



(12) 发明专利

(10) 授权公告号 CN 114078478 B

(45) 授权公告日 2022. 09. 23

(21) 申请号 202111339469.4

G10L 15/22 (2006.01)

(22) 申请日 2021.11.12

G10L 15/16 (2006.01)

G10L 15/06 (2013.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 114078478 A

(56) 对比文件

CN 114187903 A, 2022.03.15

CN 112053687 A, 2020.12.08

(43) 申请公布日 2022.02.22

(73) 专利权人 北京百度网讯科技有限公司

地址 100085 北京市海淀区上地十街10号

百度大厦2层

审查员 张彩

(72) 发明人 吴震 王潇 苏显泽 瞿琴

吴玉芳

(74) 专利代理机构 北京鸿德海业知识产权代理

有限公司 11412

专利代理师 田宏宾

(51) Int. Cl.

G10L 15/26 (2006.01)

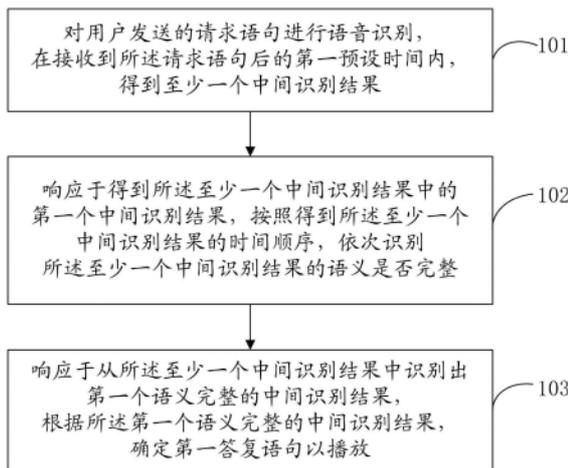
权利要求书4页 说明书13页 附图5页

(54) 发明名称

语音交互的方法、装置、电子设备及存储介质

(57) 摘要

本公开提供了语音交互的方法、装置、电子设备及可读存储介质,涉及计算机技术领域,具体涉及语音技术、自然语言处理等人工智能技术领域。一种具体实现方案为:对用户发送的请求语句进行语音识别,在接收到请求语句后的第一预设时间内,得到至少一个中间识别结果;其中,第一预设时间的结束时刻,早于请求语句的语音活动检测的尾点时刻;响应于得到第一个中间识别结果,按照得到至少一个中间识别结果的时间顺序,依次识别至少一个中间识别结果的语义是否完整;响应于从至少一个中间识别结果中识别出第一个语义完整的中间识别结果,根据第一个语义完整的中间识别结果,确定第一答复语句以播放。



1. 一种语音交互的方法,包括:

对用户发送的请求语句进行语音识别,在接收到所述请求语句后的第一预设时间内,得到至少一个中间识别结果;其中,所述第一预设时间的结束时刻,早于所述请求语句的语音活动检测的尾点时刻;

响应于得到所述至少一个中间识别结果中的第一个中间识别结果,按照得到所述至少一个中间识别结果的时间顺序,依次识别所述至少一个中间识别结果的语义是否完整;

响应于从所述至少一个中间识别结果中识别出第一个语义完整的中间识别结果,根据所述第一个语义完整的中间识别结果,确定第一答复语句以播放;

其中,所述依次识别所述至少一个中间识别结果的语义是否完整,包括:依次针对所述至少一个中间识别结果中的各中间识别结果,

基于所述各中间识别结果作为历史最终识别结果中前缀的第一概率和作为所述历史最终识别结果的第二概率,确定所述各中间识别结果语义完整的第三概率;根据所述第三概率是否大于预设阈值,确定所述各中间识别结果的语义是否完整;或者,

将所述各中间识别结果的词向量和热度、以及所述各中间识别结果作为历史最终识别结果中前缀的第一概率和作为所述历史最终识别结果的第二概率输入神经网络模型,经所述神经网络模型输出所述各中间识别结果语义完整的第四概率;根据所述第四概率是否大于预设阈值,确定所述各中间识别结果的语义是否完整;

其中,所述历史最终识别结果中前缀,为所述历史最终识别结果中前面一部分内容、而非完整的所述历史最终识别结果。

2. 根据权利要求1所述的方法,其中,所述依次识别所述至少一个中间识别结果的语义是否完整,还包括:

依次针对所述至少一个中间识别结果中的各中间识别结果,利用语义完整性模型,获取所述各中间识别结果作为历史最终识别结果中前缀的第一概率,以及所述各中间识别结果作为所述历史最终识别结果的第二概率;所述语义完整性模型基于用户日志中的最终识别结果进行统计计算得到。

3. 根据权利要求1所述的方法,其中,所述依次识别所述至少一个中间识别结果的语义是否完整,还包括:

依次针对所述至少一个中间识别结果中的各中间识别结果,获取所述各中间识别结果的词向量;

获取所述各中间识别结果作为历史最终识别结果中前缀的第一概率,以及所述各中间识别结果作为所述历史最终识别结果的第二概率;

获取所述各中间识别结果的热度。

4. 根据权利要求3所述的方法,其中,所述获取所述各中间识别结果作为历史最终识别结果中前缀的第一概率,以及所述各中间识别结果作为所述历史最终识别结果的第二概率,包括:

利用语义完整性模型,获取所述各中间识别结果作为历史最终识别结果中前缀的第一概率,以及所述各中间识别结果作为所述历史最终识别结果的第二概率;所述语义完整性模型基于用户日志中的最终识别结果进行统计计算得到。

5. 根据权利要求3或4所述的方法,其中,所述神经网络模型的训练包括:

基于用户日志中的最终识别结果构造训练集;所述训练集包括多个训练样本,所述多个训练样本包括正样本和负样本,所述正样本具有语义完整的标注信息,所述负样本具有语义不完整的标注信息;

分别针对所述多个训练样本中各训练样本,获取所述各训练样本作为历史最终识别结果中前缀的第一概率,以及所述各训练样本作为所述历史最终识别结果的第二概率;

获取所述各训练样本的热度;

将所述各训练样本的所述词向量、所述第一概率、所述第二概率和所述热度输入待训练的神经网络模型,经所述待训练的神经网络模型输出所述各训练样本语义完整的预测概率;

根据所述预测概率是否大于预设阈值,确定所述各训练样本的语义是否完整,得到预测结果;

基于所述各训练样本的预测结果与对应的标注信息之间的差异,对所述待训练的神经网络模型进行训练,直至满足预设训练完成条件。

6. 根据权利要求1-4中任一项所述的方法,还包括:

在接收到所述请求语句后的第二预设时间内,得到所述请求语句的最终识别结果,其中,所述第二预设时间的结束时刻,晚于所述请求语句的语音活动检测的尾点时刻;

响应于从所述至少一个中间识别结果中未识别出语义完整的中间识别结果,根据所述最终识别结果,确定第二答复语句以播放。

7. 一种语音交互的装置,包括:

语音识别单元,用于对用户发送的请求语句进行语音识别,在接收到所述请求语句后的第一预设时间内,得到至少一个中间识别结果;其中,所述第一预设时间的结束时刻,早于所述请求语句的语音活动检测的尾点时刻;

语义完整识别单元,用于响应于得到所述至少一个中间识别结果中的第一个中间识别结果,按照得到所述至少一个中间识别结果的时间顺序,依次识别所述至少一个中间识别结果中各中间识别结果的语义是否完整;

第一确定单元,用于响应于从所述至少一个中间识别结果中识别出第一个语义完整的中间识别结果,根据所述第一个语义完整的中间识别结果,确定第一答复语句以播放;

其中,所述语义完整识别单元包括:

第一确定模块,用于依次针对所述至少一个中间识别结果中的各中间识别结果,基于所述各中间识别结果作为历史最终识别结果中前缀的第一概率和作为所述历史最终识别结果的第二概率,确定所述各中间识别结果语义完整的第三概率;第二确定模块,用于根据所述第三概率是否大于预设阈值,确定所述各中间识别结果的语义是否完整;或者,

神经网络模型,用于依次针对所述至少一个中间识别结果中的各中间识别结果,接收输入的所述各中间识别结果的词向量和热度、以及所述各中间识别结果作为历史最终识别结果中前缀的第一概率和作为所述历史最终识别结果的第二概率,输出所述各中间识别结果语义完整的第四概率;第三确定模块,用于根据所述第四概率是否大于预设阈值,确定所述各中间识别结果的语义是否完整;

其中,所述历史最终识别结果中前缀,为所述历史最终识别结果中前面一部分内容、而非完整的所述历史最终识别结果。

8. 根据权利要求7所述的装置,其中,所述语义完整识别单元还包括:

语义完整性模型,用于依次针对所述各中间识别结果,获取所述各中间识别结果作为历史最终识别结果中前缀的第一概率,以及所述各中间识别结果作为所述历史最终识别结果的第二概率;所述语义完整性模型基于用户日志中的最终识别结果进行统计计算得到。

9. 根据权利要求7所述的装置,其中,所述语义完整识别单元还包括:

第一获取模块,用于依次针对所述各中间识别结果,获取所述各中间识别结果的词向量;

第二获取模块,用于获取所述各中间识别结果作为历史最终识别结果中前缀的第一概率,以及所述各中间识别结果作为所述历史最终识别结果的第二概率;

第三获取模块,用于获取所述各中间识别结果的热度。

10. 根据权利要求9所述的装置,其中,所述第二获取模块具体用于:利用语义完整性模型,获取所述各中间识别结果作为历史最终识别结果中前缀的第一概率,以及所述各中间识别结果作为所述历史最终识别结果的第二概率;所述语义完整性模型基于用户日志中的最终识别结果进行统计计算得到。

11. 根据权利要求9或10所述的装置,还包括:

构造单元,用于基于用户日志中的最终识别结果构造训练集;所述训练集包括多个训练样本,所述多个训练样本包括正样本和负样本,所述正样本具有语义完整的标注信息,所述负样本具有语义不完整的标注信息;

第一获取单元,用于分别针对所述多个训练样本中各训练样本,获取所述各训练样本作为历史最终识别结果中前缀的第一概率,以及所述各训练样本作为所述历史最终识别结果的第二概率;

第二获取单元,用于获取所述各训练样本的热度;

待训练的神经网络模型,用于接收输入的所述各训练样本的所述词向量、所述第一概率、所述第二概率和所述热度,输出所述各训练样本语义完整的预测概率;

第二确定单元,用于根据所述预测概率是否大于预设阈值,确定所述各训练样本的语义是否完整,得到预测结果;

训练单元,用于基于所述各训练样本的预测结果与对应的标注信息之间的差异,对所述待训练的神经网络模型进行训练,直至满足预设训练完成条件。

12. 根据权利要求7-10中任一项所述的装置,其中,

所述语音识别单元,还用于在接收到所述请求语句后的第二预设时间内,得到所述请求语句的最终识别结果,其中,所述第二预设时间的结束时刻,晚于所述请求语句的语音活动检测的尾点时刻;

所述第一确定单元,还用于响应于从所述至少一个中间识别结果中未识别出语义完整的中间识别结果,根据所述最终识别结果,确定第二答复语句以播放。

13. 一种电子设备,包括:

至少一个处理器;以及

与所述至少一个处理器通信连接的存储器;其中,

所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行权利要求1-6中任一项所述的方法。

14. 一种存储有计算机指令的非瞬时计算机可读存储介质,其中,所述计算机指令用于使所述计算机执行根据权利要求1-6中任一项所述的方法。

15. 一种计算机程序设备,包括计算机程序,所述计算机程序在被处理器执行时实现根据权利要求1-6中任一项所述的方法。

语音交互的方法、装置、电子设备及存储介质

技术领域

[0001] 本公开涉及计算机技术领域,具体涉及语音技术、自然语言处理等人工智能技术领域,尤其涉及语音交互的方法、装置、电子设备及存储介质。

背景技术

[0002] 随着人工智能技术的进步,人机语音交互(简称:语音交互)也得到了迅速发展和广泛应用,例如,可广泛应用于智能电视、智能音箱、虚拟现实(virtual reality,VR)眼镜等智能设备,以及各种语音助手应用(APP)中。

[0003] 在传统的人机语音交互中,语音识别和后续调用对话服务进行响应是串行进行的,即在语音活动检测(Voice Activity Detection,VAD)的尾点后才将进行语音识别并根据该语音识别结果调用对话服务进行响应。由此导致语音交互的响应时间较长,影响用户体验。针对该问题,现有技术提出了流式提前拉取对话资源的方案,通过提前进行流式语音识别,将VAD过程与后续调用对话服务进行响应的过程并行化,从而减少语音交互的响应时间。

[0004] 本公开的发明人通过研究发现,上述流式提前拉取对话资源的方案,虽然可以减少语音交互的响应时间,但是在流式语音识别的过程中,并不知道用户何时会停止说话,需要在语音识别结果变化时,不停的调用对话服务,获取对应的结果并缓存起来,由此导致对话服务的请求量增加,造成大量对话服务计算资源的浪费,尤其是针对一些需要调用收费资源服务应用程序接口(Application Programming Interface,API)的情况,例如对话服务需要调用天气服务商、音频内容提供商的接口以获取响应资源时,需要按次付费,会进一步增加经济成本。

发明内容

[0005] 本公开提供了一种语音交互的方法、装置、电子设备及存储介质。

[0006] 根据本公开的一方面,提供了一种语音交互的方法,包括:

[0007] 对用户发送的请求语句进行语音识别,在接收到所述请求语句后的第一预设时间内,得到至少一个中间识别结果;其中,所述第一预设时间的结束时刻,早于所述请求语句的语音活动检测的尾点时刻;

[0008] 响应于得到所述至少一个中间识别结果中的第一个中间识别结果,按照得到所述至少一个中间识别结果的时间顺序,依次识别所述至少一个中间识别结果的语义是否完整;

[0009] 响应于从所述至少一个中间识别结果中识别出第一个语义完整的中间识别结果,根据所述第一个语义完整的中间识别结果,确定第一答复语句以播放。

[0010] 根据本公开的另一方面,提供了一种语音交互的装置,包括:

[0011] 语音识别单元,用于对用户发送的请求语句进行语音识别,在接收到所述请求语句后的第一预设时间内,得到至少一个中间识别结果;其中,所述第一预设时间的结束时

刻,早于所述请求语句的语音活动检测的尾点时刻;

[0012] 语义完整识别单元,用于响应于得到所述至少一个中间识别结果中的第一个中间识别结果,按照得到所述至少一个中间识别结果的时间顺序,依次识别所述至少一个中间识别结果的语义是否完整;

[0013] 第一确定单元,用于响应于从所述至少一个中间识别结果中识别出第一个语义完整的中间识别结果,根据所述第一个语义完整的中间识别结果,确定第一答复语句以播放。

[0014] 根据本公开的又一方面,提供了一种电子设备,包括:

[0015] 至少一个处理器;以及

[0016] 与所述至少一个处理器通信连接的存储器;其中,

[0017] 所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行如上所述的方面和任一可能的实现方式的方法。

[0018] 根据本公开的又一方面,提供了一种存储有计算机指令的非瞬时计算机可读存储介质,所述计算机指令用于使所述计算机执行如上所述的方面和任一可能的实现方式的方法。

[0019] 根据本公开的又一方面,提供了一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时实现如上所述的方面和任一可能的实现方式的方法。

[0020] 根据本公开的又一方面,提供了一种人工智能设备,包括如上所述的电子设备。

[0021] 由上述技术方案可知,本公开实施例对用户发送的请求语句进行语音识别,在接收到所述请求语句后的第一预设时间内得到至少一个中间识别结果,在得到所述至少一个中间识别结果中的第一个中间识别结果起,即按照得到所述至少一个中间识别结果的时间顺序,依次识别所述至少一个中间识别结果中的语义是否完整,响应于从所述至少一个中间识别结果中识别出第一个语义完整的中间识别结果,即根据所述第一个语义完整的中间识别结果,确定第一答复语句以播放。由此,本公开实施例通过提前进行流式语音识别,在识别出第一个语义完整的中间识别结果时,才根据所述第一个语义完整的中间识别结果调用对话服务确定第一答复语句,既可以通过提前拉取对话资源减少语音交互的响应时间,又可以避免基于语义不完整的中间识别结果大量调用对话服务进行对话模型的计算和资源拉取,能够大幅度减少对对话服务的无效调用,减少对话服务的请求量,从而节省对话服务的计算资源和存储资源、以及收费资源服务,从而降低成本。

[0022] 应当理解,本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征,也不用于限制本公开的范围。本公开的其它特征将通过以下的说明书而变得容易理解。

附图说明

[0023] 附图用于更好地理解本方案,不构成对本公开的限定。其中:

[0024] 图1是根据本公开第一实施例的示意图;

[0025] 图2是根据本公开第二实施例的示意图;

[0026] 图3是根据本公开第三实施例的示意图;

[0027] 图4是根据本公开第四实施例的示意图;

[0028] 图5是根据本公开实施例的一个语音交互过程的示意图;

- [0029] 图6是根据本公开第五实施例的示意图；
[0030] 图7是根据本公开第六实施例的示意图；
[0031] 图8是用来实现本公开实施例的语音交互的方法的电子设备的框图。

具体实施方式

[0032] 以下结合附图对本公开的示范性实施例做出说明,其中包括本公开实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本公开的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0033] 显然,所描述的实施例是本公开一部分实施例,而不是全部的实施例。基于本公开中的实施例,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全部其他实施例,都属于本公开保护的范围。

[0034] 需要说明的是,本公开实施例中所涉及的终端设备可以包括但不限于手机、个人数字助理(Personal Digital Assistant,PDA)、无线手持设备、平板电脑(Tablet Computer)、车辆上的计算设备等智能设备;显示设备可以包括但不限于个人电脑、电视、车辆上耦合的显示器等具有显示功能的设备。

[0035] 另外,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。

[0036] 现有流式提前拉取对话资源的方案,虽然可以显著减少语音交互的响应时间,但是在流式语音识别的过程中,并不知道用户何时会停止说话,需要在语音识别结果变化时,不停的调用对话服务,例如针对用户的请求语句“我想听稻香”,对话服务会依次收到:

[0037] A.我

[0038] B.我想

[0039] C.我想听

[0040] D.我想听稻

[0041] E.我想听稻香(VAD的尾点时刻前得到的中间识别结果,对应图5中的时刻1.5)

[0042] F.我想听稻香(VAD的尾点时刻后得到的最终识别结果,对应图5中的时刻3)

[0043] 其中,由于中间识别结果A、B、C、D、E都发生了变化,每一个与之前的均不同,均需要调用对话服务,获取对应的结果并缓存起来,导致对话服务的请求量增加,对话服务的请求量约是传统方案的n-1倍,其中n是所有中间识别结果和最终识别结果的平均长度。

[0044] 与传统的人机语音交互方案相比,该流式提前拉取对话资源的方案虽然可以在收到最终识别结果F时,直接将中间识别结果E的对话结果返回,以减少语音交互的响应时间,但是之前的中间识别结果A、B、C、D对对话服务的请求结果,并不会被采用,而对话服务依然需要进行对话模型的计算和资源获取,会造成大量的对话服务计算资源浪费,尤其是针对一些需要调用收费资源服务API的情况,例如调用天气服务商、音频内容提供商的接口以获取响应资源时,是按次收费的,会进一步增加相应的经济成本。

[0045] 因此,亟需提供一种语音交互处理方法,以在减少语音交互的响应时间的情况下,减少对对话服务的无效调用,从而降低成本。

[0046] 图1是根据本公开第一实施例的示意图,如图1所示。

[0047] 101、对用户发送的请求语句进行语音识别,在接收到所述请求语句后的第一预设时间内,得到至少一个中间识别结果。

[0048] 其中,请求语句是用户发送的想要得到回复的语句,例如可以是“播放儿歌”、“我想听稻香”等不同需求的话语,本公开实施例对此不做限定。

[0049] 其中,第一预设时间的结束时刻,早于所述请求语句的语音活动检测的尾点时刻(VADend)。第一预设时间是指,在用户发送的请求语句的尾音落下之后开始计时的一段时间,即该第一预设时间为语音交互的装置检测到的静音状态持续的时间。第一预设时间的长度可以根据需要设置。在实际使用时,为了减少用户的等待时间,第一预设时间的长度可以设置为5ms、10ms、20ms等,这里只是举例说明,不能作为本公开中,对第一预设时间长度的限定。

[0050] 其中,中间识别结果,是在用户发送的请求语句的尾音落下之后,第一预设时间内,对查询语句进行语音识别所得到的结果。

[0051] 102、响应于得到所述至少一个中间识别结果中的第一个中间识别结果,按照得到所述至少一个中间识别结果的时间顺序,依次识别所述至少一个中间识别结果的语义是否完整,即是否具有完整的语义,表达的意思是否完整。

[0052] 103、响应于从所述至少一个中间识别结果中识别出第一个语义完整的中间识别结果,根据所述第一个语义完整的中间识别结果,确定第一答复语句以播放。

[0053] 其中,第一答复语句,是根据第一个语义完整的中间识别结果去调用对话服务,对话服务根据该第一个语义完整的中间识别结果及所处的对话状态,判定用户的意图,进行对话模型的计算和资源获取,生成要给用户播报的答复语句。

[0054] 需要说明的是,101~103的执行主体的部分或全部可以为位于本地终端即服务提供商的终端设备的应用,或者还可以为设置在位于本地终端的应用中的插件或软件开发工具包(Software Development Kit,SDK)等功能单元,或者还可以为位于网络侧服务器中的处理引擎,或者还可以为位于网络侧的分布式系统,本实施例对此不进行特别限定。

[0055] 可以理解的是,所述应用可以是安装在终端上的本地程序(nativeApp),或者还可以是终端上的浏览器的一个网页程序(webApp),本实施例对此不进行特别限定。

[0056] 这样,通过提前进行流式语音识别,在识别出第一个语义完整的中间识别结果时,才根据该第一个语义完整的中间识别结果调用对话服务确定第一答复语句,既可以通过提前拉取对话资源减少语音交互的响应时间,又可以避免基于语义不完整的中间识别结果大量调用对话服务进行对话模型的计算和资源拉取,能够大幅度减少对对话服务的无效调用,减少对话服务的请求量,从而节省对话服务的计算资源和存储资源、以及收费资源服务,从而降低成本。

[0057] 图2是根据本公开第二实施例的示意图,如图2所示。在图1所示实施例的基础上,在102中,依次识别所述至少一个中间识别结果的语义是否完整,可以包括:

[0058] 201、依次针对所述至少一个中间识别结果的各中间识别结果,利用语义完整性模型,获取各中间识别结果作为历史最终识别结果中前缀的第一概率,以及各中间识别结果作为历史最终识别结果的第二概率。

[0059] 其中的语义完整性模型基于用户日志中的最终识别结果进行统计计算得到,例

如,统计大规模线上用户日志中,最终识别结果“我想听稻香”总共出现过多少次,最终识别结果“我想订明天早上六点五十的闹铃”总共出现过多少次,等等,对线上大规模用户日志中的各最终识别结果总共出现的次数分别进行统计计算,得到语义完整性模型。语义完整性模型可以按照一定更新周期,例如一周,基于线上用户日志的更新而更新,本公开实施例对语义完整性模型的是否更新、以及更新周期不做限制。

[0060] 本公开实施例中,历史最终识别结果是线上大规模用户日志中的最终识别结果,线上各用户每次进行语音交互发出请求语句时,会得到至少一个中间识别结果和一个最终识别结果,用于实现本公开实施例的语音交互的装置会针对各用户生成用户日志,记录该用户的每次语音交互中的至少一个中间识别结果和一个最终识别结果。

[0061] 本公开实施例中,中间识别结果作为历史最终识别结果中前缀的第一概率,即该中间识别结果仅作为历史最终识别结果中前面一部分内容、而非完整的历史最终识别结果出现的概率。中间识别结果作为历史最终识别结果的第二概率,即该中间识别结果单独作为历史最终识别结果(即作为一个完整的历史最终识别结果)出现的概率。

[0062] 202、基于所述第一概率和所述第二概率确定各中间识别结果语义完整的第三概率。

[0063] 本公开实施例中,第一概率越高,第二概率越低,则该中间识别结果语义完整的第三概率就越低。例如对于中间结果“我想听”,在用户的表述中,很少会单独作为一个请求语句出现,几乎都是作为请求语句的前缀出现,则中间结果“我想听”的第三概率就比较低,以将该中间结果“我想听”判定为语义不完整。

[0064] 203、根据所述第三概率是否大于预设阈值,确定各中间识别结果的语义是否完整。

[0065] 本公开实施例中,可以在第三概率大于预设阈值时,确定该中间识别结果的语义完整,所表达的意思比较完整,是一个完整的语句,从而即可将该中间识别结果作为完整的语句使用,进而根据该中间识别结果确定第一答复语句;否则,在第三概率小于或等于预设阈值时,确定中间识别结果的语义不完整,所表达的意思不够完整,可能不是一个完整的句子,该中间识别结果不可靠。其中的预设阈值可以根据实际需求设定,例如可以设定为0.5,并可以根据实际需求调整。

[0066] 基于本实施例,可以预先通过对大规模线上用户日志中的最终识别结果进行统计计算得到语义完整性模型,然后利用语义完整性模型,可以快速得到各中间识别结果作为历史最终识别结果中前缀的第一概率、以及各中间识别结果作为历史最终识别结果的第二概率,由该第一概率和第二概率精确确定各中间识别结果语义完整的第三概率,从而利用历史统计计算的方法,快速、客观的确定各中间识别结果的语义是否完整。

[0067] 图3是根据本公开第三实施例的示意图,如图3所示。在图1所示实施例的基础上,在102中,依次识别所述至少一个中间识别结果的语义是否完整,可以包括:

[0068] 301、依次针对所述至少一个中间识别结果的各中间识别结果,获取各中间识别结果的词向量。

[0069] 可选地,在本实施例的一个可能的实现方式中,可以通过词到向量(Word to the vector)方式,将中间识别结果从文本转换为词向量。

[0070] 302、获取各中间识别结果作为历史最终识别结果中前缀的第一概率,以及各中间

识别结果作为历史最终识别结果的第二概率。

[0071] 本公开实施例中,历史最终识别结果是线上大规模用户日志中的最终识别结果,线上各用户每次进行语音交互发出请求语句时,会得到至少一个中间识别结果和一个最终识别结果,用于实现本公开实施例的语音交互的装置会针对各用户生成用户日志,记录该用户的每次语音交互中的至少一个中间识别结果和一个最终识别结果。

[0072] 本公开实施例中,中间识别结果作为历史最终识别结果中前缀的第一概率,即该中间识别结果仅作为历史最终识别结果中前面一部分内容、而非完整的历史最终识别结果出现的概率。中间识别结果作为历史最终识别结果的第二概率,即该中间识别结果单独作为历史最终识别结果(即作为一个完整的历史最终识别结果)出现的概率。

[0073] 303、获取各中间识别结果的热度。

[0074] 本公开实施例中,中间识别结果的热度,即某一时间段内或者过去的所有历史时间内,用户通过用于实现本公开实施例的语音交互的装置,发送包含该中间识别结果的请求语句的使用度,可以通过大规模线上用户日志中的中间识别结果和最终识别结果统计计算得到。

[0075] 304、将所述词向量、所述第一概率、所述第二概率和所述热度输入神经网络模型,经神经网络模型输出各中间识别结果语义完整的第四概率。

[0076] 本公开实施例中的神经网络模型,可以是任何基于深度学习方式的神经网络模型,例如深度神经网络(Deep Neural Network,DNN),长短期记忆(Long Short Term Memory,LSTM)、基于多头注意力机制的模型(transformer)等,本公开实施例对此不做限制。

[0077] 305、根据所述第四概率是否大于预设阈值,确定各中间识别结果的语义是否完整。

[0078] 本公开实施例中,可以在第四概率大于预设阈值时,确定该中间识别结果的语义完整,所表达的意思比较完整,是一个完整的语句,从而即可将该中间识别结果作为完整的语句使用,进而根据该中间识别结果确定第一答复语句;否则,在第四概率小于或等于预设阈值时,确定中间识别结果的语义不完整,所表达的意思不够完整,可能不是一个完整的句子,该中间识别结果不可靠。其中的预设阈值可以根据实际需求设定,例如可以设定为0.5,并可以根据实际需求调整。

[0079] 基于本实施例,基于中间识别结果的词向量、该中间识别结果作为历史最终识别结果中前缀的第一概率和作为历史最终识别结果的第二概率、以及热度,通过神经网络模型,基于深度学习的方式,来预测该中间识别结果语义完整的第四概率,可以避免由于统计计算方式存在数据稀疏导致无法确定中间识别结果语义是否完整、或者导致确定结果不准确的问题,可以解决各中间识别结果的泛化性问题,实现对中间识别结果语义是否完整的确定,并提高中间识别结果语义是否完整的确定结果的准确性。

[0080] 可选地,在本实施例的一个可能的实现方式中,在操作302可以通过操作201实现,即可以利用语义完整性模型,获取各中间识别结果作为历史最终识别结果中前缀的第一概率,以及各中间识别结果作为历史最终识别结果的第二概率,其中的语义完整性模型基于用户日志中的最终识别结果进行统计计算得到。

[0081] 基于本实施例,可以预先通过对大规模线上用户日志中的最终识别结果进行统计

计算得到语义完整性模型,然后利用语义完整性模型,可以快速、客观的确定各中间识别结果作为历史最终识别结果中前缀的第一概率、以及各中间识别结果作为历史最终识别结果的第二概率,基于该第一概率和第二概率、以及词向量和热度,结合神经网络模型来预测该中间识别结果语义完整的第四概率,可以在保证关于中间识别结果语义是否完整的确定结果的准确性的情况下,进一步提高中间识别结果语义是否完整的确定结果的速度和客观性。

[0082] 可选地,可以预先通过多个具有标注信息的训练样本训练得到所述神经网络模型。例如,在本实施例的一个可能的实现方式中,可以通过如下方式训练得到所述神经网络模型:

[0083] 基于用户日志中的最终识别结果构造训练集,其中,该训练集包括多个训练样本,该多个训练样本包括正样本和负样本,其中的正样本具有语义完整的标注信息(例如1),负样本具有语义不完整的标注信息(例如0)。例如,对于用户日志中的最终识别结果“我想听稻香”,可以构造出5个训练样本:a.我;b.我想;c.我想听;d.我想听稻香;e.我想听稻香。其中a、b、c、d会被作为负样本(即语义不完整的样本)标注0,e作为正样本(即语义完整的样本)标注1;

[0084] 分别针对所述多个训练样本中各训练样本,获取各训练样本作为历史最终识别结果中前缀的第一概率,以及各训练样本作为所述历史最终识别结果的第二概率;

[0085] 获取各训练样本的热度;

[0086] 将各训练样本的所述词向量、所述第一概率、所述第二概率和所述热度输入待训练的神经网络模型,经待训练的神经网络模型输出各训练样本语义完整的预测概率;

[0087] 根据预测概率是否大于预设阈值,确定各训练样本的语义是否完整,得到预测结果;

[0088] 基于各训练样本的预测结果与对应的标注信息之间的差异,对待训练的神经网络模型进行训练,即调整神经网络模型的网络参数,直至满足预设训练完成条件。

[0089] 其中,在训练得到所述神经网络模型的过程中,对各训练样本的处理操作与上述对各中间处理结果的对应操作的内容类似,可参考上述各实施例的记载,此处不再赘述。

[0090] 本公开实施例中,上述训练得到所述神经网络模型的过程可以是迭代操作,即迭代执行上述训练得到所述神经网络模型的过程,直至满足预设训练完成条件,即可由待训练的神经网络模型得到所述神经网络模型。其中的预设训练完成条件,例如可以包括但不限于以下任意一项或多项:上述多个训练样本的预测结果与对应的标注信息之间的差异的均值,小于预设差异阈值(例如0.2),或者迭代执行上述训练得到所述神经网络模型的过程的次数达到预设次数(例如2000次),等,本公开实施例对此不做限制。

[0091] 基于本实施例,可以基于用户日志中的最终识别结果构造训练集,采用深度学习的方式来训练得到神经网络模型,具有泛化性,能够避免由于数据稀疏问题导致的限制性,可以用于识别任何文本的语义完整性,能够有效识别各中间识别结果的语义是否完整,提高识别结果的准确性。

[0092] 图4是根据本公开第四实施例的示意图,如图4所示。在图1-图3所示任一实施例的基础上,该实施例还可以包括:

[0093] 401、在接收到所述请求语句后的第二预设时间内,得到所述请求语句的最终识别

结果。

[0094] 其中,第二预设时间的结束时刻,晚于所述请求语句的语音活动检测的尾点时刻。第二预设时间,是在用户发送的请求语句的尾音落下之后开始计时的一段时间,这一段计时时间的长度大于第一预设时间的长度。可以理解的是,第二预设时间,也是语音交互的装置检测到的静音状态的持续时间。

[0095] 402、响应于从所述至少一个中间识别结果中未识别出语义完整的中间识别结果,根据所述最终识别结果,确定第二答复语句以播放。

[0096] 基于本实施例,可以在从所述至少一个中间识别结果中未识别出语义完整的中间识别结果时,根据请求语句的最终识别结果确定第二答复语句并播放,以确保语音交互结果的准确性,避免错误的语音交互结果影响用户体验。

[0097] 图5是根据本公开实施例的一个语音交互过程的示意图,如图5所示。以一个具体的语音交互过程为例,对本公开实施例做进一步的说明。

[0098] 如图5所示,在一次完成的语音交互过程中,从用户发送请求语句的语音落到播放交互装置返回的答复语句的语音起,会经历如下各阶段:

[0099] 用户发出“我想听稻香”的请求语句,该请求语句的语音信号的尾音(即用户语音落的时刻)时刻为1;

[0100] 在时刻1之后即进入VAD检测,即连续检测到第二预设时间 t_1 的静音,直到到达时刻2,即VADend点,语音交互的装置才会认为用户已经说完;

[0101] 在时刻1之后经过第一预设时间到达时刻1.5,即得到至少一个中间识别结果,包括:A.我;B.我想;C.我想听;D.我想听稻;E.我想听稻香;

[0102] 依次识别上述各中间识别结果A、B、C、D、E的语义是否完整,识别出中间识别结果A、B、C、D的语义不完整,中间识别结果E的语义完整;

[0103] 根据中间识别结果E调用对话服务进行对话模型的计算和资源获取,从时刻1.5经历时间 t_3 后,在时刻4得到第一答复文本,即语音合成(Text To Speech, tts)文本;

[0104] 对该第一答复文本进行语音合成,从时刻4经历时间 t_4 后,在时刻5得到第一答复语句“好的”,开始调用相应的系统接口,播放该第一答复语句,经历时间 t_5 到达时刻6,用户听到第一答复语句的播放音频;

[0105] 同时,从时刻1经过第二预设时间达到时刻2(VAD end点),语音交互的装置持续检测到静音,认为用户已经说完,继续进行语音识别,到达时刻3,才得到最终识别结果:F.我想听稻香。

[0106] 具体实现时,时刻6可能早于时刻3,也可以由于第二预设时间相对较短使得时刻6晚于时刻3,但是,由于中间识别结果E比最终识别结果F更早得到,会提前拉取对话资源从而减少语音交互的响应时间,使得用户可以更早的听到回复语句的音频,从而提高用户体验;同时,仅针对语义完整的中间识别结果调用对话服务可以避免基于语义不完整的中间识别结果大量调用对话服务进行对话模型的计算和资源拉取,能够大幅度减少对对话服务的无效调用,减少对话服务的请求量,从而节省对话服务的计算资源和存储资源、以及收费资源服务,从而降低成本。

[0107] 本实施例中,通过提前进行流式语音识别,在识别出第一个语义完整的中间识别结果时,才根据所述第一个语义完整的中间识别结果调用对话服务确定第一答复语句,既

可以通过提前拉取对话资源减少语音交互的响应时间,又可以避免基于语义不完整的中间识别结果大量调用对话服务进行对话模型的计算和资源拉取,能够大幅度减少对对话服务的无效调用,减少对话服务的请求量,从而节省对话服务的计算资源和存储资源、以及收费资源服务,从而降低成本。

[0108] 另外,预先通过对大规模线上用户日志中的最终识别结果进行统计计算得到语义完整性模型,然后利用语义完整性模型,可以快速得到各中间识别结果作为历史最终识别结果中前缀的第一概率、以及各中间识别结果作为历史最终识别结果的第二概率,由该第一概率和第二概率精确确定各中间识别结果语义完整的第三概率,从而利用历史统计计算的方法,快速、客观的确定各中间识别结果的语义是否完整。

[0109] 另外,基于中间识别结果的词向量、该中间识别结果作为历史最终识别结果中前缀的第一概率和作为历史最终识别结果的第二概率、以及热度,通过神经网络模型,基于深度学习的方式,来预测该中间识别结果语义完整的第四概率,可以避免由于统计计算方式存在数据稀疏导致无法确定中间识别结果语义是否完整、或者导致确定结果不准确的问题,可以解决各中间识别结果的泛化性问题,实现对中间识别结果语义是否完整的确定,并提高中间识别结果语义是否完整的确定结果的准确性。

[0110] 另外,可以基于用户日志中的最终识别结果构造训练集,采用深度学习的方式来训练得到神经网络模型,具有泛化性,能够避免由于数据稀疏问题导致的限制性,可以用于识别任何文本的语义完整性,能够有效识别各中间识别结果的语义是否完整,提高识别结果的准确性。

[0111] 另外,可以在从所述至少一个中间识别结果中未识别出语义完整的中间识别结果时,根据请求语句的最终识别结果确定第二答复语句并播放,以确保语音交互结果的准确性,避免错误的语音交互结果影响用户体验。

[0112] 需要说明的是,对于前述的各方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本公开并不受所描述的动作顺序的限制,因为依据本公开,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定是本公开所必须的。

[0113] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述的部分,可以参见其他实施例的相关描述。

[0114] 图6是根据本公开第五实施例的示意图,如图6所示。本实施例的语音交互的装置600可以包括语音识别单元601、语义完整识别单元602和第一确定单元603。其中,语音识别单元601,用于对用户发送的请求语句进行语音识别,在接收到所述请求语句后的第一预设时间内,得到至少一个中间识别结果,其中,第一预设时间的结束时刻,早于所述请求语句的语音活动检测的尾点时刻;语义完整识别单元602,用于响应于得到所述至少一个中间识别结果中的第一个中间识别结果,按照得到所述至少一个中间识别结果的时间顺序,依次识别所述至少一个中间识别结果的语义是否完整;第一确定单元603,用于响应于从所述至少一个中间识别结果中识别出第一个语义完整的中间识别结果,根据所述第一个语义完整的中间识别结果,确定第一答复语句以播放。

[0115] 需要说明的是,本实施例的语音交互的装置600的部分或全部可以为位于本地终

端即服务提供商的终端设备的应用,或者还可以为设置在位于本地终端的应用中的插件或软件开发工具包(Software Development Kit,SDK)等功能单元,或者还可以为位于网络侧服务器中的处理引擎,或者还可以为位于网络侧的分布式系统,本实施例对此不进行特别限定。

[0116] 可以理解的是,所述应用可以是安装在终端上的本地程序(nativeApp),或者还可以是终端上的浏览器的一个网页程序(webApp),本实施例对此不进行特别限定。

[0117] 图7是根据本公开第六实施例的示意图,如图7所示。在图6所示实施例的基础上,本实施例语音交互的装置700中,语义完整识别单元602可以包括:语义完整性模型701,用于依次针对所述至少一个中间识别结果中的各中间识别结果,获取各中间识别结果作为历史最终识别结果中前缀的第一概率,以及各中间识别结果作为历史最终识别结果的第二概率,其中的语义完整性模型基于用户日志中的最终识别结果进行统计计算得到;第一确定模块702,用于基于所述第一概率和所述第二概率确定所述各中间识别结果语义完整的第三概率;第二确定模块703,用于根据所述第三概率是否大于预设阈值,确定各中间识别结果的语义是否完整。

[0118] 可选地,在本实施例的一个可能的实现方式中,语义完整识别单元602可以包括(图中未示出):第一获取模块,用于依次针对所述至少一个中间识别结果中的各中间识别结果,获取各中间识别结果的词向量;第二获取模块,用于获取各中间识别结果作为历史最终识别结果中前缀的第一概率,以及各中间识别结果作为历史最终识别结果的第二概率;第三获取模块,用于获取各中间识别结果的热度;神经网络模型,用于接收输入的所述词向量、所述第一概率、所述第二概率和所述热度,输出各中间识别结果语义完整的第四概率;第三确定模块,用于根据第四概率是否大于预设阈值,确定各中间识别结果的语义是否完整。

[0119] 可选地,在本实施例的一个可能的实现方式中,第二获取模块具体用于:利用语义完整性模型,获取各中间识别结果作为历史最终识别结果中前缀的第一概率,以及各中间识别结果作为所述历史最终识别结果的第二概率,其中的语义完整性模型基于用户日志中的最终识别结果进行统计计算得到。

[0120] 可选地,再参见图7,本实施例语音交互的装置700中还可以包括构造单元604、第一获取单元605、第二获取单元606、待训练的神经网络模型607、第二确定单元608和训练单元609。其中,构造单元604,用于基于用户日志中的最终识别结果构造训练集,该训练集包括多个训练样本,所述多个训练样本包括正样本和负样本,正样本具有语义完整的标注信息,负样本具有语义不完整的标注信息;第一获取单元605,用于分别针对所述多个训练样本中各训练样本,获取各训练样本作为历史最终识别结果中前缀的第一概率,以及各训练样本作为所述历史最终识别结果的第二概率;第二获取单元606,用于获取各训练样本的热度;待训练的神经网络模型607,用于接收输入的各训练样本的词向量、第一概率、第二概率和热度,输出各训练样本语义完整的预测概率;第二确定单元608,用于根据预测概率是否大于预设阈值,确定各训练样本的语义是否完整,得到预测结果;训练单元609,用于基于各训练样本的预测结果与对应的标注信息之间的差异,对待训练的神经网络模型607进行训练,直至满足预设训练完成条件,得到上述神经网络模型。

[0121] 可选地,在本实施例的一个可能的实现方式中,语音识别单元601,还用于在接收

到所述请求语句后的第二预设时间内,得到所述请求语句的最终识别结果,其中,第二预设时间的结束时刻,晚于所述请求语句的语音活动检测的尾点时刻;相应地,该实施例中,第一确定单元603,还用于响应于从所述至少一个中间识别结果中未识别出语义完整的中间识别结果,根据所述最终识别结果,确定第二答复语句以播放。

[0122] 本实施例中,通过提前进行流式语音识别,在识别出第一个语义完整的中间识别结果时,才根据所述第一个语义完整的中间识别结果调用对话服务确定第一答复语句,既可以通过提前拉取对话资源减少语音交互的响应时间,又可以避免基于语义不完整的中间识别结果大量调用对话服务进行对话模型的计算和资源拉取,能够大幅度减少对对话服务的无效调用,减少对话服务的请求量,从而节省对话服务的计算资源和存储资源、以及收费资源服务,从而降低成本。

[0123] 另外,预先通过对大规模线上用户日志中的最终识别结果进行统计计算得到语义完整性模型,然后利用语义完整性模型,可以快速得到各中间识别结果作为历史最终识别结果中前缀的第一概率、以及各中间识别结果作为历史最终识别结果的第二概率,由该第一概率和第二概率精确确定各中间识别结果语义完整的第三概率,从而利用历史统计计算的方法,快速、客观的确定各中间识别结果的语义是否完整。

[0124] 另外,基于中间识别结果的词向量、该中间识别结果作为历史最终识别结果中前缀的第一概率和作为历史最终识别结果的第二概率、以及热度,通过神经网络模型,基于深度学习的方式,来预测该中间识别结果语义完整的第四概率,可以避免由于统计计算方式存在数据稀疏导致无法确定中间识别结果语义是否完整、或者导致确定结果不准确的问题,可以解决各中间识别结果的泛化性问题,实现对中间识别结果语义是否完整的确定,并提高中间识别结果语义是否完整的确定结果的准确性。

[0125] 另外,可以基于用户日志中的最终识别结果构造训练集,采用深度学习的方式来训练得到神经网络模型,具有泛化性,能够避免由于数据稀疏问题导致的限制性,可以用于识别任何文本的语义完整性,能够有效识别各中间识别结果的语义是否完整,提高识别结果的准确性。

[0126] 另外,可以在从所述至少一个中间识别结果中未识别出语义完整的中间识别结果时,根据请求语句的最终识别结果确定第二答复语句并播放,以确保语音交互结果的准确性,避免错误的语音交互结果影响用户体验。

[0127] 本公开的技术方案中,所涉及的用户个人信息的获取,存储和应用等,均符合相关法律法规的规定,且不违背公序良俗。

[0128] 根据本公开的实施例,本公开还提供了一种电子设备、一种可读存储介质和一种计算机程序产品,进一步地,还提供了一种包括所提供的电子设备的人工智能设备。

[0129] 图8示出了可以用来实施本公开的实施例的示例电子设备800的示意性框图。电子设备旨在表示各种形式的数字计算机,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例,并且不意在限制本文中描述的和/或者要求的本公开的实现。

[0130] 如图8所示,设备800包括计算单元801,其可以根据存储在只读存储器(ROM) 802中

的计算机程序或者从存储单元808加载到随机访问存储器 (RAM) 803中的计算机程序,来执行各种适当的动作和处理。在RAM803中,还可存储设备800操作所需的各种程序和数据。计算单元801、ROM802以及RAM803通过总线804彼此相连。输入/输出 (I/O) 接口805也连接至总线804。

[0131] 设备800中的多个部件连接至I/O接口805,包括:输入单元806,例如键盘、鼠标等;输出单元807,例如各种类型的显示器、扬声器等;存储单元808,例如磁盘、光盘等;以及通信单元809,例如网卡、调制解调器、无线通信收发机等。通信单元809允许设备800通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据。

[0132] 计算单元801可以是各种具有处理和计算能力的通用和/或专用处理组件。计算单元801的一些示例包括但不限于中央处理单元 (CPU)、图形处理单元 (GPU)、各种专用的人工智能 (AI) 计算芯片、各种运行机器学习模型算法的计算单元、数字信号处理器 (DSP)、以及任何适当的处理器、控制器、微控制器等。计算单元801执行上文所描述的各个方法和处理,例如语音交互的方法。例如,在一些实施例中,语音交互的方法可被实现为计算机软件程序,其被有形地包含于机器可读介质,例如存储单元808。在一些实施例中,计算机程序的部分或者全部可以经由ROM 802和/或通信单元809而被载入和/或安装到设备800上。当计算机程序加载到RAM803并由计算单元801执行时,可以执行上文描述的语音交互的方法的一个或多个步骤。备选地,在其他实施例中,计算单元801可以通过其他任何适当的方式(例如,借助于固件)而被配置为执行语音交互的方法。

[0133] 本文中以上描述的系统和技术和各种实施方式可以在数字电子电路系统、集成电路系统、场可编程门阵列 (FPGA)、专用集成电路 (ASIC)、专用标准产品 (ASSP)、芯片上系统的系统 (SOC)、负载可编程逻辑设备 (CPLD)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0134] 用于实施本公开的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器或控制器,使得程序代码当由处理器或控制器执行时使流程图和/或框图中所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行,作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务器上执行。

[0135] 在本公开的上下文中,机器可读介质可以是有形的介质,其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备,或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器 (RAM)、只读存储器 (ROM)、可擦除可编程只读存储器 (EPROM或快闪存储器)、光纤、便捷式紧凑盘只读存储器 (CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。

[0136] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0137] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)和互联网。

[0138] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。服务器可以是云服务器,也可以为分布式系统的服务器,或者是结合了区块链的服务器。

[0139] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本发公开中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本公开公开的技术方案所期望的结果,本文在此不进行限制。

[0140] 上述具体实施方式,并不构成对本公开保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等,均应包含在本公开保护范围之内。

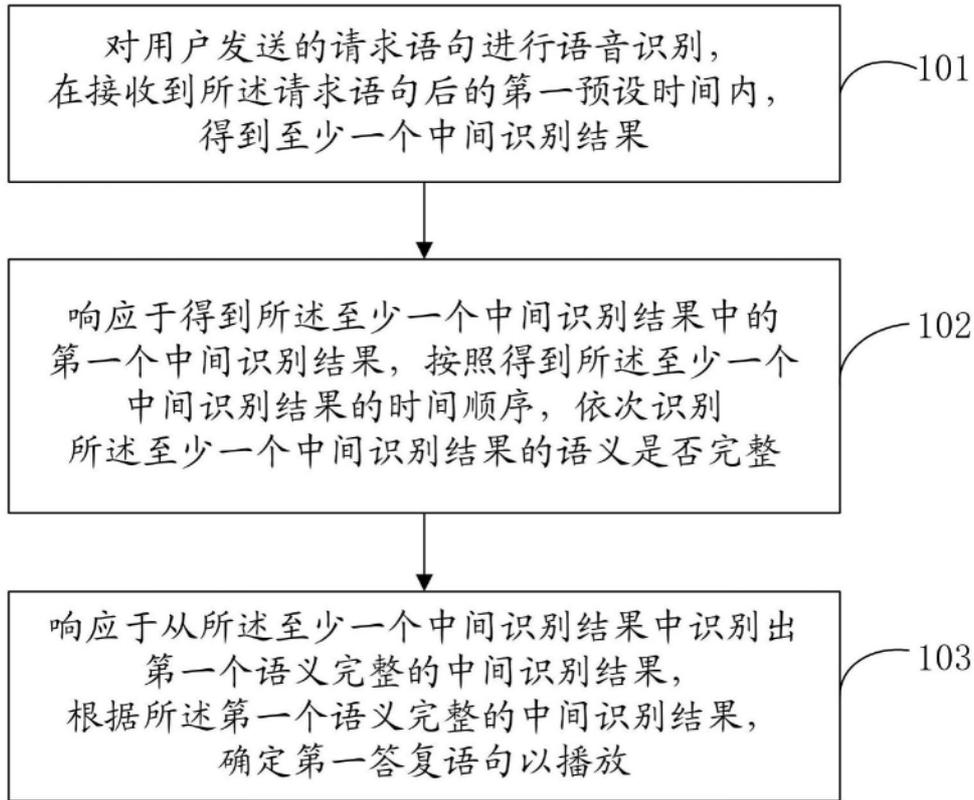


图1

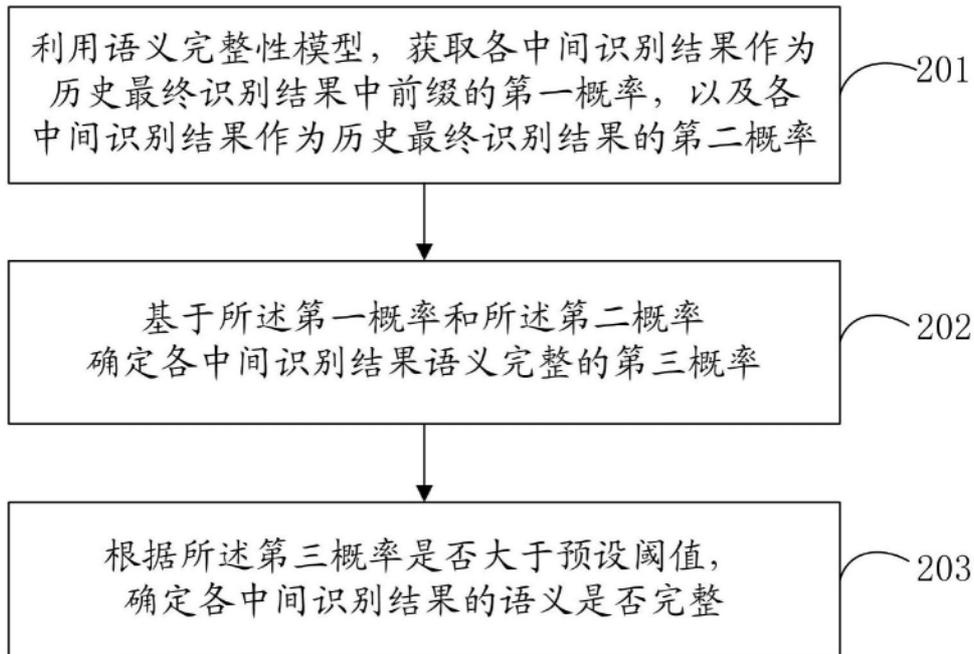


图2

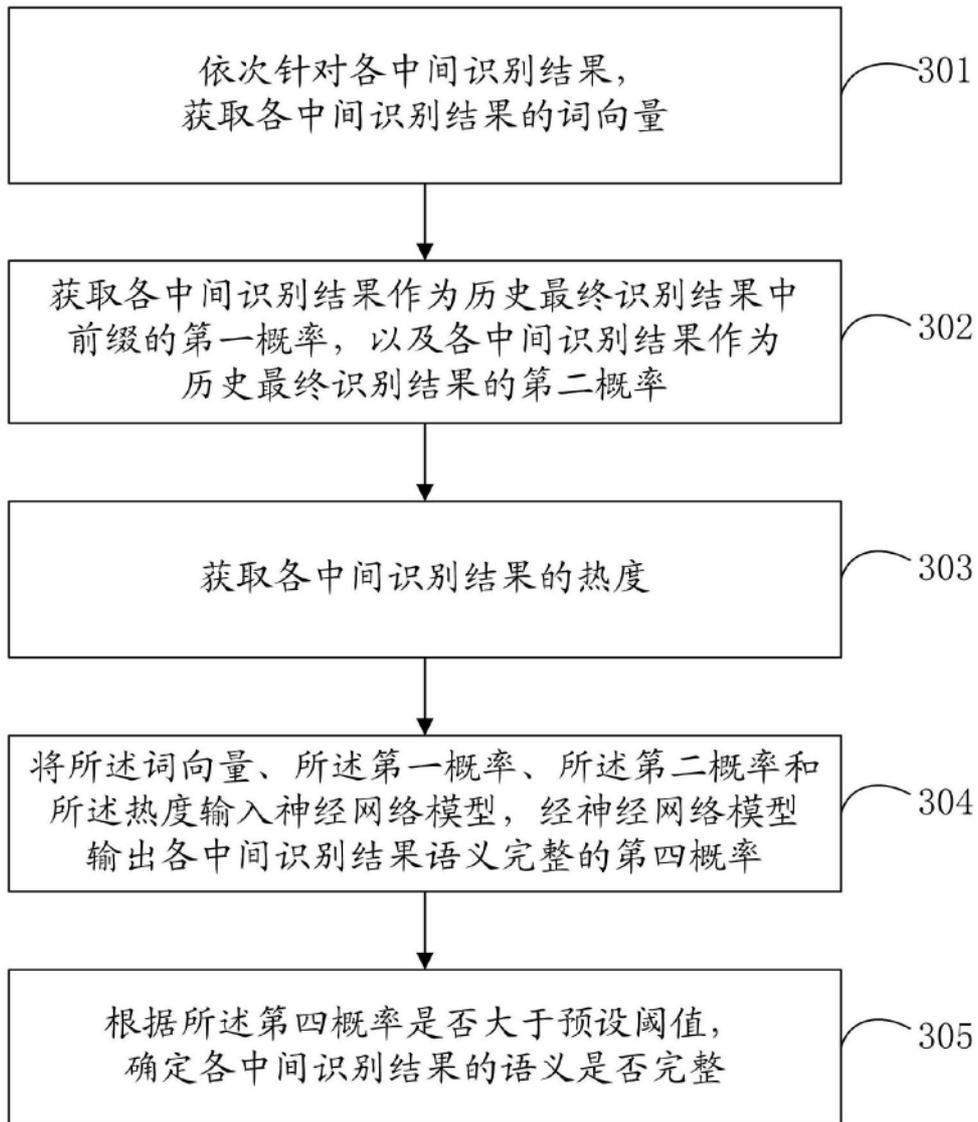


图3

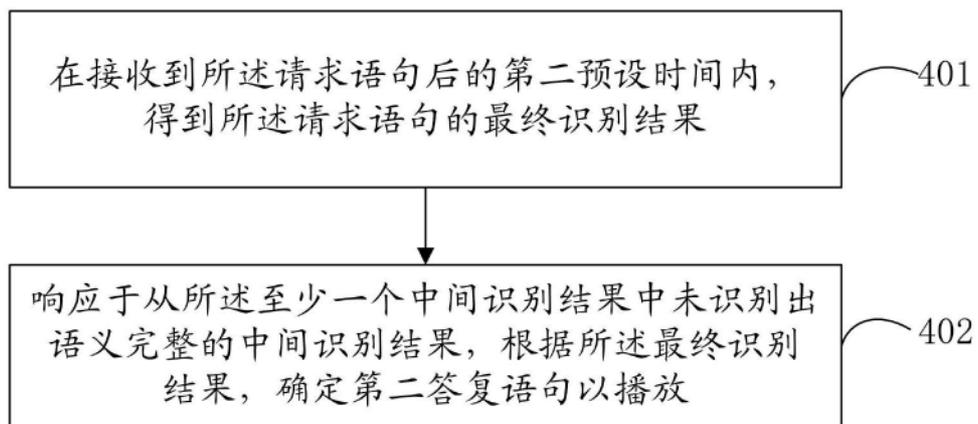


图4

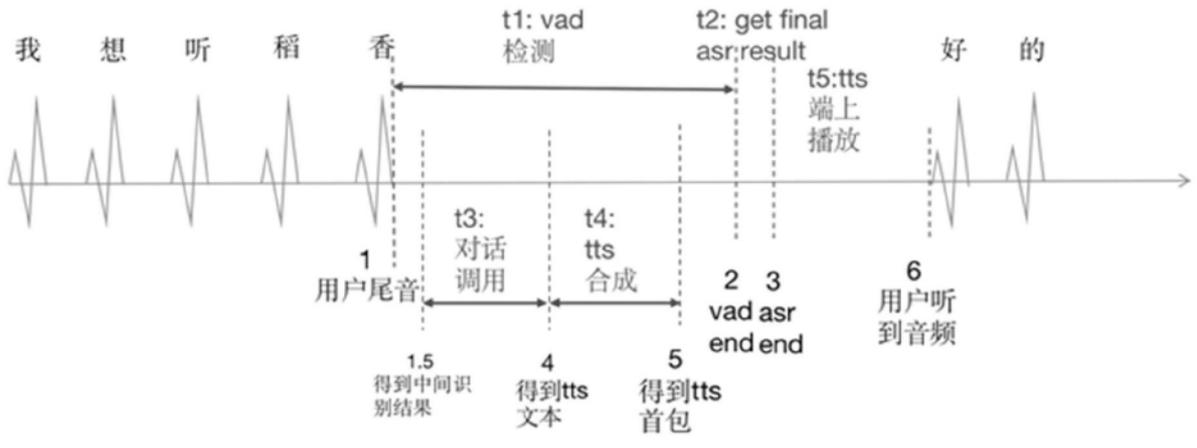


图5

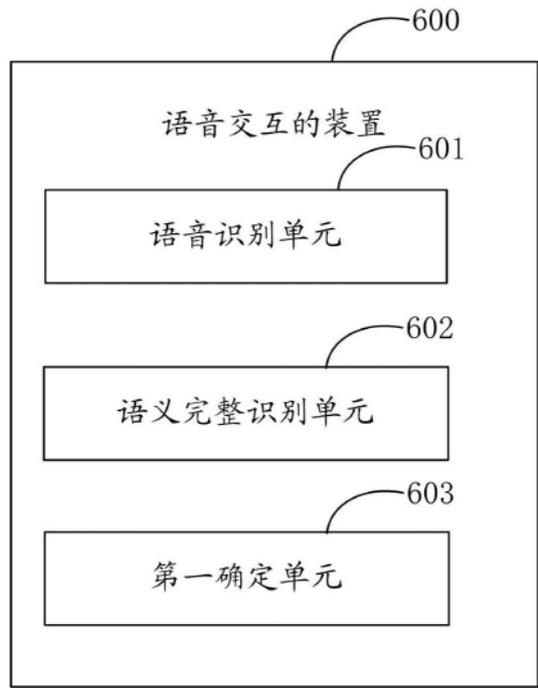


图6

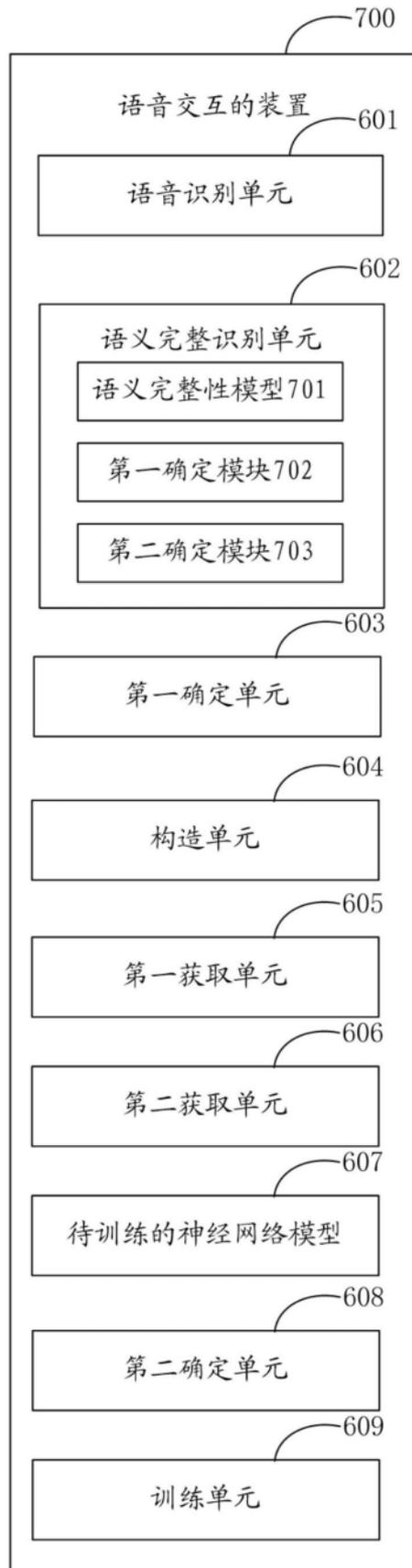


图7

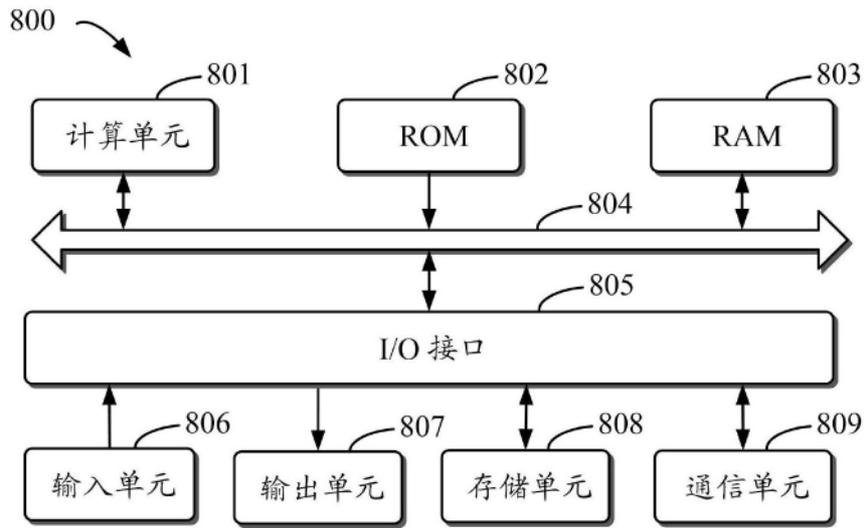


图8