



(12) 发明专利

(10) 授权公告号 CN 111833874 B

(45) 授权公告日 2023.12.05

(21) 申请号 202010662473.3

G06F 3/16 (2006.01)

(22) 申请日 2020.07.10

G06F 9/4401 (2018.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 111833874 A

(56) 对比文件

CN 103871408 A, 2014.06.18

CN 107564517 A, 2018.01.09

(43) 申请公布日 2020.10.27

CN 108335695 A, 2018.07.27

(73) 专利权人 上海茂声智能科技有限公司

CN 109243431 A, 2019.01.18

地址 201306 上海市浦东新区环湖西二路
888号C楼

CN 109272991 A, 2019.01.25

CN 109686368 A, 2019.04.26

(72) 发明人 陈晓松 李旭滨

CN 109871238 A, 2019.06.11

US 2017188437 A1, 2017.06.29

(74) 专利代理机构 上海领誉知识产权代理有限公司 31383

US 9047857 B1, 2015.06.02

专利代理师 车超平 王琰

审查员 董小东

(51) Int. Cl.

G10L 15/22 (2006.01)

G10L 15/08 (2006.01)

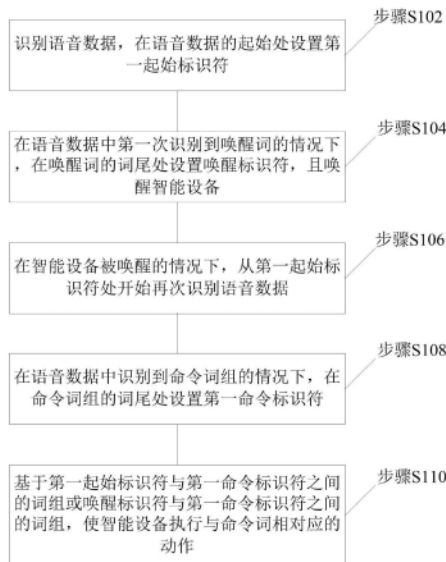
权利要求书2页 说明书10页 附图4页

(54) 发明名称

一种基于标识符的人机交互方法、系统、设备和存储介质

(57) 摘要

本发明涉及一种基于标识符的人机交互方法、系统、设备和存储介质,其中,该基于标识符的人机交互方法包括:识别语音数据,在语音数据的起始处设置第一起始标识符;在语音数据中第一次识别到唤醒词的情况下,在唤醒词的词尾处设置唤醒标识符,且唤醒智能设备;在智能设备被唤醒的情况下,从第一起始标识符处开始再次识别语音数据;在语音数据中识别到命令词组的情况下,在命令词组的词尾处设置第一命令标识符;基于第一起始标识符与第一命令标识符之间的词组或唤醒标识符与第一命令标识符之间的词组,使智能设备执行与命令词组相对应的动作。通过本发明,增加one-shot交互方式的同时,保留了原始的“先唤醒再识别”的交互方式,提高了人机交互体验。



1. 一种基于标识符的人机交互方法,其特征在于,所述方法包括:
 - 识别语音数据,在所述语音数据的起始处设置第一起始标识符;
 - 在所述语音数据中第一次识别到唤醒词的情况下,在所述唤醒词的词尾处设置唤醒标识符,且唤醒智能设备;
 - 在所述智能设备被唤醒的情况下,从所述第一起始标识符处开始再次识别所述语音数据;
 - 在所述语音数据中识别到命令词组的情况下,在所述命令词组的词尾处设置第一命令标识符;
 - 基于所述第一起始标识符与所述第一命令标识符之间的词组或所述唤醒标识符与所述第一命令标识符之间的词组,使所述智能设备执行与所述命令词组相对应的执行动作。
2. 根据权利要求1所述的方法,其特征在于,在所述智能设备被唤醒的情况下,从所述第一起始标识符开始再次识别所述语音数据之后,所述方法还包括:
 - 在再次识别到所述唤醒词的情况下,继续识别所述语音数据。
3. 根据权利要求1所述的方法,其特征在于,在所述智能设备被唤醒的情况下,从所述第一起始标识符开始再次识别所述语音数据之后,所述方法还包括:
 - 在所述唤醒标识符之后设置第二起始标识符,其中,所述唤醒标识符与所述第二起始标识符之间具有一定的时间间隔;
 - 在所述语音数据中识别到所述命令词组的情况下,在所述命令词组的词尾处设置第二命令标识符;
 - 基于所述第二起始标识符与所述第二命令标识符之间的词组,使所述智能设备执行与所述命令词组相对应的执行动作。
4. 根据权利要求1所述的方法,其特征在于,在所述识别语音数据之前,所述方法还包括:
 - 获取唤醒词和命令词;
 - 将所述命令词作为单一命令词组,且将所述唤醒词与所述命令词组合为复合命令词组;
 - 基于所述单一命令词组与所述复合命令词组构建命令词集,其中,所述单一命令词组与所述复合命令词组所对应的执行动作相同。
5. 一种基于标识符的人机交互系统,其特征在于,包括:
 - 语音采集单元,用于采集语音数据;
 - 语音识别单元,用于识别所述语音数据;
 - 标识单元,用于在所述语音数据的起始处设置第一起始标识符,且在所述语音数据中第一次出现唤醒词的情况下,在所述唤醒词的词尾处设置唤醒标识符,在所述语音数据中出现命令词组的情况下,在所述命令词组的词尾处设置第一命令标识符;
 - 唤醒单元,用于在检测到所述唤醒标识符的情况下,唤醒智能设备。
6. 根据权利要求5所述的系统,其特征在于,还包括:
 - 维持单元,用于在所述智能设备被唤醒的情况下,使所述智能设备维持在唤醒状态。
7. 根据权利要求5所述的系统,其特征在于,还包括:
 - 命令词集构建单元,用于将命令词作为单一命令词组以及将唤醒词与命令词组合为复

合命令词组,并基于所述单一命令词组与所述复合命令词组构建命令词集;

存储单元,用于存储所述命令词集。

8.根据权利要求5所述的系统,其特征在于,还包括;

执行单元,用于基于所述第一起始标识符与所述第一命令标识符之间的词组或所述唤醒标识符与所述第一命令标识符之间的词组,使所述智能设备执行与所述命令词组相对应的执行动作。

9.一种计算机设备,包括存储器、处理器以及存储在所述存储器上并可在所述处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序时实现如权利要求1~4中任一项所述的方法。

10.一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现如权利要求1~4中任一项所述的方法。

一种基于标识符的人机交互方法、系统、设备和存储介质

技术领域

[0001] 本申请涉及人机交互领域,特别是涉及一种基于标识符的人机交互方法、系统、设备和存储介质。

背景技术

[0002] 人机交互是智能设备的最为重要的功能,随着智能设备的发展,人们对人机交互体验的要求也越来越高。现有的人机交互方式一般有两种方式:“先唤醒再识别”人机交互方式:即用户先说出语音唤醒词,智能设备接收并响应语音唤醒,然后用户再说出语音命令词,智能设备再应答语音命令,如在用户说出“你好”的情况下,智能设备被唤醒,然后用户再说出“开灯”,智能设备执行开灯的操作;one-shot人机交互方式:即用户不间断地说出语音唤醒词和语音命令词,然后智能设备接收并应答语音命令,如在用户说出“你好开灯”的情况下,智能设备被唤醒且执行开灯的操作。

[0003] 相关技术中,“先唤醒再识别”人机交互方式在使用的情况下,必须先说出语音唤醒词,待智能设备唤醒之后,用户才能说出命令词,操控智能设备执行操作,这种方式的局限性比较大,人机交互体验较差。one-shot人机交互方式也只是在所有的命令词前添加一个唤醒词,将唤醒词和命令词作为一个复合命令词,从而使用户与智能设备进行交互,其本质上与“先唤醒再识别”人机交互方式非常相似。因为在用户使用智能设备的情况下,不同的用户有不同的人机交互习惯,所以智能设备仅支持“先唤醒再识别”人机交互方式或仅只支持one-shot人机交互方式都会影响用户的人机交互体验。

[0004] 目前针对相关技术中因智能设备仅能支持“先唤醒再识别”人机交互方式或one-shot人机交互方式,而导致影响用户体验的问题,尚未提出有效的解决方案。

发明内容

[0005] 本申请实施例提供了一种基于标识符的人机交互方法、系统、设备和存储介质,以至少解决相关技术中因智能设备仅能支持“先唤醒再识别”人机交互方式或one-shot人机交互方式,而导致影响用户体验的问题。

[0006] 本发明的第一个方面,提供了一种基于标识符的人机交互方法,所述方法包括:

[0007] 识别语音数据,在所述语音数据的起始处设置第一起始标识符;

[0008] 在所述语音数据中第一次识别到唤醒词的情况下,在所述唤醒词的词尾处设置唤醒标识符,且唤醒智能设备;

[0009] 在所述智能设备被唤醒的情况下,从所述第一起始标识符处开始再次识别所述语音数据;

[0010] 在所述语音数据中识别到命令词组的情况下,在所述命令词组的词尾处设置第一命令标识符;

[0011] 基于所述第一起始标识符与所述第一命令标识符之间的词组或所述唤醒标识符与所述第一命令标识符之间的词组,使所述智能设备执行与所述命令词组相对应的执行动

作。

[0012] 进一步地,在所述的人机交互系统中,在所述智能设备被唤醒的情况下,从所述起始标识符开始再次识别所述语音数据之后,所述方法还包括:

[0013] 在再次识别到所述唤醒词的情况下,继续识别所述语音数据。

[0014] 进一步地,在所述的人机交互系统中,在所述智能设备被唤醒的情况下,从所述第一起始标识符开始再次识别所述语音数据之后,所述方法还包括:

[0015] 在所述唤醒标识符之后设置第二起始标识符,其中,所述唤醒标识符与所述第二起始标识符之间具有一定的时间间隔;

[0016] 在所述语音数据中识别到命令词组的情况下,在所述命令词组的词尾处设置第二命令标识符;

[0017] 基于所述第二起始标识符与所述第二命令标识符之间的词组,使所述智能设备执行与所述命令词组相对应的执行动作。

[0018] 进一步地,在所述的人机交互系统中,在所述识别语音数据之前,所述方法还包括:

[0019] 获取唤醒词和命令词;

[0020] 将所述命令词作为单一命令词组,且将所述唤醒词与所述命令词组合为复合命令词组;

[0021] 基于所述单一命令词组与所述复合命令词组构建命令词集,其中,所述单一命令词组与所述组合命令词组所对应的执行动作相同。

[0022] 本发明的第二个方面,提供了一种基于标识符的人机交互系统,包括;

[0023] 语音采集单元,用于采集语音数据;

[0024] 语音识别单元,用于识别所述语音数据;

[0025] 标识单元,用于在所述语音数据的起始处设置第一起始标识符,且在所述语音数据中第一次出现唤醒词的情况下,在所述唤醒词的词尾处设置唤醒标识符,在所述语音数据中出现命令词组的情况下,在所述命令词组的词尾处设置第一命令标识符;

[0026] 唤醒单元,用于在检测到所述唤醒标识符的情况下,唤醒智能设备。

[0027] 进一步地,在所述的人机交互系统中,还包括:

[0028] 维持单元,用于在所述智能设备被唤醒的情况下,使所述智能设备维持在唤醒状态。

[0029] 进一步地,在所述的人机交互系统中,还包括:

[0030] 命令词集构建单元,用于将命令词作为单一命令词组以及将唤醒词与命令词组合为复合命令词组,并基于所述单一命令词组与所述复合命令词组构建命令词集;

[0031] 存储单元,用于存储所述命令词集。

[0032] 进一步地,在所述的人机交互系统中,还包括:

[0033] 执行单元,用于基于所述第一起始标识符与所述第一命令标识符之间的词组或所述唤醒标识符与所述第一命令标识符之间的词组,使所述智能设备执行与所述命令词组相对应的执行动作。

[0034] 进一步地,在所述的人机交互系统中,所述标识单元还用于在所述唤醒标识符之后设置第二起始标识符,且在所述语音数据中识别到命令词组的情况下,在所述命令词组

的词尾处设置第二命令标识符,其中,所述唤醒标识符与所述第二起始标识符之间具有一定的时间间隔。

[0035] 进一步地,在所述的人机交互系统中,所述执行单元还用于执行根据第二起始标识符与第二命令标识符之间的命令词组相对应的执行动作。

[0036] 本发明的第三个方面,提供了一种计算机设备,包括存储器、处理器以及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现如上任一项所述的方法。

[0037] 本发明的第四个方面,提供了一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现如上任一项所述的方法。

[0038] 相比于相关技术,本申请实施例提供的一种基于标识符的人机交互方法、系统、设备和存储介质,通过识别语音数据,在语音数据的起始处设置第一起始标识符;在语音数据中第一次识别到唤醒词的情况下,在唤醒词的词尾处设置唤醒标识符,且唤醒智能设备;在智能设备被唤醒的情况下,从第一起始标识符处开始再次识别语音数据;在语音数据中识别到命令词组的情况下,在命令词组的词尾处设置第一命令标识符;基于第一起始标识符与第一命令标识符之间的词组或唤醒标识符与第一命令标识符之间的词组,使智能设备执行与命令词组相对应的执行动作,在原始的“先唤醒再识别”的人机交互方式上,增加了one-shot人机交互方式,解决了因智能设备仅能支持“先唤醒再识别”人机交互方式或one-shot人机交互方式而导致的影响用户体验的问题,从而提高了用户的人机交互体验。

附图说明

[0039] 此处所说明的附图用来提供对本申请的进一步理解,构成本申请的一部分,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。在附图中:

[0040] 图1是根据本发明实施例的基于标识符的人机交互方法的流程图一;

[0041] 图2是根据本发明实施例的基于标识符的人机交互方法的流程图二;

[0042] 图3是根据本发明实施例的基于标识符的人机交互方法的流程图三;

[0043] 图4是根据本发明实施例的基于标识符的人机交互方法的流程图四;

[0044] 图5是根据本发明实施例的基于标识符的人机交互系统的结构框图一;

[0045] 图6是根据本发明实施例的基于标识符的人机交互系统的结构框图二;

[0046] 图7是根据本发明实施例的基于标识符的人机交互系统的结构框图三。

具体实施方式

[0047] 为了使本申请的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本申请进行描述和说明。应当理解,此处所描述的具体实施例仅仅用以解释本申请,并不用于限定本申请。基于本申请提供的实施例,本领域普通技术人员在没有作出创造性劳动的前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0048] 显而易见地,下面描述中的附图仅仅是本申请的一些示例或实施例,对于本领域的普通技术人员而言,在不付出创造性劳动的前提下,还可以根据这些附图将本申请应用于其他类似情景。此外,还可以理解的是,虽然这种开发过程中所作出的努力可能是复杂并且冗长的,然而对于与本申请公开的内容相关的本领域的普通技术人员而言,在本申请揭

露的技术内容的基础上进行的一些设计,制造或者生产等变更只是常规的技术手段,不当理解为本申请公开的内容不充分。

[0049] 在本申请中提及“实施例”意味着,结合实施例描述的特定特征、结构或特性可以包含在本申请的至少一个实施例中。在说明书中的各个位置出现该短语并不一定均是指相同的实施例,也不是与其它实施例互斥的独立的或备选的实施例。本领域普通技术人员显式地和隐式地理解的是,本申请所描述的实施例在不冲突的情况下,可以与其它实施例相结合。

[0050] 除非另作定义,本申请所涉及的技术术语或者科学术语应当为本申请所属技术领域内具有一般技能的人士所理解的通常意义。本申请所涉及的“一”、“一个”、“一种”、“该”等类似词语并不表示数量限制,可表示单数或复数。本申请所涉及的术语“包括”、“包含”、“具有”以及它们任何变形,意图在于覆盖不排他的包含;例如包含了一系列步骤或模块(单元)的过程、方法、系统、产品或设备没有限定于已列出的步骤或单元,而是可以还包括没有列出的步骤或单元,或可以还包括对于这些过程、方法、产品或设备固有的其它步骤或单元。本申请所涉及的“连接”、“相连”、“耦接”等类似的词语并非限定于物理的或者机械的连接,而是可以包括电气的连接,不管是直接的还是间接的。本申请所涉及的“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。字符“/”一般表示前后关联对象是一种“或”的关系。本申请所涉及的术语“第一”、“第二”、“第三”等仅仅是区别类似的对象,不代表针对对象的特定排序。

[0051] 本发明提供一种基于标识符的人机交互方法、系统、设备以及存储介质,可以应用于人机交互领域,尤其是可以应用在可以进行人机交互的智能设备上,如智能音箱。

[0052] 图1是根据本发明实施例的基于标识符的人机交互方法的流程图一,请参见图1,该方法包括:

[0053] 步骤S102,识别语音数据,在语音数据的起始处设置第一起始标识符;

[0054] 步骤S104,在语音数据中第一次识别到唤醒词的情况下,在唤醒词的词尾处设置唤醒标识符,且唤醒智能设备;

[0055] 步骤S106,在智能设备被唤醒的情况下,从第一起始标识符处开始再次识别语音数据;

[0056] 步骤S108,在语音数据中识别到命令词组的情况下,在命令词组的词尾处设置第一命令标识符;

[0057] 步骤S110,基于第一起始标识符与第一命令标识符之间的词组或唤醒标识符与第一命令标识符之间的词组,使智能设备执行与命令词组相对应的执行动作。

[0058] 其中,在智能设备没有被唤醒的情况下,智能设备一直处于待唤醒状态。此外,在智能设备执行与命令词组相对应的执行动作之后,智能设备可以切换到待唤醒状态,使智能设备等待下一次被唤醒;也可以继续保持唤醒状态,等待识别用户的语音数据。

[0059] 在一些实施例中,语音数据可以包括唤醒词和命令词组。例如,在智能音箱待识别的语音数据为“你好开灯”的情况下,智能音箱在“你好开灯”的起始处设置第一起始标识符a,在“你好”的词尾处设置唤醒标识符b,即“a你好b开灯”;此时智能音箱由于识别到唤醒词“你好”,其由待唤醒状态切换到唤醒状态,然后智能音箱从第一起始标识符a处再次识别该

语音数据,在识别到“开灯”的情况下,在“开灯”的词尾处设置第一命令标识符c,即“a你好b开灯c”;基于唤醒标识符b与第一命令标识符c之间的词组,即“开灯”,或者基于第一起始标识符a与第一命令标识符c之间的词组,即“你好开灯”,智能音箱执行与“开灯”或“你好开灯”相对应的执行动作。

[0060] 在一些实施例中,在智能音箱维持在唤醒状态的情况下,若智能音箱接收到命令词组,则智能音箱可以执行与该命令词组相对应的动作。

[0061] 例如,在智能音箱处于唤醒状态的情况下,若智能音箱接收到的命令词组为“开灯”,此时智能音箱在“开灯”的起始处设置第一起始标识符a,在“开灯”的词尾处设置第一命令标识符c,智能音箱基于第一起始标识符a与第一命令标识符c之间的词组执行与词组相对应的执行动作。

[0062] 通过步骤S102至步骤S110,解决了现有的智能设备不能同时支持“先唤醒再识别”和“one-shot”这两种人机交互方式而导致影响用户交互体验的问题。

[0063] 图2是根据本发明实施例的基于标识符的人机交互方法的流程图二,请参见图2,在智能设备被唤醒的情况下,从第一起始标识符开始再次识别语音数据之后,该方法还包括如下步骤:

[0064] 步骤S202,在再次识别到唤醒词的情况下,继续识别语音数据。

[0065] 在一些实施例中,在智能设备被唤醒的情况下,由于需要从第一起始标识符再次识别语音数据,在识别到唤醒标识符的情况下,会再次识别到唤醒词,此时可以将再次识别到的唤醒词作为无效唤醒词,并将再次识别到的唤醒词过滤掉,然后继续识别语音数据。

[0066] 例如,在智能音箱识别的语音数据为“你好开灯”的情况下,在第一次识别到唤醒词“你好”的情况下,在该语音数据的起始处设置第一起始标识符a,在唤醒词的词尾处设置唤醒标识符b,即“a你好b开灯”,且智能音箱被唤醒;然后智能设备从第一起始标识符a对该语音数据进行再次识别;在唤醒标识符b处识别到唤醒词“你好”,此时智能音箱将唤醒词“你好”作为无效唤醒词,并将再次识别到的唤醒词“你好”过滤掉,并继续识别语音数据。

[0067] 在一些实施例中,在智能设备被唤醒且再次识别语音数据的情况下,若此时识别到唤醒词,智能设备不执行任何动作,继续处于唤醒状态。

[0068] 例如,在智能音箱识别的语音数据为“你好”的情况下,在“你好”的起始处设置第一起始标识符a,在“你好”的词尾处设置唤醒标识符b,即“a你好b”,且智能音箱被唤醒;智能音箱从第一起始标识符a处再次识别该语音数据,在智能音箱识别到唤醒标识符b的情况下,智能音箱继续识别语音数据。

[0069] 在一些实施例中,在智能设备被唤醒且再次识别语音数据的情况下,若此时识别到唤醒词,但是没有识别到命令词组,此时智能设备可以对该语音数据进行回复,且使智能设备维持唤醒状态。

[0070] 例如,在智能音箱识别的语音数据为“你好”的情况下,在“你好”的起始处设置第一起始标识符a,在“你好”的词尾处设置唤醒标识符b,即“a你好b”,且智能音箱被唤醒;智能音箱从第一起始标识符a处再次识别该语音数据,若智能音箱识别到唤醒标识符b之后无任何命令词组,此时智能音箱可以回复“在的”,且使智能音箱维持在唤醒状态,等待用户继续输入下一语音数据。

[0071] 在一些实施例中,在智能设备再次识别语音数据且识别到唤醒标识符的情况下,

若在唤醒标识符之后再次识别到唤醒词,则继续识别语音数据。例如,在智能音箱接收到的语音数据为“你好···你好开灯”的情况下,在该语音数据的起始处设置第一起始标识符a,在第一次出现的唤醒词的词尾处设置唤醒标识符b,即“a你好b···你好开灯”,且唤醒智能音箱;然后从第一起始标识符处再次识别该语音数据;在识别到唤醒标识符b之前的唤醒词“你好”的情况下,智能音箱可以经过一定的时间间隔继续识别该语音数据,且在一定的时间间隔后在识别到第二个唤醒词“你好”情况下,继续识别该语音数据。

[0072] 通过步骤S202,解决了在智能设备唤醒的情况下,若再次识别到语音数据中的唤醒词,智能设备会被多次唤醒的问题,从而提高了用户的交互体验。

[0073] 图3是根据本发明实施例的基于标识符的人机交互方法的流程图三,请参见图3,在智能设备被唤醒的情况下,从第一起始标识符开始再次识别语音数据之后,该方法还包括:

[0074] 步骤S302,在唤醒标识符之后,设置第二起始标识符,其中,唤醒标识符与第二起始标识符之间具有一定的时间间隔;

[0075] 步骤S304,在语音数据中识别到命令词组的情况下,在命令词组的词尾处设置第二命令标识符;

[0076] 步骤S306,基于第二起始标识符与第二命令标识符之间的词组,使智能设备执行与命令词组相对应的执行动作。

[0077] 其中,唤醒标识符与第二起始标识符之间具有一定的时间间隔可以设置为任意的时间间隔,如可以设置为10秒、20秒、一分钟以及两分钟。

[0078] 在一些实施例中,第二起始标识符与第二命令标识符之间的语音数据可以不包括唤醒词。例如,在智能音箱识别的语音数据为“你好···开灯”的情况下,此时智能音箱在语音数据的起始处设置第一起始标识符a,在唤醒词“你好”的词尾处设置唤醒标识符b,即“a你好b···开灯”,其中“···”用于指示一定的时间间隔;然后智能音箱在一定的时间间隔之后的“开灯”的起始处设置第二起始标识符d,在一定的时间间隔之后的语音数据“开灯”的词尾处设置第二命令标识符f,即“a你好b···d开灯f”,然后基于第二起始标识符d与第二命令标识符f之间的词组“开灯”,使智能音箱执行与“开灯”相对应的执行动作。

[0079] 在一些实施例中,第二起始标识符与第二命令标识符之间的语音数据可以包括唤醒词,例如,在智能音箱识别到的语音数据为“你好···你好开灯”的情况下,智能音箱在语音数据的起始处设置第一起始标识符a,在唤醒词的词尾处设置唤醒标识符b,即“a你好b···开灯”,且唤醒智能音箱,“···”用于指示一定的时间间隔;经过一定的时间间隔之后,在“你好开灯”的起始处设置第二起始标识符d,在“你好开灯”的词尾处设置第二命令标识符f,智能设备根据第二起始标识符d与第二命令标识符f之间的词组“你好开灯”使智能设备执行与词组相对应的执行动作。

[0080] 通过步骤S302至步骤S306,实现了智能设备在被唤醒后能够在一段时间内仍然维持接收语音数据的状态,并能够接收的语音数据,较好地提高了用户体验。

[0081] 图4是根据本发明实施例的基于标识符的人机交互方法的流程图四,请参见图4,在识别语音数据之前,所述方法还包括:

[0082] 步骤S402,获取唤醒词和命令词;

[0083] 步骤S404,将命令词作为单一命令词组,且将唤醒词与命令词组合为复合命令词

组；

[0084] 步骤S406,基于单一命令词组与复合命令词组构建命令词集,其中,单一命令词组与复合命令词组所对应的执行动作相同。

[0085] 其中,获取唤醒词可以为获取能够唤醒智能设备的至少一个唤醒词,如天猫精灵的唤醒词为“你好天猫”、“天猫精灵”,获取命令词为获取用户的所有的常用命令词,如“播放音乐”、“开灯”、“播报天气”等常用命令词。

[0086] 其中,将唤醒词和命令词组合为复合命令词组可以为“唤醒词+命令词”,也可以为“命令词+唤醒词”,如在唤醒词为“你好”、命令词为“开灯”的情况下,可以将唤醒词和命令词组合为“你好开灯”或“开灯你好”。

[0087] 例如,在语音数据包括复合命令词组“你好开灯”的情况下,则智能设备执行开灯动作。

[0088] 通过步骤S402至步骤S406,解决了智能设备在接收到复合命令词组的情况下,执行的动作与复合命令词组相对应的执行动作不同的问题,进而提高了人机交互体验。

[0089] 对应于上述的一种基于标识符的人机交互方法,在本实施例中,还提供了一种基于标识符的人机交互系统,该系统用于实现上述实施例及优选实施例,已经进行过说明的不在赘述。

[0090] 图5是根据本发明实施例的基于标识符的人机交互系统的结构框图一,请参见图5,该系统包括:

[0091] 语音采集单元51,用于采集语音数据;

[0092] 语音识别单元52,用于识别语音数据;

[0093] 标识单元53,用于在语音数据的起始处设置第一起始标识符,且在语音数据中第一次出现唤醒词的情况下,在唤醒词的词尾处设置唤醒标识符,在语音数据中出现命令词组的情况下,在命令词组的词尾处设置第一命令标识符;

[0094] 唤醒单元54,用于在检测到唤醒标识符的情况下,唤醒智能设备;

[0095] 执行单元56,用于基于第一起始标识符与第一命令标识符之间的词组或唤醒标识符与第一命令标识符之间的词组,使智能设备执行与词组相对应的执行动作。

[0096] 其中,语音采集单元51用于采集用户的语音数据并将采集的语音数据发送到语音识别单元52;语音识别单元52对该语音数据进行识别;在语音识别单元52第一次识别到唤醒词的情况下,语音识别单元52发送标识指令到标识单元53,标识单元53在语音数据的起始处设置第一起始标识符,在唤醒词的词尾处设置唤醒标识符;唤醒单元54根据该唤醒标识符,唤醒智能设备;且在语音识别单元52识别到命令词组的情况下,语音识别单元52发送标识指令到标识单元53,标识单元53在命令词组的词尾处设置第一命令标识符;执行单元56根据标识单元53标识的第一起始标识符和第一命令标识符之间的词组执行与词组相对应的执行动作。

[0097] 在一些实施例中,标识单元53可以只在语音识别单元52第一次识别到唤醒词的词尾处设置唤醒标识符;标识单元53也可以在语音识别单元52每识别到一次唤醒词,在每一个唤醒词的词尾处均设置唤醒标识符,如第一次识别的唤醒词的词尾处设置第一唤醒标识符,在第二次识别的唤醒词的词尾处设置第二唤醒标识符,在第n次识别的唤醒词的词尾处设置第n唤醒标识符。

[0098] 优选地,标识单元53只在第一次识别到的唤醒词的词尾处设置唤醒标识符。

[0099] 其中,执行单元56根据语音识别单元52识别到的命令词组去执行与命令词组相对应的执行动作,如在语音识别单元52识别到的命令词组为“开灯”的情况下,执行单元56根据该命令词组“开灯”执行相对应的执行动作。

[0100] 在一些实施例中,标识单元53还用于在唤醒标识符之后设置第二起始标识符,其中,唤醒标识符与第二起始标识符之间具有一定的时间间隔;并且在语音数据中识别到命令词组的情况下,标识单元53还用于在命令词组的词尾处设置第二命令标识符,从而使执行单元56能够根据第二起始标识符与第二命令标识符之间的词组执行与词组相对应的执行动作。

[0101] 通过语音采集单元51、语音识别单元52、标识单元53、唤醒单元54以及执行单元56组成的基于标识符的人机交互系统,实现了智能设备能够同时兼容“先唤醒再识别”和“one-shot”人机交互方式的效果,极大提高了人机交互体验。

[0102] 图6是根据本发明实施例的基于标识符的人机交互系统的结构框图二,请参见图6,该系统还包括:

[0103] 维持单元55,用于在智能设备被唤醒的情况下,使智能设备维持在唤醒状态。

[0104] 在一些实施例中,维持单元55能够在智能设备被唤醒的情况下,维持单元能够使智能设备在唤醒状态维持一定的时间间隔;若在该时间间隔内,用户仍然未说出语音数据,则智能设备切换到待唤醒状态。

[0105] 例如,在用户说出唤醒词“你好”使智能音箱被唤醒的情况下,维持单元55能够使智能音箱维持五分钟的唤醒状态;若在五分钟内,智能音箱再次接收到语音数据,则智能音箱可以按照预设逻辑继续响应用户;若在五分钟之后,智能音箱仍然未接收到用户的语音数据,则此时智能音箱从唤醒状态切换为待唤醒状态。

[0106] 通过维持单元55能够使智能设备维持唤醒状态,从而解决了用户在智能设备被唤醒的情况下,不能及时说出语音数据而导致智能设备切换为待唤醒状态的问题,进而提高了用户的人机交互体验。

[0107] 图7是根据本发明实施例的基于标识符的人机交互系统的结构框图三,请参见图7,该系统还包括:

[0108] 命令词集构建单元57,用于将命令词作为单一命令词组以及将唤醒词和命令词组合为复合命令词组,并基于单一命令词组与复合命令词组构建命令词集;

[0109] 存储单元58,用于存储命令词集。

[0110] 其中,单一命令词组所对应的执行动作与复合命令词组所对应的执行动作设置为同一执行动作。例如,在单一命令词组为“开灯”、复合命令词组为“你好开灯”的情况下,单一命令词组“开灯”与复合命令词组“你好开灯”所对应的执行动作均为开灯。

[0111] 其中,存储单元58不仅可以用于存储命令词集,也可以用于存储唤醒词。

[0112] 通过命令词集构建单元57和存储单元58,能够更好的让智能设备兼容“先唤醒再识别”和“one-shot”人机交互方式的效果,提高人机交互体验。

[0113] 另外,本申请实施例的基于标识符的人机交互方法可以由计算机设备来实现,计算机设备包括但不限于处理器以及存储有计算机程序指令的存储器。

[0114] 在一些实施例中,处理器可以包括中央处理器(CPU),或者特定集成电路

(Application Specific Integrated Circuit, 简称为ASIC), 或者可以被配置成实施本申请实施例的一个或多个集成电路。

[0115] 在一些实施例中, 存储器可以包括用于数据或指令的大容量存储器。举例来说而非限制, 存储器可包括硬盘驱动器 (Hard DiSk Drive, 简称为HDD)、软盘驱动器、固态驱动器 (Solid State Drive, 简称为SSD)、闪存、光盘、磁光盘、磁带或通用串行总线 (UniverSal Serial BuS, 简称为USB) 驱动器或者两个或更多个以上这些的组合。在合适的情况下, 存储器可包括可移除或不可移除 (或固定) 的介质。在合适的情况下, 存储器可在数据处理装置的内部或外部。在特定实施例中, 存储器是非易失性 (Non-Volatile) 存储器。在特定实施例中, 存储器包括只读存储器 (Read-Only Memory, 简称为ROM) 和随机存取存储器 (Random AcceSS Memory, 简称为RAM)。在合适的情况下, 该ROM可以是掩模编程的ROM、可编程ROM (Programmable Read-Only Memory, 简称为PROM)、可擦除PROM (EraSable Programmable Read-Only Memory, 简称为EPROM)、电可擦除PROM (Electrically EraSable Programmable Read-Only Memory, 简称为EEPROM)、电可改写ROM (Electrically Alterable Read-Only Memory, 简称为EAROM) 或闪存 (FLASH) 或者两个或更多个以上这些的组合。在合适的情况下, 该RAM可以是静态随机存取存储器 (Static Random-AcceSS Memory, 简称为SRAM) 或动态随机存取存储器 (Dynamic Random AcceSS Memory, 简称为DRAM), 其中, DRAM可以是快速页模式动态随机存取存储器 (FaSt Page Mode Dynamic Random AcceSS Memory, 简称为FPMDRAM)、扩展数据输出动态随机存取存储器 (Extended Date Out Dynamic Random AcceSS Memory, 简称为EDODRAM)、同步动态随机存取内存 (SynchrouS Dynamic Random-AcceSS Memory, 简称SDRAM) 等。

[0116] 存储器可以用来存储或者缓存需要处理和/或通信使用的各种数据文件, 以及处理器所执行的可能的计算机程序指令。

[0117] 处理器通过读取并执行存储器中存储的计算机程序指令, 以实现上述实施例中的任意一种基于标识符的人机交互方法。

[0118] 在其中一些实施例中, 计算机设备还可包括通信接口和总线。其中, 处理器、存储器、通信接口通过总线连接并完成相互间的通信。

[0119] 通信接口用于实现本申请实施例中各模块、装置、单元和/或设备之间的通信。通信接口还可以实现与其他部件例如: 外接设备、图像/数据采集设备、数据库、外部存储以及图像/数据处理工作站等之间进行数据通信。

[0120] 总线包括硬件、软件或两者, 将计算机设备的部件彼此耦接在一起。总线包括但不限于以下至少之一: 数据总线 (Data BuS)、地址总线 (AddreSS BuS)、控制总线 (Control BuS)、扩展总线 (ExpanSion BuS)、局部总线 (Local BuS)。举例来说而非限制, 总线可包括图形加速接口 (Accelerated GraphicS Port, 简称为AGP) 或其他图形总线、增强工业标准架构 (Extended InduSty Standard Architecture, 简称为EISA) 总线、前端总线 (Front Side BuS, 简称为FSB)、超传输 (Hyper TranSport, 简称为HT) 互连、工业标准架构 (InduSty Standard Architecture, 简称为ISA) 总线、无线带宽 (InfiniBand) 互连、低引脚数 (Low Pin Count, 简称为LPC) 总线、存储器总线、微信道架构 (Micro Channel Architecture, 简称为MCA) 总线、外围组件互连 (Peripheral Component Interconnect, 简称为PCI) 总线、PCI-ExpreSS (PCI-X) 总线、串行高级技术附件 (Serial Advanced

Technology Attachment, 简称为SATA) 总线、视频电子标准协会局部 (Video ElectronicS StandardS ASSociation Local BuS, 简称为VLB) 总线或其他合适的总线或者两个或更多个以上这些的组合。在合适的情况下, 总线可包括一个或多个总线。尽管本申请实施例描述和示出了特定的总线, 但本申请考虑任何合适的总线或互连。

[0121] 该计算机设备可以基于语音数据、第一起始标识符、唤醒标识符以及第一命令标识符执行本申请实施例的基于标识符的人机交互方法, 从而实现结合图1描述的方法。

[0122] 另外, 结合上述实施例中的基于标识符的人机交互方法, 本申请实施例可提供一种计算机可读存储介质来实现。该计算机可读存储介质上存储有计算机程序指令; 该计算机程序指令被处理器执行时能够实现上述实施例中的任意一种基于标识符的人机交互方法。

[0123] 以上所述实施例的各技术特征可以进行任意的组合, 为使描述简洁, 未对上述实施例中的各个技术特征所有可能的组合都进行描述, 然而, 只要这些技术特征的组合不存在矛盾, 都应当认为是本说明书记载的范围。

[0124] 以上所述实施例仅表达了本申请的几种实施方式, 其描述较为具体和详细, 但并不能因此而理解为对发明专利范围的限制。应当指出的是, 对于本领域的普通技术人员来说, 在不脱离本申请构思的前提下, 还可以做出若干变形和改进, 这些都属于本申请的保护范围。因此, 本申请专利的保护范围应以所附权利要求为准。

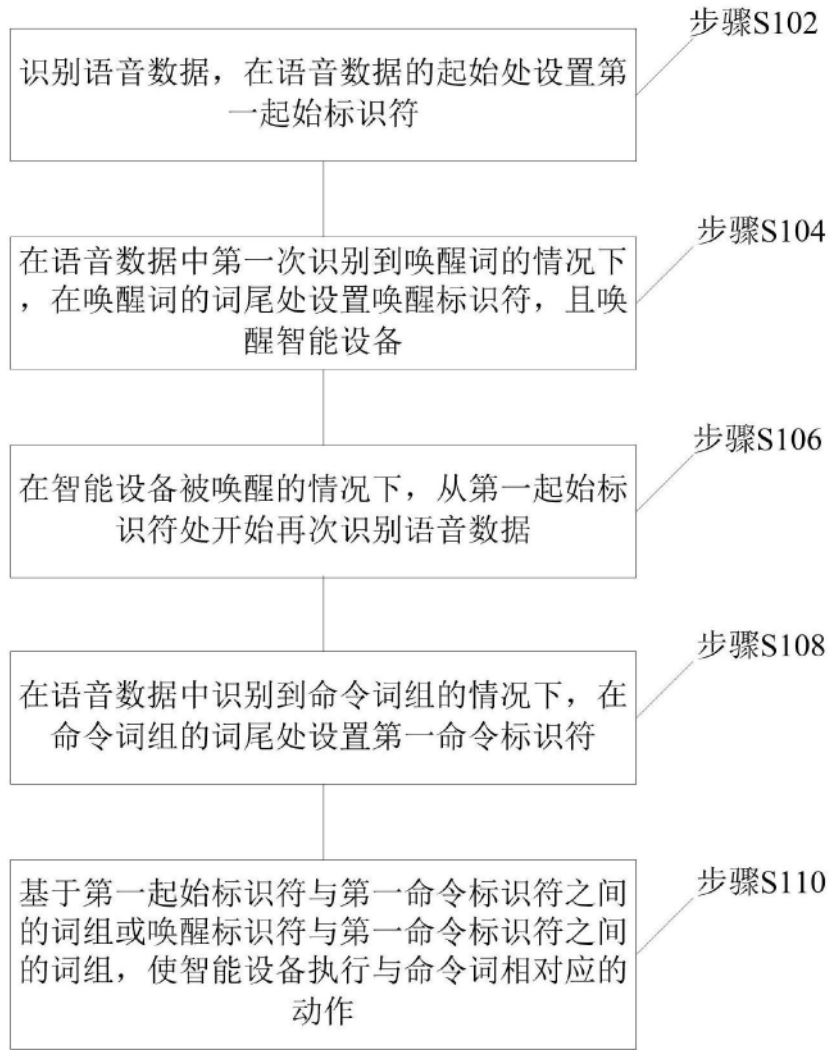


图1

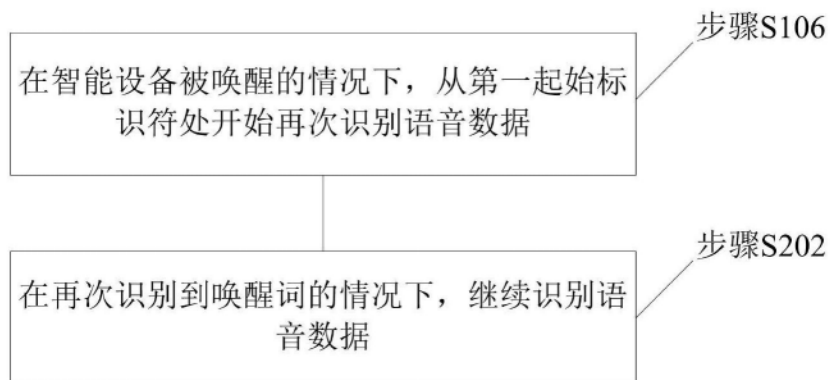


图2

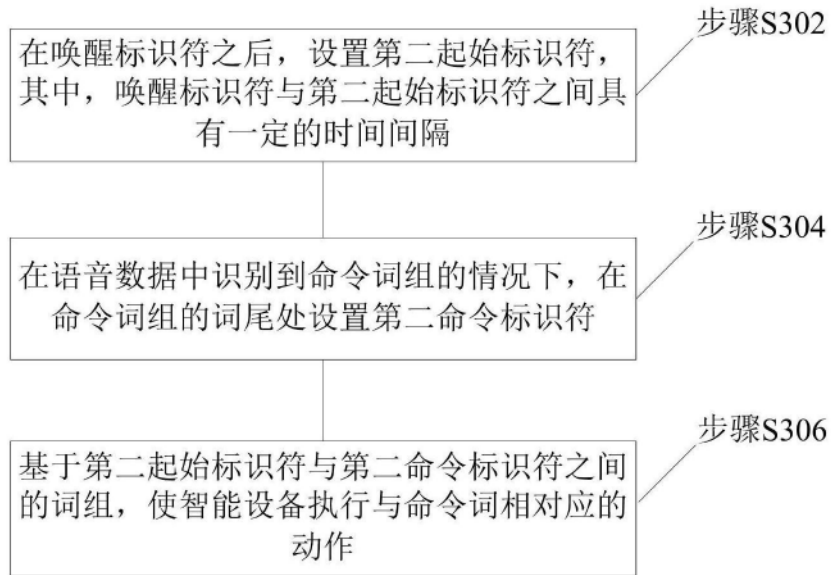


图3

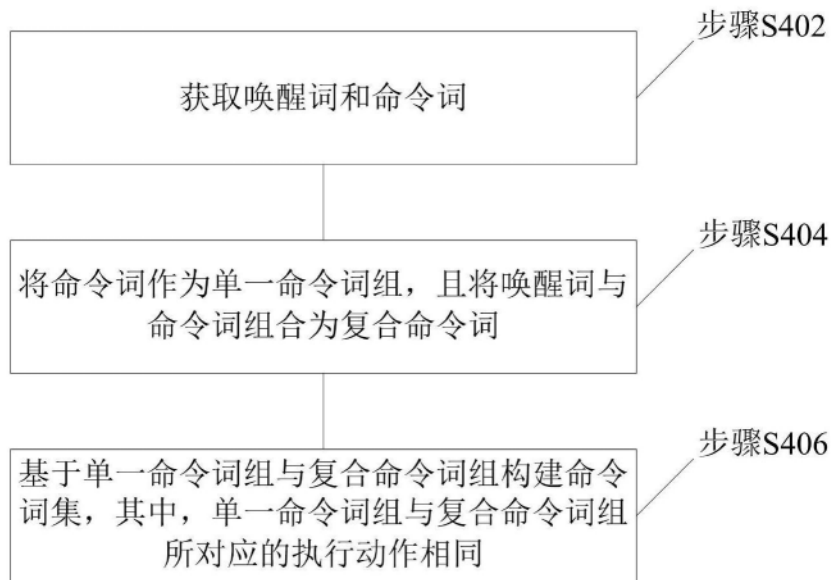


图4

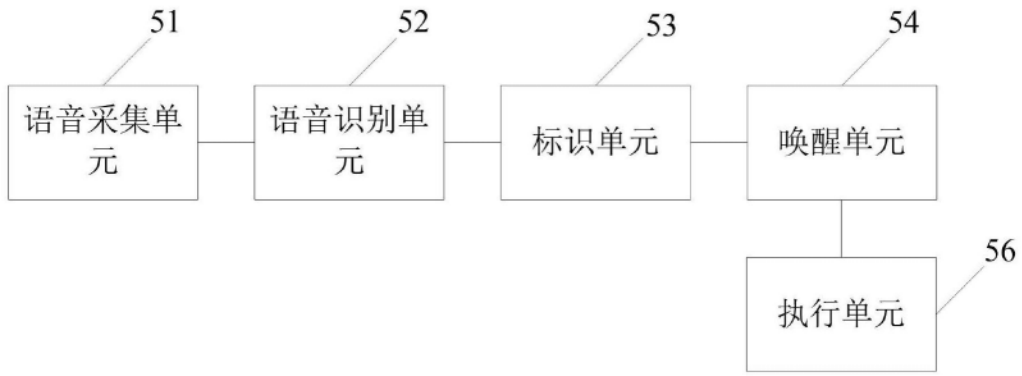


图5

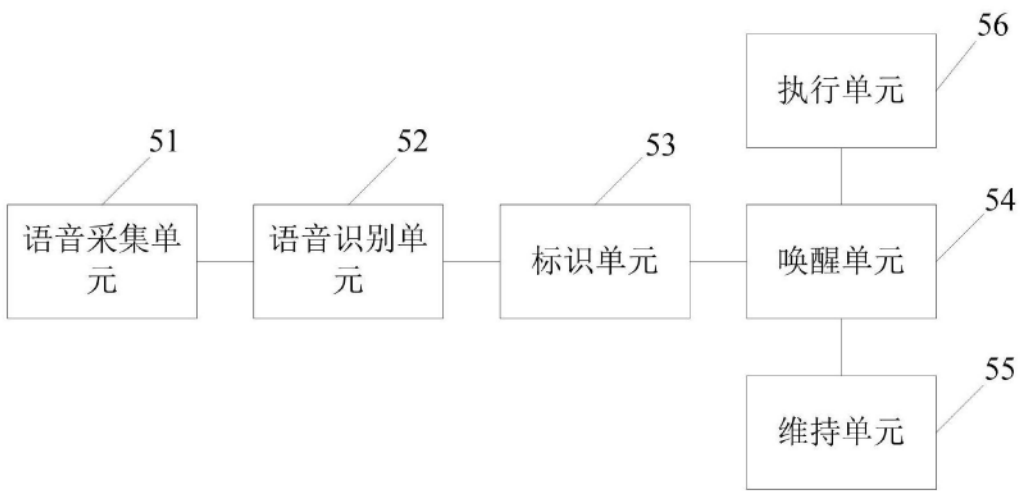


图6

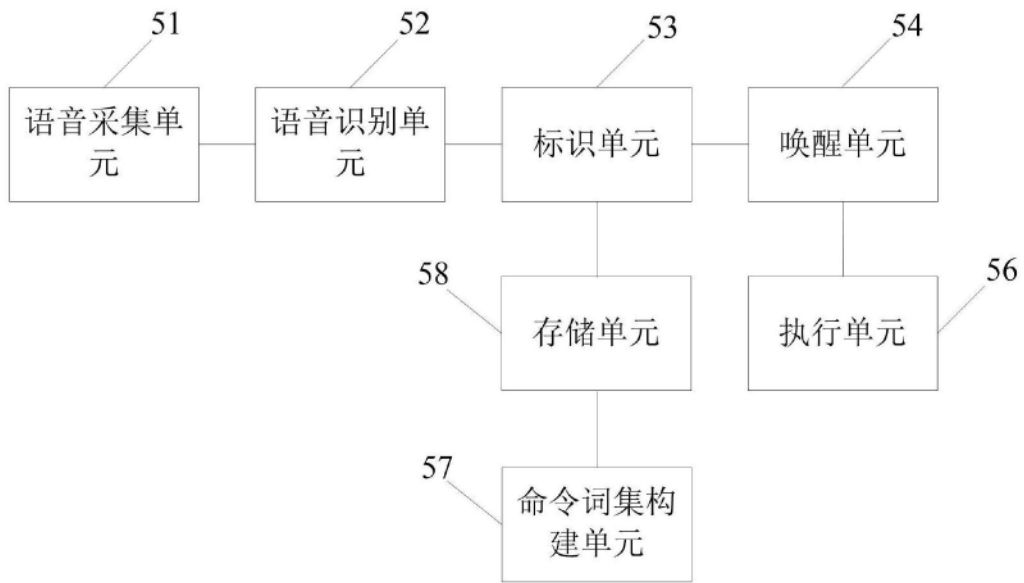


图7