



(12) 发明专利

(10) 授权公告号 CN 117409779 B

(45) 授权公告日 2024.02.23

(21) 申请号 202311714796.2

G10L 15/30 (2013.01)

(22) 申请日 2023.12.14

(56) 对比文件

(65) 同一申请的已公布的文献号

申请公布号 CN 117409779 A

CN 112382294 A, 2021.02.19

CN 107578776 A, 2018.01.12

CN 112306560 A, 2021.02.02

(43) 申请公布日 2024.01.16

CN 109493856 A, 2019.03.19

(73) 专利权人 四川蜀天信息技术有限公司

CN 110223691 A, 2019.09.10

地址 625000 四川省雅安市雨城区和兴街3

CN 110853644 A, 2020.02.28

号姚桥农贸市场1层等8处

CN 110968353 A, 2020.04.07

(72) 发明人 王帅 谭克强 赵磊 蒋维中

CN 111028834 A, 2020.04.17

周舒婷

CN 116246616 A, 2023.06.09

(74) 专利代理机构 北京超凡宏宇知识产权代理

CN 117133283 A, 2023.11.28

有限公司 11463

US 2016155443 A1, 2016.06.02

专利代理师 曹延鹏

US 2020265838 A1, 2020.08.20

US 2023298578 A1, 2023.09.21

(51) Int. Cl.

审查员 颜博

G10L 15/22 (2006.01)

G10L 15/26 (2006.01)

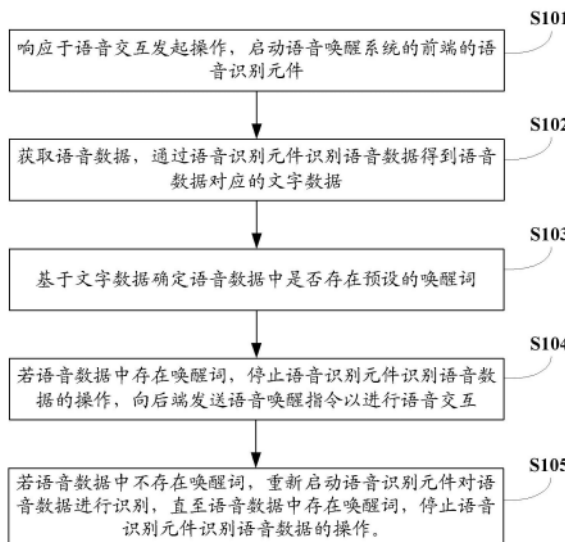
权利要求书2页 说明书11页 附图4页

(54) 发明名称

一种语音唤醒方法、装置、系统及可读介质

(57) 摘要

本发明提供了一种语音唤醒方法、装置、系统及可读介质,涉及语音识别技术领域,应用于语音唤醒系统的前端,前端与语音唤醒系统的后端通信连接,方法包括:响应于语音交互发起操作,启动语音唤醒系统的前端的语音识别元件;获取语音数据,通过语音识别元件识别语音数据得到语音数据对应的文字数据;基于文字数据确定语音数据中是否存在预设的唤醒词;若语音数据中存在唤醒词,停止语音识别元件识别语音数据的操作,向后端发送语音唤醒指令以进行语音交互;若语音数据中不存在唤醒词,重新启动语音识别元件对语音数据进行识别,直至语音数据中存在唤醒词,停止语音识别元件识别语音数据的操作,减少唤醒阶段中服务器负载压力,提高机器唤醒响应效率。



1. 一种语音唤醒方法,其特征在于,应用于语音唤醒系统的前端,所述前端与所述语音唤醒系统的后端通信连接,所述方法包括:

响应于语音交互发起操作,启动所述语音唤醒系统的前端的语音识别元件;

获取语音数据,通过所述语音识别元件识别所述语音数据得到所述语音数据对应的文字数据;

基于所述文字数据确定所述语音数据中是否存在预设的唤醒词;

若所述语音数据中存在所述唤醒词,停止所述语音识别元件识别所述语音数据的操作,向所述后端发送语音唤醒指令以进行语音交互;

若所述语音数据中不存在所述唤醒词,重新启动所述语音识别元件对所述语音数据进行识别,直至所述语音数据中存在所述唤醒词,停止所述语音识别元件识别所述语音数据的操作;

在所述停止所述语音识别元件识别所述语音数据的操作,向所述后端发送语音唤醒指令以进行语音交互之前,所述方法还包括:

基于所述唤醒词的字符占比数以及所述文字数据中是否含有预先设置的动作词集合中的任一动作词,确定所述唤醒词是否有效;

若所述唤醒词无效,则重新启动所述语音识别元件对所述语音数据进行识别,直至所述唤醒词有效;其中,所述字符占比数用于确定所述唤醒词是否有效。

2. 根据权利要求1所述的方法,其特征在于,在所述启动所述语音唤醒系统的前端的语音识别元件之前,所述方法还包括:

若未开启语音数据的获取权限,对所述获取权限的开启进行请求;

在所述语音数据的获取权限开启后,获取预先设置的配置信息;所述配置信息包括唤醒词、语种信息和语音识别函数;

基于所述配置信息对所述语音识别元件进行初始化。

3. 根据权利要求2所述的方法,其特征在于,所述通过所述语音识别元件识别所述语音数据得到所述语音数据对应的文字数据,包括:

确定所述语音数据对应的语种;

基于所述语种和所述语音识别函数将所述语音数据转化为所述文字数据。

4. 根据权利要求1所述的方法,其特征在于,确定所述唤醒词是否有效,包括:

判断所述唤醒词所处位置后的所述文字数据中是否存在任一所述动作词;

若所述唤醒词所处位置后的所述文字数据中不存在任一所述动作词,则确定所述唤醒词的字符占比数是否大于或等于预先设置的占比数阈值;

若所述唤醒词的字符占比数大于或等于所述占比数阈值,则所述唤醒词有效。

5. 根据权利要求1所述的方法,其特征在于,当所述语音数据对应的语种为中文时,所述基于所述文字数据确定所述语音数据中是否存在唤醒词,包括:

确定所述唤醒词对应的第一拼音和所述文字数据对应的第二拼音;

基于所述第一拼音和所述第二拼音进行正则比对;

若所述第二拼音中包括所述第一拼音,所述语音数据中存在所述唤醒词。

6. 根据权利要求1所述的方法,其特征在于,在所述向所述后端发送语音唤醒指令之后,所述方法还包括:

基于预先设置的时间阈值确定所述语音交互是否结束；
若所述语音交互结束，重新启动所述语音识别元件。

7. 一种语音唤醒装置，其特征在于，应用于语音唤醒系统的前端，所述前端与所述语音唤醒系统的后端通信连接，所述装置包括用于执行权利要求1至6任一项所述的语音唤醒方法的各个模块，所述模块包括语音识别元件启动模块、语音数据识别模块、唤醒词判断模块和指令发送模块，其中：

所述语音识别元件启动模块，用于响应于语音交互发起操作，启动所述语音唤醒系统的前端的语音识别元件；

所述语音数据识别模块，用于获取语音数据，通过所述语音识别元件识别所述语音数据得到所述语音数据对应的文字数据；

所述唤醒词判断模块，用于基于所述文字数据确定所述语音数据中是否存在预设的唤醒词；

所述指令发送模块，用于若所述语音数据中存在所述唤醒词，停止所述语音识别元件识别所述语音数据的操作，向所述后端发送语音唤醒指令以进行语音交互；若所述语音数据中不存在所述唤醒词，重新启动所述语音识别元件对所述语音数据进行识别，直至所述语音数据中存在所述唤醒词，停止所述语音识别元件识别所述语音数据的操作；

所述指令发送模块，还用于基于所述唤醒词的字符占比数以及所述文字数据中是否含有预先设置的动作词集合中的任一动作词，确定所述唤醒词是否有效；若所述唤醒词无效，则重新启动所述语音识别元件对所述语音数据进行识别，直至所述唤醒词有效；其中，所述字符占比数用于确定所述唤醒词是否有效。

8. 一种语音唤醒系统，其特征在于，包括：前端和后端，所述前端和后端通信连接；

所述前端用于，响应于语音交互发起操作，启动所述语音唤醒系统的前端的语音识别元件；获取语音数据，通过所述语音识别元件识别所述语音数据得到所述语音数据对应的文字数据；基于所述文字数据确定所述语音数据中是否存在预设的唤醒词；若所述语音数据中存在所述唤醒词，停止所述语音识别元件识别所述语音数据的操作，向所述后端发送语音唤醒指令以进行语音交互；若所述语音数据中不存在所述唤醒词，重新启动所述语音识别元件对所述语音数据进行识别，直至所述语音数据中存在所述唤醒词，停止所述语音识别元件识别所述语音数据的操作；在所述停止所述语音识别元件识别所述语音数据的操作，向所述后端发送语音唤醒指令以进行语音交互之前，还包括：基于所述唤醒词的字符占比数以及所述文字数据中是否含有预先设置的动作词集合中的任一动作词，确定所述唤醒词是否有效；若所述唤醒词无效，则重新启动所述语音识别元件对所述语音数据进行识别，直至所述唤醒词有效；其中，所述字符占比数用于确定所述唤醒词是否有效；

所述后端用于，对所述前端发送的语音唤醒指令进行响应；获取所述语音数据，基于所述语音数据进行语音交互。

9. 一种计算机可读介质，其特征在于，所述计算机可读介质具有处理器可执行的非易失的程序代码，所述程序代码使所述处理器执行所述权利要求1至6任一项所述的语音唤醒方法。

一种语音唤醒方法、装置、系统及可读介质

技术领域

[0001] 本发明涉及语音识别技术领域,尤其是涉及一种语音唤醒方法、装置、系统及可读介质。

背景技术

[0002] 近年来,具有语音操作功能的智能设备逐渐进入了人们的生活,这类设备在使用语音操作之前,一般需要通过特定的唤醒词来唤醒语音操作功能。现有语音服务的唤醒方法中,一般是将用户端采集的语音数据传输至云服务器或服务器进行语音识别,检测到识别文字中含有预设唤醒词则唤醒。但是,该唤醒方法中由于网络抖动、网络阻塞等问题的存在导致云服务器或服务器无法实时处理用户输入的语音数据,机器唤醒响应不及时。而且,通过云服务器或服务器进行语音识别,为了保证不错过对用户不定时输入唤醒指令的识别任务,云服务器或服务器需要独占一定数量的资源不间断的对语音数据进行实时的语音识别,不仅增大服务器的负载压力,还增加了不必要的资源消耗和系统开销,造成资源浪费和高耗能。因此,基于终端低功耗的唤醒技术方案显得尤为重要。

[0003] 现阶段,基于终端低功耗的语音唤醒产品中较为典型的语音唤醒功能的实现主要包括两类,其中一类是基于嵌入式语音唤醒芯片实现,其需要在制作智能终端产品时预先将语音唤醒芯片内嵌于智能终端设备中并激活语音唤醒功能,如天猫精灵音箱、小度智能耳机、科沃斯扫地机等;另一类是将语音唤醒SDK嵌入App中并激活语音唤醒功能,如天猫精灵App、小爱同学App、百度语音助手App等。第一类基于嵌入式语音唤醒芯片实现语音唤醒由于芯片架构的不同导致其存在芯片软件适配度差的问题,而且针对那些出厂时没有嵌入语音唤醒芯片的智能终端来说,无法通过简单的处理手段将该芯片嵌入到原智能终端设备中故而无法实现语音唤醒功能。而第二类基于内嵌语音唤醒SDK的App,由于不同终端设备所支持的操作环境以及硬件架构具备一定的差异性,则对应的语音唤醒SDK需要做定制化的更改以满足其操作系统、硬件架构适配性的要求,导致其不同操作系统下应用的语音唤醒软件之间可迁移性、灵活性不高。

发明内容

[0004] 有鉴于此,本发明的目的在于提供一种语音唤醒方法、装置及系统,以减小唤醒阶段的负载压力,同时解决现有技术中不同智能终端设备进行语音唤醒时遇到的适配性差的问题。

[0005] 第一方面,本发明实施例提供了一种语音唤醒方法,应用于语音唤醒系统的前端,前端与语音唤醒系统的后端通信连接,方法包括:响应于语音交互发起操作,启动语音唤醒系统的前端的语音识别元件;获取语音数据,通过语音识别元件识别语音数据得到语音数据对应的文字数据;基于文字数据确定语音数据中是否存在预设的唤醒词;若语音数据中存在唤醒词,停止语音识别元件识别语音数据的操作,向后端发送语音唤醒指令以进行语音交互;若语音数据中不存在唤醒词,重新启动语音识别元件对语音数据进行识别,直至语

音数据中存在唤醒词,停止语音识别元件识别语音数据的操作。

[0006] 在本发明较佳的实施例中,在上述启动语音唤醒系统的前端的语音识别元件之前,方法还包括:若未开启语音数据的获取权限,对获取权限的开启进行请求;在语音数据的获取权限开启后,获取预先设置的配置信息;配置信息包括唤醒词、语种信息和语音识别函数;基于配置信息对语音识别元件进行初始化。

[0007] 在本发明较佳的实施例中,上述通过语音识别元件识别语音数据得到语音数据对应的文字数据,包括:确定语音数据对应的语种;基于语种和语音识别函数将语音数据转化为文字数据。

[0008] 在本发明较佳的实施例中,在上述停止语音识别元件识别语音数据的操作,向后端发送语音唤醒指令以进行语音交互之前,方法还包括基于唤醒词的字符占比数以及文字数据中是否含有预先设置的动作词集合中的任一动作词,确定唤醒词是否有效;若唤醒词无效,则重新启动语音识别元件对语音数据进行识别,直至唤醒词有效;其中,字符占比数用于确定唤醒词是否有效。

[0009] 在本发明较佳的实施例中,上述确定唤醒词是否有效,包括:判断唤醒词所处位置后的文字数据中是否存在任一动作词;若唤醒词所处位置后的文字数据中不存在任一动作词,则确定唤醒词的字符占比数是否大于或等于预先设置的占比数阈值;若唤醒词的字符占比数大于或等于占比数阈值,则唤醒词有效。

[0010] 在本发明较佳的实施例中,当语音数据对应的语种为中文时,上述基于文字数据确定语音数据中是否存在唤醒词,包括:确定唤醒词对应的第一拼音和文字数据对应的第二拼音;基于第一拼音和第二拼音进行正则比对;若第二拼音中包括第一拼音,语音数据中存在唤醒词。

[0011] 在本发明较佳的实施例中,在向上述后端发送语音唤醒指令之后,方法还包括:基于预先设置的时间阈值确定语音交互是否结束;若语音交互结束,重新启动语音识别元件。

[0012] 第二方面,本发明实施例还提供一种语音唤醒装置,应用于语音唤醒系统的前端,前端与语音唤醒系统的后端通信连接,装置包括用于执行上述的语音唤醒方法的各个模块,模块包括语音识别元件启动模块、语音数据识别模块、唤醒词判断模块和指令发送模块,其中:语音识别元件启动模块,用于响应于语音交互发起操作,启动语音唤醒系统的前端的语音识别元件;语音数据识别模块,用于获取语音数据,通过语音识别元件识别语音数据得到语音数据对应的文字数据;唤醒词判断模块,用于基于文字数据确定语音数据中是否存在预设的唤醒词;指令发送模块,用于若语音数据中存在唤醒词,停止语音识别元件识别语音数据的操作,向后端发送语音唤醒指令以进行语音交互;若语音数据中不存在唤醒词,重新启动语音识别元件对语音数据进行识别,直至语音数据中存在唤醒词,停止语音识别元件识别语音数据的操作。

[0013] 第三方面,本发明实施例还提供一种语音唤醒系统,包括:前端和后端,前端和后端通信连接;前端用于,响应于语音交互发起操作,启动语音唤醒系统的前端的语音识别元件;获取语音数据,通过语音识别元件识别语音数据得到语音数据对应的文字数据;基于文字数据确定语音数据中是否存在预设的唤醒词;若语音数据中存在唤醒词,停止语音识别元件识别语音数据的操作,向后端发送语音唤醒指令以进行语音交互;若语音数据中不存在唤醒词,重新启动语音识别元件对语音数据进行识别,直至语音数据中存在唤醒词,停止

语音识别元件识别语音数据的操作。后端用于,对前端发送的语音唤醒指令进行响应;获取语音数据,基于语音数据进行语音交互。

[0014] 第四方面,本发明实施例还提供一种计算机可读介质,计算机可读介质具有处理器可执行的非易失的程序代码,程序代码使处理器执行上述第一方面的语音唤醒方法。

[0015] 本发明实施例带来了以下有益效果:

[0016] 本发明实施例提供了一种语音唤醒方法、装置、系统及可读介质,相对于前端将采集的语音数据传输至后端进行识别并判断是否含有唤醒词的方法,本申请中利用前端轻量级的语音识别元件对语音数据进行识别,并基于是否在语音数据中识别出预设的唤醒词判断是唤醒后端进行语音交互,还是重启语音识别元件对用户后续输入的语音数据进行识别,直至识别到预设的唤醒词后停止工作。在语音识别元件的基础上通过重启机制和判断机制的双机制协同工作实现对后端的语音唤醒,如此在前端进行唤醒词的语音识别以减少在唤醒阶段中服务器的负载压力和资源开销,提高机器唤醒响应效率。再者,本申请中是以不同的智能终端设备支持的语音识别元件做基础,不需要针对该设备的支持操作系统以及硬件架构做适配性调整。因而本申请改造后得到的语音识别元件的可迁移性和灵活度较高,且改造后的语音识别元件的体积小、量级轻,几乎不影响应用程序的性能,响应速率快。

[0017] 本公开的其他特征和优点将在随后的说明书中阐述,或者,部分特征和优点可以从说明书推知或毫无疑问地确定,或者通过实施本公开的上述技术即可得知。

[0018] 为使本公开的上述目的、特征和优点能更明显易懂,下文特举较佳实施例,并配合所附图,作详细说明如下。

附图说明

[0019] 为了更清楚地说明本发明具体实施方式或现有技术中的技术方案,下面将对具体实施方式或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施方式,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0020] 图1为本发明实施例提供的一种语音唤醒方法的流程图;

[0021] 图2为本发明实施例提供的另一种语音唤醒方法的流程图;

[0022] 图3为本发明实施例提供的又一种语音唤醒方法的流程图;

[0023] 图4为本发明实施例提供的基于文字数据确定语音数据中是否存在唤醒词的流程图;

[0024] 图5为本发明实施例提供的一种语音唤醒装置的结构示意图;

[0025] 图6为本发明实施例提供的一种语音唤醒系统的结构示意图;

[0026] 图7为本发明实施例提供的一种电子设备的结构示意图。

具体实施方式

[0027] 为使本发明实施例的目的、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。

[0028] 近年来,具有语音操作功能的智能设备逐渐进入了人们的生活,这类设备在使用语音操作之前,一般需要通过特定的唤醒词来唤醒语音操作功能。

[0029] 现有语音服务的唤醒方法中,一般是将用户端采集的语音数据传输至后端服务器进行语音识别,检测到识别文字中含有预设唤醒词则唤醒。但是,该唤醒方法中由于网络抖动、网络阻塞等问题的存在导致后端无法实时处理用户输入的语音数据,导致唤醒响应不及时,而且,通过后端服务器进行语音识别,在海量用户的使用场景下会增大服务器的负载压力。

[0030] 基于此,本发明实施例提供的一种语音唤醒方法、装置、系统及可读介质,可以通过响应于语音交互发起操作,启动语音唤醒系统的前端的语音识别元件,再获取语音数据,通过语音识别元件识别语音数据得到语音数据对应的文字数据,然后基于文字数据确定语音数据中是否存在唤醒词,在语音数据中存在唤醒词时,停止语音识别元件识别语音数据的操作,向后端发送语音唤醒指令以进行语音交互。通过前端的语音识别元件先对语音数据中是否存在唤醒词进行确定,在确定存在唤醒词后再将语音唤醒指令发送至后端,减小了唤醒阶段的负载压力,提高了唤醒响应效率。在语音数据中不存在唤醒词,重新启动语音识别元件对语音数据进行识别,直至语音数据中存在唤醒词,停止语音识别元件识别语音数据的操作。在语音识别原件的基础上通过重启机制和判断机制的双机制协同工作实现对后端的语音唤醒,以达到在不同的智能终端设备时不需要针对该设备的支持操作系统以及硬件架构做适配性调整,该语音识别元件的可迁移性和灵活度较高。

[0031] 为便于对本实施例进行理解,首先对本发明实施例所公开的一种语音唤醒方法进行详细介绍。

[0032] 实施例1

[0033] 本发明实施例提供一种语音唤醒方法,应用于语音唤醒系统的前端,前端与语音唤醒系统的后端通信连接,图1为本发明实施例提供的一种语音唤醒方法的流程图。如图1所示,该语音唤醒方法可以包括如下步骤:

[0034] 步骤S101,响应于语音交互发起操作,启动语音唤醒系统的前端的语音识别元件。

[0035] 其中,语音交互可以应用于浏览器中,也可以应用于与人工智能机器人,此处不作限定。

[0036] 以应用于浏览器为例,语音交互发起操作可以为用户打开浏览器进入浏览器的交互页面,在用户打开浏览器进入浏览器的交互页面后语音识别元件就会启动。

[0037] 以应用于与人工智能机器人为例,语音交互发起操作可以为给人工智能机器人通电,在通电后语音识别元件就会启动。

[0038] 步骤S102,获取语音数据,通过语音识别元件识别语音数据得到语音数据对应的文字数据。

[0039] 其中,语音数据可以由音频采集设备进行采集并发送至语音识别元件,音频采集设备可以是麦克风等。

[0040] 其中,语音识别元件在接收到语音数据后,会发起调用表征开始识别语音数据的函数的第一调用请求,在第一调用请求通过后,在语音识别元件开始对语音数据进行识别时,会发起调用表征识别语音数据的函数的第二调用请求,在第二调用请求通过后,语音识别元件会开始对语音数据进行识别。

[0041] 进一步地,表征识别语音数据的函数可以将语音数据转化为对应的文字数据,并将文字数据进行逻辑断句。例如识别语音数据函数可以采用现有技术的断点检测算法,通过监测语音数据静默状态进行断句。

[0042] 其中,语音识别元件识别在启动前会根据配置信息进行初始化,配置信息可以包括唤醒词、语种信息和语音数据转化为文字数据等。

[0043] 具体地,通过语音识别元件识别语音数据得到语音数据对应的文字数据,可以包括:确定语音数据对应的语种;基于语种和语音识别函数将语音数据转化为文字数据。

[0044] 其中,语音识别元件可以先根据语音数据确定对应的语种,根据该语种调用对应的语音识别函数对语音数据进行识别,从而将语音数据转化为文字数据。

[0045] 步骤S103,基于文字数据确定语音数据中是否存在预设的唤醒词。

[0046] 其中,在语音识别元件将语音数据转化为文字数据后,会发起调用表征识别唤醒词的函数的第三调用请求,在第三调用请求通过后,会基于文字数据确定语音数据中是否存在预设的唤醒词。

[0047] 步骤S104,若语音数据中存在唤醒词,停止语音识别元件识别语音数据的操作,向后端发送语音唤醒指令以进行语音交互。

[0048] 其中,通过前端对唤醒词进行预先判定,在确定存在唤醒词后再向后端发送语音唤醒指令以进行语音交互,提高了唤醒效率效率,减小了唤醒阶段的负载压力。

[0049] 具体地,在向后端发送语音唤醒指令之后,方法还可以包括:基于预先设置的时间阈值确定语音交互是否结束;若语音交互结束,重新启动语音识别元件。

[0050] 其中,在预先设置的时间阈值内若一直未检测到语音数据的输入,则可以认为语音交互结束。

[0051] 在后端接收语音唤醒指令与用户开始进行语音交互时,语音识别元件就会关闭,此时用户输入的语音数据将由后端的服务器进行解析并响应,也就是说用户在此后持续产生的语音数据不会被语音识别元件识别,直至语音交互结束后,再重新启动语音识别元件,来继续对语音数据进行识别来确定是否需要唤醒后端。解决了现有语音识别元件在得到当前语音识别结果后直接停止工作进行休眠等待重启,而不能持续识别用户输入的语音数据的问题,减小了耗能。

[0052] 步骤S105,若语音数据中不存在唤醒词,重新启动语音识别元件对语音数据进行识别,直至语音数据中存在唤醒词,停止语音识别元件识别语音数据的操作。

[0053] 本发明实施例提供的语音唤醒方法,利用前端设置的轻量级的语音识别元件对语音数据进行识别,并基于是否在语音数据中识别出预设的唤醒词判断是唤醒后端进行语音交互,还是重启语音识别元件对用户后续输入的语音数据进行识别,直至识别到唤醒词后停止工作,在语音识别元件的基础上通过重启机制和判断机制的双机制协同工作实现对后端的语音唤醒,如此在前端进行语音识别以减少在唤醒阶段中服务器的负载压力,提高机器唤醒响应效率。而且,本申请中是以不同的智能终端设备支持的语音识别元件做基础,不需要针对该设备的支持操作系统以及硬件架构做适配性调整。因而本申请改造后得到的语音识别元件的可迁移性和灵活度较高,且改造后的语音识别元件的体积小、量级轻,几乎不影响应用程序的性能,响应速率快。

[0054] 实施例2

[0055] 本发明实施例还提供另一种语音唤醒方法;该方法在上述实施例方法的基础上实现;该方法重点描述在上述启动语音唤醒系统的前端的语音识别元件之前还需执行的步骤。

[0056] 图2为本发明实施例提供的另一种语音唤醒方法的流程图,如图2所示,该语音唤醒方法还可以包括如下步骤:

[0057] 步骤S201,确定是否开启语音数据的获取权限。

[0058] 其中,语音数据的获取权利可以是麦克风权限。

[0059] 在实际应用过程中,若步骤S201的判断结果为是,即语音数据的获取权限已开启,则执行步骤S202,若步骤S201的判断结果为否,即语音数据的获取权限未开启,则执行步骤S204。

[0060] 步骤S202,获取预先设置的配置信息。

[0061] 其中,预先设置的配置信息包括唤醒词、语种信息和语音识别函数。

[0062] 其中,唤醒词可以例如“你好”、“小潘小潘”等,语种信息可以例如中文、英文等的相关信息,语音识别函数为对语音数据进行识别的函数,可以根据语音数据的语种调用对应语音识别函数进行语音数据的识别。

[0063] 其中,配置信息中还可以包括表征开始识别语音数据的函数、表征识别语音数据的函数以及表征识别唤醒词的函数。

[0064] 步骤S203,基于配置信息对语音识别元件进行初始化。

[0065] 针对来着不同国家的用户,其常用语种并不相同,若仅设置单一语种语音识别则对于使用其他语种的用户来说,鉴于发音错误等缘由会造成很多漏唤醒,进而降低唤醒成功率。为了降低语音交互门槛,满足用户对于不同语种的个性化需求,本申请中能够根据浏览器中浏览界面或人工智能机器人使用的语种信息对语音识别元件可以应用的语种进行自动化更新,有利于针对不同国家用户调整语音识别元件加载的识别言语的语种信息,提高唤醒成功率。

[0066] 作为一种实现方式,可以针对不同国家地区的不同方言,制作对应的配置包,在应用于不同方言的情况下,可以根据实际需求进行对应配置包的预下载得到对应的配置信息,以供在启动语音识别元件前利用对应的配置信息进行初始化,大大提高了灵活度以及唤醒响应效率。

[0067] 步骤S204,对获取权限的开启进行请求

[0068] 其中,可以通过网页弹窗或语音提示等进行请求。

[0069] 本发明实施例提供的语音唤醒方法,在未开启语音数据的获取权限的情况下,会对获取权限的开启进行请求,在开启了语音数据的获取权限的情况下,会根据配置信息对语音识别元件进行初始化,从而实现根据浏览器中设备浏览界面或人工智能机器人使用的语种信息对语音识别元件可以应用的语种进行自动化更新,有利于针对不同国家用户调整语音识别元件加载的识别言语的语种信息,提高唤醒成功率。

[0070] 实施例3

[0071] 本发明实施例还提供又一种语音唤醒方法;该方法在上述实施例方法的基础上实现;该方法重点描述在上述停止语音识别元件识别语音数据的操作,向后端发送语音唤醒指令以进行语音交互之前还需执行的步骤。

[0072] 图3为本发明实施例提供的又一种语音唤醒方法的流程图,如图3所示,该语音唤醒方法还可以包括如下步骤:

[0073] 步骤S301,基于唤醒词的字符占比数以及文字数据中是否含有预先设置的动作词集合中的任一动作词,确定唤醒词是否有效。

[0074] 其中,字符占比数用于确定唤醒词是否有效,预先设置的动作词集合中包括多个动作词,动作词用于对后端进行操作指令的下达,例如,查询、搜索、打开、关闭等。

[0075] 具体地,确定唤醒词是否有效,可以包括:判断唤醒词所处位置后的文字数据中是否存在任一动作词;若唤醒词所处位置后的文字数据中不存在任一动作词,则确定唤醒词的字符占比数是否大于或等于预先设置的占比数阈值;若唤醒词的字符占比数大于或等于占比数阈值,则唤醒词有效。

[0076] 其中,可以通过上述表征识别语音数据的函数对文字数据进行逻辑断句,得到至少一个语句文本。

[0077] 在实际使用过程中,用户输入唤醒词时可能同时会说出自己的指令,如“小X帮我打开电视”。本申请通过判断在唤醒词所处的位置之后的文字数据中是否存在任一动作词,如果存在动作词,则可以认为用户输入唤醒词后又继续下达了操作指令,则可以认为唤醒词是有效的。

[0078] 但是,在实际应用场景中有的用户有唤醒需求时,在向音频采集设备输入唤醒词后会等待设备反馈有效唤醒信息(例如有效唤醒后设备回复用户“在呢”)后才会再输入自己的需求,此时在语句文本中不存在动作词因而仅仅通过唤醒词之后的文字数据中有无动作词判断是否为有效唤醒会造成漏唤醒。为了解决上述问题,如果不存在动作词,则会通过唤醒词字符占比数对唤醒词是否有效继续进行进一步判断。

[0079] 由于在确定存在唤醒词前在判断机制和重启机制的协同工作下会使得语音识别元件一直处于开启状态,当用户为了容易记忆唤醒词,而将唤醒词设置为较为常用的词语时,容易出现误唤醒的情形。例如唤醒词为“小X”,且不存在动作词时,若输入语音数据为“小X很好用”此时意图并不是唤醒,而当输入语音数据为“小X小X在吗”此时意图意图是唤醒,针对这种不存在动作词,且唤醒词和需求指令分多次输入的场景,本申请通过确定唤醒词的字符占比数来降低误唤醒的概率,在字符占比数小于占比数阈值时,认为唤醒词是无效的,字符占比数大于或等于占比数阈值时,认为唤醒词是有效的。

[0080] 其中,占比数阈值由用户设置的唤醒词的字数来决定。

[0081] 具体地,确定唤醒词的字符占比数可以包括:确定唤醒词的字数占唤醒词所处的语句文本的字数的百分比;将该百分比作为唤醒词的字符数占比。

[0082] 在实际应用过程中,若步骤S301的判断结果为否,即唤醒词无效,则执行步骤S302,若步骤S301的判断结果为是,即唤醒词有效,则执行步骤S303。

[0083] 步骤S302,重新启动语音识别元件对语音数据进行识别,直至唤醒词有效。

[0084] 步骤S303,停止语音识别元件识别语音数据的操作,向后端发送语音唤醒指令以进行语音交互。

[0085] 进一步地,为了更进一步的确定唤醒词是否有效,本申请中还可以基于唤醒词在语句文本中所处的位置、字符占比数以及文字数据中是否含有预先设置的动作词集合中的任一动作词,确定所述唤醒词是否有效。例如,当多媒体教室中某个智能终端唤醒词设置为

“小X”时,老师在上课途中想要请某位同学读课文而被音频采集设备采集到“我们请小X同学帮我们念一下本段课文”时存在误唤醒,且诸如此类的语音数据还包括“我们将唤醒词设置为 小X”等。为了过滤掉此类误唤醒情况,可以将唤醒词处于语句文本的句首作为唤醒词有效的又一个判定条件,在唤醒词在任一语句文本的句首的情况下,会对唤醒词是否有效进行进一步判断。具体方法包括但不限于以下几种:

[0086] (1)对文字数据进行逻辑断句,得到至少一个语句文本;基于唤醒词在语句文本中所处的位置,确定唤醒词是否在任一语句文本的句首;若唤醒词在任一语句文本的句首,判断位置后的文字数据中是否存在任一动作词;若位置后的文字数据中不存在任一动作词,则确定唤醒词的字符占比数是否大于或等于预先设置的占比数阈值;若唤醒词的字符占比数大于或等于占比数阈值,则唤醒词有效。

[0087] (2)对文字数据进行逻辑断句,得到至少一个语句文本;基于每个语句文本确定是否存在任一动作词;若语句文本中存在任一动作词,则确定唤醒词在语句文本中所处的位置,确定唤醒词是否在任一语句文本的句首;若唤醒词在任一语句文本的句首,则唤醒词有效。

[0088] (3)对文字数据进行逻辑断句,得到至少一个语句文本;基于每个语句文本确定唤醒词的字符占比数是否大于或等于占比数阈值;若唤醒词的字符占比数大于或等于占比数阈值,则确定唤醒词在语句文本中所处的位置,确定唤醒词是否在任一语句文本的句首;若唤醒词在任一语句文本的句首,则唤醒词有效。

[0089] 本发明实施例提供的语音唤醒方法,可以通过对唤醒词所处的位置之后的文字数据是否存在任一动作词来进行唤醒词是否有效的进一步判断,在不存在任一动作词的情况下,通过唤醒词的字符占比数与占比数阈值的大小关系进行唤醒词是否有效的最后判断,在唤醒词有效的情况下对后端进行唤醒,降低了后端被误唤醒的概率,提高了唤醒的准确性。

[0090] 实施例4

[0091] 本发明实施例还提供另一种语音唤醒方法;该方法在上述实施例方法的基础上实现;该方法重点描述上述基于文字数据确定语音数据中是否存在唤醒词的具体实现方式。

[0092] 图4为本发明实施例提供的基于文字数据确定语音数据中是否存在唤醒词的流程图,如图4所示,当语音数据对应的语种为中文时,该基于文字数据确定语音数据中是否存在唤醒词可以包括如下步骤:

[0093] 步骤S401,确定唤醒词对应的第一拼音和文字数据对应的第二拼音。

[0094] 其中,可以利用拼音识别元件对唤醒词和文字数据进行处理,得到第一拼音和第二拼音。

[0095] 步骤S402,基于第一拼音和第二拼音进行正则比对。

[0096] 步骤S403,若第二拼音中包括第一拼音,语音数据中存在唤醒词。

[0097] 由于唤醒词一般都是一个词语或者比较短小的句子,因而语境匮乏缺少对应的对话场景,造成识别得到的文字数据可能出现音同字不同的情况,若是都以文字数据进行匹配那么就会出现漏唤醒的情况。

[0098] 本发明实施例提供的语音唤醒方法,可以通过将唤醒词和文字数据转化成拼音的形式进行正则对比,避免了在识别得到的文字数据出现音同字不同的情况下,出现漏唤醒

的情况,提高了唤醒响应效率。

[0099] 实施例5

[0100] 对应于上述方法实施例,本发明实施例提供了一种语音唤醒装置,应用于语音唤醒系统的前端,前端与语音唤醒系统的后端通信连接,图5为本发明实施例提供的一种语音唤醒装置的结构示意图,如图5所示,该语音唤醒装置可以包括:

[0101] 语音识别元件启动模块501,用于响应于语音交互发起操作,启动语音唤醒系统的前端的语音识别元件。

[0102] 语音数据识别模块502,用于获取语音数据,通过语音识别元件识别语音数据得到语音数据对应的文字数据。

[0103] 唤醒词判断模块503,用于基于文字数据确定语音数据中是否存在预设的唤醒词。

[0104] 指令发送模块504,用于若语音数据中存在唤醒词,停止语音识别元件识别语音数据的操作,向后端发送语音唤醒指令以进行语音交互;若语音数据中不存在唤醒词,重新启动语音识别元件对语音数据进行识别,直至语音数据中存在唤醒词,停止语音识别元件识别语音数据的操作。

[0105] 在一些实施例中,语音识别元件启动模块,还用于若未开启语音数据的获取权限,对获取权限的开启进行请求;在语音数据的获取权限开启后,获取预先设置的配置信息;配置信息包括唤醒词、语种信息和语音识别函数;基于配置信息对语音识别元件进行初始化。

[0106] 在一些实施例中,语音数据识别模块,还用于确定语音数据对应的语种;基于语种和语音识别函数将语音数据转化为文字数据。

[0107] 在一些实施例中,指令发送模块,还用于基于唤醒词的字符占比数以及文字数据中是否含有预先设置的动作词集合中的任一动作词,确定唤醒词是否有效;若唤醒词无效,则重新启动语音识别元件对语音数据进行识别,直至唤醒词有效;其中,字符占比数用于确定唤醒词是否有效。

[0108] 在一些实施例中,指令发送模块,还用于判断唤醒词所处位置后的文字数据中是否存在任一动作词;若位置后的文字数据中不存在任一动作词,则确定唤醒词的字符占比数是否大于或等于预先设置的占比数阈值;若唤醒词的字符占比数大于或等于占比数阈值,则唤醒词有效。

[0109] 在一些实施例中,唤醒词判断模块,还用于确定唤醒词对应的第一拼音和文字数据对应的第二拼音;基于第一拼音和第二拼音进行正则比对;若第二拼音中包括第一拼音,语音数据中存在唤醒词。

[0110] 在一些实施例中,指令发送模块,还用于基于预先设置的时间阈值确定语音交互是否结束;若语音交互结束,重新启动语音识别元件。

[0111] 本发明实施例所提供的装置,其实现原理及产生的技术效果和前述方法实施例相同,为简要描述,装置实施例部分未提及之处,可参考前述方法实施例中相应内容。

[0112] 实施例6

[0113] 本发明实施例还提供了一种语音唤醒系统,用于运行上述语音唤醒方法;参见图6所示的一种语音唤醒系统的结构示意图,该语音唤醒系统可以包括前端601和后端602,所述前端601和后端602通信连接。

[0114] 所述前端601用于,响应于语音交互发起操作,启动所述语音唤醒系统的前端601

的语音识别元件;获取语音数据,通过所述语音识别元件识别所述语音数据得到所述语音数据对应的文字数据;基于所述文字数据确定所述语音数据中是否存在预设的唤醒词;若所述语音数据中存在唤醒词,停止所述语音识别元件识别所述语音数据的操作,向所述后端602发送语音唤醒指令以进行语音交互;若语音数据中不存在唤醒词,重新启动语音识别元件对语音数据进行识别,直至语音数据中存在唤醒词,停止语音识别元件识别语音数据的操作。

[0115] 所述后端602用于,对所述前端601发送的语音唤醒指令进行响应;获取所述语音数据,基于所述语音数据进行语音交互。

[0116] 实施例7

[0117] 本发明实施例还提供了一种电子设备,用于运行上述语音唤醒方法;参见图7所示的一种电子设备的结构示意图,该电子设备包括存储器700和处理器701,其中,存储器700用于存储一条或多条计算机指令,一条或多条计算机指令被处理器701执行,以实现上述语音唤醒方法。

[0118] 进一步地,图7所示的电子设备还包括总线702和通信接口703,处理器701、通信接口703和存储器700通过总线702连接。

[0119] 其中,存储器700可能包含高速随机存取存储器(RAM,Random Access Memory),也可能还包括非不稳定的存储器(non-volatile memory),例如至少一个磁盘存储器。通过至少一个通信接口703(可以是有线或者无线)实现该系统网元与至少一个其他网元之间的通信连接,可以使用互联网,广域网,本地网,城域网等。总线702可以是ISA总线、PCI总线或EISA总线等。总线可以分为地址总线、数据总线、控制总线等。为便于表示,图7中仅用一个双向箭头表示,但并不表示仅有一根总线或一种类型的总线。

[0120] 处理器701可能是一种集成电路芯片,具有信号的处理能力。在实现过程中,上述方法的各步骤可以通过处理器701中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器701可以是通用处理器,包括中央处理器(Central Processing Unit,简称CPU)、网络处理器(Network Processor,简称NP)等;还可以是数字信号处理器(Digital Signal Processor,简称DSP)、专用集成电路(Application Specific Integrated Circuit,简称ASIC)、现场可编程门阵列(Field-Programmable Gate Array,简称FPGA)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。可以实现或者执行本发明实施例中的公开的各方法、步骤及逻辑框图。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。结合本发明实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成,或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器,闪存、只读存储器,可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器700,处理器701读取存储器700中的信息,结合其硬件完成前述实施例的方法的步骤。

[0121] 本发明实施例还提供了一种计算机可读介质,该计算机可读介质具有处理器可执行的非易失的程序代码,该程序代码使处理器执行上述语音唤醒方法,具体实现可参见方法实施例,在此不再赘述。

[0122] 本发明实施例所提供的进行语音唤醒方法的计算机程序产品,包括存储了处理器可执行的非易失的程序代码的计算机可读存储介质,所述程序代码包括的指令可用于执行

前面方法实施例中所述的方法,具体实现可参见方法实施例,在此不再赘述。

[0123] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统、装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0124] 在本发明所提供的几个实施例中,应该理解到,所揭露的系统、装置和方法,可以通过其它的方式实现。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,又例如,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通信接口,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。

[0125] 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。

[0126] 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。

[0127] 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个处理器可执行的非易失的计算机可读取存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品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)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0128] 最后应说明的是:以上所述实施例,仅为本发明的具体实施方式,用以说明本发明的技术方案,而非对其限制,本发明的保护范围并不局限于此,尽管参照前述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:任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,其依然可以对前述实施例所记载的技术方案进行修改或可轻易想到变化,或者对其中部分技术特征进行等同替换;而这些修改、变化或者替换,并不使相应技术方案的本质脱离本发明实施例技术方案的精神和范围,都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以权利要求的保护范围为准。

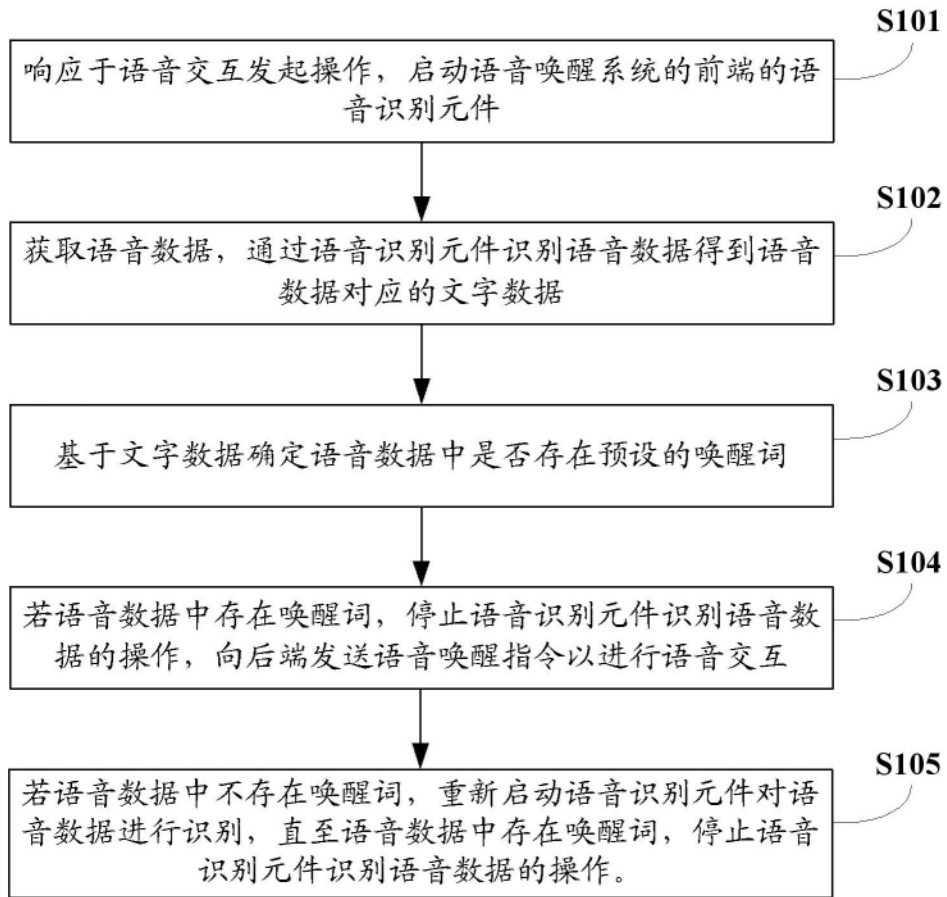


图1

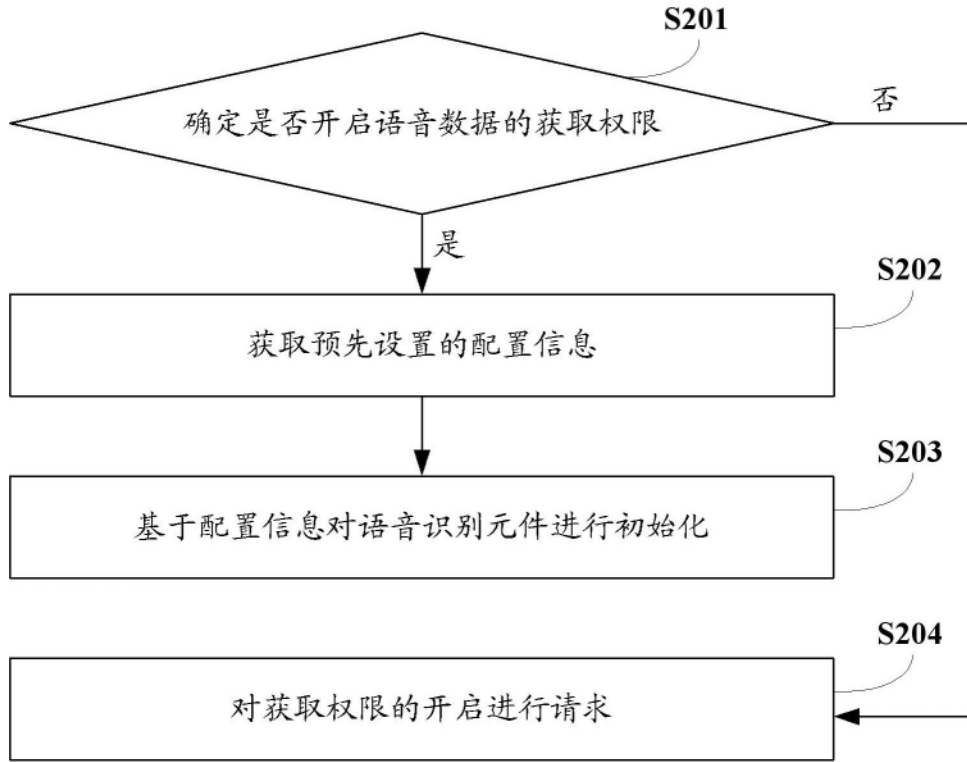


图2

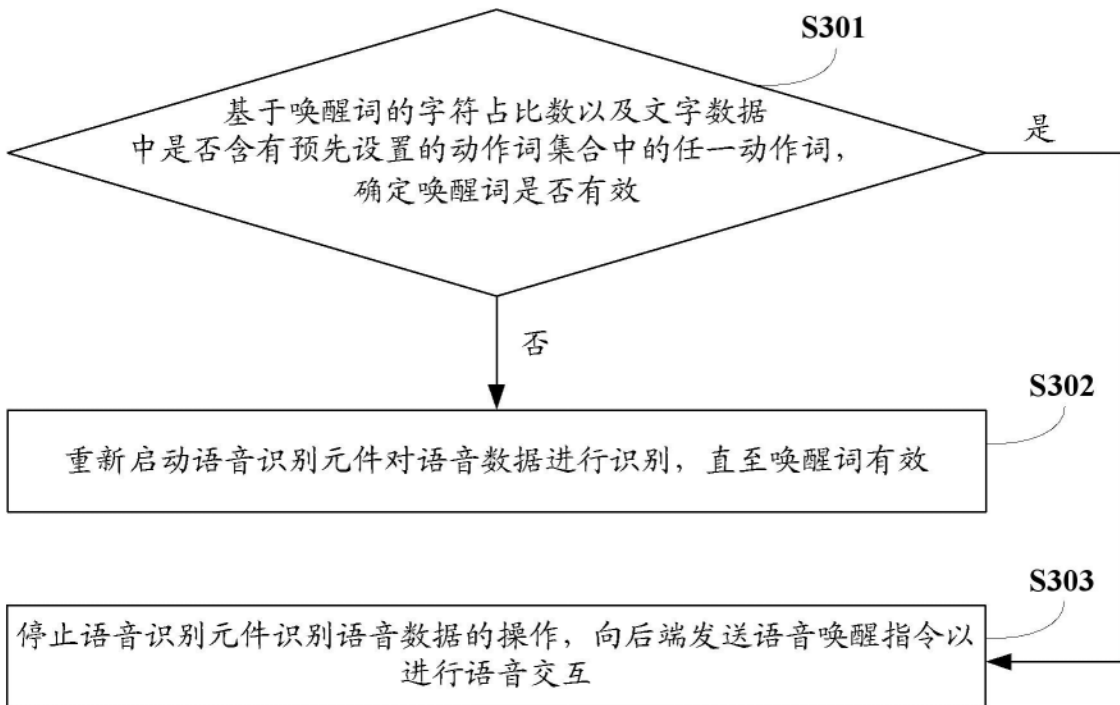


图3

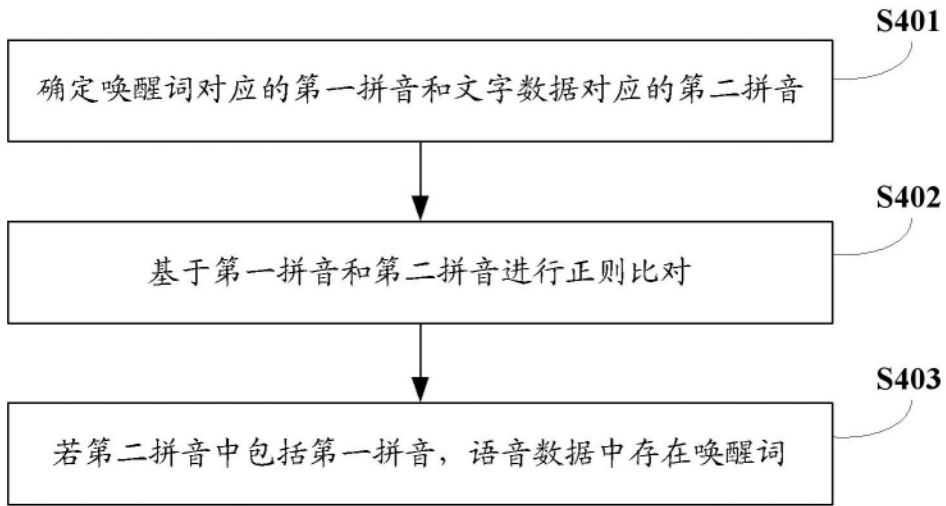


图4

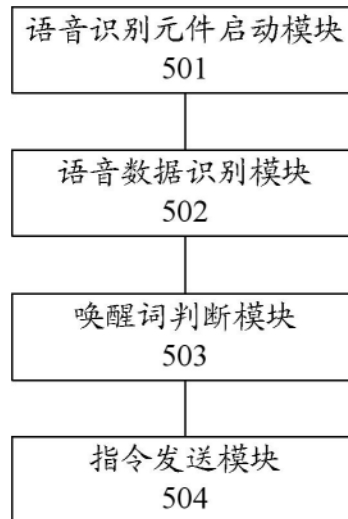


图5

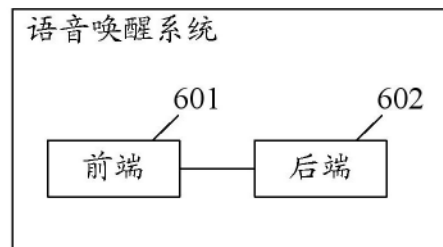


图6

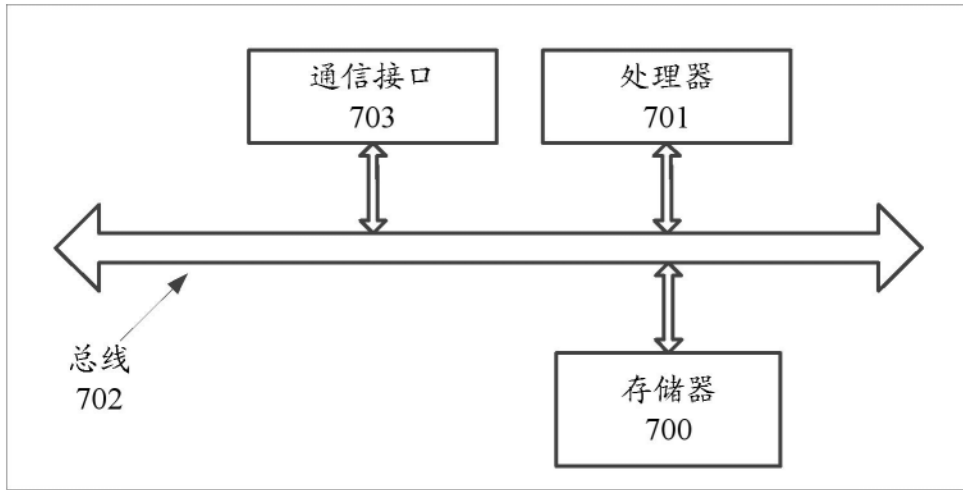


图7