



(12) 发明专利

(10) 授权公告号 CN 112055974 B

(45) 授权公告日 2022.06.10

(21) 申请号 201980029406.3

(22) 申请日 2019.02.27

(65) 同一申请的已公布的文献号
申请公布号 CN 112055974 A

(43) 申请公布日 2020.12.08

(30) 优先权数据
1803408.2 2018.03.02 GB

(85) PCT国际申请进入国家阶段日
2020.10.30

(86) PCT国际申请的申请数据
PCT/FI2019/050156 2019.02.27

(87) PCT国际申请的公布数据
W02019/166698 EN 2019.09.06

(73) 专利权人 诺基亚技术有限公司

地址 芬兰埃斯波

(72) 发明人 L·拉克索宁

(74) 专利代理机构 北京市中咨律师事务所
11247

专利代理师 杨晓光

(51) Int.Cl.

H04S 7/00 (2006.01)

H04R 3/12 (2006.01)

G06F 3/16 (2006.01)

(56) 对比文件

US 2016232901 A1, 2016.08.11

CN 104604257 A, 2015.05.06

审查员 杨梦琳

权利要求书4页 说明书14页 附图7页

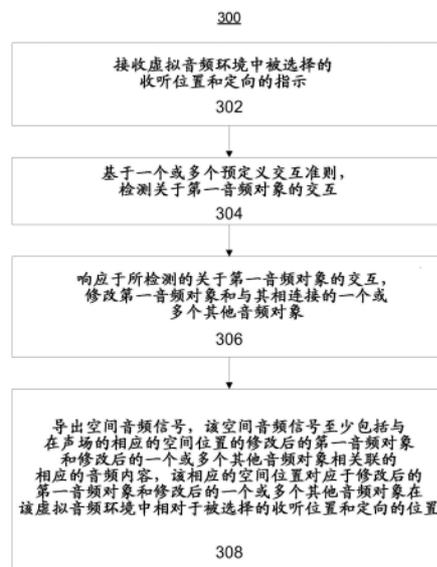
(54) 发明名称

音频处理

(57) 摘要

根据示例性实施例,提供了一种渲染空间音频信号的方法,该空间音频信号表示在包括一个或多个音频对象的可选视点音频环境中的声场,其中,每个音频对象与相应的音频内容和在该音频环境中的相应位置相关联,该方法包括:接收该音频环境中被选择的收听位置和定向的指示;基于一个或多个预定义交互准则,检测关于第一音频对象的交互;响应于所检测的关于第一音频对象的交互,修改第一音频对象和与其相连接的一个或多个其他音频对象;以及导出至少包括以下项的空间音频信号:与在该声场的第一空间位置的修改后的第一音频对象相关联的音频内容,该第一空间位置对应于修改后的第一音频对象在该音频环境中相对于所述被选择的收听位置和定向的位置;以及与在该声场的相应的其他空间位置的修改后的一个或多个其他音频对象相关联的音频内容,该相应的其他空间位置对应于修改后的一个或多个其他音频对象在该音频环境中相对于所

述被选择的收听位置和定向的位置。



1. 一种用于渲染空间音频信号的方法,所述空间音频信号表示在包括一个或多个音频对象的可选视点音频环境中的声场,其中,每个音频对象与相应的音频内容和在所述音频环境中的相应位置相关联,其中,所述渲染空间音频信号被配置为基于跟踪真实世界空间中的收听位置而改变,所述方法包括:

接收所述音频环境中被选择的收听位置和定向的指示;

基于一个或多个预定义交互准则,检测关于第一音频对象的交互,其中,所述一个或多个预定义交互准则规定所述第一音频对象与所述被选择的收听位置和/或被选择的定向之间的交互;

响应于被检测的交互,修改所述第一音频对象和与其相连接的一个或多个其他音频对象;以及

导出至少包括以下项的所述空间音频信号:

与在所述声场的第一空间位置的修改后的第一音频对象相关联的音频内容,所述第一空间位置对应于所述修改后的第一音频对象在所述音频环境中相对于所述被选择的收听位置和定向的位置;以及

与在所述声场的相应的其他空间位置的修改后的一个或多个其他音频对象相关联的音频内容,所述相应的其他空间位置对应于所述修改后的一个或多个其他音频对象在所述音频环境中相对于所述被选择的收听位置和定向的位置。

2. 根据权利要求1所述的方法,其中,所述检测包括:检测所述第一音频对象与所述一个或多个其他音频对象之间的交互。

3. 根据权利要求1所述的方法,其中,所述检测包括:检测所述第一音频对象与所述被选择的收听位置和定向之间的交互。

4. 根据权利要求3所述的方法,其中,检测所述第一音频对象与所述被选择的收听位置和定向之间的交互包括应用以下交互准则中的一个:

检测所述被选择的收听位置与所述第一音频对象的位置之间的距离低于预定义阈值距离;

检测所述被选择的收听位置与所述第一音频对象的位置之间的距离低于预定义阈值距离,以及检测所述被选择的定向朝向所述第一音频对象的位置;

检测所述被选择的收听位置与被连接到所述第一音频对象的另一音频对象的位置之间的距离低于所述预定义阈值距离;以及

检测被选择的定向朝向所述另一音频对象的位置。

5. 根据权利要求1所述的方法,其中,音频对象与相应的元数据相关联,所述元数据包括内容部分和格式部分,其中,所述内容部分包括所述音频内容,所述格式部分包括一个或多个渲染参数,所述一个或多个渲染参数定义依赖于所述音频对象在所述音频环境中相对于所述被选择的收听位置和定向的位置的所述音频内容的期望渲染。

6. 根据权利要求1所述的方法,其中,所述修改包括:

响应于所述被检测的交互,识别要被应用于所述第一音频对象的第一修改,并对所述第一音频对象应用所述第一修改;

响应于所述被检测的交互,识别要被修改的一个或多个其他音频对象,识别要被应用于相应的其他音频对象的一个或多个其他修改;以及

对所述相应的其他音频对象应用所述一个或多个其他修改，
其中，所述第一修改和所述一个或多个其他修改中的每一个包括以下中的一个：
激活所述第一音频对象和所述一个或多个其他音频对象中的相应的音频对象；
去激活所述第一音频对象和所述一个或多个其他音频对象中的相应的音频对象；
改变源自所述第一音频对象和所述一个或多个其他音频对象中的相应的音频对象的
音频内容的相对幅度；以及

改变所述第一音频对象和所述一个或多个音频对象中的相应的音频对象在所述音频
环境中的位置。

7. 根据权利要求6所述的方法，其中，所述第一音频对象与第一交互元数据相关联，并
且其中，所述一个或多个其他音频对象与相应的其他交互元数据相关联，其中，

所述第一交互元数据包括所述一个或多个预定义交互准则的定义、要被应用于所述第
一音频对象的第一修改的定义、以及第一其他音频对象的标识；以及

第一其他交互元数据包括要被应用于所述第一其他音频对象的第一其他修改的定义。

8. 根据权利要求6所述的方法，其中，所述音频环境包括与控制元数据相关联的音频控
制对象，其中，所述第一音频对象与第一交互元数据相关联，并且其中，所述一个或多个其
他音频对象与相应的其他交互元数据相关联，其中，

所述控制元数据包括所述一个或多个预定义交互准则的定义、以及所述第一音频对象
的标识；

所述第一交互元数据包括要被应用于所述第一音频对象的第一修改的定义、以及第一
其他音频对象的标识；以及

第一其他交互元数据包括要被应用于所述第一其他音频对象的第一其他修改的定义。

9. 根据权利要求7或8所述的方法，其中，

所述第一其他交互元数据包括第二其他音频对象的标识；以及

第二其他交互元数据包括要被应用于所述第二其他音频对象的第二其他修改的定义。

10. 根据权利要求6所述的方法，其中，所述第一音频对象与第一交互元数据相关联，并
且其中，所述第一交互元数据包括：所述一个或多个预定义交互准则的定义、要被应用于所
述第一音频对象的第一修改的定义、所述一个或多个其他音频对象的相应的标识、以及要
被应用于所述一个或多个其他音频对象的相应的一个或多个其他修改的相应定义。

11. 根据权利要求6所述的方法，其中，所述音频环境包括与控制元数据相关联的音频
控制对象，并且其中，所述第一音频对象与第一交互元数据相关联，其中，所述控制元数据
包括：所述一个或多个预定义交互准则的定义、要被应用于所述第一音频对象的第一修改
的定义、所述一个或多个其他音频对象的相应的标识、以及要被应用于所述一个或多个其
他音频对象的相应的一个或多个其他修改的相应定义。

12. 根据权利要求1所述的方法，其中，接收所述被选择的收听位置和定向的所述指示
包括接收定义以下中的一个或多个的用户输入：

所述音频环境中被选择的位置；

所述音频环境中当前位置的改变；

所述音频环境中相对于一个或多个预定义参考方向的被选择的定向；

所述音频环境中当前定向相对于一个或多个预定义参考方向的改变。

13. 一种用于渲染空间音频信号的装置,所述空间音频信号表示在包括一个或多个音频对象的可选视点音频环境中的声场,其中,每个音频对象与相应的音频内容和在所述音频环境中的相应位置相关联,其中,所述渲染空间音频信号被配置为基于跟踪真实世界空间中的收听位置而改变,所述装置包括:

至少一个处理器;以及

包括用于一个或多个程序的计算机代码的至少一个存储器,所述至少一个存储器和所述计算机代码被配置为与所述至少一个处理器一起使所述装置至少:

接收音频环境中被选择的收听位置和定向的指示;

基于一个或多个预定义交互准则,检测关于第一音频对象的交互,其中,所述一个或多个预定义交互准则规定所述第一音频对象与所述被选择的收听位置和/或被选择的定向之间的交互;

响应于被检测的交互,修改所述第一音频对象和与其相连接的一个或多个其他音频对象;以及

导出至少包括以下项的空间音频信号:

与在所述声场的第一空间位置的修改后的第一音频对象相关联的音频内容,所述第一空间位置对应于所述修改后的第一音频对象在所述音频环境中相对于所述被选择的收听位置和定向的位置;以及

与在所述声场的相应的其他空间位置的修改后的一个或多个其他音频对象相关联的音频内容,所述相应的其他空间位置对应于所述修改后的一个或多个其他音频对象在所述音频环境中相对于所述被选择的收听位置和定向的位置。

14. 根据权利要求13所述的装置,其中,所述装置被使得渲染所述空间音频信号,所述空间音频信号表示在包括一个或多个音频对象的可选视点音频环境中的声场,其中,每个音频对象与相应的音频内容和在所述音频环境中的相应位置相关联。

15. 根据权利要求14所述的装置,被使得进行修改的所述装置进一步包括被使得进行以下操作的构件:

响应于所述被检测的交互,识别要被应用于所述第一音频对象的第一修改,并对所述第一音频对象应用所述第一修改;

识别要被修改的一个或多个其他音频对象;

响应于所述被检测的交互,识别要被应用于相应的其他音频对象的一个或多个其他修改;以及

对所述相应的其他音频对象应用所述一个或多个其他修改,

其中,所述第一修改和所述一个或多个其他修改中的每一个包括以下中的一个:

激活所述第一音频对象和所述一个或多个其他音频对象中的相应的音频对象;

去激活所述第一音频对象和所述一个或多个其他音频对象中的相应的音频对象;

改变源自所述第一音频对象和所述一个或多个其他音频对象中的相应的音频对象的音频内容的相对幅度;以及

改变所述第一音频对象和所述一个或多个音频对象中的相应的音频对象在所述音频环境中的位置。

16. 根据权利要求15所述的装置,其中,所述第一音频对象与第一交互元数据相关联,

并且其中,所述一个或多个其他音频对象与相应的其他交互元数据相关联,其中,

所述第一交互元数据包括所述一个或多个预定义交互准则的定义、要被应用于所述第一音频对象的第一修改的定义、以及第一其他音频对象的标识;以及

第一其他交互元数据包括要被应用于所述第一其他音频对象的第一其他修改的定义。

17. 根据权利要求15所述的装置,其中,所述音频环境包括与控制元数据相关联的音频控制对象,其中,所述第一音频对象与第一交互元数据相关联,并且其中,所述一个或多个其他音频对象与相应的其他交互元数据相关联,其中,

所述控制元数据包括所述一个或多个预定义交互准则的定义、以及所述第一音频对象的标识;

所述第一交互元数据包括要被应用于所述第一音频对象的第一修改的定义、以及第一其他音频对象的标识;以及

第一其他交互元数据包括要被应用于所述第一其他音频对象的第一其他修改的定义。

18. 根据权利要求15所述的装置,其中,所述第一音频对象与第一交互元数据相关联,并且其中,所述第一交互元数据包括:所述一个或多个预定义交互准则的定义、要被应用于所述第一音频对象的第一修改的定义、所述一个或多个其他音频对象的相应的标识、以及要被应用于所述一个或多个其他音频对象的相应的一个或多个其他修改的相应定义。

19. 根据权利要求15所述的装置,其中,所述音频环境包括与控制元数据相关联的音频控制对象,并且其中,所述第一音频对象与第一交互元数据相关联,其中,所述控制元数据包括:所述一个或多个预定义交互准则的定义、要被应用于所述第一音频对象的第一修改的定义、所述一个或多个其他音频对象的相应的标识、以及要被应用于所述一个或多个其他音频对象的相应的一个或多个其他修改的相应定义。

音频处理

技术领域

[0001] 本发明的示例和非限制性实施例涉及渲染自由视点音频以呈现给用户。具体地，本发明的各种实施例涉及实现由自由视点音频环境内的用户与音频源之间的交互引起的渲染给用户的声场的改变。

背景技术

[0002] 自由视点音频通常允许用户在虚拟音频环境中四处移动和改变他/她的定向(即，旋转位置)，以及根据他/她在其中的位置和定向体验为虚拟音频环境定义的声场。虽然在本公开中主要使用术语“自由视点音频”来指代这种虚拟音频环境，但是相同的音频概念也可被称为自由收听点音频、六自由度(6DoF)音频或体积音频。在一些示例中，自由视点音频可被提供为仅音频环境，例如，作为独立的虚拟音频系统或作为增强现实(AR)或混合现实(MR)环境的一部分。在其他示例中，自由视点音频可被提供为诸如虚拟现实(VR)环境的视听环境的一部分。

[0003] 通常，虚拟音频环境的声场可以依赖于为虚拟音频环境定义的多个音频源或音频对象。一般来说，给定的音频源/对象由相应的音频内容(例如作为一个或多个数字音频信号而提供)定义，并由分配给该给定的音频源/对象的元数据补充，其中元数据可以定义音频内容和/或给定的音频源/对象的各种特性，包括其位置。虚拟音频环境内的音频源可以被表示为例如相应的基于通道的床和音频对象，被表示为一阶或更高阶的全景环绕声(Ambisonics)(FOA/HOA)和音频对象，仅被表示为相应的音频对象，或者通过使用任何等效的空间音频表示来表示。在一些情况下，例如可以结合音频对象来使用参数化的沉浸式音频表示。参数化的沉浸式音频表示可以部分地由参数组成，这些参数例如针对至少一个方向上的一组时频图块描述了定向与非定向(或扩散)音频之间的能量比、扩展相干性、环绕相干性、或相对于虚拟音频环境中的参考位置和旋转的距离。

[0004] 虚拟音频环境可以包括在虚拟音频环境的各个位置处的大量音频源，将这些音频源渲染给用户可以例如取决于用户相对于音频源的位置和定向。通常，鉴于用户的当前定向，在用户他/她当前位置可用的声场涉及一种空间声音，该空间声音包括一个或多个定向声源，可能还包括环境声部分，其可以经由耳机或通过使用多通道音频再现系统来再现给用户，例如作为双耳(立体声)音频信号。在虚拟音频环境中移动的用户可涉及用户的位置相对于一个或多个声源的改变和/或用户的定向相对于一个或多个声源的改变。因此，例如，当在虚拟音频环境中移动时，用户可以移动得更接近一个或多个音频源，用户可以与一个或多个音频源接触，用户可以移动得远离一个或多个音频源，用户可以转向背对或朝向一个或多个音频源，和/或由于用户的位置和/或定向的改变而新的音频源出现或消失，所有这些都导致渲染给用户的声场的特性发生变化。

[0005] 使用户接近或接触虚拟音频环境中的音频源的用户他/她的移动用作虚拟音频环境中用户与音频源的交互的示例，而其他类型的交互同样是可能的(例如，用户触摸、伸手去够、抓住、移动音频源本身或与音频源相关联的对象等，特别是在VR场景中)。虚拟音频环

境中的一个或多个音频源可被设置为对用户与其的交互做出反应。作为这方面的一些示例,虚拟音频环境中(例如,AR,MR或VR场景中)的音频源可对用户接近或离开音频源周围,用户转向朝向或背对音频源的位置,和/或用户以其他方式与音频源进行交互做出反应。音频源对用户与其的交互的响应的直接示例包括启动或终止播放与音频源相关联的音频内容,或修改已播放的音频内容的特性(例如,幅度)。例如可以在分配给音频源的元数据中定义这种反应。

[0006] 设置至少一些对用户交互做出反应的虚拟音频环境中的音频源使得能够定义更加全面的虚拟音频环境,例如,更像真实世界音频环境的虚拟音频环境,在许多情况下,这是虚拟音频环境的理想特性。然而,经由其元数据元素来设计虚拟音频环境中的大量音频源的这种交互,从而提供合理的真实世界行为模型在很多情况下是不可行的,甚至是不可能的,因为这样做需要时间和精力。因此,期望能够以更灵活和通行的方式定义和实现由用户与虚拟音频环境中的音频源的交互而引起的反应的机制,例如以便使得能够更有效地针对例如AR、MR或VR系统而定义和实现更逼真的虚拟音频环境。

发明内容

[0007] 根据一个示例性实施例,提供了一种用于渲染空间音频信号的方法,该空间音频信号表示在包括一个或多个音频对象的可选视点音频环境中的声场,其中,每个音频对象与相应的音频内容和在该音频环境中的相应位置相关联,该方法包括:接收该音频环境中被选择的收听位置和定向的指示;基于一个或多个预定义交互准则,检测关于第一音频对象的交互;响应于所述被检测的交互,修改第一音频对象和与其相连接的一个或多个其他音频对象;以及导出至少包括以下项的空间音频信号:与在该声场的第一空间位置的修改后的第一音频对象相关联的音频内容,该第一空间位置对应于修改后的第一音频对象在该音频环境中相对于所述被选择的收听位置和定向的位置;以及与在该声场的相应的其他空间位置的修改后的一个或多个其他音频对象相关联的音频内容,该相应的其他空间位置对应于修改后的一个或多个其他音频对象在该音频环境中相对于所述被选择的收听位置和定向的位置。

[0008] 根据另一示例性实施例,提供了一种用于渲染空间音频信号的装置,其中,该空间音频信号表示在包括一个或多个音频对象的可选视点音频环境中的声场,其中,每个音频对象与相应的音频内容和在该音频环境中的相应位置相关联,该装置被配置为:接收该音频环境中被选择的收听位置和定向的指示;基于一个或多个预定义交互准则,检测关于第一音频对象的交互;响应于所述被检测的交互,修改第一音频对象和与其相连接的一个或多个其他音频对象;以及导出至少包括以下项的空间音频信号:与在该声场的第一空间位置的修改后的第一音频对象相关联的音频内容,该第一空间位置对应于修改后的第一音频对象在该音频环境中相对于所述被选择的收听位置和定向的位置;以及与在该声场的相应的其他空间位置的修改后的一个或多个其他音频对象相关联的音频内容,该相应的其他空间位置对应于修改后的一个或多个其他音频对象在该音频环境中相对于所述被选择的收听位置和定向的位置。

[0009] 根据另一示例性实施例,提供了一种用于渲染空间音频信号的装置,其中,该空间音频信号表示在包括一个或多个音频对象的可选视点音频环境中的声场,其中,每个音频

对象与相应的音频内容和在该音频环境中的相应位置相关联,该装置包括:用于接收该音频环境中被选择的收听位置和定向的指示的构件;用于基于一个或多个预定义交互准则,检测关于第一音频对象的交互的构件;用于响应于所述被检测的交互,修改第一音频对象和与其相连接的一个或多个其他音频对象的构件;以及用于导出至少包括以下项的空间音频信号的构件:与在该声场的第一空间位置的修改后的第一音频对象相关联的音频内容,该第一空间位置对应于修改后的第一音频对象在该音频环境中相对于所述被选择的收听位置和定向的位置;以及与在该声场的相应的其他空间位置的修改后的一个或多个其他音频对象相关联的音频内容,该相应的其他空间位置对应于修改后的一个或多个其他音频对象在该音频环境中相对于所述被选择的收听位置和定向的位置。

[0010] 根据另一示例性实施例,提供了一种用于渲染空间音频信号的装置,其中,该空间音频信号表示在包括一个或多个音频对象的可选视点音频环境中的声场,其中,每个音频对象与相应的音频内容和在该音频环境中的相应位置相关联,其中,该装置包括至少一个处理器;以及包括计算机程序代码的至少一个存储器,该计算机程序代码在被该至少一个处理器执行时使该装置:接收该音频环境中被选择的收听位置和定向的指示;基于一个或多个预定义交互准则,检测关于第一音频对象的交互;响应于所述被检测的交互,修改第一音频对象和与其相连接的一个或多个其他音频对象;以及导出至少包括以下项的空间音频信号:与在该声场的第一空间位置的修改后的第一音频对象相关联的音频内容,该第一空间位置对应于修改后的第一音频对象在该音频环境中相对于所述被选择的收听位置和定向的位置;以及与在该声场的相应的其他空间位置的修改后的一个或多个其他音频对象相关联的音频内容,该相应的其他空间位置对应于修改后的一个或多个其他音频对象在该音频环境中相对于所述被选择的收听位置和定向的位置。

[0011] 根据另一示例性实施例,提供了一种计算机程序,该计算机程序包括计算机可读程序代码,该计算机可读程序代码被配置为当所述程序代码在计算设备上执行时,使得至少执行根据前述示例性实施例所述的方法。

[0012] 根据示例性实施例的计算机程序可以被体现在易失性或非易失性计算机可读记录介质上,例如作为计算机程序产品,该计算机程序产品包括在其上存储程序代码的至少一个计算机可读非暂时性介质,该程序在被装置执行时,使得该装置至少执行在上面针对根据本发明的示例性实施例的计算机程序而描述的操作。

[0013] 在本专利申请中提出的本发明的示例性实施例不应被解释为对所附权利要求的适用性构成限制。动词“包括”及其派生词在本专利申请中用作开放式限制,其不排除也存在未叙述的特征。除非另有明确说明,否则下文所描述的特征可以相互任意组合。

[0014] 本发明的一些特征在所附权利要求中列出。然而,当结合附图阅读以下一些示例性实施例的描述时,将最佳地理解本发明的各方面,同时包括其构造和操作方法,以及其附加的目的和优点。

附图说明

[0015] 在附图的各个图中,通过示例而非限制的方式示出了本发明的实施例,其中:

[0016] 图1A示出根据示例的用于对用户虚拟音频环境中被选择的位置渲染声场的装置的一些逻辑实体的框图;

[0017] 图1B示出根据示例的用于对用户虚拟音频环境中被选择的位置渲染声场的装置的一些逻辑实体的框图；

[0018] 图2示出根据示例的描述用于形成或修改表示虚拟音频环境中的声场的空间音频信号的方法的流程图；

[0019] 图3示出根据示例的描述用于形成或修改表示虚拟音频环境中的声场的空间音频信号的方法步骤的流程图；

[0020] 图4A示出根据示例的描述用于形成或修改表示虚拟音频环境中的声场的空间音频信号的方法步骤的流程图；

[0021] 图4B示出根据示例的描述用于形成或修改表示虚拟音频环境中的声场的空间音频信号的方法步骤的流程图；

[0022] 图5A示出根据示例的描述在单个设备中的音频信号渲染装置的一些逻辑实体的布置的框图；

[0023] 图5B示出根据示例的描述在两个设备中的音频信号渲染装置的一些逻辑实体的布置的框图；

[0024] 图5C示出根据示例的描述在三个设备中的音频信号渲染装置的一些逻辑实体的布置的框图；

[0025] 图6示出根据示例的装置的一些元件的框图。

具体实施方式

[0026] 在整个本公开中，术语“(虚拟) 音频环境”用于指代覆盖多个位置或地点并且具有为其定义的多个音频对象的虚拟环境。这样的虚拟环境可以跨越例如在其每个维度上具有相应的预定义大小的二维或三维空间。虚拟音频环境中包括的音频对象分别在该环境中具有其相应的位置。音频对象在虚拟音频环境中的位置可以是固定的，或者它可以随时间改变或被改变。音频对象还可以具有相对于音频环境中的一个或多个参考点(或参考方向)的定向。与位置一样，音频对象的定向也可以是固定的，或者它可以随时间改变或被改变。定向可用于定义由音频对象发出的声音的方向。如果没有为音频对象定义定向，则相应的音频对象可以应用预定义的默认定向，或者它可被视为全向音频对象。

[0027] 在本公开中所使用的术语“音频对象”并不指代某一音频标准或音频格式的元素，而是用作指代虚拟音频环境中的音频实体的通用术语。音频对象可以例如可替代地被称为音频源或音频条目。音频对象与音频内容和在虚拟音频环境中的位置相关联。在示例中，可以经由包括音频内容和至少定义音频对象的一个或多个空间和操作特性的一个或多个属性或参数的数据结构来定义该音频对象。作为这方面的示例，可以用定义音频对象在虚拟音频环境中的(当前)位置的一个或多个属性或参数来提供音频对象。作为另一示例，可以用定义音频内容的格式的一个或多个属性(例如，长度/时长、采样速率、音频通道数量、为其应用的音频编码格式等)来提供音频对象。音频对象的音频内容可被提供为例如数字音频信号，而音频对象的一个或多个属性可被提供为与音频对象和/或音频内容相关联的元数据。可以使用任何适用的预定义格式来提供元数据。随后在本公开中通过多个示例描述了关于虚拟音频环境的音频对象的其他方面。

[0028] 在示例中，与音频对象相关联的元数据可以包括内容部分和格式部分。其中，内容

部分可用于描述音频中包含的内容,并且它例如可包括与音频对象相关联的音频内容。格式部分可用于描述允许与音频对象相关联的音频内容的期望(正确)渲染的音频对象的技术特性。

[0029] 在本文中,术语“与.....相关联”用于描述音频对象与音频内容之间的关系以及音频对象与元数据(以及其中包括的任何参数或属性)之间的关系。然而,此关系也可被描述为为音频对象定义的音频内容和/或元数据,或者被描述为被分配给音频对象的音频内容和/或元数据。

[0030] 通常,这样的虚拟音频环境用作允许用户在虚拟音频环境中四处移动和/或改变他/她的定向以及根据他/她在虚拟音频环境中的位置和定向来体验其中可用的声场的可选视点音频环境或自由视点音频环境的示例。因此,虽然以下描述主要使用术语“虚拟音频环境”,但是该术语将以非限制性方式来解释,其包含各种类型的可选视点音频环境,其可被提供为或被称为例如自由收听点音频、六自由度(6DoF)音频或体积音频。通常,音频环境被提供为增强现实(AR)、混合现实(MR)或虚拟现实(VR)系统的一部分,而独立的音频环境也是可能的。

[0031] 根据非限制性示例,虚拟音频环境可被提供为AR或MR系统的一部分,或者虚拟音频环境可用作AR或MR系统。图1A示出了在AR或MR系统的上下文中用于对用户虚拟音频环境102中被选择的位置渲染声场的装置的一些逻辑实体的框图。在该使用场景中,音频渲染引擎104被设置为提供预定义的真实世界空间的位置或地点与虚拟音频环境102中的对应位置或地点之间的预定义映射。用户在真实世界空间中的位置和/或定向使用本领域中已知的适用的用户跟踪构件108来跟踪,该用户跟踪构件108操作以提取用户的在真实世界中的位置和/或定向的一个或多个指示,其被提供为音频渲染引擎104的输入数据。音频渲染引擎104操作以根据预定义映射,基于输入数据,导出用户在音频环境102中的位置和/或定向。音频渲染引擎104基于依赖于所导出的用户在音频环境102中的位置和/或定向的音频环境102的一个或多个音频对象,创建用于再现给用户的一个或多个空间音频信号,以使得所创建的一个或多个空间音频信号表示依赖于所导出的用户在音频环境102中的位置和/或定向的声场。音频渲染引擎104进一步通过音频再现构件106提供所创建的一个或多个空间音频信号以再现给用户。可以提供这种用于向用户再现音频环境以提供AR或MR系统的装置例如用于建筑物或其一部分(例如,百货商店、购物中心、旅馆、办公楼、博物馆、医院等)或用于预定义的室外区域(例如,公园、游乐场等)。

[0032] 在另一非限制性示例中,虚拟音频环境可被提供为还涉及视觉组件的VR系统的一部分。图1B示出了在VR系统的上下文中用于对用户虚拟音频环境102中被选择的位置渲染声场的装置的一些逻辑实体的框图。在这样的使用场景中,虚拟音频环境中的位置通常不与任何物理(真实世界)位置相对应,但是音频渲染引擎104被设置为提供虚拟世界的位置或地点与虚拟音频环境102中的相应的位置或地点之间的预定义映射。用户在虚拟世界中的位置和/或定向是经由通过用户输入构件208而接收到的用户命令或控制来跟踪或定义的,而音频渲染引擎104操作以根据预定义映射,基于用户命令和/或控制,导出用户在虚拟音频环境102中的位置和/或定向。如在前面针对图1A的装置所描述的,音频渲染引擎104基于依赖于所导出的用户在音频环境102中的位置和/或定向的音频环境102的一个或多个音频对象,创建用于再现给用户的一个或多个空间音频信号,以使得所创建的一个或多个空

间音频信号表示依赖于所导出的用户在音频环境102中的位置和/或定向的声场。音频渲染引擎104进一步通过音频再现构件106提供所创建的一个或多个空间音频信号以再现给用户。可以提供这种VR系统例如用于虚拟世界用于所捕获或计算机生成的内容。作为这方面的示例,可以经由VR系统向用户提供在虚拟博物馆或展览馆中的游览。

[0033] 通常,音频渲染引擎104操作以形成表示反映用户在虚拟音频环境102中的当前位置和/或定向的声场的空间音频信号,该空间音频信号被提供给音频再现构件106以播放给用户。空间音频信号例如可以包括双通道双耳(立体声)音频信号(用于耳机收听)或根据合适的多通道布局的多通道信号(用于经由扬声器系统收听)。由空间音频信号表示的声场可涉及在声场的相应的空间位置处的零个或多个定向声源,使得根据用户在虚拟音频环境102中的位置和/或定向,这些声源对应于零个或多个当前活动的音频对象的相应位置。零个或多个定向声源中的每一个可以以相应的相对幅度(例如,响度、信号水平)在声场中渲染,该相对幅度可以至少部分地被设置或调整以反映虚拟音频环境102中的当前用户位置与相应的音频对象的位置之间的距离,例如以使得被应用于某个声源的幅度的衰减随着用户与该声源之间距离的增加而增加。用于在声场的期望空间位置设置声源以及用于形成包括在该声场的相应的空间位置的多个声源的组合空间音频信号的各种技术在本领域中是已知的,在本文中可以使用此类合适的技术。例如,在双耳呈现的情况下,可以使用头部相关传递函数(HRTF)滤波或另一对应的技术。

[0034] 要被渲染给用户的声场的空间特性例如可以由于以下原因而变化:用户的位置和/或定向的变化(即,由于用户的移动),一个或多个当前活动的音频对象的移动,虚拟音频场景其整体的移动(相对于用户的位置),一个或多个其他音频对象的激活,和/或一个或多个当前活动的音频源的去激活。因此,音频渲染引擎104可操作以规则地更新被渲染给用户的声场的空间特性(例如,其中的声源的空间位置),以根据虚拟音频环境102中当前活动音频对象的当前位置,反映用户在虚拟音频环境102中的当前位置和/或定向。

[0035] 虚拟音频环境102的音频对象中的至少一些是交互式对象,其被设置为对用户直接与相应的音频对象进行交互做出响应和/或对用户经由一个或多个中间音频对象与相应的音频对象进行交互做出响应。在本文中,音频对象的响应涉及在音频对象方面应用的修改。修改例如可以在被分配给音频对象或与其相关联的元数据中定义。修改例如可涉及音频对象在虚拟音频环境102中的位置和/或音频内容的特性,如将通过随后在本公开中提供的示例更详细地描述的。音频对象的响应可以由音频渲染引擎102通过创建或修改表示依赖于所导出的用户在虚拟音频环境102中的位置和/或定向以及基于被应用于该音频对象的修改的声场的一个或多个空间音频信号来实现。

[0036] 在此方面,虚拟音频环境102可以包括被设置用于以下一种或多种类型的交互和/或响应的音频对象:

[0037] -单独交互(Individual interaction):音频对象可被设置为对用户与音频对象的交互进行响应,以使得该响应与用户与虚拟音频环境102的另一音频对象交互(同时或过去)无关。

[0038] -关联交互(Connected interaction):第一音频对象可被设置为对用户与音频对象的交互进行响应,一个或多个第二对象可被设置为对在第一音频对象中被唤起(invoked)的响应进行响应。因此,在第一音频对象中被唤起的响应可被视为第一音频对象

与被设置为相应地进行反应的一个或多个第二音频对象的交互。此外,相应的一个或多个第三音频对象可被设置为对在多个第二音频对象中的相应的音频对象中被唤起的响应进行响应,以此类推。从概念上讲,可以考虑两种不同类型的关联交互:

[0039] ○第一级关联交互,可以在第二音频对象中被唤起,作为响应于用户与第一音频对象的交互而在第一音频对象中引起的响应的结果;

[0040] ○第二级关联交互,可以在第三或后续音频对象中被唤起,作为由于由用户与第一音频对象的交互引起的一系列两个或更多个关联交互(第一级关联交互以及随后的一个或多个第二级关联交互)而在另一音频对象中引起的响应的结果。

[0041] 因此,第一级关联交互可以在与其相关联的一个或多个其他音频对象中唤起相应的一系列零个或多个第二级关联交互。如果音频对象被设置用于单独交互和关联交互两者,则由这两个交互引起的相应的响应可彼此相似或可彼此不同。

[0042] -组交互(Group interaction):第一音频对象和一个或多个第二对象可被设置为共同对用户与第一音频对象的交互进行响应。如果音频对象被设置为用于单独交互和组交互两者,则由这两个交互引起的相应的反应优选地彼此不同。此外,如果音频对象被设置为用于关联交互和组交互两者,则由这两个交互引起的相应的反应也优选地彼此不同。

[0043] 可以经由元数据来控制关联交互和组交互中的每一个。可以在被分配给虚拟音频环境102的音频对象或与其相关联的元数据中和/或被分配给虚拟音频环境102中包括的专用音频控制对象或与其相关联的元数据中提供定义相应的交互的信息。音频控制对象例如可被提供为虚拟音频环境102的特定对象类型或与空音频内容相关联的音频对象。

[0044] 用户与音频对象之间的交互的检测可涉及关于虚拟音频环境102中用户与音频对象之间的空间关系的考虑。这可涉及音频渲染引擎104根据相对于音频对象在虚拟音频环境102中的位置的用户在虚拟音频环境102中的位置和/或定向,确定用户是否在该音频对象的附近。音频渲染引擎104可以响应于在虚拟音频环境102中用户在音频对象的附近而考虑用户与该音频对象的交互。在以下方面提供了这方面的非限制性示例:

[0045] -响应于用户的位置与音频对象的位置之间的距离低于预定义阈值距离,该用户可被视为在该音频对象的附近。

[0046] -响应于用户的位置与音频对象的位置之间的距离低于预定义阈值距离以及该音频对象朝向该用户的位置,该用户可被视为在该音频对象的附近。

[0047] -响应于用户的位置与音频对象的位置之间的距离低于预定义阈值距离以及该用户朝向该音频对象的位置,该用户可被视为在该音频对象的附近。

[0048] -响应于用户的位置与音频对象的位置之间的距离低于预定义阈值距离、该用户朝向该音频对象的位置、以及该音频对象朝向该用户的位置,该用户可被视为在该音频对象的附近。

[0049] 从上面的示例可以看出,可以分开考虑可针对音频对象而定义的定向和用户的相对于声源的位置的定向。

[0050] 用户的位置与音频对象的位置之间也可存在另一空间关系:虚拟音频环境102可以根据用户的位置与音频对象的位置之间的距离来定义用于源自音频对象的声音的默认衰减,以使得该衰减随着距离的增加而增加。可以针对多个频率子带而单独地(并且可不同地)定义默认衰减,或者默认衰减在整个频带上可以相同。在示例中,默认衰减定义被应用

于虚拟音频环境102的所有音频对象的默认音频呈现级别。在其他示例中,可以针对音频对象子集而共同定义默认衰减,或者可以针对虚拟音频环境102的一个或多个(甚至所有)音频对象而单独定义默认衰减。通常,适用于给定音频对象的默认衰减并不考虑用户相对于给定音频对象的接近度以及可由此产生的任何响应。

[0051] 在一些示例中,用于检测用户与音频对象之间的交互的空间关系特性因此可基于音频对象针对用户的默认音频渲染音量水平。当音频对象针对在当前用户位置的用户的默认音频渲染高于针对该音频对象而定义的阈值(例如,在与该音频对象相关联的交互元数据中)时,交互被检测到,并且根据因此被观测到的交互程度,将与音频对象相关联的音频内容渲染给用户。根据该替代的实现,所述对象的默认音频渲染从该点开始,仅用于检测交互是否被维持,直到该交互结束。仅当音频对象的默认音频渲染音量水平例如由于至少音频对象或用户的位置变化或由于任何其他原因而低于当前阈值时,默认音频渲染才再次被用于提供针对用户的整个音频呈现的音频对象贡献。

[0052] 除了空间关系之外,用户与音频对象之间的交互的确定还可涉及关于时间方面的考虑。作为这方面的示例,音频渲染引擎104可以响应于至少在预定时间段内在虚拟音频环境102中用户在音频对象的附近而考虑用户与该音频对象的交互。

[0053] 作为对时间方面的考虑的替代或补充,对用户与音频对象之间的交互的检测可涉及关于寻址该音频对象或与该音频对象相关联的对象的用户动作的考虑。作为这方面的示例,音频渲染引擎104可以响应于接收到指示在虚拟音频环境102中用户在音频对象的附近时寻址该音频对象或与该音频对象相关联的对象的用户动作(例如,一个或多个预定义用户动作中的一个)的一个或多个控制输入而考虑用户与该音频对象的交互。引起提供这种控制输入的事件不在本公开的范围之内。然而,在AR或MR系统中,这样的控制输入可以响应于用户触摸或伸手去够与音频对象相关联的真实世界对象(例如,可以由用户跟踪构件108来检测和指示)而被触发,而在VR系统中,这样的控制输入(例如,通过用户输入构件208)可以响应于用户触摸或伸手去够与音频对象相关联的虚拟世界对象而被触发。

[0054] 在下文中,描述了响应于用户与音频对象的交互而在该音频对象中被唤起的响应的一些非限制说明性示例。即使以下示例以单数形式提及响应,一般而言,用户与音频对象的交互可以在交互下的音频对象中唤起响应的组合或一系列响应和/或两个或多个独立的响应。

[0055] 通常,响应包括与用户相关的音频对象的状态的某种变化或修改。作为示例,响应于用户与音频对象的交互而在该音频对象中被唤起的响应可以包括音频对象的激活。这可涉及基于与表示声场的空间音频信号中的音频对象相关联的音频内容而引入定向声音分量,以使得该声音分量的空间位置对应于相对于用户在虚拟音频环境102中的当前位置和/或定向的音频对象在虚拟音频环境102中的当前位置。作为另一示例,响应于用户与音频对象的交互而在该音频对象中被唤起的响应可以包括音频对象的去激活,这可涉及基于与来自声场中的音频对象相关联的音频内容,移除被渲染的定向声音分量。

[0056] 作为另一示例,响应于用户与音频对象的交互而在该音频对象中被唤起的响应可以包括对应于该音频对象的声场的定向声音分量的幅度变化(例如,信号水平或响度的变化)。该变化可涉及减小幅度(例如,增大衰减或减小增益)或增大幅度(例如,减小衰减或增大增益)。在此方面,可以关于为音频对象定义的默认衰减(在上文中描述)的操作所产生的

渲染幅度来定义和引入示例性幅度变化。作为这方面的示例,可由用户移动接近音频对象(并由此与该音频对象进行交互)所产生的幅度变化可导致与该音频对象相关联的音频内容以明显比由默认衰减定义的高的信号水平被渲染给用户。在另一示例中,用户移动接近音频对象(并由此与其进行交互)可导致与该音频对象相关联的音频内容以恒定的信号水平或更低的信号水平被渲染给用户,尽管用户与该音频对象之间的距离变得更小。换句话说,如果没有任何唤起来自音频对象的响应的用户交互,则由该音频对象表示的定向声音分量的渲染级别(例如,其音量)将反映自由视点音频渲染的一般特性,例如,其中移动接近音频源通常将导致被感知的响度(而不是该音频源它本身的信号水平)的增加。用户与音频对象之间的交互因此可以改变该动态。

[0057] 作为又一示例,响应于用户与音频对象的交互而在该音频对象中被唤起的响应可以包括该音频对象的位置的变化,该位置变化导致对应于该音频对象的声场的定向声音分量的空间位置的变化。该位置变化可涉及从当前位置到被定义的目标位置的一次性变化。作为这方面的一些示例,音频对象的位置变化可被定义为直接从当前位置变化到目标位置,或者它可被定义为在特定时间段内经由一个或多个中间位置变化到目标位置。在另一示例中,该位置变化可涉及两个或多个位置之间的重复或连续变化,例如,随机地或以预定义的时间间隔地。

[0058] 由于前述的用户与音频对象的交互而在该音频对象中被唤起的反应的说明性示例也用作响应于在虚拟音频环境102的另一音频对象中被唤起的响应而在音频对象中被唤起的反应的适用示例。

[0059] 如上文所描述的,音频渲染引擎104操作以在虚拟音频环境102内用户的当前位置将声场渲染给用户作为空间音频信号,其中,该空间音频信号的空间特性被规则地(例如,以预定义的间隔)更新以反映相对于虚拟音频环境102中当前活动音频对象的当前位置的用户在虚拟音频环境102中的当前位置和/或定向。在下文中,我们也将用户在虚拟音频环境102中的位置和定向称为虚拟音频环境102中被选择的位置和定向。

[0060] 在形成或修改空间音频信号时,音频渲染引擎104例如可以根据由图2中的流程图所示的方法300进行操作。方法300例如可以由音频渲染引擎104来实现。方法300开始于接收虚拟音频环境102中被选择的收听位置和定向的指示,如框302所示。例如,如上文所描述的,可以从用户跟踪构件108或从控制构件208接收这样的指示。

[0061] 方法300还涉及检测关于虚拟音频环境102的第一音频对象的交互,如框304所示。在示例中,关于第一音频对象的交互包括基于一个或多个预定义交互准则,第一音频对象与虚拟音频环境102中被选择的收听位置和定向之间的交互。在另一示例中,关于第一音频对象的交互涉及第一音频对象与虚拟音频环境102的一个或多个其他音频对象之间的交互。

[0062] 方法300还包括响应于检测到关于第一音频对象的交互,修改第一音频对象和与其相连接一个或多个其他音频对象,如框306所示。第一音频对象与一个或多个其他音频对象之间的连接(link)例如可以经由与第一音频对象相关联的交互元数据来定义。

[0063] 最后,如框308所示,方法300进行到导出包括以下项中的至少一个的空间音频信号:

[0064] -与在声场的第一空间位置的修改后的第一音频对象相关联的音频内容,该第一

空间位置对应于修改后的第一音频对象在虚拟音频环境102中相对于被选择的收听位置和定向的位置；

[0065] -与在声场的第二空间位置的修改后的第二音频对象相关联的音频内容,该第二空间位置对应于修改后的第二音频对象在虚拟音频环境102中相对于被选择的收听位置和定向的位置。

[0066] 在示例中,空间音频信号的导出涉及创建空间音频信号,其包括与在声场的相应的空间位置的修改后的第一和/或第二音频对象相关联的相应的音频内容,例如以使得源自修改后的第一和/或第二音频对象的音频内容是该声场的唯一定向声源。在另一示例中,空间音频信号的导出涉及修改空间音频信号,以使得它包括与在声场的相应的空间位置的修改后的第一和/或第二音频对象相关联的相应的音频内容,例如以使得所得到的修改后的空间音频信号除了源自修改后的第一和/或第二音频对象的音频内容之外还包括一个或多个其他定向声音源。

[0067] 作为提供在上文中参考框306所描述的操作的示例,方法300可以包括图3的流程图中所示的方法步骤300'。根据方法步骤300',方法300可涉及响应于所检测的关于第一音频对象的交互,识别要被应用于第一音频对象的第一修改,如框310所示,以及对第一音频对象应用第一修改,如框312所示。

[0068] 根据方法步骤300',方法300还涉及响应于所检测的关于第一音频对象的交互,识别要被修改的一个或多个其他音频对象,识别要被应用于一个或多个其他音频对象的相应的一个或多个其他修改,以及对一个或多个其他音频对象应用所识别的相应的一个或多个其他修改,如框314、316和318所示。第一修改与一个或多个其他修改之间的关系例如可以是在上文中所描述的关联交互和组交互中的一个。下面描述关于与框310至316有关的操作的说明性示例。

[0069] 在上文中所描述的各种示例中,交互的检测和作为检测与第一音频对象的交互的结果而在第一和/或第二音频对象中被应用的修改依赖于元数据。为了清楚的描述,在下文中我们将这样的元数据称为交互元数据。根据交互的类型,交互元数据可被提供为与第一音频对象相关联的元数据,与一个或多个其他音频对象中的一个相关联的元数据,与音频控制对象相关联的元数据,或者与第一音频对象、一个或多个其他音频对象和音频控制对象中的一个或多个相关联的元数据。交互元数据可以定义以下各方面:

[0070] -定义与第一音频对象的交互的一个或多个交互准则;

[0071] -响应于检测到交互而要被应用于第一音频对象的一个或多个响应或修改;

[0072] -一个或多个其他音频对象的标识;

[0073] -要被应用于所识别的一个或多个其他音频对象的相应的一个或多个响应或修改。

[0074] 可以根据其中虚拟音频环境102被提供的基础应用或框架的要求来选择用于实现交互元数据中的以上定义的确切格式或语法。

[0075] 针对不同类型的交互(关联交互、组交互),在与第一音频对象相关联的交互元数据、与一个或多个其他音频对象相关联的交互元数据、和/或与音频控制对象相关联的交互元数据中的一个或多个之间分布交互元数据的非限制性示例在下面提供。

[0076] 作为基于与第一音频对象和一个或多个其他音频对象相关联的相应的交互元数

据而提供在上文中参考框304和306所描述的操作的示例,方法300可以包括图4A的流程图所示的方法步骤400,该方法步骤400可以被认为是将方法300呈现为第一音频对象与一个或多个其他音频对象之间的关联交互的实现。根据方法步骤400,方法300可涉及基于在与第一音频对象相关联的第一交互元数据中定义的一个或多个预定义交互准则,检测关于第一音频对象的交互,如框404所示。在示例中,关于第一音频对象的交互包括第一音频对象与虚拟音频环境102中被选择的收听位置和定向之间的交互,而在另一示例中,关于第一音频对象的交互包括第一音频对象与虚拟音频环境102的一个或多个其他音频对象之间的交互。

[0077] 根据方法步骤400,方法300还可涉及从与第一音频对象相关联的第一元数据识别要被应用于第一音频对象的第一修改和第一其他音频对象,如框410所示。如框412所示,该方法还可涉及对第一音频对象应用第一修改。根据方法步骤400,方法300还可涉及从与第一其他音频对象相关联的第一其他交互元数据识别要被应用于第一其他音频对象的第一其他修改以及可选的第二其他音频对象,如框416-1所示。该方法还可涉及对第一其他音频对象应用第一其他修改,如框418-1所示。

[0078] 如果第一其他元数据包括第二其他音频对象的标识,则根据方法步骤400,方法300还可继续从与第二其他音频对象相关联的第二其他交互元数据识别要被应用于第二其他音频对象的第二其他修改以及可选地继续识别第三其他音频对象。如框416-2所示。该方法还继续对第一其他音频对象应用第一其他修改,如框418-2所示。

[0079] 作为示例,在方法步骤400的上下文中,第一音频对象可以与第一交互元数据相关联,该第一交互元数据包括规定第一音频对象与被选择的位置和定向之间的交互的一个或多个交互准则的定义、响应于所检测的交互而要被应用于第一音频对象的第一修改的定义、以及第一其他音频对象的标识,而第一其他音频对象可以与第一其他交互元数据相关联,该第一其他交互元数据包括响应于所检测的与第一音频对象的交互而要被应用于第一其他音频对象的第一其他修改的定义。此外,第一其他交互元数据可以选择性地包括第二其他音频对象的标识,并且第二其他音频对象可以与第二其他交互元数据相关联,该第二其他交互元数据包括要被应用于第二其他音频对象的第二其他修改的定义、以及另外可选的第三其他音频对象的标识。音频对象的这种连接性(linkage)可被应用以提供涉及第一音频对象和一个或多个其他音频对象直到任何期望数量的其他音频对象的关联交互。

[0080] 作为提供在上文中参考框304和306所描述的操作的示例,方法300可以包括在图4B的流程图中所示的方法步骤500,该方法步骤500可以被认为是将方法300呈现为涉及第一音频对象和一个或多个其他音频对象的组交互的实现。在此方面,该方法包括基于在与第一音频对象相关联的第一交互元数据中定义的一个或多个预定义交互准则,检测关于第一音频对象的交互,如框504所示。在示例中,关于第一音频对象的交互包括第一音频对象与虚拟音频环境102中被选择的收听位置和定向之间的交互,而在另一示例中,关于第一音频对象的交互涉及第一音频对象与虚拟音频环境102的一个或多个其他音频对象之间的交互。

[0081] 根据方法步骤500,方法300还可涉及从与第一音频对象相关联的第一元数据识别要被应用于第一音频对象的第一修改、一个或多个其他音频对象、以及相应的一个或多个其他修改,如框510所示,根据方法步骤500,方法300还可涉及对第一音频对象应用第一修

改,如框512所示,以及对一个或多个其他音频对象应用相应的一个或多个其他修改,如框518所示。

[0082] 作为示例,在方法步骤500的上下文中,第一音频对象可以与第一交互元数据相关联,该第一交互元数据包括定义第一音频对象与被选择的位置和定向之间的交互的一个或多个交互准则的定义、响应于所检测的交互而要被应用于第一音频对象的第一修改的定义、一个或多个其他音频对象的相应定义、以及响应于所检测的交互而要被应用于一个或多个其他对象的一个或多个其他修改的相应定义。

[0083] 在方法步骤500的上下文中的示例中,第一音频对象可被提供为与非空音频内容相关联的常规音频对象。在另一示例中,第一音频对象可被提供为与空音频内容相关联的音频控制对象,由此,该音频控制对象作用于第一音频对象和一个或多个其他音频对象的组交互的控制点。

[0084] 返回参考图1A和1B中所示的用于在虚拟音频环境102中被选择的位置渲染声场的装置,图5A至5C示出了描绘将图1A和1B中所示的元件布置到多个设备中的相应的非限制性示例的框图。在此方面,图5A示出了其中音频环境102、音频渲染引擎104和音频再现构件106被设置在单个设备620中的装置。设备620例如可以是膝上型计算机、台式计算机、电视机、游戏机、或其他类型的家庭娱乐设备等。在此,音频环境102可被提供为存储在设备620中提供的一个或多个存储器中的信息,音频渲染引擎104可被提供为执行存储在存储器中的计算机程序的处理器,音频再现构件106可被提供为在设备610中提供的扬声器装置。

[0085] 图5B示出了图5A的布置的变型,其中,音频环境102和音频渲染引擎104在第一设备621中提供,音频再现构件106在第二设备622中提供。其中,音频环境102和音频渲染引擎104可以按照上面针对设备620描述的来提供,第一设备621可同样地例如被提供为便携式计算机、台式计算机、电视机、游戏机、或其他类型的家庭娱乐设备等。托管音频再现构件106的第二设备622可以包括通过有线或无线通信链路可连接到第一设备621的扬声器装置或耳机。在另一场景中,第一设备621可被提供为诸如平板计算机、移动电话(例如,智能电话)、便携式媒体播放设备、便携式游戏设备等的移动设备,而第二设备622可被提供为通过无线或有线通信链路可连接到第一设备621的耳机或头戴式耳机。在又一场景中,第一设备621可被提供为服务器设备(例如,音频渲染服务器),而第二设备622可被提供为被提供有使能经由通信网络连接到第一设备621的适用的无线通信构件的耳机或头戴式耳机。

[0086] 图5C示出了图5B的布置的变型,其中,音频渲染引擎104在第一设备621中提供,音频再现构件106在第二设备622中提供,虚拟音频环境102在第三设备623中提供。在上文中针对图5B的布置而概述的示例性场景也适用于图5C的布置,除了第一设备621经由通信网络可连接到第三设备623以使得音频渲染引擎104能够访问其中的虚拟音频环境102以外。

[0087] 在一些示例性实施例中,提供了例如在诸如6DoF音频源交互的情形下的虚拟音频环境102的音频对象之间的关联交互,从而在结合源自与用户直接交互的音频对象的音频内容的音频播放期间,使得音频对象(诸如与期望或目标音频内容分离的音频对象)的无缝交互场景显得自然。一些示例性实施例在处理不同类型的音频对象或源(与交互元数据相关联的音频对象和与交互元数据不相关联的音频对象)时实现了自然交互。应当理解,本发明的实施例使得能够满足逻辑交互响应而无需满足交互元数据的期望,特别是对于关联交互、跨接等。

[0088] 图6示出了示例性装置700的一些组件的框图。装置700可以包括图6中未示出的其他组件、元件或部件。装置700例如可用于实现音频渲染引擎104。

[0089] 装置700包括处理器716和用于存储数据和计算机程序代码717的存储器715。存储器715和存储在其中的计算机程序代码717的一部分可进一步被设置为与处理器716一起实现在上文中在音频渲染引擎104的上下文中描述的功能。

[0090] 装置700包括用于与其他设备通信的通信部件712。通信部件712包括使能与其他装置的有线或无线通信的至少一个通信装置。通信部分712的通信装置也可被称为相应的通信构件。

[0091] 装置700还可以包括用户I/O(输入/输出)组件718,这些组件可被设置为可与处理器716和计算机程序代码717的一部分一起提供用户接口以用于接收来自装置700的用户的输入和/或向装置700的用户提供输出以控制由装置700实施的音频渲染引擎104的操作的至少一些方面。用户I/O组件718可以包括诸如显示器、触摸屏、触摸板、鼠标、键盘、和/或一个或多个键或按钮布置等的硬件组件。用户I/O组件718也可被称为外围设备。处理器716可被设置为例如根据计算机程序代码717的一部分,并且还可能根据经由用户I/O组件718接收的用户输入和/或根据经由通信部件712接收的信息来控制设备700的操作。

[0092] 虽然处理器716被描绘为单个组件,但是它也可被实现为一个或多个单独的处理组件。类似地,虽然存储器715被描绘为单个组件,但是它也可被实现为一个或多个单独的组件,这些组件中的一些或全部可被集成/移动和/或可以提供永久/半永久/动态/缓存存储。

[0093] 存储在存储器715中的计算机程序代码717可以包括计算机可执行指令,该计算机可执行指令在被加载到处理器716中时控制装置700的操作的一个或多个方面。作为示例,计算机可执行指令可被提供为一个或多个指令的一个或多个序列。处理器716能够通过从存储器715读取包括在其中的一个或多个指令的一个或多个序列来加载和执行计算机程序代码717。一个或多个指令的一个或多个序列可被配置为当被处理器716执行时,使得装置700执行在上文中在音频渲染引擎104的上下文中描述的操作、过程和/或功能。

[0094] 因此,装置700可以包括至少一个处理器716以及包括用于一个或多个程序的计算机程序代码717的至少一个存储器715,至少一个存储器715和计算机程序代码717被配置为与至少一个处理器716一起使得装置700执行在上文中在音频渲染引擎104的上下文中描述的操作、过程和/或功能。

[0095] 存储在存储器715中的计算机程序例如可被提供为相应的计算机程序产品,该计算机程序产品包括在其上存储计算机程序代码717的至少一个计算机可读非暂时性介质,该计算机程序代码在被装置700执行时,使得装置700至少执行在上文中在音频渲染引擎104(或其一个或多个组件)的上下文中描述的操作、过程和/或功能。计算机可读非暂时性介质可以包括存储设备或记录介质,诸如CD-ROM、DVD、蓝光光盘、或有形地体现计算机程序的其他制品。作为另一示例,计算机程序可被提供为被配置为可靠地传送计算机程序的信号。

[0096] 不应将对处理器的提及理解为仅涵盖可编程处理器,还可以涵盖专用电路,诸如现场可编程门阵列(FPGA)、专用电路(ASIC)、信号处理器等。在上述说明中描述的特征也可通过明确描述的组合以外的其他组合来使用。

[0097] 虽然已经参考某些特征描述了一些功能,但是这些功能也可通过其他特征来执行,无论是否描述。虽然已经参考某些实施例描述了一些特征,但是这些特征也可在其他实施例中提供,无论是否描述。

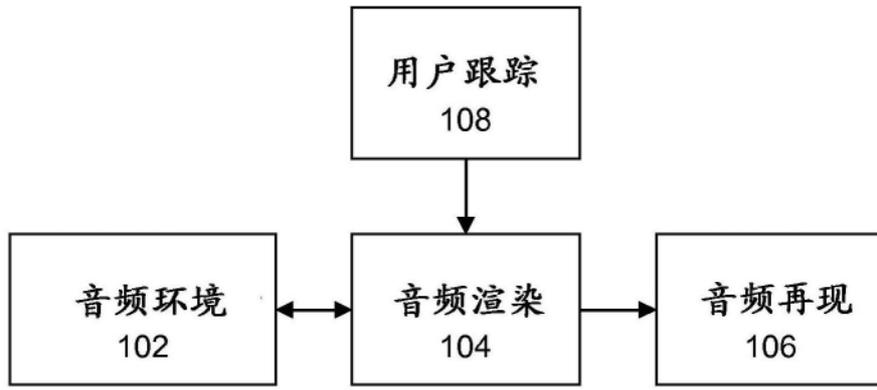


图1A

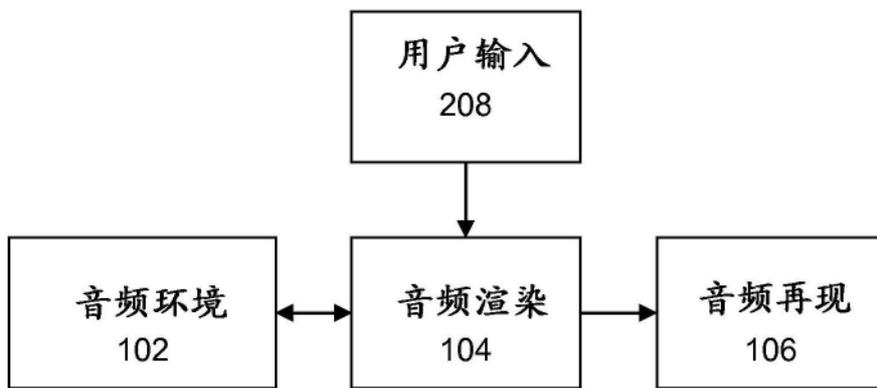


图1B

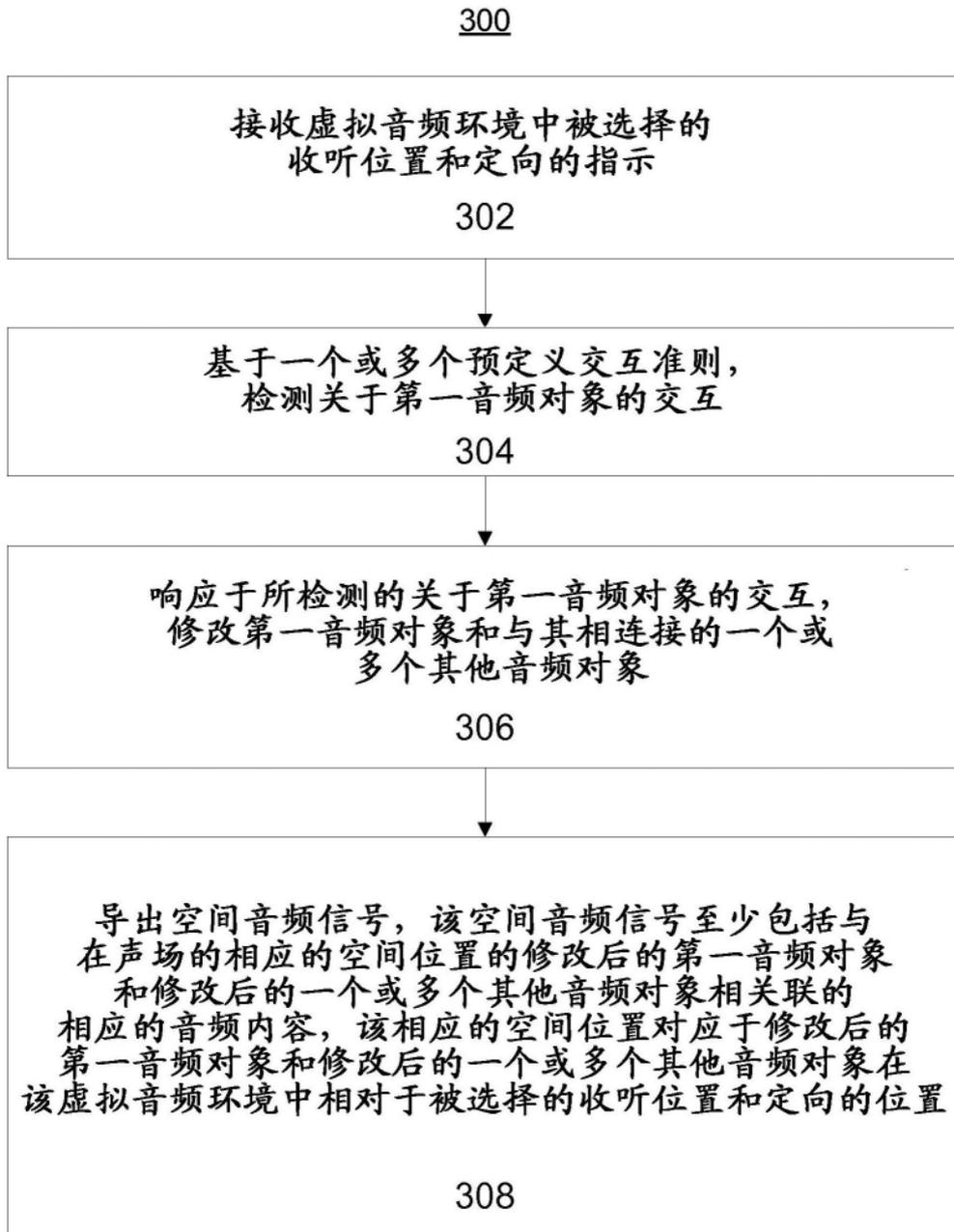


图2



图3

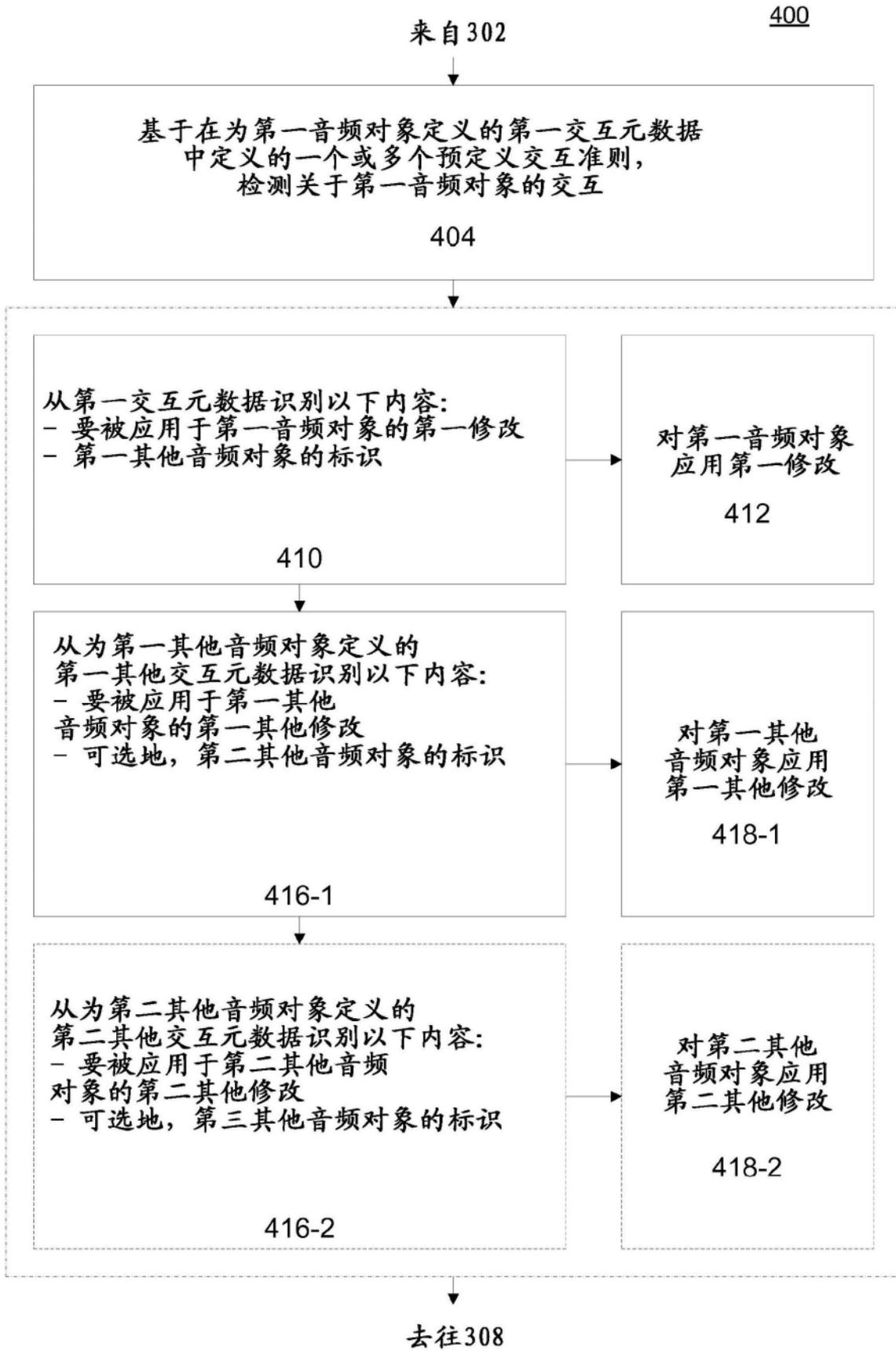


图4A

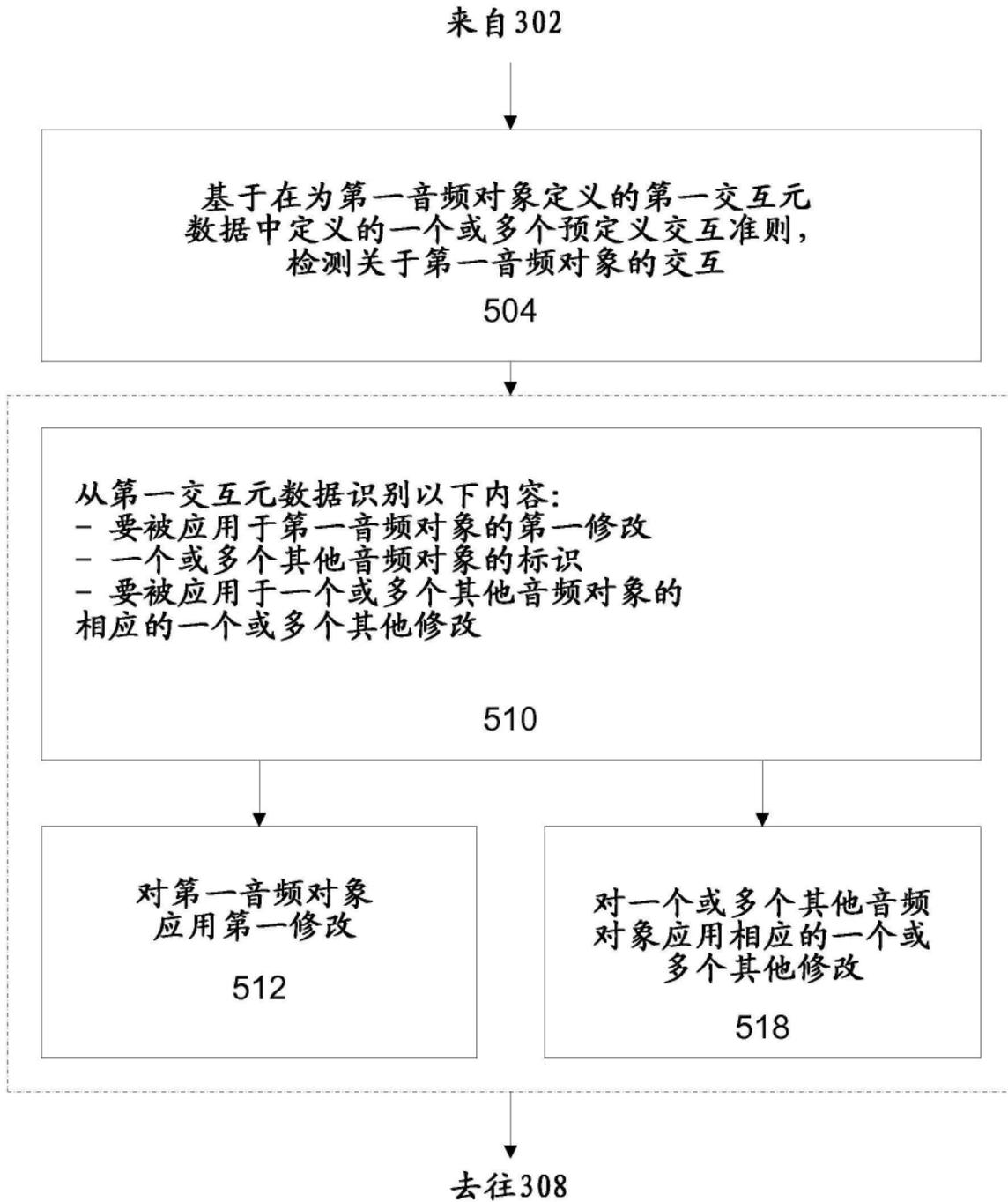


图4B

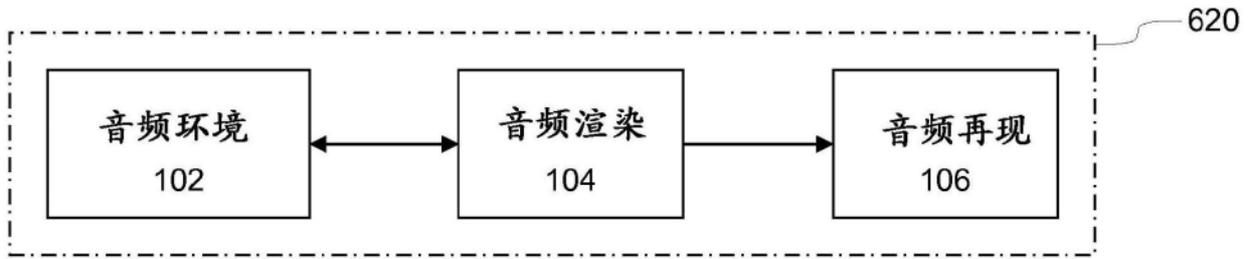


图5A



图5B



图5C

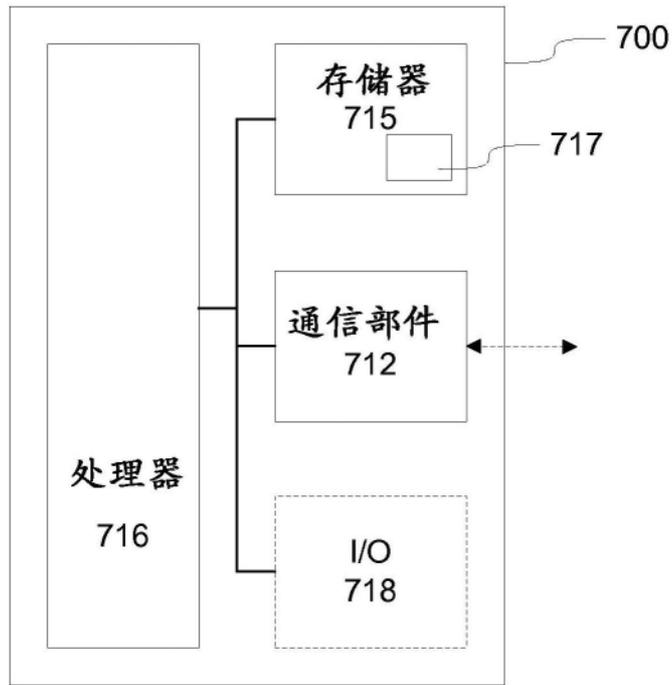


图6