



(12)发明专利申请

(10)申请公布号 CN 111314187 A

(43)申请公布日 2020.06.19

(21)申请号 202010071543.8

(22)申请日 2020.01.21

(71)申请人 星络智能科技有限公司

地址 518000 广东省深圳市南山区粤海街道海珠社区海德三道1698号航天科技广场A座31A01

(72)发明人 付杰

(51)Int.Cl.

H04L 12/28(2006.01)

G10L 15/22(2006.01)

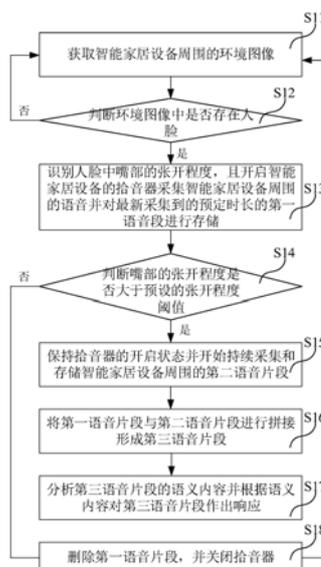
权利要求书2页 说明书6页 附图3页

(54)发明名称

存储介质、智能家居设备及其唤醒方法

(57)摘要

本申请公开了一种存储介质、智能家居设备及其唤醒方法,该唤醒方法包括:获取智能家居设备周围的环境图像;判断环境图像中是否存在人脸;在判断到环境图像中存在人脸时,识别人脸中嘴部的张开程度,且开启智能家居设备的拾音器采集智能家居设备周围的语音并对最新采集到的预定时长的第一语音段进行存储;判断嘴部的张开程度是否大于预设的张开程度阈值;在判断到嘴部的张开程度大于预设的张开程度阈值时,保持拾音器的开启状态并开始持续采集和存储智能家居设备周围的第二语音片段,将第一语音片段与第二语音片段进行拼接形成第三语音片段;分析第三语音片段的语义内容并根据语义内容对第三语音片段作出响应。本申请能够方便用户与智能家居设备的语音交互。



1. 一种智能家居设备的唤醒方法,其特征在于,所述唤醒方法包括:

获取所述智能家居设备周围的环境图像;

判断所述环境图像中是否存在人脸;

在判断到所述环境图像中存在人脸时,识别所述人脸中嘴部的张开程度,且开启所述智能家居设备的拾音器采集所述智能家居设备周围的语音并对最新采集到的预定时长的第一语音段进行存储;

判断所述嘴部的张开程度是否大于预设的张开程度阈值;

在判断到所述嘴部的张开程度大于所述预设的张开程度阈值时,保持所述拾音器的开启状态并开始持续采集和存储所述智能家居设备周围的第二语音片段,将所述第一语音片段与所述第二语音片段进行拼接形成第三语音片段;

分析所述第三语音片段的语义内容并根据所述语义内容对所述第三语音片段作出响应。

2. 根据权利要求1所述的唤醒方法,其特征在于,所述唤醒方法还包括:

在判断到所述嘴部的张开程度小于或者等于所述预设的张开程度阈值时,删除所述第一语音片段,且在删除所述第一语音片段后返回所述获取所述智能家居设备周围的环境图像的步骤。

3. 根据权利要求1所述的唤醒方法,其特征在于,所述唤醒方法还包括:

在判断到所述嘴部的张开程度小于或者等于所述预设的张开程度阈值时,关闭所述拾音器。

4. 根据权利要求1所述的唤醒方法,其特征在于,所述将所述第一语音片段与所述第二语音片段进行拼接形成第三语音片段的步骤包括:

将完整的所述第一语音片段与所述第二语音片段进行拼接形成所述第三语音片段。

5. 根据权利要求1所述的唤醒方法,其特征在于,所述将所述第一语音片段与所述第二语音片段进行拼接形成第三语音片段的步骤包括:

对所述第一语音片段进行语音端点检测,以获取最接近所述第一语音片段的语音结束时间点的语音间隔时间点;

生成从所述语音间隔时间点到所述语音结束时间点的第四语音片段;

将所述第四语音片段与所述第二语音片段进行拼接形成所述第三语音片段。

6. 根据权利要求1所述的唤醒方法,其特征在于,所述将所述第一语音片段与所述第二语音片段进行拼接形成第三语音片段的步骤包括:

识别所述第一语音片段并判断所述第一语音片段中是否存在预设的关键词;

在判断到所述第一语音片段中存在预设的关键词时,生成从所述关键词到所述第一语音片段的语音结束时间点的第五语音片段,并将所述第五语音片段与所述第二语音片段进行拼接形成所述第三语音片段。

7. 根据权利要求6所述的唤醒方法,其特征在于,所述将所述第一语音片段与所述第二语音片段进行拼接形成第三语音片段的步骤还包括:

在判断到所述第一语音片段中不存在预设的关键词时,将完整的所述第一语音片段与所述第二语音片段进行拼接形成所述第三语音片段。

8. 根据权利要求1所述的唤醒方法,其特征在于,所述开启所述智能家居设备的拾音器

采集所述智能家居设备周围的语音并对最新采集到的预定时长的第一语音段进行存储的步骤包括：

从当前时间点开始计算，对预定时长之前存储的语音数据进行删除。

9. 一种智能家居设备，其特征在于，所述智能家居设备包括处理器和与所述处理器电连接的存储器，所述存储器用于存储计算机程序，所述处理器用于调用所述计算机程序以执行权利要求1-8任意一项所述的方法。

10. 一种存储介质，其特征在于，所述存储介质存储有计算机程序，所述计算机程序能够被处理器执行以实现权利要求1-8任意一项所述的方法。

存储介质、智能家居设备及其唤醒方法

技术领域

[0001] 本发明涉及智能家居技术领域,特别是涉及一种存储介质、智能家居设备及其唤醒方法。

背景技术

[0002] 智能家居是在互联网影响之下物联化的体现。智能家居通过物联网技术将家中的各种设备(如音视频设备、照明系统、窗帘控制、空调控制、安防系统、数字影院系统、影音服务器、影柜系统、网络家电等)连接到一起,提供家电控制、照明控制、电话远程控制、室内外遥控、防盗报警、环境监测、暖通控制、红外转发以及可编程定时控制等多种功能和手段。与普通家居相比,智能家居不仅具有传统的居住功能,兼备建筑、网络通信、信息家电、设备自动化,提供全方位的信息交互功能,甚至为各种能源费用节约资金。

[0003] 语音交互是智能家居设备的常见功能,现有的智能家居设备在唤醒时通常需要特定的唤醒词,在输入唤醒词之后智能家居设备才会被唤醒,唤醒后再与智能家居设备进行语音交互,实现对家中设备的控制等人机交互,这种方式导致每次语音交互前都要输入唤醒词,操作不方便,效率低。

发明内容

[0004] 本申请主要解决的技术问题是提供一种存储介质、智能家居设备及其唤醒方法,能够实现智能家居设备的免唤醒词唤醒,方便用户与智能家居设备的语音交互。

[0005] 为解决上述技术问题,本申请实施例采用的一个技术方案是:提供一种智能家居设备的唤醒方法,该唤醒方法包括:获取智能家居设备周围的环境图像;判断环境图像中是否存在人脸;在判断到环境图像中存在人脸时,识别人脸中嘴部的张开程度,且开启智能家居设备的拾音器采集智能家居设备周围的语音并对最新采集到的预定时长的第一语音段进行存储;判断嘴部的张开程度是否大于预设的张开程度阈值;在判断到嘴部的张开程度大于预设的张开程度阈值时,保持拾音器的开启状态并开始持续采集和存储智能家居设备周围的第二语音片段,将第一语音片段与第二语音片段进行拼接形成第三语音片段;分析第三语音片段的语义内容并根据语义内容对第三语音片段作出响应。

[0006] 其中,唤醒方法还包括:在判断到嘴部的张开程度小于或者等于预设的张开程度阈值时,删除第一语音片段,并返回获取智能家居设备周围的环境图像的步骤。

[0007] 其中,唤醒方法还包括:在判断到嘴部的张开程度小于或者等于预设的张开程度阈值时,关闭拾音器。

[0008] 其中,将第一语音片段与第二语音片段进行拼接形成第三语音片段的步骤包括:将完整的第一语音片段与第二语音片段进行拼接形成第三语音片段。

[0009] 其中,将第一语音片段与第二语音片段进行拼接形成第三语音片段的步骤包括:对第一语音片段进行语音端点检测,以获取最接近第一语音片段的语音结束时间点的语音间隔时间点;生成从语音间隔时间点到语音结束时间点的第四语音片段;将第四语音片段

与第二语音片段进行拼接形成第三语音片段。

[0010] 其中,将第一语音片段与第二语音片段进行拼接形成第三语音片段的步骤包括:识别第一语音片段并判断第一语音片段中是否存在预设的关键词;在判断到第一语音片段中存在预设的关键词时,生成从关键词到第一语音片段的语音结束时间点的第五语音片段,并将第五语音片段与第二语音片段进行拼接形成第三语音片段。

[0011] 其中,将第一语音片段与第二语音片段进行拼接形成第三语音片段的步骤还包括:在判断到第一语音片段中不存在预设的关键词时,将完整的第一语音片段与第二语音片段进行拼接形成第三语音片段。

[0012] 其中,开启智能家居设备的拾音器采集智能家居设备周围的语音并对最新采集到的预定时长的第一语音段进行存储的步骤包括:从当前时间点开始计算,对预定时长之前存储的语音数据进行删除。

[0013] 为解决上述技术问题,本申请实施例采用的另一个技术方案是:提供一种智能家居设备,该智能家居设备包括处理器和与处理器电连接的存储器,存储器用于存储计算机程序,处理器用于调用计算机程序以执行上述的方法。

[0014] 为解决上述技术问题,本申请实施例采用的又一个技术方案是:提供一种存储介质,该存储介质存储有计算机程序,计算机程序能够被处理器执行以实现上述的方法。

[0015] 本申请实施例通过获取智能家居设备周围的环境图像;判断环境图像中是否存在人脸;在判断到环境图像中存在人脸时,识别人脸中嘴部的张开程度,且开启智能家居设备的拾音器采集智能家居设备周围的语音并对最新采集到的预定时长的第一语音段进行存储;判断嘴部的张开程度是否大于预设的张开程度阈值;在判断到嘴部的张开程度大于预设的张开程度阈值时,保持拾音器的开启状态并开始持续采集和存储智能家居设备周围的第二语音片段,将第一语音片段与第二语音片段进行拼接形成第三语音片段;分析第三语音片段的语义内容并根据语义内容对第三语音片段作出响应,能够实现智能家居设备的免唤醒词唤醒,方便用户与智能家居设备的语音交互。

附图说明

[0016] 图1是本申请实施例智能家居设备的唤醒方法的流程示意图;

[0017] 图2是本申请实施例将第一语音片段与第二语音片段进行拼接形成第三语音片段的第一种实施方式的流程示意图;

[0018] 图3是本申请实施例将第一语音片段与第二语音片段进行拼接形成第三语音片段的第二种实施方式的流程示意图;

[0019] 图4是本申请实施例将第一语音片段与第二语音片段进行拼接形成第三语音片段的第三种实施方式的流程示意图;

[0020] 图5是本申请实施例智能家居设备的硬件结构示意图。

具体实施方式

[0021] 下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述。可以理解的是,此处所描述的具体实施例仅用于解释本申请,而非对本申请的限定。另外还需要说明的是,为了便于描述,附图中仅示出了与本申请相关的部分而非全部结

构。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范畴。

[0022] 本申请中的术语“第一”、“第二”等是用于区别不同对象，而不是用于描述特定顺序。此外，术语“包括”和“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，而是可选地还包括没有列出的步骤或单元，或可选地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。

[0023] 在本文中提及“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。本领域技术人员显式地和隐式地理解的是，本文所描述的实施例可以与其它实施例相结合。

[0024] 请参阅图1，图1是本申请实施例智能家居设备的唤醒方法的流程示意图。

[0025] 在本实施例中，智能家居设备的唤醒方法可以包括以下步骤：

[0026] 步骤S11：获取智能家居设备周围的环境图像。

[0027] 其中，智能家居设备可以是智能面板、智能音箱等。智能家居设备包括摄像头，通过智能家居设备的摄像头获取智能家居设备周围的环境图像。可选地，智能面板可以是安装在屋内墙体上的智能面板，或者是摆放在桌面的摆放式智能面板等。

[0028] 步骤S12：判断环境图像中是否存在人脸。

[0029] 其中，智能家居设备还包括处理器，处理器判断摄像头获取的环境图像中是否存在人脸。

[0030] 在步骤S12中，若是，即在判断到环境图像中存在人脸时，执行步骤S13-S14。

[0031] 步骤S13：识别人脸中嘴部的张开程度，且开启智能家居设备的拾音器采集智能家居设备周围的语音并对最新采集到的预定时长的第一语音段进行存储。

[0032] 其中，在判断到存在人脸时进一步识别人脸中嘴部的张开程度，与此同时，开启智能家居设备的拾音器采集语音，并以队列的形式存储最新采集到的预定时长的第一语音段，即第一语音段在不断的更新中，对预定时长前存储的语音数据进行删除，对最近存储的数据进行保存。

[0033] 其中，开启智能家居设备的拾音器采集智能家居设备周围的语音并对最新采集到的预定时长的第一语音段进行存储的步骤包括：从当前时间点开始计算，对预定时长之前存储的语音数据进行删除。

[0034] 步骤S14：判断嘴部的张开程度是否大于预设的张开程度阈值。

[0035] 在步骤S14中，若是，即在判断到嘴部的张开程度大于预设的张开程度阈值时，执行步骤S15-S17。

[0036] 在步骤S14中，若否，即在判断到嘴部的张开程度小于预设的张开程度阈值时，执行步骤S18，并在执行完步骤S18后返回步骤S11。

[0037] 步骤S15：保持拾音器的开启状态并开始持续采集和存储智能家居设备周围的第二语音片段。

[0038] 其中，如果识别到嘴部的张开程度大于预设张开程度阈值，则开始正常的语音存储方式，开始持续采集和存储智能家居设备周围的第二语音片段，不再对预定时间之前存

储的语音数据进行删除。

[0039] 步骤S16:将第一语音片段与第二语音片段进行拼接形成第三语音片段。

[0040] 其中,为了防止在识别到嘴部的张开程度大于预设的张开程度阈值之前的语音遗漏,将第一语音片段与第二语音片段进行拼接形成第三语音片段。

[0041] 其中,将第一语音片段与第二语音片段进行拼接形成第三语音片段的步骤包括:将完整的第一语音片段与第二语音片段进行拼接形成第三语音片段。

[0042] 步骤S17:分析第三语音片段的语义内容并根据语义内容对第三语音片段作出响应。

[0043] 例如,第三语音片段的语义内容为“今天天气怎样?”。那么智能家居设备对此作出的响应可以是“晴天,15-18摄氏度,湿度65%”

[0044] 步骤S18:删除第一语音片段,并关闭拾音器。

[0045] 在本实施例中,先是判断获取的周围环境图像中是否存在人脸,判断到存在人脸时,开始以队列的形式存储最近的一段语音,判断人脸中嘴部的张开程度,然后在张开程度大于阈值时,开始持续存储完整的语音,一方面可以不需要唤醒词进行准确的唤醒,另一方面,可以避免语音的遗漏。

[0046] 请参阅图2,图2是本申请实施例将第一语音片段与第二语音片段进行拼接形成第三语音片段的的第一种实施方式的流程示意图。

[0047] 在本实施例中,将第一语音片段与第二语音片段进行拼接形成第三语音片段的步骤具体可以包括:

[0048] 步骤S21:对第一语音片段进行语音端点检测,以获取最接近第一语音片段的语音结束时间点的语音间隔时间点。

[0049] 其中,第一语音片段的内容并非所有的内容都有效,可以通过语音端点检测语音间隔时间段,去除冗余的部分,然后再与第二语音片段进行拼接。

[0050] 步骤S22:生成从语音间隔时间点到语音结束时间点的第四语音片段。

[0051] 步骤S23:将第四语音片段与第二语音片段进行拼接形成第三语音片段。

[0052] 请参阅图3,图3是本申请实施例将第一语音片段与第二语音片段进行拼接形成第三语音片段的第二种实施方式的流程示意图。

[0053] 在本实施例中,将第一语音片段与第二语音片段进行拼接形成第三语音片段的步骤具体可以包括:

[0054] 步骤S31:识别第一语音片段并判断第一语音片段中是否存在预设的关键词。

[0055] 其中,第一语音片段并非所有内容都为有效内容,可能对于智能家居设备而言,只有关键词出现后的内容有效,关键词可以是多个,形成关键词数据库,例如,“我想”、“我查查”、“我想知道”等等。将第一语音片段中关键词之后的内容与第二语音片段进行拼接,获取整个有效的语音片段。

[0056] 在步骤S31中,若是,即在判断到第一语音片段中存在预设的关键词时,执行步骤S32。

[0057] 在步骤S31中,若否,即在判断到第一语音片段中不存在预设的关键词时,执行步骤S33。

[0058] 步骤S32:生成从关键词到第一语音片段的语音结束时间点的第五语音片段,并将

第五语音片段与第二语音片段进行拼接形成第三语音片段。

[0059] 步骤S33:将完整的第一语音片段与第二语音片段进行拼接形成第三语音片段。

[0060] 请参阅图4,图4是本申请实施例将第一语音片段与第二语音片段进行拼接形成第三语音片段的第三种实施方式的流程示意图。

[0061] 在本实施例中,将第一语音片段与第二语音片段进行拼接形成第三语音片段的步骤具体可以包括:

[0062] 步骤S41:识别第一语音片段并判断第一语音片段中是否存在预设的关键词。

[0063] 在步骤S41中,若是,即在判断到第一语音片段中存在预设的关键词时,执行步骤S42。

[0064] 在步骤S41中,若否,即在判断到第一语音片段中不存在预设的关键词时,执行步骤S43-S45。

[0065] 步骤S42:生成从关键词到第一语音片段的语音结束时间点的第五语音片段,并将第五语音片段与第二语音片段进行拼接形成第三语音片段。

[0066] 步骤S43:对第一语音片段进行语音端点检测,以获取最接近第一语音片段的语音结束时间点的语音间隔时间点。

[0067] 步骤S44:生成从语音间隔时间点到语音结束时间点的第四语音片段。

[0068] 步骤S45:将第四语音片段与第二语音片段进行拼接形成第三语音片段。

[0069] 请参阅图5,图5是本申请实施例智能家居设备的硬件结构示意图。

[0070] 在本实施例中,智能家居设备50包括处理器51和与处理器52电连接的存储器52,存储器52用于存储计算机程序,处理器52用于调用该计算机程序以执行上述任意一实施例所描述的方法。

[0071] 本申请实施例还提供一种存储介质,该存储介质存储有计算机程序,该计算机程序被处理器执行时能够实现上述任一实施例的方法。

[0072] 其中,该计算机程序可以以软件产品的形式存储在上述存储介质中,包括若干指令用以使得一个设备或处理器执行本申请各个实施方式方法的全部或部分步骤。

[0073] 存储介质是计算机存储器中用于存储某种不连续物理量的媒体。而前述的存储介质可以为:U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0074] 在本申请所提供的几个实施方式中,应该理解到,所揭露的方法及设备,可以通过其它的方式实现。例如,以上所描述的设备实施方式仅仅是示意性的,例如,模块或单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。

[0075] 作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施方式方案的目的。

[0076] 另外,在本发明各个实施方式中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0077] 本申请实施例通过获取智能家居设备周围的环境图像;判断环境图像中是否存在人脸;在判断到环境图像中存在人脸时,识别人脸中嘴部的张开程度,且开启智能家居设备的拾音器采集智能家居设备周围的语音并对最新采集到的预定时长的第一语音段进行存储;判断嘴部的张开程度是否大于预设的张开程度阈值;在判断到嘴部的张开程度大于预设的张开程度阈值时,保持拾音器的开启状态并开始持续采集和存储智能家居设备周围的第二语音片段,将第一语音片段与第二语音片段进行拼接形成第三语音片段;分析第三语音片段的语义内容并根据语义内容对第三语音片段作出响应,能够实现智能家居设备的免唤醒词唤醒,方便用户与智能家居设备的语音交互。

[0078] 以上仅为本申请的实施方式,并非因此限制本申请的专利范围,凡是利用本申请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他相关的技术领域,均同理包括在本申请的专利保护范围内。

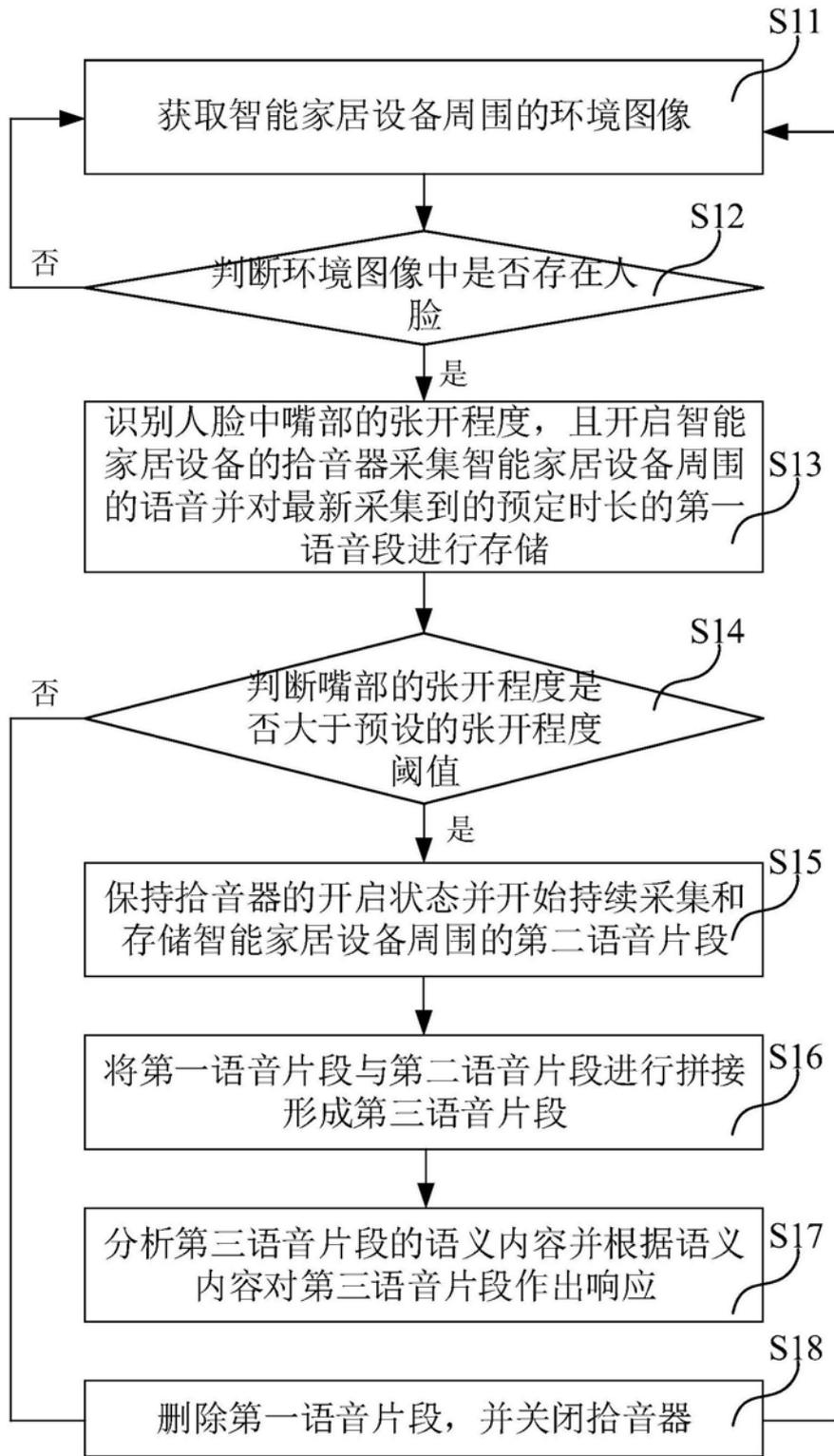


图1

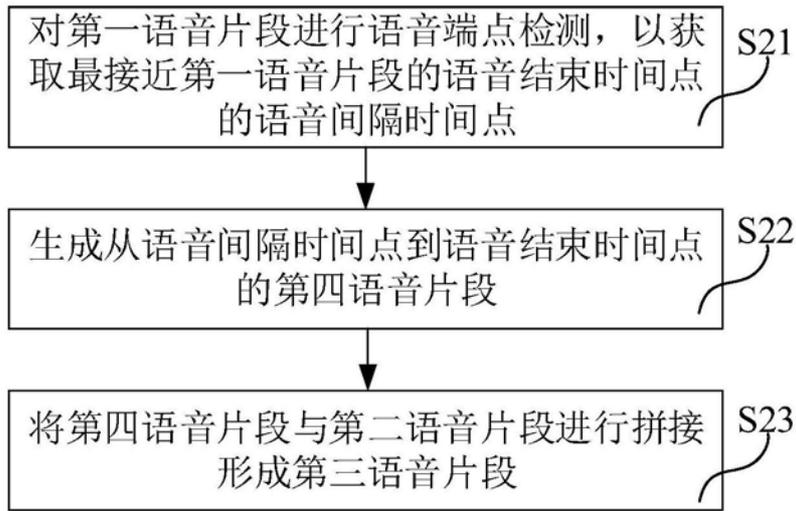


图2

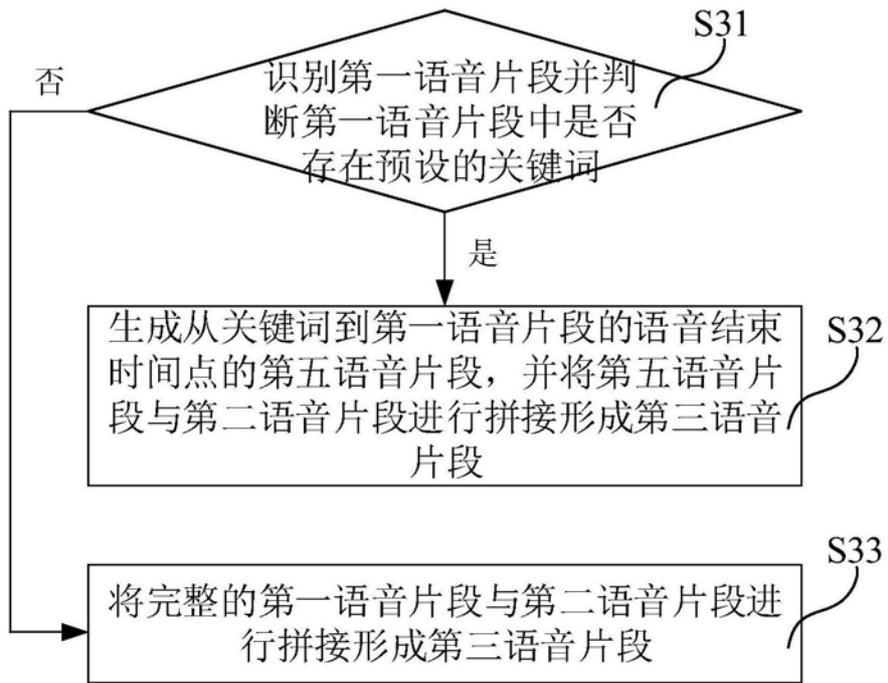


图3

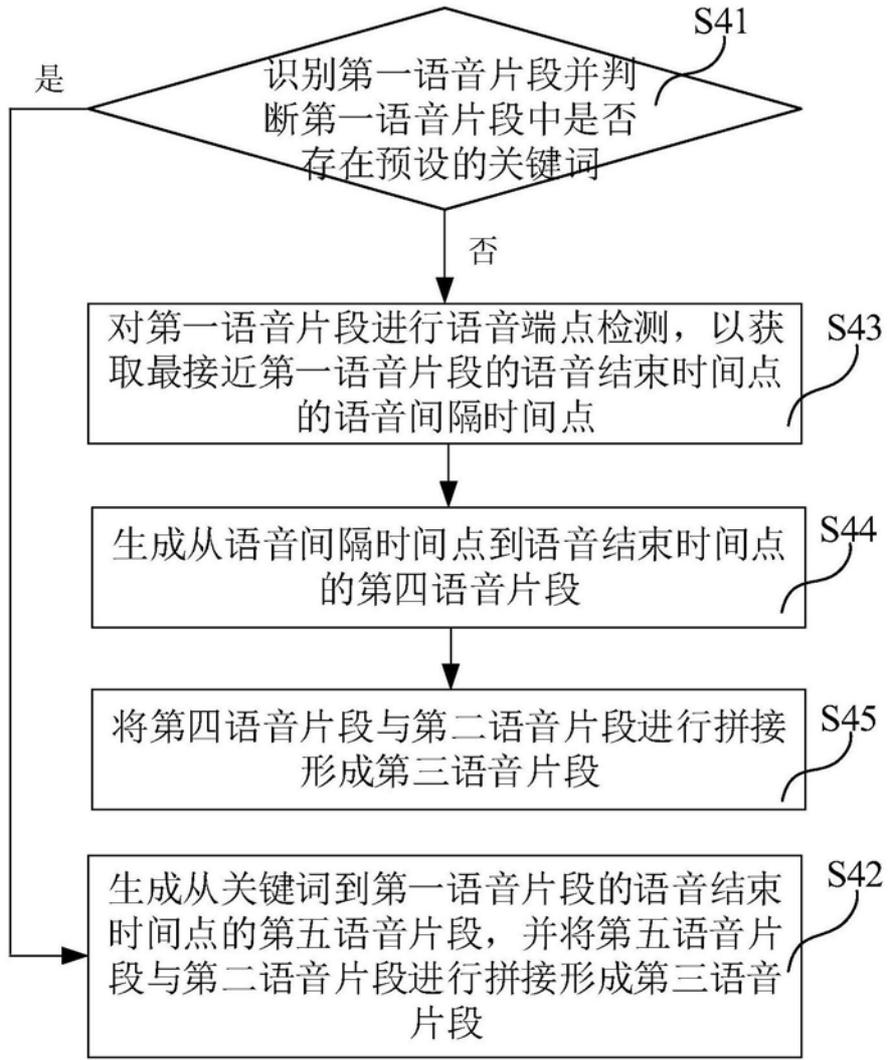


图4

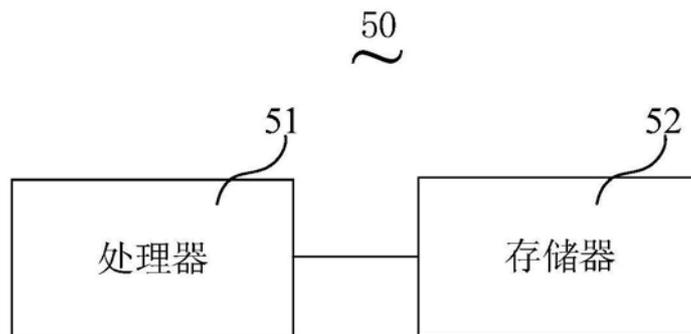


图5