



(12)发明专利申请

(10)申请公布号 CN 111354360 A

(43)申请公布日 2020.06.30

(21)申请号 202010187179.1

(22)申请日 2020.03.17

(71)申请人 北京百度网讯科技有限公司  
地址 100085 北京市海淀区上地十街10号  
百度大厦2层

(72)发明人 何桂

(74)专利代理机构 北京同立钧成知识产权代理有限公司 11205  
代理人 张娜 刘芳

(51)Int.Cl.  
G10L 15/22(2006.01)  
G10L 15/08(2006.01)

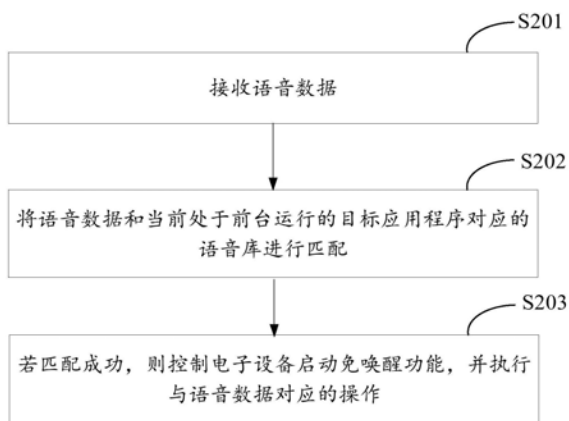
权利要求书2页 说明书10页 附图4页

(54)发明名称

语音交互处理方法、装置和电子设备

(57)摘要

本申请公开了一种语音交互处理方法、装置和电子设备,涉及语音技术领域。具体实现方案为:通过接收语音数据,并将该语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,其中,该语音库中包括多个与目标应用程序关联的控制命令;若匹配成功,则控制电子设备启动免唤醒功能,并执行与语音数据对应的操作。可以看出,本申请实施例中将接收到的语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,若匹配成功,则该电子设备将不需要唤醒,直接执行与语音数据对应的操作,这样,电子设备可以及时为用户提供相应的服务,从而可以提高用户的体验。



1. 一种语音交互处理方法,其特征在于,包括:  
接收语音数据;  
将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,所述语音库中包括多个与所述目标应用程序关联的控制命令;  
若匹配成功,则控制电子设备启动免唤醒功能,并执行与所述语音数据对应的操作。
2. 根据权利要求1所述的方法,其特征在于,所述方法还包括:  
若监测到所述目标应用程序关闭或者处于后台运行状态,则控制所述电子设备退出免唤醒功能。
3. 根据权利要求1所述的方法,其特征在于,所述将所述语音数据和所述目标应用程序对应的语音库进行匹配之前,所述方法还包括:  
建立所述目标应用程序对应的语音库。
4. 根据权利要求1-3任一项所述的方法,其特征在于,所述方法还包括:  
接收更新指令,所述更新指令中包括目标应用程序的标识信息和新的控制命令;  
根据所述标识信息和所述新的控制命令,对所述目标应用程序的语音库进行更新。
5. 根据权利要求1-3任一项所述的方法,其特征在于,将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,包括:  
确定目标对象的朝向;  
若所述目标对象朝向所述电子设备,则将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配。
6. 根据权利要求5所述的方法,其特征在于,所述若所述目标对象朝向所述电子设备,则将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,包括:  
若所述目标对象朝向所述电子设备,则确定所述目标对象与所述电子设备之间的距离是否小于预设值;  
若所述目标对象与所述电子设备之间的距离小于所述预设值,则将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配。
7. 根据权利要求1-3任一项所述的方法,其特征在于,所述控制电子设备启动免唤醒功能,包括:  
控制所述电子设备关闭唤醒引擎,并打开识别引擎。
8. 一种语音交互处理装置,其特征在于,包括:  
接收模块,用于接收语音数据;  
处理模块,用于将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,所述语音库中包括多个与所述目标应用程序关联的控制命令;  
所述处理模块,还用于在匹配成功时,控制所述语音交互处理装置启动免唤醒功能,并执行与所述语音数据对应的操作。
9. 根据权利要求8所述的装置,其特征在于,所述处理模块,还用于在监测到所述目标应用程序关闭或者处于后台运行状态时,控制所述语音交互处理装置退出免唤醒功能。
10. 根据权利要求8所述的装置,其特征在于,所述处理模块,还用于建立所述目标应用程序对应的语音库。
11. 根据权利要求8-10任一项所述的装置,其特征在于,

所述接收模块,还用于接收更新指令,所述更新指令中包括目标应用程序的标识信息和新的控制命令;

所述处理模块,还用于根据所述标识信息和所述新的控制命令,对所述目标应用程序的语音库进行更新。

12. 根据权利要求8-10任一项所述的装置,其特征在于,所述处理模块,具体用于:  
确定目标对象的朝向;

若所述目标对象朝向所述语音交互处理装置,则将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配。

13. 根据权利要求12所述的装置,其特征在于,所述处理模块,具体用于:

若所述目标对象朝向所述语音交互处理装置,则确定所述目标对象与所述语音交互处理装置之间的距离是否小于预设值;

若所述目标对象与所述语音交互处理装置之间的距离小于所述预设值,则将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配。

14. 根据权利要求8-10任一项所述的装置,其特征在于,所述处理模块,具体用于:  
控制所述语音交互处理装置关闭唤醒引擎,并打开识别引擎。

15. 一种电子设备,其特征在于,包括:

至少一个处理器;以及

与所述至少一个处理器通信连接的存储器;其中,

所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行权利要求1-7中任一项所述的方法。

16. 一种存储有计算机指令的非瞬时计算机可读存储介质,其特征在于,所述计算机指令用于使所述计算机执行权利要求1-7中任一项所述的方法。

## 语音交互处理方法、装置和电子设备

### 技术领域

[0001] 本申请涉及数据技术领域,尤其涉及语音技术领域。

### 背景技术

[0002] 随着人工智能技术的不断发展,用户在日常生活中越来越多的通过语音交互方式从电子设备获得服务。

[0003] 目前,大多数电子设备都需要用户对其进行唤醒,在唤醒电子设备之后,用户才能正常发起语音查询(query)请求,即用户与电子设备进行对话。这种用户每次在发起语音查询请求之前,都必须唤醒电子设备的情况,造成用户体验感较差。基于此,现有技术中还提供一种免唤醒的方案:电子设备针对预设的一些免唤醒词采集大量的语料数据,并根据这些语料数据对初始唤醒模型进行不断训练,得到唤醒模型。这样,电子设备在获取到语音数据时,将获取的语音数据与训练得到的唤醒模型进行匹配,在匹配到预设的免唤醒词时,则唤醒电子设备。

[0004] 然而,上述的免唤醒方案中,由于预设的免唤醒词的数量有限,可能会造成在语音数据中不包括预设的免唤醒词时,电子设备无法及时提供相应的服务,造成用户体验不好。

### 发明内容

[0005] 本申请实施例提供一种语音交互处理方法、装置和电子设备,可以使得电子设备能够及时为用户提供相应的服务,由此可以提高用户的体验。

[0006] 第一方面,本申请实施例提供一种语音交互处理方法,该语音交互处理方法可以包括:

[0007] 接收语音数据;

[0008] 将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,所述语音库中包括多个与所述目标应用程序关联的控制命令;

[0009] 若匹配成功,则控制电子设备启动免唤醒功能,并执行与所述语音数据对应的操作。

[0010] 可以看出,与现有技术不同的是,本申请实施例在将接收到的语音数据和当前处于前台运行的目标应用程序对应的语音库匹配成功之后,不需要唤醒电子设备,而是直接对语音数据进行识别,以执行与语音数据对应的操作,这样,不仅可以简化用户的操作,而且电子设备可以及时为用户提供相应的服务,由此可以提高用户的体验。

[0011] 第二方面,本申请实施例还提供一种语音交互处理装置,该装置包括:

[0012] 接收模块,用于接收语音数据;

[0013] 处理模块,用于将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,所述语音库中包括多个与所述目标应用程序关联的控制命令;

[0014] 所述处理模块,还用于在匹配成功时,控制所述语音交互处理装置启动免唤醒功能,并执行与所述语音数据对应的操作。

[0015] 第三方面,本申请实施例还提供一种电子设备,该电子设备可以包括:

[0016] 至少一个处理器;以及与所述至少一个处理器通信连接的存储器;其中,所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行上述第一方面任一种可能的实现方式中所述的语音交互处理方法。

[0017] 第四方面,本申请实施例还提供一种存储有计算机指令的非瞬时计算机可读存储介质,所述计算机指令用于使所述计算机执行上述第一方面任一种可能的实现方式中所述的语音交互处理方法。

[0018] 上述申请中的一个实施例具有如下优点或有益效果:在接收到语音数据后,将语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,该语音库中包括多个与目标应用程序关联的控制命令,若匹配成功,则控制电子设备启动免唤醒功能,并执行与语音数据对应的操作。因此,本申请实施例在将接收到的语音数据和当前处于前台运行的目标应用程序对应的语音库匹配成功之后,不需要唤醒电子设备,而是直接对语音数据进行识别,以执行与语音数据对应的操作,这样,不仅可以简化用户的操作,而且电子设备可以及时为用户提供相应的服务,由此可以提高用户的体验。另外,本申请实施例提供的方案,可以避免现有技术中需要采集大量的语料数据对初始唤醒模型进行不断训练,才能得到唤醒模型,由此可以降低语音处理的成本。

[0019] 上述可选方式所具有的其他效果将在下文中结合具体实施例加以说明。

## 附图说明

[0020] 附图用于更好地理解本方案,不构成对本申请的限定。其中:

[0021] 图1为本申请实施例提供的语音交互处理方法的一种应用场景示意图;

[0022] 图2是本申请实施例一提供的语音交互处理方法的流程示意图;

[0023] 图3是本申请实施例二提供的语音交互处理方法的流程示意图;

[0024] 图4为语音库进行更新的示意图;

[0025] 图5是本申请实施例三提供的语音交互处理方法的流程示意图;

[0026] 图6是本申请实施例提供的语音交互处理装置60的结构示意图;

[0027] 图7是用来实现本申请实施例的语音交互处理方法的电子设备的框图。

## 具体实施方式

[0028] 以下结合附图对本申请的示范性实施例做出说明,其中包括本申请实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本申请的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0029] 在本申请的实施例中,“至少一个”是指一个或者多个,“多个”是指两个或两个以上。“和/或”,描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B这三种情况,其中A,B可以是单数或者复数。在本申请的文字描述中,字符“/”一般表示前后关联对象是一种“或”的关系。

[0030] 在本申请的实施例中,电子设备可以为支持信息通信,并能够对语音数据进行处

理的各种电子设备,包括但不限于智能手机、平板电脑、便携式计算机和台式计算机等。

[0031] 为了通过语音交互方式从电子设备获得服务,首先需要将电子设备唤醒,但是这种方式造成用户的操作较为繁琐,使得用户的体验不好。基于此,目前还有一种方式是电子设备针对预设的一些免唤醒词采集大量的语料数据,并根据这些语料数据对初始唤醒模型进行不断训练,得到唤醒模型。这样,电子设备在获取到语音数据时,将获取的语音数据与训练得到的唤醒模型进行匹配,在匹配到预设的免唤醒词时,则唤醒电子设备。但是,由于预设的免唤醒词的数量是有限的,而且每个人想从电子设备中获得的相关服务可能会不一样,这样,就会造成用户在使用电子设备获得相关服务,且输入电子设备的语音数据中不包括上述预设的免唤醒词时,电子设备将无法及时提供相应的服务,从而造成用户的体验不好。

[0032] 为了使得电子设备能够根据用户的需求,及时为用户提供相应的服务,可以由用户自定义一些免唤醒词,也即用户根据自己的需求,在电子设备中预先存储一些免唤醒词,以达到电子设备免唤醒的目的。但是采用该方法,虽然在一定程度上可以提高电子设备在免唤醒的情况下,为用户提供相关服务的及时性,但是,由于电子设备能够提供的服务经常会发生变化,在电子设备提供的服务发生变化时,用户就需要重新设置一些免唤醒词,使得用户操作较为繁琐,导致用户体验不高。

[0033] 基于上述论述,为了使得电子设备能够及时为用户提供相应的服务,以提高用户的体验,本申请实施例提供了一种语音交互处理方法,在接收到语音数据后,将语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,该语音库中包括多个与目标应用程序关联的控制命令,若匹配成功,则控制电子设备启动免唤醒功能,并执行与语音数据对应的操作。可以看出,与现有技术不同的是,本申请实施例在将接收到的语音数据和当前处于前台运行的目标应用程序对应的语音库匹配成功之后,不需要唤醒电子设备,而是直接对语音数据进行识别,以执行与语音数据对应的操作,这样,不仅可以简化用户的操作,而且电子设备可以及时为用户提供相应的服务,由此可以提高用户的体验。另外,本申请实施例提供的方案,可以避免现有技术中需要采集大量的语料数据对初始唤醒模型进行不断训练,才能得到唤醒模型,由此可以降低语音处理的成本。

[0034] 可以理解的是,本申请实施例提供的语音交互处理方法可以应用于语音识别的场景中,尤其应用于通过免唤醒方式控制电子设备,以获得相关服务的场景中。下面,以通过语音方式控制电子设备播放某个电影场景为例,详细说明如何进行语音交互处理。

[0035] 图1为本申请实施例提供的语音交互处理方法的一种应用场景示意图,如图1所示,当用户想要通过电子设备播放某个电影时,用户可以向电子设备输入“播放电影C”的语音数据,电子设备在接收到该语音数据后,会确定当前处于前台运行的目标应用程序(application,APP),若确定出当前处于前台运行的目标应用程序为某个视频类应用程序,则查询“播放电影C”是否在该视频类应用程序的语音库中。假设“播放电影C”在该视频类应用程序的语音库中,则电子设备将不需要唤醒电子设备,也即将会关闭唤醒引擎,直接打开识别引擎,对接收到的“播放电影C”进行识别,以开始播放电影C。这样,通过免唤醒方式即可为用户提供相应的服务,避免了现有技术中用户输入的语音数据中必须包含预设的免唤醒词才能使终端提供服务的现象,由此可以提高电子设备提供服务的及时性,从而可以提高用户的体验。

[0036] 下面,将通过具体的实施例对本申请提供的语音交互处理方法进行详细地说明。可以理解的是,下面这几个具体的实施例可以相互结合,对于相同或相似的概念或过程可能在某些实施例不再赘述。

[0037] 图2是本申请实施例一提供的语音交互处理方法的流程示意图,该语音交互处理方法可以由软件和/或硬件装置执行,例如,该硬件装置可以为语音交互处理装置,该语音交互处理装置可以设置在电子设备或者车载终端中。示例的,请参见图2所示,该语音交互处理方法可以包括:

[0038] S201、接收语音数据。

[0039] S202、将语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配。

[0040] 其中,语音库中包括多个与目标应用程序关联的控制命令。

[0041] 在本步骤中,电子设备在接收到用户输入的语音数据之后,可以确定当前在电子设备中,处于前台运行的目标应用程序,并将接收到的语音数据和该目标应用程序的语音库进行匹配,也即查询该语音数据是否在目标应用程序的语音库中。

[0042] 其中,目标应用程序可以包括能够为用户提供服务的各种平台,例如,可以为视频类应用程序、图像处理类应用程序、即时通信类应用程序和导航应用程序等等。

[0043] S203、若匹配成功,则控制电子设备启动免唤醒功能,并执行与语音数据对应的操作。

[0044] 在本步骤中,若语音数据和目标应用程序对应的语音库匹配成功,也即目标应用程序对应的语音库中存在该语音数据时,则控制电子设备启动免唤醒功能,也即不需要唤醒电子设备,直接对接收到的语音数据进行识别,以执行与该语音数据对应的操作。

[0045] 在本实施例中,由于电子设备在将接收到的语音数据和当前处于前台运行的目标应用程序对应的语音库匹配成功后,即可控制电子设备执行与语音数据对应的操作,从而避免了现有技术中,在电子设备被唤醒之后,用户只能在预设时间段内与电子设备进行交互,在预设时间段之后,需要重新唤醒电子设备的现象,从而实现了用户与电子设备之间的多次交互。

[0046] 本申请实施例提供的语音交互处理方法,在接收到语音数据后,将语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,该语音库中包括多个与目标应用程序关联的控制命令,若匹配成功,则控制电子设备启动免唤醒功能,并执行与语音数据对应的操作。因此,本申请实施例在将接收到的语音数据和当前处于前台运行的目标应用程序对应的语音库匹配成功之后,不需要唤醒电子设备,而是直接对语音数据进行识别,以执行与语音数据对应的操作,这样,不仅可以简化用户的操作,而且电子设备可以及时为用户提供相应的服务,由此可以提高用户的体验。另外,本申请实施例提供的方案,可以避免现有技术中需要采集大量的语料数据对初始唤醒模型进行不断训练,才能得到唤醒模型,由此可以降低语音处理的成本。

[0047] 进一步地,对于电子设备中运行的应用程序,需要预先建立与其对应的语音库,该语音库中包括有多个与该应用程序关联或者相关的控制命令。示例性的,该语音库可以为预先注册的,或者可以为应用程序开发者预先输入并存储的。

[0048] 需要进行说明的是,不同的应用程序对应的语音库中的控制命令,可能相同,也可能不同。例如:对于视频类应用程序,其对应的语音库中可能存储有“播放电影A”、“播放”、

“暂停”、“下一页”和“上一页”等等。对于图像类应用程序,其对应的语音库中可能存储有“查看图片”、“删除”、“放大”、“下一页”和“上一页”等等。

[0049] 在本实施例中,可以根据应用程序能够支持的服务建立语音库,由此可以提高语音数据匹配的准确性。另外,用户还可以根据需要提供注册支持免唤醒功能的应用程序列表,从而提高了语音交互处理的灵活性。

[0050] 图3是本申请实施例二提供的语音交互处理方法的流程示意图,本实施例在上述各实施例的基础上,对目标应用程序对应的语音库进行更新的过程进行详细说明。如图3所示,该语音交互处理方法可以包括:

[0051] S301:接收更新指令,该更新指令中包括目标应用程序的标识信息和新的控制命令。

[0052] 在本步骤中,由于电子设备能够提供的服务经常会发生变化,如视频类应用程序中的视频经常会改变,为了能够提高语音数据和语音库匹配的精确度,需要经常对应用程序对应的语音库进行更新。

[0053] 具体地,用户可以通过语音方式向电子设备发送更新指令,也可以通过触摸某个按键或者按压某个物理按键的方式发送该更新指令,对于更新指令的具体发送方式,本申请实施例不做限制。

[0054] S302:根据标识信息和新的控制命令,对目标应用程序的语音库进行更新。

[0055] 在本步骤中,电子设备接收到的更新指令中包括有目标应用程序的标识信息和新的控制命令,电子设备可以根据该标识信息,确定出是对该标识信息对应的目标应用程序的语音库进行更新,然后根据新的控制命令,对目标应用程序的语音库进行更新。其中,更新包括将新的控制命令添加在语音库中,或者根据新的控制命令,删除语音库中的控制命令等。

[0056] 在一种可能的实现方式中,在对语音库进行更新时,电子设备可以先查询语音库中是否已经包含了该新的控制命令,若语音库中已经包含了该新的控制命令,则将不对该语音库做更新处理。

[0057] 图4为语音库进行更新的示意图,如图4所示,假设用户通过语音方式向电子设备发送更新指令,如用户向电子设备输入“在应用程序A中添加播放视频1”,电子设备根据该更新指令,将可以获知是对应用程序A的语音库进行更新,而且需要在该语音库中添加控制命令“播放视频1”。电子设备可以向用户显示添加确认页面,如图4所示,若用户点击“确定”,则电子设备可以将控制命令“播放视频1”添加在应用程序A对应的语音库中。其中,电子设备在添加该控制命名之前,可以先判断应用程序A对应的语音库中是否已经存在“播放视频1”,若存在,则不会对语音库进行更新,若不存在,则将该控制命令添加在语音库中。

[0058] 在另一种可能的实现方式中,电子设备还可以根据目标应用程序能够提供的服务,周期性的更新其对应的语音库。例如:对于视频类应用程序,可以根据更新的视频或者更新的视频处理功能,更新语音库中的控制命令,对于图片处理类应用程序,可以根据更新的图片或者更新的图片处理功能,更新语音库中的控制命令等等。

[0059] 在本实施例中,由于电子设备在接收到更新指令之后,会根据该更新指令对语音库中的控制命令进行更新,从而可以提高语音数据和语音库进行匹配的准确性,而且使得建立的语音库中的控制命令更丰富。



[0060] 图5是本申请实施例三提供的语音交互处理方法的流程示意图,本实施例在上述各实施例的基础上,对将语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配的过程进行详细说明。如图5所示,该语音交互处理方法可以包括:

[0061] S501:接收语音数据。

[0062] S501和S201的执行过程类似,此处不再赘述。

[0063] S502:确定目标对象的朝向。

[0064] S503:若目标对象朝向电子设备,则将语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配。

[0065] 具体地,在实际应用中,用户有时可能并非想通过电子设备获得相关服务,而仅是在聊天过程中,提到了目标应用程序的语音库中的某个控制命令,例如:用户在聊天过程中,提到“我想看XX电影”,此时,电子设备可能就会误识别,从而执行该控制命令相关的操作。

[0066] 为了解决这一问题,在本申请实施例中,电子设备在将语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配时,可以先确定目标对象的朝向,若目标对象朝向电子设备,则将语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配。

[0067] 具体地,上述目标对象可以为用户,也可以理解为向电子设备输入语音数据的用户。电子设备在接收到用户输入的语音数据后,可以通过声源定位技术,确定用户相对于电子设备的角度,若确定出的角度小于预设值,则可以确定出用户朝向电子设备,说明用户是想通过电子设备获取相关服务。此时,电子设备可以将接收到的语音数据和目标应用程序对应的语音库进行匹配。

[0068] 当然,电子设备还可以通过其他方式确定目标对象的朝向,例如可以根据摄像头拍摄的包含目标对象的图片,并对图片进行识别,从而确定出目标对象是否朝向电子设备等。对于确定目标对象的朝向的具体方式,本申请实施例不做限制。

[0069] 在本实施例中,电子设备通过确定目标对象的朝向,并在目标对象朝向电子设备时,将语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,从而可以降低电子设备误识别的概率,提高用户的体验。

[0070] 为了进一步降低电子设备误识别的概率,在确定出目标对象朝向电子设备后,还可以判断目标对象与电子设备之间的距离是否小于预设值,若判断出目标对象与电子设备之间的距离小于预设值,则将语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配。

[0071] 具体地,上述目标对象可以为用户,在具体的实现过程中,可以通过红外感应的方式确定用户与电子设备之间的距离,若确定出用户与电子设备之间的距离小于预设值,则说明用户目前可能正在操作电子设备,也即用户想要通过电子设备获取相应的服务,此时,电子设备会将接收到的语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配。

[0072] 其中,上述预设值可以根据实际情况或者经验进行设置,例如可以设置为10cm或者15cm等,对于预设值的具体取值,本申请实施例在此不做限制。

[0073] 在本实施例中,可以在目标对象朝向电子设备,且目标对象与电子设备之间的距

离小于预设值时,将接收到的语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,从而可以更好的降低误识别的概率,提高用户的体验。

[0074] S504:若匹配成功,则控制电子设备启动免唤醒功能,并执行与语音数据对应的操作。

[0075] 在一种可能的实现方式中,控制电子设备启动免唤醒功能,可以是控制电子设备关闭唤醒引擎,并打开识别引擎。也即,在语音数据和目标应用程序的语音库匹配成功之后,电子设备将关闭唤醒引擎,此时,电子设备不需要通过唤醒引擎来唤醒,而是通过打开识别引擎,对接收到的语音数据进行识别,以执行该语音数据对应的操作。

[0076] 在本实施例中,在语音数据和目标应用程序的语音库匹配成功后,控制电子设备关闭唤醒引擎,并打开识别引擎,这样,电子设备在免唤醒状态下,可以直接对语音数据进行识别,从而执行与语音数据对应的操作,由此可以提高电子设备为用户提供相关服务的及时性,提高了用户的体验。

[0077] 进一步地,若语音数据和当前处于前台运行的目标应用程序对应的语音库匹配失败,也即目标应用程序对应的语音库中不存在该语音数据时,则可以忽略该语音数据,电子设备将不执行任何操作。

[0078] 本申请实施例提供的语音交互处理方法,在接收到语音数据后,将语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,该语音库中包括多个与目标应用程序关联的控制命令,若匹配成功,则控制电子设备启动免唤醒功能,并执行与语音数据对应的操作。因此,本申请实施例在将接收到的语音数据和当前处于前台运行的目标应用程序对应的语音库匹配成功之后,不需要唤醒电子设备,而是直接对语音数据进行识别,以执行与语音数据对应的操作,这样,不仅可以简化用户的操作,而且电子设备可以及时为用户提供相应的服务,由此可以提高用户的体验。另外,本申请实施例提供的方案,可以避免现有技术中需要采集大量的语料数据对初始唤醒模型进行不断训练,才能得到唤醒模型,由此可以降低语音处理的成本。

[0079] 在上述各实施例的基础上,若电子设备监测到目标应用程序关闭或者处于后台运行状态,则将控制该电子设备退出免唤醒功能。

[0080] 其中,控制该电子设备退出免唤醒功能,可以理解为控制电子设备打开唤醒引擎,也即电子设备需要通过唤醒词进行唤醒之后,才能继续为用户提供相应的服务。

[0081] 在本实施例中,电子设备在监测到目标应用程序关闭或者处于后台运行状态时,将控制该电子设备退出免唤醒功能,由此不仅可以提高电子设备执行相关操作的准确性,而且可以有效的防止误触发。

[0082] 另外,示例性的,电子设备在打开唤醒引擎后,还可以关闭识别引擎,从而可以节省内存,降低功耗。

[0083] 图6是本申请实施例提供的语音交互处理装置60的结构示意图,示例的,请参见图6所示,该语音交互处理装置60可以包括:

[0084] 接收模块601,用于接收语音数据;

[0085] 处理模块602,用于将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,所述语音库中包括多个与所述目标应用程序关联的控制命令;

[0086] 所述处理模块602,还用于在匹配成功时,控制所述语音交互处理装置启动免唤醒

功能,并执行与所述语音数据对应的操作。

[0087] 可选地,所述处理模块602,还用于在监测到所述目标应用程序关闭或者处于后台运行状态时,控制所述语音交互处理装置退出免唤醒功能。

[0088] 可选地,所述处理模块602,还用于建立所述目标应用程序对应的语音库。

[0089] 可选地,所述接收模块601,还用于接收更新指令,所述更新指令中包括目标应用程序的标识信息和新的控制命令;

[0090] 所述处理模块602,还用于根据所述标识信息和所述新的控制命令,对所述目标应用程序的语音库进行更新。

[0091] 可选地,所述处理模块602,具体用于:

[0092] 确定目标对象的朝向;

[0093] 若所述目标对象朝向所述语音交互处理装置,则将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配。

[0094] 可选地,所述处理模块602,具体用于:

[0095] 若所述目标对象朝向所述语音交互处理装置,则确定所述目标对象与所述语音交互处理装置之间的距离是否小于预设值;

[0096] 若所述目标对象与所述语音交互处理装置之间的距离小于所述预设值,则将所述语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配。

[0097] 可选地,所述处理模块602,具体用于:

[0098] 控制所述电子设备关闭唤醒引擎,并打开识别引擎。

[0099] 本申请实施例提供的语音交互处理装置60,可以执行上述任一实施例中的语音交互处理方法的技术方案,其实现原理以及有益效果与语音交互处理方法的实现原理及有益效果类似,可参见语音交互处理方法的实现原理及有益效果,此处不再进行赘述。

[0100] 根据本申请的实施例,本申请还提供了一种电子设备和一种可读存储介质。

[0101] 如图7所示,是根据本申请实施例的语音交互处理方法的电子设备的框图。电子设备旨在表示各种形式的数字计算机,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例,并且不意在限制本文中描述的和/或者要求的本申请的实现。

[0102] 如图7所示,该电子设备包括:一个或多个处理器701、存储器702,以及用于连接各部件的接口,包括高速接口和低速接口。各个部件利用不同的总线互相连接,并且可以被安装在公共主板上或者根据需要以其它方式安装。处理器可以对在电子设备内执行的指令进行处理,包括存储在存储器中或者存储器上以在外部输入/输出装置(诸如,耦合至接口的显示设备)上显示GUI的图形信息的指令。在其它实施方式中,若需要,可以将多个处理器和/或多条总线与多个存储器和多个存储器一起使用。同样,可以连接多个电子设备,各个设备提供部分必要的操作(例如,作为服务器阵列、一组刀片式服务器、或者多处理器系统)。图7中以一个处理器701为例。

[0103] 存储器702即为本申请所提供的非瞬时计算机可读存储介质。其中,所述存储器存储有可由至少一个处理器执行的指令,以使所述至少一个处理器执行本申请所提供的语音

交互处理方法。本申请的非瞬时计算机可读存储介质存储计算机指令,该计算机指令用于使计算机执行本申请所提供的语音交互处理方法。

[0104] 存储器702作为一种非瞬时计算机可读存储介质,可用于存储非瞬时软件程序、非瞬时计算机可执行程序以及模块,如本申请实施例中的语音交互处理方法对应的程序指令/模块(例如,附图6所示的接收模块601和处理模块602)。处理器701通过运行存储在存储器702中的非瞬时软件程序、指令以及模块,从而执行服务器的各种功能应用以及数据处理,即实现上述方法实施例中的语音交互处理方法。

[0105] 存储器702可以包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需要的应用程序;存储数据区可存储根据语音交互处理方法的电子设备的使用所创建的数据等。此外,存储器702可以包括高速随机存取存储器,还可以包括非瞬时存储器,例如至少一个磁盘存储器件、闪存器件、或其他非瞬时固态存储器件。在一些实施例中,存储器702可选包括相对于处理器701远程设置的存储器,这些远程存储器可以通过网络连接至语音交互处理方法的电子设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0106] 语音交互处理方法的电子设备还可以包括:输入装置703和输出装置704。处理器701、存储器702、输入装置703和输出装置704可以通过总线或者其他方式连接,图7中以通过总线连接为例。

[0107] 输入装置703可接收输入的数字或字符信息,以及产生与语音交互处理方法的电子设备的用户设置以及功能控制有关的键信号输入,例如触摸屏、小键盘、鼠标、轨迹板、触摸板、指示杆、一个或者多个鼠标按钮、轨迹球、操纵杆等输入装置。输出装置704可以包括显示设备、辅助照明装置(例如,LED)和触觉反馈装置(例如,振动电机)等。该显示设备可以包括但不限于,液晶显示器(LCD)、发光二极管(LED)显示器和等离子体显示器。在一些实施方式中,显示设备可以是触摸屏。

[0108] 此处描述的系统和技术各种实施方式可以在数字电子电路系统、集成电路系统、专用ASIC(专用集成电路)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0109] 这些计算程序(也称作程序、软件、软件应用、或者代码)包括可编程处理器的机器指令,并且可以利用高级过程和/或面向对象的编程语言、和/或汇编/机器语言来实施这些计算程序。如本文使用的,术语“机器可读介质”和“计算机可读介质”指的是用于将机器指令和/或数据提供给可编程处理器的任何计算机程序产品、设备、和/或装置(例如,磁盘、光盘、存储器、可编程逻辑装置(PLD)),包括,接收作为机器可读信号的机器指令的机器可读介质。术语“机器可读信号”指的是用于将机器指令和/或数据提供给可编程处理器的任何信号。

[0110] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视

器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0111] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)和互联网。

[0112] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。

[0113] 根据本申请实施例的技术方案,在接收到语音数据后,将语音数据和当前处于前台运行的目标应用程序对应的语音库进行匹配,该语音库中包括多个与目标应用程序关联的控制命令,若匹配成功,则控制电子设备启动免唤醒功能,并执行与语音数据对应的操作。因此,本申请实施例在将接收到的语音数据和当前处于前台运行的目标应用程序对应的语音库匹配成功之后,不需要唤醒电子设备,而是直接对语音数据进行识别,以执行与语音数据对应的操作,这样,不仅可以简化用户的操作,而且电子设备可以及时为用户提供相应的服务,由此可以提高用户的体验。另外,本申请实施例提供的方案,可以避免现有技术中需要采集大量的语料数据对初始唤醒模型进行不断训练,才能得到唤醒模型,由此可以降低语音处理的成本。

[0114] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本发申请中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本申请公开的技术方案所期望的结果,本文在此不进行限制。

[0115] 上述具体实施方式,并不构成对本申请保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本申请的精神和原则之内所作的修改、等同替换和改进等,均应包含在本申请保护范围之内。

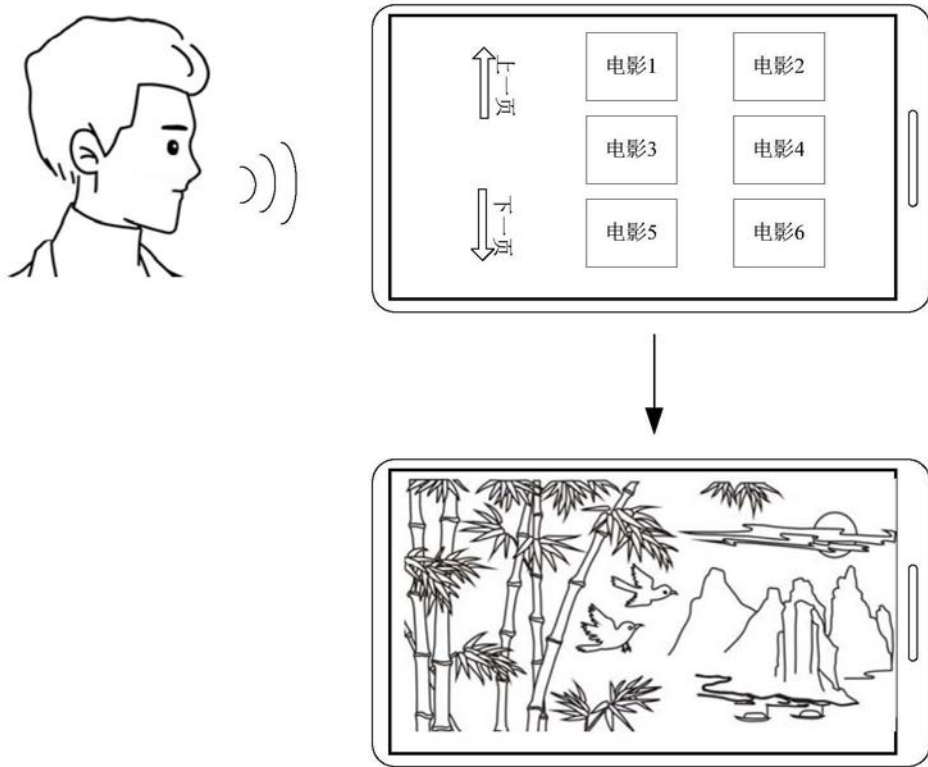


图1

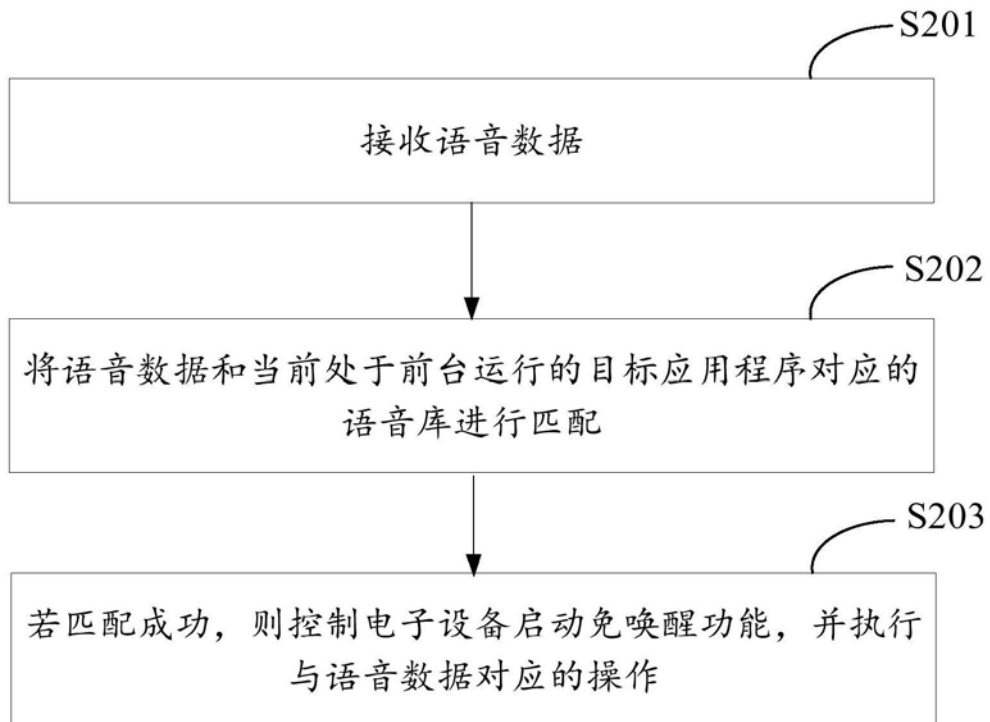


图2

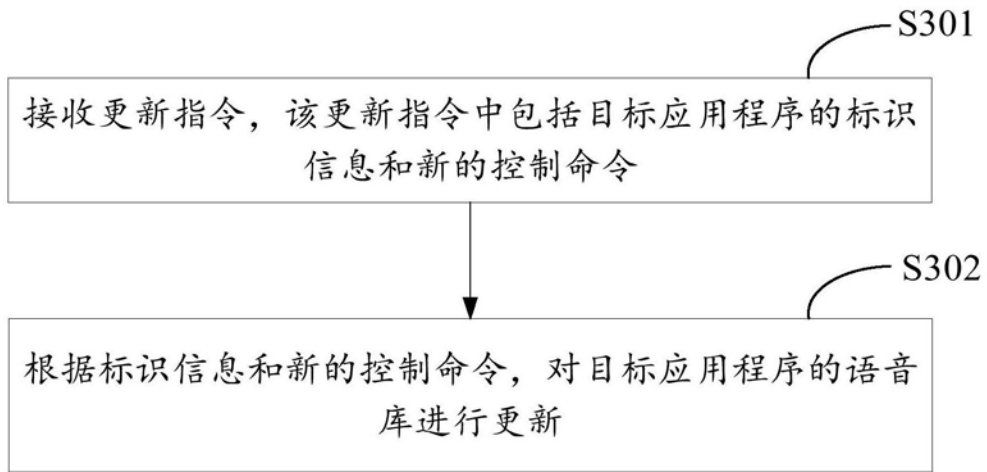


图3



图4

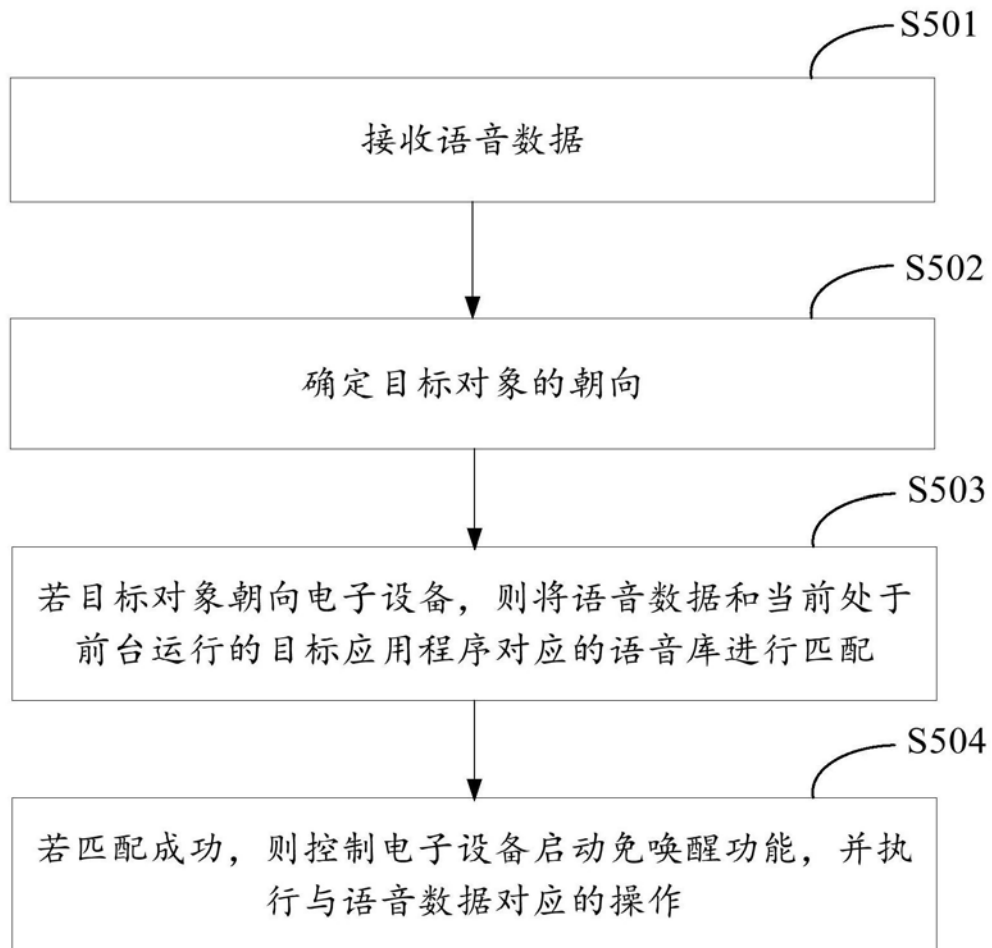


图5



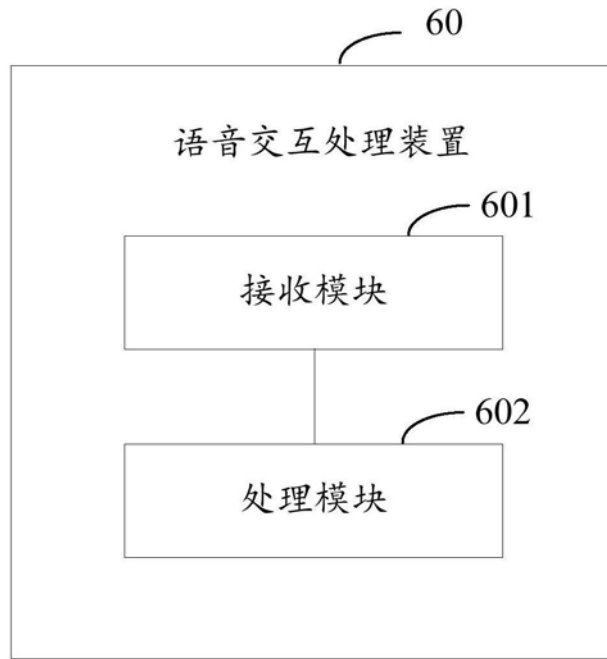


图6

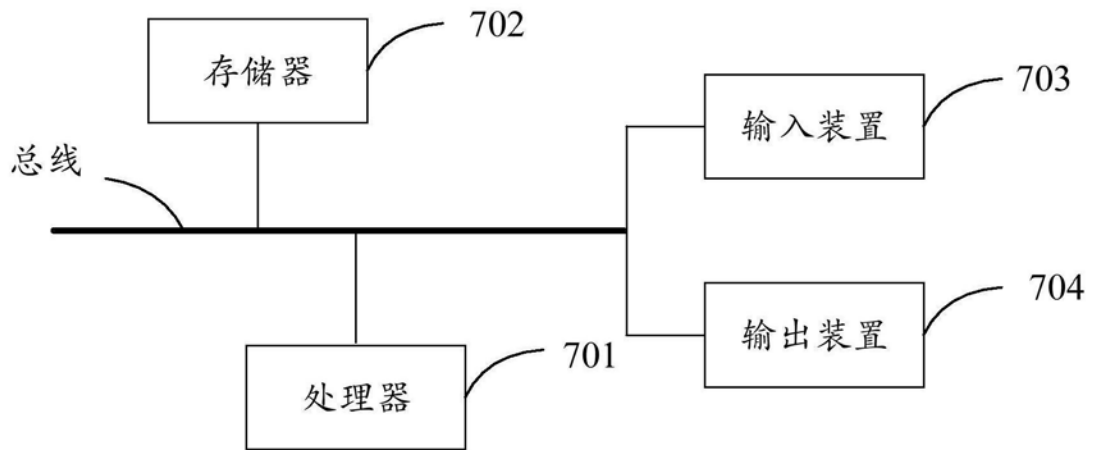


图7