



(12)发明专利申请

(10)申请公布号 CN 106683684 A

(43)申请公布日 2017. 05. 17

(21)申请号 201611101545.7

(22)申请日 2016.12.05

(71)申请人 上海木谷机器人技术有限公司

地址 200336 上海市长宁区广顺路33号2幢
402室

(72)发明人 蒋化冰 齐鹏举 方园 米万珠
舒剑 吴琨 管伟 罗璇 罗承雄
张海建

(74)专利代理机构 北京太合九思知识产权代理
有限公司 11610

代理人 刘戈

(51) Int. Cl.

G10L 21/0216(2013.01)

G10L 25/21(2013.01)

G10L 15/00(2013.01)

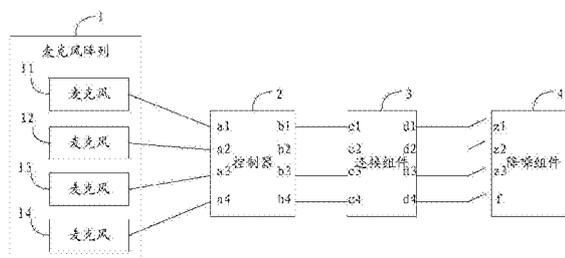
权利要求书1页 说明书6页 附图2页

(54)发明名称

音频信号处理系统和音频信号处理方法

(57)摘要

本发明实施例提供一种音频信号处理系统和音频信号处理方法,该系统包括:由N个麦克风组成的麦克风阵列,控制器,连接组件和降噪组件;其中,N个麦克风呈圆环状分布,降噪组件的输入端包括N-1个主麦克接口和1个辅麦克接口;N个麦克风分别与控制器连接,用于将采集的N路音频信号输入控制器;控制器通过连接组件与降噪组件的输入端连接,用于对N路音频信号进行信号强度比较,根据比较结果控制连接组件的输出端与降噪组件的输入端之间的连接关系,从而实现了根据音频信号强度自适应地切换主、辅麦克接口的输入信号,以自适应当前的智能语音交互场景,保证良好的降噪效果,有助于提高语音识别结果的准确性。



1. 一种音频信号处理系统,其特征在于,包括:

由N个麦克风组成的麦克风阵列,控制器,连接组件,以及降噪组件,N为大于2的整数;其中,

所述N个麦克风呈圆环状分布;所述降噪组件的输入端包括N-1个主麦克接口和1个辅麦克接口;

所述N个麦克风分别与所述控制器连接,用于将采集的N路音频信号输入所述控制器;

所述控制器通过所述连接组件与所述降噪组件的输入端连接,用于对所述N路音频信号进行信号强度比较,根据比较结果控制所述连接组件的输出端与所述降噪组件的输入端之间的连接关系。

2. 根据权利要求1所述的系统,其特征在于,所述控制器具体用于:控制所述连接组件将信号强度最低的一路音频信号输入至所述1个辅麦克接口,将其他的N-1路音频信号分别输入至所述N-1个主麦克接口。

3. 根据权利要求1所述的系统,其特征在于,所述连接组件包括N个开关,所述N个开关的输入端与所述控制器的N个音频信号输出端分别对应固定连接。

4. 根据权利要求1所述的系统,其特征在于,所述麦克风阵列为全向麦克风阵列。

5. 根据权利要求1所述的系统,其特征在于,所述麦克风阵列被水平设置在机器人机身上。

6. 根据权利要求1所述的系统,其特征在于,还包括:

分别与所述N个麦克风和所述控制器连接的桥接器;

所述桥接器,用于将所述N路音频信号转换为一路音频信号传输给所述控制器;

所述控制器还用于:将所述一路音频信号还原为所述N路音频信号。

7. 根据权利要求1至6中任一项所述的系统,其特征在于,还包括:

防尘防风装置;

所述麦克风阵列安装在所述防尘防风装置内。

8. 根据权利要求1至6中任一项所述的系统,其特征在于,还包括:

语音识别组件和交互组件;

所述语音识别组件分别与所述降噪组件的输出端和所述控制器连接,用于对降噪后的音频信号进行语音识别,将语音识别结果输入给所述控制器;

所述控制器还用于根据所述语音识别结果控制所述交互组件进行相应的交互反馈。

9. 一种音频信号处理方法,其特征在于,包括:

获取N路音频信号,所述N路音频信号是由N个呈圆环状分布的麦克风分别采集的,N为大于2的整数;

对所述N路音频信号进行信号强度比较;

根据信号强度比较结果,确定所述N路音频信号中的N-1路主音频信号和1路辅音频信号;

以所述1路辅音频信号对所述N-1路主音频信号进行降噪处理。

音频信号处理系统和音频信号处理方法

技术领域

[0001] 本发明涉及信号处理技术领域,尤其涉及一种音频信号处理系统和音频信号处理方法。

背景技术

[0002] 语音交互,已经遍及我们生活的各个领域,如手机、电视、车载、空调等等领域。能够进行正常语音交互的前提是准确识别语音信号。以手机终端为例,由于环境噪声的存在,手机终端的麦克风在采集用户的有效声源信号的同时,也会采集到环境噪声信号,这些环境噪声信号会对有效声源信号的准确识别造成干扰。

[0003] 为了提高语音识别的准确性,一种方式是采用麦克风阵列以抑制噪声的干扰,从而提升语音识别效果,其中一种简单而常用的是双麦克风阵列结构。以手机终端为例,双麦克风阵列由主麦克风和辅麦克风组成,分别设置在手机终端的不同位置,一般地,靠近用户说话的位置处设置主麦克风,远离用户说话的位置设置辅麦克风。其中,主麦克风用于采集语音信号和环境噪声,辅麦克风采集环境噪声,两者信号进行相减运算可以抑制环境噪声对语音信号的干扰。

[0004] 目前,在诸如手机终端等场景中,用户在使用手机终端时,声音源方位相对固定,从而双麦克风位置固定,即哪个作为主麦克风哪个作为辅麦克风已经固定设置。但是,随着各种智能交互产品的不断问世,语音交互的场景发生了很大改变,声音源相对智能交互产品的方位不再固定不变,此时,固定设置某个麦克风作为主麦克风、另一麦克风作为辅麦克风已经不能灵活适应智能语音交互场景的需求,很可能导致声源信号的识别准确性大大降低。

发明内容

[0005] 有鉴于此,本发明实施例提供一种音频信号处理系统和音频信号处理方法,能够自适应语音交互场景,有助于提高语音识别结果的准确性。

[0006] 本发明实施例提供一种音频信号处理系统,包括:

[0007] 由N个麦克风组成的麦克风阵列,控制器,连接组件,以及降噪组件,N为大于2的整数;其中,

[0008] 所述N个麦克风呈圆环状分布;所述降噪组件的输入端包括N-1个主麦克接口和1个辅麦克接口;

[0009] 所述N个麦克风分别与所述控制器连接,用于将采集的N路音频信号输入所述控制器;

[0010] 所述控制器通过所述连接组件与所述降噪组件的输入端连接,用于对所述N路音频信号进行信号强度比较,根据比较结果控制所述连接组件的输出端与所述降噪组件的输入端之间的连接关系。

[0011] 本发明实施例提供一种音频信号处理方法,包括:

[0012] 获取N路音频信号,所述N路音频信号是由N个呈圆环状分布的麦克风分别采集的,N为大于2的整数;

[0013] 对所述N路音频信号进行信号强度比较;

[0014] 根据信号强度比较结果,确定所述N路音频信号中的N-1路主音频信号和1路辅音频信号;

[0015] 以所述1路辅音频信号对所述N-1路主音频信号进行降噪处理。

[0016] 本发明实施例提供的音频信号处理系统和音频信号处理方法,该系统中包括由N个麦克风组成的圆环形麦克风阵列,控制器,连接组件,以及降噪组件,其中,控制器通过连接组件与降噪组件的N个输入端连接。当N个麦克风分别采集到音频信号时,将采集到的N路音频信号输入控制器,控制器对N路音频信号进行信号强度比较,根据比较结果自适应地进行N路音频信号的主、辅角色定位。具体地,控制器根据N路音频信号的信号强度比较结果控制连接组件的输出端与降噪组件的主、辅麦克接口之间的连接关系,从而实现了根据音频信号强度自适应地切换主、辅麦克接口的输入信号,以自适应当前的智能语音交互场景,也保证了降噪组件实现更佳的降噪效果,有助于提高语音识别结果的准确性。

附图说明

[0017] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例,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,还可以根据这些附图获得其他的附图。

[0018] 图1为本发明实施例提供的音频信号处理系统实施例一的结构示意图;

[0019] 图2为本发明实施例提供的音频信号处理系统实施例二的结构示意图;

[0020] 图3为本发明实施例提供的音频信号处理方法实施例一的流程图。

具体实施方式

[0021] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,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,都属于本发明保护的范围。

[0022] 在本发明实施例中使用的术语是仅仅出于描述特定实施例的目的,而非旨在限制本发明。在本发明实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式,除非上下文清楚地表示其他含义,“多种”一般包含至少两种,但是不排除包含至少一种的情况。

[0023] 应当理解,本文中使用的术语“和/或”仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B这三种情况。另外,本文中字符“/”,一般表示前后关联对象是一种“或”的关系。

[0024] 应当理解,尽管在本发明实施例中可能采用术语第一、第二、第三等来描述XXX,但这些XXX不应限于这些术语。这些术语仅用来将XXX彼此区分开。例如,在不脱离本发明实施例范围的情况下,第一XXX也可以被称为第二XXX,类似地,第二XXX也可以被称为第一XXX。

[0025] 取决于语境,如在此所使用的词语“如果”、“若”可以被解释成为“在……时”或“当……时”或“响应于确定”或“响应于检测”。类似地,取决于语境,短语“如果确定”或“如果检测(陈述的条件或事件)”可以被解释成为“当确定时”或“响应于确定”或“当检测(陈述的条件或事件)时”或“响应于检测(陈述的条件或事件)”。

[0026] 还需要说明的是,术语“包括”、“包含”或者其任何其他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商品或者系统不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种商品或者系统所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的商品或者系统中还存在另外的相同要素。

[0027] 图1为本发明实施例提供的音频信号处理系统实施例一的结构示意图,如图1所示,该系统包括:

[0028] 由N个麦克风组成的麦克风阵列1,控制器2,连接组件3,以及降噪组件4,N为大于2的整数。

[0029] 其中,如图1所示,该N个麦克风呈圆环状分布,图中示意出了该麦克风阵列包括11、12、13、14四个麦克风。

[0030] 在本实施例中,该圆环形状的麦克风阵列为全向麦克风阵列,即通过该圆环形状的麦克风阵列可以接收全方位的音频信号。

[0031] 在实际应用中,这N个麦克风可以是均匀分布的,也可以根据实际应用情况非均匀分布,比如结合实际应用场景,如果用户在某些方位说话的可能性更大,则在该方位上,可以布设比较多的麦克风,在不易说话的方位,布设少量麦克风。

[0032] 本实施例中提供的上述音频信号处理系统可以适用于多种智能交互产品中,本实施例中,以应用在智能移动机器人中为例,此时,麦克风阵列1可以被水平设置在机器人机身上的适当位置,比如麦克风阵列1设置在机器人头部。

[0033] 当该机器人被启动语音交互功能后,麦克风阵列1用于采集用户交互语音。

[0034] 如图1所示,N个麦克风分别与控制器2连接,用于将采集的N路音频信号输入控制器2。在一种可选连接方式中,N个麦克风可以直接与控制器2连接,此时,如图1所示,在包含11、12、13、14四个麦克风的举例中,这四个麦克风分别连接到控制器2的a1、a2、a3、a4四个输入端。

[0035] 本实施例中,控制器2进一步通过连接组件3与降噪组件4的输入端连接,其中,降噪组件4的输入端包括N-1个主麦克接口和1个辅麦克接口。具体地,如图1中示意的,控制器2的四个输出端b1、b2、b3、b4分别与连接组件3的四个输入端c1、c2、c3、c4对应固定连接,连接组件3的四个输出端d1、d2、d3、d4则可以切换与降噪组件4的N-1个主麦克接口和1个辅麦克接口之间的连接关系,图1中,以z1、z2、z3示意主麦克接口,以f示意辅麦克接口。

[0036] 可选地,该连接组件3可以实现为N个开关。

[0037] 在实际应用中,当某个用户开始说话时,触发N个麦克风采集音频信号,将采集的N路音频信号通过控制器2的输入端输入给控制器2,进而,控制器2对N路音频信号进行信号强度比较,根据比较结果控制连接组件3的输出端与降噪组件4的输入端之间的连接关系,将根据信号强度比较结果确定的主音频信号输入至降噪组件4的主麦克接口,辅音频信号输入至降噪组件4的辅麦克接口。

[0038] 具体地,控制器2确定信号强度最低的一路音频信号为辅音频信号,其他N-1路音频信号为主音频信号,从而,控制连接组件3将信号强度最低的一路音频信号输入至1个辅麦克风接口,将其他的N-1路音频信号分别输入至N-1个主麦克风接口。

[0039] 其中,主麦克风接口和辅麦克风接口可以被视为是对音频信号的信号属性的表征。具体来说,可以将主麦克风接口输入的音频信号视为声音源信号,即含有有效语音成分更多的信号,将从辅麦克风接口输入的音频信号视为噪声信号,从而触发降噪组件4根据辅麦克风接口输入的音频信号对从主麦克风接口输入的音频信号的降噪处理。

[0040] 具体地,控制器2在接收到N路音频信号后,可以通过对这N路音频信号分别进行一定的信号处理,比如放大、滤波等处理,求取这N路音频信号的信号强度,进行比较。

[0041] 本实施例中,可选地,控制器2可以使用各种应用专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理设备(DSPD)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、微中控元件、微处理器或其他电子元件实现。

[0042] 本实施例提供的音频信号处理系统中包括由N个麦克风组成的圆环形麦克风阵列,控制器,连接组件,以及降噪组件,其中,控制器通过连接组件与降噪组件的N个输入端连接。当N个麦克风分别采集到音频信号时,将采集到的N路音频信号输入控制器,控制器对N路音频信号进行信号强度比较,根据比较结果自适应地进行N路音频信号的主、辅角色定位。具体地,控制器根据N路音频信号的信号强度比较结果控制连接组件的输出端与降噪组件的主、辅麦克风接口之间的连接关系,从而实现了根据音频信号强度自适应地切换主、辅麦克风接口的输入信号,以自适应当前的智能语音交互场景,也保证了降噪组件实现更佳的降噪效果,有助于提高语音识别结果的准确性。

[0043] 图2为本发明实施例提供的音频信号处理系统实施例二的结构示意图,如图2所示,在图1所示实施例基础上,可选地,该系统还包括:

[0044] 分别与N个麦克风和控制器2连接的桥接器5,其中:

[0045] 桥接器5,用于将N路音频信号转换为一路音频信号传输给控制器2;

[0046] 控制器2还用于:将一路音频信号还原为N路音频信号。

[0047] 由于支持多路音频信号输入的控制器成本较贵,且接口数量很有限,为了降低成本以及避免控制器接口数量的限制,本实施例中提供了上述桥接器5。

[0048] 如图2所示,由于桥接器5具有N个音频信号输入端口和一个音频信号输出端口,其分别接收N路音频信号,将N路音频信号转换为一路后,输入给控制器2,控制器2为了接收N路音频信号只需要设置一个输入接口a1即可,实现方便、成本低廉。而桥接器可以选择一个CPLD或者FPGA就可以方便实现。

[0049] 可选地,该系统还包括:防尘防风装置6。其中,麦克风阵列1安装在防尘防风装置6内。

[0050] 本实施例中,为了物理上尽量保证环境因素对语音识别结果的不利影响,在麦克风阵列的组装工艺上提供了防尘防风装置6,以尽量降低环境因素对语音识别结果的不利影响。

[0051] 其中,防尘防风装置6中比如包括防风棉、防尘网等结构,以降低风声、粉尘对麦克风阵列的影响。

[0052] 可选地,该系统还包括:语音识别组件7和交互组件8。

[0053] 其中,语音识别组件7分别与降噪组件4的输出端和控制器2连接,用于对降噪后的音频信号进行语音识别,将语音识别结果输入给控制器2。

[0054] 控制器2还用于根据语音识别结果控制交互组件8进行相应的交互反馈。

[0055] 本发明实施例提供的音频信号处理系统一般适用于智能语音交互的产品中,为了实现智能语音交互功能,在通过降噪组件4对输入的N路音频信号进行了降噪处理后,降噪后的音频信号输入给语音识别组件7,以完成用户输入语音的语音识别处理。同时,为了实现基于语音的智能交互,以机器人为例,需要基于语音识别结果向用户进行相应的反馈。本实施例中,以机器人为例,该交互组件8比如可以是语音播放器,控制器2可以基于语音识别结果通过语音播放器向用户反馈应答语音;再比如可以是显示屏,控制器2可以基于语音识别结果通过显示屏向用户反馈某种业务操作界面;再比如还可以是运动部件,控制器2可以基于语音识别结果通过控制运动部件使机器人执行相应的反馈动作,等等。

[0056] 图3为本发明实施例提供的音频信号处理方法实施例一的流程图,本实施例提供的该音频信号处理方法可以由一音频信号处理系统来执行,该音频信号处理系统可以实现为硬件,或者实现为软件和硬件的组合,该音频信号处理系统可以集成设置比如移动机器人等语音交互设备中,比如可以是图1、图2所示的系统结构。如图3所示,该方法包括如下步骤:

[0057] 步骤101、获取N路音频信号,N路音频信号是由N个呈圆环状分布的麦克风分别采集的。

[0058] 其中,N为大于2的整数。

[0059] 步骤102、对N路音频信号进行信号强度比较。

[0060] 步骤103、根据信号强度比较结果,确定N路音频信号中的N-1路主音频信号和1路辅音频信号。

[0061] 步骤104、以该1路辅音频信号对N-1路主音频信号进行降噪处理。

[0062] 具体地,根据信号强度比较结果,确定N路音频信号中的N-1路主音频信号和1路辅音频信号,包括:

[0063] 根据信号强度的比较结果,确定信号强度最低的一路音频信号为辅音频信号,其他N-1路音频信号为主音频信号。

[0064] 本实施例提供的音频信号处理方法的具体适用场景和详细过程,可以参见前述系统实施例中的说明,在此不赘述。

[0065] 以上所描述的系统实施例仅仅是示意性的,其中所述作为分离部件说明的单元(诸如各种组件、装置等)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模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性的劳动的情况下,即可以理解并实施。

[0066] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到各实施方式可借助加必需的通用硬件平台的方式来实现,当然也可以通过硬件。基于这样的理解,上述技术方案本质上或者说对现有技术做出贡献的部分可以以产品的形式体现出来,该计算机产品可以存储在计算机可读存储介质中,如ROM/RAM、磁碟、光盘等,包括若干指令用以使得一台计算机装置(可以是个人计算机,服务器,或者网络装置等)执行各个实施例或者实施例

的某些部分所述的方法。

[0067] 最后应说明的是：以上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精神和范围。

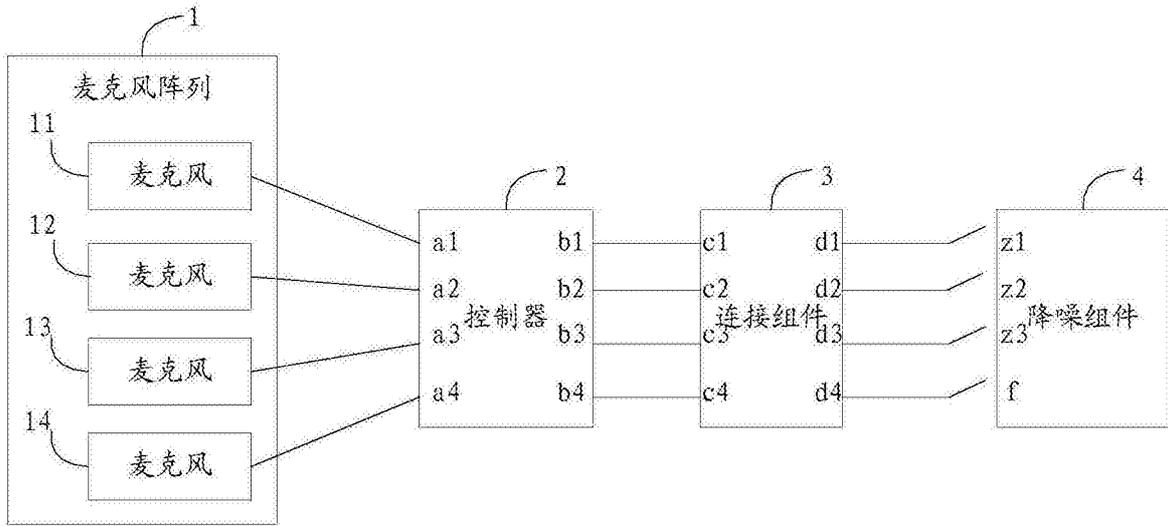


图1

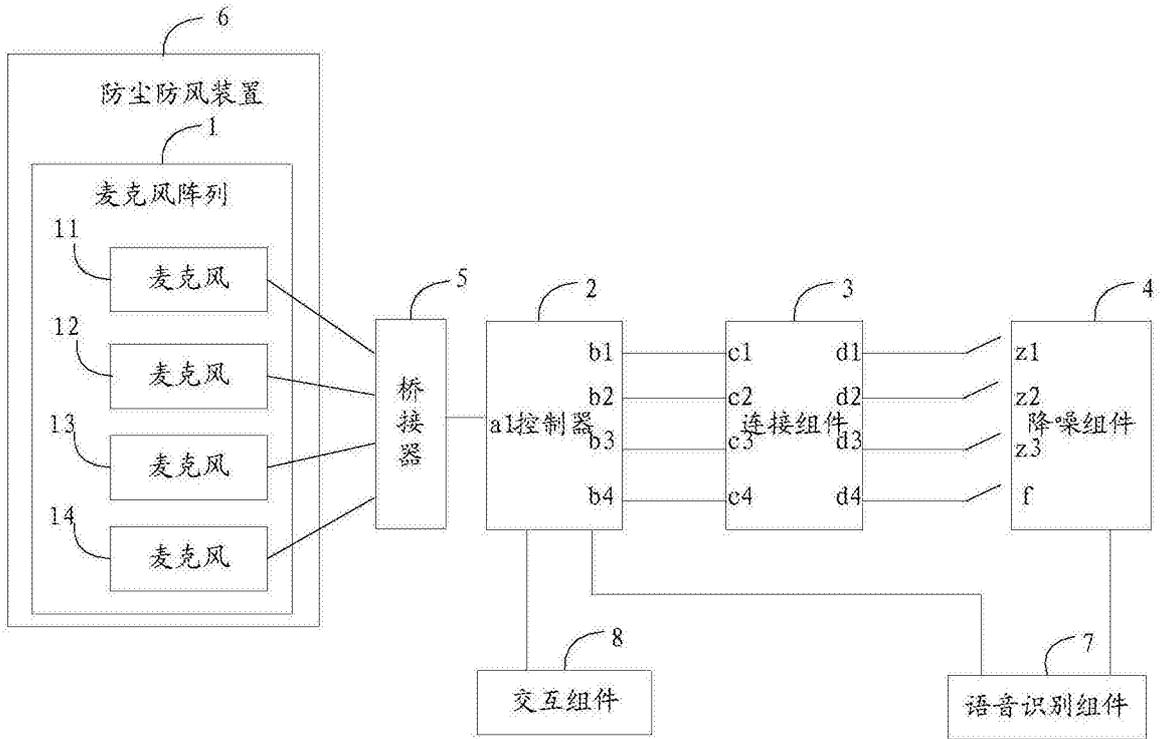


图2

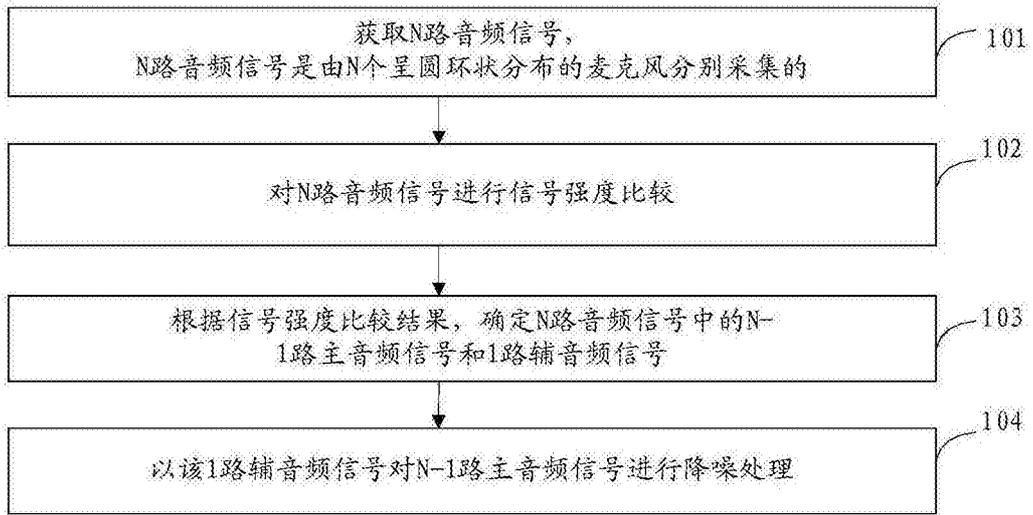


图3