[NPL 3]
ISO/IEC JTC1/SC29/WG11 N14747, August 2014, Sapporo, Japan, "Text of ISO/IEC 23008-3/DIS, 3D Audio"

[Summary]

[Technical Problem]

[0007]
However, the technology described above fails to obtain sound of sufficiently high quality.

[0008]
For example, in the MPEG-H 3D Audio standard, information indicative of a degree of extent of a sound image called spread is included in metadata of an audio object and a process for extending a sound image is performed on the basis of the spread. However, in the process for extending a sound image, there is a constraint that the extent of a sound image is symmetrical in the upward and downward direction and the leftward and rightward direction with respect to the center at the position of the audio object. Therefore, a process that takes a directionality (radial direction) of sound from the audio object into consideration cannot be performed and sound of sufficiently high quality cannot be obtained.

[0009]
Embodiments of the present disclosure make it possible to obtain sound of higher quality.

[Solution to Problem]

[0009a]
According to a first aspect of the present disclosure, there is provided an audio processing apparatus including: an acquisition unit configured to acquire metadata including position information indicative of a position of an audio object and sound image information configured from a vector of three dimensions and representative of an extent of a sound

image from the position; a vector calculation unit configured to: determine a three-dimensional vector representative of the position of the audio object, wherein a start point of the three-dimensional vector is the origin in a three-dimensional coordinate system whose origin is given by a position of a user; and calculate, based on the three-dimensional vector and a ratio between a horizontal direction angle and a vertical direction angle of a region representative of the extent of the sound image determined by the sound image information, at least one spread vector indicative of a position in the region, wherein the number of spread vectors is determined in advance; and a gain calculation unit configured to calculate, based on the at least one spread vector, a gain of each of audio signals supplied to two or more sound outputting units positioned in the proximity of the position indicated by the position information.

[0010]

Disclosed herein is an audio processing apparatus that includes an acquisition unit configured to acquire metadata including position information indicative of a position of an audio object and sound image information configured from a vector of at least two or more dimensions and representative of an extent of a sound image from the position, a vector calculation unit configured to calculate, based on a horizontal direction angle and a vertical direction angle of a region representative of the extent of the sound image determined by the sound image information, a spread vector indicative of a position in the region, and a gain calculation unit configured to calculate, based on the spread vector, a gain of each of audio signals supplied to two or more sound outputting units positioned in the proximity of the position indicated by the position information.

[0011]

The vector calculation unit may calculate the spread vector

based on a ratio between the horizontal direction angle and the vertical direction angle.

[0012]

The vector calculation unit may calculate the number of spread vectors determined in advance.

[0013]

The vector calculation unit may calculate a variable arbitrary number of spread vectors.

[0014]

The sound image information may be a vector indicative of a center position of the region.

[0015]

The sound image information may be a vector of two or more dimensions indicative of an extent degree of the sound image from the center of the region.

[0016]

The sound image information may be a vector indicative of a relative position of a center position of the region as viewed from a position indicated by the position information.

[0017]

The gain calculation unit may calculate, the gain for each spread vector in regard to each of the sound outputting units, calculate an addition value of the gains calculated in regard to the spread vectors for each of the sound outputting units, quantize the addition value into a gain of two or more values for each of the sound outputting units, and calculate a final gain for each of the sound outputting units based on the quantized addition value.

[0018]

The gain calculation unit may select the number of meshes each of which is a region surrounded by three ones of the sound outputting units and which number is to be used for calculation of the gain and calculate the gain for each of the spread vectors based on a result of the selection of the

number of meshes and the spread vector.

[0019]

The gain calculation unit may select the number of meshes to be used for calculation of the gain, whether or not the quantization is to be performed and a quantization number of the addition value upon the quantization and calculate the final gain in response to a result of the selection.

[0020]

The gain calculation unit may select, based on the number of the audio objects, the number of meshes to be used for calculation of the gain, whether or not the quantization is to be performed and the quantization number.

[0021]

The gain calculation unit may select, based on an importance degree of the audio object, the number of meshes to be used for calculation of the gain, whether or not the quantization is to be performed and the quantization number.

[0022]

The gain calculation unit may select the number of meshes to be used for calculation of the gain such that the number of meshes to be used for calculation of the gain increases as the position of the audio object is positioned nearer to the audio object that is high in the importance degree.

[0023]

The gain calculation unit may select, based on a sound pressure of the audio signal of the audio object, the number of meshes to be used for calculation of the gain, whether or not the quantization is to be performed and the quantization number.

[0024]

The gain calculation unit may select, in response to a result of the selection of the number of meshes, three or more ones of the plurality of sound outputting units including the sound outputting units that are positioned at different heights from

each other, and calculate the gain based on one or a plurality of meshes formed from the selected sound outputting units.

[0024a]

According to a second aspect of the present disclosure, there is provided an audio processing method including: acquiring metadata including position information indicative of a position of an audio object and sound image information configured from a vector of three dimensions and representative of an extent of a sound image from the position; determining a three-dimensional vector representative of the position of the audio object, wherein a start point of the three-dimensional vector is the origin in a three-dimensional coordinate system whose origin is given by a position of a user; and calculating, based on the three-dimensional vector and a ratio between a horizontal direction angle and a vertical direction angle of a region representative of the extent of the sound image determined by the sound image information, at least one spread vector indicative of a position in the region, wherein the number of spread vectors is determined in advance; and calculating, based on the at least one spread vector, a gain of each of audio signals supplied to two or more sound outputting units positioned in the proximity of the position indicated by the position information.

[0024b]

According to a third aspect of the present disclosure, there is provided a non-transitory computer readable storage medium having stored thereon computer readable instructions, which when executed by a computer, cause the computer to execute operations, the operations including: acquiring metadata including position information indicative of a position of an audio object and sound image information configured from a vector of three dimensions and representative of an extent of a sound image from the position; determining a three-

dimensional vector representative of the position of the audio object, wherein a start point of the three-dimensional vector is the origin in a three-dimensional coordinate system whose origin is given by a position of a user; and calculating, based on the three-dimensional vector and a ratio between a horizontal direction angle and a vertical direction angle of a region representative of the extent of the sound image determined by the sound image information, at least one spread vector indicative of a position in the region, wherein the number of spread vectors is determined in advance; and calculating, based on the at least one spread vector, a gain of each of audio signals supplied to two or more sound outputting units positioned in the proximity of the position indicated by the position information.

[0025]

Disclosed herein is an audio processing method or a program that includes the steps of acquiring metadata including position information indicative of a position of an audio object and sound image information configured from a vector of at least two or more dimensions and representative of an extent of a sound image from the position, calculating, based on a horizontal direction angle and a vertical direction angle of a region representative of the extent of the sound image determined by the sound image information, a spread vector indicative of a position in the region, and calculating, based on the spread vector, a gain of each of audio signals supplied to two or more sound outputting units positioned in the proximity of the position indicated by the position information.

[0026]

As per embodiments of the present disclosure, metadata including position information indicative of an audio object and sound image information configured from a vector of at least two or more dimensions and representative of an extent

of a sound image from the position is acquired. Then, based on a horizontal direction angle and a vertical direction angle regarding a region representative of the extent of the sound image determined by the sound image information, a spread vector indicative of a position in the region is calculated. Further, based on the spread vector, a gain of each of audio signals supplied to two or more sound outputting units positioned in the proximity of the position indicated by the position information is calculated.

[Advantageous Effect of Invention]

[0027]

As per embodiments of the present disclosure, sound of higher quality can be obtained.

[0028]

It is to be noted that the effect described here is not necessarily limitative, but any of effects described in the present disclosure may be exhibited.

[Brief Description of Drawings]

[0029]

[FIG. 1]

FIG. 1 is a view illustrating VBAP.

[FIG. 2]

FIG. 2 is a view illustrating a position of a sound image.

[FIG. 3]

FIG. 3 is a view illustrating a spread vector.

[FIG. 4]

FIG. 4 is a view illustrating a spread center vector method.

[FIG. 5]

Claims:

1. An audio processing apparatus comprising:

an acquisition unit configured to acquire metadata including position information indicative of a position of an audio object and sound image information configured from a vector of three dimensions and representative of an extent of a sound image from the position;

a vector calculation unit configured to:

determine a three-dimensional vector representative of the position of the audio object, wherein a start point of the three-dimensional vector is the origin in a three-dimensional coordinate system whose origin is given by a position of a user; and

calculate, based on the three-dimensional vector and a ratio between a horizontal direction angle and a vertical direction angle of a region representative of the extent of the sound image determined by the sound image information, at least one spread vector indicative of a position in the region, wherein the number of spread vectors is determined in advance; and

a gain calculation unit configured to calculate, based on the at least one spread vector, a gain of each of audio signals supplied to two or more sound outputting units positioned in the proximity of the position indicated by the position information.

2. An audio processing method comprising:

acquiring metadata including position information indicative of a position of an audio object and sound image information configured from a vector of three dimensions and representative of an extent of a sound image from the position;

determining a three-dimensional vector representative of the position of the audio object, wherein a start point of the three-dimensional vector is the origin in a three-dimensional coordinate system whose origin is given by a position of a user; and

calculating, based on the three-dimensional vector and a ratio between a horizontal direction angle and a vertical direction angle of a region representative of the extent of the sound image determined by the sound image information, at least one spread vector indicative of a position in the region, wherein the number of spread vectors is determined in advance; and

calculating, based on the at least one spread vector, a gain of each of audio signals supplied to two or more sound outputting units positioned in the proximity of the position indicated by the position information.

3. A non-transitory computer readable storage medium having stored thereon computer readable instructions, which when executed by a computer, cause the computer to execute operations, the operations comprising:

acquiring metadata including position information indicative of a position of an audio object and sound image information configured from a vector of three dimensions and representative of an extent of a sound image from the position;

determining a three-dimensional vector representative of the position of the audio object, wherein a start point of the three-dimensional vector is the origin in a three-dimensional coordinate system whose origin is given by a position of a user; and

calculating, based on the three-dimensional vector and a ratio between a horizontal direction angle and a vertical direction angle of a region representative of the extent of the sound image determined by the sound image information, at least one spread vector indicative of a position in the region, wherein the number of spread vectors is determined in advance; and

calculating, based on the at least one spread vector, a gain of each of audio signals supplied to two or more sound outputting units positioned in the proximity of the position indicated by the position information.