



(12)发明专利申请

(10)申请公布号 CN 110827818 A

(43)申请公布日 2020.02.21

(21)申请号 201911138882.7

(22)申请日 2019.11.20

(71)申请人 腾讯科技(深圳)有限公司

地址 518000 广东省深圳市南山区高新区
科技中一路腾讯大厦35层

(72)发明人 孔秀哲

(74)专利代理机构 北京派特恩知识产权代理有
限公司 11270

代理人 王姗姗 张颖玲

(51) Int. Cl.

G10L 15/22(2006.01)

G10L 15/26(2006.01)

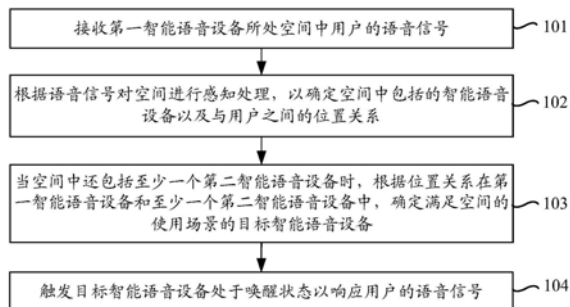
权利要求书4页 说明书26页 附图9页

(54)发明名称

一种智能语音设备的控制方法、装置、设备
及存储介质

(57)摘要

本发明提供了一种智能语音设备的控制方法、装置、电子设备及存储介质;方法包括:接收第一智能语音设备所处空间中用户的语音信号;根据语音信号对空间进行感知处理,以确定空间中包括的智能语音设备以及与用户之间的位置关系;当空间中还包括至少一个第二智能语音设备时,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,以及触发目标智能语音设备处于唤醒状态以响应用户的语音信号。通过本发明,能够在多个智能语音设备的复杂环境中实现对用户的语音信号的智能化响应,从而提高用户的体验感。



1. 一种智能语音设备的控制方法,其特征在于,所述方法包括:
接收第一智能语音设备所处空间中用户的语音信号;
根据所述语音信号对所述空间进行感知处理,以确定所述空间中包括的智能语音设备以及与所述用户之间的位置关系;
当所述空间中还包括至少一个第二智能语音设备时,根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语音设备,以及
触发所述目标智能语音设备处于唤醒状态以响应所述用户的语音信号。
2. 根据权利要求1所述的方法,其特征在于,所述根据所述语音信号对所述空间进行感知处理,以确定所述空间中包括的智能语音设备以及与所述用户之间的位置关系,包括:
对于所述空间中的任一智能语音设备所接收的所述用户的语音信号,执行以下处理:
对所述智能语音设备从多个方向接收的所述用户的语音信号进行解析处理,得到所述智能语音设备从多个方向接收所述用户的语音信号的能量值;
将最大能量值对应的方向确定为所述用户相对于所述智能语音设备的方向,并
根据语音信号的能量值随距离衰减的关系,以及所述最大能量值相对于所述用户的基准语音信号的能量值的衰减值,确定所述衰减值所对应的距离为所述智能语音设备与所述用户之间的距离。
3. 根据权利要求1所述的方法,其特征在于,所述根据所述语音信号对所述空间进行感知处理,以确定所述空间中包括的智能语音设备以及与所述用户之间的位置关系,包括:
对于所述空间中的任一智能语音设备所接收的所述用户的语音信号,执行以下处理:
对所述语音信号进行解析处理,得到所述智能语音设备与所述用户之间的第一距离、以及所述用户相对于所述智能语音设备的第一方向;
响应于接收到的语音信号,对所述空间进行障碍物探测,得到所述智能语音设备与所述用户之间的第二距离;
对所述空间进行障碍物识别,得到所述用户相对于所述智能语音设备的第二方向;
当所述第一距离与所述第二距离的距离差值大于距离误差阈值,和/或,所述第一方向与所述第二方向之间的方向误差大于方向误差阈值时,将所述第一距离与所述第二距离的加权值,确定为所述智能语音设备与所述用户之间的距离,并将所述第一方向与所述第二方向的平均值确定为所述用户相对于所述智能语音设备的方向。
4. 根据权利要求1所述的方法,其特征在于,所述根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语音设备,包括:
对于根据所述空间中任一智能设备接收的语音信号进行感知处理而得到的位置关系,执行以下处理:
当所述位置关系保持不变的时间超出时间阈值时,确定所述用户处于静止状态;
根据所述位置关系包括的所述智能语音设备与所述用户之间距离,将所述空间中与所述用户距离最小的智能语音设备确定为目标智能语音设备。
5. 根据权利要求1所述的方法,其特征在于,所述根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语

音设备,包括:

对于根据所述空间中任一智能设备接收的语音信号进行感知处理而得到的位置关系,执行以下处理:

当所述位置关系发生变化时,确定所述用户处于运动状态;

根据所述位置关系包括的所述用户相对于所述智能语音设备的方向,将所述方向的变化方向确定为所述用户相对于所述智能语音设备的移动方向;

根据所述位置关系中包括所述用户与所述智能设备之间的距离,将所述距离的倒数与所述移动方向的向量相乘,得到所述智能语音设备与所述用户之间的位置关系的匹配度;

在所述第一智能语音设备和所述至少一个第二智能语音设备中,将匹配度最高的智能语音设备确定为目标智能语音设备;

其中,当所述用户相对于所述智能语音设备的方向的变化为趋近所述智能语音设备时,所述移动方向取值为正,当所述用户相对于所述智能语音设备的方向的变化为远离所述智能语音设备时,所述移动方向取值为负。

6. 根据权利要求1所述的方法,其特征在于,所述根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语音设备,包括:

确定所述第一智能语音设备和所述至少一个第二智能语音设备中处于唤醒状态的智能语音设备;

当所述处于唤醒状态的智能语音设备与所述用户之间的距离未超出临界距离时,确定所述处于唤醒状态的智能语音设备为目标智能语音设备;

其中,所述临界距离为所述用户和所述智能语音设备能够彼此正确感知对方发出的语音信号时的最大距离。

7. 根据权利要求1所述的方法,其特征在于,所述根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语音设备,包括:

当所述第一智能语音设备和所述至少一个第二智能语音设备中存在正在与所述用户交互的智能语音设备,且

与所述用户之间的距离未超出临界距离时,确定正在与所述用户交互的智能语音设备为目标智能语音设备。

8. 根据权利要求1所述的方法,其特征在于,所述根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语音设备,包括:

确定在接收到所述语音信号之前处于唤醒状态的智能语音设备与所述用户之间的位置关系的变化趋势;

当根据所述位置关系的变化趋势确定所述处于唤醒状态的智能语音设备将超出临界距离时,在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备;

所述触发所述目标智能语音设备处于唤醒状态,包括:

当确定与用户的位置关系匹配度最高的智能语音设备为目标智能语音设备时,触发所

述处于唤醒状态的智能语音设备处于待机状态,实时唤醒所述目标智能语音设备。

9. 根据权利要求1所述的方法,其特征在于,所述根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语音设备,包括:

确定在接收到所述语音信号之前处于唤醒状态的智能语音设备与所述用户之间的位置关系的变化趋势;

当根据所述位置关系的变化趋势确定所述处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备;

所述触发所述目标智能语音设备处于唤醒状态,包括:

在所述处于唤醒状态的智能语音设备未超出所述临界距离之前,预先唤醒所述目标智能语音设备。

10. 根据权利要求1所述的方法,其特征在于,所述触发所述目标智能语音设备处于唤醒状态之后,所述方法还包括:

触发所述目标智能语音设备之外的处于唤醒状态的智能语音设备实时切换到待机状态;或者,

等待预设时间段,并在所述等待预设时间段内,针对所述目标智能语音设备之外的处于唤醒状态的智能语音设备,确定与所述用户之间的位置关系的变化趋势,当确定所述处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,触发所述处于唤醒状态的智能语音设备切换到待机状态。

11. 根据权利要求1-10任一项所述的方法,其特征在于,所述根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语音设备,包括:

基于所述第一智能语音设备以及至少一个第二智能语音设备所绑定的用户账号,确定所述账号对应多个智能语音设备时,在所述多个智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

12. 一种智能语音设备的控制装置,其特征在于,所述装置包括:

接收模块,用于接收第一智能语音设备所处空间中用户的语音信号;

感知模块,用于根据所述语音信号对所述空间进行感知处理,以确定所述空间中包括的智能语音设备以及与所述用户之间的位置关系;

处理模块,用于当所述空间中还包括至少一个第二智能语音设备时,根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语音设备,以及

触发模块,用于触发所述目标智能语音设备处于唤醒状态以响应所述用户的语音信号。

13. 一种智能语音设备,其特征在于,包括:

存储器,用于存储可执行指令;

处理器,用于执行所述存储器中存储的可执行指令时,实现权利要求1至11任一项所述的智能语音设备的控制方法。

14. 一种用于控制智能语音设备的服务器,其特征在于,包括:
存储器,用于存储可执行指令;
处理器,用于执行所述存储器中存储的可执行指令时,实现权利要求1至11任一项所述的智能语音设备的控制方法。

15. 一种存储介质,其特征在于,存储有可执行指令,用于引起处理器执行时,实现权利要求1至11任一项所述的智能语音设备的控制方法。

一种智能语音设备的控制方法、装置、设备及存储介质

技术领域

[0001] 本发明涉及人工智能技术,尤其涉及一种智能语音设备的控制方法、装置、电子设备及存储介质。

背景技术

[0002] 人工智能(Artificial Intelligence, AI)是计算机科学的一个综合技术,通过研究各种智能机器的设计原理与实现方法,使机器具有感知、推理与决策的功能。人工智能技术是一门综合学科,涉及领域广泛,例如自然语言处理技术以及机器学习/深度学习等几大方向,随着技术的发展,人工智能技术将在更多的领域得到应用,并发挥越来越重要的价值。

[0003] 随着计算机技术的发展,智能语音设备成为人工智能领域的重要应用之一,智能语音设备可以通过智能对话与即时问答的智能交互,有助于用户解决各种问题、即智能语音设备可以回答用户的问题,还可以满足用户提出的需求,例如用户需求为播放一首XX歌曲,则智能语音设备可以为用户播放该XX歌曲。

[0004] 然而,随着接入语音服务(VS, Voice Service)的智能语音设备越来越多,出现了不少在同一场景(同一个家里或同一个房间)下存在有多台智能语音设备的情况,这种情况下,如果用户唤醒智能语音设备,并发起语音请求,则会有多台智能语音设备同时响应并回复用户发起的语音请求,从而极大地降低了用户的体验感。

发明内容

[0005] 本发明实施例提供一种智能语音设备的控制方法、装置、电子设备及存储介质,能够在多个智能语音设备的复杂环境中实现对用户的语音信号的智能化响应,从而提高用户的体验感。

[0006] 本发明实施例的技术方案是这样实现的:

[0007] 本发明实施例提供一种智能语音设备的控制方法,所述方法包括:

[0008] 接收第一智能语音设备所处空间中用户的语音信号;

[0009] 根据所述语音信号对所述空间进行感知处理,以确定所述空间中包括的智能语音设备以及与所述用户之间的位置关系;

[0010] 当所述空间中还包括至少一个第二智能语音设备时,根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语音设备,以及

[0011] 触发所述目标智能语音设备处于唤醒状态以响应所述用户的语音信号。

[0012] 上述技术方案中,所述根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语音设备,包括:

[0013] 基于所述用户的语音信号的声纹特征,识别所述用户的语音信号所对应的用户;

[0014] 基于所述第一智能语音设备以及至少一个第二智能语音设备所绑定的用户账号,

确定所述账号所对应的用户为所述语音信号所对应的用户时,将所述账号对应的智能语音设备确定为可唤醒的智能语音设备;

[0015] 在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0016] 本发明实施例提供一种智能语音设备的控制装置,所述装置包括:

[0017] 接收模块,用于接收第一智能语音设备所处空间中用户的语音信号;

[0018] 感知模块,用于根据所述语音信号对所述空间进行感知处理,以确定所述空间中包括的智能语音设备以及与所述用户之间的位置关系;

[0019] 处理模块,用于当所述空间中还包括至少一个第二智能语音设备时,根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语音设备,以及

[0020] 触发模块,用于触发所述目标智能语音设备处于唤醒状态以响应所述用户的语音信号。

[0021] 上述技术方案中,所述感知模块还用于对于所述空间中的任一智能语音设备所接收的所述用户的语音信号,执行以下处理:

[0022] 对所述智能语音设备从多个方向接收的所述用户的语音信号进行解析处理,得到所述智能语音设备从多个方向接收所述用户的语音信号的能量值;

[0023] 将最大能量值对应的方向确定为所述用户相对于所述智能语音设备的方向,并

[0024] 根据语音信号的能量值随距离衰减的关系,以及所述最大能量值相对于所述用户的基准语音信号的能量值的衰减值,确定所述衰减值所对应的距离为所述智能语音设备与所述用户之间的距离。

[0025] 上述技术方案中,所述感知模块还用于对于所述空间中的任一智能语音设备所接收的所述用户的语音信号,执行以下处理:

[0026] 对所述语音信号进行解析处理,得到所述智能语音设备与所述用户之间的第一距离、以及所述用户相对于所述智能语音设备的第一方向;

[0027] 响应于接收到的语音信号,对所述空间进行障碍物探测,得到所述智能语音设备与所述用户之间的第二距离;

[0028] 对所述空间进行障碍物识别,得到所述用户相对于所述智能语音设备的第二方向;

[0029] 当所述第一距离与所述第二距离的距离差值大于距离误差阈值,和/或,所述第一方向与所述第二方向之间的方向误差大于方向误差阈值时,将所述第一距离与所述第二距离的加权值,确定为所述智能语音设备与所述用户之间的距离,并将所述第一方向与所述第二方向的平均值确定为所述用户相对于所述智能语音设备的方向。

[0030] 上述技术方案中,所述处理模块还用于对于根据所述空间中任一智能设备接收的语音信号进行感知处理而得到的位置关系,执行以下处理:

[0031] 当所述位置关系保持不变的时间超出时间阈值时,确定所述用户处于静止状态;

[0032] 根据所述位置关系包括的所述智能语音设备与所述用户之间距离,将所述空间中与所述用户距离最小的智能语音设备确定为目标智能语音设备。

[0033] 上述技术方案中,所述处理模块还用于对于根据所述空间中任一智能设备接收的

语音信号进行感知处理而得到的位置关系,执行以下处理:

[0034] 当所述位置关系发生变化时,确定所述用户处于运动状态;

[0035] 根据所述位置关系包括的所述用户相对于所述智能语音设备的方向,将所述方向的变化方向确定为所述用户相对于所述智能语音设备的移动方向;

[0036] 根据所述位置关系中包括所述用户与所述智能设备之间的距离,将所述距离的倒数与所述移动方向的向量相乘,得到所述智能语音设备与所述用户之间的位置关系的匹配度;

[0037] 在所述第一智能语音设备和所述至少一个第二智能语音设备中,将匹配度最高的智能语音设备确定为目标智能语音设备;

[0038] 其中,当所述用户相对于所述智能语音设备的方向的变化为趋近所述智能语音设备时,所述移动方向取值为正,当所述用户相对于所述智能语音设备的方向的变化为远离所述智能语音设备时,所述移动方向取值为负。

[0039] 上述技术方案中,所述处理模块还用于确定所述第一智能语音设备和所述至少一个第二智能语音设备中处于唤醒状态的智能语音设备;

[0040] 当所述处于唤醒状态的智能语音设备与所述用户之间的距离未超出临界距离时,确定所述处于唤醒状态的智能语音设备为目标智能语音设备;

[0041] 其中,所述临界距离为所述用户和所述智能语音设备能够彼此正确感知对方发出的语音信号时的最大距离。

[0042] 上述技术方案中,所述处理模块还用于当所述第一智能语音设备和所述至少一个第二智能语音设备中存在正在与所述用户交互的智能语音设备,且

[0043] 与所述用户之间的距离未超出临界距离时,确定正在与所述用户交互的智能语音设备为目标智能语音设备。

[0044] 上述技术方案中,处理模块还用于确定在接收到所述语音信号之前处于唤醒状态的智能语音设备与所述用户之间的位置关系的变化趋势;

[0045] 当根据所述位置关系的变化趋势确定所述处于唤醒状态的智能语音设备将超出临界距离时,在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备;

[0046] 所述触发模块还用于当确定与用户的位置关系匹配度最高的智能语音设备为目标智能语音设备时,触发所述处于唤醒状态的智能语音设备处于待机状态,实时唤醒所述目标智能语音设备。

[0047] 上述技术方案中,所述处理模块还用于确定在接收到所述语音信号之前处于唤醒状态的智能语音设备与所述用户之间的位置关系的变化趋势;

[0048] 当根据所述位置关系的变化趋势确定所述处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备;

[0049] 所述触发模块还用于在所述处于唤醒状态的智能语音设备未超出所述临界距离之前,预先唤醒所述目标智能语音设备。

[0050] 上述技术方案中,所述处理模块还用于确定在接收到所述语音信号之前处于唤醒状态的智能语音设备与所述用户之间的位置关系的变化趋势;

[0051] 当根据所述位置关系的变化趋势确定所述处于唤醒状态的智能语音设备不会超出临界距离时,确定所述处于唤醒状态的智能语音设备为目标智能语音设备。

[0052] 上述技术方案中,所述处理模块还用于获取在接收到所述语音信号之前处于唤醒状态的智能语音设备的历史数据,并结合所述已唤醒的智能语音设备的位置关系的变化趋势、使用时长以及唤醒次数,通过人工智能模型预测处于唤醒状态的智能语音设备的预计使用时长;

[0053] 在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备确定为目标智能语音设备;

[0054] 所述触发模块还用于当所述预计使用时长达到时,实时唤醒所述目标智能语音设备;或者,

[0055] 在所述预计使用时长到达之前,预先唤醒所述目标智能语音设备。

[0056] 上述技术方案中,所述装置还包括:

[0057] 切换模块,用于触发所述目标智能语音设备之外的处于唤醒状态的智能语音设备实时切换到待机状态;或者,

[0058] 等待预设时间段,并在所述等待预设时间段内,针对所述目标智能语音设备之外的处于唤醒状态的智能语音设备,确定与所述用户之间的位置关系的变化趋势,当确定所述处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,触发所述处于唤醒状态的智能语音设备切换到待机状态。

[0059] 上述技术方案中,所述装置还包括:

[0060] 响应模块,用于当所述目标智能语音设备与在接收到所述语音信号之前处于唤醒状态的智能语音设备不是同一设备,且在所述处于唤醒状态的智能语音设备最后一次响应所述用户的语音信号的过程中,所述处于唤醒状态的智能语音设备与所述用户之间的距离超出临界距离时,触发所述目标智能语音设备重新响应所述最后一次语音信号。

[0061] 上述技术方案中,处理模块还用于基于所述第一智能语音设备以及至少一个第二智能语音设备所绑定的用户账号,确定所述账号对应多个智能语音设备时,在所述多个智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0062] 上述技术方案中,所述处理模块还用于基于所述用户的语音信号的声纹特征,识别所述用户的语音信号所对应的用户;

[0063] 基于所述第一智能语音设备以及至少一个第二智能语音设备所绑定的用户账号,确定所述账号所对应的用户为所述语音信号所对应的用户时,将所述账号对应的智能语音设备确定为可唤醒的智能语音设备;

[0064] 在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0065] 本发明实施例提供一种智能语音设备,包括:

[0066] 存储器,用于存储可执行指令;

[0067] 处理器,用于执行所述存储器中存储的可执行指令时,实现本发明实施例提供的智能语音设备的控制方法。

[0068] 本发明实施例提供一种用于控制智能语音设备的服务器,包括:

- [0069] 存储器,用于存储可执行指令;
- [0070] 处理器,用于执行所述存储器中存储的可执行指令时,实现本发明实施例提供的智能语音设备的控制方法。
- [0071] 本发明实施例提供一种存储介质,存储有可执行指令,用于引起处理器执行时,实现本发明实施例提供的智能语音设备的控制方法。
- [0072] 本发明实施例具有以下有益效果:
- [0073] 通过根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,触发目标智能语音设备响应用户的语音信号,从而避免在同一场景下的智能语音设备都响应用户的语音信号,提高用户的体验感。

附图说明

- [0074] 图1是本发明实施例提供的智能语音设备的控制方法的一个可选的应用场景10示意图;
- [0075] 图2是本发明实施例提供的电子设备500的结构示意图;
- [0076] 图3A-3C是本发明实施例提供的智能语音设备的控制方法的流程示意图;
- [0077] 图4是本发明实施例提供的智能语音设备的控制方法的流程示意图;
- [0078] 图5是本发明实施例提供的用户唤醒智能设备的示意图;
- [0079] 图6是本发明实施例提供的智能语音设备的应用场景示意图;
- [0080] 图7是本发明实施例提供的智能语音设备与云端交互的应用场景示意图;
- [0081] 图8是本发明实施例提供的智能语音设备1上传给云端的语音数据的波形图;
- [0082] 图9是本发明实施例提供的智能语音设备1上传给云端的语音数据的频谱图;
- [0083] 图10是本发明实施例提供的智能语音设备2上传给云端的语音数据的波形图;
- [0084] 图11是本发明实施例提供的智能语音设备2上传给云端的语音数据的频谱图;
- [0085] 图12是本发明实施例提供的智能语音设备与云端交互的另一应用场景示意图。

具体实施方式

- [0086] 为了使本发明的目的、技术方案和优点更加清楚,下面将结合附图对本发明作进一步地详细描述,所描述的实施例不应视为对本发明的限制,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它实施例,都属于本发明保护的范围。
- [0087] 在以下的描述中,所涉及的术语“第一\第二”仅仅是是区别类似的对象,不代表针对对象的特定排序,可以理解地,“第一\第二”在允许的情况下可以互换特定的顺序或先后次序,以使这里描述的本发明实施例能够以除了在这里图示或描述的以外的顺序实施。
- [0088] 除非另有定义,本文所使用的所有的技术和科学术语与属于本发明的技术领域的技术人员通常理解的含义相同。本文中所使用的术语只是为了描述本发明实施例的目的,不是旨在限制本发明。
- [0089] 对本发明实施例进行进一步详细说明之前,对本发明实施例中涉及的名词和术语进行说明,本发明实施例中涉及的名词和术语适用于如下的解释。
- [0090] 1) 语音助手:一款智能型的终端应用,通过智能对话与即时问答的智能交互,以帮忙用户解决各种问题,主要帮忙用户解决生活类问题。

[0091] 2) 云端:又称云平台,是一种采用应用程序虚拟化技术(Application Virtualization)的软件平台,集软件搜索、下载、使用、管理、备份等多种功能为一体。通过该平台,各类常用软件都能够在独立的虚拟化环境中被封装起来,从而使应用软件不会与系统产生耦合,达到绿色使用软件的目的。

[0092] 3) 能量值:语音数据的能量值,当智能语音设备的语音数据的能量值越大,则说明该智能语音设备接收用户的语音信息越清晰,也就是该智能语音设备距离用户越近。通过波形图以及频谱图可以表示能量值,当波形图上波形的振幅越大,则语音数据的能量值越大、即语音数据的能量值与波形的振幅成正比关系;当频谱图上的高频区越活跃,则语音数据的能量值越大,当频谱图上的高频区越活跃,则语音数据的能量值越大、即语音数据的能量值与高频区的活跃度成正比关系。

[0093] 4) 语音识别:让机器通过识别和理解过程把语音信号转变为相应的文本或命令的过程。

[0094] 为至少解决相关技术的上述技术问题,本发明实施例提供一种智能语音设备的控制方法、装置、电子设备和存储介质,能够让目标智能语音设备处于唤醒状态以响应用户的语音信号,避免在同一场景下的智能语音设备都响应用户的语音信号,从而提高用户的体验感。下面说明本发明实施例提供的电子设备的示例性应用,实现本发明实施例提供的智能语音设备控制方案的电子设备可以是服务器,例如部署在云端的服务器,根据在同一空间中的第一智能语音设备以及至少一个第二智能语音设备提供的用户的语音信号,确定第一智能语音设备以及至少一个第二智能语音设备与用户之间的位置关系,并根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,并触发目标智能语音设备处于唤醒状态以响应用户的语音信号。

[0095] 实现本发明实施例提供的智能语音设备控制方案的电子设备也可是笔记本电脑,平板电脑,台式计算机,移动设备(例如,移动电话,个人数字助理)等各种具有智能语音功能的用户终端(智能语音设备),以第一智能语音设备为手持终端为例,其根据接收到的用户的语音信号以及在同一空间中的至少一个第二智能语音设备提供的用户的语音信号,确定智能语音设备与用户之间的位置关系,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,并触发目标智能语音设备处于唤醒状态以响应用户的语音信号。

[0096] 参见图1,图1是本发明实施例提供的智能语音设备的控制方法的一个可选的应用场景10示意图,终端200(示例性示出智能语音设备200-1、智能语音设备200-2、智能语音设备200-3)通过网络300连接服务器100,网络300可以是广域网或者局域网,又或者是二者的组合。

[0097] 终端200可以被用来接收用户的语音信号,例如,当用户发出语音信号后,终端自动采集用户的语音信号。

[0098] 在一些实施例中,终端200本地执行本发明实施例提供的智能语音设备的控制方法来完成根据用户的语音信号,对空间进行感知处理,以确定空间中包括的智能语音设备以及与用户之间的位置关系,并根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,触发目标智能语音设备处于唤醒状态以响应用户的语音信号,例如,在智能语音设备200-1(第一智能语音设备)上安装语

音助手,用户发出语音信号后,智能语音设备200-1采集用户的语音信号,并接收智能语音设备200-2和智能语音设备200-3(第二智能语音设备)采集的用户的语音信号,根据语音信号确定智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系,根据位置关系确定满足空间的使用场景的目标智能语音设备(第一智能语音设备以及第二智能语音设备中任一个智能语音设备),并通过语音助手触发目标智能语音设备处于唤醒状态以响应用户的语音信号。

[0099] 终端200也可以通过网络300向服务器100发送用户的语音信号,并调用服务器100提供的智能语音设备的控制功能,服务器100通过本发明实施例提供的智能语音设备的控制方法进行控制处理,例如,在终端200(智能语音设备)上安装语音助手,用户发出语音信号后,终端200通过语音助手采集用户的语音信号,并通过网络300向服务器100发送用户的语音信号,服务器100基于用户的语音信号,确定智能语音设备以及与用户之间的位置关系,并根据位置关系确定满足空间的使用场景的目标智能语音设备,并向目标智能语音设备发送控制指令,触发目标智能语音设备处于唤醒状态,并通过语音助手响应用户的语音信号。

[0100] 继续说明本发明实施例提供的实现智能语音设备控制方案的电子设备的结构,参见图2,图2是本发明实施例提供的电子设备500的结构示意图,图2所示的电子设备500包括:至少一个处理器510、存储器550、至少一个网络接口520和用户接口530。电子设备500中的各个组件通过总线系统540耦合在一起。可理解,总线系统540用于实现这些组件之间的连接通信。总线系统540除包括数据总线之外,还包括电源总线、控制总线和状态信号总线。但是为了清楚说明起见,在图2中将各种总线都标为总线系统540。

[0101] 处理器510可以是一种集成电路芯片,具有信号的处理能力,例如通用处理器、数字信号处理器(DSP, Digital Signal Processor),或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等,其中,通用处理器可以是微处理器或者任何常规的处理器等。

[0102] 用户接口530包括使得能够呈现媒体内容的一个或多个输出装置531,包括一个或多个扬声器和/或一个或多个视觉显示屏。用户接口530还包括一个或多个输入装置532,包括有助于用户输入的用户接口部件,比如键盘、鼠标、麦克风、触屏显示屏、摄像头、其他输入按钮和控件。

[0103] 存储器550包括易失性存储器或非易失性存储器,也可包括易失性和非易失性存储器两者。其中,非易失性存储器可以是只读存储器(ROM, Read Only Memory),易失性存储器可以是随机存取存储器(RAM, Random Access Memory)。本发明实施例描述的存储器550旨在包括任意适合类型的存储器。存储器550可选地包括在物理位置上远离处理器510的一个或多个存储设备。

[0104] 在一些实施例中,存储器550能够存储数据以支持各种操作,这些数据的示例包括程序、模块和数据结构或者其子集或超集,下面示例性说明。

[0105] 操作系统551,包括用于处理各种基本系统服务和执行硬件相关任务的系统程序,例如框架层、核心库层、驱动层等,用于实现各种基础业务以及处理基于硬件的任务;

[0106] 网络通信模块552,用于经由一个或多个(有线或无线)网络接口520到达其他计算设备,示例性的网络接口520包括:蓝牙、无线相容性认证(WiFi)、和通用串行总线(USB,

Universal Serial Bus)等;

[0107] 显示模块553,用于经由一个或多个与用户接口530相关联的输出装置531(例如,显示屏、扬声器等)使得能够呈现信息(例如,用于操作外围设备和显示内容和信息的用户接口);

[0108] 输入处理模块554,用于对一个或多个来自一个或多个输入装置532之一的一个或多个用户输入或互动进行检测以及翻译所检测的输入或互动。

[0109] 在一些实施例中,本发明实施例提供的智能语音设备的控制装置可以采用软硬件结合的方式实现,作为示例,本发明实施例提供的智能语音设备的控制装置可以是采用硬件译码处理器形式的处理器,其被编程以执行本发明实施例提供的智能语音设备的控制方法,例如,硬件译码处理器形式的处理器可以采用一个或多个应用专用集成电路(ASIC, Application Specific Integrated Circuit)、DSP、可编程逻辑器件(PLD, Programmable Logic Device)、复杂可编程逻辑器件(CPLD, Complex Programmable Logic Device)、现场可编程门阵列(FPGA, Field-Programmable Gate Array)或其他电子元件。

[0110] 在另一些实施例中,本发明实施例提供的智能语音设备的控制装置可以采用软件方式实现,图2示出了存储在存储器550中的智能语音设备的控制装置555,其可以是程序和插件等形式的软件,并包括一系列的模块,包括接收模块5551、感知模块5552、处理模块5553、触发模块5554、切换模块5555、以及响应模块5556;其中,接收模块5551、感知模块5552、处理模块5553、触发模块5554、切换模块5555、以及响应模块5556用于实现本发明实施例提供的智能语音设备的控制方法。

[0111] 下面结合本发明实施例上述提供的智能语音设备的示例性应用和实施,以第一智能语音设备为执行主体为例,说明本发明实施例提供的智能语音设备的控制方法。参见图3A,图3A是本发明实施例提供的智能语音设备的控制方法的流程示意图,结合图3A示出的步骤进行说明。

[0112] 在步骤101中,接收第一智能语音设备所处空间中用户的语音信号。

[0113] 当用户发出语音信号后,例如唤醒词“ABAB”,第一智能语音设备可以采集用户的语音信号,并通过局域网广播或者其他近距离通信方式传递采集到的用户的语音信号,第一智能语音设备也可以接收第一智能语音设备所处空间中第二智能语音设备(即,空间中除了第一智能语音设备的任一个智能语音设备,数量可以为一个或多个)采集的用户的语音信号。

[0114] 在步骤102中,根据语音信号对空间进行感知处理,以确定空间中包括的智能语音设备以及与用户之间的位置关系。

[0115] 当第一智能语音设备接收到用户的语音信号后、即第一智能语音设备采集的用户的语音信号以及第二智能语音设备发送的用户的语音信号,根据语音信号对空间进行感知处理,确定空间中包括的智能语音设备以及与用户之间的位置关系、即第一智能语音设备、第二智能语音设备分别与用户之间的位置关系。

[0116] 参见图3B,图3B是本发明实施例提供的一个可选的流程示意图,在一些实施例中,图3B示出图3A中的步骤102可以通过图3B示出的步骤1021至步骤1023实现。

[0117] 根据语音信号对空间进行感知处理,以确定空间中包括的智能语音设备以及与用户之间的位置关系,包括:对于空间中的任一智能语音设备所接收的用户的语音信号,执行

以下处理：

[0118] 在步骤1021中,对智能语音设备从多个方向接收的用户的语音信号进行解析处理,得到智能语音设备从多个方向接收用户的语音信号的能量值。

[0119] 在步骤1022中,将最大能量值对应的方向确定为用户相对于智能语音设备的方向。

[0120] 在步骤1023中,根据语音信号的能量值随距离衰减的关系,以及最大能量值相对于用户的基准语音信号的能量值的衰减值,确定衰减值所对应的距离为智能语音设备与用户之间的距离。

[0121] 在第一智能语音设备接收到其他智能语音设备发来的用户的语音信号(语音信号携带设备标识,用于唯一标识设备)后,会对任一智能语音设备(第一智能语音设备以及第二智能语音设备)所接收的用户的语音信号都进行方向以及距离的识别处理。智能语音设备可以设备多方向的麦克风阵列,用于从多个方向接收用户的语音设备,从而可以对智能语音设备从多个方向接收的用户的语音信号进行解析处理,得到智能语音设备从多个方向接收用户的语音信号的能量值,其中,不同方向的语音信号的能量值不同,智能语音设备越靠近用户方向的语音信号的能量值越大,因此,将语音信号最大能量值对应的方向确定为用户相对于智能语音设备的方向。在确定用户相对于智能语音设备的方向后,可以根据语音信号的能量值随距离衰减的关系,以及最大能量值相对于用户的基准语音信号的能量值的衰减值,确定衰减值所对应的距离为智能语音设备与用户之间的距离,其中,用户的基准语音信号的能量值为固定值,或者为其他精确语音检测设备实时检测到的用户的语音信号的标准能量值。

[0122] 在一些实施例中,根据语音信号对空间进行感知处理,以确定空间中包括的智能语音设备以及与用户之间的位置关系,包括:对于空间中的任一智能语音设备所接收的用户的语音信号,执行以下处理:对语音信号进行解析处理,得到智能语音设备与用户之间的第一距离、以及用户相对于智能语音设备的第一方向;响应于接收到的语音信号,对空间进行障碍物探测,得到智能语音设备与用户之间的第二距离;对空间进行障碍物识别,得到用户相对于智能语音设备的第二方向;当第一距离与第二距离的距离差值大于距离误差阈值,和/或,第一方向与第二方向之间的方向误差大于方向误差阈值时,将第一距离与第二距离的加权值,确定为智能语音设备与用户之间的距离,并将第一方向与第二方向的平均值确定为用户相对于智能语音设备的方向。

[0123] 在第一智能语音设备接收到其他智能语音设备发来的用户的语音信号(语音信号携带设备标识,用于唯一标识设备)后,会对任一智能语音设备(第一智能语音设备以及第二智能语音设备)所接收的用户的语音信号都进行方向以及距离的识别处理。由于仅通过语音信号确定智能语音设备与用户之间的距离、以及用户相对于智能语音设备的方向可能不太准确,因此,还可以通过其他方式确定智能语音设备与用户之间的距离、以及用户相对于智能语音设备的方向,并结合两种方法得到的智能语音设备与用户之间的距离、以及用户相对于智能语音设备的方向,最后得到准确的智能语音设备与用户之间的距离、以及用户相对于智能语音设备的方向。

[0124] 即首先对智能语音设备从多个方向接收的用户的语音信号进行解析处理,得到智能语音设备从多个方向接收用户的语音信号的能量值;将最大能量值对应的方向确定为用

户相对于智能语音设备的第一方向,并根据语音信号的能量值随距离衰减的关系,以及最大能量值相对于用户的基准语音信号的能量值的衰减值,确定衰减值所对应的距离为智能语音设备与用户之间的第一距离。然后,响应于接收到的语音信号,可以触发其他设备对空间进行障碍物探测,得到智能语音设备与用户之间的第二距离,其中,其他设备可以是声波探测设备(如超声波探测)、图像采集分析设备(如摄像头采集,识别人的轮廓)、生物信号探测设备(如红外探测)等用于探测智能语音设备与用户之间的距离。例如声波探测设备可以发出声波,并接收障碍物反射回来的声波,根据声波来回的时间,确定智能语音设备与用户之间的距离;图像采集分析设备可以采集当前空间中的障碍物的图像,并根据图像识别方法识别出用户,以及确定出智能语音设备与用户之间的距离;生物信号探测设备可以探测生物信号,例如探测当前空间中的用户,并根据探测出的用户确定出智能语音设备与用户之间的距离。其中,其他设备可以集成在智能语音设备中,也可以是能被智能语音设备感知和使用的独立装置。在对空间进行障碍物探测,得到智能语音设备与用户之间的第二距离的同时,还可以通过其他设备对空间进行障碍物识别,得到用户相对于智能语音设备的第二方向,其中,其他设备可以是声波探测设备(如超声波探测)、图像采集分析设备(如摄像头采集,识别人的轮廓)、生物信号探测设备(如红外探测)等用于探测用户相对于智能语音设备的方向。例如声波探测设备可以发出声波,并接收障碍物反射回来的声波,根据返回声波的方向,确定用户相对于智能语音设备的方向;图像采集分析设备可以采集当前空间中的障碍物的图像,并根据图像识别方法识别出用户,以及确定出用户相对于智能语音设备的方向;生物信号探测设备可以探测生物信号,例如探测当前空间中的用户,并根据探测出的用户确定出用户相对于智能语音设备的方向。

[0125] 在通过语音信号解析确定出智能语音设备与用户之间的第一距离以及用户相对于智能语音设备的第一方向(第一方法)、通过其他设备确定出智能语音设备与用户之间的第二距离以及用户相对于智能语音设备的第二方向(第二方法)后,当确定第一距离与第二距离的距离差值大于距离误差阈值,和/或,第一方向与第二方向之间的方向误差大于方向误差阈值时,将第一距离与第二距离的加权值,确定为智能语音设备与用户之间的距离,并将第一方向与第二方向的平均值确定为用户相对于智能语音设备的方向,从而融合第一方法以及第二方法,提高智能语音设备与用户之间的位置关系的准确性。其中,距离误差阈值、方向误差阈值以及权重为用户预先设置的经验值,例如,第一距离与第二距离的权重可以根据具体情况进行设定的,当侧重于第一方法时,可以对第一距离设置较高的权重、第二距离设置较低的权重,如第一距离设置的权重为0.6、第二距离设置的权重为0.4,则智能语音设备与用户之间的距离=第一距离*0.6+第二距离*0.4。

[0126] 需要说明的是,其他设备可以设置为持续感知,因此,其他设备可以实时响应于接收到的语音信号。其他设备也可以设置为定期开启或关闭,从而可以达到省电的目的。

[0127] 当然,也可以仅采用第二方法确定出智能语音设备与用户之间的距离以及用户相对于智能语音设备的方向,不考虑第一方法确定出智能语音设备与用户之间的距离以及用户相对于智能语音设备的方向,即仅通过响应于接收到的语音信号,对空间进行障碍物探测,得到智能语音设备与用户之间的第二距离;对空间进行障碍物识别,得到用户相对于智能语音设备的第二方向。

[0128] 在步骤103中,当空间中还包括至少一个第二智能语音设备时,根据位置关系在第

一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备。

[0129] 当确定出智能语音设备与用户之间的位置关系(方向以及距离)后,当空间中只有第一智能语音设备时,确定第一智能语音设备为目标智能语音设备,当空间中还包括至少一个第二智能语音设备时,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,从而使得目标智能语音设备响应用户的语音信号,避免同一空间中多个智能语音设备同时响应用户的语音信号。

[0130] 参见图3B,图3B是本发明实施例提供的一个可选的流程示意图,在一些实施例中,图3B示出图3A中的步骤103可以通过图3B示出的步骤1031至步骤1032实现。

[0131] 根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:对于根据空间中任一智能设备接收的语音信号进行感知处理而得到的位置关系,执行以下处理:

[0132] 在步骤1031中,当位置关系保持不变的时间超出时间阈值时,确定用户处于静止状态。

[0133] 在步骤1032中,根据位置关系包括的智能语音设备与用户之间距离,将空间中与用户距离最小的智能语音设备确定为目标智能语音设备。

[0134] 在第一智能语音设备确定出智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系后,可以对任一智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系都进行目标智能语音设备的确定处理。当任一智能语音设备与用户之间的位置关系在预设的时长内保持不变时,可以确定用户处于静止状态,此时可以根据智能语音设备与用户之间距离来确定目标智能语音设备,即将空间中与用户距离最小的智能语音设备确定为目标智能语音设备。

[0135] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:对于根据空间中任一智能设备接收的语音信号进行感知处理而得到的位置关系,执行以下处理:当位置关系发生变化时,确定用户处于运动状态;根据位置关系包括的用户相对于智能语音设备的方向,将方向的变化方向确定为用户相对于智能语音设备的移动方向;根据位置关系中包括用户与智能设备之间的距离,将距离的倒数与移动方向的向量相乘,得到智能语音设备与用户之间的位置关系的匹配度;在第一智能语音设备和至少一个第二智能语音设备中,将匹配度最高的智能语音设备确定为目标智能语音设备;其中,当用户相对于智能语音设备的方向的变化为趋近智能语音设备时,移动方向取值为正,当用户相对于智能语音设备的方向的变化为远离智能语音设备时,移动方向取值为负。

[0136] 在第一智能语音设备确定出智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系后,可以对任一智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系都进行目标智能语音设备的确定处理。一般情况下,用户在发出语音信号的过程中是处于状态的,例如当用户在发出语音信号之前距离第一智能语音设备较近,在发出语音信号之后距离第一智能语音设备较远。因此,为了确定出合适的目标智能语音设备来响应用户的语音信号,可以确定用户是否处于运动状态。当智能语音设备与用户之间的位置关系发生变化时,例如位置关系持续变化时间超出时间阈值,则确

定用户处于运动状态,此时可以根据智能语音设备与用户之间距离以及用户相对于所述智能语音设备的方向来确定目标智能语音设备。位置关系包括智能语音设备与用户之间距离以及用户相对于所述智能语音设备的方向,可以根据位置关系中用户相对于智能语音设备的方向,将方向的变化方向确定为用户相对于智能语音设备的移动方向,例如当用户相对于智能语音设备的方向的变化为趋近智能语音设备时,移动方向取值为正,当用户相对于智能语音设备的方向的变化为远离智能语音设备时,移动方向取值为负,或者将方向的变化方向的夹角确定为用户相对于智能语音设备的移动方向。将距离的倒数与移动方向的向量相乘,得到智能语音设备与用户之间的位置关系的匹配度,则匹配度越高的智能语音设备,越能满足用户需求。因此,在第一智能语音设备和至少一个第二智能语音设备中,将匹配度最高的智能语音设备确定为目标智能语音设备,该目标智能语音设备最能满足用户需求。

[0137] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:确定第一智能语音设备和至少一个第二智能语音设备中处于唤醒状态的智能语音设备;当处于唤醒状态的智能语音设备与用户之间的距离未超出临界距离时,确定处于唤醒状态的智能语音设备为目标智能语音设备;其中,临界距离为用户和智能语音设备能够彼此正确感知对方发出的语音信号时的最大距离。

[0138] 在确定空间中还包括至少一个第二智能语音设备时,第一智能语音设备接收第二智能语音设备发送的语音信号可以携带唤醒状态标识,该唤醒状态标识用于标识第二智能语音设备是否处于唤醒状态,当第一智能语音设备广播语音信号时,在语音信号中携带第一智能语音设备的唤醒状态标识,根据任一智能语音设备的唤醒状态标识可以确定第一智能语音设备和至少一个第二智能语音设备中处于唤醒状态的智能语音设备,当处于唤醒状态的智能语音设备与用户之间的距离未超出临界距离时,说明处于唤醒状态的智能语音设备还能感知用户发出的语音信号,用户也能感知处于唤醒状态的智能语音设备发出的语音信号,为了提高用户体验的连续性,可以确定处于唤醒状态的智能语音设备为目标智能语音设备,即使该处于唤醒状态的智能语音设备并不是距离用户最近的智能语音设备,只要该处于唤醒状态的智能语音设备可以满足用户体验,即可让处于唤醒状态的智能语音设备继续响应用户的语音信号。例如,2个智能语音设备距离用户都比较近,且相对远的智能语音设备处于唤醒状态,则用户感知智能语音设备输出的语音信号不会明显降低,因此,可以继续保持相对远的设备响应用户的语音信号,直至当前使用的智能语音设备的距离超出临界距离时,才唤醒距离最近的智能语音设备。

[0139] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:当第一智能语音设备和至少一个第二智能语音设备中存在正在与用户交互的智能语音设备,且与用户之间的距离未超出临界距离时,确定正在与用户交互的智能语音设备为目标智能语音设备。

[0140] 在确定空间中还包括至少一个第二智能语音设备时,第一智能语音设备接收第二智能语音设备发送的语音信号可以携带交互状态标识,该交互状态标识用于标识第二智能语音设备是否处于与用户交互的状态,当第一智能语音设备广播语音信号时,在语音信号中携带第一智能语音设备的交互状态标识,根据任一智能语音设备的交互状态标识可以确

定第一智能语音设备和至少一个第二智能语音设备中正在与用户交互的智能语音设备,当正在与用户交互的智能语音设备与用户之间的距离未超出临界距离时,说明正在与用户交互的智能语音设备还能感知用户发出的语音信号,用户也能感知正在与用户交互的智能语音设备发出的语音信号,为了提高用户体验的连续性,可以确定正在与用户交互的智能语音设备为目标智能语音设备,即使该正在与用户交互的智能语音设备并不是距离用户最近的智能语音设备,只要该正在与用户交互的智能语音设备可以满足用户体验,即可让正在与用户交互的智能语音设备继续响应用户的语音信号。例如,如2个智能语音设备距离用户都比较近,且相对远的智能语音设备与用户处于交互状态,则继续保持相对远的智能语音设备响应用户的语音信号,直至相对远的智能语音设备超出临界距离,避免切换响应用户语音的延迟时,才唤醒距离最近的智能语音设备。

[0141] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:确定在接收到语音信号之前处于唤醒状态的智能语音设备与用户之间的位置关系的变化趋势;当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备将超出临界距离时,在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0142] 在第一智能语音设备确定出智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系后,可以对任一智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系都进行目标智能语音设备的确定处理。首先确定在接收到语音信号之前处于唤醒状态的智能语音设备与用户之间的位置关系的变化趋势,当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备不会超出临界距离时,则确定处于唤醒状态的智能语音设备为目标智能语音设备;当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备将超出临界距离时,说明处于唤醒状态的智能语音设备可能无法满足用户需求,因此,在第一智能语音设备和至少一个第二智能语音设备中,根据上述匹配度的确定方法确定出智能语音设备与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0143] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:确定在接收到语音信号之前处于唤醒状态的智能语音设备与用户之间的位置关系的变化趋势;当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0144] 在第一智能语音设备确定出智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系后,可以对任一智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系都进行目标智能语音设备的确定处理。首先确定在接收到语音信号之前处于唤醒状态的智能语音设备与用户之间的位置关系的变化趋势,当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,则说明处于唤醒状态的智能语音设备将可能不会满足用户需求,因此,在第一智能语音设备和至少一个第二智能语音设备中,根据上述匹配度的确定方法确定出智能语音设备与

用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0145] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:确定在接收到语音信号之前处于唤醒状态的智能语音设备与用户之间的位置关系的变化趋势;当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备不会超出临界距离时,确定处于唤醒状态的智能语音设备为目标智能语音设备。

[0146] 当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备不会在预设时长内超出临界距离或者不会在预设时长内超出临界距离时,说明处于唤醒状态的智能语音设备还能满足用户需求,为了提高用户体验连续性,可以将处于唤醒状态的智能语音设备确定为目标智能语音设备。

[0147] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:获取在接收到语音信号之前处于唤醒状态的智能语音设备的历史数据,并结合已唤醒的智能语音设备的位置关系的变化趋势、使用时长以及唤醒次数,通过人工智能模型预测处于唤醒状态的智能语音设备的预计使用时长;在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备确定为目标智能语音设备。

[0148] 在确定目标智能语音设备之前,需要获取在接收到语音信号之前处于唤醒状态的智能语音设备的历史数据,并结合已唤醒的智能语音设备的位置关系的变化趋势、使用时长以及唤醒次数,通过人工智能模型预测处于唤醒状态的智能语音设备的预计使用时长,其中,已唤醒的智能语音设备的位置关系的变化趋势表征用户距离已唤醒的智能语音设备越近,则已唤醒的智能语音设备的预计使用时长越长;已唤醒的智能语音设备的使用时长越长,则预计使用时长越长;已唤醒的智能语音设备唤醒次数越多,则预计使用时长越长。并在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备确定为目标智能语音设备,以便后续在预计使用时长到达时,触发目标智能语音设备处于唤醒状态并响应用户的语音信号。

[0149] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:基于第一智能语音设备以及至少一个第二智能语音设备所绑定的用户账号,确定账号对应多个智能语音设备时,在多个智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0150] 在第一智能语音设备确定出智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系后,可以对任一智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系都进行目标智能语音设备的确定处理。基于第一智能语音设备以及至少一个第二智能语音设备所绑定的用户账号,确定账号对应多个智能语音设备,则说明需要从多个智能语音设备中确定出目标智能语音设备以响应用户的语音信号。因此,在多个智能语音设备中,可以根据上述智能语音设备与用户之间的位置关系的匹配度的确定方法,确定匹配度最高的智能语音设备为目标智能语音设备。

[0151] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:基于用户的语音信号的声纹特

征,识别用户的语音信号所对应的用户;基于第一智能语音设备以及至少一个第二智能语音设备所绑定的用户账号,确定账号所对应的用户为语音信号所对应的用户时,将账号对应的智能语音设备确定为可唤醒的智能语音设备;在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0152] 在第一智能语音设备确定出智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系后,可以对任一智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系都进行目标智能语音设备的确定处理。基于用户的语音信号的声纹特征,识别用户的语音信号所对应的用户,当基于第一智能语音设备以及至少一个第二智能语音设备所绑定的用户账号,确定账号所对应的用户为语音信号所对应的用户时,可以将账号对应的智能语音设备确定为可唤醒的智能语音设备,则说明需要从多个可唤醒的智能语音设备中确定出目标智能语音设备以响应用户的语音信号。因此,在多个可唤醒的智能语音设备中,可以根据上述智能语音设备与用户之间的位置关系的匹配度的确定方法,确定匹配度最高的智能语音设备为目标智能语音设备。

[0153] 在步骤104中,触发目标智能语音设备处于唤醒状态以响应用户的语音信号。

[0154] 在确定目标智能语音设备后,第一智能语音设备可以触发目标智能语音设备处于唤醒状态,并响应用户的语音信号,其中,在触发目标智能语音设备处于唤醒状态之前,目标智能语音设备也可能处于唤醒状态。第一智能语音设备可以通过局域网广播或用其他近距离通信方式触发目标智能语音设备处于唤醒状态以响应用户的语音信号。

[0155] 在一些实施例中,在确定在接收到语音信号之前处于唤醒状态的智能语音设备与用户之间的位置关系的变化趋势,当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备将超出临界距离时,在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备时,触发目标智能语音设备处于唤醒状态,包括:当确定与用户的位置关系匹配度最高的智能语音设备为目标智能语音设备时,触发处于唤醒状态的智能语音设备处于待机状态,实时唤醒目标智能语音设备。

[0156] 在确定处于唤醒状态的智能语音设备将超出临界距离时,可以触发处于唤醒状态的智能语音设备处于待机状态,并实时唤醒目标智能语音设备,避免处于唤醒状态的智能语音设备仍处于唤醒状态,可以达到省电的效果,还可以避免多个智能语音设备都处于唤醒状态,降低用户体验感。

[0157] 在一些实施例中,在确定在接收到所述语音信号之前处于唤醒状态的智能语音设备与用户之间的位置关系的变化趋势,当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备时,触发目标智能语音设备处于唤醒状态,包括:在处于唤醒状态的智能语音设备未超出临界距离之前,预先唤醒目标智能语音设备。

[0158] 在确定处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,可以在处于唤醒状态的智能语音设备未超出临界距离之前,预先唤醒目标智能语音设备,实现智能语音设备的无缝衔接,避免在处于唤醒状态的智能语音设备未超出临界距离时,再来唤醒

目标智能语音设备,浪费唤醒的时间。

[0159] 在一些实施例中,在获取在接收到语音信号之前处于唤醒状态的智能语音设备的历史数据,并结合已唤醒的智能语音设备的位置关系的变化趋势、使用时长以及唤醒次数,通过人工智能模型预测处于唤醒状态的智能语音设备的预计使用时长,在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备确定为目标智能语音设备时,触发目标智能语音设备处于唤醒状态,包括:当预计使用时长达到时,实时唤醒目标智能语音设备;或者,在预计使用时长到达之前,预先唤醒目标智能语音设备。

[0160] 在通过人工智能模型预测处于唤醒状态的智能语音设备的预计使用时长后,可以在预计使用时长达到时,实时唤醒目标智能语音设备,达到省电的效果;或者,在预计使用时长到达之前,预先唤醒目标智能语音设备,实现智能语音设备的无缝衔接,避免在预计使用时长到达时,再来唤醒目标智能语音设备,浪费唤醒的时间。

[0161] 在一些实施例中,触发目标智能语音设备处于唤醒状态之后,方法还包括:触发目标智能语音设备之外的处于唤醒状态的智能语音设备实时切换到待机状态;或者,等待预设时间段,并在等待预设时间段内,针对目标智能语音设备之外的处于唤醒状态的智能语音设备,确定与用户之间的位置关系的变化趋势,当确定处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,触发处于唤醒状态的智能语音设备切换到待机状态。

[0162] 为了避免同一空间下,有多个智能语音设备处于唤醒状态,可以触发目标智能语音设备之外的处于唤醒状态的智能语音设备实时切换到处于待机状态。或者,为了避免智能语音设备来回切换唤醒状态,可以触发目标智能语音设备处于唤醒状态之后,等待预设时间段,并在等待预设时间段内,针对目标智能语音设备之外的处于唤醒状态的智能语音设备,确定与用户之间的位置关系的变化趋势,当确定处于唤醒状态的智能语音设备在预设时长内将超出临界距离时,再来触发处于唤醒状态的智能语音设备切换到待机状态。

[0163] 参见图3C,图3C是本发明实施例提供的一个可选的流程示意图,在一些实施例中,图3C示出,在步骤105中,当目标智能语音设备与在接收到语音信号之前处于唤醒状态的智能语音设备不是同一设备,且在处于唤醒状态的智能语音设备最后一次响应用户的语音信号的过程中,处于唤醒状态的智能语音设备与用户之间的距离超出临界距离时,触发目标智能语音设备重新响应最后一次语音信号。

[0164] 为了避免用户没有清楚地获得智能语音设备响应用户的语音信号,可以在目标智能语音设备与在接收到语音信号之前处于唤醒状态的智能语音设备不是同一设备,且处于唤醒状态的智能语音设备最后一次响应用户的语音信号时与用户之间的距离超出临界距离时,则说明处于唤醒状态的智能语音设备最后一次响应用户的语音信号时,用户可能没有感知到处于唤醒状态的智能语音设备最后一次响应用户的语音信号的信息,因此可以触发目标智能语音设备重新响应最后一次语音信号,避免信号遗漏。

[0165] 下面以实现本发明实施例提供的智能语音设备控制方案的执行主体为服务器为例,说明本发明实施例提供的智能语音设备的控制方法。参见图4,图4是本发明实施例提供的智能语音设备的控制方法的流程示意图,结合图4示出的步骤进行说明。

[0166] 在步骤201中,第一智能语音设备以及至少一个第二智能语音设备接收用户的语音信号。

[0167] 当用户发出语音信号后,例如唤醒词“ABAB”,第一智能语音设备以及至少一个第二智能语音设备可以采集用户的语音信号。

[0168] 在步骤202中,第一智能语音设备以及至少一个第二智能语音设备向服务器发送用户的语音信号。

[0169] 在步骤203中,服务器接收第一智能语音设备以及至少一个第二智能语音设备向服务器发送的用户的语音信号。

[0170] 在步骤204中,服务器根据语音信号对空间进行感知处理,以确定空间中包括的智能语音设备以及与用户之间的位置关系。

[0171] 在服务器接收到第一智能语音设备以及至少一个第二智能语音设备的用户的语音信号后,可以根据语音信号对空间进行感知处理,以确定空间中包括的智能语音设备以及与用户之间的位置关系。

[0172] 在一些实施例中,根据语音信号对空间进行感知处理,以确定空间中包括的智能语音设备以及与用户之间的位置关系,包括:对于空间中的任一智能语音设备所接收的用户的语音信号,执行以下处理:

[0173] 对智能语音设备从多个方向接收的用户的语音信号进行解析处理,得到智能语音设备从多个方向接收用户的语音信号的能量值;将最大能量值对应的方向确定为用户相对于智能语音设备的方向,并根据语音信号的能量值随距离衰减的关系,以及最大能量值相对于用户的基准语音信号的能量值的衰减值,确定衰减值所对应的距离为智能语音设备与用户之间的距离。

[0174] 在服务器接收到第一智能语音设备以及第二智能语音设备发来的用户的语音信号(语音信号携带设备标识,用于唯一标识设备)后,会对任一智能语音设备(第一智能语音设备以及第二智能语音设备)所接收的用户的语音信号都进行方向以及距离的识别处理。智能语音设备可以设备多方向的麦克风阵列,用于从多个方向接收用户的语音设备,从而智能语音设备得到从多个方向接收用户的语音信号的能量值,并将从多个方向接收用户的语音信号的能量值发送至服务器,服务器根据多个方向接收用户的语音信号的能量值进行方向以及距离的识别处理。

[0175] 在一些实施例中,根据语音信号对空间进行感知处理,以确定空间中包括的智能语音设备以及与用户之间的位置关系,包括:对于空间中的任一智能语音设备所接收的用户的语音信号,执行以下处理:对语音信号进行解析处理,得到智能语音设备与用户之间的第一距离、以及用户相对于智能语音设备的第一方向;响应于接收到的语音信号,对空间进行障碍物探测,得到智能语音设备与用户之间的第二距离;对空间进行障碍物识别,得到用户相对于智能语音设备的第二方向;当第一距离与第二距离的距离差值大于距离误差阈值,和/或,第一方向与第二方向之间的方向误差大于方向误差阈值时,将第一距离与第二距离的加权值,确定为智能语音设备与用户之间的距离,并将第一方向与第二方向的平均值确定为用户相对于智能语音设备的方向。

[0176] 在步骤205中,当空间中还包括至少一个第二智能语音设备时,服务器根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备。

[0177] 当服务器确定出智能语音设备与用户之间的位置关系(方向以及距离)后,当空间

中只有第一智能语音设备时,确定第一智能语音设备为目标智能语音设备,当空间中还包括至少一个第二智能语音设备时,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,从而使得目标智能语音设备响应用户的语音信号,避免同一空间中多个智能语音设备同时响应用户的语音信号。

[0178] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:对于根据空间中任一智能设备接收的语音信号进行感知处理而得到的位置关系,执行以下处理:当位置关系保持不变的时间超出时间阈值时,确定用户处于静止状态;根据位置关系包括的智能语音设备与用户之间距离,将空间中与用户距离最小的智能语音设备确定为目标智能语音设备。

[0179] 在服务器确定出智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系后,可以对任一智能语音设备(第一智能语音设备以及第二智能语音设备)与用户之间的位置关系都进行目标智能语音设备的确定处理。

[0180] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:对于根据空间中任一智能设备接收的语音信号进行感知处理而得到的位置关系,执行以下处理:当位置关系发生变化时,确定用户处于运动状态;根据位置关系包括的用户相对于智能语音设备的方向,将方向的变化方向确定为用户相对于智能语音设备的移动方向;根据位置关系中包括用户与智能设备之间的距离,将距离的倒数与移动方向的向量相乘,得到智能语音设备与用户之间的位置关系的匹配度;在第一智能语音设备和至少一个第二智能语音设备中,将匹配度最高的智能语音设备确定为目标智能语音设备;其中,当用户相对于智能语音设备的方向的变化为趋近智能语音设备时,移动方向取值为正,当用户相对于智能语音设备的方向的变化为远离智能语音设备时,移动方向取值为负。

[0181] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:确定第一智能语音设备和至少一个第二智能语音设备中处于唤醒状态的智能语音设备;当处于唤醒状态的智能语音设备与用户之间的距离未超出临界距离时,确定处于唤醒状态的智能语音设备为目标智能语音设备;其中,临界距离为用户和智能语音设备能够彼此正确感知对方发出的语音信号时的最大距离。

[0182] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:当第一智能语音设备和至少一个第二智能语音设备中存在正在与用户交互的智能语音设备,且与用户之间的距离未超出临界距离时,确定正在与用户交互的智能语音设备为目标智能语音设备。

[0183] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:确定在接收到语音信号之前处于唤醒状态的智能语音设备与用户之间的位置关系的变化趋势;当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备将超出临界距离时,在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0184] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设

备中,确定满足空间的使用场景的目标智能语音设备,包括:确定在接收到语音信号之前处于唤醒状态的智能语音设备与用户之间的位置关系的变化趋势;当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0185] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:确定在接收到语音信号之前处于唤醒状态的智能语音设备与用户之间的位置关系的变化趋势;当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备不会超出临界距离时,确定处于唤醒状态的智能语音设备为目标智能语音设备。

[0186] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:获取在接收到语音信号之前处于唤醒状态的智能语音设备的历史数据,并结合已唤醒的智能语音设备的位置关系的变化趋势、使用时长以及唤醒次数,通过人工智能模型预测处于唤醒状态的智能语音设备的预计使用时长;在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备确定为目标智能语音设备。

[0187] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:基于第一智能语音设备以及至少一个第二智能语音设备所绑定的用户账号,确定账号对应多个智能语音设备时,在多个智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0188] 在一些实施例中,根据位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,包括:基于用户的语音信号的声纹特征,识别用户的语音信号所对应的用户;基于第一智能语音设备以及至少一个第二智能语音设备所绑定的用户账号,确定账号所对应的用户为语音信号所对应的用户时,将账号对应的智能语音设备确定为可唤醒的智能语音设备;在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0189] 在步骤206中,服务器触发目标智能语音设备处于唤醒状态以响应用户的语音信号。

[0190] 在服务器确定目标智能语音设备后,可以根据目标智能语音设备的地址向目标智能语音设备发送唤醒指令,目标智能语音设备接收唤醒指令,进入唤醒状态,并响应用户的语音信号,从而触发目标智能语音设备处于唤醒状态以响应用户的语音信号。

[0191] 在一些实施例中,在确定在接收到语音信号之前处于唤醒状态的智能语音设备与用户之间的位置关系的变化趋势,当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备将超出临界距离时,在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备时,触发目标智能语音设备处于唤醒状态,包括:当确定与用户的位置关系匹配度最高的智能语音设备为目标智能语音设备时,触发处于唤醒状态的智能语音设备处于待机状态,实时唤醒目标智能

语音设备。

[0192] 在一些实施例中,在确定在接收到所述语音信号之前处于唤醒状态的智能语音设备与用户之间的位置关系的变化趋势,当根据位置关系的变化趋势确定处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备时,触发目标智能语音设备处于唤醒状态,包括:在处于唤醒状态的智能语音设备未超出临界距离之前,预先唤醒目标智能语音设备。

[0193] 在一些实施例中,在获取在接收到语音信号之前处于唤醒状态的智能语音设备的历史数据,并结合已唤醒的智能语音设备的位置关系的变化趋势、使用时长以及唤醒次数,通过人工智能模型预测处于唤醒状态的智能语音设备的预计使用时长,在第一智能语音设备和至少一个第二智能语音设备中,确定与用户之间的位置关系的匹配度最高的智能语音设备确定为目标智能语音设备时,触发目标智能语音设备处于唤醒状态,包括:当预计使用时长达到时,实时唤醒目标智能语音设备;或者,在预计使用时长到达之前,预先唤醒目标智能语音设备。

[0194] 在一些实施例中,触发目标智能语音设备处于唤醒状态之后,方法还包括:触发目标智能语音设备之外的处于唤醒状态的智能语音设备实时切换到待机状态;或者,等待预设时间段,并在等待预设时间段内,针对目标智能语音设备之外的处于唤醒状态的智能语音设备,确定与用户之间的位置关系的变化趋势,当确定处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,触发处于唤醒状态的智能语音设备切换到待机状态。

[0195] 在一些实施例中,触发目标智能语音设备处于唤醒状态之后,方法还包括:当目标智能语音设备与在接收到语音信号之前处于唤醒状态的智能语音设备不是同一设备,且处于唤醒状态的智能语音设备最后一次响应用户的语音信号的过程中,处于唤醒状态的智能语音设备与用户之间的距离超出临界距离时,触发目标智能语音设备重新响应最后一次语音信号。

[0196] 至此已经结合本发明实施例提供的电子设备为智能语音设备和服务器时的示例性应用和实施,说明本发明实施例提供的智能语音设备的控制方法,下面继续说明本发明实施例提供的智能语音设备的控制装置555中各个模块配合实现智能语音设备的控制的方案。

[0197] 接收模块5551,用于接收第一智能语音设备所处空间中用户的语音信号;

[0198] 感知模块5552,用于根据所述语音信号对所述空间进行感知处理,以确定所述空间中包括的智能语音设备以及与所述用户之间的位置关系;

[0199] 处理模块5553,用于当所述空间中还包括至少一个第二智能语音设备时,根据所述位置关系在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定满足所述空间的使用场景的目标智能语音设备,以及

[0200] 触发模块5554,用于触发所述目标智能语音设备处于唤醒状态以响应所述用户的语音信号。

[0201] 上述技术方案中,所述感知模块5552还用于对于所述空间中的任一智能语音设备所接收的所述用户的语音信号,执行以下处理:对所述智能语音设备从多个方向接收的所述用户的语音信号进行解析处理,得到所述智能语音设备从多个方向接收所述用户的语音

信号的能量值;将最大能量值对应的方向确定为所述用户相对于所述智能语音设备的方向,并根据语音信号的能量值随距离衰减的关系,以及所述最大能量值相对于所述用户的基准语音信号的能量值的衰减值,确定所述衰减值所对应的距离为所述智能语音设备与所述用户之间的距离。

[0202] 上述技术方案中,所述感知模块5552还用于对于所述空间中的任一智能语音设备所接收的所述用户的语音信号,执行以下处理:对所述语音信号进行解析处理,得到所述智能语音设备与所述用户之间的第一距离、以及所述用户相对于所述智能语音设备的第一方向;响应于接收到的语音信号,对所述空间进行障碍物探测,得到所述智能语音设备与所述用户之间的第二距离;对所述空间进行障碍物识别,得到所述用户相对于所述智能语音设备的第二方向;当所述第一距离与所述第二距离的距离差值大于距离误差阈值,和/或,所述第一方向与所述第二方向之间的方向误差大于方向误差阈值时,将所述第一距离与所述第二距离的加权值,确定为所述智能语音设备与所述用户之间的距离,并将所述第一方向与所述第二方向的平均值确定为所述用户相对于所述智能语音设备的方向。

[0203] 上述技术方案中,所述处理模块5553还用于对于根据所述空间中任一智能设备接收的语音信号进行感知处理而得到的位置关系,执行以下处理:当所述位置关系保持不变的时间超出时间阈值时,确定所述用户处于静止状态;根据所述位置关系包括的所述智能语音设备与所述用户之间距离,将所述空间中与所述用户距离最小的智能语音设备确定为目标智能语音设备。

[0204] 上述技术方案中,所述处理模块5553还用于对于根据所述空间中任一智能设备接收的语音信号进行感知处理而得到的位置关系,执行以下处理:当所述位置关系发生变化时,确定所述用户处于运动状态;根据所述位置关系包括的所述用户相对于所述智能语音设备的方向,将所述方向的变化方向确定为所述用户相对于所述智能语音设备的移动方向;根据所述位置关系中包括所述用户与所述智能设备之间的距离,将所述距离的倒数与所述移动方向的向量相乘,得到所述智能语音设备与所述用户之间的位置关系的匹配度;在所述第一智能语音设备和所述至少一个第二智能语音设备中,将匹配度最高的智能语音设备确定为目标智能语音设备;其中,当所述用户相对于所述智能语音设备的方向的变化为趋近所述智能语音设备时,所述移动方向取值为正,当所述用户相对于所述智能语音设备的方向的变化为远离所述智能语音设备时,所述移动方向取值为负。

[0205] 上述技术方案中,所述处理模块5553还用于确定所述第一智能语音设备和所述至少一个第二智能语音设备中处于唤醒状态的智能语音设备;当所述处于唤醒状态的智能语音设备与所述用户之间的距离未超出临界距离时,确定所述处于唤醒状态的智能语音设备为目标智能语音设备;其中,所述临界距离为所述用户和所述智能语音设备能够彼此正确感知对方发出的语音信号时的最大距离。

[0206] 上述技术方案中,所述处理模块5553还用于当所述第一智能语音设备和所述至少一个第二智能语音设备中存在正在与所述用户交互的智能语音设备,且与所述用户之间的距离未超出临界距离时,确定正在与所述用户交互的智能语音设备为目标智能语音设备。

[0207] 上述技术方案中,处理模块5553还用于确定在接收到所述语音信号之前处于唤醒状态的智能语音设备与所述用户之间的位置关系的变化趋势;当根据所述位置关系的变化趋势确定所述处于唤醒状态的智能语音设备将超出临界距离时,在所述第一智能语音设备

和所述至少一个第二智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备;

[0208] 所述触发模块5554还用于当确定与用户的位置关系匹配度最高的智能语音设备为目标智能语音设备时,触发所述处于唤醒状态的智能语音设备处于待机状态,实时唤醒所述目标智能语音设备。

[0209] 上述技术方案中,所述处理模块5553还用于确定在接收到所述语音信号之前处于唤醒状态的智能语音设备与所述用户之间的位置关系的变化趋势;当根据所述位置关系的变化趋势确定所述处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备;

[0210] 所述触发模块5554还用于在所述处于唤醒状态的智能语音设备未超出所述临界距离之前,预先唤醒所述目标智能语音设备。

[0211] 上述技术方案中,所述处理模块5553还用于确定在接收到所述语音信号之前处于唤醒状态的智能语音设备与所述用户之间的位置关系的变化趋势;当根据所述位置关系的变化趋势确定所述处于唤醒状态的智能语音设备不会超出临界距离时,确定所述处于唤醒状态的智能语音设备为目标智能语音设备。

[0212] 上述技术方案中,所述处理模块5553还用于获取在接收到所述语音信号之前处于唤醒状态的智能语音设备的历史数据,并结合所述已唤醒的智能语音设备的位置关系的变化趋势、使用时长以及唤醒次数,通过人工智能模型预测处于唤醒状态的智能语音设备的预计使用时长;在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备确定为目标智能语音设备;所述触发模块还用于当所述预计使用时长达到时,实时唤醒所述目标智能语音设备;或者,在所述预计使用时长到达之前,预先唤醒所述目标智能语音设备。

[0213] 上述技术方案中,所述装置还包括:

[0214] 切换模块5555,用于触发所述目标智能语音设备之外的处于唤醒状态的智能语音设备实时切换到待机状态;或者,

[0215] 等待预设时间段,并在所述等待预设时间段内,针对所述目标智能语音设备之外的处于唤醒状态的智能语音设备,确定与所述用户之间的位置关系的变化趋势,当确定所述处于唤醒状态的智能语音设备将在预设时长内超出临界距离时,触发所述处于唤醒状态的智能语音设备切换到待机状态。

[0216] 上述技术方案中,所述智能语音设备的控制装置555还包括:

[0217] 响应模块5556,用于当所述目标智能语音设备与在接收到所述语音信号之前处于唤醒状态的智能语音设备不是同一设备,且在所述处于唤醒状态的智能语音设备最后一次响应所述用户的语音信号的过程中,所述处于唤醒状态的智能语音设备与所述用户之间的距离超出临界距离时,触发所述目标智能语音设备重新响应所述最后一次语音信号。

[0218] 上述技术方案中,处理模块5553还用于基于所述第一智能语音设备以及至少一个第二智能语音设备所绑定的用户账号,确定所述账号对应多个智能语音设备时,在所述多个智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0219] 上述技术方案中,所述处理模块5553还用于基于所述用户的语音信号的声纹特征,识别所述用户的语音信号所对应的用户;基于所述第一智能语音设备以及至少一个第二智能语音设备所绑定的用户账号,确定所述账号所对应的用户为所述语音信号所对应的用户时,将所述账号对应的智能语音设备确定为可唤醒的智能语音设备;在所述第一智能语音设备和所述至少一个第二智能语音设备中,确定与所述用户之间的位置关系的匹配度最高的智能语音设备为目标智能语音设备。

[0220] 本发明实施例还提供一种存储有可执行指令的存储介质,其中存储有可执行指令,当可执行指令被处理器执行时,将引起处理器执行本发明实施例提供的智能语音设备的控制方法,例如,如图3A-3C示出的智能语音设备的控制方法,或如图4示出的智能语音设备的控制方法。

[0221] 在一些实施例中,存储介质可以是FRAM、ROM、PROM、EPROM、EEPROM、闪存、磁表面存储器、光盘、或CD-ROM等存储器;也可以是包括上述存储器之一或任意组合的各种设备。

[0222] 在一些实施例中,可执行指令可以采用程序、软件、软件模块、脚本或代码的形式,按任意形式的编程语言(包括编译或解释语言,或者声明性或过程性语言)来编写,并且其可按任意形式部署,包括被部署为独立的程序或者被部署为模块、组件、子例程或者适合在计算环境中使用的其它单元。

[0223] 作为示例,可执行指令可以但不一定对应于文件系统中的文件,可以可被存储在保存其它程序或数据的文件的一部分,例如,存储在超文本标记语言(HTML,Hyper Text Markup Language)文档中的一个或多个脚本中,存储在专用于所讨论的程序的单个文件中,或者,存储在多个协同文件(例如,存储一个或多个模块、子程序或代码部分的文件)中。

[0224] 作为示例,可执行指令可被部署为在一个计算设备上执行,或者在位于一个地点的多个计算设备上执行,又或者,在分布在多个地点且通过通信网络互连的多个计算设备上执行。

[0225] 下面,将说明本发明实施例在一个实际的应用场景中的示例性应用。

[0226] 自动语音识别(ASR, Automatic Speech Recognition)技术在单个智能语音设备场景下能很好的满足用户需求,但是在多个智能语音设备共存的场景下用户的体验不是很好。

[0227] 随着智能语音设备越来越多,出现了不少在同一场景(同一个家里或同一个房间)下存在有多台智能语音设备的情况,这种情况下,如果用户唤醒智能语音设备,并发起语音请求,则会有多台智能语音设备同时响应并回复用户发起的语音请求,从而极大地降低了用户的体验感。

[0228] 为了解决上述问题,通过本发明实施例提供一种智能语音设备的控制方法(基于空间感知的单设备响应方法(VSSP, Voice Service Spatial Perception)),该方法可以根据智能语音设备从用户侧接收到的语音的能量大小及VS账号体系等综合纬度评判出物理空间上离用户最近的一台设备。因此,在上述情况下,即使用户对着智能语音设备发起语音请求时,也只有离用户最近的智能语音设备会给出响应,离用户较远的其他智能语音设备将不会响应用户的请求,并自动进入待机状态等待下一次唤醒,避免声音混乱,从而能更好地适用于多智能语音设备共存的场景。

[0229] 图5是本发明实施例提供的用户唤醒智能设备的示意图,如图5所示,智能语音设

备1和智能语音设备2处于同一个环境下,其中,智能语音设备1的物理空间相对于智能语音设备2的物理空间距离用户更近一些,当用户说出“ABAB”的唤醒词时,智能语音设备1和智能语音设备2均被唤醒,并等待响应用户的语音请求。

[0230] 图6是本发明实施例提供的智能语音设备的应用场景示意图,如图6所示,当用户发起真正的语音请求时,例如“今天天气怎么样”,智能语音设备1和智能语音设备2均会收到该用户的语音请求。此时,由于智能语音设备1相对于智能语音设备2距离用户更近,因此,只有智能语音设备1(目标智能语音设备)会回复用户的语音请求,并播报回复语,例如“深圳今天晴,气温.....”,而智能语音设备2不会回复用户的语音请求,并从被唤醒的状态切换到待机状态,等待下一次用户的唤醒。

[0231] 图7是本发明实施例提供的智能语音设备与云端交互的应用场景示意图,如图7所示,在图6的应用场景下,智能语音设备1和智能语音设备2接收到用户的语音请求后,会将对应用户的语音请求(语音数据)发送给云端,云端会收到来自智能语音设备1的请求和来自智能语音设备2的请求,其中,智能语音设备1的请求和智能语音设备2的请求都是相同的“今天天气怎么样”的语音流。此时,云端会判断智能语音设备1和智能语音设备2是否是同一个登录账号,如果智能语音设备1和智能语音设备2的登录账号相同,则说明智能语音设备1和智能语音设备2很大概率是属于同一个用户的2台设备。而且如果云端收到的智能语音设备1的请求和智能语音设备2的请求的时间相近,则智能语音设备1和智能语音设备2处于同一个环境下的概率也非常大。此时,云端会进行VSSP处理,比较智能语音设备1上传的语音数据的能量值和智能语音设备2上传的语音数据的能量值,根据语音数据的能量值判断出智能语音设备1离用户的距离更近,因此,云端会给智能语音设备1(目标智能语音设备)下发播报指令,智能语音设备1根据播报指令播报回复语,例如“深圳今天晴,气温.....”,并继续响应用户的请求。同时,云端会给智能语音设备2下发待机指令,使智能语音设备2进入待机状态,等待下一次用户的唤醒。

[0232] 图8是本发明实施例提供的智能语音设备1上传给云端的语音数据的波形图,图9是本发明实施例提供的智能语音设备1上传给云端的语音数据的频谱图,图10是本发明实施例提供的智能语音设备2上传给云端的语音数据的波形图,图11是本发明实施例提供的智能语音设备2上传给云端的语音数据的频谱图,如图8和图10所示的波形图可知,在图7的应用场景下,当“今天天气怎么样”上传到云端后,云端通过语音识别可以确定图8中的波形能量值比图10的波形能量值大很多;如图9和图11所示的频谱图可知,在图9的应用场景下,当“今天天气怎么样”上传到云端后,云端通过语音识别可以确定在相同时间内图9中的频谱能量值比图11的频谱能量值大很多、即由图9以及图11所示频谱图中方框的地方可知,图9中语音数据的高频区比图11中语音数据的高频区更能活跃。因此,由图8-11可知,智能语音设备1在相同时间收到的语音数据的能量值大于智能语音设备2的语音数据的能量值,则说明智能语音设备1相对于智能语音设备3距离用户的物理空间更近,需要触发VSSP方法,云端控制只让智能语音设备1响应用户的语音请求。

[0233] 图12是本发明实施例提供的智能语音设备与云端交互的另一应用场景示意图,如图12所述,智能语音设备1和智能语音设备2是相同的登录账号,智能语音设备3是不同于智能语音设备1和智能语音设备2的账号,且智能语音设备与用户的距离分别为:智能语音设备3>智能语音设备2>智能语音设备1,此时VSSP方法只会在智能语音设备1和智能语音设备

2之间生效,智能语音设备1(目标智能语音设备)和智能语音设备3将会同时响应用户的请求,智能语音设备2由于VSSP方法会自动进入待机状态,从而可以适用于同一个空间中多智能语音设备共存,但是智能语音设备属于不同用户的场景(比如办公室等公共场所),不会因为智能语音设备同处同一场景,而导致其他距离更远的智能语音设备无法同时使用,达到智能语音设备共享的效果。

[0234] 为了验证本发明实施例所达到的效果,将采用VSSP方法的智能语音设备(叮当智能屏)与现有的智能语音设备(度度在家、天天精灵)进行比较,其中,在同一场景下,分别设置两个叮当智能屏、两个度度在家、两个天天精灵,且登录同一账号,其比较结果,如表1所示:

[0235] 表1

| 产品型号 | 设备1距离用户1米 | 设备2距离用户3米 |
|-------|-----------|-----------|
| 度度在家 | 响应 | 响应 |
| 天天精灵 | 响应 | 响应 |
| 叮当智能屏 | 响应 | 不响应 |

[0237] 由表1可知,在相同的应用场景下,度度在家和天天精灵在2个设备共存的场景下,都会出现同时唤醒,同时响应用户的语音请求,并同时回复,从而导致声音混乱的现象。而开启VSSP方法的叮当智能屏在2个设备共存的场景下,虽然会出现同时唤醒的情况,但是只有最近的一台设备(距离用户1米的设备1)回复用户的语音请求,用户的体验感比较好。

[0238] 当多个智能语音设备在同一个局域网下,在多个智能语音设备(智能语音设备1、智能语音设备2……)收到用户的语音请求时,也可以不借助云端,而智能语音设备1自身在局域网下会发送加密的账号信息以及语音数据的能量值,并接收到其他智能语音设备(智能语音设备2……)发送的加密的账号信息以及语音数据的能量值,并将其他智能语音设备发送的加密的账号信息以及语音数据的能量值与自身的加密的账号信息以及语音数据的能量值在本地进行比对。如果接收的账号信息和自身账号信息一致,且接收到语音数据的能量值比自身语音数据的能量值大,则智能语音设备1将自动进入待机状态。

[0239] 另外,智能语音设备还可以从用户侧接收到的语音信号的其他的纬度评判出物理空间上最符合用户需求的一台设备。例如,1)可以考虑语音信号的位置关系(即方向和距离中至少一种),当用户处于静止状态时,优先考虑距离的维度,并根据距离确定出目标智能语音设备来响应用户的语音信号;当用户处于运动状态时,可以综合考虑移动方向和距离匹配的程度,例如两者匹配程度加权,来确定目标智能语音设备来响应用户的语音信号。2)除了使用语音信号的能量值来表示用户与智能语音设备之间的距离,还可以使用红外传感装置、超声波装置、摄像头等装置感知位置关系,这些装置可以是集成在智能语音设备中的,也可以被智能语音设备感知和使用的独立装置。3)当存在多个唤醒的智能语音设备时,唤醒的智能语音设备可以根据用户的身份信息进行过滤,即唤醒的智能语音设备仅限于绑定到同一用户账号的。4)如果当前用户与智能语音设备有交互(例如用户边走边说),则在保障用户交互的前提下进行确定目标智能语义设备。

[0240] 综上所述,本发明实施例通过位置关系在第一智能语音设备和至少一个第二智能语音设备中,确定满足空间的使用场景的目标智能语音设备,从而避免在同一场景下的智能语音设备都响应用户的语音信号,提高用户的体验感。

[0241] 以上所述,仅为本发明的实施例而已,并非用于限定本发明的保护范围。凡在本发明的精神和范围之内所作的任何修改、等同替换和改进等,均包含在本发明的保护范围之内。

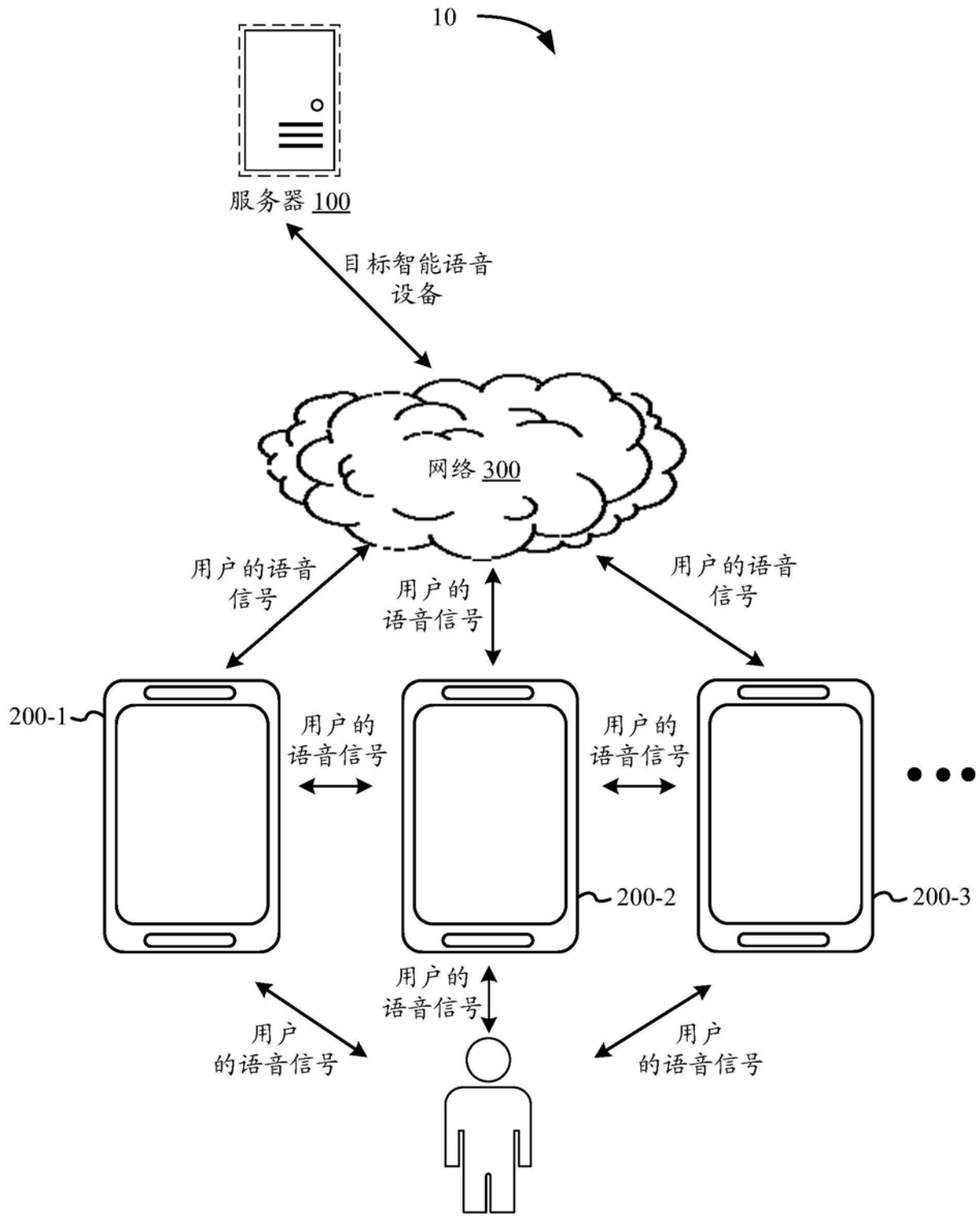


图1

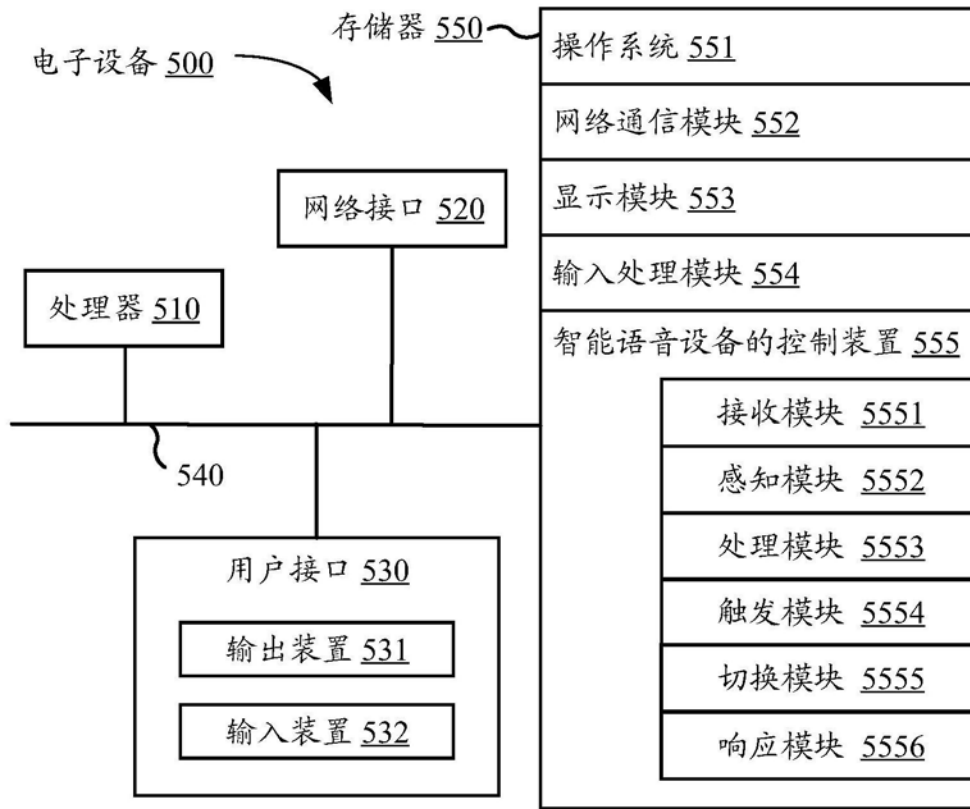


图2

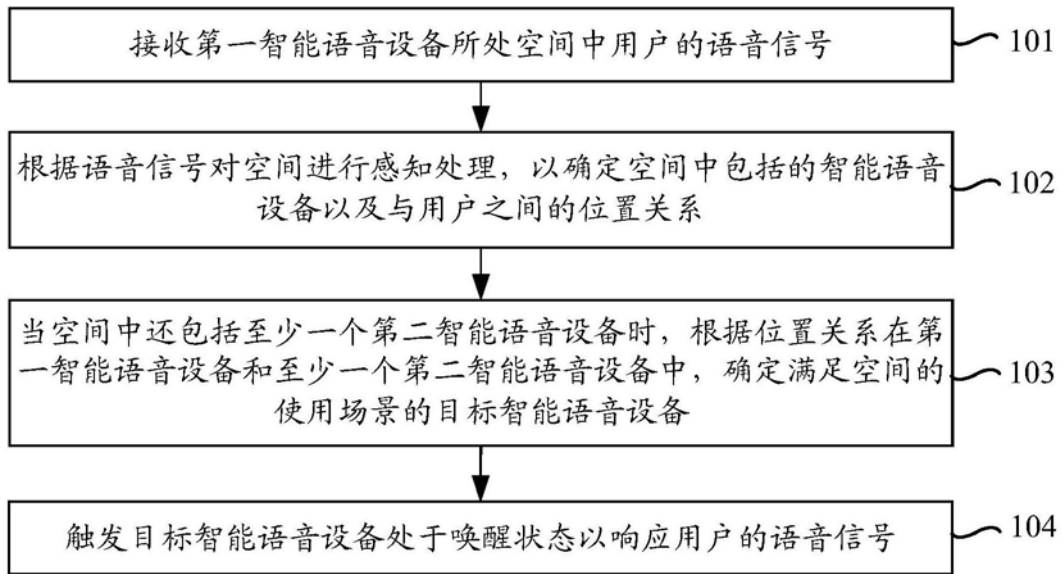


图3A

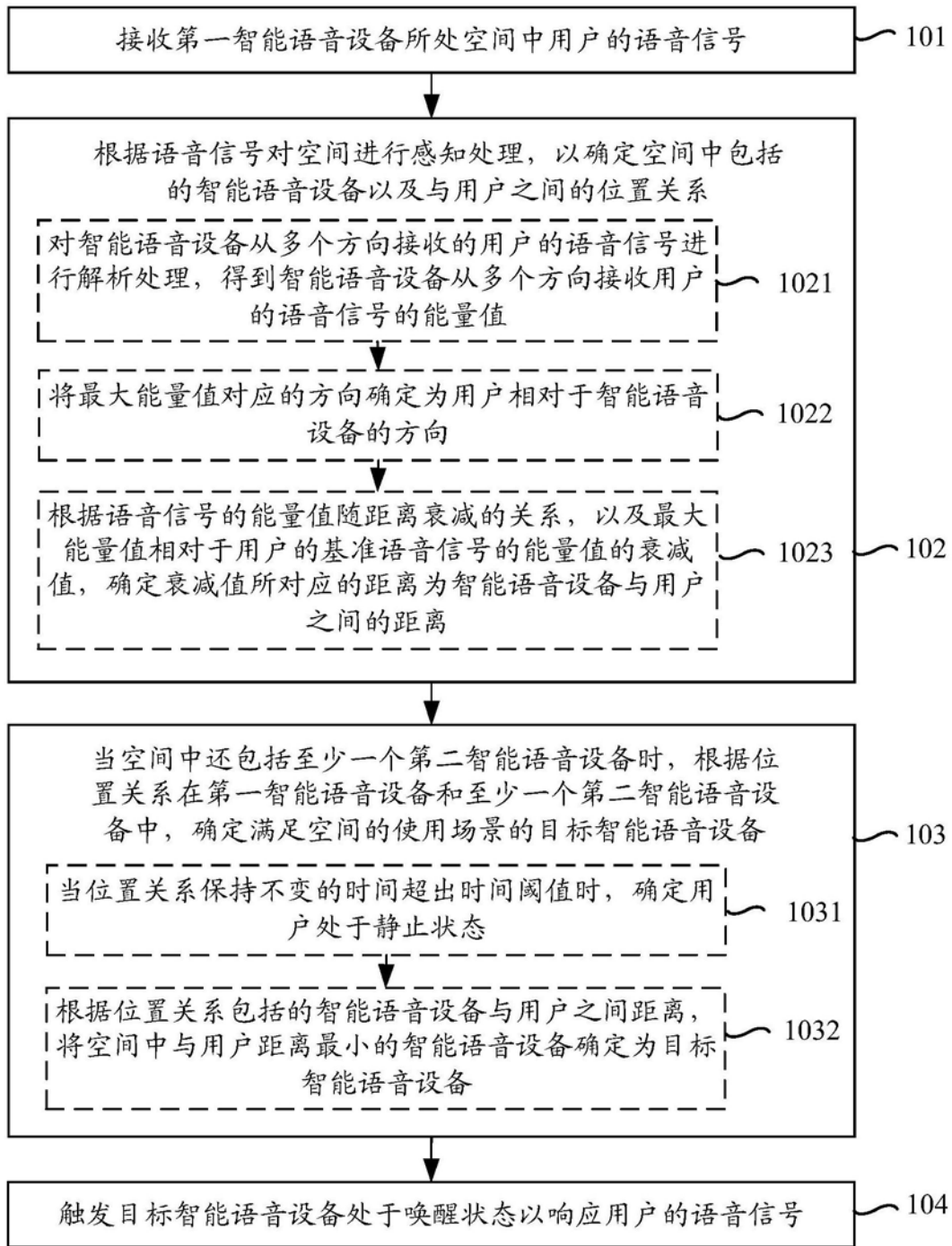


图3B

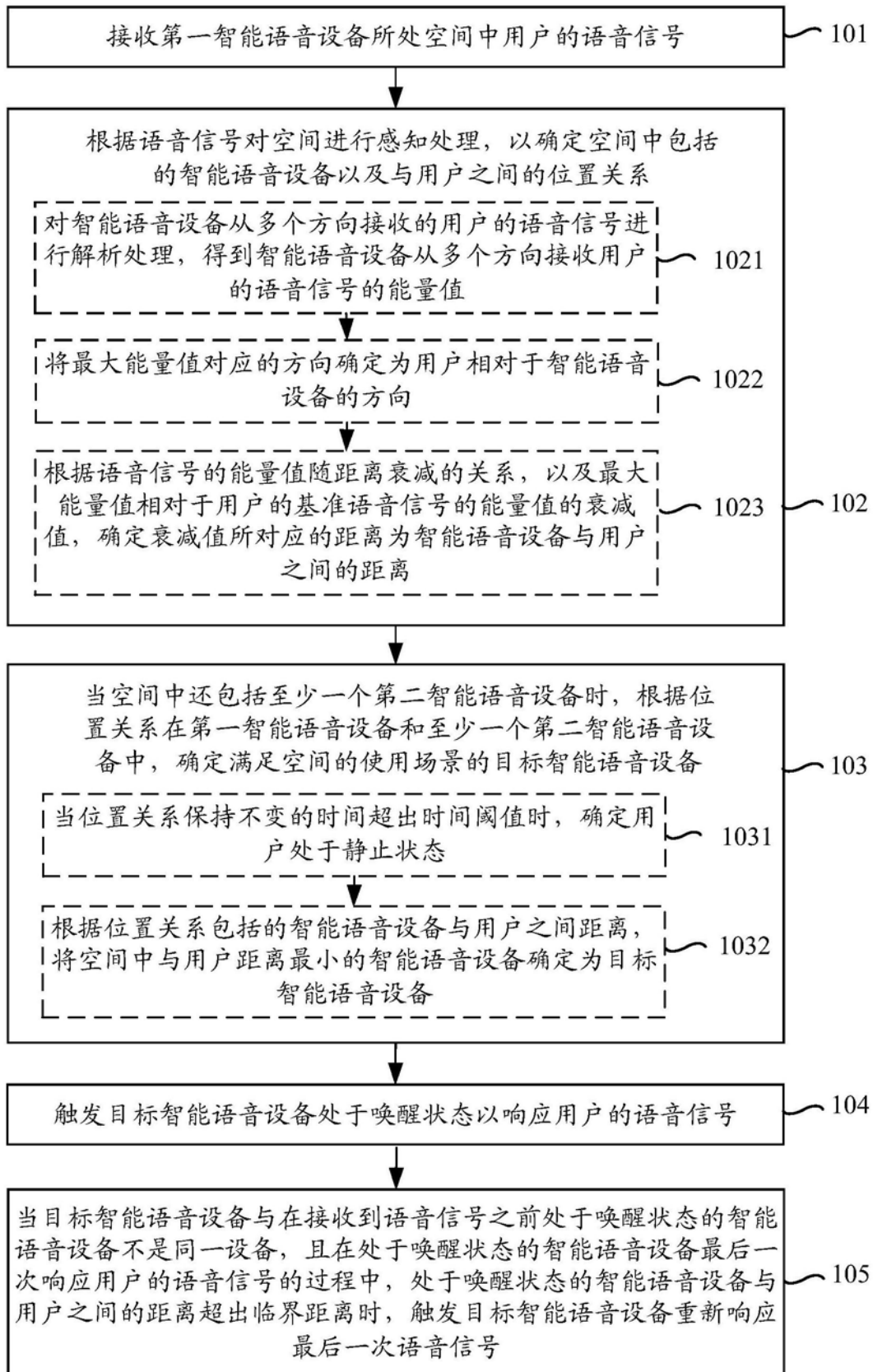


图3C

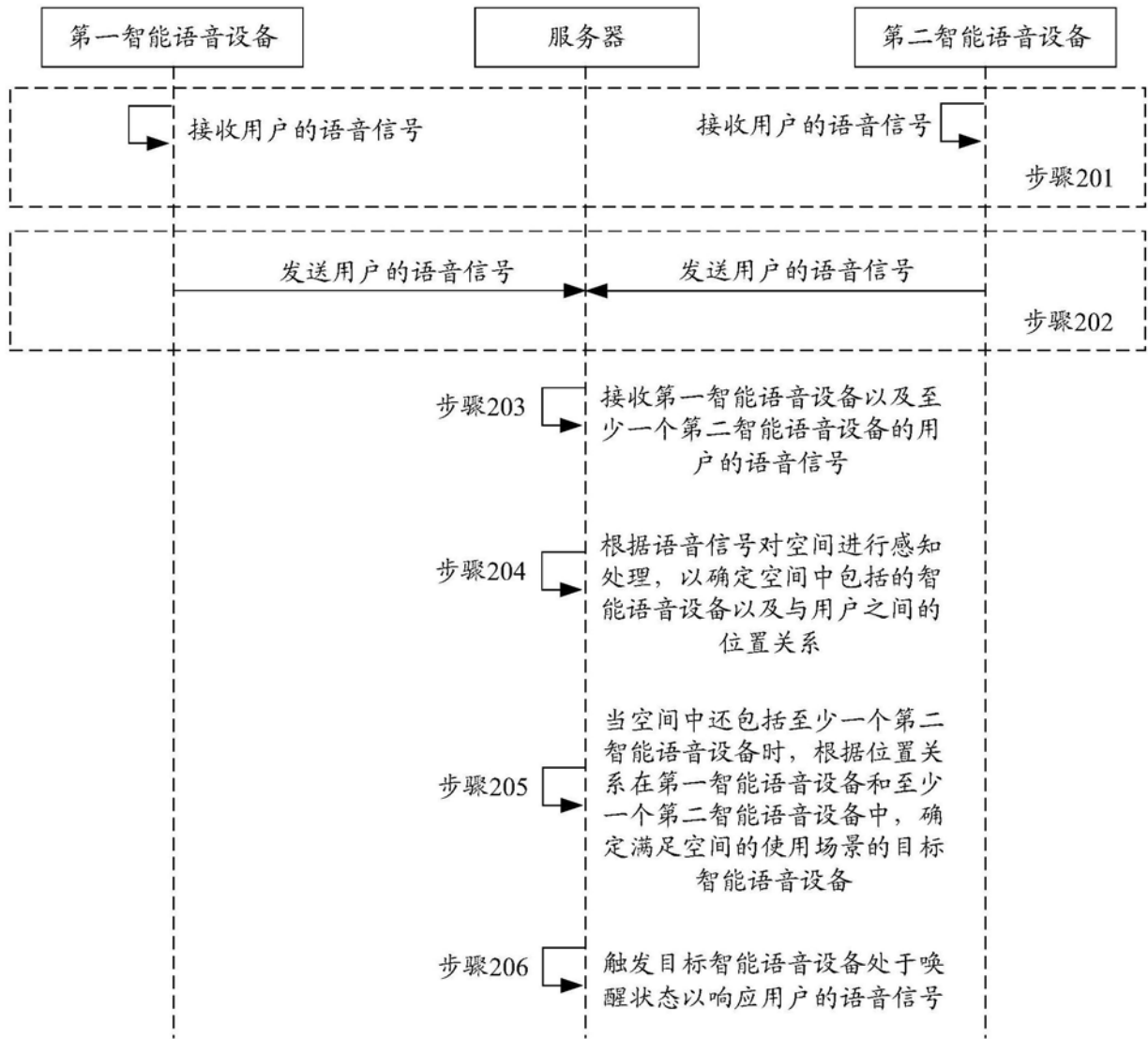


图4

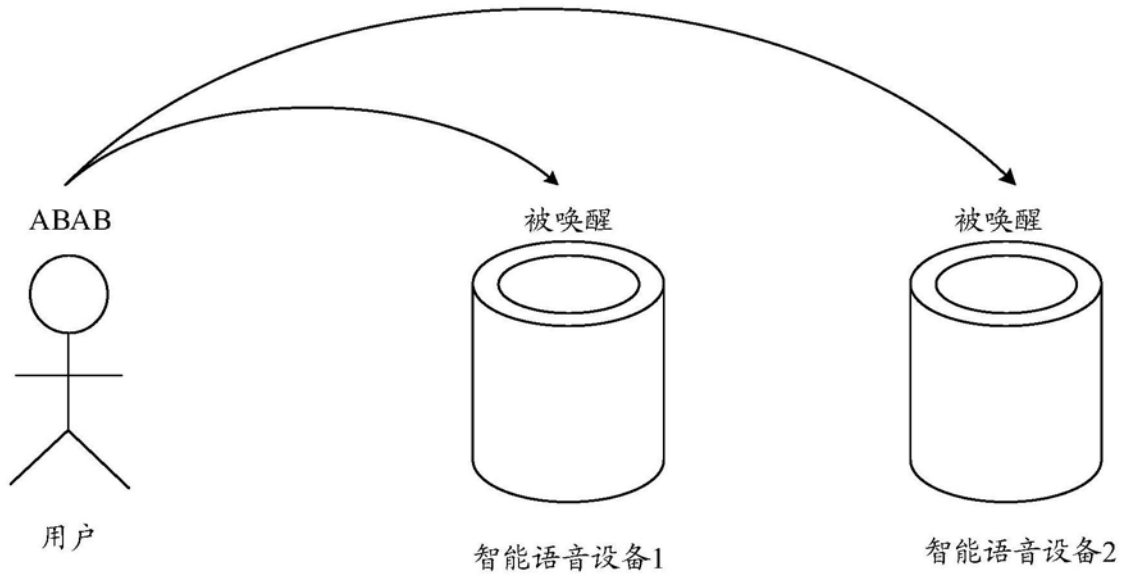


图5

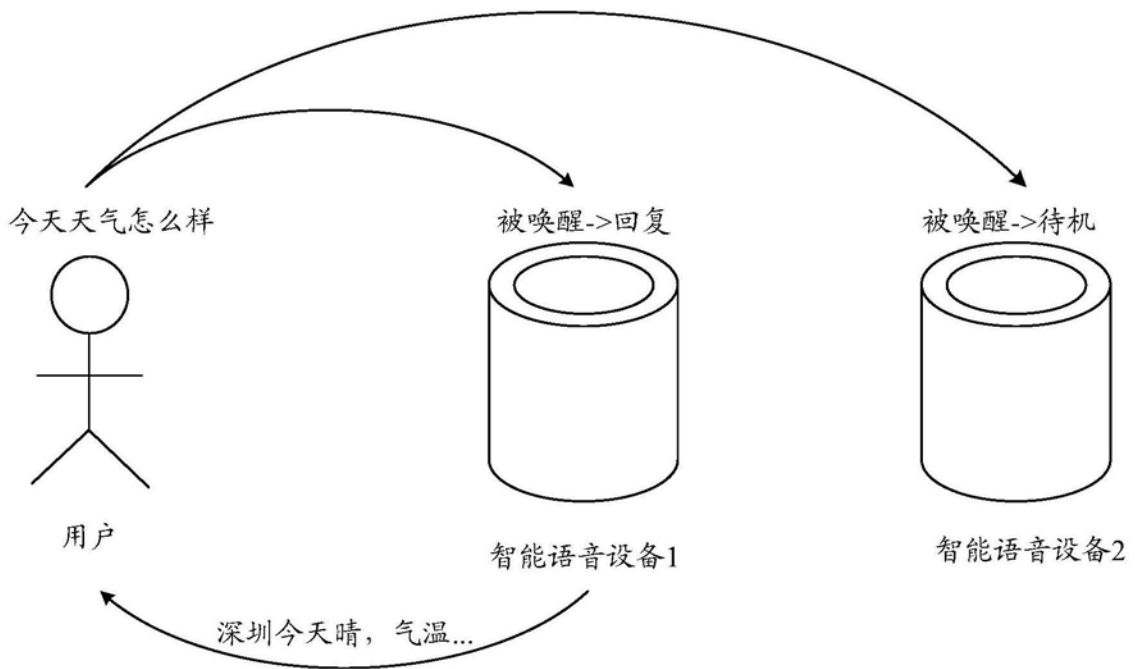


图6

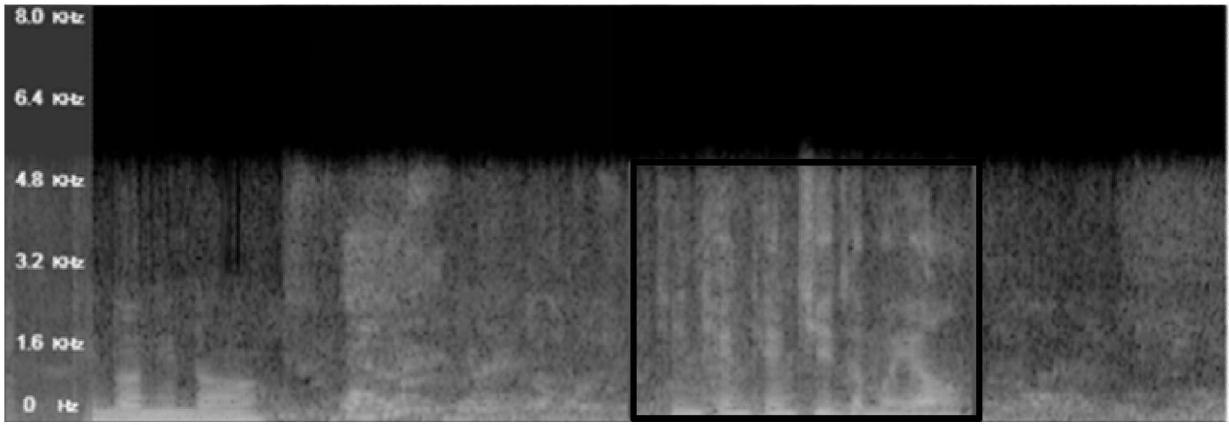


图9

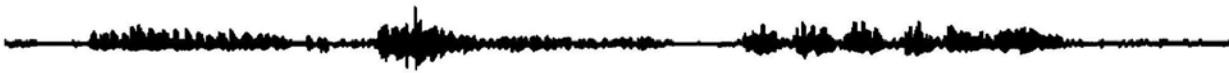


图10

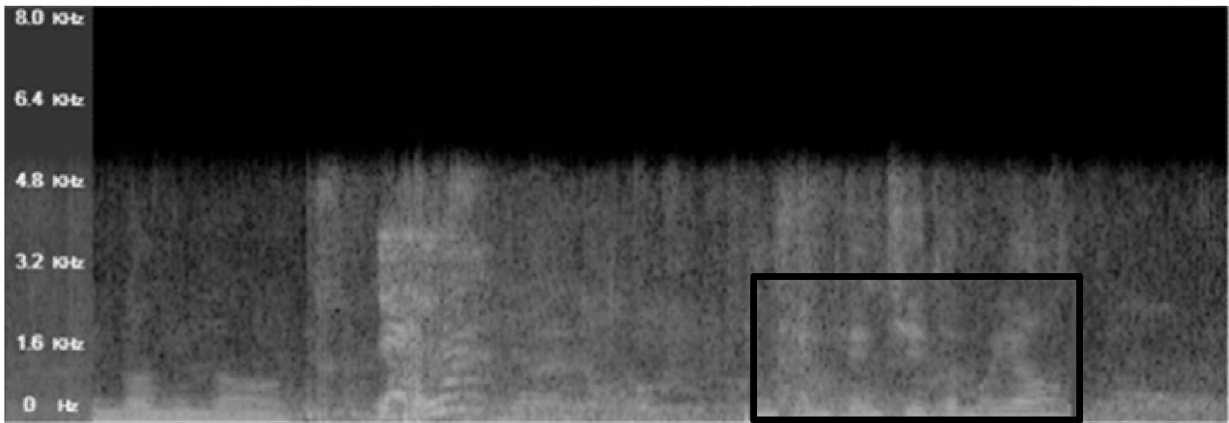


图11

