



(12) 发明专利申请

(10) 申请公布号 CN 112969995 A

(43) 申请公布日 2021. 06. 15

(21) 申请号 201980071834.2

(22) 申请日 2019.10.28

(30) 优先权数据

10-2018-0130007 2018.10.29 KR

(85) PCT国际申请进入国家阶段日

2021.04.29

(86) PCT国际申请的申请数据

PCT/KR2019/014292 2019.10.28

(87) PCT国际申请的公布数据

W02020/091350 EN 2020.05.07

(71) 申请人 三星电子株式会社

地址 韩国京畿道水原市

(72) 发明人 金灿佑 李璟懋 卢在英 张东韩

曹根硕 邢知远

(74) 专利代理机构 北京铭硕知识产权代理有限公司 11286

代理人 张思宇 张军

(51) Int.Cl.

G06F 3/16 (2006.01)

G06F 3/01 (2006.01)

G06F 21/32 (2006.01)

G10L 17/24 (2006.01)

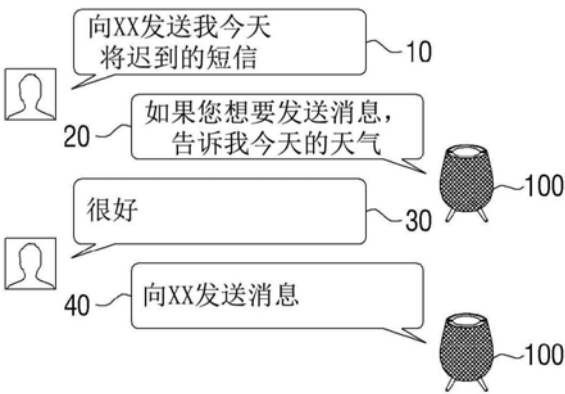
权利要求书2页 说明书24页 附图11页

(54) 发明名称

电子装置及其控制方法

(57) 摘要

提供了一种电子装置和控制方法,包括输入接口、通信接口、包括至少一个命令的存储器、以及至少一个处理器,其中,所述至少一个处理器被配置为:控制电子装置并执行所述至少一个命令以进行以下操作:通过输入接口接收用户语音,通过对所述用户语音进行分析来确定所述用户语音是否是是需要用户确认的任务相关的语音,当确定所述用户语音是是需要用户确认的任务相关的语音时生成用于用户确认的询问,以及当通过输入接口输入与所述询问相应的用户响应时,执行与所述用户语音相应的任务。实施例可使用根据机器学习、神经网络和深度学习算法中的至少一个学习的人工智能模型。



1. 一种电子装置,包括:

输入接口;

通信接口;

存储器,包括至少一个命令;以及

至少一个处理器,连接到输入接口、通信接口和存储器,并且被配置为控制电子装置,

其中,所述至少一个处理器执行所述至少一个命令以被配置为:

通过输入接口接收用户语音,

通过对所述用户语音进行分析来识别所述用户语音是否是与需要用户确认的任务相关的语音,

基于确定所述用户语音是与需要用户确认的任务相关的语音,获得针对用户确认的询问,以及

基于通过输入接口输入了与所述询问相应的用户响应,执行与所述用户语音相应的任务。

2. 根据权利要求1所述的电子装置,其中,所述询问是与所述与用户语音相应的任务不相关的询问。

3. 根据权利要求1所述的电子装置,其中,所述至少一个处理器执行所述至少一个命令以进一步被配置为:

通过对所述用户语音进行分析来识别与所述用户语音相关的任务和用于执行该任务的实体,

基于识别出的任务和实体获得所述用户语音的安全分数,以及

基于所述用户语音的所述安全分数来确定所述用户语音是否是与需要用户确认的任务相关的语音。

4. 根据权利要求3所述的电子装置,其中,所述至少一个处理器执行所述至少一个命令以进一步被配置为:

基于在所述用户语音的所述安全分数的基础上确定所述用户语音是与需要用户确认的任务相关的语音,基于所述安全分数获得所述询问。

5. 根据权利要求4所述的电子装置,其中,所述至少一个处理器执行所述至少一个命令以进一步被配置为:

随着所述安全分数变高,获得与所述与用户语音相应的任务具有低相关性的询问。

6. 根据权利要求4所述的电子装置,其中,所述至少一个处理器执行所述至少一个命令以进一步被配置为:

基于所述安全分数是阈值或更大值,提供用于用户认证的认证消息。

7. 根据权利要求1所述的电子装置,其中,所述至少一个处理器执行所述至少一个命令以进一步被配置为:

提取包括在所述用户语音中的文本中的至少一个文本,以及

获得用于诱导针对提取出的所述至少一个文本的用户话语的询问。

8. 根据权利要求1所述的电子装置,其中,所述至少一个处理器执行所述至少一个命令以进一步被配置为:

通过对使用电子装置的用户的声音进行分析,获得与另一用户的声音特征信息区分开的关于文本的声音特征信息,以及

将所述关于所述文本的声音特征信息存储在存储器中。

9. 根据权利要求8所述的电子装置,其中,所述至少一个处理器执行所述至少一个命令以进一步被配置为:

基于确定所述用户声音是与需要用户确认的任务相关的声音,获得包括存储的所述文本的询问,以及

基于在所述用户响应中获得了所述关于所述文本的声音特征信息,执行与所述用户声音相应的任务。

10. 根据权利要求1所述的电子装置,其中,需要用户确认的任务包括汇款任务、产品购买任务、电子邮件发送任务、消息发送任务或电话呼叫任务中的至少一个。

11. 一种电子装置的控制方法,包括:

接收用户声音;

通过对所述用户声音进行分析来识别所述用户声音是否是是需要用户确认的任务相关的声音;

基于确定所述用户声音是与需要用户确认的任务相关的声音,获得针对用户确认的询问;以及

基于输入了与所述询问相应的用户响应,执行与所述用户声音相应的任务。

12. 根据权利要求11所述的控制方法,其中,所述询问是与所述与用户声音相应的任务不相关的询问。

13. 根据权利要求11所述的控制方法,其中,确定所述用户声音是与需要用户确认的任务相关的声音的步骤包括:

通过对所述用户声音进行分析来识别与所述用户声音相关的任务和用于执行该任务的实体;

基于识别出的任务和实体获得所述用户声音的安全分数;以及

基于所述用户声音的所述安全分数来确定所述用户声音是否是是需要用户确认的任务相关的声音。

14. 根据权利要求13所述的控制方法,

其中,询问是基于所述安全分数生成的,并且

其中,所述安全分数基于确定所述用户声音是与需要用户确认的任务相关的声音。

15. 根据权利要求14所述的控制方法,其中,随着所述安全分数变高,与所述与用户声音相应的任务具有低相关性的询问被生成。

电子装置及其控制方法

技术领域

[0001] 本公开涉及一种电子装置及其控制方法。更具体地,本公开涉及一种提供用于用户确认的问询以执行与用户语音相应的任务的电子装置及其控制方法。

背景技术

[0002] 近来,已经在各个领域中使用人工智能系统。人工智能系统是机器执行学习和决策制定的系统,并且与现有的基于规则的智能系统不同,其自身变得智能。随着更频繁地使用人工智能系统,提高了识别率,并且可更准确地理解用户的偏好。因此,现有的基于规则的智能系统已逐渐被基于深度学习的人工智能系统所取代。

[0003] 这种人工智能技术通常由机器学习(例如,深度学习)和使用机器学习的元素技术组成。机器学习是其本身对输入数据的特征进行分类/学习的算法技术,并且元素技术是使用诸如深度学习等的机器学习算法来模拟人脑的诸如识别、决策等的功能的技术,并且由诸如语言理解、视觉理解、推理/预测、知识表示、运动控制等的技术领域组成。

[0004] 应用人工智能技术的各种领域如下。语言理解是识别和应用/处理人类语言/字符的技术,并且包括自然语言处理、机器翻译、对话系统、提问与回答、语音识别/合成等。视觉理解是识别和处理诸如人类视觉的事物的技术,并且包括对象识别、对象跟踪、图像搜索、人类识别、场景理解、空间理解、图像增强等。推理/预测是决定以及逻辑推理和预测信息的技术,并且包括基于知识/概率的推断、最优化预测、基于偏好的规划、推荐等。知识表示是将人类经验信息自动化和处理为知识数据的技术,并且包括知识构建(数据创建/分类)、知识管理(数据利用)等。运动控制是控制车辆的自动驾驶和机器人的运动的技术,并且包括运动控制(导航、碰撞、驾驶)、操纵控制(行为控制)等。

[0005] 近来,电子装置也已经被配置为使用用户语音来执行各种任务。例如,电子装置可通过用户语音执行各种任务,诸如消息发送任务、电子邮件发送任务、汇款任务、外部装置控制任务、产品购买任务、内容再现任务等。

[0006] 电子装置可通过用户语音执行各种任务,但存在电子装置可通过用户无意的用户语音或除从用户以外发出的语音执行用户无意的任务的问题。作为示例,电子装置通过由用户在与另一用户对话期间发出的语音或者从外部装置(诸如,电视机(TV))输出的语音来执行用户无意的任务。

[0007] 具体地,在用户无意的语音是用于执行与用户隐私或用户安全性相关的任务的情况下,电子装置通过用户无意的语音执行错误的任务,从而导致诸如侵犯隐私或安全性漏洞的问题。因此,需要一种用于防止通过用户无意的语音执行任务的方法。

[0008] 上述信息仅作为背景技术信息被呈现,并且被用于帮助理解本公开。关于上述内容中的任何一个是否可作为关于本公开的现有技术适用,没有做出任何确定,也没有做出断言。

发明内容

[0009] 技术问题

[0010] 本公开的各方面是为了至少解决上述问题和/或缺点,并且至少提供以下描述的优点。因此,本公开的一方面在于提供一种用于提供用于用户确认的问询以执行与用户语音相应的任务的设备和方法及其控制方法。

[0011] 本公开的另一方面在于提供一种能够确定与用户语音相关的任务是否是是需要用户确认的任务并且基于确定结果生成和提供用于用户确认的问询的电子装置及其控制方法。

[0012] 另外的方面将部分地在下面的描述中阐述,并且部分地将从描述中显而易见,或者可通过实践所呈现的实施例来学习。

[0013] 问题的解决方案

[0014] 根据本公开的一方面,提供了一种电子装置。所述电子装置包括输入接口、通信接口、包括至少一个命令指令的存储器、以及连接到输入接口、通信接口和存储器并且被配置为控制电子装置的至少一个处理器,其中,所述至少一个处理器执行所述至少一个命令以被配置为:通过输入接口接收用户语音,通过对所述用户语音进行分析来确定所述用户语音是否是是需要用户确认的任务相关的语音,当确定所述用户语音是是需要用户确认的任务相关的语音时,生成用于用户确认的问询,并且当通过所述输入接口输入了与所述问询相应的用户响应时,执行与所述用户语音相应的所述任务。

[0015] 根据本公开的另一方面,提供了一种电子装置的控制方法。所述控制方法包括:接收用户语音,通过对所述用户语音进行分析来确定所述用户语音是否是是需要用户确认的任务相关的语音,当确定所述用户语音是是需要用户确认的任务相关的语音时,生成用于用户确认的问询,以及当输入与所述问询相应的用户响应时,执行与所述用户语音相应的任务。

[0016] 本发明的有益效果

[0017] 根据上述实施例,即使当用户无意的语音被输入到电子装置时,执行用于执行任务的用户确认,使得用户可保护隐私并增强安全性。

[0018] 根据以下结合附图公开了本公开的各种实施例的详细描述,本公开的其他方面、优点和显著特征对于本领域技术人员将变得显而易见。

附图说明

[0019] 通过以下结合附图的描述,本公开的某些实施例的上述和/或其他方面、特征和优点将更加明显,其中:

[0020] 图1是根据本公开的实施例的响应于用户语音执行任务的电子装置的使用示图;

[0021] 图2是示出根据本公开的实施例的电子装置的组件的框图;

[0022] 图3是示出根据本公开的实施例的电子装置的详细组件的框图;

[0023] 图4是示出根据本公开的实施例的人工智能代理系统的对话系统的框图;

[0024] 图5是用于描述根据本公开的实施例的在用户语音是是需要用户确认的任务相关的语音的情况下提供用于用户确认的问询的电子装置的控制方法的流程图;

[0025] 图6A和图6B是用于描述根据本公开的各种实施例的生成用于用户确认的问询的

示例的示图；

[0026] 图7A和图7B是用于描述根据本公开的各种实施例的生成用于用户确认的问询的示例的示图；

[0027] 图8是用于描述根据本公开的实施例的生成用于用户确认的问询的示例的示图；

[0028] 图9A和图9B是用于描述根据本公开的各种实施例的生成用于用户确认的问询的示例的示图；

[0029] 图10是示出根据本公开的实施例的由人工智能系统生成用于用户确认的问询的示例的序列图；

[0030] 图11是示出根据本公开的实施例的由人工智能系统生成用于用户确认的问询的另一示例的序列图；

[0031] 图12是示出根据本公开的实施例的由人工智能系统生成用于用户确认的问询的另一示例的序列图；以及

[0032] 图13是用于描述根据本公开的实施例的电子装置或外部服务器根据用户语音的安全分数将用户语音转换为文本的示例的序列图。

[0033] 在整个附图中，相同的附图标号将被理解为指代相同的部分、组件和结构。

具体实施方式

[0034] 提供参考附图的以下描述以帮助全面理解由权利要求及其等同物限定的本公开的各种实施例。其包括各种具体细节以帮助理解，但这些仅被认为是示例性的。因此，本领域普通技术人员将认识到的是，在不脱离本公开的范围和精神的情况下，可对本文所描述的各种实施例进行各种改变和修改。此外，为了清楚和简明，可省略对公知的功能和结构的描述。

[0035] 在以下描述和权利要求书中使用的术语和词语不限于书面含义，而是仅被用于使得能够清楚且一致地理解本公开。因此，对于本领域技术人员显而易见的是，提供本公开的各种实施例的以下描述仅出于说明目的，而不是出于限制由所附权利要求及其等同物限定的本公开的目的。

[0036] 应当理解的是，除非上下文另有明确指示，否则单数形式包括复数指代。因此，例如，对“组件表面”的引用包括对一个或更多个这样的表面的引用。

[0037] 在本公开中，表述“具有”、“可具有”、“包括”、“可包括”等指示相应特征（例如，数值、功能、操作、诸如部件的组件等）的存在，并且不排除另外的特征的存在。

[0038] 在本公开中，表述“A或B”、“A和/或B中的至少一个”、“A和/或B中的一个或更多个”等可包括一起列举的项的所有可能组合。例如，“A或B”、“A和B中的至少一个”或“A或B中的至少一个”可指示以下全部：1) 包括至少一个A的情况，2) 包括至少一个B的情况，或3) 包括至少一个A和至少一个B两者的情况。

[0039] 在本公开中使用的表述“第一”、“第二”等可指示各种组件而不考虑组件的顺序和/或重要性如何，它们将仅被用于将一个组件与其他组件区分开，并且不限制相应的组件。

[0040] 当提及任何组件（例如，第一组件）（可操作地或通信地）结合到或连接到另一组件（例如，第二组件）时，应当理解的是，任何组件直接结合到另一组件或可通过另一组件（例

如,第三组件)结合到另一组件。当提及任何组件(例如,第一组件)“直接结合”或“直接连接”到另一组件(例如,第二组件)时,应理解的是,在任何组件与另一组件(例如,第二组件)之间不存在另一组件(例如,第三组件)。

[0041] 根据情况,在本公开中使用的表述“被配置(或设置)为”可由表述“适合于”、“具有……的能力”、“被设计为”、“适于”、“被制造为”或“能够”代替。术语“被配置(或设置)为”可能不一定表示在硬件中“专门被设计为”。相反,表述“被配置为…的装置”可表示该装置可与其他装置或组件一起“执行”。例如,“被配置(或设置)为执行A、B和C的处理器”可表示用于执行相应操作的专用处理器(例如,嵌入式处理器)或可通过执行存储在存储器装置中的一个或多个软件程序来执行相应操作的通用处理器(例如,中央处理器(CPU)或应用处理器)。

[0042] 根据本公开的各种实施例的电子装置可包括例如智能电话、平板个人计算机(PC)、移动电话、视频电话、电子书阅读器、桌上型PC、膝上型PC、上网本计算机、工作站、服务器、个人数字助理(PDA)、便携式多媒体播放器(PMP)、MP3播放器、医疗装置、相机或可穿戴装置中的至少一个。可穿戴装置可包括配件型可穿戴装置(例如,手表、戒指、手镯、脚链、项链、眼镜、隐形眼镜或头戴式装置(HMD)、织物或衣物集成型可穿戴装置(例如,电子服装)、身体附着型可穿戴装置(例如,皮肤垫或纹身)和身体植入型可穿戴装置中的至少一种。在一些实施例中,电子装置可包括例如电视机(TV)、数字视频盘(DVD)播放器、音频播放器、冰箱、空调、吸尘器、烤箱、微波炉、洗衣机、空气净化器、机顶盒、家庭自动化控制面板、安全控制面板、媒体盒(例如,三星电子株式会社的HomeSync™、苹果公司的TV™或谷歌的TV™)、游戏控制台(例如,Xbox™、PlayStation™)、电子词典、电子钥匙、摄像机或数码相框中的至少一个。

[0043] 在其他实施例中,电子装置可包括各种医疗装置(例如,各种便携式医疗测量装置(诸如血糖仪、心率计、血压计、体温计等)、磁共振血管造影术(MRA)、磁共振成像(MRI)、计算机断层扫描(CT)、拍摄装置、超声装置等)、导航装置、全球导航卫星系统(GNSS)、事件数据记录器(EDR)、飞行数据记录器(FDR)、汽车信息娱乐装置等、船舶电子装置(例如,船舶导航装置、陀螺罗盘等)、航空电子装置、安全装置、汽车头单元、工业或家用机器人、无人机、金融机构的自动柜员机(ATM)、商店的销售点(POS)或物联网(IoT)装置(例如,灯泡、各种传感器、喷洒系统、火灾报警器、恒温器、路灯、烤面包机、健身器材、热水箱、加热器、锅炉等)中的至少一种。

[0044] 在本公开中,术语“用户”可指使用电子装置的人或使用电子装置的装置(例如,人工智能电子装置)。

[0045] 在下文中,将参照附图详细描述本公开。

[0046] 图1是根据本公开的实施例的根据用户语音执行任务的电子装置的使用示图。

[0047] 参照图1,电子装置100可接收用于激活人工智能代理程序的触发语音。作为示例,触发语音可包括诸如“Bixby”、“Siri”等的触发词语。在这种情况下,人工智能代理程序可包括对话系统,其中,所述对话系统可将用户对用户语音的响应处理或提供为自然语言,并且执行针对用户语音的任务。在这种情况下,除了用于激活人工智能代理程序的触发词语之外,电子装置100还可在设置在电子装置100中的特定按钮被选择之后接收用户语音。电子装置100可响应于触发语音激活人工智能代理程序。

[0048] 电子装置100可在人工智能代理程序被激活之后接收用户语音。在这种情况下,用户语音可以是用于执行特定任务的语音。

[0049] 作为示例,如在图1中所示出的,电子装置100可获得用户语音10“向XX发送我今天将迟到的短信”。

[0050] 电子装置100可对用户语音进行分析以确定(识别)用户语音是否是是需要用户确认的任务相关的用户语音。具体地,电子装置100可通过自动语音识别模块获得针对用户语音10的文本,并且基于获得的文本来确定(识别)用户语音是否是是需要用户确认的任务相关的用户语音。

[0051] 作为实施例,电子装置100可确定(识别)获得的用户语音是否是用于执行与用户隐私相关的预定任务(例如,汇款任务、产品购买任务、电子邮件发送任务、消息发送任务、电话呼叫任务等)或需要认证过程的用户语音。也就是说,电子装置100可确定(识别)用户语音是否相应于与用户隐私相关或需要认证过程的任务,以确定(识别)用户语音是否是是需要用户确认的任务相关的用户语音。

[0052] 作为另一示例,电子装置100可基于在用户语音的基础上获得的文本来识别与用户语音相关的任务和用于执行任务的实体。另外,电子装置100可基于识别出的任务和实体获得用户语音的安全分数,并且基于用户语音的安全分数确定用户语音是否是是需要用户确认的任务相关的用户语音。这里,在安全分数是预定值或更大的情况下,电子装置100可确定用户语音是是需要用户确认的任务相关的用户语音。

[0053] 当确定用户语音是是需要用户确认的任务相关的用户语音时,电子装置100可生成(获得)用于用户确认的问询。在这种情况下,电子装置100可生成(获得)与用户语音不相关的问询。作为示例,如在图1中所示出的,电子装置100可生成(获得)完全与消息发送任务不相关的问询20“如果您想要发送消息,告诉我今天的天气”。

[0054] 作为实施例,电子装置100可基于安全分数生成(获得)用于用户确认的问询。具体地,随着安全分数变高,电子装置100可生成(获得)与相应于用户语音的任务具有低相关性的问询。也就是说,电子装置100可生成(获得)与当前问询不相关的问询。另外,随着安全分数变低,电子装置100可生成(获得)与相应于用户语音的任务具有高相关性的问询。作为示例,电子装置100可在包括在用户语音中的文本中提取与任务相关的至少一个文本,并且生成(获得)用于诱导针对提取出的至少一个文本的用户话语的问询。

[0055] 另外,当安全分数是阈值或更大时,电子装置100可提供用于用户认证的认证消息以及用户确认。作为示例,电子装置100可基于预先注册的用户信息来提供用于用户认证的认证消息。在这种情况下,在认证消息中可包括要求至少一个信息的问询。例如,在认证消息中可包括可要求用于请求认证的用户信息的问询,并且在认证消息中可包括可另外要求除了用于请求认证的用户信息之外的至少一个信息的问询。作为示例,电子装置100可提供包括出于用户认证的目的而要求包括用户信息的特定词语的话语的问询的认证消息,并且除了要求包括用户信息的特定词语的话语的问询之外,还可提供还包括要求在显示器上拖动特定形状(例如,直线、星形等)或要求生物信息认证(例如,指纹、虹膜等)的消息的认证消息。

[0056] 作为另一示例,电子装置100可对平时使用电子装置100的用户的语音进行分析,以获得与另一用户的语音特征信息区分的关于文本的语音特征信息,并且将所述关于所

述文本的语音特征信息存储在存储器中。另外,当确定用户语音是与需要用户确认的任务相关的语音时,电子装置100可出于用户认证的目的生成包括存储的所述文本的问询。

[0057] 根据另一实施例,电子装置100可随机选择并生成存储的多个问询中的一个问询。

[0058] 根据另一实施例,电子装置100可通过显示器显示多个对象,并且生成要求选择所述多个对象中的一个对象的问询。

[0059] 根据另一实施例,电子装置100可基于电子装置100周围的情况来生成问询。具体地,电子装置100可通过各种方法确定电子装置100所处的场所(例如,家、车辆、公共场所等)。作为示例,电子装置100可基于连接到电子装置100的网络信息、全球定位系统(GPS)信息、关于连接到电子装置100的外部装置的信息等来确定电子装置100所处的场所。可选地,电子装置100可基于通过麦克风获取的外部噪声、环境噪声矢量、是否存在与周围人的对话等来确定电子装置100所处的场所或者电子装置100周围是否存在人。当确定电子装置100位于家中或者电子装置100周围没有人时,电子装置100可生成包括个人信息的问询。然而,当确定电子装置100位于外面或者电子装置100周围存在人时,电子装置100可生成不包括个人信息的问询。

[0060] 根据另一实施例,在可能无法计算安全分数或者在电子装置100中计算的安全分数在阈值范围内的情况下,电子装置100可通过在外部服务器中提供的学习的人工智能模型来计算更准确的安全分数。在这种情况下,与包括在电子装置100中的计算安全分数的方法相比,学习的人工智能模型(即基于用户语音和输入到各种电子装置100的安全分数学习的人工智能模型)可计算更准确的安全分数。

[0061] 当问询被生成时,电子装置100可输出生成的问询20。在这种情况下,电子装置100可通过自然语言生成模块以自然语言形式输出所生成的问询。

[0062] 电子装置100可接收用户对问询的响应。作为示例,如在图1中所示出的,电子装置100可接收用户响应30“很好”。

[0063] 电子装置100可确定用户响应是否是问询相应的用户响应。作为示例,在问询是要求用户信息的问询的情况下,电子装置100可确定在用户响应中是否包括准确的用户信息,在问询是要求特定词语的话语的问询的情况下,电子装置100可确定在用户响应中是否包括特定词语,以及在问询是要求选择显示器上显示的多个对象中的一个对象的问询的情况下,电子装置100可确定是否选择了要求在所述多个对象中选择对象。

[0064] 当与问询相应的用户响应被输入时,电子装置100可执行与用户语音相应的任务,并且提供所执行的任务的结果。作为示例,电子装置100可执行文本发送任务作为与用户语音10相应的任务,并且输出指示执行文本发送任务的结果的引导消息40“向XX发送消息”。

[0065] 根据上述实施例,电子装置100可通过用户确认执行与用户语音相应的任务,以保护用户隐私并增强安全性。

[0066] 在上述实施例中已经描述了人工智能代理系统被存储在电子装置100中并且电子装置100直接确定用户语音是否是是需要用户确认的任务相关的用户语音并生成问询的情况,但这仅是示例。上述操作中的一些可由外部服务器或其他装置来实现。作为示例,外部服务器可获得针对用户语音的文本,确定用户语音是否是是需要用户确认的任务相关的用户语音,或者生成问询。

[0067] 如上所述,电子装置100可使用人工智能代理来提供对用户语音的响应。在这种情

况下,作为用于提供基于人工智能(AI)的服务(例如,语音识别服务、秘书服务、翻译服务、搜索服务等)的专用程序的人工智能代理可由现有的通用处理器(例如,中央处理器(CPU))或单独的AI专用处理器(例如,图形处理器(GPU)等)来运行。具体地,人工智能代理可控制将在以下描述的各种模块(例如,对话系统)。

[0068] 具体地,在预定用户语音(例如,“Bixby”等)被输入或设置在电子装置100中的按钮(例如,用于运行人工智能代理的按钮)被按压的情况下,可对人工智能代理进行操作。另外,人工智能代理可对用户语音进行分析以确定用户语音是否是是需要用户确认的任务相关的用户语音,并且基于确定结果生成和提供问询。

[0069] 在预定用户语音(例如,“Bixby”等)被输入或设置在电子装置100中的按钮(例如,用于运行人工智能代理的按钮)被按压的情况下,可对人工智能代理进行操作。另外,人工智能代理可在输入用户语音或者按压设置在电子装置100中的按钮(例如,用于运行人工智能代理的按钮)之前处于被运行的状态。在这种情况下,在输入预定的用户语音(例如,“Bixby”等)或按压设置在电子装置100中的按钮(例如,用于运行人工智能代理的按钮)之后,电子装置100的人工智能代理可执行针对用户语音的任务。例如,在人工智能由AI专用处理器来运行的情况下,可在输入预定用户语音(例如,“Bixby”等)或按压设置在电子装置100中的按钮(例如,用于运行人工智能代理的按钮)之前由通用处理器来运行电子装置100的功能,并且可在输入预定用户语音(例如,“Bixby”等)或按压设置在电子装置100中的按钮(例如,用于运行人工智能代理的按钮)之后由AI专用处理器运行电子装置100的功能。

[0070] 另外,人工智能代理可在输入用户语音或按压设置在电子装置100中的按钮(例如,用于运行人工智能代理的按钮)之前处于待机状态。这里,待机状态是感测预定义用户输入被接收以控制人工智能代理的操作开始的状态。当在人工智能代理处于待机状态时输入预定用户语音(例如,“Bixby”等)或者按压设置在电子装置100中的按钮(例如,用于运行人工智能代理的按钮)时,电子装置100可运行人工智能代理并使用运行的人工智能代理执行针对用户语音的任务。

[0071] 另外,人工智能代理可在输入用户语音(例如,“Bixby”等)或按压设置在电子装置100中的按钮(例如,用于运行人工智能代理的按钮)之前处于终止的状态。当在人工智能代理终止的状态下输入预定用户语音(例如,“Bixby”等)或者按压设置在电子装置100中的按钮(例如,用于运行人工智能代理的按钮)时,电子装置100可运行人工智能代理并使用运行的人工智能代理执行针对用户语音的任务。

[0072] 人工智能代理可控制以下将描述的各种装置或模块。将在以下对此进行详细描述。另外,在下文中将通过各种实施例描述在电子装置100和服务器之间使用各种学习模型对用户语音进行分析以确定用户语音是否是是需要用户确认的任务相关的用户语音并基于确定结果生成和提供问询的具体示例。

[0073] 图2是示出根据本公开的实施例的电子装置的组件的框图。

[0074] 参照图2,电子装置100可包括输入接口110、通信接口120、存储器130和至少一个处理器140。然而,电子装置不限于包括上述组件,并且可根据期望的电子装置的类型添加、省略或组合一些组件。

[0075] 输入接口110可接收用于控制电子装置100的用户输入。作为示例,输入接口110可接收用于控制电子装置100的用户的各种用户操纵,诸如用户触摸、用户语音等。具体地,输

入接口110可接收用于执行任务的用户语音。

[0076] 通信接口120是用于执行与任意数量的外部装置的通信的组件,并且可在需要时执行与外部电子装置的各种通信。通信接口120与外部装置的通信连接可包括通过第三装置(例如,中继器、集线器、接入点、服务器、网关等)的通信。无线通信可包括使用例如长期演进(LTE)、高级LTE(LTE-A)、码分多址(CDMA)、宽带CDMA(WCDMA)、通用移动通信系统(UMTS)、无线宽带(WiBro)或全球移动通信系统(GSM)中的至少一个的蜂窝通信。根据实施例,无线通信可包括例如无线保真(WiFi)、蓝牙、蓝牙低功耗(BLE)、Zigbee、近场通信(NFC)、磁安全传输、射频(RF)或体域网(BAN)中的至少一个。有线通信可包括例如通用串行总线(USB)、高清晰度多媒体接口(HDMI)、推荐标准232(RS-232)、电力线通信或普通老式电话服务(POTS)中的至少一个。执行无线通信或有线通信的网络可包括电信网络(例如,计算机网络(例如,局域网(LAN)或广域网(WAN))、因特网或电话网络中的至少一个。

[0077] 另外,通信接口120可执行与用于提供人工智能服务的外部服务器的通信。具体地,通信接口120可将用户语音发送到外部服务器,并且从外部服务器接收用于用户确认的询问。

[0078] 存储器130可存储与电子装置100的至少一个其他组件相关的命令或数据。具体地,存储器130可由非易失性存储器、易失性存储器、闪存、硬盘驱动器(HDD)、固态驱动器(SDD)等实现。存储器130可由处理器140访问,并且存储器130中的数据读出、记录、校正、删除、更新等可由处理器140来执行。在本公开中,术语“存储器”可包括存储器130、处理器140中的只读存储器(ROM)(未示出)、随机存取存储器(RAM)(未示出)或安装在电子装置100中的存储卡(未示出)(例如,微安全数字(SD)卡或记忆棒)。另外,用于配置将在显示器的显示区域上显示的各种屏幕的程序、数据等可被存储在存储器130中。

[0079] 此外,存储器130可存储用于操作对话系统的人工智能代理。具体地,电子装置100可使用人工智能代理来生成自然语言作为对用户话语的响应。在这种情况下,人工智能代理可以是用于提供基于AI的服务(例如,语音识别服务器、秘书服务、翻译服务、搜索服务等)的专用程序。具体地,人工智能代理可由现有的通用处理器(例如,CPU)或单独的AI专用处理器(例如,GPU等)来执行。

[0080] 另外,存储器130可包括配置如在图4所示出的对话系统的多个组件(或模块)。具体地,存储器130可包括命令分类模块430和询问生成模块440。将参照图4对此进行详细描述。

[0081] 处理器140可电连接到存储器130以控制电子装置100的一般操作和功能。具体地,处理器140可执行存储在存储器130中的至少一个命令,以通过输入接口110接收用户语音并对用户语音进行分析,从而确定用户语音是否是是需要用户确认的任务相关的语音。另外,当确定用户语音是是需要用户确认的任务相关的语音时,处理器140可生成并提供用于用户确认的询问。此外,当通过输入接口110输入与询问相应的用户响应时,处理器130可执行与用户语音相应的任务。

[0082] 具体地,处理器140可确定与用户语音相关的任务是否是是需要用户确认的任务,诸如与用户隐私相关或要求安全性的任务。作为示例,处理器140可确定与用户语音相关的任务是否是汇款任务、产品购买任务、电子邮件发送任务、消息发送任务或电话呼叫任务中的一个。

[0083] 另外,处理器140可对通过输入接口110输入的用户语音进行分析以识别与用户语音相关的任务和用于执行任务的实体,基于识别出的任务和实体获得用户语音的安全分数,并且基于用户语音的安全分数确定用户语音是否是与需要用户确认的任务相关的语音。

[0084] 此外,当基于用户语音的安全分数,用户语音是与需要用户确认的任务相关的语音时,处理器140可基于安全分数生成问询。作为示例,处理器140可随着安全分数变高生成与相应于用户语音的任务具有低相关性的问询,并且随着安全分数变低生成与相应于用户语音的任务具有高相关性的问询。

[0085] 另外,当安全分数是阈值或更大时,处理器140可提供用于用户认证的认证消息。

[0086] 此外,处理器140可提取包括在用户语音中的文本中的至少一个文本,并且生成用于诱导针对提取出的所述至少一个文本的用户话语的问询。

[0087] 此外,处理器140可对使用电子装置100的用户的语音进行分析,以获得与另一用户的语音特征信息区分开的关于文本的语音特征信息,并且将所述关于所述文本的语音特征信息存储在存储器130中。此外,当确定用户语音是与需要用户确认的任务相关的语音时,处理器140可生成包括存储的所述文本的问询。此外,当在用户响应中获得了所述关于所述文本的语音特征信息时,电子装置100可执行与用户语音相应的任务。

[0088] 图3是详细示出根据本公开的实施例的电子装置的详细组件的框图。

[0089] 参照图3,电子装置100可包括输入接口110、通信接口120、存储器130、显示器150、扬声器160、传感器170和处理器140。已经参照图2描述了在图3中所示出的输入接口110、通信接口120、存储器130和处理器140,因此省略了对其的重合描述。

[0090] 输入接口110可接收用于控制电子装置100的用户输入。具体地,输入接口110可接收用于执行特定任务的用户语音。如在图3中所示出的,输入接口110可包括用于接收用户语音的麦克风111、用于接收使用用户的手、触控笔等的用户触摸的触摸面板113、用于接收用户操纵的按钮115等。然而,在图3中所示出的输入接口110仅是示例,并且可由其他输入装置(例如,键盘、鼠标、运动输入等)来实现。

[0091] 显示器150可根据处理器140的控制来显示各种类型的信息。具体地,显示器150可显示包括用于问询的多个对象的用户界面(UI)。另外,显示器150可显示包括用户与人工智能代理之间的对话的消息窗口。显示器150可由触摸屏与触摸面板113一起来实现。

[0092] 扬声器160是输出各种警报或语音消息以及由音频处理器执行各种处理(诸如解码、放大、噪声过滤等)的各种音频数据的组件。具体地,扬声器160可将针对用户语音的问询或指导消息输出为具有自然语言形式的语音消息。用于输出音频的组件可以由扬声器实现,但这仅是示例,并且用于输出音频的组件可由可输出音频数据的输出端来实现。

[0093] 传感器170可感测电子装置100的各种类型的状态信息。作为示例,传感器170可包括可感测电子装置100的运动信息的运动传感器(例如,陀螺仪传感器、加速度传感器等),并且可包括可感测位置信息的传感器(例如,GPS传感器)、可感测电子装置100周围的环境信息的传感器(例如,温度传感器、湿度传感器、大气压力传感器等)、可检测电子装置100的用户信息的传感器(例如,血压传感器、血糖传感器、脉搏率传感器等)。传感器170还可包括用于捕捉电子装置100的外部的图像的图像传感器等。

[0094] 图4是示出根据本公开的实施例的人工智能代理系统的对话系统的框图。

[0095] 参照图4,对话系统400是用于通过自然语言与虚拟人工智能代理执行对话的组件。根据本公开的实施例,对话系统400可被存储在电子装置100的存储器130中。然而,这仅是示例,并且包括在对话系统400中的至少一个组件可被包括在至少一个外部服务器中。

[0096] 对话系统400可包括自动语音识别(ASR)模块410、自然语言理解(NLU)模块420、命令分类模块430、问询生成模块440、对话管理器(DM)模块450、自然语言生成器(NLG)模块460和从文本到语音(TTS)模块470。对话系统400还可包括路径规划器模块、动作规划器模块等。

[0097] 自动语音识别(ASR)模块410可将从电子装置100接收的用户语音转换为文本数据。例如,自动语音识别模块410可包括话语识别模块。话语识别模块可包括声学模型和语言模型。例如,声学模型可包括与发声有关的信息,并且语言模型可包括关于单位音素信息的组合和单位音素信息的信息。话语识别模块可使用与发声有关的信息和关于单位音素信息的信息将用户话语转换为文本数据。关于声学模型和语言模型的信息可被存储在例如自动语音识别数据库(ASR DB) 415中。

[0098] 自然语言理解模块420可执行语法分析或语义分析以抓取用户的意图。句法分析可将用户输入划分为语法单元(例如,词语、短语、词素等)并且抓取划分出的单元具有哪些语法元素。可使用语义匹配、规则匹配、公式匹配等来执行语义分析。因此,自然语言理解模块420可从用户输入获得表达意图所需的域、意图或实体(或参数、槽等)。

[0099] 自然语言理解模块420可使用被划分为抓取意图所需的域、意图和实体的匹配规则来确定用户意图执行的任务和实体。例如,一个域(例如,消息)可包括用于确定任务的多个意图(例如,消息发送、消息删除等),并且一个意图可包括多个实体(例如,发送目标、发送时间、发送内容等)。多个规则可包括例如一个或更多个基本元素参数。匹配规则可被存储在自然语言理解数据库(NLU DB) 425中。

[0100] 自然语言理解模块420可使用诸如词素、短语等的语言特征(例如,语法元素)来抓取从用户输入中提取的词语的含义,并且将所抓取的词语的含义与领域和意图进行匹配以确定用户意图执行的任务。例如,自然语言理解模块420可计算在每一个域和意图中包括多少个从用户输入提取出的词语,并且确定用户意图执行的任务。根据实施例,自然语言理解模块420可使用作为用于抓取用户意图执行的任务的基础的词语来确定用户语音的实体。根据实施例,自然语言理解模块420可使用存储用于抓取与用户语音相关的任务的语言特征的自然语言识别数据库425来确定与用户语音相关的任务。在这种情况下,个人信息等可被包括在自然语言识别数据库425中。

[0101] 自然语言理解模块420可基于与用户语音相关的任务和实体来生成路径规则。例如,自然语言理解模块420可基于用户输入的意图来选择将运行的应用,并且确定将在选择的应用中执行的任务。自然语言理解模块420可通过确定与所确定的任务相应的实体来生成路径规则。根据实施例,由自然语言理解模块420生成的路径规则可包括关于将运行的应用、将在应用中运行的任务以及运行任务所需的实体的信息。

[0102] 自然语言理解模块420可基于与用户语音相关的任务和实体来生成一个路径规则或多个路径规则。例如,自然语言理解模块420可从路径规划器模块接收与电子装置100相应的路径规则集,并且将与用户语音相关的任务和实体映射到接收到的路径规则集以确定路径规则。在这种情况下,路径规则可包括关于用于执行应用的功能的任务(或操作)的信

息或关于运行任务所需的实体的信息。另外，路径规则可包括应用的操作顺序。电子装置100可接收路径规则，根据路径规则选择应用，并且在所选择的应用中运行包括在路径规则中的任务。

[0103] 自然语言理解模块420可通过基于与用户语音相关的任务和实体确定将运行的应用、将在应用中运行的任务以及运行任务所需的实体来生成一个路径规则或多个路径规则。例如，自然语言理解模块420可通过使用电子装置100的信息根据用户语音的意图以本体或图形模型形式排列将运行的应用和将在应用中运行的任务来生成路径规则。生成的路径规则可通过例如路径规划器模块被存储在路径规则数据库中。生成的路径规则可被添加到数据库425的路径规则集。

[0104] 自然语言理解模块420可选择生成的所述多个路径规则中的至少一个路径规则。作为示例，自然语言理解模块420可选择所述多个路径规则中的最佳路径规则。作为另一示例，自然语言理解模块420可在基于用户语音仅指定一些任务的情况下选择所述多个路径规则。自然语言理解模块420可通过用户的另外的输入来确定所述多个路径规则中的一个路径规则。

[0105] 命令分类模块430可基于与用户语音相关的任务和通过自然语言理解模块420获取的实体来确定用户语音是否是是需要用户确认的任务相关的语音。具体地，命令分类模块430可获得与用户语音相关的任务的安全分数和通过自然语言理解模块420获得的实体的安全分数。在这种情况下，任务的安全分数和实体的安全分数可被预先存储在安全分数DB 435中。另外，可基于任务是否需要安全分数、任务的使用频率等来预先确定任务的安全分数。作为示例，如在表1中所示出的，任务的安全分数可被存储在安全分数或隐私等级DB 435中。

[0106] [表1]

[0107]	任务的种类	安全分数
	汇款任务	5
	产品购买任务	
	消息发送任务和电子邮件发送任务	3
	电话呼叫任务	2
	内容再现任务、电子邮件阅读任务等	0

[0108] 另外，可基于针对每一个任务的实体的重要性、量和使用频率来确定实体的安全分数。作为示例，在汇款任务的情况下，电子装置100可使用针对汇款时使用的金额的日志来确定安全分数。作为示例，包括关于由用户主要使用的金额的信息的实体可具有低安全分数，并且包括关于由用户不主要使用的金额的信息的实体可具有高安全分数。也就是说，因为用户的日志信息的使用，所以可根据使用电子装置100的用户的用户信息针对相同金额生成不同的安全分数。作为另一示例，在消息发送任务、电子邮件发送任务和电话任务的情况下，可根据从用户的联系人信息应用获得的针对其他用户的联系频率和所述其他用户的信息(例如，VIP和朋友)来生成不同的安全分数。另外，即使当任务由相同用户执行时，也可根据附加项目(照片文件、视频文件、音频文件或文档)的类型或大小来生成和存储不同

的安全分数。作为示例,如在表2中所示出的,实体的安全分数可被存储在DB 435中。

[0109] [表2]

[0110]	任务的种类	实体	安全分数
[0111]	汇款任务	0美元~100美元	1
		100美元~1000美元	2
		1000美元	3
	消息发送任务、电子邮件发送任务和电话呼叫任务	发送目标是家庭组	1
		发送目标是公司组	3
		发送目标是朋友组	2
		附件是风景照片	2
		附件是人像照片	3

[0112] 任务的安全分数和实体的安全分数可被存储在DB 425中,但这仅是示例,并且任务的安全分数和实体的安全分数可由命令分类模块430来计算。例如,可将关于任务和实体的信息输入到预先学习的人工智能模型,使得任务的安全分数和实体的安全分数可被获得。命令分类模块430可基于获得的任务的安全分数和获得的实体的安全分数来获得用户语音的安全分数。在这种情况下,命令分类模块430可通过等式1来获得用户语音的安全分数。

[0113] 用户语音的安全分数=

[0114] (任务的安全分数)* w_1 +(实体的安全分数)* w_2 ...等式1

[0115] 作为示例, w_1 和 w_2 可分别由诸如0.8和0.2的系数来实现,但这仅是示例,并且可根据用户信息等使用不同的系数来计算用户语音的安全分数。另外,等式1仅是示例,并且可根据实体的种类、任务的类型、用户信息等通过不同的等式来计算用户语音的安全分数。

[0116] 另外,在无法确定与用户语音相应的任务或实体使得安全分数无法被计算或者当确定由诸如等式1的线性等式计算的安全分数在阈值范围内时,命令分类模块430可请求外部服务器计算安全分数以计算更准确的安全分数。在这种情况下,外部服务器可使用具有比线性等式的准确度更高准确度的学习的人工智能模型来计算安全分数。

[0117] 另外,命令分类模块430可基于获得的用户语音的安全分数来确定与用户语音相关的任务是否需要用户确认的任务。具体地,命令分类模块430可基于用户语音的安全分数是否是预定值(例如,2)或更大来确定与用户语音相关的任务是否需要用户确认的任务。作为示例,在与第一用户语音相关的任务是汇款任务并且汇款金额是\$200的情况下,命令分类模块430可确定第一用户语音的安全分数是 $5*0.8+2*0.2=4.4$,并且因为第一用户语音的安全分数是预定值或更大,因此可确定与第一用户语音相关的任务是用户确认的任务。作为示例,在与第二用户语音相关的任务是电话呼叫任务并且电话呼叫目标是家庭的情况下,命令分类模块430可确定第二用户语音的安全分数是 $2*0.8+1*0.2=1.8$,并且因为第二用户语音的安全分数小于预定值,因此可确定与第二用户语音相关的任务不是需要用户确认的任务。

[0118] 在上述实施例中,命令分类模块430基于安全分数来确定与用户语音相关的任务是否是需要用户确认的任务,但这仅是示例,并且命令分类模块430可确定与用户语音相关的任务是否是预定义任务,以确定与用户语音相关的任务是否是需要用户确认的任务。作为示例,在与用户语音相关的任务是预先存储的汇款任务、产品购买任务、电子邮件发送任务、消息发送任务或电话呼叫任务中的一个的情况下,命令分类模块430可确定与用户语音相关的任务是需要用户确认的任务。

[0119] 当确定与用户语音相关的任务是需要用户确认的任务时,命令分类模块430可将与用户语音相关的信息(例如,文本信息、安全分数信息等)输出到问询生成模块440,并且当确定与用户语音相关的任务不是需要用户确认的任务时,命令分类模块430可将与用户语音相关的信息(例如,文本信息等)输出到对话管理器模块450。

[0120] 问询生成模块440可生成用于用户确认以执行与用户语音相关的任务的问询。在这种情况下,问询生成模块440可生成与和用户语音相关的任务不相关的问询。例如,在与用户语音相关的任务是汇款任务的情况下,问询生成模块440可询问与汇款任务不相关的问询,诸如天气问询、文本话语诱导问询等。

[0121] 另外,问询生成模块440可生成用于向用户诱导具有低话语频率的答案的问询。作为示例,问询生成模块440生成用于诱导包括用户当前不频繁使用的文本的响应的问询(例如,“如果您现在想要发送电子邮件,请告诉我济州岛最高山的名称”)。作为另一示例,问询生成模块440可生成要求任意数字或文本话语的问询(例如,“如果您现在想要发送电子邮件,请读出xxxx(随机数)”)。作为另一示例,问询生成模块440可生成要求用户的个人信息的问询(例如,“如果您现在想要发送电子邮件,请告诉我您的出生日期”)。在这种情况下,可通过用户日志信息或文本语料库来计算话语频率。

[0122] 另外,问询生成模块440可基于用户语音的安全分数来生成问询。具体地,问询生成模块440可基于用户语音的安全分数来确定问询的复杂度。作为示例,问询生成模块440可随着安全分数变高生成要求复杂响应的问询,并且可随着安全分数变低生成要求简单响应的问询(例如,是/否)。

[0123] 具体地,问询生成模块440可基于用户语音的安全分数来确定用户语音的任务与问询之间的相关性。作为示例,随着安全分数变高,问询生成模块440可生成与用户语音的任务不相关的问询。也就是说,问询生成模块440可生成诱导在语义上远离用户语音的任务(远离连续空间中的命令域树或命令矢量)的话语的问询。作为示例,问询生成模块440可使用人工智能模型(诸如,可在矢量空间中映射文本的skip-gram模型)来表达矢量空间中的文本。另外,问询生成模块440可使用与在矢量空间上表达的文本相应的矢量角度之间的相似性(即,余弦相似性)或两个矢量之间的距离来生成诱导在语义上远离用户语音的任务的话语的问询。作为示例,当两个矢量之间的距离是90度或者两个矢量之间的距离大时,电子装置100可确定与所述两个矢量相应的文本彼此不相关。

[0124] 随着安全分数变低,问询生成模块440可生成与用户语音的任务相关的问询。另外,问询生成模块440可基于用户语音的安全分数生成需要用户认证的问询。作为示例,在用户语音的安全分数是预定值或更大的情况下,问询生成模块440可生成用于用户认证的要求用户的个人信息的问询。

[0125] 另外,因为问询生成模块440需要将对问询的用户响应与问询的正确答案进行比

较,所以问询生成模块440可使用诸如现有的用户的响应等的日志信息来生成诱导预期具有高准确度的话语的问询。

[0126] 另外,问询生成模块440可基于电子装置100周围的情况来生成问询。具体地,问询生成模块440可通过各种方法确定电子装置100所处的场所(例如,家、车辆、公共场所等)。作为示例,问询生成模块440可基于连接到电子装置100的网络信息、GPS信息、关于连接到电子装置100的外部装置的信息等来确定电子装置100所处的场所。具体地,问询生成模块440可基于连接到电子装置100的网络信息(例如,互联网协议(IP)地址)、通过GPS检测到的位置信息等来确定电子装置100的位置是家庭还是公共场所。另外,在基于通过通信模块(诸如蓝牙)搜索的关于外部装置的信息(例如,装置的类型、装置的名称等)搜索的外部装置的数量是一个的情况下,问询生成模块440可确定电子装置100的位置是家,并且在基于关于外部装置的信息搜索的匿名外部装置的数量是多个的情况下,问询生成模块440可确定电子装置100的位置是公共场所。

[0127] 可选地,问询生成模块440可基于通过麦克风获取的外部噪声、环境噪声矢量、是否存在与周围人的对话等来确定电子装置100所处的场所或者电子装置100周围是否存在人。具体地,问询生成模块440可通过将通过麦克风获得的音频输入到学习的人工智能模型或基于隐马尔可夫模型(HMM)的声学模型来确定当前噪声的场所是车辆、图书馆、卫生间还是地铁。可选地,问询生成模块440可通过测量通过麦克风获得的音频的大小Db来确定电子装置100所处的场所。另外,电子装置100可使用环境噪声矢量来确定电子装置100所处的场所。具体地,电子装置100可根据噪声的种类来生成并存储矢量变换模型(诸如,说话者识别矢量变换)。问询生成模块440可使用语音识别声学模型感测除话语部分之外的噪声部分,使用预先存储的矢量转换模型将噪声部分的声源段转换为环境矢量,将先前生成的针对每一个噪声的矢量与环境矢量进行比较以获得分数(例如,余弦距离),然后基于获得的分数确定电子装置100当前所处的场所或环境。另外,问询生成模块440可基于预先注册的说话者信息来确定是否存在与周围人的对话或关于进行对话的人的信息,以确定电子装置100当前所处的场所或环境或者电子装置100周围是否存在人。

[0128] 问询生成模块440可基于确定的电子装置100的位置或者是否存在另一人来生成问询。具体地,当确定电子装置100位于家中或者电子装置100周围不存在人时,问询生成模块440可生成包括个人信息的问询。然而,当确定电子装置100位于外部或者电子装置100周围存在人时,问询生成模块440可生成不包括个人信息的问询。

[0129] 另外,问询生成模块440可通过各种方法生成问询。将参照图6A至图9B对此进行详细描述。

[0130] 对话管理器模块450可执行由自然语言理解模块420确定的任务。也就是说,对话管理器模块450可基于根据基于自然语言理解模块420获得的任务和实体来执行任务,并且生成对用户语音的响应。

[0131] 另外,对话管理器模块450可确定由自然语言理解模块420抓取的用户的意图是否清楚。例如,对话管理器模块450可基于实体的信息是否足够来确定用户的意图是否清楚。对话管理器模块450可确定在自然语言理解模块420中抓取的实体是否足以执行任务。根据实施例,在用户的意图不清楚的情况下,对话管理器模块450可执行用于请求用户输入所需信息的反馈。例如,对话管理器模块450可执行用于请求用户输入关于实体的信息以抓取用

户意图的反馈。另外,对话管理器模块450可生成并输出用于确认包括由自然语言理解模块420改变的文本的用户询问的消息。

[0132] 根据实施例,对话管理器模块450可包括内容提供器模块。在内容提供器模块可基于由自然语言理解模块420抓取的任务和实体来执行操作的情况下,内容提供器模块可生成执行与用户输入相应的任务的结果。

[0133] 根据另一实施例,对话管理器模块450可使用存储在知识DB 455中的知识库来提供对用户语音的响应。在这种情况下,知识库可被包括在电子装置100中,但这仅是示例,并且知识库可被包括在外部服务器中。

[0134] 自然语言生成器模块(NLG模块)460可将通过问询生成模块430或对话管理器模块450输出的信息改变为文本形式。改变为文本形式的信息可具有自然语言话语的形式。改变为文本形式的信息可以是例如关于另外的输入的信息、用于引导完成与用户输入相应的操作的信息、或者用于引导用户的另外的输入的信息(例如,用户输入的反馈信息)。改变为文本形式的信息可被显示在电子装置100的显示器150上,或者可由从文本到语音模块(TTS模块)470被改变为语音形式。

[0135] 从文本到语音模块(TTS模块)470可将具有文本形式的信息改变为具有语音形式的信息。从文本到语音模块470可从自然语言生成模块440接收具有文本形式的信息,将具有文本形式的信息改变为具有语音形式的信息,并且将具有语音形式的信息输出到扬声器。

[0136] 自然语言理解模块420、命令分类模块430、问询生成模块440和对话管理器模块450可被实现为至少一个模块。作为示例,自然语言理解模块420、命令分类模块430、问询生成模块440和对话管理器模块450可被实现为一个模块,以确定用户的任务和实体,基于确定的用户的任务和实体来确定与用户语音相关的任务是否需要用户确认的任务,并且基于确定结果来生成问询或获得与用户语音相应的响应(例如,路径规则)。作为另一示例,命令分类模块430和问询生成模块440可被包括在对话管理器模块450中。

[0137] 图5是用于描述根据本公开的实施例的在用户语音是与需要用户确认的任务相关的语音的情况下提供用于用户确认的问询的电子装置的控制方法的流程图。

[0138] 参照图5,在操作S510,电子装置100可接收用户语音。在这种情况下,电子装置100可通过麦克风111接收用户语音,但这仅是示例,并且电子装置100可以从外部装置接收用户语音。

[0139] 在操作S520,电子装置100可确定用户语音是否是是与需要用户确认的任务相关的语音。具体地,电子装置100可基于用户语音是否是涉及用户隐私或要求安全性的任务相关的语音来确定用户语音是否是是与需要用户确认的任务相关的语音。

[0140] 当在操作S520确定用户语音是与需要用户确认的任务相关的语音时,在操作S530,电子装置100可获得(生成)用于用户确认的问询。在这种情况下,用于用户确认的问询可以是与用户语音相关的任务不相关的问询。电子装置100可通过各种方法生成用于用户确认的问询,将在以下参照图6A至图9B对此进行描述。电子装置100可输出用于用户确认的问询。在这种情况下,电子装置100可通过扬声器160输出问询,但这仅是示例,并且电子装置100可通过显示器150输出问询。

[0141] 在操作S540,电子装置100可确定是否输入了与问询相应的用户响应。

[0142] 当在操作S540-Y输入了与问询相应的用户响应时,在操作S550,电子装置100可执行与用户语音相应的任务。然而,当在操作S540未输入与问询相应的用户响应时,电子装置100可以不执行与用户语音相应的任务或者输出用于附加确认或附加反馈的问询。

[0143] 当在操作S520确定用户语音不是与需要用户确认的任务相关的语音时,在操作S550,电子装置100可立即执行与用户语音相应的任务。

[0144] 在下文中,将参照图6A至图9B描述生成用于用户确认的问询的各种示例。

[0145] 图6A和图6B是用于描述根据本公开的各种实施例的基于安全分数生成问询的示例的示意图。

[0146] 参照图6A,电子装置100可接收用户语音610“向我的妻子发送我今天将一起吃饭的消息”。

[0147] 电子装置100可计算接收到的用户语音610的安全分数。参照表1、表2和等式1,电子装置100可将用户语音610的安全分数确定2.8。因此,电子装置100可确定用户语音的安全分数是第一阈值(例如,3)或更小,以生成具有低复杂度的问询。

[0148] 作为示例,电子装置100可生成问询620“如果您想要发送消息,请告诉我“蓝色大海””。

[0149] 另外,当接收到用户响应630“蓝色大海”时,电子装置100可执行作为与用户语音相关的任务的消息发送任务,并且输出作为执行任务的结果的引导消息640“已经发送消息”。

[0150] 参照图6B,电子装置100可接收用户语音650“向我的妻子汇款一百万韩元”。

[0151] 电子装置100可计算接收到的用户语音650的安全分数。参照表1、表2和等式1,电子装置100可确定用户语音650的安全分数是4.4。因此,电子装置100可确定用户语音的安全分数是第二阈值(例如,4)或更大,以生成用于用户认证的问询。作为示例,电子装置100可生成问询660“如果您想要汇款,请告诉我您的妻子的生日和账户密码的前两位数”以要求用户的个人信息。在这种情况下,为了防止个人信息的泄漏,电子装置100可生成仅要求一些个人信息而不是全部个人信息的问询。

[0152] 另外,当接收到用户响应670“五月15日、46”时,电子装置100可确定用户响应670是否是问询660的正确答案,当用户响应670是问询660的正确答案时,执行作为与用户语音相关的任务的汇款任务,并且输出作为执行任务的结果的引导消息680“已经向您的妻子汇款一百万元”。

[0153] 图7A和图7B是用于描述根据本公开的各种实施例的使用包括在用户语音中的一些文本生成问询的示例的示意图。

[0154] 参照图7A,电子装置100可接收用户语音710“向经理发送包括我的当前位置信息的“我将迟到”的消息”。

[0155] 电子装置100可计算接收到的用户语音710的安全分数。参照表1、表2和等式1,电子装置100可确定用户语音710的安全分数是3。因此,电子装置100可确定用户语音的安全分数是第三阈值(例如,2)或更大,以询问用于用户确认的问询。

[0156] 在这种情况下,电子装置100可生成要求通过自动语音识别模块410获得的多个文本中的至少一个文本的话语的问询。在这种情况下,电子装置100可生成要求所述多个文本中的被确定为关键词的至少一个文本的话语的问询。作为示例,电子装置100可生成问询

720“如果您想要向经理发送消息,请告诉我“包括位置的发送””。在这种情况下,电子装置100可使用通过自然语言理解(NLU)模块获得的结果值来生成要求所述多个文本中的被确定为关键词的至少一个文本的话语的问询。具体地,在自然理解处理中执行命名实体识别(NER)的情况下,电子装置100可通过将执行NER的结果彼此组合来生成要求组合文本的话语的问询。作为示例,在获得执行NER的结果{接收者:“经理”,命令:“发送文本”,消息:“迟到”,附加内容:“位置信息”}的情况下,电子装置100可生成要求文本的组合(诸如,除了消息之外的接收者和附加内容)的话语的问询。作为另一示例,在不使用NLU结果的情况下,电子装置100可通过使用预定规则组合通过自动语音识别(ASR)的结果获得的至少一个文本来生成问询。作为示例,电子装置100可通过对文本的词素或文本的语音的标记部分进行分析,来生成要求话语“经理、当前位置信息、消息”的问询,其中,所述话语“经理、当前位置信息、消息”是基于ASR结果获得的文本中的名词。另外,电子装置100可生成要求话语“位置信息”的问询,其中,所述话语“位置信息”是基于ASR结果获得的文本中用户先前未发音的文本。

[0157] 另外,当接收到用户响应730“包括位置的发送”时,电子装置100可响应于用户响应730通过传感器170获得当前位置信息,执行用于将消息与获得的位置信息一起发送给经理的消息发送任务,并且输出作为执行任务的结果的引导消息740“已经完成消息发送”。

[0158] 电子装置100还可基于用户语音的安全分数来确定问询的复杂度。例如,随着安全分数变高,电子装置可生成要求更多个词语(例如,“经理、迟到、位置、包括、发送”)的话语的问询,并且随着安全分数变低,电子装置可生成要求更少个词语(例如,“经理、发送”)的话语的问询。

[0159] 参照图7B,电子装置100可接收用户语音750“向经理发送包括我的当前位置信息的“我将迟到”的消息”。

[0160] 电子装置100可计算接收到的用户语音750的安全分数。参照表1、表2和等式1,电子装置100可确定用户语音750的安全分数是3。因此,电子装置100可确定用户语音的安全分数是第三阈值(例如,2)或更大,以询问用于用户确认的问询。

[0161] 在这种情况下,电子装置100可生成要求关于通过自然语言理解模块420获得的实体的信息的问询。在这种情况下,电子装置100可生成要求关于“发送目标”的信息的问询,其中,所述“发送目标”是通过自然语言理解模块420获得的实体。作为示例,电子装置100可生成问询760“如果您想要向经理发送消息,请告诉我经理的姓名”。

[0162] 另外,当接收到用户响应770“Hong Kil-dong”时,电子装置100可响应于用户响应770通过传感器170获得当前位置信息,执行用于将消息与获得的位置信息一起发送给经理的消息发送任务,并且输出作为执行任务的结果的引导消息780“已经完成消息发送”。

[0163] 图8是用于描述根据本公开的实施例的根据电子装置的性能生成问询的方法的示图。

[0164] 如上所述,在电子装置100不包括显示器150的情况下,电子装置100可生成具有语音形式的要求用户响应的问询。然而,在电子装置100包括显示器150的情况下,电子装置100可在显示器150上显示多个对象,并且生成诱导选择所述多个对象中的一个对象的问询。

[0165] 参照图8,电子装置100可获得用户语音810“在xx商业中心订购三包尿布”。

[0166] 电子装置100可计算接收到的用户语音810的安全分数。电子装置100可确定用户语音810的安全分数是第三阈值(例如,2)或更大,以询问用于用户确认的询问。

[0167] 在这种情况下,电子装置100可显示包括多个对象的UI 830。在这种情况下,包括在UI 830中的对象可包括与用户语音不相关的文本或图像。作为示例,如在图8中所示出的,电子装置100可显示包括诸如“梨”、“苹果”和“菠萝”的水果对象的UI 830。

[0168] 另外,电子装置100可生成要求选择所述多个对象中的至少一个对象的询问。作为示例,电子装置100可生成询问820“请在显示的UI中选择苹果”。

[0169] 另外,当选择苹果的用户触摸被输入时,电子装置100可响应于用户触摸执行产品购买任务,并且输出作为执行任务的结果的引导消息840“已经订购三包尿布”。

[0170] 即使电子装置100包括显示器150,电子装置100也可仅在特定情况下(例如,用户当前靠近电子装置100的情况)在显示器上显示UI。作为示例,当通过分析通过相机捕捉的图像或通过麦克风接收的语音确定用户位于电子装置附近,或者用户处于他/她佩戴电子装置100的状态或处于他/她使用电子装置100的状态时,电子装置100可在显示器上显示UI。

[0171] 在上述实施例中,该询问是要求选择显示器上显示的多个对象中的至少一个对象的询问,但这仅是示例,并且电子装置100可生成要求与显示器150上显示的所述至少一个对象相应的文本的话语的询问。作为示例,电子装置100可以在显示器150上显示对象“苹果”并生成询问“请告诉我屏幕上显示的水果的名称”。在上述实施例中,以音频形式输出询问,但这仅是示例,并且可在显示器150上输出询问。

[0172] 当确定存在用户正在使用的另一电子装置时,电子装置100可使用另一电子装置生成用于用户确认的询问。电子装置100可基于可由另一电子装置获得的传感器信息来生成用于用户确认的询问,并且接收由另一电子装置获得的传感器信息以确定用户是否对询问做出响应。

[0173] 作为确定是否存在用户正在使用的另一电子装置的方法,电子装置100可感测被确定为由用户拥有的另一电子装置连接到电子装置100,并且当电子装置100从另一电子装置接收用户的装置使用信息时,确定用户正在使用另一电子装置。

[0174] 电子装置100可使用另一电子装置的用户简档信息、账户信息、认证信息和从生物传感器获得的生物信息来确定用户是否拥有另一电子装置。装置使用信息可包括指示用户是否穿戴了另一电子装置的信息和指示用户是否正在使用另一电子装置的信息。

[0175] 作为示例,指示用户是否穿戴了另一电子装置的信息可以是在检测到可穿戴装置和/或电子装置100的运动之后可穿戴装置的可附接/可拆卸结构处于紧固状态并且时间在参考时间内的情况、在可穿戴装置和/或电子装置100中检测到生物信号之后时间在参考时间内的情况、或者可穿戴装置的可附接/可拆卸结构处于紧固状态并且在可穿戴装置中检测到生物信号的情况。可使用设置在可穿戴装置的可附接/可拆卸结构中的传感器的感测值来识别可穿戴装置的可附接/可拆卸结构是否处于紧固状态。设置在可附接/可拆卸结构中的传感器可包括例如导电传感器、霍尔传感器、磁传感器等。可使用设置在可穿戴装置和/或电子装置100中的运动传感器或加速度传感器来识别可穿戴装置和/或电子装置100的运动。可穿戴装置和/或电子装置100可包括检测生物信号的生物传感器以检测生物信号。生物传感器可包括例如心率传感器、脉搏传感器、血压传感器、汗液传感器、体温传感

器、虹膜传感器、指纹传感器等。

[0176] 作为示例,在检测到用户输入之后时间在参考时间内的情况下,在接收到更新屏幕的用户命令的情况下,或者在屏幕处于打开状态的情况下,电子装置100可确定相应装置正在被使用。作为示例,电子装置100可生成问询(或请求)“请摇动你佩戴的智能手表”。另外,电子装置100可使用智能手表中的运动传感器等接收关于智能手表的运动的信息,以确定用户是否响应。作为示例,电子装置100可生成问询(或请求)“请您在您当前正在使用的装置上执行指纹认证”。然后,在从用户正在使用的另一电子装置的指纹传感器识别并认证指纹的情况下,电子装置100可接收相应的认证信息以确定用户是否做出响应。

[0177] 图9A和图9B是用于描述根据本公开的各种实施例的学习包括在用户语音中的一些文本的语音特征信息并使用具有学习的语音特征信息的文本生成问询的示例的示图。

[0178] 当从用户输入用户语音时,电子装置100可使用用户语音来执行说话者识别。当通过说话者识别确认用户是使用电子装置100的用户时,电子装置100可收集由使用电子装置100的用户发出的用户语音。

[0179] 另外,当使用电子装置100的用户使用说话者识别服务时,电子装置100可通过将使用电子装置100的用户的语音的语音特征与针对其他用户的语音的注册语音特征进行比较来获得各个用户的语音之间的相似性。具体地,电子装置100可获得关于用户语音的语音特征信息。在这种情况下,电子装置100可通过对针对用户语音中包括的每一个音素的频率特性进行分析来获得语音特征信息。具体地,电子装置100可使用诸如倒频谱、线性预测系数(LPC)、mel频率倒谱系数(MFCC)、滤波器组能量等的特征提取技术来获得用户语音中包括的每一个音素的频率特性。在这种情况下,针对每一个音素的频率特性可由语音特征矢量来表示,并且可由二维矢量或更多维矢量的多维矢量表示。

[0180] 另外,电子装置100可通过将针对多个用户中的每一个用户获得的每一个音素的语音特征信息进行比较来获得相似性。另外,电子装置100可获得在针对每一个音素的语音特征信息的相似性中具有高相似性差异的音素,以提取说话者特征信息。也就是说,电子装置100可在用户之间对针对每一个音素的语音特征信息进行比较,以排除不反映说话者的特征的具有高相似性的语音,并且将反映说话者的特征的具有低相似性的语音存储为说话者特征信息。另外,电子装置100可将针对多个用户中的每一个用户获得的每一个音素的语音特征信息(即,语音特征矢量)输入到学习的人工智能模型(例如,深度神经网络(DNN)模型),以提取指示说话者的特征的说话者特征矢量。在这种情况下,说话者语音特征矢量可被表示为二维矢量或更多维矢量的多维矢量。

[0181] 另外,电子装置100可将高权重分配给具有最大相似度差异的部分的音素。另外,当在同一音素中连续存在大差异的相似性分数时,电子装置100可将高权重分配给具有大差异的相似性分数的音素。

[0182] 参照图9A,在用户语音910“播放我昨天观看的戏剧”中包括的文本中的文本“昨天”与在用户说出用户语音910的同时由另一用户说出的“昨天”被区分开的情况下(也就是说,在相似度分数差为阈值或更大的情况下),电子装置100可将关于文本“昨天”的信息915和关于文本的语音特征的信息存储在存储器130中。

[0183] 在新用户被注册的情况下,电子装置100可诱导现有用户的语音特征的文本的话语。也就是说,电子装置100可将现有用户的语音特征与新用户的语音特征彼此进行比较和

存储,以进一步增强现有用户的语音特征的识别能力。

[0184] 电子装置100可使用由使用电子装置100的用户发音的文本之中的具有存储在存储器130中并且与另一用户的话语区分开的语音特征的文本来生成问询。

[0185] 参照图9B,电子装置100可接收用户语音920“购买xx产品”。

[0186] 电子装置100可计算接收到的用户语音920的安全分数。在用户语音920的安全分数是第三阈值(例如,2)或更大的情况下,电子装置100可询问用于用户确认的问询。具体地,在安全分数是需要用户认证的第四阈值(例如,3.5)或更高的情况下,电子装置100可生成用于用户认证的问询。

[0187] 具体地,电子装置100可使用具有在使用电子装置100的用户与另一用户之间区分开的语音特征的文本来生成问询。例如,电子装置100可使用文本“昨天”来生成要求“您昨天遇到的朋友”的话语的问询930“执行认证。语音密码是“您昨天遇到的朋友””,其中,所述文本“昨天”是具有存储在存储器130中并且与另一用户区分开的语音特征的文本。在这种情况下,在语音密码中可包括至少一个其他文本以及具有与另一用户区分开的语音特征的文本。另外,每当生成问询时,可改变包括语音密码的问询930。作为示例,可生成包括文本“昨天”的各种语音密码(诸如“昨天您遇到的朋友”、“昨天天气”、“昨天您吃的食物”等)。也就是说,每当生成问询时,可改变语音密码以防止通过记录的语音执行用户认证。

[0188] 另外,当接收到用户响应940“我昨天遇到的朋友”时,电子装置100可将针对包括在用户响应940中的“昨天”的语音特征与预先存储的针对“昨天”的语音特征进行比较,以执行用户认证。在针对包括在用户响应940中的“昨天”的语音特征与预先存储的针对“昨天”的语音特征之间的相似度是阈值或更大的情况下,电子装置100可执行用户认证并执行与用户语音920相关的任务。另外,电子装置100可输出作为认证结果的引导消息950“已认证”。

[0189] 根据如上所述的本公开的各种实施例,电子装置100可在执行与用户语音相应的任务之前执行用户确认或用户认证,以防止执行用户无意的任务。

[0190] 电子装置100可使用存储在电子装置100中的人工智能代理程序来执行如上所述的操作,但这仅是示例,并且电子装置100可结合外部服务器执行如上所述的操作。

[0191] 在下文中,将参照图10至图12描述结合外部服务器生成用于用户确认的问询的示例。

[0192] 图10是示出根据本公开的实施例的由人工智能系统生成用于用户确认的问询的示例的序列图。在图10中,示出了用于描述外部服务器确定与用户语音相应的任务是否需要用户确认的任务的示例的示意图。

[0193] 参照图10,电子装置100可接收触发语音(S1005)。在这种情况下,触发语音可包括用于运行存储在电子装置100中的人工智能代理程序的触发词语。根据另一实施例,为了执行用于语音识别的人工智能代理程序,电子装置100可接收用于选择设置在电子装置100中的特定按钮的用户命令。

[0194] 电子装置100可响应于触发语音来激活人工智能代理(S1010)。在这种情况下,人工智能代理可以是用于向用户提供交互服务的虚拟秘书。

[0195] 电子装置100可接收用户语音(S1015)。在这种情况下,用户语音可包括用于执行特定任务的文本。

[0196] 电子装置100可将用户语音发送到服务器1000 (S1020)。

[0197] 服务器1000可将用户语音转换为文本 (S1025)。在这种情况下,服务器1000可使用自动语音识别 (ASR) 模块将用户语音转换为文本。

[0198] 服务器1000可识别(确定)与用户语音相应的任务 (S1030)。具体地,服务器1000可通过自然语言理解 (NLU) 模块对用户语音进行分析,以确定与用户语音相应的任务和实体。

[0199] 服务器1000可识别(确定)与用户语音相应的任务是否需要用户确认的任务 (S1035)。也就是说,服务器1000可确定与用户语音相应的任务是否是用户隐私相关或需要安全性的任务。

[0200] 当确定与用户语音相应的任务是需要用户确认的任务时,服务器1000可将问询生成命令发送到电子装置100 (S1040)。在这种情况下,服务器1000可将关于用户语音的信息与问询生成命令一起发送。

[0201] 电子装置100可响应于问询生成命令获得(生成)用于用户确认的问询 (S1045)。在这种情况下,问询可以是与相应于用户语音的任务不相关的任务。

[0202] 电子装置100可接收对用户问询的用户响应 (S1050),并且基于用户响应来执行任务 (S1055)。

[0203] 图11是示出根据本公开的实施例的由人工智能系统生成用于用户确认的问询的另一示例的序列图。在图11中,示出了用于描述外部服务器确定与用户语音相应的任务是否需要用户确认的任务并生成问询的示例的示图。

[0204] 参照图11,图11的S1105、S1110、S1115、S1120、S1125、S1130和S1135相应于在图10中所描述的S1005、S1010、S1015、S1020、S1025、S1030和S1035,因此省略其详细描述。

[0205] 当识别(确定)与用户语音相应的任务是需要用户确认的任务时,服务器1000可获得(生成)问询 (S1140)。在这种情况下,服务器1000可生成与相应于用户语音的任务不相关的问询。

[0206] 服务器1000可将问询发送到电子装置100 (S1145)。

[0207] 电子装置100可提供问询 (S1150)。在这种情况下,电子装置100可通过扬声器160以音频形式输出问询,但这仅是示例,并且电子装置100可通过显示器150输出问询。

[0208] 电子装置100可接收用户响应 (S1155),并且将接收到的用户响应发送到服务器1000 (S1160)。

[0209] 服务器1000可识别(确定)从电子装置100接收的用户响应是否是用户响应 (S1165)。当从电子装置100接收的用户响应是与问询相应的用户响应时,服务器1000可生成任务执行命令(指示与任务执行命令一起执行任务的结果的引导消息)。当从电子装置100接收的用户响应不是与问询相应的用户响应时,服务器1000可生成不可能消息,其中,所述不可能消息包括指示无法执行任务的消息。

[0210] 服务器1000可将任务执行命令或不可能消息发送到电子装置100 (S1170)。

[0211] 电子装置100可基于任务执行命令来执行任务或者提供不可能消息 (S1175)。

[0212] 图12是示出根据本公开的实施例的由人工智能系统生成用于用户确认的问询的另一示例的序列图。在图12中,示出了用于描述外部服务器将用户语音转换为文本的示例的示图。

[0213] 参照图12,图12的S1205、S1210、S1215、S1220和S1225相应于在图10中所描述的

S1005、S1010、S1015、S1020和S1025,因此省略其详细描述。

[0214] 服务器1000可将获得的文本发送到电子装置100 (S1230)。

[0215] 电子装置100可识别(确定)与用户语音相应的任务(S1235)。具体地,电子装置100可通过自然语言理解(NLU)模块分析用户语音,以确定与用户语音相应的任务和实体。

[0216] 电子装置100可识别(确定)与用户语音相应的任务是否需要用户确认的任务(S1240)。也就是说,电子装置100可确定与用户语音相应的任务是否是用户隐私相关或需要安全性的任务。

[0217] 当确定与用户语音相应的任务是否需要用户确认的任务时,电子装置100可响应于问询生成命令获得(生成)用于用户确认的问询(S1245)。在这种情况下,问询可以是与相应于用户语音的任务不相关的任务。

[0218] 电子装置100可接收针对用户问询的用户响应(S1250),并且基于用户响应来执行任务(S1255)。

[0219] 图13是用于描述根据本公开的实施例的电子装置或外部服务器根据用户语音的安全分数将用户语音转换为文本的示例的序列图。

[0220] 参照图13,图13的S1305、S1310、S1315、S1320、S1325和S1330相应于在图12中所示出的S1205、S1210、S1215、S1220、S1225和S1230,因此省略其详细描述。

[0221] 电子装置100可基于获得的文本使用第一方式来计算用户语音的安全分数(S1335)。在这种情况下,第一种方式是线性回归方式或基于一般规则的轻量级方式。作为示例,可按照如在等式1中的方式计算用户语音的安全分数。

[0222] 电子装置100可确定安全分数是否未被计算或在阈值范围内(S1340)。也就是说,在可能无法通过在电子装置100中存在的自然语言理解模块确定与用户语音相应的任务或实体的情况下,或者在安全分数基于阈值范围(例如,0.5)在阈值范围(例如,0.45至0.55)内的情况下,电子装置100可确定安全分数是否未被计算或在阈值范围内,以通过外部服务器1000获得准确的安全分数。

[0223] 在计算出的安全分数在阈值范围之外的情况下,电子装置100可基于计算出的安全分数获得(生成)问询(S1360)。以上已经描述了使用安全分数生成问询的方法,因此省略其详细描述。

[0224] 在未计算安全分数或安全分数在阈值范围内的情况下,电子装置100可请求服务器1000重新计算安全分数(S1345)。在这种情况下,电子装置100可将与用户语音相应的文本信息与请求一起发送。

[0225] 服务器1000可使用第二方式重新计算用户语音的安全分数(S1350)。在这种情况下,第二方式可以是使用人工智能模型(例如,DNN模型)的方式,其中,所述人工智能模型被学习以通过输入与用户语音相应的文本来计算用户语音的安全分数。服务器1000可将重新计算的安全分数发送到电子装置100 (S1355),并且电子装置100可基于重新计算的安全分数获得(生成)问询(S1360)。

[0226] 电子装置100可接收对用户问询的用户响应(S1365),并且基于用户响应来执行任务(S1370)。

[0227] 根据如上所述的各种实施例的电子装置100可结合外部服务器确定与用户语音相应的任务是否需要用户确认的任务,并且基于确定结果生成用于用户确认的问询。

[0228] 在图13中已经描述了服务器1000将用户语音转换为文本的情况,但这仅是示例,并且可通过设置在电子装置100中的ASR模块410将用户语音转换为文本。

[0229] 在上述实施例中已经描述了一次输入与需要用户确认的任务相关的用户语音的情况,但这仅是示例,并且可连续地或周期性地输入与需要用户确认的任务相关的用户语音。

[0230] 在这种情况下,电子装置100可确定用户语音是否由同一用户输入。具体地,在人工智能代理通过触发语音被激活的状态下,在通过输入第一用户语音开始用于用户语音识别的会话之后的预定时间内再次输入第二用户语音的情况下,电子装置100可确定第一用户语音和第二用户语音是否是同一用户发出的语音。当第一用户语音和第二用户语音是由同一用户输入的时,电子装置100可针对第二用户语音执行处理,并且将用于用户语音识别的会话维持预定时间。然而,在第一用户语音和第二用户语音不是由同一说话者发出的语音的情况下,电子装置100可以不处理第二用户语音或者可提供拒绝响应。当同一用户在预定时间内未输入用户语音时,电子装置100可结束用于用户语音识别的会话。

[0231] 另外,电子装置100可将针对第一用户语音的安全分数和针对第二用户语音的安全分数彼此进行比较。在针对第二用户语音的安全分数高于针对第一用户语音的安全分数的情况下,电子装置100可生成用于针对第二用户语音的附加确认(或附加认证)的问询。作为示例,电子装置100可生成诱导句子或词语的话语,其中,所述句子或词语包括具有与另一用户的语音特征区分开的语音特征的文本。另外,电子装置100可使用在维持用于用户语音识别的会话的同时实时地临时累积的数据来执行用户认证。

[0232] 在本公开中所使用的术语“~器/者”或“模块”可包括由硬件、软件或固件配置的单元,并且可与诸如例如逻辑、逻辑块、组件、电路等的术语兼容地使用。术语“~器/者”或“模块”可以是执行一个或更多个功能的整体配置的组件或最小单元或者其一部分。例如,模块可由专用集成电路(ASIC)来配置。

[0233] 本公开的各种实施例可通过包括存储在机器可读存储介质(例如,计算机可读存储介质)中的指令的软件来实现。机器可以是存储介质调用存储的指令的装置,并且可根据调用的指令进行操作,并且可包括根据所公开的实施例的电子装置(例如,电子装置100)。在由处理器执行命令的情况下,处理器可直接执行与命令相应的功能,或者其他组件可在处理器的控制下执行与命令相应的功能。命令可包括由编译器或解释器创建或执行的代码。机器可读存储介质可以以非暂时性存储介质的形式来提供。这里,术语“非暂时性”意味着存储介质是有形的而不包括信号,并且不对数据是半永久还是临时存储在存储介质中进行区分。

[0234] 根据实施例,可在计算机程序产品中包括和提供根据在文档中公开的各种实施例的方法。计算机程序产品可作为产品在卖方和买方之间被交易。计算机程序产品可以以可由机器读取的存储介质(例如,光盘只读存储器(CD-ROM))的形式被发布或通过应用商店(例如,Play Store™)在线发布。在在线发布的情况下,计算机程序产品中的至少部分可被至少临时存储在存储介质中,诸如制造商的服务器、应用商店的服务器或中继服务器的存储器中,或者被临时创建。

[0235] 根据不同实施例的组件(例如,模块或程序)中的每一个可包括单个实体或多个实体,并且可省略上述相应子组件中的一些,或者可在不同实施例中进一步包括其他子组件。

可选地或另外地,一些组件(例如,模块或程序)可被集成到一个实体中,并且可以相同或相似的方式执行在被集成之前由各个相应组件执行的功能。由根据不同实施例的模块、程序或其他组件执行的操作可以以顺序方式、并行方式、迭代方式或启发式方式来执行,至少一些操作可以以不同的顺序来执行或被省略,可对操作进行组合,或者可添加其他操作。

[0236] 虽然已经参考本公开的各种实施例示出和描述了本公开,但本领域技术人员将理解的是,在不脱离由所附权利要求及其等同物限定的本公开的精神和范围的情况下,可在其中进行形式和细节上的各种改变。

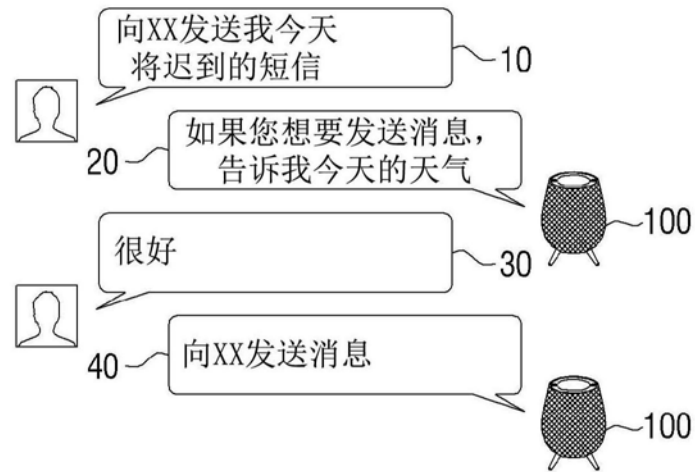


图1

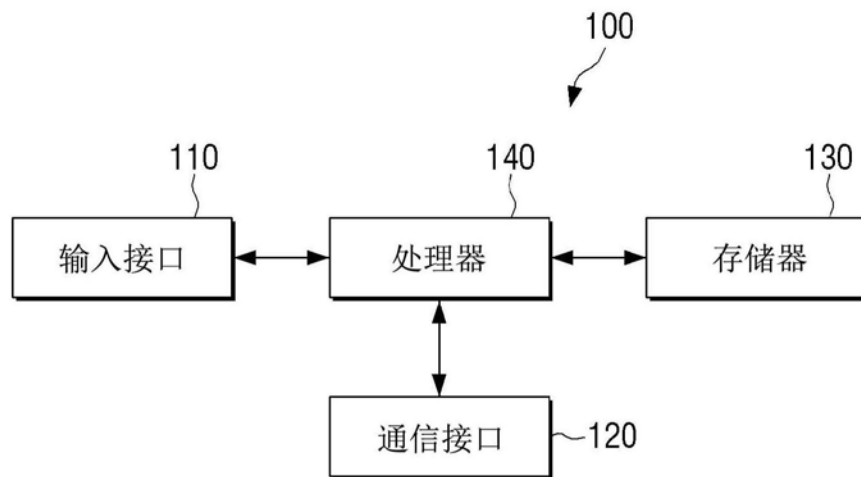


图2

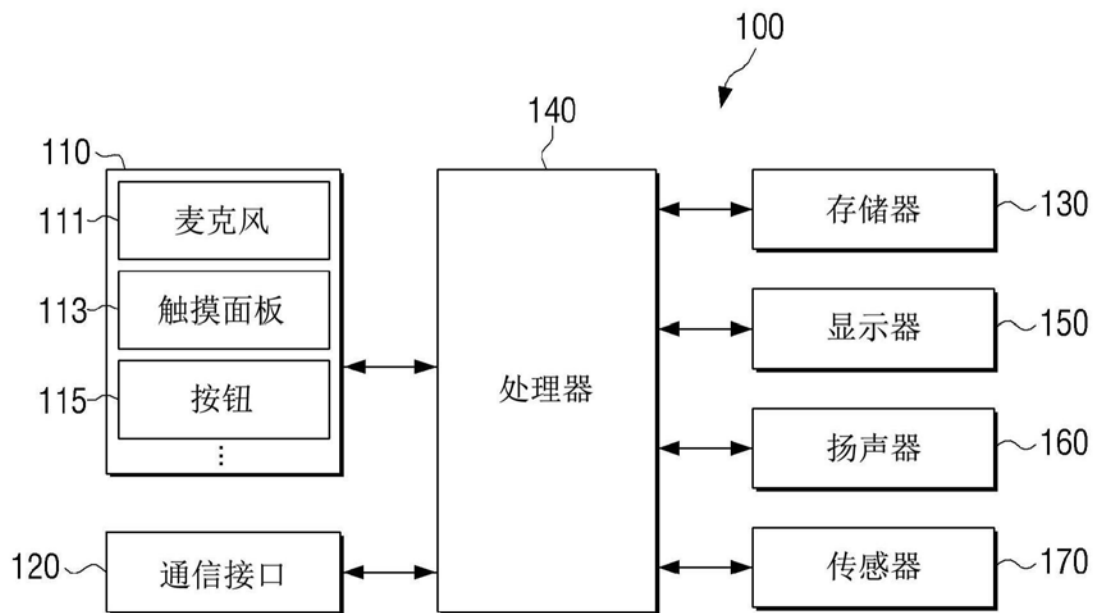


图3

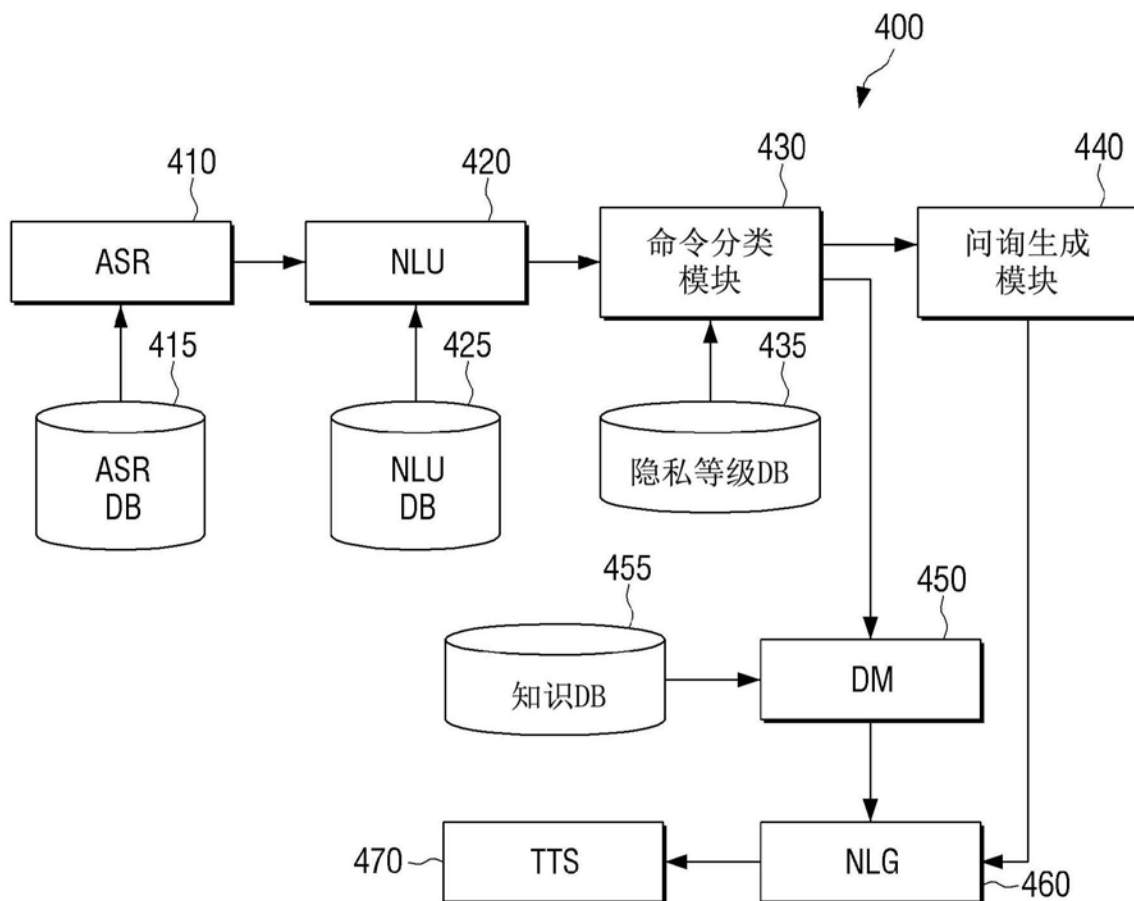


图4

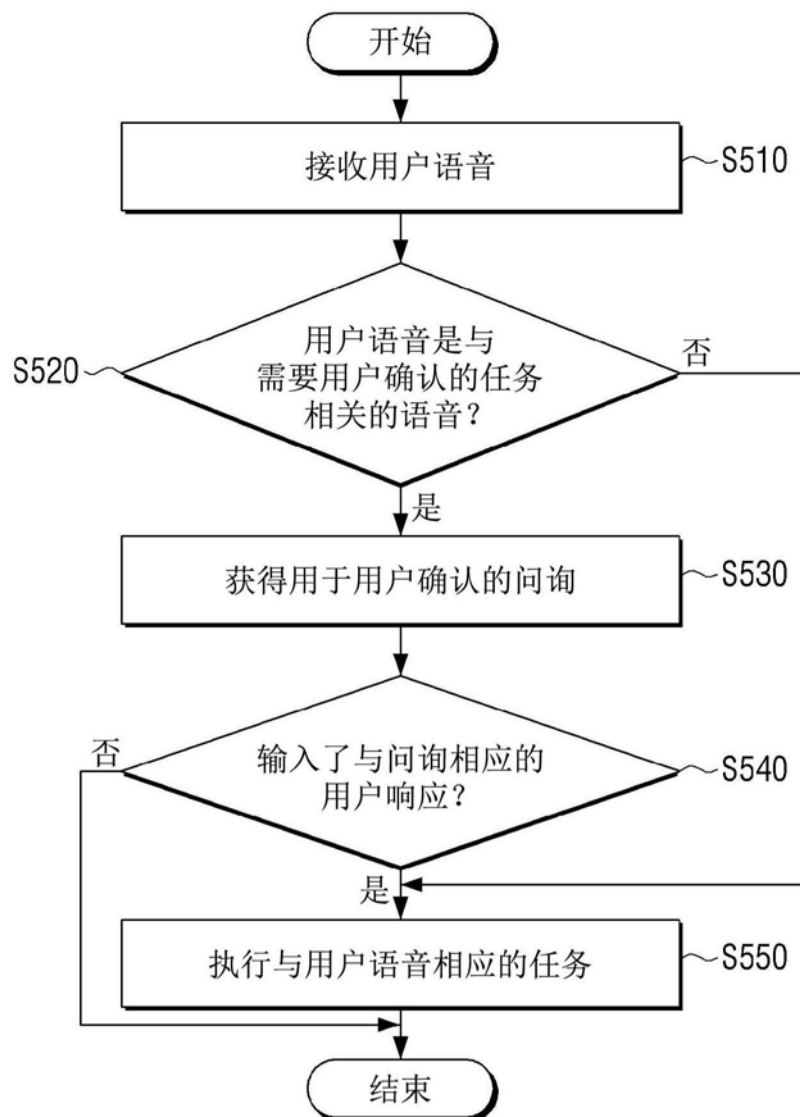


图5

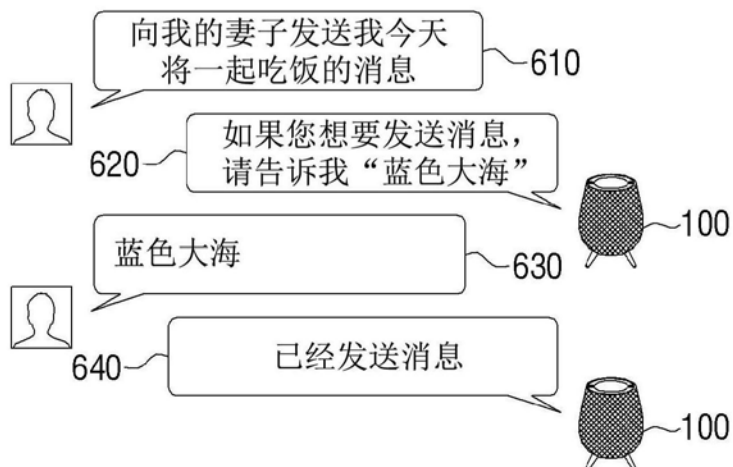


图6a

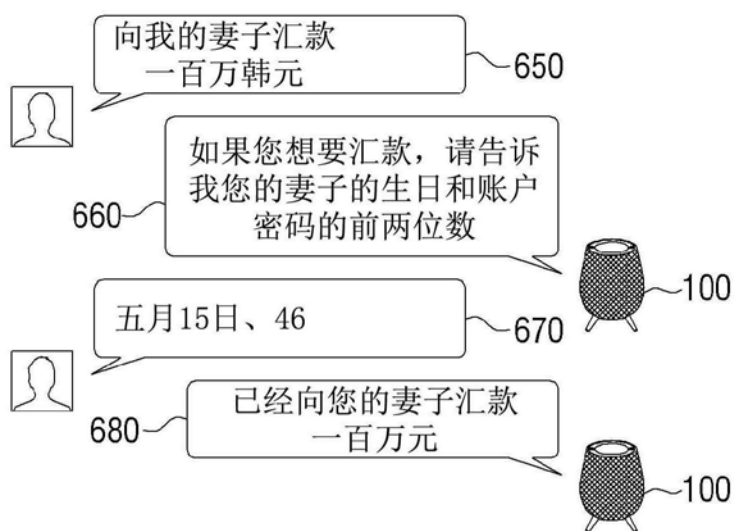


图6b

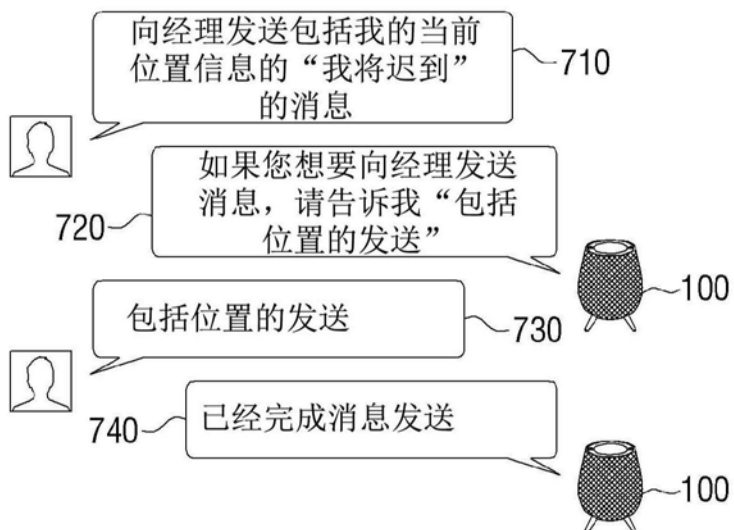


图7a

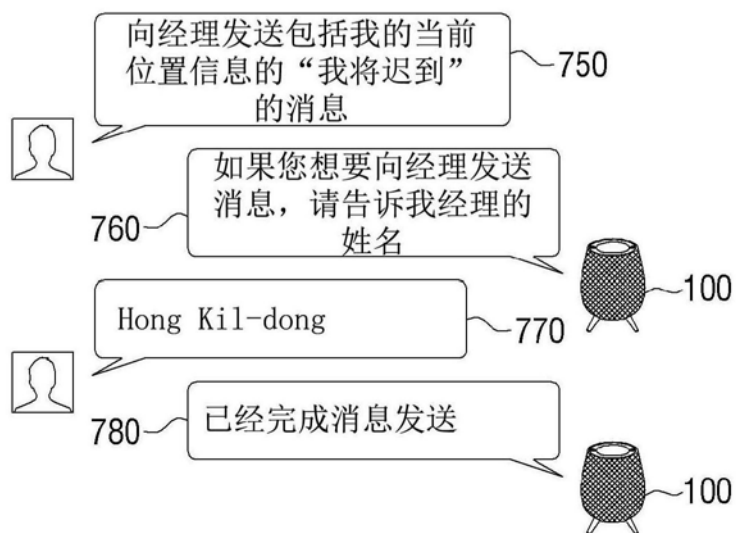


图7b

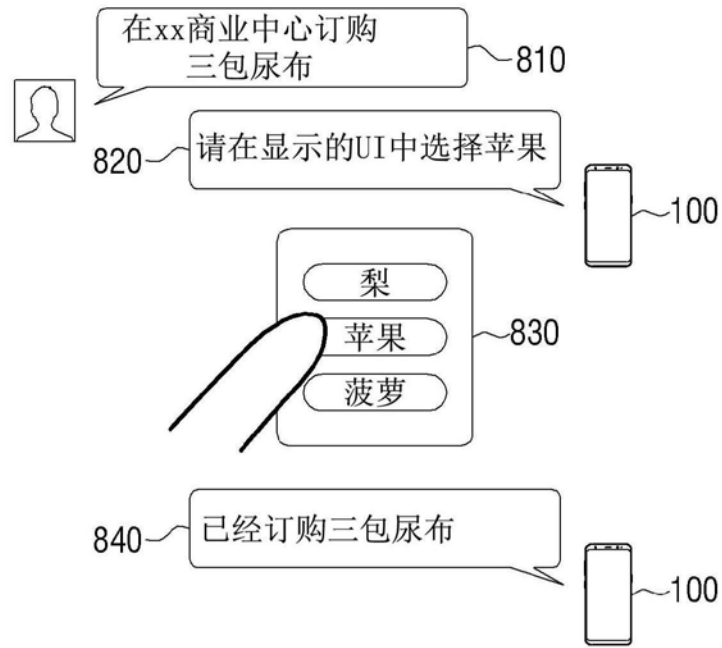


图8

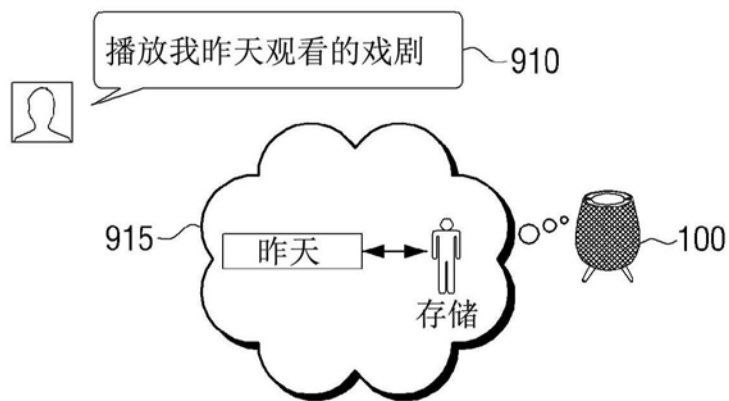


图9a

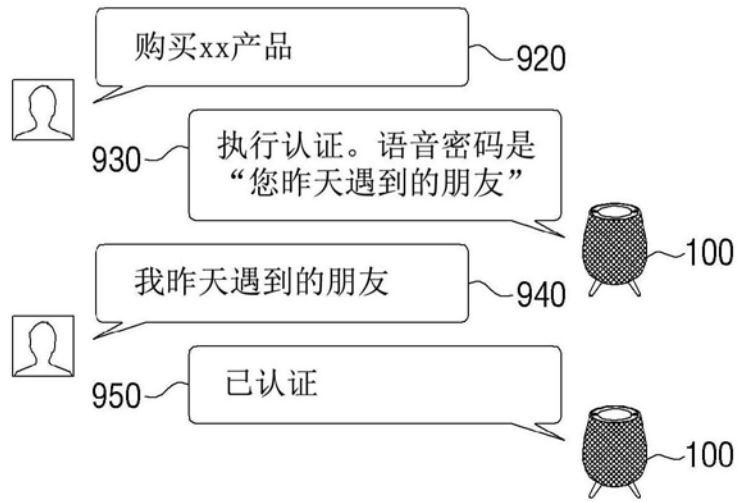


图9b

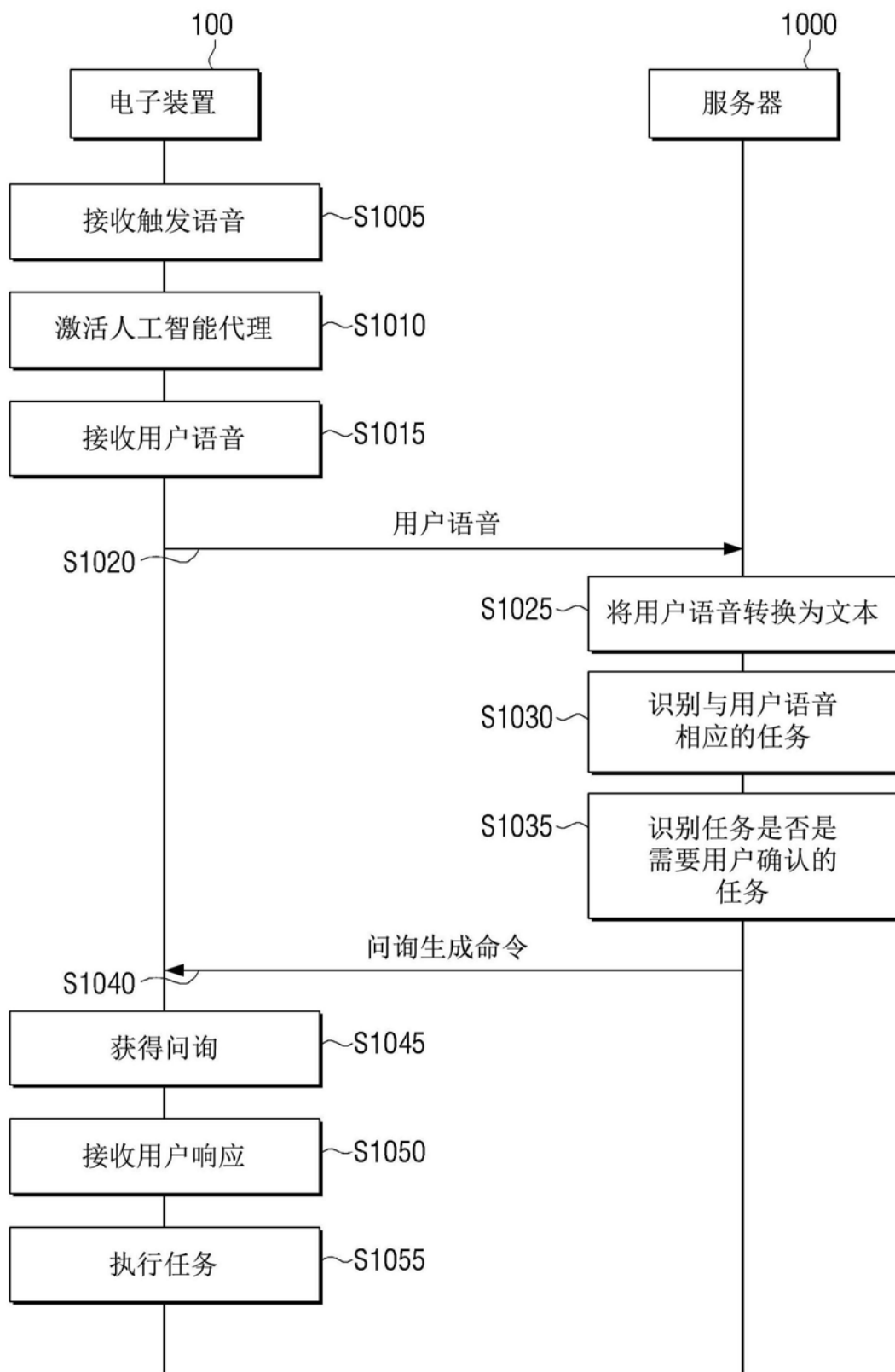


图10

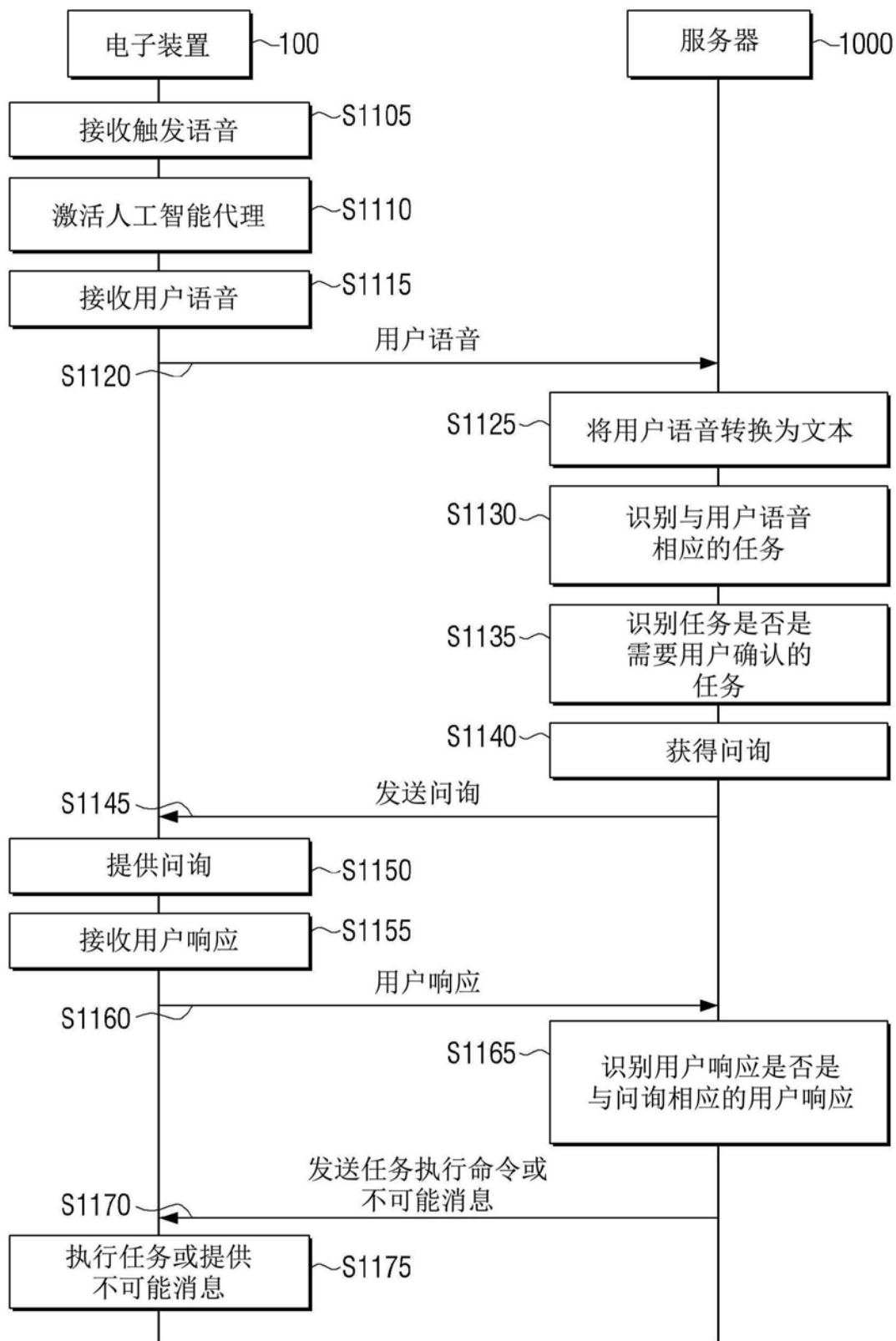


图11

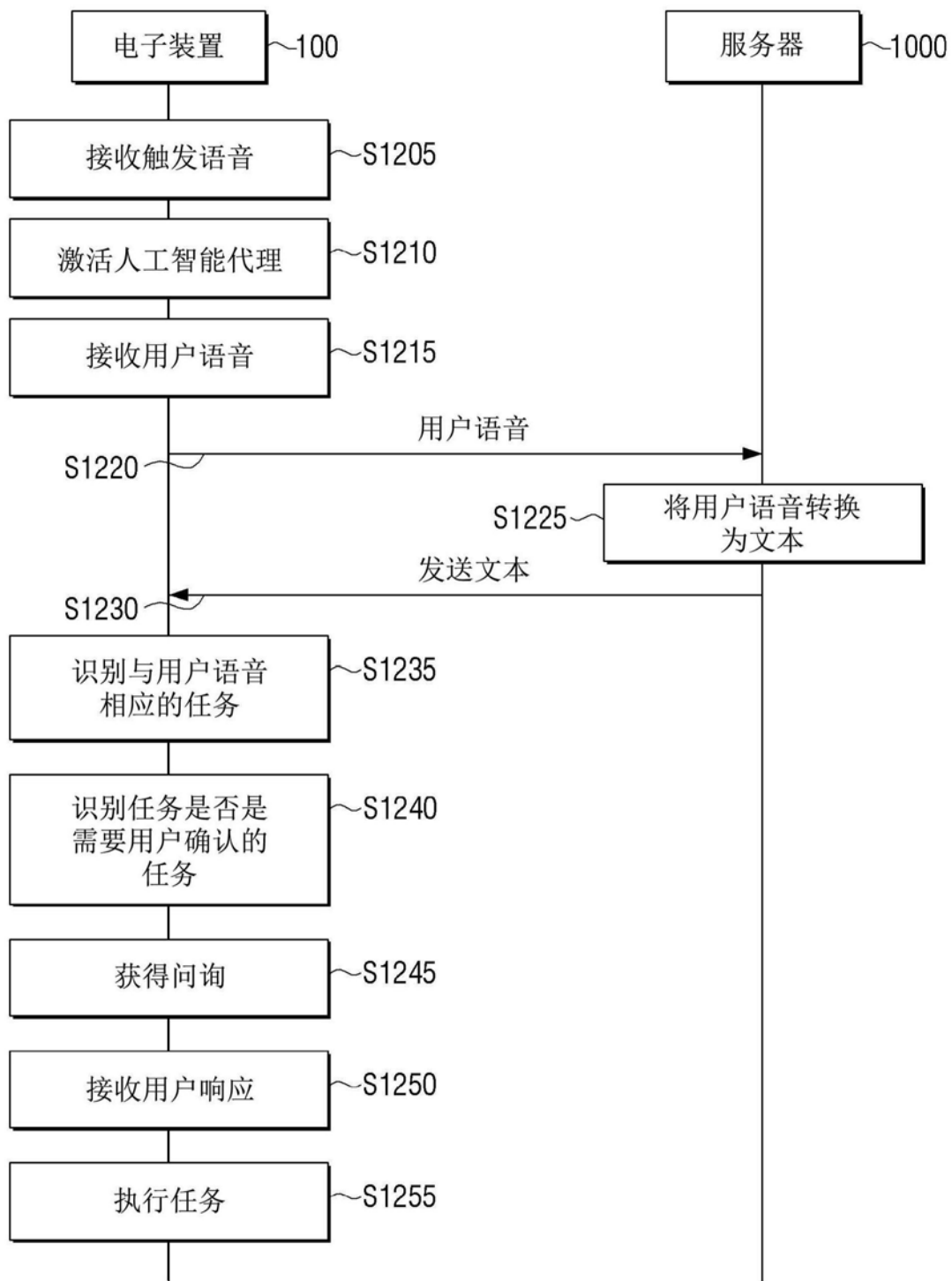


图12

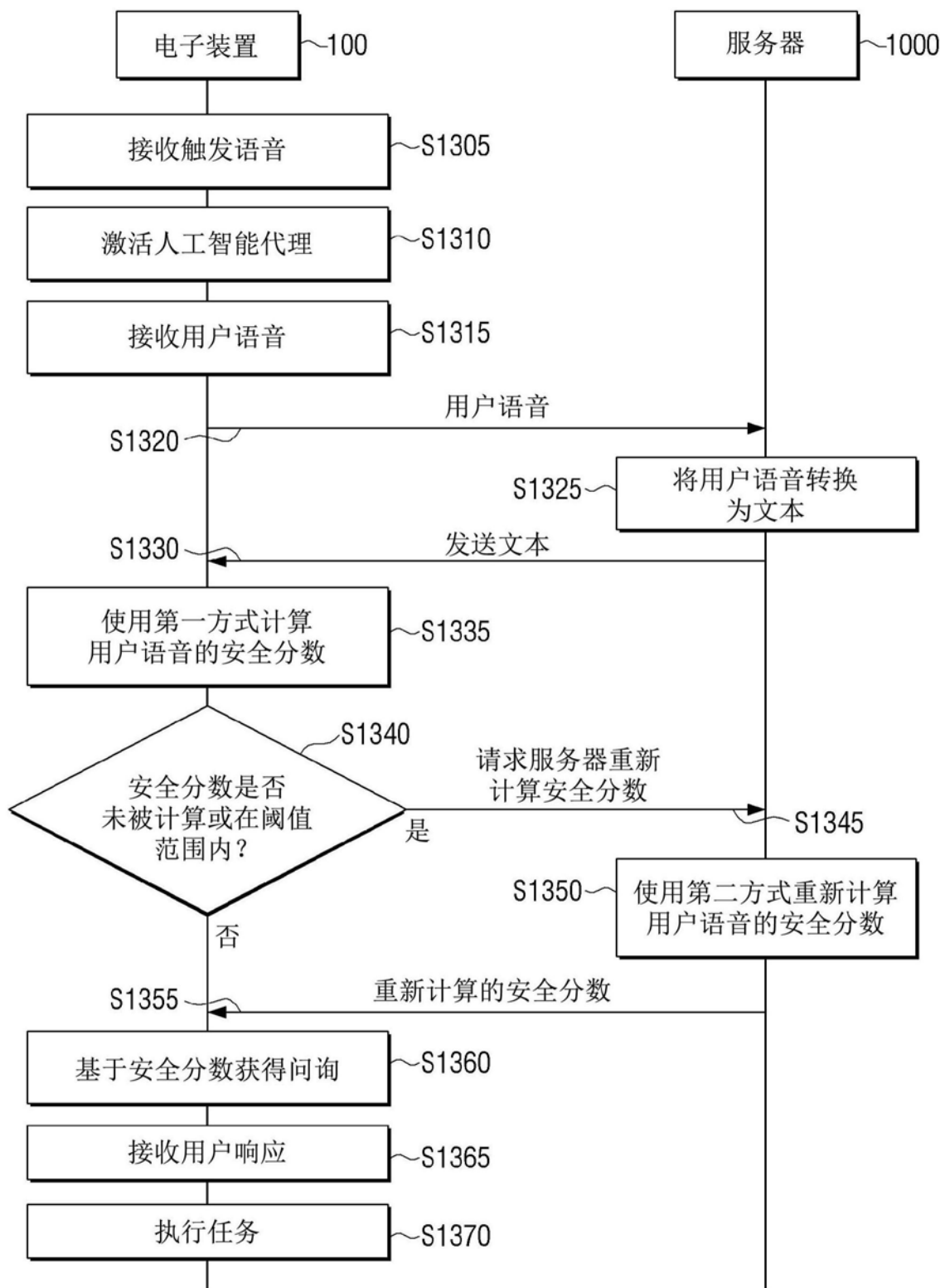


图13