



(12) 发明专利申请

(10) 申请公布号 CN 115331672 A

(43) 申请公布日 2022. 11. 11

(21) 申请号 202211255037.X

(22) 申请日 2022.10.13

(71) 申请人 深圳市声扬科技有限公司  
地址 518000 广东省深圳市前海深港合作区前湾一路1号A栋201室(入驻深圳市前海商务秘书有限公司)

(72) 发明人 陈东鹏 宋政斌

(74) 专利代理机构 深圳市智圈知识产权代理事务所(普通合伙) 44351  
专利代理师 李璇

(51) Int. Cl.  
G10L 15/22 (2006.01)  
G10L 15/26 (2006.01)  
G10L 17/02 (2013.01)  
G10L 19/26 (2013.01)

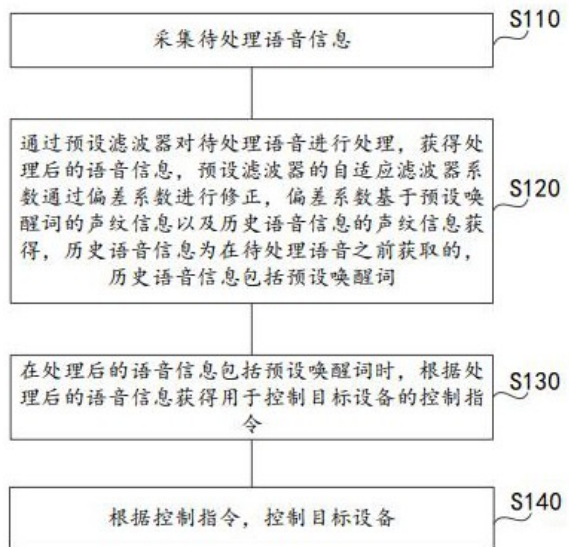
权利要求书2页 说明书12页 附图4页

(54) 发明名称

设备控制方法、装置、电子设备及存储介质

(57) 摘要

本申请公开了一种设备控制方法、装置、电子设备及存储介质,方法包括:采集待处理语音信息,通过预设滤波器对待处理语音信息进行处理,获得处理后的语音信息,预设滤波器的自适应滤波器系数通过偏差系数进行修正,偏差系数基于预设唤醒词的声纹信息以及历史语音信息的声纹信息获得,在处理后的语音信息包括预设唤醒词时,根据处理后的语音信息获得用于控制目标设备的控制指令,根据控制指令,控制目标设备。通过上述方法,通过偏差系数对预设滤波器的自适应滤波器系数进行修正,使得通过该预设滤波器处理后的语音信息更加准确,有效提升了语音处理性能,从而提高了设备控制的准确性。



1. 一种设备控制方法,其特征在于,所述方法包括:

采集待处理语音信息;

通过预设滤波器对所述待处理语音进行处理,获得处理后的语音信息,所述预设滤波器的自适应滤波器系数通过偏差系数进行修正,所述偏差系数基于预设唤醒词的声纹信息以及历史语音信息的声纹信息获得,所述历史语音信息在所述待处理语音之前获取的,所述历史语音信息包括所述预设唤醒词;

在所述处理后的语音信息包括所述预设唤醒词时,根据所述处理后的语音信息获得用于控制目标设备的控制指令;

根据所述控制指令,控制所述目标设备。

2. 根据权利要求1所述的方法,其特征在于,所述通过预设滤波器对所述待处理语音进行处理,获得处理后的语音信息之前,所述方法还包括:

对所述待处理语音信息进行降噪以及回声消除处理;

从处理后的所述待处理语音中提取人声语段,得到第一有效声音信息;

所述通过预设滤波器对所述待处理语音进行处理,获得处理后的语音信息,包括:

通过所述预设滤波器对所述第一有效声音信息进行处理,获得处理后的语音信息。

3. 根据权利要求1所述的方法,其特征在于,所述历史语音信息的声纹信息的获取方法包括:

获取所述历史语音信息;

从所述历史语音信息中提取人声语段,得到第二有效声音信息;

通过多个预置算法分别对所述第二有效声音信息进行处理,得到每个所述预置算法分别对应的算法处理结果;每个所述预置算法通过对应该预置算法的配置系数对所述预设滤波器的自适应滤波器系数配置后获得,每个所述预置算法对应的配置系数基于所述偏差系数获得;

根据所述多个预置算法各自对应的算法处理结果,得到所述历史语音信息的声纹信息。

4. 根据权利要求3所述的方法,其特征在于,所述历史语音信息的声纹信息包括所述多个预置算法各自的算法处理结果对应的声纹信息;所述偏差系数的获得方法包括:

分别确定所述多个预置算法各自的算法处理结果对应的声纹信息与所述预设唤醒词的声纹信息的声纹相似度;

将声纹相似度的最大值确定为目标声纹相似度;

根据所述目标声纹相似度,得到所述偏差系数。

5. 根据权利要求4所述的方法,其特征在于,所述根据所述目标声纹相似度,得到所述偏差系数,包括:

计算所述目标声纹相似度的平方,作为第一计算结果;

将预设常数与所述第一计算结果的差值作为所述偏差系数。

6. 根据权利要求1所述的方法,其特征在于,所述预设滤波器的自适应滤波器系数的修正的方法包括:

获取所述预设滤波器的修正前的自适应滤波器系数、迭代步长以及误差项,所述修正前的自适应滤波器系数是在通过所述预设滤波器对所述历史语音信息进行处理之前,所述

预设滤波器中的自适应滤波器系数；

根据所述迭代步长、所述误差项、所述偏差系数以及所述历史语音信息对应的特征信息,对所述修正前的自适应滤波器系数进行修正,得到修正后的自适应滤波器系数。

7. 根据权利要求6所述的方法,其特征在于,所述根据所述迭代步长、所述误差项、所述偏差系数以及所述历史语音信息对应的特征信息,对所述修正前的自适应滤波器系数进行修正,得到修正后的自适应滤波器系数,包括:

计算所述迭代步长、所述误差项、所述偏差系数以及所述历史语音信息对应的特征信息的积,作为第二计算结果;

计算所述第二计算结果与所述修正前的自适应滤波器系数的和,作为所述修正后的自适应滤波器系数。

8. 一种设备控制装置,其特征在于,所述装置包括:

语音采集模块,用于采集待处理语音信息;

语音处理模块,用于通过预设滤波器对所述待处理语音进行处理,获得处理后的语音信息,所述预设滤波器的自适应滤波器系数通过偏差系数进行修正,所述偏差系数基于预设唤醒词的声纹信息以及历史语音信息的声纹信息获得,所述历史语音信息在所述待处理语音之前获取的,所述历史语音信息包括所述预设唤醒词;

指令识别模块,用于在所述处理后的语音信息包括所述预设唤醒词时,根据所述处理后的语音信息获得用于控制目标设备的控制指令;

设备控制模块,用于根据所述控制指令,控制所述目标设备。

9. 一种电子设备,其特征在于,包括:

一个或多个处理器;

存储器;

一个或多个程序,其中所述一个或多个程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行,所述一个或多个程序配置用于执行如权利要求1-7中任意一项所述的方法。

10. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储有程序代码,所述程序代码可被处理器调用执行如权利要求1-7中任意一项所述的方法。

## 设备控制方法、装置、电子设备及存储介质

### 技术领域

[0001] 本申请涉及人工智能技术领域,更具体地,设计一种设备控制方法、装置、电子设备及计算机可读存储介质。

### 背景技术

[0002] 随着人工智能技术的发展,语音识别技术作为人工智能技术中的一个重要领域,其应用场景也越来越多,例如,用户可通过语音控制电子设备执行某些动作,如用户可以控制电子设备播放音乐或者查询天气等。

[0003] 目前,用户可以向电子设备发送包括唤醒词和控制指令的语音信息,电子设备根据唤醒词,确定语音信息的发送者是已认证的用户之后,电子设备可以根据语音信息的控制指令,控制对应的设备输出动作。

[0004] 但是,采用现有的方法对设备进行控制的准确性较差。

### 发明内容

[0005] 鉴于此,本申请实施例提出了一种设备控制方法、装置、电子设备及计算机可读存储介质,以改善上述问题。

[0006] 第一方面,本申请实施例提供了一种设备控制方法,该方法包括:采集待处理语音信息;通过预设滤波器对待处理语音进行处理,获得处理后的语音信息,预设滤波器的自适应滤波器系数通过偏差系数进行修正,偏差系数基于预设唤醒词的声纹信息以及历史语音信息的声纹信息获得,历史语音信息在待处理语音之前获取的,历史语音信息包括预设唤醒词;在处理后的语音信息包括预设唤醒词时,根据处理后的语音信息获得用于控制目标设备的控制指令;根据控制指令,控制目标设备。

[0007] 第二方面,本申请实施例提供了一种设备控制装置,装置包括:语音采集模块,用于采集待处理语音信息;语音处理模块,通过预设滤波器对待处理语音进行处理,获得处理后的语音信息,预设滤波器的自适应滤波器系数通过偏差系数进行修正,偏差系数基于预设唤醒词的声纹信息以及历史语音信息的声纹信息获得,历史语音信息在待处理语音之前获取的,历史语音信息包括预设唤醒词;指令识别模块,用于在处理后的语音信息包括预设唤醒词时,根据处理后的语音信息获得用于控制目标设备的控制指令;设备控制模块,用于根据控制指令,控制目标设备。

[0008] 第三方面,本申请实施例提供了一种电子设备,包括:一个或多个处理器;存储器;一个或多个程序,其中一个或多个程序被存储在存储器中并被配置为由一个或多个处理器执行,一个或多个程序配置用于执行上述的方法。

[0009] 第四方面,本申请实施例提供了一种计算机可读存储介质,计算机可读存储介质存储有程序代码,程序代码可被处理器调用执行上述的方法。

[0010] 本申请提供的一种设备控制方法、装置、电子设备及计算机可读存储介质,在本申请中,预设滤波器的自适应滤波器系数通过偏差系数进行修正,包括修正后的自适应滤波

器系数的预设滤波器的语音处理能力较强,使得通过预设滤波器对待处理语音信息处理后获得的处理后的语音信息包括较少的噪声、且更加准确,从而使得根据待处理语音信息获得的控制指令更加准确,提高了设备控制的准确性。

### 附图说明

[0011] 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请的一些实施例,对于本领域技术人员来讲,在不付出创造性劳动的前提下,还可以

图1示出了本申请实施例提出的一种设备控制方法的实施环境示意图;

图2示出了本申请一个实施例提出的一种设备控制方法的流程图;

图3示出了本申请实施例中历史声纹信息的一种获取方法的流程图;

图4示出了本申请实施例中偏差系数的一种获取方法的流程图;

图5示出了本申请实施例中自适应滤波器系数的一种修正方法的流程图;

图6示出了本申请一个实施例提出的一种设备控制装置的结构框图;

图7示出了本申请提出的一种电子设备的结构框图。

### 具体实施方式

[0012] 现在将参考附图更全面地描述示例实施方式。然而,示例实施方式能够以多种形式实施,且不应被理解为限于在此阐述的范例;相反,提供这些实施方式使得本申请将更加全面和完整,并将示例实施方式的构思全面地传达给本领域的技术人员。

[0013] 此外,所描述的特征、结构或特性可以以任何合适的方式结合在一个或更多实施例中。在下面的描述中,提供许多具体细节从而给出对本申请的实施例的充分理解。然而,本领域技术人员将意识到,可以实践本申请的技术方案而没有特定细节中的一个或更多,或者可以采用其它的方法、组元、装置、步骤等。在其它情况下,不详细示出或描述公知方法、装置、实现或者操作以避免模糊本申请的各方面。

[0014] 附图中所示的方框图仅仅是功能实体,不一定必须与物理上独立的实体相对应。即,可以采用软件形式来实现这些功能实体,或在一个或多个硬件模块或集成电路中实现这些功能实体,或在不同网络和/或处理器装置和/或微控制器装置中实现这些功能实体。

[0015] 附图中所示的流程图仅是示例性说明,不是必须包括所有的内容和操作/步骤,也不是必须按所描述的顺序执行。例如,有的操作/步骤还可以分解,而有的操作/步骤可以合并或部分合并,因此实际执行的顺序有可能根据实际情况改变。

[0016] 需要说明的是:在本文中提及的“至少一种”是指一种、两种或两种以上。

[0017] 以下对本申请实施例的技术方案的实现细节进行详细阐述:

请参阅图1,图1示出了本申请提出的一种设备控制方法的实施环境示意图,在该实施环境中,包括车辆100以及设置在该车辆100上的设备控制装置110,设备控制装置110包括车载采集装置111(包括至少一个车载麦克风)、处理器112和存储器113。车载采集装置111可以放置在车辆100的内部空间,车载采集装置111的放置方式包括内嵌入车辆内饰、悬浮于车辆内饰、与车辆有线连接和与车辆无线连线等,车载采集装置111可以用于采集者待处理语音信息。

[0018] 在该场景中,设备控制装置110可以作为本申请实施例中的设备控制方法的执行主体,所需要控制的设备可以是指车辆100中的各个功能模块,例如,车辆100中的灯光可以作为一个设备,通过用户(可以是指司机)的语音信息,对车辆的灯光进行控制。

[0019] 应当理解,上述描述的应用场景和对上述应用场景中各设备的描述尽是一种示例,本实施例中对比不作任何限定。例如,实施环境还可以是智能家居系统,智能家居系统包括控制中心和多个智能家居设备,控制中心作为执行主体,用于执行本申请的设备控制方法,智能家居设备作为被控设备,在控制中心的控制下,输出对应的动作。

[0020] 为了方便表述,下述各个实施例中,以电子设备为执行本申请实施例中的设备控制方法的执行主体进行解释。

[0021] 请参阅图2,图2示出了本申请一个实施例提出的一种设备控制方法的流程图,方法可以用于电子设备,例如可以是图1中的设备控制装置110,方法包括:

S110:采集待处理语音信息。

[0022] 在本实施例中,待处理语音信息是指电子设备采集用户所发出的语音,待处理语音信息中可能包含电子设备所处实际环境的杂音,如其他人声干扰杂音、空气流动产生的噪音等,为未经加工的原声音。其中,待处理语音信息的声源可以为移动声源,也可以是固定声源。

[0023] 作为一种实施方式,电子设备可以包括音频采集装置,例如:智能手机、耳机、车载终端、智能音箱等,实时采集待处理语音,也可以通过外部设备传输的预先采集的待处理语音,并将其发送至电子设备。

[0024] S120:通过预设滤波器对待处理语音进行处理,获得处理后的语音信息,预设滤波器的自适应滤波器系数通过偏差系数进行修正,偏差系数基于预设唤醒词的声纹信息以及历史语音信息的声纹信息获得,历史语音信息在待处理语音之前获取的,历史语音信息包括预设唤醒词。

[0025] 在本实施例中,预设滤波器是指通过训练样本训练获得的自适应滤波器,训练样本可以是自适应滤波器应用的场景中实时采集的音频信息。滤波器的训练过程实质是对滤波器的自适应滤波器系数进行调整的过程,当训练结束时,获得准确性较高的滤波器的自适应滤波器系数。

[0026] 预设滤波器对待处理语音信息进行处理可以是指,通过预设滤波器中的自适应滤波器系数,对待处理语音信息进行滤波处理,得到滤波后的结果作为处理后的语音信息,经过预设滤波器处理后的语音信息包括较少的噪声和杂音。

[0027] 声纹信息(声纹)是指通过声谱仪显示的,携带言语信息的语音声波图谱的通称,是各种声学特征图谱的集合。声纹信息是一种生物特征,跟发音人所说的具体语句无关,即不管发音人说哪句话,所得到的声纹是相同的,声纹不仅具有特定性,而且有相对稳定性的特点。声纹信息可以包括但不限于MFCC(Mel Frequency Cepstral Coefficients,梅尔倒谱系数)特征、LPCC(Linear Prediction Cepstrum Coefficient,线性预测倒谱系数)特征。其中,MFCC特征,利用了人耳听觉频率非线性特性,将频谱转化为基于Mel频率的非线性频谱,然后转换到倒谱域上,充分模拟了人的听觉特性,而且没有任何前提假设,MFCC特征具有识别性能和抗噪能力;LPCC特征,是线性预测系数在倒谱域上的表示方式,是基于语音信号为自回归信号的假设,利用线性预测分析获得倒谱参数,实验中LPC阶数为线性预测倒谱

参数,体现了每个人特定的声道特性。

[0028] 预设唤醒词是指用于唤醒电子设备的设备控制功能的词组,预设唤醒词可以是用户自行设定的词组也可以是电子设备规定的一个或几个词组,预设唤醒词可以是任何语种和形式的字符,例如预设唤醒词可以是小红、小杜等。用户可以通过电子设备输入包括预设唤醒词的预设语音信息,电子设备对预设语音信息进行处理,得到预设唤醒词以及预设唤醒词的声纹信息,电子设备存储预设唤醒词以及预设唤醒词的声纹信息。

[0029] 唤醒电子设备的设备控制功能之后,电子设备可以接收并处理包括控制指令的语音信息,唤醒电子设备的设备控制功能之前,电子设备不接收,也不会处理包括控制指令的语音信息。

[0030] 历史语音信息可以是待处理语音信息之前的语音信息,例如,历史语音信息可以是前一次采集到的,包括预设唤醒词的待处理语音信息。得到历史语音信息之后,电子设备对历史语音信息进行处理,得到历史语音信息的声纹信息。

[0031] 偏差系数可以是用于对自适应滤波器系数进行修正的参数,通过预设唤醒词的声纹信息以及历史语音信息的声纹信息获得。

[0032] 在实际应用中,由于声音信号和环境噪声的不确定性,常常无法得到语音信息中的关键内容,从而造成实际复杂声音环境的偏差。造成实际复杂声音环境的偏差的原因可以是音频采集装置的麦克风阵列使用固定的麦克风阵列波束来划分音区,这种音区的划分是基于声源和麦克风阵列中不同麦克风的距离差或时间差的范围来确定,比如实际的距离差或时间差在某范围内的声源就在目标音区,否则就在干扰音区,但这样的计算只有在理想条件下才是准确的,在实际复杂场景中(如障碍物、声波的漫射反射现象等)会导致和理论值的偏差,并且各偏差是相互独立且不同的。

[0033] 作为一种实施方式,电子设备通过麦克风阵列采集到的待处理语音信息转换为电子设备可以处理的音频数据,利用经过训练后的预设滤波器对音频数据进行处理。由于预设滤波器会不断迭代更新自适应滤波器系数,使得对音频数据的处理越来越接近真实语音,修正实际复杂声音环境中的偏差。例如,预设滤波器对应有目标方向,使用某个方向上的滤波器相关参数来训练该目标方向对应的预设滤波器,使训练得到的预设滤波器可以增强该目标方向上的语音信号,抑制其他方向上的干扰噪音。

[0034] 进一步的,S120之前,方法可以包括:对待处理语音信息进行降噪以及回声消除处理;从处理后的待处理语音中提取人声语段,得到第一有效声音信息;相应的,S120包括:通过预设滤波器对第一有效声音信息进行处理,获得处理后的语音信息。

[0035] 其中,麦克风阵列采集到的待处理语音信息中可能包含实际环境的杂音并且用户声音可能因反射多次反馈至麦克风中,如不处理会影响后续声纹提取的效果,因此在对待处理语音通过预设滤波器处理前,先对待处理语音进行降噪以及回声消除处理。

[0036] 示例性的,对于噪音的处理可在电子设备中设置降噪处理的第三方应用程序,也可安装消声器等物理降噪的方式;对于回声的处理可在电子设备中安装回声抑制器等消除回声。应当理解,上述的降噪及回声消除处理的方式仅为示意性的,还可以有更多的处理方式,上述的列举不应当作为本方案对待处理语音信息进行降噪以及回声消除处理的限制。

[0037] 从处理后的待处理语音中提取人声语段,得到有效声音信息。电子设备可以通过VAD(Voice Activity Detection,语音活动检测)提取人声语段,提取后的人声语段为第一

有效声音信息。其中,VAD可以识别和消除声音中的静音期,只向电子设备传输非静音期的声音,节省了电子设备的带宽资源,提高声音提取的有效性。

[0038] 还需要说明的是,当电子设备的语音识别/对话功能开启后,麦克风阵列会持续采集待处理的语音信息并VAD也会持续对处理后语音提取人声语段。电子设备的语音识别/对话功能的开启方式具体可以是,用户通过电子设备的控件操作选择进入语音识别功能,也可以是用户发出含唤醒词的语音被电子设备成功识别后,电子设备自动开启语音识别/对话功能。应当理解上述的电子设备语音识别/对话功能的开启方式仅为示意性的,在本申请实施例中不作为具体的限定。

[0039] 预设滤波器对经过降噪、回声消除以及提取人声片段处理后的第一有效声音信息进行处理,得到处理后的语音信息。关于预设滤波器对第一有效声音的具体处理过程可以参阅前文的具体描述,此处不做一一赘述。

[0040] S130:在处理后的语音信息包括预设唤醒词时,根据处理后的语音信息获得用于控制目标设备的控制指令。

[0041] 作为一种实施方式,对处理后的语音信息通过唤醒词检测模型进行预设唤醒词检测,若经过预设滤波器处理后的语音信息中包括能够激活电子设备的预设唤醒词时,电子设备从休眠状态切换到唤醒状态,并获取语音信息中控制目标设备的控制指令,将音频形式的控制指令转化为目标设备能够识别的信号。若经过预设滤波器处理后的语音信息中没有检测到预设唤醒词,则唤醒词检测模型继续检测该语音信息是否含有唤醒词或麦克风阵列继续采集语音信息。

[0042] 其中,唤醒词检测模型用于从语音信息中识别出预设唤醒词,该唤醒词识别模型包括但不限于DNN(Deep Neural Networks,深度神经网络)、CNN(Convolutional NeuralNetwork,卷积神经网络)、LSTM(Long Short-Term Memory,长短期记忆网络)等至少一种声学模型。当将待识别的语音信息输入到唤醒词识别模型中,唤醒词识别模型输出该语音信息中包含某一预设唤醒词的概率分数,当该概率分数大于预设分数阈值,则确定该语音信息中包含该预设唤醒词。

[0043] S140:根据控制指令,控制目标设备。

[0044] 作为一种实施方式,根据控制指令,控制目标设备,控制方式可以是该电子设备控制自身进行后续操作,例如用户对车载终端发出导航目的地的语音请求,车载终端则打开导航应用程序,并导航到目的地;也可以是该电子设备控制其他与该电子设备通过网络或蓝牙相连的设备,例如智能家居总控系统获取到用户的控制指令后,发送控制请求至空调、扫地机器人等智能家居设备,智能家居设备执行用户指令的操作。

[0045] 本实施例提供的一种设备控制方法,在本申请中,预设滤波器的自适应滤波器系数通过偏差系数进行修正,包括修正后的自适应滤波器系数的预设滤波器的语音处理能力较强,使得通过预设滤波器对待处理语音信息处理后获得的处理后的语音信息包括较少的噪声、且更加准确,从而使得根据待处理语音信息获得的控制指令更加准确,提高了设备控制的准确性。

[0046] 请参阅图3,图3示出了本申请实施例中历史声纹信息的一种获取方法的流程图,方法可以用于电子设备,例如可以是图1中的设备控制装置110,方法包括:

S210:获取历史语音信息。



[0047] 关于获取历史语音信息方法的具体过程可以参阅前文步骤S110至S120中对采集待处理语音信息以及对待处理语音信息进行降噪以及回声消除处理的过程,此处不作一一赘述。

[0048] S220:从历史语音信息中提取人声语段,得到第二有效声音信息。

[0049] 关于从历史语音信息中提取人声语段得到第二有效声音的具体过程可以参阅前文对待处理语音信息进行人声语段的提取过程,此处不一一赘述。

[0050] S230:通过多个预置算法分别对第二有效声音信息进行处理,得到每个预置算法分别对应的算法处理结果;每个预置算法通过对应该预置算法的配置系数对预设滤波器的自适应滤波器系数配置后获得,每个预置算法对应的配置系数基于偏差系数获得。

[0051] 在本实施例中,一个预置算法是通过一个配置系数对预设滤波器的自适应滤波器系数配置后获得,各个配置系数可以是基于偏差系数获得的。

[0052] 例如,预设算法包括两个时,两个配置系数可以是偏差系数的绝对值以及偏差系数的绝对值的相反数,又如,预设算法包括四个时,四个配置系数可以是偏差系数的绝对值、偏差系数的绝对值的相反数、偏差系数的绝对值的2倍以及零。

[0053] 通过每个预置算法对第二有效声音信息进行处理,得到每个预置算法各自对应的算法处理结果,例如,根据四个预置算法对第二有效声音信息进行处理,得到四个预置算法各自对应的算法处理结果,此时,算法处理结果包括四个。

[0054] 作为一种实施方式,多个预置算法包括第一算法、第二算法以及第三算法,对应的多个配置系数可以包括第一数值、第二数值以及第三数值,S230可以包括:通过第一算法对第二有效声音信息进行处理,获得第一结果;通过第二算法对第二有效声音信息进行处理,获得第二结果;通过第三算法对第二有效声音信息进行处理,获得第三结果。

[0055] 在本实施例中,第一算法是指通过第一数值对预设滤波器的自适应滤波器系数配置后得到,此处第一数值可以配置为零,相当于第一算法的偏差系数为零,第一算法是直接基于预设滤波器对历史语音进行处理之前的预设滤波器对应的自适应滤波器系数。第一结果是指通过预设滤波器的第一算法处理后得到的语音信息。

[0056] 第二算法是指通过第二数值对预设滤波器的自适应滤波器系数配置后得到,此处第二数值可以配置为偏差系数绝对值取正值。第二结果是指通过预设滤波器的第二算法处理后得到的语音信息。

[0057] 第三算法是指通过第三数值对预设滤波器的自适应滤波器系数配置后得到,此处第三数值可以配置为偏差系数绝对值取负值。第三结果是指通过预设滤波器的第三算法处理后得到的语音信息。

[0058] S240:根据多个预置算法各自对应的算法处理结果,得到历史语音信息的声纹信息。

[0059] 得到多个预置算法各自对应的算法处理结果,对该多个预置算法各自对应的算法处理结果分别计算,得到每个算法处理结果对应的声纹信息,得到的各算法处理结果对应的声纹信息即为历史语音信息的声纹信息。

[0060] 例如,上文所述,通过第一算法对所述第二有效声音信息进行处理,获得第一结果;通过第二算法对第二有效声音信息进行处理,获得第二结果;通过第三算法对第二有效声音信息进行处理,获得第三结果,然后对第一算法、第二算法、第三算法得到的第一结果、

第二结果以及第三结果分别计算,得到对应的三个声纹信息,得三个的各声纹信息为历史语音信息的声纹信息。

[0061] 本实施例提供的一种历史声纹信息的获取方法,在本申请中,通过对历史语音信息进行获取并提取人声语段,再对提取后声音信息通过算法进行处理,对处理后的结果计算声纹信息,以此得到历史语音信息的声纹信息,使得历史语音信息的声纹信息更加多元化,提高了历史语音信息的声纹信息的准确性。

[0062] 请参阅图4,图4示出了本申请实施例中偏差系数的一种获取方法的流程图,方法可以用于电子设备,例如可以是图1中的设备控制装置110,方法包括:

S310:分别确定多个预置算法各自的算法处理结果对应的声纹信息与预设唤醒词的声纹信息的声纹相似度。

[0063] 作为一种实施方式,电子设备中预先存储有唤醒词的声纹信息。当获得多个预置算法各自的算法处理结果对应的声纹信息之后,将各算法处理结果对应的声纹信息与预先存储有唤醒词的声纹信息一一进行比对,得到各声纹信息对应的声纹相似度。

[0064] 例如,历史声纹信息包括上述第一结果、第二结果以及第三结果分别对应的声纹信息,分别确定第一结果的声纹信息、第二结果的声纹信息以及第三结果的声纹信息与预设唤醒词的声纹信息的声纹相似度。

[0065] 应当理解的是,各声纹信息的声纹相似度的值的范围应该为大于等于0并小于等于1,声纹信息的声纹相似度所代表的意义应是声纹相似度越大,所对应的声纹信息与预设唤醒词的声纹信息越相似。

[0066] S320:将声纹相似度的最大值确定为目标声纹相似度。

[0067] 作为一种实施方式,将对各声纹相似度中最大值的声纹信息作为目标声纹信息,将目标声纹信息的声纹相似度作为目标声纹相似度。

[0068] S330:根据目标声纹相似度,得到偏差系数。

[0069] 在获得目标声纹相似度之后,根据目标声纹相似度,确定预设滤波器的偏差系数。

[0070] 作为一种实施方式,S330可以包括:计算目标声纹相似度的平方,作为第一计算结果;将预设常数与第一计算结果的差值作为偏差系数。

[0071] 示例性的,若目标声纹相似度为0.8,则目标相似度的平方为0.64作为第一计算结果。

[0072] 在本实施例中,预设常数应为不等于零的正数,例如预设常数可以设置为1。

[0073] 得到预设常数与第一计算结果之后,可以将预设常数减第一计算结果的计算结果为偏差系数的值。

[0074] 作为另一种实施方式,在第一次通过VAD提取人声语音提取有效声音信息的情况下,电子设备中还没有存储历史语音信息的声纹信息,因此无法得到目标声纹相似度及第一运算结果,此种情况下可以设置一个预设值作为偏差系数。该预设值的数值范围应大于等于0并小于等于1。

[0075] 本实施例提供的一种偏差系数的获取方法,在本申请中,通过确定历史语音信息的声纹信息与预设唤醒词的声纹信息相似度的最大值,以此来得到偏差系数,使得通过相似度最大值确定的偏差系数更加准确,为后续提高预设滤波器的处理性能做了准备。

[0076] 请参阅图5,图5示出了本申请实施例中自适应滤波器系数的一种修正方法的流程

图,方法可以用于电子设备,例如可以是图1中的设备控制装置110,方法包括:

S410:获取预设滤波器的修正前的自适应滤波器系数、迭代步长以及误差项,修正前的自适应滤波器系数是在通过预设滤波器对历史语音信息进行处理之前,预设滤波器中的自适应滤波器系数。

[0077] 在本实施例中,预设滤波器的修正前的自适应滤波器系数可以是指预设滤波器对历史语音信息进行处理之前,预设滤波器中的自适应滤波器系数。可以将预设滤波器的修正前的自适应滤波器系数记为 $h(n)$ ,将迭代步长记为 $u(n)$ ,将误差项记为 $e(n)$ 。

[0078] S420:根据迭代步长、误差项、偏差系数以及历史语音信息对应的特征信息,对修正前的自适应滤波器系数进行修正,得到修正后的自适应滤波器系数。

[0079] 电子设备得到历史语音信息之后,对历史语音信息进行特征提取,得到历史语音信息的特征信息,可以通过语音特征提取模型对历史语音信息进行特征提取,语音特征提取模型可以是对神经网络模型进行训练获得的。

[0080] 在得到迭代步长、误差项、偏差系数以及历史语音信息对应的特征信息之后,进一步对修正前的自适应滤波器系数进行修正,得到修正后的自适应滤波器系数。

[0081] 进一步的,S420可以包括:计算迭代步长、误差项、偏差系数以及历史语音信息对应的特征信息的积,作为第二计算结果;计算第二计算结果与修正前的自适应滤波器系数的和,作为预设滤波器的修正后的自适应滤波器系数。

[0082] 在本实施例中,将偏差系数记为 $c$ ,将历史语音信息对应的特征信息记为 $x(n)$ ,将预设滤波器的修正后的自适应滤波器系数记为 $h(n+1)$ 。其中, $x(n)$ 也可以是指历史语音信息对应的第二有效声音信息对应的特征信息。

[0083] 作为一种实施方式,按照下述公式进行计算,作为第二计算结果: $c*u(n)*e(n)*x(n)$ ,其中, $n$ 为迭代次数。

[0084] 预设滤波器的修正后的自适应滤波器系数 $h(n+1)$ 按照下述公式进行计算: $h(n+1)=h(n)+c*u(n)*e(n)*x(n)$ 。

[0085] 可以理解的是,在处理后的语音信息包括预设唤醒词时,根据处理后的语音信息对自适应滤波器系数 $h(n+1)$ 进行修正,并将修正后的自适应滤波器系数 $h(n+2)$ 作为下一次进行语音处理的自适应滤波器系数。其中,根据处理后的语音信息对自适应滤波器系数 $h(n+1)$ 进行修正的过程如下:

电子设备得到待处理语音信息对应的第一有效声音信息;再通过第四算法(将0作为偏差系数对自适应滤波器进行配置得到该第四算法)对第一有效声音信息进行处理,获得第四结果,通过第五算法(根据第一有效声音信息得到新的偏差系数,并通过新的偏差系数的绝对值对预设滤波器进行配置得到该第五算法)对第一有效声音信息进行处理,获得第五结果,通过第六算法(根据第一有效声音信息得到新的偏差系数,并通过新的偏差系数的绝对值的相反数对预设滤波器进行配置得到该第六算法)对第一有效声音信息进行处理,获得第六结果,根据第四结果、第五结果、第六结果,分别确定第四结果的声纹信息、第五结果的声纹信息以及第六结果的声纹信息;

再分别确定第四结果的声纹信息、第五结果的声纹信息以及第六结果的声纹信息与预设唤醒词的声纹信息的声纹相似度,将声纹相似度的最大值确定为新的目标声纹相似度,计算新的目标声纹相似度的平方,作为第三计算结果,将预设常数与第三计算结果的差

值作为新的偏差系数;获取预设滤波器的迭代步长以及误差项,并根据新的偏差系数、迭代步长、误差项、第一有效声音信息对应的特征信息的积,作为第四计算结果;计算第四计算结果与本次得到的自适应滤波器系数的和,作为修正后的自适应滤波器系数 $h(n+2)$ ,从而得到修正后的预设滤波器。

[0086] 本实施例提供的一种自适应滤波器系数的修正方法,在本申请中,通过偏差系数对预设滤波器的自适应滤波器系数进行修正,使得修正后的预设滤波器能够在处理下一次语音信息时的能力增强,经过处理后的语音信息误差更小、更加准确。

[0087] 以下介绍本申请的装置实施例,可以用于执行本申请上述方法实施例中的方法。对于本申请装置实施例中未披露的细节,请参照本申请上述方法实施例。

[0088] 请参阅图6,图6示出了本申请一个实施例提出的一种设备控制装置的结构框图,该设备控制装置500包括:语音采集模块510、语音处理模块520、指令识别模块530、设备控制模块540。

[0089] 语音采集模块510,用于采集待处理语音信息。

[0090] 语音处理模块520,用于通过预设滤波器对待处理语音进行处理,获得处理后的语音信息,预设滤波器的自适应滤波器系数通过偏差系数进行修正,偏差系数基于预设唤醒词的声纹信息以及历史语音信息的声纹信息获得,历史语音信息在待处理语音之前获取的,历史语音信息包括预设唤醒词。

[0091] 指令识别模块530,用于在处理后的语音信息包括预设唤醒词时,根据处理后的语音信息获得用于控制目标设备的控制指令。

[0092] 设备控制模块540,用于根据控制指令,控制目标设备。

[0093] 在本实施例中,该设备控制装置500还包括:降噪及回声消除模块、人声语段提取模块;降噪及回声消除模块用于对待处理语音信息进行降噪及回声消除处理;人声语段提取模块用于从处理后的待处理语音中提取人声语段,得到第一有效声音信息;相应的,语音处理模块520用于通过预设滤波器对第一有效声音信息进行处理,获得处理后的语音信息。

[0094] 进一步的,语音处理模块520还包括:历史声纹信息获取模块、偏差系数计算模块、滤波器系数修正模块,其中:

历史声纹信息获取模块,用于对历史语音信息中提取的人声片段通过第一算法、第二算法、第三算法进行处理,获得第一结果、第二结果以及第三结果,并得到历史语音信息的声纹信息;

偏差系数计算模块,用于确定历史语音信息的声纹信息与预设唤醒词的声纹信息的相似度,并根据声纹相似度的最大值计算偏差系数;

滤波器系数修正模块,用于获取预设滤波器的修正前的自适应滤波器系数、迭代步长以及误差项,并根据历史语音信息对应的特征信息以及修正前的自适应滤波器系数,得到修正后的自适应滤波器系数。

[0095] 需要说明的是,本申请中的装置实施例与前述方法实施例是相互对应的,装置实施例中具体的原理可以参见前述方法实施例中的内容,此处不再赘述。

[0096] 图7是根据本申请实施例提出的一种电子设备600的结构框图。本申请实施例提供一种包括可以执行前述方法的处理器602的电子设备600,该电子设备600可以为终端设备,终端设备可以是智能手机、平板电脑、计算机、穿戴式设备、智能家居、智能机器人等。

[0097] 电子设备600还包括存储器604。其中,该存储器604中存储有可以执行前述实施例内容的程序,而处理器602可以执行该存储器604中存储的程序。

[0098] 其中,处理器602可以包括一个或者多个用于处理数据的核以及消息矩阵单元。处理器602利用各种接口和线路连接整个电子设备600内的各个部分,通过运行或执行存储在存储器604内的指令、程序、代码集或指令集,以及调用存储在存储器604内的数据,执行电子设备600的各种功能和处理数据。可选地,处理器602可以采用数字信号处理(Digital Signal Processing,DSP)、现场可编程门阵列(Field-Programmable Gate Array,FPGA)、可编程逻辑阵列(Programmable Logic Array,PLA)中的至少一种硬件形式来实现。处理器602可集成中央处理器(Central Processing Unit,CPU)、图像处理(Graphics Processing Unit,GPU)和调制解调器等中的一种或几种的组合。其中,CPU主要处理操作系统、用户界面和应用程序等;GPU用于负责显示内容的渲染和绘制;调制解调器用于处理无线通信。可以理解的是,上述调制解调器也可以不集成到处理器602中,单独通过一块通信芯片进行实现。

[0099] 存储器604可以包括随机存储器(Random Access Memory,RAM),也可以包括只读存储器(Read-Only Memory)。存储器604可用于存储指令、程序、代码、代码集或指令集。存储器604可包括存储程序区和存储数据区,其中,存储程序区可存储用于实现操作系统的指令、用于实现至少一个功能的指令、用于实现下述各个方法实施例的指令等。存储数据区还可以存储电子设备600在使用中所获取的数据(如,待推荐数据以及操作方式)等。

[0100] 电子设备600还可以包括网络模块以及屏幕,网络模块用于接收以及发送电磁波,实现电磁波与电信号的相互转换,从而与通讯网络或者其他设备进行通讯,例如和音频播放设备进行通讯。网络模块可包括各种现有的用于执行这些功能的电路元件,例如,天线、射频收发器、数字信号处理器、加密/解密芯片、用户身份模块(SIM)卡、存储器等等。网络模块可与各种网络如互联网、企业内部网、无线网络进行通讯或者通过无线网络与其他设备进行通讯。上述的无线网络可包括蜂窝式电话网、无线局域网或者城域网。屏幕可以进行界面内容的显示以及进行数据交互。

[0101] 在一些实施例中,电子设备600还可以包括有:外设接口606和至少一个外围设备。处理器602、存储器604和外设接口606之间可以通过总线或信号线相连。各个外围设备可以通过总线、信号线或电路板与外设接口连接。具体地,外围设备包括:射频组件608、定位组件612、摄像头614、音频组件616、显示屏618以及电源622等中的至少一种

外设接口606可被用于将I/O(Input/Output,输入/输出)相关的至少一个外围设备连接到处理器602和存储器604。在一些实施例中,处理器602、存储器604和外设接口606被集成在同一芯片或电路板上;在一些其他实施例中,处理器602、存储器604和外设接口606中的任意一个或两个可以在单独的芯片或电路板上实现,本申请实施例对此不加以限定。

[0102] 射频组件608用于接收和发射RF(Radio Frequency,射频)信号,也称电磁信号。射频组件608通过电磁信号与通信网络以及其他通信设备进行通信。射频组件608将电信号转换为电磁信号进行发送,或者,将接收到的电磁信号转换为电信号。可选地,射频组件608包括:天线系统、RF收发器、一个或多个放大器、调谐器、振荡器、数字信号处理器、编解码芯片组、用户身份模块卡等等。射频组件608可以通过至少一种无线通信协议来与其它终端进行

通信。该无线通信协议包括但不限于：万维网、城域网、内联网、各代移动通信网络(2G、3G、4G及5G)、无线局域网和/或WiFi(Wireless Fidelity,无线保真)网络。在一些实施例中,射频组件708还可以包括NFC(Near Field Communication,近距离无线通信)有关的电路,本申请对此不加以限定。

[0103] 定位组件612用于定位电子设备600的当前地理位置,以实现导航或LBS(LocationBased Service,基于位置的服务)。定位组件612可以是基于美国的GPS(Global Positioning System,全球定位系统)、北斗系统或伽利略系统的定位组件。

[0104] 摄像头614用于采集图像或视频。可选地,摄像头614包括前置摄像头和后置摄像头。通常,前置摄像头设置在电子设备600的前面板,后置摄像头设置在电子设备600的背面。在一些实施例中,后置摄像头为至少两个,分别为主摄像头、景深摄像头、广角摄像头、长焦摄像头中的任意一种,以实现主摄像头和景深摄像头融合实现背景虚化功能、主摄像头和广角摄像头融合实现全景拍摄以及VR(Virtual Reality,虚拟现实)拍摄功能或者其他融合拍摄功能。在一些实施例中,摄像头614还可以包括闪光灯。闪光灯可以是单色温闪光灯,也可以是双色温闪光灯。双色温闪光灯是指暖光闪光灯和冷光闪光灯的组合,可以用于不同色温下的光线补偿。

[0105] 音频组件616可以包括麦克风和扬声器。麦克风用于采集用户及环境的声波,并将声波转换为电信号输入至处理器602进行处理,或者输入至射频组件608以实现语音通信。出于立体声采集或降噪的目的,麦克风可以为多个,分别设置在电子设备600的不同部位。麦克风还可以是阵列麦克风或全向采集型麦克风。扬声器则用于将来自处理器602或射频组件608的电信号转换为声波。扬声器可以是传统的薄膜扬声器,也可以是压电陶瓷扬声器。当扬声器是压电陶瓷扬声器时,不仅可以将电信号转换为人类可听见的声波,也可以将电信号转换为人类听不见的声波以进行测距等用途。在一些实施例中,音频组件614还可以包括耳机插孔。

[0106] 显示屏618用于显示UI(User Interface,用户界面)。该UI可以包括图形、文本、图标、视频及其它们的任意组合。当显示屏618是触摸显示屏时,显示屏618还具有采集在显示屏618的表面或表面上方的触摸信号的能力。该触摸信号可以作为控制信号输入至处理器602进行处理。此时,显示屏618还可以用于提供虚拟按钮和/或虚拟键盘,也称软按钮和/或软键盘。在一些实施例中,显示屏618可以为一个,设置电子设备600的前面板;在另一些实施例中,显示屏618可以为至少两个,分别设置在电子设备600的不同表面或呈折叠设计;在又一些实施例中,显示屏618可以是柔性显示屏,设置在电子设备600的弯曲表面上或折叠面上。甚至,显示屏618还可以设置成非矩形的不规则图形,也即异形屏。显示屏618可以采用LCD(Liquid Crystal Display,液晶显示屏)、OLED(Organic Light-Emitting Diode,机发光二极管)等材质制备。

[0107] 电源622用于为电子设备600中的各个组件进行供电。电源622可以是交流电、直流电、一次性电池或可充电电池。当电源622包括可充电电池时,该可充电电池可以是有线充电电池或无线充电电池。有线充电电池是通过有线线路充电的电池,无线充电电池是通过无线线圈充电的电池。该可充电电池还可以用于支持快充技术。

[0108] 本申请实施例还提供一种计算机可读存储介质。该计算机可读介质中存储有程序代码,程序代码可被处理器调用执行上述方法实施例中所述的方法。

[0109] 计算机可读存储介质可以是诸如闪存、EEPROM(电可擦除可编程只读存储器)、EPROM、硬盘或者ROM之类的电子存储器。可选地,计算机可读存储介质包括非易失性计算机可读介质(non-transitory computer-readable storage medium)。计算机可读存储介质具有执行上述方法中的任何方法步骤的程序代码的存储空间。这些程序代码可以从一个或者多个计算机程序产品中读出或者写入到这一个或者多个计算机程序产品中。程序代码可以例如以适当形式进行压缩。

[0110] 本申请实施例还提供了一种计算机程序产品或计算机程序,该计算机程序产品或计算机程序包括计算机指令,该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令,处理器执行该计算机指令,使得该计算机设备执行上述各种可选实现方式中描述的方法。

[0111] 综上所述,本申请公开了一种设备控制方法、装置、电子设备及计算机可读存储介质,在本申请中,预设滤波器的自适应滤波器系数通过偏差系数进行修正,包括修正后的自适应滤波器系数的预设滤波器的语音处理能力较强,使得通过预设滤波器对待处理语音信息处理后获得的处理后的语音信息包括较少的噪声、且更加准确,从而使得根据待处理语音信息获得的控制指令更加准确,提高了设备控制的准确性。

[0112] 应当注意,尽管在上文详细描述中提及了用于动作执行的设备的若干模块或者单元,但是这种划分并非强制性的。实际上,根据本申请的实施方式,上文描述的两个或更多模块或者单元的特征和功能可以在一个模块或者单元中具体化。反之,上文描述的一个模块或者单元的特征和功能可以进一步划分为由多个模块或者单元来具体化。

[0113] 通过以上的实施方式的描述,本领域的技术人员易于理解,这里描述的示例实施方式可以通过软件实现,也可以通过软件结合必要的硬件的方式来实现。因此,根据本申请实施方式的技术方案可以以软件产品的形式体现出来,该软件产品可以存储在一个非易失性存储介质(可以是CD-ROM,U盘,移动硬盘等)中或网络上,包括若干指令以使得一台计算设备(可以是个人计算机、服务器、触控终端、或者网络设备等)执行根据本申请实施方式的方法。

[0114] 本领域技术人员在考虑说明书及实践这里公开的实施方式后,将容易想到本申请的其它实施方案。本申请旨在涵盖本申请的任何变型、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申请的一般性原理并包括本申请未公开的本技术领域中的公知常识或惯用技术手段。

[0115] 应当理解的是,本申请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。本申请的范围仅由所附的权利要求来限制。

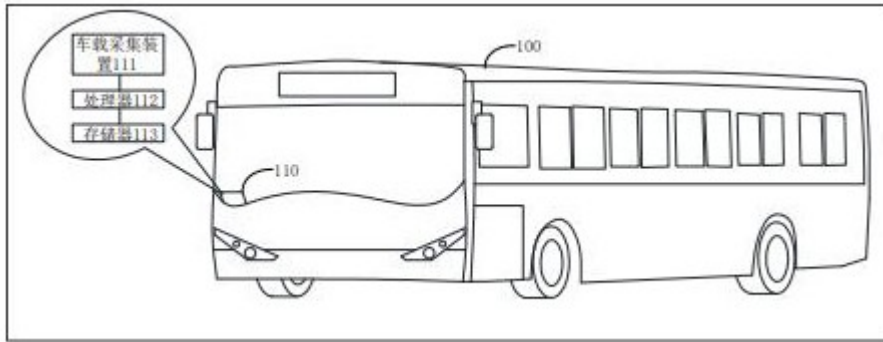


图1

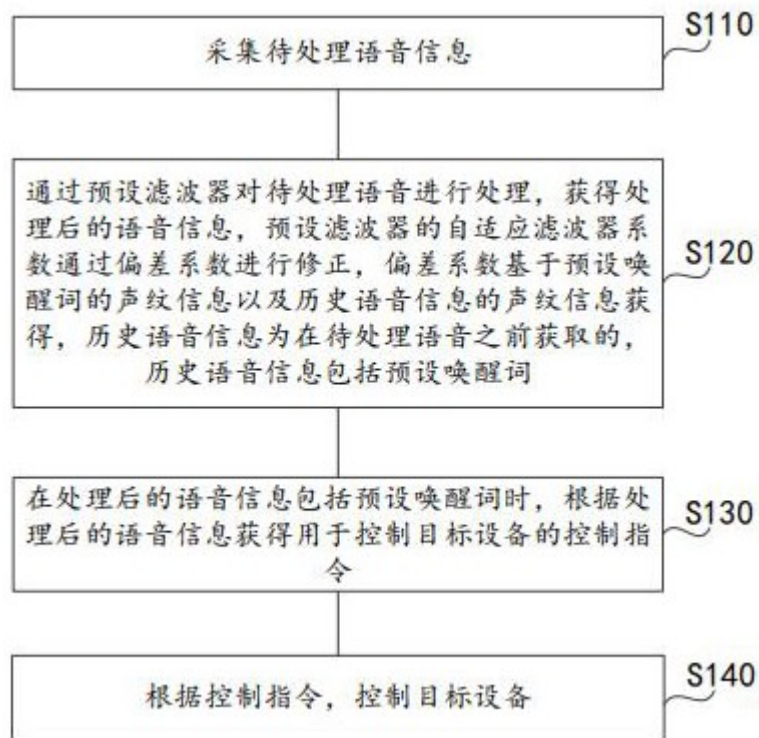


图2



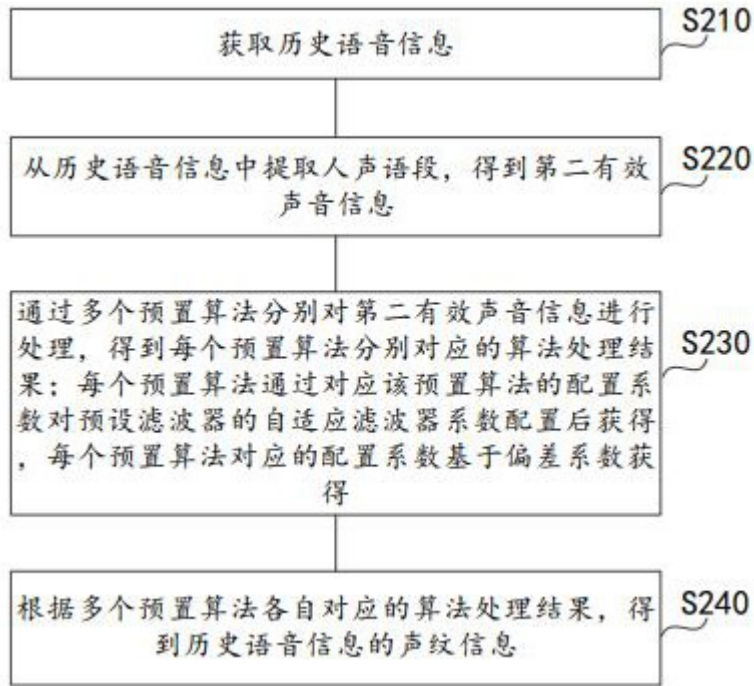


图3

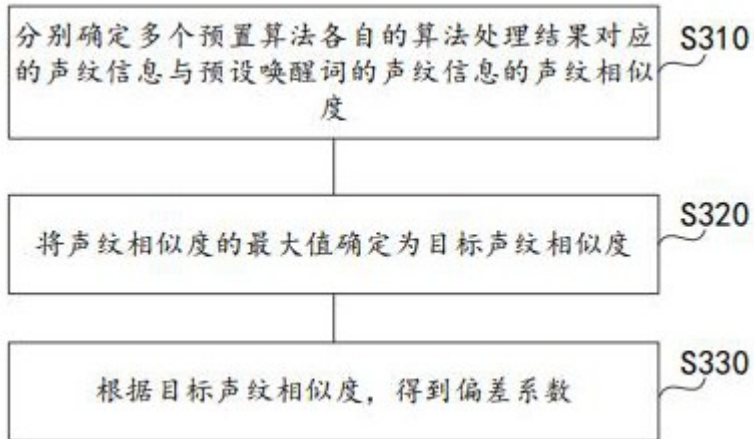


图4

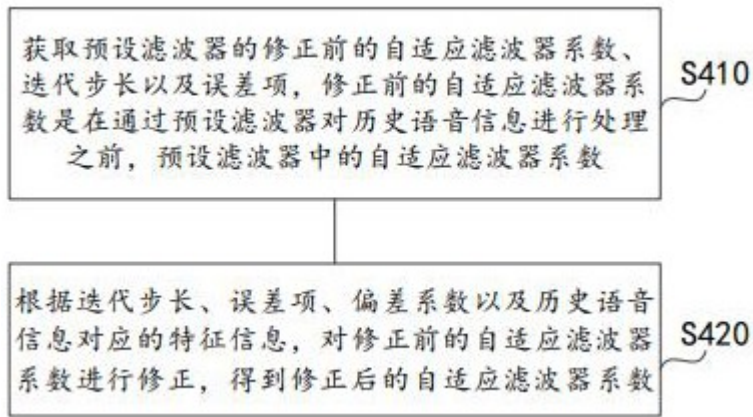


图5

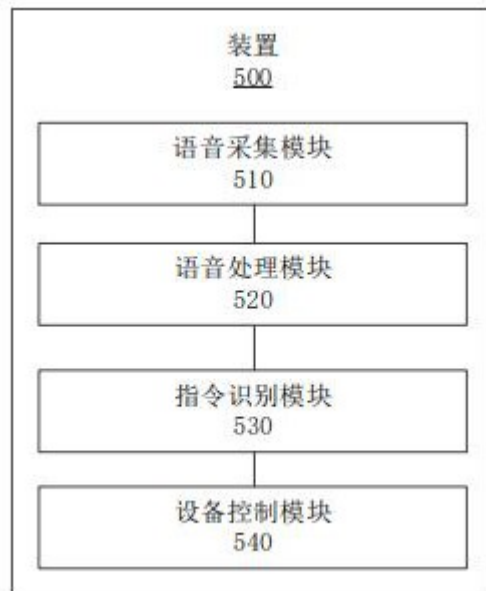


图6

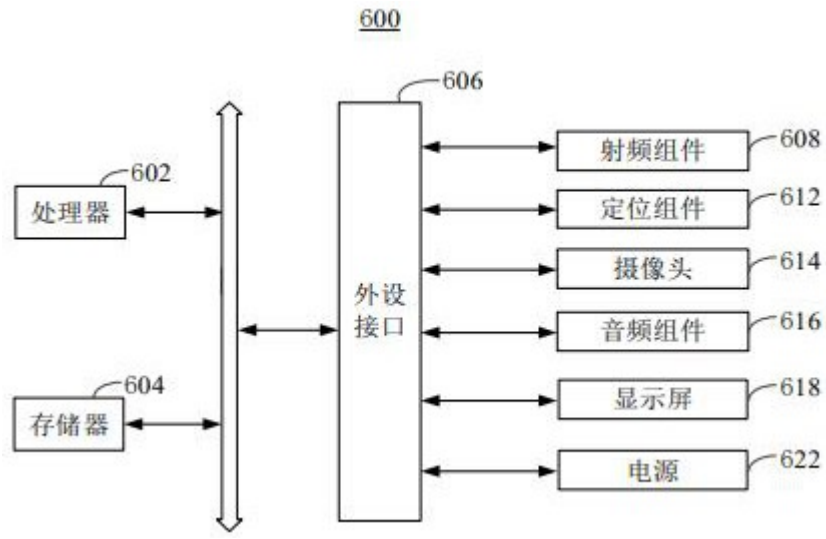


图7