



(12) 发明专利

(10) 授权公告号 CN 111418011 B

(45) 授权公告日 2023.05.12

(21) 申请号 201880076788.0

(22) 申请日 2018.09.28

(65) 同一申请的已公布的文献号
申请公布号 CN 111418011 A

(43) 申请公布日 2020.07.14

(30) 优先权数据
15/718,911 2017.09.28 US

(85) PCT国际申请进入国家阶段日
2020.05.27

(86) PCT国际申请的申请数据
PCT/US2018/053517 2018.09.28

(87) PCT国际申请的公布数据
W02019/067959 EN 2019.04.04

(73) 专利权人 搜诺思公司
地址 美国加利福尼亚州

(72) 发明人 塞义德·巴盖里·塞雷斯基
罗米·卡德里

(74) 专利代理机构 中科专利商标代理有限责任
公司 11021
专利代理师 潘军

(51) Int.Cl.
G10L 21/0208 (2006.01)

(56) 对比文件
CN 105472191 A, 2016.04.06
CN 1748250 A, 2006.03.15
JP 2007235875 A, 2007.09.13
US 2009010445 A1, 2009.01.08
S. Doclo et al.. Combined acoustic
echo and noise reduction using GSVD-based
optimal filtering. 2000 IEEE International
Conference on Acoustics, Speech, and
Signal Processing. Proceedings. 2002, 全文.
(续)

审查员 薛彦杰

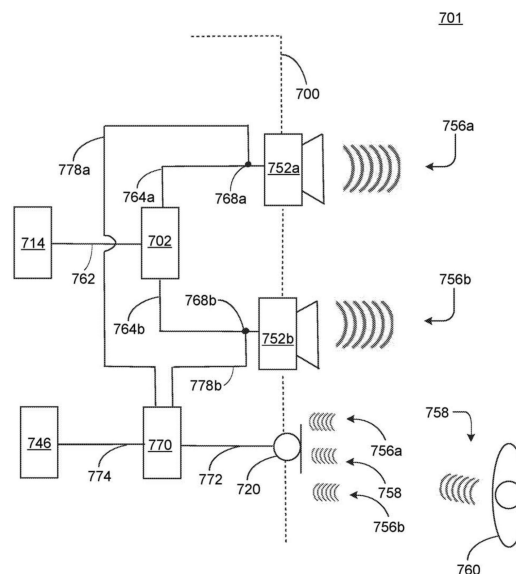
权利要求书2页 说明书31页 附图11页

(54) 发明名称

多声道声学回声消除

(57) 摘要

一种操作回放设备的方法,包括:接收源音频内容,源音频内容包括第一声道音频流和第二声道音频流。方法还包括:通过回放设备的第一扬声器驱动器和第二扬声器驱动器回放第一声道音频流和第二声道音频流,从而产生第一声道音频输出和第二声道音频输出。由回放设备的麦克风接收所捕获的音频流,并且所捕获的音频流的部分与第一声道音频输出和第二声道音频输出相对应。将第一声道音频流和第二声道音频流组合为复合音频信号,并且对复合音频信号执行声学回声消除,以产生声学回声消除输出,然后将声学回声消除输出应用于所捕获的音频流以增加所捕获的音频流的信噪比。



CN 111418011 B

[接上页]

(56) 对比文件

Karim Helwani et al..Source-domain
adaptive filtering for MIMO systems with

application to acoustic echo
cancellation.2010 IEEE International
Conference on Acoustics, Speech and
Signal Processing.2010,全文.

1. 一种操作回放设备的方法,所述回放设备具有第一扬声器驱动器、至少第二扬声器驱动器和一个或多个麦克风,所述方法包括:

通过所述回放设备的网络接口接收源音频流,所述源音频流包括要由所述回放设备回放的源音频内容,其中,所述源音频内容包括第一声道音频流和第二声道音频流;

通过经由第一扬声器驱动器回放第一声道音频流来产生第一声道音频输出;

通过经由第二扬声器驱动器回放第二声道音频流来产生第二声道音频输出;

通过一个或多个麦克风接收所捕获的音频流,所捕获的音频流包括与第一声道音频输出相对应的第一部分和与第二声道音频输出相对应的第二部分,其中所捕获的音频流具有第一信噪比;

对第一声道音频流和第二声道音频流执行奇异值分解,以产生信号分量的组合集;

基于一个或多个参数选择信号分量的组合集的子集;

将信号分量的组合集的子集和所捕获的音频流变换到短时傅立叶变换域;

对信号分量的组合集的子集执行声学回声消除,以产生第一声学回声消除输出;以及

将第一声学回声消除输出应用于所捕获的音频流。

2. 根据权利要求1所述的方法,其中,对第一声道音频流和第二声道音频流执行奇异值分解包括对第一声道音频流和第二声道音频流的并集执行奇异值分解。

3. 根据权利要求1所述的方法,其中,对第一声道音频流和第二声道音频流执行奇异值分解包括对第一声道音频流和第二声道音频流分别执行奇异值分解。

4. 根据权利要求1至3中任一项所述的方法,其中:

所捕获的音频流包括与由用户发出的声命令相对应的第三部分,以及

将第一声学回声消除输出应用于所捕获的音频流将所述第一部分和所述第二部分从所捕获的音频流中最小化或消除。

5. 根据权利要求1至3中任一项所述的方法,其中,所述一个或多个参数包括高于第一阈值能量含量的能量含量。

6. 根据权利要求1至3中任一项所述的方法,其中,所述一个或多个参数包括高于第一阈值方差的所计算的方差。

7. 根据权利要求1至3中任一项所述的方法,

其中,所述回放设备包括第三扬声器驱动器,

其中,所述源音频内容包括第一声道音频流、第二声道音频流和第三声道音频流,

其中,所述方法还包括通过经由第三扬声器驱动器回放第三声道音频流来产生第三声道音频输出。

8. 根据权利要求7所述的方法,其中:

由一个或多个麦克风接收的所记录的音频流还包括与所述第三声道音频输出相对应的第三部分。

9. 根据权利要求7所述的方法,还包括:

对所述第一声道音频流、第二声道音频流和第三声道音频流执行奇异值分解,以产生信号分量的组合集。

10. 根据权利要求6所述的方法,还包括:

当以下之一发生时,执行声学回声消除:

回放功能由回放设备启动,或者
在启动回放功能之后,回放设备接收到取消静音命令。

11.根据权利要求6所述的方法,其中,对第一声道音频流执行奇异值分解是由回放设备的处理器执行的,

其中,对第二声道音频流执行奇异值分解是由回放设备的处理器执行的,

其中,基于一个或多个参数来选择信号分量的组合集的子集是由回放设备的处理器执行的,

其中,对所述信号分量的组合集的子集执行声学回声消除是由回放设备的处理器执行的,以及

其中,将第一声学回声消除输出应用于所捕获的音频流是由回放设备的处理器执行的。

12.一种存储指令的有形非暂时性计算机可读介质,所述指令在由处理器执行时使回放设备执行任何一项前述权利要求。

13.一种回放设备,包括:

多个扬声器驱动器;以及

处理器,被配置为使所述回放设备执行任何前述权利要求所述的方法。

多声道声学回声消除

[0001] 相关申请的交叉引用

[0002] 本申请要求于2017年9月28日递交的题为“Multi-Channel Acoustic Echo Cancellation”的美国专利申请No.15/718,911的优先权,其内容通过引用整体并入本文。

技术领域

[0003] 本公开涉及消费者产品,更具体地,涉及与媒体播放或者其某个方面有关的方法、系统、产品、特征、服务和其他元素。

背景技术

[0004] 访问和收听外放设置的数字音频的选项是有限的,直到2003年SONOS公司申请了其首批专利申请中的一件题为“Method for Synchronizing Audio Playback between Multiple Networked Devices”的专利申请,并在2005年开始出售媒体回放系统为止。Sonos无线高保真(HiFi)系统使人们可以通过一个或多个联网回放设备体验来自许多源的音乐。通过安装在智能电话、平板电脑或计算机上的软件控制应用,人们能够在具有联网回放设备的任何房间中播放他或她期望的内容。另外,使用控制器,例如,能够将不同的歌曲流传输到具有回放设备的每个房间,能够将房间组合在一起进行同步回放,或者可以在所有房间中同步收听相同的歌曲。

[0005] 鉴于对数字媒体的兴趣日益增长,仍然需要开发一种消费者易于使用的技术以进一步增强收听体验。

附图说明

[0006] 参考以下说明书、所附权利要求和附图,可以更好地理解所公开的技术的特征、方面和优点,在附图中:

[0007] 图1示出其中可以实践某些实施例的媒体回放系统配置;

[0008] 图2是根据本文描述的方面的示例回放设备的功能框图;

[0009] 图3是根据本文描述的方面的示例控制设备的功能框图;

[0010] 图4是根据本文描述的方面的示例控制器界面;

[0011] 图5是根据本文描述的方面的多个网络设备的示意图;

[0012] 图6是根据本文描述的方面的网络麦克风设备的功能框图;

[0013] 图7是回放设备可以位于其中的示例环境的示意图;

[0014] 图8是根据本文描述的方面的操作回放设备的方法的流程图;

[0015] 图9是回放设备的音频管线的框图;

[0016] 图10是操作回放设备的方法的流程图;

[0017] 图11是操作回放设备的方法的流程图;以及

[0018] 图12是操作回放设备的方法的流程图。

[0019] 附图用于示出示例实施例的目的,但可以理解,本发明不限于附图所示的布置和

手段。

具体实施方式

[0020] I. 概述

[0021] 一些回放设备被配置为播放音乐和/或其他音频内容,且还具有语音控制能力。这些回放设备可以被配置为通过麦克风从用户接收声命令,分析所接收的声命令,并且翻译声命令以由回放设备执行。例如,如果用户说出声命令以“播放滚石乐队”,则回放设备可以通过麦克风捕获该命令,分析并翻译该命令,并且相应地播放来自滚石乐队的音乐。

[0022] 然而,在一些情况下,用户可能在回放设备已经回放媒体时发出声命令,并且麦克风可以捕获与用户的声命令和回放设备输出的媒体二者相对应的音频。在这种情况下,回放设备输出的媒体可能包含“噪声”,该“噪声”可以使用户对回放设备的声命令模糊不清。在其他情况下,由回放设备回放媒体的麦克风捕获的音频本身可以解释为向回放设备发出的语音命令,并且回放设备可能投入计算资源以确定由麦克风捕获的该伪“命令”的性质。

[0023] 因此,对所捕获的信号执行声学回声消除可以“过滤出”由回放设备回放的媒体的输出,从而增加由麦克风捕获的信号的信噪比。本质上,声学回声消除过程从麦克风捕获的音频信号中移除由回放设备播放的不需要的音频,从而使所捕获的音频上的语音命令更加清晰。

[0024] 然而,使用具有一个以上扬声器的回放设备(例如,多声道回放系统)执行声学回声消除可能具有相关联的缺点。具体地,通常将音频信号流发送给回放设备的两个或更多个扬声器驱动器中的每一个,并且必须通过声学回声消除从所捕获的音频中分别滤除每个音频信号流。该过滤过程通常需要大量的计算资源,这通常超出了典型回放设备的CPU的计算能力。

[0025] 本文提供的示例涉及与前文所述的声学回声消除过程相比,使用更少的计算资源来允许执行多声道声学回声消除的方法、回放设备和系统。在一些实施例中,可以使用回放设备的现有CPU来执行多声道声学回声消除。

[0026] 虽然本文描述的一些示例可以涉及由诸如“用户”和/或其他实体之类的给定行动者执行的功能,但是应该理解,这仅仅出于解释的目的。除非权利要求本身的语言明确要求,否则不应将权利要求解释为要求任何此类示例行动者进行动作。本领域普通技术人员将理解,本公开包括许多其他实施例。

[0027] II. 示例操作环境

[0028] 图1示出媒体回放系统100的示例配置,在媒体回放系统100中可以实践或实施本文公开的一个或多个实施例。如图所示的媒体回放系统100与具有若干房间和空间(例如,主卧室、书房、餐厅和客厅)的示例家居环境相关联。如图1的示例所示,媒体回放系统100包括回放设备102-124、控制设备126和128、以及有线或无线网络路由器130。

[0029] 可以在以下部分中找到关于示例媒体回放系统100的不同组件以及不同组件可以如何交互以向用户提供媒体体验的进一步讨论。虽然本文的讨论可能总体上涉及示例媒体回放系统100,但是本文描述的技术不限于尤其如图1所示的家居环境内的应用。例如,本文描述的技术可以在可能期望多区域音频的环境中有用,例如,诸如餐馆、商场或机场之类的商业环境、诸如运动型多用途车(SUV)、公共汽车或小汽车之类的载运工具、船舶或船只、飞

机等。

[0030] a. 示例回放设备

[0031] 图2示出示例回放设备200的功能框图,示例回放设备200可以被配置为图1的媒体回放系统100的回放设备102-124中的一个或多个。回放设备200可以包括处理器202、软件组件204、存储器206、音频处理组件208、音频放大器210、扬声器212、包括无线接口216和有线接口218的网络接口214、以及麦克风220。在一种情况下,回放设备200可以不包括扬声器212,而是包括用于将回放设备200连接到外部扬声器的扬声器接口。在另一种情况下,回放设备200可以既不包括扬声器212也不包括音频放大器210,而是包括用于将回放设备200连接到外部音频放大器或视听接收器的音频接口。

[0032] 在一个示例中,处理器202是时钟驱动计算组件,其被配置为根据存储器206中存储的指令处理输入数据。存储器206可以是有形计算机可读介质,其被配置为存储可由处理器202执行的指令。例如,存储器206可以是数据存储设备,其可以加载有可由处理器202执行以实现某些功能的一个或多个软件组件204。在一个示例中,所述功能可以包括回放设备200从音频源或另一回放设备获取音频数据。在另一示例中,所述功能可以包括回放设备200向网络上的另一设备或回放设备发送音频数据。在又一示例中,所述功能可以包括将回放设备200与一个或多个回放设备配对以创建多声道音频环境。

[0033] 某些功能可以包括回放设备200与一个或多个其他回放设备同步对音频内容的回放。在同步回放期间,收听者将优选地不能够感知到回放设备200和一个或多个其他回放设备对音频内容的回放之间的时间延迟差异。通过引用整体并入本文的题为“System and method for synchronizing operations among a plurality of independently clocked digital data processing devices”的美国专利No.8,234,395提供了回放设备之间的音频回放同步的详细示例。

[0034] 存储器206还可以被配置为存储与回放设备200相关联的数据,例如,回放设备200是其一部分的一个或多个区域和/或区域组、回放设备200可访问的音频源、或回放设备200(或某个其他回放设备)可以与之相关联的回放队列。数据可以被存储为一个或多个状态变量,所述状态变量被周期性地更新并用于描述回放设备200的状态。存储器206还可以包括与媒体系统的其他设备的状态相关联的数据,并且间或地在设备之间共享,使得设备中的一个或多个具有与系统相关联的最新数据。其他实施例也是可能的。

[0035] 音频处理组件208可以包括一个或多个数模转换器(DAC)、音频预处理组件、音频增强组件或数字信号处理器(DSP)等。在一个实施例中,音频处理组件208中的一个或多个可以是处理器202的子组件。在一个示例中,音频处理组件208可以处理和/或有意地改变音频内容以产生音频信号。然后,所产生的音频信号可被提供至音频放大器210进行放大,并通过扬声器212回放。具体地,音频放大器210可以包括被配置为将音频信号放大到用于驱动扬声器212中的一个或多个的电平的设备。扬声器212可以包括单独的换能器(例如,“驱动器”)或包括具有一个或多个驱动器的壳体的完整的扬声器系统。例如,扬声器212的特定驱动器可以包括例如低音喇叭(例如,针对非常低的频率)、中频段驱动器(例如,针对中频)和/或高频扬声器(例如,针对高频)。在一些情况下,一个或多个扬声器212中的每个换能器可以由音频放大器210的各个对应的音频放大器来驱动。除了产生用于由回放设备200回放的模拟信号之外,音频处理组件208还可以被配置为处理要向一个或多个其他回放设备发

送以进行回放的音频内容。

[0036] 可以例如通过音频线路输入的输入连接(例如,自动检测3.5mm音频线路输入连接)或网络接口214从外部源接收要由回放设备200处理和/或回放的音频内容。

[0037] 网络接口214可以被配置为促进回放设备200与数据网络上的一个或多个其他设备之间的数据流。这样,回放设备200可以被配置为通过数据网络从与回放设备200通信的一个或多个其他回放设备、局域网内的网络设备、或通过诸如互联网之类的广域网从音频内容源接收音频内容。在一个示例中,回放设备200发送和接收的音频内容和其他信号可以以包含基于互联网协议(IP)的源地址和基于IP的目的地址的数字数据包数据的形式来发送。在这种情况下,网络接口214可以被配置为解析数字数据包数据,使得去往回放设备200的数据被回放设备200正确地接收和处理。

[0038] 如图所示,网络接口214可以包括无线接口216和有线接口218。无线接口216可以为回放设备200提供网络接口功能,以根据通信协议(例如,任何无线标准,包括IEEE 802.11a、802.11b、802.11g、802.11n、802.11ac、802.15、4G移动通信标准等)与其他设备(例如,与回放设备200相关联的数据网络内的其他回放设备、扬声器、接收器、网络设备、控制设备)无线通信。有线接口218可以为回放设备200提供网络接口功能,以根据通信协议(例如,IEEE 802.3)通过有线连接与其他设备通信。虽然图2中所示的网络接口214包括无线接口216和有线接口218,但是在一些实施例,网络接口214可以仅包括无线接口或仅包括有线接口。

[0039] 麦克风220可以被布置为检测回放设备200的环境中的声音。例如,麦克风可以安装在回放设备的外壳的外壁上。麦克风可以是现在已知或以后开发的任何类型的麦克风,例如,电容式麦克风、驻极体电容式麦克风或动态麦克风。麦克风可以对扬声器212的频段的一部分敏感。扬声器212中的一个或多个可以与麦克风220相反地操作。在一些方面,回放设备200可以没有麦克风220。

[0040] 在一个示例中,回放设备200和另一回放设备可以配对,以播放音频内容的两个单独的音频分量。例如,回放设备200可以被配置为播放左声道音频分量,而另一回放设备可以被配置为播放右声道音频分量,从而产生或增强音频内容的立体声效果。配对的回放设备(也称为“绑定的回放设备”)还可以与其他回放设备同步播放音频内容。

[0041] 在另一示例中,回放设备200可以与一个或多个其他回放设备声音合并以形成单个合并的回放设备。合并的回放设备可以被配置为与非合并的回放设备或配对的回放设备不同地处理和再现声音,这是因为合并的回放设备可以具有可以通过其呈现音频内容的附加的扬声器驱动器。例如,如果回放设备200是被设计为呈现低频段音频内容的回放设备(即,低音喇叭),则回放设备200可以与被设计为呈现全频段音频内容的回放设备合并。在这种情况下,当与低频回放设备200合并时,全频段回放设备可以被配置为仅呈现音频内容的中高频分量,而低频段回放设备200则呈现音频内容的低频分量。合并的回放设备还可以与单个回放设备或另一合并的回放设备配对。

[0042] 举例来说,SONOS公司目前提供(或已经提供)销售某些回放设备,包括“PLAY:1”、“PLAY:3”、“PLAY:5”、“PLAYBAR”、“CONNECT:AMP”、“CONNECT”、和“SUB”。任何其他过去、现在和/或将来的回放设备可以附加地或备选地用于实现本文公开的示例实施例的回放设备。此外,应当理解,回放设备不限于图2示出的示例或SONOS的产品供应。例如,回放设备可以

包括有线或无线耳机。在另一示例中,回放设备可以包括个人移动媒体回放设备的扩展基座,或与其交互。在又一示例中,回放设备可以是诸如电视、照明器材或在室内外使用的一些其他设备之类的另一设备或组件的组成部分。

[0043] b. 示例回放区域配置

[0044] 再次参考图1的媒体回放系统100,环境可以具有一个或多个回放区域,每个回放区域具有一个或多个回放设备。媒体回放系统100可以建立有一个或多个回放区域,之后可以添加或移除一个或多个区域,以达到图1所示的示例配置。每个区域可以根据不同的房间或空间(例如书房、浴室、主卧室、卧室、厨房、餐厅、客厅和/或阳台)被赋予名称。在一种情况下,单个回放区域可以包括多个房间或空间。在另一种情况下,单个房间或空间可以包括多个回放区域。

[0045] 如图1所示,阳台、餐厅、厨房、浴室、书房和卧室区域均具有一个回放设备,而客厅和主卧室区域均具有多个回放设备。在客厅区域中,回放设备104、106、108和110可以被配置为作为单独的回放设备、作为一个或多个绑定的回放设备、作为一个或多个合并的回放设备或其任何组合来同步播放音频内容。类似地,在主卧室的情况下,回放设备122和124可以被配置为作为单独的回放设备、作为绑定的回放设备、或作为合并的回放设备来同步播放音频内容。

[0046] 在一个示例中,图1的环境中的一个或多个回放区域可以分别播放不同的音频内容。例如,用户可以在阳台区域中烧烤并收听正由回放设备102播放的嘻哈音乐,而另一用户可以正在厨房区域中准备食物并收听正由回放设备114播放的古典音乐。在另一示例中,回放区域可以与另一回放区域同步地播放相同的音频内容。例如,用户可以在书房区域中,其中回放设备118正在播放与阳台区域中的回放设备102正在播放的摇滚音乐相同的摇滚音乐。在这种情况下,回放设备102和118可以同步播放摇滚音乐,使得用户可以在不同回放区域之间移动时无缝地(或至少基本上无缝地)享受被外放播放的音频内容。可以以类似于如先前引用的美国专利No.8,234,395中所述的回放设备之间的同步的方式来实现回放区域之间的同步。

[0047] 如上文所建议的,可以动态地修改媒体回放系统100的区域配置,并且在一些实施例中,媒体回放系统100支持多种配置。例如,如果用户将一个或多个回放设备物理地移动到区域中或从区域中移出,则可以重新配置媒体回放系统100以适应改变。例如,如果用户将回放设备102从阳台区域物理地移动到书房区域,则书房区域现在可以包括回放设备118和回放设备102。可以通过诸如控制设备126和128之类的控制设备,将回放设备102与书房区域配对或分组,和/或重新命名(如果需要)。另一方面,如果一个或多个回放设备被移动到家居环境中还不是回放区域的特定区域,则可以针对该特定区域创建新的回放区域。

[0048] 此外,媒体回放系统100的不同回放区域可以被动态组合为区域组或分成单独的回放区域。例如,餐厅区域和厨房区域可以组合为用于宴会的区域组,使得回放设备112和114可以同步呈现音频内容。另一方面,如果用户期望在客厅空间中听音乐,而另一用户期望看电视,则客厅区域可以被分成包括回放设备104的电视区域和包括回放设备106、108和110的收听区域。

[0049] c. 示例控制设备

[0050] 图3示出示例控制设备300的功能框图,示例控制设备300可以被配置为媒体回放

系统100的控制设备126和128之一或二者。如图所示,控制设备300可以包括处理器302、存储器304、网络接口306、用户界面308、麦克风310和软件组件312。在一个示例中,控制设备300可以是媒体回放系统100的专用控制器。在另一示例中,控制设备300可以是可以在其上安装媒体回放系统控制器应用软件的网络设备,例如,iPhone™、iPad™或任何其他智能电话、平板计算机或网络设备(例如,联网计算机,如PC或Mac™)。

[0051] 处理器302可以被配置为执行与有助于用户对媒体播放系统100进行访问、控制和配置相关的功能。存储器304可以是数据存储设备,其可以加载有可由处理器302执行以实现这些功能的一个或多个软件组件。存储器304还可以被配置为存储媒体回放系统控制器应用软件和与媒体回放系统100和用户相关联的其他数据。

[0052] 在一个示例中,网络接口306可以基于行业标准(例如,红外;无线电;包括IEEE 802.3的有线标准;包括IEEE 802.11a、802.11b、802.11g、802.11n、802.11ac、802.15、4G移动通信标准的无线标准等)。网络接口306可以提供用于控制设备300与媒体回放系统100中的其他设备通信的手段。在一个示例中,可以通过网络接口306在控制设备300和其他设备之间传送数据和信息(例如,状态变量)。例如,媒体回放系统100中的回放区域和区域组配置可以由控制设备300通过网络接口306从回放设备或另一网络设备接收,或者由控制设备300通过网络接口306向另一回放设备或网络设备发送。在一些情况下,该另一网络设备可以是另一控制设备。

[0053] 诸如音量控制和音频回放控制之类的回放设备控制命令也可以通过网络接口306从控制设备300向回放设备传送。如上文所建议的,还可以由用户使用控制设备300来执行对媒体回放系统100的配置的改变。配置改变可以包括:向区域增加或从区域中移除一个或多个回放设备、向区域组增加或从区域组中移除一个或多个区域、形成绑定或合并的播放器、将一个或多个回放设备与绑定或合并的播放器中分离等。因此,控制设备300有时可以被称为控制器,无论控制设备300是专用控制器还是在其上安装有媒体回放系统控制器应用软件的网络设备。

[0054] 控制设备300可以包括麦克风310。麦克风310可以被布置为检测控制设备300的环境中的声音。麦克风310可以是现在已知或以后开发的任何类型的麦克风,例如,电容式麦克风、驻极体电容式麦克风或动态麦克风。麦克风可以对频段的一部分敏感。两个或更多个麦克风310可以被布置为捕获音频源(例如,语音、可听声音)的位置信息和/或帮助滤除背景噪声。

[0055] 控制设备300的用户界面308可以被配置为通过提供诸如图4中所示的控制器界面400之类的控制器界面来帮助用户访问和控制媒体回放系统100。控制器界面400包括回放控制区410、回放区域区420、回放状态区430、回放队列区440和音频内容源区450。所示的控制器界面400仅是可以在诸如图3的控制设备300(和/或图1的控制设备126和128)之类的网络设备上提供、并由用户访问以控制媒体回放系统(例如,媒体回放系统100)的用户界面的一个示例。备选地,可以在一个或多个网络设备上实现变化的格式、样式和交互序列的其他用户界面,以提供对媒体回放系统的类似的控制访问。

[0056] 回放控制区410可以包括可选择(例如,通过触摸或通过使用光标)图标,以使所选择的回放区域或区域组中的回放设备播放或暂停、快进、回退、跳到下一个、跳到前一个、进入/退出随机播放模式、进入/退出重复模式、进入/退出交叉淡入淡出模式(cross fade

mode)。回放控制区410还可以包括用于修改均衡设置、回放音量等的可选择图标。

[0057] 回放区域区420可以包括媒体回放系统100内的回放区域的表示。在一些实施例中,回放区域的图形表示可以是可选择的,以调出附加的可选择图标来管理或配置媒体回放系统中的回放区域,例如,创建绑定的区域、创建区域组、分离区域组、重命名区域组等。

[0058] 例如,如图所示,可以在回放区域的每个图形表示内提供“分组”图标。在特定区域的图形表示内提供的“分组”图标可以是可选择的,以调出用于选择媒体回放系统中要与特定区域分在一组的一个或多个其他区域的选项。一旦被分组,已经与特定区域分在一组的区域中的回放设备将被配置为与特定区域中的回放设备同步地播放音频内容。类似地,可以在区域组的图形表示内提供“分组”图标。在这种情况下,“分组”图标可以是可选择的,以调出用于不选择要从区域组中移除的该区域组中的一个或多个区域的选项。通过诸如控制器界面400之类的用户界面对区域进行分组和取消分组的其他交互和实现也是可能的。当回放区域或区域组配置被修改时,可以动态地更新回放区域在回放区域区420中的表示。

[0059] 回放状态区430可以包括在所选择的回放区域或区域组中当前正在播放、先前播放或安排为接下来播放的音频内容的图形表示。可以在用户界面上可视地区分所选择的回放区域或区域组,例如,在回放区域区420和/或回放状态区430内。图形表示可以包括曲目名称、艺术家姓名、专辑名称、专辑年份、曲目长度以及当通过控制器界面400控制媒体回放系统时用户知道了会有用的其他相关信息。

[0060] 回放队列区440可以包括与所选择的回放区域或区域组相关联的回放队列中的音频内容的图形表示。在一些实施例中,每个回放区域或区域组可以与回放队列相关联,该回放队列包含与由该回放区域或区域组回放的零个或多个音频项相对应的信息。例如,回放队列中的每个音频项可以包括统一资源标识符(URI)、统一资源定位符(URL)或一些其他标识符,其可以由回放区域或区域组中的回放设备用于从本地音频内容源或联网音频内容源查找和/或获取音频项,可能供回放设备回放。

[0061] 在一个示例中,可以将播放列表添加到回放队列,在这种情况下,可以将与播放列表中的每个音频项对应的信息添加到回放队列。在另一示例中,回放队列中的音频项可以被保存为播放列表。在另外的示例中,当回放区域或区域组正在持续播放流媒体音频内容(例如,互联网收音机,其可以持续播放直到被停止),而不是具有回放持续时间的分立音频项时,回放队列可以是空的或被填充但是“未使用”。在备选实施例中,回放队列可以包括互联网收音机和/或其他流媒体音频内容项,并且当回放区域或区域组正在播放这些项时处于“使用中”。其他示例也是可能的。

[0062] 当回放区域或区域组被“分组”或“取消分组”时,可以清除与受影响的回放区域或区域组相关联的回放队列,或者重新关联。例如,如果包括第一回放队列的第一回放区域与包括第二回放队列的第二回放区域被分在一组,则所建立的区域组可以具有相关联的回放队列,其最初是空的,包含来自第一回放队列的音频项(例如,如果第二回放区域被添加到第一回放区域),或包含来自第二回放队列的音频项(例如,如果第一回放区域被添加到第二回放区域),或包含来自第一回放队列和第二回放队列二者的音频项的组合。随后,如果所建立的区域组被取消分组,则所得到的第一回放区域可以与先前的第一回放队列重新关联,或者与新的回放队列相关联,该新的回放队列是空的,或者包含与来自在所建立的区域组被取消分组之前所建立的区域组相关联的回放队列的音频项。类似地,所得到的第二回

放区域可以与先前的第二回放队列重新关联,或者与新的回放队列相关联,该新的回放队列是空的,或者包含来自在与所建立的区域组被取消分组之前所建立的区域组相关联的回放队列的音频项。其他示例也是可能的。

[0063] 返回参考图4的控制器界面400,音频内容在回放队列区440中的图形表示可以包括曲目标题、艺术家姓名、曲目长度以及与回放队列中的音频内容相关联的其他相关信息。在一个示例中,音频内容的图形表示可以是可选择的,以调出附加的可选择图标来管理和/或操纵回放队列和/或回放队列中表示的音频内容。例如,可以将所表示的音频内容从回放队列中移除,将所表示的音频内容移动到回放队列内的不同位置,或者选择所表示的音频内容以立即播放,或者在任何当前播放的音频内容之后进行播放等。与回放区域或区域组相关联的回放队列可以存储于该回放区域或区域组中的一个或多个回放设备上、不在该回放区域或区域组中的回放设备上和/或一些其他指定设备上的存储器中。

[0064] 音频内容源区450可以包括可选择的音频内容源的图形表示,可以从音频内容源中获取音频内容,并由所选择的回放区域或区域组来播放。有关音频内容源的讨论可参见以下部分。

[0065] d. 示例音频内容源

[0066] 如前所述,区域或区域组中的一个或多个回放设备可以被配置为从各种可用音频内容源中获取回放音频内容(例如,根据音频内容的对应URI或URL)。在一个示例中,回放设备可以直接从对应的音频内容源(例如,线路输入连接)中获取音频内容。在另一示例中,可以在网络上,通过一个或多个其他回放设备或网络设备向回放设备提供音频内容。

[0067] 示例音频内容源可以包括:媒体回放系统(例如,图1的媒体回放系统100)中的一个或多个回放设备的存储器、一个或多个网络设备(例如,控制设备、支持网络的个人计算机、或者网络附接存储器(NAS)等)上的本地音乐库、通过互联网(例如,云)提供音频内容的流媒体音频服务、或者通过回放设备或网络设备上的线路输入连接连接至媒体回放系统的音频源等。

[0068] 在一些实施例中,可以在诸如图1的媒体回放系统100之类的媒体回放系统中定期添加音频内容源,或从中移除音频内容源。在一个示例中,每当添加、移除或更新一个或多个音频内容源时,可以执行对音频项编索引。对音频项编索引可以包括:扫描由媒体回放系统中的回放设备可访问的网络上共享的所有文件夹/目录中的可识别音频项,并且生成或更新包含元数据(例如,标题、艺术家、专辑、曲目长度等)及其他关联信息(例如,找到的每个可识别音频项的URI或URL)的音频内容数据库。用于管理和维护音频内容源的其他示例也是可能的。

[0069] 以上关于回放设备、控制器设备、回放区域配置和媒体内容源的讨论仅提供了操作环境的一些示例,在该操作环境中可以实现下面描述的功能和方法。本文未明确描述的媒体回放系统、回放设备和网络设备的配置和其他操作环境也可以适用且适于功能和方法的实现。

[0070] e. 示例多个联网设备

[0071] 图5示出了示例多个设备500,其可以被配置为基于语音控制提供音频回放体验。本领域普通技术人员将理解,图5中所示的设备仅用于说明目的,并且包括不同和/或附加设备的变型是可能的。如图所示,多个设备500包括计算设备504、506和508;网络麦克风设

备(NMD) 512、514和516;回放设备(PBD) 532、534、536和538;以及控制器设备(CR) 522。

[0072] 多个设备500中的每个设备可以是具有网络功能的设备,其可以根据一个或多个网络协议(例如,NFC、蓝牙、以太网和IEEE 802.11等),通过一种或多种类型的网络(例如,广域网(WAN)、局域网(LAN)和个域网(PAN)等)与多个设备中的一个或多个其他设备建立通信。

[0073] 如图所示,计算设备504、506和508可以是云网络502的一部分。云网络502可以包括附加的计算设备。在一个示例中,计算设备504、506和508可以是不同的服务器。在另一示例中,计算设备504、506和508中的两个或更多个可以是单个服务器的模块。类似地,计算设备504、506和508中的每一个可以包括一个或多个模块或服务器。本文中为了便于说明,计算设备504、506和508中的每一个可以被配置为在云网络502内执行特定功能。例如,计算设备508可以是用于流媒体音乐服务的音频内容源。

[0074] 如图所示,计算设备504可以被配置为通过通信路径542与NMD 512、514和516接口连接。NMD 512、514和516可以是一个或多个“智能家居”系统的组件。在一种情况下,NMD 512、514和516可以物理地分布在整个家中,类似于图1所示的设备分布。在另一种情况下,NMD 512、514和516中的两个或更多个可以物理地位置彼此相对靠近。通信路径542可以包括一种或多种类型的网络,例如包括互联网的WAN、LAN和/或PAN等。

[0075] 在一个示例中,NMD 512、514和516中的一个或多个可以是配置为主要用于音频检测的设备。在另一示例中,NMD 512、514和516中的一个或多个可以是具有各种主要实用程序的设备的组件。例如,如上面结合图2和图3所讨论的,NMD 512、514和516中的一个或多个可以是回放设备200的麦克风220或控制设备300的麦克风310。此外,在一些情况下,NMD 512、514和516中的一个或多个可以是回放设备200或控制设备300。在示例中,NMD 512、514和/或516中的一个或多个可以包括布置在麦克风阵列中的多个麦克风。

[0076] 如图所示,计算设备506可以被配置为通过通信路径544与CR 522和PBD 532、534、536和538接口连接。在一个示例中,CR 522可以是网络设备,例如图2的回放设备200。因此,CR 522可以被配置为提供图4的控制器界面400。类似地,PBD 532、534、536和538可以是回放设备,例如图3的控制设备300。这样,PBD 532、534、536和538可以物理地分布在整个家中,如图1所示。为了说明目的,PBD 536和538可以是绑定区域530的一部分,而PBD 532和534可以是它们各自区域的一部分。如上所述,PBD 532、534、536和538可以被动态地绑定、分组、解除绑定和取消分组。通信路径544可以包括一种或多种类型的网络,例如包括互联网的WAN、LAN和/或PAN等。

[0077] 在一个示例中,与NMD 512、514和516一样,CR 522和PBD 532、534、536和538也可以是一个或多个“智能家居”系统的组件。在一种情况下,PBD 532、534、536和538与NMD 512、514和516分布在相同的家中。此外,如上文建议的,PBD 532、534、536和538中的一个或多个可以是NMD 512、514和516中的一个或多个。

[0078] NMD 512、514和516可以是局域网的一部分,并且通信路径542可以包括通过WAN(通信路径,未示出)将NMD 512、514和516的局域网链接到计算设备504的接入点。同样地,NMD 512、514和516中的每一个可以通过该接入点彼此通信。

[0079] 类似地,CR 522和PBD 532、534、536和538可以是局域网和/或本地回放网络的一部分(如前面部分中讨论的),并且通信路径544可以包括通过WAN将CR 522和PBD 532、534、

536和538的局域网和/或本地回放网络链接到计算设备506的接入点。这样,CR 522和PBD 532、534、536和538中的每一个也可以通过该接入点彼此通信。

[0080] 在一个示例中,通信路径542和544可以包括相同的接入点。在示例中,NMD 512、514和516、CR 522和PBD 532、534、536和538中的每一个可以通过家庭的相同接入点来访问云网络502。

[0081] 如图5所示,NMD 512、514和516、CR 522和PBD 532、534、536和538中的每一个还可以通过通信方式546与一个或多个其他设备直接通信。如本文所述的通信方式546可以包括根据一个或多个网络协议通过一种或多种类型的网络在设备之间的一种或多种形式的通信,和/或可以包括通过一个或多个其他网络设备的通信。例如,通信方式546可以包括蓝牙™(IEEE 802.15)、NFC、无线直连和/或专有无线等中的一个或多个。

[0082] 在一个示例中,CR 522可以通过蓝牙™与NMD 512通信,并且可以通过另一局域网与PBD 534通信。在另一示例中,NMD 514可以通过另一局域网与CR 522通信,并且可以通过蓝牙与PBD 536通信。在又一示例中,PBD 532、534、536和538中的每一个可以根据生成树协议通过本地回放网络彼此通信,同时分别通过不同于本地回放网络的局域网与CR 522通信。其他示例也是可能的。

[0083] 在一些情况下,NMD 512、514和516、CR 522和PBD 532、534、536和538之间的通信方式可以根据设备间的通信类型、网络状况和/或时延要求而改变。例如,当NMD 516首次被引入具有PBD 532、534、536和538的家中时,可以使用通信方式546。在一种情况下,NMD 516可以通过NFC向PBD 538发送与NMD 516相对应的标识信息,并且作为响应,PBD 538可以通过NFC(或某种其他形式的通信)向NMD 516发送局域网信息。然而,一旦在家中配置了NMD 516,NMD 516和PBD 538之间的通信方式可能改变。例如,NMD 516可以随后通过通信路径542、云网络502和通信路径544与PBD 538通信。在另一示例中,NMD和PBD可能从不通过本地通信方式546通信。在另一示例中,NMD和PBD可以主要通过本地通信方式546通信。其他示例也是可能的。

[0084] 在说明性示例中,NMD 512、514和516可以被配置为接收用于控制PBD 532、534、536和538的语音输入。可用的控制命令可以包括先前讨论的任何媒体回放系统控制,例如回放音量控制、回放传输控制、音乐源选择和分组等。在一个实例中,NMD 512可以接收用于控制PBD 532、534、536和538中的一个或多个的语音输入。响应于接收到语音输入,NMD 512可以通过通信路径542向计算设备504发送语音输入以进行处理。在一个示例中,计算设备504可以将语音输入转换为等效的文本命令,并解析该文本命令以识别命令。然后,计算设备504可以随后向计算设备506发送文本命令。在另一示例中,计算设备504可以将语音输入转换为等效的文本命令,然后向计算设备506发送文本命令。然后,计算设备506可以解析文本命令以识别一个或多个回放命令。

[0085] 例如,如果文本命令是“在区域1中播放来自流媒体服务1的艺术家1的曲目1”,则计算设备506可以标识(i)从流媒体服务1可获得的艺术家1的曲目1的URL,以及(ii)区域1中的至少一个回放设备。在该示例中,来自流媒体服务1的艺术家1的曲目1的URL可以是指向计算设备508的URL,并且区域1可以是绑定区域530。这样,当识别出URL以及PBD 536和538之一或两者时,计算设备506可以通过通信路径544向PBD 536和538之一或两者发送所识别的URL,用于回放。作为响应,PBD 536和538之一或两者可以根据接收到的URL从计算设

备508中获取音频内容,并且开始播放来自流媒体服务1的艺术家1的曲目1。

[0086] 在又一示例中,计算设备504可以执行一些处理以识别用户的意图或相关命令,并且向计算设备506提供与语音输入相关的媒体内容的相关信息。例如,计算设备504可以对语音输入执行话语到文本转换,并分析语音输入,以得到命令或意图(例如,播放、暂停、停止、音量增大、音量减小、跳过、下一个、分组、取消分组)以及关于如何执行命令的其他信息。计算设备504或计算设备506可以确定哪些PBD命令对应于由计算设备504确定的命令或意图。可以从计算设备504向计算设备506发送从语音输入确定的命令或意图和/或与执行命令相关的其他信息。计算设备504上的处理可以由应用、模块、附加软件、具有本地联网麦克风系统软件平台的集成件、和/或本地联网麦克风系统软件平台来执行。

[0087] 本领域普通技术人员将了解,以上仅是一个说明性示例,并且其他实现也是可能的。在一种情况下,如上所述,多个设备500中的一个或多个执行的操作可以由多个设备500中的一个或多个其他设备来执行。例如,从语音输入到文本命令的转换可以替代地、部分地或完全由另一个或多个设备来执行,例如NMD 512、计算设备506、PBD 536和/或PBD 538。类似地,URL的标识可以替代地、部分地或完全由另一个或多个设备执行,例如,NMD 512、计算设备504、PBD 536和/或PBD 538。

[0088] f. 示例网络麦克风设备

[0089] 图6示出示例网络麦克风设备600的功能框图,示例网络麦克风设备600可以被配置为图5的NMD 512、514和516中的一个或多个。如图所示,网络麦克风设备600包括处理器602、存储器604、麦克风阵列606、网络接口608、用户界面610、软件组件612和扬声器614。本领域普通技术人员将了解,其他网络麦克风设备配置和布置也是可能的。例如,备选地,网络麦克风设备可以不包括扬声器614,或者具有单个麦克风而不是麦克风阵列606。

[0090] 处理器602可以包括一个或多个处理器和/或控制器,其可以采用通用或专用处理器或控制器的形式。例如,处理器602可以包括微处理器、微控制器、专用集成电路、数字信号处理器等。存储器604可以是数据存储设备,其可以加载有可由处理器602执行以实现这些功能的一个或多个软件组件。因此,存储器604可以包括一个或多个非暂时性计算机可读存储介质,其示例可以包括:易失性存储介质(例如,随机存取存储器、寄存器、高速缓存等)、以及非易失性存储介质(例如,只读存储器、硬盘驱动器、固态驱动器、闪存和/或光存储设备等)。

[0091] 麦克风阵列606可以是多个麦克风,其被布置为检测网络麦克风设备600的环境中的声音。麦克风阵列606可以包括现在已知或以后开发的任何类型的麦克风,例如,电容式麦克风、驻极体电容式麦克风或动态麦克风等。在一个示例中,麦克风阵列可以被布置为检测来自相对于网络麦克风设备的一个或多个方向的音频。麦克风阵列606可以对频段的一部分敏感。在一个示例中,麦克风阵列606的第一子集可以对第一频段敏感,而麦克风阵列的第二子集可以对第二频段敏感。麦克风阵列606还可以被布置为捕获音频源(例如,语音、可听声音)的位置信息和/或帮助过滤背景噪声。值得注意的是,在一些实施例中,麦克风阵列可以仅由单个麦克风组成,而不是由多个麦克风组成。

[0092] 网络接口608可以被配置为促进各种网络设备(例如,参考图5,其中的CR 522、PBD 532-538、云网络502中的计算设备504-508,以及其他网络麦克风设备等)之间的无线和/或有线通信。这样,网络接口608可以采用任何合适的形式来执行这些功能,其示例可以包括:

以太网接口、串行总线接口(例如,FireWire、USB 2.0)、适于促进无线通信的芯片组和天线、和/或提供有线和/或无线通信的任何其他接口。在一个示例中,网络接口608可以基于行业标准(例如,红外;无线电;包括IEEE 802.3的有线标准;包括IEEE 802.11a、802.11b、802.11g、802.11n、802.11ac、802.15、4G移动通信标准的无线标准等)。

[0093] 网络麦克风设备600的用户界面610可以被配置为促进与网络麦克风设备的用户交互。在一个示例中,用户界面610可以包括物理按钮、设置在触敏屏幕和/或表面上的图形界面等中的一个或多个,以供用户直接向网络麦克风设备600提供输入。用户界面610还可以包括灯和扬声器614中的一个或多个,以向用户提供视觉和/或音频反馈。在一个示例中,网络麦克风设备600还可以被配置为通过扬声器614回放音频内容。

[0094] III. 示例系统

[0095] 本文描述的实施例涉及对由回放设备的至少一个麦克风捕获的音频流执行多声道声学回声消除,该回放设备能够通过第一扬声器驱动器和第二扬声器驱动器回放音频内容。图7示出了根据所公开的实施例的这样的回放设备可以位于其中的示例环境701。例如,示例环境701可以是家庭的客厅或卧室,并且回放设备700可以能够通过两个或更多个扬声器(例如,第一扬声器驱动器752a和第二扬声器驱动器752b)在一个或多个方向上输出音频内容。音频内容可以由计算设备508通过云网络502提供。例如,计算设备可以是诸如Spotify、Amazon Music、Pandora之类的音乐服务提供商,并且云网络可以是诸如互联网之类的广域网(WAN)。

[0096] 回放设备700具有被配置为捕获和/或记录音频的一个或多个麦克风720。一个或多个麦克风720可以靠近回放设备700。例如,麦克风720可以物理地在回放设备上和/或物理地在回放设备中,或者有线或无线地连接到回放设备700。在一些实施例中,麦克风(未示出)可以位于房间中远离回放设备700的位置。例如,麦克风可以位于如网络设备上,网络设备可以是控制器设备、NMD或另一音频回放设备。在一些实施例中,回放设备700(或网络设备)处记录的音频可以包括源音频内容,并且可以用于确定回放设备的估计频率响应(即,自响应),并且自响应可用于校准音频回放设备。

[0097] a. 使用复合音频信号执行声学回声消除的示例

[0098] 图8是根据公开的实施例的执行多声道声学回声消除的方法800的流程图。本文公开的方法和其他过程可以包括一个或多个操作、功能或动作。尽管以连续顺序示出了各个框,但是这些框也可以并行执行,和/或以与本文描述的顺序不同的顺序执行。而且,根据所需的实现,可以将各个框组合成更少的框,分成更多的框,和/或移除框。

[0099] 另外,对于本文公开的方法和其他过程和方法,流程图示出了本实施例的一种可能实现的功能和操作。在这方面,每个框可以表示模块、段或程序代码的一部分,其包括可由处理器执行以实现过程中的特定逻辑功能或步骤的一个或多个指令。程序代码可以存储在任何类型的计算机可读介质上,例如,包括磁盘或硬盘驱动器的储存设备。计算机可读介质可以包括非暂时性计算机可读介质,例如,短时间存储数据的计算机可读介质,如寄存器存储器、处理器高速缓存和随机存取存储器(RAM)。计算机可读介质还可以包括非暂时性介质,例如,辅存或持久性长期储存器,如只读存储器(ROM)、光盘或磁盘、紧凑盘只读存储器(CD-ROM)等。计算机可读介质还可以是任何其他易失性或非易失性存储系统。计算机可读介质可以被认为是计算机可读存储介质,例如有形的储存设备。另外,图中的每个框可以表

示被硬连线以执行过程中的特定逻辑功能的电路。

[0100] 在一个示例中,回放设备700(可以等同于回放设备200)至少部分地执行所公开的用于多声道声学回声消除的功能。在另一示例中,计算设备504-508至少部分地执行所公开的用于多声道声学回声消除的功能。在又一示例中,控制设备300执行用于多声道声学回声消除的功能。在另一示例中,用于执行多声道声学回声消除的功能可以至少部分地由一个或多个NMD 512-516执行。其他布置也是可能的。

[0101] 在图8所示的实施例中,示出了用于操作图7的回放设备700的方法800。简言之,在802,回放设备700接收源音频流。回放设备700的网络接口714(与图2的网络接口214相对应)接收源音频流,并且源音频流包括要由回放设备700回放的源音频内容。源音频内容包括第一声道音频流和第二声道音频流。

[0102] 在804,回放设备200可以通过第一扬声器驱动器752a回放第一声道音频流,从而产生第一声道音频输出756a,如图7所示。在806,回放设备200可以通过第二扬声器驱动器752b回放第二声道音频流,从而产生第二声道音频输出756b,如图7所示。在808,麦克风720从回放设备700的环境701接收或捕获所捕获的音频流,并且将指示所捕获的音频的数据存储在所捕获的音频流中。所捕获的音频流可以包括与第一声道音频输出756a相对应的第一部分、与第二声道音频输出756b相对应的第二部分、以及与用户760发出的声命令758相对应的第三部分。所捕获的音频流具有第一信噪比。

[0103] 在810,将第一声道音频流和第二声道音频流组合为复合音频信号。在811,将复合音频信号和所捕获的音频流变换到短时傅立叶变换域。在812,对复合音频信号执行声学回声消除,并且对复合音频信号执行声学回声消除产生声学回声消除输出。在814,将声学回声消除输出应用于所捕获的音频流,从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比,第二信噪比比第一信噪比大。

[0104] 在802处开始,网络接口214通过无线连接、有线连接和/或另一合适的传输方式来接收源音频流。源音频流可以包括要由回放设备200回放的源音频内容或者与要由回放设备200回放的源音频内容相对应。在图7的实施例中,一个或多个通信链路762将网络接口214连接到回放设备200的一个或多个处理器702,并且一个或多个通信链路762将源音频流提供给一个或多个处理器702。源音频流可以包括任意数量的合适的音频声道流,或者可以被划分为任意数量的合适的音频声道流,以在回放设备200上回放源音频内容。在图7的实施例中,源音频流可以包括第一声道音频流和第二声道音频流。一个或多个第一声道通信链路764a将回放设备200的一个或多个处理器702连接到第一扬声器驱动器752a,并且一个或多个第一声道通信链路764a将第一声道音频流提供给第一扬声器驱动器752a。另外,一个或多个第二声道通信链路764b将回放设备700的一个或多个处理器702连接到第二扬声器驱动器752b,并且一个或多个第二声道通信链路764b将第二声道音频流提供给第二扬声器驱动器752b。在其他实施例中,源音频内容还可以包括第三声道音频流和/或附加的(例如,第四、第五)音频流,并且一个或多个对应的声道通信链路可以将回放设备700的一个或多个处理器702连接到对应的扬声器驱动器。一个或多个通信链路可以采取印刷电路板上的迹线的形式。例如,如果回放设备200包括两个扬声器(或扬声器驱动器),则第一换能器或第一扬声器驱动器752a(例如,右扬声器)可以被配置为回放第一声道音频流,并且第二换能器或第二扬声器驱动器752b(例如,左扬声器)可以被配置为回放第二声道音频流。因

此,在该示例配置中,第一扬声器驱动器752a和第二扬声器驱动器752b可以协作以双声道立体声回放源音频内容。作为另一示例,如果回放设备700包括三个扬声器,则回放设备700的第一扬声器驱动器752a(例如,右扬声器)可以回放(或可以被配置为回放)第一声道音频流,回放设备700的第二扬声器驱动器752b(例如,左扬声器)可以回放(或可以被配置为回放)第二声道音频流,并且回放设备的第三扬声器驱动器(例如,中置扬声器、中音扬声器、低音喇叭)可以回放(或可以被配置为回放)第三声道音频流。在另一示例配置中,回放设备700的第一扬声器驱动器752a(例如,中置扬声器)可以回放(或可以被配置为回放)第一声道音频流,回放设备700的第二扬声器驱动器752b(例如,右扬声器)可以回放(或可以被配置为回放)第二声道音频流,并且回放设备的第三扬声器(例如,左扬声器)可以回放(或可以被配置为回放)第三声道音频流。在这些示例配置中,第一、第二和第三扬声器驱动器可以协作以三声道立体声回放源音频内容。可以设想任何数量的附加扬声器和包括源音频内容的相应附加音频流。

[0105] 源音频流(和/或第一声道音频流、第二声道音频流、第三声道音频流、以及与音频流有关联的任何其他附加的声道音频流)可以被分段为一个或多个数据块。例如,块可以采取音频内容的数字样本的数据包的形式。这些数据块可以存储在执行多声道声学回声消除的回放设备700上和/或存储在与多声道声学回声消除相关联的计算设备上。

[0106] 图9是被配置为接收或回放源音频流的音频管线900的框图。音频管线900包括源954和信号处理器902,其可以是与回放设备700相关联的任何信号处理器,例如处理器702。第一数模转换器966a和第一扬声器驱动器952a通过一个或多个第一声道通信链路964a耦接到信号处理器902。一个或多个第一声道通信链路964a可以与图7的第一声道通信链路764a的全部或部分相对应,并且第一扬声器驱动器952a可以与图7的第一扬声器驱动器752a相对应。第二数模转换器966b和第二扬声器驱动器952b通过一个或多个第二声道通信链路964b耦接到信号处理器902(和/或另一信号处理器(未示出)),并且一个或多个第二声道通信链路964b可以与图7的第二声道通信链路764b的全部或部分相对应,第二扬声器驱动器952b可以与图7的第二扬声器驱动器752b相对应。在所示的实施例中,第三数模转换器966c和第三扬声器驱动器952c通过一个或多个第三声道通信链路964c耦接到信号处理器902(和/或另一信号处理器)。然而,在一些实施例中,更少或更多的数模转换器和相关联的扬声器驱动器可以通过一个或多个通信链路耦接到信号处理器902(和/或另一信号处理器)。

[0107] 如图9所示,音频管线900设置在回放设备700上或者并入回放设备700中。然而,音频管线900可以在控制器设备、计算设备、NMD或其组合上,或者并入控制器设备、计算设备、NMD或其组合中。在音频管线900驻留在回放设备700上的情况下,通信链路可以采取印刷电路板上的迹线的形式。在音频管道驻留在其组合上的情况下,通信链路可以采用有线或无线网络(例如以太网或WiFi网络)的形式。在一些实施例中,音频管线900的全部或部分可以设置在回放设备的外壳上或者并入回放设备的外壳中。

[0108] 源954可以是诸如存储器或硬盘驱动器的存储设备,其存储源音频内容。备选地,源954可以是诸如音乐服务提供商之类的计算设备,其存储源音频内容并将其提供给音频回放设备。源音频内容可以采取在时域中定义音频内容的数字样本的音频文件的形式。源954可以位于将网络接口714连接到回放设备700的一个或多个处理器702的一个或多个通

信链路962(与图7的通信链路762相对应)上,或者沿着该一个或多个通信链路962。

[0109] 在音频回放设备输出音频信号之前,信号处理器902可以将一种或多种滤波算法应用于源音频内容。滤波算法可以基于回放设备的音量设置、回放设备的先前校准、设备朝向、内容类型等中的一项或多项而变化。此外,信号处理器902可以包括采样率转换器、比特深度转换器和声道上/下混频器中的一个或多个。采样率转换器可以改变源音频内容的采样率。采样率可以定义代表每单位时间的源音频内容的采样数量。比特深度转换器可以改变源音频内容信号的比特深度。比特深度可以是用于表示数字样本的比特数。声道上/下混频器可以混合来自不同声道(例如,立体声的左声道和右声道)的源音频内容。信号处理器902也可以执行其他功能。

[0110] 在一些实施例中,信号处理器902可以在数字域中处理源音频内容,并输出经处理的数字信号。第一数模转换器966a、第二数模转换器966b和第三数模转换器966c可以将信号处理器902的数字信号(例如,第一声道音频流、第二声道音频流和第三声道音频流的数字信号)转换为模拟信号(例如,第一声道模拟信号、第二声道模拟信号和第三声道模拟信号)。模拟信号可以输出到第一扬声器驱动器954a、第二扬声器驱动器954b和/或第三扬声器驱动器954c中的相应一个,其将模拟信号转换为可听音频。例如,音频输出可以包括:与第一扬声器驱动器954a正在回放的第一声道音频流相对应的第一声道音频输出756a(在图7中示出)、与第二扬声器驱动器954b正在回放的第二声道音频流相对应的第二声道音频输出756b(在图7中示出)、以及与第三扬声器驱动器956c正在回放的第三声道音频流相对应的第三声道音频输出。

[0111] 可以在图9的分接头967和/或分接头968a、968b、968c处接收在多声道声学回声消除中使用的源音频内容(例如,第一声道音频流和第二声道音频流)。每个分接头967、968a-968c可以是通信链路的端部,该通信链路在执行声学回声消除的处理器(例如,处理器770)与第一、第二或第三声道通信链路964a-964c中的相应一个之间延伸。分接头968a可以设置在信号处理器902与第一数模转换器966a之间的第一声道通信链路964a的一部分上,并且分接头968a可以设置为尽可能靠近第一数模转换器966a。类似地,分接头968b可以设置在信号处理器902与第二数模转换器966b之间的第二声道通信链路964b的一部分上,并且分接头968b可以设置为尽可能靠近第二数模转换器966b。另外,分接头968c可以设置在信号处理器902和第三数模转换器966c之间的第三声道通信链路964c的一部分上,并且分接头968c可以设置为尽可能靠近第三数模转换器966c。在源音频内容是在分接头967处接收的情况下,在执行多通道声学回声消除之前,可能需要将原本由信号处理器902应用的处理应用于源音频流,如下面所讨论的。

[0112] 如图7的示例配置所示,分接头768a(与图9的分接头968a相对应)是通信链路778a的端部,通信链路778a在第一声道通信链路764a的一部分与处理器770之间延伸,并且通信链路778a向执行声学回声消除的处理器770提供第一声道音频流。在图7中还示出,分接头768b(与图9的分接头968b相对应)是通信链路778b的端部,通信链路778b在第二声道通信链路764b的一部分与处理器770之间延伸,并且通信链路778b向执行声学回声消除的处理器770提供第二声道音频流。在其他实施例中,与图9的分接头968b相对应的分接头可以是通信链路(未示出)的端部,该通信链路在第三声道通信链路的一部分(与图9的第三声道通信链路964c相对应)与处理器770之间延伸,并且该通信链路向执行声学回声消除的处理器

770提供第三声道音频流。如果在回放设备700中包括其他扬声器驱动器,则可以将其他通信链路链接到处理器770。

[0113] 在808,回放设备700的麦克风接收所捕获的音频,所捕获的音频包括由与麦克风相关联的模数转换器从模拟信号转换的数字信号。所捕获的音频可以由回放设备700、网络设备控制器设备NMD或另一音频回放设备的麦克风以任何合适的方式来记录、捕获和/或存储。在一些实施例中,麦克风720可以是麦克风阵列中的一个。所捕获的音频流可以包括在麦克风的可检测频率范围内的任何噪声或可听事件。例如,如图7所示,用户760可以说出指向回放设备700的声命令758,并且声命令758被麦克风720捕获,使得所捕获的音频流包括向回放设备200发出的声命令758。在一些实施例中,特别在图7的实施例中,声命令758可以在回放设备700正在回放源音频流时发出。因此,所捕获的音频流可以包括与声命令758相对应的部分,并且还包含与第一扬声器驱动器752a播放的第一声道音频输出756a相对应的第一部分、与第二扬声器驱动器752b播放的第二声道音频输出756b相对应的第二部分,并且在一些实施例中,还包括与第三扬声器驱动器播放的第三声道音频输出相对应的部分。所捕获的音频流因此可以具有第一信噪比。

[0114] 第一部分、第二部分、与第三声道音频输出相对应的部分以及与所捕获的音频流中的声命令758相对应的部分可以是回放设备700输出的音频内容的时域表示(例如,第一声道音频输出、第二声道音频输出和第三声道音频输出)。所捕获的音频流可以被分段为一个或多个数据块,例如数据包。所捕获的音频流可以存储在回放设备200上或被传递给另一网络设备,例如计算设备、另一回放设、控制设备或NMD。

[0115] 正在执行多声道声学回声消除的回放设备700可以记录或捕获所捕获的音频流的全部或一部分。所捕获的音频流可以通过与回放设备700共位的至少一个麦克风720(例如,麦克风阵列606中的一个或多个麦克风)记录,或者所捕获的音频流可以通过与回放设备空间上位于不同位置的一个或多个麦克风记录。例如,另一回放设备可以记录所捕获的音频流,网络设备可以接收该所捕获的音频流,和/或NMD可以记录该所捕获的音频流。所捕获的音频流可以包括通过处理器的一个或多个操作从所捕获的音频流中导出的信号。

[0116] 如810处所示,可以将第一声道音频流、第二声道音频流(以及在一些实施例中,第三声道音频流)组合或混合为复合音频信号。例如,复合音频信号可以是第一声道音频流、第二声道音频流和第三声道音频流之和。在一些实施例中,可以将第一声道音频流、第二声道音频流和第三声道音频流从时域转换到频域(通过傅立叶变换或其他已知方法),并且复合音频信号可以是第一声道音频流、第二声道音频流和第三声道音频流在频域中之和。可以在频域中进一步处理复合音频信号,或者可以以任何合适或已知的方式将复合音频信号变换到时域中以进行进一步处理。

[0117] 可以通过设置在回放设备的外壳上或回放设备的外壳内的处理器将第一、第二和第三声道音频流组合或混合为复合音频信号,处理器例如是处理器702或音频处理组件208、和/或与回放设备相关联的任何其他处理器组件。备选地,可以通过远离回放设备200布置的处理器将第一、第二和第三声道音频流组合或混合为复合音频信号。

[0118] 在812,与回放设备700相关联的一个或多个处理器对复合音频信号执行声学回声消除,并且对复合音频信号执行声学回声消除产生或导致声学回声消除输出。在814,与回放设备200相关联的一个或多个处理器将声学回声消除输出应用于所捕获的音频流,从而

将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比,其中第二信噪比比第一信噪比大。

[0119] 执行声学回声消除的一个或多个处理器可以是与将声学回声消除输出应用于所捕获的音频流的处理器相同的处理器。在图7的实施例中,一个或多个处理器可以包括处理器770。在其他实施例中,一个或多个处理器可以包括处理器702、音频处理组件208(在图2示出)和/或与回放设备700或任何其他设备相关联的任何其他处理器组件。在一些实施例中,执行声学回声消除的处理器可以与将声学回声消除输出应用于所捕获的音频流的处理器不同,并且二者中的任一个可以是处理器770的处理器702。在图7的实施例中,一个或多个麦克风通信链路772将回放设备700的处理器770连接到麦克风720,并且一个或多个麦克风通信链路772将所捕获的音频流提供给处理器770。因此,经由通信链路772发送的所捕获的音频流可以具有第一信噪比。在具有两个或更多个麦克风的实施例中,一个或多个另外的麦克风通信链路将每个附加的麦克风连接到回放设备700的处理器770(或任何其他处理器),以将与每个麦克风相关联的所捕获的音频流提供给处理器770。

[0120] 可以以任何方式执行执行声学回声消除来产生声学回声消除输出并将声学回声消除输出应用于所捕获的音频流。一般来说,声学回声消除涉及两个输入:(1) 麦克风记录的信号(这里,是所捕获的音频流);以及(2) 通常从音频回放管线中的一个点(例如图9的示例)获取的参考信号(或多个参考信号),该信号最接近表示预期由扬声器输出的模拟音频(这里,是复合音频信号)。然后,声学回声消除尝试确定传递函数(也称为滤波器),该传递函数以最小误差将参考信号变换为所记录的麦克风信号。对声学回声消除产生的输出进行逆变换并将其与麦克风记录的信号混合,导致参考信号与麦克风记录的信号产生偏差。因此,用户的语音输入(声命令758)的“信号”保留在经处理的所捕获的音频流中,而来自回放设备700的扬声器的音频输出(第一声道音频输出756a和第二声道音频输出756b)的“噪声”被减少或移除。也就是说,在将声学回声消除输出应用于所捕获的音频流以产生第二信噪比之后,第二信噪比将比与所捕获的音频流在应用声学回声消除输出之前的第一信噪比大。

[0121] 将传递函数应用于参考信号可以解决可能导致声学回声消除变得不稳定的因素(例如,房间中的回声、音频输出换能器的非线性响应)。在一些实施例中,对于每个声道到每个麦克风(可以是麦克风阵列之一)和每个扬声器驱动器到每个麦克风,都可以存在传递函数。

[0122] 为了降低处理要求,可以在短时傅立叶变换[“STFT”]域中执行与声学回声消除相关联的处理。因此,在811,可以将参考信号(这里,是复合音频信号)和麦克风720记录的信号(这里,是所捕获的音频流)转换到STFT域。可以通过与回放设备700相关联的一个或多个处理器将复合音频信号和所捕获的音频流变换到STFT,并且可以在执行声学回声消除和/或将声学回声消除输出应用于所捕获的音频流之前,将复合音频信号和所捕获的音频流变换到STFT域。

[0123] 另外,声学回声消除算法可以包括或涉及一个或多个自适应跨频带(cross-band)滤波器,其涉及进一步降低计算复杂度的优化子过程。跨频带滤波器消除了对快速傅立叶变换算法的卷积的需要,其在保持处理能力的同时实现准确的响应。

[0124] 更详细地介绍声学回声消除,参考信号(这里,是复合音频信号)可以在频域中,并

且例如可以使用傅立叶变换将麦克风捕获或记录的信号(这里,是所捕获的音频流)从时域转换到频域。复合音频信号和所捕获的音频流可以分别被表示为在频域中具有幅度和相位的数据向量。传递函数可以是复合音频信号向量(S)与所捕获的音频流向量(M)之间的差。

[0125] 可以基于自适应回声消除算法来计算传递函数。自适应回声消除的特征方程式可以表示为:

$$[0126] \quad M = [S * H + X] \quad (1)$$

[0127] 其中,M是频域中的复数向量,表示所捕获的音频流的幅度和相位;

[0128] S是频域中的复数向量,表示复合音频信号向量的幅度和相位;

[0129] H是频域中的复数传递函数,表示在M中不存在任何所记录的干扰的情况下S与M之间的差;以及

[0130] X表示所记录的干扰,例如M中的静态噪声(例如,嗡嗡声)或背景噪声(例如,发言)。如果M中不存在所记录的干扰,则 $M = S * X$ 且 $X = 0$ 。

[0131] H的实部,例如,S到M的传递函数的幅度分量可以计算为:

$$[0132] \quad \text{mag}H_n = (\text{mag}S_n - \text{mag}M_n) a_n + (1 - a_n) (\text{mag}H_{n-1}) \quad (2)$$

[0133] 其中a是信噪比,例如,信号可以是S,并且干扰可以是所记录的干扰X。a可以表示为在0和1之间归一化的对数值的函数;

[0134] $\text{mag}S_n$ 是复合音频信号向量的幅度分量向量;

[0135] $\text{mag}M_n$ 是所捕获的音频流向量的幅度分量向量;以及

[0136] n是迭代。

[0137] 每次迭代时,使用来自一次迭代的所捕获的音频流向量 $\text{mag}M$ 的幅度分量向量、复合音频信号 $\text{mag}S$ 的幅度分量向量以及从S到M的传递函数 $\text{mag}H$ 的幅度分量来计算针对下一次迭代的S到M的传递函数的幅度分量。

[0138] 可以为每次迭代执行以下计算:

$$[0139] \quad \text{mag}M_n - \text{mag}S_n * \text{mag}H_n \quad (3)$$

[0140] 其中, $\text{mag}M_n$ 是所捕获的音频流向量的幅度分量向量;

[0141] $\text{mag}S_n$ 是复合音频信号向量的幅度分量向量;

[0142] $\text{mag}H_n$ 是从S到M的传递函数的幅度分量。

[0143] 如本领域普通技术人员将理解的,当该等式的结果为零或基本为零时,传递函数可以收敛,其指示从所捕获的音频流中移除了复合音频信号,仅留下来自用户的声命令,从而增加了所捕获的音频流中的信噪比。本领域普通技术人员将认识到,等式的结果可以备选地收敛在零的可接受范围内,并且该范围将允许减少所捕获的音频流中的复合音频信号,以允许在第一扬声器驱动器752a、第二扬声器驱动器752b回放的音频信号的情况下,检测到来自用户的声命令。

[0144] 因此,参考图7的实施例,一个或多个通信链路774可以将处理器774连接到另一组件776(其可以是处理器702或另一处理器,其分析所捕获的音频流以确定用户760是否已经发出声命令758和/或向回放设备700发出与声命令758相对应的命令)。因此,将声学回声消除输出应用于所捕获的音频流,从而在一个或多个通信链路774中产生第二信噪比,第二信噪比与应用声学回声消除输出之前的所捕获的音频流(例如,在沿一个或多个通信链路772发送之前的所捕获的音频流)相关联的第一信噪比大。因此,在第一声道音频输出756a和第二

声道音频输出756b的“噪声”被减少或移除的同时,用户的声命令758保留在所捕获的音频流中以供组件776处理。

[0145] 信噪比的增加可以取决于若干因素中的任何一个或全部,例如,第一声道音频输出756a和第二声道音频输出756b的回放音量、以及回放设备700所位于的房间的大小、尺寸、声学等。在一些实施例中,第二信噪比与第一信噪比之间的差可以在10db(或大约10dB)到20dB(或大约20dB)的范围内。

[0146] 通过将第一声道音频流和第二声道音频流(以及可选地,第三和其他音频流)组合为复合音频信号,可以仅对复合音频信号执行声学回声消除过程,而不对第一、第二和第三声道音频流中的每一个执行声学回声消除过程。因此,简化了声学回声消除,从而使用较不复杂的算法实现了适当大的信噪比,而无需对当前使用的回放设备CPU进行成本昂贵的升级。

[0147] b. 对信号并行地执行声学回声消除的示例

[0148] 在一些实施例中,可以对第一和第二(以及第三等)声道音频流并行地执行声学回声消除。也就是说,处理器可以对第一声道音频流执行声学回声消除以产生第一声学回声消除输出,处理器可以对第二声道音频流执行声学回声消除以产生第二声学回声消除输出,并且可选地,处理器可以对第三声道音频流执行声学回声消除以产生第三声学回声消除输出。一个或多个处理器可以与第二声道音频流以及可选地,第三声道音频流并行地对第一声道音频流执行声学回声消除。可以以任何已知的方式执行并行处理。然后,一个或多个处理器可以将第一声学回声消除输出、第二声学消除输出,以及可选地,第三声学消除输出应用于所捕获的音频流,从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比,并且第二信噪比比第一信噪比大。

[0149] c. 使用奇异值分解执行声学回声消除的示例

[0150] 在其他实施例中,可以将奇异值分解(“SVD”)合并到声学回声消除过程中,以降低计算复杂度,同时实现改善的信噪比。图10中示出了用于结合SVD操作回放设备700的示例性方法1000。

[0151] 在1002,回放设备700(例如,图7中所示的回放设备700)接收源音频流,并且回放设备700具有第一扬声器驱动器752a、至少第二扬声器驱动器752b和至少一个麦克风,如前所述。如前所述,回放设备700可以包括第三扬声器驱动器和任意数量的附加扬声器驱动器。通过回放设备700的网络接口714接收源音频流,并且源音频流包括要由回放设备700回放的源音频内容,如先前所解释的。源音频内容包括第一声道音频流、第二声道音频流,以及可选地,第三声道音频流或附加声道音频流,如先前所解释的。

[0152] 在1004,回放设备700通过第一扬声器驱动器752a回放第一声道音频流,从而产生第一声道音频输出756a,如前所述。在1006,回放设备700通过第二扬声器驱动器752b回放第二声道音频流,从而产生第二声道音频输出756b,如前所述。同样,如前所述,回放设备700可以通过第三扬声器驱动器回放第三声道音频流,从而产生第三声道音频输出。回放设备200的另外的扬声器驱动器可以回放另外的声道音频流,从而产生另外的声道音频输出。

[0153] 在1008,麦克风720接收或捕获音频流。如前所述,并参考图7的实施例,所捕获的音频流通过通信链路772发送到处理器770。所捕获的音频流可以包括:与第一声道音频输出756a相对应的第一部分、与第二声道音频输出756b相对应的第二部分、以及与用户760发

出的声命令758相对应的部分,并且所捕获的音频流具有第一信噪比。如前所述,所捕获的音频流还可以包括与第三声道音频输出相对应的第三部分和与附加声道音频输出相对应的附加部分。

[0154] 在1010,与回放设备700相关联的一个或多个处理器对第一声道音频流、第二声道音频流,以及可选地,第三声道音频流(也就是说,对第一、第二和第三声道音频流的并集)执行SVD以产生信号分量的组合集。在一些实施例中,与回放设备700相关联的一个或多个处理器可以同时第一声道音频流、第二声道音频流,以及可选地,第三声道音频流执行SVD。

[0155] 在其他实施例中,与回放设备700相关联的一个或多个处理器可以对第一声道音频流单独执行SVD以产生第一信号分量集,并且对第二声道音频流单独执行SVD以产生第二信号分量集。可选地,与回放设备700相关联的一个或多个处理器还可以对第三声道音频流单独执行SVD以产生第三信号分量集,并且如果适当,对附加声道音频流执行SVD以产生相应的附加信号分量集。

[0156] 在1012,与回放设备700相关联的一个或多个处理器基于一个或多个参数来选择或确定信号分量的组合集的第一子集。在一些实施例中,与回放设备700相关联的一个或多个处理器可以基于一个或多个参数来选择或确定信号分量的组合集的第二子集(和另外的子集)。

[0157] 在其他实施例中,与回放设备700相关联的一个或多个处理器可以基于一个或多个参数来选择或确定第一信号分量集的子集,并且与回放设备700相关联的一个或多个处理器可以基于一个或多个参数来选择或确定第二信号分量集的子集。可选地,与回放设备700相关联的一个或多个处理器可以基于一个或多个参数来选择或确定第三(和附加)信号分量集的子集。第二信号分量集的子集或第三(或附加)信号分量集的子集的一个或多个参数中的任一个或全部可以与第一信号分量集的子集的一个或多个参数中的任一个或全部相对应。

[0158] 在SVD中,数据是在统计域中而不是在时域或频域中表示,并且数据被投影到基于统计准则的轴集合上,因此并不固定。数据在SVD中的表示有效地将数据分离为单独的源,以揭示投影中的显著结构。在一些示例中,对对于第一、第二和第三声道音频流的数据集执行SVD,可以揭示在某一频率或时间的指示符(例如,峰值),并且这些指示符可以与信号在对应频率或时间的强分量(例如,强能量内容)相对应。然后,可以忽略或丢弃与不想要的源相对应的投影(例如,弱能量内容)。然后,可以将需要的数据投影回原始域或空间,从而减少原始数据集中包括的信息量。如本领域普通技术人员将意识到的,对原始数据的经滤波的子集(例如,能量含量高于阈值能量含量的子集和/或所计算的方差高于阈值方差的子集)执行声学回声消除比对原始数据的整个集合执行声学回声消除需要更少的处理资源。

[0159] 在如何执行SVD的示例中,SVD定理提供:

$$[0160] \quad A = USV^T$$

[0161] 其中

$$[0162] \quad U^T U = I$$

$$[0163] \quad V^T V = I \text{ (即, } U \text{ 和 } V \text{ 是正交的)}$$

[0164] A是MxN矩阵,它可以与第一声道音频流、第二声道音频流和/或第三声道音频流

(和/或任何另外的声道音频流)之一相对应或相关联。A的行和列可以与数据和实验条件相关联。例如,A中的数据可以与N个音频波形相对应或者包括N个音频波形,每个音频波形的长度可以为M个样本。U是MxM矩阵,其列为左奇异向量,S是MxN矩阵,具有奇异值并且是对角的,并且 V^T 具有右奇异向量的行。SVD表示原始数据在协方差矩阵为对角线的坐标系中的扩展。

[0165] 计算SVD包括找到 AA^T 和 $A^T A$ 的特征值和特征向量。 $A^T A$ 的特征向量组成V的列,且 AA^T 的特征向量组成U的列。另外,S中的奇异值是 AA^T 或 $A^T A$ 的特征值的平方根。奇异值是S的对角线条目,并按降序排列。奇异值始终是实数。如果A是实矩阵,则U和V也是实数。

[0166] S的值中在阈值以下的值可以被忽略或设置为0,因此在U和V的部分中的对应值也可以被忽略或设置为0。因此,当投影回原始域或空间时,数据集中包括的信息量将减少,从而简化了声学回声消除过程。

[0167] 可以基于若干参数中的任一个或全部来选择或确定信号分量的组合集的子集(或第一、第二和第三信号分量集的子集)。这些参数可以与阈值相对应,使得在阈值(与相对强或剧烈的分量相对应)以上的信号分量被选择并且包括在对应的子集中。例如,一个参数可以是能量含量,并且可以为子集选择能量含量高于第一阈值能量含量的信号分量。在一些实施例中,一个参数可以是方差,并且可以为子集选择所计算的方差高于第一阈值方差的信号分量。在一些实施例中,可以为子集选择能量含量高于第一阈值能量含量,并且所计算的方差高于第一阈值方差的所有信号分量。在一些实施例中,可以从第一、第二或第三信号分量集中的任何一个或全部中选择任何给定的子集。例如,仅第三信号分量集可以具有能量含量高于第一阈值能量含量的信号分量,并且第一信号分量集和第二信号分量集可以具有能量含量小于或等于第一阈值能量含量的信号分量。

[0168] 因此,在一些实施例中,信号分量的组合集的第一子集可以具有以下至少一项:(a)高于第一阈值能量含量的能量含量,或(b)高于第一阈值方差的所计算的方差。另外,信号分量的组合集的第二子集可以具有以下至少一项:(a)高于第二阈值能量含量的能量含量,或(b)高于第二阈值方差的所计算的方差。第一和第二阈值能量含量中的每一个可以不同,并且第一和第二阈值方差中的每一个可以不同。

[0169] 在其他实施例中,第一信号分量集的子集可以具有以下至少一项:(a)高于第一阈值能量含量的能量含量,或(b)高于第一阈值方差的所计算的方差,并且第二信号分量集的子集可以具有以下至少一项:(a)高于第二阈值能量含量的能量含量,或(b)高于第二阈值方差的所计算的方差。在一些实施例中,第三信号分量集的子集可以具有(a)高于第三阈值能量含量的能量含量,或(b)高于第三阈值方差的所计算的方差。第一、第二和第三阈值能量含量中的每一个可以相同,或者其中的一个或多个可以是唯一的。另外,第一、第二和第三阈值方差中的每一个可以相同,或者其中的一个或多个可以是唯一的。

[0170] 在1014,与回放设备700相关联的一个或多个处理器对信号分量的组合集的第一子集执行声学回声消除,并且对信号分量的组合集的第一子集执行声学回声消除产生第一声学回声消除输出。在一些实施例中,与回放设备700相关联的一个或多个处理器还可以对信号分量的组合集的第二子集执行声学回声消除。与回放设备700相关联的一个或多个处理器可以同时信号分量的组合集的第一子集和信号分量的组合集的第二子集(以及任何其他子集)执行声学回声消除,以产生第一声学回声消除输出。然而,与回放设备700相关联

的一个或多个处理器可以以任何顺序对信号分量的组合集的第一子集和信号分量的组合集的第二子集执行声学回声消除,以产生第一声学回声消除输出、第二声学回声消除输出以及其他的声学回声消除输出。

[0171] 在一些实施例中,与回放设备700相关联的一个或多个处理器对第一信号分量集的子集执行声学回声消除,并且对第一信号集的子集执行声学回声消除产生第一声学回声消除输出。另外,与回放设备700相关联的一个或多个处理器对第二信号分量集的子集执行声学回声消除,并且对第二信号子集的子集执行声学回声消除产生第二声学回声消除输出。另外,与回放设备700相关联的一个或多个处理器可以对第三信号分量集的子集执行声学回声消除,并且对第三信号量子集的子集执行声学回声消除产生第三声学回声消除输出。此外,与回放设备700相关联的一个或多个处理器可以对另外的信号分量集的子集执行声学回声消除,并且对另外的信号子集的子集执行声学回声消除产生另外的声学回声消除输出。如前所述,处理器770可以对第一信号分量集的子集、第二信号分量集的子集以及第三信号分量集的子集中的任何一个或全部执行声学回声消除。

[0172] 在1016,与回放设备700相关联的一个或多个处理器将第一声学回声消除输出应用于所捕获的音频流,从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比,其中第二信噪比比第一信噪比大。在一些实施例中,与回放设备700相关联的一个或多个处理器可以将第一声学回声消除输出应用于所记录的音频流,将第二声学回声消除输出应用于所捕获的音频流,以及可选地,将第三声学回声消除输出应用于所捕获的音频流,从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比。在其他实施例中,与回放设备700相关联的一个或多个处理器还可以将附加声学回声消除输出应用于所记录的音频流,以将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比。

[0173] 与回放设备700相关联的一个或多个处理器可以以任何合适的方式将第一声学回声消除输出(以及可选地,第二声学回声消除输出和第三(以及附加)声学回声消除输出)应用于所记录的音频流,例如前述关于复合音频信号所描述的方式。在一些实施例中,第一声学回声消除输出、第二声学回声消除输出和第三(和附加)声学回声消除输出中的两个或更多个可以同时应用于所记录的音频流。在其他示例中,第一声学回声消除输出、第二声学回声消除输出以及可选的第三(和附加)声学回声消除输出可以以并行、串行或其任何组合的方式应用于所记录的音频流。

[0174] 如前所述,执行声学回声消除的一个或多个处理器可以与将声学回声消除输出应用于所捕获的音频流的处理器是相同的处理器。在一些实施例中,一个或多个处理器可以包括处理器770、处理器702、音频处理组件208(在图2中示出)和/或与回放设备700或任何其他设备相关联的任何其他处理器组件。在一些实施例中,执行声学回声消除的一个或多个处理器可以与将声学回声消除输出应用于所捕获的音频流的一个或多个处理器不同。

[0175] 可以例如根据所公开技术的一个或多个实施例来执行声学回声消除。与回放设备700相关联的一个或多个处理器770可以将第一声学回声消除输出(以及第二声学回声消除输出和第三声学回声消除输出)应用于所捕获的音频流,以将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比。也就是说,参考图7的实施例,将第一和第二(以及其他)声学回声消除输出应用于所捕获的音频流会在一个或多个通信链路774中产生第二信噪比,并且第二信噪比与应用声学回声消除输出之前的所捕获的音频流(例如,在一个或多个通

信链路772中)相关联的第一信噪比大。因此,在减少或移除第一声道音频输出756a和第二声道音频输出756b的“噪声”的同时,用户的声命令758保留在所捕获的音频流中,以供组件776进行处理。

[0176] 为了降低处理要求,可以在短时傅立叶变换[“STFT”]域中执行与声学回声消除相关联的处理。也就是说,可以将信号分量和所捕获的音频流(和/或与声学回声消除相关联的任何其他信号)的组合集的子集变换到STFT域。在一些实施例中,可以将第一信号分量集的子集、第二信号分量集的子集、第三信号分量集的子集以及所捕获的音频流(和/或与声学回声消除相关联的任何其他信号)变换到STFT域。另外,可以在声学回声消除操作中对任何合适的信号集使用自适应跨频带滤波器。此外,在将第一和第二(以及可选地,第三)声学回声消除输出应用于所捕获的音频流的过程中应用的滤波器可以是复合滤波器,包括针对每个声道到每个麦克风(其可以是麦克风阵列中的一个麦克风)和/或每个扬声器驱动器到每个麦克风的组合传递函数。

[0177] 因此,通过对第一、第二和第三声道音频流执行SVD,并且通过基于SVD选择信号分量的子集,可以识别出最相关的信号分量,并且丢弃最不显著的信号分量。相对于对第一、第二和第三声道音频流中的每一个执行声学回声消除,这大大简化了声学回声消除操作。因此,现有的回放设备CPU可用于声学回声消除操作,以实现所捕获的音频流的信噪比的令人满意的增加,并且信噪比的增加可以在一定范围内(或高于阈值),以允许在由第一扬声器驱动器752a、第二扬声器驱动器752b回放的音频信号的“噪声”的情况下,检测到来自用户760的声命令758。在一些示例中,(并且取决于第一声道音频输出756a和第二声道音频输出756b的回放音量和/或回放设备700所位于的房间的大小、尺寸、声学等),第二信噪比与第一信噪比之间的差可以在10db(或大约10db)至20db(或大约20db)的范围内。因此,可以提高与回放设备700相关联的语音控制的性能,而不会招致更强大的CPU的成本。

[0178] d. 使用参考声道执行声学回声消除的示例

[0179] 在其他实施例中,可以将相关参考声道并入声学回声消除过程中,并且在图11中示出了用于操作并入有相关参考声道的回放设备700的示例性方法1100。在1102,回放设备700接收源音频流,并且回放设备具有第一扬声器驱动器752a、至少第二扬声器驱动器752b和至少一个麦克风720,如前所述。如先前所解释的,回放设备700可以包括第三扬声器驱动器和任意数量的附加扬声器驱动器。通过回放设备700的网络接口714接收源音频流,并且源音频流包括要由回放设备700回放的源音频内容,如先前所解释的。源音频内容包括第一声道音频流、第二声道音频流,以及可选地,第三声道音频流或附加声道音频流,如先前所解释的。在一些实施例中,第一声道音频流可以是相关参考声道。然而,在其他实施例中,第二声道音频流、第三声道音频流或其他音频流中的任何一个可以是相关参考声道。

[0180] 在1104,回放设备700通过第一扬声器驱动器752a回放第一声道音频流,从而产生第一声道音频输出756a,如前所述。在1106,回放设备700通过第二扬声器驱动器752b回放第二声道音频流,从而产生第二声道音频输出756b,如前所述。同样,如前所述,回放设备700可以通过第三扬声器驱动器回放第三声道音频流,从而产生第三声道音频输出。回放设备700的另外的扬声器驱动器可以回放另外的声道音频流,从而产生另外的声道音频输出。

[0181] 在1108,麦克风720接收或捕获音频流。如前所述,所捕获的音频流可以通过通信链路772发送给处理器770。所捕获的音频流包括:与第一声道音频输出756a相对应的第一

部分和与第二声道音频输出756b相对应的第二部分,以及与用户760发出的声命令758相对应的部分,并且所捕获的音频流具有第一信噪比。如前所述,所捕获的音频流还可以包括与第三声道音频输出相对应的第三部分和与附加声道音频输出相对应的附加部分。

[0182] 在1110,确定相关参考声道与第二声道音频流之间的互相关,以产生经相关的第二声道信号,并且经相关的第二声道信号是第二声道音频流相对于相关参考声道的特有部分。因此,仅将对第二声道音频流(相对于相关参考声道)特有的信号包括在经相关的第二声道信号中。

[0183] 在一些实施例中,确定相关参考信道和第三声道音频流之间的互相关以产生经相关的第三声道信号,并且经相关的第三声道信号是第三声道音频流相对于相关参考声道的特有部分。因此,仅将对第三声道音频流(相对于相关参考声道)特有的信号包括在经相关的第三声道信号中。可以确定相关参考信道与任何附加声道音频流之间的互相关,以产生另外的经相关的声道信号。

[0184] 如果相关参考声道是第二声道音频流,则确定相关参考声道与第一声道音频流之间的互相关以产生经相关的第二声道信号。确定相关参考通道与第三声道音频流之间的互相关以产生经相关的第三声道信号,并且经相关的第三声道信号是第三声道音频流相对于相关参考声道的特有部分。类似地,如果相关参考声道是第三声道音频流,则确定相关参考声道与第一声道音频流之间的互相关以产生经相关的第二声道信号,并且确定相关参考声道与第二声道音频流之间的互相关以产生经相关的第三声道信号。

[0185] 可以以任何方式执行互相关,并且可以通过设置在回放设备700的壳体上或设置在回放设备700的壳体内的处理器(例如,处理器702、处理器770、音频处理组件208(在图2中示出)和/或与回放设备700相关联的任何其他处理器组件)来执行任何一个或所有互相关。备选地,可以由远离回放设备700设置的处理器执行任何一个或所有互相关。

[0186] 在1112,与回放设备700相关联的一个或多个处理器对相关参考声道执行声学回声消除,并且对相关参考声道执行声学回声消除产生了第一声学回声消除输出。在1114处,与回放设备700相关联的一个或多个处理器对经相关的第二声道信号执行声学回声消除,并且对经相关的第二声道信号执行声学回声消除产生了第二声学回声消除输出。另外,与回放设备700相关联的一个或多个处理器可以对经相关的第三声道信号执行声学回声消除,并且对经相关的第三声道信号执行声学回声消除产生了第三声学回声消除输出。与回放设备700相关联的一个或多个处理器可以对附加的经相关的声道信号执行声学回声消除,并且对附加的经相关的声道信号执行声学回声消除可以产生附加声学回声消除输出。如前所解释的,处理器770可以对相关参考声道、经相关的第二声道信号和经相关的第三声道信号中的任何一个或全部执行声学回声消除。

[0187] 在1116,与回放设备700相关联的一个或多个处理器将第一声学回声消除输出应用于所记录的音频流,并且将第二声学回声消除输出应用于所捕获的音频流,从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比,并且第二信噪比比第一信噪比大。

[0188] 在一些实施例中,与回放设备700相关联的一个或多个处理器可以将第一声学回声消除输出应用于所记录的音频流,将第二声学回声消除输出应用于所捕获的音频流,并且将第三声学回声消除输出应用于所捕获的音频流,从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比。在其他实施例中,与回放设备700相关联的一个或多个处理器

还可以将附加声学回声消除输出应用于所记录的音频流,以将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比。

[0189] 与回放设备700相关联的一个或多个处理器可以以任何方式将第一声学回声消除输出、第二声学回声消除输出、以及可选地将第三(和附加)声学回声消除输出应用于所记录的音频流。例如,第一声回声消除输出、第二声回声消除输出、以及可选地第三(和附加)声回声消除输出中的两个或更多个可以同时应用于所记录的音频流。在其他示例中,第一声学回声消除输出、第二声学回声消除输出以及可选的第三(和附加)声学回声消除输出可以以并行、串行或其任何组合的方式应用于所记录的音频流。

[0190] 如前所述,执行声学回声消除的一个或多个处理器可以与将声学回声消除输出应用于所捕获的音频流的处理器是相同的处理器。在一些实施例中,一个或多个处理器可以包括处理器770、处理器702、音频处理组件208(在图2中示出)和/或与回放设备700或任何其他设备相关联的任何其他处理器组件。在一些实施例中,执行声学回声消除的一个或多个处理器可以与将声学回声消除输出应用于所捕获的音频流的一个或多个处理器不同。

[0191] 可以例如根据所公开技术的一个或多个实施例来执行声学回声消除。另外,与回放设备700相关联的一个或多个处理器可以以任何已知方式将第一声学回声消除输出、第二声学回声消除输出、以及可选地将第三(和附加)声学回声消除输出应用于所捕获的音频流,以将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比。也就是说,参考图7的实施例,将第一和第二(以及其他)声学回声消除输出应用于所捕获的音频流会在一个或多个通信链路774中产生第二信噪比,并且第二信噪比与应用声学回声消除输出之前的所捕获的音频流(例如,在一个或多个通信链路772中)相关联的第一信噪比大。因此,在第一声道音频输出756a和第二声道音频输出756b的“噪声”被减少或移除的同时,用户的声命令758保留在所捕获的音频流中以供组件776处理。

[0192] 如前所述,可以在短时傅立叶变换[“STFT”]域中执行与声学回声消除相关联的处理。也就是说,可以将相关参考声道、经相关的第二声道信号、经相关的第三声道信号以及所捕获的音频流(和/或与声学回声消除相关联的任何其他信号)变换到STFT域。另外,可以在声学回声消除操作中对任何合适的信号集使用自适应跨频带滤波器。此外,在将第一和第二(以及可选地,第三)声学回声消除输出应用于所捕获的音频流的过程中应用的滤波器可以是复合滤波器,包括针对每个声道到每个麦克风(其可以是麦克风阵列中的一个麦克风)和/或每个扬声器驱动器到每个麦克风的组合传递函数。

[0193] 因此,通过确定相关参考声道与第二声道音频流之间的互相关以及相关参考声道与第三声道音频流之间的互相关,可以相对于对第一声道音频流、第二声道音频流和第三声道音频流中的每一个执行声学回声消除减少声学回声消除操作的计算复杂度。因此,现有的回放设备CPU可以用于声学回声消除操作,以实现所捕获的音频流的信噪比令人满意的提高(例如,在10db到20dB范围内的提高),从而改善与回放设备700相关联的语音控制的性能而不会招致更强大的CPU的成本。

[0194] e. 使用互相关执行声学回声消除的示例

[0195] 在其他实施例中,可以将没有固定参考(例如,相关参考声道)的互相关并入声学回声消除过程,并且在图12中示出了用于操作并入有相关参考声道的回放设备700的示例性方法1200。

[0196] 在1202,回放设备700接收源音频流,并且回放设备700具有第一扬声器驱动器752a、至少第二扬声器驱动器752b和至少一个麦克风720,如前所述。如前所述,回放设备700可以包括第三扬声器驱动器和任意数量的附加扬声器驱动器。可以通过回放设备700的网络接口714接收源音频流,并且源音频流可以包括要由回放设备700回放的源音频内容,如先前所解释的。如前所述,源音频内容可以包括第一声道音频流、第二声道音频流,以及可选地,第三声道音频流或附加声道音频流。

[0197] 在1204,回放设备700可以通过第一扬声器驱动器752a回放第一声道音频流,从而产生第一声道音频输出756a,如前所述。在1206,回放设备700可以通过第二扬声器驱动器752b回放第二声道音频流,从而产生第二声道音频输出756b,如前所述。同样,如前所述,回放设备700可以通过第三扬声器驱动器回放第三声道音频流,从而产生第三声道音频输出。回放设备700的另外的扬声器驱动器可以回放另外的声道音频流,从而产生另外的声道音频输出。

[0198] 在1208,麦克风720接收或捕获可以被麦克风720接收或捕获的音频流。如前所述,所捕获的音频流可以通过通信链路772发送给处理器770。所捕获的音频流包括:与第一声道音频输出756a相对应的第一部分、与第二声道音频输出756b相对应的第二部分、以及由用户760发出的声命令758相对应的部分,并且所捕获的音频流具有第一信噪比。如前所述,所捕获的音频流还可以包括与第三声道音频输出相对应的第三部分和与附加声道音频输出相对应的附加部分。

[0199] 在1210,与回放设备700相关联的一个或多个处理器可以基于一个或多个参数来确定或选择第一声道音频流、第二声道音频流以及可选地,第三声道音频流(以及任何附加声道音频流)的第一子集。此外,与回放设备700相关联的一个或多个处理器可以基于一个或多个参数来确定或选择第一声道音频流、第二声道音频流以及可选地,第三声道音频流(以及任何附加声道音频流)的第二子集。还可以确定或选择其他子集。

[0200] 在一些实施例中,与回放设备700相关联的一个或多个处理器基于一个或多个参数来选择或确定第一声道音频流、第二声道音频流以及可选地,第三声道音频流的第一子集可以包括或包含将第一声道音频流、第二声道音频流以及可选地,第三声道音频流的互相关。选择或确定第一子集可以基于互相关所产生的一个或多个参数。例如,选择或确定第一子集可以包括从第一声道音频流、第二声道音频流以及第三声道音频流中的任何一个或所有选择或确定能量含量高于第一阈值能量含量和/或相关性高于第一阈值相关性的信号分量。

[0201] 在一些实施例中,基于一个或多个参数来选择或确定第一声道音频流、第二声道音频流以及可选地,第三声道音频流的第二子集可以包括或包含将第一声道音频流、第二声道音频流以及可选地,第三声道音频流的互相关。选择或确定第二子集可以基于互相关所产生的一个或多个参数。例如,选择或确定第二子集可以包括从第一声道音频流、第二声道音频流以及第三声道音频流中的任何一个或所有选择或确定能量含量高于第二阈值能量含量和/或相关性高于第二阈值相关性的信号分量。第二阈值能量含量可以与第一阈值能量含量不同(例如,小于)和/或第二阈值相关性可以与第一阈值相关性不同(例如,小于)。

[0202] 在一些实施例中,第一子集可以是对第一声道音频流、第二声道音频流以及可选

地,第三声道音频流公共的信号。然后,第二子集可以是第一声道音频流特有的信号,第三子集可以是第二声道音频流特有的信号,并且第四子集可以是第三声道音频流特有的信号。

[0203] 可以以任何方式执行互相关,并且可以通过设置在回放设备700的壳体上或设置在回放设备700的壳体内的处理器(例如,处理器702、处理器770、音频处理组件208(在图2中示出)和/或与回放设备700相关联的任何其他处理器组件)来执行任何一个或所有互相关。备选地,可以由远离回放设备700设置的处理器执行任何一个或所有互相关。

[0204] 在1212,与回放设备700相关联的一个或多个处理器对第一子集执行声学回声消除,并且对第一子集执行声学回声消除产生第一声学回声消除输出。可选地,与回放设备700相关联的一个或多个处理器可以对第二子集执行声学回声消除,并且对第二子集执行声学回声消除产生第二声学回声消除输出。另外,与回放设备700相关联的一个或多个处理器可以可选地对第三(和另外的)子集执行声学回声消除,以产生第三(和另外的)声学回声消除输出。如前所述,处理器770可以对第一子集、第二子集和第三子集中的任何一个或全部执行声学回声消除。

[0205] 在1214,与回放设备700相关联的一个或多个处理器将第一声学回声消除输出应用于所记录的音频流,从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比,并且第二信噪比比第一信噪比大。在一些实施例中,与回放设备700相关联的一个或多个处理器可以将第一声学回声消除输出应用于所记录的音频流,将第二声学回声消除输出应用于所捕获的音频流,以及可选地将第三(和其他)声学回声消除输出应用于所捕获的音频流,以将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比。

[0206] 在一些实施例中,与回放设备700相关联的一个或多个处理器可以以任何方式将第一声学回声消除输出、第二声学回声消除输出、以及可选地将第三(和附加)声学回声消除输出应用于所记录的音频流。例如,第一声学回声消除输出、第二声学回声消除输出、以及可选地第三(和附加)声学回声消除输出中的两个或更多个可以同时应用于所记录的音频流。在其他示例中,第一声学回声消除输出、第二声学回声消除输出以及可选的第三(和附加)声学回声消除输出可以以并行、串行或其任何组合的方式应用于所记录的音频流。

[0207] 如前所述,执行声学回声消除的一个或多个处理器可以与将声学回声消除输出应用于所捕获的音频流的处理器是相同的处理器。在一些实施例中,一个或多个处理器可以包括处理器770、处理器702、音频处理组件208(在图2中示出)和/或与回放设备700或任何其他设备相关联的任何其他处理器组件。在一些实施例中,执行声学回声消除的一个或多个处理器可以与将声学回声消除输出应用于所捕获的音频流的一个或多个处理器不同。

[0208] 可以例如根据所公开技术的一个或多个实施例来执行声学回声消除。与回放设备700相关联的一个或多个处理器可以以任何已知的方式将第一声学回声消除输出,以及可选地第二声学回声消除输出和第三(和附加)声学回声消除输出应用于所捕获的音频流,以将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比。也就是说,参考图7的实施例,将第一和第二(以及其他)声学回声消除输出应用于所捕获的音频流会在一个或多个通信链路774中产生第二信噪比,并且第二信噪比与应用声学回声消除输出之前的所捕获的音频流(例如,在一个或多个通信链路772中)相关联的第一信噪比大。在一些实施例中,第二信噪比与第一信噪比之间的差可以在10dB(或大约10dB)至20dB(或大约20dB)的范围内。

因此,在第一声道音频输出756a和第二声道音频输出756b的“噪声”被减少或移除的同时,用户的声命令758保留在所捕获的音频流中以供组件776处理。

[0209] 如前所述,可以在短时傅立叶变换[“STFT”]域中执行与声学回声消除相关联的处理。也就是说,可以将声道音频流的第一、第二和第三子集和所捕获的音频流(和/或与声学回声消除相关联的任何其他信号)转换到STFT域。另外,可以在声学回声消除操作中对任何合适的信号集使用自适应跨频带滤波器。此外,在将第一和第二(以及可选地,第三)声学回声消除输出应用于所捕获的音频流的过程中应用的滤波器可以是复合滤波器,包括针对每个声道到每个麦克风(其可以是麦克风阵列中的一个麦克风)和/或每个扬声器驱动器到每个麦克风的组合传递函数。

[0210] 因此,通过确定第一声道音频流、第二声道音频流和任何附加的音频流之间的互相关,声学回声消除操作的计算复杂度可以相对于对第一声道音频流、第二声道音频流和第三声道音频流中的每一个执行声学回声消除降低,如前所述。

[0211] 可以以任何合适的方式来触发执行本文所述的声学回声消除(和/或将声学回声消除输出应用于所捕获的音频流)。例如,与回放设备700相关联的一个或多个处理器可以检测到回放功能由回放设备700启动。在其他示例中,与回放设备700相关联的一个或多个处理器可以在启动回放功能之后检测到回放设备接收到取消静音命令。

[0212] 本文提供的示例涉及方法、回放设备以及有形非暂时性计算机可读介质。操作回放设备的方法的实施例可以包括:通过回放设备的网络接口接收源音频流,源音频流包括要由回放设备回放的源音频内容。回放设备包括第一扬声器驱动器和至少第二扬声器驱动器,并且还包含一个或多个麦克风。源音频内容包括第一声道音频流和第二声道音频流。方法还可以包括:通过经由第一扬声器驱动器回放第一声道音频流来产生第一声道音频输出,并且通过经由第二扬声器驱动器回放第二声道音频流来产生第二声道音频输出。方法还可以包括:通过一个或多个麦克风接收所捕获的音频流,所捕获的音频流包括与第一声道音频输出相对应的第一部分,并且还包括与第二声道音频输出相对应的第二部分,其中所捕获的音频流具有第一信噪比。方法可以另外包括:将第一声道音频流和第二声道音频流组合为复合音频信号,并且将复合音频信号和所捕获的音频流转换到短时傅立叶变换域。方法还包括:对复合音频信号执行声学回声消除,并且对复合音频信号执行声学回声消除产生了声学回声消除输出。方法可以另外包括:将声学回声消除输出应用于所捕获的音频流,从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比,其中第二信噪比比第一信噪比大。

[0213] 在另一方面,提供了一种非暂时性计算机可读介质。非暂时性计算机可读介质上存储有指令,所述指令可由计算设备执行以使计算设备执行功能。功能包括:通过回放设备的网络接口接收源音频流,源音频流包括要由回放设备回放的源音频内容,并且回放设备包括第一扬声器驱动器和至少第二扬声器驱动器,并且还包括至少一个麦克风。源音频内容包括第一声道音频流和第二声道音频流。功能还可以包括:通过经由第一扬声器驱动器回放第一声道音频流来产生第一声道音频输出,并且通过经由第二扬声器驱动器回放第二声道音频流来产生第二声道音频输出。功能还可以包括:通过一个或多个麦克风接收所捕获的音频流,所捕获的音频流包括与第一声道音频输出相对应的第一部分,并且还包括与第二声道音频输出相对应的第二部分,其中所捕获的音频流具有第一信噪比。功能可以另

外包括：将第一声道音频流和第二声道音频流组合为复合音频信号，并且将复合音频信号和所捕获的音频流转换到短时傅立叶变换域。功能还包括：对复合音频信号执行声学回声消除，并且对复合音频信号执行声学回声消除产生了声学回声消除输出。功能还可以包括：将声学回声消除输出应用于所捕获的音频流，从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比，其中第二信噪比比第一信噪比大。

[0214] 在又一方面，提供了一种回放设备。回放设备包括第一扬声器驱动器、至少第二扬声器驱动器、一个或多个麦克风，处理器以及存储器。存储器上存储有指令，所述指令可由处理器执行以使回放设备执行功能。功能包括：通过回放设备的网络接口接收源音频流，源音频流包括要由回放设备回放的源音频内容。源音频内容包括第一声道音频流和第二声道音频流。功能还可以包括：通过经由第一扬声器驱动器回放第一声道音频流来产生第一声道音频输出，并且通过经由第二扬声器驱动器回放第二声道音频流来产生第二声道音频输出。功能还可以包括：通过一个或多个麦克风接收所捕获的音频流，所捕获的音频流包括与第一声道音频输出相对应的第一部分，并且还包含与第二声道音频输出相对应的第二部分，其中所捕获的音频流具有第一信噪比。功能可以另外包括：将第一声道音频流和第二声道音频流组合为复合音频信号。功能还包括：对复合音频信号执行声学回声消除，并且对复合音频信号执行声学回声消除产生了声学回声消除输出。功能还可以包括：将声学回声消除输出应用于所捕获的音频流，从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比，其中第二信噪比比第一信噪比大。

[0215] 在另一方面，所捕获的音频流包括与用户发出的声命令相对应的第三部分，并且其中，将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比导致所捕获的音频流中的第一部分和第二部分被消除或最小化。

[0216] 在另一方面，功能还包括：检测触发以对复合音频信号执行声学回声消除，其中，检测触发包括检测：(a) 回放功能由回放设备启动或 (b) 在启动回放功能之后，回放设备接收到取消静音命令。

[0217] 操作回放设备的方法的另一方面可以包括：通过回放设备的网络接口接收源音频流，源音频流包括要由回放设备回放的源音频内容。回放设备包括第一扬声器驱动器和至少第二扬声器驱动器，并且还包含一个或多个麦克风。源音频内容包括第一声道音频流和第二声道音频流。方法还可以包括：通过经由第一扬声器驱动器回放第一声道音频流来产生第一声道音频输出，并且通过经由第二扬声器驱动器回放第二声道音频流来产生第二声道音频输出。方法还可以包括：通过一个或多个麦克风接收所捕获的音频流，所捕获的音频流包括与第一声道音频输出相对应的第一部分和与第二声道音频输出相对应的第二部分，并且还包含与第二声道音频输出相对应的第二部分，其中所捕获的音频流具有第一信噪比。方法还可以包括：对第一声道音频流和第二声道音频流执行奇异值分解，以产生信号分量的组合集。方法可以另外包括：基于一个或多个参数来选择信号分量的组合集的子集。方法还可以包括：对信号分量的组合集的子集执行声学回声消除，其中，对信号分量的组合集的子集执行声学回声消除产生第一声学回声消除输出。方法可以另外包括：将第一声学回声消除输出应用于所捕获的音频流，从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比，其中第二信噪比比第一信噪比大。

[0218] 在另一方面，提供了一种非暂时性计算机可读介质。非暂时性计算机可读介质上

存储有指令,所述指令可由计算设备执行以使计算设备执行功能。功能包括:通过回放设备的网络接口接收源音频流,源音频流包括要由回放设备回放的源音频内容。回放设备包括第一扬声器驱动器和至少第二扬声器驱动器,并且还包括一个或多个麦克风。源音频内容包括第一声道音频流和第二声道音频流。功能还可以包括:通过经由第一扬声器驱动器回放第一声道音频流来产生第一声道音频输出,并且通过经由第二扬声器驱动器回放第二声道音频流来产生第二声道音频输出。功能还可以包括:通过一个或多个麦克风接收所捕获的音频流,所捕获的音频流包括与第一声道音频输出相对应的第一部分和与第二声道音频输出相对应的第二部分,并且还包括与第二声道音频输出相对应的第二部分,其中所捕获的音频流具有第一信噪比。功能还可以包括:对第一声道音频流和第二声道音频流执行奇异值分解,以产生信号分量的组合集。功能可以另外包括:基于一个或多个参数来选择信号分量的组合集的子集。功能还可以包括:对信号分量的组合集的子集执行声学回声消除,其中,对第一信号集的子集执行声学回声消除产生第一声学回声消除输出。功能还可以包括:将第一声学回声消除输出应用于所捕获的音频流,从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比,其中,第二信噪比比第一信噪比大。

[0219] 在又一方面,提供了一种回放设备。回放设备包括第一扬声器驱动器、至少第二扬声器驱动器、一个或多个麦克风、处理器以及存储器。存储器上存储有指令,所述指令可由处理器执行以使回放设备执行功能。功能包括:通过回放设备的网络接口接收源音频流,源音频流包括要由回放设备回放的源音频内容。源音频内容包括第一声道音频流和第二声道音频流。功能还可以包括:通过经由第一扬声器驱动器回放第一声道音频流来产生第一声道音频输出,并且通过经由第二扬声器驱动器回放第二声道音频流来产生第二声道音频输出。功能还可以包括:通过一个或多个麦克风接收所捕获的音频流,所捕获的音频流包括与第一声道音频输出相对应的第一部分和与第二声道音频输出相对应的第二部分,其中所捕获的音频流具有第一信噪比。功能还可以包括:对第一声道音频流和第二声道音频流执行奇异值分解,以产生信号分量的组合集。功能可以另外包括:基于一个或多个参数来选择信号分量的组合集的子集。功能还可以包括:对信号分量的组合集的子集执行声学回声消除,其中,对信号分量的组合集的子集执行声学回声消除产生第一声学回声消除输出。功能还可以包括:将第一声学回声消除输出应用于所捕获的音频流,从而将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比,其中,第二信噪比比第一信噪比大。

[0220] 在另一方面,所捕获的音频流包括与用户发出的声命令相对应的第三部分,并且其中,将所捕获的音频流的信噪比从第一信噪比增加到第二信噪比导致所捕获的音频流中的第一部分和第二部分被消除或最小化。

[0221] 在又一方面,功能可以另外包括:检测触发以对第一信号分量集的子集或第二信号分量集的子集执行声学回声消除,其中,检测触发包括:a)回放功能由回放设备启动,或者(b)回放功能启动后,回放设备接收到取消静音命令。

[0222] IV. 结论

[0223] 以上描述尤其公开了各种示例系统、方法、装置和尤其包括在硬件上执行的固件和/或软件的制品。应当理解的是,这些示例仅是示意性的,而不应当被认为是限制性的。例如,可以想到,这些固件、硬件和/或软件方面或组件中的任意一个或全部可以专门在硬件中实现、专门在软件中实现、专门在固件中实现、或在硬件、软件和/或固件的任意组合中实

现。因此,所提供的示例不是实现这些系统、方法、装置和/或制品的唯一方式。

[0224] 此外,本文对“实施例”的提及意味着结合实施例描述的特定特征、结构或特性可以包括在本发明的至少一个示例性实施例中。在说明书中各处出现该短语不一定都指代相同的实施例,也不是与其他实施例互斥的分离的或备选的实施例。因此,本领域技术人员应当显式地和隐式地理解的是,本文所描述的实施例可以与其他实施例组合。

[0225] 主要在说明性的环境、系统、过程、步骤、逻辑块、处理以及直接或间接地与耦接到网络的数据处理设备的操作相类似的其他象征性表示的方面上,提出本说明书。本领域技术人员通常使用这些处理描述和表示,以向本领域技术人员的其他技术人员传播他们的工作内容。阐述了各种具体细节,以提供本公开的透彻理解。然而,本领域技术人员应理解,不需要特定、具体细节就可以实施本公开。在其他实例中,没有描述熟知的方法、过程、组件和电路,以避免不必要地使实施例的方面模糊不清。因此,本公开的范围由随附权利要求、而不是以上实施例的描述来界定。

[0226] 当随附权利要求中的任一项权利要求被理解成涵盖纯软件和/或固件实现时,在此明确限定至少一个示例中的至少一个元素以包括存储软件和/或固件的非暂时性有形介质,如存储器、DVD、CD、蓝光等。



图1

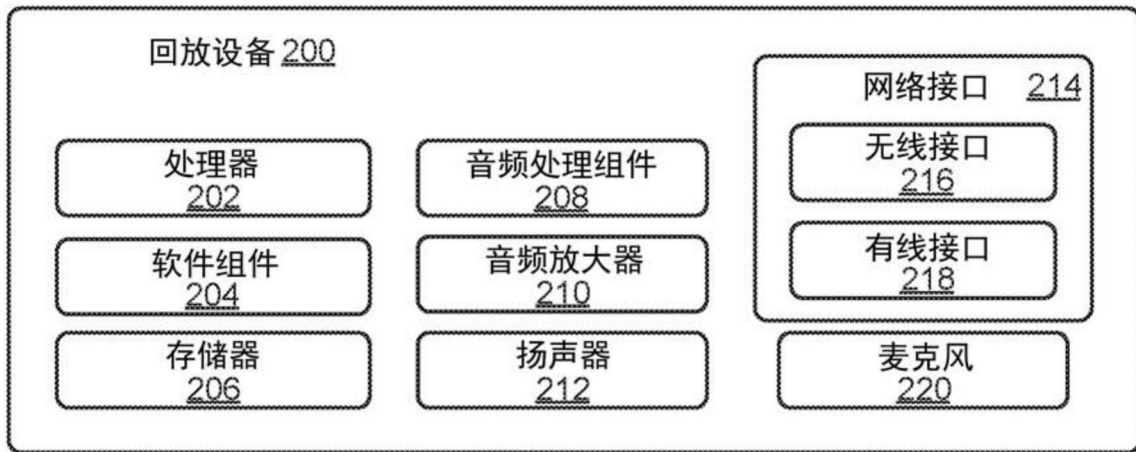


图2

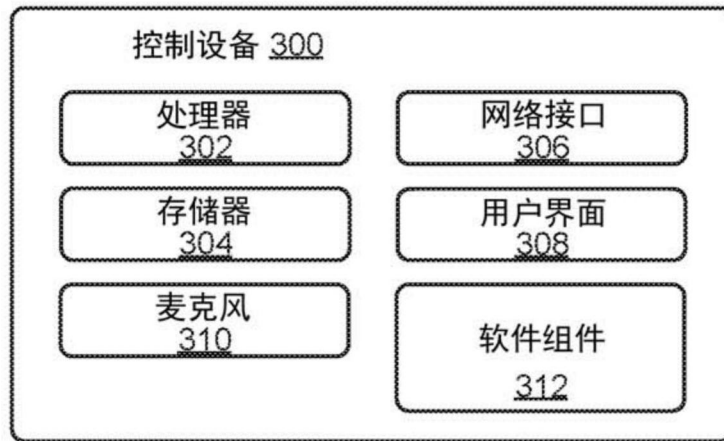


图3

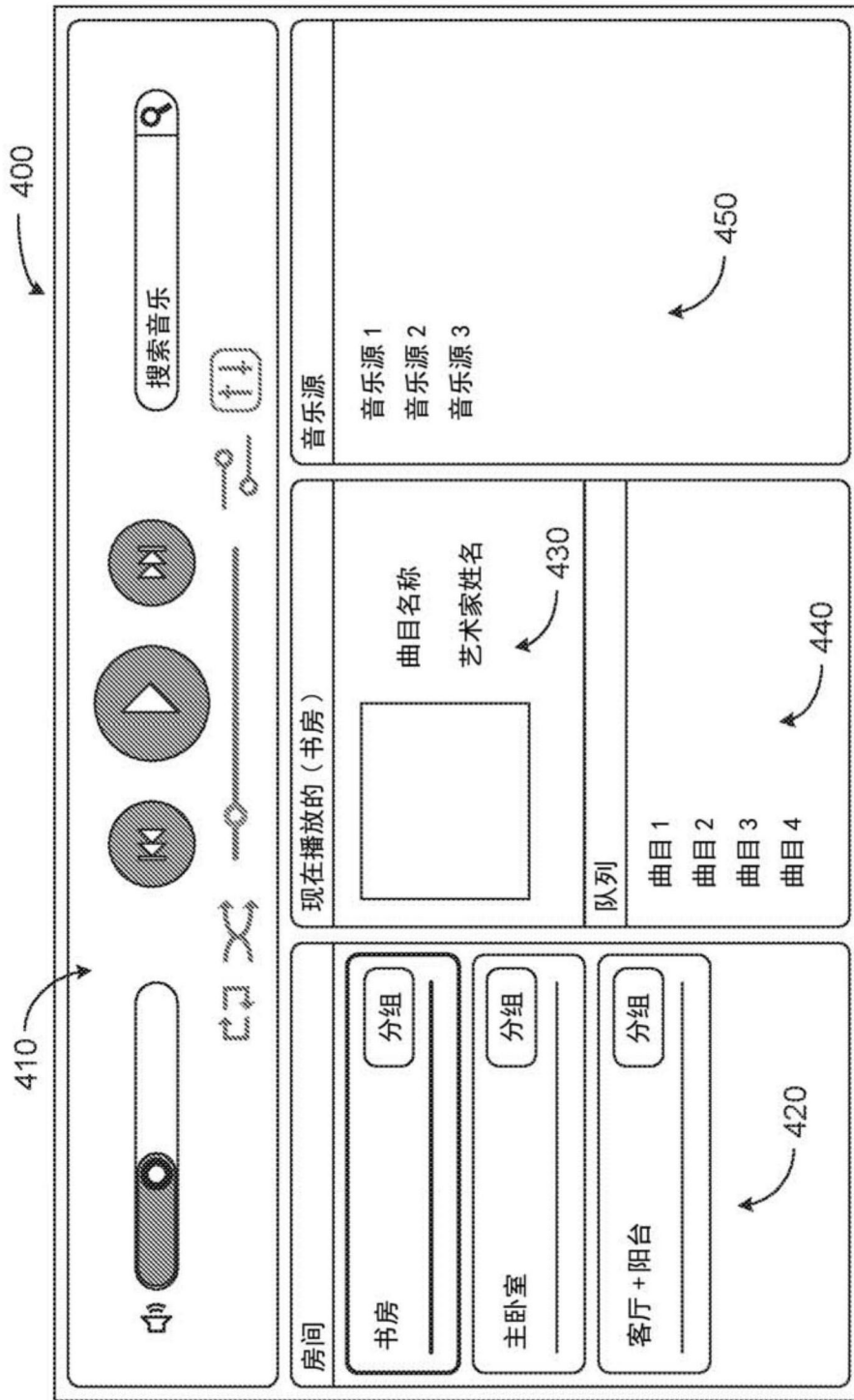


图4

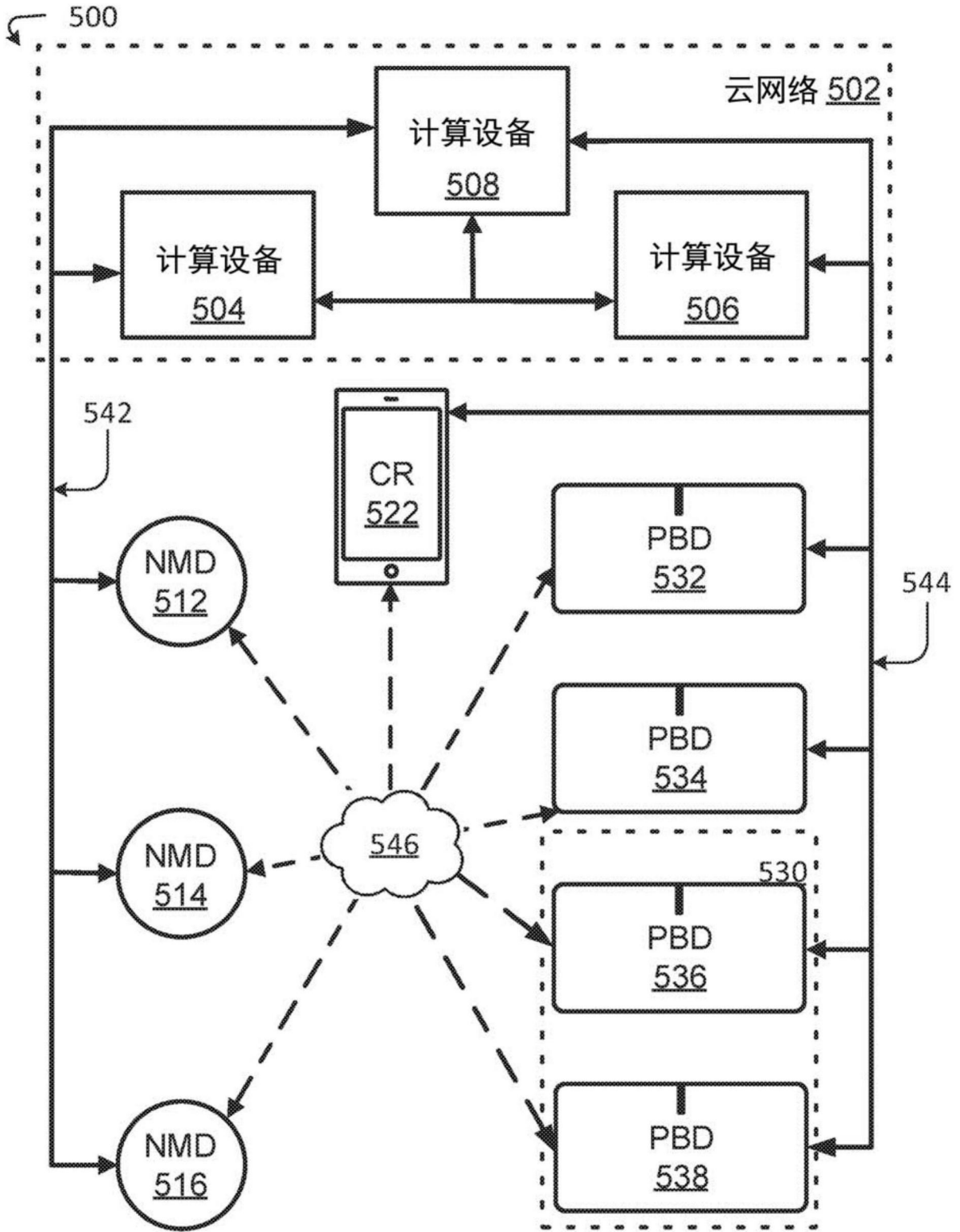


图5



图6

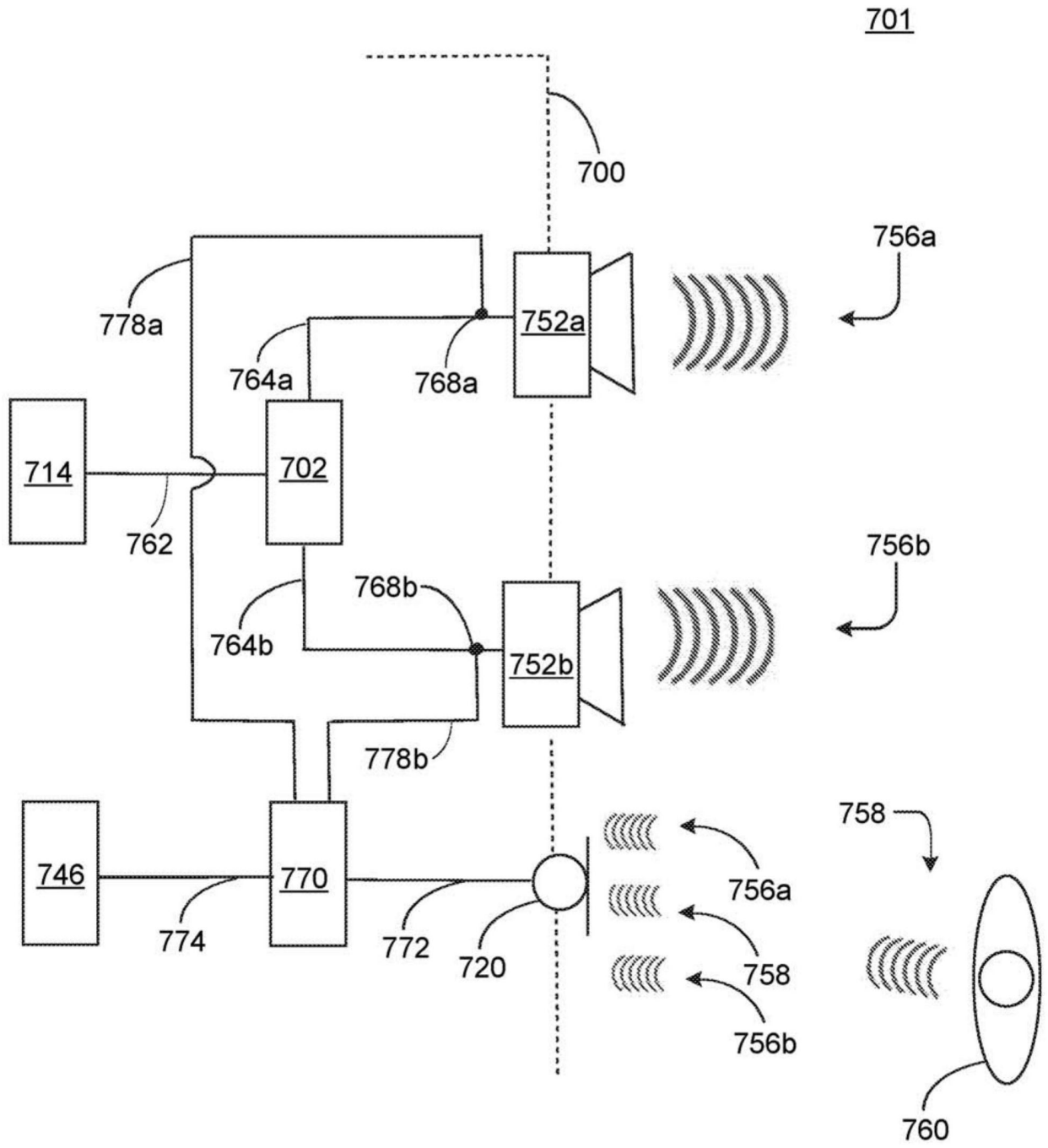


图7

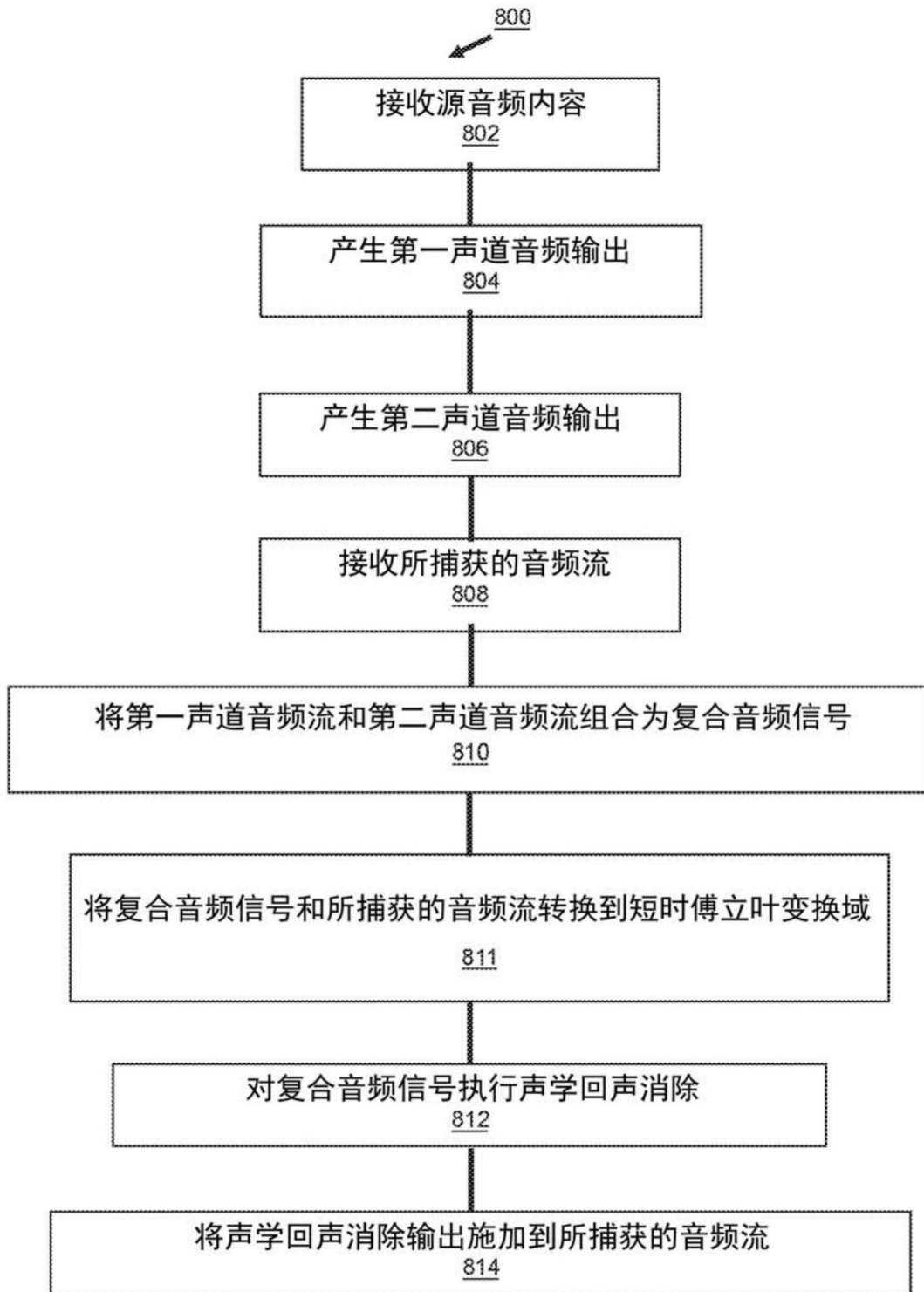


图8

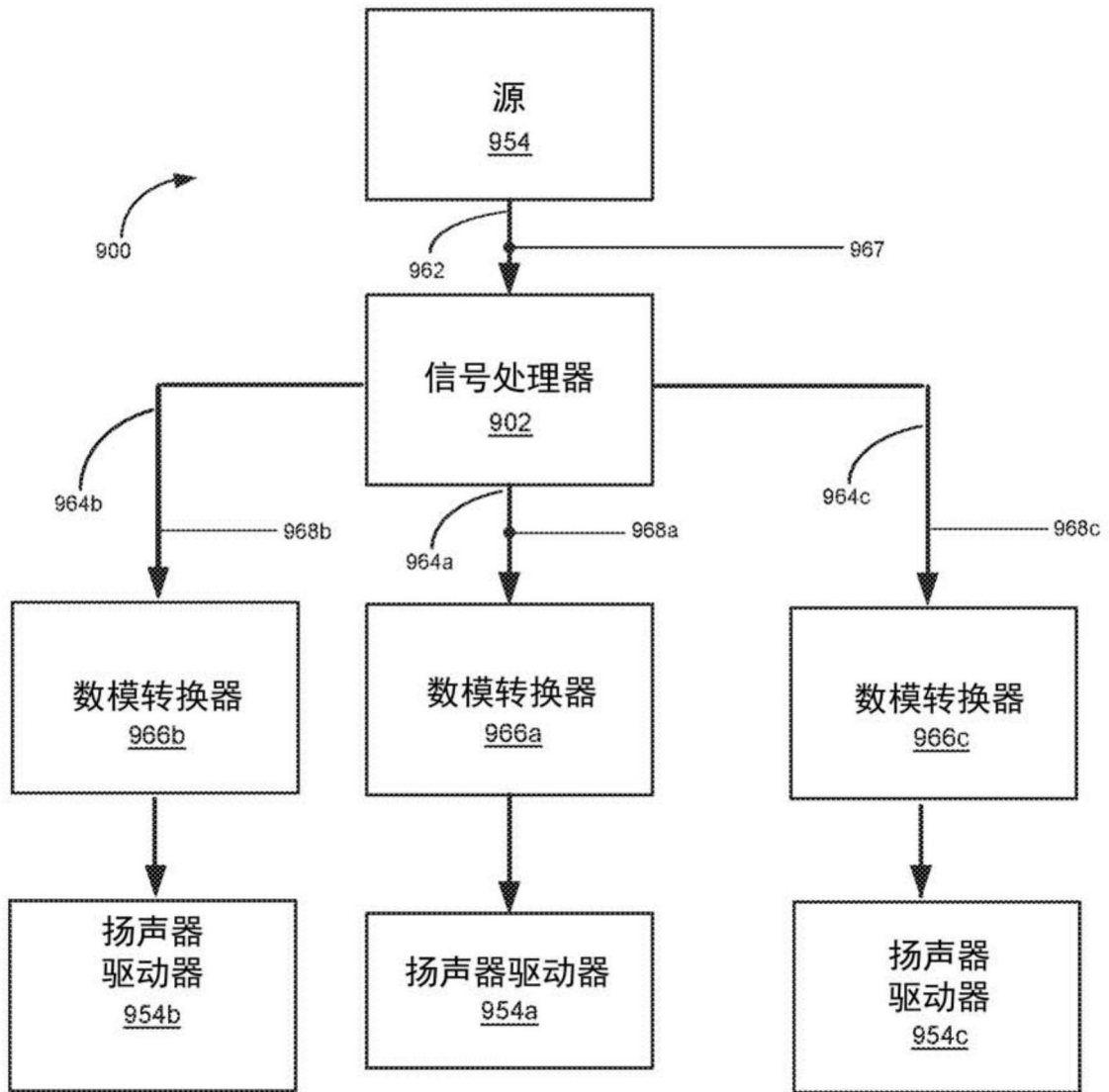


图9

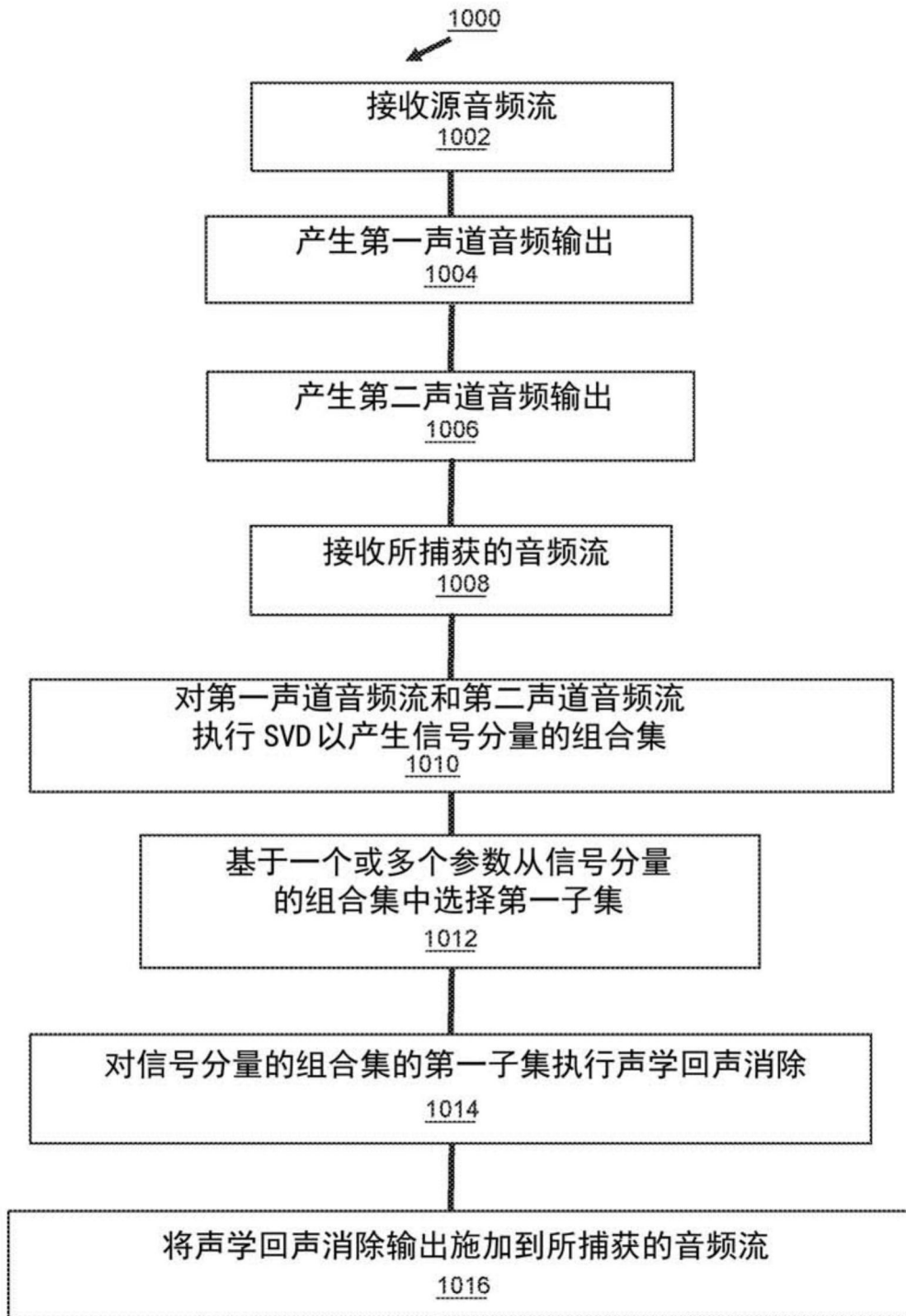


图10

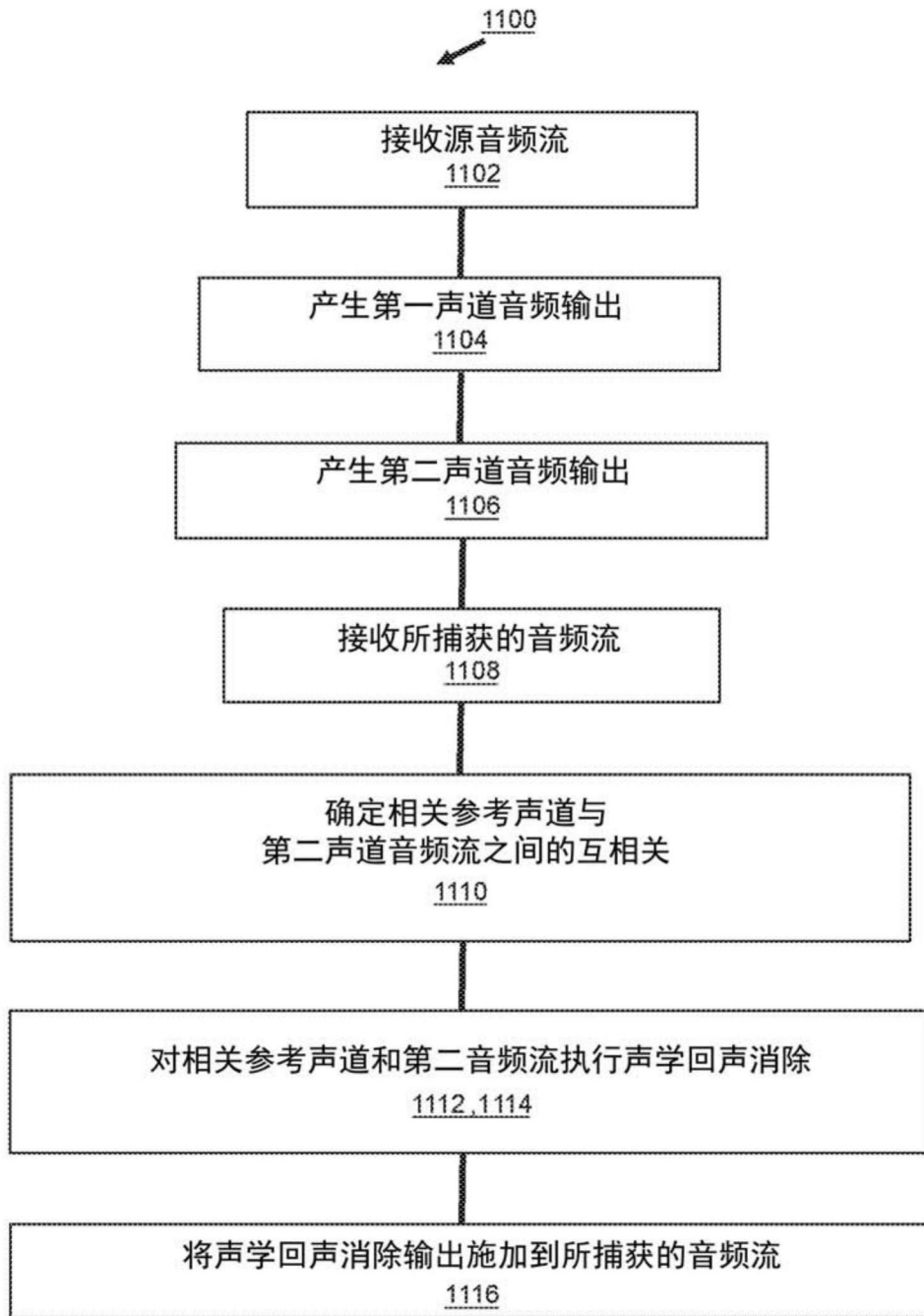


图11

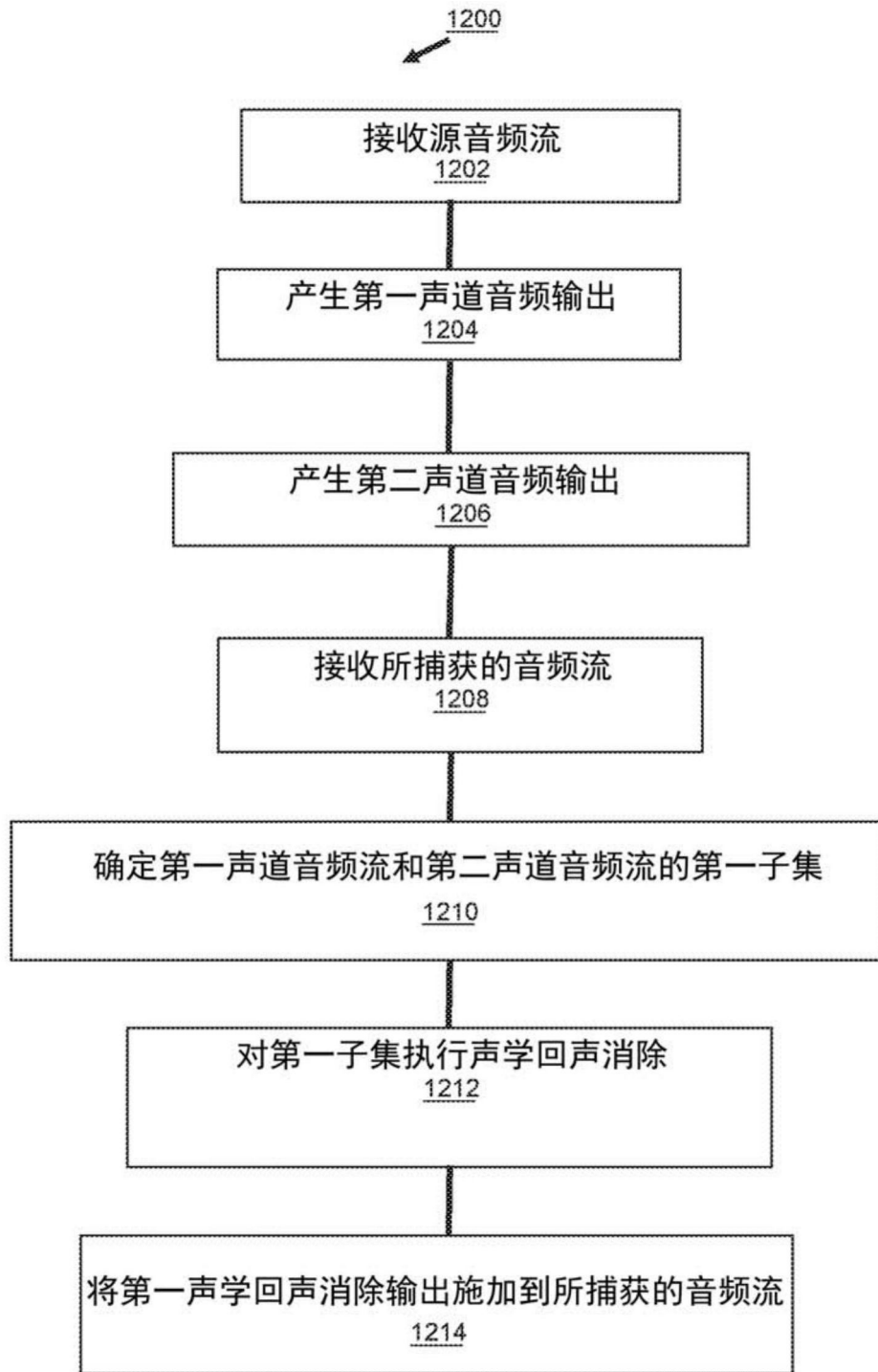


图12