



(12) 发明专利申请

(10) 申请公布号 CN 112599144 A

(43) 申请公布日 2021.04.02

(21) 申请号 202011412928.2

(22) 申请日 2020.12.03

(71) 申请人 OPPO(重庆)智能科技有限公司
地址 401120 重庆市渝北区玉峰山镇玉龙大道188号

(72) 发明人 彭翊

(74) 专利代理机构 深圳市隆天联鼎知识产权代理有限公司 44232
代理人 刘抗美

(51) Int. Cl.

G10L 21/0208 (2013.01)

G10L 21/0272 (2013.01)

G10L 21/0308 (2013.01)

G06F 17/12 (2006.01)

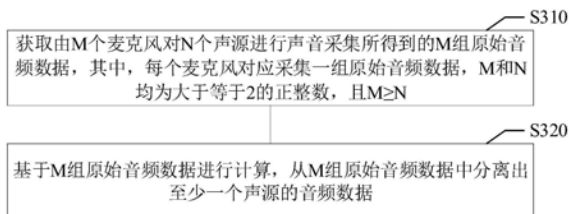
权利要求书2页 说明书11页 附图4页

(54) 发明名称

音频数据处理方法、音频数据处理装置、介质与电子设备

(57) 摘要

本公开提供一种音频数据处理方法、音频数据处理装置、计算机可读存储介质与电子设备，涉及音频处理技术领域。该音频数据处理方法包括：获取由M个麦克风对N个声源进行声音采集所得到的M组原始音频数据，其中，每个麦克风对应采集一组原始音频数据，M和N均为大于等于2的正整数，且M≥N；基于所述M组原始音频数据进行计算，从所述M组原始音频数据中分离出至少一个声源的音频数据。本公开可以对多个麦克风采集的原始音频数据进行有效处理，以从原始音频数据中获取对应声源的音频数据。



1. 一种音频数据处理方法,其特征在于,包括:

获取由M个麦克风对N个声源进行声音采集所得到的M组原始音频数据,其中,每个麦克风对应采集一组原始音频数据,M和N均为大于等于2的正整数,且 $M \geq N$;

基于所述M组原始音频数据进行计算,从所述M组原始音频数据中分离出至少一个声源的音频数据。

2. 根据权利要求1所述的方法,其特征在于,所述基于所述M组原始音频数据进行计算,从所述M组原始音频数据中分离出至少一个声源的音频数据,包括:

获取每个麦克风的位姿系数;

基于所述每个麦克风的位姿系数和所述M组原始音频数据,构建以每个声源的音频数据为变量的方程;

求解所述方程,得到至少一个声源的音频数据。

3. 根据权利要求2所述的方法,其特征在于,所述M个麦克风为固定麦克风。

4. 根据权利要求1所述的方法,其特征在于,当所述M个麦克风包括至少一个非固定麦克风时,所述基于所述M组原始音频数据进行计算,从所述M组原始音频数据中分离出至少一个声源的音频数据,包括:

将所述非固定麦克风采集的原始音频数据,确定为位于所述非固定麦克风处的声源的音频数据。

5. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

提供至少一个音频控件,每个音频控件对应于一个声源;

根据用户通过所述音频控件输入的音频调节指令,对所述音频控件对应的声源的音频数据进行播放调节。

6. 根据权利要求5所述的方法,其特征在于,所述对所述音频控件对应的声源的音频数据进行播放调节,包括:

确定调节时间范围;

在所述调节时间范围内,对所述音频控件对应的声源的音频数据进行播放音量的调节。

7. 根据权利要求5所述的方法,其特征在于,在进行播放调节后,所述方法还包括:

将各声源的音频数据合成为目标音频并播放。

8. 一种音频数据处理装置,其特征在于,包括:

音频数据获取模块,用于获取由M个麦克风对N个声源进行声音采集所得到的M组原始音频数据,其中,每个麦克风对应采集一组原始音频数据,M和N均为大于等于2的正整数,且 $M \geq N$;

音频数据分离模块,用于基于所述M组原始音频数据进行计算,从所述M组原始音频数据中分离出至少一个声源的音频数据。

9. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现权利要求1至7任一项所述的方法。

10. 一种电子设备,其特征在于,包括:

处理器;以及

存储器,用于存储所述处理器的可执行指令;

其中,所述处理器配置为经由执行所述可执行指令来执行权利要求1至7任一项所述的方法。

音频数据处理方法、音频数据处理装置、介质与电子设备

技术领域

[0001] 本公开涉及音频处理技术领域,尤其涉及一种音频数据处理方法、音频数据处理装置、计算机可读存储介质与电子设备。

背景技术

[0002] 随着移动终端的多样化发展,现有的终端设备上常常配置有多个麦克风,用于对采集到的音频信号进行降噪处理。然而,当用户在不同环境或不同应用场景下进行拍摄或录音时,每个麦克风都会采集到除对应声源之外的其他干扰音频数据。而由于现有技术中缺乏有效的对声源的音频数据进行分离的方法,因此,用户只能对包含干扰数据的整体音频数据进行调节,难以针对实际需要,针对某一种声源的音频数据进行有效调节,从而也无法提供多样化的音频需求,影响用户体验。

发明内容

[0003] 本公开提供了一种音频数据处理方法、音频数据处理装置、计算机可读存储介质与电子设备,进而至少在一定程度上改善现有技术缺乏对声源的音频数据进行有效分离的问题。

[0004] 本公开的其他特性和优点将通过下面的详细描述变得显然,或部分地通过本公开的实践而习得。

[0005] 根据本公开的第一方面,提供一种音频数据处理方法,包括:获取由M个麦克风对N个声源进行声音采集所得到的M组原始音频数据,其中,每个麦克风对应采集一组原始音频数据,M和N均为大于等于2的正整数,且 $M \geq N$;基于所述M组原始音频数据进行计算,从所述M组原始音频数据中分离出至少一个声源的音频数据。

[0006] 根据本公开的第二方面,提供一种音频数据处理装置,包括:音频数据获取模块,用于获取由M个麦克风对N个声源进行声音采集所得到的M组原始音频数据,其中,每个麦克风对应采集一组原始音频数据,M和N均为大于等于2的正整数,且 $M \geq N$;音频数据分离模块,用于基于所述M组原始音频数据进行计算,从所述M组原始音频数据中分离出至少一个声源的音频数据。

[0007] 根据本公开的第三方面,提供一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现上述音频数据处理方法。

[0008] 根据本公开的第四方面,提供一种电子设备,包括:处理器;以及存储器,用于存储所述处理器的可执行指令;其中,所述处理器配置为经由执行所述可执行指令来执行上述音频数据处理方法。

[0009] 本公开的技术方案具有以下有益效果:

[0010] 根据上述音频数据处理方法、音频数据处理装置、计算机可读存储介质与电子设备,获取由M个麦克风对N个声源进行声音采集所得到的M组原始音频数据,其中,每个麦克风对应采集一组原始音频数据,M和N均为大于等于2的正整数,且 $M \geq N$;基于M组原始音频数

据进行计算,从M组原始音频数据中分离出至少一个声源的音频数据。一方面,本示例性实施例提出一种新的音频数据处理方法,针对具有多个麦克风的终端设备,可以基于麦克风采集的原始音频数据进行计算,从多组原始音频数据中分离出至少一个声源的音频数据,实现了对多麦克风终端设备采集声源的音频数据的有效处理,且分离过程简单、准确性较高;另一方面,本示例性实施例可以从每组原始音频数据中分离出对应的声源的音频数据,为后期用户针对某一声源的音频数据进行调整提供了便捷,也使得音频数据的调整更加具有针对性和灵活性。

[0011] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性和解释性的,并不能限制本公开。

附图说明

[0012] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本公开的实施例,并与说明书一起用于解释本公开的原理。显而易见地,下面描述中的附图仅仅是本公开的一些实施例,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。

[0013] 图1示出本示例性实施方式的一种系统架构的示意图;

[0014] 图2示出本示例性实施方式的电子设备的示意图;

[0015] 图3示出本示例性实施方式的一种音频数据处理方法的流程图;

[0016] 图4示出本示例性实施方式的一种音频数据处理方法的子流程图;

[0017] 图5示出本示例性实施方式的另一种音频数据处理方法的子流程图;

[0018] 图6示出本示例性实施方式的一种音频数据处理装置的结构框图。

具体实施方式

[0019] 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公开将更加全面和完整,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。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方式中。在下面的描述中,提供许多具体细节从而给出对本公开的实施方式的充分理解。然而,本领域技术人员将意识到,可以实践本公开的技术方案而省略特定细节中的一个或更多,或者可以采用其它的方法、元件、装置、步骤等。在其它情况下,不详细示出或描述公知技术方案以避免喧宾夺主而使得本公开的各方面变得模糊。

[0020] 此外,附图仅为本公开的示意性图解,并非一定是按比例绘制。图中相同的附图标记表示相同或类似的部分,因而将省略对它们的重复描述。附图中所示的一些方框图是功能实体,不一定必须与物理或逻辑上独立的实体相对应。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。

[0021] 图1示出了本公开示例性实施方式的一种系统架构的示意图。如图1所示,该系统架构100可以包括:终端设备110和服务器120。终端设备110可以是配置有多个麦克风的各种电子设备,包括但不限于手机、平板电脑、数码相机、个人电脑等。应该理解,图1中的终端

设备和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端和服务器。比如服务器120可以是多个服务器组成的服务器集群等。

[0022] 本公开实施方式所提供的音频数据处理方法可以由终端设备110执行,例如在终端设备110采集多组原始音频数据后,直接根据多组原始音频数据进行计算,实现声源音频数据的分离处理;也可以由服务器120执行,例如终端设备110采集多组原始音频数据后,上传到服务器120,由服务器120对多组原始音频数据进行计算,实现声源音频数据的分离处理等。本公开对此不做限定。

[0023] 本公开的示例性实施方式提供一种用于实现音频数据处理方法的电子设备,其可以是图1中的终端110或服务器120。该电子设备至少包括处理器和存储器,存储器用于存储处理器的可执行指令,处理器配置为经由执行可执行指令来执行音频数据处理方法。

[0024] 电子设备可以以各种形式来实施,例如可以包括手机、平板电脑、笔记本电脑、个人数字助理(Personal Digital Assistant,PDA)、导航装置、可穿戴设备、无人机等移动设备,以及台式电脑、智能电视等固定设备。

[0025] 下面以图2中的终端设备200为例,对电子设备的构造进行示例性说明。本领域技术人员应当理解,除了特别用于移动目的的部件之外,图2中的构造也能够应用于固定类型的设备。在另一些实施方式中,终端设备200可以包括比图示更多或更少的部件,或者组合某些部件,或者拆分某些部件,或者不同的部件布置。图示的部件可以以硬件、软件或软件和硬件的组合实现。各部件间的接口连接关系只是示意性示出,并不构成对终端设备200的结构限定。在另一些实施方式中,终端设备200也可以采用与图2不同的接口连接方式,或多种接口连接方式的组合。

[0026] 如图2所示,终端设备200具体可以包括:处理器210、内部存储器221、外部存储器接口222、USB接口230、充电管理模块240、电源管理模块241、电池242、天线1、天线2、移动通信模块250、无线通信模块260、音频模块270、扬声器271、受话器272、麦克风273、耳机接口274、传感器模块280、显示屏幕290、摄像模组291、指示器292、马达293、按键294以及用户标识模块(Subscriber Identification Module,SIM)卡接口295等。

[0027] 处理器210可以包括一个或多个处理单元,例如:处理器210可以包括应用处理器(Application Processor,AP)、调制解调处理器、图形处理器(Graphics Processing Unit,GPU)、图像信号处理器(Image Signal Processor,ISP)、控制器、编码器、解码器、数字信号处理器(Digital Signal Processor,DSP)、基带处理器和/或神经网络处理器(Neural-Network Processing Unit,NPU)等。编码器可以对图像或视频数据进行编码(即压缩),形成码流数据;解码器可以对图像或视频的码流数据进行解码(即解压缩),以还原出图像或视频数据。

[0028] 在一些实施方式中,处理器210可以包括一个或多个接口。接口可以包括集成电路(Integrated Circuit,I2C)接口、集成电路内置音频(Integrated Circuit Sound,I2S)接口、脉冲编码调制(Pulse Code Modulation,PCM)接口、通用异步收发传输器(Universal Asynchronous Receiver/Transmitter,UART)接口、移动产业处理器接口(Mobile Industry Processor Interface,MIPI)、通用输入输出(General-Purpose Input/Output,GPIO)接口、用户标识模块(Subscriber Identity Module,SIM)接口和/或通用串行总线(Universal Serial Bus,USB)接口等。通过不同的接口和终端设备200的其

他部件形成连接。

[0029] USB接口230是符合USB标准规范的接口,具体可以是MiniUSB接口, MicroUSB接口, USBTypeC接口等。USB接口230可以用于连接充电器为终端设备200充电,也可以连接耳机,通过耳机播放音频,还可以用于终端设备200连接其他电子设备,例如连接电脑、外围设备等。

[0030] 充电管理模块240用于从充电器接收充电输入。充电管理模块240为电池242充电的同时,还可以通过电源管理模块241为设备供电。

[0031] 电源管理模块241用于连接电池242、充电管理模块240与处理器210。电源管理模块241接收电池242和/或充电管理模块240的输入,为终端设备200的各个部分供电,还可以用于监测电池的状态。

[0032] 终端设备200的无线通信功能可以通过天线1、天线2、移动通信模块250、无线通信模块260、调制解调处理器以及基带处理器等实现。天线1和天线2用于发射和接收电磁波信号。移动通信模块250可以提供应用在终端设备200上的包括2G/3G/4G/5G等无线通信的解决方案。

[0033] 无线通信模块260可以提供应用在终端设备200上的包括无线局域网(Wireless Local Area Networks, WLAN)(如无线保真(Wireless Fidelity, Wi-Fi)网络)、蓝牙(Bluetooth, BT)、近距离无线通信技术(Near Field Communication, NFC)、红外技术(Infrared, IR)等无线通信解决方案。无线通信模块260经由天线2接收电磁波,将电磁波信号调频以及滤波处理,将处理后的信号发送到处理器210。无线通信模块260还可以从处理器210接收待发送的信号,对其进行调频,放大,经天线2转为电磁波辐射出去。

[0034] 在一些实施方式中,终端设备200的天线1和移动通信模块250耦合,天线2和无线通信模块260耦合,使得终端设备200可以通过无线通信技术与网络以及其他设备通信。无线通信技术可以包括全球移动通讯系统(Global System for Mobile communications, GSM),通用分组无线服务(General Packet Radio Service, GPRS),码分多址接入(Code Division Multiple Access, CDMA),宽带码分多址(Wideband Code Division Multiple Access, WCDMA),时分码分多址(Time Division-Synchronous Code Division Multiple Access, TD-SCDMA),长期演进(Long Term Evolution, LTE)和/或IR技术等。

[0035] 终端设备200通过GPU、显示屏幕290及应用处理器等实现显示功能。GPU用于执行数学和几何计算,以实现图形渲染,并连接显示屏幕290和应用处理器。处理器210可包括一个或多个GPU,其执行程序指令以生成或改变显示信息。终端设备200可以包括一个或多个显示屏幕290,用于显示图像,视频等。

[0036] 终端设备200可以通过ISP、摄像模组291、编码器、解码器、GPU、显示屏幕290及应用处理器等实现拍摄功能。摄像模组291用于捕获静态图像或视频,通过感光元件采集光信号,转换为电信号。ISP用于处理摄像模组291反馈的数据,将电信号转换成数字图像信号。

[0037] 外部存储器接口222可以用于连接外部存储卡,例如Micro SD卡,实现扩展终端设备200的存储能力。内部存储器221可以用于存储计算机可执行程序代码,可执行程序代码包括指令。内部存储器221可以包括存储程序区和存储数据区。存储数据区可存储终端设备200使用过程中所创建的数据(比如图像,视频)等。处理器210通过运行存储在内部存储器221的指令和/或存储在设置于处理器中的存储器的指令,执行终端设备200的各种功能应

用以及数据处理。

[0038] 终端设备200可以通过音频模块270、扬声器271、受话器272、麦克风273、耳机接口274及应用处理器等实现音频功能。例如音乐播放、录音等。音频模块270用于将数字音频信息转换成模拟音频信号输出,也用于将模拟音频输入转换为数字音频信号。音频模块270还可以用于对音频信号编码和解码。扬声器271,用于将音频电信号转换为声音信号。受话器272,用于将音频电信号转换成声音信号。麦克风273,用于将声音信号转换为电信号,在本示例性实施例中,终端设备200中可以包括多个麦克风273,例如两个、三个或其他更多的麦克风等,各麦克风可以根据需要设置在终端设备的特定位置,例如当配置有两个麦克风时,可以将主麦克风设置在终端设备的下方,将辅麦克风设置在终端设备的上方等。耳机接口274用于连接有线耳机。

[0039] 传感器模块280可以包括触摸传感器2801、压力传感器2802、陀螺仪传感器2803、气压传感器2804等。触摸传感器2801用于感应外部输入的触摸事件。压力传感器2802用于感受压力信号,可以将压力信号转换成电信号,用于实现压力触控等功能。陀螺仪传感器2803可以用于确定终端设备200的运动姿态。气压传感器2804用于测量气压,可通过计算海拔高度,辅助定位和导航。此外,根据实际需要,还可以在传感器模块280中设置其他功能的传感器,例如深度传感器、加速度传感器、距离传感器等。

[0040] 指示器292可以是指示灯,可以用于指示充电状态,电量变化,也可以用于指示消息,未接来电,通知等。

[0041] 马达293可以产生振动提示,例如来电、闹钟、接收信息等的振动提示,也可以用于触摸振动反馈等。

[0042] 按键294包括开机键,音量键等。终端设备200可以接收按键输入,产生与终端设备200的用户设置以及功能控制有关的键信号输入。

[0043] 终端设备200可以支持一个或多个SIM卡接口295,用于连接SIM卡,使终端设备200通过SIM卡和网络交互,实现通话以及数据通信等功能。

[0044] 下面对本公开示例性实施方式的音频数据处理方法和音频数据处理装置进行具体说明。

[0045] 图3示出了本示例性实施方式中一种音频数据处理方法的流程,包括以下步骤S310~S320:

[0046] 步骤S310,获取由M个麦克风对N个声源进行声音采集所得到的M组原始音频数据,其中,每个麦克风对应采集一组原始音频数据,M和N均为大于等于2的正整数,且 $M \geq N$ 。

[0047] 麦克风可以将外部的声音信号转换为电信号,以使终端能够对电信号进行相应处理。现有的终端设备中通常可以配置两个及两个以上的麦克风,用于采集多组音频信号,例如在智能手机的下方和上方分别设置一个麦克风,其中下方的麦克风可以作为主麦克风,在用户进行语音通话时采集包含用户通话音频的音频数据,上方的麦克风可以作为辅麦克风,在用户开启免提通话时采集包含用户通话音频的音频数据等,或者通过辅麦克风采集的音频数据对用户通话的音频数据进行降噪处理。但是,无论是哪一麦克风采集的音频数据都会包含声源之外的其他音频数据。

[0048] 在本示例性实施例中,终端设备可以配置有M个麦克风,且每个麦克风可以对应采集一组原始音频数据,该原始音频数据中可以包括声源的音频数据和干扰音频数据。其中,

声源的音频数据可以认为是麦克风主要采集的声音数据,其可以是正对麦克风方向采集的声音,或者在麦克风预设角度或范围内采集的声音等,例如用户在打电话时,嘴部朝向智能手机的某一个麦克风或者某几个麦克风,则该用户说话的声音可以作为一个声源的音频数据。声源可以是一种声音也可以是多种声音,例如在辩论赛中,某一方辩论队中多个人的声音可以共同作为一个声源的音频数据。不同的麦克风采集的原始音频数据中的声源的音频数据可以相同,即M个麦克风可以对N个声源进行声音采集,其中M大于N,例如在一用户说话时,采用三个麦克风采集三组原始音频数据,这三组原始音频数据中包含的声源音频数据相同,均为该用户说话时的音频数据;不同的麦克风采集的原始音频数据中的声源音频数据也可以不同,即M个麦克风可以对N个声源进行声音采集,其中M等于N,例如在第一用户和第二用户对话时,通过配置有两个麦克风的智能手机进行录音,第一麦克风和第二麦克风可以分别采集一组原始音频数据,这两组原始音频数据中的声源音频数据可以不同,具体为第一用户说话的音频数据和第二用户说话的音频数据。

[0049] 在本示例性实施例中,原始音频数据可以由终端设备直接通过麦克风采集音频数据得到,例如通过智能终端的录音功能,对周围环境的声信号进行采集等;另外,原始音频数据的采集,也可以在视频数据采集时同步进行,例如对周围环境进行摄像,同步进行录音,得到包含音频数据的视频数据,基于该音频数据进行处理等等。

[0050] 步骤S320,基于M组原始音频数据进行计算,从M组原始音频数据中分离出至少一个声源的音频数据。

[0051] 由于原始音频数据中含有除声源的音频数据之外的其他干扰音频数据,若用户直接对原始音频数据进行调整,则在调整其中声源的音频数据的同时,也会对干扰音频数据进行调整,导致难以针对某一声源的音频数据进行精准调节。例如通过第一麦克风和第二麦克风在嘈杂的环境中对第一用户和第二用户说话过程进行录音,采集到第一原始音频数据和第二原始音频数据,其中第一原始音频数据中包含的声源的音频数据为第一用户说话的音频数据,第二原始音频数据中包含的声源的音频数据为第二用户说话的音频数据,若用户为了调高第一用户说话的声音,调节第一原始音频数据的音量,则第一原始音频数据中的干扰音频数据的声音也会随之增大。因此,本示例性实施例基于对采集到的多组原始音频数据进行计算,具体的,可以基于每个麦克风的位姿系数、原始音频数据,构建以每个声源的音频数据为变量的方程,通过求解该方程,从多组原始音频数据中分离出至少一个声源的音频数据。在得到每个声源的音频数据后,本示例性实施例可以根据用户输入的针对一个或几个声源的音频数据的调节指令,对声源的音频数据进行音量、音色等调节,得到目标音频进行播放。另外,本示例性实施例还可以在分离对声源的音频数据后,对各个声源进行识别,自适应调节满足特定条件的声源的音频数据,例如在分离出环境音和人声之后,识别出声源为人声的音频数据,自适应调节人声的音频数据的音量等。

[0052] 在一示例性实施例中,如图4所示,上述步骤S320可以包括以下步骤:

[0053] 步骤S410,获取每个麦克风的位姿系数;

[0054] 步骤S420,基于每个麦克风的位姿系数和M组原始音频数据,构建以每个声源的音频数据为变量的方程;

[0055] 步骤S430,求解方程,得到至少一个声源的音频数据。

[0056] 在终端设备中,当每个麦克风被配置完成后,可以使用特定位置的音频信号对每

个麦克风进行位姿系数的标定。在确定每个麦克风的位姿系数后,可以基于每个麦克风的位姿系数和M组原始音频数据,构建以每个声源的音频数据为变量的方程。

[0057] 下面以三个麦克风为例进行说明,设第一麦克风接收到的原始音频数据为R,其中包括第一声源的音频数据为x;第二麦克风接收到的原始音频数据为L,其中包括第二声源的音频数据为y;第三麦克风接收到的原始音频数据为K,其中包括第三声源的音频数据为z。进一步,构建关于声源的音频数据的方程,通过以下方程表示:

$$[0058] \quad \begin{cases} a_1 \cdot x + b_1 \cdot y + c_1 \cdot z = R \\ a_2 \cdot x + b_2 \cdot y + c_2 \cdot z = L \\ a_3 \cdot x + b_3 \cdot y + c_3 \cdot z = K \end{cases}$$

[0059] 方程中 a_1 、 a_2 、 a_3 、 b_1 、 b_2 、 b_3 、 c_1 、 c_2 、 c_3 为与各麦克风相对位置有关的位姿系数。通过对上述方程的计算,可以求解得到三个声源对应的音频数据x、y、z。

[0060] 如果是四个麦克风,同理,设第四麦克风接收到的原始音频数据为F,其中包括第四声源的音频数据为f,构建以下方程:

$$[0061] \quad \begin{cases} a_1 \cdot x + b_1 \cdot y + c_1 \cdot z + d_1 \cdot f = R \\ a_2 \cdot x + b_2 \cdot y + c_2 \cdot z + d_2 \cdot f = L \\ a_3 \cdot x + b_3 \cdot y + c_3 \cdot z + d_3 \cdot f = K \\ a_4 \cdot x + b_4 \cdot y + c_4 \cdot z + d_4 \cdot f = F \end{cases}$$

[0062] 计算可以求解得到四个声源对应的音频数据x、y、z、f。

[0063] 在一示例性实施例中,上述M个麦克风为固定麦克风。

[0064] 当M个麦克风为固定麦克风时,每个麦克风被配置完成之后,其相对位置关系可以被确定,因此,可以在测试环境下,通过多组测试音频数据和测试声源的音频数据构建方程,求解得到每个麦克风对应的位姿系数。

[0065] 在一示例性实施例中,当M个麦克风包括至少一个非固定麦克风时,上述步骤S320可以包括:

[0066] 将非固定麦克风采集的原始音频数据,确定为位于非固定麦克风处的声源的音频数据。

[0067] 当M个麦克风中包括非固定麦克风时,例如外界耳机配置的麦克风或其他设备等,由于其相对于终端设备的位姿系数不固定,无法提前标定相关的位姿系数,因此,本示例性实施例可以直接将非固定麦克风采集的原始音频数据作为非固定麦克风处声源的音频数据。

[0068] 举例说明,在上述包含四个麦克风的实施例中,当第四麦克风为非固定麦克风时,可以得到以下方程:

$$[0069] \quad \begin{cases} a_1 \cdot x + b_1 \cdot y + c_1 \cdot z + d_1 \cdot f = R \\ a_2 \cdot x + b_2 \cdot y + c_2 \cdot z + d_2 \cdot f = L \\ a_3 \cdot x + b_3 \cdot y + c_3 \cdot z + d_3 \cdot f = K \\ f = F \end{cases}$$

[0070] 在上述方程组中, a_1 、 a_2 、 a_3 、 b_1 、 b_2 、 b_3 、 c_1 、 c_2 、 c_3 、 d_1 、 d_2 、 d_3 为与各麦克风相对位置有关的位姿系数。其中,除了x、y、z待求解, d_1 、 d_2 、 d_3 也未知,考虑到第四麦克风为非固定麦克风,可以忽略f对第一麦克风、第二麦克风和第三麦克风的影响,将上述方程简化为:

$$[0071] \quad \begin{cases} a_1 \cdot x + b_1 \cdot y + c_1 \cdot z = R \\ a_2 \cdot x + b_2 \cdot y + c_2 \cdot z = L \\ a_3 \cdot x + b_3 \cdot y + c_3 \cdot z = K \\ f = F \end{cases}$$

[0072] 进一步,求解方程即可以得到四个声源对应的音频数据 x 、 y 、 z 、 f 。

[0073] 综上,本示例性实施方式中,获取由 M 个麦克风对 N 个声源进行声音采集所得到的 M 组原始音频数据,其中,每个麦克风对应采集一组原始音频数据, M 和 N 均为大于等于2的正整数,且 $M \geq N$;基于 M 组原始音频数据进行计算,从 M 组原始音频数据中分离出至少一个声源的音频数据。一方面,本示例性实施例提出一种新的音频数据处理方法,针对具有多个麦克风的终端设备,可以基于麦克风采集的原始音频数据进行计算,从多组原始音频数据中分离出至少一个声源的音频数据,实现了对多麦克风终端设备采集声源的音频数据的有效处理,且分离过程简单、准确性较高;另一方面,本示例性实施例可以从每组原始音频数据中分离出对应的声源的音频数据,为后期用户针对某一声源的音频数据进行调整提供了便捷,也使得音频数据的调整更加具有针对性和灵活性

[0074] 在一示例性实施例中,如图5所示,方法还可以包括以下步骤:

[0075] 步骤S510,提供至少一个音频控件,每个音频控件对应于一个声源;

[0076] 步骤S520,根据用户通过音频控件输入的音频调节指令,对音频控件对应的声源的音频数据进行播放调节。

[0077] 在本示例性实施例中,可以在终端设备中提供至少一个音频控件,例如滑块、选项等,每个音频控件对应一个声源,用户能够通过音频控件输入音频调节指令,该指令可以是调高音量、调低音量等基础指令,还可以是改变音色或加入混音等自定义指令等等。当终端设备接收到用户输入的音频调节指令后,可以根据该音频调节指令对音频控件对应的声源进行播放调节处理。在本示例性实施例中,用户可以通过多种方式输入上述音频调节指令,例如触控操作或语音指令等,本公开对此不做具体限定。

[0078] 在录制球赛的应用场景中,可以采用配置有四个麦克风的终端设备进行视频和音频的录制,分别采集声源为左侧声音、右侧声音、目标声音(球场内特定对象的声音,例如某一运动员的声音)以及解说声音的四组原始音频数据,通过本示例性实施例进行四种声源的音频数据的分离后,可以得到包含四条音轨的目标音频。当录制时距离目标较远时,可调高目标声音对应的音轨的音量;当需要关闭解说时,可将解说声音的音轨音量调为零,以关闭解说音轨。

[0079] 在录制辩论赛的应用场景中,使用终端设备的后置摄像头横屏进行视频录制,该终端设备配置有三个麦克风,可以录制三组原始音频数据。其中,第一麦克风对应右方辩论队,第二麦克风对应左方辩论队,第三麦克风对应主持人。通过本示例性实施例将各个声源的音频数据分离之后,可以根据需要对三个声源(右方辩论队、左方辩论队或主持人)的音频数据进行音量调整,从而消除由于现场距离差异导致的三方音量差异较大的问题,获得更好的临场效果的视频。

[0080] 在录制微录或小视频的应用场景中,使用配置四个麦克风的终端设备的前置摄像头竖屏录制视频,使第一麦克风对应录制用户的动作声音,例如滑板的声音、跑步的声音等,第二麦克风和第三麦克风对应录制背景环境音,第四麦克风对应录制用户声音,通过本

示范性实施例将各个声源的音频数据进行分离后,用户可以根据需要调整相应声源的音频数据的音量大小,例如增大用户声音或降低背景环境音等等。

[0081] 在一示范性实施例中,上述对音频控件对应的声源的音频数据进行播放调节,包括:

[0082] 确定调节时间范围;

[0083] 在调节时间范围内,对音频控件对应的声源的音频数据进行播放音量的调节。

[0084] 调节时间范围可以是音频数据中的任意一时间段,例如录制第一用户与第二用户说话的20分钟的音频数据中,第一用户说话的时间范围为前10分钟,第二用户说话的时间范围为后10分钟,则可以将前10分钟作为调节时间范围,也可以将后10分钟作为调节时间范围,还可以将其他任意时间段作为调节时间范围,具体的时间范围可以根据用户需求进行自定义设置。另外,调节时间范围可以基于用户输入的特定操作确定,例如用户可以拖动滑块控件来确定调节时间范围,或者输入起始时间点和结束时间点来确定调节时间范围等。在确定调节时间范围后,本示范性实施例可以对处于调节时间范围内的,音频控件对应的声源的音频数据进行播放音量的调节,从而提高音频数据调节的针对性。

[0085] 在一示范性实施例中,在进行播放调节之后,上述音频数据处理方法还可以包括:

[0086] 将各声源的音频数据合成为目标音频并播放。

[0087] 通过上述声源的音频数据的分离以及用户的自定义调节,最终可以生成基于用户需求的多样化的目标音频,进行播放。合成的目标音频可以作为新的音频文件存储在终端设备中,也可以覆盖原有的音频文件。用户在向其他用户发送音频文件时,可以选择发送原有的音频文件,或者经过调节后的目标音频文件等。

[0088] 本公开的示范性实施方式还提供一种音频数据处理装置。如图6所示,该音频数据处理装置600可以包括:音频数据获取模块610,用于获取由M个麦克风对N个声源进行声音采集所得到的M组原始音频数据,其中,每个麦克风对应采集一组原始音频数据,M和N均为大于等于2的正整数,且 $M \geq N$;音频数据分离模块620,用于基于M组原始音频数据进行计算,从M组原始音频数据中分离出至少一个声源的音频数据。

[0089] 在一示范性实施例中,音频数据分离模块包括:位姿系数获取单元,用于获取每个麦克风的位姿系数;方程构建单元,用于基于每个麦克风的位姿系数和M组原始音频数据,构建以每个声源的音频数据为变量的方程;方程求解单元,用于求解方程,得到至少一个声源的音频数据。

[0090] 在一示范性实施例中,M个麦克风为固定麦克风。

[0091] 在一示范性实施例中,当M个麦克风包括至少一个非固定麦克风时,音频数据分离模块包括:非固定麦克风声源确定单元,用于将非固定麦克风采集的原始音频数据,确定为位于非固定麦克风处的声源的音频数据。

[0092] 在一示范性实施例中,上述音频数据处理装置还包括:音频控件提供模块,用于提供至少一个音频控件,每个音频控件对应于一个声源;播放调节模块,用于根据用户通过音频控件输入的音频调节指令,对音频控件对应的声源的音频数据进行播放调节。

[0093] 在一示范性实施例中,播放调节模块包括:时间范围确定单元,用于确定调节时间范围;播放音量调节单元,用于在调节时间范围内,对音频控件对应的声源的音频数据进行播放音量的调节。

[0094] 在一示例性实施例中,上述音频数据处理装置还包括:播放模块,用于将各声源的音频数据合成为目标音频并播放。

[0095] 上述装置中各模块的具体细节在方法部分实施方式中已经详细说明,未披露的细节内容可以参见方法部分的实施方式内容,因而不再赘述。

[0096] 所属技术领域的技术人员能够理解,本公开的各个方面可以实现为系统、方法或程序产品。因此,本公开的各个方面可以具体实现为以下形式,即:完全的硬件实施方式、完全的软件实施方式(包括固件、微代码等),或硬件和软件方面结合的实施方式,这里可以统称为“电路”、“模块”或“系统”。

[0097] 本公开的示例性实施方式还提供了一种计算机可读存储介质,其上存储有能够实现本说明书上述方法的程序产品。在一些可能的实施方式中,本公开的各个方面还可以实现为一种程序产品的形式,其包括程序代码,当程序产品在终端设备上运行时,程序代码用于使终端设备执行本说明书上述“示例性方法”部分中描述的根据本公开各种示例性实施方式的步骤,例如可以执行图3、图4或图5中任意一个或多个步骤。

[0098] 本公开的示例性实施方式还提供了一种用于实现上述方法的程序产品,其可以采用便携式紧凑盘只读存储器(CD-ROM)并包括程序代码,并可以在终端设备,例如个人电脑上运行。然而,本公开的程序产品不限于此,在本文件中,可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0099] 程序产品可以采用一个或多个可读介质的任意组合。可读介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以为但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。

[0100] 计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了可读程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。可读信号介质还可以是可读存储介质以外的任何可读介质,该可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。

[0101] 可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于无线、有线、光缆、RF等等,或者上述的任意合适的组合。

[0102] 可以以一种或多种程序设计语言的任意组合来编写用于执行本公开操作的程序代码,程序设计语言包括面向对象的程序设计语言—诸如Java、C++等,还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、作为一个独立的软件包执行、部分在用户计算设备上部分在远程计算设备上执行、或者完全在远程计算设备或服务器上执行。在涉及远程计算设备的情形中,远程计算设备可以通过任意种类的网络,包括局域网(LAN)或广域网(WAN),连接到用户计算设备,或者,可以连接到外部计算设备(例如利用因特网服务提供商来通过因特网连接)。

[0103] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本公开的其

他实施方式。本公开旨在涵盖本公开的任何变型、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公开的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施方式仅被视为示例性的,本公开的真正范围和精神由权利要求指出。

[0104] 应当理解的是,本公开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。本公开的范围仅由所附的权利要求来限定。

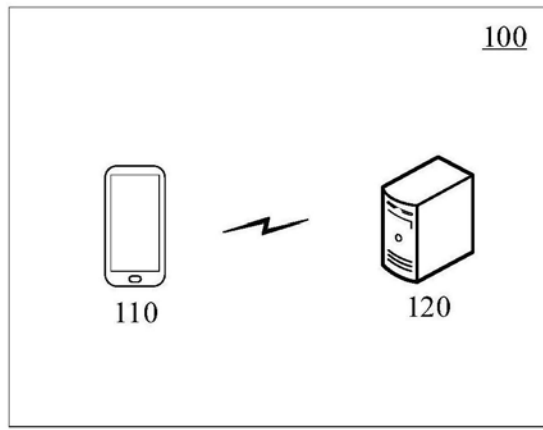


图1

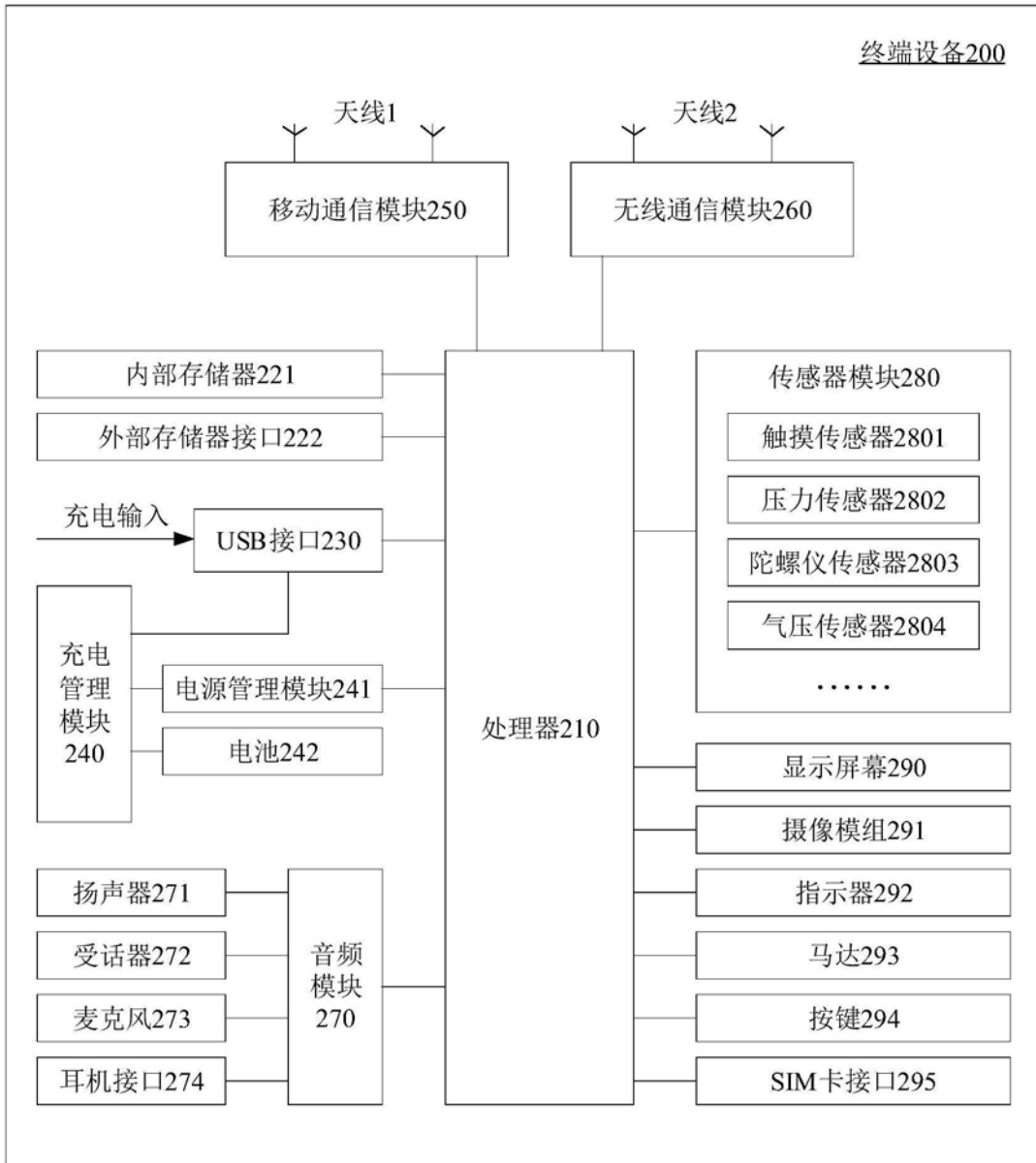


图2

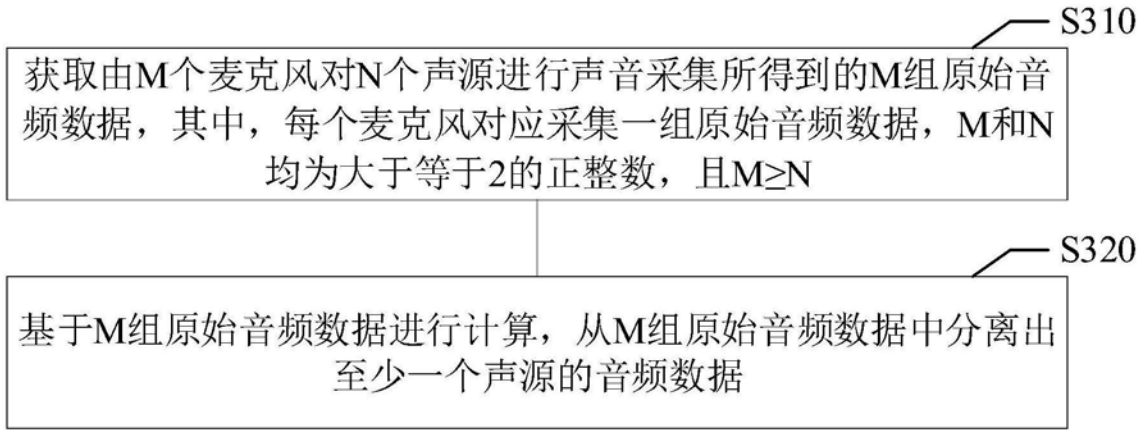


图3

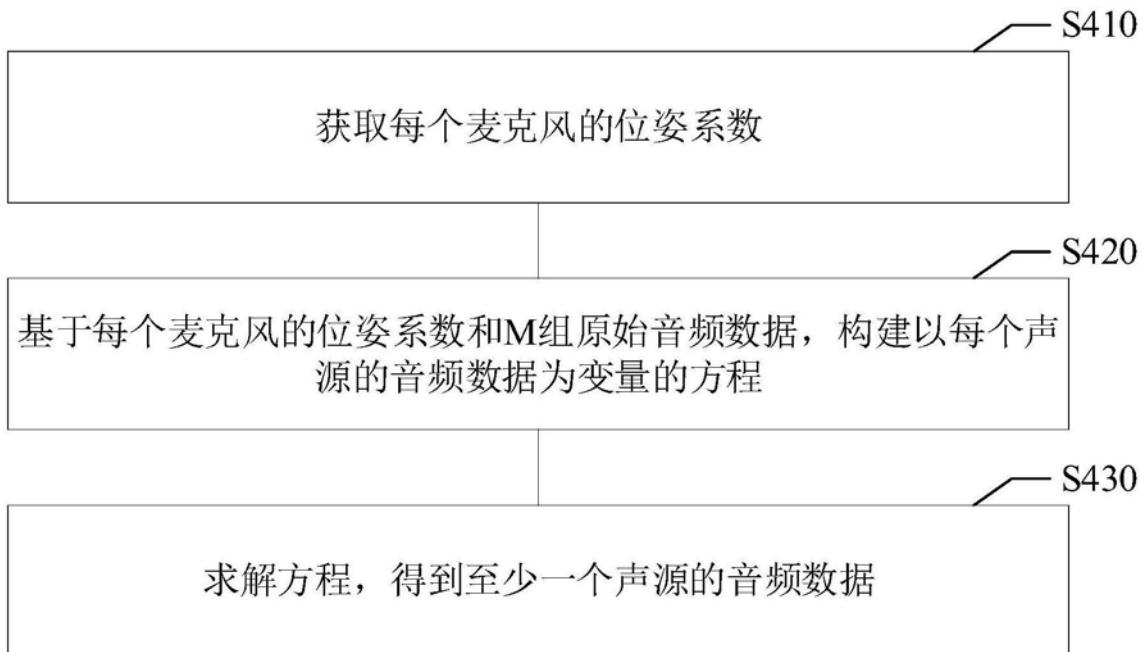


图4

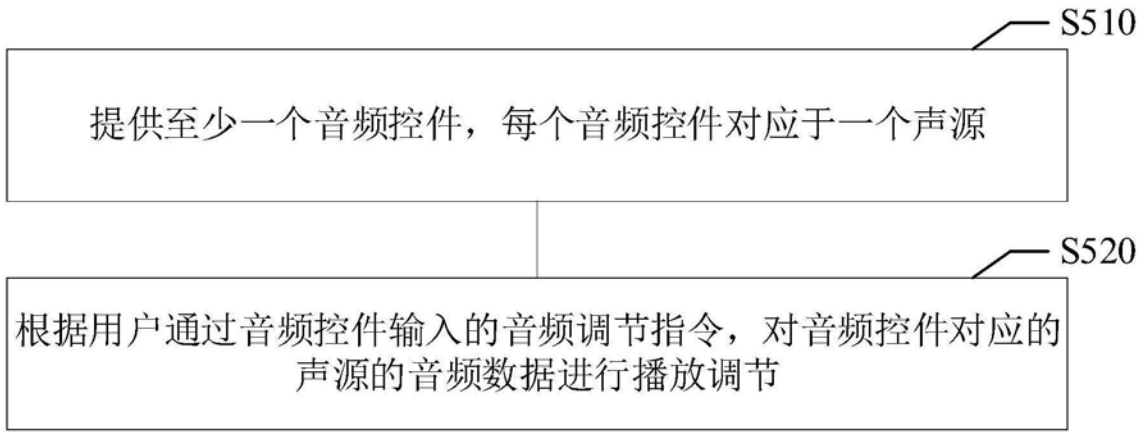


图5

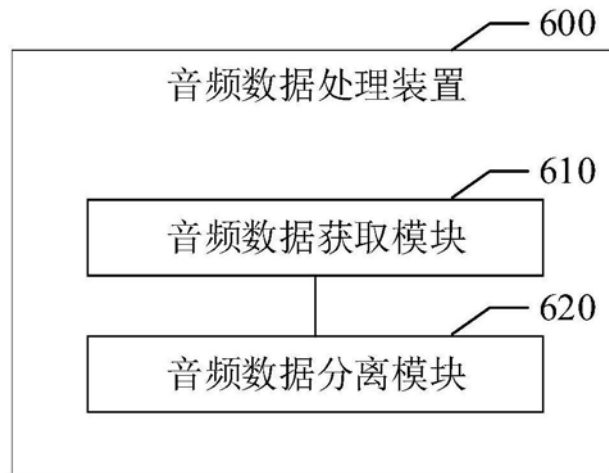


图6