



(12)发明专利

(10)授权公告号 CN 105075292 B

(45)授权公告日 2017. 07. 25

(21)申请号 201480009029.4

专利权人 杜比国际公司

(22)申请日 2014.03.10

(72)发明人 安东尼奥·马特奥斯舒莱

(65)同一申请的已公布的文献号

尼古拉斯·R·泰辛戈斯

申请公布号 CN 105075292 A

(74)专利代理机构 北京集佳知识产权代理有限公司

(43)申请公布日 2015.11.18

公司 11227

代理人 朱胜 李春晖

(30)优先权数据

P201330461 2013.03.28 ES

61/833,581 2013.06.11 US

(51)Int.Cl.

H04S 3/00(2006.01)

(85)PCT国际申请进入国家阶段日

2015.08.14

(56)对比文件

CN 102576562 A, 2012.07.11,

CN 101874414 A, 2010.10.27,

WO 2013006322 A1, 2013.01.10,

US 2006206221 A1, 2006.09.14,

US 2010092014 A1, 2010.04.15,

(86)PCT国际申请的申请数据

PCT/US2014/022793 2014.03.10

(87)PCT国际申请的公布数据

WO2014/159272 EN 2014.10.02

审查员 吕薇

(73)专利权人 杜比实验室特许公司

地址 美国加利福尼亚州

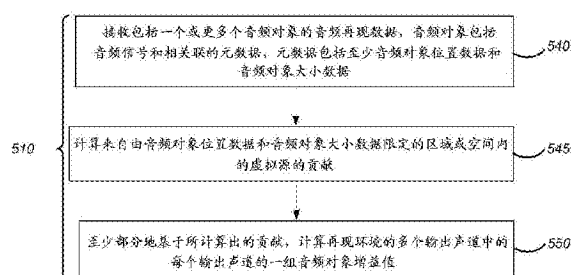
权利要求书4页 说明书16页 附图14页

(54)发明名称

用于创作和渲染音频再现数据的方法和设备

(57)摘要

可以针对音频对象在其内可以移动的空间限定多个虚拟源位置。用于渲染音频数据的建立步骤可以包括:接收再现扬声器位置数据并且根据再现扬声器位置数据和每个虚拟源位置预先计算每个虚拟源的增益值。增益值可以在“运行时”期间被存储和使用,在“运行时”期间针对再现环境的扬声器渲染音频再现数据。在运行时期,针对每个音频对象,可以计算来自自由音频对象位置数据和音频对象大小数据限定的区域或空间内的虚拟源位置的贡献。可以至少部分地基于所计算出的贡献计算再现环境的每个输出声道的一组增益值。每个输出声道可以与再现环境的至少一个再现扬声器相对应。



1. 一种用于创作和渲染音频再现数据的方法,包括:

接收包括一个或多个音频对象的音频再现数据,所述音频对象包括音频信号和相关联的元数据,所述元数据包括至少音频对象位置数据和音频对象大小数据;

针对所述一个或多个音频对象中的音频对象,计算来自所述音频对象位置数据和所述音频对象大小数据限定的音频对象区域或空间内相应虚拟源位置处的虚拟源的虚拟源增益值;以及

至少部分地基于所计算出的虚拟源增益值来计算多个输出声道中的每个输出声道的一组音频对象增益值,其中,每个输出声道与再现环境的至少一个再现扬声器相对应,并且所述虚拟源位置中的每个虚拟源位置与所述再现环境内的相应静止位置相对应。

2. 根据权利要求1所述的方法,其中,计算一组音频对象增益值的步骤包括:计算所述音频对象区域或空间内的虚拟源的虚拟源增益值的加权平均值。

3. 根据权利要求2所述的方法,其中,用于所述加权平均值的权重取决于所述音频对象的位置、所述音频对象的大小和所述音频对象区域或空间内的每个虚拟源位置。

4. 根据权利要求1所述的方法,还包括:

接收包括再现扬声器位置数据的再现环境数据。

5. 根据权利要求4所述的方法,还包括:

根据所述再现环境数据限定多个虚拟源位置;以及

针对所述虚拟源位置中的每个虚拟源位置,计算所述多个输出声道中的每个输出声道的虚拟源增益值。

6. 根据权利要求5所述的方法,其中,所述虚拟源位置中的至少一些虚拟源位置与所述再现环境外部的的位置相对应。

7. 根据权利要求5所述的方法,其中,所述虚拟源位置沿x轴、y轴和z轴被均匀地间隔开。

8. 根据权利要求5所述的方法,其中,所述虚拟源位置具有沿x轴和y轴的第一均匀间距和沿z轴的第二均匀间距。

9. 根据权利要求7或8所述的方法,其中,所述计算所述多个输出声道中的每个输出声道的一组音频对象增益值的步骤包括:独立计算来自沿x轴、y轴和z轴的虚拟源的虚拟源增益值。

10. 根据权利要求5所述的方法,其中,所述虚拟源位置被非均匀地间隔开。

11. 根据权利要求4所述的方法,还包括:将所计算出的虚拟源增益值存储在存储系统中。

12. 根据权利要求11所述的方法,其中,计算一组音频对象增益值的步骤包括:

从所述存储系统检索所计算出的、与音频对象位置和音频对象大小相对应的虚拟源增益值;以及

在所计算出的虚拟源增益值之间进行插值。

13. 根据权利要求12所述的方法,其中,在所计算出的虚拟源增益值之间进行插值的步骤包括:

确定所述音频对象位置附近的多个邻近虚拟源位置;

确定所计算出的、每个所述邻近虚拟源位置的虚拟源增益值;

确定所述音频对象位置和每个所述邻近虚拟源位置之间的多个距离;以及根据所述多个距离在所计算出的虚拟源增益值之间进行插值。

14. 根据权利要求1所述的方法,其中,所述音频对象区域或空间是矩形、矩形棱柱、圆形、球形、椭圆形或椭圆体中至少之一。

15. 根据权利要求1所述的方法,其中,所述再现环境包括影院音响系统环境。

16. 根据权利要求1所述的方法,还包括:对至少一些所述音频再现数据进行去相关。

17. 根据权利要求1所述的方法,还包括:对用于具有超过阈值的音频对象大小的音频对象的音频再现数据进行去相关。

18. 根据权利要求1所述的方法,其中,所述再现环境数据包括再现环境边界数据,所述方法还包括:

确定所述音频对象区域或空间包括再现环境边界外部的的外部区域或空间;以及至少部分地基于所述外部区域或空间来施加衰落因子。

19. 根据权利要求18所述的方法,还包括:

确定音频对象在距再现环境边界的阈值距离内;以及不向在所述再现环境的相对边界上的再现扬声器提供扬声器馈送信号。

20. 根据权利要求1所述的方法,还包括:

接收包括再现扬声器位置数据和再现环境边界数据的再现环境数据;

确定所述音频对象区域或空间包括再现环境边界外部的的外部区域或空间;以及至少部分地基于所述外部区域或空间确定衰落因子;

其中,计算所述多个输出声道中的每个输出声道的一组增益值也基于所述相关联的元数据和所述衰落因子。

21. 根据权利要求20所述的方法,其中,所述衰落因子与所述外部区域成比例。

22. 根据权利要求20所述的方法,还包括:

确定音频对象处在距再现环境边界的阈值距离内;以及不向在所述再现环境的相对边界上的再现扬声器提供扬声器馈送信号。

23. 根据权利要求20所述的方法,还包括:

根据所述再现环境数据限定多个虚拟源位置;以及

针对每个所述虚拟源位置计算多个输出声道中的每个输出声道的虚拟源增益。

24. 根据权利要求23所述的方法,其中,所述虚拟源位置被均匀地间隔开。

25. 一种用于创作和渲染音频再现数据的设备,所述设备包括:

用于接收包括一个或更多个音频对象的音频再现数据的装置,所述音频对象包括音频信号和相关联的元数据,所述元数据包括至少音频对象位置数据和音频对象大小数据;

用于针对所述一个或更多个音频对象中的音频对象,计算来自所述音频对象位置数据和所述音频对象大小数据限定的音频对象区域或空间内相应虚拟源位置处的虚拟源的虚拟源增益值的装置;以及

用于至少部分地基于所计算出的虚拟源增益值来计算多个输出声道中的每个输出声道的一组音频对象增益值的装置,其中,每个输出声道与再现环境的至少一个再现扬声器相对应,并且所述虚拟源位置中的每个虚拟源位置与所述再现环境内的相应静止位置相对应。

26. 根据权利要求25所述的设备, 其中, 所述用于计算一组音频对象增益值的装置计算所述音频对象区域或空间内的虚拟源的虚拟源增益值的加权平均值。

27. 根据权利要求26所述的设备, 其中, 用于所述加权平均值的权重取决于所述音频对象的位置、所述音频对象的大小和所述音频对象区域或空间内的每个虚拟源位置。

28. 根据权利要求25所述的设备, 还包括: 用于接收包括再现扬声器位置数据的再现环境数据的装置。

29. 根据权利要求28所述的设备, 还包括:

用于根据所述再现环境数据限定多个虚拟源位置的装置; 以及

用于针对所述虚拟源位置中的每个虚拟源位置, 计算所述多个输出声道中的每个输出声道的虚拟源增益值的装置。

30. 根据权利要求28所述的设备, 其中, 所述虚拟源位置中的至少一些虚拟源位置与所述再现环境外部的的位置相对应。

31. 根据权利要求28所述的设备, 其中, 所述虚拟源位置沿x轴、y轴和z轴被均匀地间隔开。

32. 根据权利要求28所述的设备, 其中, 所述虚拟源位置具有沿x轴和y轴的第一均匀间距和沿z轴的第二均匀间距。

33. 根据权利要求31或32所述的设备, 其中, 所述用于计算一组音频对象增益值的装置独立计算来自沿x轴、y轴和z轴的虚拟源的虚拟源增益值。

34. 一种用于创作和渲染音频再现数据的设备, 包括:

接口系统; 以及

逻辑系统, 包括:

用于从所述接口系统接收包括一个或更多个音频对象的音频再现数据的装置, 所述音频对象包括音频信号和相关联的元数据, 所述元数据包括至少音频对象位置数据和音频对象大小数据;

用于针对一个或更多个所述音频对象中的音频对象, 计算来自所述音频对象位置数据和所述音频对象大小数据限定的音频对象区域或空间内相应虚拟源位置处的虚拟源的虚拟源增益值的装置; 以及

用于至少部分地基于所计算出的虚拟源增益值来计算多个输出声道中的每个输出声道的一组音频对象增益值的装置, 其中, 每个输出声道与再现环境的至少一个再现扬声器相对应, 并且所述虚拟源位置中的每个虚拟源位置与所述再现环境内的相应静止位置相对应。

35. 根据权利要求34所述的设备, 其中, 所述用于计算一组音频对象增益值的装置计算所述音频对象区域或空间内的虚拟源的虚拟源增益值的加权平均值。

36. 根据权利要求35所述的设备, 其中, 用于所述加权平均值的权重取决于所述音频对象的位置、所述音频对象的大小和所述音频对象区域或空间内的每个虚拟源位置。

37. 根据权利要求34所述的设备, 其中, 所述逻辑系统还包括用于从所述接口系统接收包括再现扬声器位置数据的再现环境数据的装置。

38. 根据权利要求37所述的设备, 其中, 所述逻辑系统还包括:

用于根据所述再现环境数据限定多个虚拟源位置的装置; 以及

用于针对所述虚拟源位置中的每个虚拟源位置,计算所述多个输出声道中的每个输出声道的虚拟源增益值的装置。

39.根据权利要求38所述的设备,其中,所述虚拟源位置中的至少一些虚拟源位置与所述再现环境外部的的位置相对应。

40.根据权利要求38所述的设备,其中,所述虚拟源位置沿x轴、y轴和z轴被均匀地间隔开。

41.根据权利要求38所述的设备,其中,所述虚拟源位置具有沿x轴和y轴的第一均匀间距和沿z轴的第二均匀间距。

42.根据权利要求40或41所述的设备,其中,所述用于计算所述多个输出声道中的每个输出声道的一组音频对象增益值的装置独立计算来自沿x轴、y轴和z轴的虚拟源的虚拟源增益值。

43.根据权利要求34所述的设备,还包括存储器装置,其中,所述接口系统包括在所述逻辑系统和所述存储器装置之间的接口。

44.根据权利要求43所述的设备,其中,所述接口系统包括网络接口。

45.根据权利要求43所述的设备,还包括用户接口,其中,所述逻辑系统还包括用于经由所述用户接口接收用户输入的装置,所述用户输入包括但不限于输入音频对象大小数据。

46.根据权利要求45所述的设备,其中,所述逻辑系统还包括用于对所述输入音频对象大小数据进行缩放的装置。

用于创作和渲染音频再现数据的方法和设备

[0001] 相关申请的交叉引用

[0002] 本申请要求于2013年3月28日提交的西班牙专利申请No.P201330461以及于2013年6月11日提交的美国临时专利申请No.61/833,581的优先权,这两个专利申请中的每个专利申请在此通过引用以其全部内容合并到本文中。

技术领域

[0003] 本公开内容涉及音频再现数据的创作和渲染。具体地,本公开内容涉及针对再现环境(如影院音响再现系统)来创作和渲染音频再现数据。

背景技术

[0004] 由于在1927年引入了有声电影,所以出现了用于捕获电影声轨的艺术意图并且在影院环境中对其进行重放的技术的稳定发展。在二十世纪三十年代,唱片的同步声音由电影的变积式(variable area)声音代替,在二十世纪四十年代使用剧场声学考虑和改进的扬声器设计连同早期引入的多轨记录和可操纵的重放(使用控制音调来移动声音)一起对此进行了进一步改进。在二十世纪五六十年代,电影的涂磁道(magnetic striping)允许剧院中的多声道回放、将环绕声声道和高达五个屏幕声道引入高级剧院中。

[0005] 在二十世纪七十年代,杜比提出了在后期制作中和在胶片上降噪以及对与3个屏幕声道和单声道环绕声道的混合(mixes)进行编码和分配的经济有效的方法。在二十世纪八十年代,使用杜比频谱记录(SR)降噪和验证程序(如THX)进一步提高了影院音响的品质。在二十世纪九十年代期间,杜比使用提供分离的左屏幕声道、中心屏幕声道、右屏幕声道、左环绕阵列和右环绕阵列以及用于低频效果的超低音扬声器声道的5.1声道格式将数字声音引入电影。2010年提出的杜比环绕声7.1通过将现有的左环绕声道和右环绕声道分成四个“区”增加了环绕声道的数量。

[0006] 随着声道数量增加以及扬声器布局从平面二维(2D)阵列转变成包括高度的三维(3D)阵列,创作和渲染声音的任务正变得越来越复杂。期待改进的方法和装置。

发明内容

[0007] 本公开内容中描述的主题的一些方面可以在用于渲染包括没有参照任何特定再现环境创建的音频对象的音频再现数据的工具中来实现。如本文中所使用的,术语“音频对象”可以指代音频信号和相关联的元数据的流。元数据可以至少表示音频对象的位置和音频对象的表观大小。然而,元数据还可以表示渲染约束数据、内容类型数据(例如,会话、效果等)、增益数据、轨迹数据等。一些音频对象可以是静止的,而其他音频对象可以具有随时间变化的元数据:这样的音频对象可以移动,可以改变大小和/或可以具有随时间变化的其他属性。

[0008] 当音频对象在再现环境中被回放或被监视时,可以至少根据位置元数据和大小元数据来渲染音频对象。渲染步骤可以包括:计算一组输出声道中的每个声道的一组音频对

象增益值。每个输出声道可以与再现环境中的一个或多个再现扬声器相对应。

[0009] 本文中描述的一些实现包括可以在渲染任何特定音频对象之前发生的“建立”步骤。在本文中还可以被称为第一级或级1的建立步骤可以包括：在音频对象可以在其内移动的空间中限定多个虚拟源位置。如本文中所使用的，“虚拟源位置”是静止点源的位置。根据这样的实现，建立步骤可以包括：接收再现扬声器位置数据并且根据再现扬声器位置数据和虚拟源位置预先计算每个虚拟源的虚拟源增益值。如本文中所使用的，术语“扬声器位置数据”可以包括表示再现环境的一些或所有扬声器的位置的位置数据。位置数据可以被设置为再现扬声器位置的绝对坐标，例如笛卡尔坐标、球面坐标等。可替代地或另外，位置数据可以被设置为相对于其他再现环境位置（例如再现环境的声学“最佳听音位置 (sweet spots)”）的坐标（例如，如笛卡尔坐标或角坐标）。

[0010] 在一些实现中，虚拟源增益值可以在“运行时”期间被存储和使用，在该“运行时”期间，针对再现环境的扬声器渲染音频再现数据。在运行时期，针对每个音频对象，可以计算来自自由音频对象位置数据和音频对象大小数据限定的区域或空间内的虚拟源位置的贡献。计算来自虚拟源位置的贡献的步骤可以包括：计算在建立步骤期间针对由音频对象的大小和音频对象的位置限定的音频对象区域或空间内的虚拟源位置确定的多个预先计算出的虚拟源增益值的加权平均值。可以至少部分地基于所计算出的虚拟源贡献来计算再现环境的每个输出声道的一组音频对象增益值。每个输出声道可以与再现环境的至少一个再现扬声器相对应。

[0011] 因此，本文中描述的一些方法包括：接收包括一个或多个音频对象的音频再现数据。音频对象可以包括音频信号和相关联的元数据。元数据可以包括至少音频对象位置数据和音频对象大小数据。所述方法可以包括：计算来自自由音频对象位置数据和音频对象大小数据限定的音频对象区域或空间内的虚拟源的贡献。所述方法可以包括：至少部分地基于所计算出的贡献来计算多个输出声道中的每个输出声道的一组音频对象增益值。每个输出声道可以与再现环境中的至少一个再现扬声器相对应。例如，再现环境可以是影院音响系统环境。

[0012] 计算来自虚拟源的贡献的步骤可以包括：计算所述音频对象区域或空间内的虚拟源的虚拟源增益值的加权平均值。所述加权平均值的权重可以取决于音频对象的位置、音频对象的大小和/或所述音频对象区域或空间内的每个虚拟源位置。

[0013] 所述方法还可以包括：接收包括再现扬声器位置数据的再现环境数据。所述方法还可以包括：根据再现环境数据限定多个虚拟源位置，并且针对每个虚拟源位置计算多个输出声道中的每个输出声道的虚拟源增益值。在一些实现中，每个虚拟源位置可以与再现环境内的位置相对应。然而，在一些实现中，至少一些虚拟源位置可以与再现环境外部的的位置相对应。

[0014] 在一些实现中，虚拟源位置可以沿x轴、y轴和z轴被均匀地间隔开。然而，在一些实现中，在所有方向上，间距可以不同。例如，虚拟源位置可以具有沿x轴和y轴的第一均匀间距和沿z轴的第二均匀间距。计算多个输出声道中的每个输出声道的一组音频对象增益值的步骤可以包括：独立计算来自沿x轴、y轴和z轴的虚拟源的贡献。在替代实现中，虚拟源位置可以被非均匀地间隔开。

[0015] 在一些实现中，计算多个输出声道中的每个输出声道的音频对象增益值的步骤可

以包括:确定要在位置 x_0 、 y_0 、 z_0 处渲染的各种大小的音频对象的增益值($g_1(x_0, y_0, z_0; s)$)。例如,音频对象增益值($g_1(x_0, y_0, z_0; s)$)可以被表达为:

$$[0016] \quad \left[\sum_{x_{vs}, y_{vs}, z_{vs}} [w(x_{vs}, y_{vs}, z_{vs}; x_0, y_0, z_0; s) g_1(x_{vs}, y_{vs}, z_{vs})]^p \right]^{\frac{1}{p}},$$

[0017] 其中, (x_{vs}, y_{vs}, z_{vs}) 表示虚拟源位置, $g_1(x_{vs}, y_{vs}, z_{vs})$ 表示虚拟源位置 x_{vs}, y_{vs}, z_{vs} 的声道1的增益值, 并且 $w(x_{vs}, y_{vs}, z_{vs}; x_0, y_0, z_0; s)$ 表示至少部分地基于音频对象的位置 (x_0, y_0, z_0) 、音频对象的大小和虚拟源位置 (x_{vs}, y_{vs}, z_{vs}) 确定的 $g_1(x_{vs}, y_{vs}, z_{vs})$ 的一个或多个加权函数。

[0018] 根据一些这样的实现, $g_1(x_{vs}, y_{vs}, z_{vs}) = g_1(x_{vs}) g_1(y_{vs}) g_1(z_{vs})$, 其中, $g_1(x_{vs})$ 、 $g_1(y_{vs})$ 和 $g_1(z_{vs})$ 表示x、y和z的独立的增益函数。在一些这样的实现中, 加权函数可以因式分解为(factor as):

$$[0019] \quad w(x_{vs}, y_{vs}, z_{vs}; x_0, y_0, z_0; s) = w_x(x_{vs}; x_0; s) w_y(y_{vs}; y_0; s) w_z(z_{vs}; z_0; s),$$

[0020] 其中, $w_x(x_{vs}; x_0; s)$, $w_y(y_{vs}; y_0; s)$ 和 $w_z(z_{vs}; z_0; s)$ 表示 x_{vs} 、 y_{vs} 和 z_{vs} 的独立的加权函数。根据一些这样的实现, p可以是音频对象大小的函数。

[0021] 一些这样的方法可以包括:将所计算出的虚拟源增益值存储在存储系统中。计算来自音频对象区域或空间内的虚拟源的贡献的步骤可以包括:从存储系统检索所计算出的、与音频对象位置和音频对象大小相对应的虚拟源增益值, 并且在所计算出的虚拟源增益值之间进行插值。在所计算出的虚拟源增益值之间进行插值的步骤可以包括:确定音频对象位置附近的多个邻近虚拟源位置;确定所计算出的、每个邻近虚拟源位置的虚拟源增益值;确定音频对象位置和每个邻近虚拟源位置之间的多个距离;以及根据多个距离在所计算出的虚拟源增益值之间进行插值。

[0022] 在一些实现中, 再现环境数据可以包括再现环境边界数据。所述方法可以包括:确定音频对象区域或空间包括再现环境边界外部的的外部区域或空间, 并且至少部分地基于所述外部区域或空间来施加衰落因子(fade-out factor)。一些方法可以包括:确定音频对象可以在距再现环境边界的阈值距离内, 并且不向在再现环境的相对边界上的再现扬声器提供扬声器馈送信号。在一些实现中, 音频对象区域或空间可以是矩形、矩形棱柱、圆形、球形、椭圆形和/或椭圆柱。

[0023] 一些方法可以包括对至少一些音频再现数据进行去相关。例如, 所述方法可以包括:对用于具有超过阈值的音频对象大小的音频对象的音频再现数据进行去相关。

[0024] 本文中描述了替代方法。一些这样的方法包括:接收包括再现扬声器位置数据和再现环境边界数据的再现环境数据, 并且接收包括一个或多个音频对象和相关联的元数据的音频再现数据。元数据可以包括音频对象位置数据和音频对象大小数据。所述方法可以包括:确定由音频对象位置数据和音频对象大小数据限定的音频对象区域或空间包括再现环境边界外部的的外部区域或空间, 并且至少部分地基于外部区域或空间来确定衰落因子。所述方法可以包括:至少部分地基于相关联的元数据和衰落因子来计算多个输出声道中的每个输出声道的一组增益值。每个输出声道可以与再现环境中的至少一个再现扬声器相对应。衰落因子可以与外部区域成比例。

[0025] 所述方法还可以包括:确定音频对象可以在距再现环境边界的阈值距离内, 并且

不向在再现环境的相对边界上的再现扬声器提供扬声器馈送信号。

[0026] 所述方法还可以包括:计算来自音频对象区域或空间内的虚拟源的贡献。所述方法可以包括:根据再现环境数据限定多个虚拟源位置,以及针对每个虚拟源位置计算多个输出声道中的每个输出声道的虚拟源增益。虚拟源位置可以被均匀地间隔开或者可以被非均匀地间隔开,这取决于具体实现。

[0027] 一些实现可以在存储有软件的一个或更多个非暂态介质中体现。软件可以包括用于对用于接收包括一个或更多个音频对象的音频再现数据的一个或更多个装置进行控制的指令。音频对象可以包括音频信号和相关联的元数据。元数据可以包括至少音频对象位置数据和音频对象大小数据。软件可以包括用于下述操作的指令:针对一个或更多个音频对象中的音频对象计算来自自由音频对象位置数据和音频对象大小数据限定的区域或空间内的虚拟源的贡献,并且至少部分地基于所计算出的贡献来计算多个输出声道中的每个输出声道的一组音频对象增益值。每个输出声道可以与再现环境的至少一个再现扬声器相对应。

[0028] 在一些实现中,计算来自虚拟源的贡献的步骤可以包括:计算来自音频对象区域或空间内的虚拟源的虚拟源增益值的加权平均值。所述加权平均值的权重可以取决于音频对象的位置、音频对象的大小和/或音频对象区域或空间内的每个虚拟源位置。

[0029] 软件可以包括用于接收包括再现扬声器位置数据的再现环境数据的指令。软件可以包括用于下述操作的指令:根据再现环境数据限定多个虚拟源位置,并且针对每个虚拟源位置计算多个输出声道中的每个输出声道的虚拟源增益值。每个虚拟源位置可以与再现环境内的位置相对应。在一些实现中,至少一些虚拟源位置可以与再现环境外部的的位置相对应。

[0030] 根据一些实现,虚拟源位置可以被均匀地间隔开。在一些实现中,虚拟源位置可以具有沿x轴和y轴的第一均匀间距和沿z轴的第二均匀间距。计算多个输出声道中的每个输出声道的一组音频对象增益值的步骤可以包括:独立计算来自沿x轴、y轴和z轴的虚拟源的贡献。

[0031] 本文中描述了各种装置和设备。一些这样的设备可以包括接口系统和逻辑系统。接口系统可以包括网络接口。在一些实现中,设备可以包括存储器装置。接口系统可以包括逻辑系统和存储器装置之间的接口。

[0032] 逻辑系统可以适用于:从接口系统接收包括一个或更多个音频对象的音频再现数据。音频对象可以包括音频信号和相关联的元数据。元数据可以包括至少音频对象位置数据和音频对象大小数据。逻辑系统可以适用于:针对一个或更多个音频对象中的音频对象来计算来自自由音频对象位置数据和音频对象大小数据限定的音频对象区域或空间内的虚拟源的贡献。逻辑系统可以适用于:至少部分地基于所计算出的贡献来计算多个输出声道中的每个输出声道的一组音频对象增益值。每个输出声道可以与再现环境中的至少一个再现扬声器相对应。

[0033] 计算来自虚拟源的贡献的步骤可以包括:计算音频对象区域或空间内的虚拟源的虚拟源增益值的加权平均值。用于加权平均值的权重可以取决于音频对象的位置、音频对象的大小和音频对象区域或空间内的每个虚拟源位置。逻辑系统可以适用于:从接口系统接收包括再现扬声器位置数据的再现环境数据。

[0034] 逻辑系统可以适用于：根据再现环境数据限定多个虚拟源位置，并且针对每个虚拟源位置计算多个输出声道中的每个输出声道的虚拟源增益值。每个虚拟源位置可以与再现环境内的位置相对应。然而，在一些实现中，至少一些虚拟源位置可以与再现环境外部的的位置相对应。虚拟源位置可以被均匀地间隔开或可以被非均匀地间隔开，这取决于实现。在一些实现中，虚拟源位置可以具有沿x轴和y轴的第一均匀间距和沿z轴的第二均匀间距。计算多个输出声道中的每个输出声道的一组音频对象增益值的步骤可以包括：独立计算来自沿x轴、y轴和z轴的虚拟源的贡献。

[0035] 设备还可以包括用户接口。逻辑系统可以适用于：经由用户接口接收用户输入（如音频对象大小数据）。在一些实现中，逻辑系统可以适用于对输入的音频对象大小数据进行缩放。

[0036] 在附图和下面的描述中阐述了在本说明书中描述的主题的一个或更多个实现的细节。根据描述、附图和权利要求，其他特征、方面和优点将变得明显。注意，下面的图的相对尺寸可能并未按比例绘出。

附图说明

[0037] 图1示出了具有杜比环绕声5.1配置的再现环境的示例；

[0038] 图2示出了具有杜比环绕声7.1配置的再现环境的示例；

[0039] 图3示出了具有Hamasaki 22.2环绕声配置的再现环境的示例；

[0040] 图4A示出了描绘在虚拟再现环境中的不同高度处的扬声器区的图形用户接口(GUI)的示例；

[0041] 图4B示出了另一再现环境的示例；

[0042] 图5A是提供音频处理方法的概述的流程图；

[0043] 图5B是提供建立处理的示例的流程图；

[0044] 图5C是提供根据预先计算的虚拟源位置的增益值计算接收的音频对象的增益值的运行时处理的示例的流程图；

[0045] 图6A示出了与再现环境有关的虚拟源位置的示例；

[0046] 图6B示出了与再现环境有关的虚拟源位置的替代示例；

[0047] 图6C至图6F示出了将近场声像调节(panning)技术和远场声像调节技术应用于不同位置处的音频对象的示例；

[0048] 图6G示出了在边长等于1的正方形的每个角落处具有一个扬声器的再现环境的示例；

[0049] 图7示出了来自音频对象位置数据和音频对象大小数据限定的区域内的虚拟源的贡献的示例；

[0050] 图8A和图8B示出了再现环境内的两个位置处的音频对象；

[0051] 图9是概述了至少部分地基于音频对象的多大区域或空间延伸到再现环境的边界外部来确定衰落因子的方法的流程图；

[0052] 图10是提供创作装置和/或渲染装置的部件的示例的框图；

[0053] 图11A是表示可以用于音频内容创建的一些部件的框图；以及

[0054] 图11B是表示可以用于再现环境中的音频回放的一些部件的框图。

[0055] 各个附图中的相同附图标记和名称表示相同的要素。

具体实施方式

[0056] 下面的描述涉及出于描述本公开内容的一些新颖方面的目的的一些实现以及可以实现这些新颖方面的背景的示例。然而,可以按照各种不同的方式来应用本文中的教示。例如,当根据具体的再现环境描述了各种实现时,本文中的教示能够广泛地适用于其他已知的再现环境以及在以后可能引入的再现环境。此外,可以在各种创作工具和/或渲染工具中实现所描述的实现,各种创作和/或渲染工具可以以多种硬件、软件、固件等来实现。因此,本公开内容的教示并不意在限于附图中示出的实现和/或本文中描述的实现,而是具有广泛的适用性。

[0057] 图1示出了具有杜比环绕声5.1配置的再现环境的示例。在二十世纪九十年代开发了杜比环绕声5.1,但该配置仍然广泛部署在影院音响系统环境中。投影仪105可以被配置成:将(例如,用于电影的)视频图像投影在屏幕150上。音频再现数据可以与视频图像同步并且由声音处理器110进行处理。功率放大器115可以将扬声器馈送信号提供给再现环境100的扬声器。

[0058] 杜比环绕声5.1配置包括左环绕阵列120和右环绕阵列125,这两个阵列中的每个阵列包括由单个声道成组驱动(gang-driven)的一组扬声器。杜比环绕声5.1配置还包括左屏幕声道130、中心屏幕声道135和右屏幕声道140的独立声道。为了低频效应(LFE)设置了用于超低音扬声器145的独立声道。

[0059] 在2010年,杜比通过引入杜比环绕声7.1提供了对数码影院音响(digital cinema sound)的增强。图2示出了具有杜比环绕声7.1配置的再现环境的示例。数字投影仪205可以被配置成接收数字视频数据并且将视频图像投影到屏幕150上。视频再现数据可以由声音处理器210进行处理。功率放大器215可以向再现环境200的扬声器提供扬声器馈送信号。

[0060] 杜比环绕声7.1配置包括左侧环绕阵列220和右侧环绕阵列225,这两个环绕阵列中的每个阵列可以由单个声道来驱动。与杜比环绕声5.1一样,杜比环绕声7.1配置包括左屏幕声道230、中心屏幕声道235、右屏幕声道240和超低音扬声器245的独立声道。然而,杜比环绕声7.1通过将杜比环绕声5.1的左环绕声道和右环绕声道分成下述四个区来增加环绕声道的数量:除了左侧环绕阵列220和右侧环绕阵列225以外,还包括用于左后环绕扬声器224的独立声道和用于右后环绕扬声器226的独立声道。增加再现环境200内环绕区的数量可以显著提高声音的定位。

[0061] 为了创建更身临其境的环境,一些再现环境可以被配置有由数量增加的声道驱动的数量增加的扬声器。此外,一些再现环境可以包括在各个高度处部署的扬声器,一些高度可以位于再现环境的座位区上方。

[0062] 图3示出了具有Hamasaki 22.2环绕声配置的再现环境的示例。日本的NHK科技研究实验室开发了Hamasaki 22.2作为超高清电视的环绕声部件。Hamasaki 22.2提供24个扬声器声道,这24个扬声器声道可以用于驱动在三个层中布置的扬声器。再现环境300的上扬声器层310可以由9个声道来驱动。中间扬声器层320可以由10个声道来驱动。低扬声器层330可以由5个声道来驱动,这5个声道中的2个声道用于超低音扬声器345a和超低音扬声器345b。

[0063] 因此,现代的趋势是不仅包括更多扬声器和更多声道,而且还包括不同高度处的扬声器。随着声道数量增加和扬声器布局从2D阵列转变成3D阵列,定位和渲染声音的任务变得越来越困难。因此,本受让人(assignee)已经开发了各种工具和相关的用户接口,所述各种工具和相关的用户接口增加了功能性和/或降低了3D音频声音系统的创作复杂度。参照于2012年4月20日提交的题为“System and Tools for Enhanced 3D Audio Authoring and Rendering”(“创作和渲染应用”)的美国临时专利申请No.61/636,102(在此通过引用将其全部内容合并到本文中)的图5A至图19D详细描述了一些这样的工具。

[0064] 图4A示出了描绘虚拟再现环境中的不同高度处的扬声器区的图形用户接口(GUI)的示例。可以根据来自逻辑系统的指令、根据从用户输入装置接收的信号等将GUI 400例如显示在显示装置上。下面参照图10描述一些这样的装置。

[0065] 如本文中参考虚拟再现环境(如虚拟再现环境404)所使用的,术语“扬声器区”一般指代可以具有或不具有与实际再现环境的再现扬声器一对一对应关系的逻辑结构。例如,“扬声器区位置”可以与或不与电影再现环境的特定再现扬声器位置相对应。替代地,术语“扬声器区位置”一般可以指代虚拟再现环境的区域。在一些实现中,虚拟再现环境的扬声器区可以例如通过使用虚拟技术(如Dolby Headphone™(有时称为Mobile Surround™))与虚拟扬声器相对应,该虚拟技术使用一组双声道立体声耳机来实时创建虚拟环绕声环境。在GUI 400中,在第一高度处存在七个扬声器区402a并且在第二高度处存在两个扬声器区402b,从而在虚拟再现环境404中形成总共九个扬声器区。在该示例中,扬声器区1至3位于虚拟再现环境404的前方区域405中。前方区域405可以与例如屏幕150所处的电影再现环境的区域相对应,与电视机屏幕所处的家庭的区域相对应等。

[0066] 在此,扬声器区4一般与左区域410中的扬声器相对应,并且扬声器区5与虚拟再现环境404的右区域415中的扬声器相对应。扬声器区6与左后区域412相对应,并且扬声器区7与虚拟再现环境404的右后区域414相对应。扬声器区8与上区域420a中的扬声器相对应,扬声器区9与上区域420b中的扬声器相对应,上区域420b可以是虚拟天花板区域。因此,如创作和渲染应用中更详细描述,图4A中所示的扬声器区1至9的位置可以与或不与实际再现环境的再现扬声器的位置相对应。此外,其他实现可以包括较多或较少扬声器区和/或高度。

[0067] 在创作和渲染应用中描述的各种实现中,用户接口(如GUI 400)可以用作创作工具和/或渲染工具的一部分。在一些实现中,创作工具和/或渲染工具可以通过存储在一个或多个非暂态介质上的软件来实现。创作工具和/或渲染工具可以(至少部分地)由硬件、固件等如下面参照图10描述的逻辑系统和其他装置来实现。在一些创作实现中,相关联的创作工具可以用于创建相关联的音频数据的元数据。元数据可以例如包括表示三维空间中的音频对象的位置和/或轨迹的数据、扬声器区约束数据等。可以相对于虚拟再现环境404的扬声器区402而不是相对于实际再现环境的特定扬声器布局来创建元数据。渲染工具可以接收音频数据和相关联的元数据,并且可以针对再现环境计算音频增益和扬声器馈送信号。可以根据幅值相移处理(amplitude panning process)来计算这样的音频增益和扬声器馈送信号,该幅值相移处理可以创建声音来自再现环境中的位置P的感知。例如,可以根据下面的等式将扬声器馈送信号提供给再现环境的再现扬声器1至再现扬声器N:

[0068] $x_i(t) = g_i x(t), i = 1, \dots, N$ (等式1)

[0069] 在等式1中, $x_i(t)$ 表示要被施加至扬声器*i*的扬声器馈送信号, g_i 表示相应声道的增益系数(gain factor), $x(t)$ 表示音频信号, t 表示时间。可以例如根据以下中描述的幅值相移方法来确定增益系数: V.Pulkki, “Compensating Displacement of Amplitude-Panned Virtual Sources” (关于虚拟、合成和娱乐音频的音频工程师协会(AES)国际会议), 第二章, 第3页至第4页, 其在此通过引用被合并到本文中。在一些实现中, 增益可能是频率相关的。在一些实现中, 通过用 $x(t-\Delta t)$ 代替 $x(t)$ 可以引入时间延迟。

[0070] 在一些渲染实现中, 参考扬声器区402创建的音频再现数据可以被映射至各种再现环境的扬声器位置, 各种再现环境可以处于杜比环绕声5.1配置、杜比环绕声7.1配置、Hamasaki 22.2配置或其他配置中。例如, 参照图2, 渲染工具可以将扬声器区4和扬声器区5的音频再现数据映射至具有杜比环绕声7.1配置的再现环境的左侧环绕阵列220和右侧环绕阵列225。可以将扬声器区1、2和3的音频再现数据分别映射至左屏幕声道230、右屏幕声道240和中心屏幕声道235。可以将扬声器区6和扬声器区7的音频再现数据映射至左后环绕扬声器224和右后环绕扬声器226。

[0071] 图4B示出了另一再现环境的示例。在一些实现中, 渲染工具可以将扬声器区1、2和3的音频再现数据映射至再现环境450的相应屏幕扬声器455。渲染工具可以将扬声器区4和5的音频再现数据映射至左侧环绕阵列460和右侧环绕阵列465, 并且可以将用于扬声器区8和9的音频再现数据映射至左头顶扬声器470a和右头顶扬声器470b。可以将用于扬声器区6和7的音频再现数据映射至左后环绕扬声器480a和右后环绕扬声器480b。

[0072] 在一些创作实现中, 创作工具可以用于创建音频对象的元数据。如上所述, 术语“音频对象”可以指代音频数据信号和相关联的元数据的流。元数据可以表示音频对象的3D位置、音频对象的表观大小、渲染约束以及内容类型(例如, 对话、效果)等。取决于实现, 元数据可以包括其他类型的数据, 例如增益数据、轨迹数据等。一些音频对象可以是静止的, 而另一些音频对象可以运动。可以根据相关联的元数据来创作或渲染音频对象细节, 该相关联的元数据在其他事物中可以表示音频对象在给定时间点在三维空间中的位置。当在再现环境中监视或回放音频对象时, 可以根据与再现环境的再现扬声器布局有关的音频对象的位置元数据和大小元数据来渲染音频对象。

[0073] 图5A是提供音频处理方法的概述的流程图。下面参照图5B以及下列等等来描述更详细的示例。这些方法可以包括比本文中所示和描述的块更多或更少的块, 而并不一定按照本文中所示的次序来执行。这些方法可以至少部分地由诸如如图10至图11B中所示和下面描述的设备等设备来执行。在一些实现中, 这些方法可以至少部分地通过存储在一个或多个非暂态介质中的软件来实现。软件可以包括用于控制一个或多个装置来执行本文中描述的方法的指令。

[0074] 在图5A所示的示例中, 方法500开始于确定与特定再现环境有关的虚拟源位置的虚拟源增益值的建立步骤(块505)。图6A示出了与再现环境有关的虚拟源位置的示例。例如, 块505可以包括: 确定与再现环境600a的再现扬声器位置625有关的虚拟源位置605的虚拟源增益值。虚拟源位置605和再现扬声器位置625仅仅是示例。在图6A所示的示例中, 虚拟源位置605沿x轴、y轴和z轴被均匀地间隔开。然而, 在替代实现中, 虚拟源位置605可以被不同地间隔开。例如, 在一些实现中, 虚拟源位置605可以具有沿x轴和y轴的第一均匀间距和沿z轴的第二均匀间距。在其他实现中, 虚拟源位置605可以被非均匀地间隔开。

[0075] 在图6A所示的示例中,再现环境600a和虚拟源空间602a是共同延伸的,以使得每个虚拟源位置605与再现环境600a内的位置相对应。然而,在替代实现中,再现环境600和虚拟源空间602可以不是共同延伸的。例如,至少一些虚拟源位置605可以与再现环境600外部的的位置相对应。

[0076] 图6B示出了与再现环境有关的虚拟源位置的替代示例。在该示例中,虚拟源空间602b延伸到再现环境600b外部。

[0077] 返回至图5A,在该示例中,在渲染任意特定音频对象之前发生块505的建立步骤。在一些实现中,可以将块505中确定的虚拟源增益值存储在存储系统中。在根据至少一些虚拟源增益值来计算接收的音频对象的音频对象增益值的“运行时”步骤期间(块510)可以使用存储的虚拟源增益值。例如,块510可以包括:至少部分地基于与音频对象区域或空间内的虚拟源位置相对应的虚拟源增益值来计算音频对象增益值。

[0078] 在一些实现中,方法500可以包括可选块515(其涉及对音频数据进行去相关)。块515可以是运行时步骤的一部分。在一些这样的实现中,块515可以包括频域中的卷积。例如,块515可以包括:对每个扬声器馈送信号施加有限脉冲响应(“FIR”)滤波器。

[0079] 在一些实现中,可以根据音频对象大小和/或创作者的艺术意图执行或不执行块515的步骤。根据一些这样的实现,通过表示(例如,通过相关联的元数据中包括的去相关标记)当音频对象大小大于或等于大小阈值时应当打开去相关而如果音频对象大小在大小阈值以下则应当关闭去相关,创作工具可以将音频对象大小与去相关相联系。在一些实现中,可以根据关于大小阈值和/或其他输入值的用户输入来控制(例如,增加、减少或禁用)去相关。

[0080] 图5B是提供建立步骤的示例的流程图。因此,图5B中所示的所有块是可以在图5A的块505中执行的步骤的示例。在此,建立步骤开始于接收到再现环境数据(块520)。再现环境数据可以包括再现扬声器位置数据。再现环境数据还可以包括表示再现环境的边界(如墙壁、天花板等)的数据。如果再现环境是电影院,则再现环境数据还可以包括电影屏幕位置的表示。

[0081] 再现环境数据还可以包括表示输出声道与再现环境的再现扬声器的相关性的数据。例如,再现环境可以具有杜比环绕声7.1配置,例如如图2所示和上面描述的配置。因此,再现环境数据还可以包括表示Lss声道与左侧环绕扬声器220之间的相关性、Lrs声道与左后环绕扬声器224之间的相关性等的的数据。

[0082] 在该示例中,块525包括根据再现环境数据限定虚拟源位置605。虚拟源位置605可以被限定在虚拟源空间内。在一些实现中,虚拟源空间可以与音频对象可以在其中移动的空间相对应。如图6A和图6B所示,在一些实现中,虚拟源空间602可以与再现环境600的空间共同延伸,而在其他实现中,至少一些虚拟源位置605可以与再现环境600外部的的位置相对应。

[0083] 此外,取决于具体实现,虚拟源位置605可以在虚拟源空间602内均匀间隔或不均匀间隔。在一些实现中,虚拟源位置605可以沿所有方向均匀间隔。例如,虚拟源位置605可以形成 $N_x \times N_y \times N_z$ 虚拟源位置605的直角坐标网。在一些实现中,N的值可以在5至100的范围内。N的值可以至少部分地取决于再现环境中的再现扬声器的数量:可以期望在每个再现扬声器位置之间包括两个或更多个虚拟源位置605。

[0084] 在其他实现中,虚拟源位置605可以具有沿x轴和y轴的第一均匀间距和沿z轴的第二均匀间距。虚拟源位置605可以形成 $N_x \times N_y \times M_z$ 虚拟源位置605的直角坐标网。例如,在一些实现中,与沿x轴或y轴相比,沿z轴可以存在较少的虚拟源位置605。在一些这样的实现中,N的值可以在10至100的范围内,而M的值可以在5至10的范围内。

[0085] 在该示例中,块530包括计算每个虚拟源位置605的虚拟源增益值。在一些实现中,块530包括:针对每个虚拟源位置605计算再现环境的多个输出声道中的每个声道的虚拟源增益值。在一些实现中,块530可以包括:应用基于矢量的幅值相移(“VBAP”)算法、成对声像调节算法或类似的算法来计算位于每个虚拟源位置605处的点源的增益值。在其他实现中,块530可以包括:应用可分离算法来计算位于每个虚拟源位置605处的点源的增益值。如本文中所使用的,“可分离”算法是下述算法:特定扬声器的增益可以被表达为可以针对虚拟源位置的每个坐标分别计算的两个或更多个因子的乘积。示例包括在各种现有混合控制台声像调节器(包括但不限于Pro Tools™软件)和由AMS Neve提供的数字电影控制台中实现的声像调节器中实现的算法。下面提供一些二维示例。

[0086] 图6C至图6F示出了将近场声像调节技术和远场声像调节技术应用于不同位置处的音频对象的示例。首先参照图6C,音频对象基本上在虚拟再现环境400a的外部。因此,在该实例中将应用一个或更多个远场声像调节方法。在一些实现中,远场声像调节方法可以基于本领域普通技术人员已知的基于矢量的幅值相移(VBAP)等式。例如,远场声像调节方法可以基于以下中描述的VBAP等式:V.Pulkki,“Compensating Displacement of Amplitude-Panned Virtual Sources”(关于虚拟、合成和娱乐音频的AES国际会议),第2.3章第4页,其在此通过引用被合并到本文中。在替代实现中,其他方法(例如涉及相应的声学平面或球面波的合成的方法)可以用于对远场音频对象和近场音频对象进行声像调节。在以下中描述了相关方法:D.de Vries,“Wave Field Synthesis”(AES专著,1999年),其在此通过引用被合并到本文中。

[0087] 现在参照图6D,音频对象610在虚拟再现环境400a内部。因此,在该实例中将应用一个或更多个近场声像调节方法。一些这样的近场声像调节方法将使用包围虚拟再现环境400a中的音频对象610的很多扬声器区。

[0088] 图6G示出了在边长等于1的正方形的每个角落处具有一个扬声器的再现环境的示例。在该示例中,x-y轴的原点(0,0)与左(L)屏幕扬声器130重合。因此,右(R)屏幕扬声器140的坐标为(1,0),左环绕(Ls)扬声器120的坐标为(0,1),右环绕(Rs)扬声器125的坐标为(1,1)。音频对象位置615(x,y)是至左扬声器的右边的x单位以及距屏幕150的y单位。在该示例中,四个扬声器中的每个扬声器接收与每个扬声器沿x轴和y轴的距离成比例的因子 $\cos/\sin$ 。根据一些实现,可以如下计算增益:

[0089] 如果 $l=L, Ls$,则 $G_l(x) = \cos(\pi/2 \cdot x)$

[0090] 如果 $l=R, Rs$,则 $G_l(x) = \sin(\pi/2 \cdot x)$

[0091] 如果 $l=L, R$,则 $G_l(y) = \cos(\pi/2 \cdot y)$

[0092] 如果 $l=Ls, Rs$,则 $G_l(y) = \sin(\pi/2 \cdot y)$

[0093] 总增益是乘积: $G_l(x,y) = G_l(x)G_l(y)$ 。通常,这些函数取决于所有扬声器的所有坐标。然而, $G_l(x)$ 不取决于源的y-位置,并且 $G_l(y)$ 不取决于其x-位置。为了说明简单计算,假设音频对象位置615为(0,0),则左扬声器的位置为 $G_L(x) = \cos(0) = 1$, $G_L(y) =$

$\cos(0) = 1$ 。总增益为乘积： $G_L(x, y) = G_L(x) G_L(y) = 1$ 。类似的计算产生 $G_{Ls} = G_{Rs} = G_R = 0$ 。

[0094] 当音频对象进入或离开虚拟再现环境400a时,可以期望在不同的声像调节模式之间进行混合。例如,当音频对象610从图6C中所示的音频对象位置615移动至图6D中所示的音频对象位置615时或者当音频对象610从图6D中所示的音频对象位置615移动至图6C中所示的音频对象位置615时,可以应用根据近场声像调节方法和远场声像调节方法计算出的增益的混合。在一些实现中,成对声像法则(pair-wise panning law)(例如,能量保持正弦或幂定律(energy-preserving sine or power law))可以用于在根据近场声像调节方法和远场声像调节方法计算的增益之间进行混合。在替代实现中,成对声像法则可以是幅值保持而不是能量保持,以使得和等于1而不是平方和等于1。还可以对由此得到的经处理的信号进行混合,以例如使用两种声像调节方法独立地处理音频信号并且使两个由此得到的音频信号交叉衰落。

[0095] 现在返回至图5B,无论在块530中使用的算法如何,都可以将由此得到的增益值存储在存储系统中(块535),以在运行时操作期间使用。

[0096] 图5C是提供根据虚拟源位置的预先计算的增益值计算接收的音频对象的增益值的运行时步骤的示例的流程图。图5C中所示的所有块是可以在图5A的块510中执行的步骤的示例。

[0097] 在该示例中,运行时步骤开始于接收到包括一个或多个音频对象的音频再现数据(块540)。在该示例中,音频对象包括音频信号和相关联的元数据,元数据包括至少音频对象位置数据和音频对象大小数据。参考图6A,例如,音频对象610至少部分地由音频对象位置615和音频对象空间620a来限定。在该示例中,接收的音频对象大小数据表示音频对象空间620a与矩形棱柱的空间相对应。然而,在该示例中,如图6B所示,接收的音频对象大小数据表示音频对象空间620b与球形的空间相对应。这些大小和形状仅仅是示例;在替代实现中,音频对象可以具有多个其他大小和/或形状。在一些替代示例中,音频对象的区域或空间可以是矩形、圆形、椭圆形、椭圆体或球扇形(spherical sector)。

[0098] 在该实现中,块545包括:计算来自音频对象位置数据和音频对象大小数据限定的区域或空间内的虚拟源的贡献。在图6A和图6B所示的示例中,块545可以包括:计算来自音频对象空间620a或音频对象空间620b内的虚拟源位置605处的虚拟源的贡献。如果音频对象的元数据随时间发生变化,则可以根据新的元数据值再次执行块545。例如,如果音频对象大小和/或音频对象位置发生变化,则不同的虚拟源位置605可以落入音频对象空间620内,和/或在先前计算中使用的虚拟源位置605可以距音频对象位置615的距离不同。在块545中,将根据新的音频对象大小和/或位置来计算相应的虚拟源贡献。

[0099] 在一些示例中,块545可以包括:从存储系统接收所计算出的、与音频对象位置和音频对象大小相对应的虚拟源位置的虚拟源增益值,并且在计算出的虚拟源增益值之间进行插值。在计算出的虚拟源增益值之间进行插值的步骤可以包括:确定音频对象位置附近的多个邻近虚拟源位置;确定每个邻近虚拟源位置的所计算出的虚拟源增益值;确定音频对象位置和每个邻近虚拟源位置之间的多个距离;并且根据所述多个距离在计算的虚拟源增益值之间进行插值。

[0100] 计算来自虚拟源的贡献的步骤可以包括:计算由音频对象的大小限定的区域或空

间内的虚拟源位置的所计算出的虚拟源增益值的加权平均值。加权平均值的权重可以取决于例如音频对象的位置、音频对象的大小和所述区域或空间内的每个虚拟源位置。

[0101] 图7示出了来自由音频对象位置数据和音频对象大小数据限定的区域内的虚拟源的贡献的示例。图7描述了垂直于z轴获得的音频环境200a的横截面。因此，从沿着z轴的、观察者向下看向音频环境200a的角度绘制了图7。在该示例中，音频环境200a是具有杜比环绕声7.1配置（例如如图2中所示和上面描述的配置）的影院音响系统环境。因此，再现环境200a包括左侧环绕扬声器220、左后环绕扬声器224、右侧环绕扬声器225、右后环绕扬声器226、左屏幕声道230、中心屏幕声道235、右屏幕声道240和超低音扬声器245。

[0102] 音频对象610具有由音频对象空间620b表示的大小，在图7中示出了音频对象空间620b的矩形横截面区域。假设在图7、图12中所示的时间的时刻下的音频对象位置615，虚拟源位置605包括在由x-y平面中的音频对象空间620b包围的区域中。取决于音频对象空间620b沿z方向的延伸和虚拟源位置605沿z轴的间距，在音频对象空间620b内可以包括或不包括另外的虚拟源位置605s。

[0103] 图7示出了来自由音频对象610的大小限定的区域或空间内的虚拟源位置605的贡献。在该示例中，用来描述每个虚拟源位置605的圆形的直径与来自相应的虚拟源位置605的贡献相对应。最靠近音频对象位置615的虚拟源位置605a被示为最大，表示来自相应的虚拟源的贡献最大。第二最大贡献来自次靠近音频对象位置615的虚拟源位置605b处的虚拟源。虚拟源位置605c做出了较小贡献，该虚拟源位置605c距音频对象位置615较远但仍然在音频对象空间620b内。在音频对象空间620b外部的虚拟源位置605d被示为最小，这表示在该示例中相应的虚拟源未做出贡献。

[0104] 返回图5C，在该示例中，块550包括：至少部分地基于计算出的贡献来计算多个输出声道中的每个输出声道的一组音频对象增益值。每个输出声道可以与再现环境的至少一个再现扬声器相对应。块550可以包括对由此得到的音频对象增益值进行归一化。例如，对于图7中所示的实现，每个输出声道可以对应于单个扬声器或一组扬声器。

[0105] 计算多个输出声道中的每个输出声道的音频对象增益值的步骤可以包括：确定要在位置 x_0, y_0, z_0 处渲染的各种大小的音频对象的增益值($g_l^{size}(x_0, y_0, z_0; s)$)。在本文中，有时可以将该音频对象增益值称为“音频对象大小贡献”。根据一些实现，音频对象增益值($g_l^{size}(x_0, y_0, z_0; s)$)可以表达为：

$$[0106] \quad \left[\sum_{x_{vs}, y_{vs}, z_{vs}} [w(x_{vs}, y_{vs}, z_{vs}; x_0, y_0, z_0; s) g_l(x_{vs}, y_{vs}, z_{vs})]^p \right]^{1/p} \quad (\text{等式 2})$$

[0107] 在等式2中， (x_{vs}, y_{vs}, z_{vs}) 表示虚拟源位置， $g_l(x_{vs}, y_{vs}, z_{vs})$ 表示虚拟源位置 x_{vs}, y_{vs}, z_{vs} 的声道l的增益值， $w(x_{vs}, y_{vs}, z_{vs}; x_0, y_0, z_0; s)$ 表示至少部分地基于音频对象的位置 (x_0, y_0, z_0) 、音频对象的大小和虚拟源位置 (x_{vs}, y_{vs}, z_{vs}) 确定的 $g_l(x_{vs}, y_{vs}, z_{vs})$ 的权重。

[0108] 在一些示例中，分量p可以具有1与10之间的值。在一些实现中，p可以是音频对象大小s的函数。例如，如果s相对较大，则在一些实现中p可以相对较小。根据一些实现，可以如下来确定p：

[0109] 如果 $s \leq 0.5$ ，则 $p=6$

[0110] 如果 $s > 0.5$, 则 $p = 6 + (-4) (s - 0.5) / (s_{\max} - 0.5)$

[0111] 其中, $s_{\max}$ 与内部放大的大小 s_{internal} 的最大值(下面描述)相对应, 并且其中, 音频对象大小 $s = 1$ 可以与具有大小(例如, 直径)等于再现环境的边界之一的长度(例如, 等于再现环境的一面墙壁的长度)的音频对象相对应。

[0112] 如果虚拟源位置沿着轴均匀分布并且如果例如如上所述加权函数和增益函数是可分离的, 则部分地根据用于计算虚拟源增益值的算法, 可以简化等式2。如果满足这些条件, 则 $g_i(x_o, y_o, z_o; s)$ 可以被表达为 $g_{1x}(x_{vs}) g_{1y}(y_{vs}) g_{1z}(z_{vs})$, 其中 $g_{1x}(x_{vs})$ 、 $g_{1y}(y_{vs})$ 和 $g_{1z}(z_{vs})$ 表示虚拟源位置的x坐标、y坐标和z坐标的独立的增益函数。

[0113] 类似地, $w(x_{vs}, y_{vs}, z_{vs}; x_o, y_o, z_o; s)$ 可以分解成因子为 $w_x(x_{vs}; x_o; s) w_y(y_{vs}; y_o; s) w_z(z_{vs}; z_o; s)$, 其中, $w_x(x_{vs}; x_o; s)$ 、 $w_y(y_{vs}; y_o; s)$ 和 $w_z(z_{vs}; z_o; s)$ 表示虚拟源位置的x坐标、y坐标和z坐标的独立的加权函数。在图7中示出了一个这样的示例。在该示例中, 可以独立于被表达为 $w_y(y_{vs}; y_o; s)$ 的加权函数720来计算被表达为 $w_x(x_{vs}; x_o; s)$ 的加权函数710。在一些实现中, 加权函数710和加权函数720可以是高斯函数, 而加权函数 $w_z(z_{vs}; z_o; s)$ 可以是余弦函数和高斯函数的乘积。

[0114] 如果 $w(x_{vs}, y_{vs}, z_{vs}; x_o, y_o, z_o; s)$ 可以分解成因子 $w_x(x_{vs}; x_o; s) w_y(y_{vs}; y_o; s) w_z(z_{vs}; z_o; s)$, 则等式2简化成:

[0115] $[f_1^x(x_o; s) f_1^y(y_o; s) f_1^z(z_o; s)]^{1/p}$, 其中

[0116] $f_1^x(x_o; s) = \sum_{x_{vs}} [g_1(x_{vs}) w(x_{vs}; x_o; s)]^p$,

[0117] $f_1^y(y_o; s) = \sum_{y_{vs}} [g_1(y_{vs}) w(y_{vs}; y_o; s)]^p$ 以及

[0118] $f_1^z(z_o; s) = \sum_{z_{vs}} [g_1(z_{vs}) w(z_{vs}; z_o; s)]^p$.

[0119] 函数 f 可以包括关于虚拟源的所有需要的信息。如果沿每个轴对可能的对象位置进行离散化, 则可以将每个函数 f 表示为矩阵。在块505的建立步骤期间(参见图5A)可以预先计算每个函数 f 并且将每个函数 f 例如作为矩阵或作为查找表存储在存储系统中。在运行时(块510), 可以从存储系统检索查找表或矩阵。运行时步骤可以包括: 考虑到音频对象位置和音频对象大小, 在这些矩阵的最靠近的相应值之间进行插值。在一些实现中, 插值可以是线性的。

[0120] 在一些实现中, 音频对象大小贡献 g_1^{size} 可以与音频对象位置的“音频对象近场增益(audio object neargain)”结果结合。如本文中所使用的, “音频对象近场增益”是基于音频对象位置615所计算出的增益。可以使用用于计算每个虚拟源增益值的同一算法来执行增益计算。根据一些这样的实现, 可以例如作为音频对象大小的函数在音频对象大小贡献和音频对象近场增益结果之间执行交叉衰落计算。这样的实现可以提供音频对象的平滑声像调节和平滑增长, 并且可以允许最小音频对象大小和最大音频对象大小之间的平滑转变。在一个这样的实现中,

[0121] $g_i^{\text{total}}(x_o, y_o, z_o; s) = \alpha(s) g_i^{\text{neargain}}(x_o, y_o, z_o; s) + \beta(s) \tilde{g}_i^{\text{size}}(x_o, y_o, z_o; s)$, 其中 s

$< S_{\text{xfade}}$, $\alpha = \cos((s/S_{\text{xfade}})(\pi/2))$, $\beta = \sin((s/S_{\text{xfade}})(\pi/2))$ $s \geq S_{\text{xfade}}$, $\alpha = 0$, $\beta = 1$,

[0122] 并且其中, $\tilde{g}_i^{\text{size}}$ 表示先前计算的 g_1^{size} 的归一化版本。在一些这样的实现中, S_{xfade}

$=0.2$ 。然而,在替代实现中, S_{xfade} 可以具有其他值。

[0123] 根据一些实现,在其可能值的范围的较大部分中可以增大音频对象大小值。在一些创作实现中,例如,用户可以面临音频对象大小值 $S_{\text{user}} \in [0, 1]$,该音频对象大小值 $S_{\text{user}} \in [0, 1]$ 被映射成由算法用于较大范围例如范围 $[0, S_{\text{max}}]$ 的实际大小,其中, $S_{\text{max}} > 1$ 。该映射可以确保:当用户将大小设置成最大时,增益变得真正独立于对象的位置。根据一些这样的实现,可以根据连接多对点 $(S_{\text{user}}, S_{\text{internal}})$ 的分段线性函数(piece-wise linear function)来进行这样的映射,其中, S_{user} 表示用户选择的音频对象大小, S_{internal} 表示由算法确定的相应的音频对象大小。根据一些这样的实现,可以根据连接多对点 $(0, 0)$ 、 $(0.2, 0.3)$ 、 $(0.5, 0.9)$ 、 $(0.75, 1.5)$ 和 $(1, S_{\text{max}})$ 的分段线性函数来进行映射。在一个这样的实现中, $S_{\text{max}} = 2.8$ 。

[0124] 图8A和图8B示出了再现环境内的两个位置中的音频对象。在这些示例中,音频对象空间620b是半径小于再现环境200a的长度或宽度的二分之一的球形。根据杜比7.1来配置再现环境200a。在图8A中描绘的时间的时刻下,音频对象位置615相对更靠近再现环境200a的中心。在图8B中描绘的时间下,音频对象位置615已经移动靠近再现环境200a的边界。在该示例中,边界是电影院的左墙壁并且与左侧环绕扬声器220的位置重合。

[0125] 出于美学原因,期望可以对接近再现环境的边界的音频对象的音频对象增益计算进行修改。在图8A和图8B中,例如,当音频对象位置615在距再现环境的左边界805的阈值距离内时,无扬声器馈送信号被提供至再现环境的相对边界上的扬声器(在此,右侧环绕扬声器225)。在图8B所示的示例中,如果音频对象位置615也大于距屏幕的阈值距离,则当音频对象位置615在距再现环境的左边界805的阈值距离(其可以是不同的阈值距离)内时,无扬声器馈送信号被提供至与左屏幕声道230、中心屏幕声道235、右屏幕声道240相对应的扬声器或超低音扬声器245。

[0126] 在图8B所示的示例中,音频对象空间620b包括左边界805外部的区域或空间。根据一些实现,增益计算的衰落因子可以至少部分地基于多少左边界805处在音频对象空间620b内和/或音频对象的多少区域或空间延伸到这样的边界的外部。

[0127] 图9是略述至少部分地基于音频对象的多少区域或空间延伸到再现环境的边界的外部来确定衰落因子的方法的流程图。在块905中,接收再现环境数据。在该示例中,再现环境数据包括再现扬声器位置数据和再现环境边界数据。块910包括接收包括一个或多个音频对象和相关联的元数据的音频再现数据。在该示例中,元数据包括至少音频对象位置数据和音频对象大小数据。

[0128] 在该实现中,块915包括:确定由音频对象位置数据和音频对象大小数据限定的音频对象区域或空间包括再现环境边界外部的区域或空间。块915还可以包括:确定多大比例的音频对象区域或空间处在再现环境边界的外部。

[0129] 在块920中,确定衰落因子。在该示例中,衰落因子可以至少部分地基于外部的区域。例如,衰落因子可以与外部区域成比例。

[0130] 在块925中,可以至少部分地基于相关联的元数据(在该示例中,音频对象位置数据和音频对象大小数据)和衰落因子来计算多个输出声道中的每个输出声道的一组音频对象增益值。每个输出声道可以与再现环境的至少一个再现扬声器相对应。

[0131] 在一些实现中,音频对象增益计算可以包括:计算来自音频对象区域或空间内的虚拟源的贡献。虚拟源可以与可以参考再现环境数据限定的多个虚拟源位置相对应。虚拟

源位置可以被均匀地间隔开或被不均匀地间隔开。对于每个虚拟源位置,可以计算多个输出声道中的每个输出声道的虚拟源增益值。如上所述,在一些实现中,在建立步骤期间可以计算并存储这些虚拟源增益值,然后对这些虚拟源增益值进行检索以供在运行时操作期间使用。

[0132] 在一些实现中,可以将衰落因子应用于与再现环境内的虚拟源位置相对应的所有虚拟源增益值。在一些实现中,可以如下修改 g_l^{size} :

$$[0133] \quad g_l^{size} = \left[g_l^{bound} + (\text{衰落因子}) \times g_l^{inside} \right]^{1/p}, \quad \text{其中}$$

[0134] 如果 $d_{bound} \geq s$,则衰落因子=1,

[0135] 如果 $d_{bound} < s$,则衰落因子= d_{bound}/s ,

[0136] 其中, d_{bound} 表示音频对象位置和再现环境的边界之间的最小距离,并且 g_l^{bound} 表示沿边界的虚拟源的贡献。例如,参考图8B, g_l^{bound} 可以表示在音频对象空间620b内并且靠近边界805的虚拟源的贡献。在该示例中,与图6A的情况一样,不存在位于再现环境外部的虚拟源。

[0137] 在替代实现中,可以如下修改 g_l^{size} :

$$[0138] \quad g_l^{size} = \left[g_l^{outside} + (\text{衰落因子}) \times g_l^{inside} \right]^{1/p},$$

[0139] 其中, $g_l^{outside}$ 表示基于位于再现环境外部但是在音频对象区域或空间内的虚拟源的音频对象增益。例如,参考图8B, $g_l^{outside}$ 可以表示在音频对象空间620b内并且在边界805外部的虚拟源的贡献。在该示例中,与图6B的情况一样,在再现环境内部和再现环境外部两者均存在虚拟源。

[0140] 图10是提供创作和/或渲染设备的部件的示例的框图。在该示例中,装置1000包括接口系统1005。接口系统1005可以包括网络接口(例如无线网络接口)。可替代地或另外,接口系统1005可以包括通用串行总线(USB)接口或其他这样的接口。

[0141] 装置1000包括逻辑系统1010。逻辑系统1010可以包括处理器(如通用单芯片处理器或通用多芯片处理器)。逻辑系统1010可以包括数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或其他可编程逻辑装置、分立门或晶体管逻辑或分立硬件部件或者其组合。逻辑系统1010可以被配置成控制装置1000的其他部件。虽然在图10中示出了装置1000的部件之间没有接口,但是逻辑系统1010可以被配置有用于与其他部件进行通信的接口。其他部件可以被配置成或不被配置成当需要时彼此进行通信。

[0142] 逻辑系统1010可以被配置成:执行音频创作和/或渲染功能,包括但不限于本文中描述的类型音频创作和/或渲染功能。在一些这样的实现中,逻辑系统1010可以被配置成:(至少部分地)根据存储在一个或多个非暂态介质中的软件进行操作。非暂态介质可以包括与逻辑系统1010相关联的存储器(如随机存取存储器(RAM)和/或只读存储器(ROM))。非暂态介质可以包括存储系统1015的存储器。存储系统1015可以包括一个或多个适当类型的非暂态存储介质(如闪存存储器、硬盘驱动器等)。

[0143] 显示系统1030可以包括一个或多个适当类型的显示器,这取决于装置1000的表现形式。例如,显示系统1030可以包括液晶显示器、等离子显示器、双稳态显示器等。

[0144] 用户输入系统1035可以包括被配置成接收来自用户的输入的一个或多个装置。

在一些实现中,用户输入系统1035可以包括覆盖显示系统1030的显示器的触摸屏。用户输入系统1035可以包括鼠标、跟踪球、手势检测系统、操纵杆、存在于显示系统1030上的一个或更多个GUI和/或菜单、按钮、键盘、开关等。在一些实现中,用户输入系统1035可以包括麦克风1025:用户可以经由麦克风1025向装置1000提供语音命令。逻辑系统可以被配置成:用于语音识别并且用于根据这样的语音命令控制装置1000的至少一些操作。

[0145] 电力系统1040可以包括一个或更多个适当的储能装置(如镍镉电池或锂离子电池)。电力系统1040可以被配置成接收来自电气插座的电力。

[0146] 图11A是表示可以用于音频内容创建的一些部件的框图。系统1100可以例如用于混音室(mixing studio)和/或配音阶段中的音频内容创建。在该示例中,系统1100包括音频和元数据创作工具1105和渲染工具1110。在该实现中,音频和元数据创作工具1105和渲染工具1110分别包括音频连接接口1107和1112,音频连接接口1107和1112可以被配置成经由AES/EBU、MADI、会话等进行通信。音频和元数据创作工具1105和渲染工具1110分别包括网络接口1109和1117,网络接口1109和1117可以被配置成通过TCP/IP或任何其他合适的协议来发送和接收元数据。接口1120被配置成将音频数据输出至扬声器。

[0147] 系统1100可以例如包括现有创作系统(如Pro Tools™)系统,其将元数据创建工具(即,如本文中描述的声像调节器)作为插件运行。声像调节器也可以在与渲染工具1110连接的独立系统(例如,PC或混合控制台)上运行,或者可以在与渲染工具1110相同的物理装置上运行。在后一种情况下,声像调节器和渲染器可以例如通过共享的存储器来使用本地连接。也可以将声像调节器GUI设置在平板装置、膝上型电脑等上。渲染工具1110可以包括渲染系统,该渲染系统包括被配置成用于实现与图5A至图5C和图9中描述的方法一样的渲染方法的声音处理器。渲染系统可以包括例如包括用于音频输入/输出的接口和适当的逻辑系统的个人计算机、膝上型电脑等。

[0148] 图11B是表示可以用于再现环境(例如,电影院)中的音频回放的一些部件的框图。在该示例中,系统1150包括影院服务器1155和渲染系统1160。影院服务器1155和渲染系统1160分别包括网络接口1157和1162,网络接口1157和1162可以被配置成通过TCP/IP或任何其他适当的协议来发送和接收音频对象。接口1164被配置成将音频数据输出至扬声器。

[0149] 对本领域的普通技术人员而言,对在本公开内容中描述的实现的的各种修改容易是显而易见的。在不背离本公开内容的精神或范围的情况下,本文中限定的一般原理可以适用于其他实现。因此,权利要求并不意在限于本文中所示的实现,但是意在使本公开内容、本文中公开的原理和新颖的特征与最广的范围一致。

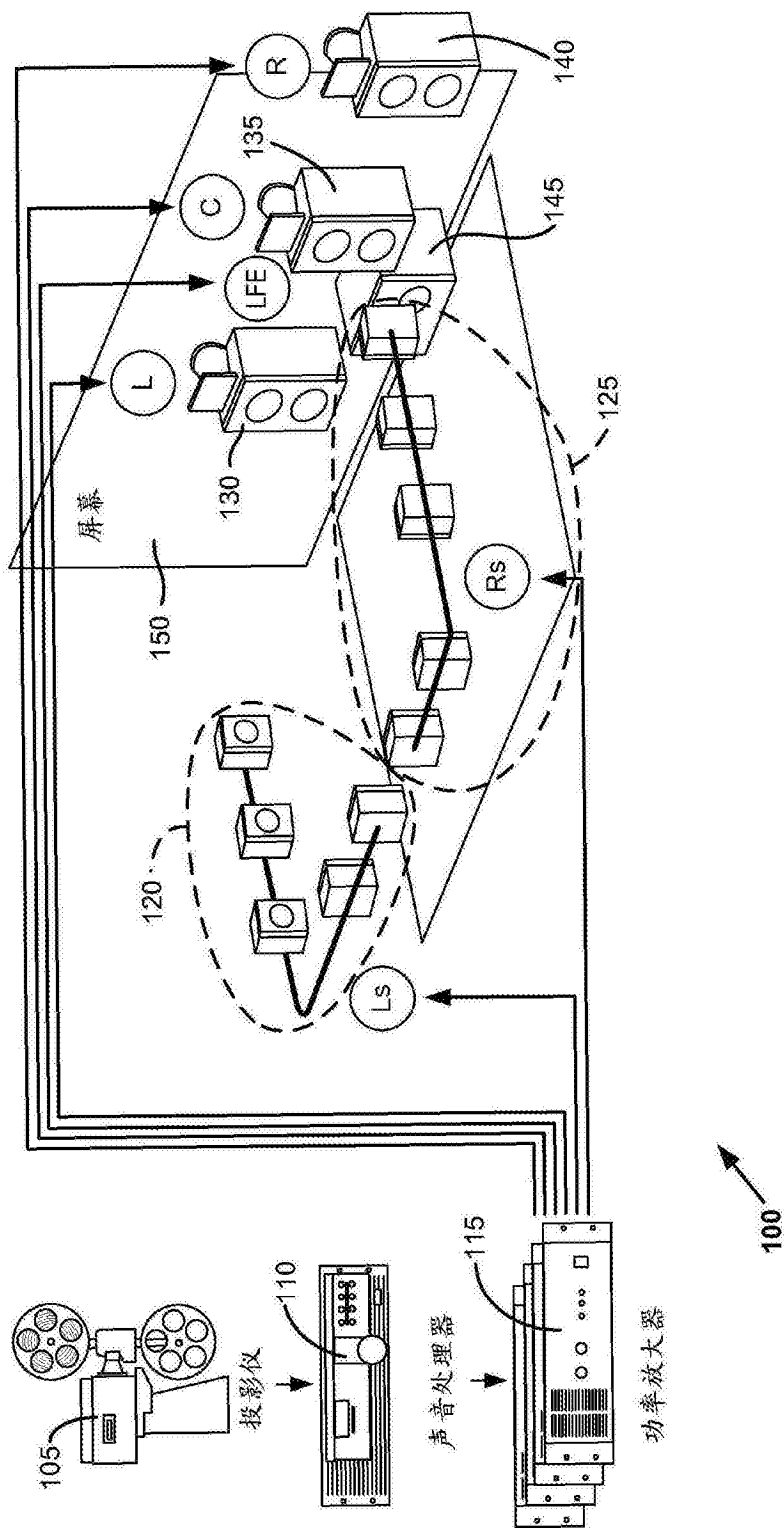


图1

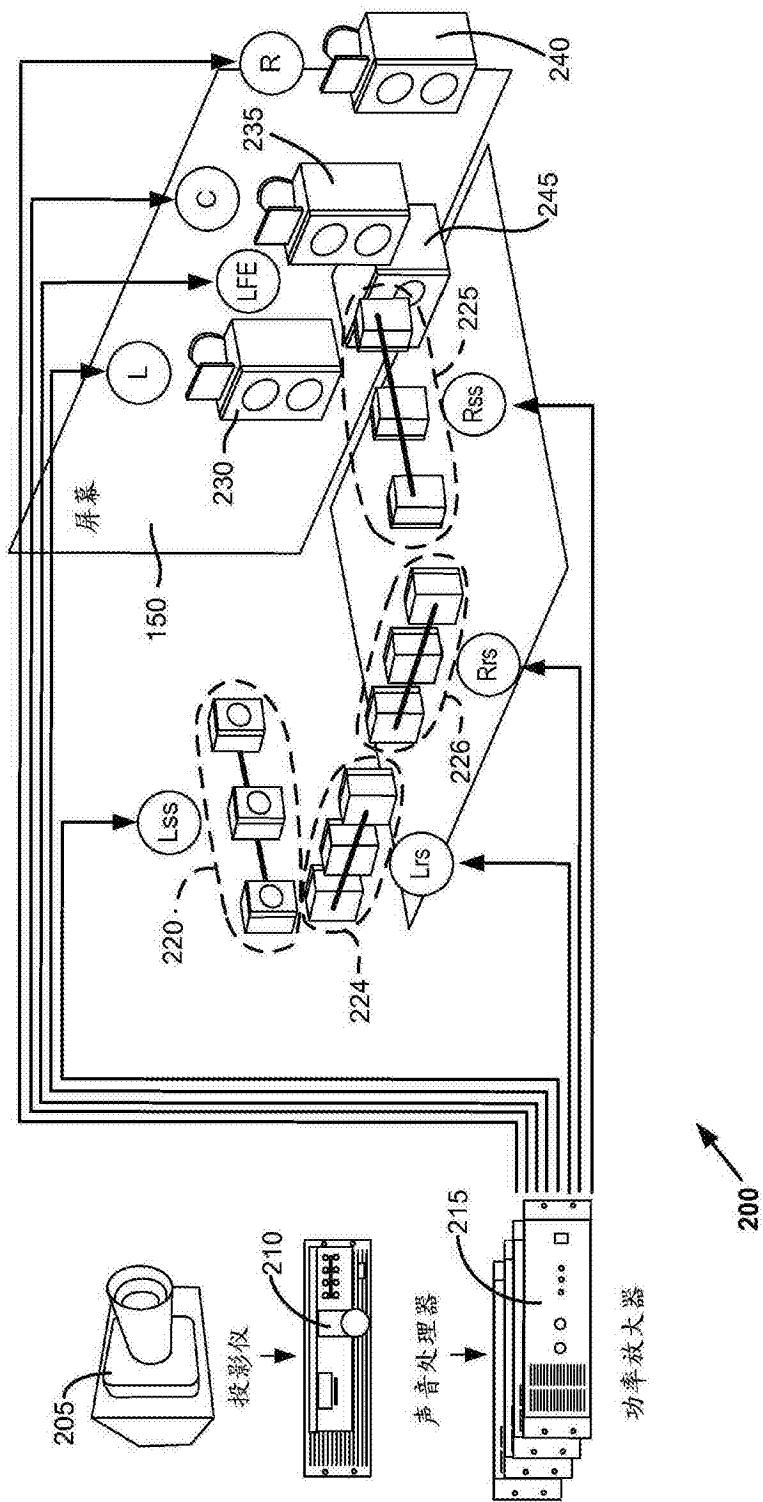


图2

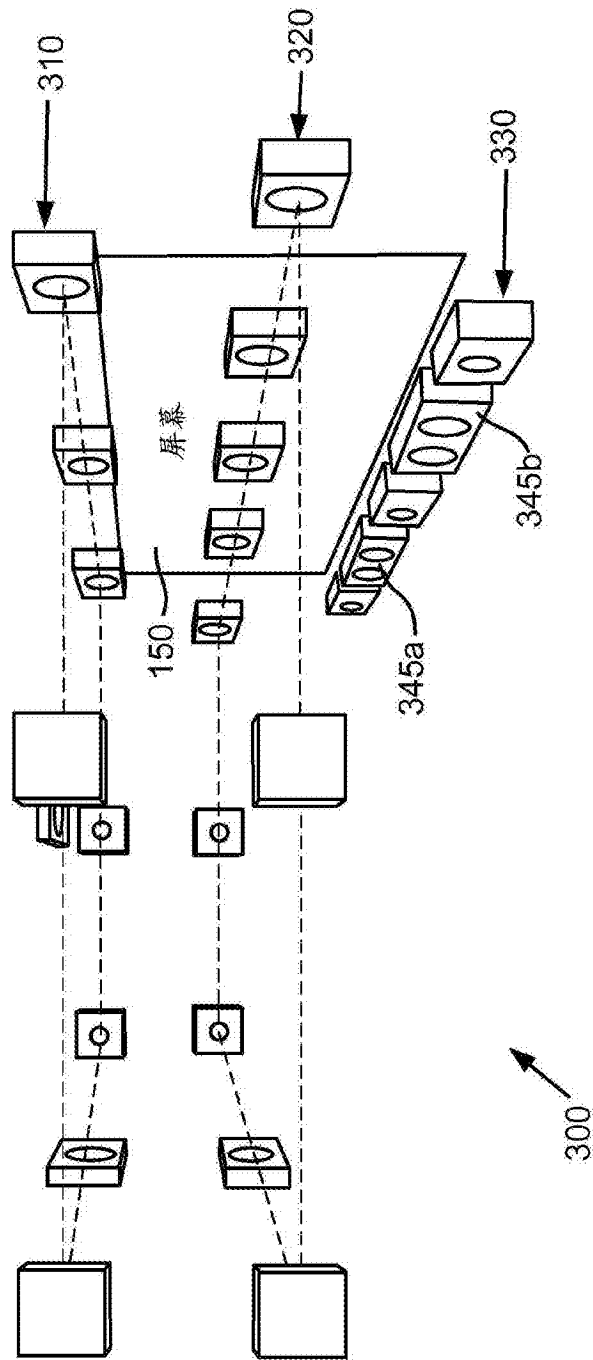


图3

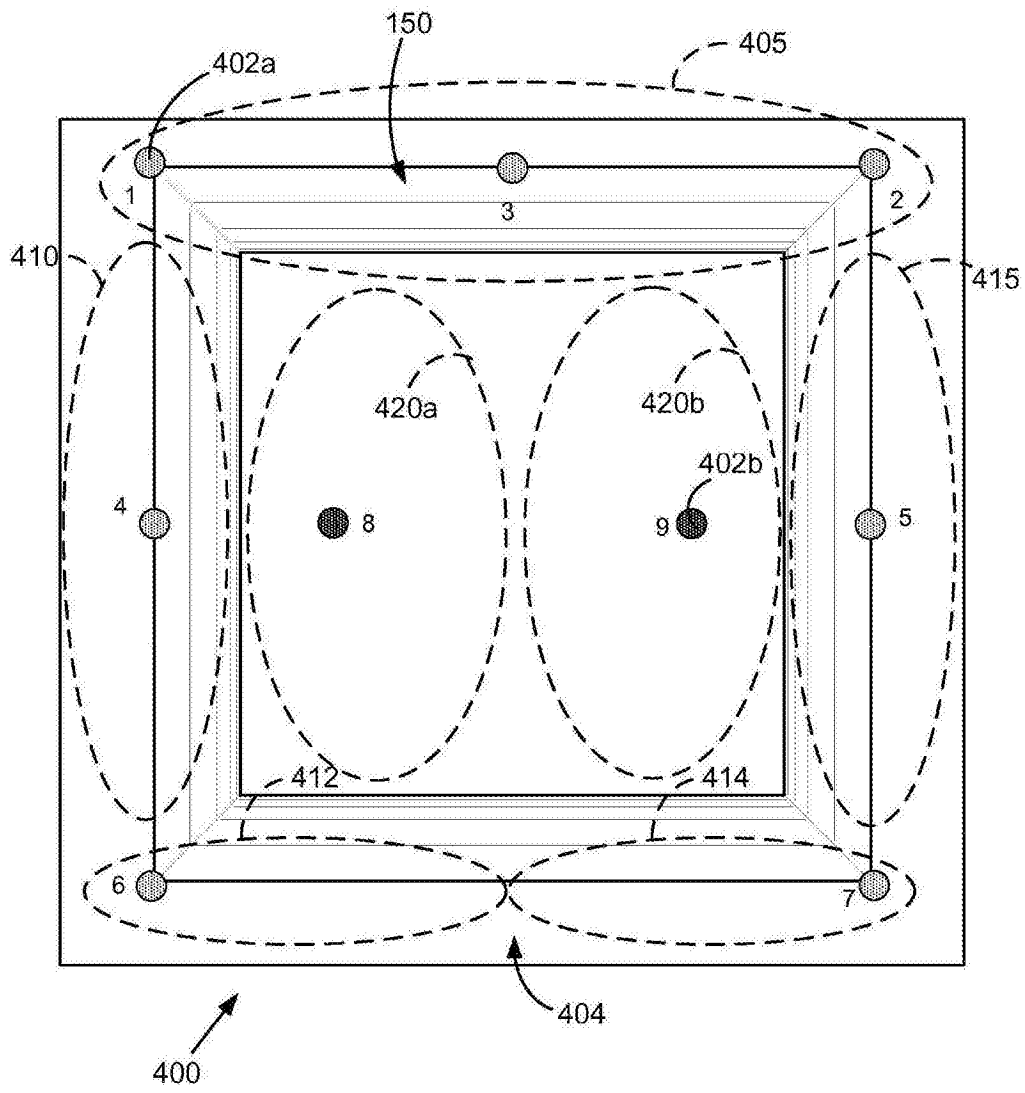


图4A

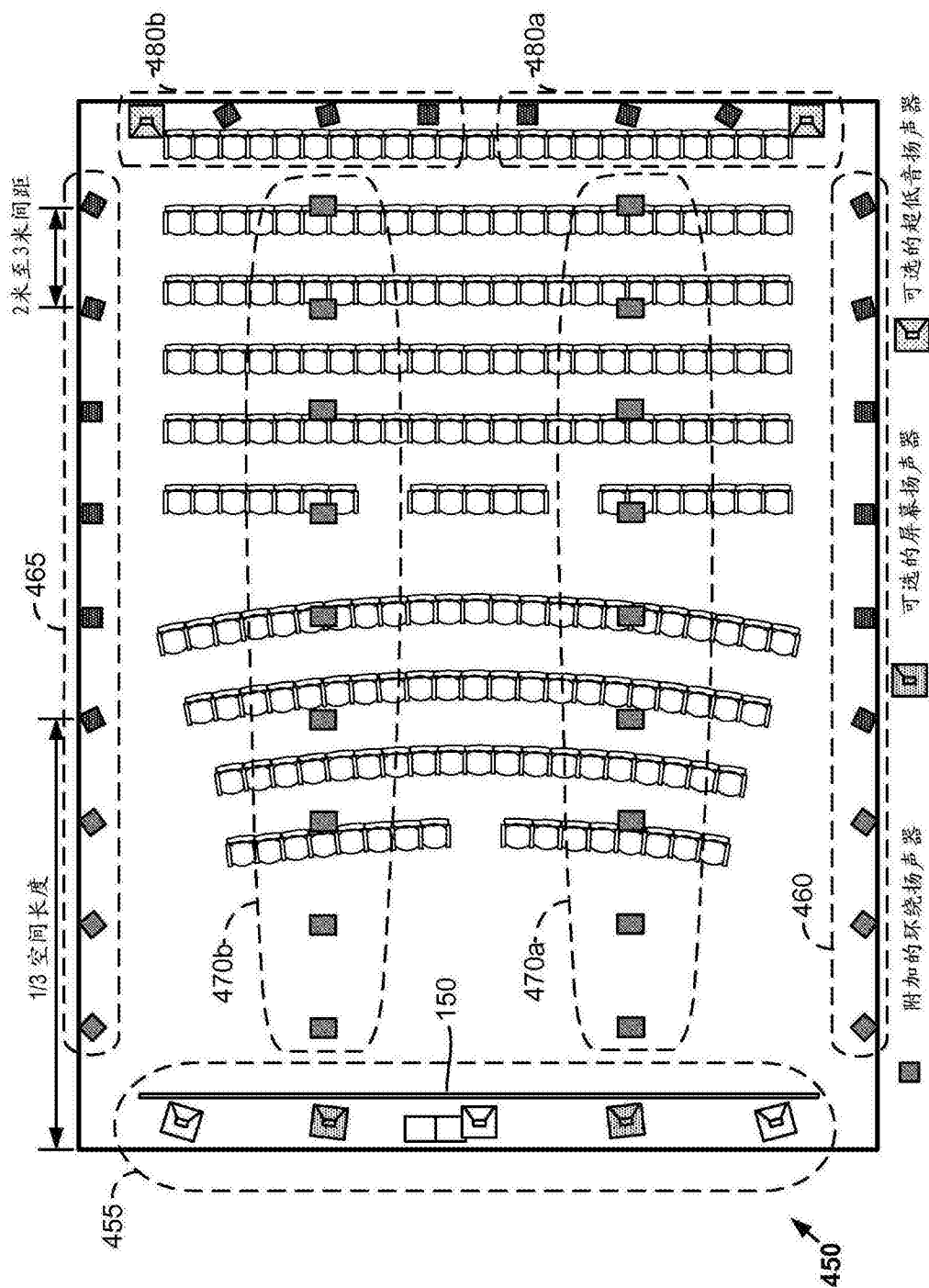


图4B

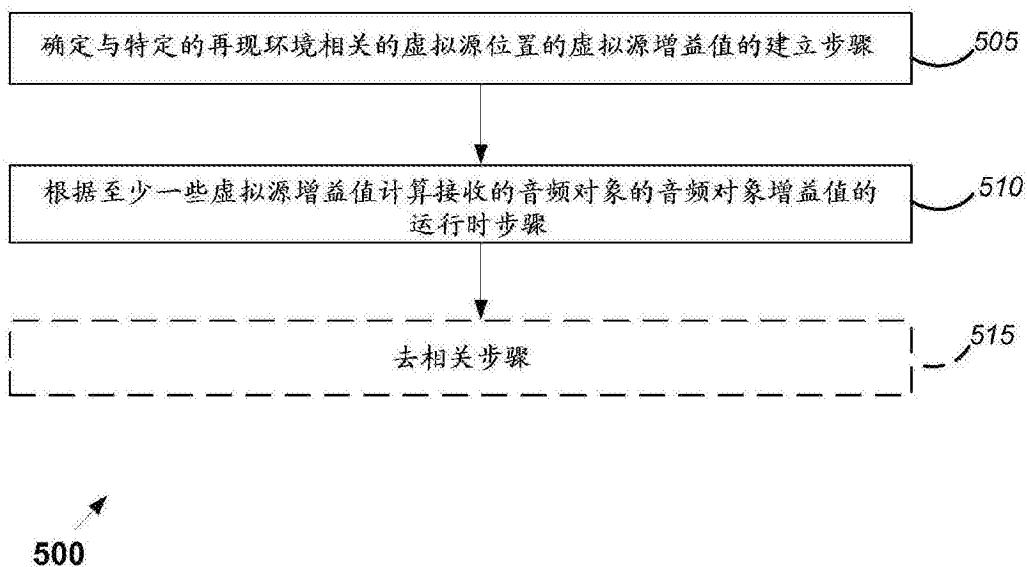


图5A

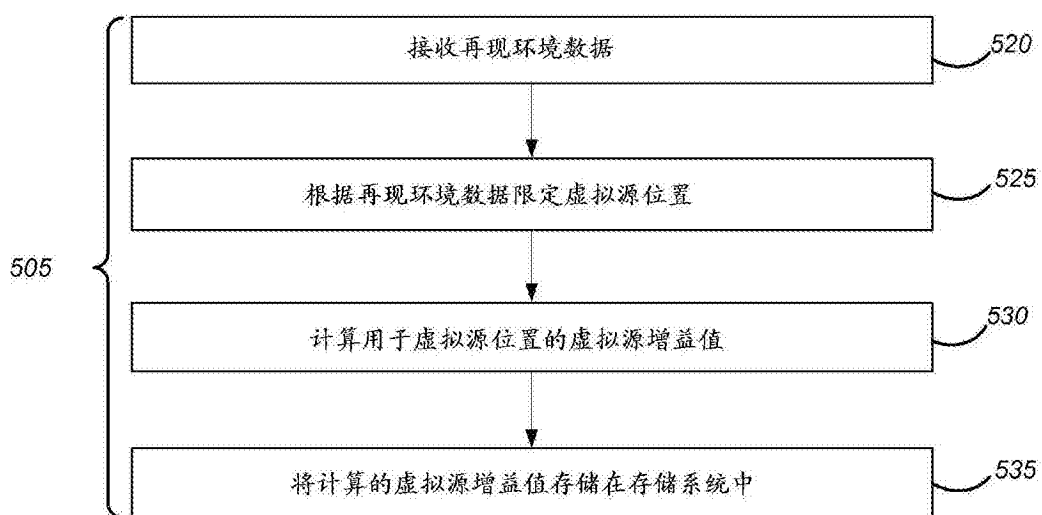


图5B

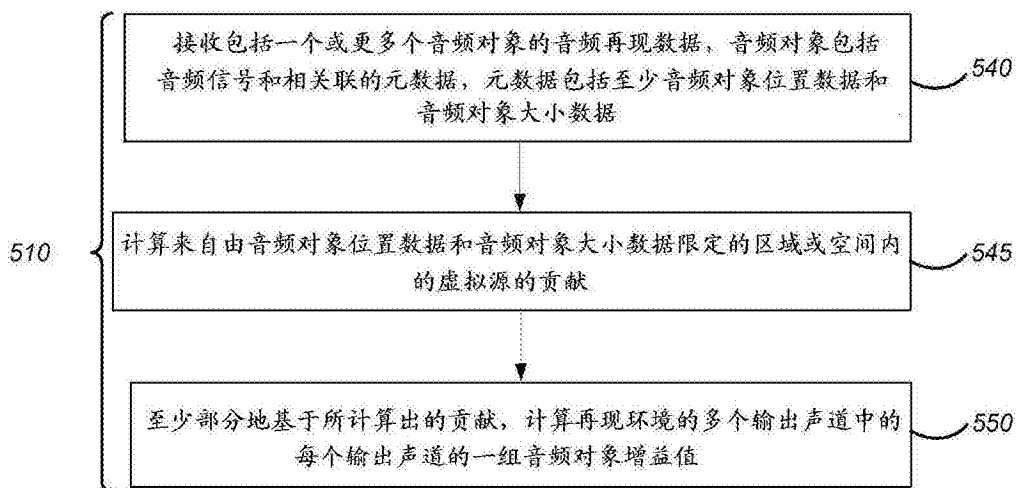


图5C

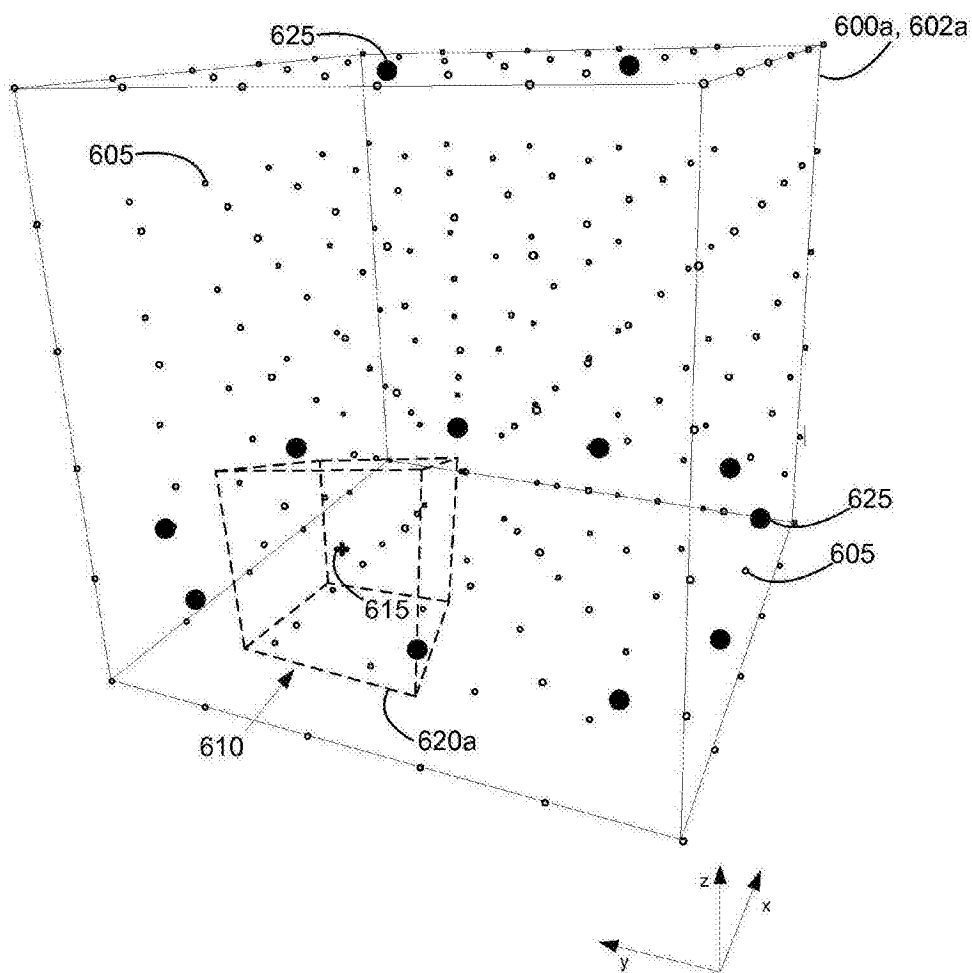


图6A

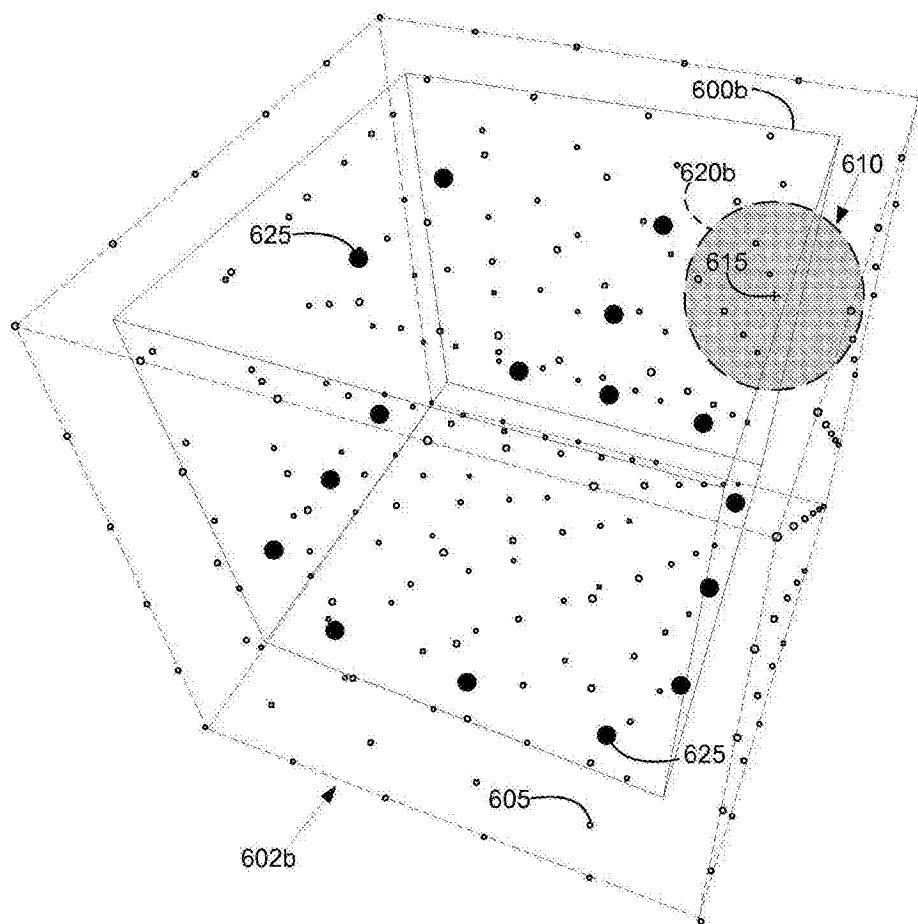


图6B

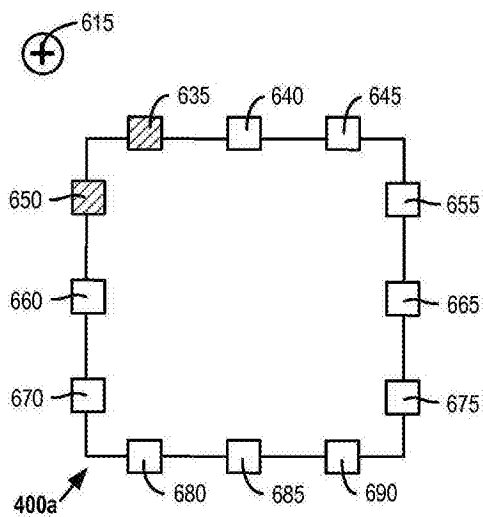


图6C

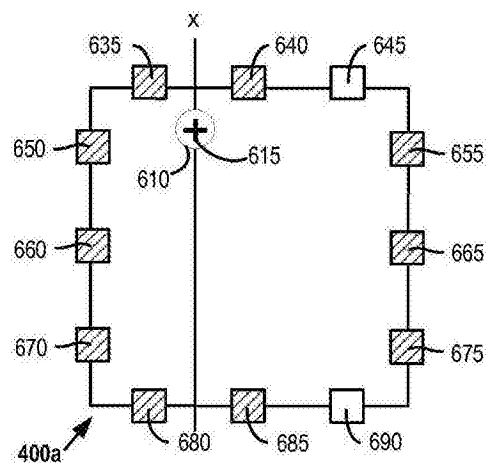


图6D

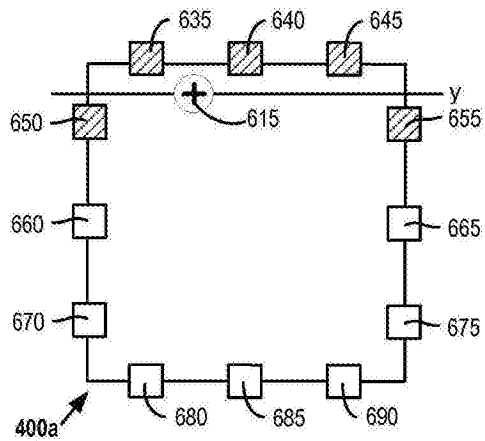


图6E

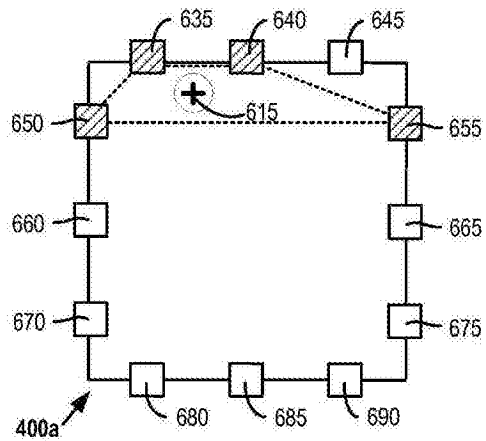


图6F

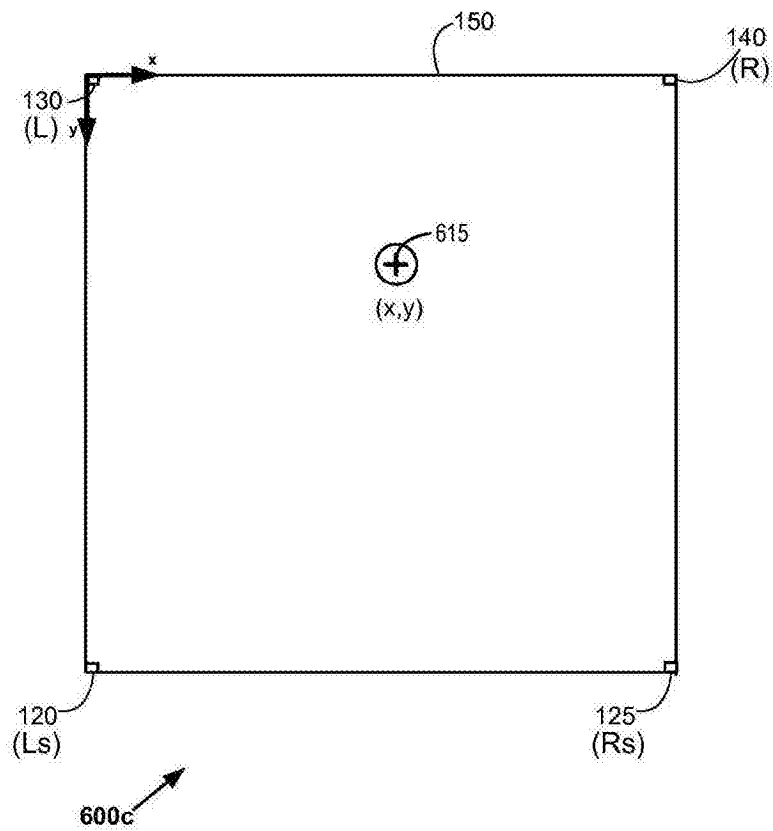


图6G

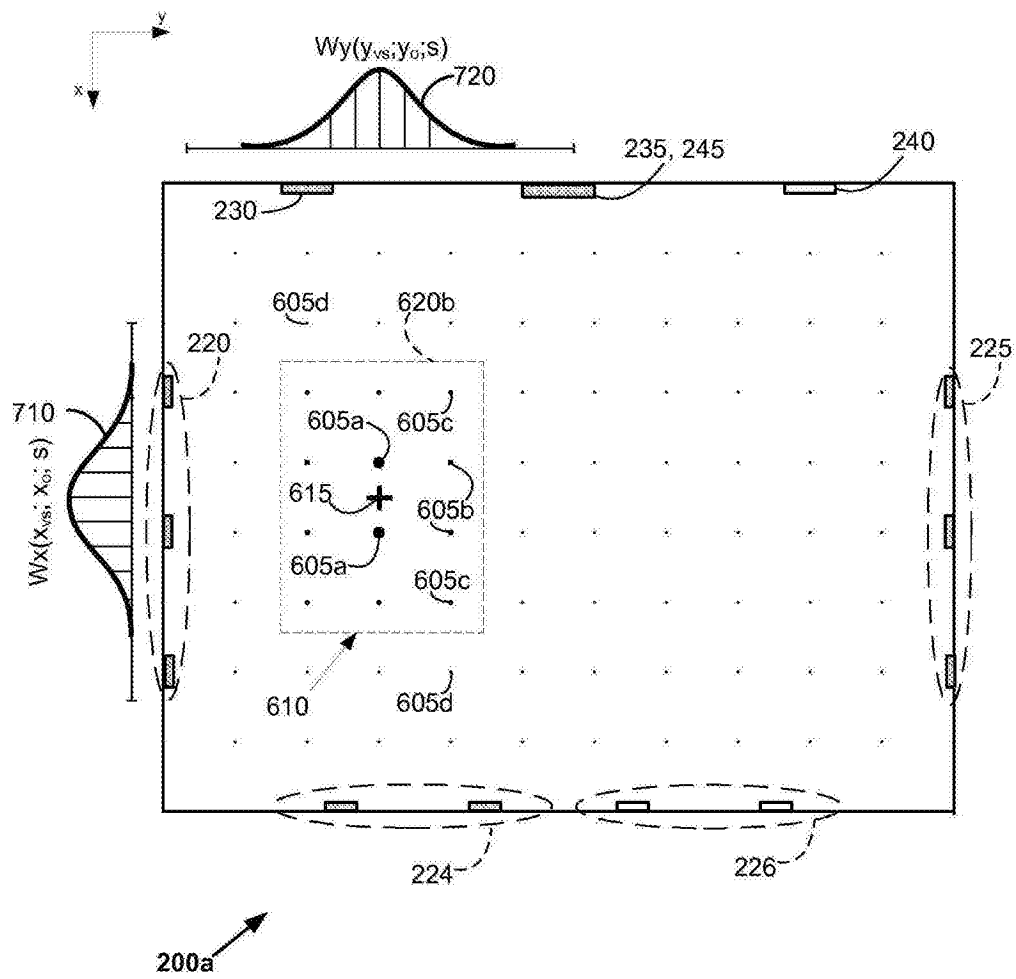


图7

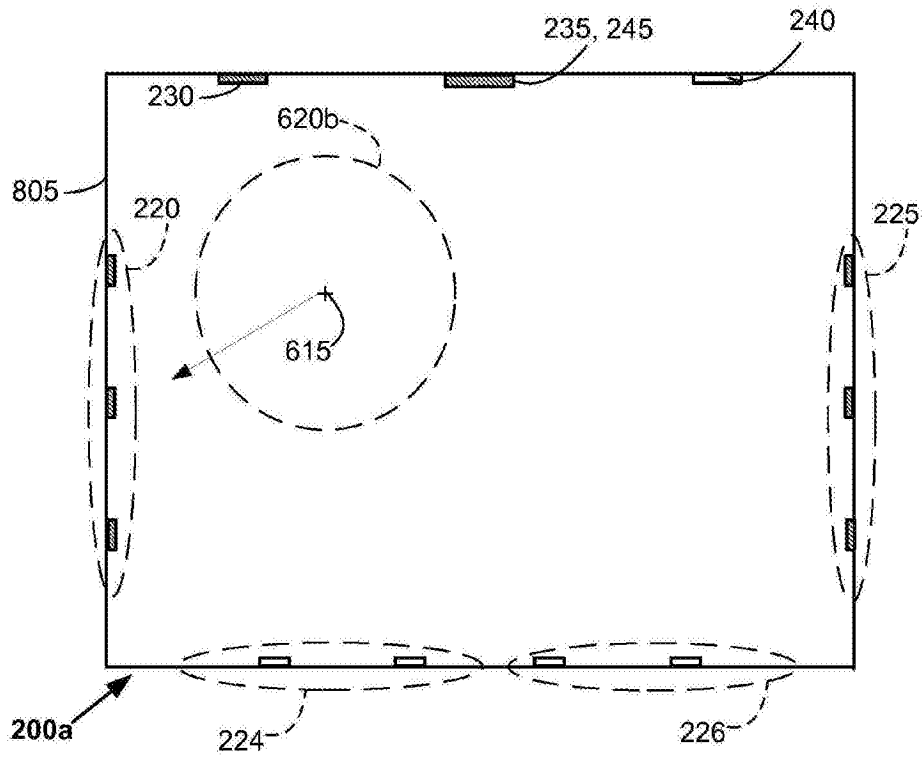


图8A

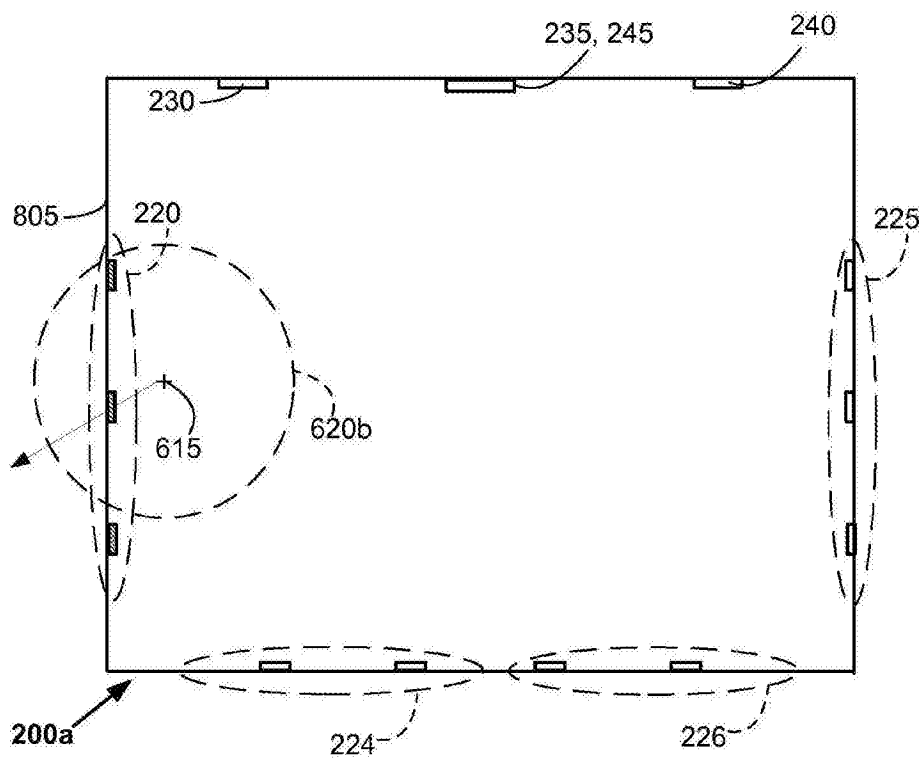


图8B

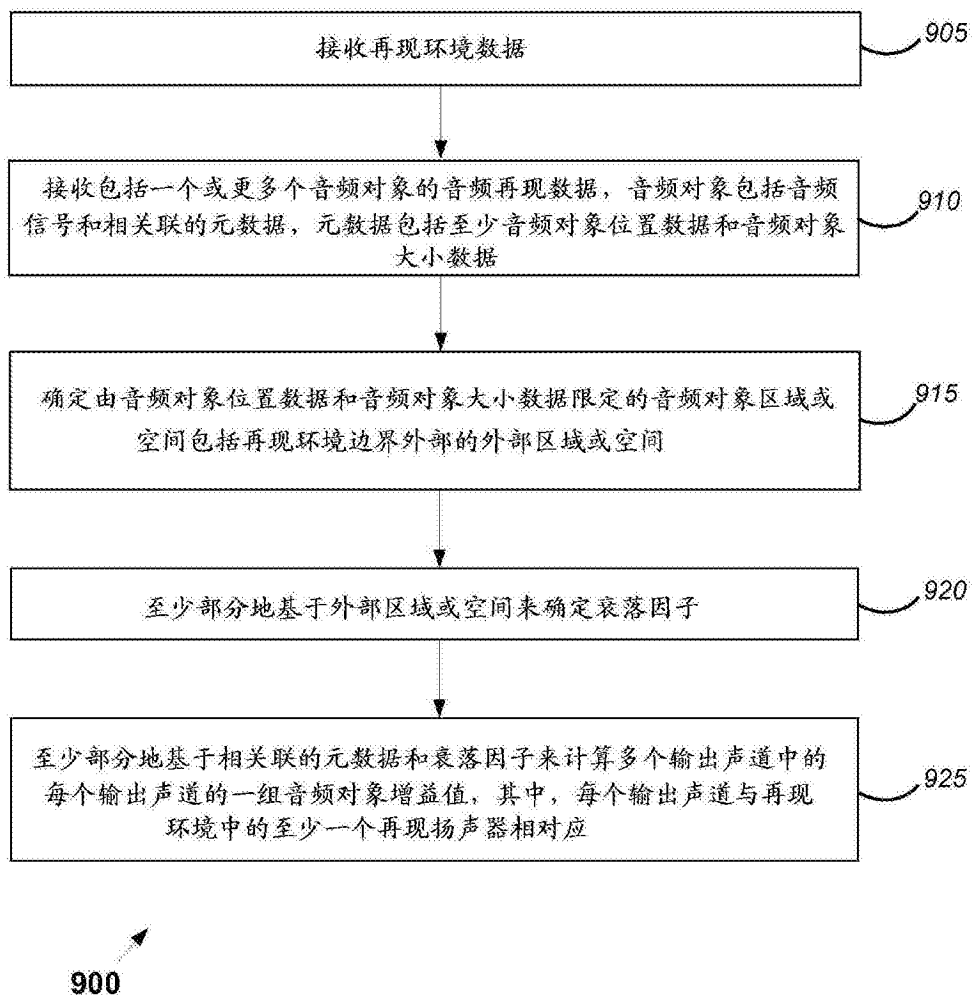


图9

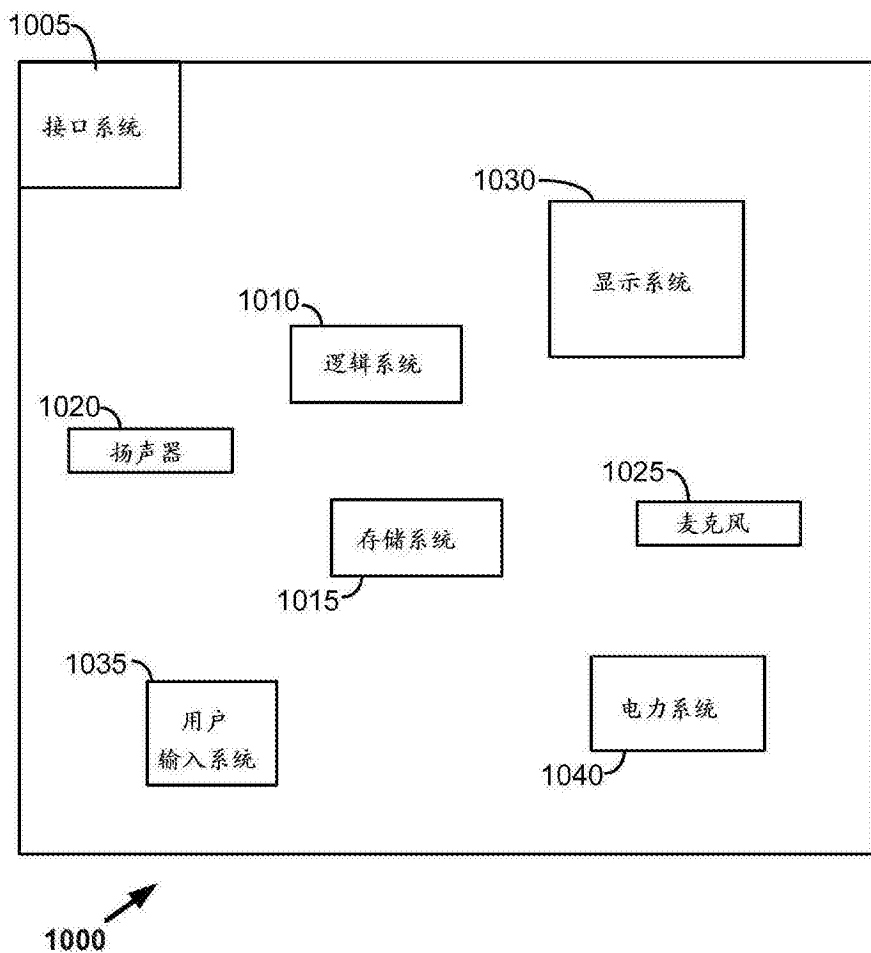


图10

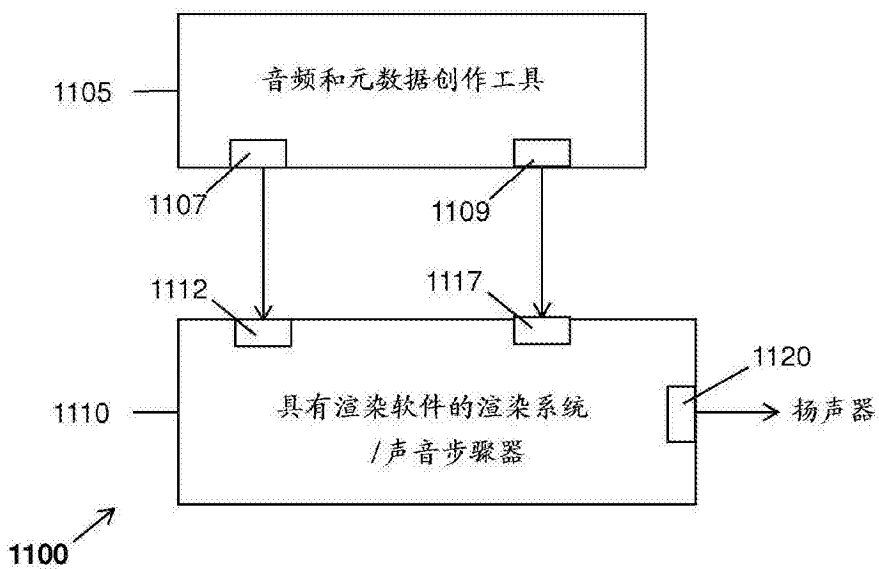


图11A

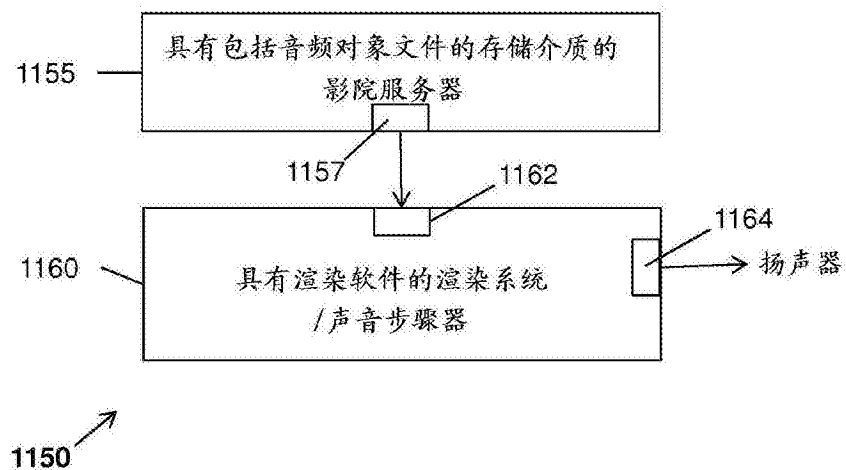


图11B