



(12) 发明专利申请

(10) 申请公布号 CN 104658546 A

(43) 申请公布日 2015. 05. 27

(21) 申请号 201310586824. 7

(22) 申请日 2013. 11. 19

(71) 申请人 腾讯科技(深圳)有限公司

地址 518000 广东省深圳市福田区振兴路赛格科技园 2 栋东 403 室

(72) 发明人 王永鑫 李斌 罗程

(74) 专利代理机构 广州华进联合专利商标代理有限公司 44224

代理人 何平 邓云鹏

(51) Int. Cl.

G10L 21/0216(2013. 01)

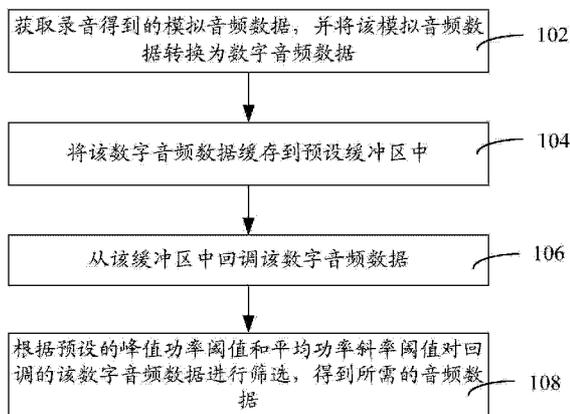
权利要求书2页 说明书6页 附图2页

(54) 发明名称

录音处理方法和装置

(57) 摘要

本发明涉及一种录音处理方法和装置。该方法包括以下步骤:获取录音得到的模拟音频数据,并将所述模拟音频数据转换为数字音频数据;将所述数字音频数据缓存到预设缓冲区中;从所述缓冲区中回调所述数字音频数据;根据预设的峰值功率阈值和平均功率斜率阈值对回调的所述数字音频数据进行筛选,得到所需的音频数据。上述录音处理方法和装置,将录音得到的模拟音频数据转换为数字音频数据后,缓存在缓冲区中,通过从缓冲区中回调数字音频数据,并根据峰值功率阈值和平均功率斜率阈值进行筛选,去掉了环境噪声的影响,得到所需的音频数据的质量较高。



1. 一种录音处理方法,包括以下步骤:
获取录音得到的模拟音频数据,并将所述模拟音频数据转换为数字音频数据;
将所述数字音频数据缓存到预设缓冲区中;
从所述缓冲区中回调所述数字音频数据;
根据预设的峰值功率阈值和平均功率斜率阈值对回调的所述数字音频数据进行筛选,得到所需的音频数据。
2. 根据权利要求1所述的录音处理方法,其特征在于,所述获取录音得到的模拟音频数据,并将所述模拟音频数据转换为数字音频数据的步骤包括:
通过音频队列接口进行录音得到模拟的音频数据,并通过所述音频队列接口将所述模拟音频数据转换为脉冲编码调制的音频数据。
3. 根据权利要求1所述的录音处理方法,其特征在于,在所述获取录音得到的模拟音频数据,并将所述模拟音频数据转换为数字音频数据的步骤之前,还包括:
根据所需录音的质量预先设置所述缓冲区的容量大小。
4. 根据权利要求3所述的录音处理方法,其特征在于,所述方法还包括:
通过调整所述缓冲区的容量大小来调节回调所述数字音频数据的频率。
5. 根据权利要求1所述的录音处理方法,其特征在于,所述从所述缓冲区中回调所述数字音频数据的步骤包括:
当所述数字音频数据将所述缓冲区填充完毕后,通过回调接口从该缓冲区中回调该数字音频数据。
6. 根据权利要求1所述的录音处理方法,其特征在于,所述根据预设的峰值功率阈值和平均功率斜率阈值对回调的所述数字音频数据进行筛选,得到所需的音频数据的步骤包括:
过滤峰值功率小于所述峰值功率阈值的回调的数字音频数据;
过滤平均功率斜率小于所述平均功率斜率阈值的回调的数字音频数据,得到所需的音频数据。
7. 一种录音处理装置,其特征在于,包括:
获取模块,用于获取录音得到的模拟音频数据,并将所述模拟音频数据转换为数字音频数据;
缓存模块,用于将所述数字音频数据缓存到预设缓冲区中;
回调模块,用于从所述缓冲区中回调所述数字音频数据;
筛选模块,用于根据预设的峰值功率阈值和平均功率斜率阈值对回调的所述数字音频数据进行筛选,得到所需的音频数据。
8. 根据权利要求7所述的录音处理装置,其特征在于,所述获取模块还用于通过音频队列接口进行录音得到模拟的音频数据,并通过所述音频队列接口将所述模拟音频数据转换为脉冲编码调制的音频数据。
9. 根据权利要求7所述的录音处理装置,其特征在于,所述装置还包括:
设置模块,用于在所述获取录音得到的模拟音频数据,并将所述模拟音频数据转换为数字音频数据之前,根据所需录音的质量预先设置所述缓冲区的容量大小。
10. 根据权利要求9所述的录音处理装置,其特征在于,所述装置还包括:

调整模块,用于通过调整所述缓冲区的容量大小来调节回调所述数字音频数据的频率。

11. 根据权利要求 7 所述的录音处理装置,其特征在于,所述回调模块还用于当所述数字音频数据将所述缓冲区填充完毕后,通过回调接口从该缓冲区中回调该数字音频数据。

12. 根据权利要求 7 所述的录音处理装置,其特征在于,所述筛选模块还用于过滤峰值功率小于所述峰值功率阈值的回调的数字音频数据;以及过滤平均功率斜率小于所述平均功率斜率阈值的回调的数字音频数据,得到所需的音频数据。

录音处理方法和装置

技术领域

[0001] 本发明涉及音频数据处理领域,特别是涉及一种录音处理方法和装置。

背景技术

[0002] 随着语音技术的发展,以及人们在生活中需要通过录音记录他人所说的,如培训所讲的课程、教学所讲的内容等等,录音设备应运而生。录音设备的种类很多,如录音笔、录音机、带录音功能的移动终端、带录音功能的计算机等。以带录音功能的移动终端为例,传统的移动终端开启录音功能进行录音时,移动终端并不会区分所需录制的声音与周围的杂声,将所需录制的声音及杂声一起记录下来,其录音因受到周围环境噪声的干扰,其录制的声音质量较差。

发明内容

[0003] 基于此,有必要针对传统录音设备录音的质量较差的问题,提供一种能提高录音质量的录音处理方法。

[0004] 此外,还有必要提供一种能提高录音质量的录音处理装置。

[0005] 一种录音处理方法,包括以下步骤:

[0006] 获取录音得到的模拟音频数据,并将所述模拟音频数据转换为数字音频数据;

[0007] 将所述数字音频数据缓存到预设缓冲区中;

[0008] 从所述缓冲区中回调所述数字音频数据;

[0009] 根据预设的峰值功率阈值和平均功率斜率阈值对回调的所述数字音频数据进行筛选,得到所需的音频数据。

[0010] 一种录音处理装置,包括:

[0011] 获取模块,用于获取录音得到的模拟音频数据,并将所述模拟音频数据转换为数字音频数据;

[0012] 缓存模块,用于将所述数字音频数据缓存到预设缓冲区中;

[0013] 回调模块,用于从所述缓冲区中回调所述数字音频数据;

[0014] 筛选模块,用于根据预设的峰值功率阈值和平均功率斜率阈值对回调的所述数字音频数据进行筛选,得到所需的音频数据。

[0015] 上述录音处理方法和装置,将录音得到的模拟音频数据转换为数字音频数据后,缓存在缓冲区中,通过从缓冲区中回调数字音频数据,并根据峰值功率阈值和平均功率斜率阈值进行筛选,去掉了环境噪声的影响,得到所需的音频数据的质量较高。

附图说明

[0016] 图1为一个实施例中录音处理方法的流程图;

[0017] 图2为根据预设的峰值功率阈值和平均功率斜率阈值对回调的该数字音频数据进行筛选,得到所需的音频数据的步骤的流程图;

- [0018] 图 3 为一个实施例中录音处理装置的结构框图；
[0019] 图 4 为另一个实施例中录音处理装置的结构框图；
[0020] 图 5 为终端的模块图。

具体实施方式

[0021] 为了使本发明的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本发明进行进一步详细说明。应当理解，此处所描述的具体实施例仅仅用以解释本发明，并不用于限定本发明。

[0022] 除非上下文另有特定清楚的描述，本发明中的元件和组件，数量既可以单个的形式存在，也可以多个的形式存在，本发明并不对此进行限定。本发明中的步骤虽然用标号进行了排列，但并不用于限定步骤的先后次序，除非明确说明了步骤的次序或者某步骤的执行需要其他步骤作为基础，否则步骤的相对次序是可以调整的。可以理解，本文中所使用的术语“和 / 或”涉及且涵盖相关联的所列项目中的一者或一者以上的任何和所有可能的组合。

[0023] 如图 1 所示，为一个实施例中录音处理方法的流程图。该录音处理方法可应用于具备录音功能的录音设备上，该录音设备可为计算机、移动终端等。该移动终端可为智能手机、平板电脑、个人数字助理等。该录音处理方法，包括以下步骤：

[0024] 步骤 102，获取录音得到的模拟音频数据，并将该模拟音频数据转换为数字音频数据。

[0025] 在一个实施例中，该获取录音得到的模拟音频数据，并将该模拟音频数据转换为数字音频数据的步骤包括：

[0026] 通过 AudioQueue（音频队列）接口进行录音得到模拟的音频数据，并通过该 AudioQueue 接口将该模拟音频数据转换为脉冲编码调制（PCM, Pulse Code Modulation）的音频数据。

[0027] 具体的，该 AudioQueue 接口是 iOS 操作系统提供的录音及播放功能的框架接口。通过设置 AudioQueue 接口，可以通过 AudioQueue 接口将原始的模拟音频数据转换为脉冲编码调制的音频数据。脉冲编码调制是对模拟信号先抽样，再对样值幅度量化，编码的过程，如此可将一个时间连续、取值连续的模拟信号变换成时间离散、取值离散的数字信号后在信道中传输。利用 iOS 操作系统自带的 AudioQueue 接口进行录音，操作方便。

[0028] 步骤 104，将该数字音频数据缓存到预设缓冲区中。

[0029] 采用 AudioQueue 接口进行录音，可预先设置 AudioQueue 接口的 Buffer（缓冲区）的容量大小。

[0030] 在一个实施例中，在该获取录音得到的模拟音频数据，并将该模拟音频数据转换为数字音频数据的步骤之前，还包括：根据所需录音的质量预先设置该缓冲区的容量大小。

[0031] 具体的，所需录音的质量的录音效果可通过采样率表示，采样率越高录音质量越高，音频失真越少。缓冲区的容量大小根据需要设定，本实施例中，缓冲区的容量大小设为 1024 字节。采样率为 16000，但不限于此。此外，缓冲区的容量大小受到音频处理的精确度和 CPU（Central Processing Unit, 中央处理器）的处理能力的限制。音频处理的精确度越高，缓冲区的容量大小越小；缓冲区的容量大小越小，则 CPU 处理能力要求越高。

[0032] 步骤 106, 从该缓冲区中回调该数字音频数据。

[0033] 具体的, 通过回调接口实时从该缓冲区中回调该数字音频数据。该回调接口需预先在操作系统(如 iOS 系统)中进行注册。例如回调接口 MyInputBufferHandler 注册代码如下:

[0034] AudioQueueNewInput (&mRecordFormat, MyInputBufferHandler, this/*userData*/, NULL/*run loop*/, NULL/*run loop mode*/, 0/*flags*/, &mQueue)。

[0035] 回调接口 MyInputBufferHandler 实现从 AudioQueue 接口的 Buffer 回调数字音频数据的代码为:

[0036] void AQRecorder::MyInputBufferHandler (void*inUserData, AudioQueueRef inAQ, AudioQueueBufferRef inBuffer, const AudioTimeStamp*inStartTime, UInt32inNumPackets, const AudioStreamPacketDescription*inPacketDesc)。

[0037] 在一个实施例中, 从该缓冲区中回调该数字音频数据的步骤包括: 当数字音频数据将缓冲区填充完毕后, 通过回调接口从该缓冲区中回调该数字音频数据。当填充完毕后, 从缓冲区中回调数据, 可有效控制回调处理的时机, 减少回调的次数, 降低系统的运行负担。

[0038] 在一个实施例中, 上述录音处理方法还包括: 通过调整该缓冲区的容量大小来调节回调该数字音频数据的频率。

[0039] 具体的, 缓冲区的容量大小越小, 回调该数字音频数据的频率越高, 缓冲区的容量大小越大, 回调该数字音频数据的频率越低。缓冲区的容量大小越小, 回调频率越高, 实时性越好, 运算强度随之增加。

[0040] 步骤 108, 根据预设的峰值功率阈值和平均功率斜率阈值对回调的该数字音频数据进行筛选, 得到所需的音频数据。

[0041] 具体的, 峰值功率 (Peak Power) 和平均功率 (Average Power) 斜率可由回调接口得到: AQRecorder*aqr=(AQRecorder*) inUserData。

[0042] 峰值功率: (aqr->levels[0]).mPeakPower;

[0043] 平均功率: (aqr->levels[0]).mAveragePower。

[0044] 峰值功率 (Peak Power) 阈值和平均功率 (Average Power) 斜率阈值可根据需要设定, 如峰值功率阈值可为 0.2、0.4、0.6 等, 平均功率斜率阈值可为 $\tan 10^\circ$ 、 $\tan 15^\circ$ 等。根据多种场景调优(如安静、嘈杂环境), 调优的峰值功率阈值为 0.2, 平均功率斜率阈值为 $\tan 10^\circ$, 但不限于此。

[0045] 在一个实施例中, 如图 2 所示, 该根据预设的峰值功率阈值和平均功率斜率阈值对回调的该数字音频数据进行筛选, 得到所需的音频数据的步骤包括:

[0046] 步骤 202, 过滤峰值功率小于该峰值功率阈值的回调的数字音频数据。

[0047] 步骤 204, 过滤平均功率斜率小于该平均功率斜率阈值的回调的数字音频数据, 得到所需的音频数据。

[0048] 具体的, 首先通过峰值功率过滤掉, 其次再以平均功率斜率过滤, 提高了过滤效率。

[0049] 上述录音处理方法, 将录音得到的模拟音频数据转换为数字音频数据后, 缓存在缓冲区中, 通过从缓冲区中回调数字音频数据, 并根据峰值功率阈值和平均功率斜率阈值

进行筛选,去掉了环境噪声的影响,得到所需的音频数据的质量较高。

[0050] 如图 3 所示,为一个实施例中录音处理装置的结构框图。该录音处理装置可应用于具备录音功能的录音设备上,该录音设备可为计算机、移动终端等。该移动终端可为智能手机、平板电脑、个人数字助理等。该录音处理装置,包括:获取模块 320、缓存模块 340、回调模块 360 和筛选模块 380。

[0051] 获取模块 320 用于获取录音得到的模拟音频数据,并将该模拟音频数据转换为数字音频数据。

[0052] 在一个实施例中,获取模块 320 还用于通过 AudioQueue 接口进行录音得到模拟的音频数据,并通过该 AudioQueue 接口将该模拟音频数据转换为脉冲编码调制(PCM, Pulse Code Modulation)的音频数据。

[0053] 具体的,该 AudioQueue 接口是 iOS 操作系统提供的录音及播放功能的框架接口。通过设置 AudioQueue 接口,可以通过 AudioQueue 接口将原始的模拟音频数据转换为脉冲编码调制的音频数据。脉冲编码调制是对模拟信号先抽样,再对样值幅度量化,编码的过程,如此可将一个时间连续、取值连续的模拟信号变换成时间离散、取值离散的数字信号后在信道中传输。利用 iOS 操作系统自带的 AudioQueue 接口进行录音,操作方便。

[0054] 缓存模块 340 用于将该数字音频数据缓存到预设缓冲区中。采用 AudioQueue 接口进行录音,可预先设置 AudioQueue 接口的 Buffer (缓冲区)的容量大小。

[0055] 回调模块 360 用于从该缓冲区中回调该数字音频数据。具体的,回调模块 360 通过回调接口实时从该缓冲区中回调该数字音频数据。该回调接口需预先在操作系统(如 iOS 系统)中进行注册,例如回调接口 MyInputBufferHandler。

[0056] 在一个实施例中,回调模块 360 还用于当数字音频数据将缓冲区填充完毕后,通过回调接口从该缓冲区中回调该数字音频数据。当填充完毕后,从缓冲区中回调数据,可有效控制回调处理的时机,减少回调的次数,降低系统的运行负担。

[0057] 筛选模块 380 用于根据预设的峰值功率阈值和平均功率斜率阈值对回调的该数字音频数据进行筛选,得到所需的音频数据。

[0058] 具体的,峰值功率(Peak Power)和平均功率(Average Power)斜率可由回调接口得到。峰值功率(Peak Power)阈值和平均功率(Average Power)斜率阈值可根据需要设定,如峰值功率阈值可为 0.2、0.4、0.6 等,平均功率斜率阈值可为 $\tan 10^\circ$ 、 $\tan 15^\circ$ 等。根据多种场景调优(如安静、嘈杂环境),调优的峰值功率阈值为 0.2,平均功率斜率阈值为 $\tan 10^\circ$,但不限于此。

[0059] 在一个实施例中,该筛选模块 380 还用于过滤峰值功率小于该峰值功率阈值的回调的数字音频数据;以及过滤平均功率斜率小于该平均功率斜率阈值的回调的数字音频数据,得到所需的音频数据。先通过峰值功率阈值筛选效率高。

[0060] 上述录音处理装置,将录音得到的模拟音频数据转换为数字音频数据后,缓存在缓冲区中,通过从缓冲区中回调数字音频数据,并根据峰值功率阈值和平均功率斜率阈值进行筛选,去掉了环境噪声的影响,得到所需的音频数据的质量较高。

[0061] 如图 4 所示,为另一个实施例中的录音处理装置的结构框图。该录音处理装置,除了包括获取模块 320、缓存模块 340、回调模块 360 和筛选模块 380,还包括设置模块 310 和调整模块 312。

[0062] 其中,设置模块 310 用于在该获取录音得到的模拟音频数据,并将该模拟音频数据转换为数字音频数据之前,根据所需录音的质量预先设置该缓冲区的容量大小。具体的,所需录音的质量的录音效果可通过采样率表示,采样率越高录音质量越高,音频失真越少。本实施例中,缓冲区的容量大小设为 1024 字节。采样率为 16000,但不限于此。此外,缓冲区的容量大小受到音频处理的精确度和 CPU(中央处理器)的处理能力的限制。音频处理的精确度越高,缓冲区的容量大小越小;缓冲区的容量大小越小,则 CPU 处理能力要求越高。

[0063] 调整模块 312 用于通过调整该缓冲区的容量大小来调节回调该数字音频数据的频率。

[0064] 具体的,缓冲区的容量大小越小,回调该数字音频数据的频率越高,缓冲区的容量大小越大,回调该数字音频数据的频率越低。缓冲区的容量大小越小,回调频率越高,实时性越好,运算强度随之增加。

[0065] 上述录音处理方法和装置除了应用于 iOS 操作系统,利用 AudioQueue 接口进行录音,并设置其 Buffer 进行处理外,还可应用于其他的操作系统,如 android、window phone 等,利用其他操作系统对应的录音接口录音,采用模数转换器对原始录音得到的音频数据进行模数转换等,然后采用相同的配置 Buffer 的机制实现对录音进行过滤处理,以提高录音的质量。

[0066] 图 5 为能实现本发明实施例的一个计算机系统 1000 的模块图。该计算机系统 1000 只是一个适用于本发明的计算机系统环境的示例,不能认为是提出了对本发明的使用范围的任何限制。计算机系统 1000 也不能解释为需要依赖于或具有图示的示例性的计算机系统 1000 中的一个或多个部件的组合。

[0067] 图 5 中示出的计算机系统 1000 是一个适合用于本发明的计算机系统的例子。具有不同子系统配置的其它架构也可以使用。例如有大众所熟知的带录音功能的台式机、笔记本、个人数字助理、智能电话、平板电脑、便携式媒体播放器、机顶盒等类似设备可以适用于本发明的一些实施例。但不限于以上所列举的设备。

[0068] 如图 5 所示,计算机系统 1000 包括处理器 1010、存储器 1020 和系统总线 1022。包括存储器 1020 和处理器 1010 在内的各种系统组件连接到系统总线 1022 上。处理器 1010 是一个用来通过计算机系统中基本的算术和逻辑运算来执行计算机程序指令的硬件。存储器 1020 是一个用于临时或永久性存储计算程序或数据(例如,程序状态信息)的物理设备。系统总线 1020 可以为以下几种类型的总线结构中的任意一种,包括存储器总线或存储控制器、外设总线和局部总线。处理器 1010 和存储器 1020 可以通过系统总线 1022 进行数据通信。其中存储器 1020 包括只读存储器(ROM)或闪存(图中都未示出),以及随机存取存储器(RAM),RAM 通常是指加载了操作系统和应用程序的主存储器。

[0069] 计算机系统 1000 还包括显示接口 1030(例如,图形处理单元)、显示设备 1040(例如,液晶显示器)、音频接口 1050(例如,声卡)以及音频设备 1060(例如,扬声器)。显示设备 1040 和音频设备 1060 是用于体验多媒体内容的媒体设备。

[0070] 计算机系统 1000 一般包括一个存储设备 1070。存储设备 1070 可以从多种计算机可读介质中选择,计算机可读介质是指可以通过计算机系统 1000 访问的任何可利用的介质,包括移动的和固定的两种介质。例如,计算机可读介质包括但不限于,闪速存储器(微型 SD 卡),CD-ROM,数字通用光盘(DVD)或其它光盘存储、磁带盒、磁带、磁盘存储或其它磁存储

设备,或者可用于存储所需信息并可由计算机系统 1000 访问的任何其它介质。

[0071] 计算机系统 1000 还包括输入装置 1080 和输入接口 1090 (例如, I/O 控制器)。用户可以通过输入装置 1080,如键盘、鼠标、显示装置 1040 上的触摸面板设备,输入指令和信息到计算机系统 1000 中。输入装置 1080 通常是通过输入接口 1090 连接到系统总线 1022 上的,但也可以通过其它接口或总线结构相连接,如通用串行总线(USB)。

[0072] 计算机系统 1000 可在网络环境中与一个或者多个网络设备进行逻辑连接。网络设备可以是个人电脑、服务器、路由器、智能电话、平板电脑或者其它公共网络节点。计算机系统 1000 通过局域网(LAN)接口 1100 或者移动通信单元 1110 与网络设备相连接。局域网(LAN)是指在有限区域内,例如家庭、学校、计算机实验室、或者使用网络媒体的办公楼,互联组成的计算机网络。WiFi 和双绞线布线以太网是最常用的构建局域网的两种技术。WiFi 是一种能使计算机系统 1000 间交换数据或通过无线电波连接到无线网络的技术。移动通信单元 1110 能在一个广阔的地理区域内移动的同时通过无线电通信线路接听和拨打电话。除了通话以外,移动通信单元 1110 也支持在提供移动数据服务的 2G,3G 或 4G 蜂窝通信系统中进行互联网访问。

[0073] 应当指出的是,其它包括比计算机系统 1000 更多或更少的子系统的计算机系统也能适用于发明。例如,计算机系统 1000 可以包括能在短距离内交换数据的蓝牙单元,用于照相的图像传感器,以及用于测量加速度的加速计。

[0074] 如上面详细描述,适用于本发明的计算机系统 1000 能执行录音处理方法的指定操作。计算机系统 1000 通过处理器 1010 运行在计算机可读介质中的软件指令的形式来执行这些操作。这些软件指令可以从存储设备 1070 或者通过局域网接口 1100 从另一设备读入到存储器 1020 中。存储在存储器 1020 中的软件指令使得处理器 1010 执行上述的录音处理方法。此外,通过硬件电路或者硬件电路结合软件指令也能同样实现本发明。因此,实现本发明并不限于任何特定硬件电路和软件的组合。

[0075] 以上所述实施例仅表达了本发明的几种实施方式,其描述较为具体和详细,但并不能因此而理解为对本发明专利范围的限制。应当指出的是,对于本领域的普通技术人员来说,在不脱离本发明构思的前提下,还可以做出若干变形和改进,这些都属于本发明的保护范围。因此,本发明的保护范围应以所附权利要求为准。

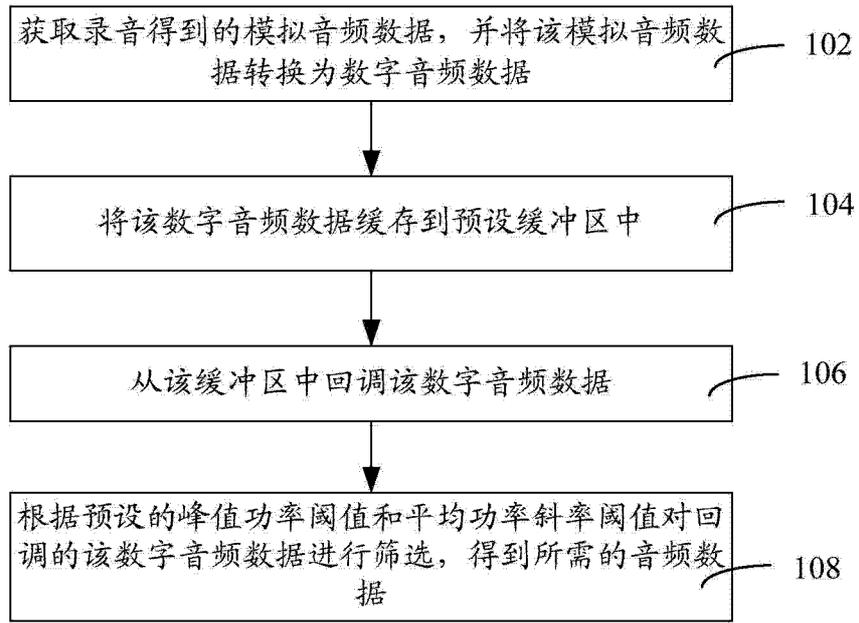


图 1

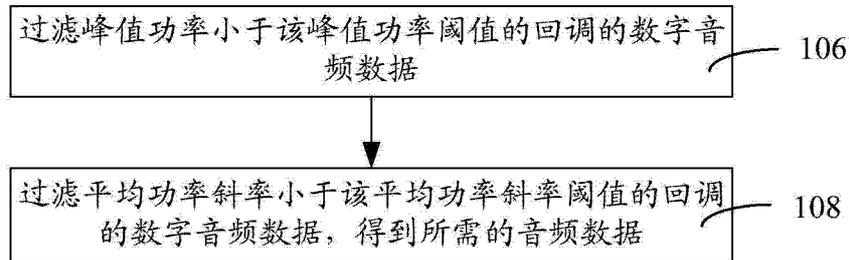


图 2

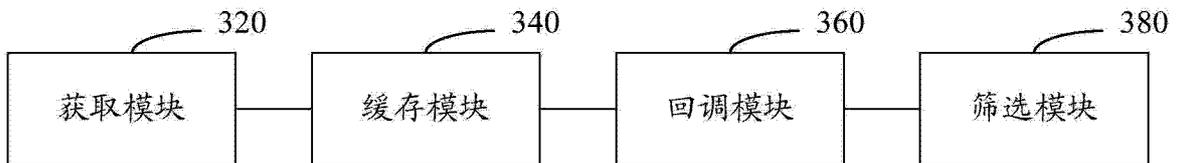


图 3

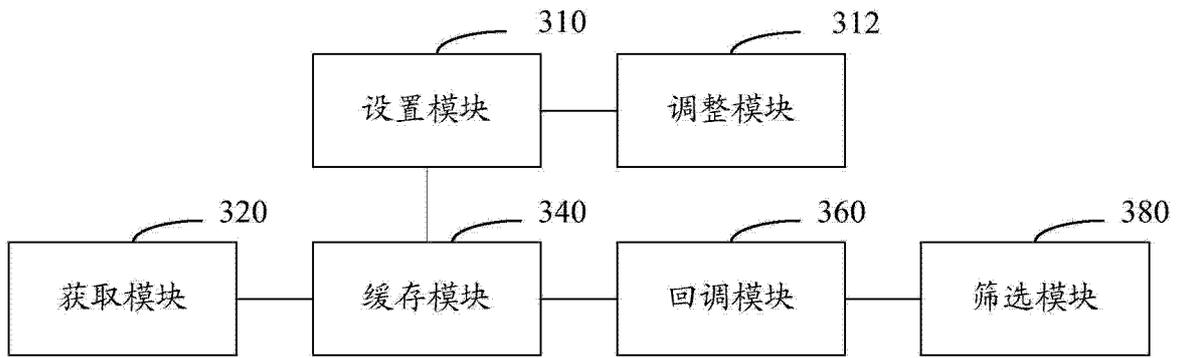


图 4

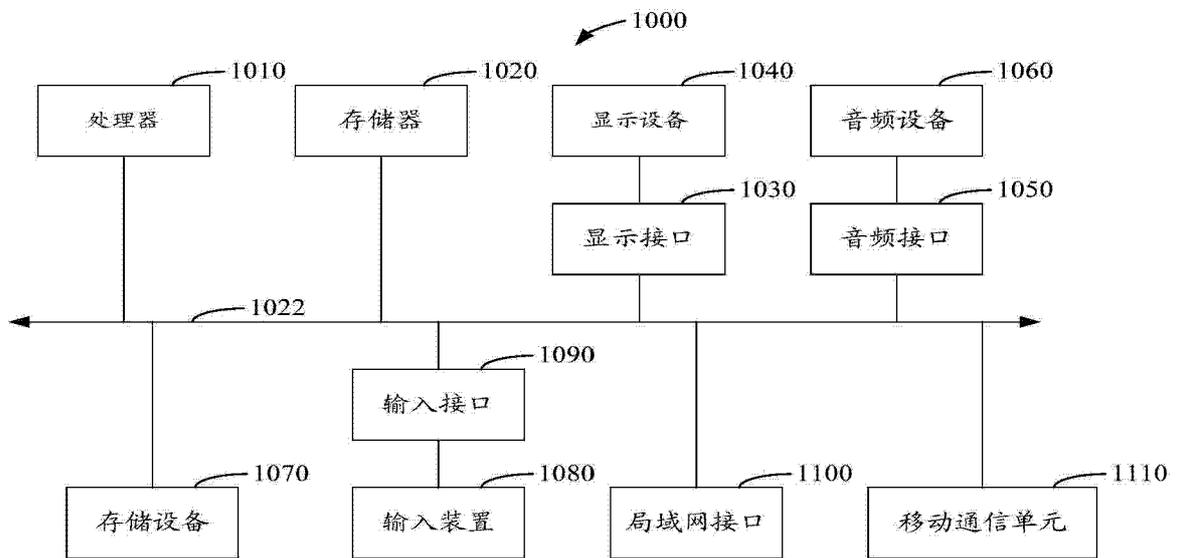


图 5