

[DESCRIPTION]

[Title]

INFORMATION PROCESSING DEVICE AND METHOD, AND PROGRAM

[Technical Field]

[0001]

The present technique relates to an information processing device and method and a program and particularly relates to an information processing device and method and a program that can reproduce contents on the basis of the intention of a contents producer.

[Background Art]

[0002]

For example, in a free viewpoint space, it is assumed that each object to be disposed in the space by using an absolute coordinate system is placed in a fixed layout (for example, see PTL 1).

[0003]

In this case, the direction of each object viewed from any listening position is uniquely determined on the basis of the relationship between the coordinate position and the face orientation of a listener and an object in an absolute space, the gain of each object is uniquely determined on the basis of a distance from the listening position, and the sound of each object is reproduced.

[Citation List]

[Patent Literature]

[0004]

[PTL 1]

WO 2019/198540

[Summary]

[Technical Problem]

[0005]

The artistic quality of contents and points to be emphasized to listeners are also significant.

[0006]

In some cases, objects desirably move to the front, for example, musical instruments or players to be emphasized in the contents of music or athletes to be emphasized in the contents of sports.

[0007]

In some cases, each object is supposed to be disposed in a different layout, for example, an object is supposed to be always disposed at a fixed position with respect to a listener.

[0008]

In light of such circumstances, the attraction of contents may be insufficiently found out in the simply physical relationship between a listener and an object.

[0009]

The present technique has been made in view of such circumstances and is designed to reproduce contents on the basis of the intention of a contents producer.

[Solution to Problem]

[0010]

An information processing device according to an aspect of the present technique includes: a listener-position information acquisition unit that acquires listener position information on the viewpoint of a listener; a reference-viewpoint information acquisition unit that acquires position information on a first reference viewpoint, object position information on a first object at the first reference viewpoint, position information on a second reference viewpoint, object position information on the first object at the second reference viewpoint, and object position information on a second object; and an object position calculation unit that calculates position information on the first object at the viewpoint of the listener on the basis of the listener position information, the position information on the first reference viewpoint, the object position information on the first object at the first reference viewpoint, the position information on the second reference viewpoint, and the object position information on the first object at the second reference viewpoint.

[0011]

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

[0012]

In an aspect of the present technique, listener position information on the viewpoint of a listener is acquired; position information on a first reference viewpoint, object position information on a first object at the first reference viewpoint, position information on a second reference viewpoint, and object position information on the first object at the second reference viewpoint are

acquired, object position information on a second object is acquired; and position information on the first object at the viewpoint of the listener is calculated on the basis of the listener position information, the position information on the first reference viewpoint, the object position information on the first object at the first reference viewpoint, the position information on the second reference viewpoint, and the object position information on the first object at the second reference viewpoint.

[Brief Description of Drawings]

[0013]

[Fig. 1]

Fig. 1 is an explanatory drawing of an absolute-coordinate base interpolation object and a polar-coordinate base fixed object.

[Fig. 2]

Fig. 2 illustrates a configuration example of a content reproducing system.

[Fig. 3]

Fig. 3 is an explanatory drawing of the polar-coordinate base interpolation object.

[Fig. 4]

Fig. 4 illustrates a configuration example of the content reproducing system.

[Fig. 5]

Fig. 5 is an explanatory drawing of the absolute-coordinate base fixed object.

[Fig. 6]

Fig. 6 illustrates a configuration example of the content reproducing system.

[Fig. 7]

Fig. 7 is an explanatory drawing of the interpolation of object absolute-coordinate position information.

[Fig. 8]

Fig. 8 is an explanatory drawing of an internal ratio in a triangular mesh on the viewpoint side.

[Fig. 9]

Fig. 9 is an explanatory drawing of the calculation of an object position.

[Fig. 10]

Fig. 10 is an explanatory drawing of the calculation of object polar-coordinate position information by interpolation.

[Fig. 11]

Fig. 11 indicates an example of system configuration information.

[Fig. 12]

Fig. 12 indicates an example of a bit stream format.

[Fig. 13]

Fig. 13 indicates an example of the bit stream format.

[Fig. 14]

Fig. 14 indicates a metadata example of the polar-coordinate base fixed object.

[Fig. 15]

Fig. 15 indicates a metadata example of the polar-coordinate base interpolation object.

[Fig. 16]

Fig. 16 indicates a metadata example of the absolute-coordinate base fixed object.

[Fig. 17]

Fig. 17 illustrates a configuration example of the content reproducing system.

[Fig. 18]

Fig. 18 is a flowchart for explaining provision.

[Fig. 19]

Fig. 19 is a flowchart for explaining the generation of reproduction audio data.

[Fig. 20]

Fig. 20 is a flowchart for explaining the generation of polar-coordinate position information.

[Fig. 21]

Fig. 21 illustrates a configuration example of a computer.

[Description of Embodiments]

[0014]

Embodiments to which the present technique is applied will be described below with reference to the accompanying drawings.

[0015]

<First Embodiment>

<Present Technique>

The present technique is designed to obtain a more flexible object layout by preparing multiple types of objects having different coordinate representations such as an origin position and a coordinate form, thereby reproducing contents on the basis of the intention of a contents producer.

[0016]

For example, in a free viewpoint space where Artistic Intent is used, any object position can be generated by interpolation using absolute coordinates with respect to a plurality of reference viewpoints or polar coordinate data based on the orientation of a listener.

[0017]

In addition to such an object, if an object is to be always disposed at a fixed position with respect to a listener, an object layout cannot be expressed regardless of the listener because the orientation of the listener is considered in the foregoing technique.

[0018]

If an object is to be disposed at an absolute position in a free viewpoint space regardless of the position of a listener, the foregoing technique needs to calculate backward an absolute fixed object at each reference viewpoint as relative layout information, resulting in an undesirable situation in light of computational

complexity and accuracy.

[0019]

Thus, in the present technique, four types (kinds) of objects are available: an absolute-coordinate base interpolation object, a polar-coordinate base fixed object, a polar-coordinate base interpolation object, and an absolute-coordinate base fixed object. This can reproduce contents on the basis of the intention of a contents producer.

[0020]

For example, in free viewpoint audio using Audio Artistic Intent according to the present technique, information on object layouts at a plurality of reference viewpoints is produced in advance on the assumption that the reference viewpoints are produced by a contents producer.

[0021]

A listener can freely move to positions other than the reference viewpoints.

[0022]

If the listener is located at a position different from the reference viewpoint, interpolation is performed on the basis of position information on objects at the reference viewpoints surrounding the position of the listener, thereby calculating position information on an object corresponding to the current position of the listener.

[0023]

Thus, for example, when the listener moves from a position P11 to a position P11' in a free viewpoint space as indicated by an arrow Q11 in Fig. 1, the position of the same object moves from a position P12 to a position P12' according to the movement of the listener.

[0024]

Hence, spatial acoustics can be reproduced at a free viewpoint position on the basis of the intention of a contents producer.

[0025]

Hereinafter, such an object will be referred to as an absolute-coordinate base interpolation object.

[0026]

An absolute-coordinate base interpolation object is located at a position in a free viewpoint space with respect to each of the reference viewpoints. If a listener is located at a position different from the reference viewpoint, the position of the absolute-coordinate base interpolation object is determined by interpolation on the basis of the positions of absolute-coordinate base interpolation objects determined with respect to multiple reference viewpoints surrounding the listener.

[0027]

Thus, the position of the absolute-coordinate base interpolation object changes depending upon the position or orientation of the listener in the free viewpoint

space.

[0028]

In consideration of a guide support system that depends upon, for example, the acoustic AR (Augmented Reality) of sound or the arrival direction of sound among systems in which such an absolute-coordinate base interpolation object is used, an object needs to be fixed with respect to a listener regardless of the position and orientation of the listener in a free viewpoint space. Hereinafter, such an object will be referred to as a polar-coordinate base fixed object.

[0029]

For example, when the listener moves from a position P11 to a position P11' in a free viewpoint space as indicated by an arrow Q12 in Fig. 1, the polar-coordinate base fixed object moves from, for example, a position P13 to a position P13', accordingly. At this point, when viewed from the listener, the polar-coordinate base fixed object is always located at the same position, e.g., on the left ahead of the listener before and after the movement.

[0030]

The representation by the absolute-coordinate base interpolation object is based on mapping of an object on absolute coordinates in consideration of the orientation of the listener.

[0031]

Thus, the representation cannot deal with an object having the property of being always located at the same position with respect to the listener. In the present technique, for the purpose of fixing an object with respect to a listener, an absolute-coordinate base interpolation object and a polar-coordinate base fixed object are combined to deal with an object designed for the purpose.

[0032]

For example, a content reproducing system for dealing with an absolute-coordinate base interpolation object and a polar-coordinate base fixed object in this manner is configured as illustrated in Fig. 2.

[0033]

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

[0034]

The server 11 includes a configuration information transmission unit 21 and an encoded data transmission unit 22.

[0035]

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

[0036]

In the content reproducing system, a plurality of listening positions on a predetermined common absolute coordinate space are specified (set) as the

positions of reference viewpoints (hereinafter also referred to as reference viewpoint positions) by a contents producer.

[0037]

In this case, the contents producer specifies (sets), as reference viewpoints in advance, a position to serve as the listening position of the listener on the common absolute coordinate space during the reproduction of contents, and a desirable orientation of the face of the listener at the position, that is, a desirable viewpoint for listening to the sound of the contents.

[0038]

In the server 11, system configuration information and object polar-coordinate encoded data are prepared in advance. The system configuration is information on the reference viewpoints and the object polar-coordinate encoded data indicates the positions of absolute-coordinate base interpolation objects for the reference viewpoints.

[0039]

In this case, the object polar-coordinate encoded data on an absolute-coordinate base interpolation object at each reference viewpoint is obtained by encoding object polar-coordinate position information indicating the relative position of the absolute-coordinate base interpolation object viewed from each reference viewpoint.

[0040]

In the object polar-coordinate position information, the relative position of the absolute-coordinate base interpolation object viewed from the reference viewpoint, that is, with respect to the reference viewpoint is represented by polar coordinates. The same absolute-coordinate base interpolation object is disposed, for each reference viewpoint, at a different absolute position in the common absolute coordinate space.

[0041]

The configuration information transmission unit 21 transmits the system configuration information to the client 12 via a network or the like immediately after the start of an operation of the content reproducing system, that is, immediately after a connection to, for example, the client 12 is established. The system configuration information may be properly resent to the client 12 at a proper timing after the connection is established.

[0042]

The encoded data transmission unit 22 selects two or more of the reference viewpoints on the basis of the viewpoint selection information supplied from the configuration information transmission unit 21 and transmits object polar-coordinate encoded data on absolute-coordinate base interpolation objects at the selected reference viewpoints to the client 12 via a network or the like.

[0043]

In this case, the viewpoint selection information is information on the reference

viewpoints selected by, for example, the client 12.

[0044]

Thus, the object polar-coordinate encoded data on the absolute-coordinate base interpolation objects at the selected reference viewpoints requested from the client 12 is acquired in the encoded data transmission unit 22 and is transmitted to the client 12.

[0045]

Hereinafter, it is assumed that three reference viewpoints are selected (specified) by the viewpoint selection information.

[0046]

Furthermore, in the server 11, object polar-coordinate encoded data on a polar-coordinate base fixed object is prepared in addition to the object polar-coordinate encoded data on the absolute-coordinate base interpolation objects.

[0047]

The object polar-coordinate encoded data on the polar-coordinate base fixed object is obtained by encoding object polar-coordinate position information indicating the relative positions of absolute-coordinate base interpolation objects viewed from a listener, that is, with respect to the listener. The object polar-coordinate position information is position information represented by polar coordinates.

[0048]

The object polar-coordinate position information on the polar-coordinate base fixed object is different from the object polar-coordinate position information on the absolute-coordinate base interpolation object in that the object polar-coordinate position information is obtained with an origin at (with respect to) the position of a listener instead of (an origin at) a reference viewpoint, more specifically, with respect to the position and orientation of the listener.

[0049]

Even if the position and orientation of the listener are changed, the object polar-coordinate position information on the polar-coordinate base fixed object is not changed, so that a piece of object polar-coordinate encoded data is prepared for one polar-coordinate base fixed object.

[0050]

In the encoded data transmission unit 22, the object polar-coordinate encoded data on the polar-coordinate base fixed object is acquired and is transmitted to the client 12.

[0051]

The client 12 includes a listener-position information acquisition unit 41, a viewpoint selection unit 42, a configuration information acquisition unit 43, an encoded data acquisition unit 44, a decoding unit 45, a coordinate transformation unit 46, a coordinate-axis transformation unit 47, an object position calculation unit 48, and a polar-coordinate transformation unit 49.

[0052]

The listener-position information acquisition unit 41 acquires listener position information on the absolute position of a listener (listening position) on the common absolute coordinate space in response to an operation for specification by a user (listener) and supplies the information to the viewpoint selection unit 42, the object position calculation unit 48, and the polar-coordinate transformation unit 49.

[0053]

For example, in the listener position information, the position of the listener in the common absolute coordinate space is represented by absolute coordinates. Hereinafter, the coordinate system of absolute coordinates indicated by the listener position information will be also referred to as a common absolute coordinate system.

[0054]

The viewpoint selection unit 42 selects three reference viewpoints surrounding the listening position 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 on the selection result to the configuration information acquisition unit 43.

[0055]

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

[0056]

In the following example, the client 12 includes the viewpoint selection unit 42 that selects the reference viewpoint on the basis of the listener position information and the system configuration information. The viewpoint selection unit 42 may be provided in the server 11.

[0057]

The encoded data acquisition unit 44 receives the object polar-coordinate encoded data transmitted from the server 11 and supplies the information to the decoding unit 45. In other words, the encoded data acquisition unit 44 acquires the object polar-coordinate encoded data from the server 11.

[0058]

The decoding unit 45 decodes the object polar-coordinate encoded data supplied from the encoded data acquisition unit 44.

[0059]

The decoding unit 45 supplies the object polar-coordinate position information on the absolute-coordinate base interpolation object to the coordinate transformation

unit 46, the object polar-coordinate position information being obtained by decoding.

[0060]

Furthermore, the decoding unit 45 outputs the object polar-coordinate position information on the polar-coordinate base fixed object as polar-coordinate position information to a rendering unit, which is not illustrated, the object polar-coordinate position information being obtained by decoding.

[0061]

The coordinate transformation unit 46 performs coordinate transformation on the object polar-coordinate position information supplied from the decoding unit 45 and supplies object absolute-coordinate position information obtained by the coordinate transformation to the coordinate-axis transformation unit 47.

[0062]

In the coordinate transformation unit 46, coordinate transformation is performed to transform polar coordinates into absolute coordinates. Thus, the object polar-coordinate position information on polar coordinates indicating the position of the absolute-coordinate base interpolation object viewed from a reference viewpoint is converted into object absolute-coordinate position information on absolute coordinates indicating the position of the absolute-coordinate base interpolation object in the absolute coordinate system with the origin located at the position of the reference viewpoint.

[0063]

The coordinate-axis transformation unit 47 performs coordinate-axis transformation 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.

[0064]

In this case, the coordinate-axis transformation is processing performed using coordinate transformation (coordinate-axis transformation) and an offset shift in combination. The coordinate-axis transformation obtains the object absolute-coordinate position information on the absolute coordinates of the absolute-coordinate base interpolation object projected to the common absolute coordinate space. In other words, the object absolute-coordinate position information obtained by coordinate-axis transformation is the absolute coordinates of the common absolute coordinate system (position information represented by absolute coordinates), the absolute coordinates indicating the absolute position of the absolute-coordinate base interpolation object on the common absolute coordinate space.

[0065]

The object position calculation unit 48 performs interpolation 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 unit 47 and supplied the final object absolute-coordinate position information obtained by the interpolation to the polar-coordinate transformation unit 49. In this case, the final object absolute-coordinate position information is information on the position of the absolute-coordinate base interpolation object in the common absolute coordinate system when the viewpoint of the listener is located at the listening position indicated by the listener position information.

[0066]

In the object position calculation unit 48, the absolute position of the absolute-coordinate base interpolation object, which corresponds to the listening position, in the common absolute coordinate space, that is, the absolute coordinates of the common absolute coordinate system are calculated as the final object absolute-coordinate position information from the listening position indicated by the listener position information and the positions of three reference viewpoints indicated by the viewpoint selection information. At this point, the object position calculation unit 48 acquires the system configuration information from the configuration information acquisition unit 43 or acquires the viewpoint selection information from the viewpoint selection unit 42 as necessary.

[0067]

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 then outputs the polar-coordinate position information obtained by the transformation to the rendering unit, which is not illustrated, in the subsequent stage.

[0068]

In the polar-coordinate transformation unit 49, polar-coordinate transformation is performed to transform the object absolute-coordinate position information on the absolute coordinates of the common absolute-coordinate system into polar-coordinate position information on polar coordinates indicating the relative position of an object (absolute-coordinate base interpolation object) viewed from the listening position.

[0069]

For example, rendering such as VBAP (Vector Based Amplitude Panning) requires polar-coordinate position information on the relative position of an object with respect to the listening position, as position information on the object.

[0070]

Thus, for the absolute-coordinate base interpolation object, the object absolute-coordinate position information on the viewpoint of the listener is determined by interpolation. The object absolute-coordinate position information is converted

into polar-coordinate position information and is supplied to the rendering unit.
[0071]

In contrast, the object polar-coordinate position information on the polar-coordinate base fixed object is polar coordinates viewed from the viewpoint of the listener and thus is supplied as it is to the rendering unit without being interpolated or transformed into polar coordinates.

[0072]

The polar-coordinate base fixed object allows an object layout without depending upon the position and orientation of the listener.

[0073]

By preparing two different types of objects: the absolute-coordinate base interpolation object and the polar-coordinate base fixed object, contents can be reproduced on the basis of the intention of a contents producer.

[0074]

The polar-coordinate base fixed object is fixed regardless of the position of the listener in the free viewpoint space. However, according to the intention of a contents producer, an object fixed regardless of the orientation of the listener may be disposed at a different position for each reference viewpoint.

[0075]

In this case, each viewpoint has a different object layout with the listener located at the center as illustrated in, for example, Fig. 3. In rendering at the viewpoint, an object may be fixed with respect to the listener. In Fig. 3, parts corresponding to those of Fig. 1 are denoted by the same reference signs and a description thereof is omitted as necessary.

[0076]

For example, in the example of Fig. 3, when the listener is located at a position P11 in a free viewpoint space, an object is fixed at a position P21 regardless of the orientation of the listener.

[0077]

If the listener is located at the same position, the object viewed from the listener is always located at the same position (direction), e.g., on the left ahead of the listener regardless of the orientation of the listener.

[0078]

When the listener moves from the position P11 to a position P11', the object at the position P21 moves to a position P21', accordingly. If the listener is located at the position P11', the object viewed from the listener is always located at the same position regardless of the orientation of the listener.

[0079]

At this point, the relative position of the object viewed from the listener at the position P11 is different from that viewed from the listener at the position P11'.

[0080]

Hereinafter, such an object will be referred to as a polar-coordinate base

interpolation object.

[0081]

In the example of Fig. 3, the absolute-coordinate base interpolation object and the polar-coordinate base interpolation object are prepared with reference to Fig. 1, so that contents can be reproduced on the basis of the intention of a contents producer.

[0082]

If the absolute-coordinate base interpolation object and the polar-coordinate base interpolation object are prepared, the content reproducing system is configured as illustrated in, for example, Fig. 4. In Fig. 4, parts corresponding to those of Fig. 2 are denoted by the same reference signs and a description thereof is omitted as necessary.

[0083]

In the content reproducing system illustrated in Fig. 4, the server 11 includes the configuration information transmission unit 21 and the encoded data transmission unit 22 as in the example of Fig. 2.

[0084]

In the server 11, the object polar-coordinate encoded data on the polar-coordinate base interpolation object is prepared for each reference viewpoint in addition to the object polar-coordinate encoded data on the absolute-coordinate base interpolation object.

[0085]

The object polar-coordinate encoded data on the polar-coordinate base interpolation object at each reference viewpoint is obtained by encoding the object polar-coordinate position information indicating the relative position of the polar-coordinate base interpolation object viewed from each reference viewpoint.

[0086]

The object polar-coordinate position information on the polar-coordinate base interpolation object has the same coordinate expression as the object polar-coordinate position information on the absolute-coordinate base interpolation object.

[0087]

In other words, the object polar-coordinate position information on the polar-coordinate base interpolation object and the object polar-coordinate position information on the absolute-coordinate base interpolation object are both polar coordinates indicating a relative position viewed from the reference viewpoint serving as an origin.

[0088]

As described above, the object polar-coordinate position information on the polar-coordinate base interpolation object has the same coordinate expression as the object polar-coordinate position information on the absolute-coordinate base interpolation object.

[0089]

In the client 12, however, the object polar-coordinate position information on the polar-coordinate base interpolation object is subjected to different processing from the absolute-coordinate base interpolation object, thereby obtaining the polar-coordinate base interpolation object.

[0090]

For three reference viewpoints indicated by the viewpoint selection information, the encoded data transmission unit 22 transmits the object polar-coordinate encoded data on the absolute-coordinate base interpolation object and the object polar-coordinate encoded data on the polar-coordinate base interpolation object to the client 12.

[0091]

In the example of Fig. 4, the client 12 further includes an object position calculation unit 71 in addition to the configuration of Fig. 2.

[0092]

In the decoding unit 45 of the client 12, the object polar-coordinate position information on the absolute-coordinate base interpolation object is supplied to the coordinate transformation unit 46 as in Fig. 2, the object polar-coordinate position information being obtained by decoding.

[0093]

Moreover, the decoding unit 45 supplies the object polar-coordinate position information on the polar-coordinate base interpolation object at each reference viewpoint to the object position calculation unit 71, the object polar-coordinate position information being obtained by decoding.

[0094]

The object position calculation unit 71 performs interpolation on the basis of the listener position information supplied from the listener-position information acquisition unit 41 and the object polar-coordinate position information on the polar-coordinate base interpolation object from the decoding unit 45. At this point, the object position calculation unit 71 acquires the system configuration information from the configuration information acquisition unit 43 or acquires the viewpoint selection information from the viewpoint selection unit 42 as necessary.

[0095]

Thus, the object polar-coordinate position information on the polar-coordinate base interpolation object for the position of the listener (listening position) is obtained as polar-coordinate position information.

[0096]

The polar-coordinate position information is polar coordinates indicating the relative position of the polar-coordinate base interpolation object viewed from the listening position when the viewpoint of the listener is located at the listening position indicated by the listener position information.

[0097]

The object position calculation unit 71 outputs the polar-coordinate position information on the polar-coordinate base interpolation object to the rendering unit, which is not illustrated, in the subsequent stage, the polar-coordinate position information being obtained by interpolation.

[0098]

In the object position calculation unit 71, the object polar-coordinate position information on reference viewpoints is used directly, that is, as polar coordinates for interpolation without being transformed into absolute coordinates, and the polar-coordinate position information indicating relative positions viewed from the listening position is generated.

[0099]

As described above, the polar-coordinate base interpolation object is subjected to different processing from the absolute-coordinate base interpolation object. This can obtain the polar-coordinate base interpolation object that is viewed at a different position by the listener according to the listening position but is always fixed at the same position (direction) when viewed from the listener at the same listening position regardless of the orientation (viewpoint) of the listener.

[0100]

For example, for the purpose of replicating a live venue, an object representing background noise or the like at the venue is disposed at an absolute position in a free viewpoint space regardless of the position of a listener as illustrated in, for example, Fig. 5. In Fig. 5, parts corresponding to those of Fig. 1 are denoted by the same reference signs and a description thereof is omitted.

[0101]

In this example, a free viewpoint space has an object always located at the same position in the free viewpoint space regardless of the viewpoint of the listener, in addition to the absolute-coordinate base interpolation object and the polar-coordinate base fixed object. Hereinafter, such an object always located at the same position (fixed position) in a free viewpoint space (common absolute coordinate space) will be referred to as an absolute-coordinate base fixed object.

[0102]

For example, in the example of Fig. 5, when the listener moves from a position P11 to a position P11', the absolute-coordinate base interpolation object moves from a position P12 to a position P12' and the polar-coordinate base fixed object moves from a position P13 to a position P13'. Even when the listener moves from the position P11 to the position P11', the absolute-coordinate base fixed object stays at a position P31.

[0103]

Such a behavior of the absolute-coordinate base fixed object may be represented by the absolute-coordinate base interpolation object. In this case, however, object polar-coordinate position information is determined by calculating

backward an absolute fixed object at each reference viewpoint as relative layout information, resulting in many disadvantages such as excessive calculation and a demerit in terms of accuracy.

[0104]

Hence, in the present technique, the absolute-coordinate base fixed object is additionally prepared as an object at a fixed absolute coordinate position, and the absolute-coordinate base fixed object is combined with the absolute-coordinate base interpolation object and the polar-coordinate base fixed object, achieving a reduction in computational complexity and a merit in terms of accuracy.

[0105]

Specifically, for example, if the absolute-coordinate base interpolation object and the absolute-coordinate base fixed object are prepared, the content reproducing system is configured as illustrated in, for example, Fig. 6. In Fig. 6, parts corresponding to those of Fig. 2 are denoted by the same reference signs and a description thereof is omitted as necessary.

[0106]

In the example of Fig. 6, the server 11 and the client 12 are identical in configuration to those in Fig. 2. In the server 11, object absolute-coordinate encoded data on the absolute-coordinate base fixed object is prepared in addition to the absolute-coordinate base interpolation object and the object polar-coordinate encoded data.

[0107]

The object absolute-coordinate encoded data on the absolute-coordinate base fixed object is obtained by encoding the object absolute-coordinate position information on absolute coordinates indicating the absolute position of the absolute-coordinate base fixed object in the common absolute coordinate space (common absolute coordinate system).

[0108]

The object absolute-coordinate position information on the absolute-coordinate base fixed object corresponds to the object absolute-coordinate position information on the absolute-coordinate base interpolation object, the object absolute-coordinate position information being obtained by the object position calculation unit 48 of the client 12.

[0109]

For three reference viewpoints indicated by the viewpoint selection information, the encoded data transmission unit 22 transmits the object polar-coordinate encoded data on the absolute-coordinate base interpolation object and the object absolute-coordinate encoded data on the absolute-coordinate base fixed object to the client 12.

[0110]

Thus, in the decoding unit 45 of the client 12, the object polar-coordinate encoded data and the object absolute-coordinate encoded data are decoded.

[0111]

The decoding unit 45 supplies the decoded object polar-coordinate position information on the absolute-coordinate base interpolation object to the coordinate transformation unit 46 and the decoded object absolute-coordinate position information on the absolute-coordinate base fixed object to the polar-coordinate transformation unit 49.

[0112]

The absolute-coordinate base fixed object is always fixed at the same position in the common absolute coordinate space regardless of the viewpoint of the listener, thereby eliminating the need for interpolation unlike the absolute-coordinate base interpolation object.

[0113]

The object absolute-coordinate position information on the absolute-coordinate base fixed object may be handled as the object absolute-coordinate position information on the absolute-coordinate base interpolation object, the object absolute-coordinate position information being obtained by the object position calculation unit 48. Thus, the object absolute-coordinate position information is directly supplied to the polar-coordinate transformation unit 49 after being decoded.

[0114]

The polar-coordinate transformation unit 49 performs polar-coordinate transformation on the object absolute-coordinate position information supplied from the decoding unit 45 as well as the object absolute-coordinate position information supplied from the object position calculation unit 48, on the basis of the listener position information from the listener-position information acquisition unit 41.

[0115]

The polar-coordinate transformation unit 49 outputs the polar-coordinate position information on the absolute-coordinate base interpolation object and the polar-coordinate position information on the absolute-coordinate base fixed object to the rendering unit, which is not illustrated, in the subsequent stage, the polar-coordinate position information being obtained by the polar-coordinate transformation.

[0116]

In the above example, a specific combination of the absolute-coordinate base interpolation object, the polar-coordinate base fixed object, the polar-coordinate base interpolation object, and the absolute-coordinate base fixed object was described. Any combination of the objects may be adopted in the present disclosure.

[0117]

<Example of Interpolation>

Specific examples of interpolation performed in the object position calculation

unit 48 and interpolation performed in the object position calculation unit 71 will be described below.

[0118]

An example of interpolation performed in the object position calculation unit 48 on the basis of the object absolute-coordinate position information at each reference viewpoint will be described below.

[0119]

For example, as illustrated on the left side of Fig. 7, it is assumed that object absolute-coordinate position information at any listening position F is determined by interpolation.

[0120]

In this example, three reference viewpoints A, B, and C surround the listening position F and interpolation is performed using information on the reference viewpoints A to C.

[0121]

Hereinafter, at the listening position F has an X coordinate and a Y coordinate are represented as (x_f, y_f) in the common absolute coordinate system, that is, the XYZ coordinate system.

[0122]

Likewise, at the positions of the reference viewpoint A, the reference viewpoint B, and the reference viewpoint C, X coordinates and Y coordinates are represented as (x_a, y_a) , (x_b, y_b) , and (x_c, y_c) , respectively.

[0123]

In this case, as illustrated on the right side of Fig. 7, an object position F' at the listening position F is determined on the basis of the coordinates of an object position A', an object position B', and an object position C' that correspond to the reference viewpoint A, the reference viewpoint B, and the reference viewpoint C.

[0124]

In this case, for example, the object position A' indicates the position of the object when the viewpoint is located at the reference viewpoint A, that is, the position of the absolute-coordinate base interpolation object, which is indicated by the object absolute-coordinate position information on the reference viewpoint A, in the common absolute coordinate system.

[0125]

Furthermore, the object position F' indicates the position of the absolute-coordinate base interpolation object in the common absolute coordinate system when the listener is located at the listening position F, that is, a position indicated by the object absolute-coordinate position information serving as the output of the object position calculation unit 48.

[0126]

Hereinafter, at the object position A', the object position B', and the object position C', X coordinates and Y coordinates are represented as (x'_a, y'_a) , (x'_b, y'_b) ,

and (x_c', y_c') , respectively, and an X coordinate and a Y coordinate at the object position F' are represented as (x_f', y_f') .

[0127]

Furthermore, in the following description, a triangular region surrounded by any three reference viewpoints such as the reference viewpoints A to C, that is, a triangular region formed by the three reference viewpoints will be referred to as a triangular mesh.

[0128]

Since the common absolute coordinate space includes multiple reference viewpoints, a plurality of triangular meshes can be formed with the vertices at the reference viewpoints.

[0129]

Similarly, in the following description, a triangular region surrounded (formed) by object positions such as the object positions A' to C' indicated by object absolute-coordinate position information at any three reference viewpoints will be referred to as a triangular mesh.

[0130]

Coordinates indicating any position in the common absolute coordinate system (XYZ coordinate system) can be obtained from coordinates at the position in the xyz coordinate system and information on the reference viewpoints included in the system configuration information, more specifically, the positions of the reference viewpoints and the orientations of the listener at the reference viewpoints.

[0131]

The xyz coordinate system is an absolute coordinate system with the origin (fiducial) at the position of the reference viewpoint. To simplify the description, it is assumed that a Z coordinate value in the XYZ coordinate system is equal to a z coordinate value in the xyz coordinate system.

[0132]

According to the Ceva's theorem, by properly determining the internal ratios of the sides of a triangular mesh, any listening position in the triangular mesh formed by three reference viewpoints is uniquely determined at the intersection point of lines from the three vertices of the triangular mesh to the internally dividing points of the three sides that are not adjacent to the vertices.

[0133]

This is established for all triangular meshes regardless of the shapes of the triangular meshes from a proof formula, by determining the configurations of the internal ratios of the three sides of the triangular meshes.

[0134]

Hence, by determining the internal ratios of a triangular mesh including a listening position near a viewpoint, that is, with respect to a reference viewpoint and applying the internal ratios to an object, that is, a triangular mesh at the

position of the object, a proper object position can be determined with respect to any listening position.

[0135]

Referring to Figs. 8 and 9, an example of interpolation for determining object absolute-coordinate position information on the object position F' for the listening position F in Fig. 7 will be described below.

[0136]

For example, as illustrated in Fig. 8, the X coordinates and Y coordinates of internally dividing points are first determined in a triangular mesh having reference viewpoints A to C with the listening position F included in the triangular mesh.

[0137]

A point D denotes an intersection point of a straight line passing through the listening position F and the reference viewpoint C and a line AB from the reference viewpoint A to the reference viewpoint B, and (x_d, y_d) denotes coordinates indicating the position of the point D on the XY plane. In other words, the point D is an internally dividing point on the line AB (side AB).

[0138]

In this case, regarding an X coordinate and a Y coordinate that indicate the position of any point on a line CF from the reference viewpoint C to the listening position F and an X coordinate and a Y coordinate that indicate the position of any point on the line AB, relationships are established as in formula (1) below.

[0139]

[Math. 1]

$$\begin{aligned} \text{Line CF: } Y &= \alpha_1 X - \alpha_1 x_c + y_c, \text{ where } \alpha_1 = (y_c - y_f) / (x_c - x_f) \\ \text{Line AB: } Y &= \alpha_2 X - \alpha_2 x_a + y_a, \text{ where } \alpha_2 = (y_b - y_a) / (x_b - x_a) \quad \cdot \cdot \cdot (1) \end{aligned}$$

[0140]

Since the point D is the intersection point of the straight line passing through the reference viewpoint C and the listening position F and the line AB, the coordinates (x_d, y_d) of the point D on the XY plane can be determined from formula (1). The coordinates (x_d, y_d) are expressed in formula (2) below.

[0141]

[Math. 2]

$$\begin{aligned} x_d &= (\alpha_1 x_c - y_c - \alpha_2 x_a + y_a) / (\alpha_1 - \alpha_2) \\ y_d &= \alpha_1 x_d - \alpha_1 x_c + y_c \quad \cdot \cdot \cdot (2) \end{aligned}$$

[0142]

Thus, as expressed in formula (3) below, an internal ratio (m, n) at the point D of the line AB, that is, a division ratio at the point D can be obtained 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.

[0143]

[Math. 3]

$$\begin{aligned} m &= \sqrt{(x_a - x_d)^2 + (y_a - y_d)^2} \\ n &= \sqrt{(x_b - x_d)^2 + (y_b - y_d)^2} \end{aligned} \quad \cdot \cdot \cdot (3)$$

[0144]

Likewise, a point E denotes an intersection point of a straight line passing through the listening position F and the reference viewpoint B and a line AC from the reference viewpoint A to the reference viewpoint C, and (x_e, y_e) denotes coordinates indicating the position of the point E on the XY plane. In other words, the point E is an internally dividing point on the line AC (side AC).

[0145]

In this case, regarding an X coordinate and a Y coordinate that indicate the position of any point on a line BF from the reference viewpoint B to the listening position F and an X coordinate and a Y coordinate that indicate the position of any point on the line AC, relationships are established as in formula (4) below.

[0146]

[Math. 4]

$$\begin{aligned} \text{Line BF : } Y &= \alpha_3 X - \alpha_3 x_b + y_b, \text{ where } \alpha_3 = (y_b - y_f) / (x_b - x_f) \\ \text{Line AC : } Y &= \alpha_4 X - \alpha_4 x_a + y_a, \text{ where } \alpha_4 = (y_c - y_a) / (x_c - x_a) \end{aligned} \quad \cdot \cdot \cdot (4)$$

[0147]

Since the point E is the intersection point of the straight line passing through the reference viewpoint B and the listening position F and the line AC, the coordinates (x_e, y_e) of the point E on the XY plane can be determined from formula (4). The coordinates (x_e, y_e) are expressed in formula (5) below.

[0148]

[Math. 5]

$$\begin{aligned} x_e &= (\alpha_3 x_b - y_b - \alpha_4 x_a + y_a) / (\alpha_3 - \alpha_4) \\ y_e &= \alpha_3 x_e - \alpha_3 x_b + y_b \end{aligned} \quad \cdot \cdot \cdot (5)$$

[0149]

Thus, as expressed in formula (6) below, an internal ratio (k, l) at the point E of the line AC, that is, a division ratio at the point E can be obtained 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.

[0150]

[Math. 6]

$$\begin{aligned} k &= \sqrt{(x_a - x_e)^2 + (y_a - y_e)^2} \\ l &= \sqrt{(x_c - x_e)^2 + (y_c - y_e)^2} \end{aligned} \quad \cdot \cdot \cdot (6)$$

[0151]

Thereafter, the determined ratios of the two sides, specifically, the internal ratio (m, n) and the internal ratio (k, l) are applied to an object-side triangular mesh as illustrated in Fig. 9, so that coordinates (x_f', y_f') at the object position F' on the XY plane are determined.

[0152]

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

[0153]

Similarly, a point E' on a line A'C' connecting the object position A' and an object position C' corresponds to the point E.

[0154]

Moreover, an object position F' at the intersection point of 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' corresponds to the listening position F.

[0155]

In this case, it is assumed that the point D' of the line A'B' has the same internal ratio (m, n) as the point D. At this point, as expressed in formula (7) below, the coordinates (x_d', y_d') of the point D' on the XY plane can be obtained on the basis of the internal 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'.

[0156]

[Math. 7]

$$\begin{aligned} x_d' &= (nx_a' + mx_b') / (m+n) \\ y_d' &= (ny_a' + my_b') / (m+n) \end{aligned} \quad \cdot \cdot \cdot (7)$$

[0157]

Moreover, it is assumed that the point E' of the line A'C' has the same internal ratio (k, l) as the point E. At this point, as expressed in formula (8) below, the coordinates (x_e', y_e') of the point E' on the XY plane can be obtained on the basis of the internal 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'.

[0158]

[Math. 8]

$$\begin{aligned} x_e' &= (lx_a' + kx_c') / (k+l) \\ y_e' &= (ly_a' + ky_c') / (k+l) \end{aligned} \quad \cdot \cdot \cdot (8)$$

[0159]

Hence, regarding an X coordinate and a Y coordinate that indicate the position of any point on a line B'E' from the object position B' to the point E' and an X coordinate and a Y coordinate that indicate the position of any point on a line C'D' from the object position C' to the point D', relationships are established as in

formula (9) below.

[0160]

[Math. 9]

Line B'E' : $Y = \alpha_5 X + y_b' - \alpha_5 x_b'$, where $\alpha_5 = (y_e' - y_b') / (x_e' - x_b')$

Line C'D' : $Y = \alpha_6 X + y_c' - \alpha_6 x_c'$, where $\alpha_6 = (y_d' - y_c') / (x_d' - x_c')$

• • • (9)

[0161]

Since the target object position F' is the intersection point of the line B'E' and the line C'D', the coordinates (x_f', y_f') of the object position F' can be obtained from the relationships of formula (9) according to formula (10) below.

[0162]

[Math. 10]

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

$$y_f' = \alpha_6 x_f' + y_c' - \alpha_6 x_c' \quad \cdot \cdot \cdot (10)$$

[0163]

This processing obtains the coordinates (x_f', y_f') of the object position F' on the XY plane.

[0164]

Subsequently, the coordinates (x_f', y_f', z_f') of the object position F' in the XYZ coordinate system are determined 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. In other words, a Z coordinate z_f' of the object position F' in the XYZ coordinate system is determined.

[0165]

For example, a triangle with the vertices at the object position A', the object position B', and the object position C' in the XYZ coordinate system (common absolute coordinate system) is determined on a three-dimensional 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 determined. Furthermore, a point with an X coordinate and a Y coordinate of (x_f', y_f') is determined on the three-dimensional plane A'B'C', and a z coordinate at the point is z_f' .

[0166]

Specifically, a vector with the initial point at the object position A' and the final point at the object position B' in the XYZ coordinate system is denoted as a vector $A'B' = (x_{ab}', y_{ab}', z_{ab}')$.

[0167]

Similarly, a vector with the initial point at the object position A' and the final point at the object position C' in the XYZ coordinate system is denoted as a vector

$$A'C' = (x_{ac}', y_{ac}', z_{ac}').$$

[0168]

The vector A'B' and the vector 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'. In other words, the vector A'B' and the vector A'C' can be obtained by formula (11) below.

[0169]

[Math. 11]

$$\text{Vector A'B'} : (x_{ab}', y_{ab}', z_{ab}') = (x_b' - x_a', y_b' - y_a', z_b' - z_a')$$

$$\text{Vector A'C'} : (x_{ac}', y_{ac}', z_{ac}') = (x_c' - x_a', y_c' - y_a', z_c' - z_a') \cdot \cdot \cdot (11)$$

[0170]

A normal vector (s, t, u) of the three-dimensional plane A'B'C' is the outer product of the vector A'B' and the vector A'C' and can be determined by formula (12) below.

[0171]

[Math. 12]

$$(s, t, u) = (y_{ab}'z_{ac}' - z_{ab}'y_{ac}', z_{ab}'x_{ac}' - x_{ab}'z_{ac}', x_{ab}'y_{ac}' - y_{ab}'x_{ac}') \cdot \cdot \cdot (12)$$

[0172]

Hence, from the normal vector (s, t, u) and the coordinates (x_a', y_a', z_a') of the object position A', a plane equation of the three-dimensional plane A'B'C' is determined as expressed in formula (13) below.

[0173]

[Math. 13]

$$s(X - x_a') + t(Y - y_a') + u(Z - z_a') = 0 \cdot \cdot \cdot (13)$$

[0174]

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 been already determined, the substitution of the X coordinate x_f' and the Y coordinate y_f' into X and Y of the plane equation of formula (13) can determine the Z coordinate z_f' as expressed in formula (14) below.

[0175]

[Math. 14]

$$z_f' = (-s(x_f' - x_a') - t(y_f' - y_a')) / u + z_a' \cdot \cdot \cdot (14)$$

[0176]

The calculation obtains the coordinates (x_f', y_f', z_f') of the target object position F'. In the object position calculation unit 48, the object absolute-coordinate position information indicating the obtained coordinates (x_f', y_f', z_f') of the object position

F' is outputted.

[0177]

Interpolation performed in the object position calculation unit 71 will be described below.

[0178]

For example, as indicated in Fig. 10, the object polar-coordinate position information on the polar-coordinate base interpolation object is determined by interpolation at the listening position F surrounded by the three reference viewpoints: the reference viewpoint A, the reference viewpoint B, and the reference viewpoint C.

[0179]

In Fig. 10, parts corresponding to those of Fig. 8 are denoted by the same reference characters (signs) and a description thereof is omitted as necessary.

[0180]

In other words, also in the example of Fig. 10, the same calculation as in Fig. 8 determines the coordinates (x_d, y_d) of the point D, the coordinates (x_e, y_e) of the point E, the coordinates (x_f, y_f) of the listening position F, an internal ratio (m, n), and an internal ratio (k, l).

[0181]

Moreover, the object polar-coordinate position information on the polar-coordinate base interpolation object at the reference viewpoint A, the reference viewpoint B, and the reference viewpoint C is denoted as (Az(a), El(a), Rad(a)), (Az(b), El(b), Rad(b)), and (Az(c), El(c), Rad(c)).

[0182]

For example, Az(a), El(a), and Rad(a) are a horizontal angle, a vertical angle, and a radius that constitute polar coordinates.

[0183]

In this case, from the object polar-coordinate position information (Az(a), El(a), Rad(a)) and (Az(b), El(b), Rad(b)) at the reference viewpoint A and the reference viewpoint B and the internal ratio (m, n), the object polar-coordinate position information (Az(d), El(d), Rad(d)) on the polar-coordinate base interpolation object at the point D is determined by formula (15) below.

[0184]

[Math. 15]

$$\begin{aligned} \text{Az}(d) &= (m \cdot \text{Az}(b) + n \cdot \text{Az}(a)) / (m+n) \\ \text{El}(d) &= (m \cdot \text{El}(b) + n \cdot \text{El}(a)) / (m+n) \\ \text{Rad}(d) &= (m \cdot \text{Rad}(b) + n \cdot \text{Rad}(a)) / (m+n) \quad \cdot \cdot \cdot (15) \end{aligned}$$

[0185]

Furthermore, from the object polar-coordinate position information (Az(d), El(d), Rad(d)) at the point D, the object polar-coordinate position information (Az(c), El(c), Rad(c)) at the reference viewpoint C, the coordinates (x_c, y_c) of the reference

viewpoint C, the coordinates (x_d, y_d) of the point D, and the coordinates (x_f, y_f) of the listening position F, the object polar-coordinate position information ($Az(f)$, $El(f)$, $Rad(f)$) on the polar-coordinate base interpolation object at the listening position F is determined by formula (16) below.

[0186]

[Math. 16]

$$Az(f) = (o * Az(c) + p * Az(d)) / (o + p)$$

$$El(f) = (o * El(c) + p * El(d)) / (o + p)$$

$$Rad(f) = (o * Rad(c) + p * Rad(d)) / (o + p)$$

where

$$o = \text{SQRT}((x_d - x_f)^2 + (y_d - y_f)^2 + (z_d - z_f)^2)$$

$$p = \text{SQRT}((x_c - x_f)^2 + (y_c - y_f)^2 + (z_c - z_f)^2) \quad \cdot \cdot \cdot (16)$$

[0187]

As described above, in the object position calculation unit 71, interpolation is performed on the basis of the object polar-coordinate position information at the three reference viewpoints according to formula (16) while keeping polar coordinates, so that the object polar-coordinate position information on the polar-coordinate base interpolation object at the listening position is calculated.

[0188]

The interpolation performed in the object position calculation unit 71 is not limited to the example described with reference to Fig. 10. The interpolation may be implemented by any kinds of processing, for example, interpolation using a vector operation or an operation using a neural network.

[0189]

<Example of System Configuration Information>

Fig. 11 indicates an example of the bit stream format of the system configuration information when the absolute-coordinate base interpolation object, the polar-coordinate base fixed object, the polar-coordinate base interpolation object, and the absolute-coordinate base fixed object can be handled.

[0190]

In the example of Fig. 11, “NumOfObjs” represents the number of objects constituting contents. In this case, the number of objects means the total number of the absolute-coordinate base interpolation objects, the polar-coordinate base fixed objects, the polar-coordinate base interpolation objects, and the absolute-coordinate base fixed objects.

[0191]

Moreover, “NumOfRefViewPoint” represents the number of reference viewpoints.

[0192]

The system configuration information includes pieces of reference viewpoint information including reference viewpoint position information and listener

orientation information, as many as the number of reference viewpoints
“NumOfRefViewPoint”.

[0193]

The reference viewpoint position information is the absolute coordinates of the common absolute coordinate system that indicates the position of the reference viewpoint. In this example, the reference viewpoint position information includes an X coordinate “RefViewX[i]”, a Y coordinate “RefViewY[i]”, and a Z coordinate “RefViewZ[i]” that represent the position of the reference viewpoint in the common absolute coordinate system.

[0194]

The listener orientation information is a rotation angle (horizontal angle) of the face of the listener in the horizontal direction, the rotation angle indicating a desired orientation of the face of the listener at the reference viewpoint, that is, an assumed orientation of the face of the listener at the reference viewpoint.

[0195]

In this example, a horizontal angle “RefYaw[i]” of the face of the listener is included as the listener orientation information. The listener orientation information may further include a vertical angle (pitch angle) indicating the orientation of the face of the listener in the vertical direction in addition to a horizontal angle (yaw angle) of the face of the listener.

[0196]

The system configuration information further includes information
“ObjectOverLapMode[i]” indicating a reproduction mode in which the listener and the positions of the objects overlap each other, that is, the listener (listening position) and the objects are located at the same position according to the number of objects “NumOfObjs”. Furthermore, “ProhibitRadius” represents the normalized value of a distance from the object to the listener if a space is normalized by 1.0.

[0197]

“InpterporationMode” represents an interpolation mode permitted for the client 12. Moreover, “NonInterpolatePolarObjFlag” is a flag indicating the presence or absence of the polar-coordinate base fixed object. Specifically, a value “1” of the flag “NonInterpolatePolarObjFlag” indicates the presence of the polar-coordinate base fixed object, whereas a value “0” indicates the absence of the polar-coordinate base fixed object.

[0198]

If the value of the flag “NonInterpolatePolarObjFlag” is “1”, “NumOfObjs_NIPO” representing the number of polar-coordinate base fixed objects is stored in the system configuration information.

[0199]

“NonInterpolateCartesianObjFlag” is a flag indicating the presence or absence of the absolute-coordinate base fixed object. Specifically, a value “1” of the flag

“NonInterpolateCartesianObjFlag” indicates the presence of the absolute-coordinate base fixed object, whereas a value “0” indicates the absence of the absolute-coordinate base fixed object.

[0200]

If the value of the flag “NonInterpolateCartesianObjFlag” is “1”, “NumOfObjs_NICO” representing the number of absolute-coordinate base fixed objects is stored in the system configuration information.

[0201]

“InterpolatePolarObjFlag” is a flag indicating the presence or absence of the polar-coordinate base interpolation object. Specifically, a value “1” of the flag “InterpolatePolarObjFlag” indicates the presence of the polar-coordinate base interpolation object, whereas a value “0” indicates the absence of the polar-coordinate base interpolation object.

[0202]

If the value of the flag “InterpolatePolarObjFlag” is “1”, “NumOfObjs_IPO” representing the number of polar-coordinate base interpolation objects is stored in the system configuration information.

[0203]

“NumOfAncBytes” represents the size of an extended information region and “AncByteData[i]” represents extended-region byte data.

[0204]

For example, the system configuration information configured as in Fig. 11 is transmitted from the server 11 to the client 12.

[0205]

<Example of Bit Stream Format>

Fig. 12 indicates an example of the bit stream format in batch transmission of information on the positions of the objects, that is, the object polar-coordinate encoded data or the object absolute-coordinate encoded data on the condition that the positions of the objects do not change with time, that is, the objects do not move.

[0206]

In this example, “fva_structure_info_polar0” represents the system configuration information. The system configuration information is not always supposed to be included when being separately transmitted.

[0207]

In a bit stream, pieces of metadata “object_metadata0” on the absolute-coordinate base interpolation objects at the reference points are included as many as the number of reference viewpoints “NumOfRefViewPoint”.

[0208]

The metadata “object_metadata0” includes the object polar-coordinate position information on the absolute-coordinate base interpolation object, more specifically, the object polar-coordinate encoded data or gain information (gain

amount) on the absolute-coordinate base interpolation object.

[0209]

Since the information is particularly transmitted by batch in this example, the metadata “object_metadata0” on the absolute-coordinate base interpolation objects is stored for all the reference viewpoints.

[0210]

If the value of a flag “NonInterpolatePolarObjFlag” included in the system configuration information is “1”, metadata “object_metadata_nonintpPolar0” on the polar-coordinate base fixed objects is stored in the bit stream.

[0211]

Likewise, if the value of the flag “NonInterpolateCartesianObjFlag” included in the system configuration information is “1”, metadata “object_metadata_nonintpCarte0” on the absolute-coordinate base fixed objects is stored in the bit stream.

[0212]

Furthermore, if the value of a flag “InterpolatePolarObjFlag” included in the system configuration information is “1”, metadata “object_metadata_intpPolar0” on the polar-coordinate base interpolation objects is stored in the bit stream.

[0213]

In this example, as in the case of the absolute-coordinate base interpolation object, pieces of the metadata “object_metadata_intpPolar0” on the polar-coordinate base interpolation objects at the reference points are stored as many as the number of reference viewpoints “NumOfRefViewPoint”.

[0214]

In the example of Fig. 12, it is assumed that the positions of the objects do not change with time.

[0215]

Fig. 13 indicates an example of the bit stream format when the object polar-coordinate encoded data or the object absolute-coordinate encoded data on the objects is transmitted for each frame of audio data on the objects so as to correspond to a change of the position of each object with time.

[0216]

In Fig. 13, “fva_structure_info_polar_present” represents a configuration information presence flag indicating the presence or absence of the system configuration information in the bit stream. A value “1” of the configuration information presence flag particularly indicates that the system configuration information is included (stored). In contrast, a value “0” of the configuration information presence flag indicates that the system configuration information is not included.

[0217]

If the value of the configuration information presence flag “fva_structure_info_polar_present” is “1”, the system configuration information

“fva_structure_info_polar()” is included in the bit stream.

[0218]

Also in the example of Fig. 13, the system configuration information is not always supposed to be included. The system configuration information may be transmitted at regular or irregular intervals. In other words, the system configuration information may be transmitted for one frame without being transmitted for another frame.

[0219]

Since the information is transmitted for each frame in this example, metadata “object_metadata()” on the absolute-coordinate base interpolation objects is stored only for three reference viewpoints indicated by the viewpoint selection information received (acquired) from the client 12.

[0220]

Moreover, as in the example of Fig. 12, if the value of a flag “NonInterpolatePolarObjFlag” included in the system configuration information is “1”, metadata “object_metadata_nonintpPolar()” on the polar-coordinate base fixed objects is stored in the bit stream.

[0221]

If the value of a flag “NonInterpolateCartesianObjFlag” included in the system configuration information is “1”, metadata “object_metadata_nonintpCarte()” on the absolute-coordinate base fixed objects is stored in the bit stream.

[0222]

Furthermore, if the value of a flag “InterpolatePolarObjFlag” included in the system configuration information is “1”, metadata “object_metadata_intpPolar()” on the polar-coordinate base interpolation objects is stored in the bit stream only for three reference viewpoints indicated by the viewpoint selection information.

[0223]

Referring to Figs. 14 to 16, examples of metadata on the polar-coordinate base fixed objects, metadata on the polar-coordinate base interpolation objects, and metadata on the absolute-coordinate base fixed objects will be described below.

[0224]

Fig. 14 indicates an example of the bit stream format of the metadata “object_metadata_nonintpPolar()” on the polar-coordinate base fixed objects indicated in Figs. 12 and 13.

[0225]

In this example, the object polar-coordinate position information (object polar-coordinate encoded data) and the gain amount (gain information) of the polar-coordinate base fixed objects are stored according to the number of polar-coordinate base fixed objects “NumOfObjs_NIPO” included in the system configuration information.

[0226]

Specifically, “PosAzi[i]”, “PosEle[i]”, and “PosRad[i]” represent a horizontal angle,

a vertical angle, and a radius that constitute the object polar-coordinate position information on the polar-coordinate base fixed object. Moreover, “Gain[i]” represents a gain amount for a gain adjustment to audio data on the polar-coordinate base fixed object, more specifically, encoded gain information obtained by encoding the gain information.

[0227]

Fig. 15 indicates an example of the bit stream format of the metadata “object_metadata_intpPolar()” on the polar-coordinate base interpolation objects indicated in Figs. 12 and 13.

[0228]

In this example, the object polar-coordinate position information (object polar-coordinate encoded data) and the gain amount (gain information) of the polar-coordinate base interpolation objects are stored according to the number of polar-coordinate base interpolation objects “NumOfObjs_IPO” included in the system configuration information.

[0229]

Specifically, “PosAzi[i]”, “PosEle[i]”, and “PosRad[i]” represent a horizontal angle, a vertical angle, and a radius that constitute the object polar-coordinate position information on the polar-coordinate base interpolation object. Moreover, “Gain[i]” represents a gain amount for a gain adjustment to audio data on the polar-coordinate base interpolation object, more specifically, encoded gain information.

[0230]

Fig. 16 indicates an example of the bit stream format of the metadata “object_metadata_nonintpCarte()” on the absolute-coordinate base fixed objects indicated in Figs. 12 and 13.

[0231]

In this example, the object absolute-coordinate position information (object absolute-coordinate encoded data) and the gain amount (gain information) of the absolute-coordinate base fixed objects are stored according to the number of absolute-coordinate base fixed objects “NumOfObjs_NICO” included in the system configuration information.

[0232]

Specifically, “PosX[i]”, “PosY[i]”, and “PosZ[i]” represent an X coordinate, a Y coordinate, and a Z coordinate of the common absolute coordinate system (XYZ coordinate system), the X, Y, and Z coordinates constituting the object absolute-coordinate position information on the absolute-coordinate base fixed object. Moreover, “Gain[i]” represents a gain amount for a gain adjustment to audio data on the absolute-coordinate base fixed object, more specifically, encoded gain information.

[0233]

<Configuration Example of Content Reproducing System>

A more specific embodiment of the content reproducing system to which the present technique is applied will be described below.

[0234]

Fig. 17 illustrates a configuration example of the content reproducing system to which the present technique is applied. In Fig. 17, parts corresponding to those of Fig. 4 are denoted by the same reference signs and a description thereof is omitted as necessary.

[0235]

The content reproducing system in Fig. 17 includes the server 11 that distributes contents and the client 12 that receives contents distributed from the server 11.

[0236]

The server 11 includes a configuration information recording unit 101, the configuration information transmission unit 21, a recording unit 102, the encoded data transmission unit 22, and a transmission buffer 103.

[0237]

The configuration information recording unit 101 records, for example, the system configuration information prepared as indicated in Fig. 11 and supplies the recorded system configuration information to the configuration information transmission unit 21. The recording unit 102 may partly serve as the configuration information recording unit 101.

[0238]

The recording unit 102 records, for example, encoded audio data obtained by encoding audio data on the objects, the object polar-coordinate encoded data, the object absolute-coordinate encoded data, and the encoded gain information on each object at each reference viewpoint. The data and information constitute the contents.

[0239]

The recording unit 102 supplies, for example, the recorded encoded audio data, the recorded object polar-coordinate encoded data, the recorded object absolute-coordinate encoded data, and the recorded encoded gain information to the encoded data transmission unit 22 in response to a request or the like.

[0240]

The transmission buffer 103 temporarily holds, for example, the encoded audio data, the object polar-coordinate encoded data, the object absolute-coordinate encoded data, and the encoded gain information that are supplied from the encoded data transmission unit 22.

[0241]

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

[0242]

The communication unit 111 corresponds to the configuration information

acquisition unit 43 and the encoded data acquisition unit 44 in, for example, Fig. 2 and transmits and receives various kinds of data through communications with the server 11.

[0243]

For example, the communication unit 111 transmits the viewpoint selection information, which is supplied from the viewpoint selection unit 42, to the server 11 and receives the system configuration information and the bit stream that are transmitted from the server 11. In other words, the communication unit 111 acts as a reference-viewpoint information acquisition unit that acquires, from the server 11, the system configuration information, the object polar-coordinate encoded data, the object absolute-coordinate encoded data, and the encoded gain information that are included in the bit stream.

[0244]

The position calculation unit 112 generates the polar-coordinate position information on the positions of all kinds of objects on the basis of the object polar-coordinate position information and the object absolute-coordinate position information that are supplied from the decoding unit 45 and the system configuration information supplied from the communication unit 111, and supplies the polar-coordinate position information to the rendering unit 113.

[0245]

Moreover, the position calculation unit 112 performs a gain adjustment to audio data on all kinds of objects and supplies, to the rendering unit 113, the audio data having been subjected to the gain adjustment, the audio data being supplied from the decoding unit 45.

[0246]

The position calculation unit 112 includes the coordinate transformation unit 46, the coordinate-axis transformation unit 47, the object position calculation unit 48, the polar-coordinate transformation unit 49, and the object position calculation unit 71.

[0247]

The rendering unit 113 performs, for example, rendering such as VBAP on the basis of the polar-coordinate position information and the audio data that are supplied from the polar-coordinate transformation unit 49, the object position calculation unit 71, and the decoding unit 45, generates reproduction audio data for reproducing a sound of contents, and outputs the audio data.

[0248]

<Explanation of Provision>

The operations of the server 11 and the client 12 in Fig. 17 will be described below.

[0249]

For example, in response to a request to set up a network session from the client 12 after initialization, the server 11 performs processing for setting up the

network session with the client 12.

[0250]

Thereafter, the server 11 receives information on the start of the session from the client 12. In response to a request to transmit the system configuration information from the client 12, the server 11 starts provision, that is, processing for providing contents.

[0251]

Referring to the flowchart of Fig. 18, the provision by the server 11 will be described below.

[0252]

In step S11, the configuration information transmission unit 21 reads system configuration information on requested contents from the configuration information recording unit 101 and transmits the read system configuration information to the client 12.

[0253]

For example, the system configuration information is transmitted to the client 12 via a network or the like immediately after a session is set up, that is, immediately after connection between the server 11 and the client 12 is established and before encoded audio data or the like is transmitted.

[0254]

When the system configuration information is transmitted, viewpoint selection information on three reference viewpoints corresponding to the position of the listener is transmitted from the client 12.

[0255]

In step S12, the configuration information transmission unit 21 receives the viewpoint selection information transmitted from the client 12 and supplies the viewpoint selection information and the system configuration information to the encoded data transmission unit 22.

[0256]

In step S13, the encoded data transmission unit 22 loads data on the absolute-coordinate base interpolation object constituting contents, into the transmission buffer 103 on the basis of the viewpoint selection information and the system configuration information that are supplied from the configuration information transmission unit 21.

[0257]

Specifically, regarding three reference viewpoints indicated by the viewpoint selection information, the encoded data transmission unit 22 reads object polar-coordinate encoded data and encoded gain information from the recording unit 102 for each absolute-coordinate base interpolation object, supplies the data and information to the transmission buffer 103, and causes the transmission buffer 103 to hold the data and information. The encoded data transmission unit 22 also reads encoded audio data on each absolute-coordinate base interpolation

object from the recording unit 102, supplies the data to the transmission buffer 103, and causes the transmission buffer 103 to hold the data.

[0258]

In step S14, the encoded data transmission unit 22 determines the presence or absence of a polar-coordinate base fixed object as an object constituting the contents, on the basis of the system configuration information. In this case, if the value of the flag “NonInterpolatePolarObjFlag” included in the system configuration information is “1”, it is determined that the polar-coordinate base fixed object is present.

[0259]

If it is determined that the polar-coordinate base fixed object is present in step S14, the encoded data transmission unit 22 in step S15 loads data on the polar-coordinate base fixed object constituting the contents, into the transmission buffer 103 on the basis of the system configuration information.

[0260]

Specifically, the encoded data transmission unit 22 reads, from the recording unit 102, object polar-coordinate encoded data, encoded gain information, and encoded audio data on each polar-coordinate base fixed object constituting the contents, supplies the data and information to the transmission buffer 103, and causes the transmission buffer 103 to hold the data and information.

[0261]

After the processing of step S15, the process advances to step S16.

[0262]

If it is determined in step S14 that the polar-coordinate base fixed object is absent, the processing of step S15 is not performed. The process then advances to step S16.

[0263]

In step S16, the encoded data transmission unit 22 determines the presence or absence of an absolute-coordinate base fixed object as an object constituting the contents, on the basis of the system configuration information. In this case, if the value of the flag “NonInterpolateCartesianObjFlag” included in the system configuration information is “1”, it is determined that the absolute-coordinate base fixed object is present.

[0264]

If it is determined that the absolute-coordinate base fixed object is present in step S16, the encoded data transmission unit 22 in step S17 loads data on the absolute-coordinate base fixed object constituting the contents, into the transmission buffer 103 on the basis of the system configuration information.

[0265]

Specifically, the encoded data transmission unit 22 reads, from the recording unit 102, object absolute-coordinate encoded data, encoded gain information, and encoded audio data on each absolute-coordinate base fixed object constituting the

contents, supplies the data and information to the transmission buffer 103, and causes the transmission buffer 103 to hold the data and information.

[0266]

After the processing of step S17, the process advances to step S18.

[0267]

If it is determined in step S16 that the absolute-coordinate base fixed object is absent, the processing of step S17 is not performed. The process then advances to step S18.

[0268]

In step S18, the encoded data transmission unit 22 determines the presence or absence of a polar-coordinate base interpolation object as an object constituting the contents, on the basis of the system configuration information. In this case, if the value of the flag “InterpolatePolarObjFlag” included in the system configuration information is “1”, it is determined that the polar-coordinate base interpolation object is present.

[0269]

If it is determined in step S18 that the polar-coordinate base interpolation object is present, the process advances to step S19.

[0270]

In step S19, the encoded data transmission unit 22 loads data on the polar-coordinate base interpolation object constituting the contents, into the transmission buffer 103 on the basis of the viewpoint selection information and the system configuration information.

[0271]

Specifically, regarding three reference viewpoints indicated by the viewpoint selection information, the encoded data transmission unit 22 reads object polar-coordinate encoded data and encoded gain information from the recording unit 102 for each polar-coordinate base interpolation object, supplies the data and information to the transmission buffer 103, and causes the transmission buffer 103 to hold the data and information. The encoded data transmission unit 22 also reads encoded audio data on each polar-coordinate base interpolation object from the recording unit 102, supplies the data to the transmission buffer 103, and causes the transmission buffer 103 to hold the data.

[0272]

After the processing of step S19, the process advances to step S20.

[0273]

If it is determined in step S18 that the polar-coordinate base interpolation object is absent, the processing of step S19 is not performed. The process then advances to step S20.

[0274]

In step S20, the encoded data transmission unit 22 multiplexes the data on the objects to generate a bit stream, the data being loaded into the transmission

buffer 103 in the processing of steps S13 to S19. In this case, the system configuration information may be also multiplexed. This generates a bit stream in the format of, for example, Fig. 13.

[0275]

In step S21, the encoded data transmission unit 22 transmits the generated bit stream to the client 12. This completes the distribution of the contents to the client 12.

[0276]

In step S22, the encoded data transmission unit 22 determines whether to terminate the processing.

[0277]

For example, in the case of a request to stop the transmission of contents from the client 12 or at the completion of the transmission of all the pieces of data on contents, it is determined that the processing is to be terminated.

[0278]

In step S22, if it is determined that the processing is not to be terminated, the process returns to step S12 to repeat the foregoing processing.

[0279]

If it is determined in step S22 that the processing is to be terminated, the units of the server 11 stop performing processing, thereby terminating the provision.

For example, at the completion of the transmission of all the pieces of data on contents, the server 11 transmits information on the completion of the data transmission to the client 12 and terminates the provision.

[0280]

As described above, the server 11 generates a bit stream including data on the objects constituting the contents, that is, necessary kinds of objects from among data on the four kinds of objects, and transmits the bit stream to the client 12. This can reproduce contents on the basis of the intention of a contents producer.

[0281]

<Explanation of Generation of Reproduction Audio Data>

Moreover, the client 12 requests the server 11 to set up a network session after initialization. When receiving a response from the server 11, the client 12 transmits a request to transmit the system configuration information through the communication unit 111.

[0282]

When the system configuration information is transmitted from the server 11 in response to a transmission request, the client 12 starts generating the reproduction audio data.

[0283]

Referring to the flowchart of Fig. 19, the generation of the reproduction audio data by the client 12 will be described below.

[0284]

In step S51, the communication unit 111 receives the system configuration information transmitted from the server 11 and supplies the information to the viewpoint selection unit 42, the coordinate-axis transformation unit 47, the object position calculation unit 48, and the object position calculation unit 71. In this case, the system configuration information may be decoded by the decoding unit 45 as necessary.

[0285]

In step S52, the listener-position information acquisition unit 41 acquires the listener position information in response to a listener's operation or the like and supplies the information to the viewpoint selection unit 42, the object position calculation unit 48, the object position calculation unit 71, and the polar-coordinate transformation unit 49.

[0286]

In step S53, the viewpoint selection unit 42 selects three 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 the viewpoint selection information on the selection result to the communication unit 111.

[0287]

For example, in step S53, three reference viewpoints surrounding the listening position indicated by the listener position information are selected from a plurality of reference viewpoints indicated by the system configuration information.

[0288]

Moreover, the communication unit 111 requests the start of the transmission of a bit stream from the server 11.

[0289]

In step S54, the communication unit 111 transmits, to the server 11, the viewpoint selection information supplied from the viewpoint selection unit 42.

[0290]

In response to the transmission, the bit stream generated in step S20 of Fig. 18 is transmitted from the server 11.

[0291]

In step S55, the communication unit 111 receives the bit stream transmitted from the server 11 and supplies the bit stream to the decoding unit 45.

[0292]

In step S56, the client 12 generates the polar-coordinate position information for each object.

[0293]

Referring to the flowchart of Fig. 20, the generation of the polar-coordinate position information in step S56 will be described below.

[0294]

In step S81, the decoding unit 45 extracts data on the absolute-coordinate base interpolation object from the bit stream supplied from the communication unit 111 and decodes the data.

[0295]

The decoding unit 45 supplies the decoded object polar-coordinate position information on the absolute-coordinate base interpolation object to the coordinate transformation unit 46 and supplies decoded gain information on the absolute-coordinate base interpolation object to the object position calculation unit 48.

[0296]

Moreover, the decoding unit 45 supplies decoded audio data on the absolute-coordinate base interpolation object to the polar-coordinate transformation unit 49.

[0297]

In step S82, the coordinate transformation unit 46 performs coordinate transformation on the supplied object polar-coordinate position information on the absolute-coordinate base interpolation object from the decoding unit 45 and supplies the obtained object absolute-coordinate position information to the coordinate-axis transformation unit 47.

[0298]

This obtains the object absolute-coordinate position information indicating the position of the absolute-coordinate base interpolation object in the absolute coordinate system with the origin located at the position of the reference viewpoint.

[0299]

In step S83, the coordinate-axis transformation unit 47 performs coordinate-axis transformation 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.

[0300]

The coordinate-axis transformation unit 47 performs coordinate-axis transformation on the object absolute-coordinate position information on the absolute-coordinate base interpolation object for each reference viewpoint and supplies the obtained object absolute-coordinate position information on the position of the absolute-coordinate base interpolation object in the common absolute coordinate system to the object position calculation unit 48. In the coordinate-axis transformation, the reference viewpoint information on the reference viewpoints included in the system configuration information, that is, the reference viewpoint position information and the listener orientation information are used.

[0301]

In step S84, the object position calculation unit 48 performs interpolation 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 unit 47, and the gain information supplied from the decoding unit 45.

[0302]

For example, the object position calculation unit 48 performs the same calculations as the foregoing formulas (1) to (6) on the basis of the reference viewpoint position information included in the system configuration information and the listener position information and determines an internal ratio (m, n) and an internal ratio (k, l).

[0303]

The object position calculation unit 48 then performs the same calculations as the foregoing formulas (7) to (14) on the basis of the determined internal ratios (m, n) and (k, l) and the object absolute-coordinate position information on the reference viewpoints, so that the final object absolute-coordinate position information on the absolute-coordinate base interpolation object is obtained by interpolation.

[0304]

The object position calculation unit 48 also performs interpolation on the gain information as on the object absolute-coordinate position information and determines final gain information on the absolute-coordinate base interpolation object.

[0305]

The object position calculation unit 48 supplies the interpolated final object absolute-coordinate position information on the absolute-coordinate base interpolation object and the interpolated gain information to the polar-coordinate transformation unit 49.

[0306]

In step S85, the polar-coordinate transformation unit 49 performs polar-coordinate transformation on the supplied object absolute-coordinate position information on the absolute-coordinate base interpolation object 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 generates the polar-coordinate position information.

[0307]

This obtains the polar-coordinate position information on polar coordinates indicating the relative position of the absolute-coordinate base interpolation object viewed from the listening position.

[0308]

The polar-coordinate transformation unit 49 makes a gain adjustment to the supplied audio data on the absolute-coordinate base interpolation object from the decoding unit 45 on the basis of the supplied gain information on the absolute-coordinate base interpolation object from the object position calculation unit 48.

[0309]

The polar-coordinate transformation unit 49 supplies the obtained polar-coordinate position information and audio data on the absolute-coordinate base interpolation object to the rendering unit 113.

[0310]

In step S86, the decoding unit 45 determines whether the bit stream supplied from the communication unit 111 includes data on the polar-coordinate base fixed object.

[0311]

For example, in step S86, if the data is included in the bit stream or optionally if the value of the flag “NonInterpolatePolarObjFlag” of the system configuration information supplied from the communication unit 111 is “1”, it is determined that data on the polar-coordinate base fixed object is present.

[0312]

If it is determined that data on the polar-coordinate base fixed object is present in step S86, the decoding unit 45 in step S87 extracts the data on the polar-coordinate base fixed object from the bit stream supplied from the communication unit 111 and decodes the data.

[0313]

The decoding unit 45 properly makes a gain adjustment to the decoded audio data on the polar-coordinate base fixed object on the basis of the decoded gain information on the polar-coordinate base fixed object.

[0314]

Moreover, the decoding unit 45 supplies, to the rendering unit 113, the decoded object polar-coordinate position information on the polar-coordinate base fixed object and the gain-adjusted audio data on the polar-coordinate base fixed object. In this case, the object polar-coordinate position information on the polar-coordinate base fixed object is supplied as the polar-coordinate position information on the polar-coordinate base fixed object to the rendering unit 113 without being transformed.

[0315]

After the processing of step S87, the process advances to step S88.

[0316]

If it is determined in step S86 that the polar-coordinate base fixed object is absent, the process advances to step S88.

[0317]

In step S88, the decoding unit 45 determines whether the bit stream supplied from the communication unit 111 includes data on the absolute-coordinate base fixed object.

[0318]

For example, in step S88, if the data is included in the bit stream or optionally if the value of the flag “NonInterpolateCartesianObjFlag” of the system

configuration information supplied from the communication unit 111 is “1”, it is determined that data on the absolute-coordinate base fixed object is present.

[0319]

If it is determined that data on the absolute-coordinate base fixed object is present in step S88, the decoding unit 45 in step S89 extracts the data on the absolute-coordinate base fixed object from the bit stream supplied from the communication unit 111 and decodes the data.

[0320]

The decoding unit 45 supplies the decoded object absolute-coordinate position information on the absolute-coordinate base fixed object, the gain information, and the audio data to the polar-coordinate transformation unit 49.

[0321]

In step S90, the polar-coordinate transformation unit 49 performs polar-coordinate transformation on the supplied object absolute-coordinate position information on the absolute-coordinate base fixed object from the decoding unit 45 on the basis of the listener position information supplied from the listener-position information acquisition unit 41 and generates the polar-coordinate position information.

[0322]

This obtains the polar-coordinate position information on polar coordinates indicating the relative position of the absolute-coordinate base fixed object viewed from the listener (the listening position).

[0323]

The polar-coordinate transformation unit 49 makes a gain adjustment to the supplied audio data on the absolute-coordinate base fixed object from the decoding unit 45 on the basis of the supplied gain information on the absolute-coordinate base fixed object from the decoding unit 45.

[0324]

The polar-coordinate transformation unit 49 supplies the obtained polar-coordinate position information and audio data on the absolute-coordinate base fixed object to the rendering unit 113.

[0325]

After the processing of step S90, the process advances to step S91.

[0326]

If it is determined in step S88 that the absolute-coordinate base fixed object is absent, the process advances to step S91.

[0327]

In step S91, the decoding unit 45 determines whether the bit stream supplied from the communication unit 111 includes data on the polar-coordinate base interpolation object.

[0328]

For example, in step S91, if the data is included in the bit stream or optionally if

the value of the flag “InterpolatePolarObjFlag” of the system configuration information supplied from the communication unit 111 is “1”, it is determined that data on the polar-coordinate base interpolation object is present.

[0329]

If it is determined that data on the polar-coordinate base interpolation object is present in step S91, the decoding unit 45 in step S92 extracts the data on the polar-coordinate base interpolation object from the bit stream supplied from the communication unit 111 and decodes the data.

[0330]

The decoding unit 45 supplies the decoded object polar-coordinate position information on the polar-coordinate base interpolation object, the gain information, and the audio data to the object position calculation unit 71.

[0331]

In step S93, the object position calculation unit 71 performs interpolation 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 supplied object polar-coordinate position information on the polar-coordinate base interpolation object from the decoding unit 45, and the gain information supplied from the decoding unit 45.

[0332]

For example, the object position calculation unit 71 performs the same calculations as the foregoing formulas (1) to (3) on the basis of the reference viewpoint position information included in the system configuration information and the listener position information and determines an internal ratio (m , n).

[0333]

The object position calculation unit 71 then performs the same calculations as the foregoing formulas (15) and (16) on the basis of the determined internal ratio (m , n), the reference viewpoint position information, the object polar-coordinate position information on the polar-coordinate base interpolation object at each reference viewpoint, and the listener position information, so that the polar-coordinate position information on the polar-coordinate base interpolation object is determined by interpolation.

[0334]

The object position calculation unit 71 also performs interpolation on the gain information as on the polar-coordinate position information and makes a gain adjustment to the audio data on the polar-coordinate base interpolation object on the basis of the obtained final gain information on the polar-coordinate base interpolation object.

[0335]

The object position calculation unit 71 supplies the polar-coordinate position information and audio data on the polar-coordinate base interpolation object to the rendering unit 113, the position information and audio data being obtained

by interpolation and a gain adjustment.

[0336]

After the processing of step S93, the generation of the polar-coordinate position information is completed, and then the process advances to step S57 of Fig. 19.

[0337]

If it is determined in step S91 that the polar-coordinate base interpolation object is absent, the generation of the polar-coordinate position information is completed without performing the processing of step S92 and step S93. The process then advances to step S57 of Fig. 19.

[0338]

Referring to the description of the flowchart of Fig. 19 again, the rendering unit 113 in step S57 performs rendering such as VBAP on the basis of the supplied polar-coordinate position information and audio data on each object and outputs the obtained reproduction audio data.

[0339]

In step S57, rendering is performed on the basis of the polar-coordinate position information and audio data on the absolute-coordinate base interpolation object and the absolute-coordinate base fixed object from the polar-coordinate transformation unit 49, the polar-coordinate position information and audio data on the polar-coordinate base fixed object from the decoding unit 45, and the polar-coordinate position information and audio data on the polar-coordinate base interpolation object from the object position calculation unit 71.

[0340]

For example, through a speaker in the subsequent stage of the rendering unit 113, a sound of contents is reproduced on the basis of the reproduction audio data.

[0341]

In step S58, the client 12 determines whether to terminate the current processing. For example, in step S58, in response to an instruction to stop the reproduction of contents from the listener or at the reproduction of all the pieces of received data on contents, it is determined that the processing is to be terminated.

[0342]

If it is determined in step S58 that the processing is not to be terminated, the process returns to step S52 to repeat the foregoing processing.

[0343]

If it is determined in step S58 that the processing is to be terminated, the client 12 finishes the session with the server 11 and stops the processing performed in the units, so that the generation of the reproduction audio data is terminated.

[0344]

As described above, the client 12 performs proper processing on all kinds of objects included in the received bit stream and generates the reproduction audio

data. This can reproduce the contents on the basis of the intention of a contents producer, allowing the listener to sufficiently feel attraction to the contents.

[0345]

As described above, the present technique combines an object flexibly located according to the position of a listener and a listener-oriented object fixed regardless of the position of the listener and encodes data on all kinds of objects so as to efficiently transmit and reproduce the data.

[0346]

This can represent an object position determined by interpolation at any position in an object layout that is designed by Audio Artistic Intent in a three-dimensional space according to the intention of a contents producer. The present technique can also achieve an object always fixed regardless of the orientation of the listener and an object disposed at an absolute position in a free viewpoint space.

[0347]

Thus, the present technique can achieve applications such as a guide support system that depends upon, for example, the acoustic AR or the arrival direction of sound and achieve a world of contents reproduction in the form of Artistic Intent, in which objects are disposed at proper positions by interpolation, in a free viewpoint space including fixed background noise.

[0348]

<Configuration Example of Computer>

The series of processing can be performed by hardware or software. If the series of processing is performed by software, a program constituting the software is installed on a computer. In this case, the computer includes, for example, a computer built in dedicated hardware and a general-purpose personal computer on which various programs are installed to be able to execute various functions.

[0349]

Fig. 21 is a block diagram illustrating a hardware configuration example of a computer that executes a program to perform the series of processing.

[0350]

In the computer, a central processing unit (CPU) 501, read-only memory (ROM) 502, and random access memory (RAM) 503 are connected to one another via a bus 504.

[0351]

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.

[0352]

The input unit 506 includes a keyboard, a mouse, a microphone, and an imaging element. The output unit 507 includes a display and a speaker. The recording unit 508 includes a hard disk and a nonvolatile memory. The communication

unit 509 includes a network interface. The drive 510 drives a removable recording medium 511, e.g., a magnetic disk, an optical disc, a magneto-optical disk, or a semiconductor memory.

[0353]

In the computer configured thus, the CPU 501 loads, for example, a program recorded in the recording unit 508 into the RAM 503 through the input/output interface 505 and the bus 504 and executes the program, so that the series of processing is performed.

[0354]

The program to be executed by the computer (the CPU 501) can be provided in such a manner as to be recorded on, for example, the removable recording medium 511 serving as a packaged medium. The program can also be provided through a wired or wireless transmission medium, e.g., a local area network, the Internet, or digital satellite broadcasting.

[0355]

In the computer, the program can be installed on the recording unit 508 through the input/output interface 505 by loading the removable recording medium 511 into the drive 510. Furthermore, the program can be received by the communication unit 509 through a wired or wireless transfer medium and installed on the recording unit 508. In addition, the program can be installed in advance on the ROM 502 or the recording unit 508.

[0356]

Note that the program executed by a computer may be a program that performs processing chronologically in the order described in the present specification or may be a program that performs processing in parallel or at a necessary timing, e.g., a calling time.

[0357]

Embodiments of the present technique are not limited to the foregoing embodiment and can be changed in various manners without departing from the gist of the present technique.

[0358]

For example, the present technique may be configured as cloud computing in which one function is shared and cooperatively processed by a plurality of devices via a network.

[0359]

In addition, the steps described in the flowchart can be executed by one device or a plurality of devices in a shared manner.

[0360]

Furthermore, if one step includes a plurality of processes, the plurality of processes included in the one step can be executed by one device or a plurality of devices in a shared manner.

[0361]

The present technique can also be configured as follows.

[0362]

(1) An information processing device including: a listener-position information acquisition unit that acquires listener position information on the viewpoint of a listener;

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

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

(2) The information processing device according to (1), wherein the first reference viewpoint and the second reference viewpoint are viewpoints set in advance by a contents producer.

(3) The information processing device according to (1) or (2), wherein the first reference viewpoint and the second reference viewpoint are viewpoints selected on the basis of the listener position information.

(4) The information processing device according to any one of (1) to (3), wherein the object position calculation unit calculates the position information on the first object at the viewpoint of the listener by interpolation.

(5) The information processing device according to (4), wherein the object position information on the first object at the first reference viewpoint is coordinate information indicating the relative position of the first object with respect to the first reference viewpoint, and

the object position information on the first object at the second reference viewpoint is coordinate information indicating the relative position of the first object with respect to the second reference viewpoint.

(6) The information processing device according to (5), wherein the object position information on the first object is information indicating a position represented by polar coordinates.

(7) The information processing device according to (6), wherein the object position calculation unit performs the interpolation on the basis of the polar-coordinate object position information on the first object.

(8) The information processing device according to any one of (4) to (6), wherein the object position calculation unit transforms the polar-coordinate object position information on the first object into object absolute-coordinate position information indicating the absolute position of the first object in a common

absolute coordinate space, and performs the interpolation on the basis of the object absolute-coordinate position information represented by absolute coordinates.

(9) The information processing device according to (4), wherein for at least three reference viewpoints including the first reference viewpoint and the second reference viewpoint, the reference-viewpoint information acquisition unit acquires the position information on the reference viewpoints and the object position information on the first object at the reference viewpoints, and the object position calculation unit performs the interpolation on the basis of the listener position information, the position information on three of the reference viewpoints, and the object position information on the first object at the three reference viewpoints.

(10) The information processing device according to any one of (1) to (9), wherein the object position information on the second object is coordinate information indicating the relative position of the second object with respect to the position of the listener.

(11) The information processing device according to any one of (1) to (9), wherein the object position information on the second object is coordinate information indicating the absolute position of the second object in the common absolute coordinate space, and the object position calculation unit transforms the object position information on the second object into position information on the second object at the viewpoint of the listener.

(12) The information processing device according to any one of (1) to (11), wherein the object position calculation unit calculates position information on the first object at the viewpoint of the listener on the basis of the listener position information, the position information on the first reference viewpoint, the object position information on the first object at the first reference viewpoint, listener orientation information on the set orientation of the face of the listener at the first reference viewpoint, the position information on the second reference viewpoint, the object position information on the first object at the second reference viewpoint, and the listener orientation information at the second reference viewpoint.

(13) The information processing device according to (12), wherein the reference-viewpoint information acquisition unit acquires configuration information including the position information and the listener orientation information on at least three reference viewpoints including the first reference viewpoint and the second reference viewpoint.

(14) The information processing device according to (13), wherein the configuration information includes information on the number of reference viewpoints and information on the number of first objects and second objects.

(15) An information processing method that causes an information processing

device to:

acquire listener position information on the viewpoint of a listener;

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

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

(16) A program that causes a computer to: acquire listener position information on the viewpoint of a listener;

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

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

[Reference Signs List]

[0363]

- 11 Server
- 12 Client
- 21 Configuration information transmission unit
- 22 Encoded data transmission unit
- 41 Listener-position information acquisition unit
- 42 Viewpoint selection unit
- 45 Decoding unit
- 48 Object position calculation unit
- 71 Object position calculation unit
- 111 Communication unit
- 112 Position calculation unit
- 113 Rendering unit

[CLAIMS]

[Claim 1]

An information processing device comprising: a listener-position information acquisition unit that acquires listener position information on a viewpoint of a listener;

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

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

[Claim 2]

The information processing device according to claim 1, wherein the first reference viewpoint and the second reference viewpoint are viewpoints set in advance by a contents producer.

[Claim 3]

The information processing device 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.

[Claim 4]

The information processing device according to claim 1, wherein the object position calculation unit calculates the position information on the first object at the viewpoint of the listener by interpolation.

[Claim 5]

The information processing device according to claim 4, wherein the object position information on the first object at the first reference viewpoint is coordinate information indicating a relative position of the first object with respect to the first reference viewpoint, and the object position information on the first object at the second reference viewpoint is coordinate information indicating a relative position of the first object with respect to the second reference viewpoint.

[Claim 6]

The information processing device according to claim 5, wherein the object position information on the first object is information indicating a position represented by polar coordinates.

[Claim 7]

The information processing device according to claim 6, wherein the object

position calculation unit performs the interpolation on the basis of the polar-coordinate object position information on the first object.

[Claim 8]

The information processing device according to claim 4, wherein the object position calculation unit transforms the polar-coordinate object position information on the first object into object absolute-coordinate position information indicating an absolute position of the first object in a common absolute coordinate space, and performs the interpolation on a basis of the object absolute-coordinate position information represented by absolute coordinates.

[Claim 9]

The information processing device according to claim 4, wherein for at least three reference viewpoints including the first reference viewpoint and the second reference viewpoint, the reference-viewpoint information acquisition unit acquires the position information on the reference viewpoints and the object position information on the first object at the reference viewpoints, and the object position calculation unit performs the interpolation on a basis of the listener position information, the position information on three of the reference viewpoints, and the object position information on the first object at the three reference viewpoints.

[Claim 10]

The information processing device according to claim 1, wherein the object position information on the second object is coordinate information indicating a relative position of the second object with respect to the position of the listener.

[Claim 11]

The information processing device according to claim 1, wherein the object position information on the second object is coordinate information indicating an absolute position of the second object in a common absolute coordinate space, and the object position calculation unit transforms the object position information on the second object into position information on the second object at the viewpoint of the listener.

[Claim 12]

The information processing device according to claim 1, wherein the object position calculation unit calculates position information on the first object at the viewpoint of the listener on a basis of the listener position information, the position information on the first reference viewpoint, the object position information on the first object at the first reference viewpoint, listener orientation information on a set orientation of a face of the listener at the first reference viewpoint, the position information on the second reference viewpoint, the object position information on the first object at the second reference viewpoint, and the listener orientation information at the second reference viewpoint.

[Claim 13]

The information processing device according to claim 12, wherein the reference-viewpoint information acquisition unit acquires configuration information including the position information and the listener orientation information on at least three reference viewpoints including the first reference viewpoint and the second reference viewpoint.

[Claim 14]

The information processing device according to claim 13, wherein the configuration information includes information on the number of reference viewpoints and information on the number of first objects and second objects.

[Claim 15]

An information processing method that causes an information processing device to:

acquire listener position information on a viewpoint of a listener;

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

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

[Claim 16]

A program that causes a computer to: acquire listener position information on a viewpoint of a listener;

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

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