



(12)发明专利申请

(10)申请公布号 CN 109844857 A

(43)申请公布日 2019.06.04

(21)申请号 201780063647.0

(74)专利代理机构 北京市柳沈律师事务所
11105

(22)申请日 2017.10.16

代理人 陈丽来

(30)优先权数据

62/409,287 2016.10.17 US

(51)Int.Cl.

G10L 15/22(2006.01)

(85)PCT国际申请进入国家阶段日

2019.04.15

(86)PCT国际申请的申请数据

PCT/US2017/056828 2017.10.16

(87)PCT国际申请的公布数据

WO2018/075417 EN 2018.04.26

(71)申请人 哈曼国际工业有限公司

地址 美国康涅狄格州

(72)发明人 D.欧文斯

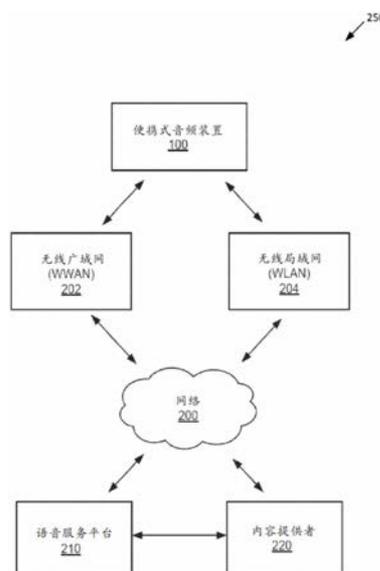
权利要求书3页 说明书14页 附图5页

(54)发明名称

具有语音能力的便携式音频装置

(57)摘要

一种便携式音频装置,包括能够与WLAN连接的网卡和连接到WWAN的无线调制解调器。所述便携式音频装置经由所述网卡和WLAN或所述无线调制解调器和WWAN与语音服务平台和/或内容提供者通信。如果所述便携式音频装置无法访问所述WLAN,所述便携式音频装置可通过经由所述无线调制解调器和WWAN与所述语音服务平台通信来处理 and 响应语音查询。所述便携式音频装置还包括电池,所述电池为所述便携式音频装置的各种硬件和软件部件提供电力以执行诸如高级语音功能等各种功能。所述便携式音频装置提供了真正的便携性,并且可以在诸如在家庭或建筑物环境内或在家庭或建筑物环境之外的任何环境中使用。



1. 一种便携式音频装置,其包括:
 - 一组一个或多个远场麦克风,所述一组一个或多个远场麦克风被配置成获取语音查询;
 - 一组一个或多个扬声器;
 - 无线调制解调器,所述无线调制解调器被配置成连接到无线广域网(WWAN);
 - 存储器,所述存储器包括音频处理应用;以及
 - 处理器,所述处理器联接到所述存储器并且在执行所述音频处理应用时被配置成:
 - 基于所述语音查询产生数据请求;
 - 经由所述无线调制解调器向所述WWAN传送所述数据请求;
 - 经由所述无线调制解调器从所述WWAN接收与所述数据请求相关联的数据响应;
 - 基于所述数据响应产生音频信号;以及
 - 将所述音频信号传送到所述一组扬声器以产生音频输出。
2. 如权利要求1所述的便携式音频装置,其还包括电池,所述电池用于对所述一组麦克风、所述一组扬声器和所述无线调制解调器中的至少一者供电。
3. 如权利要求1所述的便携式音频装置,其中所述处理器被配置成通过使用自然语言处理技术处理所述语音查询来基于所述语音查询产生所述数据请求。
4. 如权利要求1所述的便携式音频装置,其中所述一组扬声器包括用于第一频率范围的第一驱动器和用于第二频率范围的第二驱动器。
5. 如权利要求1所述的便携式音频装置,其中:
 - 所述语音查询包括对媒体项或信息项的请求;并且
 - 所述数据响应包括所述媒体项或信息项。
6. 如权利要求1所述的便携式音频装置,其还包括订户身份模块(SIM)装置,其中所述处理器还被配置成基于经由所述SIM装置接收的信息与所述WWAN通信。
7. 如权利要求1所述的便携式音频装置,其还包括:网卡,所述网卡用于连接到无线局域网(WLAN),其中所述处理器还被配置成:
 - 在产生所述数据请求之前,连续地监视将由所述一组麦克风获取的唤醒命令;以及
 - 响应于检测到所述唤醒命令,确定经由所述无线调制解调器连接到所述WWAN并且不由所述网卡连接到所述WLAN。
8. 如权利要求1所述的便携式音频装置,其中所述处理器还被配置成:
 - 在产生所述数据请求之前,在所述便携式音频装置处于低功率状态时,连续地监视将由所述一组麦克风获取的唤醒命令;以及
 - 响应于检测到所述唤醒命令,将所述便携式音频装置从所述低功率状态转换到正常功率状态。
9. 如权利要求8所述的便携式音频装置,其中:
 - 在所述低功率状态中,用于监视所述唤醒命令的所述便携式音频装置的第一组部件被供电,其中不用于监视所述唤醒命令的所述便携式音频装置的第二组部件被断电;以及
 - 在所述正常功率状态中,所述第二组部件中的至少一个部件被供电。
10. 一种包括指令的计算机可读存储介质,所述指令在由处理器执行时使所述处理器通过执行以下步骤来处理语音查询:

- 经由便携式音频装置的一组一个或多个远场麦克风来获取所述语音查询；
基于所述语音查询产生数据请求；
经由所述便携式音频装置中所包括的无线调制解调器将所述数据请求传送到无线广域网 (WWAN) ；
经由所述无线调制解调器从所述WWAN接收与所述数据请求相关联的数据响应；
基于所述数据响应产生音频信号；以及
将所述音频信号传送到包括在所述便携式音频装置中的一组扬声器以产生音频输出。
11. 如权利要求10所述的计算机可读存储介质，其中所述便携式音频装置包括电池，所述电池用于在低功率状态中对所述便携式音频装置的第一组部件供电且在正常功率状态中对所述便携式音频装置的第二组部件供电。
12. 如权利要求10所述的计算机可读存储介质，其中基于所述语音查询产生所述数据请求包括使用自然语言处理技术来处理所述语音查询。
13. 如权利要求10所述的计算机可读存储介质，其中所述一组扬声器包括分频器、用于第一频率范围的第一驱动器和用于第二频率范围的第二驱动器。
14. 如权利要求10所述的计算机可读存储介质，其中：
所述语音查询包括对基于音频的媒体项或信息项的请求；以及
所述数据响应包括所述基于音频的媒体项或信息项。
15. 如权利要求10所述的计算机可读存储介质，其中所述便携式音频装置包括订户身份模块 (SIM) 装置，所述便携式音频装置被配置成基于经由所述SIM装置接收的信息与所述WWAN通信。
16. 如权利要求10所述的计算机可读存储介质，其中所述便携式音频装置包括用于连接到无线局域网 (WLAN) 的网络适配器，还包括在经由所述无线调制解调器传送所述数据请求之前，确定所述WLAN不在范围内的步骤。
17. 如权利要求10所述的计算机可读存储介质，其还包括以下步骤：
在产生所述数据请求之前，在所述便携式音频装置处于低功率状态时，连续地监视将由所述一组麦克风获取的唤醒命令；以及
响应于检测到所述唤醒命令，将所述便携式音频装置从所述低功率状态转换到正常功率状态。
18. 如权利要求17所述的计算机可读存储介质，其中：
在所述低功率模式中，所述一组麦克风中的第一麦克风子组被供电以监视所述唤醒命令，其中所述一组麦克风中的第二麦克风子组被断电；以及
在所述正常功率模式中，所述第二麦克风子组被供电。
19. 一种便携式音频装置，其包括：
一组一个或多个麦克风，所述一组一个或多个麦克风被配置成获取语音查询；
一组一个或多个扬声器；
无线调制解调器，所述无线调制解调器被配置成连接到无线广域网 (WWAN) ；
存储器，所述存储器包括音频处理应用；以及
处理器，所述处理器联接到所述存储器并且在执行所述音频处理应用时被配置成：
基于所述语音查询产生数据请求；

经由所述无线调制解调器向所述WWAN传送所述数据请求；
经由所述无线调制解调器从所述WWAN接收与所述数据请求相关联的数据响应；
基于所述数据响应产生音频信号；以及
将所述音频信号传送到所述一组扬声器以基于所述音频信号产生声音，
其中所述便携式音频装置不包括视频显示屏。

20. 如权利要求19所述的便携式音频装置，其中所述便携式音频装置不包括视频处理能力。

21. 如权利要求19所述的便携式音频装置，其还包括连接到无线局域网 (WLAN) 的网卡，其中所述处理器还被配置成在连接到所述无线广域网 (WWAN) 之前，确定所述WWAN的信号强度大于所述WLAN的信号强度。

22. 如权利要求21所述的便携式音频装置，其中所述处理器还被配置成：
检测触发事件；以及

响应于检测到所述触发事件，确定所述WLAN的信号强度大于所述WWAN的信号强度并与所述WLAN建立网络连接。

23. 如权利要求22所述的便携式音频装置，其中所述触发事件包括以下至少一者：预定时间段的期满、当前连接的网络的信号强度下降到预定阈值水平以下、所述用户的调度事件，以及所述便携式音频装置的移动超过预定阈值距离。

具有语音能力的便携式音频装置

[0001] 相关申请的交叉引用

[0002] 本申请要求2016年10月17日提交并且具有序列号62/409,287的题为“PORTABLE AUDIO DEVICE WITH VOICE CAPABILITIES (具有语音能力的便携式音频装置)”的美国临时专利申请的权益。所述相关申请的主题在此以引用的方式并入本文中。

技术领域

[0003] 本公开的实施方案总体上涉及音频装置,并且更具体地,涉及具有语音能力的便携式音频装置。

背景技术

[0004] 能够接收和处理语音查询的语音助理装置(VAD)正变得越来越受用户欢迎。流行的VAD的示例包括亚马逊回声(Amazon Echo)和谷歌家庭(Google Home)。VAD为用户提供了方便的基于语音的接口,以从互联网请求特定项(诸如数字媒体项或信息项)并经由VAD的扬声器接收来自VAD的响应。例如,用户可以提交针对特定的基于音频的媒体项(诸如音乐曲目、流播客或有声读物)的口头查询。VAD处理口头命令并连接到局域网(LAN)以从互联网检索所请求的媒体项,然后经由扬声器输出所请求的媒体项。为了处理用户查询,VAD还可利用与一个或多个内容提供者交互以检索所请求的媒体项或信息项的语音服务平台(VSP)。

[0005] 上述语音助理装置的一个缺点是装置仅在连接到LAN(诸如家庭、学校或工作场所)时是可用的。虽然无线LAN(诸如Wi-Fi)通常与VAD一起使用,但VAD必须在无线局域网(WLAN)的相对短的范围以内可用。此外,VAD通常经由连接到房屋或建筑物的墙壁插座的AC电源插头来供电。因此,由于所需的WLAN连接的短的本地覆盖区域和电源插头和插座的要求,常规的VAD基本上系留到房屋或建筑物的有限区域。缺乏常规VAD的便携性限制了对家庭或建筑物环境的VAD的使用,由此VAD通常在这样的有限环境之外不可用。

[0006] 如上所述,在家庭或建筑物之外的环境中使用VAD的更有效的技术将是有用的。

发明内容

[0007] 各种实施方案阐述一种便携式音频装置,所述便携式音频装置包括被配置成获取语音查询的一组一个或多个远场麦克风、被配置成连接到无线广域网(WWAN)的无线调制解调器、包括音频处理应用的存储器,以及联接到所述存储器的处理器,并且当执行音频处理应用时,所述处理器被配置成执行一组步骤。所述一组步骤包括基于所述语音查询产生数据请求并且经由所述无线调制解调器向所述WWAN传送所述数据请求。所述一组步骤还包括经由所述无线调制解调器从所述WWAN接收与所述数据请求相关联的数据响应。所述一组步骤还包括基于所述数据响应产生音频信号且将所述音频信号传送到扬声器组以产生音频输出。

[0008] 其他实施方案包括但不限于包括用于执行所公开的技术的一个或多个方面的指

令的计算机可读介质,以及用于执行所公开的技术的一个或多个方面的方法。

[0009] 所公开的方法的至少一个优点是,即使当便携式音频装置不在WLAN的范围内并且没有从墙壁插座接收电力时,便携式音频装置也能够处理语音查询并提供对语音查询的响应。因此,便携式音频装置提供了真正的便携性并且可以在任何环境中(诸如在家庭或建筑物环境内或在家庭或建筑物环境之外)使用。

附图说明

[0010] 为了可以详细地理解上述一个或多个实施方案的所述特征的方式,可以通过参考某些具体实施方案来对以上简要概述的一个或多个实施方案进行更具体的描述,其中一些实施方案在附图中示出。然而,应注意,附图仅示出了典型实施方案,且因此不应被视为以任何方式限制其范围,因为各种实施方案的范围也包含其他实施方案。

[0011] 图1是被配置成实现各种实施方案的一个或多个方面的便携式音频装置的概念框图;

[0012] 图2是根据各种实施方案的一个或多个方面的其中可以实现图1的便携式音频装置的通信环境的概念图示;

[0013] 图3是根据各种实施方案的用于执行便携式音频装置的功能的方法步骤的流程图;

[0014] 图4是根据各种实施方案的用于处理由便携式音频装置接收的语音查询的方法步骤的流程图;以及

[0015] 图5是根据各种实施方案的用于确定是否连接到WWAN或WLAN的方法步骤的流程图。

具体实施方式

[0016] 在以下描述中,阐述了许多具体细节以提供对某些具体实施方案的更透彻理解。然而,对于本领域技术人员将显而易见的是,可以在没有这些具体细节中的一个或多个的情况下或在附加的具体细节的情况下实践其他实施方案。

[0017] 图1是被配置成实现各种实施方案的一个或多个方面的便携式音频装置100的概念框图。如图所示,便携式音频装置100包括处理器102、输入/输出(I/O)装置104(包括网卡150和无线调制解调器155)、订户身份模块(SIM)装置106、存储器110、一组一个或多个扬声器120、一组一个或多个麦克风130、运动装置175以及通过总线160互连的一个或多个电池140。总线160共同地表示通信地连接便携式音频装置100的许多内部装置(在图1中示出)的所有系统、外围装置和芯片组总线。存储器110包括被配置成与数据库114交互的音频处理应用112。电池140作为整体向便携式音频装置100提供电力。在一些实施方案中,电池140基于便携式音频装置100的功率状态向包括处理器102、网卡150、无线调制解调器155、SIM装置106、存储器110、一组扬声器120、一组麦克风130或其任意组合的各种硬件和/或软件部件提供电力。

[0018] 如本文所述,便携式音频装置100包括能够经由无线广域网(WWAN)与语音服务平台通信的无线调制解调器155。在一些实施方案中,当便携式音频装置不能够访问本地网络(例如,无线局域网或“WLAN”)时,便携式音频装置100可以通过经由WWAN与语音服务平台通

信来处理 and 响应语音查询。因此,便携式音频装置100允许在诸如便携式扬声器的专用音频装置中实现高级语音功能,而无需便携式音频装置100连接到用户的本地网络或移动装置(例如,智能电话)。此外,便携式音频装置100可选地使得能够例如响应于用户发出的语音请求从一个或多个内容提供者接收内容。便携式音频装置100还包括电池140,其提供便携式音频装置100的硬件和软件部件所需的所有功率以执行本文描述的功能(诸如高级语音功能)。因此,便携式音频装置100不需要来自房屋或建筑物的墙壁插座的电力来提供本文描述的高级语音功能。以此方式,便携式音频装置提供真正的便携性,并且可在诸如在家庭或建筑物环境内或在家庭或建筑物环境之外的任何环境中使用。

[0019] 通常,处理器102被配置成协调便携式音频装置100的整体操作。处理器102可以是配置成处理数据和/或执行程序代码的任何技术上可行的形式的处理装置。处理器102可以包括例如但不限于片上系统(SoC)、中央处理单元(CPU)、图形处理单元(GPU)、专用集成电路(ASIC)、数字信号处理器(DSP)、现场可编程门阵列(FPGA)等。处理器102包括一个或多个处理核。在操作中,处理器102是便携式音频装置100的主处理器,控制和协调其他系统部件的操作。

[0020] I/O装置104可以包括输入装置、输出装置和既能够接收输入也能够提供输出的装置。例如但不限于,I/O装置104可以包括向移动计算装置、移动电话、(一个或多个)扬声器120、(一个或多个)麦克风130、远程数据库、其他类型的音频装置、其他类型的计算装置等发送数据和/或从移动计算装置、移动电话、(一个或多个)扬声器120、(一个或多个)麦克风130、远程数据库、其他类型的音频装置、其他类型的计算装置等接收数据的有线和/或无线通信装置等。在一些实施方案中,I/O装置104可以包括用于与一个或多个其他装置无线通信的Bluetooth®装置。

[0021] 另外,在一些实施方案中,I/O装置104包括网卡150和无线调制解调器155。在一些实施方案中,网卡150与第一类型的网络(非蜂窝网络)连接,并且无线调制解调器155与第二类型的网络(蜂窝网络)连接,第一类型的网络和第二类型的网络包括不同类型的网络。网卡150可以包括网络接口控制器(NIC)、网络适配器、LAN适配器、物理网络接口或可将便携式音频装置100连接到诸如局域网(LAN)或无线局域网(WLAN)的非蜂窝网络的类似的计算机硬件部件,用于与非蜂窝网络通信。无线调制解调器155可以包括移动宽带调制解调器、长期演进(LTE)调制解调器或可将便携式音频装置100连接到诸如无线广域网(WWAN)的蜂窝网络的类似的计算机硬件部件,用于与蜂窝网络通信。

[0022] SIM装置106与存储与一个或多个WWAN相关联的订户信息的一个或多个订户识别模块(SIM)卡对接。例如,SIM卡可以包括用于安全地存储用于使用便携式音频装置100识别和认证订户的国际移动订户身份(IMSI)号码及其相关密钥的集成电路。在一些实施方案中,SIM装置106与一个或多个I/O装置104(诸如无线调制解调器155)通信,以使(一个或多个)I/O装置104能够认证并连接到WWAN。在一些实施方案中,便携式音频装置100例如基于经由SIM装置106从SIM卡读取的订户信息经由无线调制解调器155连接到WWAN。

[0023] (一个或多个)麦克风130被配置成从周围环境获取声学数据并将与声学数据相关联的信号传送到音频处理应用112以进行处理。例如,由(一个或多个)麦克风130获取的声学数据可由便携式音频装置100(例如,由音频处理应用112)处理,以检测旨在将便携式音频装置100从低功率状态唤醒的“唤醒”命令。唤醒命令可以包括由音频处理应用112识别的

包括唯一名称、词语或短语的特定的“唤醒词”。便携式音频装置100可以预先配置有存储到存储器110的特定唤醒命令。唤醒命令也可以由用户改变为另一个唤醒词。响应于接收和检测到唤醒命令,音频处理应用112可以退出低功率状态(睡眠状态)并且将便携式音频装置100转换到用于接收和处理语音查询的正常功率状态(唤醒状态)。在便携式音频装置100被唤醒(转换到正常功率状态)之后,(一个或多个)麦克风130可以从用户获取包括语音查询的声学数据,并且将与语音查询相关联的信号传送到音频处理应用112以用于处理。

[0024] 在各种实施方案中,所述一组一个或多个麦克风130可以包括能够获取声学数据的任何类型的换能器,包括(例如且不限于)差分麦克风、压电麦克风、光学麦克风等。在一些实施方案中,麦克风130可以实现微波技术来检测特定环境中的声音(例如,语音)。在进一步的实施方案中,麦克风130可以实现远场技术以从特定距离(例如,高达25英尺远)获取/检测用户的语音(声学数据)。因此,用户需要相对靠近便携式音频装置100以供麦克风130从用户获取/检测唤醒命令和/或语音查询。例如,所述一组麦克风130可包括多个远场麦克风130。远场麦克风不同于近场麦克风,近场麦克风要求用户相对靠近近场麦克风说话(例如,在12英寸内),以便麦克风获得足够精确的声学数据用于语音处理。相比之下,远场麦克风可从用户在比可由近场麦克风获得的更大的距离下获取来自用户的足够准确的声学数据(口述语音)。例如,所述一组麦克风130可包括可从各种距离(例如,高达15英尺远、20英尺远或25英尺远)获取/检测用户的语音以用于语音处理目的的一个或多个远场麦克风。

[0025] 存储器110可以包括存储器模块或存储器模块的集合。存储器110内的音频处理应用112由处理器102执行以实现便携式音频装置100的整体功能,并且因此整体地协调便携式音频装置100的操作。由音频处理应用112执行的处理可以包括,例如但不限于,过滤、语音识别、使用自然语言处理(NLP)、统计分析、启发式处理、声学处理和/或其他类型的数据处理和分析。

[0026] 例如,在初始状态中,音频处理应用112可以在低功率状态下操作以节省电池140的功率。当处于低功率状态时,音频处理应用112可以处于“始终收听模式”,其连续地监视经由一个或多个麦克风130获取的声学数据以连续地监视唤醒命令。在一些实施方案中,当处于低功率状态时,音频处理应用112可以仅给监视唤醒命令所需的便携式音频装置100的硬件和/或软件部件供电并激活,而对监视唤醒命令不需要的便携式音频装置100的所有其他硬件和/或软件部件断电并去激活。例如,当处于低功率状态时,音频处理应用112可以仅对所述一组麦克风130、处理器102和存储器110供电并激活。在其他实施方案中,音频处理应用112可以对所述一组麦克风130中的第一麦克风子组供电,其中所述一组麦克风中的第二麦克风子组被断电。例如,在低功率状态期间,可以仅对所述一组麦克风130中一个或两个麦克风供电。当处于低功率状态时,音频处理应用112还可以断电并去激活I/O装置104(包括网卡150和无线调制解调器155)、SIM装置106、(一个或多个)扬声器120或其任意组合。

[0027] 当音频处理应用112检测到由一个或多个麦克风130获取的声学数据包括唤醒命令时,音频处理应用112可以通过对获取和处理语音查询所需的便携式音频装置100的硬件和/或软件部件供电并激活来将便携式音频装置100从低功率状态(睡眠状态)转换到正常功率状态(唤醒状态)。例如,在正常功率状态中,音频处理应用112可对在低功率状态期间

未供电的麦克风组130中的麦克风的第二子组、I/O装置104(包括网卡150和无线调制解调器155)、SIM装置106、(一个或多个)扬声器120或其任意组合供电并激活。因此,在低功率状态中,便携式音频装置100的较低数量的硬件和/或软件部件被供电,因此与正常功率状态(其相对于低功率状态需要便携式音频装置100的较高数量的硬件和/或软件部件被供电)相比,需要来自电池140的较低量的功率。

[0028] 在便携式音频装置100被唤醒(转换到正常功率状态)之后,音频处理应用112连续地监视经由(一个或多个)麦克风130接收的声学数据以检测语音查询。然后,(一个或多个)麦克风130可以获取包括语音查询的声学数据,并将与语音查询相关联的信号传送到音频处理应用112以进行处理。语音查询可以由音频处理应用112处理以产生包括IP数据流的数据请求,所述数据请求包括所接收的语音查询的基于计算机的表示。

[0029] 在一些实施方案中,音频处理应用112通过使用自然语言处理技术处理语音查询来基于语音查询产生数据请求。自然语言处理(NLP)可以包括人工智能和计算语言学的形式,以便于人类(自然)语言和计算机语言之间的交互。自然语言处理可以接收自然语言语音查询并将语音查询转换为计算机硬件和软件部件可以理解的基于计算机的数据请求。可以使用存储到数据库114的词典(术语词典)和语法/语法结构将自然语言处理应用于所接收的语音查询。

[0030] 然后,音频处理应用112经由一个或多个I/O装置104向语音服务平台和/或内容提供者传送数据请求。语音服务平台可以处理所接收的数据请求以产生数据响应。语音服务平台还可利用内容提供者的服务来产生数据响应。然后,音频处理应用112可以从语音服务平台和/或内容提供者接收针对数据请求的数据响应。基于所接收的数据响应,音频处理应用112产生包括音频信号的音频响应。然后,音频处理应用112将音频响应(音频信号)传送到所述一组扬声器120。

[0031] 所述一组扬声器120被配置成基于从便携式音频装置100和/或便携式音频装置100中包括的音频装置(例如,功率放大器)接收的音频信号产生声音(音频输出)。在一些实施方案中,所述一组扬声器120包括多个全范围扬声器。在进一步的实施方案中,所述一组扬声器120包括电子分频器和用于产生不同频率范围的多个不同类型的驱动器,每个驱动器用于产生特定的频率范围。例如,所述一组扬声器120可以包括1.25英寸和1.75英寸之间的第一驱动器和直径在3.25英寸和3.75英寸之间的第二驱动器(诸如低音喇叭)。在这些实施方案中,电子分频器包括电子滤波器电路,以将接收到的音频信号分离成两个或更多个频率范围,并且将音频信号传送到针对不同频率范围设计的不同驱动器。例如,一组扬声器120可以包括用于第一频率范围(例如,诸如小于120Hz的低频范围)的第一驱动器和用于第二频率范围(例如,更高频率范围,诸如大于120Hz)的第二驱动器,第一频率范围和第二频率范围是不同的频率范围。

[0032] 在一些实施方案中,从用户接收的语音查询可以包括回放特定媒体项的请求(例如,“播放音乐曲目X”)。在一些实施方案中,特定媒体项可以包括仅包括音频数据并且不包含视频数据的基于音频的媒体项(诸如音乐曲目、流播客、有声读物等)。语音查询由音频处理应用112处理(例如,使用NLP技术)以产生用于特定媒体项的数据请求(IP数据流)。音频处理应用112经由一个或多个I/O装置104将数据请求传送到语音服务平台。语音服务平台定位并检索特定媒体项(例如,通过利用内容提供者的服务)并且将特定媒体项作为数据响

应传送到音频处理应用112。替代地,语音服务平台可以将数据请求传送给内容提供者,内容提供者定位并检索特定媒体项,并将特定媒体项作为数据响应传送给音频处理应用112。音频处理应用112对接收到的数据响应(诸如解压缩、解码等)执行回放操作以产生包括音频信号的音频响应。音频信号被发送到(一个或多个)扬声器120,所述(一个或多个)扬声器120基于音频信号产生/再现声音(音频输出)以回放特定媒体项。

[0033] 作为另一示例,从用户接收的语音查询可以包括对特定信息项的请求。例如,信息项可以包括实时信息(诸如当前天气、交通、新闻、体育得分、股票报价等)或一般信息(诸如与历史、科学、健康、爱好等有关的一般事实)。例如,对信息项的语音查询可以包括请求,诸如:今天是什么天气?昨天红袜队比赛谁赢了?佛罗里达的州议会大厦是什么?第一次世界大战是什么时候结束的?语音查询由音频处理应用112处理(例如,使用NLP技术)以产生用于特定信息项的数据请求(IP数据流)。音频处理应用112经由一个或多个I/O装置104将数据请求传送到语音服务平台。语音服务平台定位并检索特定信息项(例如,通过利用内容提供者的服务)并且将特定信息项作为数据响应传送到音频处理应用112。替代地,语音服务平台可将数据请求传送到定位并检索特定信息项的内容提供者,且将所述特定信息项作为数据响应传送到音频处理应用112。音频处理应用112对数据响应执行转换操作,以基于数据响应产生音频响应。例如,音频响应可以包括对语音查询的口头语音响应的音频剪辑,所述语音响应包括表示数据响应的音频。音频信号被输出到(一个或多个)扬声器120,所述(一个或多个)扬声器120基于音频信号产生声音(音频输出)以回放对语音查询的语音响应。

[0034] 在各种实施方案中,便携式音频装置100的部件(例如,处理器102、(一个或多个)扬声器120、(一个或多个)麦克风130、电池140等)被布置在便携式外壳中,使得便携式音频装置100能够被携带到诸如房屋或建筑物环境的外部的不同的位置。在其他实施方案中,(一个或多个)扬声器120及(一个或多个)麦克风130可联接到便携式音频装置100的一个或多个部件,但与所述一个或多个部件分离。然而,本文中所揭示的实施方案涵盖被构造为实施便携式音频装置100的功能性的任何技术上可行的系统。

[0035] 在一些实施方案中,便携式音频装置100包括被构造为用于仅执行本文描述的功能的专用音频装置。在这些实施方案中,便携式音频装置100不包括视频显示屏或视频处理能力,而是仅提供音频能力,包括音频捕获、音频记录、音频处理、音频输出等。在一些实施方案中,便携式音频装置100不包括任何类型的电话呼叫能力。在进一步的实施方案中,便携式音频装置100不包括接收用户输入的任何硬按钮,由此通过音频/语音输入接收所有用户输入。通过不包括诸如视频显示屏、视频能力、电话呼叫能力、硬按钮或其任意组合之类的特征,可以显著降低便携式音频装置100的成本、复杂性和功率要求以产生高效且稳健的专用音频装置。在这方面,当便携式音频装置100持续需要来自电池140的电力来监视低功率模式中的唤醒命令时,视频显示屏和视频和电话呼叫能力的消除也可以帮助节省电池140的电力。另外,通过不包括通常易于损坏(例如,破裂)的视频显示屏,可以改善便携式音频装置100的整体耐用性。

[0036] 图2是根据各种实施方案的一个或多个方面的其中可以实现图1的便携式音频装置100的通信环境250的概念图示。如图所示,便携式音频装置100可以经由无线广域网(WWAN) 202和/或经由无线局域网(WLAN) 204与语音服务平台210和/或内容提供者220通信。

语音服务平台210和内容提供者220可经由通信网络200(诸如互联网)与无线广域网(WWAN) 202和/或无线局域网(WLAN) 204通信。

[0037] 在各种实施方案中,WWAN 202可以包括长期演进(LTE)网络、全球移动通信系统(GSM)网络、码分多址(CDMA)网络、全球微波接入互操作性(WiMAX)网络和/或任何其他类型的无线广域网。WLAN 204可包括任何类型的无线局域网,诸如Wi-Fi网络(例如,IEEE802.11网络)。通常,WLAN 204可以包括第一类型的网络(非蜂窝网络),并且WWAN 202可以包括第二类型的网络(蜂窝网络),第一类型的网络和第二类型的网络包括不同类型的网络。WWAN 202提供比WLAN 204大得多的覆盖范围并且需要不同的技术。特别地,WWAN 202利用移动通信蜂窝网络技术,所述移动通信蜂窝网络技术被区域性地、全国地或全球地提供,并且由无线服务提供者提供。相反,WLAN 204不利用移动通信蜂窝网络技术,并且通常在家庭、学校或工作场所环境中提供相对较短的覆盖范围。

[0038] 在一些实施方案中,响应于接收和检测到唤醒命令,音频处理应用112确定是否使用无线调制解调器155连接到WWAN 202。具体地,响应于接收到唤醒命令,音频处理应用112然后可以确定是使用无线调制解调器155连接到WWAN 202还是使用网卡150连接到WLAN 204。例如,音频处理应用112可以首先确定便携式音频装置100是否在WLAN 204的覆盖范围内。如果是,则音频处理应用112可以使用网卡150连接到WLAN 204。如果不是,则音频处理应用112可以使用无线调制解调器155连接到WWAN 202。

[0039] 在这些实施方案中,音频处理应用112确定是连接到WWAN 202还是WLAN 204,并响应于接收和检测到唤醒命令并立即与WWAN 202或WLAN 204建立网络连接。因此,在接收并检测到唤醒命令之后不久,便携式音频装置100就准备好使用与WWAN 202或WLAN 204的网络连接来接收和处理语音查询。例如,响应于接收和检测到唤醒命令,音频处理应用112可以通过确定便携式音频装置100在WLAN 204的覆盖范围之外来确定连接到WWAN 202并与WWAN 202建立网络连接。例如,响应于接收和检测到唤醒命令,音频处理应用112可以通过确定便携式音频装置100在WLAN 204的覆盖范围之内来确定连接到WLAN 204并与WLAN 204建立网络连接。在进一步的实施方案中,便携式音频装置100还可以无缝地在WWAN 202和WLAN 204之间切换,或者在服务和/或音频输出中具有短的中断。

[0040] 如上所述,当处于低功率模式时,网卡150、无线调制解调器155和SIM装置106可以断电。在一些实施方案中,响应于在低功率模式中检测到唤醒命令,音频处理应用112确定是否连接到WWAN 202或WLAN 204,并且仅对与WWAN 202或WLAN 204建立网络连接所需的部件供电。例如,如果音频处理应用112确定连接到WWAN 202,则音频处理应用112可以仅对无线调制解调器155和/或SIM装置106供电并且使网卡150断电。因此,响应于在低功率状态中检测到唤醒命令并且确定连接到WWAN 202时,音频处理应用112可以仅对无线调制解调器155和/或SIM装置106供电并且使网卡150断电。作为另一示例,如果音频处理应用112确定连接到WLAN 204,则音频处理应用112可以仅对网卡150供电并且不对无线调制解调器155和/或SIM装置106供电(即,保持无线调制解调器155和/或SIM装置106断电)。因此,响应于在低功率状态中检测到唤醒命令并且确定连接到WLAN 204时,音频处理应用112可以仅对网卡150供电并且不对无线调制解调器155和/或SIM装置106供电。

[0041] 一旦建立到WWAN 202或WLAN 204的网络连接,便携式音频装置100就可以从用户接收语音查询并且处理语音查询以产生数据请求。然后,便携式音频装置100经由WWAN 202

或WLAN 204和通信网络200(诸如互联网)将数据请求传送到语音服务平台210和/或内容提供者220。然后,便携式音频装置100经由WWAN 202或WLAN 204和通信网络200从语音服务平台210和/或内容提供者220接收数据响应。通常,语音服务平台210可以与内容提供者220协同工作,以产生用于来自便携式音频装置100的数据请求的数据响应。语音服务平台210可以提供智能个人助理服务,诸如Amazon Alexa、Google Assistant、Apple Siri和Microsoft Cortana。内容提供者220可以包括网络搜索和内容服务器,例如Microsoft Bing服务器、Google服务器或Yahoo!服务器。

[0042] 例如,在检测到由用户发出的语音查询时,音频处理应用112可以将语音查询转换为数据请求,并且经由WWAN 202或WLAN 204将数据请求传送到语音服务平台210。然后,语音服务平台210可以处理数据请求以产生数据响应,所述数据响应经由WWAN 202或WLAN 204传送到便携式音频装置100。在一些实施方案中,在处理数据请求之后,语音服务平台210可产生传送到内容提供者220的响应。然后,内容提供者220可以向便携式音频装置100传送内容(例如,媒体项和/或其他类型的数据),所述内容包括针对数据请求的数据响应。然后,便携式音频装置100将基于从内容提供者220接收的内容(数据响应)来产生包括通过(一个或多个)扬声器120播放的音频信号的音频响应。

[0043] 图3是根据各种实施方案的用于执行便携式音频装置100的功能的方法步骤的流程图。尽管结合图1-图2的系统描述了方法步骤,但是本领域技术人员将理解,被配置成以任何顺序执行方法步骤的任何系统都在本公开的范围内。在一些实施方案中,图3的方法300由在便携式音频装置100上执行的音频处理应用112执行。

[0044] 如图所示,方法300开始于步骤310,其中音频处理应用112最初启用便携式音频装置100的低功率状态以节省电池140的功率。在低功率状态中,音频处理应用112可以处于“始终收听模式”,其连续地监视经由一个或多个麦克风130获取的声学数据以检测唤醒命令。在一些实施方案中,当处于低功率状态时,音频处理应用112可以仅给监视唤醒命令所需的便携式音频装置100的硬件和/或软件部件供电并激活,而对监视唤醒命令不需要的便携式音频装置100的所有其他硬件和/或软件部件断电并去激活。在一些实施方案中,当处于低功率状态时,音频处理应用112对用于监视唤醒命令的便携式音频装置100的第一组部件供电,其中便携式音频装置100的第二组部件不用于监视唤醒命令被断电。

[0045] 然后,音频处理应用112确定(在步骤320)是否由一个或多个麦克风130获取并检测到唤醒命令。如果否,则方法300在步骤320处继续,由此音频处理应用112继续监视唤醒命令。如果获取并检测到唤醒命令,则方法300在步骤330处继续。在步骤330处,响应于检测到唤醒命令,音频处理应用112退出低功率状态(睡眠状态)并且将便携式音频装置100转换到正常功率状态(唤醒状态)。在一些实施方案中,在正常功率状态中,音频处理应用112对在低功率状态期间断电的便携式音频装置100的第二组部件供电。在其他实施方案中,在正常功率状态中,音频处理应用112对便携式音频装置100的第二组部件中的至少一个部件供电。

[0046] 在步骤340处,还响应于检测到唤醒命令,音频处理应用112确定是否连接到WWAN 202或WLAN 204,并与WWAN 202或WLAN 204建立网络连接。例如,音频处理应用112可以确定便携式音频装置100是否在WLAN 204的覆盖范围内。如果是,则音频处理应用112可以确定连接到WLAN 204并且使用网卡150与WLAN 204建立网络连接。如果不是,则音频处理应用

112可以确定连接到WWAN 202并且使用无线调制解调器155与WWAN 202建立网络连接。在一些实施方案中,如果音频处理应用112确定连接到WWAN 202,则音频处理应用112可以仅对无线调制解调器155和/或SIM装置106供电并且使网卡150断电。此外,如果音频处理应用112确定连接到WLAN 204,则音频处理应用112可以仅对网卡150供电并且不对无线调制解调器155和/或SIM装置106供电。

[0047] 在与WWAN 202或WLAN 204建立网络连接之后,音频处理应用112确定(在步骤350)是否已经由一个或多个麦克风130获取并检测到来自用户的语音查询。如果否,则方法300在步骤370处继续。如果是,则在步骤360处,音频处理应用112处理语音查询并且经由所述一组扬声器120向用户提供音频响应。下面结合图4更详细地讨论步骤360。

[0048] 在步骤370,音频处理应用112确定是否超过预定时间阈值,由此音频处理应用112没有检测到语音查询。通常,如果在某个时间段内没有检测到语音查询,则音频处理应用112可以将便携式音频装置100从正常功率状态转换到低功率状态,以节约电池140的功率。例如,如果音频处理应用112超过5分钟未检测到语音查询,则便携式音频装置100可以返回到低功率状态。如果尚未超过时间阈值,则方法300在步骤350处继续,由此音频处理应用112连续地监视语音查询。如果已经超过时间阈值,则方法300在步骤310处继续,由此音频处理应用112将便携式音频装置100转换回低功率状态。

[0049] 图4是根据各种实施方案的用于处理由便携式音频装置100接收的语音查询的方法步骤的流程图。尽管结合图1-图2的系统描述了方法步骤,但是本领域技术人员将理解,被配置成以任何顺序执行方法步骤的任何系统都在本公开的范围内。在一些实施方案中,图4的方法400由在便携式音频装置100上执行的音频处理应用112执行。方法400可以包括图3的步骤360。

[0050] 如图所示,方法400从图3的步骤350开始,由此接收来自用户的语音查询。在步骤410处,音频处理应用112基于语音查询产生数据请求(例如,使用NLP技术)。数据请求可以包括IP数据流并且表示所接收的语音查询。然后,音频处理应用112经由WWAN 202或WLAN 204向语音服务平台210和/或内容提供者220传送(在步骤420处)数据请求。如图2中所示,语音服务平台210和内容提供者220可经由通信网络200(诸如互联网)与WWAN 202或WLAN 204通信。语音服务平台210可以处理数据请求以产生数据响应。语音服务平台还可利用内容提供者220的服务来产生数据响应。

[0051] 然后,音频处理应用112从语音服务平台210和/或内容提供者220接收(在步骤430处)用于数据请求的数据响应。基于所接收的数据响应,音频处理应用112然后产生(在步骤440)包括音频信号的音频响应。然后,音频处理应用112将音频响应传送(在步骤450)到再现/回放音频响应的(一个或多个)扬声器120。然后,方法400返回到图3的步骤370。

[0052] 图5是根据各种实施方案的用于确定是否连接到WWAN或WLAN的方法步骤的流程图。尽管结合图1-图2的系统描述了方法步骤,但是本领域技术人员将理解,被配置成以任何顺序执行方法步骤的任何系统都在本公开的范围内。在一些实施方案中,图5的方法500由在便携式音频装置100上执行的音频处理应用112执行。

[0053] 方法500包括用于确定是连接到WWAN还是WLAN的替代或补充实施方案(而不是结合图1-图4描述的实施方案)。在所述实施方案中,图5的方法500可以与图3的方法300和图4的方法400并行地执行。方法500可包括在后台执行以连续地监视WWAN或WLAN的信号强度并

基于它们各自的信号强度连接到WWAN或WLAN的单独方法。以这种方式,方法500可连续地监视WWAN或WLAN的信号强度,且在正常功率模式中从用户接收语音查询的同时无缝地转换WWAN与WLAN之间的网络连接。在这样的实施方案中,当在图3的方法300的步骤320-是处接收到唤醒命令时启动/触发图5的方法500。此外,在这样的实施方案中,不执行图3的方法300的步骤340(确定是否连接到WWAN 202或WLAN 204)。相反,如果需要,图5的方法500可以在后台连续地并行执行,以连续地监视WWAN或WLAN的信号强度以及WWAN或WLAN之间的转换。

[0054] 如图所示,当音频处理应用112接收到唤醒命令时,方法500开始于步骤510。如上所述,在一些实施方案中,步骤510对应于图3的方法300的步骤320-是。在步骤520处,音频处理应用112然后确定WWAN 202的第一信号强度和WLAN 204的第二信号强度。第一信号强度可以包括经由无线调制解调器155确定的便携式音频装置100和WWAN 202之间的信号强度。第二信号强度可以包括经由网卡150确定的便携式音频装置100和WLAN 204之间的信号强度。在步骤530处,音频处理应用112然后比较WWAN 202的第一信号强度和WLAN 204的第二信号强度,以确定哪个信号强度更大。

[0055] 在步骤540处,音频处理应用112基于第一信号强度和第二信号强度的比较来与WWAN 202或WLAN 204建立网络连接。例如,如果在步骤530处确定WWAN 202的第一信号强度大于WLAN 204的第二信号强度,则音频处理应用112可以与WWAN 202建立网络连接。相反,如果在步骤530确定WLAN 204的第二信号强度大于WWAN 202的第一信号强度,则音频处理应用112可以与WLAN 204建立网络连接。

[0056] 在步骤550处,音频处理应用112然后确定是否检测到触发事件。触发事件的检测使得音频处理应用112再次基于它们各自的信号强度来评估和确定是否连接到WWAN 202或WLAN 204。如果在步骤550处检测到触发事件,则方法500在步骤520处继续,由此音频处理应用112确定WWAN 202和WLAN 204的信号强度,比较信号强度,并且根据哪个网络具有更大的信号强度来与WWAN 202或WLAN 204建立网络连接。如果在步骤550未检测到触发事件,则方法500在步骤550处继续,由此音频处理应用112继续监视触发事件。

[0057] 在一些实施方案中,触发事件可以包括预定时间段的期满。在这样的实施方案中,方法500可以连续地监视WWAN 202和WLAN 204的信号强度,并且基于它们在预定时间间隔处的相应信号强度来与WWAN 202或WLAN 204建立网络连接。

[0058] 在一些实施方案中,触发事件包括确定当前连接的网络的信号强度已经下降到预定阈值水平以下。例如,如果便携式音频装置100当前与WWAN 202连接,但是音频处理应用112检测到WWAN 202的信号强度已经低于阈值水平,则将检测到触发事件。在这些实施方案中,方法500可以连续地监视便携式音频装置100当前连接到的网络(WWAN 202或WLAN 204)的信号强度,以检测当前信号强度是否低于阈值水平。如果当前信号强度低于阈值水平,则音频处理应用112确定WWAN 202和WLAN 204的信号强度,比较信号强度,并与具有较大信号强度的WWAN 202或WLAN 204建立网络连接。

[0059] 在一些实施方案中,触发事件包括检测用户的调度事件。在这些实施方案中,触发事件基于对用户的调度的了解。例如,调度事件可以指定用户要在预定时间段内(例如,在接下来的5分钟内)离开特定位置X。在检测到调度事件(包括触发事件)时,方法500可随后确定WWAN 202和WLAN 204的信号强度,比较信号强度,并与具有较大信号强度的WWAN 202

或WLAN 204建立网络连接。

[0060] 在一些实施方案中,触发事件包括检测便携式音频装置100的超过预定阈值距离的移动。便携式音频装置100的移动可以由运动装置175(在图1中示出)检测。运动装置175可包括GPS或其他位置敏感装置。在这些实施方案中,当运动装置175检测到便携式音频装置100已经移动了预定阈值距离时,检测到触发事件。如果便携式音频装置100已经移动了预定阈值距离,则方法500可以确定WWAN 202和WLAN 204的信号强度,比较信号强度,并且与具有较大信号强度的WWAN 202或WLAN 204建立网络连接。

[0061] 在进一步的实施方案中,默认网络(WWAN 202或WLAN 204)可以由音频处理应用112或用户设置。在这些实施方案中,如果可用的话,音频处理应用112可以首先尝试与默认网络建立网络连接。如果默认网络不可用,则音频处理应用112可尝试与其他网络建立网络连接。例如,如果WLAN选项是新的受密码保护的的网络,则WWAN 202可以被设置为默认网络(这需要注册过程),所述WLAN是拥挤的、运行缓慢的、丢弃的或以其他方式有问题的,和/或用户具有无限的WWAN 202数据计划(因此,价格不是WWAN 202的因素)。在另一示例中,如果用户不具有无限的WWAN 202数据计划(因此价格是WWAN 202的因素),则WLAN 204可以被设置为默认网络。

[0062] 总而言之,本文所描述的实施方案涉及一种便携式音频装置,其包括能够与WLAN连接的网卡和连接到WWAN的无线调制解调器。便携式音频装置可经由网卡和WLAN或无线调制解调器和WWAN与语音服务平台通信。在一些实施方案中,当便携式音频装置无法访问WLAN时,便携式音频装置可通过经由无线调制解调器和WWAN与语音服务平台通信来处理 and 响应语音查询。便携式音频装置100还包括电池140,所述电池提供便携式音频装置100的各种硬件和软件部件所需的功率以执行本文描述的高级语音功能和其他功能。为了节省电池电量,可以将便携式音频装置100置于低功率状态,同时连续地监视唤醒命令。当接收到唤醒命令时,便携式音频装置100可以转换到正常功率状态以接收和处理来自用户的语音查询。

[0063] 所公开的方法的至少一个优点是,即使当便携式音频装置不在WLAN的范围内并且没有从房屋或建筑物的墙壁插座接收电力时,便携式音频装置也能够处理语音查询并提供对语音查询的响应。特别地,便携式音频装置能够接收语音查询,与语音服务平台通信,并且当便携式音频装置不在诸如Wi-Fi网络的WLAN的范围内时,提供对语音查询的响应。此外,便携式音频装置的电池允许便携式音频装置在WLAN范围之外的区域中提供高级语音功能。因此,便携式音频装置提供了真正的便携性并且可以在任何环境中(诸如在家庭或建筑物环境内或在家庭或建筑物环境之外)使用。

[0064] 1. 在一些实施方案中,一种便携式音频装置包括:一组一个或多个远场麦克风,其被配置成获取语音查询;一组一个或多个扬声器;无线调制解调器,其被配置成连接到无线广域网(WWAN);存储器,其包括音频处理应用;以及处理器,其联接到所述存储器并且在执行所述音频处理应用时被配置成:基于所述语音查询产生数据请求;经由所述无线调制解调器向所述WWAN传送所述数据请求;经由所述无线调制解调器从所述WWAN接收与所述数据请求相关联的数据响应;基于所述数据响应产生音频信号;以及将所述音频信号传送到所述一组扬声器以产生音频输出。

[0065] 2. 如条款1所述的便携式音频装置,其还包括电池,所述电池用于对所述一组麦克

风、所述一组扬声器和所述无线调制解调器中的至少一者供电。

[0066] 3. 如条款1或2所述的便携式音频装置,其中所述处理器被配置成通过使用自然语言处理技术处理所述语音查询来基于所述语音查询产生所述数据请求。

[0067] 4. 如条款1-3中任一项所述的便携式音频装置,其中所述一组扬声器包括用于第一频率范围的第一驱动器和用于第二频率范围的第二驱动器。

[0068] 5. 如条款1-4中任一项所述的便携式音频装置,其中:所述语音查询包括对媒体项或信息项的请求;并且所述数据响应包括所述媒体项或所述信息项。

[0069] 6. 如条款1-5中任一项所述的便携式音频装置,其还包括订户身份模块(SIM)装置,其中,所述处理器还被配置成基于经由所述SIM装置接收的信息与所述WWAN通信。

[0070] 7. 如条款1-6中任一项所述的便携式音频装置,其还包括连接到无线局域网(WLAN)的网卡,其中所述处理器还被配置成:在产生所述数据请求之前,连续地监视将由所述一组麦克风获取的唤醒命令;以及响应于检测到所述唤醒命令,确定经由所述无线调制解调器连接到所述WWAN,并且不经由所述网卡连接到所述WLAN。

[0071] 8. 如条款1-7中任一项所述的便携式音频装置,其中,所述处理器还被配置成:在产生所述数据请求之前,当所述便携式音频装置处于低功率状态时,连续地监视将由所述一组麦克风获取的唤醒命令;以及响应于检测到所述唤醒命令,将所述便携式音频装置从所述低功率状态转换到正常功率状态。

[0072] 9. 如条款1-8中任一项所述的便携式音频装置,其中:在所述低功率状态下,用于监视所述唤醒命令的所述便携式音频装置的第一组部件被供电,其中不用于监视所述唤醒命令的所述便携式音频装置的第二组部件被断电;并且在所述正常功率状态下,所述第二组部件中的至少一个部件被供电。

[0073] 10. 在一些实施方案中,一种计算机可读存储介质包括指令,所述指令当由处理器执行时,使所述处理器通过执行以下步骤处理语音查询:经由便携式音频装置的一组一个或多个远场麦克风获取所述语音查询;基于所述语音查询产生数据请求;经由包括在所述便携式音频装置中的无线调制解调器向无线广域网(WWAN)传送所述数据请求;经由所述无线调制解调器从所述WWAN接收与所述数据请求相关联的数据响应;基于所述数据响应产生音频信号;以及将所述音频信号传送到包括在所述便携式音频装置中的一组扬声器以产生音频输出。

[0074] 11. 如条款10所述的计算机可读存储介质,其中所述便携式音频装置包括电池,所述电池用于在低功率状态中对所述便携式音频装置的第一组部件供电且在正常功率状态中对所述便携式音频装置的第二组部件供电。

[0075] 12. 如条款10或11所述的计算机可读存储介质,其中基于所述语音查询产生所述数据请求包括使用自然语言处理技术来处理所述语音查询。

[0076] 13. 如条款10-12中任一项所述的计算机可读存储介质,其中所述一组扬声器包括分频器、用于第一频率范围的第一驱动器和用于第二频率范围的第二驱动器。

[0077] 14. 如条款10-13中的任一项所述的计算机可读存储介质,其中:所述语音查询包括对基于音频的媒体项或信息项的请求;并且所述数据响应包括所述基于音频的媒体项或信息项。

[0078] 15. 如条款10-14中任一项所述的计算机可读存储介质,其中所述便携式音频装置

包括订户身份模块 (SIM) 装置, 所述便携式音频装置被配置成基于经由所述 SIM 装置接收的信息与所述 WWAN 通信。

[0079] 16. 如条款 10-15 中任一项所述的计算机可读存储介质, 其中所述便携式音频装置包括连接到无线局域网 (WLAN) 的网络适配器, 还包括在经由所述无线调制解调器传送所述数据请求之前确定所述 WLAN 不在范围内的步骤。

[0080] 17. 如条款 10-16 中任一项所述的计算机可读存储介质, 其还包括以下步骤: 在产生所述数据请求之前, 在所述便携式音频装置处于低功率状态时连续地监视将由所述一组麦克风获取的唤醒命令; 以及响应于检测到所述唤醒命令, 将所述便携式音频装置从所述低功率状态转换到正常功率状态。

[0081] 18. 如条款 10-17 中任一项所述的计算机可读存储介质, 其中: 在所述低功率模式中, 所述一组麦克风中的第一麦克风子组被供电以用于监视所述唤醒命令, 其中所述一组麦克风中的第二麦克风子组被断电; 在所述正常功率模式中, 所述第二麦克风子组被供电。

[0082] 19. 在一些实施方案中, 一种便携式音频装置, 包括: 一组一个或多个麦克风, 其被配置成获取语音查询; 一组一个或多个扬声器; 无线调制解调器, 其被配置成连接到无线广域网 (WWAN); 存储器, 其包括音频处理应用; 以及处理器, 其联接到存储器并且在执行音频处理应用时被配置成: 基于所述语音查询产生数据请求; 经由所述无线调制解调器向所述 WWAN 传送所述数据请求; 经由所述无线调制解调器从所述 WWAN 接收与所述数据请求相关联的数据响应; 基于所述数据响应产生音频信号; 以及将所述音频信号传送到所述一组扬声器以基于所述音频信号产生声音, 其中所述便携式音频装置不包括视频显示屏。

[0083] 20. 如条款 19 所述的便携式音频装置, 其中所述便携式音频装置不包括视频处理能力。

[0084] 21. 如条款 19 或 20 所述的便携式音频装置, 其还包括连接到无线局域网 (WLAN) 的网卡, 其中所述处理器还被配置成, 在连接到所述无线广域网 (WWAN) 之前, 确定所述 WWAN 的信号强度大于所述 WLAN 的信号强度。

[0085] 22. 如条款 19-21 中任一项所述的便携式音频装置, 其中所述处理器还被配置成: 检测触发事件; 以及响应于检测到所述触发事件, 确定所述 WLAN 的信号强度大于所述 WWAN 的信号强度并与所述 WLAN 建立网络连接。

[0086] 23. 如条款 19-22 中任一项所述的便携式音频装置, 其中所述触发事件包括以下至少一者: 预定时间段的期满、当前连接的网络的信号强度下降到预定阈值水平以下、所述用户的调度事件, 以及所述便携式音频装置的移动超过预定阈值距离。

[0087] 已经出于说明的目的给出了对各种实施方案的描述, 但是并不旨在穷举或限制于所公开的实施方案。在不脱离所描述的实施方案的范围和精神的情况下, 许多修改和变化对于本领域的普通技术人员将是显而易见的。

[0088] 本实施方案的各方面可以体现为系统、方法或计算机程序产品。因此, 本公开的各方面可采取完全硬件实施方案、完全软件实施方案 (包括固件、驻留软件、微代码等) 或结合软件方面和硬件方面的实施方案的形式, 这些方面在这里通常都被称为“电路”、“部件”、“模块”或“系统”。此外, 本公开的各方面可以采取体现在一个或多个其上体现有计算机可读程序代码的计算机可读介质中的计算机程序产品的形式。

[0089] 可以采用一个或多个计算机可读介质的任意组合。计算机可读介质可以是计算机

可读信号介质或者计算机可读存储介质。计算机可读存储介质可以是例如但不限于电子、磁、光、电磁、红外或半导体系统、设备或装置,或者前述的任何合适的组合。计算机可读存储介质的更具体示例(非穷举列表)将包括以下:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器(EPROM或快闪存储器)、光纤、便携式光盘只读存储器(CD-ROM)、光学存储装置、磁存储装置或上述的任何合适的组合。在本文件的上下文中,计算机可读存储介质可以是能够包含或存储由指令执行系统、设备或装置使用或与之结合使用的程序的任何有形介质。

[0090] 以上参照根据本公开的实施方案的方法、设备(系统)和计算机程序产品的流程图和/或框图描述了本公开的各方面。应当理解,流程图图示和/或框图的每个框以及流程图图示和/或框图中的框的组合可以由计算机程序指令实现。这些计算机程序指令可以被提供给通用计算机、专用计算机或其他可编程数据处理设备的处理器以产生机器,使得经由计算机或其他可编程数据处理设备的处理器执行的指令能够实现流程图和/或框图的一个或多个框中指定的功能/动作。这样的处理器可以是但不限于通用处理器、专用处理器、特定应用处理器、或现场可编程处理器或门阵列。

[0091] 附图中的流程图和框图显示了根据本公开的各种实施方案的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这方面,流程图或框图中的每个框可以表示模块、代码段或代码部分,所述模块、代码段或代码部分包括用于实现(一个或多个)指定逻辑功能的一个或多个可执行指令。还应当注意,在有些作为替代的实现方式中,框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,连续示出的两个框实际上可以基本上同时执行,或者这些框有时可以根据所涉及的功能而以相反的顺序执行。还将注意,框图和/或流程图图示的每个框以及框图和/或流程图图示中的框的组合可以由执行指定功能或动作的基于专用硬件的系统或专用硬件和计算机指令的组合来实现。

[0092] 虽然前述内容针对本公开的实施方案,但可在不脱离其基本范围的情况下设计出本公开的其他和进一步的实施方案,并且其范围由所附权利要求来确定。

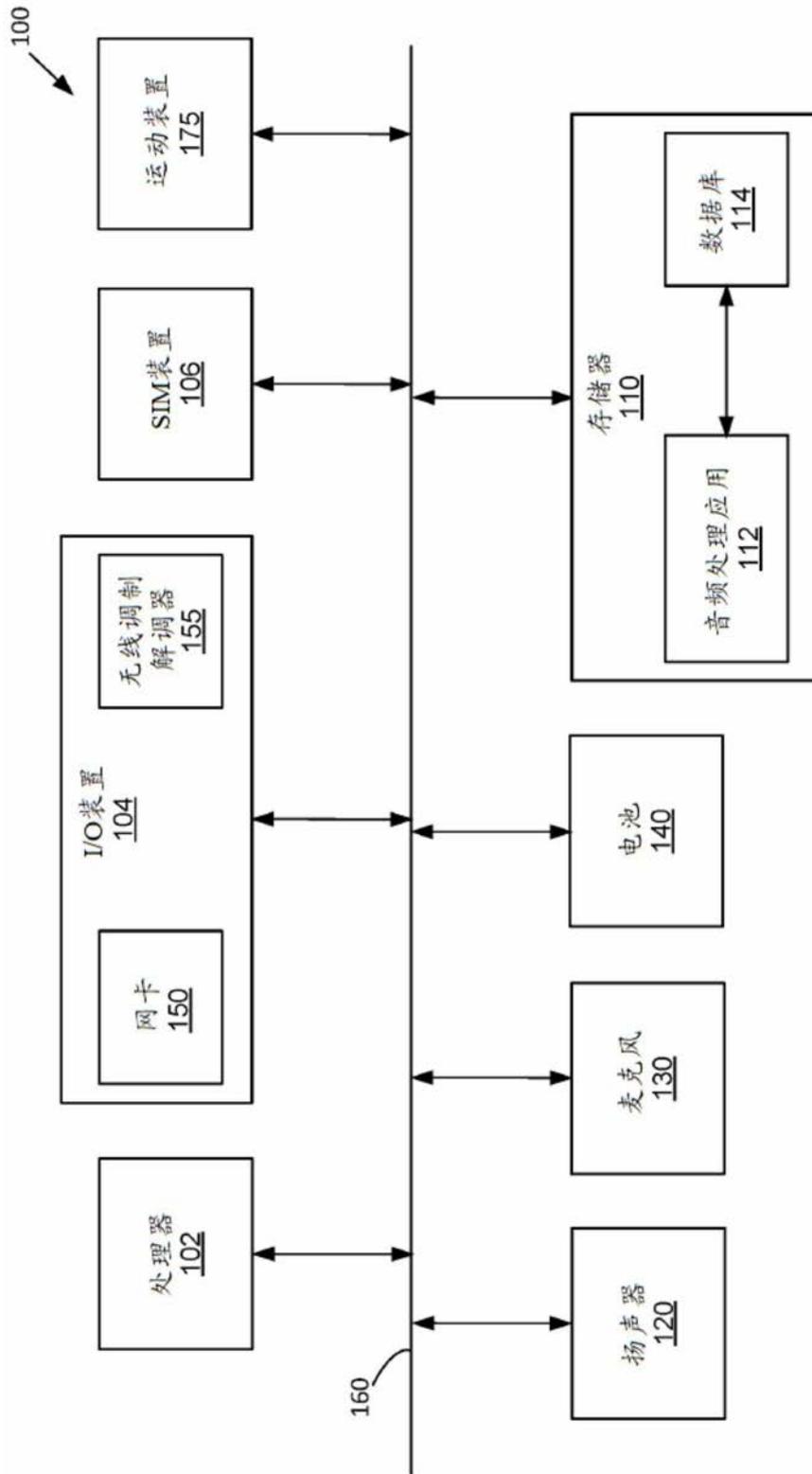


图1

250

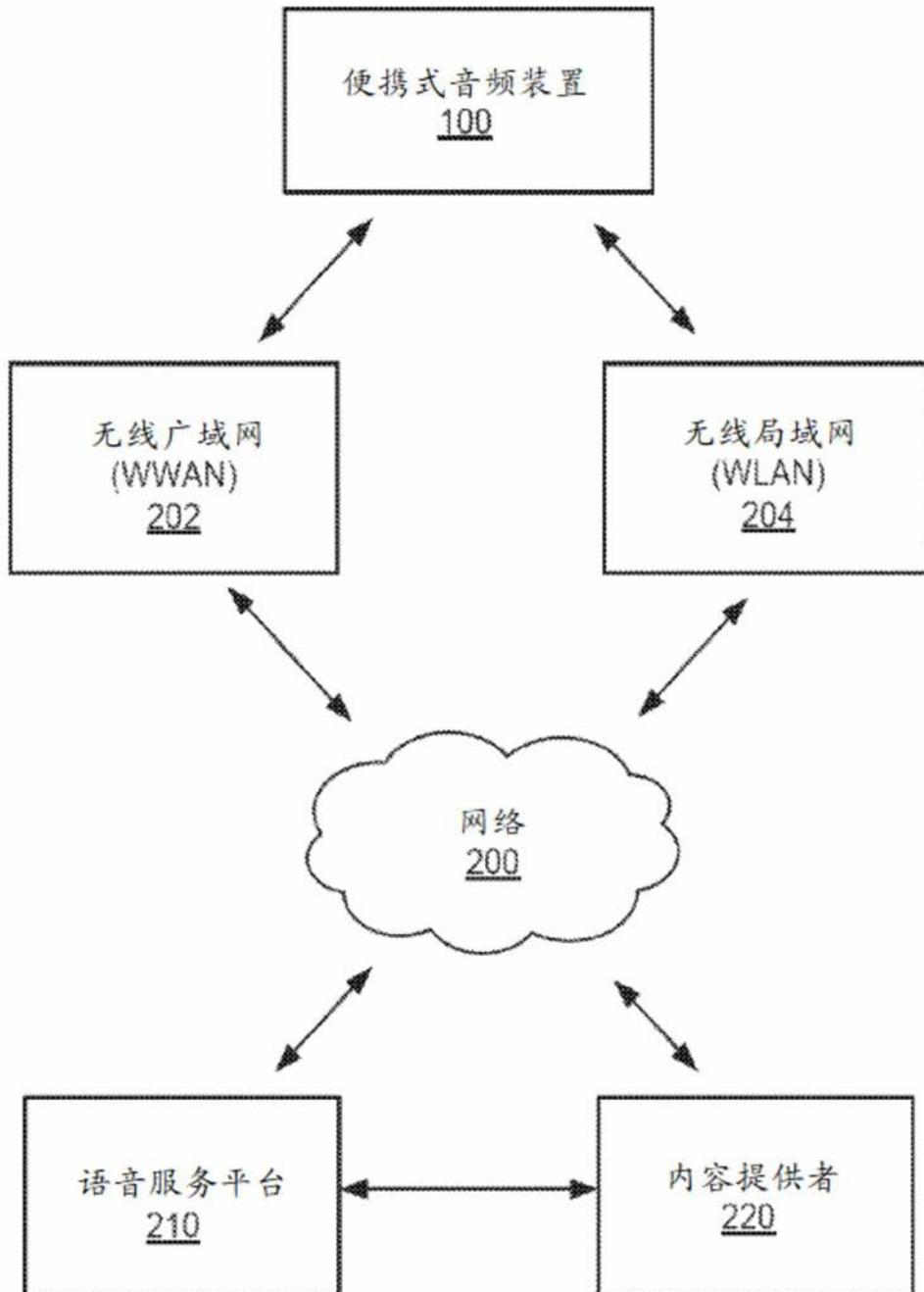


图2

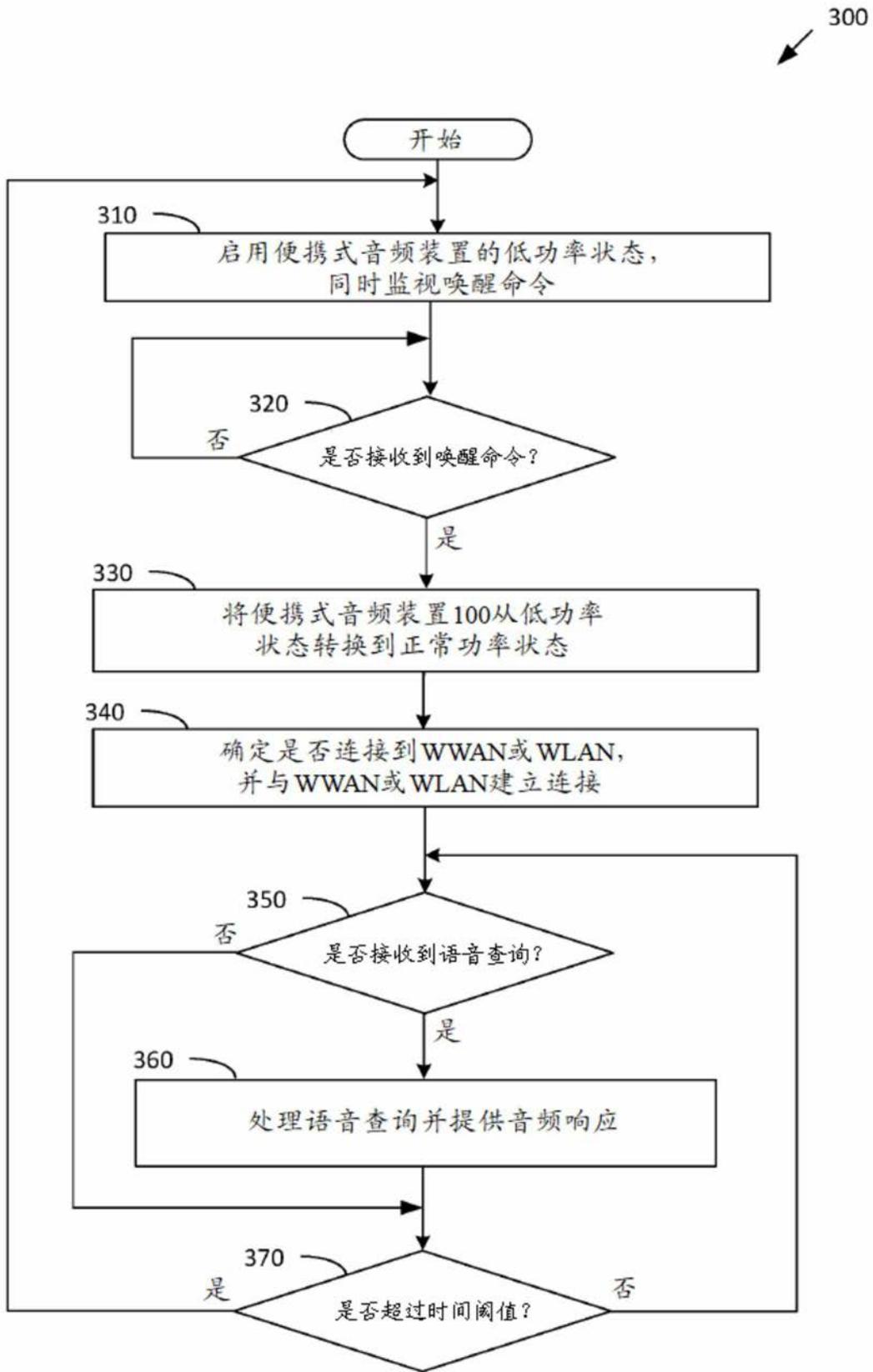


图3

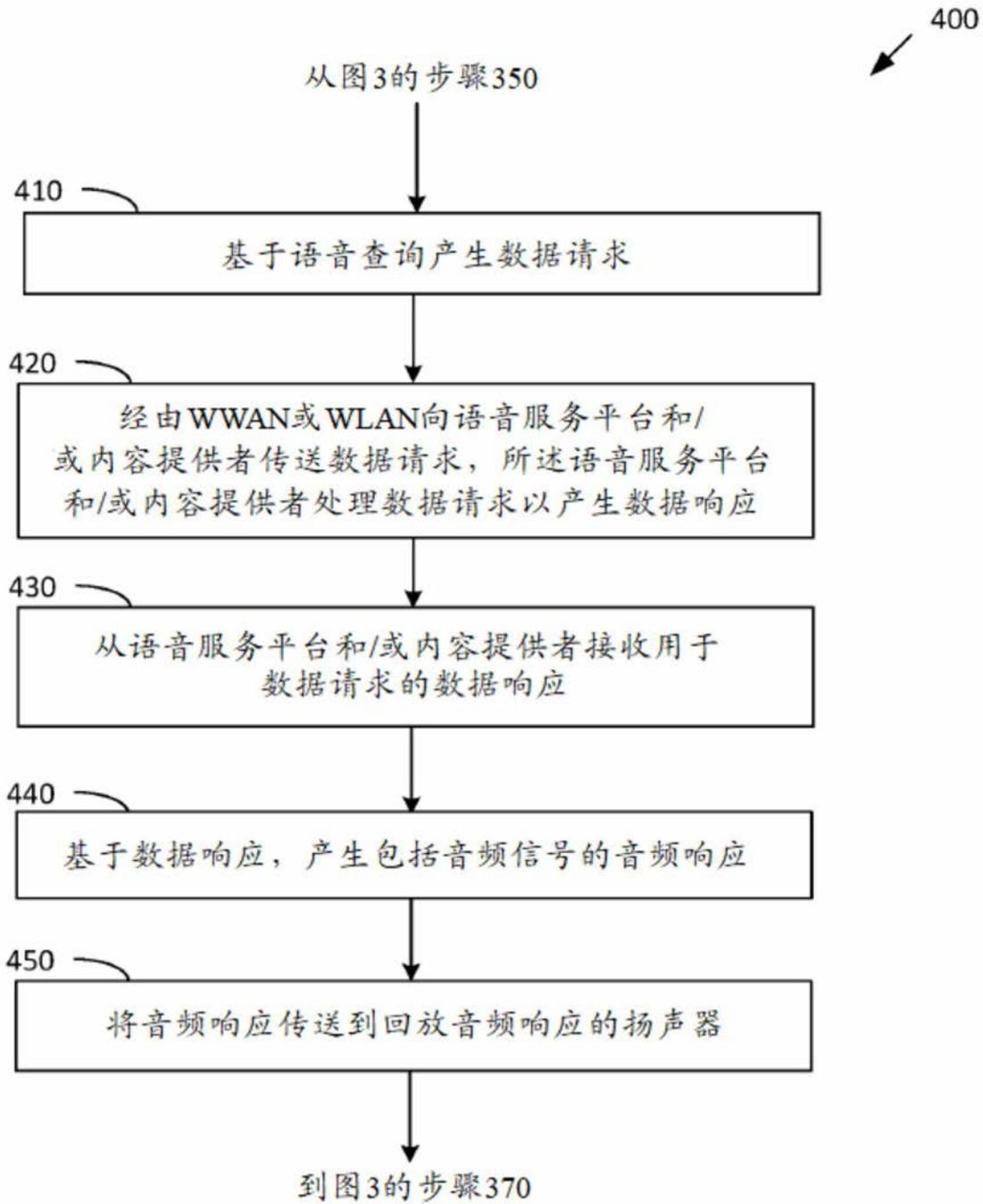


图4

500

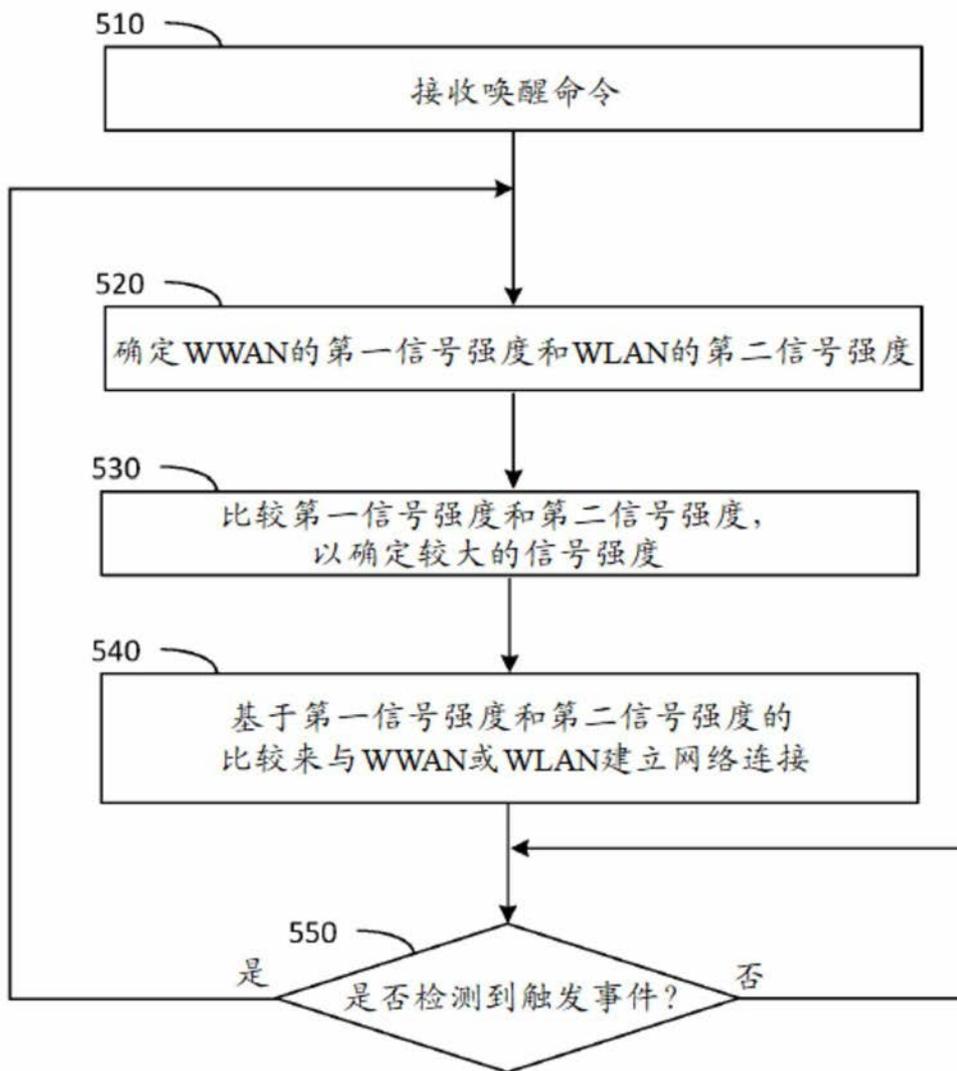


图5