



(12)发明专利申请

(10)申请公布号 CN 111479196 A

(43)申请公布日 2020.07.31

(21)申请号 202010302650.7

(22)申请日 2017.02.21

(30)优先权数据

- 62/298,418 2016.02.22 US
- 62/298,425 2016.02.22 US
- 62/298,350 2016.02.22 US
- 62/298,388 2016.02.22 US
- 62/298,410 2016.02.22 US
- 62/298,433 2016.02.22 US
- 62/298,439 2016.02.22 US
- 62/298,393 2016.02.22 US
- 62/312,350 2016.03.23 US
- 15/098,718 2016.04.14 US
- 15/098,805 2016.04.14 US
- 15/131,776 2016.04.18 US
- 15/223,218 2016.07.29 US

(62)分案原申请数据

201780025028.2 2017.02.21

(71)申请人 搜诺思公司

地址 美国加利福尼亚州

(72)发明人 西蒙·贾维斯

尼古拉斯·A·J·米林顿

凯斯·科尔宾 马克·普拉格

罗米·卡德里 克里斯多夫·巴茨

衍年·威利·陈

(74)专利代理机构 中科专利商标代理有限责任公司 11021

代理人 王蓉

(51)Int.Cl.

H04R 3/00(2006.01)

H04R 3/12(2006.01)

H04R 27/00(2006.01)

H04R 29/00(2006.01)

G06F 3/16(2006.01)

G10L 15/22(2006.01)

G10L 21/02(2013.01)

H04L 12/28(2006.01)

H04S 7/00(2006.01)

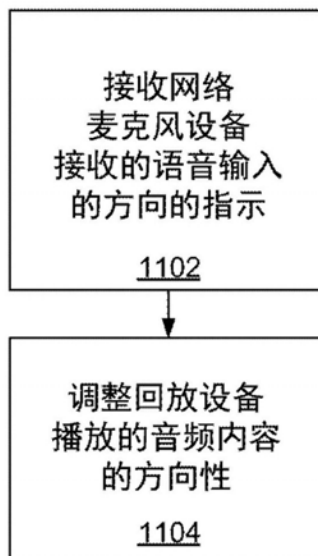
权利要求书2页 说明书50页 附图17页

(54)发明名称

媒体回放系统的语音控制

(57)摘要

接收包括命令字、一个或多个媒体变量实例和一个或多个区域变量实例的语音输入。确定对应于命令字的媒体回放系统命令。识别对应于一个或多个媒体变量实例的媒体内容。媒体回放系统可以基于一个或多个区域变量实例对媒体内容执行媒体回放系统命令。



1. 一种方法,包括:

通过计算设备来使媒体回放系统在第二回放区域中回放媒体内容;

通过所述计算设备来接收语音输入,所述语音输入包括命令字和一个或多个区域变量实例;

通过所述计算设备来确定与所述命令字相对应的媒体回放系统命令以及基于所述一个或多个区域变量实例在第二回放区域中回放所述媒体内容的意图;以及

在所述确定之后,通过所述计算设备来使所述媒体回放系统在所述第二回放区域中回放所述媒体内容,其中,所述第一回放区域包括第一回放设备,并且所述第二回放区域包括第二回放设备。

2. 根据权利要求1所述的方法,还包括:响应于所述确定,通过所述计算设备来形成包括所述第一回放区域和所述第二回放区域在内的区域组,其中,在接收所述语音输入之前,所述第二回放区域不是所述区域组的成员。

3. 根据权利要求2所述的方法,还包括:在所述形成之后,通过所述计算设备来使所述第一回放设备从所述区域组中移除,以使得在所述第一回放设备离开所述区域组之后,所述第二回放设备继续在所述第二回放区域中回放所述媒体内容。

4. 根据权利要求1所述的方法,其中,所述媒体回放系统命令包括用于移动所述媒体内容的回放的命令,并且其中,所述方法还包括:响应于用于移动所述媒体内容的回放的所述命令,通过所述计算设备来使所述第一回放设备停止在所述第一回放区域中对所述媒体内容的回放。

5. 根据权利要求1所述的方法,还包括:响应于所述确定,通过所述计算设备来使所述媒体回放系统将所述第一回放区域的回放队列的至少一部分复制到所述第二回放区域的回放队列。

6. 根据权利要求1所述的方法,其中,所述语音输入是第二语音输入,其中,所述命令字是第二命令字,并且其中,所述方法还包括:

通过所述计算设备来接收第一语音输入,其中,所述第一语音输入包括第一命令字和一个或多个媒体变量实例;

通过所述计算设备来确定与所述第一命令字相对应的媒体回放系统命令以及基于所述一个或多个媒体变量实例回放所述媒体内容的意图;以及

通过所述计算设备来基于所述一个或多个媒体变量实例识别所述媒体内容,其中,所述媒体回放系统响应于与所述第一命令字相对应的媒体回放系统命令而在所述第一回放区域中回放所述媒体内容。

7. 根据权利要求1所述的方法,其中,所述语音输入还包括一个或多个媒体变量实例,并且其中,所述方法还包括:通过所述计算设备来基于所述一个或多个媒体变量实例识别所述媒体内容。

8. 根据权利要求1所述的方法,其中,通过所述计算设备来使所述媒体回放系统在所述第二回放区域中回放所述媒体内容包括:使所述媒体回放系统与所述第一回放区域中的所述媒体内容的回放同步地在所述第二回放区域中回放所述媒体内容。

9. 一种计算设备,包括:

一个或多个处理器;

存储指令的存储器,所述指令在由所述一个或多个处理器执行时,使所述计算设备执行根据权利要求1至8中的一项所述的方法。

10.一种计算机可读存储介质,所述计算机可读存储介质中存储有指令,所述指令用于使计算设备执行根据权利要求1至8中的任一项所述的方法。

11.一种计算机程序,所述计算机程序被存储在计算机可读存储介质中,所述计算机程序用于使计算设备执行根据权利要求1至8中的任一项所述的方法。

## 媒体回放系统的语音控制

[0001] 本申请是申请日为2017年2月21日、申请号为201780025028.2的中国发明专利申请“媒体回放系统的语音控制”的分案申请。

[0002] 相关申请的交叉引用

[0003] 本申请根据PCT第8条要求以下各项的优先权权益：(i) 2016年7月29日提交的题为“Voice Control of a Media Playback System”的美国非临时专利申请No.15/223,218、(ii) 2016年4月14日提交的题为“Music Service Selection”的美国非临时专利申请No.15/098,718、(iii) 2016年4月14日提交的题为“Room-Corrected Voice Detection”的美国非临时专利申请No.15/098,805、(iv) 2016年4月18日提交的题为“Action based on User ID”的美国非临时专利申请No.15/131,776、(v) 2016年3月23日提交的题为“Voice Control of a Media Playback System”的美国临时专利申请No.62/312,350、(vi) 2016年2月22日提交的题为“Audio Response Playback”的美国临时专利申请No.62/298,418、(vii) 2016年2月22日提交的题为“Music Service Selection”的美国临时专利申请No.62/298,425、(viii) 2016年2月22日提交的题为“Metadata Exchange Involving a Networked Playback System and a Networked MicrophoneSystem”的美国临时专利申请No.62/298,350、(ix) 2016年2月22日提交的题为“Handling of Loss of Pairing Between Networked Devices”的美国临时专利申请No.62/298,388、(x) 2016年2月22日提交的题为“Default Playback Device(s)”的美国临时专利申请No.62/298,410、(xi) 2016年2月22日提交的题为“Room-Corrected Voice Detection”的美国临时专利申请No.62/298,433、(xii) 2016年2月22日提交的题为“Content Mixing”的美国临时专利申请No.62/298,439、以及(xiii) 2016年2月22日提交的题为“Action Based on User ID”的美国临时专利申请No.62/298,393。这些专利申请中的每件专利申请的内容通过引用全部合并于此。

### 技术领域

[0004] 本公开涉及消费者产品，更具体地，涉及与媒体回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

[0005] 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无线高保真系统使人们可以通过一个或多个联网回放设备体验来自多个源的音乐。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，人们可以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，可以将不同的歌曲流传输到具有回放设备的每个房间，可以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。

[0006] 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

[0007] 参考以下说明书、所附权利要求和附图,将更好地理解本公开的技术的特征、方面和优点,在附图中:

[0008] 图1示出了可以实施某些实施例的示例媒体回放系统配置;

[0009] 图2示出了根据本文描述的方面的示例回放设备的功能框图;

[0010] 图3示出了根据本文描述的方面的示例控制设备的功能框图;

[0011] 图4示出了根据本文描述的方面的示例控制器界面;

[0012] 图5示出了根据本文描述的方面的示例多个网络设备;

[0013] 图6示出了根据本文描述的方面的示例网络麦克风设备的功能框图;

[0014] 图7是关于向网络麦克风设备提供环境的声学特性(acoustics)的示例流程图;

[0015] 图8是关于向网络麦克风设备提供环境的声学特性的另一示例流程图;

[0016] 图9是关于解释网络麦克风设备接收的语音输入的示例流程图;

[0017] 图10是关于确定环境的声学特性的示例流程图;

[0018] 图11是关于对回放设备播放的音频内容施加方向性的另一示例流程图;

[0019] 图12示出了根据本文描述的方面的示例音乐控制方法的流程图;

[0020] 图13A-13D示出了根据本文描述的方面的配置过程的示例用户界面;

[0021] 图14示出了根据一些实施例的示例方法;

[0022] 图15示出了根据一些实施例的另一示例方法;

[0023] 图16是关于通过网络麦克风设备识别流传输音乐服务的示例流程图;以及

[0024] 图17是关于通过网络麦克风设备识别流传输音乐服务的另一示例流程图。

[0025] 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

### I. 概述

[0027] 通常,语音命令可以是控制本文讨论的任何媒体回放系统控制的命令。例如,语音命令可以是媒体回放系统通过该媒体回放系统的一个或多个回放设备来播放媒体内容的命令。在另一示例中,语音命令可以是修改媒体回放系统的一个或多个媒体回放设备的回放设置的命令。回放设置可以包括例如回放音量、回放传输控制、音乐源选择、分组等。

[0028] 在一个方面中,本文描述的示例涉及向网络麦克风设备(NMD)提供NMD的操作环境的声学特性。NMD可以在解释NMD接收的语音命令时使用环境的声学特性。

[0029] NMD可以通过麦克风阵列接收语音输入(例如,讲话)并基于语音输入执行功能的设备。例如,NMD可以接收语音命令并解释接收到的语音命令。然后,NMD可以基于语音命令执行功能。具体地,NMD可以接收“播放来自‘流传输服务1’的‘艺术家1’的‘曲目1’”的语音命令,确定语音输入是播放音频内容的命令,然后促使播放所请求的来自‘流传输服务1’的‘艺术家1’的‘曲目1’。其他布置也是可能的。

[0030] NMD的操作环境可以具有特定声学特性。声学特性定义了声音在环境中的行进方

式。可以以许多方式确定环境的声学特性。作为一个示例,具有扬声器和麦克风的回放设备可以通过扬声器输出测试音调,通过麦克风接收测试音调,并且分析接收到的测试音调来确定环境的声学特性。在另一示例中,另一网络设备(例如,控制设备、第一回放设备或甚至NMD自身)可以包括从第二回放设备的扬声器接收测试音调的麦克风,并分析接收到的测试音调以确定环境的声学特性。测试音调可以是不同频率的音频声音。这些声学特性可以存储在通信网络中的计算设备上或回放设备上,以供回放设备在回放音频内容时使用。

[0031] 在一些示例中,可以基于环境的已知特性(如房间的大小、房间天花板的高度和房间中的家具)来推断声学特性。数据库可以存储具有不同特性的房间的声学特性。存储在数据库中的声学特性可以是基于对具有特定特征的房间进行的预先分析而确定的。用户可以输入房间的特性,并且数据库可以输出环境的声学特性。声学特性还可以存储在回放设备自身或数据库中,以供回放设备使用。

[0032] 题为“Playback Device Calibration”的美国非临时专利No.14/481,511、题为“Audio Settings Based on Environment”的美国非临时专利No.14/216,306、题为“Hybrid Test Tone for Space-Averaged Rom Audio Calibration Using a Moving Microphone”的美国非临时专利No.14/805,140、题为“Multi-Channel Pairing in Media System”的美国非临时专利No.14/825,961、以及题为“System and Method for Device Playback Calibration”的美国专利No.9,106,192也描述了用于校准回放设备的各种示例,其内容均通过引用整体并入本文。

[0033] NMD通常可以通过麦克风响应来处理语音输入。麦克风响应是麦克风关于频率的灵敏度的指示。在实施例中,NMD还可以使用环境的声学特性来处理语音输入以校正语音输入中的失真。这样可以更好地解释语音输入。

[0034] 回放设备可以向NMD提供环境的声学特性。由于媒体回放系统可以具有多个回放设备,所以提供声学特性的回放设备可以是最靠近NMD的回放设备。作为另一示例,提供声学特性的回放设备可以是在同一区域的回放设备。在一些情况下,回放设备可以与NMD绑定(分组或配对)。绑定可以表示回放设备和NMD彼此非常靠近,例如,在同一房间或同一区域中,和/或同步播放内容。当NMD与回放设备绑定时,NMD可以被布置为从绑定的回放设备接收环境的声学特性。此外,如果NMD被移动并且绑定到另一回放设备,则其所绑定到的该另一回放设备可以提供NMD所处环境的新的声学特性。

[0035] NMD可以使用其所处环境的声学特性来改进其操作。例如,NMD可以基于声学特性来定义滤波器。NMD可以将滤波器应用于网络麦克风设备接收的语音输入,以校正环境的声学特性,例如,语音输入中的频谱、空间和时间失真。这样可以更好地解释语音输入。

[0036] 在许多情况下,NMD可以放置在媒体回放系统附近。媒体回放系统可以包括回放音频内容的多个回放设备。这些回放设备可以分布在家中的各个位置,并且这些位置可以被称为区域,例如家中的卧室或起居室。由此,NMD可能需要在存在也是由媒体回放系统回放的音频内容(如音乐)时解释语音输入。

[0037] 在一些情况下,回放设备可以向NMD提供对正在回放的内容的指示,使得NMD不仅可以将滤波器应用于语音输入以校正环境的声学特性,还可以使用正在回放的音频内容以更好地分隔语音输入。

[0038] 附加地或备选地,NMD可以能够确定语音输入的方向性。方向性可以定义语音输入

来自的方向。该方向性可被提供至媒体回放系统。媒体回放系统包括多个扬声器。媒体回放系统可以使用该方向性来调整多个扬声器的音频输出。例如,媒体回放系统也可以将一个或多个回放设备产生的音频声音引导至语音输入来自的方向。这可以是收听者所在的位置。其他布置也是可能的。

[0039] 在另一方面中,本文描述的一些示例涉及通过语音输入控制媒体回放系统。本文提供的一些示例可以涉及一种方法。该方法可以包括接收包括命令字和一个或多个媒体变量实例的语音输入,以及确定语音输入对应于音乐控制。该方法还可以包括确定对应于命令字的命令,以及处理语音输入以识别对应于意图的媒体回放系统命令。可以处理语音输入以识别与一个或多个音乐变量实例相关的媒体内容,并且可以确定媒体内容的一个或多个媒体项。可以使媒体回放系统对一个或多个媒体项执行命令。

[0040] 在其他示例中,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。这些功能包括接收包括命令字和一个或多个媒体变量实例的语音输入,以及确定语音输入对应于音乐控制。这些功能还可以包括确定对应于命令字的命令,以及处理语音输入以识别对应于意图的媒体回放系统命令。可以处理语音输入以识别与一个或多个音乐变量实例相关的媒体内容,并且可以确定媒体内容的一个或多个媒体项。可以使媒体回放系统对一个或多个媒体项执行命令。

[0041] 根据另一些示例,提供了一种系统。该系统包括处理器和存储器。存储器上存储有可由装置执行以使系统执行功能的指令。这些功能包括接收包括命令字和一个或多个媒体变量实例的语音输入,以及确定语音输入对应于音乐控制。这些功能还可以包括确定对应于命令字的命令,以及处理语音输入以识别对应于意图的媒体回放系统命令。可以处理语音输入以识别与一个或多个音乐变量实例相关的媒体内容,并且可以确定媒体内容的一个或多个媒体项。可以使媒体回放系统对一个或多个媒体项执行命令。

[0042] 在又一方面中,本文描述的示例包括媒体回放系统(或者其可能的一个或多个组件)接收语音命令,并基于用户识别确定媒体回放系统执行的适当动作。

[0043] 收听外放媒体内容可以一种涉及家人、朋友和访客的社交活动。媒体内容可以包括,例如,热线广播、有声读物(books)、来自电视的音频、存储在本地驱动器上的音乐、来自媒体源(例如,Pandora® Radio、Spotify®、Slacker®、Radio、Google Play™、iTunes Radio)的音乐以及其他可听材料。例如,在家中,人们可以在聚会和其他社交聚会上以外放方式播放音乐。在这种环境下,人们可能希望在一个收听区域或同时在多个收听区域中播放音乐,使得每个收听区域中的音乐可以是同步的,而没有可听见的回声或干扰(glitch)。当人们可以使用语音命令来控制音频回放设备或系统时,可以进一步丰富这种体验。例如,人们可能希望改变音频内容、播放列表或收听区域,将音乐曲目添加到播放列表或回放队列,或者改变回放设置(例如,播放、暂停、下一曲目、前一曲目、回放音量和EQ设置等)。

[0044] 收听外放媒体内容也可以是个人体验。例如,个人可以在工作前的早晨、锻炼时、晚餐时、或者白天在家中或工作时的其他时间为自己以外放方式播放音乐。针对这些个人体验,个人可以选择将音频内容的回放限制在单个收听区域或收听处。当个人可以使用语音命令来选择收听区域、音频内容、回放设置及其他设置时,可以进一步丰富这种体验。

[0045] 对尝试执行语音命令的人进行识别也可以是体验的重要元素。可能希望基于这个

人是谁以及这个人希望媒体回放设备或系统做什么来执行语音命令。作为说明,在家中中进行聚会或社交聚会时,主人或家庭所有者可能希望阻止某些访客使用语音命令来改变音频内容、收听区域或回放设置。在一些情况下,主人或家庭所有者可能希望允许某些访客使用语音命令来改变音频内容、收听区域或回放设置,同时阻止其他访客进行此类改变。基于用户简档或语音配置设置的用户识别可以帮助区分家庭所有者的语音和访客的语音。

[0046] 在另一示例中,用户识别可以用于区分成人的语音和儿童的语音。在一些情况下,家庭所有者可能希望阻止儿童使用语音命令来收听不适合儿童的音频内容。在其他情况下,家庭所有者可能希望阻止儿童改变收听区域或回放设置。例如,家庭所有者可能希望以特定音量收听音频内容并且阻止儿童改变音频内容的音量。用户识别可以帮助设置家长控制设置或约束设置,以阻止儿童访问特定内容或者改变收听区域或回放设置。例如,基于用户简档或语音配置设置的用户识别可以帮助确定这个儿童是谁,允许这个儿童听什么,或者允许这个儿童改变什么设置。

[0047] 在又一示例中,用户识别可以用于防止无意的语音命令。例如,家庭所有者可能希望防止来自电视的音频或任何其他音频内容无意中触发语音命令。本文描述了与上述示例类似和不同的许多其他示例,并且示出了基于语音识别的不同类型的动作。

[0048] 本文描述的一些实施例包括媒体回放系统(或者其可能的一个或多个组件)接收语音命令,并基于用户识别确定媒体回放系统执行的适当动作。

[0049] 在示例配置中,媒体回放系统包括单独的或与诸如媒体回放系统服务器之类的计算设备组合的一个或多个媒体回放设备。在另一示例配置中,媒体回放系统可以包括联网麦克风系统服务器和一个或多个NMD或与之通信。在又一示例配置中,媒体回放系统服务器和/或联网麦克风系统服务器可以是基于云的服务器系统。这些设备和/或服务器中的任一个或组合可以接收用于媒体回放系统的语音命令。

[0050] 在示例操作中,一个或多个功能可以由联网麦克风系统单独地或与媒体回放系统组合地执行。例如,接收语音命令可以包括联网麦克风系统通过一个或多个NMD接收语音命令,以及向媒体回放系统发送语音命令以进行进一步处理。然后,媒体回放系统可以将语音命令转换为等效的文本命令,并解析文本命令以识别命令。在另一情况下,联网麦克风系统可将语音命令转换为等效的文本命令,并向媒体回放系统发送文本命令以解析文本命令和识别命令。

[0051] 语音命令可以是控制本文讨论的任何媒体回放系统控制的命令。例如,在一些实施例中,语音命令可以是媒体回放系统通过该媒体回放系统的一个或多个回放设备来播放媒体内容的命令。在一些实施例中,语音命令可以是修改媒体回放系统的一个或多个媒体回放设备的回放设置的命令。回放设置可以包括例如回放音量、回放传输控制、音乐源选择、分组等。

[0052] 在接收到语音命令之后,媒体回放系统的计算设备可以确定是否是从媒体回放系统的注册用户接收到语音命令。例如,媒体回放系统可以注册到家中的一个或多个用户或特定用户。媒体回放系统的计算设备可以被配置为基于存储在计算设备中的用户简档将语音命令与注册用户相关联。注册用户可以具有创建并存储在计算设备中的用户简档。用户简档可以包含特定于用户的信息。例如,用户简档可以包含关于用户的年龄、位置、偏好回放设置、偏好播放列表、偏好音频内容、对该用户设置的访问约束的信息以及识别用户语音

的信息等。

[0053] 媒体回放系统的计算设备可以被配置为基于用户设置的语音配置设置将语音命令与用户相关联。例如,媒体回放系统可以要求用户提供一个语音输入或一系列语音输入。然后,媒体回放系统的计算设备可以处理语音输入,将语音输入与用户相关联,并存储信息,使得媒体回放系统可以识别来自用户的语音命令。

[0054] 在一些示例中,媒体回放系统的计算设备可以被配置为确定与语音命令相关联的置信度水平,这可以进一步帮助确定是从注册用户接收到语音命令。可以基于用户历史、位置单独地或者与通常在用户简档中找到的任何其他信息组合地,来确定置信度水平。

[0055] 响应于确定是从注册用户接收到语音命令,媒体回放系统的计算设备可以配置用于媒体回放系统的指令。指令可以基于语音命令中的内容和注册用户的用户简档中的信息。附加地或备选地,指令可以基于语音命令中的内容和存储在计算设备上的语音配置设置。

[0056] 如本文的示例中所述,语音命令可包括各种内容。在一个示例中,语音命令中的内容可以包括一个或多个回放设备播放媒体内容的命令。在一些情况下,基于一个或多个回放设备播放媒体内容的命令和注册用户的用户简档中的信息,媒体回放系统的计算设备可以配置指令,使得一个或多个回放设备从注册用户的偏好媒体源(例如,音乐流传输服务)获得媒体内容。然后,计算设备可以配置指令,使得媒体回放系统通过媒体回放系统的一个或多个回放设备播放媒体内容。在其他情况下,基于一个或多个回放设备播放媒体内容的命令和注册用户的用户简档中的信息,计算设备可以包括指令,以便(i) 将一个或多个回放设备配置有注册用户的一个或多个偏好回放设置,和(ii) 使得一个或多个回放设备通过媒体回放系统按照注册用户的偏好回放设置来播放媒体内容。

[0057] 在另一示例中,语音命令中的内容可以包括一个或多个回放设备播放媒体内容的命令,但是可以不识别媒体回放系统的特定收听区域或回放区域。基于语音命令中的内容和注册用户的用户简档中的信息,计算设备可以配置指令,使得一个或多个回放设备通过媒体回放系统的特定回放区域内的一个或多个媒体回放设备来播放媒体内容。

[0058] 在又一示例中,语音命令中的内容可以包括使媒体回放系统修改回放设置的命令。基于语音命令中的内容和注册用户的用户简档中的信息,计算设备可以配置指令,使得媒体回放系统修改媒体回放系统的一个或多个回放设备的回放设置。

[0059] 其他示例可以涉及媒体回放系统确定优先顺序,以解决从不同用户接收的冲突语音命令。例如,媒体回放系统可以分配优先顺序,其中从注册的访客接收的语音命令具有比非注册的访客更高的优先级。例如,冲突语音命令可以包括从用户接收的播放歌曲的语音命令和随后从另一用户接收的停止播放歌曲的语音命令。本文描述了与上述示例类似和不同的许多其他示例。

[0060] 在另一示例中,媒体回放系统可以基于接收到的与注册用户或注册的访客用户相关联的唤醒词或唤醒短语来采取动作。唤醒词或唤醒短语(例如,“嘿Sonos”)可以用于触发时间段,在该时间段内系统将基于接收到的唤醒词接受来自用户的附加命令。例如,主人或授权访客可以发送将歌曲添加到播放队列的语音命令(例如,“嘿Sonos,让我们队列(queue up)歌曲”),这可以开启一个时间段(例如,五分钟),用于主人或授权访客发送将特定歌曲添加到播放队列的附加语音命令。本文描述了与上述示例类似和不同的许多其他示例。

[0061] 在配置了用于媒体回放系统的一个指令或一组指令之后,计算设备的一些实施例可以将一个指令或一组指令发送到媒体回放系统的一个或多个回放设备。

[0062] 在又一示例中,媒体回放系统的计算设备可以确定是否是从儿童接收到语音命令。计算设备可以被配置为基于用户简档或访客简档中的信息来区分成人和儿童。具体地,计算设备可以基于用户语音的音调或频率来区分成人和儿童。

[0063] 响应于确定是从儿童接收到语音命令,可以防止一个或多个回放设备播放可能不适合该儿童的给定媒体内容。在一些情况下,可以防止计算设备和/或一个或多个回放设备基于儿童的语音命令的内容来修改回放设置。

[0064] 在又一示例中,动作可以基于确定是否是从访客用户而不是媒体回放系统的注册用户接收到语音命令。在一个示例中,注册用户可能已经为访客用户创建了访客简档。访客简档可以包括用户简档中包括的任何信息。在另一示例中,媒体回放系统的计算设备可以确定语音命令不是从注册用户接收的,然后可以询问注册用户该语音命令是否来自注册用户的访客。

[0065] 响应于确定是从访客用户接收到语音命令,媒体回放系统的计算设备可以(1)为访客用户分配约束设置,(2)基于语音命令中的内容和为访客用户分配的约束设置来配置用于一个或多个回放设备的指令,以及(3)向一个或多个回放设备发送指令。约束设置可以是限制对媒体回放系统的控制的任何设置。

[0066] 在另一方面,本文描述的示例涉及基于命令来识别和访问合适的流传输服务(例如,流传输音频曲目)。

[0067] 可以基于NMD提供的语音命令来识别和访问流传输服务。例如,NMD可以接收语音命令“播放Pandora®70年代的摇滚乐电台”,确定该语音是播放来自特定流传输服务的特定电台的命令,然后促使对来自该服务的电台的回放。在其他实现中,流传输服务的选择可以基于经由控制器设备的用户界面的命令输入(例如,文本输入)。其它类型的命令也是可能的。

[0068] 在另一示例中,命令可能未明确指定用户期望访问哪个流传输服务。例如,用户可以发出更普通的命令,例如“播放Queen”或“播放70年代摇滚乐”。在这种情况下,现有系统在区分用户期望播放的内容类型方面(即,艺术家Queen还是专辑Queen、70年代摇滚乐电台还是70年代摇滚乐的“最佳”专辑)的能力有限。另外,这种系统不能将用户想要的内容类型与各种流传输服务的内容类型能力进行匹配(即,哪些服务能够播放广播电台、艺术家、专辑)。

[0069] 鉴于不断增加的内容量和可用流传输服务数量,需要一种系统,其能够智能地基于用户命令选择与提供给用户的内容类型相匹配的所期望的流传输服务。

[0070] 在示例中,可以由包括多个支持网络的设备的网络配置来完成对流传输服务的选择。网络配置可以包括接收、处理和分析命令的NMD、回放设备、计算设备和/或控制器设备(例如,平板电脑、智能电话)。该配置还可以包括基于对接收到的命令的处理和分析从一个或多个音乐内容服务器检索和/或请求音频内容。然后,NMD、控制器设备和/或任何数量的回放设备可以获得音频内容,以基于命令提供音频回放体验。网络配置也可以采用其他形式。

[0071] 在另一示例中,选择可以基于单独或组合的多个标准。在一种情况下,流传输服务

的选择可以取决于命令所指示的内容类型(例如,歌曲、流派(genre)、无线电台)以及特定流传输服务是否支持所指示的内容类型。在这种情况下,可以利用内容类型逻辑来将命令或命令的部分与内容类型相关联。然后,可以将通过逻辑识别的内容类型映射到具有可用内容类型的流传输服务。可以以各种其他方式执行对通过命令指示的内容类型的分析。

[0072] 在又一示例中,流传输服务的选择可以部分地基于用户历史,这可以考虑用户的流传输服务偏好。用户偏好可以基于每个区域、最常播放的内容类型等。另外,可以包含各种形式的“外部”数据,包括但不限于地理数据、人口统计数据 and 天气类型数据。可能存在其他类型的选择影响标准。

[0073] 除了选择流传输服务之外,对用户命令的处理可以导致输出替代的指示。在这样的示例中,系统可以输出能够播放命令所指示的内容类型的流传输服务的建议。在另一示例中,系统可以输出“内容不可用”的指示。可以通过网络麦克风设备或控制器在NMD、控制器处或在一个或多个回放设备处输出这样的指示。

[0074] 各种选择标准可以用作算法的输入,以确定各种流传输服务的置信度指标(metric)。置信度指标可以是特定流传输服务是否是用户可能希望收听的流传输服务的指示。例如,置信度水平可以是分配给流传输服务的概率值或百分比(例如,1-100)。在一个示例中,可以提供具有最高置信度指标的流传输服务以用于流传输。在另一情况下,如果所计算的最高置信度指标未超过阈值置信度值或者前N个置信度水平在彼此的指定范围内,则可以触发错误状态。在这种情况下,错误状态可以使得网络麦克风设备(1)输出“内容不可用”的指示,(2)要求用户重复/进一步指定命令,(3)使得音频由优选的合作伙伴或默认服务来播放等。置信度指标可以以多种其他方式使用。

[0075] 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。

[0076] II. 示例操作环境

[0077] 图1示出了媒体回放系统100的示例配置,在媒体回放系统100中可以实施或实现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、餐厅和起居室)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。

[0078] 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如何交互,以便向用户提供媒体体验的进一步讨论。虽然本文的讨论会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如图1所示的家居环境等等内的应用。例如,本文描述的技术可以在可能需要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。

[0079] a. 示例回放设备

[0080] 图2示出了示例回放设备200的功能框图,该示例回放设备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。

[0081] 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,功能可以包括回放设备200从音频源或另一回放设备检索音频数据。在另一示例中,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,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。

[0082] 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对音频内容的同步回放。在同步回放期间,收听者将优选地不能够感知回放设备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。通过引用合并于此的题为“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,更详细地提供了回放设备之间的音频回放同步的一些示例。

[0083] 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所属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)可能与之相关联的回放队列。数据可以存储为一个或多个状态变量,该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,使得一个或多个设备具有与系统相关联的最新数据。其他实施例也是可能的。

[0084] 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一个或多个扬声器212的电平的设备。扬声器212可包括单独的换能器(例如,“驱动器”)或具有一个或多个驱动器的包括壳体的完整扬声器系统。例如,扬声器212的特定驱动器可以包括例如低音炮(例如,针对低频)、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的每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。

[0085] 可以例如通过音频线路输入连接(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。

[0086] 网络接口214可以被配置为促进数据网络上回放设备200与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、在局域网内从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的目的地址的数字分组数据的形式来发送。在这种情况下，网络接口214可以被配置为解析数字分组数据，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。

[0087] 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，数据网络内与回放设备200相关联的其他回放设备、扬声器、接收器、网络设备、控制设备）进行无线通信。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进行通信。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，但是在一些实施例中，网络接口214可以仅包括无线接口或仅包括有线接口。

[0088] 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壳体的外壁上。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220的部分频段敏感。一个或多个扬声器220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。

[0089] 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。

[0090] 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用于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，低音炮），则回放设备200可以与被设计为呈现全频率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，全频率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。

[0091] 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”。任意其他过去、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。

[0092] b. 示例回放区域配置

[0093] 返回参考图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被赋予名称,例如书房、浴室、主卧室、卧室、厨房、餐厅、起居室和/或阳台。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。

[0094] 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,而起居室和主卧室区域均具有多个回放设备。在起居室区域中,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。

[0095] 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回放设备102播放的嘻哈音乐,而另一用户可以在厨房区域中准备食物并收听回放设备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,使得用户在不同回放区域之间移动时可以无缝地(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方式来实现回放区域之间的同步。

[0096] 如上所述,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以适应这些改变。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。

[0097] 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114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,如果用户希望在起居室空间中听音乐,而另一用户希望看电视,则起居室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。

#### [0098] c. 示例控制设备

[0099] 图3示出了示例控制设备300的功能框图,该示例控制设备300可以被配置成媒体回放系统100的控制设备126和128中的一个或两者。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、平板电脑或网络设备(例如,联网计算机,如PC或Mac™)。

[0100] 处理器302可以被配置为执行与有助于用户对媒体回放系统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。

[0101] 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经由网络接口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。

[0102] 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。如上所述,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:在区域中增加或从中移除一个或多个回放设备、在区域组中增加或从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从绑定的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是专用控制器还是安装有媒体回放系统控制器应用软件的网络设备。

[0103] 控制设备300可以包括麦克风310。麦克风310可以被布置成检测控制设备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。麦克风可以对部分频段敏感。两个或更多个麦克风310可以被布置为捕捉音频源(例如,语音、可听声音)的位置信息和/或帮助滤除背景噪声。

[0104] 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。所示的用户界面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不同格式、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可比较的控制访问。

[0105] 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,其用于使得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/退出交叉播放模式(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。

[0106] 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,以便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。

[0107] 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,以便调出用于选择媒体回放系统中的、将与特定区域分在一组的一个或多个其他区域的选项。在分组之后,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,以调出用于取消选择区域组中的要从该区域组中移除的一个或多个区域的选项。通过诸如用户界面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。

[0108] 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标题、艺术家姓名、专辑名称、专辑年份、曲目长度以及当经由用户界面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。

[0109] 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,该回放队列包含对应于由该回放区域或区域组回放的零个或多个音频项的信息。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/或检索音频项,可能供回放设备回放。

[0110] 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。在另一示例中,当回放区域或区域组正在持续播放流传输音频内容(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/或其他流传输音频内容项,并且当回放区域或区域组正在播放这些内容项时处于“使用中”。其他示例也是可能的。

[0111] 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,或者包含来自在所建立的区域组被取消分组之前与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,或者包含来自在所建立的区域组被取消分组之前与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。

[0112] 返回参考图4的用户界面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。

[0113] 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检索音频内容,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。

#### [0114] d. 示例音频内容源

[0115] 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检索回放音频内容(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)中检索音频内容。在另一示例中,可以通过网络,经由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。

[0116] 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))上的本地音乐库、通过互联网(例如,云)提供音频内容的流传输音频服务、或者回放设备或网络设备上通过线路输入连接来连接到媒体回放系统的音频源等。

[0117] 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:扫描通过媒体回放系统中的回放设备可访问的网络共享的所有文件夹/目录中的可识别音频项,以及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。

[0118] 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。

#### [0119] e. 示例多个联网设备

[0120] 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。

[0121] 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。

[0122] 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可以是用于流传输音乐服务的音频内容源。

[0123] 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。

[0124] 在一个示例中,NMD 512、514和516中的一个或多个可以是被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(utility)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或网络设备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或网络设备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。

[0125] 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的网络设备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的回放设备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。

[0126] 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,如上所述,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。

[0127] 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。

[0128] 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进行通信。

[0129] 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。

[0130] 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TM (IEEE 802.15)、NFC、无线直连和/或专有无线等中的一个或多个。

[0131] 在一个示例中,CR 522可以通过蓝牙TM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。

[0132] 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首先被引入具有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发送对应于NMD 516的标识信息,并且作为响应,PBD 538可以通过NFC (或一些其他形式的通信) 向NMD 516发送局域网信息。然而,在家中配置了NMD 516之后,NMD 516和PBD 538之间的通信方式可能会改变。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进行通信。在另一示例中,NMD和PBD可以主要通过本地通信方式546进行通信。其他示例也是可能的。

[0133] 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。在一种情况下,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。

[0134] 例如,如果文本命令是“在区域1中播放来自流传输服务1的艺术家1的曲目1”,则计算设备506可以识别(i) 从流传输服务1可得的艺术家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,来自流传输服务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中的一个或两个时,计算设备506可以通过通信路径544向PBD 536和538中的一个或两个发送所识别的URL,用于回放。作为响应,PBD 536和538中的一个或两个可以根据接收到的URL从计算设备508中检索音频内容,并且开始播放来自流传输服务1的艺术家1的曲目1。

[0135] 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。

[0136] 本领域普通技术人员将理解,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。

[0137] f. 示例网络麦克风设备

[0138] 图6示出了示例网络麦克风设备600的功能框图,该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理解,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。

[0139] 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,处理单元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执行以实现上述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)。

[0140] 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相对于网络麦克风设备来自一个或多个方向的音频。麦克风阵列606可以对部分频段敏感。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还可以被布置为捕捉音频源(例如,语音、可听声音)的位置信息和/或帮助滤除背景噪声。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。

[0141] 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等)、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)。

[0142] 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608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,以便向用户提供视觉和/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。

[0143] III. 用于房间校正语音检测的示例系统、装置和方法

[0144] 在许多情况下,网络麦克风设备可以放置在回放设备附近。例如,回放设备可以与网络麦克风设备放置在同一房间中。

[0145] 本文描述的示例涉及向NMD提供NMD操作的环境的声学特性。环境可以是家中的房间,例如卧室或起居室。环境的声学特性可以定义声音在房间中行进的方式。NMD可以使用环境的声学特性来解释对NMD说出的语音输入。在许多情况下,网络麦克风设备(NMD)可以放置在回放设备附近。例如,回放设备可以与NMD放置在同一房间中。

[0146] 声学特性通常由环境中的表面定义。例如,房间中的硬表面可以反射声音。另一方面,软表面可以吸收声音。环境中这些不同类型的表面的存在及其布置将影响房间的声学特性和NMD解释语音输入的能力。

[0147] NMD可能需要考虑这些声学特性以准确地恢复说出的语音输入。另外,在一些情况下,可以在NMD接收语音输入的同时由回放设备同时回放音频内容。NMD可以使用内容回放和/或环境的声学特性的知识来解释语音输入。

[0148] 图7-11给出了可以在所公开的操作环境中实现的实施例。

[0149] 图7是可以执行以改进对NMD接收的语音输入的解释的功能的流程图。在一些示例中,所描述的这些功能中的一个或多个可以由回放设备执行。在其他示例中,诸如504-508之类的计算设备还可以结合回放设备或代替回放设备来执行这些功能中的一个或多个。计算设备可以与回放设备相关联,并执行与回放设备相关联的处理。

[0150] 在702处,可以获得环境的声学特性。声学特性可以与回放设备操作的环境相关联。已经以各种方式确定了声学特性。例如,回放系统可以已经通过某个校准阶段确定了环境的声学特性。媒体回放系统的回放设备可以具有麦克风和扬声器。扬声器可以输出一个或多个音调,并且一个或多个扬声器可以使用相应的麦克风接收音调。音调可以是音频的一个或多个频率。每个回放设备可以输出音调。基于多个回放设备接收的音调,可以确定环境的声学特性。在另一示例中,另一网络设备(如控制设备、第一回放设备或者甚至NMD自身)可以包括接收第二回放设备的扬声器输出的测试音调的麦克风。其他网络设备(替代回放设备或作为回放设备的补充)可以分析接收到的测试音调以确定环境的声学特性。其他布置也是可能的。

[0151] 声学特性可以由音调的频谱响应、空间响应和时间响应来定义。频谱响应可以是对麦克风处接收的声能的分析。空间响应可以是对麦克风处接收的声能的方向的分析。时间响应可以是对麦克风处接收的声能的回响的分析。回放系统可以分析这些响应,并且还可以考虑接收音调的方向以确定环境的声学特性。声学特性特征的指示可以存储在回放设备和/或计算设备504-508中的一个或多个上。

[0152] 在另一示例中,可以基于环境的已知特征(如房间的大小、房间天花板的高度和房间中的家具)来预定义声学特性。计算设备504-508中的一个或多个维护的数据库可以存储

具有不同类型特征的房间的声学特性。存储在计算设备上的声学特性可以是基于对具有特定特征的房间进行的预先分析而确定的。用户可以在媒体回放系统的控制器设备上输入房间的特征,并且控制器设备可以访问该数据库以确定房间的声学特性。然后可以将这些声学特性提供给位于环境中的回放设备或存储在计算设备上。

[0153] 作为示例,声学特性可以指示房间的左侧具有硬表面、右侧具有软表面并且房间是矩形形状。从本质上讲,声学特性可以从声学角度描绘房间的特征。

[0154] 媒体回放系统可以包括多个回放设备。每个回放设备可以具有回放设备操作的环境的声学特性。在704处,可以识别一个或多个NMD。一个或多个NMD可以在同一环境中。在一些示例中,向NMD发送声学特性的回放设备可以是最靠近NMD的回放设备。在其他示例中,向NMD发送声学特性的回放设备可以是与NMD在同一区域中的回放设备。可以在回放设备和/或NMD的设置期间的一些校准过程中指示接近度。例如,NMD可以向回放设备发送其存在的指示。该存在可以由回放设备中的状态变量来指示。回放设备可以访问该状态变量以识别NMD。类似地,NMD可以具有识别回放设备的存在的类似状态变量。

[0155] 在其他示例中,媒体回放系统的回放设备可以与NMD绑定(配对或分组)。绑定也可以指示回放设备和NMD彼此非常靠近,例如在同一房间或同一区域中,或同步播放音频内容。回放设备和NMD之间的绑定在某些方面可以类似于回放设备和回访设备的绑定方式。该绑定可以反映在回放设备存储的状态变量中。当NMD与回放设备绑定时,回放设备可以被布置为识别NMD。此外,如果NMD被移动并绑定到另一回放设备,则可以更新另一回放设备中的状态变量以反映与NMD的绑定。类似地,NMD可以更新其状态变量以反映存在新的回放设备。

[0156] 在一些示例中,绑定的回放设备和NMD可以使用各种联网技术来减少它们之间的时延。例如,可以使用本地网络连接而不是WAN网络连接(LAN或蓝牙)进行通信。作为另一示例,当回放设备绑定到NMD时,通信可以回到本地网络中的不同频段,例如,从“典型的”2.4Ghz通信切换到5Ghz通信。作为又一示例,通信可以切换到2.4或5Ghz频谱上的保留通道,或者当回放设备和NMD绑定时可以减少其他网络流量,例如,回放设备可以停止用作其他网络流量的中继节点。其他布置也是可能的。

[0157] 在706处,可以提供声学特性,使得将声学特性应用于语音输入。在一个实施例中,声学特性可以作为消息通过通信网络546中的一个或多个从回放设备发送到NMD。在其他实施例中,回放设备可以使计算设备上的声学特性通过通信链路542被发送到NMD。在其他实施例中,声学特性可以在与回放设备相关联的计算设备上,并且回放设备可以使与NMD相关联的计算设备能够访问声学特性。其他布置也是可能的。

[0158] 在一些实施例中,回放设备可能不需要向NMD提供声学特性。相反,NMD自身可能能够从计算设备检索声学特性。NMD可以提供其接近(绑定、配对或其他方式)的回放设备的指示,并且计算设备可以提供环境的声学特性。其他布置也是可能的。

[0159] NMD通常可以通过麦克风响应来处理语音输入。麦克风响应是麦克风关于频率的灵敏度的指示。

[0160] 在实施例中,NMD可以对其接收的语音输入应用声学特性,以校正语音输入中的失真。这样,NMD可以能够更好地解释语音输入。NMD可以由自身来应用此声学特性,和/或将处理卸载至计算设备,在这种情况下,NMD上的处理可以是基于云的。

[0161] NMD可以在对语音输入应用声学特性时基于声学特性来定义滤波器。滤波器可以

包括环境的频谱响应、空间响应和时间响应。NMD可以在解释语音输入之前将滤波器应用于NMD接收到的语音输入,以校正失真。可以基于以下推导来确定滤波器:

$$[0162] \quad X_a(w) \times P(w) \times h(w) \times m(w) = Y_a(w)$$

[0163] 其中, $X_a$ 是校准音调, $P$ 是回放设备的扬声器响应, $h$ 是房间响应(例如,房间的声学特性), $m$ 是麦克风响应, $Y_a$ 是处理后的响应,其可以是与校准过程中发送的音调对应的接收到的音调。符号 $\times$ 表示频域中的卷积函数。由于 $X_a$ 、 $P$ 、 $m$ 和 $Y_a$ 是已知的,房间响应(例如,滤波器)可以被计算为:

$$[0164] \quad h^{-1}(w) : X_a(w) \times P(w) \times m(w) = Y_a(w) \times h^{-1}(w)$$

[0165] 然后,如果接收到语音输入 $Y_b$ ,则可以应用房间响应 $h$ (例如,滤波器),将语音输入 $X_b$ 确定为:

$$[0166] \quad X_b(w) = Y_b(w) \times h^{-1}(w) \times P^{-1}(w) \times m^{-1}(w)$$

[0167] 在一些实施例中,回放系统可以周期性地确定环境的声学特性以考虑环境的变化。在这种情况下,一个或多个回放设备可以周期性地执行操作702-706,使得NMD可以将当前声学特性应用于接收到的语音输入,以正确地解释它。

[0168] 图8是可以执行以改进对NMD的语音输入的恢复的功能的另一流程图。在一些示例中,所描述的这些功能中的一个或多个可以由回放设备执行。在其他示例中,诸如504-508之类的计算设备还可以结合回放设备或代替回放设备来执行这些功能中的一个或多个。计算设备可以与回放设备相关联,并执行与回放设备相关联的处理。

[0169] 在802处,可以获得环境的声学特性;在804处,可以识别NMD;并且在806处,可以提供声学特性。声学特性可以被提供给NMD和/或与NMD相关联的计算设备。

[0170] 在一些情况下,NMD可以在媒体回放系统也在回放音频内容的环境中操作。NMD可以在正在播放音频内容的同时接收语音输入。

[0171] 在808处,可以提供音频内容。音频内容可以被提供给NMD和/或与NMD相关联的计算设备。例如,内容可以是回放设备正在播放的音乐流。回放设备可以将音频内容提供给该NMD附近、与该NMD在同一区域中的NMD,或者提供给可以绑定至该NMD(或分在一组)的NMD。在一些实施例中,可以通过计算设备将内容提供给NMD,例如,当NMD的处理基于云时。在其他实施例中,例如,当NMD和/或回放设备的处理基于云时,可以通过与回放设备相关联的计算设备将内容提供给与NMD相关联的计算设备。

[0172] 在808处,可以将声学特性(以及可选地音频内容)应用于语音输入。例如,NMD(或与NMD相关联的计算设备)可以将滤波器应用于NMD接收到的语音输入以解释语音输入。例如,NMD(或与NMD相关联的计算设备)可以使用与语音输入一起正在回放的音频内容,以更好地分离语音输入。为了解释语音输入的目的,回放设备正在播放的音频内容实际上可能是噪声。在这一方面,可以从接收到的语音输入中减去音频内容,以便更好地分离接收到的语音输入。

[0173] 图9是NMD和/或相关联的计算设备在解释语音输入时执行的功能的流程图。在902处,可以获得环境的声学特性的指示。在904处,可选地,还可以接收回放设备正在播放的音频内容。在存在NMD或者绑定到NMD的情况下,回放设备可以在同一区域中。在906处,例如,NMD可以接收语音输入。语音输入可以是语音命令。通过NMD上的某个按钮按压或语音输入中指示其为语音输入的命令字,NMD可以知道它已经接收到语音输入。在908处,可以应用/

使用滤波器和/或回放设备正在播放的音频内容中的一个或多个来解释接收到的语音输入。

[0174] 图10是例如使NMD能够确定环境的声学特性,而不是由回放设备提供该声学特性的功能的流程图。这些功能可以由回放设备和/或与回放设备相关联的计算设备来执行。

[0175] NMD可以发送用于一个或多个回放设备播放音调的指示。NMD向其发送指示的回放设备可以包括NMD附近的那些回放设备和/或绑定到NMD(或分在一组)或在同一区域中的那些回放设备。

[0176] 在1002处,可以接收指示以播放音频音调。可以从NMD或与NMD相关联的计算设备接收指示。作为响应,在1004处,回放设备可以输出音频音调。NMD可以使用麦克风阵列606接收音调。麦克风阵列606可以提供音调幅度的指示。另外,麦克风阵列606可以提供音调方向的指示。可以基于从多个回放设备接收音调来确定方向。基于从多个回放设备接收音调,NMD和/或相关联的计算设备然后可以确定环境的声学特性。这样,NMD可能不需要从回放设备获得该声学特性。

[0177] 麦克风阵列606可以使NMD能够确定语音输入来自的方向。该方向可以被回放设备用于改善音频内容的回放。例如,媒体回放系统也可以将一个或多个回放设备产生的音频声音引导至语音输入所来自的相同方向。这可以是收听者所在的位置。其他布置也是可能的。

[0178] 图11是回放设备和/或相关联的计算设备执行的、与使用通过NMD确定的方向性相关联的功能的流程图。

[0179] 在1102处,可以接收方向的指示。其可以从NMD和/或相关联的计算设备接收的。方向可以指示NMD接收语音输入的位置以及因此用户在环境中可能所处的位置。回放设备可以具有多个扬声器,扬声器的输出可以被控制以影响回放设备回放的音频内容的方向性。在1104处,回放设备可以使用该方向性来调整多个扬声器的音频输出。音频输出可以是音乐内容的回放。回放设备可以根据NMD指示的方向性来调整扬声器输出的音频信号的相位,以产生音频输出的方向性。在这一方面,音频内容可以指向用户,因此改善了收听音频体验。附加地或备选地,回放设备还可以调整音频输出的幅度(例如,音量)。在一些情况下,NMD也可以产生音频输出。音频输出可以是对语音输入的语音响应,或某种类型的其他声音,例如哔哔声或音调。再附加地或备选地,回放系统可以使用方向性来识别可能适合于替代NMD或作为补充而输出该音频输出的回放设备。例如,所识别的回放设备可能位于收听者的正前方,使收听者更容易听到音频输出。其他布置也是可能的。

[0180] IV. 用于媒体回放系统的语音控制的示例系统、装置和方法

[0181] 本文描述的示例可以涉及控制媒体回放系统。具体地,图12示出了用于播放音频响应的方法的示例流程图1200。方法1200呈现了可以在操作环境内实现的方法的实施例,该操作环境包括,例如,图1的媒体回放系统100、图2的回放设备200中的一个或多个、图3的控制设备300中的一个或多个、图5的系统500中的多个设备中的一个或多个。方法1200可以包括,如本文所述,通过各种通信路径在联网麦克风系统和媒体回放系统之间发送和接收信息,和/或如2016年2月22日提交的题为“Metadata exchange involving a networked playback system and a networked microphone system”的专利申请No.62/298,350所述,使用元数据交换通道。该申请由此通过引用全部合并于此。方法1200可以包括一个或多

个操作、功能或动作。

[0182] 在一些示例中，NMD 512、514、516中的一个或多个可以连接到一个或多个网络或通信路径542、546。NMD可以被配置为与媒体回放系统(例如，计算设备506、CR522、PBD 532、534、536、538)集成或接口连接。NMD可以包括在联网麦克风系统中(例如，NMD 512、514、516、504)。

[0183] 在框1202处，为了配置NMD以使NMD与媒体回放系统接口连接或关联，用户可以选择例如控制设备(例如，CR 522)来启动与NMD或联网麦克风系统或媒体回放系统相关联的应用。应用可以是用于控制联网麦克风系统和/或媒体回放系统的特征或设置的应用。联网麦克风系统可以具有一个或多个专用附加功能，可用于向联网麦克风系统提供附加或增强的功能。

[0184] 在框1204处，可以显示联网麦克风系统的可用的附加功能以供用户选择。图13A中示出了示例用户界面1300，其显示可用的附加功能。NMD可以具有一个或多个可用的附加特征1302、1304、1306。特征1302可以是音乐控制特征，其可以例如控制媒体回放系统的回放。在框1206处，控制设备可以接收用于将音乐控制特征或技能添加到NMD的用户输入。例如，用户可以通过选择特征1302来选择添加音乐控制特征1302。音乐控制特征1302可以是用户将音乐控制功能或技能添加到联网麦克风系统的应用。在选择了添加音乐控制特征1302时，显示器可以改变以示出用户界面1308，提示需要用户的账户信息。

[0185] 在框1208处，控制设备可以通过用户界面1308和输入字段1310和1312接收用于音乐控制特征的账户信息。账户信息可以是用户具有的媒体回放系统和/或音乐服务账户。账户信息可以包括账户的用户名和密码。可以在字段1310中输入用户名，并且可以在字段1312中输入密码。用户可以选择按钮(例如，提交按钮1314)来提交用户名和密码，用于向媒体回放系统(例如，计算设备506)和/或音乐服务(例如，计算设备508)进行认证。可以通过通信路径542、544、546和/或元数据交换通道中的任何一个来发送账户信息，用于在计算设备504、计算设备506和/或计算设备508处进行认证。一旦被认证，与账户相关联的用户信息可以从媒体回放系统被发送到联网麦克风系统。用户信息可以是自定义名称(例如，自定义区域名、自定义播放列表名、自定义歌曲名、自定义专辑名、自定义艺术家名等)、与用户信息相关联的家庭标识符、PBD标识符、和/或区域标识符。自定义名称可以是用户提供的任何名称。例如，媒体回放系统可以提供公共区域名称的列表，供用户在对区域命名时进行选择，和/或向用户提供键入区域名称的选项。控制设备300的界面可以显示公共区域名称的列表和/或显示字段，用户可以通过由控制设备300接收到的输入(例如，语音、文本)在字段中输入自定义区域名称。自定义名称信息可以被发送到系统500中的任何设备或与之共享。

[0186] 在框1210处，可以为NMD确定一个默认区域(例如，一个或多个回放设备)或多个默认区域(例如，不同区域、包括至少两个区域的组中的至少两个回放设备)。可以如2016年2月22日提交的题为“Default Playback Device(s)”的专利申请No. 62/298,410中所描述的那样确定默认区域或回放设备。该申请由此通过引用全部合并于此。

[0187] 在一些方面，用户可以在初始配置或设置期间指定默认区域。图13C中所示的用户界面1316示出了示例用户界面，其使用户能够选择默认回放区域与NMD相关联。用户界面1316可以显示可用回放区域1318的列表。可用回放区域的显示名称可包括与在框1208处输入的用户账户相关联的自定义名称。例如，可用回放区域1318的列表可以包括“起居室”和

“厨房”以及“Nick的房间”的自定义区域名。所显示的列表可以是可选择的,以使用户能够指定用于响应于NMD或其他控制设备接收到的命令而回放语音响应和/或音乐的默认区域。图13C示出了用户选择厨房作为默认区域的示例。用户可以选择按钮1320(例如,提交按钮)以确认选择默认区域,并提交对默认区域的选择。可以显示图13D中所示的确认屏幕1322,以向用户确认设置过程完成,并且用户可以选择按钮1326以关闭确认屏幕1322。

[0188] 在框1212处,可以从例如媒体回放系统向联网麦克风系统发送用户特定回放信息。用户特定信息可以是自定义回放信息,例如自定义区域名称、自定义播放列表和/或自定义播放列表名称。在一些方面,用户特定信息和/或用户账户可以与家庭标识符(HHI)相关联。可以通过例如元数据交换通道和/或媒体回放系统和联网麦克风系统之间的任何其他通信路径,从媒体回放系统(例如,计算设备506、CR 522、PBD 532、PBD 534、PBD 536和/或PBD 538)向联网麦克风系统(例如,计算设备504、NMD 512、NMD 514和/或NMD 516)发送用户特定回放信息。

[0189] 计算设备504可以将自定义回放信息存储为动态变量,该动态变量可以是能够与不同变量类型动态关联的变量。例如,“Nick的房间”可以是自定义区域名称,并且可以被存储为与区域变量类型相关联的动态变量。作为另一示例,“Nick的收藏”可以是具有用户创建的自定义名称的自定义播放列表。名称“Nick的收藏”可以被存储为动态变量,并与播放列表变量类型相关联。以这种方式,联网麦克风系统可以知道并且可以在语音输入中识别出与用户相关联的自定义信息。

[0190] 动态变量可以存储在表或其他数据结构中,并与不同的变量类型动态关联。例如,每个动态变量可以与一个或多个变量类型相关联。动态变量可以和与用户账户相关联的标识符一起存储。例如,自定义区域名称“Nick的房间”可以与“Nick的房间”区域中的媒体回放系统的区域标识符和/或PBD的标识符相关联。作为另一示例,自定义区域名称“Nick的房间”可以和区域标识符标签一起存储,和/或“Nick的房间”区域中的PBD的标识符可以和“Nick的房间”区域的标签一起存储。动态变量可以连续地、周期性地或非周期性地更新,以包括用户所添加或移除的、或者与用户的账户相关联的新的自定义名称。自定义名称可以是用户提供的任何名称,其可能存在或不存在于数据库中。

[0191] 在一些示例中,每个动态变量可以与标识符一起存储或与之相关联,该标识符可以用于识别媒体回放系统命令中的动态变量。例如,区域名称“Nick的房间”可以与特定于媒体回放系统的区域标识符一起存储,并且当命令请求对“Nick的房间”中的回放设备执行动作时,可以使用媒体回放系统命令提供“Nick的房间”的区域标识符,作为“Nick的房间”的补充或替代。

[0192] 在框1214处,用户可以通过说出可被NMD接收的命令或请求来提供语音输入。联网麦克风系统可以将语音输入从语音转换为文本,并解析单词以确定语音输入的语法。口头命令可以具有特定语法,联网麦克风系统可以将其识别为在音乐回放控制的域或范畴中。例如,用户可以说“在厨房和Nick的房间中播放甲壳虫(The Beatles)”。联网麦克风系统可以将单词“播放”识别为对应于意图的命令或直接识别为与音乐回放相关联的意图,并且可以在框1216处将口头命令识别为在音乐回放控制的域或范畴。

[0193] 在另一示例中,存在或包括一个或多个媒体变量实例和/或一个或多个区域变量实例可以指示命令字“播放”对应于“播放”意图。联网麦克风系统可以通过搜索可包括音乐

元数据的音乐目录和确定例如艺术家名为“甲壳虫”，来确定“甲壳虫”对应于媒体变量实例。联网麦克风系统可以基于与区域变量类型相关联的动态变量和/或公共区域名称来确定“厨房”和/或“Nick的房间”对应于区域名称。命令字“播放”与媒体变量“甲壳虫”和/或“Nick的房间”的组合可以使联网麦克风系统能够确定语音输入对应于播放所请求音乐的音乐控制意图。

[0194] 因为口头命令被识别为在音乐回放控制的域或范畴中，所以联网麦克风系统可以优先考虑与著名艺术家“甲壳虫(The Beatles)”相关的音乐内容的搜索结果，并且不优先考虑或排除与同音字“the beetles”相关的音乐内容的搜索结果，该同音字“the beetles”可能不与任何著名艺术家关联。换句话说，响应于识别出语音输入在音乐域中，用于语音到文本识别的一组单词或词汇可以被改变为特定于该音乐域，其可以包括一般在词典中找不到的单词，和/或在词典中或者是与口头语言不同的语言中的单词。

[0195] 例如，响应于确定语音输入在音乐域中，可以更新用于语音到文本识别的一组单词或词汇，以包括媒体项目的元数据信息(例如，艺术家姓名、曲目名、专辑名、歌曲名)。在一些方面，语音输入可以包括例如自定义名称(例如，区域、播放列表)、媒体变量、和/或艺术家姓名、专辑名和/或曲目名的字母和/或符号的口述。作为另一示例，在处理音乐控制命令之前，可以改变语音到文本转换的所得文本中的单词，以使用媒体特定单词。媒体变量实例的语音到文本转换可以使用在音乐元数据、媒体目录和/或自定义或本地媒体标识符(例如，播放列表名、曲目名、专辑名、艺术家名等)中找到的单词，替代联网麦克风系统所使用的标准单词或作为其补充。类似地，区域变量的语音到文本转换可以使用自定义区域名称。

[0196] 在框1218处，可以处理包含音乐控制命令的语音输入。联网麦克风系统可以具有各种预定义的语法，这些语法可以与用户的意图(例如，播放、暂停、添加到队列、分组、通过控制设备300可进行的一个或多个其他传输控制)相关联。每个意图可以对应于一个或多个媒体回放系统命令，其可以与意图相同或不同，或者与意图相似或不相似。例如，将音乐回放从第一区域移动到第二区域的意图可以对应于将当前正在播放的媒体项和/或回放队列从第一区域移动到第二区域的媒体回放系统命令。在另一示例中，移动音乐的意图可以对应于媒体回放系统将第一区域的回放队列复制到第二区域的回放队列，和/或将第一区域的状态变量复制到第二区域的状态变量的回放队列复制命令。

[0197] 作为又一示例，移动音乐的意图可以对应于两个媒体回放系统命令。这两个命令可以是第二区域与第一区域分在一组，然后从组中移除第一区域，从而实际上将第一区域的状态传送到第二区域。

[0198] 媒体回放系统命令可以包括应用程序接口(API)，其响应于确定意图对应于媒体回放系统命令而被调用。联网麦克风系统和/或媒体回放系统可以具有用户意图和媒体回放系统命令之间的定义的映射或对应关系。

[0199] 在一些示例中，可以对云网络502中的计算设备(例如，计算设备504、计算设备506、计算设备508)中存储的数据执行媒体回放系统命令。例如，将媒体项(例如，曲目、专辑、播放列表)添加到另一播放列表或回放队列的意图可以被添加至存储在云网络502中的播放列表或回放队列。可以响应于存储在云网络502中的播放列表或回放队列的改变来更新存储在PBD 532、534、536、538上的回放队列，使得回放队列的一部分与云网络502中的播放列表或回放队列的一部分或全部相匹配。

[0200] 某些单词、语法和/或短语可以与相同的意图相关联。例如，语音输入中包括命令字“播放”、“收听”或“听”可以对应于让媒体回放系统回放媒体内容的用户意图。每个意图可以具有不同类型的预定义变量或间隙(slot)，其中对变量或间隙执行命令或意图指定的动作。变量或间隙可以位于各种短语的预定义位置或定位处。例如，“播放”命令语法可以具有用户希望回放的媒体的媒体变量，并且还可以具有用户希望相关联的回放设备回放媒体内容的位置或区域的位置或区域变量。在“在厨房中播放甲壳虫”的口头命令的示例中，媒体或音乐变量的实例可以是“甲壳虫”，并且区域变量的实例可以是“厨房”。在识别对应于媒体变量实例和/或区域变量实例的相关对象时，联网麦克风系统和/或媒体回放系统分开地处理媒体变量和区域变量，和/或对它们进行不同的处理。

[0201] 可以与“播放”意图相关联的语法或短语的另一示例可以是语法“让我在[区域变量]中收听[媒体变量]”。这种语法的一个示例可以是“让我在Emily的房间中收听Paul Simon”，其中“让我...收听...”可能与“播放”意图相关联，“Paul Simon”可能是媒体变量的实例，而“Emily的房间”可能是区域变量的实例。联网麦克风系统可以解析语音输入，并确定存储在联网麦克风系统中的哪个语法与语音输入匹配以识别语音输入的意图。语法可以存储在联网麦克风系统中的任何设备中。

[0202] 用于“播放”意图或命令的语法或短语的又一示例可以是语法“我想在[区域变量]中收听[媒体变量]”。单词“收听”或句式“我想...收听...”可能与“播放”意图相关联。“播放”意图的其他语法是可能的。

[0203] 另一示例命令或意图可以与将媒体内容添加到可存储在媒体回放系统中(例如，存储在PBD、计算设备506和/或CR 522中)的队列相关。示例的添加到队列语法可以是“将[媒体变量]添加到[区域变量]中的队列”。与本文描述的其他方面类似，区域变量可以是可选的，并且系统可以基于各种技术或方法来原因命令所应用的区域(例如，使用默认区域，使用最后使用的区域，基于用户存在信息，使用积极播放媒体的区域)。可以将与媒体变量对应的所选择的媒体内容添加到区域中的队列。

[0204] 作为另一示例，命令或意图可以是下一首播放命令，其可以使所选择的媒体内容被添加到队列的顶部，以在区域中接下来播放。该命令的示例语法可以是“下一首播放[媒体变量]”。类似于本文描述的其他方面，区域变量可以是可选的。

[0205] 命令或意图的另一示例可以是移动或传送命令，其可以将当前播放的音乐和/或区域的回放队列从一个区域移动或传送到另一区域。例如，用户可以说出“将音乐移动到[区域变量]”的语音输入，其中命令字“移动”或“传送”可以对应于将回放状态移动到另一区域的意图。

[0206] 本文描述的命令和意图是示例，其他意图或命令是可能的。例如，可用于通过控制设备300控制媒体回放系统(如本文所述)的每种控制可以具有可用于控制该系统的对应意图。例如，控制命令的名称可以对应于意图。如果控制命令涉及一个或多个媒体项，则命令的语法可以包括一个或多个媒体变量。如果控制命令涉及一个或多个区域，则命令的语法可以包括一个或多个区域。与不同意图一起使用的其他变量也是可能的。

[0207] 可用于通过控制设备300控制媒体回放系统的控制的示例可以包括传输控制命令。这些命令或意图可以与当前正在播放的媒体项相关，例如，传输命令(例如，停止、暂停、跳过、倒回、快进、返回、音量等)或者与将当前播放的媒体项保存或添加到另一回放队列或

播放列表相关的命令。用于对正在播放的媒体项采取的意图或动作的语法可以更为简单，并且可以对应于传输控制的名称。例如，用于暂停音乐回放的意图的语音输入可以是语音输入“暂停”。

[0208] 可以在不同的命令语法中使用不同类型的变量。对于媒体变量，媒体变量可以是语法中的变量或间隙，其中用户可以通过语音输入在变量或间隙处指定用户希望收听或播放的媒体内容。媒体变量可以是与音乐相关的各种特征或特性（例如，媒体变量的类型），包括但不限于：专辑名称、艺术家姓名、歌曲名称、播放列表名称、自定义播放列表名称、流派（例如，流行音乐、古典音乐、乡村音乐、摇滚音乐、R&B等）、情绪（例如，浪漫、运动、高效（productive））、音乐节奏（例如，欢快、慢节拍）、电台名称、作曲家姓名、音乐时代（例如，巴洛克、浪漫、古典、20世纪）、时间段（例如，80年代、90年代）、播放列表创建者的姓名、排名（例如，最佳、前40名）和/或其他音乐识别特征。音乐变量可以具有自定义名称，该自定义名称可以由用户账户的用户创建的自定义播放列表名称和/或其他自定义名称（例如，自定义歌曲名称、自定义专辑名称、自定义艺术家姓名）。

[0209] 对于区域变量，区域变量可以是语法中的变量或间隙，其中用户可以通过语音输入在该变量或间隙处指定执行所请求的动作或意图（例如，播放所请求的音乐）的位置或区域。用户可以在语音输入中包括区域实例，也可以不包括。在这种情况下，用户未指定区域，例如，通过简单地说“播放一些甲壳虫”，联网麦克风系统和/或媒体回放系统可以确定在默认区域中和/或基于其他输入（例如，用户存在信息、上下文信息（context information）、位置信息）在其他区域中“播放一些甲壳虫”。区域变量可以包括用户提供的自定义区域名称的动态变量。作为另一示例，自定义区域名称可以是例如“Nick的房间”或“三楼会议室”。

[0210] 在一些示例中，语法可以包括媒体服务或应用或者其他媒体相关服务、产品或应用（例如，媒体回放系统）执行语音输入的媒体服务变量。系统可以识别用于所有媒体相关内容的默认回放系统或区域，或者将不同的回放系统或区域与不同的服务相关联。例如，用户可以说“在卧室里播放Spotify上的Josh Groban”。系统（例如，联网麦克风系统和/或媒体回放系统）可以将“Spotify”识别为媒体服务变量的实例，将“Josh Groban”识别为音乐变量的实例，并将“卧室”识别为区域变量的实例。系统可以在媒体服务Spotify®的媒体目录中搜索与“Josh Groban”相关的媒体内容，如本文所讨论的。

[0211] 一些类型的变量可以是标量，因为标量变量可以形成为包含同一变量的多个实例的向量或集合。标量变量的向量可以具有用户说出的格式或语法：“[第一标量变量]和[第二标量变量]”、“[第一标量变量]、[第二标量变量]和[第三标量变量]”、或者“[第一标量变量][第二标量变量][第三标量变量]”。例如，区域变量可以是标量变量，并且用户可以指定在多个区域“厨房、起居室和餐厅”中播放媒体内容。在一些方面，可以为向量或标量变量的预定义集合命名。例如，名为“楼下”的预定义向量可以与家庭环境中位于楼下的所有区域相关联。在图1所示的示例环境中，“楼下”可以是“起居室”、“餐厅”、“书房”、“主卧室”、“卧室”和“浴室”区域。控制设备300可以示出区域列表，并且用户可以从区域列表中选择这些区域以与名称或标签“楼下”相关联。

[0212] 在一些示例中，媒体变量可以是标量变量。可以在单个命令短语中说出一个以上相同类型的音乐变量，并且可以独立于其他媒体变量实例或者与其他媒体变量实例组合地，针对相应的媒体项来处理每个媒体变量实例。例如，语音输入可以是“让我们听甲壳虫

和海滩男孩 (the Beach Boys) 的音乐”，其可以对应于语法“让我们听 [第一媒体变量] 和 [第二媒体变量] 的音乐”。在一个方面，第一媒体变量实例“甲壳虫”可以独立于“海滩男孩”来进行处理。在独立于“海滩男孩”而处理“甲壳虫”时，可以针对任何与甲壳虫相关的媒体项来处理“甲壳虫”，并且可以针对任何与“海滩男孩”相关的媒体项来处理“海滩男孩”，将参考框1218更详细地描述。在另一方面，第一媒体变量实例“甲壳虫”和第二媒体变量实例“海滩男孩”可以被处理为一种组合，其中系统可以处理与“甲壳虫”和“海滩男孩”两者相关的媒体项 (例如，甲壳虫和海滩男孩表演和/或被列为艺术家的歌曲)。

[0213] 联网麦克风系统和/或媒体回放系统可以将向量或区域变量实例集合与分组命令相关联。例如，联网麦克风系统和/或媒体回放系统可以识别房间变量的标量变量语法，并确定该语法对应于对区域变量的指定实例进行分组的媒体回放系统命令。作为示例，用户可以指定在“起居室、厨房和Nick的房间”中播放媒体内容。系统500 (例如，联网麦克风系统和/或媒体回放系统) 可以基于在短语中说出“Nick的房间”的位置或地方，将“Nick的房间”识别为区域变量，并在存储动态变量的表中搜索对应于“Nick的房间”的自定义区域名称。搜索可以由计算设备504和/或计算设备506执行，并且可以基于对表的搜索，在计算设备504和计算设备506之间传输自定义区域名称和/或标识符。可以通过元数据交换通道和/或计算设备504和计算设备506之间的任何其他通信路径来进行传输。

[0214] 可被媒体回放系统识别为分组命令的语法的另一示例可以是“添加 [区域变量]”或“也在 [区域变量] 中播放”。例如，联网麦克风系统可以将单词“添加”或短语“也在.....中播放”识别为对应于分组意图或命令。系统可以仅基于在语法的命令字位置中包括“添加”或“也在.....中播放”的命令字的语音输入，或者基于包括命令字和区域变量实例的语音输入，来确定意图。这种类型的分组可以基于上下文，例如，哪些区域中已经正在播放音乐的上下文。媒体回放系统可以通过任何通信路径和/或元数据交换通道向联网麦克风系统发送该信息。在一些方面，媒体回放系统可以接收信息，并且理解也在指定区域中播放当前正在一个或多个区域中播放的音乐，和/或将指定区域包括在现有的组中。

[0215] 在一些示例中，音乐变量可以是标量变量。例如，用户可以指定“播放后街男孩 (Backstreet Boys) 90年代的音乐”。“后街男孩”可以是艺术家的姓名，“90年代”可以是音乐的时间段。另一个示例可以是“播放麦当娜 (Madonna) 的 American Pie”，其中“American Pie”可以是曲目名称，而“麦当娜”可以是艺术家的姓名。

[0216] 可以通过在音乐数据库中搜索与特定音乐变量相关的信息来处理音乐变量。数据库可以是计算设备504、计算设备506和/或计算设备508处的音乐数据库。例如，用户可以说“播放 American Pie”。联网麦克风系统可以在计算设备504中搜索与“American Pie”相关的任何音乐信息。计算设备504可以返回如下结果，例如，具有与该音乐变量相同名称的专辑的艺术家、与该音乐变量匹配或类似的专辑名称、名称为该音乐变量的曲目、具有该音乐变量的广播电台、名称为该音乐变量的播放列表、与该音乐变量相关的内容的流传输服务提供商标识符、和/或原始语音到文本转换结果。使用“American Pie”的示例，则搜索结果可能会返回艺术家“Don McLean”、名为“American Pie”的专辑、名为“American Pie”的曲目、名为“American Pie”的广播电台 (例如，针对“American Pie”的 Pandora 广播电台的标识符)、针对曲目“American Pie”的音乐服务 (例如，流传输音乐服务，如 Spotify® 或

**Pandora®**) 曲目标识符 (例如, 针对“American Pie”的**Spotify®** 曲目标识符、URI和/或URL)、和/或“American Pie”的原始语音到文本结果。联网麦克风系统可以向媒体回放系统提供来自数据库搜索的一组结果。可以通过元数据交换通道和/或在联网麦克风系统和媒体回放系统之间建立的任何其他通信路径来提供结果。

[0217] 在一些示例中, 口头命令可以指定要播放的特定媒体内容 (例如, 歌曲、艺术家) 或一组媒体内容 (例如, 专辑), 例如, “播放专辑American Pie”, 在这种情况下, 搜索结果可以返回名为“American Pie”的专辑和名为“American Pie”的专辑的音乐服务标识符。

[0218] 作为另一示例, 用户可以说出命令“播放foofoo”, 其中音乐变量是“foofoo”。“foofoo”可能不对应于任何音乐特征或特性, 因此, 数据库可能没有任何对应于音乐变量“foofoo”的艺术家、专辑和/或曲目。在此示例中, 返回的结果可能只是“foofoo”的语音到文本转换的结果。语音到文本转换或原始文本可以用于搜索自定义名称 (例如, 自定义专辑名称、自定义歌曲名称、自定义艺术家姓名)。原始文本可以用于搜索存储在设备 (例如, CR 522) 上, 或存储在支持网络的存储设备 (例如, 网络附接存储 (NAS) 设备) 上的内容。支持网络的存储设备可以能够通过通信路径542、544、546与媒体回放系统和/或联网麦克风系统进行通信。在一些情况下, 用户可以将自定义内容存储在计算设备508上。支持网络的存储设备中的内容可以在系统500中的任何设备上被索引, 并且可以是基于原始文本可搜索的。

[0219] 作为又一示例, 用户可以说出命令“播放甲壳虫”。在音乐上下文之外, “甲壳虫” (Beatles) 可能会被解释为对应于“甲壳虫” (beetles), 因为这些词是同音异义词。由于在框716中可以将口头命令识别为在音乐控制域, 因此联网麦克风系统可以知道在语音命令“播放甲壳虫”中“甲壳虫”可能对应于名为“甲壳虫”的艺术家或其他音乐内容。搜索数据库可以产生如下结果: 艺术家甲壳虫、针对不同流传输服务的甲壳虫的艺术家标识符、基于艺术家甲壳虫的推荐播放列表的标识符、和/或语音“甲壳虫”的语音到文本转换的原始结果。

[0220] 作为另一示例, 用户说出的语音输入可以是“播放爵士乐”。系统500可以基于对存储流派列表的数据库的搜索来将“爵士乐”识别为流派。系统可以搜索对应于爵士乐流派的相关媒体内容。例如, 搜索结果可以是各种音乐服务的数据库中“爵士乐”流派的标识符和/或对应于“爵士乐”流派的播放列表的播放列表名称或播放列表标识符。

[0221] 在一些示例中, 语音输入可以包括依次说出的两个命令或短语。联网麦克风系统可以识别出依次说出的两个命令可能是相关的。第一命令可以包含第一变量类型, 第二命令可以包含第一变量类型的子集。例如, 用户可以提供语音输入“播放一些古典音乐”, 然后是“播放一些浪漫音乐”。两个命令之间可能会有暂停。系统500可以具有与音乐变量的古典实例相关联的各个时代的古典音乐。系统500可以认识出“浪漫”指的是古典音乐的浪漫时代, 并且可以将命令处理为播放一些浪漫时代的音乐的命令, 而不是播放一些浪漫情绪的音乐的命令。作为另一示例, 用户可以说“播放一些\*NSync”, 然后短暂停顿, 然后说“播放No Strings Attached”。系统500可以识别出“No Strings Attached”是艺术家\*NSync的专辑, 并且可以播放该专辑, 而不是另一位艺术家的名为“No Strings Attached”的专辑。

[0222] 在一些示例中, 该命令可以包括用于媒体回放系统名称的系统名称变量。媒体回放系统可以与名称 (例如, 媒体回放系统制造商的名称 (例如, **SONOS®**)、自定义名称和/或其他名称) 相关联。语音输入可以包括系统名称, 用于识别与用户账户相关联的执行命令或意图的特定系统。在一些情况下, 在框1216中, 口头命令中媒体回放系统名称变量的存

在,可以用于自动将语音输入识别为在音乐控制域中。

[0223] 在框1220处,媒体回放系统可以通过执行每个被解析的命令和变量来执行口头音乐控制命令。媒体回放系统可以从联网麦克风系统和/或一个或多个媒体回放系统命令接收搜索结果。媒体回放系统可以在来自联网麦克风系统的一个或多个消息中接收媒体回放系统命令和/或媒体信息(例如,搜索结果)。基于搜索结果,媒体回放系统(例如,计算设备506、控制设备522和/或回放设备532、534、536、538)可以决定响应于口头命令所要播放的媒体内容或进行回放的队列。当接收到除原始文本到语音转换之外的搜索结果时,可以使用各种技术来确定响应于口头命令而播放什么。

[0224] 在一些方面,媒体回放系统可以基于各种数据点和/或偏好给不同的结果权重。例如,不同结果的权重可以基于以下任何项:流行度、用户回放历史、与用户账户相关联的音乐服务、媒体回放系统可用的音乐服务、存储在媒体回放系统相关联的存储设备上的音乐和/或用户偏好。例如,不同艺术家的歌曲可能具有相同的名称。如果搜索结果返回不同艺术家的具有相同名称的歌曲,则可以选择更流行的歌曲,和/或给其更大的权重。可以基于各种源,例如用户播放计数、无线电播放计数、音乐排行榜排名、销售的拷贝和/或其他可用数据来确定歌曲的流行度。作为另一示例,用户最近播放的媒体项可以具有较大的权重。

[0225] 在一些示例中,用户可以针对媒体变量指示用于选择媒体项的优先顺序。例如,用户可以指示自定义播放列表是优先的,接着是广播电台、组织的播放列表、曲目、专辑、艺术家。系统可以基于用户偏好给予优先级较高的媒体变量类型较大的权重。用户偏好可以存储在用户简档中和/或以其他方式与用户账户相关联,使得用户偏好可以被发送到联网麦克风系统和/或媒体回放系统。一些媒体变量类型可以被赋予相同的优先级或权重。

[0226] 在一些示例中,如果用户未提供用户偏好,则可以使用默认优先级排序。系统可以使用不同媒体变量类型的优先级的任意组合。例如,系统可以优先考虑对应于专辑的媒体内容,然后是艺术家和曲目、再然后是对应于原始文本的结果。

[0227] 作为另一示例,媒体变量实例可以是艺术家的姓名,并且艺术家可以对应于播放列表和广播电台。对于媒体变量实例,播放广播电台可以优于播放列表。

[0228] 作为又一示例,媒体回放系统可以过滤掉与用户和/或媒体回放系统不具有访问权的音乐服务相关联的结果。在过滤掉不可访问的结果之后,媒体回放系统可以基于本文描述的各种方法或技术来选择要播放的媒体内容。

[0229] 本文描述的选择方法是说明性的,并且可以使用其他示例方法或技术,用于响应于“播放”命令来选择要播放的媒体项。

[0230] 作为对响应于接收到“播放”命令或意图而选择要播放的内容的响应,媒体回放系统可以使用搜索结果中提供的标识符和/或其他识别信息来识别内容的源。标识符可以是网络地址或链接,其中,媒体回放系统可以检索相应的内容,例如,网络存储位置或统一资源定位符(URL)和/或元数据(例如,艺术家姓名、专辑名称、曲目号、广播电台名称、播放列表名称、媒体服务名称等)。

[0231] 在一些示例中,命令或意图可以是传输命令(例如,暂停、暂停指定时间后播放、跳过、后退、倒回、快进)。如上所述,系统500可以确定对应于命令或意图的媒体回放系统命令,并执行相应的媒体回放系统命令。联网麦克风系统可以通过向媒体回放系统发送消息来使媒体回放系统执行命令,该消息包括识别命令对应的意图(例如,播放、暂停等)和命令

的相关对象(例如,诸如区域名称、区域标识符、组标识符、媒体标识符之类的变量)的命令信息。例如,传输命令可以包括要在其中执行命令的区域的区域变量的实例(例如,Nick的房间、厨房或区域的其他标识符)。

[0232] 响应于基于语音输入在PBD上引起动作,媒体回放系统可以更新和/或存储与PBD执行的动作相关的状态信息。例如,PBD可以更新状态变量以指示区域的状态,例如,区域或特定回放设备当前正在播放特定媒体项,和/或特定媒体项被添加到存储在PBD上的队列。状态变量可以被系统500中的任何设备访问,以便实现系统500中的各种设备的无缝回放和/或控制。

[0233] 虽然本文已经关于媒体内容(例如,音乐内容、视频内容)描述了方法和系统,但是本文描述的方法和系统可以应用于可以具有可由媒体回放系统播放的相关音频的各种内容。例如,可以响应于语音输入来播放可能不是音乐目录的一部分的预先录制的声音。一个示例是语音输入“夜莺的声音是什么样的?”。联网麦克风系统对该语音输入的响应可能不是具有标识符的音乐内容,而可能是短音频剪辑。媒体回放系统可以接收与回放短音频剪辑相关联的信息(例如,存储地址、链接、URL、文件)以及用于播放短音频剪辑的媒体回放系统命令。其他示例是可能的,包括播客、新闻剪辑、通知声音、警报等。

[0234] V. 用于基于用户识别的动作的示例系统、装置和方法

[0235] 本文描述的示例包括媒体回放系统(或者其可能的一个或多个组件)接收语音命令,并基于用户识别(或至少基于说出语音命令的用户)确定媒体回放系统执行的适当动作。在一些示例中,媒体回放系统可以包括计算设备506(其被配置为媒体回放系统服务器)以及PBD532、534、536和538中的一个或多个。在一些实施例中,媒体回放系统可以包括联网麦克风系统或与联网麦克风系统通信,该联网麦克风系统包括计算设备504(其被配置为联网麦克风系统服务器)以及NMD512、514和516中的一个或多个。

[0236] 通常,应当理解,本文描述的一个或多个功能可以由联网麦克风系统单独地或与媒体回放系统组合地执行。还应理解,计算设备506执行的一个或多个功能可以由媒体回放系统的PBD 532、534、536和538中的一个或多个和/或CR 522来执行。

[0237] 如上所述,语音命令的示例包括用于控制先前讨论的任何媒体回放系统控制的命令。例如,语音命令可以是媒体回放系统通过该媒体回放系统的一个或多个回放设备来播放媒体内容的命令。在另一示例中,语音命令可以是触发时间段或窗口的命令,在该时间段或窗口中接收与初始语音命令相关联的附加语音命令。在又一示例中,语音命令可以是修改媒体回放系统的一个或多个媒体回放设备的回放设置的命令。回放设置可以包括例如回放音量、回放传输控制、音乐源选择、分组等。

[0238] 媒体内容的示例包括:热线广播、有声读物、来自电视的音频、存储在本地驱动器上的音乐、或来自媒体源的音乐等。媒体源的示例包括Pandora® Radio、Spotify®、Slacker®、Radio、Google Play™和iTunes Radio等。

[0239] 用户识别的示例包括将用户识别为注册用户、访客用户、儿童或未知用户。

[0240] 示例注册用户包括通过用户简档和/或语音配置设置等与媒体回放系统链接或相关联的一个或多个用户。示例用户简档可以包括关于用户的年龄、位置、偏好的回放设置、偏好的播放列表、偏好的音频内容、对该用户设置的访问约束的信息以及识别用户语音的

信息、用户历史等。识别用户语音的示例信息包括用户语音的音调或频率、年龄、性别和用户历史等。示例语音配置设置可以包括让用户提供语音输入或一系列语音输入以供媒体回放系统识别和关联用户的设置。

[0241] 示例访客用户包括通过注册用户的用户简档、或者注册用户或注册用户许可的访客用户创建的访客简档与媒体回放系统链接或相关联的一个或多个用户。示例访客简档可以包括用户简档中包括的任何类型的信息。

[0242] 在一些示例中,在他或她自己的家中具有他或她自己的媒体回放系统的访客可以将与他或她自己的媒体回放系统相关联的用户简档存储在计算设备506中。当访客来到主人家中并尝试使用语音命令来控制主人的媒体回放系统时,连接到主人的回放系统的计算设备506可以能够访问访客的用户简档设置,包括但不限于:(i) 访客具有用户账户的音乐服务、(ii) 访客的播放列表、(iii) 主人是否已给访客授权了控制主人的媒体回放系统的访问权、和/或(iv) 访客的用户简档中其他可能信息。

[0243] 可以识别儿童用户,例如通过用户简档中的信息(如果儿童是媒体回放系统的注册用户之一)、访客简档中的信息和/或用户语音的音调或频率。

[0244] 在一些示例中,接收语音命令包括媒体回放系统通过PBD 532、534、536和538中的一个或多个和/或计算设备506(其被配置为媒体回放系统服务器)接收语音命令。计算设备506可以将语音命令转换为等效的文本命令,并解析文本命令以识别命令。

[0245] 在另外的示例中,一个或多个功能可以由联网麦克风系统单独地或与媒体回放系统组合地执行。例如,接收语音命令包括联网麦克风系统通过NMD 512、514或516中的一个或多个接收语音命令,以及向计算设备506和/或PBD 532、534、536和538中的一个或多个发送语音命令,用于进一步处理。计算设备506可以将语音命令转换为等效的文本命令,并解析文本命令以识别命令。在一些情况下,联网麦克风系统可以将语音命令转换为等效的文本命令,并向计算设备506和/或PBD 532、534、536和538中的一个或多个发送文本命令,以解析文本命令和识别命令。

[0246] 在接收到语音命令之后,计算设备506和/或PBD 532、534、536和538中的一个或多个可以确定是否是从媒体回放系统的注册用户接收到该语音命令。在一些示例中,确定是否是从注册用户接收到语音命令可以包括:计算设备506和/或PBD 532、534、536和538中的一个或多个确定是否存在与语音命令相关联的存储在媒体回放系统上的用户简档。例如,计算设备506和/或PBD 532、534、536和538中的一个或多个可以尝试将语音命令与识别用户语音的信息匹配,该信息可以包括在存储在媒体回放系统上的用户简档中。单独地或与媒体回放系统组合地,联网麦克风系统可以通过与计算设备506通信来确定是否是从媒体回放系统的注册用户接收到语音命令。

[0247] 在其他示例中,确定是否是从注册用户接收到语音命令可以包括:计算设备506和/或PBD 532、534、536和538中的一个或多个确定语音命令是否与媒体回放系统的语音配置设置中的语音输入匹配。例如,用户可以已经通过提供一个语音输入或一系列语音输入以供媒体回放系统识别并与用户关联,来预先配置媒体回放系统以识别用户语音。一个语音输入或一系列语音输入可以存储在计算设备506和/或PBD 532、534、536和538中的一个或多个上。在一些实施例中,一个语音输入或一系列语音输入可以存储在联网麦克风系统上。

[0248] 在又一示例中,确定是否是从注册用户接收到语音命令可以包括:计算设备506、CR 522和/或PBD 532、534、536和538中的一个或多个单独地或组合地,确定与接收到的语音命令相关联的置信度水平。可以基于用户历史、位置单独地或者与通常在用户简档中找到的任何其他信息组合地,来确定置信度水平。

[0249] 例如,媒体回放系统可以在厨房中从注册用户接收第一语音命令,并基于接收到的语音命令确定置信度。媒体回放系统可以从NMD 512-513、CR 522和PBD 532-538中的任何一个或多个接收第一语音命令。此外,媒体回放系统可以在用户家中的另一房间中从注册用户接收相同的语音命令,并基于接收到的语音命令确定置信度。媒体回放系统可以从NMD 512-513、CR 522和PBD 532-538中的任何一个或多个接收第二语音命令。然后,媒体回放系统可以基于从用户整个家中的不同计算设备(例如,CR 522)、NMD和/或PBD接收的命令来确定新的置信度。相应地,则媒体回放系统可以具有从注册用户接收语音命令的较高置信度水平。

[0250] 在另一情况下,媒体回放系统可以从注册用户接收语音命令,并基于用户历史确定置信度水平。在操作中,媒体回放系统可以从NMD 512-513、CR 522和PBD 532-538中的任何一个或多个接收语音命令。在接收到语音命令之后,如果接收到的语音命令包括艺术家、播放列表、流派、或在通常与注册用户相关联的用户简档中找到的任何其他信息,则计算设备506、CR 522、和/或PBD 532、534、536和538中的一个或多个可以单独地或组合地确定较高的置信度水平。例如,如果注册用户通常听Michael Jackson的歌曲,则媒体回放系统可以具有关于从注册用户接收到播放Michael Jackson的“Thriller”的语音命令的较大置信度水平。与上述类似和不同的许多其他示例是可能的。

[0251] 在其他情况下,媒体回放系统可以基于在用户简档中找到的注册用户的语音命令模式来建立置信度水平。例如,媒体回放系统可以从注册用户接收播放Britney Spears的特定歌曲的语音命令,并基于接收到的语音命令确定置信度水平。每当媒体回放系统接收到相同的语音命令或类似的语音命令时,例如播放Britney Spears的另一首歌的命令,媒体回放系统可以建立较高的置信度水平,因此可以具有关于从注册用户接收到语音命令的较大置信度水平。

[0252] 通常,如前所述,应当理解,本文描述的一个或多个功能可以由联网麦克风系统单独地或与媒体回放系统组合地执行。还应理解,计算设备506执行的一个或多个功能可以由CR 522和/或媒体回放系统的PBD 532、534、536和538中的一个或多个和/或可能由NMD 512、514和516中的一个或多个来执行。

[0253] 在一些示例中,确定置信度水平包括媒体回放系统通过计算设备506(其被配置为媒体回放系统服务器)、CR 522和/或PBD 532、534、536和538中的一个或多个(单独或相互组合地)来确定置信度水平。例如,CR 522和/或PBD 532、534、536和538中的一个或多个可以(i)确定与接收到的语音命令相关联的置信度水平,(ii)基于所确定的置信度水平,确定是从注册用户接收到语音命令,以及(iii)向计算设备506(其被配置为媒体回放系统服务器)发送用于执行语音命令的指令。在另一实例中,CR 522和/或PBD 532、534、536和538中的一个或多个可以(i)确定与接收到的语音命令相关联的置信度水平,以及(ii)向计算设备506发送与置信度水平相关联的数据,用于进一步处理。然后,计算设备506可以(i)基于所确定的置信度水平确定是从注册用户接收到语音命令,以及(ii)向CR 522和/或PBD

532、534、536和538中的一个或多个发送用于执行语音命令的指令。

[0254] 在其他示例中,确定置信度水平包括:媒体回放系统单独地或与联网麦克风系统组合地确定置信度水平。例如,媒体回放系统可以通过CR 522和/或媒体回放系统的PBD 532、534、536和538中的一个或多个和/或可能的NMD 512、514和516中的一个或多个来接收语音命令。响应于接收到的语音命令,媒体回放系统可以向NMD 512、514或516中的一个或多个发送与置信度水平相关联的数据。然后,联网麦克风可以(i)确定与接收到的数据相关联的置信度水平,以及(ii)执行命令或向媒体回放系统发送用于执行命令的指令。响应于确定是从注册用户接收到语音命令,计算设备506可以配置用于媒体回放系统的一个或多个PBD的指令。指令可以基于语音命令中的内容和注册用户的用户简档中的信息。附加地或备选地,该指令可以基于语音命令中的内容和存储在计算设备506、PBD 532、534、536和538中的一个或多个、或联网麦克风系统上的语音配置设置。

[0255] 在一些示例中,语音命令中的内容可以包括用于使媒体回放系统的一个或多个PBD播放媒体内容的命令。在一些实施例中,基于使媒体回放系统播放媒体内容的命令和注册用户的用户简档中的信息,计算设备506可以配置一个指令或一组指令,以使得一个或多个PBD从注册用户的偏好媒体源获得媒体内容。

[0256] 在另一示例中,基于使媒体回放系统播放媒体内容的命令和注册用户的用户简档中的信息,计算设备506可以配置指令,以使得媒体回放系统通过媒体回放系统的一个或多个PBD来播放媒体内容。例如,计算设备506可以包括以下指令:(i)将媒体回放系统配置有注册用户的偏好回放设置中的一个或多个,以及(ii)使得一个或多个PBD在注册用户的偏好回放设置下播放媒体内容。

[0257] 偏好回放设置可以是存储在注册用户的用户简档中的偏好回放设置。附加地或备选地,偏好回放设置可以基于存储在注册用户的用户简档中的用户历史。用户历史可以包括用户播放媒体内容的常用或先前使用的回放设置。

[0258] 在又一示例中,语音命令中的内容可以包括使媒体回放系统播放媒体内容的命令,但是可以不识别媒体回放系统的特定回放区域。基于注册用户的用户简档中的内容和信息,例如用户历史,计算设备506可以(i)配置使得媒体回放系统通过媒体回放系统的特定回放区域内的一个或多个PBD来播放媒体内容的指令,和(ii)实现所配置的指令,以通过一个或多个PBD播放媒体内容。

[0259] 在又一示例中,语音命令中的内容可以包括使媒体回放系统修改回放设置的命令。基于使媒体回放系统修改回放设置的命令和注册用户的用户简档中的信息,计算设备506可以(i)配置用于使媒体回放系统修改用于媒体回放系统的一个或多个PBD的回放设置的指令,和(ii)实现所配置的指令,以修改一个或多个PBD的回放设置。

[0260] 其他的示例可以包括媒体回放系统确定是否是从儿童接收到语音命令。例如,如果儿童是媒体回放系统的注册用户之一,则计算设备506可以基于用户简档中的信息来区分成人和儿童。在另一情况下,计算设备506可以基于用户语音的音调或频率来区分成人和儿童。

[0261] 在又一情况下,确定是否是从儿童接收到语音命令可以包括:计算设备506、CR 522和/或PBD 532、534、536和538中的一个或多个单独地或组合地,确定与接收到的语音命令相关联的置信度。如上所述,可以基于用户历史、位置单独地或者与通常在用户简档中找

到的任何其他信息组合地,来确定置信度水平。

[0262] 在示例操作中,媒体回放系统可以从儿童可能所在的特定房间(例如,儿童的卧室、游戏室、地下室等)中的NMD或PBD接收语音命令。因为语音命令是从儿童可能所在的房间中的设备(NMD或PBD)接收的,所以媒体回放系统可以具有关于从儿童接收到语音命令的较高置信度水平。

[0263] 在另一示例中,媒体回放系统可以接收针对特定类型内容的语音命令,并且基于内容的类型,确定关于从儿童接收到语音命令的较高置信度水平。例如,如果媒体回放系统接收到播放卡通节目或电影的原声音乐的语音命令,则媒体回放系统可以具有关于从儿童接收到语音命令的较高置信度水平。与上述类似和不同的许多其他示例是可能的。

[0264] 响应于确定从儿童接收到语音命令,一些示例可以包括阻止一个或多个PBD播放可能不适合儿童的给定媒体。在一些情况下,可以防止计算设备506和/或一个或多个PBD基于儿童语音命令的内容来修改回放设置。例如,计算设备506和/或一个或多个PBD可以忽略增加一个或多个PBD的音量的儿童语音命令。

[0265] 在一些情况下,媒体回放设备可以基于确定语音命令来自访客用户而不是媒体回放系统的注册用户来采取动作。例如,计算设备506可以存储可与特定访客相关联的预先创建的访客简档。在另一示例中,计算设备506可以确定语音命令不是从注册用户接收的,然后可以询问注册用户该语音命令是否来自访客。然后,注册用户可以具有阻止计算设备506和/或一个或多个PBD执行语音命令的全部或部分内容的选项。

[0266] 在又一示例中,确定是否是从访客用户接收到语音命令可以包括:计算设备506、CR 522和/或PBD 532、534、536和538中的一个或多个单独地或组合地,确定与接收到的语音命令相关联的置信度水平。如上所述,可以基于用户历史、位置单独地或者与通常在用户简档中找到的任何其他信息组合地,来确定置信度水平。

[0267] 响应于确定从访客用户接收到语音命令,计算设备506可以(1)分配访客用户的约束设置,(2)基于语音命令中的内容和分配的访客用户的约束设置来配置用于一个或多个PBD的指令,以及(3)向一个或多个PBD发送指令,以用于执行。在一些示例中,分配访客用户的约束设置可以包括:计算设备506将语音命令与存储在计算设备506和/或一个或多个PBD上的特定访客简档相匹配。访客简档可以包括约束设置,以及关于特定访客用户的语音的信息,例如访客语音的频率或音调,以及先前描述的其他信息。约束设置可以是限制对媒体回放系统的控制的任何设置。

[0268] 其他示例包括媒体回放系统确定优先顺序,以解决从不同用户接收的冲突语音命令。例如,冲突语音命令可以是用户接收的播放歌曲的语音命令和随后从另一用户接收的停止播放歌曲的语音命令。其他示例也是可能的,例如,从用户接收的增加一个或多个PBD 532、534、536和538的音量的语音命令,以及随后从另一用户接收的减小音量的语音命令。

[0269] 具体地,媒体回放系统(通过NMD 512-516、CR 522、PBD 532-538和/或计算设备506中的一个或多个)可以从注册用户或主人接收在回放区域中播放歌曲的语音命令。随后,媒体回放系统可能从非注册用户或访客接收到停止在该回放区域中播放歌曲的冲突语音命令。为解决该冲突,媒体回放系统可以应用优先顺序,其中从注册用户接收的语音命令具有比非注册用户或访客更高的优先级。

[0270] 在另一示例中,媒体回放系统可以分配优先顺序,其中从注册的访客接收的语音命令具有比非注册的访客更高的优先级。在一些情况下,从一个注册的访客接收的语音命令可以具有比另一个注册的访客更高的优先级。附加地或备选地,从成人接收的语音命令可以具有比儿童更高的优先级。

[0271] 在又一示例中,媒体回放系统接收到的控制器发布的命令(例如,由CR 522或被配置为控制媒体回放系统的另一计算设备发出的命令)可以具有比注册用户更低的优先级,但是可以具有被比非注册用户或访客更高的优先级。在一些情况下,一些注册的访客可以具有比控制器发出的命令更高的优先级。确定和分配优先顺序的其他示例是可能的。

[0272] 另外,媒体回放系统可以基于接收到与注册用户相关联的唤醒词或唤醒短语来采取动作。唤醒词或短语可以包括存储在注册用户的简档中的特定单词或短语(例如,“嘿,Sonos”)。在一些情况下,不同的用户可以给媒体回放系统配置不同的唤醒词或短语。在其他情况下,媒体回放系统可以被配置有用于所有(或任何)用户的相同唤醒词或短语。

[0273] 在一些示例中,注册用户可以具有通用唤醒词或短语,其触发时间段或窗口,以用于媒体回放系统从注册用户、访客和/或非注册用户接收与唤醒词或短语相关联的附加语音命令。例如,注册用户或主人可以发送在播放队列中添加歌曲的语音命令(例如,“嘿,Sonos,让我们队列歌曲”),这可以打开一个时间段或窗口(例如,五分钟),在此期间,注册用户或主人可以发送附加的语音命令,以将特定歌曲添加到播放队列中(例如,“添加Michael Jackson的Thriller”)。在另一示例中,注册用户或主人可以发送语音命令(例如,“嘿Sonos,开放对我的家居系统的控制”),其授权家中所有访客在用户定义或默认的时间段或窗口内、或者在特定时间段内发送语音命令来将歌曲添加到播放队列、播放歌曲、或改变音量等(例如,“嘿Sonos,在接下来的4小时内开放对我的家居系统的控制”或“嘿Sonos,从现在开始到周六下午2点,开放对我的家居系统的控制”)。在一些情况下,注册用户或主人可以发送语音命令(例如,“嘿Sonos,对我的起居室的控制限于授权访客”),其仅授权一些访客在一个时间段或窗口内发送语音命令来控制回放区域中的一个或多个PBD 532、534、536和538和/或计算设备506。

[0274] 在另一示例中,注册用户可以具有用于不同语音命令的不同唤醒词或短语,其触发时间段或窗口,以用于媒体回放系统接收与唤醒词或短语相关联的附加语音命令。例如,注册用户或主人可以具有用户特定的唤醒词或短语以发送语音命令将歌曲添加到播放队列(例如,“嘿Sonos,让我们队列歌曲”,“哟,Sonos,队列歌曲”,“阿尔法(Alpha)歌曲队列”等),并且可以具有不同的用户特定的唤醒词或短语,以授权家中的访客控制媒体回放设备(例如,“嘿Sonos,开放访问”,“派对时间到了”等)。

[0275] 在又一示例中,注册用户或主人可具有用户特定或通用的唤醒词或短语,以发送授权家中的某些访客在一个时间段或窗口内具有对媒体回放系统的约束性控制的语音命令。题为“Systems,Methods,Apparatus,and Articles of Manufacture to Provide a Crowd-Sourced Playlist with Guest Access”的美国专利公开No.2013/0346859更详细地提供了对媒体回放系统进行约束性控制的一些示例,该专利通过引用将其整体并入本文。

[0276] 在另一示例中,注册用户或主人可具有用户特定或通用的唤醒词或短语,以发送语音命令,用于授权家中的注册访客在一个时间段或窗口内具有对媒体回放系统的开放控

制或约束性控制,同时防止未注册访客进行控制。在一些情况下,注册用户或主人可具有用户特定或通用的唤醒词或短语,以发送语音命令,用于授权家中的成人在一个时间段或窗口内具有对媒体回放系统的开放控制或约束性控制,同时防止儿童进行控制。与上述类似和不同的许多其他示例是可能的。

[0277] 在一些情况下,注册用户或主人可以指定媒体回放系统接收附加语音命令的时间段或窗口。例如,注册用户或主人可以发送语音命令(例如,“嘿,Sonos,我的家居系统开放控制一小时”),其授权访客在指定时间段内(例如,一小时)发送附加语音命令来控制媒体回放系统。与上述类似和不同的许多其他示例是可能的。

[0278] 在其他情况下,注册用户或主人可以关闭或切断用于接收与初始唤醒词或短语相关联的附加语音命令的时间段或窗口。例如,如果注册用户或主人说出具有唤醒词或短语的语音命令,该语音命令开启用于在一小时内接收附加语音命令的时间段或窗口,则注册用户或主人可以在一小时的时间段到期之前发送另一语音命令(例如,“嘿,Sonos,队列歌曲完成”)以切断该一小时的时间段或窗口。与上述类似和不同的许多其他示例是可能的。

[0279] 进一步地,示例可以涉及媒体回放系统基于从注册访客用户接收到的唤醒词或唤醒短语来采取动作。注册的访客用户可以具有被存储在访客简档中的唤醒词或短语。响应于确定从访客用户接收到唤醒词或唤醒短语,媒体回放系统可以(i)确定是否存在与访客用户相关联的约束设置,(ii)基于唤醒词或短语以及针对访客用户分配的约束设置,配置用于一个或多个PBD的指令,以及(iii)向一个或多个PBD发送该指令以供执行(例如,开启一个时间段或窗口,以用于接收与唤醒词命令相关联的附加语音命令)。

[0280] 在一些情况下,例如,如果媒体回放系统已经从注册用户或主人接收到具有唤醒词或短语的语音命令,而且用于接收附加命令的时间段或窗口尚未到期,则媒体回放系统可以抑制基于从注册的访客用户接收唤醒词或短语来采取动作。

[0281] 在其他情况下,媒体回放系统可以基于从注册的访客用户接收唤醒词或唤醒短语而采取动作,并且如果随后媒体回放系统接收到来自注册用户或主人的语音命令,则可以随后关闭或切断用于接收附加语音命令的时间段或窗口。在一些实施例中,注册访客可以在时间段或窗口到期之前将其关闭或切断。在其他实施例中,如果注册访客是儿童,则成人可以在时间段或窗口到期之前将其关闭或切断。与上述类似和不同的许多其他示例是可能的。

[0282] 在配置用于媒体回放系统的指令之后,一些示例可以包括将指令发送到媒体回放系统的一个或多个PBD以执行指令。在一些示例中,媒体回放系统可以将指令发送到计算设备506。在其他示例中,媒体回放系统可以向联网麦克风系统发送指令。

[0283] 图14中所示的方法1400呈现了可以在包括或包含以下各项的操作环境内实现的方法的实施例:例如,图1的媒体回放系统100、图2的一个或多个回放设备200、图3的一个或多个控制设备300、图4的用户界面和/或图5中所示的配置。方法1400可以包括一个或多个操作、功能或动作,如框1402-1406中的一个或多个所示。

[0284] 方法1400开始于框1402,其包括接收针对媒体回放系统的语音命令。在一些实施例中,接收语音命令包括媒体回放系统通过PBD 532、534、536和538中的一个或多个和/或计算设备506(其被配置为媒体回放系统服务器)接收语音命令。在一个示例中,计算设备506可以将语音命令转换为等效的文本命令,并解析该文本命令以识别命令。

[0285] 在一个示例中,一个或多个功能可以由联网麦克风系统单独地或与媒体回放系统组合地执行。在一些实施例中,接收语音命令包括联网麦克风系统通过NMD 512、514或516中的一个或多个接收语音命令,以及向计算设备506和/或PBD 532、534、536和538中的一个或多个发送语音命令,用于进一步处理。在另一示例中,计算设备506和/或PBD 532、534、536和538中的一个或多个可以将语音命令转换为等效的文本命令,并解析文本命令以识别命令。在另一示例中,联网麦克风系统可以将语音命令转换为等效的文本命令,并向计算设备506和/或PBD 532、534、536和538中的一个或多个发送文本命令,以解析文本命令和识别命令。

[0286] 接下来,方法1400进行到框1404,其包括确定是否是从媒体回放系统的注册用户接收到语音命令。在一些示例中,确定是否是从注册用户接收到语音命令可以包括:计算设备506确定是否存在与语音命令相关联的存储在媒体回放系统上的用户简档。例如,计算设备506可以尝试将语音命令与用户简档中识别用户语音的信息进行匹配。

[0287] 在另一示例中,确定是否是从注册用户接收到语音命令可以包括:确定语音命令是否与媒体回放系统的语音配置设置中存储的语音输入匹配。例如,用户可以已经通过提供一个语音输入或一系列语音输入以供媒体回放系统识别并与用户关联,来预先配置媒体回放系统以识别用户语音。语音配置设置可以存储在计算设备506和/或PBD 532、534、536和538中的一个或多个上。备选地,计算设备506可以与联网麦克风系统通信以存储语音配置设置。

[0288] 在又一示例中,确定是否是从注册用户接收到语音命令可以包括:确定与接收到的语音命令相关联的置信度水平。置信度水平可以是与说出命令的人相关联的置信度水平,例如,从普通注册用户接收命令的置信度水平、从特定注册用户接收命令的置信度水平、从注册用户以外的其他人接收命令的置信度水平、从注册访客接收命令的置信度水平、从儿童接收命令的置信度水平、和/或从特定儿童接收命令的置信度水平。置信度水平也可以是与请求的内容相关联的置信度水平,例如,该请求是播放“AC/DC”而不是例如“Hayseed Dixie”的请求的置信度水平,这两个乐队非常不同,但名字听起来很相似。可以基于用户历史、位置单独地或者与通常在用户简档中找到的任何其他信息组合地,来确定置信度水平。在操作中,置信度水平的确定可以由CR 522、PBD 532-538、NMD 512-516和/或计算设备504-508中的任何一个或多个单独地或组合地执行。

[0289] 在示例操作中,媒体回放系统可以在厨房中从注册用户接收语音命令,并基于接收到的语音命令确定置信度。媒体回放设备可以从CR 522、NMD 512-516和/或PBD 532-538中的任何一个或多个接收语音命令。接下来,媒体回放系统可以在用户家中的另一房间中从注册用户接收相同的语音命令,并基于接收到的语音命令确定置信度。然后,媒体回放系统可以至少部分地基于接收到语音命令的房间,基于从用户的整个家中的不同房间中的不同设备接收的命令来确定新的置信度水平。相应地,则媒体回放系统可以具有从注册用户接收语音命令的较高置信度水平。

[0290] 在另一示例中,媒体回放系统可以从注册用户接收语音命令,并基于用户历史确定置信度水平。具体地,如果接收的语音命令包括艺术家、播放列表、流派或在通常与注册用户相关联的用户简档中找到的任何其他信息,则媒体回放系统可以确定较高的置信度水平。例如,如果注册用户通常听Michael Jackson的歌曲,则媒体回放系统可以具有关于从

注册用户接收到“播放Thriller”的语音命令的较大置信度水平。同样,如果注册用户通常听Michael Jackson的歌曲或大约1980年代的歌曲,则媒体回放系统可以具有关于“播放Thriller”的语音命令是播放艺术家Michael Jackson的“Thriller”而不是Fall Out Boy乐队的歌曲“Thriller”的命令的较大置信度水平。与上述类似和不同的许多其他示例是可能的。

[0291] 在又一示例中,媒体回放系统可以基于在用户简档中找到的注册用户的语音命令模式来建立置信度水平。例如,媒体回放系统可以从注册用户接收播放Britney Spears的特定歌曲的语音命令,并基于接收到的语音命令确定置信度水平。每当媒体回放系统接收到相同的语音命令或类似的语音命令时,例如播放Britney Spears的另一首歌的命令,媒体回放系统可以建立较高的置信度水平,并且可以具有关于从注册用户接收到语音命令的较大置信度水平。

[0292] 最后,方法1400进行到框1406,其包括响应于确定是从注册用户接收到语音命令,基于来自语音命令的内容和注册用户的用户简档中的信息来配置用于媒体回放系统的指令。

[0293] 在一些示例中,语音命令中的内容可以包括用于使媒体回放系统的一个或多个PBD播放媒体内容的命令。在一些情况下,基于使一个或多个PBD播放媒体内容的命令和注册用户的用户简档中的信息,计算设备506可以配置指令,使得媒体播放系统从注册用户的偏好媒体源获得媒体或音频内容。

[0294] 在其他情况下,基于使媒体回放系统播放媒体内容的命令和注册用户的用户简档中的信息,媒体回放系统可以配置指令,以使得媒体回放系统通过媒体回放系统的一个或多个PBD来播放媒体内容。具体地,计算设备506可以包括以下指令:(i)将媒体回放系统配置有注册用户的偏好回放设置中的一个或多个,以及(ii)使得媒体回放系统的一个或多个PBD在注册用户的偏好回放设置下播放媒体内容。偏好回放设置可以是存储在注册用户的用户简档中的偏好回放设置。附加地或备选地,偏好回放设置可以基于存储在注册用户的用户简档中的用户历史。用户历史可以包括用户播放媒体内容的常用或先前使用的回放设置。

[0295] 在一些情况下,语音命令中的内容可以包括使媒体回放系统的一个或多个PBD播放媒体内容的命令,但是可以不识别媒体回放系统的特定收听区域或回放区域。基于该内容和注册用户的用户简档中的信息,例如用户历史,计算设备506可以配置一个指令或一组指令,使得媒体回放系统通过媒体回放系统的特定回放区域内的一个或多个媒体回放设备播放媒体内容。

[0296] 在另一情况下,语音命令中的内容可以包括使媒体回放系统修改回放设置的命令。基于使媒体回放系统修改回放设置的命令和注册用户的用户简档中的信息,计算设备506可以(i)配置用于使媒体回放系统修改用于媒体回放系统的一个或多个PBD的回放设置的一个或一组指令,和(ii)实现所配置的一个或一组指令,以修改一个或多个PBD的回放设置。

[0297] 其他的示例可以包括媒体回放系统确定是否是从儿童接收到语音命令。例如,如果儿童是媒体回放系统的注册用户之一,则计算设备506可以基于用户简档中的信息来区分成人和儿童。在另一示例中,计算设备506可以基于用户语音的音调或频率来区分成人和

儿童。

[0298] 在又一示例中,确定是否是从儿童接收到语音命令可以包括:确定与接收到的语音命令相关联的置信度水平。如上所述,可以基于用户历史、位置单独地或者与通常在用户简档中找到的任何其他信息组合地,来确定置信度水平。

[0299] 在一些情况下,媒体回放系统可以通过儿童可能所在的特定房间(例如,儿童的卧室、游戏室、地下室等)中的设备(例如,NMD 512-516或PBD 532-538中的任何一个)接收语音命令。因为命令是从儿童可能所在的房间中的设备接收的,所以媒体回放系统可以具有关于从儿童接收到语音命令的较高置信度水平。

[0300] 在其他情况下,媒体回放系统可以接收语音命令,并基于语音命令的内容确定是从儿童接收到命令的置信度水平。例如,如果媒体回放系统接收到播放卡通节目或电影的原声音乐的语音命令,则媒体回放系统可以具有关于从儿童接收到语音命令的较高置信度水平。与上述类似和不同的许多其他示例是可能的。

[0301] 响应于确定从儿童接收到语音命令,一些示例可以包括阻止媒体回放系统的一个或多个PBD播放可能不适合儿童的给定媒体。一些示例可以涉及防止计算设备506和/或一个或多个PBD基于儿童语音命令的内容来修改回放设置。例如,计算设备506可以忽略增加一个或多个PBD的音量的儿童语音命令。

[0302] 另外,其他示例可以涉及基于确定是从访客用户而不是媒体回放系统的注册用户接收到语音命令的动作。在一些情况下,计算设备506可以存储可与特定访客相关联的预先创建的访客简档。在其他情况下,计算设备506可以确定语音命令不是从注册用户接收的,然后可以询问注册用户该语音命令是否来自访客。

[0303] 此外,确定是否是从访客用户接收到语音命令可以包括:媒体回放系统确定与接收到的语音命令相关联的置信度水平。如上所述,可以基于用户历史、位置单独地或者与通常在用户简档中找到的任何其他信息组合地,来确定置信度水平。

[0304] 响应于确定从访客用户接收到语音命令,计算设备506可以(1)分配访客用户的约束设置,(2)基于语音命令中的内容和分配的访客用户的约束设置来配置用于一个或多个PBD的指令,以及(3)向一个或多个PBD发送指令,以用于执行。在一些实施例中,分配访客用户的约束设置可以包括:计算设备506将语音命令与存储在计算设备506上的特定访客简档相匹配。

[0305] 另外,示例可以涉及媒体回放系统应用优先顺序来解决从不同用户接收的冲突语音命令。例如,冲突语音命令可以是用户接收的播放歌曲的语音命令和随后从另一用户接收的停止播放歌曲的语音命令。其他示例也是可能的,例如,从用户接收的增加一个或多个回放设备(例如,PBD 532、534、536和538)的音量的语音命令,以及随后从另一用户接收的减小音量的语音命令。具体地,媒体回放系统可以从注册用户或主人接收在回放区域中播放歌曲的语音命令。随后,媒体回放系统可能从非注册用户或访客接收到停止在该回放区域中播放歌曲的冲突语音命令。为解决该冲突,媒体回放系统可以应用优先顺序,其中从注册用户接收的语音命令具有比来自非注册用户或访客的语音命令更高的优先级。

[0306] 在一些示例中,媒体回放系统可以分配优先顺序,其中从注册的访客接收的语音命令具有比来自非注册的访客的语音命令更高的优先级。在一种情况下,从一个注册的访客接收的语音命令可以具有比另一个注册的访客更高的优先级。在另一情况下,从成人接

收的语音命令可以具有比儿童更高的优先级。

[0307] 在其他情况下,媒体回放系统接收的控制器发布的命令(例如,从CR 522或被配置为控制媒体回放系统的其他计算设备接收的命令,或者可能是从计算设备506接收的命令)可以具有低于注册用户的优先级,但可以具有高于非注册用户或非注册访客的优先级。在一些实施例中,一些注册的访客可以具有高于控制器发布的命令的优先级。确定和分配优先顺序的其他示例是可能的。

[0308] 在配置用于媒体回放系统的指令之后,一些实施例可以向媒体回放系统的一个或多个PBD发送指令以执行指令。在一些实施例中,计算设备506可以向联网麦克风系统发送指令。

[0309] 图15中所示的方法1500呈现了可以在包括或包含以下各项的操作环境内实现的方法的实施例:例如,图1的媒体回放系统100、图2的一个或多个回放设备200、图3的一个或多个控制设备300、图4的用户界面和/或图5中所示的配置。方法1500可以包括一个或多个操作、功能或动作,如框1502-1506中的一个或多个所示。

[0310] 方法1500开始于框1502,其包括接收与针对媒体回放系统的语音命令相关联的唤醒词或唤醒短语。如上所述,唤醒词或短语可以是存储在用户简档中的特定单词或短语(例如,“嘿,Sonos”)。在一些实施例中,媒体回放系统可以接收与注册用户的语音命令相关联的通用唤醒词或短语(例如,“嘿Sonos”)。附加地或备选地,媒体回放系统可以接收与注册的访客用户的语音命令相关联的通用唤醒词或短语。在一些情况下,媒体回放系统可以被配置为针对不同注册用户具有不同唤醒词或短语。

[0311] 具体地,注册用户可以具有针对不同语音命令的不同的用户特定唤醒词或短语。例如,媒体回放系统可以接收唤醒词或短语以将歌曲添加到播放队列(例如,“嘿Sonos,让我们队列歌曲”,“哟,Sonos,队列歌曲”,“阿尔法歌曲队列”等),并且可以接收不同的用户特定的唤醒词或短语,以授权家中的访客控制媒体回放设备(例如,“嘿Sonos,开放访问”,“派对时间到了”等)。

[0312] 接下来,方法1500进行到框1504,其包括确定是否是从媒体回放系统的注册用户接收到与语音命令相关联的唤醒词。在一些实施例中,确定是否是从注册用户接收到与语音命令相关联的唤醒词可以类似于在方法1400的框1404中描述的确定是否是从注册用户接收到语音命令。

[0313] 最后,方法1500进行到框1506,其包括响应于确定是从注册用户接收到与语音命令相关联的唤醒词,基于接收到的唤醒词、来自语音命令的内容和注册用户的用户简档中的信息来配置用于媒体回放系统的指令。

[0314] 在一些示例中,用于媒体回放系统的指令可以包括开启一个时间段或窗口,用于媒体回放系统从注册用户、访客和/或非注册用户接收与接收到的唤醒词相关联的附加语音命令。例如,响应于确定是从注册用户接收到将歌曲添加到播放队列的唤醒词,媒体回放系统可以开启一个时间段(例如,五分钟),用于注册用户发送将特定歌曲添加到播放队列中的附加语音命令(例如,“添加Michael Jackson的Thriller”)。

[0315] 在另一示例中,响应于确定是从注册用户接收到授权所有访客控制媒体回放系统的唤醒词,媒体回放系统可以开启一个时间段(例如,一小时),用于允许家中所有访客发送将歌曲添加到播放队列、播放歌曲或改变音量等用于用户定义或默认的时间段或窗口的其

他功能的语音命令。

[0316] 接下来,方法1500进行到框1506,其包括响应于确定是从注册用户接收到唤醒词,基于接收到的唤醒词或短语、语音命令中的内容和注册用户的用户简档中的信息来确定唤醒词是否与约束设置相关联。

[0317] 在一些示例中,媒体回放系统可以基于注册用户或注册的访客用户的用户简档中的约束设置来配置指令。从注册用户接收的唤醒词可以与某些访客的约束设置相关联。例如,注册用户或主人可以发送语音命令(例如,“嘿Sonos,对我的起居室的控制限于授权访客”),其授权注册的访客在一个时间段或窗口内发送附加语音命令以控制回放区域中的一个或多个PBD 532、534、536和538和/或计算设备506,同时阻止未注册的访客发送附加语音命令。在另一种情况下,接收到的唤醒词可以与儿童的约束设置相关联。与上述类似和不同的许多其他示例是可能的,包括但不限于本文其他地方描述的示例。

[0318] 在其他示例中,从注册用户接收的唤醒词可以与约束设置相关联,该约束设置允许某些访客在一个时间段或窗口内对媒体回放系统具有约束性控制。题为“Systems, Methods, Apparatus, and Articles of Manufacture to Provide a Crowd-Sourced Playlist with Guest Access”的美国专利公开No.2013/0346859更详细地提供了对媒体回放系统进行约束性控制的一些示例,该专利通过引用将其整体并入本文。

[0319] 响应于确定从访客用户接收到唤醒词或唤醒短语,媒体回放系统可以(i)确定是否存在与访客用户相关联的约束设置,(ii)基于唤醒词或短语以及针对访客用户分配的约束设置,配置用于一个或多个PBD的指令,以及(iii)向一个或多个PBD发送该指令以供执行(例如,开启一个时间段或窗口,以用于接收与唤醒词命令相关联的附加语音命令)。

[0320] 在一些示例中,例如,如果媒体回放系统已经从注册用户或主人接收到具有唤醒词或短语的语音命令,而且用于接收附加命令的时间段或窗口尚未到期,则媒体回放设备可以通过一个或多个PBD 532、534、536和538和/或计算设备506,抑制基于从注册的防客用户接收到唤醒词或短语来采取动作。

[0321] 在配置用于媒体回放系统的指令之后,一些示例可以包括向媒体回放系统的一个或多个PBD发送命令以执行指令。在一些示例中,计算设备506可以向媒体回放系统的一个或多个PBD发送一个命令或一组命令。

[0322] 在一些情况下,在配置用于媒体回放系统执行的指令之后,注册用户或主人可以关闭或切断用于接收与指令相关联的附加语音命令的时间段或窗口。例如,如果注册用户或主人发送具有唤醒词或短语的语音命令,该语音命令开启用于在一小时内接收附加语音命令的时间段或窗口,则注册用户或主人可以在一小时的时间段到期之前发送另一语音命令(例如,“嘿,Sonos,队列歌曲完成”)以切断该一小时的时间段或窗口。与上述类似和不同的许多其他示例是可能的。

[0323] 其他示例可以包括媒体回放系统基于从注册的访客用户接收唤醒词或唤醒短语而采取动作,并且如果随后媒体回放系统接收到来自注册用户或主人的语音命令,则可以随后关闭或切断用于接收附加语音命令的时间段或窗口。在一些实施例中,注册访客可以在时间段或窗口到期之前将其关闭或切断。在其他实施例中,如果注册访客是儿童,则成人可以在时间段或窗口到期之前将其关闭或切断。与上述类似和不同的许多其他示例是可能的。

[0324] VI. 用于音乐服务选择的示例系统、装置和方法

[0325] 本文描述的示例涉及基于命令来识别和访问合适的流传输服务(例如,流传输音频曲目)。

[0326] 图16和图17中所示的方法1600和1700呈现了可以在操作环境内实现的方法的实施例,该操作环境包括,例如,图1的媒体回放系统100、图2的回放设备200中的一个或多个以及图3的控制设备300中的一个或多个。方法1600和1700可以包括一个或多个操作、功能或动作,如框1602-1614和1702-1708中的一个或多个所示。

[0327] 图16是与用于识别流传输音乐服务的过程相关的示例流程图。在1602处,NMD 512-516或CR 522可以接收命令的指示,其可以指示从流传输服务提供用于回放的音频内容。在一些情况下,NMD 512-516接收的命令可以采取语音命令的形式,而CR 522接收的命令可以是用户界面上的文本命令输入。

[0328] 通常,接收的命令可以包括与一个或多个音频内容类型有关的信息。在一些情况下,命令可以包括艺术家、歌曲、专辑或流派名称(即,“播放Led Zeppelin”,“播放70年代的摇滚乐”)。另外,命令可以包括可以进一步指示内容类型的前缀和/或后缀类型信息(例如,“最佳.....”、“.....无线电”、“.....播放列表”)。例如,命令“播放Led Zeppelin无线电”可以指示用户希望以无线电格式收听特定艺术家的音乐。接收的命令还可以包括指示内容类型的各种其他形式的信息。

[0329] 可以以各种方式处理在1602处接收的命令指示。在一个实现中,命令的处理可以通过云网络502完成。在这种情况下,由NMD512-516接收的语音命令可以使语音输入通过通信网络546发送到计算设备504-508中的一个或多个以进行处理。云计算设备可以将语音输入转换为等效的文本命令,并解析文本命令以识别命令。在另一配置中,云计算设备可以仅将语音输入转换为等效的文本格式,并向第二计算设备发送该等效文本以进行解析和命令识别。在其他情况下,NMD 512-516可以在通过通信网络546传输之前将语音输入转换为文本,或者将语音输入转换为文本并且执行解析以识别命令。在CR522接收文本命令的情况下,可以通过通信网络546向计算设备504-508之一发送文本输入以进行解析和命令识别。在另一情况下,CR 522可以执行对文本输入的解析以识别命令。

[0330] 在另一实现中,命令的处理可以通过本地网络在本地完成。在这种情况下,NMD 512-516接收的语音命令可以使语音输入通过本地网络被发送到一个或多个本地计算设备以进行处理。本地计算设备可以将语音输入转换为等效的文本命令,并解析文本命令以识别命令。在另一配置中,本地计算设备可以仅将语音输入转换为等效的文本格式,并向第二本地计算设备发送该等效的文本格式以进行解析和命令识别。在其他情况下,NMD 512-16可以在通过本地网络传输之前将语音输入转换为文本,或者将语音输入转换为文本并且执行解析以识别命令。在CR 522接收文本命令的情况下,可以通过本地网络向本地计算设备发送文本输入以进行解析和命令识别。在另一情况下,CR 522可以执行对文本输入的解析以识别命令。可以存在用于处理命令的其他配置。

[0331] 在1604处,计算设备可以识别命令所指示的内容类型。内容类型的标识可以通过云网络502或通过本地网络在本地完成。

[0332] 在一个实现中,计算设备可以使用内容类型逻辑来将命令或命令的部分与内容类型相关联。使用“播放Led Zeppelin无线电”的上述示例,计算设备可以将内容类型识别为

“艺术家/无线电台”。在另一示例中，命令“播放电子舞曲”可以使内容类型被识别为“流派”。可以对各种其他内容类型进行类似的识别。对内容类型的识别可以例如通过输入诸如“舞曲”之类的命令关键词来完成，并且数据库可以将关键词映射到诸如流派之类的内容类型指示。在一些示例中，数据库可以驻留在计算设备上或网络麦克风设备上。

[0333] 如果确定在1608处没有识别出内容类型，则该方法可以直接进行到1614。内容类型可能因多种原因而不可识别，这些原因包括用户输入错误、语音输入质量差、背景噪声、或者根本不知道这样的内容类型。例如，可能无法识别指示鲜为人知的艺术家姓名的命令的内容类型。

[0334] 在1614处，指示可以由计算设备输出，并且通过通信网络546发送到NMD512-516、PBD 532-538或CR 522中的任何一个或全部，以指示“内容不可用”。然后可以向用户可听地或可视地呈现没有可用内容的指示。例如，NMD和PBD可以输出可听指示，而CR可以能够输出可听和可视指示。附加地或备选地，所发送的指示可以使得向用户输出指示他或她重新输入命令的建议。例如，建议可以是让用户指定一些附加识别特性以便帮助识别内容类型。

[0335] 然而，如果在1606处确定已经识别出内容类型，则该方法可以进行到1608以识别能够播放在1604处识别的内容类型的流传输服务。通常，特定流传输服务可能与其他流传输服务显著不同，不仅在于它们提供的音频内容不同，而且在于它们呈现内容的方式不同。例如，每个流传输服务可以拥有流式传输某些艺术家或专辑的音乐内容的相对专有权利。在另一种情况下，一些流传输服务，例如Pandora®，可能只以无线电台格式进行流式传输，而其他一些流传输服务，例如Spotify®，可能能够根据艺术家、歌曲、专辑或无线电台按请求来流式传输音乐。鉴于这一事实，显然并非所有的流传输服务都能够流式传输在704处识别的内容类型。

[0336] 在一种情况下，计算设备可以通过将所识别的内容类型的元数据与查找表进行比较来识别合适的流传输服务，该查找表可以包含可用内容的条目，以及各种流传输服务能够以什么格式来提供内容。在一些情况下，计算设备可以将查询引导到所有可用的流传输服务。在其他情况下，计算设备可以仅查询可用流传输服务的子集。这种子集可以由计算设备基于多个因素单独或组合地进行选择，这些因素包括用户注册的流传输服务、自用户上次使用流传输服务以来的天数、流传输服务流行度、用户设置等。例如，如果用户仅在Pandora®、Spotify®和Deezer®上注册，则计算设备可以仅查询这些流传输服务以确定哪些是合适的。

[0337] 这样的查找表可以存储在计算设备上的存储器中、或存储在诸如计算设备之类的外部位置、或者存储在音乐服务处。考虑到各种查找表可能分布于多个音乐服务，计算设备可以同时或顺序地查询每个音乐服务以便找到匹配。识别合适的流传输服务的其他方式也是可能的。

[0338] 在一个实现中，在1608处对流传输服务的识别还可以包括确定用户注册的流传输服务的当前可用回放容量。通常，一些流传输服务可能会限制在任意给定时间注册用户可用的活动流的数量。例如，Spotify®可能只允许每个注册账户单个活动流。在一种情况下，计算设备可以通过查询用户注册的服务的使用状态（即，有多少活动流），然后将使用状态与容量约束数据（即，Spotify=仅1个活动流）进行比较，来确定当前可用的回放容量。在另

一示例中,流传输服务可以响应于查询而输出二进制值以指示是否有可用的流。可以以其他方式确定可用回放容量。

[0339] 在该实现中,在1608处,计算设备可以将注册的服务识别为支持命令指示的内容类型,并进一步确定注册的服务没有可用的流。例如,如果用户及其配偶共享Spotify®账户,并且当用户发出命令“播放Eye of the Tiger”时,音乐正被流式传输到在健身房的配偶的智能电话设备,则计算设备可以将Spotify®识别为能够播放这首歌以及流不可用。这种情况可以使计算设备识别能够支持该内容类型的另一流传输服务,例如Apple Music®。

[0340] 在另一种情况下,计算设备可能无法识别可支持该内容类型的另一流传输服务。例如,如果用户请求了由单个流传输服务独家提供的内容,计算设备仅考虑流传输服务的子集等,则可能发生上述情况。在这种情况下,计算设备可以使当前活动流“被盗”以用于提供对应于命令的内容。使用上述示例,如果没有其他流传输服务能够支持“Eye of the Tiger”,则可以取消在健身房的配偶的流并将其提供给用户。

[0341] 在一种情况下,在可用的流传输服务被识别(1610处)为能够支持所识别的内容类型的情况下,该过程可以进行到1614以使PBD 532-538的任何组合回放音频内容。在一种情况下,可以通过向服务API查询内容并使内容被流式传输来访问音乐服务。根据来自PBD 532-538或计算设备504-506的请求,音频内容可以直接从计算设备508或从与流传输音乐服务相关联的各种其他计算设备直接地流式传输到PBD 532-538。还存在发起和引起对流媒体内容的回放的其他方式。

[0342] 在另一种情况下,如果在710处识别的流传输服务当前不可用(即,未安装应用,用户未注册),则在714处,计算设备可以输出与使得音乐服务能够被使用的建议动作步骤有关的指示。该指示可以被发送NMD 512-516、PBD 532-538或CR 522的任何组合,并且可以产生可听和/或视觉建议,指示所识别的音乐服务能够支持内容类型,和/或呈现关于如何注册、下载或以其他方式使用音乐服务的指令。

[0343] 图17是与图16的框1610和1612中用于识别流传输音乐服务的示例过程相关的另一示例流程图。在1702处,计算设备可以至少部分地基于命令识别的内容类型来产生要针对流传输服务确定的一个或多个置信度指标。通常,置信度指标可以是针对一个或多个流传输服务计算的数值或百分比值(例如,1-100)。这种置信度指标可以反映选择特定音乐服务来提供流传输音频将向用户提供他或她所期望的内容的可能性。例如,被分配了置信度指标80的流传输服务可能比置信度指标为45的流传输服务更适合于进行内容提供,其中较高的数字表示较高的适用置信度水平。

[0344] 所计算的流传输服务的置信度指标可以基于诸如内容类型、回放容量、使用历史、外部数据等的数字标准。这样的标准可以由各种数据类型构成,并且可以从诸如NMD、CR、PBD、计算设备、音乐服务和各种外部源之类的各种源检索。数据可以被合计并存储在中央位置(例如,与计算设备504或506相关联的数据库),或者以分布式方式存储。

[0345] 在一种情况下,置信度指标可以考虑流传输服务支持命令指示的内容类型的适用性。确定各种流传输服务提供内容类型的适用性可以包括,将与所识别的内容类型有关的元数据映射到查找表或者查询具有各种音乐服务的表,并给匹配的字段的数量赋值。作为

示例,指定“播放Jackson 5播放列表”的命令可以具有内容类型“艺术家/播放列表”。在这种情况下,具有两个流传输服务(例如,**Pandora®**和**Spotify®**)的查找表都可能在艺术家字段中包含对Jackson 5的引用。但是,只有**Spotify®**可能包含标识Jackson 5的播放列表字段,因为**Pandora®**不支持播放列表的内容类型。在这种情况下,**Pandora®**可以被提供内容类型值2,**Spotify®**被提供值1。

[0346] 附加地或备选地,可以采用字段匹配的强度。使用上述示例并假设**Spotify®**服务不包含与Jackson 5对应的艺术家字段但具有Michael Jackson的艺术家条目,其可以由计算设备利用音乐元数据被识别为Jackson 5的前成员。在这种情况下,**Spotify®**服务可能不会在艺术家字段被赋值0,而是该服务被提供小于1的调整值。确定流传输服务内容类型适用性值的其他形式是可能的。

[0347] 在另一种情况下,可以部分地基于各种形式的历史使用数据来计算置信度指标。各种历史数据类型可以从诸如NMD、CR、PBD、计算设备、音乐服务和各种外部源之类的各种源检索。数据可以被合计并存储在中央位置(例如,与计算设备504或506相关联的数据库),或者以分布式方式存储。

[0348] 例如,使用数据可以指示用户访问给定流传输服务的频率。作为另一示例,使用数据可以包括基于时间的数据,以识别用户通常在一天中的不同时间、一周中的各天和一年中的各个月所使用的服务。例如,用户可能更喜欢在早上收听**iHeartRadio®**,在晚上收听**Tidal®**。作为另一示例,在多区域环境中,这些使用数据可以基于每个区域指示关于流传输服务的用户偏好。例如,如果用户通常在浴室区域中90%的时间访问**Spotify®**,在生活区域中80%的时间访问**Deezer®**,则与**Spotify®**和**Deezer®**对应的置信度指标可能会有很大差异,具体取决于用户打算将音频流式传输到哪个区域。在这种情况下,如果用户希望在浴室中听音乐,则**Spotify®**将是更好的选择。还可以存在各种其他类型的历史使用数据。

[0349] 另外,可以部分地基于各种“外部”数据类型来确定置信度指标。这样的数据可以包括宏观类型数据,其可以考虑地理位置或人口统计数据等。例如,这种宏观数据可以指示特定流传输服务在世界的某些区域或国家的子区域中不可用或不受欢迎,这可以导致较低的置信度指标。在这种情况下,流传输服务的置信度指标可以根据地理位置而变化。“外部”数据的类型还可以包括天气数据,可以考虑天气数据,例如,用户喜欢在凉爽干燥的夏季夜晚在露台上收听**Spotify®**。另外,可以考虑日历数据,以识别假日,以及通常在这些日子流式传输的音乐服务。可以存在其他形式的外部数据。

[0350] 此外,可以组合各种标准以确定给定流传输服务的置信度指标。例如,用户可能更喜欢使用**Pandora®**在起居室中收听各种古典音乐,但可能仅使用**Apple Music®**在卧室中收听他们最喜欢的艺术家的完整专辑。在这种情况下,内容类型(流派、艺术家)可以与关于特定音乐服务的使用位置的用户历史相结合,以确定特定音乐服务的置信度指标。可以

存在许多其他组合。

[0351] 计算机实现的算法可以将上述标准映射到流传输服务的置信度指标。例如,可以将一个或多个标准映射到输出置信度指标的表中。在一些实施例中,算法可以涉及对各种标准(例如,内容类型、使用历史和/或“外部数据”)进行加权。算法可以基于相对重要性为各种标准分配不同的权重。例如,可以认为用户偏好比人口统计数据更有影响力并且为其提供更大的权重。输入的权重可以由系统、通过用户设置定义,或者可基于用户反馈动态调整。每个加权输入可以被输入到表中,例如,用于映射到置信度指标,然后将其组合,以整体形成音乐服务的合计置信度指标。

[0352] 在1704处,计算设备504或506可以确定给定流传输服务是否满足置信度条件。可以以多种方式满足置信度条件。在一种情况下,如果给定流传输服务的置信度指标超过置信度水平阈值(其可以是默认系统设置或者可由用户调整),则可以满足置信度条件。例如,如果三个置信度指标被计算为服务1=85、服务2=83、服务3=25且阈值置信度为80,则在1706处,可以使得输出服务1以进行流式传输或者向用户建议服务1(如上文参考1614所讨论的)。

[0353] 在另一种情况下,仅当(1)流传输服务的所计算的置信度指标大于阈值置信度水平,且(2)所计算的两个最高置信度指标不在彼此的阈值范围内时,才满足置信度条件。例如,如果在上面刚刚提到的情况下置信度水平范围是3,则由于服务1和服务2的置信度指标,将不满足置信度条件。置信度条件可以附加地或备选地包含各种其他规则。

[0354] 如果在1708处发现不满足置信度条件,则计算设备504或506可以在1708处触发错误状态。错误状态可以触发引起某些事件,例如,输出内容不可用的指示,如参考1616所讨论的。在另一种情况下,触发的错误状态可以使得内容通过默认流传输服务或优选的合作伙伴服务进行流式传输。在又一情况下,错误状态可以使计算设备向NMD 512-516、PBD 532-538或CR 522中的一个或全部输出指示,以引起对用于获得与已接收的命令有关的更多信息的指令或查询的听觉或视觉呈现。

[0355] 在一种情况下,可以指示用户提供诸如艺术家或专辑之类的附加内容类型。例如,如果初始命令是“播放无线电台”,则可以向用户指示“请提供流派”。

[0356] 在另一种情况下,可以向用户询问可帮助进一步调整初始命令从而调整各种流传输服务的置信度指标的一个问题或一系列问题。例如,如果初始命令针对的是宽泛的流派“电子乐”,则可以询问用户他们是否喜欢特定的子流派,例如“鼓和贝斯”或“迷幻曲”。附加地或备选地,可以询问用户与音乐无关的问题,例如“你在做什么?”或“你感觉如何?”来推断用户可能希望收听的子流派。例如,如果用户对上述问题回答“阅读”和/或“放松”,则可以推断出用户对子流派“迷幻曲”感兴趣。询问的其他示例是可能的。

[0357] 用户对指令或问题的响应可以采用可被NMD接收的语音输入或者可通过CR的图形界面接收的文本输入的形式。用户响应可以使过程1600重复,这可能导致满足置信度条件或者触发附加错误状态。

[0358] VII. 其他示例特征

[0359] (特征1)一种用于回放系统的方法,所述回放系统包括至少一个回放设备和包括麦克风设备的网络设备,所述方法包括:获得回放设备所处环境的声学特性;识别所述回放设备的环境中的网络麦克风设备;向所述网络麦克风设备提供所述声学特性;以及由所述

网络麦克风设备将所述声学特性应用于所述网络麦克风设备接收的语音输入。

[0360] (特征2) 根据特征1所述的方法,还包括:

[0361] 向所述网络麦克风设备提供由所述回放设备在所述环境中回放的音频内容;由所述麦克风设备将所提供的音频内容应用于所述网络麦克风设备接收的语音输入。

[0362] (特征3) 根据特征1或2所述的方法,其中,所识别的网络麦克风设备处于以下中的至少一种情形:被绑定到所述回放设备;以及和所述回放设备在同一区域中。

[0363] (特征4) 根据前述任一特征所述的方法,其中,向所述麦克风设备发送所述声学特性的所述回放设备是所述媒体回放系统中的多个回放设备中最靠近所述麦克风设备的回放设备。

[0364] (特征5) 根据前述任一特征所述的方法,还包括:从所述网络麦克风设备接收对所述语音输入的方向的指示;以及基于接收到的对所述语音输入的方向的指示,调整所述回放设备播放的音频内容的方向性。

[0365] (特征6) 根据前述任一特征所述的方法,其中,使得所述网络麦克风设备将所述声学特性应用于所述网络麦克风设备接收的语音输入包括:使得所述回放设备将基于所述声学特性的滤波器应用于接收到的语音输入。

[0366] (特征7) 根据前述任一特征所述的方法,其中,在校准阶段中获得所述环境的所述声学特性,在所述校准阶段中:一个或多个回放设备输出一个或多个音调,所述网络设备的麦克风接收所述一个或多个回放设备输出的所述音调;以及分析接收到的音调以确定所述环境的所述声学特性。

[0367] (特征8) 根据前述任一特征所述的方法,其中,向所述麦克风设备提供所述声学特性包括以下中的一个:将所述声学特性作为消息发送到所述麦克风设备;以及为所述麦克风设备提供对所述声学特性的访问权。

[0368] (特征9) 一种计算机可读存储介质,包括用于由处理器执行的指令,所述指令在被执行时使得所述处理器实现根据任一前述特征的方法。

[0369] (特征10) 一种媒体回放系统,包括至少一个回放设备和包括麦克风设备的网络设备,所述媒体回放系统被配置为执行任一前述特征所述的方法。

[0370] (特征11) 一种用于计算设备的方法,所述方法包括:接收语音输入,所述语音输入包括:命令字、一个或多个媒体变量实例和一个或多个区域变量实例;确定对应于所述命令字的媒体回放系统命令;识别对应于所述一个或多个媒体变量实例的媒体内容;以及使得媒体回放系统基于所述一个或多个区域变量实例对所述媒体内容执行所述媒体回放系统命令。

[0371] (特征12) 根据特征11所述的方法,还包括:在确定所述媒体回放系统命令对应于所述命令字之前,确定所述语音输入对应于音乐控制,其中,基于对应于音乐控制的可用命令来确定对应于所述命令字的所述媒体回放系统命令。

[0372] (特征13) 根据特征11或12所述的方法,还包括:从所识别的媒体内容中确定一个或多个媒体项;以及发送指示所述一个或多个媒体项的网络存储位置的标识符。

[0373] (特征14) 根据特征11至13中任一项所述的方法,还包括:基于所述命令字在第一位置来识别所述语音输入中的所述命令字;基于所述一个或多个媒体变量实例在第二位置,来识别所述语音输入中的所述一个或多个媒体变量实例;以及基于所述一个或多个区

域变量实例在第三位置来识别所述语音输入中的所述一个或多个区域变量实例。

[0374] (特征15) 根据前述特征11至14中任一项所述的方法,其中,确定对应于所述命令字的所述媒体回放系统命令包括:确定对应于所述命令字的意图;以及确定对应于所述意图的所述媒体回放系统命令。

[0375] (特征16) 根据前述特征11至15中任一项所述的方法,其中,接收到的语音输入的所述一个或多个区域变量实例指示媒体回放系统的一个或多个区域,所述一个或多个区域包括一个或多个回放设备。

[0376] (特征17) 根据前述特征11至16中任一项所述的方法,其中,使得所述媒体回放系统对所述媒体内容执行所述媒体回放系统命令包括:从所述计算设备向所述媒体回放系统发送消息,所述消息包括:识别所述媒体回放系统命令的命令信息,识别对应于所述一个或多个媒体变量实例的媒体内容的媒体信息;以及对应于所述一个或多个区域变量实例的一个或多个区域标识符。

[0377] (特征18) 一种存储指令的有形的非暂时性计算机可读介质,所述指令当被计算设备的一个或多个处理器执行时,使得所述计算设备执行前述特征11-17中任一项的方法。

[0378] (特征19) 一种计算设备,包括:一个或多个处理器;存储指令的存储器,所述指令在被所述一个或多个处理器执行时,使得所述计算设备执行特征11至17中任一项的方法。

[0379] (特征20) 一种方法,包括:接收用于媒体回放系统的语音命令;确定是否是从所述媒体回放系统的注册用户接收到所述语音命令;以及如果是从注册用户接收到所述语音命令,则基于来自所述语音命令的内容和所述注册用户的用户简档中的信息来配置用于所述媒体回放系统的指令。

[0380] (特征21) 根据特征20所述的方法,还包括向所述媒体回放系统的一个或多个回放设备发送所述指令。

[0381] (特征23) 根据特征21或22所述的方法,其中,所述语音命令是所述媒体回放系统播放媒体内容的命令。

[0382] (特征24) 根据特征23所述的方法,其中,所配置的指令指示所述媒体回放系统从所述注册用户的偏好媒体源获得所述媒体内容。

[0383] (特征25) 根据特征23所述的方法,其中,所配置的指令使得所述媒体回放系统通过所述媒体回放系统的一个或多个媒体回放设备来播放所述媒体内容。

[0384] (特征26) 根据特征23所述的方法,其中,所配置的指令包括用于以下操作的指令:将所述媒体回放系统配置有所述注册用户的偏好回放设置中的一个或多个偏好回放设置;以及使得所述媒体回放系统通过所述媒体回放系统按照所述注册用户的偏好回放设置来播放所述媒体内容。

[0385] (特征27) 根据特征26所述的方法,其中,所述注册用户的偏好回放设置包括以下中的一个或多个:偏好回放音量和偏好音频均衡设置。

[0386] (特征28) 根据特征21或22所述的方法,其中,所述语音命令是所述媒体回放系统修改回放设置的命令,并且所配置的指令使得所述媒体回放系统修改所述媒体回放系统的一个或多个媒体回放设备的回放设置。

[0387] (特征29) 根据前述特征20至28中任一项所述的方法,还包括:如果并非是从注册用户接收到所述语音命令,则确定是否是从访客用户接收到所述语音命令;以及如果是从

访客用户接收到所述语音命令,则:为所述访客用户分配约束设置;基于来自所述语音命令的内容以及为所述访客用户分配的约束设置,配置用于所述媒体回放系统的指令;以及向所述媒体回放系统发送所述指令。

[0388] (特征30) 根据前述特征20至29中任一项所述的方法,其中,所述媒体回放系统包括回放网络和一个或多个回放设备。

[0389] (特征31) 根据前述特征20至30中任一项所述的方法,还包括如果并非是从注册用户接收到所述语音命令,则忽略所述语音命令。

[0390] (特征32) 一种编码有指令的有形的非暂时性计算机可读介质,其中当一个或多个处理器执行时,所述指令使计算设备执行根据前述特征20至31中任一项所述的方法。

[0391] (特征33) 一种媒体回放系统,包括:一个或多个处理器;以及根据特征32的有形的非暂时性计算机可读介质。

[0392] (特征34) 一种方法,包括:接收内容的指示;基于接收到的指示来识别至少一种内容类型;从多个音乐服务中确定支持所述至少一种内容类型的至少一个音乐服务;基于支持所述至少一种内容类型的所述至少一个音乐服务,使得所述至少一个音乐服务发送与所述内容类型相关联的音频内容。

[0393] (特征35) 根据特征34所述的方法,其中,通过网络麦克风设备来接收内容的指示。

[0394] (特征36) 根据特征34或35所述的方法,其中,确定所述至少一个音乐服务包括确定与所述多个音乐服务中的所述至少一个音乐服务对应的置信度指标。

[0395] (特征37) 根据特征34或35所述的方法,其中,确定所述至少一个音乐服务包括确定是否满足置信度水平条件。

[0396] (特征38) 根据特征37所述的方法,其中,所述置信度指标基于对所述多个音乐服务中的音乐服务的访问的历史。

[0397] (特征39) 根据特征38所述的方法,其中,所述置信度指标还基于使用特定音乐服务的特定位置。

[0398] (特征40) 根据特征34至39中任一项所述的方法,还包括:接收内容的第二指示;基于接收到的指示来识别至少一种内容类型;确定所述多个音乐服务中没有一个音乐服务支持所述至少一种内容类型;以及使得错误状态被触发。

[0399] (特征41) 根据前述特征34至40中任一项所述的方法,其中,所述内容类型选自包括艺术家、流派、歌曲、专辑和无线电台的组。

[0400] (特征42) 根据前述特征34至41中任一项所述的方法,其中,确定所述至少一个音乐服务包括防问查找表,所述查找表包含所述多个流传输服务的可用内容的条目。

[0401] (特征43) 根据前述特征24至42中任一项所述的方法,还包括仅查询所述多个音乐服务中用户注册过的音乐服务。

[0402] (特征44) 根据前述特征34至43中任一项所述的方法,还包括在识别所述内容类型之后,向用户询问一个或多个问题以进一步指定在内容的初始指示中指示的所述内容类型;接收指示更具体的内容类型的进一步用户输入;以及基于所述更具体的内容类型确定所述一个或多个音乐服务。

[0403] (特征45) 一种网络设备,包括:网络接口,被配置为通过网络与多个联网设备通信;包括指令的处理器,所述指令在被执行时使得所述处理器执行根据前述特征34至44中

任一项所述的方法。

[0404] (特征46)一种计算机可读存储介质,包括用于由处理器执行的指令,所述指令在被执行时使得所述处理器实现根据特征34至44中任一项的方法。

[0405] VIII. 结论

[0406] 以上描述公开了各种示例系统、方法、装置和包括在硬件上执行的固件和/或软件等组件的制品。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。

[0407] 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。

[0408] 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存储设备。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,辅存或持久性长期存储设备,如只读存储器(ROM)、光盘或磁盘、紧凑盘只读存储器(CD-ROM)。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,例如有形的存储设备。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。

[0409] 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。

[0410] 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合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。

[0411] 当随附的任一项权利要求被理解成涵盖纯软件和/或固件的实现时,在此明确限定至少一个示例中的至少一个单元以包括存储软件和/或固件的非暂时性有形介质,如存储器、DVD、CD、蓝光等。

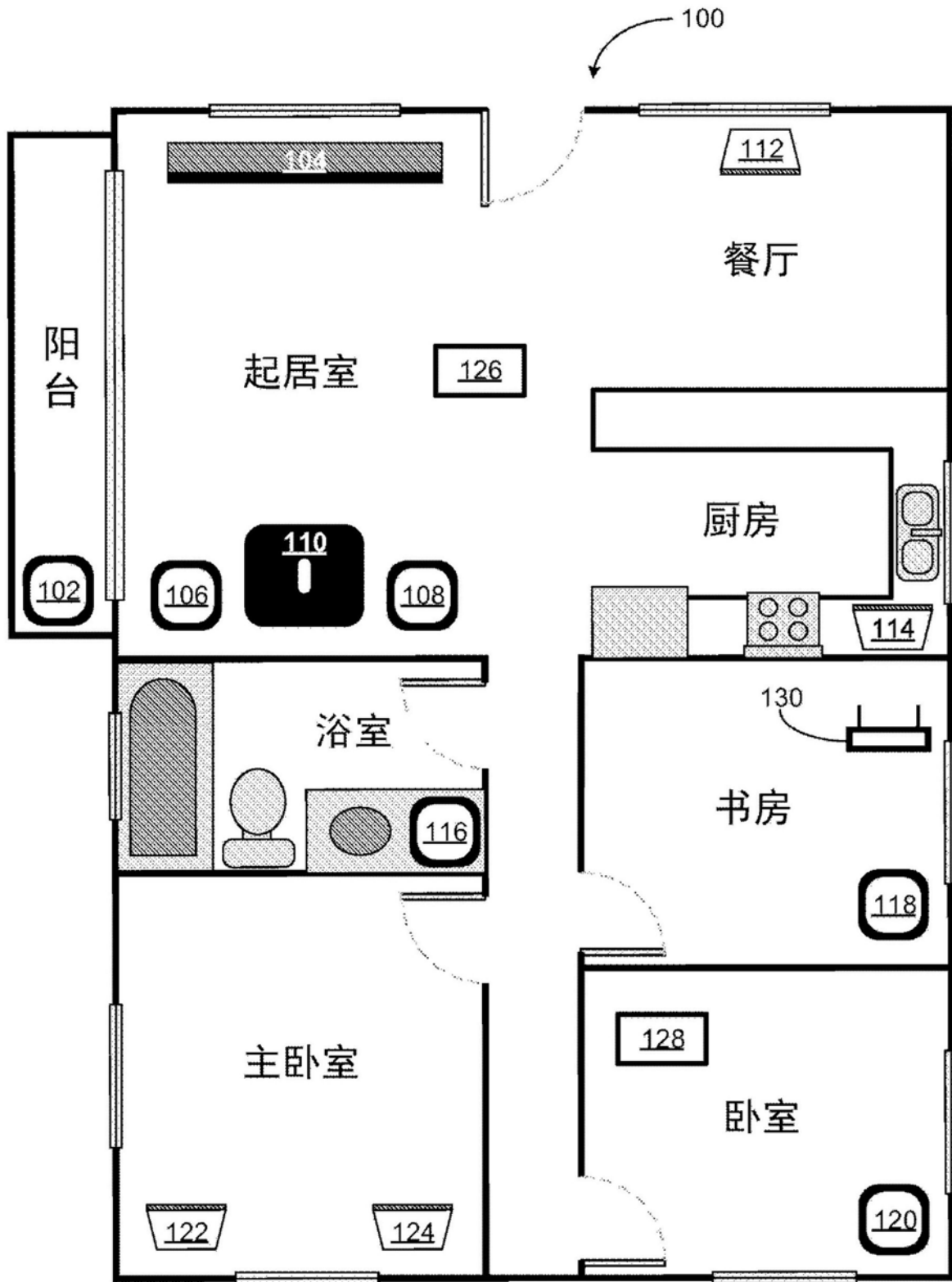


图1

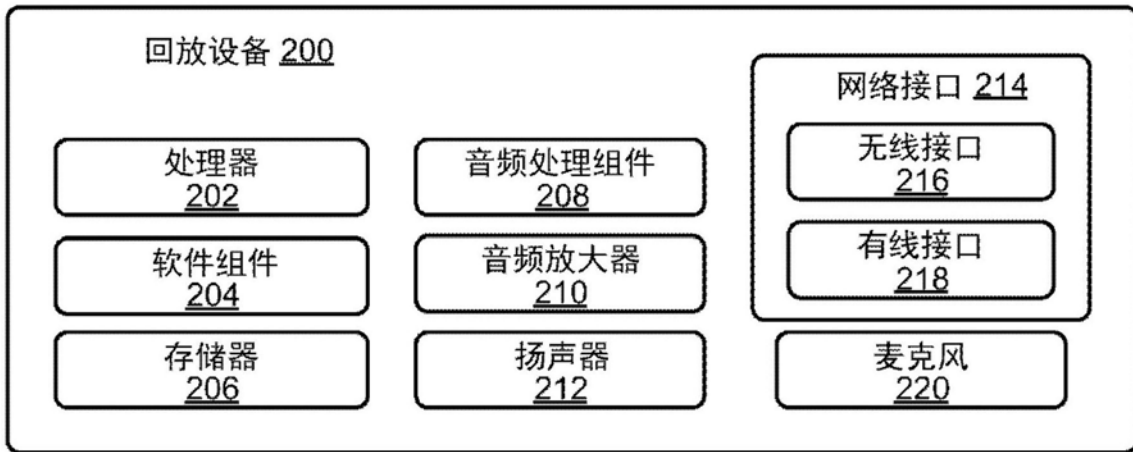


图2

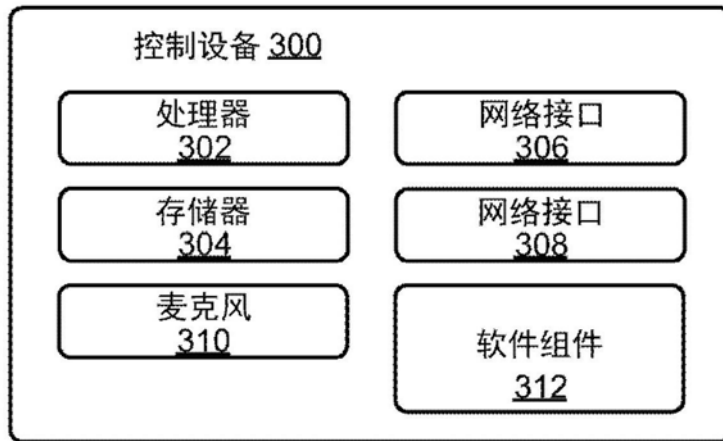


图3

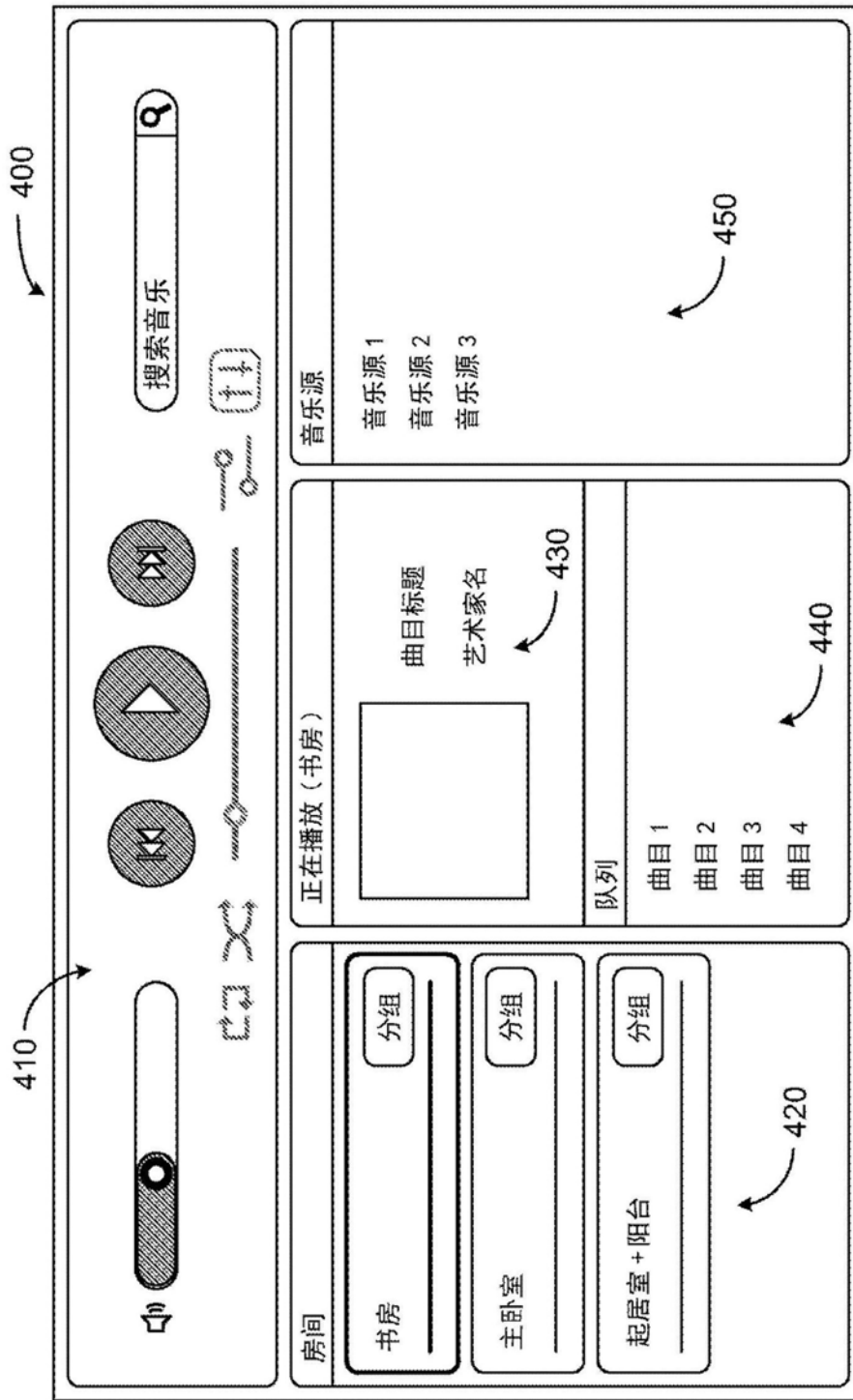


图4

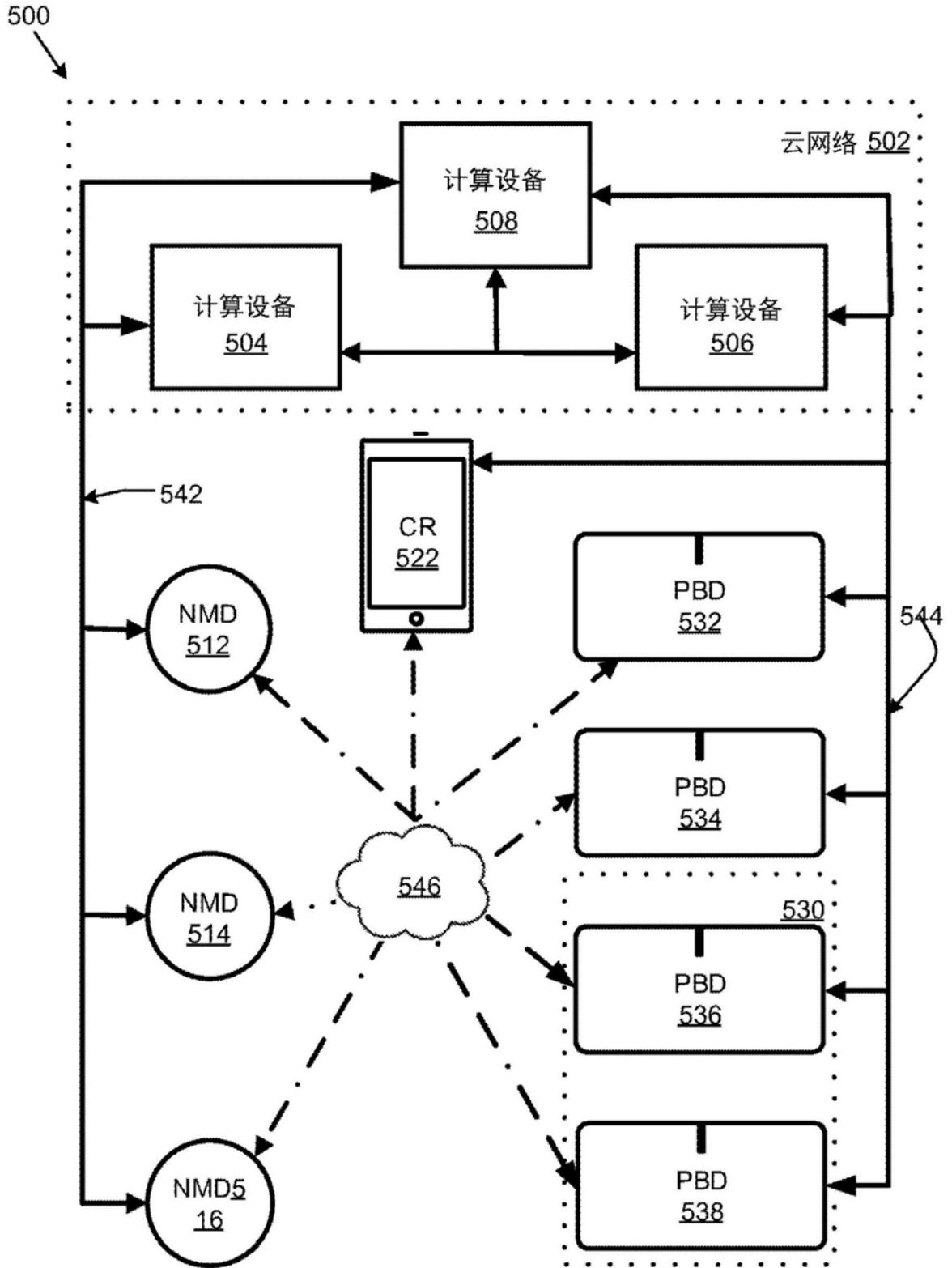


图5

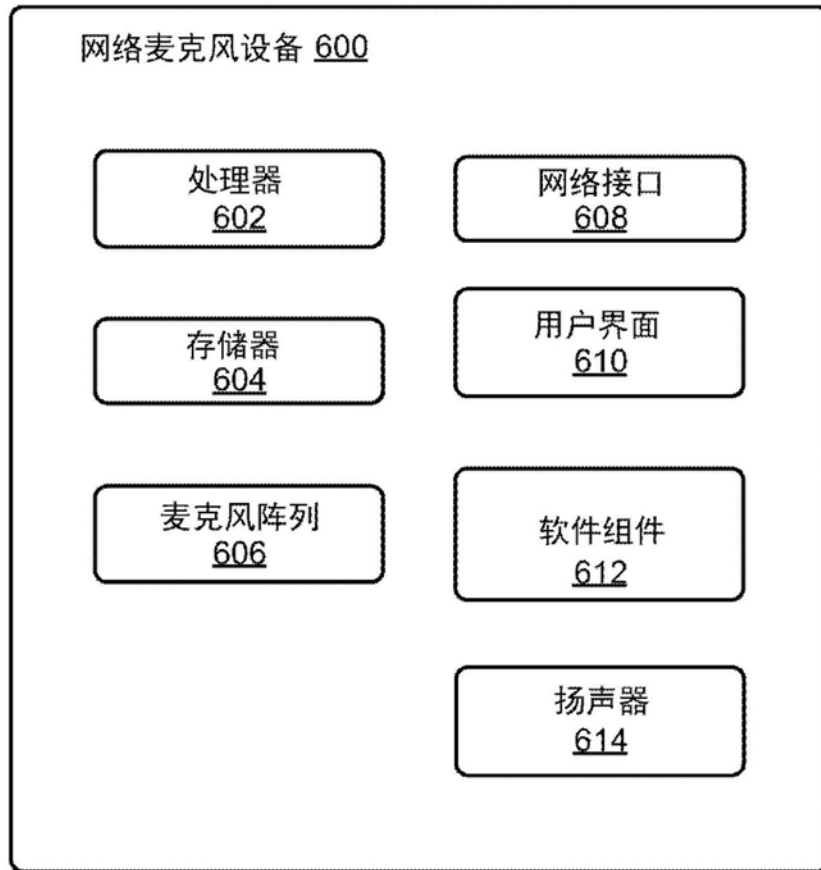


图6

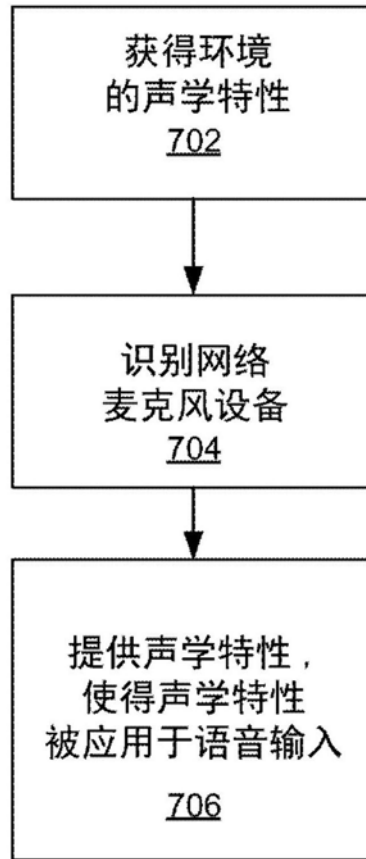


图7



图8

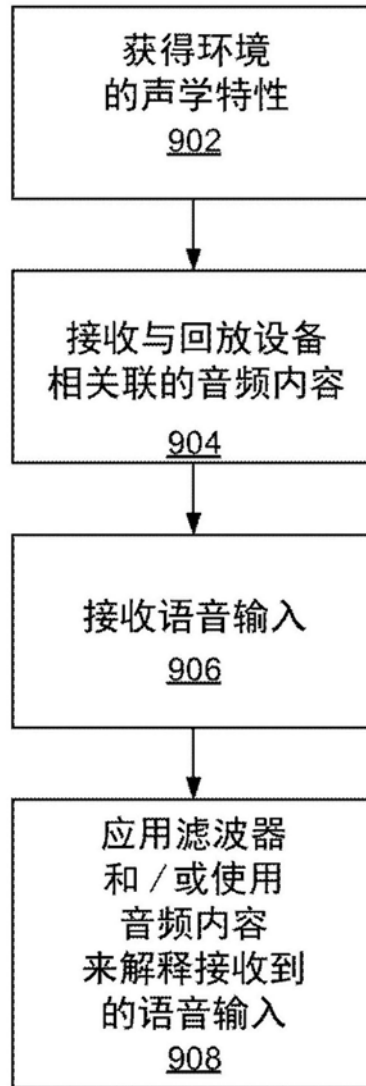


图9



图10

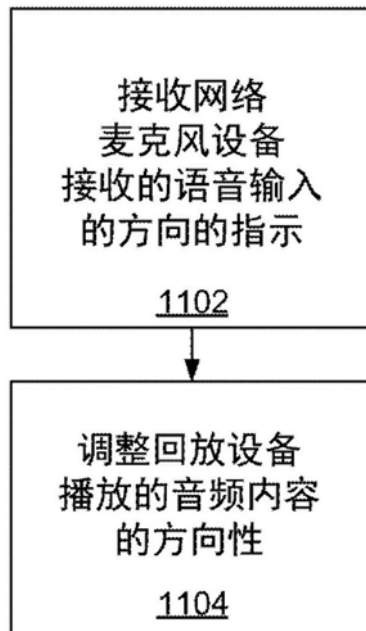


图11

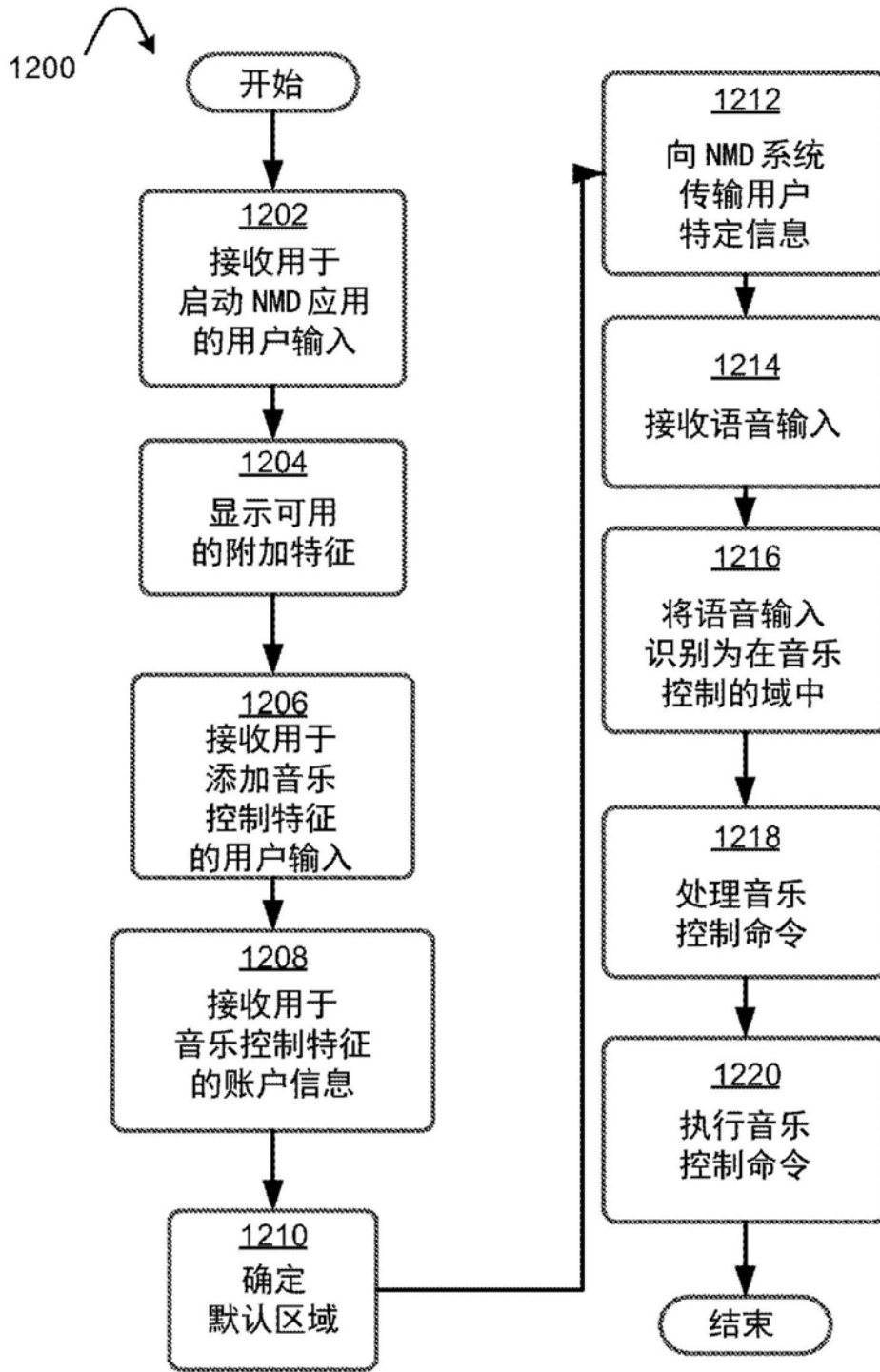


图12

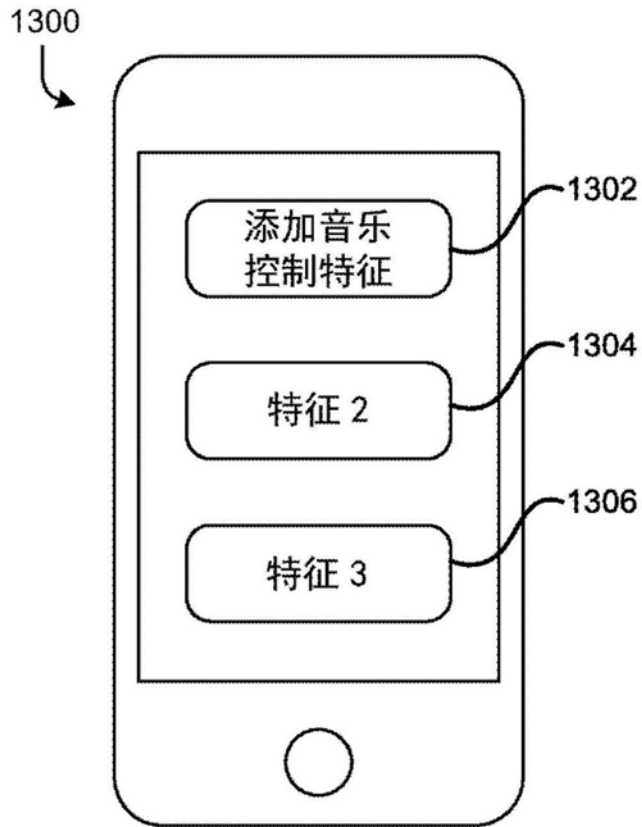


图13A

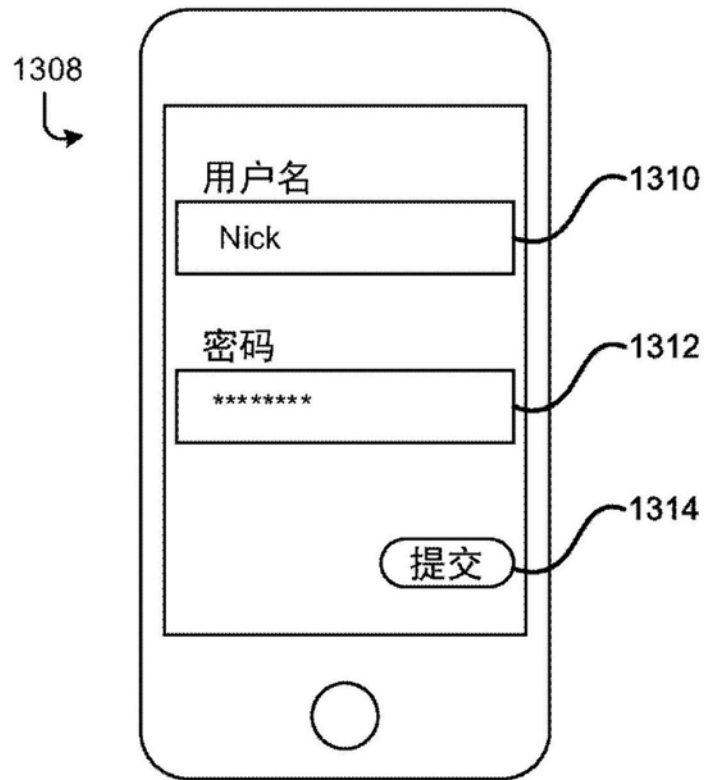


图13B

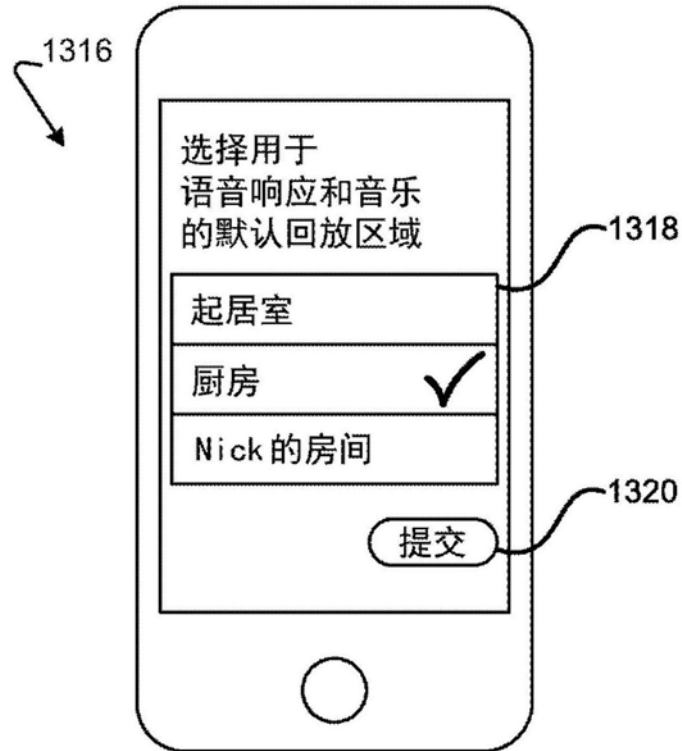


图13C

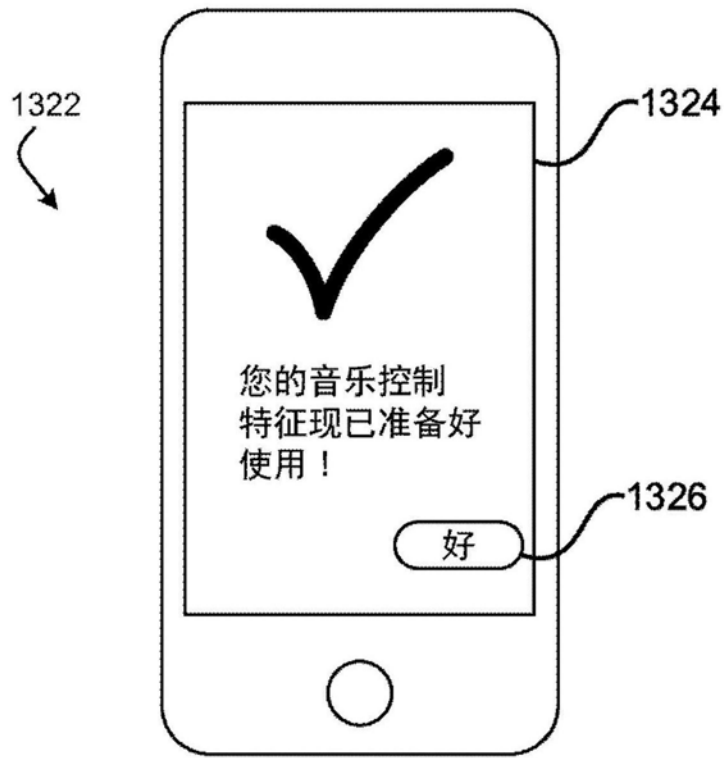


图13D

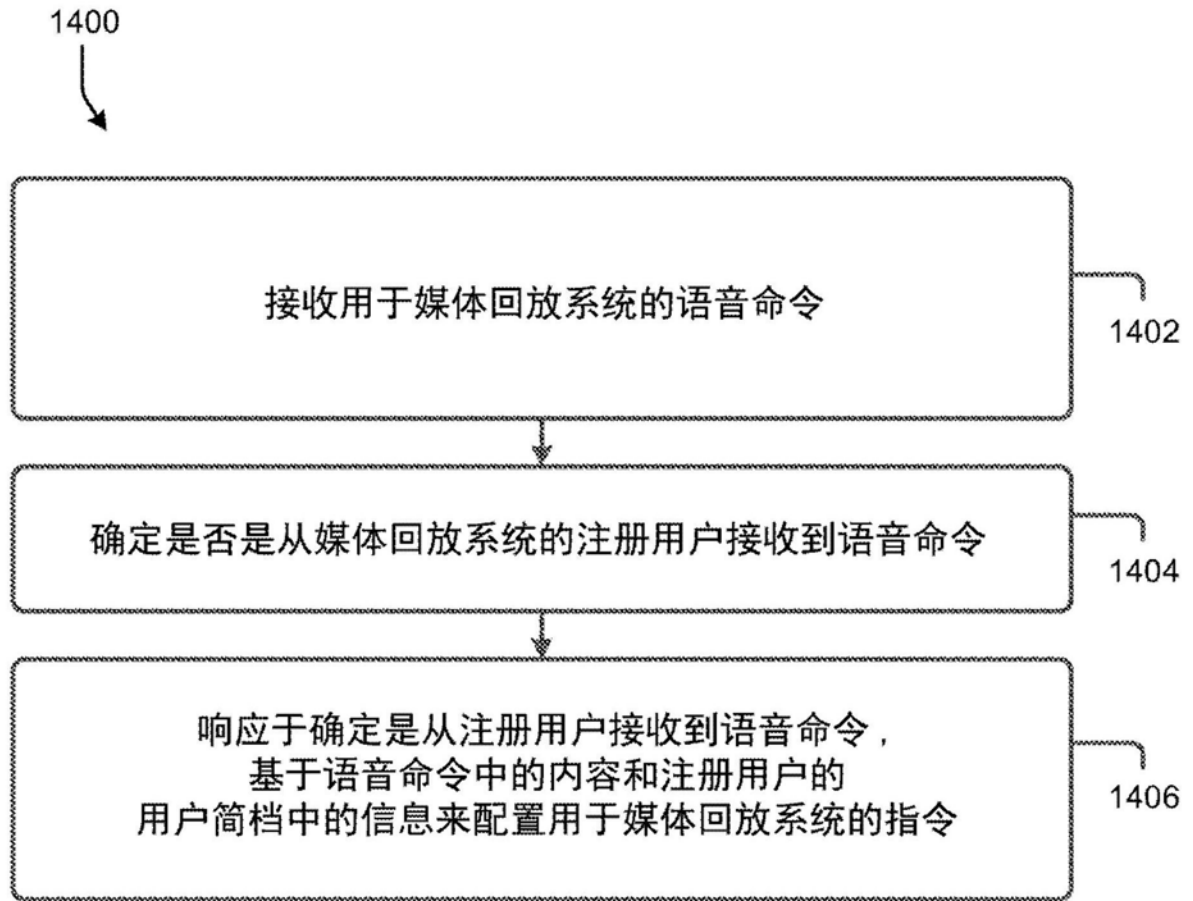


图14

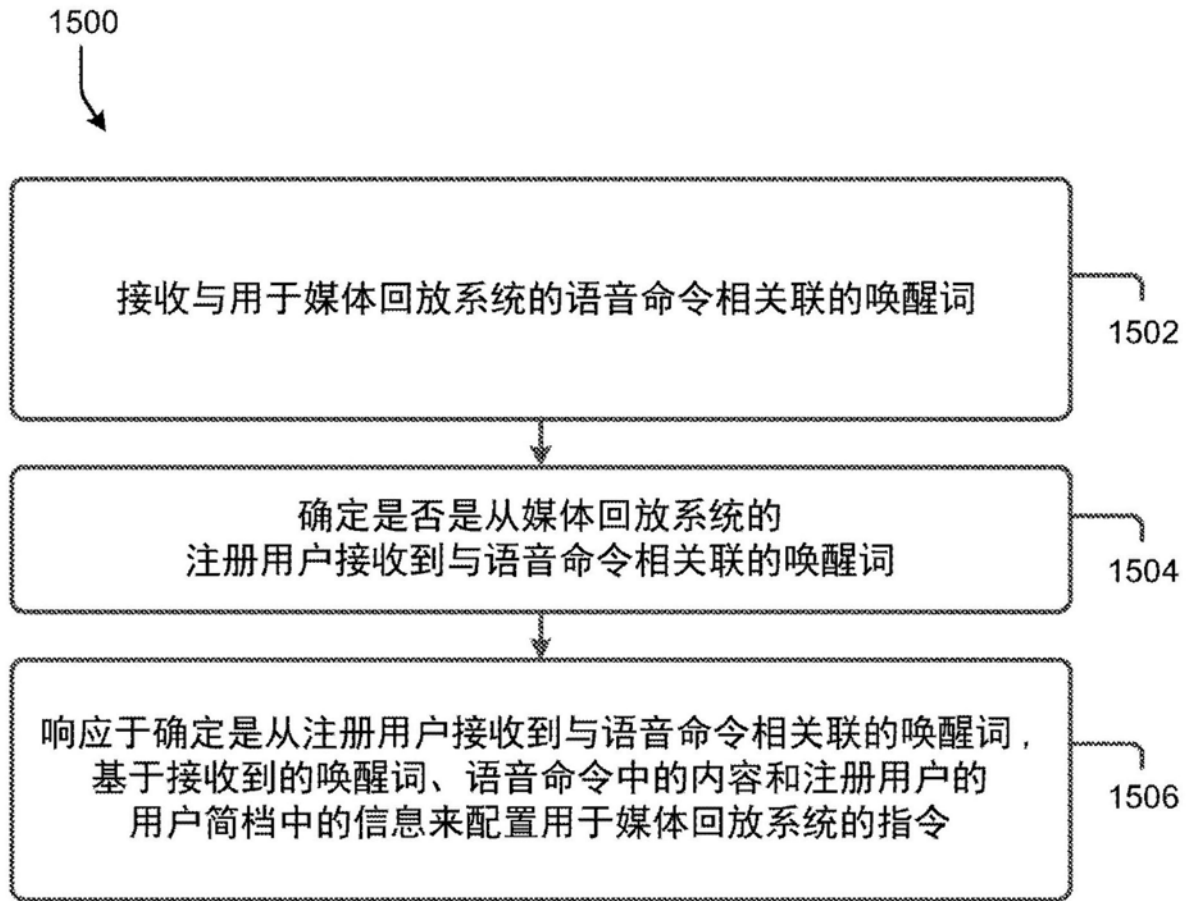


图15

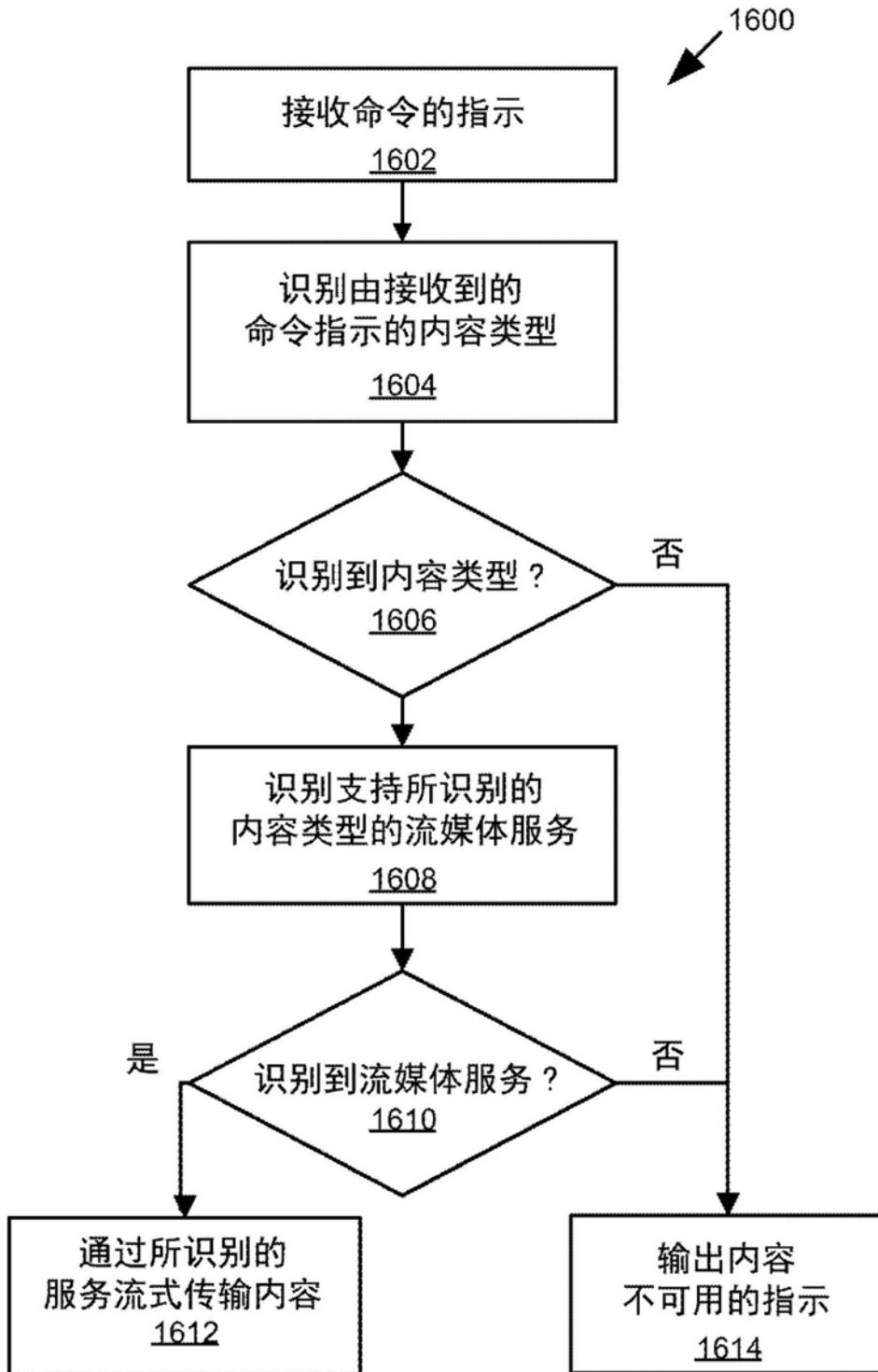


图16

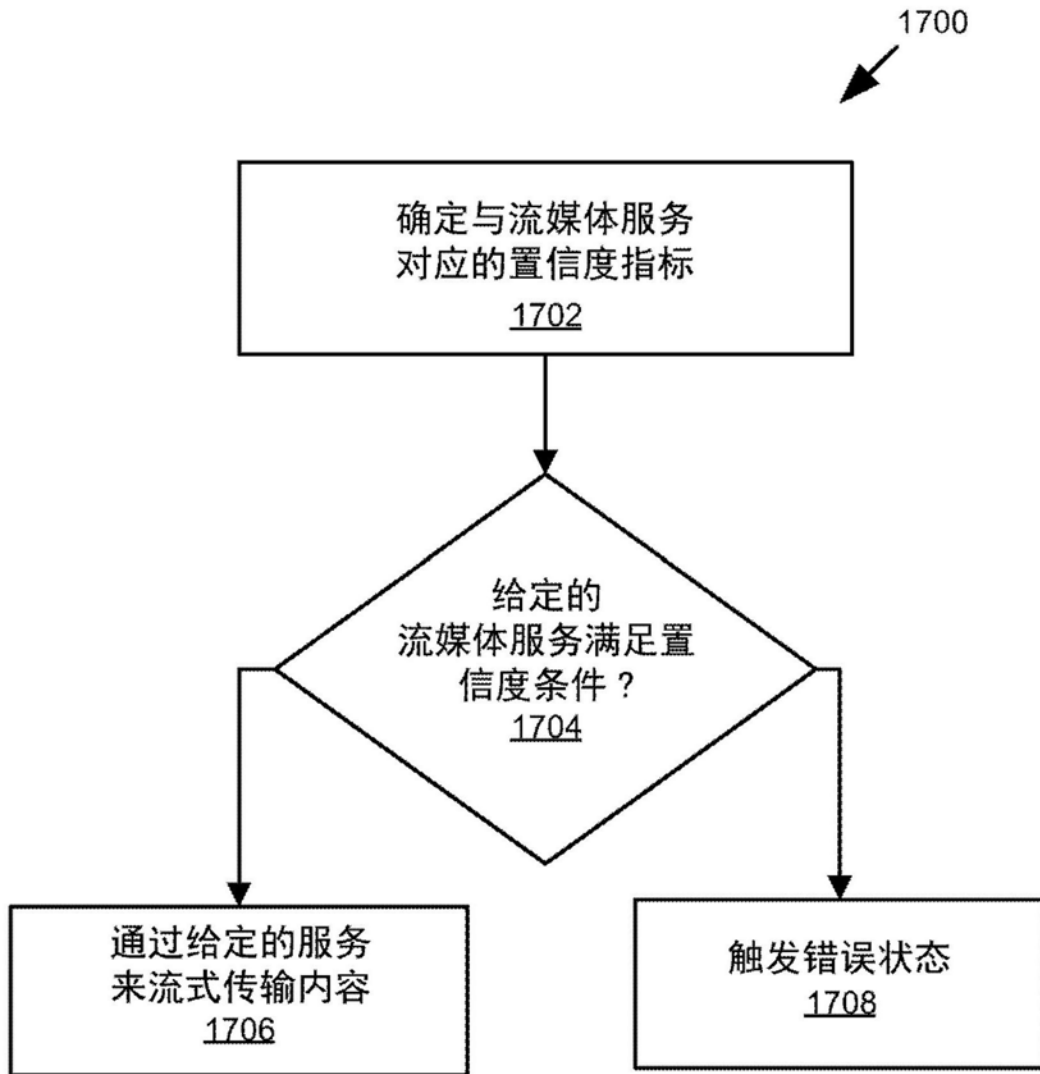


图17