



(12)发明专利申请

(10)申请公布号 CN 109599111 A

(43)申请公布日 2019.04.09

(21)申请号 201910000688.6

(22)申请日 2019.01.02

(71)申请人 百度在线网络技术(北京)有限公司

地址 100085 北京市海淀区上地十街10号

百度大厦三层

(72)发明人 陈果果 牛飞 王芑 邢仁泰

张涛

(74)专利代理机构 北京同立钧成知识产权代理

有限公司 11205

代理人 荣甜甜 刘芳

(51)Int.Cl.

G10L 15/22(2006.01)

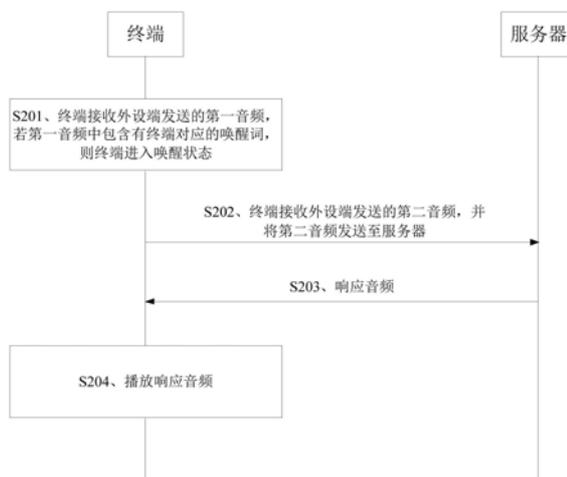
权利要求书3页 说明书11页 附图6页

(54)发明名称

语音交互方法、装置和存储介质

(57)摘要

本发明提供一种语音交互方法、装置和存储介质,该方法包括:接收外设端发送的第一音频,若第一音频中包含有终端对应的唤醒词,则终端进入唤醒状态;接收外设端发送的第二音频,并将第二音频发送至服务器,以使服务器根据第二音频向终端返回响应音频;接收服务器发送的响应音频,并播放响应音频。本发明通过外设端进行收音,进而使得终端与服务器之间进行交互获取响应音频,丰富了外设端与终端的交互功能,提高了用户体验。



1. 一种语音交互方法,应用于终端,其特征在于,包括:

接收外设端发送的第一音频,若所述第一音频中包含有终端对应的唤醒词,则所述终端进入唤醒状态;

接收所述外设端发送的第二音频,并将所述第二音频发送至服务器,以使所述服务器根据所述第二音频向所述终端返回响应音频;

接收所述服务器发送的所述响应音频,并播放所述响应音频。

2. 根据权利要求1所述的方法,其特征在于,在所述终端将所述第二音频发送至服务器之后,还包括:

接收所述服务器发送的停止发送消息,所述停止发送消息用于指示所述终端停止向所述服务器发送音频,所述停止发送消息是所述服务器在接收到所述第二音频之后的第一预设时长内,未接收到所述终端发送的第三音频发送的。

3. 根据权利要求2所述的方法,其特征在于,所述外设端为持续收音的外设端;

所述接收所述外设端发送的第二音频,并将所述第二音频发送至服务器之后,还包括:

接收所述外设端发送的第三音频,并将所述第三音频存储在所述终端中,所述第三音频为:接收到所述第二音频的所述第一预设时长后接收到的音频。

4. 根据权利要求1所述的方法,其特征在于,所述播放所述响应音频之后,还包括:

若在第二预设时长内未接收到所述外设端发送的新的音频,则进入休眠状态,并向所述外设端发送休眠消息,所述休眠消息用于指示所述外设端结束收音。

5. 根据权利要求1所述的方法,其特征在于,所述接收所述外设端发送的第一音频之前,还包括:

向外设端发送收音指令,所述收音指令用于指示所述外设端开始收音。

6. 一种语音交互方法,应用于外设端,其特征在于,包括:

向终端发送第一音频,若所述第一音频中包含有终端对应的唤醒词,则所述终端进入唤醒状态;

向所述终端发送第二音频,以使所述终端将所述第二音频发送至服务器,使得所述服务器根据所述第二音频向所述终端返回响应音频。

7. 根据权利要求6所述的方法,其特征在于,所述向所述终端发送第二音频之后,还包括:

向所述终端发送第三音频,所述第三音频的发送时间与所述第二音频的发送时间之差大于第一预设时长。

8. 根据权利要求7所述的方法,其特征在于,所述方法还包括:

接收所述终端发送的休眠消息;

结束收音。

9. 根据权利要求6所述的方法,其特征在于,所述向所述终端发送第一音频之前,还包括:

接收所述终端发送的收音指令,并根据所述收音指令开始收音。

10. 一种语音交互装置,其特征在于,包括:

第一音频发送模块,用于接收外设端发送的第一音频,若所述第一音频中包含有语音交互装置对应的唤醒词,则所述语音交互装置进入唤醒状态;

第二音频发送模块,用于接收所述外设端发送的第二音频,并将所述第二音频发送至服务器,以使所述服务器根据所述第二音频向所述语音交互装置返回响应音频;

播放模块,用于接收所述服务器发送的所述响应音频,并播放所述响应音频。

11. 根据权利要求10所述的装置,其特征在于,所述语音交互装置还包括:停止发送消息接收模块;

所述停止发送消息接收模块,用于接收所述服务器发送的停止发送消息,所述停止发送消息用于指示所述语音交互装置停止向所述服务器发送音频,所述停止发送消息是所述服务器在接收到所述第二音频之后的第一预设时长内,未接收到所述语音交互装置发送的第三音频发送的。

12. 根据权利要求11所述的装置,其特征在于,所述外设端为持续收音的外设端;所述语音交互装置还包括:第三音频接收模块;

所述第三音频接收模块,用于接收所述外设端发送的第三音频,并将所述第三音频存储在所述语音交互装置中,所述第三音频的接收时间与所述第二音频的接收时间之差大于所述第一预设时长。

13. 根据权利要求10所述的装置,其特征在于,所述语音交互装置还包括:休眠模块;

所述休眠模块,用于若在第二预设时长内未接收到所述外设端发送的新的收音,则进入休眠状态,并向所述外设端发送休眠消息,所述休眠消息用于指示所述外设端结束收音。

14. 根据权利要求10所述的装置,其特征在于,所述语音交互装置还包括:收音指令发送模块;

所述收音指令发送模块,用于向外设端发送收音指令,所述收音指令用于指示所述外设端开始收音。

15. 一种语音交互装置,其特征在于,包括:

收音指令接收模块,用于接收终端发送的收音指令,并根据所述收音指令开始收音;

第一音频发送模块,用于向所述终端发送第一音频,若所述第一音频中包含有终端对应的唤醒词,则所述终端进入唤醒状态;

第二音频发送模块,用于向所述终端发送第二音频,以使所述终端将所述第二音频发送至服务器,使得所述服务器根据所述第二音频向所述终端返回响应音频。

16. 根据权利要求15所述的装置,其特征在于,所述语音交互装置还包括:第三音频发送模块;

所述第三音频发送模块,用于向所述终端发送第三音频,所述第三音频的发送时间与所述第二音频的发送时间之差大于第一预设时长。

17. 根据权利要求16所述的装置,其特征在于,所述语音交互装置还包括:结束收音模块;

所述结束收音模块,用于接收所述终端发送的休眠消息;结束收音。

18. 根据权利要求15所述的装置,其特征在于,所述语音交互装置还包括:收音指令接收模块;

所述收音指令接收模块,用于接收终端发送的收音指令,并根据所述收音指令开始收音。

19. 一种终端,其特征在于,包括:至少一个处理器和存储器;

所述存储器存储计算机执行指令；

所述至少一个处理器执行所述存储器存储的计算机执行指令，使得所述终端执行权利要求1-5任一项所述的方法。

20. 一种外设端，其特征在于，包括：至少一个处理器和存储器；

所述存储器存储计算机执行指令；

所述至少一个处理器执行所述存储器存储的计算机执行指令，使得所述外设端执行权利要求6-9任一项所述的方法。

21. 一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储有计算机执行指令，当所述计算机执行指令被处理器执行时，实现权利要求1-5任一项所述的方法。

22. 一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储有计算机执行指令，当所述计算机执行指令被处理器执行时，实现权利要求6-9任一项所述的方法。

语音交互方法、装置和存储介质

技术领域

[0001] 本发明涉及语音交互技术领域,尤其涉及一种语音交互方法、装置和存储介质。

背景技术

[0002] 蓝牙(Bluetooth)是一种无线技术标准,可实现固定设备、移动设备和楼宇个人域网之间的短距离数据交换;终端与蓝牙设备进行连接后,根据蓝牙设备的品类,可对蓝牙设备进行对应的操作;如蓝牙设备为蓝牙音箱时,终端可以通过蓝牙设备播放音乐。

[0003] 现有技术中,终端与蓝牙设备之间的交互功能单一,不符合目前设备智能化的方向,用户体验差。

发明内容

[0004] 本发明提供一种语音交互方法、装置和存储介质,通过外设端进行收音,进而使得终端与服务器之间进行交互,丰富了外设端与终端的交互功能,提高了用户体验。

[0005] 本发明的第一方面提供一种语音交互方法,应用于终端,包括:

[0006] 接收所述外设端发送的第一音频,若所述第一音频中包含有终端对应的唤醒词,则所述终端进入唤醒状态;

[0007] 接收所述外设端发送的第二音频,并将所述第二音频发送至服务器,以使所述服务器根据所述第二音频向所述终端返回响应音频;

[0008] 接收所述服务器发送的所述响应音频,并播放所述响应音频。

[0009] 可选的,在所述终端将所述第二音频发送至服务器之后,还包括:

[0010] 接收所述服务器发送的停止发送消息,所述停止发送消息用于指示所述终端停止向所述服务器发送收音,所述停止发送消息是所述服务器在接收到所述第二音频之后的第一预设时长内,未接收到所述终端发送的第三音频发送的。

[0011] 可选的,所述外设端为持续收音的外设端;

[0012] 所述接收所述外设端发送的第二音频,并将所述第二音频发送至服务器之后,还包括:

[0013] 接收所述外设端发送的第三音频,并将所述第三音频存储在所述终端中,所述第三音频的接收时间与所述第二音频的接收时间之差大于所述第一预设时长。

[0014] 可选的,所述播放所述响应音频之后,还包括:

[0015] 若在第二预设时长内未接收到所述外设端发送的新的音频,则进入休眠状态,并向所述外设端发送休眠消息,所述休眠消息用于指示所述外设端结束收音。

[0016] 可选的,所述接收所述外设端发送的第一音频之前,还包括:

[0017] 向外设端发送收音指令,所述收音指令用于指示所述外设端开始收音。

[0018] 本发明的第二方面提供一种语音交互方法,应用于外设端,包括:

[0019] 向所述终端发送第一音频,若所述第一音频中包含有终端对应的唤醒词,则所述终端进入唤醒状态;

- [0020] 向所述终端发送第二音频,以使所述终端将所述第二音频发送至服务器,使得所述服务器根据所述第二音频向所述终端返回响应音频。
- [0021] 可选的,所述向所述终端发送第二音频之后,还包括:
- [0022] 向所述终端发送第三音频,所述第三音频的发送时间与所述第二音频的发送时间之差大于第一预设时长。
- [0023] 可选的,所述方法还包括:接收所述终端发送的休眠消息;
- [0024] 结束收音。
- [0025] 可选的,所述向所述终端发送第一音频之前,还包括:
- [0026] 接收所述终端发送的收音指令,并根据所述收音指令开始收音。
- [0027] 本发明的第三方面提供一种语音交互装置,包括:
- [0028] 第一音频发送模块,用于接收所述外设端发送的第一音频,若所述第一音频中包含有语音交互装置对应的唤醒词,则所述语音交互装置进入唤醒状态;
- [0029] 第二音频发送模块,用于接收所述外设端发送的第二音频,并将所述第二音频发送至服务器,以使所述服务器根据所述第二音频向所述语音交互装置返回响应音频;
- [0030] 播放模块,用于接收所述服务器发送的所述响应音频,并播放所述响应音频。
- [0031] 可选的,所述装置还包括:停止发送消息接收模块;
- [0032] 所述停止发送消息接收模块,用于接收所述服务器发送的停止发送消息,所述停止发送消息用于指示所述语音交互装置停止向所述服务器发送收音,所述停止发送消息是所述服务器在接收到所述第二音频之后的第一预设时长内,未接收到所述语音交互装置发送的第三音频发送的。
- [0033] 可选的,所述外设端为持续收音的外设端。
- [0034] 可选的,所述装置还包括:第三音频接收模块;
- [0035] 所述第三音频接收模块,用于接收所述外设端发送的第三音频,并将所述第三音频存储在所述语音交互装置中,所述第三音频的接收时间与所述第二音频的接收时间之差大于所述第一预设时长。
- [0036] 可选的,所述装置还包括:休眠模块;
- [0037] 所述休眠模块,用于若在第二预设时长内未接收到所述外设端发送的新的收音,则进入休眠状态,并向所述外设端发送休眠消息,所述休眠消息用于指示所述外设端结束收音。
- [0038] 可选的,所述装置还包括:收音指令发送模块;
- [0039] 所述收音指令发送模块,用于向外设端发送收音指令,所述收音指令用于指示所述外设端开始收音。
- [0040] 本发明的第四方面提供一种语音交互装置,包括:
- [0041] 第一音频发送模块,用于向所述终端发送第一音频,若所述第一音频中包含有终端对应的唤醒词,则所述终端进入唤醒状态;
- [0042] 第二音频发送模块,用于向所述终端发送第二音频,以使所述终端将所述第二音频发送至服务器,使得所述服务器根据所述第二音频向所述终端返回响应音频。
- [0043] 可选的,所述装置还包括:第三音频发送模块;
- [0044] 所述第三音频发送模块,用于向所述终端发送第三音频,所述第三音频的发送时

间与所述第二音频的发送时间之差大于第一预设时长。

[0045] 可选的,所述装置还包括:结束收音模块;

[0046] 所述结束收音模块,用于接收所述终端发送的休眠消息;结束收音。

[0047] 可选的,所述装置还包括:收音指令接收模块;

[0048] 所述收音指令接收模块,用于接收终端发送的收音指令,并根据所述收音指令开始收音。

[0049] 本发明的第五方面提供一种终端,包括:至少一个处理器和存储器;

[0050] 所述存储器存储计算机执行指令;

[0051] 所述至少一个处理器执行所述存储器存储的计算机执行指令,使得所述终端执行上述第一方面的语音交互方法。

[0052] 本发明的第六方面提供一种外设端,包括:至少一个处理器和存储器;

[0053] 所述存储器存储计算机执行指令;

[0054] 所述至少一个处理器执行所述存储器存储的计算机执行指令,使得所述外设端执行上述第二方面的语音交互方法。

[0055] 本发明的第七方面提供一种计算机可读存储介质,所述计算机可读存储介质上存储有计算机执行指令,当所述计算机执行指令被处理器执行时,实现上述第一方面的语音交互方法。

[0056] 本发明的第八方面提供一种计算机可读存储介质,所述计算机可读存储介质上存储有计算机执行指令,当所述计算机执行指令被处理器执行时,实现上述第二方面的语音交互方法。

[0057] 本发明提供一种语音交互方法、装置和存储介质,该方法包括:接收外设端发送的第一音频,若第一音频中包含有终端对应的唤醒词,则终端进入唤醒状态;接收外设端发送的第二音频,并将第二音频发送至服务器,以使服务器根据第二音频向终端返回响应音频;接收服务器发送的响应音频,并播放响应音频。本发明通过外设端进行收音,进而使得终端与服务器之间进行交互,丰富了外设端与终端的交互功能,提高了用户体验。

附图说明

[0058] 图1为本发明提供的语音交互方法适用的场景示意图;

[0059] 图2为本发明提供的语音交互方法的流程示意图一;

[0060] 图3为本发明提供的语音交互方法的流程示意图二;

[0061] 图4为本发明提供的终端的界面示意图;

[0062] 图5为本发明提供的语音交互方法的流程示意图三;

[0063] 图6为本发明提供的一语音交互装置的结构示意图一;

[0064] 图7为本发明提供的一语音交互装置的结构示意图二;

[0065] 图8为本发明提供的一语音交互装置的结构示意图三;

[0066] 图9为本发明提供的又一语音交互装置的结构示意图一;

[0067] 图10为本发明提供的又一语音交互装置的结构示意图二;

[0068] 图11为本发明提供的又一语音交互装置的结构示意图三。

具体实施方式

[0069] 为使本发明的目的、技术方案和优点更加清楚，下面将结合本发明的实施例，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

[0070] 现有的技术中的蓝牙外设端多种多样，如蓝牙耳机、蓝牙音箱、蓝牙键盘、运动手环等，这些蓝牙外设端在使用前，需要与终端建立蓝牙连接；示例性的，蓝牙音箱与终端建立的蓝牙连接的过程为：长按蓝牙音箱的电源键，使得蓝牙音箱打开，在终端上搜索蓝牙音箱的名称，输入配对密码，则可建立蓝牙连接。

[0071] 在建立连接后，终端可以通过蓝牙音箱播放终端上的歌曲或者其他音频，该音频可以是存储在终端的本地文件夹中的，也可以是终端与服务器交互获取的即时音频；终端将需要播放的文件发送给蓝牙音箱，蓝牙音箱即能播放文件对应的音频。

[0072] 但现有技术中终端与蓝牙外设端之间的交互功能过于单一，只能实现在终端的控制下被动播放，不能与用户进行交互，用户体验差；且现有技术中能够与用户进行交互的设备为智能设备，其前提是可以与服务器建立连接，外设端的配置成本高。

[0073] 正是为了解决上述终端与蓝牙外设端之间的交互功能过于单一的问题，且在丰富二者之间的交互功能的同时，降低蓝牙外设端的配置成本；本发明提供了一种语音交互方式。图1为本发明提供的语音交互方法适用的场景示意图，如图1所示，本发明提供的语音交互方法的适用的场景中包括：外设端、终端和服务器。

[0074] 其中，外设端可以与终端建立蓝牙连接，具体的该蓝牙连接可以是现有技术中的基于经典蓝牙的数据通信，在终端的系统设置界面引导用户选择指定设备并完成配对；或者，终端可以与外设端建立DMA (DuerOS Mobile Accessories) 连接，示例性的，终端在想要与外设端建立DMA连接时，可以直接在终端的应用程序的界面完成外设端的扫描、配对和连接，不需要返回到终端的系统设置界面进行设置，再到应用程序的界面完成连接。对应的，本实施例中建立普通蓝牙连接时，外设端为普通的蓝牙设备；在终端与之建立DMA连接时，外设端为DMA设备，即支持DMA蓝牙协议的设备。具体的，当终端与外设端建立的是普通蓝牙连接时，具体的方式可以参照现有技术中的蓝牙连接方式；终端与外设端建立的是DMA连接的过程，具体在下述实施例中进行说明。

[0075] 本发明中的终端和服务器之间可以为无线连接或者有线连接，本发明中的终端可以为手机、个人数字助理 (PersonalDigital Assistant, PDA)、平板电脑、便携设备 (例如，便携式计算机、袖珍式计算机或手持式计算机) 等移动设备；也可以是台式计算机等固定设备。

[0076] 下面从终端和服务器之间交互的角度，对本发明提供的语音交互方法进行说明，图2为本发明提供的语音交互方法的流程示意图一，如图2所示，本实施例提供的语音交互方法可以包括：

[0077] S201，终端接收外设端发送的第一音频，若第一音频中包含有终端对应的唤醒词，则终端进入唤醒状态。

[0078] 本实施例中，终端对接收到外设端发送的第一音频逐个进行解析，具体的，该解析过程可以为：终端将获取的第一音频转化为文字。

[0079] 终端判断接收到的第一音频中是否有预设唤醒词,该唤醒词用于唤醒终端,具体的,是唤醒终端与服务器之间的交互。对应的,终端判断第一音频对应的文字数据中是否有唤醒词,或者终端判断第一音频对应的音频数据中是否有唤醒词对应的音频数据。在终端确定第一音频中有唤醒词时,即进入唤醒状态,即终端将携带有唤醒词的第一音频之后的收音发送给服务器。

[0080] 示例性的,唤醒词为“小度”,则第一音频对应的文字数据中具有“小度”时,或者第一音频对应的音频数据中具有“小度”对应的音频数据时,确定该第一音频中携带有唤醒词“小度”,则终端进入唤醒状态。

[0081] 本实施例中,可以想到的是,为了节省终端的内存,可以在终端中预设固定时长,在终端向外设端发送收音指令外设端开始收音后,若终端在固定时长内未接收到携带有唤醒词的音频,则可以与外设端断开连接,或者向外设端发送停止收音指令,以指示外设端停止收音;在真正需要外设端进行收音时,可以重新向外设端发送收音指令。

[0082] S202,终端接收外设端发送的第二音频,并将第二音频发送至服务器。

[0083] 终端在进入唤醒状态后,由于外设端一直在收音,终端可以接收到外设端发送的第二音频,本实施例中可以将终端进入唤醒状态后的音频作为有效音频。终端在接收到有效收音即第二音频之后,为了获取该第二音频对应的响应音频,可以将该第二音频发送给服务器。

[0084] 示例性的,用户在说完“小度”后,外设端将该第一音频发送给终端,终端进入唤醒状态,接着用户继续说“今天北京的天气如何”,外设端将该第二音频发送给终端,终端进一步的将该第二音频“今天北京的天气如何”发送给服务器,以得到该第二音频的响应数据。

[0085] S203,服务器根据第二音频向终端返回响应音频。

[0086] 本实施例中,服务器在接收到第二音频后,可以对第二音频进行解析。具体的,服务器对第二音频进行解析的过程可以为:将第二音频转化为文字,将文字进行切分处理,获取该文字对应的多个词语;再根据每个词语的词性获取目标词语,再根据目标词语的语义,获取该第二音频对应的响应音频。

[0087] 本实施例中可以采用切词工具如神经语言程序学(Neuro-Linguistic Programming,NLP)工具对第二音频对应的文字进行分词处理,获取文字对应的多个词语,如第二音频对应的文字为“今天北京的天气如何”,采用切词工具将该文字切分成多个词语,具体的切分后的词语可以是“今天”、“北京”、“的”、“天气”和“如何”。

[0088] 本实施例中,可选的,可根据获取的多个词语的词性,获取有效信息对应的目标词语,如将切分后的会话消息中的量词、副词、形容词等去掉,获取有效信息对应的目标词语,如名词和动词等,如将上述切分结果中的“如何”和“的”去掉,获取有效信息对应的目标词语,“今天”、“北京”和“天气”。服务器根据获取的目标词语确定用户问的是今天北京的天气,于是服务器可以向用户返回关于今天北京天气的响应音频,如“今天北京晴天,气温20度”。

[0089] 值得注意的是,当第二音频对应的文字为多句文字时,服务器可以先将文字进行分句处理,再对每个子句进行分词处理,再根据每个子句的中目标词语的语义获取每个子句对应的响应音频,将第二音频对应的多个响应音频按照子句在文字中的先后顺序发送给终端。

[0090] 示例性的,用户的第二音频对应的文字为“去北京有什么好玩的吗?哪里住宿性价比高?”,服务器将文字分隔为两个子句“北京有什么好玩的地方”和“哪里住宿性价比高”。再分别获取每个子句对应的目标词语,如“北京”、“好玩的”、“地方”和“住宿”、“性价比高”,则分别获取每个子句对应的响应音频,如分别为“北京好玩的地方有故宫、长城……”以及“在北京住宿你可以选择xx酒店”。

[0091] S204,终端接收服务器发送的响应音频,并播放响应音频。

[0092] 本实施例中,终端获取第二音频对应的响应音频后,自动播放该响应音频。可以想到的是,当响应音频对应的为多个子句的音频时,按照接收响应音频的时间先后顺序逐个对响应音频进行播放。

[0093] 本实施例中,为了节省终端的电量或者用户正在使用终端导致终端不方便播放响应音频时,可以在接收到该响应音频后,将该响应音频发送给外设端,由外设端对响应音频进行播放。具体的,该种实施方式下,外设端可以为具有音频播放功能的外设端,如蓝牙音箱、运动手环等。

[0094] 本实施例中采用外设端进行收音,相对于现有技术中的终端直接与服务器进行交互获取响应音频的方式;一方面,由于终端的收音效果有限,距离其一定距离可能不能准确收音,或者收音效果差,本实施例中采用外设端如带有Mic的车载支架,其收音效果更好;另一方面,还使得终端与蓝牙设备的交互更为多样化,提高用户体验。

[0095] 本实施例提供一种语音交互方法、装置和存储介质,该方法包括:向外设端发送收音指令,收音指令用于指示外设端开始收音;接收外设端发送的第一音频,若第一音频中包含有终端对应的唤醒词,则终端进入唤醒状态;接收外设端发送的第二音频,并将第二音频发送至服务器,以使服务器根据第二音频向终端返回响应音频;接收服务器发送的响应音频,并播放响应音频。本实施例通过外设端进行收音,进而使得终端与服务器之间进行交互,丰富了外设端与终端的交互功能,提高了用户体验。

[0096] 下面结合图3对本发明提供的语音交互方法进行进一步说明,图3为本发明提供的语音交互方法的流程示意图二,如图3所示,本实施例提供的语音交互方法可以包括:

[0097] S301,终端向外设端发送收音指令,收音指令用于指示外设端开始收音。

[0098] 本实施例中,终端向外设端发送收音指令之前,需与外设端建立DMA连接。

[0099] 现有技术中,终端与外设蓝牙设备之间建立蓝牙连接为:终端通过现有的蓝牙扫描方式,即蓝牙低功耗(Bluetooth Low Energy,ble)扫描获取可以连接的蓝牙设备,与蓝牙设备间先建立ble连接;该连接建立后,蓝牙设备向终端返回响应消息,该响应消息指示终端可以通过支持rfcomm协议的rfcomm链路进行与蓝牙设备的连接,终端在收到该响应消息后断开与蓝牙设备的ble连接,重新通过rfcomm链路与蓝牙设备进行连接。现有技术中的连接方式会导致ble链路正常状态下,影响进行rfcomm连接的成功率和速度。

[0100] 本实施例中的外设端为支持DMA协议的外设端,具体的,本实施例中对终端与外设端建立DMA连接方式做简要说明:终端在扫描的过程中,支持DMA协议的DMA外设端向终端发送广播包,该广播包中包含有指示该外设端支持DMA连接的标识信息,则终端直接通过rfcomm链路与外设端进行连接,解决了现有技术中的ble链路正常状态下,影响进行rfcomm连接的成功率和速度的问题。

[0101] 本实施例中的外设端具有收音功能,具体的,外设端可以为具有麦克Mic的车载支

架,具有收音功能的蓝牙音箱、蓝牙耳机、发光二极管(Light-Emitting Diode,LED)灯等设备。

[0102] 终端与外设端建立蓝牙连接或者DMA连接后,用户有语音交互需求时,如用户想要询问天气、播放歌曲时,可以在终端的界面上进行操作,以触发终端向外设端发送收音指令。图4为本发明提供的终端的界面示意图,如图4所示,终端在与外设端建立连接后,可在终端界面上显示外设端的名称,如外设端A;以及“开始收音”控件,用户通过点击或者其他操作选择该“开始收音”控件,以触发终端向外设端发送收音指令,具体的,该收音指令用于指示外设端开始收音。

[0103] S302,外设端接收终端发送的收音指令,并根据收音指令开始收音。

[0104] 本实施例中,外设端接收到终端的收音指令后,开始收音。

[0105] 具体的,本实施例中的外设端可以为持续收音的外设端,如具有Mic的车载支架,该车载支架为DMA设备时,由于该车载支架为持续收音的品类,即其会将接收到收音指令后的音频均发送给终端;相应的,终端中设置有存储音频的文件夹或者目录,在文件夹或者目录专门用于存储外设端发送的音频。

[0106] 可以想到的是,若终端触发收音指令后,外设端开始收音,外设端获取的音频为用户的唤醒音频。示例性的,用户在选择“开始收音”控件后,开始说话,如终端与服务器进行交互的唤醒词为“小度”,则用户会说“小度”,以唤醒终端与服务器进行交互。

[0107] 但在外设端收音的过程中,用户由于其他事宜未能及时输出唤醒词,则外设端获取的音频可以为周围环境中的汽车声,或者用户与其他用户进行交谈的会话,该种情况下,由于终端未接收到携带有唤醒词的音频,终端和服务器之间的交互未被唤醒,但外设端仍会收音,并发送给终端。

[0108] S303,终端接收外设端发送的第一音频,若第一音频中包含有终端对应的唤醒词,则终端进入唤醒状态。

[0109] S304,终端接收外设端发送的第二音频,并将第二音频发送至服务器。

[0110] S305,终端接收服务器发送的停止发送消息,停止发送消息用于指示终端停止向服务器发送音频。

[0111] 本实施例中,服务器中设置有第一预设时长,服务器在接收到终端发送的第二音频后,第一预设时间内若未再接收到终端发送的第三音频,则确定用户说话完毕,则根据第二音频获取对应的响应音频,且向终端发送停止发送消息。

[0112] 具体的,终端在接收到服务器发送的停止发送消息后,不再向服务器发送新的音频。可以想到的是,若本实施例中的外设端为持续收音的外设端,则其在向终端发送第二音频的第一预设时长之后,还能获取第三音频,则向终端继续发送第三音频,而终端接收到该第三音频后,不再将该第三音频发送给服务器进行处理。若本实施例中的外设端为可控收音的外设端,则终端在接收到服务器发送的停止发送消息后,可以向外设端发送停止发送消息,以便外设端停止收音,这样可以减小外设端的功耗。

[0113] S306,外设端向终端发送第三音频。

[0114] 本实施例中的外设端为持续收音的外设端,具体的,其在接收到终端的收音指令后,开始收音;即使在终端接收到服务器发送的停止发送消息后,还是会向终端发送第三音频,其中,第三音频的接收时间与第二音频的接收时间之差大于第一预设时长。

[0115] 可以想到的是,外设端属于持续收音的外设端,只要终端未向外设端发送停止收音的指令,外设端一直会将音频发送给终端,本实施例中,外设端在将第三音频发送给终端后,还会陆续将第四音频、第五音频发送给终端。

[0116] S307,终端将第三音频存储在终端中。

[0117] 本实施例中的终端中设置有存储音频的文件夹或者目录,终端将接收到的第三音频存储在对应的文件夹或者目录中。其中,由于终端接收到服务器的停止收音的消息,则不再将后续接收到的音频发送给服务器进行处理。

[0118] 可以想到的是,为了继续实现服务器与终端的交互,用户可以在终端上设置将第三音频发送给服务器,以使服务器对第三音频进行解析,为终端返回第三音频对应的响应音频。

[0119] S308,服务器根据第二音频向终端返回响应音频。

[0120] S309,终端接收服务器发送的响应音频,并播放响应音频。

[0121] 本实施例中的S303-S304、S308-S309的具体实施方式可参考上述实施例中S201-S202、S203-S204中的相关描述,在此不做赘述。

[0122] 本实施例中的终端与外设端建立DMA连接,解决了现有技术中的ble链路正常状态下,影响进行rfcomm连接的成功率和速度的问题;且服务器中预先设置有第一预设时长,在终端向服务器发送第二音频后的第一预设时长内,若终端未向服务器发送第三音频,则服务器向终端发送停止收音信息,指示终端不再向服务器发送新的收音,服务器对第二音频进行解析且返回响应音频,丰富了外设端与终端的交互功音频能,提高了用户体验。

[0123] 下面结合图5对本发明提供的语音交互方法进行进一步说明,图5为本发明提供的语音交互方法的流程图三,如图5所示,本实施例提供的语音交互方法可以包括:

[0124] S501,终端向外设端发送收音指令,收音指令用于指示外设端开始收音。

[0125] S502,外设端接收终端发送的收音指令,并根据收音指令开始收音。

[0126] S503,终端接收外设端发送的第一音频,若第一音频中包含有终端对应的唤醒词,则终端进入唤醒状态。

[0127] S504,终端接收外设端发送的第二音频,并将第二音频发送至服务器。

[0128] S505,终端接收服务器发送的停止发送消息,停止发送消息用于指示终端停止向服务器发送音频。

[0129] S506,服务器根据第二音频向终端返回响应音频。

[0130] S507,终端接收服务器发送的响应音频,并播放响应音频。

[0131] S508,终端若在第二预设时长内未接收到外设端发送的新的音频,则进入休眠状态,并向外设端发送休眠消息。

[0132] 本实施例中,终端中存储有第二预设时长,在终端在接收到第二音频后的第二预设时长内若未接收到外设端发送的新的音频,则确定用户没有新的语音交互的需求,则终端进入休眠状态,具体的还向外设端发送休眠消息。休眠消息用于指示外设端结束收音。

[0133] 其中,终端进入休眠状态可以是终端需要再次接收到携带有唤醒词的音频时,才能进入与服务器交互的唤醒状态。值得注意的是,本实施例中的第二预设时长大于,服务器从接收到第二音频至返回响应音频的时长。

[0134] S509,外设端结束收音。

[0135] 外设端在接收到终端发送的休眠消息后,确定用户没有新的语音交互的需求,则结束收音。等待下次终端的收音指令。

[0136] 本实施例中的S501-S505、S506-S507的具体实施方式可参考上述实施例中S301-S305、S308-S309中的相关描述,在此不做赘述。

[0137] 本实施例中,若终端在接收到第二音频后的第二预设时长内未接收到外设端发送的新的音频,则进入休眠状态,并向外设端发送休眠消息,休眠消息用于指示外设端结束收音。本实施例在用户无新的语音交互的需求时终端进入休眠状态,且外设端停止收音,以节省终端和外设端的功耗。

[0138] 图6为本发明提供的一语音交互装置的结构示意图一,如图6所示,该语音交互装置为终端,该语音交互装置600包括:第一音频发送模块601、第二音频发送模块602和播放模块603。

[0139] 第一音频发送模块601,用于接收外设端发送的第一音频,若第一音频中包含有语音交互装置对应的唤醒词,则语音交互装置进入唤醒状态。

[0140] 第二音频发送模块602,用于接收外设端发送的第二音频,并将第二音频发送至服务器,以使服务器根据第二音频向语音交互装置返回响应音频。

[0141] 播放模块603,用于接收服务器发送的响应音频,并播放响应音频。

[0142] 本实施例提供的语音交互装置与上述语音交互方法实现的原理和技术效果类似,在此不作赘述。

[0143] 可选的,图7为本发明提供的一语音交互装置的结构示意图二,如图7所示,该语音交互装置600包括:停止发送消息接收模块604、第三音频接收模块605和休眠模块606和收音指令发送模块607。

[0144] 停止发送消息接收模块604,用于接收服务器发送的停止发送消息,停止发送消息用于指示语音交互装置停止向服务器发送音频,停止发送消息是服务器在接收到第二音频之后的第一预设时长内,未接收到语音交互装置发送的第三音频发送的。

[0145] 可选的,外设端为持续收音的外设端。

[0146] 第三音频接收模块605,用于接收外设端发送的第三音频,并将第三音频存储在语音交互装置中,第三音频的接收时间与第二音频的接收时间之差大于第一预设时长。

[0147] 休眠模块606,用于若在第二预设时长内未接收到外设端发送的新的音频,则进入休眠状态,并向外设端发送休眠消息,休眠消息用于指示外设端结束收音。

[0148] 收音指令发送模块607,用于向外设端发送收音指令,所述收音指令用于指示所述外设端开始收音。

[0149] 图8为本发明提供的一语音交互装置的结构示意图三,如图8所示,该语音交互装置800包括:存储器801和至少一个处理器802。

[0150] 存储器801,用于存储程序指令。

[0151] 处理器802,用于在程序指令被执行时实现本实施例中的语音交互方法,具体实现原理可参见上述实施例,本实施例此处不再赘述。

[0152] 该语音交互装置800还可以包括及输入/输出接口803。

[0153] 输入/输出接口803可以包括独立的输出接口和输入接口,也可以为集成输入和输出的集成接口。其中,输出接口用于输出数据,输入接口用于获取输入的数据,上述输出的

数据为上述方法实施例中输出的统称,输入的数据为上述方法实施例中输入的统称。

[0154] 本发明还提供一种可读存储介质,可读存储介质中存储有执行指令,当语音交互装置的至少一个处理器执行该执行指令时,当计算机执行指令被处理器执行时,实现上述实施例中的语音交互方法。

[0155] 本发明还提供一种程序产品,该程序产品包括执行指令,该执行指令存储在可读存储介质中。语音交互装置的至少一个处理器可以从可读存储介质读取该执行指令,至少一个处理器执行该执行指令使得语音交互装置实施上述的各种实施方式提供的语音交互方法。

[0156] 图9为本发明提供的又一语音交互装置的结构示意图一,如图9所示,该语音交互装置为外设端,该语音交互装置900包括:第一音频发送模块901和第二音频发送模块902。

[0157] 第一音频发送模块901,用于向所述终端发送第一音频,若所述第一音频中包含有终端对应的唤醒词,则所述终端进入唤醒状态。

[0158] 第二音频发送模块902,用于向所述终端发送第二音频,以使所述终端将所述第二音频发送至服务器,使得所述服务器根据所述第二音频向所述终端返回响应音频。

[0159] 本实施例提供的语音交互装置与上述语音交互方法实现的原理和技术效果类似,在此不作赘述。

[0160] 可选的,图10为本发明提供的又一语音交互装置的结构示意图二,如图10所示,该语音交互装置900包括:第三音频发送模块903、结束收音模块904和收音指令接收模块905。

[0161] 所述第三音频发送模块903,用于向所述终端发送第三音频,所述第三音频的发送时间与所述第二音频的发送时间之差大于第一预设时长。

[0162] 所述结束收音模块904,用于接收所述终端发送的休眠消息;结束收音。

[0163] 收音指令接收模块905,用于接收终端发送的收音指令,并根据所述收音指令开始收音。

[0164] 图11为本发明提供的又一语音交互装置的结构示意图三,如图11所示,该语音交互装置1100包括:存储器1101和至少一个处理器1102。

[0165] 存储器1101,用于存储程序指令。

[0166] 处理器1102,用于在程序指令被执行时实现本实施例中的语音交互方法,具体实现原理可参见上述实施例,本实施例此处不再赘述。

[0167] 该语音交互装置1100还可以包括及输入/输出接口1103。

[0168] 输入/输出接口1103可以包括独立的输出接口和输入接口,也可以为集成输入和输出的集成接口。其中,输出接口用于输出数据,输入接口用于获取输入的数据,上述输出的数据为上述方法实施例中输出的统称,输入的数据为上述方法实施例中输入的统称。

[0169] 本发明还提供一种可读存储介质,可读存储介质中存储有执行指令,当语音交互装置的至少一个处理器执行该执行指令时,当计算机执行指令被处理器执行时,实现上述实施例中的语音交互方法。

[0170] 本发明还提供一种程序产品,该程序产品包括执行指令,该执行指令存储在可读存储介质中。语音交互装置的至少一个处理器可以从可读存储介质读取该执行指令,至少一个处理器执行该执行指令使得语音交互装置实施上述的各种实施方式提供的语音交互方法。

[0171] 在本发明所提供的几个实施例中,应该理解到,所揭露的装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0172] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0173] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用硬件加软件功能单元的形式实现。

[0174] 上述以软件功能单元的形式实现的集成的单元,可以存储在一个计算机可读存储介质中。上述软件功能单元存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)或处理器(英文:processor)执行本发明各个实施例所述方法的部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(英文:Read-Only Memory,简称:ROM)、随机存取存储器(英文:Random Access Memory,简称:RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0175] 在上述网络设备或者终端设备的实施例中,应理解,处理器可以是中央处理单元(英文:Central Processing Unit,简称:CPU),还可以是其他通用处理器、数字信号处理器(英文:Digital Signal Processor,简称:DSP)、专用集成电路(英文:Application Specific Integrated Circuit,简称:ASIC)等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本申请所公开的方法的步骤可以直接体现为硬件处理器执行完成,或者用处理器中的硬件及软件模块组合执行完成。

[0176] 最后应说明的是:以上各实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述各实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分或者全部技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

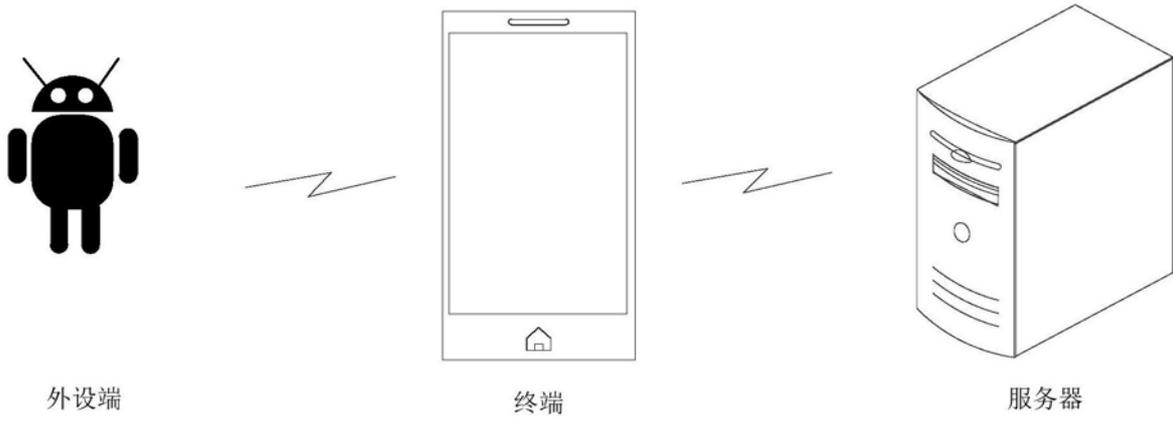


图1

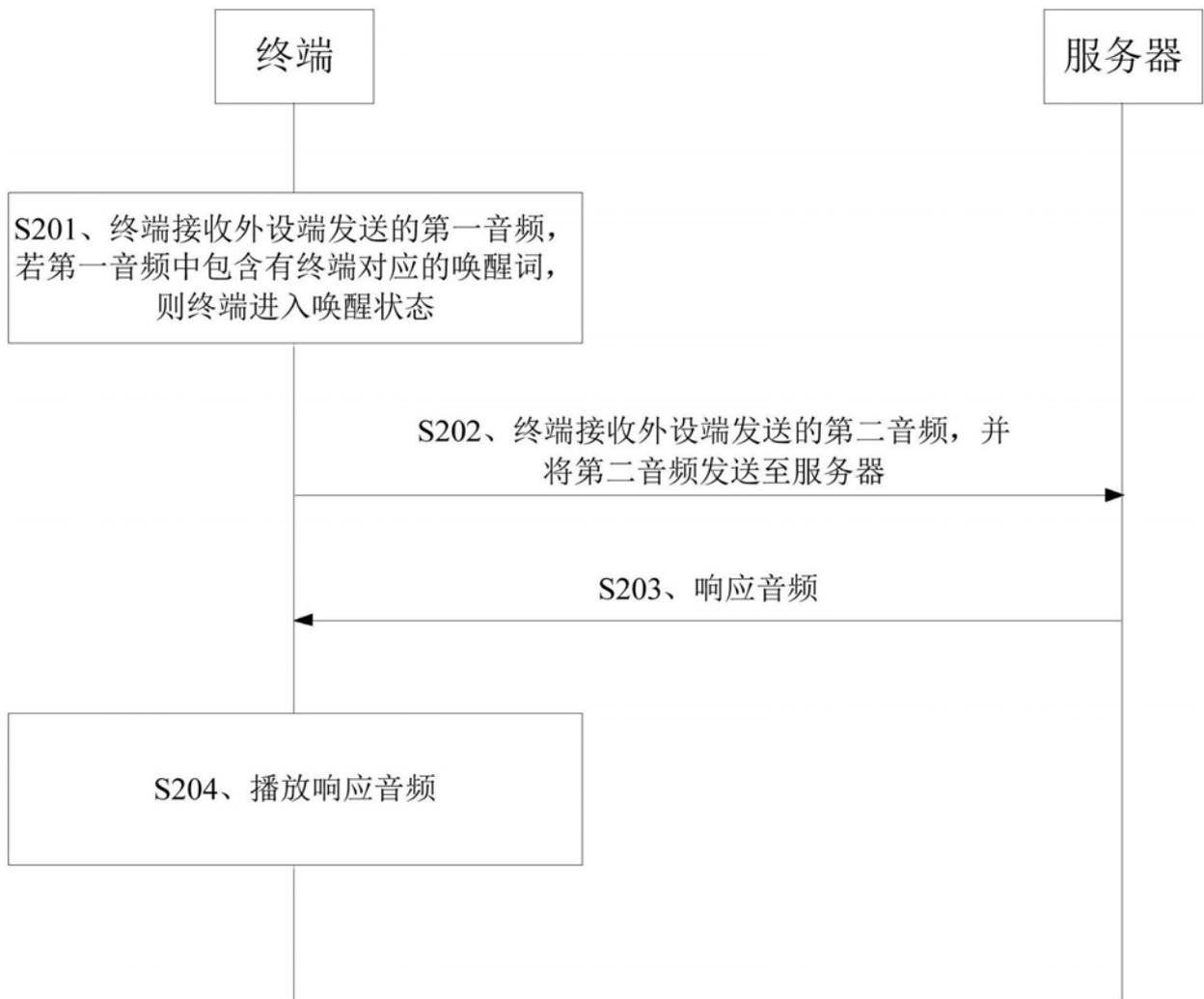


图2

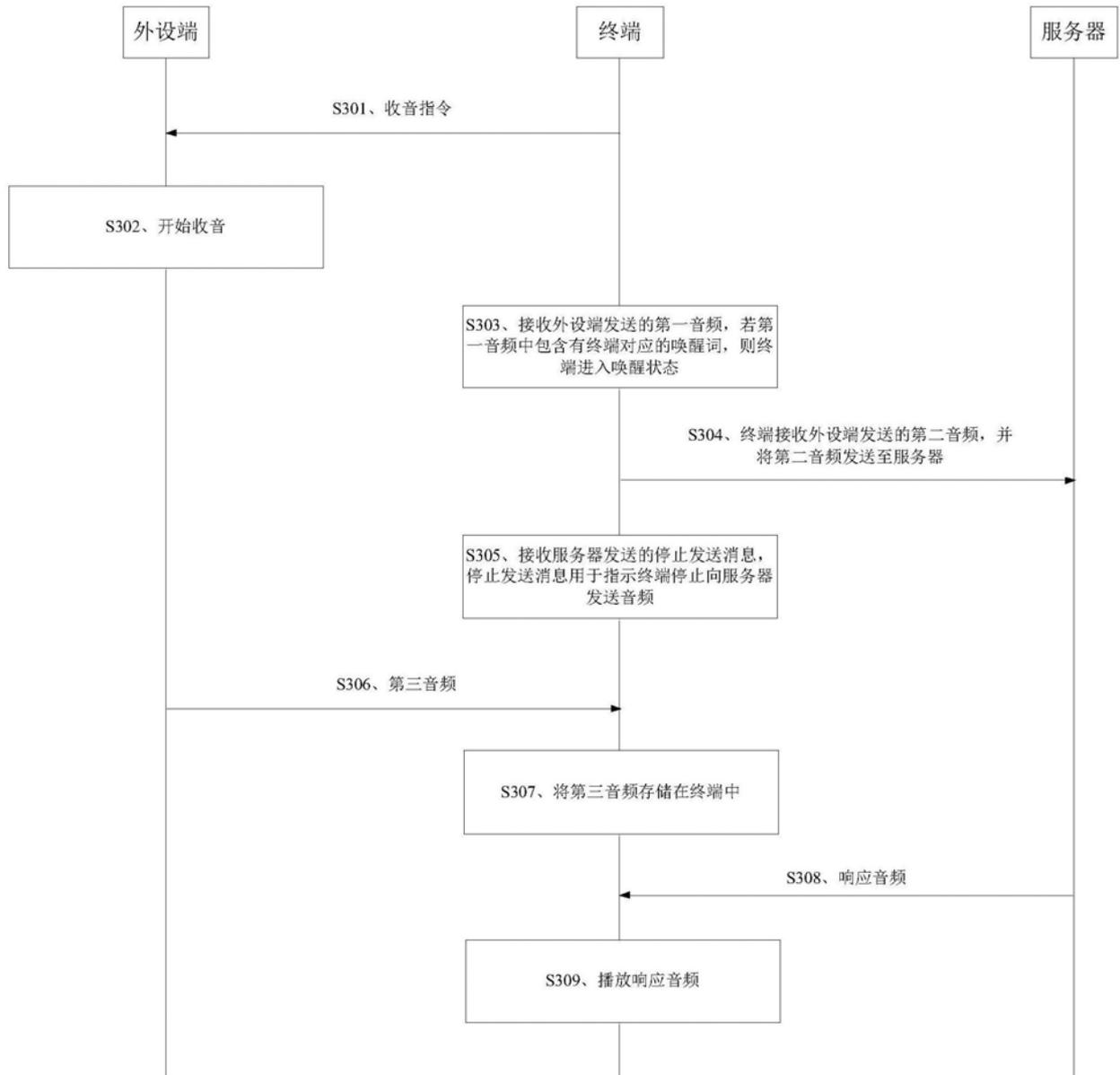


图3

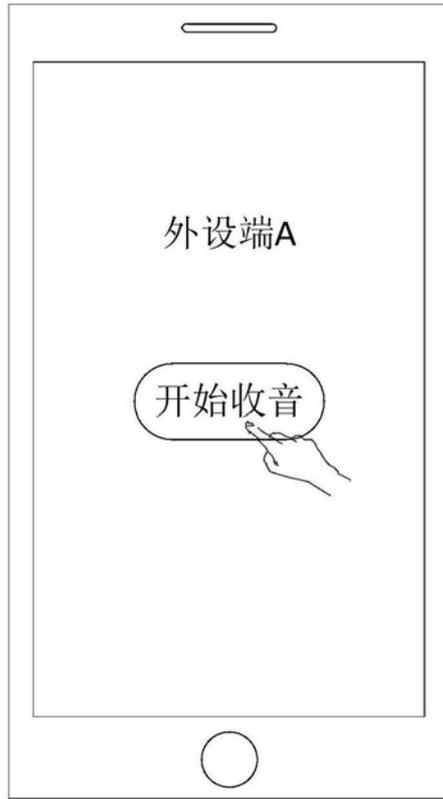


图4

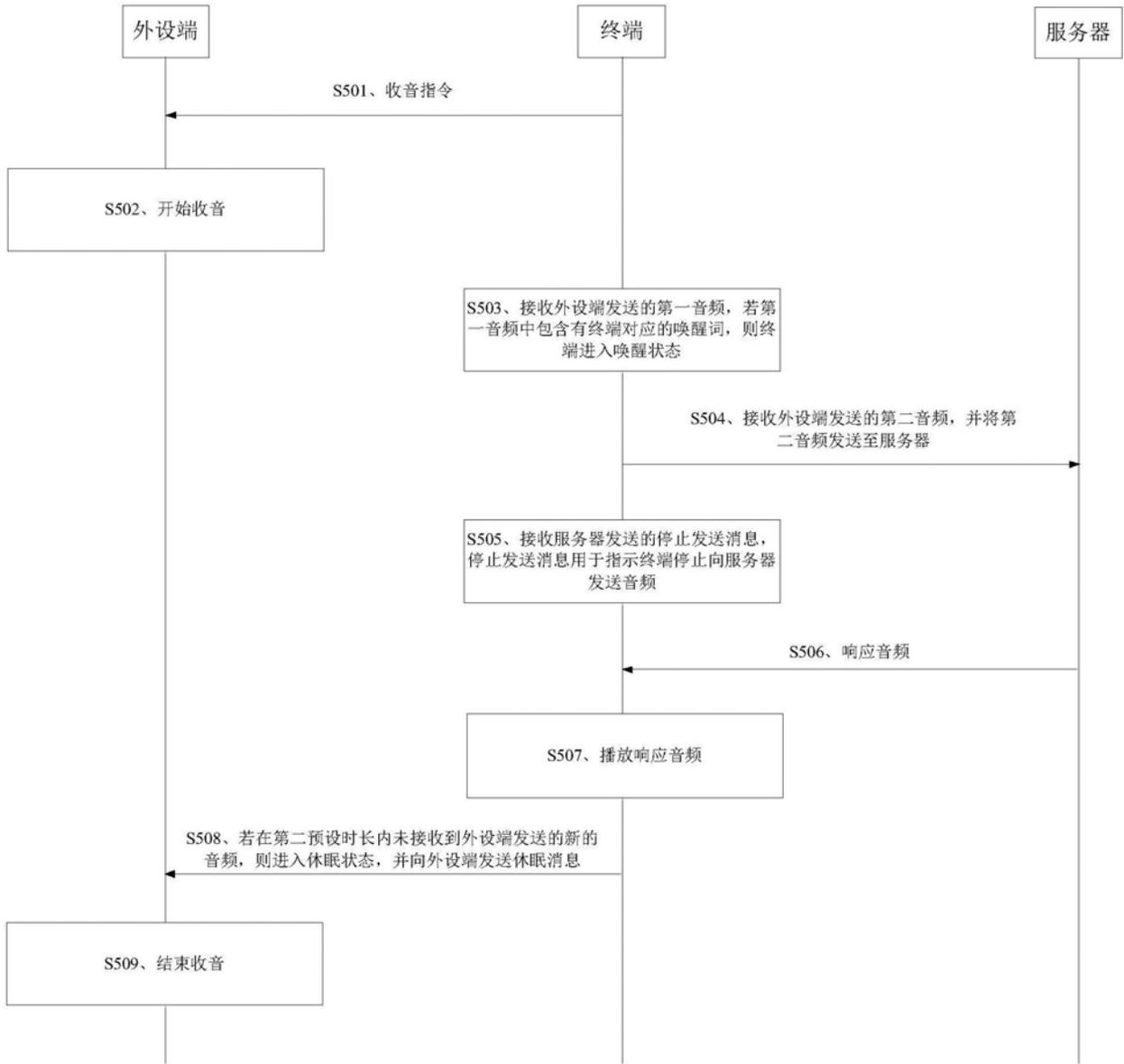


图5

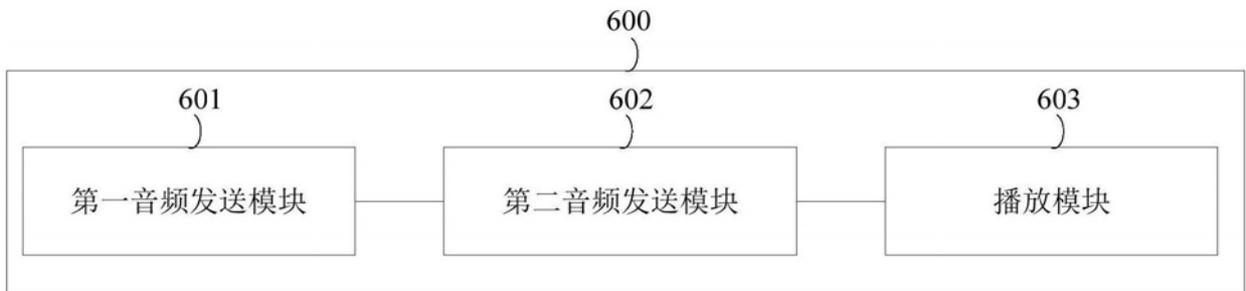


图6

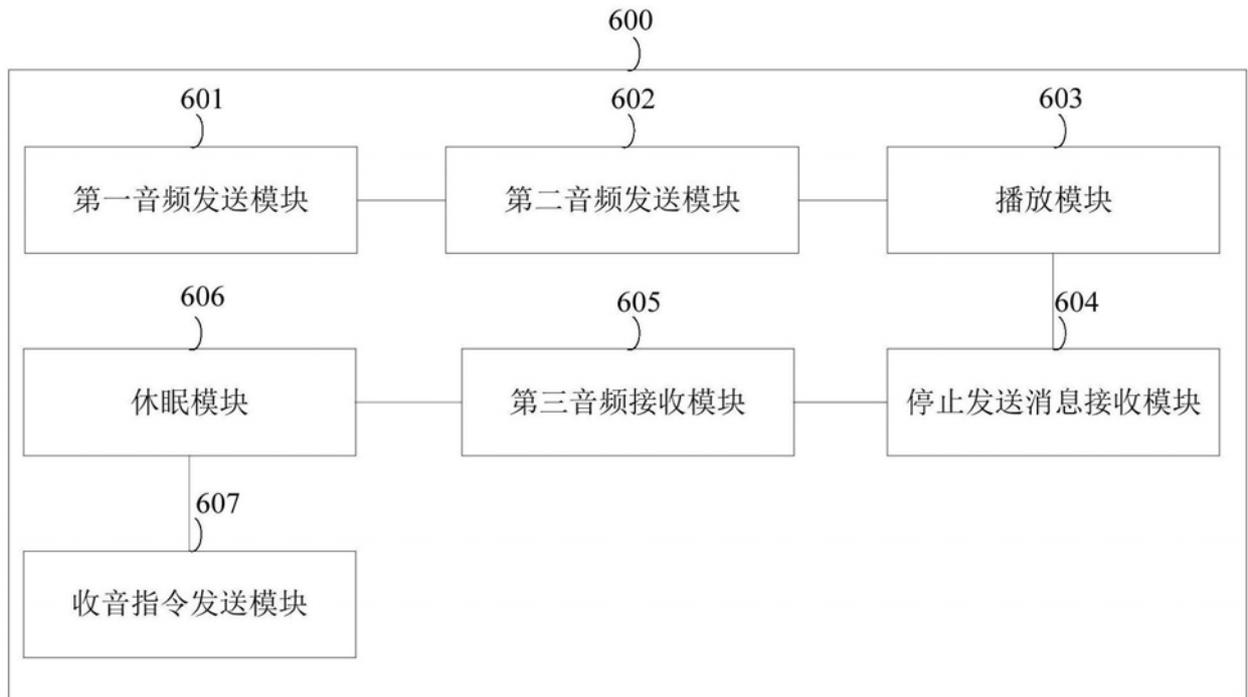


图7

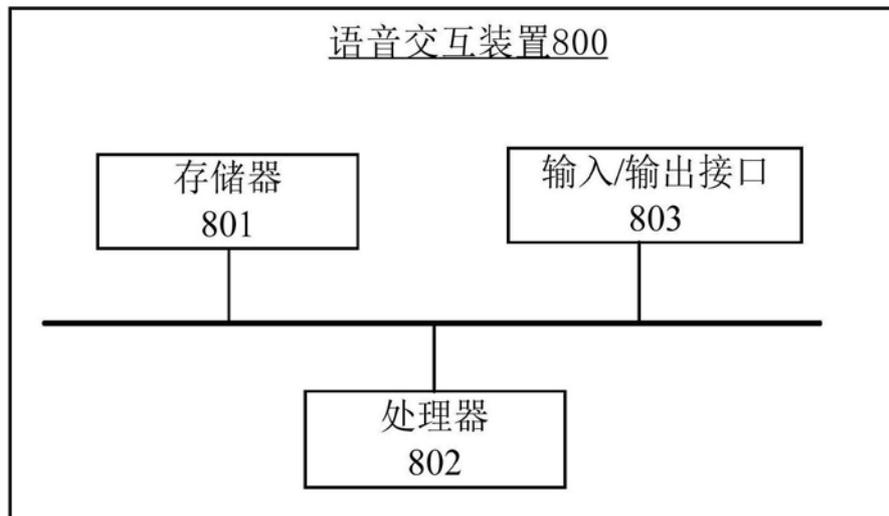


图8

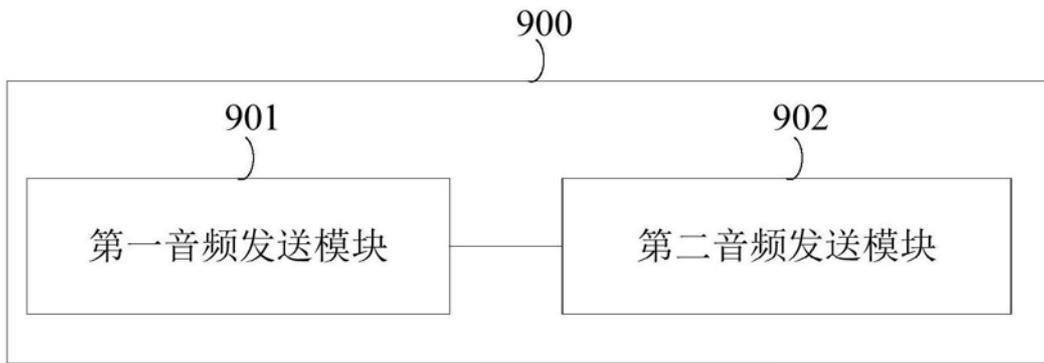


图9

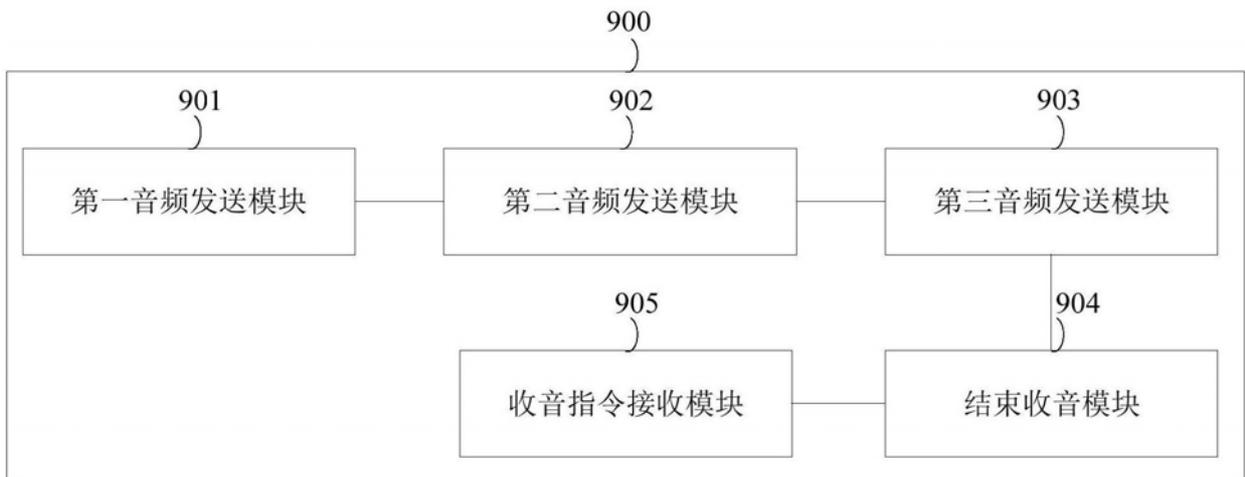


图10

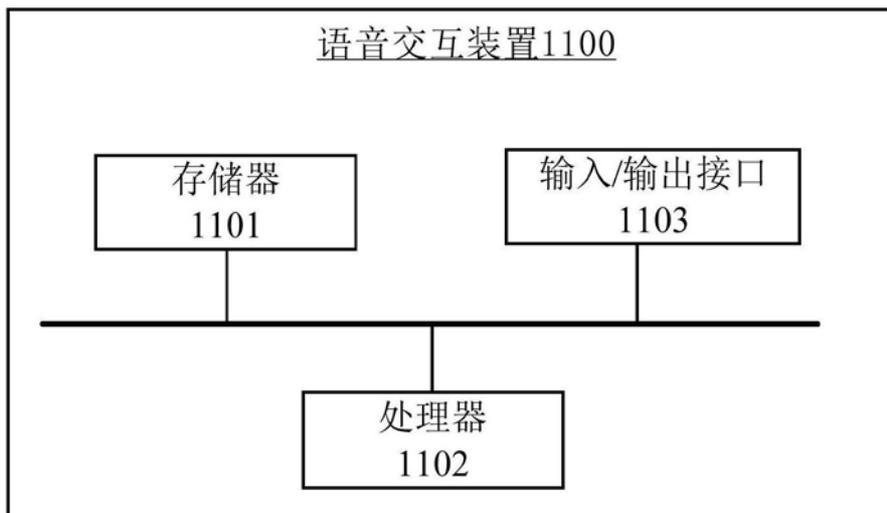


图11