



(10) 授权公告号 CN 112420032 B

(45) 授权公告日 2024. 06. 28

(21) 申请号 202010519019.2

(22) 申请日 2020.06.09

(65) 同一申请的已公布的文献号

申请公布号 CN 112420032 A

(43) 申请公布日 2021.02.26

(30) 优先权数据

10-2019-0127219 2019.10.14 KR

62/889,232 2019.08.20 US

(73) 专利权人 三星电子株式会社

地址 韩国京畿道水原市灵通区三星路129号

(72) 发明人 张赛博 高炫睦 李景薰

库纳尔·沙玛

拉加文德·哈努曼塔斯·拉马塞特蒂

(74) 专利代理机构 北京英赛嘉华知识产权代理有限公司 11204

专利代理师 王达佐 杨莘

(51) Int.Cl.

G10L 15/22 (2006.01)

G10L 15/18 (2013.01)

G10L 15/06 (2013.01)

G10L 15/26 (2006.01)

(56) 对比文件

CN 107004410 A, 2017.08.01

CN 107210033 A, 2017.09.26

审查员 郝丽娜

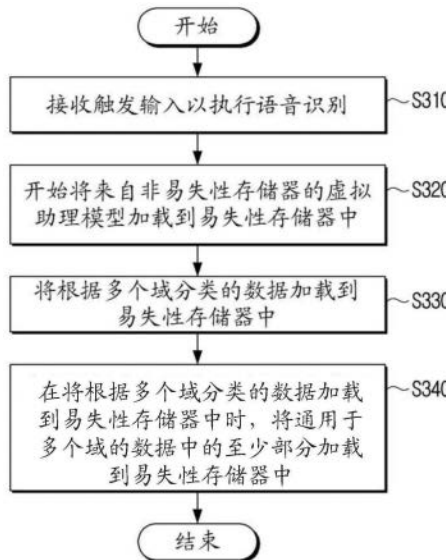
权利要求书2页 说明书20页 附图9页

(54) 发明名称

电子设备及用于控制电子设备的方法

(57) 摘要

公开了能够有效地进行语音识别和自然语言理解的电子设备及其控制方法。电子设备包括麦克风、非易失性存储器、易失性存储器和处理器,其中非易失性存储器配置为存储虚拟助理模型数据,虚拟助理模型数据包括根据多个域分类的数据和通用于多个域的数据,处理器配置为:基于通过麦克风接收触发输入以执行对用户语音的语音识别、启动将虚拟助理模型数据从非易失性存储器加载到易失性存储器中、将根据多个域分类的数据中的第一数据加载到易失性存储器中、并且在将第一数据加载到易失性存储器中的同时,将通用于多个域的数据的至少一部分加载到易失性存储器中。



1. 一种电子设备, 包括:

麦克风;

非易失性存储器, 配置为存储虚拟助理模型数据, 所述虚拟助理模型数据包括根据多个域分类的数据和通用于多个域的数据;

易失性存储器; 以及

处理器, 配置为:

基于通过所述麦克风接收用于对用户语音执行语音识别的触发输入, 启动将所述虚拟助理模型数据从所述非易失性存储器加载到所述易失性存储器中;

将所述根据多个域分类的数据中的、对应于与通过所述麦克风接收的第一用户语音有关的至少一个第一域的第一数据加载到所述易失性存储器中, 以及在将所述第一数据加载到所述易失性存储器中的同时, 将所述通用于多个域的数据中的至少一部分加载到所述易失性存储器中。

2. 根据权利要求1所述的电子设备, 其中, 所述处理器进一步配置为:

基于通过所述麦克风接收所述第一用户语音, 通过对所述第一用户语音进行语音识别, 获得与所述第一用户语音对应的第一文本; 以及

基于所获得的第一文本, 识别所述多个域中与所述第一用户语音对应的所述至少一个第一域。

3. 根据权利要求2所述的电子设备, 其中, 所述处理器进一步配置为: 基于所识别的所述至少一个第一域, 将所述根据多个域分类的数据中的、与所识别的所述至少一个第一域对应的数据作为所述第一数据加载到所述易失性存储器中。

4. 根据权利要求3所述的电子设备, 其中, 所述处理器进一步配置为:

对加载的、与所识别的所述至少一个第一域对应的所述第一数据进行处理; 以及

基于经处理的第一数据, 对所述第一文本进行自然语言理解。

5. 根据权利要求3所述的电子设备, 其中, 所述处理器进一步配置为: 将加载的、与所识别的所述至少一个第一域对应的所述第一数据在所述易失性存储器中保持预定时长。

6. 根据权利要求5所述的电子设备, 其中, 所述处理器进一步配置为:

基于在所述易失性存储器中保持与所识别的所述至少一个第一域对应的所述第一数据的同时通过所述麦克风接收第二用户语音, 识别与所述第二用户语音对应的至少一个第二域; 以及

将所述根据多个域分类的数据中的、与所识别的至少一个第二域对应且不同于所述第一数据的第二数据加载到所述易失性存储器中。

7. 根据权利要求1所述的电子设备, 其中, 所述处理器进一步配置为:

基于所述电子设备被开启, 将所述通用于多个域的数据中的与用于识别所述触发输入的模块对应的数据加载到所述易失性存储器中; 以及

基于使用加载的与用于识别所述触发输入的模块对应的数据对所述触发输入进行识别, 确定对所述用户语音执行所述语音识别。

8. 一种用于控制电子设备的方法, 所述电子设备包括易失性存储器和非易失性存储器, 所述非易失性存储器存储虚拟助理模型数据, 所述虚拟助理模型数据包括根据多个域分类的数据和通用于多个域的数据, 所述方法包括:

基于接收用于对用户语音进行语音识别的触发输入,启动将所述虚拟助理模型数据从所述非易失性存储器加载到所述易失性存储器中;

将所述根据多个域分类的数据中的、对应于与通过麦克风接收的第一用户语音有关的至少一个第一域的第一数据加载到所述易失性存储器中;以及

在将所述第一数据加载到所述易失性存储器中的同时,将所述通用于多个域的数据中的至少一部分加载到所述易失性存储器中。

9. 根据权利要求8所述的方法,还包括:

基于接收所述第一用户语音,通过对所述第一用户语音进行语音识别,获得与所述第一用户语音对应的第一文本;以及

基于所获得的第一文本,识别所述多个域中与所述第一用户语音对应的所述至少一个第一域。

10. 根据权利要求9所述的方法,其中,加载所述第一数据包括:

基于所识别的所述至少一个第一域,将所述根据多个域分类的数据中的、与所识别的所述至少一个第一域对应的数据作为所述第一数据加载到所述易失性存储器中。

11. 根据权利要求10所述的方法,还包括:

对加载的、与所识别的所述至少一个第一域对应的所述第一数据进行处理;以及

基于经处理的第一数据,对所述第一文本进行自然语言理解。

12. 根据权利要求10所述的方法,还包括:

将加载的、与所识别的所述至少一个第一域的对应的所述第一数据在所述易失性存储器中保持预定时长。

13. 根据权利要求12所述的方法,还包括:

基于在所述易失性存储器中保持与所识别的所述至少一个第一域对应的所述第一数据的同时接收第二用户语音,识别与所述第二用户语音对应的至少一个第二域;以及

将所述根据多个域分类的数据中的、与所识别的至少一个所述第二域对应且不同于所述第一数据的第二数据加载到所述易失性存储器中。

14. 根据权利要求8所述的方法,还包括:

基于所述电子设备被开启,将所述通用于多个域的数据中的与用于识别所述触发输入的模块对应的数据加载到所述易失性存储器中;以及

基于使用加载的与用于识别所述触发输入的模块对应的数据对所述触发输入进行识别,确定对所述用户语音执行所述语音识别。

15. 一种非暂时性计算机可读记录介质,其上记录有指令,所述指令能够由至少一个处理器执行以用于执行根据权利要求8所述的方法。

电子设备及用于控制电子设备的方法

[0001] 相关申请的交叉引用

[0002] 本申请基于并要求2019年10月14日在韩国知识产权局提交的第10-2019-0127219号韩国专利申请的优先权并要求2019年8月20日在美国专利商标局提交的第62/889232号美国临时专利申请的权益,这些申请公开内容通过引用整体并入本文。

技术领域

[0003] 本公开涉及一种电子设备和用于控制该电子设备的方法,更具体地,涉及能够有效地执行语音识别和自然语言理解过程的电子设备及其控制方法。

背景技术

[0004] 近年来,人工智能(AI)助理领域的技术发展已加速,并且使用语音识别和自然语言理解的技术已应用于各种电子设备。在这种情况下,不仅提高了用户语音的响应性能,而且作为需要解决的主要任务之一,还提高了用户语音的响应速度。

[0005] 为了解决用户语音的响应速度问题,存在背景中预加载用于执行语音识别和自然语言理解的所有数据并等待接收用户语音的方法。然而,利用这种方法,存在的问题在于,大的负载不能添加到存储器,并且随着要处理的信息变化更大,负载增加更多。

[0006] 还存在以下的方法,即在接收到用户的触发输入的情况下在加载和初始化用于处理用户语音所必需的所有数据之后处理用户语音。然而,利用这种方法,存在的问题在于,输出对用户语音的响应的的时间被延迟了与加载和初始化数据所需的时间一样多的时间。

[0007] 因此,需要用于在减少存储器的负载的同时提高用户语音的处理速度的技术。

发明内容

[0008] 提供了能够有效地控制与语音识别和自然语言理解相关的数据的加载过程以及用户语音的处理过程的电子设备及其控制方法。

[0009] 附加方面将部分地在随后的描述中阐述,并且部分地将通过描述显而易见,或者可以通过所呈现的实施方式的实践来获知。

[0010] 根据本公开的一方面,电子设备包括:麦克风、非易失性存储器、易失性存储器和处理器,其中非易失性存储器配置为存储虚拟助理模型数据,虚拟助理模型数据包括根据多个域分类的数据和通用于多个域的数据,处理器配置为:基于通过麦克风接收用于对用户语音进行语音识别的触发输入,启动将虚拟助理模型数据从非易失性存储器加载到易失性存储器中,将根据多个域分类的数据中的第一数据加载到易失性存储器中,以及在将第一数据加载到易失性存储器中的同时,将通用于多个域的数据中的至少一部分加载到易失性存储器中。

[0011] 处理器还可配置为:基于通过麦克风接收第一用户语音,通过对第一用户语音进行语音识别,获得与第一用户语音对应的第一文本;以及基于所获得的第一文本,识别多个域中与第一用户语音对应的至少一个第一域。

[0012] 处理器可进一步配置为:基于所识别的至少一个第一域,将根据多个域分类的数据中的、与所识别的至少一个第一域对应的数据作为第一数据加载到易失性存储器中。

[0013] 处理器可进一步配置为:对加载的、与所识别的至少一个第一域对应的第一数据进行处理;以及基于经处理的第一数据,对第一文本进行自然语言理解。

[0014] 处理器可进一步配置为:将加载的、与所识别的至少一个第一域的对应的第一数据在易失性存储器中保持预定时长。

[0015] 处理器可进一步配置为:基于在易失性存储器中保持与所识别的至少一个第一域对应的第一数据的同时、通过麦克风接收第二用户语音,识别与第二用户语音对应的至少一个第二域;以及将根据多个域分类的数据中的、与所识别的至少一个第二域对应且不同于第一数据的第二数据加载到易失性存储器中。

[0016] 处理器还可配置为:基于电子设备被开启,将通用于多个域的数据中的与用于识别触发输入的模块对应的数据加载到易失性存储器中;以及基于使用加载的、与用于识别触发输入的模块对应的数据识别触发输入,确定对用户语音执行语音识别。

[0017] 根据本公开的另一方面,提供了用于控制电子设备的方法,其中电子设备包括易失性存储器和非易失性存储器,非易失性存储器存储虚拟助理模型数据,虚拟助理模型数据包括根据多个域分类的数据和通用于多个域的数据,该方法包括:基于接收用于对用户语音进行语音识别的触发输入,启动将虚拟助理模型数据从非易失性存储器加载到易失性存储器中;将根据多个域分类的数据中的第一数据加载到易失性存储器中;以及在将第一数据加载到易失性存储器中的同时,将通用于多个域的数据中的至少一部分加载到易失性存储器中。

[0018] 方法还可包括:基于接收第一用户语音,通过对第一用户语音进行语音识别,获得与第一用户语音对应的第一文本;以及基于所获得的第一文本,识别多个域中与第一用户语音对应的至少一个第一域。

[0019] 加载第一数据可以包括:基于所识别的至少一个第一域,将根据多个域分类的数据中与所识别的至少一个第一域对应的数据作为第一数据加载到易失性存储器中。

[0020] 方法还可以包括:对加载的与所识别的至少一个第一域对应的第一数据进行处理;基于处理后的第一数据对第一文本进行自然语言理解。

[0021] 方法还可以包括将加载的与所识别的至少一个第一域对应的第一数据在易失性存储器中保持预定时长。

[0022] 方法还可以包括:基于在易失性存储器中保持与所识别的至少一个第一域对应的第一数据的同时接收第二用户语音,识别与第二用户语音对应的至少一个第二域;以及将根据多个域分类的数据中的、与所识别的至少一个第二域对应且不同于第一数据的第二数据加载到易失性存储器中。

[0023] 方法还可以包括:基于电子设备被开启,将通用于多个域的数据中的与用于识别触发输入的模块对应的数据加载到易失性存储器中;以及基于使用加载的、与用于识别触发输入的模块对应的数据识别触发输入,确定对用户语音执行语音识别。

[0024] 根据本发明的另一方面,一种非暂时性计算机可读记录介质具有存储于其上的指令,该指令可由至少一个处理器执行以执行上述方法。

[0025] 根据本发明的另一方面,一种电子设备包括非易失性存储器、易失性存储器以及

处理器,非易失性存储器配置为存储虚拟助理模型数据,虚拟助理模型数据包括根据多个域分类的数据和通用于多个域的数据,处理器配置为:基于确定接收到对用户语音执行语音识别的触发输入,将根据多个域分类的数据中的第一数据加载到易失性存储器中,并且在将第一数据加载到易失性存储器中的同时,将通用于多个域的数据中的至少一部分加载到易失性存储器中。

[0026] 处理器还可配置为:基于接收第一用户语音,通过对第一用户语音进行语音识别,获得与第一用户语音对应的第一文本;以及基于所获得的第一文本,识别多个域中与第一用户语音对应的至少一个第一域。

[0027] 处理器可进一步配置为基于所识别的至少一个第一域,将根据多个域分类的数据中的、与所识别的至少一个第一域对应的数据作为第一数据加载到易失性存储器中。

[0028] 处理器可进一步配置为:对加载的、与所识别的至少一个第一域对应的第一数据进行处理;以及基于经处理的第一数据,对第一文本进行自然语言理解。

[0029] 处理器可进一步配置为:基于在易失性存储器中保持与所识别的至少一个第一域对应的第一数据的同时接收第二用户语音,识别与第二用户语音对应的至少一个第二域;以及将根据多个域分类的数据中的、与所识别的至少一个第二域对应且不同于第一数据的第二数据加载到易失性存储器中。

附图说明

[0030] 通过下面结合附图进行的描述,本公开的某些实施方式的以上和其它方面、特征和优点将变得更加显而易见,在附图中:

[0031] 图1是简要示出根据实施方式的电子设备的配置的框图;

[0032] 图2是示出根据实施方式的存储在非易失性存储器中的虚拟助理模型数据的结构的模拟图;

[0033] 图3是示出根据实施方式的控制电子设备的过程的流程图;

[0034] 图4A至图4C是示出根据实施方式与加载虚拟助理模型数据的过程相关的各种实施方式的图;

[0035] 图5是示出根据实施方式的电子设备的详细配置的图;

[0036] 图6A和图6B是示出与依次接收触发输入、第一用户语音和第二用户语音的示例相关联的实施方式的图;

[0037] 图7是更详细地示出了以下过程的流程图,在该过程中基于第一用户语音被接收,然后第二用户语音被顺序地接收,通过将第一用户语音相关联的至少一个域和与第二用户语音相关联的至少一个域进行比较来确定是否加载根据多个域分类的数据;

[0038] 图8是示出由根据实施方式的电子设备基于第一用户语音被接收然后第二用户语音被顺序地接收来处理第二用户语音的过程的图;以及

[0039] 图9是示出根据实施方式的电子设备的详细配置的框图。

具体实施方式

[0040] 将参考附图更详细地描述各实施方式,使得本领域技术人员能够容易地实现本公开。

[0041] 图1是简要示出根据实施方式的电子设备100的配置的框图。

[0042] 如图1所示,电子设备100包括麦克风110、非易失性存储器120、易失性存储器130和处理器140(例如,至少一个处理器)。

[0043] 麦克风110可以获取外部产生的声音或语音的信号。麦克风110可以获取由外部产生的声音或语音引起的振动,并将所获取的振动转换为电信号。

[0044] 麦克风110可以获取由用户发声生成的用户语音的信号。本文中的术语“用户语音”可以与诸如“用户声音”或“用户发声”的术语互换。所获取的信号可转换成数字信号并存储在易失性存储器130中。麦克风110可以包括模数(A/D)转换器和/或可以与位于麦克风外部的A/D转换器相关联地操作。经由麦克风110接收的至少一些用户语音可输入到语音识别和自然语言理解模型。具体地,在对应于虚拟助理模型的触发输入之后,经由麦克风110接收的用户语音可以由输入到语音识别和自然语言理解模型的用户语音来指定。

[0045] 例如,如果“你好,ABC”是预定的触发词并且通过麦克风110接收到用户语音“你好,ABC,今天的天气如何?”,那么在接收“你好,ABC”之后通过麦克风110接收的“今天的天气如何?”可以指定为作为语音识别和自然语言理解模型的输入的用户语音。在该示例中,虚拟助理模型可配置为包括至少识别触发词的模型、语音识别模型和自然语言理解模型。

[0046] 在下文中,对“用户语音”的引用可以指在接收到触发输入之后,作为通过麦克风110接收的用户语音而输入到语音识别和自然语言理解模型的用户语音。

[0047] 非易失性存储器120指的是即使在电源断电的情况下也能保持存储的信息的存储器。例如,非易失性存储器可以包括闪存、可编程只读存储器(PROM)、磁阻随机存取存储器(MRAM)或电阻RAM(RRAM)中的至少一个。

[0048] 易失性存储器130指的是需要连续供电以维持所存储的信息的存储器。例如,易失性存储器130可以包括动态随机存取存储器(DRAM)和静态RAM(SRAM)中的至少一个。

[0049] 在下文中,假设易失性存储器130是与处理器140分离的配置,但是这是为了清楚地描述根据实施方式的电子设备100的操作,并且易失性存储器130可以实现为作为根据另一实施方式的处理器140的组件而包括在处理器140中。

[0050] 根据实施方式的非易失性存储器120可以存储与语音识别模型和自然语言理解模型相关联的数据。与存储在非易失性存储器120中的语音识别模型和自然语言理解模型相关联的数据可加载到易失性存储器130中。为了便于描述,用于(例如,需要)提供诸如语音识别模型和自然语言理解模型的虚拟助理服务的、与多个模型相关联的数据被称为“虚拟助理模型数据”。

[0051] 具体地,虚拟助理模型数据的一部分可以被分类为“根据多个域分类的数据”。这里,域是指根据与用户语音或用户语音的主题相对应的控制命令的类型来分类的一种类别。例如,多个域可以包括诸如“天气”、“日程”、“消息”、“图库”、“设置”等域。根据多个域分类的数据可以构造成用于通过如下所述的自然语言理解(NLU)模型来理解自然语言,并且以数据库的形式存储在非易失性存储器120中。

[0052] 同时,虚拟助理模型数据的另一部分可分类为“通用于多个域的数据”。这里,“通用于多个域”是指与用于执行用户语音的语音识别和自然语言理解的多个模型相关联的数据,而不管何种域对应于用户语音。

[0053] 多个模型可以包括识别来自用户语音的触发词的语音唤醒(WoV)模型、执行语音

识别的自动语音识别 (ASR) 模型、以及执行自然语言理解的自然语言理解 (NLU) 模型等,并且每个模型可以包括配置为执行用于(例如,必须用于)语音识别和自然语言理解的特定功能的多个模块。因此,根据本公开的模型的结构可以以包括对模型中包含的模块的结构的意义来进行。

[0054] 下面参考图2更详细地描述存储在非易失性存储器120中的虚拟助理模型数据的结构。

[0055] 处理器140控制电子设备100的整体操作。具体地,如上所述,处理器140可以连接到电子设备100的组件,包括如上所述的麦克风110、非易失性存储器120和易失性存储器130,以控制电子设备100的整体操作。处理器140可以包括中央处理单元 (CPU)、图形处理单元 (GPU)、主处理单元 (MPU) 等中的至少一个。

[0056] 处理器140可以将存储在非易失性存储器120中的虚拟助理模型数据加载到易失性存储器130中。加载是以下操作或包括以下操作:即调用、复制或将存储在非易失性存储器120中的数据移动到易失性存储器130并存储该数据以使处理器140可以访问该数据的操作。

[0057] 如果存储在非易失性存储器120中的数据之中的、与用于执行语音识别和自然语言理解的多个模型相关联的数据被加载到易失性存储器130中,则处理器140可以初始化多个模型。初始化是指当(或基于)与多个模型相关联的数据从非易失性存储器120加载到易失性存储器130中时,用于(例如,必须用于)执行多个模型中的每一个的过程的数据是以资源或数据库的形式从与多个模型相关联的数据中调用的。也就是说,初始化是指或包括将存储在非易失性存储器120中的数据重新配置为可访问文件或数据库的形式并存储在易失性存储器130中,使得加载到易失性存储器130并被执行的软件程序可以利用存储在非易失性存储器120中的数据。将与多个模型相关联的数据加载到易失性存储器130中可以包括加载与多个模型相关联的数据,然后初始化多个模型。

[0058] 如果(或基于)与多个模型相关联的数据被加载到易失性存储器130中,则处理器140可以通过多个模型执行语音识别和自然语言理解。具体地,当(或基于)通过麦克风110接收到用户语音时,处理器140可以执行语音识别并获取与接收到的用户语音对应的文本。当获取与用户语音对应的文本时,处理器140可以通过对所获取的文本执行自然语言理解来获取关于用户意图的信息。

[0059] 更具体地,如果通过麦克风110获取的用于用户语音的信号被转换成数字信号形式的信号并存储在如上所述的易失性存储器130中,则处理器140可以使用存储在易失性存储器130中的信号作为ASR模型的输入数据,来获取与用户语音对应的文本,并将所获取的文本存储在易失性存储器130中。处理器140可以通过使用存储在易失性存储器130中的文本作为NLU模型的输入数据,来获取关于用户意图的信息。下面参考图2更详细地描述包括ASR模型和NLU模型的多个模型。

[0060] 根据各种实施方式,处理器140可以有效地控制将存储在非易失性存储器120中的虚拟助理模型数据加载到易失性存储器130中的过程,以及对通过麦克风110接收的用户语音进行处理的过程,这将在下面参考图3、图4A至图4C、图5、图6A至图6B以及图7到图9进行描述。现在将参考图2描述根据实施方式存储在非易失性存储器120中的数据的结构。

[0061] 图2是示出根据实施方式的存储在非易失性存储器120中的虚拟助理模型数据的

结构的模拟图。

[0062] 参照图2,存储在非易失性存储器120中的虚拟助理模型数据可包括与NLU模型相关联的数据、与ASR模型相关联的数据以及与WoV模型相关联的数据。NLU模型相关数据可以被分类为“按域的数据”,诸如对应于第一域的数据、对应于第二域的数据、对应于第三域的数据等。此外,“与NLU模块相关联的数据”可以为例如与域分类器相关联的数据、与数据加载器/处理器相关联的数据、与NLU引擎相关联的数据、与标准处理器相关联的数据等。

[0063] 可以根据多个域对NLU模型相关数据中的每个域的数据进行分类,从而可将通用于多个域的数据之中的NLU模块相关数据与ASR模型相关数据和WoV模型相关数据区分开。因此,为了便于描述,被分类成NLU模型相关数据中用于每个域的数据和其它数据的数据被称为“根据多个域分类的数据10”和“通用于多个域的数据20”。

[0064] 电子设备100可以通过多个模型对用户语音执行语音识别和自然语言理解,并且下面描述多个模块以及包括在多个模型中的多个模型。

[0065] NLU模型可以包括多个模块,例如域分类器、数据加载器、数据处理器、NLU引擎、标准处理器等。

[0066] 域分类器是指用于对与用户语音相关联的域进行分类的模块。域分类器可以对多个域中与通过麦克风接收的用户语音相关联的至少一个域进行分类。

[0067] 数据加载器是指将存储在非易失性存储器120中的虚拟助理模型数据的根据多个域分类的数据10加载到易失性存储器130中的模块。特别地,数据加载器可以选择性地将与通过域分类器识别的至少一个域对应的数据加载到易失性存储器130中。下面描述根据本公开的选择性加载过程。

[0068] 数据处理器是指用于处理根据多个域分类的数据10的模块。具体地,数据处理器可以将根据多个域分类的数据10转换为图形格式,或者将数据10序列化为参考数据。例如,数据处理器可以基于加载到非易失性存储器120中的数据生成n图形或JavaScript对象符号(JSON)等。根据实施方式的数据处理器可以处理根据多个域而分类的数据10中、通过数据加载器加载到易失性存储器130中的数据。

[0069] NLU引擎是指使用根据多个域分类的数据10来执行自然语言理解的模块。具体地,当(或基于)如下所述通过ASR模型获取对应于用户语音的文本时,NLU引擎可以基于所获取的文本获取关于与用户语音对应的用户意图的信息。此外,在通过NLU引擎获取关于用户意图的信息时,可以使用根据多个域分类的数据10,更具体地,可以使用通过数据加载器加载到易失性存储器130中并由数据处理器处理的数据。

[0070] NLU引擎可以是基于规则的规则或基于规则的引擎,或者可以是基于人工神经网络的深度神经网络(DNN)引擎。根据实施方式,NLU引擎可用于包括二者。NLU引擎可以对通过ASR模型获取(或获得)的文本进行句法分析和语义分析,以获取(或获得)关于用户意图的信息。

[0071] 具体地,NLU引擎可以对所获取的文本的语法单元(例如,单词、短语、语素等中的至少一个)进行分类,并识别所分类的语法单元具有哪个语法元素。NLU引擎可基于所识别的语法元素来确定文本的含义。NLU引擎还可以通过将文本的所确定的含义与识别的域、包括在识别的域中的多个意图以及多个参数进行匹配,来获取关于用户意图的信息。例如,NLU引擎可以通过将文本的所确定的含义与作为识别的域的“警报”、作为识别的域中包括

的多个意图的“警报设置”和“警报释放”、作为表示用户意图的参数的“时间”、“重复次数”、“警报声音”等进行匹配,来获取关于用户意图的信息。

[0072] 标准处理器是指将用户语音中包括的非结构化信息转换为结构化信息的模块。具体地,如果用户语音中包括非结构化信息,则标准处理器可以获取与非结构化信息相对应的结构化信息。例如,如果用户语音包括非结构化信息,例如“请让我知道星期五的天气”,则标准处理程序可以获取例如“年:2019,月:4,日:19,周:星期五”的信息作为与非结构化信息对应的结构化信息。如果获取了与非结构化信息对应的结构化信息,则NLU引擎可以基于所获取的结构化信息来获取关于用户意图的信息。

[0073] ASR模型是指对用户语音执行语音识别的模型。ASR模型可以包括声学模型(AM)、发音模型(PM)、语言模型(LM)等中的至少一个。

[0074] AM可以提取接收的用户语音的声学特征并获取音素序列。PM可以包括发音词典(发音辞典),并且将所获取的音素序列映射到单词以获取单词序列。LM可以将概率分配给所获取的单词序列。也就是说,ASR模型可以通过诸如AM、PM和LM的人工智能(AI)模型来获取对应于用户语音的文本。ASR模型可以包括端到端语音识别模型,其中AM、PM和LM的分量被组合到单个神经网络中。

[0075] WoV模型是指用于从用户语音中识别触发词的模型。WoV模型识别(或确定)接收的用户语音是否包括预定的触发词,并且如果识别出(或基于识别或确定)用户语音包括预定的触发词,则可以确定出在接收到触发词之后对通过麦克风接收的用户语音执行语音识别。

[0076] 根据一个或多个实施方式使用如上所述的虚拟助理模型数据执行各种操作的过程可以由流控制器控制。流控制器是指控制加载过程和数据处理的整体操作以控制关联的软件模块的模块。具体地,在加载如图2所示的虚拟助理模型数据之前,可以将流控制器加载到易失性存储器130中,以确定虚拟助理模型数据的加载顺序和加载目标数据,将虚拟助理模型数据加载到易失性存储器130中,并通过多个模型和多个模块执行各种操作。下面参考图5更详细地描述流控制器。

[0077] 如上所述,与多个模型和多个模块关联的数据可以包括软件数据和/或被实现为执行多个模型和多个模块的编程语言的指令。多个模型和多个模块中的至少一个可以通过AI模型来实现。

[0078] 已描述了多个模型和包括在多个模型中的多个模块中的每个的功能,但是这仅仅是为了阐明和描述各种实施方式,并且本公开不限于如上所述假定的名称和分类标准。

[0079] 将基于对数据的结构和多个模块的描述来描述各种实施方式。

[0080] 图3是示出根据实施方式的控制电子设备100的过程的流程图。

[0081] 参照图3,在操作S310中,电子设备100可以接收用于执行语音识别的触发输入(例如,预定的触发、短语、单词等)。在图1的描述中,包括预定触发词的用户语音被描述为触发输入的示例,但是触发输入不限于此。下面参考图9描述各种类型的触发输入。

[0082] 一旦(或基于)接收到触发输入,在操作S320中,电子设备100就可开始将虚拟助理模型数据从非易失性存储器120加载到易失性存储器中。具体地,当接收到触发输入时,电子设备100可通过用于将存储在非易失性存储器中的虚拟助理模型数据加载到易失性存储器中的模块,开始将存储在非易失性存储器中的与NLU模型和ASR模型相关联的数据加载到

易失性存储器中。在加载包括在虚拟助理模型数据中的各种数据的过程中,可以应用如下所述的各种实施方式。

[0083] 开始将与ASR模型和NLU模型关联的数据加载到易失性存储器中不限于将与ASR模型和NLU模型关联的所有数据加载到易失性存储器中。也就是说,本公开可应用于以下情况,即在接收触发输入之前将与ASR模型和NLU模型关联的一些数据预先加载到易失性存储器中,并且在接收触发输入的情况下(或基于接收触发输入)将未被加载到易失性存储器中的与ASR模型和NLU模型相关联的数据中的其它数据加载到易失性存储器中。

[0084] 在操作S330,电子设备100可以根据多个域分类的数据加载到易失性存储器中。当在操作S320中(或基于)虚拟助理模型数据中的、与数据加载器相关联的数据在虚拟助理模型数据的加载开始之后被加载到易失性存储器时,电子设备100可以通过数据加载器将根据多个域分类的数据加载到易失性存储器中。

[0085] 根据实施方式,电子设备100可以选择性地将根据多个域分类的数据中的一些加载到易失性存储器中。这里,选择性地加载根据多个域分类的数据的一部分意味着或包括仅加载根据多个域分类的数据中的、通过麦克风接收的与用户语音相关联的数据,这将在下面关于从用户语音的接收中选择性地加载根据多个域分类的数据的过程进行详细描述。

[0086] 具体地,在操作S310中接收到执行语音识别的触发输入之后,电子设备100可以经由麦克风接收第一用户语音。如果在操作S320中加载虚拟助理模型数据之后,将与虚拟助理模型数据的ASR模型相关联的数据加载到易失性存储器中,则电子设备100可以使用第一用户语音作为ASR模型的输入数据来获取与第一用户语音对应的第一文本。为了便于描述,通过麦克风110接收的用户语音在本文被称为第一用户语音,以与下面参考图6A至图6B和图7至图8描述的第二用户语音区别开。

[0087] 当(或基于)虚拟助理模型数据中与域分类器相关联的数据被加载到易失性存储器中时,电子设备100可以通过使用所获取的第一文本作为域分类器的输入数据,来识别多个域中与第一用户语音相关联的至少一个域。如果识别出与第一用户语音相关联的至少一个域,则电子设备100可以根据多个域分类的数据中的、与被识别为与第一用户语音相关联的至少一个域对应的数据加载到易失性存储器中。

[0088] 当根据多个域分类的数据被加载到易失性存储器中时,在操作S340中,电子设备100可以将通用于多个域的数据的至少一部分加载到易失性存储器中。换句话说,在将根据多个域分类的数据并行加载到易失性存储器中之后,电子设备100可以将通用于多个域的数据中的至少一些与根据多个域分类的数据一起并行加载。例如,当将根据多个域分类的数据加载到易失性存储器中时,电子设备100可以将通用于多个域的数据的、与ASR模型相关联的数据加载到易失性存储器中。

[0089] 根据实施方式,在如上所述识别与第一用户语音相关联的至少一个域之后,将对应于所识别的至少一个域的数据加载到易失性存储器中。在加载对应于所识别的域的数据的同时,电子设备100可将通用于多个域的数据中的至少一些加载到易失性存储器中。例如,电子设备100可以在将与识别出的至少一个域对应的数据加载到易失性存储器中的同时,将通用于多个域的数据中的、与NLU引擎相关联的至少一些数据加载到易失性存储器中。

[0090] 参考图4A至图4C,更具体地描述了虚拟助理模型数据的并行加载和选择性加载的

各种实施方式。

[0091] 图4A至图4C是示出根据实施方式的与加载虚拟助理模型数据的过程有关的各种实施方式的图。

[0092] 图4A至图4C以对应于每个模型和模块的一种类型的块示出了根据各种实施方式的基于时间坐标轴的虚拟助理模型数据的并行加载过程、多个模型以及通过多个模型进行处理的过程。根据实施方式,近似地示出了每个块的长度,并且根据本公开的模型和模块的加载过程和模型以及通过模型和模块的处理的时间不必对应于如图4A至图4C所示的块的长度。

[0093] 图4A是示出NLU模型中包括的多个模块之间的并行加载过程的图。如图4A所示,NLU模型包括域分类器411、数据加载器412、数据处理器413、标准处理器421和NLU引擎422。

[0094] 如上参考图3所述,电子设备100可以将将在通用于多个域的数据中的至少一些加载到易失性存储器中,同时将与被识别为对应于用户语音的至少一个域对应的数据加载到易失性存储器中。例如,如图4A所示,电子设备100通过数据加载器412将根据多个域分类的数据之中的与被识别为关联于用户语音的至少一个域对应的数据加载到易失性存储器中。在加载与所识别的域对应的数据期间,可并行地将通用于多个域的数据中的、与NLU引擎相关联的数据加载到易失性存储器中。

[0095] 电子设备100可以首先将与域分类器411相关联的数据加载到易失性存储器中,并且通过域分类器411识别与用户语音对应的至少一个域,以选择性地将与所识别的域对应的数据加载到易失性存储器中。

[0096] 虽然上面已经描述了根据多个域分类的数据与通用于多个域的数据之间的并行加载,但是根据一个或多个其它实施方式的电子设备100可以将通用于多个域的数据中的一些和一些其它数据并行加载到易失性存储器中。

[0097] 例如,如图4A所示,电子设备100可以在将与域分类器411相关联的数据加载到易失性存储器中的同时,并行地加载与标准处理器421相关联的数据和与NLU引擎相关联的数据。电子设备100还可以在将数据加载器412加载到易失性存储器中的同时,并行地将与NLU引擎相关联的数据加载到易失性存储器中。电子设备100可以在将数据处理器413加载到易失性存储器中的同时,将通用于多个域的数据中的与NLU引擎相关联的数据加载到易失性存储器中。

[0098] 虽然上面主要描述了数据加载过程,但是应当理解,电子设备100可以在通过多个模型和多个模块执行处理过程的同时,将通用于多个域的数据中的至少一个并行地加载到易失性存储器中。

[0099] 例如,与标准处理器421相关联的数据和与NLU引擎相关联的数据可以在执行域分类器411的处理过程以及域分类器411的加载过程的同时被加载到易失性存储器中。当通过数据处理器413运行被加载到易失性存储器中的通用于多个域的数据时,与NLU引擎相关联的数据可以被加载到易失性存储器。

[0100] 如上所述,当通过域分类器411识别与用户语音关联的至少一个域时,数据加载器412和数据处理器413可以根据所识别的域执行加载和处理。就这一点而言,数据加载器412和数据处理器413可以被称为“依赖于域分类器411的模块”。在NLU模型中包括的多个模块中,标准处理器421和NLU引擎422可以不是(例如,可以几乎不需要)通过依赖于域分类器

411而被加载到易失性存储器中,并且在与域分类器411相关联的处理和加载期间可期望进行预加载。在这种意义上,域分类器411、数据加载器412和数据处理器413之外的其余模块可以被称为“不依赖于域分类器411的模块”。

[0101] 参照图4A,电子设备100可以将与域分类器411相关联的数据以及模块410中的至少一个依赖于域分类器411的模块加载到易失性存储器中,并且在通过每个模块执行处理的同时,电子设备100可以将与模块420中至少一个不依赖于域分类器411的模块相关联的数据加载到易失性存储器中。在依赖于域分类器411的模块和不依赖于域分类器411的模块之间的并行加载的实施方式同样可以应用于以下示例,其中依赖于域分类器411的模块和不依赖于域分类器411的模块包括与图4A中所示的示例不同的模块。

[0102] 在NLU模型中包括的多个模块中标准处理器421与NLU引擎422之间没有相互顺序的关系,并且如图4A所示的实施方式仅是根据本公开的各种实施方式中的一个实施方式。例如,根据另一实施方式,电子设备100可以在将与数据加载器412相关联的数据或与数据处理器413相关联的数据加载到易失性存储器中的同时,将与标准处理器421相关联的数据加载到易失性存储器中。

[0103] 图4B包括图4A中描述的NLU模型的加载和处理过程之外,还包括根据实施方式的ASR模型430的加载和处理过程。

[0104] 如以上参考图4A所述,电子设备100可以通过域分类器411识别对应于用户语音的至少一个域。然而,为了执行域分类器411的处理,可以(例如,需要)预先执行处理ASR模型430的过程。因此,在通过域分类器411识别对应于用户语音的至少一个域之前,电子设备100可以将与ASR模型430相关联的数据加载到易失性存储器中,并通过ASR模型430获取对应于用户语音的文本。当(或基于)通过ASR模型430获取(或获得)对应于用户语音的文本时,电子设备100可以将所获取的文本输入到域分类器411以识别对应于用户语音的至少一个域。

[0105] 当将与ASR模型430相关联的数据加载到易失性存储器中并通过ASR模型430获取对应于用户语音的文本时,电子设备100可以并行地将与域分类器411相关联的数据加载到易失性存储器中。换句话说,电子设备100可以在完成与ASR模型相关联的数据的加载过程和ASR模型的处理过程之前,将与域分类器411相关联的数据加载到易失性存储器中,以使得当(或基于)根据ASR模型430的处理结果获得文本时,执行域分类器411的处理过程。尽管图4B示出了与ASR模型430相关联的数据和与域分类器411相关联的数据的加载过程同时启动,但是应当理解,一个或多个其它实施方式不限于此。

[0106] 同时,如图4B所示,可以以重叠的方式应用参考图4A描述的与NLU模型的加载过程相关联的各种实施方式,同时执行与ASR模型相关联的数据的加载过程和ASR模型的处理。例如,如图4B所示,在执行与ASR模型相关联的数据的加载过程和ASR模型的处理的情况下,电子设备100可以将与标准处理器422相关联的数据和与NLU引擎相关联的数据、连同与域分类器411相关联的数据并行地加载到易失性存储器中。

[0107] 图4C包括WoV模型440的加载和处理过程,以进一步描述实施方式。

[0108] 如上所述,WoV模型440指的是用于从用户语音中识别触发词的模型。WoV模型可以识别用户语音是否包括预定的触发词,并且如果(或基于)识别出用户语音包括预定的触发词,则可以确定出在接收到触发词之后对通过麦克风接收的用户语音执行语音识别。

[0109] 如果用户语音被识别为包括预定的触发词,则WoV模型440可调用如下所述的初始化器来启动将存储在非易失性存储器120中的与ASR模型430和NLU模型410、420相关联的数据加载到易失性存储器130中。为此,可以在加载ASR模型430和NLU模型之前,将WoV模型440加载到易失性存储器中。

[0110] 特别地,当电子设备100被开启并因此用于驱动电子设备的操作系统(OS)被加载时,电子设备100可以将与WoV模型440相关联的数据加载到易失性存储器中。也就是说,根据实施方式的电子设备100可以从ASR模型430和WoV模型440中单独地分类具有较少使用率的存储和资源的WoV模型440,并且在接收到触发输入之前将WoV模型440预加载到后台。然而,应理解一个或多个其它实施方式不限于此,且WoV模型440可响应(或基于)接收到触发输入而加载到易失性存储器中。

[0111] 在图4A到图4C中,已经示出了并行处理加载与多个模块相关联的数据的过程和根据多个模块中的每个的处理的过程中的多达三个过程。然而,应当理解,这仅是一个示例,并且一个或多个其它实施方式不限于此,根据非易失性存储器与易失性存储器之间的带宽,可以并行处理更多的过程。

[0112] 图5是示出根据实施方式的电子设备100的详细配置的图。

[0113] 如上所述,非易失性存储器120可以存储虚拟助理模型数据,并且虚拟助理模型数据可以包括根据多个域分类的数据和通用于多个域的数据。

[0114] 处理器140可将存储在非易失性存储器120中的虚拟助理模型数据加载到易失性存储器130中,并访问加载到易失性存储器130中的虚拟助理模型数据。具体地,处理器140可以将非易失性存储器120中存储的数据中的、与用于执行语音识别和自然语言理解的多个模型相关联的数据加载到易失性存储器130中,初始化加载到易失性存储器130中的多个模型,并且通过多个模型执行语音识别和自然语言理解。

[0115] 处理器140可将与存储在非易失性存储器120中的流控制器相关联的数据加载到易失性存储器130中,并通过流控制器控制总体操作。也即,流控制器指的是控制与数据加载过程和数据处理过程相关联的总体操作的模块,并且可以称为流控制模块等。

[0116] 流控制器可以确定虚拟助理模型数据的加载顺序和加载目标数据。流控制器可以包括初始化器,并且可以通过初始化器将存储在非易失性存储器120中的数据加载到易失性存储器130中。也就是说,初始化器可以指将存储在非易失性存储器120中的数据加载到易失性存储器130中的模块,并且也可以称为加载器、初始化模块等。

[0117] 流控制器可以调用多个模块来执行多个模块中的每个的操作、从多个模块接收对应于调用的响应、并且通过基于接收的响应再次调用多个模块来控制各种操作。

[0118] 这里,加载是指这样的操作,其中存储在非易失性存储器120中的数据被获取并存储在易失性存储器130中,以便处理器140可访问。具体地,处理器140可以通过流控制器基于存储在非易失性存储器120中的索引信息和/或寻址信息,来识别非易失性存储器120中存储的虚拟助理模型数据中的、待被加载的目标数据。处理器140可以通过读取被识别为通过初始化器被加载的数据并将所读取的数据写入易失性存储器130中,将存储在非易失性存储器120中的数据加载到易失性存储器130中。

[0119] 易失性存储器130可以实现为例如动态随机存取存储器(DRAM)。DRAM包括多个单元,并且多个单元中的每个可以包括用作开关的晶体管和存储位数据的电容器。处理器140

可以控制包括在DRAM的多个单元中的每个单元中的晶体管,从而通过对与每个晶体管对应的电容器中的电子进行充电或放电来执行写入或读取操作。

[0120] 下面进一步描述根据实施方式的通过处理器140经由流控制器调用多个模块来控制电子设备100的操作的过程。

[0121] 处理器140可以接收触发输入以执行语音识别。当(或基于)接收到触发输入时,处理器140可以确定是否通过WoV模型执行语音识别。

[0122] 根据实施方式,处理器140可以从ASR模型和NLU模型中单独地分类具有低使用率的存储器和资源的WoV模型,以便在接收到触发输入之前将WoV模型加载到后台中。也就是说,如果电子设备100被开启并且因此用于驱动电子设备的操作系统(OS)被加载,则处理器140可以将与WoV模型相关联的数据加载到易失性存储器130中。

[0123] 如果接收到触发输入,则WoV模型可以确定在接收到触发输入之后,对通过麦克风110接收的用户语音执行语音识别。如果确定出在接收到触发输入之后对通过麦克风110接收的用户语音执行(或将要执行)语音识别,则WoV模型可调用如上所述的初始化器以启动将存储在非易失性存储器120中的虚拟助理模型数据加载到易失性存储器130中。

[0124] 换句话说,如果接收到触发输入,则处理器140可以启动将虚拟助理模型数据从非易失性存储器120加载到易失性存储器130中。具体地,当接收到触发输入时,电子设备100可以通过用于将存储在非易失性存储器中的虚拟助理模型数据加载到易失性存储器中的模块,来启动加载存储在非易失性存储器中的、与ASR模型和NLU模型有关的数据。此外,在加载包括在虚拟助理模型数据中的各种数据的过程中,可以应用如下所述的各种实施方式。

[0125] 处理器140可根据多个域分类的数据加载到易失性存储器130中,并且在将根据多个域分类的数据加载到易失性存储器130中时,将通用于多个域的数据的至少一部分加载到易失性存储器130中。

[0126] 处理器140可以将通用于多个域的数据的至少一些连同根据多个域分类的数据一起并行加载到易失性存储器130中,而不是在加载根据多个域分类的数据之后加载通用于多个域的数据。下面更详细地描述根据实施方式的虚拟助理模型数据的并行加载过程。

[0127] 处理器140可以通过麦克风110接收第一用户语音。如果(或基于)接收到第一用户语音,则处理器140可以通过ASR模型获取第一文本。ASR模型可以通过对第一用户语音执行语音识别,来获取对应于第一用户语音的第一文本。

[0128] 如果(或基于)获得(或获取)对应于第一用户语音的第一文本,则处理器140可以通过域分类器识别与第一用户语音关联的至少一个域。例如,如图5所示,根据多个域分类的数据可以包括第一域相关数据、第二域相关数据和第三域相关数据,并且域分类器可以将第一域和第二域识别为与第一用户语音相关联的至少一个域。

[0129] 与第一用户语音相关联的至少一个域是指多个域中非常可能与第一用户语音相关联的域。如果获取了对应于第一用户语音的第一文本,则域分类器可以基于第一文本来识别具有与第一用户语音相关联的高概率的n个域。这里,可以预先在考虑到自然语言理解的准确性、数据加载时间等的情况下确定n。

[0130] 如果识别了与第一用户语音相关联的至少一个域,则处理器140可以将根据多个域分类的数据中的、与被识别为与第一用户语音相关联的至少一个域对应的数据加载到易

失性存储器130中。处理器140可以将与数据管理器相关联的数据加载到易失性存储器130中,并且经由数据加载器,处理器140可以根据多个域分类的数据中的、与被识别为与第一用户语音相关联的至少一个域对应的数据加载到易失性存储器130中。例如,如果第一域和第二域被识别为与第一用户语音相关联的至少一个域,则处理器140可以根据多个域分类的数据中的与第一域和第二域对应的数据加载到易失性存储器130中。

[0131] 当将对应于所识别的至少一个域的数据加载到易失性存储器130中时,处理器140可将通用于多个域的数据中的至少一些加载到易失性存储器130中。例如,当将对应于根据多个域分类的数据的第一域和第二域的数据加载到易失性存储器130中时,处理器140可以将与NLU引擎相关联的数据的至少一些加载到易失性存储器130中。

[0132] 如果(或基于)根据多个域分类的数据中的与被识别为与第一用户语音相关联的域对应的数据被加载到易失性存储器中,则处理器140可以处理所加载的数据。具体而言,处理器140可以将与数据处理器相关联的数据加载到易失性存储器中,并通过数据处理器处理与根据多个域分类的数据中的与被识别为与第一用户语音相关联的域对应的数据。

[0133] 如果(或基于)与被识别为与第一用户语音相关联的域对应的数据被处理,则处理器140可以基于所处理的数据来执行对第一文本的自然语言理解。处理器140可以加载与NLU引擎相关联的数据、通过NLU引擎执行对与第一用户语音对应的第一文本的自然语言理解、并获取与第一用户语音对应的用户意图有关的信息。

[0134] 处理器140可以加载非易失性存储器120中存储的数据中的、与标准处理器相关联的数据,并且通过标准处理器将包括在第一用户语音中的非结构化信息转换为结构化信息。

[0135] 上面已经描述了处理器140的用于执行语音识别和自然语言理解的多个模型以及用于通过多个模型执行各种操作的过程,但是应当理解,多个模型以及多个模型中的每个的名称和功能可以在实现本公开的目的的范围内根据设计而改变。多个模型和多个模块中的至少两个模块可以以端到端的方式实现为组合,在任何情况下,一个或多个实施方式可以在实现本公开的目的的范围内应用。

[0136] 此外,上面已经描述了用于对接收的用户语音进行的语音识别和自然语言理解过程,但是应当理解,根据本公开可以基于语音识别和自然语言理解的结果来执行用于获取与用户语音对应的响应的自然语言生成过程,并且还可以执行用于将所获取的响应转换为输出语音形式的文本到语音(TTS)过程。

[0137] 根据如上所述的各种实施方式,可以并行加载虚拟助理模型数据,即,用于执行语音识别和自然语言理解的数据,并且可以有效地处理(或执行)语音识别和自然语言理解的整个过程。

[0138] 通过仅将与被识别为与接收的用户语音相关联的域对应的数据加载到易失性存储器130中,而不将根据多个域分类的所有数据加载到易失性存储器130中,可以缩短数据加载时间。由于在执行自然语言理解的过程中减少了参考数据的量,所以也可缩短自然语言理解过程的时间。

[0139] 通过预先将WoV模型加载到易失性存储器130中,如果电子设备被开启,则可进一步提高电子设备100处理用户语音的速度。

[0140] 参考图6A至图6B和图7,描述了接收第一用户语音,然后接收第二用户语音的实施

方式。在下文中,可省略参照图1至图3、图4A至图4C和图5对以上提供的那些内容的冗余描述。

[0141] 图6A和图6B是示出与依次接收触发输入、第一用户语音和第二用户语音的示例相关联的实施方式的图。

[0142] 具体而言,图6A是示出电子设备100中接收触发输入、第一用户语音和第二用户语音的特定示例的图,图6B是示出由电子设备100处理触发输入、第一用户语音和第二用户语音的过程的流程图。

[0143] 参照图6A,在操作S610中,电子设备100可以通过麦克风接收触发输入。如果(或基于)接收到触发输入,则在操作S620中,电子设备100可以确定在接收到触发输入之后对通过麦克风接收到的用户语音执行语音识别。

[0144] 例如,如图6A所示,电子设备100可以经由麦克风接收包括预定触发词(例如“你好,ABC”)的触发输入。电子设备100可以确定在接收到触发输入之后,对经由麦克风接收的第一用户语音和第二用户语音执行语音识别。

[0145] 在接收到触发输入(或触发输入的接收)之后,在操作S630中,电子设备100可以经由麦克风接收第一用户语音。如果(或基于)接收到第一用户语音,则在操作S640中,电子设备100可以获取与第一用户语音对应的第一文本,并基于所获取的第一文本识别与第一用户语音关联的至少一个域。如果(或基于)与第一用户语音关联的至少一个域被识别,则在操作S650中,电子设备100可以将根据多个域分类的数据中的与被识别为与第一用户语音关联的至少一个域对应的数据加载到易失性存储器中。

[0146] 例如,电子设备100可以接收“请向我展示专利会议”作为第一用户语音。如果接收到第一用户语音,则电子设备100可获取对应于第一用户语音的第一文本,并基于所获取的第一文本将域“提醒”、“日程”和“消息”识别为与第一用户语音相关联的至少一个域。如果(或基于)域“提醒”、“日程”和“消息”被识别,则电子设备100可以选择性地将根据多个域分类的数据中的、与域“提醒”、“日程”和“消息”对应的数据加载到易失性存储器中。

[0147] 在接收到第一用户语音之后,在操作S660中,电子设备100可以经由麦克风接收第二用户语音。第二用户语音可以是在电子设备100对第一用户语音的响应输出之后由用户发声生成的用户语音。例如,第二用户语音可以是在电子设备100对第一用户语音“请向我展示专利会议”的响应“日程的标题是什么”已输出之后,由用户发声生成的用户语音。

[0148] 如果(或基于)接收到第二用户语音,则电子设备100可以获取对应于第二用户语音的第二文本,并且在操作S670中,基于所获取的第二文本以与接收第一用户语音的示例相同的方式,识别与第二用户语音相关联的至少一个域,并且在操作S680中,加载根据多个域分类的数据中的、与被识别为与第二用户语音相关联的至少一个域对应的数据。

[0149] 例如,如图6A所示,电子设备100可以接收第二用户输入“请在警报中找到专利会议”。如果接收到第二用户语音,则电子设备100可以获取对应于第二用户语音的第二文本。电子设备100可以基于所获取的第二文本将域“提醒”、“日程”和“警报”识别为与第二用户语音相关联的至少一个域。

[0150] 在该示例中,被识别为与第二用户语音相关联的至少一个域中的域“提醒”和“日程”包括在被识别为与第一用户语音相关联的至少一个域中。因此,如果根据多个域,对应于域“提醒”和“日程”的数据保持在易失性存储器中,则电子设备100不必另外加载根据多

个域分类的数据中对应于域“提醒”和“日程”的数据。因此,仅选择性地加载根据多个域分类的数据中对应于域“警报”的数据是有效的。

[0151] 根据实施方式,电子设备100可以选择性地仅加载根据多个域分类的数据中的与第二用户语音相关联的至少一个域中的、与没有包括在与第一用户语音相关联的至少一个域中的域对应的数据。将进一步参考图7和图8来描述参考图6A和图6B的实施方式。

[0152] 图7是示出当(或基于)接收到第一用户语音并且随后顺序地接收到第二用户语音时,通过比较与第一用户语音相关联的至少一个域和与第二用户语音相关联的至少一个域来确定是否加载根据多个域分类的数据的过程的流程图。

[0153] 如上所述,电子设备100可以接收第一用户语音,并识别与所接收的第一用户语音相关联的至少一个域。当(或基于)识别出与第一用户语音相关联的至少一个域时,电子设备100可以将根据多个域分类的数据中的、与第一用户语音相关联的至少一个域相对应的数据加载到易失性存储器中。

[0154] 如果根据多个域分类的数据中的与第一用户语音相关联的至少一个域对应的数据被加载到易失性存储器中,则电子设备100可以将加载的与对应于第一用户语音相关联的至少一个域的数据在易失性存储器中保持预定的时长。在下文中,描述了在易失性存储器中保持与第一用户语音相关联的至少一个域对应的数据的示例。下面参考图8描述用于在易失性存储器中保持数据的方法和用于确定用于保持数据的预定周期的方法。

[0155] 如图7所示,在操作S710中,电子设备100可以接收第二用户语音。在接收到第一用户语音之后,电子设备100可以经由麦克风接收第二用户语音。如果(或基于)接收到第二用户语音,则电子设备100可以获取或获得对应于第二用户语音的第二文本。如果(或基于)获得对应于第二用户语音的第二文本,则在操作S720中,电子设备100可以基于所获取的第二文本,来识别与第二用户语音相关联的至少一个域。

[0156] 如果(或基于)识别出与第二用户语音相关联的至少一个域,则在操作S730中,电子设备100可以将与第一用户语音相关联的至少一个域和与第二用户语音相关联的至少一个域进行比较。具体地,电子设备100可以将与第一用户语音相关联的至少一个域和与第二用户语音相关联的至少一个域进行比较,以识别与第二用户语音相关联的至少一个域是否被包括在与第一用户语音相关联的至少一个域中。

[0157] 作为(或基于)与第一用户语音相关联的至少一个域和与第二用户语音相关联的至少一个域之间的比较的结果,在操作S740中,电子设备100可以确定是否加载根据多个域分类的数据。

[0158] 在操作S740-N中,如果(或基于)确定不加载根据多个域分类的数据,则电子设备100可以不加载根据多个域分类的数据的、与第二用户语音相关联的至少一个域相对应的数据。具体地,如果与第二用户语音相关联的至少一个域被识别为包括在与第一用户语音相关联的至少一个域中,则电子设备100可以不进一步将根据多个域分类的数据加载到易失性存储器中。

[0159] 如果在操作S470-Y中确定加载根据多个域分类的数据,则在操作S750中,电子设备100可以加载根据多个域分类的数据的、与第二用户语音相关联的至少一个域对应的数据。具体而言,如果识别出与第二用户语音相关联的至少一个域中存在没有包括在与第一用户语音相关联的至少一个域中的域,则电子设备100可以加载根据多个域分类的数据中

的、与关联于第二用户语音的至少一个数据并且不同于在易失性存储器中保持的数据对应的数据。

[0160] 电子设备100可以将与第二用户语音相关联的至少一个域中、没有被包括在与第一用户语音相关联的至少一个域中的域识别为新域,并且选择性地将根据多个域分类的数据中仅对应于所识别的新域的数据加载到易失性存储器中。

[0161] 图8是示出根据实施方式的电子设备100基于接收的第一用户语音以及随后顺序接收的第二用户语音来处理第二用户语音的过程的图。

[0162] 在参考图8描述处理第二用户语音的过程时,假设先前执行了参考图5描述的过程。

[0163] 处理器140可以接收第二用户语音。当(或基于)接收到第二用户语音时,处理器140可以通过域分类器识别与第二用户语音相关联的至少一个域。例如,域分类器可以将第二域和第三域识别为与第二用户语音相关联的至少一个域。

[0164] 如以上参考图7所描述的,处理器140可以将加载到易失性存储器130中的、与第一用户语音相关联的至少一个域对应的数据在易失性存储器130中保持预定时长。预定时长的开始点和结束点(或长度)可以以多种方式确定,并且根据各种实施方式也可以由电子设备100的制造商或用户来改变。

[0165] 例如,预定时长的开始点可以确定为第一用户语音的接收完成时的时间或电子设备100对第一用户语音的响应被输出时的时间。此外,如果在从预定时长的开始时间起经过预定时间的时间内没有接收到第二用户语音,则可以将经过预定时长的时间点设置为预定时长的结束点。如果(或基于)在从预定时间点起的预定时长的时间内接收到第二用户语音,则可以通过第二用户语音的接收完成时的时间或电子设备100相对于第二用户语音的响应输出时的时间再次估计预定时长。

[0166] 如上所述,易失性存储器130可以实现为DRAM。在DRAM的情况下,由于通过对电容器中的电子充电来存储数据,所以根据电容器的特征,在经过一定时长之后,随着充电的电子被放电,所存储的数据丢失。因此,为了保持存储在DRAM中的数据,应用刷新技术来周期性地重写数据。也就是说,处理器140可以通过周期性地刷新存储在易失性存储器130中的数据来将存储的数据保持预定的时长。

[0167] 虽然已主要描述了保持对应于与第一用户语音相关联的至少一个域的数据,但是应当理解,根据各种实施方式,可以在易失性存储器130中保持加载到易失性存储器130中用于处理第一用户语音的数据中的、用于(例如,需要用于)对第二用户语音进行处理的数据中的至少一些数据。

[0168] 在下文中描述在易失性存储器中保持与第一用户语音相关联的至少一个域相对应的数据的情况。也即,如图8所示,描述在易失性存储器130中保持根据多个域分类的数据中的第一域相关数据和第二域相关数据的示例。

[0169] 如果(或基于)如上所述识别出与第二用户语音相关联的至少一个域,则处理器140可以通过流控制器将与第一用户语音相关联的至少一个域和与第二用户语音相关联的至少一个域进行比较。具体地,处理器140可以将与第一用户语音相关联的至少一个域和与第二用户语音相关联的至少一个域进行比较,以识别与第二用户语音相关联的至少一个域是否被包括在与第一用户语音相关联的至少一个域中。如果对应于第一用户语音的域和对

应于第二用户语音的域在数量上是多个,则处理器140可以逐个域地顺序进行比较或者同时比较整个域。

[0170] 作为与第一用户语音相关联的至少一个域和与第二用户语音相关联的至少一个域之间的比较结果,处理器140可以确定是否通过流控制器加载根据多个域分类的数据。

[0171] 如果(或基于)确定不加载根据多个域的数据,则处理器140可以不加载根据多个域分类的数据的与第二用户语音相关联的至少一个域对应的数据。具体地,如果与第二用户语音相关联的至少一个域被识别为包括在与第一用户语音相关联的至少一个域中,则处理器140可以不进一步将根据多个域分类的数据加载到易失性存储器中。

[0172] 如果(或基于)确定要加载根据多个域分类的加载数据,则处理器140可加载根据多个域的分类数据的、与第二用户语音相关联的至少一个域对应的数据。具体地,如果识别出存在与第二用户语音相关联的至少一个域中的、没有包括在与第一用户语音相关联的至少一个域中的域,则处理器140可以将根据多个域分类的数据中的、对应于与第二用户语音相关联的至少一个域的并且不同于在易失性存储器中保持的数据加载到易失性存储器中。

[0173] 处理器140可以将与第二用户语音相关联的至少一个域中的、与第一用户语音相关联的至少一个域中的域识别为新域,并且选择性地将根据多个域分类的数据中的与所识别的新域相对应的数据加载到易失性存储器中。

[0174] 例如,如图7所示,处理器140可以将第三域识别为与第二用户语音相关联的至少一个域中的、没有包括在与第一用户语音相关联的至少一个域中的新域。处理器140可以选择性地仅将根据多个域分类的数据中的与被识别为新域的第三域相对应的数据加载到易失性存储器130中。可以由数据加载器执行根据多个域分类的数据的加载,并且可以由流控制器调用数据加载器,如参考图5所描述的。

[0175] 当(或基于)根据上述处理将根据多个域分类的数据中与第二用户语音相关联的至少一个域对应的所有数据加载到易失性存储器130中时,NLU引擎可以通过基于根据多个域分类的数据中与第二域对应的数据以及与第三域对应的数据进行自然语言理解,来获取与第二用户语音对应的用户意图有关的信息。如图4所示,在NLU引擎执行自然语言理解之前,可以通过数据处理器执行数据处理。

[0176] 如果处理器140加载对应于与第二用户语音相关联的至少一个域的数据,则处理器140可以控制不将与已经加载到易失性存储器中的数据中的与关联于第二用户语音的域不对应的数据加载到易失性存储器中。这里,不将数据加载到存储器意味着已写入数据的存储区被写入0(或空数据),或者没有新数据被写入相应的存储区。

[0177] 尽管已经描述了,如果根据多个域分类的数据被加载到易失性存储器130中,则处理加载的数据并且使用处理的数据执行自然语言理解,但是应当理解,一个或多个其它实施方式不限于此。例如,根据另一实施方式,处理器140可将加载到易失性存储器130中并由数据处理器处理的数据存储在非易失性存储器120中,从而减少用于(例如需要用于)稍后处理数据的时间。

[0178] 上面已经描述了在已接收到第一用户语音之后处理接收的第二用户语音的过程,但是应当理解,一个或多个其它实施方式不限于此。例如,根据另一实施方式,在接收到第二用户语音之后,可以以相同的方式将处理第二用户语音的过程应用于接收的第三用户语音。

[0179] 根据如上所述的各种实施方式,如果顺序地接收第一用户语音和第二用户语音,则在用于执行第二用户语音的自然语言理解的根据多个域分类的数据的加载过程中,可以通过利用已加载和保持在易失性存储器130中的数据来改进用于第二用户语音的数据加载时间和处理速度。

[0180] 图9是示出根据实施方式的电子设备100的详细配置的框图。

[0181] 如图1所示,电子设备100包括麦克风110、非易失性存储器120、易失性存储器130和处理器140,并且还可包括通信器150、输出器160和输入器170。然而,所示出的配置仅是示例性的,并且可添加额外组件或配置,或可根据各种实施方式省略一个或多个组件或配置。

[0182] 通信器150包括电路并且可以与服务器或外部设备通信。具体地,处理器140可以从通过通信器150连接的服务器或外部设备接收各种数据或信息,并且可以向服务器或外部设备发送各种数据或信息。此外,通信器150可以包括Wi-Fi模块151、蓝牙模块152、无线通信模块153和近场通信(NFC)模块154中的至少一个。

[0183] 处理器140可以经由通信器150从服务器接收如上所述的与语音识别和自然语言理解相关的数据,并且更新存储在非易失性存储器120中的数据。

[0184] 处理器140可以通过通信器150将根据本公开的语音识别和自然语言理解过程中的至少一些的结果发送到服务器,并从服务器接收用于一些其它过程的结果。也就是说,一个或多个实施方式可以以电子设备100中的设备上形式来实现,并且根据本公开的过程中的至少一部分可以经由至少一个服务器来实现。

[0185] 根据实施方式的触发输入可以通过通信器150从用于控制电子设备100的远程控制设备接收。具体地,如果从远程控制设备生成用于执行语音识别的控制信号,则处理器140可以通过通信器150接收控制信号。

[0186] 输出器160可以包括电路,并且处理器140可以通过输出器160输出电子设备100可以执行的各种功能。输出器160可以包括显示器161、扬声器162和指示器163中的至少一个。

[0187] 显示器161可以在处理器140的控制下输出图像数据。根据实施方式的显示器161可以通过处理器140的控制来显示用户界面(UI)。显示器161可以实现为液晶显示器(LCD)面板、有机发光二极管(OLED)显示器等,并且显示器161也可以实现为柔性显示器、透明显示器等。然而,显示器161不限于特定类型。扬声器162可以在处理器140的控制下输出音频数据,并且指示器163可以在处理器140的控制下打开或关闭。

[0188] 具体地,在各种实施方式中,输出器160可以在处理器140的控制下,输出根据接收到的触发输入的处理结果、对应于接收的用户语音的响应、用于电子设备100的用户的引导消息等。显示器161可以在处理器140的控制下,显示表示数据加载进度或者语音识别和自然语言理解的处理进度的进度条,并且还可以显示指示根据多个域分类的数据是否被保持在易失性存储器130中的信息等。

[0189] 输入器170包括电路,并且处理器140可以经由输入器170接收用于控制电子设备100的操作的用户命令。具体地,输入器170可以包括麦克风171、输入按钮172、相机等,并且输入器170可以以包括在触摸显示器中的形式来实现。

[0190] 输入器170可以接收触发输入以执行语音识别。虽然上面已描述了通过麦克风171以包括预设触发词的用户语音的形式接收触发输入,但是应当理解,一个或多个其它实施

方式不限于此。例如,根据另一实施方式,可以通过电子设备100上的输入按钮172接收触发输入。输入按钮172可以是物理地设置在电子设备100的外部上的物理按钮,或者是以显示在触摸显示器上的用户界面(UI)的形式实现的软按钮。

[0191] 根据如上所述的各种实施方式,可以并行加载虚拟助理模型数据,即用于执行语音识别和自然语言理解的数据,从而有效地处理语音识别和自然语言理解的整个过程。

[0192] 可以通过选择性地仅加载与被识别为与接收的用户语音相关的域对应的数据,而不将根据多个域分类的所有数据加载到易失性存储器130中,来缩短数据加载时间。

[0193] 此外,由于在根据多个域分类的数据中,只有与被识别为与接收的用户语音相关联的域对应的数据被选择性地加载到易失性存储器130中,所以对加载的数据的处理时间也可缩短,并且在执行自然语言理解的过程中参考的数据量可减少,从而也缩短了自然语言理解过程的时间。

[0194] 如果顺序地接收到第一用户语音和第二用户语音,则在用于执行第二用户语音的自然语言理解的根据多个域分类的数据的加载过程中,可以通过利用已加载和保持在易失性存储器130中的数据来改进数据的加载时间和第二用户语音的处理速度。

[0195] 当电子设备开启时,WoV模型被预先加载到易失性存储器130中,从而进一步提高了电子设备100处理与用户语音相关联的速度。

[0196] 在现有技术中,以例如域分类器、规则引擎、标准处理器、DNN引擎和其他各种模块、数据加载器和数据处理器的顺序依次执行相关数据的加载和初始化每个模块的过程。同时,与现有技术相比,根据如上所述的各种实施方式的选择性加载过程和并行加载过程提供了具有显著改进的处理速度的AI辅助功能。

[0197] 如上所述,电子设备100的控制方法、处理器的控制过程及其各种实施方式可以实现为程序并提供给电子设备100。特别地,包括电子设备100的控制方法的程序可以存储并提供在非暂时性计算机可读介质中。

[0198] 非暂时性计算机可读介质是指半永久地存储数据而不是在非常短的时间内存储数据的介质,例如寄存器、高速缓存、存储器等,并且可由装置读取。详细地,上述各种应用、指令或程序可以存储并且可以提供在非暂时性计算机可读介质中,例如,光盘(CD)、数字多功能盘(DVD)、硬盘、蓝光盘、通用串行总线(USB)、存储卡、只读存储器(ROM)等。

[0199] 多个模块中的至少一个模块可以通过AI模型来实现。可以通过非易失性存储器120、易失性存储器130和处理器140来执行与AI相关联的功能。

[0200] 处理器140可以包括一个或多个处理器140。此时,一个或多个处理器140可以是例如中央处理单元(CPU)、应用处理器(AP)等的通用处理器;例如图形处理单元(GPU)、可视处理单元(VPU)的图形处理单元;和/或例如神经处理单元(NPU)的AI专用处理器。

[0201] 一个或多个处理器140根据存储在非易失性存储器120和易失性存储器130中的预定操作规则或人工智能(AI)模型,来控制输入数据的处理。通过训练或学习来提供预定的操作规则或人工智能模型。

[0202] 这里,提供通过对多个学习数据应用学习算法的学习装置,来产生具有期望特性的预定操作规则或AI模型。可以在执行根据实施方式的AI的设备本身中执行学习,和/或可通过单独的服务器/系统来实现。

[0203] AI模型可包括多个神经网络层。每一层具有多个权重值,并通过计算前一层和多

个权重的运算来执行层运算。神经网络的示例包括但不限于卷积神经网络 (CNN)、深度神经网络 (DNN)、递归神经网络 (RNN)、受限Boltzmann机器 (RBM)、深度置信度网络 (DBN)、双向递归深度神经网络 (BRDNN)、生成对抗网络 (GAN) 和深度Q网络。

[0204] 学习算法是用于使用多个学习数据来训练预定目标设备 (例如, 机器人) 以使得、允许或控制目标设备进行确定或预测的方法。学习算法的示例包括但不限于监督学习、无监督学习、半监督学习或强化学习。

[0205] 尽管已参考各种示例性实施方式示出并描述了本公开, 但是应理解, 各种示例性实施方式旨在说明性的, 而不是限制性的。本领域普通技术人员将理解, 在不脱离包括所附权利要求及其等同的本公开的精神和范围的情况下, 可以对形式和细节进行各种改变。

100

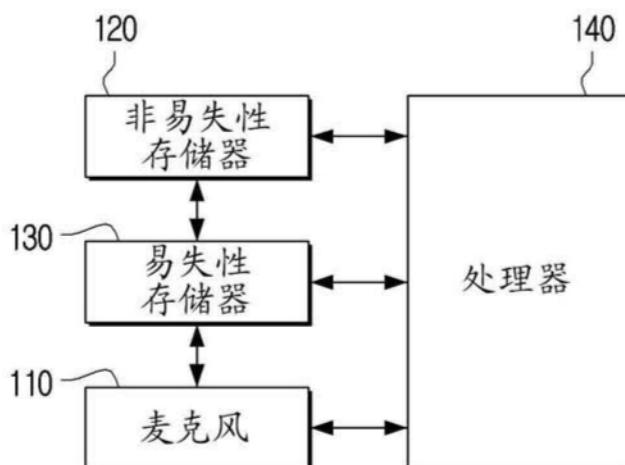


图1

根据多个域分类的数据(10)

通用于多个域的数据(20)

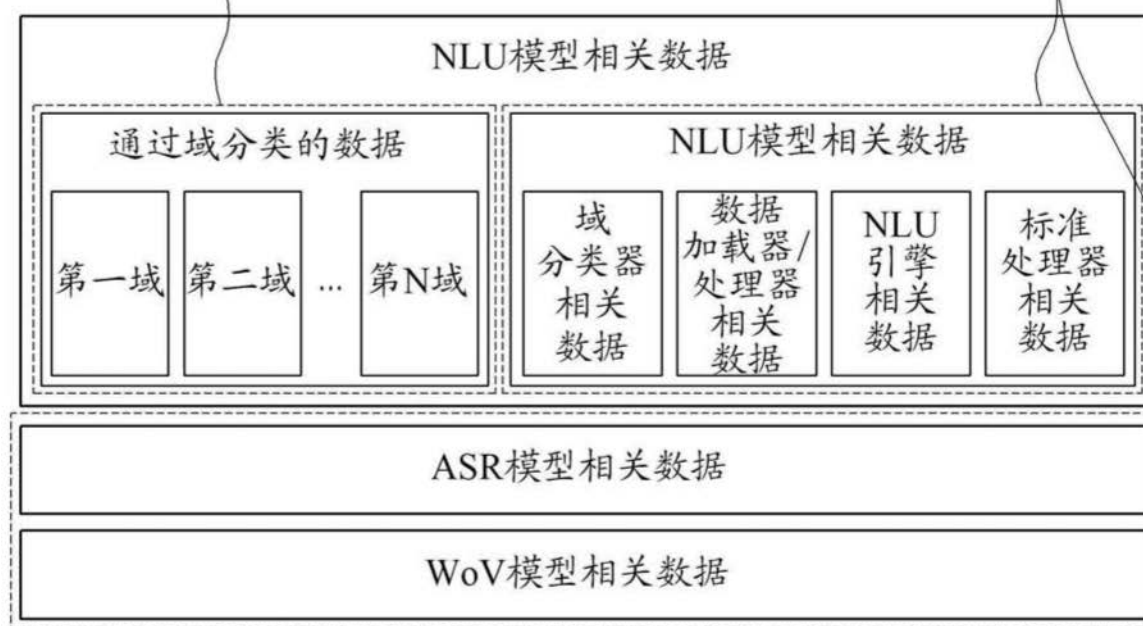


图2

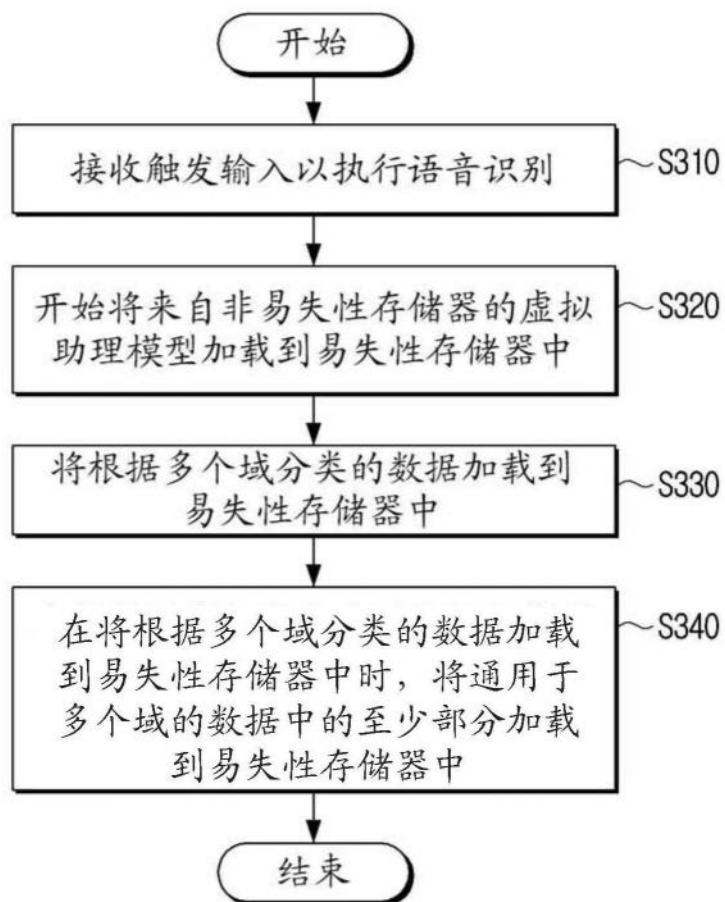


图3

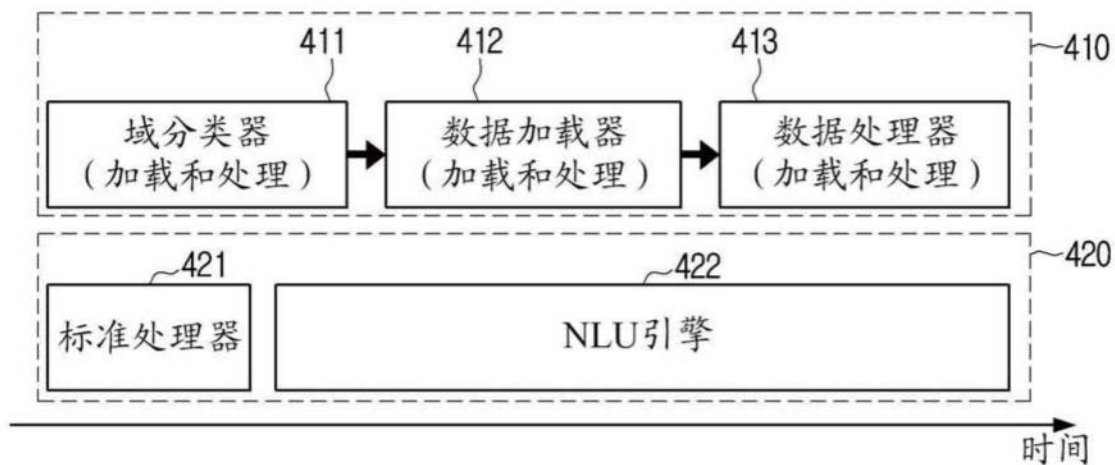


图4A

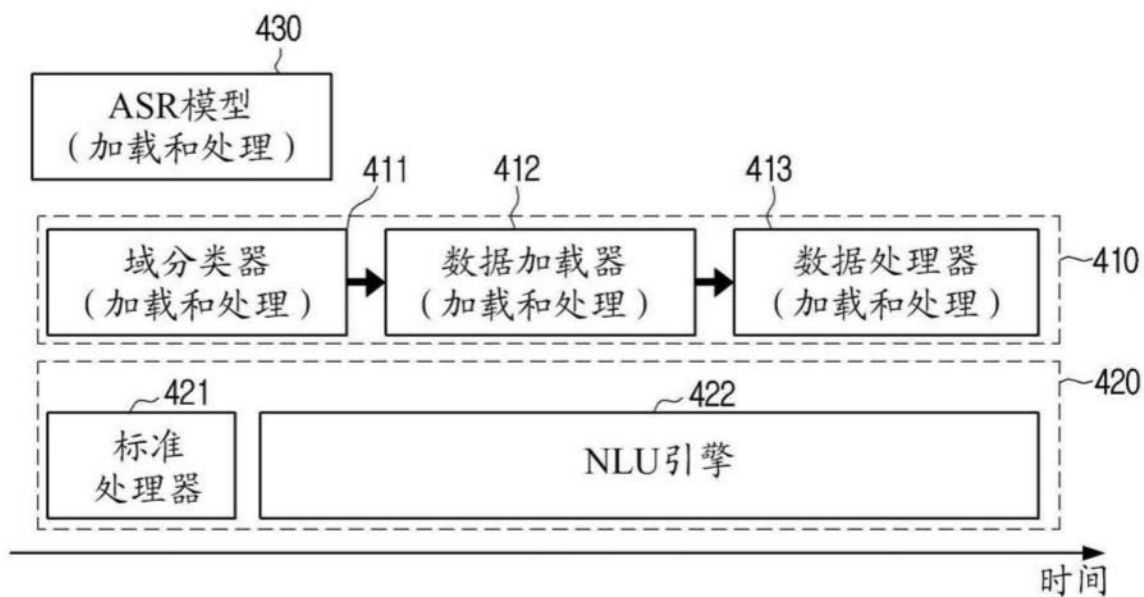


图4B

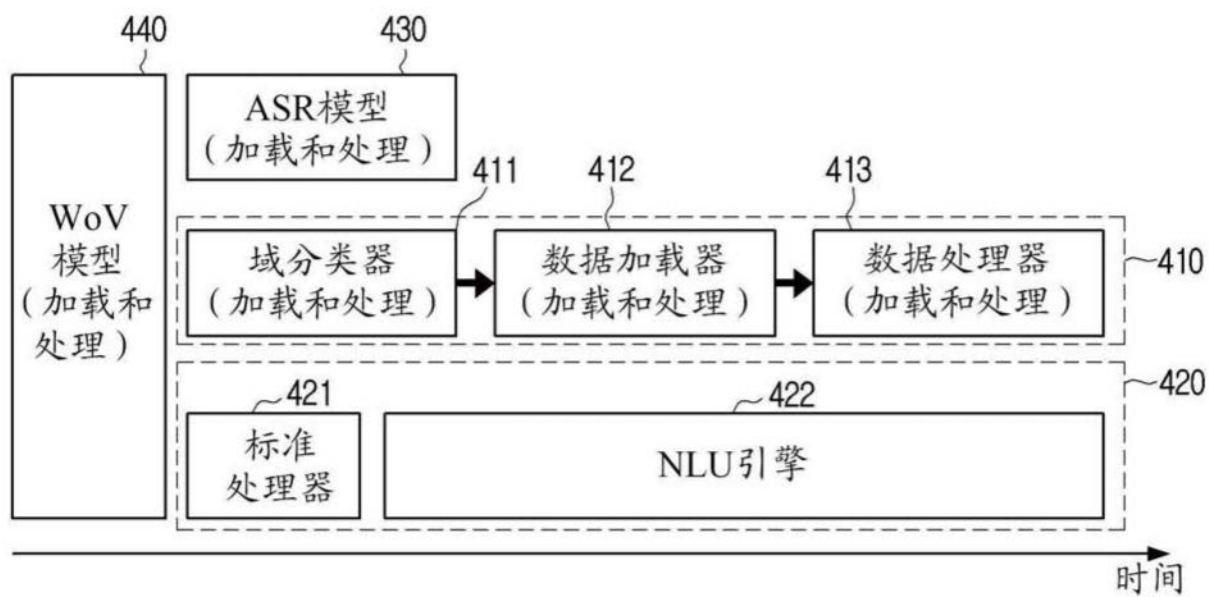


图4C

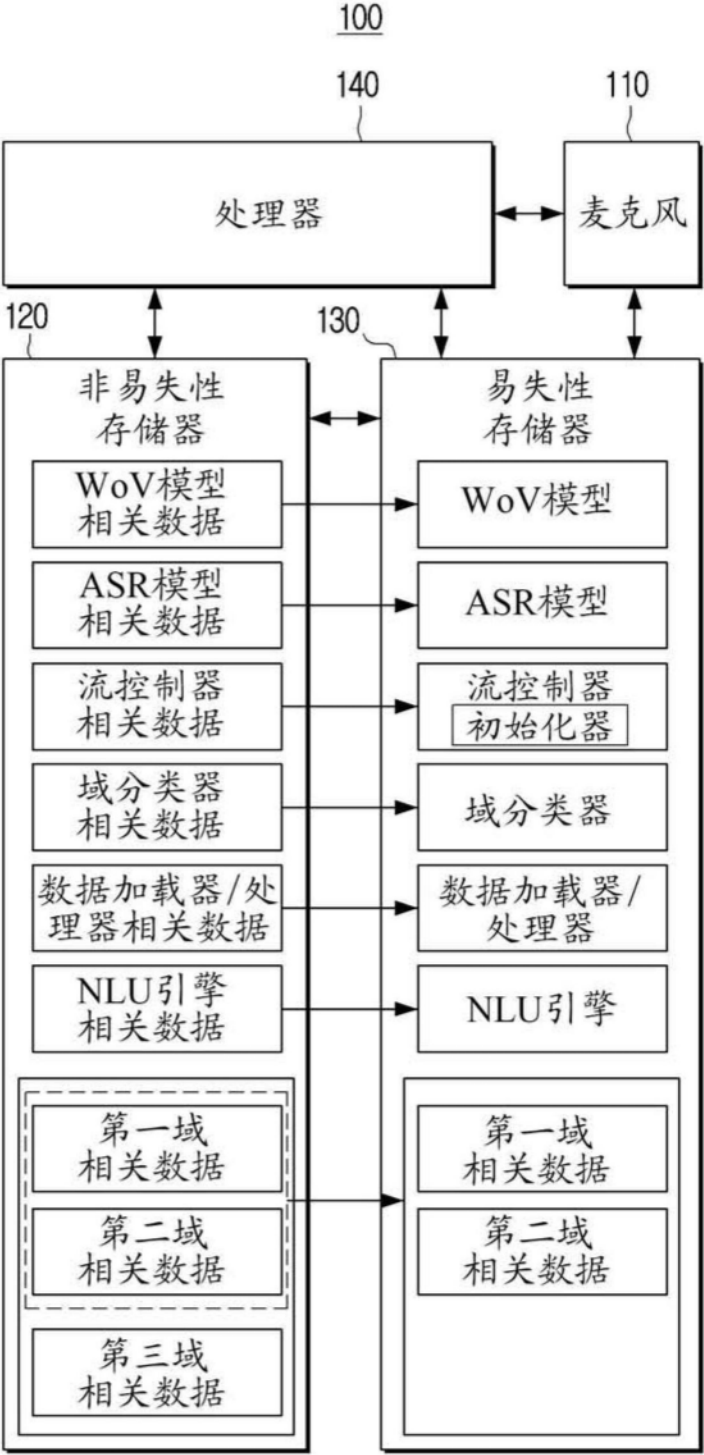


图5

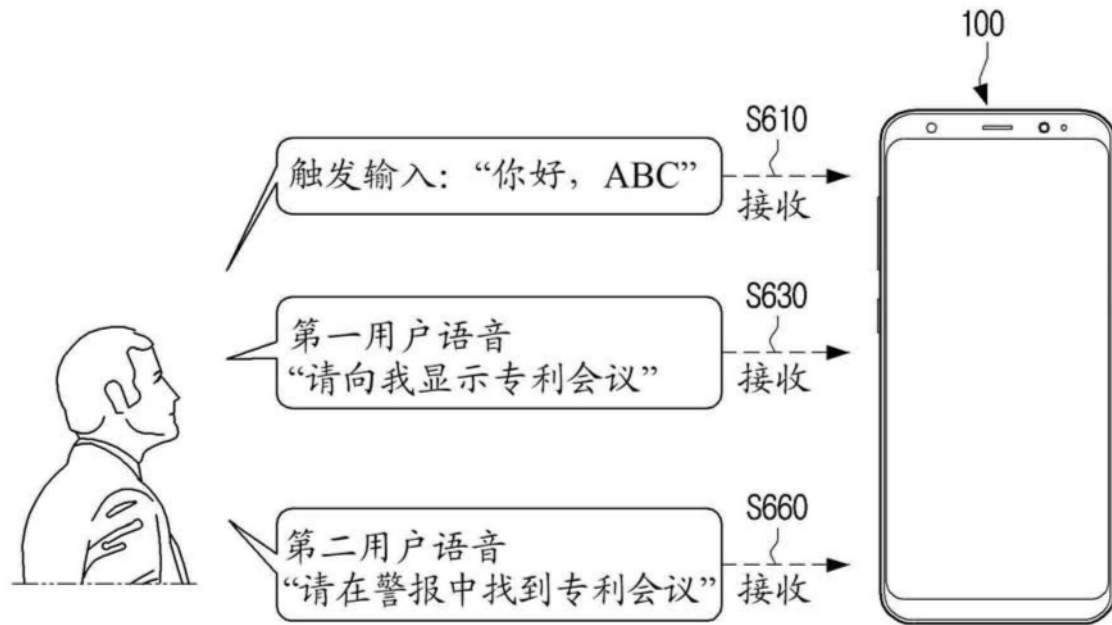


图6A

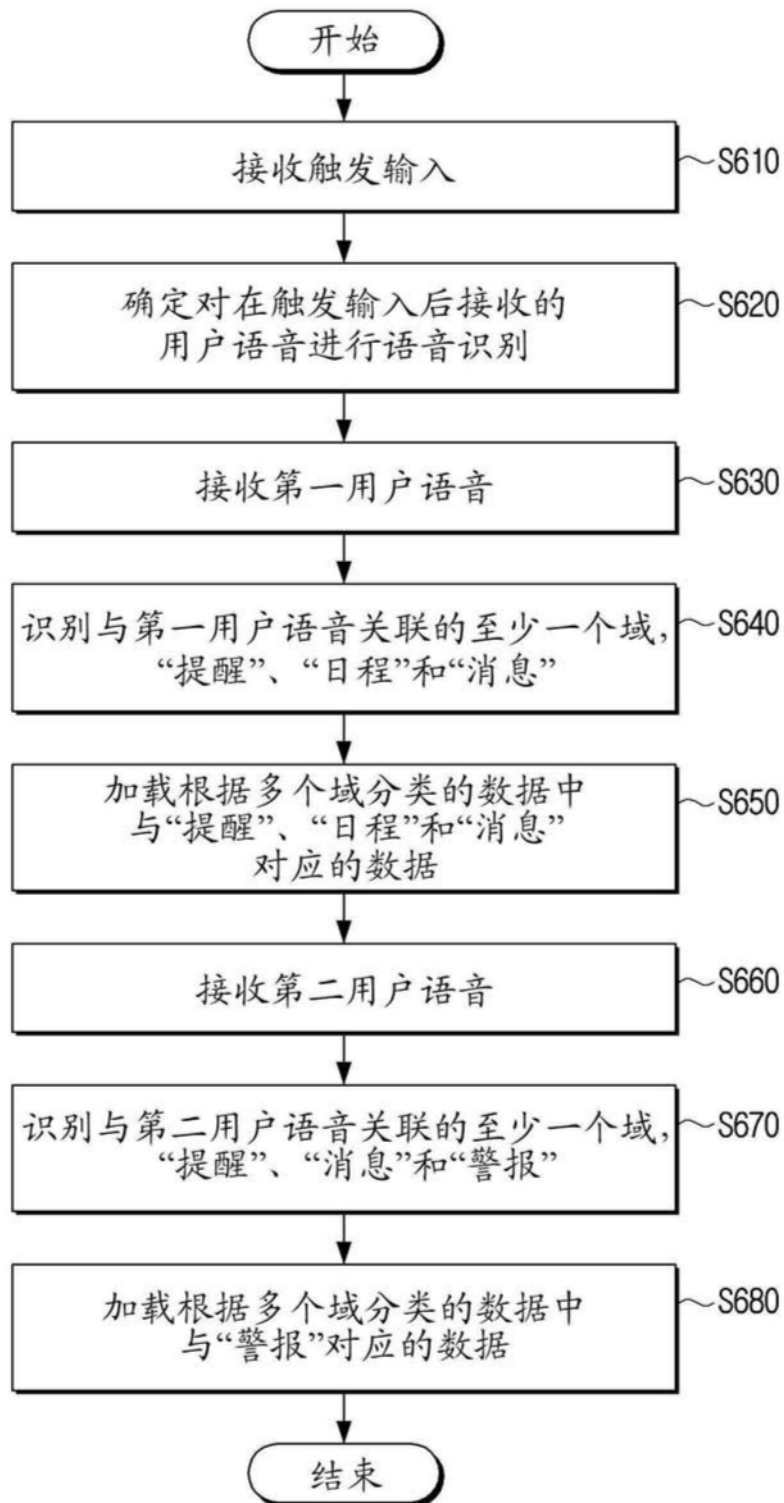


图6B

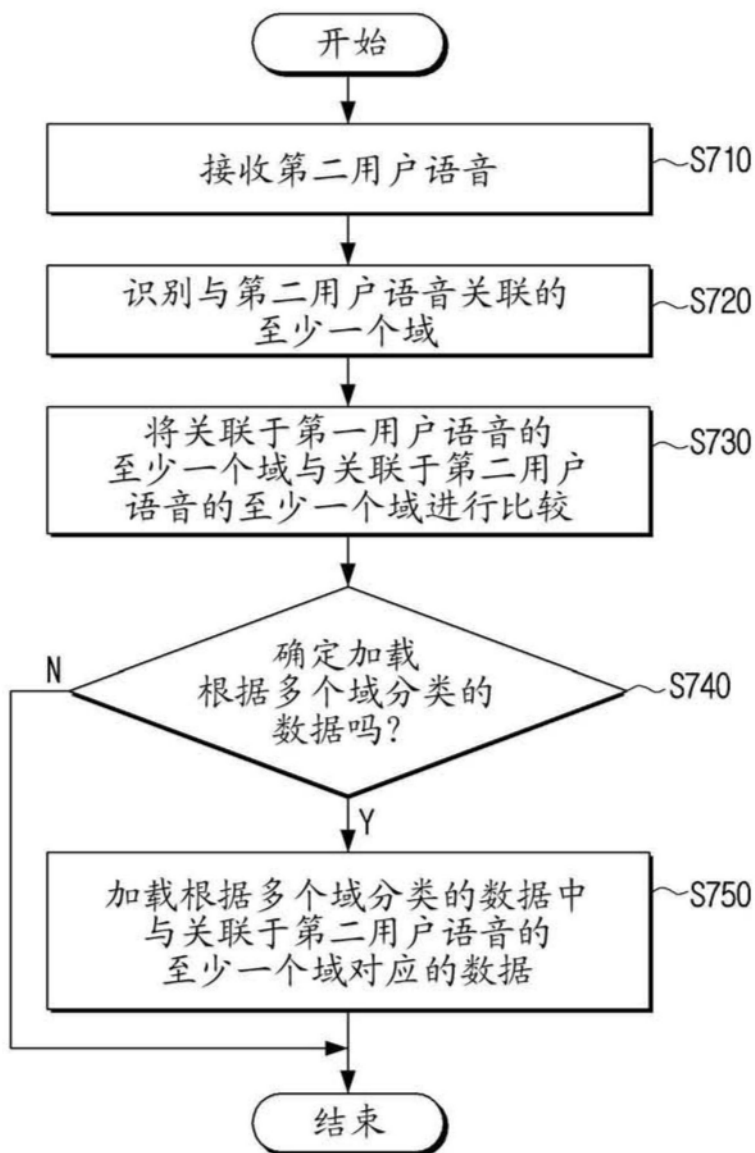


图7

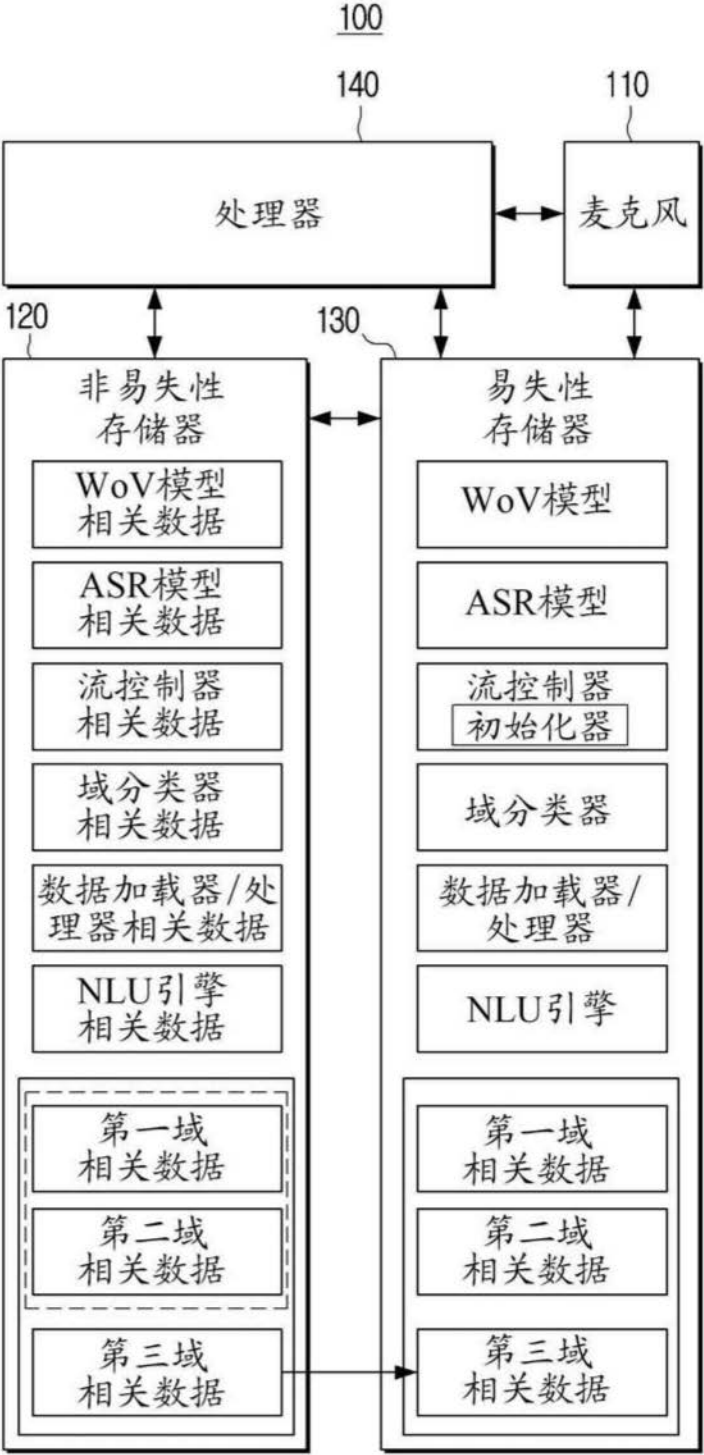


图8

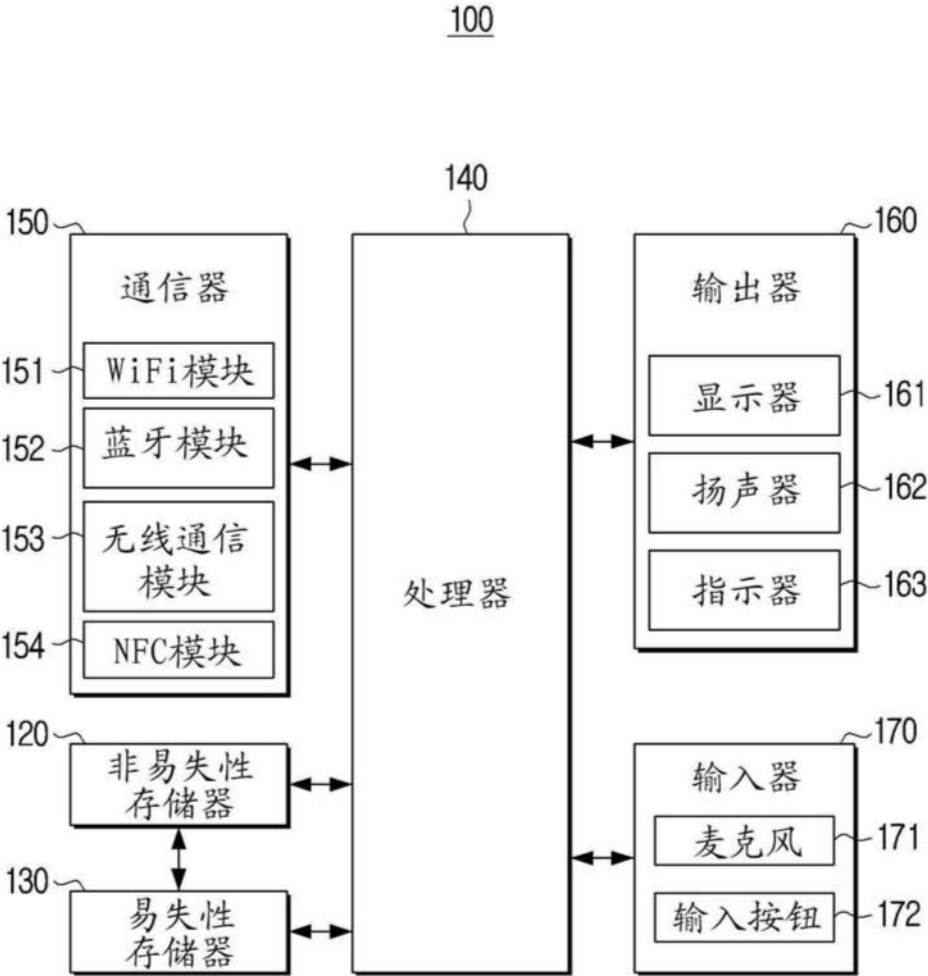


图9