

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2019 年 1 月 10 日 (10.01.2019)



(10) 国际公布号
WO 2019/007247 A1

(51) 国际专利分类号:
G10L 15/22 (2006.01)

(21) 国际申请号: PCT/CN2018/093225

(22) 国际申请日: 2018 年 6 月 28 日 (28.06.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710539395.6 2017年7月4日 (04.07.2017) CN

(71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

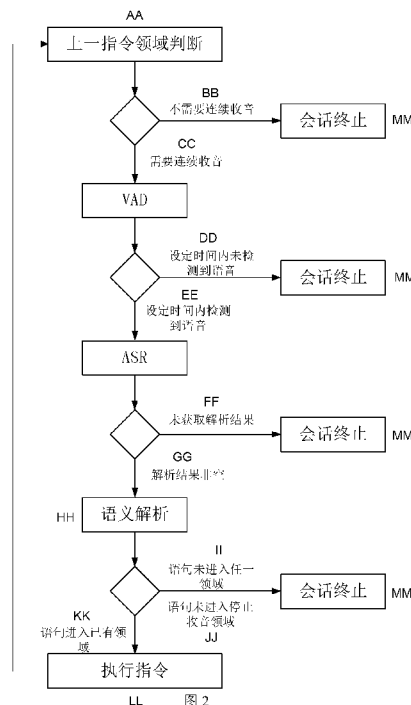
(72) 发明人: 刘广兴(LIU, Guangxing); 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 许毅(XU, Yi); 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS,

(54) Title: HUMAN-MACHINE CONVERSATION PROCESSING METHOD AND APPARATUS, AND ELECTRONIC DEVICE

(54) 发明名称: 人机会话的处理方法、装置及电子设备



AA JUDGEMENT OF THE FIELD OF THE LAST INSTRUCTION
BB NO NEED FOR CONTINUOUS VOICE RECEPTION
CC NEED FOR CONTINUOUS VOICE RECEPTION
DD NO VOICE IS DETECTED WITHIN A SET TIME
EE VOICE IS DETECTED WITHIN A SET TIME
FF NO PARSING RESULT IS ACQUIRED
GG A PARSING RESULT IS NOT EMPTY
HH SEMANTIC PARSING
II A SENTENCE HAS NOT ENTERED ANY FIELD
JJ THE SENTENCE HAS NOT ENTERED THE FIELD OF STOPPING VOICE RECEPTION
KK THE SENTENCE ENTERS AN EXISTING FIELD
LL EXECUTE AN INSTRUCTION
MM A CONVERSATION ENDS

(57) Abstract: Provided are a human-machine conversation processing method and apparatus, and an electronic device. The method comprises: after a device completing the last voice command, identifying the content of the last voice command to determine whether a user has the need to input the voice command again based on the last voice command; if it is determined that the user has the need to input a voice command again, initiating voice activity detection; otherwise, ending the conversation. The solution of the embodiment of the present invention can meet the requirement that a user wants to actively have continuous conversation with a device without

JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

repeatedly waking up the device, thereby improving the user experience, and improving the conversation efficiency.

(57) 摘要: 本发明实施例提供一种人机会话的处理方法、装置及电子设备, 其中, 方法包括: 在设备完成上一条语音指令后, 对上一条语音指令的内容进行识别, 确定用户是否有基于上一条语音指令而再次输入语音指令的需求; 如果确定用户有再次输入语音指令的需求, 则启动语音活动检测; 否则, 结束本次会话。本发明实施例的方案可以在不需要反复唤醒设备的基础上, 满足用户想主动与设备进行连续会话的需求, 改善用户使用体验, 提高会话效率。

人机会话的处理方法、装置及电子设备

本申请要求 2017 年 07 月 04 日递交的申请号为 201710539395.6、发明名称为“人机会话的处理方法、装置及电子设备”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5

技术领域

本申请涉及计算机技术领域，尤其涉及一种人机会话的处理方法、装置及电子设备。

背景技术

10 在人机会话的场景下，用户向设备输入一条语音指令后，设备会执行用户输入的语音指令，例如增大音量、播放视频等，指令执行完毕后，还可以通过 TTS (Text To Speech, 文字转语音) 向用户进行反馈，比如播放“音量已经增大”、“视频已打开”等。当设备完成一条语音指令后，认为整个会话终止了，便进入休眠状态。

但是，在连续会话场景下，如果用户还有进一步的语音指令想输入，则需要重新唤醒设备。重新唤醒设备无论从时间上还是程序上都会导致较大的使用不便，例如，用户
15 要重新输入语音唤醒词，并且，唤醒设备也会花一定的时间，从而严重影响使用体验。

发明内容

20 本发明提供了一种人机会话的处理方法、装置及电子设备，在不需要反复唤醒设备的基础上，满足用户想主动与设备进行连续会话的需求，改善用户使用体验，提高会话效率。

为达到上述目的，本发明的实施例采用如下技术方案：

第一方面，提供了一种人机会话的处理方法，包括：

25 在设备完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求；

如果确定用户有再次输入语音指令的需求，则启动语音活动检测；否则，结束本次会话。

第二方面，提供了另一种人机会话的处理方法，包括：

对所接收的语音指令的内容进行识别；

30 判断用户是否有再次输入语音指令的需求；

根据判断结果，执行人机会话操作。

第三方面，提供了一种人机会话的处理装置，包括：

指令识别模块，用于在设备完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求；

- 5 语音检测模块，用于如果确定用户有再次输入语音指令的需求，则启动语音活动检测；否则，结束本次会话。

第四方面，提供了另一种人机会话的处理装置，包括：

内容识别模块，用于对所接收的语音指令的内容进行识别；

需求判断模块，用于判断用户是否有再次输入语音指令的需求；

- 10 执行操作模块，用于根据判断结果，执行人机会话操作。

第五方面，提供了一种电子设备，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

- 15 在设备完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求；

如果确定用户有再次输入语音指令的需求，则启动语音活动检测；否则，结束本次会话。

第六方面，提供了另一种电子设备，包括：

存储器，用于存储程序；

- 20 处理器，耦合至所述存储器，用于执行所述程序，以用于：

对所接收的语音指令的内容进行识别；

判断用户是否有再次输入语音指令的需求；

根据判断结果，执行人机会话操作。

- 25 本发明提供的人机会话的处理方法、装置及电子设备，在设备完成上一条语音指令后，通过对用户是否会输入下一条语音指令进行预测判断，从而能够提高设备执行用户的连续语音指令的效率，提升用户体验。

上述说明仅是本申请技术方案的概述，为了能够更清楚了解本申请的技术手段，而可依照说明书的内容予以实施，并且为了让本申请的上述和其它目的、特征和优点能够更明显易懂，以下特举本申请的具体实施方式。

附图说明

通过阅读下文优选实施方式的详细描述，各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的，而并不认为是对本申请的限制。而且在整个附图中，用相同的参考符号表示相同的部件。在附图中：

- 5 图 1 为本发明实施例的人机会话的处理的逻辑示意图一；
图 2 为本发明实施例的人机会话的处理的逻辑示意图二；
图 3 为本发明实施例的人机会话的处理的系统结构图；
图 4a 为本发明实施例的人机会话的处理方法流程图一；
图 4b 为本发明实施例的人机会话的处理方法流程图一；
10 图 5a 为本发明实施例的人机会话的处理装置结构图一；
图 5b 为本发明实施例的人机会话的处理装置结构图二；
图 6 为本发明实施例的人机会话的处理装置结构图三；
图 7 为本发明实施例的电子设备的结构示意图一；
图 8 为本发明实施例的电子设备的结构示意图二。

15

具体实施方式

下面将参照附图更详细地描述本公开的示例性实施例。虽然附图中显示了本公开的示例性实施例，然而应当理解，可以以各种形式实现本公开而不应被这里阐述的实施例所限制。相反，提供这些实施例是为了能够更透彻地理解本公开，并且能够将本公开的范围完整的传达给本领域的技术人员。

20

在现有人机会话的场景中，用户向机器输入一条语音指令后，机器会执行用户输入的语音指令，例如增大音量、播放视频等，指令执行完毕后，还可以通过 TTS 向用户进行反馈，比如播放“音量已经增大”。当然，通过 TTS 向用户进行反馈并不是必要的操作，设备可以执行语音反馈；或者设备会针对该语音指令进行回答，例如用户输入“讲个笑话”，则设备会选取笑话通过 TTS 来进行回答，再例如用户输入“今天天气如何”，设备会通过 TTS 播报天气预报。为了方便描述，无论是设备通过 TTS 向用户回答了问题还是进行了反馈还是仅执行了语音指令对应的操作而没有进行反馈，我们将这几种情形都统称为“设备完成语音指令”。

25

本发明改变了现有技术中，在设备完成语音指令后，马上结束会话的处理流程，其核心思想在于，在设备完成语音指令后，先对前一次用户输入的语音指令的内容进行判

30

断，确认用户是否还会输入下一条语音指令，如果判定为用户还会输入下一条语音指令，则在设备完成语音指令后，进入语音活动检测（Voice Activity Detection，VAD）流程，如果判定为用户不会输入下一条语音指令，则终止会话，从而提高用户连续会话的效率。

如图 1 所示，为本发明实施例的人机会话的处理的逻辑示意图。在该逻辑图中，人机会话的基础流程依次为用户唤醒、VAD、语音输入、ASR（Automatic Speech Recognition，自动语音识别）、语义解析、指令执行、系统反馈、TTS，此流程构成一个闭环。如果通过对上一条语音指令的内容进行判断后，用户需要有连续发语音指令的需求，则在 TTS 之后可再启动 VAD 进行语音检测，重复原有会话流程。

另外，本申请还要解决在考虑连续会话场景下，在上述流程中的各环节判断整个会话终止的问题。

现有的判断会话终止的方案为从用户唤醒到最终 TTS 反馈整个流程完结作为会话终止判断条件，基本上不考虑连续会话场景，在会话流程中的各环节一旦出现异常情况，会进行错误类型判断，并通过 TTS 进行异常反馈，TTS 播报之后认为一个会话单元终止。在一些特殊情况下，如语音系统主动向用户提问时，TTS 播报后会重新从 VAD 环节开始进行会话流程。

图 2 为在连续会话场景下，终止会话的处理逻辑图，在该逻辑中，大体包括 5 个步骤：领域判断、VAD、ASR、语义解析、执行语音指令。

领域判断：判断上一语音指令完成后是否需要进入连续会话（连续会话）状态。如果根据上一语音指令的内容，判断用户还有进一步发出语音指令的需求，则启动 VAD；否则，结束本次会话。

VAD：启动 VAD 后，在设定时间内如果检测到语音信号，则将语音信号发送至 ASR 进行语音解析，形成文本；如果没有检测到语音信号，则结束本次会话。

ASR：对语音信号进行文本解析，如果解析得到文本内容，则将文本内容进行语义解析；如果文本解析后未得到文本内容，则终止本次会话。

语义解析：对文本内容进行语义解析，判断文本中的语句是否进入预置的领域内，如果语句未进入任一领域，或者语句进入停止收音（终止会话）领域，则终止本次会话；如果语句进入已有的领域，则根据已有的领域形成语音指令。

执行指令：根据确定的语音指令，控制相应设备执行语音指令并通过 TTS 反馈。

语音指令执行完成后，流程重新指向第一个步骤，即继续对上一指令的领域进行判断，确定用户是否需要连续会话。

基于图 1 和图 2 所示的人机会话的处理方法的逻辑示意图，本发明实施例提供了一种人机会话的处理系统，用以提高人机会话场景下，用户主动连续会话的效率。如图 3 所示，该系统包括：设备 310 和服务器 320。

设备 310 包括：

- 5 人机会话过程中的人机交互设备，如麦克风、音响等以及执行语音指令的操作设备，如媒体播放设备、空调、电视、冰箱等。

设备 310，用于在人机会话过程中，与人进行交互，包括语音信号采集、TTS 反馈等，以及执行语音指令的具体操作。

- 10 服务器 2 具有控制 VAD 启动、ASR、语义解析、形成语音控制指令并反馈至设备等各环节的逻辑处理功能。

如图 3 所示，服务器 2 具体包括：人机会话的处理装置 321 和领域库 322；

人机会话的处理装置 321，包括：

- 15 指令识别模块，用于在设备 310 完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求；在对语音指令的内容进行识别时，需要将语音指令在领域库 322 中进行需求类型的判断。领域库 322 中预置存储了多个不同领域内的多种意图。

所谓领域，既人机交互系统中实现用户的某一类需求的功能。领域识别即判断用户的某一语音属于哪一类需求的过程。

- 20 意图，既人机交互系统中实现某一领域下用户某一单一明确需求的功能。意图识别即判断用户的某一语音属于某一领域下哪一个明确需求的过程。

通过将用户的上一条语音指令的内容在领域库 322 中识别判断，就可以判定用户是否还存在再次输入语音指令的需求。

语音检测模块，用于如果确定用户有再次输入语音指令的需求，则启动语音活动检测 VAD；否则，结束本次会话。

- 25 在设备 310 一侧，用于接收语音信号的设备如麦克风始终是开启状态，但只有服务器 320 确定启动 VAD 的检测流程后，麦克风接收的语音信号才会被传送至人机会话的处理装置 321 中，作为人机会话流程中的语音指令信号。而 VAD 的检测流程只有在每次会话开始时启动，用以检测语音信号，当系统认为无语音输入时，会自动关闭 VAD，直到再次用户唤醒设备才再次启动 VAD。因此，当上述指令识别模块判定出用户还有再次输入语音指令的需求时，会触发语音检测模块启动 VAD。
- 30

语音检测模块，还用于在启动 VAD 后，如果在指定的检测时间内没有检测到语音信号，则结束本次会话；否则，触发语音识别模块对检测到的语音信号进行自动语音识别 ASR。

时间计算模块，用于：

- 5 统计用户在各次会话过程中，用户从设备唤醒成功到发出语音指令的第一平均时间；统计用户在各次会话过程中，从启动 VAD 到用户发出语音指令的第二平均时间；根据第一平均时间和所述第二平均时间计算得到指定的检测时间。

- 10 VAD 检测的时间对 VAD 的效果影响比较大。VAD 的检测时间应该根据用户在连续会话情景下的发音时间来决定，不同的用户发音习惯不同。这就需要动态的调整 VAD 的执行时间，具体可根据用户使用音箱的习惯动态的优化 VAD 检测的时间。这些习惯统计出的时间包括：用户从设备唤醒成功到发出语音指令的平均时间，以及从启动 VAD 到用户发出语音指令的平均时间。

进一步地，时间计算模块根据第一平均时间和第二平均时间计算得到指定的检测时间，可包括：

- 15 根据 $T4 = T3 + (T2 - T3) / 2$

计算得到所述指定的检测时间 T4；其中，T2 为第一平均时间 T1 和预设冗余时间之和，T3 为第二平均时间。

语音识别模块，还用于在对检测到的语音信号进行 ASR 后，如果未识别出文本内容，则结束本次会话；否则，触发语义解析模块对识别出的文本内容进行语义解析。

- 20 语义解析模块，用于在对识别出的文本内容进行语义解析后，如果解析得到的语义未进入任一预置的领域，或者解析得到的语义明确为结束本次会话，则结束本次会话；否则，根据解析得到的语义所进入的领域生成语音指令，并触发指令执行模块根据语音指令控制相应设备执行操作。

- 25 本发明实施例提供的人机会话的处理系统，在设备完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求；如果确定用户有再次输入语音指令的需求，则启动语音活动检测 VAD；否则，结束本次会话。进一步地，在连续会话过程中，根据各环节执行情况，预先设置结束会话的条件，并在判定条件形成后，结束本次会话，实现连续会话的完整流程。

下面通过多个实施例来进一步说明本申请的技术方案。

实施例一

基于上述根据语音指令内容判定是否进行连续会话的方案思想，如图 4a 所示，其为本发明实施例示出的人机会话的处理方法流程图一，该方法的执行主体为图 3 中所示的人机会话的处理装置。如图 4a 所示，该人机会话的处理方法包括如下步骤：

- 5 S410，在设备完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求。

在现有的人机会话场景中，用户在输入一条语音指令后，往往还想基于该语音指令的内容以及执行结果，再次输入新的语音指令，例如，上一条语音输入是“帮我搜索变形金刚”，系统搜索完毕后，通过屏幕展示搜索结果列表，此时用户很可能会再输入“播放第一个”搜索结果的语音指令。

10

为了在满足用户连续会话需求的前提下，提高人机会话效率，本实施例中在设备每完成一条语音指令后，会对上一条已完成的语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求。从而确定是否要启动连续会话的流程。

领域判断，即判断某一领域指令完成后是否需要进入连续会话状态。连续会话的触发需要根据用户上一个语音指令的需求类型及执行情况进行判定。

15

判断连续会话的依据来自于但不局限于以下几个原则：

1.是否明确不需要连续会话；如控制类指令（“我要把灯打开”）基本上是用户的一次操作，这类场景下不需要连续会话；

2.是否经常有上下文诉求：如天气领域（“今天天气怎么样？”——“明天呢”）；

20

3.是否是一个连续多指令需求：如电影查询（“帮我搜索变形金刚”——“播放第一个”——“全屏播放”）。

基于上述原则来判定是否需要进行连续会话，如果需要进行连续会话，即认为用户有基于上一条语音指令而再次输入语音指令的需求。

如果确定用户有再次输入语音指令的需求，则执行步骤 S420，否则执行 S430。

25

S420，启动语音活动检测 VAD；

S430，结束本次会话。

当确定用户有再次输入语音指令的需求时，可以控制设备侧开启 VAD 的检测流程，以采集用户可能输入的语音指令，并上传给服务器侧的人机会话的处理装置进行识别处理。

30

当确定用户没有再次输入语音指令的需求时，可以结束本次会话，控制设备进入待

机状态。

本实施例中的对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求的判定过程，可在服务器侧处理完成，即通过服务器控制前端设备是否开启连读会话流程。

5 进一步地，上述方法还包括：

在启动 VAD 后，如果在指定的检测时间内没有检测到语音信号，则结束本次会话；否则，对检测到的语音信号进行自动语音识别 ASR。

当确定用户有再次输入语音指令的需求时，服务器侧控制设备侧启动 VAD 流程，并将麦克风检测的语音信号传送至服务器侧进行解析。如果在指定的检测时间内没有检测到人声输入，则认为会话结束，VAD 流程关闭；如在指定的检测时间内检测到人声输入，则将检测到的声音上传到服务器端进行 ASR 处理。

人声检测需要屏蔽掉噪音干扰，稳态的噪音比较好识别和屏蔽，如频率稳定不变的空调噪音，电机噪音；但是动态的噪音比较难屏蔽，如歌声，电视机噪音等频率变化较大且包含人声录音的噪音。因此 VAD 的检测时间对 VAD 的效果影响比较大。

15 VAD 的检测时间应该根据用户在连续会话下的发音时间来决定，不同的用户发音习惯不同。动态的调整 VAD 是根据用户使用音箱的习惯动态的优化 VAD 检测的时间，优化的策略如下：

统计用户在各次会话过程中，用户从设备唤醒成功到发出语音指令的第一平均时间；统计用户在各次会话过程中，从启动 VAD 到用户发出语音指令的第二平均时间；根据第一平均时间和第二平均时间计算得到指定的检测时间。

具体的，在用户的各次会话过程中，实时计算用户在正常唤醒设备的场景下，从设备唤醒成功到发出指令发出的平均时间 T1（第一平均时间）。

25 设置初始连续会话下的 VAD 的检测时间 $T2=T1+3(s)$ ，由于连续会话状态下，用户的发音一般会较慢于设备正常唤醒下的发音，顾初始的连续会话的检测时间默认为在用户唤醒设备状态下的评估时间基础上增加 3 秒（冗余时间）。

设置的初始连续会话下的 VAD 的检测时间为一个较长的容错时间，需要根据用户在连续会话状态的实际平均发音时间进行收敛，顾在用户的各次会话过程中，还需要实时计算连续会话状态下，从启动 VAD 到用户发出语音指令的平均时间 T3（第二平均时间）。

30 在初始连续会话下的 VAD 的检测时间 T2 基础上，利用用户在连续会话状态下的实

际平均发音时间 T3 对 T2 进行修正，从而可得到比较合理的指定的检测时间 T4。

进一步地，根据第一平均时间和第二平均时间计算得到指定的检测时间的处理步骤，可包括：

通过如下公式：

$$5 \qquad \qquad \qquad T4=T3+(T2-T3)/2 \qquad \qquad \qquad (1)$$

计算得到上述的指定的检测时间 T4；其中，T2 为第一平均时间 T1 和预设冗余时间之和，T3 为所述第二平均时间。

进一步地，上述方法还包括：

在对检测到的语音信号进行 ASR 后，如果未识别出文本内容，则结束本次会话；否则，对识别出的文本内容进行语义解析。

在 ASR 阶段依然有判断连续会话终止的条件。当有噪音在 VAD 环节被误判为人声输入时，ASR 可能无法识别此段声音，此时识别结果可能返回为空，故可以认为在连续会话场景下，ASR 操作结果返回为空时此次会话结束。但也存在用户的语音指令由于客观因素影响导致 ASR 无法识别返回为空的情况，故在 ASR 环节终止本次会话的同时，

15 可通过 TTS 提示用户设备没有听清，有需要请再次唤醒。

当 ASR 执行完成后，识别出文本内容，可将文本内容继续传输至下一处理环节，即进行语义解析。

进一步地，上述方法还包括：

在对识别出的文本内容进行语义解析后，如果解析得到的语义未进入任一预置的领域，或者解析得到的语义明确为结束本次会话，则结束本次会话；否则，根据解析得到的语义所进入的领域生成语音指令，并根据语音指令控制相应设备执行操作。

20 通过理解用户语音指令的实际语义可进行连续会话终止的判断。

语义解析可分为三个环节，领域识别，意图识别，执行逻辑判断，其中，用以判断连续对话是否终止的判断基本上在领域识别环节就可完成，具体可包括但不限于如下两种判断条件：

25 当经 ASR 解析的文本中的语义未落到任何一个领域（领域库中预置有多个不同的领域）时，会话终止。如果连续会话场景下，干扰噪音被误判进入语义解析阶段，杂乱没有逻辑的语句不容易被语义理解到某一领域中，此时会话终止；

当解析得到的语义落入到“停止词”领域时，表示用户明确使用指令来停止连续会话的场景，此时会话终止。“停止词”领域中的语料如“好的”、“谢谢”、“没事儿

了”……。

当语义落入除“停止词”以外的其他领域，正常执行该领域指令，即根据落入到的领域中的具体意图，生成语音指令，然后根据语音指令控制设备执行相应操作。

指令执行成功后，可重新回到起始步骤 S410 进行上一条语音指令的领域判断，确定

5 用户是否有基于上一条语音指令而再次输入语音指令的需求，以及启动 VAD 流程。

指令执行失败，则可跳过步骤 S410，直接开启 VAD，从步骤 S420 开始进入连续对话情景。

本发明实施例提供的人机会话的处理方法，在设备完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令

10 的需求；如果确定用户有再次输入语音指令的需求，则启动语音活动检测；否则，结束本次会话，从而有效的提高在连续会话场景下的会话处理效率。

进一步地，在连续会话执行过程中，通过对各个环节预制结束本次会话的判断条件，可在条件满足时，结束本次会话流程，保证连续会话的完整性。

15 实施例二

图 4b 为本发明实施例提供的另一种人机会话的处理方法，该方法在图 4a 所示方法的基础上，进行了少许改变。如图 4b 所示，该人机会话的处理方法包括如下步骤：

S440，对所接收的语音指令的内容进行识别。

在现有的人机会话场景中，用户在输入一条语音指令后，往往还想基于该语音指令

20 的内容以及执行结果，再次输入新的语音指令，例如，上一条语音输入是“帮我搜索变形金刚”，系统搜索完毕后，通过屏幕展示搜索结果列表，此时用户很可能会再输入“播放第一个”搜索结果的语音指令。

为了在满足用户连续会话需求的前提下，提高人机会话效率，本实施例中在设备每收到语音指令后，先会对该语音指令的内容进行识别，具体判识别程可参见步骤 S410

25 的相关内容。

S450，判断用户是否有再次输入语音指令的需求；

在对语音指令的内容进行识别后，判断用于是否存在基于该语音指令再次输入语音指令的需求。具体判断过程可参见步骤 S410 的相关内容。

例如，可针对不同用户（声纹识别）的个性化习惯对用户是否进行再次进行语音输入的结果进行历史数据的统计，根据统计结果计算用户再次输入语音指令的概率。如果

30

得到的概率大于预设的概率阈值，则判定用户有再次输入语音指令的需求；否则，判定用户没有再次输入语音指令的需求。

S460，根据判断结果，执行人机会话操作。

5 通过用户输入的语音指令的内容，判断用户是否有再次输入语音的需求，并根据判断结果指导后续的处理操作。例如，如果确定用户有再次输入语音指令的需求，则在设备完成本次语音指令后执行启动语音活动检测，以实现连续会话；否则，在设备完成本次语音指令后结束本次会话。

另外，需要说明的是，实施例一中的方法中的步骤也可在本实施例中的方法步骤中执行，在此对步骤原理不做赘述。

10 本发明实施例提供的人机会话的处理方法，对所接收的语音指令的内容进行识别，判断用户是否有再次输入语音指令的需求；然后根据判断结果，执行人机会话操作，从而有效的提高在连续会话场景下的会话处理效率。

实施例三

15 如图 5a 所示，为本发明实施例的人机会话的处理装置结构图一，该人机会话的处理装置可用于执行如图 4a 所示的方法步骤，其包括：

指令识别模块 510，用于在设备完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求；

20 语音检测模块 520，用于如果确定用户有再次输入语音指令的需求，则启动语音活动检测 VAD；否则，结束本次会话。

进一步地，如图 5b 所示，上述装置还包括语音识别模块 530；

语音检测模块 520，还用于在启动 VAD 后，如果在指定的检测时间内没有检测到语音信号，则结束本次会话；否则，触发语音识别模块 530 对检测到的语音信号进行自动语音识别 ASR。

25 进一步地，如图 5b 所示，上述装置还包括时间计算模块 540，用于：统计用户在各次会话过程中，用户从设备唤醒成功到发出语音指令的第一平均时间；统计用户在各次会话过程中，从启动 VAD 到用户发出语音指令的第二平均时间；根据第一平均时间和第二平均时间计算得到上述指定的检测时间。

30 进一步地，上述时间计算模块 540 根据第一平均时间和第二平均时间计算得到指定的检测时间，包括：

根据 $T4 = T3 + (T2 - T3) / 2$

计算得到指定的检测时间 $T4$ ；其中， $T2$ 为第一平均时间 $T1$ 和预设冗余时间之和， $T3$ 为第二平均时间。

进一步地，如图 5b 所示，上述装置还包括语义解析模块 550；

5 语音识别模块 540，还用于在对检测到的语音信号进行 ASR 后，如果未识别出文本内容，则结束本次会话；否则，触发语义解析模块 550 对识别出的文本内容进行语义解析。

进一步地，如图 5b 所示，上述装置还包括指令执行模块 560；

语义解析模块 550，还用于在对识别出的文本内容进行语义解析后，如果解析得到的语义未进入任一预置的领域，或者解析得到的语义明确为结束本次会话，则结束本次会话；否则，根据解析得到的语义所进入的领域生成语音指令，并触发指令执行模块 560 根据语音指令控制相应设备执行操作。

本发明实施例提供的人机会话的处理装置，在设备完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求；如果确定用户有再次输入语音指令的需求，则启动语音活动检测 VAD；否则，结束本次会话，从而有效的提高在连续会话场景下的会话处理效率。

进一步地，在连续会话执行过程中，通过对各个环节预制结束本次会话的判断条件，可在条件满足时，结束本次会话流程，保证连续会话的完整性。

20 实施例四

如图 6 所示，为本发明实施例提供的人机会话的处理装置结构图二，该人机会话的处理装置可用于执行如图 4b 所示的方法步骤，其包括：

内容识别模块 610，用于对所接收的语音指令的内容进行识别；

需求判断模块 620，用于判断用户是否有再次输入语音指令的需求；

25 执行操作模块 630，用于根据判断结果，执行人机会话操作。

本发明实施例提供的人机会话的处理装置，对所接收的语音指令的内容进行识别，判断用户是否有再次输入语音指令的需求；然后根据判断结果，执行人机会话操作，从而有效的提高在连续会话场景下的会话处理效率。

30 实施例五

前面实施例三描述了人机会话的处理装置的整体架构，该装置的功能可借助一种电子设备实现完成，如图 7 所示，其为本发明实施例的电子设备的结构示意图，具体包括：存储器 710 和处理器 720。

存储器 710，用于存储程序。

5 除上述程序之外，存储器 710 还可被配置为存储其它各种数据以支持在电子设备上的操作。这些数据的示例包括用于在电子设备上操作的任何应用程序或方法的指令，联系人数据，电话簿数据，消息，图片，视频等。

存储器 710 可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器（SRAM），电可擦除可编程只读存储器（EEPROM），可擦除可编程只读存储器（EPROM），可编程只读存储器（PROM），只读存储器（ROM），磁存储器，快闪存储器，磁盘或光盘。

处理器 720，耦合至存储器 710，用于执行存储器 710 中的程序，以用于：

在设备完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求；

15 如果确定用户有再次输入语音指令的需求，则启动语音活动检测 VAD；否则，结束本次会话。

上述的具体处理操作已经在前面实施例中进行了详细说明，在此不再赘述。

进一步，如图 7 所示，电子设备还可以包括：通信组件 730、电源组件 740、音频组件 750、显示器 760 等其它组件。图 7 中仅示意性给出部分组件，并不意味着电子设备只包括图 7 所示组件。

通信组件 730 被配置为便于电子设备和其他设备之间有线或无线方式的通信。电子设备可以接入基于通信标准的无线网络，如 WiFi，2G 或 3G，或它们的组合。在一个示例性实施例中，通信组件 730 经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，通信组件 730 还包括近场通信（NFC）模块，以促进短程通信。例如，在 NFC 模块可基于射频识别（RFID）技术，红外数据协会（IrDA）技术，超宽带（UWB）技术，蓝牙（BT）技术和其他技术来实现。

电源组件 740，为电子设备的各种组件提供电力。电源组件 740 可以包括电源管理系统，一个或多个电源，及其他与为电子设备生成、管理和分配电力相关联的组件。

音频组件 750 被配置为输出和/或输入音频信号。例如，音频组件 750 包括一个麦克风（MIC），当电子设备处于操作模式，如呼叫模式、记录模式和语音识别模式时，麦

麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器 710 或经由通信组件 730 发送。在一些实施例中，音频组件 750 还包括一个扬声器，用于输出音频信号。

- 5 显示器 760 包括屏幕，其屏幕可以包括液晶显示器（LCD）和触摸面板（TP）。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与触摸或滑动操作相关的持续时间和压力。

实施例六

- 10 前面实施例四描述了人机会话的处理装置的整体架构，该装置的功能可借助一种电子设备实现完成，如图 8 所示，其为本发明实施例的电子设备的结构示意图，具体包括：存储器 810 和处理器 820。

存储器 810，用于存储程序。

- 15 除上述程序之外，存储器 810 还可被配置为存储其它各种数据以支持在电子设备上的操作。这些数据的示例包括用于在电子设备上操作的任何应用程序或方法的指令，联系人数据，电话簿数据，消息，图片，视频等。

- 20 存储器 810 可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器（SRAM），电可擦除可编程只读存储器（EEPROM），可擦除可编程只读存储器（EPROM），可编程只读存储器（PROM），只读存储器（ROM），磁存储器，快闪存储器，磁盘或光盘。

处理器 820，耦合至存储器 810，用于执行存储器 810 中的程序，以用于：

对所接收的语音指令的内容进行识别；

判断用户是否有再次输入语音指令的需求；

根据判断结果，执行人机会话操作。

- 25 上述的具体处理操作已经在前面实施例中进行了详细说明，在此不再赘述。

进一步，如图 8 所示，电子设备还可以包括：通信组件 830、电源组件 840、音频组件 850、显示器 860 等其它组件。图 8 中仅示意性给出部分组件，并不意味着电子设备只包括图 8 所示组件。

- 30 通信组件 830 被配置为便于电子设备和其他设备之间有线或无线方式的通信。电子设备可以接入基于通信标准的无线网络，如 WiFi，2G 或 3G，或它们的组合。在一个示

例性实施例中，通信组件 830 经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，通信组件 830 还包括近场通信（NFC）模块，以促进短程通信。例如，在 NFC 模块可基于射频识别（RFID）技术，红外数据协会（IrDA）技术，超宽带（UWB）技术，蓝牙（BT）技术和其他技术来实现。

5 电源组件 840，为电子设备的各种组件提供电力。电源组件 840 可以包括电源管理系统，一个或多个电源，及其他与为电子设备生成、管理和分配电力相关联的组件。

10 音频组件 850 被配置为输出和/或输入音频信号。例如，音频组件 850 包括一个麦克风（MIC），当电子设备处于操作模式，如呼叫模式、记录模式和语音识别模式时，麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器 810 或经由通信组件 830 发送。在一些实施例中，音频组件 850 还包括一个扬声器，用于输出音频信号。

15 显示器 860 包括屏幕，其屏幕可以包括液晶显示器（LCD）和触摸面板（TP）。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与触摸或滑动操作相关的持续时间和压力。

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：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

20 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。

25

权利要求书

1.一种人机会话的处理方法，其特征在于，包括：

在设备完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求；

5 如果确定用户有再次输入语音指令的需求，则启动语音活动检测；否则，结束本次会话。

2.根据权利要求1所述的方法，其特征在于，所述方法还包括：

在启动所述语音活动检测后，如果在指定的检测时间内没有检测到语音信号，则结束本次会话；否则，对检测到的语音信号进行自动语音识别。

10 3.根据权利要求2所述的方法，其特征在于，所述方法还包括：

统计所述用户在各次会话过程中，用户从设备唤醒成功到发出语音指令的第一平均时间；

统计所述用户在各次会话过程中，从启动所述语音活动检测到用户发出语音指令的第二平均时间；

15 根据所述第一平均时间和所述第二平均时间计算得到所述指定的检测时间。

4.根据权利要求3所述的方法，其特征在于，所述根据所述第一平均时间和所述第二平均时间计算得到所述指定的检测时间，包括：

根据 $T4=T3+(T2-T3)/2$

计算得到所述指定的检测时间 $T4$ ；其中，所述 $T2$ 为所述第一平均时间 $T1$ 和预设
20 冗余时间之和，所述 $T3$ 为所述第二平均时间。

5.根据权利要求2所述的方法，其特征在于，所述方法还包括：

在对检测到的所述语音信号进行所述自动语音识别后，如果未识别出文本内容，则结束本次会话；否则，对识别出的文本内容进行语义解析。

6.根据权利要求5所述的方法，其特征在于，所述方法还包括：

25 在对识别出的所述文本内容进行语义解析后，如果解析得到的语义未进入任一预置的领域，或者解析得到的语义明确为结束本次会话，则结束本次会话；否则，根据解析得到的语义所进入的领域生成语音指令，并根据语音指令控制相应设备执行操作。

7.一种人机会话的处理方法，其特征在于，包括：

对所接收的语音指令的内容进行识别；

30 判断用户是否有再次输入语音指令的需求；

根据判断结果，执行人机会话操作。

8.根据权利要求7所述的方法，其特征在于，所述根据判断结果，执行人机会话操作包括：

如果确定用户有再次输入语音指令的需求，则在设备完成本次语音指令后执行启动
5 语音活动检测；否则，在设备完成本次语音指令后结束本次会话。

9.根据权利要求7所述的方法，其特征在于，所述判断用户是否有再次输入语音指令的需求包括：

计算用户再次输入语音指令的概率；

如果所述概率大于预设的概率阈值，则判定用户有再次输入语音指令的需求；否则，
10 判定用户没有再次输入语音指令的需求。

10.一种人机会话的处理装置，其特征在于，包括：

指令识别模块，用于在设备完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求；

语音检测模块，用于如果确定用户有再次输入语音指令的需求，则启动语音活动检测；
15 否则，结束本次会话。

11.一种人机会话的处理装置，其特征在于，包括：

内容识别模块，用于对所接收的语音指令的内容进行识别；

需求判断模块，用于判断用户是否有再次输入语音指令的需求；

执行操作模块，用于根据判断结果，执行人机会话操作。

20 12.一种电子设备，其特征在于，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

在设备完成上一条语音指令后，对上一条语音指令的内容进行识别，确定用户是否有基于上一条语音指令而再次输入语音指令的需求；

25 如果确定用户有再次输入语音指令的需求，则启动语音活动检测；否则，结束本次会话。

13.一种电子设备，其特征在于，包括：

存储器，用于存储程序；

处理器，耦合至所述存储器，用于执行所述程序，以用于：

30 对所接收的语音指令的内容进行识别；

判断用户是否有再次输入语音指令的需求；
根据判断结果，执行人机会话操作。

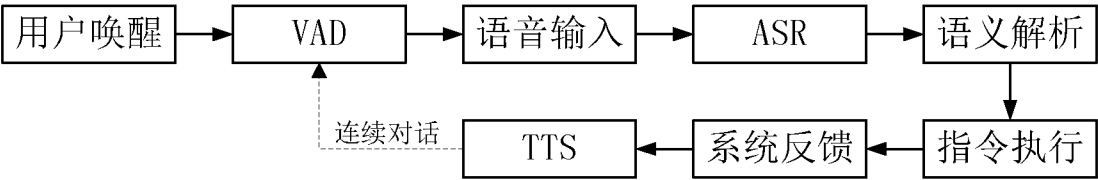


图 1

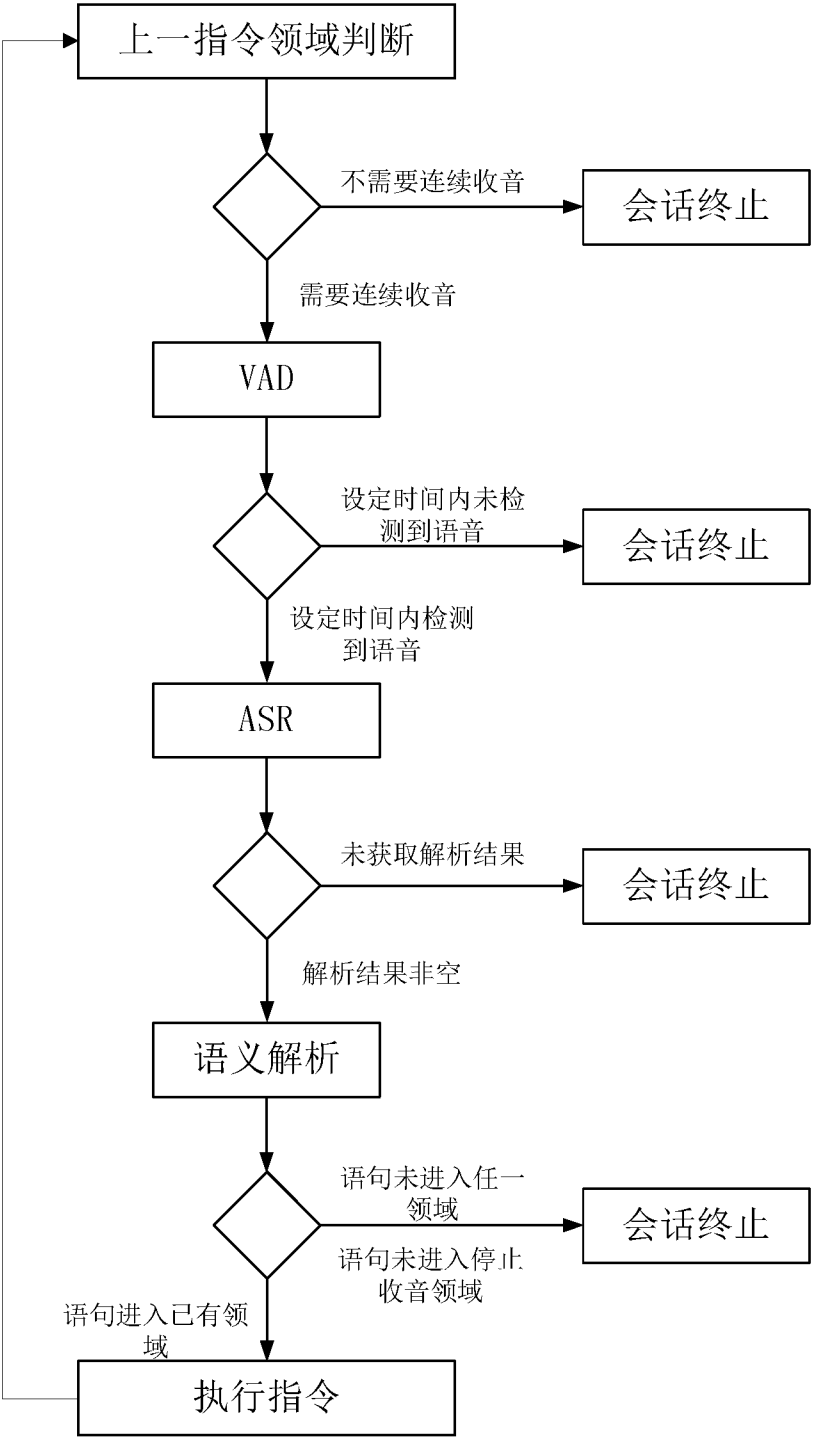


图 2

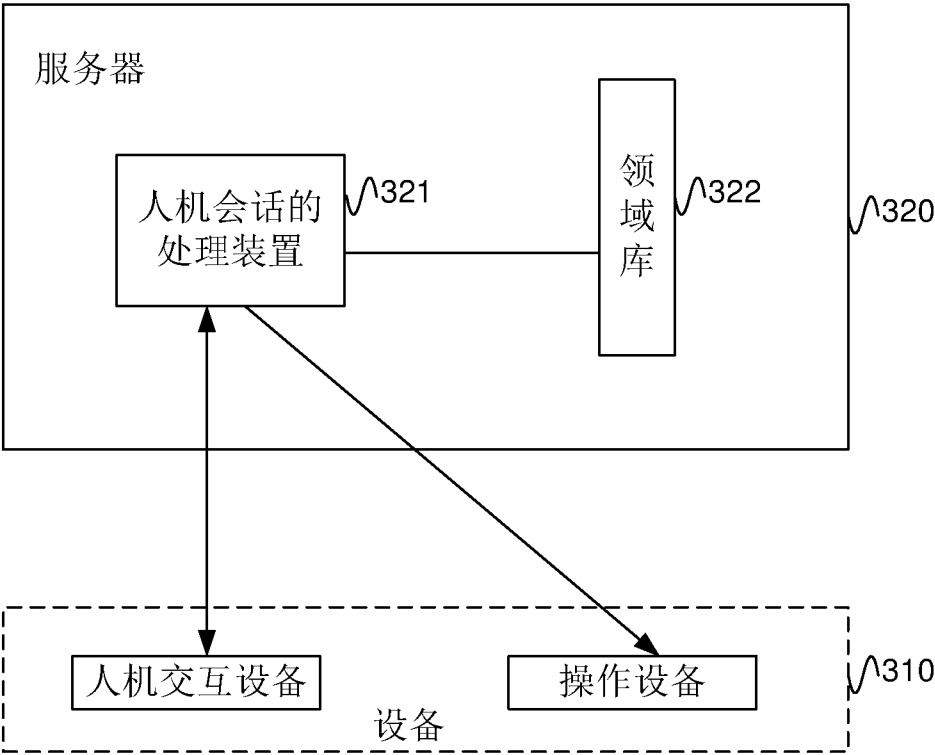


图 3

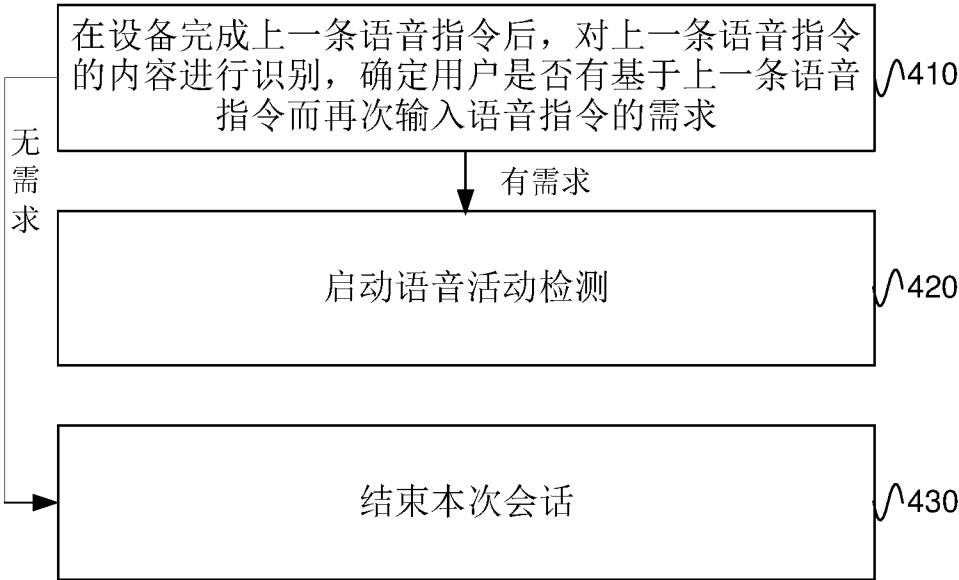


图 4a

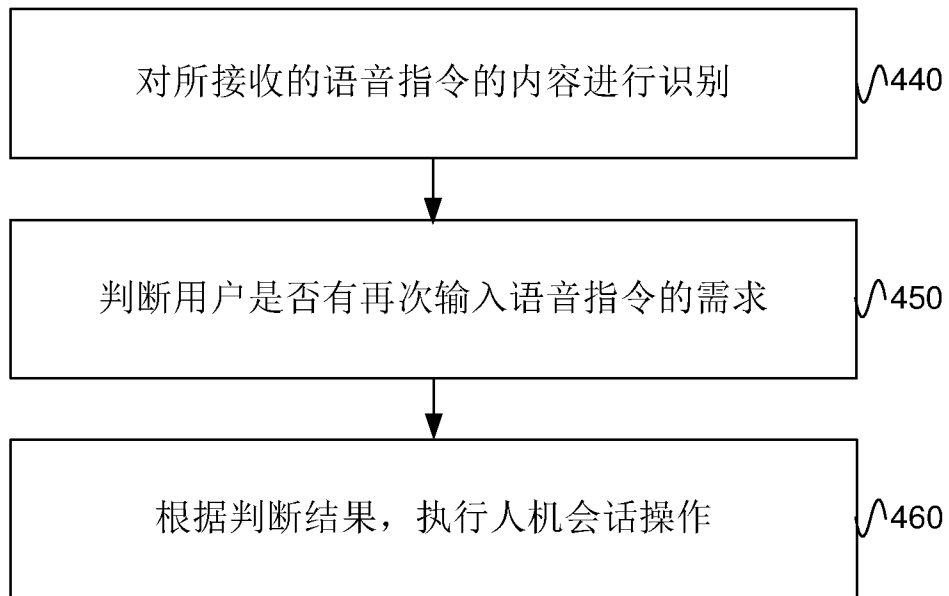


图 4b

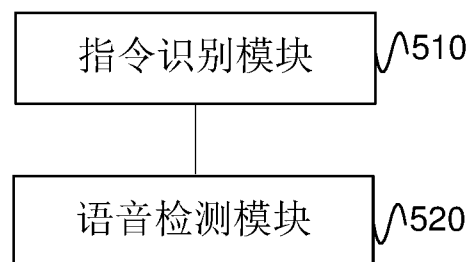


图 5a

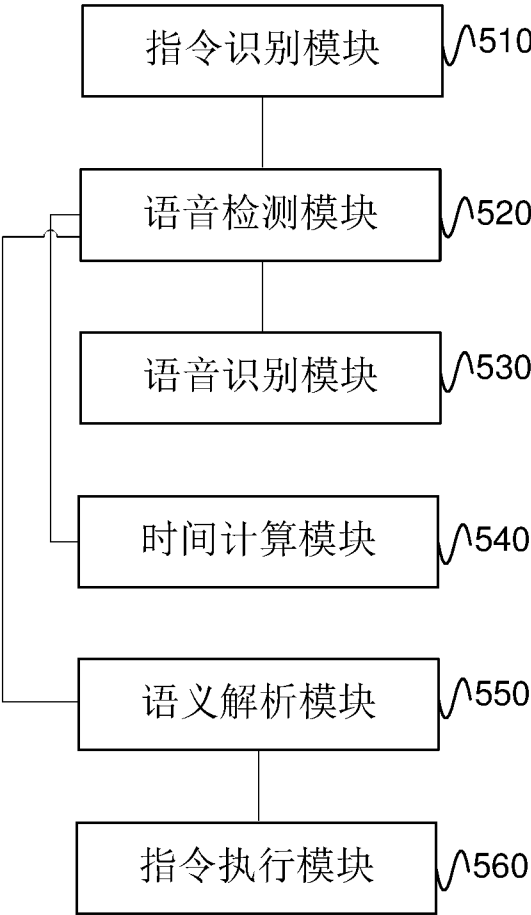


图 5b

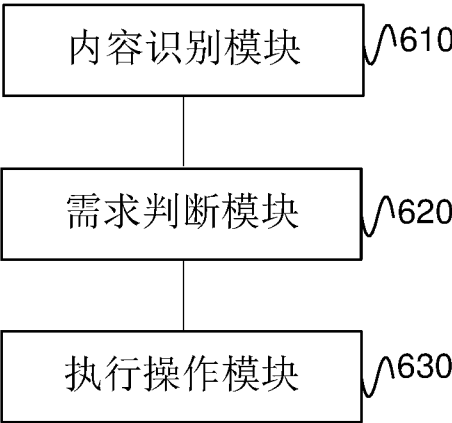


图 6

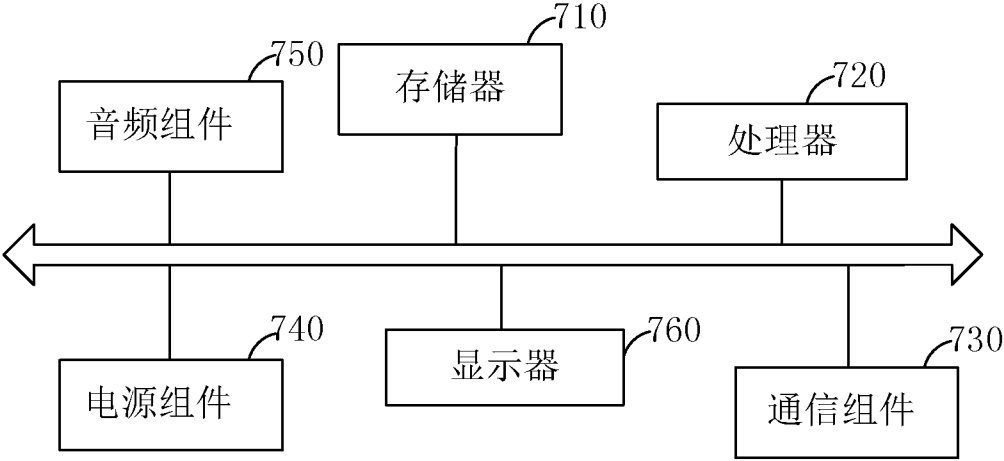


图 7

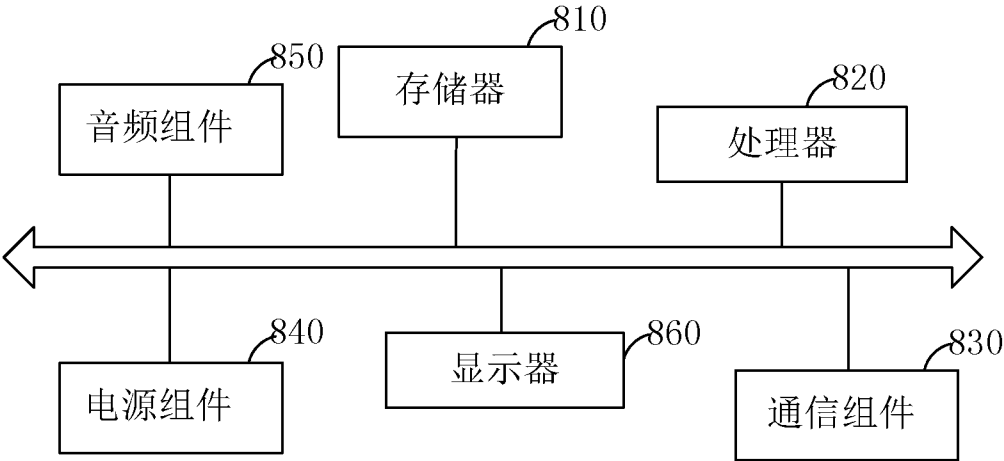


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/093225

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/22(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC: 语音, 控制, 指令, 命令, 声控, 唤醒, 开启, 休眠, 重复, 检测, 需要, 第二条, 第二次, 再次, 再一次, 下一次, 下一条, 上一条, 上次, 上下文, 多条, 多次, 多指令, 连续, 继续, 后续, 会话, 人机交互, 识别, 判断, 推测, 预测, 预估, 确定, 还需, 仍需, 输入, 还要, 还有, 需求, 意图, 避免, 输入, voice, speech, acoustic, control, command, awake, wake?up, sleep, repeat, need, second, next, context, more, continue, recognition, predict, pre?estimate, determine, intention, avoid, input, VAD

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 106205612 A (BEIJING GUANGNIAN WUXIAN SCIENCE AND TECHNOLOGY CO., LTD.) 07 December 2016 (2016-12-07) description, paragraphs [0059]-[0077] and [0105]-[0113]	7, 9, 11
Y	CN 106653021 A (SHANGHAI ZHIZHEN INTELLIGENT NETWORK TECHNOLOGY CO., LTD.) 10 May 2017 (2017-05-10) description, paragraphs [0039]-[0070], [0101]-[0120] and [0143]	1-6, 8, 10, 12-13
Y	CN 106205612 A (BEIJING GUANGNIAN WUXIAN SCIENCE AND TECHNOLOGY CO., LTD.) 07 December 2016 (2016-12-07) description, paragraphs [0059]-[0077] and [0105]-[0113]	1-6, 8, 10, 12-13
A	CN 104505093 A (FOSHAN SHUNDE MIDEA ELECTRICAL HEATING APPLIANCES MANUFACTURING CO., LTD. ET AL.) 08 April 2015 (2015-04-08) entire document	1-13
A	CN 106205615 A (WANG, ZHENGRONG) 07 December 2016 (2016-12-07) entire document	1-13

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

03 September 2018

Date of mailing of the international search report

20 September 2018

Name and mailing address of the ISA/CN

State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/093225

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2014278435 A1 (NUANCE COMMUNICATIONS, INC.) 18 September 2014 (2014-09-18) entire document	1-13

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/093225

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	106205612	A	07 December 2016	None	
CN	106653021	A	10 May 2017	None	
CN	104505093	A	08 April 2015	None	
CN	106205615	A	07 December 2016	None	
US	2014278435	A1	18 September 2014	None	

国际检索报告

国际申请号

PCT/CN2018/093225

A. 主题的分类

G10L 15/22 (2006.01) i

按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

G10L

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))

CNPAT, CNKI, WPI, EPODOC: 语音, 控制, 指令, 命令, 声控, 唤醒, 开启, 休眠, 重复, 检测, 需要, 第二条, 第二次, 再次, 再一次, 下一条, 上一条, 上次, 上下文, 多条, 多次, 多指令, 连续, 继续, 后续, 会话, 人机交互, 识别, 判断, 推测, 预测, 预估, 确定, 还需, 仍需, 输入, 还要, 还有, 需求, 意图, 避免, 输入, voice, speech, acoustic, control, command, awake, wake?up, sleep, repeat, need, second, next, context, more, continue, recognition, predict, pre?estimate, determine, intention, avoid, input, VAD

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 106205612 A (北京光年无限科技有限公司) 2016年 12月 7日 (2016 - 12 - 07) 说明书第[0059]-[0077], [0105]-[0113]段	7, 9, 11
Y	CN 106653021 A (上海智臻智能网络科技股份有限公司) 2017年 5月 10日 (2017 - 05 - 10) 说明书第[0039]-[0070], [0101]-[0120], [0143]段	1-6, 8, 10, 12-13
Y	CN 106205612 A (北京光年无限科技有限公司) 2016年 12月 7日 (2016 - 12 - 07) 说明书第[0059]-[0077], [0105]-[0113]段	1-6, 8, 10, 12-13
A	CN 104505093 A (佛山市顺德区美的电热器制造有限公司等) 2015年 4月 8日 (2015 - 04 - 08) 全文	1-13
A	CN 106205615 A (王峥嵘) 2016年 12月 7日 (2016 - 12 - 07) 全文	1-13

☒ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2018年 9月 3日

国际检索报告邮寄日期

2018年 9月 20日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局 (ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10) 62019451

受权官员

毕雅超

电话号码 86-(10)-53961777

C. 相关文件		
类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	US 2014278435 A1 (NUANCE COMMUNICATIONS, INC.) 2014年 9月 18日 (2014 - 09 - 18) 全文	1-13

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/093225

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	106205612	A	2016年 12月 7日	无	
CN	106653021	A	2017年 5月 10日	无	
CN	104505093	A	2015年 4月 8日	无	
CN	106205615	A	2016年 12月 7日	无	
US	2014278435	A1	2014年 9月 18日	无	

表 PCT/ISA/210 (同族专利附件) (2015年1月)