



(12) 发明专利

(10) 授权公告号 CN 109841214 B

(45) 授权公告日 2021.06.01

(21) 申请号 201811590014.8

G10L 15/18 (2013.01)

(22) 申请日 2018.12.25

G10L 25/03 (2013.01)

G10L 25/27 (2013.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 109841214 A

(43) 申请公布日 2019.06.04

(73) 专利权人 百度在线网络技术(北京)有限公司

地址 100085 北京市海淀区上地十街10号  
百度大厦三层

(72) 发明人 彭汉迎 欧阳能钧

(74) 专利代理机构 北京同立钧成知识产权代理有限公司 11205

代理人 荣甜甜 刘芳

(56) 对比文件

WO 2017138934 A1,2017.08.17

CN 107622767 A,2018.01.23

CN 103678456 A,2014.03.26

CN 102819009 A,2012.12.12

CN 107533842 A,2018.01.02

CN 108899044 A,2018.11.27

CN 106653041 A,2017.05.10

CN 107465986 A,2017.12.12

CN 107277699 A,2017.10.20

US 2016039356 A1,2016.02.11

审查员 何元

(51) Int. Cl.

G10L 15/22 (2006.01)

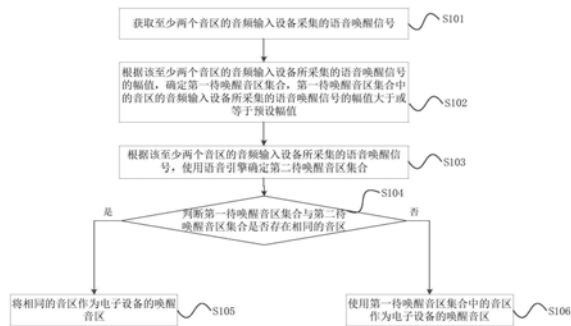
权利要求书2页 说明书11页 附图4页

(54) 发明名称

语音唤醒处理方法、装置和存储介质

(57) 摘要

本发明提供一种语音唤醒处理方法、装置和存储介质,电子设备在获取到至少两个音区的音频输入设备采集的语音唤醒信号之后,可以基于该至少两个音区的音频输入设备采集的语音唤醒信号的幅值所得到的待唤醒音区,纠正使用语音引擎所识别的待唤醒音区,避免采集到同一用户发出语音唤醒信号的多个音频输入设备所在的音区均被唤醒,可以提高电子设备所得到的语音唤醒结果的准确性。因此,本发明能够解决因车载终端的各音区的声音隔离度不足,导致车载终端的语音唤醒的准确性较低的技术问题。



1. 一种语音唤醒处理方法,其特征在于,所述方法应用于电子设备,所述电子设备包括设置在每个音区的音频输入设备,所述方法包括:

获取至少两个音区的音频输入设备采集的语音唤醒信号;

根据所述至少两个音区的音频输入设备所采集的语音唤醒信号的幅值,确定第一待唤醒音区集合,所述第一待唤醒音区集合中的音区的音频输入设备所采集的语音唤醒信号的幅值大于或等于预设幅值;

根据所述至少两个音区的音频输入设备所采集的语音唤醒信号,使用语音引擎确定第二待唤醒音区集合;

若所述第一待唤醒音区集合与所述第二待唤醒音区集合存在相同的音区,则将所述相同的音区作为所述电子设备的唤醒音区。

2. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

若所述第一待唤醒音区集合与所述第二待唤醒音区集合不存在相同的音区,则将所述第一待唤醒音区集合中的音区作为所述电子设备的唤醒音区。

3. 根据权利要求1所述的方法,其特征在于,所述根据所述至少两个音区的音频输入设备所采集的语音唤醒信号的幅值,确定第一待唤醒音区集合,包括:

判断所述至少两个音区的音频输入设备所采集的语音唤醒信号的幅值中的最大幅值是否大于或等于预设幅值;

若是,则将最大幅值对应的音频输入设备所在的第一音区的标识添加至所述第一待唤醒音区集合;

根据所述第一音区与每个第二音区的声音隔离度,更新每个所述第二音区的音频输入设备所采集的语音唤醒信号的幅值,所述第二音区为所述至少两个音区中除所述第一音区之外的音区;

判断更新后的第二音区的音频输入设备所采集的语音唤醒信号的幅值中是否存在大于或等于所述预设幅值的幅值;

若是,则将大于或等于所述预设幅值的幅值对应的音频输入设备所在的第二音区的标识添加至所述第一待唤醒音区集合。

4. 根据权利要求1所述的方法,其特征在于,所述获取至少两个音区的音频输入设备采集的语音唤醒信号,包括:

获取所述至少两个音区的音频输入设备采集的语音信号;

根据预设唤醒词,确定所述至少两个音区的音频输入设备所采集的语音信号为语音唤醒信号。

5. 根据权利要求1所述的方法,其特征在于,所述将所述相同的音区作为所述电子设备的唤醒音区之后,所述方法还包括:

对所述电子设备的唤醒音区的音频输入设备后续采集的语音信号,进行语义识别和处理。

6. 根据权利要求1-5任一项所述的方法,其特征在于,所述电子设备为车载终端。

7. 根据权利要求6所述的方法,其特征在于,一个音区覆盖车辆的一个座位。

8. 一种语音唤醒处理装置,其特征在于,所述装置应用于电子设备,所述电子设备包括设置在每个音区的音频输入设备,所述装置包括:

获取模块,用于获取至少两个音区的音频输入设备采集的语音唤醒信号;

第一处理模块,用于根据所述至少两个音区的音频输入设备所采集的语音唤醒信号的幅值,确定第一待唤醒音区集合,所述第一待唤醒音区集合中的音区的音频输入设备所采集的语音唤醒信号的幅值大于或等于预设幅值;

第二处理模块,用于根据所述至少两个音区的音频输入设备所采集的语音唤醒信号,使用语音引擎确定第二待唤醒音区集合;

唤醒模块,用于在所述第一待唤醒音区集合与所述第二待唤醒音区集合存在相同的音区时,将所述相同的音区作为所述电子设备的唤醒音区。

9. 一种语音唤醒处理装置,其特征在于,包括:至少一个处理器和存储器;

所述存储器存储计算机执行指令;

所述至少一个处理器执行所述存储器存储的计算机执行指令,使得所述语音唤醒处理装置执行权利要求1-7任一项所述的方法。

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上存储有计算机执行指令,当所述计算机执行指令被处理器执行时,实现权利要求1-7任一项所述的方法。

## 语音唤醒处理方法、装置和存储介质

### 技术领域

[0001] 本发明涉及智能设备技术领域,尤其涉及一种语音唤醒处理方法、装置和存储介质。

### 背景技术

[0002] 车载终端可以通过语音与用户进行交互,为用户提供导航、播放音乐、拨打电话等服务。现有的车载终端的麦克风是对准主驾驶位来设计的,可以为车辆的驾驶者提供语音服务。这种设计方式使车载终端无法为车辆上的其他成员提供服务。

[0003] 目前,推出了一种可以满足车辆全员使用的车载四音区方案,该方案中,将车辆内部的空间分为前后左右四个音区,每个音区覆盖车辆上的一个位置,车载终端在每个音区设置有独立的麦克风阵列,可以与该音区对应位置的用户进行语音交互。具体地,设置在每个音区的麦克风阵列可以对该音区对应位置的用户发出的语音信号进行采集。车载终端的语音引擎可以根据每个音区的麦克风阵列所采集的语音信号,识别哪个音区被用户唤醒,并对唤醒音区的麦克风阵列后续所采集的语音信号进行处理和响应。

[0004] 然而,由于车辆内部的空间较小,虽然通过麦克风阵列模组算法可以对四个位置制造一定的声音隔离度,但是任一位置的用户在语音唤醒车载终端时,还是会有一部分语音信号被其它音区的麦克风阵列采集,影响车载终端的语音唤醒的准确性。

### 发明内容

[0005] 本发明提供一种语音唤醒处理方法、装置和存储介质,用于解决因车载终端的各音区的声音隔离度不足,导致车载终端的语音唤醒的准确性较低的技术问题。

[0006] 本发明的第一方面提供一种语音唤醒处理方法,所述方法应用于电子设备,所述电子设备包括设置在每个音区的音频输入设备,所述方法包括:

[0007] 获取至少两个音区的音频输入设备采集的语音唤醒信号;

[0008] 根据所述至少两个音区的音频输入设备所采集的语音唤醒信号的幅值,确定第一待唤醒音区集合,所述第一待唤醒音区集合中的音区的音频输入设备所采集的语音唤醒信号的幅值大于或等于预设幅值;

[0009] 根据所述至少两个音区的音频输入设备所采集的语音唤醒信号,使用语音引擎确定第二待唤醒音区集合;

[0010] 若所述第一待唤醒音区集合与所述第二待唤醒音区集合存在相同的音区,则将所述相同的音区作为所述电子设备的唤醒音区。

[0011] 可选的,所述方法还包括:

[0012] 若所述第一待唤醒音区集合与所述第二待唤醒音区集合不存在相同的音区,则将所述第一待唤醒音区集合中的音区作为所述电子设备的唤醒音区。

[0013] 可选的,所述根据所述至少两个音区的音频输入设备所采集的语音唤醒信号的幅值,确定第一待唤醒音区集合,包括:

[0014] 判断所述至少两个音区的音频输入设备所采集的语音唤醒信号的幅值中的最大幅值是否大于或等于预设幅值；

[0015] 若是,则将最大幅值对应的音频输入设备所在的第一音区的标识添加至所述第一待唤醒音区集合；

[0016] 根据所述第一音区与每个第二音区的声音隔离度,更新每个所述第二音区的音频输入设备所采集的语音唤醒信号的幅值,所述第二音区为所述至少两个音区中除所述第一音区之外的音区；

[0017] 判断更新后的第二音区的音频输入设备所采集的语音唤醒信号的幅值中是否存在大于或等于所述预设幅值的幅值；

[0018] 若是,则将大于或等于所述预设幅值的幅值对应的音频输入设备所在的第二音区的标识添加至所述第一待唤醒音区集合。

[0019] 可选的,所述获取至少两个音区的音频输入设备采集的语音唤醒信号,包括：

[0020] 获取所述至少两个音区的音频输入设备采集的语音信号；

[0021] 根据预设唤醒词,确定所述至少两个音区的音频输入设备所采集的语音信号为语音唤醒信号。

[0022] 可选的,所述将所述相同的音区作为所述电子设备的唤醒音区之后,所述方法还包括：

[0023] 对所述电子设备的唤醒音区的音频输入设备后续采集的语音信号,进行语义识别和处理。

[0024] 可选的,所述电子设备为车载终端。

[0025] 可选的,一个音区覆盖所述车辆的一个座位。

[0026] 本发明的第二方面提供一种语音唤醒处理装置,所述装置应用于电子设备,所述电子设备包括设置在每个音区的音频输入设备,所述装置包括：

[0027] 获取模块,用于获取至少两个音区的音频输入设备采集的语音唤醒信号；

[0028] 第一处理模块,用于根据所述至少两个音区的音频输入设备所采集的语音唤醒信号的幅值,确定第一待唤醒音区集合,所述第一待唤醒音区集合中的音区的音频输入设备所采集的语音唤醒信号的幅值大于或等于预设幅值；

[0029] 第二处理模块,用于根据所述至少两个音区的音频输入设备所采集的语音唤醒信号,使用语音引擎确定第二待唤醒音区集合；

[0030] 唤醒模块,用于在所述第一待唤醒音区集合与所述第二待唤醒音区集合存在相同的音区时,将所述相同的音区作为所述电子设备的唤醒音区。

[0031] 可选的,所述唤醒模块,还用于在所述第一待唤醒音区集合与所述第二待唤醒音区集合不存在相同的音区时,将所述第一待唤醒音区集合中的音区作为所述电子设备的唤醒音区。

[0032] 可选的,所述第一处理模块,包括：

[0033] 第一判断单元,用于判断所述至少两个音区的音频输入设备所采集的语音唤醒信号的幅值中的最大幅值是否大于或等于预设幅值；

[0034] 第一添加单元,用于在所述至少两个音区的音频输入设备所采集的语音唤醒信号的幅值中的最大幅值大于或等于预设幅值时,将最大幅值对应的音频输入设备所在的第一

音区的标识添加至所述第一待唤醒音区集合；

[0035] 更新单元,用于根据所述第一音区与每个第二音区的声音隔离度,更新每个所述第二音区的音频输入设备所采集的语音唤醒信号的幅值,所述第二音区为所述至少两个音区中除所述第一音区之外的音区；

[0036] 第二判断单元,用于判断更新后的第二音区的音频输入设备所采集的语音唤醒信号的幅值中是否存在大于或等于所述预设幅值的幅值；

[0037] 第二添加单元,用于在更新后的第二音区的音频输入设备所采集的语音唤醒信号的幅值中存在大于或等于所述预设幅值的幅值时,将大于或等于所述预设幅值的幅值对应的音频输入设备所在的第二音区的标识添加至所述第一待唤醒音区集合。

[0038] 可选的,所述获取模块,具体用于获取所述至少两个音区的音频输入设备采集的语音信号,并根据预设唤醒词,确定所述至少两个音区的音频输入设备所采集的语音信号为语音唤醒信号。

[0039] 可选的,所述装置还包括：

[0040] 第三处理模块,用于在所述唤醒模块将所述相同的音区作为所述电子设备的唤醒音区之后,对所述电子设备的唤醒音区的音频输入设备后续采集的语音信号,进行语义识别和处理。

[0041] 可选的,所述电子设备为车载终端。

[0042] 可选的,一个音区覆盖所述车辆的一个座位。

[0043] 本发明的第三方面提供一种语音唤醒处理装置,包括:至少一个处理器和存储器；

[0044] 所述存储器存储计算机执行指令；

[0045] 所述至少一个处理器执行所述存储器存储的计算机执行指令,使得所述语音唤醒处理装置执行上述语音唤醒处理方法。

[0046] 本发明的第四方面提供一种计算机可读存储介质,所述计算机可读存储介质上存储有计算机执行指令,当所述计算机执行指令被处理器执行时,实现上述语音唤醒处理方法。

[0047] 本发明提供的语音唤醒处理方法、装置和存储介质,对于处于音区之间声音隔离度不足的电子设备来说,电子设备在获取到至少两个音区的音频输入设备采集的语音唤醒信号之后,可以基于该至少两个音区的音频输入设备采集的语音唤醒信号的幅值所得到的待唤醒音区,纠正使用语音引擎所识别的待唤醒音区,避免采集到同一用户发出语音唤醒信号的多个音频输入设备所在的音区均被唤醒,可以提高电子设备所得到的语音唤醒结果的准确性。

## 附图说明

[0048] 图1为本发明提供的车载终端的应用场景示意图；

[0049] 图2为本发明提供的一种语音唤醒处理方法的流程示意图；

[0050] 图3为本发明提供的另一种语音唤醒处理方法的流程示意图；

[0051] 图4为本发明提供的一种语音唤醒处理装置的结构示意图；

[0052] 图5为本发明提供的另一种语音唤醒处理装置的结构示意图；

[0053] 图6为本发明提供的又一种语音唤醒处理装置的结构示意图；

[0054] 图7为本发明提供的又一种语音唤醒处理装置的结构示意图。

### 具体实施方式

[0055] 为使本发明的目的、技术方案和优点更加清楚，下面将结合本发明的实施例，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

[0056] 图1为本发明提供的车载终端的应用场景示意图。如图1所示，为了满足车辆全员可以使用车载终端，目前推出了一种车载四音区方案。该方案中，将车辆内部的空间分为四个音区，每个音区覆盖车辆上的一个位置，每个音区设置有独立的麦克风阵列，可以采集该音区对应位置的用户的语音信号。该麦克风阵列可以设置在任一便于采集用户的语音信号的位置，例如该音区对应位置的上方车顶处。

[0057] 例如，音区1覆盖车辆主驾驶位置，音区1的麦克风阵列用于采集车辆主驾驶位置的用户发出的语音信号；音区2覆盖车辆副驾驶位置，音区2的麦克风阵列用于采集车辆副驾驶位置的用户发出的语音信号；音区3包括车辆后排左边的位置，音区3的麦克风阵列用于采集车辆后排左边的位置的用户发出的语音信号；音区4包括车辆后排右边的位置，音区4的麦克风阵列用于采集车辆后排右边的位置的用户发出的语音信号。

[0058] 可以理解，为了便于示例，图1仅示出了车载终端的麦克风阵列的位置，本发明对车载终端的构成，以及，车载终端在车辆中的位置，以及，麦克风阵列的位置并不限定。

[0059] 相应地，车载终端的语音引擎可以根据每个音区的麦克风阵列所采集的语音信号，识别哪个音区被用户唤醒，并对唤醒音区的麦克风阵列后续所采集的语音信号进行处理和响应。通过这种方式，可以使车载终端也可以为车辆上的其他位置的成员提供语音服务，不再仅仅局限于为车辆的驾驶者提供语音服务。

[0060] 然而，由于车辆内部的空间较小，虽然通过麦克风阵列模组算法可以对四个位置制造一定的声音隔离度，但是任一位置的用户在语音唤醒车载终端时，还是会有一部分语音信号被其它音区的麦克风阵列采集，影响车载终端的语音唤醒的准确性。例如，音区1的用户在语音输入语音唤醒信号时，可能会使音区1的麦克风阵列、音区2的麦克风阵列、音区3的麦克风阵列都采集到该用户语音输入的语音唤醒信号，导致车载终端的语音引擎误识别为音区1、音区2、音区3均被唤醒，致使车载终端的音区1、音区2、音区3同时响应用户，用户体验较差。

[0061] 考虑到上述问题，本发明提供了一种语音唤醒处理方法，用以解决因车载终端的各音区的声音隔离度不足，导致车载终端的语音唤醒的准确性较低的技术问题。

[0062] 可以理解，本发明提供的语音唤醒处理方法可以适用于任一可以为多个音区的用户提供语音服务的电子设备。例如，车载终端、智能家居等。示例性的，当电子设备为车载终端时，上述所说的一个音区可以覆盖车辆上的一个位置。当电子设备为智能家居时，上述所说的一个音区可以对应智能家居所在空间的一个区域等。

[0063] 下面结合具体地实施例对本发明的技术方案进行详细说明。下面这几个具体的实施例可以相互结合，对于相同或相似的概念或过程可能在某些实施例不再赘述。

[0064] 图2为本发明提供的一种语音唤醒处理方法的流程示意图。本发明的执行主体可

以是语音唤醒装置,还可以是集成了语音唤醒装置的电子设备,下述以集成了语音唤醒装置的电子设备(简称:电子设备)为例,对本发明提供的语音唤醒处理方法进行说明和介绍。在本实施例中,电子设备包括设置在每个音区的音频输入设备。这里所说的音频输入设备可以为任一具有音频采集功能的设备,例如,麦克风阵列。

[0065] 如图2所示,该方法可以包括:

[0066] S101、获取至少两个音区的音频输入设备采集的语音唤醒信号。

[0067] 可以理解,该至少两个音区的音频输入设备采集的语音唤醒信号:可以是同一用户发出的语音唤醒信号,也可以是不同音区的用户发出的语音唤醒信号。以音区A的音频输入设备和音区B的音频输入设备为例,音区A的音频输入设备采集的语音唤醒信号和音区B的音频输入设备采集的语音唤醒信号可以是来自同一用户的语音唤醒信号,或者,音区A的音频输入设备采集的语音唤醒信号是来自音区A的用户发出的语音唤醒信号、音区B的音频输入设备采集的语音唤醒信号是来自音区B用户发出的语音唤醒信号。

[0068] 本实施例不限定上述电子设备获取至少两个音区的音频输入设备采集的语音唤醒信号的方式。例如,电子设备可以获取该至少两个音区的音频输入设备采集的语音信号,并根据预设唤醒词,确定该至少两个音区的音频输入设备所采集的语音信号是否为语音唤醒信号。例如,当该至少两个音区的音频输入设备所采集的语音信号携带有预设唤醒词时,确定该至少两个音区的音频输入设备所采集的语音信号为语音唤醒信号,电子设备可以进一步地执行后续操作。当该至少两个音区的音频输入设备所采集的语音信号不携带有预设唤醒词时,则电子设备可以结束此次流程。

[0069] S102、根据该至少两个音区的音频输入设备所采集的语音唤醒信号的幅值,确定第一待唤醒音区集合,第一待唤醒音区集合中的音区的音频输入设备所采集的语音唤醒信号的幅值大于或等于预设幅值。

[0070] 由于发出语音唤醒信号的用户距离音频输入设备越近,音频输入设备所采集到的语音唤醒信号的幅值越大。也就是说,幅值越大,说明用户距离音频输入设备越近,用户唤醒该音频输入设备所在音区的可能性越大。幅值越小,说明用户距离音频输入设备越远,用户唤醒该音频输入设备所在音区的可能性越小。

[0071] 因此,电子设备可以将每个音区的音频输入设备所采集的语音唤醒信号的幅值与预设幅值相比较,来甄别哪些音区被用户唤醒。其中,上述预设幅值的大小可以根据所有音区的音频输入设备所能采集到的语音唤醒信号的最大幅值确定。例如,该预设幅值可以等于所有音区的音频输入设备所能采集到的语音唤醒信号的最大幅值,或者,该预设幅值可以略小于所有音区的音频输入设备所能采集到的语音唤醒信号的最大幅值。示例性的,以电子设备为车载终端为例,该预设幅值可以为2000。

[0072] 若某一音区的音频输入设备所采集的语音唤醒信号的幅值大于或等于预设幅值,说明该音区的音频输入设备所采集的语音唤醒信号是该音区的用户发出的。即,该音区的用户在唤醒该音区。则电子设备可以将该音区作为待唤醒音区,并将该音区的标识添加至第一待唤醒音区集合。即第一待唤醒音区集合中所包括的音区为根据音区的音频输入设备所采集的语音唤醒信号的幅值确定的待唤醒音区。可以理解,本发明中所涉及的音区的标识可以为音区的索引编号,音区的位置信息等任一能够唯一标识该音区的标识,对此不限定。



[0073] S103、根据该至少两个音区的音频输入设备所采集的语音唤醒信号,使用语音引擎确定第二待唤醒音区集合。

[0074] 具体的,电子设备还可以使用语音引擎,对该至少两个音区的音频输入设备所采集的语音唤醒信号进行识别,看哪些音区被用户唤醒,并将这些音区的标识组成的集合作为第二待唤醒音区集合。即第二待唤醒音区集合中所包括的音区为通过语音引擎所确定的待唤醒音区。

[0075] 关于电子设备如何根据音区的音频输入设备采集的语音唤醒信号,使用语音引擎确定待唤醒音区,可以参见现有技术,对此不再赘述。

[0076] S104、判断第一待唤醒音区集合与第二待唤醒音区集合是否存在相同的音区。若是,则执行S105,若否,则执行S106。

[0077] S105、将相同的音区作为电子设备的唤醒音区。

[0078] S106、使用第一待唤醒音区集合中的音区作为电子设备的唤醒音区。

[0079] 电子设备在得到第一待唤醒音区集合和第二待唤醒音区集合之后,可以将第一待唤醒音区集合和第二待唤醒音区集合进行比较,判断两个待唤醒音区集合是否存在相同的音区(即是否存在交集)。若存在相同的音区,则确定该相同的音区被用户唤醒,将该相同的音区作为电子设备的唤醒音区。进一步地,若第一待唤醒音区集合与第二待唤醒音区集合不存在相同的音区,则直接使用第一待唤醒音区集合中的音区作为电子设备的唤醒音区。

[0080] 当电子设备处于多个音区之间的声音隔离度小于或等于预设隔离度的场景中时,可能会导致电子设备设置在该多个音区中的音频输入设备采集到来自相同声源的语音信号。也就是说,设置在多个音区的音频输入设备所采集的语音唤醒信号可能是来自同一用户的语音唤醒信号。这里所说的预设隔离度阈值可以根据各音区之间的声音隔离度确定。

[0081] 由于语音唤醒信号的幅值可以反映用户距离各音区的音频输入设备的远近,因此,电子设备基于语音唤醒信号的幅值可以更准确的判断出到底是哪个音区被唤醒。进而,使用基于语音唤醒信号的幅值所得到的待唤醒音区,来纠正电子设备基于语音引擎识别的待唤醒音区,可以提高电子设备所得到的语音唤醒结果的准确性。

[0082] 进一步地,上述电子设备在确定电子设备的唤醒音区之后,可以对电子设备的唤醒音区的音频输入设备后续采集的语音信号,进行语义识别和处理,以与该音区的用户进行语音交互,为该音区的用户提供语音服务。

[0083] 继续参照图1,以电子设备为车载终端为例,由于各音区之间的声音隔离度不足,导致音区1的用户在语音输入语音唤醒信号时,会使音区1的音频输入设备、音区2的音频输入设备、音区3的音频输入设备都采集到该用户语音输入的语音唤醒信号。

[0084] 则车载终端可以根据音区1的音频输入设备采集的语音唤醒信号的幅值、音区2的音频输入设备采集的语音唤醒信号的幅值、音区3的音频输入设备采集的语音唤醒信号的幅值,得到第一待唤醒音区集合。假定音区1的音频输入设备采集的语音唤醒信号的幅值大于或等于预设幅值,则第一待唤醒音区集合包括:音区1的标识。

[0085] 同时,车载终端可以根据音区1的音频输入设备采集的语音唤醒信号、音区2的音频输入设备采集的语音唤醒信号、音区3的音频输入设备采集的语音唤醒信号,使用语音引擎得到的第二待唤醒音区集合。假定第二待唤醒音区集合包括:音区1的标识、音区2的标识和音区3的标识。

[0086] 在该示例下,第一待唤醒音区集合和第二待唤醒音区集合存在相同的音区,且该相同的音区为音区1,则电子设备可以将该音区1作为电子设备的唤醒音区。相比电子设备将语音引擎所识别的待唤醒音区作为唤醒音区,本实施例的方法,可以提高车载终端所得到的语音唤醒结果的准确性,避免了车载终端因各音区的声音隔离度不足,导致车载终端的语音唤醒的准确性较低的技术问题。

[0087] 本发明提供的语音唤醒处理方法,对于处于音区之间声音隔离度不足的电子设备来说,电子设备在获取到至少两个音区的音频输入设备采集的语音唤醒信号之后,可以基于该至少两个音区的音频输入设备采集的语音唤醒信号的幅值所得到的待唤醒音区,纠正使用语音引擎所识别的待唤醒音区,避免采集到同一用户发出语音唤醒信号的多个音频输入设备所在的音区均被唤醒,可以提高电子设备所得到的语音唤醒结果的准确性。

[0088] 图3为本发明提供的另一种语音唤醒处理方法的流程示意图。如图3所示,在上述实施例的基础上,上述步骤S102可以包括如下步骤:

[0089] S201、判断至少两个音区的音频输入设备所采集的语音唤醒信号的幅值中的最大幅值是否大于或等于预设幅值。若是,则执行S202,若否,则结束流程。

[0090] 通过判断至少两个音区的音频输入设备所采集的语音唤醒信号的幅值中的最大幅值是否大于或等于预设幅值,来甄别哪个音区可能在被唤醒。

[0091] 继续参照图1,以电子设备为车载终端为例,假定音区1的用户在语音输入语音唤醒信号唤醒音区1,音区3的用户在语音输入语音唤醒信号唤醒音区3。由于各音区之间的声音隔离度不足,导致音区1的音频输入设备、音区2的音频输入设备、音区3的音频输入设备都采集到语音唤醒信号。

[0092] 假定音区1的音频输入设备采集的语音唤醒信号的幅值大于音区3的音频输入设备采集的语音唤醒信号的幅值、音区3的音频输入设备采集的语音唤醒信号的幅值大于音区2的音频输入设备采集的语音唤醒信号的幅值。则电子设备可以比较音区1的音频输入设备采集的语音唤醒信号的幅值是否大于或等于预设幅值。

[0093] S202、将最大幅值对应的音频输入设备所在的第一音区的标识添加至第一待唤醒音区集合。

[0094] 继续参照上述示例,若音区1的音频输入设备采集的语音唤醒信号的幅值大于或等于预设幅值,则车载终端可以将音区1的标识添加到第一待唤醒音区集合。此时,音区1即为第一音区。

[0095] S203、根据第一音区与每个第二音区的声音隔离度,更新每个第二音区的音频输入设备所采集的语音唤醒信号的幅值,第二音区为至少两个音区中除第一音区之外的音区。

[0096] 具体的,电子设备在确定第一音区可能在被用户唤醒之后,可以进一步地根据第一音区的音频输入设备采集的语音唤醒信号的幅值,以及,第一音区与第二音区的声音隔离度,计算出第一音区的用户在发出语音唤醒信号时,该用户的声音被第二音区的音频输入设备采集到的幅值大小 $Q$ 。然后,电子设备可以将第二音区的音频输入设备所采集的语音唤醒信号的幅值减去该 $Q$ ,以消除第一音区的用户的声音对第二音区的影响,得到该第二音区的音频输入设备所采集的语音唤醒信号的实际幅值。

[0097] 例如,电子设备可以根据下述公式(1),得到每个第二音区的音频输入设备所采集

的语音唤醒信号的幅值：

$$[0098] \quad W_j = W_j - W_i * (10^p / 20) \quad (1)$$

[0099] 其中,  $W_j$  表示第  $j$  个音区 (即第二音区) 的音频输入设备所采集的语音唤醒信号的幅值,  $W_i$  表示第  $i$  个音区 (即第一音区) 的音频输入设备所采集的语音唤醒信号的幅值,  $p$  表示第  $j$  个音区与第  $i$  个音区的声音隔离度。

[0100] 当电子设备处于多个音区之间的声音隔离度小于或等于预设隔离度的场景中时, 第一音区的音频输入设备和第二音区的音频输入设备所采集的语音唤醒信号可能是来自同一用户的语音唤醒信号, 或者说, 混合有同一用户的语音唤醒信号。因此, 通过本步骤所示的方式, 可以消除第一音区的用户的语音对第二音区的影响。

[0101] 继续参照上述示例, 车载终端设备可以根据音区1与音区2的声音隔离度, 消除音区1的用户的语音对音区2的影响, 得到音区2的音频输入设备所采集的语音唤醒信号的实际幅值。车载终端设备可以根据音区1与音区3的声音隔离度, 消除音区1的用户的语音对音区3的影响, 得到音区3的音频输入设备所采集的语音唤醒信号的实际幅值。

[0102] S204、判断更新后的第二音区的音频输入设备所采集的语音唤醒信号的幅值中是否存在大于或等于预设幅值的幅值, 若是, 则执行S205, 若否, 则结束。

[0103] 通过判断更新后的第二音区的音频输入设备所采集的语音唤醒信号的幅值是否大于或等于预设幅值, 来甄别是否还有哪个音区可能在被唤醒。

[0104] S205、将大于或等于预设幅值的幅值对应的音频输入设备所在的第二音区的标识添加至第一待唤醒音区集合。

[0105] 继续参照上述示例, 在本示例中, 音区3的音频输入设备所采集的语音唤醒信号的幅值大于或等于预设幅值, 则电子设备可以将音区3的标识也添加至第一待唤醒音区集合。即, 电子设备通过各音区的音频输入设备采集的语音唤醒信号的幅值, 得到的第一待唤醒音区集合包括音区1和音区3。

[0106] 本发明提供的语音唤醒处理方法, 电子设备在获取到至少两个音区的音频输入设备采集的语音唤醒信号之后, 可以基于该至少两个音区的音频输入设备采集的语音唤醒信号的幅值, 以及, 最大幅值对应的音频输入设备所在的音区与其他音区的隔离度, 可以准确的识别出哪些音区被用户唤醒, 提高了电子设备根据各音区的音频输入设备采集的语音唤醒信号的幅值, 识别音区被用户唤醒的准确度。

[0107] 可以理解, 虽然上述实施例均以电子设备为执行主体对本发明提供的语音唤醒处理方法进行了说明和介绍。但是, 本领域技术人员可以理解的是, 具体实现时, 上述电子设备也可以通过电子设备上所安装的语音交互系统或者应用程序等, 实现上述方法实施例, 对此不再赘述。

[0108] 图4为本发明提供的一种语音唤醒处理装置的结构示意图, 如图4所示, 该语音唤醒处理装置可以应用于前述所说的电子设备。例如, 该语音唤醒处理装置可以为应用于上述电子设备中设备的芯片。

[0109] 如图4所示, 该装置包括: 获取模块11、第一处理模块12、第二处理模块13和唤醒模块14。其中,

[0110] 获取模块11, 用于获取至少两个音区的音频输入设备采集的语音唤醒信号;

[0111] 第一处理模块12, 用于根据所述至少两个音区的音频输入设备所采集的语音唤醒

信号的幅值,确定第一待唤醒音区集合,所述第一待唤醒音区集合中的音区的音频输入设备所采集的语音唤醒信号的幅值大于或等于预设幅值;

[0112] 第二处理模块13,用于根据所述至少两个音区的音频输入设备所采集的语音唤醒信号,使用语音引擎确定第二待唤醒音区集合;

[0113] 唤醒模块14,用于在所述第一待唤醒音区集合与所述第二待唤醒音区集合存在相同的音区时,将所述相同的音区作为所述电子设备的唤醒音区。

[0114] 可选的,所述唤醒模块14,还用于在所述第一待唤醒音区集合与所述第二待唤醒音区集合不存在相同的音区时,将所述第一待唤醒音区集合中的音区作为所述电子设备的唤醒音区。

[0115] 可选的,所述获取模块11,具体用于获取所述至少两个音区的音频输入设备采集的语音信号,并根据预设唤醒词,确定所述至少两个音区的音频输入设备所采集的语音信号为语音唤醒信号。

[0116] 可选的,所述电子设备为车载终端。

[0117] 可选的,一个音区覆盖所述车辆的一个座位。

[0118] 图5为本发明提供的另一种语音唤醒处理装置的结构示意图,如图5所示,在上述图4所示的框图的基础上,上述第一处理模块12可以包括:

[0119] 第一判断单元121,用于判断所述至少两个音区的音频输入设备所采集的语音唤醒信号的幅值中的最大幅值是否大于或等于预设幅值;

[0120] 第一添加单元122,用于在所述至少两个音区的音频输入设备所采集的语音唤醒信号的幅值中的最大幅值大于或等于预设幅值时,将最大幅值对应的音频输入设备所在的第一音区的标识添加至所述第一待唤醒音区集合;

[0121] 更新单元123,用于根据所述第一音区与每个第二音区的声音隔离度,更新每个所述第二音区的音频输入设备所采集的语音唤醒信号的幅值,所述第二音区为所述至少两个音区中除所述第一音区之外的音区;

[0122] 第二判断单元124,用于判断更新后的第二音区的音频输入设备所采集的语音唤醒信号的幅值中是否存在大于或等于所述预设幅值的幅值;

[0123] 第二添加单元125,用于在更新后的第二音区的音频输入设备所采集的语音唤醒信号的幅值中存在大于或等于所述预设幅值的幅值时,将大于或等于所述预设幅值的幅值对应的音频输入设备所在的第二音区的标识添加至所述第一待唤醒音区集合。

[0124] 图6为本发明提供的又一种语音唤醒处理装置的结构示意图,如图6所示,在上述图4所示的框图的基础上,上述装置还可以包括:

[0125] 第三处理模块15,用于在所述唤醒模块将所述相同的音区作为所述电子设备的唤醒音区之后,对所述电子设备的唤醒音区的音频输入设备后续采集的语音信号,进行语义识别和处理。

[0126] 本发明提供的语音唤醒处理,可以执行上述方法实施例中电子设备的动作,其实现原理和技术效果类似,在此不再赘述。

[0127] 需要说明的是,应理解以上模块可以以软件通过处理元件调用的形式实现;也可以以硬件的形式实现。例如,获取模块可以为单独设立的处理元件,也可以集成在上述装置的某一个芯片中实现,此外,也可以以程序代码的形式存储于上述装置的存储器中,由上述

装置的某一个处理元件调用并执行以上处理单元的功能。此外这些单元全部或部分可以集成在一起,也可以独立实现。这里所述的处理元件可以是一种集成电路,具有信号的处理能力。在实现过程中,上述方法的各步骤或以上各个单元可以通过处理器元件中的硬件的集成逻辑电路或者软件形式的指令完成。

[0128] 例如,以上这些单元可以是配置成实施以上方法的一个或多个集成电路,例如:一个或多个专用集成电路(application specific integrated circuit,ASIC),或,一个或多个微处理器(digital signal processor,DSP),或,一个或者多个现场可编程门阵列(field programmable gate array,FPGA)等。再如,当以上某个单元通过处理元件调度程序代码的形式实现时,该处理元件可以是通用处理器,例如中央处理器(central processing unit,CPU)或其它可以调用程序代码的处理器。再如,这些单元可以集成在一起,以片上系统(system-on-a-chip,SOC)的形式实现。

[0129] 图7为本发明提供的又一种语音唤醒处理装置的结构示意图,该语音唤醒处理装置例如可以是电子设备,比如车载终端、智能家居等。如图7所示,该语音唤醒处理装置700包括:存储器701和至少一个处理器702。

[0130] 存储器701,用于存储程序指令。

[0131] 处理器702,用于在程序指令被执行时实现本实施例中的语音唤醒处理方法,具体实现原理可参见上述实施例,本实施例此处不再赘述。

[0132] 该语音唤醒处理装置700还可以包括及输入/输出接口703。

[0133] 输入/输出接口703可以包括独立的输出接口和输入接口,也可以为集成输入和输出的集成接口。其中,输出接口用于输出数据,输入接口用于获取输入的数据,上述输出的数据为上述方法实施例中输出的统称,输入的数据为上述方法实施例中输入的统称。

[0134] 本发明还提供一种可读存储介质,可读存储介质中存储有执行指令,当语音唤醒处理装置的至少一个处理器执行该执行指令时,当计算机执行指令被处理器执行时,实现上述实施例中的语音唤醒处理方法。

[0135] 本发明还提供一种程序产品,该程序产品包括执行指令,该执行指令存储在可读存储介质中。语音唤醒处理装置的至少一个处理器可以从可读存储介质读取该执行指令,至少一个处理器执行该执行指令使得语音唤醒处理装置实施上述的各种实施方式提供的语音唤醒处理方法。

[0136] 在本发明所提供的几个实施例中,应该理解到,所揭露的装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述单元或模块的划分,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。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0137] 所述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例方案的目的。

[0138] 另外,在本发明各个实施例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,也可以采用硬件加软件功能单元的形式实现。

[0139] 上述以软件功能单元的形式实现的集成的单元,可以存储在一个计算机可读取存储介质中。上述软件功能单元存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)或处理器(英文:processor)执行本发明各个实施例所述方法的部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(英文:Read-Only Memory,简称:ROM)、随机存取存储器(英文:Random Access Memory,简称:RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0140] 在上述实施例中,应理解,处理器可以是中央处理单元(英文:Central Processing Unit,简称:CPU),还可以是其他通用处理器、数字信号处理器(英文:Digital Signal Processor,简称:DSP)、专用集成电路(英文:Application Specific Integrated Circuit,简称:ASIC)等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本发明所公开的方法的步骤可以直接体现为硬件处理器执行完成,或者用处理器中的硬件及软件模块组合执行完成。

[0141] 最后应说明的是:以上各实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述各实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分或者全部技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

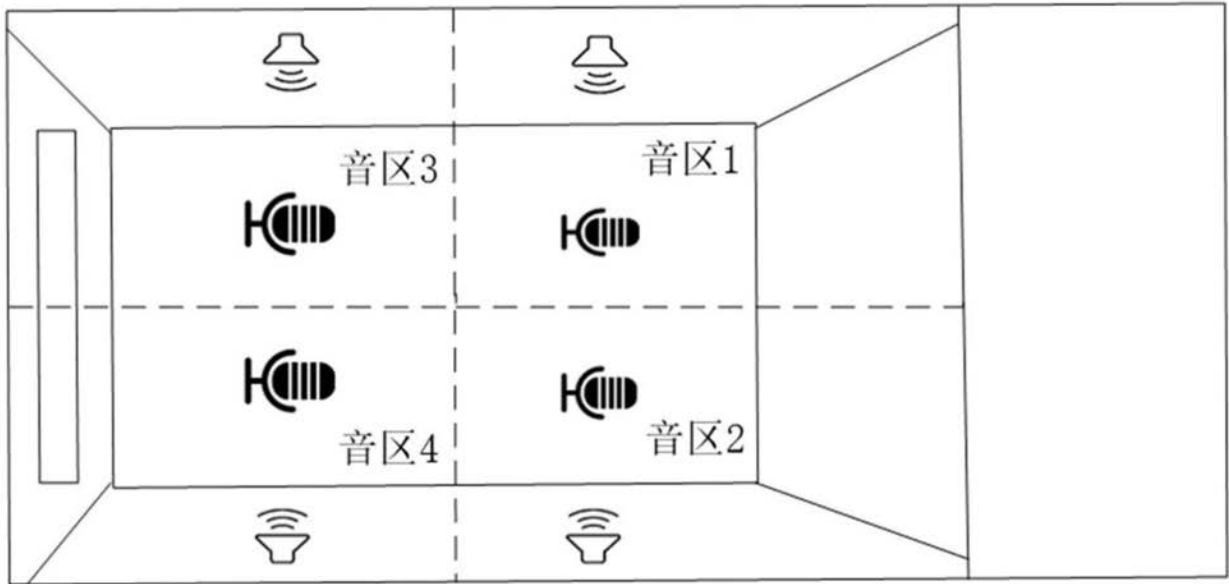


图1

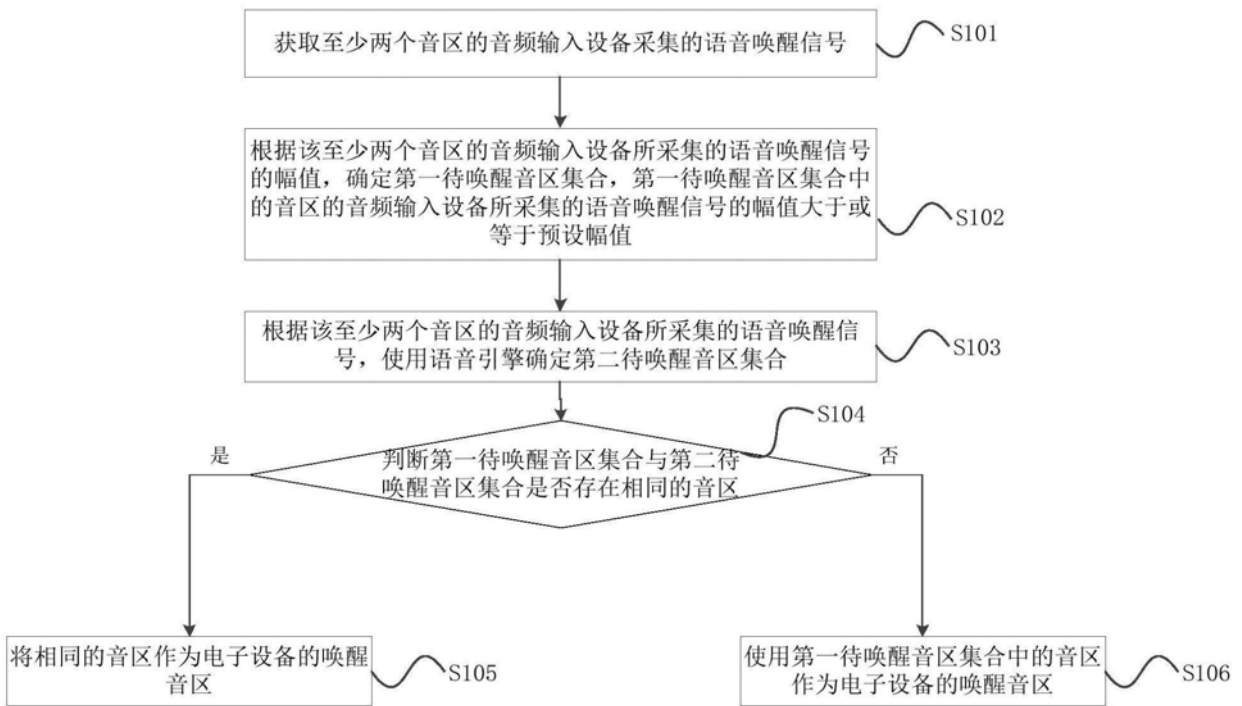


图2

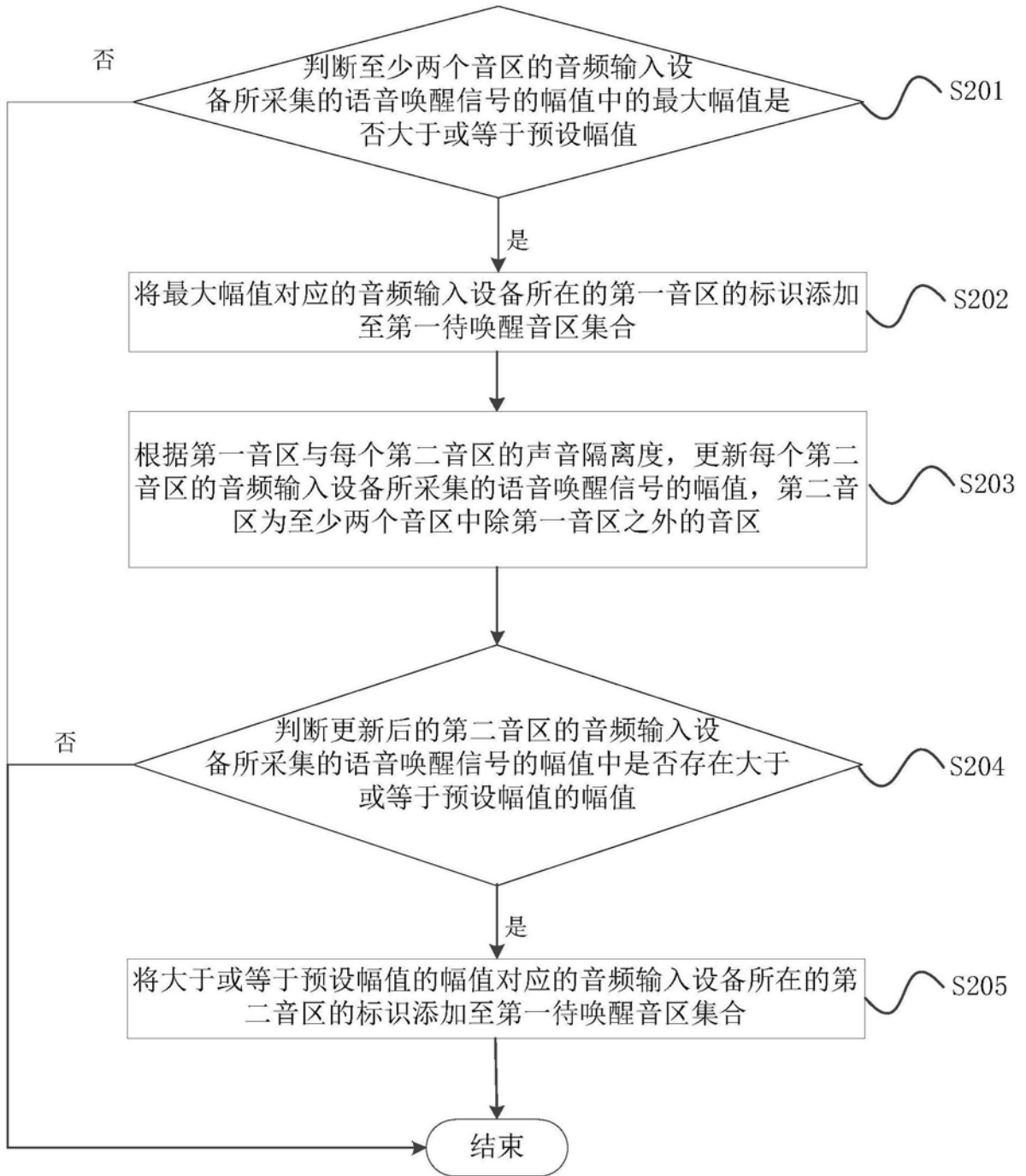


图3



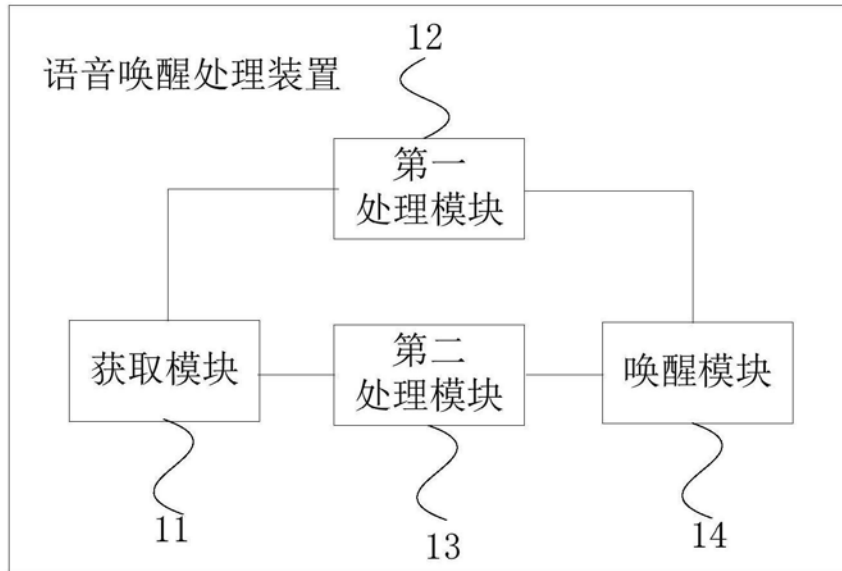


图4

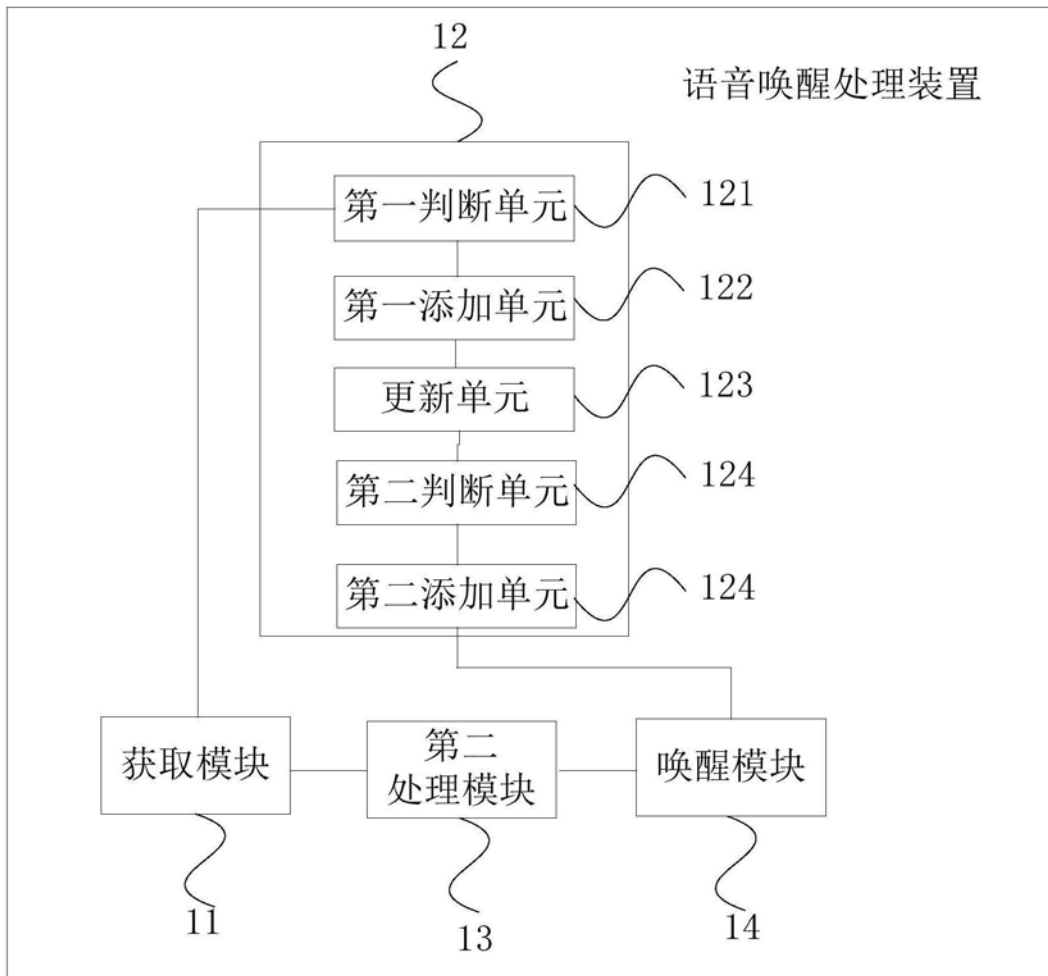


图5

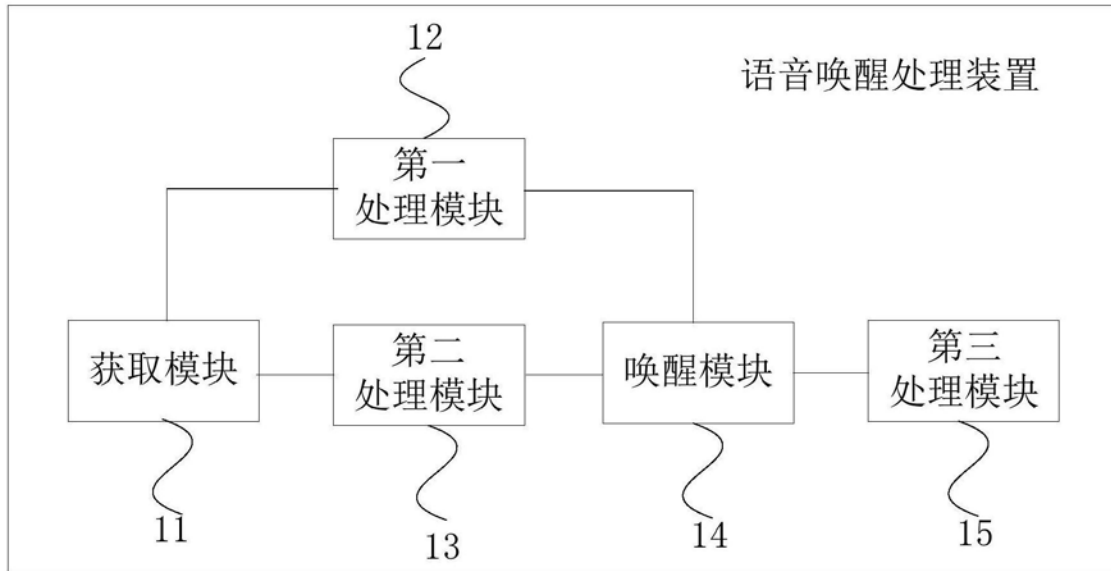


图6

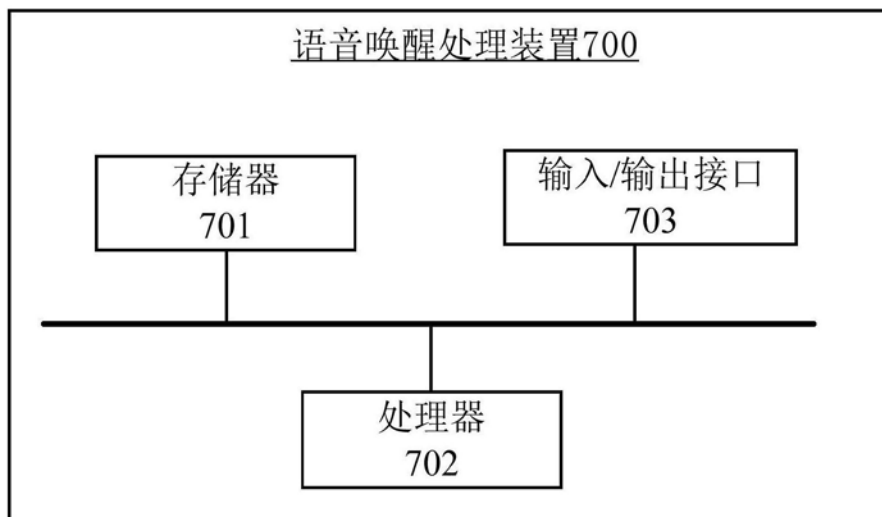


图7