



(12)发明专利申请

(10)申请公布号 CN 111061452 A

(43)申请公布日 2020.04.24

(21)申请号 201911300685.0

(22)申请日 2019.12.17

(71)申请人 北京小米智能科技有限公司
地址 100085 北京市海淀区清河朱房路临
66号F栋2单元1层101-103室

(72)发明人 方彦彬

(74)专利代理机构 北京名华博信知识产权代理
有限公司 11453

代理人 姜超

(51)Int.Cl.

G06F 3/16(2006.01)

G06F 3/0481(2013.01)

G06F 3/0487(2013.01)

G10L 15/22(2006.01)

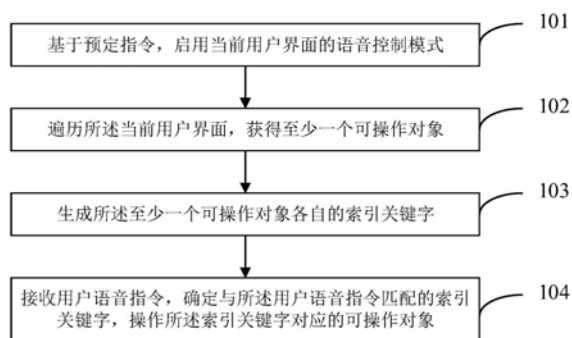
权利要求书3页 说明书11页 附图5页

(54)发明名称

用户界面的语音控制方法和装置

(57)摘要

本公开是关于一种用户界面的语音控制方法和装置。涉及智能设备控制技术,解决了语音助手实现方案配置繁琐、系统资源消耗大、应用范围窄的问题。该方法包括:基于预定指令,启用当前用户界面的语音控制模式;遍历所述当前用户界面,获得至少一个可操作对象;生成所述至少一个可操作对象各自的索引关键字;接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。本公开提供的技术方案适用于智能设备的语音控制操作,实现了高效、准确、易用的跨App通用语音控制。



1. 一种用户界面的语音控制方法,其特征在于,包括:
基于预定指令,启用当前用户界面的语音控制模式;
遍历所述当前用户界面,获得至少一个可操作对象;
生成所述至少一个可操作对象各自的索引关键字;
接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。

2. 根据权利要求1所述的用户界面的语音控制方法,其特征在于,所述遍历所述当前用户界面,获得至少一个可操作对象的步骤包括:

从服务器获取在所述当前用户界面上待显示的内容;
确定所述待显示的内容中的至少一个可操作对象。

3. 根据权利要求1所述的用户界面的语音控制方法,其特征在于,所述生成所述至少一个可操作对象各自的索引关键字的步骤包括:

为各可操作对象生成唯一的索引关键字,所述索引关键字包含以下形式中的任一或任意多种:

所述可操作对象的全部文字描述信息、所述可操作对象的部分文字描述信息、所述可操作对象的编号。

4. 根据权利要求3所述的用户界面的语音控制方法,其特征在于,所述生成所述至少一个可操作对象各自的索引关键字的步骤之后,还包括:

在所述索引关键字为所述可操作对象的编号的情况下,在所述当前用户界面下所述可操作对象的显示位置标注所述编号。

5. 根据权利要求3所述的用户界面的语音控制方法,其特征在于,所述接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象的步骤包括:

识别所述用户语音指令,提取所述用户语音指令中的语音操作信息和/或语音对象信息,所述语音操作信息指示操作,所述语音对象信息指示操作指向的对象;

根据所述语音对象信息,确定所述用户语音指令指向的可操作对象;

根据所述语音操作信息,确定对所述可操作对象所执行的操作;

根据所述可操作对象和/或所述操作,执行所述用户语音指令。

6. 根据权利要求5所述的用户界面的语音控制方法,其特征在于,所述根据所述语音对象信息,确定所述用户语音指令指向的可操作对象的步骤包括:

确定所述语音对象信息匹配的索引关键字;

确定所述索引关键字对应的可操作对象为所述语音对象信息指向的可操作对象。

7. 根据权利要求5所述的用户界面的语音控制方法,其特征在于,所述根据所述语音操作信息,确定对所述可操作对象所执行的操作的步骤包括:

查询预设的操作列表,在所述操作列表中包含多种操作;

确定与所述语音操作信息匹配的操作,以所述操作作为对所述可操作对象所执行的操作。

8. 一种用户界面的语音控制装置,其特征在于,包括:

模式启动模块,用于基于预定指令,启用当前用户界面的语音控制模式;

操作对象获取模块,用于遍历所述当前用户界面,获得至少一个可操作对象;
索引生成模块,用于生成所述至少一个可操作对象各自的索引关键字;
指令执行模块,用于接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,
操作所述索引关键字对应的可操作对象。

9. 根据权利要求8所述的用户界面的语音控制装置,其特征在于,所述操作对象获取模块包括:

待显示内容获取子模块,用于从服务器获取在所述当前用户界面上待显示的内容;
操作对象确定子模块,用于确定所述待显示的内容中的至少一个可操作对象。

10. 根据权利要求8所述的用户界面的语音控制装置,其特征在于,所述索引生成模块包括:

关键字生成子模块,用于为遍历确定的各可操作对象生成唯一的索引关键字,所述索引关键字包含以下形式中的任一或任意多种:

所述可操作对象的全部文字描述信息、所述可操作对象的部分文字描述信息、所述可操作对象的编号。

11. 根据权利要求10所述的用户界面的语音控制装置,其特征在于,所述索引生成模块还包括:

标注子模块,用于在所述索引关键字为所述可操作对象的编号的情况下,在所述当前用户界面下所述可操作对象的显示位置标注所述编号。

12. 根据权利要求10所述的用户界面的语音控制装置,其特征在于,所述指令执行模块包括:

语音识别子模块,用于识别所述用户语音指令,所述用户语音指令包括语音操作信息和/或语音对象信息,所述语音操作信息指示操作,所述语音对象信息指示操作指向的对象;

对象确定子模块,用于根据所述语音对象信息,确定所述用户语音指令指向的可操作对象;

操作确定子模块,用于根据所述语音操作信息,确定对所述可操作对象所执行的操作;

指令执行子模块,用于根据所述可操作对象和/或所述操作,执行所述用户语音指令。

13. 根据权利要求12所述的用户界面的语音控制装置,其特征在于,所述对象确定子模块包括:

索引查询单元,用于查询所述索引关键字,确定所述语音对象信息匹配的索引关键字;

指向对象确定单元,用于确定所述索引关键字对应的可操作对象为所述语音对象信息指向的可操作对象。

14. 根据权利要求12所述的用户界面的语音控制装置,其特征在于,所述操作确定子模块包括:

列表查询单元,用于查询预设的操作列表,在所述操作列表中包含多种操作;

操作确定单元,用于确定与所述语音操作信息匹配的操作,以所述操作作为对所述可操作对象所执行的操作。

15. 一种计算机装置,其特征在于,包括:

处理器;

用于存储处理器可执行指令的存储器；

其中,所述处理器被配置为:

基于预定指令,启用当前用户界面的语音控制模式;

遍历所述当前用户界面,获得至少一个可操作对象;

生成所述至少一个可操作对象各自的索引关键字;

接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。

16.一种非临时性计算机可读存储介质,当所述存储介质中的指令由移动终端的处理器执行时,使得移动终端能够执行一种用户界面的语音控制方法,所述方法包括:

基于预定指令,启用当前用户界面的语音控制模式;

遍历所述当前用户界面,获得至少一个可操作对象;

生成所述至少一个可操作对象各自的索引关键字;

接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。

用户界面的语音控制方法和装置

技术领域

[0001] 本公开涉及智能设备控制技术,尤其涉及一种用户界面的语音控制方法和装置。

背景技术

[0002] 随着自动语音识别(ASR)及自然语言处理(NLP)技术的发展,越来越多的语音辅助类软件开始出现,例如iPhone的siri,微软的小冰等。语音助手的功能也由最初简单的对话,开始结合终端设备功能向用户提供更多的帮助,例如可通过一语音助手打开某个App、查看天气、播放歌曲等。即便如此,还是不能完全解放双手,完全通过语音释放手机的各项能力。

[0003] 一种常见的语音助手实现方式是通过记录App的用户界面(UI)布局,把语音命令转换为模拟手势,模拟用户手势操作,以对指定控件进行操作。

[0004] 但App的布局可能发生变化,由于语音控制定位是根据App布局确定的,如果不及时更新App布局与语音操作信息的对应关系,则会发生误操作。且不同的App都需要单独定制相关的模拟手势操作方式,逐个App定制模拟手势会造成系统资源消耗过大。

[0005] 某些App也会对外提供特殊接口供语音助手调用,这样通过语音助手即可调用该App的语音支持组件,实现通过语音控制该App。但此种方案需要App向语音助手定制接口,这样App会增加额外的工作量,运算压力大;如App只是提供部分主要功能的接口,则会导致语音助手对App的操作能力受限。且不是所有App都会提供接口,同样导致可用语音助手操作的App数量较少。

[0006] 综上,语音助手对App内用户界面的控制方式不统一,导致语音控制的准确性和效率均较为低下。

发明内容

[0007] 为克服相关技术中存在的问题,本公开提供一种用户界面的语音控制方法和装置。

[0008] 根据本公开实施例的第一方面,提供一种用户界面的语音控制方法,其特征在于,包括:

[0009] 基于预定指令,启用当前用户界面的语音控制模式;

[0010] 遍历所述当前用户界面,获得至少一个可操作对象;

[0011] 生成所述至少一个可操作对象各自的索引关键字;

[0012] 接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。

[0013] 优选的,所述遍历所述当前用户界面,获得至少一个可操作对象的步骤包括:

[0014] 从服务器获取在所述当前用户界面上待显示的内容;

[0015] 确定所述待显示的内容中的至少一个可操作对象。

[0016] 优选的,所述生成所述至少一个可操作对象各自的索引关键字的步骤包括:

[0017] 为各可操作对象生成唯一的索引关键字,所述索引关键字包含以下形式中的任一或任意多种:

[0018] 所述可操作对象的全部文字描述信息、所述可操作对象的部分文字描述信息、所述可操作对象的编号。

[0019] 优选的,所述生成所述至少一个可操作对象各自的索引关键字的步骤之后,还包括:

[0020] 在所述索引关键字为所述可操作对象的编号的情况下,在所述当前用户界面下所述可操作对象的显示位置标注所述编号。

[0021] 优选的,所述接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象的步骤包括:

[0022] 识别用户语音指令,所述用户语音指令包括语音操作信息和/或语音对象信息,所述语音操作信息指示操作,所述语音对象信息指示操作指向的对象;

[0023] 根据所述语音对象信息,确定所述用户语音指令指向的可操作对象;

[0024] 根据所述语音操作信息,确定对所述可操作对象所执行的操作;

[0025] 根据所述可操作对象和/或所述操作,执行所述用户语音指令。

[0026] 优选的,所述根据所述语音对象信息,确定所述用户语音指令指向的可操作对象的步骤包括:

[0027] 查询所述索引关键字,确定所述语音对象信息匹配的索引关键字;

[0028] 确定所述索引关键字对应的可操作对象为所述语音对象信息指向的可操作对象。

[0029] 优选的,所述根据所述语音操作信息,确定对所述可操作对象所执行的操作的步骤包括:

[0030] 查询预设的操作列表,在所述操作列表中包含多种操作;

[0031] 确定与所述语音操作信息匹配的操作,以所述操作作为对所述可操作对象所执行的操作

[0032] 根据本公开的实施例的第二方面,提供了一种用户界面的语音控制装置,包括:

[0033] 模式启动模块,用于基于预定指令,启用当前用户界面的语音控制模式;

[0034] 操作对象获取模块,用于遍历所述当前用户界面,获得至少一个可操作对象;

[0035] 索引生成模块,用于生成所述至少一个可操作对象各自的索引关键字;

[0036] 指令执行模块,用于接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。

[0037] 优选的,所述操作对象获取模块包括:

[0038] 待显示内容获取子模块,用于从服务器获取在所述当前用户界面上待显示的内容;

[0039] 操作对象确定子模块,用于确定所述待显示的内容中的至少一个可操作对象。

[0040] 优选的,所述索引生成模块包括:

[0041] 关键字生成子模块,用于为遍历确定的各可操作对象生成唯一的索引关键字,所述索引关键字包含以下形式中的任一或任意多种:

[0042] 所述可操作对象的全部文字描述信息、所述可操作对象的部分文字描述信息、所述可操作对象的编号。

- [0043] 优选的,所述索引生成模块还包括:
- [0044] 标注子模块,用于在所述索引关键字为所述可操作对象的编号的情况下,在所述当前用户界面下所述可操作对象的显示位置标注所述编号。
- [0045] 优选的,所述指令执行模块包括:
- [0046] 语音识别子模块,用于识别用户语音指令,所述用户语音指令包括语音操作信息和/或语音对象信息,所述语音操作信息指示操作,所述语音对象信息指示操作指向的对象;
- [0047] 对象确定子模块,用于根据所述语音对象信息,确定所述用户语音指令指向的可操作对象;
- [0048] 操作确定子模块,用于根据所述语音操作信息,确定对所述可操作对象所执行的操作;
- [0049] 指令执行子模块,用于根据所述可操作对象和/或所述操作,执行所述用户语音指令。
- [0050] 优选的,所述对象确定子模块包括:
- [0051] 索引查询单元,用于查询所述索引关键字,确定所述语音对象信息匹配的索引关键字;
- [0052] 指向对象确定单元,用于确定所述索引关键字对应的可操作对象为所述语音对象信息指向的可操作对象。
- [0053] 优选的,所述操作确定子模块包括:
- [0054] 列表查询单元,用于查询预设的操作列表,在所述操作列表中包含多种操作;
- [0055] 操作确定单元,用于确定与所述语音操作信息匹配的操作,以所述操作作为对所述可操作对象所执行的操作。
- [0056] 根据本公开的实施例的第三方面,提供了一种计算机装置,包括:
- [0057] 处理器;
- [0058] 用于存储处理器可执行指令的存储器;
- [0059] 其中,所述处理器被配置为:
- [0060] 基于预定指令,启用当前用户界面的语音控制模式;
- [0061] 遍历所述当前用户界面,获得至少一个可操作对象;
- [0062] 生成所述至少一个可操作对象各自的索引关键字;
- [0063] 接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。
- [0064] 根据本公开的实施例的第四方面,提供了一种非临时性计算机可读存储介质,当所述存储介质中的指令由移动终端的处理器执行时,使得移动终端能够执行一种用户界面的语音控制方法,所述方法包括:
- [0065] 基于预定指令,启用当前用户界面的语音控制模式;
- [0066] 遍历所述当前用户界面,获得至少一个可操作对象;
- [0067] 生成所述至少一个可操作对象各自的索引关键字;
- [0068] 接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。

[0069] 本公开的实施例提供的技术方案可以包括以下有益效果：在需要启动语音控制时，基于预定指令，启用当前用户界面的语音控制模式，遍历所述当前用户界面，获得至少一个可操作对象，然后生成所述至少一个可操作对象各自的索引关键字，再接收用户语音指令，确定与所述用户语音指令匹配的索引关键字，操作所述索引关键字对应的可操作对象。基于实时生成的索引关键字完成语音控制，与当前应用环境完全匹配，实现了高效、准确、易用的跨App通用语音控制，解决了语音助手实现方案配置繁琐、系统资源消耗大、应用范围窄的问题。

[0070] 应当理解的是，以上的一般描述和后文的细节描述仅是示例性和解释性的，并不能限制本公开。

附图说明

[0071] 此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本发明的实施例，并与说明书一起用于解释本发明的原理。

[0072] 图1是根据一示例性实施例示出的一种语音控制方法的流程图。

[0073] 图2是根据一示例性实施例示出的一种语音控制方法的流程图。

[0074] 图3是根据一示例性实施例示出的一种语音控制方法的流程图。

[0075] 图4是根据一示例性实施例示出的一种语音控制方法的流程图。

[0076] 图5是根据一示例性实施例示出的确定可操作对象的流程图。

[0077] 图6是根据一示例性实施例示出的确定执行的操作的流程图。

[0078] 图7是根据一示例性实施例示出的一种语音控制装置的框图。

[0079] 图8是根据一示例性实施例示出的一种语音控制装置的框图。

[0080] 图9是根据一示例性实施例示出的索引生成模块703的框图。

[0081] 图10是根据一示例性实施例示出的索引生成模块703的框图。

[0082] 图11是根据一示例性实施例示出的指令执行模块704的框图。

[0083] 图12是根据一示例性实施例示出的对象确定子模块1102的框图。

[0084] 图13是根据一示例性实施例示出的操作确定子模块1103的框图。

[0085] 图14是根据一示例性实施例示出的一种装置的框图(移动终端的一般结构)。

具体实施方式

[0086] 这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本发明相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本发明的一些方面相一致的装置和方法的例子。

[0087] 一种常见的语音助手实现方式是通过记录App的用户界面(UI)布局，把语音命令转换为模拟手势，模拟用户手势操作，以对指定控件进行操作。

[0088] 但App的布局可能发生变化，由于语音控制定位是根据App布局确定的，如果不及时更新App布局与语音操作信息的对应关系，则会发生误操作。且不同的App都需要单独定制相关的模拟手势操作方式，逐个App定制模拟手势会造成系统资源消耗过大。

[0089] 某些App也会对外提供特殊接口供语音助手调用，这样通过语音助手即可调用该

App的语音支持组件,实现通过语音控制该App。但此种方案需要App向语音助手定制接口,这样App会增加额外的工作量,运算压力大;如App只是提供部分主要功能的接口,刚会导致语音助手对App的操作能力受限。且不是所有App都会提供接口,同样导致可用语音助手操作的App数量较少。

[0090] 为了解决上述问题,本公开的示例性实施例提供了一种语音控制方法和装置。通过动态遍历生成各App的语音索引,实现了实时的与实际应用环境相配合的语音控制配置,解决了语音助手实现方案存在配置繁琐、系统资源消耗大、应用范围窄的问题。

[0091] 本公开的一示例性实施例提供了一种用户界面的语音控制方法,使用该方法实现语音控制的流程如图1所示,包括:

[0092] 步骤101、基于预定指令,启用当前用户界面的语音控制模式。

[0093] 本公开实施例中,可在系统内建立一种语音控制模式,在这一模式下,启动语音控制。

[0094] 语音控制模式的启动,可通过预先设定的指令来实现。指令可有多种形式,例如通过某一按键/按键组合、特定的语音指令(例如“启动语音控制”这句话)、特定手势等。这些指令可固定配置于系统中,也可由用户自定义。

[0095] 本公开实施例中,所述当前用户界面包含以下界面中的任一或任意多种:

[0096] 系统界面、锁屏界面、应用内界面。

[0097] 其中,所述系统界面为操作系统的界面,在操作系统界面包含多种不同界面的情况下,在任意界面均可适用本公开实施例。

[0098] 所述锁屏界面为屏幕锁定情况下的界面。

[0099] 所述应用内界面为系统内安装的不同应用、软件、组件等的可视化界面,在启动应用时即激活了应用内界面。

[0100] 步骤102、遍历所述当前用户界面,获得至少一个可操作对象。

[0101] 本公开实施例中,所述可操作对象可为界面下的图标、按钮等。可为屏幕中的任意位置。具体的,通过遍历当前终端页面,找到所有可点击的对象作为可操作对象。

[0102] 本步骤中,在需要进行语音控制时,即启动实时的遍历,确定可操作对象,为准确进行语音控制提供配置基础。

[0103] 步骤103、生成所述至少一个可操作对象各自的索引关键字。

[0104] 本公开实施例中,每个可操作对象都具有一个索引关键字。根据索引关键字,即可确定对应的可操作对象。在用户语音指令中,包含有指示索引关键字的部分,根据该部分的识别结果,确定索引关键字后,即可确定对应的可操作对象,进而执行用户语音指令。

[0105] 步骤104、接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。

[0106] 本步骤中,执行用户语音指令,对当前用户界面进行操作控制。具体的,识别用户语音指令可获取一识别结果,识别结果包含对象和操作两方面的信息。根据索引关键字,确定对象后,即可执行该用户语音指令。

[0107] 本公开的一示例性实施例还提供了一种用户界面的语音控制方法,在语音控制模式启动的情况之下,由于用户的操作,当前用户界面显示的内容会发生变化。在这一场景下获取可操作对象的具体流程如图2所示,包括:

- [0108] 步骤201、从服务器获取在所述当前用户界面上待显示的内容。
- [0109] 本步骤中,根据用户操作,获取在所述当前用户界面上待显示的内容。
- [0110] 具体的,可根据用户操作,例如滑动屏幕、点按某一图标激活新的页面等,获取即将显示于当前用户界面上的待显示的内容。
- [0111] 步骤202、确定所述待显示的内容中的至少一个可操作对象。
- [0112] 本步骤中,遍历待显示的内容,获得至少一个可操作对象。
- [0113] 在待显示内容加载到当前用户界面显示后,即可将生成的索引关键字应用于当前用户界面,进而执行用户语音指令。
- [0114] 本公开的一示例性实施例还提供了一种用户界面的语音控制方法,基于可操作对象的相关信息建立遍历确定的各可操作对象的索引关键字,具体流程如图3所示,包括:
- [0115] 步骤301、为遍历确定的各可操作对象生成唯一的索引关键字。
- [0116] 本公开实施例中,所述索引关键字包含以下形式中的任一或任意多种:
- [0117] 所述可操作对象的全部文字描述信息、所述可操作对象的部分文字描述信息、所述可操作对象的编号。
- [0118] 可操作对象一般具有文字描述(例如App名称),该文字描述的长短不定。可使用可操作对象的全部文字描述信息作为索引关键字,例如“短信”。
- [0119] 在文字描述较长时,可从中摘取部分文字描述信息作为索引关键字。或为该可操作对象进行编号,编号可为数字,以该编号作为该可操作对象的索引关键字。
- [0120] 在可操作对象没有相应的文字描述时,也可为该可操作对象进行编号,以该编号作为该可操作对象的索引关键字。
- [0121] 优选的,在可操作对象在文字描述较长时,还可同时设置该可操作对应的文字描述和编号为索引关键字。这样,在用户后续发语音指令指向该文字描述或编号的情况下,均可索引至该可操作对象。
- [0122] 步骤302、在所述索引关键字为所述可操作对象的编号的情况下,在所述当前用户界面下所述可操作对象的显示位置标注所述编号。
- [0123] 本步骤为可选步骤,在为可操作对象生成编号作为索引关键字的情况下,在当前用户界面可操作对象的位置标注所述编号,即在可操作对象位置显示该编号,在语音控制期间,该标号即为该可操作对象的索引关键字。用户可直接看到可操作对象对应的编号,进而准确的发出语音指令。
- [0124] 本发明实施例中,考虑到用户语音指令中对同一可操作对象语言描述上的多样性,还可以基于索引关键字建立语音索引,在所述语音索引中指示所述索引关键字与所述语音对象信息之间的关联关系。对于语音索引,一般均支持语音识别。
- [0125] 例如,将“短信”这个词与“短信”的语音建立关联,在语音识别结果转化的文字为“短信”时,即可根据语音索引确定索引关键字为“短信”,进而确定可操作对象;或将“收藏”这个索引关键字与“我的收藏”/“收藏夹”等语音识别结果相关联,允许用一种或多种与索引关键字并不完全相同的语音识别结果选定同一可操作对象。
- [0126] 本公开的一示例性实施例还提供了一种用户界面的语音控制方法,在实时生成索引关键字后,基于所述索引关键字,执行用户语音指令,操作与所述用户语音指令对应的可操作对象,具体流程如图4所示,包括:

[0127] 步骤401、识别用户语音指令。

[0128] 本步骤中,在进入语音控制模式等启动语音控制的场景之下后,即开始识别用户语音指令。

[0129] 所述用户语音指令包括语音操作信息和/或语音对象信息,所述语音操作信息指示操作,所述语音对象信息指示操作指向的对象。

[0130] 步骤402、根据所述语音对象信息,确定所述用户语音指令指向的可操作对象。

[0131] 本步骤具体如图5所示,包括:

[0132] 步骤501、查询所述索引关键字,确定所述语音对象信息匹配的索引关键字。

[0133] 本步骤中,首先根据语音对象信息的识别结果,匹配得到对应的索引关键字。

[0134] 步骤502、确定所述索引关键字对应的可操作对象为所述语音对象信息指向的可操作对象。

[0135] 本步骤中,在确定索引关键字后,再根据索引关键字归属可操作对象的对应关系,最终在语音对象信息和可操作对象之间确定关联,选定可操作对象。

[0136] 步骤403、根据所述语音操作信息,确定对所述可操作对象所执行的操作。

[0137] 本步骤具体如图6所示,包括:

[0138] 步骤601、查询预设的操作列表,在所述操作列表中包含多种操作。

[0139] 本公开实施例中,可维护一操作列表,在操作列表中维护有语音操作信息可进行的虚拟操作,例如:点(短按),压(长按),浏览(慢速滑动),返回(返回上一级界面),桌面(返回桌面),声音调整(音量操作),输入(输入法输入)等操作。

[0140] 步骤602、确定与所述语音操作信息匹配的操作,以所述操作作为对所述可操作对象所执行的操作。

[0141] 本步骤中,根据对语音操作信息的识别结果,匹配相应的操作,进而确定对所述可操作对象所执行的操作。

[0142] 需要说明的是,部分语音操作信息可单独使用,不需要与语音对象信息构成语音指令,即部分操作是没有指向特写对象的,例如“返回上一级”的操作。

[0143] 步骤402和步骤403为可选步骤,在语音指令包括语音对象信息时才需要执行步骤402,在语音指令包括语音操作信息时才需要执行步骤403。步骤402与步骤403可由不同组件进行,也可由相同组件进行,并无严格时序关系。

[0144] 步骤404、根据所述可操作对象和/或所述操作,执行所述用户语音指令。

[0145] 本步骤中,根据对语音指令分析得到的可操作对象和/或操作,执行语音指令。

[0146] 具体的,在语音指令包括语音操作信息和语音对象信息的情况下,对语音指令解析后可得到可操作对象+操作的结果,即对相应的可操作对象执行相应的操作。

[0147] 在语音指令仅包括语音操作信息的情况下,对语音指令解析后可得到操作,仅执行该操作即可。

[0148] 在语音指令仅包括语音对象信息的情况下,如系统已预先配置了默认语音操作,则对根据语音对象信息确定的可操作对象执行该默认语音操作。例如默认语音操作为“打开”,语音对象信息的语音内容为“微信”,则执行打开微信的操作。

[0149] 在执行语音指令时,由于应用环境是通过实时遍历当前用户界面生成的,因此语音指令仅与当前用户界面相关,比如当前的前台应用是淘宝,那么语音指令“打开购物车”

的结果就是打开淘宝的购物车,而非京东、亚马逊等其他应用。

[0150] 本公开的实施例中,还可由用户自定义特殊语音命令,显示当前App支持的直达语音服务,该服务面向App开放,由App定制。在启动App时,即可应用相应的直达语音服务。

[0151] 本公开的一示例性实施例还提供了一种用户界面的语音控制装置,其结构如图7所示,包括:

[0152] 模式启动模块701,用于基于预定指令,启用当前用户界面的语音控制模式;

[0153] 操作对象获取模块702,用于遍历所述当前用户界面,获得至少一个可操作对象;

[0154] 索引生成模块703,用于生成所述至少一个可操作对象各自的索引关键字;

[0155] 指令执行模块704,用于接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。

[0156] 优选的,所述操作对象获取模块的结构如图8所示,包括:

[0157] 待显示内容获取子模块801,用于从服务器获取在所述当前用户界面上待显示的内容;

[0158] 操作对象确定子模块802,用于确定所述待显示的内容中的至少一个可操作对象。

[0159] 优选的,所述索引生成模块703的结构如图9所示,包括:

[0160] 关键字生成子模块901,用于为遍历确定的各可操作对象生成唯一的索引关键字,所述索引关键字包含以下形式中的任一或任意多种:

[0161] 所述可操作对象的全部文字描述信息、所述可操作对象的部分文字描述信息、所述可操作对象的编号。

[0162] 优选的,所述索引生成模块703的结构如图10所示,还包括:

[0163] 标注子模块902,用于在所述索引关键字为所述可操作对象的编号的情况下,在所述当前用户界面下所述可操作对象的显示位置标注所述编号。

[0164] 优选的,所述指令执行模块704的结构如图11所示,包括:

[0165] 语音识别子模块1101,用于识别用户语音指令,所述用户语音指令包括语音操作信息和/或语音对象信息,所述语音操作信息指示操作,所述语音对象信息指示操作指向的对象;

[0166] 对象确定子模块1102,用于根据所述语音对象信息,确定所述用户语音指令指向的可操作对象;

[0167] 操作确定子模块1103,用于根据所述语音操作信息,确定对所述可操作对象所执行的操作;

[0168] 指令执行子模块1104,用于根据所述可操作对象和/或所述操作,执行所述用户语音指令。

[0169] 优选的,所述对象确定子模块1102的结构如图12所示,包括:

[0170] 索引查询单元1201,用于查询所述索引关键字,确定所述语音对象信息匹配的索引关键字;

[0171] 指向对象确定单元1202,用于确定所述索引关键字对应的可操作对象为所述语音对象信息指向的可操作对象。

[0172] 优选的,所述操作确定子模块1103的结构如图13所示,包括:

[0173] 列表查询单元1301,用于查询预设的操作列表,在所述操作列表中包含多种操作;

[0174] 操作确定单元1302,用于确定与所述语音操作信息匹配的操作,以所述操作作为对所述可操作对象所执行的操作。

[0175] 本公开的一示例性实施例还提供了一种计算机装置,包括:

[0176] 处理器;

[0177] 用于存储处理器可执行指令的存储器;

[0178] 其中,所述处理器被配置为:

[0179] 基于预定指令,启用当前用户界面的语音控制模式;

[0180] 遍历所述当前用户界面,获得至少一个可操作对象;

[0181] 生成所述至少一个可操作对象各自的索引关键字;

[0182] 接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。

[0183] 图14是根据一示例性实施例示出的一种用于语音控制的装置1400的框图。例如,装置1400可以是移动电话,计算机,数字广播终端,消息收发设备,游戏控制台,平板设备,医疗设备,健身设备,个人数字助理等。

[0184] 参照图14,装置1400可以包括以下一个或多个组件:处理组件1402,存储器1404,电力组件1406,多媒体组件1408,音频组件1410,输入/输出(I/O)的接口1412,传感器组件1414,以及通信组件1416。

[0185] 处理组件1402通常控制装置1400的整体操作,诸如与显示,电话呼叫,数据通信,相机操作和记录操作相关联的操作。处理组件1402可以包括一个或多个处理器1420来执行指令,以完成上述的方法的全部或部分步骤。此外,处理组件1402可以包括一个或多个模块,便于处理组件1402和其他组件之间的交互。例如,处理组件1402可以包括多媒体模块,以方便多媒体组件1408和处理组件1402之间的交互。

[0186] 存储器1404被配置为存储各种类型的数据以支持在设备1400的操作。这些数据的示例包括用于在装置1400上操作的任何应用程序或方法的指令,联系人数据,电话簿数据,消息,图片,视频等。存储器1404可以由任何类型的易失性或非易失性存储设备或者它们的组合实现,如静态随机存取存储器(SRAM),电可擦除可编程只读存储器(EEPROM),可擦除可编程只读存储器(EPROM),可编程只读存储器(PROM),只读存储器(ROM),磁存储器,快闪存储器,磁盘或光盘。

[0187] 电力组件1406为装置1400的各种组件提供电力。电力组件1406可以包括电源管理系统,一个或多个电源,及其他与为装置1400生成、管理和分配电力相关联的组件。

[0188] 多媒体组件1408包括在所述装置1400和用户之间的提供一个输出接口的屏幕。在一些实施例中,屏幕可以包括液晶显示器(LCD)和触摸面板(TP)。如果屏幕包括触摸面板,屏幕可以被实现为触摸屏,以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界,而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中,多媒体组件1408包括一个前置摄像头和/或后置摄像头。当设备1400处于操作模式,如拍摄模式或视频模式时,前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

[0189] 音频组件1410被配置为输出和/或输入音频信号。例如,音频组件1410包括一个麦

克风(MIC),当装置1400处于操作模式,如呼叫模式、记录模式和语音识别模式时,麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器1404或经由通信组件1416发送。在一些实施例中,音频组件1410还包括一个扬声器,用于输出音频信号。

[0190] I/O接口1412为处理组件1402和外围接口模块之间提供接口,上述外围接口模块可以是键盘,点击轮,按钮等。这些按钮可包括但不限于:主页按钮、音量按钮、启动按钮和锁定按钮。

[0191] 传感器组件1414包括一个或多个传感器,用于为装置1400提供各个方面的状态评估。例如,传感器组件1414可以检测到设备1400的打开/关闭状态,组件的相对定位,例如所述组件为装置1400的显示器和小键盘,传感器组件1414还可以检测装置1400或装置1400一个组件的位置改变,用户与装置1400接触的存在或不存在,装置1400方位或加速/减速和装置1400的温度变化。传感器组件1414可以包括接近传感器,被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件1414还可以包括光传感器,如CMOS或CCD图像传感器,用于在成像应用中使用。在一些实施例中,该传感器组件1414还可以包括加速度传感器,陀螺仪传感器,磁传感器,压力传感器或温度传感器。

[0192] 通信组件1416被配置为便于装置1400和其他设备之间有线或无线方式的通信。装置1400可以接入基于通信标准的无线网络,如WiFi,2G或3G,或它们的组合。在一个示例性实施例中,通信组件1416经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中,所述通信组件1416还包括近场通信(NFC)模块,以促进短程通信。例如,在NFC模块可基于射频识别(RFID)技术,红外数据协会(IrDA)技术,超宽带(UWB)技术,蓝牙(BT)技术和其他技术来实现。

[0193] 在示例性实施例中,装置1400可以被一个或多个应用专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理设备(DSPD)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、控制器、微控制器、微处理器或其他电子元件实现,用于执行上述方法。

[0194] 在示例性实施例中,还提供了一种包括指令的非临时性计算机可读存储介质,例如包括指令的存储器1404,上述指令可由装置1400的处理器1420执行以完成上述方法。例如,所述非临时性计算机可读存储介质可以是ROM、随机存取存储器(RAM)、CD-ROM、磁带、软盘和光数据存储设备等。

[0195] 一种非临时性计算机可读存储介质,当所述存储介质中的指令由移动终端的处理器执行时,使得移动终端能够执行一种用户界面的语音控制方法,所述方法包括:

[0196] 基于预定指令,启用当前用户界面的语音控制模式;

[0197] 遍历所述当前用户界面,获得至少一个可操作对象;

[0198] 生成所述至少一个可操作对象各自的索引关键字;

[0199] 接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。

[0200] 本公开的实施例提供了一种语音控制方法和装置,在需要启动语音控制时,基于预定指令,启用当前用户界面的语音控制模式,遍历所述当前用户界面,获得至少一个可操作对象,然后生成所述至少一个可操作对象各自的索引关键字,再接收用户语音指令,确定与所述用户语音指令匹配的索引关键字,操作所述索引关键字对应的可操作对象。基于实时生成的索引关键字完成语音控制,与当前应用环境完全匹配,实现了高效、准确、易用的

跨App通用语音控制,解决了语音助手实现方案配置繁琐、系统资源消耗大、应用范围窄的问题。

[0201] 可应用于android系统的手机等移动终端上,以语音控制完成大多数操作。

[0202] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本发明的其它实施方案。本申请旨在涵盖本发明的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本发明的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本发明的真正范围和精神由下面的权利要求指出。

[0203] 应当理解的是,本发明并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本发明的范围仅由所附的权利要求来限制。

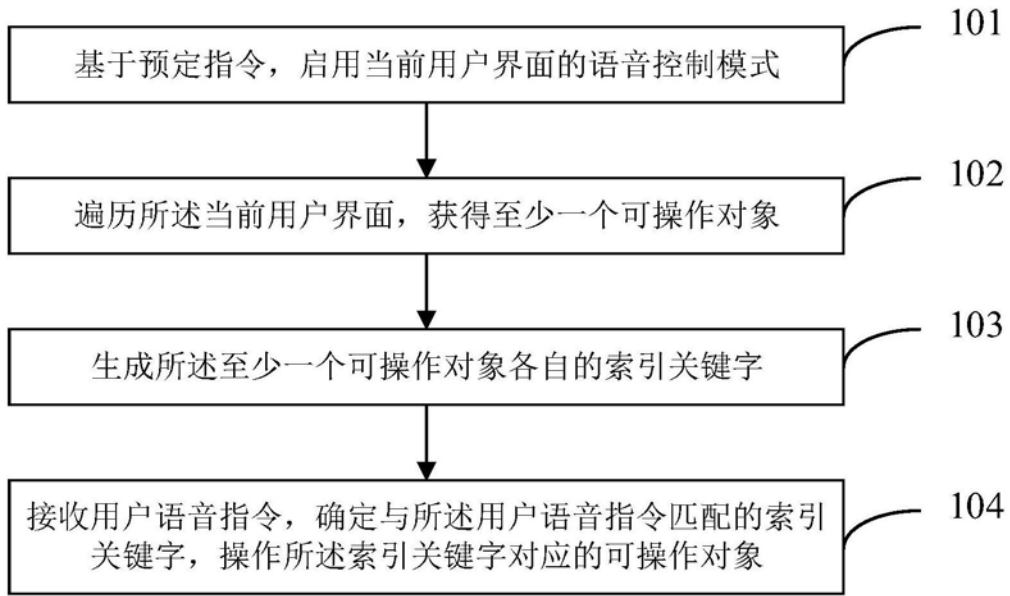


图1

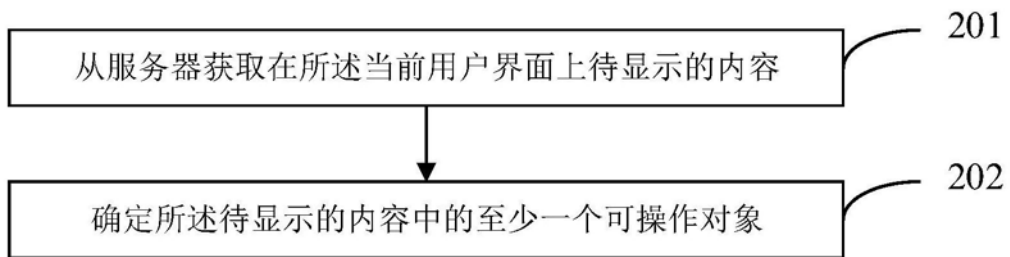


图2

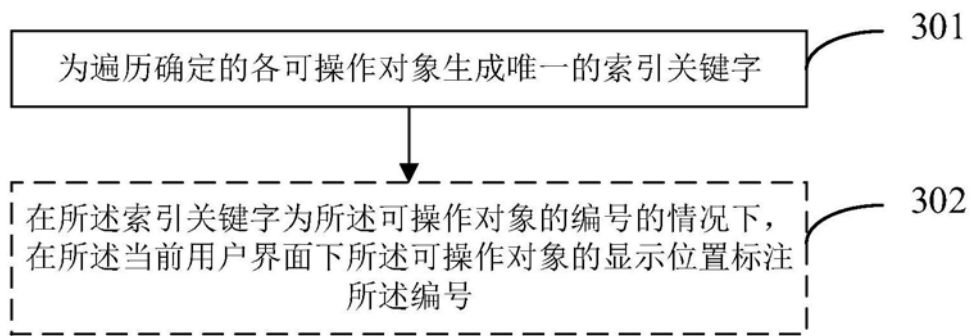


图3

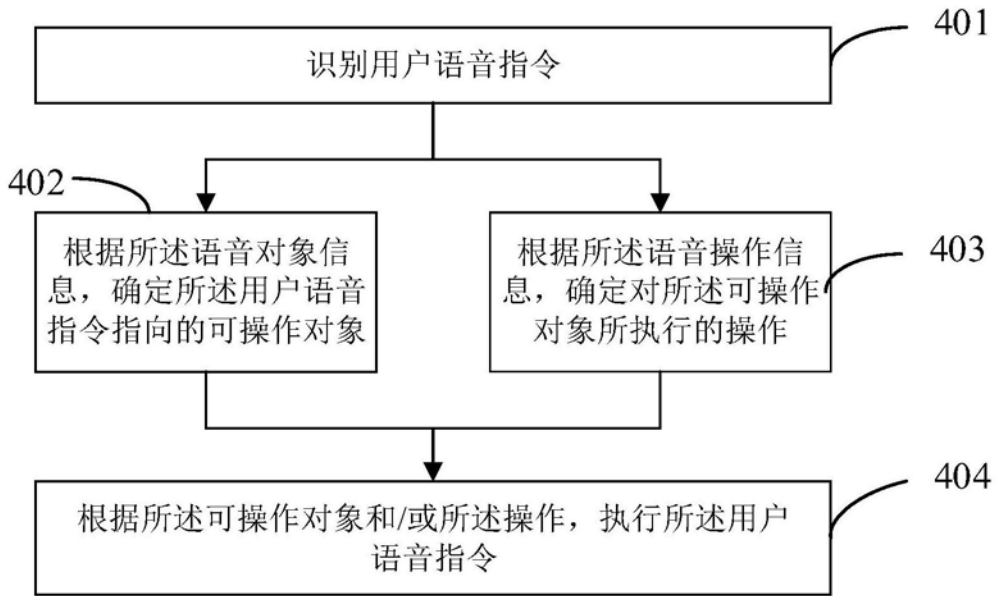


图4

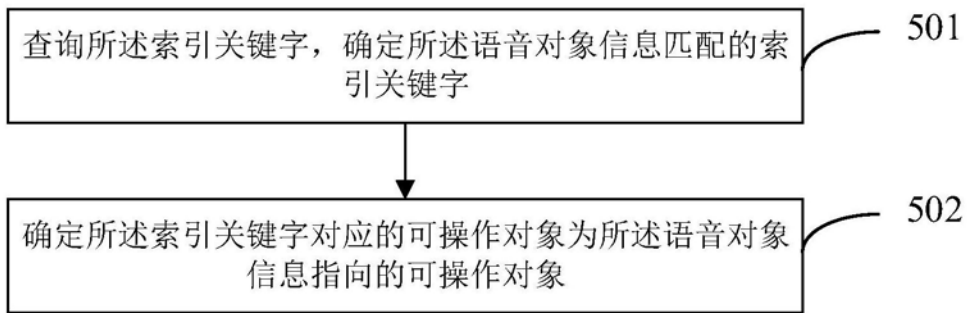


图5

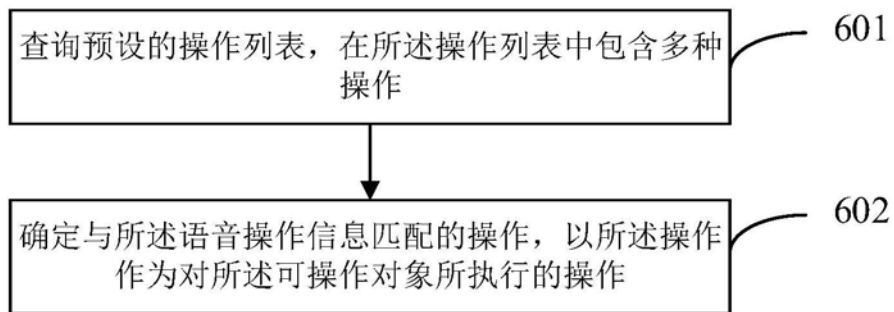


图6

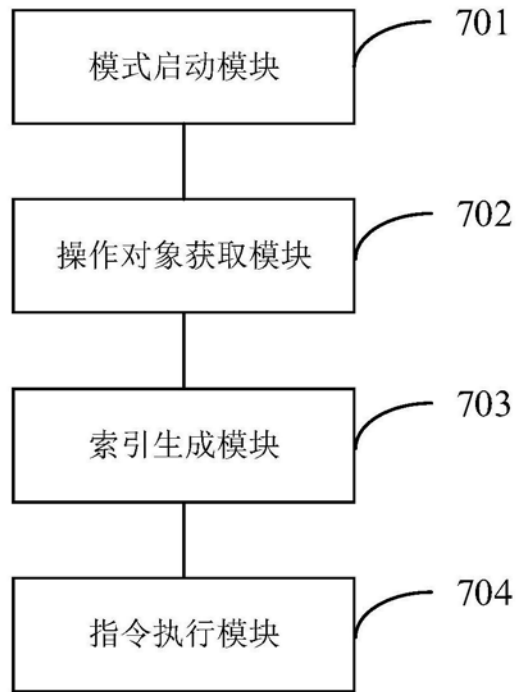


图7

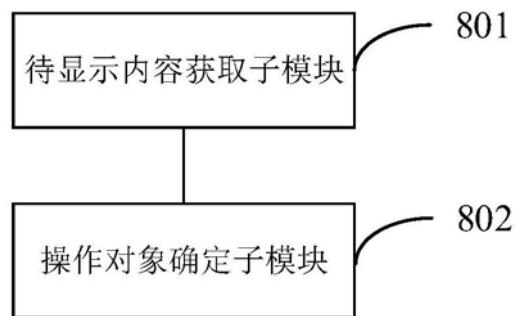


图8

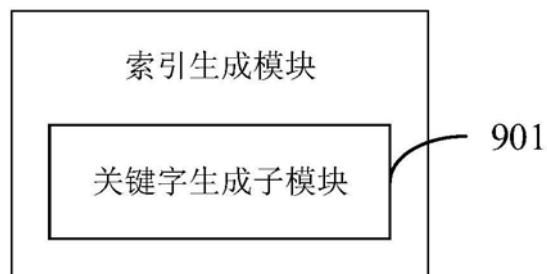


图9

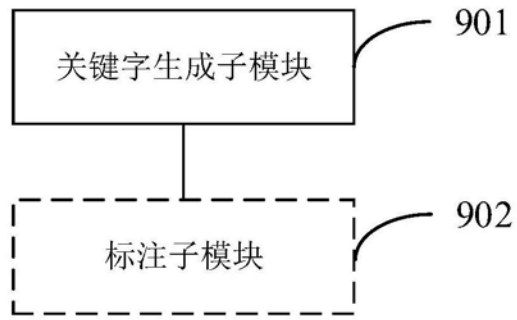


图10

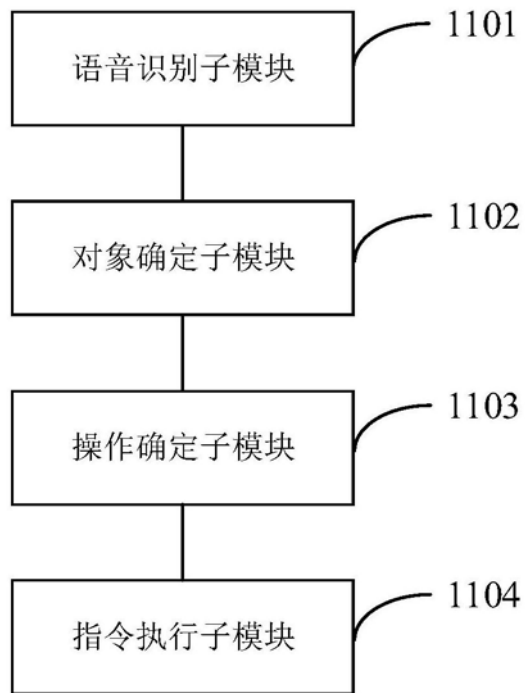


图11

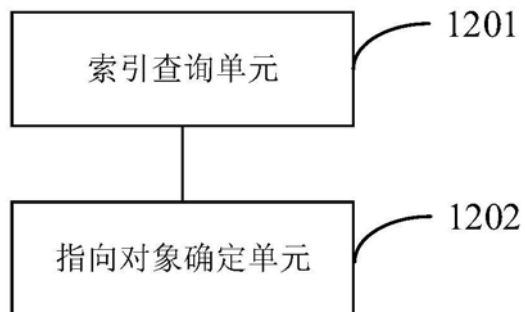


图12

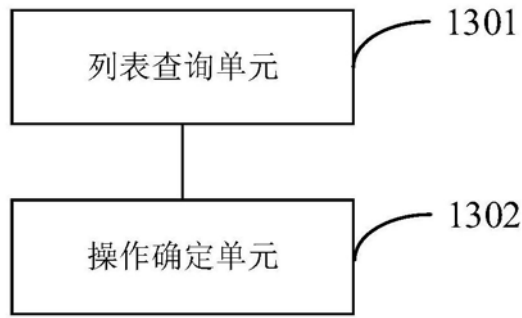


图13

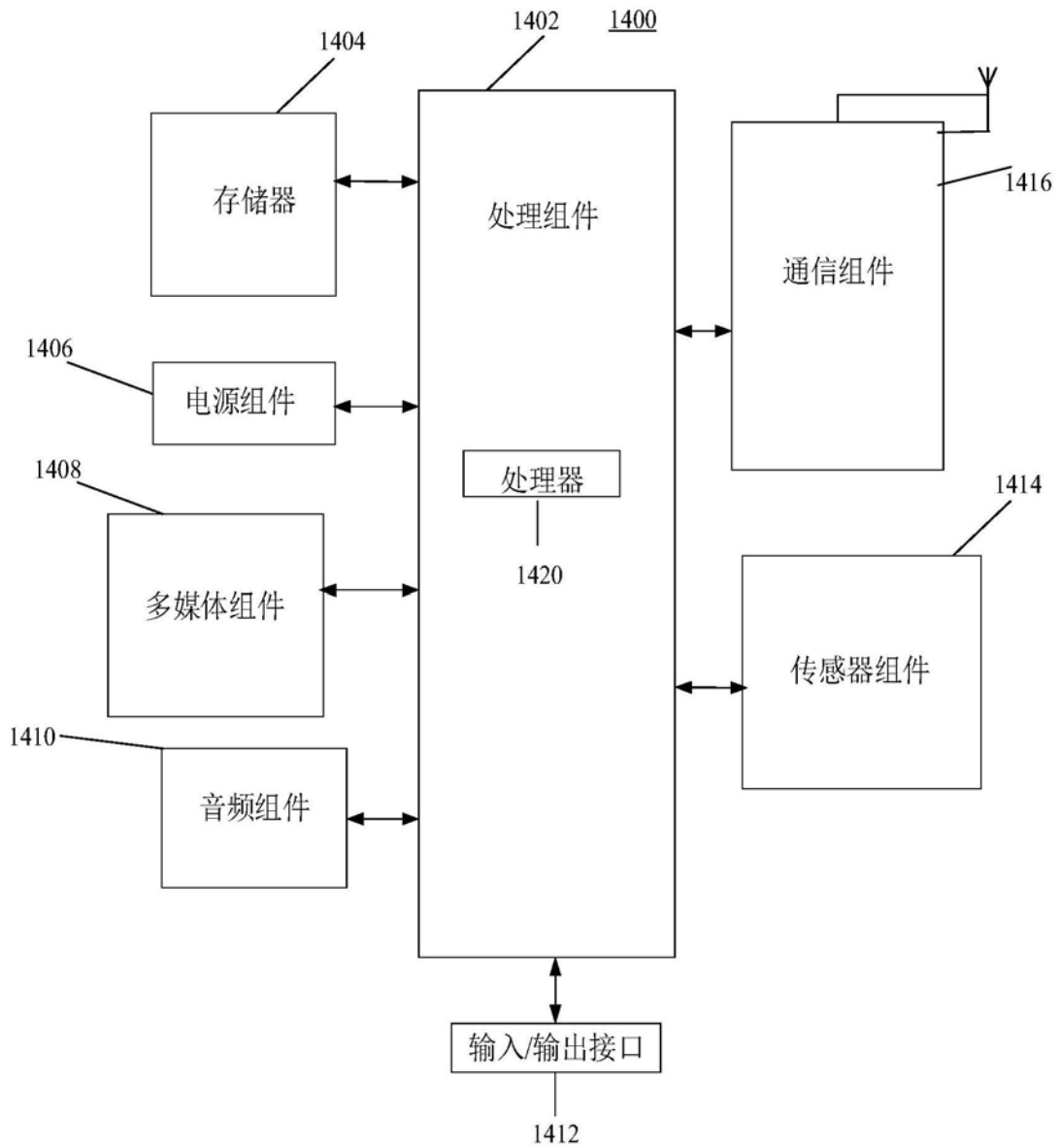


图14