



(12)发明专利申请

(10)申请公布号 CN 109712620 A

(43)申请公布日 2019.05.03

(21)申请号 201811585283.5

G10L 15/28(2013.01)

(22)申请日 2018.12.24

(71)申请人 广东美的白色家电技术创新中心有限公司

地址 528311 广东省佛山市顺德区北滘镇北滘居委会工业大道美的全球创新中心4栋2楼

申请人 美的集团股份有限公司

(72)发明人 徐小峰 张晨 黄海斌 刘欣

(74)专利代理机构 深圳市威世博知识产权代理事务所(普通合伙) 44280

代理人 钟子敏

(51)Int.Cl.

G10L 15/22(2006.01)

G10L 15/26(2006.01)

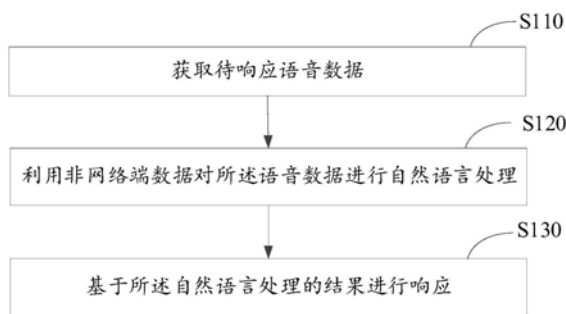
权利要求书2页 说明书9页 附图4页

(54)发明名称

语音交互方法、语音交互设备和存储装置

(57)摘要

本申请公开了一种语音交互方法、语音交互设备和存储装置。其中,所述方法包括:获取待响应语音数据;利用非网络端数据对所述待响应语音数据进行自然语言处理;基于所述自然语言处理的结果进行响应。上述方案,能够降低网络状况对设备语音交互的影响。



1. 一种语音交互方法,其特征在于,包括:
 - 获取待响应语音数据;
 - 利用非网络端数据对所述待响应语音数据进行自然语言处理;
 - 基于所述自然语言处理的结果进行响应。
2. 根据权利要求1所述的方法,其特征在于,所述对所述待响应语音数据进行自然语言处理,包括:
 - 将所述待响应语音数据转换成文本数据;
 - 对所述文本数据进行自然语言处理。
3. 根据权利要求2所述的方法,其特征在于,所述将所述待响应语音数据转换成文本数据,包括:
 - 基于本地词汇库中的词汇将所述待响应语音数据转换成文本数据;
 - 其中,所述本地词汇库中包括以下至少一者:设定领域的词汇和通用词汇。
4. 根据权利要求2所述的方法,在所述将所述待响应语音数据转换成文本数据之后,所述方法还包括:
 - 判断所述文本数据是否属于设定领域的内容;
 - 所述对所述文本数据进行自然语言处理,包括:
 - 响应于所述文本数据属于所述设定领域的内容,对所述文本数据进行自然语言处理。
5. 根据权利要求4所述的方法,所述判断所述文本数据是否属于设定领域的内容,包括:
 - 利用本地分类模型对所述文本数据进行分析,并基于分析结果确定所述文本数据是否属于设定领域的内容。
6. 根据权利要求2所述的方法,其特征在于,所述对所述文本数据进行自然语言处理,包括:
 - 对所述文本数据进行语义理解,得到语义结果;
 - 基于所述语义结果生成第一响应数据。
7. 根据权利要求6所述的方法,其特征在于,所述基于所述语义结果生成第一响应数据,包括:
 - 生成与所述语义结果匹配的执行指令;
 - 所述基于所述自然语言处理的结果进行响应,包括:
 - 执行所述执行指令。
8. 根据权利要求6所述的方法,其特征在于,所述基于所述语义结果生成第一响应数据,包括:
 - 生成与所述语义结果匹配的回复数据,其中,所述回复数据为文本回复数据或语音回复数据;
 - 所述基于所述自然语言处理的结果进行响应,包括:
 - 显示所述文本回复数据或语音播放所述语音回复数据。
9. 根据权利要求8所述的方法,其特征在于,所述生成与所述语义结果匹配的回复数据,包括:
 - 确定与所述语义结果匹配的回复内容;

对所述回复内容进行自然语言生成,得到文本回复数据;

将所述文本回复数据转换成语音回复数据。

10. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

判断利用非网络端数据是否能够对所述待响应语言数据进行自然语言处理;

响应于利用非网络端数据能够对所述待响应语言数据进行自然语言处理,依赖网络端数据对所述待响应语音数据进行自然语言处理。

11. 根据权利要求10所述的方法,其特征在于,所述判断利用非网络端数据是否能够对所述待响应语言数据进行自然语言处理,包括:

检测在预设时间内是否有进行对所述待响应语音数据的响应动作;

所述依赖网络端数据对所述待响应语音数据进行自然语言处理,包括:

将所述待响应语音数据发送给网络端设备;

接收所述网络端设备反馈的第二响应数据,其中,所述第二响应数据为所述网络端设备对所述待响应语音数据进行自然语言处理得到。

12. 一种语音交互设备,其特征在于,包括相互耦接的存储器和处理器;其中,所述处理器用于执行所述存储器存储的程序指令实现权利要求1至11任一项所述的方法。

13. 一种存储装置,其特征在于,存储有能够被处理器运行的程序指令,所述程序指令用于实现权利要求1至11任一项所述的方法。

语音交互方法、语音交互设备和存储装置

技术领域

[0001] 本申请涉及语音处理领域,特别是涉及一种语音交互方法、语音交互设备和存储装置。

背景技术

[0002] 随着信息技术的不断发展,用户交互技术得到了广泛的应用。而语音交互作为继键盘交互、触摸屏交互后的新一代用户交互模式,具有方便快捷的特点,逐渐被广泛应用于各领域中。例如,对于家电领域,家智能家居设备如音箱等在其传统功能上,还可加上语音交互模式,以实现与用户之间的互动交流。

[0003] 目前,设备在与用户进行语音交互时,需要对其采集到的用户语音数据进行处理进而对其进行响应。然而,由于对语音数据进行处理的过程较为复杂,通常是发送给网络端专门的语音处理平台来实现,并通过接受语音处理平台反馈的处理结果做出响应。这种依赖线上处理的语音交互模式,会受限于设备当前所处的网络情况,若当前网络情况较差,则设备可能无法及时针对用户语音数据进行响应或者直接无法做出响应,故极大影响设备与用户之间的正常语音交互。

发明内容

[0004] 本申请主要解决的技术问题是提供一种语音交互方法、语音交互设备和存储装置,能够降低网络状况对语音交互的影响。

[0005] 为了解决上述问题,本申请第一方面提供了一种语音交互方法,其特征在于,包括:获取待响应语音数据;利用非网络端数据对所述语音数据进行自然语言处理;基于所述自然语言处理的结果进行响应。

[0006] 为了解决上述问题,本申请第二方面提供了一种语音交互设备,包括相互耦接的存储器和处理器;其中,所述处理器用于执行所述存储器存储的程序指令实现上述的方法。

[0007] 为了解决上述问题,本申请第三方面提供了一种存储装置,存储有处理器可运行的程序指令,所述程序指令用于执行上述的方法。

[0008] 上述方案中,利用非网络端数据对待响应语音数据的自然语言处理,该处理过程无需依赖网络端数据,故可实现离线式的语音交互,降低甚至避免了网络状况对设备语音交互的影响。而且,由于网络端的数据交互需要耗费一定时间,故相比于依赖网络端数据,设备直接依赖非网络端数据对语音数据进行自然语言处理,进而基于该自然语言处理的结果进行响应,可加快语言交互的响应速度。

附图说明

[0009] 图1是本申请语音交互方法一实施例的流程示意图;

[0010] 图2是本申请语音交互方法另一实施例中步骤S120的流程示意图;

[0011] 图3a是本申请语音交互方法再一实施例中步骤S223的流程示意图

- [0012] 图3b是本申请语音交互设备一实施例的结构示意图；
- [0013] 图4a是本申请语音交互方法再一实施例的流程示意图；
- [0014] 图4b是本申请语音交互系统一实施例的结构示意图；
- [0015] 图5是本申请语音交互设备另一实施例的结构示意图；
- [0016] 图6是本申请存储装置一实施例的结构示意图。

具体实施方式

[0017] 下面结合说明书附图,对本申请实施例的方案进行详细说明。

[0018] 以下描述中,为了说明而不是为了限定,提出了诸如特定系统结构、接口、技术之类的具体细节,以便透彻理解本申请。

[0019] 本文中术语“系统”和“网络”在本文中常被可互换使用。本文中术语“和/或”,仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B这三种情况。另外,本文中字符“/”,一般表示前后关联对象是一种“或”的关系。此外,本文中的“多”表示两个或者多于两个。

[0020] 请参阅图1,图1是本申请语音交互方法一实施例的流程示意图。本实施例中,该方法由语音交互设备执行,该语音交互设备可以为具有处理能力的任意设备,例如为智能家居设备、计算机、平板电脑、手机等。该智能家居设备可以为空调、电饭煲、微波炉、电冰箱、音箱等家居电子设备。具体地,该方法包括以下步骤:

[0021] S110:获取待响应语音数据。

[0022] 具体地,语音交互设备设有语音采集电路,该语音采集电路具体可包括麦克风等声音采集器以及用于对声音采集器采集到的电信号进行滤波和放大处理的滤波放大电路。语音交互设备通过该语音采集电路检测到所处环境中的语音数据,以作为待响应语音数据。以语音交互设备为智能空调为例,智能空调可通过该语音采集电路检测到用户发出的语音数据,并将该语音数据作为待响应语音数据,进而执行下述S120以实现与用户之间的交互。

[0023] 可以理解的是,为降低语音交互设备的电路复杂度,进而减少语音交互设备的体积,语音交互设备也可从其他语音采集设备处获得待响应语音数据。例如,该语音交互设备与语音采集设备可通过USB、蓝牙等方式连接,语音采集设备在检测到所处环境中的语音数据后,通过USB数据线传输或蓝牙传输等通信方式将该检测到的语音数据作为待响应语音数据传输至语音交互设备,使得语音交互设备获得待响应语音数据。

[0024] S120:利用非网络端数据对所述语音数据进行自然语言处理。

[0025] 其中,所述非网络端数据为不是通过互联网、局域网、移动通信网络等通信网络获得的远程数据,该非网络端数据可以包括语音交互设备的本地数据、与语音交互设备进行电连接的处理设备(如通过USB接口、总线接口等输入输出接口方式与语音交互设备电连接的处理设备)存储的数据。

[0026] 本实施例利用非网络端数据对待响应语音数据进行自然语言处理,即在实现对待响应语音数据进行自然语言处理的过程中无需依赖任意网络端数据,该过程仅需依赖语音交互设备的非网络端数据仅可实现。由此,语音交互设备即使在离线情况下也可自行进行待响应语音数据的处理,无需受到网络情况的影响。该自然语言处理过程可包括待响应语

音数据进行识别得到文本数据,进而对文本数据进行语义理解,并根据语义结果产生对应的响应数据。

[0027] S130:基于所述自然语言处理的结果进行响应。

[0028] 例如,S120对待响应语音数据进行自语言处理,得到对应响应数据,如执行指令或用于告知用户的回复数据。语言交互设备执行该执行指令,或者语言播放或显示该回复数据,以响应用户输入的待响应语言数据,实现与用户之间的交互。

[0029] 可以理解的是,为实现语音交互,需要对待响应语音数据进行自然语言处理,然后基于自然语言处理结果进行响应。而实现自然语言处理需要较大的处理能力,故本申请语音交互设备配置有能够不依赖网络端数据即可实现自然语言处理的处理能力,进而可实现本申请的可离线式的语音交互。

[0030] 本实施例利用非网络端数据对待响应语音数据进行自然语言处理,该处理过程无需依赖网络端数据,故可实现离线式的语音交互,降低甚至避免了网络状况对设备语音交互的影响。而且,由于网络端的数据交互需要耗费一定时间,故相比于依赖网络端数据,语音交互设备直接依赖非网络端数据对待响应语音数据进行自然语言处理,进而基于该自然语言处理的结果进行响应,可加快语言交互的响应速度。

[0031] 下面,对本申请利用非网络端数据所实现的对待响应语音数据进行自然语言处理的过程进行具体举例说明。

[0032] 请结合参阅图2和图3b,图2是本申请语音交互方法另一实施例中S120的流程示意图,图3b是本申请语音交互设备一实施例的结构示意图。本实施例中,上述S120中的对所述待响应语音数据进行自然语言处理,具体包括以下子步骤:

[0033] S221:将待响应语音数据转换成文本数据。

[0034] 为了实现后续的自然语言处理,语音交互设备先将S110获取的待响应语音数据通过语音识别相关技术进行转换,以得到相应的文本数据。

[0035] 具体地,语音交互设备可以基于本地词汇库中的词汇将待响应语音数据转换成文本数据。例如,语音交互设备在本地存在有声学模型、语言模型和发音词典,其中,声学模型是对声学、语音学、环境的变量、说话人性别、口音等的差异的知识表示。语言模型是对一组字序列构成的知识表示。发音字典(lexicon)包含了从单词(words)到音素(phones)之间的映射,作用是用来连接声学模型和语言模型的。上述预存的声学模型、语言模型和发音词典中的至少一者包括有若干词汇(例如发音词典包括若干词语-拼音的映射),其中,该用于存储词汇的数据库或数据文件统称为本地词汇库。具体地,语言交互设备先将待响应语音数据与声学模型进行模式匹配,再结合发音词典和语言模型最终识别得到的文本数据。

[0036] 由于语音交互设备中预存的词汇量的多少会影响该语音识别的准确度,例如,若语音交互设备未预存待响应语音数据中某一词汇,则语音交互设备无法对该语音词汇进行正确的文本转换。因此,理论上,语音交互设备预存的词汇量越多,其语音识别体验越好。但另一方面,语音交互设备预存的词汇量越多,对语音交互设备的处理能力的要求越高,也即语音交互设备可能需要配置更高处理能力的处理器。故,可综合考虑对语音识别的准确度要求和语音交互设备所能支持的处理能力,确定语音交互设备预存的词汇量。在一应用场景中,可为语音交互设备的本地词汇库预存中等规模词汇量,例如,该本地词汇库中的词汇数量超过1000且少于10000,如为1000个、2000个、5000个或10000个等。

[0037] 另外,本地词汇库中包括以下至少一者:设定领域的词汇和通用词汇。由此,语音交互设备可相应识别设定领域和通用领域的语音数据。其中,设定领域,也可称为垂直领域,指预先设定的具有一定行业背景的范围,设定领域的词汇为与该领域关联性较强的词汇,例如该设定领域为家电领域,或者进一步为家电的空调领域、微波炉领域等,空调领域的词汇包括打开、空调、制冷、制暖等。通用领域是相对垂直领域而言,其并不仅针对某一行业领域,而是涉及多行业领域或者多行业领域可通用的,其通用词汇为多行业领域均可能使用的词汇,例如一些口语词:今天、明天、请、帮我、一下、天气、如何等。

[0038] 在一应用场景中,考虑词汇量决定语音交互设备所需的处理能力,可将语音交互设备限定为针对设定领域的语音交互,进而降低所需预存的词汇量。例如,语音交互设备的本地词汇库中存储有设定领域的词汇以及在该设定领域的交流中可能出现的通用词汇如上述的口语词。例如,语音交互设备为风扇,其设定领域为风扇领域,故其设定领域的词汇包括开关、风速之类,同时由于口语的随机性,也可能涉及很多无意义的辅助词汇,例如:“帮我打开风扇”,“给我打开风扇”,“请打开风扇一下”,这些句子中的“帮我”、“给我”、“请”、“一下”,这些通用词汇对于理解没有任何意义,但对于语音识别来说,必须得涵盖这类的通用词语。通常,对于一个设定领域的交互来说,其说法词汇往往是有限的,一般汉语词汇量到一定量如2000多个就能基本畅通交流。因此,语音交互设备只需预存其设定领域的词汇以及相关的一些通用词汇达到一定词汇量,即可实现设定领域的交互。

[0039] 本实施例中,该语音转换成文本的步骤由语音交互设备内的自动语音识别(Automatic Speech Recognition,ASR)模块31实现。该ASR模块31可存储有上述声学模型、发音词典、语言模型等本地数据文件,用于将语言数据转换成计算机能识别的符号序列——文本。

[0040] 可以理解的是,在其他实施例中,上述用于语音识别的声学模型、发音词典、语言模型等本地数据文件也可不存储于语音交互设备本地中,而是存储于与语音交互设备电连接的其他处理设备中。

[0041] S222:判断文本数据是否属于设定领域的内容。若是,则执行S223,否则结束流程。

[0042] 具体地,语音交互设备对经S110获得的文本数据进行分析以确定文本数据涉及的内容是否属于设定领域。本实施例中,语音交互设备在本地预设有分类模型32,本S222可具体包括:利用本地分类模型32对所述文本数据进行分析,并基于分析结果确定所述文本数据是否属于设定领域的内容。例如,语音交互设备预存有利用深度学习等算法建立的分类模型,ARS模块31将转换得到的文本数据输入至该分类模型32进行处理,得到文本数据属于设定领域的可能性结果,该可能性结果可以包括文本数据属于预设领域时的执行度和/或得分。语音交互设备将得到的属于设定领域的可能性结果与预设阈值进行比较,若可能性结果大于该预设阈值,则可确定文本数据属于设定领域的内容。

[0043] 可以理解的是,在其他实施例中,语音交互设备并不限定在设定领域的语音交互时,也可不执行S222,也即语音交互设备执行S221之后无需判断直接执行S223。

[0044] S223:对文本数据进行自然语言处理。

[0045] 具体地,语音交互设备在判断待处理自然语言属于设定领域的内容的情况下,响应于该判断结果而对所述文本数据进行自然语言处理。其中,语音交互设备可采用现有的自然语言处理方式对文本数据进行处理,以得到与该文本数据匹配的响应数据,该响应数

据例如为执行指令,或者用于告知用户的回复数据。

[0046] 本实施例中,如图3a所示,该步骤S223具体包括以下子步骤:

[0047] S223a:对文本数据进行语义理解,得到语义结果。

[0048] 具体地,先对待响应语音数据对应的文本数据进行自然语言理解,也即将文本数据转换成计算机能够理解的槽位信息,利用该槽位信息即可得到文本数据传达的语义信息。本实施例结合图3b进行举例说明,语音交互设备可设有自然语言理解(Natural Language Understanding,NLU)模块33,该子步骤S223a可以由该NLU模块33实现。例如,NLU模块33接收到分类模型32输出的分类结果,并在分类模型32输出的分配结果为文本数据属于设定领域内容时,执行对文本数据进行语义理解。

[0049] 在一实施例中,子步骤S223a可包括:对所述文本数据进行槽位填充,得到槽位信息,其中所述槽位信息用于表示所述文本数据的语义结果。槽位(也称为语义槽,slot)是指语义理解的表示,通常由一些参数对描述,例如:待响应语音数据对应的文本数据为:“今天深圳天气怎么样”,其有以下槽位信息:{领域:天气;时间:今天;地址:深圳}。其中,针对语音交互设备针对的领域不同,且槽位可设为不同。例如,语音交互设备针对的设定领域为家电领域,其槽位信息可包括:设备控制、厨房、场景等。

[0050] 本实施例中,语音交互设备先过滤到非设定领域的文本数据,故只需对设定领域的文本数据进行语义理解,由此可降低语音交互设备实现语义理解的处理能力需求。

[0051] S223b:基于所述语义结果生成第一响应数据。

[0052] 具体地,该第一响应数据可以为产生匹配的执行指令并执行该执行指令或者产生用于交互的回复数据。其中,语音交互设备可设有对话管理(Dialog Management,DM)模块34(也可称为应用逻辑模块34)、自然语言生成(Natural Language Generation,NLG)模块35和文本到语音合成(Text To Speech,TTS)模块36。该子步骤S223b的执行可以由DM模块34、NLG模块35和TTS模块36的组合或部分模块实现。

[0053] 在一应用场景中,语音交互设备执行该S223b包括:生成与所述语义结果匹配的执行指令;对应地,语音交互设备在执行上述S130时即具体包括执行所述执行指令;和/或,生成与所述语义结果匹配的回复数据,其中,所述回复数据可以为文本回复数据或语音回复数据;例如,所述生成与所述语义结果匹配的文本回复数据,包括:确定与所述语义结果匹配的回复内容;对所述回复内容进行自然语言生成,得到文本回复数据;所述生成与所述语义结果匹配的语音回复数据,包括:确定与所述语义结果匹配的回复内容;对所述回复内容进行自然语言生成,得到文本回复数据;将所述文本回复数据转换成语音回复数据。对应地,语音交互设备在执行上述S130时即具体包括:显示所述文本回复数据或语音播放所述语音回复数据。

[0054] 其中,该执行指令可以由DM模块34直接生成,并输出至语音交互设备的执行单元37执行该执行指令,该执行单元37例如为风扇的扇叶驱动器、空调的压缩机电机等。该回复数据可先由DM模块34确定要回复的内容,并利用NLG模块35将要回复的内容转换成生成文本类型的回复数据,若采用显示方式进行回复,则将该文本类型的回复数据输出至语音交互设备的显示单元(图3b未示),以由显示单元显示该回复数据;若采用语音播放方式进行回复,则由NLG模块35将该文本类型的回复数据输出至TTS模块36,以由TTS模块将文本类型的回复数据进行文本语音转换,得到语音类型的回复数据,进而将该语音类型的回复数据

输出至语音交互设备的声音播放单元38来语音播放该回复数据。可以理解的是的,语音交互设备可既产生执行指令也产生回复数据。例如,空调接收到用户发出的语音数据“帮我打开下空调”,空调对其进行自然语言处理并生成空调打开指令和回复数据“好的,已为您打开空调”,空调响应于该空调打开指令而启动自身,并通过语音播报该回复数据“好的,已为您打开空调”。

[0055] 可以理解的是,图3b所述的语音交互设备所包括的模块或单元可以为不同的程序部分,也可为由不同电路实现的功能模组。

[0056] 本实施例,语音交互设备过滤掉对非设定领域的语音数据,只进行设定领域的语音数据进行自然语言处理,可降低语音交互设备实现语音识别及其他自然语言处理过程中所需要的处理能力,进而降低语音交互设备的实现难度和成本。而且,由于设定领域的相关词汇量有限,故在符合语音交互设备的处理能力的词汇量范围内,可增加较多的相关通用词汇,进而可对该设定领域的自然语言进行较为可靠准确的理解,故可一定程度上实现真正的自然语言交互。

[0057] 请参阅图4a,图4a是本申请语音交互方法再一实施例的流程示意图。本实施例中,该方法由语音交互设备执行,该语音交互设备可以为具有处理能力的任意设备,例如为智能家居设备、计算机、平板电脑、手机等具体地,该方法包括以下步骤:

[0058] S410:获取待响应语音数据。

[0059] S420:利用非网络端数据对所述待响应语音数据进行自然语言处理。

[0060] S430:基于利用非网络端数据得到的自然语言处理的结果进行响应。

[0061] 其中,步骤S410-S430的说明请参考上面实施例的相关描述,在此不做赘述。

[0062] S440:判断利用非网络端数据是否能够对待响应语音数据进行自然语言处理。若是,则执行S450,否则,可结束流程,以并接收到新的待响应数据时,继续执行本申请语音交互方法。

[0063] 本实施例中,S420步骤是依赖非网络端数据对所述待响应语音数据进行自然语言处理,如上面实施例所述可知,待响应语音数据的自然语言处理的实现需要依靠语音交互设备中预先存储的一些相关词汇或其他数据,由于语音交互设备的处理能力有限,故其预存的相关数据也是有限,故可能存在利用语音交互设备的本地预存数据无法对语音数据进行语音识别和后续自然语言处理的情况。故,语音交互设备在执行S420后,监测自身是否已基于待响应语音数据进行自然语言处理并基于处理结果进行响应,例如,检测在预设时间内是否有进行对所述待响应语音数据的响应动作(如执行指令或语音播放回复数据)。若在预设时间内未检测到对所述待响应语音数据的响应动作,则确定为检测到利用非网络端数据无法对所述待响应语音数据进行自然语言处理,也即目前无法实现离线式语音交互,此时,语音交互设备则考虑依赖网络端对所述待响应语音数据进行自然语言处理。

[0064] S450:依赖网络端数据对所述待响应语音数据进行自然语言处理。

[0065] 语音交互设备可先检测当前是否可实现网络通信,若不可以,则发起网络连接请求,以请求建立通信链路以进行后续的数据交互。在确定可实现网络通信后,语音交互设备依赖网络端数据来对所述待响应语音数据进行自然语言处理。例如,如图4b所示,语音交互系统包括相互连接的语音交互设备41和网络端设备42,所述语音交互设备41为如上面所述的语音交互设备,网络端设备42可以为远程服务器等,用于对接收到的语音数据进行语音

识别以及自然语言处理,以产生第二响应数据。具体地,语音交互设备41将所述待响应语音数据发送给网络端设备42;语音交互设备41接收所述网络端设备42反馈的第二响应数据;其中,所述第二响应数据为所述网络端设备对所述待响应语音数据进行自然语言处理得到,例如上述的执行指令或回复数据。其中,网络端设备基于所述待响应语音数据进行自然语言处理的过程可参考语音交互设备自身基于所述待响应语音数据进行自然语言处理的过程,故在此不做赘述。

[0066] S460:基于依赖网络端数据得到的自然语言处理的结果进行响应。

[0067] 例如,语音交互设备在接收网络端反馈的第二响应数据后,该语音交互设备41执行该第二响应数据,例如执行对应指令或者语音播放回复数据等。具体,可参考上述S130步骤的相关描述。

[0068] 本实施例中,在离线式语音交互无法实现时,语音交互设备可进一步依赖网络端数据来实现基于所述待响应语音数据进行自然语言处理并进行响应,也即通过线上实现语音交互,由此采用离线和线上的结合方式,可保证语音交互的可靠性。

[0069] 请参阅图5,图5是本申请语音交互设备一实施例的结构示意图。本实施例中,该语音交互设备50包括存储器51、处理器52和通信电路53。其中,通信电路53、存储器51分别耦接处理器52。具体地,语音交互设备50的各个组件可通过总线耦合在一起,或者语音交互设备50的处理器分别与其他组件一一连接。该语音交互设备50可以为计算机、手机、智能家居设备等具有一定处理能力的设备。其中,智能家居设备可以为空调、电饭煲、微波炉、电冰箱、音箱等家居电子设备。

[0070] 通信电路53用于与其他设备通信。例如,当语音交互设备需要线上提供的自然语言处理功能时,通信电路53可与网络端设备通信。

[0071] 存储器51用于存储处理器52执行的程序指令以及处理器52在处理过程中的数据如非网络端数据,其中,该存储器51包括非易失性存储部分,用于存储上述程序指令。

[0072] 处理器52控制语音交互设备50的操作,处理器52还可以称为CPU (Central Processing Unit,中央处理单元)。处理器52可能是一种集成电路芯片,具有信号的处理能力。处理器52还可以是通用处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现成可编程门阵列(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。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。另外,处理器52可以由多个成电路芯片共同实现。

[0073] 在本实施例中,处理器52通过调用存储器51存储的程序指令,用于执行上述任一方法实施例方法的步骤。

[0074] 例如,处理器52用于获取待响应语音数据;利用非网络端数据对所述待响应语音数据进行自然语言处理;基于所述自然语言处理的结果进行响应(例如控制上述的执行单元和语音播放单元等部件执行与上述自然语言处理结果匹配的操作)。

[0075] 在一些实施例中,处理器52执行所述对所述待响应语音数据进行自然语言处理包括:将所述待响应语音数据转换成文本数据;对所述文本数据进行自然语言处理。

[0076] 进一步地,存储器51存储有词汇库,处理器52执行所述将所述待响应语音数据转换成文本数据,可包括:基于本地词汇库中的词汇将所述待响应语音数据转换成文本数据。

[0077] 其中,所述本地词汇库中的词汇数量可以为超过1000且少于10000。

[0078] 其中,所述本地词汇库中可以包括以下至少一者:设定领域的词汇和通用词汇。

[0079] 在一些实施例中,处理器52在执行所述将所述待响应语音数据转换成文本数据之后,还用于:判断所述文本数据是否属于设定领域的内容;处理器52执行所述对所述文本数据进行自然语言处理,包括:响应于所述文本数据属于所述设定领域的内容,对所述文本数据进行自然语言处理。

[0080] 在一些实施例中,存储器51存储有分类模型,处理器52执行所述判断所述文本数据是否属于设定领域的内容,包括:利用本地分类模型对所述文本数据进行分析,并基于分析结果确定所述文本数据是否属于设定领域的内容。

[0081] 在一些实施例中,处理器52执行所述对所述文本数据进行自然语言处理,包括:对所述文本数据进行语义理解,得到语义结果;基于所述语义结果生成第一响应数据。

[0082] 在一些实施例中,处理器52执行所述基于所述语义结果生成第一响应数据,可包括:生成与所述语义结果匹配的执行指令;所述基于所述自然语言处理的结果进行响应,包括:执行所述执行指令。

[0083] 在一些实施例中,所述基于所述语义结果生成第一响应数据,包括:生成与所述语义结果匹配的回复数据,其中,所述回复数据为文本回复数据或语音回复数据;所述基于所述自然语言处理的结果进行响应,包括:显示所述文本回复数据或语音播放所述语音回复数据。

[0084] 其中,处理器52执行所述生成与所述语义结果匹配的语音回复数据,可包括:确定与所述语义结果匹配的回复内容;对所述回复内容进行自然语言生成,得到文本回复数据;将所述文本回复数据转换成语音回复数据。

[0085] 在一些实施例中,处理器52还用于判断利用非网络端数据是否能够对所述待响应语言数据进行自然语言处理;响应于利用非网络端数据能够对所述待响应语言数据进行自然语言处理,依赖网络端数据对所述待响应语音数据进行自然语言处理。

[0086] 进一步地,处理器52执行所述判断利用非网络端数据是否能够对所述待响应语言数据进行自然语言处理,可包括:检测在预设时间内是否有进行对所述待响应语音数据的响应动作;处理器52执行所述依赖网络端数据对所述待响应语音数据进行自然语言处理,可包括:通过通信电路53将所述待响应语音数据发送给网络端设备;通过通信电路53接收所述网络端设备反馈的第二响应数据,其中,所述第二响应数据为所述网络端设备对所述待响应语音数据进行自然语言处理得到。

[0087] 可以理解的是,在另一实施例中,若语音交互设备无需与网络端进行通信,则该语音交互设备可不包括上述通信电路53。

[0088] 请参阅图6,本申请还提供一种存储装置的实施例。本实施例中,该存储装置60存储有处理器可运行的程序指令61,该程序指令61用于执行上述实施例中的方法。

[0089] 该存储装置60具体可以为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)、磁碟或者光盘等可以存储程序指令的介质,或者也可以为存储有该程序指令的服务器,该服务器可将存储的程序指令发送给其他设备运行,或者也可以自运行该存储的程序指令。

[0090] 在一些实施例中,存储装置60还可以为如图5所示的存储器。

[0091] 上述方案中,利用非网络端数据对待响应语音数据的自然语言处理,该处理过程

无需依赖网络端数据,故可实现离线式的语音交互,降低甚至避免了网络状况对设备语音交互的影响。而且,由于网络端的数据交互需要耗费一定时间,故相比于依赖网络端数据,设备直接依赖非网络端数据对语音数据进行自然语言处理,进而基于该自然语言处理的结果进行响应,可加快语言交互的响应速度。

[0092] 在本申请所提供的几个实施例中,应该理解到,所揭露的方法和装置,可以通过其它的方式实现。例如,以上所描述的装置实施方式仅仅是示意性的,例如,模块或单元的划分,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。

[0093] 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方式方案的目的。

[0094] 另外,在本申请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,也可以采用软件功能单元的形式实现。

[0095] 集成的单元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,可以存储在一个计算机可读取存储介质中。基于这样的理解,本申请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全部或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)或处理器(processor)执行本申请各个实施方式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。

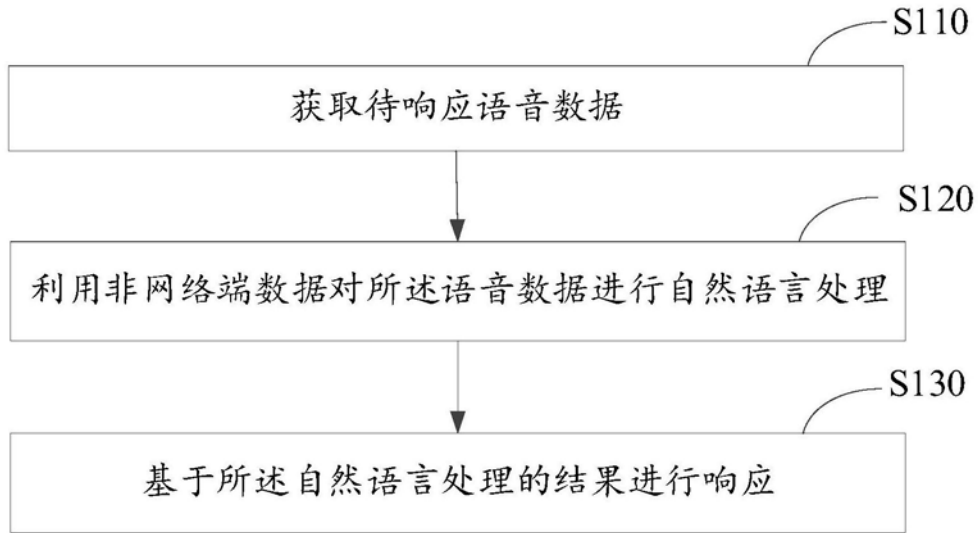


图1

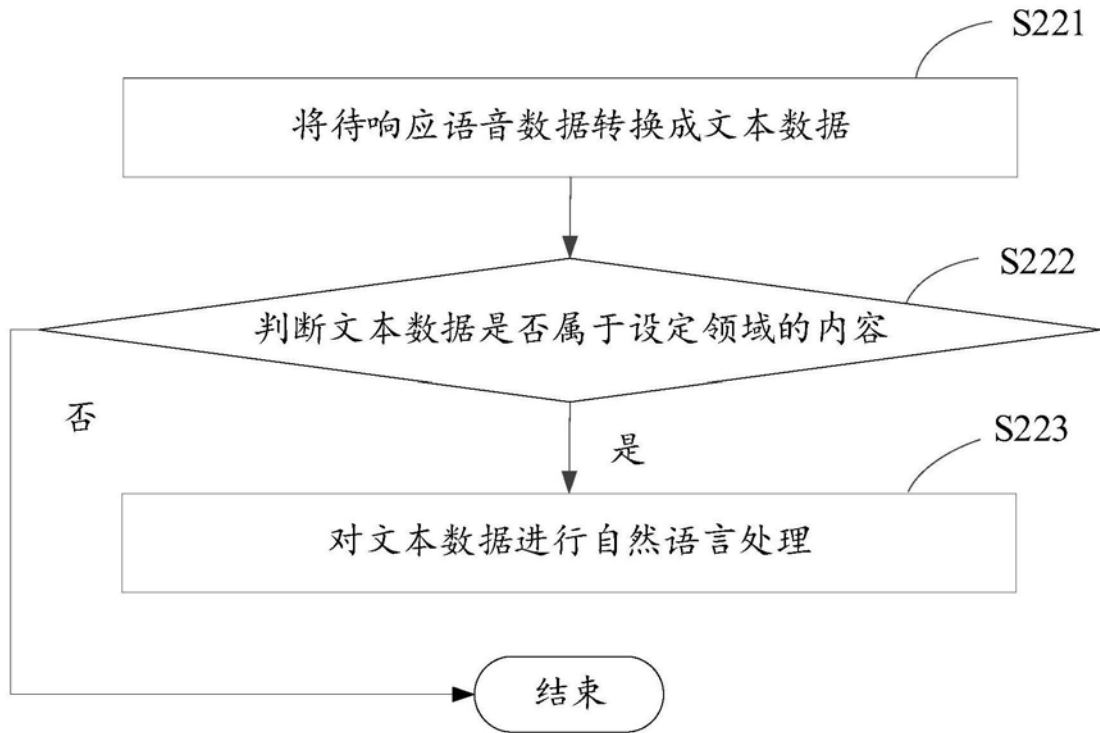


图2

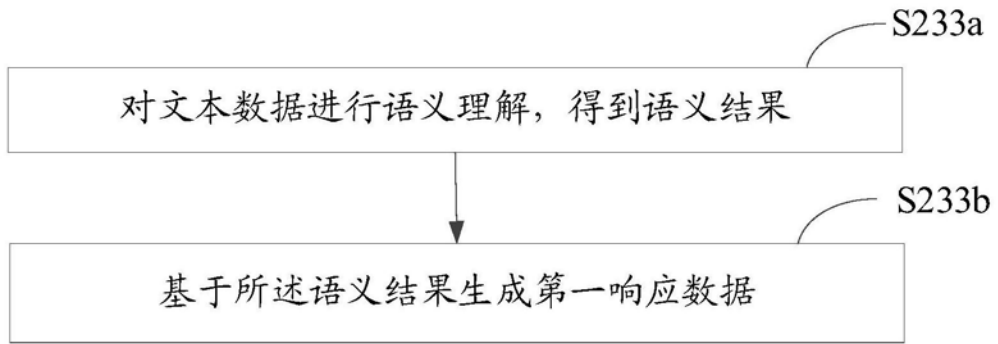


图3a

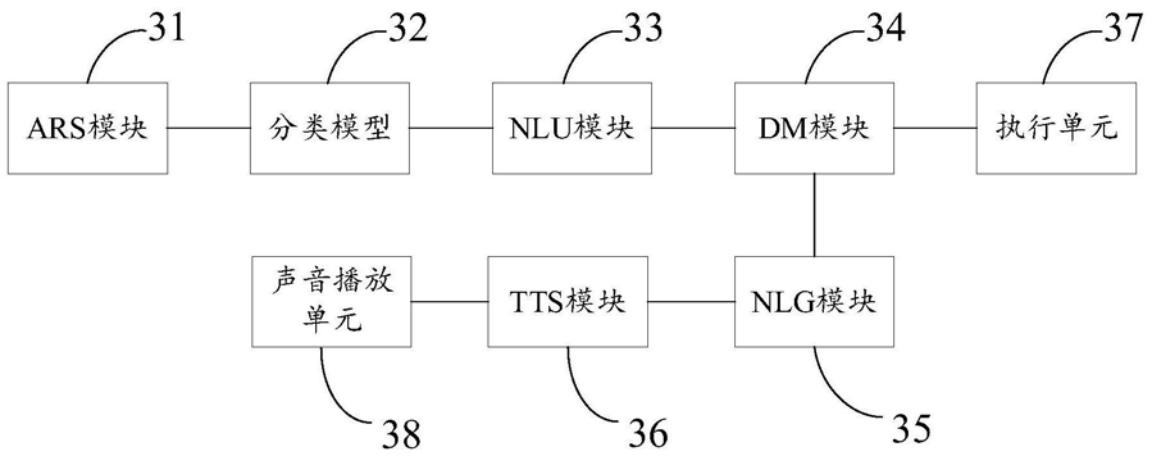


图3b

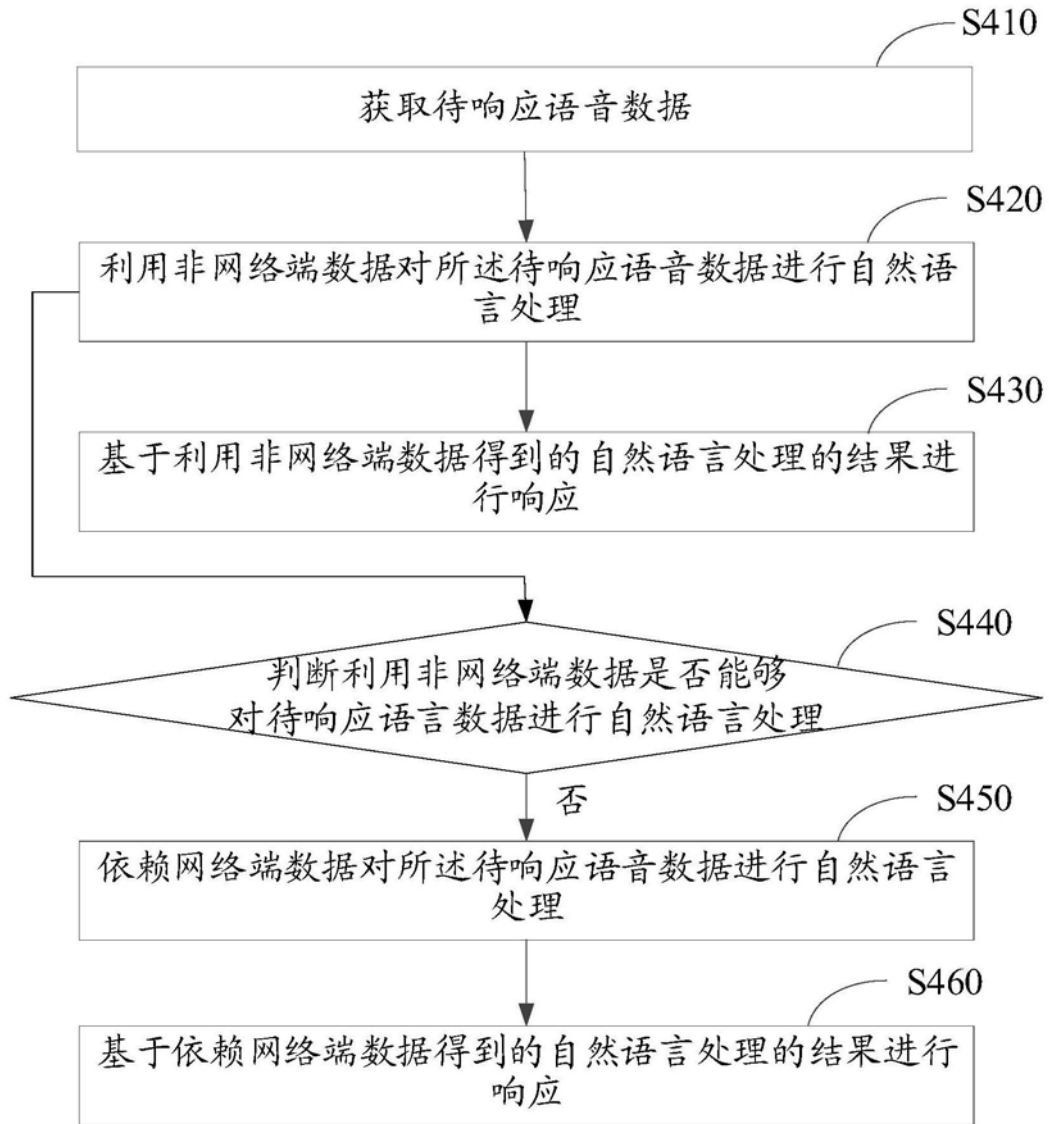


图4a

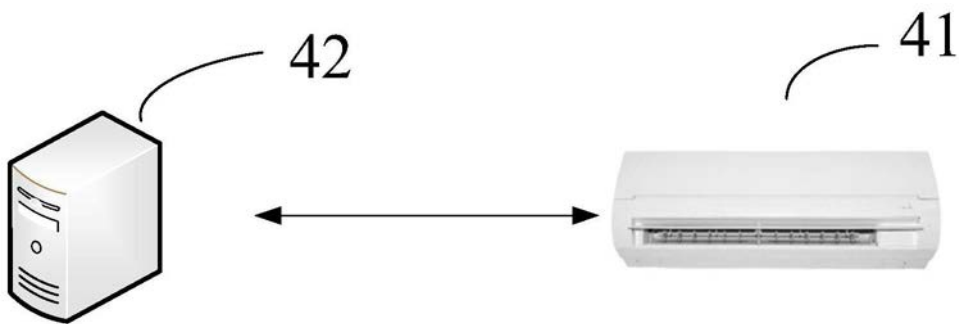


图4b

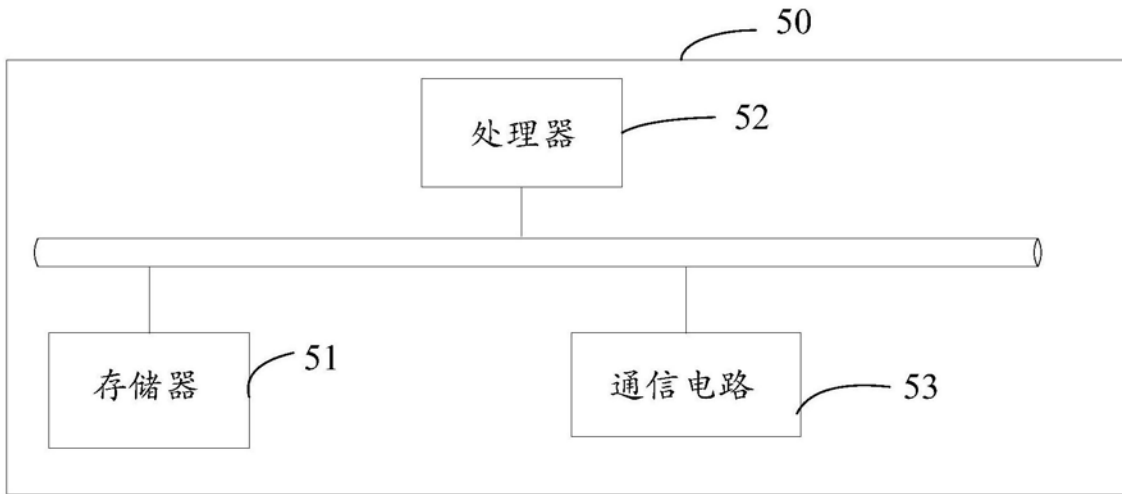


图5

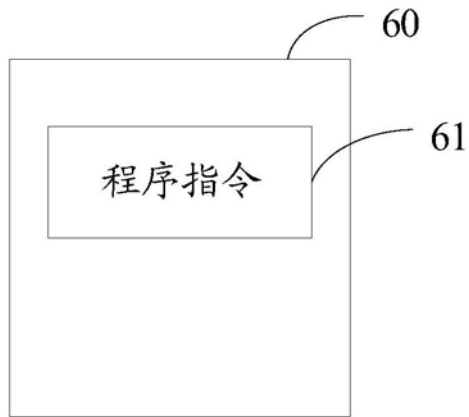


图6