



(12) 发明专利申请

(10) 申请公布号 CN 120019673 A

(43) 申请公布日 2025. 05. 16

(21) 申请号 202380071658.9

(22) 申请日 2023.09.28

(30) 优先权数据

63/417,398 2022.10.19 US

(85) PCT国际申请进入国家阶段日

2025.04.08

(86) PCT国际申请的申请数据

PCT/JP2023/035546 2023.09.28

(87) PCT国际申请的公布数据

W02024/084920 JA 2024.04.25

(71) 申请人 松下电器(美国)知识产权公司

地址 美国加利福尼亚

(72) 发明人 横本成悟 石川智一 宇佐见阳

中桥康太 江原宏幸 山田摩里子

宫阪修二

(74) 专利代理机构 永新专利商标代理有限公司

72002

专利代理师 安香子

(51) Int.Cl.

H04S 7/00 (2006.01)

H04R 3/00 (2006.01)

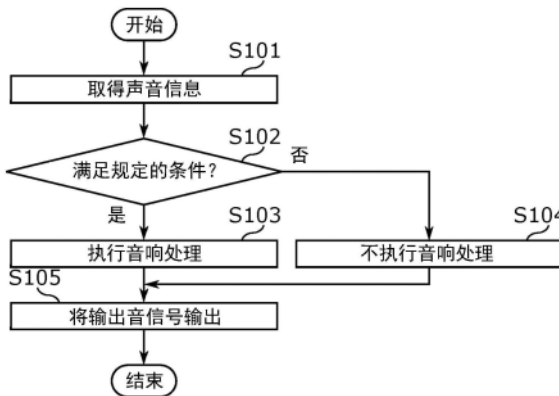
权利要求书2页 说明书27页 附图12页

(54) 发明名称

音响处理方法、音响处理装置以及程序

(57) 摘要

信息处理方法包括:取得使用收音装置收集从音源发出的声音而得到的声音信号的步骤(S101);对于上述声音信号,执行使上述收音装置与上述音源的相对位置在时域中反复变化的音响处理的步骤(S103);以及输出已执行上述音响处理的输出声音信号的步骤(S105)。



1. 一种音响处理方法,其中,包括:
取得使用收音装置收集从音源发出的声音而得到的声音信号的步骤;
对所述声音信号,执行使所述收音装置与所述音源的相对位置在时域中反复变化的音响处理的步骤;以及
输出已执行所述音响处理的输出声音信号的步骤。
2. 如权利要求1所述的音响处理方法,其中,
在执行所述音响处理的步骤中,
判定所述声音信号的声压在时域中的变化是否满足与该变化有关的规定的条件,
在判定为满足所述规定的条件的情况下执行所述音响处理,
在判定为不满足所述规定的条件的情况下不执行所述音响处理。
3. 如权利要求1所述的音响处理方法,其中,
在执行所述音响处理的步骤中,
使用所述声音信号推测所述收音装置与所述音源的位置关系,
判定推测出的所述位置关系是否满足与该位置关系有关的规定的条件,
在判定为满足所述规定的条件的情况下执行所述音响处理,
在判定为不满足所述规定的条件的情况下不执行所述音响处理。
4. 如权利要求1所述的音响处理方法,其中,
所述声音信号包含与收集声音时的状况有关的收音状况信息,
在执行所述音响处理的步骤中,
判定所述声音信号中包含的所述收音状况信息是否满足与该收音状况信息有关的规定的条件,
在判定为满足所述规定的条件的情况下执行所述音响处理,
在判定为不满足所述规定的条件的情况下不执行所述音响处理。
5. 如权利要求1所述的音响处理方法,其中,
在执行所述音响处理的步骤中,
使用所述声音信号推测所述收音装置与所述音源的位置关系,
在与推测出的所述位置关系相应的处理条件下执行所述音响处理。
6. 一种音响处理方法,用来输出使从虚拟声音空间内的音源对象发出的声音被感知为在所述虚拟声音空间内的收听点处收听到的输出声音信号,其中,包括:
取得包含从所述音源对象发出的声音的声音信号的步骤;
受理使所述收听点与所述音源对象的相对位置变化的指示的步骤,该指示包含通过该指示使所述相对位置变化的第1变化量;
对所述声音信号执行使所述相对位置变化所述第1变化量、并且使所述相对位置在时域中反复变化第2变化量的音响处理的步骤;以及
输出已执行所述音响处理的所述输出声音信号的步骤。
7. 如权利要求6所述的音响处理方法,其中,
所述音源对象模拟了实空间的用戶,
所述音响处理方法还包括从设置在所述实空间中的检测所述用戶的传感器取得检测结果的步骤,

所述第2变化量基于所述检测结果来计算。

8. 如权利要求6所述的音响处理方法,其中,
所述音源对象模拟了实空间的用戶,

所述音响处理方法还包括从设置在所述实空间中的检测所述用户的传感器取得检测结果的步骤,

所述第2变化量与所述检测结果独立地计算。

9. 如权利要求6所述的音响处理方法,其中,
所述第2变化量与所述第1变化量独立地计算。

10. 如权利要求6所述的音响处理方法,其中,
所述第1变化量越大,则所述第2变化量被计算为越大的数值。

11. 如权利要求6所述的音响处理方法,其中,
所述第1变化量越小,则所述第2变化量被计算为越大的数值。

12. 如权利要求1或6所述的音响处理方法,其中,
还包括取得对所述声音信号的控制信息的步骤,
在执行所述音响处理的步骤中,
在所述控制信息表示执行所述音响处理的情况下执行所述音响处理。

13. 一种音响处理装置,其中,具备:

取得部,取得使用收音装置收集从音源发出的声音而得到的声音信号;

处理部,对所述声音信号,执行使所述收音装置与所述音源的相对位置在时域中反复变化的音响处理;以及

输出部,输出已执行所述音响处理的输出声音信号。

14. 一种音响处理装置,用来输出使从虚拟声音空间内的音源对象发出的声音被感知为在所述虚拟声音空间内的收听点处收听到的输出声音信号,其中,具备:

取得部,取得包含从所述音源对象发出的声音的声音信号;

受理部,受理使所述收听点与所述音源对象的相对位置变化的指示,该指示包含通过该指示使所述相对位置变化的第1变化量;

处理部,对所述声音信号执行使所述相对位置变化所述第1变化量、并且使所述相对位置在时域中反复变化第2变化量的音响处理;以及

输出部,输出已执行所述音响处理的所述输出声音信号。

15. 一种程序,用来使计算机执行权利要求1或6所述的音响处理方法。

音响处理方法、音响处理装置以及程序

技术领域

[0001] 本公开涉及音响处理方法、音响处理装置及程序。

背景技术

[0002] 以往,已知有用来在虚拟的三维空间内使用户感知立体的声音的与音响再现有关的技术(例如,参照专利文献1)。此外,为了使得在这样的三维空间内如从音源对象到达用户那样感知声音,需要进行根据原来的声音信息生成输出音信息的处理。这里,为了使收听声音的用户更好地感受三维空间内的临场感,会进行使声音的定位感增加的音响处理。例如,已知有带来定位感以使得从自坐标波动附加装置输入的音源坐标的方向听到声音的立体音响处理装置(参照专利文献1)。

[0003] 现有技术文献

[0004] 专利文献

[0005] 专利文献1:日本特开2005—295416号公报

发明内容

[0006] 发明要解决的课题

[0007] 在为了增加声音的定位感而赋予波动的情况下,有时不能适当地执行赋予该波动的音响处理。所以,在本公开中,对用来更适当地执行音响处理的音响处理方法等进行说明。

[0008] 用来解决课题的手段

[0009] 有关本公开的一技术方案音响处理方法包括:取得使用收音装置收集从音源发出的声音而得到的声音信号;对所述声音信号,执行使所述收音装置与所述音源的相对位置在时域中反复变化的音响处理的步骤;以及输出已执行所述音响处理的输出声音信号的步骤。

[0010] 此外,有关本公开的另一技术方案音响处理方法,是用来输出使从虚拟声音空间内的音源对象发出的声音感知为如在所述虚拟声音空间内的收听点收听到的输出声音信号的音响处理方法,包括:取得包含从所述音源对象发出的声音的声音信号的步骤;受理使所述收听点与所述音源对象的相对位置变化的指示的步骤,该指示包含通过该指示使所述相对位置变化的第1变化量;对所述声音信号执行使所述相对位置变化所述第1变化量、并且使所述相对位置在时域中反复变化第2变化量的音响处理的步骤;以及输出已执行所述音响处理的所述输出声音信号的步骤。

[0011] 此外,有关本公开的一技术方案音响处理装置具备:取得部,取得使用收音装置收集从音源发出的声音而得到的声音信号;处理部,对所述声音信号,执行使所述收音装置与所述音源的相对位置在时域中反复变化的音响处理;以及输出部,输出已执行所述音响处理的输出声音信号。

[0012] 此外,有关本公开的另一技术方案音响处理装置,是用来输出使从虚拟声音空

间内的音源对象发出的声音感知为如在所述虚拟声音空间内的收听点收听到的输出声音信号的音响处理装置,包括:取得部,取得包含从所述音源对象发出的声音的声音信号;受理部,受理使所述收听点与所述音源对象的相对位置变化的指示,该指示包含通过该指示使所述相对位置变化的第1变化量;处理部,对所述声音信号执行使所述相对位置变化所述第1变化量、并且使所述相对位置在时域中反复变化第2变化量的音响处理;以及输出部,输出已执行所述音响处理的所述输出声音信号。

[0013] 此外,本公开的一技术方案还能够作为用于使计算机执行如上所述的音响处理方法的程序实现。

[0014] 另外,这些包含性或具体的技术方案也可以由系统、装置、方法、集成电路、计算机程序或计算机可读的CD-ROM等的非暂时性的记录介质实现,也可以由系统、装置、方法、集成电路、计算机程序及记录介质的任意的组合来实现。

[0015] 发明效果

[0016] 根据本公开,能够更适当地执行音响处理。

附图说明

[0017] 图1是表示有关实施方式的音响再现系统的使用事例的概略图。

[0018] 图2A是用来说明有关实施方式的音响再现系统的使用事例的图。

[0019] 图2B是用来说明有关实施方式的音响再现系统的使用事例的图。

[0020] 图3是表示有关实施方式的音响再现系统的功能结构的框图。

[0021] 图4是表示有关实施方式的取得部的功能结构的框图。

[0022] 图5是表示有关实施方式的处理部的功能结构的框图。

[0023] 图6是用来说明有关实施方式的音响再现系统的另一例的图。

[0024] 图7是用来说明有关实施方式的音响再现系统的另一例的图。

[0025] 图8是用来说明有关实施方式的音响再现系统的另一例的图。

[0026] 图9是用来说明有关实施方式的音响再现系统的另一例的图。

[0027] 图10是用来说明有关实施方式的音响再现系统的另一例的图。

[0028] 图11是用来说明有关实施方式的音响再现系统的另一例的图。

[0029] 图12是用来说明有关实施方式的音响再现系统的另一例的图。

[0030] 图13是用来说明有关实施方式的音响再现系统的另一例的图。

[0031] 图14是用来说明有关实施方式的音响再现系统的另一例的图。

[0032] 图15是用来说明有关实施方式的音响再现系统的另一例的图。

[0033] 图16是表示有关实施方式的音响处理装置的动作的流程图。

[0034] 图17是用来说明有关实施方式的音响处理的频率特性的图。

[0035] 图18是用来说明有关实施方式的音响处理的波动的大小的图。

[0036] 图19是用来说明有关实施方式的,音响处理的波动的周期及角度的图。

[0037] 图20是表示有关实施方式的另一例的处理部的功能结构的框图。

[0038] 图21是表示有关实施方式的另一例的音响处理装置的动作的流程图。

具体实施方式

[0039] (作为公开的基础的认识)

[0040] 以往,已知有用在虚拟的三维空间内(以下,有时称作三维声场或虚拟声音空间)使用户感知立体的声音的与音响再现有关的技术(例如,参照专利文献1)。通过使用该技术,用户能够像音源对象存在于虚拟空间内的规定位置,并且声音从其方向到来那样感知该声音。为了像这样使声像定位在虚拟的三维空间内的规定位置,例如对于音源对象的声音的信号,需要进行产生被感知为立体的声音那样的两耳间的声音的到来时间差及两耳间的声音的等级差(或声压差)等计算处理。这样的计算处理通过应用立体音响滤波器来进行。立体音响滤波器是如果对原来的声音信息应用该滤波器后的输出音信号被再现,则会有立体感地感知声音的方向、距离等的位置、音源的大小、空间的大小等的信息处理用滤波器。

[0041] 作为这样的立体音响滤波器的应用的计算处理的一例,已知有对目标音信号卷积用来使其感知为从规定方向到来的声音的头相关传递函数的处理。通过对于从音源对象的位置到用户位置的声音的到来方向,以充分小的角度实施该头相关传递函数的卷积的处理,用户感受到的临场感提高。

[0042] 此外,近年来积极地进行与虚拟现实(VR:Virtual Reality)有关的技术的开发。在虚拟现实中,由于三维声场中的声音的定位感也带来影像的临场感,所以会执行用来增加定位感的音响处理。在为了增加声音的定位感而赋予波动的情况下,从该效果的观点来看,不需要对全部的声音一样地赋予波动。换言之,存在波动的赋予有效地发挥作用的条件的。通过仅在满足这样的条件的情况下进行波动的赋予,不再需要不必要地准备处理资源,所以可以说是优选的。

[0043] 更具体的本公开的概要如下。

[0044] 有关本公开的第1技术方案音响处理方法包括:取得使用收音装置收集从音源发出的声音而得到的声音信号的步骤;对声音信号,执行使收音装置与音源的相对位置在时域中反复变化的音响处理的步骤;以及输出已执行音响处理的输出声音信号的步骤。

[0045] 根据这样的音响处理方法,在如使用收音装置收集到的声音信号那样,是收音装置的配置位置相对于音源的位置不相对变化等的丧失临场感的条件的情况下,通过音响处理使收音装置与音源的相对位置在时域中反复变化而赋予波动,从而能够重建丧失的临场感。这样,从重建临场感的观点来看能够更适当地执行音响处理。

[0046] 此外,有关第2技术方案音响处理方法在第1技术方案所述的音响处理方法中,在执行音响处理的步骤中,判定声音信号的声压在时域中的变化是否满足与该变化有关的规定条件;在判定为满足规定的条件的情况下执行音响处理;在判定为不满足规定的条件的情况下不执行音响处理。

[0047] 根据这样的音响处理方法,能够通过是否满足与声音信号的声压在时域中的变化有关的规定条件的判定,来改变是否执行音响处理。

[0048] 此外,有关第3技术方案音响处理方法在第1或第2技术方案所述的音响处理方法中,在执行音响处理的步骤中,使用声音信号推测收音装置与音源的位置关系;判定推测出的位置关系是否满足与该位置关系有关的规定条件;在判定为满足规定的条件的情况下执行音响处理;在判定为不满足规定的条件的情况下不执行音响处理。

[0049] 根据这样的音响处理方法,能够通过是否满足与使用声音信号推测出的收音装置和音源的位置关系有关的规定的条件的判定,来改变是否执行音响处理。

[0050] 此外,有关第4技术方案的音响处理方法在第1~第3技术方案的任一个技术方案所述的音响处理方法中,声音信号包含与收集声音时的状况有关的收音状况信息;在执行音响处理的步骤中,判定声音信号中包含的收音状况信息是否满足与该收音状况信息有关的规定的条件;在判定为满足规定的条件的情况下执行音响处理;在判定为不满足规定的条件的情况下不执行音响处理。

[0051] 根据这样的音响处理方法,能够通过与声音信号中包含的收音状况信息有关的规定的条件的判定,来改变是否执行音响处理。

[0052] 此外,有关第5技术方案的音响处理方法在第1~第4技术方案的任一个技术方案所述的音响处理方法中,在执行音响处理的步骤中,使用声音信号推测收音装置与音源的位置关系;在与推测出的位置关系相应的处理条件下执行音响处理。

[0053] 根据这样的音响处理方法,能够在与使用声音信号推测出的收音装置和音源的位置关系相应的处理条件下执行音响处理。

[0054] 此外,有关本公开的第6技术方案的音响处理方法,是用来输出使从虚拟声音空间内的音源对象发出的声音被感知为在虚拟声音空间内的收听点收听到的输出声音信号的音响处理方法,包括:取得包含从音源对象发出的声音的声音信号的步骤;受理使收听点与音源对象的相对位置变化的指示的步骤,该指示包含通过该指示使相对位置变化的第1变化量;对声音信号执行使相对位置变化第1变化量、并且使相对位置在时域中反复变化第2变化量的音响处理的步骤;以及输出已执行音响处理的输出声音信号的步骤。

[0055] 根据这样的音响处理方法,当使从虚拟声音空间内的音源对象发出的声音被感知为在虚拟声音空间内的收听点收听到时,与基于使收听点和音源对象的相对位置变化的指示产生的第1变化量的相对位置的变化独立地,在声音信号中已经丧失了临场感的情况下,通过由音响处理使收听点和音源对象的相对位置在时域中以第2变化量反复变化地赋予波动,能够重建丧失的临场感。这样,从重建临场感的观点来看能够更适当地执行音响处理。

[0056] 此外,有关第7技术方案的音响处理方法在第6技术方案所述的音响处理方法中,音源对象模拟实空间的用户;音响处理方法还包括从设置在实空间中的检测用户的传感器取得检测结果的步骤;第2变化量基于检测结果来计算。

[0057] 根据这样的音响处理方法,作为第2变化量,能够根据从检测与音源对象对应的实空间的用户的传感器取得的检测结果来计算第2变化量。

[0058] 此外,有关第8技术方案的音响处理方法在第6技术方案所述的音响处理方法中,音源对象模拟实空间的用户;音响处理方法还包括从设置在实空间中的检测用户的传感器取得检测结果的步骤;第2变化量与检测结果独立地计算。

[0059] 根据这样的音响处理方法,能够与从检测与音源对象对应的实空间的用户的传感器取得的检测结果独立地计算第2变化量。

[0060] 此外,有关第9技术方案的音响处理方法在第1技术方案所述的音响处理方法中,第2变化量与第1变化量独立地计算。

[0061] 根据这样的音响处理方法,能够计算与第1变化量独立的第2变化量。

[0062] 此外,有关第10技术方案的音响处理方法在第6技术方案所述的音响处理方法中,

第1变化量越大,则第2变化量被计算为越大的数值。

[0063] 根据这样的音响处理方法,能够计算出第1变化量越大则越大的第2变化量。

[0064] 此外,有关第11技术方案的音响处理方法在第6技术方案所述的音响处理方法中,第1变化量越小,第2变化量被计算为越大的数值。

[0065] 根据这样的音响处理方法,能够计算出第1变化量越小则越大的第2变化量。

[0066] 此外,有关第12技术方案的音响处理方法在第1~第11技术方案的任一个技术方案所述的音响处理方法中,还包括取得对声音信号的控制信息的步骤;在执行音响处理的步骤中,在控制信息表示执行音响处理的情况下执行音响处理。

[0067] 根据这样的音响处理方法,能够在所取得的控制信息表示执行音响处理的情况下执行音响处理。

[0068] 此外,有关本公开的第13技术方案的音响处理装置具备:取得部,取得使用收音装置收集从音源发出的声音而得到的声音信号;处理部,对声音信号,执行使收音装置与音源的相对位置在时域中反复变化的音响处理;以及输出部,输出已执行音响处理的输出声音信号。

[0069] 根据这样的音响处理装置,能够起到与上述所记载的音响处理方法同样的效果。

[0070] 此外,有关本公开的第14技术方案的音响处理装置,是用来输出使从虚拟声音空间内的音源对象发出的声音被感知为在虚拟声音空间内的收听点处收听到的输出声音信号的音响处理装置,包括:取得部,取得包含从音源对象发出的声音的声音信号;受理部,受理使收听点与音源对象的相对位置变化的指示,该指示包含通过该指示使相对位置变化的第1变化量;处理部,对声音信号执行使相对位置变化第1变化量、并且使相对位置在时域中反复变化第2变化量的音响处理;以及输出部,输出已执行音响处理的输出声音信号。

[0071] 根据这样的音响处理装置,能够起到与上述所记载的音响处理方法同样的效果。

[0072] 进而,这些包含性或具体的技术方案也可以由系统、装置、方法、集成电路、计算机程序或计算机可读的CD-ROM等的非暂时性的记录介质实现,也可以由系统、装置、方法、集成电路、计算机程序及记录介质的任意的组合来实现。

[0073] 以下,参照附图对实施方式具体地进行说明。另外,以下说明的实施方式都表示包含性或具体的例子。在以下的实施方式中表示的数值、形状、材料、构成要素、构成要素的配置位置及连接形态、步骤、步骤的顺序等是一例,不是限定本公开的意思。此外,关于以下的实施方式的构成要素中的在独立权利要求中没有记载的构成要素,设为任意的构成要素进行说明。另外,各图是示意图,并不一定是严密地图示的。此外,在各图中,对于实质上相同的结构赋予相同的标号,有将重复的说明省略或简略化的情况。

[0074] 此外,在以下的说明中,有时对要素赋予第1、第2及第3等序数。这些序数是为了识别要素而对要素赋予的,并不一定对应于有意义的顺序。这些序数也可以适当替换,也可以新赋予,也可以去除。

[0075] (实施方式)

[0076] [概要]

[0077] 首先,对有关实施方式的音响再现系统的概要进行说明。图1是表示有关实施方式的音响再现系统的使用事例的概略图。在图1中,表示使用音响再现系统100的用户99。

[0078] 图1所示的音响再现系统100与立体影像再现装置200同时被使用。通过同时视听

立体的图像及立体的声音,图像增强听觉上的临场感,并且声音增强视觉上的临场感,能够体会到就像处于拍摄图像及声音的现场一样。例如,已知在显示有人进行讲话的图像(运动图像)的情况下,即使在讲话声音的声像的定位与该人的嘴边偏离的情况下,用户99也会感知为从该人的口发出的讲话声音。这样,通过视觉信息来修正声像的位置等,通过配合图像和声音会提高临场感。

[0079] 立体影像再现装置200是佩戴在用户99的头部上的图像显示设备。因而,立体影像再现装置200与用户99的头部一体地移动。例如,如图所示,立体影像再现装置200是由用户99的耳和鼻支承的眼镜型的设备。

[0080] 立体影像再现装置200通过对应于用户99的头部的运动使显示的图像变化,使用户99感知为其在三维图像空间内运动头部。即,当三维图像空间内的物体位于用户99的正面时,如果用户99朝向右方则该物体向用户99的左方移动,如果用户99朝向左方则该物体向用户99的右方移动。这样,立体影像再现装置200相对于用户99的运动,使三维图像空间向与用户99的运动相反的方向移动。

[0081] 立体影像再现装置200分别显示在用户99的左右眼分别产生了相应于视差的偏差的两个图像。用户99能够基于所显示的图像的相应于视差的偏差来感知图像上的物体的三维的位置。另外,在将音响再现系统100用于睡眠引导用的治愈音的再现等用户99闭上眼睛而使用的情况下等,不需要同时使用立体影像再现装置200。即,立体影像再现装置200不是本公开的必须的构成要素。作为立体影像再现装置200,除了专用的影像显示设备以外,也有使用用户99拥有的智能电话、平板电脑装置等通用的便携终端的情况。

[0082] 在这样的通用的便携终端上,除了用来显示影像的显示器以外,还搭载有用来检测终端的姿势、运动的各种传感器。进而,还搭载信息处理用的处理器,能够与网络连接而与云服务器等服务器装置进行信息的收发。即,也能够通过智能电话与没有信息处理功能的通用的耳机等的组合来实现立体影像再现装置200及音响再现系统100。

[0083] 也可以如该例那样,将检测头部的运动的功能、影像的提示功能、提示用的影像信息处理功能、声音提示功能及提示用的声音信息处理功能适当地配置到1个以上的装置中而实现立体影像再现装置200及音响再现系统100。在不需要立体影像再现装置200的情况下,只要能够将检测头部的运动的功能、声音提示功能及提示用的声音信息处理功能适当地配置到1个以上的装置中即可,例如,也可以通过具有提示用的声音信息处理功能的计算机或智能电话等的处理装置、具有检测头部的运动的功能及声音提示功能的耳机等来实现音响再现系统100。

[0084] 音响再现系统100是佩戴在用户99的头部上的声音提示设备。因而,音响再现系统100与用户99的头部一体地移动。例如,本实施方式的音响再现系统100是所谓的头戴耳机型的设备。另外,对于音响再现系统100的形态没有特别限定,例如也可以是分别独立地佩戴在用户99的左右耳的两个耳塞型的设备。

[0085] 音响再现系统100通过对应于用户99的头部的运动使提示的声音变化,使用户99感知为如用户99在三维声场内运动头部那样。因此,如上述那样,音响再现系统100相对于用户99的运动,使三维声场向与用户99的运动相反的方向移动。

[0086] 这里,以提高用户99收听的声音的临场感的目的,会执行对声音赋予波动的音响处理。例如,图2A及图2B是用来说明有关实施方式的音响再现系统的使用事例的图。在图2A

中表示了正在进行所谓的视频通话的用户。在图2A的左图中,声音在如耳机那样嘴(音源)与耳机的麦克风(收音装置)的位置几乎不变的条件下被收集。但是,在右图的通话对方处,相对于在影像上运动的用户,音源和收音装置的位置几乎不移动,会产生由此带来的违和感。在这样的情况下,通过应用与在影像上运动的用户运动匹配的声音的波动、或者与会话中的用户的一般的运动匹配的声音的波动,会降低声音的违和感,增加临场感。

[0087] 此外,在图2B中表示了正在将所谓的虚拟直播用的歌声进行立体收音的用户。正在收集声音的用户也可以是与作为收听者的用户99不同的用户。例如,可以想到歌手、艺术家等。在图2B的左图中,通过用户朝向固定的麦克风演唱来收集歌声。使用收集到的声音,在右图的虚拟影像上再现声音,与在虚拟空间内的直播会场中模拟了用户的化身一边跳舞一边唱歌的影像一起视听,从而实现虚拟直播。此时,如果追随于化身的运动,指定虚拟声音空间内的音源对象(化身的头部)的位置作为声音的再现位置,则即使位置匹配,也不会再现实际的用户所应该有的波动的微小的运动,导致声音的临场感减少。在本公开中,通过像这样对声音赋予本来应该有的波动,进行增加声音的临场感的音响处理。此外,作为产生同样的课题的其他状况,有以下的情况:在如图2A的视频通话中即使使用能够将用户的波动也包含在内收集声音的收音装置,也会为了使得成为容易由听者听清的声音而应用被称作AGC(自动音量调整)的机械性声音处理等,在声音中波动被抑制,反而产生违和感。在本公开中还包括通过再次赋予通过这样的机械性声音处理而被抑制的波动来降低声音的违和感,增加临场感的情况。

[0088] 另一方面,波动的赋予是通过对输出的输出音信号进行滤波处理,使声音在时域中反复移动来进行的。该处理由于需要在时域的连续的两个时间点分别应用不同的滤波器,所以较复杂,在预计难以获得波动的效果的情况下优选不应用音响处理。

[0089] [构成]

[0090] 接着,参照图3对有关本实施方式的音响再现系统100的结构进行说明。图3是表示有关实施方式的音响再现系统的功能结构的框图。

[0091] 如图3所示,有关本实施方式的音响再现系统100具备信息处理装置101、通信模块102、检测器103和驱动器104。

[0092] 信息处理装置101是音响处理装置的一例,是用来进行音响再现系统100中的各种信号处理的运算装置。信息处理装置101例如也可以由计算机等的具备处理器和存储器、以由处理器执行存储在存储器中的程序的形式实现。通过该程序的执行,发挥以下说明的与各功能部有关的功能。

[0093] 信息处理装置101具有取得部111、处理部121及信号输出部141。以下,对于信息处理装置101具有的各功能部的详细情况,与信息处理装置101以外的结构的详细情况一起进行说明。

[0094] 通信模块102是用来受理向音响再现系统100的声音信息的输入的接口装置。通信模块102例如具备天线和信号变换器,通过无线通信从外部的装置接收声音信息。更详细地讲,通信模块102使用天线接收表示被变换为用于无线通信的形式后的声音信息的无线信号,通过信号变换器进行从无线信号向声音信息的再变换。由此,音响再现系统100通过无线通信从外部的装置取得声音信息。由通信模块102取得的聲音信息被取得部111取得。这样,声音信息被输入到信息处理装置101。另外,音响再现系统100与外部的装置的通信也可

以通过有线通信进行。

[0095] 音响再现系统100取得的声音信息是使用收音装置收集从音源发出的声音而得到的声音信号。声音信息例如也可以以MPEG—H 3D Audio (ISO/IEC 23008—3)、MPEG—I等规定的形式被进行了编码。作为一例,编码后的声音信息中包含关于由音响再现系统100再现的规定音的信息、以及与使该声音的声像在三维声场内定位在规定位置(即感知为从规定方向到来的声音)时的定位位置有关的信息和其他元数据。例如,声音信息中包含与包括第1规定音及第2规定音的多个声音有关的信息,使声像定位,以使得将再现各个声音时的声像感知为从三维声场内的不同的位置到来的声音。

[0096] 通过该立体的声音,例如能够与使用立体影像再现装置200辨识的图像一起提高视听的内容等的临场感。另外,声音信息中,也可以仅包含关于规定音的信息。在此情况下,也可以另行取得与规定位置有关的信息。此外,如上述那样,声音信息包含与第1规定音有关的第1声音信息及与第2规定音有关的第2声音信息,但也可以分别取得单独地包含它们的多个声音信息,通过同时再现来使声像定位在三维声场内的不同的位置。这样,对输入的声音信息的形态没有特别限定,只要音响再现系统100具备与各种形态的声音信息对应的取得部111即可。

[0097] 声音信息中包含的元数据包含用来控制用于赋予波动的音响处理的控制信息。控制信息是用来指定是否执行音响处理的信息。例如,在控制信息中指定了执行音响处理的情况下,进而可以判定是否满足规定的条件,在满足规定的条件的情况下执行音响处理,也可以与是否满足规定的条件的判定无关地执行音响处理。另一方面,在控制信息中指定了不执行音响处理的情况下,不执行音响处理。这样,既可以通过是否满足规定的条件的判定和是否在控制信息中指定了执行音响处理的两个契机来执行音响处理,也可以通过是否指定了执行音响处理的1个契机来执行音响处理。控制信息也可以不包含在元数据中。例如,也可以是,控制信息能够通过音响再现系统100的动作设定来指定,存储在存储部中。并且,控制信息也可以在音响再现系统100的启动时取得并如上述那样使用。

[0098] 此外,在元数据中也可以包含收音状况信息。收音状况信息是声音信息中包含的与规定音的收集有关的混响等级及噪声等级。关于收音状况信息的详细情况在后面叙述。

[0099] 声音信息也可以作为比特流取得。对将声音信息作为比特流取得的情况下的比特流的构造例进行说明。在比特流中,例如包含声音信号和元数据。声音信号是表示与声音的频率及强弱有关的信息等的、表现了声音的声音数据。元数据也可以包含上述的信息以外的空间信息。空间信息是与听取基于声音信号的声音的听者所处的空间有关的信息。具体而言,空间信息是与使该声音的声像定位在声音空间(例如三维声场内)中的规定的位置时的、即被听者感知为从规定方向到达的声音时的该规定的位置(定位位置)有关的信息。在空间信息中,例如包括音源对象信息和表示听者的位置的位置信息。

[0100] 音源对象信息是产生基于声音信号的声音的、即表示再现声音信号的物体的对象的信息,并且是与配置在对应于配置该物体的实空间的虚拟的空间即声音空间中的虚拟的对象(音源对象)有关的信息。音源对象信息例如包含表示配置在声音空间中的音源对象的位置的信息、与音源对象的朝向有关的信息、与音源对象发出的声音的指向性有关的信息、表示音源对象是否属于生物的信息及表示音源对象是否是运动体的信息等。例如,声音信号与音源对象信息所表示的1个以上的音源对象对应。

[0101] 作为比特流的数据构造的一例,比特流例如由元数据(控制信息)和声音信号构成。

[0102] 声音信号和元数据既可以保存在一个比特流中,也可以分别保存在多个比特流中。同样,声音信号和元数据既可以保存在一个文件中,也可以分别保存在多个文件中。

[0103] 比特流既可以按每个音源存在,也可以按每个再现时间存在。在按每个再现时间存在比特流的情况下,也可以同时将多个比特流并行地处理。

[0104] 元数据既可以按每个比特流赋予,也可以作为对多个比特流进行控制的信息一起赋予。此外,元数据也可以按每个再现时间赋予。

[0105] 声音信号和元数据在分别保存在多个比特流或多个文件中的情况下,可以包含表示与一个或部分比特流或文件关联的其他的比特流或文件的信息,也可以包含表示与全部比特流或文件分别关联的其他的比特流或文件的信息。这里,关联的比特流或文件例如是在音响处理时有可能被同时使用的比特流或文件。此外,在关联的比特流或文件中也可以包括将表示关联的其他的比特流或文件的信息一起记述的比特流或文件。

[0106] 这里,表示关联的其他的比特流或文件的信息,例如是表示该其他的比特流的识别符、表示其他文件的文件名、URL (Uniform Resource Locator: 统一资源定位符) 或URI (Uniform Resource Identifier: 统一资源标识符) 等。在此情况下,取得部111基于表示关联的其他的比特流或文件的信息,确定或取得比特流或文件。此外,也可以在比特流内包含有表示关联的其他比特流的信息,并且在比特流内包含表示与其他的比特流或文件关联的比特流或文件的信息。并且,包含表示关联的比特流或文件的信息的文件例如也可以是被用于内容的分发的清单文件等的控制文件。

[0107] 另外,全部的元数据或部分元数据也可以从声音信号的比特流以外取得。例如,既可以将对音响进行控制的元数据和对影像进行控制的元数据的某一方从比特流以外取得,也可以将两者的元数据都从比特流以外取得。此外,在对影像进行控制的元数据包含在由声音信号再现系统(对应于音响再现系统100)取得的比特流中的情况下,声音信号再现系统也可以具备将能够用于影像的控制的元数据输出给显示图像的显示装置或再现立体影像的立体影像再现装置(例如,实施方式中的立体影像再现装置200)的功能。

[0108] 进而,对元数据中包含的信息的例子进行说明。

[0109] 元数据也可以是为了记述由声音空间表现的场景而使用的信息。这里,场景是指表示使用元数据由声音信号再现系统建模的声音空间中的三维影像及音响事件的全部要素的集合体的用语。即,这里所述的元数据,不仅包括对音响处理进行控制的信息,也可以包括对影像处理进行控制的信息。当然,元数据中既可以包括仅对音响处理和影像处理的某一方进行控制的信息,也可以包括在两者的控制中使用的信息。

[0110] 声音信号再现系统通过使用比特流中包含的元数据以及追加取得的交互性的听者的位置信息等对声音信号进行音响处理,生成虚拟的音响效果。在本实施方式中说明音响效果中进行初期反射处理、障碍物处理、衍射处理、遮断处理和混响处理的情况,但也可以使用元数据进行其他的音响处理。例如,可以考虑声音信号再现系统附加距离衰减效果、定位、多普勒效应等的音响效果。此外,也可以将对音响效果的全部或一部分的开启关闭进行切换的信息、优先级信息作为元数据来附加。

[0111] 此外,作为一例,被编码的元数据包括:与包含音源对象及障碍物对象的声音空间

有关的信息、以及与使该声音的声像在声音空间内定位在规定位置(即,使得被感知为从规定方向到达的声音)时的定位位置有关的信息。这里,障碍物对象是在音源对象发出的声音到达听者为止的期间例如将声音遮挡或将声音反射等而可能给听者感知的声音带来影响的对象。障碍物对象除了静止物体以外,还可能包括人等的动物或机械等的运动体。此外,在声音空间中存在多个音源对象的情况下,对于任意的音源对象而言,其他的音源对象可能成为障碍物对象。建材或无生物等作为不发出声音的对象的非发声对象及发出声音的音源对象也都可能成为障碍物对象。

[0112] 在元数据中,包含分别表示声音空间的形状、存在于声音空间中的障碍物对象的形状信息及位置信息、存在于声音空间中的音源对象的形状信息及位置信息、以及声音空间中的听者的位置及朝向的信息的全部或一部分。

[0113] 声音空间是封闭空间或开放空间的哪种都可以。此外,元数据中包含表示例如地板、墙壁或天花板等的在声音空间中能够将声音反射的构造物的反射率、以及存在于声音空间中的障碍物对象的反射率的信息。这里,反射率是反射音与入射音的能量比,按声音的每个频带设定。当然,反射率也可以不依赖于声音的频带而一律地设定。在声音空间为开放空间的情况下,例如也可以使用一律地设定的衰减率、衍射音或初期反射音等的参数。

[0114] 在上述说明中,列出了反射率,但作为元数据中包含的与障碍物对象或音源对象有关的参数也可以包含反射率以外的信息。例如,反射率以外的信息也可以包含与对象的材料有关的信息作为与音源对象及非发声对象双方有关的元数据。具体而言,反射率以外的信息也可以包含扩散率、透射率、吸音率等参数。

[0115] 作为与音源对象有关的信息,也可以包含音量、放射特性(指向性)、再现条件、从一个对象发出的音源的数量及种类、以及指定对象中的音源区域的信息等。再现条件中例如也可以设定是持续地流出的声音还是事件触发的声音。对象中的音源区域既可以通过听者的位置与对象的位置的相对关系来设定,也可以以对象为基准来设定。在对象中的音源区域通过听者的位置与对象的位置的相对关系来设定的情况下,以听者看的对象的面为基准,能够使听者感知为,从听者看时从对象的右侧发出声音A、从左侧发出声音B。在对象中的音源区域以对象为基准来设定的情况下,不论听者看的方向如何都可以固定从对象的哪个区域发出哪个声音。例如,可以使听者感知为,当从正面看对象时从右侧流出高音,从左侧流出低音。在此情况下,在听者绕到对象的背面的情况下,能够使听者感知为,从背面看时从右侧流出低音、从左侧流出高音。

[0116] 作为与空间有关的元数据,可以包括到初期反射音为止的时间、混响时间、直接音与扩散音的比率等。在直接音与扩散音的比率为零的情况下,能够使听者仅感知直接音。

[0117] 使用图4说明取得部111的一例。图4是表示有关实施方式的取得部的功能结构的框图。如图4所示,本实施方式的取得部111例如具备编码声音信息输入部112、解码处理部113及感测信息输入部114。

[0118] 编码声音信息输入部112是被输入取得部111所取得的被编码的(换言之进行了encode)声音信息的处理部。编码声音信息输入部112将被输入的声音信息输出给解码处理部113。

[0119] 解码处理部113是通过将从编码声音信息输入部112输出的声音信息解码(换言之decode),以用于以后的处理的形式生成与声音信息中包含的规定音有关的信息和与规定

位置有关的信息的处理部。

[0120] 关于感测信息输入部114,以下与检测器103的功能一起进行说明。

[0121] 检测器103是用来检测用户99的头部的运动速度的装置。检测器103是将陀螺仪传感器、加速度传感器等用于运动的检测的各种传感器组合而构成的。在本实施方式中,检测器103内置在音响再现系统100中,但例如也可以内置在与音响再现系统100同样对应于用户99的头部的运动而动作的立体影像再现装置200等外部的装置中。在此情况下,检测器103也可以不包含在音响再现系统100中。此外,作为检测器103,也可以使用外部的摄像装置等拍摄用户99的头部的运动,通过对所拍摄的图像进行处理来检测用户99的运动。

[0122] 检测器103例如被一体地固定于音响再现系统100的壳体,检测壳体的运动的速度。包含上述的壳体的音响再现系统100由于在用户99佩戴之后与用户99的头部一体地移动,所以检测器103作为结果能够检测用户99的头部的运动速度。

[0123] 检测器103,例如作为用户99的头部的运动量,既可以检测以在三维空间内相互正交的3轴中的至少一个轴为旋转轴的旋转量,也可以检测以上述3轴中的至少一个轴为位移方向的位移量。此外,检测器103也可以检测旋转量及位移量双方作为用户99的头部的运动量。

[0124] 感测信息输入部114从检测器103取得用户99的头部的运动速度。更具体地讲,感测信息输入部114取得每单位时间由检测器103检测到的用户99的头部的运动量作为运动速度。这样,感测信息输入部114从检测器103取得旋转速度及位移速度的至少一方。这里取得的用户99的头部的运动量用于决定三维声场内的用户99的位置及姿势(换言之是坐标及朝向)。在音响再现系统100中,基于所决定的用户99的坐标及朝向来决定声像的相对的位置,再现声音。因此,能够根据用户99的头部的运动量来变更三维声场内的收听点。换言之,感测信息输入部114能够受理使收听点与声像(音源对象)的相对位置变化的指示,该指示包含根据指示使相对位置变化的第1变化量。另外,相对位置是指表示由收音装置或收听点与声像(音源对象)的相对距离及相对方向中的至少一方表现的、相对于一方的另一方的位置的概念。

[0125] 处理部121基于所决定的用户99的坐标及朝向,对于规定音,基于上述用户99的坐标及朝向决定使用户99感知为从三维声场内的哪个方向到来的声音,对声音信息进行处理以使再现的输出音信息成为这样的声音。并且,处理部121与上述的处理一起执行用来赋予波动的音响处理。这里赋予的波动,包括音源对象与收音装置的距离在时域中反复变化的相对距离的波动、以及音源对象与收音装置的方向在时域中反复变化的相对方向的波动。

[0126] 图5是表示有关实施方式的处理部的功能结构的框图。处理部121如图5所示,作为用来执行音响处理的功能部分而具备判定部122、存储部123及执行部124。另外,处理部121作为与上述声音信息的处理有关的功能部分而具有未图示的其他的功能部分。

[0127] 判定部122进行用来决定是否执行音响处理的判定。判定部122例如通过判定是否满足规定的条件,在满足该规定的条件的情况下决定为执行音响处理,在不满足该规定的条件的情况下决定为不执行音响处理。关于规定的条件的详细情况在后面叙述。表示规定的条件的信息例如被存储部123存储在存储设备中。

[0128] 存储部123是进行对存储有信息的存储设备(未图示)保存信息及读出信息的处理的存储控制器。

[0129] 执行部124根据判定部122的判定结果执行音响处理。

[0130] 信号输出部141是生成输出音信号并将所生成的输出音信号向驱动器104输出的功能部。

[0131] 信号输出部141针对与用来决定声音的固定位置并使得定位在该位置的处理一起根据判定结果执行音响处理之后的声音信息,生成作为数字数据的输出声音信号。并且,信号输出部141通过基于输出声音信号进行从数字信号向模拟信号的信号变换等,生成波形信号,基于波形信号使驱动器104产生声波,对用户99提示声音。驱动器104例如具有振动板、磁铁及音圈等的驱动机构。驱动器104根据波形信号使驱动机构动作,通过驱动机构使振动板振动。这样,驱动器104通过与输出声音信号相应的振动板的振动而产生声波(是指“再现”输出音信号,即用户99的感知不包含在“再现”的意思中),声波在空气中传播而传递到用户99的耳朵,用户99感知到声音。

[0132] [有关本实施方式的音响再现系统的另一例]

[0133] 在上述的例子中,说明了有关本实施方式的音响再现系统100是声音提示设备,具备信息处理装置101、通信模块102、检测器103和驱动器104,但也可以将音响再现系统100的功能用多个装置实现,也可以用一个装置实现。使用图6~图15进行说明。图6~图15是用来说明有关实施方式的音响再现系统的另一例的图。

[0134] 例如,也可以是信息处理装置601包含在声音提示设备602中,声音提示设备602进行音响处理和声音的提示双方。此外,也可以是信息处理装置601和声音提示设备602分担实施在本公开中说明的音响处理,也可以是经由网络与信息处理装置601或声音提示设备602连接的服务器实施在本公开中说明的音响处理的一部分或整体。

[0135] 另外,在上述说明中,称作信息处理装置601,但在信息处理装置601对将在声音信号或音响处理中使用的空间信息的至少一部分的数据编码而生成的比特流进行解码来实施音响处理的情况下,也可以将信息处理装置601称作解码装置,也可以将音响再现系统100(即,图中的立体音响再现系统600)称作解码处理系统。

[0136] 这里,对音响再现系统100作为解码处理系统发挥功能的例子进行说明。

[0137] <编码装置的例>

[0138] 图7是表示作为本公开的编码装置的一例的编码装置700的构成的功能框图。

[0139] 输入数据701是包含对编码器702输入的空间信息及/或声音信号的作为编码对象的数据。关于空间信息的详细情况在后面进行说明。

[0140] 编码器702将输入数据701编码,生成编码数据703。编码数据703例如是通过编码处理生成的比特流。

[0141] 存储器704保存编码数据703。存储器704例如可以是硬盘或SSD(Solid—State Drive:固态驱动器),也可以是其他的存储装置。

[0142] 另外,在上述说明中作为存储在存储器704中的编码数据703的一例而举出了通过编码处理生成的比特流,但也可以是比特流以外的数据。例如,编码装置700也可以在存储器704中存储将比特流变换为规定的数据格式而生成的变换后的数据。变换后的数据例如也可以是保存有一个或多个比特流的文件或复用流。这里,文件是具有例如ISOBMFF(ISO Base Media File Format:ISO基础媒体文件格式)等的文件格式的文件。此外,编码数据703也可以是将上述的比特流或文件分割而生成的多个包的形式。在将由编码器702生成的

比特流变换为与比特流不同的数据的情况下,编码装置700也可以具备未图示的变换部,也可以由CPU(Central Processing Unit:中央处理器)进行变换处理。

[0143] <解码装置的例>

[0144] 图8是表示作为本公开的解码装置的一例的解码装置800的构成的功能框图。

[0145] 存储器804例如保存有与由编码装置700生成的编码数据703相同的数据。存储器804读出所保存的数据,作为解码器802的输入数据803输入。输入数据803例如是作为解码对象的比特流。存储器804例如既可以是硬盘或SSD,也可以是其他的存储装置。

[0146] 另外,解码装置800也可以不是将存储器804所存储的数据原样作为输入数据803,而是将对所读出的数据进行变换而生成的变换后的数据作为输入数据803。变换前的数据例如也可以是保存有一个或多个比特流的复用数据。这里,复用数据也可以是具有例如ISOBMFF等的文件格式的文件。此外,变换前的数据也可以是将上述的比特流或文件分割而生成的多个包的形式。在将与从存储器804读出的比特流不同的数据变换为比特流的情况下,解码装置800也可以具备未图示的变换部,也可以由CPU进行变换处理。

[0147] 解码器802将输入数据803解码,生成对听者提示的声音信号801。

[0148] <编码装置的另一例>

[0149] 图9是表示作为本公开的编码装置的另一例的编码装置900的构成的功能框图。在图9中,对于具有与图7的构成相同的功能的构成赋予与图7的构成相同的标号,关于这些构成省略说明。

[0150] 编码装置700具备存储编码数据703的存储器704,相对于此,编码装置900在具备将编码数据703对外部发送的发送部901这一点上与编码装置700不同。

[0151] 发送部901基于编码数据703或将编码数据703变换而生成的其他数据形式的数据将发送信号902对其他的装置或服务器发送。在发送信号902的生成中使用的数据例如是在编码装置700中说明的比特流、复用数据、文件或包。

[0152] <解码装置的另一例>

[0153] 图10是表示作为本公开的解码装置的另一例的解码装置1000的构成的功能框图。在图10中,对于具有与图8的构成相同的功能的构成赋予与图8的构成相同的标号,关于这些构成省略说明。

[0154] 解码装置800具备读出输入数据803的存储器804,相对于此,解码装置1000在具备从外部接收输入数据803的接收部1001这一点上与解码装置800不同。

[0155] 接收部1001将接收信号1002接收而取得接收数据,输出对解码器802输入的输入数据803。接收数据也可以与对解码器802输入的输入数据803相同,也可以是与输入数据803不同的数据形式的数据。在接收数据是与输入数据803不同的数据形式的数据的情况下,接收部1001也可以将接收数据变换为输入数据803,也可以是解码装置1000具备的未图示的变换部或CPU将接收数据变换为输入数据803。接收数据例如是在编码装置900中说明的比特流、复用数据、文件或包。

[0156] <解码器的功能说明>

[0157] 图11是表示作为图8或图10中的解码器802的一例的解码器1100的构成的功能框图。

[0158] 输入数据803是被编码的比特流,包含作为被编码的声音信号的编码声音数据和

在音响处理中使用的元数据。

[0159] 空间信息管理部1101取得输入数据803中包含的元数据,解析元数据。元数据包含记述有配置在声音空间中的作用于声音的要素的信息。空间信息管理部1101管理将元数据解析而得到的音响处理所需要的空间信息,对渲染部1103提供空间信息。另外,在本公开中将在音响处理中使用的信息称作空间信息,但也可以是其以外的称呼。关于在该音响处理中使用的信息,例如也可以称作声音空间信息,也可以称作场景信息。此外,在音响处理中使用的信息随时间而变化的情况下,也可以将对渲染部1103输入的空间信息称作空间状态、声音空间状态、场景状态等。

[0160] 此外,也可以将空间信息按每个声音空间或按每个场景管理。例如,在将不同的房间表现为虚拟空间的情况下,既可以将各个房间作为不同的声音空间的场景来管理,也可以尽管是相同的空间但作为根据表现的场合而不同的场景来管理空间信息。在空间信息的管理中,也可以赋予识别各个空间信息的识别符。空间信息的数据既可以包含在作为输入数据803的一形态的比特流中,也可以是比特流包含空间信息的识别符,从比特流以外取得空间信息的数据。在比特流中仅包含空间信息的识别符的情况下,也可以在渲染时使用空间信息的识别符,取得存储在音响信号处理装置的存储器或外部的服务器中的空间信息的数据作为输入数据。

[0161] 另外,空间信息管理部1101管理的信息并不限定于在比特流中包含的信息。例如,输入数据803中,作为不包含在比特流中的数据,也可以包含从提供VR或AR的软件应用或服务服务器取得的表示空间特性或构造的数据。此外,例如输入数据803中,作为不包含在比特流中的数据,也可以包含表示听者或对象的特性或位置等的信息。此外,输入数据803中,作为表示听者的位置的信息,也可以包含由包括解码装置的终端所具备的传感器取得的信息、或表示基于由传感器取得的信息推测出的终端的位置的信息。即,空间信息管理部1101也可以与外部的系统或服务服务器通信而取得空间信息及听者的位置。此外,空间信息管理部1101也可以从外部的系统取得时钟同步信息,执行与渲染部1103的时钟同步的处理。另外,上述的说明中的空间也可以是虚拟地形成的空间、即VR空间,也可以是实空间或与实空间对应的虚拟空间、即AR空间或MR(Mixed Reality:混合现实)空间。此外,虚拟空间也可以被称作声场或声音空间。此外,上述的说明中的表示位置的信息既可以是表示空间内的位置的坐标值等的信息,也可以是表示相对于规定的基准位置的相对位置的信息,也可以是表示空间内的位置的移动或加速度的信息。

[0162] 声音数据解码器1102将输入数据803中包含的编码声音数据解码,取得声音信号。

[0163] 立体音响再现系统600取得的编码声音数据例如是被以MPEG—H 3DAudio(ISO/IEC 23008—3)等的规定的形式编码的比特流。另外,MPEG—H 3D Audio只不过是生成比特流中包含的编码声音数据时能够使用的编码方式的一例,也可以包含为以其他的编码方式编码的比特流和编码声音数据。例如,使用的编码方式也可以是MP3(MPEG—1Audio Layer—3)、AAC(Advanced Audio Coding:高级音频编码)、WMA(Windows Media Audio:视窗媒体音频)、AC3(Audio Codec—3)、Vorbis等的非可逆编解码,也可以是ALAC(Apple Lossless Audio Codec:Apple无损音频编解码器)、FLAC(Free Lossless Audio Codec:无损音频编解码器)等的可逆编解码,也可以使用上述以外的任意的编码方式。例如,也可以是PCM(pulse code modulation:脉冲编码调制)数据为编码声音数据的一种。在此情况下,

解码处理例如在该PCM数据的量化比特数是N的情况下,也可以是将N比特的二进制数变换为渲染部1103能够处理的数形式(例如浮点形式)的处理。

[0164] 渲染部1103以声音信号和空间信息为输入,使用空间信息对声音信号实施音响处理,输出音响处理后的声音信号801。

[0165] 空间信息管理部1101在开始渲染之前,读入输入信号的元数据,检测由空间信息规定的对象或声音等的渲染项目,发送给渲染部1103。在渲染开始后,空间信息管理部1101掌握空间信息及听者的位置的随时间的变化,将空间信息更新并进行管理。并且,空间信息管理部1101将更新后的空间信息发送给渲染部1103。渲染部1103基于输入数据中包含的声音信号和从空间信息管理部1101接收到的空间信息,生成附加了音响处理的声音信号并输出。

[0166] 空间信息的更新处理和附加了音响处理的声音信号的输出处理既可以由相同的线程执行,也可以将空间信息管理部1101和渲染部1103分别分配给独立的线程。在空间信息的更新处理和附加了音响处理的声音信号的输出处理由不同的线程进行处理的情况下,既可以单独地设定线程的启动频率,也可以并行地执行处理。

[0167] 通过空间信息管理部1101和渲染部1103以不同的独立的线程执行处理,能够优先地对渲染部1103分配运算资源,所以在一点延迟都不能容许的出声处理的情况下,例如即使是在1个采样(0.02msec)延迟的情况下也会发生小的噪声的出声处理也能够安全地实施。此时,对空间信息管理部1101限制运算资源的分配。但是,空间信息的更新是比声音信号的输出处理低频率的处理(例如,如收听者的脸的朝向的更新那样的处理)。因此,不像声音信号的输出处理那样必须瞬间响应,所以即使限制运算资源的分配,也不会对给收听者带来的音响上的品质产生大的影响。

[0168] 空间信息的更新既可以按每预先设定的时间或期间定期地执行,也可以在满足了预先设定的条件的情况下执行。此外,空间信息的更新既可以由听者或声音空间的管理者手动地执行,也可以以外部系统的变化为契机来执行。例如,在收听者操作控制器而将自身的化身的站立位置瞬间地歪斜、或将时刻瞬时地前进或回退、或虚拟空间的管理者突然实施了变更场景的环境那样的呈现的情况下,配置有空间信息管理部1101的线程也可以除了定期的启动以外还作为单发的中断处理而启动。

[0169] 执行空间信息的更新处理的信息更新线程承担的作用例如是基于收听者佩戴的VR护目镜的位置或朝向,将配置在虚拟空间内的收听者的化身的位置或朝向更新的处理、以及在虚拟空间内移动的物体的位置的更新等,在以几十Hz左右的比较低的频率启动的处理线程内提供。也可以以这样的发生频率低的处理线程进行反映直接音的性质的处理。这是因为,直接音的性质变动的频率比用于音频输出的音频处理帧的发生频率低。反而这样做能够使该处理的运算负荷相对减小,而且如果以不必要的快频率更新信息则会发生产生脉冲性的噪声的风险,所以还能够避免该风险。

[0170] 图12是表示作为图8或图10中的解码器802的另一例的解码器1200的构成的功能框图。

[0171] 图12在输入数据803不是包含编码声音数据而是包含未被编码的声音信号这一点上与图11不同。输入数据803包括包含元数据的比特流和声音信号。

[0172] 空间信息管理部1201与图11的空间信息管理部1101相同,所以省略说明。

[0173] 渲染部1202与图11的渲染部1103相同,所以省略说明。

[0174] 另外,在上述说明中将图12的构成称作解码器,但也可以称作实施音响处理的音响处理部。此外,也可以将包括音响处理部的装置称作音响处理装置而不是解码装置。此外,也可以将音响信号处理装置(信息处理装置601)称作音响处理装置。

[0175] <编码装置的物理构成>

[0176] 图13是表示编码装置的物理构成的一例的图。此外,图13所示的编码装置是上述的编码装置700及900等的一例。

[0177] 图13的编码装置具备处理器、存储器和通信IF。

[0178] 处理器例如是CPU(Central Processing Unit:中央处理器)、DSP(Digital Signal Processor:数字信号处理器)或GPU(Graphics Processing Unit:图形处理器),也可以通过该CPU、DSP或GPU执行存储在存储器中的程序来实施本公开的编码处理。此外,处理器也可以是进行包括本公开的编码处理的对声音信号的信号处理的专用电路。

[0179] 存储器例如由RAM(Random Access Memory:随机访问存储器)或ROM(Read Only Memory:只读存储器)构成。存储器也可以包括硬盘等的磁存储介质或SSD(Solid State Drive:固态驱动器)等的半导体存储器等。此外,也可以包括嵌入在CPU或GPU中的内部存储器在内而称作存储器。

[0180] 通信IF(Inter Face:接口)例如是与Bluetooth(注册商标)或WIGIG(注册商标)等的通信方式对应的通信模块。编码装置具有经由通信IF与其他通信装置进行通信的功能,发送被编码的比特流。

[0181] 通信模块例如由与通信方式对应的信号处理电路和天线构成。在上述的例子中,作为通信方式而举出Bluetooth(注册商标)或WIGIG(注册商标)为例,但也可以与LTE(Long Term Evolution:长期演进)、NR(New Radio:新空口)或Wi-Fi(注册商标)等的通信方式对应。此外,通信IF也可以不是上述那样的无线通信方式,而是Ethernet(注册商标)、USB(Universal Serial Bus:通用串行总线)、HDMI(注册商标)(High-Definition Multimedia Interface:高清多媒体接口)等的有线的通信方式。

[0182] <音响信号处理装置的物理构成>

[0183] 图14是表示音响信号处理装置的物理构成的一例的图。另外,图14的音响信号处理装置也可以是解码装置。此外,也可以将这里说明的构成的一部分装备在声音提示装置602中。此外,图14所示的音响信号处理装置是上述的音响信号处理装置601的一例。

[0184] 图14的音响信号处理装置具备处理器、存储器、通信IF、传感器和扬声器。

[0185] 处理器例如是CPU(Central Processing Unit:中央处理器)、DSP(Digital Signal Processor:数字信号处理器)或GPU(Graphics Processing Unit:图形处理器),也可以通过该CPU、DSP或GPU执行存储在存储器中的程序来实施本公开的音响处理或解码处理。此外,处理器也可以是进行包括本公开的音响处理的对声音信号的信号处理的专用电路。

[0186] 存储器例如由RAM(Random Access Memory:随机访问存储器)或ROM(Read Only Memory:只读存储器)构成。存储器也可以包括硬盘等的磁存储介质或SSD(Solid State Drive:固态驱动器)等的半导体存储器等。此外,也可以包括嵌入在CPU或GPU中的内部存储器在内而称作存储器。

[0187] 通信IF(Inter Face:接口)例如是与Bluetooth(注册商标)或WIGIG(注册商标)等的通信方式对应的通信模块。图2I所示的音响信号处理装置具有经由通信IF与其他通信装置进行通信的功能,取得解码对象的比特流。将所取得的比特流例如保存在存储器中。

[0188] 通信模块例如由与通信方式对应的信号处理电路和天线构成。在上述的例子中,作为通信方式而举出Bluetooth(注册商标)或WIGIG(注册商标)为例,但也可以与LTE(Long Term Evolution:长期演进)、NR(New Radio:新空口)或Wi-Fi(注册商标)等的通信方式对应。此外,通信IF也可以不是上述那样的无线通信方式,而是Ethernet(注册商标)、USB(Universal Serial Bus:通用串行总线)、HDMI(注册商标)(High-Definition Multimedia Interface:高清多媒体接口)等的有线的通信方式。

[0189] 传感器进行用来推测听者的位置或朝向的感测。具体而言,传感器基于听者的头部等身体的一部分或整体的位置、朝向、运动、速度、角速度、或加速度等中的某一个或多个检测结果来推测听者的位置及/或朝向,生成表示听者的位置及/或朝向的位置信息。另外,位置信息也可以是表示实空间中的听者的位置及/或朝向的信息,也可以是表示以规定的时间点的听者的位置及/或朝向为基准的听者的位置及/或朝向的位移的信息。此外,位置信息也可以是表示与立体音响再现系统或具备传感器的外部装置的相对的位置及/或朝向的信息。

[0190] 传感器例如也可以是相机等的摄像装置或LiDAR(Light Detection And Ranging:激光雷达)等的测距装置,也可以拍摄听者的头部的运动,通过对所拍摄的图像进行处理来检测听者的头部的运动。此外,作为传感器例如也可以使用利用毫米波等的任意的频带的无线进行位置推测的装置。

[0191] 另外,图14所示的音响信号处理装置也可以从具备传感器的外部的设备经由通信IF取得位置信息。在此情况下,音响信号处理装置也可以不包括传感器。这里,外部的设备例如是在图6中说明的声音提示装置602或佩戴在听者的头部上的立体影像再现装置等。此时,传感器例如将陀螺仪传感器及加速度传感器等各种传感器组合而构成。

[0192] 例如作为听者的头部的运动的速度,传感器既可以检测以在声音空间内相互正交的3轴的至少1个为旋转轴的旋转的角速度,也可以检测以上述3轴的至少1个为位移方向的位移的加速度。

[0193] 例如作为听者的头部的运动的量,传感器既可以检测以在声音空间内相互正交的3轴的至少1个为旋转轴的旋转量,也可以检测以上述3轴的至少1个为位移方向的位移量。具体而言,传感器检测6DoF(位置(x,y,z)及角度(yaw,pitch,roll))作为听者的位置。传感器将陀螺仪传感器及加速度传感器等用于运动的检测的各种传感器组合而构成。

[0194] 另外,传感器只要能够检测听者的位置即可,也可以由相机或GPS(Global Positioning System:全球定位系统)接收机等实现。也可以使用利用LiDAR(Laser Imaging Detection and Ranging:激光雷达)等实施自位置推测而得到的位置信息。例如,在由智能电话实现声音信号再现系统的情况下传感器被内置在智能电话中。

[0195] 此外,传感器中也可以包括检测图14所示的音响信号处理装置的温度的热电偶等的温度传感器、以及检测音响信号处理装置所具备的或与音响信号处理装置连接的电池的剩余量的传感器等。

[0196] 扬声器例如具有振动板、磁铁或音圈等的驱动机构和放大器,将音响处理后的声

音信号作为声音提示给听者。扬声器根据经由放大器放大的声音信号(更具体地讲,表示声音的波形的波形信号)使驱动机构动作,通过驱动机构使振动板振动。这样,对应于声音信号而振动的振动板产生声波,声波在空气中传播而传递到听者的耳朵,听者感知声音。

[0197] 另外,这里举出图14所示的音响信号处理装置具备扬声器并经由该扬声器提示音响处理后的声音信号的情况为例进行了说明,但声音信号的提示机构并不限定于上述的构成。例如,也可以对由通信模块连接的外部的声音提示装置602输出音响处理后的声音信号。由通信模块进行的通信既可以是有线也可以是无无线。此外,作为另一例,也可以是图14所示的音响信号处理装置具备输出声音的模拟信号的端子,将耳塞等的线缆与端子连接而从耳塞等提示声音信号。在上述的情况下,作为声音提示装置602的佩戴在听者的头部或身体的一部分上的耳机、耳塞、头戴显示器、颈部扬声器、可穿戴扬声器或由被固定的多个扬声器构成的环绕扬声器等再现声音信号。

[0198] <渲染部的功能说明>

[0199] 图15是表示图11及图12的渲染部1103及1202的详细的构成的一例的功能框图。

[0200] 渲染部由解析部和合成部构成,对输入信号中包含的声音数据附加音响处理并输出。

[0201] 以下,对输入信号中包含的信息进行说明。

[0202] 输入信号例如由空间信息、传感器信息和声音数据构成。输入信号也可以包含由声音数据和元数据(控制信息)构成的比特流,在此情况下,也可以在元数据中包含空间信息。

[0203] 空间信息是与立体音响再现系统所形成的声音空间(三维声场)有关的信息,由与声音空间中包含的对象有关的信息以及与听者有关的信息构成。对象中存在发出声音而成为音源的音源对象和不发出声音的非发声对象。非发声对象作为将音源对象所发出的声音反射的障碍物对象发挥功能,但也有作为音源对象将其他音源对象所发出的声音反射的障碍物对象发挥功能的情况。

[0204] 作为对音源对象和非发声对象共同赋予的信息,有位置信息、形状信息、对象将声音反射时的音量的衰减率等。

[0205] 位置信息由欧几里得空间的例如X轴、Y轴、Z轴这3轴的坐标值表示,但也可以不一定是三维信息。例如,也可以是由X轴、Y轴这两轴的坐标值表示的二维信息。对象的位置信息由用网格、体素表现的形状的代表位置决定。

[0206] 形状信息也可以包含与表面的材料有关的信息。

[0207] 此外,也可以包含表示对象是否属于生物的信息、表示对象是否是移动体的信息等。在对象是移动体的情况下,位置信息也可以随着时间而移动,变化后的位置信息或变化量被传送给渲染部。

[0208] 与音源对象有关的信息除了上述的对音源对象和非发声对象共同赋予的信息以外,还包含声音数据和为了将声音数据放射到声音空间内而需要的信息。

[0209] 声音数据是表示与声音的频率及强弱有关的信息等的表现了被听者感知的声音的数据。声音数据典型的是PCM信号,但也可以是使用MP3等的编码方式压缩的数据。在此情况下,至少需要在该信号到达合成部之前进行解码,所以也可以在渲染部中包括未图示的解码部。或者,也可以由声音数据解码器1102解码。

[0210] 只要对于1个音源对象设定至少1个声音数据即可,也可以设定多个声音数据。此外,也可以赋予识别各个声音数据的识别信息,作为与音源对象有关的信息而保持声音数据的识别信息。

[0211] 作为为了将声音数据放射到声音空间内而需要的信息,例如可以包含在再现声音数据时作为基准的基准音量的信息、表示声音数据的性质(也称作特性)的信息、与音源对象的位置有关的信息、与音源对象的朝向有关的信息、与音源对象发出的声音的指向性有关的信息等。基准音量的信息例如是将声音数据放射到声音空间中时的音源位置处的声音数据的振幅值的有效值,也可以作为分贝(dB)值而用浮点表示。

[0212] 例如在基准音量为0dB的情况下,也可以表示不使声音数据所表示的信号电平的音量增减而以原来的音量从与上述位置有关的信息所指示的位置对声音空间放射声音,在-6dB的情况下,也可以表示使声音数据所表示的信号电平的音量成为约一半并从与上述位置有关的信息所指示的位置对声音空间放射声音。这些信息也可以针对1个声音数据或针对多个声音数据一起赋予。

[0213] 表示声音数据的性质的信息例如可以是与音源的音量有关的信息,并且是表示其时间序列上的变动的信息。例如,在声音空间是虚拟会议室、音源是讲话者的情况下,音量在短时间内断续地变迁。如果将其更简单地表现,则可以说有声部分和无声部分交替地产生。

[0214] 此外,在声音空间是音乐厅、音源是演奏者的情况下,将音量维持一定时间长度。此外,在声音空间是战场、音源是爆炸物的情况下,爆炸声的音量仅在一瞬间变大,然后是无声并持续。这样,音源的音量的信息不仅包含声音的大小的信息,还包含声音的大小的变迁的信息,也可以将这样的信息设为表示声音数据的性质的信息。

[0215] 这里,声音的大小的变迁的信息也可以是将频率特性以时间序列表示的数据。也可以是表示有声的区间的持续时间长度的数据。也可以是表示有声的区间的持续时间长度和无声的区间的时间长度的时间序列的数据。也可以是以多组时间序列列举了可看作声音信号的振幅为恒常(可看作大致一定)的持续时间和其间的该信号的振幅值的数据的数据等。也可以是可看作声音信号的频率特性为恒常的持续时间的数据。也可以是以多组时间序列列举了可看作声音信号的频率特性为恒常的持续时间和其间的该频率特性的数据的数据等。

[0216] 作为数据的形式,例如也可以是表示波谱图的大致形状的数据。此外,也可以将作为上述频率特性的基准的音量设为上述基准音量。表示基准音量的信息和声音数据的性质的信息除了用于计算使听者感知的直接音或反射音的音量以外,还在用于进行是否使听者感知的选择的选择处理中使用。关于表示声音数据的性质的信息的其他例或向具体的选择处理的使用方式在后面叙述。

[0217] 与朝向有关的信息典型地由yaw、pitch、roll表现。或者,也可以省略roll的旋转,而由方位角(yaw)、仰角(pitch)表现。朝向信息也可以随着时间而变化,在发生了变化的情况下被传送到渲染部。

[0218] 与听者有关的信息是与声音空间中的听者的位置信息及朝向有关的信息。位置信息由欧几里得空间的XYZ轴的位置表示,但也可以不一定是三维信息,也可以是二维信息。与朝向有关的信息典型地由yaw、pitch、roll表现。或者,也可以省略roll的旋转,而由方位

角(yaw)、仰角(pitch)表现。位置信息和朝向信息也可以随着时间而变化,在发生了变化的情况下被传送到渲染部。

[0219] 传感器信息是包含由听者佩戴的传感器检测到的旋转量或位移量等和听者的位置及朝向的信息。传感器信息被传送到渲染部,渲染部基于传感器信息更新听者的位置及朝向的信息。传感器信息例如也可以使用由便携终端利用GPS、相机或LiDAR (Laser Imaging Detection and Ranging)等实施自身位置推测而得到的位置信息。此外,也可以检测从传感器以外经由通信模块从外部取得的信息作为传感器信息。也可以从传感器取得表示声音信号处理装置的温度的信息及表示电池的剩余量的信息。也可以实时地取得声音信号处理装置或声音信号提示装置的运算资源(CPU能力、存储器资源、PC性能)等。

[0220] 解析部承担与上述例子中的取得部111同样的功能。即,进行输入信号的解析,由处理部121取得需要的信息。

[0221] 合成部承担与上述例子中的处理部121和信号输出部141同样的功能。基于直接音的声音信号和解析部计算出的直接音到来时刻及直接音到来时音量的信息,对被输入的声音信号进行加工而生成直接音。此外,基于解析部计算出的反射音到来时刻和反射音到来时音量的信息,对被输入的声音信号进行加工而生成反射音。合成部将所生成的直接音与反射音合成并输出。

[0222] [动作]

[0223] 接着,参照图16~图19,对在上述中所说明的音响再现系统100的动作进行说明。图16是表示有关实施方式的音响再现系统的动作的流程图。此外,图17是用来说明有关实施方式的音响处理的频率特性的图。图18是用来说明有关实施方式的音响处理的波动的大小的图。图19是用来说明有关实施方式的音响处理的波动的周期及角度的图。

[0224] 另外,假设在图16所示的各步骤之前,通过基于控制信息的判定而成为了执行音响处理的设定来进行说明。如图16所示,首先,由取得部111取得声音信息(声音信号)(S101)。接着,判定部122判定是否执行音响处理。具体而言,判定部122读出存储在存储部123中的规定的条件,通过判定是否满足规定的条件,判定是否执行音响处理(S102)。

[0225] 关于规定的条件,以下说明几个例子。

[0226] 首先,在所取得的声音信息中的规定音的声压在时域中的变化为规定的阈值以下的情况下,认为声音信息的规定音不包含波动,波动的赋予是适当的。作为可以说是适合于进行音响处理的条件,如果设定与声压在时域中的变化有关的条件,则能够在表示上述的阈值以下的声压在时域中的变化的情况下判定为满足规定的条件。

[0227] 这里,在图17中表示了在各频率的声音从音源(各虚线圆的中心)发出的情况下,在水平面内的各方向上以相同的声压到达的距离之差。在图17所示的各图中,表示了在该频率下声音的各方向上的传播特性之差,可以说形状越不规则,则越容易反映音源的波动。换言之,为了根据声压在时域中的变化来判断音源的波动,可以将规定音按每个频率分解,以更容易反映音源的波动的频率判定是否表示声压在时域中的变化。例如,如果是图中所示的1000Hz以上的频率,则形状从圆形变化为不规则的形状,可以说容易反映波动。此外,如果是图中所示的4000Hz以上的频率,则形状从圆形变化为更不规则的形状,可以说更容易反映波动。

[0228] 相反,如图17所示,在赋予波动的情况下,也可以说即使对小于1000Hz的频率执行

音响处理也难以得到波动的效果。因此,在音响处理中,也可以仅对1000Hz以上的频率执行音响处理,也可以仅对4000Hz以上的频率执行音响处理。或者,也可以执行频率越大则波动越大的音响处理。

[0229] 此外,使用所取得的声音信息中的规定位置或规定音的声压推测收音装置与音源的位置关系,在推测出的位置关系为规定的阈值以下的情况下,可以认为正在使用耳机的麦克风那样的近讲式收音装置,所以可认为在声音信息的规定音中不包含波动,波动的赋予是适当的。作为可以说适合于进行音响处理的条件,如果设定与推测出的位置关系有关的条件,则能够在表示上述的阈值以下的情况判定为满足规定的条件。

[0230] 这里,在图18中表示了在XYZ的3轴中标绘了人的头部的运动的结果。在图18中,在上段表示Y轴方向(上下方向)的头部的运动的标绘,在中段表示Z轴方向(前后方向)的头部的运动的标绘,在下段表示X轴方向(左右方向)的头部的运动的标绘。如图中所示,可知人的头部在X轴方向(左右方向)上有 $\pm 0.2\text{m}$ 的运动,在Y轴方向(上下方向)上有 $\pm 0.02\text{m}$ 的运动,在Z轴方向(前后方向)上有 $\pm 0.05\text{m}$ 的运动。

[0231] 即,如果没有这样的大小的运动,则可以想到推测出的位置关系是使用耳机的麦克风那样的近讲式收音装置那样的规定的阈值以下。

[0232] 相反,如图18所示,在赋予波动的情况下,也可以在X轴方向(左右方向)上再现 $\pm 0.2\text{m}$ 的运动,在Y轴方向(上下方向)上再现 $\pm 0.02\text{m}$ 的运动,在Z轴方向(前后方向)上再现 $\pm 0.05\text{m}$ 的运动而执行音响处理。这样,也能够对应于收音装置与音源的位置关系的处理条件下执行音响处理。

[0233] 此外,在图19中表示了对于人的头部的运动,在Yaw、Pitch、Roll的3个旋转轴中标绘了旋转角的结果。在图19中,在上段表示了Yaw角的旋转角,在中段表示了Pitch角的旋转角,在下段表示了Roll角的旋转角。如图中所示,可知人的头部以3~4s周期,Yaw角有 ± 20 度的旋转,Pitch角有 ± 10 度的旋转,Yaw角有 ± 3 度的旋转。

[0234] 即,如果没有这样的周期及角度的运动,则可以想到推测出的位置关系是使用耳机的麦克风那样的近讲式收音装置那样的规定的阈值以下。

[0235] 相反,如图19所示,在赋予波动的情况下,也可以以3~4s周期,在Yaw角再现 ± 20 度的旋转,在Pitch角再现 ± 10 度的旋转,在Yaw角再现 ± 3 度的旋转而执行音响处理。这样,也能够对应于收音装置与音源的位置关系的处理条件下执行音响处理。

[0236] 此外,使用与收集声音时的状况有关的收音状况信息,在收音状况信息所表示的混响等级及/或噪声等级是规定的阈值以下的情况下,可以想到正在使用耳机的麦克风那样的近讲式收音装置,所以可认为在声音信息的规定音中不包含波动,波动的赋予是适当的。作为可以说适合于进行音响处理的条件,如果设定与收音状况信息所表示的混响等级及/或噪声等级有关的条件,则能够在表示上述的阈值以下的混响等级及/或噪声等级的情况下判定为满足规定的条件。

[0237] 除此以外,也可以使用利用耳机的麦克风那样的近讲式收音装置收集声音之类的与收音设备有关的信息(型号等的确定设备的信息或是否需要赋予波动等的表示设备的特性的信息),在由该信息表示使用耳机的麦克风那样的近讲式收音装置的情况下判定为满足规定的条件。

[0238] 回到图16,在判定部122判定为满足上述规定的条件的情况下(S102中“是”),执行

部124执行音响处理(S103)。另一方面,在判定部122判定为不满足上述规定的条件的情况下(S102中“否”),执行部124不执行音响处理(S104)。接着,信号输出部141生成输出声音信号并输出(S105)。

[0239] [另一例]

[0240] 以下,使用图20及图21对有关实施方式的另一例的音响再现系统进行说明。图20是表示有关实施方式的另一例的处理部的功能结构的框图。图21是表示有关实施方式的另一例的音响处理装置的动作的流程图。另外,在以下的有关另一例的说明中,有时通过将上述实施方式的一部分说明的“收音装置”改称作“收听点”而省略说明。

[0241] 与上述实施方式的音响再现系统100相比,有关实施方式的另一例的音响再现系统在代替处理部121而具备处理部121a这一点上不同。

[0242] 处理部121a中代替判定部122而具有计算部125。计算部125计算第1变化量及第2变化量。第1变化量是基于使收听点与音源对象的相对位置变化的指示的变化量,与所谓的VR空间内的移动的移动量对应。并且,如果限于虚拟声音空间,则是随着收听点的移动的收听点与音源对象的相对位置的变化量。关于第1变化量,通过从作为传感器的检测器103取得检测结果而取得此时的相对位置的变化指示即变化量。即,在本例中,取得部111(特别是感测信息输入部114)受理包含第1变化量的指示。

[0243] 在本实施方式中,除了这样的相对位置的变化以外,还发生由波动引起的收听点的变化,所以分为第1变化量和第2变化量来计算。另外,通过将第2变化量设为0,能够不经过由判定部122进行的处理来区分执行音响处理和不执行音响处理。第2变化量既可以通过检测结果来计算,也可以与检测结果独立地计算。例如,第2变化量也可以通过使用由检测结果表示的音源对象与收听点的相对位置的变化速度、或作为变化量的第1变化量的函数计算。或者,第2变化量也可以不使用(独立于)音源对象与收听点的相对位置的变化速度或作为变化量的第1变化量,而只是根据控制信息及收音状况信息等的内容制作时对内容赋予的信息等唯一地计算。

[0244] 此外,在第1变化量大的情况下,存在音源对象相对于停止的收听点大幅移动的情况。在这样的情况下,第1变化量越大则该音源对象的波动也越大是自然的。即,可以是第1变化量越大则第2变化量越大。因此,在音响处理中,相当于波动的大小的第2变化量根据第1变化量成为只要第1变化量越大则越大的第2变化量即可。

[0245] 另一方面,在音响处理中,作为相当于波动的大小的第2变化量根据第1变化量而变化的例子,相反地也有第1变化量越大则设为越小的(例如0的)第2变化量为适当的情况。具体而言,例如在第1变化量大(或相对位置的变化速度快)的情况下,即使赋予波动也不怎么能看到临场感的增加效果。这是因为,由于由波动引起的变化和相对位置的变化同步地重叠或相互抵消,难以使听者感知到被赋予了波动。在这样的情况下,只要第1变化量越大则设为越小的(例如0的)第2变化量即可。

[0246] 以下,说明有关本例的音响再现系统的动作。另外,假设在图21所示的各步骤之前,通过基于控制信息的判定而成为了执行音响处理的设定而进行说明。如图21所示,首先,由取得部111取得声音信息(声音信号)(S201)。接着,计算部125计算第1变化量(S202)。此外,计算部125计算第2变化量(S203)。是否执行音响处理(是否赋予波动)可以根据是否将第2变化量计算为0来设定。并且,执行部124作为音响处理执行使相对位置变化第1变化

量、并且使相对位置在时域中以第2变化量反复变化的音响处理(S204)。然后,信号输出部141生成输出声音信号并输出(S205)。

[0247] (其他实施方式)

[0248] 以上,对实施方式进行了说明,但本公开并不限定于上述的实施方式。

[0249] 例如,在上述的实施方式中说明的音响再现系统既可以作为具备全部构成要素的一个装置实现,也可以对多个装置分配各功能,通过该多个装置协同来实现。在后者的情况下,作为相当于音响处理装置的装置,也可以使用智能电话、平板电脑终端或PC等的音响处理装置。例如,在具有生成附加了音响效果的音响信号的作为渲染器的功能的音响再现系统100中,也可以由服务器承担渲染器的功能的全部或一部分。即,取得部111、处理部121、信号输出部141的全部或一部分也可以存在于未图示的服务器中。在此情况下,音响再现系统100例如将计算机或智能电话等的音响处理装置、佩戴在用户99上的头戴显示器(HMD)、耳机等的声音提示设备和未图示的服务器组合来实现。另外,计算机、声音提示设备和服务器既可以通过相同的网络可通信地连接,也可以通过不同的网络连接。在通过不同的网络连接的情况下,通信发生延迟的可能性高,所以也可以仅在计算机、声音提示设备和服务器通过相同的网络可通信地连接的情况下许可服务器中的处理。此外,也可以根据音响再现系统100受理的比特流的数据量决定服务器是否承担渲染器的全部或一部分的功能。

[0250] 此外,本公开的音响再现系统也可以作为与仅具备驱动器的再现装置连接,对于该再现装置,作为仅再现基于所取得的声音信息生成的输出音信号的音响处理装置来实现。在此情况下,音响处理装置既可以作为具备专用的电路的硬件实现,也可以作为用来使通用的处理器执行特定的处理的软件实现。

[0251] 此外,在上述的实施方式中,也可以将特定的处理部所执行的处理通过其他的处理部执行。此外,也可以变更多个处理的顺序,也可以将多个处理并行地执行。

[0252] 此外,在上述的实施方式中,各构成要素也可以通过执行适合于各构成要素的软件程序来实现。各构成要素也可以通过CPU或处理器等的程序执行部将记录在硬盘或半导体存储器等的记录介质中的软件程序读出并执行来实现。

[0253] 此外,各构成要素也可以由硬件实现。例如,各构成要素也可以是电路(或集成电路)。这些电路既可以作为整体构成1个电路,也可以是分别不同的电路。此外,这些电路分别既可以是通用的电路,也可以是专用的电路。

[0254] 此外,本公开的全局性或具体的形态也可以由装置、装置、方法、集成电路、计算机程序或计算机可读的CD-ROM等的记录介质实现。此外,本公开的整体性或具体的形态也可以由装置、装置、方法、集成电路、计算机程序及记录介质的任意的组合实现。

[0255] 例如,本公开也可以作为由计算机执行的声音信号再现方法实现,也可以作为用来使计算机执行声音信号再现方法的程序实现。本公开也可以作为记录有这样的程序的计算机可读的非暂时性的记录介质实现。

[0256] 除此以外,对于各实施方式实施本领域技术人员想到的各种变形而得到的形态或在不脱离本公开的主旨的范围中将各实施方式的构成要素及功能任意地组合而实现的形态也包含在本公开中。

[0257] 另外,本公开中的被编码的声音信息可以换言之是包含声音信号和元数据的比特流,声音信号是关于由音响再现系统100再现的规定音的信息,元数据是与使该规定音的声

像在三维声场内定位在规定位置时的定位位置有关的信息。也可以由音响再现系统100作为例如以MPEG—H 3D Audio (ISO/IEC 23008—3) 等规定的形式编码的比特流而取得声音信息。作为一例,被编码的声音信号包含关于由音响再现系统100再现的规定音的信息。这里所述的规定音,是存在于三维声场中的音源对象发出的声音或自然环境音,例如可以包含机械音或包括人在内的动物的声音等。另外,在三维声场中存在多个音源对象的情况下,音响再现系统100取得与多个音源对象分别对应的多个声音信号。

[0258] 另一方面,元数据例如是在音响再现系统100中为了控制对声音信号的音响处理而使用的信息。元数据也可以是为了记述由虚拟空间(三维声场)表现的场景而使用的信息。这里所述的场景,是指表示使用元数据由音响再现系统100建模的虚拟空间中的三维影像及音响事件的全部要素的集合体的用语。即,这里所述的元数据,不仅包括对音响处理进行控制的信息,也可以包括对影像处理进行控制的信息。当然,元数据中既可以包括仅对音响处理和影像处理的某一方进行控制的信息,也可以包括在两者的控制中使用的信息。在本公开中音响再现系统100取得的比特流中,有时包含这样的元数据。或者,音响再现系统100也可以如后述那样与比特流分开地以单体取得元数据。

[0259] 音响再现系统100通过使用比特流中包含的元数据及追加取得的交互性的用户99的位置信息等对声音信号进行音响处理,生成虚拟的音响效果。例如,可以考虑附加初始反射音生成、后期混响音生成、衍射音生成、距离衰减效果、定位、声像定位处理或多普勒效应等音响效果。此外,也可以将切换音响效果的全部或一部分的开启/关闭的信息作为元数据来附加。

[0260] 另外,全部的元数据或一部分元数据也可以从声音信息的比特流以外取得。例如,既可以将对音响进行控制的元数据和对影像进行控制的元数据的某一方从比特流以外取得,也可以将两者的元数据都从比特流以外取得。

[0261] 此外,在由音响再现系统100取得的比特流中包含对影像进行控制的元数据的情况下,音响再现系统100也可以具备将能够用于影像的控制的元数据输出给显示图像的显示装置或再现立体影像的立体影像再现装置的功能。

[0262] 此外,作为一例,被编码的元数据包括:与包含发出声音的音源对象及障碍物对象的三维声场有关的信息、以及与使该声音的声像在三维声场内定位在规定位置(即,使得被感知为从规定方向到达的声音)时的定位位置有关的信息、即与规定方向有关的信息。这里,障碍物对象是在音源对象发出的声音到达用户99为止的期间例如将声音遮挡或将声音反射等而可能给用户99感知的声音带来影响的对象。障碍物对象除了静止物体以外,还可能包括人等的动物或机械等的运动体。此外,在三维声场中存在多个音源对象的情况下,对于任意的音源对象而言,其他的音源对象可能成为障碍物对象。此外,建材或无生物等非发声源对象及发出声音的音源对象也都可能成为障碍物对象。

[0263] 作为构成元数据的空间信息,也可以不仅是三维声场的形状,还包含分别表示存在于三维声场中的障碍物对象的形状及位置和存在于三维声场中的音源对象的形状及位置的信息。三维声场是封闭空间或开放空间的哪种都可以,在元数据中,包含例如在地板、墙壁或天花板等的在三维声场中能够将声音反射的构造物的反射率以及存在于三维声场中的障碍物对象的反射率的信息。这里,反射率是反射音相对于入射音的能量比,按声音的每个频带设定。当然,反射率也可以不取决于声音的频带而一律地设定。在三维声场为开放

空间的情况下,例如也可以使用一律地设定的衰减率、衍射音或初期反射音等参数。

[0264] 在上述说明中,作为元数据中包含的与障碍物对象或音源对象有关的参数而列出了反射率,但元数据也可以包含反射率以外的信息。例如,作为与音源对象及非发声源对象这两者有关的元数据,也可以包含与对象的材料有关的信息。具体而言,元数据也可以包含扩散率、透射率或吸音率等参数。

[0265] 作为与音源对象有关的信息,也可以包含音量、放射特性(指向性)、再现条件、从一个对象发出的音源的数量和种类、或指定对象中的音源区域的信息等。再现条件中例如也可以设定是持续地流出的声音还是事件触发的声音。对象中的音源区域既可以通过用户99的位置与对象的位置的相对关系来设定,也可以以对象为基准来设定。在通过用户99的位置与对象的位置的相对关系来设定的情况下,以用户99看对象的面为基准,能够使用户99感知为,从用户99看时从对象的右侧发出声音X、从左侧发出声音Y。在以对象为基准来设定的情况下,不论用户99看的方向如何都可以固定从对象的哪个区域发出哪个声音。例如,可以使用户99感知为,当从正面看对象时从右侧流出高音,从左侧流出低音。在此情况下,如果用户99绕到对象的背面,能够使用户99感知为,从背面看从右侧流出低音、从左侧流出高音。

[0266] 作为与空间有关的元数据,可以包括到初期反射音为止的时间、混响时间、直接音与扩散音的比率等。在直接音与扩散音的比率为零的情况下,能够使用户99仅感知直接音。

[0267] 此外,也可以将表示三维声场中的用户99的位置及朝向的信息作为初始设定预先作为元数据包含在比特流中,也可以不包含在比特流中。在表示用户99的位置及朝向的信息没有包含在比特流中的情况下,从比特流以外的信息取得表示用户99的位置及朝向的信息。例如,如果是VR空间中的用户99的位置信息则可以从提供VR内容的应用取得,如果是用来提示声音作为AR的用户99的位置信息则使用例如便携终端利用GPS、相机或LiDAR(Laser Imaging Detection and Ranging)等实施自位置推测而得到的位置信息。另外,声音信号和元数据既可以保存在一个比特流中,也可以分别保存在多个比特流中。同样,声音信号和元数据既可以保存在一个文件中,也可以分别保存在多个文件中。

[0268] 在声音信号和元数据分别保存在多个比特流中的情况下,也可以在保存有声音信号和元数据的多个比特流中的一个或一部分的比特流中包含表示关联的其他比特流的信息。此外,也可以在保存有声音信号和元数据的多个比特流的各比特流的元数据或控制信息中包含表示关联的其他比特流的信息。在声音信号和元数据分别保存在多个文件中的情况下,也可以在保存有声音信号和元数据的多个文件中的一个或一部分的文件中包含表示关联的其他的比特流或文件的信息。此外,也可以在保存有声音信号和元数据的多个比特流的各比特流的元数据或控制信息中包含表示关联的其他的比特流或文件的信息。

[0269] 这里,关联的比特流或文件分别例如是在音响处理时有可能同时被使用的比特流或文件。此外,表示关联的其他比特流的信息既可以一起记述在保存有声音信号和元数据的多个比特流中的一个比特流的元数据或控制信息中,也可以分割地记述在保存有声音信号和元数据的多个比特流中的两个以上的比特流的元数据或控制信息中。同样,表示关联的其他的比特流或文件的信息既可以一起记述在保存有声音信号和元数据的多个文件中的一个文件的元数据或控制信息中,也可以分割地记述在保存有声音信号和元数据的多个文件中的两个以上的文件的元数据或控制信息中。此外,一起记述有表示关联的其他的比

特流或文件的信息的控制文件也可以与保存有声音信号和元数据的多个文件分开而生成。此时,控制文件也可以不保存声音信号和元数据。

[0270] 这里,表示关联的其他的比特流或文件的信息,例如是表示该其他的比特流的识别符、表示其他文件的文件名、URL (Uniform Resource Locator:统一资源定位符) 或URI (Uniform Resource Identifier:统一资源标识符) 等。在此情况下,取得部111基于表示关联的其他的比特流或文件的信息,确定或取得比特流或文件。此外,也可以是表示关联的其他比特流的信息包含在保存有声音信号和元数据的多个比特流中的至少一部分比特流的元数据或控制信息中,并且表示关联的其他文件的信息包含在保存有声音信号和元数据的多个文件中的至少一部分文件的元数据或控制信息中。这里,包含表示关联的比特流或文件的信息的文件,例如是用于分发内容的清单文件等的控制文件。

[0271] 工业实用性

[0272] 本公开在使用户感知立体的声音等的音响再现时具有实用性。

[0273] 标号说明

[0274] 99 用户

[0275] 100 音响再现系统

[0276] 101 信息处理装置

[0277] 102 通信模块

[0278] 103 检测器

[0279] 104 驱动器

[0280] 111 取得部

[0281] 112 编码声音信息输入部

[0282] 113 解码处理部

[0283] 114 感测信息输入部

[0284] 121、121a处理部

[0285] 122 判定部

[0286] 123 存储部

[0287] 124 执行部

[0288] 125 计算部

[0289] 141 信号输出部

[0290] 200 立体影像再现装置

[0291] 600 立体音响再现系统

[0292] 601 信息处理装置

[0293] 602 声音提示设备

[0294] 700、900 编码装置

[0295] 701、803 输入数据

[0296] 702 编码器

[0297] 703 编码数据

[0298] 704、804 存储器

[0299] 800、1000 解码装置

- [0300] 801 声音信号
- [0301] 802、1100、1200 解码器
- [0302] 901 发送部
- [0303] 902 发送信号
- [0304] 1001 接收部
- [0305] 1002 接收信号
- [0306] 1101、1201 空间信息管理部
- [0307] 1102 声音数据解码器
- [0308] 1103、1202 渲染部

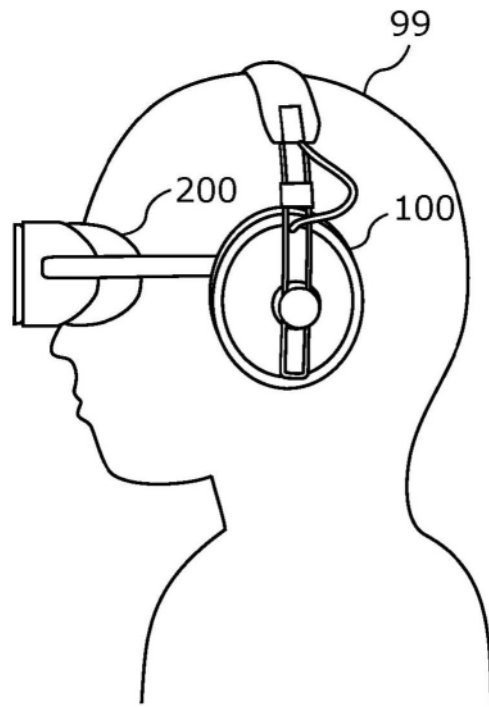


图1

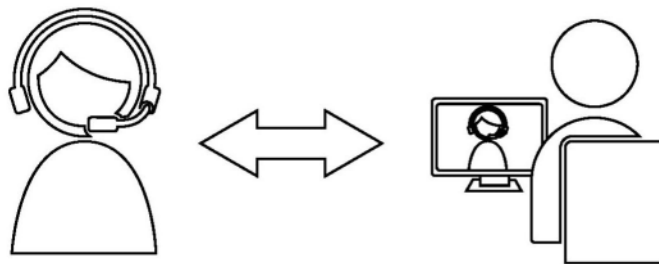


图2A

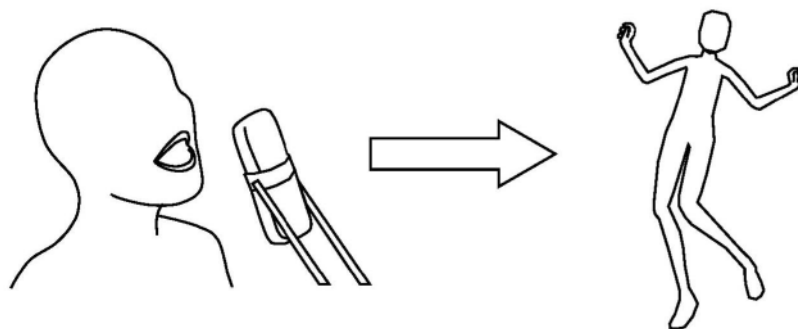


图2B



图3

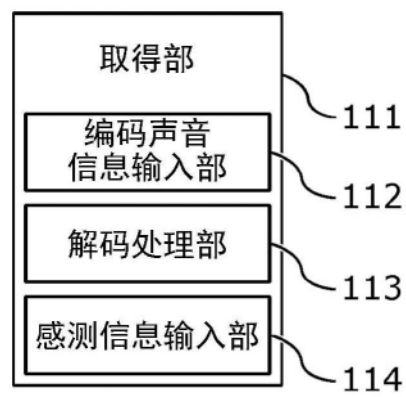


图4

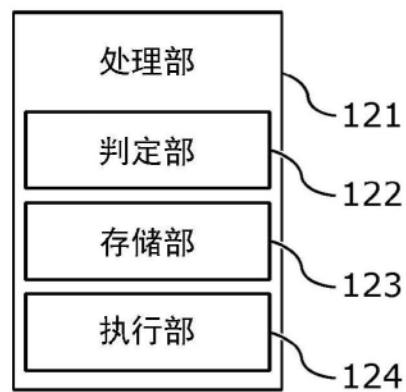


图5

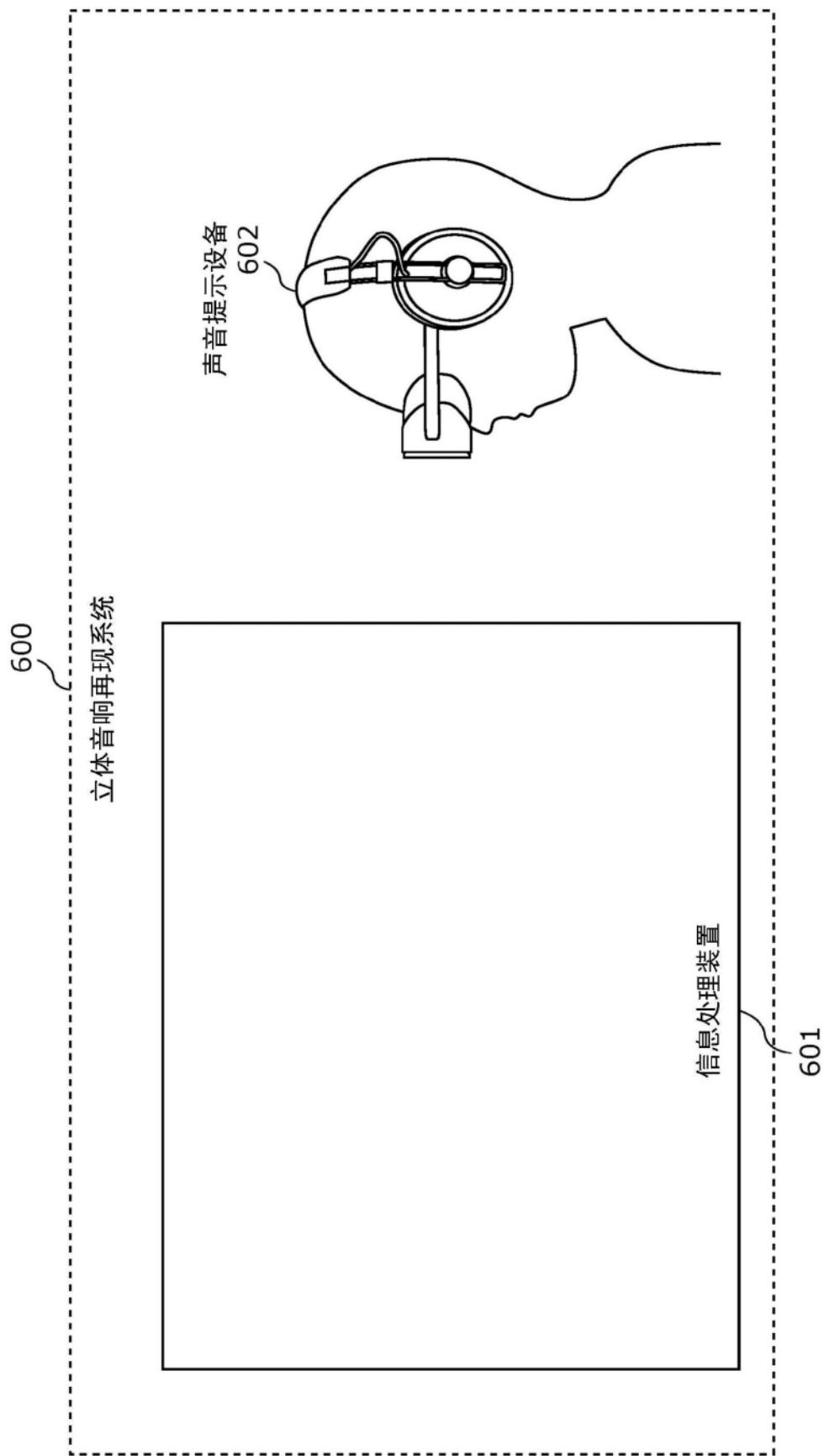


图6

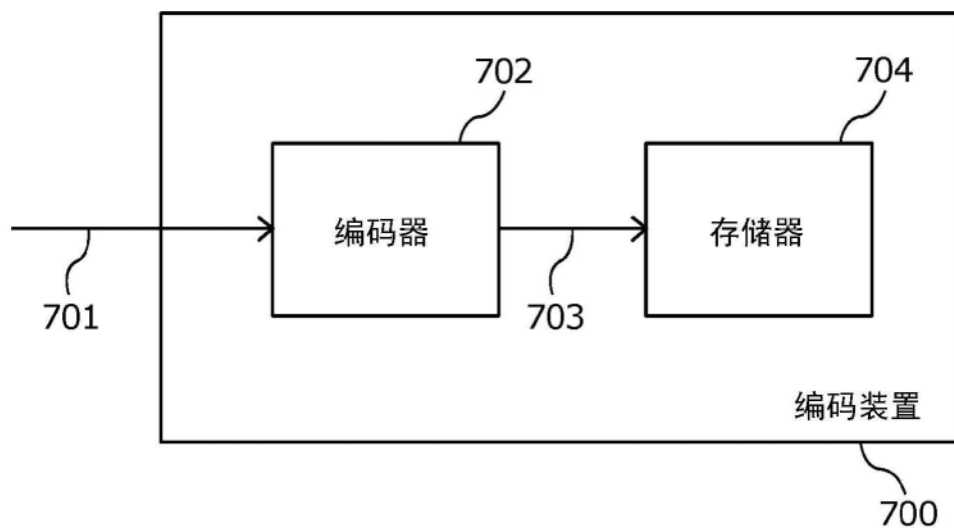


图7

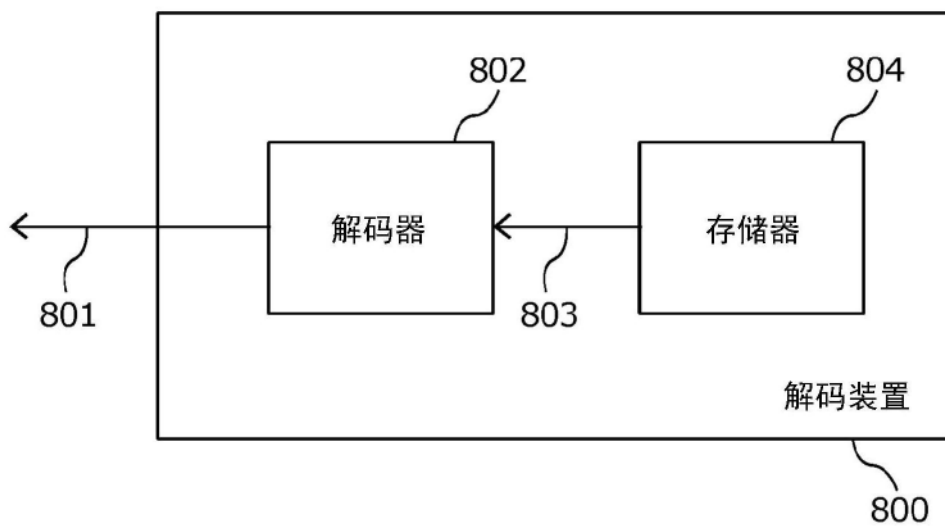


图8

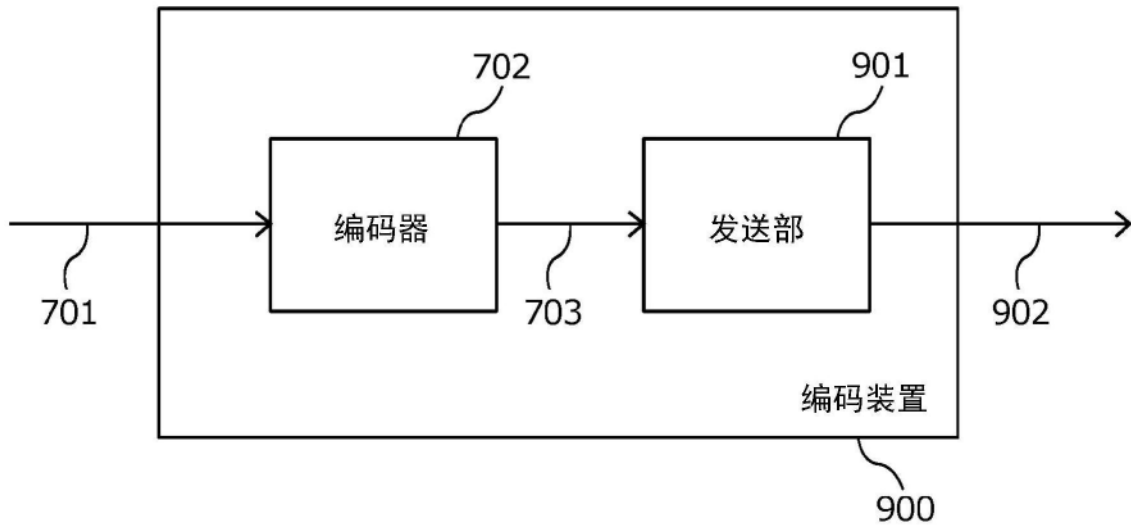


图9

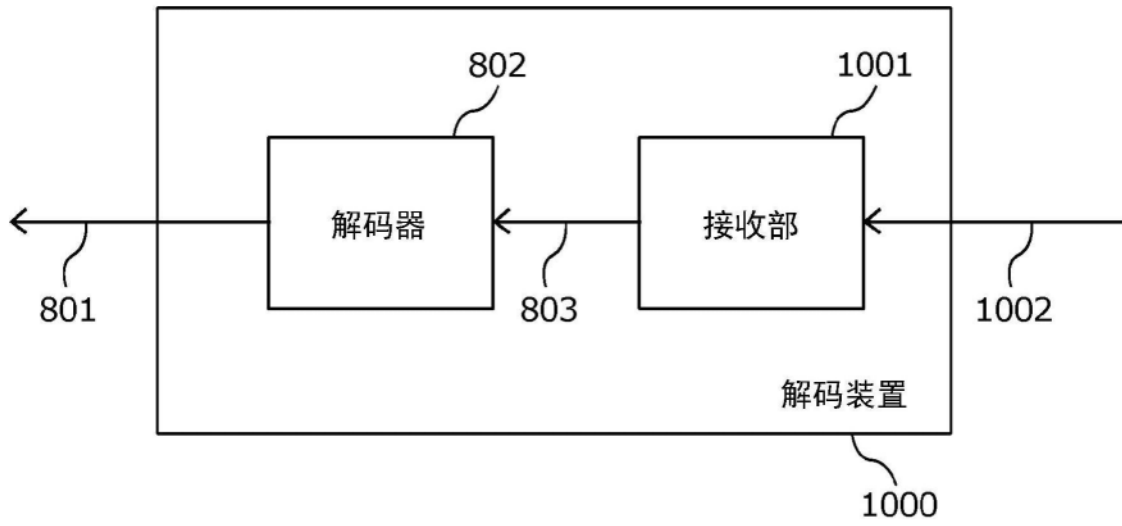


图10

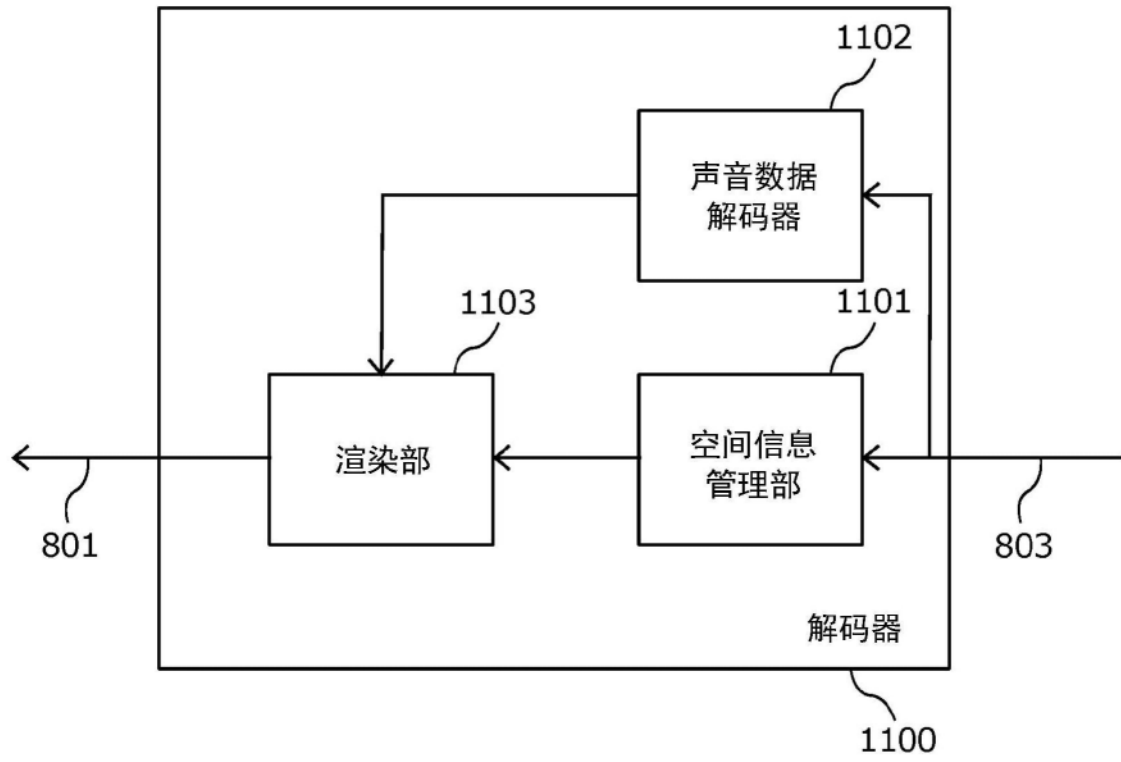


图11

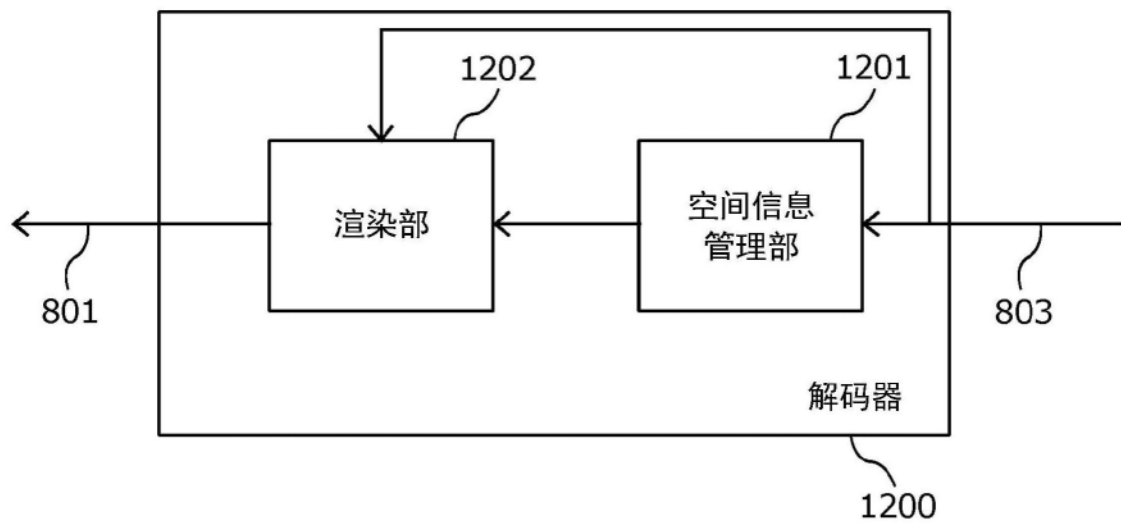


图12

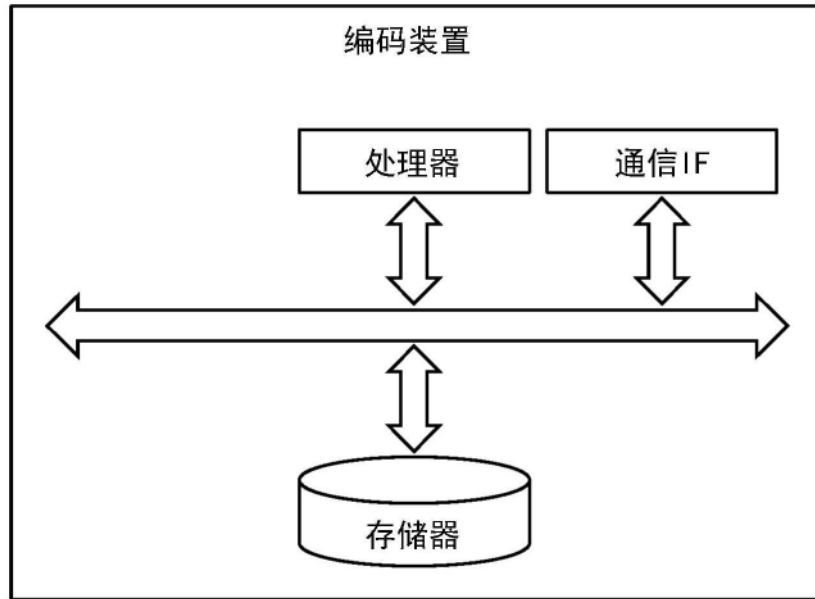


图13

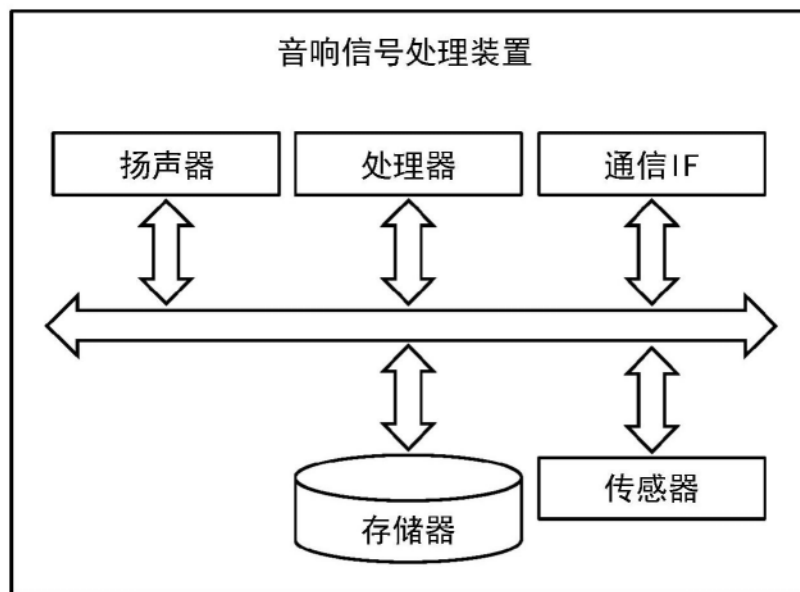


图14

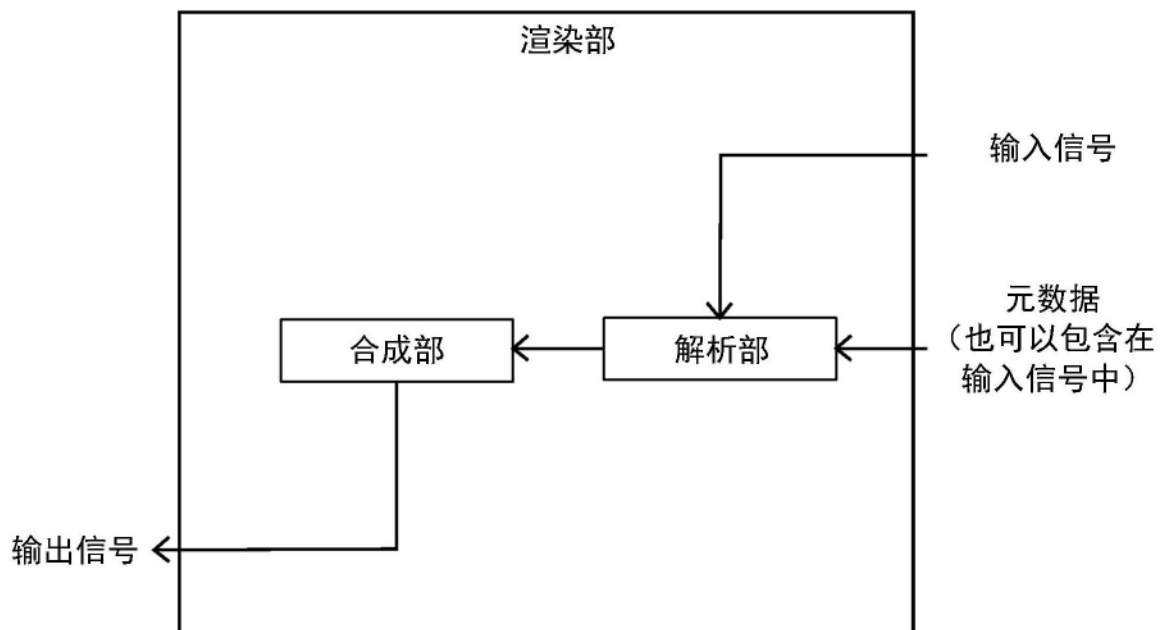


图15

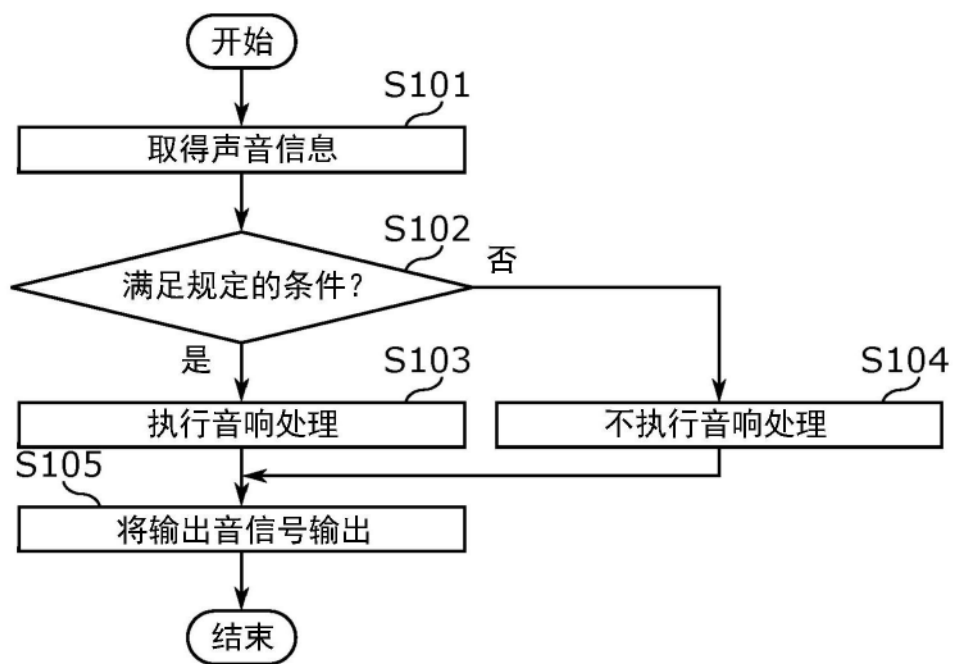


图16

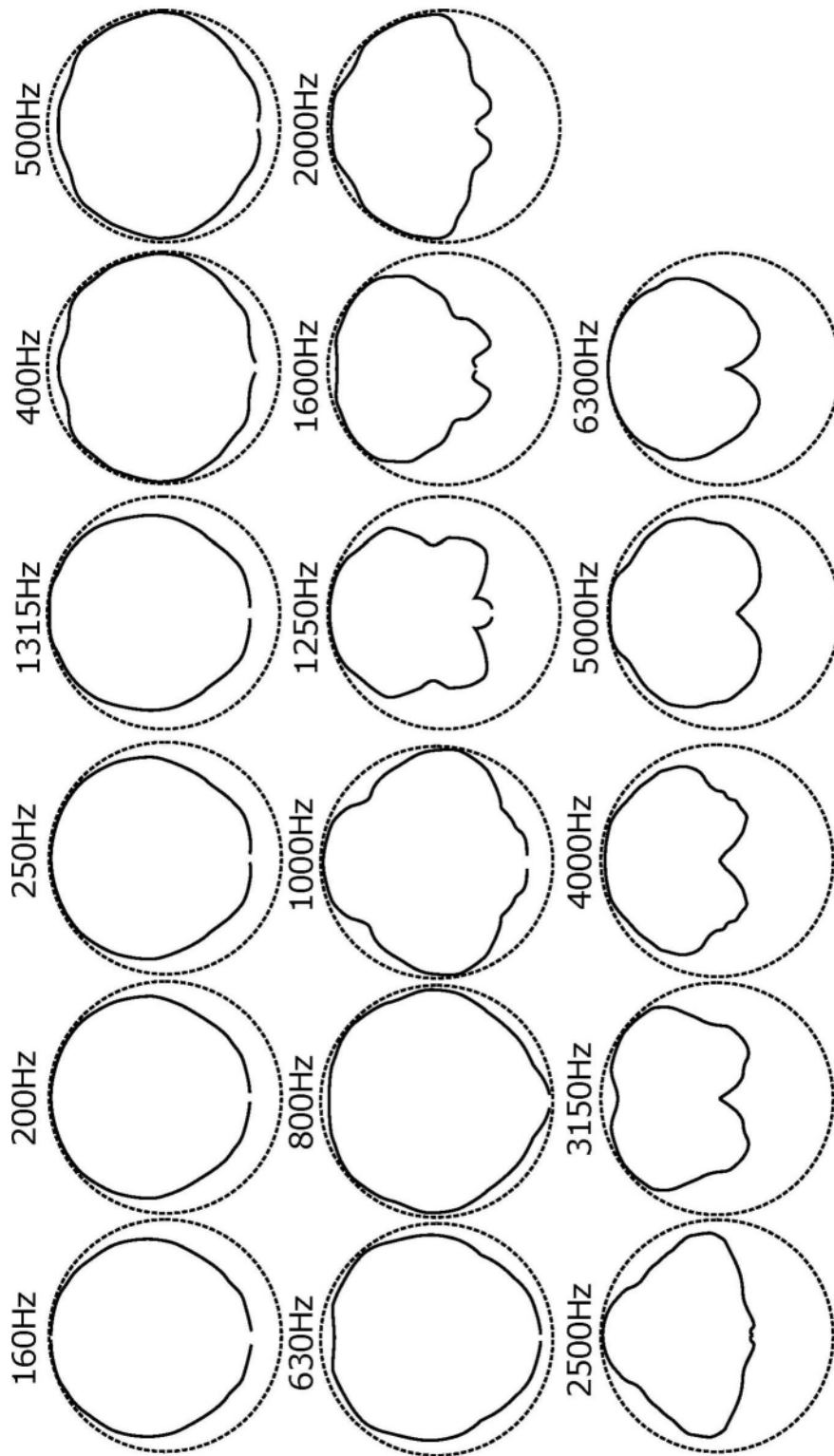


图17

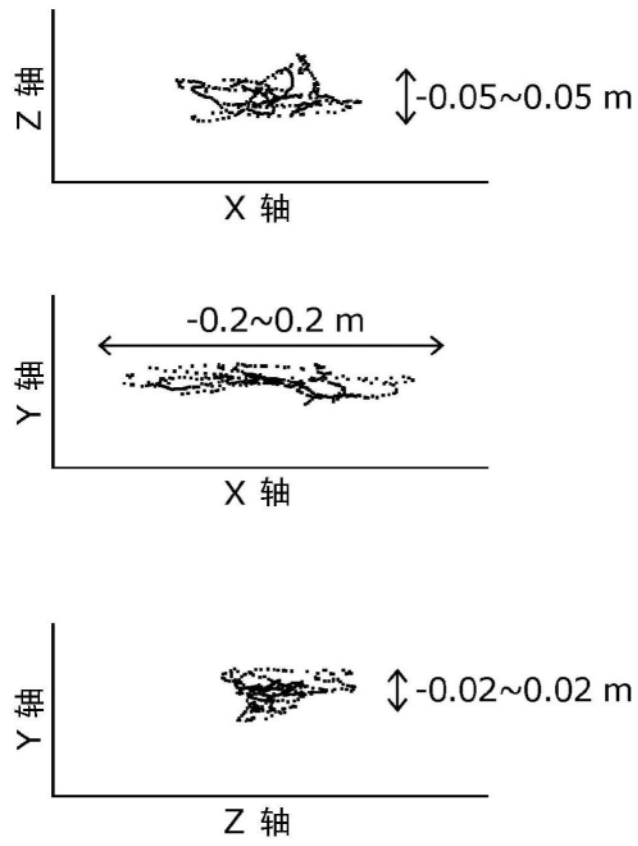


图18

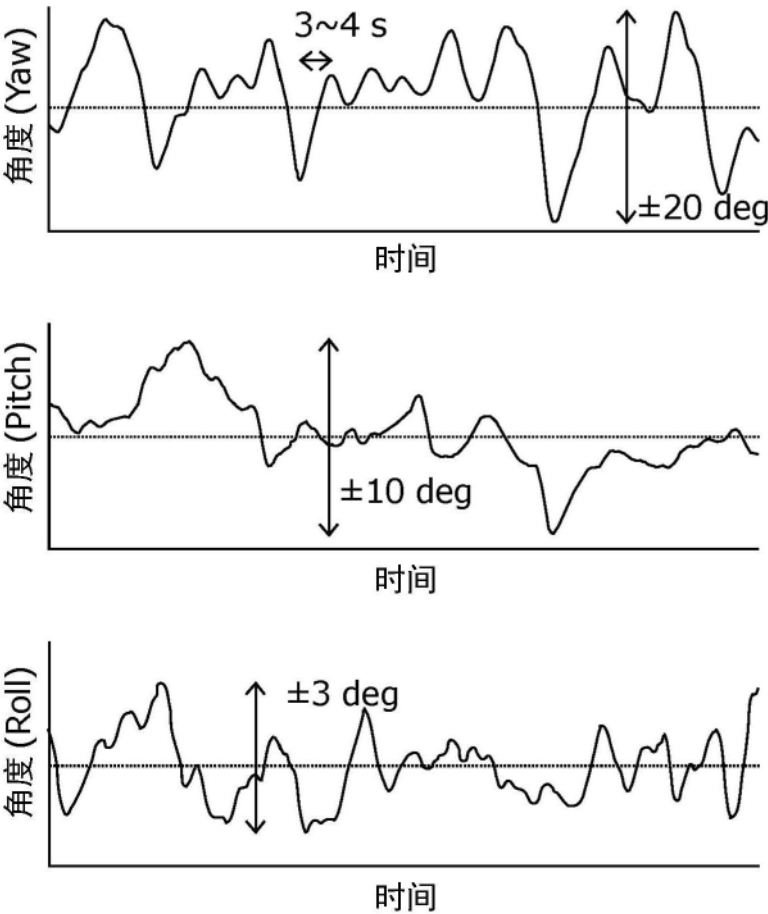


图19

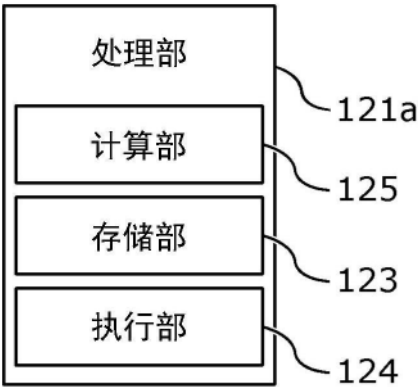


图20

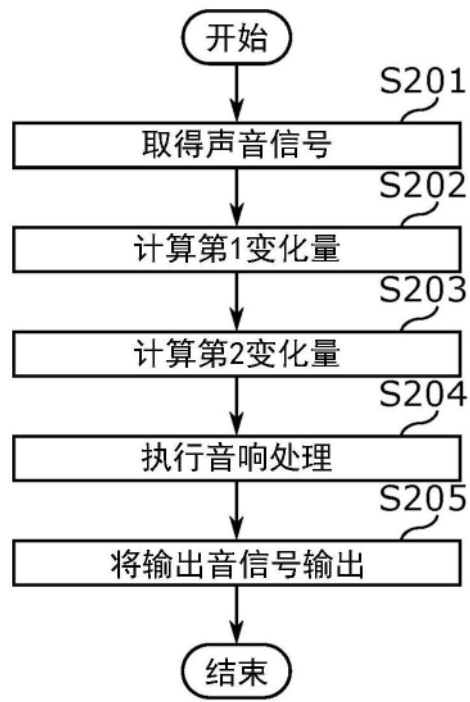


图21