] - \text{bin_range_per_band}[i_band-1]) & (i_bin < \text{bin_range_per_band_fadein}[i_band]) \\ 1.0 & (i_bin \geq \text{bin_range_per_band_fadein}[i_band]) \end{cases}$$

25

• • • (10)

[0655]

Furthermore, the weight $\text{model_weight}_{i_band}[i_bin]$ of

the output value of the mixture model of the higher frequency band "i_band" for the predetermined bin "i_bin" can be obtained by the following Expression (11).

[0656]

5 [Math. 11]

$$\text{model_weight}_{i_band}[i_bin] = 1 - \text{model_weight}_{i_band-1}[i_bin] \quad \cdot \cdot \cdot (11)$$

[0657]

Furthermore, the output value $F_{i_bin}(x)$ of the mixture model for the bin "i_bin" can be obtained by calculating the following Expression (12) on the basis of the weight and the output value $F_{i_band+j-1}(x)$ of the mixture model for the band "i_band + j-1" for the bin "i_bin".

[0658]

15 [Math. 12]

$$F_{i_bin}(x) = \left(\sum_{j=0}^1 \text{model_weight}_{i_band+j-1}[i_bin] * F_{i_band+j-1}(x) \right) * \text{scale_factor}[i_bin] + \text{offset}[i_bin] \quad \cdot \cdot \cdot (12)$$

[0659]

Note that $\text{scale_factor}[i_bin]$ and $\text{offset}[i_bin]$ in Expression (12) indicate the scale factor and the minimum value (offset value) of the bin "i_bin".

[0660]

The directivity data calculator 82 calculates Expression (12) to calculate the output value of the mixture model of each bin, that is, the directivity gain of each bin of each data point. In this way, the data amount of the model data can be reduced.

[0661]

<Fifth embodiment>

<Utilization of symmetry of data>

In the third embodiment, utilization of data
5 symmetry is described.

[0662]

In a case where the amount of transmission data is
reduced by utilizing such symmetry, it is also possible
to utilize the symmetry in the front-rear direction in
10 addition to the up-and-down symmetry and the left-and-
right symmetry of the directivity data described above,
and to utilize a combination of symmetries in the front-
rear direction, the up-and-down direction, and the left-
and-right direction.

15 [0663]

In such a case, Syntax of the model data is, for
example, as illustrated in Fig. 37.

[0664]

The model data illustrated in Fig. 37 includes the
20 number of frequency points "bin_count" indicating the
number of bins, and the frequency "bin_freq[i]" at the
center of the bin is stored by the number of frequency
points "bin_count".

[0665]

25 In addition, the model data also stores the number
of bands "band_count", and symmetry information
"use_symmetry[j]", the number of mixtures "mix_count[j]",
and bin information "bin_range_per_band[j]" related to
the use of the symmetry of the directivity data are
30 stored by the number of bands "band_count", that is, for
each band.

[0666]

The symmetry information "use_symmetry[j]" is similar to the symmetry information "use_symmetry" illustrated in Fig. 30, but in this example, the values "5" to "7" of the symmetry information "use_symmetry[j]" are used without being reserved as described later.

[0667]

Furthermore, the number of mixtures "mix_count[j]" and the bin information "bin_range_per_band[j]" are similar to those illustrated in Fig. 31, and are information indicating the number of distributions constituting the mixture model of bands and a bin (bin) for the original directivity data before modeling.

[0668]

In the example illustrated in Fig. 30, the number of mixtures "mix_count[j]" and the bin information "bin_range_per_band[j]" are stored for each piece of operation-related information and the like. However, since the number of mixtures and the bin information are the same, the number of mixtures and the bin information are stored in a portion other than the operation-related information in the model data in the example of Fig. 37.

[0669]

In the example of Fig. 37, the value of the symmetry information "use_symmetry[j]" for each band is any value of "0" to "7".

[0670]

The values "4", "3", "2", "1", and "0" of the symmetry information "use_symmetry[j]" indicate that the up-and-down and left-and-right symmetry operation is performed, the left-and-right symmetry operation is

performed, the up-and-down symmetry operation is performed, any symmetry and rotation are utilized, and any symmetry and rotation operation is not performed, as in the example of Fig. 30.

5 [0671]

Values "7", "6", and "5" of the symmetry information "use_symmetry[j]" indicate that an up-and-down and front-and-back symmetry operation is performed, a front-and-back and left-and-right symmetry operation is performed, and a front-and-back symmetry operation is performed.

[0672]

In a case where the number of bands "band_count" is larger than 0 ($j > 0$), the crossfade flag "fade_flag" in each band is stored in the model data.

[0673]

The crossfade flag "fade_flag" is the same as that described with reference to Fig. 35. That is, in a case where the value of the crossfade flag "fade_flag" is "1", the crossfade between the bands is performed, and in a case where the value is "0", the crossfade between the bands is not performed.

[0674]

In addition, in a case where the value of the crossfade flag "fade_flag" is "1", an upper limit bin index "bin_range_per_band_fadein[j]" for the band is stored in the model data.

[0675]

In addition, a start bin "start_bin" is stored in the model data.

[0676]

The original directivity data before modeling may not substantially include data for a bin having a lower frequency among the bins indicated by the frequency "bin_freq[i]". That is, the directivity gain of a bin with a low frequency may be 0.

5 [0677]

The start bin "start_bin" is information indicating the bin having the lowest frequency in which the directivity gain that is not 0 is included as data among the bins indicated by the frequency "bin_freq[i]".

10 [0678]

Furthermore, operation-related information for a rotation operation or a symmetry operation is stored in the model data according to the value of the symmetry information "use_symmetry[j]".

15 [0679]

In a case where the value of the symmetry information "use_symmetry[j]" is "7", the operation-related information "FrontBackVerticalSymmetricDir()" for the up-and-down and front-and-back symmetry operation is described in the model data. In a case where the value of the symmetry information "use_symmetry[j]" is "6", the operation-related information "FrontBackLeftRightSymmetricDir()" for the front-and-back and left-and-right symmetry operation is described in the model data.

20 [0680]

Furthermore, in a case where the value of the symmetry information "use_symmetry[j]" is "5", the operation-related information "FrontBackSymmetricDir()" for the front-and-back symmetry operation is described in

30

the model data.

[0681]

In a case where the value of the symmetry information "use_symmetry[j]" is "4", the operation-
 5 related information "LeftRightVerticalLineSymmetricDir()" is described in the model data. In a case where the value of the symmetry information "use_symmetry[j]" is "3", the operation-related information "LeftRightLineSymmetricDir()" is described in the model
 10 data.

[0682]

Further, in a case where the value of the symmetry information "use_symmetry[j]" is "2", the operation-related information "VerticalLineSymmetricDir()" is
 15 described in the model data.

[0683]

In a case where the value of the symmetry information "use_symmetry[j]" is "1", the operation-related information "SymmetricDir()" is described in the
 20 model data. In a case where the value of the symmetry information "use_symmetry[j]" is "0", the information "NonSymmetricDir()" is described in the model data.

[0684]

Furthermore, information regarding the dynamic
 25 range "DynamicRangeForDir()" is described in the model data.

[0685]

The information "DynamicRangeForDir()" stores a scale factor "scale_factor[i]" and a minimum value
 30 "offset[i]" for each bin in which the center frequency is equal to or higher than the center frequency of the bin

indicated by the start bin "start_bin".

[0686]

Fig. 38 illustrates an example of Syntax of the information "NonSymmetricDir()" for obtaining the
 5 directivity data in the model data illustrated in Fig. 37.

[0687]

In the example illustrated in Fig. 38,
 "kappa[j][k]", "weight[j][k]", "gamma_azi[j][k]", and
 10 "gamma_elev[j][k]" as model parameters and a selection
 flag "dist_flag[j][k]" are stored by the number of
 mixtures "mix_count[k]".

[0688]

Here, "gamma_azi[j][k]" and "gamma_elev[j][k]"
 15 indicate a horizontal direction angle (azimuth angle) and
 a vertical direction angle (elevation angle) indicating
 the direction of the vector γ_1 .

[0689]

In the example of Fig. 34, the vector γ_1 is
 20 expressed by "gamma_x[j][k]", "gamma_y[j][k]", and
 "gamma_z[j][k]", but in Fig. 38, the vector γ_1 is
 expressed by an azimuth angle and an elevation angle.

[0690]

In addition, "beta[j][k]" and "gamma1_azi[j][k]" as
 25 model parameters are also stored according to the value
 of the selection flag "dist_flag[j][k]".

[0691]

"gamma1_azi[j][k]" is an angle (rotation angle) in
 the horizontal direction indicating the relative
 30 direction of the major axis vector γ_2 and the minor axis
 vector γ_3 when viewed from the vector γ_1 .

[0692]

That is, in this example, the major axis vector γ_2 and the minor axis vector γ_3 can be obtained from the vector γ_1 and the angle "gamma1_azi[j][k]".

5 [0693]

Fig. 39 illustrates an example of Syntax of the operation-related information

"LeftRightLineSymmetricDir()".

[0694]

10 In this example, as in the case of
 "NonSymmetricDir()" in Fig. 38, "kappa[j][k]",
 "weight[j][k]", "gamma_azi[j][k]", and "gamma_elev[j][k]"
 as model parameters, and a selection flag
 "dist_flag[j][k]" are stored by the number of mixtures
 15 "mix_count[k]".

[0695]

In addition, "beta[j][k]" and "gamma1_azi[j][k]" as
 model parameters are also stored according to the value
 of the selection flag "dist_flag[j][k]".

20 [0696]

Further, the operation-related information
 "LeftRightLineSymmetricDir()" stores "sym_flag[k]" for
 each distribution (mixture) such as Kent distribution or
 vMF distribution constituting a mixture model
 25 representing the distribution of the directivity gain in
 the band by the number of mixtures "mix_count[k]".

[0697]

"sym_flag[k]" is flag information indicating
 whether or not to perform an operation such as symmetry
 30 or rotation on a target distribution. For example, a
 value "00" of the flag information "sym_flag[k]"

indicates that an operation such as symmetry or rotation is not performed, and the value "01" of the flag information "sym_flag[k]" indicates that a symmetry operation is performed.

5 [0698]

Therefore, for example, in a case where the value of the flag information "sym_flag[k]" of the predetermined distribution stored in the operation-related information "LeftRightLineSymmetricDir()" is

10 "01", a left-and-right symmetry operation is performed on the distribution.

[0699]

The data format (Syntax) of the operation-related information "FrontBackVerticalSymmetricDir()",

15 "FrontBackLeftRightSymmetricDir()",

"FrontBackSymmetricDir()",

"LeftRightVerticalLineSymmetricDir()",

"VerticalLineSymmetricDir()", and "SymmetricDir()" in the model data is similar to "LeftRightLineSymmetricDir()" in

20 Fig. 39.

[0700]

In this case, the flag information "sym_flag[k]" in each piece of operation-related information is flag information indicating whether or not to perform an operation corresponding to the operation-related information.

25

[0701]

Specifically, for example, in a case where the value of the flag information "sym_flag[k]" of the predetermined distribution (mixture) stored in the operation-related information

30

"VerticalLineSymmetricDir()" is "01", the up-and-down symmetry operation is performed on the distribution.

[0702]

Furthermore, for example, in addition to the
5 information stored in the operation-related information
"LeftRightLineSymmetricDir()" illustrated in Fig. 39,
information necessary for the rotation operation and the
symmetry operation is also stored in the operation-
related information "SymmetricDir()" according to the
10 value of the flag information "sym_flag[k]".

[0703]

Specifically, for example, the rotation axis
azimuth angle "sym_azi", the rotation axis elevation
angle "sym_elev", the rotation angle "sym_rotation", the
15 yaw angle "sym_yaw", the pitch angle "sym_pitch", and the
roll angle "sym_roll" described with reference to Fig. 31
are appropriately stored in the operation-related
information. Then, a rotation operation or a symmetry
operation is performed for each distribution constituting
20 the mixture model according to the value of the flag
information "sym_flag[k]". In this case, a combination
of operations to be executed, such as only a rotation
operation, only a symmetry operation, and both the
rotation operation and the symmetry operation, can be
25 designated by a value of the flag information
"sym_flag[k]".

[0704]

Note that the configuration of the operation-
related information "SymmetricDir()" may be similar to
30 the configuration of the example illustrated in Fig. 31,
and the presence or absence of execution of the rotation

operation and the symmetry operation may be defined by the operation count information "sym_operation_count" and the operation flag "sym_operation_flag".

[0705]

5 Further, in a case where the operation-related information "FrontBackVerticalSymmetricDir()", "FrontBackLeftRightSymmetricDir()", or "FrontBackSymmetricDir()" is stored in the model data, that is, in a case where the value of the symmetry
10 information "use_symmetry[j]" is "7", "6", or "5", the directivity data calculator 82 performs a symmetry operation at the time of decoding the directivity data.

[0706]

Specifically, in a case where the value of the
15 symmetry information "use_symmetry[j]" is "7", the directivity data calculator 82 performs an up-and-down and front-and-back symmetry operation on the distribution in which the value of the flag information "sym_flag[k]" is "01", and obtains a new distribution.

20 [0707]

Then, the directivity data calculator 82 calculates the directivity data (directivity gain) from the new distribution and the like. In addition, thereafter, the crossfade between bands is also appropriately performed
25 according to the value of the crossfade flag "fade_flag" for each band.

[0708]

Here, the up-and-down and front-and-back symmetry operation is an operation of obtaining an up-and-down and
30 front-and-back symmetric distribution by performing an up-and-down symmetry operation and a front-and-back

symmetry operation on a distribution to be operated.

[0709]

The up-and-down symmetry operation performed in this case is a symmetry operation in which the front
5 horizontal plane (horizontal plane) viewed from the sound source is the cross-section SF11 illustrated in Fig. 33. In other words, the up-and-down symmetry operation is realized by performing the symmetry operation described with reference to Fig. 33 with the horizontal plane as
10 the cross section SF11.

[0710]

In addition, the front-and-back symmetry operation is a symmetry operation in which a plane obtained by rotating the front median plane (median plane) viewed
15 from the sound source by 90 degrees in the horizontal direction is a cross section SF11 illustrated in Fig. 33. In other words, the front-and-back symmetry operation is realized by performing the symmetry operation described with reference to Fig. 33 with a plane obtained by
20 rotating the front median plane by 90 degrees in the horizontal direction as the cross-section SF11.

[0711]

In a case where the value of the symmetry information "use_symmetry[j]" is "6", the directivity
25 data calculator 82 performs a front-and-back and left-and-right symmetry operation on the distribution in which the value of the flag information "sym_flag[k]" is "01" to obtain a new distribution, and calculates the directivity data using the obtained distribution.

30 [0712]

The front-and-back and left-and-right symmetry

operation is an operation of obtaining a symmetrical distribution in the front-back and left-and-right directions by performing the front-and-back symmetry operation and the left-and-right symmetry operation on the distribution to be operated. The left-and-right symmetry operation performed in this case is a symmetry operation in which the front median plane (median plane) viewed from the sound source is a cross section SF11 illustrated in Fig. 33.

10 [0713]

Further, for example, in a case where the value of the symmetry information "use_symmetry[j]" is "5", the directivity data calculator 82 performs a front-and-back symmetry operation on the distribution in which the value of the flag information "sym_flag[k]" is "01" to obtain a new distribution, and calculates the directivity data using the obtained distribution.

[0714]

Note that the distribution such as the vMF distribution and the Kent distribution on which the symmetry operations including the left-and-right symmetry operation, the up-and-down symmetry operation, and the front-and-back symmetry operation are performed is effective over the entire spherical surface where the directivity data is defined at the time of decoding (at the time of restoration). In addition, a boundary may be defined in the distribution to be operated or the distribution obtained by the operation, and the directivity gain may be discontinuous at the boundary.

20 25 30 [0715]

In addition, in the fifth embodiment, regarding the

operation of symmetry or rotation defined by the symmetry information "use_symmetry[j]" for each band, whether or not the operation of symmetry or rotation is actually performed for each distribution (mixture) such as the
5 Kent distribution constituting the mixture model of the bands is defined by the flag information "sym_flag[k]".
[0716]

However, the invention is not limited thereto, and an operation of symmetry or rotation to be executed for
10 each distribution (mixture) such as Kent distribution constituting the band mixture model may be defined.
[0717]

In such a case, for example, for each distribution, it is conceivable to appropriately store 1-bit symmetry
15 information "use_symmetry" and 3-bit flag information "sym_flag[k]" in operation-related information or the like of model data, and define an operation to be performed for each distribution.
[0718]

20 In this example, for example, 1-bit symmetry information "use_symmetry" is flag information indicating whether or not to perform an operation such as symmetry or rotation.
[0719]

25 Specifically, for example, in a case where the value of the symmetry information "use_symmetry" is "1", an operation such as symmetry or rotation is performed, and in a case where the value of the symmetry information "use_symmetry" is "0", an operation such as symmetry or
30 rotation is not performed.
[0720]

Furthermore, in a case where the value of the symmetry information "use_symmetry" is "0", since the operation of symmetry or rotation is not performed on the target distribution, the flag information "sym_flag[k]" for the distribution is not stored in the operation-related information or the like.

[0721]

On the other hand, in a case where the value of the symmetry information "use_symmetry" is "1", since the operation of symmetry or rotation is performed on the target distribution, the flag information "sym_flag[k]" for the distribution is stored in the operation-related information and the like.

[0722]

Then, the directivity data calculator 82 performs an operation according to the value of the flag information "sym_flag[k]", and a new distribution is obtained.

[0723]

At this time, for the values "0", "1", "2", "3", "4", "5", "6", and "7" of the flag information "sym_flag[k]", for example, no operation, any symmetry, rotation operation, an up-and-down symmetry operation, a left-and-right symmetry operation, an up-and-down and left-and-right symmetry operation, a front-and-back symmetry operation, a front-and-back and left-and-right symmetry operation, and an up-and-down and front-and-back symmetry operation may be assigned.

[0724]

<Others>

When calculating the rough directivity data

(directivity data) such as step S52 in Fig. 12 and step S117 in Fig. 20, the directivity data calculator 82 calculates a mixture model $F'(x; \theta)$ for each band on the basis of the model parameter.

5 [0725]

At this time, the directivity data calculator 82 performs weighted addition on a plurality of distributions constituting the mixture model, such as the Kent distribution, the vMF distribution, and the complex
10 Bingham distribution obtained from the model parameter, by using the weight φ_i of the distributions, that is, the weight[j][k] and the weight[i_band][i_mix] described above, to obtain the mixture model $F'(x; \theta)$ (directivity data).

15 [0726]

The value of the weight φ_i of each distribution is determined such that the sum of the weights φ_i of the plurality of distributions constituting the mixture model is 1, but the value of each weight φ_i may be a positive
20 value or a negative value.

[0727]

For example, by setting the weight φ_i of some distributions to a negative value, it is possible to provide not only a steep peak but also a dip in the
25 mixture model, as in the relationship between the bandpass and the band reject filter in the filter.

[0728]

For example, in a case where the weight φ_i of one distribution such as the Kent distribution or the vMF
30 distribution constituting the mixture model is a positive value, when the distribution is multiplied by the weight

φ_i , the distribution after the multiplication of the weight φ_i is as indicated by the arrow Q101 in Fig. 40, for example.

[0729]

5 Note that, in Fig. 40, the lateral direction indicates a predetermined direction on the spherical surface in the distribution such as the Kent distribution defined on the spherical surface, and the vertical direction indicates a value at each position of the
10 distribution, that is, the directivity gain.

[0730]

 In the example indicated by the arrow Q101, it can be seen that there is an upward convex peak in the figure in the distribution after multiplication by the weight
15 φ_i .

[0731]

 On the other hand, in a case where the weight φ_i of one distribution such as the Kent distribution or the vMF distribution constituting the mixture model is a negative
20 value, when the distribution is multiplied by the weight φ_i , the distribution after the multiplication by the weight φ_i is as indicated by the arrow Q102, for example. In this example, it can be seen that there is a downward convex dip in the distribution after multiplication by
25 the weight φ_i in the drawing.

[0732]

 Therefore, when the weight φ_i of each distribution including a negative value is appropriately determined on the condition that the sum of the weights φ_i of all the
30 distributions constituting the mixture model for the band is 1, the degree of freedom is further increased, and the

mixture models having more various shapes can be expressed.

[0733]

As described above, even in a case where the weight φ_i of any distribution is set to a negative value, when the sum of the weights φ_i of all the distributions is set to 1 (1.0), the generality is not lost.

[0734]

Furthermore, in a case where a negative value can also be taken as the value of the weight φ_i , for example, high-order 1 bit of a weight of 10 bits, `weight[j][k]`, as the weight φ_i in Fig. 25, Fig. 31, Fig. 34, or the like is used as the sign bit. The same applies to, for example, the weight `weight[i_band][i_mix]` in Fig. 5.

15 [0735]

<Configuration example of computer>

Note that, the above-described series of processes may be executed by hardware or software. In a case where the series of processes are executed by the software, a program constituting the software is installed on a computer. Here, examples of the computer include a computer incorporated in dedicated hardware, and for example, a general-purpose personal computer capable of executing various functions by installing various programs.

25

[0736]

Fig. 41 is a block diagram illustrating a configuration example of hardware of a computer that executes the above-described series of processes by a program.

30

[0737]

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.

[0738]

5 Moreover, an input/output interface 505 is 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.

10 [0739]

 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
15 memory, and the like. The communication unit 509 includes a network interface and the like. The drive 510 drives a removable recording medium 511 such as a magnetic disk, an optical disc, a magneto-optical disk, or a semiconductor memory.

20 [0740]

 In the computer configured as described above, the CPU 501 loads, for example, a 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,
25 so as to execute the above-described series of processes.

[0741]

 The program executed by the computer (CPU 501) can be provided by being recorded on the removable recording medium 511 as a package medium, or the like, for example.
30 Furthermore, the program may be provided via a wired or wireless transmission medium, such as a local area

network, the Internet, or digital satellite broadcasting.
[0742]

In the computer, the program can be installed in
the recording unit 508 via the input/output interface 505
5 by mounting the removable recording medium 511 to the
drive 510. Furthermore, the program can be received by
the communication unit 509 via the wired or wireless
transmission medium to be installed on the recording unit
508. In addition, the program can be installed in the
10 ROM 502 or the recording unit 508 in advance.
[0743]

Note that, the program executed by the computer may
be a program that is processed in time series in the
order described in the present specification, or a
15 program that is processed in parallel or at a necessary
timing such as when a call is made.
[0744]

Furthermore, the embodiments of the present
technology are not limited to the above-described
20 embodiments, and various modifications are possible
without departing from the scope of the present
technology.
[0745]

For example, the present technology may be
25 configured as cloud computing in which one function is
shared by a plurality of devices via a network to process
together.
[0746]

In addition, each step described in the above
30 flowcharts can be executed by one device or shared and
performed by a plurality of devices.

[0747]

Moreover, in a case where a plurality of processing steps is included in one step, the plurality of processing included in the one step can be performed by one device or shared and performed by a plurality of devices.

[0748]

Moreover, the present technology may also have following configurations.

10 [0749]

(1)

An information processing device including an acquisition unit configured to acquire model data obtained by modeling directivity data representing directivity of a sound source, and

a calculator configured to calculate the directivity data on the basis of the model data.

(2)

The information processing device according to Item (1), in which

the model data includes a model parameter constituting a mixture model, the model parameter being obtained by modeling the directivity data with the mixture model including one or more distributions.

25 (3)

The information processing device according to Item (2), in which

the one or more distributions include at least any one of a vMF distribution or a Kent distribution.

30 (4)

The information processing device according to Item

(2) or (3), in which

the directivity data includes a directivity gain
for each of a plurality of frequency bins, and

the model data includes the model parameter
5 constituting the mixture model representing a
distribution of the directivity gain for each band that
is a frequency band including one or more of the
frequency bins.

(5)

10 The information processing device according to Item
(4), in which

the model data includes a scale factor indicating a
dynamic range of the directivity gain in the frequency
bin and a minimum value of the directivity gain in the
15 frequency bin.

(6)

The information processing device according to any
one of Items (1) to (5), in which

the model data includes difference information
20 indicating a difference between the directivity data
before modeling and the directivity data after modeling,
and

the information processing device further includes
an addition unit configured to add the difference
25 information to the directivity data calculated by the
calculator.

(7)

The information processing device according to Item
(6), in which

30 the difference information is Huffman encoded.

(8)

The information processing device according to any one of Items (1) to (7), in which

the directivity data includes a directivity gain for each of a plurality of frequency bins, and

5 the information processing device further includes an interpolation processing unit configured to calculate the directivity gain of the new frequency bin by performing an interpolation process on the basis of the directivity data calculated by the calculator.

10 (9)

The information processing device according to any one of Items (1) to (8), in which

the directivity data includes a directivity gain at each of a plurality of data points, and

15 the information processing device further includes an interpolation processing unit configured to calculate the directivity gain at the new data point by performing an interpolation process on the basis of the directivity data calculated by the calculator.

20 (10)

The information processing device according to any one of Items (1) to (9), further including

a directivity convolution unit configured to convolve the directivity data and audio data.

25 (11)

The information processing device according to Item (10), further including

an HRTF convolution unit configured to convolve the audio data in which the directivity data is convolved and
30 an HRTF.

(12)

The information processing device according to Item
(2), in which

the one or more distributions include a complex
Bingham distribution or a complex watson distribution.

5 (13)

The information processing device according to Item
(1), in which

the model data includes a spherical harmonic
coefficient obtained by modeling the directivity data by
10 spherical harmonic function expansion as a model
parameter.

(14)

The information processing device according to Item
(1), in which

15 the model data includes a model parameter obtained
by modeling the directivity data by one or more methods
different from each other.

(15)

The information processing device according to Item
20 (14), in which

the methods include at least any one of a method of
modeling with a mixture model including one or more
distributions or a method of modeling by spherical
harmonic function expansion.

25 (16)

The information processing device according to Item
(14) or (15), in which

the model data further includes difference
information indicating a difference between the
30 directivity data after modeling by the one or more
methods and the directivity data before modeling.

(17)

The information processing device according to Item
(16), in which
the difference information is Huffman encoded.

5 (18)

The information processing device according to Item
(17), in which
each of a real part and an imaginary part of the
difference information is individually Huffman encoded.

10 (19)

The information processing device according to Item
(14) or (15), in which
the model data includes difference code data
obtained by Huffman encoding at least any one of a
15 difference between positions or a difference between
frequencies in a space of difference information
indicating a difference between the directivity data
after modeling by the one or more methods and the
directivity data before modeling.

20 (20)

The information processing device according to Item
(19), in which
the model data includes the difference code data
obtained by individually Huffman encoding each of a real
25 part and an imaginary part of a difference of the
difference information.

(21)

The information processing device according to Item
(14) or (15), in which
30 the model data includes the model parameter
obtained by modeling the directivity data by a

predetermined method, and another model parameter
obtained by modeling a difference between the directivity
data after modeling by the predetermined method and the
directivity data before modeling by a method different
5 from the predetermined method.
(22)

The information processing device according to Item
(14) or (15), in which
the model data includes the model parameter
10 obtained by modeling the directivity data by a
predetermined method, and another model parameter
obtained by modeling a ratio between the directivity data
after modeling by the predetermined method and the
directivity data before modeling by a method different
15 from the predetermined method.
(23)

The information processing device according to Item
(14) or (15), in which
the model data includes a model parameter obtained
20 by further modeling the model parameter obtained by
modeling the directivity data.
(24)

The information processing device according to any
one of Items (14) to (23), in which
25 the model data includes the model parameter
obtained by modeling the directivity data by a method
different for each frequency band.
(25)

The information processing device according to any
30 one of Items (1) to (24), in which
the directivity data includes a directivity gain at

each of a plurality of data points, and

the model data includes information indicating a method of disposing the data points and information for identifying an arrangement position of the data points.

5 (26)

The information processing device according to Item (25), in which

the model data includes priority information indicating priority of the directivity data for each type
10 of the sound source.

(27)

The information processing device according to Item (26), in which

the number of data points changes according to the
15 priority, and

the calculator identifies an arrangement position of the data points using the priority information.

(28)

The information processing device according to Item
20 (19), in which

the directivity data includes a directivity gain for each frequency bin at each of a plurality of data points, and

the model data includes the difference code data of
25 at least any one of a difference between the data points or a difference between the frequency bins of the difference information indicating a difference between the directivity gain of the directivity data after modeling by the one or more methods and the directivity
30 gain of the directivity data before modeling after a rearrangement of the difference information.

(29)

The information processing device according to Item (28), in which

the rearrangement is a rearrangement in a
5 predetermined order, an order of priority of the data points or the frequency bins, an ascending order of the difference information, or a descending order of the difference information.

(30)

10 The information processing device according to Item (4), in which

the model data includes a parameter obtained by parameterizing at least any one of a scale factor indicating a dynamic range of the directivity gain in
15 each of the frequency bins or a minimum value of the directivity gain in each of the frequency bins.

(31)

The information processing device according to any one of Items (2) to (5), in which

20 the model data includes operation-related information for a rotation operation or a symmetry operation, and

the calculator calculates the model parameter rotated or symmetrically moved by performing the rotation
25 operation or the target operation on the model parameter on the basis of the operation-related information, and calculates the directivity data using the distribution obtained by the rotated or symmetrically moved model parameter.

30 (32)

The information processing device according to Item

(4) or (5), in which

the calculator calculates the directivity gain of the predetermined frequency bin by performing weighted addition on an output value of the mixture model of a
5 predetermined band and an output value of the mixture model of another band adjacent to the predetermined band.
(33)

The information processing device according to any one of Items (2) to (5), in which

10 the calculator calculates the directivity data by performing weighted addition on a plurality of the distributions obtained from the model parameter by using a weight including a negative value.
(34)

15 An information processing method including
by an information processing device
acquiring model data obtained by modeling
directivity data representing directivity of a sound
source, and
20 calculating the directivity data on the basis of
the model data.
(35)

A program for causing a computer to execute the steps of
25 acquiring model data obtained by modeling
directivity data representing directivity of a sound
source, and
calculating the directivity data on the basis of
the model data.
30 (36)

An information processing device including

a modeling unit configured to model directivity data representing directivity of a sound source with a mixture model including one or more distributions, and
a model data generation unit configured to generate
5 model data including a model parameter constituting the mixture model, the model parameter being obtained by the modeling.

(37)

An information processing method including
10 by an information processing device
modeling directivity data representing directivity of a sound source with a mixture model including one or more distributions, and
generating model data including model parameter
15 constituting the mixture model, the model parameter being obtained by the modeling.

(38)

A program for causing a computer to execute the steps of
20 modeling directivity data representing directivity of a sound source with a mixture model including one or more distributions, and
generating model data including model parameter constituting the mixture model, the model parameter being
25 obtained by the modeling.

(39)

An information processing device including
an acquisition unit configured to acquire
difference directivity data obtained by obtaining at
30 least any one of a difference between data points or a difference between frequency bins of a directivity gain

for directivity data representing directivity of a sound source, the directivity data including the directivity gain of each of a plurality of the frequency bins at a plurality of the data points, and

5 a calculator configured to calculate the directivity data on the basis of the difference directivity data.

(40)

 The information processing device according to Item
10 (39), in which
 the difference directivity data is Huffman encoded,
and

 the calculator decodes the difference directivity data that is Huffman encoded.

15 (41)

 The information processing device according to Item (40), in which

 each of a real part and an imaginary part of the difference directivity data is individually Huffman
20 encoded.

(42)

 The information processing device according to any one of Items (39) to (41), in which

 the difference directivity data is obtained by
25 obtaining at least any one of the difference between the data points or the difference between the frequency bins after the directivity gains are rearranged.

(43)

 The information processing device according to Item
30 (42), in which

 the rearrangement is a rearrangement in a

predetermined order, an order of priority of the data points or the frequency bins, an ascending order of the directivity gains, or a descending order of the directivity gains.

5 (44)

An information processing method including
by an information processing device
acquiring difference directivity data obtained by
obtaining at least any one of a difference between data
10 points or a difference between frequency bins of a
directivity gain for directivity data representing
directivity of a sound source, the directivity data
including the directivity gain of each of a plurality of
the frequency bins at a plurality of the data points, and
15 calculating the directivity data on the basis of
the difference directivity data.

(45)

A program for causing a computer to execute the
steps of
20 acquiring difference directivity data obtained by
obtaining at least any one of a difference between data
points or a difference between frequency bins of a
directivity gain for directivity data representing
directivity of a sound source, the directivity data
25 including the directivity gain of each of a plurality of
the frequency bins at a plurality of the data points, and
calculating the directivity data on the basis of
the difference directivity data.

30 REFERENCE SIGNS LIST

[0750]

	11	Server
	21	Modeling unit
	22	Model data generation unit
	23	Audio data encoding unit
5	51	Information processing device
	61	Acquisition unit
	62	Distribution model decoding unit
	63	Audio data decoding unit
	64	Rendering processing unit
10	82	Directivity data calculator
	83	Difference information decoding unit
	84	Addition unit
	85	Frequency interpolation processing unit
	88	Temporal interpolation processing unit
15	89	Directivity convolution unit
	90	HRTF convolution unit

CLAIMS

1. An information processing device comprising:
an acquisition unit configured to acquire model
5 data obtained by modeling directivity data representing
directivity of a sound source; and
a calculator configured to calculate the
directivity data on a basis of the model data.
- 10 2. The information processing device according to
claim 1, wherein
the model data includes a model parameter
constituting a mixture model, the model parameter being
obtained by modeling the directivity data with the
15 mixture model including one or more distributions.
3. The information processing device according to
claim 2, wherein
the one or more distributions include at least any
20 one of a vMF distribution or a Kent distribution.
4. The information processing device according to
claim 2, wherein
the directivity data includes a directivity gain
25 for each of a plurality of frequency bins, and
the model data includes the model parameter
constituting the mixture model representing a
distribution of the directivity gain for each band that
is a frequency band including one or more of the
30 frequency bins.

5. The information processing device according to claim 4, wherein
the model data includes a scale factor indicating a dynamic range of the directivity gain in the frequency
5 bin and a minimum value of the directivity gain in the frequency bin.
6. The information processing device according to claim 1, wherein
10 the model data includes difference information indicating a difference between the directivity data before modeling and the directivity data after modeling, and
the information processing device further comprises
15 an addition unit configured to add the difference information to the directivity data calculated by the calculator.
7. The information processing device according to claim 6, wherein
20 the difference information is Huffman encoded.
8. The information processing device according to claim 1, wherein
25 the directivity data includes a directivity gain for each of a plurality of frequency bins, and
the information processing device further comprises an interpolation processing unit configured to calculate the directivity gain of the new frequency bin by
30 performing an interpolation process on a basis of the directivity data calculated by the calculator.

9. The information processing device according to claim 1, wherein
the directivity data includes a directivity gain at
5 each of a plurality of data points, and
the information processing device further comprises
an interpolation processing unit configured to calculate
the directivity gain at the new data point by performing
an interpolation process on a basis of the directivity
10 data calculated by the calculator.
10. The information processing device according to claim 1, further comprising:
a directivity convolution unit configured to
15 convolve the directivity data and audio data.
11. The information processing device according to claim 10, further comprising:
an HRTF convolution unit configured to convolve the
20 audio data in which the directivity data is convolved and
an HRTF.
12. The information processing device according to claim 2, wherein
25 the one or more distributions include a complex
Bingham distribution or a complex watson distribution.
13. The information processing device according to claim 1, wherein
30 the model data includes a spherical harmonic
coefficient obtained by modeling the directivity data by

spherical harmonic function expansion as a model parameter.

14. The information processing device according to
5 claim 1, wherein
the model data includes a model parameter obtained by modeling the directivity data by one or more methods different from each other.

10 15. The information processing device according to claim 14, wherein
the methods include at least any one of a method of modeling with a mixture model including one or more distributions or a method of modeling by spherical
15 harmonic function expansion.

16. The information processing device according to claim 14, wherein
the model data further includes difference
20 information indicating a difference between the directivity data after modeling by the one or more methods and the directivity data before modeling.

17. The information processing device according to
25 claim 16, wherein
the difference information is Huffman encoded.

18. The information processing device according to claim 17, wherein
30 each of a real part and an imaginary part of the difference information is individually Huffman encoded.

19. The information processing device according to claim 14, wherein
the model data includes difference code data
5 obtained by Huffman encoding at least any one of a difference between positions or a difference between frequencies in a space of difference information indicating a difference between the directivity data after modeling by the one or more methods and the
10 directivity data before modeling.
20. The information processing device according to claim 19, wherein
the model data includes the difference code data
15 obtained by individually Huffman encoding each of a real part and an imaginary part of a difference of the difference information.
21. The information processing device according to
20 claim 14, wherein
the model data includes the model parameter obtained by modeling the directivity data by a predetermined method, and another model parameter obtained by modeling a difference between the directivity
25 data after modeling by the predetermined method and the directivity data before modeling by a method different from the predetermined method.
22. The information processing device according to
30 claim 14, wherein
the model data includes the model parameter

obtained by modeling the directivity data by a predetermined method, and another model parameter obtained by modeling a ratio between the directivity data after modeling by the predetermined method and the
5 directivity data before modeling by a method different from the predetermined method.

23. The information processing device according to claim 14, wherein
10 the model data includes a model parameter obtained by further modeling the model parameter obtained by modeling the directivity data.

24. The information processing device according to
15 claim 14, wherein the model data includes the model parameter obtained by modeling the directivity data by a method different for each frequency band.

20 25. The information processing device according to claim 1, wherein the directivity data includes a directivity gain at each of a plurality of data points, and the model data includes information indicating a
25 method of disposing the data points and information for identifying an arrangement position of the data points.

26. The information processing device according to claim 25, wherein
30 the model data includes priority information indicating priority of the directivity data for each type

of the sound source.

27. The information processing device according to claim 26, wherein

5 the number of data points changes according to the priority, and

the calculator identifies an arrangement position of the data points using the priority information.

10 28. The information processing device according to claim 19, wherein

the directivity data includes a directivity gain for each frequency bin at each of a plurality of data points, and

15 the model data includes the difference code data of at least any one of a difference between the data points or a difference between the frequency bins of the difference information indicating a difference between the directivity gain of the directivity data after
20 modeling by the one or more methods and the directivity gain of the directivity data before modeling after a rearrangement of the difference information.

29. The information processing device according to claim 28, wherein

25 the rearrangement is a rearrangement in a predetermined order, an order of priority of the data points or the frequency bins, an ascending order of the difference information, or a descending order of the
30 difference information.

30. The information processing device according to claim 4, wherein

the model data includes a parameter obtained by parameterizing at least any one of a scale factor
5 indicating a dynamic range of the directivity gain in each of the frequency bins or a minimum value of the directivity gain in each of the frequency bins.

31. The information processing device according to
10 claim 2, wherein

the model data includes operation-related information for a rotation operation or a symmetry operation, and

the calculator calculates the model parameter
15 rotated or symmetrically moved by performing the rotation operation or the target operation on the model parameter on a basis of the operation-related information, and calculates the directivity data using the distribution obtained by the rotated or symmetrically moved model
20 parameter.

32. The information processing device according to claim 4, wherein

the calculator calculates the directivity gain of
25 the predetermined frequency bin by performing weighted addition on an output value of the mixture model of a predetermined band and an output value of the mixture model of another band adjacent to the predetermined band.

30 33. The information processing device according to claim 2, wherein

the calculator calculates the directivity data by performing weighted addition on a plurality of the distributions obtained from the model parameter by using a weight including a negative value.

5

34. An information processing method comprising:

by an information processing device

acquiring model data obtained by modeling

directivity data representing directivity of a sound

10 source; and

calculating the directivity data on a basis of the model data.

35. A program for causing a computer to execute the

15 steps of:

acquiring model data obtained by modeling

directivity data representing directivity of a sound

source; and

calculating the directivity data on a basis of the

20 model data.

36. An information processing device comprising:

a modeling unit configured to model directivity

data representing directivity of a sound source with a

25 mixture model including one or more distributions; and

a model data generation unit configured to generate

model data including a model parameter constituting the mixture model, the model parameter being obtained by the modeling.

30

37. An information processing method comprising:

by an information processing device
modeling directivity data representing directivity
of a sound source with a mixture model including one or
more distributions; and

5 generating model data including model parameter
constituting the mixture model, the model parameter being
obtained by the modeling.

38. A program for causing a computer to execute the
10 steps of:

 modeling directivity data representing directivity
of a sound source with a mixture model including one or
more distributions; and

 generating model data including model parameter
15 constituting the mixture model, the model parameter being
obtained by the modeling.

39. An information processing device comprising:
 an acquisition unit configured to acquire
20 difference directivity data obtained by obtaining at
least any one of a difference between data points or a
difference between frequency bins of a directivity gain
for directivity data representing directivity of a sound
source, the directivity data including the directivity
25 gain of each of a plurality of the frequency bins at a
plurality of the data points; and
 a calculator configured to calculate the
directivity data on a basis of the difference directivity
data.

30

40. The information processing device according to

claim 39, wherein
the difference directivity data is Huffman encoded,
and
the calculator decodes the difference directivity
5 data that is Huffman encoded.

41. The information processing device according to
claim 40, wherein
each of a real part and an imaginary part of the
10 difference directivity data is individually Huffman
encoded.

42. The information processing device according to
claim 39, wherein
15 the difference directivity data is obtained by
obtaining at least any one of the difference between the
data points or the difference between the frequency bins
after the directivity gains are rearranged.

20 43. The information processing device according to
claim 42, wherein
the rearrangement is a rearrangement in a
predetermined order, an order of priority of the data
points or the frequency bins, an ascending order of the
25 directivity gains, or a descending order of the
directivity gains.

44. An information processing method comprising:
by an information processing device
30 acquiring difference directivity data obtained by
obtaining at least any one of a difference between data

points or a difference between frequency bins of a
directivity gain for directivity data representing
directivity of a sound source, the directivity data
including the directivity gain of each of a plurality of
5 the frequency bins at a plurality of the data points; and
calculating the directivity data on a basis of the
difference directivity data.

45. A program for causing a computer to execute the
10 steps of:
acquiring difference directivity data obtained by
obtaining at least any one of a difference between data
points or a difference between frequency bins of a
directivity gain for directivity data representing
15 directivity of a sound source, the directivity data
including the directivity gain of each of a plurality of
the frequency bins at a plurality of the data points; and
calculating the directivity data on a basis of the
difference directivity data.