



(12)发明专利申请

(10)申请公布号 CN 108337362 A

(43)申请公布日 2018.07.27

(21)申请号 201711427997.9

G10L 21/0208(2013.01)

(22)申请日 2017.12.26

(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 徐嘉南 陈果果 钱庆庚

(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 杨贝贝 刘芳

(51) Int. Cl.

H04M 1/725(2006.01)

G10L 15/06(2013.01)

G10L 15/18(2013.01)

G10L 15/22(2006.01)

G10L 15/26(2006.01)

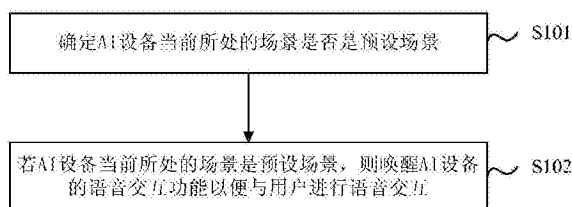
权利要求书4页 说明书14页 附图4页

(54)发明名称

语音交互方法、装置、设备和存储介质

(57)摘要

本申请提供一种语音交互方法、装置、设备和存储介质,该方法应用在AI设备中,确定AI设备当前所处的场景是否是预设场景,若AI设备当前所处的场景是预设场景,则唤醒AI设备的语音交互功能以便与用户进行语音交互。通过场景直接接触发语音交互过程,免去了通过物理唤醒或者唤醒词进行唤醒的过程,简化语音交互的使用流程,降低语音交互的学习成本,提高用户体验。



1. 一种语音交互方法,其特征在于,应用于人工智能AI设备,包括:  
确定AI设备当前所处的场景是否是预设场景;  
若所述AI设备当前所处的场景是所述预设场景,则唤醒所述AI设备的语音交互功能以便与用户进行语音交互。
2. 根据权利要求1所述的方法,其特征在于,所述确定AI设备当前所处的场景是否是预设场景,包括:  
检测所述AI设备的操作状态是否发生变化;  
若发生变化,则确定所述AI设备在操作状态变化后所处的场景是否是预设场景;  
或者,  
接收用户在所述AI设备上输入的场景设置指令;  
根据所述场景设置指令,确定所述AI设备当前所处的场景是否是预设场景;  
或者,  
根据预设周期,周期性检测确定所述AI设备当前所处的场景是否是预设场景;  
或者,  
检测所述AI设备的麦克风是否处于开启状态;  
若所述麦克风处于开启状态,则确定所述AI设备当前所处的场景是否是预设场景。
3. 根据权利要求1或2所述的方法,其特征在于,所述预设场景包括通话场景,则确定AI设备当前所处的场景是否是预设场景,包括:  
检测所述AI设备是否在通话过程中或接收到通话请求;  
若是,则确定所述AI设备当前所处的场景是预设场景;  
或者,  
所述预设场景包括媒体文件播放场景,则所述确定AI设备当前所处的场景是否是预设场景,包括:  
检测所述AI设备是否正在播放媒体文件,所述媒体文件包括图像文件、音频文件以及视频文件中的至少一个;  
若是,则确定所述AI设备当前所处的场景是预设场景;  
或者,  
所述预设场景包括移动场景,则所述确定AI设备当前所处的场景是否是预设场景,包括:  
检测所述AI设备的移动速度,并判断所述移动是否是否大于预设值;  
若是,则确定所述AI设备当前所处的场景是预设场景;  
或者,  
所述预设场景包括信息场景,则所述确定AI设备当前所处的场景是否是预设场景,包括:  
检测所述AI设备是否接收到短消息或者通知消息;  
若是,则确定所述AI设备当前所处的场景是预设场景。
4. 根据权利要求1或2所述的方法,其特征在于,所述唤醒所述AI设备的语音交互功能以便与用户进行语音交互,包括:  
获取用户的语音数据;

根据所述语音数据以及预先设置的与所述AI设备当前所处的场景对应的指令集合进行语音交互。

5. 根据权利要求4所述的方法,其特征在於,所述获取用户的语音数据,包括:

控制所述AI设备的麦克风,采集用户的语音数据;

或者,

控制与所述AI设备连接的蓝牙或者耳机麦克风采集用户的语音获取用户的语音数据;

或者,

接收其他设备发送用户的语音数据。

6. 根据权利要求5所述的方法,其特征在於,所述方法还包括:

采用声学模型和语义理解模型,对所述语音数据进行识别和理解,得到语义理解结果;

当所述语义理解结果的置信度大于预设阈值时,执行所述语义理解结果指示的操作。

7. 根据权利要求6所述的方法,其特征在於,所述采用声学模型和语义理解模型,对所述语音数据进行识别和理解之前,所述方法还包括:

对所述语音数据进行噪声消除和回声消除处理。

8. 根据权利要求6所述的方法,其特征在於,所述采用预先获取的声学模型和语义理解模型,对所述语音数据进行识别和理解,得到语义理解结果,包括:

对所述语音数据采用所述声学模型进行匹配,识别出语义数据;

根据所述语义理解模型对所述语义数据进行理解分析,得到所述语义理解结果。

9. 根据权利要求6至8任一项所述的方法,其特征在於,所述方法还包括:

根据所述AI设备当前所处的场景、所述AI设备当前所处的场景对应的指令集合以及所述AI设备的状态,评估所述语义理解结果的置信度;

判断所述语义理解结果的置信度是否大于预设阈值;

当所述语义理解结果的置信度小于所述预设阈值时,放弃执行所述语义理解结果指示的操作。

10. 根据权利要求6至8任一项所述的方法,其特征在於,所述执行所述语义理解结果指示的操作,包括:

将所述语义理解结果以指定指令的方式输出至软件接口进行执行。

11. 一种语音交互装置,其特征在於,包括:

第一处理模块,用于确定语音交互装置当前所处的场景是否是预设场景;

第二处理模块,用于若所述语音交互装置当前所处的场景是所述预设场景,则唤醒所述语音交互装置的语音交互功能以便与用户进行语音交互。

12. 根据权利要求11所述的装置,其特征在於,所述第一处理模块具体用于:

检测所述语音交互装置的操作状态是否发生变化;

若发生变化,则确定所述语音交互装置在操作状态变化后所处的场景是否是预设场景;

或者,

接收用户在所述语音交互装置上输入的场景设置指令;

根据所述场景设置指令,确定所述语音交互装置当前所处的场景是否是预设场景;

或者,

根据预设周期,周期性检测确定所述语音交互装置当前所处的场景是否是预设场景;  
或者,  
检测所述语音交互装置的麦克风是否处于开启状态;  
若所述麦克风处于开启状态,则确定所述语音交互装置当前所处的场景是否是预设场景。

13. 根据权利要求11或12所述的装置,其特征在于,所述预设场景包括通话场景,则所述第一处理模块还用于:

检测所述语音交互装置是否在通话过程中或接收到通话请求;  
若是,则确定所述语音交互装置当前所处的场景是预设场景;  
或者,  
所述预设场景包括媒体文件播放场景,则所述第一处理模块还用于:  
检测所述语音交互装置是否正在播放媒体文件,所述媒体文件包括图像文件、音频文件以及视频文件中的至少一个;  
若是,则确定所述语音交互装置当前所处的场景是预设场景;  
或者,  
所述预设场景包括移动场景,则所述第一处理模块还用于:  
检测所述语音交互装置的移动速度,并判断所述移动是否是否大于预设值;  
若是,则确定所述语音交互装置当前所处的场景是预设场景;  
或者,  
所述预设场景包括信息场景,则所述第一处理模块还用于:  
检测所述语音交互装置是否接收到短消息或者通知消息;  
若是,则确定所述语音交互装置当前所处的场景是预设场景。

14. 根据权利要求11或12所述的装置,其特征在于,所述第二处理模块具体用于:  
获取用户的语音数据;  
根据所述语音数据以及预先设置的与所述语音交互装置当前所处的场景对应的指令集合进行语音交互。

15. 根据权利要求14所述的装置,其特征在于,所述第二处理模块还具体用于:  
控制所述语音交互装置的麦克风,采集用户的语音数据;  
或者,  
控制与所述语音交互装置连接的蓝牙或者耳机麦克风采集用户的语音获取用户的语音数据;  
或者,  
接收其他设备发送用户的语音数据。

16. 根据权利要求15所述的装置,其特征在于,所述装置还包括:  
第三处理模块,用于采用声学模型和语义理解模型,对所述语音数据进行识别和理解,得到语义理解结果;  
第四处理模块,用于当所述语义理解结果的置信度大于预设阈值时,执行所述语义理解结果指示的操作。

17. 根据权利要求16所述的装置,其特征在于,所述第三处理模块在采用声学模型和语

义理解模型,对所述语音数据进行识别和理解之前,还用于对所述语音数据进行噪声消除和回声消除处理。

18. 根据权利要求16所述的装置,其特征在于,所述第三处理模块具体用于:

对所述语音数据采用所述声学模型进行匹配,识别出语义数据;

根据所述语义理解模型对所述语义数据进行理解分析,得到所述语义理解结果。

19. 根据权利要求16至18任一项所述的装置,其特征在于,所述第四处理模块具体用于:

根据所述语音交互装置当前所处的场景、所述语音交互装置当前所处的场景对应的指令集合以及所述语音交互装置的状态,评估所述语义理解结果的置信度;

判断所述语义理解结果的置信度是否大于预设阈值;

当所述语义理解结果的置信度小于所述预设阈值时,放弃执行所述语义理解结果指示的操作。

20. 根据权利要求16至18任一项所述的装置,其特征在于,所述第四处理模块还用于:

将所述语义理解结果以指定指令的方式输出至软件接口进行执行。

21. 一种人工智能AI设备,其特征在于,包括:存储器和处理器;

所述存储器用于存储计算机指令;所述处理器用于运行所述存储器存储的所述计算机指令实现权利要求1至10任一项所述的语音交互方法。

22. 一种存储介质,其特征在于,包括:可读存储介质和计算机指令,所述计算机指令存储在所述可读存储介质中;所述计算机指令用于实现权利要求1至10任一项所述的语音交互方法。

## 语音交互方法、装置、设备和存储介质

### 技术领域

[0001] 本申请实施例涉及人工智能(Artificial Intelligence, AI)技术领域,尤其涉及一种语音交互方法、装置、设备和存储介质。

### 背景技术

[0002] 随着智能终端设备的不断发展,能够进行语音交互的设备逐渐增多,语音交互在用户日常生活中的应用也逐渐增多,围绕如何提高产品的易用性的产品设计正不断受到关注。

[0003] 目前,常用的语音交互过程包括以下几种:第一种方式,用户点击终端设备上的控制按钮或者home键启动语音交互过程,用户说出期望的操作,终端设备采集用户的语音数据实现与设备进行语音交互。第二种方式,采用固定的唤醒词启动语音交互过程,用户需要先知晓该终端设备的语音交互使用的唤醒词语,用户说出该唤醒词,终端设备根据采集到的固定的唤醒词启动语音交互过程,在启动语音交互过程后采集用户的语音数据进行语音交互,例:用户说出“小度小度”来唤醒手机的语音交互功能。第三种方式,采用oneshot技术实现语音交互,使用唤醒词加期望动作启动语音交互,即用户同时说出固定唤醒词以及期望终端设备执行的内容,终端设备根据唤醒词启动语音交互过程,直接根据采集到的用户期望终端设备执行的内容进行语音交互,例:用户说出“小度小度,今天天气怎么样”与手机进行语音交互。

[0004] 在上述几种语音交互方案中,在进行语音交互时,用户需要采用控制按钮、home键等物理唤醒,或者用户说出固定唤醒词的方式使终端设备启动语音交互功能,且使用唤醒词唤醒存在一定的误差,导致目前的语音交互的使用流程繁琐,成功率较低,致使用户使用语音交互的使用频率比较低。

### 发明内容

[0005] 本申请实施例提供一种语音交互方法、装置、设备和存储介质,用于解决上述几种语音交互方案中,语音交互的使用流程繁琐,成功率较低,致使用户使用语音交互的使用频率比较低的问题。

[0006] 本申请第一方面提供一种语音交互方法,包括:

[0007] 确定AI设备当前所处的场景是否是预设场景;

[0008] 若所述AI设备当前所处的场景是所述预设场景,则唤醒所述AI设备的语音交互功能以便与用户进行语音交互。

[0009] 在一种具体的实现方式中,所述确定AI设备当前所处的场景是否是预设场景,包括:

[0010] 检测所述AI设备的操作状态是否发生变化;

[0011] 若发生变化,则确定所述AI设备在操作状态变化后所处的场景是否是预设场景。

[0012] 在一种具体的实现方式中,所述确定AI设备当前所处的场景是否是预设场景,包

括：

[0013] 接收用户在所述AI设备上输入的场景设置指令；

[0014] 根据所述场景设置指令，确定所述AI设备当前所处的场景是否是预设场景。

[0015] 在一种具体的实现方式中，所述确定AI设备当前所处的场景是否是预设场景，包括：

[0016] 根据预设周期，周期性检测确定所述AI设备当前所处的场景是否是预设场景。

[0017] 在一种具体的实现方式中，所述确定AI设备当前所处的场景是否是预设场景，包括：

[0018] 检测所述AI设备的麦克风是否处于开启状态；

[0019] 若所述麦克风处于开启状态，则确定所述AI设备当前所处的场景是否是预设场景。

[0020] 在一种具体的实现方式中，所述预设场景包括通话场景，则确定AI设备当前所处的场景是否是预设场景，包括：

[0021] 检测所述AI设备是否在通话过程中或接收到通话请求；

[0022] 若是，则确定所述AI设备当前所处的场景是预设场景。

[0023] 在一种具体的实现方式中，所述预设场景包括媒体文件播放场景，则所述确定AI设备当前所处的场景是否是预设场景，包括：

[0024] 检测所述AI设备是否正在播放媒体文件，所述媒体文件包括图像文件、音频文件以及视频文件中的至少一个；

[0025] 若是，则确定所述AI设备当前所处的场景是预设场景。

[0026] 在一种具体的实现方式中，所述预设场景包括移动场景，则所述确定AI设备当前所处的场景是否是预设场景，包括：

[0027] 检测所述AI设备的移动速度，并判断所述移动是否是否大于预设值；

[0028] 若是，则确定所述AI设备当前所处的场景是预设场景。

[0029] 可选的，所述预设场景包括信息场景，则所述确定AI设备当前所处的场景是否是预设场景，包括：

[0030] 检测所述AI设备是否接收到短消息或者通知消息；

[0031] 若是，则确定所述AI设备当前所处的场景是预设场景。

[0032] 可选的，所述唤醒所述AI设备的语音交互功能以便与用户进行语音交互，包括：

[0033] 获取用户的语音数据；

[0034] 根据所述语音数据以及预先设置的与所述AI设备当前所处的场景对应的指令集合进行语音交互。

[0035] 可选的，所述获取用户的语音数据，包括：

[0036] 控制所述AI设备的麦克风，采集用户的语音数据；

[0037] 或者，

[0038] 控制与所述AI设备连接的蓝牙或者耳机麦克风采集用户的语音获取用户的语音数据；

[0039] 或者，

[0040] 接收其他设备发送用户的语音数据。

- [0041] 进一步地,所述方法还包括:
- [0042] 采用声学模型和语义理解模型,对所述语音数据进行识别和理解,得到语义理解结果;
- [0043] 当所述语义理解结果的置信度大于预设阈值时,执行所述语义理解结果指示的操作。
- [0044] 可选的,所述采用声学模型和语义理解模型,对所述语音数据进行识别和理解之前,所述方法还包括:
- [0045] 对所述语音数据进行噪声消除和回声消除处理。
- [0046] 可选的,所述采用预先获取的声学模型和语义理解模型,对所述语音数据进行识别和理解,得到语义理解结果,包括:
- [0047] 对所述语音数据采用所述声学模型进行匹配,识别出语义数据;
- [0048] 根据所述语义理解模型对所述语义数据进行理解分析,得到所述语义理解结果。
- [0049] 可选的,所述方法还包括:
- [0050] 根据所述AI设备当前所处的场景、所述AI设备当前所处的场景对应的指令集合以及所述AI设备的状态,评估所述语义理解结果的置信度;
- [0051] 判断所述语义理解结果的置信度是否大于预设阈值;
- [0052] 当所述语义理解结果的置信度小于所述预设阈值时,放弃执行所述语义理解结果指示的操作。
- [0053] 可选的,所述执行所述语义理解结果指示的操作,包括:
- [0054] 将所述语义理解结果以指定指令的方式输出至软件接口进行执行。
- [0055] 本申请第二方面提供一种语音交互装置,包括:
- [0056] 第一处理模块,用于确定语音交互装置当前所处的场景是否是预设场景;
- [0057] 第二处理模块,用于若所述语音交互装置当前所处的场景是所述预设场景,则唤醒所述语音交互装置的语音交互功能以便与用户进行语音交互。
- [0058] 可选的,所述第一处理模块具体用于:
- [0059] 检测所述语音交互装置的操作状态是否发生变化;
- [0060] 若发生变化,则确定所述语音交互装置在操作状态变化后所处的场景是否是预设场景。
- [0061] 可选的,所述第一处理模块具体用于:
- [0062] 接收用户在所述语音交互装置上输入的场景设置指令;
- [0063] 根据所述场景设置指令,确定所述语音交互装置当前所处的场景是否是预设场景。
- [0064] 可选的,所述第一处理模块具体用于:
- [0065] 根据预设周期,周期性检测确定所述语音交互装置当前所处的场景是否是预设场景。
- [0066] 可选的,所述第一处理模块具体用于:
- [0067] 检测所述语音交互装置的麦克风是否处于开启状态;
- [0068] 若所述麦克风处于开启状态,则确定所述语音交互装置当前所处的场景是否是预设场景。



- [0069] 可选的,所述预设场景包括通话场景,则所述第一处理模块还用于:
- [0070] 检测所述语音交互装置是否在通话过程中或接收到通话请求;
- [0071] 若是,则确定所述语音交互装置当前所处的场景是预设场景。
- [0072] 可选的,所述预设场景包括媒体文件播放场景,则所述第一处理模块还用于:
- [0073] 检测所述语音交互装置是否正在播放媒体文件,所述媒体文件包括图像文件、音频文件以及视频文件中的至少一个;
- [0074] 若是,则确定所述语音交互装置当前所处的场景是预设场景。
- [0075] 可选的,所述预设场景包括移动场景,则所述第一处理模块还用于:
- [0076] 检测所述语音交互装置的移动速度,并判断所述移动是否是否大于预设值;
- [0077] 若是,则确定所述语音交互装置当前所处的场景是预设场景。
- [0078] 可选的,所述预设场景包括信息场景,则所述第一处理模块还用于:
- [0079] 检测所述语音交互装置是否接收到短消息或者通知消息;
- [0080] 若是,则确定所述语音交互装置当前所处的场景是预设场景。
- [0081] 可选的,所述第二处理模块具体用于:
- [0082] 获取用户的语音数据;
- [0083] 根据所述语音数据以及预先设置的与所述语音交互装置当前所处的场景对应的指令集合进行语音交互。
- [0084] 可选的,所述第二处理模块还具体用于:
- [0085] 控制所述语音交互装置的麦克风,采集用户的语音数据;
- [0086] 或者,
- [0087] 控制与所述语音交互装置连接的蓝牙或者耳机麦克风采集用户的语音获取用户的语音数据;
- [0088] 或者,
- [0089] 接收其他设备发送用户的语音数据。
- [0090] 可选的,所述装置还包括:
- [0091] 第三处理模块,用于采用声学模型和语义理解模型,对所述语音数据进行识别和理解,得到语义理解结果;
- [0092] 第四处理模块,用于当所述语义理解结果的置信度大于预设阈值时,执行所述语义理解结果指示的操作。
- [0093] 可选的,所述第三处理模块在采用声学模型和语义理解模型,对所述语音数据进行识别和理解之前,还用于对所述语音数据进行噪声消除和回声消除处理。
- [0094] 可选的,所述第三处理模块具体用于:
- [0095] 对所述语音数据采用所述声学模型进行匹配,识别出语义数据;
- [0096] 根据所述语义理解模型对所述语义数据进行理解分析,得到所述语义理解结果。
- [0097] 可选的,所述第四处理模块具体用于:
- [0098] 根据所述语音交互装置当前所处的场景、所述语音交互装置当前所处的场景对应的指令集合以及所述语音交互装置的状态,评估所述语义理解结果的置信度;
- [0099] 判断所述语义理解结果的置信度是否大于预设阈值;
- [0100] 当所述语义理解结果的置信度小于所述预设阈值时,放弃执行所述语义理解结果

指示的操作。

[0101] 可选的,所述第四处理模块还用于:

[0102] 将所述语义理解结果以指定指令的方式输出至软件接口进行执行。

[0103] 本申请第三方面提供一种AI设备,包括:存储器和处理器;

[0104] 所述存储器用于存储计算机指令;所述处理器用于运行所述存储器存储的所述计算机指令实现第一方面任一实现方式提供的语音交互方法。

[0105] 本申请第四方面提供一种存储介质,包括:可读存储介质和计算机指令,所述计算机指令存储在所述可读存储介质中;所述计算机指令用于实现第一方面任一实现方式提供的语音交互方法。

[0106] 本发明实施例第五方面提供一种程序产品,该程序产品包括计算机指令(即计算机程序),该计算机指令存储在可读存储介质中。AI设备的至少一个处理器可以从可读存储介质读取该计算机指令,至少一个处理器执行该计算机指令使得AI设备实施前述第一方面任一实施方式提供的语音交互方法。

[0107] 本申请实施例提供的语音交互方法、装置、设备和存储介质,确定AI设备当前所处的场景是否是预设场景,若AI设备当前所处的场景是预设场景,则唤醒AI设备的语音交互功能以便与用户进行语音交互。通过AI设备所处场景直接触发语音交互过程,免去了通过物理唤醒或者唤醒词进行唤醒的过程,简化语音交互的使用流程,降低语音交互的学习成本,提高用户体验。同时不需要专门的语音交互唤醒硬件或者唤醒词唤醒过程,降低了技术开发成本。

## 附图说明

[0108] 为了更清楚地说明本申请实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍,显而易见地,下面描述中的附图是本申请的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其它的附图。

[0109] 图1为本申请实施例提供的语音交互方法的构思示意图;

[0110] 图2为本申请实施例提供的语音交互方法实施例一的流程图;

[0111] 图3为本申请实施例提供的语音交互方法中预设场景设置示意图;

[0112] 图4为本申请实施例提供的语音交互方法实施例二的流程图;

[0113] 图5为本申请实施例提供的语音交互方法实施例三的流程图;

[0114] 图6为本申请实施例提供的语音交互装置实施例一的结构示意图;

[0115] 图7为本申请实施例提供的语音交互装置实施例二的结构示意图。

## 具体实施方式

[0116] 为使本申请实施例的目的、技术方案和优点更加清楚,下面将结合本申请实施例中的附图,对本申请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申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其它实施例,都属于本申请保护的范围。

[0117] 随着语音交互应用的场景及设备逐渐增多,围绕如何提高产品的易用性的产品设

计正不断受到关注。目前采用的物理按键唤醒语音交互过程,或者采用用户说出固定的唤醒词唤醒语音交互过程的方案,存在以下几个问题:

[0118] (1)、物理唤醒及唤醒词唤醒都需要不断引导用户,但是实际用户使用频度仍然非常低,导致用户教育成本较大。

[0119] (2)、用户需求满足被唤醒动作隔断,需要经过2到3步,并且存在成功率的影响(唤醒成功且需求识别成功),导致使用流程上较为繁琐。

[0120] (3)除了满足用户需求的投入外,物理按键及唤醒词都需要额外的研发成本和时间成本,导致技术上开发成本较大。

[0121] 基于上述存在的问题,本申请提供一种语音交互方法,通过在特定的场景下直接进行语音交互过程,即场景触发,终端设备在预设场景下,用户可直接说出相关动作指令,终端设备可免去唤醒过程直接对用户说出的相关动作指令进行识别以及执行,实现更便捷的语音交互,简化流程,提高用户体验,同时降低技术开发成本。

[0122] 本申请实施例提供的语音交互方法可应用在所有的人工智能(Artificial Intelligence, AI)设备中,例如:手机、电脑、平板电脑、可穿戴设备、机器人、智能家电、服务器等终端设备,工业生产设备,医疗器械设备,安防设备等,对此本方案不做限制。

[0123] 图1为本申请实施例提供的语音交互方法的构思示意图,如图1所示,该方案中,本方案的主要构思是:预先在预先在AI设备中设置一个或多个预设场景,AI设备在确定自身是否处于其中任一个预设场景时,唤醒语音交互功能,采集用户的语音数据,完成与用户之间的语音交互过程。预设场景的设置可以在AI设备出厂前进行设置,也可以在用户使用过程中,根据自身需求进行手动设置。

[0124] 下面通过几个具体实施例对该语音交互方法进行说明。

[0125] 图2为本申请实施例提供的语音交互方法实施例一的流程图,如图2所示,本实施例提供的语音交互方法具体包括以下步骤:

[0126] S101:确定AI设备当前所处的场景是否是预设场景。

[0127] 在本步骤中,根据在AI设备中设置一个或多个预设场景,AI设备在确定自身是否处于其中任一个预设场景,然后决定是否与用户进行语音交互。如果AI设备未处于预设场景,则不唤醒语音交互功能。

[0128] 该预设场景可在AI设备出厂是进行配置,也可以是在应用过程中,根据用户的操作进行设置,对此本方案不做限制。例如:图3为本申请实施例提供的语音交互方法中预设场景设置示意图,如图3所示,AI设备的设置中可提供用于设置语音交互场景的接口,并提供多个接口以供用户选择,图中,在语音交互场景设置界面中,提供了多个常用的场景,并对应提供了开关操作,用户可根据实际情况选择,在那些场景下自动进行语音交互,例如图中用户选择了移动场景,那么在AI设备处于移动场景下则自动执行语音交互流程。

[0129] 该预设场景可以是AI设备所处的环境,或者AI设备自身的状态的场景,例如:以手机为例,通话场景,包括用户正在通话过程中,或者接收到来电等;媒体文件播放场景,包括手机正在播放视频、播放音频,或者播放图像等;信息场景,包括手机接收到短信息或者接收到应用程序的通知消息等;闹钟场景;拍摄场景,包括手机正在拍照或者录像等;移动场景,包括手机为车载状态,导航状态,或者用户正在跑步或者走路等;充电场景等。

[0130] 在该步骤的具体实现中,AI设备可以根据自身状态变化、用户的操作、周期、或者

AI设备的麦克风的的状态等方式确定是否处于预设场景。具体的确定AI设备是否处于预设场景的方式至少包括以下几种：

[0131] 第一种方式,检测AI设备的操作状态是否发生变化;若发生变化,则确定所述AI设备在操作状态变化后所处的场景是否是预设场景。

[0132] 该方案中, AI设备通过操作状态变化触发确定是否处于预设场景的过程,这里的操作状态即包括AI设备被用户操作导致的变化,也包括AI设备本身的状态发生变化,例如:同样的以手机为例,手机在待机时,接到来电,这个时候手机的状态发生变化,则可确定手机在状态变化后的来电持续,或者接通电话,或者拒接等场景是否是预设场景。又或者,手机接收到短信,或者某个应用程序的通知,这时手机的状态发生变化,则可检测状态变化后的接收到消息是否是预设场景等。又或者,用户将手机进行解锁,手机从待机黑屏状态变成屏幕点亮的可操作状态,这时手机的状态发生变化,则可确定手机在状态变化后是否处于预设场景等。

[0133] 第二种方式,接收用户在AI设备上输入的场景设置指令,根据场景设置指令,确定AI设备当前所处的场景是否是预设场景。

[0134] 该方案中,当用户在AI设备上进行操作,即输入场景设置指令则触发AI设备确定当前所处的场景是否是预设场景。例如,用户打开手机中的地图软件并开启导航功能,则手机进入了导航场景,则可确定该手机的导航场景是否是预设场景。又或者,用户输入电话号码并拨出,手机开始拨打电话,则手机进行通话场景,则可确定该通话场景是否是预设场景等。

[0135] 第三种方式,根据预设周期,周期性检测确定AI设备当前所处的场景是否是预设场景。

[0136] 该方案中,可在出厂前在AI设备中设置确定AI设备所处场景是否是预设场景的预设周期,或者根据用户在使用过程中的操作,设置用于确定AI设备所处场景是否是预设场景的周期。AI设备在运行过程中,按照该预设周期周期性的确定当前所处场景是否为预设场景。

[0137] 第四种方式,检测AI设备的麦克风是否处于开启状态,若麦克风处于开启状态,则确定AI设备当前所处的场景是否是预设场景。

[0138] 该方案中,通过麦克风的的状态决定是不是要确定AI设备的当前所处的场景是不是预设场景,例如:AI设备检测到麦克风为开启状态,则可认为用户有语音交互的需求,因此可触发确定当前所处的场景是不是预设场景。若麦克风是关闭状态,则可认为用户暂时没有语音交互的需求。

[0139] 除了上述几种触发AI设备判断当前所处场景是不是预设场景之外,还可以设置其他的方式或者功能来触发该过程,对此本方案不做限制。

[0140] 在上述几种方式的基础上,对于不同的场景, AI设备在进行当前所处的场景是不是预设场景时,需要进行的检测方式不尽相同,下面通过几个例子进行说明。

[0141] 例如:预设场景包括通话场景,那么在确定AI设备当前所处场景是否是预设场景的过程中, AI设备需要检测是否在通话过程中,或者是否接收到通话请求,或者在通话拨出过程中,如果是在任一个状态下,均可确定该AI设备当前所处的场景是预设场景。

[0142] 预设场景包括媒体文件播放场景,则在确定AI设备当前所处场景是否是预设场景

的过程中, AI设备需要检测是否正在播放图像、音频文件或者视频文件, 例如: 播放相册中的照片, 播放音乐, 播放语音消息, 播报通知消息等, 或者正在播放视频, 在线观看视频等, 如果确定了AI设备正在播放媒体文件, 则可以确定AI设备当前所处的场景是预设场景。

[0143] 预设场景包括移动场景, 则在确定AI设备当前所处场景是否是预设场景的过程中, AI设备需要检测移动速度, 判断AI设备的移动速度是否大于预设值, 例如, AI设备在车载状态下, 用户必然不能通过手动操作去执行一些指令, 这时可设置一定的移动速度的阈值, 在速度大于该阈值时, 确定AI设备处于移动场景, 即当前所处的场景是预设场景。

[0144] 预设场景包括信息场景, 则在确定AI设备当前所处场景是否是预设场景的过程中, 需要检测AI设备是否接收到短消息或者通知消息, 例如接收到别的用户发送的短信, 又或者接收到某个应用软件发送的通知消息, 则可以确定AI设备当前所处的场景是预设场景。

[0145] 预设场景包括闹钟场景, 则在确定AI设备当前所处场景是否是预设场景的过程中, 需要检测闹钟是否正在提醒状态, 例如时间到达用户设置的时间, AI设备播放铃声或者震动, 则确定AI设备当前所处的场景是预设场景。

[0146] 预设场景包括拍摄场景, 则在确定AI设备当前所处场景是否是预设场景的过程中, AI设备需要检测是否正在进行拍照, 或者录像等过程中, 若是, 则确定AI设备当前所处场景是预设场景等。

[0147] 在该方案的具体实现中, 预设场景不限于一个, 可在AI设备中设置触发语音交互过程的一个或者多个场景, 对此本方案不做限制。

[0148] S102: 若AI设备当前所处的场景是预设场景, 则唤醒AI设备的语音交互功能以便与用户进行语音交互。

[0149] 在本步骤中, 按照上述的任一方式确定出AI设备当前所处的场景是预设场景时, 则可直接唤醒AI设备的语音交互功能与用户进行语音交互, 对用户说出的内容进行采集识别, 确定其中的指令并执行。

[0150] 本实施例提供的语音交互方法, 通过预先在AI设备中设置可触发语音交互过程的预设场景, AI设备在使用过程中确定当前所处的场景是不是预设场景, 如果确定出当前所处的场景是预设场景, 则唤醒语音交互过程, 即通过AI设备所处场景直接触发语音交互过程, 免去了通过物理唤醒或者唤醒词进行唤醒的过程, 简化语音交互的使用流程, 降低语音交互的学习成本, 提高用户体验, 同时不需要专门的语音交互唤醒硬件或者唤醒词唤醒过程, 降低了技术开发成本。

[0151] 图4为本申请实施例提供的语音交互方法实施例二的流程图。如图4所示, 在上述实施例一的基础上, S102中的若AI设备当前所处的场景是预设场景, 则唤醒AI设备的语音交互功能以便与用户进行语音交互, 一种具体实现中可包括以下步骤:

[0152] S1021: 若AI设备当前所处的场景是预设场景, 则获取用户的语音数据。

[0153] 在本步骤中, 可知唤醒AI设备的语音交互功能以便与用户进行语音交互首先要做的就是获取用户的语音数据, 即需要采集用户说出的内容, 具体实现中至少可以通过以下几种方式实现:

[0154] 第一种方式, 控制AI设备的麦克风, 采集用户的语音数据。

[0155] 如果AI设备的麦克风是开启状态, 则AI设备可直接通过麦克风采集用户说出的内

容,得到语音数据;如果AI设备的麦克风没有打开,则AI设备控制打开麦克风,然后采集用户的语音数据。

[0156] 第二种方式,控制与AI设备连接的蓝牙或者耳机麦克风采集用户的语音获取用户的语音数据。

[0157] 如果AI设备连接了蓝牙耳机、无线耳机或者有线的耳机,则AI设备可通过耳机的麦克风采集用户说出的内容,得到用户的语音数据。

[0158] 第三种方式,接收其他设备发送用户的语音数据。

[0159] 如果AI设备连接了其他的可穿戴设备,例如智能手表、智能手环等,或者连接了其他的AI设备,也可以通过其他AI设备采集语音数据发送给该AI设备,对此本方案不做限制。

[0160] S1022:根据语音数据以及预先设置的与AI设备当前所处的场景对应的指令集合进行语音交互。

[0161] 在本步骤中,AI设备获取到用户的语音数据之后,根据当前所处的场景,以及该场景对应的指令集合进行语音交互,其含义是,该方案中每个预设场景均设置有对应的指令集合。对语音数据进行分析识别确定用户指示的操作集合时,需要与场景对应的指令集合进行比对,实现精确的语音交互。

[0162] 下面以手机为例,通过几个具体的实例,对不同的场景对应不同的指令集合进行说明。

[0163] 通话场景:在通话场景下,用户可能需要的操作包括挂断电话、接听电话、忽略来电、挂断并发送信息给对方、在通话过程中进行信息的分享等等,因此可在通话场景下设置包括上述至少一个操作对应的指令的集合,例如:除了简单的挂断、接听、忽略等指令外,用户在通话过程中,对方需要获取某个第三方用户的联系方式,由于在通过过程中查看通讯录并不方便,且需要再说对方记录,较为繁琐,则用户可以说出需求“将某某的电话号码发送给某某”,手机在通话场景触发语音交互过程后,根据用户的语音数据,调用通讯录中某某的联系方式以短信或者网络信息等方式发送给正在通话的对方,这时候可在指令集合中设置通讯录调用指令,短消息发送指令等,即在对每个场景创建对应的指令集合,具体的指令类型可根据实际情况进行设置,对此本方案不做限制。

[0164] 媒体文件播放场景:例如,手机正在播放音乐,用户可能需要的操作包括:切换至上一首、切换至下一首、随机播放、播放、暂停、快进、快退、收藏、单曲循环、列表播放、声音大一点、声音小一点、静音等,因此可对播放音乐的场景设置实现上述至少一个操作的指令集合。

[0165] 手机正在播放视频,则用户可能需要的操作包括:调亮屏幕、调暗屏幕、上一集、下一集、播放、暂停、快进、快退、收藏、回放、列表播放、声音大一点、声音小一点、静音等,因此可对播放视频的场景设置实现上述至少一个操作的指令集合。

[0166] 手机正在以幻灯片的模式播放相册中的图像,则用户可能需要的操作包括:停止播放,上一张,下一张,退出播放等,可对播放图像的场景设置实现上述操作的指令集合。

[0167] 闹钟场景:在闹钟场景下,用户可能需要的操作包括关闭闹钟,几分钟后提醒,或者下次提醒等,据此可对闹钟场景设置实现上述操作的指令集合

[0168] 信息场景:在手机收到通知或者短信息时,用户可能需要的操作包括,播报通知、播报信息、存储信息、删除信息,可对该信息场景设置对应该些操作中的至少一个的指令集

合。

[0169] 拍摄场景:以拍照场景为例,在手机启动拍照模式时,用户可能需要的操作包括拍照、茄子、保存图片、不保存、删除照片等操作,因此可对该拍照场景设置对应该些操作中的至少一个的指令集合。

[0170] 导航场景:在手机启动导航模式时,用户可能需要的操作包括:重新选择路线、变更目的地、结束导航、查找目的地最近的停车场等等。同样的可对导航场景设置对应该些操作中的至少一个的指令集合。

[0171] 以上述几个简单场景为例,可知道AI设备在不同场景下进行语音交互的目的不同,要实现的功能和执行的的操作也不同,因此可在该语音交互方法的具体实现中,为每个预设场景设置不同的指令集合,以便在场景触发了语音交互过程之后,能够根据用户的语音数据,准确的判断用户需要执行的操作,并执行。

[0172] 本实施例提供的语音交互方法,AI设备所处场景直接触发语音交互过程,免去了通过物理唤醒或者唤醒词进行唤醒的过程,简化语音交互的使用流程,降低语音交互的学习成本,提高用户体验。同时在不同的场景下,设置不同的指令集合,针对不同场景设置的不同的指令可更精确的实现用户需要的操作,提高用户体验。

[0173] 图5为本申请实施例提供的语音交互方法实施例三流程图。如图5所示,在上述两个实施例的基础上,AI设备当前所处场景为预设场景触发语音交互过程,获取到语音数据之后,根据语音数据以及预先设置的与AI设备当前所处的场景对应的指令集合进行语音交互的具体实现步骤包括:

[0174] S201:采用声学模型和语义理解模型,对语音数据进行识别和理解,得到语义理解结果。

[0175] 在本步骤之前,首先需要根据处理的能力词得到语言数据资料,即进行大量场景及人群的训练数据收集。对收集后的数据进行标注及处理,区分意图类型。再对训练数据进行声学模型训练,形成固定声学模型,(也称为语音识别模型、或者语音识别模型参数)和语音理解模型(也称为语音理解模型参数)。实际应用时,AI设备对语音数据采用所述声学模型进行匹配,识别出语义数据,并根据语义理解模型对所述语义数据进行理解分析,得到语义理解结果。其含义是对语音数据进行声学模型匹配,识别出相应的语音数据,如果能成功识别,再通过对比语义理解模型,分析出用户输入语句的语义,得到当前语句的语义理解结果。

[0176] 在一种可能的实现方式中,可在进行S201之前,对语音数据进行噪声消除和回声消除处理。具体的,如果AI设备自带芯片支持噪声消除及回声消除处理会先进行语音数据优化。即使硬件设备不具备噪声消除及回声消除能力,获取声音数据后,会再次使用开发工具(例如:度秘SDK)中内置的噪声消除算法和回声消除算法对声音进行优化处理。

[0177] S202:根据AI设备当前所处的场景、AI设备当前所处的场景对应的指令集合以及AI设备的状态,评估语义理解结果的置信度。

[0178] 在本步骤中,以AI设备为例,AI设备采用当前所处场景对应的指令集合,以及AI设备当前的状态,对上述采用语义理解模型理解语音数据得到的语义理解结果的置信度进行评估,即评估语义理解结果的可信度。语义理解结果会包含多种维度,如语句意图、动作词、专有名词等。以通话场景为例,比如“忽略电话”,就包含动作:“忽略”以及专业名词“电话”,

整体语句意图为忽略当前设备来电。特别的,当用户表述为“忽略来电”、“忽略接听这个电话”这类泛华性质的表达时,则需要根据AI设备状态和场景对应的指令集合进行相应的意图及的动作识别准确度评估,得到上述语义理解结果的置信度。

[0179] S203:判断语义理解结果的置信度是否大于预设阈值。

[0180] 在本步骤中,可在AI设备中预先设置置信度的阈值,即根据多次模拟测试,确定出语义理解结果的准确度的门限,也就是评估出的置信度的门限,在用户输入了语音数据之后,按照上述方法得到了语义理解结果以及对应的置信度,则需要判断置信度是否大于预设阈值。

[0181] S204:当语义理解结果的置信度大于预设阈值时,执行语义理解结果指示的操作。

[0182] S205:当语义理解结果的置信度小于预设阈值时,放弃执行语义理解结果指示的操作。

[0183] 在上述两个步骤中,即评估结果中的置信度大于设定的阈值,即视为语义理解成功,执行语义理解结果指示的操作即可,具体的执行方式中,可将语义理解结果以指定指令的方式输出至软件接口进行执行,以便AI设备根据该指令调用相应的软件或者硬件进行执行。否则视为不成功,放弃上述得到的语义理解结果,或者还可以重新获取语音数据进行上述过程。

[0184] 在上述方案的具体实现中,不同于用户主动操作按键触发AI设备进入语音采集状态,本申请提供的语音交互方案中,一旦AI设备进入预设场景时,AI设备自动进入语音采集状态,例如开启mic调用(这个操作本身由合作硬件厂商处理,SDK提出约定要求和定义),实时接收用户语音输入。即获取语音数据。同时,该方案中,不同于唤醒按钮或唤醒词,唤醒后仅提供单一指令(如启动语音接收)仅与AI设备某一应用功能配合。本申请提供的语义交互过程,AI设备的语音交互过程与当前所处场景是深度关联的,识别出的场景,以及该场景对应的指令需要与硬件定义相关的操作(如对硬件的屏幕控制、喇叭控制、通信控制等),相应的提供了一整套标准化的硬件调用接口规范,可以帮助硬件厂商快速控制AI设备实现需要的操作。

[0185] 综上所述,本申请各个实施例提供的语音交互方法,免去了唤醒AI设备进行语音采集的动作,让用户用自然语言表达的方式,使用语音交互。能够很好的降低用户语音交互的学习成本,对于提高用户体验和粘性会有较大帮助。同时有助于用户培养语音交互的习惯,提高其他场景和功能的使用频率,简化语音交互的使用流程,降低语音交互的学习成本,提高用户体验。同时不需要专门的语音交互唤醒硬件或者唤醒词唤醒过程,降低了技术开发成本。

[0186] 图6为本申请实施例提供的语音交互装置实施例一的结构示意图,如图6所示,本实施例提供的语音交互装置10包括:

[0187] 第一处理模块11,用于确定语音交互装置当前所处的场景是否是预设场景;

[0188] 第二处理模块12,用于若所述语音交互装置当前所处的场景是所述预设场景,则唤醒所述语音交互装置的语音交互功能以便与用户进行语音交互。

[0189] 本实施例提供的语音交互装置,用于实现前述任一方法实施例中的方法实施例的技术方案,其实现原理和技术效果类似,通过场景直接触发语音交互过程,免去了通过物理唤醒或者唤醒词进行唤醒的过程,简化语音交互的使用流程,降低语音交互的学习成本,提



高用户体验。

[0190] 在该语音交互装置10的一种具体实现中,所述第一处理模块11具体用于:

[0191] 检测所述语音交互装置的操作状态是否发生变化;

[0192] 若发生变化,则确定所述语音交互装置在操作状态变化后所处的场景是否是预设场景。

[0193] 可选的,所述第一处理模块11具体用于:

[0194] 接收用户在所述语音交互装置上输入的场景设置指令;

[0195] 根据所述场景设置指令,确定所述语音交互装置当前所处的场景是否是预设场景。

[0196] 可选的,所述第一处理模块11具体用于:

[0197] 根据预设周期,周期性检测确定所述语音交互装置当前所处的场景是否是预设场景。

[0198] 可选的,所述第一处理模块11具体用于:

[0199] 检测所述语音交互装置的麦克风是否处于开启状态;

[0200] 若所述麦克风处于开启状态,则确定所述语音交互装置当前所处的场景是否是预设场景。

[0201] 可选的,所述预设场景包括通话场景,则所述第一处理模块11还用于:

[0202] 检测所述语音交互装置是否在通话过程中或接收到通话请求;

[0203] 若是,则确定所述语音交互装置当前所处的场景是预设场景。

[0204] 可选的,所述预设场景包括媒体文件播放场景,则所述第一处理模块11还用于:

[0205] 检测所述语音交互装置是否正在播放媒体文件,所述媒体文件包括图像文件、音频文件以及视频文件中的至少一个;

[0206] 若是,则确定所述语音交互装置当前所处的场景是预设场景。

[0207] 可选的,所述预设场景包括移动场景,则所述第一处理模块11还用于:

[0208] 检测所述语音交互装置的移动速度,并判断所述移动是否是否大于预设值;

[0209] 若是,则确定所述语音交互装置当前所处的场景是预设场景。

[0210] 可选的,所述预设场景包括信息场景,则所述第一处理模块11还用于:

[0211] 检测所述语音交互装置是否接收到短消息或者通知消息;

[0212] 若是,则确定所述语音交互装置当前所处的场景是预设场景。

[0213] 可选的,所述第二处理模块12具体用于:

[0214] 获取用户的语音数据;

[0215] 根据所述语音数据以及预先设置的与所述语音交互装置当前所处的场景对应的指令集合进行语音交互。

[0216] 可选的,所述第二处理模块12还具体用于:

[0217] 控制所述语音交互装置的麦克风,采集用户的语音数据;

[0218] 或者,

[0219] 控制与所述语音交互装置连接的蓝牙或者耳机麦克风采集用户的语音获取用户的语音数据;

[0220] 或者,

- [0221] 接收其他设备发送用户的语音数据。
- [0222] 上述实现方式提供的语音交互装置,用于实现前述任一方法实施例的技术方案,其实现原理和技术效果类似,在此不再赘述。
- [0223] 图7为本申请实施例提供的语音交互装置实施例二的结构示意图,如图7所示,该语音交互装置10还包括:
- [0224] 第三处理模块13,用于采用声学模型和语义理解模型,对所述语音数据进行识别和理解,得到语义理解结果;
- [0225] 第四处理模块14,用于当所述语义理解结果的置信度大于预设阈值时,执行所述语义理解结果指示的操作。
- [0226] 在一种具体实现中,所述第三处理模块13在采用声学模型和语义理解模型,对所述语音数据进行识别和理解之前,还用于对所述语音数据进行噪声消除和回声消除处理。
- [0227] 可选的,所述第三处理模块13具体用于:
- [0228] 对所述语音数据采用所述声学模型进行匹配,识别出语义数据;
- [0229] 根据所述语义理解模型对所述语义数据进行理解分析,得到所述语义理解结果。
- [0230] 可选的,所述第四处理模块14具体用于:
- [0231] 根据所述语音交互装置当前所处的场景、所述语音交互装置当前所处的场景对应的指令集合以及所述语音交互装置的状态,评估所述语义理解结果的置信度;
- [0232] 判断所述语义理解结果的置信度是否大于预设阈值;
- [0233] 当所述语义理解结果的置信度小于所述预设阈值时,放弃执行所述语义理解结果指示的操作。
- [0234] 可选的,所述第四处理模块14还用于:
- [0235] 将所述语义理解结果以指定指令的方式输出至软件接口进行执行。
- [0236] 上述实现方式提供的语音交互装置,用于实现前述任一方法实施例的技术方案,其实现原理和技术效果类似,在此不再赘述。
- [0237] 本申请还提供一种AI设备,包括:存储器和处理器;处理器的数量为至少一个。
- [0238] 所述存储器用于存储计算机指令;所述处理器用于运行所述存储器存储的所述计算机指令实现前述任一方法实施例提供的语音交互方法。
- [0239] 本申请还提供一种存储介质,包括:可读存储介质和计算机指令,所述计算机指令存储在所述可读存储介质中;所述计算机指令用于实现前述任一方法实施例提供的语音交互方法。
- [0240] 本申请实施例还提供一种程序产品,该程序产品包括计算机指令(即计算机程序),该计算机指令存储在可读存储介质中。AI设备的至少一个处理器可以从可读存储介质读取该计算机指令,至少一个处理器执行该计算机指令使得AI设备实施前述任一实施例提供的语音交互方法。
- [0241] 在上述AI设备的具体实现中,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。结合本申请实施例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。

[0242] 本领域普通技术人员可以理解：实现上述各方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成。前述的程序可以存储于一计算机可读取存储介质中。该程序在执行时，执行包括上述各方法实施例的步骤；而前述的存储介质包括：只读存储器（英文：read-only memory，缩写：ROM）、RAM、快闪存储器、硬盘、固态硬盘、磁带（英文：magnetic tape）、软盘（英文：floppy disk）、光盘（英文：optical disc）及其任意组合。

[0243] 最后应说明的是：以上各实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述各实施例对本申请进行了详细的说明，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申请各实施例技术方案的范围。

可唤醒语音交互的预设场景{通话场景、移动场景、媒体文件播放场景、信息场景、闹钟场景、导航场景、充电场景……}

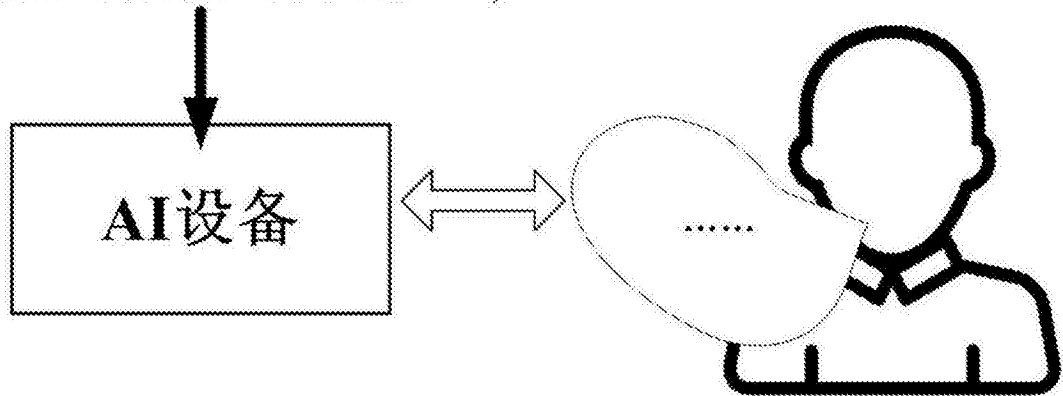


图1

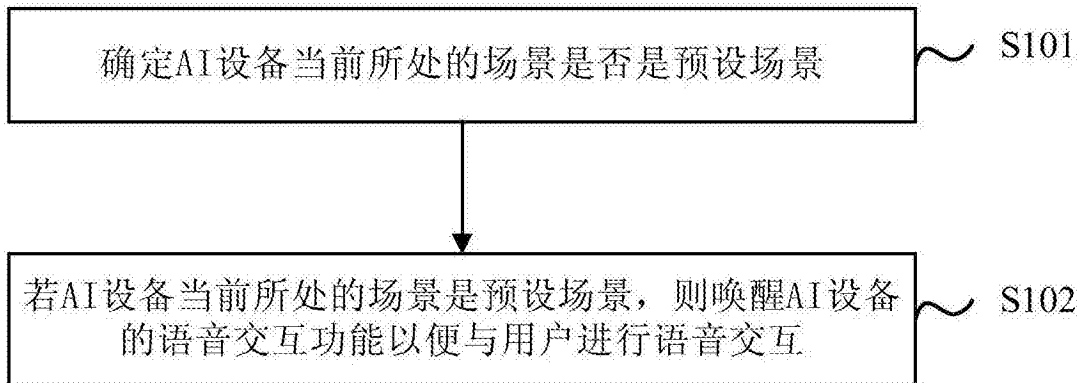


图2

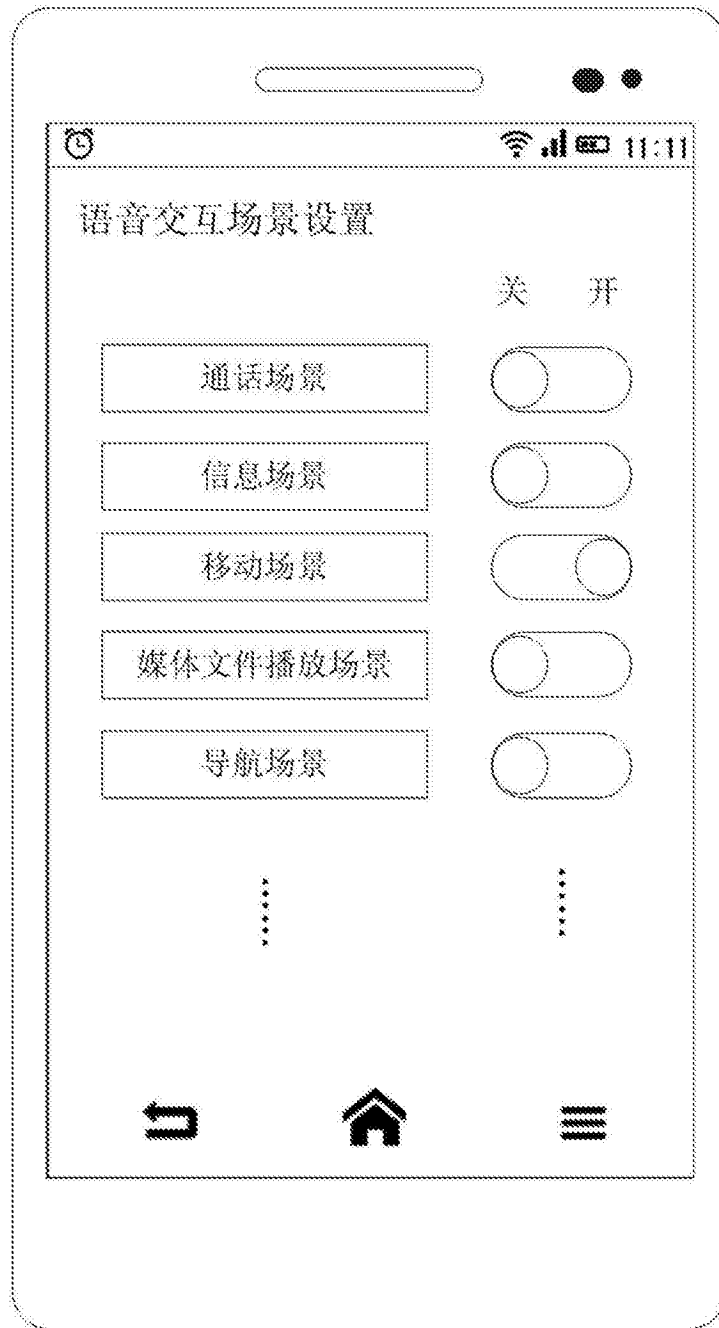


图3

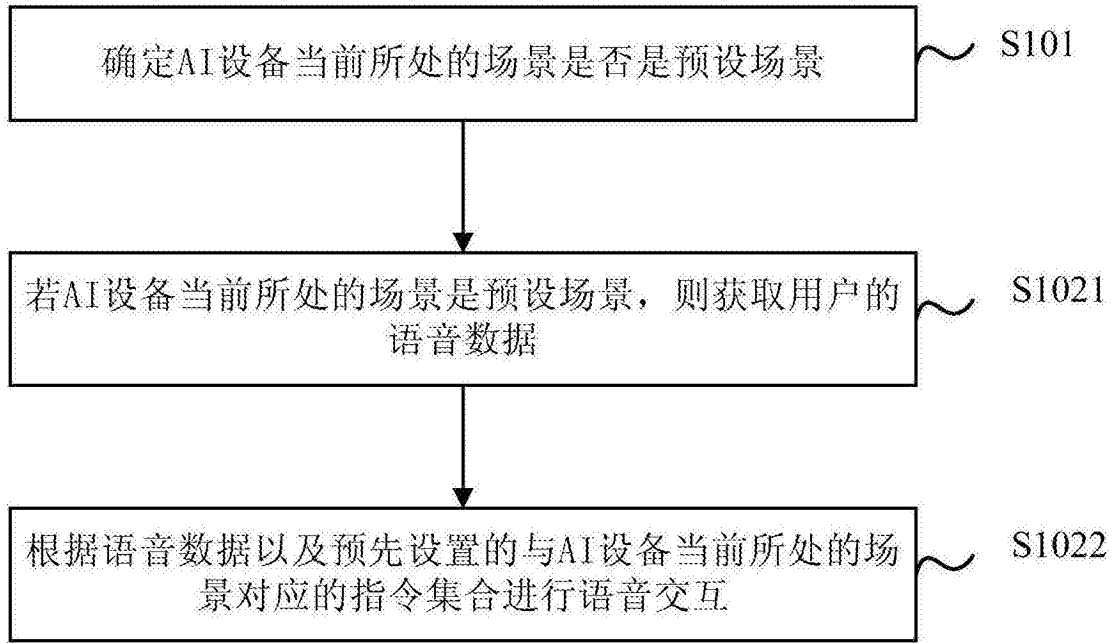


图4

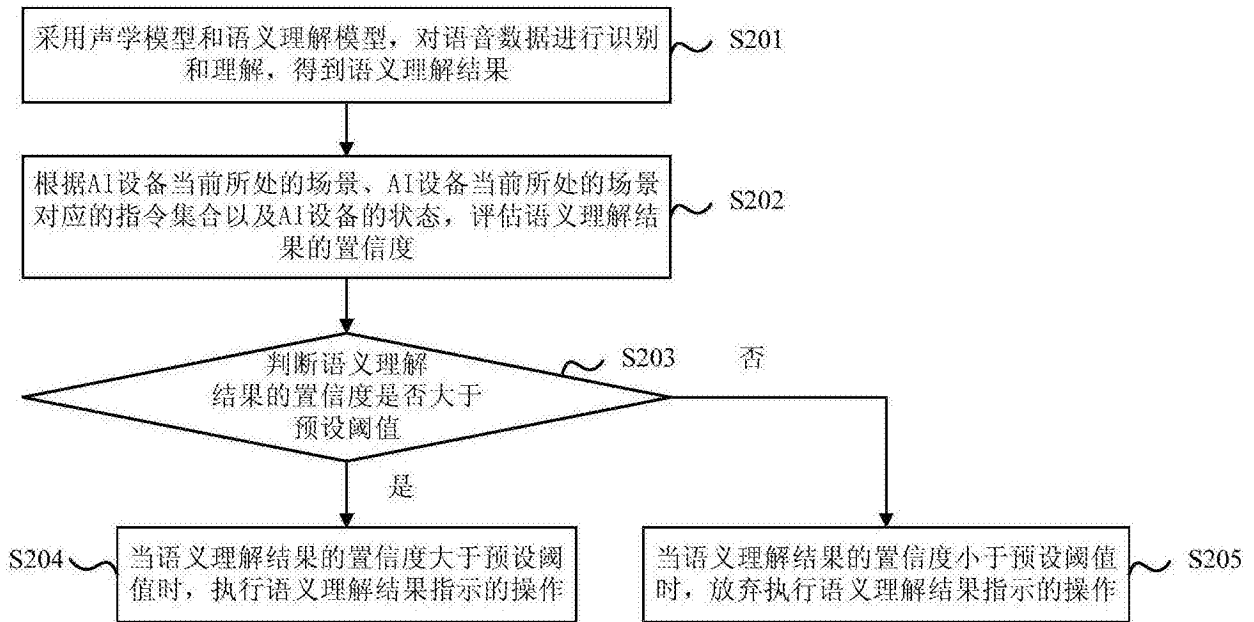


图5

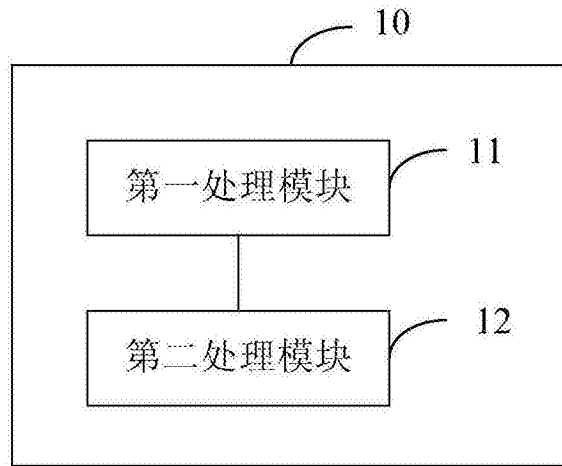


图6

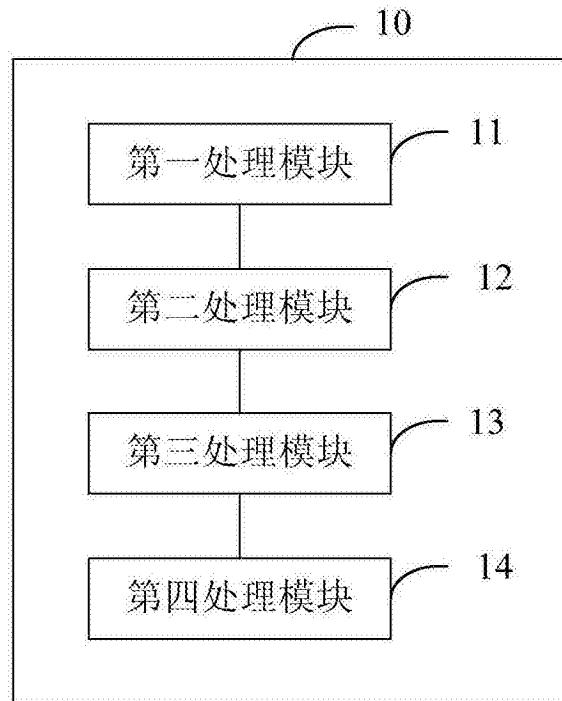


图7