



(12) 发明专利申请

(10) 申请公布号 CN 116830604 A

(43) 申请公布日 2023. 09. 29

(21) 申请号 202180092766.5

(22) 申请日 2021.12.02

(30) 优先权数据

63/121,108 2020.12.03 US

63/202,003 2021.05.21 US

(85) PCT国际申请进入国家阶段日

2023.08.02

(86) PCT国际申请的申请数据

PCT/US2021/061669 2021.12.02

(87) PCT国际申请的公布数据

W02022/120091 EN 2022.06.09

(71) 申请人 杜比实验室特许公司

地址 美国加利福尼亚州

(72) 发明人 J·B·兰多 A·J·西斐德

(74) 专利代理机构 北京市汉坤律师事务所

11602

专利代理师 魏小微 吴丽丽

(51) Int.Cl.

H04S 7/00 (2006.01)

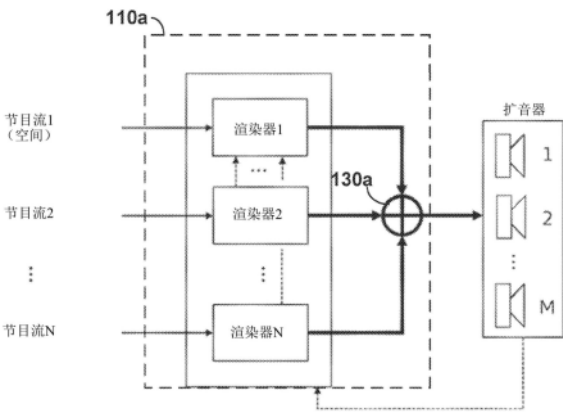
权利要求书3页 说明书46页 附图50页

(54) 发明名称

动态应用的渲染配置的渐进计算和应用

(57) 摘要

一些示例涉及通过以下操作来渲染接收到的音频数据:根据与第一组扬声器激活相对应的第一渲染配置确定环境中的一组扩音器的第一相对激活;接收指示从第一渲染配置到第二渲染配置的转变的第一渲染转变指示;以及确定与第二渲染配置的简化版本相对应的第二组扬声器激活。一些示例涉及:执行从第一组扬声器激活到第二组扬声器激活的第一转变;确定与第二渲染配置的完整版本相对应的第三组扬声器激活;以及执行到第三组扬声器激活的第二转变而无需完成第一转变。



1. 一种音频处理方法,包括:

由控制系统并且经由接口系统接收音频数据,所述音频数据包括一个或多个音频信号和相关联的空间数据,所述空间数据指示与音频信号相对应的预期感知空间位置;

由所述控制系统渲染所述音频数据以经由环境的一组扩音器再现,从而产生第一经渲染的音频信号,其中,渲染所述音频数据以供再现涉及根据第一渲染配置确定所述环境中的一组扩音器的第一相对激活,所述第一渲染配置对应于第一组扬声器激活;

经由所述接口系统向所述环境的所述一组扩音器中的至少一些扩音器提供所述第一经渲染的音频信号;

由所述控制系统并且经由所述接口系统接收第一渲染转变指示,所述第一渲染转变指示指示从所述第一渲染配置到第二渲染配置的转变;

由所述控制系统确定与所述第二渲染配置的简化版本相对应的第二组扬声器激活;

由所述控制系统执行从所述第一组扬声器激活到所述第二组扬声器激活的第一转变;

由所述控制系统确定与所述第二渲染配置的完整版本相对应的第三组扬声器激活;以及

由所述控制系统执行到所述第三组扬声器激活的第二转变,而无需完成所述第一转变。

2. 如权利要求1所述的方法,其中,所述第一组扬声器激活、所述第二组扬声器激活和所述第三组扬声器激活是频率相关的扬声器激活。

3. 如权利要求2所述的方法,其中,所述频率相关的扬声器激活涉及在至少第一频带中应用感知空间位置的模型,所述模型产生与收听者的左耳和右耳处的音频对象位置相对应的双耳响应。

4. 如权利要求3所述的方法,其中,所述频率相关的扬声器激活涉及在至少第二频带中应用感知空间位置模型,所述模型将从一组扩音器播放的音频信号的感知空间位置置于由所述扩音器的相关联的激活增益加权的所述一组扩音器的位置的质心处。

5. 如权利要求1至4中任一项所述的方法,其中,所述第一组扬声器激活、所述第二组扬声器激活或所述第三组扬声器激活中的至少一者是优化成本的结果,所述成本是以下各项的函数:

当在所述环境中的所述一组扩音器上回放时播放的音频信号的感知空间位置的模型;

所述音频信号的预期感知空间位置与所述一组扩音器中的每个扩音器的位置的接近度的度量;以及

一个或多个附加动态可配置功能,其中,所述一个或多个附加动态可配置功能基于以下各项中的一项或多项:扩音器与一个或多个收听者的接近度;扩音器与吸引力位置的接近度,其中,吸引力是有利于更接近所述吸引力位置的相对更高的扩音器激活的因素;扩音器与排斥力位置的接近度,其中,排斥力是有利于更接近所述排斥力位置的相对更低的扩音器激活的因素;每个扩音器相对于所述环境中的其他扩音器的能力;所述扩音器关于其他扩音器的同步;唤醒词性能;或回声消除器性能。

6. 如权利要求1至5中任一项所述的方法,进一步包括:

由所述控制系统并且经由所述接口系统接收第二渲染转变指示,所述第二渲染转变指示指示到第三渲染配置的转变;

由所述控制系统确定与所述第三渲染配置相对应的第四组扬声器激活;以及
由所述控制系统执行到所述第四组扬声器激活的第三转变,而无需完成所述第一转变或所述第二转变。

7.如权利要求6所述的方法,进一步包括:

由所述控制系统并且经由所述接口系统接收第三渲染转变指示,所述第三渲染转变指示指示到第四渲染配置的转变;

由所述控制系统确定与所述第四渲染配置相对应的第五组扬声器激活;以及

由所述控制系统执行到所述第五组扬声器激活的第四转变,而无需完成所述第一转变、所述第二转变或所述第三转变。

8.如权利要求1至5中任一项所述的方法,进一步包括:

由所述控制系统并且经由所述接口系统顺序地接收第二至第(N)渲染转变指示;

由所述控制系统确定与所述第二至第(N)渲染转变指示相对应的第四至第(N+2)组扬声器激活;

由所述控制系统顺序地执行从所述第四组扬声器激活到第(N+1)组扬声器激活的第三至第(N)转变;以及

由所述控制系统执行到所述第(N+2)组扬声器激活的第(N+1)转变,而无需完成所述第一至第(N)转变中的任一项。

9.如权利要求1至5中任一项所述的方法,进一步包括:

由所述控制系统并且经由所述接口系统接收第二渲染转变指示,所述第二渲染转变指示指示到第三渲染配置的转变;

由所述控制系统确定与所述第三渲染配置的简化版本相对应的第四组扬声器激活;

由所述控制系统执行从所述第三组扬声器激活到所述第四组扬声器激活的第三转变;

由所述控制系统确定与所述第三渲染配置的完整版本相对应的第五组扬声器激活;以及

由所述控制系统执行到所述第五组扬声器激活的第四转变,而无需完成所述第一转变、所述第二转变或所述第三转变。

10.如权利要求1至5中任一项所述的方法,进一步包括:

由所述控制系统并且经由所述接口系统顺序地接收第二至第(N)渲染转变指示;

由所述控制系统为所述第二至第(N)渲染转变指示中的每一个确定第一组扬声器激活和第二组扬声器激活,对于所述第二至第(N)渲染转变指示中的每一个,所述第一组扬声器激活对应于渲染配置的简化版本,而所述第二组扬声器激活对应于渲染配置的完整版本;

由所述控制系统顺序地执行从第四组扬声器激活到第(2N)组扬声器激活的第三至第(2N-1)转变;以及

由所述控制系统执行到第(2N+1)组扬声器激活的第(2N)转变,而无需完成第一至第(2N)转变中的任一项。

11.如权利要求1至10中任一项所述的方法,其中,单个渲染器实例渲染所述音频数据以供再现。

12.如权利要求1至11中任一项所述的方法,其中,渲染所述音频数据以供再现包括根据所述渲染配置确定单组插值激活,并且应用所述单组插值激活来产生单组经渲染的音频

信号。

13. 如权利要求12所述的方法,其中,所述单组经渲染的音频信号被馈送到一组扩音器延迟线中,所述一组扩音器延迟线包括用于多个扩音器中的每个扩音器的一个扩音器延迟线。

14. 如权利要求1至13中任一项所述的方法,其中,渲染所述音频数据以供再现还涉及在频域中确定和实施扩音器延迟。

15. 如权利要求14所述的方法,其中,在所述频域中确定和实施扬声器延迟涉及确定和实施变换块延迟和由频域滤波器系数应用的子块延迟的组合,所述子块延迟是允许不是频域变换块大小的确切倍数的延迟的残差相位项。

16. 如权利要求14或权利要求15所述的方法,其中,渲染所述音频数据以供再现还涉及实施具有单独读取偏移的一组变换块延迟线。

17. 如权利要求14至16中任一项所述的方法,其中,渲染所述音频数据以供再现还涉及实施子块延迟滤波。

18. 如权利要求17所述的方法,其中,实施所述子块延迟滤波涉及跨所述频域变换的块实施多抽头滤波器。

19. 如权利要求1至18中任一项所述的方法,其中,所述渲染所述音频数据以供再现包括为每个渲染配置确定和应用插值扬声器激活和交叉渐变窗口。

20. 如权利要求19所述的方法,其中,渲染所述音频数据以供再现还涉及实施具有单独延迟线读取偏移的一组变换块延迟线,其中,交叉渐变窗口选择至少部分地基于所述延迟线读取偏移,并且其中,如果所述延迟线读取偏移不相同,则所述交叉渐变窗口被设计为具有单位幂和。

21. 如权利要求1至20中任一项所述的方法,其中,所述渲染所述音频数据以供再现是在频域中执行的。

22. 如权利要求1至21中任一项所述的方法,其中,所述第一组扬声器激活用于三维空间中对应的多个位置中的每一个。

23. 如权利要求1至22中任一项所述的方法,其中,所述空间数据包括位置元数据。

24. 如权利要求1至21中任一项所述的方法,其中,所述第一组扬声器激活对应于基于声道的音频格式。

25. 如权利要求24所述的方法,其中,所述预期感知空间位置包括所述基于声道的音频格式的声道。

26. 一种装置,所述装置被配置成执行如权利要求1至25中任一项所述的方法。

27. 一种系统,所述系统被配置成执行如权利要求1至25中任一项所述的方法。

28. 一种或多种非暂态介质,其上存储有软件,所述软件包括用于控制一个或多个设备执行如权利要求1至25中任一项所述的方法的指令。

动态应用的渲染配置的渐进计算和应用

[0001] 发明人:Joshua B.Lando和Alan J.Seefeldt

[0002] 相关申请的交叉引用

[0003] 本申请要求以下申请的优先权:于2020年12月3日提交的美国临时申请63/121,108以及于2021年5月21日提交的美国临时申请63/202,003,这些申请中的每一个通过援引以其全文并入本文。

技术领域

[0004] 本公开涉及用于渲染音频以供一组扬声器中的一些或所有扬声器(例如,每个激活的扬声器)回放的系统和方法。

背景技术

[0005] 音频设备,包括但不限于智能音频设备,已经被广泛部署并且正在成为许多家庭的共同特征。尽管用于控制音频设备的现有系统和方法提供了益处,但改进的系统和方法将仍是期望的。

[0006] 符号和术语

[0007] 贯穿本公开,包括在权利要求书中,术语“扬声器(speaker)”、“扩音器(loudspeaker)”和“音频再现换能器”同义地用于表示任何发声换能器(或一组换能器)。典型的耳机组包括两个扬声器。扬声器可以被实施为包括多个换能器(例如,低音扬声器和高音扬声器),换能器可以由单个公共扬声器馈送或多个扬声器馈送驱动。在一些示例中,(多个)扬声器馈送可以在耦接到不同换能器的不同电路分支中经历不同处理。

[0008] 贯穿本公开,包括在权利要求中,在广义上使用“对”信号或数据执行操作的表达(例如,对信号或数据进行滤波、缩放、变换或应用增益)来表示直接对信号或数据执行操作或对信号或数据的已处理版本(例如,在对其执行操作之前已经历了初步滤波或预处理的信号版本)执行操作。

[0009] 贯穿本公开,包括在权利要求中,在广义上使用表达“系统”来表示设备、系统或子系统。例如,实施解码器的子系统可以被称为解码器系统,并且包括这样的子系统的系统(例如,响应于多个输入而生成X个输出信号的系统,其中,子系统生成M个输入,而其他X-M个输入是从外部源接收的)也可以被称为解码器系统。

[0010] 贯穿本公开,包括在权利要求中,在广义上使用术语“处理器”来表示可编程或以其他方式可配置(例如,用软件或固件)为对数据(例如,音频、视频或其他图像数据)执行操作的系统或设备。处理器的示例包括现场可编程门阵列(或其他可配置集成电路或芯片组)、被编程和/或以其他方式被配置成对音频或其他声音数据执行流水线式处理的数字信号处理器、可编程通用处理器或计算机、以及可编程微处理器芯片或芯片组。

[0011] 贯穿本公开,包括在权利要求中,术语“耦接(couples)”或“耦接(coupled)”用于意指直接或间接连接。因此,如果第一设备耦接到第二设备,则连接可以通过直接连接或者通过经由其他设备和连接的间接连接实现。

[0012] 如本文所使用的,“智能设备”是可以在某种程度上交互地和/或自主地操作的电子设备,其通常被配置用于经由比如蓝牙、Zigbee、近场通信、Wi-Fi、光保真(Li-Fi)、3G、4G、5G等各种无线协议与一个或多个其他设备(或网络)进行通信。若干种值得注意的智能设备类型是智能电话、智能汽车、智能恒温器、智能门铃、智能锁、智能冰箱、平板手机和平板电脑、智能手表、智能手环、智能钥匙链和智能音频设备。术语“智能设备”还可以是指展现出比如人工智能等普适计算的一些性质的设备。

[0013] 在本文中,使用表达“智能音频设备”来表示智能设备,其是单一用途音频设备或多用途音频设备(例如,实施虚拟助理功能的至少一些方面的音频设备)。单一用途音频设备是包括或耦接到至少一个麦克风(并且可选地还包括或耦接到至少一个扬声器和/或至少一个相机)并且很大程度上或主要被设计为实现单一用途的设备(例如,电视(TV))。例如,尽管TV通常可以播放(并且被认为能够播放)来自节目素材的音频,但在大多数实例中,现代TV运行某种操作系统,应用程序(包括看电视的应用程序)在操作系统上本地运行。从这个意义上说,具有(多个)扬声器和(多个)麦克风的单一用途音频设备通常被配置成运行本地应用程序和/或服务以直接使用(多个)扬声器和(多个)麦克风。一些单一用途音频设备可以被配置成组合在一起以实现在一定区或用户配置区域上播放音频。

[0014] 一种常见类型的多用途音频设备是实施虚拟助理功能的至少一些方面的音频设备,尽管虚拟助理功能的其他方面可以由比如一个或多个服务器等一个或多个其他设备来实施,多用途音频设备被配置用于与一个或多个服务器通信。这样的多用途音频设备在本文中可以被称为“虚拟助理”。虚拟助理是包括或耦接到至少一个麦克风(并且可选地还包括或耦接到至少一个扬声器和/或至少一个相机)的设备(例如,智能扬声器或语音助理集成设备)。在一些示例中,虚拟助理可以提供将多个设备(不同于虚拟助理)用于某种意义上支持云的应用程序或以其他方式未在虚拟助理本身中或之上完全实施的应用程序的能力。换句话说,虚拟助理功能的至少一些方面(例如,言语识别功能)可以(至少部分地)由一个或多个服务器或其他设备实施,虚拟助理可以经由网络(比如互联网)与一个或多个服务器或其他设备通信。虚拟助理有时可以一起工作,例如,以离散和有条件地定义的方式。例如,两个或更多个虚拟助理可以在其中之一(例如,最确信已经听到唤醒词的虚拟助理)对唤醒词做出响应的意义上一起工作。在一些实施方式中,连接的虚拟助理可以形成一种星座,星座可以由一个主应用程序管理,主应用程序可以是(或实施)虚拟助理。

[0015] 在本文中,“唤醒词”在广义上用于表示任何声音(例如,人类说出的词或其他声音),其中智能音频设备被配置成响应于检测到(“听到”)声音(使用包括在智能音频设备中或耦接到智能音频设备的至少一个麦克风,或至少一个其他麦克风)而唤醒。在该上下文中,“唤醒”表示设备进入等待(换句话说,正在收听)声音命令的状态。在一些实例中,本文中所谓的“唤醒词”可以包括超过一个词,例如,短语。

[0016] 在本文中,表达“唤醒词检测器”表示被配置成连续搜索实时声音(例如,言语)特征与训练模型之间的对齐的设备(或表示包括用于将设备配置成连续搜索实时声音特征与训练模型之间的对齐的指令的软件)。通常,每当唤醒词检测器确定检测到唤醒词的概率超过预定义阈值,就会触发唤醒词事件。例如,阈值可以是被调整以在错误接受率与错误拒绝率之间给出合理折衷的预定阈值。在唤醒词事件之后,设备可能会进入一种状态(可以被称为“唤醒”状态或“注意力”状态),在状态下设备会收听命令并且将接收到的命令传递给更

大、计算更密集的认可器。

[0017] 如本文所使用的,术语“节目流”和“内容流”是指一个或多个音频信号的集合,并且在一些实例中是指视频信号的集合,信号集合的至少部分是旨在要一起听到的。示例包括音乐、电影声轨、电影、电视节目、电视节目的音频部分、播客、实时语音通话、来自智能助理的合成语音响应等的选择。在一些实例中,内容流可以包括音频信号的至少一部分的多个版本,例如,多于一种语言的相同对话。在这样的实例中,一次旨在再现音频数据或其部分的仅一个版本(例如,与单一语言相对应的版本)。

发明内容

[0018] 本公开的至少一些方面可以经由方法来实施。一些这样的方法可以涉及音频处理。例如,一些方法可以涉及由控制系统并且经由接口系统接收音频数据。音频数据可以包括一个或多个音频信号和相关联的空间数据。空间数据可以指示与音频信号相对应的预期感知空间位置。在一些示例中,空间数据可以是或者可以包括位置元数据。根据一些示例,空间数据可以是、可以包括或者可以对应于基于声道的音频格式的声道。

[0019] 在一些示例中,方法可以涉及由控制系统渲染音频数据以经由环境的一组扩音器再现,从而产生第一渲染的音频信号。在一些这样的示例中,渲染音频数据以供再现可以涉及根据第一渲染配置确定环境中的一组扩音器的第一相对激活。第一渲染配置可以对应于第一组扬声器激活。在一些示例中,方法可以涉及经由接口系统向环境的一组扩音器中的至少一些扩音器提供第一渲染的音频信号。

[0020] 根据一些示例,方法可以涉及由控制系统并且经由接口系统接收第一渲染转变指示。第一渲染转变指示可以例如指示从第一渲染配置到第二渲染配置的转变。

[0021] 在一些示例中,方法可以涉及由控制系统确定第二组扬声器激活。根据该示例,第二组扬声器激活对应于第二渲染配置的简化版本。然而,在其他示例中,第二组扬声器激活可以对应于第二渲染配置的完整、全保真版本。

[0022] 根据一些示例,方法可以涉及由控制系统执行从第一组扬声器激活到第二组扬声器激活的第一转变。在一些示例中,方法可以涉及由控制系统确定第三组扬声器激活。根据该示例,第三组扬声器激活对应于第二渲染配置的完整版本。在一些示例中,方法可以涉及由控制系统执行到第三组扬声器激活的第二转变,而无需完成第一转变。在一些示例中,单个渲染器实例可以渲染音频数据以供再现。

[0023] 在一些示例中,第一组扬声器激活、第二组扬声器激活和第三组扬声器激活可以是频率相关的扬声器激活。根据一些这样的示例,频率相关的扬声器激活可以与在至少第一频带中应用感知空间位置的模型相对应和/或通过在至少第一频带中应用感知空间位置的模型而产生,模型产生与收听者的左耳和右耳处的音频对象位置相对应的双耳响应。

[0024] 在一些示例中,频率相关的扬声器激活可以与在至少第二频带中应用感知空间位置模型相对应和/或通过在至少第二频带中应用感知空间位置模型而产生,模型将从一组扩音器播放的音频信号的感知空间位置置于由扩音器的相关联的激活增益加权的一组扩音器的位置的质心处。

[0025] 根据一些示例,第一组扬声器激活、第二组扬声器激活和/或第三组扬声器激活可以至少部分地基于成本函数。在一些这样的示例中,第一组扬声器激活、第二组扬声器激活

和/或第三组扬声器激活可以是优化成本的结果,成本是以下各项的函数:当在环境中的一组扩音器上回放时播放的音频信号的感知空间位置的模型;音频信号的预期感知空间位置与一组扩音器中的每个扩音器的位置的接近度的度量;和/或一个或多个附加动态可配置功能。在一些这样的示例中,一个或多个附加动态可配置功能可以基于以下各项中的一项或多项:扩音器与一个或多个收听者的接近度;扩音器与吸引力位置的接近度(其中,吸引力可以是有利于更接近吸引力位置的相对更高的扩音器激活的因素);扩音器与排斥力位置的接近度(其中,排斥力可以是有利于更接近排斥力位置的相对更低的扩音器激活的因素);每个扩音器相对于环境中的其他扩音器的能力;扩音器关于其他扩音器的同步;唤醒词性能;和/或回声消除器性能。

[0026] 根据一些示例,方法可以涉及由控制系统并且经由接口系统接收第二渲染转变指示。根据一些这样的示例,第二渲染转变指示可以指示到第三渲染配置的转变。在一些这样的示例中,方法可以涉及由控制系统确定与第三渲染配置相对应的第四组扬声器激活。在一些这样的示例中,方法可以涉及由控制系统执行到第四组扬声器激活的第三转变,而无需完成第一转变或第二转变。在一些示例中,方法可以涉及由控制系统并且经由接口系统接收第三渲染转变指示。在一些这样的示例中,第三渲染转变指示可以指示到第四渲染配置的转变。在一些这样的示例中,方法可以涉及由控制系统确定与第四渲染配置相对应的第五组扬声器激活。在一些这样的示例中,方法可以涉及由控制系统执行到第五组扬声器激活的第四转变,而无需完成第一转变、第二转变或第三转变。

[0027] 在一些示例中,方法可以涉及由控制系统并且经由接口系统顺序地接收第二至第(N)渲染转变指示。在一些这样的示例中,方法可以涉及由控制系统确定与第二至第(N)渲染转变指示相对应的第四至第(N+2)组扬声器激活。在一些这样的示例中,方法可以涉及由控制系统顺序地执行从第四组扬声器激活到第(N+1)组扬声器激活的第三至第(N)转变。在一些这样的示例中,方法可以涉及由控制系统执行到第(N+2)组扬声器激活的第(N+1)转变,而无需完成第一至第(N)转变中的任一项。

[0028] 根据一些示例,方法可以涉及由控制系统并且经由接口系统接收第二渲染转变指示。在一些实例中,第二渲染转变指示可以指示到第三渲染配置的转变。在一些这样的示例中,方法可以涉及由控制系统确定与第三渲染配置的简化版本相对应的第四组扬声器激活。在一些这样的示例中,方法可以涉及由控制系统执行从第三组扬声器激活到第四组扬声器激活的第三转变。在一些这样的示例中,方法可以涉及由控制系统确定与第三渲染配置的完整版本相对应的第五组扬声器激活。在一些这样的示例中,方法可以涉及由控制系统执行到第五组扬声器激活的第四转变,而无需完成第一转变、第二转变或第三转变。

[0029] 在一些示例中,方法可以涉及由控制系统并且经由接口系统顺序地接收第二至第(N)渲染转变指示。在一些这样的示例中,方法可以涉及由控制系统为第二至第(N)渲染转变指示中的每一个确定第一组扬声器激活和第二组扬声器激活。在一些这样的示例中,第一组扬声器激活可以对应于渲染配置的简化版本,而第二组扬声器激活可以对应于第二至第(N)渲染转变指示中的每一个的渲染配置的完整版本。在一些这样的示例中,方法可以涉及由控制系统顺序地执行从第四组扬声器激活到第(2N)组扬声器激活的第三至第(2N-1)转变。在一些这样的示例中,方法可以涉及由控制系统执行到第(2N+1)组扬声器激活的第(2N)转变,而无需完成第一至第(2N)转变中的任一项。

[0030] 根据一些示例,渲染音频数据以供再现可以涉及根据渲染配置确定单组插值激活,并且应用单组插值激活来产生单组经渲染的音频信号。在一些这样的示例中,单组经渲染的音频信号可以被馈送到一组扩音器延迟线中。在一些这样的示例中,一组扩音器延迟线可以包括用于多个扩音器中的每个扩音器的一个扩音器延迟线。

[0031] 在一些示例中,渲染音频数据以供再现可以在频域中执行。在一些这样的示例中,渲染音频数据以供再现可以涉及在频域中确定和实施扩音器延迟。在一些这样的示例中,在频域中确定和实施扬声器延迟可以涉及确定和实施变换块延迟和由频域滤波器系数应用的子块延迟的组合。在一些这样的示例中,子块延迟可以是允许不是频域变换块大小的确切倍数的延迟的残差相位项。在一些示例中,渲染音频数据以供再现可以涉及实施具有单独读取偏移的一组变换块延迟线。

[0032] 在一些示例中,渲染音频数据以供再现可以涉及实施子块延迟滤波。在一些这样的示例中,实施子块延迟滤波可以涉及跨频域变换的块实施多抽头滤波器。

[0033] 根据一些示例,渲染音频数据以供再现可以涉及为每个渲染配置确定和应用插值扬声器激活和交叉渐变窗口。在一些这样的示例中,渲染音频数据以供再现可以涉及实施具有单独延迟线读取偏移的一组变换块延迟线。在一些这样的示例中,交叉渐变窗口选择可以至少部分地基于延迟线读取偏移。在一些这样的示例中,如果延迟线读取偏移不相同,则交叉渐变窗口可以被设计为具有单位累和。

[0034] 在一些示例中,第一组扬声器激活可以用于三维空间中对应的多个位置中的每一个。然而,根据一些示例,第一组扬声器激活可以对应于基于声道的音频格式。在一些这样的示例中,预期感知空间位置可以与基于声道的音频格式的声道相对应。

[0035] 本文描述的一些或所有操作、功能和/或方法可以由一个或多个设备根据存储在一种或多种非暂态介质上的指令(例如,软件)来执行。这种非暂态介质可以包括如本文描述的存储器设备,包括但不限于随机存取存储器(RAM)设备、只读存储器(ROM)设备等。因此,本公开中描述的主题的一些创新方面可以在其上存储有软件的一种或多种非暂态介质中实施。

[0036] 本公开的至少一些方面可以经由装置来实施。例如,一个或多个设备可以能够至少部分地执行本文公开的方法。在一些实施方式中,装置可以包括接口系统和控制系统。控制系统可以包括一个或多个通用单芯片或多芯片处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或其他可编程逻辑设备、离散门或晶体管逻辑、离散硬件部件或其组合。在一些示例中,该装置可以是上文引用的音频设备之一。然而,在一些实施方式中,该装置可以是另一种类型的设备,如移动设备、膝上型计算机、服务器等。

[0037] 在以下附图和说明中阐述了本说明书所描述的主题的一个或多个实施方式的细节。从描述、附图和权利要求中,其他特征、方面和优点将变得显而易见。注意,以下附图的相对尺寸可能不是按比例来绘制的。

附图说明

[0038] 图1A是示出了能够实施本公开的各个方面的装置的部件的示例的框图。

[0039] 图1B是实施例的最小版本的框图。

[0040] 图2A描绘了具有附加特征的另一个(更有能力的)实施例。

[0041] 图2B是概述了可以由如图1A、图1B或图2A中示出的装置或系统等装置或系统执行的方法的一个示例的流程图。

[0042] 图2C和图2D是图示了示例的一组扬声器激活和对象渲染位置的图。

[0043] 图2E是概述了可以由如图1A中示出的装置或系统等装置或系统执行的方法的一个示例的流程图。

[0044] 图2F是示例实施例中的扬声器激活的图。

[0045] 图2G是示例实施例中的对象渲染位置的图。

[0046] 图2H是示例实施例中的扬声器激活的图。

[0047] 图2I是示例实施例中的对象渲染位置的图。

[0048] 图2J是示例实施例中的扬声器激活的图。

[0049] 图2H是示例实施例中的扬声器激活的图。

[0050] 图2I是示例实施例中的对象渲染位置的图。

[0051] 图2J是示例实施例中的扬声器激活的图。

[0052] 图2K是示例实施例中的对象渲染位置的图。

[0053] 图3A和图3B示出了连接的生活空间的平面图示例。

[0054] 图4A和图4B示出了提供空间音乐混音和语音助理响应的同时回放的多流渲染器的示例。

[0055] 图5A、图5B和图5C图示了所公开的多流渲染器的第三示例用例。

[0056] 图6示出了图1B中示出的多流渲染器的频域/变换域示例。

[0057] 图7示出了图2A中示出的多流渲染器的频域/变换域示例。

[0058] 图8示出了具有音频流响度估计器的多流渲染系统的实施方式。

[0059] 图9A示出了被配置用于多个经渲染的流的交叉渐变的多流渲染系统的示例。

[0060] 图9B是示例实施例中指示扬声器激活的点的图。

[0061] 图10是根据一个示例的指示扬声器激活的点之间的三线性插值的图。

[0062] 图11A和图11B示出了提供渲染配置计算服务的示例。

[0063] 图12A和图12B示出了渲染配置转变的示例。

[0064] 图13呈现了与用于管理渲染配置之间的转变的替代性实施方式相对应的块。

[0065] 图14呈现了与用于管理渲染配置之间的转变的另一实施方式相对应的块。

[0066] 图15呈现了根据一个示例的与频域渲染器相对应的块。

[0067] 图16呈现了与用于管理渲染配置之间的转变的另一实施方式相对应的块。

[0068] 图17示出了具有单位幂和的交叉渐变窗口对的示例。

[0069] 图18A和图18B呈现了具有单位和的交叉渐变窗口对的示例。

[0070] 图19呈现了与用于管理第一至第L组扬声器激活之间的转变的另一实施方式相对应的块。

[0071] 图20A示出了具有单位幂和的交叉渐变窗口对的示例。

[0072] 图20B示出了具有单位和的交叉渐变窗口对的示例。

[0073] 图21、图22、图23、图24和图25是呈现了没有、具有一些或所有匹配的读取偏移的交叉渐变窗口的示例的图。

[0074] 图26、图27、图28、图29和图30图示了与图21至图25相同的情况，但限制了3种激活

频节目流的同时回放的许多其他情况。多流渲染系统的一些实施方式可以被配置成执行以下操作中的一项或多项：

[0085] ●在多个任意放置的扩音器上同时渲染和回放多个音频节目流，其中，该节目流中的至少一个是空间混音。

[0086] ○术语“节目流”是指旨在作为一个整体被收听的一个或多个音频信号的集合。示例包括音乐选择、电影声轨、播客、现场语音通话、来自智能助理的合成语音响应等。

[0087] ○“空间混音”是旨在在收听者的左耳和右耳处传递不同信号（不仅仅是单声道）的节目流。用于空间混音的音频格式的示例包括立体声、5.1和7.1环绕声、对象音频格式（如杜比全景声(Dolby Atmos)和高保真度立体声响复制(Ambisonics)）。

[0088] ○“渲染”节目流是指跨多个扩音器主动分布相关联的一个或多个音频信号以实现特定感知印象的过程。

[0089] ●根据附加节目流中的一个或多个的渲染动态地修正至少一个空间混音的渲染。对空间混音的渲染的这种修正的示例包括但不限于

[0090] ○根据与一个或多个附加节目流中的至少一个的渲染相关联的扩音器的相对激活来修正多个扩音器的相对激活。

[0091] ○根据一个或多个附加节目流中的至少一个的渲染的空间性质来扭曲空间混音的预期空间平衡。

[0092] ○根据该一个或多个附加节目流中的至少一个的响度或可听度来修正该空间混音的响度或可听度。

[0093] 图1A是示出了能够实施本公开的各个方面的装置的部件的示例的框图。根据一些示例，装置100可以是或者可以包括被配置用于执行本文公开的方法中的至少一些方法的智能音频设备。在其他实施方式中，装置100可以是或者可以包括被配置用于执行本文公开的方法中的至少一些方法的另一种设备，如膝上型计算机、蜂窝电话、平板设备、智能家居中枢等。在一些这样的实施方式中，装置100可以是或者可以包括服务器。在一些实施方式中，装置100可以被配置成实施本文中可以被成为“音频会话管理器”的设备。

[0094] 在该示例中，装置100包括接口系统105和控制系统110。在一些实施方式中，接口系统105可以被配置用于与正在执行或被配置用于执行软件应用程序的一个或多个设备通信。这样的软件应用程序在本文中有时可以被成为“应用程序”或简称为“app”。在一些实施方式中，接口系统105可以被配置用于交换控制信息和与应用程序有关的相关联的数据。在一些实施方式中，接口系统105可以被配置用于与音频环境的一个或多个其他设备通信。在一些示例中，音频环境可以是家庭音频环境。在一些实施方式中，接口系统105可以被配置用于与音频环境的音频设备交换控制信息和相关联的数据。在一些示例中，控制信息和相关联的数据可以与装置100被配置用于与其通信的一个或多个应用程序有关。

[0095] 在一些实施方式中，接口系统105可以被配置用于接收音频节目流。音频节目流可以包括被安排由环境的至少一些扬声器再现的音频信号。音频节目流可以包括如声道数据和/或空间元数据等空间数据。在一些实施方式中，接口系统105可以被配置用于从环境中的一个或多个麦克风接收输入。

[0096] 接口系统105可以包括一个或多个网络接口和/或一个或多个外部设备接口（如一个或多个通用串行总线(USB)接口）。根据一些实施方式，接口系统105可以包括一个或多个

无线接口。接口系统105可以包括用于实施用户接口的一个或多个设备,如一个或多个麦克风、一个或多个扬声器、显示系统、触摸传感器系统和/或手势传感器系统。在一些示例中,接口系统105可以包括控制系统110与存储器系统(如图1A中示出的可选存储器系统115)之间的一个或多个接口。然而,在一些实例中,控制系统110可以包括存储器系统。

[0097] 例如,控制系统110可以包括通用单芯片或多芯片处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或其他可编程逻辑设备、离散门或晶体管逻辑、和/或离散硬件部件。

[0098] 在一些实施方式中,控制系统110可以驻留在超过一个设备中。例如,控制系统110的一部分可以驻留在本文描绘的环境之一内的设备中,并且控制系统110的另一部分可以驻留在环境之外的设备中,如服务器、移动设备(例如,智能电话或平板电脑)等。在其他示例中,控制系统110的一部分可以驻留在本文描绘的环境之一内的设备中,并且控制系统110的另一部分可以驻留在环境的一个或多个其他设备中。例如,控制系统功能可以跨环境的多个智能音频设备分布,或者可以由编排设备(如本文中可以被称为智能家居中枢的设备)和环境的一个或多个其他设备共享。在一些这样的示例中,接口系统105也可以驻留在多于一个设备中。

[0099] 在一些实施方式中,控制系统110可以被配置用于至少部分地执行本文公开的方法。根据一些示例,控制系统110可以被配置用于实施管理在多个扬声器上回放多个音频流的方法。

[0100] 本文描述的一些或所有方法可以由一个或多个设备根据存储在一个或多个非暂态介质上的指令(例如,软件)来执行。这种非暂态介质可以包括如本文描述的存储器设备,包括但不限于随机存取存储器(RAM)设备、只读存储器(ROM)设备等。该一个或多个非暂态介质可以例如驻留在图1A中示出的可选存储器系统115和/或控制系统110中。因此,可以在其上存储有软件的一个或多个非暂态介质中实施本公开中描述的主题的各个创新方面。例如,软件可以包括用于控制至少一个设备来处理音频数据的指令。例如,软件可以由如图1A的控制系统110等控制系统的一个或多个部件执行。

[0101] 在一些示例中,装置100可以包括图1A中示出的可选麦克风系统120。可选麦克风系统120可以包括一个或多个麦克风。在一些实施方式中,一个或多个麦克风可以是另一个设备(如扬声器系统的扬声器、智能音频设备等)的一部分或与其相关联。在一些示例中,装置100可以不包括麦克风系统120。然而,在一些这样的实施方式中,装置100仍然可以被配置成经由接口系统110接收音频环境中的一个或多个麦克风的麦克风数据。

[0102] 根据一些实施方式,装置100可以包括图1A中示出的可选扩音器系统125。可选扩音器系统125可以包括一个或多个扩音器,该扩音器在本文中也可以被称为“扬声器”。在一些示例中,可选扩音器系统125的至少一些扩音器可以任意地定位。例如,可选扩音器系统125的至少一些扬声器可以放置在不与任何标准规定的扩音器布局相对应的位置,如杜比5.1、杜比5.1.2、杜比7.1、杜比7.1.4、杜比9.1、滨崎22.2等。在一些这样的示例中,可选扬声器系统125的至少一些扩音器可以放置在对空间方便的位置(例如,在有空间容纳扩音器的位置),但不在任何标准规定的扩音器布局中。在一些示例中,装置100可以不包括扩音器系统125。

[0103] 在一些实施方式中,装置100可以包括图1A中示出的可选传感器系统129。可选传

传感器系统129可以包括一个或多个相机、触摸传感器、手势传感器、运动检测器等。根据一些实施方式,可选传感器系统129可以包括一个或多个相机。在一些实施方式中,相机可以是独立式相机。在一些示例中,可选传感器系统129的一个或多个相机可以驻留在智能音频设备中,该智能音频设备可以是单一用途音频设备或虚拟助理。在一些这样的示例中,可选传感器系统129的一个或多个相机可以驻留在TV、移动电话或智能扬声器中。在一些示例中,装置100可以不包括传感器系统129。然而,在一些这样的实施方式中,装置100仍然可以被配置成经由接口系统110接收音频环境中的一个或多个传感器的传感器数据。

[0104] 在一些实施方式中,装置100可以包括图1A中示出的可选显示系统135。可选显示系统135可以包括一个或多个显示器,如一个或多个发光二极管(LED)显示器。在一些实例中,可选显示系统135可以包括一个或多个有机发光二极管(OLED)显示器。在其中装置100包括显示系统135的一些示例中,传感器系统129可以包括接近显示系统135的一个或多个显示器的触摸传感器系统和/或手势传感器系统。根据一些这样的实施方式,控制系统110可以被配置用于控制显示系统135来呈现一个或多个图形用户界面(GUI)。

[0105] 根据一些这样的示例,装置100可以是或者可以包括智能音频设备。在一些这样的实施方式中,装置100可以是或者可以包括唤醒词检测器。例如,装置100可以是或者可以包括虚拟助理。

[0106] 图1B是实施例的最小版本的框图。描绘了N个节目流($N \geq 2$),其中第一个明确标记为空间,其对应的音频信号集合馈送通过对应的渲染器,该渲染器各自单独被配置用于通过公共的一组M个任意间隔开的扩音器($M \geq 2$)来回放其对应的节目流。渲染器在本文中也可以被称为“渲染模块”。渲染模块和混音器130a可以经由软件、硬件、固件或其某种组合来实施。在该示例中,渲染模块和混音器130a经由控制系统110a实施,该控制系统110a是上文参考图1A描述的控制系统110的实例。N个渲染器中的每一个输出跨所有N个渲染器相加的一组M个扩音器馈送,用于在M个扩音器上同时回放。根据该实施方式,将有关收听环境内M个扩音器的布局的信息提供给所有渲染器,由从扩音器框馈送返回的虚线指示,以便渲染器可以正确地被配置用于通过扬声器进行回放。该布局信息可以从或可以不从一个或多个扬声器本身发送,取决于特定实施方式。根据一些示例,布局信息可以由被配置用于确定M个扩音器中的每一个在收听环境中的相对位置的一个或多个智能扬声器提供。一些这样的自动定位方法可以基于到达方向(DOA)方法或到达时间(TOA)方法。在其他示例中,该布局信息可以由另一个设备确定和/或由用户输入。在一些示例中,可以将关于收听环境内M个扩音器中的至少一些扩音器的能力的扩音器规格信息提供给所有渲染器。这种扩音器规格信息可以包括阻抗、频率响应、灵敏度、额定功率、单独驱动器的数量和位置等。根据该示例,来自附加节目流中的一个或多个附加节目流的渲染的信息被馈送到主要空间流的渲染器中,使得可以根据该信息动态地修正该渲染。该信息由从渲染框2到渲染框N向上回到渲染框1的虚线表示。

[0107] 图2A描绘了具有附加特征的另一个(更有能力的)实施例。在该示例中,渲染模块和混音器130b经由控制系统110b实施,该控制系统110b是上文参考图1A描述的控制系统110的实例。在该版本中,在所有N个渲染器之间上下行进的虚线表示N个渲染器中的任何一个都可以有助于剩余N-1个渲染器中的任何一个的动态修正的想法。换句话说,N个节目流中的任何一个节目流的渲染可以根据剩余N-1个节目流中的任何一个的一个或多个渲染的

组合而动态地修正。另外地,节目流中的任何一个或多个节目流可以是空间混音,并且任何节目流的渲染,无论它是否是空间的,都可以根据其他节目流中的任何节目流来动态修正。例如,如上所述,可以将扩音器布局信息提供给N个渲染器。在一些示例中,可以将扩音器规格信息提供给N个渲染器。在一些实施方式中,麦克风系统120a可以在收听环境内包括一组K个麦克风($K \geq 1$)。在一些示例中,(多个)麦克风可以附接到扩音器中的一个或多个扩音器或与其相关联。这些麦克风可以将其捕获的音频信号(由实线表示)和附加配置信息(例如,其位置)(由虚线表示)两者馈送回到一组N个渲染器中。N个渲染器中的任何渲染器可以然后根据该附加麦克风输入进行动态修正。本文中提供了各种示例。

[0108] 从麦克风输入得到的并且随后用于动态修正N个渲染器中的任何渲染器的信息的示例包括但不限于:

[0109] ●对系统用户的特定词语或短语话语的检测。

[0110] ●对系统的一个或多个用户的位置的估计。

[0111] ●对收听空间中特定位置处的N个节目流的任何组合的响度的估计。

[0112] ●对收听环境中其他环境声音(如背景噪声)的响度的估计。

[0113] 图2B是概述了可以由如图1A、图1B或图2A中示出的装置或系统等装置或系统执行的方法的一个示例的流程图。与本文描述的其他方法一样,不必以所指示的顺序来执行方法200的框。此外,这样的方法可以包括比所示出和/或所描述的框更多或更少的框。方法200的框可以由一个或多个设备执行,该设备可以是(或可以包括)控制系统,如图1A、图1B和图2A中示出的和上文描述的控制系统110、控制系统110a或控制系统110b,或者其他公开的控制系统的示例之一。

[0114] 在该实施方式中,框205涉及经由接口系统接收第一音频节目流。在该示例中,第一音频节目流包括被安排由环境的至少一些扬声器再现的第一音频信号。此处,第一音频节目流包括第一空间数据。根据该示例,第一空间数据包括声道数据和/或空间元数据。在一些示例中,框205涉及经由接口系统接收第一音频节目流的控制系统的第一渲染模块。

[0115] 根据该示例,框210涉及渲染第一音频信号以经由环境的扬声器再现,从而产生第一经渲染的音频信号。例如,如上所述,方法200的一些示例涉及接收扩音器布局信息。例如,如上所述,方法200的一些示例涉及接收扩音器规格信息。在一些示例中,第一渲染模块可以至少部分地基于扩音器布局信息和/或扩音器规格信息产生第一经渲染的音频信号。

[0116] 在该示例中,框215涉及经由接口系统接收第二音频节目流。在该实施方式中,第二音频节目流包括被安排由环境的至少一些扬声器再现的第二音频信号。根据该示例,第二音频节目流包括第二空间数据。第二空间数据包括声道数据和/或空间元数据。在一些示例中,框215涉及经由接口系统接收第二音频节目流的控制系统的第二渲染模块。

[0117] 根据该实施方式,框220涉及渲染第二音频信号以经由环境的扬声器再现,从而产生第二经渲染的音频信号。在一些示例中,第二渲染模块可以至少部分地基于接收到的扩音器布局信息和/或接收到的扩音器规格信息产生第二经渲染的音频信号。

[0118] 在一些实例中,环境中的一些或所有扬声器可以任意地定位。例如,环境的至少一些扬声器可以放置在不对应于任何标准规定的扬声器布局(如杜比5.1、杜比7.1、滨崎22.2等)的位置。在一些这样的示例中,环境的至少一些扬声器可以放置在相对于环境的家具、墙等而言方便的位置(例如,在有空间容纳扬声器的位置),但不采用任何标准的规定的扬

声器布局。

[0119] 因此,一些实施方式框210或框220可以涉及灵活渲染到任意定位的扬声器。一些这样的实施方式可以涉及质心振幅平移(CMAP)、灵活虚拟化(FV)或两者的组合。从高水平来看,这两种技术渲染一组的一个或多个音频信号,每个音频信号具有相关联的期望的感知空间位置,用于在一组两个或更多个扬声器上回放,其中,该组扬声器的相对激活是通过扬声器回放的该音频信号的感知空间位置的模型以及音频信号的期望感知空间位置与扬声器位置的接近度的函数。该模型确保收听者在其预期空间位置附近听到音频信号,并且接近度项控制使用哪些扬声器来实现该空间印象。具体地,接近度项有利于激活靠近音频信号的期望感知空间位置的扬声器。对于CMAP和FV两者,该函数关系可以方便地从成本函数中得到,该成本函数写为两个项的和,一个项用于空间方面并且一个项用于接近度:

$$[0120] \quad C(\mathbf{g}) = C_{\text{spatial}}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) + C_{\text{proximity}}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) \quad (1)$$

[0121] 此处,集合 $\{\vec{s}_i\}$ 表示一组M个扩音器的位置, $\vec{o}$ 表示音频信号的期望感知空间位置,并且 $\mathbf{g}$ 表示扬声器激活的M维向量。对于CMAP,向量中的每个扬声器激活表示每个扬声器的增益,而对于FV,每个扬声器激活表示滤波器(在第二种情况下, $\mathbf{g}$ 可以等效地被视为特定频率处的复值的向量,并且跨多个频率计算不同的 $\mathbf{g}$ 以形成滤波器)。扬声器激活的最佳向量可以通过跨激活最小化成本函数找到:

$$[0122] \quad \mathbf{g}_{\text{opt}} = \min_{\mathbf{g}} C(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) \quad (2a)$$

[0123] 在成本函数的某些定义下,可能很难控制由上述最小化产生的最佳激活的绝对水平,尽管 $\mathbf{g}_{\text{opt}}$ 的分量之间的相对水平是适当的。为了解决该问题,可以执行 $\mathbf{g}_{\text{opt}}$ 的后续归一化,以便控制激活的绝对水平。例如,可以期望将向量归一化为具有单位长度,这符合常用的恒定功率平移规则:

$$[0124] \quad \bar{\mathbf{g}}_{\text{opt}} = \frac{\mathbf{g}_{\text{opt}}}{\|\mathbf{g}_{\text{opt}}\|} \quad (2b)$$

[0125] 在一些示例中,灵活渲染算法的确切行为可以取决于成本函数的 C_{spatial} 和 $C_{\text{proximity}}$ 这两个项的特定构建。对于CMAP, C_{spatial} 可以从模型得到,该模型将从一组扩音器播放的音频信号的感知空间位置放置在由扩音器的相关联的激活增益 g_i (向量 $\mathbf{g}$ 的元素)加权的这些扩音器的位置的质心:

$$[0126] \quad \vec{o} = \frac{\sum_{i=1}^M g_i \vec{s}_i}{\sum_{i=1}^M g_i} \quad (3)$$

[0127] 然后等式3可以被操纵成表示期望音频位置与由激活的扩音器产生的音频位置之间的平方误差的空间成本:

$$[0128] \quad C_{\text{spatial}}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) = \|(\sum_{i=1}^M g_i) \vec{o} - \sum_{i=1}^M g_i \vec{s}_i\|^2 = \|\sum_{i=1}^M g_i (\vec{o} - \vec{s}_i)\|^2 \quad (4)$$

[0129] 对于FV,成本函数的空间项被不同地定义。目标是在收听者的左耳和右耳处产生与音频对象位置 $\vec{o}$ 相对应的双耳响应 $\mathbf{b}$ 。概念上, $\mathbf{b}$ 是滤波器的 2×1 向量(每只耳朵一个滤波器),但更方便地将其视为特定频率下复值的 2×1 向量。继续以特定频率进行该表示,可以

按对象位置从一组HRTF索引中取得期望的双耳响应:

$$[0130] \quad \mathbf{b} = \text{HRTF}\{\vec{o}\} \quad (5)$$

[0131] 同时,由扩音器在收听者的耳朵处产生的 2×1 双耳响应 $\mathbf{e}$ 可以被建模为 $2 \times M$ 声学传输矩阵 $\mathbf{H}$ 乘以复扬声器激活值的 $M \times 1$ 向量 $\mathbf{g}$:

$$[0132] \quad \mathbf{e} = \mathbf{H}\mathbf{g} \quad (6)$$

[0133] 声学传输矩阵 $\mathbf{H}$ 可以基于扩音器位置的集合 $\{\vec{s}_i\}$ 相对于收听者位置建模。最后,成本函数的空间分量可以定义为期望的双耳响应(等式5)与扩音器产生的双耳响应(等

[0134] 式6)之间的平方误差:

$$[0135] \quad C_{\text{spatial}}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) = (\mathbf{b} - \mathbf{H}\mathbf{g})^*(\mathbf{b} - \mathbf{H}\mathbf{g}) \quad (7)$$

[0136] 方便地,等式4和7两者中定义的针对CMAP和FV的成本函数的空间项都可以重新布置为矩阵二次函数,作为扬声器激活 $\mathbf{g}$ 的函数:

$$[0137] \quad C_{\text{spatial}}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) = \mathbf{g}^* \mathbf{A} \mathbf{g} + \mathbf{B} \mathbf{g} + \mathbf{C} \quad (8)$$

[0138] 其中, $\mathbf{A}$ 表示 $M \times M$ 方阵, $\mathbf{B}$ 表示 $1 \times M$ 向量,并且 $\mathbf{C}$ 表示标量。在该示例中,矩阵 $\mathbf{A}$ 的秩为2,并且因此当 $M > 2$ 时,存在无限多个空间误差项等于零的扬声器激活 $\mathbf{g}$ 。引入成本函数的第二项 $C_{\text{proximity}}$ 去除了该不确定性,并且产生了与其他可能的解相比具有感知上有益性质的特定解。对于CMAP和FV两者, $C_{\text{proximity}}$ 可以被构建使得位置 $\vec{s}_i$ 远离期望的音频信号位置 $\vec{o}$ 的扬声器的激活比位置接近期望位置的扬声器的激活受到更多惩罚。该构建产生了稀疏的扬声器激活最佳集合,其中,仅接近期望的音频信号的位置的扬声器才会被显著激活,并且实际上导致音频信号的空间再现,这对于在该组扬声器周围的收听者移动可能在感知上更加稳健。

[0139] 为此,成本函数的第二项 $C_{\text{proximity}}$ 可以定义为扬声器激活的绝对值平方的距离加权和。这可以以矩阵形式简洁地表示为:

$$[0140] \quad C_{\text{proximity}}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) = \mathbf{g}^* \mathbf{D} \mathbf{g} \quad (9a)$$

[0141] 其中, $\mathbf{D}$ 表示期望的音频位置与每个扬声器之间距离惩罚的对角矩阵:

$$[0142] \quad \mathbf{D} = \begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_M \end{bmatrix}, \quad d_i = \text{distance}(\vec{o}, \vec{s}_i) \quad (9b)$$

[0143] 距离惩罚函数可以采取许多形式,但以下是有用的参数化:

$$[0144] \quad \text{distance}(\vec{o}, \vec{s}_i) = \alpha d_0^2 \left(\frac{\|\vec{o} - \vec{s}_i\|}{d_0} \right)^\beta \quad (9c)$$

[0145] 其中, $\|\vec{o} - \vec{s}_i\|$ 表示期望的音频位置与扬声器位置之间的欧几里得距离,并且 α 和 β 表示可调参数。参数 α 指示惩罚的全局强度; d_0 与距离惩罚的空间范围相对应(在大约 d_0 距离处或更远的扩音器将受到惩罚),并且 β 解释了在距离 d_0 处惩罚起始的突然性。

[0146] 将等式8和9a中定义的成本函数的两个项相结合,得出总体成本函数:

$$[0147] \quad C(\mathbf{g}) = \mathbf{g}^* \mathbf{A} \mathbf{g} + \mathbf{B} \mathbf{g} + \mathbf{C} + \mathbf{g}^* \mathbf{D} \mathbf{g} = \mathbf{g}^* (\mathbf{A} + \mathbf{D}) \mathbf{g} + \mathbf{B} \mathbf{g} + \mathbf{C} \quad (10)$$

[0148] 对于该示例,将该成本函数关于 $\mathbf{g}$ 的导数设置为零并且求解 $\mathbf{g}$ 产生最佳扬声器激活解:

$$[0149] \quad \mathbf{g}_{opt} = \frac{1}{2} (\mathbf{A} + \mathbf{D})^{-1} \mathbf{B} \quad (11)$$

[0150] 通常,等式11中的最优解可以产生值为负的扬声器激活。对于灵活渲染器的CMAP构建,这样的负激活可能是不期望的,并且因此在一些示例中,等式(11)可以在所有激活保持为正的情况下最小化。

[0151] 图2C和图2D是图示了示例的一组扬声器激活和对象渲染位置的图。在这些示例中,扬声器激活和对象渲染位置与4、64、165、-87和-4度的扬声器位置相对应。图2C示出了扬声器激活245a、250a、255a、260a和265a,其包括针对这些特定扬声器位置的等式11的最优解。图2D将单独扬声器位置绘制为正方形267、270、272、274和275,其分别与扬声器激活245a、250a、255a、260a和265a相对应。图2D还将大量可能的对象角度的理想对象位置(换句话说,要渲染音频对象的位置)示出为点276a,并且将针对这些对象的对应实际渲染位置示出为点278a,通过虚线279a连接到理想对象位置。

[0152] 一类实施例涉及用于渲染音频以供多个协调的(编排的)智能音频设备中的至少一个(例如,所有或一些)进行回放的方法。例如,用户的家庭中(的系统中)存在的一组智能音频设备可以被编排来处理各种同时用例,包括灵活渲染(根据实施例)音频以供智能音频设备中的所有或一些(即,由智能音频设备的所有或一些(多个)扬声器)回放。考虑了与系统的许多交互,这些交互需要对渲染进行动态修正。这种修正可以但不一定专注于空间保真度。

[0153] 一些实施例是用于渲染音频以供一组智能音频设备中的至少一个(例如,所有或一些)智能音频设备回放(或供另一组扬声器的至少一个(例如,所有或一些)扬声器回放)的方法。渲染可以包括成本函数的最小化,其中,成本函数包括至少一个动态扬声器激活项。这种动态扬声器激活项的示例可以包括(但不限于):

- [0154] ●扬声器与一个或多个收听者的接近度;
- [0155] ●扬声器与吸引力或排斥力的接近度;
- [0156] ●扬声器关于一些位置(例如,收听者位置或婴儿房)的可听度;
- [0157] ●扬声器的能力(例如,频率响应和失真);
- [0158] ●扬声器关于其他扬声器的同步;
- [0159] ●唤醒词性能;以及
- [0160] ●回声消除器性能。

[0161] (多个)动态扬声器激活项可以启用各种行为中的至少一种,包括将音频的空间呈现扭曲远离特定智能音频设备,使得该特定智能音频设备的麦克风可以更好地听到说话者或者使得可以更好地从智能音频设备的(多个)扬声器听到次级音频流。

[0162] 一些实施例实施渲染以供协调的(编排的)多个智能音频设备的(多个)扬声器回放。其他实施例实施渲染以供另一组扬声器的(多个)扬声器回放。

[0163] 将灵活渲染方法(根据一些实施例实施)与一组无线智能扬声器(或其他智能音频

设备)配对可以产生非常有能力并且易于使用的空间音频渲染系统。在考虑与这种系统的交互时,显然期望对空间渲染进行动态修正,以便针对在系统使用期间可能出现的其他目标进行优化。为了实现该目标,一类实施例增强了现有的灵活渲染算法(其中扬声器激活是先前公开的空间项和接近度项的函数),具有一个或多个附加动态可配置功能,该附加动态可配置功能取决于正在渲染的音频信号、扬声器组和/或其他外部输入的一个或多个属性。根据一些实施例,等式1中给出的现有灵活渲染的成本函数可以例如根据以下等式增加这些一个或多个附加依赖项:

$$[0164] \quad C(\mathbf{g}) = C_{\text{spatial}}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) + C_{\text{proximity}}(\mathbf{g}, \vec{o}, \{\vec{s}_i\}) + \sum_j C_j(\mathbf{g}, \{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j) \quad (12)$$

[0165] 在等式12中,项 $C_j(\mathbf{g}, \{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j)$ 表示附加成本项,并且 $\{\hat{o}\}$ 表示正在渲染的音频信号(例如,基于对象的音频节目)的一组一个或多个性质, $\{\hat{s}_i\}$ 表示正在渲染音频的扬声器的一组一个或多个性质,并且 $\{\hat{e}\}$ 表示一个或多个附加外部输入。每个项 $C_j(\mathbf{g}, \{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j)$ 返回成本作为与音频信号、扬声器和/或外部输入的一个或多个属性的组合相关的激活 g 的函数,一般由集合 $\{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j$ 表示。应当理解,在该示例中,集合 $\{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j$ 至少包含来自 $\{\hat{o}\}$ 、 $\{\hat{s}_i\}$ 或 $\{\hat{e}\}$ 中的任何一个的一个元素。

[0166] $\{\hat{o}\}$ 的示例包括但不限于:

- [0167] ● 音频信号的期望的感知空间位置;
- [0168] ● 音频信号的水平(可能随时间变化);和/或
- [0169] ● 音频信号的频谱(可能随时间变化)。

[0170] $\{\hat{s}_i\}$ 的示例包括但不限于:

- [0171] ● 扩音器在收听空间中的位置;
- [0172] ● 扩音器的频率响应;
- [0173] ● 扩音器的回放水平限制;
- [0174] ● 扬声器内部动态处理算法的参数,如限制器增益;
- [0175] ● 从每个扬声器到其他扬声器的声学传输的测量或估计;
- [0176] ● 扬声器上回声消除器性能的测量;和/或
- [0177] ● 扬声器关于彼此的相对同步。

[0178] $\{\hat{e}\}$ 的示例包括但不限于:

- [0179] ● 回放空间中一个或多个收听者或说话者的位置;
- [0180] ● 从每个扩音器到收听位置的声学传输的测量或估计;
- [0181] ● 从说话者到一组扩音器的声学传输的测量或估计;
- [0182] ● 回放空间中其他一些地标的位置;和/或
- [0183] ● 从每个扬声器到回放空间中一些其他地标的声学传输的测量或估计;

[0184] 使用等式12中定义的新成本函数,可以通过如先前在等式2a和2b中指定的关于 g 的最小化和可能的后归一化来找到最佳激活组。

[0185] 图2E是概述了可以由如图1A中示出的装置或系统等装置或系统执行的方法的一个示例的流程图。与本文描述的其他方法一样,不必以所指示的顺序来执行方法280的框。此外,这样的方法可以包括比所示出和/或所描述的框更多或更少的框。方法280的框可以由一个或多个设备执行,该设备可以是(或者可以包括)控制系统,如图1A中示出的控制系统110。

[0186] 在该实施方式中,框285涉及由控制系统并且经由接口系统接收音频数据。在该示例中,音频数据包括一个或多个音频信号和相关联的空间数据。根据该实施方式,空间数据指示与音频信号相对应的预期感知空间位置。在一些实例中,预期感知空间位置可以是明确的,例如,如由杜比全景声位置元数据等位置元数据所指示的。在其他实例中,预期感知空间位置可以是隐式的,例如,预期感知空间位置可以是与根据杜比5.1、杜比7.1的声道或其他基于声道的音频格式相关联的假定位置。在一些示例中,框285涉及经由接口系统接收音频数据的控制系统的渲染模块。

[0187] 根据该示例,框290涉及由控制系统渲染音频数据以经由环境的一组扩音器再现,从而产生经渲染的音频信号。在该示例中,渲染音频数据中包括的一个或多个音频信号中的每一个涉及通过优化成本函数来确定环境中一组扩音器的相对激活。根据该示例,成本是当在环境中的该组扩音器上回放时音频信号的感知空间位置的模型的函数。在该示例中,成本也是音频信号的预期感知空间位置与该组扩音器中的每个扩音器的位置的接近度的度量的函数。在该实施方式中,成本也是一个或多个附加动态可配置功能的函数。在该示例中,动态可配置功能基于以下各项中的一项或多项:扩音器与一个或多个收听者的接近度;扩音器与吸引力位置的接近度,其中,吸引力是有利于更接近吸引力位置的相对更高的扩音器激活的因素;扩音器与排斥力位置的接近度,其中,排斥力是有利于更接近排斥力位置的相对更低的扩音器激活的因素;每个扩音器相对于该环境中的其他扩音器的能力;该扩音器关于其他扩音器的同步;唤醒词性能;或回声消除器性能。

[0188] 在该示例中,框295涉及经由接口系统向环境的该组扩音器中的至少一些扩音器提供经渲染的音频信号。

[0189] 根据一些示例,感知空间位置的模型可以产生与收听者的左耳和右耳处的音频对象位置相对应的双耳响应。替代性地或另外地,感知空间位置的模型可以将从一组扩音器播放的音频信号的感知空间位置放置在由扩音器的相关联的激活增益加权的该组扩音器的位置的质心。

[0190] 在一些示例中,一个或多个附加动态可配置功能可以至少部分地基于一个或多个音频信号的水平。在一些实例中,一个或多个附加动态可配置功能可以至少部分地基于一个或多个音频信号的频谱。

[0191] 方法280的一些示例涉及接收扩音器布局信息。在一些示例中,一个或多个附加动态可配置功能可以至少部分地基于每个扩音器在环境中的位置。

[0192] 方法280的一些示例涉及接收扩音器规格信息。在一些示例中,一个或多个附加动态可配置功能可以至少部分地基于每个扩音器的能力,该能力可以包括以下各项中的一项或多项:频率响应、回放水平限制或一个或多个扩音器动态处理算法的参数。

[0193] 根据一些示例,一个或多个附加动态可配置功能可以至少部分地基于从每个扩音器到其他扩音器的声学传输的测量或估计。替代性地或另外地,一个或多个附加动态可配

置功能可以至少部分地基于环境中一个或多个人的收听者或扬声器位置。替代性地或另外地,一个或多个附加动态可配置功能可以至少部分地基于从每个扩音器到收听者或扬声器位置的声学传输的测量或估计。声学传输的估计可以例如至少部分地基于墙、家具或可以驻留在每个扩音器与收听者或扬声器位置之间的其他物体。

[0194] 替代性地或另外地,一个或多个附加动态可配置功能可以至少部分地基于环境中一个或多个非扩音器对象或地标的对象位置。在一些这样的实施方式中,一个或多个附加动态可配置功能可以至少部分地基于从每个扩音器到对象位置或地标位置的声学传输的测量或估计。

[0195] 可以通过采用一个或多个适当定义的附加成本项来实施灵活渲染以实现许多新的和有用的行为。下文列出的所有示例行为都是为了在某些被认为不期望的条件下惩罚某些扩音器。最终结果是这些扩音器在一组音频信号的空间渲染中被较少激活。在许多这些情况下,人们可能会考虑简单地调低不期望的扩音器,而与对空间渲染进行任何修正无关,但这种策略可能会显著降低音频内容的整体平衡。例如,混音的某些分量可能变得完全听不见。另一方面,对于所公开的实施例,将这些惩罚集成到渲染的核心优化中允许渲染适应并使用剩余较少惩罚的扬声器来执行最佳可能的空间渲染。这是更优雅、更加能适应的、更有效的解决方案。

[0196] 示例用例包括但不限于:

[0197] ●在收听区域周围提供更平衡的空间呈现

[0198] ○已经发现,空间音频最好跨与预期收听区域距离大致相同的扩音器来呈现。可以构建成本,使得明显比扩音器到收听区域的平均距离更近或更远的扩音器被惩罚,因此减少该扩音器的激活;

[0199] ●将音频移动离开或朝向收听者或说话者

[0200] ○如果系统的用户正在试图与系统的或与系统相关联的智能语音助理说话,则创建惩罚更接近说话者的扩音器的成本可能是有益的。通过该方式,这些扩音器更少激活,允许该扩音器相关联的麦克风更好地听到说话者;

[0201] ○为了为单个收听者提供更亲密的体验,最小化收听空间中针对其他人的回放水平,远离收听者的位置的扬声器可能会受到严重惩罚,以便仅最接近收听者的扬声器被最显著地激活;

[0202] ●将音频移动离开或朝向地标、区或区域

[0203] ○收听空间附近的某些位置可以被认为是敏感的,如婴儿房、婴儿床、办公室、阅读区域、学习区域等。在这种情况下,可以构建惩罚在该位置、区或区域附近使用扬声器的成本;

[0204] ○替代性地,对于上文的相同情况(或类似情况),扬声器系统可以已经生成从每个扬声器到婴儿房中的声学传输的测量,特别是当扬声器(带有附接或关联的麦克风)之一驻留在该婴儿房内时。在这种情况下,可以构建惩罚使用测量的到房间的声学传输高的扬声器的成本,而不是使用扬声器与婴儿房的物理接近度;和/或

[0205] ●扬声器的能力的最佳使用

[0206] ○不同扩音器的能力可以显著变化。例如,一种流行的智能扬声器仅包含具有有限低频能力的单个1.6”全音域驱动器。另一方面,另一个智能扬声器包含更有能力的3”低

音扬声器。这些能力通常反映在扬声器的频率响应中,并且如此,可以在成本项中利用与扬声器相关联的一组响应。在特定频率下,相对于其他扬声器能力较弱的扬声器(如根据其频率响应测得的)被惩罚并且因此被更少程度地激活。在一些实施方式中,这种频率响应值可以用智能扩音器存储,并且然后报告给负责优化灵活渲染的计算单元;

[0207] ○许多扬声器包含超过一个驱动器,每个驱动器负责播放不同的频率范围。例如,一种流行的智能扬声器是双向设计,包含用于较低频率的低音扬声器和用于较高频率的高音扬声器。通常,这种扬声器包含用于将全音域回放音频信号划分为适当的频率范围并且发送到相应的驱动器的分频电路。替代性地,这种扬声器可以为每个单独驱动器提供灵活渲染器回放访问,以及关于每个单独驱动器的能力(如频率响应)的信息。通过应用如上文描述的成本项,在一些示例中,灵活渲染器可以基于其不同频率下的相对能力在两个驱动器之间自动建立分频器;

[0208] ○上文描述的频率响应的示例使用专注于扬声器的固有能力上,但可能没有准确地反映放置在收听环境中的扬声器的能力。在某些情况下,如在预期收听位置测得的扬声器频率响应可以通过一些校准程序获得。可以使用这种测量替代预先计算的响应以更好地优化扬声器的使用。例如,某个扬声器在特定频率下可以天生就非常有能力,但由于其放置(例如,在墙或一件家具后面)可能会在预期的收听位置产生非常有限的响应。捕获该响应并且馈送到适当成本项中的测量可以防止这种扬声器的显著激活;

[0209] ○频率响应仅为扩音器的回放能力的一个方面。许多较小的扩音器开始失真,并且然后随着回放水平的增加而达到其偏移限制,特别是对于较低频率。为了减少这种失真,许多扩音器实施动态处理,该动态处理将回放水平限制在可以随频率变化的某些限制阈值以下。在扬声器靠近或处于这些阈值而其他参与灵活渲染的扬声器没有靠近或处于这些阈值的情况下,在限制扬声器中降低信号水平并且将该能量转移到其他负担较轻的扬声器是有意义的。根据一些实施例,可以通过适当地配置相关联的成本项来自动实现这种行为。这种成本项可以涉及以下各项中的一项或多项:

[0210] ■监测与扩音器的限制阈值相关的全局回放音量。例如,音量水平接近其限制阈值的扩音器可能会受到更多惩罚;

[0211] ■监测与扩音器限制阈值(可能随频率变化)相关的动态信号水平(也可能随频率变化)。例如,监测的信号水平接近其限制阈值的扩音器可能会受到更多惩罚;

[0212] ■直接监测扩音器的动态处理的参数,如限制增益。在一些这样的示例中,参数指示更多限制的扩音器可能会受到更多惩罚;和/或

[0213] ■监测放大器向扩音器传递的实际瞬时电压、电流和功率,以确定扩音器是否在线性范围内操作。例如,更少线性地操作的扩音器可能会受到更多惩罚;

[0214] ○带有集成式麦克风和交互式语音助理的智能扬声器通常采用某种类型的回声消除来降低由录音麦克风拾取的扬声器播放的音频信号水平。该降低越大,扬声器听到和理解空间中说话者的机会就越大。如果回声消除器的残差始终很高,这可以指示扬声器正在被驱动到非线性区域中,在该非线性区域中回声路径的预测变得具有挑战性。在这种情况下,将信号能量从扬声器转移出去可以是有意义的,并且如此,考虑回声消除器性能的成本项可以是有意义的。这样的成本项可以将高成本分配给其相关联的回声消除器性能不佳的扬声器;

[0215] ○为了在多个扩音器上渲染空间音频时实现可预测的成像,通常需要跨越时间合理地同步在一组扩音器上的回放。对于有线扩音器来说,这是给定的,但对于大量无线扩音器来说,同步可能具有挑战性并且最终结果可变。在这种情况下,每个扩音器报告其与目标同步的相对程度可以是可能的,并且然后该程度可以馈送到同步成本项中。在一些这样的示例中,具有较低同步程度的扩音器可能会受到更多惩罚,并且因此被排除在渲染之外。另外地,某些类型的音频信号可能不需要紧密同步,例如,旨在扩散或非定向的音频混音的分量。在一些实施方式中,可以如此用元数据对分量进行标记,并且可以修正同步成本项,使得惩罚减少。

[0216] 接下来描述实施例的示例。

[0217] 与等式 9 a 和 9 b 中定义的接近度成本类似,将每个新的成本函数项 $C_j(\mathbf{g}, \{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j)$ 表达为扬声器激活的绝对值平方的加权和也是方便的:

$$[0218] \quad C_j(\mathbf{g}, \{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j) = \mathbf{g}^* \mathbf{W}_j (\{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j) \mathbf{g} \quad (13a)$$

[0219] 其中, $\mathbf{W}_j$ 表示权重 $w_{ij} = w_{ij}(\{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j)$ 的对角矩阵,描述与项 j 的激活扬声器 i 相关联的成本:

$$[0220] \quad \mathbf{W}_j = \begin{bmatrix} w_{1j} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_{Mj} \end{bmatrix} \quad (13b)$$

[0221] 将等式 13a 和 13b 与等式 10 中给出的 CMAP 和 FV 成本函数的矩阵二次版本相结合,产生等式 12 中给出的 (一些实施例的) 通用扩展成本函数的潜在有益实施方式:

$$[0222] \quad C(\mathbf{g}) = \mathbf{g}^* \mathbf{A} \mathbf{g} + \mathbf{B} \mathbf{g} + \mathbf{C} + \mathbf{g}^* \mathbf{D} \mathbf{g} + \sum_j \mathbf{g}^* \mathbf{W}_j \mathbf{g} = \mathbf{g}^* (\mathbf{A} + \mathbf{D} + \sum_j \mathbf{W}_j) \mathbf{g} + \mathbf{B} \mathbf{g} + \mathbf{C} \quad (14)$$

[0223] 有了新成本函数项的该定义,总体成本函数仍然是矩阵二次的,并且可以通过等式 14 的微分找到最佳组激活 $\mathbf{g}_{opt}$ 以产生

$$[0224] \quad \mathbf{g}_{opt} = \frac{1}{2} (\mathbf{A} + \mathbf{D} + \sum_j \mathbf{W}_j)^{-1} \mathbf{B} \quad (15)$$

[0225] 将权重项 w_{ij} 中的每一个视为扩音器中的每一个的给定连续惩罚值 $p_{ij} = p_{ij}(\{\{\hat{o}\}, \{\hat{s}_i\}, \{\hat{e}\}\}_j)$ 的函数是有用的。在一个示例实施例中,该惩罚值是从 (将被渲染的) 对象到所考虑的扩音器的距离。在另一个示例实施例中,该惩罚值表示给定扩音器无法再现一些频率。基于该惩罚值,权重项 w_{ij} 可以参数化为:

$$[0226] \quad w_{ij} = \alpha_j f_j \left(\frac{p_{ij}}{\tau_j} \right) \quad (16)$$

[0227] 其中, α_j 表示前置因子 (其考虑了权重项的全局强度), 其中, τ_j 表示惩罚阈值 (约为或超过该惩罚阈值, 权重项变得显著), 并且其中, $f_j(x)$ 表示单调增加函数。例如, 有了 $f_j(x) = x^{\beta_j}$, 权重项具有以下形式:

$$[0228] \quad w_{ij} = \alpha_j \left(\frac{p_{ij}}{\tau_j} \right)^{\beta_j} \quad (17)$$

[0229] 其中, α_j 、 β_j 、 τ_j 是可调参数, 其分别指示惩罚的全局强度、惩罚起始的突然性和惩罚的程度。在设置这些可调值时应小心谨慎, 使得成本项 C_j 相对于任何其他附加成本项以及 C_{spatial} 和 $C_{\text{proximity}}$ 的相对影响适合用于实现期望的结果。例如, 根据经验, 如果一个人希望特定惩罚明显地支配其他惩罚, 则将其强度 α_j 设置为下一个最大惩罚强度的大约十倍可能是适当的。

[0230] 如果所有扩音器都被惩罚, 则通常可以方便地在后处理中从所有权重项中减去最小惩罚, 使得扬声器中的至少一个不会被惩罚:

$$[0231] \quad w_{ij} \rightarrow w'_{ij} = w_{ij} - \min_i (w_{ij}) \quad (18)$$

[0232] 如上所述, 使用本文所描述的新成本函数项 (以及根据其他实施例采用的类似的新成本函数项) 可以实现许多可能的用例。接下来, 用以下三个示例来描述更具体的细节: 将音频移动朝向收听者或说话者、将音频移动离开收听者或说话者以及将音频移动离开地标。

[0233] 在第一个示例中, 在本文中将被称为“吸引力”的东西用于将音频拉动朝向某个位置, 在一些示例中, 该位置可以是收听者或说话者的位置、地标位置、家具位置等。该位置在本文中可以被称为“吸引力位置”或“吸引子位置”。如本文所使用的, “吸引力”是有利于更接近吸引力位置的相对更高的扩音器激活的因素。根据该示例, 权重 w_{ij} 采用等式17的形式, 连续惩罚值 p_{ij} 由第 i 个扬声器距固定吸引子位置 $\vec{l}_j$ 的距离给出, 并且阈值 τ_j 由所有扬声器的这些距离中的最大值给出:

$$[0234] \quad p_{ij} = \|\vec{l}_j - \vec{s}_i\|, \text{ 并且} \quad (19a)$$

$$[0235] \quad \tau_j = \max_i \|\vec{l}_j - \vec{s}_i\| \quad (19b)$$

[0236] 为了说明将音频“拉动”朝向收听者或说话者的用例, 特别地设置 $\alpha_j = 20$, $\beta_j = 3$, 并且将 $\vec{l}_j$ 设置为与180度的收听者/说话者位置相对应的向量。 α_j 、 β_j 和 $\vec{l}_j$ 的这些值仅为示例。在一些实施方式中, α_j 可以在1至100的范围内并且 β_j 可以在1至25的范围内。

[0237] 图2F是示例实施例中的扬声器激活的图。在该示例中, 图2F示出了扬声器激活245b、250b、255b、260b和265b, 其包括图1和图2中相同扬声器位置的成本函数的最优解, 加上由 w_{ij} 表示的吸引力。图2G是示例实施例中的对象渲染位置的图。在该示例中, 图2G示出了针对大量可能对象角度的对应理想对象位置276b和针对那些对象的对应实际渲染位置278b, 通过虚线279b连接到理想对象位置276b。实际渲染位置278b朝向固定位置 $\vec{l}_j$ 的倾斜取向说明了吸引子权重对成本函数的最优解的影响。

[0238] 在第二个和第三个示例中, “推斥力”用于将音频“推动”离开某个位置, 该位置可以是收听者位置、说话者位置或其他位置, 如地标位置、家具位置等。在一些示例中, 推斥力可以用于将音频推动离开收听环境的区域或区, 如办公区域、阅读区域、床或卧室区域 (例如, 婴儿床或卧室) 等。根据一些这样的示例, 特定位置可以用作区或区域的代表。例如, 表示婴儿床的位置可以是婴儿的头部的估计位置、与婴儿相对应的估计声源位置等。该位置

在本文中可以被称为“推斥力位置”或“推斥位置”。如本文所使用的,“推斥力”是更有利于更接近推斥力位置的相对更低的扩音器激活的因素。根据该示例,相对于固定推斥位置 $\vec{l}_j$ 定义 p_{ij} 和 τ_j ,类似于等式19中的吸引力:

$$[0239] \quad p_{ij} = \max_i \|\vec{l}_j - \vec{s}_i\| - \|\vec{l}_j - \vec{s}_i\|, \text{ 并且} \quad (19c)$$

$$[0240] \quad \tau_j = \max_i \|\vec{l}_j - \vec{s}_i\| \quad (19d)$$

[0241] 为了说明将音频“推动”离开收听者或说话者的用例,特别地设置 $\alpha_j=5, \beta_j=2$,并且将 $\vec{l}_j$ 设置为与180度的收听者/说话者位置相对应的向量。 α_j 、 β_j 和 $\vec{l}_j$ 的这些值仅为示例。如上所述,在一些示例中, α_j 可以在1至100的范围内并且 β_j 可以在1至25的范围内。图2H是示例实施例中的扬声器激活的图。根据该示例,图2H示出了扬声器激活245c、250c、255c、260c和265c,其包括如先前图的相同扬声器位置的成本函数的最优解,加上由 w_{ij} 表示的推斥力。图2I是示例实施例中的对象渲染位置的图。在该示例中,图2I示出了针对大量可能对象角度的理想对象位置276c和针对那些对象的对应实际渲染位置278c,该对应实际渲染位置通过虚线279c连接到理想对象位置276c。实际渲染位置278c离开固定位置 $\vec{l}_j$ 的倾斜取向说明了推斥子权重对成本函数的最优解的影响。

[0242] 第三个示例用例是将音频“推动”离开声学上敏感的地标,如通向睡着的婴儿的房间的门。与最后的示例类似,将 $\vec{l}_j$ 设置为与180度的门位置(绘图的底部中心)相对应的向量。为了实现更强的推斥力并且将声场完全倾斜到主要收听空间的前部,设置 $\alpha_j=20, \beta_j=5$ 。图2J是示例实施例中的扬声器激活的图。再次,在该示例中,图2J示出了扬声器激活245d、250d、255d、260d和265d,其包括对同一组扬声器位置的最优解,加上更强的推斥力。图2K是示例实施例中的对象渲染位置的图。并且再次,在该示例中,图2K示出了针对大量可能对象角度的理想对象位置276d和针对那些对象的对应实际渲染位置278d,该对应实际渲染位置通过虚线279d连接到理想对象位置276d。实际渲染位置278d的倾斜取向说明了更强的推斥子权重对成本函数的最优解的影响。

[0243] 现在返回图2B,在该示例中,框225涉及至少部分地基于第二音频信号、第二经渲染的音频信号或其特性中的至少一项来修正针对第一音频信号的渲染过程,以产生经修正的第一经渲染的音频信号。本文中公开了修正渲染过程的各种示例。例如,经渲染的信号的“特性”可以包括在预期收听位置处估计或测量的响度或可听度,无论是在静默中或在存在一个或多个附加的经渲染的信号的情况下。特性的其他示例包括如相关联的节目流的组成信号的预期空间位置等与该信号的渲染相关联的参数、信号渲染于其上的扩音器的位置、根据组成信号的预期空间位置的扩音器的相对激活、以及与用于生成该经渲染的信号的渲染算法相关联的任何其他参数或状态。在一些示例中,框225可以由第一渲染模块执行。

[0244] 根据该示例,框230涉及至少部分地基于第一音频信号、第一经渲染的音频信号或其特性中的至少一项来修正针对第二音频信号的渲染过程,以产生经修正的第二经渲染的音频信号。在一些示例中,框230可以由第二渲染模块执行。

[0245] 在一些实施方式中,修正针对第一音频信号的渲染过程可以涉及将第一音频信号的渲染扭曲远离第二经渲染的音频信号的渲染位置和/或响应于第二音频信号或第二经渲染

染的音频信号中的一个或多个的响度而修正第一经渲染的音频信号中的一个或多个的响度。替代性地或另外地,修正针对第二音频信号的渲染过程可以涉及将第二音频信号的渲染扭曲远离第一经渲染的音频信号的渲染位置和/或响应于第一音频信号或第一经渲染的音频信号中的一个或多个的响度而修正第二经渲染的音频信号中的一个或多个的响度。下文参考图3以及以下提供了一些示例。

[0246] 然而,其他类型的渲染过程修正在本公开的范围。例如,在一些实例中,修正第一音频信号或第二音频信号的渲染过程可以涉及执行频谱修正、基于可听度的修正或动态范围修正。这些修正可以与基于响度的渲染修正有关或者可以与其无关,取决于特定示例。例如,在所述主要空间流在开敞布置生活区域中渲染并且包括烹饪提示的次要流在相邻厨房中渲染的情况下,可能期望确保烹饪提示在厨房中仍然是可听的。这可以通过在没有干扰性第一信号的情况下估计厨房中经渲染的烹饪提示流的响度、然后估计厨房中存在的第二信号的响度、并且最后跨多个频率动态地修正两个流的响度和动态范围来完成,从而确保第二信号在厨房中的可听度。

[0247] 在图2B中示出的示例中,框235涉及混合至少经修正的第一经渲染的音频信号和经修正的第二经渲染的音频信号以产生经混合的音频信号。例如,框235可以由图2A中示出的混音器130b执行。

[0248] 根据该示例,框240涉及将经混合的音频信号提供给环境的至少一些扬声器。方法200的一些示例涉及由扬声器回放经混合的音频信号。

[0249] 如图2B中示出的,一些实施方式可以提供多于2个渲染模块。一些这样的实施方式可以提供N个渲染模块,其中,N是大于2的整数。因此,一些这样的实施方式可以包括一个或多个附加渲染模块。在一些这样的示例中,一个或多个附加渲染模块中的每一个可以被配置用于经由接口系统接收附加音频节目流。附加音频节目流可以包括被安排由环境的至少一个扬声器再现的附加音频信号。一些这样的实施方式可以涉及渲染附加音频信号以经由环境的至少一个扬声器再现从而产生附加的经渲染的音频信号,以及至少部分地基于第一音频信号、第一经渲染的音频信号、第二音频信号、第二经渲染的音频信号或其特性中的至少一项来修正附加音频信号的渲染过程以产生经修正的附加的经渲染的音频信号。根据一些这样的示例,混合模块可以被配置用于将经修正的附加的经渲染的音频信号与至少经修正的第一经渲染的音频信号和经修正的第二经渲染的音频信号混合,以产生经混合的音频信号。

[0250] 如上文参考图1A和图2A所描述的,一些实施方式可以包括包含收听环境中的一个或多个麦克风的麦克风系统。在一些这样的示例中,第一渲染模块可以被配置用于至少部分地基于来自麦克风系统的第一麦克风信号来修正第一音频信号的渲染过程。“第一麦克风信号”可以从单个麦克风或从2个或更多个麦克风接收,取决于特定实施方式。在一些这样的实施方式中,第二渲染模块可以被配置用于至少部分地基于第一麦克风信号来修正第二音频信号的渲染过程。

[0251] 如上文参考图2A所述,在一些实例中,一个或多个麦克风的位置可以是已知的并且可以提供给控制系统。根据一些这样的实施方式,控制系统可以被配置用于基于第一麦克风信号估计第一声源位置以及至少部分地基于该第一声源位置来修正第一音频信号或第二音频信号中的至少一项的渲染过程。例如,可以基于来自具有已知位置的三个或更多

个麦克风或麦克风组中的每一个麦克风的DOA数据,根据三角测量过程来估计第一声源位置。替代性地或另外地,可以根据从两个或更多个麦克风接收的信号的振幅来估计第一声源位置。可以假定产生最高振幅信号的麦克风离第一声源位置最近。在一些这样的示例中,第一声源位置可以设置为最近的麦克风的位置。在一些这样的示例中,第一声源位置可以与区的位置相关联,其中,区是由通过预训练的分类器(如高斯混音器模型)处理来自两个或更多个麦克风的信号来选择的。

[0252] 在一些这样的实施方式中,控制系统可以被配置用于确定第一麦克风信号是否与环境噪声相对应。一些这样的实施方式可以涉及至少部分地基于第一麦克风信号是否与环境噪声相对应来修正第一音频信号或第二音频信号中的至少一项的渲染过程。例如,如果控制系统确定第一麦克风信号与环境噪声相对应,则修正第一音频信号或第二音频信号的渲染过程可以涉及增加经渲染的音频信号的水平,使得在预期收听位置处存在噪声时信号的感知响度基本上等于不存在噪声时信号的感知响度。

[0253] 在一些示例中,控制系统可以被配置用于确定第一麦克风信号是否与人类语音相对应。一些这样的实施方式可以涉及至少部分地基于第一麦克风信号是否与人类语音相对应来修正第一音频信号或第二音频信号中的至少一项的渲染过程。例如,如果控制系统确定第一麦克风信号与人类语音(如唤醒词)相对应,则修正第一音频信号或第二音频信号的渲染过程可以涉及与由远离第一声源位置的扬声器再现的经渲染的音频信号的响度相比,减小由在第一声源位置附近的扬声器再现的经渲染的音频信号的响度。修正第一音频信号或第二音频信号的渲染过程可以替代性地或另外涉及修正渲染过程以将相关联的节目流的组成信号的预期位置扭曲远离第一声源位置和/或与远离第一声源位置的扬声器相比,对在第一声源位置附近的扬声器的使用进行惩罚。

[0254] 在一些实施方式中,如果控制系统确定第一麦克风信号与人类语音相对应,则控制系统可以被配置用于在与第一声源位置不同的环境的位置附近的一个或多个扬声器中再现第一麦克风信号。在一些这样的示例中,控制系统可以被配置用于确定第一麦克风信号是否与儿童的哭声相对应。根据一些这样的实施方式,控制系统可以被配置用于在与看护者(如父母、亲戚、监护人、儿童看护服务提供者、老师、护士等)的估计位置相对应的环境位置附近的一个或多个扬声器中再现第一麦克风信号。在一些示例中,估计看护者的估计位置的过程可以由如“<唤醒词>,不要吵醒婴儿”等语音命令触发。控制系统将能够根据实施虚拟助理的最近智能音频设备的位置、通过基于由三个或更多个本地麦克风提供的DOA信息的三角测量等来估计说话者(看护者)的位置。根据一些实施方式,控制系统将具有婴儿房位置(和/或其中的收听设备)的先验知识,然后能够执行适当的处理。

[0255] 根据一些这样的示例,控制系统可以被配置用于确定第一麦克风信号是否与命令相对应。在一些实例中,如果控制系统确定第一麦克风信号与命令相对应,则控制系统可以被配置用于确定对命令的回复并且控制在第一声源位置附近的至少一个扬声器来再现回复。在一些这样的示例中,控制系统可以被配置用于在控制在第一声源位置附近的至少一个扬声器再现回复之后,恢复到第一音频信号或第二音频信号的未修正的渲染过程。

[0256] 在一些实施方式中,控制系统可以被配置用于执行命令。例如,控制系统可以是或者可以包括被配置成根据命令控制音频设备、电视、家用电器等的虚拟助理。

[0257] 通过图1A、图1B和图2A中示出的最小且更有能力的多流渲染系统的该定义,可以

针对许多有用的场景实现对多个节目流的同时回放的动态管理。现在将参考图3A和图3B描述若干个示例。

[0258] 首先检查先前讨论的涉及同时在客厅中回放空间电影声轨和在连接的厨房中回放烹饪提示的示例。空间电影声轨是上文提及的“第一音频节目流”的示例，并且烹饪提示音频是上文提及的“第二音频节目流”的示例。图3A和图3B示出了连接的生活空间的平面图示例。在该示例中，生活空间300包括在左上方处的客厅、在下方中央处的厨房和在右下方处的卧室。跨生活空间分布的方框和圆圈305a至305h表示放置在对于空间便利的位置的一组8个扩音器，但不遵守任何标准规定的布局（任意放置）。在图3A中，仅回放空间电影声轨，并且考虑到扩音器的能力和布局，客厅310和厨房315中的所有扩音器被用于围绕坐在沙发325上、面向电视330的收听者320a创建优化的空间再现。电影声轨的该最佳再现在视觉上由位于激活扩音器边界内的云状线335a表示。

[0259] 在图3B中，在厨房315中的单个扩音器305g上为第二收听者320b同时渲染和回放烹饪提示。该第二节目流的再现在视觉上由从扩音器305g传出的云状线340表示。如果在不修正电影声轨的渲染的情况下同时回放这些烹饪提示，如图3A所示，那么来自从厨房315中或附近的扬声器传出的电影声轨的音频会干扰第二收听者理解烹饪提示的能力。相反，在该示例中，根据烹饪提示的渲染动态地修正空间电影声轨的渲染。具体地，电影声轨的渲染从烹饪提示的渲染位置（厨房315）附近的扬声器移开，其中图3B中从厨房附近的扬声器被推开的较小云状线335b在视觉上表示了该移位。在一些实施方式中，如果在播放电影声轨时烹饪提示回放停止，那么电影声轨的渲染可以动态地移回在图3A中看到的其原始的最佳配置。可以通过许多所公开方法来实现空间电影声轨的渲染中的这种动态移位。

[0260] 许多空间音频混音包括被设计为在收听空间中的特定位置处播放的多个组成音频信号。例如，杜比5.1和7.1环绕声混音分别由6个和8个信号构成，意指在收听者周围的指定规范位置的扬声器上回放。基于对象的音频格式（例如，杜比全景声（Dolby Atmos））由具有相关联的元数据的组成音频信号构成，该元数据描述了将要渲染音频的收听空间中可能的时变3D位置。假设空间电影声轨的渲染器能够在关于任意扩音器组的任何位置渲染单独的音频信号，则可以通过扭曲空间混音内音频信号的预期位置来实现对图3A和图3B中描绘的渲染的动态移位。例如，与音频信号相关联的2D或3D坐标可以被推动远离厨房中扬声器的位置，或者替代性地被拉动朝向客厅的左上角。这种扭曲的结果是在厨房附近的扬声器较少被使用，因为空间混音的音频信号的经扭曲的位置现在离该位置更远了。虽然该方法确实实现了让第二音频流对于第二收听者来说更加易懂的目标，但这样做的代价是显著改变了针对第一收听者的电影声轨的预期空间平衡。

[0261] 用于实现向空间渲染的动态移位的第二种方法可以通过使用灵活渲染系统来实现。在一些这样的实施方式中，灵活渲染系统可以是CMAP、FV或两者的混合，如上文所描述的。一些这样的灵活渲染系统试图再现空间混音，该空间混音的所有组成信号被认为来自其预期位置。在一些示例中，在对混音的每个信号这样做的时候，优先考虑激活接近该信号的期望位置的扩音器。在一些实施方式中，附加项可以动态添加到渲染的优化中，这会基于其他标准惩罚某些扩音器的使用。例如，可以将可以被称为“排斥力”的东西动态地放置在厨房的位置，以高度惩罚在该位置附近使用扩音器并且有效地将空间电影声轨的渲染推开。如本文所使用的，术语“排斥力”可以指与在收听环境的特定位置或区中相对较低的扬

声器激活相对应的因素。换句话说,短语“排斥力”可以指有利于激活距与“排斥力”相对应的特定位置或区域相对较远的扬声器的因素。然而,根据一些这样的实施方式,渲染器可以仍试图使用剩余的、较少被惩罚的扬声器来再现混音的预期空间平衡。如此,与简单地扭曲混音的组成信号的预期位置的方法相比,该技术可以被认为是实现渲染的动态移位的优越的方法。

[0262] 可以使用图1B中描绘的多流渲染器的最小版本来实现所描述的将空间电影声轨的渲染从厨房中的烹饪提示移开的场景。然而,可以通过采用图2A中描绘的更有能力的系统来实现对该场景的改进。虽然将空间电影声轨的渲染移位确实改进了厨房中烹饪提示的可懂度,但电影声轨在厨房中可以仍然是明显可听的。根据两个流的瞬时条件,烹饪提示可能会被电影声轨掩蔽;例如,电影声轨中的响亮时刻掩蔽烹饪提示中的轻柔时刻。要处理该问题,可以根据空间电影声轨的渲染添加对烹饪提示的渲染的动态修正。例如,可以执行用于跨频率和时间动态地改变音频信号以便在存在干扰信号的情况下保持其感知响度的方法。在该场景中,可以生成对在厨房位置处经移位的电影声轨的感知响度的估计,并且将该估计作为干扰信号馈送到这样的过程中。然后可以动态地修正烹饪提示的时间和频率变化水平以维持其感知响度高于该干扰,由此更好地维持针对第二收听者的可懂度。厨房中电影声轨的响度的所需估计可以来自该声轨的渲染的扬声器馈送、来自在厨房中或厨房附近的麦克风的信号或其组合。维持烹饪提示的感知响度的该过程通常将提高烹饪提示的水平,并且在一些情况下,整体响度可能会变得高得令人反感。要解决该问题,可以采用又另一种渲染修正。可以根据厨房中经响度修正的烹饪提示变得太大声而将干扰空间电影声轨动态地调低。最后,一些外部噪声源可能同时干扰两个节目流的可听度;例如,在烹饪期间在厨房中可能使用搅拌机。在客厅和厨房两者中该环境噪声源的响度估计可以由连接到渲染系统的麦克风生成。例如,可以将该估计添加到对厨房中声轨的响度的估计中,以影响烹饪提示的响度修正。同时,可以根据环境噪声估计来另外地修正客厅中声轨的渲染,以在存在该环境噪声的情况下维持客厅中声轨的感知响度,由此更好地维持针对客厅中的收听者的可听度。

[0263] 如可以看到的,所公开的多流渲染器的该示例用例对两个节目流采用了大量相互关联的修正,从而优化其同时回放。总之,对流的这些修正可以列出如下:

[0264] ●空间电影声轨

[0265] ○根据在厨房中渲染的烹饪提示将空间渲染从厨房移开

[0266] ○根据在厨房中渲染的烹饪提示的响度来动态降低响度

[0267] ○根据对来自厨房的干扰性搅拌机噪声在客厅中的响度的估计来动态提高响度

[0268] ●烹饪提示

[0269] ○根据对电影声轨和厨房中搅拌机噪声两者的响度的组合估计来动态提高响度

[0270] 所公开的多流渲染器的第二个示例用例涉及同时发生的空间节目流(如音乐)的回放以及智能语音助理对用户的一些询问的响应。对于回放通常被约束为单个设备上的单声道或立体声回放的现有智能扬声器,与语音助理的交互通常由以下阶段构成:

[0271] 1) 播放音乐

[0272] 2) 用户说出语音助理唤醒词

[0273] 3) 智能扬声器识别该唤醒词并且显著调低(躲避)该音乐

[0274] 4) 用户向智能助理说出命令(即“播放下一首歌曲”)

[0275] 5) 智能扬声器识别该命令,通过借由扬声器播放在躲避的音乐之上混合的某个语音响应(即“好的,播放下一首歌曲”)来确认该命令,并且然后执行该命令

[0276] 6) 智能扬声器将音乐上调回到原始音量

[0277] 图4A和图4B示出了提供空间音乐混音和语音助理响应的同时回放的多流渲染器的示例。当在多个经编排的智能扬声器上播放空间音频时,一些实施例提供了对上述事件链的改进。具体地,可以将空间混音从被选择为适合于中继来自语音助理的响应的一个或多个扬声器移开。为语音助理响应创建该空间意指与上文列出的现有情况相比,空间混音可以被调低更少,或者可以根本不被调低。图4A和图4B描绘了该场景。在该示例中,经修正的事件链可以发生如下:

[0278] 1) 正在为用户在众多经编排的智能扬声器上播放空间音乐节目流(图4A中的云状线335c)。

[0279] 2) 用户320c说出语音助理唤醒词。

[0280] 3) 一个或多个智能扬声器(例如,扬声器305d和/或扬声器305f)识别该唤醒词,并且使用来自与一个或多个智能扬声器相关联的麦克风的相关联的记录来确定用户320c的位置或者用户320c最接近哪个(哪些)扬声器。

[0281] 4) 在预期语音助理响应节目流在该位置附近被渲染时,将空间音乐混音的渲染从在先前步骤中确定的位置移开(图4B中的云状线335d)。

[0282] 5) 用户对智能助理(例如,对运行智能助理/虚拟助理软件的智能扬声器)说出命令。

[0283] 6) 智能扬声器识别该命令,合成对应的响应节目流,并且在用户的位置附近渲染响应(图4B中的云状线440)。

[0284] 7) 当语音助理响应完成时,空间音乐节目流的渲染将切换回其原始状态(图4A中的云状线335c)。

[0285] 除了优化空间音乐混音和语音助理响应的同时回放,空间音乐混音的移位也可以改进步骤5中一组扬声器理解收听者的能力。这是因为音乐已经从在收听者附近的扬声器中移出,由此将语音提高到相关麦克风的其他比例。

[0286] 与针对具有空间电影混音和烹饪提示的先前场景所描述的场景类似,当前场景可以进一步优化,超出通过根据语音助理响应将空间混音的渲染移位而提供的优化。将空间混音移位本身可能不足以使语音助理响应对于用户来说是完全易懂的。简单的解决方案是也将空间混音降低固定的量,尽管少于当前状态所需的量。替代性地,可以根据空间音乐混音节目流的响度来动态提高语音助理响应节目流的响度,以维持响应的可听度。作为扩展,如果响应流上的该提高过程变得太大,则空间音乐混音的响度也可能被动态削减。

[0287] 图5A、图5B和图5C图示了所公开的多流渲染器的第三示例用例。该示例涉及管理空间音乐混音节目流和舒适噪声节目流的同时回放,同时试图确保婴儿在相邻的房间里保持睡眠,但能够听到婴儿是否哭泣。图5A描绘了起点,其中,空间音乐混音(由云状线335e表示)跨客厅310和厨房315中的所有扬声器为聚会上的许多人以最佳方式播放。在图5B中,婴儿510现在正在试图在右下方所绘制的相邻卧室505中睡觉。为帮助确保这一点,空间音乐混音从卧室动态地移开,以最小化在卧室中的泄漏,如云状线335f所描绘的,同时仍为聚会

上的人们维持合理的体验。同时,包含舒缓白噪声的第二节目流(由云状线540表示)从婴儿房的扬声器305h中播放,以掩蔽来自相邻房间中音乐的任何剩余泄漏。在一些示例中,要确保完全掩蔽,可以根据对泄漏到婴儿房中的空间音乐的响度的估计来动态地修正该白噪声流的响度。该估计可以从空间音乐的渲染的扬声器馈送、来自婴儿房的麦克风的信号或其组合中生成。同样,如果空间音乐混音的响度变得太大,则可以根据响度经修正的噪声使空间音乐混音的响度动态衰减。这类似于第一个场景的空间电影混音与烹饪提示之间的响度处理。最后,婴儿房里的麦克风(例如,与在一些实施方式中可以是智能扬声器的扬声器305h相关联的麦克风)可以被配置成记录来自婴儿的音频(消除可能从空间音乐和白噪声中拾取的声音),并且如果(通过机器学习、经由模式匹配算法等)检测到哭泣,则这些经处理的麦克风信号的组合然后可以充当可以在客厅310中的收听者320d附近同时回放的第三节目流,该收听者320d可以是父母或其他看护者。图5C用云状线550描绘了该附加流的再现。在这种情况下,空间音乐混音可以另外地从在父母附近的播放婴儿的哭声的扬声器移开,如相对于图5B的云状线335f的形状的云状线335g的经修正的形状所示出的,并且婴儿的哭声的节目流可以根据空间音乐流来修正响度,使得婴儿的哭声对于收听者320d保持可听。该示例中考虑的优化三个节目流的同时回放的互连修正可以总结如下:

[0288] ●客厅里的空间音乐混音

[0289] ○将空间渲染从婴儿房移开,以减少进入该房间的传播

[0290] ○根据在婴儿房中渲染的白噪声的响度来动态降低响度

[0291] ○根据婴儿的哭声在父母附近的扬声器上渲染,将空间渲染从父母移开

[0292] ●白噪声

[0293] ○根据对渗透进婴儿房中的音乐流的响度的估计来动态提高响度

[0294] ●记录婴儿的哭声

[0295] ○根据对在父母或其他看护者的位置处的音乐混音的响度的估计来动态提高响度。

[0296] 接下来描述可以如何实施一些提到的实施例的示例。

[0297] 在图1B中,渲染块1...N中的每一个都可以实施为如先前提到的CMAP、FV或混合渲染器等任何单流渲染器的相同实例。以该方式构造多流渲染器具有一些方便并且有用的性质。

[0298] 首先,如果渲染是在该层级布置中完成的,并且每个单流渲染器实例都被配置成在频域/变换域(例如,QMF)中操作,那么流的混合也可以发生在频域/变换域中,并且逆变换只需要运行一次,用于M个声道。与在时域中运行 $N \times M$ 逆变换和混合相比,这是显著的效率改进。

[0299] 图6示出了图1B中示出的多流渲染器的频域/变换域示例。在该示例中,在每个节目流被渲染模块1至N中的对应渲染模块接收之前,正交镜像分析滤波器组(QMF)被应用于节目流1至N中的每一个。根据该示例,渲染模块1至N在频域中操作。在混音器630a混合了渲染模块1至N的输出之后,逆合成滤波器组635a将混音转换到时域,并且将时域中的经混合的扬声器馈送信号提供给扩音器1至M。在该示例中,正交镜像滤波器组、渲染模块1至N、混音器630a和逆滤波器组635a是控制系统110c的部件。

[0300] 图7示出了图2A中示出的多流渲染器的频域/变换域示例。如在图6中,在每个节目

流被渲染模块1至N中的对应渲染模块接收之前,正交镜像滤波器组(QMF)被应用于节目流1至N中的每一个。根据该示例,渲染模块1至N在频域中操作。在该实施方式中,来自麦克风系统120b的时域麦克风信号也被提供给正交镜像滤波器组,使得渲染模块1至N接收频域中的麦克风信号。在混音器630b混合了渲染模块1至N的输出之后,逆滤波器组635b将混音变换到时域,并且将时域中的经混合的扬声器馈送信号提供给扩音器1至M。在该示例中,正交镜像滤波器组、渲染模块1至N、混音器630b和逆滤波器组635b是控制系统110d的部件。

[0301] 频域中层级方法的另一个益处在于计算每个音频流的感知响度,并且使用该信息动态地修正其他音频流中的一个或多个音频流。为了说明该实施例,考虑上文参考图3A和图3B描述的先前提到的示例。在这种情况下,有两个音频流($N=2$),空间电影声轨和烹饪提示。还可以有在厨房中的搅拌机产生的由K个麦克风中的一个或多个麦克风拾取的环境噪声。

[0302] 在每个音频流s被单独渲染并且每个麦克风i被捕获并且变换到频域之后,可以计算充当每个音频流s或麦克风信号i的感知响度的时变估计的源激励信号 E_s 或 E_i 。在该示例中,这些源激励信号是针对c个扩音器的跨时间t的b个频带、经由音频流的变换系数 X_s 或麦克风信号的变换系数 X_i 从经渲染的流或所捕获的麦克风来计算的,并且使用与频率相关的时间常数 λ_b 进行平滑:

$$[0303] \quad E_s(b, t, c) = \lambda_b E_s(b, t-1, c) + (1 - \lambda_b) |X_s(b, t, c)|^2 \quad (20a)$$

$$[0304] \quad E_i(b, t, c) = \lambda_b E_i(b, t-1, c) + (1 - \lambda_b) |X_i(b, t, c)|^2 \quad (20b)$$

[0305] 在该示例中,原始源激励是对每个流在特定位置处的感知响度的估计。对于空间流,该位置在图3B中的云状线335b中间,而对于烹饪提示流,该位置在云状线340中间。由麦克风拾取的搅拌机噪声的位置可以例如基于最靠近该搅拌机噪声的来源的(多个)麦克风的(多个)特定位置。

[0306] 必须将原始源激励变换到将被其修正的(多个)音频流的收听位置,以估计该原始源激励在每个目标音频流的收听位置处作为噪声的可感知程度。例如,如果音频流1是电影声轨并且音频流2是烹饪提示,则 $\hat{E}_{12}$ 将为经变换的(噪声)激励。针对每个频带b,根据每个扩音器c,通过将可听度比例系数从源音频流s应用到目标音频流x或将 A_{xi} 从麦克风i应用到目标音频流x来计算该变换。 A_{xs} 和 A_{xi} 的值可以通过使用距离比率或对实际可听度(可以随着时间而变化)的估计来确定。

$$[0307] \quad \hat{E}_{xs}(b, t, c) = A_{xs}(b, t, c) E_s(b, t, c) \quad (21a)$$

$$[0308] \quad \hat{E}_{xi}(b, t, c) = A_{xi}(b, t, c) E_i(b, t, c) \quad (21b)$$

[0309] 在等式13a中, $\hat{E}_{xs}$ 表示不参考麦克风输入针对源音频流计算的原始噪声激励。在等式13b中, $\hat{E}_{xi}$ 表示参考麦克风输入计算的原始噪声激励。根据该示例,然后跨流1到N、麦克风1到K和输出声道1到M对原始噪声激励 $\hat{E}_{xs}$ 或 $\hat{E}_{xi}$ 求和,以获得针对目标流x的总噪声估计 $\hat{E}_x$:

$$[0310] \quad \hat{E}_x(b, t) = \sum_{c=1}^M \left(\sum_{s=1}^N \hat{E}_{xs}(b, t, c) + \sum_{i=1}^K \hat{E}_{xi}(b, t, c) \right) \quad (22)$$

[0311] 根据一些替代性实施方式,通过省略等式14中的项 $\sum_{i=1}^K \hat{E}_{xi}(b, t, c)$,可以在不参考麦克风输入的情况下获得总噪声估计。

[0312] 在该示例中,对总原始噪声估计进行平滑以避免由过快地修正目标流引起的可感知的伪影。根据该实施方式,平滑基于使用快速启动和缓慢释放的概念,类似于音频压缩机。针对目标流x的经平滑的噪声估计 $\bar{E}_x$ 在该示例中计算为:

$$[0313] \quad \bar{E}_x(b, t) = \lambda(b, t) \hat{E}_x(b, t) + (1 - \lambda(b, t)) \bar{E}_x(b, t - 1) \quad (23)$$

$$[0314] \quad \lambda(b, t) = \begin{cases} \lambda_{fast}, & \hat{E}_x(b, t) > \bar{E}_x(b, t) \\ \lambda_{slow}, & \hat{E}_x(b, t) \leq \bar{E}_x(b, t) \end{cases} \quad (24)$$

[0315] 一旦有了针对流x的完整噪声估计 $\bar{E}_x(b, t)$,就可以重用先前计算的源激励信号 $E_x(b, t, c)$ 来确定时变增益集合 $G_x(b, t, c)$ 以应用于目标音频流x,从而确保该目标音频流在噪声中保持可听。可以使用各种技术中的任一种来计算这些增益。

[0316] 在一个实施例中,响度函数 $L\{\cdot, \cdot\}$ 可以应用于激励,以对人类的响度感知中的各种非线性进行建模,并且计算描述感知响度在频率上的时变分布的特定响度信号。将 $L\{\cdot, \cdot\}$ 应用于针对噪声估计和经渲染的音频流x的激励给出了针对每个信号的特定响度的估计:

$$[0317] \quad L_{xn}(b, t) = L\{\bar{E}_x(b, t)\} \quad (25a)$$

$$[0318] \quad L_x(b, t, c) = L\{E_x(b, t, c)\} \quad (25b)$$

[0319] 在等式17a中, L_{xn} 表示针对噪声的特定响度的估计,并且在等式17b中, L_x 表示针对经渲染的音频流x的特定响度的估计。这些特定响度信号表示当信号被单独听到时的感知响度。然而,如果两个信号混合,则可能出现掩蔽。例如,如果噪声信号比流x信号响亮得多,则该噪声信号将掩蔽流x信号,从而相对于单独听到的信号的感知响度降低该信号的感知响度。这种现象可以用采用两个输入的部分响度函数 $PL\{\cdot, \cdot\}$ 来建模。第一输入是感兴趣信号的激励,并且第二输入是竞争(噪声)信号的激励。该函数返回表示在存在竞争信号的情况下感兴趣信号的感知响度的部分特定响度信号 PL 。在存在噪声信号的情况下,流x信号的部分特定响度然后可以跨频带b、时间t和扩音器c直接从激励信号中计算:

$$[0320] \quad PL_x(b, t, c) = PL\{E_x(b, t, c), \bar{E}_x(b, t)\} \quad (26)$$

[0321] 为了在存在噪声的情况下维持音频流x信号的可听度,可以计算增益 $G_x(b, t, c)$ 以应用于音频流x,从而提高响度,直到该信号在噪声之上可听,如等式8a和8b中示出的。替代性地,如果噪声来自另一个音频流s,则可以计算两个增益集合。在一个这样的示例中,第一个增益集合 $G_x(b, t, c)$ 将应用于音频流x以提高其响度并且第二个增益集合 $G_s(b, t)$ 将应用于竞争音频流s以降低其响度,使得增益的组合确保音频流x的可听度,如等式9a和9b中示出的。在等式的两个集合中, $\overline{PL}_x(b, t, c)$ 表示在应用补偿增益后在存在噪声的情况下源信

号的部分特定响度。

$$[0322] \quad \overline{PL}_x(b, t, c) = PL\{G_x(b, t, c)^2 E_x(b, t, c), \bar{E}_x(b, t)\} \quad (27a)$$

[0323] 使得

$$[0324] \quad L_x(b, t, c) - \overline{PL}_x(b, t, c) < L_x(b, t, c) - PL_x(b, t, c) \quad (27b)$$

$$[0325] \quad \overline{PL}_x(b, t, c) = PL\{G_x(b, t, c)^2 E_x(b, t, c), G_s(b, t)^2 \bar{E}_x(b, t)\} \quad (28a)$$

[0326] 再次,使得

$$[0327] \quad L_x(b, t, c) - \overline{PL}_x(b, t, c) < L_x(b, t, c) - PL_x(b, t, c) \quad (28b)$$

[0328] 在一些示例中,在应用于音频流之前,可以跨频率使用平滑函数 $S\{\cdot\}$ 进一步平滑原始增益,以再次避免可听伪影。 $\bar{G}_x(b, t, c)$ 和 $\bar{G}_s(b, t)$ 表示目标音频流x和竞争音频流s的最终补偿增益:

$$[0329] \quad \bar{G}_x(b, t, c) = S\{G_x(b, t, c)\} \quad (29a)$$

$$[0330] \quad \bar{G}_s(b, t) = S\{G_s(b, t)\} \quad (29b)$$

[0331] 在一个实施例中,这些增益可以直接应用于音频流的所有经渲染的输出声道。在另一个实施例中,这些增益可以替代地在被渲染之前应用于音频流的对象,例如,使用美国专利申请公开号2019/0037333A1中描述的方法,该专利申请公开通过引用并入本文。这些方法涉及基于音频对象的空间元数据计算与多个预定义声道覆盖区中的每一个相关的每个音频对象的平移系数。可以基于所计算的平移系数和音频对象将音频信号转换为与预定义声道覆盖区相关的子混音。每个子混音可以指示多个音频对象的与预定义声道覆盖区之一相关的分量的总和。子混音增益可以通过对每个子混音应用音频处理来生成,并且可以控制应用到每个音频对象的对象增益。对象增益可以是针对每个音频对象的平移系数和与每个预定义声道覆盖区相关的子混音增益的函数。将增益应用于对象具有一些优势,特别是在与流的其他处理结合时。

[0332] 图8示出了具有音频流响度估计器的多流渲染系统的实施方式。根据该示例,图8的多流渲染系统还被配置用于实施如等式12a-21b中所描述的响度处理,以及每个单流渲染器内的补偿增益应用。在该示例中,在每个节目流被渲染模块1和2中的对应渲染模块接收之前,正交镜像滤波器组(QMF)被应用于节目流1和2中的每一个。在替代性示例中,在每个节目流被渲染模块1至N中的对应渲染模块接收之前,正交镜像滤波器组(QMF)可以被应用于节目流1至N中的每一个。根据该示例,渲染模块1和2在频域中操作。在该实施方式中,响度估计模块805a计算针对节目流1的响度估计,例如,如上文参考等式12a至17b所描述的。类似地,在该示例中,响度估计模块805b计算针对节目流2的响度估计。

[0333] 在该实施方式中,来自麦克风系统120c的时域麦克风信号也被提供给正交镜像滤波器组,使得响度估计模块805c接收频域中的麦克风信号。在该实施方式中,响度估计模块805c计算针对麦克风信号的响度估计,例如,如上文参考等式12b至17a所描述的。在该示例中,响度处理模块810被配置用于实施如等式18至21b中所描述的响度处理,以及针对每个单流渲染模块的补偿增益应用。在该实施方式中,响度处理模块810被配置用于改变节目流

1的音频信号和节目流2的音频信号,以便在存在一个或多个干扰信号的情况下保持该音频信号的感知响度。在一些实例中,控制系统可以确定麦克风信号与应该将节目流提高到其之上的环境噪声相对应。然而,在一些示例中,控制系统可以确定麦克风信号与唤醒词、命令、儿童的哭声或可能需要由智能音频设备和/或一个或多个收听者听到的其他这样的音频相对应。在一些这样的实施方式中,响度处理模块810可以被配置用于改变麦克风信号,以便在存在节目流1的干扰音频信号和/或节目流2的干扰音频信号的情况下保持该麦克风信号的感知响度。此处,响度处理模块810被配置成向渲染模块1和2提供适当的增益。

[0334] 在混音器630c混合了渲染模块1至N的输出之后,逆滤波器组635c将混音转换到时域,并且将时域中的经混合的扬声器馈送信号提供给扩音器1至M。在该示例中,正交镜像滤波器组、渲染模块1至N、混音器630c和逆滤波器组635c是控制系统110e的部件。

[0335] 图9A示出了被配置用于多个经渲染的流的交叉渐变的多流渲染系统的示例。在一些这样的实施例中,多个经渲染的流的交叉渐变用于在动态更改渲染配置时提供流畅的体验。一个示例是前述的同时回放空间节目流(如音乐)的用例,其中智能语音助理对收听者的一些询问做出响应,如上文参考图4A和图4B所描述的。在这种情况下,使用替代性空间渲染配置将额外的单流渲染器实例化并且同时在其之间交叉渐变是有用的,如图9A中示出的。

[0336] 在该示例中,在节目流被渲染模块1a和1b接收之前,QMF应用于节目流1。类似地,在节目流被渲染模块2a和2b接收之前,QMF应用于节目流2。在一些实例中,渲染模块1a的输出可以与在唤醒词检测之前对节目流1的期望再现相对应,而渲染模块1b的输出可以与在唤醒词检测之后对节目流1的期望再现相对应。类似地,渲染模块2a的输出可以与在唤醒词检测之前对节目流2的期望再现相对应,而渲染模块2b的输出可以与在唤醒词检测之后对节目流2的期望再现相对应。在该实施方式中,渲染模块1a和1b的输出被提供给交叉渐变模块910a,并且渲染模块2a和2b的输出被提供给交叉渐变模块910b。例如,交叉渐变时间可以在数百毫秒到若干秒的范围内。

[0337] 在混音器630d混合了交叉渐变模块910a和910b的输出之后,逆滤波器组635d将混音转换到时域,并且将时域中的经混合的扬声器馈送信号提供给扩音器1至M。在该示例中,正交镜像滤波器组、渲染模块、交叉渐变模块、混音器630d和逆滤波器组635d是控制系统110f的部件。

[0338] 在一些实施例中,可以预先计算在单流渲染器1a、1b、2a和2b中的每一个中使用的渲染配置。这对于像智能语音助理等用例特别方便和高效,因为空间配置通常是已知先验的,并且不依赖于系统的其他动态方面。在其他实施例中,预先计算渲染配置可以是不可能或不可取的,在这种情况下,必须在系统运行时动态地计算每个单流渲染器的完整配置。

[0339] (根据一些实施例)实施动态成本灵活渲染的实际考虑之一是复杂性。在一些情况下,考虑到对象位置(每个要渲染的音频对象的可以由元数据指示的位置)每秒可能变化多次,实时求解每个音频对象的每个频带的唯一成本函数可能不可行。以存储器为代价降低复杂性的替代性方法是使用查找表,该查找表对所有可能对象位置的三维空间进行采样。该采样不必在所有维度上都相同。图9B是示例实施例中指示扬声器激活的点的图。在该示例中,对x维度和y维度采样15个点,并且对z维度采样5个点。根据该示例,每个点表示M个扬声器激活,音频环境中的M个扬声器中的每一个有一个扬声器激活。例如,对于质心振幅平

移(CMAP)渲染,扬声器激活可以是每个扬声器单个增益,或者对于灵活虚拟化(FV)或混合CMAP/FV渲染,可以是跨多个N个频率的复值向量。其他实施方式可以包括更多样本或更少样本。例如,在一些实施方式中,扬声器激活的空间采样可能不一致。一些实施方式可以涉及比图9B中示出的更多或更少的x,y平面中的扬声器激活样本。一些这样的实施方式可能仅在一个x,y平面中确定扬声器激活样本。根据该示例,每个点表示针对CMAP或FV解的M个扬声器激活。在一些实施方式中,如图9B中示出的一组扬声器激活可以存储在数据结构中,该数据结构在本文中可以被称为“表”(或者“笛卡尔表”,如图9B所指示的)。

[0340] 期望渲染位置不一定与已经为其计算扬声器激活的位置相对应。在运行时,为了确定每个扬声器的实际激活,可以实施某种形式的插值。在一些这样的示例中,可以使用离期望渲染位置最近的8个点的扬声器激活之间的三线性插值。图10是根据一个示例的指示扬声器激活的点之间的三线性插值的图。根据该示例,在图10中示出的矩形棱柱的顶点处或附近的实心圆1003对应于离已经为其计算扬声器激活的期望渲染位置最近的8个点的位置。在该实例中,期望渲染位置是图10中呈现的矩形棱柱内的点。在该示例中,连续线性插值的过程包括对顶平面中的每对点进行插值以确定第一插值点1005a和第二插值点1005b、对底平面中的每对点进行插值以确定第三插值点1010a和第四插值点1010b、对第一插值点1005a和第二插值点1005b进行插值以确定顶平面中的第五插值点1015、对第三插值点1010a和第四插值点1010b进行插值以确定底平面中的第六插值点1020、以及对第五插值点1015和第六插值点1020进行插值以确定顶平面与底平面之间的第七插值点1025。尽管三线性插值是有效的插值方法,但本领域技术人员将理解,三线性插值只是可以用于实施本公开的各方面的一种可能的插值方法,并且其他示例可以包括其他插值方法。例如,一些实施方式可以涉及比图9B中示出的更多或更少的x,y平面中的插值。一些这样的实施方式可以涉及仅一个x,y平面中的插值。在一些实施方式中,期望渲染位置的扬声器激活将被简单地设置为离已经为其计算扬声器激活的期望渲染位置最近的位置的扬声器激活。

[0341] 如上文参考图9B所述,在一些示例中,每个数据结构或表可以与特定的渲染配置相对应。在一些示例中,每个数据结构或表可以与特定渲染配置的一个版本(例如,简化版本或完整版本)相对应。渲染配置的复杂性和计算对应扬声器激活所需的时间与表中点(扬声器激活)的数量相关,在三维示例中,该数量可以计算为5个维度的乘积:x、y和z点的数量、音频环境中扩音器的数量M以及所涉及的频带的数量N。在一些示例中,表大小与复杂性之间的这种相关性是由于需要最小化潜在的唯一成本函数来填充表中的每个点而产生的。在一些示例中,复杂性和计算扬声器激活表的时间也可以与渲染的保真度相关。

[0342] 图11A和图11B示出了提供渲染配置计算服务的示例。高质量渲染配置计算对于一些音频设备(如图11A和图11B中示出的智能音频设备1105a、1105b和1105c)可能太复杂。

[0343] 在图11A中示出的示例中,智能音频设备1105a-1105c被配置成向渲染配置计算服务1115a发送请求1110a。在该示例中,渲染配置计算服务1115a是基于云的服务,在一些实例中,该基于云的服务可以由数据中心的一个或多个服务器来实施。根据该示例,智能音频设备1105a-1105c被配置成经由互联网将请求1110a从音频环境1100a发送到渲染配置计算服务1115a。此处,渲染配置计算服务1115a响应于请求1110a经由互联网向智能音频设备1105a-1105c提供渲染配置1120a。

[0344] 根据图11B中示出的示例,智能音频设备1105a-1105c被配置成向渲染配置计算服

务1115b发送请求1110b。在该示例中,渲染配置计算服务1115b由音频环境1100b的另一个更有能力的设备实施。此处,音频设备1105d是比智能音频设备1105a-1105c相对更有能力的智能音频设备。在一些示例中,本地设备可以是本文中被称为编排设备或“智能家居中枢”的设备。根据该示例,智能音频设备1105a-1105c被配置成经由如Wi-Fi网络等本地网络向渲染配置计算服务1115b发送请求1110b。在该实施方式中,渲染配置计算服务1115b响应于请求1110b经由本地网络向智能音频设备1105a-1105c提供渲染配置1120b。

[0345] 在图11A和图11B中示出的示例中,网络通信给计算渲染配置的过程增加了附加的延迟。在动态的经编排的音频系统中,使渲染配置改变以最小的延迟生效是期望的性质。当渲染配置改变响应于实时事件(如响应于唤醒词话语而执行空间回避的渲染配置)时,尤其如此。对于高度动态的应用,延迟可以说比质量更重要,因为在对事件做出反应之前,系统中的任何感知滞后都会对用户体验产生负面影响。因此,在许多用例中,可能非常期望获得具有低延迟(理想情况下在几百毫秒的范围内)的有效渲染配置。一些公开的示例包括用于以低延迟实现渲染配置改变的方法。一些公开的示例实施了用于计算渲染配置的相对较低复杂性版本的方法,以满足延迟要求并且然后在渲染配置的较高质量版本可用时渐进地转变到该较高质量版本。

[0346] 动态的灵活渲染器的另一个期望的性质是不管系统的当前状态如何都能平滑地转变到新渲染配置的能力。例如,如果在系统处于渲染配置的较低复杂性版本与渲染配置的较高质量版本之间的转变的中间时请求第三配置改变,则一些公开的实施方式被配置成开始向新渲染配置的转变,而无需等待第一渲染配置转变完成。本公开包括用于平滑地处理这种转变的方法。

[0347] 所公开的示例通常适用于将一组扬声器激活(例如,一组M个扬声器激活)应用于音频信号集合以生成输出(例如,M个输出)的任何渲染器,无论渲染器是在时域还是频域中操作。通过确保当前一组扬声器激活与新的一组扬声器激活之间的渲染配置转变是平滑和连续的而且是可任意中断的,渲染配置之间转变的定时可以有效地与计算新渲染配置的扬声器激活所花费的时间解耦。一些这样的实施方式使得能够渐进地转变到新的、较低质量/复杂性的渲染配置,并且然后转变到对应的较高质量/复杂性的渲染配置,后者可以异步计算和/或由除了执行渲染的设备之外的设备计算。能够支持平滑、连续且可任意中断的转变的一些实施方式还具有不管可能正在进行的任何先前转变如何都允许一组新的目标渲染激活在任何时间动态地更新的期望性质。

[0348] 渲染配置转变

[0349] 本文公开了用于实施平滑、动态且可任意中断的渲染配置转变(本文中也可以被称为渲染器配置转变或简称为“转变”)的各种方法。图12A和图12B示出了渲染配置转变的示例。在图12A中示出的示例中,音频系统最初再现根据渲染配置1渲染的音频。例如,音频环境中的一个或多个控制系统可以根据渲染配置1将输入音频数据渲染为扩音器馈送信号。例如,渲染配置1可以对应于如上文参考图9B描述的数据结构(“表”)。

[0350] 根据该示例,在时间A时接收到渲染转变指示。在一些示例中,渲染转变指示可以由如智能扬声器或“智能家居中枢”等编排设备接收,该编排设备被配置成协调或编排音频环境中的音频设备的渲染。在该示例中,渲染转变指示是当前渲染配置应该转变到渲染配置2的指示。渲染转变指示可以例如对应于(例如,可以响应于)音频系统中检测到的事件,

如唤醒词话语、来电呼叫、第二内容流(如与上文参考图3B描述的“烹饪提示”示例相对应的内容流)将由音频环境的一个或多个音频设备渲染的指示等。渲染配置2可以例如对应于响应于唤醒词话语而执行空间回避的渲染配置、用于响应于用户输入而将音频“推”离音频环境位置的渲染配置等。根据该示例,在转变时间 t 期间,音频系统从渲染配置1转变到渲染配置2。在转变时间 t 之后,音频系统已经完全转变到渲染配置2。

[0351] 图12B示出了中断的渲染器配置转变的示例。在图12B中示出的示例中,音频系统最初再现根据渲染配置1渲染的音频。根据该示例,在时间C时接收到第一渲染转变指示。在该示例中,第一渲染转变指示是当前渲染配置应该转变到渲染配置2的指示。根据该示例,音频系统将需要转变时间 t_1 (从时间C到时间E)来从渲染配置1完全转变到渲染配置2。然而,在转变时间 t_1 过去之前并且在音频系统已经完全转变到渲染配置2之前,在该示例中,在时间D时接收到第二渲染转变指示。此处,第二渲染转变指示是当前渲染配置应该转变到渲染配置3的指示。根据该示例,在接收到第二渲染转变指示时中断从渲染配置1到渲染配置2的转变。在该示例中,音频系统需要转变时间 t_2 (从时间D到时间F)来完全转变到渲染配置3。

[0352] 本文公开了用于管理渲染配置之间的转变的各种机制。在上文参考图9A描述的示例中,多个经渲染的流的交叉渐变用于在动态改变渲染配置时提供平滑转变。

[0353] 图13呈现了与用于管理渲染配置之间的转变的替代性实施方式相对应的块。与其他公开的实施方式一样,图13中呈现的元件的数量、布置和类型仅作为示例。根据该示例,渲染器1315、QMF 1310和逆QMF 1335经由控制系统110g植入,该控制系统是上文参考图1A描述的控制系统110的实例。此处,控制系统110g被示出为接收节目流,在该实例中,该节目流包括节目流音频数据1305a和节目流元数据1305b。在该示例中,对于给定的节目流,仅实施了单个渲染器实例,与维持两个或更多个渲染器实例的实施方式相比,这是更高效的实施方式。

[0354] 根据该实施方式,渲染器1315维持激活数据结构1312a和1312b,在该示例中,该激活数据结构是查找表A和B。数据结构1312a和1312b中的每一个对应于渲染配置或渲染配置的一个版本(如简化版本或完整版本)。表A对应于当前渲染配置,而表B对应于目标渲染配置。例如,数据结构1312a和1312b可以对应于一组扬声器激活,如图9B中示出的和上文描述的一组扬声器激活。在图13中示出的示例中,目标渲染配置对应于已经由控制系统110g接收的渲染转变指示。换句话说,渲染转变指示是控制系统110g从当前渲染配置(对应于表A)转变到新渲染配置的指示。在一些实例中,表B可以对应于新渲染配置的简化版本。在其他示例中,表B可以对应于新渲染配置的完整版本。

[0355] 在该实施方式中,当渲染器1315正在渲染音频对象时,渲染器1315为音频对象的位置计算两组扬声器激活:此处,针对音频对象的位置的一组扬声器激活基于表A,而针对音频对象的位置的另一组扬声器激活基于表B。根据该示例,三线性插值模块1313a和1313b被配置成确定每个扬声器的实际激活。在一些这样的示例中,三线性插值模块1313a和1313b被配置成根据最近的8个点的扬声器激活之间的三线性插值过程来确定每个扬声器的实际激活,如上文参考图10描述的。其他实施方式可以使用其他类型的插值。在仍其他实施方式中,每个扬声器的实际激活可以根据多于或少于最近的8个点的扬声器激活来确定。

[0356] 在该示例中,控制系统110g的模块1314被配置成至少部分地基于交叉渐变时间

1311来确定两组扬声器激活之间的幅度归一化插值。根据该示例,控制系统110g的模块1314被配置成基于从三线性插值模块1313a接收的表A的插值以及基于从三线性插值模块1313b接收的表B的插值来确定单个扬声器激活表。在一些示例中,渲染转变指示还指示交叉渐变时间1311,其对应于上文参考图12A和图12B描述的转变时间。然而,在其他示例中,控制系统110g将例如通过访问存储的交叉渐变时间确定交叉渐变时间1311。在一些实例中,交叉渐变时间1311可以根据对包括控制系统110g的设备的用户输入来进行配置。

[0357] 在该实施方式中,控制系统110g的模块1316被配置成根据幅度归一化插值在频域中计算每个音频对象的最后一组扬声器激活。在该示例中,逆滤波器组1335将最后一组扬声器激活转换到时域,并将时域中的扬声器馈送信号提供给扩音器1至M。在替代性实施方式中,渲染器1315可以在时域中操作。

[0358] 图14呈现了与用于管理渲染配置之间的转变的另一实施方式相对应的块。与其他公开的实施方式一样,图14中呈现的元件的数量、布置和类型仅作为示例。在该示例中,图14包括可以如上所述起作用的图13的块。然而,在该示例中,控制系统110g还被配置成实施表组合模块1405。

[0359] 在一些实例中,例如,如上文参考图12B描述的,可以中断从第一渲染配置到第二渲染配置的转变。在一些这样的示例中,可以在系统正在从第一渲染配置转变到第二渲染配置的时间间隔期间接收第二渲染转变指示。第二渲染转变指示可以例如指示到第三渲染配置的转变。

[0360] 根据一些示例,表组合模块1405可以被配置成处理渲染配置转变的中断。在一些这样的示例中,如果渲染配置转变被另一渲染转变指示中断,则表组合模块1405可以被配置成组合上述数据结构1312a和1312b(先前的两个查找表A和B),以创建新的当前查找表A'(数据结构1412a)。在一些这样的实施方式中,控制系统110g被配置成用新的目标查找表B'(数据结构1412b)代替数据结构1312b。新的目标查找表B'可以例如与第三渲染配置相对应。

[0361] 在一些示例中,表组合模块1405的块1414可以被配置成通过在A表和B表的每个点处应用相同的幅度归一化插值机制来实施组合操作,然后实施插值过程:在该实施方式中,插值过程涉及分别由块1413a和1413b对表A和B的内容执行的单独三线性插值过程。插值过程可以至少部分地基于先前渲染配置转变被中断的时间(如图14的先前交叉渐变中断时间1411所指示的)。

[0362] 根据一些实施方式,新的当前查找表A'(数据结构1412a)与一个或多个先前的表(A和/或B)相比可以可选地具有减小的维度。A'的维度选择允许在转变到与新的目标查找表B'相关联的新渲染配置期间在一次性计算的复杂性与质量之间进行权衡。在目标表B被新的目标表B'代替之后,在一些实例中,表组合模块1405可以被配置成继续表之间的实时插值,直到渲染配置转变完成。在一些实施方式中,表组合模块1405(例如,结合其他公开的特征)可以被配置成不间断地处理多个渲染配置转变中断。

[0363] 渲染配置的渐进计算和应用是一种满足良好用户体验所需的低延迟目标而不牺牲灵活渲染配置(例如,对异构的一组任意放置的音频设备的回放进行优化的灵活渲染配置)的质量的方式。给定与特定渲染配置相关联的一组约束,一些公开的方法涉及并行计算较低复杂性的解(例如,渲染配置的较低复杂性版本)和较高复杂性的解(例如,渲染配置的

较高复杂性版本)。例如,可以选择渲染配置的较低复杂性版本的维度,使得可以以最小的延迟来计算渲染配置的较低复杂性版本。在一些实例中,可以并行计算较高复杂性的解。根据一些这样的示例,一旦渲染配置的较低复杂性版本可用,渲染配置转变就可以开始例如使用上文描述的方法。稍后,当渲染配置的较高复杂性版本变得可用时,到渲染配置的较低复杂性版本的转变可以被到渲染配置的较高复杂性版本的转变中断(如果渲染配置转变尚未完成),例如,如上所述。

[0364] 下表1提供了高复杂性和低复杂性渲染配置查找表的一些示例维度,以说明其计算中的差异的数量级。

[0365]

配置类型	x	y	z	频率(N)	最小化成本函数
低复杂性	3	3	2	1	18
高复杂性	9	9	5	20	8,100

[0366] 表1

[0367] 频域实施方式的附加示例

[0368] 图15呈现了根据一个示例的与频域渲染器相对应的块。与其他公开的实施方式一样,图15中呈现的元件的数量、布置和类型仅作为示例。根据该示例,频域渲染器1515、QMF 1510和逆QMF 1535经由控制系统110h实施,该控制系统是上文参考图1A描述的控制系统110的实例。此处,控制系统110h被示出为接收节目流,在该实例中,该节目流包括节目流音频数据1505a和节目流元数据1505b。除非另有说明,否则读者可以假设图15的块1510、1512、1513、1516和1535提供图13的块1310、1312、1313、1316和1335的功能。

[0369] 在该示例中,频域渲染器1515被配置成将一组扬声器激活应用于节目流音频数据1505a以生成M个输出,扩音器1至M中的每一个有一个输出。除了扬声器激活之外,在该示例中,频域渲染器1515被配置成将变化的延迟应用于M个输出,例如以将声音从每个扩音器到达收听位置的时间对准。例如,在M个扬声器激活由时域滤波器系数表示的情况下,这些延迟可以被实施为样本延迟和群延迟的任何组合。在其中在频域中实施渲染的一些示例中,M个扬声器激活可以由N个频域滤波器系数表示,并且延迟可以由变换块延迟(在图15的示例中经由变换块延迟线模块1518实施)与由频域滤波器系数应用的残差线性相位项(子块延迟)(在图15的示例中经由子块延迟模块1520实施)的组合来表示。根据一些示例,子块延迟可以被实施为滤波器组的每个频带的简单复数乘法器,每个频带的复数值是根据斜率等于子块延迟的负值的线性相位项选择的。为了得到更精确的结果,子块延迟可以用跨滤波器组的块操作的更高精度的多抽头滤波器来实施。

[0370] 图16呈现了与用于管理渲染配置之间的转变的另一实施方式相对应的块。与其他公开的实施方式一样,图16中呈现的元件的数量、布置和类型仅作为示例。根据该示例,频域渲染器1615、QMF 1610和逆QMF 1635经由控制系统110h实施,该控制系统是上文参考图1A描述的控制系统110的实例。此处,控制系统110h被示出为接收节目流,在该实例中,该节目流包括节目流音频数据1605a和节目流元数据1605b。

[0371] 为了支持可以连续地和任意地中断而没有间断的渲染配置转变,可以使用前述对扬声器激活进行插值的方法,如上文参考图13描述的。在图16中示出的示例中,如在上文参考图13描述的示例中,存在2种激活渲染配置,每种激活渲染配置都有其自己的数据结构(数据结构1612a和1612b,在图16中也被称为查找表A和B)。因此,除非另有说明,否则读者

可以假设图16的块1610、1612a、1612b、1613a、1613b、1616和1635提供图13的块1310、1312a、1312b、1313a、1313b、1316和1335的功能。此外,除非另有说明,否则读者可以假设图16的块1618提供图15的变换块延迟线模块1518的功能。

[0372] 例如,控制系统110g的模块1614可以被配置成至少部分地基于交叉渐变时间1611来确定两组扬声器激活之间的幅度归一化插值。根据该示例,控制系统110的模块1614被配置成:基于从三线性插值模块1613a接收的表A的插值以及基于从三线性插值模块1613b接收的表B的插值,确定单个扬声器激活表。在该实施方式中,控制系统110g的模块1616被配置成根据由模块1614确定的扬声器激活表在频域中计算每个音频对象的最后一组扬声器激活。此处,模块1616向变换块延迟线模块1618输出扬声器馈送,扩音器1至M中的每一个有一个扬声器馈送。

[0373] 根据该示例,变换块延迟线模块1618应用一组延迟线,每个扬声器馈送一个延迟线。如在上文参考图15描述的示例中,延迟可以由变换块延迟(在图16的示例中经由变换块延迟线1618实施)和根据频域滤波器系数应用的残差线性相位项(子块延迟,在本文中也可以被称为子块延迟滤波器)的组合来表示。在该示例中,子块延迟是允许不是频域变换块大小的确切倍数的延迟的残差相位项。

[0374] 根据该示例,每种激活渲染配置还具有其自己的对应延迟和读取偏移。此处,读取偏移A用于与表A相对应的渲染配置(或渲染配置版本),而读取偏移B用于与表B相对应的渲染配置(或渲染配置版本)。根据一些示例,“读取偏移A”对应于与渲染配置A相关联的一组M个读取偏移,M个声道中的每一个有一个读取偏移。在这样的示例中,“读取偏移B”对应于与渲染配置B相关联的一组M个读取偏移。在这样的实施方式中,对延迟的比较和使用单位幂和或者单位振幅和进行的选择可以在每声道的基础上进行。如上所述,根据一些示例,在从每个延迟线读取之后,使用附加滤波级来实施与对应于表A和表B的渲染配置相关联的子块延迟。在该示例中,与查找表A相对应的激活渲染配置的子块延迟由子块延迟模块1620a实施,而与查找表B相对应的激活渲染配置的子块延迟由子块延迟模块1620b实施。

[0375] 在该示例中,交叉渐变模块1625对每种配置的多组延迟的扬声器馈送(子块延迟模块1620a和子块延迟模块1620b的输出)进行交叉渐变,以产生单组M个输出扬声器馈送。在一些示例中,交叉渐变模块1625可以被配置成对每种渲染配置应用交叉渐变窗口。根据一些实施方式,交叉渐变模块1625可以被配置成至少部分地基于延迟线读取偏移A和B来选择交叉渐变窗口。

[0376] 可以使用许多可能的对称交叉渐变窗口对。因此,根据特定实施方式,交叉渐变模块1625可以被配置成以不同方式选择交叉渐变窗口对。在一些实施方式中,如果延迟线读取偏移A和B在可以根据变换块大小样本确定的程度上不相同,则交叉渐变模块1625可以被配置成选择交叉渐变窗口以具有单位幂和。实际上来说,如果渲染配置A和B的总延迟在彼此的变换块大小样本内,则读取偏移A和B将看起来是相同的。例如,如果变换块包括64个样本,则在48 kHz的采样率下,对应的时间间隔大约为1.333毫秒。相差超过阈值量。根据一些示例,这种条件可以如下表示:

[0377] $w(i)^2 + (1 - w(i))^2 = 1$ (30)

[0378] 在等式30中,i表示与时间相关的块索引,但是在频域中。满足等式30的标准的w(i)的一个示例是:

$$[0379] \quad w(i) = \cos\left(i * \frac{\pi}{2}\right) \quad (31)$$

[0380] 图17示出了具有单位幂和的交叉渐变窗口对的示例。在该示例中,图17中呈现的一对窗口 $w(i)$ 和 $(1-w(i))$ 基于等式31。

[0381] 然而,如图16中示出的,如果对于给定的输出声道读取偏移A和B相等,则这意味着与渲染配置A和B的给定声道相关联的总延迟(组合的块和子块延迟)是类似的(例如,在变换块大小的样本数量内)。在这种情况下,交叉渐变窗口应该具有单位振幅和(本文中也被称为“单位和”)而不是单位幂和,因为被组合的信号很可能是高度相关的。满足单位和标准的 $w(i)$ 的一些示例如下:

$$[0382] \quad w(i) = i \quad (32)$$

$$[0383] \quad w(i) = \cos\left(i * \frac{\pi}{2}\right)^2 \quad (33)$$

[0384] 图18A和图18B呈现了具有单位振幅和的交叉渐变窗口对的示例。图18A示出了根据等式32的具有单位和的交叉渐变窗口对的示例。图18B示出了根据等式33的具有单位和的交叉渐变窗口对的示例。

[0385] 先前的交叉渐变窗口的示例很简单。然而,图16的交叉渐变模块1625可能需要更一般化的窗口设计方法,以便能够处理可以连续地和任意地中断而没有间断的渲染配置转变。在渲染配置转变的每次中断期间,可能需要从块延迟线进行附加读取。在如图16中示出的实施方式中,如果从渲染配置A到渲染配置B的转变因为接收到指示转变到渲染配置C的第二渲染转变指示而中断,则将涉及3个块延迟线读取和三个读取偏移,并且交叉渐变模块1625应该实施3部分交叉渐变窗口。

[0386] 图19呈现了与用于管理第一至第L组扬声器激活之间的转变的另一实施方式相对应的块。在该示例中,L是大于二的整数。与其他公开的实施方式一样,图19中呈现的元件的数量、布置和类型仅作为示例。根据该示例,频域渲染器1915、QMF 1910和逆QMF 1935经由控制系统110i实施,该控制系统是上文参考图1A描述的控制系统110的实例。此处,控制系统110i被示出为接收节目流,在该实例中,该节目流包括节目流音频数据1903a和节目流元数据1903b。

[0387] 在一些示例中,L组扬声器激活可以对应于L种渲染配置。然而,如本文别处所述,一些实施方式可以涉及与单个渲染配置的多个版本相对应的多组扬声器激活。例如,第一组扬声器激活可以用于渲染配置的简化版本,而第二组扬声器激活可以用于渲染配置的完整版本。在一些这样的示例中,单个渲染转变指示可以导致到渲染配置的简化版本的第一转变和到渲染配置的完整版本的第二转变。

[0388] 因此,在参考图19描述的示例(以及本文中公开的其他示例)中,在转变指示、扬声器激活集合与渲染配置之间可能有也可能没有相同的数值关系。如果存在响应于接收到的渲染转变指示的简化到完整的转变,则可以为渲染转变指示确定两组扬声器激活,并且可能有两个转变对应于渲染转变指示。然而,如果没有响应于渲染转变指示的简化到完整的转变,则可以为渲染转变指示确定一组扬声器激活,并且可能只有一个转变对应于渲染转变指示。

[0389] 为了简单起见,以下讨论将假设图19的L组扬声器激活对应于L种渲染配置。为了支持可以连续地和任意地中断而没有间断的渲染配置转变,可以使用上述对扬声器激活进行插值的方法(例如,如上文参考图13、图14和图16所描述的)。因此,除非另有说明,否则读者可以假设图19的块1910、1912a、1912b、1913a、1913b、1916和1935提供图16的块1610、1612a、1612b、1613a、1613b、1616和1635的功能。此外,除非另有说明,否则读者可以假设图19的块1918提供图16的块1618的功能。

[0390] 在图16中示出的示例中,存在2种激活渲染配置,每种激活渲染配置都具有对应的数据结构(数据结构1612a和1612b,在图16中也被称为查找表A和B),并且每种激活渲染配置都具有对应的子块延迟模块(子块延迟模块1620a和1620b)。在图19中示出的示例中,存在L种激活渲染配置,每种激活渲染配置都具有对应的子块延迟模块(子块延迟模块1920a-1920l)。

[0391] 然而,在图19中示出的示例中,仍然只存在两种当前渲染配置数据结构:数据结构1912a对应于当前渲染配置(可以是转变渲染配置),而数据结构1912b对应于目标渲染配置(在该示例中是渲染配置L)。这种效率是通过上文参考图14描述的実施方式的可伸缩性来实现的:在图19中示出的示例中,控制系统110i还被配置成实施表组合模块1905,该表组合模块可以如上文参考图14的表组合模块1405描述的那样起作用。因此,在该示例中,不管中断的次数和激活渲染配置的数量如何,只实施两个激活查找表。

[0392] 根据图19中示出的示例,交叉渐变模块1925被配置成实施L部分交叉渐变窗口。L部分交叉渐变窗口集合的设计应该考虑L个块延迟线读取中的哪些具有匹配的读取偏移以及哪些具有不同的读取偏移(只要给定变换块大小就可以确定)。交叉渐变窗口集合设计还应该考虑在给定中断(i_0)时作为先前交叉渐变窗口集合的一部分的(L-1)种渲染配置的值,以确保平滑转变。

[0393] 在一些实施方式中,可以使用以下方法来设计L部分交叉渐变窗口集合。

[0394] 1. 给定对称的一对衰减和上升窗口函数 $w(i)$ 和 $1-w(i)$,最初前L-1个窗口(与(多种)当前渲染配置相对应的窗口)可以被设置为衰减函数 $w(i)$,而最后的窗口L(与目标渲染配置相对应的窗口)可以被设置为上升函数 $1-w(i)$ 。

[0395] 2. L-1个衰减窗口可以由最后一次中断(i_0)时的先前窗口值来缩放。如上所述, i 表示与时间相关的块索引,但是在频域中。因此, i_0 可以是对应于时间点的频域中的样本或样本块。

[0396] 3. L个窗口的集合($l_{1..L}(i)$)可以通过对具有匹配读取偏移的窗口进行求和来分组为K个组。K的范围可以为1到L。如果没有一个读取偏移匹配,则将有L个组。如果所有读取偏移都匹配,则将有1个组。

[0397] 4. K个窗口组($k_{1..K}(i)$)中的每一个的幂之和例如如下计算:

$$p(i) = \sum_{j=1}^K (k_j(i))^2 \quad (34)$$

[0399] 5. 然后将归一化函数计算为一除以K个窗口组的幂之和:

$$n(i) = \frac{1}{p(i)} \quad (35)$$

[0401] 6.归一化函数的平方根可以用于缩放L个窗口的原始集合,这确保了窗口组的幂之和将是单位的:

$$[0402] \quad l_{1...L}(i) = \sqrt{n(i)}(l_{1...L}(i)). \quad (36)$$

[0403] 在一些实施方式中,等式36中示出的幂和约束的更一般化的公式取代了等式31的公式。使用这种方法意味着 $w(i)$ 的选择将影响在直接、非中断转变中使用的简单幂保持交叉渐变窗口对的形状。

[0404] 一种选择 $w(i)$ 的方法是组合等式32和33:

$$[0405] \quad w(i) = \alpha * i + (1 - \alpha) * \cos\left(i * \frac{\pi}{2}\right)^2 \quad (37)$$

[0406] 通过为 α (如 $\alpha = 0.75$) 选择适当的值并将等式37用作先前段落中描述的L部分交叉渐变设计算法中的基本窗口函数 $w(i)$,可以得到类似于等式31的幂保持交叉渐变窗口对和类似于等式33的线性交叉渐变窗口对。

[0407] 图20A示出了根据等式31和37的具有单位幂和的交叉渐变窗口对的示例,其中 $\alpha = 0.75$ 。其他示例可以具有其他 α 值,如0.60、0.65、0.70、0.80、0.85等。

[0408] 图20B示出了根据等式33和37的具有单位和的交叉渐变窗口对的示例,其中 $\alpha = 0.75$ 。其他示例可以具有其他 α 值。

[0409] 图21、图22、图23、图24和图25是呈现了没有、具有一些或所有匹配的读取偏移的交叉渐变窗口的示例的图。在这些示例中,交叉渐变窗口是使用公式34-37设计的,y轴表示 $w(i)$,并且x轴表示 i 。如上所述, i 表示与时间相关的块索引,但是在频域中。

[0410] 在这些示例中,在与 $0 \leq i < i_1$ 相对应的时间间隔期间,发生了从渲染配置A到渲染配置B的第一渲染配置转变。第一渲染配置转变可以例如已经响应于第一渲染转变指示。根据这些示例,在与 i_1 相对应的时间处或附近,第一渲染配置转变因为接收到指示到渲染配置C的转变的第二渲染转变指示而中断。

[0411] 在这些示例中,在与 $i_1 \leq i < i_2$ 相对应并且不需要完成到渲染配置B的转变的时间间隔期间,发生了到渲染配置C的第二渲染配置转变。

[0412] 根据这些示例,在与 i_2 相对应的时间处或附近,第二渲染配置转变因为接收到指示到渲染配置D的转变的第三渲染转变指示而中断。在这些示例中,在与 $i_2 \leq i < 1$ 相对应并且不需要完成到渲染配置C的转变的时间间隔期间,发生了到渲染配置D的第三渲染配置转变。在这些示例中,第三渲染配置转变在与 $i = 1$ 相对应的时间完成。

[0413] 在图21中呈现的示例中,渲染配置A-D的读取偏移中没有有一个匹配。因此,所有渲染配置转变的交叉渐变窗口都被选择为具有单位幂和。

[0414] 根据图22中呈现的示例,渲染配置A-D的所有读取偏移都匹配。因此,所有渲染配置转变的交叉渐变窗口都被选择为具有单位和。

[0415] 在图23中呈现的示例中,只有渲染配置A和B的读取偏移匹配。因此,在该示例中,用于第一渲染配置转变(从渲染配置A到渲染配置B)的交叉渐变窗口对被选择为具有单位和,而用于第二和第三渲染配置转变的交叉渐变窗口对被选择为具有单位幂和。

[0416] 根据图24中呈现的示例,渲染配置A和B的读取偏移匹配,并且渲染配置C和D的读取偏移匹配。然而,渲染配置B和C的读取偏移不匹配。因此,在该示例中,用于第一渲染配置

转变(从渲染配置A到渲染配置B)和第三渲染配置转变(从渲染配置C到渲染配置D)的交叉渐变窗口对被选择为具有单位 $\sum$,而用于第二渲染配置转变的交叉渐变窗口对被选择为具有单位 $\sum$ 。

[0417] 在图25中呈现的示例中,渲染配置A和C的读取偏移匹配,并且渲染配置B和D的读取偏移匹配。然而,渲染配置A和B的读取偏移不匹配,并且渲染配置C和D的读取偏移不匹配。因此,在该示例中,用于第一渲染配置转变(从渲染配置A到渲染配置B)的交叉渐变窗口对被选择为具有单位 $\sum$ 。然后,到渲染配置C的后续转变被选择为在应用于渲染配置B的窗口与应用于渲染配置A和C的窗口之和之间具有单位 $\sum$ 。最后,对于到渲染配置D的转变,交叉渐变窗口被设计为在应用于配置A和C的求和窗口与应用于配置B和D的求和窗口之间具有单位 $\sum$ 。

[0418] 使用如图19中示出的频域渲染器可以允许任意数量的中断,但在一些实例中可能无法很好地缩放。因此,在一些实施方式中,可以限制需要延迟线读取和交叉渐变的激活渲染配置的数量,以便保持复杂性受限。图26、图27、图28、图29和图30图示了与图21至图25相同的情况,但限制了3种激活渲染配置。在这些示例中,通过消除与 i_2 相对应的时间处渲染配置B的贡献,激活渲染配置的数量被限制为三种。在一些替代性实施方式中,通过消除与 i_2 相对应的时间处渲染配置A的贡献,激活渲染配置的数量可以被限制为三种。在一些实例中,这种情况可以通过将激活渲染配置的数量限制为三种来概括,其中,除了在与 i_2 相对应的时间处具有最大幅度的相关联的交叉渐变窗口的两种其他渲染配置(在该示例中为B、C和D)之外,总是包括最新的渲染配置。注意,在这些示例中, $\sum$ 和在所有情况下都保持统一,因为剩余渲染配置的贡献是归一化的。

[0419] 图31是概述方法的示例的流程图。与本文描述的其他方法一样,不必以所指示的顺序来执行方法3100的框。此外,这样的方法可以包括比所示出和/或所描述的框更多或更少的框。在一些示例中,可以同时执行这样的方法的一个或多个框。方法3100的框可以由一个或多个设备执行,该设备可以是(或可以包括)控制系统,如图1A、图1B和图2A中示出的和上文描述的控制系统110、控制系统110a或控制系统110b,或者由其他公开的控制系统的示例之一执行。

[0420] 在该实施方式中,框3105涉及由控制系统并且经由接口系统接收音频数据。此处,音频数据包括一个或多个音频信号和相关联的空间数据。在该实例中,空间数据指示与音频信号相对应的预期感知空间位置。根据一些示例(例如,对于如Dolby Atmos™等音频对象实施方式),空间数据可以是或者可以包括位置元数据。然而,在一些实例中,第一组扬声器激活可以对应于基于声道的音频格式。在一些这样的实例中,预期感知空间位置可以对应于基于声道的音频格式的声道(例如,可以对应于左声道、右声道、中声道等)。

[0421] 在该示例中,框3110涉及由控制系统渲染音频数据以经由环境的一组扩音器再现,从而产生第一经渲染的音频信号。在该实施方式中,渲染音频数据以供再现涉及根据第一渲染配置确定环境中的一组扩音器的第一相对激活。在该示例中,第一渲染配置对应于第一组扬声器激活。根据一些示例,第一组扬声器激活可以用于三维空间中对应的多个位置中的每一个。在该实施方式中,框3115涉及经由接口系统向环境的该组扩音器中的至少一些扩音器提供第一经渲染的音频信号。

[0422] 在该示例中,框3120涉及由控制系统并且经由接口系统接收第一渲染转变指示。

在该实施方式中,第一渲染转变指示指示从第一渲染配置到第二渲染配置的转变。

[0423] 在该实施方式中,框3125涉及由控制系统确定与第二渲染配置的简化版本相对应的第二组扬声器激活。根据该示例,框3130涉及由控制系统执行从第一组扬声器激活到第二组扬声器激活的第一转变。

[0424] 在该实施方式中,框3135涉及由控制系统确定与第二渲染配置的完整版本相对应的第三组扬声器激活。在一些实例中,框3135可以与框3125和/或框3130同时执行。在该示例中,框3140涉及由控制系统执行到第三组扬声器激活的第二转变,而无需完成第一转变。

[0425] 根据一些示例,方法3100可以涉及由控制系统并且经由接口系统接收第二渲染转变指示。在一些这样的示例中,第二渲染转变指示指示到第三渲染配置的转变。在一些示例中,方法3100可以涉及由控制系统确定与第三渲染配置的简化版本相对应的第四组扬声器激活。在一些示例中,方法3100可以涉及由控制系统执行从第三组扬声器激活到第四组扬声器激活的第三转变。在一些示例中,方法3100可以涉及由控制系统确定与第三渲染配置的完整版本相对应的第五组扬声器激活,并且由控制系统执行到第五组扬声器激活的第四转变,而无需完成第一转变、第二转变或第三转变。

[0426] 在一些示例中,方法3100可以涉及由控制系统并且经由接口系统顺序地接收第二至第(N)渲染转变指示。一些这样的方法可以涉及由控制系统为第二至第(N)渲染转变指示中的每一个确定第一组扬声器激活和第二组扬声器激活。第一组扬声器激活可以对应于渲染配置的简化版本,而第二组扬声器激活可以对应于第二至第(N)渲染转变指示中的每一个的渲染配置的完整版本。在一些示例中,方法3100可以涉及由控制系统顺序地执行从第四组扬声器激活到第(2N)组扬声器激活的第三至第(2N-1)转变。在一些示例中,方法3100可以涉及由控制系统执行到第(2N+1)组扬声器激活的第(2N)转变,而无需完成第一至第(2N)转变中的任一项。在一些实施方式中,单个渲染器实例可以渲染音频数据以供再现。

[0427] 然而,不一定是所有渲染转变指示都将涉及响应于接收到的渲染转变指示的简化到完整的转变的情况。如在以上示例中,如果存在响应于接收到的渲染转变指示的简化到完整的渲染转变,则可以为渲染转变指示确定两组扬声器激活,并且可能有两个转变对应于渲染转变指示。然而,如果没有响应于渲染转变指示的简化到完整的转变,则可以为渲染转变指示确定一组扬声器激活,并且可能只有一个转变对应于渲染转变指示。

[0428] 在一些示例中,方法3100可以涉及由控制系统并且经由接口系统接收第二渲染转变指示。第二渲染转变指示可以指示到第三渲染配置的转变。在一些这样的示例中,方法3100可以涉及由控制系统确定与第三渲染配置相对应的第四组扬声器激活,并且由控制系统执行到第四组扬声器激活的第三转变,而无需完成第一转变或第二转变。

[0429] 根据一些示例,方法3100可以涉及由控制系统并且经由接口系统接收第三渲染转变指示。第三渲染转变指示可以指示到第四渲染配置的转变。在一些实例中,方法3100可以涉及由控制系统确定与第四渲染配置相对应的第五组扬声器激活,并且由控制系统执行到第五组扬声器激活的第四转变,而无需完成第一转变、第二转变或第三转变。

[0430] 在一些示例中,方法3100可以涉及由控制系统并且经由接口系统顺序地接收第二至第(N)渲染转变指示,并且由控制系统确定与第二至第(N)渲染转变指示相对应的第四至第(N+2)组扬声器激活。根据一些示例,方法3100可以涉及由控制系统顺序地执行从第四组扬声器激活到第(N+1)组扬声器激活的第三至第(N)转变,并且由控制系统执行到第(N+2)

组扬声器激活的第(N+1)转变,而无需完成第一至第(N)转变中的任一种。

[0431] 根据一些实施方式,第一组扬声器激活、第二组扬声器激活和第三组扬声器激活是频率相关的扬声器激活。在一些这样的示例中,应用频率相关的扬声器激活可以涉及在第一频带中应用感知空间位置的模型,该模型产生与收听者的左耳和右耳处的音频对象位置相对应的双耳响应。替代性地或另外地,应用频率相关的扬声器激活可以涉及在至少第二频带中应用感知空间位置模型,该模型将从一组扩音器播放的音频信号的感知空间位置置于由扩音器的相关联的激活增益加权的该组扩音器的位置的质心处。

[0432] 在一些示例中,第一组扬声器激活、第二组扬声器激活或第三组扬声器激活中的至少一者可以是优化成本的结果,该成本是当在环境中的一组扩音器上回放时播放的音频信号的感知空间位置的模型的函数。在一些实例中,成本可以是音频信号的预期感知空间位置与该组扩音器中的每个扩音器的位置的接近度的度量的函数。替代性地或另外地,成本可以是基于以下各项中的一项或多项的一个或多个附加动态可配置功能的度量的函数:扩音器与一个或多个收听者的接近度;扩音器与吸引力位置的接近度,其中,吸引力是更有利于更接近该吸引力位置的相对更高的扩音器激活的因素;扩音器与排斥力位置的接近度,其中,排斥力是更有利于更接近该排斥力位置的相对更低的扩音器激活的因素;每个扩音器相对于该环境中的其他扩音器的能力;该扩音器关于其他扩音器的同步;唤醒词性能;和/或回声消除器性能。

[0433] 根据一些实施方式,渲染音频数据以供再现可以涉及根据渲染配置确定单组插值激活,并且应用该单组插值激活来产生单组经渲染的音频信号。在一些这样的示例中,单组经渲染的音频信号可以被馈送到一组扩音器延迟线中。该组扩音器延迟线可以包括用于多个扩音器中的每个扩音器的一个扩音器延迟线。

[0434] 在一些示例中,渲染音频数据以供再现可以在频域中执行。因此,在一些实例中,渲染音频数据以供再现可以涉及在频域中确定和实施扩音器延迟。根据一些这样的示例,在频域中确定和实施扬声器延迟可以涉及确定和实施变换块延迟和由频域滤波器系数应用的子块延迟的组合。在一些示例中,子块延迟可以是允许不是频域变换块大小的确切倍数的延迟的残差相位项。因此,在一些示例中,渲染音频数据以供再现可以涉及实施子块延迟滤波。在一些实施方式中,渲染音频数据以供再现可以涉及实施具有单独读取偏移的一组块延迟线。

[0435] 在一些示例中,渲染音频数据以供再现可以涉及为每个渲染配置确定和应用插值扬声器激活和交叉渐变窗口。根据一些这样的示例,渲染音频数据以供再现可以涉及实施具有单独延迟线读取偏移的一组块延迟线。在一些这样的示例中,交叉渐变窗口选择可以至少部分地基于延迟线读取偏移。在一些实例中,其中,如果延迟线读取偏移相差超过阈值量,则交叉渐变窗口可以被设计为具有单位幂和。根据一些示例,如果延迟线读取偏移相同或相差不超过阈值量,则交叉渐变窗口可以被设计为具有单位和。

[0436] 如上所述,在一些实施方式中,单个渲染器实例可以渲染音频数据以供再现。

[0437] 图32是概述另一种方法的示例的流程图。与本文描述的其他方法一样,不必以所指示的顺序来执行方法3200的框。此外,这样的方法可以包括比所示出和/或所描述的框更多或更少的框。在一些示例中,可以同时执行这样的方法的一个或多个框。方法3200的框可以由一个或多个设备执行,该设备可以是(或可以包括)控制系统,如图1A、图1B和图2A中示

出的和上文描述的控制系统110、控制系统110a或控制系统110b,或者由其他公开的控制系统示例之一执行。

[0438] 在该实施方式中,框3205涉及由控制系统并且经由接口系统接收音频数据。在该示例中,音频数据包括一个或多个音频信号和相关联的空间数据。此处,空间数据指示与音频信号相对应的预期感知空间位置。

[0439] 在该示例中,框3210涉及由控制系统渲染音频数据以经由环境的一组扩音器再现,从而产生第一经渲染的音频信号。根据该示例,渲染音频数据以供再现涉及根据第一渲染配置确定环境中的一组扩音器的第一相对激活。

[0440] 在该实施方式中,第一渲染配置对应于第一组扬声器激活。在一些实例中,第一组扬声器激活可以用于三维空间中对应的多个位置中的每一个。在一些示例中,空间数据可以是或者可以包括位置元数据。然而,在一些示例中,第一组扬声器激活可以对应于基于声道的音频格式。在一些这样的示例中,预期感知空间位置包括基于声道的音频格式的声道。在图32中示出的实施方式中,框3215涉及经由接口系统向环境的该组扩音器中的至少一些扩音器提供第一经渲染的音频信号。

[0441] 根据该示例,框3220涉及由控制系统并且经由接口系统顺序地接收第一至第(L-1)渲染转变指示。在该实例中,第一至第(L-1)渲染转变指示中的每一个指示从当前渲染配置到新渲染配置的转变。

[0442] 在该实施方式中,框3225涉及由控制系统确定与第一至第(L-1)渲染转变指示相对应的第二至第(L)组扬声器激活。根据该示例,框3230涉及由控制系统顺序地执行从第一组扬声器激活到第(L-1)组扬声器激活的第一至第(L-2)转变。在该实施方式中,框3225涉及由控制系统执行到第(L)组扬声器激活的第(L-1)转变,而无需完成第一至第(L-2)转变中的任一项。

[0443] 在一些实施方式中,单个渲染器实例可以渲染音频数据以供再现。在一些实例中,渲染音频数据以供再现可以在频域中执行。根据一些示例,渲染音频数据以供再现可以涉及根据渲染配置确定单组插值激活,并且应用该单组插值激活来产生单组经渲染的音频信号。

[0444] 在一些这样的示例中,单组经渲染的音频信号可以被馈送到一组扩音器延迟线中。该组扩音器延迟线可以例如包括用于多个扩音器中的每个扩音器的一个扩音器延迟线。

[0445] 参考图33描述了示例实施例。与本文提供的其他图一样,图33中示出的元件的类型和数量仅作为示例提供。其他实施方式可以包括更多、更少和/或不同类型和数量的元件。图33描绘了收听环境的平面图,该收听环境在该示例中为生活空间。根据该示例,环境3300包括在左上方处的客厅3310、在下方中央处的厨房3315、以及在右下方的卧室3322。跨生活空间分布的方框和圆圈表示一组扩音器3305a-3305h,该组扩音器中的至少一些扩音器在一些实施方式中可以是智能扬声器,放置在对空间方便的位置,但不遵循任何标准规定的布局(任意地放置)。在一些示例中,扩音器3305a-3305h可以被协调以实施一个或多个公开的实施例。在该示例中,环境3300包括分布在整个环境中的相机3311a-3311e。在一些实施方式中,环境3300中的一个或多个智能音频设备还可以包括一个或多个相机。该一个或多个智能音频设备可以是单一用途音频设备或虚拟助理。在一些这样的示例中,可选传

传感器系统130的一个或多个相机可以驻留在电视3330中或上、在移动电话中或在智能扬声器(如扩音器3305b、3305d、3305e或3305h中的一个或多个)中。尽管在本公开内容中呈现的环境3300的每个描绘中都未示出相机3311a-3311e,但在一些实施方式中,每个环境3300仍然可以包括一个或多个相机。

[0446] 本公开的一些方面包括一种被配置(例如,被编程)成执行所公开的方法的一个或多个示例的系统或设备,以及一种存储用于实施所公开的方法或其步骤的一个或多个示例的代码的有形计算机可读介质(例如,磁盘)。例如,一些公开的系统可以是或者包括可编程通用处理器、数字信号处理器或微处理器,该可编程通用处理器、数字信号处理器或微处理器用软件或固件编程为和/或以其他方式被配置成对数据执行各种操作中的任一个,包括所公开的方法或其步骤的实施例。这样的通用处理器可以是或者包括计算机系统,该计算机系统包括输入设备、存储器和处理子系统,该处理子系统被编程(和/或以其他方式被配置)为响应于向其断言的数据而执行所公开的方法(或其步骤)的一个或多个示例。

[0447] 一些实施例可以被实施为可配置的(例如,可编程的)数字信号处理器(DSP),该数字信号处理器被配置(例如,被编程和以其他方式被配置)为对(多个)音频信号执行需要的处理,包括对所公开的方法的一个或多个示例的执行。可替代地,所公开的系统(或其元件)的实施例可以被实施为通用处理器(例如,个人计算机(PC)或其他计算机系统或微处理器,其可以包括输入设备和存储器),该通用处理器用软件或固件编程为和/或以其他方式被配置成执行各种操作中的任一个,包括所公开的方法的一个或多个示例。可替代地,本发明系统的一些实施例的元件被实施为被配置(例如,被编程)成执行所公开的方法的一个或多个示例的通用处理器或DSP,并且该系统还包括其他元件(例如,一个或多个扩音器和/或一个或多个麦克风)。被配置成执行所公开的方法的一个或多个示例的通用处理器可以耦接到输入设备(例如,鼠标和/或键盘)、存储器和显示设备。

[0448] 本公开的另一方面是一种计算机可读介质(例如,磁盘或其他有形存储介质),该计算机可读介质存储用于控制一个或多个设备执行所公开方法或其步骤的一个或多个示例的代码。

[0449] 将从以下枚举的示例实施例(EEE)中理解各种特征和方面:

[0450] EEE1.一种音频处理方法,包括:

[0451] 由控制系统并且经由接口系统接收音频数据,所述音频数据包括一个或多个音频信号和相关联的空间数据,所述空间数据指示与音频信号相对应的预期感知空间位置;

[0452] 由所述控制系统渲染所述音频数据以经由环境的一组扩音器再现,从而产生第一经渲染的音频信号,其中,渲染所述音频数据以供再现涉及根据第一渲染配置确定环境中的一组扩音器的第一相对激活,所述第一渲染配置对应于第一组扬声器激活;

[0453] 经由所述接口系统向所述环境的所述一组扩音器中的至少一些扩音器提供所述第一经渲染的音频信号;

[0454] 由所述控制系统并且经由所述接口系统顺序地接收第一至第(L-1)渲染转变指示,所述第一至第(L-1)渲染转变指示中的每一个指示从当前渲染配置到新渲染配置的转变;

[0455] 由所述控制系统确定与所述第一至第(L-1)渲染转变指示相对应的第二至第(L)组扬声器激活;

[0456] 由所述控制系统顺序地执行从所述第一组扬声器激活到所述第(L-1)组扬声器激活的第一至第(L-2)转变;以及

[0457] 由所述控制系统执行到所述第(L)组扬声器激活的第(L-1)转变,而无需完成所述第一至第(L-2)转变中的任一项。

[0458] EEE2.如权利要求EEE1所述的方法,其中,单个渲染器实例渲染所述音频数据以供再现。

[0459] EEE3.如权利要求EEE1或权利要求EEE2所述的方法,其中,渲染所述音频数据以供再现包括根据所述渲染配置确定单组插值激活,并且应用所述单组插值激活来产生单组经渲染的音频信号。

[0460] EEE4.如权利要求EEE3所述的方法,其中,所述单组经渲染的音频信号被馈送到一组扩音器延迟线中,所述一组扩音器延迟线包括用于多个扩音器中的每个扩音器的一个扩音器延迟线。

[0461] EEE5.如权利要求EEE1至EEE4中任一项所述的方法,其中,所述渲染所述音频数据以供再现是在频域中执行的。

[0462] EEE6.如权利要求EEE1至EEE5中任一项所述的方法,其中,所述第一组扬声器激活用于三维空间中对应的多个位置中的每一个。

[0463] EEE7.如权利要求EEE1至EEE6中任一项所述的方法,其中,所述空间数据包括位置元数据。

[0464] EEE8.如权利要求EEE1至EEE5中任一项所述的方法,其中,所述第一组扬声器激活对应于基于声道的音频格式。

[0465] EEE9.如权利要求EEE8所述的方法,其中,所述预期感知空间位置包括所述基于声道的音频格式的声音。

[0466] EEE10.一种装置,所述装置被配置成执行如权利要求EEE1至EEE9中任一项所述的方法。

[0467] EEE11.一种系统,所述系统被配置成执行如权利要求EEE1至EEE9中任一项所述的方法。

[0468] EEE12.一种或多种非暂态介质,其上存储有软件,所述软件包括用于控制一个或多个设备执行如权利要求EEE1至EEE9中任一项所述的方法的指令。

[0469] 虽然在本文中已经描述了具体实施例和应用,但是对于本领域普通技术人员而言显而易见的是,在不脱离在本文中描述的并要求保护的范围的情况下,可以对本文所述的实施例和应用进行许多改变。应当理解,虽然已经示出和描述了某些形式,但是本公开的范围不限于所描述和示出的具体实施例或所描述的具体方法。

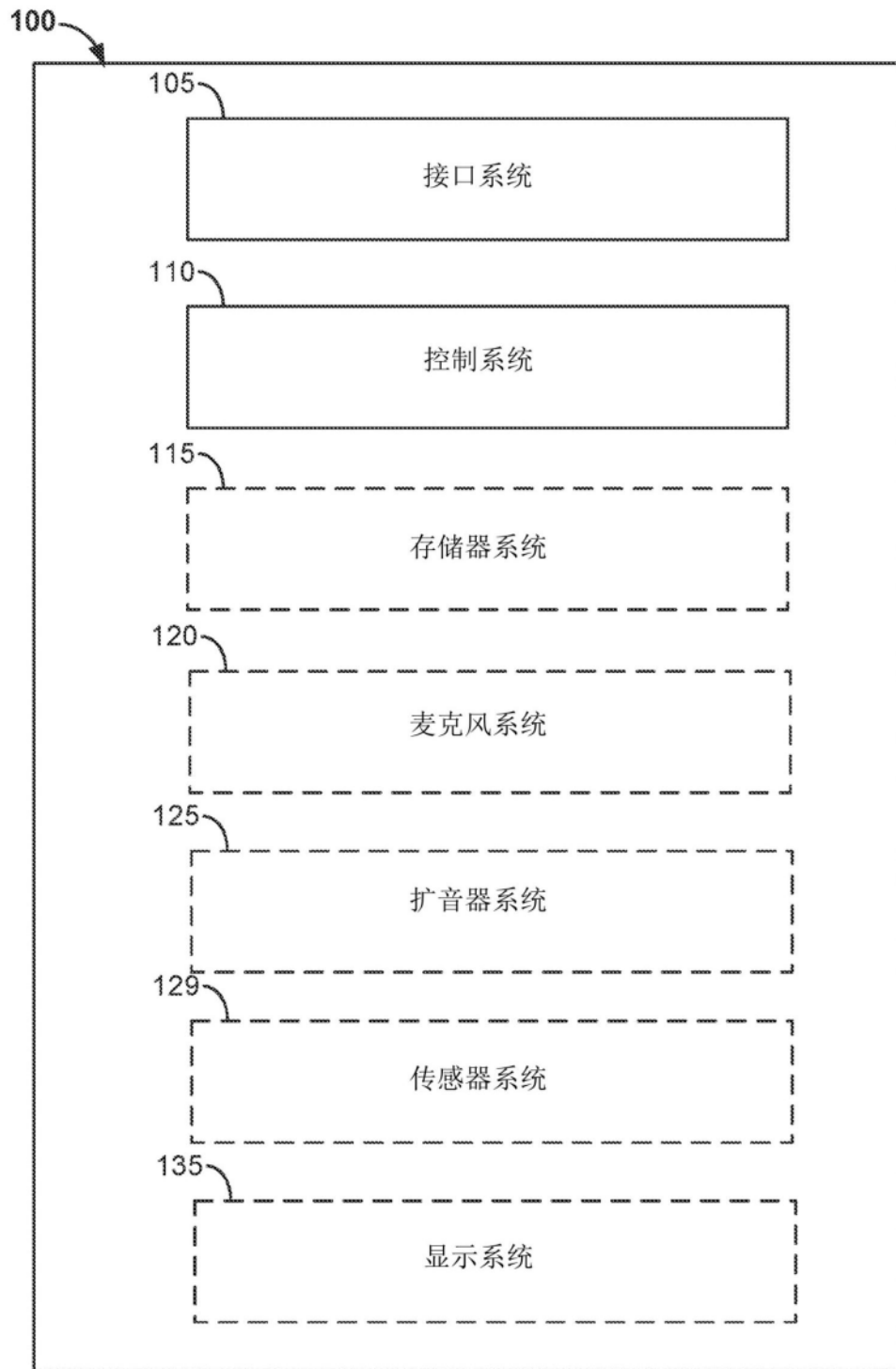


图1A

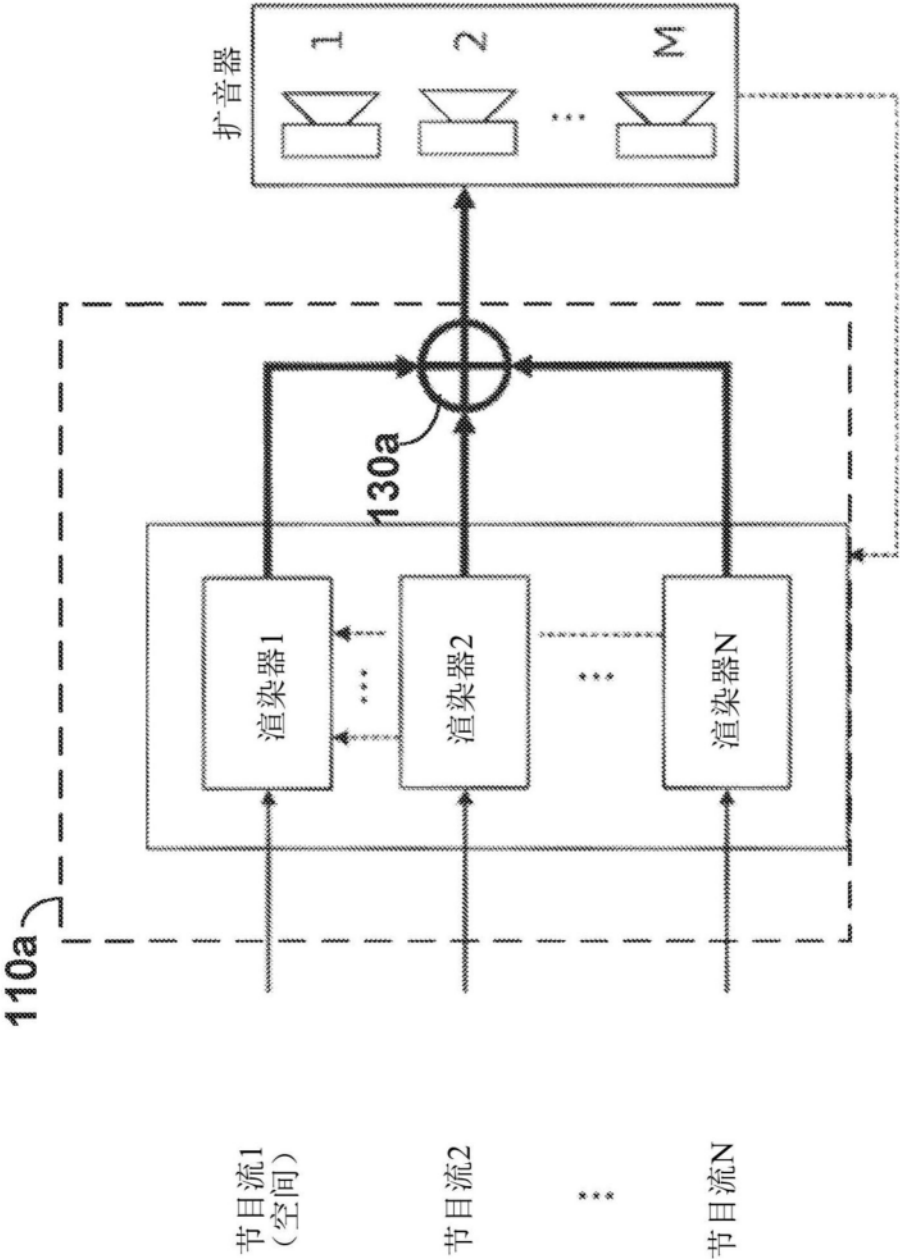


图1B

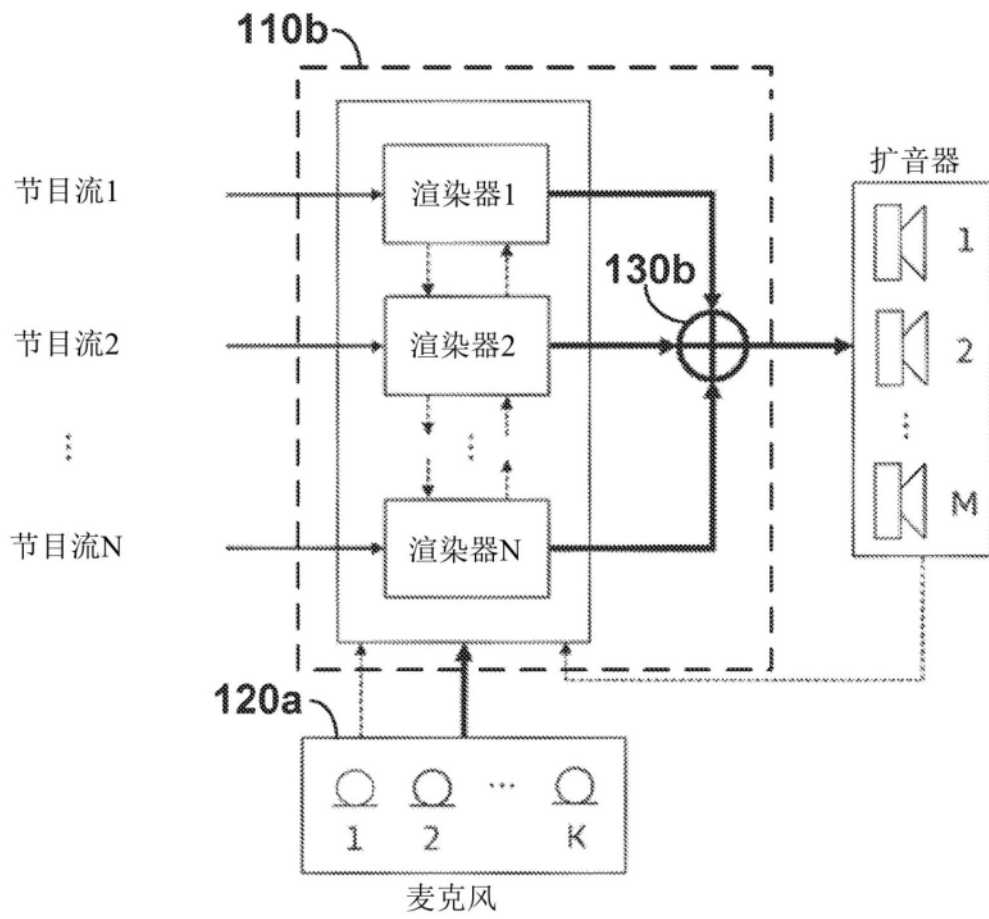


图2A

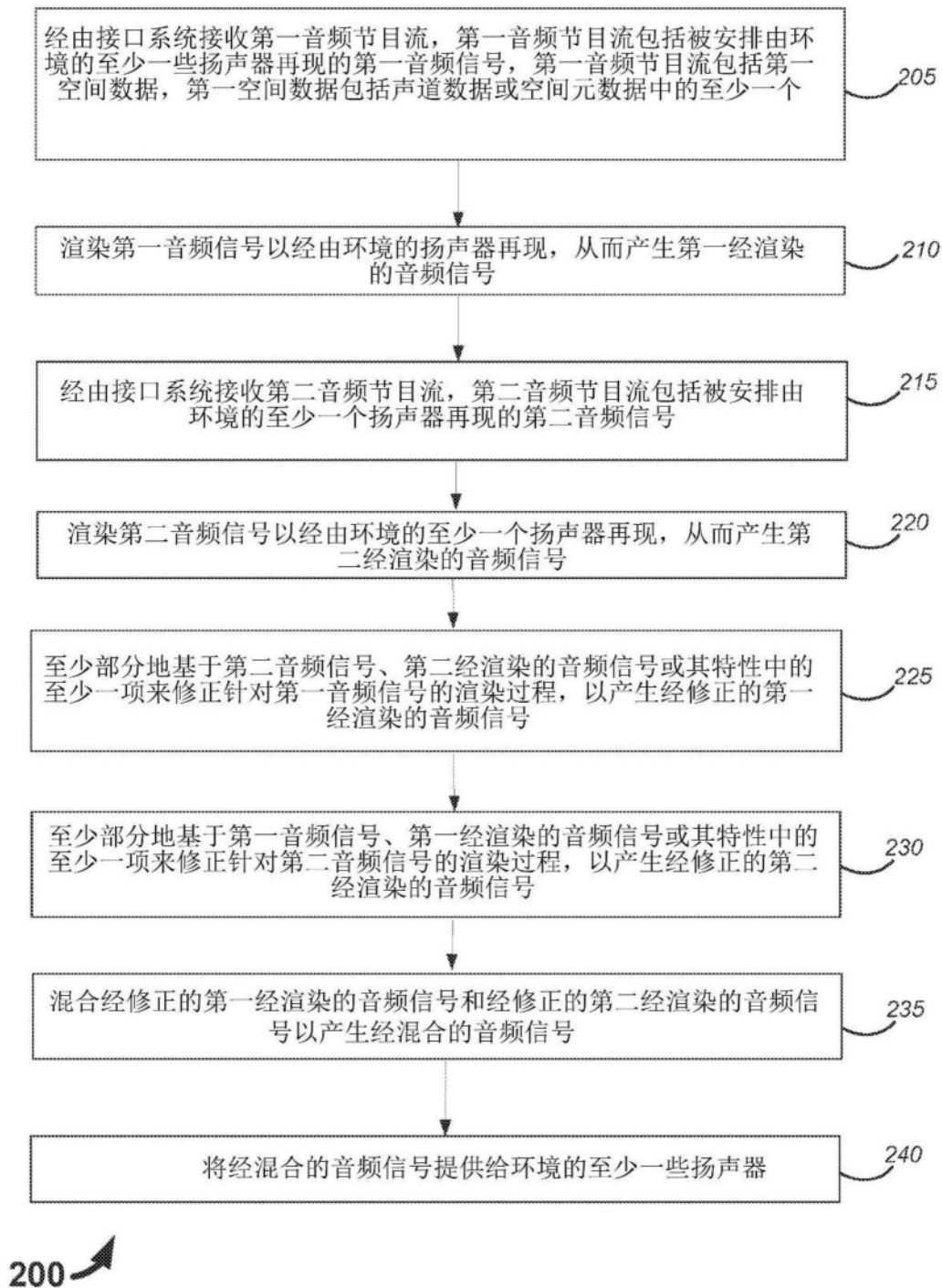


图2B

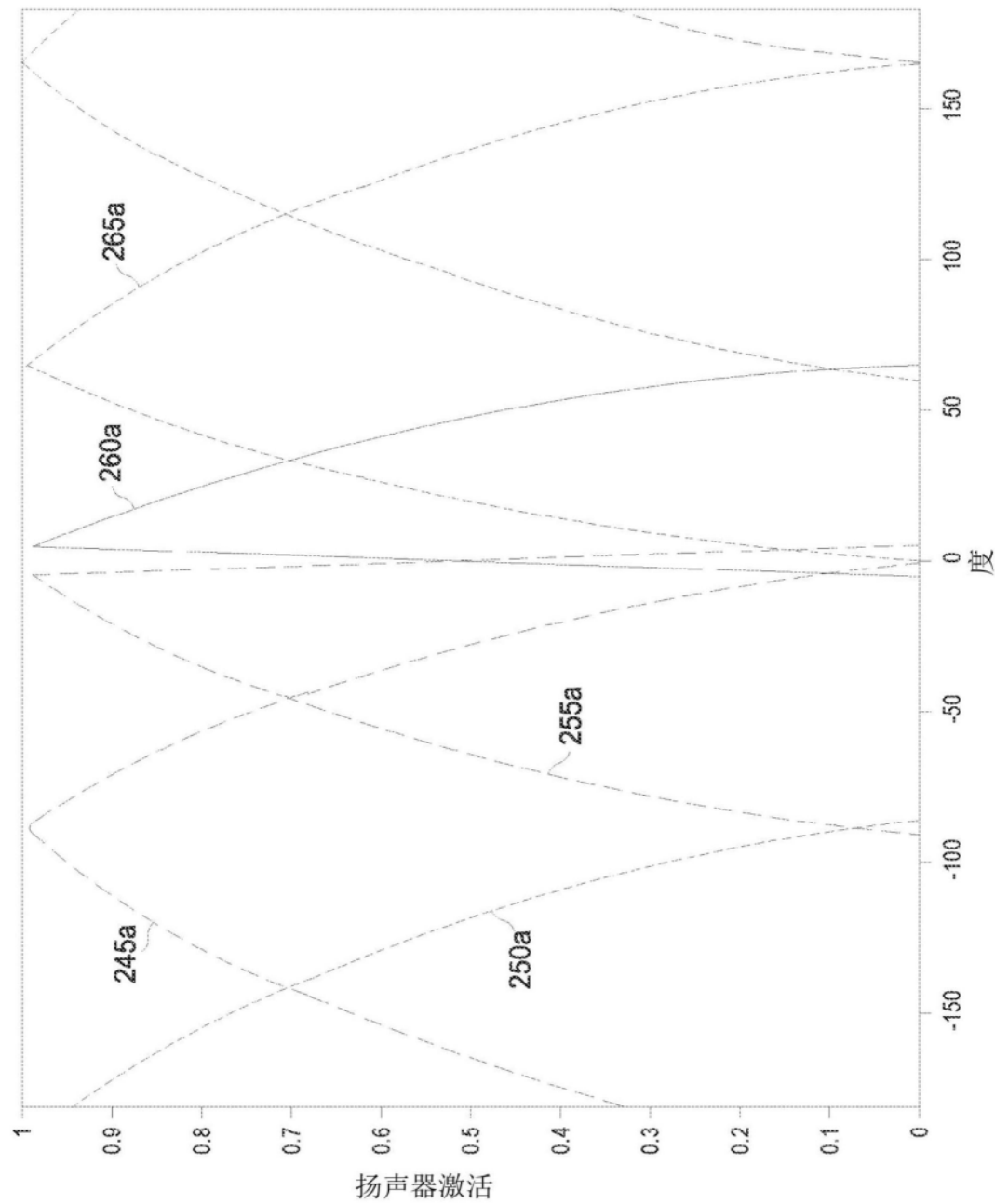


图2C

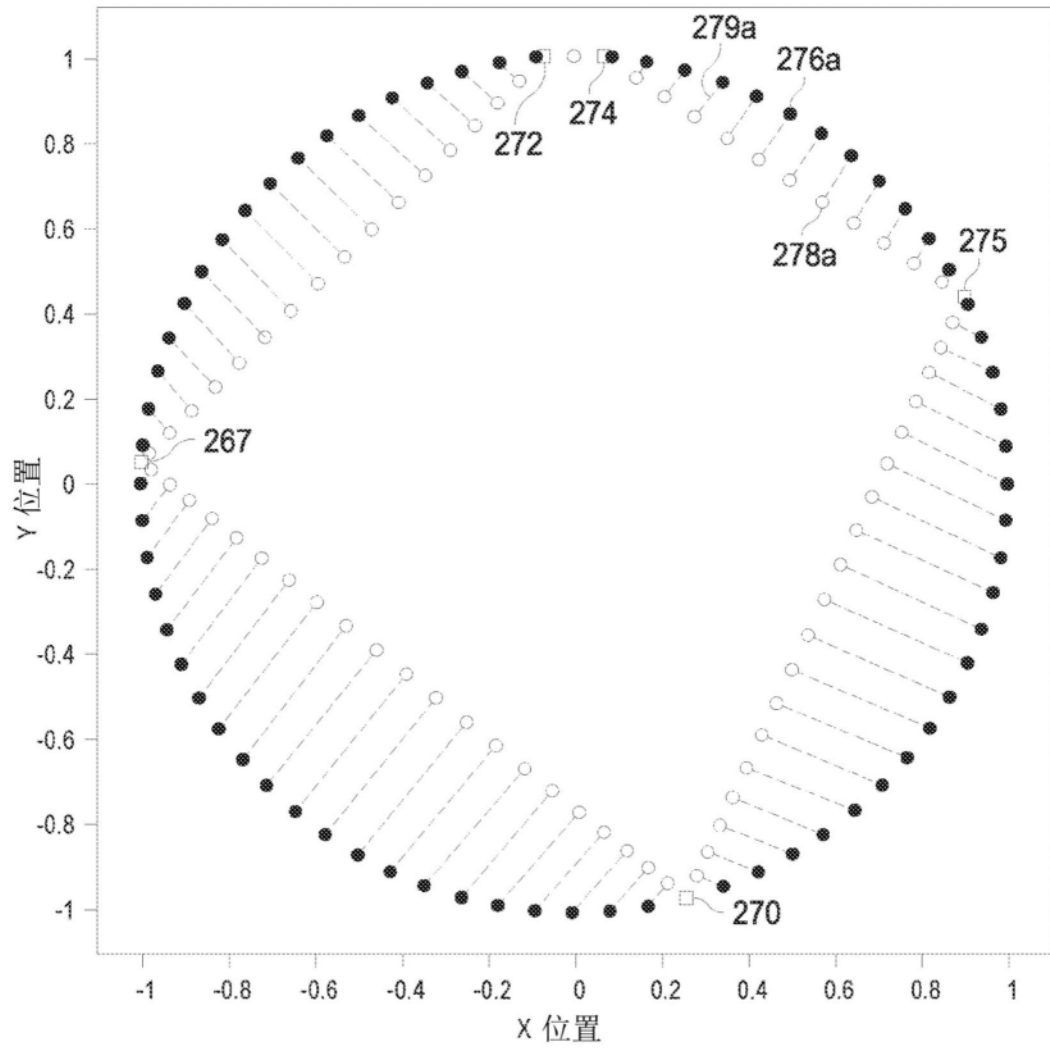


图2D

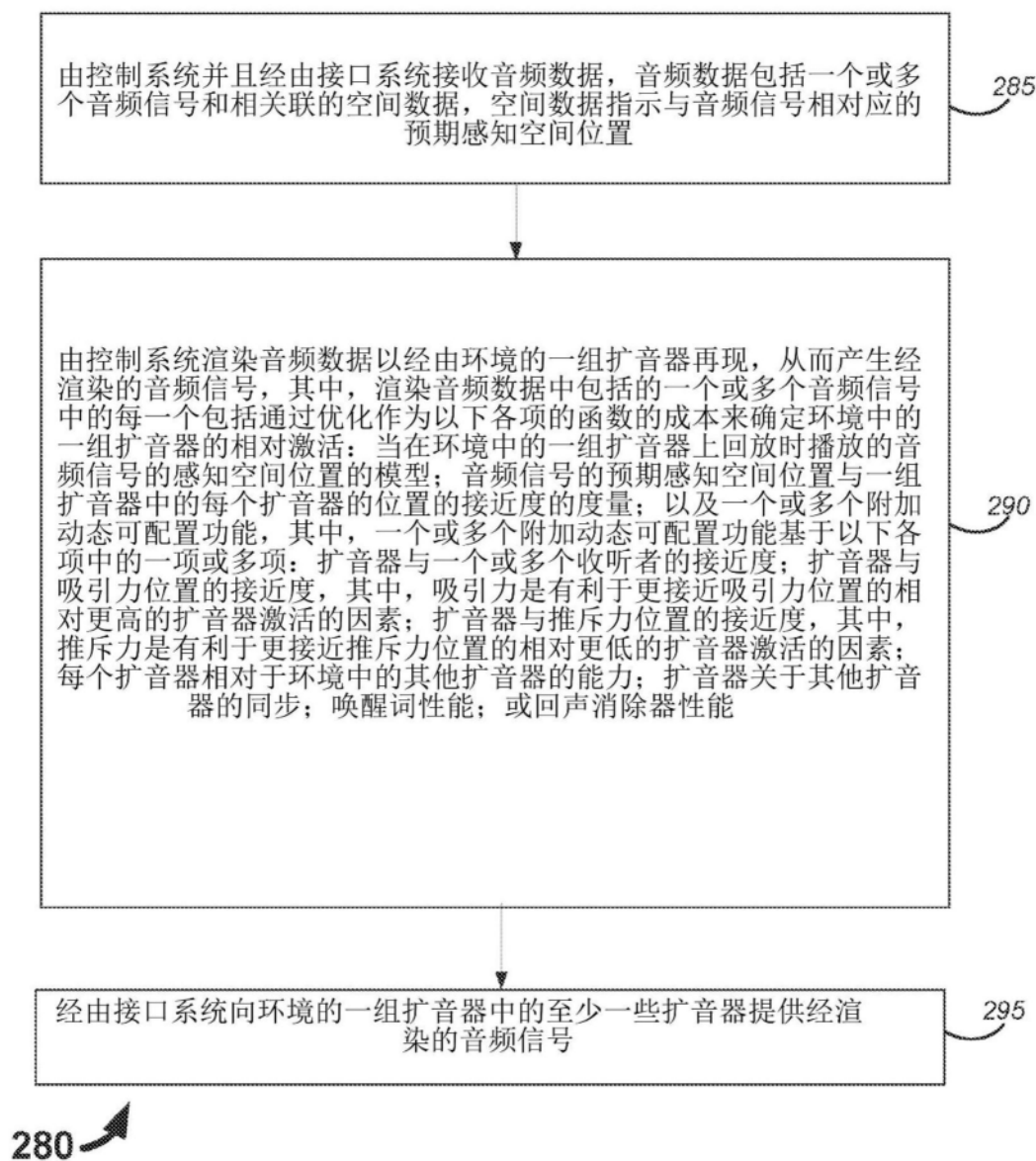


图2E

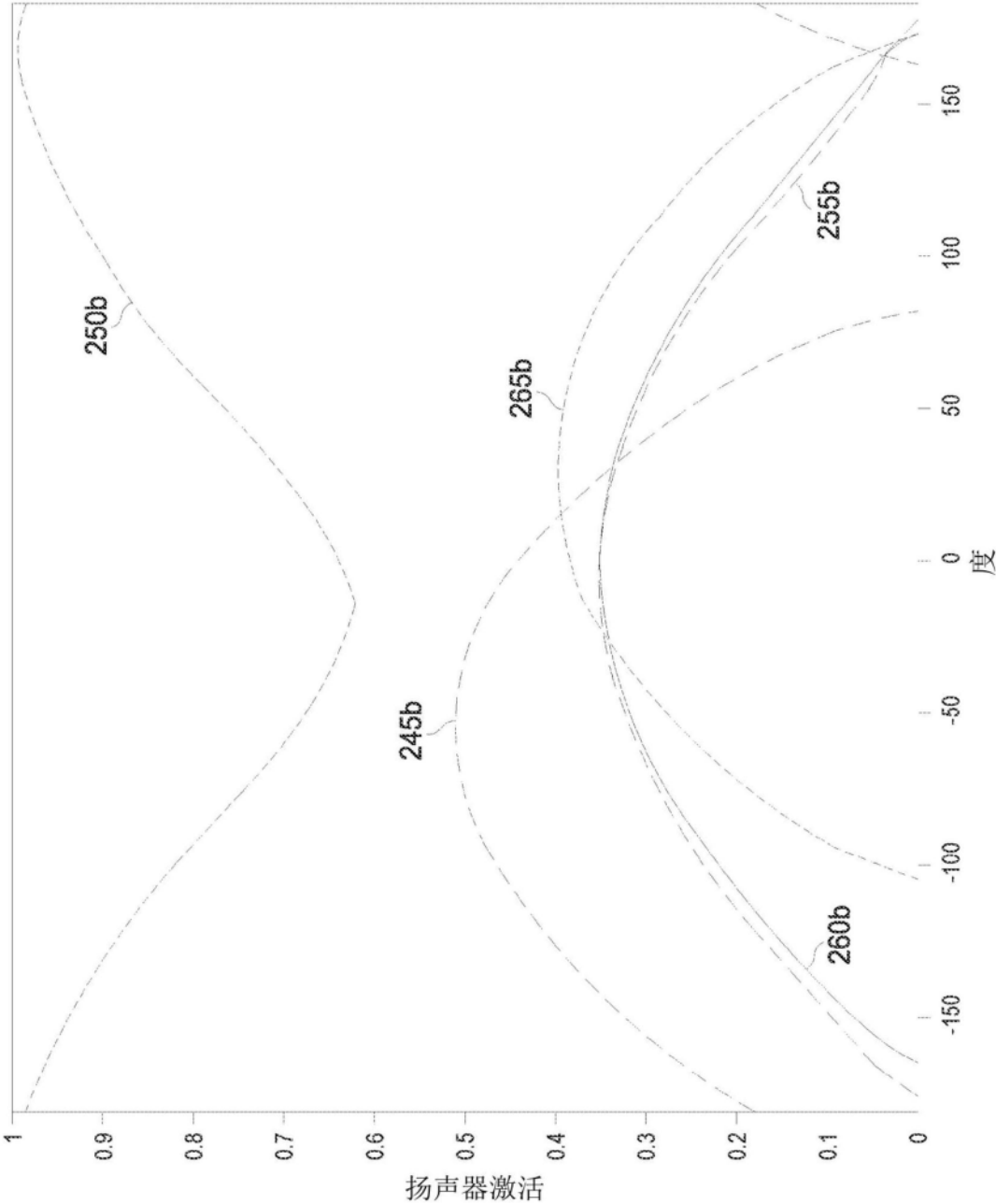


图2F

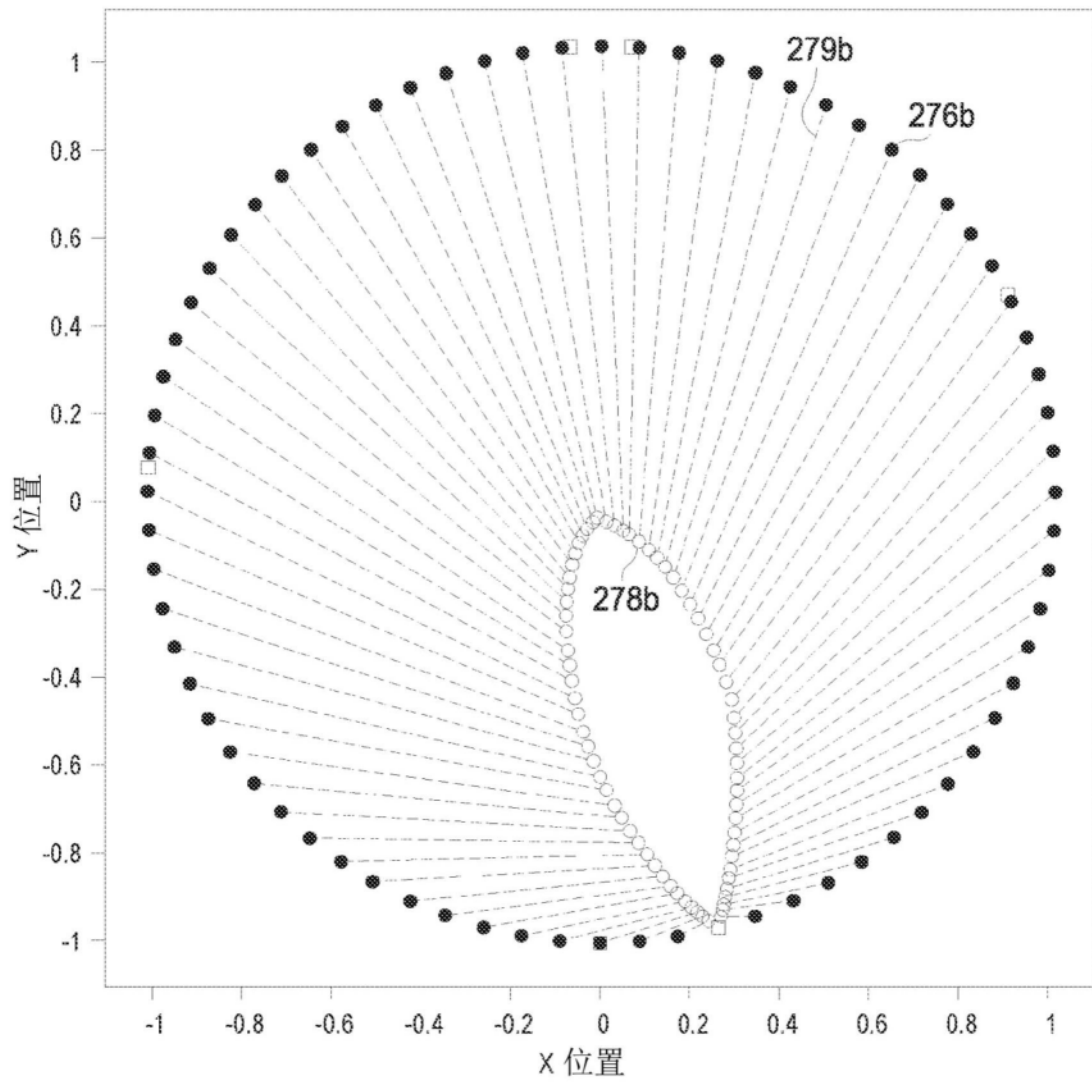


图2G

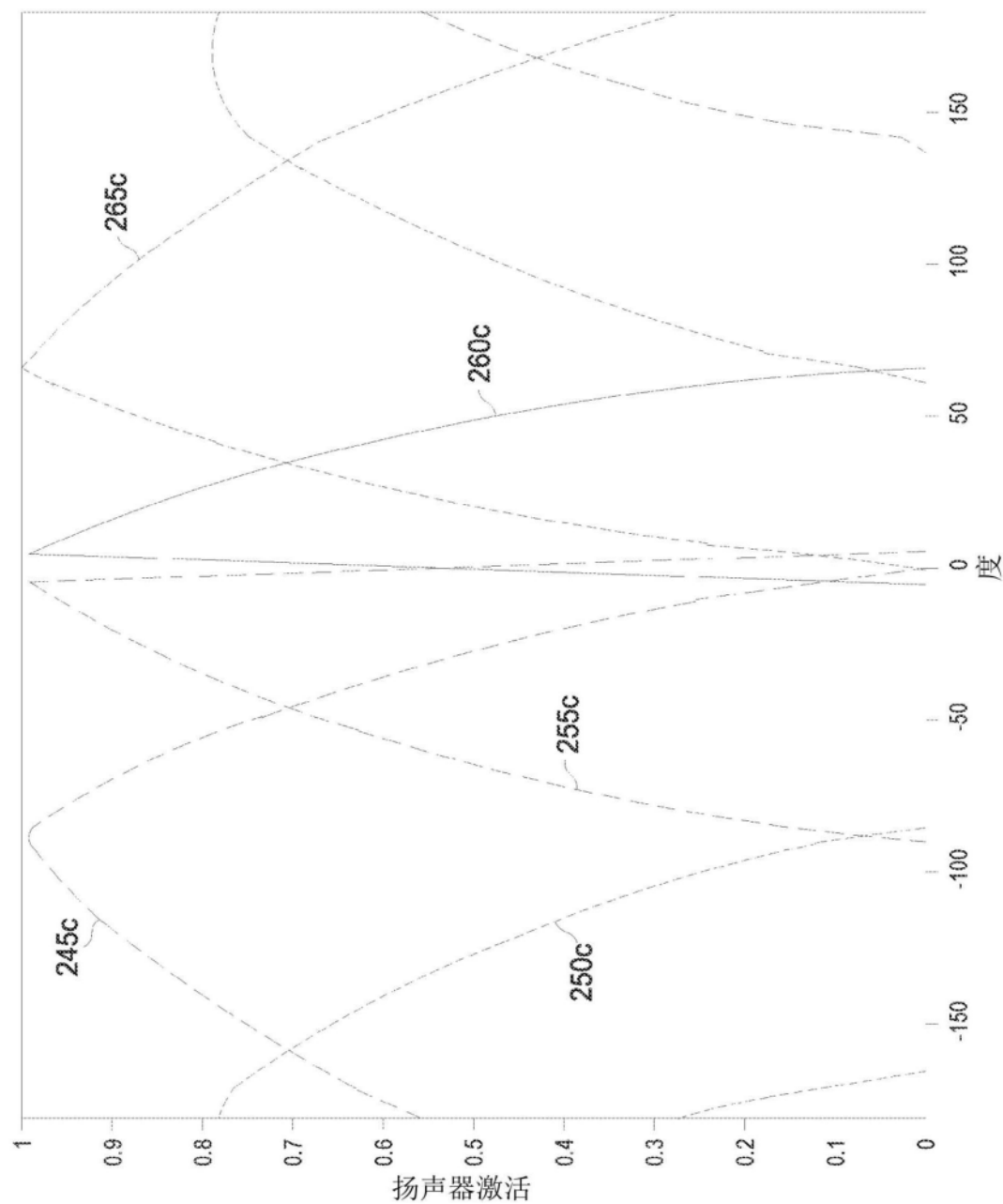


图2H

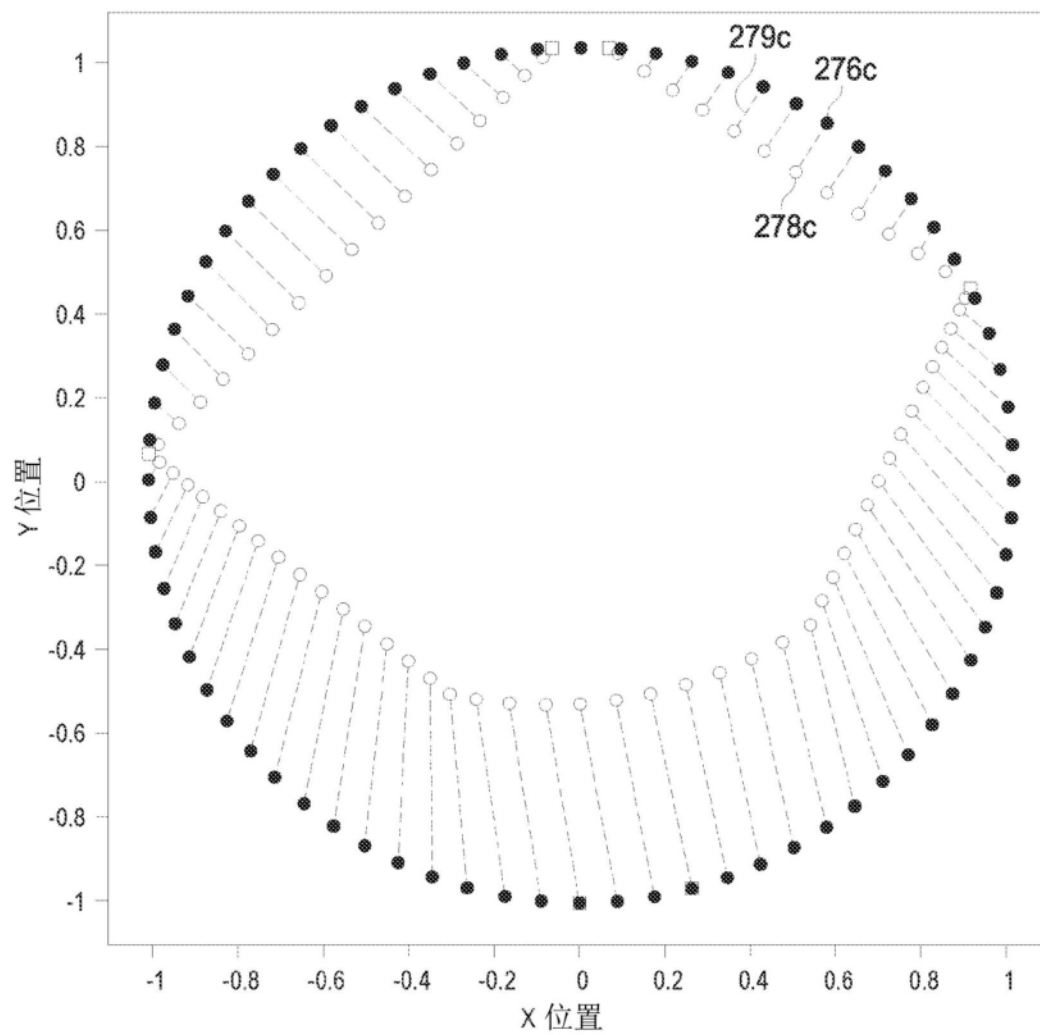


图2I

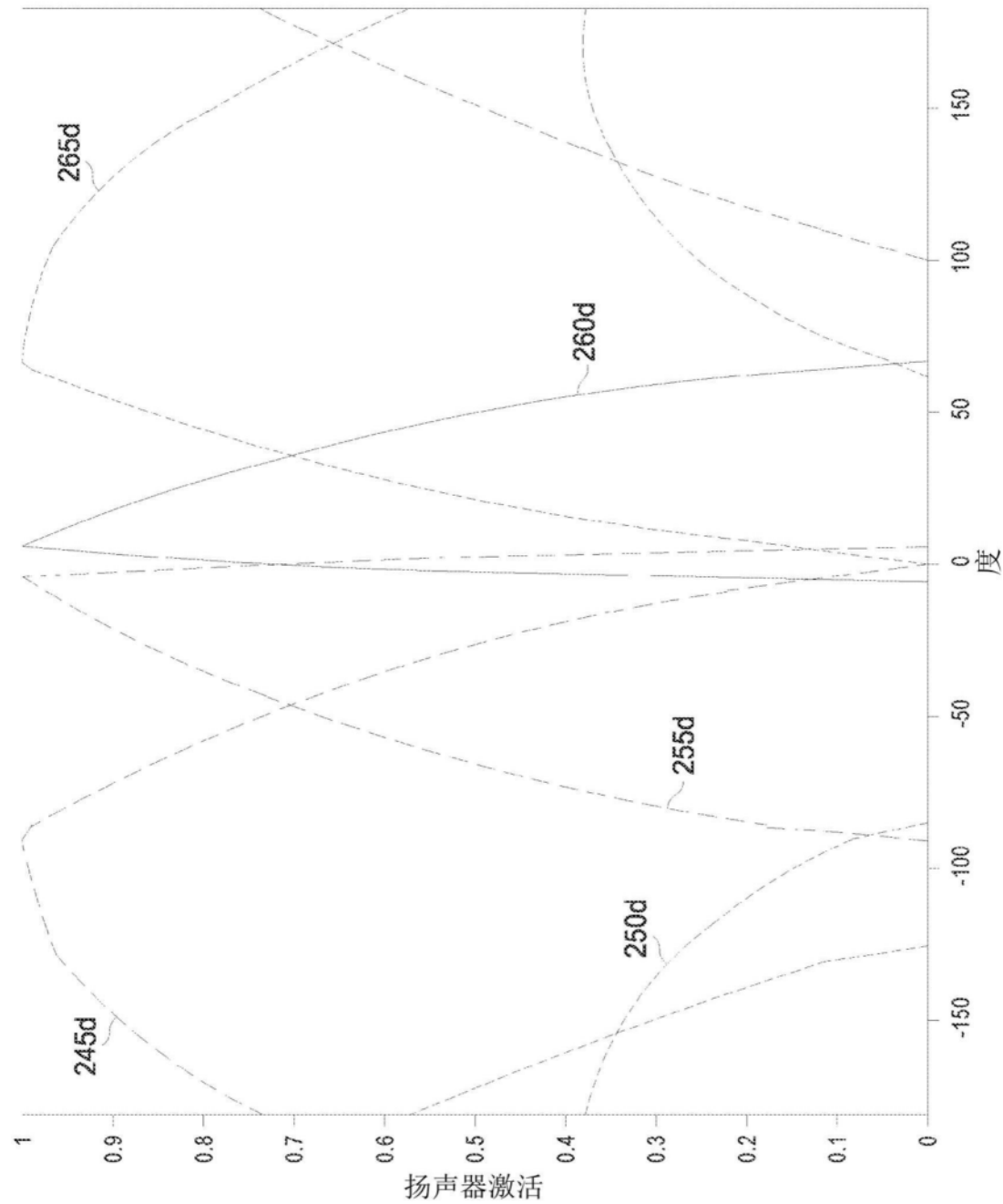


图2J

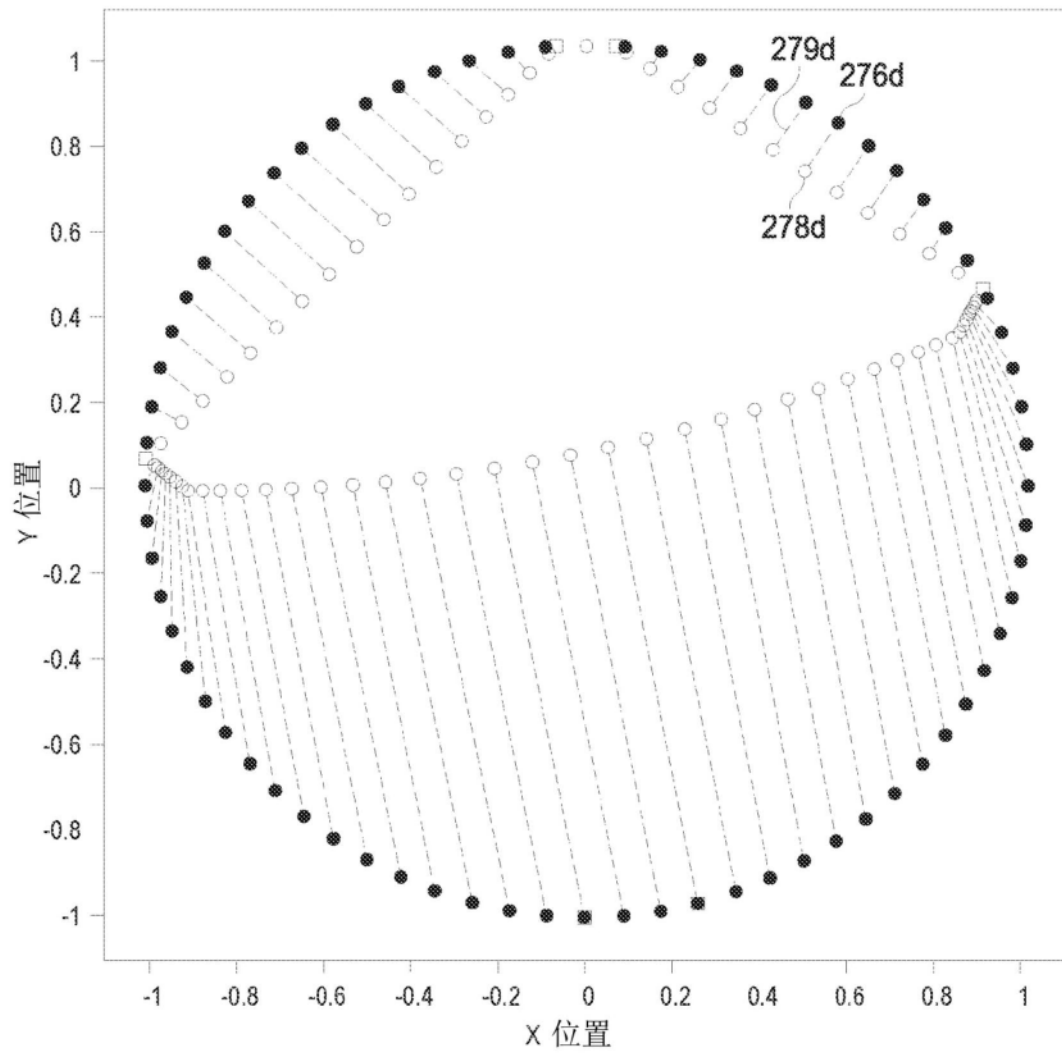


图2K

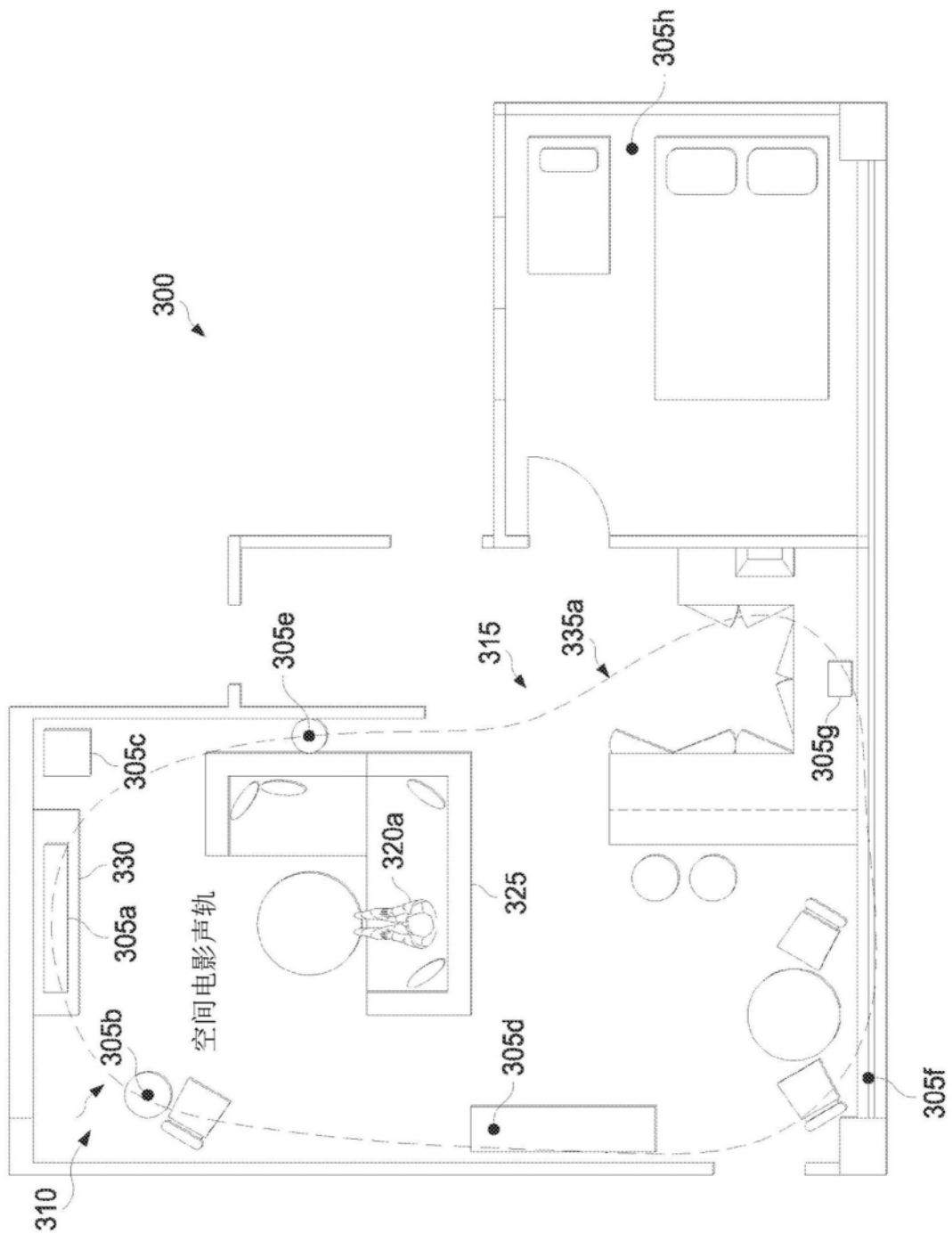


图3A

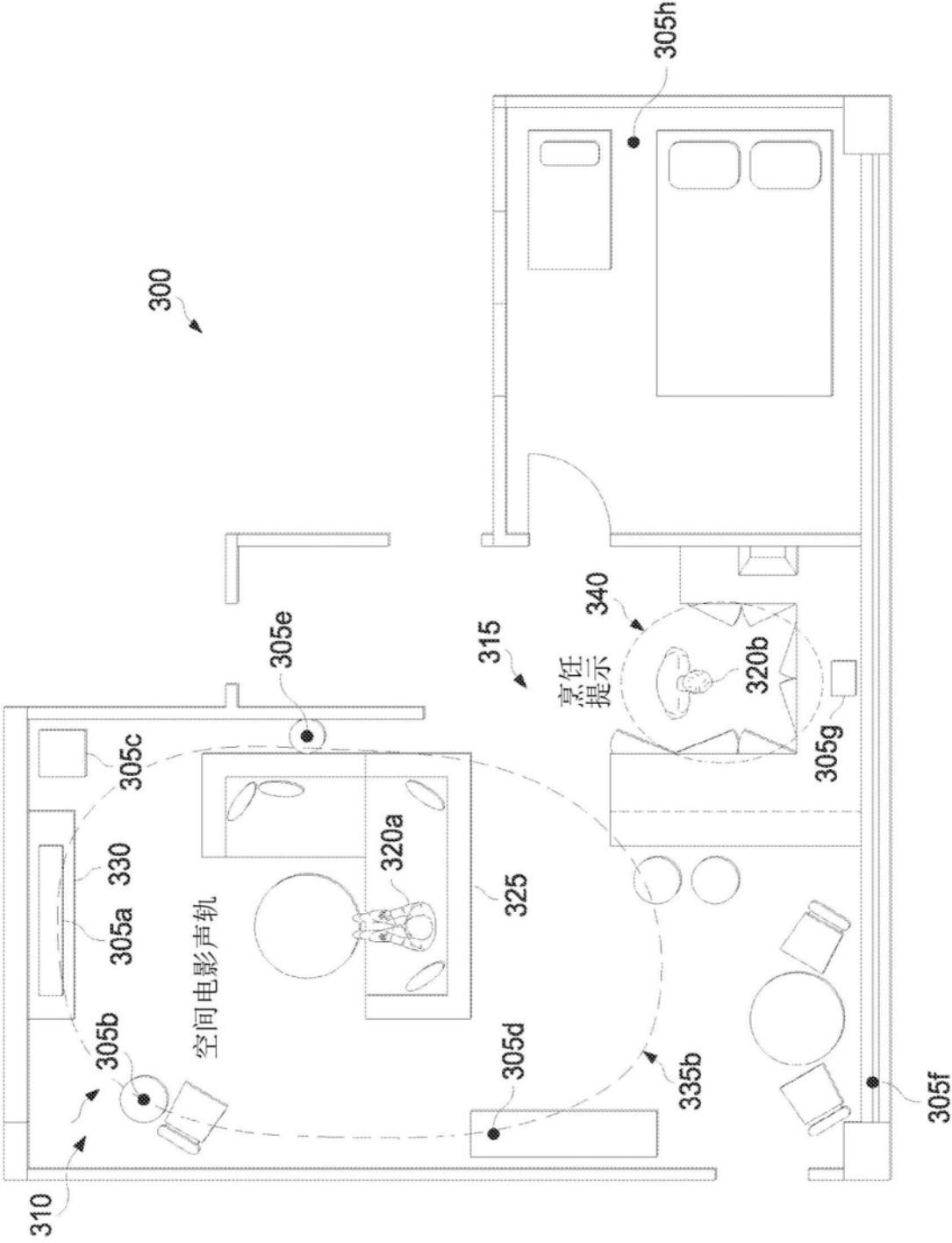


图3B

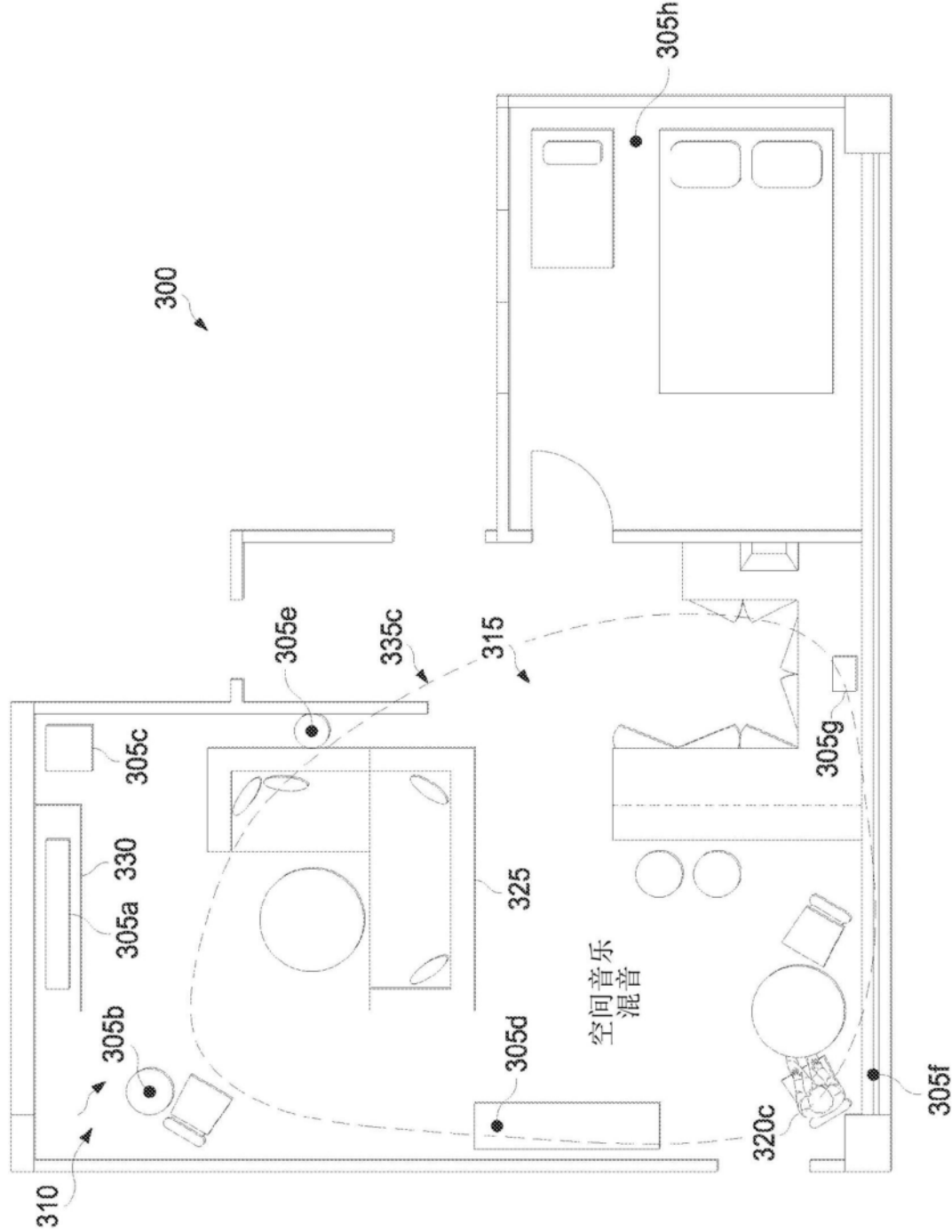


图4A

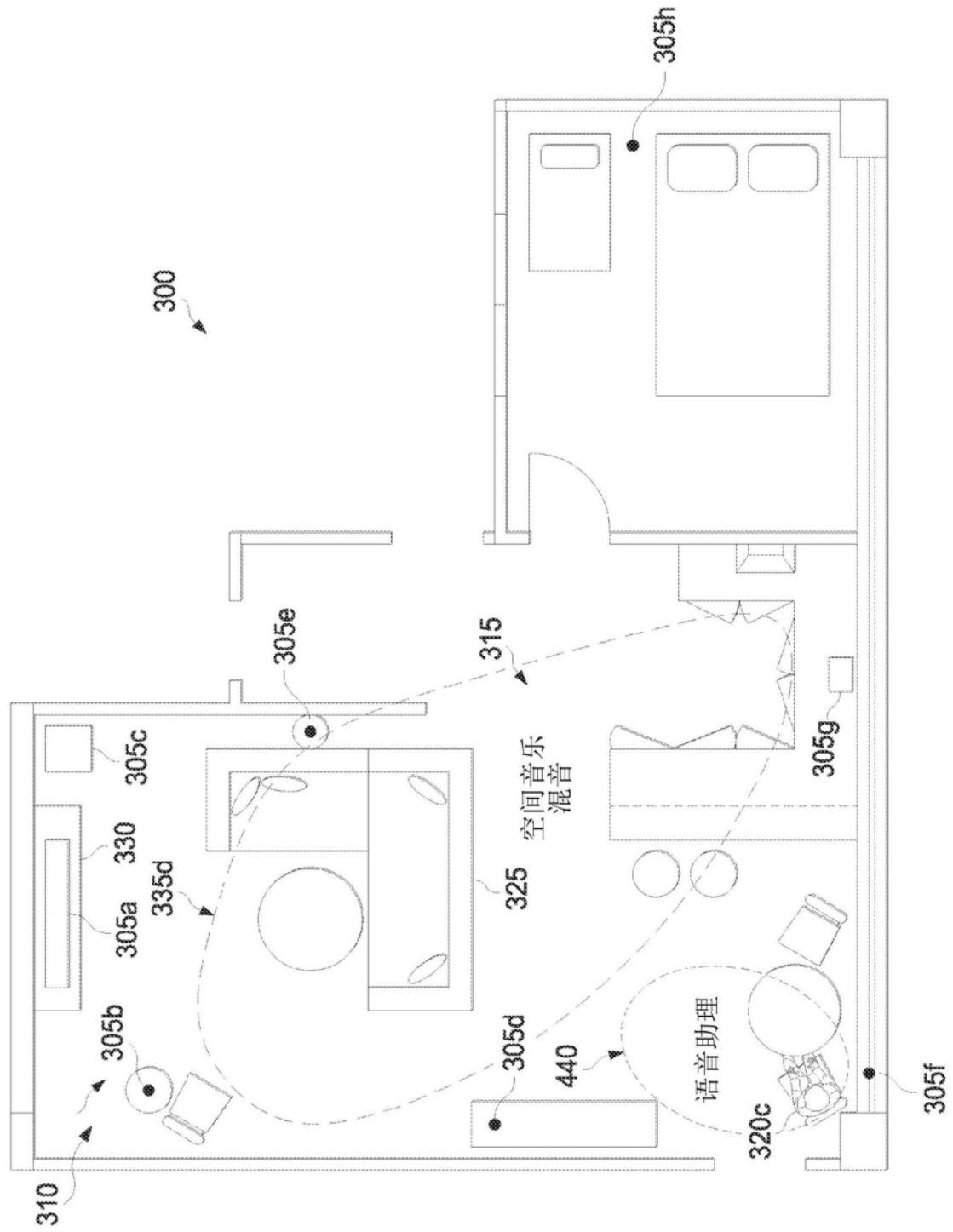


图4B

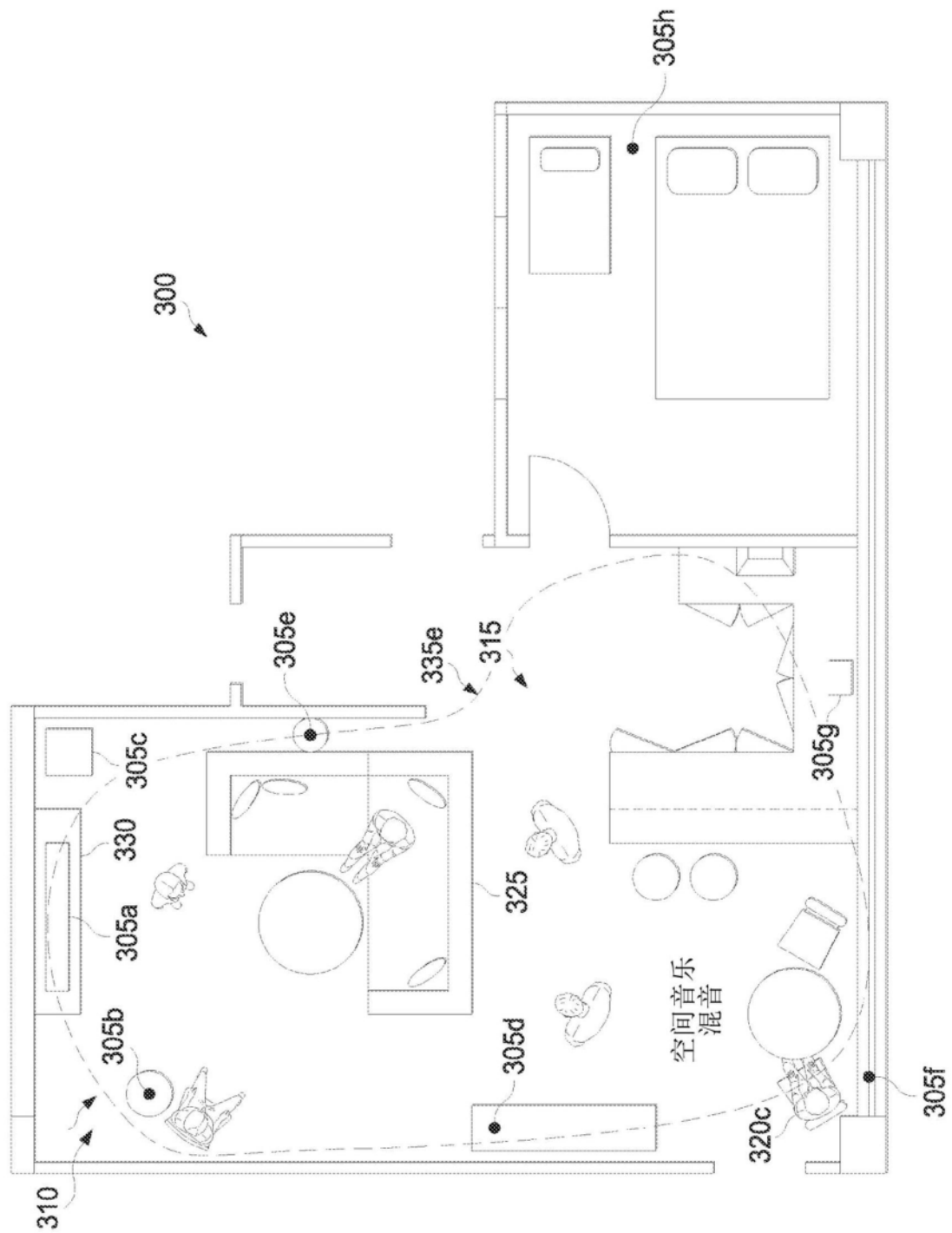


图5A

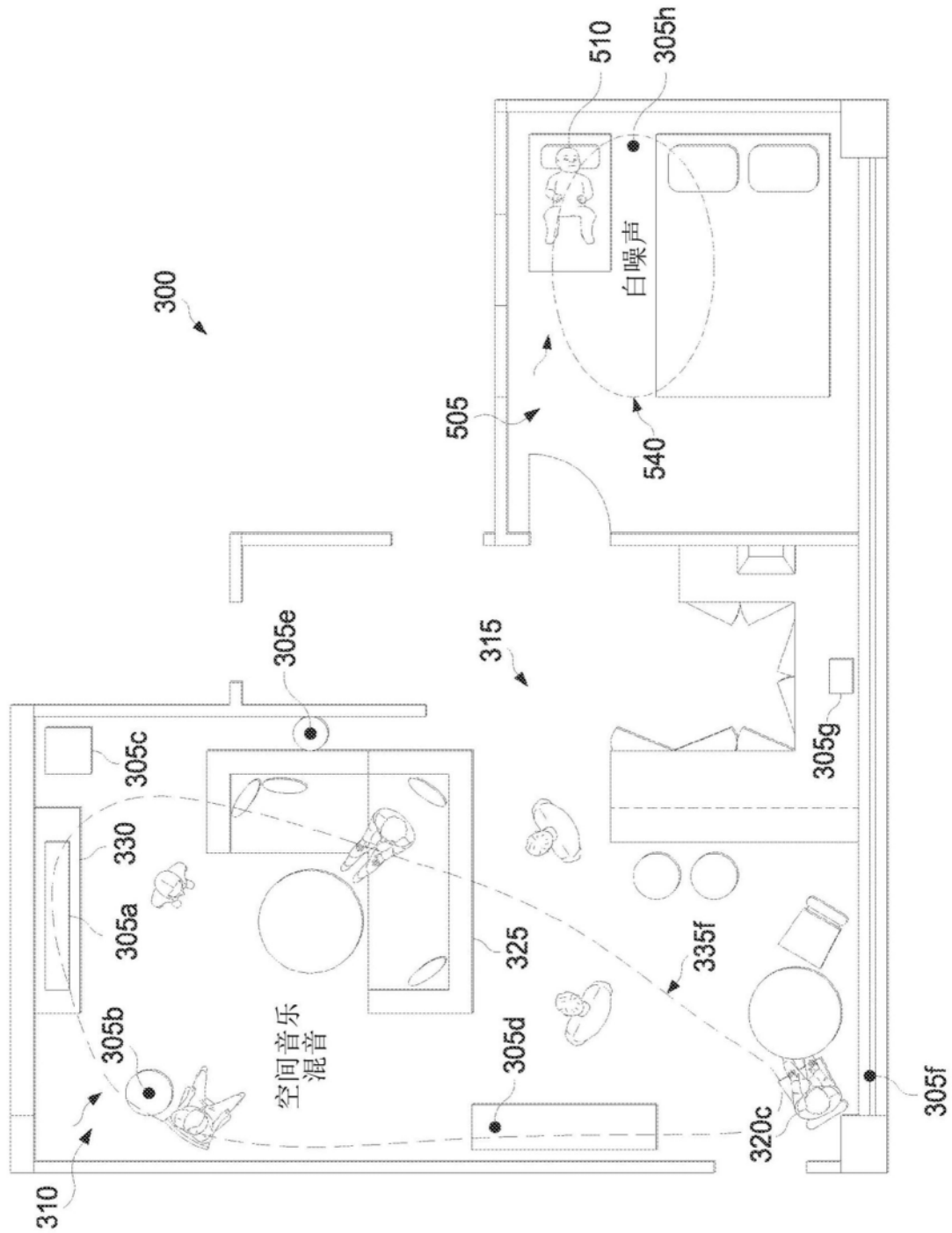


图5B

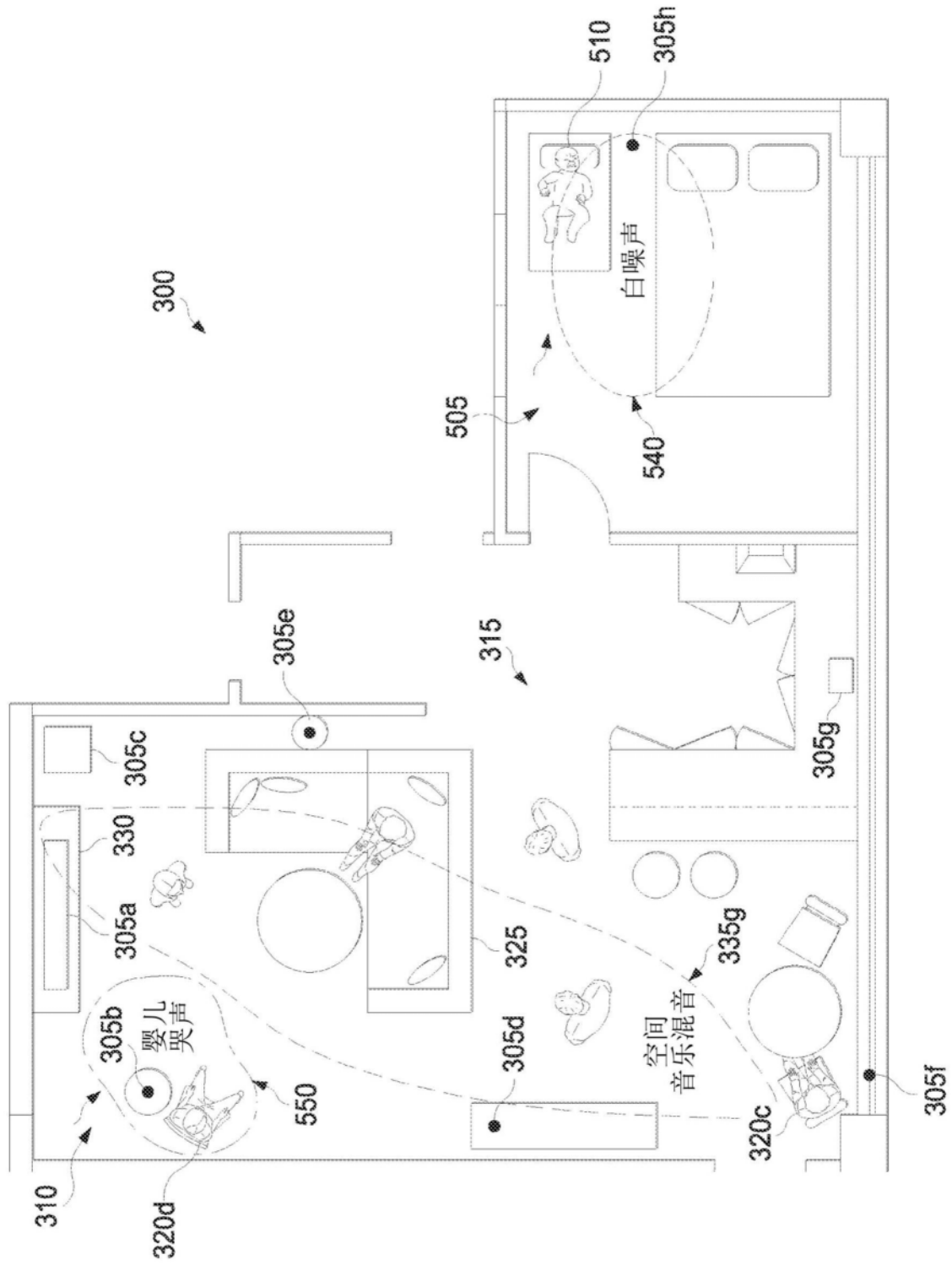


图5C

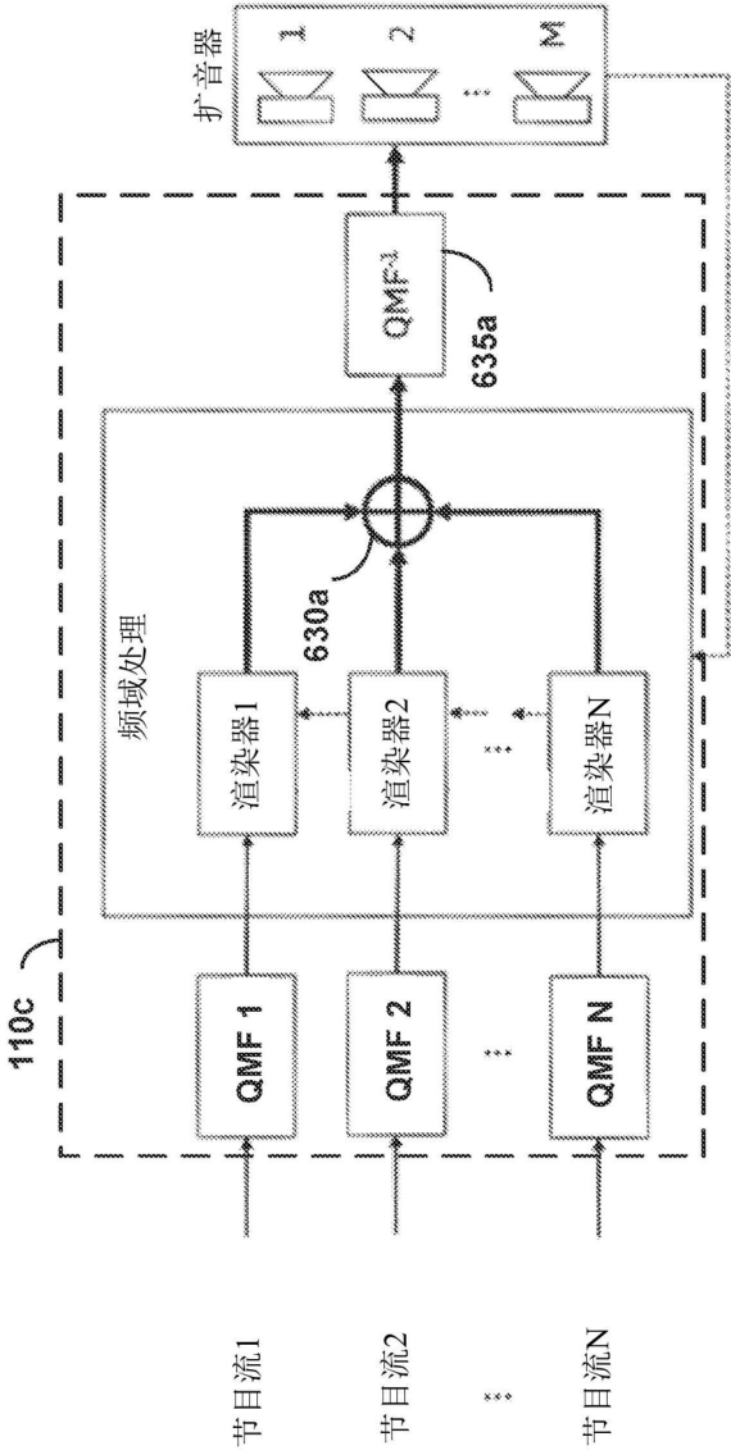


图6

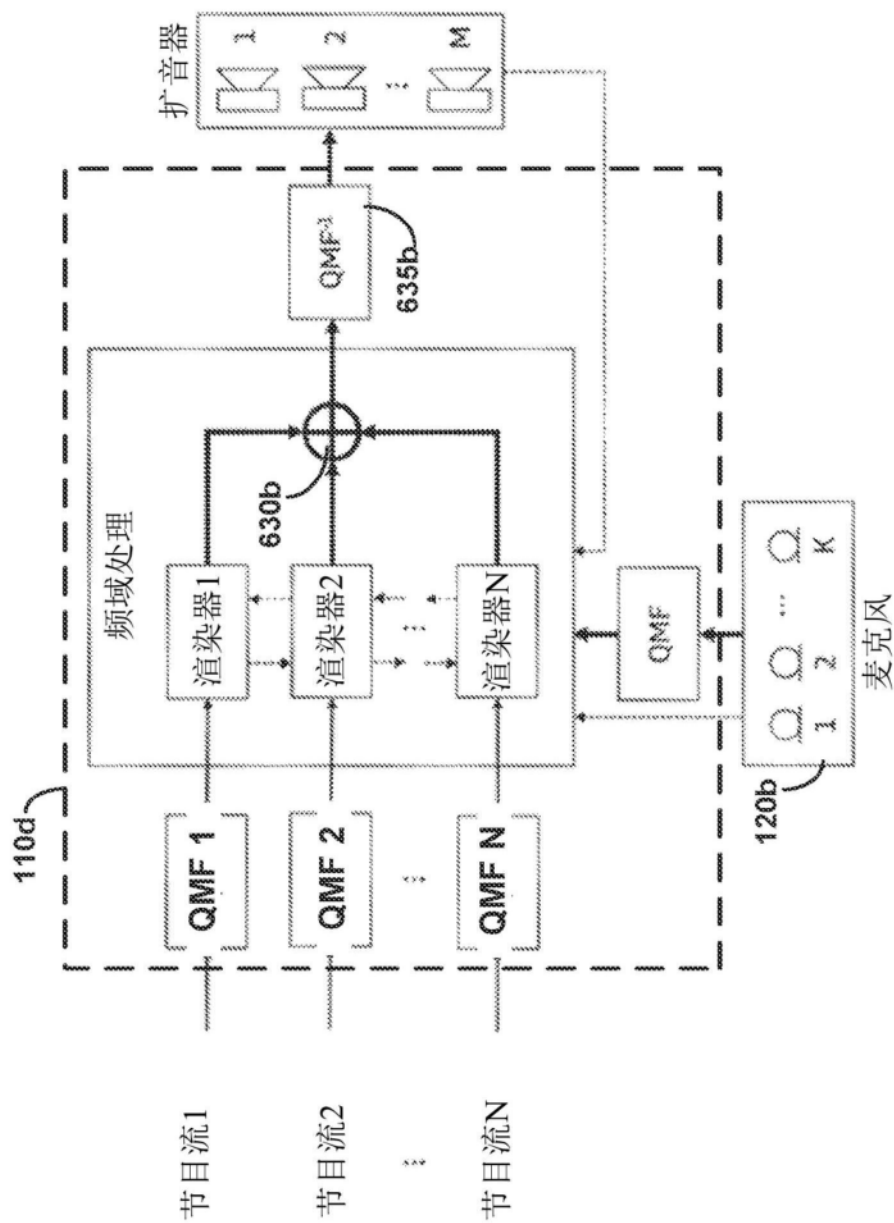


图7

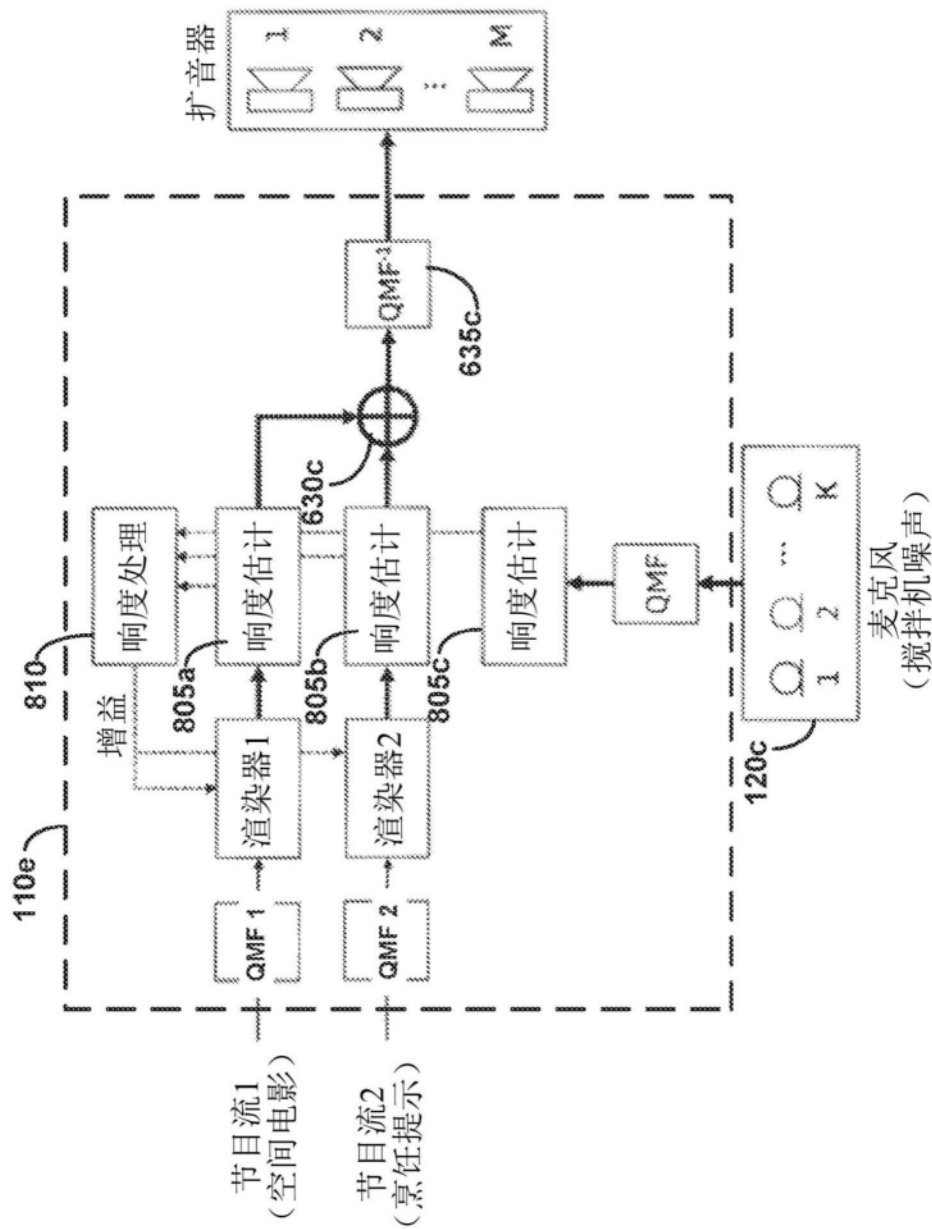


图8

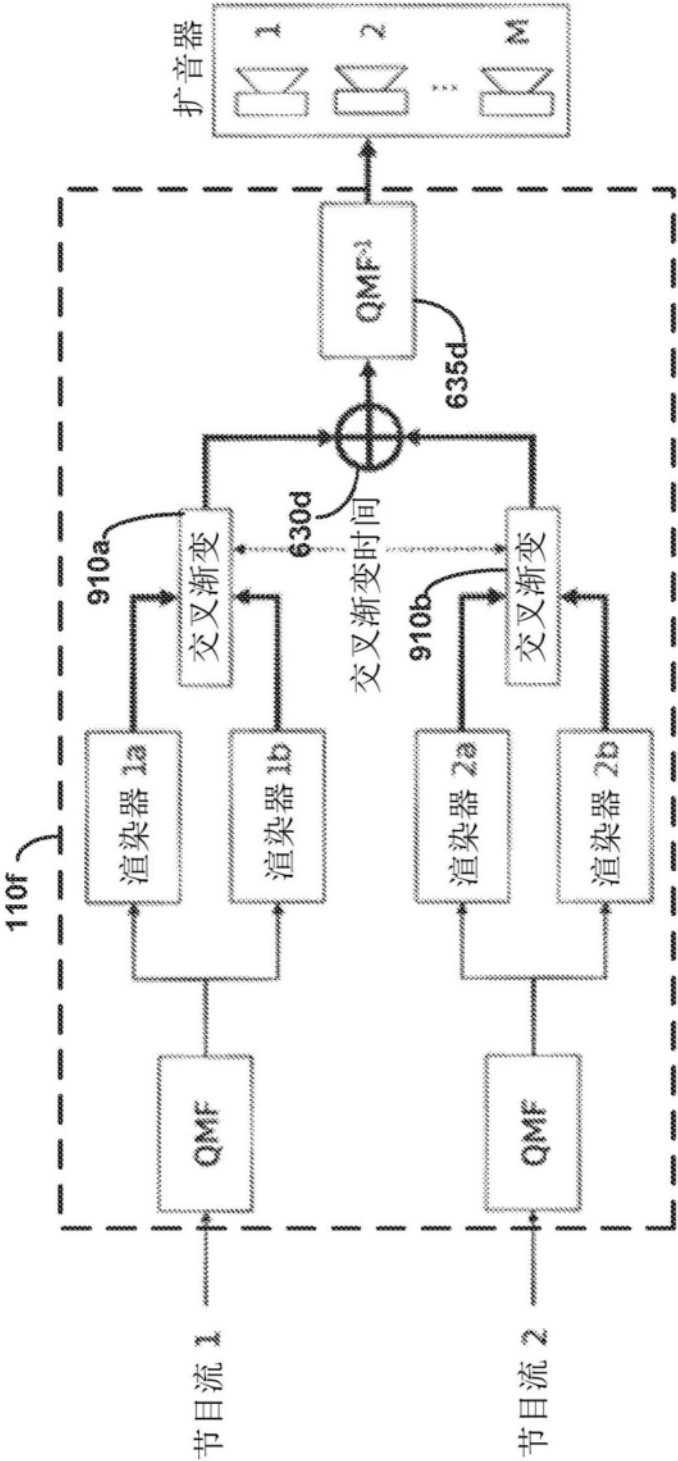


图9A

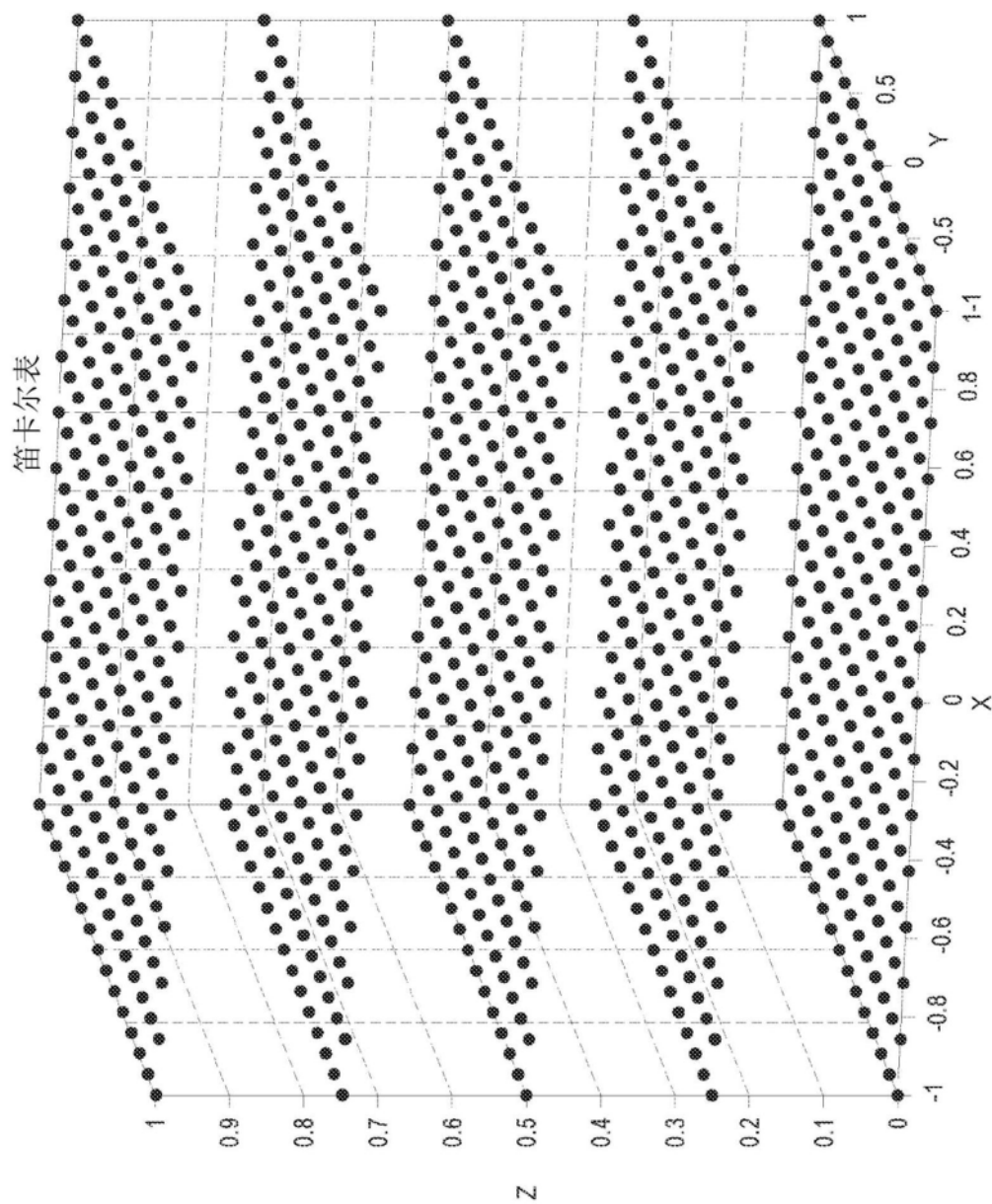


图9B

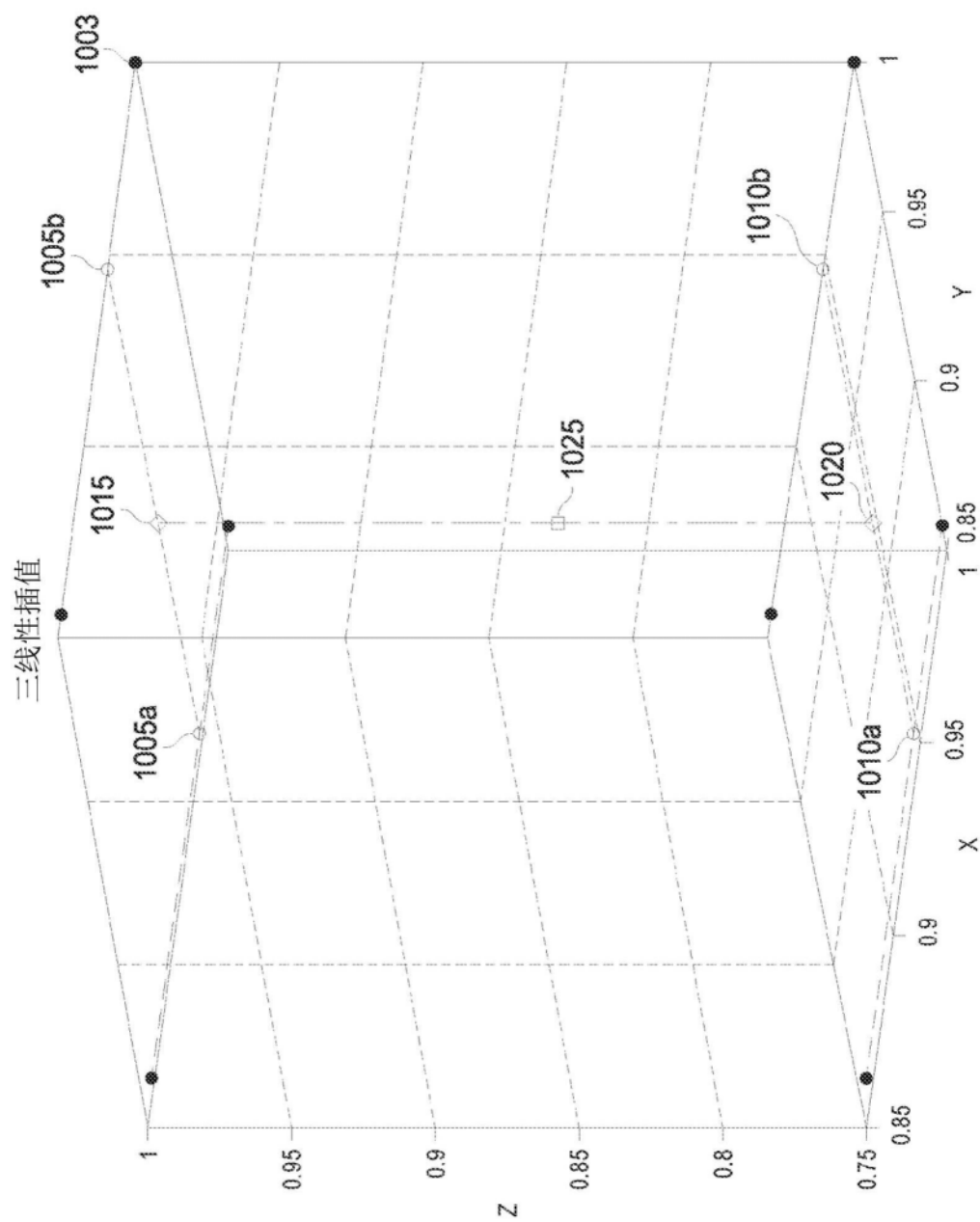


图10

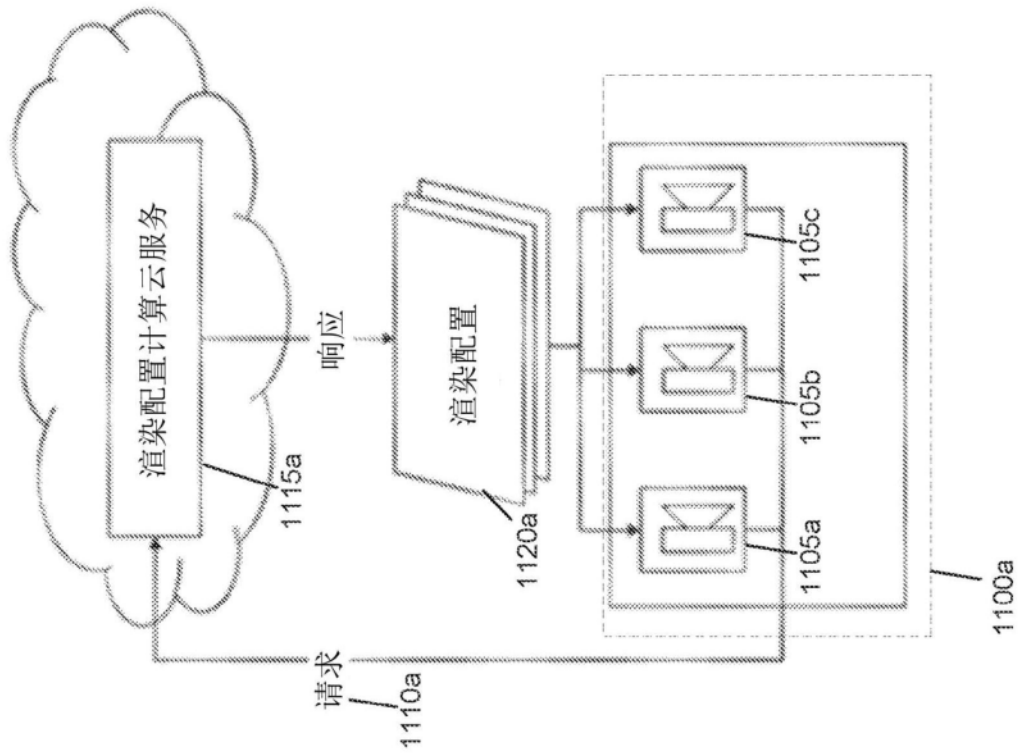


图11A

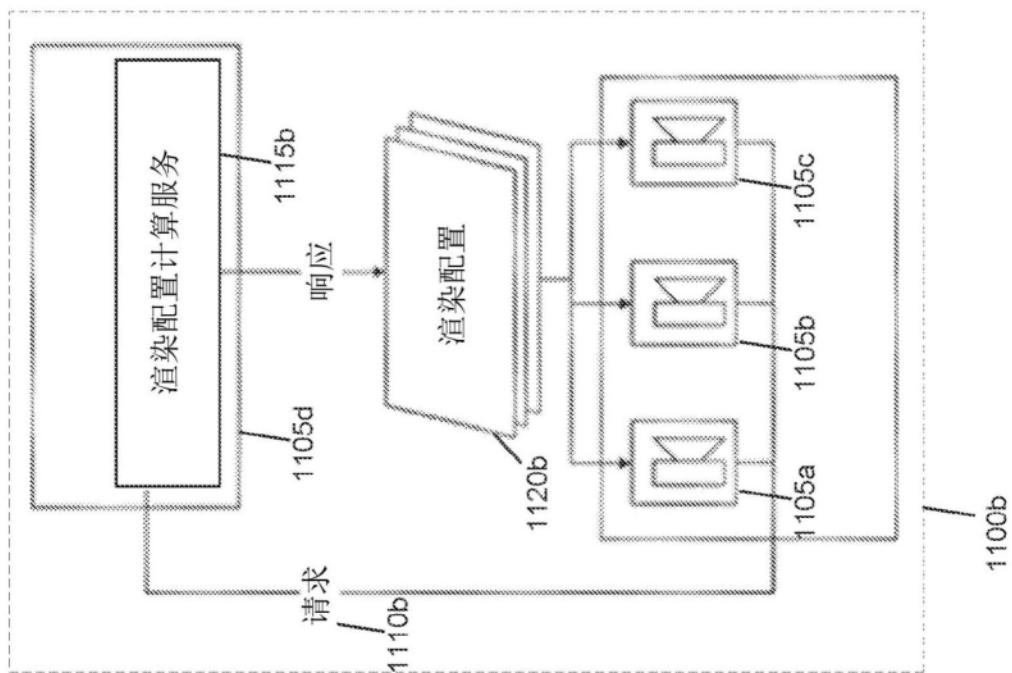


图11B

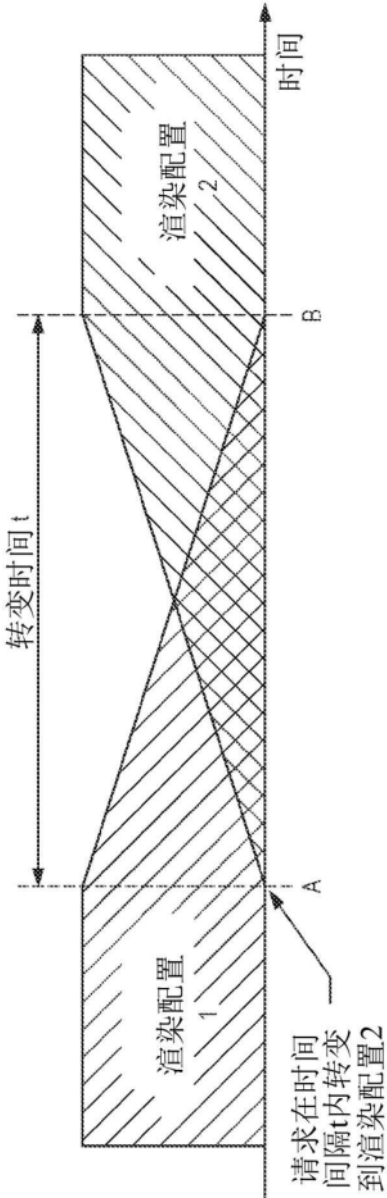


图12A

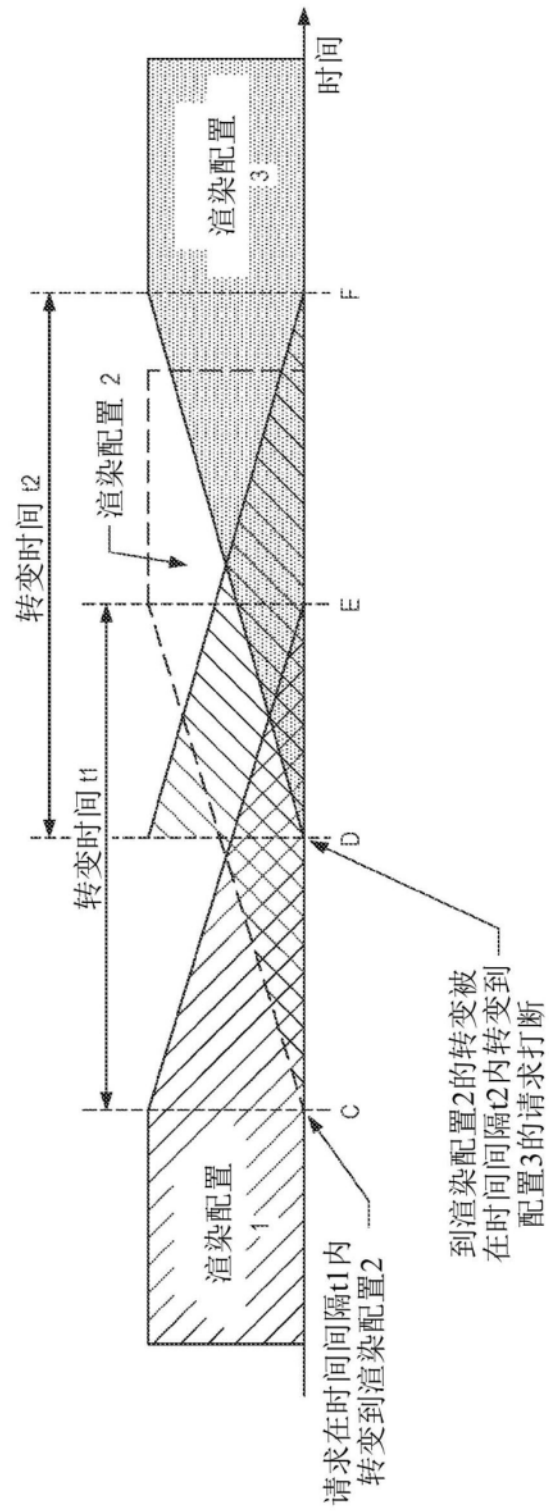


图12B

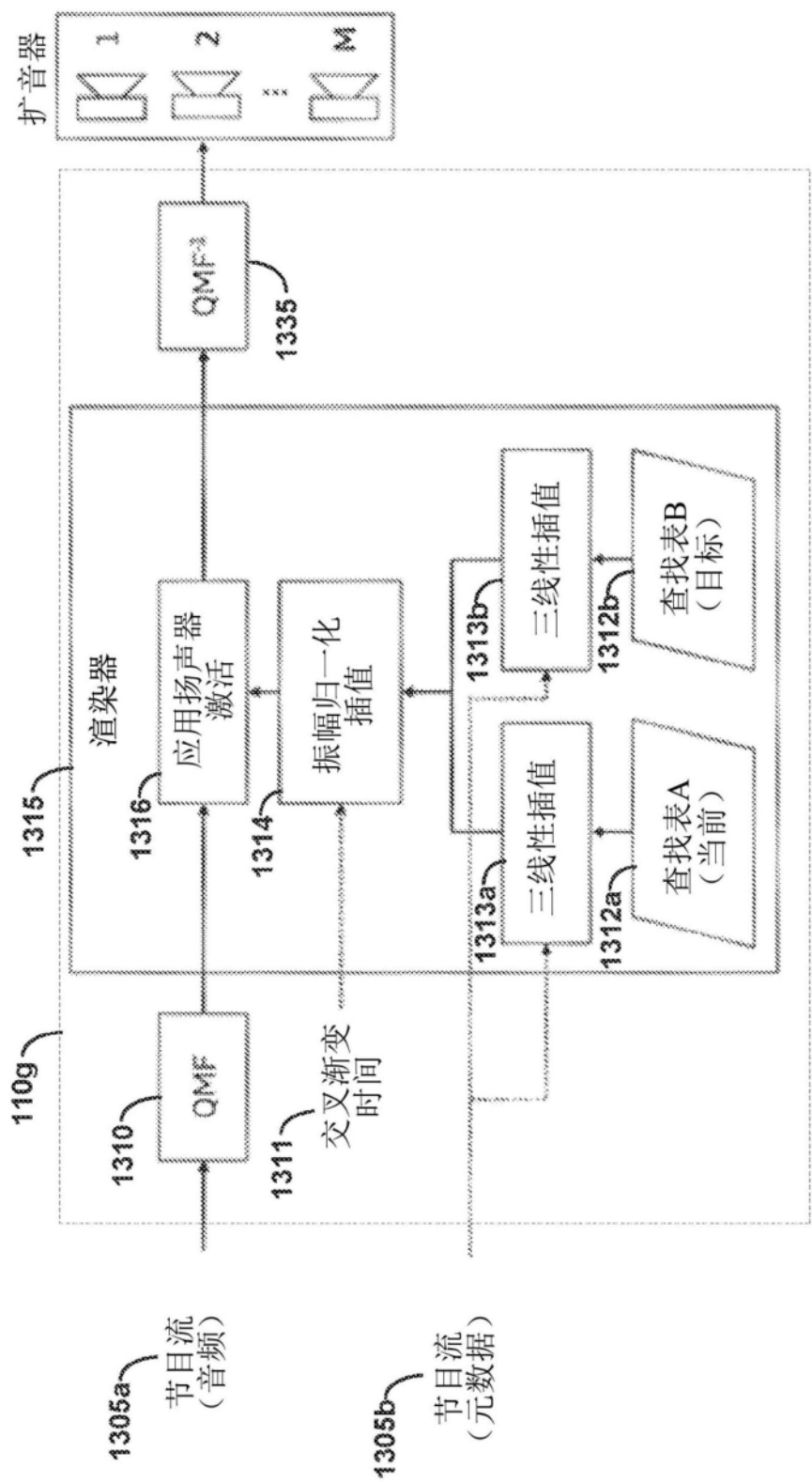


图13

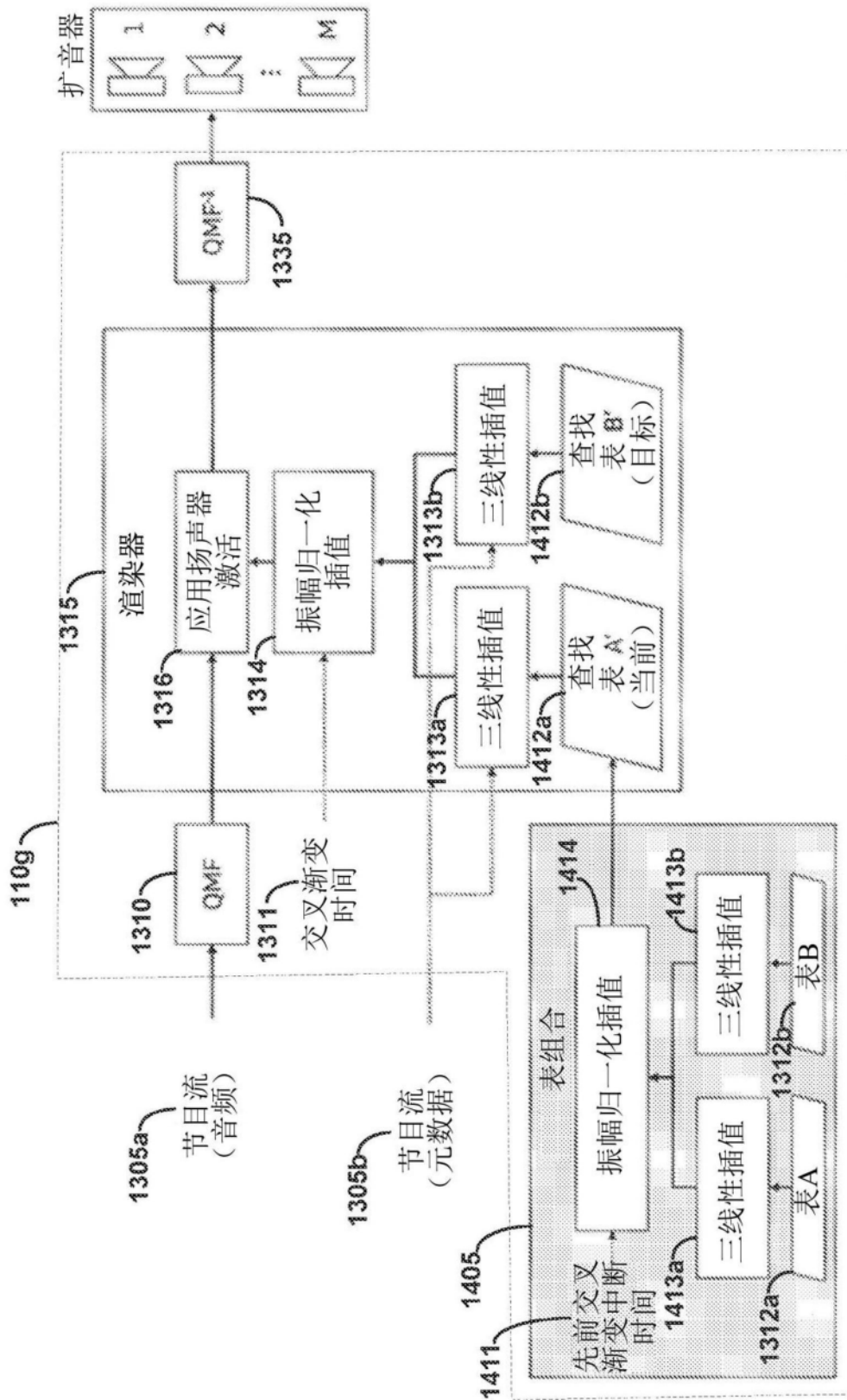


图14

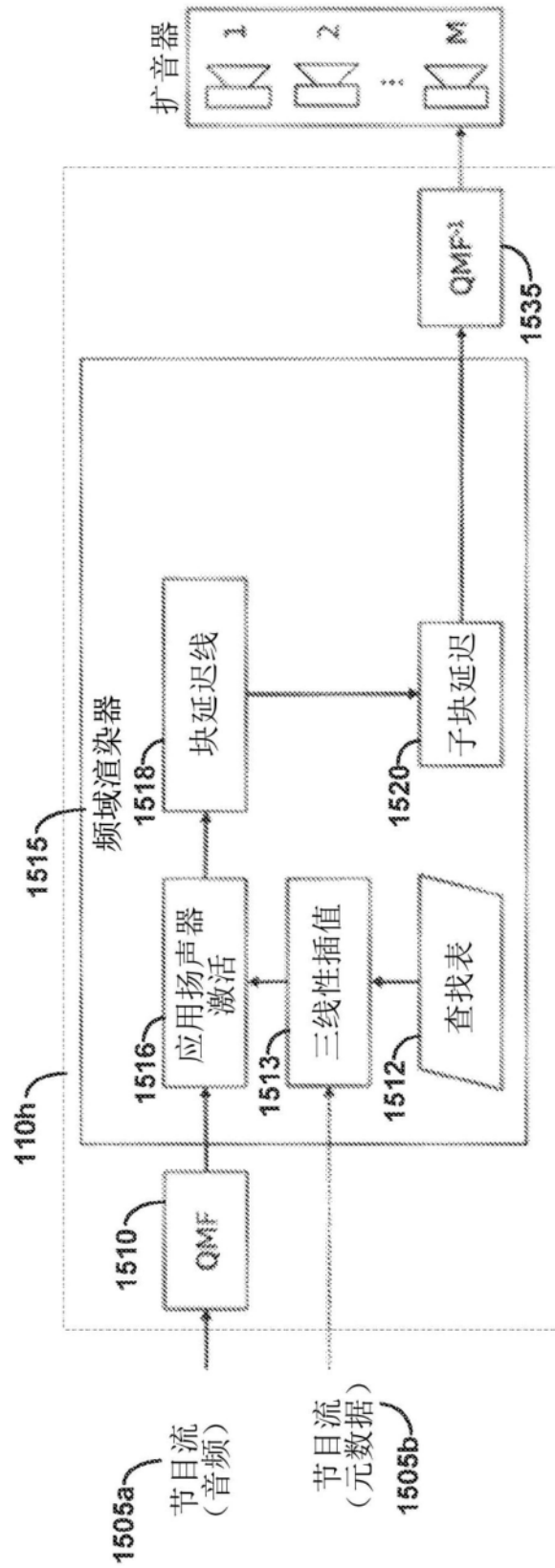


图15

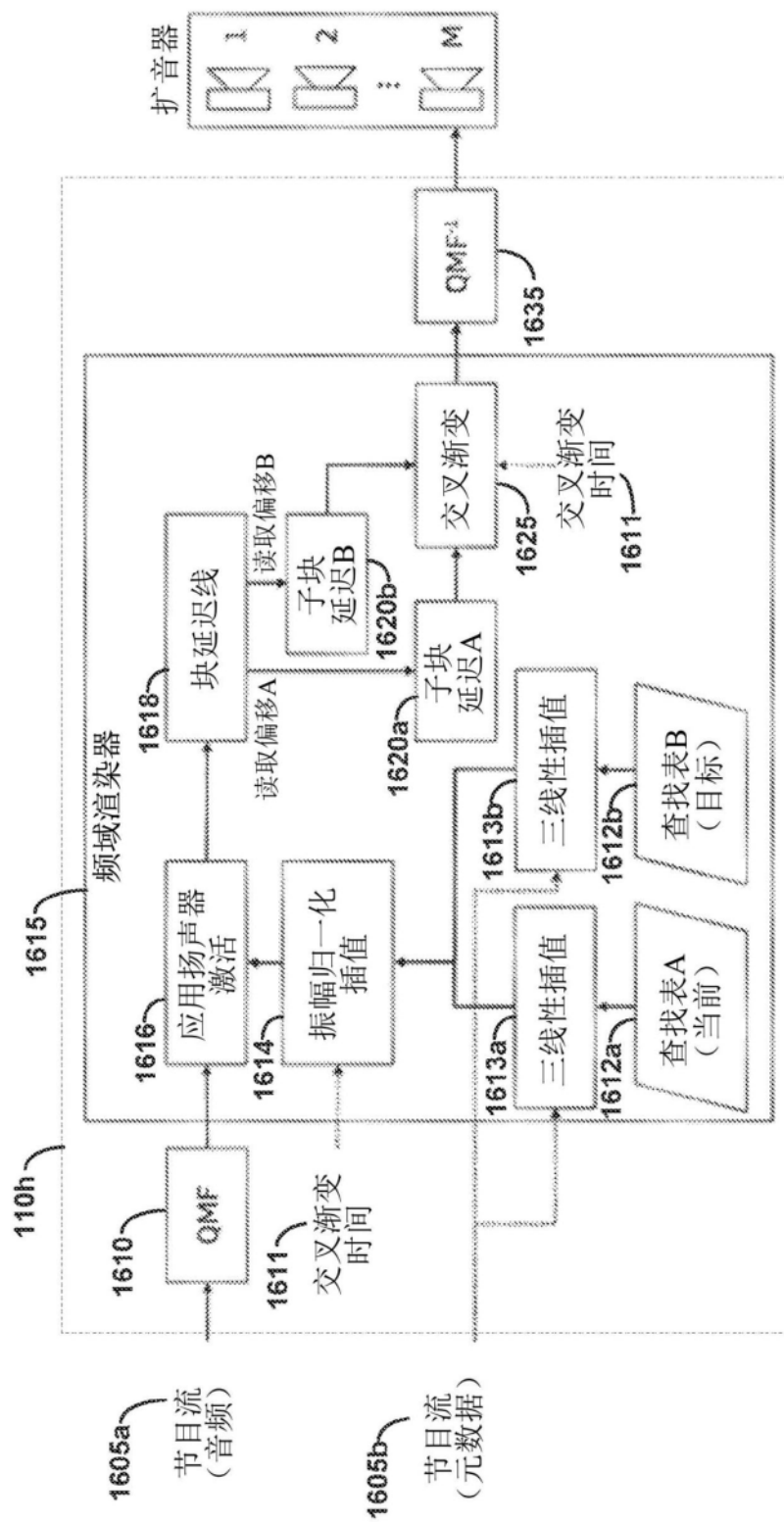


图16

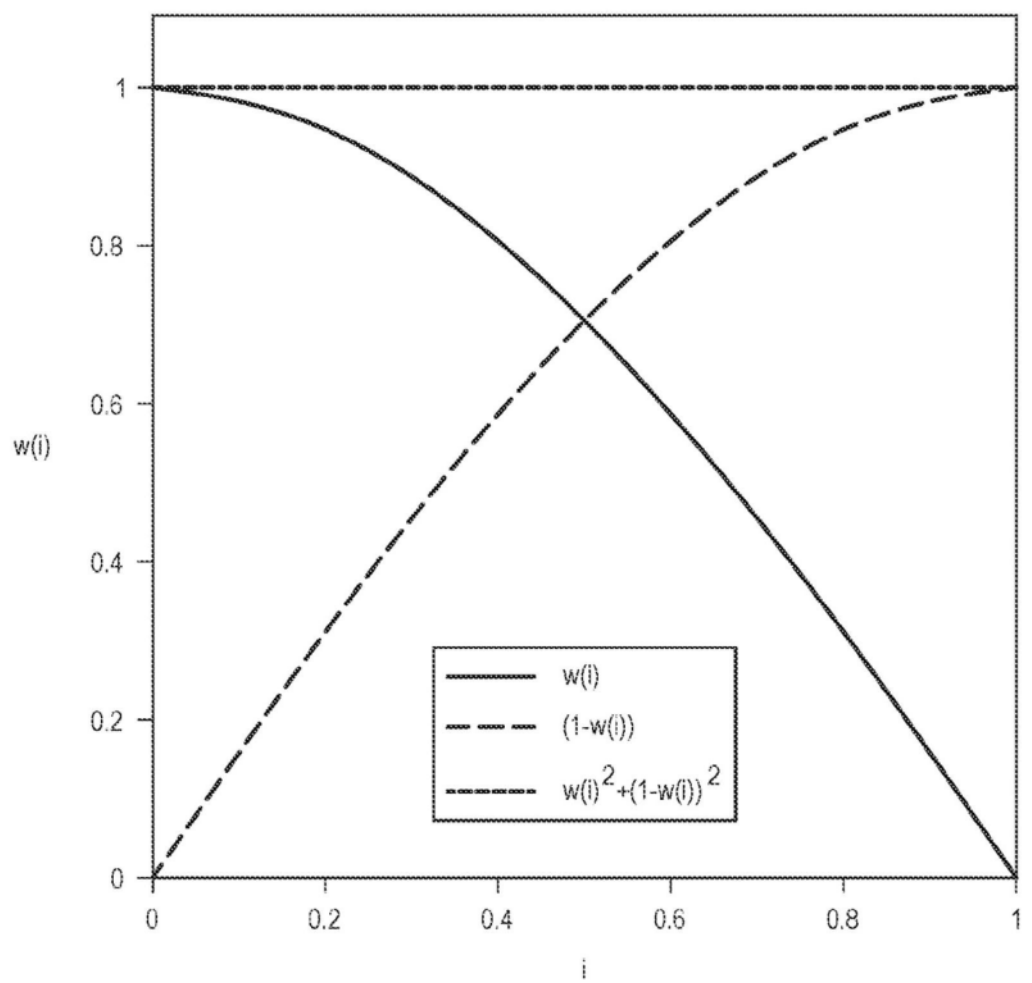


图17

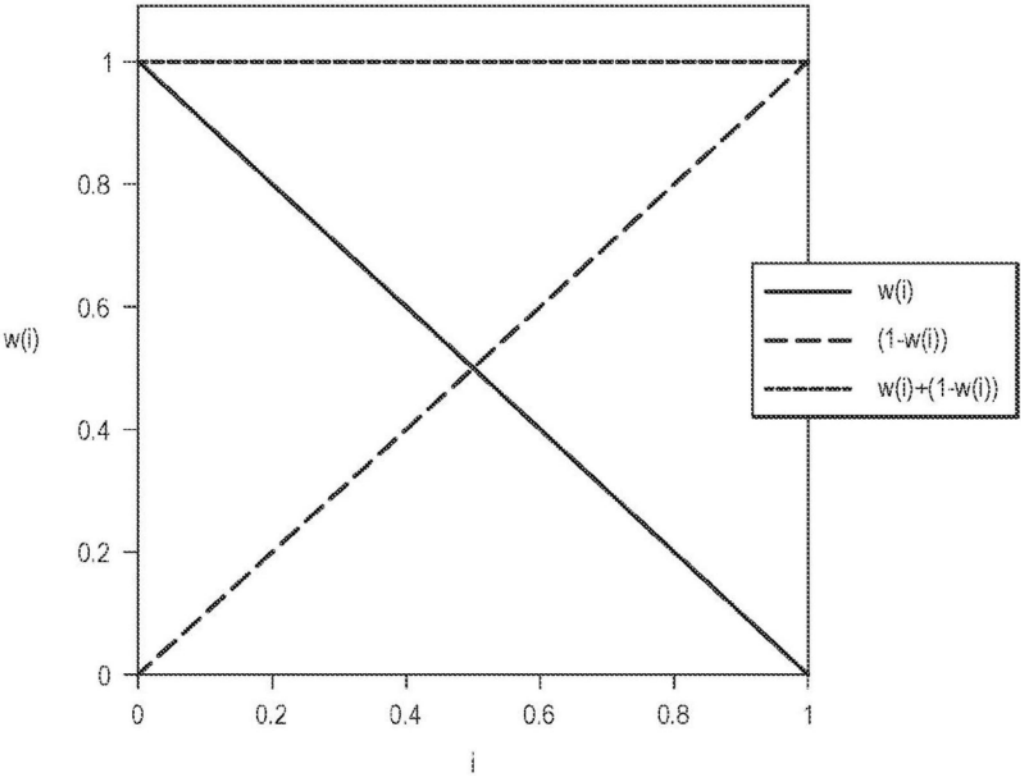


图18A

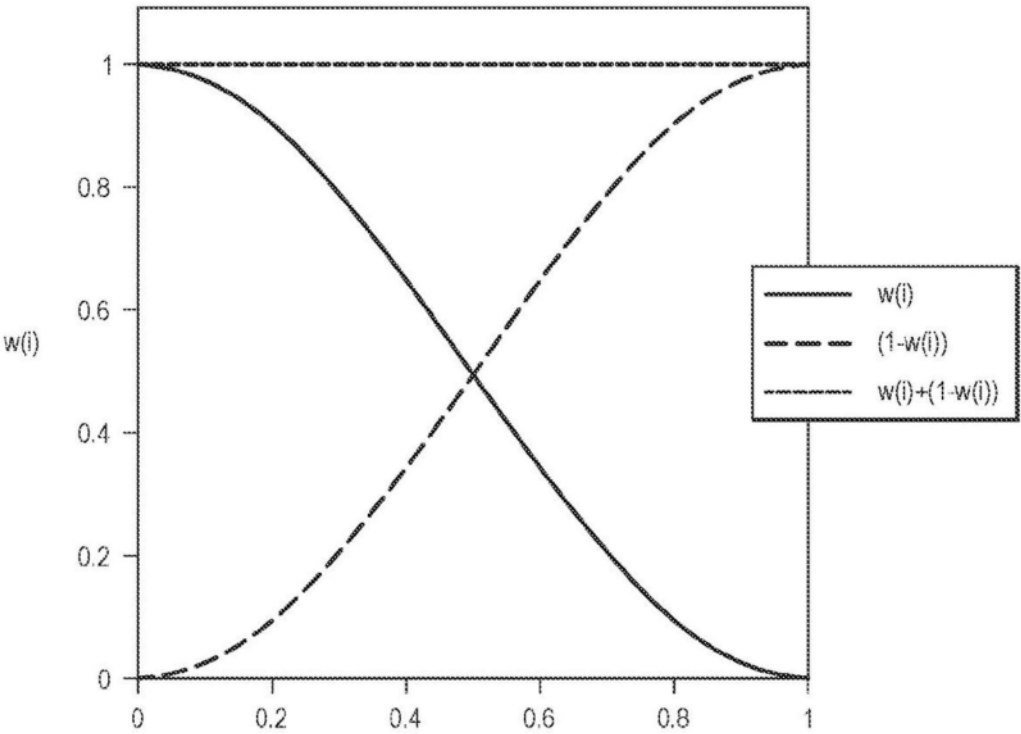


图18B

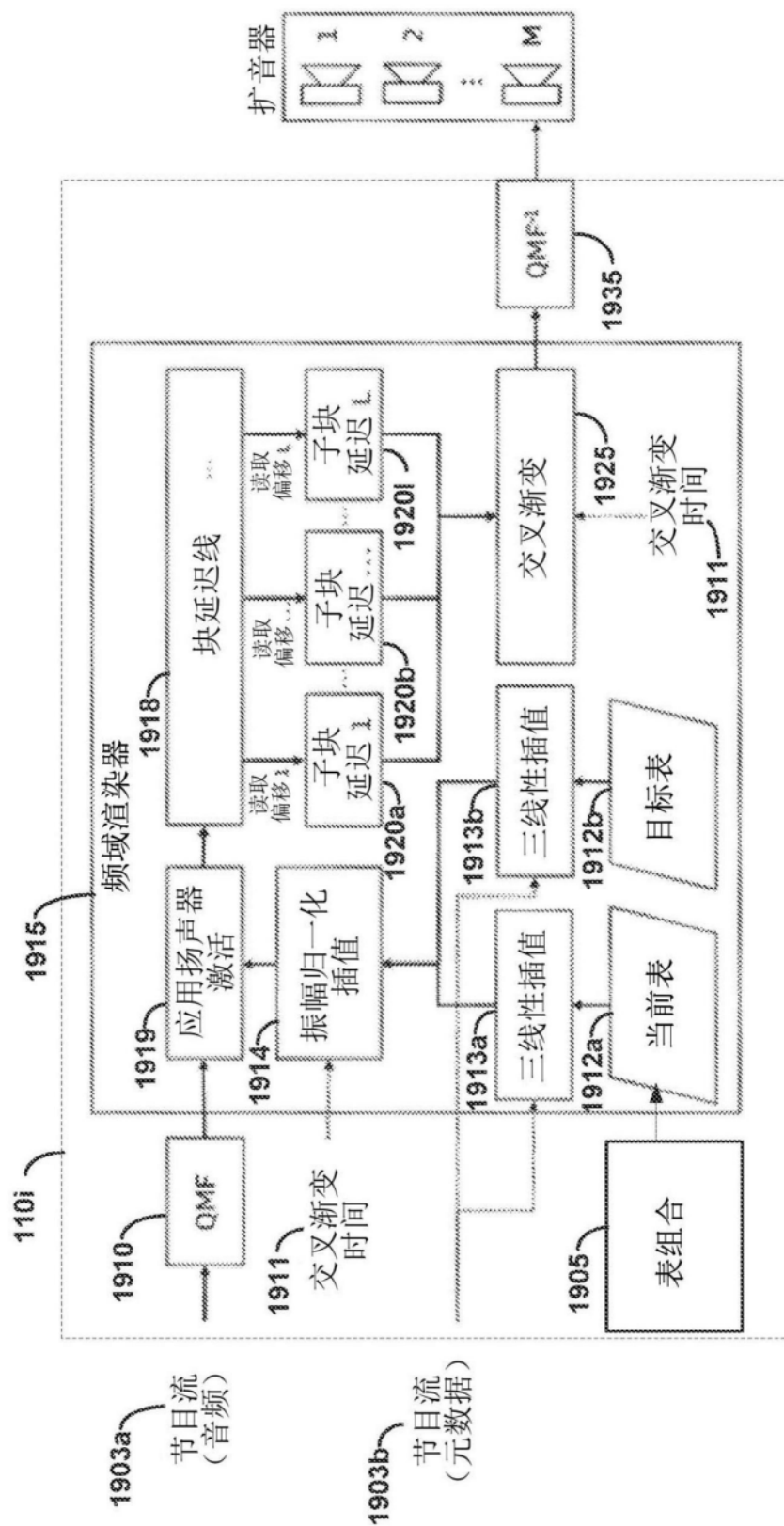


图19

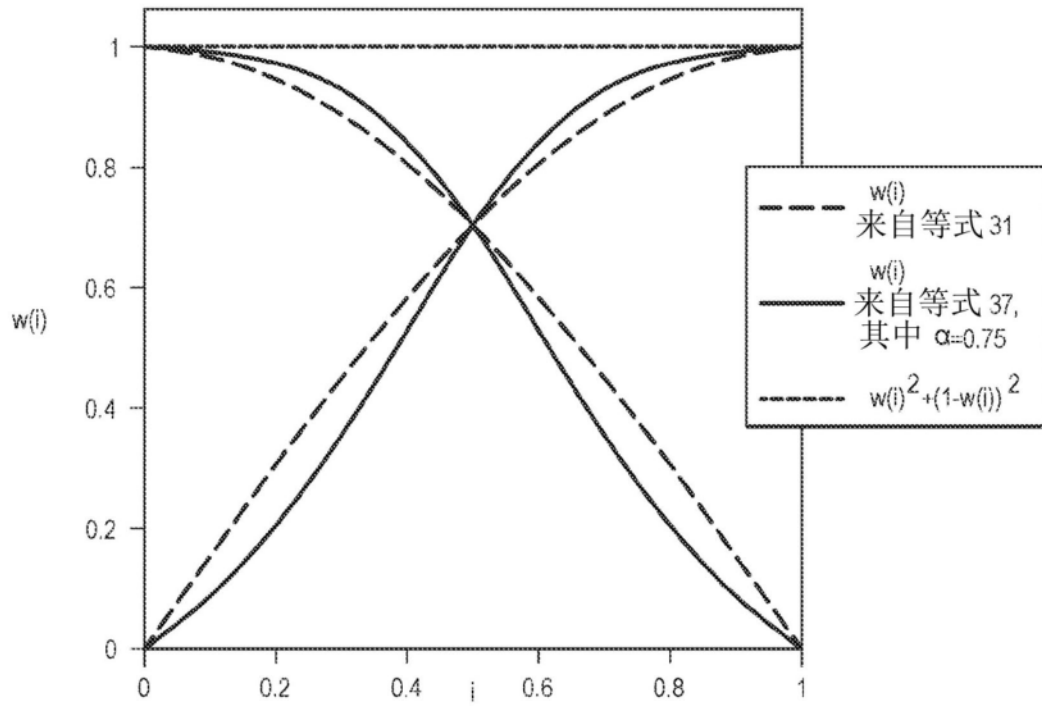


图20A

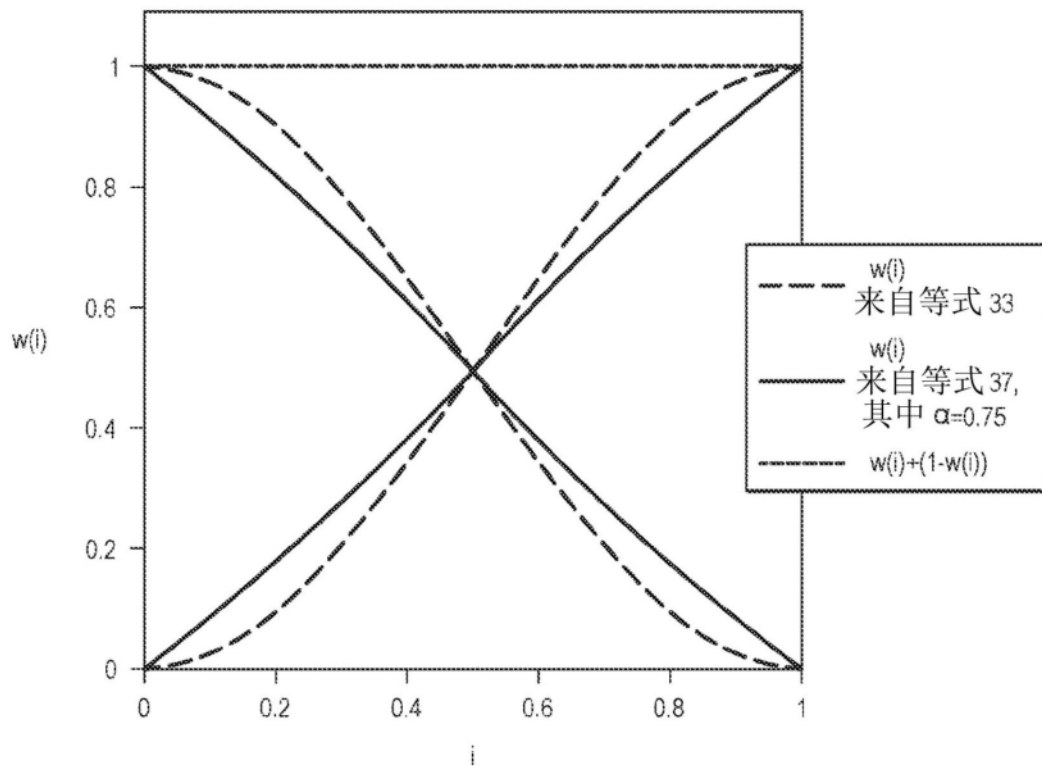


图20B

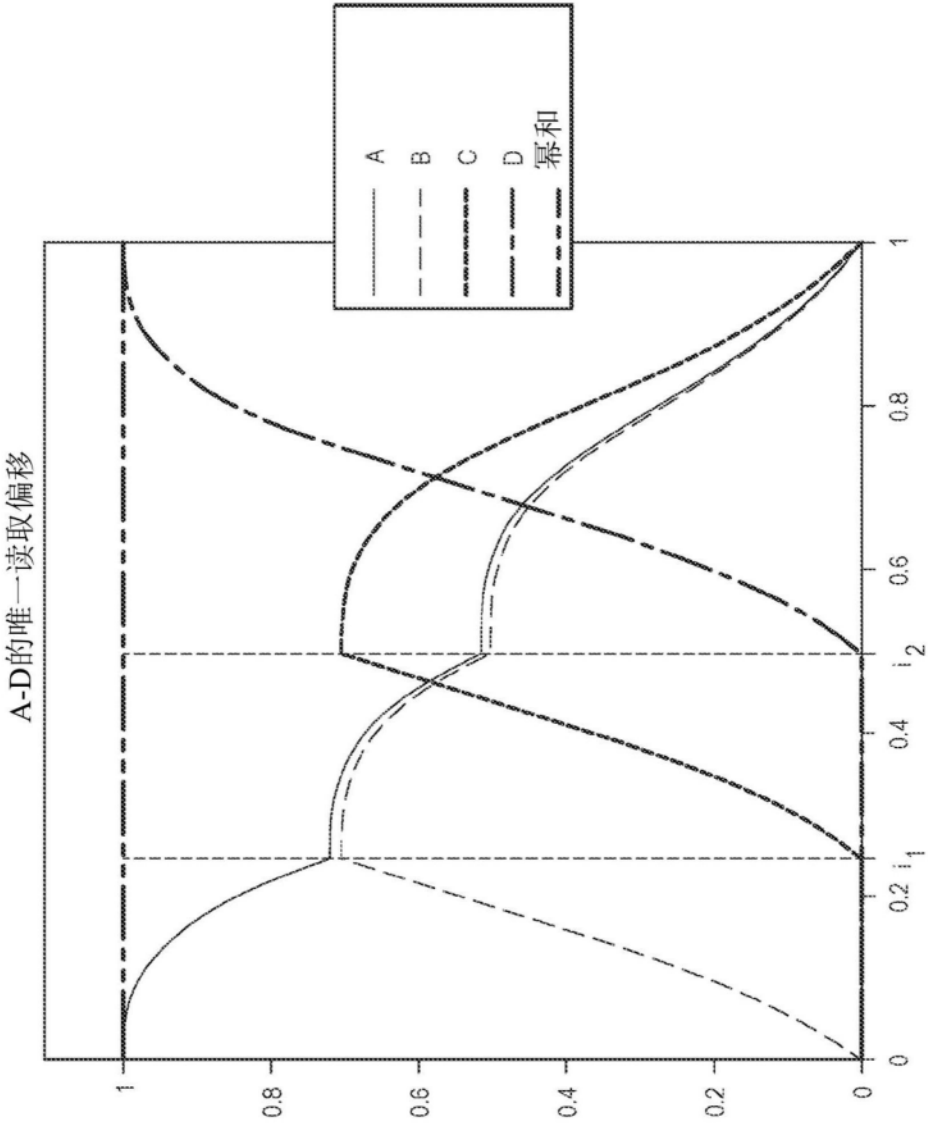


图21

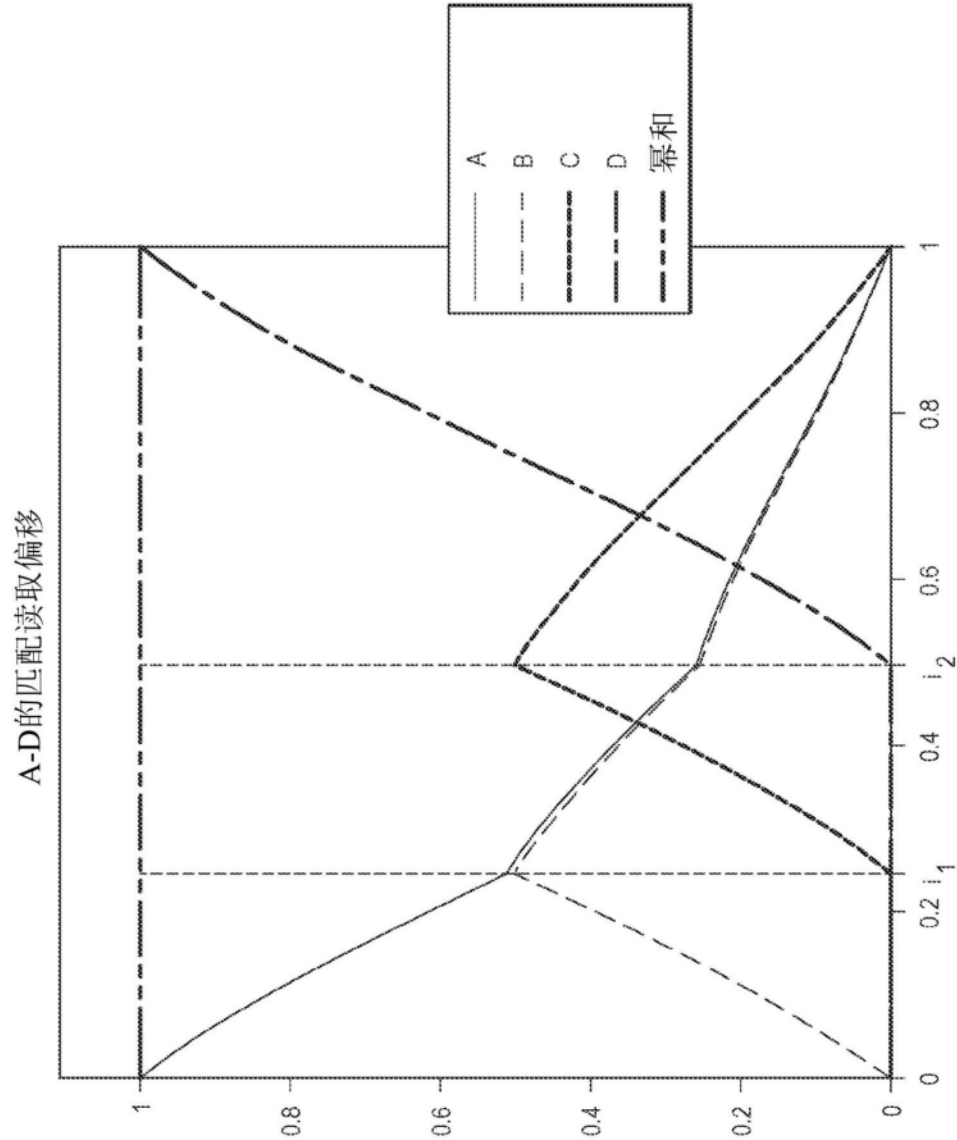


图22

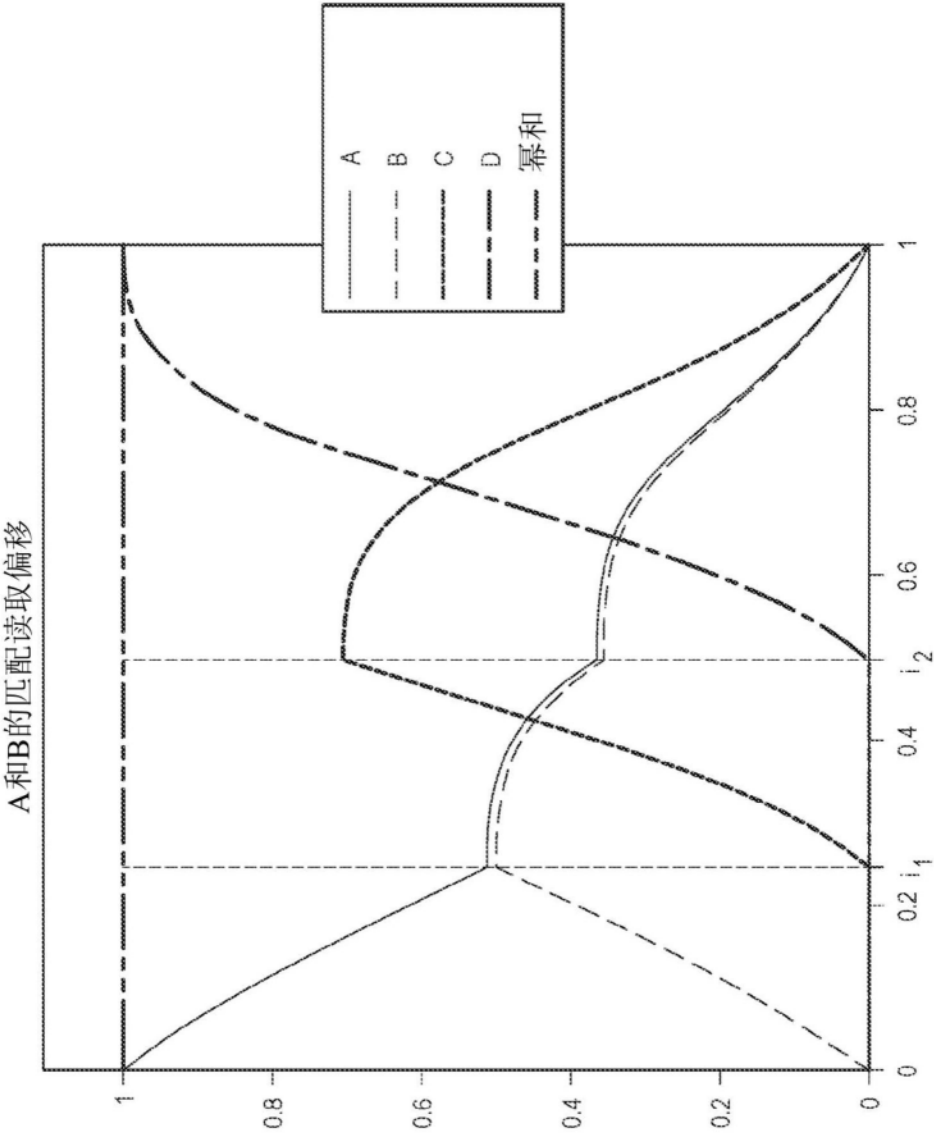


图23

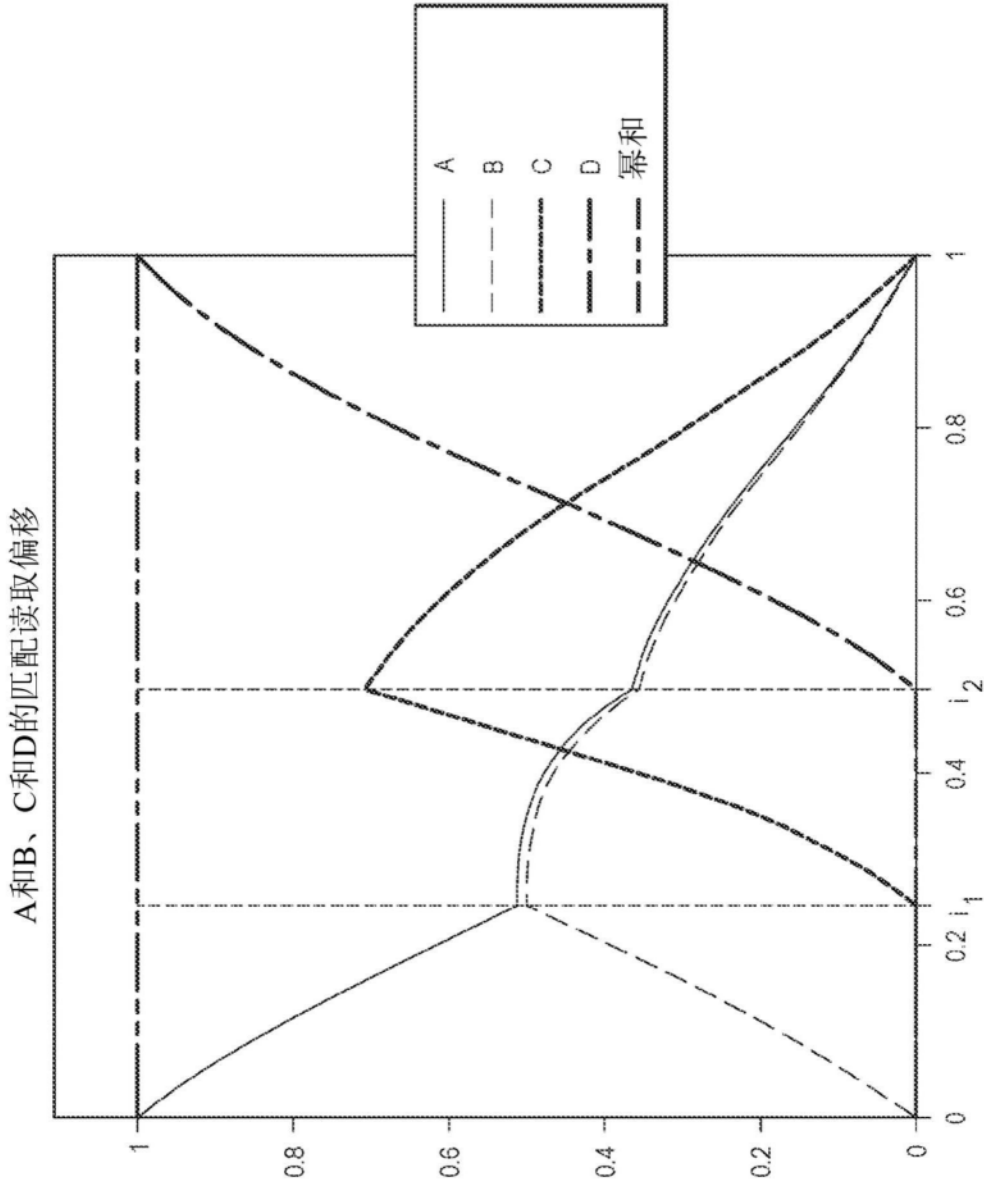


图24

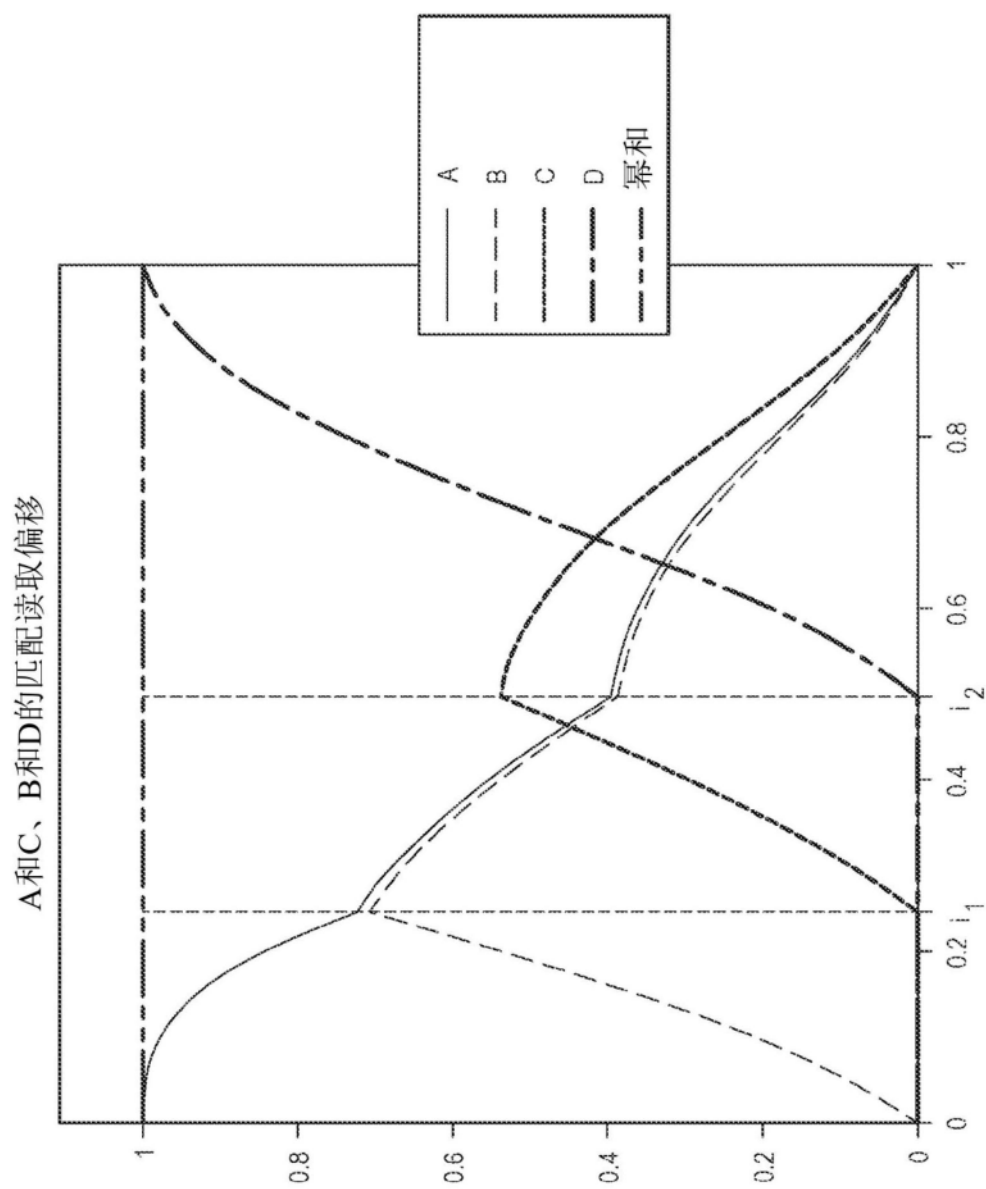


图25

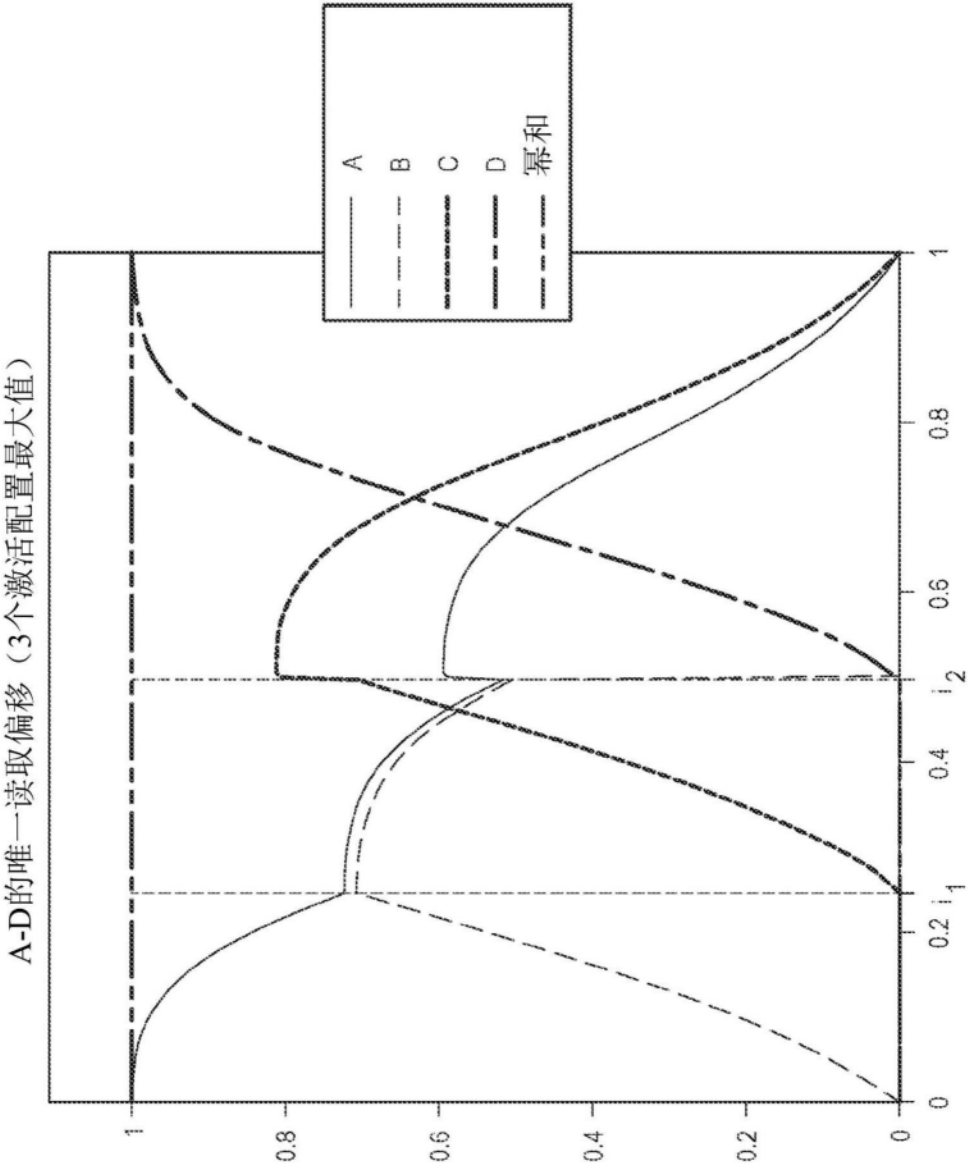


图26

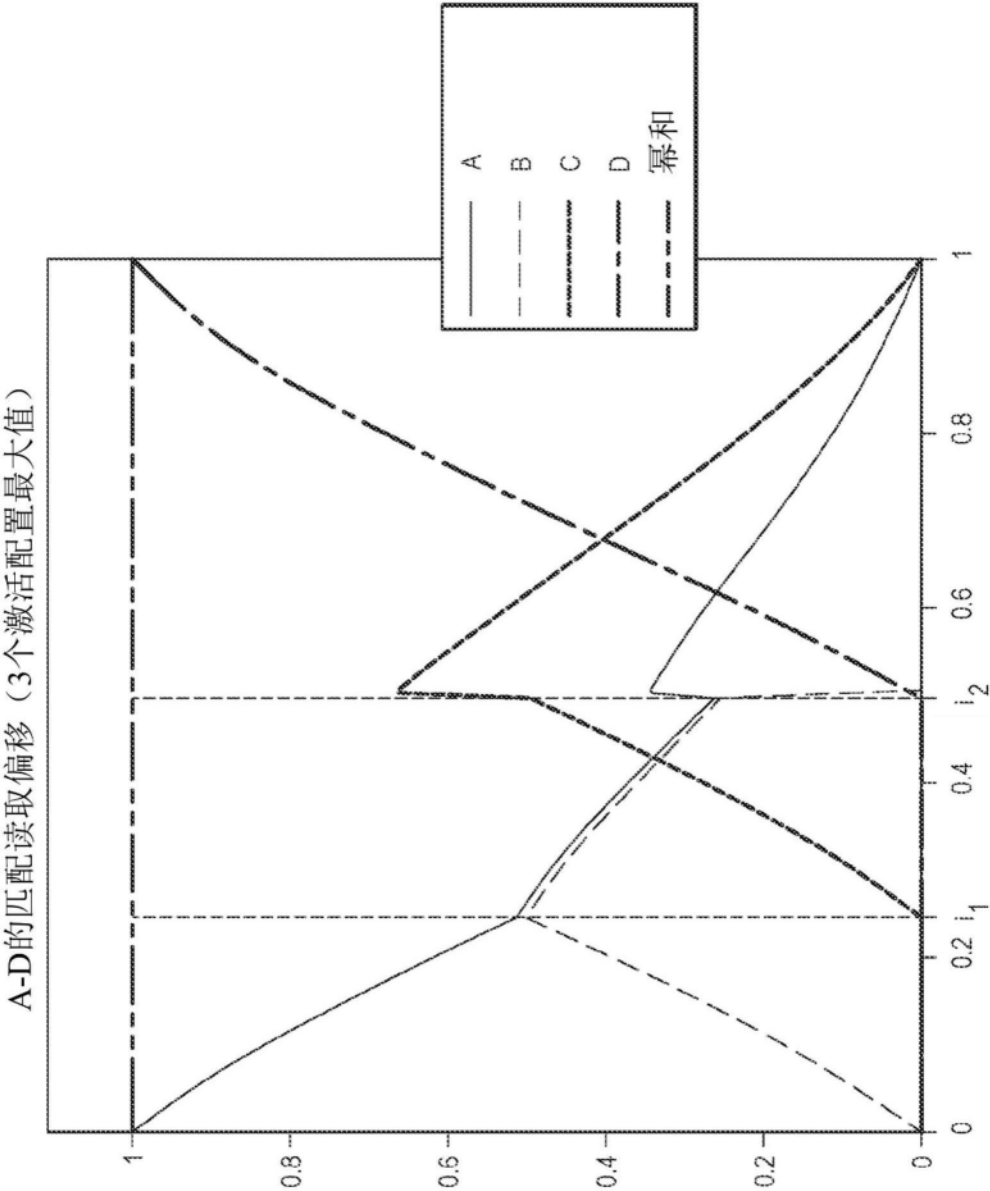


图27

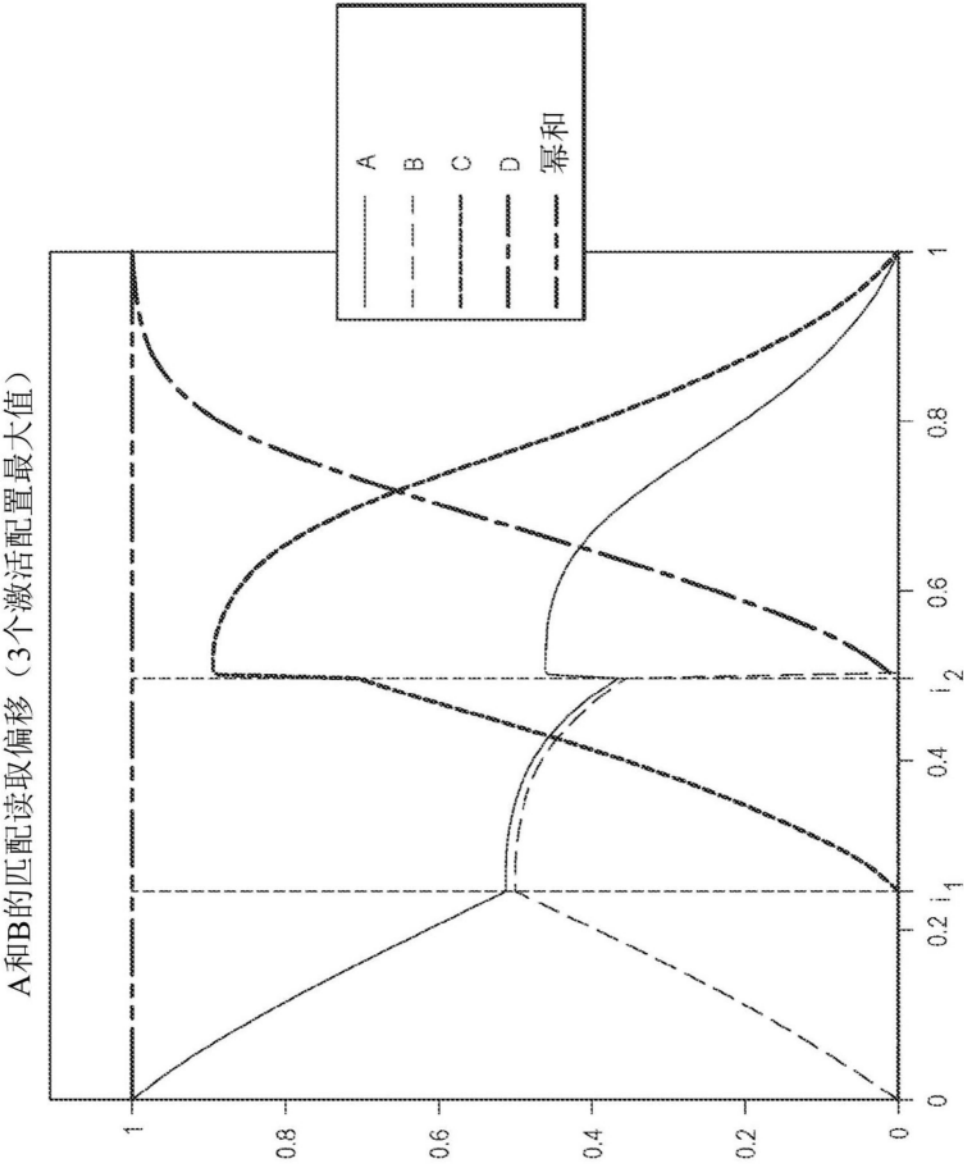


图28

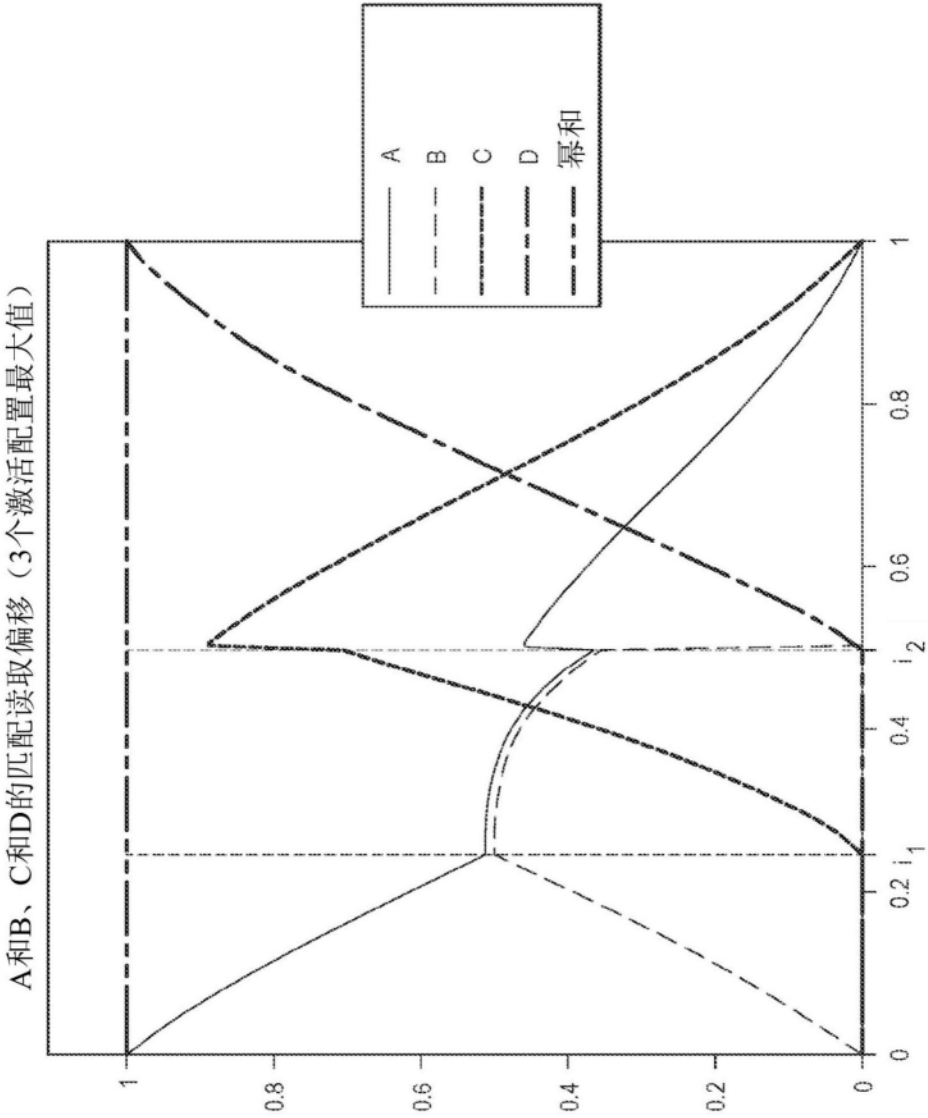


图29

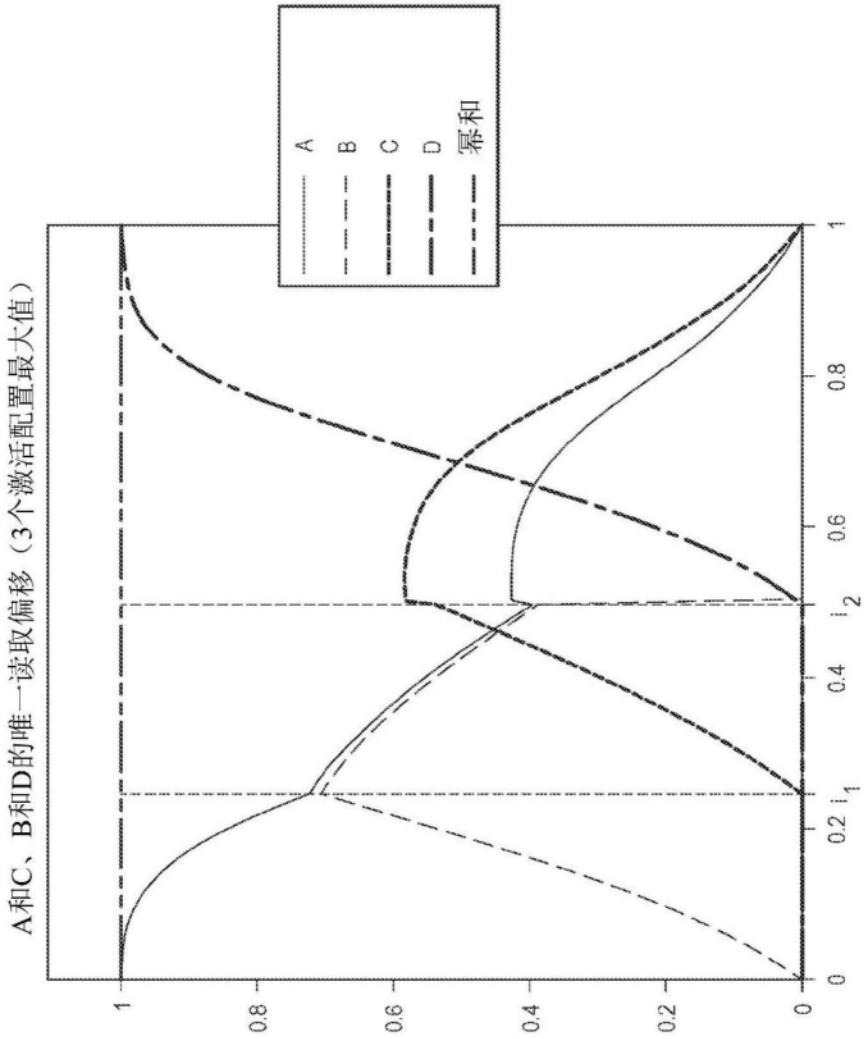


图30

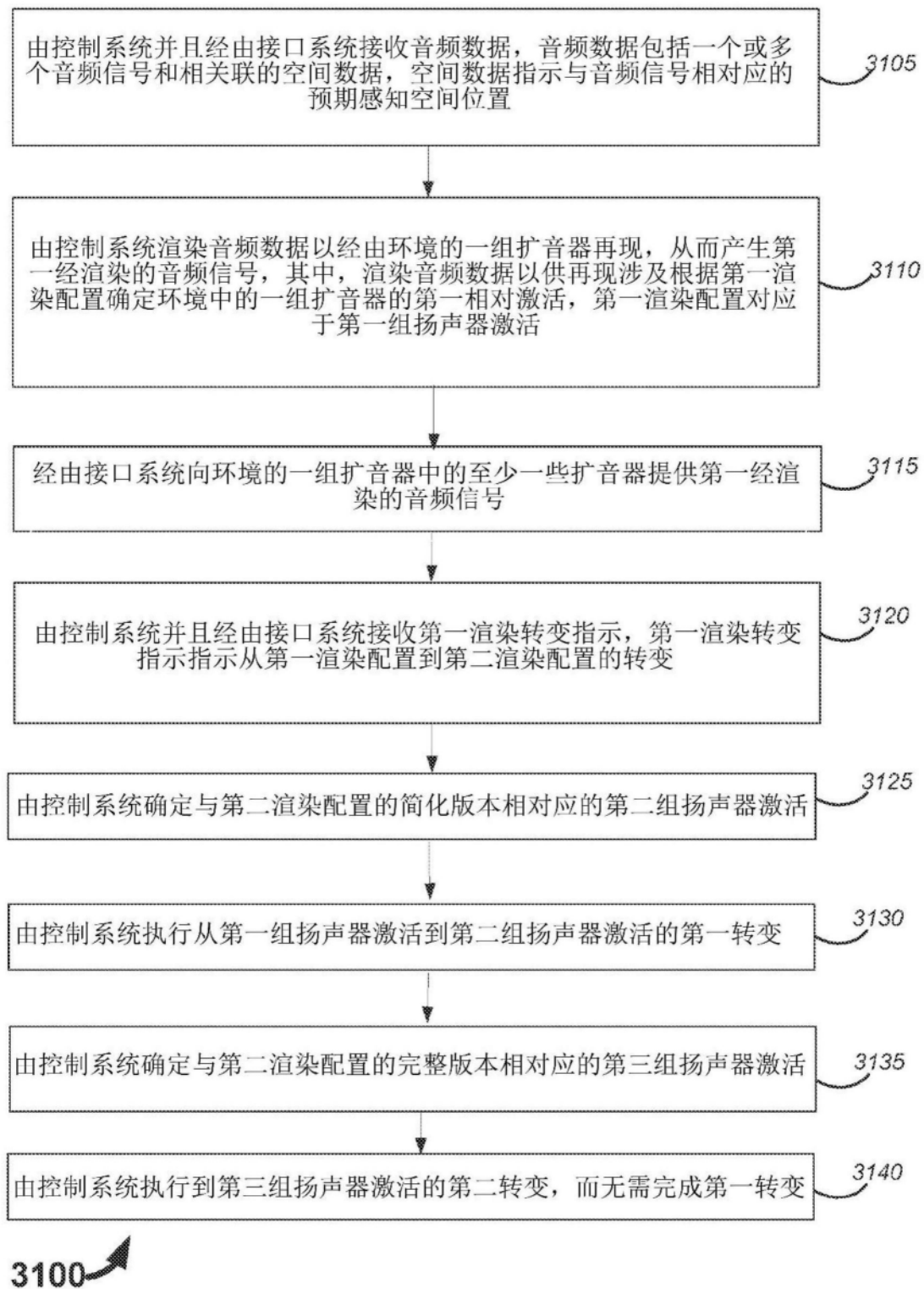


图31

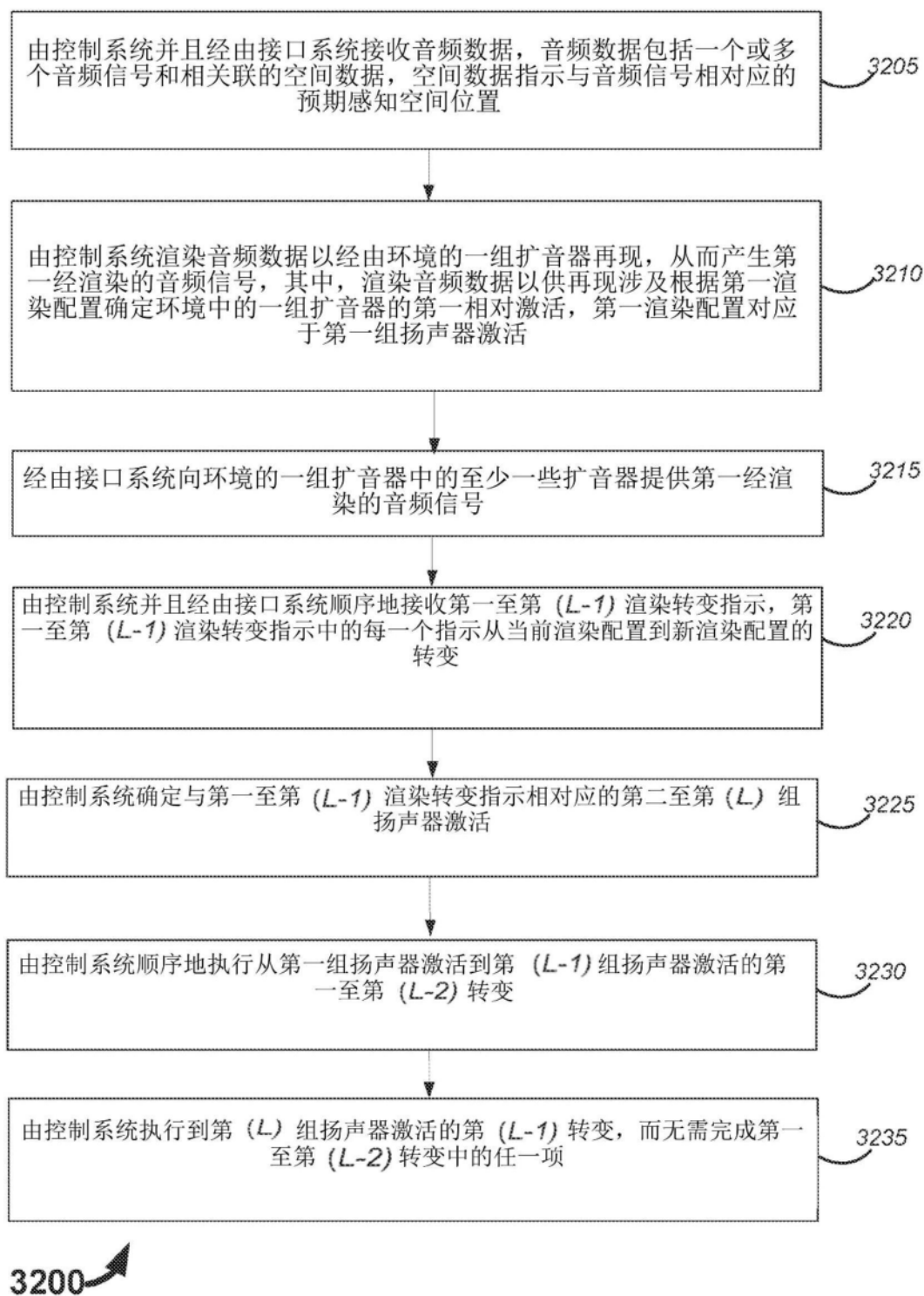


图32

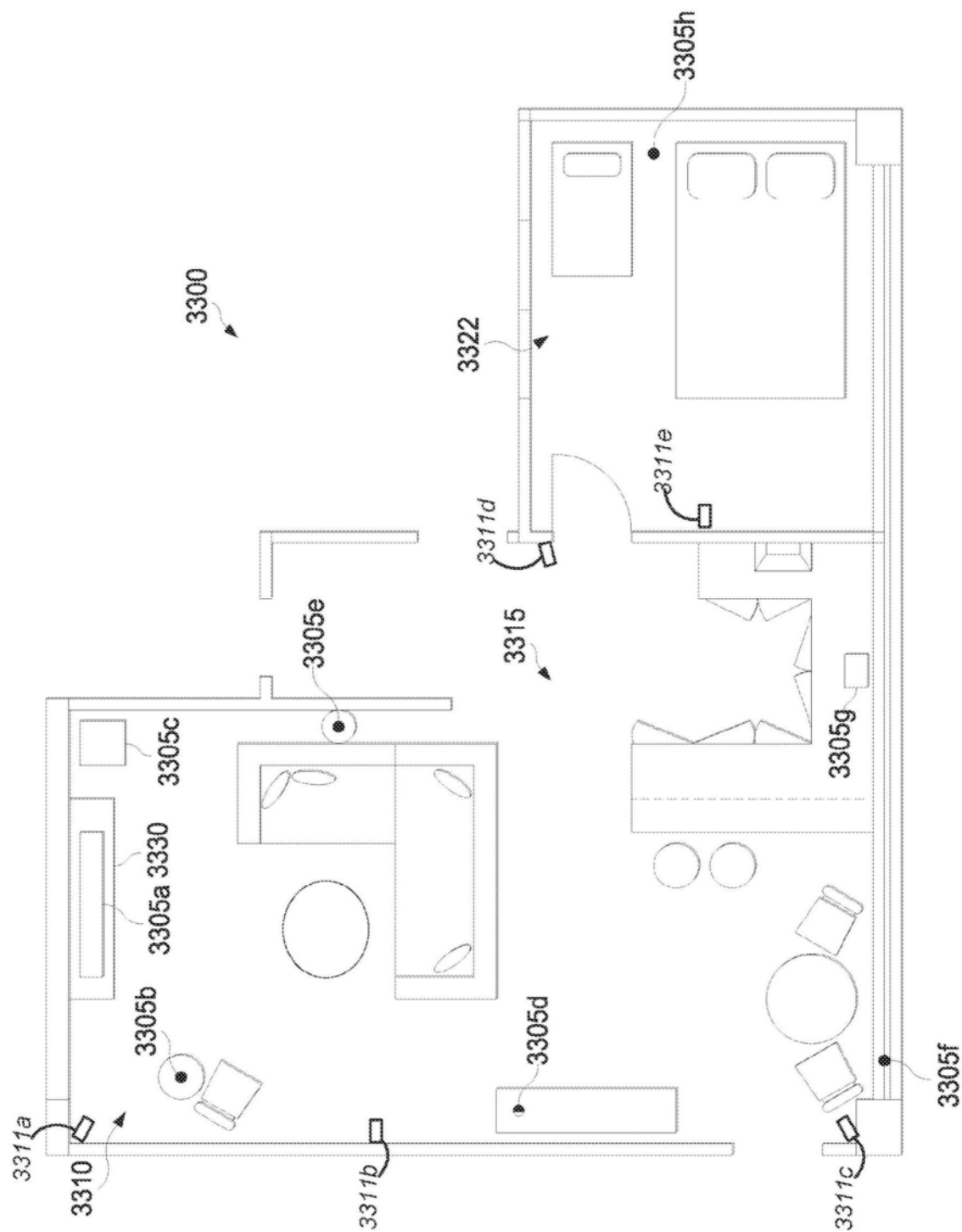


图33