



(12) 发明专利

(10) 授权公告号 CN 112562697 B

(45) 授权公告日 2024. 11. 08

(21) 申请号 202011538529.0

(22) 申请日 2016.06.09

(65) 同一申请的已公布的文献号  
申请公布号 CN 112562697 A

(43) 申请公布日 2021.03.26

(30) 优先权数据  
2015-126650 2015.06.24 JP  
2015-148683 2015.07.28 JP

(62) 分案原申请数据  
201680034827.1 2016.06.09

(73) 专利权人 索尼公司  
地址 日本东京都

(72) 发明人 山本优树 知念彻 辻实

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

专利代理师 王萍 王鹏

(51) Int.Cl.  
G10L 19/008 (2013.01)  
H04S 7/00 (2006.01)  
H04S 5/02 (2006.01)

(56) 对比文件  
CN 103650535 A, 2014.03.19

审查员 毛健

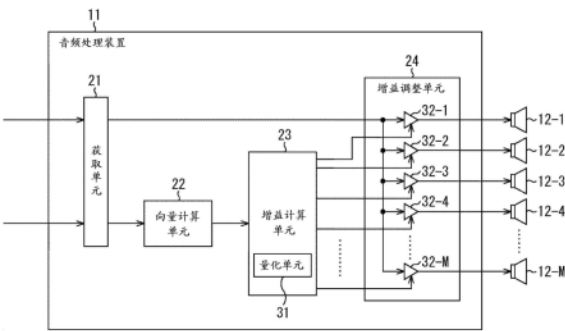
权利要求书2页 说明书36页 附图20页

(54) 发明名称

音频处理装置和方法以及计算机可读存储介质

(57) 摘要

本技术涉及能够获取更高质量的声音的用于处理声音的装置、方法和程序。该获取单元获取对象的音频信号和元数据。该向量计算单元基于指示声像的扩散且包括在对象的元数据中的水平方向角和竖直方向角来计算表示指示声像的扩散的区域内的位置的扩展向量,其中扩展向量的数量是预先确定的并且不依赖于声像的扩散。该增益计算单元基于扩展向量通过VBAP来计算每个扬声器的音频信号的VBAP增益。该特征可以应用于声音处理装置。



1. 一种音频处理装置,包括:

获取单元,被配置成获取包括指示音频对象的位置的位置信息以及由三维的向量构成并且表示声像距所述位置的扩散的声像信息的元数据;

向量计算单元,被配置成:

确定表示所述音频对象的位置的三维向量,其中,所述三维向量的起点是三维坐标系中的原点,所述三维坐标系的原点由用户的位置给定;以及

基于所述三维向量以及表示由所述声像信息确定的声像的扩散的区域的水平方向角和竖直方向角之间的比率来计算指示所述区域中的位置的至少一个扩展向量,其中所述扩展向量的数量是预先确定的;以及

增益计算单元,被配置成基于所述至少一个扩展向量来计算提供给位于所述位置信息指示的位置附近的两个或更多个声音输出单元的音频信号中的每个音频信号的增益,

其中,所述增益计算单元:

计算关于所述声音输出单元中的每个声音输出单元的每个扩展向量的增益,

计算关于所述声音输出单元中的每个声音输出单元的扩展向量所计算的增益的相加值,

将所述相加值归一化,以及

基于归一化的相加值来计算所述声音输出单元中的每个声音输出单元的最终增益。

2. 一种音频处理方法,包括以下步骤:

获取包括指示音频对象的位置的位置信息以及由三维的向量构成并且表示声像距所述位置的扩散的声像信息的元数据;

确定表示所述音频对象的位置的三维向量,其中,所述三维向量的起点是三维坐标系中的原点,所述三维坐标系的原点由用户的位置给定;以及

基于所述三维向量以及表示由所述声像信息确定的声像的扩散的区域的水平方向角和竖直方向角之间的比率来计算指示所述区域中的位置的至少一个扩展向量,其中所述扩展向量的数量是预先确定的;以及

基于所述至少一个扩展向量来计算提供给位于所述位置信息指示的位置附近的两个或更多个声音输出单元的音频信号中的每个音频信号的增益,

其中,计算所述增益包括:

计算关于所述声音输出单元中的每个声音输出单元的每个扩展向量的增益,

计算关于所述声音输出单元中的每个声音输出单元的扩展向量所计算的增益的相加值,

将所述相加值归一化,以及

基于归一化的相加值来计算所述声音输出单元中的每个声音输出单元的最终增益。

3. 一种计算机可读存储介质,其上存储有使计算机执行以下处理的程序,所述处理包括以下步骤:

获取包括指示音频对象的位置的位置信息以及由三维的向量构成并且表示声像距所述位置的扩散的声像信息的元数据;

确定表示所述音频对象的位置的三维向量,其中,所述三维向量的起点是三维坐标系中的原点,所述三维坐标系的原点由用户的位置给定;以及

基于所述三维向量以及表示由所述声像信息确定的声像的扩散的区域的水平方向角和竖直方向角之间的比率来计算指示所述区域中的位置的至少一个扩展向量,其中所述扩展向量的数量是预先确定的;以及

基于所述至少一个扩展向量来计算提供给位于所述位置信息指示的位置附近的两个或更多个声音输出单元的音频信号中的每个音频信号的增益,

其中,计算所述增益包括:

计算关于所述声音输出单元中的每个声音输出单元的每个扩展向量的增益,

计算关于所述声音输出单元中的每个声音输出单元的扩展向量所计算的增益的相加值,

将所述相加值归一化,以及

基于归一化的相加值来计算所述声音输出单元中的每个声音输出单元的最终增益。

## 音频处理装置和方法以及计算机可读存储介质

[0001] 本发明申请是申请日期为2016年6月9日、申请号为“201680034827.1”、发明名称为“用于处理声音的装置、方法及程序”的发明专利申请的分案申请。

### 技术领域

[0002] 本技术涉及音频处理装置、方法及程序,并且更具体地涉及通过其能够获得更高质量的聲音的音频处理装置、方法及程序。

### 背景技术

[0003] 通常,作为使用多个扬声器来控制声像的定位的技术,已知VBAP(基于向量的幅值相移)(例如,参考NPL 1)。

[0004] 在VBAP中,通过从三个扬声器输出声音,声像可以被定位在由这三个扬声器限定的三角形的内侧的一个任意点处。

[0005] 然而,在现实世界中,认为声像不是定位于一个点处,而是定位于具有一定扩散程度的部分空间中。例如,当从声带产生人声时,认为声音的振动传播到脸部、身体等,因此认为声音是从作为整个人体的部分空间发出的。

[0006] 作为如上所述将声音定位在这种部分空间中的技术,即,作为延伸声像的技术,通常已知MDAP(多方向幅值相移)(例如,参考NPL 2)。此外,MDAP也用于MPEG-H 3D(运动图像专家组-高质量三维)音频标准的渲染处理单元(例如,参考NPL 3)。

[0007] [引用列表]

[0008] [非专利文献(NPL)]

[0009] [NPL 1]

[0010] Ville Pulkki,“Virtual Sound Source Positioning Using Vector Base Amplitude Panning”,美国电声学协会期刊,第45卷,第6期,第456-466页,1997年

[0011] [NPL 2]

[0012] Ville-Pulkki,“Uniform Spreading of Amplitude Panned Virtual Sources”,会议记录1999,关于信号处理对于音频和声学的应用的IEEE研讨会,新帕尔茨,纽约,1999年10月17-20日

[0013] [NPL 3]

[0014] ISO/IEC JTC1/SC29/WG11 N14747,2014年8月,札幌,日本,“Text of ISO/IEC 23008-3/DIS,3D Audio”

### 发明内容

[0015] [技术问题]

[0016] 然而,上述技术不能获得足够高质量的声音。

[0017] 例如,在MPEG-H 3D音频标准中,指示声像的扩散程度的信息(被称为扩展)包括在音频对象的元数据中,并且基于该扩展执行用于延伸声像的处理。然而,在延伸声像的处理

中,存在声像的扩散相对于音频对象的位置的中心在上下方向和左右方向上对称的约束。因此,不能执行考虑到来自音频对象的声音的方向性(径向方向)的处理,从而不能获得足够高质量的声音。

[0018] 鉴于上述情况作出了本技术,使得可以获得更高质量的声音。

[0019] [问题的解决方案]

[0020] 根据本技术的一个方面的音频处理装置包括:获取单元,被配置成获取包括指示音频对象的位置的位置信息和由至少二维或更多维的向量构成并表示声像距该位置的扩散的声像信息的元数据;向量计算单元,被配置成基于表示由声像信息确定的声像的扩散的区域的水平方向角和竖直方向角来计算指示该区域中的位置的扩展向量;以及增益计算单元,被配置成基于扩展向量来计算提供给位于由位置信息指示的位置附近的两个或更多个声音输出单元的音频信号中的每个音频信号的增益。

[0021] 向量计算单元可以基于水平方向角与竖直方向角之间的比率来计算扩展向量。

[0022] 向量计算单元可以计算预先确定的数量的扩展向量。

[0023] 向量计算单元可以计算任意可变数量的扩展向量。

[0024] 声像信息可以是指示该区域的中心位置的向量。

[0025] 声像信息可以是指示声像距该区域的中心的扩散程度的二维或更多维的向量。

[0026] 声像信息可以是指示从位置信息指示的位置观看到的区域的中心位置的相对位置的向量。

[0027] 增益计算单元可以计算关于声音输出单元中的每个声音输出单元的每个扩展向量的增益,计算针对声音输出单元中的每个声音输出单元的扩展向量所计算的增益的相加值,将相加值量化成关于声音输出单元中的每个声音输出单元的两个或更多个值的增益,以及基于量化的相加值来计算关于声音输出单元中的每个声音输出单元的最终增益。

[0028] 增益计算单元可以选择要用于计算增益网格的数量,网格中的每个网格是由声音输出单元中的三个声音输出单元包围的区域,并且增益计算单元基于对网格的数量的选择结果和扩展向量来计算扩展向量中的每个扩展向量的增益。

[0029] 增益计算单元可以选择要用于计算增益的网格的数量、是否要执行量化以及量化时的所述相加值的量化数量,并且响应于选择结果来计算最终增益。

[0030] 增益计算单元可以基于音频对象的数量来选择要用于计算增益的网格的数量、是否要执行量化以及量化数量。

[0031] 增益计算单元可以基于音频对象的重要性程度来选择要用于计算增益的网格的数量、是否要执行量化以及量化数量。

[0032] 增益计算单元可以选择要用于计算增益的网格的数量,使得要用于计算增益的网格的数量随着音频对象的位置位于更靠近重要性程度高的音频对象而增加。

[0033] 增益计算单元可以基于音频对象的音频信号的声压来选择要用于计算增益的网格的数量、是否要执行量化以及量化数量。

[0034] 增益计算单元可以响应于对网格数量的选择结果来选择包括位于彼此不同高度处的声音输出单元的多个声音输出单元中的三个或更多个声音输出单元,并且基于由所选择的聲音输出单元形成的一个或多个网格来计算增益。

[0035] 根据本技术的一个方面的音频处理方法或程序包括以下步骤:获取包括指示音频

对象的位置的位置信息和由至少二维或更多维的向量构成并且表示声像距该位置的扩散的声像信息的元数据;基于表示由声像信息确定的声像的扩散的区域的水平方向角和竖直方向角来计算指示该区域中的位置的扩展向量;基于扩展向量来计算提供给位于位置信息指示的位置附近的两个或更多个声音输出单元的音频信号中的每个音频信号的增益。

[0036] 在本技术的一个方面中,获取包括指示音频对象的位置的位置信息和由至少二维或更多维的向量构成并且表示声像距该位置的扩散的声像信息的元数据。然后,基于表示由声像信息确定的声像的扩散的区域的水平方向角和竖直方向角来计算指示该区域中的位置的扩展向量。此外,基于扩展向量来计算提供给位于位置信息指示的位置附近的两个或更多个声音输出单元的音频信号中的每个音频信号的增益。

[0037] 根据本技术的一个方面的音频处理装置包括:获取单元,被配置成获取包括指示音频对象的位置的位置信息和由至少二维或更多维的向量构成并表示声像距该位置的扩散的声像信息的元数据;向量计算单元,被配置成基于表示由声像信息确定的声像的扩散的区域的水平方向角和竖直方向角来计算指示该区域中的位置的扩展向量,其中扩展向量的数量是预先确定的并且不依赖于声像的扩散;以及增益计算单元,被配置成基于扩展向量来计算提供给位于由位置信息指示的位置附近的两个或更多个声音输出单元的音频信号中的每个音频信号的增益。

[0038] 根据本技术的一个方面的音频处理方法或程序包括以下步骤:获取包括指示音频对象的位置的位置信息和由至少二维或更多维的向量构成并且表示声像距该位置的扩散的声像信息的元数据;基于表示由声像信息确定的声像的扩散的区域的水平方向角和竖直方向角来计算指示该区域中的位置的扩展向量,其中扩展向量的数量是预先确定的并且不依赖于声像的扩散;基于扩展向量来计算提供给位于位置信息指示的位置附近的两个或更多个声音输出单元的音频信号中的每个音频信号的增益。

[0039] [发明的有益效果]

[0040] 利用本技术的一个方面,可以获得更高质量的声音。

[0041] 应当注意,这里描述的效果不一定是限制性的,而是可以表现出本公开中所描述的效果中的任一效果。

## 附图说明

[0042] 图1是示出了VBAP的图。

[0043] 图2是示出了声像的位置的图。

[0044] 图3是扩展向量的图。

[0045] 图4是示出了扩展中心向量法的图。

[0046] 图5是示出了扩散辐射向量法的图。

[0047] 图6是示出了音频处理装置的配置的示例的图。

[0048] 图7是示出了再现处理的流程图。

[0049] 图8是示出了扩展向量计算处理的流程图。

[0050] 图9是示出了基于扩展三维向量进行的扩展向量计算处理的流程图。

[0051] 图10是示出了基于扩展中心向量进行的扩展向量计算处理的流程图。

[0052] 图11是示出了基于扩展端向量进行的扩展向量计算处理的流程图。

- [0053] 图12是示出了基于扩展辐射向量进行的扩展向量计算处理的流程图。
- [0054] 图13是示出了基于扩展向量位置信息进行的扩展向量计算处理的流程图。
- [0055] 图14是示出了网格的数量的切换的图。
- [0056] 图15是示出了网格的数量的切换的图。
- [0057] 图16是示出了网格的形成的图。
- [0058] 图17是示出了音频处理装置的配置的示例的图。
- [0059] 图18是示出了再现处理的流程图。
- [0060] 图19是示出了音频处理装置的配置的示例的图。
- [0061] 图20是示出了再现处理的流程图。
- [0062] 图21是示出了VBAP增益计算处理的流程图。
- [0063] 图22是示出了计算机的配置的示例的图。

### 具体实施方式

- [0064] 在下文中,参考附图描述应用本技术的实施方式。
- [0065] <第一实施方式>
- [0066] <VBAP和延伸声像的处理>
- [0067] 当获取音频对象的音频信号和诸如音频对象的位置信息的元数据来进行渲染时,本技术使得能够获得更高质量的声音。应当注意,在下面的描述中,音频对象被简称为对象。
- [0068] 首先,下面描述VBAP和以MPEG-H 3D音频标准延伸声像的处理。
- [0069] 例如,如图1所示,假定欣赏具有声音的运动图像、音乐作品等的内容的用户U11正在收听从三个扬声器SP1至SP3输出的三声道声音作为内容的声音。
- [0070] 在刚刚描述的这种情况下,使用输出不同声道的声音的三个扬声器SP1至SP3的位置的信息将声像定位在位置p处得到检验。
- [0071] 例如,位置p由三维向量(在下文中也称为向量p)表示,该三维向量的起点是三维坐标系中的原点0,三维坐标系的原点0由用户U11的头部的的位置给定。此外,如果其起点由原点0给定并且在朝向扬声器SP1至SP3的位置的方向上取向的三维向量分别被表示为向量 $I_1$ 至 $I_3$ ,则向量p可以由向量 $I_1$ 到 $I_3$ 的线性和表示。
- [0072] 换言之,向量p可以被表示为 $p = g_1 I_1 + g_2 I_2 + g_3 I_3$ 。
- [0073] 这里,如果与向量 $I_1$ 至 $I_3$ 相乘的系数 $g_1$ 至 $g_3$ 分别被计算和确定为从扬声器SP1至SP3输出的声音的增益,则声像可以被定位在位置p处。
- [0074] 使用三个扬声器SP1至SP3的位置信息来确定系数 $g_1$ 至 $g_3$ 并且以上述方式控制声像的定位位置的技术被称为三维VBAP。特别地,在下面的描述中,将关于每个扬声器确定的增益例如系数 $g_1$ 至 $g_3$ 称为VBAP增益。
- [0075] 在图1的示例中,声像可以被定位在包括扬声器SP1、SP2和SP3的位置的球体上的三角形形状的区域TR11中的任意位置处。这里,区域TR11是以原点0为中心并且经过扬声器SP1至SP3的位置的球体的表面上的区域,并且是由扬声器SP1至SP3包围的三角形区域。
- [0076] 如果使用这种三维VBAP,则可以将声像定位在空间中的任意位置。应当注意,例如在'Ville Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude

Panning”,美国电化学协会期刊,第45卷,第6期,第456-466页,1997年’等中详细描述了VBAP。

[0077] 现在,描述根据MPEG-H 3D音频标准延伸声像的处理。

[0078] 在MPEG-H 3D音频标准中,从编码装置输出通过以下处理而获得的比特流:对通过对每个对象的音频信号进行编码所获得的编码音频数据和通过对每个对象的元数据进行编码所获得的编码元数据进行复用。

[0079] 例如,元数据包括:指示对象在空间中的位置的位置信息、指示对象的重要性程度的重要性信息、以及作为指示对象的声像的扩散程度的信息的扩展。

[0080] 这里,指示声像的扩散程度的扩展是从0度到180度的任意角度,并且编码装置可以关于每个对象指定对于音频信号的每个帧而值不同的扩展。

[0081] 此外,对象的位置由水平方向角方位角、竖直方向角仰角和距离半径表示。具体地,对象的位置信息由水平方向角方位角、竖直方向角仰角和距离半径的值来配置。

[0082] 例如,考虑三维坐标系,如图2所示,在三维坐标系中,欣赏从未示出的扬声器输出的对象的声音的用户的位置被确定为原点0,并且图2中的右上方向、左上方向和向上方向被确定为彼此竖直的x轴、y轴和z轴。此时,如果一个对象的位置被表示为位置OBJ11,则声像可以被定位在三维坐标系中的位置OBJ11处。

[0083] 此外,如果将位置OBJ11和原点0互连的线性线被表示为线L,则由线性线L和x轴在xy平面上限定的图2中的水平方向上的角度 $\theta$ (方位角)为表示位置OBJ11处的对象在水平方向上的位置的水平方向角方位角,并且水平方向角方位角具有满足以下的任意值: $-180^\circ \leq \text{方位角} \leq 180^\circ$

[0084] 例如,x轴方向的正方向被确定为方位角 $=0^\circ$ 而x轴方向的负方向被确定为方位角 $=+180^\circ$ 或 $=-180^\circ$ 。此外,围绕原点0的逆时针方向被确定为方位角的正方向,而围绕原点0的顺时针方向被确定为方位角的负方向。

[0085] 此外,由线性线L和xy平面限定的角度,即,图2中的竖直方向上的角度 $\gamma$ (仰角)是表示位于位置OBJ11处的对象在竖直方向上的位置的竖直方向角仰角,并且竖直方向角仰角具有满足以下的任意值: $-90^\circ \leq \text{仰角} \leq 90^\circ$ 。例如,xy平面上的位置是仰角 $=0^\circ$ ,以及图2中的向上方向是竖直方向角仰角的正方向,而图2中的向下方向是竖直方向角仰角的负方向。

[0086] 此外,线性线L的长度,即,从原点0到位置OBJ11的距离是到用户的距离半径,并且距离半径具有0或更大的值。具体地,距离半径具有满足以下的值: $0 \leq (\text{半径}) \leq \infty$ 。在下面的描述中,距离半径也称为径向方向的距离。

[0087] 应当注意,在VBAP中,从所有扬声器或对象到用户的距离半径相等,并且将距离半径归一化为1来执行计算是通用方法。

[0088] 以该方式包括在元数据中的对象的位置信息由水平方向角方位角、竖直方向角仰角和距离半径的值来配置。

[0089] 在下面的描述中,水平方向角方位角、竖直方向角仰角和距离半径也分别简称为方位角、仰角和半径。

[0090] 此外,在接收包括编码音频数据和编码元数据的比特流的解码装置中,在执行编码音频数据和编码元数据的解码之后,响应于元数据中所包括的扩展的值来执行用于延伸



声像的渲染处理。

[0091] 具体地,解码装置首先将由对象的元数据中所包括的位置信息指示的空间中的位置确定为位置 $p$ 。如上所述,位置 $p$ 对应于图1中的位置 $p$ 。

[0092] 然后,解码装置设置18个扩展向量 $p_1$ 至 $p_{18}$ ,使得例如将位置 $p$ 设置成位置 $p$ =中心位置 $p_0$ ,如图3所示,这些扩展向量在围绕中心位置 $p_0$ 的单位球体上在上下方向和左右方向上对称。应当注意,在图3中,与图1的情况下的部分对应的部分由相同的附图标记表示,并且将适当地省略对这些部分的描述。

[0093] 在图3中,五个扬声器 $SP_1$ 至 $SP_5$ 布置在以原点 $O$ 为中心半径为1的单位球体的球面上,以及由位置信息指示的位置 $p$ 为中心位置 $p_0$ 。在下面的描述中,位置 $p$ 也被具体地称为对象位置 $p$ ,并且其起点为原点 $O$ 且其终点为对象位置 $p$ 的向量也被称为向量 $p$ 。此外,其起点为原点 $O$ 且其终点为中心位置 $p_0$ 的向量也称为向量 $p_0$ 。

[0094] 在图3中,其起点为原点 $O$ 并且由虚线绘制的箭头标记表示扩展向量。然而,虽然在图3中实际上有18个扩展向量,但是出于图3的可视性仅绘制了八个扩展向量。

[0095] 这里,扩展向量 $p_1$ 至 $p_{18}$ 中的每一个是其终点位置位于以中心位置 $p_0$ 为中心的球面上的圆形的区域 $R_{11}$ 内的向量。具体地,由向量 $p_0$ 和其终点位置位于由区域 $R_{11}$ 表示的圆形的圆周上的扩展向量限定的角度是由扩展指示的角度。

[0096] 因此,当扩展的值增加时,每个扩展向量的终点位置被设置在离中心位置 $p_0$ 较远的位置处。换言之,区域 $R_{11}$ 在尺寸上增加。

[0097] 区域 $R_{11}$ 表示声像距对象的位置的扩散。换言之,区域 $R_{11}$ 是指示对象的声像延伸的扩散的区域。此外,可以认为,由于认为对象的声音是从整个对象发出的,所以区域 $R_{11}$ 表示对象的形状。在下面的描述中,与区域 $R_{11}$ 一样,指示对象的声像延伸的扩散的区域也被称为指示声像的扩散的区域。

[0098] 此外,在扩展的值为0的情况下,18个扩展向量 $p_1$ 至 $p_{18}$ 的终点位置等于中心位置 $p_0$ 。

[0099] 应当注意,在下面的描述中,扩展向量 $p_1$ 至 $p_{18}$ 的终点位置也分别特定地被称为位置 $p_1$ 至 $p_{18}$ 。

[0100] 在如上所述确定在单位球面上沿向上下方向和左右方向对称的扩展向量之后,解码装置关于向量 $p$ 和扩展向量,即关于位置 $p$ 和位置 $p_1$ 至 $p_{18}$ 中的每一个通过VBAP计算声道的扬声器中的每一个的VBAP增益。此时,计算扬声器的VBAP增益,使得声像被定位在诸如位置 $p$ 和位置 $p_1$ 的位置中的每个位置处。

[0101] 然后,解码装置将关于每个扬声器的位置所计算的VBAP增益相加。例如,在图3的示例中,关于扬声器 $SP_1$ 和位置 $p_1$ 至 $p_{18}$ 所计算的位置 $p$ 的VBAP增益被相加。

[0102] 此外,在为各个扬声器计算的加法处理之后,解码装置对VBAP增益进行归一化。具体地,执行归一化,使得所有扬声器的VBAP增益的平方和变成1。

[0103] 然后,解码装置将对象的音频信号乘以通过归一化所获得的扬声器的VBAP增益,以获得各个扬声器的音频信号,并将关于各个扬声器所获得的音频信号提供给扬声器,使得扬声器输出声音。

[0104] 因此,例如,在图3的示例中,声像被定位,使得从整个区域 $R_{11}$ 输出声音。换言之,声像被延伸到整个区域 $R_{11}$ 。

[0105] 在图3中,当不执行用于延伸声像的处理时,对象的声像被定位在位置p处,因此在这种情况下,基本上从扬声器SP2和扬声器SP3输出声音。相反,当执行用于延伸声像的处理时,声像被延伸到整个区域R11,因此当声音再现时,从扬声器SP1至SP4输出声音。

[0106] 顺便提及,当执行如上所述的用于延伸声像的这种处理时,与不执行延伸声像的处理的替代情况相比,渲染时的处理量增加。因此,出现能够由解码装置处理的对象的数量减少的情况,或者出现由包含硬件规模小的渲染器的解码装置不能执行渲染的情况。

[0107] 因此,在当渲染时执行用于延伸声像的处理的情况下,期望能够以尽可能小的处理量进行渲染。

[0108] 此外,由于存在上述18个扩展向量在围绕中心位置 $p_0$ =位置p的单位球面上在上下方向以及左右方向上对称的约束,所以不能执行考虑对象的声音的方向性(辐射方向)或者对象的形状的处理。因此,不能获得足够高质量的声音。

[0109] 此外,在MPEG-H 3D音频标准中,由于一种处理被规定为在渲染时用于延伸声像的处理,所以在渲染器的硬件规模小的情况下,不能执行用于延伸声像的处理。换言之,不能执行音频的再现。

[0110] 此外,在MPEG-H 3D音频标准中,不能执行切换执行渲染的处理,使得可以通过渲染器的硬件规模所允许的处理量来获得具有最高质量的声音。

[0111] 考虑到如上所述的这种情况,本技术使得可以在渲染时减少处理量。此外,本技术使得可以通过表示对象的方向性或形状来获得足够高质量的声音。此外,本技术使得可以响应于渲染器等硬件规模选择适当的处理作为进行渲染时的处理,以在允许的处理量的扩散内获得具有最高质量的声音。

[0112] 下面描述本技术的概要。

[0113] <处理量的减少>

[0114] 首先,描述在渲染时处理量的减少。

[0115] 在不延伸声像的常见VBAP处理(渲染处理)中,执行下面具体描述的处理A1至A3:

[0116] (处理A1)

[0117] 对于三个扬声器计算与音频信号相乘的VBAP增益。

[0118] (处理A2)

[0119] 执行归一化,使得三个扬声器的VBAP增益的平方和变成1。

[0120] (处理A3)

[0121] 对象的音频信号乘以VBAP增益。

[0122] 这里,在处理A3中,由于对于三个扬声器中的每一个执行音频信号乘以VBAP增益的乘法处理,因此刚刚描述的这种乘法处理最多执行三次。

[0123] 另一方面,在执行用于延伸声像的处理的VBAP处理(渲染处理)中,执行下面具体描述的处理B1至B5:

[0124] (处理B1)

[0125] 对于向量p计算要与三个扬声器中的每个扬声器的音频信号相乘的VBAP增益。

[0126] (处理B2)

[0127] 对于18个扩展向量计算要与三个扬声器中的每个扬声器的音频信号相乘的VBAP增益。

[0128] (处理B3)

[0129] 对于每个扬声器使关于向量所计算的VBAP增益相加。

[0130] (处理B4)

[0131] 执行归一化,使得所有扬声器的VBAP增益的平方和变成1。

[0132] (处理B5)

[0133] 对象的音频信号乘以VBAP增益。

[0134] 当执行用于延伸声像的处理时,由于输出声音的扬声器的数量为三个或更多个,所以执行处理B5中的乘法处理三次或更多次。

[0135] 因此,如果将执行用于延伸声像的处理的情况与不执行用于延伸声像的处理的另一情况进行相互比较,则当执行用于延伸声像的处理时,处理量特别是通过处理B2和B3增加了一定量,并且处理B5中的处理量也大于处理A3中的处理量。

[0136] 因此,本技术使得可以通过对关于每个扬声器确定的向量的VBAP增益的总和进行量化来减少上述处理B5中的处理量。

[0137] 具体地,通过本技术执行如下所述的这种处理。应当注意,关于每个向量例如向量p或关于每个扬声器确定的扩展向量所计算的VBAP增益的和(相加值)也被称为VBAP增益相加值。

[0138] 首先,在执行处理B1至B3并且关于每个扬声器获得VBAP增益相加值之后,然后对VBAP增益相加值进行二值化。在二值化中,例如,每个扬声器的VBAP增益相加值为0和1中之一。

[0139] 可以采用诸如四舍五入、上限(上舍入)、下限(截取)或阈值处理的任何方法作为用于对VBAP增益相加值进行二值化的方法。

[0140] 在以这种方式对VBAP增益相加值进行二值化之后,基于二值化的VBAP增益相加值执行上述处理B4。因此,每个扬声器的最终VBAP增益是除了0之外的一个增益。换言之,如果对VBAP增益相加值进行二值化,则每个扬声器的VBAP增益的最终值为0或预定值。

[0141] 例如,如果作为二值化的结果三个扬声器的VBAP增益相加值为1而另一扬声器的VBAP增益相加值为0,则三个扬声器的VBAP增益的最终值为 $1/3^{(1/2)}$ 。

[0142] 在以这种方式获得扬声器的最终VBAP增益之后,执行将扬声器的音频信号乘以最终VBAP增益的处理作为处理B5'来替代上述处理B5。

[0143] 如果以如上所述的这种方式执行二值化,则由于每个扬声器的VBAP增益的最终值变成0和预定值中之一,所以在处理B5'中,仅需要执行乘法处理一次,因此能够减少处理量。换言之,虽然处理B5要求执行乘法处理三次或更多次,但是处理B5'仅要求执行乘法处理一次。

[0144] 应当注意,尽管这里给出了对VBAP增益相加值进行二值化的情况作为示例,但是也可以以其他方式将VBAP增益相加值量化成三个值或更多个值中之一。

[0145] 例如,在VBAP增益相加值为三个值中之一的情况下,在执行上述处理B1至B3并且对于每个扬声器获得VBAP增益相加值之后,VBAP增益相加值被量化成0、0.5和1中之一。然后,执行处理B4和处理B5'。在这种情况下,处理B5'中的乘法处理的次数最多为2次。

[0146] 在以这种方式对VBAP增益相加值进行x值转换的情况下,即在将VBAP增益相加值量化成x增益中之一(其中x等于或大于2)情况下,则在处理B5'中执行乘法处理的次数最大

变成  $(x-1)$ 。

[0147] 应当注意,虽然在前面的描述中描述了在执行用于延伸声像的处理时对VBAP增益相加值进行量化以减少处理量的示例,但是同样在不执行用于延伸声像的处理的情况下,可以类似地通过对VBAP增益进行量化来减少处理量。具体地,如果对对于向量 $p$ 确定的每个扬声器的VBAP增益进行量化,则通过归一化后的VBAP增益进行的音频信号的乘法处理的执行次数可以减少。

[0148] <用于表示对象的形状和对象的声音的方向性的处理>

[0149] 现在,描述通过本技术来表示对象的形状和对象的声音的方向性的处理。

[0150] 在下文中,描述了包括扩展三维向量法、扩展中心向量法、扩展端向量法、扩展辐射向量法和任意扩展向量法的五种方法。

[0151] (扩展三维向量法)

[0152] 首先,描述了扩展三维向量法。

[0153] 在扩展三维向量法中,作为三维向量的扩展三维向量被存储在比特流中并与比特流一起被发送。这里,假设扩展三维向量例如被存储在每个对象的每个音频信号的帧的元数据中。在这种情况下,指示声像的扩散程度的扩展不存储在元数据中。

[0154] 例如,扩展三维向量是包括以下三个因子的三维向量:指示声像在水平方向上的扩散程度的 $s3\_azimuth$ 、指示声像在竖直方向上的扩散程度的 $s3\_elevation$ 、指示声像的半径方向上的深度的 $s3\_radius$ 。

[0155] 具体地,扩展三维向量 = ( $s3\_azimuth, s3\_elevation, s3\_radius$ )。

[0156] 这里, $s3\_azimuth$ 指示声像在位置 $p$ 的水平方向上即在上述水平方向角方位角的方向上的扩展角。具体地, $s3\_azimuth$ 指示由向量 $p$ (向量 $p0$ )和从原点 $o$ 朝向指示声像的扩散的区域的水平方向侧中的末端的向量限定的角度。

[0157] 类似地, $s3\_elevation$ 指示声像在位置 $p$ 的竖直方向上即在上述竖直方向角仰角的方向上的扩展角。具体地, $s3\_elevation$ 指示由向量 $p$ (向量 $p0$ )和从原点 $o$ 朝向指示声像的扩散的区域的竖直方向侧中的末端的向量限定的角度。此外, $s3\_radius$ 指示在上述距离半径的方向上即在单位球面的法线方向上的深度。

[0158] 应当注意, $s3\_azimuth$ 、 $s3\_elevation$ 和 $s3\_radius$ 具有等于或大于0的值。此外,虽然这里的扩展三维向量是指示由对象的位置信息指示的位置 $p$ 的相对位置的信息,但是扩展三维向量还可以另外是指示绝对位置的信息。

[0159] 在扩展三维向量法中,使用如上所述的这种扩展三维向量来执行渲染。

[0160] 具体地,在扩展三维向量法中,通过基于扩展三维向量计算下面给出的表达式(1)来计算扩展的值:

[0161] [表达式1]

[0162] 扩展: $\max(s3\_azimuth, s3\_elevation) \dots (1)$

[0163] 应当注意,表达式(1)中的 $\max(a, b)$ 指示返回值 $a$ 和 $b$ 中的较高值的函数。因此, $s3\_azimuth$ 和 $s3\_elevation$ 的较高值被确定为扩展的值。

[0164] 然后,基于以这种方式获得的扩展的值和包括在元数据中的位置信息,与MPEG-H 3D音频标准的情况类似地计算18个扩展向量 $p1$ 至 $p18$ 。

[0165] 因此,将由包括在元数据中的位置信息指示的对象的位置 $p$ 确定为中心位置 $p0$ ,并

且确定18个扩展向量p1至p18,使得它们在以中心位置p0为中心的球面上在左右方向上和上下方向上对称。

[0166] 此外,在扩展三维向量法中,其起点为原点0且其终点为中心位置p0的向量p0被确定为扩展向量p0。

[0167] 此外,每个扩展向量由水平方向角方位角、竖直方向角仰角和距离半径表示。在下文中,特别地扩展向量p<sub>i</sub> (其中i=0至18)的水平方向角方位角和竖直方向角仰角分别被表示为a(i)和e(i)。

[0168] 在以这种方式获得扩展向量p0至p18之后,基于s3\_azimuth与s3\_elevation之间的比率将扩展向量p1至p18变成(校正为)最终扩展向量。

[0169] 具体地,在s3\_azimuth大于s3\_elevation的情况下,执行下面的表达式(2)的计算,以将作为扩展向量p1至p18的仰角的e(i)变成e'(i):

[0170] [表达式2]

$$e'(i) = e(0) + (e(i) - e(0)) \times s3\_elevation / s3\_azimuth \quad \dots (2)$$

[0172] 应当注意,对于扩展向量p0,不执行仰角的校正。

[0173] 相反,在s3\_azimuth小于s3\_elevation的情况下,执行下面的表达式(3)的计算,以将作为扩展向量p1至p18的方位角的a(i)变成a'(i):

[0174] [表达式3]

$$a'(i) = a(0) + (a(i) - a(0)) \times s3\_azimuth / s3\_elevation \quad \dots (3)$$

[0176] 应当注意,对于扩展向量p0,不执行方位角的校正。

[0177] 将s3\_azimuth和s3\_elevation中较大的一个确定为扩展来以如上所述的这种方式确定扩展向量的处理是以下处理:将指示声像在单位球面上的扩散的区域暂时设置为由s3\_azimuth和s3\_elevation中较大的一个的角度限定的半径的圆以通过类似于常规处理的处理来确定扩展向量。

[0178] 此外,随后响应于s3\_azimuth和s3\_elevation之间的大小关系通过表达式(2)或表达式(3)来校正扩展向量的处理是以下处理:校正指示声像的扩散的区域即扩展向量,使得指示声像在单位球面上的扩散的区域变成由扩展三维向量指定的原始的s3\_azimuth和s3\_elevation限定的区域。

[0179] 因此,之后上述处理全部变成以下处理:基于扩展三维向量即基于s3\_azimuth和s3\_elevation来计算指示声像在单位球面上的扩散的区域—其具有圆形形状或椭圆形形状—的扩展向量。

[0180] 在以这种方式获得扩展向量之后,扩展向量p0至p18此后用于执行上述处理B2、处理B3、处理B4和处理B5',以生成要提供给扬声器的音频信号。

[0181] 应当注意,在处理B2中,对于扩展向量p0至p18的19个扩展向量中的每一个计算每个扬声器的VBAP增益。这里,由于扩展向量p0为向量p,所以可以认为对于扩展向量p0计算VBAP增益的处理是执行处理B1。此外,在处理B3之后,根据需要执行每个VBAP增益相加值的量化。

[0182] 通过以这种方式由扩展三维向量将指示声像的扩散的区域设置成任意形状的区域,可以表示对象的形状和对象的声音的方向性,从而可以通过渲染获得更高质量的声音。

[0183] 此外,虽然这里描述了将值s3\_azimuth和s3\_elevation中的较高值用作扩展的值

的示例,但是另外可以将值s3\_azimuth和s3\_elevation中的较低值用作扩展的值。

[0184] 在这种情况下,当s3\_azimuth大于s3\_elevation时,对每个扩展向量的方位角a(i)进行校正,但是当s3\_azimuth小于s3\_elevation时,对每个扩展向量的仰角e(i)进行校正。

[0185] 此外,虽然这里给出了其中确定扩展向量p0至p18即预先确定的19个扩展向量并且对于这些扩展向量计算VBAP增益的示例,但是要计算的扩展向量的数量是可变的。

[0186] 在刚刚描述的这种情况下,可以例如响应于s3\_azimuth和s3\_elevation之间的比率来确定要生成的扩展向量的数量。根据刚刚描述的这种处理,例如,在对象水平伸长并且对象的声音在竖直方向上的扩散小的情况下,如果忽略在竖直方向上并列的扩展向量并且扩展向量基本上在水平方向上并列,则可以适当地表示声音水平方向上的扩散。

[0187] (扩展中心向量法)

[0188] 现在,描述扩展中心向量法。

[0189] 在扩展中心向量法中,作为三维向量的扩展中心向量被存储在比特流中并与比特流一起被发送。这里,假定扩展中心向量例如被存储在每个对象的每个音频信号的帧的元数据中。在这种情况下,指示声像的扩散程度的扩展也被存储在元数据中。

[0190] 扩展中心向量是指示指示对象的声像的扩散的区域的中心位置p0的向量。例如,扩展中心向量是由以下三个因子构成的三维向量:指示中心位置p0的水平方向角的方位角、指示中心位置p0的竖直方向角的仰角、和指示中心位置p0在径向方向上的距离的半径。

[0191] 具体地,扩展中心向量=(方位角,仰角,半径)。

[0192] 当渲染处理时,由扩展中心向量指示的位置被确定为中心位置p0,并且计算扩展向量p0至p18作为扩展向量。这里,例如,如图4所描绘的,扩展向量p0是其起点为原点0且其终点为中心位置p0的向量p0。应当注意,在图4中,与图3的情况下的部分对应的部分由相同附图标记表示,并且适当地省略对它们的描述。

[0193] 此外,在图4中,用虚线绘制的箭头表示扩展向量,以及同样在图4中,为了使图容易看出,仅描绘了九个扩展向量。

[0194] 然而,在图3所描绘的例子中,位置p=中心位置P0,在图4的示例中,中心位置p0是与位置p不同的位置。在该示例中,可以看出,相对于为对象的位置的位置p,指示声像的扩散并以中心位置p0为中心的区域R21从图3的示例中的左侧移位至图4中的左侧。

[0195] 如果能够以这种方式通过扩展中心向量指定任意位置作为指示声像的扩散的区域的中心位置p0,则可以以更高的准确度表示对象的声音的方向性。

[0196] 在扩展中心向量法中,如果获得扩展向量p0至p18,则之后对于向量p执行处理B1,并且对于扩展向量p0至p18执行处理B2。

[0197] 应当注意,在处理B2中,可以对于19个扩展向量中的每一个计算VBAP增益,或者可以仅对于除了扩展向量p0之外的扩展向量p1至p18计算VBAP增益。在下文中,假定也对于扩展向量p0计算VBAP增益来给出描述。

[0198] 此外,在计算每个向量的VBAP增益之后,执行处理B3、处理B4和处理B5'以生成要提供给扬声器的音频信号。应当注意,在处理B3之后,根据需要执行VBAP增益相加值的量化。

[0199] 通过如上所述的这种扩展中心向量法,也可以通过渲染获得足够高质量的声音。

[0200] (扩展端向量法)

[0201] 现在,描述扩展端向量法。

[0202] 在扩展端向量法中,作为五维向量的扩展端向量被存储在比特流中并与比特流一起被发送。这里,假定例如扩展端向量被存储在每个对象的每个音频信号的帧的元数据中。在这种情况下,指示声像的扩散程度的扩展不存储在元数据中。

[0203] 例如,扩展端向量是表示指示对象的声像的扩散的区域的向量,并且是由以下五个因子构成的向量:扩展左端方位角、扩展右端方位角、扩展上限端面仰角、扩展下端仰角和扩展半径。

[0204] 这里,构成扩展端向量的扩展左端方位角和扩展右端方位角分别表示水平方向角方位角的值,水平方向角方位角表示指示声像的扩散的区域的水平方向上的左端和右端的绝对位置。换言之,扩展左端方位角和扩展右端方位角分别指示以下角度:指示声像相对于指示声像的扩散的区域的中心位置 $p_0$ 在向左方向和向右方向上的扩散程度。

[0205] 同时,扩展上端仰角和扩展下端仰角分别表示指示竖直方向角仰角的值,竖直方向角仰角指示声像的扩散的区域的竖直方向上的上端和下端的绝对位置。换言之,扩展上端仰角和扩展下端仰角分别指示以下角度:表示声像相对于指示声像的扩散的区域的中心位置 $p_0$ 在向上方向和向下方向上的扩散程度的角度。此外,扩展半径表示声像在径向方向上的深度。

[0206] 应当注意,虽然这里扩展端向量是指示空间中的绝对位置的信息,但是扩展端向量可以另外是指示由对象的位置信息指示的位置 $p$ 的相对位置的信息。

[0207] 在扩展端向量法中,使用如上所述的这种扩展端向量进行渲染。

[0208] 具体地,在扩展端向量法中,基于扩展端向量计算下面的表达式(4),以计算中心位置 $p_0$ :

[0209] [表达式4]

[0210] 方位角: (扩展左端方位角+扩展右端方位角)/2

[0211] 仰角: (扩展上端仰角+扩展下端仰角)/2

[0212] 半径:扩展半径

[0213] ... (4)

[0214] 具体地,指示中心位置 $p_0$ 的水平方向角方位角是扩展左端方位角和扩展右端方位角之间的中间(平均)角度,以及指示中心位置 $p_0$ 的竖直方向角仰角是扩展上端仰角和扩展下端仰角之间的中间(平均)角度。此外,指示中心位置 $p_0$ 的距离半径是扩展半径。

[0215] 因此,在扩展端向量法中,中心位置 $p_0$ 有时成为与位置信息所表示的对象的位置 $p$ 不同的位置。

[0216] 此外,在扩展端向量法中,通过计算下面的表达式(5)来计算扩展的值:

[0217] [表达式5]

[0218] 扩展:  $\max((\text{扩展左端方位角}-\text{扩展右端方位角})/2, (\text{扩展上端仰角}-\text{扩展下端仰角})/2)$

[0219] ... (5)

[0220] 应当注意,表达式(5)中的 $\max(a, b)$ 表示返回值 $a$ 和 $b$ 中的较高值的函数。因此,由扩展端向量指示的作为与指示对象的声像的扩散的区域中的水平方向上的半径相对应的

角度的(扩展左端方位角-扩展右端方位角)/2和作为与该区域中的竖直方向上的半径对应的角度的(扩展上端方位角-扩展右端方位角)/2的值中的较高值被确定为扩展的值。

[0221] 然后,基于中心位置 $p_0$ (向量 $p_0$ )和以这种方式获得的扩展的值,与MPEG-H 3D音频标准的情况类似地计算18个扩展向量 $p_1$ 至 $p_{18}$ 。

[0222] 因此,确定18个扩展向量 $p_1$ 至 $p_{18}$ ,使得它们在以中心位置 $p_0$ 为中心的球面上在上下方向和左右方向上对称。

[0223] 此外,在扩展端向量法中,将其起点为原点0且其终点为中心位置 $p_0$ 的向量 $p_0$ 确定为扩展向量 $p_0$ 。

[0224] 同样在扩展端向量法中,与扩展三维向量法的情况类似,每个扩展向量由水平方向角方位角、竖直方向角仰角和距离半径表示。换言之,扩展向量 $p_i$ (其中 $i=0$ 至18)的水平方向角方位角和竖直方向角仰角分别由 $a(i)$ 和 $e(i)$ 表示。

[0225] 在以这种方式获得扩展向量 $p_0$ 至 $p_{18}$ 之后,基于(扩展左端方位角-扩展右端方位角)和(扩展上端仰角-扩展下端仰角)之间的比率来改变(校正)扩展向量 $p_1$ 至 $p_{18}$ 以确定最终的扩展向量。

[0226] 具体地,如果(扩展左端方位角-扩展右端方位角)大于(扩展上端仰角-扩展下端仰角),则执行下面给出的表达式(6)的计算,并且将作为每个扩展向量 $p_1$ 至 $p_{18}$ 的仰角的 $e(i)$ 变成 $e'(i)$ :

[0227] [表达式6]

[0228] 
$$e'(i) = e(0) + (e(i) - e(0)) \times (\text{扩展上端仰角} - \text{扩展下端仰角}) / (\text{扩展左端方位角} - \text{扩展右端方位角}) \quad \dots (6)$$

[0229] 应当注意,对于扩展向量 $p_0$ ,不执行仰角的校正。

[0230] 另一方面,当(扩展左端方位角-扩展右端方位角)小于(扩展上端仰角-扩展下端仰角)时,执行下面给出的表达式(7)的计算,并且将作为每个扩展向量 $p_1$ 至 $p_{18}$ 的方位角的 $a(i)$ 变成 $a'(i)$ :

[0231] [表达式7]

[0232] 
$$a'(i) = a(0) + (a(i) - a(0)) \times (\text{扩展左端方位角} - \text{扩展右端方位角}) / (\text{扩展上端仰角} - \text{扩展下端仰角})$$

[0233] 
$$\dots (7)$$

[0234] 应当注意,对于扩展向量 $p_0$ ,不执行方位角的校正。

[0235] 应当注意,如上所述的扩展向量的计算方法基本上与扩展三维向量法的情况下计算方法相似。

[0236] 因此,之后上述处理全部是以下处理:根据扩展端向量计算指示圆形或椭圆形的声像在由扩展端向量限定的单位球面上的扩散的区域的扩展向量。

[0237] 在以这种方式获得扩展向量之后,使用向量 $p$ 和扩展向量 $p_0$ 至 $p_{18}$ 来执行上述的处理B1、处理B2、处理B3、处理B4和处理B5',从而生成要提供给扬声器的音频信号。

[0238] 应当注意,在处理B2中,对于19个扩展向量计算每个扬声器的VBAP增益。此外,在处理B3之后,根据需要执行VBAP增益相加值的量化。

[0239] 通过以这种方式由扩展端向量将指示声像的扩散的区域设置成具有位于任意位置的中心位置 $p_0$ 的任意形状的区域,可以表示对象的形状和对象的声音的方向性,从而通



过渲染可以获得更高质量的声音。

[0240] 此外,虽然这里描述了将值(扩展左端方位角-扩展右端方位角)/2和(扩展上端仰角-扩展下端仰角)/2中的较高值用作扩展的值的例子,但是另外可以将这些值中的较低值用作扩展的值。

[0241] 此外,虽然将对于扩展向量p0计算VBAP增益的情况描述为示例,但是可以对于扩频向量p0不计算VBAP增益。假定对于扩展向量p0也计算VBAP增益给出了下面的描述。

[0242] 可替代地,类似于扩展三维向量法的情况,可以例如响应于(扩展左端方位角-扩展右端方位角)与(扩展上端仰角-扩展下端仰角)之间的比例来计算要生成的扩展向量的数量。

[0243] (扩散辐射向量法)

[0244] 此外,描述了扩展辐射向量法。

[0245] 在扩展辐射向量法中,作为三维向量的扩展辐射向量被存储在比特流中并与比特流一起被发送。这里,假定例如将扩展辐射向量存储在每个对象的每个音频信号的帧的元数据中。在这种情况下,指示声像的扩散程度的扩展也存储在元数据中。

[0246] 扩展辐射向量是表示指示对象的声像的扩散的区域的中心位置p0到对象的位置p的相对位置的向量。例如,扩展辐射向量是由以下三个因子构成的三维向量:从位置p观看,指示中心位置p0的水平方向角的方位角、指示中心位置p0的竖直方向角的仰角、以及指示中心位置p0的径向方向上的距离的半径。

[0247] 换言之,扩散辐射向量=(方位角,仰角,半径)。

[0248] 当渲染处理时,将由通过将扩展辐射向量和向量p相加而获得的向量指示的位置确定为中心位置p0,并且计算扩展向量p0至p18作为扩展向量。这里,例如,如图5所示,扩展向量p0是其起点为原点0且其终点为中心位置p0的向量p0。应当注意,在图5中,与图5的情况中的部分对应的部分由相同的附图标记表示,并且将适当地省略对这些部分的描述。

[0249] 此外,在图5中,用虚线绘制的箭头表示扩展向量,以及同样在图5中,为了使图容易看出,仅描绘了九个扩展向量。

[0250] 然而,在图3所描绘的示例中,位置p=中心位置P0,而在图5所描绘的示例中,中心位置p0是与位置p不同的位置。在该示例中,通过向量p和由箭头标记B11指示的扩展辐射向量的向量相加而获得的向量的末端位置为中心位置p0。

[0251] 此外,可以认识到,指示声像的扩散并以中心位置p0为中心的区域R31相对于为对象的位置的位置p超过图3的示例中的左侧被移位至图5中的左侧。

[0252] 如果能够以这种方式使用扩展辐射向量和位置p将任意位置指定为指示声像的扩散的区域的中心位置p0,则可以更准确地表示对象的声音的方向性。

[0253] 在扩展辐射向量法中,如果获得扩展向量p0至p18,则对于向量p执行处理B1,而对于扩展向量p0至p18执行处理B2。

[0254] 应当注意,在处理B2中,可以对于19个扩展向量计算VBAP增益,或者可以仅对于除了扩展向量p0之外的扩展向量p1至p18计算VBAP增益。在下面的描述中,假定对于扩展向量p0也计算VBAP增益。

[0255] 此外,如果计算每个向量的VBAP增益,则执行处理B3、处理B4和处理B5'以生成要提供给扬声器的音频信号。应当注意,在处理B3之后,根据需要执行每个VBAP增益相加值的

量化。

[0256] 此外,通过如上所述的这种扩散辐射向量法,可以通过渲染获得足够高质量的声音。

[0257] (任意扩展向量法)

[0258] 随后,描述任意扩展向量法。

[0259] 在任意扩展向量法中,指示用于计算VBAP增益的扩展向量的数量的扩展向量数量信息和指示每个扩展向量的终点位置的扩展向量位置信息被存储在比特流中并与比特流一起被发送。这里,假定扩展向量数量信息和扩展向量位置信息例如被存储在每个对象的每个音频信号的帧的元数据中。在这种情况下,指示声像的扩散程度的扩展不存储在元数据中。

[0260] 在渲染处理时,基于每片扩展向量位置信息,计算其起点为原点0且其终点为由扩展向量位置信息指示的位置的向量作为扩展向量。

[0261] 此后,对于向量p执行处理B1,并且对于每个扩展向量执行处理B2。此外,在计算每个向量的VBAP增益之后,执行处理B3、处理B4和处理B5'以生成要提供给扬声器的音频信号。应当注意,在处理B3之后,根据需要执行每个VBAP增益相加值的量化。

[0262] 根据如上所述的这种任意扩展向量法,可以任意地指定声像要被延伸的扩散以及该扩散的形状,因此可以通过渲染获得足够高质量的声音。

[0263] <处理的切换>

[0264] 在本技术中,可以响应于渲染器的硬件规模等选择适当的处理作为进行渲染时的处理,并且在可允许的处理量的扩散内获得最高质量的声音。

[0265] 具体地,在本技术中,为了能够在多个处理之间进行切换,用于处理切换的索引被存储在比特流中,并且与比特流一起从编码装置被发送至解码装置。换言之,用于切换处理的索引值索引被添加至比特流语法。

[0266] 例如,响应于索引值索引的值来执行下面的处理。

[0267] 具体地,当索引值索引=0时,解码装置,更具体地,解码装置中的渲染器执行与在常规MPEG-H 3D音频标准的情况下执行的渲染类似的渲染。

[0268] 另一方面,例如,当索引值索引=1时,在根据常规MPEG-H 3D音频标准指示18个扩展向量的索引的组合中,预定组合的索引被存储在比特流中并与比特流一起被发送。在这种情况下,渲染器计算与由存储在比特流中并与比特流一起发送的每个索引指示的扩展向量有关的VBAP增益。

[0269] 此外,例如,当索引值索引=2时,指示在处理中要使用的扩展向量的数量的信息以及根据常规MPEG-H 3D音频标准指示18个扩展向量中的哪一个由要用于处理的扩展向量指示的索引被存储在比特流中并与比特流一起发送。

[0270] 此外,例如,当索引值索引=3时,根据上述任意扩展向量法执行渲染处理,以及例如,当索引值索引=4时,在渲染处理中执行上述VBAP增益相加值的二值化。此外,例如,当索引值索引=5时,根据上述扩展中心向量法执行渲染处理。

[0271] 此外,可以不指定用于在编码装置中切换处理的索引值索引,而是可以由解码装置中的渲染器选择处理。

[0272] 在刚刚描述的这种情况下,例如,基于对象的元数据中包括的重要性信息来切换

处理似乎是一个值得推荐的想法。具体地,例如,对于其由重要性信息指示的重要性高(等于或高于预定值)的对象,执行上述由索引值索引=0指示的处理。对于其由重要性信息指示的重要性低(低于预定值)的对象,可以执行上述由索引值索引=4指示的处理。

[0273] 通过以这种方式适当地切换进行渲染时的处理,可以响应于渲染器的硬件规模等在可允许的处理量的扩散内获得最高质量的声音。

[0274] <音频处理装置的配置的示例>

[0275] 随后,描述上述本技术的更具体的实施方式。

[0276] 图6是描绘了应用本技术的音频处理装置的配置的示例的图。

[0277] 分别对应于M个声道的扬声器12-1至12-M连接至图6中所描绘的音频处理装置11。音频处理装置11基于从外部提供的对象的音频信号和元数据来生成不同声道的音频信号,并将音频信号提供给扬声器12-1至12-M,使得扬声器12-1至12-M再现声音。

[0278] 应当注意,在下面的描述中,在不需要将扬声器12-1至12-M彼此特定区分的情况下,它们中的每一个仅被称为扬声器12。每个扬声器12是基于提供给其的音频信号输出声音的声音输出单元。

[0279] 扬声器12被布置成包围欣赏内容等的用户。例如,扬声器12布置在上述单位球面上。

[0280] 音频处理装置11包括获取单元21、向量计算单元22、增益计算单元23和增益调整单元24。

[0281] 获取单元21获取来自外部的对象的音频信号以及每个对象的音频信号的每一帧的元数据。例如,通过由解码装置对从编码装置输出的比特流中包括的编码音频数据和编码元数据进行解码来获得音频数据和元数据。

[0282] 获取单元21将所获取的音频信号提供给增益调整单元24,并将所获取的元数据提供给向量计算单元22。这里,元数据例如根据需要包括指示对象的位置的位置信息、指示每个对象的重要性程度的重要性信息、指示对象的声像的空间扩散的扩展等。

[0283] 向量计算单元22基于从获取单元21提供至此的元数据来计算扩展向量,并将扩展向量提供给增益计算单元23。此外,根据需要,向量计算单元22将由元数据中所包括的位置信息指示的每个对象的位置即指示位置p的向量p提供给增益计算单元23。

[0284] 增益计算单元23基于从向量计算单元22提供的扩展向量和向量p通过VBAP来计算与每个声道对应的扬声器12的VBAP增益,并将VBAP增益提供给增益调整单元24。此外,增益计算单元23包括用于对每个扬声器的VBAP增益进行量化的量化单元31。

[0285] 增益调整单元24基于从增益计算单元23提供的每个VBAP增益对从获取单元21提供的对象的音频信号进行增益调整,并且将作为增益调整的结果而获得的M个声道的音频信号提供给扬声器12。

[0286] 增益调整单元24包括放大单元32-1至32-M。放大单元32-1至32-M将从获取单元21提供的音频信号乘以从增益计算单元23提供的VBAP增益,并将通过乘法获得的音频信号提供给扬声器12-1至12-M以便再现声音。

[0287] 应当注意,在下面的描述中,在不需要将放大单元32-1至32-M彼此特定区分的情况下,它们中的每一个也仅被称为放大单元32。

[0288] <再现处理的描述>

[0289] 现在,描述图6中描绘的音频处理装置11的操作。

[0290] 如果从外部提供音频信号和对象的元数据,则音频处理装置11执行再现处理以再现对象的声音。

[0291] 在下面,参照图7的流程图描述音频处理装置11的再现处理。应当注意的是,针对音频信号的每一帧执行该再现处理。

[0292] 在步骤S11处,获取单元21从外部获取对象的一帧的音频信号和元数据,并且将该音频信号提供至放大单元32,同时它将元数据提供至向量计算单元22。

[0293] 在步骤S12处,向量计算单元22基于从获取单元21提供的元数据执行扩展向量计算处理,并且将作为扩展向量计算处理的结果而获得的扩展向量提供至增益计算单元23。此外,根据需要,向量计算单元22还向增益计算单元23提供向量p。

[0294] 应当注意的是,尽管在下文中描述了扩展向量计算处理的细节,但是在扩展向量计算处理中,扩展向量通过扩展三维向量方法、扩展中心向量方法、扩展端向量方法、扩展辐射向量方法或任意扩展向量法来计算。

[0295] 在步骤S13处,增益计算单元23基于表示预先保存的扬声器12的位置的位置信息以及从向量计算单元22提供的向量p和扩展向量来计算各个扬声器12的VBAP增益。

[0296] 具体地,关于扩展向量和向量p中的每一个,计算每个扬声器12的VBAP增益。因此,对于扩展向量和向量p中的每一个,获得位于对象的位置附近(即,位于由向量指示的位置附近)的一个或多个扬声器12的VBAP增益。应当注意的是,虽然必定计算扩展向量的VBAP增益,但是如果通过步骤S12处的处理并未将向量p从向量计算单元22提供至增益计算单元23,则不计算向量p的VBAP增益。

[0297] 在步骤S14处,增益计算单元23将关于每个向量计算的VBAP增益相加,以计算每个扬声器12的VBAP增益相加值。具体地,计算针对同一扬声器12所计算的向量的VBAP增益的相加值(总和)作为VBAP增益相加值。

[0298] 在步骤S15处,量化单元31决定是否要执行VBAP增益相加值的二值化。

[0299] 例如,可以基于在上文中描述的索引值索引来决定或者可以基于作为元数据的由重要性信息指示的对象的重要性程度来决定是否要执行二值化。

[0300] 如果基于索引值索引来执行决定,则例如可以将从比特流读出的索引值索引提供至增益计算单元23。可替代地,如果基于重要性信息来执行决定,则可以将重要性信息从向量计算单元22提供至增益计算单元23。

[0301] 如果在步骤S15处决定要执行二值化,则在步骤S16处,量化单元31对针对每个扬声器12决定的VBAP增益的相加值(即,VBAP增益相加值)进行二值化。此后,处理前进至步骤S17。

[0302] 相反,如果在步骤S15处决定不执行二值化,则跳过步骤S16处的理,并且处理前进至步骤S17。

[0303] 在步骤S17处,增益计算单元23对每个扬声器12的VBAP增益进行归一化,使得所有扬声器12的VBAP增益的平方和可以变为1。

[0304] 具体地,对针对每个扬声器12所决定的VBAP增益的相加值执行归一化,使得所有相加值的平方和可以变为1。增益计算单元23将通过归一化获得的扬声器12的VBAP增益提供至与各个扬声器12对应的放大单元32。

[0305] 在步骤S18处,放大单元32将从获取单元21提供的音频信号与从增益计算单元23提供的VBAP增益相乘,并且将所得到的值提供至扬声器12。

[0306] 然后在步骤S19处,放大单元32使扬声器12基于提供至其的音频信号来再现声音,从而结束再现处理。因此,对象的声像被定位在再现空间中的期望的部分空间中。

[0307] 以如上所述这样的方式,音频处理装置11基于元数据来计算扩展向量,针对每个扬声器12计算每个向量的VBAP增益,并且针对每个扬声器12决定和归一化VBAP增益的相加值。通过以这种方式计算关于扩展向量的VBAP增益,可以表示对象的声像的空间扩散,特别是对对象的形状或声音的方向性,并且可以获得更高质量的声音。

[0308] 此外,通过根据需要对VBAP增益的相加值进行二值化,不仅可以减少渲染时的处理量,而且可以响应于音频处理装置11的处理能力(硬件规模)执行适当的处理以获得尽可能高质量的声音。

[0309] <扩展向量计算处理的描述>

[0310] 在此,参照图8的流程图来描述与图7的步骤S12处的处理对应的扩展向量计算处理。

[0311] 在步骤S41处,向量计算单元22基于扩展三维向量来决定是否要计算扩展向量。

[0312] 例如,与图7的步骤S15处的情况类似,可以基于索引值索引来决定使用哪种方法计算扩展向量,或者可以基于由重要性信息指示的对象的重要性程度来决定使用哪种方法计算扩展向量。

[0313] 如果在步骤S41处决定要基于扩展三维向量计算扩展向量,即,如果决定要通过扩展三维方法计算扩展向量,则处理前进至步骤S42。

[0314] 在步骤S42处,向量计算单元22基于扩展三维向量来执行扩展向量计算处理,并且将所得到的向量提供至增益计算单元23。应当注意的是,在下文中描述基于扩展三维向量的扩展向量计算处理的细节。

[0315] 在扩展向量被计算之后,扩展向量计算处理结束,此后,处理前进至图7的步骤S13。

[0316] 另一方面,如果在步骤S41处决定不要基于扩展三维向量来计算扩展向量,则处理前进至步骤S43。

[0317] 在步骤S43处,向量计算单元22决定是否要基于扩展中心向量来计算扩展向量。

[0318] 如果在步骤S43处决定要基于扩展中心向量来计算扩展向量,即,如果决定要通过扩展中心向量方法来计算扩展向量,则处理前进至步骤S44。

[0319] 在步骤S44处,向量计算单元22基于扩展中心向量执行扩展向量计算处理,并且将所得到的向量提供至增益计算单元23。应当注意的是,在下文中描述基于扩展中心向量的扩展向量计算处理的细节。

[0320] 在扩展向量被计算之后,扩展向量计算处理结束,此后,处理前进至图7的步骤S13。

[0321] 另一方面,如果在步骤S43处决定不要基于扩展中心向量来计算扩展向量,则处理前进至步骤S45。

[0322] 在步骤S45处,向量计算单元22决定是否要基于扩展端向量来计算扩展向量。

[0323] 如果在步骤S45处决定要基于扩展端向量来计算扩展向量,即,如果决定要通过扩

展端向量方法来计算扩展向量,则处理前进至步骤S46。

[0324] 在步骤S46处,向量计算单元22基于扩展端向量执行扩展向量计算处理,并且将所得到的向量提供至增益计算单元23。应当注意的是,在下文中描述基于扩展端向量的扩展向量计算处理的细节。

[0325] 在扩展向量被计算之后,扩展向量计算处理结束,此后,处理前进至图7的步骤S13。

[0326] 此外,如果在步骤S45处决定不要基于扩展端向量来计算扩展向量,则处理前进至步骤S47。

[0327] 在步骤S47处,向量计算单元22决定是否要基于扩展辐射向量来计算扩展向量。

[0328] 如果在步骤S47处决定要基于扩展辐射向量来计算扩展向量,即,如果决定要通过扩展辐射向量方法来计算扩展向量,则处理前进至步骤S48。

[0329] 在步骤S48处,向量计算单元22基于扩展辐射向量来执行扩展向量计算处理,并且将所得到的向量提供至增益计算单元23。应当注意的是,在下文中描述基于扩展辐射向量的扩展向量计算处理的细节。

[0330] 在扩展向量被计算之后,扩展向量计算处理结束,此后,处理前进至图7的步骤S13。

[0331] 另一方面,如果在步骤S47处决定不要基于扩展辐射向量来计算扩展向量,即,如果决定要通过扩展辐射向量方法来计算扩展向量,则处理前进至步骤S49。

[0332] 在步骤S49处,向量计算单元22基于扩展向量位置信息来执行扩展向量计算处理,并且将所得到的向量提供至增益计算单元23。应当注意的是,在下文中描述基于扩展向量位置信息的扩展向量计算处理的细节。

[0333] 在扩展向量被计算之后,扩展向量计算处理结束,此后,处理前进至图7的步骤S13。

[0334] 音频处理装置11以这种方式通过多种方法中的适当方法来计算扩展向量。通过以这种方式经过适当的方法计算扩展向量,响应于渲染器的硬件规模等,可以获得在可允许的处理量的范围内的最高质量的声音。

[0335] <基于扩展三维向量的扩展向量计算处理的说明>

[0336] 现在,参照图8来描述与在上文描述的步骤S42、S44、S46、S48和S49处的处理对应的处理的细节。

[0337] 首先,参照图9的流程图来描述与图8的步骤S42对应的基于扩展三维向量的扩展向量计算处理。

[0338] 在步骤S81处,向量计算单元22将由包括在从获取单元21提供的元数据中的位置信息指示的位置确定为对象位置 $p$ 。换言之,表示位置 $p$ 的向量是向量 $p$ 。

[0339] 在步骤S82处,向量计算单元22基于包括在从获取单元21提供的元数据中的扩展三维向量来计算扩展。具体地,向量计算单元22计算在上文中给出的表达式(1)以计算扩展。

[0340] 在步骤S83处,向量计算单元22基于向量 $p$ 和扩展来计算扩展向量 $p_0$ 至 $p_{18}$ 。

[0341] 在此,向量 $p$ 被确定为表示中心位置 $p_0$ 的向量 $p_0$ ,并且向量 $p$ 被原样确定为扩展向量 $p_0$ 。此外,与MPEG-H 3D音频标准的情况类似,作为扩展向量 $p_1$ 至 $p_{18}$ ,向量被计算成在以

下区域内沿上下方向和左右方向对称,所述区域以中心位置p0为中心并且通过由单位球体上的扩展指示的角度来限定。

[0342] 在步骤S84处,向量计算单元22基于扩展三维向量是否满足 $s3\_azimuth \geq s3\_elevation$ ,即, $s3\_azimuth$ 是否大于 $s3\_elevation$ 做出决定。

[0343] 如果在步骤S84处决定满足 $s3\_azimuth \geq s3\_elevation$ ,则在步骤S85处,向量计算单元22改变扩展向量p1至p18的仰角。具体地,向量计算单元22执行在上文中描述的表达式(2)的计算,以校正扩展向量的仰角以获得最终扩展向量。

[0344] 在获得最终扩展向量之后,向量计算单元22将扩展向量p0至p18提供至增益计算单元23,从而结束基于扩展三维向量的扩展向量计算处理。由于图8的步骤S42处的处理随其结束,因此其后的处理前进至图7的步骤S13。

[0345] 另一方面,如果在步骤S84处决定不满足 $s3\_azimuth \geq s3\_elevation$ ,则在步骤S86处,向量计算单元22改变扩展向量p1至p18的方位角。具体地,向量计算单元22执行在上文中给定的表达式(3)的计算,以校正扩展向量的方位角,从而获得最终扩展向量。

[0346] 在获得最终扩展向量之后,向量计算单元22将扩展向量p0至p18提供至增益计算单元23,从而结束基于扩展三维向量的扩展向量计算处理。因此,由于图8的步骤S42处的处理结束,因此其后的处理前进至图7的步骤S13。

[0347] 音频处理装置11以如上所述这样的方式通过扩展三维向量方法来计算每个扩展向量。因此,变得可以表示对象的形状以及对象的声音的方向性,并且获得较高质量的声音。

[0348] <基于扩展中心向量的扩展向量计算处理的说明>

[0349] 现在,参照图10的流程图来描述与图8的步骤S44对应的基于扩展中心向量的扩展向量计算处理。

[0350] 应当注意的是,步骤S111处的处理与图9的步骤S81处的处理类似,因此,省略对其的描述。

[0351] 在步骤S112处,向量计算单元22基于扩展中心向量以及包括在从获取单元21提供的元数据中的扩展来计算扩展向量p0至p18。

[0352] 具体地,向量计算单元22将由扩展中心向量指示的位置设置为中心位置p0,并且将表示中心位置p0的向量设置为扩展向量p0。此外,向量计算单元22将扩展向量p1至p18确定成使得它们在以下区域中沿上下方向和左右方向对称的定位,所述区域以中心位置p0为中心并且通过由单位球体上的扩展指示的角度来限定。与MPEG-H 3D音频标准的情况基本上类似地确定扩展向量p1至p18。

[0353] 向量计算单元22将通过上述处理获得的扩展向量p0至p18和向量p提供至增益计算单元23,从而结束基于扩展中心向量的扩展向量计算处理。因此,图8的步骤S44处的处理结束,此后,处理前进至图7的步骤S13。

[0354] 音频处理装置11以如上所述这样的方式通过扩展中心向量方法来计算向量p和扩展向量。因此,变得可以表示对象的形状以及对象的声音的方向性,并且获得较高质量的声音。

[0355] 应当注意的是,在基于扩展中心向量的扩展向量计算处理中,扩展向量p0可以不被提供至增益计算单元23。换言之,可以不关于扩展向量p0来计算VBAP增益。

[0356] <基于扩展端向量的扩展向量计算处理的说明>

[0357] 此外,参照图11的流程图来描述与图8的步骤S46对应的基于扩展端向量的扩展向量计算处理。

[0358] 应当注意的是,步骤S141处的处理与图9的步骤S81处的处理类似,因此,省略对其的描述。

[0359] 在步骤S142处,向量计算单元22基于包括在从获取单元21提供的元数据中的扩展端向量来计算中心位置 $p_0$ ,即,向量 $p_0$ 。具体地,向量计算单元22计算在上文中给出的表达式(4)以计算中心位置 $p_0$ 。

[0360] 在步骤S143处,向量计算单元22基于扩展端向量来计算扩展。具体地,向量计算单元22计算在上文中给出的表达式(5)以计算扩展。

[0361] 在步骤S144处,向量计算单元22基于中心位置 $p_0$ 和扩展来计算扩展向量 $p_0$ 至 $p_{18}$ 。

[0362] 在此,表示中心位置 $p_0$ 的向量 $p_0$ 被原样设置为扩展向量 $p_0$ 。此外,与MPEG-H 3D音频标准的情况类似,扩展向量 $p_1$ 至 $p_{18}$ 被计算成使得它们在以下区域内沿上下方向和左右方向对称的定位,所述区域以中心位置 $p_0$ 为中心并且通过由单位球体上的扩展指示的角度来限定。

[0363] 在步骤S145处,向量计算单元22决定是否满足(扩展左端方位角-扩展右端方位角) $\geq$ (扩展上端仰角-扩展下端仰角),即,(扩展左端方位角-扩展右端方位角)是否大于(扩展上端仰角-扩展下端仰角)。

[0364] 如果在步骤S145处决定满足(扩展左端方位角-扩展右端方位角) $\geq$ (扩展上端仰角-扩展下端仰角),则在步骤S146处,向量计算单元22改变扩展向量 $p_1$ 至 $p_{18}$ 的仰角。具体地,向量计算单元22执行在上文中给出的表达式(6)的计算,以校正扩展向量的仰角以获得最终扩展向量。

[0365] 在获得最终扩展向量之后,向量计算单元22将扩展向量 $p_0$ 至 $p_{18}$ 和向量 $p$ 提供至增益计算单元23,从而结束基于扩展端向量的扩展向量计算处理。因此,图8的步骤S46处的处理结束,此后,处理前进至图7的步骤S13。

[0366] 另一方面,如果在步骤S145处决定不满足(扩展左端方位角-扩展右端方位角) $\geq$ (扩展上端仰角-扩展下端仰角),则在步骤S147处,向量计算单元22改变扩展向量 $p_1$ 至 $p_{18}$ 的方位角。具体地,向量计算单元22执行在上文中给出的表达式(7)的计算,以校正扩展向量的方位角以获得最终扩展向量。

[0367] 在获得最终扩展向量之后,向量计算单元22将扩展向量 $p_0$ 至 $p_{18}$ 和向量 $p$ 提供至增益计算单元23,从而结束基于扩展端向量的扩展向量计算处理。因此,图8的步骤S46处的处理结束,此后,处理前进至图7的步骤S13。

[0368] 音频处理装置11如上所述通过扩展端向量方法来计算扩展向量。因此,变得可以表示对象的形状以及对象的声音的方向性,并且获得较高质量的声音。

[0369] 应当注意的是,在基于扩展端向量的扩展向量计算处理中,扩展向量 $p_0$ 可以不被提供至增益计算单元23。换言之,可以不关于扩展向量 $p_0$ 来计算VBAP增益。

[0370] <基于扩展辐射向量的扩展向量计算处理的说明>

[0371] 现在,参照图12的流程图来描述与图8的步骤S48对应的基于扩展辐射向量的扩展向量计算处理。



[0372] 应当注意的是,步骤S171处的处理与图9的步骤S81处的处理类似,因此,省略对该处理的描述。

[0373] 在步骤S172处,向量计算单元22基于包括在从获取单元21提供的元数据中的扩展和扩展辐射向量来计算扩展向量 $p_0$ 至 $p_{18}$ 。

[0374] 具体地,向量计算单元22将由通过将表示对象位置 $p$ 的向量 $p$ 与辐射向量相加而获得的向量指示的位置设置为中心位置 $p_0$ 。指示该中心位置 $p_0$ 的向量是向量 $p_0$ ,并且向量计算单元22将向量 $p_0$ 原样设置为扩展向量 $p_0$ 。

[0375] 此外,向量计算单元22将扩展向量 $p_1$ 至 $p_{18}$ 确定成使得它们在以下区域中沿上下方向和左右方向对称的定位,所述区域以中心位置 $p_0$ 为中心并且通过由单位球体上的扩展指示的角度来限定。与MPEG-H 3D音频标准的情况基本上类似地确定扩展向量 $p_1$ 至 $p_{18}$ 。

[0376] 向量计算单元22将通过上述处理获得的扩展向量 $p_0$ 至 $p_{18}$ 和向量 $p$ 提供至增益计算单元23,从而结束基于扩展辐射向量的扩展向量计算处理。因此,由于图8的步骤S48处的处理结束,因此其后的处理前进至图7的步骤S13。

[0377] 音频处理装置11以如上所述这样的方式通过扩展辐射向量方法来计算向量 $p$ 和扩展向量。因此,变得可以表示对象的形状以及对象的声音的方向性,并且获得较高质量的声音。

[0378] 应当注意的是,在基于扩展辐射向量的扩展向量计算处理中,扩展向量 $p_0$ 可以不被提供至增益计算单元23。换言之,可以不关于扩展向量 $p_0$ 来计算VBAP增益。

[0379] <基于扩展向量位置信息的扩展向量计算处理的说明>

[0380] 现在,参照图13的流程图来描述与图8的步骤S49对应的基于扩展向量位置信息的扩展向量计算处理。

[0381] 应当注意的是,步骤S201处的处理与图9的步骤S81处的处理类似,因此,省略对其的描述。

[0382] 在步骤S202处,向量计算单元22基于包括在从获取单元21提供的元数据中的扩展向量数量信息和扩展向量位置信息来计算扩展向量。

[0383] 具体地,向量计算单元22计算以下向量,该向量在原点 $O$ 处具有起点并且在由作为扩展向量的扩展向量位置信息指示的位置处具有终点。在此,计算等于由扩展向量数量信息指示的数量的扩展向量的数量。

[0384] 向量计算单元22将通过上述处理获得的扩展向量和向量 $p$ 提供至增益计算单元23,从而结束基于扩展向量位置信息的扩展向量计算处理。因此,由于图8的步骤S49处的处理结束,因此其后的处理前进至图7的步骤S13。

[0385] 音频处理装置11以如上所述这样的方式通过任意扩展向量方法来计算向量 $p$ 和扩展向量。因此,变得可以表示对象的形状以及对象的声音的方向性,并且获得较高质量的声音。

[0386] <第二实施方式>

[0387] <渲染处理的处理量减少>

[0388] 顺便提及,如上所述,VBAP被称为用于使用多个扬声器来控制声像的定位的技术,即,用于执行渲染处理的技术。

[0389] 在VBAP中,通过从三个扬声器输出声音,可以将声像定位在根据三个扬声器构成

的三角形的内侧上的任意点处。在下文中,特别是根据这样的三个扬声器构成的三角形被称为网格。

[0390] 由于通过VBAP对每个对象执行渲染处理,因此在对象的数量大(如例如,在游戏中)的情况下,渲染处理的处理量大。因此,小硬件规模的渲染器可能不能对所有对象执行渲染,因此,可以再现仅有限数量的对象的声音。这可能在声音再现时损害呈现或声音质量。

[0391] 因此,本技术使得可以减少渲染处理的处理量,同时抑制呈现或声音质量的劣化。

[0392] 在下文中,描述了如刚刚描述的这样的技术。

[0393] 在普通VBAP处理中,即,在渲染处理中,对每个对象执行在上文中描述的处理A1至A3的处理,以产生扬声器的音频信号。

[0394] 由于针对其大体上计算VBAP增益的扬声器的数量是三个并且针对构成音频信号的每个采样来计算每个扬声器的VBAP增益,因此在处理A3中的乘法处理中,将乘法执行等于(音频信号的采样数量 $\times$ 3)的次数。

[0395] 相反,在本技术中,通过对VBAP增益执行相等的增益处理,即,VBAP增益的量化处理以及用于在适当的组合中改变要在VBAP增益计算时使用的网格数量的网格数量切换处理,渲染处理的处理量减少。

[0396] (量化处理)

[0397] 首先,描述量化处理。在此,作为量化处理的示例,描述了二值化处理和三值化处理。

[0398] 在执行二值化处理作为量化处理的情况下,在执行处理A1之后,对通过处理A1获得的每个扬声器的VBAP增益进行二值化。在二值化中,例如,每个扬声器的VBAP增益由0和1之一表示。

[0399] 应当注意的是,用于对VBAP增益进行二值化的方法可以是任何方法如四舍五入、上限(上舍入)、下限(截取)或阈值处理。

[0400] 在以这种方式对VBAP增益进行二值化之后,执行处理A2和处理A3以产生扬声器的音频信号。

[0401] 此时,在处理A2中,由于基于二值化的VBAP增益来执行归一化,因此与在上文中描述的扩展向量的量化时一样,扬声器的最终VBAP增益类似地变为除0以外的一个值。换言之,如果VBAP增益被二值化,则扬声器的最终VBAP增益的值为0或预定值。

[0402] 因此,在处理A3中的乘法处理中,可以将乘法执行(音频信号的采样数量 $\times$ 1)次,因此可以显著减少渲染处理的处理量。

[0403] 类似地,在处理A1之后,可以对获得的扬声器的VBAP增益进行三值化。在如刚刚描述的这样的情况下,通过处理A1获得的每个扬声器的VBAP增益被三值化为值0、0.5和1中的一个值。然后,其后执行处理A2和处理A3以产生扬声器的音频信号。

[0404] 因此,由于处理A3中的乘法处理中的乘法时间数量变为最大值(音频信号的采样数量 $\times$ 2),因此可以显著减少渲染处理的处理量。

[0405] 应当注意的是,尽管在此给出的描述将VBAP增益被二值化或三值化的情况作为示例,但是VBAP增益可被量化为4个或更多个值。对此进行概括,例如,VBAP增益被量化成使得其具有等于或大于2的x增益之一,或者换言之,如果对VBAP增益进行x量化次数的量化,则

处理A3中的乘法处理的次数变为最大值(x-1)。

[0406] 可以以如上所述这样的方式通过对VBAP增益进行量化来减少渲染处理的处理量。如果渲染处理的处理量以这种方式减少,则即使在对象的数量大的情况下,仍然变得可以对所有对象执行渲染,从而在声音再现时可以将呈现或声音质量的劣化抑制到低水平。换言之,可以减少渲染处理的处理量,同时抑制呈现或声音质量的劣化。

[0407] (网格数量切换处理)

[0408] 现在,描述网格数量切换处理。

[0409] 在VBAP中,如在上文中描述的,例如,参照图1,通过沿三个扬声器SP1至SP3的方向指向的向量 $I_1$ 至 $I_3$ 的线性和来表示指示处理目标的对象的声像的位置 $p$ 的向量 $p$ ,并且与前述向量相乘的系数 $g_1$ 至 $g_3$ 是扬声器的VBAP增益。在图1的示例中,由扬声器SP1至SP3包围的三角形区域TR11形成一个网格。

[0410] 在计算VBAP增益时,具体地通过以下表达式(8)通过从三角形形状的网格的逆矩阵 $L_{123}^{-1}$ 以及对象的声像的位置 $p$ 的计算来确定三个系数 $g_1$ 至 $g_3$ :

[0411] [表达式8]

$$[0412] \quad [g_1 g_2 g_3] = p L_{123}^{-1} = [p_1 p_2 p_3] \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix}^{-1} \dots (8)$$

[0413] 应当注意的是,表达式(8)中的 $p_1$ 、 $p_2$ 和 $p_3$ 指示笛卡尔坐标系上(即,图2中描绘的三维坐标系上)的表示对象的声像的位置的x坐标、y坐标和z坐标。

[0414] 此外, $I_{11}$ 、 $I_{12}$ 和 $I_{13}$ 是x分量、y分量和z分量在指向构成网格的第一扬声器SP1的向量 $I_1$ 被分解为x轴上的分量、y轴上的分量和z轴上的分量的情况下的值,并且分别对应于第一扬声器SP1的x坐标、y坐标和z坐标。

[0415] 类似地, $I_{21}$ 、 $I_{22}$ 和 $I_{23}$ 分别是x分量、y分量和z分量在指向构成网格的第二扬声器SP2的向量 $I_2$ 被分解为x轴上的分量、y轴上的分量和z轴上的分量的情况下的值。此外, $I_{31}$ 、 $I_{32}$ 和 $I_{33}$ 分别是x分量、y分量和z分量在指向构成网格的第三扬声器SP3的向量 $I_3$ 被分解为x轴上的分量、y轴上的分量和z轴上的分量的情况下的值。

[0416] 此外,如由下面的表达式(9)表示的,定义了从位置 $p$ 的三维坐标系的 $p_1$ 、 $p_2$ 和 $p_3$ 到球坐标系的坐标 $\theta$ 、 $\gamma$ 和 $r$ 的转换,其中 $r=1$ ,如下面的表达式(9)所示。在此, $\theta$ 、 $\gamma$ 和 $r$ 分别是在上文中描述的水平方向角方位角、竖直方向角仰角和距离半径。

[0417] [表达式9]

$$[0418] \quad [p_1 \ p_2 \ p_3] = [\cos(\theta) \times \cos(\gamma) \sin(\theta) \times \cos(\gamma) \sin(\gamma)] \dots (9)$$

[0419] 如上所述,在内容再现侧的空间中,即,在再现空间中,多个扬声器设置在单位球体上,并且一个网格由多个扬声器中的三个扬声器构成。此外,单位球体的整个表面基本上被多个网格覆盖,而多个网格之间没有留下间隙。此外,网格被确定成使得它们彼此不交叠。

[0420] 在VBAP中,如果从设置在单位球体的表面上的扬声器中的对构成包括对象的位置 $p$ 的一个网格的两个或三个扬声器输出声音,则可以将声像定位在位置 $p$ ,因此,除了构成网格的扬声器以外的扬声器的VBAP增益为0。

[0421] 因此,在计算VBAP增益时,可以指定包括对象的位置p的一个网格,以计算构成网格的扬声器的VBAP增益。例如,可以从所计算的VBAP增益来决定预定网格是否是包括位置p的网格。

[0422] 具体地,如果关于网格计算的三个扬声器的VBAP增益都是等于或高于0的值,则网格是包括对象的位置p的网格。相反,如果三个扬声器的VBAP增益中至少之一具有负值,则由于对象的位置p位于扬声器构成的网格之外,因此所计算的VBAP增益不是正确的VBAP增益。

[0423] 因此,在计算VBAP增益时,一个接一个地选择网格作为处理目标的网格,并且对处理目标的网格执行在上文中给出的表达式(8)的计算,以计算构成网格的每个扬声器的VBAP增益。

[0424] 然后,根据VBAP增益的计算结果,决定处理目标的网格是否是包含对象的位置p的网格,如果决定处理目标的网格是不包括位置p的网格,则将下一个网格确定为新的处理目标的网格,并且对该网格执行类似的处理。

[0425] 另一方面,如果决定处理目标的网格是包括对象的位置p的网格,则将构成网格的扬声器的VBAP增益确定为所计算的VBAP增益,而将其他扬声器的VBAP增益设置为0。因此,所有扬声器的VBAP增益被获得。

[0426] 以这种方式,在渲染处理中,同时执行用于计算VBAP增益的处理以及用于指定包括位置p的网格的处理。

[0427] 具体地,为了获得正确的VBAP增益,重复以下处理:连续地选择处理目标的网格,直到构成网格的扬声器的所有VBAP增益指示等于或高于0的值为止;以及计算网格的VBAP增益。

[0428] 因此,在渲染处理中,与单位球体的表面上的网格的数量一样,即为了获得正确的VBAP增益,指定包括位置p的网格所需的处理量增加。

[0429] 因此,在本技术中,并不是实际的再现环境中的所有扬声器用于形成(构成)网格,而是来自所有扬声器中的仅一些扬声器用于形成网格,以减少网格的总数并且减少在渲染处理时的处理量。具体地,在本技术中,执行用于改变网格总数的网格数量切换处理。

[0430] 具体地,例如,在22个声道的扬声器系统中,如图14中描绘的,将包括扬声器SPK1至SPK22的总共22个扬声器设置为单位球体的表面上的不同声道的扬声器。应当注意的是,在图14中,原点0与图2中描绘的原点0对应。

[0431] 在以这种方式将22个扬声器设置在单位球体的表面上的情况下,如果使用所有22个扬声器来形成网格,使得它们覆盖单位球体表面,则单位球体上的网格总数为40。

[0432] 相反,如图15中描绘的,例如,假定在总共22个扬声器SPK1至SPK22中,仅扬声器SPK1、SPK6、SPK7、SPK10、SPK19和SPK20总共这六个扬声器用于形成网格。应当注意的是,在图15中,用相同的附图标记表示与图14的情况中的那些对应的部分,因此适当地省略对它们的描述。

[0433] 在图15的示例中,由于22个扬声器中的仅总共6个扬声器用于形成网格,因此单位球体上的网格总数为8,网格总数可以显著减少。因此,在图15中描绘的示例中,与如图14中描绘的所有22个扬声器用于形成网格的情况相比,当计算VBAP增益时的处理量可以减少到8/40次,并且可以显著减少处理量。

[0434] 此外,在本示例中,应当注意的是,由于单位球体的整个表面被八个网格覆盖而没有间隙,因此可以在单位球体的表面上的任意位置处定位声像。然而,由于每个网格的面积随着设置在单位球体表面上的网格总数的增加而减少,因此随着网格总数增加,可以以更高的精度来控制声像的定位。

[0435] 如果通过网格数量切换处理来改变网格总数,则当选择要用于形成改变之后的数量的网格的扬声器时,期望选择如由位于原点0处的用户观察时其在竖直方向(上下方向)上的位置(即,其在竖直方向角仰角的方向上的位置)彼此不同的扬声器。换言之,期望使用三个或更多个扬声器,包括彼此位于不同高度的扬声器,来形成改变之后的数量的网格。原因是旨在抑制三维感觉(即,声音的存在)的劣化。

[0436] 例如,考虑如图16中描绘的包括设置在单位球体表面上的扬声器SP1至SP5的五个扬声器中的一些或全部用于形成网格的情况。应当注意的是,在图16中,用相同的附图标记表示与图3的情况中的那些对应的部分,因此省略对它们的描述。

[0437] 在图16中描绘的示例中所有五个扬声器SP1至SP5都用于形成覆盖在单位球体表面上的网格的情况下,网格的数量为3。具体地,包括由扬声器SP1至SP3包围的三角形形状的区域、由扬声器SP2至SP4包围的三角形形状的另一区域以及由扬声器SP2、SP4和SP5包围的三角形形状的再一区域的三个区域形成网格。

[0438] 相反,例如,如果仅使用扬声器SP1、SP2和SP5,则网格不形成三角形形状,而形成二维弧。在这种情况下,可以仅在与单位球体的扬声器SP1和SP2互连的弧上或者在与单位球体的扬声器SP2和SP5互连的弧上定位对象的声像。

[0439] 以这种方式,如果用于形成网格的所有扬声器是竖直方向上相同高度处的扬声器(即,同一层的扬声器),则由于对象的所有声像的定位位置的高度变为同一高度,因此存在劣化。

[0440] 因此,期望使用三个或更多个扬声器,包括其在竖直方向(竖直方向)上的位置彼此不同的扬声器,来形成一个或多个网格,使得可以抑制存在的劣化。

[0441] 在图16的示例中,例如,如果使用来自扬声器SP1至SP5中的扬声器SP1以及扬声器SP3至SP5,则可以形成两个网格,使得它们覆盖整个单位球体表面。在该示例中,扬声器SP1和SP5以及扬声器SP3和SP4位于彼此不同的高度处。

[0442] 在这种情况下,例如,将由扬声器SP1、SP3和SP5包围的三角形形状的区域以及由扬声器SP3至SP5包围的三角形形状的另一区域形成成为网格。

[0443] 此外,在该示例中,也可以形成包括由扬声器SP1、SP3和SP4包围的三角形形状的区域以及由扬声器SP1、SP4和SP5包围的三角形形状的另一区域的两个区域作为网格。

[0444] 在上述两个示例中,由于可以将声像定位在单位球体表面上的任意位置处,因此可以抑制存在的劣化。此外,为了形成网格,使得整个单位球体表面被多个网格覆盖,必定期望使用位于用户正上方的所谓的顶部扬声器。例如,顶部扬声器是图14中描绘的扬声器SPK19。

[0445] 与量化处理的情况类似,通过以如上所述这样的方式执行网格数量切换处理以改变网格的总数,可以减少渲染处理的处理量,此外,可以在声音再现时将呈现或声音质量的劣化抑制到低水平。换言之,可以减少渲染处理的处理量,同时抑制呈现或声音质量的劣化。

[0446] 选择是否要执行这样的网格数量切换处理或者在该网格数量切换处理中将网格的总数设置为哪个数量可以被视为选择要用于计算VBAP增益的网格的总数。

[0447] (量化处理和网格数量切换处理的组合)

[0448] 在前面的描述中,作为用于减少渲染处理的处理量的技术,描述了量化处理和网格数量切换处理。

[0449] 在执行渲染处理的渲染器侧,可以固定地使用被描述为量化处理或网格数量切换处理的一些处理,或者可以对这样的处理进行切换或者可以对这样的处理适当地进行组合。

[0450] 例如,可以基于对象的总数(在下文中被称为对象数量)、包括在对象的元数据中的重要性信息、对象的音频信号的声压等来确定要组合地执行哪些处理。此外,可以针对每个对象或音频信号的每个帧执行处理的组合,即,处理的切换。

[0451] 例如,在响应于对象数量执行处理的切换的情况下,可以执行如下所述这样的处理。

[0452] 例如,在对象数量等于或大于10的情况下,针对所有对象执行VBAP增益的二值化处理。相反,在对象数量小于10的情况下,照常仅执行在上文中描述的处理A1至处理A3。

[0453] 通过在对象数量少时照常执行处理,但是在对象数量多时以这种方式执行二值化处理,即使通过小硬件规模的渲染器也可以充分地执行渲染,并且可以获得尽可能高质量的声音。

[0454] 此外,当响应于对象数量执行处理的切换时,可以响应于对象数量执行网格数量切换处理,以适当地改变网格的总数。

[0455] 在这种情况下,例如,当对象数量等于或大于10时,可以将网格的总数设置为8,而当对象数量小于10时,可以将网格的总数设置为40。此外,可以响应于对象数量在多个阶段中改变网格的总数,使得网格的总数随着对象数量增加而减少。

[0456] 通过以这种方式响应于对象数量来改变网格的总数,可以响应于渲染器的硬件规模来调整处理量,从而获得尽可能高质量的声音。

[0457] 此外,在基于包括在对象的元数据中的重要性信息来执行处理的切换的情况下,可以执行以下处理。

[0458] 例如,当对象的重要性信息具有表示最高重要性程度的最高值时,仅照常执行处理A1至A3,但是在对象的重要性信息具有除了最高值以外的值的情况下,执行对VBAP增益的二值化处理。

[0459] 此外,例如,可以响应于对象的重要性信息的值来执行网格数量切换处理,以适当地改变网格的总数。在这种情况下,随着对象的重要性程度增加,可以增加网格的总数,并且可以在多个阶段中改变网格的总数。

[0460] 在那些示例中,可以基于每个对象的重要性信息针对每个对象切换该处理。在在此描述的处理中,可以提高具有高重要性程度的对象的声音质量,但是降低具有低重要性程度的对象的声音质量,从而减少处理量。因此,当要同时再现各种重要性程度的对象的声音时,最大限度地抑制听觉上的声音质量劣化以减少处理量,并且可以考虑这是在确保声音质量与减少处理量之间得到了很好的平衡的技术。

[0461] 以这种方式,当基于对象的重要性信息对每个对象执行处理的切换时,可以随着

对象的重要性程度增加而增加对象的总数,或者可以避免当对象的重要性程度高时执行量化处理。

[0462] 另外,对于重要性程度低的对象,即,对于其重要性信息的值低于预定值的对象,也可以针对位于靠近具有较高重要性程度的对象的位置处的对象(即,其重要性信息的值等于或高于预定值的对象)增加网格的总数或者可以不执行量化处理。

[0463] 具体地,对于其重要性信息指示最高值的对象,将网格的总数设置为40,但是对于其重要性信息不指示最高值的对象,减少网格的总数。

[0464] 在这种情况下,对于其重要性信息不是最高值的对象,可以随着对象与其重要性信息为最高值的对象之间的距离减小而增加网格的总数。通常,由于用户特别仔细地聆听高重要性程度的对象的声音,因此如果位于靠近所述对象的不同对象的声音的声音质量低,则用户将感到整个内容的声音质量不好。因此,通过还针对位于靠近具有高的重要性程度的对象处的对象来确定网格的总数,使得可以获得尽可能高质量的声音,可以抑制声音质量在听觉上的劣化。

[0465] 此外,可以响应于对象的音频信号的声压而切换处理。在此,可以通过计算音频信号的渲染目标的帧中的采样的采样值的均方值的平方根来确定音频信号的声压。具体地,可以通过计算以下表达式(10)来确定声压RMS:

[0466] [表达式10]

$$[0467] \quad RMS = 20 \times \log_{10} \left( \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} (X_n)^2} \right) \dots (10)$$

[0468] 应当注意的是,在表达式(10)中,N表示构成音频信号的帧的采样数, $x_n$ 表示帧中的第n(其中, $n=0, \dots, N-1$ )个采样的采样值。

[0469] 在响应于以这种方式获得的音频信号的声压RMS而切换处理的情况下,可以执行以下处理。

[0470] 例如,在对象的音频信号的声压RMS为-6dB或相对于0dB(其为声压RMS的满量程)的更大值的情况下,照例仅执行处理A1至A3,但是在对象的声压RMS低于-6dB的情况下,执行对VBAP增益的二值化处理。

[0471] 通常,在声音具有高声压的情况下,声音质量的劣化很可能突出,并且这样的声音通常是具有高重要性程度的对象的声音。因此,在此对于具有高声压RMS的声音的对象,防止声音质量劣化,同时对于具有低声压RMS的声音的对象,执行二值化处理,使得整体上减少处理量。由此,即使通过小硬件规模的渲染器,也可以充分地执行渲染,此外,可以获得尽可能高质量的声音。

[0472] 可替代地,可以响应于对象的音频信号的声压RMS而执行网格数量切换处理,使得适当地改变网格的总数。在这种情况下,例如,可以随着对象的声压RMS增加而增加网格的总数,并且可以在多个阶段中改变网格的总数。

[0473] 此外,可以响应于对象数量、重要性信息和声压RMS来选择量化处理或网格数量切换处理的组合。

[0474] 具体地,可以基于对象数量、重要性信息和声压RMS、是否要执行量化处理、在量化处理中要将VBAP增益量化为多少增益(即,在量化处理时的量化数量)以及要用于计算VBAP



增益的网格的总数通过根据选择的结果的处理来计算VBAP增益。在这样的情况下,例如,可以执行如下面给出的这样的处理。

[0475] 例如,在对象数量是10或更多的情况下,将网格的总数设置为10,此外,执行二值化处理。在这种情况下,由于对象数量大,因此通过减少网格的总数并且执行二值化处理来减少处理量。因此,即使在渲染器的硬件规模小的情况下,也可以执行所有对象的渲染。

[0476] 同时,在对象数量小于10并且此外重要性信息的值为最高值的情况下,仅照常执行处理A1至A3。因此,对于具有高重要性程度的对象,可以再现声音却不使声音质量劣化。

[0477] 在对象数量小于10并且此外重要性信息的值不是最高值并且此外声压RMS等于或高于-30dB的情况下,将网格的总数设置为10并且此外执行三值化处理。这使得可以将渲染处理时的处理量减少到以下程度:对于具有高声压的声音,虽然重要性程度低,但是声音的声音质量劣化不突出。

[0478] 此外,在对象数量小于10并且此外重要性信息的值不是最高值并且此外声压RMS低于-30dB的情况下,将网格的总数设置为5并且此外执行二值化处理。这使得可以针对具有低重要性程度并且具有低声压的声音充分地减少渲染处理时的处理量。

[0479] 以这种方式,当对象数量大时,减少渲染处理时的处理量,使得可以执行对所有对象的渲染,但是当对象数量小到一定程度时,选择适当的处理并且对每个对象执行渲染。因此,虽然对每个对象而言确保声音质量与减少处理装置得到了很好的平衡,但是整体上可以通过小的处理量以足够的声音质量再现声音。

[0480] <音频处理装置的配置的示例>

[0481] 现在,描述执行渲染处理同时适当地执行上述量化处理、网格数量切换处理等的音频处理装置。图17是描绘如刚刚描述的这样的音频处理装置的特定配置的示例的视图。应当注意的是,在图17中,用相同的附图标记表示与图6的情况下的那些对应的部分,因此适当地省略对它们的描述。

[0482] 图17中描绘的音频处理装置61包括获取单元21、增益计算单元23和增益调整单元71。增益计算单元23接收从获取单元21提供的对象的元数据和音频信号,针对每个对象计算扬声器12中的每个扬声器的VBAP增益,并且将所计算的VBAP增益提供至增益调整单元71。

[0483] 此外,增益计算单元23包括执行VBAP增益的量化的量化单元31。

[0484] 增益调整单元71针对每个对象将从获取单元21提供的音频信号与从增益计算单元23提供的各个扬声器12的VBAP增益相乘,以产生各个扬声器12的音频信号,并且将所述音频信号提供至扬声器12。

[0485] <再现处理的说明>

[0486] 随后,描述图17中描绘的音频处理装置61的操作。具体地,参照图18的流程图来描述通过音频处理装置61的再现处理。

[0487] 应当注意的是,在本示例中,假定针对每个帧将一个对象的音频信号和元数据或者多个对象中的每个对象提供至获取单元21,并且针对每个对象的音频信号的每个帧执行再现处理。

[0488] 在步骤S231处,获取单元21从外部获取对象的音频信号和元数据,并且将音频信号提供至增益计算单元23和增益调整单元71,同时它将元数据提供至增益计算单元23。此



外,获取单元21还针对在作为处理目标的帧中要同时再现哪些声音而获取对象的数量信息(即,对象数量的信息),并且将该信息提供至增益计算单元23。

[0489] 在步骤S232处,增益计算单元23基于从获取单元21提供的表示对象数量的信息来决定对象数量是否等于或大于10。

[0490] 如果在步骤S232处决定对象数量等于或大于10,则在步骤S233处,增益计算单元23将计算VBAP增益时要使用的网格的总数设置为10。换言之,增益计算单元23选择10作为网格的总数。

[0491] 此外,增益计算单元23响应于所选择的网格总数而从所有扬声器12中选择预定数量的扬声器12,使得在单位球体表面上形成数量等于总数的网格。然后,增益计算单元23将从所选择的扬声器12形成的单位球体表面上的10个网格确定为在计算VBAP增益时要使用的网格。

[0492] 在步骤S234处,增益计算单元23基于在步骤S233处确定的表示构成10个网格的扬声器12的位置的位置信息以及包括在从获取单元21提供的元数据中并且表示对象的位置的位置信息通过VBAP来计算每个扬声器12的VBAP增益。

[0493] 具体地,增益计算单元23按顺序使用在步骤S233处确定的网格连续地执行表达式(8)的计算,作为处理目标的网格以计算扬声器12的VBAP增益。此时,新的网格被连续地确定为处理目标的网格,直到针对构成处理目标的网格的三个扬声器12所计算的VBAP增益都指示等于或大于0的值为止,以连续地计算VBAP增益。

[0494] 在步骤S235处,量化单元31对在步骤S234处获得的扬声器12的VBAP增益进行二值化,然后处理前进至步骤S246。

[0495] 如果在步骤S232处决定对象数量小于10,则处理前进至步骤S236。

[0496] 在步骤S236处,增益计算单元23决定包括在从获取单元21提供的元数据中的对象的重要性信息的值是否为最高值。例如,如果重要性信息的值是指示重要性程度最高的值“7”,则决定重要性信息指示最高值。

[0497] 如果在步骤S236处决定重要性信息指示最高值,则处理前进至步骤S237。

[0498] 在步骤S237处,增益计算单元23基于表示扬声器12的位置的位置信息以及包括在从获取单元21提供的元数据中的位置信息来计算每个扬声器12的VBAP增益,然后处理前进至步骤S246。在此,从所有扬声器12形成的网格被连续地确定为处理目标的网格,并且通过计算表达式(8)来计算VBAP增益。

[0499] 另一方面,如果在步骤S236处决定重要性信息不指示最高值,则在步骤S238处,增益计算单元23计算从获取单元21提供的音频信号的声压RMS。具体地,针对作为处理目标的音频信号的帧执行在上文中给出的表达式(10)的计算,以计算声压RMS。

[0500] 在步骤S239处,增益计算单元23决定在步骤S238处计算的声压RMS是否等于或高于-30dB。

[0501] 如果在步骤S239处决定声压RMS等于或高于-30dB,则执行步骤S240和S241处的处理。应当注意的是,步骤S240和S241处的处理分别与步骤S233和S234处的处理类似,因此省略对它们的描述。

[0502] 在步骤S242处,量化单元31对在步骤S241处获得的每个扬声器12的VBAP增益进行三值化,然后处理前进至步骤S246。

[0503] 另一方面,如果在步骤S239处决定声压RMS低于-30dB,则处理前进至步骤S243。

[0504] 在步骤S243处,增益计算单元23将计算VBAP增益时要使用的网格的总数设置为5。

[0505] 此外,增益计算单元23响应于所选择的网格总数“5”从所有扬声器12中选择预定数量的扬声器12,并且将从所选择的扬声器12形成的单位球体表面上的五个网格确定为计算VBAP增益时要使用的网格。

[0506] 在确定计算VBAP增益时要使用的网格之后,执行步骤S244和S245处的处理,然后处理前进至步骤S246。应当注意的是,步骤S244和S245处的处理与步骤S234和S235处的处理类似,因此省略对它们的描述。

[0507] 在执行步骤S235、S237、S242或S245处的处理并且获得扬声器12的VBAP增益之后,执行步骤S246至S248处的处理,从而结束再现处理。

[0508] 应当注意的是,由于步骤S246至S248处的处理分别与在上文参照图7描述的步骤S17至S19处的处理类似,因此省略对它们的描述。

[0509] 然而,更具体地,针对各个对象基本上同时执行再现处理,并且在步骤S248处,将针对各个对象获得的扬声器12的音频信号提供至扬声器12。具体地,扬声器12基于通过将对象的音频信号相加而获得的信号来再现声音。因此,同时输出所有对象的声音。

[0510] 音频处理装置61适当地对每个对象选择性地执行量化处理和网格数量切换处理。由此,可以减少渲染处理的处理量,同时抑制呈现或声音质量的劣化。

[0511] <修改1至第二实施方式>

[0512] <音频处理装置的配置示例>

[0513] 此外,在第二实施方式的描述中,虽然描述了当不执行用于延伸声像的处理时选择性地执行量化处理或网格数量切换处理的示例,此外,当执行用于延伸声像的处理时,可以选择性地执行量化处理或网格数量切换处理。

[0514] 在这样的情况下,音频处理装置11例如以如图19中描绘的这样的方式来配置。应当注意的是,在图19中,用相同的附图标记表示与图6或图17的情况下的那些对应的部分,因此适当地省略对它们的描述。

[0515] 图19中描绘的音频处理装置11包括获取单元21、向量计算单元22、增益计算单元23和增益调整单元71。

[0516] 获取单元21针对一个或多个对象来获取对象的音频信号和元数据,并且将所获取的音频信号提供至增益计算单元23和增益调整单元71,并且将所获取的元数据提供至向量计算单元22和增益计算单元23。此外,增益计算单元23包括量化单元31。

[0517] <再现处理的说明>

[0518] 现在,参照图20的流程图来描述由图19中描绘的音频处理装置11执行的再现处理。

[0519] 应当注意的是,在本示例中,假定针对一个或多个对象,针对每个帧将对象的音频信号和元数据提供至获取单元21,并且针对每个对象的音频信号的每个帧来执行再现处理。

[0520] 此外,由于步骤S271和S272处的处理分别与图7的步骤S11和S12处的处理类似,因此省略对它们的描述。然而,在步骤S271处,将获取单元21获取的音频信号提供至增益计算单元23和增益调整单元71,并且将由获取单元21获取的元数据提供至向量计算单元22和增

益计算单元23。

[0521] 当执行步骤S271和S272处的处理时,获得扩展向量或扩展向量和向量p。

[0522] 在步骤S273处,增益计算单元23执行VBAP增益计算处理以计算每个扬声器12的VBAP增益。应当注意的是,虽然在下文中描述了VBAP增益计算处理的细节,但是在VBAP增益计算处理中,选择性地执行量化处理或网格数量切换处理,以计算每个扬声器12的VBAP增益。

[0523] 在执行步骤S273处的处理并且获得扬声器12的VBAP增益之后,执行步骤S274至S276处的处理并且再现处理结束。然而,由于那些处理分别与图7的步骤S17至S19处的处理类似,因此省略对它们的描述。然而,更具体地,针对对象基本上同时执行再现处理,并且在步骤S276处,将针对各个对象获得的扬声器12的音频信号提供至扬声器12。因此,从扬声器12同时输出所有对象的声音。

[0524] 音频处理装置11以如上所述这样的方式适当地对每个对象选择性地执行量化处理或网格数量切换处理。由此,在执行用于延伸声像的处理的情况下,也可以减少渲染处理的处理量,同时抑制呈现或声音质量的劣化。

[0525] <VBAP增益计算处理的说明>

[0526] 现在,参照图21的流程图来描述与图20的步骤S273处的处理对应的VBAP增益计算处理。

[0527] 应当注意的是,由于步骤S301至S303处的处理分别与图18的步骤S232至S234处的处理类似,因此省略对它们的描述。然而,在步骤S303处,针对扩展向量的向量或者扩展向量和向量p中的每一个来计算每个扬声器12的VBAP增益。

[0528] 在步骤S304处,增益计算单元23将针对向量计算的每个扬声器12的VBAP增益相加,以计算VBAP增益相加值。在步骤S304处,执行与图7的步骤S14处的处理类似的处理。

[0529] 在步骤S305处,量化单元31对通过步骤S304处的处理获得的每个扬声器12的VBAP增益相加值进行二值化,然后计算处理结束,此外处理前进至图20的步骤S274。

[0530] 另一方面,如果在步骤S301处决定对象数量小于10,则执行步骤S306和S307处的处理。

[0531] 应当注意的是,由于步骤S306和S307处的处理分别与图18的步骤S236和步骤S237处的处理类似,因此省略对它们的描述。然而,在步骤S307处,针对扩展向量的向量或者扩展向量和向量p中的每一个来计算每个扬声器12的VBAP增益。

[0532] 此外,在执行步骤S307处的处理之后,执行步骤S308处的处理,并且VBAP增益计算处理结束,此后处理前进至图20的步骤S274。然而,由于步骤S308处的处理与步骤S304处的处理类似,因此省略其描述。

[0533] 此外,如果在步骤S306处决定重要性信息不指示最高值,则执行步骤S309至S312处的处理。然而,由于所述处理与图18的步骤S238至S241处的处理类似,因此省略对它们的描述。然而,在步骤S312处,针对扩展向量的向量或者扩展向量和向量p中的每一个来计算每个扬声器12的VBAP增益。

[0534] 在针对向量获得扬声器12的VBAP增益之后,执行步骤S313处的处理以计算VBAP增益相加值。然而,由于步骤S313处的处理与步骤S304处的处理类似,因此省略其描述。

[0535] 在步骤S314处,量化单元31对通过步骤S313处的处理获得的每个扬声器12的VBAP

增益相加值进行三值化,并且VBAP增益计算结束,此后处理前进至图20的步骤S274。

[0536] 此外,如果在步骤S310处决定声压RMS低于-30dB,则执行步骤S315处的处理,并且将计算VBAP增益时要使用的网格的总数设置为5。应当注意的是,步骤S315处的处理与图18的步骤S243处的处理类似,因此省略其描述。

[0537] 在确定计算VBAP增益时要使用的网格之后,执行步骤S316至S318处的处理,并且结束VBAP增益计算处理,此后处理前进至图20的步骤S274。应当注意的是,步骤S316至S318处的处理与步骤S303至S305处的处理类似,因此省略对它们的描述。

[0538] 音频处理装置11以如上所述这样的方式适当地对每个对象选择性地执行量化处理或网格数量切换处理。由此,还在执行用于延伸声像的处理的情况下,可以减少渲染处理的处理量,同时抑制呈现或声音质量的劣化。

[0539] 顺便提及,虽然上述一系列处理可以由硬件执行,但是其可以以其他方式由软件来执行。在一系列处理由软件执行的情况下,将构造软件的程序安装到计算机中。在此,计算机包括并入硬件中的专用计算机,例如,可以通过安装各种程序来执行各种功能的通用个人计算机等。

[0540] 图22是描绘根据程序执行在上文中描述的一系列处理的计算机的硬件的配置的示例的框图。

[0541] 在计算机中,CPU(中央处理单元)501、ROM(只读存储器)502和RAM(随机存取存储器)503通过总线504彼此连接。

[0542] 输入/输出接口505还连接至总线504。输入单元506、输出单元507、记录单元508、通信单元509和驱动器510连接至输入/输出接口505。

[0543] 输入单元506由键盘、鼠标、麦克风、图像拾取元件等构成。输出单元507由显示单元、扬声器等构成。记录单元508由硬盘、非易失性存储器等构成。通信单元509由网络接口等构成。驱动器510驱动可移除记录介质511(如磁盘、光盘、磁光盘或半导体存储器)。

[0544] 在以如上所述这样的方式配置的计算机中,CPU 501通过输入/输出接口505和总线504将例如记录在记录单元508中的程序加载到RAM 503中,并且执行程序以执行在上文中描述的一系列处理。

[0545] 由计算机(CPU 501)执行的程序可以记录在可移除记录介质511上并且被设置为可移除记录介质511(例如,如封装介质等)。此外,可以通过诸如局域网、因特网或数字卫星广播的有线或无线传输介质来提供程序。

[0546] 在计算机中,可以通过将可移除记录介质511加载到驱动器510中,通过输入/输出接口505将程序安装到记录单元508中。可替代地,程序可以由通信单元509通过有线或无线传输介质来接收并且被安装到记录单元508中。可替代地,程序可以预先安装到ROM 502或记录单元508中。

[0547] 应当注意的是,由计算机执行的程序可以是根据本说明书中描述的顺序按照时间序列执行处理的程序或者并行执行处理或在调用程序的定时处执行处理的程序等。

[0548] 此外,本技术的实施方式不限于在上文中描述的实施方式,并且在不脱离本技术的主题的情况下,可以以各种方式被改变。

[0549] 例如,本技术可以采用云计算的配置,通过该配置,一个功能由多个装置通过网络分担和协同处理。

[0550] 此外,参考在上文中描述的流程图描述的步骤可以由单个装置执行,或者可以由多个装置分担执行。

[0551] 此外,在一个步骤包括多个处理的情况下,包括在一个步骤中的多个处理可以由单个装置执行或者可以由多个装置分担执行。

[0552] 本技术也可以采取以下配置。

[0553] (1) 一种音频处理装置,包括:

[0554] 获取单元,被配置成获取包括指示音频对象的位置的位置信息以及由至少二维或更多维的向量构成并且表示声像距所述位置的扩散的声像信息的元数据;

[0555] 向量计算单元,被配置成基于表示由所述声像信息确定的声像的扩散的区域的水平方向角和竖直方向角来计算指示所述区域中的位置的扩展向量;以及

[0556] 增益计算单元,被配置成基于所述扩展向量来计算提供给位于所述位置信息指示的位置附近的两个或更多个声音输出单元的音频信号中的每个音频信号的增益。

[0557] (2) 根据(1)所述的音频处理装置,其中,

[0558] 所述向量计算单元基于所述水平方向角与所述竖直方向角之间的比率来计算所述扩展向量。

[0559] (3) 根据(1)或(2)所述的音频处理装置,其中,

[0560] 所述向量计算单元计算预先确定的数量的扩展向量。

[0561] (4) 根据(1)或(2)所述的音频处理装置,其中,

[0562] 所述向量计算单元计算任意可变数量的扩展向量。

[0563] (5) 根据(1)所述的音频处理装置,其中,

[0564] 所述声像信息是指示所述区域的中心位置的向量。

[0565] (6) 根据(1)所述的音频处理装置,其中,

[0566] 所述声像信息是指示所述声像距所述区域的中心的扩散程度的二维或更多维的向量。

[0567] (7) 根据(1)所述的音频处理装置,其中,

[0568] 所述声像信息是指示从所述位置信息指示的位置观看到的所述区域的中心位置的相对位置的向量。

[0569] (8) 根据(1)至(7)中任一项所述的音频处理装置,其中,

[0570] 所述增益计算单元:

[0571] 计算关于所述声音输出单元中的每个声音输出单元的每个扩展向量的增益,

[0572] 计算针对所述声音输出单元中的每个声音输出单元的扩展向量所计算的增益的相加值,

[0573] 将所述相加值量化成关于所述声音输出单元中的每个声音输出单元的两个或更多个值的增益,以及

[0574] 基于量化的相加值来计算关于所述声音输出单元中的每个声音输出单元的最终增益。

[0575] (9) 根据(8)所述的音频处理装置,其中,

[0576] 所述增益计算单元选择要用于计算所述增益的网格的数量,所述网格中的每个网格是由所述声音输出单元中的三个声音输出单元包围的区域,并且所述增益计算单元基于

对网格的数量的选择结果和所述扩展向量来计算所述扩展向量中的每个扩展向量的增益。

[0577] (10) 根据 (9) 所述的音频处理装置, 其中,

[0578] 所述增益计算单元选择要用于计算所述增益的网格的数量、是否要执行量化以及量化时的所述相加值的量化数量, 以及

[0579] 响应于选择结果来计算最终增益。

[0580] (11) 根据 (10) 所述的音频处理装置, 其中,

[0581] 所述增益计算单元基于所述音频对象的数量来选择要用于计算所述增益的网格的数量、是否要执行量化以及所述量化数量。

[0582] (12) 根据 (10) 或 (11) 所述的音频处理装置, 其中,

[0583] 所述增益计算单元基于所述音频对象的重要性程度来选择要用于计算所述增益的网格的数量、是否要执行量化以及所述量化数量。

[0584] (13) 根据 (12) 所述的音频处理装置, 其中,

[0585] 所述增益计算单元选择要用于计算所述增益的网格的数量, 使得要用于计算所述增益的网格的数量随着所述音频对象的位置位于更靠近重要性程度高的音频对象而增加。

[0586] (14) 根据 (10) 至 (13) 中任一项所述的音频处理装置, 其中,

[0587] 所述增益计算单元基于所述音频对象的音频信号的声压来选择要用于计算所述增益的网格的数量、是否要执行量化以及所述量化数量。

[0588] (15) 根据 (9) 至 (14) 中任一项所述的音频处理装置, 其中,

[0589] 所述增益计算单元响应于对网格数量的选择结果来选择包括位于彼此不同高度处的声音输出单元的多个声音输出单元中的三个或更多个声音输出单元, 并且基于由所选择的聲音输出单元形成的一个或多个网格来计算增益。

[0590] (16) 一种音频处理方法, 包括以下步骤:

[0591] 获取包括指示音频对象的位置的位置信息以及由至少二维或更多维的向量构成并且表示声像距所述位置的扩散的声像信息的元数据;

[0592] 基于表示由所述声像信息确定的声像的扩散的区域的水平方向角和竖直方向角来计算指示所述区域中的位置的扩展向量; 以及

[0593] 基于所述扩展向量来计算提供给位于所述位置信息指示的位置附近的两个或更多个声音输出单元的音频信号中的每个音频信号的增益。

[0594] (17) 一种使计算机执行以下处理的程序, 所述处理包括以下步骤:

[0595] 获取包括指示音频对象的位置的位置信息以及由至少二维或更多维的向量构成并且表示声像距所述位置的扩散的声像信息的元数据;

[0596] 基于表示由所述声像信息确定的声像的扩散的区域的水平方向角和竖直方向角来计算指示所述区域中的位置的扩展向量; 以及

[0597] 基于所述扩展向量来计算提供给位于所述位置信息指示的位置附近的两个或更多个声音输出单元的音频信号中的每个音频信号的增益。

[0598] (18) 一种音频处理装置, 包括:

[0599] 获取单元, 被配置成获取包括指示音频对象的位置的位置信息的元数据; 以及

[0600] 增益计算单元, 被配置成选择要用于计算要提供给所述声音输出单元的音频信号的增益的网格的数量, 所述网格中的每个网格由三个声音输出单元包围, 并且所述增益计

算单元基于对网格的数量的选择结果和所述位置信息来计算增益。

[0601] [参考标记列表]

[0602] 11音频处理装置,21获取单元,22向量计算单元,23增益计算单元,24增益调整单元,31量化单元,61音频处理装置,71增益调整单元。

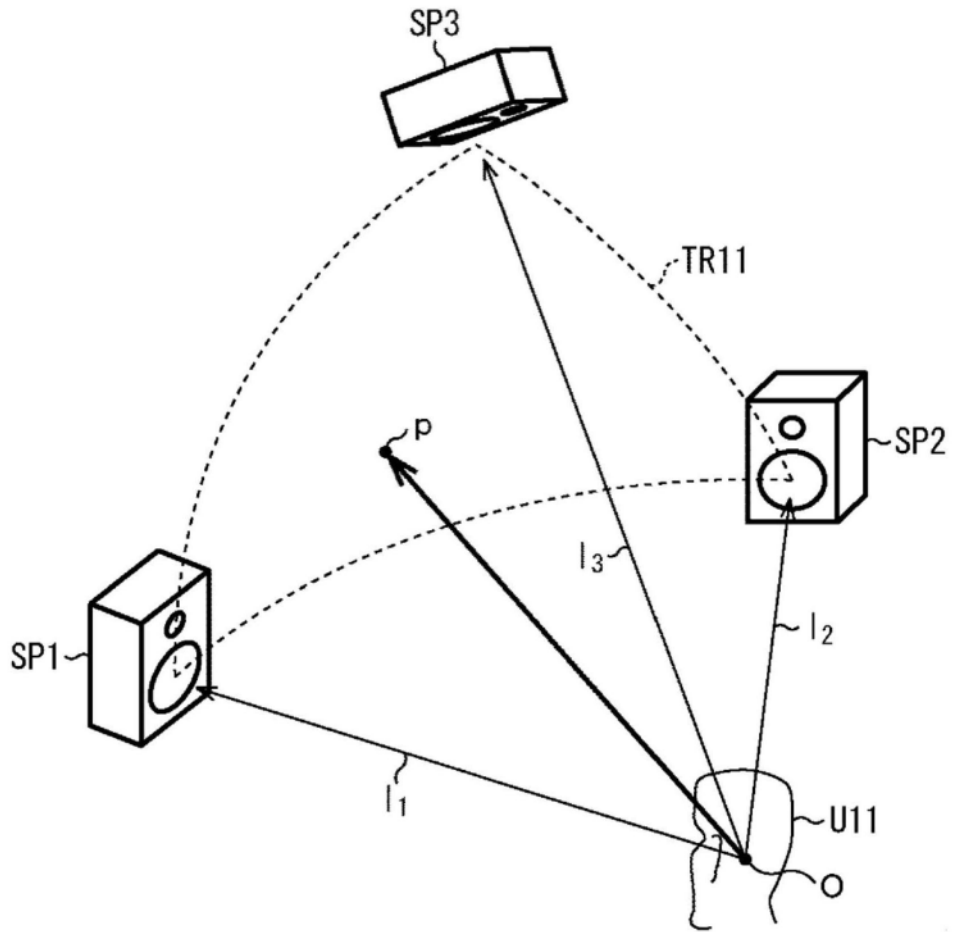


图1

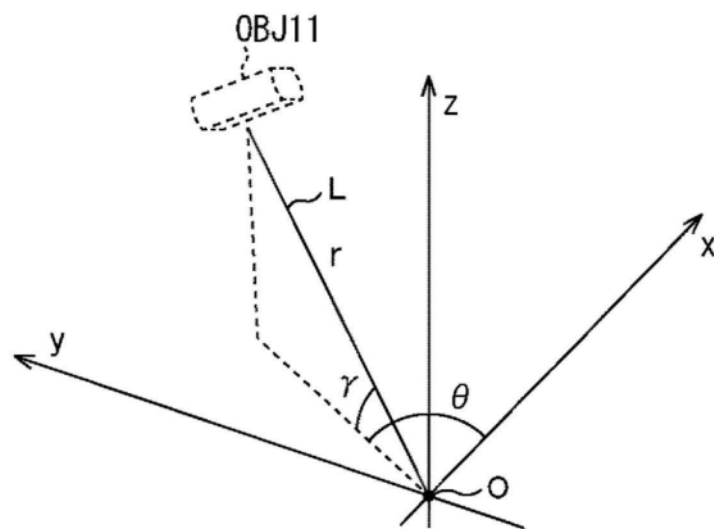


图2





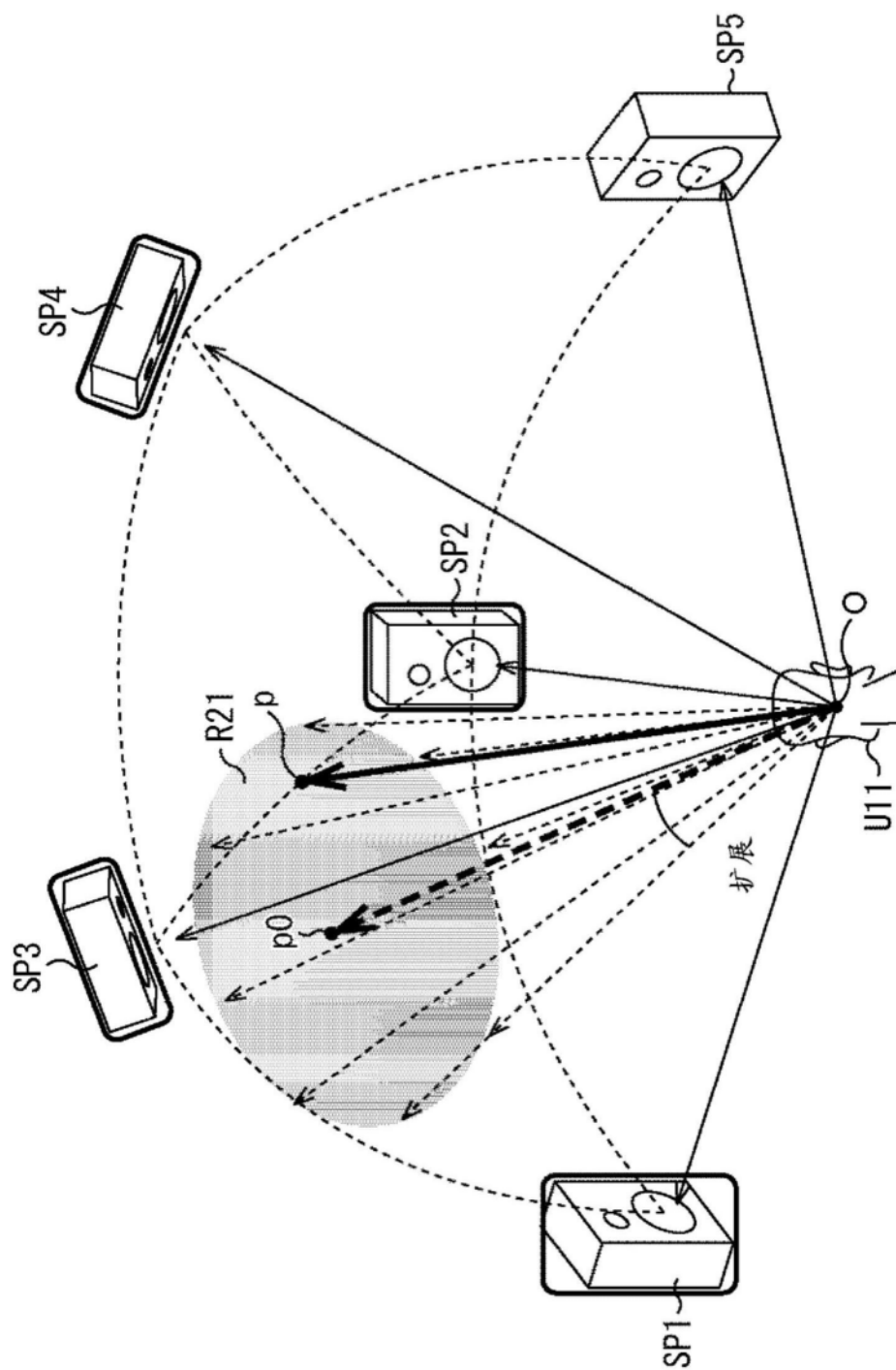


图4

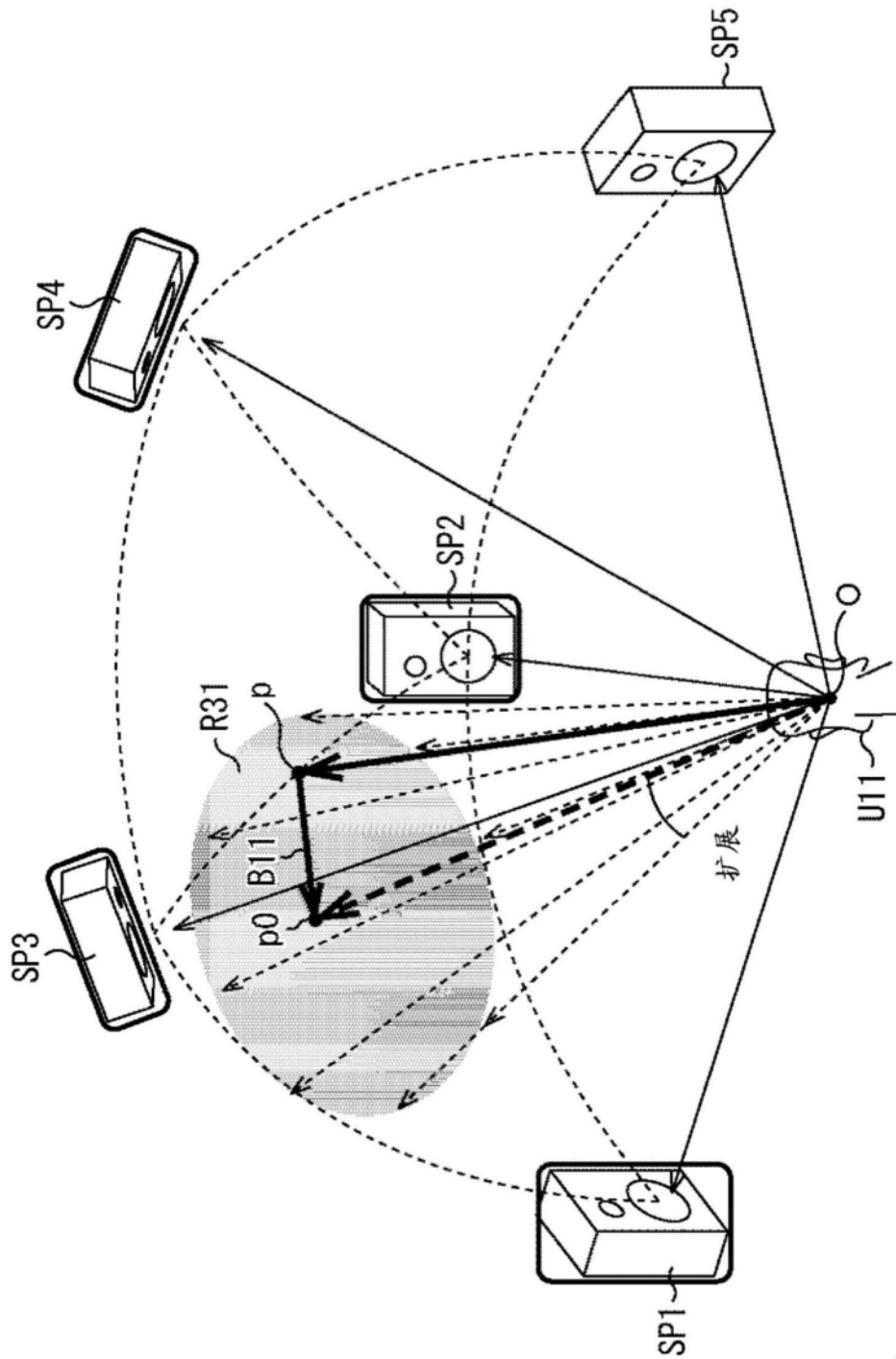


图5

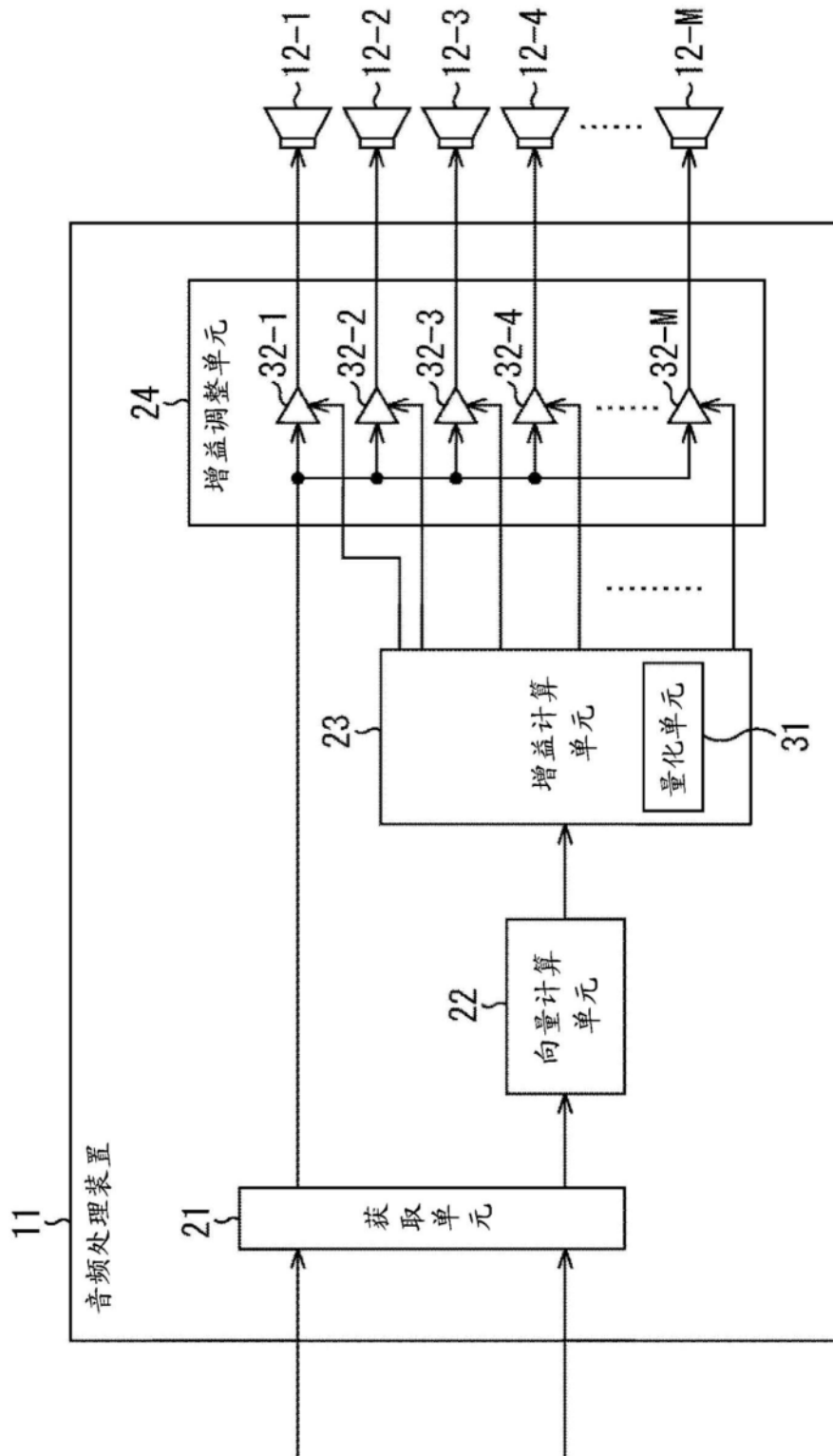


图6

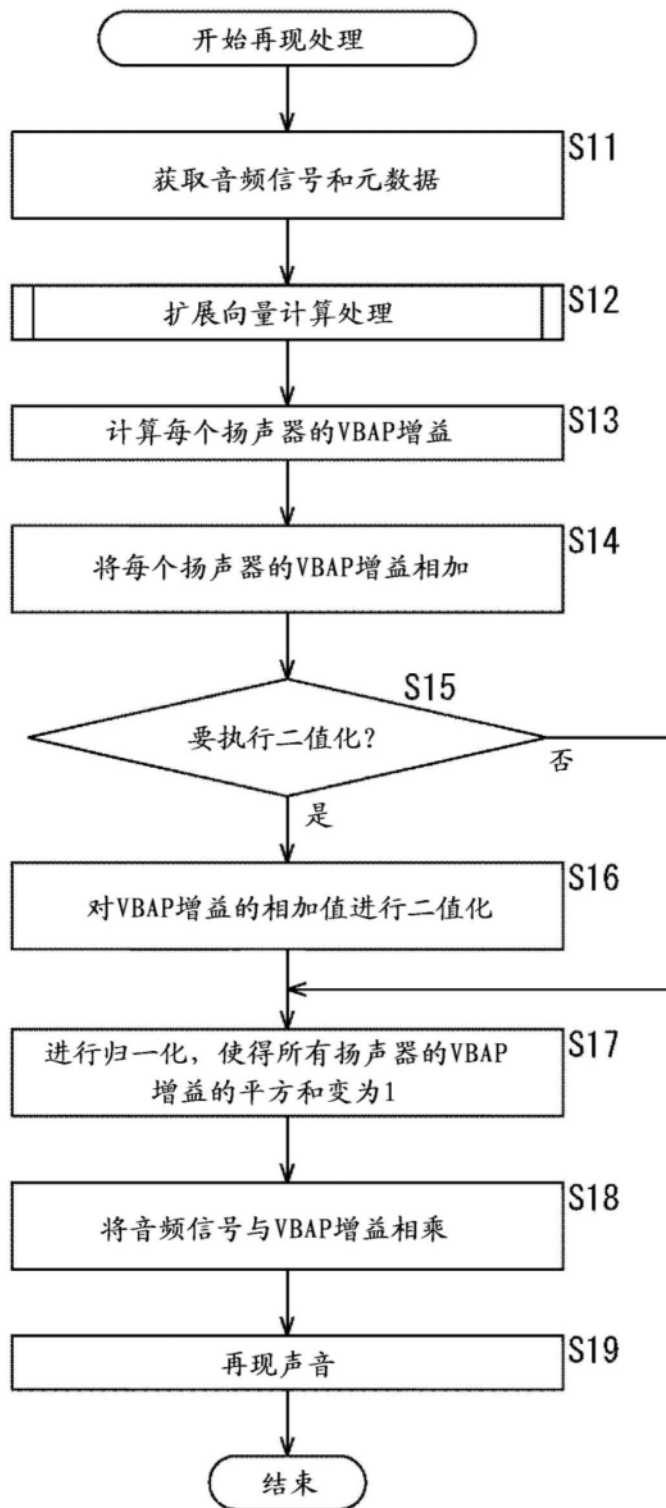


图7

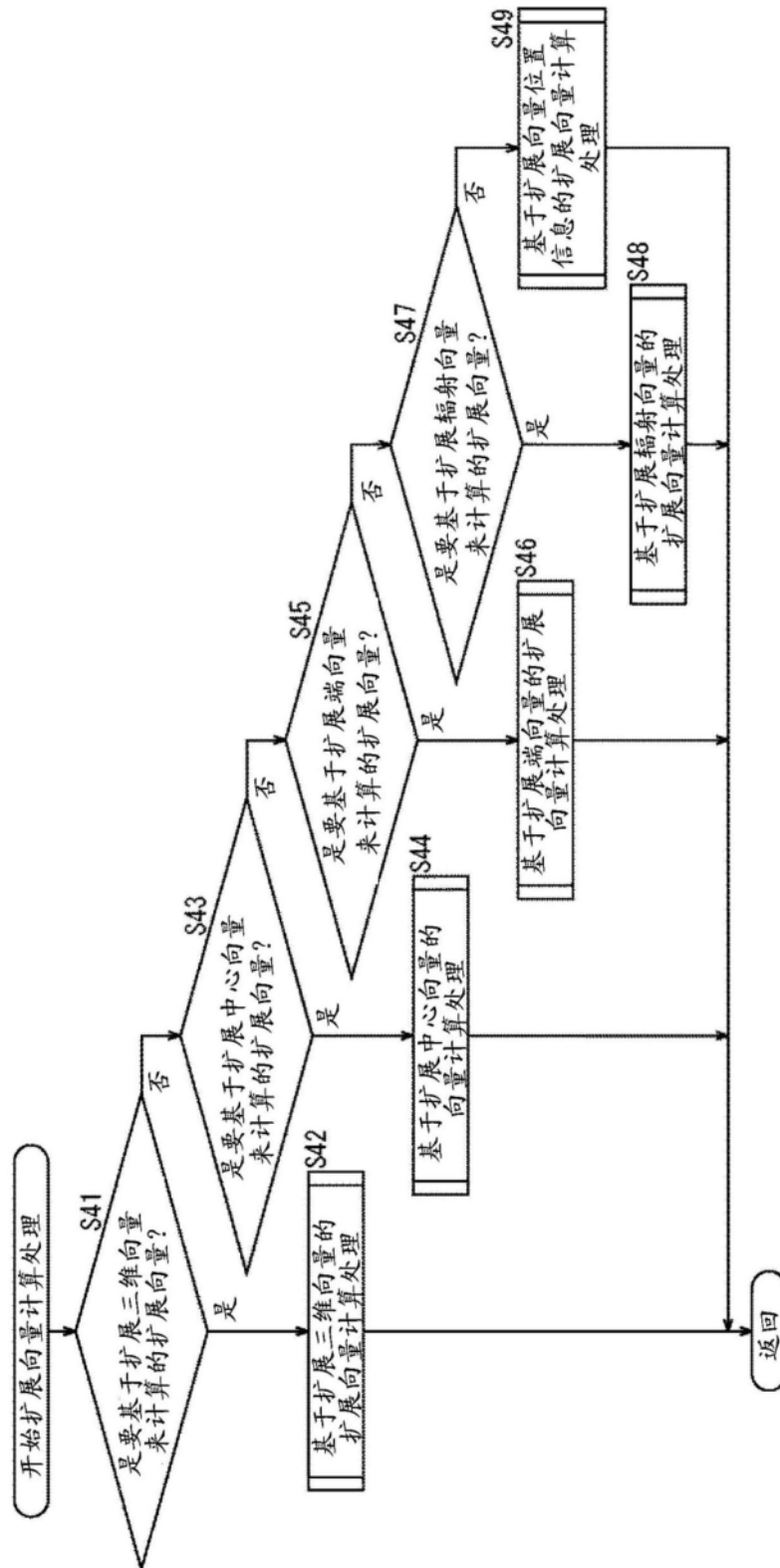


图8

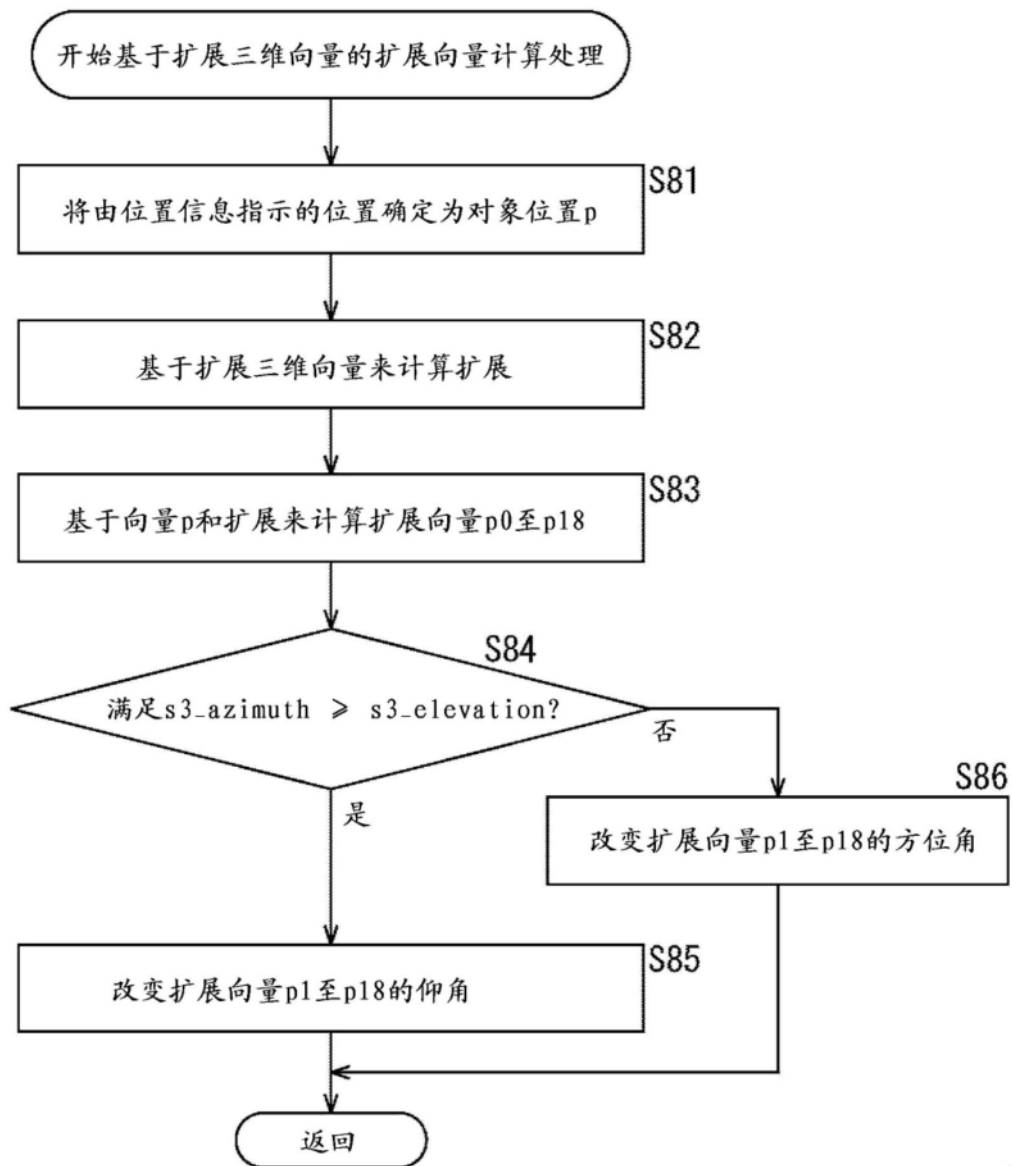


图9

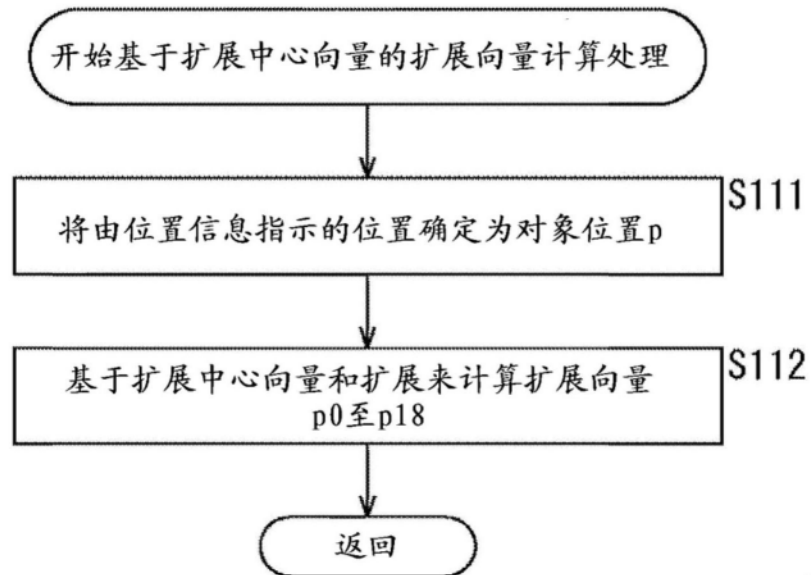


图10



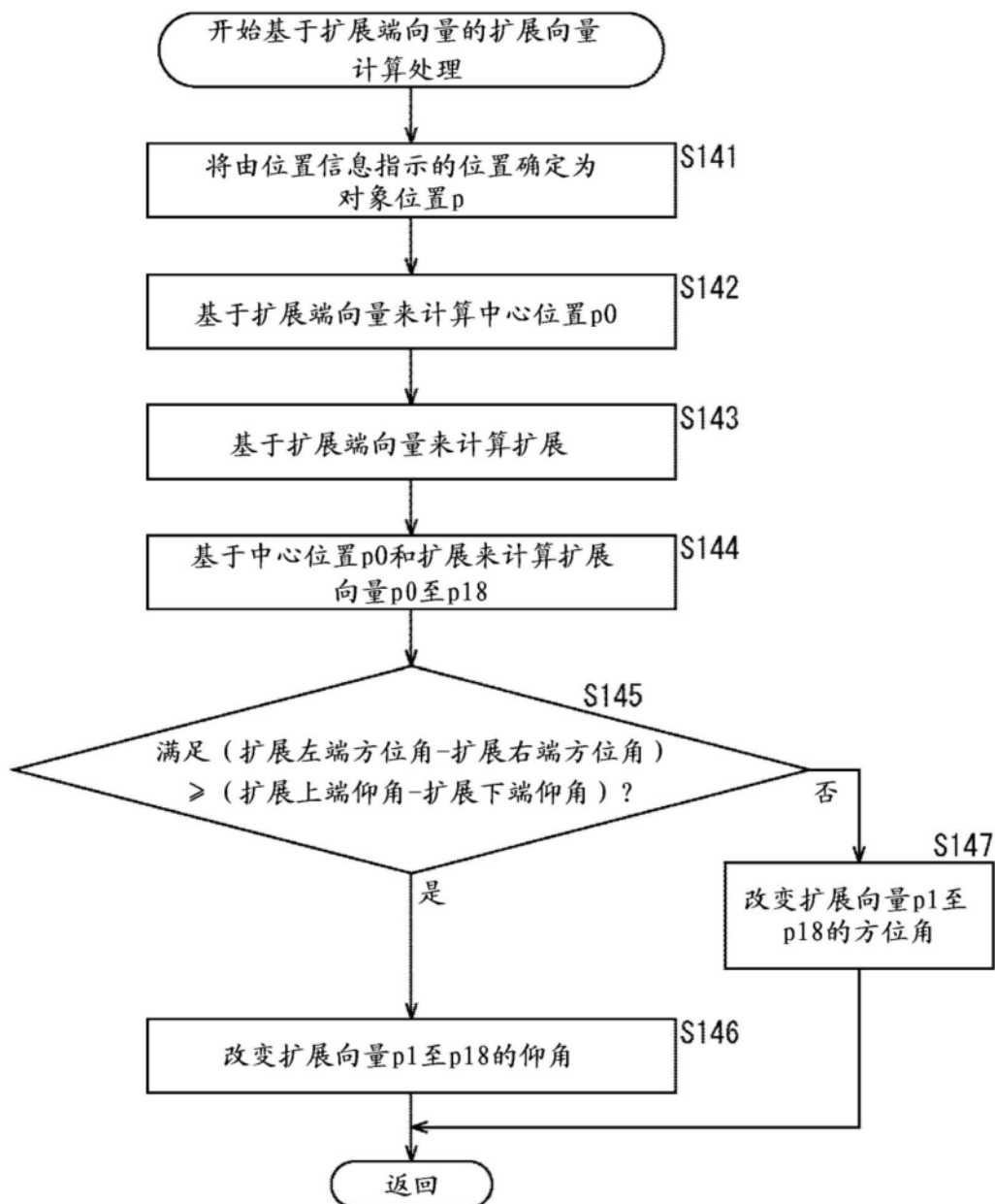


图11

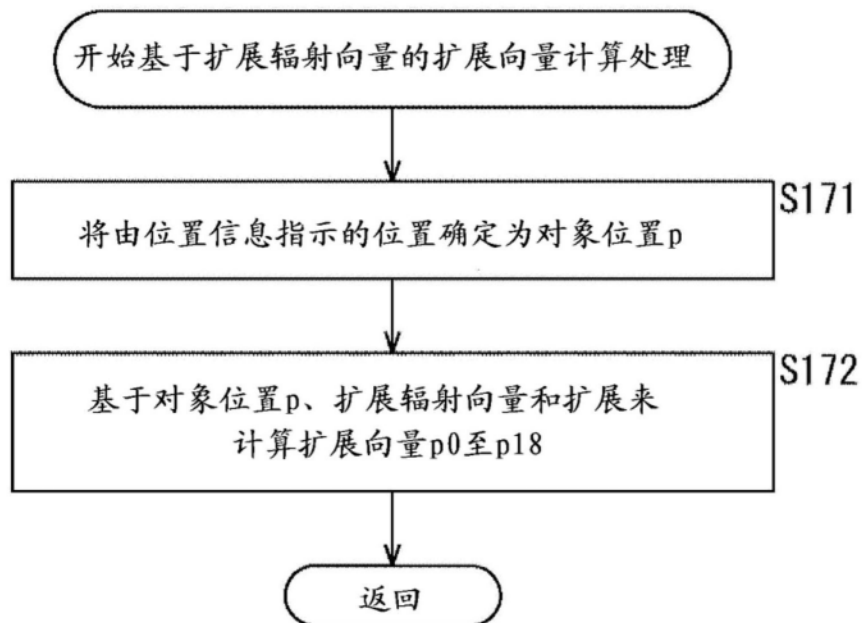


图12

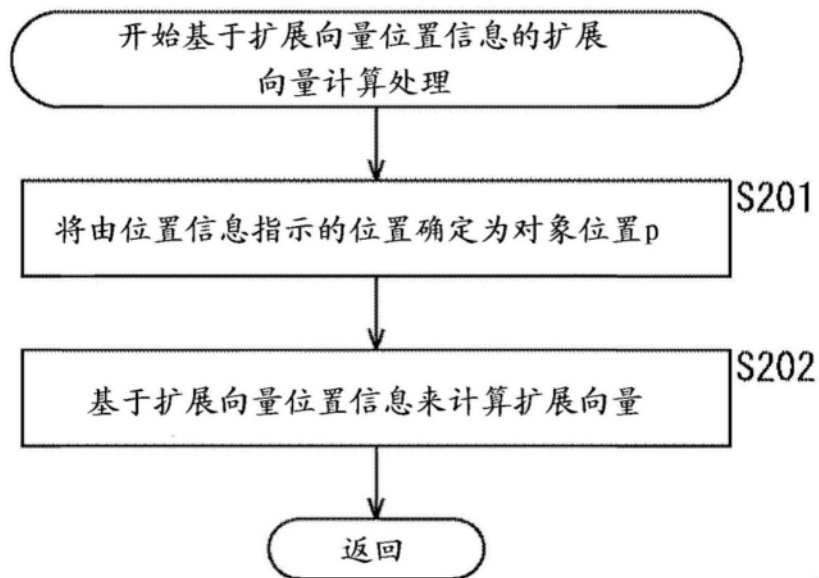


图13

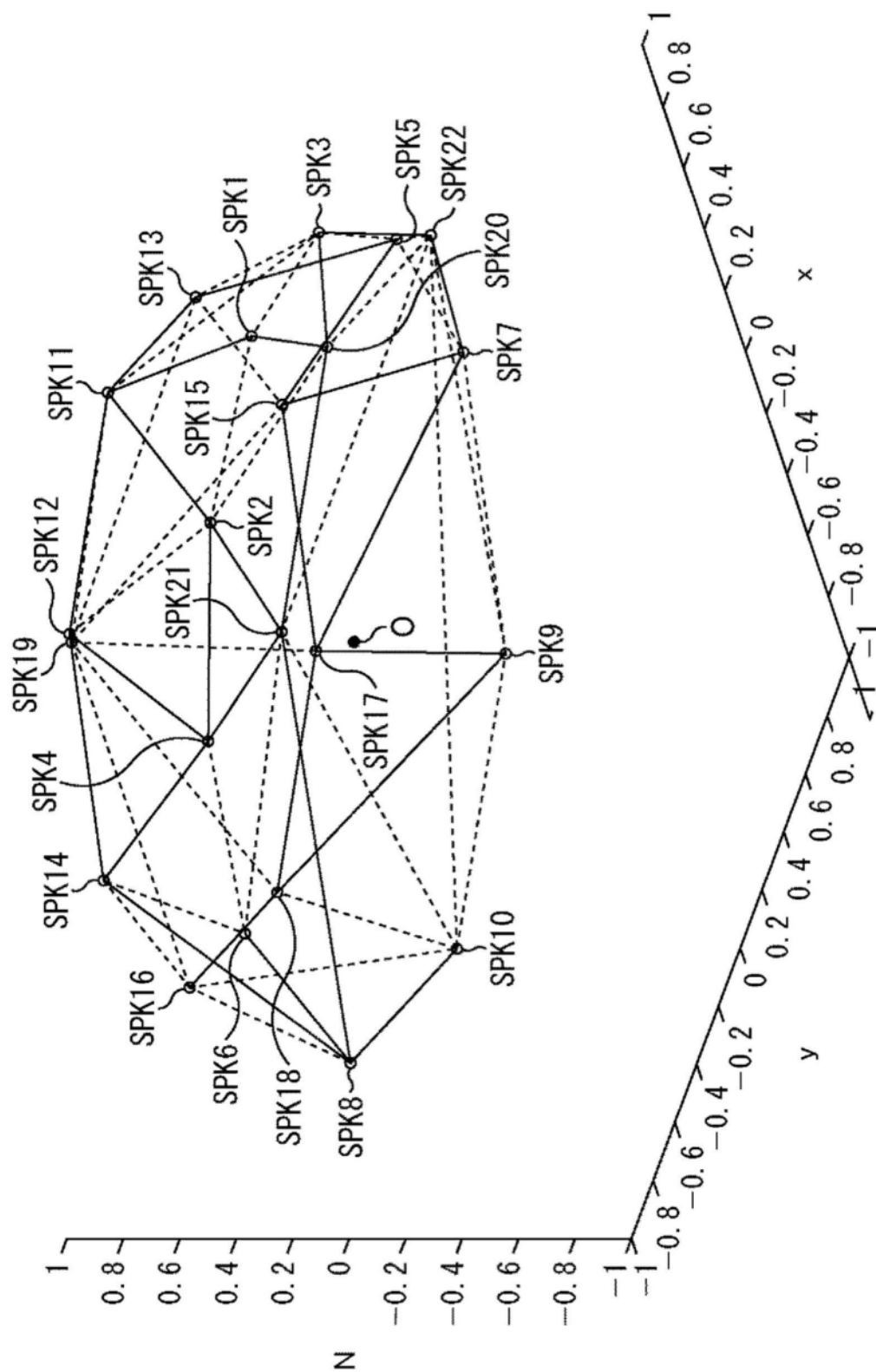


图14

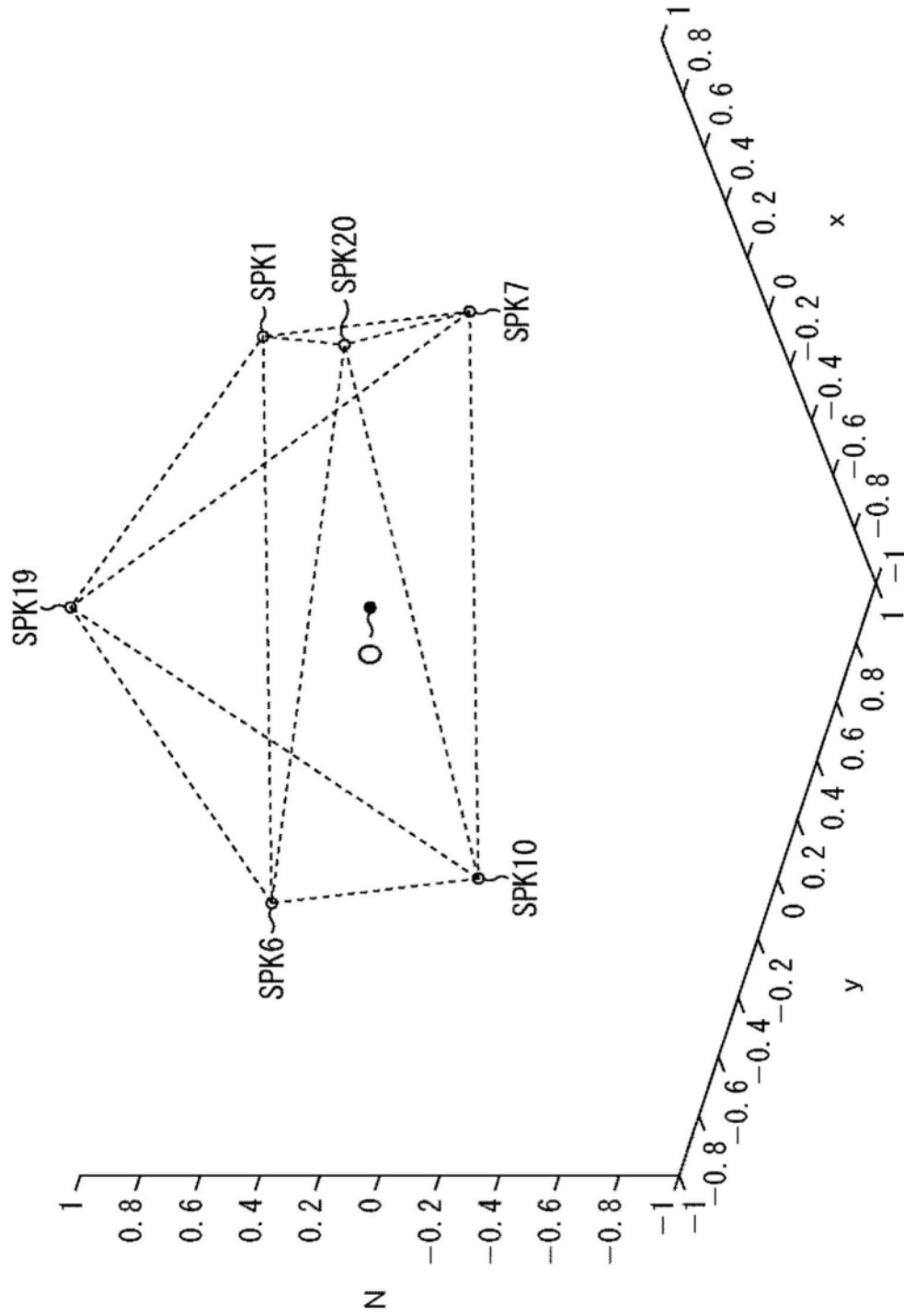


图15

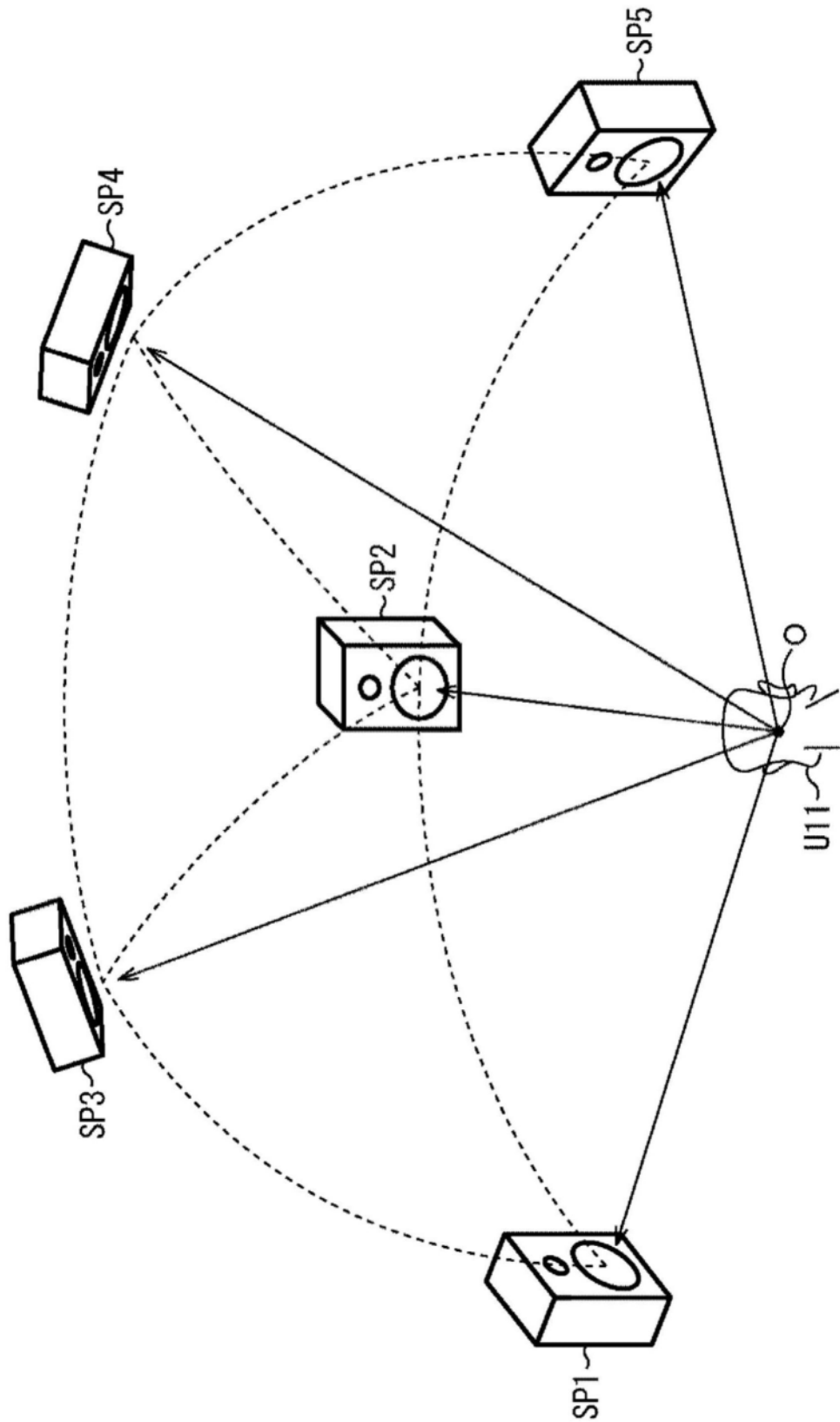


图16

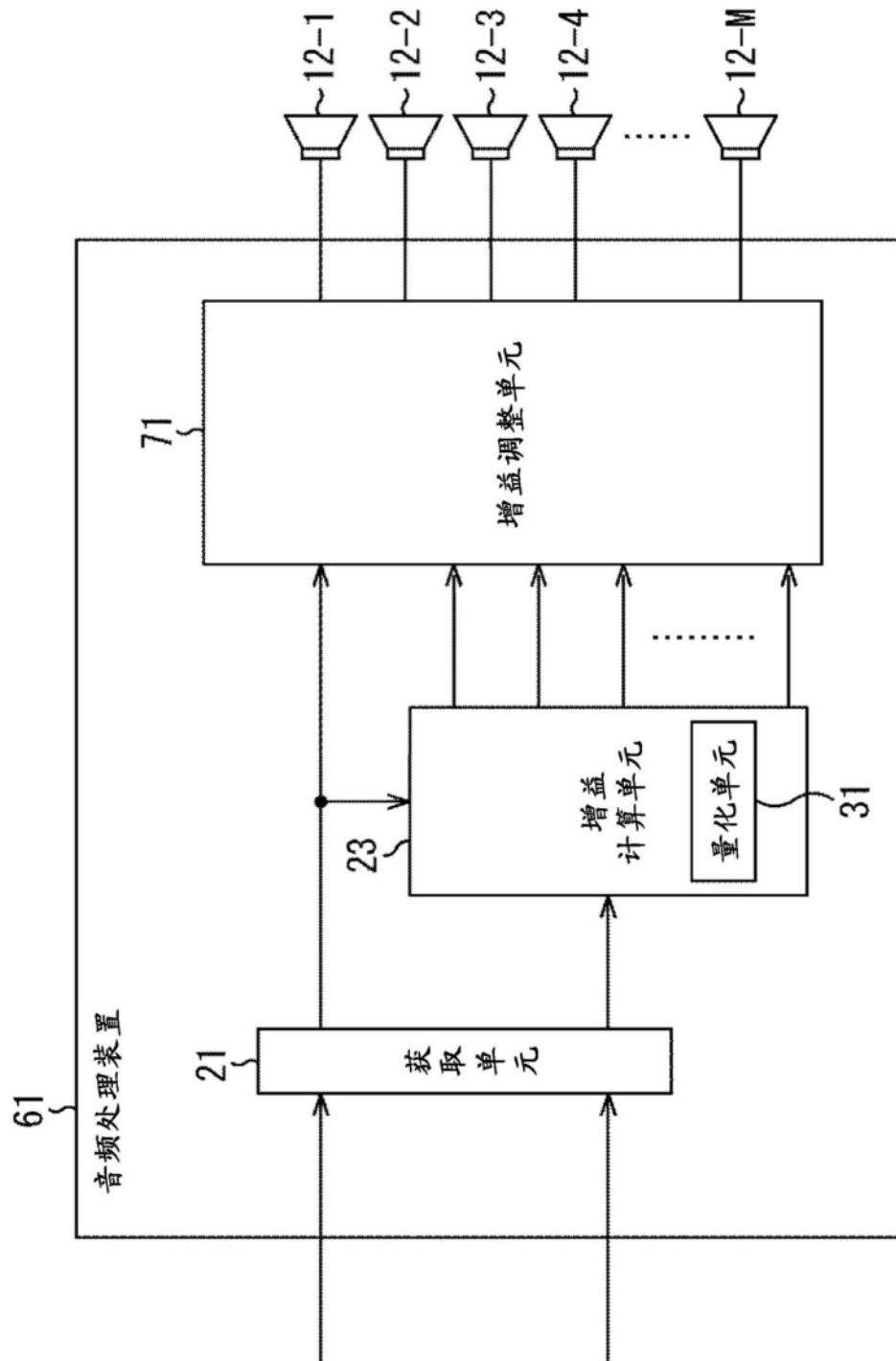


图17

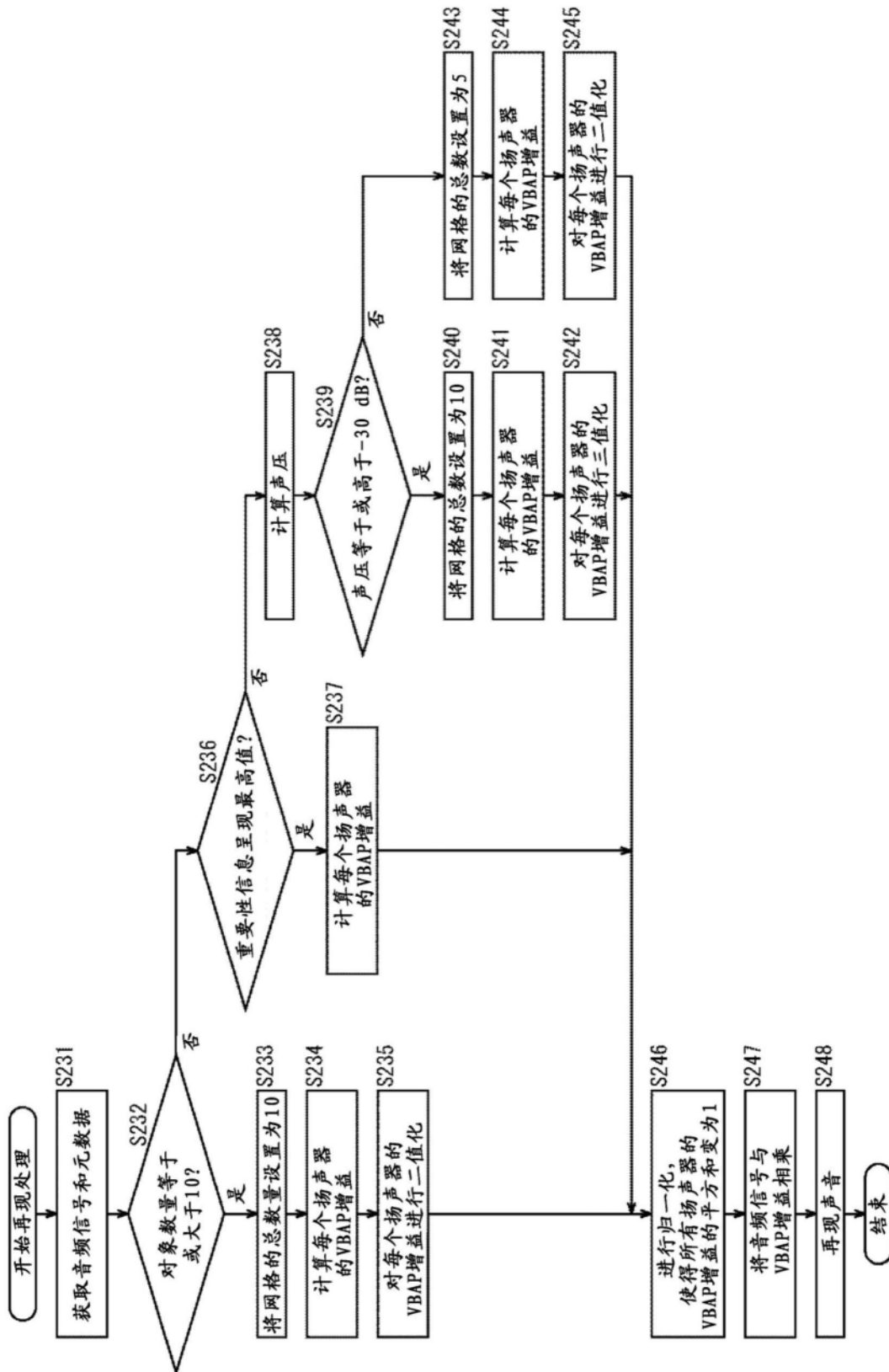


图18

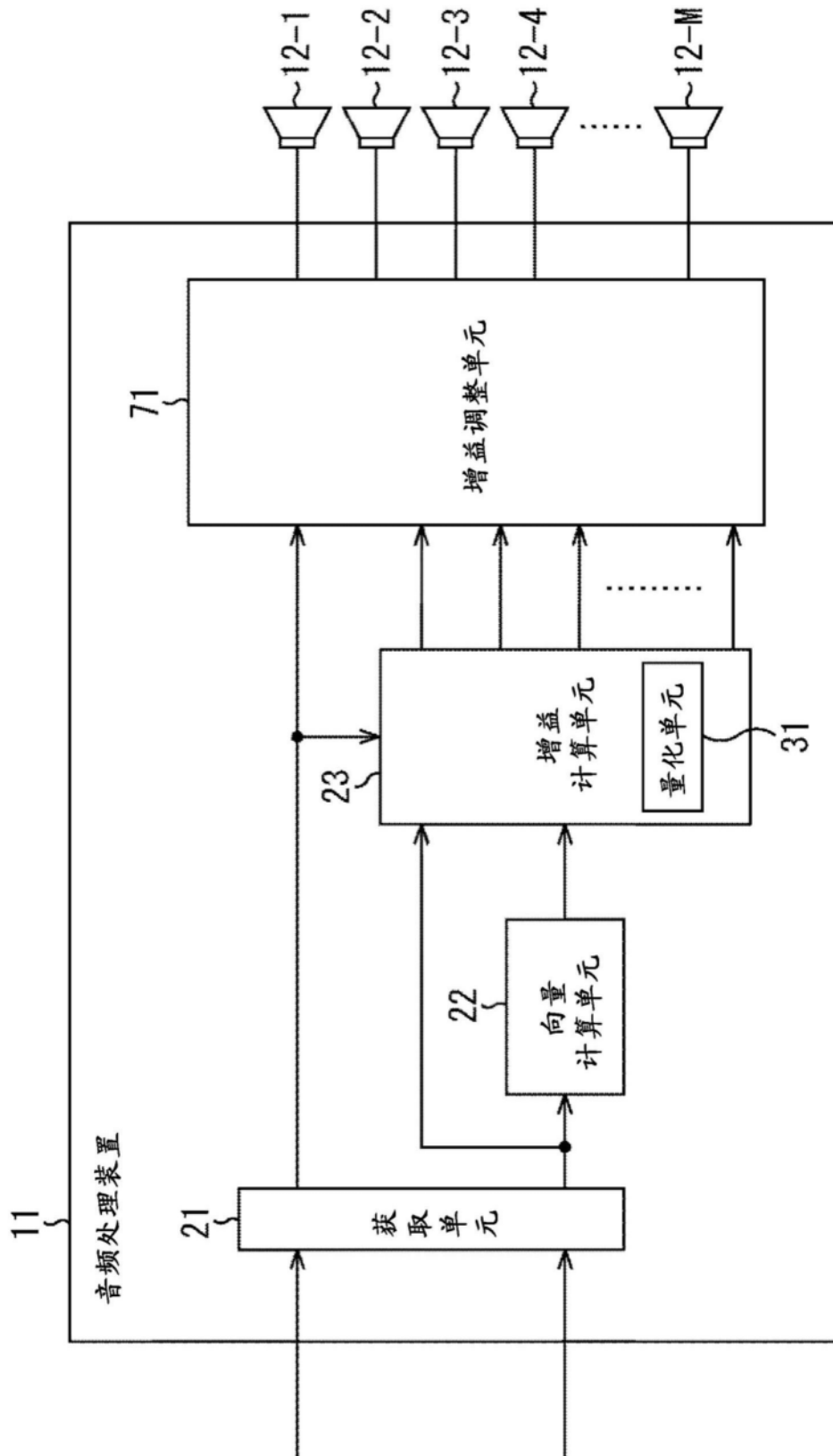


图19



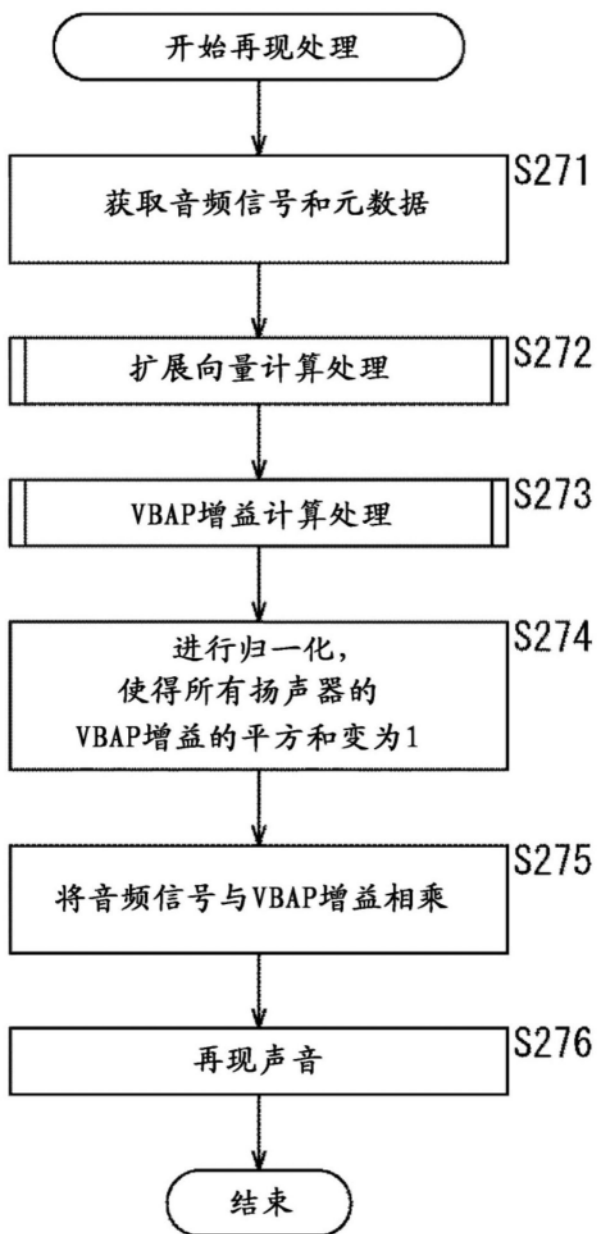


图20

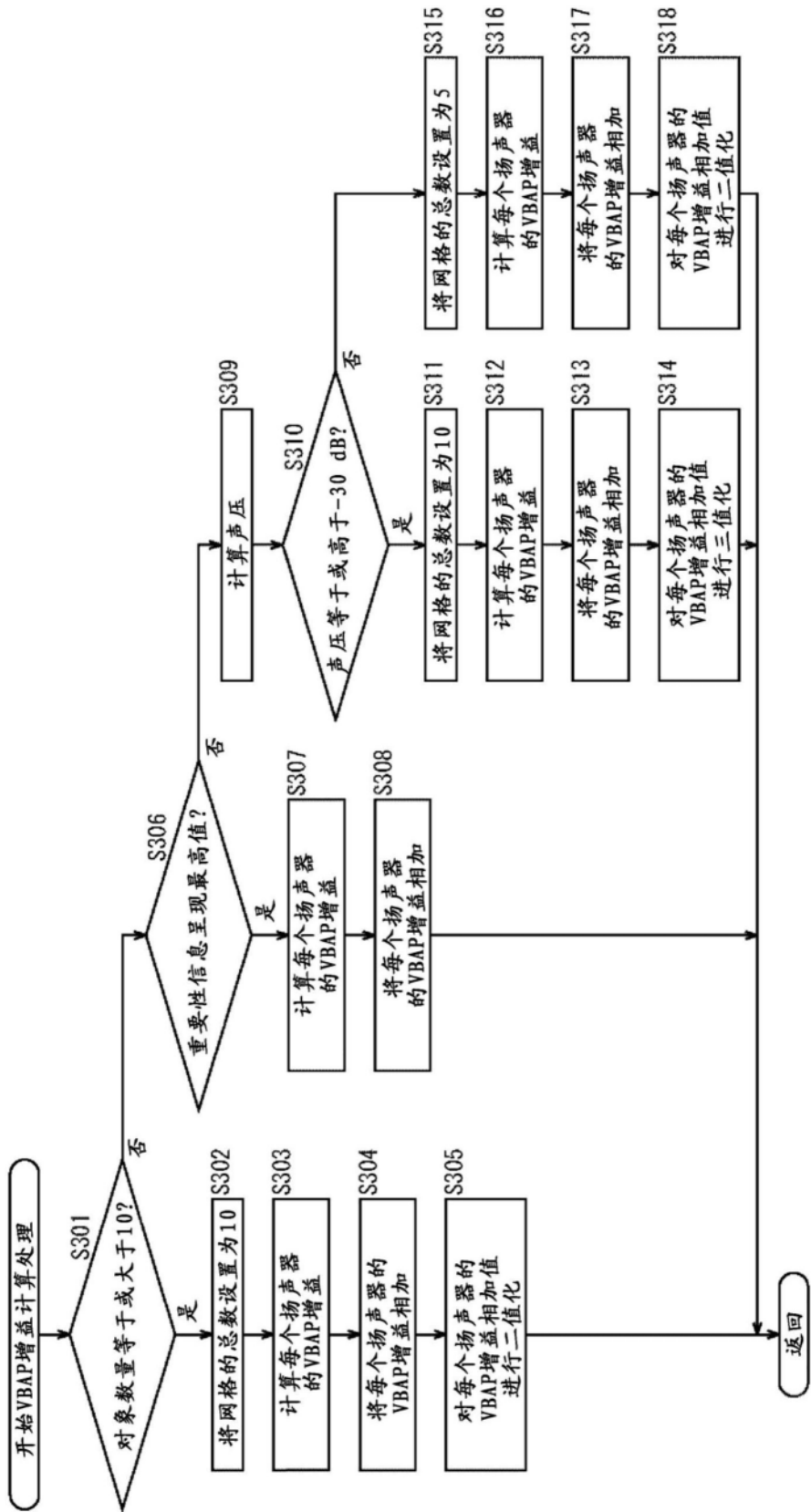


图21

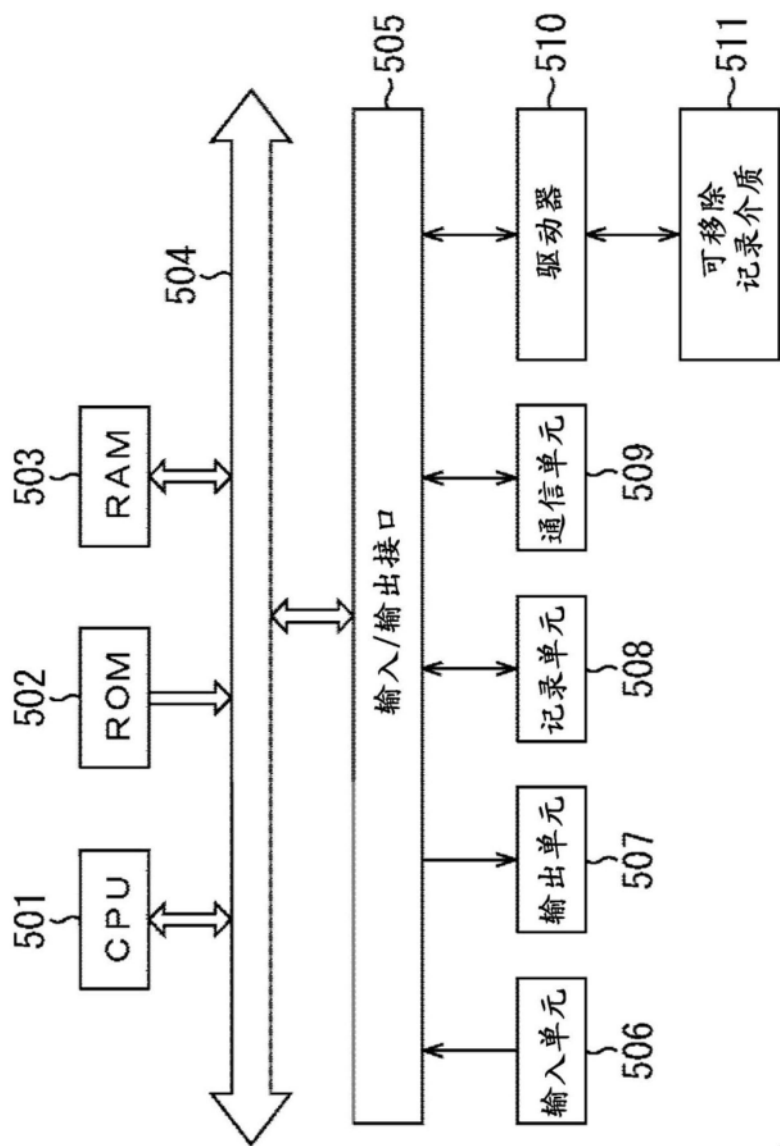


图22