



(10) 授权公告号 CN 112334872 B

(45) 授权公告日 2024.07.26

(21) 申请号 201980041799.X

(22) 申请日 2019.06.21

(65) 同一申请的已公布的文献号  
申请公布号 CN 112334872 A

(43) 申请公布日 2021.02.05

(30) 优先权数据  
10-2018-0071817 2018.06.22 KR

(85) PCT国际申请进入国家阶段日  
2020.12.21

(86) PCT国际申请的申请数据  
PCT/KR2019/007521 2019.06.21

(87) PCT国际申请的公布数据  
W02019/245331 KO 2019.12.26

(73) 专利权人 三星电子株式会社  
地址 韩国京畿道

(72) 发明人 P.特拉茨 S.莱斯基

(74) 专利代理机构 北京市柳沈律师事务所  
11105  
专利代理师 刘虹

(51) Int.Cl.  
G06F 3/16 (2006.01)  
G06F 3/00 (2006.01)  
G06F 3/023 (2006.01)  
G06F 3/0346 (2006.01)  
G06F 3/01 (2006.01)  
G10L 15/26 (2006.01)  
G10L 15/25 (2006.01)

(56) 对比文件  
US 2015317299 A1, 2015.11.05

审查员 黄旭光

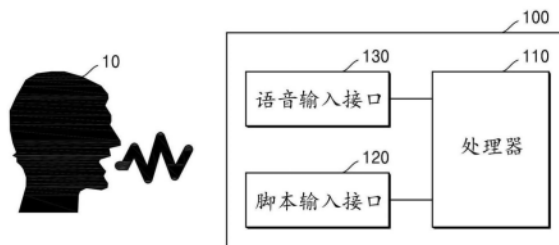
权利要求书2页 说明书20页 附图12页

(54) 发明名称

文本输入设备及其方法

(57) 摘要

提供了一种响应于用户的语音输入而输入文本的文本输入设备和相关方法。根据实施例的文本输入方法包括：提供被配置为识别脚本输入的脚本输入接口；确定是否满足指示用户的语音输入是否开始的语音激活条件；以及基于确定的结果来激活被配置为识别语音输入的语音输入接口，并且同时提供激活的语音输入接口和脚本输入接口。



1. 一种电子设备,包括:
  - 显示器;
  - 麦克风;
  - 存储器,被配置为存储至少一条指令;和
  - 至少一个处理器,其通过运行所述至少一条指令,被配置为:
    - 基于通过使用麦克风接收到的用户语音输入,激活语音输入接口;
    - 在显示器上显示用于指示语音输入接口被激活的语音输入激活图标;
    - 在显示器上显示用于显示通过语音输入接口和脚本输入接口输入的文本信息的文本显示区域,所述脚本输入接口包括分别对应于根据语言的类型而不同地显示的多个字符的多个区域;
    - 当脚本输入接口显示在显示器上、语音输入接口被激活并且语音输入激活图标显示在显示器上的时候,通过使用脚本输入接口接收脚本输入;
    - 基于脚本输入,在显示器上将脚本文本信息显示在文本显示区域中;
    - 当脚本输入接口显示在显示器上、语音输入接口被激活并且语音输入激活图标显示在显示器上的时候,通过使用语音输入接口接收语音输入;以及
    - 基于语音输入,在显示器上将语音文本信息与脚本文本信息显示在文本显示区域中,
    - 其中,以脚本输入和语音输入的输入顺序将脚本文本信息和语音文本信息显示在文本显示区域中。
2. 根据权利要求1所述的电子设备,还包括至少一个加速度计或陀螺仪,
  - 其中,处理器还被配置为基于所述至少一个加速度计或陀螺仪的检测结果来测量文本输入设备的斜率,并基于所测量的斜率来激活语音输入接口。
3. 根据权利要求1所述的电子设备,其中,基于语音输入和情境信息来提供推荐语音文本信息,并且选择所述推荐语音文本信息以显示在文本显示区域中。
4. 根据权利要求3所述的电子设备,其中,所述推荐语音文本信息被提供在脚本输入接口和文本显示区域之间。
5. 根据权利要求3所述的电子设备,其中,基于用户的使用历史信息来提供所述推荐语音文本信息。
6. 根据权利要求5所述的电子设备,其中,所述使用历史信息包括在应用中执行的功能的历史。
7. 根据权利要求5所述的电子设备,其中,所述使用历史信息包括用户的文本历史。
8. 根据权利要求5所述的电子设备,其中,所述使用历史信息包括文本信息中包括的单词的频率。
9. 根据权利要求1所述的电子设备,其中,所述至少一个处理器包括数据识别器,并且
  - 其中,数据识别器包括识别数据选择器和识别结果提供器,并且
  - 其中,识别结果提供器被配置为:
    - 将识别数据选择器选择的数据发送到服务器。
10. 根据权利要求3所述的电子设备,其中,所述推荐语音文本信息是由服务器的模型学习器或所述电子设备的模型学习器中的至少一个基于训练数据生成的。
11. 根据权利要求3所述的电子设备,其中,通过使用从服务器接收的数据识别模型来

生成所述推荐语音文本信息。

12. 根据权利要求3所述的电子设备,其中,所述推荐语音文本信息与语音文本信息相关。

13. 一种文本输入的方法,包括:

基于通过使用麦克风接收到的用户语音输入,激活语音输入接口;

在显示器上显示用于指示语音输入接口被激活的语音输入激活图标;

在显示器上显示用于显示通过语音输入接口和脚本输入接口输入的文本信息的文本显示区域,所述脚本输入接口包括分别对应于根据语言的类型而不同地显示的多个字符的多个区域;

当脚本输入接口显示在显示器上、语音输入接口被激活并且语音输入激活图标显示在显示器上的时候,通过使用脚本输入接口接收脚本输入;

基于脚本输入,在显示器上将脚本文本信息显示在文本显示区域中;

当脚本输入接口显示在显示器上、语音输入接口被激活并且语音输入激活图标显示在显示器上的时候,通过使用语音输入接口接收语音输入;以及

基于语音输入,在显示器上将语音文本信息与脚本文本信息显示在文本显示区域中,

其中,以脚本输入和语音输入的输入顺序将脚本文本信息和语音文本信息显示在文本显示区域中。

14. 一种包括指令的计算机程序,当所述程序由处理器执行时,使得所述处理器执行根据权利要求13所述的方法。

## 文本输入设备及其方法

### 技术领域

[0001] 本公开涉及响应于语音输入而输入文本的设备和方法。

### 背景技术

[0002] 随着信息技术 (IT) 的发展,已经开发了组合执行各种功能的电子设备 (诸如智能电话),并且已经发布了具有语音识别功能的电子设备以提高可操作性。语音识别功能具有优点,使得能够在不接触单独的操作按钮或触摸模块的情况下通过识别用户的语音来容易地控制设备。

[0003] 根据语音识别功能,例如,诸如智能电话的便携式终端可以执行呼叫功能,写入文本消息,并且容易地设置诸如寻路、互联网搜索或警报设置的各种功能,而无需按下单独的按钮。

### 发明内容

[0004] 技术问题

[0005] 需要一种文本输入设备和方法,其提供能够有效地识别用户语音输入的语音输入接口。

[0006] 技术方案

[0007] 根据本公开的实施例的文本输入设备包括:语音输入接口,被配置为识别语音输入;脚本输入接口,被配置为识别脚本输入;以及处理器,被配置为确定是否满足指示用户的语音输入是否开始的语音激活条件,基于确定的结果来激活语音输入接口,并且同时提供激活的语音输入接口和脚本输入接口。

[0008] 根据本公开的实施例的文本输入方法包括:提供脚本输入接口,被配置为识别脚本输入;确定是否满足指示用户的语音输入是否开始的语音激活条件;以及基于确定的结果激活被配置为识别语音输入的语音输入接口,并且同时提供激活的语音输入接口和脚本输入接口。

[0009] 根据本公开的实施例的计算机可读记录介质可以是用于执行上述方法的程序记录在其上的计算机可读记录介质。

[0010] 技术效果

[0011] 根据实施例,可以提供一种文本输入设备和方法,其提供识别用户语音输入的语音输入接口。

### 附图说明

[0012] 通过组合以下详细描述和附图,可以容易地理解本公开,并且附图标记指代结构元件。

[0013] 图1是根据实施例的文本输入设备的图。

[0014] 图2是根据实施例的文本输入设备的图。

- [0015] 图3是根据实施例的文本输入设备的图。
- [0016] 图4A是根据一些实施例的用于描述由文本输入设备执行的执行文本输入操作的方法的图。
- [0017] 图4B是根据一些实施例的用于描述由文本输入设备执行的执行文本输入操作的方法的图。
- [0018] 图5是根据一些实施例的用于描述由文本输入设备执行的、基于情境信息来执行文本输入操作的方法的图。
- [0019] 图6是根据实施例的用于描述文本输入方法的流程图。
- [0020] 图7是根据实施例的用于描述确定语音激活条件的方法的流程图。
- [0021] 图8是根据实施例的用于描述确定语音激活条件的方法的流程图。
- [0022] 图9是根据实施例的用于描述确定语音激活条件的方法的流程图。
- [0023] 图10是根据实施例的用于描述确定语音激活条件的方法的流程图。
- [0024] 图11是根据实施例的用于描述确定语音激活条件的方法的流程图。
- [0025] 图12是根据实施例的用于描述文本输入方法的流程图。
- [0026] 图13是根据一些实施例的第一设备的框图。
- [0027] 图14是根据一些实施例的第一设备的框图。
- [0028] 图15是根据一些实施例的处理器框图。
- [0029] 图16是根据一些实施例的数据训练器的框图。
- [0030] 图17是根据一些实施例的数据识别器的框图。
- [0031] 图18是示出根据一些实施例的文本输入设备和服务器彼此互操作以训练和识别数据的示例的图。

### 具体实施方式

[0032] 根据本公开的实施例的文本输入设备包括：语音输入接口，被配置为识别语音输入；脚本输入接口，被配置为识别脚本输入；以及处理器，被配置为确定是否满足指示用户的语音输入是否开始的语音激活条件，基于确定的结果来激活语音输入接口，并且同时提供激活的语音输入接口和脚本输入接口。

[0033] 在下文中，将参考附图详细描述本公开的实施例，使得本领域普通技术人员可以容易地执行本公开。然而，本公开可以以许多不同的形式来实施，并且不限于本文描述的实施例。此外，为了清楚地描述本公开，省略了与描述无关的部分，并且在整个说明书中将相同的附图标记分配给相同的元件。

[0034] 本公开的一些实施例可以由功能块配置和各种处理操作来表示。这些功能块的全部或部分可以由被配置为执行特别的功能的任何数量的硬件和/或软件元件来实现。例如，可以将本公开的功能块实现为一个或多个微处理器，或者可以实现为用于特定功能的电路配置。此外，例如，本公开的功能块可以用各种编程或脚本语言来实现。可以将功能块实现为在一个或多个处理器上执行的算法。此外，本公开可以采用用于电子环境配置、信号处理和/或数据处理的任何传统 (conventional) 技术。

[0035] 此外，附图中所示的连接线或连接构件 (members) 旨在表示元件之间的示例性功能连接、和/或物理或逻辑连接。在真实设备中，元件之间的连接可以由可替换地或添加地

各种功能连接、物理连接或电路连接来表示。

[0036] 此外,如本文所使用的术语“-器”和“模块”表示用于处理至少一个功能或操作的单元,其可以由硬件、软件或二者的组合来实现。可以用程序来实现术语“-器”和“模块”,其存储在可寻址存储介质中并且可由处理器执行。

[0037] 例如,术语“-器”和“模块”可以指元件,诸如软件元件、面向对象的软件元件、类元件和任务元件,并且可以包括进程(process)、函数、特性(attribute)、过程(procedure)、子例程、程序代码段、驱动程序、固件、微代码、电路、数据、数据库、数据结构、表、数组和变量。

[0038] 在本说明书和附图中描述的本公开的实施例仅作为具体示例而呈现,以便容易地解释本公开的技术内容并帮助理解本公开,而不旨在限制本公开的范围。即,对于本领域普通技术人员显而易见的是,基于本公开的技术构思可以进行其他修改。此外,各个实施例可以根据需要而彼此组合地操作。例如,本公开的一个实施例的一些部分和本公开的另一个实施例的一些部分可以彼此组合,使得可以操作设备。

[0039] 图1是根据实施例的文本输入设备的图。

[0040] 参照图1,文本输入设备100包括处理器110、脚本输入接口120和语音输入接口130。处理器110控制文本输入设备100的整体操作。

[0041] 用户10可以向文本输入设备100提供语音输入或脚本输入。语音输入是包括信息(例如,文本信息)的用户语音。脚本输入是包括信息(例如,文本信息)的用户动作。用户的动作可以包括与设备的接触。

[0042] 文本输入设备100可以通过使用脚本输入接口120来接收脚本输入,并且可以通过使用语音输入接口130来接收语音输入。脚本输入接口120可以包括键盘用户接口(UI),在其上显示字符、数字和符号。对于包括在脚本输入接口120中的UI,字符对于每种类型的语言可以是不同的。例如,用于韩语输入的UI和用于英语输入的UI可以单独存在。文本输入设备100可以同时提供脚本输入接口120和语音输入接口130。

[0043] 处理器110可以确定是否满足语音激活条件和语音去激活条件。语音激活条件可以是确定用户10提供语音输入的意图的结果。语音去激活条件可以是确定用户10结束提供语音输入的意图的结果。处理器110可以基于是否满足所确定的语音激活条件来自动激活语音输入接口130,或基于是否满足语音去激活条件来自动去激活语音输入接口130。

[0044] 在实施例中,处理器110可以基于通过使用麦克风接收的声音信息来确定是否满足语音激活条件。根据实施例的处理器110可以基于先前确定的条件来激活麦克风。例如,当检测到脚本输入时,处理器110可以激活麦克风。处理器110可以确定用户语音信息是否包括在通过使用麦克风接收的声音信息中,并且可以基于上述确定的结果来确定是否满足语音激活条件。

[0045] 在实施例中,处理器110可以基于文本输入设备100的斜率(slope)来确定是否满足语音激活条件。可以通过使用加速度计或陀螺仪来测量文本输入设备100的斜率。处理器110可以通过将文本输入设备100的斜率与阈值进行比较来确定是否满足语音激活条件。可替代地,处理器110可以基于文本输入设备100的斜率是否是适用于用户面部识别的斜率来确定是否满足语音激活条件。然而,基于文本输入设备100的斜率来确定语音激活条件的操作不限于上述实施例。

[0046] 在实施例中,处理器110可以基于从用户的面部到文本输入设备100的距离来确定是否满足语音激活条件。根据实施例的处理器110可以基于通过使用相机接收的图像信息来确定从用户的面部到文本输入设备100的距离。处理器110可以通过将上述距离与阈值进行比较来确定是否满足语音激活条件。可替代地,处理器110可以基于上述距离是否是适用于用户面部识别的距离来确定是否满足语音激活条件。然而,由处理器110执行的、基于距离来确定语音激活条件的操作不限于上述实施例。

[0047] 在实施例中,处理器110可以基于确定用户的唇部运动来确定是否满足语音激活条件。根据实施例的处理器110可以基于先前确定的条件来激活用户侧的相机,例如前置相机。例如,当检测到脚本输入时,处理器110可以激活相机。处理器110可以基于通过使用相机接收的图像信息来确定用户的唇部运动。例如,处理器110可以基于接收到的图像信息中所包括的用户的唇部运动是否对应于语音输入来确定是否满足语音激活条件。

[0048] 在实施例中,处理器110可以基于情境信息来确定是否满足语音激活条件。情境信息可以包括文本输入设备100的周围环境信息、文本输入设备100的状态信息、用户状态信息、用户对文本输入设备100的使用历史信息、和用户日程信息中的至少一个,但是本公开不限于此。

[0049] 文本输入设备100的周围环境信息是指距文本输入设备100一定半径内的环境信息,并且可以包括天气信息、温度信息、湿度信息、照明信息、噪声信息、声音信息、时间信息等。例如,当周围环境是拥挤区域时,文本输入设备100可以确定不满足语音激活条件。可替代地,当当前时间不是很晚时,文本输入设备100可以以更高的频率满足语音激活条件。然而,情境信息中包括的周围环境信息不限于此。

[0050] 文本输入设备100的状态信息可以包括文本输入设备100的模式信息(例如,声音模式、振动模式、静音模式、省电模式、阻止(blocking)模式、多窗口模式、自动旋转模式等)、文本输入设备100的位置信息、时间信息、通信模块激活信息(例如,Wi-Fi打开/蓝牙关闭/GPS打开/NFC打开等)、文本输入设备100的网络连接状态信息、关于在文本输入设备100中执行的的应用的信息(例如,应用标识信息、应用类型、应用使用时间、应用使用周期(cycle)等)。例如,文本输入设备100可以基于当用户开始对话时文本输入设备100的姿势、角度、相对位置等来确定是否满足语音激活条件。然而,情境信息中包括的文本输入设备的状态信息不限于此。

[0051] 用户状态信息是关于用户的特征、运动和生活模式等的信息,并且可以包括关于用户的性别、用户的行走状态、锻炼状态、驾驶状态和睡眠状态、用户的情绪状态等的信息。例如,当识别出用户在语音输入期间频繁使用的手势时,文本输入设备100可以以更高的频率满足语音激活条件。然而,包括在情境信息中的用户状态信息不限于此。

[0052] 用户对文本输入设备100的使用历史信息是关于用户使用文本输入设备100的历史的信息,并且可以包括应用的执行历史、在应用中执行的功能的历史、用户的呼叫历史、用户的文本历史以及包括在文本信息中的单词的频率。例如,文本输入设备可以在与具体的(specific)用户的文本对话期间以更高的频率来满足语音激活条件。然而,包括在情境信息中的使用历史信息不限于此。

[0053] 在确定是否满足语音激活条件时,可以根据基于预设标准的学习(learning)来确定要使用哪个情境信息。例如,可以使用监督学习和无监督学习来确定语音激活条件,监督

学习使用特定语音输入和特定情境信息作为输入值,无监督学习通过在没有单独监督的情况下自主学习确定语音激活条件所需的情境信息的类型来发现语音激活条件的模式。此外,例如,可以使用强化学习来确定语音激活条件,强化学习使用关于根据学习掌握用户意图的结果是否正确的反馈。

[0054] 在实施例中,处理器110可以通过并行地或组合地组合上述实施例来确定是否满足语音激活条件。

[0055] 处理器110可以基于确定是否满足语音激活条件的结果来激活用于识别语音输入的语音输入接口130。处理器110可以同时提供语音输入接口130和脚本输入接口120。

[0056] 处理器110基于通过使用语音输入接口130识别的语音输入来生成语音文本信息。此外,处理器110基于通过使用脚本输入接口120识别的脚本输入来生成脚本文本信息。处理器110可以通过将语音文本信息与脚本文本信息进行组合来生成文本信息。

[0057] 处理器110可以控制文本输入设备100的显示器(未示出),使得在其上显示所生成的文本信息。为了显示文本信息,处理器110可以按生成的顺序来显示脚本文本信息和语音文本信息。

[0058] 处理器110可以确定是否满足语音去激活条件。语音去激活条件可以是确定用户10结束提供语音输入的意图的结果。

[0059] 在实施例中,处理器110可以基于情境信息来确定是否满足语音去激活条件。可替代地,处理器110可以基于通过使用麦克风接收的声音信息、文本输入设备100的斜率、距用户面部的距离、以及对用户的唇部运动的确定来确定是否满足语音去激活条件。例如,处理器110可以基于文本输入设备100的斜率是否是不适于用户面部识别的斜率来确定是否满足语音去激活条件。此外,处理器110可以基于用于确定是否满足语音激活条件的要素来确定语音去激活条件。然而,由处理器110执行的确定语音去激活条件的操作不限于上述实施例。

[0060] 处理器110可以基于确定是否满足语音去激活条件的结果将语音输入接口130去激活。文本输入设备100可以同时提供脚本输入接口120和语音输入接口130,并且基于语音激活条件或语音去激活条件来控制是自动还是根据用户的选择来激活语音输入接口130。因此,可以高效地执行文本输入操作。

[0061] 文本输入设备100的示例可以包括智能电话、平板个人计算机(PC)、智能电视(TV)、移动电话、个人数字助理(PDA)、膝上型计算机、媒体播放器、微服务器、全球定位系统(GPS)设备、电子书终端、数字广播终端、导航、信息亭(kiosk)、MP3播放器、数字相机、家用电器和其他移动或非移动计算设备,但是本公开不限于此。此外,文本输入设备100的示例可以包括具有通信功能和数据处理功能的可穿戴设备,诸如手表、眼镜、发带、或戒指。然而,本公开不限于此,并且文本输入设备100可以包括能够接收用户输入并基于其而提供文本的任何类型的设备。

[0062] 此外,文本输入设备100可以经由特定网络与服务器和其他设备(未示出)通信,以便使用各种情境信息。在这种情况下,网络可以包括局域网(LAN)、广域网(WAN)、增值网络(VAN)、移动无线电通信网络、卫星通信网络或其任何组合。网络可以是使网络配置实体能够有效地彼此通信的综合数据通信网络,并且可以包括有线互联网、无线互联网和移动无线通信网络。无线通信的示例可以包括无线LAN(Wi-Fi)、蓝牙、蓝牙低功耗、紫蜂Zigbee、

Wi-Fi直连(WFD)、超宽带(UWB)、红外数据协会(IrDA)和近场通信(NFC),但是本公开不限于此。

[0063] 图2是根据实施例的文本输入设备的图。

[0064] 参考图2,文本输入设备200包括处理器210、脚本输入接口220、语音输入接口230和传感器240。

[0065] 处理器210控制文本输入设备200的整体操作。

[0066] 文本输入设备200可以通过使用脚本输入接口220来接收脚本输入,并且可以通过使用语音输入接口230来接收语音输入。文本输入设备200可以同时提供脚本输入接口220和语音输入接口230。

[0067] 传感器240包括从外部接收信息的至少一个传感器。在实施例中,传感器240可以包括接收声音信息的麦克风。在实施例中,传感器240可以包括检测文本输入设备200的运动的加速度计或陀螺仪。在实施例中,传感器240可以包括接收图像信息的相机。在实施例中,传感器240可以包括接收周围环境信息的温度传感器或湿度传感器。然而,这是示例,并且包括在传感器240中的传感器的类型不限于此。传感器240可以将接收到的信息提供给处理器210。处理器210可以基于所提供的信息来确定是否满足语音激活条件。

[0068] 图3是根据实施例的文本输入设备的图。

[0069] 参考图3,文本输入设备300包括处理器310、脚本输入接口320、语音输入接口330和辅助语音输入接口340。

[0070] 用户30可以向文本输入设备300提供语音输入或脚本输入。语音输入是包括信息(例如,文本信息)的用户语音。脚本输入是包括信息(例如,文本信息)的用户的动作。用户的动作可以包括与设备的接触。

[0071] 处理器310可以控制脚本输入接口320和语音输入接口330,使得通过使用脚本输入接口320来输入脚本输入,并且通过使用语音输入接口330来输入语音输入。

[0072] 处理器310可以确定是否满足语音激活条件和语音去激活条件。处理器310可以基于“确定是否满足语音激活条件或语音去激活条件”的结果来自动激活或去激活语音输入接口330。

[0073] 处理器310可以基于情境信息、由通过使用语音输入接口330识别的语音输入来生成推荐语音文本信息。处理器310可以通过使用辅助语音输入接口340来向用户提供推荐语音文本信息。推荐语音文本信息是基于语音输入和情境信息而预期将由用户输入的文本信息。

[0074] 情境信息可以包括文本输入设备300的周围环境信息、文本输入设备300的状态信息、用户状态信息、用户对文本输入设备300的使用历史信息、和用户日程信息中的至少一个,但是本公开不限于此。

[0075] 在生成推荐语音文本信息时,可以根据基于预设标准的学习来确定要使用哪个情境信息。例如,可以使用监督学习和无监督学习来生成推荐语音文本信息,监督学习使用特定语音输入和特定情境信息作为输入值,无监督学习通过在没有单独监督的情况下自主学习推荐语音文本信息所需的情境信息的类型来发现推荐语音文本信息的模式。此外,例如,可以使用强化学习来生成推荐语音文本信息,强化学习使用关于根据学习掌握用户意图的结果是否正确的反馈。

[0076] 处理器310可以通过使用辅助语音输入接口340来接收用户输入,所述用户输入选择所提供的推荐语音文本信息中的一条作为语音文本信息。此外,处理器310可基于通过使用脚本输入接口320而识别的脚本输入来生成脚本文本信息。处理器310可以通过将所选择的语音文本信息与脚本文本信息进行组合来生成文本信息。

[0077] 处理器310可以控制显示器(未示出),使得在其上显示所生成的文本信息。为了显示文本信息,处理器310可以按生成的顺序显示脚本文本信息和语音文本信息。

[0078] 图4A是根据一些实施例的用于描述由文本输入设备400a执行的执行文本输入操作的方法的图。

[0079] 参考图4A,文本输入设备400a包括脚本输入接口420a和语音输入接口(未示出)。用户40可以向文本输入设备400a提供语音输入或脚本输入。

[0080] 文本输入设备400a可以通过使用脚本输入接口420a来接收脚本输入,并且可以通过使用语音输入接口来接收语音输入。文本输入设备400a可以基于通过使用语音输入接口识别的语音输入来生成语音文本信息,并且可以基于通过使用脚本输入接口420a识别的脚本输入来生成脚本文本信息。文本输入设备400a可以通过将语音文本信息与脚本文本信息进行组合来生成文本信息。

[0081] 在图4A中,脚本输入接口420a以QWERTY键盘UI的形式显示,但这是示例。根据本实施例的脚本输入接口420a的形式不限于上述键盘UI。此外,尽管图4A中未示出语音输入接口,但这是示例。脚本输入接口420a可以以UI的形式显示在文本输入设备400a上。

[0082] 当检测到文本输入请求时,文本输入设备400a可以提供脚本输入接口420a。例如,当检测到“请求用于文本输入的UI”的用户输入时,文本输入设备400a可以提供脚本输入接口420a。可替代地,文本输入设备400a可以响应于来自在文本输入设备400a中运行的应用请求而提供脚本输入接口。

[0083] 当文本输入设备400a提供脚本输入接口420a时,文本输入设备400a可以设置文本显示区域401a,其是用于显示文本信息的预设区域。例如,当运行搜索应用时,可以在文本显示区域401a中包括搜索窗口。

[0084] 如上所述,文本输入设备400a可以基于通过使用脚本输入接口320识别的脚本输入来生成脚本文本信息。脚本输入可以包括字符、数字和符号中的至少一个的文本。文本输入设备400a可以在文本显示区域401a上显示生成的脚本文本信息。

[0085] 文本输入设备400a可以确定是否满足语音激活条件,该语音激活条件是确定用户提供语音输入的意图的结果。在实施例中,文本输入设备400a可以基于下述因素的组合来确定是否满足语音激活条件:是否存在脚本输入、通过使用麦克风接收的声音信息、文本输入设备400a的斜率、与用户面部的距离以及用户的唇部运动。然而,确定语音激活条件的方法不限于上述实施例。

[0086] 文本输入设备400a可以基于所确定的语音激活条件,与脚本输入接口420a一并提供用于识别语音输入的语音输入接口。

[0087] 在实施例中,文本输入设备400a可以在提供语音输入接口的同时显示语音输入接口的状态,即,是激活的还是去激活的。例如,文本输入设备400a可以基于语音输入接口的状态来改变显示在文本输入设备400a上的语音输入激活图标402a的颜色。语音输入激活图标402a可以位于脚本输入接口420a上。然而,在提供语音输入接口的同时显示语音输入接

口状态的方法不限于上述实施例。

[0088] 如上所述,文本输入设备400a可以基于通过使用语音输入接口识别的语音输入来生成语音文本信息。文本输入设备400a可以在文本显示区域401a上显示所生成的语音文本信息。

[0089] 在实施例中,当与脚本输入接口一并提供语音输入接口时,文本输入设备400a可以在文本显示区域401a上显示基于语音输入生成的语音文本信息、或基于脚本输入生成的脚本文本信息。

[0090] 在实施例中,文本输入设备400a可以按照输入的顺序在文本显示区域401a上显示语音文本信息或脚本文本信息。当上述输入同时执行时,文本输入设备400a可以根据预先指定的优先级标志来确定语音文本信息或脚本文本信息的显示顺序。根据实施例的优先级标志可以由用户指定,或者可以预先指定并存储在存储器(未示出)中。

[0091] 图4B是根据一些实施例的用于描述由文本输入设备400b执行的执行文本输入操作的方法的图。

[0092] 参考图4B,文本输入设备400b包括脚本输入接口420b、语音输入接口(未示出)和辅助语音输入接口440b。用户40可以向文本输入设备400b提供语音输入或脚本输入。

[0093] 文本输入设备400b可以通过使用脚本输入接口420b来接收脚本输入,并且可以通过使用语音输入接口来接收语音输入。

[0094] 在图4B中,脚本输入接口420b以QWERTY键盘UI的形式显示,但这是示例。根据本实施例的脚本输入接口420b的形式不限于上述键盘UI。此外,尽管图4B中未示出语音输入接口,但这是示例。语音输入接口可以以UI的形式显示在文本输入设备400b上。

[0095] 当检测到文本输入请求时,文本输入设备400b提供脚本输入接口420b。当文本输入设备400b提供脚本输入接口420b时,文本输入设备400b可以设置文本显示区域401b,文本显示区域401b是显示文本信息的预设区域。

[0096] 文本输入设备400b可以基于通过使用脚本输入接口420b识别的脚本输入来生成脚本文本信息。文本输入设备400b可以在文本显示区域401b上显示生成的脚本文本信息。

[0097] 文本输入设备400b可以确定是否满足语音激活条件,该语音激活条件是确定用户提供语音输入的意图的结果。在实施例中,文本输入设备400b可以基于下述因素的组合来确定是否满足语音激活条件:是否存在脚本输入、通过使用麦克风接收的声音信息、文本输入设备400b的斜率、与用户面部的距离以及用户的唇部运动。然而,确定语音激活条件的方法不限于上述实施例。

[0098] 文本输入设备400b可以基于所确定的语音激活条件,与脚本输入接口420b一并提供用于识别语音输入的语音输入接口。

[0099] 在实施例中,文本输入设备400b可以在提供语音输入接口的同时显示语音输入接口的状态,即,是激活的还是去激活的。例如,文本输入设备400b可以基于语音输入接口的状态来改变显示在文本输入设备400b上的语音输入激活图标402b的颜色。然而,在提供语音输入接口的同时显示语音输入接口的状态的方法不限于上述实施例。

[0100] 文本输入设备400b可以基于情境信息由通过使用语音输入接口接收的语音输入来生成推荐语音文本信息。推荐语音文本信息是基于语音输入和情境信息而预期将由用户输入的文本信息。例如,当文本输入设备400b确定语音输入中包括文本信息“text”时,文本

输入设备400b可以生成“texture”、“text”和“texted”作为推荐语音文本信息。

[0101] 文本输入设备400b可以通过使用辅助语音输入接口440b来向用户提供推荐语音文本信息。此外,文本输入设备400b可以通过使用辅助语音输入接口440b来接收用户输入,该用户输入选择所提供的推荐语音文本信息中的一条作为语音文本信息。

[0102] 文本输入设备400b可以在文本显示区域401b上显示所选择的语音文本信息。

[0103] 更具体地,当与脚本输入接口一并提供语音输入接口时,文本输入设备400b可以在文本显示区域401b上显示基于语音输入而生成的语音文本信息、或基于脚本输入而生成的脚本文本信息。

[0104] 在实施例中,文本输入设备400b可以按照输入的顺序在文本显示区域401b上显示语音文本信息或脚本文本信息。当上述输入同时执行时,文本输入设备400b可以根据预先指定的优先级标志来确定语音文本信息或脚本文本信息的显示顺序。根据实施例的优先级标志可以由用户指定,或者可以预先指定并存储在存储器(未示出)中。

[0105] 图5是根据一些实施例的用于描述由文本输入设备执行的基于情境信息来执行文本输入操作的方法的图。

[0106] 参考图5,文本输入设备500a和500b分别包括脚本输入接口520a和520b、语音输入接口(未示出)以及辅助语音输入接口540a和540b。用户可以向文本输入设备500a和500b提供语音输入或脚本输入。

[0107] 文本输入设备500a和500b可以分别通过使用脚本输入接口520a和520b来接收脚本输入,并且可以分别通过使用语音输入接口来接收语音输入。当文本输入设备500a和500b分别提供脚本输入接口520a和520b时,文本输入设备500a和500b可以分别设置文本显示区域501a和501b,文本显示区域501a和501b是用于显示文本信息的预设区域。

[0108] 文本输入设备500a和500b可以确定是否满足语音激活条件,该语音激活条件是确定用户提供语音输入的意图的结果。在实施例中,文本输入设备500a和500b可以基于下述因素的组合来确定是否满足语音激活条件:是否存在脚本输入;通过使用麦克风接收的声音信息、文本输入设备500a和500b的斜率、与用户面部的距离以及用户的唇部运动。然而,确定语音激活条件的方法不限于上述实施例。

[0109] 文本输入设备500a和500b可以分别基于所确定的语音激活条件,与脚本输入接口520a和520b一并提供用于识别语音输入的语音输入接口。

[0110] 文本输入设备500a和500b可以分别基于情境信息、由通过使用语音输入接口接收的语音输入来生成推荐语音文本信息。推荐语音文本信息是基于语音输入和情境信息而预期将由用户输入的文本信息。

[0111] 在实施例中,情境信息包括用户的性别信息。在实施例中,可以由用户的语音输入确定用户的性别信息,或者可以预先存储/学习和确定用户的性别信息。

[0112] 参考图5,当情境信息中包括的用户的性别是女性时,文本输入设备500a可以生成“未婚夫”、“丈夫”和“男朋友”作为推荐语音文本信息。当情境信息中包括的用户的性别是男性时,文本输入设备500b可以生成“未婚妻”、“妻子”和“女朋友”作为推荐语音文本信息。图5中公开的每个单词是示例,并且由文本输入设备500a和500b执行的、基于性别信息生成推荐语音文本信息的方法不限于此。

[0113] 图6是根据实施例的用于描述文本输入方法的流程图。

[0114] 在操作S610中,文本输入设备提供用于接收脚本输入的脚本输入接口。脚本输入是包括信息(例如,文本信息)的用户的动作。用户的动作可以包括与设备的接触。脚本输入接口可以包括显示字符、数字和符号的键盘UI。在实施例,当检测到“请求用于文本输入的接口”的用户输入时,文本输入设备可以提供脚本输入接口。

[0115] 在操作S620中,文本输入设备确定是否满足语音激活条件。语音激活条件可以是确定用户提供语音输入的意图的结果。在实施例,文本输入设备可以基于下述因素的组合来确定是否满足语音激活条件:是否存在脚本输入、通过使用麦克风接收的声音信息、文本输入设备的斜率、距用户面部的距离以及用户的唇部运动。可替代地,当提供脚本输入接口时,文本输入设备可以自动确定满足了语音激活条件。然而,确定语音激活条件的方法不限于上述实施例。

[0116] 在操作S630中,文本输入设备基于所确定的语音激活条件,与脚本输入接口一并提供用于识别语音输入的语音输入接口。例如,当通过使用麦克风接收的声音信息中包括用户的语音、并且满足语音激活条件时,文本输入设备可以在用户的语音存在时与脚本输入接口一并提供语音输入接口。可替代地,当提供脚本输入接口并且满足语音激活条件时,文本输入设备可以在任何时间同时自动提供脚本输入接口和语音输入接口。即,当满足语音激活条件时,文本输入设备可以激活语音输入接口,以使用语音输入和脚本输入两者来向用户提供文本输入。

[0117] 图7是根据实施例的用于描述确定语音激活条件的方法的流程图。

[0118] 在操作S710中,文本输入设备激活麦克风。当满足预设条件时,文本输入设备可以激活麦克风。例如,当存在脚本输入时,文本输入设备可以激活麦克风。在实施例,当存在特定时间内经由脚本输入接口输入的脚本输入时,文本输入设备可以确定存在脚本输入。

[0119] 在操作S720中,文本输入设备通过使用麦克风接收声音信息。声音信息是文本输入设备周围的听觉信息,其通过使用麦克风来接收。

[0120] 在操作S730中,文本输入设备确定用户的语音是否包括在接收到的声音信息中。可以基于各种算法来执行由文本输入设备执行的确定是否包括了用户的语音的方法。此外,文本输入设备可以执行噪声消除和各种信号处理操作,以便确定是否包括用户的语音。

[0121] 在操作S740中,文本输入设备基于是否包括了用户的语音来确定是否满足语音激活条件。当识别出用户的语音时,文本输入设备可以确定已经满足语音激活条件,并且经由语音向用户提供文本输入。

[0122] 图8是根据实施例的用于描述确定语音激活条件的方法的流程图。

[0123] 在操作S810中,文本输入设备测量文本输入设备的斜率。在实施例,文本输入设备可以通过使用加速度计或陀螺仪来测量文本输入设备的斜率。在实施例,当满足预设条件时,文本输入设备可以执行斜率测量操作。例如,文本输入设备可以在存在脚本输入时测量斜率。在实施例,当存在特定时间内经由脚本输入接口输入的脚本输入时,文本输入设备可以确定存在脚本输入。

[0124] 在操作S820中,文本输入设备基于所测量的斜率来确定是否满足语音激活条件。在实施例,文本输入设备可以通过将所测量的斜率与阈值进行比较来确定是否满足语音激活条件。可替代地,文本输入设备可以基于所测量的斜率是否是适合于用户面部识别的斜率来确定是否满足语音激活条件。然而,基于文本输入设备的斜率确定语音激活条件的

操作不限于上述实施例。

[0125] 图9是根据实施例的用于描述确定语音激活条件的方法的流程图。

[0126] 在操作S910中,文本输入设备测量从用户的面部到文本输入设备的距离。在实施例中,文本输入设备可以通过使用用户侧的相机(例如,前置相机)来测量上述距离。在实施例中,当满足预设条件时,文本输入设备可以激活相机。例如,当存在脚本输入时,文本输入设备可以激活相机。可替代地,当文本输入设备的斜率满足特定条件时,文本输入设备可以激活相机。然而,文本输入设备的相机激活操作和距离测量操作不限于上述实施例。

[0127] 在操作S920中,文本输入设备基于所测量的距离来确定是否满足语音激活条件。在实施例中,文本输入设备可以通过将所测量的距离与阈值进行比较来确定是否满足语音激活条件。可替代地,文本输入设备可以基于所测量的距离是否是适合于用户面部识别的距离来确定是否满足语音激活条件。然而,由文本输入设备执行的基于距离确定语音激活条件的操作不限于上述实施例。

[0128] 图10是根据实施例的用于描述确定语音激活条件的方法的流程图。

[0129] 在操作S1010中,文本输入设备激活相机以接收图像信息。在实施例中,文本输入设备可以激活用户侧的相机,例如前置相机。在实施例中,当满足预设条件时,文本输入设备可以激活相机。例如,当存在脚本输入时,文本输入设备可以激活相机。可替代地,当文本输入设备的斜率满足特定条件时,文本输入设备可以激活相机。然而,文本输入设备的相机激活操作和距离测量操作不限于上述实施例。

[0130] 在操作S1020中,文本输入设备基于所接收到的图像信息来确定用户的唇部移动。在实施例中,文本输入设备可以从图像信息识别用户的面部。

[0131] 在操作S1030中,文本输入设备基于所确定的唇部运动来确定是否满足语音激活条件。在实施例中,文本输入设备可以基于是否存在唇部运动来确定是否满足语音激活条件。可替代地,文本输入设备可以基于所确定的唇部运动是否包括文本信息来确定是否满足语音激活条件。然而,由文本输入设备执行的基于唇部运动确定语音激活条件的操作不限于上述实施例。

[0132] 图11是根据实施例的用于描述确定语音激活条件的方法的流程图。

[0133] 在操作S1110中,文本输入设备测量文本输入设备的斜率。在实施例中,文本输入设备可以通过使用加速度计或陀螺仪来测量文本输入设备的斜率。例如,当满足预设条件时,文本输入设备可以执行斜率测量操作。

[0134] 在操作S1120中,文本输入设备激活相机以接收图像信息。在实施例中,文本输入设备可以通过使用用户侧的相机(例如,前置相机)来接收图像信息。在实施例中,当满足预设条件时,文本输入设备可以激活相机。例如,当存在脚本输入时,文本输入设备可以激活相机。可替代地,当文本输入设备的斜率满足特定条件时,文本输入设备可以激活相机。

[0135] 在操作S1130中,文本输入设备测量从用户的面部到文本输入设备的距离。在实施例中,文本输入设备可以基于通过使用相机接收的图像信息来测量上述距离。

[0136] 在操作S1140中,文本输入设备激活麦克风以接收声音信息。当满足预设条件时,文本输入设备可以激活麦克风。例如,当相机被激活时,文本输入设备可以一并激活麦克风。可替代地,当文本输入设备与用户的面部之间的距离满足特定条件时,文本输入设备可以激活麦克风。

[0137] 在操作S1150中,文本输入设备确定是否满足语音激活条件。例如,当从通过使用麦克风接收的声音信息中识别出用户的语音时,文本输入设备可以确定已经满足语音激活条件。可替代地,当从通过使用相机接收的图像信息中识别出用户的唇部运动时,文本输入设备可以确定已经满足语音激活条件。可替代地,当从声音信息中识别出用户的语音、并且从通过使用相机接收的图像信息中识别出用户的唇部运动时,文本输入设备可以确定已经满足语音激活条件。文本输入设备可以通过使用各种传感器的感测操作来确定是否满足语音激活条件,并且经由语音向用户提供文本输入。

[0138] 图12是根据实施例的用于描述文本输入方法的流程图。

[0139] 在操作S1210中,文本输入设备提供用于识别脚本输入的脚本输入接口。在实施例中,当检测到请求用于文本输入的接口的用户输入时,文本输入设备可以提供脚本输入接口。

[0140] 在操作S1220中,文本输入设备确定是否满足语音激活条件。在实施例中,文本输入设备可以基于下述因素的组合来确定是否满足语音激活条件:是否存在脚本输入、通过使用麦克风接收的声音信息、文本输入设备的斜率、与用户面部的距离、以及用户的唇部运动。

[0141] 在操作S1230中,文本输入设备基于确定是否满足语音激活条件的结果,同时提供用于识别语音输入的语音输入接口和脚本输入接口。

[0142] 在操作S1240中,文本输入设备确定是否满足语音去激活条件。语音去激活条件是确定用户结束提供语音输入的意图的结果。在实施例中,文本输入设备可以基于通过使用麦克风接收的声音信息、文本输入设备的斜率、与用户面部的距离、用户的嘴唇运动等来确定是否满足语音去激活条件。例如,文本输入设备可以基于文本输入设备的斜率是否是不适合于用户面部识别的斜率来确定是否满足语音去激活条件。此外,文本输入设备可以基于用于确定是否满足语音激活条件的要素来确定语音去激活条件。

[0143] 在操作S1250中,文本输入设备基于是否满足语音去激活条件将语音输入接口去激活。文本输入设备可以同时提供脚本输入接口和语音输入接口,并且基于语音激活条件或语音去激活条件来控制是自动激活语音输入接口还是根据用户的选择来激活语音输入接口。因此,可以高效地执行文本输入操作。

[0144] 图13和图14是根据一些实施例的第一设备1000的框图。

[0145] 如图13所示,根据一些实施例的第一设备1000可以包括用户输入器1100、输出器1200、处理器1300和通信器1500。然而,图13中所示的所有元件对于第一设备1000不是必需的。第一设备1000可以包括比图13中所示的元件更多的元件,或者可以包括比图13中所示的元件更少的元件。

[0146] 例如,如图14所示,根据一些实施例的第一设备1000,除了包括用户输入器1100、输出器1200、处理器1300和通信器1500之外,还可以包括传感器1400、音频/视频(A/V)输入器1600和存储器1700。

[0147] 用户输入器1100是允许用户输入用于控制第一设备1000的数据的设备。例如,用户输入器1100可以包括键盘、圆顶开关、触摸板(例如,触摸型电容式触摸板、压力型电阻式触摸板、红外光束感测型触摸板、表面声波型触摸板、集成应变仪型触摸板、压电效应型触摸板等)、滚轮和微动开关(jog switch),但是本公开不限于此。

[0148] 用户输入器1100可以包括脚本输入接口和语音输入接口。用户输入器1100可以接收语音输入或脚本输入。此外,用户输入器1100可以接收选择推荐语音文本信息中的一条作为语音文本信息的用户输入。

[0149] 输出器1200可以输出音频信号、视频信号或振动信号。输出器1200可以包括显示器1210、音频输出器1220和振动马达1230。

[0150] 显示器1210显示由第一设备1000处理的信息。例如,显示器1210可以显示用于接收脚本输入的脚本输入接口,或显示指示语音输入接口是否被激活的语音输入激活图标。此外,显示器1210可以显示文本信息和作为显示文本信息的预设区域的文本显示区域。

[0151] 当显示器1210和触摸板形成分层结构以构成触摸屏时,显示器1210也可以用作输入设备以及输出设备。显示器1210可以包括从液晶显示器(LCD)、薄膜晶体管液晶显示器(TFT-LCD)、有机发光二极管(OLED)、柔性显示器、三维(3D)显示器和电泳显示器中选择的至少一种。根据第一设备1000的实现类型,第一设备1000可以包括两个或更多个显示器1210。可以通过使用铰链将两个或更多个显示器1210排列为彼此面对。

[0152] 音频输出器1220输出从通信器1500接收的、或存储在存储器1700中的音频数据。此外,音频输出器1220输出与由第一设备1000执行的功能相关联的音频信号(例如,呼叫信号接收声音、消息接收声音、通知声音等)。音频输出器1220可以包括扬声器、蜂鸣器等。

[0153] 振动马达1230可以输出振动信号。例如,振动马达1230可以输出与音频数据/视频数据的输出相对应的振动信号(例如,呼叫信号接收声音、消息接收声音等)。此外,当向触摸屏输入了触摸时,振动马达1230可以输出振动信号。

[0154] 处理器1300控制第一设备1000的整体操作。例如,处理器1300可以执行存储在存储器1700中的程序,以控制用户输入器1100、输出器1200、传感器1400、通信器1500和A/V输入器1600。

[0155] 具体地,处理器1300可以控制用户输入器1100,使得通过使用脚本输入接口来输入脚本输入,并且通过使用语音输入接口来输入语音输入。

[0156] 处理器1300可以确定是否满足语音激活条件和语音去激活条件。在实施例中,处理器1300可以基于下述因素的组合来确定是否满足语音激活条件和语音去激活条件:存在或不存在的脚本输入、通过使用麦克风接收的声音信息、第一设备1000的斜率、与用户面部的距离、以及用户的嘴唇运动。然而,确定语音激活条件的方法不限于上述实施例。

[0157] 处理器1300可以基于所确定的语音激活条件或所确定的语音去激活条件来自动激活或去激活包括在用户输入器1100中的语音输入接口。

[0158] 此外,处理器1300可以基于情境信息,从通过使用用户输入器1100识别的语音输入来生成推荐语音文本信息。处理器1300可以向用户提供所生成的推荐语音文本信息。

[0159] 处理器1300可以控制用户输入器1100,以通过使用包括在用户输入器1100中的辅助语音输入接口来接收选择所提供的推荐语音文本信息中的一条作为语音文本信息的用户输入。此外,处理器1300可基于通过使用脚本输入接口识别的脚本输入来生成脚本文本信息。处理器1300可以通过将所选择的语音文本信息与上述脚本文本信息组合来生成文本信息。

[0160] 处理器1300可以控制显示器1210,使得在其上显示所生成的文本信息。为了显示文本信息,处理器1300可以按生成的顺序显示脚本文本信息和语音文本信息。

[0161] 传感器1400可以感测第一设备1000的状态或第一设备1000周围的状态,并将感测到的信息发送到处理器1300。

[0162] 传感器1400可以包括地磁传感器1410、加速度传感器1420、温度/湿度传感器1430、红外传感器1440、陀螺仪传感器1450、位置传感器(例如,GPS)1460、气压传感器1470、接近传感器1480和RGB传感器(照度传感器)1490中的至少一个,但是本公开不限于此。因为可以从各个传感器的名称直观地推断它们的功能,所以将省略其详细描述。

[0163] 通信器1500可以包括一个或多个元件,用于第一设备1000和头戴式显示器(HMD)设备之间的通信,或第一设备1000和服务器之间的通信。例如,通信器1500可以包括短距离无线通信器1510、移动通信器1520和广播接收器1530。

[0164] 短距离无线通信器1510可以包括蓝牙通信器、蓝牙低功耗(BLE)通信器、近场通信器、无线局域网(WLAN)(Wi-Fi)通信器、Zigbee紫蜂通信器、红外数据协会(IrDA)通信器、Wi-Fi直连(WFD)通信器、超宽带(UWB)通信器、Ant+通信器等,但是本公开不限于此。

[0165] 移动通信器1520可以向移动通信网络上的基站、外部终端或服务器中的至少一个发送无线信号,或者从移动通信网络上的基站、外部终端或服务器中的至少一个接收无线信号。根据发送或接收的文本/多媒体消息,无线信号的示例可以包括语音呼叫信号、视频呼叫信号以及各种类型的数据。

[0166] 广播接收器1530可以经由广播信道从外部接收广播信号和/或广播相关信息。广播信道可以包括卫星信道、地波信道等。根据实现示例,第一设备1000可以不包括广播接收器1530。

[0167] 此外,通信器1500可以向HMD装置、服务器和外围装置发送用于使用情境信息的信息,或者从HMD装置、服务器和外围装置接收用于使用情境信息的信息。

[0168] 可以提供A/V输入器1600,用于输入音频信号或视频信号。为了输入音频信号或视频信号,A/V输入器1600可以包括相机1610、麦克风1620等。相机1610可以在视频呼叫模式或图像捕获模式下通过图像传感器获得图像帧,诸如静止图像或运动图像。由图像传感器捕获的图像可以由处理器1300或单独的图像处理(未示出)处理。

[0169] 由相机1610处理的图像帧可以存储在存储器1700中,或者可以通过通信器1500发送到外部。可替代地,图像帧可以用于由处理器1300执行的、确定语音激活条件和语音去激活条件的操作。相机1610可以根据终端的配置而包括两个或更多个相机。

[0170] 麦克风1620可以接收外部音频信号并将外部音频信号处理成电(electrical)语音数据。例如,麦克风1620可以从外部设备或人接收音频信号。麦克风1620可以使用各种噪声消除算法来消除在接收外部音频信号的过程中产生的噪声。通过麦克风1620接收的音频信号可以用于由处理器1300执行的、确定语音激活条件和语音去激活条件的操作。

[0171] 存储器1700可以存储用于处理和控制处理器1300的程序,并且可以存储输入到第一设备1000的数据或从第一设备1000输出的数据。

[0172] 存储器1700可以包括从闪存型、硬盘型、多媒体卡微型、卡型存储器(例如,安全数字(SD)或极端数字(XD)存储器)、随机访问存储器(RAM)、静态随机访问存储器(SRAM)、只读存储器(ROM)、电可擦除可编程ROM(EEPROM)、可编程ROM(PROM)、磁存储器、磁盘和光盘选择的至少一种类型的存储介质。

[0173] 可以将存储在存储器1700中的程序根据其功能分类为多个模块。例如,可以将程

序分类为UI模块1710、触摸屏模块1720和通知模块1730。

[0174] UI模块1710可以根据每个应用来提供与第一设备1000交互操作的专用UI或GUI。触摸屏模块1720可以检测用户在触摸屏上的触摸手势,并将关于触摸手势的信息发送到处理器1300。根据一些实施例的触摸屏模块1720可以识别和分析触摸代码。可以将触摸屏模块1720实现为包括控制器的单独硬件。

[0175] 可以在触摸屏内部或附近提供各种传感器,以便检测触摸屏上的触摸或检测在触摸屏上方的悬停。检测触摸屏上的触摸的传感器的示例可以是触觉传感器。触觉传感器可以检测处于/超越(beyond)人的敏感度的具体对象的接触。触觉传感器可以感测各种信息,诸如接触表面的粗糙度、接触对象的硬度、或接触点的温度。

[0176] 此外,检测触摸屏上的触摸的传感器的示例可以是接近传感器。

[0177] 接近传感器是指在没有机械接触的情况下,通过使用电磁力或红外光来检测接近特定检测表面的对象或处于传感器附近的对象存在或不存在的传感器。接近传感器的示例包括透射型光电传感器、直接反射型光电传感器、镜面反射型光电传感器、高频振荡型接近传感器、电容型接近传感器、磁性接近传感器和红外型接近传感器。用户的触摸手势可以包括轻击(tap)、触摸保持(touch-and-hold)、双击、拖动、平移、轻拂(flick)、拖放、滑动(swipe)等。

[0178] 通知模块1730可以输出用于通知在第一设备1000中发生事件的信号。在第一设备1000中发生的事件的示例包括呼叫信号接收、消息接收、键信号输入和日程通知。通知模块1730可以通过显示器1210以视频信号的形式输出通知信号。通知模块1730可以通过音频输出器1220以音频信号的形式输出通知信号。通知模块1730可以通过振动马达1230以振动信号的形式输出通知信号。

[0179] 图15是根据一些实施例的处理器1300的框图。

[0180] 参考图15,根据一些实施例的处理器1300可以包括数据训练器1310和数据识别器1320。

[0181] 根据实施例,可以将数据训练器1310的至少一部分和数据识别器1320的至少一部分实现为软件模块,或者可以以硬件芯片的形式制造并安装(mount)在文本输入设备上。

[0182] 数据训练器1310可以学习用于确定语音激活和去激活条件并生成推荐语音文本信息的标准。数据训练器1310可以学习关于哪些数据将被用于确定语音激活和去激活条件并生成推荐语音文本信息的标准。此外,数据训练器1310可以学习关于如何通过使用数据来确定用户的意图、如何获得和提供相关联的信息、以及如何确定语音激活和去激活条件并生成推荐语音文本信息的标准。数据训练器1310可以通过获得要用于学习的数据并将获得的数据应用于下面要描述的数据识别模型,来学习用于确定语音激活和去激活条件并生成推荐语音文本信息的标准。

[0183] 输入到数据训练器1310的训练数据的类型可以包括语音数据、文本数据等。

[0184] 数据识别器1320可以基于上述数据来确定语音激活和去激活条件或者生成推荐语音文本信息。数据识别器1320可以通过使用学习的(learned)数据识别模型从特定数据确定语音激活和去激活条件或者生成推荐语音文本信息。数据识别器1320可以根据通过学习获得的预设标准来获得特定数据,并使用所获得的数据作为输入值来使用数据识别模型。此外,数据识别器1320可以使用数据识别模型基于特定数据来确定语音激活和去激活

条件或者生成推荐语音文本信息。此外,由数据识别模型使用所获得的数据作为输入值而输出的结果值可用于更新数据识别模型。

[0185] 数据训练器1310和数据识别器1320中的至少一个可以以至少一个硬件芯片的形式制造并安装在第一设备1000上。例如,数据训练器1310和数据识别器1320中的至少一个可以以用于人工智能(AI)的专用硬件芯片的形式制造,或者可以制造为现有通用处理器(例如,中央处理单元(CPU)或应用处理器)或专用图形处理器(例如,图形处理单元(GPU))的一部分并安装在第一设备1000上。在这种情况下,用于AI的专用硬件芯片是具体用于概率计算的专用处理器。因为专用硬件芯片具有比现有通用处理器更高的并行处理性能,所以可以快速处理AI领域中的计算操作,诸如机器学习。

[0186] 数据训练器1310和数据识别器1320可以安装在单个文本输入设备上,或者可以分别安装在单独的设备上。例如,数据训练器1310和数据识别器1320中的一个可以包括在单个文本输入设备中,并且其中的另一个可以包括在服务器中。此外,由数据训练器1310构建的模型信息可以通过有线或无线通信提供给数据识别器1320,并且输入到数据识别器1320的数据可以作为附加训练数据提供给数据训练器1310。

[0187] 可以将数据训练器1310和数据识别器1320中的至少一个实现为至少一个软件模块。当数据训练器1310和数据识别器1320中的至少一个被实现为软件模块(或包括指令的程序模块)时,软件模块可以存储在非暂时性计算机可读介质中。此外,在这种情况下,可以由操作系统(OS)或者由特定应用提供至少一个软件模块。可替代地,可以由OS提供至少一个软件模块的一部分,并且可以由特定应用提供其剩余部分。

[0188] 图16是根据一些实施例的数据训练器1310的框图。

[0189] 参考图16,根据一些实施例的数据训练器1310可以包括数据获得器1310-1、预处理器1310-2、训练数据选择器1310-3、模型学习器1310-4、和模型评估器1310-5。

[0190] 数据获得器1310-1可以获得确定语音激活和去激活条件或生成推荐语音文本信息所需的数据。数据获得器1310-1可以获得例如特定用户语音和特定情境信息。

[0191] 预处理器1310-2可以预处理所获得的数据,使得所获得的数据用于学习以确定语音激活和去激活条件或生成推荐语音文本信息。预处理器1310-2可以将所获得的数据处理成预设格式,使得下面将要描述的模型学习器1310-4能够使用所获得的数据进行学习,以确定语音激活和去激活条件并生成推荐语音文本信息。

[0192] 训练数据选择器1310-3可以从预处理后的数据条目中选择学习所需的数据。所选择的数据可以提供给模型学习器1310-4。训练数据选择器1310-3可以根据用于确定语音激活和去激活条件或生成推荐语音文本信息的预设标准从预处理后的数据条目中选择学习所需的数据。此外,训练数据选择器1310-3可以根据通过由下面将要描述的模型学习器1310-4执行的学习的预设标准来选择数据。

[0193] 模型学习器1310-4可以基于训练数据来学习关于确定语音激活和去激活条件还是生成推荐语音文本信息的标准。此外,模型学习器1310-4可以学习关于哪个训练数据将被用于确定语音激活和去激活条件并生成推荐语音文本信息的标准。

[0194] 此外,模型学习器1310-4可以学习关于哪个训练数据将被用于确定语音激活和去激活条件并生成推荐语音文本信息的标准。在这种情况下,数据识别模型可以是预先构建的模型。例如,数据识别模型可以通过接收基本训练数据(例如,样本数据等)预先构建的

模型。

[0195] 可以考虑识别模型的应用领域、学习目的、设备的计算机性能等来构建数据识别模型。可以将数据识别模型设计为在计算机上模拟人脑结构。数据识别模型可以包括具有权重(weights)的多个网络节点,其模拟人类神经网络的神经元。网络节点可以形成连接关系,使得神经元模拟通过突触交换信号的神经元的突触活动。数据识别模型可以包括例如神经网络模型或从神经网络模型开发的深度学习模型。在深度学习模型中,多个网络节点可以位于不同深度(或层)处且根据卷积连接关系交换数据。例如,数据识别模型可以包括诸如深度神经网络(deep neural network,DNN)、循环神经网络(recurrent neural network,RNN)、或双向循环深度神经网络(bidirectional recurrent deep neural network,BRDNN)的模型,但是本公开不限于此。

[0196] 根据各种实施例,当存在多个预先构建的识别模型时,模型学习器1310-4可以将输入训练数据和基本训练数据之间具有大相关性的数据识别模型确定为要学习的数据识别模型。在这种情况下,可以针对每种数据类型预先分类基本训练数据,并且可以针对每种数据类型预先构建数据识别模型。例如,可以基于各种标准—诸如生成训练数据的区域、生成训练数据的时间、训练数据的大小、训练数据的类型、训练数据的创建者、和训练数据中的对象的类型—来预先对基本训练数据进行分类。

[0197] 此外,模型学习器1310-4可以通过使用例如包括误差反向传播或梯度下降的学习算法来学习数据识别模型。

[0198] 此外,模型学习器1310-4可以通过例如使用训练数据作为输入值的监督学习来学习数据识别模型。此外,模型学习器1310-4可以在没有单独监督的状况(situation)下自主学习确定情况所需的数据类型,以便通过找到用于确定状况的标准的无监督学习来学习数据识别模型。此外,模型学习器1310-4可以通过例如使用关于根据学习确定状况的结果是否正确反馈的强化学习,来学习数据识别模型。

[0199] 此外,当学习数据识别模型后,模型学习器1310-4可以存储学习的数据识别模型。在这种情况下,模型学习器1310-4可以将学习的数据识别模型存储在包括数据识别器1320的电子设备的存储器中。可替代地,模型学习器1310-4可以将学习的数据识别模型存储在包括数据识别器1320—这将在下面描述—的电子设备的存储器中。可替代地,模型学习器1310-4可以将学习的数据识别模型存储在经由有线或无线网络连接到电子设备的服务器的存储器中。

[0200] 在这种情况下,存储学习的数据识别模型的存储器还可以存储例如与电子设备的至少一个其他元件相关联的命令或数据。此外,存储器可以存储软件和/或程序。程序可以包括例如内核、中间件、应用编程接口(API)和/或应用程序(或“应用”)。

[0201] 当将评估数据输入到数据识别模型、并且从评估数据输出的识别结果不满足特定标准时,模型评估器1310-5可以使模型学习器1310-4再次学习。在这种情况下,评估数据可以是评估数据识别模型而预设的数据。

[0202] 例如,在学习的数据识别模型对评估数据的分析结果中,当评估数据的分析结果不准确的数量或比率超过预设阈值时,模型评估器1310-5可以评估为不满足特定标准。例如,在将特定标准定义为2%的比率的情况下,当学习的数据识别模型对于总共1000条评估数据中的超过20条评估数据输出不正确的识别结果时,模型评估器1310-5可以评估该学习

的数据识别模型是不合适的。

[0203] 当存在多个学习的数据识别模型时,模型评估器1310-5可以评估每个学习的数据识别模型是否满足特定标准,并将满足特定标准的模型确定为最终数据识别模型。在这种情况下,当存在满足特定标准的多个模型时,模型评估器1310-5可以将以评估分数的降序预设的任何一个模型或特定数量的模型确定为最终数据识别模型。

[0204] 数据训练器1310中的数据获得器1310-1、预处理器1310-2、训练数据选择器1310-3、模型学习器1310-4和模型评估器1310-5中的至少一个可以以硬件芯片的形式制造并安装在电子设备上。例如,数据获得器1310-1、预处理器1310-2、训练数据选择器1310-3、模型学习器1310-4和模型评估器1310-5中的至少一个可以以用于AI的专用硬件芯片的形式制造,或者可以制造为现有通用处理器(例如,CPU或应用处理器)或专用图形处理器(例如,GPU)的一部分并安装在上述各种电子设备上。

[0205] 此外,数据获得器1310-1、预处理器1310-2、训练数据选择器1310-3、模型学习器1310-4和模型评估器1310-5可以安装在单个电子设备上,或者可以分别安装在单独的电子设备上。例如,数据获得器1310-1、预处理器1310-2、训练数据选择器1310-3、模型学习器1310-4和模型评估器1310-5中的一些可以包括在电子设备中,并且上述单元中的其他单元可以包括在服务器中。

[0206] 此外,可以将数据获得器1310-1、预处理器1310-2、训练数据选择器1310-3、模型学习器1310-4和模型评估器1310-5中的至少一个实现为软件模块。当数据获得器1310-1、预处理器1310-2、训练数据选择器1310-3、模型学习器1310-4和模型评估器1310-5中的至少一个被实现为软件模块(或包括指令的程序模块)时,软件模块可以存储在非暂时性计算机可读介质中。此外,在这种情况下,可以由OS或者由特定应用提供至少一个软件模块。可替代地,可以由OS提供至少一个软件模块的一部分,并且可以由特定应用提供其剩余部分。

[0207] 图17是根据一些实施例的数据识别器1320的框图。

[0208] 参考图17,根据一些实施例的数据识别器1320可以包括数据获得器1320-1、预处理器1320-2、识别数据选择器1320-3、识别结果提供器1320-4和模型更新器1320-5。

[0209] 数据获得器1320-1可以获得状况确定所需的数据,并且预处理器1320-2可以预处理所获得的数据,以便使得能够使用针对状况确定所获得的数据。预处理器1320-2可以将获得的数据处理为预设格式,使得识别结果提供器1320-4使用获得的数据进行状况确定。

[0210] 识别数据选择器1320-3可以从预处理后的数据条目中选择状况确定所需的数据。所选择的数据可以提供给识别结果提供器1320-4。识别数据选择器1320-3可以根据用于状况确定的预设标准来选择所有或部分预处理后的数据。此外,识别数据选择器1320-3可以根据通过由下面将要描述的模型学习器1310-4执行的学习的预设标准来选择数据。

[0211] 识别结果提供器1320-4可以通过将所选择的数据应用于数据识别模型来确定状况。识别结果提供器1320-4可以根据数据识别目的来提供识别结果。识别结果提供器1320-4可以通过使用由识别数据选择器1320-3选择的数据作为输入值,来将所选择的数据应用于数据识别模型。此外,可以通过数据识别模型来确定分析结果。

[0212] 模型更新器1320-5可以基于由识别结果提供器1320-4提供的识别结果的评估来更新数据识别模型。例如,模型更新器1320-5可以向模型学习器1310-4提供由识别结果提供器1320-4提供的识别结果,使得模型学习器1310-4更新数据识别模型。

[0213] 数据识别器1320中的数据获得器1320-1、预处理器1320-2、识别数据选择器1320-3、识别结果提供器1320-4和模型更新器1320-5中的至少一个可以以硬件芯片的形式制造并安装在电子设备上。例如,数据获得器1320-1、预处理器1320-2、识别数据选择器1320-3、识别结果提供器1320-4和模型更新器1320-5中的至少一个可以以用于AI的专用硬件芯片的形式来制造,或者可以制造为现有通用处理器(例如,CPU或应用处理器)或专用图形处理器(例如,GPU)的一部分并安装在上述各种电子设备上。

[0214] 此外,数据获得器1320-1、预处理器1320-2、识别数据选择器1320-3、识别结果提供器1320-4和模型更新器1320-5可以安装在单个电子设备上,或者可以分别安装在单独的电子设备上。例如,数据获得器1320-1、预处理器1320-2、识别数据选择器1320-3、识别结果提供器1320-4和模型更新器1320-5中的一些可以包括在电子设备中,并且上述单元中的其他单元可以包括在服务器中。

[0215] 此外,可以将数据获得器1320-1、预处理器1320-2、识别数据选择器1320-3、识别结果提供器1320-4和模型更新器1320-5中的至少一个实现为软件模块。当数据获得器1320-1、预处理器1320-2、识别数据选择器1320-3、识别结果提供器1320-4和模型更新器1320-5中的至少一个被实现为软件模块(或包括指令的程序模块)时,软件模块可以存储在非暂时性计算机可读介质中。此外,在这种情况下,可以由OS或者由特定应用提供至少一个软件模块。可替代地,可以由OS提供至少一个软件模块的一部分,并且可以由特定应用提供其剩余部分。

[0216] 图18是示出根据一些实施例的文本输入设备1000和服务器2000彼此互操作以训练和识别数据的示例的图。

[0217] 参考图18,服务器2000可以学习用于确定语音激活和去激活条件以及生成推荐语音文本信息的标准,并且文本输入设备1000可以基于由服务器2000执行的学习结果来确定语音激活和去激活条件的确定以及推荐语音文本信息的生成。

[0218] 在这种情况下,服务器2000的模型学习器2340可以执行图16所示的数据训练器1310的功能。服务器2000的模型学习器2340可以学习关于哪些数据将用于确定语音激活和去激活条件并生成推荐语音文本信息的标准。此外,服务器2000的模型学习器2340可以通过使用数据来学习用于确定语音激活和去激活条件并生成推荐语音文本信息的标准。模型学习器2340可以通过获得将用于学习的数据并将获得的数据应用于下面描述的数据识别模型来学习用于确定语音激活和去激活条件并生成推荐语音文本信息的标准。

[0219] 此外,文本输入设备1000的识别结果提供器1320-4可以通过将由识别数据选择器1320-3选择的数据应用于由服务器2000生成的数据识别模型来确定语音激活和去激活条件并生成推荐语音文本信息。例如,识别结果提供器1320-4可以将由识别数据选择器1320-3选择的数据发送到服务器2000,并且服务器2000可以通过将由识别数据选择器1320-3选择的数据应用于数据识别模型来请求确定语音激活和去激活条件并生成推荐语音文本信息。此外,识别结果提供器1320-4可以从服务器2000接收由服务器2000确定的关于语音激活和去激活条件的确定以及推荐语音文本信息的生成的信息。

[0220] 可替代地,文本输入设备1000的识别结果提供器1320-4可从服务器2000接收由服务器2000生成的数据识别模型,并且可以使用接收到的数据识别模型来确定语音激活和去激活条件并生成推荐语音文本信息。在这种情况下,文本输入设备1000的识别结果提供器

1320-4可以通过将由识别数据选择器1320-3选择的数据应用于从服务器2000接收的数据识别模型来确定语音激活和去激活条件并生成推荐语音文本信息。

[0221] 此外,文本输入设备1000和服务器2000可以有效地分发和执行用于数据识别模型的学习和数据识别的操作。以这种方式,可以高效地执行数据处理,以便提供与用户意图匹配的服务,并且可以有效地保护用户的隐私。

[0222] 可以将本公开的一些实施例实现为包括存储在计算机可读存储介质中的指令的软件程序。

[0223] 例如,计算机是能够从存储介质调用存储的指令并根据调用的指令执行实施例的操作的设备。计算机可以包括根据实施例的设备、或通信地连接到设备的外部服务器。

[0224] 计算机可读存储介质可以以非暂时性存储介质的形式提供。术语“非暂时性”意味着存储介质不包括信号或电流并且是有形的,并且不区分存储介质中的数据的数据的半永久存储和临时存储。例如,非暂时性存储介质不仅可以包括诸如光盘(CD)、数字通用光盘(DVD)、硬盘、蓝光盘、通用串行总线(USB)、内部存储器、存储卡、ROM或RAM的非暂时性可读存储介质,还可以包括诸如寄存器、高速缓存或缓冲器的临时存储介质。

[0225] 此外,可以将根据实施例的方法提供为计算机程序产品。

[0226] 计算机程序产品可以包括软件程序、其上存储有软件程序的计算机可读存储介质、或在卖方和买方之间交易的产品。

[0227] 例如,计算机程序产品可以包括通过设备制造商或电子市场(例如,Google Play Store、App Store等)电子地分发的软件程序形式的产品(例如,可下载的app)。对于电子分发,软件程序的至少一部分可以存储在存储介质中或者可以临时生成。在这种情况下,存储介质可以是制造商或电子市场的服务器,或者是中继服务器的存储介质。

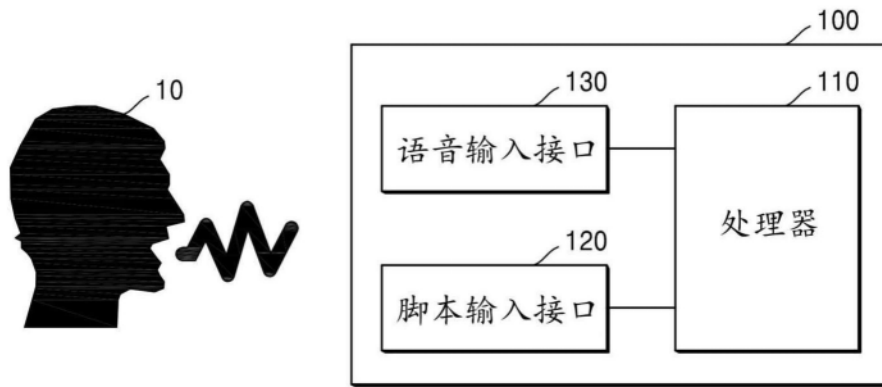


图1

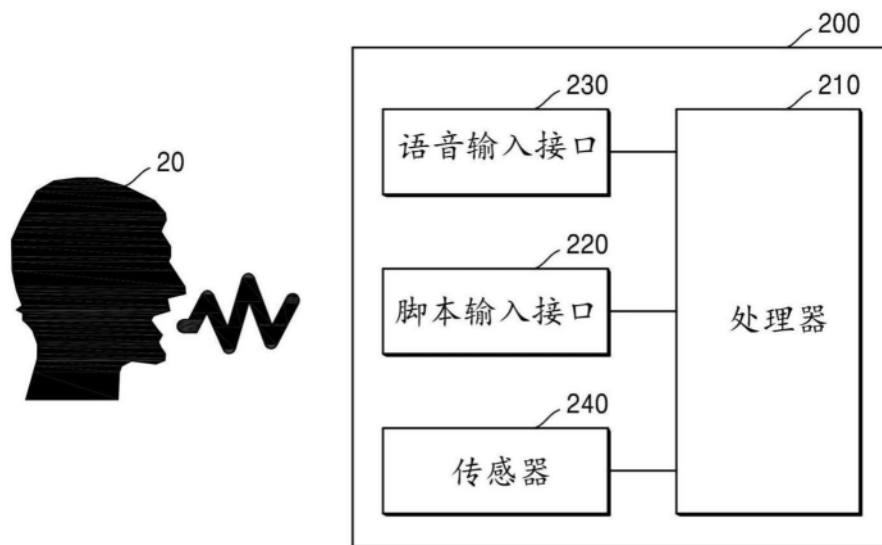


图2

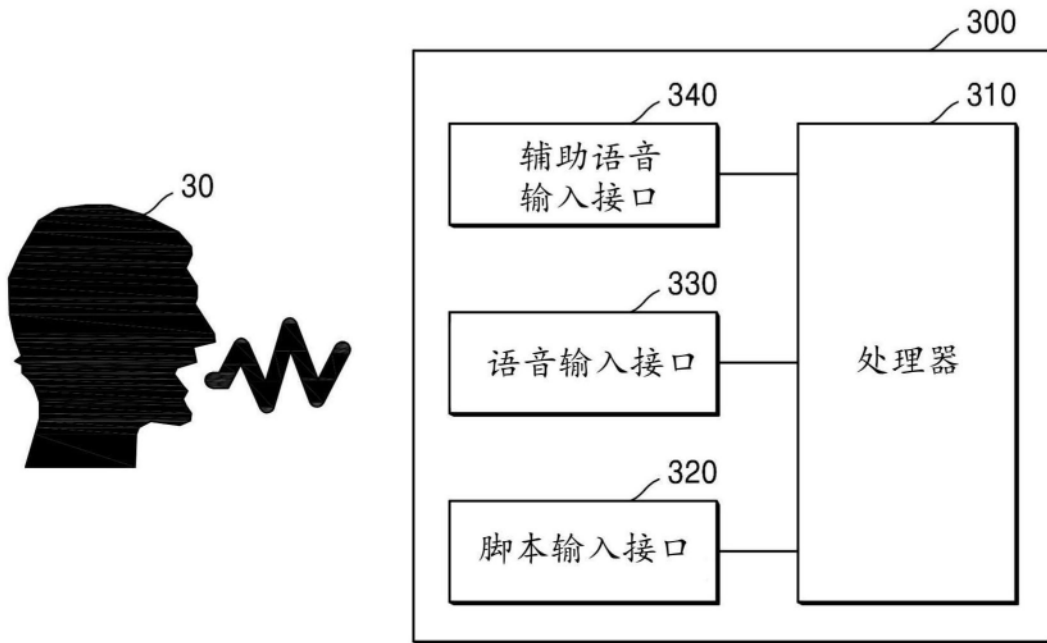


图3

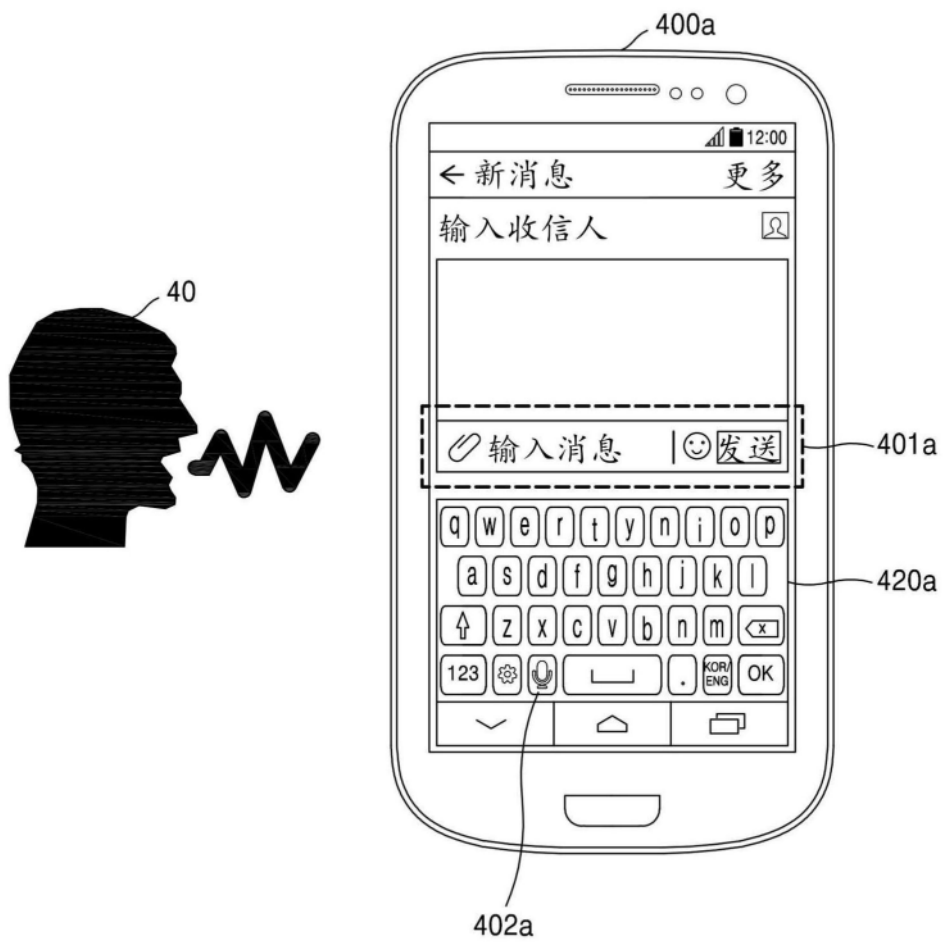


图4A

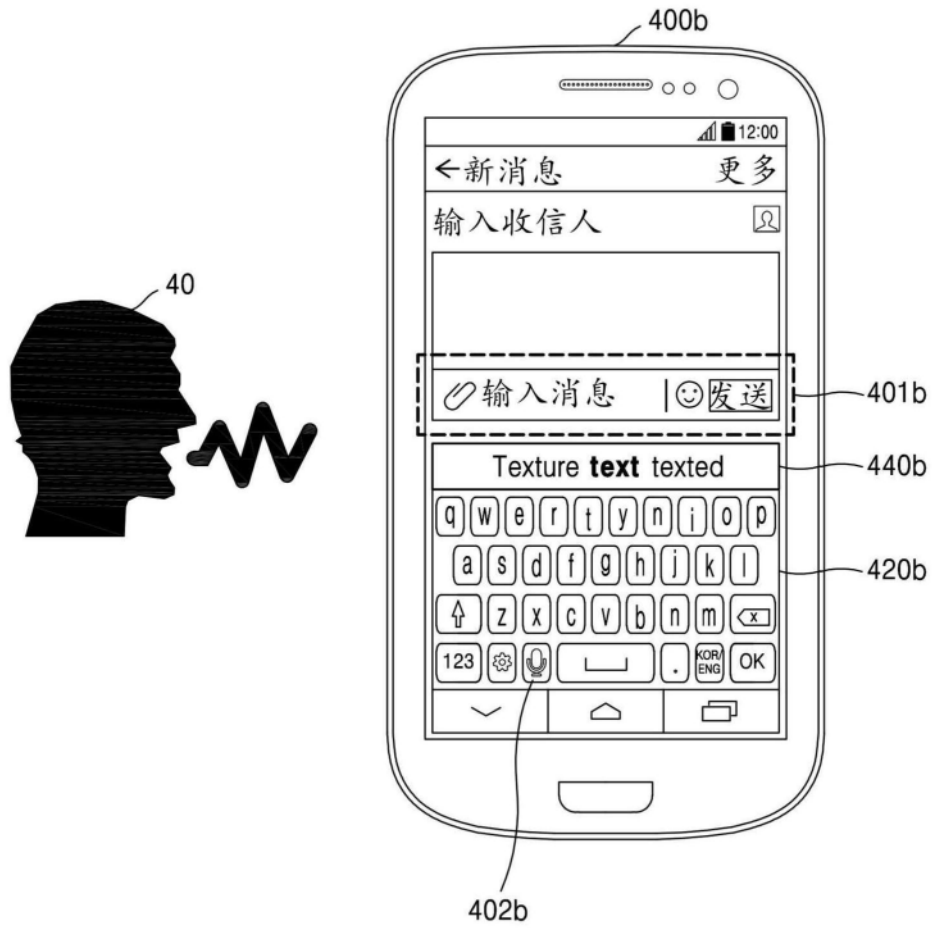


图4B

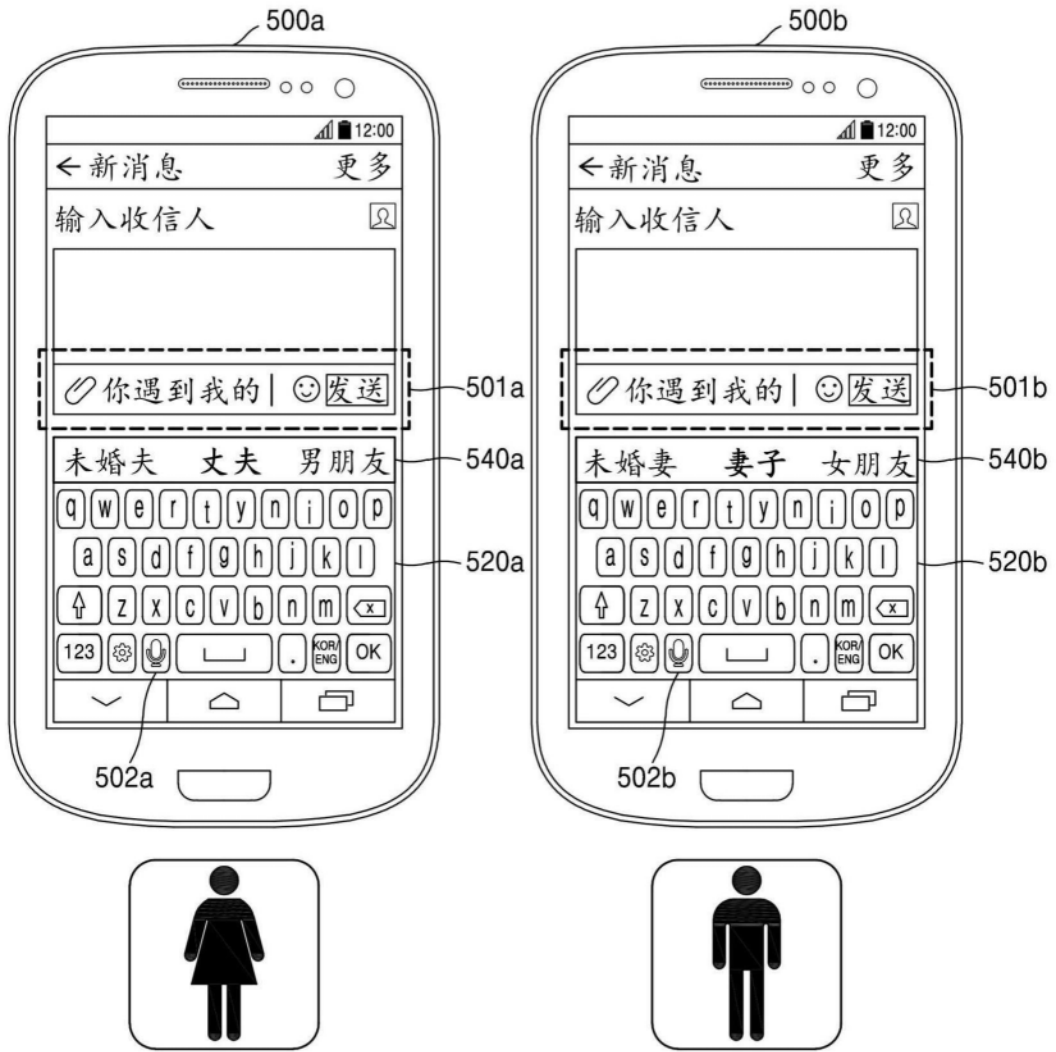


图5

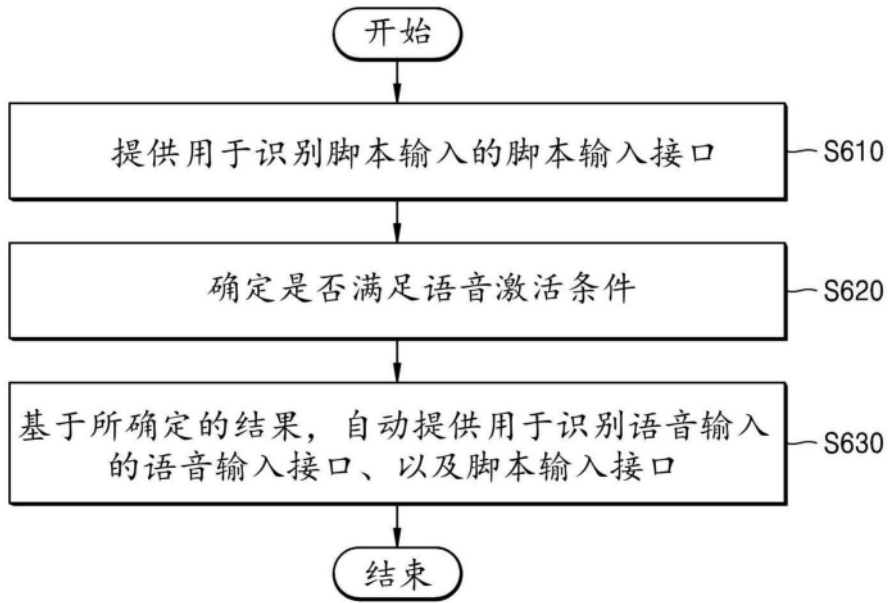


图6

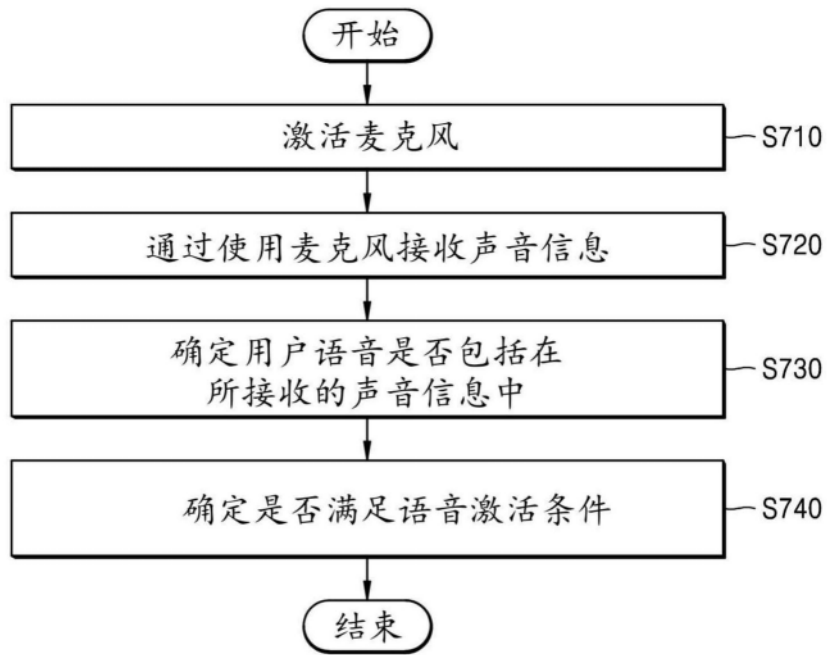


图7

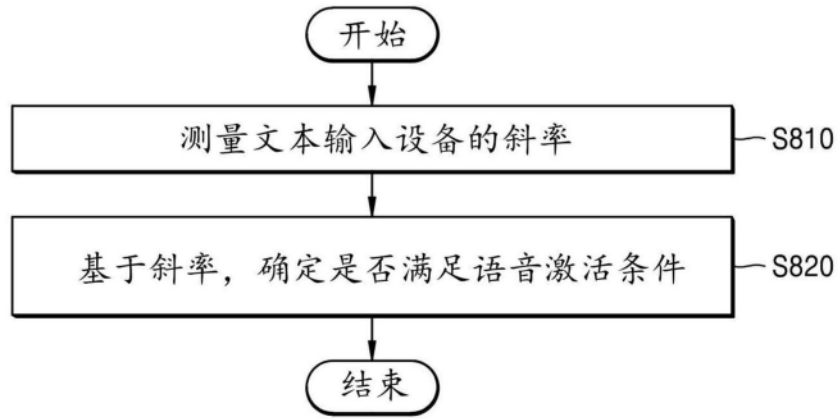


图8

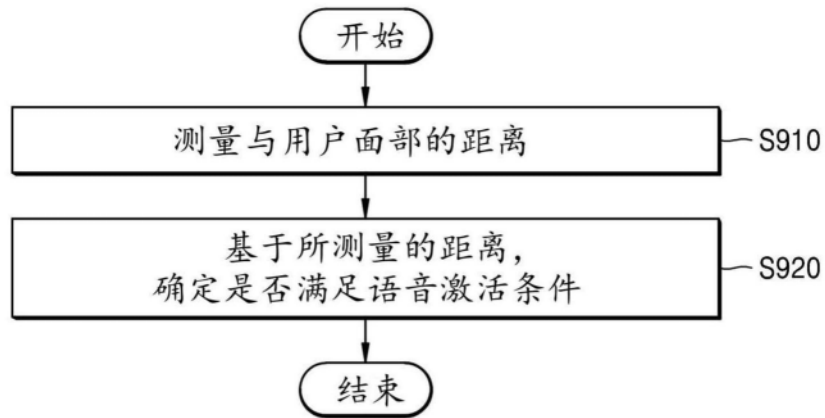


图9

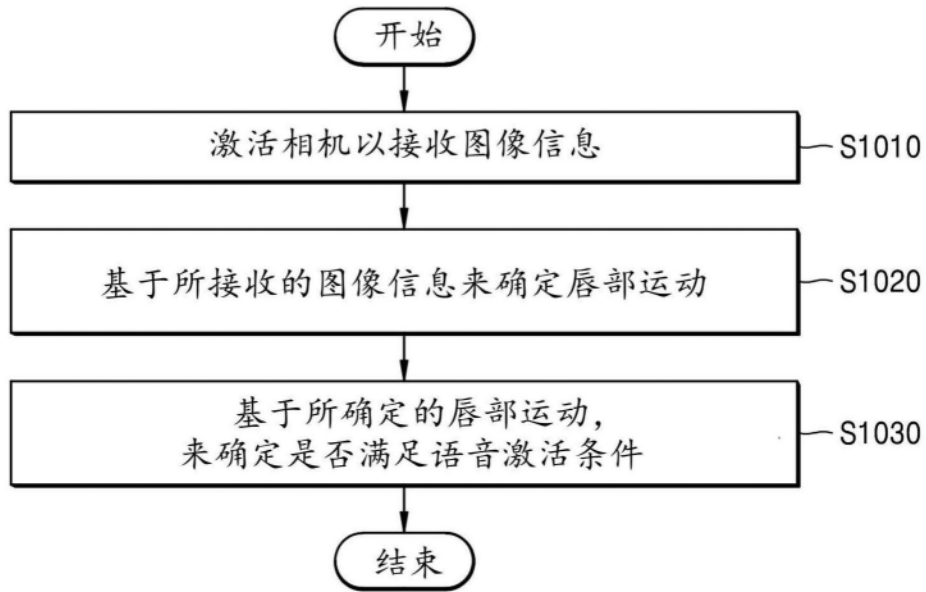


图10

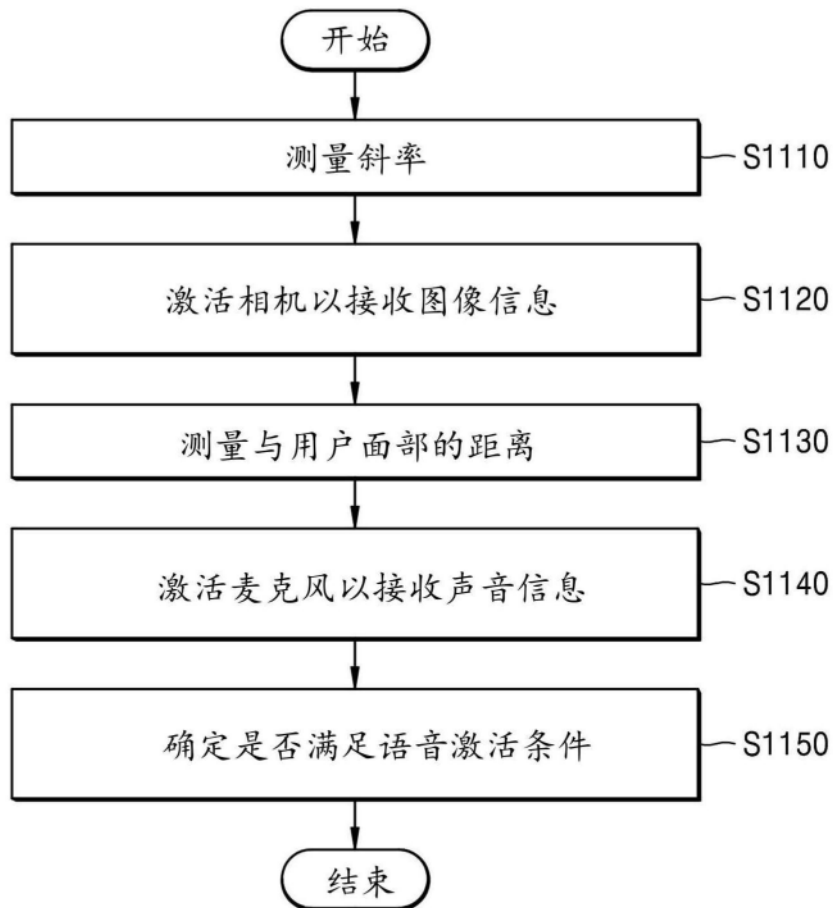


图11

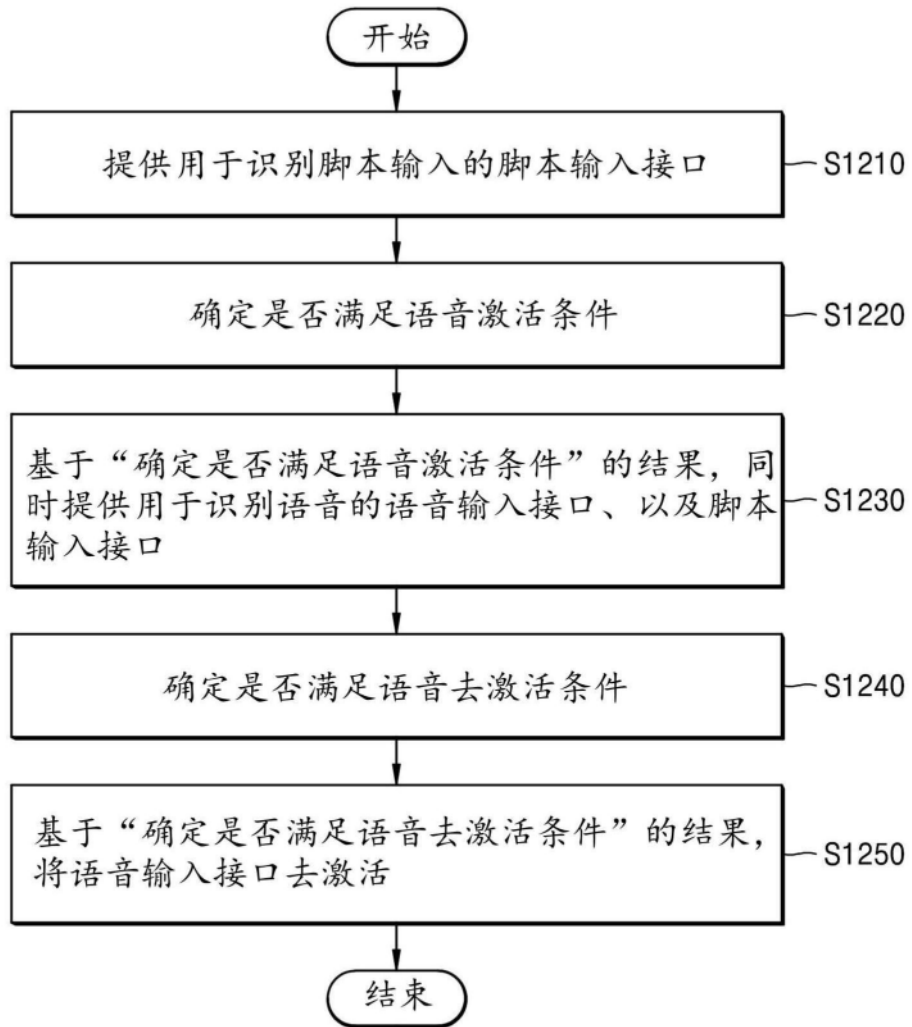


图12

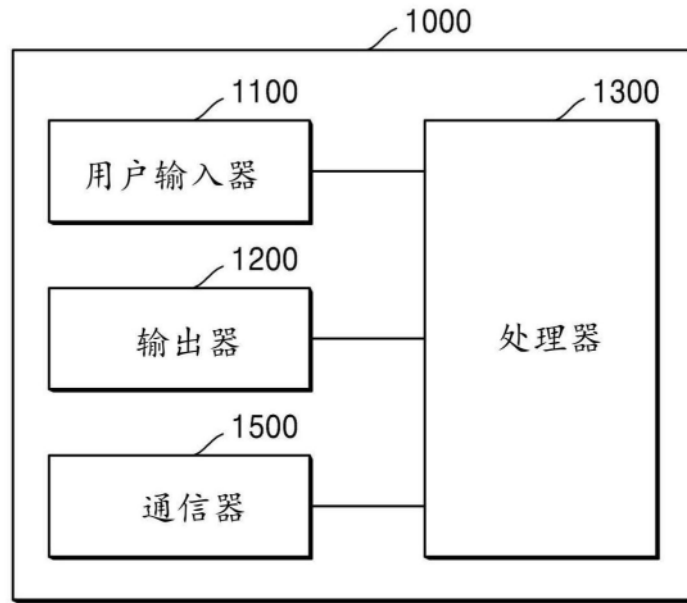


图13

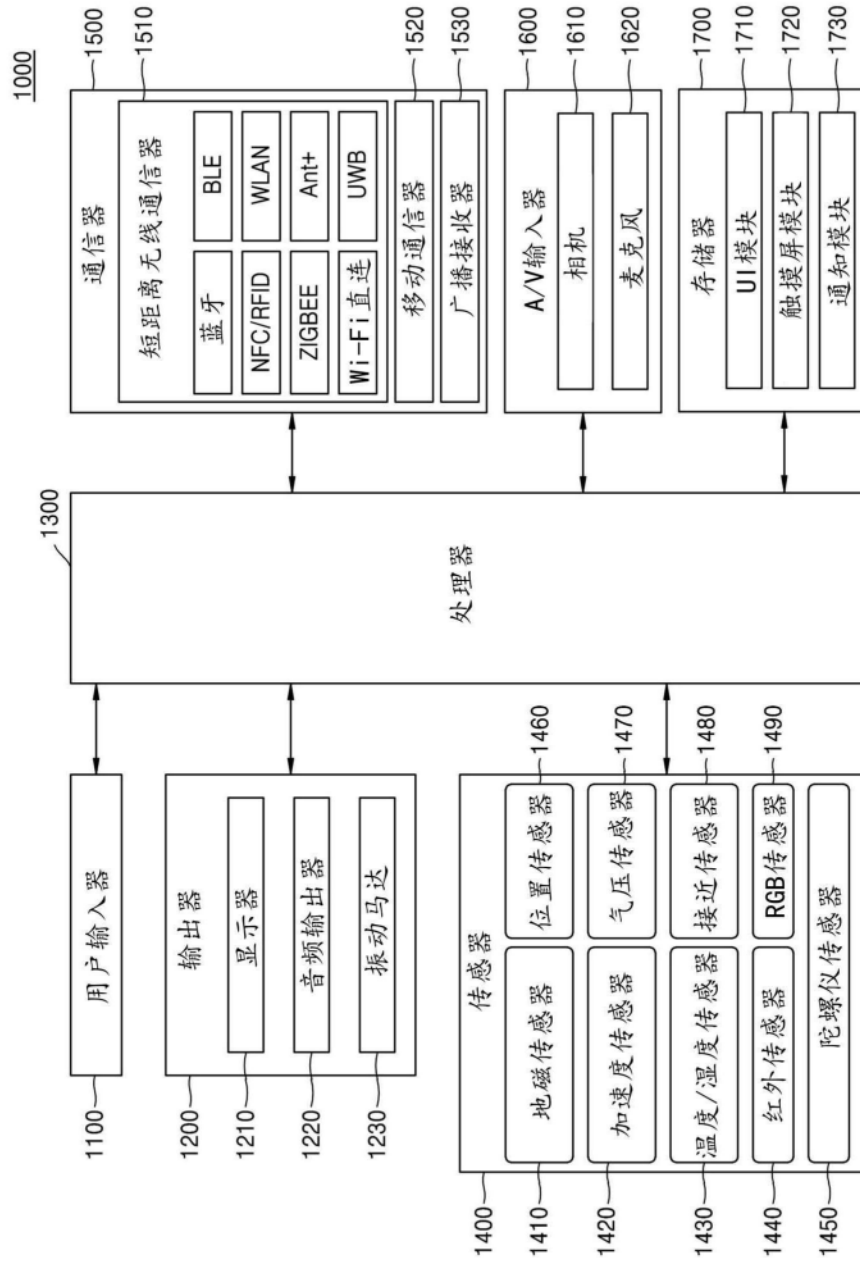


图14

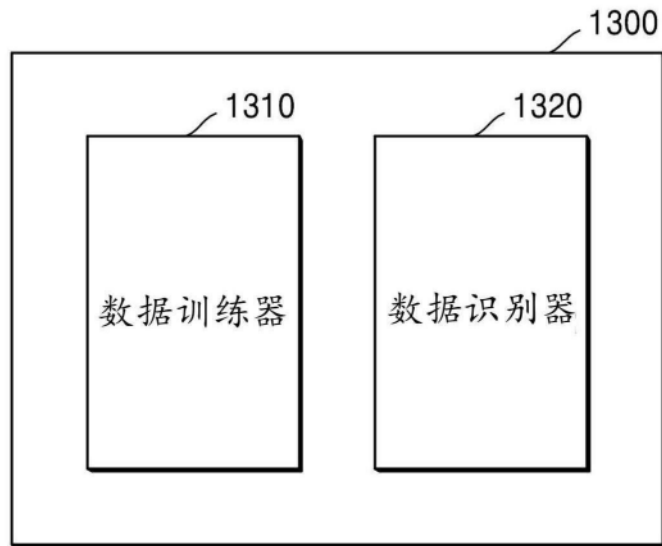


图15

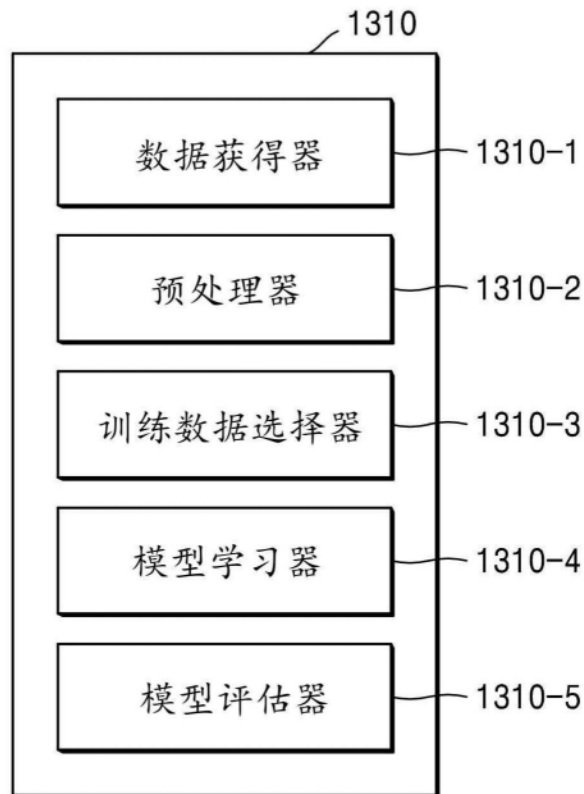


图16

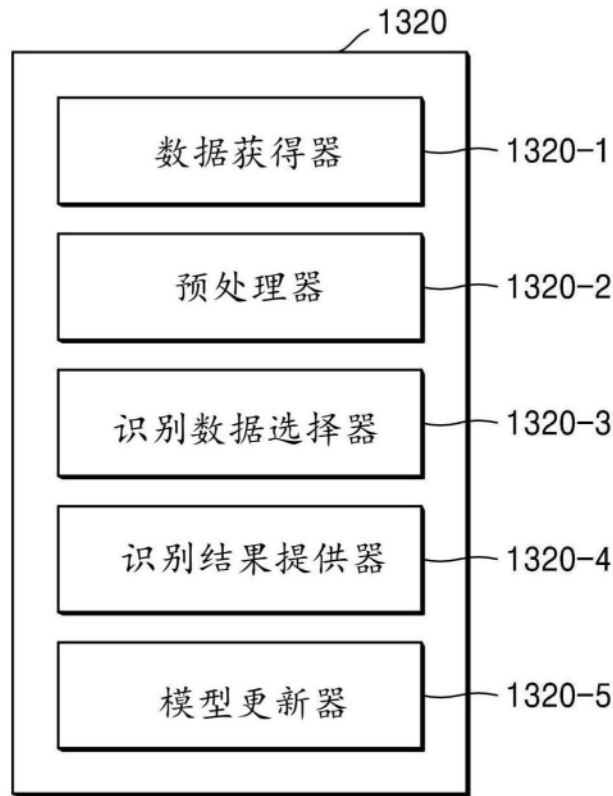


图17

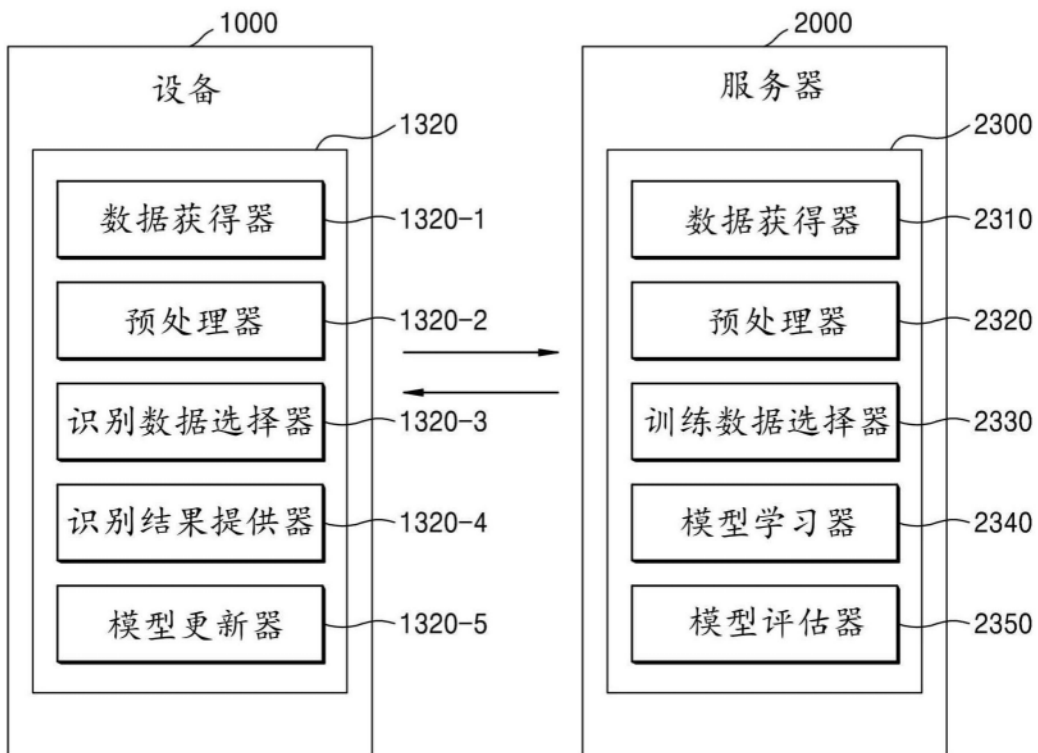


图18