



(12)发明专利申请

(10)申请公布号 CN 110286758 A

(43)申请公布日 2019.09.27

(21)申请号 201910521191.9

H04R 3/00(2006.01)

(22)申请日 2013.03.20

H04S 7/00(2006.01)

(30)优先权数据

61/619,202 2012.04.02 US

13/775,720 2013.02.25 US

(62)分案原申请数据

201380017323.5 2013.03.20

(71)申请人 高通股份有限公司

地址 美国加利福尼亚州

(72)发明人 向佩 埃里克·维瑟

(74)专利代理机构 北京律盟知识产权代理有限

责任公司 11287

代理人 杨林勳

(51)Int.Cl.

G06F 3/01(2006.01)

权利要求书2页 说明书22页 附图46页

(54)发明名称

用于声场的手势操纵的系统、方法、和设备

(57)摘要

本申请涉及用于声场的手势操纵的系统、方法、和设备。本发明描述对所产生声场的手势响应性修改。

表示的手势	对应的命令
GA1) 阻挡	CA1) 减小束强度
GA2) 招手	CA2) 增加束强度
GB1) 闭合手, 平移, 打开	CB1) 改变束方向
GB2) 将手带到一起, 平移, 使手移动分开	CB2) 改变束方向
GC1) 使手水平地朝向彼此移动	CC1) 减小束宽度
GC2) 使手水平地移动远离彼此	CC2) 增加束宽度
GD1) 将手举起到耳朵	CD1) 产生束
GD2) 摆手	CD2) 删除束
GE1) 一起举起手	CE1) 增加声场强度
GE2) 一起放下手	CE2) 减小声场强度
GF1) 使手垂直地朝向彼此移动	CF1) 压缩(例如, 场或束的)动态范围
GF2) 使手垂直地移动远离彼此	CF2) 扩展(例如, 场或束的)动态范围

1. 一种装置,包括:
一或多个处理器,其经配置以:
俘获用于辨识声场内的用户的移动的图像序列,所述声场由扬声器阵列产生;
应用所述用户的所辨识的移动来确定所述声场的所指示改变;
合成由所述扬声器阵列产生的经修改的声场以实施所述所指示改变;以及
存储器,其经配置以存储用于辨识所述用户的所述移动的所述图像序列。
2. 根据权利要求1所述的装置,其中所述用户的所述所辨识的移动表示手势,其中所述手势被解释为多个单独模式中的一个,且由所述一或多个处理器做出决定以合成与所述单独模式相关联的相应声场。
3. 根据权利要求1所述的装置,还包括用于检测和定位所关注区域的特征提取。
4. 根据权利要求3所述的装置,其中所述所关注区域包括用户的眼睛。
5. 根据权利要求3所述的装置,其中所述所关注区域包括用户的手。
6. 根据权利要求3所述的装置,其中所述所关注区域包括用户的嘴巴。
7. 根据权利要求3所述的装置,其中所述所关注区域包括用户的身体。
8. 根据权利要求3所述的装置,进一步包括辨识所检测特征的定向或在图像序列内随着时间的改变,所述图像序列表示手或每只手的轨迹、头的旋转或头的倾斜。
9. 根据权利要求1所述的装置,其中所述一或多个处理器经配置以控制所述扬声器阵列以产生不同方向上的束以独立地支持位于所述不同方向上的不同用户的手势控制。
10. 根据权利要求1所述的装置,其中话音命令用于进入手势控制模式。
11. 根据权利要求1所述的装置,其中所述用户的经分析的所辨识的移动包括面部辨识,话音辨识,或用于用户识别或用户位置的面部辨识和语音辨识两者。
12. 根据权利要求1所述的装置,其中所述用户的所辨识的移动表示手势,其中所述手势映射到命令,并且命令解释器经集成到所述一或多个处理器中并经配置以停用对当前声场配置的改变或启用对当前声场配置的改变。
13. 根据权利要求1所述的装置,其中所述俘获包括基于激光距离测量通过红外光的光学成像。
14. 根据权利要求1所述的装置,进一步包括除了对所述图像序列内的移动进行二维辨识之外还添加深度信息。
15. 根据权利要求14所述的装置,进一步包括两个或多个相机,其经配置以执行三维成像和分析。
16. 根据权利要求14所述的装置,还包括将条带图案,点图案或条带图案和点图案两者投影到所述用户的一部分上,且估计所述用户的一部分的表面点的深度。
17. 根据权利要求16所述的装置,进一步包括用于发射光的激光和用于将所述图案强加于所发射的光上的衍射光栅,以及用于俘获所述用户的受照明部分的图像的图像检测器。
18. 根据权利要求1所述的装置,其中所述用户的所辨识的移动基于用于执行空间成像的超声变换器阵列。
19. 根据权利要求1所述的装置,进一步包括屏幕上显示以提供对手势命令的反馈,其中所述反馈是显示束强度、束方向或动态范围的改变的条或拨盘。

20. 根据权利要求1所述的装置,进一步包括屏幕上显示以提供对手势命令的反馈,其中所述反馈是条或拨盘。

21. 根据权利要求20所述的装置,其中所述屏幕上显示经配置以在所述屏幕上的条上或在所述屏幕上的拨盘上显示用于所述手势命令的所述反馈,以表示所述束强度,束方向或动态范围的改变。

22. 根据权利要求20所述的装置,其中所述屏幕上显示经配置以显示无效手势命令的错误指示。

23. 根据权利要求1所述的装置,其中所述用户的所辨识的移动表示手势,其中所述手势是以下手势中的至少一者:双手手势,手与头手势,手与身体手势,以及手到耳朵的手势。

24. 根据权利要求1所述的装置,其中所述用户的所辨识的移动表示手势,其中所述手势是以下手势中的至少一者:顺时针手移动、逆时针手移动、手旋转、手抓住、以及手释放。

25. 根据权利要求1所述的装置,其中所述用户的所辨识的移动表示手势,其中所述手势是以下手势中的至少一者:手横向移动,手垂直移动,朝向用于俘获所述图像序列的传感器的手移动、远离用于俘获所述图像序列的所述传感器的手移动、手在相同方向上移动、手朝向彼此移动、以及手远离彼此移动。

26. 根据权利要求1所述的装置,其中所述用户的所辨识的移动表示手势,其中所述手势用于合成由所述扬声器阵列产生的经修改的声场以改变所述经修改的声场的音量,或控制所述经修改的声场中的束的音量。

27. 根据权利要求1所述的装置,其中所述用户的所辨识的移动表示手势,其中所述手势用于合成由所述扬声器阵列产生的经修改的声场以改变所述经修改的声场的束方向,或改变所述经修改的声场的束宽度,或改变所述经修改的声场的时间上的回声深度,或改变所述经修改的声场的动态范围扩展或压缩。

28. 根据权利要求1所述的装置,其中所述用户的所辨识的移动表示手势,其中所述手势用于合成由所述扬声器阵列产生的经修改的声场以产生或删除相对于所述扬声器阵列的轴的所指示方向上的声空值。

29. 根据权利要求1所述的装置,其中所述用户的所辨识的移动表示手势表示两个或多个手势的序列,其中所述两个或多个手势的序列用于合成由所述扬声器阵列产生的所述经修改的声场以用于菜单导航。

30. 根据权利要求1所述的装置,其中所述用户的所辨识的移动表示手势表示两个或多个手势的序列,其中所述两个或多个手势的序列用于合成由所述扬声器阵列产生的所述经修改的声场以用于经由声音的用户接口反馈。

31. 根据权利要求1所述的装置,其中所述用户的所辨识的移动表示手势表示两个或多个手势的序列,其中所述两个或多个手势的序列用于合成由所述扬声器阵列产生的所述经修改的声场以用于经由显示图标的用户接口反馈。

用于声场的手势操纵的系统、方法、和设备

[0001] 本申请为发明名称为“用于声场的手势操纵的系统、方法、和设备”的原中国发明专利申请的分案申请。原申请的中国申请号为201380017323.5；原申请的申请日为2013年3月20日。

[0002] 根据35U.S.C.§119主张优先权

[0003] 本专利申请案主张2012年4月2日申请且转让给本受让人的标题为“用于声场的手势操纵的系统、方法、设备和计算机可读媒体 (SYSTEMS, METHODS, APPARATUS, AND COMPUTER-READABLE MEDIA FOR GESTURAL MANIPULATION OF A SOUND FIELD)”的第61/619,202号临时申请案的优先权。

技术领域

[0004] 本发明涉及手势驱动的声场修改。

背景技术

[0005] 具有合理数量的元件的扬声器阵列可用以执行各种声场修改。举例来说，扬声器波束成形器可用以将声束导引到不同方向和/或产生声学空束。在另一实例中，扬声器阵列可使用波场合成来驱动以再合成真实声场的现实波前。波场合成在例如第2006/0098830A1号(罗伊德等人)、第2007/0269062A1号(罗迪佳斯特等人)和第2008/0101616A1号(梅切尔等人)美国公开专利申请案的文献中描述。扬声器阵列的其它创新使用包含再现例如环绕声、扩展声像等等效果。

发明内容

[0006] 根据一般配置的信号处理方法包含：驱动方向可控变换器以产生包含束的声场；响应于手势的表示而产生命令；响应于所述命令而产生滤波器配置；以及将空间方向性滤波器应用于输入信号以产生多通道信号。在此方法中，所述空间方向性滤波器是基于所述滤波器配置，所述手势指示所述束的方向和所述束的宽度当中的至少一者的改变，且所述命令是执行所述所指示改变的命令。还揭示具有有形特征的计算机可读存储媒体(例如，非暂时性媒体)，所述有形特征致使读取所述特征的机器执行此方法。

[0007] 根据另一一般配置的用于信号处理的设备包含：用于驱动方向可控变换器以产生束的装置；用于响应于手势的表示而产生命令的装置；用于响应于所述命令而产生滤波器配置的装置；以及用于对输入信号执行空间方向性滤波操作以产生多通道信号的装置。在此设备中，所述空间方向性滤波器是基于所述滤波器配置，所述手势指示所述束的方向和所述束的宽度当中的至少一者的改变，且所述命令是执行所述所指示改变的命令。

[0008] 根据另一一般配置的用于信号处理的设备包含：音频输出级，其经配置以驱动方向可控变换器以产生束；手势解译器，其经配置以响应于手势的表示而产生命令；命令解译器，其经配置以响应于所述命令而产生滤波器配置；以及合成器，其经配置以将空间方向性滤波器应用于输入信号以产生多通道信号。在此设备中，所述空间方向性滤波器是基于所

述滤波器配置,所述手势指示所述束的方向和所述束的宽度当中的至少一者的改变,且所述命令是执行所述所指示改变的命令。

附图说明

- [0009] 图1A展示圆锥型扬声器的实例。
- [0010] 图1B展示矩形扬声器的实例。
- [0011] 图1C展示如图1A所示的十二个扬声器的阵列的实例。
- [0012] 图1D展示如图1B所示的十二个扬声器的阵列的实例。
- [0013] 图2A和2B展示用户的阻挡手势的实例。
- [0014] 图3A和3B展示用户的另一阻挡手势的实例。
- [0015] 图4A和4B展示用户的另一阻挡手势的实例。
- [0016] 图5A和5B展示用户的重定向手势的实例。
- [0017] 图6A和6B展示用户的耳朵围拢手势的实例。
- [0018] 图7A展示根据一般配置的方法M100的流程图。
- [0019] 图7B展示方法M100的实施方案M110的流程图。
- [0020] 图7C展示方法M110的实施方案M120的流程图。
- [0021] 图8到10展示手势表示到命令的映射的实例。
- [0022] 图11说明系统对阻挡手势的响应。
- [0023] 图12说明系统对重定向手势的响应。
- [0024] 图13说明系统对耳朵围拢手势的响应。
- [0025] 图14展示用户的全局抑制手势的实例。
- [0026] 图15展示手势特征候选的三个实例。
- [0027] 图16展示由任务T400配置的空间方向性滤波器的频域实施方案SDF10的框图。
- [0028] 图17展示空间方向性滤波器的等效时域实施方案SDF20的框图。
- [0029] 图18A、18B、19A和19B展示相对于针对四个不同定向角的四元件阵列的DSB滤波器的方向的量值响应。
- [0030] 图20A和20B分别展示图19A和19B的DSB滤波器的经加权修改的束模式的实例。
- [0031] 图21A和21B分别展示八元件阵列的DSB滤波器的束模式的实例,其中滤波器的定向角为30和60度。
- [0032] 图22A和22B分别展示图21A和21B的DSB滤波器的经加权修改的束模式的实例。
- [0033] 图23A和23B分别展示具有三个和五个可选固定空间扇区的方案的实例。
- [0034] 图23C展示可由任务T300的实施方案参考的表的实例。
- [0035] 图23D展示合成器SY10的实施方案SY20的框图。
- [0036] 图24展示一组九个滤波器配置的束模式。
- [0037] 图25展示一组六个滤波器配置的束模式。
- [0038] 图26A、26B、26C和27展示成对BFNF操作的方面。
- [0039] 图28A展示根据一般配置的设备A100的框图。
- [0040] 图28B展示设备A100的实施方案A110的框图。
- [0041] 图28C展示设备A110的实施方案A120的框图。

- [0042] 图28D展示设备A100的实施方案A105的框图。
- [0043] 图29A展示根据一般配置的设备MF100的框图。
- [0044] 图29B展示设备MF100的实施方案MF110的框图。
- [0045] 图29C展示设备MF110的实施方案MF120的框图。
- [0046] 图30A展示均匀线性扬声器阵列。
- [0047] 图30B展示在扬声器之间具有对称八度间距的均匀线性阵列的一个实例。
- [0048] 图30C展示具有不对称八度间距的均匀线性阵列的实例。
- [0049] 图30D展示具有均匀间距的弯曲阵列的实例。
- [0050] 图31A展示膝上型计算机D710。
- [0051] 图31B展示显示装置TV10。
- [0052] 图31C展示显示装置TV20。
- [0053] 图32A和32B展示扩展阵列的两个实例的俯视图。
- [0054] 图32C和33展示两个不同阵列的正视图。
- [0055] 图34展示在PBE处理之前和之后的音乐信号的频谱的实例。

具体实施方式

[0056] 扬声器阵列可用以产生收听区。可能使用扬声器的线性、圆形或平面阵列来合成现实声场的波前(例如,使用波场合成)。关于线性阵列,还可能导引经再生的声场,使得其向一个方向成束,其中在其它方向上的能量低得多。本发明描述集成了手势辨识与声场修改的系统。此集成系统可用以用代表性手势驱动此种声场产生和/或修改,使得原生用户可用代表性手势起始例如阻挡声辐射或引导声束的动作。

[0057] 除非受其上下文清楚地限制,否则术语“信号”在本文中用以指示其普通意义中的任一者,包含如在导线、总线或其它传输媒体上表达的存储器位置(或存储器位置集合)的状态。除非由其上下文明确限制,否则本文中使用的术语“产生”来指示其普通意义中的任一者,例如计算或以另外方式产生。除非受其上下文清楚地限制,否则术语“计算”在本文中用以指示其普通意义中的任一者,例如运算、评估、估计及/或从多个值中进行选择。除非由其上下文明确限制,否则使用术语“获得”来指示其普通意义中的任一者,例如计算、导出、接收(例如,从外部装置)及/或检索(例如,从存储元件阵列)。除非由其上下文明确限制,否则使用术语“选择”来指示其普通意义中的任一者,例如识别、指示、应用和/或使用两者或两者以上的集合中的至少一者且少于全部。在术语“包括”用于本描述及权利要求书中的情况下,其不排除其它元件或操作。术语“基于”(如“A基于B”中)用以指示其普通意义中的任一者,包含以下情况:(i)“得自”(例如,“B是A的前体”),(ii)“至少基于”(例如,“A至少基于B”),以及在特定上下文中适当的情况下,(iii)“等于”(例如,“A等于B”)。类似地,使用术语“响应于”来指示其普通意义中的任一者,包含“至少响应于”。

[0058] 对多麦克风音频感测装置的麦克风的“位置”的参考指示所述麦克风的声敏感面的中心的位置,除非上下文另外指示。根据特定上下文,有时使用术语“通道”来指示信号路径且在其它时候指示由此路径载运的信号。除非另外指示,否则使用术语“系列”来指示两个或两个以上项目的序列。使用术语“对数”来指示基于十的对数,但将此运算扩展到其它基数在本发明的范围内。使用术语“频率分量”来指示信号的一组频率或频带中的一者,例

如信号的频域表示的样本(例如,由快速傅里叶变换产生)或信号的子带(例如,巴克(Bark)尺度或梅尔(mel)尺度子带)。

[0059] 除非另外指出,否则对具有特定特征的设备的操作的任何揭示内容还明确地希望揭示具有类似特征的方法(且反之亦然),且对根据特定配置的设备的操作的任何揭示内容还明确地希望揭示根据类似配置的方法(且反之亦然)。术语“配置”可参考由其特定上下文指示的方法、设备和/或系统来使用。术语“方法”、“过程”、“程序”和“技术”一般地且可互换地使用,除非特定上下文另外指示。术语“设备”和“装置”也一般地且可互换地使用,除非特定上下文另外指示。术语“元件”和“模块”通常用以指示较大配置的一部分。除非由其上下文明确限制,否则术语“系统”在此用以指示其普通意义中的任一者,包含“交互以用于共同目的的元件群组”。通过参考文献的一部分的任何并入也应理解为并入了在所述部分内参考的术语或变量的定义,其中这些定义出现在文献中的其它地方,以及并入了在所并入部分中参考的任何图式。

[0060] 可假定在发射声场的近场和远场区域中,波前分别为球形的和平面的。近场可界定为离开声接收器(例如,麦克风阵列)小于一个波长的空间区域。在此定义下,到所述区域的边界的距离与频率成反比变化。在例如两百、七百和两千赫兹的频率下,到一个波长边界的距离分别为大约170、49和17厘米。可有用的是改为考虑近场/远场边界处于距麦克风阵列的特定距离(例如,距阵列的麦克风或距阵列的形心50厘米,或距阵列的麦克风或距阵列的形心一米或1.5米)。除非由特定上下文另外指示,否则本文假定远场近似。

[0061] 图1A展示圆锥型扬声器的实例,且图1B展示矩形扬声器(例如,RA11x15x3.5,NXP半导体公司,荷兰埃因霍温)的实例。图1C展示如图1A所示的十二个扬声器的阵列的实例,且图1D展示如图1B所示的十二个扬声器的阵列的实例。可例如实施图1C和1D的实例以使得扬声器间距离为2.6cm且阵列的长度(31.2cm)近似等于典型膝上型计算机的宽度,但这些实例不限于这些特定尺寸,且本文描述的原理可用任何尺寸的方向可控变换器(例如,扬声器阵列)来实践。

[0062] 用于再现包含此扬声器阵列的系统的配置参数通常在字面上(例如,以科学术语)而不是描述性地表达。这些配置参数可包含波束成形器将导引声音的角度、束的宽度以及音量(参见例如第2010/0098275A1号(梅特卡夫)美国公开专利申请案的图11)。虽然这些参数从研究者的观点来看已经高度抽象,但它们对于测试工程师比对于尝试在例如电话或者平板或膝上型计算机等装置上使用所述阵列的用户来说可能更合适。

[0063] 当此扬声器阵列集成到或耦合于个人计算装置(例如,膝上型计算机、上网本或平板计算机;智能电话;机顶盒或视频重放装置;游戏控制台等等)或电视显示器时,可能希望允许用户实时调整阵列的操作。然而对于消费者应用,可能希望避免要求用户在每次声场改变时确定和指定阵列的详细配置参数,且更为用户友好的解决方案(例如,简单且较为描述性的用户接口)和集成系统可为合意的。举例来说,实施响应于用户正在做的事而实时调整阵列的接口可为合意的。

[0064] 声场修改的常规方法可能要求用户找到遥控器上的按钮,中断观看以便操纵屏幕上的用户接口,和/或学习话音命令。本发明描述集成不同的感觉信息来控制声音产生。通过配置系统以适当地响应于代表性手势,实现新的用户体验,且可开发利用声场修改算法和扬声器阵列硬件的能力的更好方式。

[0065] 如果用户不想听到来自特定方向的声音,那么指示所要改变的动作可能是用户举起他的手好像在阻挡声音一样。此动作可经表征为将手举起为防御性或保护性手势的手势,其中打开的手掌面向外。图2A和2B的序列展示用户可如何随着时间执行此手势的一个实例。指示声音是不希望的另一动作可能是用户向外移动打开的手(例如,手离开用户的移动,好像在阻挡或推开外部声音)。图3A到3B和图4A到4B的序列展示用户可如何随着时间执行此手势的两个实例。以此方式呈现具有面向外的打开手掌的手势也可称为“不想说话”手势。

[0066] 如果用户希望将声束移动到不同方向,那么指示所要改变的动作可能是用户使她或她的手朝向彼此移动(例如,好像抓住声束)且随后将手一致地移动到所要方向。图5A和5B的序列展示用户可如何随着时间执行此手势的一个实例。

[0067] 如果用户不能清楚地听到特定方向上的声音且希望声场朝向他或她的耳朵聚焦,那么指示所要改变的动作可能是用户将手围挡在耳朵边上,且可能还在所述耳朵的方向上倾斜他或她的头。此动作可经表征为将手举到用户的耳朵(例如,将手举到靠近用户的脸)的手势,其中手掌面向前且可能手指部分地向前卷曲(例如,围挡耳朵的手)。图6A和6B展示此手势随着时间的进展。

[0068] 在顶部层级,用于声场修改的集成系统可根据“辨识→分析→合成”范例来实现。此结构可经设计以辨识由用户指示的所要改变,分析当前上下文中的所指示改变,且合成经修改的声场以实施所指示改变。装置可用以俘获代表性人手势,其随后经解译为可能模式中的一者。根据单独的模式,做出决定且合成对应的声场。通常,合成将基于来自手势的细节的另外信息,其可包含手的位置、准确定向的识别等等。在一个实例中,系统驱动扬声器阵列(例如,如图1C或1D所示)以产生在第一方向上包含能量集中的声场,且辨识在声场内执行的手势,其中手势指示扬声器阵列的平面中的第二方向。

[0069] 图7A展示根据一般配置的音频信号处理方法M100的流程图,其包含声场产生任务T100、手势解译任务T200、命令解译任务T300和信号合成任务T400。声场产生任务T100驱动方向可控变换器(例如,扬声器阵列)以产生包含束的声场。将束界定为相对于变换器沿着特定角度或空间方向的能量集中。举例来说,任务T100可实施为时域和/或频域中的波束成形操作,如本文论述。

[0070] 手势解译任务T200产生执行束的方向和束的宽度当中的至少一者的改变的命令。任务T200响应于指示所述改变的手势的表示而产生此命令。任务T200可经实施以根据所述表示到所述命令的映射来产生所述命令,所述映射可体现于查找表或类似结构中。图8到10展示这些映射的实例。

[0071] 响应于由任务T200产生的命令,命令解译任务T300产生滤波器配置。滤波器配置可包含一组滤波器系数,其描述时域和/或频域中的空间方向性滤波操作(例如,波束成形操作)。或者,滤波器配置可描述对此一组滤波器系数的修改。

[0072] 信号合成任务T400通过将基于滤波器配置的空间方向性滤波器应用于输入信号而产生多通道信号。举例来说,任务T400可实施为时域和/或频域中的波束成形操作。

[0073] 任务T100驱动方向可控变换器以产生包含束的声场。任务T100通常经实施以基于由例如波束成形器或波场合成(WFS)再现技术等空间方向性滤波器产生的多通道信号而驱动变换器。举例来说,驱动信号可基于在方法M100的先前执行中由信号合成任务T400的实

例产生的多通道信号。对于其中变换器为扬声器阵列的情况,典型的是驱动信号为多通道信号,使得每一通道驱动阵列的特定对应扬声器。

[0074] 可为合意的是实施任务T100以产生驱动信号以在所产生声场的所要方向上(例如,在束的方向上)引入相长干涉,同时在所产生声场的其它方向上(例如,在束的主瓣之外的方向上)引入相消干涉。举例来说,任务T100可经实施以驱动变换器以产生在所要方向上集中的声能束且在其它方向上产生束响应的谷。图18A到22B展示下文更详细论述的束模式的实例。

[0075] 大体上,任务T100可经实施以驱动在邻近扬声器之间具有均匀或不均匀间距的线性扬声器阵列,或非线性(例如,经定形)扬声器阵列。还可能实施任务T100以针对不同频率范围以不同方式驱动变换器的部分(例如,扬声器阵列的子阵列)。此实施方案可为宽带再生提供较好的方向性。在一个此类实例中,任务T100经实施以用具有第一频率范围的第一驱动信号驱动扬声器阵列,且用具有低于第一频率范围的第二频率范围的第二驱动信号驱动阵列的交替扬声器。

[0076] 手势解译任务T200可经实施以产生通过将手势表示作为索引应用于查找表或类似结构来执行所指示改变的命令,所述查找表或类似结构将手势表示映射到对应命令。所表示手势到对应命令的此些映射的一些实例在下文论述:

[0077] A) 阻挡或推动手势(例如,“不想说话”手势)的表示可映射到在用户的方向上减小束的强度或在声场中产生局部空值的命令(例如,如图11所示)。如果用户想要改为升高局部声级,那么指示所要改变的动作可能是用户做出招手手势。此手势可经表征为呈现手掌面朝用户的举起的手(其定向可由拇指相对于例如用户的脸的向外位置指示),并且将手的手指卷曲到手掌中。或者,招手手势可经表征为举起且打开手掌向内的手,以及手的大小减小和/或距离增加。招手或拉动手势的表示可映射到在声场中产生局部束或在用户的方向上减小现存束的强度的命令。图8展示在行GA1到CA1和GA2到CA2处的此些映射的实例。

[0078] B) 如果用户想要将声束移动到不同方向,那么指示所要改变的动作可能是用户使她的手朝向彼此移动(例如,好像抓住声束)且随后将手一致地移动到所要方向。图5A和5B的序列展示用户可如何随着时间执行此手势的一个实例。或者,此所要改变可由包含以下动作的序列的手势指示:将手带到一起(例如,在围拢运动中),随后将手一起平移到新位置,随后使手移动彼此分开(或者在新位置将手保持在一起达某个周期,例如一或两秒)。此手势的单手版本可经表征为打开手掌向外的手的出现(可能随后是手指向内朝向手掌卷曲的运动),其随后移动到一侧或另一侧以指示所要方向。此手势的另一单手版本包含如下动作的序列:闭合手,随后将闭合的手平移到新位置,随后打开手(或者,在新位置中保持手静止达某个周期,例如一或两秒)。此重定向手势的表示可映射到将束从原始方向导引(例如,平移)到所指示方向中的命令(例如,如图12所示)。图8展示在行GB1到CB1和GB2到CB2处的此些映射的实例。

[0079] C) 如果用户想要将束聚焦到他自己的方向(例如,以增加电话呼叫的隐私性),那么指示所要改变的动作可能是用户展示他的手在水平方向上较宽地间隔开且随后将手带到一起。此手势的表示可映射到在用户的方向上产生束且将输入信号限定于所述束或者在用户的方向上减小现存束的宽度的命令。类似地,在水平方向上使手移动彼此远离的手势可指示增加声束的宽度的所要改变且可映射到此命令。图8展示在行GC1到CC1和GC2到CC2

处的这些映射的实例。图10展示在行GC1a到CC1a、GC1b到CC1b、GC2a到CC2a和GC2b到CC2b处的这些映射的替代实例。

[0080] D) 如果用户不能清楚地听到特定方向上的声音且想要在她的方向上聚焦声场或另外使声场更明了(例如,在噪声环境中),指示所要改变的动作可能是用户将她的手围拢在她的耳朵边上,且可能还朝向所述方向倾斜她的头。此动作可经表征为将手举到用户的耳朵(例如,将手举到用户的脸旁边)的手势,其中手掌面向前且可能手指部分地向前卷曲(例如,围拢到耳朵的手)。图6A和6B展示此手势随着时间的进展。此手势的表示可映射到在耳朵的方向上产生或强化束(例如,增加局部音量),朝向耳朵重定向现存束,或另外提升所述方向上的声场的等级的命令,例如图13中所示。替代地或另外,此手势的表示可映射到局部地或全局地压缩声场的动态范围和/或局部地或全局地增强声场的所感知响度(例如,经由低音提升和/或心理声学低音增强(PBE))的命令。图8展示在行GD1到CD1处的此映射的实例。

[0081] 相反,如果用户想要翻转此聚焦,那么指示所要改变的动作可能是用户做出擦除或摆手手势。此手势可经表征为以打开手掌向外保持(通常手指分开而不是在一起)且在手的基部围绕枢轴来回移动(即,通过围绕上臂的轴来回旋转上臂而移动)的举起的手。此手势的表示可映射到使局部音量、动态范围和或响度补偿操作返回到默认水平的命令。图8展示在行GD2到CD2处的此映射的实例。

[0082] E) 如果用户想要增加整个声场的音量,那么指示所要改变的动作可能是用户将双手一起举起(手掌向上或手掌向下),且此手势的表示可映射到增加声场的总体音量的命令。类似地,如果用户想要减小整个声场的音量,那么指示所要改变的动作可能是用户将双手一起降低(例如,手掌向下,如图14的实例中所示),且此手势的表示可映射到减小声场的总体音量的命令。图8展示在行GE1到CE1和GE2到CE2处的这些映射的实例。

[0083] F) 如果用户想要减小音量或压缩声场的动态范围(至少在用户的方向上),那么指示所要改变的动作可能是用户在垂直方向上使手朝向彼此移动或者将一只手保持到耳朵同时降低另一只手。此手势的表示可映射到减小音量或压缩声场的动态范围(局部地或全局地)的命令。图8展示在行GF1到CF1处的此映射的实例。图9展示在行GF3到CF3处的此映射的另一实例。

[0084] 类似地,如果用户想要增加音量或扩展声场的动态范围(至少在用户的方向上),那么指示所要改变的动作可能是用户在垂直方向上使手移动远离彼此或者将一只手保持到耳朵同时升高另一只手。此手势的表示可映射到增加音量或扩展声场的动态范围(局部地或全局地)的命令。图8展示在行GF2到CF2处的此映射的实例。图9展示在行GF4到CF4处的此映射的另一实例。

[0085] G) 如果用户想要相对于其它频带的等级增加或减小声场的频带的等级(例如,用于平衡),那么指示所要改变的动作可能是用户将一只手保持到耳朵同时保持另一只手直立且手掌向外并且从用户的视点来看分别顺时针或逆时针旋转此手(例如,好像转动旋钮或轮)。此手势的表示可映射到增加或减小特定频带的等级的命令。等级待调整的频带(例如,低音、中音、高音)可由旋转的手的高度指示(例如,手保持高以指示高音且保持低以指示低音)。图9展示在行GG1a到CG1a、GG2a到CG2a、GG1b到CG1b、GG2b到CG2b、GG1c到CG1c和GG2c到CG2c处的这些映射的实例。

[0086] H) 如果用户想要停用对当前声场配置的改变,那么指示所要改变的动作可能是用户呈现手掌向外的打开手以及闭合手同时顺时针旋转手。此锁定手势的表示可映射到停用对当前声场配置的改变或停用特定改变(例如,束方向的改变)同时准许其它改变(例如,局部声级的改变)的命令。类似地,如果用户想要启用对当前声场配置的改变,那么指示所要改变的动作可能是用户呈现闭合的手以及打开手同时逆时针旋转手。此解锁手势的表示可映射到启用对当前声场配置的改变的命令。图9展示在行GH1到CH1和GH2到CH2处的此些映射的实例。

[0087] 与特定实施方案中的不同手势表示相关联的动作范围可包含上文和/或在图8到10中陈述的实例中的任何一或多个者。另外或替代地,这些动作中的一些可全局地应用(例如,音量控制、频率分布和/或动态范围)。

[0088] 如下文论述,所表示的手势可包含除了手移动之外或连同手移动一起的动作。举例来说,可为合意的是实施方法M100以跟踪用户的面部表情、头部位置和/或定向和/或凝视方向以及手移动且相应地更改声场(例如,导引声束)。

[0089] 手势解译任务T200可为上下文相依的。举例来说,任务T200可经实施以仅响应于适合于当前上下文的手势的表示而产生命令。在一个此类实例中,任务T200经实施以在系统已经处于静音状态时忽略减小音量的手势的表示。在另一此类实例中,任务T200经实施以在系统已经处于所述方向上的被阻挡状态时忽略阻挡手势的表示。

[0090] 从手势表示到对应命令的映射也可为上下文相依的。举例来说,任务T200可经实施以取决于当接收到手势表示时存在的上下文而使两个或两个以上不同命令与同一手势表示相关联。另外或替代地,上下文可指示应局部地(例如,仅在用户的方向上)还是全局地应用命令。

[0091] 方法M100可经实施以从另一操作或装置获得手势表示。替代地,方法M100可经实施以通过执行所俘获场景的分析来产生手势的表示。图7B展示包含场景分析任务TA10的方法M100的实施方案M110的流程图,所述场景分析任务分析所俘获场景以产生手势的表示。

[0092] 图7C展示还包含场景俘获任务TC10的方法M110的实施方案M120的流程图,所述场景俘获任务俘获包含手势的场景。场景俘获任务TC10可经实施以使用各种俘获模态(例如,成像和/或测距)中的任何一或多个者来记录场景。举例来说,任务TC10可经实施以记录包含手势的图像或图像序列(例如,视频序列)。

[0093] 此俘获可包含通过可见光和/或红外(例如,用于基于颜色和/或亮度差的图像分段)和/或激光距离测量的光学成像。任务TC10可经实施以使用基于相机的成像系统来俘获图像序列,且任务TA10可经实施以使用图像处理技术来辨识所述序列内的对象和移动。举例来说,任务TC10可经实施以俘获声场内的用户的图像序列(例如,使用安装在个人计算装置中或上的网络摄像机)。通过两个或两个以上相机,任务TC10和TA10可甚至经实施以分别执行三维成像和分析(例如,在二维辨识操作上添加深度信息)。

[0094] 另外或替代地,任务TC10可经实施以使用超声执行此俘获。在一个此类实例中,使用超声变换器阵列来执行空间成像(例如,通过测距)。通过超声阵列中的变换器的足够密度,此空间成像的分辨率可高到足以支持微小细节的识别。另外或替代地,任务TC10可经实施以通过使用超声多普勒技术(例如,如第2010/0202656A1号(拉马克里思南等人)美国公开专利申请案中描述)俘获移动信息。

[0095] 另外或替代地,任务TC10可经实施以使用结构化发射来执行测距。举例来说,任务TC10可经实施以使用例如扩展频谱超声(例如第2012/0001875A1号(李等人,2012年1月5日公开)美国公开专利申请案中描述)的声音的结构化发射来执行测距。此测距可用以支持所俘获场景的深度成像和/或深度分段。

[0096] 另外或替代于使用声音的结构化发射,任务TC10可经实施以使用光的结构化发射用于距离映射、深度成像和/或深度分段。此实施方案可包含将已知图案(例如,周期性光图案,例如条带和/或点的图案)投影到场景上且通过俘获受照明场景的图像且确定所述图案的变形来估计场景中的表面点的深度。此方法可使用激光(例如,红外激光)来发射光,使用衍射光栅(例如,全息图)来将图案强加于发射的光上,且使用图像检测器来俘获受照明场景的图像。在此情况下,任务TA10可经实施以分析由视场中的对象反射的投影图案的图像(例如,由华盛顿州雷德蒙德微软公司的Kinect传感器执行)。

[0097] 与这些俘获模态中的任何一或多者一起(或替代地)可由俘获任务TC10使用的其它俘获方法包含使用手势手套和/或其它控制器,其可包含跟踪由控制器中的定向传感器指示的移动,其指示控制器参照外部参考方向的当前空间定向,所述外部参考方向例如为重力轴(例如,垂直于地球表面的轴)或磁轴(例如,地球的磁轴)。此定向传感器可包含一或多个惯性传感器,例如陀螺仪和/或加速度计。陀螺仪使用角动量的原理来检测围绕一条轴或围绕两条或三条(通常正交)轴中的每一者的定向改变(例如,俯仰、横摇和/或扭转的改变)。可制造为微机电系统(MEMS)装置的陀螺仪的实例包含振动陀螺仪。加速度计检测沿着一条轴或沿着两条或三条(通常正交)轴中的每一者的加速度,且可用以俘获移动方向。加速度计也可制造为MEMS装置。还可能将陀螺仪和加速度计组合到单个传感器中。另外或替代地,定向传感器可包含一或多个磁场传感器(例如,磁力计),其测量沿着一条轴或沿着两条或三条(通常正交)轴中的每一者的磁场强度。在一个实例中,控制器包含磁场传感器,其指示装置相对于磁轴(例如,地球的磁轴)的当前定向。

[0098] 场景分析任务TA10分析所俘获场景以产生手势的表示。所俘获场景的此分析可包含一或多个图像分析操作,例如图像分段(例如,按颜色和/或按深度)、肤色检测和/或背景消减。场景分析任务TA10还可包含检测随着时间的改变的一或多个操作,例如计算视频序列的邻近帧之间的差图像。

[0099] 任务TA10可经实施以对所分析场景执行一或多个子任务,例如特征提取和特征分类。特征提取可包含分析所俘获场景以检测和定位所关注区域,例如用户的手、手指、头、脸、眼睛、身体和/或肩。此分析也可包含识别所检测特征的定向和/或随着时间的改变(例如,在图像序列内),例如手或每只手的轨迹、头的旋转和/或倾斜等等。

[0100] 任务TA10的场景分析还可包含所检测区域的表征,例如计算一或多个动量(例如,Hu不变量动量)、计算k曲率(例如,用于指尖检测)、计算几何形状描述符和/或使区域匹配于一组模板当中的一或多者。此些分析操作可包含手检测、用于手指提取的凸包、隅角检测、凸性检测、用于身体部分标记的骨骼登记、映射到骨骼模型和/或深度分段。

[0101] 任务TA10可包含特征分类,例如在匹配的量度(例如,相似性量度)高于可为候选相依的阈值的条件下将特征分类为在一组手势特征候选当中最近的(例如,根据最大相似性量度)。特征的一或多个方面可包含例如以下各项中的一或多者:形状、位置(例如,用户的手彼此之间的空间关系,和/或用户的手与用户的脸和/或眼睛的空间关系)、距离(例如,

通过测距和/或通过所检测特征的大小而检测)、定向(例如,手或头的倾斜,指向的方向)和平移(例如,向左、右、上和/或下的移动)。图15展示手势特征候选的三个实例。

[0102] 特征提取也可基于过去和/或未来俘获。举例来说,任务TA10可经实施以使得俘获帧中的特定特征的检测可在当在邻近俘获帧中的附近位置处检测到同一特征的情况下视为较可能的。

[0103] 任务TA10可经实施以通过选择多个候选手势表示当中的一或多者来执行手势辨识。此手势辨识可包含在匹配的度量(例如,相似性度量)高于可为候选相依的阈值的情况下将手势分类为在一组手势候选当中最近的(例如,根据最大相似性度量)。此分类可基于隐式马尔可夫模型或其它模式辨识算法以从场景或帧内的个别特征辨识手势元素和/或辨识随着时间的手势元素的序列。额外应用可包含用于响应于菜单导航和/或用户接口反馈(例如,经由声音和/或显示图标)的复合手势(例如,两个或两个以上手势的序列)。

[0104] 任务TA10可经实施以将手势辨识为手势元素(例如,检测到的特征)序列。此序列还可包含手势元素的特定移动。如上所述,手手势可用相机来辨识和/或用超声阵列来跟踪。举例来说,图2A到2B中所示的手势可经表征为出现举起的打开手掌向外的手(可由拇指相对于例如用户的脸的向内位置指示)和/或大小增加和/或距离减小的举起的打开手掌向外的手。手移动的其它实例包含手的顺时针或逆时针旋转、横向移动、垂直移动、朝向或远离传感器的移动和相对移动(例如,手在相同方向上移动、手朝向/远离彼此移动)。手形状的改变可包含抓住和释放。手势可包含用户的头的旋转以及用户的头的横向倾斜当中的一者(例如,用于动态范围和/或音量的控制)。

[0105] 任务TA10可经实施以从分量手势元素辨识复合手势(例如,包含随着时间呈现的不同元素的序列的手势)。候选手势表示的集合还可包含双手、手与头和/或手与身体手势,例如包含手到耳朵的手势(例如,一只手到耳朵,且另一只手向上/下或内/外移动以指示音量或动态范围的改变)。

[0106] 任务TA10可经实施以将特定手势检测为手势元素的初始状态结合所述元素的一或多个方面中随着时间的改变的组合。隐式马尔可夫模型可用以将观测值序列匹配于手势状态序列(例如,序列的初始、中间和终止图像)。任务TA10可包含跟踪以检验手势序列在指定周期(例如,两到三秒)内完成。

[0107] 对于待表示的手势中的至少一些,场景分析操作(例如,任务TA10)可包含从所俘获场景提取相关联量。举例来说,此操作可经实施以量化手势元素在向上/下、左-右和/或接近/后退方向(例如,相对于传感器)上移动的程度,或两个移动手势元素(例如,用户的手)之间的初始分离和/或最终分离的程度。此量化可例如为合意的以确定对应一或多个动作的一或多个参数中的每一者的值。此些可量化参数可包含待强调的频率范围的音量、位置和/或带宽、束方向的改变、束宽度、动态范围扩展或压缩、时间上的回声深度等等。

[0108] 此参数值可应用(例如,由任务T400)于调整例如音量或动态范围的水平。在第一实例中,任务TA10根据相对于(例如,手的)初始位置的距离来确定所指示的量。此手手势可包含手的暂停以建立基本水平,且任务TA10可经实施以测量手在相对于手势的方向上(例如,上/下、左/右、内/外)的最终位置的距离。在第二实例中,任务TA10根据相对于参考位置的距离来确定所指示的量。举例来说,用户的手的参考位置可相对于用户的头来建立。在一个此类实例中,在下巴水平的手与范围的中间值相关联,其中上限和下限分别为所述水平

上方和下方大约25到30厘米。与参考的此关联也可用于其它配置参数。举例来说,可相对于保持肩宽度分隔(与当前束宽度相关联)的手的参考位置来指示所要的束宽度。

[0109] 替代地,参数(例如,音量、束宽度、动态范围)的范围可经确定为用户位置处的估计对应距离。举例来说,任务TA10可经实施以组合手之间的所观测距离或角度与手距深度传感器的所观测距离以确定在用户位置处手之间的对应距离。到用户的距离也可供例如头或手大小、肩跨度等等视觉线索来估计。

[0110] 响应于由手势解译任务T200产生的命令,命令解译任务T300产生滤波器配置。如上文论述,例如由任务T200产生的命令范围可包含以下各项中的任何一或多个者:

[0111] 产生或删除所指示方向(例如,相对于方向可控变换器的轴)上的声束;

[0112] 改变声束的宽度、音量、动态范围和/或方向(例如,相对于方向可控变换器的轴);

[0113] 改变声束相对于频率的能量分布(例如,平衡或低音提升);

[0114] 产生或删除所指示方向(例如,相对于方向可控变换器的轴)上的声空值;

[0115] 改变声空值的宽度、深度和/或方向(例如,相对于方向可控变换器的轴);以及

[0116] 锁定或解锁当前声场配置。

[0117] 另外或替代地,任务T200可产生将应用于整个声场的命令,例如音量空值、频率分布和/或动态范围。

[0118] 信号合成任务T400通过将基于滤波器配置的空间方向性滤波器应用于输入信号来产生多通道信号。图16展示由任务T400配置以产生多通道信号MCS10的每一通道MCS10-1到MCS10-N作为输入信号SS10与滤波器配置的系数 w_1 到 w_N 中的对应一者的乘积的空间方向性滤波器的频域实施方案SDF10的框图。这些乘法可串行地(即,一个接一个)和/或并行地(即,每次两个或两个以上)执行。图17展示由任务T400配置以产生多通道信号MCS10的每一通道MCS10-1到MCS10-N作为输入信号SS10与滤波器配置的滤波器系数的集合 w_1 到 w_N 中的对应一者的卷积的空间方向性滤波器的等效时域实施方案SDF20的框图。

[0119] 通过根据滤波器配置加权和/或延迟输入信号以产生多通道信号的每一通道,任务T400的此实施方案可用以获得所产生声场的所要空间分布。举例来说,空间方向性滤波器可经配置以产生多通道信号以在所得声场的所要方向上(例如,束的方向上)引入相长干涉同时在声场的其它方向上(例如,在束的主瓣之外的方向上)引入相消干涉。如下文所述,由任务T400的此实施方案产生的多通道信号可用以驱动变换器以产生在所要方向上集中的声能束且在其它方向上产生束响应中的谷。

[0120] 任务T300可经实施以根据相控阵列技术产生滤波器配置以使得多通道信号的每一通道具有相应相位(即,时间)延迟。此技术的一个实例是延迟总和波束成形(DSB)滤波器。在此情况下,空间方向性滤波器可经配置以通过将相应时间延迟应用于输入信号以产生信号MCS10的每一通道来在所要方向上引导声场。对于其中多通道信号用以驱动例如均匀间隔线性扬声器阵列的情况,可根据用于频域中的DSB滤波操作的以下表达式来计算空间方向性滤波器的通道系数 w_1 到 w_N :

$$[0121] \quad w_n = \frac{1}{N} \exp \left(-j \frac{2\pi f}{c} (n-1) d \cos \varphi_s \right), \quad (1)$$

[0122] $1 \leq n \leq N$,其中 d 为阵列中的邻近扬声器的辐射表面的中心之间的间距, N 为待驱动的扬声器的数目(可小于或等于阵列中的扬声器的数目), f 为频率区间索引, c 为声速,且 φ_s

为束相对于阵列的轴的所要角度(例如,声场的主瓣的所要方向)。对于滤波器配置的等效时域实施方案,元素 w_1 到 w_N 可实施为对应延迟。在任一域中,任务T400也可包含通过将信号MCS10的每一通道缩放 $1/N$ 的因数(或等效地,将输入信号SS10缩放 $1/N$)的信号MCS10的正规化。

[0123] 对于间距 d 等于波长 λ 的一半的频率 f_1 (其中 $\lambda=c/f_1$),表达式(1)简化为以下表达式:

$$[0124] \quad w_n = \frac{1}{N} \exp(-j\pi(n-1)\cos\varphi_s)。 \quad (2)$$

[0125] 图18A、18B、19A和19B展示针对四元件阵列在频率 f_1 处相对于此DSB滤波器的方向的量值响应(也称为束模式)的实例,其中滤波器的定向角(即,由每一图中的三角形指示的 φ_s)分别为30度、45度、60度和75度。

[0126] 应注意,图18A、18B、19A和19B中所示的滤波器束模式可在 $c/2d$ 之外的频率处不同。为了避免空间混叠,可能希望将输入信号的最大频率限于 $c/2d$ (即,使得间距 d 不大于信号的最短波长的一半)。为了引导包含高频的源分量,可为合意的是使用较靠近间隔的阵列。

[0127] 还可能实施任务T400以应用空间方向性滤波器以产生多通道信号,使得针对不同频率范围可以不同方式驱动空间方向性变换器的部分(例如,扬声器阵列的子阵列)。此实施方案可为宽带再现提供较好的方向性。在一个此类实例中,任务T400经实施以从限于 $c/4d$ 的最大频率的输入信号的频带产生第二 $N/2$ 通道多通道信号(例如,使用空间方向性滤波器的通道 w_1 到 w_N 中的交替通道),且此第二多通道信号用以驱动扬声器阵列的交替扬声器(即,具有 $2d$ 的有效间距的子阵列)。

[0128] 可为合意的是实施滤波器配置以将不同的相应权重应用于多通道信号的通道。举例来说,可为合意的是滤波器配置包含应用于滤波器系数的空间开窗函数。空间开窗函数的使用往往减小侧瓣量值和角分辨率(例如,通过加宽主瓣)。此开窗函数的实例包含(不限于)三角形和升余弦(例如,韩恩(Hann)或汉明(Hamming))窗。

[0129] 在一个实例中,任务T300经实施以产生滤波器配置以使得源空间方向性滤波器的每一通道 w_n 的系数包含空间开窗函数的相应因数 s_n 。在此情况下,表达式(1)和(2)可分别经修改为以下表达式:

$$[0130] \quad w_n(f) = s_n \exp\left[-j\frac{2\pi f}{c}(n-1)d\cos\varphi_s\right]; \quad (3a)$$

$$[0131] \quad w_n = \frac{1}{N} s_n \exp(-j\pi(n-1)\cos\varphi_s)。 \quad (3b)$$

[0132] 图20A和20B展示根据此修改的分别针对图19A和19B的四元件DSB滤波器在频率 f_1 处的束模式的实例,其中权重 S_1 到 S_4 分别具有值(2/3, 4/3, 4/3, 2/3)。

[0133] 具有较多扬声器的阵列允许较多自由度且可通常用以获得较窄的主瓣。图21A和21B展示针对八元件阵列的DSB滤波器的束模式的实例,其中滤波器的定向角度分别为30和60度。图22A和22B展示分别针对图21A和21B的八元件DSB滤波器的束模式的实例,其中由以下汉明开窗函数界定的权重 S_1 到 S_8 应用于空间方向性滤波器的对应通道的系数:

$$[0134] \quad s_n = 0.54 - 0.46 \cos \left[\frac{2\pi(n-1)}{N-1} \right]. \quad (4)$$

[0135] 任务T300可经实施以通过从集合(例如,预先计算的束配置的集合)选择适当配置、通过直接计算(例如,根据以上表达式(1)到(4)中的任一者)和/或通过修改现存滤波器配置来产生滤波器配置。用于特定命令的适当滤波器配置可取决于系统的当前状态或上下文而变化。举例来说,任务T300将产生的适当滤波器配置可取决于当前正由合成任务T400应用于空间方向性滤波器的滤波器配置。因此,可为合意的是实施任务T300以在当前声场配置的上下文中解译命令以确定将产生的适当滤波器配置。

[0136] 当前上下文可包含例如束的当前方向、宽度、音量、动态范围和/或频率分布等方面。当前上下文的其它方面可包含声场的音量、动态范围和/或频率分布的当前总体水平;用户的位置(例如,相对于变换器的轴的角度);以及可能一或多个其它用户中的每一者的位置。

[0137] 可能合意的是将目标用户环境的角跨度建模为若干离散空间扇区。图23A和23B分别展示其中扬声器阵列R10前方的空间经划分为三个空间扇区和五个空间扇区的实例。此模型的其它实例中的空间扇区的总数目包含四个、六个、七个、八个、九个和更多。在此些情况下,任务T300可经实施以产生与所指示扇区相关联的滤波器配置。

[0138] 图24展示一组九个滤波器配置的束模式。此图中每一模式的尺度和定向与图18A到22B中的模式相同,且每一模式对应于在所指示角度处定向且驱动均匀间隔的线性扬声器阵列的八通道DSB滤波器。对于其中目标跨度经划分为分别以30、45、60、75、90、105、120、135和150度为中心的九个扇区的实例,任务T300可经配置以响应于“将束向左移动”或“将束向右移动”的命令而参考当前在空间方向性滤波器中应用的滤波器配置来产生这些滤波器配置中的适当一者。

[0139] 对于其中由任务T200产生的命令范围包含改变声束的宽度的命令的应用,任务T300可经实施以参考如图23C的实例中所示的表。在此表中,符号“↓”、“↓”、“↑”和“↑”表示“将束宽度减小到最小值”、“减小束宽度”、“增加束宽度”和“将束宽度增加到最大值”。在当前正在空间方向性滤波器中应用的滤波器配置的束宽度较窄时,任务T300确定减小束宽度的命令是无效的。同样,在当前正在空间方向性滤波器中应用的滤波器配置的束宽度较宽时,任务T300确定增加束宽度的命令是无效的。无效命令可被忽略或可触发错误指示(例如,在显示屏上)。在其它情况下,任务T300产生滤波器配置以参考与当前正在空间方向性滤波器中应用的滤波器相关联的束宽度来提供所要束宽度。

[0140] 图25展示一组六个滤波器配置的束模式。此图中每一模式的尺度和定向与图18A到22B中的模式相同。模式A到F中的每一者对应于在六个度数处定向且驱动均匀间隔的线性扬声器阵列的DSB滤波器,且所述模式经布置以便增加束宽度。对于其中当前束方向为60度的实例,任务T300可经配置以响应于增加或减小束宽度的命令而参考当前在空间方向性滤波器中应用的滤波器配置来产生这些滤波器配置中的适当一者。在此特定实例中,模式A到F分别对应于八元件DSB滤波器、具有三角形窗的八元件DSB滤波器、具有汉明窗的八元件DSB滤波器、四元件DSB滤波器(例如,驱动八元件阵列的交替扬声器)、具有三角形窗的四元件DSB滤波器,以及具有汉明窗的四元件DSB滤波器。

[0141] 可为合意的是实施任务T300以产生根据超方向性波束成形器算法计算的滤波器

配置,其最大化所要方向上的增益同时最小化所有其它方向上的平均增益。超方向性波束成形器的实例包含最小方差无失真响应(MVDR)波束成形器(交叉协方差矩阵)以及线性约束最小方差(LCMV)波束成形器。其它固定或自适应波束成形技术,例如一般化侧瓣消除器(GSC)技术,也可用以计算由任务T300产生的滤波器配置中的一或多个者。

[0142] MVDR波束成形器的设计目标是以经受 $\mathbf{W}^H \mathbf{d} = 1$ 的约束 $\min_{\mathbf{W}} \mathbf{W}^H \Phi_{XX} \mathbf{W}$ 最小化输出功率,其中 \mathbf{W} 表示滤波器系数矩阵, Φ_{XX} 表示扬声器信号的正规化交叉功率谱密度矩阵,且 \mathbf{d} 表示导引向量。此束设计可表达为:

$$[0143] \quad \mathbf{W} = \frac{(\Gamma_{VV} + \mu \mathbf{I})^{-1} \mathbf{d}}{\mathbf{d}^H (\Gamma_{VV} + \mu \mathbf{I})^{-1} \mathbf{d}},$$

[0144] 其中 \mathbf{d}^T 是线性阵列的远场模型,其可表达为

$$[0145] \quad \mathbf{d}^T = [1, \exp(-j \Omega f_s c^{-1} \cos(\theta_0)), \exp(-j \Omega f_s c^{-1} 2l \cos(\theta_0)),$$

$$[0146] \quad \dots, \exp(-j \Omega f_s c^{-1} (N-1) \cos(\theta_0))],$$

[0147] 且 $\Gamma_{V_n V_m}$ 为对角元素为1且可如下表达的相干矩阵

$$[0148] \quad \Gamma_{V_n V_m} = \frac{\sin c \left(\frac{\Omega f_s l_{nm}}{c} \right)}{1 + \frac{\sigma^2}{\Phi_{VV}}} \quad \forall n \neq m。$$

[0149] 在这些等式中, μ 表示规则化参数(例如,稳定性因数), θ_0 表示束方向, f_s 表示取样速率, Ω 表示信号的角频率, c 表示声速, l 表示邻近扬声器的辐射表面的中心之间的距离, l_{nm} 表示扬声器 n 与 m 的辐射表面的中心之间的距离, Φ_{VV} 表示噪声的正规化交叉功率谱密度矩阵,且 σ^2 表示变换器噪声功率。

[0150] 任务T300和T400可经实施以产生多通道信号以用于驱动具有均匀间距的线性扬声器阵列、具有不均匀间距的线性扬声器阵列,或非线性(例如,经定形)阵列,例如具有一个以上轴的阵列。这些任务还可根据其它方向性场产生原理来实施,例如基于例如波前传播的惠更斯原理的波场合成(WFS)技术。

[0151] 对于其中扬声器阵列为非线性、具有不均匀间距和/或具有一个以上轴的实施案,可为合意的是实施任务T300以使用成对波束成形-空成形(BFNF)配置来计算滤波器系数值。此技术基于成对基础来计算系数,其中扬声器可在两个或两个以上对之间共享。图26A到27中的方法证明了可在每一频率区间独立地应用的BFNF技术。

[0152] 在此情况下,我们扩增如图26A所示的针对每一对的导引向量,其中 λ 是调节因数。针对束方向 θ_i 、频率 f 和扬声器编号 m (对内的1或2)的对 p 和源 n 的导引向量的每一元素可计算为

$$[0153] \quad d_{p,m}^n = \exp \left\{ \frac{-j \omega f_s (m-1) l_p \cos \theta_i}{c} \right\},$$

[0154] 其中 l_p 指示对 p 的扬声器之间的距离, ω 指示频率区间编号,且 f_s 指示取样频率。(此特定实例还说明对三个声场的支持,其各自基于不同的输入音频信号且在不同方向上集中,这可用以支持如本文针对不同用户独立地描述的手势控制。)使用如图26A所示的伪逆运算允许使用非方形矩阵。对于例如如图26B说明的三扬声器情况(即,两个扬声器对),

行数 $2*2=4$ 而非3,使得额外行使矩阵为非方形的。由于此方法是基于稳健的1-D DOA估计,因此不需要对阵列几何形状的完全了解,且也不需要同时使用所有扬声器的声场方向估计。图26C展示如图26A所示的BFNF的实例,其也包含正规化(例如,以防止空间混叠频率处的不良条件的逆)。

[0155] 图27展示成对正规化MVDR(最小方差无失真响应)BFNF的实例,其中获得导引向量(阵列流形向量)的方式不同于常规方法。在此情况下,由于两个对之间的扬声器的共享而消除共同信道。通过测量或通过使用正弦函数的理论计算可获得噪声相干矩阵 Γ 。

[0156] 方法M100可经实施以包含进入手势控制模式的初始手势(例如,抓住、旋转打开手)和/或话音命令的辨识。方法M100可经实施以包含用于用户识别和/或定位的脸和/或话音辨识。方法M100可包含屏幕上显示(OSD)能力以为由任务T200产生的命令提供反馈(例如,显示束强度、束方向、束宽度、动态范围等等的条或拨盘)。

[0157] 方法M100可经实施以包含产生掩蔽信号。此掩蔽可与声音阻挡手势一起使用以使例如被阻挡的声音模糊,或与束聚焦手势一起使用以掩蔽其它方向上的声音。掩蔽信号可为噪声信号,例如白噪声或粉红噪声信号。替代地,掩蔽信号可为频率特性随着时间变化的音乐信号或噪声信号,例如多路重合噪声信号。与白噪声或粉红噪声信号相比,此替代掩蔽信号的使用可较少打扰旁观者和/或较不会使用户分心。

[0158] 对于束聚焦手势,可为合意的是方法M100产生声场以使得在除了用户方向的方向上的声场的谱较类似于噪声(例如,近似为白噪声)。举例来说,可为合意的是掩蔽分量的谱补充源分量的谱。

[0159] 对于包含掩蔽声场(也可为方向性的)的产生的应用,手势也可与响应相关联以控制掩蔽场的一或多个方面(例如,掩蔽声场的产生、删除、方向控制、等级控制和/或质量控制)。

[0160] 所产生的声场可包含一个以上束(例如,两个或两个以上产生的声场的叠加)。举例来说,方法M100的实例可针对两个或两个以上用户中的每一者执行,使得由这些实例产生的相应多通道信号用以驱动同一变换器。

[0161] 图28A展示根据一般配置的设备A100的框图,其包含音频输出级A010、手势解译器GI10、命令解译器CI10以及合成器SY10。音频输出级A010经配置以驱动方向可控变换器以产生束(例如,如本文参考任务T100所述)。手势解译器GI10经配置以响应于手势的表示而产生命令(例如,如本文参考任务T200所述)。命令解译器CI10经配置以响应于命令而产生滤波器配置(例如,如本文参考任务T300所述)。合成器SY10经配置以将空间方向性滤波器应用于输入信号以产生多通道信号,其中空间方向性滤波器是基于滤波器配置(例如,如本文参考任务T400所述)。

[0162] 手势解译器GI10可经实施以使用手势表示作为进入表的索引,所述表使手势表示与如本文描述的对命令相关联。手势表示中的至少一些可伴随有使如本文所述的所指示改变量化的参数值。对于例如以下手势中的任一者,手势解译器GI10可经实施以接收对应命令的一或多个参数的相关值:声音阻挡手势—阻挡方向和/或衰减程度;静音手势—衰减程度;束重定向手势—目的地方向和/或束宽度;束聚焦手势—束宽度;举起/放下手手势—衰减或放大程度、将强调或衰减的频带;压缩/扩展手势—压缩或扩展的程度。

[0163] 设备A100可经实施以使得声场在从由所辨识手势指示的两个或两个以上固定选

项当中自动选择的方向上集中。举例来说,设备A100可经实施以使得声场在对应于左区、中心区和右区当中的一者的方向上集中。在此些情况下,设备A100可经实施以使得针对待选择的每一方向包含一组因数(例如,滤波器系数)的滤波器配置经离线计算(例如,在设计和/或制造期间)以用于在运行时由命令解译器CI10选择和由合成器SY10应用于源信号。在此情况下用于用于左、中心和右区的对应相应方向的一个实例是(45,90,135)度。其它实例包含(不限于)(30,90,150)和(60,90,120)度。图23A和23B分别展示具有三个和五个可选固定空间扇区的方案的实例。命令解译器CI10的此实施方案可经配置以根据估计用户范围在用于选定方向的不同源束宽度当中进行选择。举例来说,当用户较远离阵列时可选择较窄的束(例如,以在不同范围的用户位置处获得相似的束宽度)。

[0164] 图28B展示包含场景分析器SA10的设备A100的实施方案A110的框图。场景分析器SA10经配置以分析所俘获场景以产生手势的表示(例如,如本文参考任务TA10所述)。举例来说,场景分析器SA10可经实施以对所俘获场景执行一或多个图像分析操作,例如图像分段、肤色检测和/或背景消减,和/或检测随着时间的改变的一或多个操作,例如计算视频序列的邻近帧之间的差图像。

[0165] 场景分析器SA10可经实施以对所分析场景执行一或多个任务,例如本文所述的特征提取和/或分类。场景分析器SA10可经实施以将特定手势检测为手势元素的初始状态结合所述元素的一或多个方面中随着时间的改变的组合,如本文所述。场景分析器SA10可经实施以将手势分类为在如本文所述的一组手势候选当中最近的。

[0166] 图28C展示包含俘获装置CD10的设备A110的实施方案A120的框图,所述俘获设备俘获包含手势的场景(例如,如参考任务TC10所述)。俘获装置CD10经配置和布置以记录声场产生于其中的场景中随着时间的改变(例如,移动)。如上文论述,此装置可包含(不限于)一或多个可见光和/或红外相机、一或多个超声变换器和/或一或多个结构化光扫描器。在一个实例中,俘获装置CD10包含便携式计算机(例如,膝上型计算机、桌上型计算机、笔记本电脑或平板计算机)的网络摄像机。在另一实例中,俘获装置CD10包含智能电话的相机。

[0167] 图28D展示包含扬声器阵列R10的设备A100的实施方案A105的框图。在此实例中,音频输出级A010也经布置以基于由合成器SY10产生的多通道信号的对应通道来产生驱动信号以驱动扬声器阵列R10(或另一方向选择性变换器)。

[0168] 合成器SY10还可经实施以对输入信号执行一或多个其它音频处理操作以产生驱动信号。这些操作可包含放大和/或滤波通道中的一或多者(可能全部)。类似地,可为合意的是实施音频输出级A010以将逆滤波器应用于多通道信号以补偿在不同频率下阵列响应的差和/或实施音频输出级A010以补偿阵列的各个扬声器的响应之间的差。替代地或另外,可为合意的是实施音频输出级A010以提供与扬声器阵列(和/或与通向扬声器阵列的音频频率发射路径)的阻抗匹配。

[0169] 对于一些手势,可为合意的是设备A100以空间聚焦方式调整参数(例如,音量、动态范围和/或频率分布)。对于其它手势,可为合意的是设备A100在整个声场上调整此参数。图23D展示可用以全局地和局部地支持此些调整的合成器SY10的实施方案SY20的框图。合成器SY20包含经配置以调整输入音频信号的一或多个参数的单通道滤波器SCF10(例如,音量、动态范围、频率分布)。此调整可根据由命令解译器CI10提供的配置(例如,一或多个滤波器系数)执行。合成器SY20还包含如本文描述的空间方向性滤波器SDF10(或SDF20)的实

例,其接收由滤波器SCF10产生的信号且产生如本文所述的对应多通道驱动信号(例如,参考任务T400)。

[0170] 图29A展示根据一般配置的设备MF100的框图。设备MF100包含用于驱动方向可控变换器以产生束(例如,如本文参考任务T100所述)的装置F100。设备MF100还包含用于响应于手势的表示而产生命令(例如,如本文参考任务T200所述)的装置F200,其中所述手势指示束的方向和束的宽度当中的至少一者的改变,且其中所述命令是执行所指示改变的命令。设备MF100还包含用于响应于所述命令而产生滤波器配置(例如,如本文参考任务T300所述)的装置F300。设备MF100还包含用于对输入信号执行空间方向性滤波操作以产生多通道信号的装置F400,其中空间方向性滤波操作是基于滤波器配置(例如,如本文参考任务T400所述)。

[0171] 图29B展示设备MF100的实施方案MF110的框图,其包含用于分析所俘获场景以产生手势表示(例如,如参考任务TA10所述)的装置FA100。图29C展示设备MF110的实施方案MF120的框图,其包含用于俘获包含手势的场景的装置FC100(例如,如本文参考任务TC10所述)。

[0172] 显然注意到,本文描述的原理不限于与均匀线性扬声器阵列一起使用(例如,如图1C、1D和30A所示)。举例来说,方向性掩蔽也可与在邻近扬声器之间具有不均匀间距的线性阵列一起使用。图30B展示在扬声器之间具有对称八度间距的此阵列的一个实例,且图30C展示具有不对称八度间距的此阵列的另一实例。另外,这些原理不限于与线性阵列一起使用,且也可与元素沿着简单曲线布置的阵列一起使用,无论是具有均匀间距(例如,如图30D所示)还是具有不均匀(例如,八度)间距。本文陈述的相同原理还可分离地应用于沿着相同或不同(例如,正交)直轴或弯曲轴具有多个阵列的应用中的每一阵列。

[0173] 图31B展示将由设备A100的实施方案驱动的扬声器阵列LSL10到LSR10的实例。在此实例中,阵列布置于显示装置TV10(例如,电视机或计算机监视器)中的显示屏幕SC20下方。图31C展示此显示装置TV20中的显示屏幕SC20的任一侧上的阵列LSL10到LSR10的实例。图31A展示包含由设备A100的实施方案驱动的三个扬声器LS10、LS20、LS30的阵列的膝上型计算机D710。如图31A所示的膝上型计算机D710也可经配置以包含此阵列(例如,在底部面板PL20中在键盘后方和/或旁边和/或在顶部面板PL10中在显示屏幕SC10的边限中)。此扬声器阵列也可封闭于一或多个单独橱柜中或安装于例如汽车的交通工具的内部中。

[0174] 在线性阵列的实例中,可预期在向前方向上指向零度的主束也将在向后方向(例如,在180度处)上可听见。在扬声器或麦克风的线性阵列的上下文中常见的此现象也称为“混淆圆锥”问题。

[0175] 虽然展示在180度的范围中的方向性声场产生的特定实例,但本文描述的原理可经延伸以提供跨越平面中的任何所要角范围(例如,在两个维度中)的方向性。此延伸可包含将适当放置的扬声器添加到阵列。举例来说,可为合意的是将扬声器添加到阵列以提供前后阵列,用于同样在前后方向上的掩蔽。图32A和32B分别展示此扩展阵列的两个实例RL200和RL250的俯视图。

[0176] 这些原理也可经延伸以提供跨越空间中的任何所要角范围(3D)的方向性掩蔽。图32C和33分别展示两个阵列RL300和RL400的正视图,其可用以提供左右和上下方向上的方向性掩蔽。另外的实例包含用于在高达360度的范围内(例如,4*pi弧度的完整隐私区)的方

向性掩蔽的球形或其它3D阵列。

[0177] 当使用波束成形技术来产生宽带信号的空间模式时,变换器阵列几何形状的选择涉及低频率与高频率之间的折中。为了增强波束成形器对低频率的直接处置,较大扬声器间距是优选的。同时,如果扬声器之间的间距太大,那么阵列在高频率下再生所要效应的能力将受到较低的混叠阈值限制。为了避免空间混叠,将由阵列再生的最高频率分量的波长应大于邻近扬声器之间的距离的两倍。

[0178] 随着消费型装置变得越来越小,形状因数可约束扬声器阵列的放置。举例来说,可为合意的是膝上型计算机、上网本或平板计算机或高清晰度视频显示器具有内建的扬声器阵列。由于大小约束,扬声器可较小且不能再生所要的低音区域。替代地,可使用具有大到足以再生低音区域的增加的圆锥距离的扬声器。然而在小形状因数装置中,此物理上较大的扬声器仍可能放置太靠近而无法支持低频率下的波束成形或其它声学成像。另外,这些较大扬声器的增加的圆锥距离可实际上减小高频内容的波束成形质量。因此,可为合意的是提供处理以在其中采用波束成形的靠近间隔的扬声器阵列中产生低音信号。

[0179] 存在收听信号的较高谐波可产生听见丢失的基频的感知幻觉的生理声学现象。因此,实现对来自小扬声器的低音分量的感觉的一种方法是从低音分量产生较高谐波且重放谐波而非实际的低音分量。用于代替较高谐波以在不存在实际低频率信号的情况下实现对低音的生理声学感觉(也称为“生理声学低音增强”或PBE)的算法的描述可例如在第5,930,373号美国专利(夏修华等人,1999年7月27日发布)和第2006/0159283A1号美国公开专利申请案(马修等人,2006年7月20日公布)、第2009/0147963A1号美国公开专利申请案(史密斯,2009年6月11日公布)以及第2010/0158272A1号美国公开专利申请案(维克尔,2010年6月24日公布)中找到。此增强可尤其有用于用具有将集成的一或多个扬声器限制为物理上较小的形状因数的装置再生低频率声音。

[0180] 图34展示在PBE处理之前和之后的音乐信号的频谱的实例。在此图中,在大约200到500Hz处可见的背景(黑色)区域和线指示原始信号,且前景(白色)区域指示经增强信号。可见在低频带(例如,低于200Hz)中,PBE操作衰减实际低音的大约10dB。然而由于从大约200Hz到600Hz的经增强较高谐波,当使用小扬声器再生经增强音乐信号时,其被感知为具有比原始信号多的低音。

[0181] 可为合意的是将PBE不仅应用于减少低频可再生性限制的影响,而且减少低频率下的方向性损失的影响。举例来说,可为合意的是组合PBE与波束成形以在波束成形器可导引的范围内产生低频内容的感知。使用扬声器阵列来从经增强信号产生方向性束导致与来自不具有此增强的音频信号的输出相比具有低得多的感知频率范围的输出。另外,变为可使用较松弛的波束成形器设计来导引经增强信号,这可支持假象和/或计算复杂性的减少且允许用小扬声器阵列对低音分量的较有效导引。同时,此系统可保护小扬声器免于低频信号(例如,隆隆声)的破坏。可与如本文描述的手势声场控制结合的此些增强技术的额外描述可在标题为“用于经增强声学成像的系统、方法和设备(SYSTEMS, METHODS, AND APPARATUS FOR ENHANCED ACOUSTIC IMAGING)”的第2012/0020480A1号美国公开专利申请案(维瑟尔等人,2012年1月26日公开)中找到。

[0182] 本文揭示的方法和装置可大体上应用于任何收发和/或音频感测应用,包含此些应用的移动或其它便携式实例和/或来自远场源的信号分量的感测。举例来说,本文揭示的

配置的范围包含驻留于经配置以采用码分多址 (CDMA) 空中接口的无线电话通信系统中的通信装置。然而,所属领域的技术人员将理解,具有本文中所描述的特征的方法及设备可驻留于采用所属领域的技术人员所已知的广泛范围的技术的任何各种通信系统中,例如经由有线及/或无线(例如,CDMA、TDMA、FDMA及/或TD-SCDMA)发射信道采用IP语音 (VoIP) 的系统。

[0183] 明确预期且借此揭示本文揭示的通信装置可适用于包交换(例如,经布置以根据例如VoIP等协议载运音频发射的有线和/或无线网络)和/或电路交换的网络中。还明确预期且进而揭示本文揭示的通信装置可适用于窄带译码系统(例如,对大约四或五千赫的音频频率范围进行编码的系统)中和/或用于宽带译码系统(例如,对大于五千赫的音频频率进行编码的系统)中,包含全带宽带译码系统和分带宽带译码系统。

[0184] 提供所述配置的前述呈现以使所属领域的技术人员能够制作或使用本文揭示的方法和其它结构。本文展示和描述的流程图、框图和其它结构仅为实例,且这些结构的其它变体也在本发明的范围内。对这些配置的各种修改是可能的,且本文中所呈现的一般原理还可应用于其它配置。因此,本发明不希望限于上文所展示的配置,而是应被赋予与在本文中(包含在所申请的附加权利要求书中)以任何方式揭示的原理及新颖特征一致的最广范围,所述权利要求书形成原始揭示内容的一部分。

[0185] 所属领域的技术人员将了解,可使用多种不同技术及技艺中的任一者来表示信息及信号。举例来说,可通过电压、电流、电磁波、磁场或磁性粒子、光场或光学粒子或者其任何组合来表示可在整个以上描述中所参考的数据、指令、命令、信息、信号、位及符号。

[0186] 用于实施如本文揭示的配置的重要设计要求可包含最小化处理延迟和/或计算复杂性(通常以每秒百万指令或MIPS测量),尤其是对于计算密集型应用,例如经压缩音频或视听信息的重放(例如,根据例如本文识别的实例中的一者的压缩格式编码的文件或流)或用于宽带通信(例如,在高于八千赫的取样率下的话音通信,例如12、16、32、44.1、48或192kHz)的应用。

[0187] 如本文揭示的设备(例如,设备A100、A105、A110、A120、MF100、MF110和MF120)可在视为适合于既定应用的硬件与软件和/或与固件的任何组合中实施。举例来说,此设备的元件可制造为驻留于例如同一片上或芯片组中的两个或两个以上芯片当中的电子和/或光学装置。此装置的一个实例为固定或可编程逻辑元件(例如,晶体管或逻辑门)阵列,且这些元件中的任一者可实施为一或多个此类阵列。设备的元件中的任何两个或两个以上或甚至全部可在相同的一或多个阵列内实施。所述阵列可实施于一或多个芯片内(例如,包含两个或两个以上芯片的芯片组内)。

[0188] 本文揭示的设备的各种实施方案的一或多个元件也可完整或部分地实施为一或多个指令集,其经布置以在一或多个固定或可编程逻辑元件阵列上执行,例如微处理器、嵌入式处理器、IP核心、数字信号处理器、FPGA(现场可编程门阵列)、ASSP(专用标准产品)和ASIC(专用集成电路)。如本文揭示的设备的实施方案的各种元件中的任一者也可体现为一或多个计算机(例如,包含经编程以执行一或多个指令集或序列的一或多个阵列的机器,也称为“处理器”),且这些元件中的任何两个或两个以上或甚至全部可在相同的此类一或多个计算机内实施。

[0189] 如本文揭示的处理器或用于处理的其它装置可制造为驻留于例如同一片上或

芯片组中的两个或两个以上芯片当中的一或多个电子和/或光学装置。此装置的一个实例为固定或可编程逻辑元件(例如,晶体管或逻辑门)阵列,且这些元件中的任一者可实施为一或多个此类阵列。此一或多个阵列可实施于一或多个芯片内(例如,包含两个或两个以上芯片的芯片组内)。这些阵列的实例包含固定或可编程逻辑元件阵列,例如微处理器、嵌入式处理器、IP核心、DSP、FPGA、ASSP和ASIC。如本文揭示的处理器或用于处理的其它装置也可体现为一或多个计算机(例如,包含经编程以执行一或多个指令集或序列的一或多个阵列的机器)或其它处理器。如本文描述的处理器可用以执行任务或执行与本文所述的手势控制程序不直接相关的其它指令集,例如与其中嵌入处理器的装置或系统(例如,音频感测装置)的另一操作相关的任务。本文揭示的方法的部分还可由音频感测装置的处理器执行且将在一或多个其它处理器的控制下执行方法的另一部分。

[0190] 所属领域的技术人员将了解,结合本文揭示的配置描述的各种说明性模块、逻辑块、电路和测试以及其它操作可实施为电子硬件、计算机软件或两者的组合。这些模块、逻辑块、电路和操作可用经设计以产生本文揭示的配置的通用处理器、数字信号处理器(DSP)、ASIC或ASSP、FPGA或其它可编程逻辑装置、离散门或晶体管逻辑、离散硬件组件或其任何组合来实施或执行。举例来说,此配置可至少部分地实施为硬连线电路、制造为专用集成电路的电路配置,或加载到非易失性存储装置中的固件程序或作为机器可读代码从数据存储媒体加载或加载到数据存储媒体中的软件程序,此代码为可由例如通用处理器或其它数字信号处理单元等逻辑元件阵列执行的指令。通用处理器可以是微处理器,但在替代例中,处理器可为任何常规处理器、控制器、微控制器或状态机。处理器还可实施为计算装置的组合,例如DSP与微处理器的组合、多个微处理器、结合DSP核心的一或多个微处理器或任何其它此类配置。软件模块可驻留于非暂时性存储媒体中,例如随机存取存储器(RAM)、只读存储器(ROM)、例如快闪RAM的非易失性RAM(NVRAM)、可擦除可编程ROM(EPROM)、电可擦除可编程ROM(EEPROM)、寄存器、硬盘、可装卸式磁盘、CD ROM或此项技术中已知的任何其它形式的存储媒体。说明性存储媒体耦合到处理器,使得处理器可从存储媒体读取信息及将信息写入到存储媒体。在替代例中,存储媒体可与处理器成一体式。处理器及存储媒体可驻留于ASIC中。ASIC可驻留于用户终端中。在替代例中,处理器和存储媒体可作为离散组件驻留于用户终端中。

[0191] 应注意,本文揭示的各种方法(例如,方法M100、M110、M120及其实施方案)可由例如处理器等逻辑元件阵列执行,且如本文描述的设备的各种元件可实施为经设计以在此阵列上执行的模块。如本文使用,术语“模块”或“子模块”可指代以软件、硬件或固件形式包含计算机指令(例如,逻辑表达式)的任何方法、设备、装置、单元或计算机可读数据存储媒体。应了解,多个模块或系统可组合为一个模块或系统,且一个模块或系统可被分离成多个模块或系统以执行相同功能。当以软件或其它计算机可执行指令实施时,过程的要素本质上为用以执行例如与例程、程序、对象、组件、数据结构等有关的任务的代码段。术语“软件”应理解为包含源代码、汇编语言代码、机器代码、二进制代码、固件、宏码、微码、可由逻辑元件阵列执行的任何一或多个指令集或序列以及此类实例的任何组合。程序或代码段可存储在处理器可读存储媒体中或由体现于发射媒体或通信链路上的载波中的计算机数据信号发射。

[0192] 本文揭示的方法、方案和技术的实施方案也可有形地体现(例如,在本文列出的一

或多个计算机可读媒体中)为可由包含逻辑元件阵列的机器(例如,处理器、微处理器、微控制器或其它有线状态机)读取和/或执行的一或多个指令集。术语“计算机可读媒体”可包含可存储或传送信息的任何媒体,包含易失性、非易失性、可装卸式和不可装卸式媒体。计算机可读媒体的实例包含电子电路、半导体存储器装置、ROM、快闪存储器、可擦除ROM(EROM)、软盘或其它磁性存储装置、CD-ROM/DVD或其它光学存储装置、硬盘、光纤媒体、射频(RF)链路,或可用于存储所要信息且可被存取的任何其它媒体。计算机数据信号可包含可经由传输媒体(例如电子网络通道、光纤、空气、电磁、RF链路等)传播的任何信号。代码段可经由例如因特网或企业内部网等计算机网络来下载。在任何情况下,本发明的范围不应被解释为受此些实施例限制。

[0193] 本文描述的方法的任务中的每一者可直接以硬件、以由处理器执行的软件模块或以两者的组合体现。在如本文揭示的方法的实施方案的典型应用中,逻辑元件(例如,逻辑门)阵列经配置以执行方法的各种任务中的一者、一者以上或甚至全部。所述任务中的一或多者(可能所有)还可实施为在计算机程序产品(例如,一或多个数据存储媒体,例如磁盘、快闪或其它非易失性存储器卡、半导体存储器芯片等)中体现的代码(例如,一或多个指令集),所述计算机程序产品可由包含逻辑元件阵列(例如,处理器、微处理器、微控制器或其它有限状态机)的机器(例如,计算机)读取和/或执行。如本文揭示的方法的实施方案的任务也可由一个以上此阵列或机器执行。在这些或其它实施方案中,所述任务可在用于无线通信的装置内执行,所述装置例如为蜂窝式电话或具有此通信能力的其它装置。此装置可经配置以与电路交换和/或包交换网络通信(例如,使用一或多个协议(例如VoIP))。举例来说,此装置可包含经配置以接收和/或发射经编码帧的RF电路。

[0194] 明确地揭示本文揭示的各种方法可由便携式通信装置执行,例如手持机、头戴式耳机或便携式数字助理(PDA),且本文描述的各种设备可包含在此装置内。典型的实时(例如,在线)应用是使用此类移动装置进行的电话对话。

[0195] 在一或多个示范性实施例中,本文描述的操作可以硬件、软件、固件或其任一组合来实施。如果以软件实施,那么此些操作可作为一或多个指令或代码存储在计算机可读媒体上或经由计算机可读媒体传输。术语“计算机可读媒体”包含计算机可读存储媒体和通信(例如,传输)媒体两者。举例来说且并非限制,计算机可读存储媒体可包括存储元件阵列,例如半导体存储器(可包含(不限于)动态或静态RAM、ROM、EEPROM和/或快闪RAM)或铁电体、磁阻、双向半导体、聚合的或相变存储器;CD-ROM或其它光盘存储装置;和/或磁盘存储或其它磁性存储装置。此类存储媒体可以可由计算机存取的指令或数据结构的形式存储信息。通信媒体可包括可用以用指令或数据结构的形式载运所要程序代码且可由计算机存取的任何媒体,包含促进计算机程序从一处到另一处的转移的任何媒体。而且,适当地将任何连接称作计算机可读媒体。举例来说,如果使用同轴电缆、光纤电缆、双绞线、数字订户线(DSL)或例如红外线、无线电和/或微波的无线技术从网站、服务器或其它远程源传输软件,那么同轴电缆、光纤电缆、双绞线、DSL或例如红外线、无线电和/或微波的无线技术包含于媒体的定义中。如本文中所使用,磁盘及光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软盘及Blu-ray Disc™(蓝光光盘协会,加利福尼亚州大学城),其中磁盘通常以磁性方式再生数据,而光盘用激光以光学方式再生数据。以上各项的组合也应包含在计算机可读媒体的范围内。

[0196] 如本文描述的声信号处理设备(例如,设备A100、A105、A110、A120、MF100、MF110或MF120)可并入到电子装置中,所述装置接受语音输入以便控制某些操作,或可另外得益于所要噪声与背景噪声的分离,例如通信装置。许多应用可得益于增强或分离清晰的所要声音与源自多个方向的背景声音。这些应用可包含电子或计算装置中的人机接口,其并入有例如话音辨识及检测、语音增强及分离、话音激活式控制等能力。可为合意的是实施此声信号处理设备以适于仅提供有限处理能力的装置中。

[0197] 本文描述的模块、元件和装置的各种实施方案的元件可制造为例如驻留在同一芯片上或芯片组中的两个或两个以上芯片当中的电子和/或光学装置。此装置的一个实例为固定或可编程逻辑元件(例如,晶体管或门)的阵列。文中所描述的设备的各种实施方案的一或多个元件还可全部地或部分地实施为经布置以在一或多个固定或可编程逻辑元件阵列(例如,微处理器、嵌入式处理器、IP核心、数字信号处理器、FPGA、ASSP及ASIC)上执行的一或多个指令集。

[0198] 有可能使如本文中所述的设备的实施方案的一或多个元件用于执行并非与所述设备的操作直接相关的任务或其它指令集,例如与嵌入有所述设备的装置或系统的另一操作相关的任务。还有可能使此设备的实施方案的一或多个元件具有共同结构(例如,用于在不同时间执行对应于不同元件的代码部分的处理器、经执行以在不同时间执行对应于不同元件的任务的指令集,或在不同时间对不同元件执行操作的电子和/或光学装置的布置)。

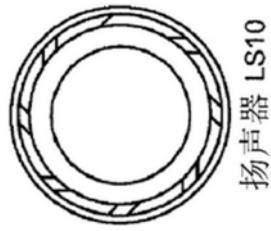


图1A



图1B

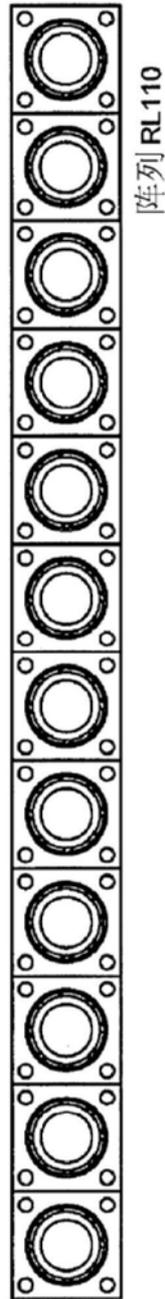


图1C

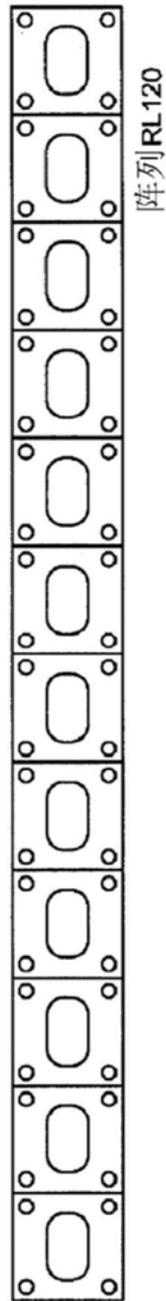


图1D



图2B



图2A

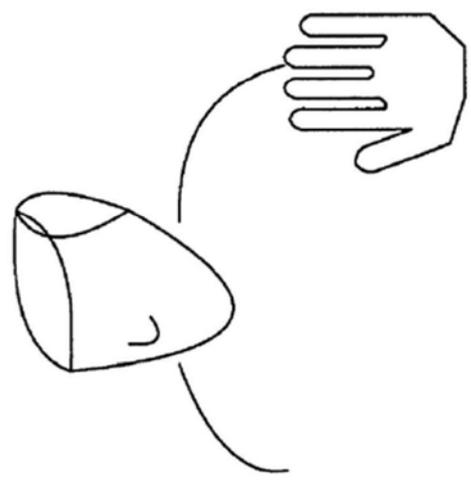
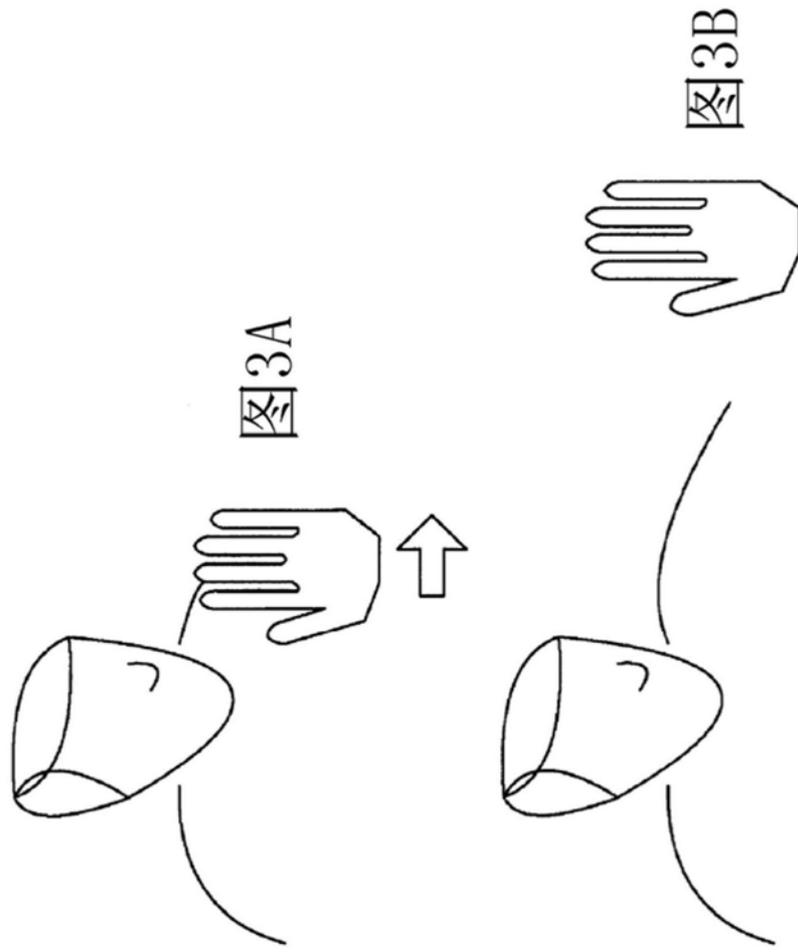


图4A



图4B



图5B

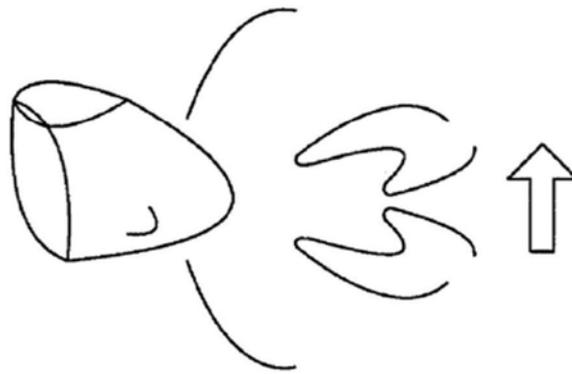


图5A

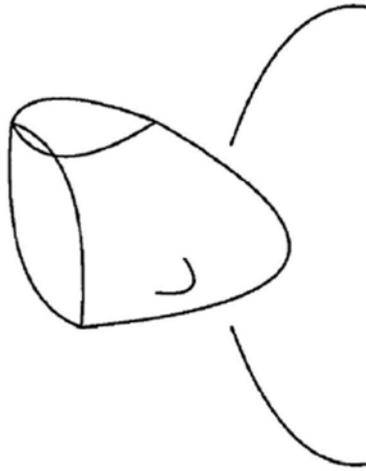


图6A



图6B

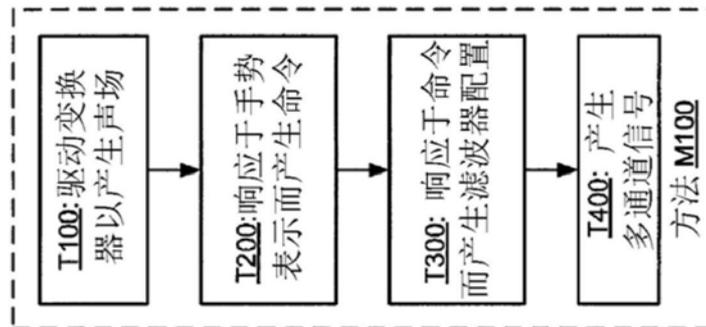


图7A

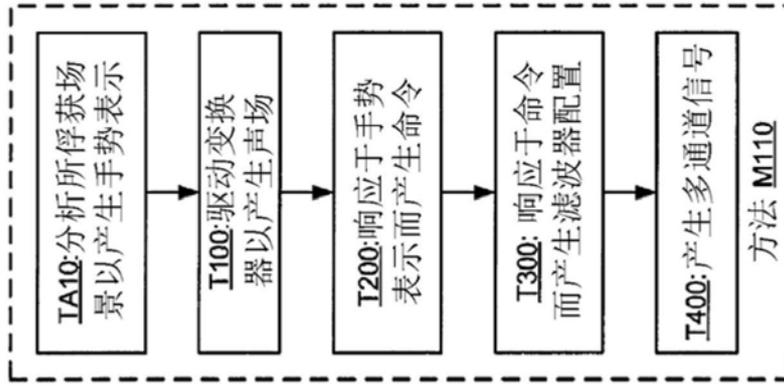


图7B

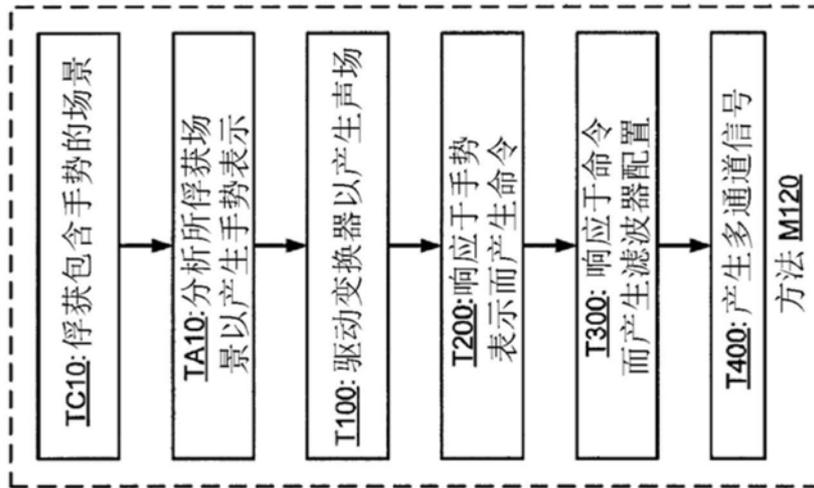


图7C

表示的手势	对应的命令
GA1) 阻挡	CA1) 减小束强度
GA2) 招手	CA2) 增加束强度
GB1) 闭合手, 平移, 打开	CB1) 改变束方向
GB2) 将手带到一起, 平移, 使手移动分开	CB2) 改变束方向
GC1) 使手水平地朝向彼此移动	CC1) 减小束宽度
GC2) 使手水平地移动远离彼此	CC2) 增加束宽度
GD1) 将手举起到耳朵	CD1) 产生束
GD2) 摆手	CD2) 删除束
GE1) 一起举起手	CE1) 增加声场强度
GE2) 一起放下手	CE2) 减小声场强度
GF1) 使手垂直地朝向彼此移动	CF1) 压缩(例如, 场或束的)动态范围
GF2) 使手垂直地移动远离彼此	CF2) 扩展(例如, 场或束的)动态范围

图8

表示的手势	对应的命令
GF3) 将一只手保持到耳朵同时放下另一只手	CF3) 减小束强度
GF4) 将一只手保持到耳朵同时举起另一只手	CF4) 增加束强度
GG1a) 将一只手保持到耳朵同时保持另一只手 打开手掌直立高于下巴水平且顺时针旋转	CG1a) 增加(例如, 场或束的)高音水平
GG2a) 将一只手保持到耳朵同时保持另一只手 打开手掌直立高于下巴水平且逆时针旋转	CG2a) 减小(例如, 场或束的)高音水平
GG1b) 将一只手保持到耳朵同时保持另一只手 打开手掌直立处于下巴水平且顺时针旋转	CG1b) 增加(例如, 场或束的)中音水平
GG2b) 将一只手保持到耳朵同时保持另一只手 打开手掌直立处于下巴水平且逆时针旋转	CG2b) 减小(例如, 场或束的)中音水平
GG1c) 将一只手保持到耳朵同时保持另一只手 打开手掌直立低于下巴水平且顺时针旋转	CG1c) 增加(例如, 场或束的)低音水平
GG2c) 将一只手保持到耳朵同时保持另一只手 打开手掌直立低于下巴水平且逆时针旋转	CG2c) 减小(例如, 场或束的)低音水平
GH1) 锁定(闭合的手的顺时针旋转)	CH1) 停用对声场配置的改变
GH2) 解锁(闭合的手的逆时针旋转)	CH2) 启用对声场配置的改变

图9

表示的手势	对应的命令
GC1a) 使手水平地朝向彼此移动到肩宽度	CC1a) 使束宽度减小一个单位(例如, 1 dB)
GC2a) 从肩宽度使手水平地移动远离彼此	CC2a) 将束宽度增加到最大值
GC1b) 从肩宽度使手水平地朝向彼此移动	CC1b) 将束宽度减小到最小值
GC2b) 使手水平地移动远离彼此到肩宽度	CC2b) 使束宽度增加一个单位(例如, 1 dB)

图10

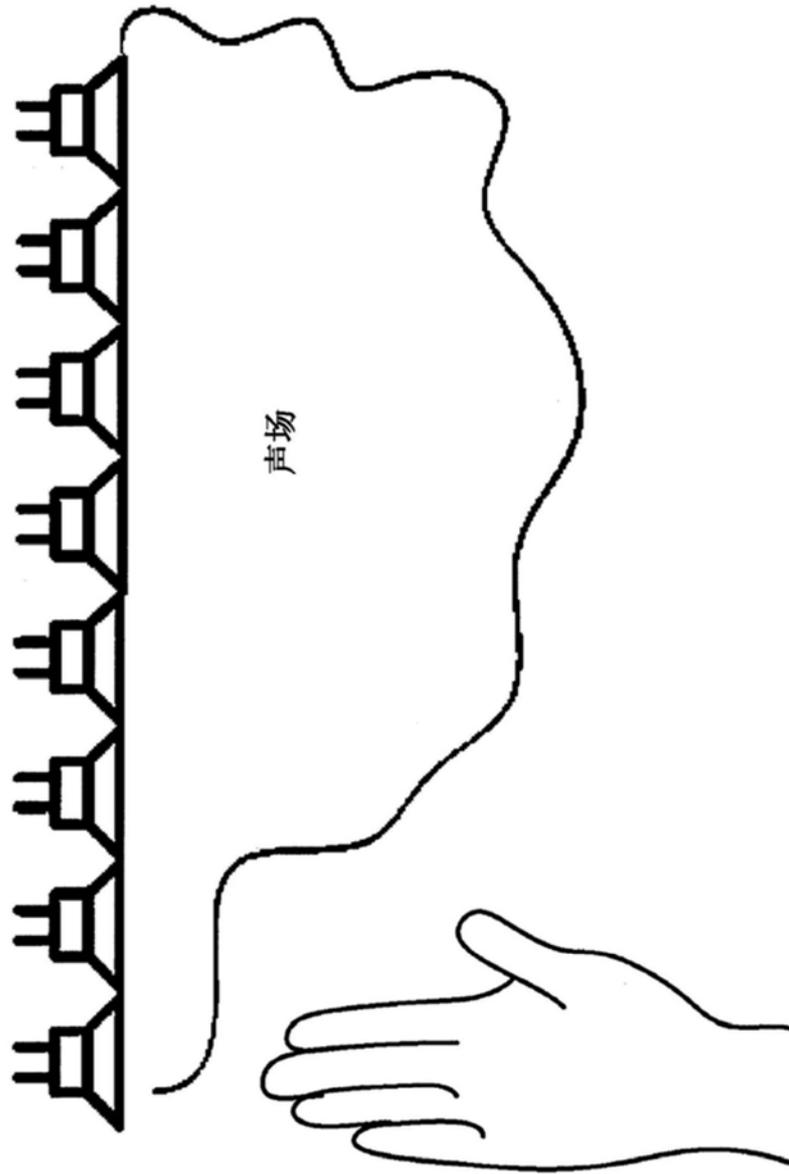


图11

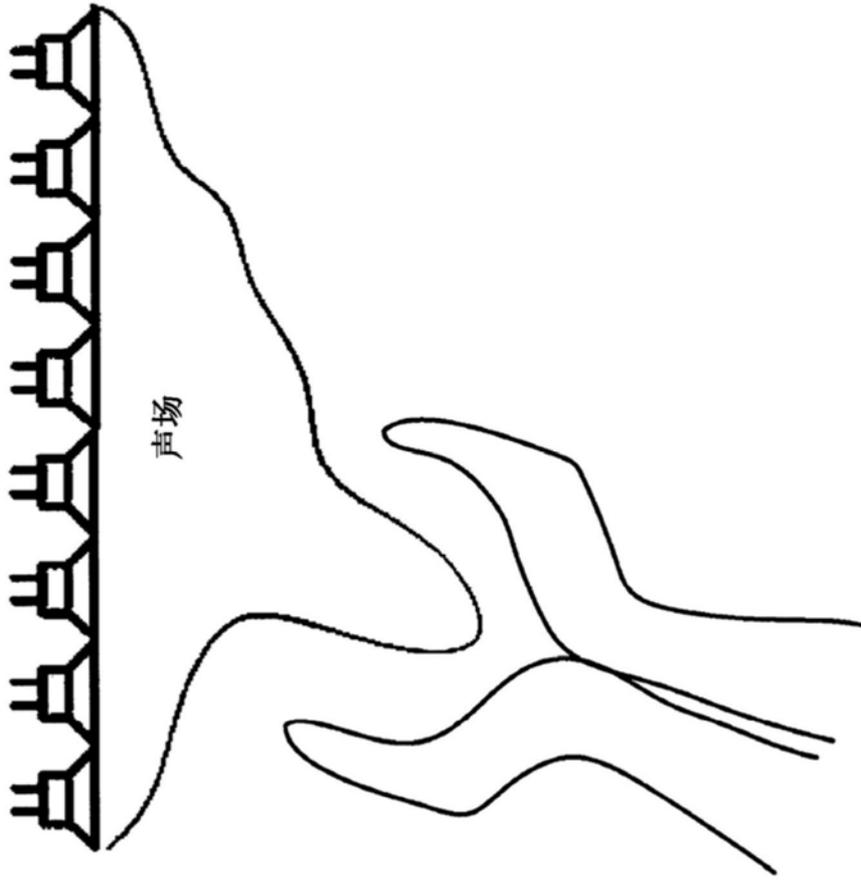


图12

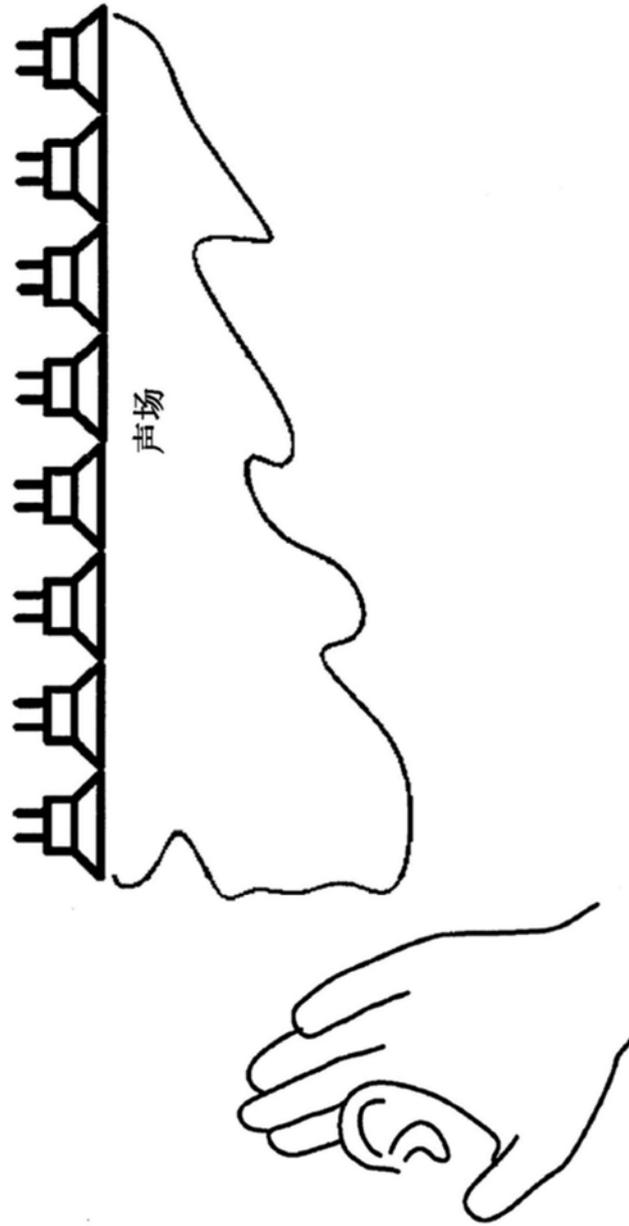


图13

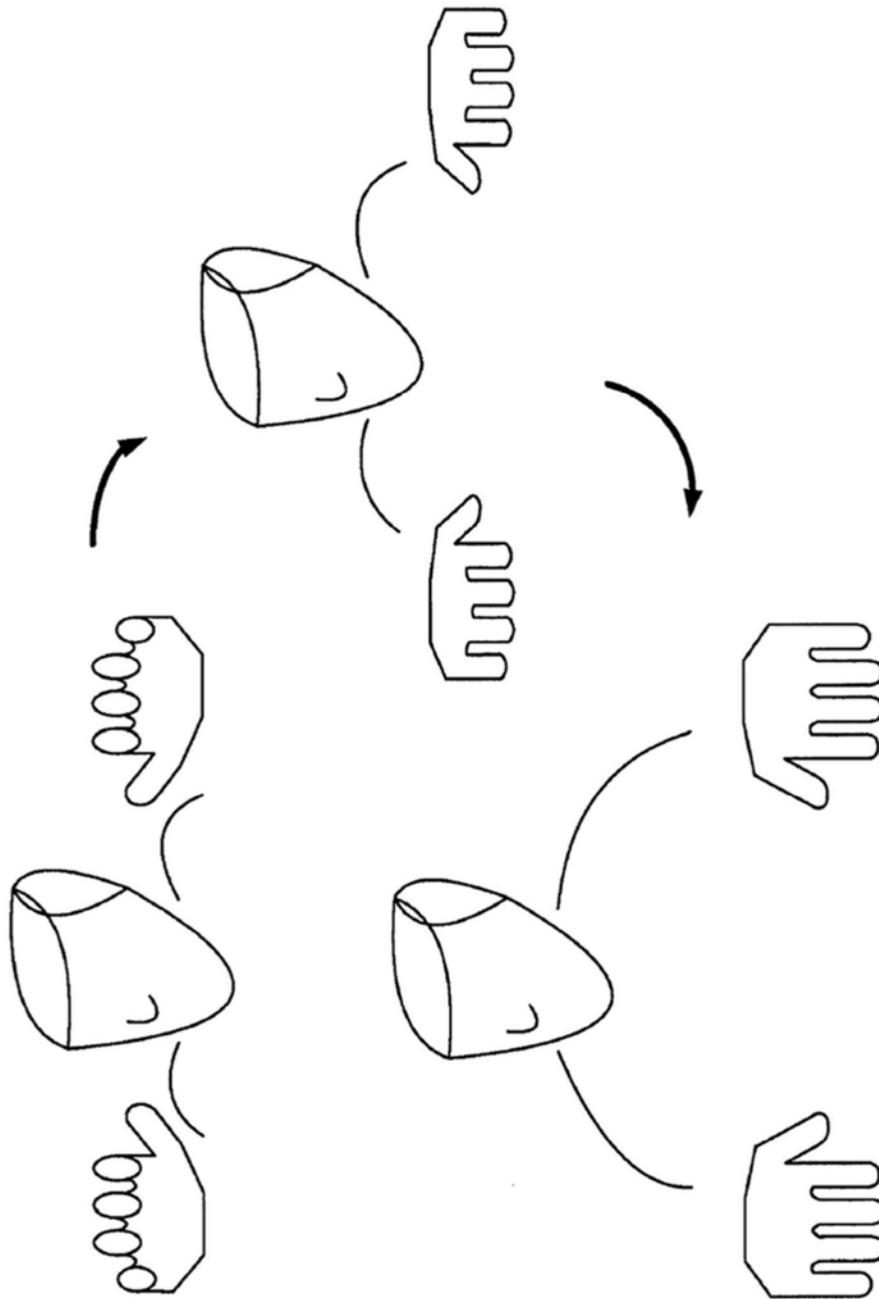


图14

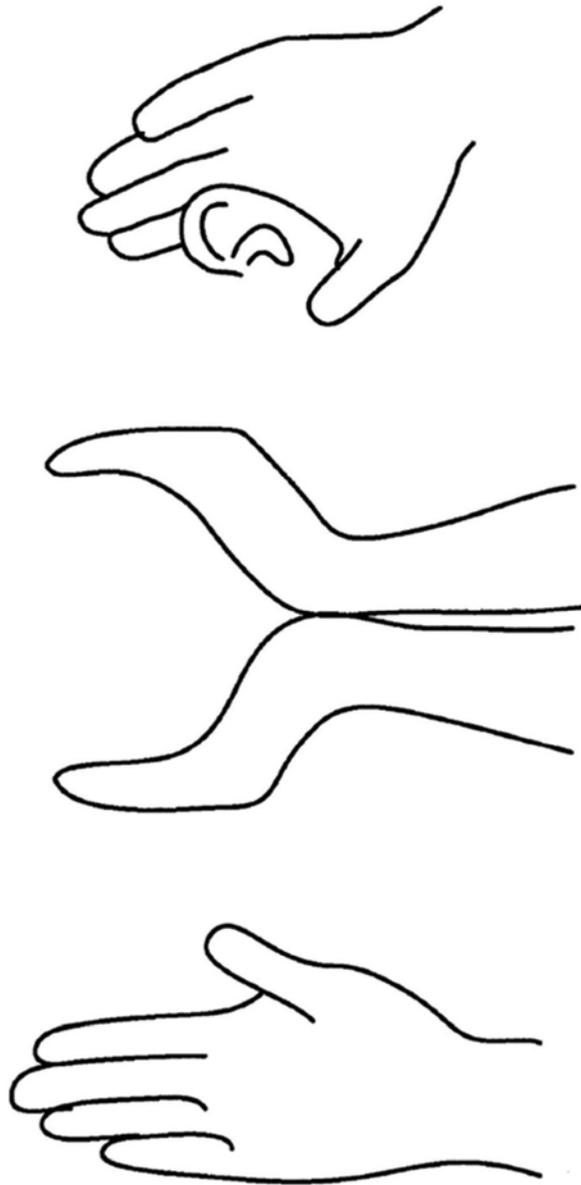


图15

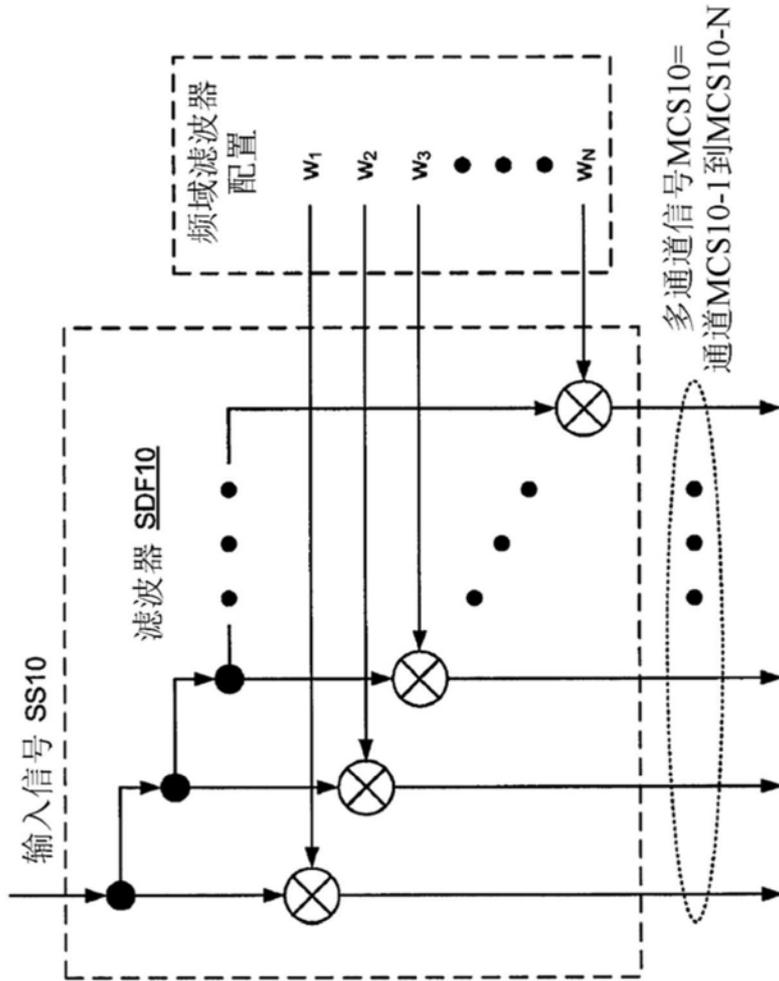


图16

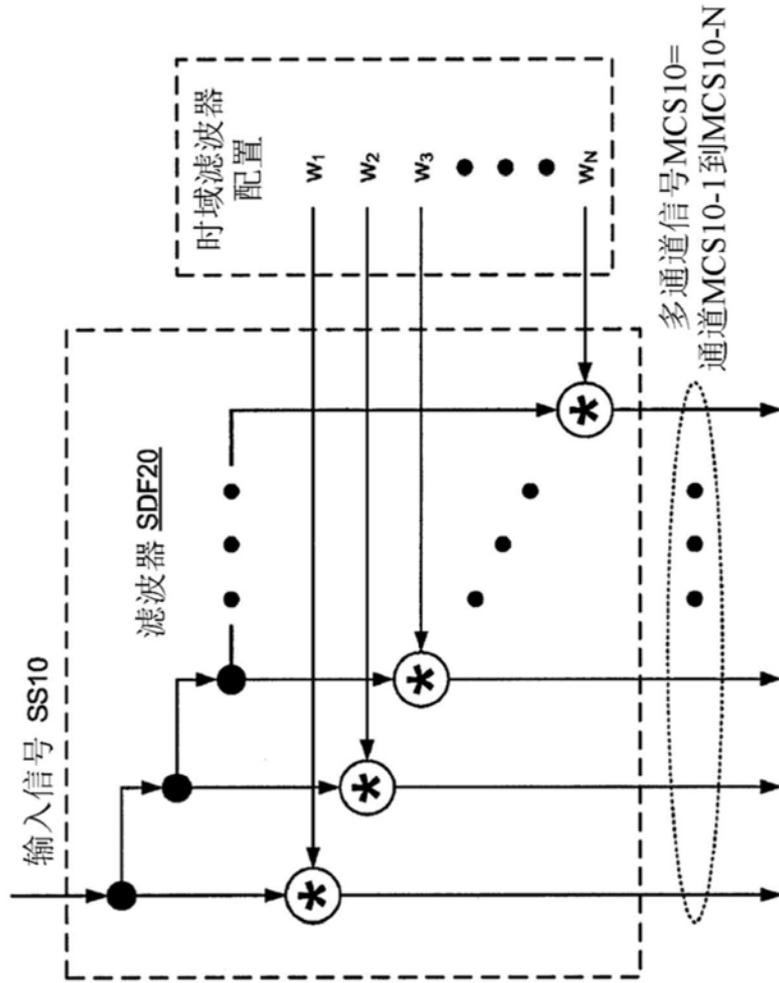


图17

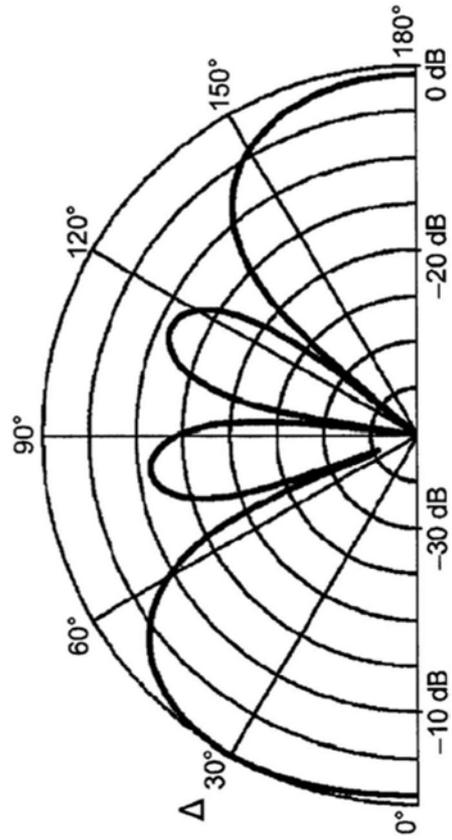


图18A

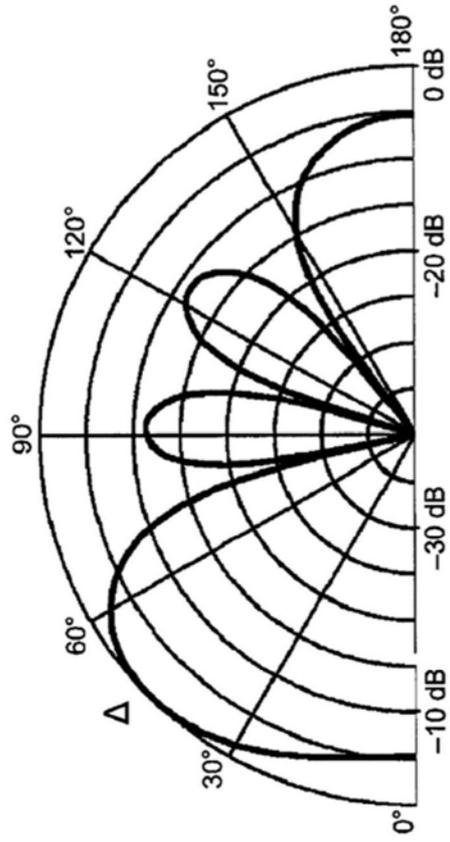


图18B

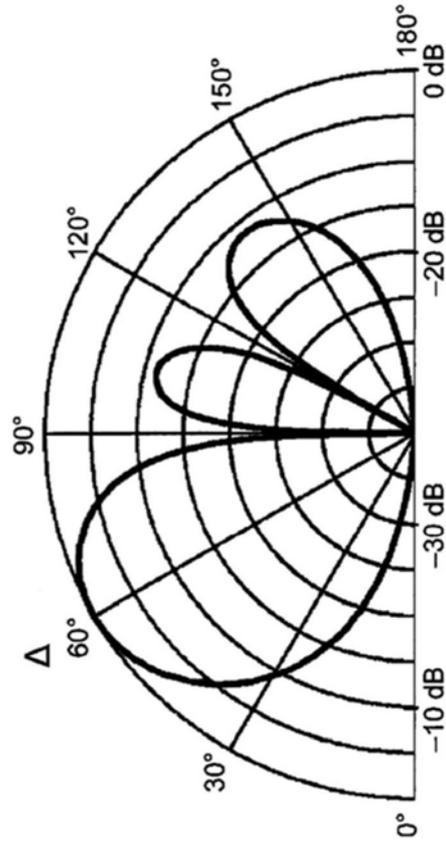


图19A

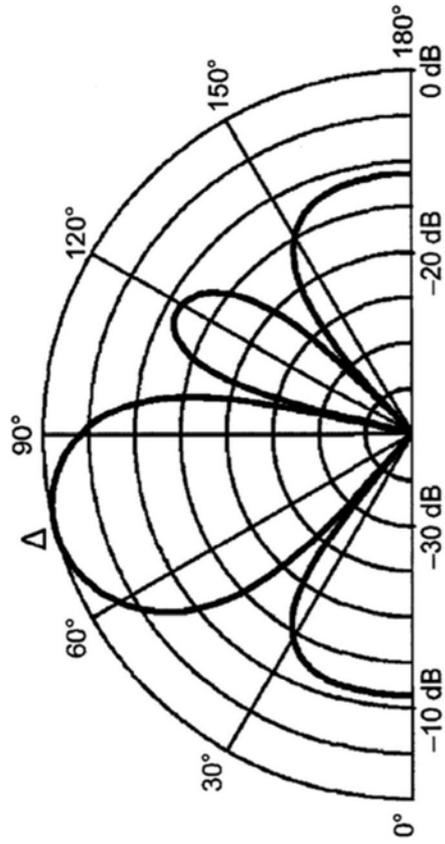


图19B

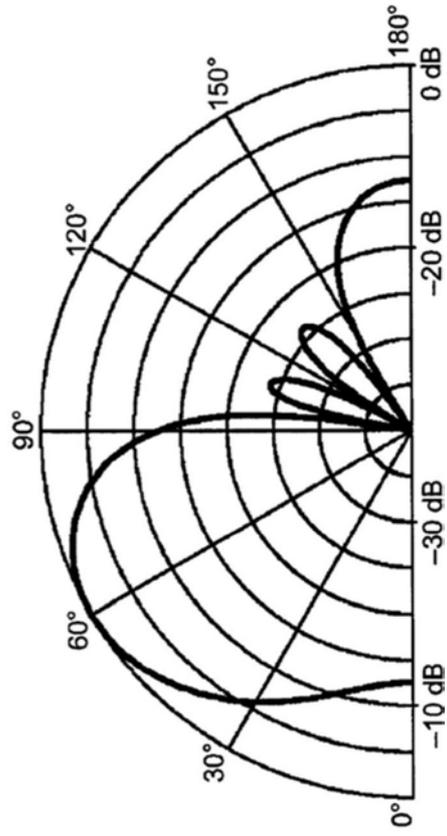


图20A

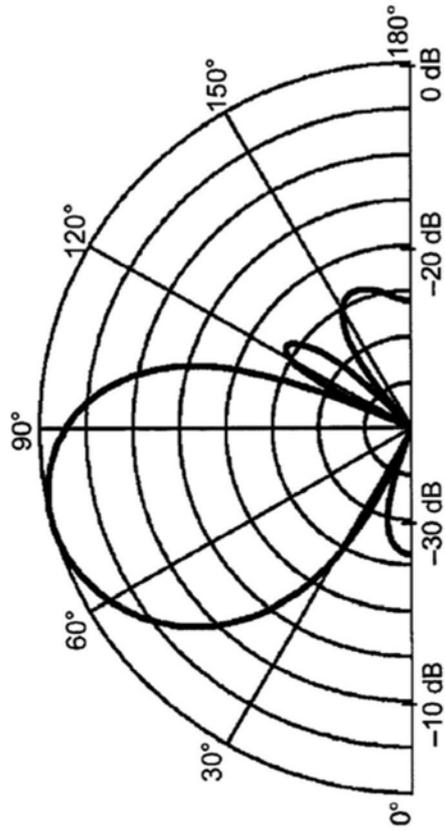


图20B

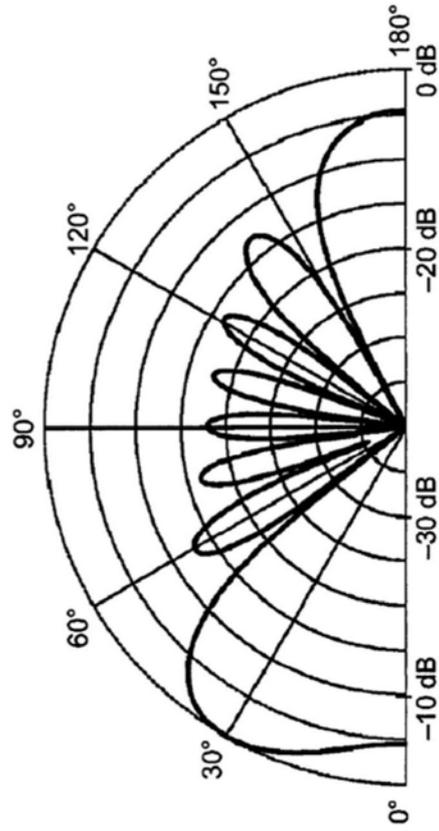


图21A

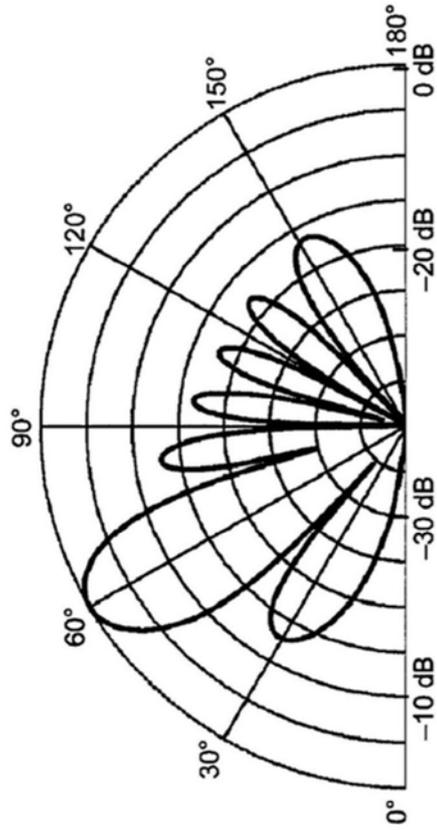


图21B

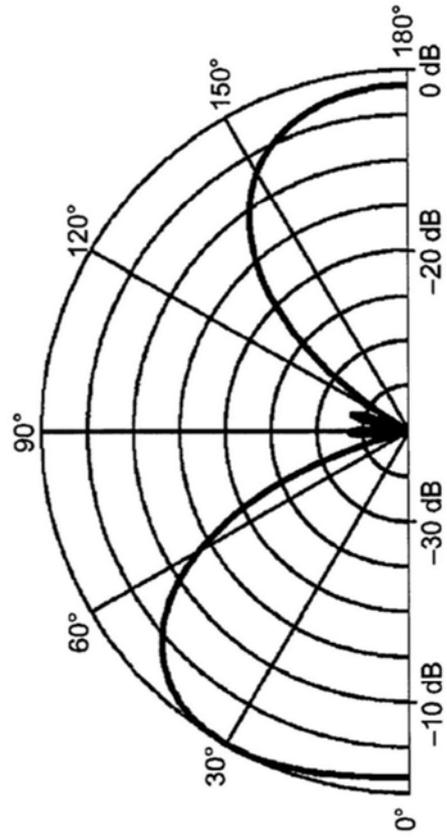


图22A

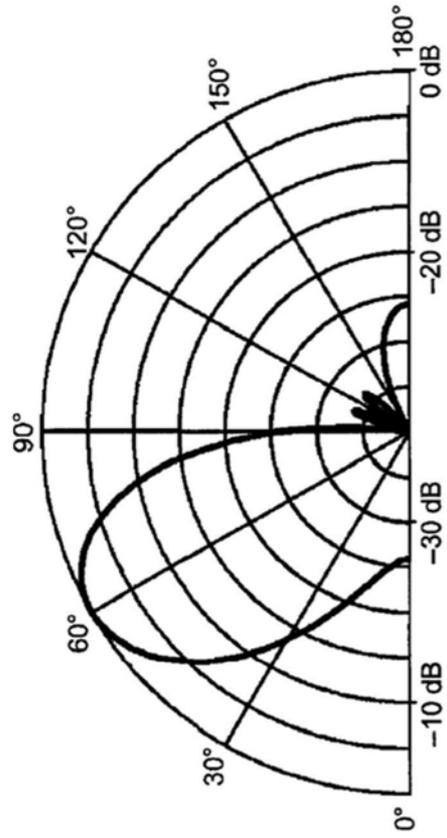


图22B

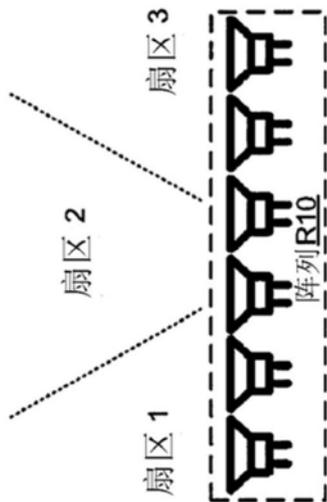


图23A

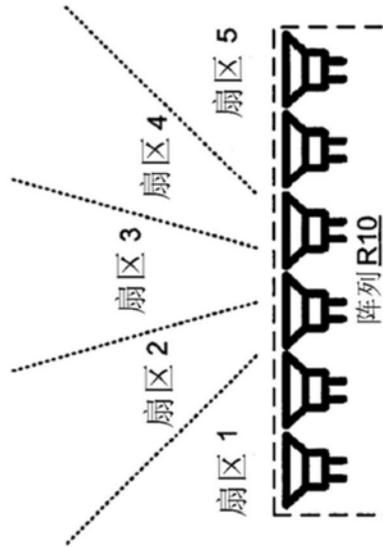


图23B

命令	⇕	窄	中	宽
	↓	X	窄	窄
	↑	X	窄	中
	⇕	宽	宽	X

图23C

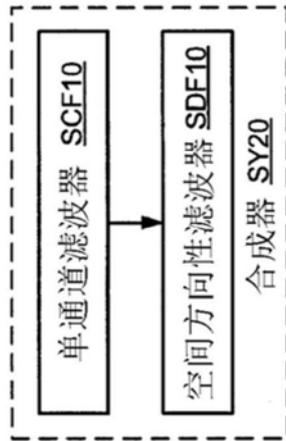


图23D

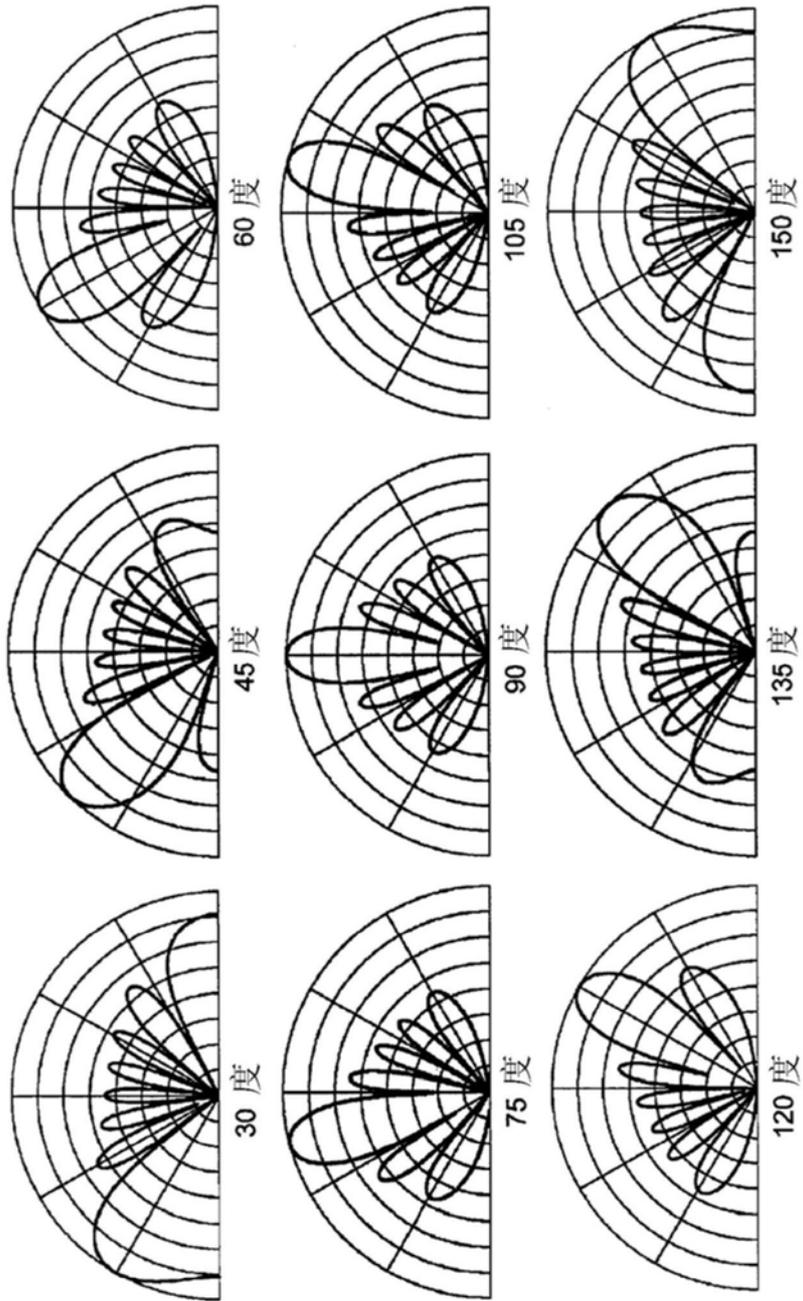


图24

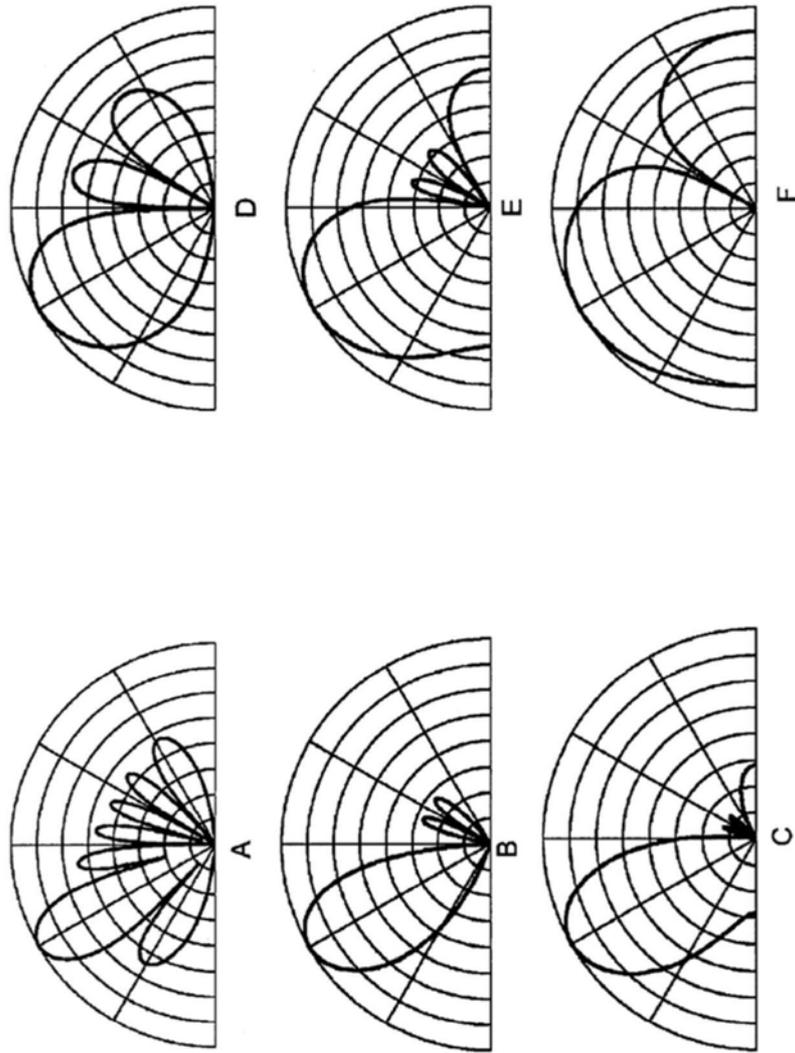


图25

导引向量 目标源编号

对编号 扬声器编号

$$\begin{bmatrix} d_{1,1}^1 & d_{1,1}^2 & d_{1,1}^3 \\ d_{1,2}^1 & d_{1,2}^2 & d_{1,2}^3 \\ d_{2,1}^1 & d_{2,1}^2 & d_{2,1}^3 \\ d_{2,2}^1 & d_{2,2}^2 & d_{2,2}^3 \end{bmatrix}^H \begin{bmatrix} d_{1,1}^1 & d_{1,1}^2 & d_{1,1}^3 \\ d_{1,2}^1 & d_{1,2}^2 & d_{1,2}^3 \\ d_{2,1}^1 & d_{2,1}^2 & d_{2,1}^3 \\ d_{2,2}^1 & d_{2,2}^2 & d_{2,2}^3 \end{bmatrix} + \lambda \cdot I \Bigg)^{-1} \begin{bmatrix} d_{1,1}^1 & d_{1,1}^2 & d_{1,1}^3 \\ d_{1,2}^1 & d_{1,2}^2 & d_{1,2}^3 \\ d_{2,1}^1 & d_{2,1}^2 & d_{2,1}^3 \\ d_{2,2}^1 & d_{2,2}^2 & d_{2,2}^3 \end{bmatrix}^H \begin{bmatrix} x_{1,1} \\ x_{1,2} \\ x_{2,1} \\ x_{2,2} \end{bmatrix}$$

图26A

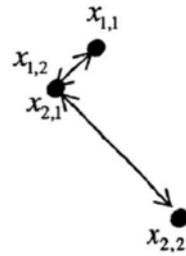


图26B

$$\frac{\left(\begin{bmatrix} d_{1,1}^1 & d_{1,1}^2 & d_{1,1}^3 \\ d_{1,2}^1 & d_{1,2}^2 & d_{1,2}^3 \\ d_{2,1}^1 & d_{2,1}^2 & d_{2,1}^3 \\ d_{2,2}^1 & d_{2,2}^2 & d_{2,2}^3 \end{bmatrix}^H \begin{bmatrix} d_{1,1}^1 & d_{1,1}^2 & d_{1,1}^3 \\ d_{1,2}^1 & d_{1,2}^2 & d_{1,2}^3 \\ d_{2,1}^1 & d_{2,1}^2 & d_{2,1}^3 \\ d_{2,2}^1 & d_{2,2}^2 & d_{2,2}^3 \end{bmatrix} + \lambda \cdot I \right)^{-1} \begin{bmatrix} d_{1,1}^1 & d_{1,1}^2 & d_{1,1}^3 \\ d_{1,2}^1 & d_{1,2}^2 & d_{1,2}^3 \\ d_{2,1}^1 & d_{2,1}^2 & d_{2,1}^3 \\ d_{2,2}^1 & d_{2,2}^2 & d_{2,2}^3 \end{bmatrix}^H \begin{bmatrix} x_{1,1} \\ x_{1,2} \\ x_{2,1} \\ x_{2,2} \end{bmatrix}}{\arg \max_i \left(\begin{bmatrix} d_{1,1}^1 & d_{1,1}^2 & d_{1,1}^3 \\ d_{1,2}^1 & d_{1,2}^2 & d_{1,2}^3 \\ d_{2,1}^1 & d_{2,1}^2 & d_{2,1}^3 \\ d_{2,2}^1 & d_{2,2}^2 & d_{2,2}^3 \end{bmatrix}^H \begin{bmatrix} d_{1,1}^1 & d_{1,1}^2 & d_{1,1}^3 \\ d_{1,2}^1 & d_{1,2}^2 & d_{1,2}^3 \\ d_{2,1}^1 & d_{2,1}^2 & d_{2,1}^3 \\ d_{2,2}^1 & d_{2,2}^2 & d_{2,2}^3 \end{bmatrix} + \lambda \cdot I \right)^{-1} \begin{bmatrix} d_{1,1}^1 & d_{1,1}^2 & d_{1,1}^3 \\ d_{1,2}^1 & d_{1,2}^2 & d_{1,2}^3 \\ d_{2,1}^1 & d_{2,1}^2 & d_{2,1}^3 \\ d_{2,2}^1 & d_{2,2}^2 & d_{2,2}^3 \end{bmatrix}^H \begin{bmatrix} d_{1,1}^i \\ d_{1,2}^i \\ d_{2,1}^i \\ d_{2,2}^i \end{bmatrix}}$$

图26C

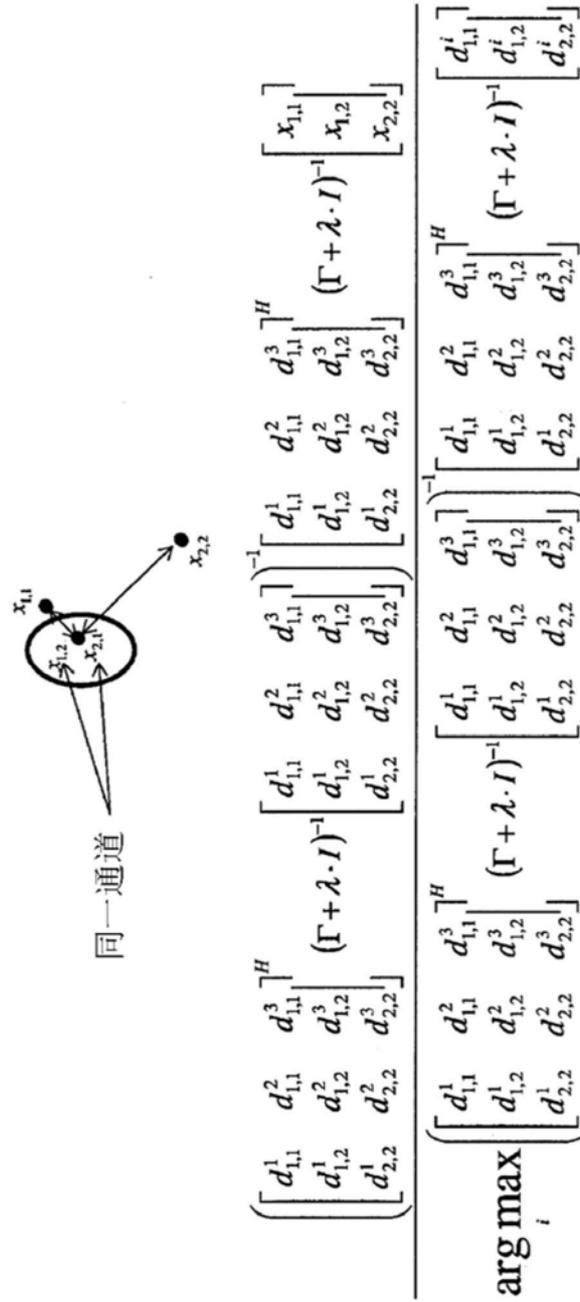


图27

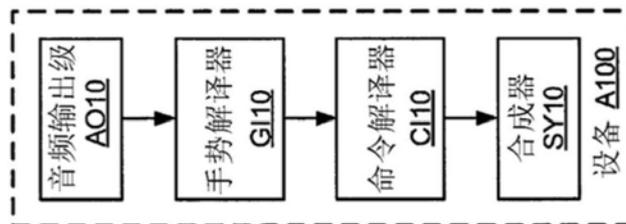


图28A

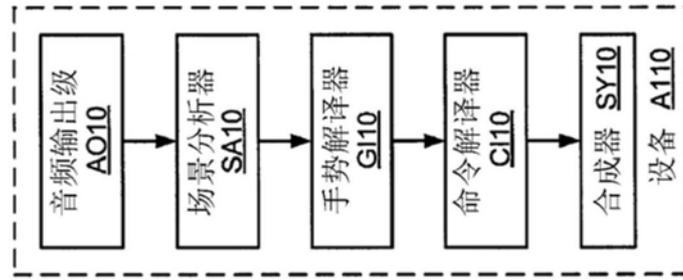


图28B

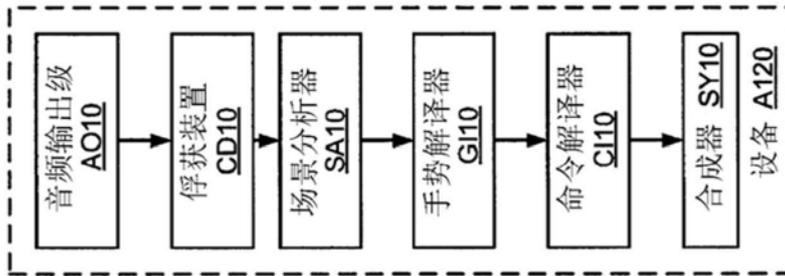


图28C

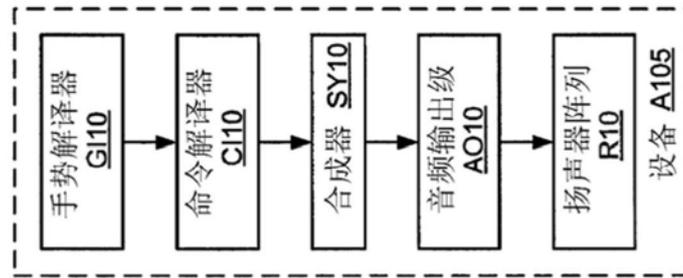


图28D

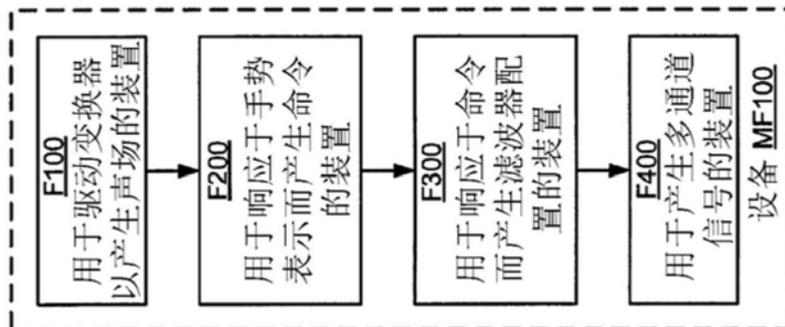


图29A

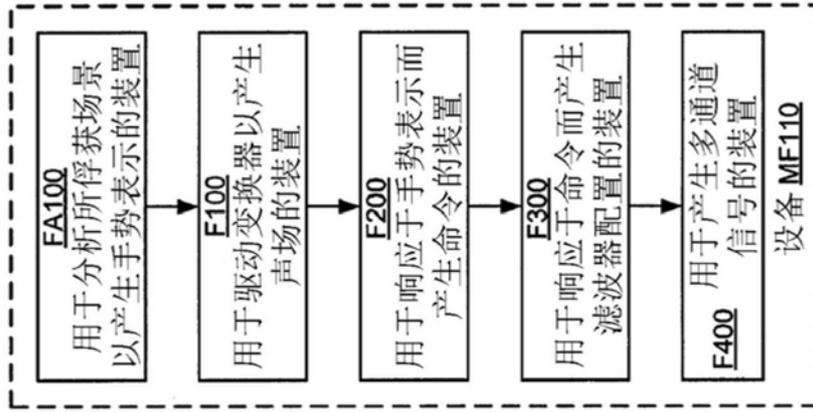


图29B

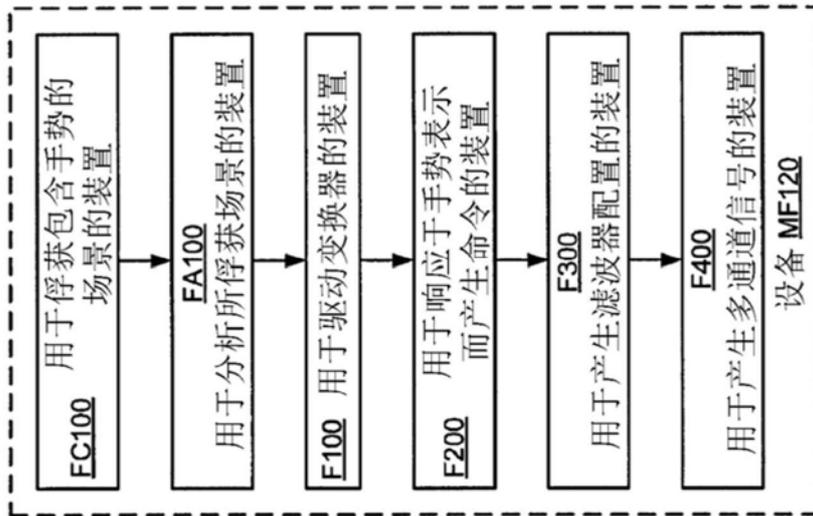


图29C

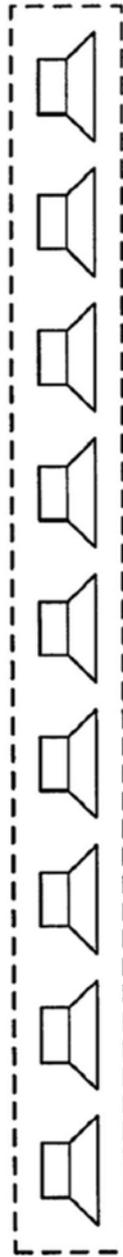


图30A



图30B



图30C

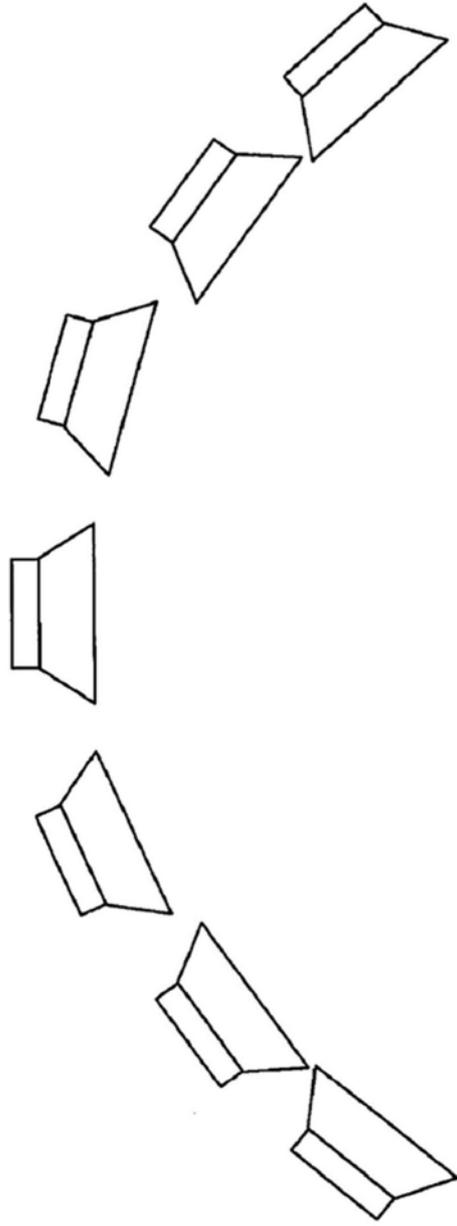


图30D

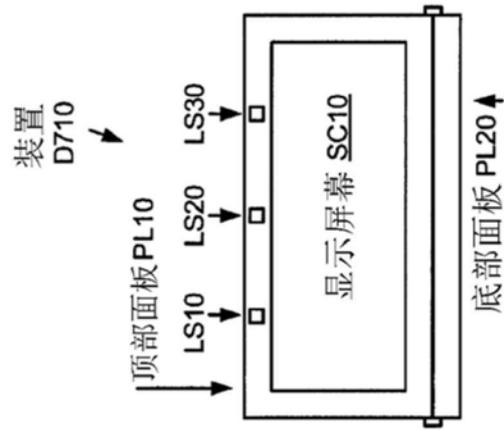


图31A

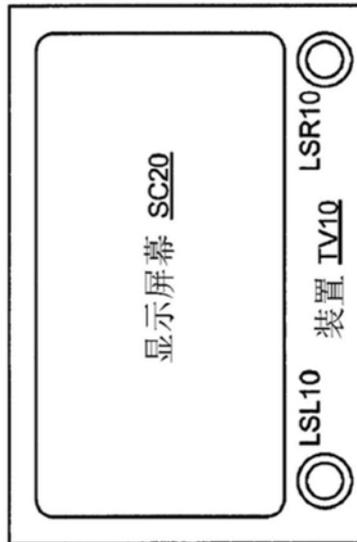


图31B



图31C

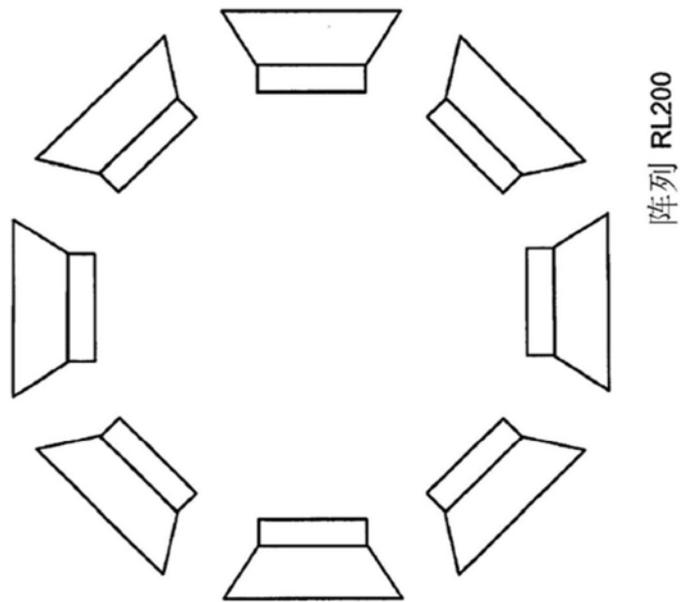


图32A

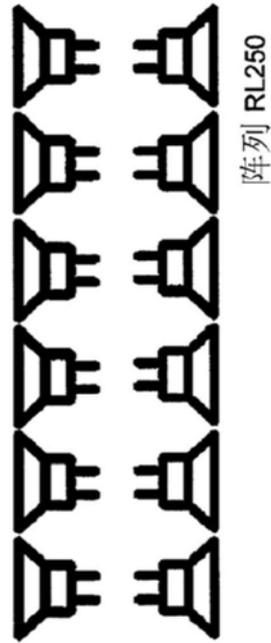


图32B

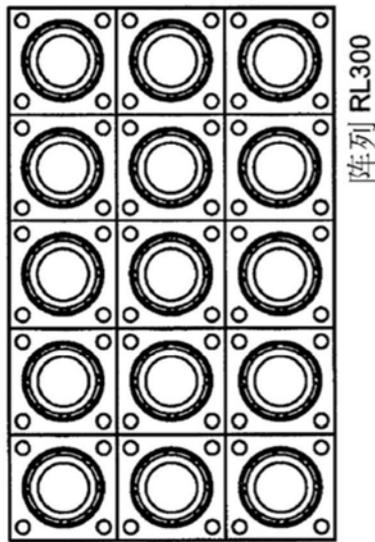


图32C

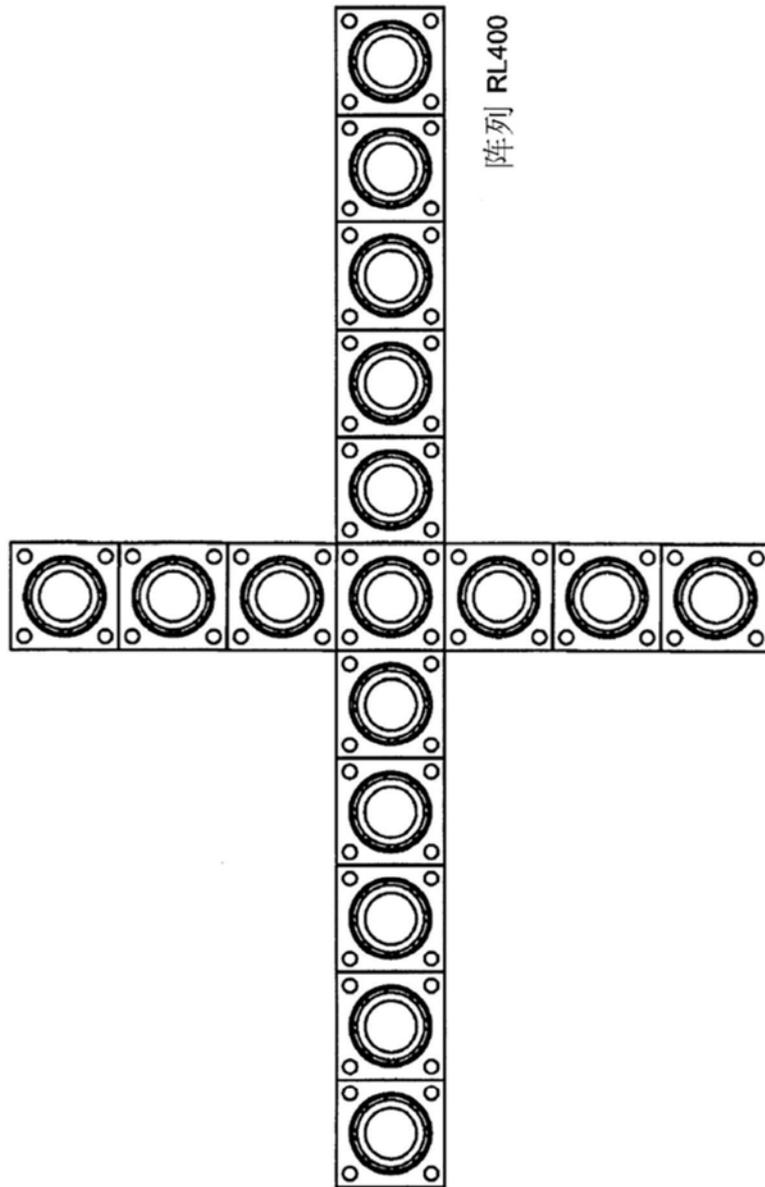


图33

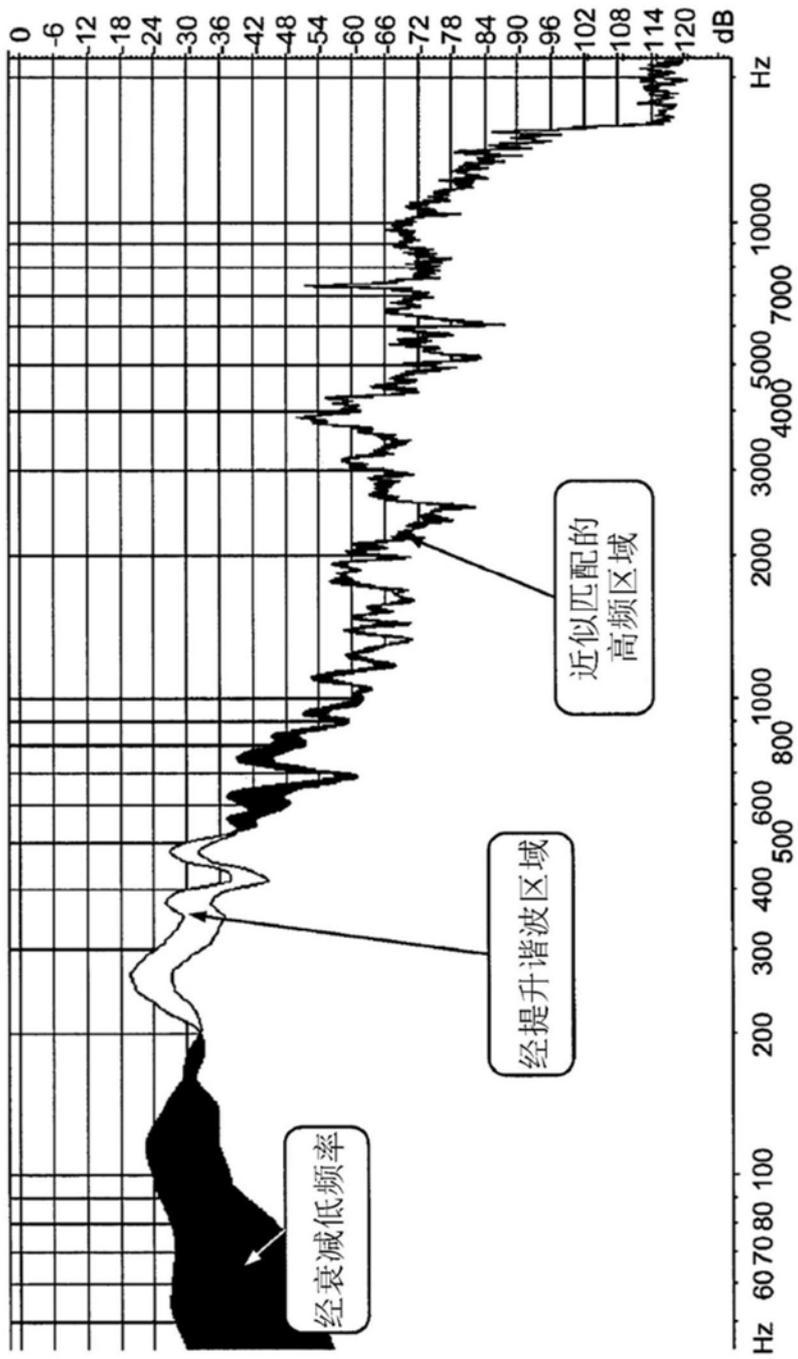


图34