



(12)发明专利申请

(10)申请公布号 CN 110010143 A

(43)申请公布日 2019.07.12

(21)申请号 201910319221.8

(22)申请日 2019.04.19

(71)申请人 出门问问信息科技有限公司  
地址 100086 北京市海淀区中关村大街19号办公A楼10层1001

(72)发明人 周舒然

(74)专利代理机构 北京知呱呱知识产权代理有限公司 11577  
代理人 贺亚明 丁彦峰

(51) Int. Cl.  
G10L 21/02(2013.01)  
G10L 21/0208(2013.01)  
G10L 21/0216(2013.01)

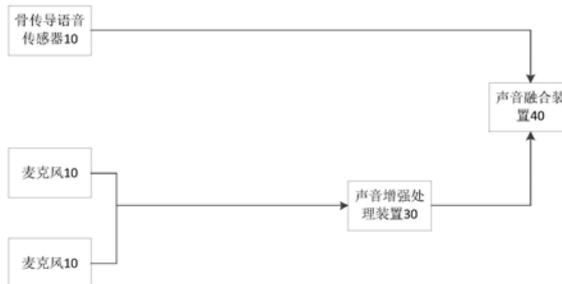
权利要求书2页 说明书6页 附图2页

(54)发明名称

一种语音信号增强系统、方法及存储介质

(57)摘要

本发明实施例涉及语音信号处理技术领域，具体公开了一种语音信号增强系统、方法及存储介质，该系统包括：骨传导语音传感器用于采集用户的语音信号，并将用户的语音信号传输至声音融合装置；麦克风用于，采集外界发出的声音信号，并将外界发出的声音信号传输至声音增强处理装置；声音增强处理装置，用于将麦克风采集的外界发出的声音信号进行降噪处理，并将经过降噪处理后的信号输入至声音融合装置；声音融合装置，用于将用户的语音信号和经过降噪处理后的信号进行融合处理，获取输出信号。通过上述方式，可以得到一个还原用户真实声音，且已经经过降噪处理的干净的信号，这个信号将作为语音识别设备的输入信号，可以供语音识别设备进行精准识别。



1. 一种语音信号增强系统,其特征在于,所述语音信号增强系统与语音识别设备建立电连接,或者包含于所述语音识别设备中,所述系统包括:骨传导语音传感器、麦克风、声音增强处理装置以及声音融合装置;

所述骨传导语音传感器用于采集用户的语音信号,并将所述用户的语音信号传输至所述声音融合装置;

所述麦克风用于,采集外界发出的声音信号,并将所述外界发出的声音信号传输至所述声音增强处理装置;

所述声音增强处理装置,用于将所述麦克风采集的外界发出的声音信号进行一次降噪处理,并将经过一次降噪处理后的信号输入至所述声音融合装置;

所述声音融合装置,用于将所述用户的语音信号和经过一次降噪处理后的信号进行融合处理,获取输出信号,用以作为所述语音识别设备的语音识别信号。

2. 根据权利要求1所述的系统,其特征在于,所述系统还包括回声消除装置,用于将所述麦克风采集的外界发出的声音信号进行回声消除处理,并将经过回声消除处理后的声音信号传输至所述声音增强处理装置。

3. 根据权利要求1所述的系统,其特征在于,所述麦克风数量为至少两个,所述声音增强处理装置具体用于,将至少两个麦克风分别采集的外界发出的声音信号进行波束成形处理后,再进行噪音抑制处理。

4. 根据权利要求1所述的系统,其特征在于,所述声音融合装置,具体用于:利用自适应滤波方法对所述用户的语音信号和经过一次降噪处理后的信号进行融合处理,获取输出信号,用以作为所述语音识别设备的语音识别信号。

5. 根据权利要求1-4任一项所述的系统,其特征在于,所述系统还包括降噪装置,用于对所述输出信号进行二次降噪处理,以便将经过所述二次降噪处理后的信号作为所述语音识别设备的语音识别信号。

6. 根据权利要求1-4任一项所述的系统,其特征在于,所述系统还包括:声音活动检测装置,用于当接收到声音传感器组中预设声音传感器传输的用户发出的语音信号时,唤醒所述声音传感器组中除所述预设声音传感器之外的声音传感器、声音增强处理装置以及声音融合装置从休眠状态进入工作状态,其中,所述声音传感器组由骨传导语音传感器以及麦克风构成。

7. 根据权利要求6所述的系统,其特征在于,所述预设声音传感器的数量为至少一个。

8. 一种语音信号增强方法,其特征在于,所述方法应用于一种语音信号增强系统,所述语音信号增强系统与语音识别设备建立电连接,或者包含于所述语音识别设备中,所述语音信号增强系统包括:骨传导语音传感器、麦克风、声音增强处理装置以及声音融合装置;所述方法包括:

所述骨传导传感器采集用户的语音信号后,传输至所述声音融合装置;

所述麦克风采集外界发出的声音信号,并传输至所述声音增强处理装置;

所述声音增强处理装置将麦克风采集的外界发出的声音信号进行一次降噪处理,并将经过一次降噪处理后的信号输入至所述声音融合装置;

所述声音融合装置将所述用户的语音信号,以及经过一次降噪处理后的信号进行声音融合处理,获取输出信号,用以作为所述语音识别设备的语音识别信号。

9. 根据权利要求8所述的方法,其特征在于,所述语音信号增强系统还包括:声音活动检测装置,所述骨传导语音传感器、麦克风构成声音传感器组;当所述声音传感器组中除预设声音传感器之外的声音传感器、声音增强处理装置以及声音融合装置进入工作状态之前,所述方法还包括:当所述声音活动检测装置接收到所述声音传感器组中预设声音传感器传输的用户发出的语音信号时,唤醒所述声音传感器组中除所述预设声音传感器之外的声音传感器、声音增强处理装置以及声音融合装置从休眠状态进入工作状态。

10. 一种计算机存储介质,其特征在于,所述计算机存储介质中包含一个或多个程序指令,所述一个或多个程序指令用于被如权利要求1-7任一项所述的语音信号增强系统执行如权利要求8或9任一项所述的语音信号增强方法步骤。

## 一种语音信号增强系统、方法及存储介质

### 技术领域

[0001] 本发明实施例涉及语音信号处理技术领域,具体涉及一种语音信号增强系统、方法及存储介质。

### 背景技术

[0002] 声音的特性是会叠加,在某些嘈杂的环境中,用户发出的语音信号会叠加外界环境中的干扰信号,例如噪音信号,或者其他来源的声音信号。而导致接收的语音信号的对象并不能清楚的识别,尤其对于需要识别语音信号的语音识别设备来说,在接收到掺杂各种噪音的语音信号时,将无法对语音信号进行有效识别,也就无法进行后续的操作。

[0003] 如此一来,必然给使用语音识别设备的用户造成困扰。那么,如何才能保证语音识别设备,即使在嘈杂的环境中,也能够从掺杂各种噪音的语音信号中,有效识别用户发出的语音信号,进而执行后续操作,成为本申请亟待解决的技术问题。

### 发明内容

[0004] 为此,本发明实施例提供一种语音信号增强系统、方法及存储介质,以解决现有技术中语音识别设备,在嘈杂的环境中无法对掺杂各种噪音的语音信号后进行有效识别的问题。

[0005] 为了实现上述目的,本发明实施例提供如下技术方案:

[0006] 第一方面,本发明实施例提供了一种语音信号增强系统,该语音信号增强系统与语音识别设备建立电连接,该系统包括:骨传导语音传感器、麦克风、声音增强处理装置以及声音融合装置;

[0007] 骨传导语音传感器用于采集用户的语音信号,并将用户的语音信号传输至声音融合装置;

[0008] 麦克风用于,采集外界发出的声音信号,并将外界发出的声音信号传输至声音增强处理装置;

[0009] 声音增强处理装置,用于将麦克风采集的外界发出的声音信号进行一次降噪处理,并将经过一次降噪处理后的信号输入至声音融合装置;

[0010] 声音融合装置,用于将用户的语音信号和一次降噪处理后的信号进行融合处理,获取输出信号,用以作为所述语音识别设备的语音识别信号。

[0011] 本发明实施例的特征还在于,系统还包括回声消除装置,用于将麦克风采集的外界发出的声音信号进行回声消除处理,并将经过回声消除处理后的声音信号传输至声音增强处理装置。

[0012] 本发明实施例的特征还在于,麦克风的数量为至少两个,声音增强处理装置具体用于,将至少两个麦克风采集的外界发出的声音信号进行波束成形处理后,再进行噪音抑制处理。

[0013] 本发明实施例的特征还在于,声音融合装置,具体用于:利用自适应滤波方法对用

户的语音信号和一次降噪处理后的信号进行融合处理,获取输出信号,用以作为所述语音识别设备的语音识别信号。

[0014] 本发明实施例的特征还在于,该系统还包括降噪装置,用于对输出信号进行二次降噪处理,以便将经过二次降噪处理后的信号作为语音识别设备的语音识别信号。

[0015] 本发明实施例的特征还在于,系统还包括:声音活动检测装置,用于当接收到声音传感器组中预设声音传感器传输的用户发出的语音信号时,唤醒声音传感器组中除预设声音传感器之外的声音传感器、声音增强处理装置以及声音融合装置从休眠状态进入工作状态,其中,声音传感器组由骨传导语音传感器以及麦克风构成。

[0016] 本发明实施例的特征还在于,预设声音传感器的数量为至少一个。

[0017] 第二方面,本发明实施例还提供了一种语音信号增强方法,该方法应用于一种语音信号增强系统,语音信号增强系统与语音识别设备建立电连接,语音信号增强系统包括:骨传导语音传感器、麦克风、声音增强处理装置以及声音融合装置;该方法包括:

[0018] 骨传导传感器采集用户的语音信号后,传输至声音融合装置;

[0019] 麦克风采集外界发出的声音信号,并传输至声音增强处理装置;

[0020] 声音增强处理装置将麦克风采集的外界发出的声音信号进行一次降噪处理,并将经过一次降噪处理后的信号输入至声音融合装置;

[0021] 声音融合装置将用户的语音信号,以及经过一次降噪处理后的信号进行声音融合处理,获取输出信号,用以作为所述语音识别设备的语音识别信号。

[0022] 本发明实施例的特征还在于,语音信号增强系统还包括回声消除装置,麦克风采集外界发出的声音信号,并传输至声音增强处理装置之前,方法还包括:

[0023] 利用回声消除装置对每一个麦克风分别采集的外界声音信号进行回声消除处理。

[0024] 本发明实施例的特征还在于,语音信号增强系统中的麦克风数量为至少两个,声音增强处理装置将麦克风采集的外界发出的声音信号进行一次降噪处理,具体包括:

[0025] 将至少两个麦克风采集的外界发出的声音信号进行波束成形处理后,再噪音抑制处理。

[0026] 本发明实施例的特征还在于,声音融合装置将用户的语音信号,以及经过一次降噪处理后的信号进行声音融合处理,获取输出信号,用以作为所述语音识别设备的语音识别信号,具体包括:

[0027] 声音融合装置利用自适应滤波方法对用户的语音信号和一次降噪处理后的信号进行融合处理,获取输出信号,用以作为所述语音识别设备的语音识别信号。

[0028] 本发明实施例的特征还在于,声音融合装置将用户的语音信号,以及经过一次降噪处理后的信号进行声音融合处理,获取输出信号之后,该方法还包括:

[0029] 利用降噪装置对输出信号进行二次降噪处理,以便将经过二次降噪处理后的信号作为语音识别设备的语音识别信号。

[0030] 本发明实施例的特征还在于,语音信号增强系统还包括:声音活动检测装置,骨传导语音传感器、麦克风构成声音传感器组;当声音传感器组中除预设声音传感器之外的声音传感器、声音增强处理装置以及声音融合装置进入工作状态之前,方法还包括:当声音活动检测装置接收到声音传感器组中预设声音传感器传输的用户发出的语音信号时,唤醒声音传感器组中除预设声音传感器之外的声音传感器、声音增强处理装置以及声音融合装置

从休眠状态进入工作状态。

[0031] 本发明实施例的特征还在于,预设声音传感器的数量为至少一个。

[0032] 第三方面,本发明实施例还提供了一种计算机存储介质,计算机存储介质中包含一个或多个程序指令,一个或多个程序指令用于被如上第一方面所述的一种语音信号增强系统执行如上第二方面所述的一种语音信号增强方法中的任一方法步骤。

[0033] 根据本发明的实施方式,具有如下优点:骨传导语音传感器主要采集用户的语音信号中的低频,而麦克风则采集外界发出的声音信号,其采集的声音信号中实际是包含了用户的语音信号和外界噪音。声音增强处理装置将外界发出的声音信号进行降噪处理后,可以滤除掉噪音信号。声音融合装置将骨传导语音传感器采集用户的语音信号中低频信号,和经过降噪处理后的信号进行融合,从而得到一个还原用户真实声音,且已经经过降噪处理的干净的信号,这个信号将作为语音识别设备的输入信号,供语音识别设备进行精准识别,并执行后续的操作。经过上述处理后的信号更加容易被语音识别设备识别,进而可以提升用户对语音识别设备的体验度。

## 附图说明

[0034] 为了更清楚地说明本发明的实施方式或现有技术中的技术方案,下面将对实施方式或现有技术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示例性的,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据提供的附图引伸获得其它的实施附图。

[0035] 图1为本发明一实施例提供的一种语音信号增强系统结构示意图;

[0036] 图2为本发明另一实施例提供的一种语音信号增强系统结构示意图;

[0037] 图3为本发明另一实施例提供的一种语音信号增强系统结构示意图;

[0038] 图4为本发明另一实施例提供的一种语音信号增强方法流程示意图。

## 具体实施方式

[0039] 以下由特定的具体实施例说明本发明的实施方式,熟悉此技术的人士可由本说明书所揭露的内容轻易地了解本发明的其他优点及功效,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。

[0040] 本发明实施例提供了一种语音信号增强系统,该语音信号增强系统与语音识别设备建立电连接,或者包含于语音识别设备中,具体如图1所示,该系统包括:骨传导语音传感器10、麦克风20、声音增强处理装置30以及声音融合装置40。

[0041] 骨传导语音传感器10用于采集用户的语音信号,并将用户的语音信号传输至声音融合装置40;

[0042] 麦克风20用于,采集外界发出的声音信号,并将外界发出的声音信号传输至声音增强处理装置30;

[0043] 声音增强处理装置30,用于将麦克风20采集的外界发出的声音信号进行一次降噪处理,并将经过一次降噪处理后的信号输入至声音融合装置40;

[0044] 声音融合装置40,用于将用户的语音信号和经过一次降噪处理后的信号进行融合

处理,获取输出信号,用以作为语音识别设备的语音识别信号。

[0045] 具体的,骨传导语音传感器10可以是骨传导麦克风20或者加速度传感器,其紧密贴合使用者的耳骨,用以充分采集用户在说话时声音延头骨的传播。声音的振动可以带动头骨和肌肉组织振动,这种振动频率刚好被骨传导语音传感器10采集。骨传导语音传感器10的有点在于,采样频率和区间相比普通麦克风20低很多,能够有效的采集低频信号,隔离高频噪音,基本覆盖大部分人声的共振峰,可以包含人声的有效信息。

[0046] 而麦克风20则用于分别采集外界发出的声音信号,外界发出的声音信号中包含了用户的语音信号,以及外界环境中的噪音。采用麦克风20的原因在于,其可以采集外侧环境造影和用户说话的声音。其涵盖了说话人声音中所有频带信号。

[0047] 而正是由于麦克风20采集的声音信号中包含外界环境中的噪音信号,因此需要通过声音增强处理装置30进行一次降噪处理。然后,将经过一次降噪处理后的信号输入至声音融合装置40中。

[0048] 可选的,在一个具体的实施例中,如果麦克风20的数量为至少两个时,声音增强处理装置30主要是用于至少两个麦克风20分别采集的外界发出的声音信号进行波束成形处理。通过波束成形处理,实现对麦克风20采集的声音信号进行空间滤波。波束成形的目的是为了进行有指向性的滤波和去除干扰。然后,再将经过空间滤波处理后的信号进行噪音抑制处理。对于波束成形处理和噪音抑制处理实际都属于现有技术,这里不做过多介绍。

[0049] 进一步可选的,如图2或者如图3所示,为了防止用户发出语音信号时,其正在听音乐或者看电影等流媒体信息时,麦克风20将会采集到一部分流媒体信号,这部分流媒体信号将成为干扰信号,进一步对最终要输出的信号进行干扰。因此,语音信号增强系统还可以包括回声消除装置50,用于结合流媒体信号的原始参考信号,对麦克风20采集到的流媒体信号进行回声消除。具体回声消除的工作原理为现有技术,这里不做过多赘述。而且,每一个回声消除装置50分别对应一个麦克风20。

[0050] 声音融合装置40,其主要是对用户的语音信号和经过一次降噪处理后的信号进行融合处理。其处理方式可以采用自适应滤波方法进行融合处理。具体的处理过程就是提取骨传导语音传感器10采集的用户的语音信号中的有效低频信号,然后将该有效低频信号替代经过声音增强装置一次降噪处理后的信号中的低频信号。同时,在替代过程中实现对所有信号的进一步滤波处理,从而使得获取的输出信号是干净的,更加贴近人声音的信号,用以作为语音识别设备的语音识别信号。

[0051] 可选的,为了使得语音识别设备的语音信号更加清楚自然,该系统还可以包括:降噪装置70,具体如图2或图3所示,降噪装置70用于对输出信号进行二次降噪处理,其主要目的是为了进一步消除环境噪音和非平稳噪音。例如嘈杂的街边噪音、引擎噪音或者风噪等。并将经过二次降噪处理后的信号最为最终输入到语音识别设备的语音识别信号。

[0052] 进一步可选的,为了降低语音增强系统所消耗的功耗,该语音增强系统中还可以包括声音活动检测装置60,具体如图2或者如图3所示。并且,仅设置声音活动检测装置60和声音传感器组中预设声音传感器处于工作状态,而其他元器件处于休眠状态。只有当声音活动检测装置60接收到声音传感器组中预设声音传感器传输的用户发出的语音信号时,才会唤醒语音增强系统中的其他部件从休眠状态进入工作状态。而声音传感器组则是由骨传导语音传感器10以及麦克风20构成。

[0053] 可选的,预设声音传感器的个数可以为至少一个。

[0054] 图2中示出了语音信号增强系统中预设声音传感器为骨传导语音传感器10。而图3中示出了语音信号增强系统中预设传感器为两个麦克风20。

[0055] 本发明实施例提供一种语音信号增强系统,骨传导语音传感器主要采集用户的语音信号中的低频,而麦克风则采集外界发出的声音信号,其采集的声音信号中实际是包含了用户的语音信号和外界噪音。声音增强处理装置将外界发出的声音信号进行降噪处理后,可以滤除掉噪音信号。声音融合装置将骨传导语音传感器采集用户的语音信号中低频信号,和经过降噪处理后的信号进行融合,从而得到一个还原用户真实声音,且已经经过降噪处理的干净的信号,这个信号将作为语音识别设备的输入信号,供语音识别设备进行精准识别,并执行后续的操作。经过上述处理后的信号更加容易被语音识别设备识别,进而可以提升用户对语音识别设备的体验度。

[0056] 如上述实施例1相对应的,本发明实施例2提供了一种语音信号增强方法,该方法应用于一种语音信号增强系统,语音信号增强系统与语音识别设备建立电连接,或者包含于语音识别设备中,语音信号增强系统包括:骨传导语音传感器、麦克风、声音增强处理装置以及声音融合装置;具体如图4所示,方法包括:

[0057] 步骤410,骨传导传感器采集用户的语音信号后,传输至声音融合装置。

[0058] 步骤420,麦克风采集外界发出的声音信号,并传输至声音增强处理装置。

[0059] 步骤430,声音增强处理装置将麦克风采集的外界发出的声音信号进行一次降噪处理,并将经过一次降噪处理后的信号输入至声音融合装置。

[0060] 步骤440,声音融合装置将用户的语音信号,以及经过一次降噪处理后的信号进行声音融合处理,获取输出信号,用以作为语音识别设备的语音识别信号。

[0061] 可选的,语音信号增强系统还包括回声消除装置,麦克风采集外界发出的声音信号,并传输至声音增强处理装置之前,方法还包括:

[0062] 利用回声消除装置对每一个麦克风分别采集的外界声音信号进行回声消除处理。

[0063] 可选的,语音信号增强系统中的麦克风数量为至少两个,声音增强处理装置将麦克风采集的外界发出的声音信号进行一次降噪处理,具体包括:

[0064] 将至少两个麦克风采集的外界发出的声音信号进行波束成形处理后,再噪音抑制处理。

[0065] 可选的,声音融合装置将用户的语音信号,以及经过一次降噪处理后的信号进行声音融合处理,获取输出信号,用以作为语音识别设备的语音识别信号,具体包括:

[0066] 声音融合装置利用自适应滤波方法对用户的语音信号和经过一次降噪处理后的信号进行融合处理,获取输出信号,用以作为语音识别设备的语音识别信号。

[0067] 可选的,语音信号增强系统还包括:降噪装置,声音融合装置将用户的语音信号,以及经过一次降噪处理后的信号进行声音融合处理,获取输出信号之后,该方法还包括:

[0068] 利用降噪装置对输出信号进行二次降噪处理,以便将经过二次降噪处理后的信号作为语音识别设备的语音识别信号。

[0069] 可选的,语音信号增强系统还包括:声音活动检测装置,骨传导语音传感器、麦克风构成声音传感器组;当声音传感器组中除预设声音传感器之外的声音传感器、声音增强处理装置以及声音融合装置进入工作状态之前,该方法还包括:当声音活动检测装置接收

到声音传感器组中预设声音传感器传输的用户发出的语音信号时,唤醒声音传感器组中除预设声音传感器之外的声音传感器、声音增强处理装置以及声音融合装置从休眠状态进入工作状态。

[0070] 可选的,预设声音传感器的数量为至少一个。

[0071] 本发明实施例提供一种语音信号增强方法中各步骤均已在上述实施例1中做了详细介绍,因此这里不做过多赘述。

[0072] 本发明实施例提供一种语音信号增强方法,骨传导语音传感器主要采集用户的语音信号中的低频,而麦克风则采集外界发出的声音信号,其采集的声音信号中实际是包含了用户的语音信号和外界噪音。声音增强处理装置将外界发出的声音信号进行降噪处理后,可以滤除掉噪音信号。声音融合装置将骨传导语音传感器采集用户的语音信号中低频信号,和经过降噪处理后的信号进行融合,从而得到一个还原用户真实声音,且已经经过降噪处理的干净的信号,这个信号将作为语音识别设备的输入信号,供语音识别设备进行精准识别,并执行后续的操作。经过上述处理后的信号更加容易被语音识别设备识别,进而可以提升用户对语音识别设备的体验度。

[0073] 与上述实施例相对应的,本发明实施例还提供了一种计算机存储介质,该计算机存储介质中包含一个或多个程序指令。其中,一个或多个程序指令用于被如实施例1所介绍的一种语音信号增强系统执行如实施例2所介绍的一种语音信号增强方法。

[0074] 虽然,上文中已经用一般性说明及具体实施例对本发明作了详尽的描述,但在本发明基础上,可以对之作一些修改或改进,这对本领域技术人员而言是显而易见的。因此,在不偏离本发明精神的基础上所做的这些修改或改进,均属于本发明要求保护的范围。

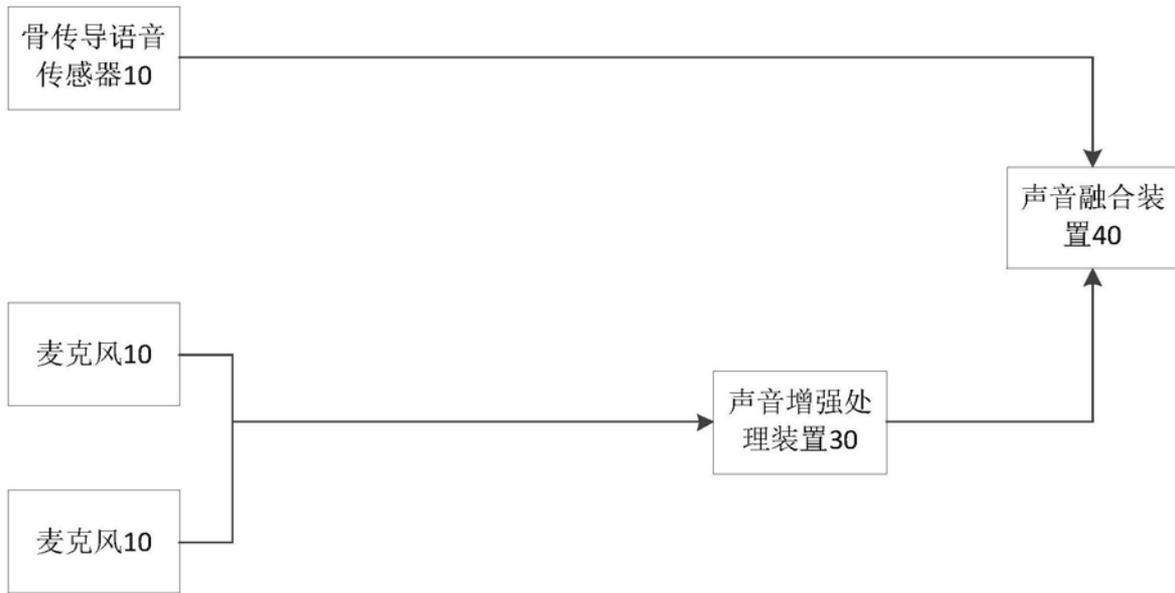


图1

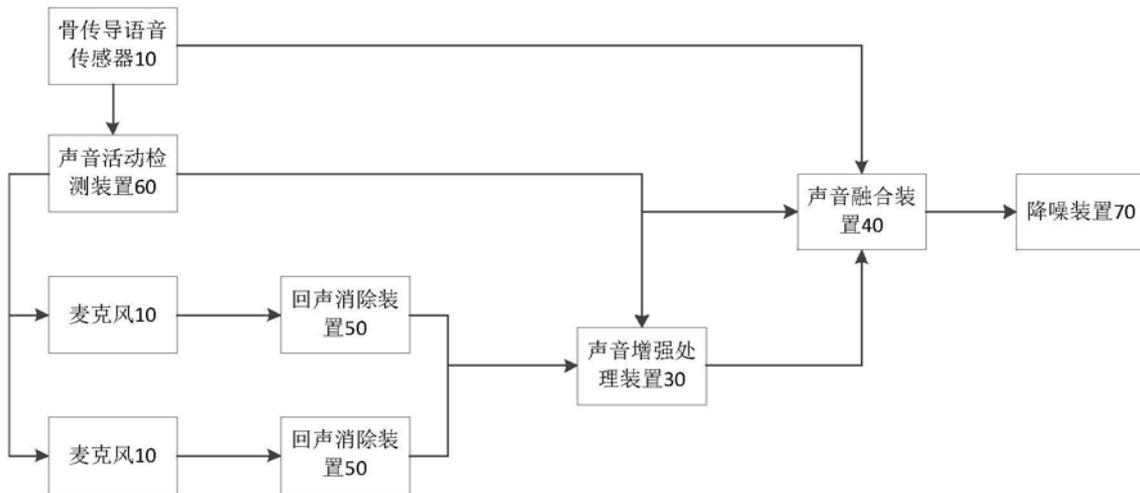


图2

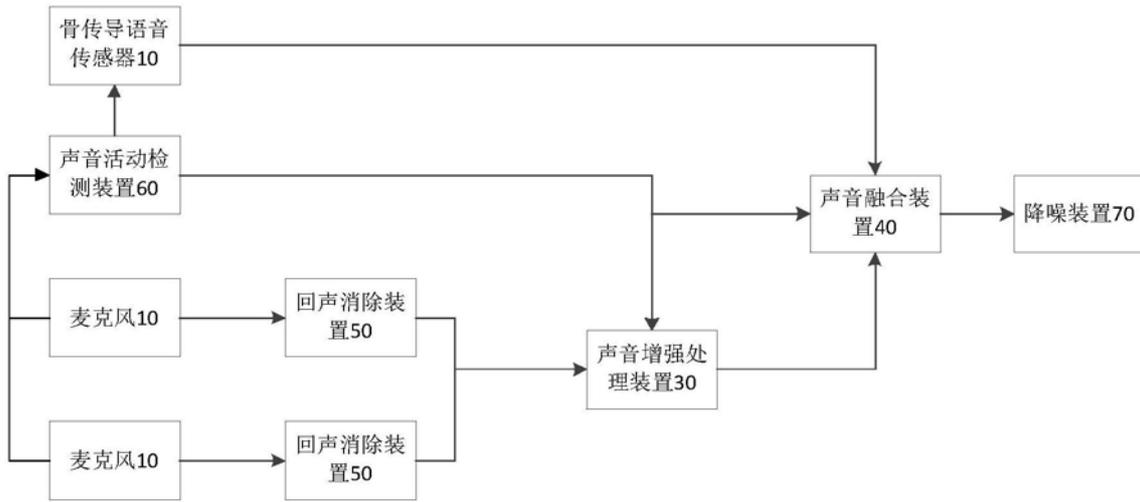


图3

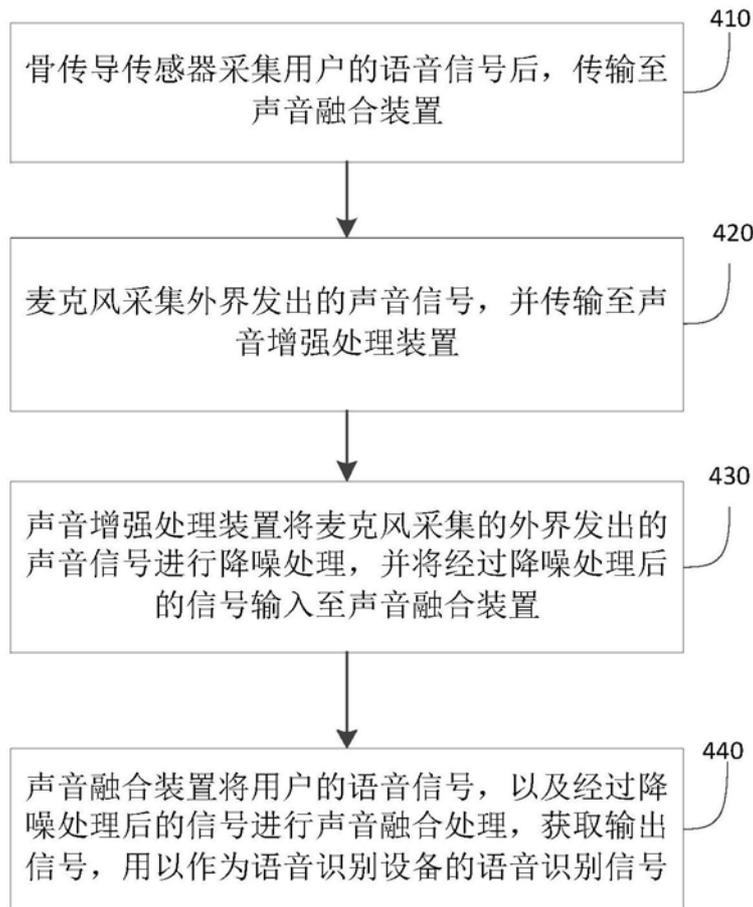


图4