



(12) 发明专利申请

(10) 申请公布号 CN 116438812 A

(43) 申请公布日 2023. 07. 14

(21) 申请号 202180076300.6

(22) 申请日 2021.11.25

(30) 优先权数据

2020-204354 2020.12.09 JP

(85) PCT国际申请进入国家阶段日

2023.05.11

(86) PCT国际申请的申请数据

PCT/JP2021/043098 2021.11.25

(87) PCT国际申请的公布数据

W02022/124084 JA 2022.06.16

(71) 申请人 索尼集团公司

地址 日本东京

(72) 发明人 中井彬人 户村朝子 远藤泰己

小泉雅彦 菅井千寻

(74) 专利代理机构 北京康信知识产权代理有限

责任公司 11240

专利代理师 刘瑞贤

(51) Int.Cl.

H04S 7/00 (2006.01)

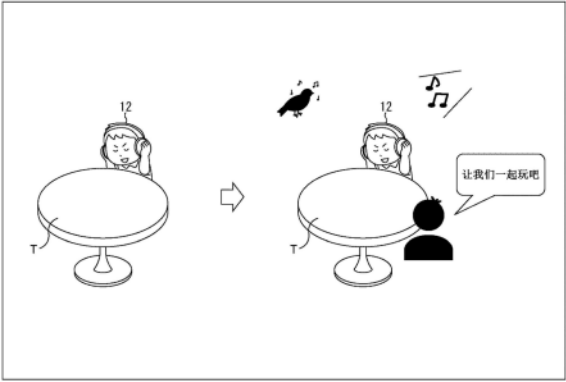
权利要求书3页 说明书24页 附图29页

(54) 发明名称

再现装置、再现方法、信息处理装置、信息处理方法及程序

(57) 摘要

本技术涉及用于在减少处理负荷的同时提供逼真内容的再现装置、再现方法、信息处理装置、信息处理方法以及程序。根据本技术的第一方面的再现装置包括再现单元，再现单元再现对象音频内容，对象音频内容包括通过使用个性化头部相关传输函数 (HRTF) 再现的第一对象和通过使用共用HRTF再现的第二对象，其中，个性化HRTF是针对收听者个性化的HRTF，共用HRTF是在多个收听者之间共享的HRTF。本技术可应用于再现包括音频数据的内容的装置。



1. 一种再现装置,包括:

再现单元,被配置为再现对象音频内容,所述对象音频内容包括通过使用个性化HRTF再现的第一对象和通过使用共用HRTF再现的第二对象,其中,所述个性化HRTF是针对收听者个性化的HRTF,所述共用HRTF是在多个收听者之间共享的HRTF。

2. 根据权利要求1所述的再现装置,进一步包括:

HRTF获取单元被配置为基于收听者的耳朵的图像以及关于测量所述共用HRTF时的扬声器的布局的布局信息获取所述个性化

HRTF,所述布局信息包括在所述对象音频内容中。

3. 根据权利要求1所述的再现装置,其中,关于所述对象音频内容的传输数据包括关于所述第一对象的音频数据以及通过对关于所述第二对象的音频数据执行使用所述共用HRTF的处理而在所述对象音频内容的传输源装置中生成的基于声道的数据。

4. 根据权利要求3所述的再现装置,其中,所述再现单元输出与通过对所述第一对象的音频数据执行使用所述个性化HRTF的处理而生成的基于声道的数据和包括在所述传输数据中的所述基于声道的数据相对应的声音。

5. 根据权利要求1所述的再现装置,其中,所述再现单元对所述第一对象的音频数据执行使用所述个性化HRTF的处理,对所述第二对象的音频数据执行使用所述共用HRTF的处理,并输出与通过对所述第一对象的音频数据执行处理而生成的基于声道的数据相对应的声音和与通过对所述第二对象的音频数据执行处理而生成的基于声道的数据相对应的声音。

6. 根据权利要求3所述的再现装置,进一步包括:传输数据获取单元,所述传输数据获取单元被配置为从为所述传输源装置准备的并且因所述第一对象的数量变化的多个所述传输数据中获取要再现的所述传输数据。

7. 一种再现方法,使再现装置执行以下操作:

再现对象音频内容,所述对象音频内容包括通过使用个性化HRTF再现的第一对象和通过使用共用HRTF再现的第二对象,其中,所述个性化HRTF是针对收听者个性化的HRTF,所述共用HRTF是在多个收听者之间共享的HRTF。

8. 一种使计算机执行以下处理的程序:

再现对象音频内容,所述对象音频内容包括通过使用个性化HRTF再现的第一对象和通过使用共用HRTF再现的第二对象,其中,所述个性化HRTF是针对收听者个性化的HRTF,所述共用HRTF是在多个收听者之间共享的HRTF。

9. 一种信息处理装置,包括:

元数据生成单元,被配置为生成包括标志的元数据,其中,所述标志指示对象是将被再现为通过使用个性化HRTF再现的第一对象还是通过使用共用HRTF再现的第二对象,所述个性化HRTF是针对收听者个性化的HRTF,所述共用HRTF是在多个收听者之间共享的HRTF;以及

内容生成单元,被配置为生成包括关于多个对象的声源数据和所述元数据的对象音频内容。

10. 根据权利要求9所述的信息处理装置,其中,所述元数据生成单元生成进一步包括关于测量所述共用HRTF时的扬声器的布局的布局信息的所述元数据。

11. 根据权利要求9所述的信息处理装置, 其中, 所述元数据生成单元生成包括关于每个对象的音像的定位的重要性的信息作为所述标志的所述元数据。

12. 根据权利要求11所述的信息处理装置, 其中, 基于对象类型、从收听位置到所述对象的位置的距离、在收听位置处的所述对象的方向以及所述对象的高度中的至少一项来设置所述重要性。

13. 一种信息处理方法, 使信息处理装置执行以下:

生成包括标志的元数据, 其中, 所述标志指示对象将被再现为通过使用个性化HRTF再现的第一对象还是通过使用共用HRTF再现的第二对象, 其中, 所述个性化HRTF是针对收听者个性化的HRTF, 所述共用HRTF是在多个收听者之间共享的HRTF; 以及

生成包括关于多个对象的声源数据和所述元数据的对象音频内容。

14. 一种使计算机执行以下处理的程序:

生成包括标志的元数据, 其中, 所述标志指示对象将被再现为通过使用个性化HRTF再现的第一对象还是通过使用共用HRTF再现的第二对象, 其中, 所述个性化HRTF是针对收听者个性化的HRTF, 所述共用HRTF是在多个收听者之间共享的HRTF, 并且生成包括关于多个对象的声源数据和所述元数据的对象音频内容。

15. 一种信息处理装置, 包括:

内容获取单元, 被配置为获取对象音频内容, 所述对象音频内容包括关于多个对象的声源数据和元数据, 所述元数据包括指示对象是被再现为通过使用个性化HRTF再现的第一对象还是通过使用共用HRTF再现的第二对象的标志, 所述个性化HRTF是针对收听者个性化的HRTF, 所述共用HRTF是在多个收听者之间共享的HRTF;

音频处理单元, 被配置为对基于所述标志被选择再现为所述第二对象的对象的音频数据执行使用共用HRTF的处理; 以及

传输数据生成单元, 被配置为生成关于所述对象音频内容的传输数据, 所述传输数据包括通过使用所述共用HRTF的处理而生成的基于声道的数据和关于所述第一对象的音频数据。

16. 根据权利要求15所述的信息处理装置, 其中, 所述传输数据生成单元针对每个对象音频内容生成因所述第一对象的数量变化的多个传输数据。

17. 一种信息处理方法, 使信息处理装置执行以下:

获取对象音频内容, 所述对象音频内容包括关于多个对象的声源数据和元数据, 所述元数据包括指示对象是将被再现为通过使用个性化HRTF再现的第一对象还是通过使用共用HRTF再现的第二对象的标志, 所述个性化HRTF是针对收听者个性化的HRTF, 所述共用HRTF是在多个收听者之间共享的HRTF;

对关于基于标志被选择再现为第二对象的对象的音频数据执行使用共用HRTF的处理; 以及

生成关于所述对象音频内容的传输数据, 所述传输数据包括通过使用所述共用HRTF的处理生成的基于声道的数据和关于所述第一对象的音频数据。

18. 一种使计算机执行以下处理的程序:

获取对象音频内容, 所述对象音频内容包括关于多个对象的声源数据和元数据, 所述元数据包括指示所述对象是将被再现为通过使用个性化HRTF再现的第一对象还是通过使

用共用HRTF再现的第二对象的标志,所述个性化HRTF是针对收听者个性化的HRTF,所述共用HRTF是在多个收听者之间共享的HRTF;

对关于基于所述标志被选择再现为所述第二对象的对象的音频数据执行使用共用HRTF的处理;以及

生成关于所述对象音频内容的传输数据,所述传输数据包括通过使用所述共用HRTF的处理而生成的基于声道的数据和关于所述第一对象的音频数据。

再现装置、再现方法、信息处理装置、信息处理方法及程序

技术领域

[0001] 本技术特别涉及再现装置、再现方法、信息处理装置、信息处理方法以及程序,通过该再现装置、再现方法、信息处理装置、信息处理方法以及程序可以在减少处理负荷的同时提供逼真(realistic,临场感)内容。

背景技术

[0002] 所谓的有声读物作为内容提供,包括讲述者和朗读者对图书的语音讲述的录音。例如,用户通过耳机收听由智能电话再现的有声读物声音。这提供了在不用眼睛追随印刷文本的情况下阅读书籍的感觉。

[0003] [引用列表]

[0004] [专利文献]

[0005] [专利文献1]

[0006] JP 2018-010654 A

[0007] [专利文献2]

[0008] JP 2009-260574 A

发明内容

[0009] [技术问题]

[0010] 在多个角色出现的场景中,仅通过书的语音叙述难以赋予逼真性。

[0011] 可通过利用头部相关传输函数(HRTF)再现例如语音的方向、距离和移动来赋予逼真性。通过使用数学地表达声音如何从声源传播到耳朵的HRTF的操作,可以立体地再现来自耳机的声音。

[0012] 由于HRTF在耳朵的形状之间变化,因此音像的精确定位需要每个用户的HRTF的个性化。针对所有声音使用个性化HRTF导致在内容的生成侧和再现侧的大量处理负荷。

[0013] 鉴于这种情况做出本技术,并且本技术在减少处理负荷的同时提供逼真的内容。

[0014] [问题的解决方案]

[0015] 根据本技术的第一方面的再现装置包括再现单元,再现单元被配置为再现对象音频内容,对象音频内容包括通过使用个性化HRTF再现的第一对象和通过使用共用HRTF再现的第二对象,其中,个性化HRTF是针对收听者个性化的HRTF,共用HRTF是在多个收听者之间共享的HRTF。

[0016] 根据本技术的第二方面的信息处理装置包括:元数据生成单元,被配置为生成包括标志的元数据,其中,标志指示对象将被再现为通过使用个性化HRTF再现的第一对象还是通过使用共用HRTF再现的第二对象,其中,个性化HRTF是针对收听者个性化的HRTF,共用HRTF是在多个收听者之间共享的HRTF;以及内容生成单元,被配置为生成包括关于多个对象的声源数据和元数据的对象音频内容。

[0017] 根据本技术的第三方面的信息处理装置包括:内容获取单元,被配置为获取对象

音频内容,对象音频内容包括关于多个对象的声源数据和元数据,元数据包括标志,标志指示对象是将被再现为通过使用个性化HRTF再现的第一对象还是通过使用共用HRTF再现的第二对象,个性化HRTF是针对收听者个性化的HRTF,共用HRTF是在多个收听者之间共享的HRTF;音频处理单元,被配置为对关于基于标志被选择再现为第二对象的对象的音频数据使用共用HRTF执行处理;传输数据生成单元,被配置为生成包括通过使用共用HRTF的处理而生成的基于声道的数据和关于第一对象的音频数据的内容的传输数据。

[0018] 在本技术的第一方面,再现对象音频内容,对象音频内容包括通过使用个性化HRTF再现的第一对象和通过使用共用HRTF再现的第二对象,其中,个性化HRTF是针对收听者个性化的HRTF,共用HRTF是在多个收听者之间共享的HRTF。

[0019] 在本技术的第二方面,生成元数据,元数据包括指示对象是否将被再现为通过使用个性化HRTF(即,针对收听者个性化的HRTF)再现的第一对象还是通过使用共用HRTF(即,在多个收听者之间共享的HRTF)再现的第二对象的标志,并生成包括关于多个对象的声源数据和元数据的对象音频内容。

[0020] 在本技术的第三方面,获取对象音频内容,对象音频内容包括关于多个对象的声源数据和元数据,元数据包括指示对象是将被再现为通过使用个性化HRTF(即,针对收听者个性化的HRTF)再现的第一对象还是通过使用共用HRTF(即,在多个收听者之间共享的HRTF)再现的第二对象的标志。此外,对关于基于标志被选择再现为第二对象的对象的音频数据执行使用共用HRTF的处理,并生成关于内容的传输数据,传输数据包括通过使用共用HRTF的处理生成的基于声道的数据和关于第一对象的音频数据。

附图说明

[0021] 图1示出了根据本技术的实施方式的内容分发系统的配置实施例。

[0022] 图2示出了用户收听有声读物内容的状态。

[0023] 图3示出了构成有声读物内容的对象的实施例。

[0024] 图4示出了HRTF的使用的实施例。

[0025] 图5示出了音像定位处理的实施例。

[0026] 图6示出了HRTF的测量的实施例。

[0027] 图7示出了当测量HRTF时扬声器的布局的实施例。

[0028] 图8示出了各层的扬声器的布局。

[0029] 图9是示出各层的扬声器的布局的另一示图。

[0030] 图10示出对象的布局的实施例。

[0031] 图11是对象的位置的示意图。

[0032] 图12示出了来自对象的声音的方向。

[0033] 图13示出了收听位置改变的实施例。

[0034] 图14是示出内容制作装置的硬件配置实施例的框图。

[0035] 图15是示出内容制作装置的功能配置实施例的框图。

[0036] 图16示出了有声读物内容的配置实施例。

[0037] 图17是示出内容管理装置的功能配置实施例的框图。

[0038] 图18示出了传输数据的实施例。

- [0039] 图19为示出再现装置的硬件配置实施例的框图；
- [0040] 图20为示出再现装置的功能配置实施例的框图；
- [0041] 图21示出了推断模型的实施例。
- [0042] 图22是用于说明内容制作装置的处理的流程图。
- [0043] 图23是用于说明内容管理装置的处理的流程图。
- [0044] 图24是用于说明再现装置的处理的流程图。
- [0045] 图25是示出内容制作装置的另一个配置实施例的框图。
- [0046] 图26示出了具有收集外部声音的功能的耳机的外观。
- [0047] 图27是示出再现装置的另一个配置实施例的框图；
- [0048] 图28是示出再现装置的另一个配置实施例的框图；
- [0049] 图29是示出内容制作装置的另一个配置实施例的框图。

具体实施方式

[0050] 下面将描述用于实现本技术的实施方式。将按照以下顺序进行描述。

- [0051] 1. 内容分发系统
- [0052] 2. 使用HRTF的音像定位原理
- [0053] 3. 有声读物内容的细节
- [0054] 4. 各装置的配置
- [0055] 5. 各装置的操作
- [0056] 6. 变形例
- [0057] 7. 其他实施例
- [0058] <<内容分发系统>>

[0059] 图1示出了根据本技术的实施方式的内容分发系统的配置实施例。

[0060] 图1中的内容分发系统包括：内容制作装置1，其是内容制作侧的装置；内容管理装置2，其是内容传输侧的装置；以及再现装置11，其是内容再现侧的装置。头戴式耳机12连接至再现装置11。

[0061] 内容管理装置2和再现装置11经由网络21（例如，互联网）彼此连接。内容制作装置1和内容管理装置2可以经由网络21彼此连接。

[0062] 在图1的内容分发系统中分发的内容是所谓的有声读物的内容，在有声读物中录音了图书的语音旁白。有声读物内容包括各种声音，例如效果声音、环境声音、BGM以及旁白和台词的人声。在下文中，有声读物内容的声音将视情况简单地描述为语音，除非需要识别声音的种类。实际上，有声读物内容的声音包括除了语音之外的各种声音。

[0063] 内容制作装置1响应于制作者的操作生成有声读物内容。由内容制作装置1生成的有声读物内容被提供给内容管理装置2。

[0064] 内容管理装置2基于由内容制作装置1生成的有声读物内容生成作为用于内容传输的数据的传输数据。在内容管理装置2中，例如，针对有声读书的内容的每个标题生成具有不同数据量的多个传输数据。内容管理装置2响应于来自再现装置11的请求向再现装置11发送具有预定数据量的传输数据。

[0065] 以类似智能电话的装置配置再现装置11。再现装置11可以用其他装置配置，例如，

平板电脑、PC、以及TV。再现装置11经由网络21接收从内容管理装置2传输的有声读物内容并再现该内容。再现装置11将通过再现有声读物内容获得的音频数据发送给头戴式耳机12并且使头戴式耳机12输出有声读物内容的语音。

[0066] 再现装置11和头戴式耳机12经由无线LAN或者预定标准(例如,蓝牙(注册商标))的通信彼此无线连接。再现装置11和头戴式耳机12可以具有有线连接。在图1的实施例中,用户收听有声读物内容所使用的装置是头戴式耳机12。可以替代地使用诸如扬声器和耳戴式耳机(内耳戴式耳机)的其他装置。

[0067] 为每个用户(有声读物内容的收听者)准备作为再现侧上的装置的再现装置11和头戴式耳机12。

[0068] 图2示出了用户收听有声读物内容的状态。

[0069] 在图2的左侧实施例中,戴有头戴式耳机12的用户在桌子T周围收听有声书籍内容。构成有声读物内容的音频数据包括对象音频数据。使用对象音频能够将角色的音声等的音像定位到由位置信息指定的位置。

[0070] 根据制作侧的意图控制每个对象的音像的位置,允许用户感受位于用户周围的音像,如图2的右侧所示。在图2的右侧实施例中,用户感觉好像从左前方听到说“让我们一起玩吧”的角色的音声并且从上方听到小鸟的歌声。此外,用户感觉好像从远处听到BGM。

[0071] 如上所述,在图1中的内容分发系统中分发的有声读物内容是声音可以再现为立体声音的内容。

[0072] 与收听未经立体声音处理的声音相比,用户可获得临场性。用户还可以感觉沉浸在有声读物内容中。

[0073] 立体声音可通过使用HRTF(头部相关传递函数)来实现,HRTF数学地表达声音如何从声源传播到耳朵。在再现装置11中执行使用HRTF的操作,使得音像定位在由制作侧设置的位置处,并且从每个位置听到每个对象的声音。

[0074] 图3示出了构成有声读物内容的对象的实施例。

[0075] 有声读物内容配置有关于多个对象(音频对象)的音频数据,多个对象包括通过使用个性化HRTF再现的对象和通过使用共用HRTF再现的对象。

[0076] 个性化HRTF是针对作为有声读物内容的收听者的用户的耳朵形状优化(个性化)的HRTF。通常,HRTF在收听者的耳朵形状和头部形状等之间而不同。个性化HRTF是表示每个用户的不同传输特性的函数。

[0077] 共用HRTF是未经优化的HRTF。在多个收听者之间共享相同的HRTF。

[0078] 通过使用个性化HRTF再现的声音允许以比通过使用共用HRTF再现的声音更高的精度定位音像,从而实现具有增强的临场性的声音。相反,通过使用共用HRTF再现的声音允许以比通过使用个性化HRTF再现的声音更低的精确度定位音像,从而实现具有粗略定位音像的声音。

[0079] 在图3的实施例中,对象#1和#2是通过使用个性化HRTF再现的个性化HRTF再现对象,并且#3和#4是通过使用共用HRTF再现的共用HRTF再现对象。

[0080] 图4示出了HRTF的使用的实施例。

[0081] 如在图4的A和B中所示,个性化HRTF用于再现需要高度定位的诸如人物台词和鸟类唱歌的对象。

[0082] 如图4的C所示,共用HRTF用于再现不需要高度定位的诸如BGM的对象。

[0083] 根据关于定位的重要性的信息切换是将对象用作个性化HRTF再现对象还是共用HRTF再现对象,所述信息由制作者设置。将设置为具有定位的重要性的对象再现为个性化HRTF再现对象。

[0084] 如上所述,在再现装置11中再现的有声读物内容是包括个性化HRTF再现对象和共用HRTF再现对象的对象音频内容。稍后将描述音频内容的数据结构的细节。

[0085] 再现装置11是用于再现对象音频内容的装置,对象音频内容包括用作第一对象的个性化HRTF再现对象和用作第二对象的共用HRTF再现对象。内容的再现在于基于对象音频数据执行用于输出每个对象声音的对象音频处理,并执行对象音频处理的至少一部分。

[0086] <<使用HRTF进行音像定位的原理>>

[0087] HRTF是传递函数,该传递函数包括由对从声源传播到鼓膜的声波反射或衍射造成的物理影响,并且由空间中不存在收听者的头部的点处的声压与如果收听者的头部的中心叠加在该相同点时传播到收听者的左鼓膜和右鼓膜的位置的声压的比率定义。HRTF叠加在听觉信号上,并且模拟声波在空间中从声源传播并到达收听者的鼓膜时声音的改变,从而定位收听者的音像。

[0088] 图5示出了音像定位处理的实施例。

[0089] 例如,在使用对象位置信息进行渲染之后音频信号被输入到卷积单元31。对象音频数据包括关于对象的位置的位置信息以及关于对象的声源数据。音频信号是包括语音、环境声音和BGM的各种声音的信号。音频信号包括作为用于左耳的信号的视频信号L和作为用于右耳的信号的视频信号R。

[0090] 卷积单元31处理输入的音频信号,使得从在图5的右侧部分中由虚线表示的左虚拟扬声器VSL和右虚拟扬声器VSR的位置听到对象声音。换言之,卷积单元31定位从头戴式耳机12输出的声音的音像,使得用户U将音像感知为来自左虚拟扬声器VSL和右虚拟扬声器VSR的声音。

[0091] 当左虚拟扬声器VSL和右虚拟扬声器VSR彼此不区分时,扬声器统称为虚拟扬声器VS。在图5的实施例中,虚拟扬声器VS位于用户U的前方,并且虚拟扬声器的数量是两个。

[0092] 卷积单元31对音频信号执行用于输出声音的音像定位,并且在音像定位之后将音频信号L和音频信号R分别输出至头戴式耳机12的左侧单元12L和右侧单元12R。构成头戴式耳机12的左侧单元12L附接至用户U的左耳,并且右侧单元12R附接至右耳。

[0093] 图6示出了HRTF的测量的实施例。

[0094] 在预定的参考环境中,虚拟头部DH的位置被设置为收听者的位置。麦克风安装在虚拟头部DH的左耳部分和右耳部分中。左真实扬声器SPL和右真实扬声器SPR被安装在音像将被定位的左虚拟扬声器和右虚拟扬声器的位置处。真实扬声器是指实际安装的扬声器。

[0095] 在虚拟头部DH的耳部收集从左真实扬声器SPL和右真实扬声器SPR输出的声音,并且提前测量传递函数作为HRTF,该传递函数表示当从左真实扬声器SPL和右真实扬声器SPR输出的声音到达虚拟头部DH的耳部时的声音的特性的变化。可针对实际就座的人测量传递函数,其中麦克风放置在人的耳朵附近,而不使用虚拟头部DH。

[0096] 如图6所示,假设从左真实扬声器SPL至虚拟头部DH的左耳的声音传递函数为M11,并且从左真实扬声器SPL至虚拟头部DH的右耳的声音传递函数为M12。此外,假设从右真实

扬声器SPR到虚拟头部DH的左耳的声音传递函数是M21,并且从右真实扬声器SPR到虚拟头部DH的右耳的声音传递函数是M22。

[0097] 图5中的HRTF数据库32存储关于作为提前测量的传递函数的HRTF的信息(关于表示HRTF的系数的信息)。

[0098] 卷积单元31从HRTF数据库32读取并获得与左虚拟扬声器VSL和右虚拟扬声器VSR的位置相对应的HRTF并且设置用于滤波器41至44的HRTF。

[0099] 滤波器41执行用于将传递函数M11应用于音频信号L的滤波并且将经滤波的音频信号L输出至加法单元45。滤波器42执行将传递函数M12应用于音频信号L的滤波,并将滤波的音频信号L输出至加法单元46。

[0100] 滤波器43执行用于将传递函数M21应用于音频信号R的滤波,并且将滤波的音频信号R输出至加法单元45。滤波器44执行用于将传递函数M22应用于音频信号R的滤波,并且将滤波的音频信号R输出至加法单元46。

[0101] 作为用于左声道的加法单元的加法单元45将由滤波器41滤波的音频信号L和由滤波器43滤波的音频信号R相加并且输出相加之后的音频信号。相加之后的音频信号传输至头戴式耳机12,并且从头戴式耳机12的左侧单元12L输出与音频信号对应的声音。

[0102] 作为用于右声道的加法单元的加法单元46将由滤波器42滤波的音频信号L和由滤波器44滤波的音频信号R相加,并输出相加之后的音频信号。相加之后的音频信号被发送到头戴式耳机12,并且与音频信号对应的声音从头戴式耳机12的右侧单元12R输出。

[0103] 以这种方式,卷积单元31通过使用与音像将被定位的位置(对象位置)对应的HRTF对音频信号执行卷积,并且定位来自头戴式耳机12的声音的音像,使得用户U将音像感知为从虚拟扬声器VS发出的声音。

[0104] <<有声读物内容的细节>>

[0105] <HRTF的测量>

[0106] 图7示出了当测量针对有声读物内容的HRTF时的扬声器(speaker)的布局的实施例。

[0107] 如图7所示,扬声器相对于测量者的位置实际设置在不同高度且不同方向上的不同位置。在这种状态下,测量与扬声器的位置相对应的HRTF。小圆圈表示扬声器的位置。

[0108] 在图7的实施例中,设置三个层,即,高HIGH层、中MID层和低LOW层,并且扬声器被设置在不同高度的各个层中。例如,MID层被设置在与测量者的耳朵相同的高度处。HIGH层被设置为高于测量者的耳朵。LOW层被设置为低于测量者的耳朵。

[0109] MID层包括MID_f层和设置在MID_f层内部的MID_c层。LOW层包括LOW_f层和设置在LOW_f层内部的LOW_c层。

[0110] 图8和图9示出了各层的扬声器的布局。环绕数字的圆圈表示各层的扬声器的位置。在图8和图9中,向上方向是测量者的正面方向。

[0111] 图8的A是示出HIGH层的扬声器的布局的顶视图。HIGH层的扬声器是相对于测量者的位置设置在半径为200cm的圆的圆周上的四个扬声器:扬声器TpFL、扬声器TpFR、扬声器TpL以及扬声器TpR。

[0112] 例如,扬声器TpFL被设置在相对于测量者的正面在左侧30°的位置处。扬声器TpFR被设置在相对于测量者的正面在右侧30°的位置处。

[0113] 图8的B是示出MID层的扬声器的布局的顶视图。

[0114] MID_f层的扬声器是相对于测量者的位置设置在半径为200cm的圆的圆周上的五个扬声器。MID_c层的扬声器是相对于测量者的位置设置在半径为40cm的圆的圆周上的五个扬声器。

[0115] 例如,MID_f层的扬声器FL和MID_c层的扬声器FL被设置在相对于测量者的正面在左侧30°的位置处。MID_f层的扬声器FR和MID_c层的扬声器FR被设置在相对于测量者的正面在右侧30°的位置处。相对于测量者的正面设置在相同方向上的五个扬声器构成MID_f层和MID_c层中的扬声器。

[0116] 如图9的A和B中所示,在LOW_f层和LOW_c层中扬声器被设置在相对于测量者的位置具有不同半径的两个圆的圆周上。

[0117] 基于从在此布局中设置的扬声器输出的声音来测量扬声器的位置的HRTF。通过在测量者的耳朵或虚拟头部的耳朵的位置处测量来自位置的脉冲响应并且在频率轴上表示脉冲响应,来确定与扬声器位置(声源位置)相对应的HRTF。

[0118] 使用虚拟头部测量的HRTF可用作共用HRTF。可替代地,在图7中的测量者的位置处可实际就座许多人并且在各个耳朵的位置处测量的HRTF的平均值可用作共用HRTF。

[0119] 使用以这种方式测量的HRTF的再现允许在方向、高度和距离上变化的不同位置处定位对象的音像。收听有声读物内容的用户可能感到沉浸在故事的世界中。

[0120] <对象的实施例>

[0121] 图10示出对象的布局的实施例。

[0122] 如图10所示,对象被布置在桌子T上的位置P1和P2以及桌子T周围的位置P3至P6处。用于再现图10中的对象的声音的音频数据包括在有声读物内容中。

[0123] 在图10的实施例中,在位置P1处的对象是积木的声音,并且在位置P2处的对象是茶倒入杯中的声音。位置P3处的对象是男孩的台词,位置P4处的对象是女孩的台词。位置P5处的对象是说人类词语的鸡的台词,位置P6处的对象是鸟唱歌。位置P6是位置P3上方的位置。

[0124] 位置P11-P14即为作为收听者的用户的位置。例如,在位置P11处的用户从后面的左侧听到在位置P3处的男孩的台词,并且从左边听到在位置P4处的女孩的台词。

[0125] 在图10中布置在用户附近的对象被再现为例如个性化HRTF再现对象。除了图10中的对象之外的诸如环境声音和BGM的对象被再现为共用HRTF再现对象。环境声音包括奔流小溪的潺潺声和风中颤抖的叶子的声音。

[0126] 图11是个性化HRTF再现对象的位置的示意图。

[0127] 在图11的实施例中,假设用户A至D位于桌子T周围的位置P11至P14处。圆圈表示参照图10描述的对象。

[0128] 例如,通过使用与LOW_c层的扬声器的位置相对应的HRTF再现位置P1和P2处的对象。通过使用与MID_c层的扬声器的位置相对应的HRTF再现位置P3、P4和P5处的对象。通过使用与HIGH层的扬声器的位置相对应的HRTF再现位置P6处的对象。位置P7用作表达位置P2处的积木对象的移动时的音像的位置。

[0129] 对象被布置在被设置为绝对位置的位置P1至P7。具体地,在不同的收听位置处,如在图12中所示在不同的方向上听到对象的声音。

[0130] 例如,在图12的实施例中,在位置P1处的积木的声音被用户A从左远方听到,并且被用户B从左近方听到。声音由用户C从右近方听到并且由用户D从右远方听到。

[0131] 通过使用参考图7所描述的扬声器布局来测量各个收听位置处的HRTF,从而固定对象的绝对位置。

[0132] 由于对象的绝对位置是固定的,如图13所示,用户可以在有声读物内容的再现期间改变收听位置。

[0133] 在图13的实施例中,在位置P11处收听有声读物内容的用户A移动至位置P12至P14中的另一个位置。每当用户A改变收听位置时,用于再现对象的HRTF被切换。

[0134] 通过改变收听位置,用户可以移动靠近喜爱的配音者的语音或者靠近语音以确认语音的内容。此外,通过改变收听位置,用户可以在以不同的方式收听内容的同时欣赏相同的有声读物内容。

[0135] <<每个装置的配置>>

[0136] <内容制作装置1的配置>

[0137] 图14是示出内容制作装置1的配置实施例的框图。

[0138] 使用计算机配置内容制作装置1。可用如图14中所示配置的单个计算机或者多个计算机配置内容制作装置1。

[0139] CPU 101、ROM 102和RAM 103经由总线104彼此连接。

[0140] 输入/输出接口105进一步连接至总线104。包括键盘和鼠标的输入单元106以及包括显示器和扬声器的输出单元107连接至输入/输出接口105。制作者使用输入单元106执行用于制作有声读物内容的各种操作。

[0141] 此外,包括硬盘和非易失性存储器的存储单元108、包括网络接口的通信单元109、以及驱动可移除介质111的驱动器110连接至输入/输出接口105。

[0142] 内容管理装置2具有与图14所示的内容制作装置1相同的配置。在下文中,适当地引用图14的配置作为内容管理装置2的配置。

[0143] 图15是示出内容制作装置1的功能配置实施例的框图。

[0144] 在内容制作装置1中,实现内容生成单元121、对象声源数据生成单元122和元数据生成单元123。通过使图14的CPU 101执行预定程序来实现图15的至少一些功能单元。

[0145] 内容生成单元121通过链接由对象声源数据生成单元122生成的对象声源数据和由元数据生成单元123生成的对象元数据生成有声读物内容。

[0146] 图16示出了有声读物内容的配置实施例。

[0147] 如图16所示,有声读物内容包括对象声源数据和对象元数据。

[0148] 在图16的实施例中,包括关于五个对象的声源数据(波形数据):对象#1至对象#5,作为对象声源数据。在该实施例中,关于包括积木的声音、茶倒入杯中的声音以及男孩的台词的声音的声源数据作为对象声源数据被包括在有声读物内容中。

[0149] 对象元数据包括位置信息、定位重要性信息、个性化布局信息和其他种类的信息。其他种类的信息包括关于对象的音量(volume)的信息。

[0150] 包括在对象元数据中的位置信息是关于对象的位置的信息。

[0151] 定位重要性信息是指示音像定位的重要性的标志。音像定位的重要性由例如制作者针对每一对象设定。音像定位的重要性被表示为例如从1至10的数值。如稍后将描述的,

基于定位重要性信息选择每个对象是将被再现为个性化HRTF再现对象还是共用HRTF再现对象。

[0152] 个性化布局信息是关于当测量HRTF时的扬声器布局的信息。参考图7和其他附图描述的扬声器布局由个性化布局信息表达。当获取个性化HRTF时,由再现装置11使用个性化布局信息。

[0153] 图15中的对象声源数据生成单元122生成关于每个对象的声源数据。关于每个对象的声源数据在在演播室等中被录音之后被获取到内容制作装置1中,并且该数据被生成成为对象声源数据。内容生成单元121将对象声源数据输出到内容生成单元121。

[0154] 元数据生成单元123生成对象元数据并且将数据输出至内容生成单元121。元数据生成单元123包括位置信息生成单元131、定位重要性信息生成单元132、以及个性化布局信息生成单元133。

[0155] 位置信息生成单元131根据制作者的设置生成关于每个对象的位置的位置信息。

[0156] 定位重要性信息生成单元132根据制作者的设置生成定位重要性信息。

[0157] 例如,根据以下内容设置定位的重要性:

- [0158] • 从收听位置到对象的位置的距离;
- [0159] • 对象的高度;
- [0160] • 对象的方向;
- [0161] • 是否集中地收听对象。

[0162] 当根据从收听位置到对象的位置的距离设置定位的重要性时,作为指示定位的重要性的值,例如,为布置在收听位置附近的对象设置高值。在这种情况下,关于用作阈值的距离,布置在收听位置附近的对象被再现为个性化HRTF再现对象,而远离收听位置布置的对象被再现为共用HRTF再现对象。

[0163] 当根据对象的高度设置定位的重要性时,作为指示定位的重要性的值,例如,为布置在高位置处的对象设置高值。在这种情况下,对于用作阈值的高度,高于参考位置布置的对象被再现为个性化HRTF再现对象,而低于参考位置布置的对象被再现为共用HRTF再现对象。

[0164] 当根据对象的方向设置定位的重要性时,作为指示定位的重要性的值,例如,为布置在接近收听者正面的方向上的对象设置高值。在这种情况下,相对于收听者的正面,沿预定范围的方向布置的对象被再现为个性化HRTF再现对象,而沿范围之外的方向布置的对象被再现为共用HRTF再现对象。

[0165] 当根据是否集中收听对象来设置定位的重要性时,作为指示定位的重要性的值,例如为集中收听对象设置高值。

[0166] 可通过使用从收听位置到对象的位置的距离、对象的高度、对象的方向、以及是否集中收听对象中的至少一些的组合来设置定位的重要性。

[0167] 以这种方式,内容制作装置1用作生成包括标志的元数据的信息处理装置,所述标志用作用于选择每个对象是将被再现为个性化HRTF再现内容还是共用HRTF再现内容的标准。

[0168] 个性化布局信息生成单元133根据制作者的设置生成个性化布局信息。

[0169] <内容管理装置2的配置>

[0170] 图17是示出内容管理装置2的功能配置实施例的框图。

[0171] 在内容管理装置2中,实现内容/共用HRTF获取单元151、对象音频处理单元152、2-ch混合单元153、传输数据生成单元154、传输数据存储单元155和传输单元156。通过使图14中的CPU 101执行预定程序来实现图17中所示的至少一些功能单元,CPU 101构成内容管理装置2。

[0172] 内容/共用HRTF获取单元151获取由内容制作装置1生成的有声读物内容。例如,内容/共用HRTF获取单元151控制通信单元109并接收和获取从内容制作装置1发送的有声读物内容。

[0173] 此外,内容/共用HRTF获取单元151获取针对有声读物内容中包括的对象的共用HRTF。例如,关于在有声读物内容的制作侧测量的共用HRTF的信息被输入到内容管理装置2。

[0174] 共用HRTF可以作为构成有声读物内容的数据包括在有声读物内容中,并且可以利用有声读物内容获取共用HRTF。

[0175] 内容/共用HRTF获取单元151基于包括在对象元数据中的定位重要性信息从构成有声读物内容的对象中选择个性化HRTF再现对象和共用HRTF再现对象。当通过从1至10的数值表示音像定位的重要性时,例如,音像定位的重要性等于或高于阈值的对象被选择为个性化HRTF再现对象,而音像定位的重要性低于阈值的对象被选择为共用HRTF再现对象。

[0176] 内容/共用HRTF获取单元151将关于被选择作为个性化HRTF再现对象的对象的音频数据输出到传输数据生成单元154。

[0177] 此外,内容/共用HRTF获取单元151将关于被选择为共用HRTF再现对象的对象的音频数据输出到对象音频处理单元152的渲染单元161。内容/共用HRTF获取单元151将获取的共用HRTF输出到双耳处理单元162。

[0178] 当生成个性化HRTF再现对象的数量变化的多个传输数据时,在改变用作阈值的值的同时,基于定位重要性信息重复地选择个性化HRTF再现对象和共用HRTF再现对象。

[0179] 对象音频处理单元152对关于从内容/共用HRTF获取单元151提供的共用HRTF再现对象的音频数据执行对象音频处理。充当立体声音处理的对象音频处理包含对对象的再现和双耳处理。

[0180] 具体地,在该实施例中,在作为发送侧的装置的内容管理装置2中执行对构成有声读物内容的对象之中的共用HRTF再现对象的对象音频处理。与将关于构成有声读书的内容的所有对象的音频数据发送到再现装置11相比,这可减少再现装置11中的处理的负荷。

[0181] 对象音频处理单元152包括渲染单元161和双耳处理单元162。

[0182] 渲染单元161基于位置信息等对共用HRTF再现对象执行渲染,并将通过渲染获得的音频数据输出到双耳处理单元162。渲染单元161基于关于共用HRTF再现对象的音频数据执行诸如VBAP(基于矢量的振幅平移)的渲染。

[0183] 双耳处理单元162使用共用HRTF对从渲染单元161提供的关于对象的音频数据执行双耳处理,并输出通过双耳处理获得的音频信号。由双耳处理单元162执行的双耳处理包括参考图5和图6描述的卷积。

[0184] 如果多个对象被选择为共用HRTF再现对象,则2-ch混合单元153对基于关于对象的音频数据生成的音频信号执行2-ch混合。

[0185] 通过执行2-ch混合,生成包括共用HRTF再现对象的音频信号L和音频信号R的基于声道的音频数据。通过2-ch混合获得的2-ch混合音频数据(基于声道的音频数据)被输出至传输数据生成单元154。

[0186] 传输数据生成单元154通过将来自内容/共用HRTF获取单元151的关于个性化HRTF再现对象的音频数据与来自2-ch混合单元153的关于共用HRTF再现对象的2-ch混合音频数据链接来生成传输数据。关于个性化HRTF再现对象的音频数据包括诸如位置信息和个性化布局信息的对象元数据。

[0187] 图18示出了传输数据的实施例。

[0188] 如图18所示,针对有声读物内容的每个标题生成个性化HRTF再现对象的数量变化的多个传输数据。

[0189] 基本上,传输数据的数据量随着个性化HRTF再现对象的数量而增加。此外,通过使用个性化HRTF执行对象音频处理,因此,具有精确定位的对象的数量随着个性化HRTF再现对象的数量而增加。这允许用户获得增强的临场感。

[0190] 在图18的实施例中,传输数据D1至D4以数据量的降序生成为有声读物内容#1的传输数据。

[0191] 传输数据D1至D3是具有大量的个性化HRTF再现对象的传输数据、具有中间数量的个性化HRTF再现对象的传输数据、以及具有小数量的个性化HRTF再现对象的传输数据。传输数据D1至D3分别包括关于共用HRTF再现对象的2-ch混合音频数据以及关于预定数量的个性化HRTF再现对象的音频数据。

[0192] 传输数据D4是不包括关于个性化HRTF再现对象的音频数据但是仅包括关于共用HRTF再现对象的2-ch混合音频数据的传输数据。

[0193] 在图17中的传输数据存储单元155存储由传输数据生成单元154生成的传输数据。

[0194] 传输单元156通过控制通信单元109与再现装置11通信,并且将存储在传输数据存储单元155中的传输数据传输至再现装置11。例如,在关于由再现装置11再现的音频内容的传输数据之中,传输对应于从再现装置11请求的质量的传输数据。当再现装置11请求高质量时,发送具有大量的个性化HRTF再现对象的传输数据。

[0195] 可通过下载或流式传输来传输传输数据。

[0196] <再现装置11的配置>

[0197] 图19为示出再现装置11的硬件配置实施例的框图。

[0198] 再现装置11被配置为使得通信单元202、存储器203、操作单元204、相机205、显示器206和音频输出单元207连接至控制单元201。

[0199] 控制单元201包括CPU、ROM和RAM。控制单元201通过执行预定程序来控制整个再现装置11的操作。

[0200] 在控制单元201中,实现应用执行单元201A。诸如用于再现有声读物内容的应用的各种应用(应用程序)由应用执行单元201A执行。

[0201] 通信单元202是适应移动通信系统的无线电通信(例如,5G通信)的通信模块。通信单元202接收从基站输出的无线电波并且经由网络21与诸如内容管理装置2的各种装置通信。通信单元202接收从内容管理装置2发送的信息并将该信息输出至控制单元201。

[0202] 存储器203包括闪速存储器。存储器203存储关于控制单元201要执行的应用的各

种信息。

[0203] 操作单元204包括叠加在显示器206上的各种按钮或触摸面板。操作单元204向控制单元201输出关于用户操作的内容的信息。

[0204] 相机205响应于用户操作来执行成像。

[0205] 显示器206包括有机EL显示器或LCD。包括用于再现有声读物内容的应用的画面的各种画面显示在显示器206上。

[0206] 音频输出单元207通过对象音频处理等将从控制单元201提供的音频数据发送至头戴式耳机12,并且从头戴式耳机12输出有声读物内容的声音。

[0207] 图20为示出再现装置11的功能配置实施例的框图。

[0208] 在再现装置11中,实现内容处理单元221、定位改进单元222、以及再现单元223。通过使图19中的应用执行单元201A执行用于再现有声读物内容的应用来实现图20中所示的功能单元中的至少一些。

[0209] 内容处理单元221包括数据获取单元231、内容存储单元232以及接收设置单元233。

[0210] 数据获取单元231通过控制通信单元202与内容管理装置2通信,并且获取关于有声读物内容的传输数据。

[0211] 例如,数据获取单元231将关于由接收设置单元233设置的质量的信息发送至内容管理装置2并且获取从内容管理装置2发送的传输数据。从内容管理装置2,发送包括关于个性化HRTF再现对象的音频数据的传输数据,该个性化HRTF再现对象的音频数据的数量根据由接收设置单元233设置的质量增加。将由数据获取单元231获取的传输数据输出至内容存储单元232并存储在内容存储单元232中。数据获取单元231用作传输数据获取单元,该传输数据获取单元从作为传输源装置的内容管理装置2准备的多个传输数据中获取要再现的传输数据。

[0212] 数据获取单元231将关于内容存储单元232中存储的传输数据中包括的个性化HRTF再现对象的音频数据输出至再现单元223的对象音频处理单元251。此外,数据获取单元231将关于共用HRTF再现对象的2-ch混合音频数据输出到2-ch混合单元252。如果关于个性化HRTF再现对象的音频数据包括在传输数据中,则数据获取单元231将包括在传输数据的对象元数据中的个性化布局信息输出到定位改进单元222。

[0213] 例如,接收设置单元233设置从内容管理装置2请求的传输数据的质量,并且将关于所设置的质量的信息输出至数据获取单元231。例如,由用户设置要从内容管理装置2请求的传输数据的质量。

[0214] 定位改进单元222包括定位改进系数获取单元241、定位改进系数存储单元242、个性化HRTF获取单元243和个性化HRTF存储单元244。

[0215] 定位改进系数获取单元241基于从数据获取单元231提供的个性化布局信息中指示的扬声器布局,获取用于改进定位的定位改进系数。定位改进系数还基于由用户选择的收听位置而适当地获取。

[0216] 例如,获取用于回响的信息作为定位改进系数。将由定位改进系数获取单元241获取的定位改进系数输出到定位改进系数存储单元242并存储。

[0217] 个性化HRTF获取单元243基于从数据获取单元231提供的个性化布局信息获取与

用户的收听位置对应的个性化HRTF。例如,从提供个性化HRTF的外部装置获取个性化HRTF。提供个性化HRTF的外部装置可为内容管理装置2或与内容管理装置2不同的装置。

[0218] 例如,个性化HRTF获取单元243控制相机205并获取通过对用作收听者的用户的耳朵成像而获得的耳朵图像。通过控制通信单元202,个性化HRTF获取单元243将个性化布局信息和耳朵图像发送到外部装置,并随后接收并获取响应于个性化布局信息和耳朵图像的发送而发送的个性化HRTF。由个性化HRTF获取单元243获取的与收听位置对应的个性化HRTF被输出到个性化HRTF存储单元244并存储在个性化HRTF存储单元244中。

[0219] 图21示出了为外部装置提供的推断模型的实施例。

[0220] 如图21所示,对于提供个性化HRTF的外部装置,准备个性化HRTF推断器以接收关于个性化布局信息中指示的扬声器布局和耳朵图像的信息作为输入,并输出个性化HRTF。

[0221] 个性化HRTF推断器是例如通过使用学习数据执行机器学习而生成的推断模型,学习数据是通过包括图7的布局的各种布局中的扬声器测量的一组HRTF和测量者的耳朵图像。在外部装置中,基于从再现装置11发送的个性化布局信息和耳朵图像来推断个性化HRTF,并将推断结果发送到再现装置11。

[0222] 例如,通过根据收听位置校正作为个性化HRTF推断器的推断结果获得的HRTF来获取与收听位置对应的个性化HRTF。关于收听位置的信息可被输入到个性化HRTF推断器,从个性化HRTF推断器输出的HRTF可直接用作与收听位置对应的个性化HRTF。

[0223] 个性化HRTF推断器可在再现装置11中准备。在这种情况下,个性化HRTF获取单元243使用耳朵图像和个性化布局信息来推断个性化HRTF。

[0224] 图20中的再现单元223包括对象音频处理单元251和2-ch混合单元252。

[0225] 对象音频处理单元251对从数据获取单元231提供的关于个性化HRTF再现对象的音频数据执行包括渲染和双耳处理的对象音频处理。

[0226] 如上所述,在内容管理装置2中执行对共用HRTF再现对象的对象音频处理,而在再现装置11中执行对个性化HRTF再现对象的对象音频处理。

[0227] 对象音频处理单元251包括渲染单元261、定位改进单元262和双耳处理单元263。

[0228] 渲染单元261基于位置信息对个性化HRTF再现对象执行渲染,并将通过渲染获得的音频数据输出到定位改进单元262。渲染单元261基于关于个性化HRTF再现对象的音频数据执行诸如VBAP的渲染。

[0229] 定位改进单元262从定位改进系数存储单元242读取定位改进系数,对从渲染单元261提供的音频数据进行诸如回响等定位改进。定位改进单元262将关于通过定位改进获得的对象的音频数据输出到双耳处理单元263。能够避免定位改进单元262的定位改进。

[0230] 双耳处理单元263从个性化HRTF存储单元244读取并获取个性化HRTF。双耳处理单元263使用从定位改进单元262提供的关于对象的音频数据的个性化HRTF执行双耳处理,并输出通过双耳处理获得的音频信号。

[0231] 2-ch混合单元252基于2-ch混合音频数据对从数据获取单元231提供的共用HRTF再现对象和从双耳处理单元263提供的音频信号执行2-ch混合。

[0232] 通过执行2-ch混合,生成包括个性化HRTF再现对象的音频分量和共用HRTF再现对象的音频分量的音频信号L和音频信号R。2-ch混合单元252将通过2-ch混合获得的2-ch混合音频数据输出至音频输出单元207,并且从头戴式耳机12输出有声读物内容的对象的声

音。

[0233] <<每个装置的操作>>

[0234] 以下将描述如此配置的装置的操作。

[0235] <内容制作装置1的操作>

[0236] 参考图22的流程图,下面将描述内容制作装置1的处理。

[0237] 在步骤S1中,对象声源数据生成单元122生成关于每个对象的声源数据。

[0238] 在步骤S2中,位置信息生成单元131响应于制作者的操作设置每个对象的位置并且生成位置信息。

[0239] 在步骤S3中,定位重要性信息生成单元132响应于制作者的操作设置每个对象的定位的重要性,并且生成定位重要性信息。

[0240] 在步骤S4中,个性化布局信息生成单元133为每个收听位置生成个性化布局信息。

[0241] 在步骤S5中,内容生成单元121通过链接由对象声源数据生成单元122生成的对象声源数据和包括由位置信息生成单元131生成的位置信息的对象元数据来生成有声读物内容。

[0242] 对于有声读物内容的每个标题执行处理。

[0243] 由于对象音频数据被包括作为关于有声读物内容的音频数据,所以与将书的语音旁白录音为立体声相比,制作者可以生成更真实的有声读物内容。

[0244] 通常,定位的重要性在对象之间变化。制作者可针对每个对象设置定位的重要性,并使用个性化HRTF或共用HRTF指定是否将再现每个对象。

[0245] <内容管理装置2的操作>

[0246] 参考图23的流程图,下面将描述内容管理装置2的处理。

[0247] 在步骤S11,内容/共用HRTF获取单元151获取有声读物内容和共用HRTF。

[0248] 在步骤S12中,内容/共用HRTF获取单元151注意到构成有声读物内容的对象之一。

[0249] 在步骤S13中,内容/共用HRTF获取单元151基于定位重要性信息确认定位被注意的对象的重要性。

[0250] 在步骤S14中,内容/共用HRTF获取单元151确定是否将注意的对象再现为个性化HRTF再现对象。

[0251] 通过比较指示被注意对象的定位的重要性的值与阈值来确定被注意对象是否将被再现为个性化HRTF再现对象。例如,如果指示被注意的对象的定位的重要性的值等于或大于阈值,则确定对象将被再现为个性化HRTF再现对象。

[0252] 如果在步骤S14确定不将对象再现为个性化HRTF再现对象,即,被注意的对象不被再现为个性化HRTF再现对象,则处理进行到步骤S15。

[0253] 在步骤S15中,对象音频处理单元152对关于从内容/共用HRTF获取单元151提供的共用HRTF再现对象的音频数据执行对象音频处理。通过使用共用HRTF对关于共用HRTF再现对象的音频数据执行对象音频处理。

[0254] 如果在步骤S14中确定被注意的对象将被再现为个性化HRTF再现对象,则跳过步骤S15的处理。

[0255] 在步骤S16中,内容/共用HRTF获取单元151确定是否已经注意到构成有声读物内容的所有对象。

[0256] 如果在步骤S16中确定尚未注意到对象中的至少一个,则处理返回到步骤S12以注意到另一对象。重复执行上述处理。

[0257] 如果在步骤S16中确定已经注意到所有对象,则2-ch混合单元153在步骤S17中确定是否已经对对象中的至少一个执行渲染。

[0258] 如果在步骤S17中确定已经对对象中的至少一个执行渲染,则在步骤S18,2-ch混合单元153对基于关于共用HRTF再现对象的音频数据生成的音频信号执行2-ch混合。关于共用HRTF再现对象的2-ch混合音频数据被提供给传输数据生成单元154,通过2-ch混合获得2-ch混合音频数据。

[0259] 如果在步骤S17中确定还没有对象被渲染,则跳过步骤S18的处理。

[0260] 在步骤S19中,传输数据生成单元154通过将关于个性化HRTF再现对象的音频数据和关于共用HRTF再现对象的2-ch混合音频数据链接来生成传输数据。在不存在共用HRTF再现对象的情况下,生成仅包括关于个性化HRTF再现对象的音频数据的传输数据。以这种方式,可生成仅包括关于个性化HRTF再现对象的音频数据的传输数据。

[0261] 在改变用作确定是否将对象再现为个性化HRTF再现对象的标准的阈值时,重复该处理,使得在内容管理装置2中生成并准备具有不同数量的个性化HRTF再现对象的多个传输数据。

[0262] 对于需要定位的重要性的对象,对象音频照原样发送,从而内容管理装置2可使再现装置11使用个性化HRTF执行对象音频处理。

[0263] 如果所有对象作为对象音频被发送,则根据再现侧的性能或通信环境,在再现装置11中可能不适当地执行对象音频处理。在内容管理装置2中对不需要定位的重要性的一些对象执行对象音频处理,从而内容管理装置2可根据例如再现装置11的性能提供传输数据并且使再现装置11执行适当的处理。

[0264] <再现装置11的操作>

[0265] 参照图24的流程图,下面描述再现装置11的处理。

[0266] 在步骤S31中,定位改进单元222获取用户的收听位置。例如,通过用户在用于再现有声读物内容的应用的屏幕上的选择来获取用户的收听位置。用户的收听位置可通过基于传感器数据分析用户的位置来获取。

[0267] 在步骤S32中,数据获取单元231从在内容管理装置2中准备的传输数据中获取对应于质量设置的传输数据。

[0268] 在步骤S33,数据获取单元231确定获取的传输数据是否包括关于个性化HRTF再现对象的音频数据。

[0269] 如果在步骤S33中确定包括关于个性化HRTF再现对象的音频数据,则在步骤S34中,数据获取单元231获取包括在传输数据的对象元数据中的个性化布局信息。

[0270] 在步骤S35,个性化HRTF获取单元243控制相机205并获取通过对用户的耳朵进行成像而获得的耳朵图像。例如,响应于用户操作,对耳朵图像进行成像。在收听有声读物内容之前,用户利用再现装置11的相机205拍摄用户的耳朵的图像。

[0271] 在步骤S36,个性化HRTF获取单元243基于个性化布局信息和耳朵图像获取个性化HRTF。

[0272] 在步骤S37中,定位改进单元222基于在个性化布局信息中指示的扬声器布局等获

得用于改进定位的定位改进系数。

[0273] 在步骤S38,对象音频处理单元251通过使用个性化HRTF对关于个性化HRTF再现对象的音频数据执行对象音频处理。

[0274] 在步骤S39,2-ch混合单元252基于2-ch混合音频数据对共用HRTF再现对象和个性化HRTF再现对象的音频信号执行2-ch混合,其中,通过对象音频处理来获得音频信号。

[0275] 在步骤S40中,2-ch混合单元252输出通过2-ch混合获得的2-ch混合音频数据,并且从头戴式耳机12输出有声读物内容的对象的声音。从头戴式耳机12输出的声音包括个性化HRTF再现对象的声音和共用HRTF再现对象的声音。

[0276] 在步骤S33中,如果确定关于个性化HRTF再现对象的音频数据不包括在传输数据中,则跳过步骤S34至S39的处理。传输数据仅包括关于共用HRTF再现对象的2-ch混合音频数据。

[0277] 在这种情况下,在步骤S40中,基于关于共用HRTF再现对象的2-ch混合音频数据输出共用HRTF再现对象的声音,2-ch混合音频数据包括在传输数据中。

[0278] 该处理允许再现装置11定位用户周围的每个对象的音像。

[0279] 通过使用针对用户个性化的HRTF再现需要定位的重要性的对象,允许再现装置11以高精度定位音像。

[0280] 如果在内容管理装置2中对所有对象执行对象音频处理,则再现装置11不能改变对象的定位。在再现侧上执行需要定位的重要性的一些对象的对象音频处理,从而再现装置11可通过使用针对用户个性化的HRTF执行再现来改变对象的定位。

[0281] 与收听未经立体声音处理的声音相比,用户可获得真实性。

[0282] 此外,在内容管理装置2中执行不需要定位的重要性的一些对象的对象音频处理,从而减少了再现装置11的处理的负荷。通常,诸如计算资源、电池耗尽和应对热量增加的措施的负荷随着对象的数量而增加。此外,随着对象数量的增加,再现装置11需要更大的存储容量来存储个性化HRTF的系数。

[0283] 换言之,再现装置11可以在减少处理负荷的同时提供逼真的有声读物内容。

[0284] <<变形例>>

[0285] <变形例1:制作侧上的传输数据准备>

[0286] 可以在内容制作装置1中生成传输数据。在这种情况下,在内容管理装置2的配置之中,为内容制作装置1提供用于生成传输数据的配置。

[0287] 图25是示出内容制作装置1的另一个配置实施例的框图。

[0288] 图25中的内容制作装置1包括内容生成单元301和传输数据生成单元302。

[0289] 内容生成单元301包括参照图15描述的配置。

[0290] 在参考图17描述的内容管理装置2的配置中,传输数据生成单元302设置有作为用于生成传输数据的配置的至少内容/共用HRTF获取单元151、对象音频处理单元152、2-ch混合单元153和传输数据生成单元154。将由传输数据生成单元302配置的传输数据提供给内容管理装置2。

[0291] 内容管理装置2的传输单元156(图17)从由内容制作装置1生成的传输数据中选择具有预定质量的传输数据并且将该传输数据传输至再现装置11。

[0292] <变形例2:传输数据的选择实施例>

[0293] 将有关质量的信息从再现装置11传输给内容管理装置2,并且由内容管理装置2选择与从再现装置11中请求的质量对应的传输数据。要发送的传输数据可以基于其他标准来选择。

[0294] 在这种情况下,有关网络环境的信息(例如,在内容管理装置2与再现装置11之间的吞吐量、再现侧的装置配置以及再现装置11的处理性能)从再现装置11传输给内容管理装置2。在内容管理装置2中,例如,根据从再现装置11中传输的信息,选择与网络环境对应的传输数据,并且将该传输数据传输给再现装置11。

[0295] <变形例3:传输数据的选择实施例>

[0296] 用户佩戴作为装置的头戴式耳机12。代替头戴式耳机12,可使用能够收集外部声音的耳戴式耳机,如图26中所示。在这种情况下,以混合风格收听有声读物内容,该混合风格是耳戴式耳机和设置在再现侧上的扬声器(真实扬声器)的组合。

[0297] 图26示出了具有收集外部声音的功能的耳戴式耳机的外观。

[0298] 图26中的耳戴式耳机是所谓的不闭合耳孔的开耳戴式(开放式)耳机。

[0299] 如在图26中的气球中放大的,耳机的右单元包括经由U形声音导管312彼此连接的驱动器单元311和环形安装部313。通过将安装部313按压在外耳孔周围来安装右单元,使得右耳被保持在安装部313与驱动器单元311之间。

[0300] 左单元具有与右单元相同的配置。左单元和右单元具有有线或无线连接。

[0301] 右单元的驱动器单元311接收从再现装置11传输的音频信号,并且从声音导管312的先端输出声音,如由箭头#1表示的,该声音与音频信号对应。用于向外耳孔输出声音的孔形成在声音导管312和安装部313的接合处。

[0302] 安装部313成形为环状。与从声音导管312的先端输出的内容声音一起,环境声音也到达外耳孔,如箭头#2所示。

[0303] 图27是示出当有声读物内容以混合样式被听到时的再现装置11的配置实施例的框图。

[0304] 在图27中所示的配置中,与参考图20描述的那些配置相同的配置由相同的参考字符表示。适当省略重复的描述。这同样适用于稍后将描述的图28。

[0305] 图27中的再现装置11的配置与图20的配置的不同之处在于,对象音频处理单元251额外设置有真实扬声器配置信息获取单元264,并且再现单元223额外包括真实扬声器输出控制单元271。

[0306] 真实扬声器配置信息获取单元264获取再现扬声器配置信息,该信息是关于设置在再现侧的真实扬声器的配置的信息。例如,再现扬声器配置信息包括关于真实扬声器的布局的信息。由再现扬声器配置信息指示的真实扬声器的布局可与个性化布局信息中指示的扬声器布局相同或不同。真实扬声器配置信息获取单元264获取的再现扬声器配置信息被输出到渲染单元261和定位改进单元262。

[0307] 在渲染单元261中,将多个个性化HRTF再现对象中的预定的个性化HRTF再现对象分配给真实扬声器,将其他对象分配给耳机,然后执行渲染。可基于从对象到收听者的距离或者定位重要性信息来选择对象中的哪一个被分配给扬声器/耳机。

[0308] 通过渲染获得的音频数据在定位改进单元262中经受定位改进。在定位改进中,例如,根据真实扬声器的布局执行回响。在通过定位改进获得的音频数据中,将关于分配给真

实扬声器的对象的音频数据提供给真实扬声器输出控制单元271,并且将关于分配给耳机的对象的音频数据提供给双耳处理单元263。

[0309] 真实扬声器输出控制单元271根据渲染单元261的分配生成具有与真实扬声器的数量一样多的声道的音频信号,并从各个真实扬声器输出分配给真实扬声器的对象的声音。

[0310] 在执行2-ch混合之后,从耳机输出分配给耳机的对象的声音,允许用户收听来自耳机的声音以及来自真实扬声器的声音。这样,能够以耳机和真实扬声器的组合的混合风格听到有声读物内容。

[0311] 在不使用头戴式耳机12或耳戴式耳机的情况下,可仅通过真实扬声器收听有声读物内容。

[0312] <变形例4:头部跟踪的实施例>

[0313] 当用户收听有声读物内容时,可执行所谓的头部跟踪以根据用户的头部定向校正HRTF。在头戴式耳机12中,安装诸如陀螺仪传感器和加速度传感器的传感器。

[0314] 图28是示出再现装置11的另一个配置实施例的框图。

[0315] 图28中的再现装置11的配置与图20的配置的不同之处在于,再现单元223额外设置有头部位置/定向获取单元281。

[0316] 头部位置/定向获取单元281基于由安装在头戴式耳机12中的传感器检测的传感器数据来检测用户的头部位置和定向。关于由头部位置/定向获取单元281检测的头部位置和定向的信息被输出到渲染单元261。

[0317] 基于由头部位置/定向获取单元281检测的头部位置和定向,渲染单元261校正由位置信息指示的对象位置,以便固定对象的绝对位置。例如,如果用户的头部相对于特定方向向右旋转30°,则渲染单元261将所有对象的位置向左旋转30°,从而校正对象的位置。

[0318] 在定位改进单元262和双耳(binaural)处理单元263中,例如,对各个个性化HRTF再现对象的音频数据执行与校正位置对应的处理。

[0319] 通过这种方式,在听到有声读物内容时,执行头部跟踪,从而允许再现装置11将每个对象的音像定位在固定位置上,而不受用户头部的运动的影响。

[0320] <变形例5:定位的重要性的确定实施例>

[0321] 每个对象的定位的重要性由制作者手动设置。定位的重要性可以由内容制作装置1自动设置。

[0322] 图29是示出内容制作装置1的另一个配置实施例的框图。

[0323] 在图29中所示的配置中,与参考图15所描述的配置相同的配置由相同的参考字符表示。适当省略重复的描述。

[0324] 图29中的内容制作装置1的配置与图15的配置的不同之处在于元数据生成单元123另外设置有声源属性信息获取单元291。

[0325] 声源属性信息获取单元291获取关于对象的声源数据的属性信息。属性信息例如包括用于存储声源数据的文件的文件名、诸如声源数据的种类的文件内容、以及从文件名和文件内容估计的信息。由声源属性信息获取单元291获取的属性信息被输出到定位重要性信息生成单元132。

[0326] 定位重要性信息生成单元132基于从声源属性信息获取单元291提供的属性信息,

设置每个对象的定位的重要性。

[0327] 例如,根据属性信息指定各对象的声源数据的种类(包括效果音和音声)。在利用强调效果音的自动设置的情况下,将高值设置为指示效果音的对象的定位的重要性的值。在利用强调音声的自动设置的情况下,将高值设置为指示针对音声的对象的定位的重要性的值。

[0328] 这样,自动设置定位的重要性,从而降低了有声读物内容的制作负荷。

[0329] <<其他实施例>>

[0330] <处理的分配实施例>

[0331] 可在再现装置11中执行针对共用HRTF再现对象和个性化HRTF再现对象的对象音频处理。在这种情况下,关于共用HRTF再现对象的音频数据与关于个性化HRTF再现对象的音频数据照原样从内容管理装置2发送到再现装置11。

[0332] 再现装置11获取共用HRTF并对共用HRTF再现对象执行对象音频处理。此外,再现装置11获取个性化HRTF并且对个性化HRTF再现对象执行对象音频处理。

[0333] 同样在这种情况下,再现装置11可以提供真实的有声读物内容。

[0334] <对象切换的实施例>

[0335] 个性化HRTF再现对象和共用HRTF再现对象可动态地切换。在这种情况下,针对每个对象动态地切换作为个性化HRTF再现对象的再现和作为共用HRTF再现对象的再现。

[0336] 例如,在内容管理装置2与再现装置11之间的网络环境较差或者在再现装置11中的处理负荷较大的情况下,选择许多对象作为共用HRTF再现对象。在内容管理装置2与再现装置11之间的适当网络环境或再现装置11中的处理负荷小的情况下,选择许多对象作为个性化HRTF再现对象。

[0337] 可根据通过流传输还是下载发送有声读物内容来切换个性化HRTF再现对象和共用HRTF再现对象。

[0338] 代替指示定位的重要性的标志,可以针对每个对象设置指示处理的优先级的标志。

[0339] 可选择地,代替指示定位的重要性的标志,标志可被设置为直接指示是将对象再现为个性化HRTF再现对象还是共用HRTF再现对象。

[0340] 以这种方式,可基于由制作侧指定的各种信息切换个性化HRTF再现对象和共用HRTF再现对象。

[0341] <交互的实施例>

[0342] • HRTF的获取

[0343] 基于耳朵图像等获取个性化HRTF。可根据用户交互获取个性化HRTF。

[0344] 个性化HRTF的获取实施例1

[0345] 例如,如果在再现侧准备包括如图7中所示设置的扬声器的空间,则作为测量者的用户可坐在扬声器的中心以测量个性化HRTF。

[0346] 个性化HRTF的获取实施例2

[0347] 可通过识别例如游戏格式的声音的方向来测量个性化HRTF。例如,当用户前方听到昆虫的声音时显示反应。

[0348] 个性化HRTF的获取实施例3

[0349] 如果准备了多个个性化HRTF样本,则可使得用户收听通过使用每个样本再现的声音并根据定位的感觉来选择使用哪个样本。当用户收听有声读物内容时,由用户选择的样本被用作个性化HRTF。

[0350] 个性化HRTF可根据游戏格式的响应、关于样本的重新选择的响应或在观看内容期间的用户响应来校正。

[0351] 个性化HRTF可利用游戏格式的个性化HRTF的获取和使用样本的个性化HRTF的获取的组合来获取。

[0352] 在这种情况下,例如,被计算为相对于用户的位置的定位误差是由位置信息指示的对象的位置的方向与用户感到音像的定位的位置的方向之间的角度差。在切换样本的同时重复计算定位误差,并且具有小定位误差的样本被选择为实际使用的个性化HRTF。

[0353] 当用户收听有声读物内容时,可根据上述方法获取个性化HRTF,并且当获取个性化HRTF时,可使用个性化HRTF执行再现。在这种情况下,例如,每个对象在有声读物的内容的前半部分中被再现为共用HRTF再现对象,并且预定对象在后半部分中被再现为个性化HRTF再现对象。

[0354] 可通过基于关于耳朵尺寸或头围的信息而不是耳朵图像的计算来获取个性化HRTF。

[0355] • 内容再现

[0356] 当再现有声读物内容时,用户可执行用于再现的操作。

[0357] 用于再现的操作包括:例如,选择沉浸式声源(对象)的放大/缩小、选择效果音的打开/关闭、旁白位置的指定、语言的切换、存在或不存在上述头部跟踪、以及选择故事。

[0358] <其他>

[0359] 上述处理可应用于包括对象音频数据的各种内容,例如,音乐内容、无线电内容、电影内容和电视节目的内容。

[0360] • 计算机的配置实施例

[0361] 上述一系列处理可以由硬件或软件执行。当通过软件执行该系列处理时,将构成软件的程序从程序记录介质安装到内置于专用硬件中的计算机或者通用个人计算机上。

[0362] 待安装的程序在被记录在图14中示出的可移除介质111中之后被提供,可移除介质111包括光盘(CD-ROM(致密盘只读存储器)、DVD(数字通用盘)等)和半导体存储器。此外,可通过有线或无线传输介质(诸如,局域网、互联网和数字广播)提供程序。程序可以预先安装在ROM 102或存储单元108中。

[0363] 由计算机执行的程序可以是按照本说明书中描述的顺序按时间序列执行处理的程序,或者可以是并行或在必要定时(例如,当被调用时)执行处理的程序。

[0364] 在本说明书中,系统意味着一组多个组成元件(装置、模块(部件)等),而不管所有组成元件是否位于同一壳体中。因此,存储在单独的壳体中并且经由网络连接的多个装置和包括存储在壳体中的多个模块的单个装置都是系统。

[0365] 而且,在本说明书中描述的有益效果仅仅是实施例性的,并不旨在进行限制,并且可以获得其他有益效果。

[0366] 本技术的实施方式不限于上述实施方式,并且在不背离本技术的主旨的情况下,可以做出各种变化。

[0367] 例如,本技术可以被配置为云计算,其中,多个装置经由网络共享和协作处理一个功能。

[0368] 此外,上述流程图中描述的每个步骤可以由一个装置执行或者由多个装置以共享方式执行。

[0369] 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一个装置执行或者由多个装置以共享方式执行。

[0370] <配置的组合实施例>

[0371] 本技术还可具有以下配置:

[0372] (1) 一种再现装置,包括再现单元,被配置为再现对象音频内容,对象音频内容包括通过使用个性化HRTF再现的第一对象和通过使用共用HRTF再现的第二对象,其中,个性化HRTF是针对收听者个性化的HRTF,共用HRTF是在多个收听者之间共享的HRTF。

[0373] (2) 根据(1)所述的再现装置,进一步包括:HRTF获取单元,被配置为基于收听者的耳朵的图像和关于测量共用HRTF时的扬声器的布局的布局信息获取个性化HRTF,布局信息包括在所述内容中。

[0374] (3) 根据(1)或(2)所述的再现装置,其中,关于内容的传输数据包括关于第一对象的音频数据以及通过使用共用HRTF对关于第二对象的音频数据执行处理而在内容的传输源装置中生成的基于声道的数据。

[0375] (4) 根据(3)所述的再现装置,其中,再现单元输出与通过使用个性化HRTF对关于第一对象的音频数据执行处理而生成的基于声道的数据和包括在传输数据中的基于声道的数据相对应的声音。

[0376] (5) 根据(1)或(2)所述的再现装置,其中,再现单元使用个性化HRTF对第一对象的音频数据执行处理,使用共用HRTF对第二对象的音频数据执行处理,并输出与通过对第一对象的音频数据执行处理而生成的基于声道的数据相对应的声音和与通过对第二对象的音频数据执行处理而生成的基于声道的数据相对应的声音。

[0377] (6) 根据(3)所述的再现装置,进一步包括传输数据获取单元,被配置为从为所述传输源装置准备的并且在所述第一对象的数量上变化的多个传输数据中获取要再现的所述传输数据。

[0378] (7) 一种使再现装置执行以下操作的再现方法:

[0379] 再现对象音频内容,所述对象音频内容包括通过使用个性化HRTF再现的第一对象和通过使用共用HRTF再现的第二对象,其中,所述个性化HRTF是针对收听者个性化的HRTF,所述共用HRTF是在多个收听者之间共享的HRTF。

[0380] (8) 一种使计算机执行以下处理的程序:

[0381] 再现对象音频内容,所述对象音频内容包括通过使用个性化HRTF再现的第一对象和通过使用共用HRTF再现的第二对象,其中,所述个性化HRTF是针对收听者个性化的HRTF,所述共用HRTF是在多个收听者之间共享的HRTF。

[0382] (9) 一种信息处理装置,包括:元数据生成单元,被配置为生成包括标志的元数据,其中,所述标志指示对象是将被再现为通过使用个性化HRTF再现的第一对象还是通过使用共用HRTF再现的第二对象,其中,所述个性化HRTF是针对收听者个性化的HRTF,所述共用HRTF是在多个收听者之间共享的HRTF;以及

[0383] 内容生成单元,被配置为生成包括关于多个对象的声源数据和元数据的对象音频内容。

[0384] (10) 根据(9)所述的信息处理装置,其中元数据生成单元生成进一步包括关于测量共用HRTF时的扬声器的布局的布局信息的元数据。

[0385] (11) 根据(9)或(10)所述的信息处理装置,其中所述元数据生成单元生成包括关于每个对象的音像的定位的重要性的信息作为标志的元数据。

[0386] (12) 根据(11)所述的信息处理装置,其中,基于对象类型、从收听位置到所述对象的位置的距离、在所述收听位置处的所述对象的方向、以及所述对象的高度中的至少一个来设置所述重要性。

[0387] (13) 一种信息处理方法,使信息处理装置执行:

[0388] 生成包括标志的元数据,其中,标志指示对象是将被再现为通过使用个性化HRTF再现的第一对象还是通过使用共用HRTF再现的第二对象,其中,个性化HRTF是针对收听者个性化的HRTF,共用HRTF是在多个收听者之间共享的HRTF;

[0389] 生成包括关于多个对象的声源数据和元数据的对象音频内容。

[0390] (14) 一种使计算机执行以下处理的程序:

[0391] 生成包括标志的元数据,并生成包括关于多个对象的声源数据和元数据的对象音频内容,其中,所述标志指示对象是将被再现为通过使用针对收听者个性化的HRTF的个性化HRTF再现的第一对象还是通过使用作为在多个收听者之间共享的HRTF的共用HRTF再现的第二对象。

[0392] (15) 一种信息处理装置,包括:

[0393] 内容获取单元,被配置为获取对象音频内容,对象音频内容包括关于多个对象的声源数据和元数据,元数据包括指示对象是将被再现为通过使用针对收听者个性化的HRTF的个性化HRTF再现的第一对象还是通过使用作为在多个收听者之间共享的HRTF的共用HRTF再现的第二对象的标志;

[0394] 音频处理单元,被配置为对关于基于标志被选择再现为第二对象的对象的音频数据使用共用HRTF执行处理;以及

[0395] 传输数据生成单元,被配置为生成关于内容的传输数据,关于内容的传输数据包括通过使用共用HRTF的处理而生成的基于声道的数据和关于第一对象的音频数据。

[0396] (16) 根据(15)所述的信息处理装置,其中所述传输数据生成单元针对每个内容生成在所述第一对象的数量上变化的多个传输数据。

[0397] (17) 一种信息处理方法,使信息处理装置执行:

[0398] 获取对象音频内容,对象音频内容包括关于多个对象的声源数据和元数据,元数据包括指示对象是将被再现为通过使用针对收听者个性化的HRTF的个性化HRTF再现的第一对象还是通过使用作为在多个收听者之间共享的HRTF的共用HRTF再现的第二对象的标志;

[0399] 对关于基于标志被选择再现为第二对象的对象的音频数据执行使用共用HRTF的处理;以及

[0400] 生成关于内容的传输数据,关于内容的传输数据包括通过使用共用HRTF的处理生成的基于声道的数据和关于第一对象的音频数据。

[0401] (18)一种使计算机执行以下处理的程序:

[0402] 获取对象音频内容,对象音频内容包括关于多个对象的声源数据和元数据,元数据包括指示对象是将被再现为通过使用针对收听者个性化的HRTF的个性化HRTF再现的第一对象还是通过使用作为在多个收听者之间共享的HRTF的共用HRTF再现的第二对象的标志;

[0403] 对关于基于标志被选择再现为第二对象的对象的音频数据执行使用共用HRTF的处理;以及

[0404] 生成关于内容的传输数据,关于内容的传输数据包括通过使用共用HRTF的处理生成的基于声道的数据和关于第一对象的音频数据。

[0405] [参考标号列表]

[0406] 1 内容制作装置

[0407] 2 内容管理装置

[0408] 11 再现装置

[0409] 12 头戴式耳机

[0410] 121 内容生成单元

[0411] 122 对象声源数据生成单元

[0412] 123 元数据生成单元

[0413] 131 位置信息生成单元

[0414] 132定位重要性信息生成单元

[0415] 133个性化布局信息生成单元

[0416] 151内容/共用HRTF获取单元

[0417] 152对象音频处理单元

[0418] 153 2-ch混合单元

[0419] 154 传输数据生成单元

[0420] 155 传输数据存储单元

[0421] 156 传输单元

[0422] 161 渲染单元

[0423] 162 双耳处理单元

[0424] 221 内容处理单元

[0425] 222 定位改进单元

[0426] 223 对象音频处理单元

[0427] 231 数据获取单元

[0428] 232 内容存储单元

[0429] 233 接收设置单元

[0430] 241 定位改进系数获取单元

[0431] 242 定位改进系数存储单元

[0432] 243个性化HRTF获取单元

[0433] 244个性化HRTF存储单元

[0434] 251对象音频处理单元

- [0435] 252 2-ch混合单元
- [0436] 261 渲染单元
- [0437] 262 定位改进单元
- [0438] 263 双耳处理单元

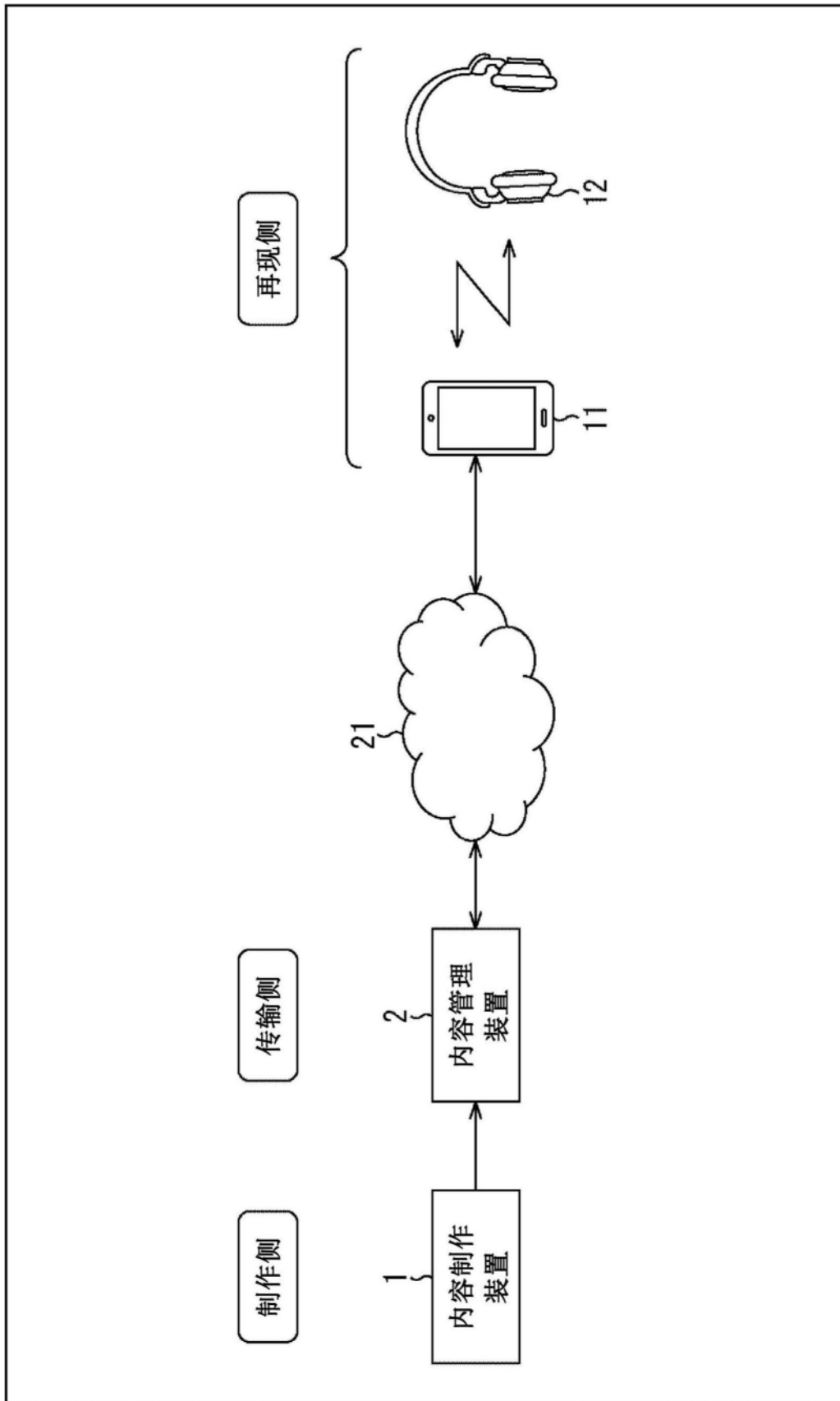


图1

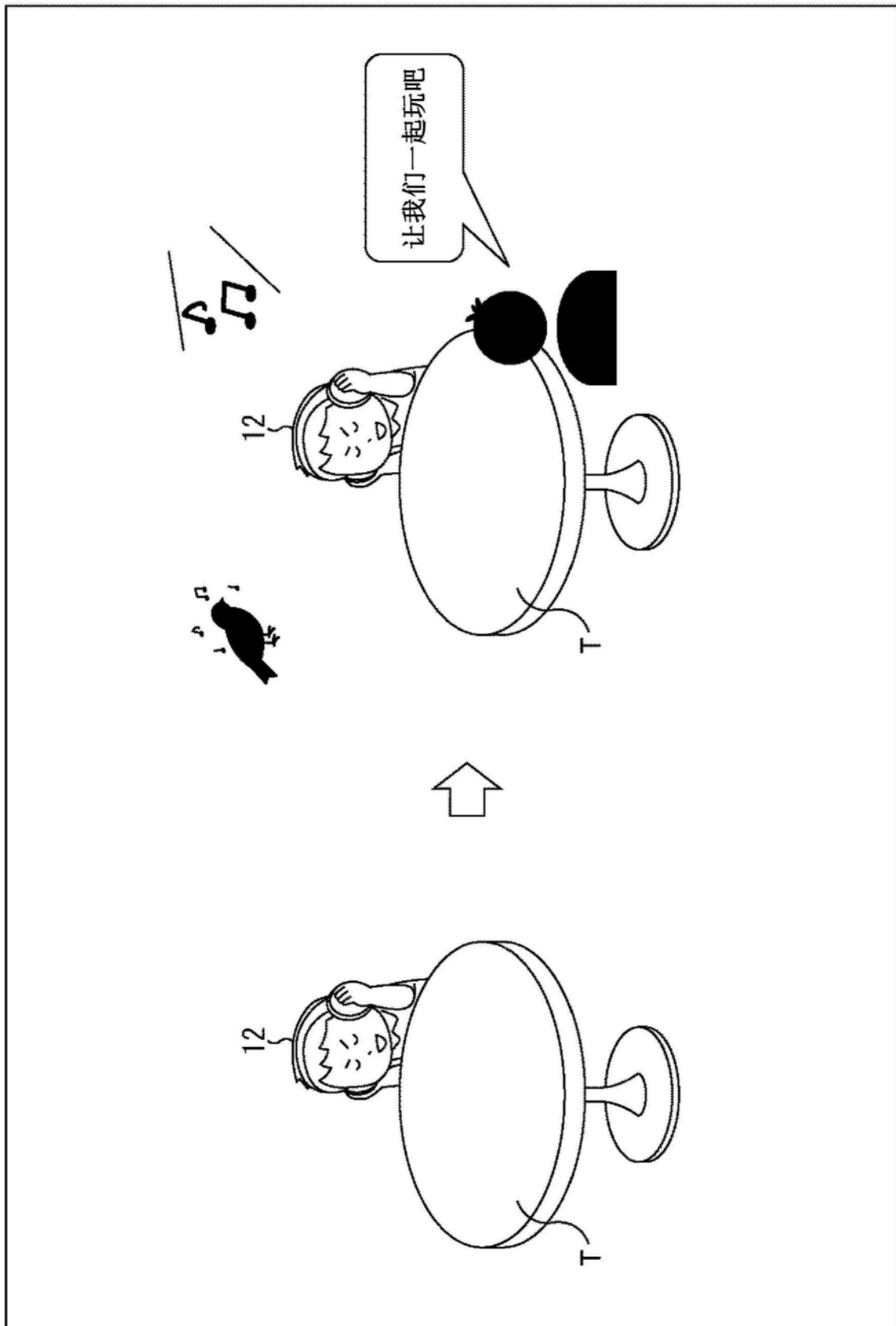


图2



图3

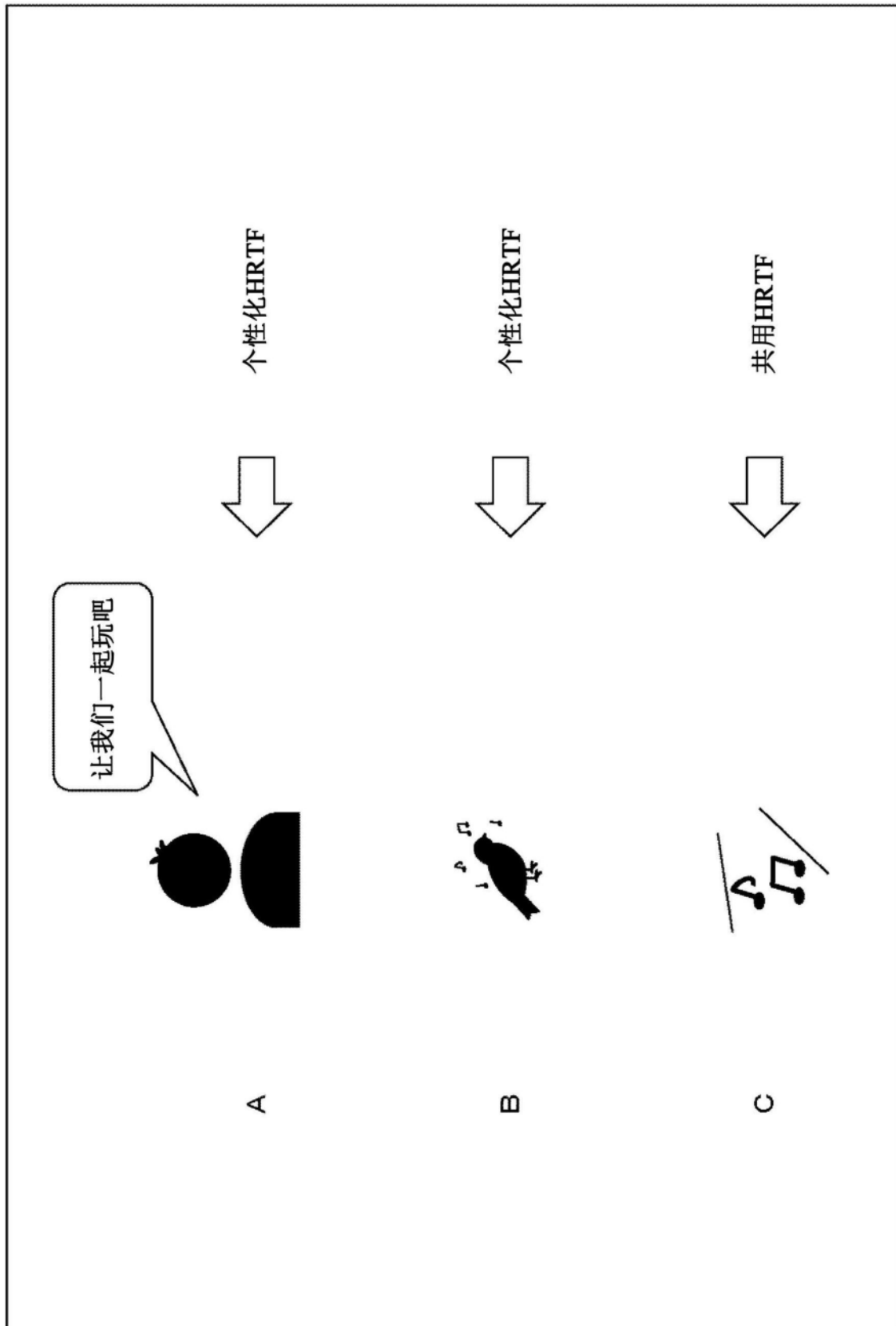


图4

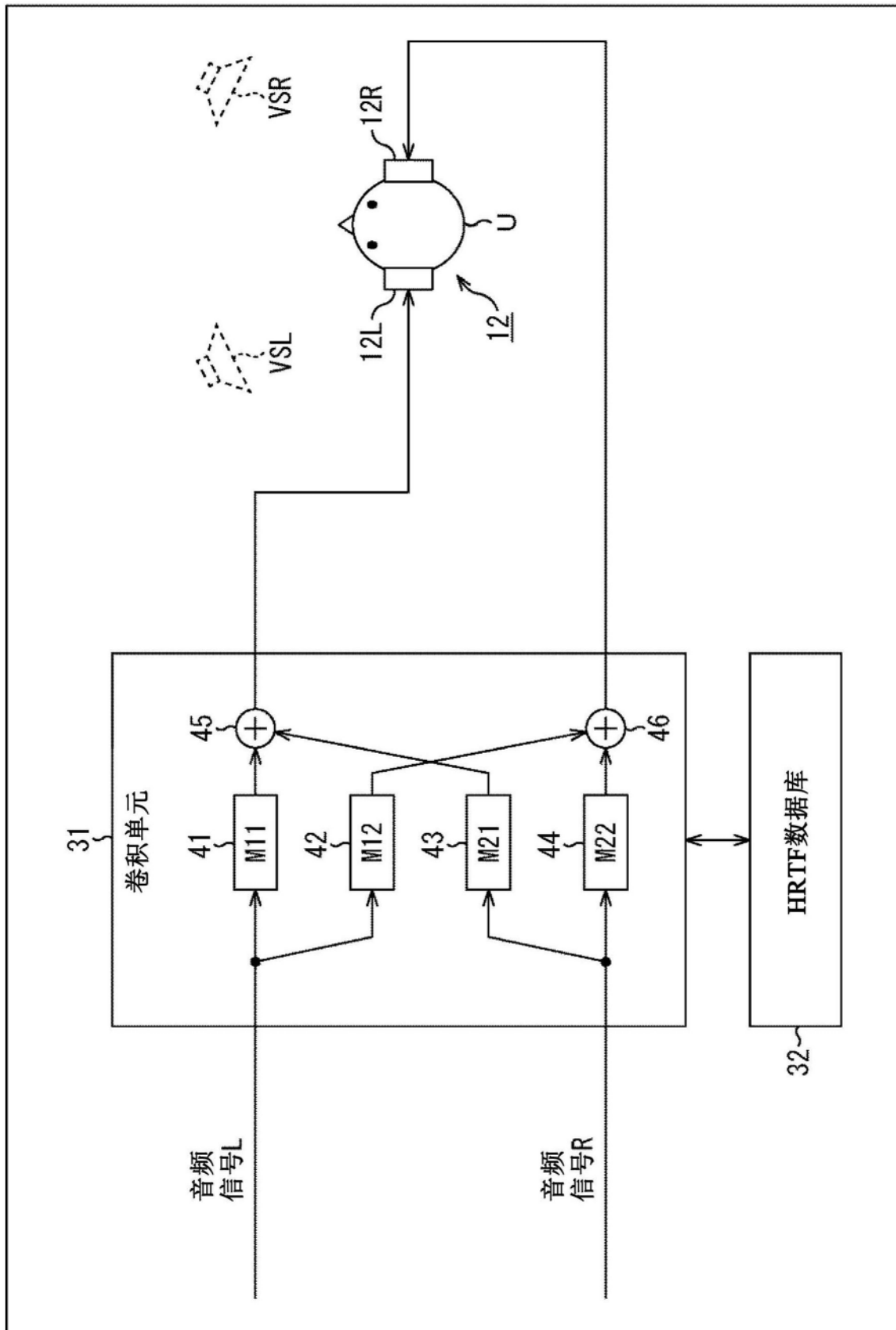


图5

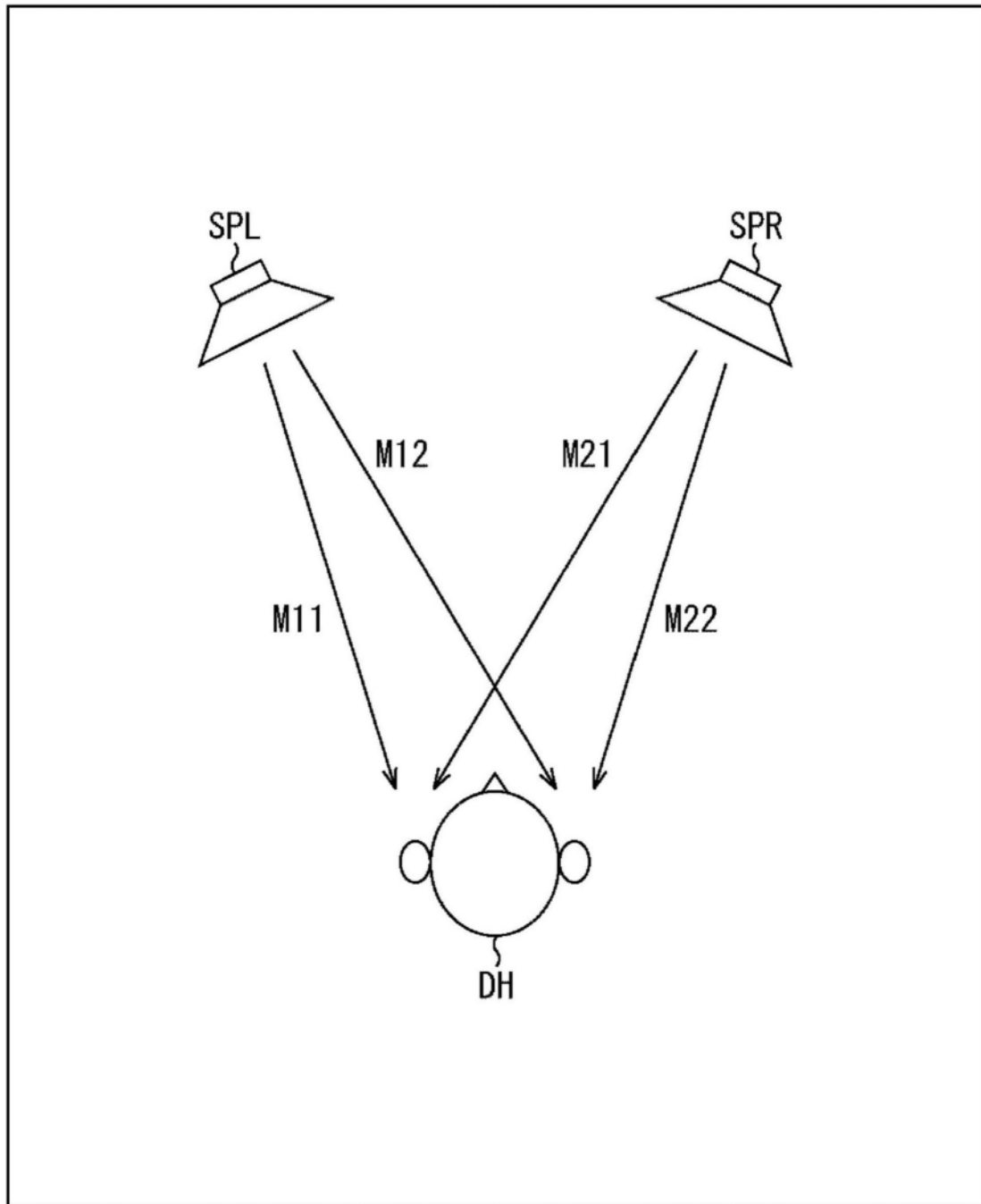


图6

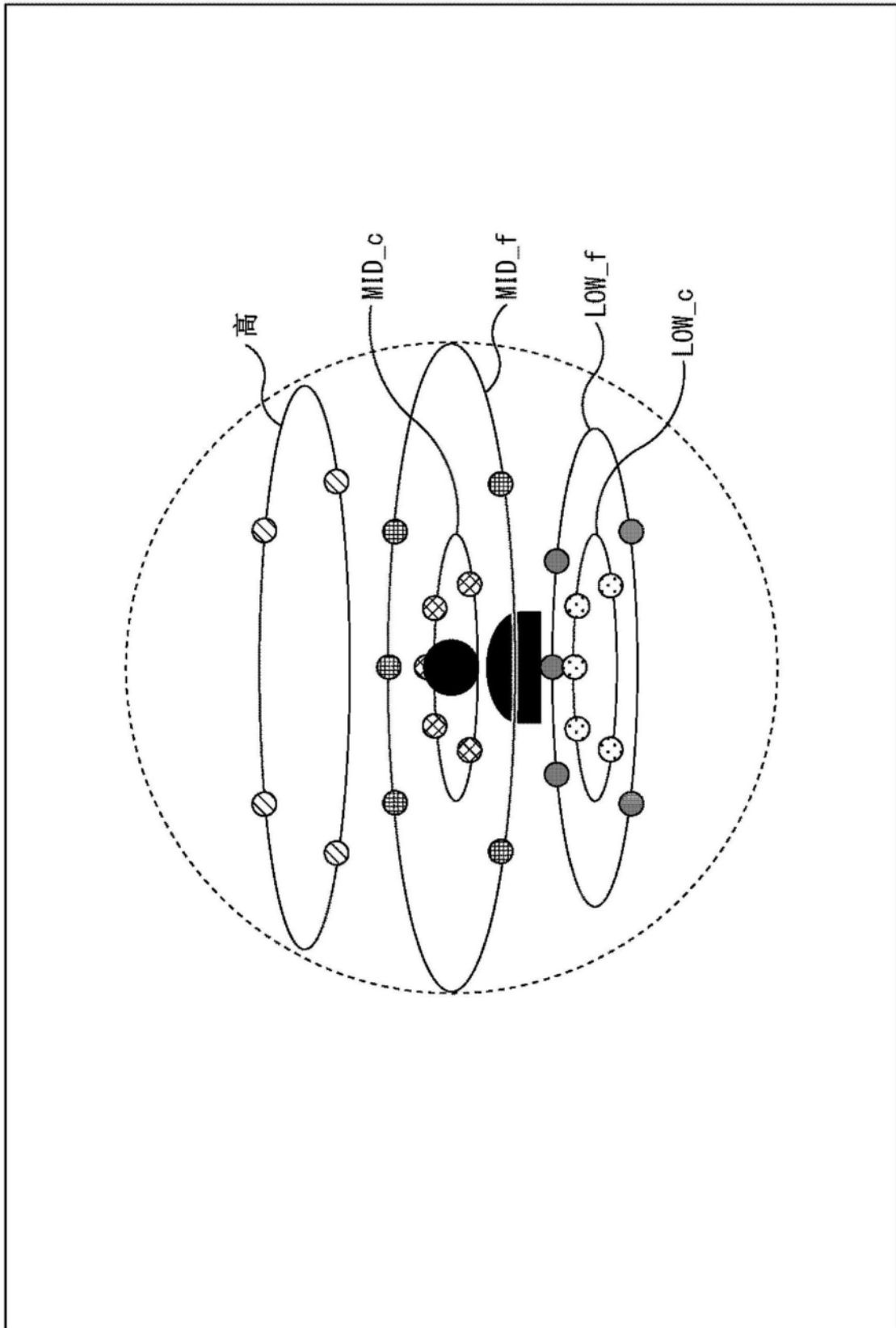


图7

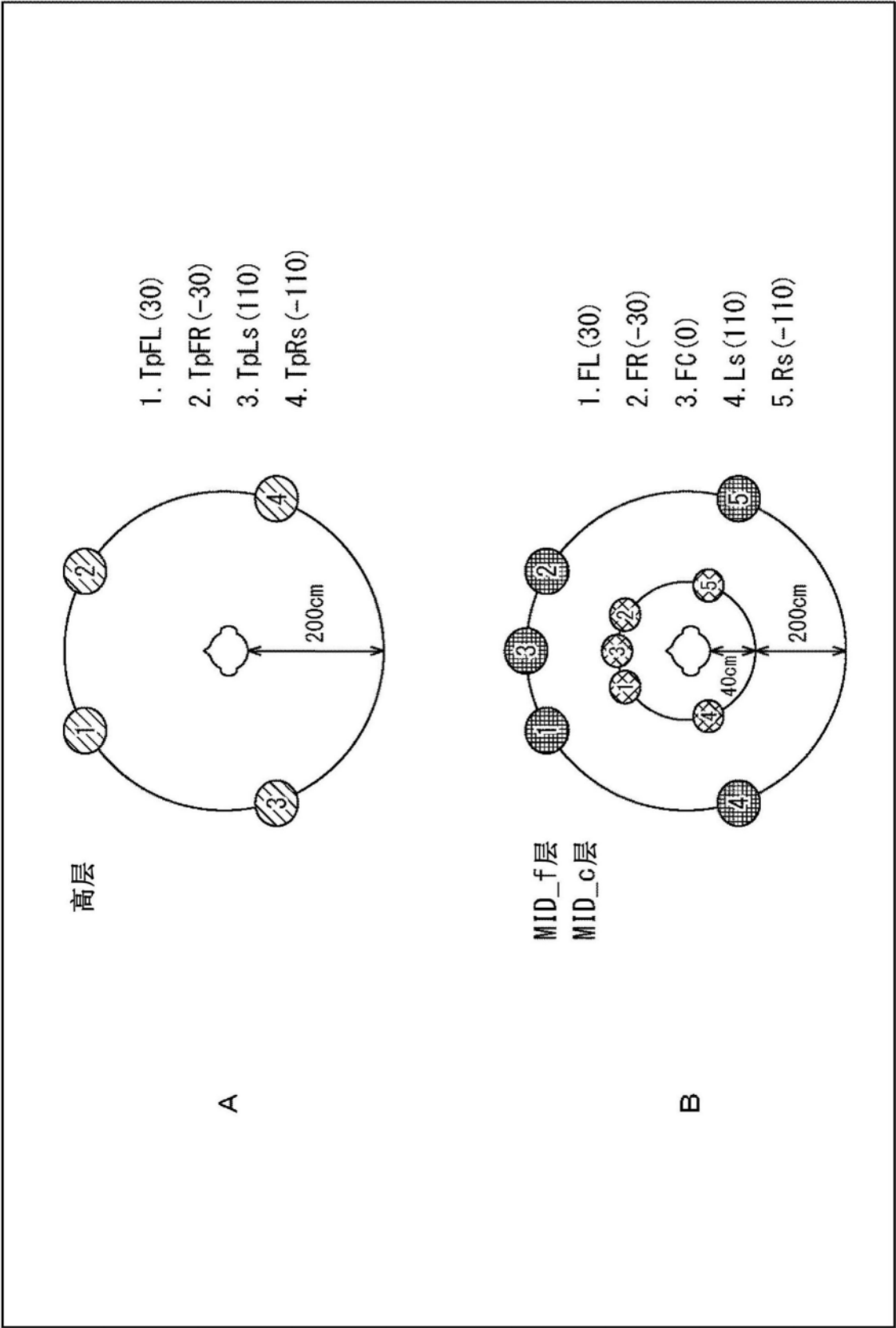


图8

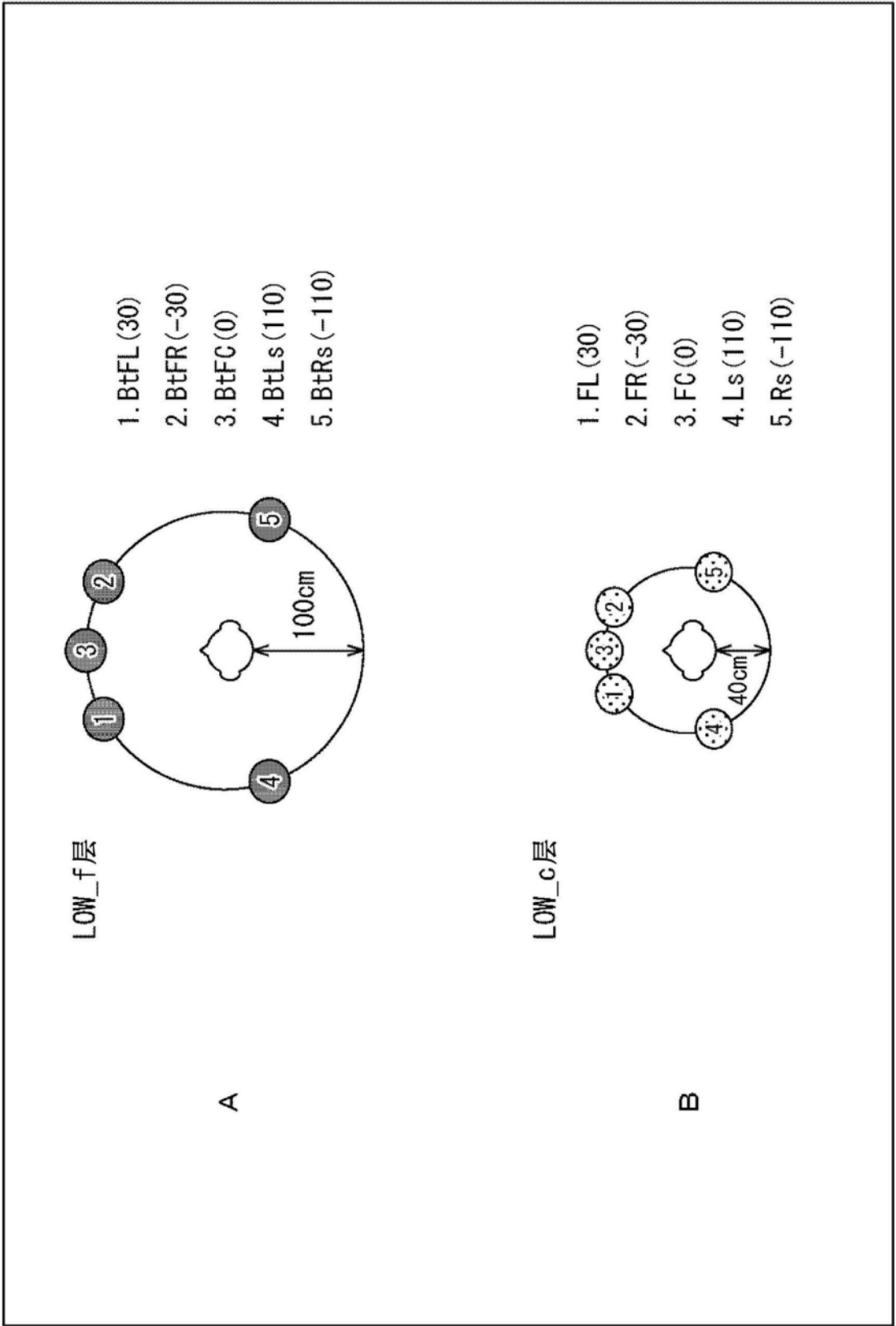


图9

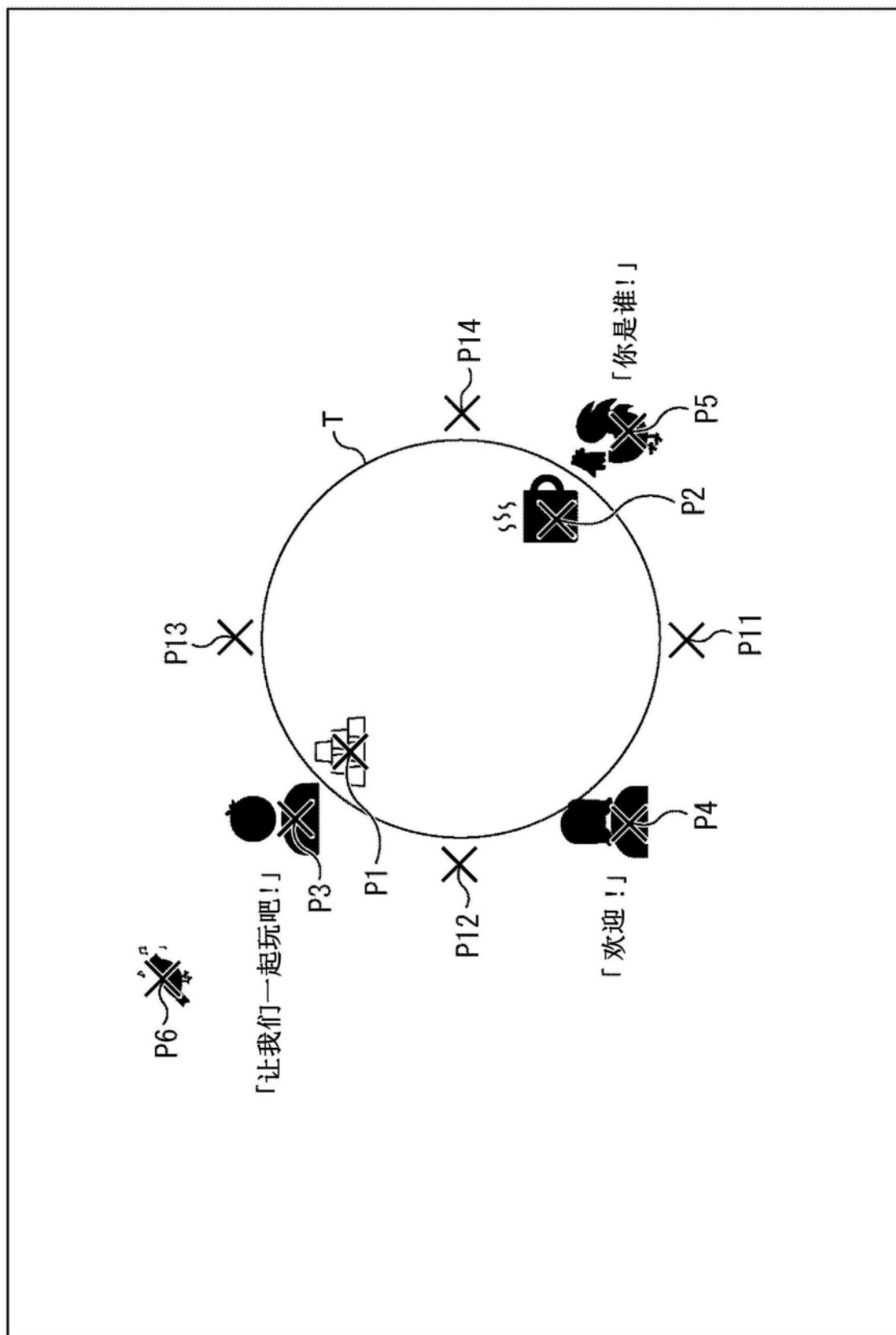


图10

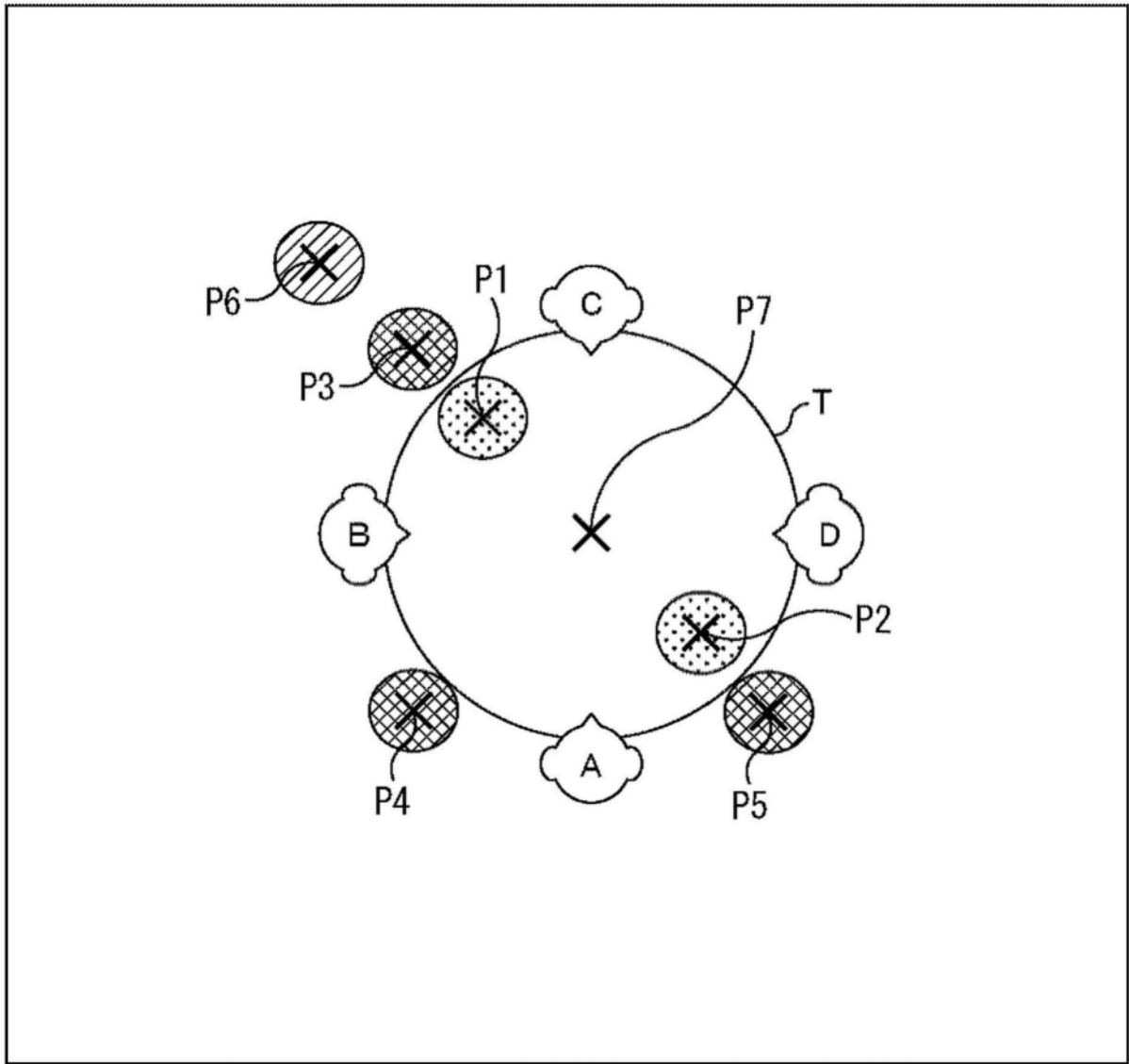


图11

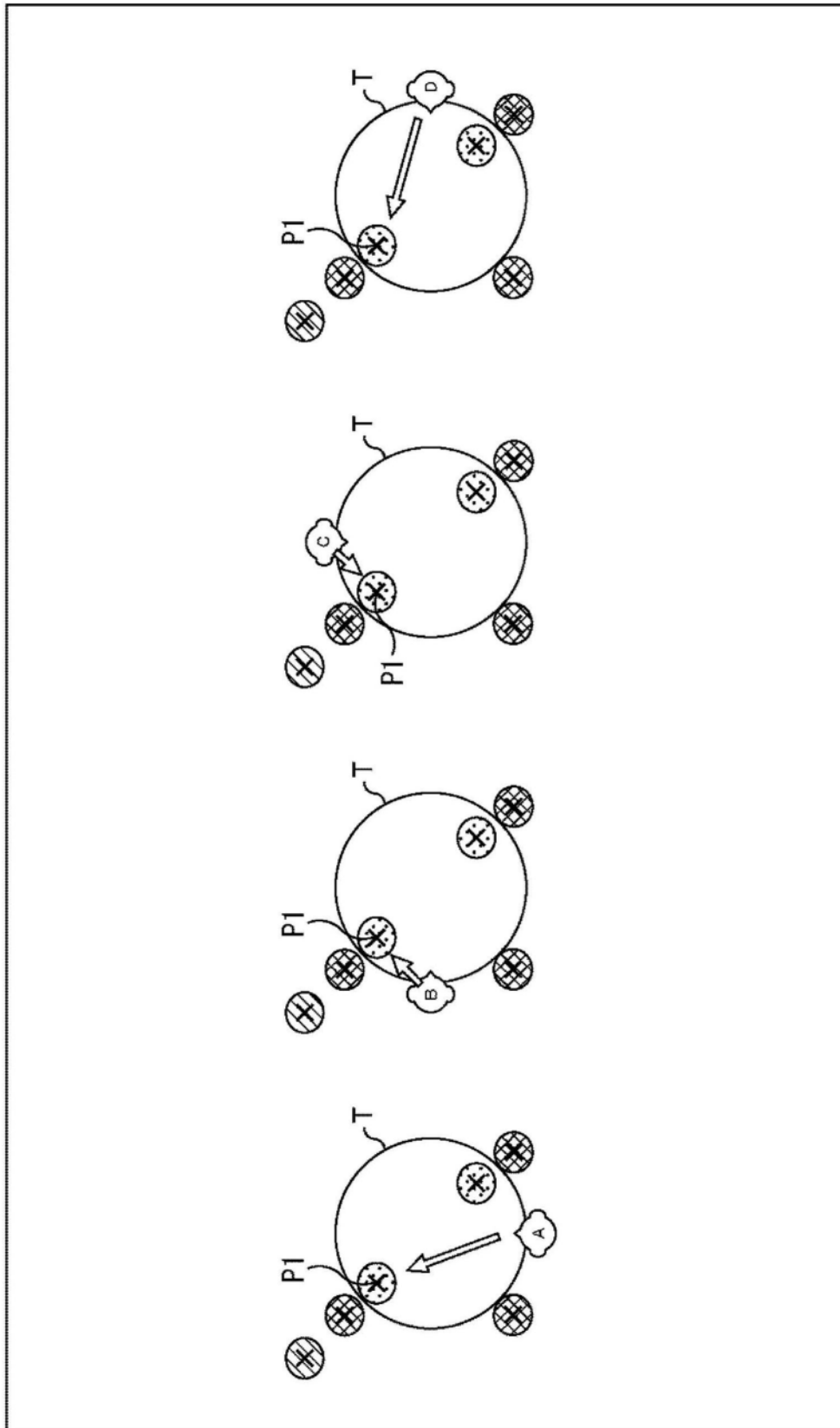


图12

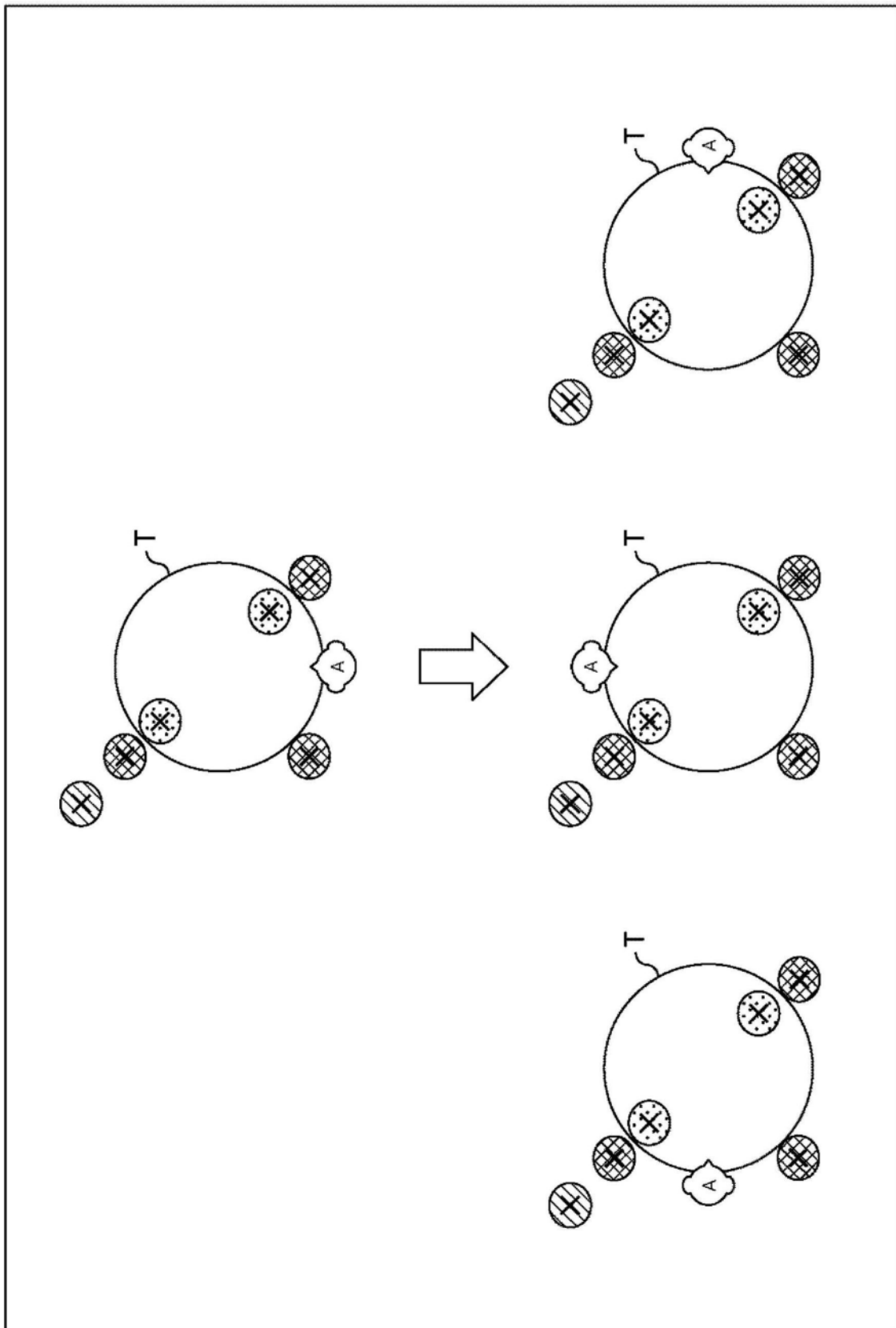


图13

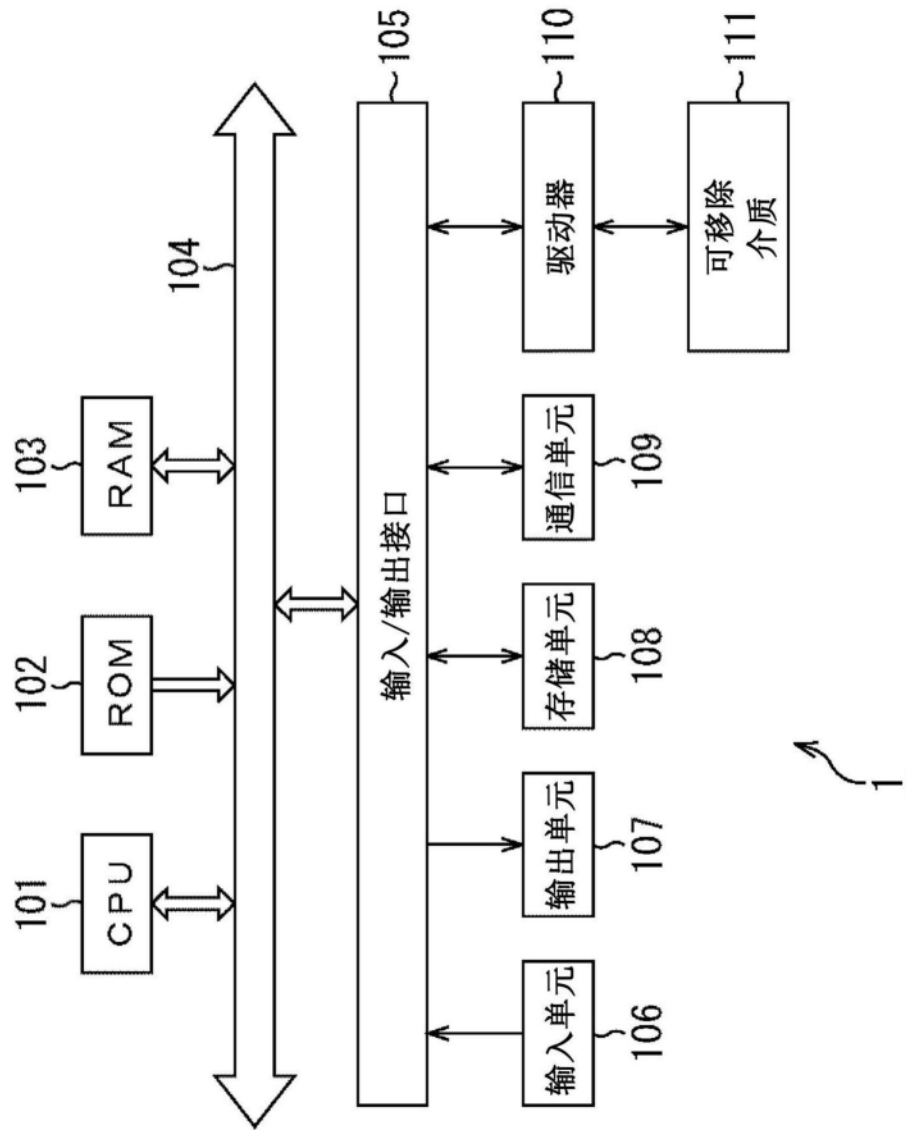


图14

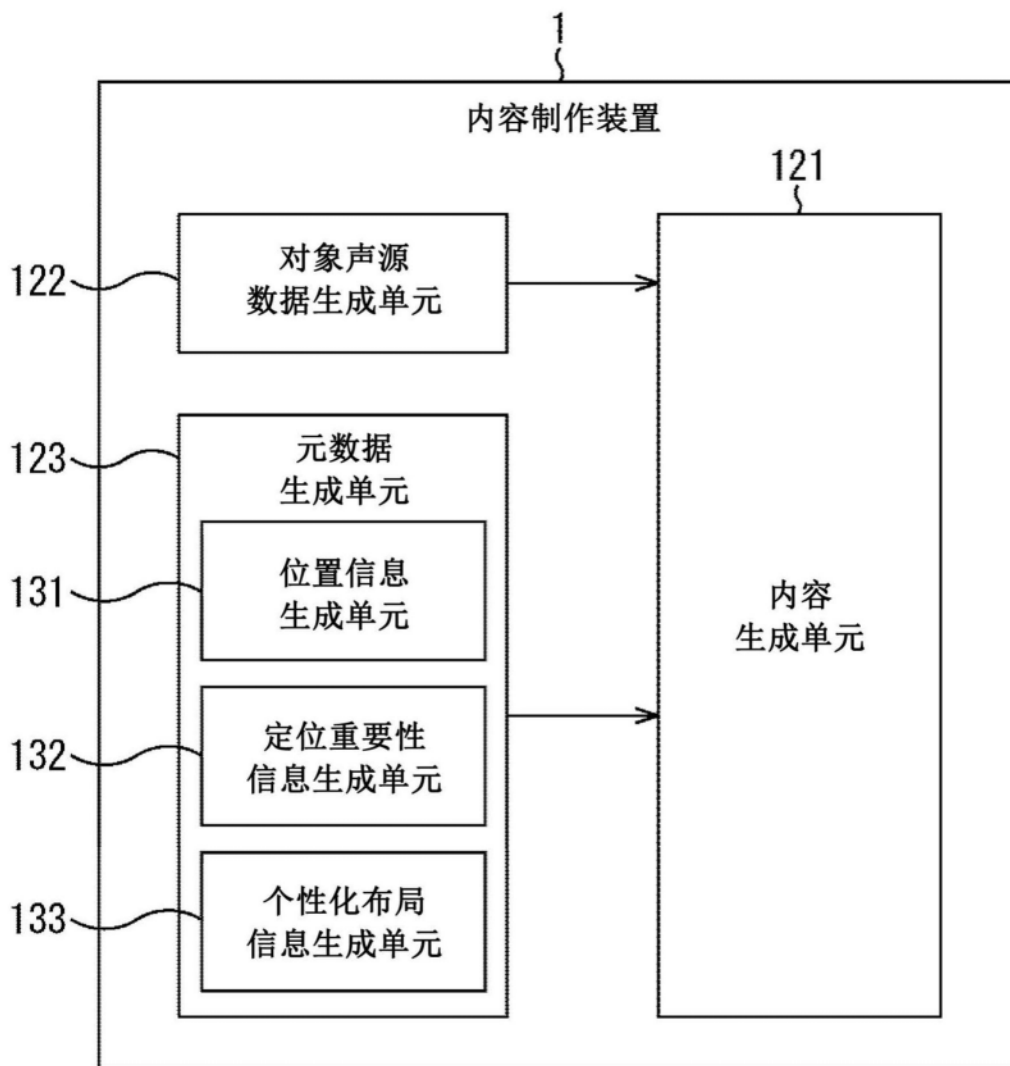


图15

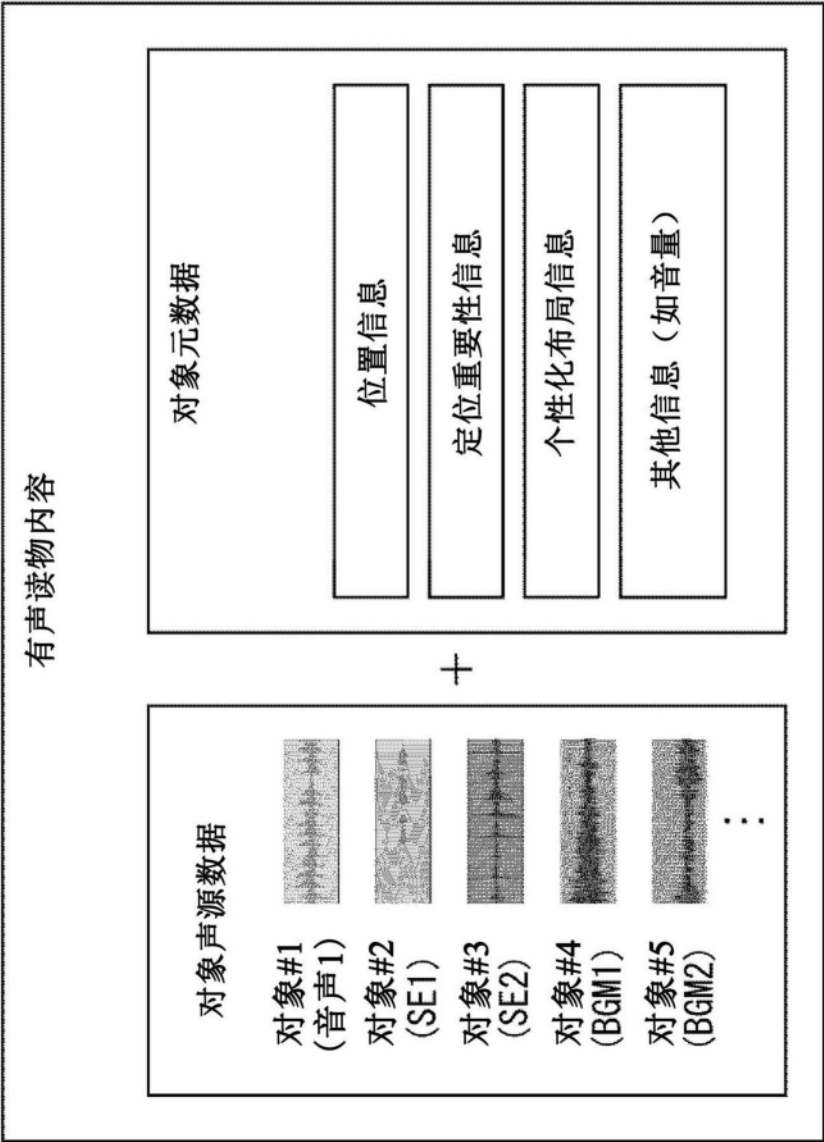


图16

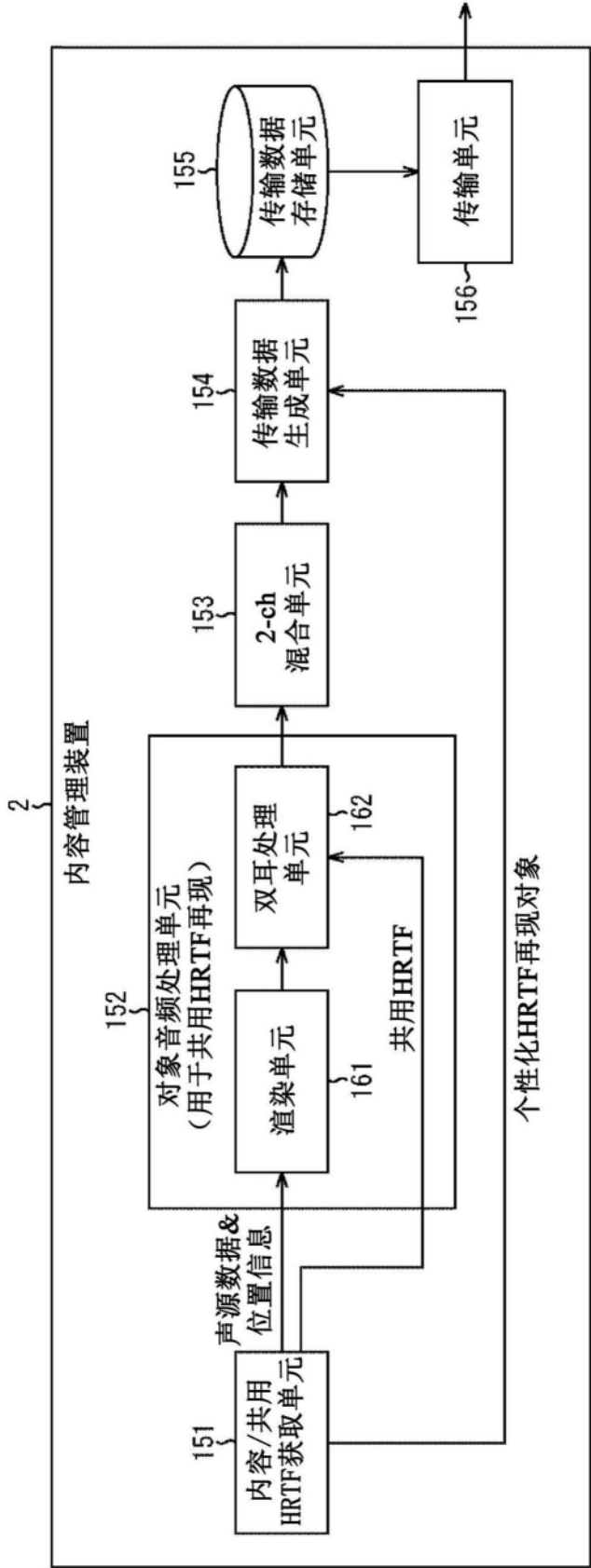


图17

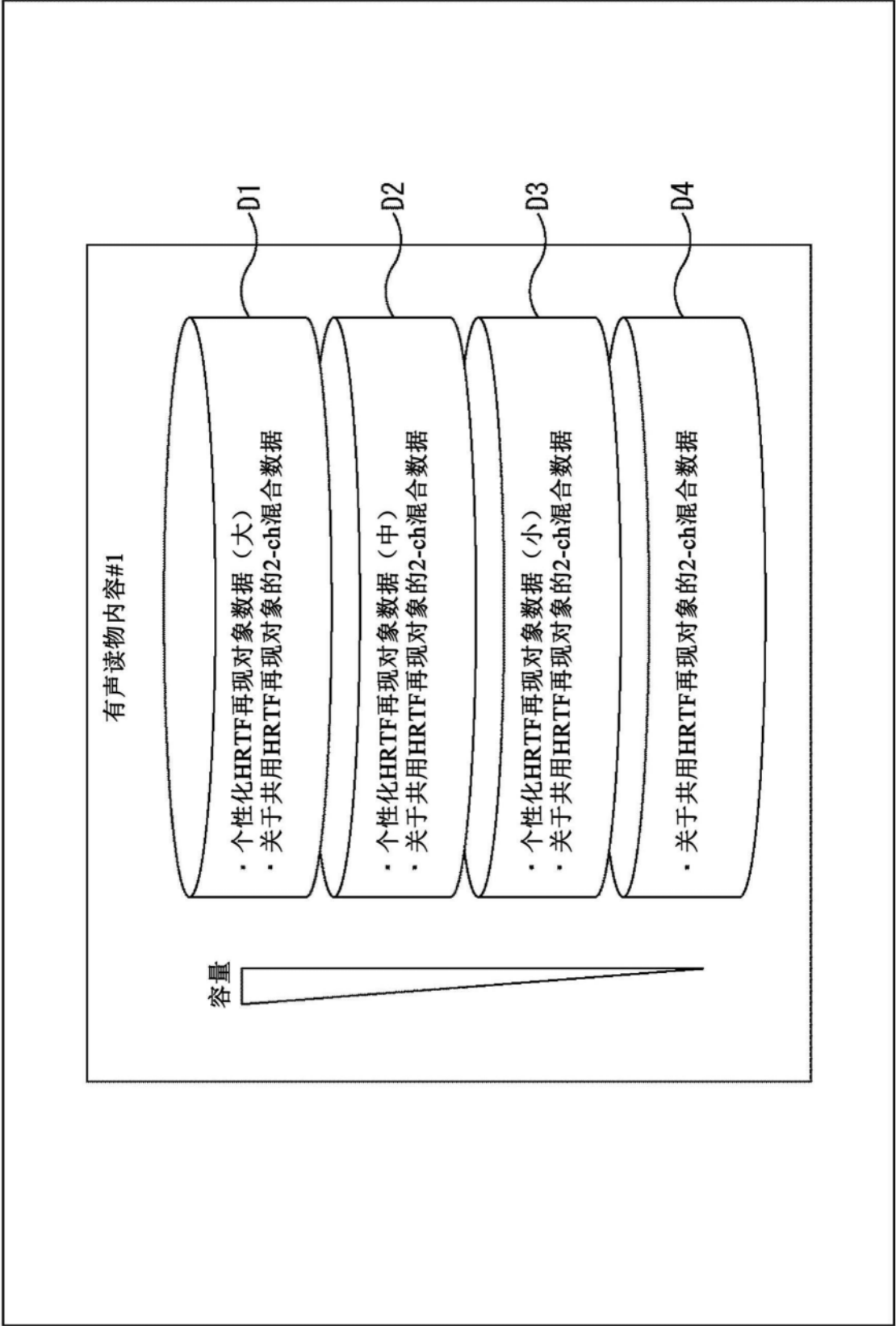


图18

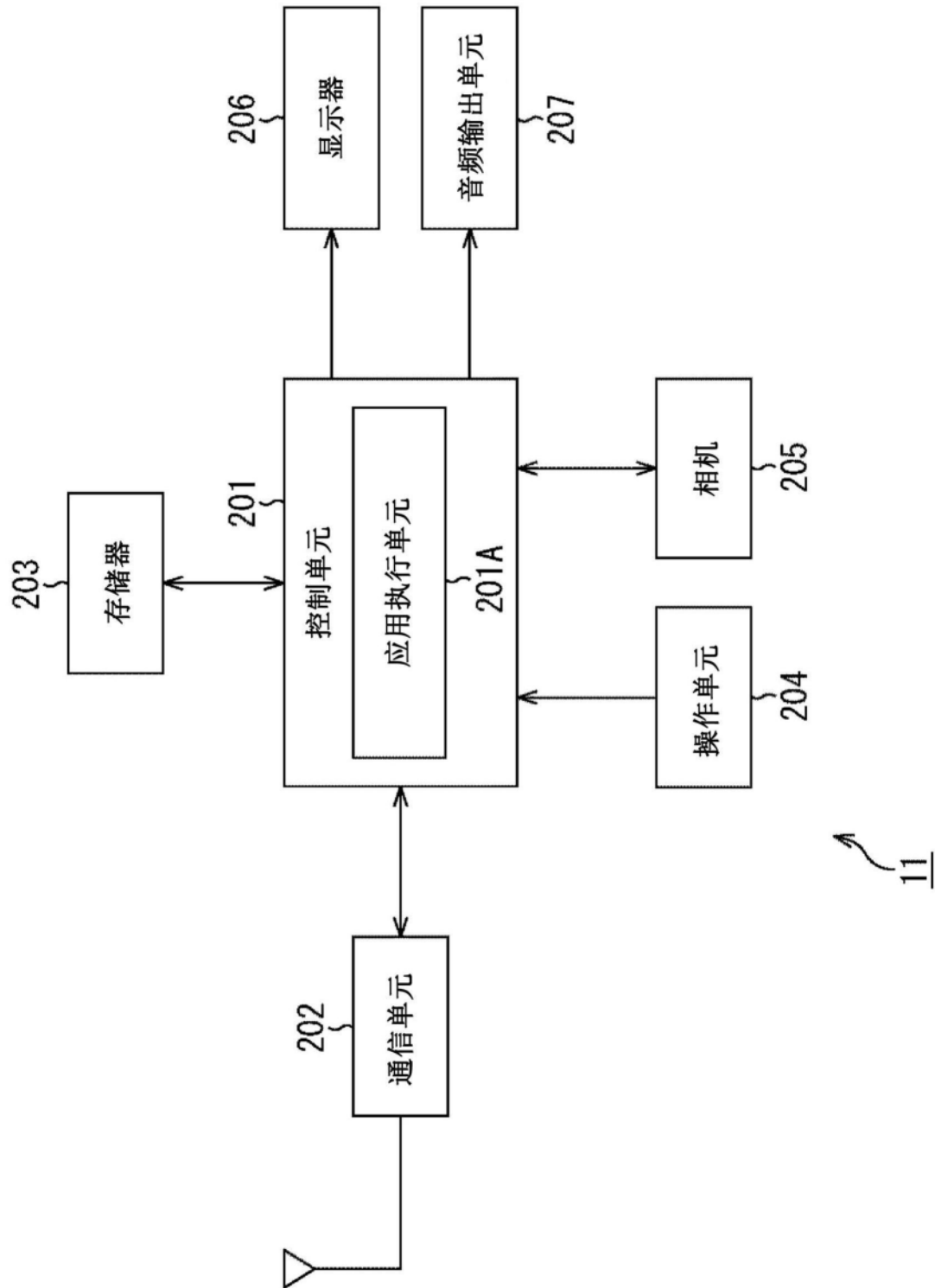


图19

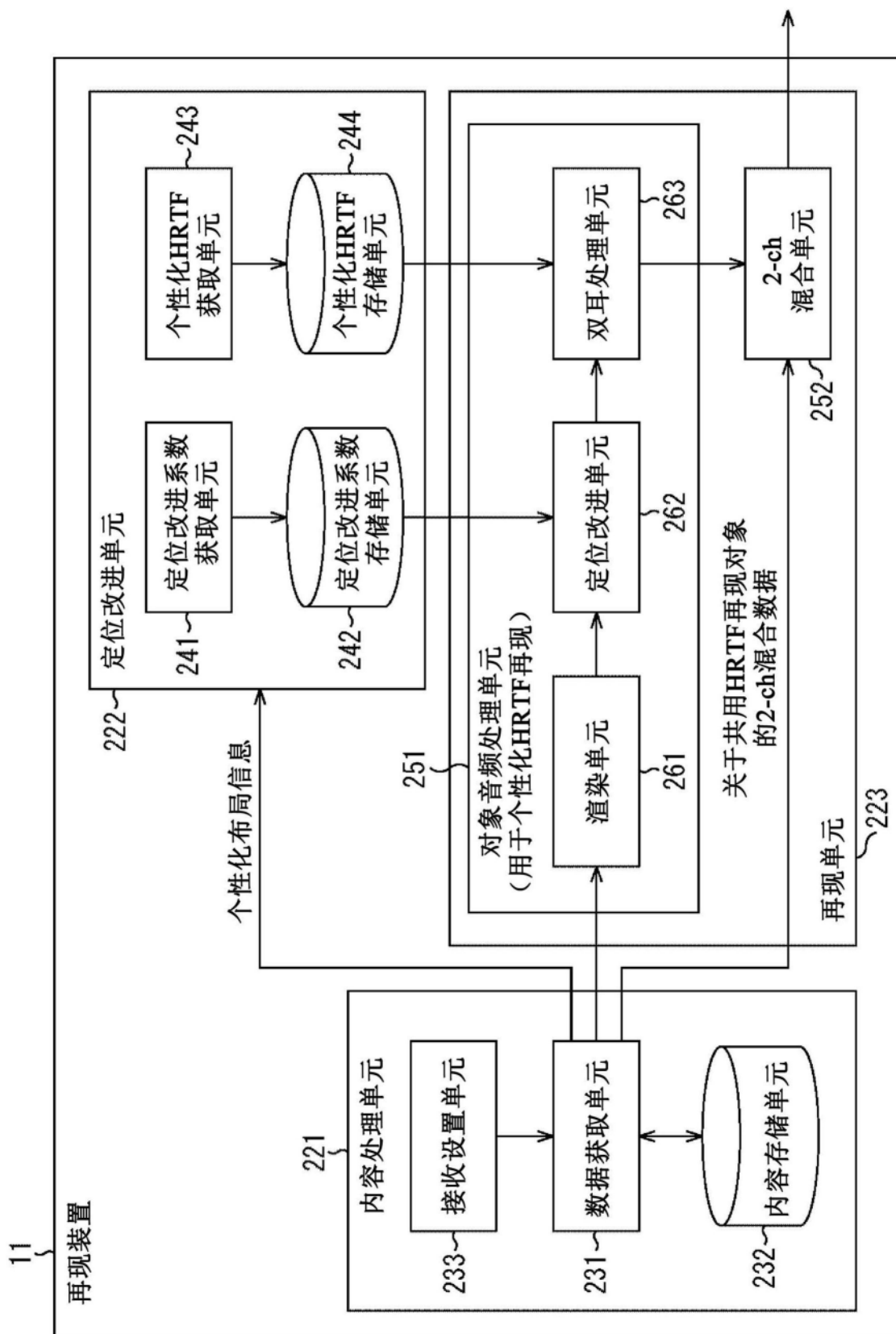


图20

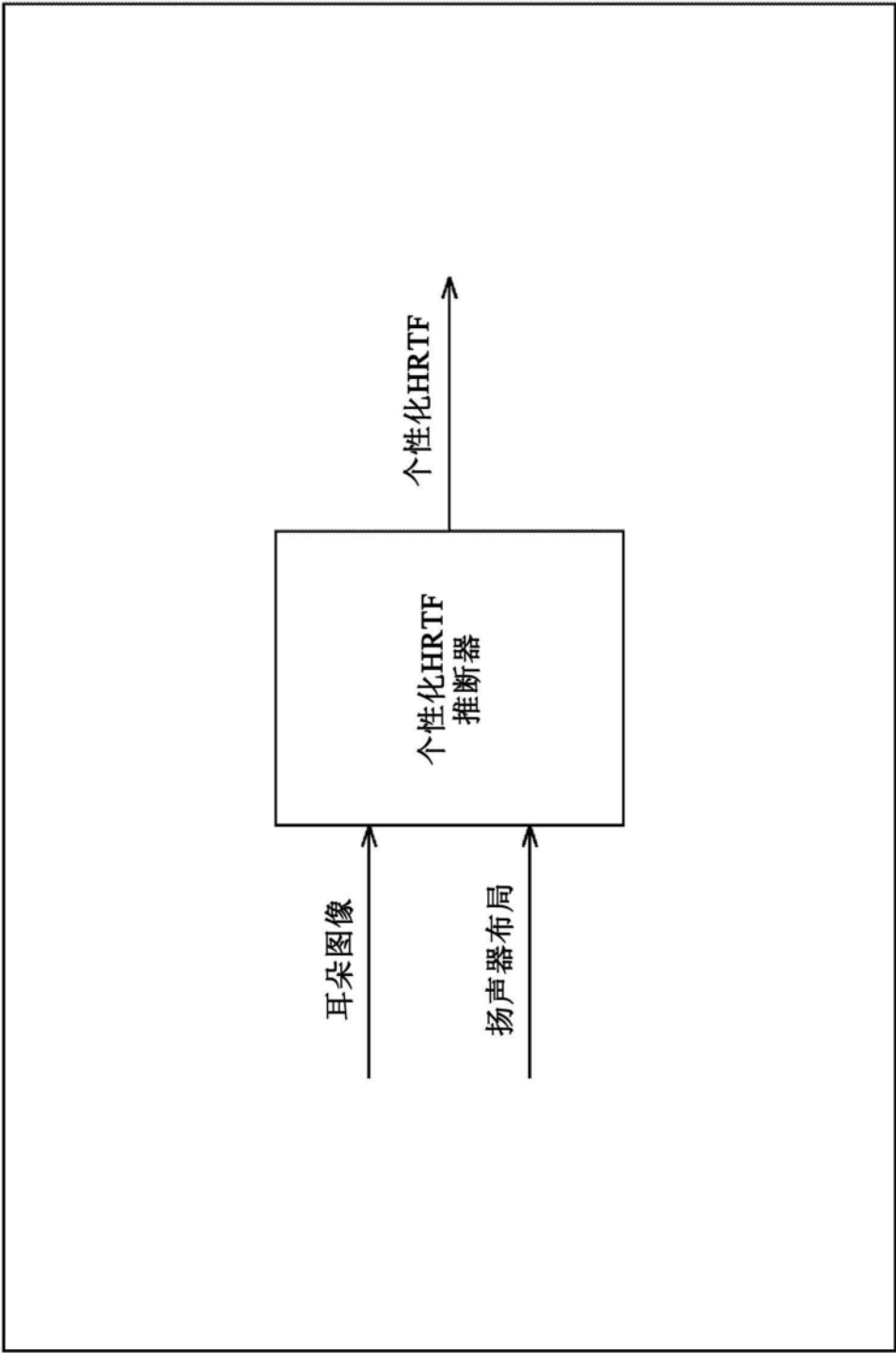


图21

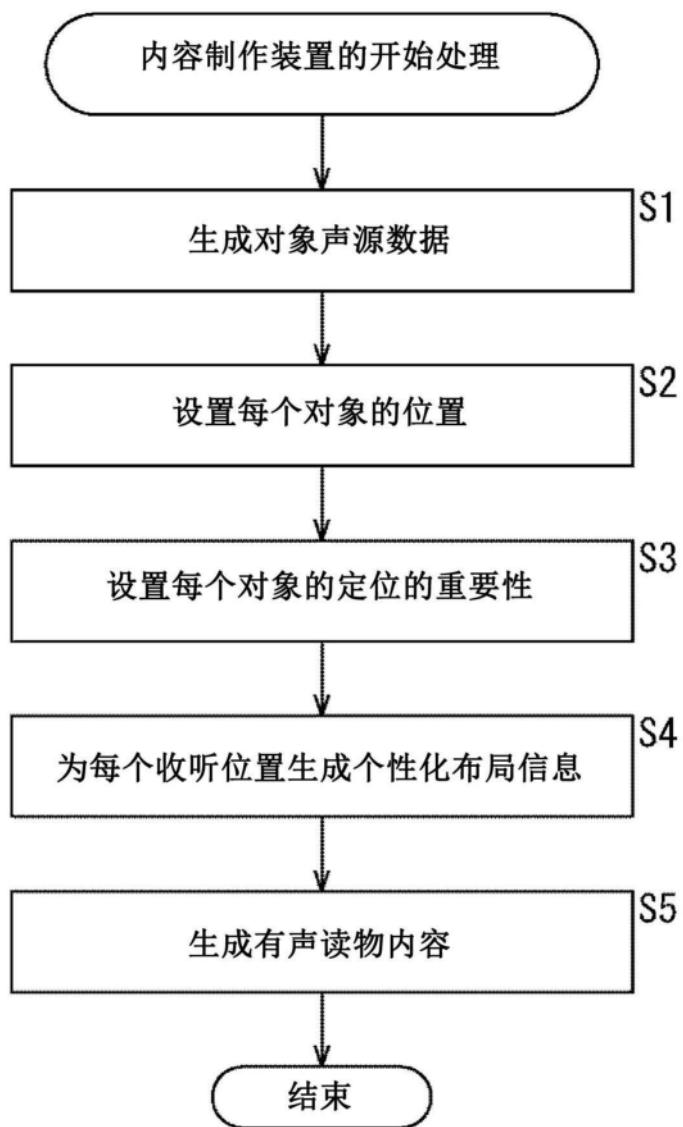


图22

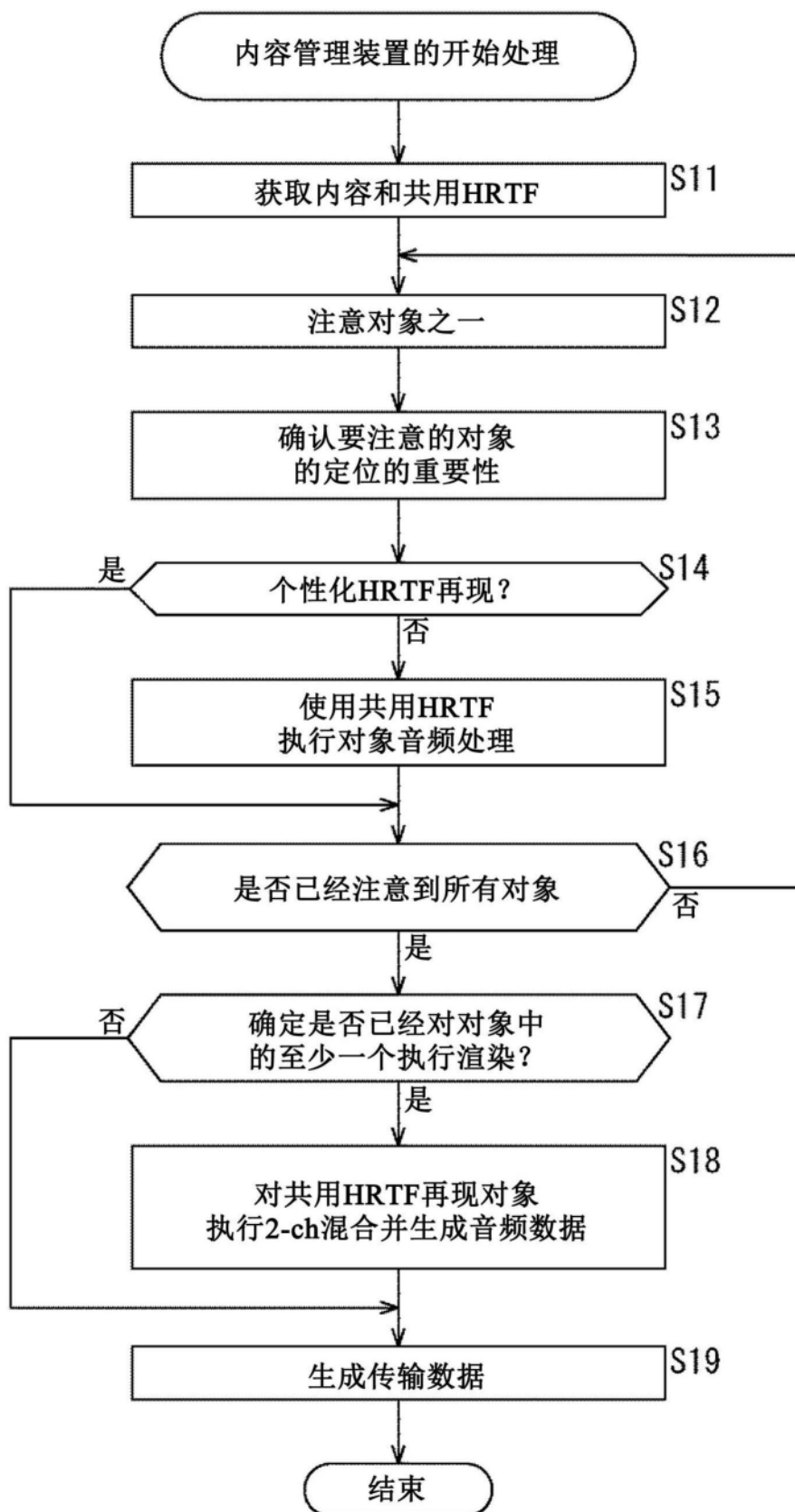


图23

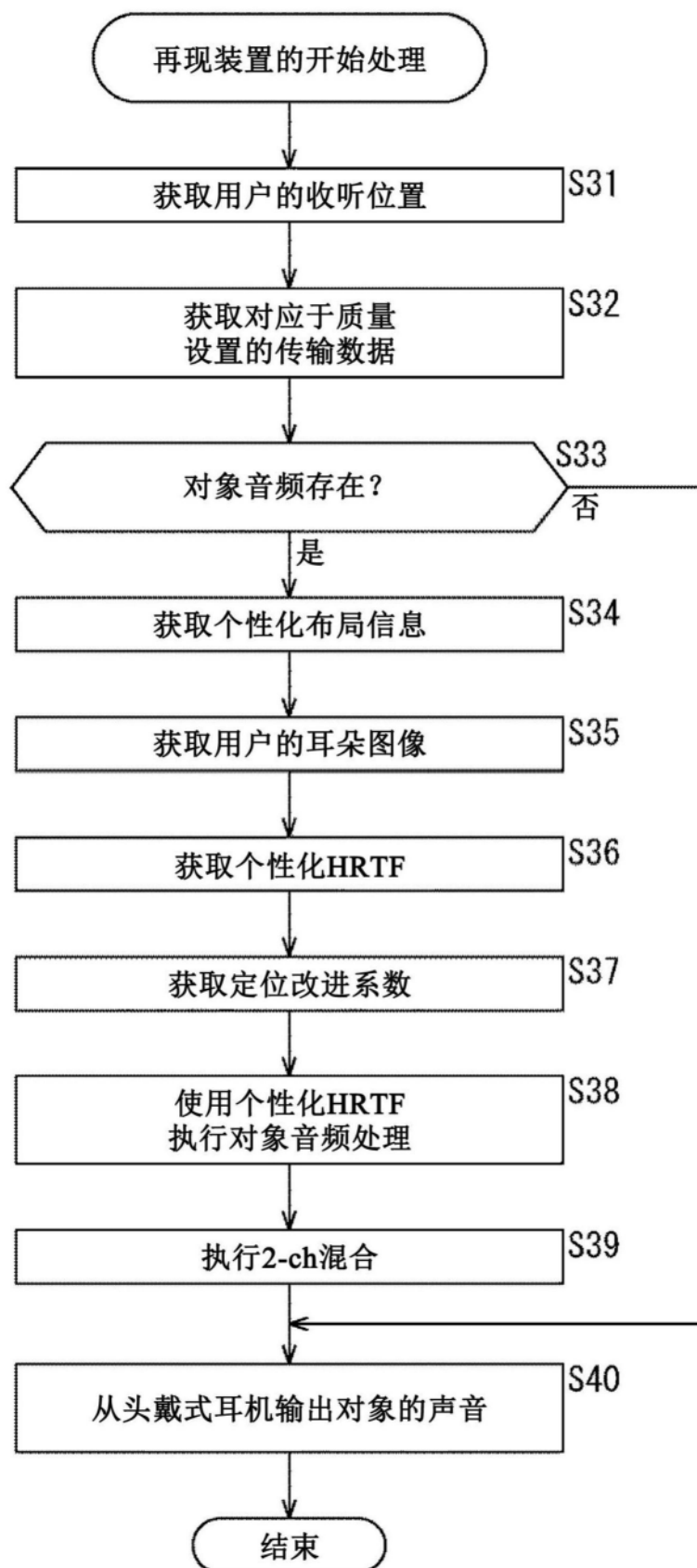


图24

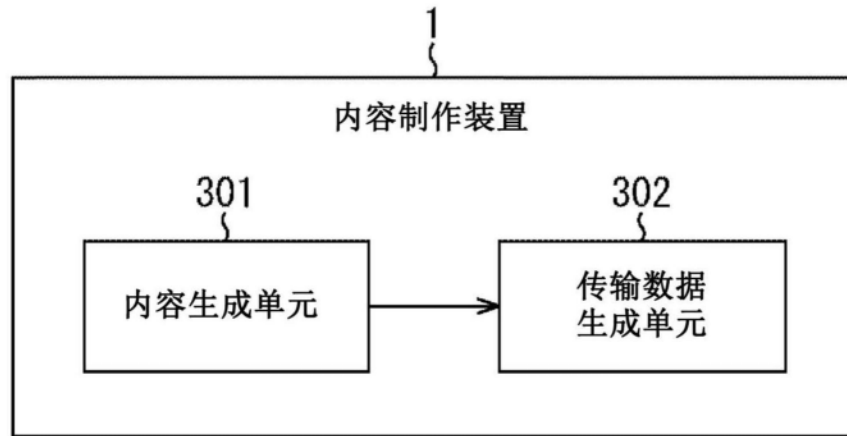


图25

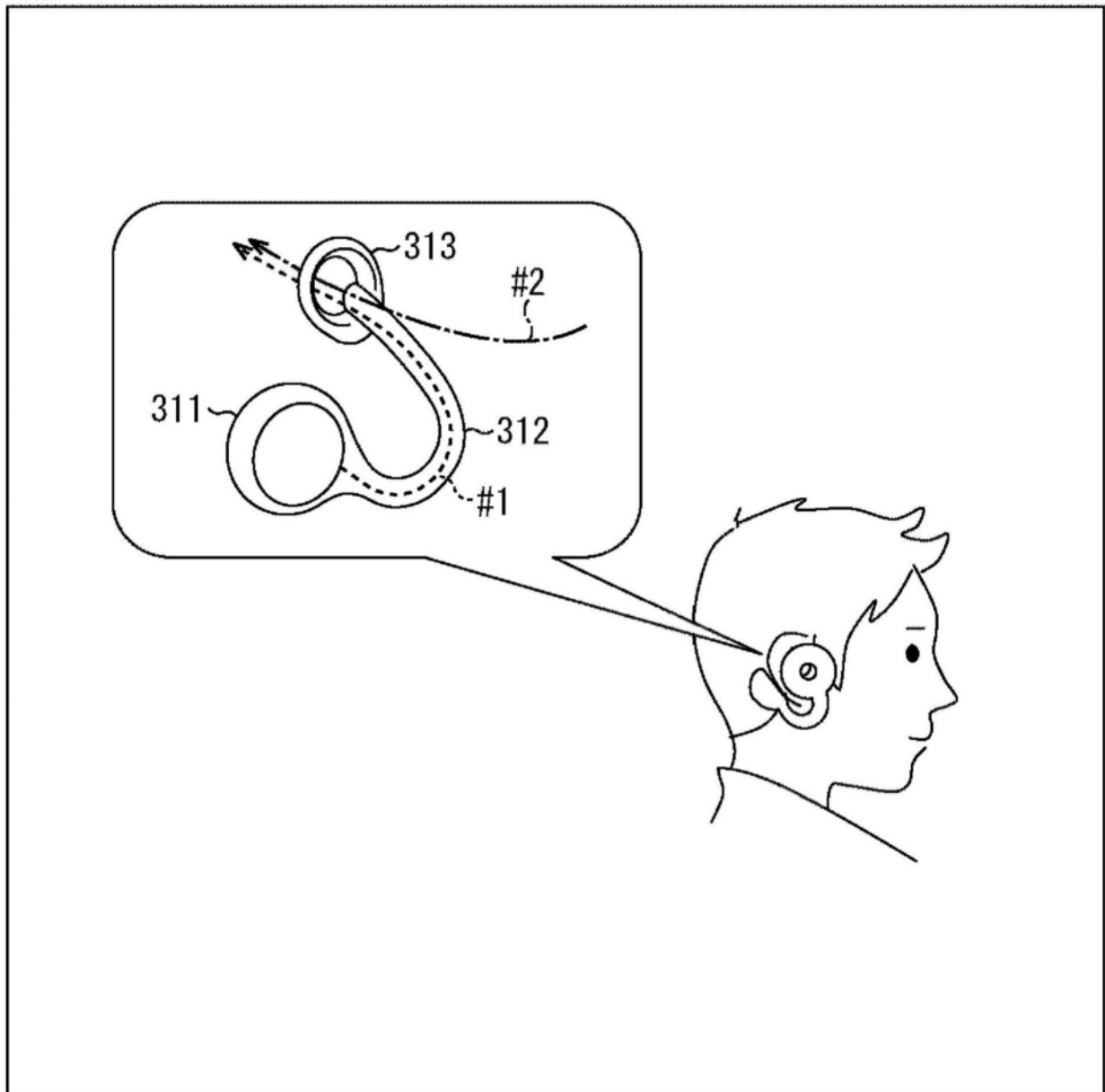


图26

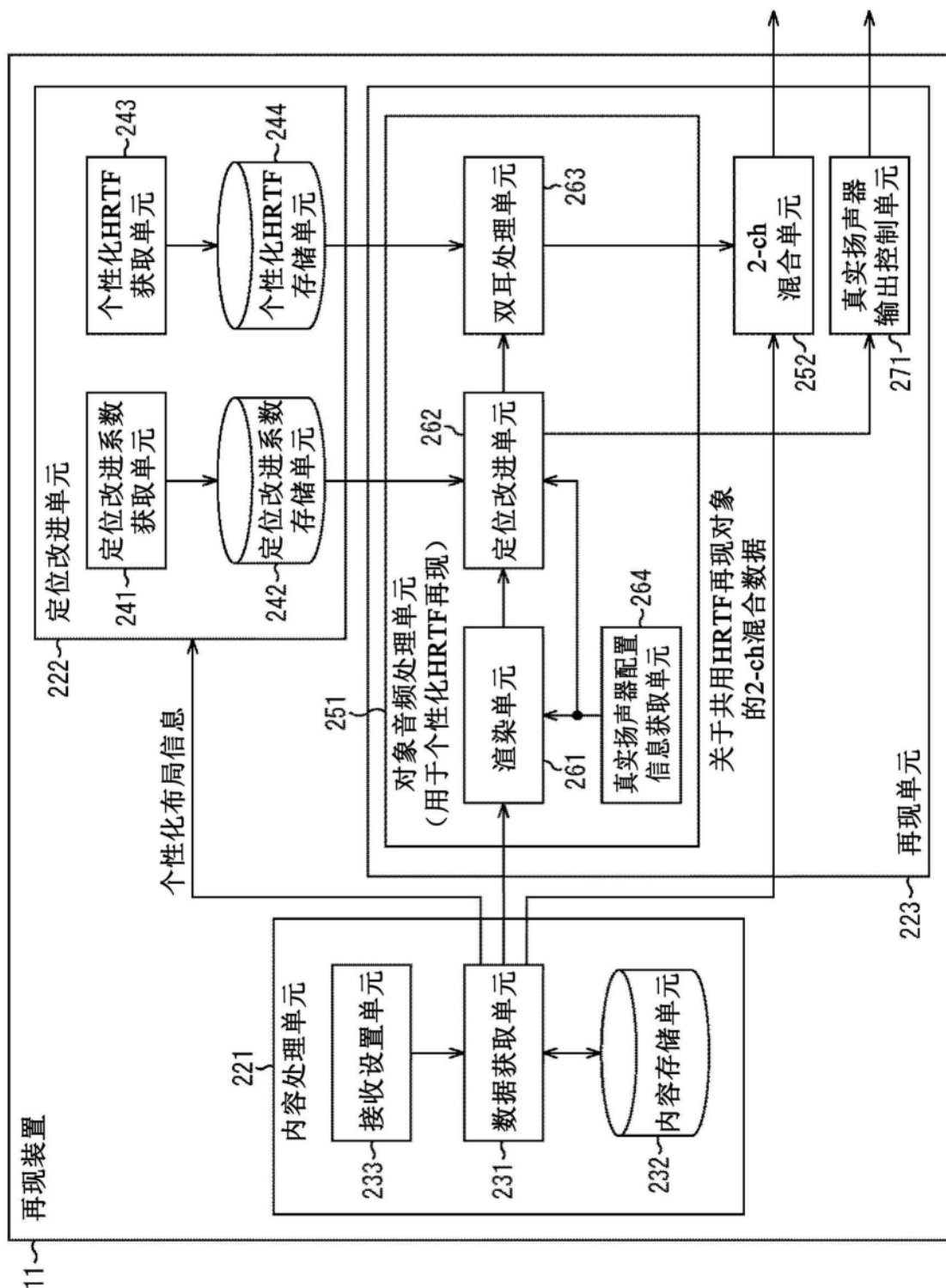


图27

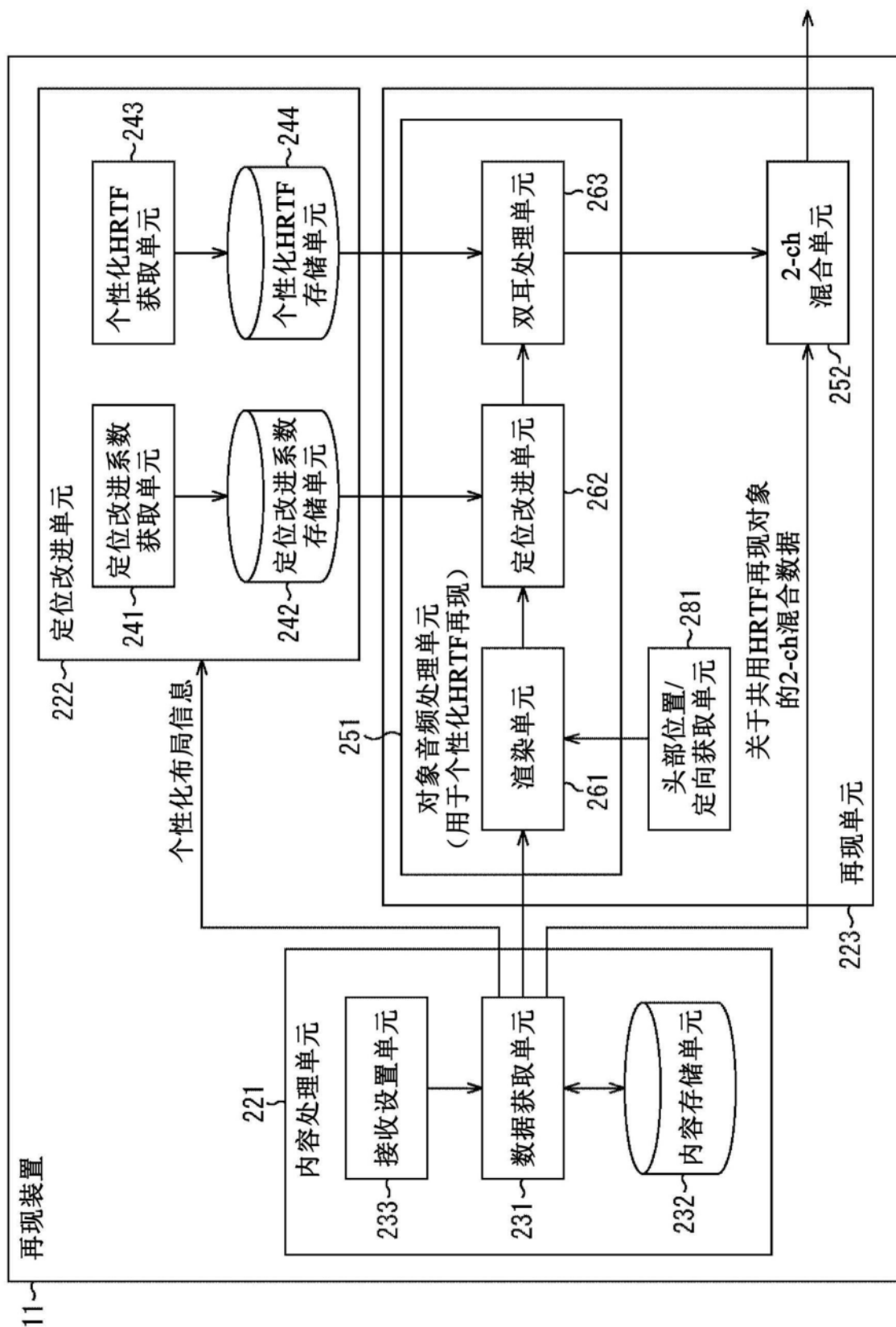


图28

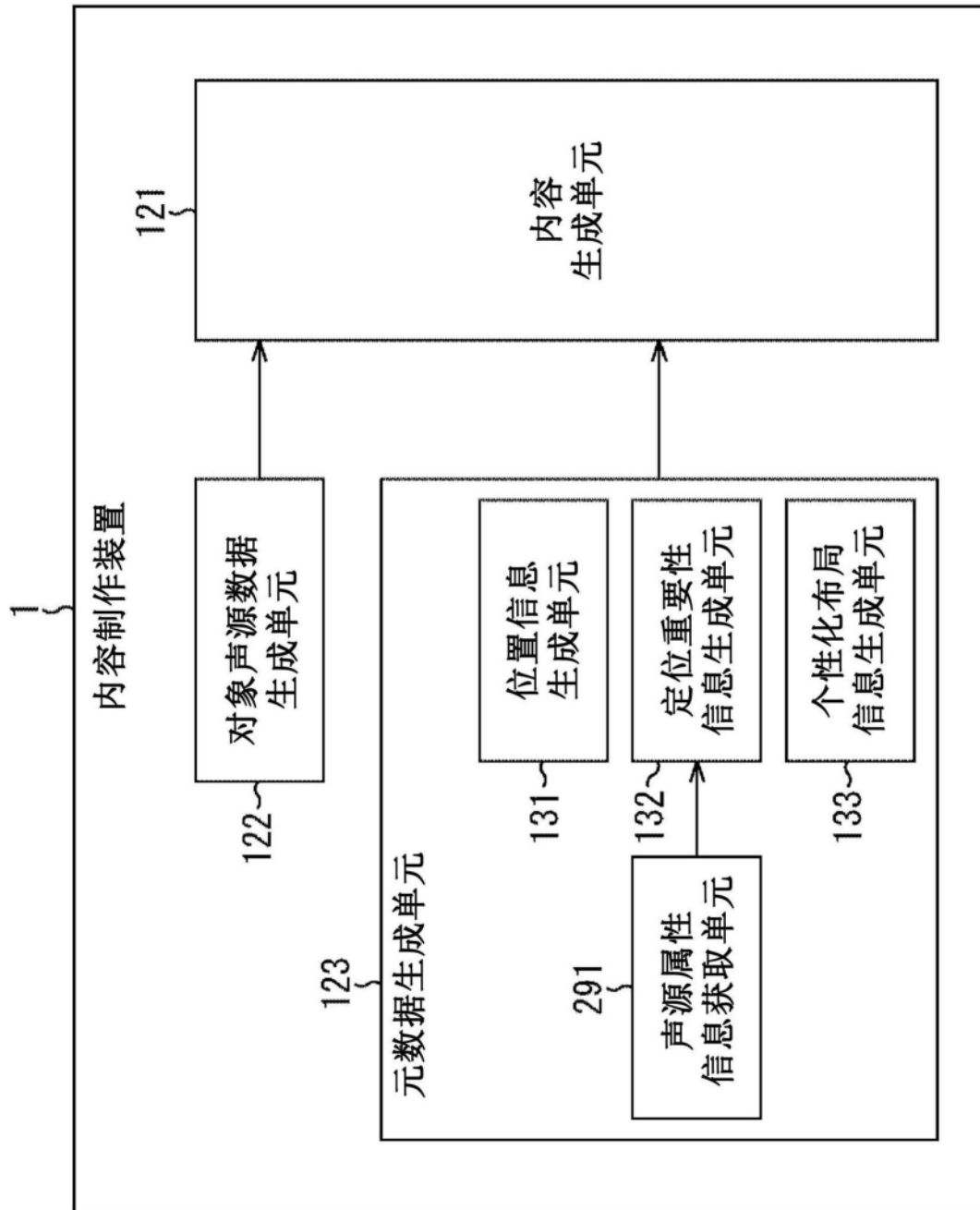


图29