



(12)发明专利

(10)授权公告号 CN 105960794 B

(45)授权公告日 2019.11.08

(21)申请号 201580006953.1

(22)申请日 2015.02.03

(65)同一申请的已公布的文献号
申请公布号 CN 105960794 A

(43)申请公布日 2016.09.21

(30)优先权数据
61/935,141 2014.02.03 US

(85)PCT国际申请进入国家阶段日
2016.08.02

(86)PCT国际申请的申请数据
PCT/US2015/014280 2015.02.03

(87)PCT国际申请的公布数据
W02015/117138 EN 2015.08.06

(73)专利权人 寇平公司
地址 美国马萨诸塞州

(72)发明人 C·帕金森 D·范 F·赫曼
J·加赛尔 M·卡达科尔

(74)专利代理机构 北京市金杜律师事务所
11256

代理人 王茂华

(51)Int.Cl.
H04M 1/60(2006.01)

(56)对比文件
US 2008300025 A1,2008.12.04,
US 2008300025 A1,2008.12.04,
US 2009186668 A1,2009.07.23,
EP 2043336 A2,2009.04.01,

审查员 高燕娃

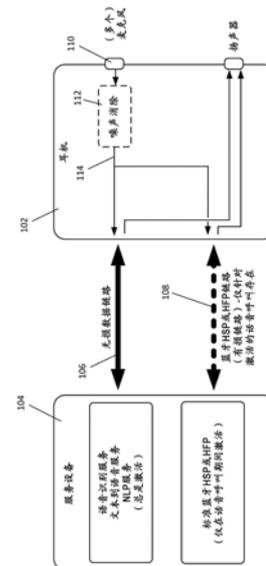
权利要求书2页 说明书6页 附图3页

(54)发明名称

用于语音命令的智能蓝牙耳机

(57)摘要

一种从用户穿戴的可穿戴设备与服务设备对接的方法,该方法包括在服务设备和可穿戴设备之间建立无损和无线数据链路。该方法还包括由可穿戴设备从可穿戴设备的一个或多个麦克风收集音频数据。该方法还包括由可穿戴设备通过无损和无线数据链路将收集的音频数据发送到服务设备。该方法还可以包括由服务设备提供与音频数据相关联的语音识别服务。



1. 一种从用户穿戴的可穿戴设备与服务设备对接的方法,所述方法包括:
在所述服务设备和所述可穿戴设备之间建立无损和无线数据链路;
由所述可穿戴设备从所述可穿戴设备的一个或多个麦克风收集音频数据;
由所述可穿戴设备通过所述无损和无线数据链路向所述服务设备发送所收集的音频数据;

从所述可穿戴设备向所述服务设备发送用于在所述服务设备处建立支持所述无损和无线数据链路所需的一个或多个组件的信息,支持所述无损和无线数据链路所需的所述一个或多个组件包括:(i) 自定义WIFI连接和自定义蓝牙简档中的一个或多个、(ii) 驱动程序以及(iii) 压缩/解压缩代码。

2. 根据权利要求1所述的方法,其中所述可穿戴设备是耳机设备。

3. 根据权利要求1所述的方法,其中所述可穿戴设备是腕表设备。

4. 根据权利要求1所述的方法,还包括由所述服务设备提供与所述音频数据相关联的语音识别服务。

5. 根据权利要求4所述的方法,其中所述语音识别服务包括宽带语音处理和(iii) 低失真语音压缩。

6. 根据权利要求1所述的方法,还包括由所述可穿戴设备提供所收集的音频数据的语音压缩。

7. 根据权利要求1所述的方法,其中所述服务设备是以下一个或多个:蜂窝电话、智能电话、平板设备、膝上型计算机、笔记本计算机、桌面型计算机、网络服务器、可穿戴移动通信设备、可穿戴移动计算机和基于云的计算实体。

8. 根据权利要求1所述的方法,还包括由所述可穿戴设备提供与所收集的音频数据相关联的噪声消除服务。

9. 根据权利要求1所述的方法,其中所述无损和无线数据链路是使用自定义蓝牙简档操作的蓝牙链路。

10. 一种在服务设备和可穿戴设备之间建立无损和无线数据链路的方法,包括:
由所述可穿戴设备在所述可穿戴设备和所述服务设备之间建立第一协议的无线链路;
由所述可穿戴设备并且使用所述第一协议的所述无线链路建立第二协议的无损无线链路;

由所述可穿戴设备向所述服务设备传送用于在所述服务设备处建立支持所述无损和无线数据链路所需的一个或多个组件的信息,支持所述无损和无线数据链路所需的所述一个或多个组件包括自定义蓝牙简档、驱动程序和压缩/解压缩代码。

11. 根据权利要求10所述的方法,其中所述第一协议的所述无线链路是有损蓝牙链路,并且所述第二协议的所述无损无线链路是无损蓝牙链路。

12. 根据权利要求11所述的方法,其中所述无损蓝牙链路基于蓝牙SPP简档。

13. 根据权利要求10所述的方法,其中第一协议的所述无损和无线链路是使用自定义蓝牙简档操作的蓝牙链路。

14. 一种可穿戴设备,包括:

至少一个麦克风;

至少一个扬声器;

话音压缩引擎；
驱动程序，被配置为在无损、无线数据信道上传输语音分组；以及
代码部署模块，被配置为向服务设备传送自定义蓝牙简档和驱动程序，以辅助在所述服务设备处实现无损链路。

15. 根据权利要求14所述的可穿戴设备，其中所述无损、无线数据信道基于蓝牙SPP简档。

16. 根据权利要求14所述的可穿戴设备，其中所述话音压缩引擎包括以下一个或多个：
(i) 子带编码器、(ii) Speex、和 (iii) ETSI分布式语音识别。

17. 根据权利要求14所述的可穿戴设备，还包括噪声消除引擎。

18. 根据权利要求17所述的可穿戴设备，其中所述噪声消除引擎接收来自两个或更多个源的音频信号，并且使用线性噪声消除算法以减少环境噪声。

19. 根据权利要求14所述的可穿戴设备，其中所述代码部署模块向所述服务设备传送小应用程序，以在所述服务设备上安装所述自定义蓝牙简档和驱动程序。

用于语音命令的智能蓝牙耳机

[0001] 相关申请

[0002] 本申请要求2014年2月3日提交的美国临时申请No. 61/935,141的权益。上述申请的全部教导通过引用合并于此。

背景技术

[0003] 被设计为与蜂窝电话或其它服务设备相配对的蓝牙耳机通常采用蓝牙免提简档 (HFP) 或蓝牙耳机简档 (HSP) 来控制音频如何从蜂窝电话被传递到耳机。HFP或HSP简档允许蜂窝电话上的传入音频数据通过近耳扬声器被直接中继到耳机以用于即时播放。同时,在耳机处从一个或多个近口麦克风收集的音频被立即传递到蜂窝电话,该蜂窝电话将收集的音频包括在当前音频电话呼叫中。

发明内容

[0004] 蓝牙耳机可以向用户提供某种形式的语音识别。这样的语音识别可以被用于控制蜂窝电话的功能和向用户提供仅通过说出命令来进行呼叫的能力。然而,最新的全部蓝牙耳机直接在蓝牙耳机本身上运行语音识别服务,或者使用基于云的识别系统。前者的语音识别服务的缺点是在耳机中需要复杂、昂贵的电子设备。后者的语音识别服务的缺点是到云的总是接通的连接的要求。

[0005] 在蓝牙设备中,语音识别服务已经利用HFP或HSP用于音频数据传输。HFP或HSP的频带是8kHz,该频带对正确的语音识别通常太窄。为了解决该问题,最近已经与诸如修改后的子带编码 (mSBC) 的压缩方法一起使用具有16kHz采样率的宽带语音 (WBS) 的新的蓝牙HFP标准 (v1.6)。

[0006] 被设计用于语音传输的HFP和HSP两者是有损的(例如,它们有时丢失语音分组或数据)。HFP和HSP通常完全不会重新传输丢失的语音分组,或者至多一次或两次重新传输它们,以限制无线电话呼叫的延迟和延续无线会话。丢失语音数据的一个或两个分组可能在解码的语音输出中几乎不引人注意。分组擦除隐藏算法进一步减少由失去语音分组引起的语音劣化。更重要的是减少蜂窝电话会话中的延迟或滞后,所以对于语音信道有损链路比高等待时间链路更可接受。

[0007] 虽然其对蜂窝电话呼叫不具有主要影响,但是丢失的分组使语音识别显著恶化。蓝牙在用于语音识别的目的时迄今没有解决分组擦除的问题的标准简档。语音信道中的有损协议在蓝牙中尚未解决。此外,HFP和HSP并未消除足够的非平稳噪声并且可以使语音传输失真,这可以使语音识别的准确性恶化。

[0008] 在本发明的实施例中,标准蓝牙耳机被改进以提供更好的语音识别和向用户传递信息。此外,本发明通过解决蓝牙中丢失数据分组的问题显著改善语音识别。

[0009] 在一些实施例中,蓝牙设备可以是另一类型的可穿戴设备,而不是耳机。这样的可穿戴设备可以包括腕戴设备,穿戴在上臂或身体的其它部分的设备。

[0010] 在一个方面中,本发明可以是用户穿戴的可穿戴设备与服务设备对接的方法。

该方法可以包括在该服务设备和该可穿戴设备之间建立无损和无线数据链路,由该可穿戴设备从该可穿戴设备的一个或多个麦克风收集音频数据。该方法还可以包括由该可穿戴设备通过该无损和无线数据链路向该服务设备发送所收集的音频数据。

[0011] 在一个实施例中,该可穿戴设备是耳机设备。在另一实施例中,该可穿戴设备是腕表设备

[0012] 一个实施例还包括由该服务设备提供与该音频数据相关联的语音识别服务。

[0013] 在一个实施例中,该语音识别服务包括宽带语音处理和(iii)低失真语音压缩。

[0014] 另一实施例还包括由该可穿戴设备提供所收集的音频数据的语音压缩。

[0015] 在一个实施例中,该服务设备是以下一个或多个:蜂窝电话、智能电话、平板设备、膝上型计算机、笔记本计算机、桌面型计算机、网络服务器、可穿戴移动通信设备、可穿戴移动计算机和基于云的计算实体。

[0016] 另一实施例还包括由该可穿戴设备提供与所收集的音频数据相关联的噪声消除服务。另一实施例还包括从该可穿戴设备向该服务设备发送用于在该服务设备处建立支持该无损和无线数据链路所需的一个或多个组件的信息。

[0017] 在一个实施例中,支持该无损和无线数据链路所需的该一个或多个组件包括:(i)自定义WIFI连接和自定义蓝牙简档中的一个或多个、(ii)驱动程序、和(iii)压缩/解压缩代码。

[0018] 在另一实施例中,该无损和无线数据链路是使用自定义蓝牙模简档操作的蓝牙链路。

[0019] 在另一方面中,本发明可以是在服务设备和可穿戴设备之间建立无损和无线数据链路的方法。该方法可以包括由该可穿戴设备在该可穿戴设备和该服务设备之间建立第一协议的无线链路。该方法还可以包括由该可穿戴设备并且使用该第一协议的该无线链路建立第二协议的无损无线链路。该方法还可以包括由该可穿戴设备向该服务设备传送用于在该服务设备处建立支持该无损和无线数据链路所需的一个或多个组件的信息。

[0020] 在一个实施例中,支持该无损和无线数据链路所需的该一个或多个组件包括自定义蓝牙简档、驱动程序和压缩/解压缩代码。在另一实施例中,第一协议的该无线链路是有损蓝牙链路,并且第二协议的该无线链路是无损蓝牙链路。在另一实施例中,该无损蓝牙链路基于蓝牙SPP简档。在另一实施例中,第一协议的该无损和无线链路是使用自定义蓝牙简档操作的蓝牙链路。

[0021] 在另一方面中,本发明可以是可穿戴设备,其包括至少一个麦克风、至少一个扬声器、语音压缩引擎、以及被配置为在无损、无线数据信道上传输语音分组的驱动程序。

[0022] 在一个实施例中,该无损、无线数据信道基于蓝牙SPP简档。在另一实施例中,该语音压缩引擎包括以下一个或多个:(i)子带编码器、(ii)Speex、和(iii)ETSI分布式语音识别。

[0023] 一个实施例还可以包括噪声消除引擎。在另一实施例中,该噪声消除引擎接收来自两个或更多个源的音频信号,并且使用线性噪声消除算法以减少环境噪声。

[0024] 一个实施例还可以包括代码部署模块,其被配置为向服务设备传送自定义蓝牙简档和驱动程序,以有助于在该服务设备处实现该无损链路。在另一实施例中,该代码部署模块向该服务设备传送小应用程序,以在该服务设备上安装该自定义蓝牙简档和驱动程序。

附图说明

[0025] 根据下文的如在附图中所示的本发明的示例实施例的更具体的描述,上文将是显而易见的,附图中相同参考标号贯穿不同视图指代相同部件。附图不一定按比例绘制,而是将重点放在说明本发明的实施例。

[0026] 图1示出了使用两个音频链路将耳机与蜂窝电话连接的示例实施例的框图。

[0027] 图2示出了根据本发明处理和传输音频信号以用于语音识别的示例实施例的框图。

[0028] 图3示出了根据本发明处理和传输音频信号以用于语音识别的另一示例实施例的框图。

具体实施方式

[0029] 本发明的示例实施例的描述如下。

[0030] 下文更详细地描述的图1是本发明的示例实施例。该实施例涉及两个主要元件-通过一个或多个无线链路连接的耳机102和服务设备104。服务设备104可以是可以实现到免提耳机的无线链路的任何设备,服务设备104包括但不限于蜂窝电话、智能电话、平板设备、膝上型计算机、笔记本计算机、桌面型计算机、网络服务器、可穿戴移动通信设备、可穿戴移动计算机或基于云的实体。可穿戴设备可以包括穿戴在用户的手腕、上臂、腿、腰或颈,或适用于支撑通信和/或计算设备的任何其它身体部分上的设备。类似地,耳机102组件可以是穿戴在用户的手腕、上臂、腿、腰或颈,或适用于支撑无线设备(例如,蓝牙或WiFi设备)的任何其它身体部分上的设备,而不是耳机。

[0031] 在本发明的实施例中,服务设备直接托管语音识别服务。为了有助于该托管,实施例建立从服务设备到耳机的新的第二数据链路。第二数据链路应该是无损的。第二数据链路可以是蓝牙数据链路。第二蓝牙数据链路可以被用于将近口麦克风输入(或如果HFP链路是激活的,则发送麦克风输入的第二副本)发送到服务设备,该服务设备运行语音识别服务/语音识别处理软件。第二蓝牙数据链路保留原始免提筒档链路,并且确保与蜂窝电话的现有固件的不间断的兼容。在采取该方式中,压缩方案可以以不被标准免提筒档支持的方式压缩服务设备和耳机之间的音频数据(例如,通过使用需要无损数据路径的压缩/解压缩方案)。

[0032] 随着该系统建立,用户可以向耳机说出命令。命令(例如,说出的音频)通过第二蓝牙数据链路被立即传送到服务设备,因此音频被传递到语音识别系统中。取决于所说的命令,语音识别器随后能够采取适当的动作,诸如向给定的电话号码发起新的呼叫。

[0033] 此外,随着该系统到位,功能不再局限于仅建立电话呼叫。耳机穿戴者可以说出自然语句以使其它重要功能进行动作,诸如“发送SMS消息给约翰,今晚我将迟到”。例如,该语句在被服务设备上的或通过无线链路被网络服务器上的语音识别和自然语言/自然语言理解引擎处理时,可以被用于创建并发送适当的SMS消息。以相同的方式,用户可以通过向耳机说话并让服务设备执行语音识别和执行适用于所识别的语音的动作,来查询电话的状态或执行基于web的查询。

[0034] 在与使用第二蓝牙数据链路来收集麦克风数据并发送到服务设备相同的时间处,该链路也可以从服务设备将音频发送回耳机以用于通过近耳扬声器重放。具体地,这用于

通过计算机生成的说出的短语,又名文本到语音(TTS),将信息传送回耳机穿戴者。

[0035] 例如,在服务设备上运行的软件可以检测到传入的SMS文本消息。通常服务设备使用鸣响提醒用户,并且可以在屏幕上显示传入的消息。在本发明的实施例中,SMS消息可以在服务器侧上被转换为语音(例如,文本到语音),并且朗读(reading)的语音音频可以通过蓝牙链路被发送以用于向用户重放。这里的结果是在用户无须操作或看到服务设备的情况下向用户大声读出传入的消息的系统。

[0036] 该技术可以与语音识别服务相结合以提供双向问答系统。例如,用户现在可以向耳机说话以询问问题,诸如“现在是什么时候?”该音频可以被语音识别服务处理,计算出答案,并且随后向用户大声说出。

[0037] 图1是示出了使用两个双向信道,无损数据链路106和有损数据链路108,来将耳机102与服务设备104连接的示例实施例的框图。在该示例实施例中,无损数据链路106是使用串行端口简档(SPP)的蓝牙链路,并且有损数据链路108是使用耳机简档(HSP)或免提简档(HFP)的蓝牙链路。在其它实施例中,无损数据链路106可以是诸如WiFi或本领域中已知的其它无线技术的另一数字数据链路。

[0038] 如下文将更加详细地描述的,虽然SPP可以为无损数据链路提供下层基础,但是简档本身不提供无损传输。因为此时,蓝牙具体地在用于语音识别的目的时不提供解决分组丢失的问题的标准简档。需要自定义的简档,或者至少需要修改后的SPP版本。

[0039] 在该示例中,无损数据链路106被建立,并且只要服务设备104和耳机102两者都被激活(即,打开),无损数据链路106就被允许保持激活。另一方面,仅当耳机102的用户进行话音呼叫时,有损数据链路108被激活。

[0040] 在该示例实施例中,耳机102上的一个或多个麦克风110收集音频数据。音频可以随后可选地通过耳机102上的噪声消除模块112,以减少背景噪声并改善语音识别。多个麦克风110的使用可以通过更有效地消除平稳和非平稳噪声两者来进一步改善总体噪声消除性能。

[0041] 如图所示,麦克风音频114可以随后被分成两个流。音频流中的一个被发送到无损数据链路106,并且另一个被发送到有损数据链路108。

[0042] 如先前所述,有损数据链路108仅在与激活的电话呼叫相关联时在耳机102和服务设备104之间建立。因此,该通信链路是间歇性的。当建立了有损数据链路108时,音频流中的一个被发送到作为正常的免提系统的一部分的服务设备104。音频通过无损数据链路106从服务设备104发送到耳机102。在服务设备操作需要呼叫通过HFP或HSP数据链路108发生的情况下,音频也可以通过有损(HFP或HSP)数据链路108从服务设备104发送到耳机102。在一些实施例中,音频可以是以计算机生成的说出的短语的形式(例如,文本到语音服务),该短语在耳机上被重放。

[0043] 如果蓝牙免提呼叫被激活,则音频也在耳机102上被回放,并且与来自无损数据链路106(在这里也被称为命令/控制链路)的任何说出的短语合并。通过暂时使来自有损数据链路108的电话呼叫语音静音,通过无损数据链路106接收的音频可以被给予优先权,或者两个音频信号可以被混合,因此用户同时听到两个音频,或者来自有损数据链路108的音频可以被暂时衰减(即,部分静音),以使更容易听到来自无损数据链路106的音频。

[0044] 图2和图3是示出了处理和传输音频语音信号以用于语音识别的示例实施例的框

图。在该示例实施例中，音频信息通过双向、无损、无线数据链路在耳机202和服务设备204之间传送。

[0045] 在图2所示的示例实施例中，音频语音信号是从两个或更多个麦克风206收集的，并由噪声消除模块208进行处理。在一个实施例中，可以使用线性算法来处理噪声消除以避免向语音信号引入任何非线性失真。图2示出使用话音压缩模块210压缩语音信号。压缩的语音信号通过例如串行端口简档(SPP)蓝牙数据链路的无损、双向、无线数据链路212被发送到服务设备204。

[0046] 服务设备204从无损数据链路212接收压缩的语音信号，并使用话音解压缩模块214将压缩的语音数据解压缩。通过无损数据路径获取的得到的话音数据可以被自动语音识别(ASR)引擎和/或自然语言处理引擎216使用。

[0047] 服务设备204可以具有要发送到耳机202的数字语音文件(例如，文本到语音(TTS)或WAVE(.wav格式))。语音数据首先被话音压缩模块压缩，并通过无损数据链路212发送到耳机。话音解压缩模块222解压缩语音数据，并将数据提供到将音频文件转换为驱动扬声器226的音频信号的TTS或WAVE播放模块224。

[0048] 图3示出使用服务设备中的ASR后端和自然语言处理(NLP)引擎在耳机302中提供前端特征提取和噪声消除的实施例。正如图2的实施例，使用两个或更多个麦克风306收集音频，噪声消除模块308减少环境噪声。数据在耳机302和服务设备304之间通过无损数据链路312传递至服务设备304处的ASR后端模块330。ASR后端模块330向NLP引擎提供处理后的语音数据。正如图2所示的实施例，TTS/WAVE文件318可以通过话音压缩模块320，无损数据链路312，话音解压缩模块322和驱动扬声器326的TTS或WAVE播放器而从服务设备304传递到耳机302。在其它实施例中，WAVE文件可以被存储在耳机上，并通过由服务设备传送的简单命令被发起以用于在耳机上重放。

[0049] 图2和图3中突出的特征是所描述的实施例可以如何提供有用的功能的示例。这些实施例可以彼此结合，或者与提供其它特征的其它实施例结合。

[0050] 以下是在所描述的实施例中可以被采用以用于语音识别的话音压缩技术的示例：

[0051] ●子带编码器(SBC)

[0052] ●蓝牙WBS mSBC

[0053] ●Speex(或基于压缩算法的其它码激励线性预测(CELP))

[0054] ●Opus

[0055] ●欧洲电信标准协会(ETSI)分布式语音识别(DSR)

[0056] 如上所述，蓝牙串行端口简档(SPP)本身并不提供无损传输。然而，所描述的实施例在结合蓝牙SPP使用时创建无损数据链路。所描述的实施例至少实现自定义蓝牙简档和驱动程序以实现无损连接所需的操作。这样的操作可以包括重传协议，诸如自动重发请求ARQ、混合ARQ(HARQ)和本领域中已知的其它丢失分组恢复技术。一些实施例包括在蓝牙链路两端中的自定义软件。该软件可以包括(多个)自定义蓝牙简档，(多个)驱动程序和压缩/解压缩代码。

[0057] 一些实施例修改蓝牙SPP以提供无损数据链路，而其它实施例提供完全自定义的蓝牙简档以提供适用于ASR的无损数据链路。还应注意，虽然示例实施例利用蓝牙来提供无线链路，但是所描述的实施例可以利用其它无线协议和接口来提供所描述的好处。

[0058] 所描述的实施例还可以提供用于在服务设备侧处安装上述自定义软件和代码的技术。在一些实施例中,服务设备侧可以包括预先安装的自定义驱动程序。在其它实施例中,一旦蓝牙链路建立,蓝牙免提设备就可以将小应用程序(或用于传送所需的驱动程序和软件的其它载体)通过上述蓝牙SPP链路下载到服务设备。

[0059] 所描述的实施例可以被容易地扩展以在蓝牙耳机上容纳显示器。在这样的扩展中,可以使用总是接通的命令和控制链路将耳机上的显示所需的信息从蜂窝电话发送到耳机。信息可以由耳机发送和渲染。备选地,信息可以由蜂窝电话渲染,并作为图像或部分图像被发送到耳机以用于显示。该后者的方法允许耳机固件是简单和灵活的-所有的辛勤工作由蜂窝电话完成。

[0060] 将清楚,这里描述的一个或多个实施例可以以许多不同形式的软件和硬件来实现。用于实现这里描述的实施例的软件代码和/或专用硬件不限制本发明。因此,在不参考具体软件代码和/或专用硬件的情况下描述实施例的操作和行为-可以理解人们将能够设计实现基于这里的描述的实施例的软件和/或硬件。

[0061] 此外,本发明的某些实施例可以被实现为执行一个或更多功能的逻辑。该逻辑可以是基于硬件的、基于软件的、或者基于硬件和基于软件的组合。一些或全部的逻辑可以被存储在一个或多个有形计算机可读存储介质上,并且可以包括可以由控制器或处理器执行的计算机可执行指令。计算机可执行指令可以包括实现本发明的一个或多个实施例的指令。有形计算机可读存储介质可以是易失性的或非易失性的,并且可以包括例如闪存、动态存储器、可移除盘以及非可移除盘。

[0062] 虽然已经参考本发明的示例实施例具体示出和描述了本发明,但是本领域技术人员将理解,可以在本发明中进行形式和细节的各种改变,而不脱离由所附权利要求涵盖的本发明的范围。

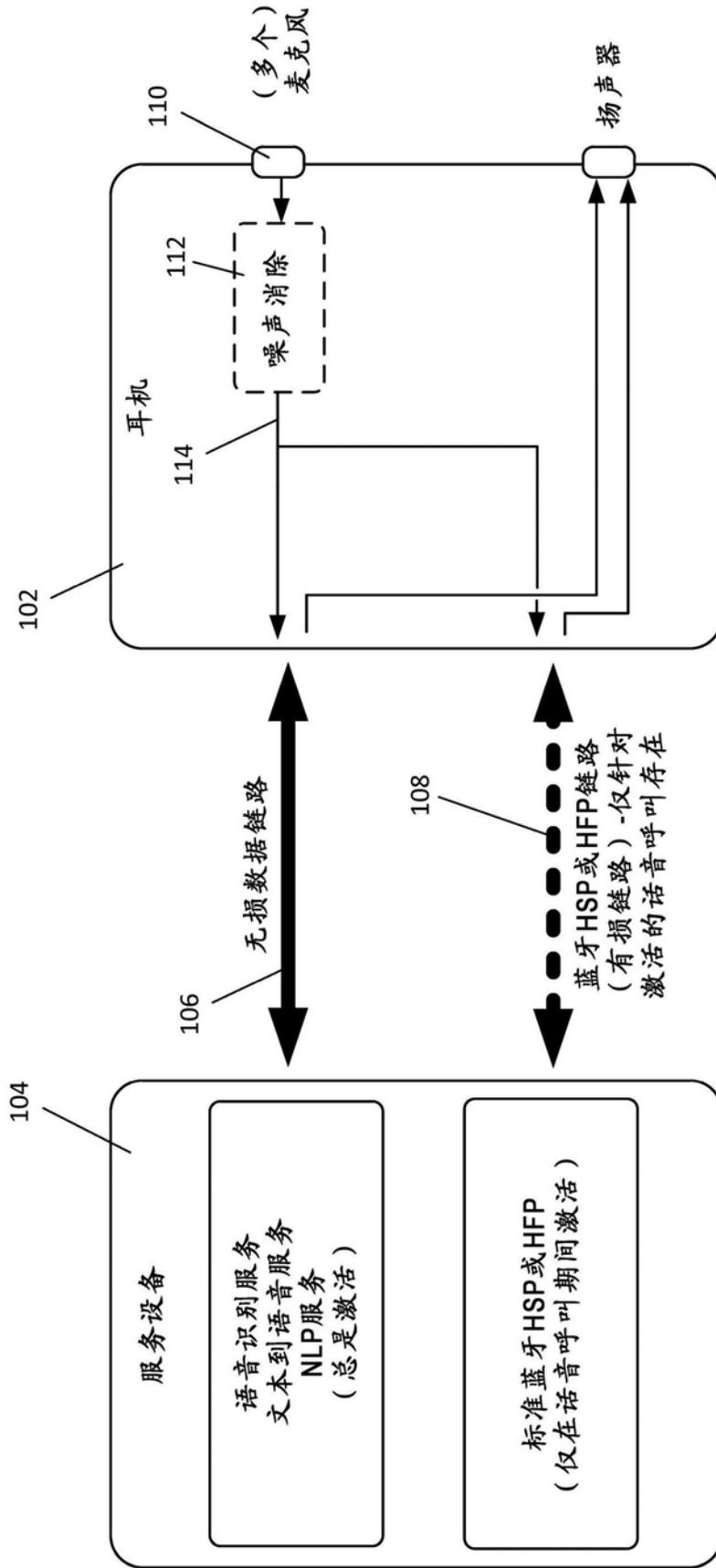


图1

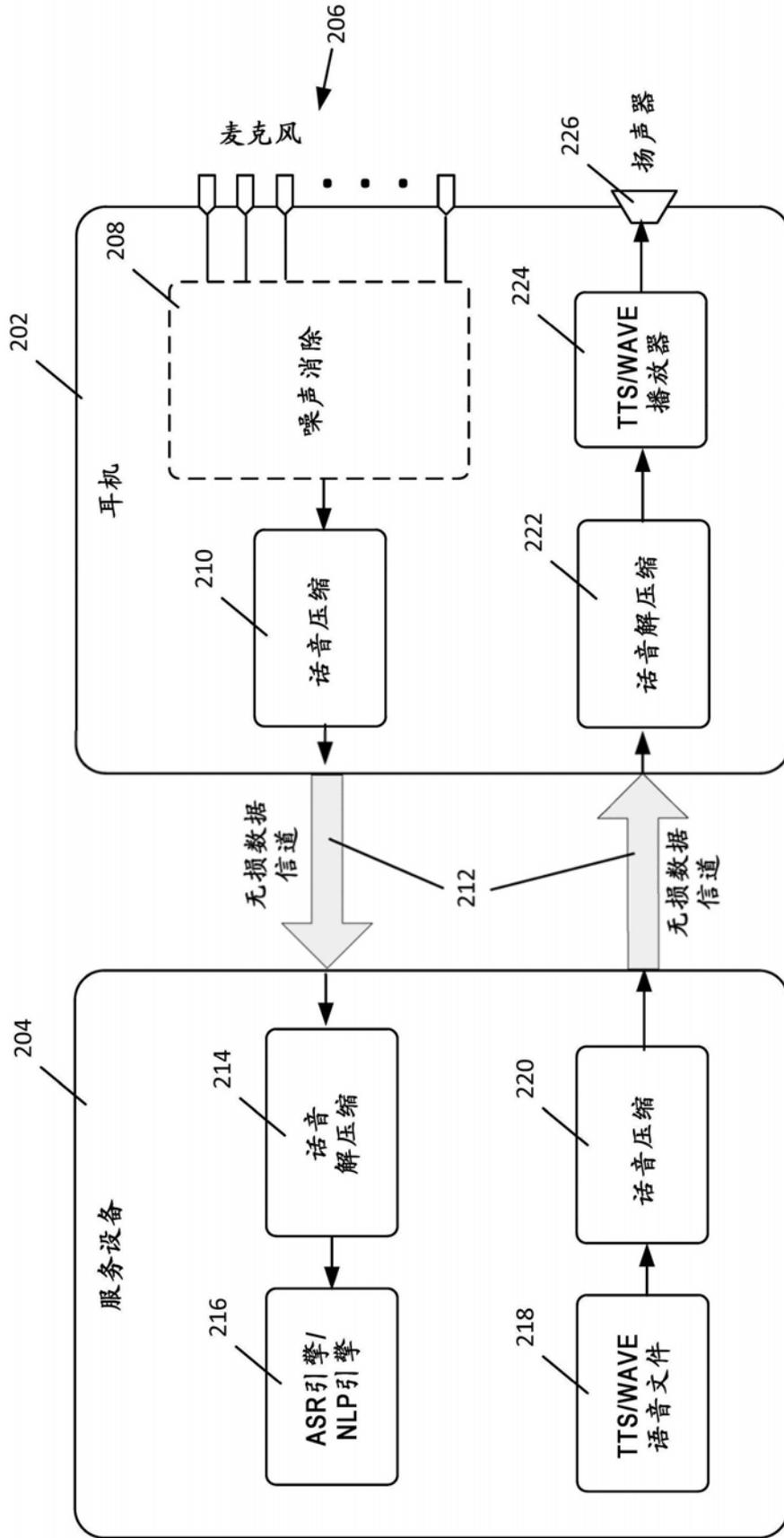


图2

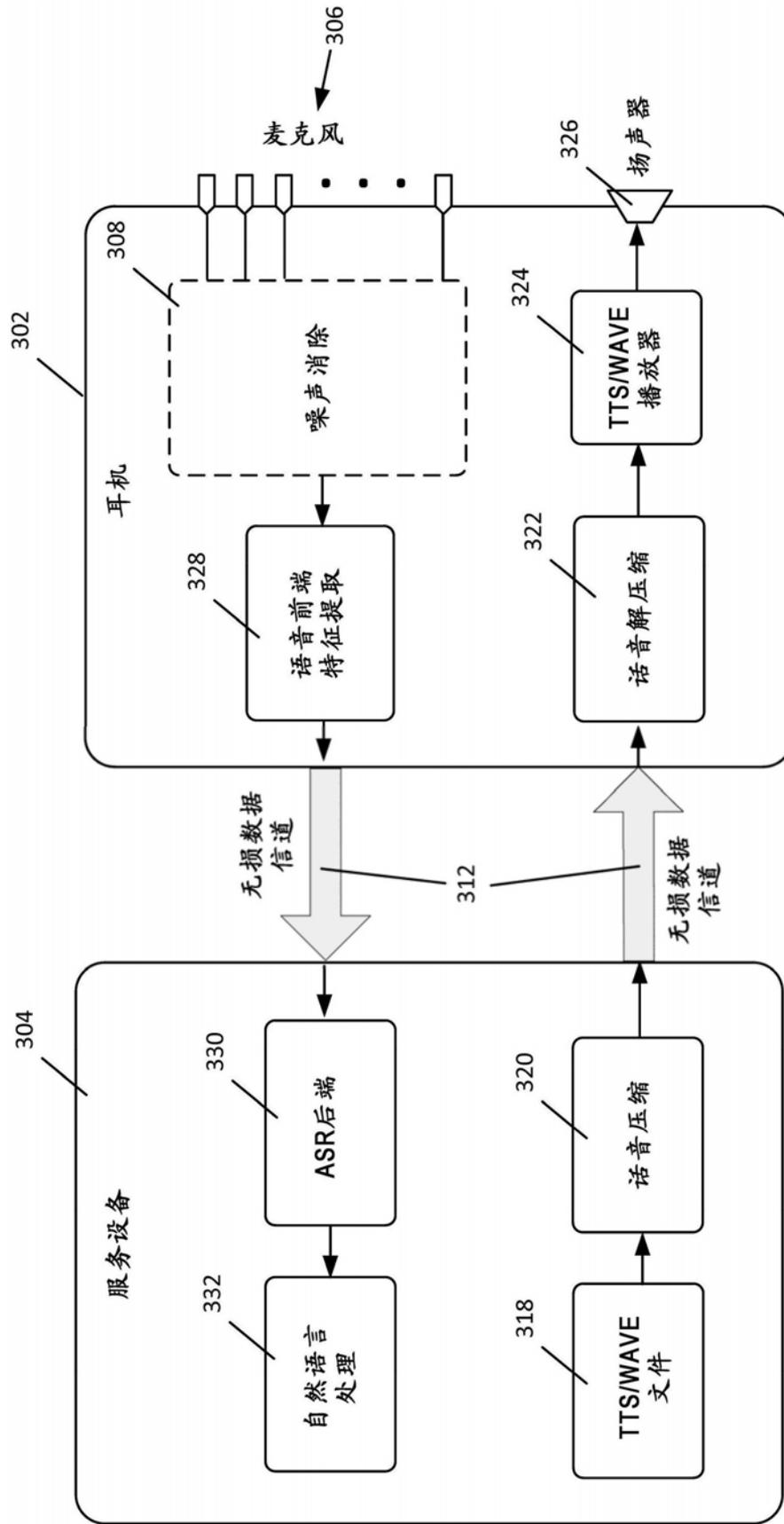


图3