DESCRIPTION

INFORMATION PROCESSING APPARATUS AND INFORMATION PROCESSING METHOD, AND PROGRAM

TECHNICAL FIELD

[0001]

The present technology relates to an information processing apparatus, an information processing method, and a program, and more particularly, to an information processing apparatus and an information processing method, and a program capable of realizing content reproduction based on an intention of a content creator.

BACKGROUND ART

[0002]

For example, in a free viewpoint space, each object arranged in the space using the absolute coordinate system is fixedly arranged (see, for example, Patent Document 1).

[0003]

In this case, the direction of each object viewed from an arbitrary listening position is uniquely obtained on the basis of the coordinate position of the listener in the absolute space, the face direction, and the relationship to the object, and the gain of each object is uniquely obtained on the basis of the distance from the listening position, and the sound of each object is reproduced.

CITATION LIST

PATENT DOCUMENT

[0004]

Patent Document 1: WO 2019/198540 A

SUMMARY OF THE INVENTION

PROBLEMS TO BE SOLVED BY THE INVENTION

[0005]

On the other hand, there are points to be emphasized as content for the artistry and the listener.

[0006]

For example, there is a case where it is desirable that an object be located forward such as, regarding music content, a musical instrument or a player at a certain listening point where the content is desired to be emphasized in terms of its substance, or regarding sports content, a player who is desired to be emphasized.

[0007]

In view of the above, there is a possibility that the mere physical relationship between the listener and the object as described above does not sufficiently convey the amusement of the content.

[0008]

The present technology has been made in view of such a situation and realizes content reproduction based on an intention of a content creator while following a free position of a listener.

SOLUTIONS TO PROBLEMS

[0009]

An information processing apparatus according to an aspect of the present technology includes: a listener position information acquisition unit that acquires

listener position information of a viewpoint of a listener; a reference viewpoint information acquisition unit that acquires position information of a first reference viewpoint and object position information of an object at the first reference viewpoint, and position information of a second reference viewpoint and object position information of the object at the second reference viewpoint; and an object position calculation unit that calculates position information of the object at the viewpoint of the listener on the basis of the listener position information, the position information of the first reference viewpoint and the object position information at the first reference viewpoint, and the position information of the second reference viewpoint and the object position information at the second reference viewpoint.

[0010]

An information processing method or program according to an aspect of the present technology includes the steps of: acquiring listener position information of a viewpoint of a listener; acquiring position information of a first reference viewpoint and object position information of an object at the first reference viewpoint, and position information of a second reference viewpoint and object position information of the object at the second reference viewpoint; and calculating position information of the object at the viewpoint of the listener on the basis of the listener position information, the position information of the first reference viewpoint and the object position information at the first reference viewpoint, and the position

information of the second reference viewpoint and the object position information at the second reference viewpoint.

[0011]

According to an aspect of the present technology, listener position information of a viewpoint of a listener is acquired; position information of a first reference viewpoint and object position information of an object at the first reference viewpoint, and position information of a second reference viewpoint and object position information of the object at the second reference viewpoint are acquired; and position information of the object at the viewpoint of the listener is calculated on the basis of the listener position information, the position information of the first reference viewpoint and the object position information at the first reference viewpoint, and the position information of the second reference viewpoint and the object position information at the second reference viewpoint.

BRIEF DESCRIPTION OF DRAWINGS

[0012]

Fig. 1 is a diagram illustrating a configuration of a content reproduction system.

Fig. 2 is a diagram illustrating a configuration of a content reproduction system.

Fig. 3 is a diagram describing a reference viewpoint.

Fig. 4 is a diagram illustrating an example of system configuration information.

Fig. 5 is a diagram illustrating an example of system configuration information.

Fig. 6 is a diagram describing coordinate transformation.

Fig. 7 is a diagram describing coordinate axis transformation processing.

Fig. 8 is a diagram illustrating an example of a transformation result by the coordinate axis transformation processing.

Fig. 9 is a diagram describing interpolation processing.

Fig. 10 is a diagram illustrating a sequence example of a content reproduction system.

Fig. 11 is a diagram describing an example of bringing an object closer to arrangement at a reference viewpoint.

Fig. 12 is a diagram describing interpolation of object absolute coordinate position information.

Fig. 13 is a diagram describing an internal division ratio in a viewpoint-side triangle mesh.

Fig. 14 is a diagram describing calculation of object position based on internal division ratio.

Fig. 15 is a diagram describing calculation of gain information based on internal division ratio.

Fig. 16 is a diagram describing selection of a triangle mesh.

Fig. 17 is a diagram illustrating a configuration of a content reproduction system.

Fig. 18 is a flowchart describing provision processing and reproduction audio data generation processing.

Fig. 19 is a flowchart describing viewpoint selection processing.

Fig. 20 is a diagram illustrating a configuration example of a computer.

MODE FOR CARRYING OUT THE INVENTION

[0013]

An embodiment to which the present technology has been applied is described below with reference to the drawings.

[0014]

<First embodiment>

<Configuration example of the content reproduction system>

The present technology has Features F1 to F6 described below.

[0015]

(Feature F1)

The feature that object arrangement and gain information at a plurality of reference viewpoints in a free viewpoint space are prepared in advance.

(Feature F2)

The feature that an object position and gain information at an arbitrary listening point are obtained on the basis of object arrangement and gain information at a plurality of reference viewpoints sandwiching or surrounding the arbitrary listening point (listening position).

(Feature F3)

The feature that, in a case where an object position and the gain amount of an arbitrary listening

point are obtained, a proportion ratio is obtained according to a plurality of reference viewpoints sandwiching or surrounding the arbitrary listening point and the arbitrary listening point, and the object position with respect to the arbitrary listening point is obtained using the proportion ratio.

(Feature F4)

The feature that object arrangement information at a plurality of reference viewpoints prepared in advance uses a polar coordinate system and is transmitted.

(Feature F5)

The feature that object arrangement information at a plurality of reference viewpoints prepared in advance uses an absolute coordinate system and is transmitted.

(Feature F6)

The feature that, in a case where an object position at an arbitrary listening point is calculated, a listener can listen with the object arrangement brought closer to any reference viewpoint by using a specific bias coefficient.

[0016]

First, a content reproduction system to which the present technology has been applied will be described.

[0017]

The content reproduction system includes a server and a client that code, transmit, and decode each piece of data.

[0018]

For example, the listener position information is transmitted from the client side to the server as necessary, and some object position information is

transmitted from the server side to the client side on the basis of the result. Then, rendering processing is performed on each object on the basis of some object position information received on the client side, and content including a sound of each object is reproduced.

[0019]

Such content reproduction system is configured as illustrated, for example, in Fig. 1.

[0020]

That is, the content reproduction system illustrated in Fig. 1 includes a server 11 and a client 12.

[0021]

The server 11 includes a configuration information sending unit 21 and a coded data sending unit 22.

[0022]

The configuration information sending unit 21 sends (transmits) system configuration information prepared in advance to the client 12, and receives viewpoint selection information or the like transmitted from the client 12 and supplies the information to the coded data sending unit 22.

[0023]

In the content reproduction system, a plurality of listening positions on a predetermined common absolute coordinate space is designated (set) in advance by a content creator as the positions of reference viewpoints (hereinafter, also referred to as the reference viewpoint positions).

[0024]

Here, the content creator designates (sets) in

advance, as the reference viewpoint, the position on the common absolute coordinate space that the content creator wants the listener to take as the listening position at the time of content reproduction, and the direction of the face that the content creator wants the listener to face at the position, that is, a viewpoint at which the content creator wants the listener to listen to the sound of the content.

[0025]

In the server 11, system configuration information that is information regarding each reference viewpoint and object polar coordinate coded data for each reference viewpoint are prepared in advance.

[0026]

Here, the object polar coordinate coded data for each reference viewpoint is obtained by coding object polar coordinate position information indicating the relative position of the object viewed from the reference viewpoint. In the object polar coordinate position information, the position of the object viewed from the reference viewpoint is expressed by polar coordinates. Note that even for the same object, the absolute arrangement position of the object in the common absolute coordinate space varies with each reference viewpoint.

[0027]

The configuration information sending unit 21 sends the system configuration information to the client 12 via a network or the like immediately after the operation of the content reproduction system is started, that is, for example, immediately after connection with the client 12 is established.

[0028]

The coded data sending unit 22 selects two reference viewpoints from among the plurality of reference viewpoints on the basis of the viewpoint selection information supplied from the configuration information sending unit 21, and sends the object polar coordinate coded data of each of the selected two reference viewpoints to the client 12 via a network or the like.

[0029]

Here, the viewpoint selection information is, for example, information indicating two reference viewpoints selected on the client 12 side.

[0030]

Therefore, in the coded data sending unit 22, the object polar coordinate coded data of the reference viewpoint requested by the client 12 is acquired and sent to the client 12. Note that the number of reference viewpoints selected by the viewpoint selection information is not limited to two, but may be three or more.

[0031]

Furthermore, the client 12 includes a listener position information acquisition unit 41, a viewpoint selection unit 42, a configuration information acquisition unit 43, a coded data acquisition unit 44, a decode unit 45, a coordinate transformation unit 46, a coordinate axis transformation processing unit 47, an object position calculation unit 48, and a polar coordinate transformation unit 49.

[0032]

The listener position information acquisition unit 41 acquires the listener position information indicating the absolute position (listening position) of the listener on the common absolute coordinate space according to the designation operation of the user (listener) or the like, and supplies the listener position information to the viewpoint selection unit 42, the object position calculation unit 48, and the polar coordinate transformation unit 49.

[0033]

For example, in the listener position information, the position of the listener in the common absolute coordinate space is expressed by absolute coordinates. Note that, hereinafter, the coordinate system of the absolute coordinates indicated by the listener position information is also referred to as a common absolute coordinate system.

[0034]

The viewpoint selection unit 42 selects two reference viewpoints on the basis of the system configuration information supplied from the configuration information acquisition unit 43 and the listener position information supplied from the listener position information acquisition unit 41, and supplies viewpoint selection information indicating the selection result to the configuration information acquisition unit 43.

[0035]

For example, the viewpoint selection unit 42 specifies a section from the position of the listener (listening position) and the assumed absolute coordinate position of each reference viewpoint, and selects two

reference viewpoints on the basis of the result of specifying the section.

[0036]

The configuration information acquisition unit 43 receives the system configuration information transmitted from the server 11 and supplies the system configuration information to the viewpoint selection unit 42 and the coordinate axis transformation processing unit 47, and transmits the viewpoint selection information supplied from the viewpoint selection unit 42 to the server 11 via a network or the like.

[0037]

Note that, here, an example in which the viewpoint selection unit 42 that selects a reference viewpoint on the basis of the listener position information and the system configuration information is provided in the client 12 will be described, but the viewpoint selection unit 42 may be provided on the server 11 side.

[0038]

The coded data acquisition unit 44 receives the object polar coordinate coded data transmitted from the server 11 and supplies the object polar coordinate coded data to the decode unit 45. That is, the coded data acquisition unit 44 acquires the object polar coordinate coded data from the server 11.

[0039]

The decode unit 45 decodes the object polar coordinate coded data supplied from the coded data acquisition unit 44, and supplies the resultant object polar coordinate position information to the coordinate transformation unit 46.

[0040]

The coordinate transformation unit 46 performs coordinate transformation on the object polar coordinate position information supplied from the decode unit 45, and supplies the resultant object absolute coordinate position information to the coordinate axis transformation processing unit 47.

[0041]

The coordinate transformation unit 46 performs coordinate transformation that transforms polar coordinates into absolute coordinates. Therefore, the object polar coordinate position information that is polar coordinates indicating the position of the object viewed from the reference viewpoint is transformed into object absolute coordinate position information that is absolute coordinates indicating the position of the object in the absolute coordinate system having the position of the reference viewpoint as the origin.

[0042]

The coordinate axis transformation processing unit 47 performs coordinate axis transformation processing on the object absolute coordinate position information supplied from the coordinate transformation unit 46 on the basis of the system configuration information supplied from the configuration information acquisition unit 43.

[0043]

Here, the coordinate axis transformation processing is processing performed by combining coordinate transformation (coordinate axis transformation) and offset shift, and the object absolute coordinate position

information indicating absolute coordinates of the object projected on the common absolute coordinate space is obtained by the coordinate axis transformation processing. That is, the object absolute coordinate position information obtained by the coordinate axis transformation processing is absolute coordinates of the common absolute coordinate system indicating the absolute position of the object on the common absolute coordinate space.

[0044]

The object position calculation unit 48 performs interpolation processing on the basis of the listener position information supplied from the listener position information acquisition unit 41 and the object absolute coordinate position information supplied from the coordinate axis transformation processing unit 47, and supplies the resultant final object absolute coordinate position information to the polar coordinate transformation unit 49. The final object absolute coordinate position information mentioned here is information indicating the position of the object in the common absolute coordinate system in a case where the viewpoint of the listener is at the listening position indicated by the listener position information.

[0045]

The object position calculation unit 48 calculates the absolute position of the object in the common absolute coordinate space corresponding to the listening position, that is, the absolute coordinates of the common absolute coordinate system, from the listening position indicated by the listener position information and the

positions of the two reference viewpoints indicated by the viewpoint selection information, and determines the absolute position as the final object absolute coordinate position information. At this time, the object position calculation unit 48 acquires the system configuration information from the configuration information acquisition unit 43 and acquires the viewpoint selection information from the viewpoint selection unit 42 as necessary.

[0046]

The polar coordinate transformation unit 49 performs polar coordinate transformation on the object absolute coordinate position information supplied from the object position calculation unit 48 on the basis of the listener position information supplied from the listener position information acquisition unit 41, and outputs the resultant polar coordinate position information to a subsequent rendering processing unit, which is not illustrated.

[0047]

The polar coordinate transformation unit 49 performs polar coordinate transformation of transforming the object absolute coordinate position information, which is absolute coordinates of the common absolute coordinate system, into polar coordinate position information, which is polar coordinates indicating a relative position of the object viewed from the listening position.

[0048]

Note that, although the example in which the object polar coordinate coded data is prepared in advance for

each reference viewpoint in the server 11 has been described above, the object absolute coordinate position information to be the output of the coordinate axis transformation processing unit 47 may be prepared in advance in the server 11.

[0049]

In such a case, the content reproduction system is configured as illustrated, for example, in Fig. 2. Note that portions in Fig. 2 corresponding to those of Fig. 1 are designated by the same reference numerals, and description is omitted as appropriate.

[0050]

The content reproduction system illustrated in Fig. 2 includes a server 11 and a client 12.

[0051]

Furthermore, the server 11 includes a configuration information sending unit 21 and a coded data sending unit 22, but in this example, the coded data sending unit 22 acquires object absolute coordinate coded data of two reference viewpoints indicated by viewpoint selection information, and sends the object absolute coordinate coded data to the client 12.

[0052]

That is, in the server 11, the object absolute coordinate coded data obtained by coding the object absolute coordinate position information to be the output of the coordinate axis transformation processing unit 47 illustrated in Fig. 1 is prepared in advance for each of the plurality of reference viewpoints.

[0053]

Therefore, in this example, the client 12 is not

provided with the coordinate transformation unit 46 or the coordinate axis transformation processing unit 47 illustrated in Fig. 1.

[0054]

That is, the client 12 illustrated in Fig. 2 includes a listener position information acquisition unit 41, a viewpoint selection unit 42, a configuration information acquisition unit 43, a coded data acquisition unit 44, a decode unit 45, an object position calculation unit 48, and a polar coordinate transformation unit 49.

[0055]

The configuration of the client 12 illustrated in Fig. 2 is different from the configuration of the client 12 illustrated in Fig. 1 on the point that the coordinate transformation unit 46 and the coordinate axis transformation processing unit 47 are not provided, and is the same as the configuration of the client 12 illustrated in Fig. 1 on the other points.

[0056]

The coded data acquisition unit 44 receives the object absolute coordinate coded data transmitted from the server 11 and supplies the object absolute coordinate coded data to the decode unit 45.

[0057]

The decode unit 45 decodes the object absolute coordinate coded data supplied from the coded data acquisition unit 44, and supplies the resultant object absolute coordinate position information to the object position calculation unit 48.

[0058]

<Regarding the present technology>

Next, the present technology will be further described.

[0059]

First, a process of creating content provided from the server 11 to the client 12 will be described.

[0060]

First, an example in which a transmission method using a polar coordinate system is used, that is, an example in which the object polar coordinate coded data is transmitted as illustrated in Fig. 1 will be described.

[0061]

Content creation using the polar coordinate system is performed for 3D audio based on a fixed viewpoint, and there is an advantage that such a creation method can be used as it is.

[0062]

A plurality of reference viewpoints at which the content creator (hereinafter, also simply referred to as a creator) wants the listener to listen to is set in the three-dimensional space according to the intention of the creator.

[0063]

Specifically, for example, as illustrated in Fig. 3, four reference viewpoints are set in a common absolute coordinate space which is a three-dimensional space. Here, four positions P11 to P14 designated by the creator are the reference viewpoints, in more detail, the positions of the reference viewpoints.

[0064]

The reference viewpoint information, which is

information regarding each reference viewpoint, includes reference viewpoint position information, which is absolute coordinates of a common absolute coordinate system indicating a standing position in the common absolute coordinate space, that is, the position of the reference viewpoint, and listener direction information indicating the direction of the face of the listener.

[0065]

Here, the listener direction information includes, for example, a rotation angle (horizontal angle) in the horizontal direction of the face of the listener at the reference viewpoint and a vertical angle indicating the direction of the face of the listener in the vertical direction.

[0066]

In Fig. 3, the arrows drawn adjacent to the respective positions P11 to P14 indicate the listener direction information at the reference viewpoint indicated by the respective positions P11 to P14, that is, the direction of the face of the listener.

[0067]

Furthermore, in Fig. 3, a region R11 indicates an example of a region where an object exists, and it can be seen that, in this example, at each reference viewpoint, the direction of the face of the listener indicated by the listener direction information is the direction of the region R11. For example, at the position P14, the direction of the face of the listener indicated by the listener direction information is backward.

[0068]

Next, the object polar coordinate position

information expressing the position of each object at each of the plurality of set reference viewpoints in a polar coordinate format and the gain amount for each object at each of the reference viewpoints are set by the creator. For example, the object polar coordinate position information includes a horizontal angle and a vertical angle of the object viewed from the reference viewpoint, and a radius indicating a distance from the reference viewpoint to the object.

[0069]

When the position and the like of the object are set for each of the plurality of reference viewpoints in this manner, Information IFP1 to Information IFP5 described below are obtained as the information regarding the reference viewpoint.

[0070]

(Information IFP1)

The number of objects

(Information IFP2)

The number of reference viewpoints

(Information IFP3)

Direction of the face of a listener at a reference viewpoint (horizontal angle and vertical angle)

(Information IFP4)

Absolute coordinate position of a reference viewpoint in an absolute space (common absolute coordinate space)

(Information IFP5)

Polar coordinate position (horizontal angle, vertical angle, and radius) and gain amount of each object viewed from Information IFP3 and Information IFP4

[0071]

Here, Information IFP3 is the above-described listener direction information and Information IFP4 is the above-described reference viewpoint position information.

[0072]

Furthermore, the polar coordinate position, which is Information IFP5, includes a horizontal angle, a vertical angle, and a radius, and is the object polar coordinate position information indicating a relative position of the object based on the reference viewpoint. Since the object polar coordinate position information is equivalent to the polar coordinate coded information of Moving Picture Experts Group (MPEG)-H, the coding system of MPEG-H can be utilized.

[0073]

Information including each piece of information from Information IFP1 to Information IFP4 among Information IFP1 to Information IFP5 is the above-described system configuration information.

[0074]

This system configuration information is transmitted to the client 12 side prior to transmission of data related to an object, that is, object polar coordinate coded data or coded audio data obtained by coding audio data of an object.

[0075]

A specific example of the system configuration information is as illustrated, for example, in Fig. 4.

[0076]

In the example illustrated in Fig. 4, “NumOfObjs”

indicates the number of objects, which is the number of objects constituting the content, that is, Information IFP1 described above, and "NumfOfRefViewPoint" indicates the number of reference viewpoints, that is, Information IFP2 described above.

[0077]

Furthermore, the system configuration information illustrated in Fig. 4 includes the reference viewpoint information corresponding to the number of reference viewpoints "NumfOfRefViewPoint".

[0078]

That is, "RefViewX[i]", "RefViewY[i]", and "RefViewZ[i]" respectively indicate the X coordinate, the Y coordinate, and the Z coordinate of the common absolute coordinate system indicating the position of the reference viewpoint constituting the reference viewpoint position information of the i-th reference viewpoint as Information IFP4.

[0079]

Furthermore, "ListenerYaw[i]" and "ListenerPitch[i]" are a horizontal angle (yaw angle) and a vertical angle (pitch angle) constituting the listener direction information of the i-th reference viewpoint as Information IFP3.

[0080]

Moreover, in this example, the system configuration information includes information "ObjectOverLapMode[i]" indicating a reproduction mode in a case where the positions of the listener and the object overlap with each other for each object, that is, the listener (listening position) and the object are at the same

position.

[0081]

Next, an example in which a transmission method using an absolute coordinate system is used, that is, an example in which object absolute coordinate coded data is transmitted as illustrated in Fig. 2 will be described.

[0082]

Also in the case of transmitting the object absolute coordinate coded data, similarly to the case of transmitting the object polar coordinate coded data, the object position with respect to each reference viewpoint is recorded as absolute coordinate position information. That is, the object absolute coordinate position information of each object is prepared by the creator for each reference viewpoint.

[0083]

However, in this example, unlike the example of the transmission method using the polar coordinate system, it is not necessary to transmit the listener direction information indicating the direction of the face of the listener.

[0084]

In the example using the transmission method using the absolute coordinate system, Information IFA1 to Information IFA4 described below are obtained as the information regarding the reference viewpoint.

[0085]

(Information IFA1)

The number of objects

(Information IFA2)

The number of reference viewpoints

(Information IFA3)

Absolute coordinate position of a reference viewpoint in an absolute space

(Information IFA4)

Absolute coordinate position and gain amount of each object when the listener is present at the absolute coordinate position indicated in Information IFA3

[0086]

Here, Information IFA1 and Information IFA2 are the same information as Information IFP1 and Information IFP2 described above, and Information IFA3 is the above-described reference viewpoint position information.

[0087]

Furthermore, the absolute coordinate position of the object indicated by Information IFA4 is the object absolute coordinate position information indicating the absolute position of the object on the common absolute coordinate space indicated by the absolute coordinates of the common absolute coordinate system

[0088]

Note that, in the transmission of the object absolute coordinate coded data from the server 11 to the client 12, the object absolute coordinate position information indicating the position of the object with accuracy corresponding to the positional relationship between the listener and the object, for example, the distance from the listener to the object, may be generated and transmitted. In this case, the information amount (bit depth) of the object absolute coordinate position information can be reduced without causing a feeling of deviation of the sound image position.

[0089]

For example, as the distance from the listener to the object is shorter, the object absolute coordinate position information (object absolute coordinate coded data) with higher accuracy, that is, the object absolute coordinate position information indicating a more accurate position is generated.

[0090]

This is because, although the position of the object is deviated depending on the quantization accuracy (quantization step width) at the time of coding, as the distance from the listener to the object is longer, the magnitude (tolerance) of the position deviation that does not cause a feeling of deviation of the localization position of the sound image is larger.

[0091]

Specifically, for example, the object absolute coordinate coded data obtained by coding the object absolute coordinate position information with the highest accuracy is prepared in advance and held in the server 11.

[0092]

Then, by extracting a part of the object absolute coordinate coded data with the highest accuracy, it is possible to obtain the object absolute coordinate coded data obtained by quantizing the object absolute coordinate position information with arbitrary quantization accuracy.

[0093]

Therefore, the coded data sending unit 22 extracts a part or all of the object absolute coordinate coded

data with the highest accuracy according to the distance from the listening position to the object, and transmits the resultant object absolute coordinate coded data with predetermined accuracy to the client 12. Note that, in such a case, it is sufficient if the coded data sending unit 22 acquires the listener position information from the listener position information acquisition unit 41 via the configuration information sending unit 21, the configuration information acquisition unit 43, and the viewpoint selection unit 42.

[0094]

Furthermore, in the content reproduction system illustrated in Fig. 2, system configuration information including each piece of information from Information IFA1 to Information IFA3 among Information IFA1 to Information IFA4 is prepared in advance.

[0095]

This system configuration information is transmitted to the client 12 side prior to transmission of data related to an object, that is, object absolute coordinate coded data or coded audio data.

[0096]

A specific example of such system configuration information is as illustrated, for example, in Fig. 5.

[0097]

In the example illustrated in Fig. 5, similarly to the example illustrated in Fig. 4, the system configuration information includes the number of objects “NumOfObjs” and the number of reference viewpoints “NumfOfRefViewPoint”.

[0098]

Furthermore, the system configuration information includes the reference viewpoint information corresponding to the number of reference viewpoints "NumfOfRefViewPoint".

[0099]

That is, the system configuration information includes the X coordinate "RefViewX[i]", the Y coordinate "RefViewY[i]", and the Z coordinate "RefViewZ[i]" of the common absolute coordinate system indicating the position of the reference viewpoint constituting the reference viewpoint position information of the i-th reference viewpoint. As described above, in this example, the reference viewpoint information does not include the listener direction information, but includes only the reference viewpoint position information.

[0100]

Moreover, the system configuration information includes reproduction mode "ObjectOverLapMode[i]" in a case where the positions of the listener and the object overlap with each other for each object.

[0101]

The system configuration information obtained as described above, the object polar coordinate coded data or the object absolute coordinate coded data of each object for each reference viewpoint, and the coded gain information obtained by coding the gain information indicating the gain amount are held in the server 11.

[0102]

Note that, hereinafter, the object polar coordinate position information and the object absolute coordinate position information are also simply referred to as

object position information in a case where it is not particularly necessary to distinguish the object polar coordinate position information and the object absolute coordinate position information. Similarly, hereinafter, the object polar coordinate coded data and the object absolute coordinate coded data are also simply referred to as object coordinate coded data in a case where it is not particularly necessary to distinguish the object polar coordinate coded data and the object absolute coordinate coded data.

[0103]

When the operation of the content reproduction system is started, the configuration information sending unit 21 of the server 11 transmits the system configuration information to the client 12 side prior to the transmission of the object coordinate coded data. Therefore, the client 12 side can understand the number of objects constituting the content, the number of reference viewpoints, the position of the reference viewpoint in the common absolute coordinate space, and the like.

[0104]

Next, the viewpoint selection unit 42 of the client 12 selects a reference viewpoint according to the listener position information, and the configuration information acquisition unit 43 sends the viewpoint selection information indicating the selection result to the server 11.

[0105]

Note that, as described above, the viewpoint selection unit 42 may be provided in the server 11, and

the reference viewpoint may be selected on the server 11 side.

[0106]

In such a case, the viewpoint selection unit 42 selects a reference viewpoint on the basis of the listener position information received from the client 12 by the configuration information sending unit 21 and the system configuration information, and supplies the viewpoint selection information indicating the selection result to the coded data sending unit 22.

[0107]

At this time, the viewpoint selection unit 42 specifies and selects, for example, two (or two or more) reference viewpoints sandwiching the listening position indicated by the listener position information. In other words, the two reference viewpoints are selected such that the listening position is located between the two reference viewpoints.

[0108]

Therefore, the object coordinate coded data for each of the plurality of selected reference viewpoints is transmitted to the client 12 side. Furthermore, in more detail, the coded data sending unit 22 transmits not only the object coordinate coded data but also the coded gain information to the client 12 regarding the two reference viewpoints indicated by the viewpoint selection information.

[0109]

On the client 12 side, the object absolute coordinate position information and the gain information at an arbitrary viewpoint of the current listener are

calculated by interpolation processing or the like on the basis of the object coordinate coded data, the coded gain information at each of the plurality of reference viewpoints received from the server 11, and the listener position information.

[0110]

Here, a specific example of calculation of final object absolute coordinate position information and gain information at an arbitrary viewpoint of the current listener will be described.

[0111]

In particular, an example of the interpolation processing using the data set of reference viewpoints of the polar coordinate system as two reference viewpoints sandwiching the listener will be described below.

[0112]

In such a case, the client 12 performs Processing PC1 to Processing PC4 described below in order to obtain final object absolute coordinate position information and gain information at the viewpoint of the listener.

[0113]

(Processing PC1)

In Processing PC1, each reference viewpoint is set as an origin from the data set at two reference viewpoints of the polar coordinate system, and the transformation into the absolute coordinate system position is performed on the object included in each data set. That is, the coordinate transformation unit 46 performs coordinate transformation as Processing PC1 with respect to the object polar coordinate position information of each object for each reference viewpoint,

and generates the object absolute coordinate position information.

[0114]

For example, as illustrated in Fig. 6, it is assumed that there is one object OBJ11 in a polar coordinate system space based on an origin O. Furthermore, a three-dimensional orthogonal coordinate system (absolute coordinate system) having the origin O as a reference (origin) and having an x axis, a y axis, and a z axis as respective axes is referred to as an xyz coordinate system.

[0115]

In this case, the position of the object OBJ11 in the polar coordinate system can be represented by polar coordinates including a horizontal angle θ, which is an angle in the horizontal direction, a vertical angle γ, which is an angle in the vertical direction, and a radius r indicating the distance from the origin O to the object OBJ11. In this example, the polar coordinates (θ, γ, r) are object polar coordinate position information of the object OBJ11.

[0116]

Note that the horizontal angle θ is an angle in the horizontal direction starting from the origin O, that is, the front of the listener. In this example, when a straight line (line segment) connecting the origin O and the object OBJ11 is LN and a straight line obtained by projecting the straight line LN on the xy plane is LN', an angle formed by the y axis and the straight line LN' is the horizontal angle θ.

[0117]

Furthermore, the vertical angle γ is an angle in the vertical direction starting from the origin O, that is, the front of the listener, and in this example, an angle formed by the straight line LN and the xy plane is the vertical angle γ. Moreover, the radius r is a distance from the listener (origin O) to the object OBJ11, that is, the length of the straight line LN.

[0118]

When the position of such object OBJ11 is expressed by coordinates (x, y, z) of the xyz coordinate system, that is, absolute coordinates, the position is indicated by Formula (1) described below.

[0119]

[Math. 1]

$$x = -r * \sin\theta * \cos\gamma$$
$$y = r * \cos\theta * \cos\gamma$$
$$z = r * \sin\gamma \quad \cdots (1)$$

[0120]

In Processing PC1, by calculating Formula (1) on the basis of the object polar coordinate position information, which is polar coordinates, the object absolute coordinate position information, which is absolute coordinates, indicating the position of the object in the xyz coordinate system (absolute coordinate system) having the position of the reference viewpoint as the origin O is calculated.

[0121]

In particular, in Processing PC1, for each of the two reference viewpoints, coordinate transformation is performed on the object polar coordinate position information of each of the plurality of objects at the

reference viewpoints.

[0122]

(Processing PC2)

In Processing PC2, for each of the two reference viewpoints, coordinate axis transformation processing is performed on the object absolute coordinate position information obtained by Processing PC1 for each object. That is, the coordinate axis transformation processing unit 47 performs the coordinate axis transformation processing as Processing PC2.

[0123]

The object absolute coordinate position information at each of the two reference viewpoints obtained by Processing PC1 described above, that is, obtained by the coordinate transformation unit 46 indicates the position in the xyz coordinate system having the reference viewpoints as the origin O. Therefore, the coordinates (coordinate system) of the object absolute coordinate position information are different for each reference viewpoint.

[0124]

Thus, the coordinate axis transformation processing of integrating the object absolute coordinate position information at each reference viewpoint into absolute coordinates of one common absolute coordinate system, that is, absolute coordinates in the common absolute coordinate system (common absolute coordinate space) is performed as Processing PC2.

[0125]

In order to perform this coordinate axis transformation processing, in addition to the data set

for each reference viewpoint, that is, the object absolute coordinate position information of each object for each reference viewpoint, absolute position information (reference viewpoint position information) of the listener and the listener direction information indicating the direction of the face of the listener are required.

[0126]

That is, the coordinate axis transformation processing requires the object absolute coordinate position information obtained by Processing PC1 and the system configuration information including the reference viewpoint position information indicating the position of the reference viewpoint in the common absolute coordinate system and the listener direction information at the reference viewpoint.

[0127]

Note that, here for the sake of brief description, only the rotation angle in the horizontal direction is used as the direction of the face indicated by the listener direction information, but information of up-and-down motion (pitch) of the face can also be added.

[0128]

Now, assuming that the common absolute coordinate system is an XYZ coordinate system having an X axis, a Y axis, and a Z axis as respective axes, and the rotation angle according to the direction of the face indicated by the listener direction information is φ, for example, the coordinate axis transformation processing is performed as illustrated in Fig. 7.

[0129]

That is, in the example illustrated in Fig. 7, as the coordinate axis transformation processing, the coordinate axis rotation of rotating the coordinate axis by the rotation angle φ, and the processing of shifting the origin of the coordinate axis from the position of the reference viewpoint to the origin position of the common absolute coordinate system, in more detail, the processing of shifting the position of the object according to the positional relationship between the reference viewpoint and the origin of the common absolute coordinate system are performed.

[0130]

In Fig. 7, a position P21 indicates the position of the reference viewpoint, and an arrow Q11 indicates the direction of the face of the listener indicated by the listener direction information at the reference viewpoint. In particular, here, the X coordinate and the Y coordinate of the position P21 in the common absolute coordinate system (XYZ coordinate system) are (Xref, Yref).

[0131]

Furthermore, a position P22 indicates the position of the object when the reference viewpoint is at the position P21. Here, the X coordinate and the Y coordinate of the common absolute coordinate system indicating the position P22 of the object are (Xobj, Yobj), and the x coordinate and the y coordinate of the xyz coordinate system indicating the position P22 of the object and having the reference viewpoint as the origin are (xobj, yobj).

[0132]

Moreover, in this example, the angle φ formed by the X axis of the common absolute coordinate system (XYZ coordinate system) and the x axis of the xyz coordinate system is the rotation angle φ of the coordinate axis transformation obtained from the listener direction information.

[0133]

Therefore, for example, the coordinate axis X (X coordinate) and the coordinate axis Y (Y coordinate) after the transformation are as indicated in Formula (2) described below.

[0134]

[Math. 2]

$$X = \text{Reference viewpoint X coordinate value} + x*\cos(\phi) + y*\sin(\phi)$$

$$Y = \text{Reference viewpoint Y coordinate value} - x*\sin(\phi) + y*\cos(\phi) \quad \ldots (2)$$

[0135]

Note that, in Formula (2), x and y represent the x axis (x coordinate) and the y axis (y coordinate) before transformation, that is, in the xyz coordinate system. Furthermore, “reference viewpoint X coordinate value” and “reference viewpoint Y coordinate value” in Formula (2) indicate an X coordinate and a Y coordinate indicating the position of the reference viewpoint in the XYZ coordinate system (common absolute coordinate system), that is, an X coordinate and a Y coordinate constituting the reference viewpoint position information.

[0136]

Given the above, in the example of Fig. 7, the X coordinate value Xobj and the Y coordinate value Yobj indicating the position of the object after the

coordinate axis transformation processing can be obtained from Formula (2).

[0137]

That is, φ in Formula (2) is set as the rotation angle φ obtained from the listener direction information at the position P21, and "Xref", "xobj", and "yobj" are substituted into "reference viewpoint X coordinate value", "x", and "y" in Formula (2), respectively, and the X coordinate value Xobj can be obtained.

[0138]

Furthermore, φ in Formula (2) is set as the rotation angle φ obtained from the listener direction information at the position P21, and "Yref", "xobj", and "yobj" are substituted into "reference viewpoint Y coordinate value", "x", and "y" in Formula (2), respectively, and the Y coordinate value Yobj can be obtained.

[0139]

Similarly, for example, when two reference viewpoints A and B are selected according to the viewpoint selection information, the X coordinate value and the Y coordinate value indicating the position of the object after the coordinate axis transformation processing for those reference viewpoints are as indicated in Formula (3) described below.

[0140]

[Math. 3]

xa = X coordinate value of reference viewpoint A + x*cos (ϕa) + y*sin (ϕa)

ya = Y coordinate value of reference viewpoint A − x*sin (ϕa) + y*cos (ϕa)

xb = X coordinate value of reference viewpoint B + x*cos (ϕb) + y*sin (ϕb)

yb = Y coordinate value of reference viewpoint B − x*sin (ϕb) + y*cos (ϕb)

... (3)

[0141]

Note that, in Formula (3), xa and ya represent the X coordinate value and the Y coordinate value of the XYZ coordinate system after the axis transformation (after the coordinate axis transformation processing) for the reference viewpoint A, and φa represents the rotation angle of the axis transformation for the reference viewpoint A, that is, the above-described rotation angle φ.

[0142]

Thus, when the x coordinate and the y coordinate constituting the object absolute coordinate position information at the reference viewpoint A obtained in Processing PC1 are substituted into Formula (3), the coordinate xa and the coordinate ya are obtained as the X coordinate and the Y coordinate indicating the position of the object in the XYZ coordinate system (common absolute coordinate system) at the reference viewpoint A. Absolute coordinates including the coordinate xa and the coordinate ya thus obtained and the Z coordinate are the object absolute coordinate position information output from the coordinate axis transformation processing unit 47.

[0143]

Note that, in this example, since only the rotation angle φ in the horizontal direction is handled, the

coordinate axis transformation is not performed for the Z axis (Z coordinate). Therefore, for example, it is sufficient if the z coordinate constituting the object absolute coordinate position information obtained in Processing PC1 is used as it is as the Z coordinate indicating the position of the object in the common absolute coordinate system.

[0144]

Similar to the reference viewpoint A, in Formula (3), xb and yb represent the X coordinate value and the Y coordinate value of the XYZ coordinate system after the axis transformation (after the coordinate axis transformation processing) for the reference viewpoint B, and φb represents the rotation angle of the axis transformation for the reference viewpoint B (rotation angle φ).

[0145]

In the coordinate axis transformation processing unit 47, the coordinate axis transformation processing as described above is performed as Processing PC2.

[0146]

Therefore, for example, when the coordinate axis transformation processing is performed on each of the four reference viewpoints illustrated in Fig. 3, the transformation result illustrated in Fig. 8 is obtained. Note that portions in Fig. 8 corresponding to those of Fig. 3 are designated by the same reference numerals, and description is omitted as appropriate.

[0147]

In Fig. 8, each circle (ring) represents one object. Furthermore, in Fig. 8, the upper side of the

drawing illustrates the position of each object on the polar coordinate system indicated by the object polar coordinate position information, and the lower side of the drawing illustrates the position of each object in the common absolute coordinate system.

[0148]

In particular, in Fig. 8, the left end illustrates the result of the coordinate axis transformation for the reference viewpoint “Origin” at the position P11 illustrated in Fig. 3, and the second from the left in Fig. 8 illustrates the result of the coordinate axis transformation for the reference viewpoint “Near” at the position P12 illustrated in Fig. 3.

[0149]

Furthermore, in Fig. 8, the third from the left illustrates the result of the coordinate axis transformation for the reference viewpoint “Far” at the position P13 illustrated in Fig. 3, and the right end in Fig. 8 illustrates the result of the coordinate axis transformation for the reference viewpoint “Back” at the position P14 illustrated in Fig. 3.

[0150]

For example, regarding the reference viewpoint “Origin”, since it is the origin viewpoint in which the position of the origin of the polar coordinate system is the position of the origin of the common absolute coordinate system, the position of the object viewed from the origin does not change before and after the transformation. On the other hand, at the remaining three reference viewpoints “Near”, “Far”, and “Back”, it can be seen that the position of the object is shifted to

the absolute coordinate position viewed from each viewpoint position. In particular, at the reference viewpoint "Back", since the direction of the face of the listener indicated by the listener direction information is backward, the object is positioned behind the reference viewpoint after the coordinate axis transformation processing.

[0151]

(Processing PC3)

In Processing PC3, the proportion ratio for the interpolation processing is obtained from the positional relationship between the absolute coordinate position of each of the two reference viewpoints, that is, the position indicated by the reference viewpoint position information included in the system configuration information and arbitrary listening position sandwiched between the positions of the two reference viewpoints.

[0152]

That is, the object position calculation unit 48 performs processing of obtaining the proportion ratio (m : n) as Processing PC3 on the basis of the listener position information supplied from the listener position information acquisition unit 41 and the reference viewpoint position information included in the system configuration information.

[0153]

Here, it is assumed that the reference viewpoint position information indicating the position of the reference viewpoint A is (x1, y1, z1), which is the first reference viewpoint, the reference viewpoint position information indicating the position of the reference

viewpoint B is (x2, y2, z2), which is the second reference viewpoint, and the listener position information indicating the listening position is (x3, y3, z3).

[0154]

In this case, the object position calculation unit 48 calculates the proportion ratio (m : n), that is, m and n of the proportion ratio by calculating Formula (4) described below.

[0155]

[Math. 4]

$$m=SQRT((x3-x1)*(x3-x1)+(y3-y1)*(y3-y1)+(z3-z1)*(z3-z1))$$
$$n=SQRT((x3-x2)*(x3-x2)+(y3-y2)*(y3-y2)+(z3-z2)*(z3-z2)) \quad \cdots (4)$$

[0156]

(Processing PC4)

Subsequently, the object position calculation unit 48 performs the interpolation processing as Processing PC4 on the basis of the proportion ratio (m : n) obtained by Processing PC3 and the object absolute coordinate position information of each object of the two reference viewpoints supplied from the coordinate axis transformation processing unit 47.

[0157]

That is, in Processing PC4, by applying the proportion ratio (m : n) obtained in Processing PC3 to the same object corresponding to the two reference viewpoints obtained in Processing PC2, the object position and the gain amount corresponding to an arbitrary listening position are obtained.

[0158]

Here, the absolute coordinate position of a predetermined object viewed from the reference viewpoint A, that is, the object absolute coordinate position information of the reference viewpoint A obtained by Processing PC2 is (xa, ya, za), and the gain amount indicated by the gain information of the predetermined object for the reference viewpoint A is g1.

[0159]

Similarly, the absolute coordinate position of the above-described predetermined object viewed from the reference viewpoint B, that is, the object absolute coordinate position information of the reference viewpoint B obtained by Processing PC2 is (xb, yb, zb), and the gain amount indicated by the gain information of the object for the reference viewpoint B is g2.

[0160]

Furthermore, the absolute coordinates indicating the position of the above-described predetermined object in the XYZ coordinate system (common absolute coordinate system) and the gain amount corresponding to an arbitrary viewpoint position between the reference viewpoint A and the reference viewpoint B, that is, the listening position indicated by the listener position information are set as (xc, yc, zc) and gain_c. The absolute coordinates (xc, yc, zc) are final object absolute coordinate position information output from the object position calculation unit 48 to the polar coordinate transformation unit 49.

[0161]

At this time, the final object absolute coordinate

position information (xc, yc, zc) and the gain amount gain_c for the predetermined object can be obtained by calculating Formula (5) described below using the proportion ratio (m : n).

[0162]

[Math. 5]

$$xc=(m*xb+n*xa)/(m+n)$$
$$yc=(m*yb+n*ya)/(m+n)$$
$$zc=(m*zb+n*za)/(m+n)$$
$$gain_c=(m*g2+n*g1)/(m+n) \quad \cdots (5)$$

[0163]

The positional relationship between the reference viewpoint A, the reference viewpoint B, and the listening position described above and the positional relationship of the same object at the respective positions of the reference viewpoint A, the reference viewpoint B, and the listening position are as illustrated in Fig. 9.

[0164]

In Fig. 9, the horizontal axis and the vertical axis indicate the X axis and the Y axis of the XYZ coordinate system (common absolute coordinate system), respectively. Note that, here for the sake of brief description, only the X-axis direction and the Y-axis direction are illustrated.

[0165]

In this example, a position P51 is a position indicated by the reference viewpoint position information (x1, y1, z1) of the reference viewpoint A, and a position P52 is a position indicated by the reference viewpoint position information (x2, y2, z2) of the reference viewpoint B.

[0166]

Furthermore, a position P53 between the reference viewpoint A and the reference viewpoint B is a listening position indicated by the listener position information (x3, y3, z3).

[0167]

In Formula (4) described above, the proportion ratio (m : n) is obtained on the basis of the positional relationship between the reference viewpoint A, the reference viewpoint B, and the listening position.

[0168]

Furthermore, a position P61 is a position indicated by the object absolute coordinate position information (xa, ya, za) at the reference viewpoint A, and a position P62 is a position indicated by the object absolute coordinate position information (xb, yb, zb) at the reference viewpoint B.

[0169]

Moreover, a position P63 between the position P61 and the position P62 is a position indicated by the object absolute coordinate position information (xc, yc, zc) at the listening position.

[0170]

By performing the calculation of Formula (5), that is, the interpolation processing in this manner, the object absolute coordinate position information indicating an appropriate object position can be obtained for an arbitrary listening position.

[0171]

Note that the example of obtaining the object position, that is, the final object absolute coordinate

position information using the proportion ratio (m : n) has been described above, but it is not limited thereto, and the final object absolute coordinate position information may be estimated using machine learning or the like.

[0172]

Furthermore, in a case where an absolute coordinate system editor is used, that is, in the case of the content reproduction system illustrated in Fig. 2, each object position of each reference viewpoint, that is, the position indicated by the object absolute coordinate position information is a position on one common absolute coordinate system. In other words, the position of the object at each reference viewpoint is expressed by absolute coordinates of the common absolute coordinate system.

[0173]

Therefore, in the content reproduction system illustrated in Fig. 2, it is sufficient if the object absolute coordinate position information obtained by the decoding of the decode unit 45 is used as the input in Processing PC3 described above. That is, it is sufficient if the calculation of Formula (4) is performed on the basis of the object absolute coordinate position information obtained by decoding.

[0174]

<Regarding operation of the content reproduction system>

Next, a flow (sequence) of processing performed in the content reproduction system described above will be described with reference to Fig. 10.

[0175]

Note that, here, an example in which the reference viewpoint is selected on the server 11 side and the object polar coordinate coded data is prepared in advance on the server 11 side will be described. That is, an example in which the viewpoint selection unit 42 is provided on the server 11 side in the example of the content reproduction system illustrated in Fig. 1 will be described.

[0176]

First, on the server 11 side, for all reference viewpoints, the polar coordinate system object position information, that is, object polar coordinate coded data is generated and held by a polar coordinate system editor, and system configuration information is also generated and held.

[0177]

Then, the configuration information sending unit 21 transmits the system configuration information to the client 12 via a network or the like.

[0178]

Then, the configuration information acquisition unit 43 of the client 12 receives the system configuration information transmitted from the server 11 and supplies the system configuration information to the coordinate axis transformation processing unit 47. At this time, the client 12 decodes (decoding) the received system configuration information and initializes the client system.

[0179]

Subsequently, when the listener position

information acquisition unit 41 acquires the listener position information and supplies the listener position information to the configuration information acquisition unit 43, the configuration information acquisition unit 43 transmits the listener position information supplied from the listener position information acquisition unit 41 to the server 11.

[0180]

Furthermore, the configuration information sending unit 21 receives the listener position information transmitted from the client 12 and supplies the listener position information to the viewpoint selection unit 42. Then, the viewpoint selection unit 42 selects reference viewpoints necessary for the interpolation processing, that is, for example, two reference viewpoints sandwiching the above-described listening position on the basis of the listener position information supplied from the configuration information sending unit 21 and the system configuration information, and supplies the viewpoint selection information indicating the selection result to the coded data sending unit 22.

[0181]

The coded data sending unit 22 prepares for transmission of the polar coordinate system object position information of the reference viewpoints necessary for the interpolation processing according to the viewpoint selection information supplied from the viewpoint selection unit 42.

[0182]

That is, the coded data sending unit 22 generates a bitstream by reading and multiplexing the object polar

coordinate coded data of the reference viewpoint indicated by the viewpoint selection information and the coded gain information. Then, the coded data sending unit 22 transmits the generated bitstream to the client 12.

[0183]

The coded data acquisition unit 44 receives and demultiplexes the bitstream transmitted from the server 11, and supplies the resultant object polar coordinate coded data and coded gain information to the decode unit 45.

[0184]

The decode unit 45 decodes the object polar coordinate coded data supplied from the coded data acquisition unit 44, and supplies the resultant object polar coordinate position information to the coordinate transformation unit 46. Furthermore, the decode unit 45 decodes the coded gain information supplied from the coded data acquisition unit 44, and supplies the resultant gain information to the object position calculation unit 48 via the coordinate transformation unit 46 and the coordinate axis transformation processing unit 47.

[0185]

The coordinate transformation unit 46 transforms the polar coordinate information into absolute coordinate position information centered on the listener for the object polar coordinate position information supplied from the decode unit 45.

[0186]

That is, for example, the coordinate transformation

unit 46 calculates Formula (1) described above on the basis of the object polar coordinate position information and supplies the resultant object absolute coordinate position information to the coordinate axis transformation processing unit 47.

[0187]

Subsequently, the coordinate axis transformation processing unit 47 performs development from the absolute coordinate position information centered on the listener to the common absolute coordinate space by coordinate axis transformation.

[0188]

For example, the coordinate axis transformation processing unit 47 performs the coordinate axis transformation processing by calculating Formula (3) described above on the basis of the system configuration information supplied from the configuration information acquisition unit 43 and the object absolute coordinate position information supplied from the coordinate transformation unit 46, and supplies the resultant object absolute coordinate position information to the object position calculation unit 48.

[0189]

The object position calculation unit 48 calculates a proportion ratio for interpolation processing from the current listener position and the reference viewpoint.

[0190]

For example, the object position calculation unit 48 calculates Formula (4) described above on the basis of the listener position information supplied from the listener position information acquisition unit 41 and the

reference viewpoint position information of the plurality of reference viewpoints selected by the viewpoint selection unit 42, and calculates the proportion ratio (m : n).

[0191]

Furthermore, the object position calculation unit 48 calculates the object position and the gain amount corresponding to the current listener position using the proportion ratio from the object position and the gain amount corresponding to the reference viewpoints sandwiching the listener position.

[0192]

For example, the object position calculation unit 48 performs interpolation processing by calculating Formula (5) described above on the basis of the object absolute coordinate position information and the gain information supplied from the coordinate axis transformation processing unit 47 and the proportion ratio (m : n), and supplies the resultant final object absolute coordinate position information and the gain information to the polar coordinate transformation unit 49.

[0193]

Then, thereafter, the client 12 performs rendering processing to which the calculated object position and gain amount are applied.

[0194]

For example, the polar coordinate transformation unit 49 performs transformation of the absolute coordinate position information into polar coordinates.

[0195]

That is, for example, the polar coordinate transformation unit 49 performs the polar coordinate transformation on the object absolute coordinate position information supplied from the object position calculation unit 48 on the basis of the listener position information supplied from the listener position information acquisition unit 41.

[0196]

The polar coordinate transformation unit 49 supplies the polar coordinate position information obtained by the polar coordinate transformation and the gain information supplied from the object position calculation unit 48 to the subsequent rendering processing unit.

[0197]

Then, the rendering processing unit performs polar coordinate rendering processing on all the objects.

[0198]

That is, the rendering processing unit performs the rendering processing in the polar coordinate system defined, for example, by MPEG-H on the basis of the polar coordinate position information and the gain information of all the objects supplied from the polar coordinate transformation unit 49, and generates reproduction audio data for reproducing the sound of the content.

[0199]

Here, for example, vector based amplitude panning (VBAP) or the like is performed as the rendering processing in the polar coordinate system defined by MPEG-H. Note that, in more detail, gain adjustment based on the gain information is performed on the audio data

before the rendering processing, but the gain adjustment may be performed not by the rendering processing unit but by the preceding polar coordinate transformation unit 49.

[0200]

When the above processing is performed on a predetermined frame and the reproduction audio data is generated, content reproduction based on the reproduction audio data is appropriately performed. Then, thereafter, the listener position information is appropriately transmitted from the client 12 to the server 11, and the above-described processing is repeatedly performed.

[0201]

As described above, the content reproduction system calculates the object absolute coordinate position information and the gain information of an arbitrary listening position by interpolation processing from the object position information of the plurality of reference viewpoints. In this way, it is possible to realize the object arrangement based on the intention of the content creator according to the listening position instead of the simple physical relationship between the listener and the object. Therefore, content reproduction based on the intention of the content creator can be realized, and the interest of the content can be sufficiently conveyed to the listener.

[0202]

<Regarding the listener and the object>

By the way, as the reference viewpoint, for example, two examples of assuming a viewpoint as a listener and assuming a viewpoint of a performer imagining to be an object are conceivable.

[0203]

In the latter case, since the listener and the object overlap at the reference viewpoint, that is, the listener and the object are at the same position, the following Cases CA1 to CA3 are conceivable.

[0204]

(Case CA1)

The listener is prohibited from overlapping with the object, or the listener is prohibited from entering a specific range

(Case CA2)

The listener is merged with the object and a sound generated from the object is output from all channels

(Case CA3)

A sound generated from overlapping objects is muted or attenuated

[0205]

For example, in the case of Case CA2, the sense of localization in the head of the listener can be recreated.

[0206]

Furthermore, in Case CA3, by muting or attenuating the sound of the object, the listener becomes a performer, and, for example, use in a karaoke mode is also conceivable. In this case, a surrounding accompaniment or the like other than the performer's singing voice surrounds the listener itself, and a feeling of singing thereinside can be obtained.

[0207]

In a case where the content creator has such intention, identifiers indicating Cases CA1 to CA3 can be

stored in a coded bitstream transmitted from the server 11 and can be transmitted to the client 12 side. For example, such an identifier is information indicating the above-described reproduction mode.

[0208]

Furthermore, in the content reproduction system described above, the listener may move around between two reference viewpoints.

[0209]

In such a case, there may be a case where some listener desires to intentionally bring an object (viewpoint) closer to the object arrangement of one (one side) of the two reference viewpoints. Specifically, for example, there may be a request for maintaining an angle that allows the listener's favorite artist to be easily seen at all times.

[0210]

Therefore, for example, the degree of bringing may be controlled by biasing the proportion processing of the internal division ratio. This can be realized by newly introducing a bias coefficient α into Formula (5) for obtaining interpolation described above, for example, as illustrated in Fig. 11.

[0211]

Fig. 11 illustrates characteristics in a case where the bias coefficient α is multiplied. In particular, the upper side in the drawing illustrates an example of bringing the object closer to the arrangement on a viewpoint X1 side, that is, the above-described reference viewpoint A side.

[0212]

On the other hand, the lower side in the drawing illustrates an example of bringing the object closer to the arrangement on a viewpoint X2 side, that is, the above-described reference viewpoint B side.

[0213]

Note that, in Fig. 11, the horizontal axis indicates the position of a predetermined viewpoint X3 in a case where the bias coefficient α is not introduced, and the vertical axis indicates the position of a predetermined viewpoint X3 in a case where the bias coefficient α is introduced. Furthermore, here, the position of the reference viewpoint A (viewpoint X1) is "0", and the position of the reference viewpoint B (viewpoint X2) is "1".

[0214]

In the example of the upper side in the drawing, for example, when the listener moves from the reference viewpoint A (viewpoint X1) side to the position of the reference viewpoint B (viewpoint X2), the smaller the bias coefficient α, the more the listener feels difficulty to reach the position of the reference viewpoint B (viewpoint X2).

[0215]

Conversely, in the example of the lower side in the drawing, for example, when the listener moves from the reference viewpoint A side to the position of the reference viewpoint B, the smaller the bias coefficient α, the more the listener feels to immediately reach the position of the reference viewpoint B.

[0216]

For example, in the case of bringing the object

closer to the arrangement on the reference viewpoint A side, the final object absolute coordinate position information (xc, yc, zc) and the gain amount gain_c can be obtained by calculating Formula (6) described below.

[0217]

On the other hand, in the case of bringing the object closer to the arrangement on the reference viewpoint B side, the final object absolute coordinate position information (xc, yc, zc) and the gain amount gain_c can be obtained by calculating Formula (7) described below.

[0218]

However, in Formulae (6) and (7), m and n of the proportion ratio (m : n) and the bias coefficient α are as indicated in Formula (8) described below.

[0219]

[Math. 6]

$$xc=(m*xb+\alpha*n*xa)/(m+\alpha*n)$$
$$yc=(m*yb+\alpha*n*ya)/(m+\alpha*n)$$
$$zc=(m*zb+\alpha*n*za)/(m+\alpha*n)$$
$$gain_c=(m*g2+\alpha*n*g1)/(m+\alpha*n) \quad \cdots(6)$$

[0220]

[Math. 7]

$$xc=(\alpha*m*xb+n*xa)/(\alpha*m+n)$$
$$yc=(\alpha*m*yb+n*ya)/(\alpha*m+n)$$
$$zc=(\alpha*m*zb+n*za)/(\alpha*m+n)$$
$$gain_c=(\alpha*m*g2+n*g1)/(\alpha*m+n) \quad \cdots(7)$$

[0221]

[Math. 8]

$$m=SQRT((x3-x1)*(x3-x1)+(y3-y1)*(y3-y1)+(z3-z1)*(z3-z1))$$
$$n=SQRT((x3-x2)*(x3-x2)+(y3-y2)*(y3-y2)+(z3-z2)*(z3-z2))$$
$$0<\alpha\leqq 1 \quad \cdots(8)$$

[0222]

Note that, in Formula (8), the reference viewpoint position information (x1, y1, z1), the reference viewpoint position information (x2, y2, z2), and the listener position information (x3, y3, z3) are similar to those in Formula (4) described above-described.

[0223]

Obtaining the final object absolute coordinate position information and the gain amount using the bias coefficient α as in Formulae (6) and (7) is to obtain the final object absolute coordinate position information and the gain amount by performing the interpolation processing by giving a weight of the bias coefficient α with respect to the object absolute coordinate position information and the gain information of a predetermined reference viewpoint.

[0224]

When the object position information of the absolute coordinates after the interpolation processing obtained in this way, that is, the object absolute coordinate position information is combined with the listener position information and transformed into the polar coordinate information (polar coordinate position information), it is possible to perform the polar coordinate rendering processing used in the existing MPEG-H in a subsequent stage.

[0225]

<Regarding the interpolation processing of the object absolute coordinate position information and the gain information>

Meanwhile, as an example in which the object position calculation unit 48 obtains the object absolute coordinate position information and the gain information at an arbitrary viewpoint position, that is, listening position by the interpolation processing, the two-point interpolation using the information of the two reference viewpoints has been described above.

[0226]

However, it is not limited thereto, and the object absolute coordinate position information and the gain information at an arbitrary listening position may be obtained by performing three-point interpolation using the information of three reference viewpoints. Furthermore, the object absolute coordinate position information and the gain information at an arbitrary listening position may be obtained by using the information of four or more reference viewpoints. Hereinafter, a specific example in a case where three-point interpolation is performed will be described.

[0227]

For example, as illustrated on the left side of Fig. 12, it is considered that the object absolute coordinate position information at an arbitrary listening position F is obtained by the interpolation processing.

[0228]

In this example, there are three reference viewpoints: reference viewpoint A, reference viewpoint B, and reference viewpoint C so as to surround the listening

position F, and here, it is assumed that the interpolation processing is performed using the information of the reference viewpoints A to C.

[0229]

Hereinafter, it is assumed that the X coordinate and the Y coordinate of the listening position F in the common absolute coordinate system, that is, the XYZ coordinate system, are (x_f, y_f).

[0230]

Similarly, it is assumed that the X coordinates and the Y coordinates of the respective positions of the reference viewpoint A, the reference viewpoint B, and the reference viewpoint C are (x_a, y_a), (x_b, y_b), and (x_c, y_c).

[0231]

In this case, as illustrated on the right side of Fig. 12, an object position F' at the listening position F is obtained on the basis of the coordinates of an object position A', an object position B', and an object position C' respectively corresponding to the reference viewpoint A, the reference viewpoint B, and the reference viewpoint C.

[0232]

Here, for example, the object position A' indicates the position of the object when the viewpoint is at the reference viewpoint A, that is, the position of the object in the common absolute coordinate system indicated by the object absolute coordinate position information of the reference viewpoint A.

[0233]

Furthermore, the object position F' indicates the position of the object in the common absolute coordinate

system when the listener is at the listening position F, that is, the position indicated by the object absolute coordinate position information to be the output of the object position calculation unit 48.

[0234]

Hereinafter, it is assumed that the X coordinates and the Y coordinates of the object position A', the object position B', and the object position C' are (x_a', y_a'), (x_b', y_b'), and (x_c', y_c'), and the X coordinate and the Y coordinate of the object position F' are (x_f', y_f').

[0235]

Furthermore, hereinafter, a triangular region surrounded by arbitrary three reference viewpoints such as the reference viewpoints A to C, that is, a region having a triangular shape formed by the three reference viewpoints is also referred to as a triangle mesh.

[0236]

Since there is a plurality of reference viewpoints in the common absolute coordinate space, a plurality of triangle meshes having the reference viewpoints as vertices can be formed in the common absolute coordinate space.

[0237]

Similarly, hereinafter, a triangular region surrounded (formed) by the object positions indicated by the object absolute coordinate position information of arbitrary three reference viewpoints such as the object positions A' to C' is also referred to as a triangle mesh.

[0238]

For example, in the example of two-point interpolation, the listener can move to an arbitrary position on a line segment connecting two reference viewpoints and listen to the sound of the content.

[0239]

On the other hand, in a case where three-point interpolation is performed, the listener can move to an arbitrary position in the region of the triangle mesh surrounded by the three reference viewpoints and listen to the sound of the content. That is, a region other than a line segment connecting two reference viewpoints in the case of two-point interpolation can be covered as the listening position.

[0240]

Also in a case where three-point interpolation is performed, similarly to the case of two-point interpolation, coordinates indicating an arbitrary position in the common absolute coordinate system (XYZ coordinate system) can be obtained from the coordinates of the arbitrary position in the xyz coordinate system, the listener direction information, and the reference viewpoint position information by Formula (2) described above.

[0241]

Note that, here, the Z coordinate value of the XYZ coordinate system is assumed to be the same as the z coordinate value of the xyz coordinate system, but in a case where the Z coordinate value and the z coordinate value are different, it is sufficient if the Z coordinate value indicating an arbitrary position is obtained by adding the Z coordinate value indicating the position of

the reference viewpoint in the XYZ coordinate system to the z coordinate value of the arbitrary position.

[0242]

It is proved by Ceva's theorem that an arbitrary listening position in a triangle mesh formed by three reference viewpoints is uniquely determined by an intersection of line segments from each of three vertices of the triangle mesh to each of internally dividing points of the three sides not adjacent to the vertices when the internal division ratio of each side of the triangle mesh is appropriately determined.

[0243]

This is established in all the triangle meshes regardless of the shape of the triangle mesh when the configuration of the internal division ratio of the three sides of the triangle mesh is determined from the proof formula.

[0244]

Therefore, when the internal division ratio of the triangle mesh including the listening position is obtained regarding the viewpoint side, that is, the reference viewpoint, and the internal division ratio is applied to the triangle mesh on the object side, that is, the object position, an appropriate object position for an arbitrary listening position can be obtained.

[0245]

Hereinafter, an example of obtaining the object absolute coordinate position information indicating the position of the object at the time of being at an arbitrary listening position using such a property of internal division ratio will be described.

[0246]

In this case, first, the internal division ratio of the side of the triangle mesh of the reference viewpoint on the XY plane of the XYZ coordinate system, which is a two-dimensional space, is obtained.

[0247]

Next, on the XY plane, the above-described internal division ratio is applied to the triangle mesh of the object positions corresponding to the three reference viewpoints, and the X coordinate and the Y coordinate of the position of the object corresponding to the listening position on the XY plane are obtained.

[0248]

Moreover, the Z coordinate of the object corresponding to the listening position is obtained on the basis of the three-dimensional plane including the positions of the three objects corresponding to the three reference viewpoints in the three-dimensional space (XYZ coordinate system) and the X coordinate and the Y coordinate of the object at the listening position on the XY plane.

[0249]

Here, an example of obtaining the object absolute coordinate position information indicating the object position F' and the gain information by the interpolation processing for the listening position F illustrated in Fig. 12 will be described with reference to Figs. 13 to 15.

[0250]

For example, as illustrated in Fig. 13, first, the X coordinate and the Y coordinate of the internally

dividing point in the triangle mesh including the reference viewpoints A to C including the listening position F are obtained.

[0251]

Now, an intersection of a straight line passing through the listening position F and the reference viewpoint C and a line segment AB from the reference viewpoint A to the reference viewpoint B is defined as a point D, and coordinates indicating the position of the point D on the XY plane are defined as (x_d, y_d). That is, the point D is an internally dividing point on the line segment AB (side AB).

[0252]

At this time, the relationship indicated in Formula (9) described below is established for the X coordinate and the Y coordinate indicating the position of an arbitrary point on a line segment CF from the reference viewpoint C to the listening position F, and the X coordinate and the Y coordinate indicating the position of an arbitrary point on the line segment AB.

[0253]

[Math. 9]

$$\text{Line segment CF: } Y=\alpha_1 X-\alpha_1 x_c+y_c, \text{ where } \alpha_1=(y_c-y_f)/(x_c-x_f)$$
$$\text{Line segment AB: } Y=\alpha_2 X-\alpha_2 x_a+y_a, \text{ where } \alpha_2=(y_b-y_a)/(x_b-x_a) \quad \cdots (9)$$

[0254]

Furthermore, since the point D is an intersection of a straight line passing through the reference viewpoint C and the listening position F and the line segment AB, the coordinates (x_d, y_d) of the point D on the XY plane can be obtained from Formula (9), and the coordinates (x_d, y_d) are as indicated in Formula (10)

described below.

[0255]

[Math. 10]

$$x_d = (\alpha_1 x_c - y_c - \alpha_2 x_a + y_a) / (\alpha_1 - \alpha_2)$$
$$y_d = \alpha_1 x_d - \alpha_1 x_c + y_c \quad \cdots (10)$$

[0256]

Therefore, as indicated in Formula (11) described below, on the basis of the coordinates (x_d, y_d) of the point D, the coordinates (x_a, y_a) of the reference viewpoint A, and the coordinates (x_b, y_b) of the reference viewpoint B, the internal division ratio (m, n) of the line segment AB by the point D, that is, the division ratio can be obtained.

[0257]

[Math. 11]

$$m = \mathrm{sqrt}((x_a - x_d)^2 + (y_a - y_d)^2)$$
$$n = \mathrm{sqrt}((x_b - x_d)^2 + (y_b - y_d)^2) \quad \cdots (11)$$

[0258]

Similarly, an intersection of a straight line passing through the listening position F and the reference viewpoint B and a line segment AC from the reference viewpoint A to the reference viewpoint C is defined as a point E, and coordinates indicating the position of the point E on the XY plane are defined as (x_e, y_e). That is, the point E is an internally dividing point on the line segment AC (side AC).

[0259]

At this time, the relationship indicated in Formula (12) described below is established for the X coordinate and the Y coordinate indicating the position of an arbitrary point on a line segment BF from the reference

viewpoint B to the listening position F, and the X coordinate and the Y coordinate indicating the position of an arbitrary point on the line segment AC.

[0260]

[Math. 12]

$$\text{Line segment BF: } Y=\alpha_3 X-\alpha_3 x_b+y_b, \text{ where } \alpha_3=(y_b-y_f)/(x_b-x_f)$$
$$\text{Line segment AC: } Y=\alpha_4 X-\alpha_4 x_a+y_a, \text{ where } \alpha_4=(y_c-y_a)/(x_c-x_a) \quad \cdots (12)$$

[0261]

Furthermore, since the point E is an intersection of a straight line passing through the reference viewpoint B and the listening position F and the line segment AC, the coordinates (x_e, y_e) of the point E on the XY plane can be obtained from Formula (12), and the coordinates (x_e, y_e) are as indicated in Formula (13) described below.

[0262]

[Math. 13]

$$x_e=(\alpha_3 x_b-y_b-\alpha_4 x_a+y_a)/(\alpha_3-\alpha_4)$$
$$y_e=\alpha_3 x_e-\alpha_3 x_b+y_b \quad \cdots (13)$$

[0263]

Therefore, as indicated in Formula (14) described below, on the basis of the coordinates (x_e, y_e) of the point E, the coordinates (x_a, y_a) of the reference viewpoint A, and the coordinates (x_c, y_c) of the reference viewpoint C, the internal division ratio (k, l) of the line segment AC by the point E, that is, the division ratio can be obtained.

[0264]

[Math. 14]

$$k=\mathrm{sqrt}((x_a-x_e)^2+(y_a-y_e)^2)$$
$$l=\mathrm{sqrt}((x_c-x_e)^2+(y_c-y_e)^2) \quad \cdots (14)$$

[0265]

Next, by applying the ratios of the two sides obtained in this manner, that is, the internal division ratio (m, n) and the internal division ratio (k, l) to the object-side triangle mesh as illustrated in Fig. 14, the coordinates (x_f', y_f') of the object position F' on the XY plane are obtained.

[0266]

Specifically, in this example, a point corresponding to the point D on a line segment A'B' connecting the object position A' and the object position B' is a point D'.

[0267]

Similarly, a point corresponding to the point E on a line segment A'C' connecting the object position A' and the object position C' is a point E'.

[0268]

Furthermore, an intersection between a straight line passing through the object position C' and the point D' and a straight line passing through the object position B' and the point E' is the object position F' corresponding to the listening position F.

[0269]

Here, it is assumed that the internal division ratio of the line segment A'B' by the point D' is the same internal division ratio (m, n) as in the case of the point D. At this time, the coordinates (x_d', y_d') of the point D' on the XY plane can be obtained on the basis of the internal division ratio (m, n), the coordinates (x_a', y_a') of the object position A', and the coordinates (x_b', y_b') of the object position B' as indicated in Formula

(15) described below.

[0270]

[Math. 15]

$$x_d' = (nx_a' + mx_b')/(m+n)$$
$$y_d' = (ny_a' + my_b')/(m+n) \quad \cdots (15)$$

[0271]

Furthermore, it is assumed that the internal division ratio of the line segment A'C' by the point E' is the same internal division ratio (k, l) as in the case of the point E. At this time, the coordinates (x_e', y_e') of the point E' on the XY plane can be obtained on the basis of the internal division ratio (k, l), the coordinates (x_a', y_a') of the object position A', and the coordinates (x_c', y_c') of the object position C' as indicated in Formula (16) described below.

[0272]

[Math. 16]

$$x_e' = (lx_a' + kx_c')/(k+l)$$
$$y_e' = (ly_a' + ky_c')/(k+l) \quad \cdots (16)$$

[0273]

Therefore, the relationship indicated in Formula (17) described below is established for the X coordinate and the Y coordinate indicating the position of an arbitrary point on a line segment B'E' from the object position B' to the point E', and the X coordinate and the Y coordinate indicating the position of an arbitrary point on a line segment C'D' from the object position C' to the point D'.

[0274]

[Math. 17]

Line segment B' E': $Y=\alpha_5 X+y_b'-\alpha_5 x_b'$, where $\alpha_5=(y_e'-y_b')/(x_e'-x_b')$
Line segment C' D': $Y=\alpha_6 X+y_c'-\alpha_6 x_c'$, where $\alpha_6=(y_d'-y_c')/(x_d'-x_c')$

$$\cdots (17)$$

[0275]

Since the target object position F' is the intersection of the line segment B'E' and the line segment C'D', the coordinates (x_f', y_f') of the object position F' can be obtained by Formula (18) described below from the relationship of Formula (17).

[0276]

[Math. 18]

$$x_f'=(-y_b'+\alpha_5 x_b'+y_c'-\alpha_6 x_c')/(\alpha_5-\alpha_6)$$
$$y_f'=\alpha_6 x_f'+y_c'-\alpha_6 x_c' \quad \cdots (18)$$

[0277]

Through the above processing, the coordinates (x_f', y_f') of the object position F' on the XY plane are obtained.

[0278]

Subsequently, the coordinates (x_f', y_f', z_f') of the object position F' in the XYZ coordinate system are obtained on the basis of the coordinates (x_f', y_f') of the object position F' on the XY plane, the coordinates (x_a', y_a', z_a') of the object position A', the coordinates (x_b', y_b', z_b') of the object position B', and the coordinates (x_c', y_c', z_c') of the object position C' in the XYZ coordinate system. That is, the Z coordinate z_f' of the object position F' in the XYZ coordinate system is obtained.

[0279]

For example, a triangle on a three-dimensional

space having the object position A', the object position B', and the object position C' as vertices in the XYZ coordinate system (common absolute coordinate space), that is, a three-dimensional plane A'B'C' including the object position A', the object position B', and the object position C' is obtained. Then, a point having the X coordinate and the Y coordinate (x_f', y_f') on the three-dimensional plane A'B'C' is obtained, and the Z coordinate of the point is z_f'.

[0280]

Specifically, a vector having the object position A' in the XYZ coordinate system as a start point and the object position B' as an end point is set as a vector A'B' = (x_{ab}', y_{ab}', z_{ab}').

[0281]

Similarly, a vector having the object position A' in the XYZ coordinate system as a start point and the object position C' as an end point is set as a vector A'C' = (x_{ac}', y_{ac}', z_{ac}').

[0282]

These vectors A'B' and A'C' can be obtained on the basis of the coordinates (x_a', y_a', z_a') of the object position A', the coordinates (x_b', y_b', z_b') of the object position B', and the coordinates (x_c', y_c', z_c') of the object position C'. That is, the vectors A'B' and A'C' can be obtained by Formula (19) described below.

[0283]

[Math. 19]

$$\text{Vector A' B'}: (x_{ab}', y_{ab}', z_{ab}') = (x_b' - x_a', y_b' - y_a', z_b' - z_a')$$
$$\text{Vector A' C'}: (x_{ac}', y_{ac}', z_{ac}') = (x_c' - x_a', y_c' - y_a', z_c' - z_a') \quad \cdots (19)$$

[0284]

Furthermore, a normal vector (s, t, u) of the three-dimensional plane A'B'C' is an outer product of the vectors A'B' and A'C', and can be obtained by Formula (20) described below.

[0285]

[Math. 20]

$$(s, t, u) = (y_{ab}' z_{ac}' - z_{ab}' y_{ac}', z_{ab}' x_{ac}' - x_{ab}' z_{ac}', x_{ab}' y_{ac}' - y_{ab}' x_{ac}') \quad \cdots (20)$$

[0286]

Therefore, from the normal vector (s, t, u) and the coordinates (x_a', y_a', z_a') of the object position A', the plane equation of the three-dimensional plane A'B'C' is as indicated in Formula (21) described below.

[0287]

[Math. 21]

$$s(X - x_a') + t(Y - y_a') + u(Z - z_a') = 0 \quad \cdots (21)$$

[0288]

Here, since the X coordinate x_f' and the Y coordinate y_f' of the object position F' on the three-dimensional plane A'B'C' have already been obtained, the Z coordinate z_f' can be obtained as indicated in Formula (21) described below by substituting the X coordinate x_f' and the Y coordinate y_f' into X and Y of the plane equation of Formula (22).

[0289]

[Math. 22]

$$z_f' = (-s(x_f' - x_a') - t(y_f' - y_a'))/u + z_a' \quad \cdots (22)$$

[0290]

Through the above calculation, the coordinates (x_f', y_f', z_f') of the target object position F' are

obtained. The object position calculation unit 48 outputs the object absolute coordinate position information indicating the coordinates (x_f', y_f', z_f') of the object position F' obtained in the above manner.

[0291]

Furthermore, similarly to the case of the object absolute coordinate position information, the gain information can also be obtained by three-point interpolation.

[0292]

That is, the gain information of the object at the object position F' can be obtained by performing the interpolation processing on the basis of the gain information of the object when the viewpoint is at each of the reference viewpoints A to C.

[0293]

For example, as illustrated in Fig. 15, it is considered to obtain a gain information G_f' of the object at the object position F' in the triangle mesh formed by the object position A', the object position B', and the object position C'.

[0294]

Now, it is assumed that the gain information of the object at the object position A' when the viewpoint is at the reference viewpoint A is G_a', the gain information of the object at the object position B' is G_b', and the gain information of the object at the object position C' is G_c'.

[0295]

In this case, first, the gain information G_d' of the object at the point D', which is the internally

dividing point of the line segment A'B' when the viewpoint is virtually at the point D, is obtained.

[0296]

Specifically, the gain information G_d' can be obtained by calculating Formula (23) described below on the basis of the internal division ratio (m, n) of the above-described line segment A'B', and the gain information G_a' of the object position A' and the gain information G_b' of the object position B'.

[0297]

[Math. 23]

$$G_d' = (m*G_b' + n*G_a') / (m+n) \quad \cdots (23)$$

[0298]

That is, in Formula (23), the gain information G_d' of the point D' is obtained by the interpolation processing based on the gain information G_a' and the gain information G_b'.

[0299]

Next, the interpolation processing is performed on the basis of the internal division ratio (o, p) of the line segment C'D' from the object position C' to the point D' by the object position F' and the gain information G_c' of the object position C' and the gain information G_d' of the point D', and the gain information G_f' of the object position F' is obtained. That is, the gain information G_f' is obtained by performing the calculation of Formula (24) described below.

[0300]

[Math. 24]

$$G_f' = (o*G_c' + p*G_d') / (o+p)$$
where
$$o = SQRT((x_d' - x_f')^2 + (y_d' - y_f')^2 + (z_d' - z_f')^2)$$
$$p = SQRT((x_c' - x_f')^2 + (y_c' - y_f')^2 + (z_c' - z_f')^2) \quad \cdots (24)$$

[0301]

The gain information G_f' thus obtained is output from the object position calculation unit 48 as the gain information of the object corresponding to the listening position F.

[0302]

By performing the three-point interpolation as described above, the object absolute coordinate position information and the gain information can be obtained for an arbitrary listening position.

[0303]

Meanwhile, in a case where the three-point interpolation is performed, when there are four or more reference viewpoints in the common absolute coordinate space, a plurality of triangle meshes can be configured by a combination of selected three of the reference viewpoints.

[0304]

For example, as illustrated on the left side of Fig. 16, it is assumed that there are reference viewpoints at five positions P91 to P95.

[0305]

In such a case, a plurality of triangle meshes such as triangle meshes MS11 to MS13 is formed (configured).

[0306]

Here, the triangle mesh MS11 is formed by positions

P91 to P93, which are reference viewpoints, the triangle mesh MS12 is formed by positions P92, P93, and P95, and the triangle mesh MS13 is formed by positions P93, P94, and P95.

[0307]

The listener can freely move in a region surrounded by the triangle meshes MS11 to MS13, that is, a region surrounded by all the reference viewpoints.

[0308]

Therefore, along with the movement of the listener, that is, the movement (change) of the listening position, the triangle mesh for obtaining the object absolute coordinate position information and the gain information at the listening position is switched.

[0309]

Note that, hereinafter, the viewpoint-side triangle mesh for obtaining the object absolute coordinate position information and the gain information at the listening position is also referred to as a selected triangle mesh. Furthermore, the object-side triangle mesh corresponding to the viewpoint-side selected triangle mesh is also appropriately referred to as a selected triangle mesh.

[0310]

The left side of Fig. 16 illustrates an example in which the listening position that was originally at the position P96 has moved thereafter to a position P96'. That is, the position P96 is the position (listening position) of the viewpoint of the listener before the movement, and the position P96' is the position of the viewpoint of the listener after the movement.

[0311]

In a case where a triangle mesh for which the three-point interpolation is performed is selected, basically, a sum (total) of distances from the listening position to the respective vertices of the triangle mesh is obtained as a total distance, and a triangle mesh having the smallest total distance among the triangle meshes including the listening position is selected as the selected triangle mesh.

[0312]

That is, basically, the selected triangle mesh is determined by condition processing of selecting the triangle mesh having the smallest total distance from the triangle meshes including the listening position. Hereinafter, the condition that the total distance is the smallest among the triangle meshes including the listening position is also particularly referred to as a viewpoint-side selection condition.

[0313]

When the three-point interpolation is performed, basically, a triangle mesh satisfying such viewpoint-side selection condition is selected as the selected triangle mesh.

[0314]

Thus, in the example illustrated on the left side of Fig. 16, when the listening position is at the position P96, the triangle mesh MS11 is selected as the selected triangle mesh, and when the listening position moves to the position P96', the triangle mesh MS13 is selected as the selected triangle mesh.

[0315]

However, when a triangle mesh having the smallest total distance is simply selected as the selected triangle mesh, discontinuous transition of the object position, that is, jump of the position of the object may occur.

[0316]

For example, as illustrated in the center of Fig. 16, it is assumed that there are triangle meshes MS21 to MS23 as object-side triangle meshes, that is, triangle meshes including object positions corresponding to each reference viewpoint.

[0317]

In this example, the triangle mesh MS21 and the triangle mesh MS22 are adjacent to each other, and the triangle mesh MS22 and the triangle mesh MS23 are also adjacent to each other.

[0318]

That is, the triangle mesh MS21 and the triangle mesh MS22 have a side common to each other, and the triangle mesh MS22 and the triangle mesh MS23 also have a side common to each other. Hereinafter, a common side of two adjacent triangle meshes is also particularly referred to as a common side.

[0319]

On the other hand, since the triangle mesh MS21 and the triangle mesh MS23 are not adjacent to each other, the two triangle meshes do not have a common side.

[0320]

Here, it is assumed that the triangle mesh MS21 is an object-side triangle mesh corresponding to the viewpoint-side triangle mesh MS11. That is, it is

assumed that a triangle mesh having each of the object positions of the same object as a vertex when the viewpoint (listening position) is at each of the positions P91 to P93, which are the reference viewpoints, is the triangle mesh MS21.

[0321]

Similarly, the triangle mesh MS22 is an object-side triangle mesh corresponding to the viewpoint-side triangle mesh MS12, and the triangle mesh MS23 is an object-side triangle mesh corresponding to the viewpoint-side triangle mesh MS13.

[0322]

For example, it is assumed that the listening position moves from the position P96 to the position P96', so that the viewpoint-side selected triangle mesh is switched from the triangle mesh MS11 to the triangle mesh MS13. In this case, on the object side, the selected triangle mesh is switched from the triangle mesh MS21 to the triangle mesh MS23.

[0323]

In the center example in the drawing, a position P101 indicates an object position when the listening position is at the position P96, the object position being obtained by performing the three-point interpolation using the triangle mesh MS21 as the selected triangle mesh. Similarly, a position P101' indicates an object position when the listening position is at the position P96', the object position being obtained by performing the three-point interpolation using the triangle mesh MS23 as the selected triangle mesh.

[0324]

Therefore, in this example, when the listening position moves from the position P96 to the position P96', the object position moves from the position P101 to the position P101'.

[0325]

However, in this case, the triangle mesh MS21 including the position P101 and the triangle mesh MS23 including the position P101' are not adjacent to each other and do not have a common side common to each other. In other words, the object position moves (transitions) across the triangle mesh MS22 present between the triangle meshes.

[0326]

Therefore, in such a case, discontinuous movement (transition) of the object position occurs. This is because the triangle mesh MS21 and the triangle mesh MS23 do not have a common side, and thus the scale (measure) of the relationship of the object positions corresponding to the respective reference viewpoints is different between the triangle meshes.

[0327]

On the other hand, when the object-side selected triangle meshes have a common side, before and after the movement of the listening position, the continuity of the scale is maintained between the selected triangle meshes before and after the movement, and the occurrence of the discontinuous transition of the object position can be suppressed.

[0328]

Therefore, in a case where the three-point

interpolation is performed, it is sufficient if not only the above-described basic condition processing, but also condition processing of selecting the viewpoint-side selected triangle mesh after the movement so that the object-side selected triangle meshes have a common side before and after the movement of the listening position is added.

[0329]

In other words, it is sufficient if the selected triangle mesh to be used for the three-point interpolation at the viewpoint after the movement is selected on the basis of the relationship between the object-side selected triangle mesh used for the three-point interpolation at the viewpoint (listening position) before the movement and the object-side triangle mesh corresponding to the viewpoint-side triangle mesh including the viewpoint position (listening position) after the movement.

[0330]

Hereinafter, the condition that the object-side triangle mesh before the movement of the listening position and the object-side triangle mesh after the movement of the listening position have a common side is also particularly referred to as an object-side selection condition.

[0331]

In a case where the three-point interpolation is performed, it is sufficient if, among the viewpoint-side triangle meshes that satisfy the object-side selection condition, a triangle mesh that further satisfies the viewpoint-side selection condition is selected as the

selected triangle mesh. However, in a case where there is no viewpoint-side triangle mesh that satisfies the object-side selection condition, a triangle mesh that only satisfies the viewpoint-side selection condition is selected as the selected triangle mesh.

[0332]

As described above, when the viewpoint-side selected triangle mesh is selected so as to satisfy not only the viewpoint-side selection condition but also the object-side selection condition, it is possible to suppress the occurrence of discontinuous movement of the object position and realize higher quality acoustic reproduction.

[0333]

In this case, for example, in the example illustrated on the left side of Fig. 16, when the listening position moves from the position P96 to the position P96', the triangle mesh MS12 is selected as the viewpoint-side selected triangle mesh with respect to the position P96', which is the listening position after the movement.

[0334]

For example, as illustrated on the right side of Fig. 16, the object-side triangle mesh MS21 corresponding to the viewpoint-side triangle mesh MS11 before the movement and the object-side triangle mesh MS22 corresponding to the viewpoint-side triangle mesh MS12 after the movement have a common side. Therefore, in this case, it can be seen that the object-side selection condition is satisfied.

[0335]

Furthermore, a position P101'' indicates an object position when the listening position is at the position P96', the object position being obtained by performing the three-point interpolation using the triangle mesh MS22 as the object-side selected triangle mesh.

[0336]

Therefore, in this example, when the listening position moves from the position P96 to the position P96', the position of the object corresponding to the listening position also moves from the position P101 to the position P101''.

[0337]

In this case, since the triangle mesh MS21 and the triangle mesh MS22 have a common side, discontinuous movement of the object position does not occur before and after the movement of the listening position.

[0338]

For example, in this example, the positions of both ends of the common side of the triangle mesh MS21 and the triangle mesh MS22, that is, the object position corresponding to the position P92, which is the reference viewpoint, and the object position corresponding to the position P93, which is the reference viewpoint, are the same position before and after the movement of the listening position.

[0339]

As described above, in the example illustrated in Fig. 16, even when the listening position is the same position at the position P96', the object position, that is, the position where the object is projected varies depending on which of the triangle mesh MS12 and the

triangle mesh MS13 is selected as the viewpoint-side selected triangle mesh.

[0340]

Therefore, by selecting a more appropriate triangle mesh from among the triangle meshes including the listening position, it is possible to suppress the occurrence of discontinuous movement of the object position, that is, the sound image position, and to realize higher quality acoustic reproduction.

[0341]

Furthermore, by combining the three-point interpolation using a triangle mesh including three reference viewpoints surrounding the listening position and selection of the triangle mesh according to the selection condition, it is possible to realize object arrangement in consideration of the reference viewpoint for an arbitrary listening position in the common absolute coordinate space.

[0342]

Note that, also in a case where the three-point interpolation is performed, similar to the case where the two-point interpolation is performed, the interpolation processing weighted on the basis of the bias coefficient α may be appropriately performed to obtain the final object absolute coordinate position information and the gain information.

[0343]

<Configuration example of the content reproduction system>

Here, a more detailed embodiment of the content reproduction system to which the present technology

described above is applied will be described.

[0344]

Fig. 17 is a diagram illustrating a configuration example of the content reproduction system to which the present technology has been applied. Note that portions in Fig. 17 corresponding to those of Fig. 1 are designated by the same reference numerals, and description is omitted as appropriate.

[0345]

The content reproduction system illustrated in Fig. 17 includes a server 11 that distributes content and a client 12 that receives distribution of content from the server 11.

[0346]

Furthermore, the server 11 includes a configuration information recording unit 101, a configuration information sending unit 21, a recording unit 102, and a coded data sending unit 22.

[0347]

The configuration information recording unit 101 records, for example, the system configuration information illustrated in Fig. 4 prepared in advance, and supplies the recorded system configuration information to the configuration information sending unit 21. Note that a part of the recording unit 102 may be the configuration information recording unit 101.

[0348]

The recording unit 102 records, for example, coded audio data obtained by coding audio data of an object constituting content, object polar coordinate coded data of each object for each reference viewpoint, coded gain

information, and the like.

[0349]

The recording unit 102 supplies the coded audio data, the object polar coordinate coded data, the coded gain information, and the like recorded in response to a request or the like to the coded data sending unit 22.

[0350]

Furthermore, the client 12 includes a listener position information acquisition unit 41, a viewpoint selection unit 42, a communication unit 111, a decode unit 45, a position calculation unit 112, and a rendering processing unit 113.

[0351]

The communication unit 111 corresponds to the configuration information acquisition unit 43 and the coded data acquisition unit 44 illustrated in Fig. 1, and transmits and receives various data by communicating with the server 11.

[0352]

For example, the communication unit 111 transmits the viewpoint selection information supplied from the viewpoint selection unit 42 to the server 11, and receives the system configuration information and the bitstream transmitted from the server 11. That is, the communication unit 111 functions as a reference viewpoint information acquisition unit that acquires the system configuration information and the object polar coordinate coded data and the coded gain information included in the bitstream from the server 11.

[0353]

The position calculation unit 112 generates the

polar coordinate position information indicating the position of the object on the basis of the object polar coordinate position information supplied from the decode unit 45 and the system configuration information supplied from the communication unit 111, and supplies the polar coordinate position information to the rendering processing unit 113.

[0354]

Furthermore, the position calculation unit 112 performs gain adjustment on the audio data of the object supplied from the decode unit 45, and supplies the audio data after the gain adjustment to the rendering processing unit 113.

[0355]

The position calculation unit 112 includes a coordinate transformation unit 46, a coordinate axis transformation processing unit 47, an object position calculation unit 48, and a polar coordinate transformation unit 49.

[0356]

The rendering processing unit 113 performs the rendering processing such as VBAP or the like on the basis of the polar coordinate position information supplied from the polar coordinate transformation unit 49 and the audio data and generates and outputs reproduction audio data for reproducing the sound of the content.

[0357]

<Description of provision processing and reproduction audio data generation processing>

Subsequently, the operation of the content reproduction system illustrated in Fig. 17 will be

described.

[0358]

That is, the provision processing by the server 11 and the reproduction audio data generation processing by the client 12 will be described below with reference to the flowchart of Fig. 18.

[0359]

For example, when distribution of predetermined content is requested from the client 12 to the server 11, the server 11 starts the provision processing and performs the processing of step S41.

[0360]

That is, in step S41, the configuration information sending unit 21 reads the system configuration information of the requested content from the configuration information recording unit 101, and transmits the read system configuration information to the client 12. For example, the system configuration information is prepared in advance, and is transmitted to the client 12 via a network or the like immediately after the operation of the content reproduction system is started, that is, for example, immediately after the connection between the server 11 and the client 12 is established and before the coded audio data or the like is transmitted.

[0361]

Then, in step S61, the communication unit 111 of the client 12 receives the system configuration information transmitted from the server 11 and supplies the system configuration information to the viewpoint selection unit 42, the coordinate axis transformation

processing unit 47, and the object position calculation unit 48.

[0362]

Note that the timing at which the communication unit 111 acquires the system configuration information from the server 11 may be any timing as long as it is before the start of reproduction of the content.

[0363]

In step S62, the listener position information acquisition unit 41 acquires the listener position information according to an operation of the listener or the like, and supplies the listener position information to the viewpoint selection unit 42, the object position calculation unit 48, and the polar coordinate transformation unit 49.

[0364]

In step S63, the viewpoint selection unit 42 selects two or more reference viewpoints on the basis of the system configuration information supplied from the communication unit 111 and the listener position information supplied from the listener position information acquisition unit 41, and supplies viewpoint selection information indicating the selection result to the communication unit 111.

[0365]

For example, in a case where two reference viewpoints are selected for the listening position indicated by the listener position information, two reference viewpoints sandwiching the listening position are selected from among the plurality of reference viewpoints indicated by the system configuration

information. That is, the reference viewpoints are selected such that the listening position is located on a line segment connecting the selected two reference viewpoints.

[0366]

Furthermore, in a case where the three-point interpolation is performed in the object position calculation unit 48, three or more reference viewpoints around the listening position indicated by the listener position information are selected from among the plurality of reference viewpoints indicated by the system configuration information.

[0367]

In step S64, the communication unit 111 transmits the viewpoint selection information supplied from the viewpoint selection unit 42 to the server 11.

[0368]

Then, the processing of SU42 is performed in the server 11. That is, in step S42, the configuration information sending unit 21 receives the viewpoint selection information transmitted from the client 12 and supplies the viewpoint selection information to the coded data sending unit 22.

[0369]

The coded data sending unit 22 reads the object polar coordinate coded data and the coded gain information of the reference viewpoint indicated by the viewpoint selection information supplied from the configuration information sending unit 21 from the recording unit 102 for each object, and also reads the coded audio data of each object of the content.

[0370]

In step S43, the coded data sending unit 22 multiplexes the object polar coordinate coded data, the coded gain information, and the coded audio data read from the recording unit 102 to generate a bitstream.

[0371]

In step S44, the coded data sending unit 22 transmits the generated bitstream to the client 12, and the provision processing ends. Therefore, the content is distributed to the client 12.

[0372]

Furthermore, when the bitstream is transmitted, the client 12 performs the processing of step S65. That is, in step S65, the communication unit 111 receives the bitstream transmitted from the server 11 and supplies the bitstream to the decode unit 45.

[0373]

In step S66, the decode unit 45 extracts the object polar coordinate coded data, the coded gain information, and the coded audio data from the bitstream supplied from the communication unit 111 and decodes the object polar coordinate coded data, the coded gain information, and the coded audio data.

[0374]

The decode unit 45 supplies the object polar coordinate position information obtained by decoding to the coordinate transformation unit 46, supplies the gain information obtained by decoding to the object position calculation unit 48, and moreover supplies the audio data obtained by decoding to the polar coordinate transformation unit 49.

[0375]

In step S67, the coordinate transformation unit 46 performs coordinate transformation on the object polar coordinate position information of each object supplied from the decode unit 45, and supplies the resultant object absolute coordinate position information to the coordinate axis transformation processing unit 47.

[0376]

For example, in step S67, for each reference viewpoint, Formula (1) described above is calculated on the basis of the object polar coordinate position information for each object, and the object absolute coordinate position information is calculated.

[0377]

In step S68, the coordinate axis transformation processing unit 47 performs coordinate axis transformation processing on the object absolute coordinate position information supplied from the coordinate transformation unit 46 on the basis of the system configuration information supplied from the communication unit 111.

[0378]

The coordinate axis transformation processing unit 47 performs coordinate axis transformation processing for each object for each reference viewpoint, and supplies the resultant object absolute coordinate position information indicating the position of the object in the common absolute coordinate system to the object position calculation unit 48. For example, in step S68, calculation similar to Formula (3) described above is performed to calculate the object absolute coordinate

position information.

[0379]

In step S69, the object position calculation unit 48 performs the interpolation processing on the basis of the system configuration information supplied from the communication unit 111, the listener position information supplied from the listener position information acquisition unit 41, the object absolute coordinate position information supplied from the coordinate axis transformation processing unit 47, and the gain information supplied from the decode unit 45.

[0380]

In step S69, the above-described two-point interpolation or three-point interpolation is performed as the interpolation processing for each object, and the final object absolute coordinate position information and the gain information are calculated.

[0381]

For example, in a case where the two-point interpolation is performed, the object position calculation unit 48 obtains the proportion ratio (m : n) by performing calculation similar to Formula (4) described above on the basis of the reference viewpoint position information included in the system configuration information and the listener position information.

[0382]

Then, the object position calculation unit 48 performs the interpolation processing of the two-point interpolation by performing calculation similar to Formula (5) described above on the basis of the obtained proportion ratio (m : n) and the object absolute

coordinate position information and the gain information of the two reference viewpoints.

[0383]

Note that by performing calculation similar to Formula (6) or (7) instead of Formula (5), the interpolation processing (two-point interpolation) may be performed by weighting the object absolute coordinate position information and the gain information of a desired reference viewpoint.

[0384]

Furthermore, for example, in a case where the three-point interpolation is performed, the object position calculation unit 48 selects three reference viewpoints for forming (configuring) a triangle mesh that satisfies viewpoint-side and object-side selection conditions on the basis of the listener position information, the system configuration information, and the object absolute coordinate position information of each reference viewpoint. Then, the object position calculation unit 48 performs the three-point interpolation on the basis of the object absolute coordinate position information and the gain information of the selected three reference viewpoints.

[0385]

That is, the object position calculation unit 48 obtains the internal division ratio (m, n) and the internal division ratio (k, l) by performing calculation similar to Formulae (9) to (14) described above on the basis of the reference viewpoint position information included in the system configuration information and the listener position information.

[0386]

Then, the object position calculation unit 48 performs the interpolation processing of the three-point interpolation by performing calculation similar to Formulae (15) to (24) described above on the basis of the obtained internal division ratio (m, n) and internal division ratio (k, l) and the object absolute coordinate position information and the gain information of each reference viewpoint. Note that also in a case where the three-point interpolation is performed, the interpolation processing (three-point interpolation) may be performed by weighting the object absolute coordinate position information and the gain information of a desired reference viewpoint.

[0387]

When the interpolation processing is performed in this manner and the final object absolute coordinate position information and the gain information are obtained, the object position calculation unit 48 supplies the obtained object absolute coordinate position information and gain information to the polar coordinate transformation unit 49.

[0388]

In step S70, the polar coordinate transformation unit 49 performs the polar coordinate transformation on the object absolute coordinate position information supplied from the object position calculation unit 48 on the basis of the listener position information supplied from the listener position information acquisition unit 41 to generate the polar coordinate position information.

[0389]

Furthermore, the polar coordinate transformation unit 49 performs the gain adjustment on the audio data of each object supplied from the decode unit 45 on the basis of the gain information of each object supplied from the object position calculation unit 48.

[0390]

The polar coordinate transformation unit 49 supplies the polar coordinate position information obtained by the polar coordinate transformation and the audio data of each object obtained by the gain adjustment to the rendering processing unit 113.

[0391]

In step S71, the rendering processing unit 113 performs the rendering processing such as VBAP or the like on the basis of the polar coordinate position information of each object supplied from the polar coordinate transformation unit 49 and the audio data, and outputs the resultant reproduction audio data.

[0392]

For example, with a speaker or the like in the subsequent stage of the rendering processing unit 113, the sound of the content is reproduced on the basis of the reproduction audio data. When the reproduction audio data is generated and output in this manner, the reproduction audio data generation processing ends.

[0393]

Note that the rendering processing unit 113 or the polar coordinate transformation unit 49 may perform processing corresponding to the reproduction mode on the audio data of the object on the basis of the listener position information and the information indicating the

reproduction mode included in the system configuration information before the rendering processing.

[0394]

In such a case, for example, attenuation processing such as gain adjustment is performed on the audio data of the object located at a position overlapping with the listening position, or the audio data is replaced with zero data and muted. Furthermore, for example, the sound of the audio data of the object located at a position overlapping with the listening position is output from all channels (speakers).

[0395]

Furthermore, the provision processing and the reproduction audio data generation processing described above are performed for each frame of content.

[0396]

However, the processing in steps S41 and S61 can be performed only at the start of reproduction of the content. Moreover, the processing of step S42 and steps S62 to S64 is not necessarily performed for each frame.

[0397]

As described above, the server 11 receives the viewpoint selection information, generates the bitstream including the information of the reference viewpoint corresponding to the viewpoint selection information, and transmits the bitstream to the client 12. Furthermore, the client 12 performs the interpolation processing on the basis of the information of each reference viewpoint included in the received bitstream, and obtains the object absolute coordinate position information and the gain information of each object.

[0398]

In this way, it is possible to realize the object arrangement based on the intention of the content creator according to the listening position instead of the simple physical relationship between the listener and the object. Therefore, content reproduction based on the intention of the content creator can be realized, and the interest of the content can be sufficiently conveyed to the listener.

[0399]

<Description of the viewpoint selection processing>

Furthermore, as described above, in the reproduction audio data generation processing described with reference to Fig. 18, in a case where the three-point interpolation is performed in step S69, three reference viewpoints for performing the three-point interpolation are selected.

[0400]

Hereinafter, the viewpoint selection processing that is processing in which the client 12 selects three reference viewpoints in a case where the three-point interpolation is performed will be described with reference to the flowchart of Fig. 19. This viewpoint selection processing corresponds to the processing of step S69 of Fig. 18.

[0401]

In step S101, the object position calculation unit 48 calculates the distance from the listening position to each of the plurality of reference viewpoints on the basis of the listener position information supplied from the listener position information acquisition unit 41 and

the system configuration information supplied from the communication unit 111.

[0402]

In step S102, the object position calculation unit 48 determines whether or not the frame (hereinafter, also referred to as a current frame) of the audio data for which the three-point interpolation is to be performed is the first frame of the content.

[0403]

In a case where it is determined in step S102 that the frame is the first frame, the processing proceeds to step S103.

[0404]

In step S103, the object position calculation unit 48 selects a triangle mesh having the smallest total distance from among triangle meshes including arbitrary three reference viewpoints among the plurality of reference viewpoints. Here, the total distance is the sum of distances from the listening position to the reference viewpoints constituting the triangle mesh.

[0405]

In step S104, the object position calculation unit 48 determines whether or not the listening position is within (included in) the triangle mesh selected in step S103.

[0406]

In a case where it is determined in step S104 that the listening position is not in the triangle mesh, since the triangle mesh does not satisfy the viewpoint-side selection condition, thereafter, the processing proceeds to step S105.

[0407]

In step S105, the object position calculation unit 48 selects a triangle mesh having the smallest total distance from among the viewpoint-side triangle meshes that have not yet been selected in the processing of steps S103 and S105 that have been performed so far for the frame to be processed.

[0408]

When a new viewpoint-side triangle mesh is selected in step S105, thereafter, the processing returns to step S104, and the above-described processing is repeatedly performed until it is determined that the listening position is within the triangle mesh. That is, a triangle mesh satisfying the viewpoint-side selection condition is searched.

[0409]

On the other hand, in a case where it is determined in step S104 that the listening position is within the triangle mesh, the triangle mesh is selected as a triangle mesh for which the three-point interpolation is performed, and thereafter, the processing proceeds to step S110.

[0410]

Furthermore, in a case where it is determined in step S102 that the frame is not the first frame, thereafter, the processing of step S106 is performed.

[0411]

In step S106, the object position calculation unit 48 determines whether or not the current listening position is in the viewpoint-side triangle mesh selected in the frame (hereinafter, also referred to as a previous

frame) immediately before the current frame.

[0412]

In a case where it is determined in step S106 that the listening position is within the triangle mesh, thereafter, the processing proceeds to step S107.

[0413]

In step S107, the object position calculation unit 48 selects the same viewpoint-side triangle mesh, which has been selected for the three-point interpolation in the previous frame, as the triangle mesh for which the three-point interpolation is performed also in the current frame. When the triangle mesh for the three-point interpolation, that is, the three reference viewpoints are selected in this manner, thereafter, the processing proceeds to step S110.

[0414]

Furthermore, in a case where it is determined in step S106 that the listening position is not in the viewpoint-side triangle mesh selected in the previous frame, thereafter, the processing proceeds to step S108.

[0415]

In step S108, the object position calculation unit 48 determines whether or not there is a triangle mesh having (including) a common side with the object-side selected triangle mesh in the previous frame among the object-side triangle meshes in the current frame. The determination processing in step S108 is performed on the basis of the system configuration information and the object absolute coordinate position information.

[0416]

In a case where it is determined in step S108 that

there is no triangle mesh having a common side, since there is no triangle mesh satisfying the object-side selection condition, thereafter, the processing proceeds to step S103. In this case, the triangle mesh satisfying only the viewpoint-side selection condition is selected for the three-point interpolation in the current frame.

[0417]

Furthermore, in a case where it is determined in step S108 that there is a triangle mesh having a common side, thereafter, the processing proceeds to step S109.

[0418]

In step S109, the object position calculation unit 48 selects a triangle mesh including the listening position and having the smallest total distance as the triangle mesh for the three-point interpolation from among the viewpoint-side triangle meshes of the current frame corresponding to the object-side triangle meshes having a common side in step S108. In this case, the triangle mesh satisfying the object-side selection condition and the viewpoint-side selection condition is selected. When the triangle mesh for the three-point interpolation is selected in this manner, thereafter, the processing proceeds to step S110.

[0419]

When it is determined in step S104 that the listening position is within the triangle mesh, the processing of step S107 is performed, or the processing of step S109 is performed, thereafter, the processing of step S110 is performed.

[0420]

In step S110, the object position calculation unit

48 performs the three-point interpolation on the basis of the object absolute coordinate position information and the gain information of the triangle mesh selected for the three-point interpolation, that is, the selected three reference viewpoints, and generates the final object absolute coordinate position information and the gain information. The object position calculation unit 48 supplies the final object absolute coordinate position information and the gain information thus obtained to the polar coordinate transformation unit 49.

[0421]

In step S111, the object position calculation unit 48 determines whether or not there is a next frame to be processed, that is, whether or not the reproduction of the content has ended.

[0422]

In a case where it is determined in step S111 that there is a next frame, since the reproduction of the content has not yet been ended, the processing returns to step S101, and the above-described processing is repeated.

[0423]

On the other hand, in a case where it is determined in step S111 that there is no next frame, the reproduction of the content has ended, and the viewpoint selection processing also ends.

[0424]

As described above, the client 12 selects an appropriate triangle mesh on the basis of the viewpoint-side and object-side selection conditions, and performs the three-point interpolation. In this way, it is

possible to suppress the occurrence of discontinuous movement of the object position and to realize higher quality acoustic reproduction.

[0425]

According to the present technology described above, it is possible to realize reproduction at each reference viewpoint according to the intention of the content creator, instead of reproduction using a physical positional relationship with respect to a conventional fixed object arrangement in the movement of the listener in a free viewpoint space.

[0426]

Furthermore, at an arbitrary listening position sandwiched between a plurality of reference viewpoints, the object position and the gain suitable for the arbitrary listening position can be generated by performing the interpolation processing on the basis of the object arrangement of the plurality of reference viewpoints. Therefore, the listener can move seamlessly between the reference viewpoints.

[0427]

Moreover, in a case where the reference viewpoint overlaps the object position, it is possible to give the listener a feeling as if the listener became the object by lowering or muting the signal level of the object. Therefore, for example, a karaoke mode, a minus one performance mode, or the like can be realized, and a feeling that the listener itself joins in the content can be obtained.

[0428]

In addition, in the interpolation processing of the

reference viewpoint, in a case where there is a reference viewpoint to which the listener wants to bring closer, the sense of movement is weighted by applying the bias coefficient α, so that the content can be reproduced with the object arrangement brought closer to the viewpoint that the listener prefers even when the listener moves.

[0429]

Furthermore, in a case where there are four or more reference viewpoints, a triangle mesh can be configured by three reference viewpoints, and the three-point interpolation can be performed. In this case, since a plurality of triangle meshes can be configured, even when the listener freely moves in a region including the triangle meshes, that is, a region surrounded by all reference viewpoints, it is possible to realize content reproduction at an appropriate object position having an arbitrary position in the region as the listening position.

[0430]

Moreover, according to the present technology, in a case of using transmission in a polar coordinate system, it is possible to realize audio reproduction of a free viewpoint space reflecting an intention of a content creator only by adding system configuration information to a conventional MPEG-H coding system.

[0431]

<Configuration example of computer>

Incidentally, the series of processing described above can be executed by hardware and it can also be executed by software. In a case where the series of processing is executed by software, a program

constituting the software is installed in a computer. Here, the computer includes a computer mounted in dedicated hardware, for example, a general-purpose a personal computer that can execute various functions by installing the various programs, or the like.

[0432]

Fig. 20 is a block diagram illustrating a configuration example of hardware of a computer in which the series of processing described above is executed by a program.

[0433]

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

[0434]

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

[0435]

The input unit 506 includes a keyboard, a mouse, a microphone, an image sensor, and the like. The output unit 507 includes a display, a speaker, and the like. The recording unit 508 includes a hard disk, a non-volatile 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 disk, a magneto-optical disk, or a semiconductor memory.

[0436]

In the computer configured in the manner described

above, the series of processing described above is performed, for example, such that the CPU 501 loads a program stored in the recording unit 508 into the RAM 503 via the input/output interface 505 and the bus 504 and executes the program.

[0437]

The program to be executed by the computer (CPU 501) can be provided by being recorded on the removable recording medium 511, for example, as a package medium or the like. Furthermore, the program can be provided via a wired or wireless transmission medium such as a local area network, the Internet, or digital satellite broadcasting.

[0438]

In the computer, the program can be installed on the recording unit 508 via the input/output interface 505 when the removable recording medium 511 is mounted on the drive 510. Furthermore, the program can be received by the communication unit 509 via a wired or wireless transmission medium and installed on the recording unit 508. In addition, the program can be pre-installed on the ROM 502 or the recording unit 508.

[0439]

Note that the program executed by the computer may be a program that is processed in chronological order along the order described in the present description or may be a program that is processed in parallel or at a required timing, e.g., when call is carried out.

[0440]

Furthermore, the embodiment of the present technology is not limited to the aforementioned

embodiments, but various changes may be made within the scope not departing from the gist of the present technology.

[0441]

For example, the present technology can adopt a configuration of cloud computing in which one function is shared and jointly processed by a plurality of apparatuses via a network.

[0442]

Furthermore, each step described in the above-described flowcharts can be executed by a single apparatus or shared and executed by a plurality of apparatuses.

[0443]

Moreover, in a case where a single step includes a plurality of pieces of processing, the plurality of pieces of processing included in the single step can be executed by a single apparatus or can be shared and executed by a plurality of apparatuses.

[0444]

Moreover, the present technology may be configured as below.

[0445]

(1)

An information processing apparatus including:

a listener position information acquisition unit that acquires listener position information of a viewpoint of a listener;

a reference viewpoint information acquisition unit that acquires position information of a first reference viewpoint and object position information of an object at

the first reference viewpoint, and position information of a second reference viewpoint and object position information of the object at the second reference viewpoint; and

an object position calculation unit that calculates position information of the object at the viewpoint of the listener on the basis of the listener position information, the position information of the first reference viewpoint and the object position information at the first reference viewpoint, and the position information of the second reference viewpoint and the object position information at the second reference viewpoint.

(2)

The information processing apparatus according to (1), in which

the first reference viewpoint and the second reference viewpoint are viewpoints set in advance by a content creator.

(3)

The information processing apparatus according to (1) or (2), in which

the first reference viewpoint and the second reference viewpoint are viewpoints selected on the basis of the listener position information.

(4)

The information processing apparatus according to any one of (1) to (3), in which

the object position information is information indicating a position expressed by polar coordinates or absolute coordinates, and

the reference viewpoint information acquisition unit acquires gain information of the object at the first reference viewpoint and gain information of the object at the second reference viewpoint.

(5)

The information processing apparatus according to (4), in which

the object position calculation unit calculates the position information of the object at the viewpoint of the listener by interpolation processing on the basis of the listener position information, the position information of the first reference viewpoint and the object position information at the first reference viewpoint, and the position information of the second reference viewpoint and the object position information at the second reference viewpoint.

(6)

The information processing apparatus according to (4) or (5), in which

the object position calculation unit calculates gain information of the object at the viewpoint of the listener by interpolation processing on the basis of the listener position information, the position information of the first reference viewpoint and the gain information at the first reference viewpoint, and the position information of the second reference viewpoint and the gain information at the second reference viewpoint.

(7)

The information processing apparatus according to (5) or (6), in which

the object position calculation unit calculates the

position information or gain information of the object at the viewpoint of the listener by performing interpolation processing by weighting the object position information or the gain information at the first reference viewpoint.

(8)

The information processing apparatus according to any one of (1) to (4), in which

the reference viewpoint information acquisition unit acquires the position information of the reference viewpoint and the object position information at the reference viewpoint for a plurality of, three or more, reference viewpoints including the first reference viewpoint and the second reference viewpoint, and

the object position calculation unit calculates the position information of the object at the viewpoint of the listener by interpolation processing on the basis of the listener position information, the position information of each of the three reference viewpoints among the plurality of the reference viewpoints, and the object position information at each of the three reference viewpoints.

(9)

The information processing apparatus according to (8), in which

the object position calculation unit calculates the gain information of the object at the viewpoint of the listener by interpolation processing on the basis of the listener position information, the position information of each of the three reference viewpoints, and gain information at each of the three reference viewpoints.

(10)

The information processing apparatus according to (9), in which

the object position calculation unit calculates the position information or gain information of the object at the viewpoint of the listener by performing interpolation processing by weighting the object position information or the gain information at a predetermined reference viewpoint among the three reference viewpoints.

(11)

The information processing apparatus according to any one of (8) to (10), in which

the object position calculation unit sets a region formed by arbitrary three reference viewpoints as a triangle mesh, and selects three reference viewpoints forming a triangle mesh satisfying a predetermined condition among a plurality of the triangular meshes as the three reference viewpoints to be used for interpolation processing.

(12)

The information processing apparatus according to (11), in which

in a case where the viewpoint of the listener moves, the object position calculation unit

sets a region formed by each of positions of the object indicated by each of the object position information at the three reference viewpoints forming the triangle mesh as an object triangle mesh, and

selects the three reference viewpoints to be used for interpolation processing at the viewpoint after movement of the listener on the basis of a relationship between the object triangle mesh corresponding to the

triangle mesh formed by the three reference viewpoints used for interpolation processing at the viewpoint before movement of the listener and the object triangle mesh corresponding to the triangle mesh including the viewpoint after movement of the listener.

(13)

The information processing apparatus according to (12), in which

the object position calculation unit uses, for interpolation processing at the viewpoint after movement of the listener, three reference viewpoints forming the triangle mesh including the viewpoint after movement of the listener corresponding to the object triangle mesh having a side common to the object triangle mesh corresponding to the triangle mesh formed by the three reference viewpoints used for interpolation processing at the viewpoint before movement of the listener.

(14)

The information processing apparatus according to any one of (1) to (13), in which

the object position calculation unit calculates the position information of the object at the viewpoint of the listener on the basis of the listener position information, the position information of the first reference viewpoint, the object position information at the first reference viewpoint, a listener direction information indicating a direction of a face of the listener set at the first reference viewpoint, the position information of the second reference viewpoint, the object position information at the second reference viewpoint, and the listener direction information at the

second reference viewpoint.

(15)

The information processing apparatus according to (14), in which

the reference viewpoint information acquisition unit acquires configuration information including the position information and the listener direction information of each of a plurality of reference viewpoints including the first reference viewpoint and the second reference viewpoint.

(16)

The information processing apparatus according to (15), in which

the configuration information includes information indicating a number of the plurality of the reference viewpoints and information indicating a number of the objects.

(17)

An information processing method including, by an information processing apparatus:

acquiring listener position information of a viewpoint of a listener;

acquiring position information of a first reference viewpoint and object position information of an object at the first reference viewpoint, and position information of a second reference viewpoint and object position information of the object at the second reference viewpoint; and

calculating position information of the object at the viewpoint of the listener on the basis of the listener position information, the position information

of the first reference viewpoint and the object position information at the first reference viewpoint, and the position information of the second reference viewpoint and the object position information at the second reference viewpoint.

(18)

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

acquiring listener position information of a viewpoint of a listener;

acquiring position information of a first reference viewpoint and object position information of an object at the first reference viewpoint, and position information of a second reference viewpoint and object position information of the object at the second reference viewpoint; and

calculating position information of the object at the viewpoint of the listener on the basis of the listener position information, the position information of the first reference viewpoint and the object position information at the first reference viewpoint, and the position information of the second reference viewpoint and the object position information at the second reference viewpoint.

REFERENCE SIGNS LIST

[0446]

11 Server

12 Client

21 Configuration information sending unit

22 Coded data sending unit

41 Listener position information acquisition unit
42 Viewpoint selection unit
44 Coded data acquisition unit
46 Coordinate transformation unit
47 Coordinate axis transformation processing unit
48 Object position calculation unit
49 Polar coordinate transformation unit
111 Communication unit
112 Position calculation unit
113 Rendering processing unit

CLAIMS

1. An information processing apparatus comprising:

a listener position information acquisition unit that acquires listener position information of a viewpoint of a listener;

a reference viewpoint information acquisition unit that acquires position information of a first reference viewpoint and object position information of an object at the first reference viewpoint, and position information of a second reference viewpoint and object position information of the object at the second reference viewpoint; and

an object position calculation unit that calculates position information of the object at the viewpoint of the listener on a basis of the listener position information, the position information of the first reference viewpoint and the object position information at the first reference viewpoint, and the position information of the second reference viewpoint and the object position information at the second reference viewpoint.

2. The information processing apparatus according to claim 1, wherein

the first reference viewpoint and the second reference viewpoint are viewpoints set in advance by a content creator.

3. The information processing apparatus according to claim 1, wherein

the first reference viewpoint and the second reference viewpoint are viewpoints selected on a basis of the listener position information.

4. The information processing apparatus according to claim 1, wherein

the object position information is information indicating a position expressed by polar coordinates or absolute coordinates, and

the reference viewpoint information acquisition unit acquires gain information of the object at the first reference viewpoint and gain information of the object at the second reference viewpoint.

5. The information processing apparatus according to claim 4, wherein

the object position calculation unit calculates the position information of the object at the viewpoint of the listener by interpolation processing on a basis of the listener position information, the position information of the first reference viewpoint and the object position information at the first reference viewpoint, and the position information of the second reference viewpoint and the object position information at the second reference viewpoint.

6. The information processing apparatus according to claim 4, wherein

the object position calculation unit calculates gain information of the object at the viewpoint of the listener by interpolation processing on a basis of the

listener position information, the position information of the first reference viewpoint and the gain information at the first reference viewpoint, and the position information of the second reference viewpoint and the gain information at the second reference viewpoint.

7. The information processing apparatus according to claim 5, wherein

the object position calculation unit calculates the position information or gain information of the object at the viewpoint of the listener by performing interpolation processing by weighting the object position information or the gain information at the first reference viewpoint.

8. The information processing apparatus according to claim 1, wherein

the reference viewpoint information acquisition unit acquires the position information of the reference viewpoint and the object position information at the reference viewpoint for a plurality of, three or more, reference viewpoints including the first reference viewpoint and the second reference viewpoint, and

the object position calculation unit calculates the position information of the object at the viewpoint of the listener by interpolation processing on a basis of the listener position information, the position information of each of the three reference viewpoints among the plurality of the reference viewpoints, and the object position information at each of the three reference viewpoints.

9. The information processing apparatus according to claim 8, wherein

the object position calculation unit calculates the gain information of the object at the viewpoint of the listener by interpolation processing on a basis of the listener position information, the position information of each of the three reference viewpoints, and gain information at each of the three reference viewpoints.

10. The information processing apparatus according to claim 9, wherein

the object position calculation unit calculates the position information or gain information of the object at the viewpoint of the listener by performing interpolation processing by weighting the object position information or the gain information at a predetermined reference viewpoint among the three reference viewpoints.

11. The information processing apparatus according to claim 8, wherein

the object position calculation unit sets a region formed by arbitrary three reference viewpoints as a triangle mesh, and selects three reference viewpoints forming a triangle mesh satisfying a predetermined condition among a plurality of the triangular meshes as the three reference viewpoints to be used for interpolation processing.

12. The information processing apparatus according to claim 11, wherein

in a case where the viewpoint of the listener

moves, the object position calculation unit

sets a region formed by each of positions of the object indicated by each of the object position information at the three reference viewpoints forming the triangle mesh as an object triangle mesh, and

selects the three reference viewpoints to be used for interpolation processing at the viewpoint after movement of the listener on a basis of a relationship between the object triangle mesh corresponding to the triangle mesh formed by the three reference viewpoints used for interpolation processing at the viewpoint before movement of the listener and the object triangle mesh corresponding to the triangle mesh including the viewpoint after movement of the listener.

13. The information processing apparatus according to claim 12, wherein

the object position calculation unit uses, for interpolation processing at the viewpoint after movement of the listener, three reference viewpoints forming the triangle mesh including the viewpoint after movement of the listener corresponding to the object triangle mesh having a side common to the object triangle mesh corresponding to the triangle mesh formed by the three reference viewpoints used for interpolation processing at the viewpoint before movement of the listener.

14. The information processing apparatus according to claim 1, wherein

the object position calculation unit calculates the position information of the object at the viewpoint of

the listener on a basis of the listener position information, the position information of the first reference viewpoint, the object position information at the first reference viewpoint, a listener direction information indicating a direction of a face of the listener set at the first reference viewpoint, the position information of the second reference viewpoint, the object position information at the second reference viewpoint, and the listener direction information at the second reference viewpoint.

15. The information processing apparatus according to claim 14, wherein

the reference viewpoint information acquisition unit acquires configuration information including the position information and the listener direction information of each of a plurality of reference viewpoints including the first reference viewpoint and the second reference viewpoint.

16. The information processing apparatus according to claim 15, wherein

the configuration information includes information indicating a number of the plurality of the reference viewpoints and information indicating a number of the objects.

17. An information processing method comprising, by an information processing apparatus:

acquiring listener position information of a viewpoint of a listener;

acquiring position information of a first reference viewpoint and object position information of an object at the first reference viewpoint, and position information of a second reference viewpoint and object position information of the object at the second reference viewpoint; and

calculating position information of the object at the viewpoint of the listener on a basis of the listener position information, the position information of the first reference viewpoint and the object position information at the first reference viewpoint, and the position information of the second reference viewpoint and the object position information at the second reference viewpoint.

18. A program causing a computer to execute processing comprising the steps of:

acquiring listener position information of a viewpoint of a listener;

acquiring position information of a first reference viewpoint and object position information of an object at the first reference viewpoint, and position information of a second reference viewpoint and object position information of the object at the second reference viewpoint; and

calculating position information of the object at the viewpoint of the listener on a basis of the listener position information, the position information of the first reference viewpoint and the object position information at the first reference viewpoint, and the position information of the second reference viewpoint

and the object position information at the second reference viewpoint.s