



(12) 发明专利

(10) 授权公告号 CN 115881126 B

(45) 授权公告日 2023.05.05

(21) 申请号 202310148494.7

G10L 15/06 (2013.01)

(22) 申请日 2023.02.22

G10L 21/0208 (2013.01)

(65) 同一申请的已公布的文献号

审查员 安鹏飞

申请公布号 CN 115881126 A

(43) 申请公布日 2023.03.31

(73) 专利权人 广东浩博特科技股份有限公司

地址 516000 广东省惠州市惠阳区秋长镇

新塘村浩博特工业园

(72) 发明人 孙红升

(74) 专利代理机构 深圳汉林汇融知识产权代理

事务所(普通合伙) 44850

专利代理师 吴洪波

(51) Int. Cl.

G10L 15/22 (2006.01)

G10L 15/02 (2006.01)

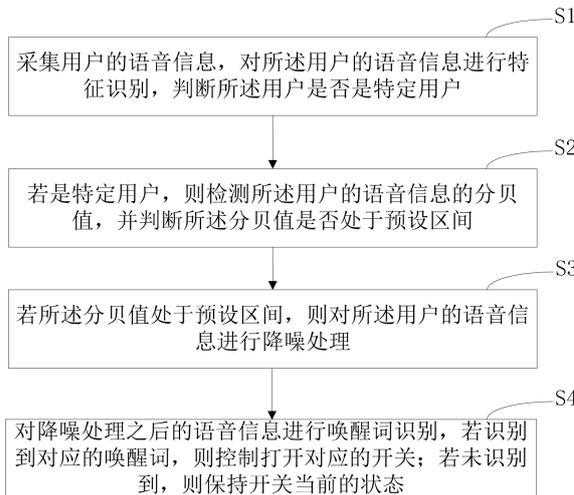
权利要求书4页 说明书9页 附图2页

(54) 发明名称

基于语音识别的开关控制方法、装置和开关设备

(57) 摘要

本发明涉及智能家居技术领域,提供了一种基于语音识别的开关控制方法、装置和开关设备,包括:采集到用户的语音信息之后,对所述用户的语音信息进行特征识别,判断所述用户是否是特定用户;若是特定用户,则检测所述用户的语音信息的分贝值,并判断分贝值是否处于预设区间;若所述分贝值处于预设区间,则对用户的语音信息进行降噪处理;对降噪处理之后的语音信息进行唤醒词识别,若识别到对应的唤醒词,则控制打开对应的开关;若未识别到,则保持开关当前的状态;其中,用于降噪处理的降噪模型以及用于唤醒识别的唤醒模型预先通过联合训练得到。本发明通过多个检测过程之后再行降噪、唤醒识别,降低了智能家居的开关被误唤醒的概率。



1. 一种基于语音识别的开关控制方法,其特征在于,包括以下步骤:

采集用户的语音信息,对所述用户的语音信息进行特征识别,判断所述用户是否是特定用户;

若是特定用户,则检测所述用户的语音信息的分贝值,并判断所述分贝值是否处于预设区间;

若所述分贝值处于预设区间,则对所述用户的语音信息进行降噪处理;

对降噪处理之后的语音信息进行唤醒词识别,若识别到对应的唤醒词,则控制打开对应的开关;若未识别到,则保持开关当前的状态;其中,用于降噪处理的降噪模型以及用于唤醒识别的唤醒模型预先通过联合训练得到;

所述采集用户的语音信息,对所述用户的语音信息进行特征识别,判断所述用户是否是特定用户的步骤之前,包括:

获取训练语音数据,对所述训练语音数据进行特征提取,得到第一语音特征向量;其中,所述训练语音数据包括对应的标签;

将所述第一语音特征向量输入至预设的编码模型中进行特征提取,得到第一编码向量;

将所述第一编码向量输入至预设的解码模型中进行解码处理,得到对应的第二语音特征向量;

将所述第二语音特征向量输入至预设的编码模型中进行特征提取,得到第二编码向量;

将所述第一编码向量、第二编码向量及对应的标签输入至分类层中进行训练,并迭代调整所述预设的编码模型、解码模型的模型参数,直至所述分类层的损失函数收敛之后,完成模型的训练;其中,将所述编码模型与所述分类层组合作为唤醒模型,用于对降噪处理之后的语音信息进行唤醒词识别;将所述编码模型与所述解码模型作为降噪模型,用于对用户的语音信息进行降噪处理。

2. 根据权利要求1所述的基于语音识别的开关控制方法,其特征在于,所述控制打开对应的开关的步骤,包括:

对所述用户的语音信息进行声源定位,获取所述用户距离开关的距离;

判断所述距离是否小于第一预设距离;

若小于所述第一预设距离,则控制立即打开对应的开关;

若不小于所述第一预设距离,则判断所述距离是否处于第一预设距离与第二预设距离之间;

若处于,则获取所述第一预设距离与第二预设距离的区间所对应的延迟时间;基于所述延迟时间,控制所述开关在延迟时间之后打开;

若不处于,则控制所述开关在预设时间之后打开;其中,所述预设时间大于所述延迟时间。

3. 根据权利要求1所述的基于语音识别的开关控制方法,其特征在于,所述控制打开对应的开关的步骤,包括:

基于所述用户的语音信息,获取所述用户的朝向区域;

根据各个区域中的开关列表,获取所述朝向区域内所有的预选开关;

获取各个所述预选开关的唤醒词,从所述预选开关中确定出唤醒词为对应识别到的唤醒词的开关,作为第一开关;

判断所述第一开关是否为关闭状态,若是,则打开所述第一开关;

获取第一开关的标识信息,对所述标识信息进行哈希计算得到对应的标识哈希值;在所述开关接收到关闭指令关闭时,将所述标识哈希值与所述用户的语音信息进行关联存储。

4. 根据权利要求1所述的基于语音识别的开关控制方法,其特征在于,所述对所述用户的语音信息进行降噪处理的步骤之前,包括:

采集当前的环境声音;

将所述环境声音输入至预设的环境检测模型中进行检测,得到对应的环境检测结果;其中,所述环境检测模型是预先训练完成的深度学习模型;

根据所述环境检测结果,在数据库中确定对应的目标模型参数集合;其中,数据库中存储有环境检测结果与模型参数集合的对应关系,所述模型参数集合中包括多个模型参数:平滑机制参数,降噪队列长度,降噪阈值大小;

获取初始的降噪模型,提取出所述目标模型参数集合中的各个目标模型参数,并将所述初始的降噪模型中的模型参数更新为对应的目标模型参数,得到更新的降噪模型;其中,所述降噪模型用于对所述用户的语音信息进行降噪处理。

5. 根据权利要求1所述的基于语音识别的开关控制方法,其特征在于,所述对降噪处理之后的语音信息进行唤醒词识别,若识别到对应的唤醒词,则控制打开对应的开关的步骤之前,包括:

获取第一语音训练样本,其中,所述第一语音训练样本为包括唤醒词的语音数据;

获取第二语音训练样本,其中,所述第二语音训练样本为不包括唤醒词的语音数据,且所述第二语音训练样本的语音数据中包括多个与唤醒词语音相近的词;

将所述第一语音训练样本输入至预设的唤醒模型中,得到所述唤醒模型被唤醒的唤醒率;

将所述第二语音训练样本输入至预设的唤醒模型中,得到所述唤醒模型被误唤醒的误唤醒率;

判断所述唤醒率是否大于对应的第一阈值,以及判断所述误唤醒率是否大于对应的第二阈值;

若所述唤醒率大于对应的第一阈值,且所述误唤醒率大于第二阈值,则增加所述唤醒模型被唤醒词唤醒的唤醒阈值以及降低平滑队列长度;

若所述唤醒率不大于对应的第一阈值,且所述误唤醒率大于第二阈值,则增加所述唤醒模型的平滑队列长度;

若所述唤醒率不大于对应的第一阈值,且所述误唤醒率不大于第二阈值,则降低所述唤醒模型被唤醒词唤醒的唤醒阈值。

6. 根据权利要求5所述的基于语音识别的开关控制方法,其特征在于,所述将所述第一语音训练样本输入至预设的唤醒模型中,得到所述唤醒模型被唤醒的唤醒率的步骤,包括:

对所述第一语音训练样本进行语音识别,识别出所述第一语音训练样本中的各个唤醒词在第一语音训练样本中的时间区间,记为唤醒词时间区间;

在预设的文档中记录所述第一语音训练样本各个唤醒词及其对应的唤醒词时间区间；
将所述第一语音训练样本输入至预设的唤醒模型中进行唤醒识别，并在所述唤醒模型开始识别时，开始计时；

在所述唤醒模型识别到各个唤醒词时，记录对应的唤醒时间区间；

将记录唤醒时间区间与所述预设的文档中记录的各个唤醒词及其对应的唤醒词时间区间进行对比；当所述唤醒时间区间与所述唤醒词时间区间重叠时，作为一次正确唤醒次数；当所述唤醒时间区间与所述唤醒词时间区间不重叠时，作为一次错误唤醒次数；

根据所述正确唤醒次数、错误唤醒次数计算所述唤醒模型被唤醒的唤醒率；其中，所述唤醒率=正确唤醒次数/(正确唤醒次数+错误唤醒次数)。

7. 一种基于语音识别的开关控制装置，其特征在于，包括：

采集单元，用于采集用户的语音信息，对所述用户的语音信息进行特征识别，判断所述用户是否是特定用户；

检测单元，用于若是特定用户，则检测所述用户的语音信息的分贝值，并判断所述分贝值是否处于预设区间；

降噪单元，用于若所述分贝值处于预设区间，则对所述用户的语音信息进行降噪处理；

控制单元，用于对降噪处理之后的语音信息进行唤醒词识别，若识别到对应的唤醒词，则控制打开对应的开关；若未识别到，则保持开关当前的状态；其中，用于降噪处理的降噪模型以及用于唤醒识别的唤醒模型预先通过联合训练得到；

所述采集单元，还用于：

获取训练语音数据，对所述训练语音数据进行特征提取，得到第一语音特征向量；其中，所述训练语音数据包括对应的标签；

将所述第一语音特征向量输入至预设的编码模型中进行特征提取，得到第一编码向量；

将所述第一编码向量输入至预设的解码模型中进行解码处理，得到对应的第二语音特征向量；

将所述第二语音特征向量输入至预设的编码模型中进行特征提取，得到第二编码向量；

将所述第一编码向量、第二编码向量及对应的标签输入至分类层中进行训练，并迭代调整所述预设的编码模型、解码模型的模型参数，直至所述分类层的损失函数收敛之后，完成模型的训练；其中，将所述编码模型与所述分类层组合作为唤醒模型，用于对降噪处理之后的语音信息进行唤醒词识别；将所述编码模型与所述解码模型作为降噪模型，用于对用户的语音信息进行降噪处理。

8. 根据权利要求7所述的基于语音识别的开关控制装置，其特征在于，还包括：

获取单元，用于获取训练语音数据，对所述训练语音数据进行特征提取，得到第一语音特征向量；其中，所述训练语音数据包括对应的标签；

第一编码单元，用于将所述第一语音特征向量输入至预设的编码模型中进行特征提取，得到第一编码向量；

第一解码单元，用于将所述第一编码向量输入至预设的解码模型中进行解码处理，得到对应的第二语音特征向量；

第二编码单元,用于将所述第二语音特征向量输入至预设的编码模型中进行特征提取,得到第二编码向量;

训练单元,用于将所述第一编码向量、第二编码向量及对应的标签输入至分类层中进行训练,并迭代调整所述预设的编码模型、解码模型的模型参数,直至所述分类层的损失函数收敛之后,完成模型的训练;其中,将所述编码模型与所述分类层组合作为唤醒模型,用于对降噪处理之后的语音信息进行唤醒词识别;将所述编码模型与所述解码模型作为降噪模型,用于对用户的语音信息进行降噪处理。

9.一种开关设备,包括存储器和处理器,所述存储器中存储有计算机程序,其特征在于,所述处理器执行所述计算机程序时实现权利要求1至6中任一项所述方法的步骤。

基于语音识别的开关控制方法、装置和开关设备

技术领域

[0001] 本发明涉及智能家居技术领域,特别涉及一种基于语音识别的开关控制方法、装置、开关设备和存储介质。

背景技术

[0002] 随着物联网技术的发展,目前的家居设备都在往智能化发展,智能家居逐渐侵入家庭生活中;目前的智能家居设备中通常采用物联网形式,基于远程通讯进行远程控制开关,或者通过简单的语音唤醒词进行开关控制;这种简单的唤醒模式容易造成误唤醒,使得开关被随意打开,给用户带来不必要的烦恼。

发明内容

[0003] 本发明的主要目的为提供一种基于语音识别的开关控制方法、装置、开关设备和存储介质,旨在降低智能家居的开关被误唤醒的概率。

[0004] 为实现上述目的,本发明提供了一种基于语音识别的开关控制方法,包括以下步骤:

[0005] 采集用户的语音信息,对所述用户的语音信息进行特征识别,判断所述用户是否是特定用户;

[0006] 若是特定用户,则检测所述用户的语音信息的分贝值,并判断所述分贝值是否处于预设区间;

[0007] 若所述分贝值处于预设区间,则对所述用户的语音信息进行降噪处理;

[0008] 对降噪处理之后的语音信息进行唤醒词识别,若识别到对应的唤醒词,则控制打开对应的开关;若未识别到,则保持开关当前的状态;其中,用于降噪处理的降噪模型以及用于唤醒识别的唤醒模型预先通过联合训练得到。

[0009] 进一步地,所述控制打开对应的开关的步骤,包括:

[0010] 对所述用户的语音信息进行声源定位,获取所述用户距离开关的距离;

[0011] 判断所述距离是否小于第一预设距离;

[0012] 若小于所述第一预设距离,则控制立即打开对应的开关;

[0013] 若不小于所述第一预设距离,则判断所述距离是否处于第一预设距离与第二预设距离之间;

[0014] 若处于,则获取所述第一预设距离与第二预设距离的区间所对应的延迟时间;基于所述延迟时间,控制所述开关在延迟时间之后打开;

[0015] 若不处于,则控制所述开关在预设时间之后打开;其中,所述预设时间大于所述延迟时间。

[0016] 进一步地,所述控制打开对应的开关的步骤,包括:

[0017] 基于所述用户的语音信息,获取所述用户的朝向区域;

[0018] 根据各个区域中的开关列表,获取所述朝向区域内所有的预选开关;

[0019] 获取各个所述预选开关的唤醒词,从所述预选开关中确定出唤醒词为对应识别到的唤醒词的开关,作为第一开关;

[0020] 判断所述第一开关是否为关闭状态,若是,则打开所述第一开关;

[0021] 获取第一开关的标识信息,对所述标识信息进行哈希计算得到对应的标识哈希值;在所述开关接收到关闭指令关闭时,将所述标识哈希值与所述用户的语音信息进行关联存储。

[0022] 进一步地,所述采集用户的语音信息,对所述用户的语音信息进行特征识别,判断所述用户是否是特定用户的步骤之前,包括:

[0023] 获取训练语音数据,对所述训练语音数据进行特征提取,得到第一语音特征向量;其中,所述训练语音数据包括对应的标签;

[0024] 将所述第一语音特征向量输入至预设的编码模型中进行特征提取,得到第一编码向量;

[0025] 将所述第一编码向量输入至预设的解码模型中进行解码处理,得到对应的第二语音特征向量;

[0026] 将所述第二语音特征向量输入至预设的编码模型中进行特征提取,得到第二编码向量;

[0027] 将所述第一编码向量、第二编码向量及对应的标签输入至分类层中进行训练,并迭代调整所述预设的编码模型、解码模型的模型参数,直至所述分类层的损失函数收敛之后,完成模型的训练;其中,将所述编码模型与所述分类层组合作为唤醒模型,用于对降噪处理之后的语音信息进行唤醒词识别;将所述编码模型与所述解码模型作为降噪模型,用于对用户的语音信息进行降噪处理。

[0028] 进一步地,所述对所述用户的语音信息进行降噪处理的步骤之前,包括:

[0029] 采集当前的环境声音;

[0030] 将所述环境声音输入至预设的环境检测模型中进行检测,得到对应的环境检测结果;其中,所述环境检测模型是预先训练完成的深度学习模型;

[0031] 根据所述环境检测结果,在数据库中确定对应的目标模型参数集合;其中,数据库中存储有环境检测结果与模型参数集合的对应关系,所述模型参数集合中包括多个模型参数:平滑机制参数,降噪队列长度,降噪阈值大小;

[0032] 获取初始的降噪模型,提取出所述目标模型参数集合中的各个目标模型参数,并将所述初始的降噪模型中的模型参数更新为对应的目标模型参数,得到更新的降噪模型;其中,所述降噪模型用于对所述用户的语音信息进行降噪处理。

[0033] 进一步地,所述对降噪处理之后的语音信息进行唤醒词识别,若识别到对应的唤醒词,则控制打开对应的开关的步骤之前,包括:

[0034] 获取第一语音训练样本,其中,所述第一语音训练样本为包括唤醒词的语音数据;

[0035] 获取第二语音训练样本,其中,所述第二语音训练样本为不包括唤醒词的语音数据,且所述第二语音训练样本的语音数据中包括多个与唤醒词语音相近的词;

[0036] 将所述第一语音训练样本输入至预设的唤醒模型中,得到所述唤醒模型被唤醒的唤醒率;

[0037] 将所述第二语音训练样本输入至预设的唤醒模型中,得到所述唤醒模型被误唤醒

的误唤醒率；

[0038] 判断所述唤醒率是否大于对应的第一阈值，以及判断所述误唤醒率是否大于对应的第二阈值；

[0039] 若所述唤醒率大于对应的第一阈值，且所述误唤醒率大于第二阈值，则增加所述唤醒模型被唤醒词唤醒的唤醒阈值以及降低平滑队列长度；

[0040] 若所述唤醒率不大于对应的第一阈值，且所述误唤醒率大于第二阈值，则增加所述唤醒模型的平滑队列长度；

[0041] 若所述唤醒率不大于对应的第一阈值，且所述误唤醒率不大于第二阈值，则降低所述唤醒模型被唤醒词唤醒的唤醒阈值。

[0042] 进一步地，所述将所述第一语音训练样本输入至预设的唤醒模型中，得到所述唤醒模型被唤醒的唤醒率的步骤，包括：

[0043] 对所述第一语音训练样本进行语音识别，识别出所述第一语音训练样本中的各个唤醒词在第一语音训练样本中的时间区间，记为唤醒词时间区间；

[0044] 在预设的文档中记录所述第一语音训练样本各个唤醒词及其对应的唤醒词时间区间；

[0045] 将所述第一语音训练样本输入至预设的唤醒模型中进行唤醒识别，并在所述唤醒模型开始识别时，开始计时；

[0046] 在所述唤醒模型识别到各个唤醒词时，记录对应的唤醒时间区间；

[0047] 将记录唤醒时间区间与所述预设的文档中记录的各个唤醒词及其对应的唤醒词时间区间进行对比；当所述唤醒时间区间与所述唤醒词时间区间重叠时，作为一次正确唤醒次数；当所述唤醒时间区间与所述唤醒词时间区间不重叠时，作为一次错误唤醒次数；

[0048] 根据所述正确唤醒次数、错误唤醒次数计算所述唤醒模型被唤醒的唤醒率；其中，所述唤醒率=正确唤醒次数/(正确唤醒次数+错误唤醒次数)。

[0049] 本发明还提供了一种基于语音识别的开关控制装置，包括：

[0050] 采集单元，用于采集用户的语音信息，对所述用户的语音信息进行特征识别，判断所述用户是否是特定用户；

[0051] 检测单元，用于若是特定用户，则检测所述用户的语音信息的分贝值，并判断所述分贝值是否处于预设区间；

[0052] 降噪单元，用于若所述分贝值处于预设区间，则对所述用户的语音信息进行降噪处理；

[0053] 控制单元，用于对降噪处理之后的语音信息进行唤醒词识别，若识别到对应的唤醒词，则控制打开对应的开关；若未识别到，则保持开关当前的状态；其中，用于降噪处理的降噪模型以及用于唤醒识别的唤醒模型预先通过联合训练得到。

[0054] 进一步地，还包括：

[0055] 获取单元，用于获取训练语音数据，对所述训练语音数据进行特征提取，得到第一语音特征向量；其中，所述训练语音数据包括对应的标签；

[0056] 第一编码单元，用于将所述第一语音特征向量输入至预设的编码模型中进行特征提取，得到第一编码向量；

[0057] 第一解码单元，用于将所述第一编码向量输入至预设的解码模型中进行解码处

理,得到对应的第二语音特征向量;

[0058] 第二编码单元,用于将所述第二语音特征向量输入至预设的编码模型中进行特征提取,得到第二编码向量;

[0059] 训练单元,用于将所述第一编码向量、第二编码向量及对应的标签输入至分类层中进行训练,并迭代调整所述预设的编码模型、解码模型的模型参数,直至所述分类层的损失函数收敛之后,完成模型的训练;其中,将所述编码模型与所述分类层组合作为唤醒模型,用于对降噪处理之后的语音信息进行唤醒词识别;将所述编码模型与所述解码模型作为降噪模型,用于对用户的语音信息进行降噪处理。

[0060] 本发明还提供一种开关设备,包括存储器和处理器,所述存储器中存储有计算机程序,所述处理器执行所述计算机程序时实现上述任一项所述方法的步骤。

[0061] 本发明还提供一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现上述任一项所述的方法的步骤。

[0062] 本发明提供的基于语音识别的开关控制方法、装置、开关设备和存储介质,采集到用户的语音信息之后,首先对所述用户的语音信息进行特征识别,判断所述用户是否是特定用户;若是特定用户,则检测所述用户的语音信息的分贝值,并判断所述分贝值是否处于预设区间;若所述分贝值处于预设区间,则对所述用户的语音信息进行降噪处理;对降噪处理之后的语音信息进行唤醒词识别,若识别到对应的唤醒词,则控制打开对应的开关;若未识别到,则保持开关当前的状态。本发明通过多个检测过程之后再行降噪、唤醒识别,降低了智能家居的开关被误唤醒的概率。

附图说明

[0063] 图1 是本发明一实施例中基于语音识别的开关控制方法步骤示意图;

[0064] 图2 是本发明一实施例中基于语音识别的开关控制装置结构框图;

[0065] 图3 是本发明一实施例的开关设备的结构示意图。

[0066] 本发明目的的实现、功能特点及优点将结合实施例,参照附图做进一步说明。

具体实施方式

[0067] 为了使本发明的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本发明进行进一步详细说明。应当理解,此处描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0068] 参照图1,本发明提供了一种基于语音识别的开关控制方法,包括以下步骤:

[0069] 步骤S1,采集用户的语音信息,对所述用户的语音信息进行特征识别,判断所述用户是否是特定用户;

[0070] 步骤S2,若是特定用户,则检测所述用户的语音信息的分贝值,并判断所述分贝值是否处于预设区间;

[0071] 步骤S3,若所述分贝值处于预设区间,则对所述用户的语音信息进行降噪处理;

[0072] 步骤S4,对降噪处理之后的语音信息进行唤醒词识别,若识别到对应的唤醒词,则控制打开对应的开关;若未识别到,则保持开关当前的状态。

[0073] 在本实施例中,上述方案应用于对智能家居中的开关设备进行控制,上述开关设

备采用语音控制实现开、关,如上述步骤S1所述的,采集用户的语音信息,室内可以有多个用户,为了避免开关设备被所有用户随意打开,本实施例中需要对用户的语音信息进行特征识别,以判断出上述用户是否为特定用户,上述特定用户指的是可以对上述开关设备进行控制的用户群体,可以是一个,也可以是多个。

[0074] 如上述步骤S2所述的,若是特定用户,则需要所述用户的语音信息的分贝值,并判断所述分贝值是否处于预设区间;可以理解的是,上述预设区间的设置限定了分贝值只能处于一个区间范围,不能过大,也不能过小;若用户的分贝值较小、较大时均可能不是正常情况下发出的语音信息,若进行开关控制,则容易造成误判。如上述步骤S3所述的,若所述分贝值处于预设区间,为了便于后续的语音识别,避免噪声造成干扰,需要对所述用户的语音信息进行降噪处理,本实施例中可以预先训练一个降噪模型用于对用户的语音信息进行降噪处理。如上述步骤S4所述的,预先训练得到一个唤醒模型,上述唤醒模型对降噪处理之后的语音信息进行唤醒词识别,若识别到对应的唤醒词,则控制开关设备打开对应的开关;若未识别到,则保持开关当前的状态。在本实施例中,开关设备采集到用户的语音信息之后,需要依次进行特征识别、分贝值检测、降噪处理之后,才会进行唤醒识别,降低了开关设备被误唤醒的概率。

[0075] 在一实施例中,所述控制打开对应的开关的步骤,包括:

[0076] 对所述用户的语音信息进行声源定位,获取所述用户距离开关的距离;

[0077] 判断所述距离是否小于第一预设距离;

[0078] 若小于所述第一预设距离,则控制立即打开对应的开关;

[0079] 若不小于所述第一预设距离,则判断所述距离是否处于第一预设距离与第二预设距离之间;

[0080] 若处于,则获取所述第一预设距离与第二预设距离的区间所对应的延迟时间;基于所述延迟时间,控制所述开关在延迟时间之后打开;数据库中存储有第一预设距离与第二预设距离构成的距离区间与延迟时间的对应关系。

[0081] 若不处于,则控制所述开关在预设时间之后打开;其中,所述预设时间大于所述延迟时间。

[0082] 在本实施例中,在控制打开对应的开关时,还需要对用户语音信息进行声源定位,获取所述用户距离开关的距离;若用户距离较远,则可以延迟一段时间打开开关;若用户距离很近,则可以立即打开开关。因此,可以设置一个第一预设距离以及第二预设距离;当用户距离小于第一预设距离时,表明用户很近,此时可以立即打开开关;当用户距离大于第二预设距离时,表明用户很远,此时可以控制开关在预设时间之后打开,上述预设时间是一个预先设定好的固定时间。

[0083] 在一实施例中,所述控制打开对应的开关的步骤,包括:

[0084] 基于所述用户的语音信息,获取所述用户的朝向区域;

[0085] 根据各个区域中的开关列表,获取所述朝向区域内所有的预选开关;

[0086] 获取各个所述预选开关的唤醒词,从所述预选开关中确定出唤醒词为对应识别到的唤醒词的开关,作为第一开关;

[0087] 判断所述第一开关是否为关闭状态,若是,则打开所述第一开关;

[0088] 获取第一开关的标识信息,对所述标识信息进行哈希计算得到对应的标识哈希

值;在所述开关接收到关闭指令关闭时,将所述标识哈希值与所述用户的语音信息进行关联存储。

[0089] 在本实施例中,由于智能家居的设备很多,而用户为了方便,会将多个智能家居的唤醒词设置为类似或者一致,此时,当用户说出唤醒词时,会造成多个智能家居的开关设备打开,因此,在本实施例中,基于所述用户的语音信息,获取所述用户的朝向区域;进而根据各个区域中的开关列表,获取所述朝向区域内所有的预选开关;获取各个所述预选开关的唤醒词,从所述预选开关中确定出唤醒词为对应识别到的唤醒词的开关,作为第一开关;判断所述第一开关是否为关闭状态,若是,则打开所述第一开关;获取第一开关的标识信息,对所述标识信息进行哈希计算得到对应的标识哈希值;在所述开关接收到关闭指令关闭时,将所述标识哈希值与所述用户的语音信息进行关联存储;基于上述标识哈希值进行关联存储,便于标识上述用户此次对开关设备的操作过程。

[0090] 在一实施例中,其中用于降噪处理的降噪模型以及用于唤醒识别的唤醒模型预先通过联合训练得到。所述采集用户的语音信息,对所述用户的语音信息进行特征识别,判断所述用户是否是特定用户的步骤S1之前,包括:

[0091] 获取训练语音数据,对所述训练语音数据进行特征提取,得到第一语音特征向量;其中,所述训练语音数据包括对应的标签;

[0092] 将所述第一语音特征向量输入至预设的编码模型中进行特征提取,得到第一编码向量;

[0093] 将所述第一编码向量输入至预设的解码模型中进行解码处理,得到对应的第二语音特征向量;

[0094] 将所述第二语音特征向量输入至预设的编码模型中进行特征提取,得到第二编码向量;

[0095] 将所述第一编码向量、第二编码向量及对应的标签输入至分类层中进行训练,并迭代调整所述预设的编码模型、解码模型的模型参数,直至所述分类层的损失函数收敛之后,完成模型的训练;其中,将所述编码模型与所述分类层组合作为唤醒模型,用于对降噪处理之后的语音信息进行唤醒词识别;将所述编码模型与所述解码模型作为降噪模型,用于对用户的语音信息进行降噪处理。

[0096] 在本实施例中,提出上述联合训练降噪模型以及唤醒模型的方案,在实际训练过程,经常遇到训练数据不够的情况,本实施例中,对训练语音数据进行特征提取,得到第一语音特征向量,所述第一语音特征向量输入至预设的编码模型中进行特征提取,得到第一编码向量;由于训练语音数据中通常包括有噪声,第一编码向量中也会携带有噪声特征;为了实现降噪,可以将所述第一编码向量输入至预设的解码模型中进行解码处理,得到对应的第二语音特征向量;第二语音特征向量是降噪之后的特征,进而将所述第二语音特征向量输入至预设的编码模型中进行特征提取,得到第二编码向量,第二编码向量相对于第一编码向量而言,更加纯净,没有噪声干扰;进一步地,将所述第一编码向量、第二编码向量及对应的标签输入至分类层中进行训练;不仅增加了训练数据量,同时还可以同时得到降噪模型以及唤醒模型,简化了训练过程,降低训练处理数据量。

[0097] 在一实施例中,所述对所述用户的语音信息进行降噪处理的步骤之前,包括:

[0098] 采集当前的环境声音;

[0099] 将所述环境声音输入至预设的环境检测模型中进行检测,得到对应的环境检测结果;其中,所述环境检测模型是预先训练完成的深度学习模型;

[0100] 根据所述环境检测结果,在数据库中确定对应的目标模型参数集合;其中,数据库中存储有环境检测结果与模型参数集合的对应关系,所述模型参数集合中包括多个模型参数:平滑机制参数,降噪队列长度,降噪阈值大小;

[0101] 获取初始的降噪模型,提取出所述目标模型参数集合中的各个目标模型参数,并将所述初始的降噪模型中的模型参数更新为对应的目标模型参数,得到更新的降噪模型;其中,所述降噪模型用于对所述用户的语音信息进行降噪处理。

[0102] 在本实施例中,针对不同的环境,需要采用不同的降噪模型,以增强降噪效果,具体地,采集当前的环境声音,通过预设的环境检测模型中进行检测,得到对应的环境检测结果;进而,在数据库中确定环境检测结果对应的目标模型参数集合,获取初始的降噪模型,根据目标模型参数集合更新降噪模型的模型参数;在本实施例中,预先得到每个环境所对应的模型参数集合,即最优的模型参数集,根据模型参数集合去调整降噪模型,使得降噪模型的降噪效率得到最大优化,提升降噪效果。

[0103] 在一实施例中,所述对降噪处理之后的语音信息进行唤醒词识别,若识别到对应的唤醒词,则控制打开对应的开关的步骤之前,包括:

[0104] 获取第一语音训练样本,其中,所述第一语音训练样本为包括唤醒词的语音数据;

[0105] 获取第二语音训练样本,其中,所述第二语音训练样本为不包括唤醒词的语音数据,且所述第二语音训练样本的语音数据中包括多个与唤醒词语音相近的词;

[0106] 将所述第一语音训练样本输入至预设的唤醒模型中,得到所述唤醒模型被唤醒的唤醒率;

[0107] 将所述第二语音训练样本输入至预设的唤醒模型中,得到所述唤醒模型被误唤醒的误唤醒率;

[0108] 判断所述唤醒率是否大于对应的第一阈值,以及判断所述误唤醒率是否大于对应的第二阈值;

[0109] 若所述唤醒率大于对应的第一阈值,且所述误唤醒率大于第二阈值,则增加所述唤醒模型被唤醒词唤醒的唤醒阈值以及降低平滑队列长度;以使得误唤醒率可以逐渐降低,最终满足要求;

[0110] 若所述唤醒率不大于对应的第一阈值,且所述误唤醒率大于第二阈值,则增加所述唤醒模型的平滑队列长度;以使得误唤醒率可以逐渐降低,且唤醒率可以逐渐增大,最终满足要求;

[0111] 若所述唤醒率不大于对应的第一阈值,且所述误唤醒率不大于第二阈值,则降低所述唤醒模型被唤醒词唤醒的唤醒阈值;以使得唤醒率可以逐渐增大,最终满足要求。

[0112] 在一实施例中,所述将所述第一语音训练样本输入至预设的唤醒模型中,得到所述唤醒模型被唤醒的唤醒率的步骤,包括:

[0113] 对所述第一语音训练样本进行语音识别,识别出所述第一语音训练样本中的各个唤醒词在第一语音训练样本中的时间区间,记为唤醒词时间区间;

[0114] 在预设的文档中记录所述第一语音训练样本各个唤醒词及其对应的唤醒词时间区间;

[0115] 将所述第一语音训练样本输入至预设的唤醒模型中进行唤醒识别,并在所述唤醒模型开始识别时,开始计时;

[0116] 在所述唤醒模型识别到各个唤醒词时,记录对应的唤醒时间区间;

[0117] 将记录唤醒时间区间与所述预设的文档中记录的各个唤醒词及其对应的唤醒词时间区间进行对比;当所述唤醒时间区间与所述唤醒词时间区间重叠时,作为一次正确唤醒次数;当所述唤醒时间区间与所述唤醒词时间区间不重叠时,作为一次错误唤醒次数;在本实施例中,采用文档方式的对比,不仅需要对比唤醒词,还需要对比唤醒词出现的时间,当时间重叠时才判定为正确唤醒。

[0118] 根据所述正确唤醒次数、错误唤醒次数计算所述唤醒模型被唤醒的唤醒率;其中,所述唤醒率=正确唤醒次数/(正确唤醒次数+错误唤醒次数)。

[0119] 参照图2,本发明一实施例中还提供了一种基于语音识别的开关控制装置,包括:

[0120] 采集单元,用于采集用户的语音信息,对所述用户的语音信息进行特征识别,判断所述用户是否是特定用户;

[0121] 检测单元,用于若是特定用户,则检测所述用户的语音信息的分贝值,并判断所述分贝值是否处于预设区间;

[0122] 降噪单元,用于若所述分贝值处于预设区间,则对所述用户的语音信息进行降噪处理;

[0123] 控制单元,用于对降噪处理之后的语音信息进行唤醒词识别,若识别到对应的唤醒词,则控制打开对应的开关;若未识别到,则保持开关当前的状态;其中,用于降噪处理的降噪模型以及用于唤醒识别的唤醒模型预先通过联合训练得到。

[0124] 在另一实施例中,还包括:

[0125] 获取单元,用于获取训练语音数据,对所述训练语音数据进行特征提取,得到第一语音特征向量;其中,所述训练语音数据包括对应的标签;

[0126] 第一编码单元,用于将所述第一语音特征向量输入至预设的编码模型中进行特征提取,得到第一编码向量;

[0127] 第一解码单元,用于将所述第一编码向量输入至预设的解码模型中进行解码处理,得到对应的第二语音特征向量;

[0128] 第二编码单元,用于将所述第二语音特征向量输入至预设的编码模型中进行特征提取,得到第二编码向量;

[0129] 训练单元,用于将所述第一编码向量、第二编码向量及对应的标签输入至分类层中进行训练,并迭代调整所述预设的编码模型、解码模型的模型参数,直至所述分类层的损失函数收敛之后,完成模型的训练;其中,将所述编码模型与所述分类层组合作为唤醒模型,用于对降噪处理之后的语音信息进行唤醒词识别;将所述编码模型与所述解码模型作为降噪模型,用于对用户的语音信息进行降噪处理。

[0130] 在本实施例中,上述装置实施例中的各个单元的具体实现,请参照上述方法实施例中所述,在此不再进行赘述。

[0131] 参照图3,本发明实施例中还提供一种开关设备,该开关设备内部结构可以如图3所示。该开关设备包括通过系统总线连接的处理器、存储器、网络接口和数据库。其中,该计算机设计的处理器用于提供计算和控制能力。该开关设备的存储器包括非易失性存储介

质、内存储器。该非易失性存储介质存储有操作系统、计算机程序和数据库。该内存储器为非易失性存储介质中的操作系统和计算机程序的运行提供环境。该开关设备的数据库用于存储本实施例中对应的数据。该开关设备的网络接口用于与外部的终端通过网络连接通信。该计算机程序被处理器执行时以实现一种基于语音识别的开关控制方法。

[0132] 本领域技术人员可以理解,图3中示出的结构,仅仅是与本发明方案相关的部分结构的框图,并不构成对本发明方案所应用于其上的开关设备的限定。

[0133] 本发明一实施例还提供一种计算机可读存储介质,其上存储有计算机程序,计算机程序被处理器执行时实现一种基于语音识别的开关控制方法。可以理解的是,本实施例中的计算机可读存储介质可以是易失性可读存储介质,也可以为非易失性可读存储介质。

[0134] 综上所述,为本发明实施例中提供的基于语音识别的开关控制方法、装置、开关设备和存储介质,采集到用户的语音信息之后,首先对所述用户的语音信息进行特征识别,判断所述用户是否是特定用户;若是特定用户,则检测所述用户的语音信息的分贝值,并判断所述分贝值是否处于预设区间;若所述分贝值处于预设区间,则对所述用户的语音信息进行降噪处理;对降噪处理之后的语音信息进行唤醒词识别,若识别到对应的唤醒词,则控制打开对应的开关;若未识别到,则保持开关当前的状态。本发明通过多个检测过程之后再行降噪、唤醒识别,降低了智能家居的开关被误唤醒的概率。

[0135] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,所述的计算机程序可存储于一非易失性计算机可读存储介质中,该计算机程序在执行时,可包括如上述各方法的实施例的流程。其中,本发明所提供的和实施例中所使用的对存储器、存储、数据库或其它介质的任何引用,均可包括非易失性和/或易失性存储器。非易失性存储器可以包括只读存储器(ROM)、可编程ROM(PROM)、电可编程ROM(EPROM)、电可擦除可编程ROM(EEPROM)或闪存。易失性存储器可包括随机存取存储器(RAM)或者外部高速缓冲存储器。作为说明而非局限,RAM通过多种形式可得,诸如静态RAM(SRAM)、动态RAM(DRAM)、同步DRAM(SDRAM)、双速据率SDRAM(SSRSDRAM)、增强型SDRAM(ESDRAM)、同步链路(Synchlink)DRAM(SLDRAM)、存储器总线(Rambus)直接RAM(RDRAM)、直接存储器总线动态RAM(DRDRAM)、以及存储器总线动态RAM(RDRAM)等。

[0136] 需要说明的是,在本文中,术语“包括”、“包含”或者其任何其它变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、装置、物品或者方法不仅包括那些要素,而且还包括没有明确列出的其它要素,或者是还包括为这种过程、装置、物品或者方法所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括该要素的过程、装置、物品或者方法中还存在另外的相同要素。

[0137] 以上所述仅为本发明的优选实施例,并非因此限制本发明的专利范围,凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其它相关的技术领域,均同理包括在本发明的专利保护范围内。

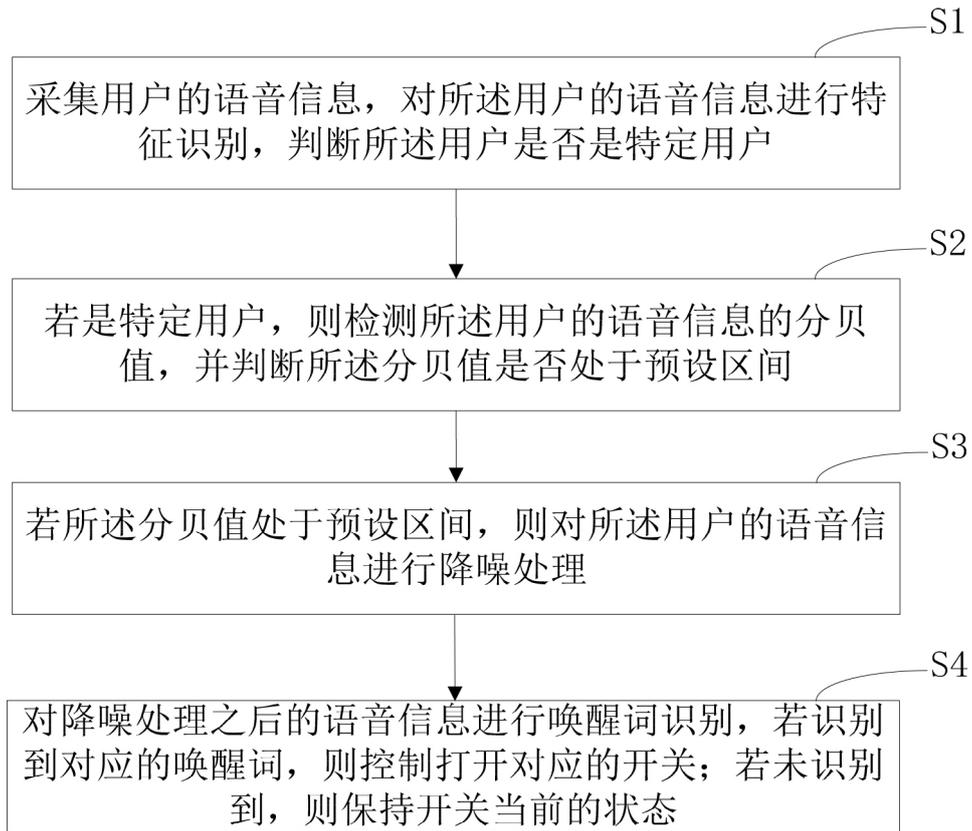


图 1

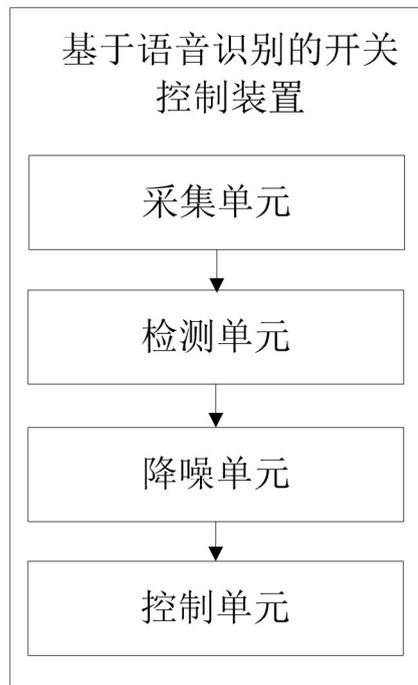


图 2

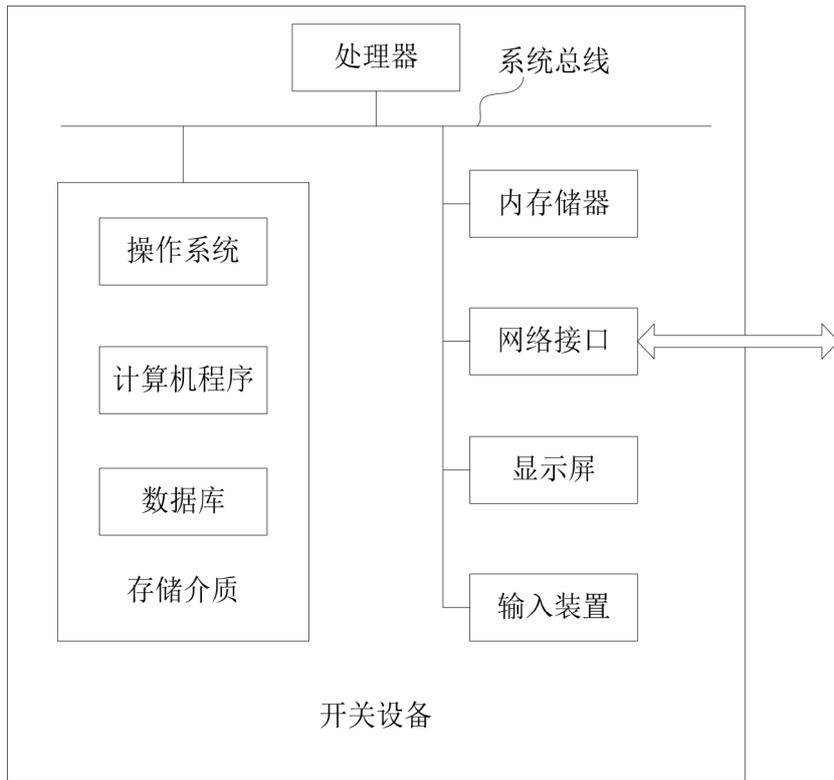


图 3