



(12)发明专利

(10)授权公告号 CN 107105367 B

(45)授权公告日 2020.07.10

(21)申请号 201710372132.0

审查员 胡向莉

(22)申请日 2017.05.24

(65)同一申请的已公布的文献号

申请公布号 CN 107105367 A

(43)申请公布日 2017.08.29

(73)专利权人 维沃移动通信有限公司

地址 523860 广东省东莞市长安镇乌沙步  
步高大道283号

(72)发明人 李凤亮 屠光明

(74)专利代理机构 北京银龙知识产权代理有限

公司 11243

代理人 许静 刘伟

(51)Int.Cl.

H04R 3/04(2006.01)

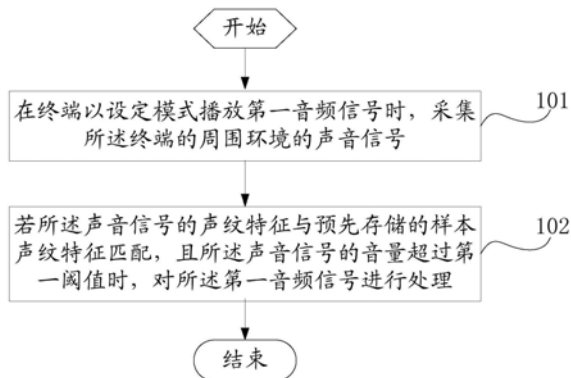
权利要求书2页 说明书14页 附图9页

(54)发明名称

一种音频信号处理方法及终端

(57)摘要

本发明提供一种音频信号处理方法及终端。所述方法包括:在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的声纹特征;若所述声纹特征的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,对所述第一音频信号进行处理,使得用户能够听到所述声音信号,以解决用户在使用终端播放音频时不能够听到他人向用户说话的问题。



1. 一种音频信号处理方法,其特征在于,应用于终端,包括:

在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的声纹信号;

若所述声纹信号的声纹特征与预先存储的样本声纹特征匹配,且所述声纹信号的音量值超过第一阈值时,对所述第一音频信号进行处理,并输出声音信号;

所述采集所述声纹信号,并输出所述声音信号,包括:

若所述声纹信号的声纹特征与预先存储的样本声纹特征匹配,且所述声纹信号的音量值超过第一阈值时,确定所述声纹信号的声源方向;

在采集所述声纹信号时,根据所述声源方向对所述声纹信号进行降噪处理,得到经过降噪处理的声纹信号;

采集所述终端的周围环境的声纹信号的步骤之前,所述方法还包括:

在用户通过预设的应用与另一方用户进行语音通信操作时,采集所述另一方用户的语音;所述语音通信操作包括通话操作或语音播放;

从采集的另一方用户的语音中提取声纹特征;

将提取出的所述声纹特征存储为样本声纹特征。

2. 根据权利要求1所述的方法,其特征在于,对所述第一音频信号进行处理,使得用户能够听到所述声音信号,包括:

降低所述第一音频信号的输出音量、降低所述第一音频信号的输出音量,同时输出所述声音信号、

暂停所述第一音频信号的播放、暂停所述第一音频信号的播放,同时输出所述声音信号、

在播放所述第一音频信号时,同时输出所述声音信号中的一种或者多种处理方法。

3. 根据权利要求2所述的方法,其特征在于,降低播放所述第一音频信号的输出音量,包括:

减小终端的播放音量,使得所述第一音频信号的输出音量降低;

或者,将所述终端的静音,使得所述第一音频信号的输出音量降低为零;

或者,利用与所述第一音频信号相位相抵消的第二音频信号对第一音频信号进行混音,使得所述第一音频信号的输出音量降低。

4. 一种终端,其特征在于,包括:

声音信号采集模块:用于在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的声纹信号;

第一音频信号处理模块:用于若所述声纹信号的声纹特征与预先存储的样本声纹特征匹配,且所述声纹信号的音量值超过第一阈值时,对所述第一音频信号进行处理;

所述第一音频信号处理模块包括第二处理单元,用于降低所述第一音频信号的输出音量,同时输出所述声音信号;

所述第二处理单元包括:

声源方向确定子单元:用于若所述声纹信号的声纹特征与预先存储的样本声纹特征匹配,且所述声纹信号的音量值超过第一阈值时,确定所述声纹信号的声源方向;

第一降噪子单元:用于在采集所述声纹信号时,根据所述声源方向对所述声纹信号进行降噪处理,得到经过降噪处理的声纹信号;

第一播放子单元:用于通过所述终端的音频输出设备播放所述经过降噪处理的声音信号;

所述终端还包括:

语音采集模块:用于在用户通过预设的应用与另一方用户进行语音通信操作时,采集所述另一方用户的语音;所述语音通信操作包括通话操作或语音播放;

声纹提取模块:用于从采集的另一方用户的语音中提取声纹特征;

声纹存储模块:用于将提取出的所述声纹特征存储为样本声纹特征。

5. 根据权利要求4所述的终端,其特征在于,所述第一音频信号处理模块还包括下述至少一种:

第一处理单元:用于降低所述第一音频信号的输出音量;

第三处理单元:用于暂停所述第一音频信号的播放;

第四处理单元:用于暂停所述第一音频信号的播放,同时输出所述声音信号;

第五处理单元:用于在播放所述第一音频信号时,同时输出所述声音信号。

6. 根据权利要求5所述的终端,其特征在于,如果所述第一音频信号处理模块包括所述第一处理单元,所述第一处理单元包括:

播放音量降低子单元:用于减小终端的播放音量,使得所述第一音频信号的输出音量降低;

或者,静音子单元:用于将所述终端的静音,使得所述第一音频信号的输出音量降低为零;

或者,混音子单元:用于利用与所述第一音频信号相位相抵消的第二音频信号对第一音频信号进行混音,使得所述第一音频信号的输出音量降低;

如果所述第一音频信号处理模块包括所述第二处理单元,所述第二处理单元包括:所述播放音量降低子单元、或者所述静音子单元、或者所述混音子单元。

7. 根据权利要求5所述的终端,其特征在于,如果所述第一音频信号处理模块包括所述第四处理单元,所述第四处理单元包括:所述声源方向确定子单元、所述第一降噪子单元、所述第一播放子单元。

8. 根据权利要求5所述的终端,其特征在于,

所述第二处理单元还包括:

录音子单元:用于将采集到的声音信号转换成录音文件;

第二降噪子单元:用于对所述录音文件进行降噪处理;

第二播放子单元:用于通过所述终端的音频输出设备播放经过降噪处理后的所述录音文件;

当所述第一音频信号处理模块包括所述第四处理单元时,所述第四处理单元包括:所述录音子单元、所述第二降噪子单元、所述第二播放子单元。

9. 一种终端,其特征在于,包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述程序时实现如权利要求1~3中任一项所述的音频信号处理方法中的步骤。

## 一种音频信号处理方法及终端

### 技术领域

[0001] 本发明涉及电子及计算机技术领域,尤其涉及一种音频信号处理方法及终端。

### 背景技术

[0002] 随着电子技术的发展,用户通过终端收听音频的方式也越来越多。除了直接使用终端自带的播放部件播放音频之外,用户可通过与终端电连接的音频输出设备收听音频,例如,将终端电连接耳机或音箱后,就可通过耳机或音箱收听音频。

[0003] 现有技术中,用户针对其所处的环境以及终端的模式,会将音频的输出音量调整到一定程度,使得自己能够清晰地收听到音频。例如,用户使用耳机收听音乐时,有时会将音乐的输出音量调得比较大;再如,用户将音箱等设备接入终端,使得音频的输出音量能够大于终端最大音量,从而提高音频收听效果。这样若他人在音频输出的过程中跟用户讲话,用户常常难以察觉或难以听到。

### 发明内容

[0004] 本发明实施例提供一种音频信号处理方法及终端,以解决用户在使用终端播放音频时,播放音频的声音干扰,不能够听清楚他人向用户说话的问题。

[0005] 第一方面,本发明实施例提供一种音频信号处理方法,应用于终端,包括:

[0006] 在终端以设定模式播放第一音频信号时,采集所述终端的周围环境的声纹信号;

[0007] 若所述声纹信号的声纹特征与预先存储的样本声纹特征匹配,且所述声纹信号的音量值超过第一阈值时,对所述第一音频信号进行处理,使得用户能够听到所述声纹信号。

[0008] 第二方面,本发明实施例提供一种终端,包括:

[0009] 声纹信号采集模块:用于在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的声纹信号;

[0010] 第一音频信号处理模块:用于若所述声纹信号的声纹特征与预先存储的样本声纹特征匹配,且所述声纹信号的音量值超过第一阈值时,对所述第一音频信号进行处理,使得用户能够听到所述声纹信号。

[0011] 第三方面,本发明实施例提供一种终端,包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述程序时实现如本发明任意一项实施例所述的音频信号处理方法中的步骤。

[0012] 本发明实施例所提供的音频信号处理方法及终端,在播放音频时,能够监听终端周围环境中的声纹信息,并检测声纹信息的声纹特征是否与预先存储的样本声纹特征匹配,并在匹配且所述声纹信息音量高于第一阈值的情况下,对第一音频信号进行处理,使得用户能够收听到声纹信号,从而,用户在使用终端收听音频时,不会因为音频信号遮盖周围他人向该用户说话的声音而使用户难以察觉到他人向自己讲话,从而他人与用户的语言交流不会受到音频播放事件的影响。

## 附图说明

[0013] 为了更清楚地说明本发明实施例的技术方案,下面将对本发明实施例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性的前提下,还可以根据这些附图获得其他的附图。

[0014] 图1为本发明一种实施例的音频信号处理方法流程示意图;

[0015] 图2为本发明另一种实施例的音频信号处理方法流程示意图;

[0016] 图3为本发明另一种实施例的音频信号处理方法流程示意图;

[0017] 图4为本发明另一种实施例的音频信号处理方法流程示意图;

[0018] 图5为本发明另一种实施例的音频信号处理方法流程示意图

[0019] 图6为本发明一种实施例的终端结构示意框图;

[0020] 图7为本发明另一种实施例的终端结构示意框图;

[0021] 图8为本发明另一种实施例的终端结构示意框图;

[0022] 图9为本发明另一种实施例的终端结构示意框图;

[0023] 图10为本发明另一种实施例的终端结构示意图。

## 具体实施方式

[0024] 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。

[0025] 本发明实施例首先提供一种音频信号处理方法,应用于终端,如图1所示,包括:

[0026] 步骤101:在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的声音信号;

[0027] 在本发明具体实施例中,所述预设方式包括,通过与终端电连接的音频输出设备播放音频的模式、和/或终端以高于一定阈值的音量播放音频的模式;更具体的,所述音频输出设备包括耳机、音箱等。

[0028] 在本发明具体实施例中,采集所述终端的周围环境的的声音信号的操作,通过终端麦克风或与终端电连接的音频输出设备的麦克风进行采集,例如,通过耳机麦克风进行采集。

[0029] 在本发明具体实施例中,所述第一音频信号为纯音频信号或包含音频信号的文件所播放的信号。

[0030] 在本发明一种实施例中,所述声音信号特指语音信号。

[0031] 步骤102:若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,对所述第一音频信号进行处理,使得用户能够听到所述声音信号。

[0032] 在本发明具体实施例中,对所述第一音频信号进行处理,包括降低所述第一音频信号的输出音量,和/或将所述声音信号夹杂在所述第一音频信号中一起输出,和/或暂停所述第一音频信号的播放。

[0033] 在本发明具体实施例中,由于用于采集终端的周围环境的声信号的声音采集设备灵敏度会因为设备不同而有所不同,某些声音采集设备能够采集到的声信号已经是音量超过第一阈值的音量了,在这种情况下,等同于所述声信号的音量超过第一阈值的情况。

[0034] 在本发明具体实施例中,所述第一阈值为不小于0的任意音量值,可通过具体试验进行测定,或可收集一般环境下,普通人对话的多个音量数据,根据多个音量数据确定所述第一阈值,从而能够排除其他人未跟用户讲话但声纹匹配的情况。

[0035] 本发明实施例中,上述终端可以为任何具备人体特征采集组件和屏幕的终端,例如:手机、平板电脑(Tablet Personal Computer)、膝上型电脑(Laptop Computer)、个人数字助理(personal digital assistant,简称PDA)、移动上网装置(Mobile Internet Device,MID)或可穿戴式设备(Wearable Device)等。

[0036] 本发明实施例所提供的音频信号处理方法,在播放音频时,能够监听终端周围环境中的声信号,并检测声信号的声纹特征是否与预先存储的样本声纹特征匹配,并在匹配且所述声信号音量高于第一阈值的情况下,对第一音频信号进行处理,使得用户能够收听到声信号,从而,用户在使用终端收听音频时,不会因为音频信号遮盖周围他人向该用户说话的声音而使用户难以察觉到他人向自己讲话,从而他人与用户的语言交流不会受到音频播放事件的影响。

[0037] 在本发明一些实施例中,对所述第一音频信号进行处理,使得用户能够听到所述声信号,包括:

[0038] 降低所述第一音频信号的输出音量、降低所述第一音频信号的输出音量,同时输出所述声信号、

[0039] 暂停所述第一音频信号的播放、暂停所述第一音频信号的播放,同时输出所述声信号、

[0040] 在播放所述第一音频信号时,同时输出所述声信号中的一种或者多种处理方法。

[0041] 在本发明一种实施例中,所述方法包括如图2所示的步骤:

[0042] 步骤201:在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的声信号。

[0043] 在本发明具体实施例中,所述预设方式包括,通过与终端电连接的音频输出设备播放音频的模式、和/或终端以高于一定阈值的音量播放音频的模式;更具体的,所述音频输出设备包括耳机、音箱等。

[0044] 在本发明具体实施例中,采集所述终端的周围环境的声信号的操作,通过终端麦克风或与终端电连接的音频输出设备的麦克风进行采集,例如,通过耳机麦克风进行采集。

[0045] 在本发明具体实施例中,所述第一音频信号为纯音频信号或包含音频信号的文件所播放的信号。

[0046] 在本发明一种实施例中,所述声信号特指语音信号。

[0047] 步骤202:若所述声信号的声纹特征与预先存储的样本声纹特征匹配,且所述声信号的音量值超过第一阈值时,降低所述第一音频信号的输出音量,使得用户能够听到

所述声音信号。在本发明具体实施例中,所述第一音频信号的输出音量,指第一音频信号最终播放出来的音量,该音量会受到终端音量、与终端电连接的声音输出设备的音量、音频文件的属性等因素的影响。

[0048] 在一种实施例中,降低所述第一音频信号的输出音量的步骤包括:

[0049] 减小终端的播放音量,使得所述第一音频信号的输出音量降低;在本发明具体实施例中,所述终端的播放音量,指终端通过其自带的音频播放部件播放音频的音量,在具体操作中可通过音量调节按钮以硬件或软件方式进行降低;

[0050] 或者,将所述终端的静音,使得所述第一音频信号的输出音量降低为零;在本发明具体实施例中,将所述终端静音的步骤,包括:降低终端音量直至终端音量为零;或,通过静音控制模块,将终端调整为静音模式;

[0051] 或者,利用与所述第一音频信号相位相抵消的第二音频信号对第一音频信号进行混音,使得所述第一音频信号的输出音量降低。在具体实施例中,利用与第一音频信号相位相抵消的第二音频信号对第一音频信号进行混音时,会削减或消除第一音频信号的输出音量,达到第一音频信号的输出音量降低的效果。

[0052] 在本发明一种实施例中,所述方法包括如图3所示的步骤:

[0053] 步骤301:在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的聲音信号;

[0054] 在本发明具体实施例中,所述预设方式包括,通过与终端电连接的音频输出设备播放音频的模式、和/或终端以高于一定阈值的音量播放音频的模式;更具体的,所述音频输出设备包括耳机、音箱等。

[0055] 在本发明具体实施例中,采集所述终端的周围环境的聲音信号的操作,通过终端麦克风或与终端电连接的音频输出设备的麦克风进行采集,例如,通过耳机麦克风进行采集。

[0056] 在本发明具体实施例中,所述第一音频信号为纯音频信号或包含音频信号的文件所播放的信号。

[0057] 在本发明一种实施例中,所述声音信号特指语音信号。

[0058] 步骤302:若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,继续采集所述声音信号,并在播放所述第一音频信号时,输出所述声音信号,使得用户能够听到所述声音信号。

[0059] 在本发明另一种实施例中,所述方法包括:

[0060] 在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的聲音信号;

[0061] 若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,继续采集所述声音信号,在播放所述第一音频信号的同时,输出所述声音信号,即将声音信号夹杂在所述第一音频信号中一起输出,使得用户能够听到所述声音信号。

[0062] 上述实施例中,若用户使用终端播放第一音频信号时,第一音频信号的输出音量较小,因而有时将从外界环境中采集的voice信号夹杂在第一音频信号中一起输出时,用户也能够较清楚地接收到voice信号,并获取voice信号中的内容。在另一些情况下,用户使用终端播放第一音频信号时,第一音频信号的输出音量较大,即使将从外界环境中采集到的声

音信号夹杂在第一音频信号中一起输出,也不能够使得用户清楚地收听到声音信号,但是用户能够意识到有他人向自己讲话,从而可以手动调低第一音频信号的输出音量或者摘下耳机,使得用户能够听清楚他人向自己讲话。

[0063] 在本发明另一种实施例中,所述方法包括:

[0064] 在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的聲音信号;

[0065] 若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,暂停所述第一音频信号的播放,使得用户无需摘下耳机或者手动静音或者降低其他音频输出设备的音量,就能够听到所述声音信号。在一些情况下,若用户在收听第一音频信号的同时接收他人向该用户讲话的内容,可能会因为漏听第一音频信号的内容而引起用户的不便,上述实施例中,能够在用户收听所述声音信号时,暂停第一音频信号的播放,使得用户不会因为有人对该用户讲话而漏听部分第一音频信号的内容。

[0066] 在本发明一种实施例中,所述方法包括:

[0067] 在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的聲音信号;

[0068] 若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,继续采集所述声音信号,并暂停所述第一音频信号的播放,同时输出所述声音信号。

[0069] 在第一音频信号暂停状况下,相比第一音频信号播放状态下能够更清楚地听到他人讲话的内容。另外在用户佩戴耳机的情况下,由于某些耳机对耳部的密封性比较好,让用户在佩戴耳机的同时收听他人说话的内容,用户可能难以听清楚,而将声音信号采集后,通过声音输出设备输出,能够提高用户收听声音信号的清晰度。在具体实施例中,输出所述声音信号,可通过终端播放部件(例如听筒或扬声器)进行输出,或通过耳机听筒进行输出,或通过音箱进行播放。

[0070] 在本发明一种实施例中,所述方法包括:

[0071] 在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的聲音信号;

[0072] 若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,继续采集所述声音信号,并降低所述第一音频信号的输出音量,同时输出所述声音信号。

[0073] 上述实施例中,能够采集所述声音信号,并在向播放采集到的所述声音信号时,降低第一音频信号的输出音量,使得用户不仅能够察觉到有人对用户讲话,还能够较为清楚地听到他人对用户讲话的内容,从而用户无需摘下耳机、或关闭其他音频输出设备、或手动调低音量。

[0074] 在本发明一种实施例中,所述方法包括图4所示的步骤:

[0075] 步骤401:在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的聲音信号;

[0076] 在本发明具体实施例中,所述预设方式包括,通过与终端电连接的音频输出设备播放音频的模式、和/或终端以高于一定阈值的音量播放音频的模式;更具体的,所述音频输出设备包括耳机、音箱等。

[0077] 在本发明具体实施例中,采集所述终端的周围环境的聲音信号的操作,通过多个麦克风进行采集,例如耳机的麦克风阵列,终端自带的多麦克风装置等。



[0078] 在本发明具体实施例中,所述第一音频信号为纯音频信号或包含音频信号的文件所播放的信号。

[0079] 在本发明一种实施例中,所述声音信号特指语音信号。

[0080] 步骤402:若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,确定所述声音信号的声源方向。通过多个麦克风装置进行声源信号的确定。

[0081] 步骤403:继续采集声音信号,在采集所述声音信号时,根据所述声源方向对所述声音信号进行降噪处理,得到经过降噪处理的声音信号。根据判断的声音信号的声源方向,利用多个麦克风,指向声源方向的麦克风进行所述声音信号的采集,从而采集到的信号中,夹杂的噪音信号较少,能够达到降噪的效果。

[0082] 步骤404:通过所述终端的音频输出设备播放降经过降噪处理的所述声音信号。

[0083] 在本发明一种实施例中,所述方法包括:

[0084] 在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的聲音信号;

[0085] 若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,确定所述声音信号的声源方向;

[0086] 继续采集所述声音信号,在采集所述声音信号时,根据所述声源方向对所述声音信号进行降噪处理,得到经过降噪处理的声音信号;

[0087] 降低所述第一音频信号的输出音量,并通过所述终端的音频输出设备播放降经过降噪处理的所述声音信号。

[0088] 上述实施例中,能够采集所述声音信号,并在向播放采集到的所述声音信号时,降低第一音频信号的输出音量,使得用户不仅能够察觉到有他人对用户讲话,还能够较为清楚地听到他人对用户讲话的内容,从而用户无需摘下耳机、或关闭其他音频输出设备、或手动调低音量。

[0089] 在本发明一种实施例中,所述方法包括:

[0090] 在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的聲音信号;

[0091] 若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,确定所述声音信号的声源方向;

[0092] 继续采集所述声音信号,在采集所述声音信号时,根据所述声源方向对所述声音信号进行降噪处理,得到经过降噪处理的声音信号;

[0093] 暂停所述第一音频信号的播放,并通过所述终端的音频输出设备播放降经过降噪处理的所述声音信号。

[0094] 上述实施例中,能够采集所述声音信号,并在向播放采集到的所述声音信号时,暂停第一音频信号的播放,使得用户不仅能够察觉到有他人对用户讲话,还能够较为清楚地听到他人对用户讲话的内容,从而用户无需摘下耳机、或关闭其他音频输出设备、或手动调低音量。

[0095] 在本发明另一种实施例中,所述方法包括:

[0096] 在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的聲音信号;

[0097] 若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,确定所述声音信号的声源方向;

[0098] 继续采集所述声音信号,在采集所述声音信号时,根据所述声源方向对所述声音信号进行降噪处理,得到经过降噪处理的声音信号;

[0099] 降低所述第一音频信号的输出音量或暂停所述第一音频信号的播放,并在第一音频信号播放时,通过所述终端的音频输出设备播放降经过降噪处理的所述声音信号。使得用户不仅能够察觉到有他人对用户讲话,还能够较为清楚地听到他人对用户讲话的内容,从而用户无需摘下耳机、或关闭其他音频输出设备、或手动调低音量。

[0100] 上述实施例中,能够采集所述声音信号,并在向播放采集到的所述声音信号时,降低第一音频信号的输出音量,使得用户不仅能够察觉到有他人对用户讲话,还能够较为清楚地听到他人对用户讲话的内容,从而用户无需摘下耳机、或关闭其他音频输出设备、或手动调低音量。

[0101] 在本发明一种实施例中,所述方法包括图5所示的步骤:

[0102] 步骤501:在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的语音信号;

[0103] 步骤502:将采集到的声音信号转换成录音文件;即录制所述声音信号;

[0104] 步骤503:对所述录音文件进行降噪处理;在具体实施例中,可通过对所述录音文件进行滤波处理等操作,在录音文件的声波信号中,识别并过滤掉噪音信号,对所述录音文件进行降噪;

[0105] 步骤504:暂停播放所述第一音频信号或继续播放第一音频信号,通过所述终端的音频输出设备播放所述经过降噪处理后的录音文件。上述实施例中,采用另外一种方式对所述声音信号进行采集并降噪,使得用户能够较为清楚地接收到他人讲话内容。

[0106] 在本发明另一种实施例中,所述方法包括:

[0107] 在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的语音信号;

[0108] 将采集到的声音信号转换成录音文件;即录制所述声音信号;

[0109] 对所述录音文件进行降噪处理;

[0110] 降低第一音频文件的输出音量,或暂停第一音频文件的播放;

[0111] 通过所述终端的音频输出设备播放所述经过降噪处理后的录音文件。使得用户不仅能够察觉到有他人对用户讲话,还能够较为清楚地听到他人对用户讲话的内容,从而用户无需摘下耳机、或关闭其他音频输出设备、或手动调低音量。

[0112] 在本发明一种实施例中,所述方法还包括:

[0113] 在用户通过预设的应用与另一方用户进行语音通信操作时,采集所述另一方用户的语音;在执行本步骤之前,用户可通过终端的设置功能授权,使得在用户进行语音通话操作或语音播放时,能够从语音中获取声纹特征;所述语音通信操作包括通话操作或语音播放;在本发明具体实施例中,所述预设的应用包括但不限于:语音通话应用、或QQ、或微信、或具有语音通话功能的其它社交软件、语音通话软件等;

[0114] 从采集的另一方用户的语音中提取声纹特征;

[0115] 将提取出的声纹特征存储为样本声纹特征;在具体操作时,若提取出的声纹特征没有进行存储,则存储所述声纹特征;若提取出的声纹已经存储为样本声纹特征,则可不存储提取出的声纹特征;

[0116] 在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的语音信号;

[0117] 若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,对所述第一音频信号进行处理,使得用户能够听到所述声音信号。

[0118] 在本发明一种具体实施例中,在采集所述终端的周围环境的聲音信号之前,需获得用户的授权,用户可通过终端的设置功能开启环境音监听,从而在终端以预设方式播放第一音频信号时,能够进入采集所述终端的周围环境的聲音信号的步骤。

[0119] 可以看出,本发明实施例提供的音频信号处理方法,能够在用户以设定的模式播放第一音频信号时,监听终端周围的环境中的聲音信号,并且能够在确定所述聲音信号的声纹特征与预设的声纹特征匹配、且聲音信号的音量高于设定音量阈值时,对播放的第一音频信号进行处理,使得用户能够察觉到所述聲音信号的存在,从而不会因为用户使用耳机或音箱或将终端外放音量调得过大而听不到他人向自己讲话。

[0120] 同时,本发明还提供一种终端,结构框图如图6所示,能够实现上述音频输出方法的细节,并实现相应效果。如图6所示,终端600包括:声音信号采集模块601、第一音频信号处理模块602,其中,

[0121] 声音信号采集模块601:用于在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的聲音信号;

[0122] 第一音频信号处理模块602:用于若所述聲音信号的声纹特征与预先存储的样本声纹特征匹配,且所述聲音信号的音量值超过第一阈值时,对所述第一音频信号进行处理,使得用户能够听到所述聲音信号。

[0123] 在本发明一些实施例中,所述第一音频信号处理模块包括:

[0124] 第一处理单元:用于降低所述第一音频信号的输出音量;或,

[0125] 第二处理单元:用于降低所述第一音频信号的输出音量,同时输出所述聲音信号;或,

[0126] 第三处理单元:用于暂停所述第一音频信号的播放;或,

[0127] 第四处理单元:用于暂停所述第一音频信号的播放,同时输出所述聲音信号;或,

[0128] 第五处理单元:用于在播放所述第一音频信号时,同时输出所述聲音信号。

[0129] 在本发明一种具体实施例中,所述终端包括声音信号采集模块、第一音频信号处理模块,其中,第一音频信号处理模块进一步包括:第一处理单元,用于降低所述第一音频信号的输出音量。

[0130] 在本发明一种具体实施例中,所述终端包括声音信号采集模块、第一音频信号处理模块,其中,第一音频信号处理模块进一步包括:第二处理单元,用于降低所述第一音频信号的输出音量,同时输出所述聲音信号。

[0131] 在本发明一种具体实施例中,所述终端包括声音信号采集模块、第一音频信号处理模块,其中,第一音频信号处理模块进一步包括:第三处理单元,用于暂停所述第一音频信号的播放。

[0132] 在本发明一种具体实施例中,所述终端包括声音信号采集模块、第一音频信号处理模块,其中,第一音频信号处理模块进一步包括:第四处理单元:用于暂停所述第一音频信号的播放,同时输出所述聲音信号。

[0133] 在本发明一种具体实施例中,所述终端包括声音信号采集模块、第一音频信号处理模块,其中,第一音频信号处理模块进一步包括:第五处理单元:用于在播放所述第一音

频信号时,同时输出所述声音信号。

[0134] 可选的,所述第一处理单元包括:

[0135] 播放音量降低子单元:用于减小终端的播放音量,使得所述第一音频信号的输出音量降低;

[0136] 或者,静音子单元:用于将所述终端的静音,使得所述第一音频信号的输出音量降低为零;

[0137] 或者,混音子单元:用于利用与所述第一音频信号相位相抵消的第二音频信号对第一音频信号进行混音,使得所述第一音频信号的输出音量降低。

[0138] 可选的,所述第二处理单元包括:

[0139] 播放音量降低子单元:用于减小终端的播放音量,使得所述第一音频信号的输出音量降低;

[0140] 或者,静音子单元:用于将所述终端的静音,使得所述第一音频信号的输出音量降低为零;

[0141] 或者,混音子单元:用于利用与所述第一音频信号相位相抵消的第二音频信号对第一音频信号进行混音,使得所述第一音频信号的输出音量降低;

[0142] 可选的,所述第二处理单元还包括:

[0143] 声源方向确定子单元:用于若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,确定所述声音信号的声源方向;

[0144] 第一降噪子单元:用于在采集所述声音信号时,根据所述声源方向对所述声音信号进行降噪处理,得到经过降噪处理的声音信号;

[0145] 第一播放子单元:用于通过所述终端的音频输出设备播放经过降噪处理的所述声音信号。

[0146] 可选的,所述第四处理单元包括:

[0147] 声源方向确定子单元:用于若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,确定所述声音信号的声源方向;

[0148] 第一降噪子单元:用于在采集所述声音信号时,根据所述声源方向对所述声音信号进行降噪处理,得到经过降噪处理的声音信号;

[0149] 第一播放子单元:用于通过所述终端的音频输出设备播放经过降噪处理的所述声音信号。

[0150] 可选的,所述第二处理单元还包括:

[0151] 录音子单元:用于将采集到的声音信号转换成录音文件;

[0152] 第二降噪子单元:用于对所述录音文件进行降噪处理;

[0153] 第二播放子单元:用于通过所述终端的音频输出设备播放经过降噪处理后的所述录音文件;

[0154] 当所述第一音频信号处理模块包括所述第四处理单元时,所述第四处理单元包括:所述录音子单元、所述第二降噪子单元、所述第二播放子单元。

[0155] 在本发明另一种实施例中,所述终端结构如图7所示,在图6的基础上,终端600包括:声音信号采集模块601、第一音频信号处理模块602、语音采集模块603、声纹提取模块604、声纹存储模块605,其中,

[0156] 声音信号采集模块601:用于在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的聲音信号;

[0157] 第一音频信号处理模块602:用于若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,对所述第一音频信号进行处理,使得用户能够听到所述声音信号

[0158] 语音采集模块603:用于在用户通过预设的应用与另一方用户进行语音通信操作时,采集所述另一方用户的语音;所述语音通信操作包括通话操作或语音播放;

[0159] 声纹提取模块604:用于从采集的另一方用户的语音中提取声纹特征;

[0160] 声纹存储模块605:用于将提取出的声纹特征存储为样本声纹特征。

[0161] 本发明实施例中,用户通过终端以预设方式播放第一音频信号时,采集终端周围环境中的声音信号,并在所述声音信号的声纹特征与预设的样本声纹特征陪时,对第一音频信号进行处理,使得用户能够听到声音信号。从而在用户收听第一音频文件时,不会因为音量的覆盖而无法听到周围他人向用户讲话,从而避免用户与他人的沟通收到终端播放第一音频信号的操作的影响。

[0162] 本发明另一种实施例参见图8,图8是本发明实施例提供的终端的结构图,能够实现上述实施例中应用程序的卸载方法的细节,并达到相同的效果。如图8所示,终端800包括:至少一个处理器801、存储器802、至少一个网络接口804和用户接口803。终端800中的各个组件通过总线系统805耦合在一起。可理解,总线系统805用于实现这些组件之间的连接通信。总线系统805除包括数据总线之外,还包括电源总线、控制总线和状态信号总线。但是为了清楚说明起见,在图8中将各种总线都标为总线系统805。

[0163] 其中,用户接口803可以包括显示器、键盘或者点击设备(例如,鼠标,轨迹球(track ball)、触感板或者触摸屏等。

[0164] 可以理解,本发明实施例中的存储器802可以是易失性存储器或非易失性存储器,或可包括易失性和非易失性存储器两者。其中,非易失性存储器可以是只读存储器(Read-Only Memory,ROM)、可编程只读存储器(Programmable ROM,PROM)、可擦除可编程只读存储器(Erasable PROM,EPROM)、电可擦除可编程只读存储器(Electrically EPROM,EEPROM)或闪存。易失性存储器可以是随机存取存储器(Random Access Memory,RAM),其用作外部高速缓存。通过示例性但不是限制性说明,许多形式的RAM可用,例如静态随机存取存储器(Static RAM,SRAM)、动态随机存取存储器(Dynamic RAM,DRAM)、同步动态随机存取存储器(Synchronous DRAM,SDRAM)、双倍数据速率同步动态随机存取存储器(Double Data Rate SDRAM,DDRSDRAM)、增强型同步动态随机存取存储器(Enhanced SDRAM,ESDRAM)、同步连接动态随机存取存储器(Synchlink DRAM,SLDRAM)和直接内存总线随机存取存储器(Direct Rambus RAM,DRRAM)。本文描述的系统和方法的存储器802旨在包括但不限于这些和任意其它适合类型的存储器。

[0165] 在一些实施方式中,存储器802存储了如下的元素,可执行模块或者数据结构,或者他们的子集,或者他们的扩展集:操作系统8021和应用程序8022。

[0166] 其中,操作系统8021,包含各种系统程序,例如框架层、核心库层、驱动层等,用于实现各种基础业务以及处理基于硬件的任务。应用程序8022,包含各种应用程序,例如媒体播放器(Media Player)、浏览器(Browser)等,用于实现各种应用业务。实现本发明实施例

方法的程序可以包含在应用程序8022中。

[0167] 在本发明实施例中,通过调用存储器802存储的程序或指令,具体的,可以是应用程序8022中存储的程序或指令,处理器801用于:在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的声音信号;

[0168] 若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,对所述第一音频信号进行处理,使得用户能够听到所述声音信号。

[0169] 上述本发明实施例揭示的方法可以应用于处理器801中,或者由处理器801实现。处理器801可能是一种集成电路芯片,具有信号的处理能力。在实现过程中,上述方法的各步骤可以通过处理器801中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器801可以是通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现场可编程门阵列(Field Programmable GateArray,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。可以实现或者执行本发明实施例中的公开的各方法、步骤及逻辑框图。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。结合本发明实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成,或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器,闪存、只读存储器,可编程只读存储器或者电可擦写可编程存储器、寄存器等领域成熟的存储介质中。该存储介质位于存储器802,处理器801读取存储器802中的信息,结合其硬件完成上述方法的步骤。

[0170] 可以理解的是,本文描述的这些实施例可以用硬件、软件、固件、中间件、微码或其组合来实现。对于硬件实现,处理单元可以实现在一个或多个专用集成电路(Application Specific Integrated Circuits,ASIC)、数字信号处理器(Digital Signal Processing,DSP)、数字信号处理设备(DSP Device,DSPD)、可编程逻辑设备(Programmable Logic Device,PLD)、现场可编程门阵列(Field-Programmable Gate Array,FPGA)、通用处理器、控制器、微控制器、微处理器、用于执行本申请所述功能的其它电子单元或其组合中。

[0171] 对于软件实现,可通过执行本文所述功能的模块(例如过程、函数等)来实现本文所述的技术。软件代码可存储在存储器中并通过处理器执行。存储器可以在处理器中或在处理器外部实现。

[0172] 可选的,处理器801还用于执行:降低所述第一音频信号的输出音量、降低所述第一音频信号的输出音量,同时输出所述声音信号、暂停所述第一音频信号的播放、暂停所述第一音频信号的播放,同时输出所述声音信号、在播放所述第一音频信号时,同时输出所述声音信号中的一种或者多种处理方法。

[0173] 可选的,处理器801还用于:将减小终端的播放音量,使得所述第一音频信号的输出音量降低;或者,将所述终端的静音,使得所述第一音频信号的输出音量降低为零;或者,利用与所述第一音频信号相位相抵消的第二音频信号对第一音频信号进行混音,使得所述第一音频信号的输出音量降低。

[0174] 可选的,处理器801还用于:若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,确定所述声音信号的声源方向;在采集所述声音信号时,根据所述声源方向对所述声音信号进行降噪处理,得到经过降噪处理的声音信号;通过所述终端的音频输出设备播放经过降噪处理的所述声音信号。

[0175] 可选的,处理器801还用于:将采集到的声音信号转换成录音文件;

[0176] 对所述录音文件进行降噪处理;

[0177] 通过所述终端的音频输出设备播放所述经过降噪处理后的录音文件。。

[0178] 可选的,处理器801还用于:在用户通过预设的应用与另一方用户进行语音通信操作时,采集所述另一方用户的语音;所述语音通信操作包括通话操作或语音播放;从采集的另一方用户的语音中提取声纹特征;将提取出的声纹特征存储为样本声纹特征。

[0179] 本发明实施例中,用户通过终端以预设方式播放第一音频信号时,采集终端周围环境中的声音信号,并在所述声音信号的声纹特征与预设的样本声纹特征陪时,对第一音频信号进行处理,使得用户能够听到声音信号。从而,在用户收听第一音频文件时,也能够感知到他人跟用户讲话的行为,不会因为音量的覆盖而无法听到周围他人向用户讲话,从而避免用户与他人的沟通收到终端播放第一音频信号的操作的影响。

[0180] 请参阅图9,图9是本发明实施例提供的终端的结构图,能够实现上述实施例中应用程序的卸载方法的细节,并达到相同的效果。如图9所示,终端900包括射频(Radio Frequency, RF)电路910、存储器920、输入单元930、显示单元940、处理器950、音频电路960、通信模块970、和电源980,还包括摄像头(图中未示出)。

[0181] 其中,输入单元930可用于接收用户输入的数字或字符信息,以及产生与终端900的用户设置以及功能控制有关的信号输入。具体地,本发明实施例中,该输入单元930可以包括触控面板931。触控面板931,也称为触摸屏,可收集用户在其上或附近的触摸操作(比如用户使用手指、触笔等任何适合的物体或附件在触控面板931上的操作),并根据预先设定的程式驱动相应的连接装置。可选的,触控面板931可包括触摸检测装置和触摸控制器两个部分。其中,触摸检测装置检测用户的触摸方位,并检测触摸操作带来的信号,将信号传送给触摸控制器;触摸控制器从触摸检测装置上接收触摸信息,并将它转换成触点坐标,再送给该处理器950,并能接收处理器950发来的命令并加以执行。此外,可以采用电阻式、电容式、红外线以及表面声波等多种类型实现触控面板931。除了触控面板931,输入单元930还可以包括其他输入设备932,其他输入设备932可以包括但不限于物理键盘、功能键(比如音量控制按键、开关按键等)、轨迹球、鼠标、操作杆等中的一种或多种。

[0182] 其中,显示单元940可用于显示由用户输入的信息或提供给用户的信息以及终端900的各种菜单界面。显示单元940可包括显示面板941,可选的,可以采用LCD或有机发光二极管(Organic Light-Emitting Diode, OLED)等形式来配置显示面板941。

[0183] 应注意,触控面板931可以覆盖显示面板941,形成触摸显示屏,当该触摸显示屏检测到在其上或附近的触摸操作后,传送给处理器950以确定触摸事件的类型,随后处理器950根据触摸事件的类型在触摸显示屏上提供相应的视觉输出。

[0184] 其中处理器950是终端900的控制中心,利用各种接口和线路连接整个手机的各个部分,通过运行或执行存储在第一存储器921内的软件程序和/或模块,以及调用存储在第二存储器922内的数据,执行终端900的各种功能和处理数据,从而对终端900进行整体监控。可选的,处理器950可包括一个或多个处理单元。

[0185] 在本发明实施例中,通过调用存储该第一存储器921内的软件程序和/或模块和/或该第二存储器922内的数据,处理器950用于:在终端以预设方式播放第一音频信号时,采集所述终端的周围环境的的声音信号;若所述声音信号的声纹特征与预先存储的样本声纹特

征匹配,且所述声音信号的音量值超过第一阈值时,对所述第一音频信号进行处理,使得用户能够听到所述声音信号。

[0186] 可选的,处理器950还用于:降低所述第一音频信号的输出音量、降低所述第一音频信号的输出音量,同时输出所述声音信号、暂停所述第一音频信号的播放、暂停所述第一音频信号的播放,同时输出所述声音信号、在播放所述第一音频信号时,同时输出所述声音信号。

[0187] 可选的,减小终端的播放音量,使得所述第一音频信号的输出音量降低;

[0188] 或者,将所述终端的静音,使得所述第一音频信号的输出音量降低为零;

[0189] 或者,利用与所述第一音频信号相位相抵消的第二音频信号对第一音频信号进行混音,使得所述第一音频信号的输出音量降低。

[0190] 可选的,处理器950还用于:若所述声音信号的声纹特征与预先存储的样本声纹特征匹配,且所述声音信号的音量值超过第一阈值时,确定所述声音信号的声源方向;在采集所述声音信号时,根据所述声源方向对所述声音信号进行降噪处理,得到经过降噪处理的声音信号;通过所述终端的音频输出设备播放经过降噪处理的所述声音信号。

[0191] 可选的,处理器950还用于:将采集到的声音信号转换成录音文件;对所述录音文件进行降噪处理;通过所述终端的音频输出设备播放所述经过降噪处理后的录音文件。

[0192] 可选的,处理器950还用于:在用户通过预设的应用与另一方用户进行语音通信操作时,采集所述另一方用户的语音;所述语音通信操作包括通话操作或语音播放;从采集的另一方用户的语音中提取声纹特征;将提取出的声纹特征存储为样本声纹特征。

[0193] 本发明实施例中,用户通过终端以预设方式播放第一音频信号时,采集终端周围环境中的声音信号,并在所述声音信号的声纹特征与预设的样本声纹特征陪时,对第一音频信号进行处理,使得用户能够听到声音信号。从而,在用户收听第一音频文件时,也能够感知到他人跟用户讲话的行为,不会因为音量的覆盖而无法听到周围他人向用户讲话,从而避免用户与他人的沟通收到终端播放第一音频信号的操作的影响。

[0194] 进一步,本发明还提供一种终端,如图10所示,终端1000包括:存储器1001、处理器1002及存储在所述存储器1001上并可在所述处理器1002上运行的计算机程序1003,所述处理器1002执行所述程序时实现如本发明任意一项实施例所述的音频信号处理方法中的步骤。

[0195] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本发明的范围。

[0196] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统、装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0197] 在本申请所提供的实施例中,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。

[0198] 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发明实施例方案的目的。

[0199] 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。

[0200] 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品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)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0201] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以权利要求的保护范围为准。

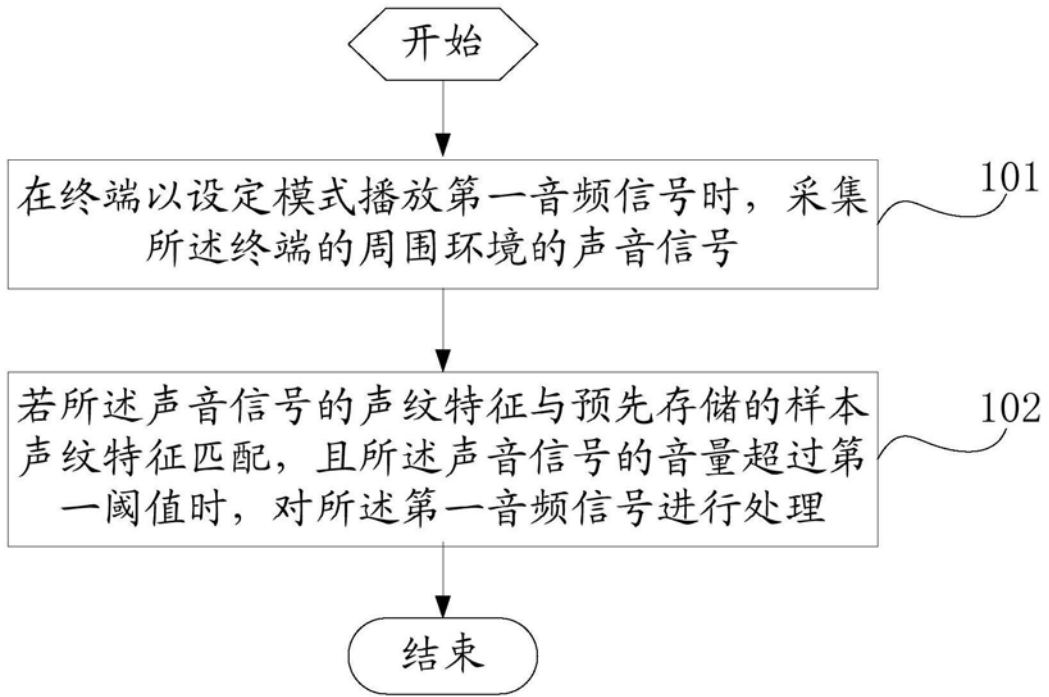


图1

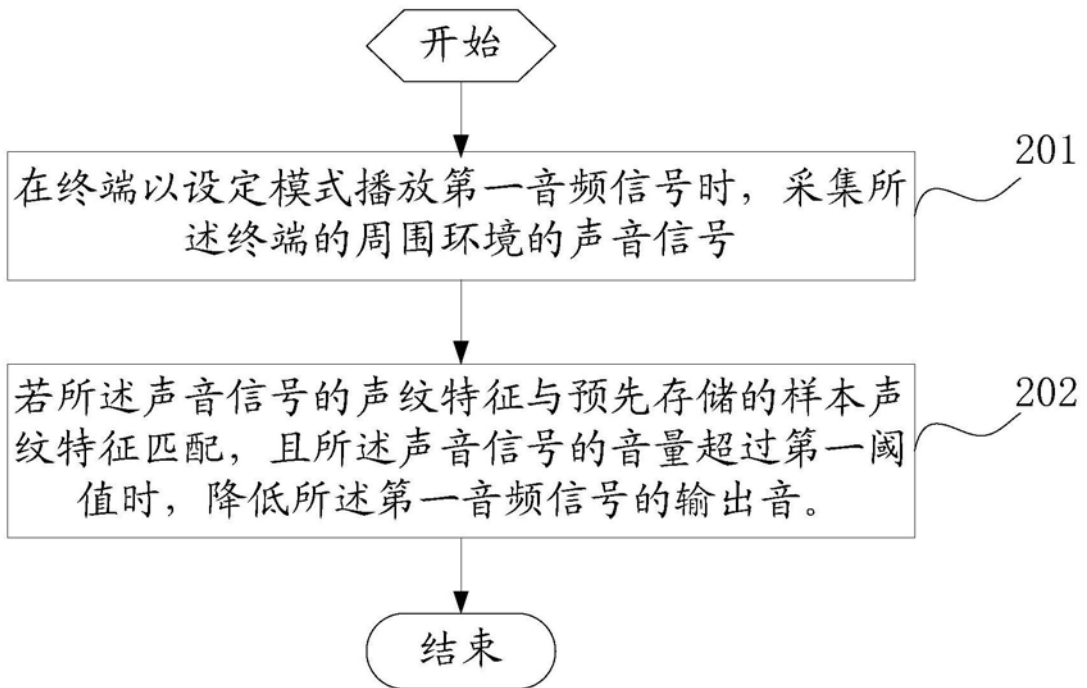


图2

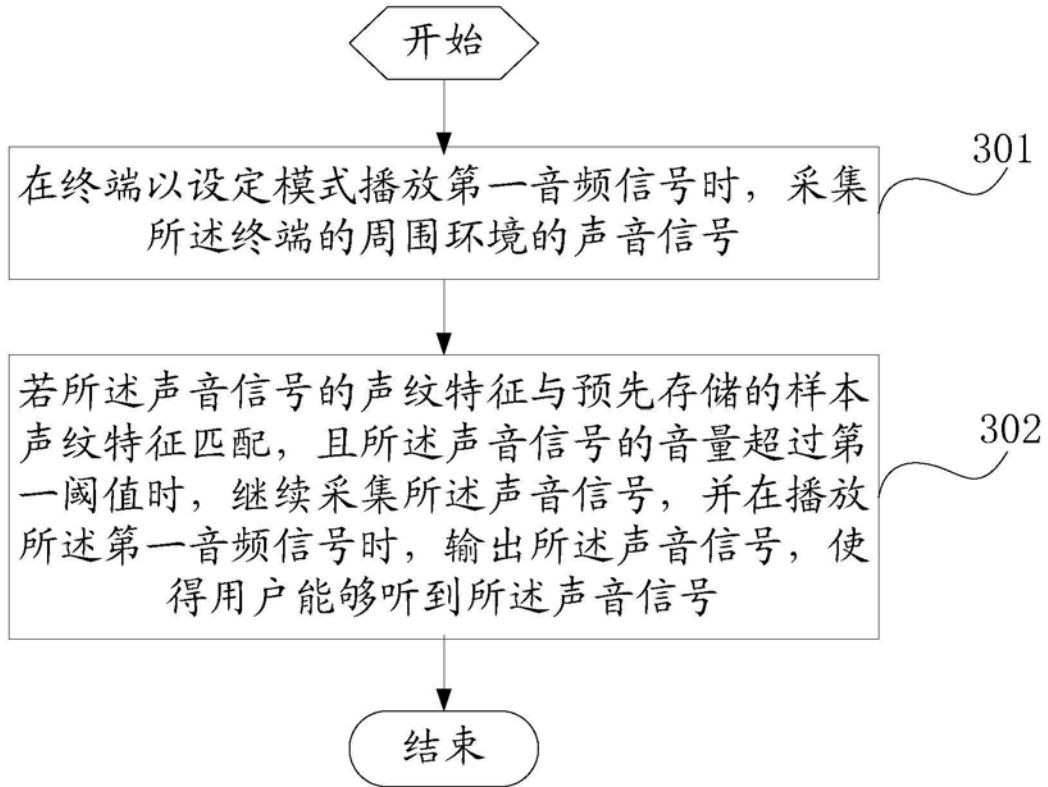


图3

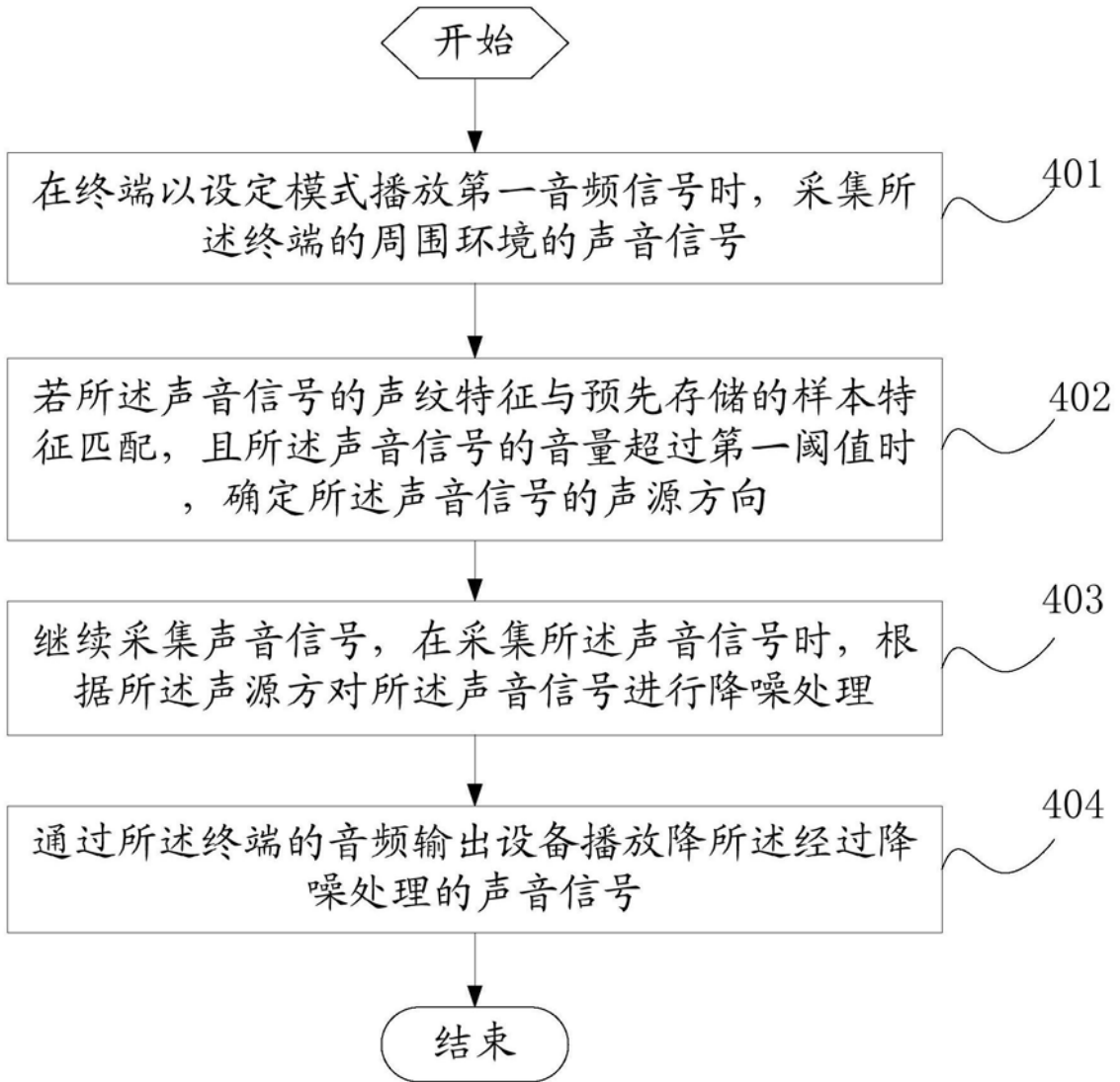


图4

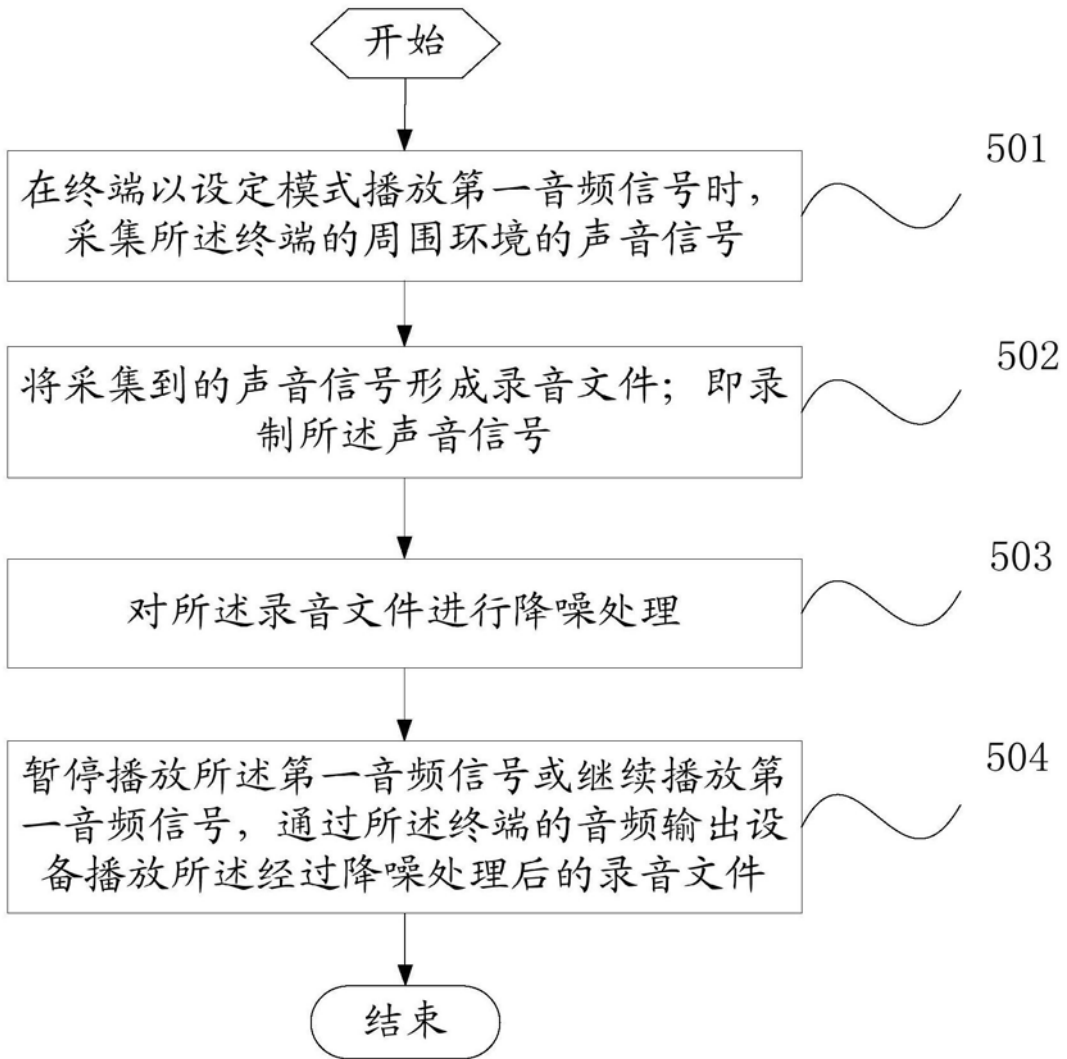


图5

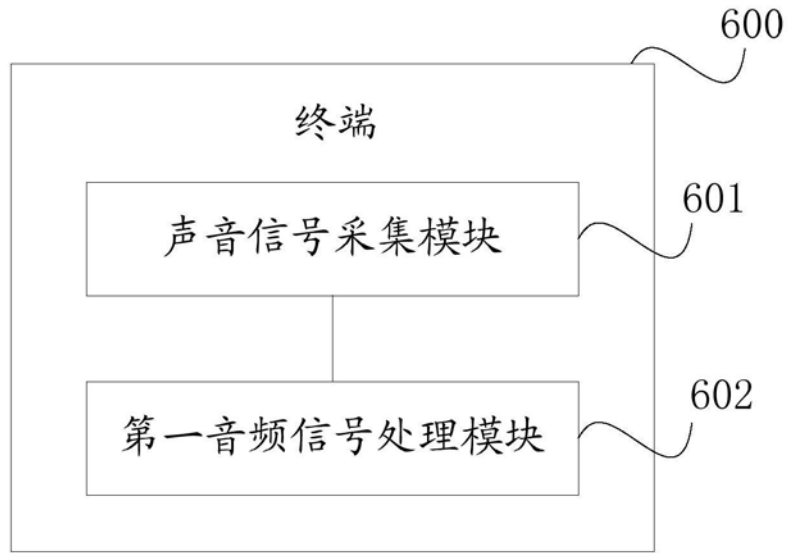


图6

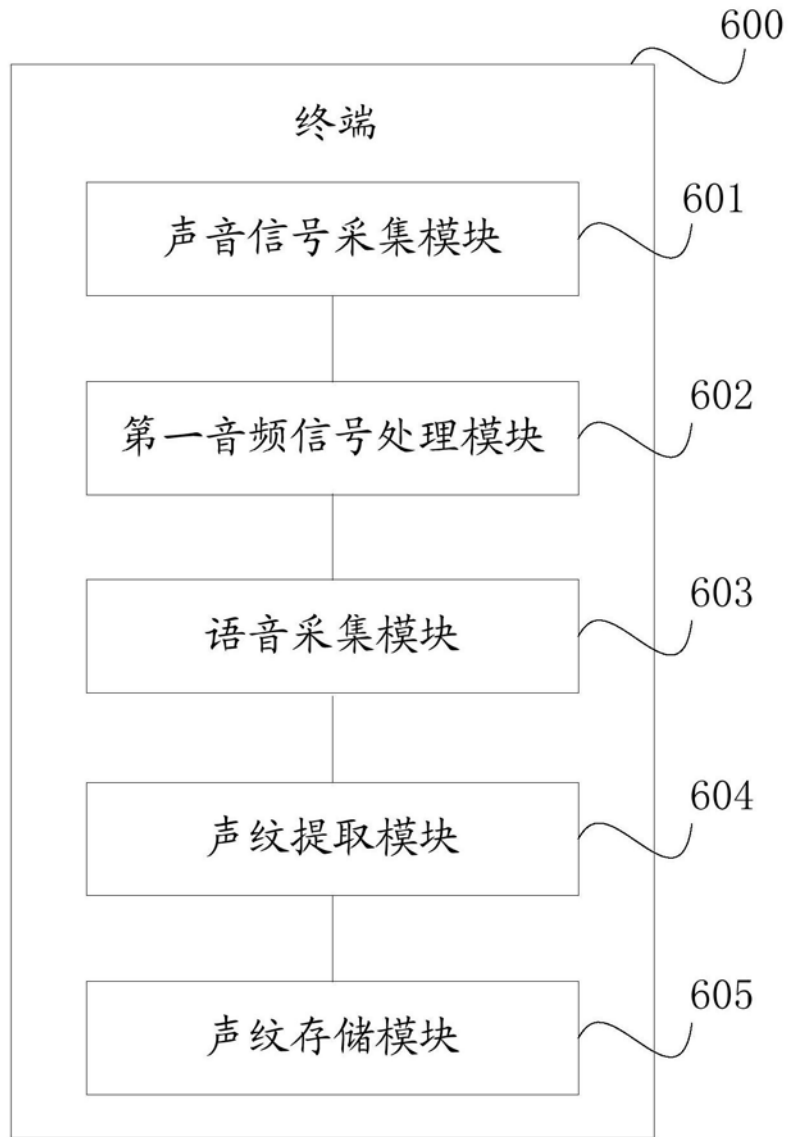


图7

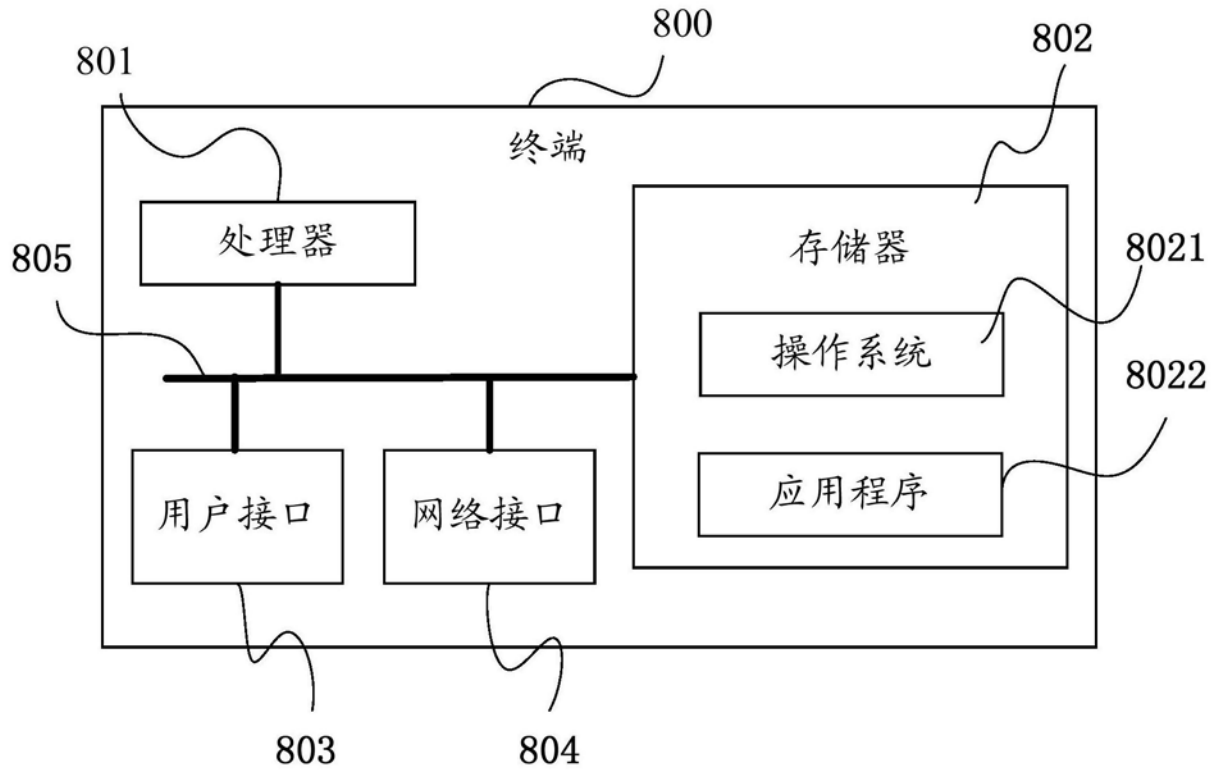


图8



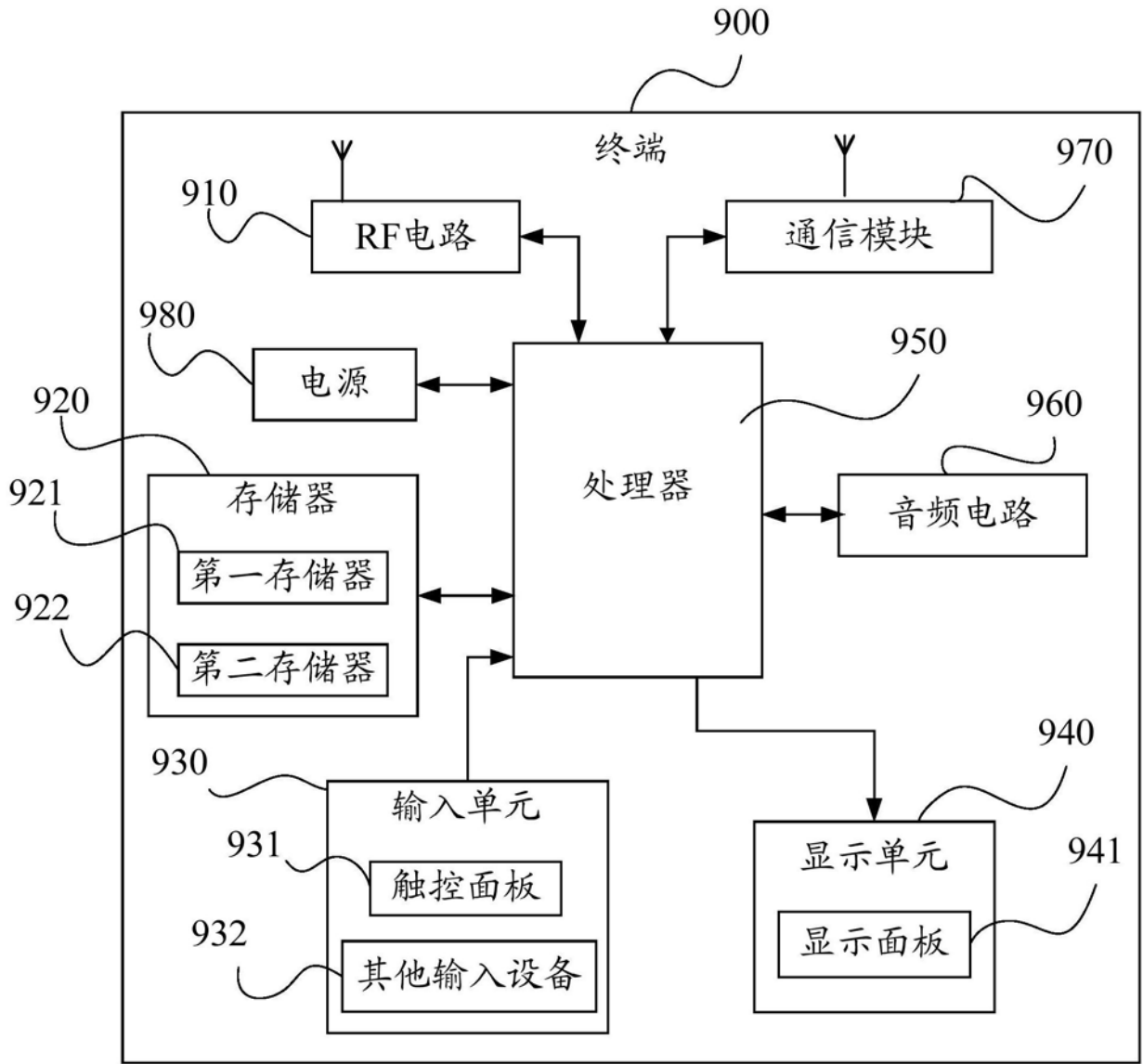


图9

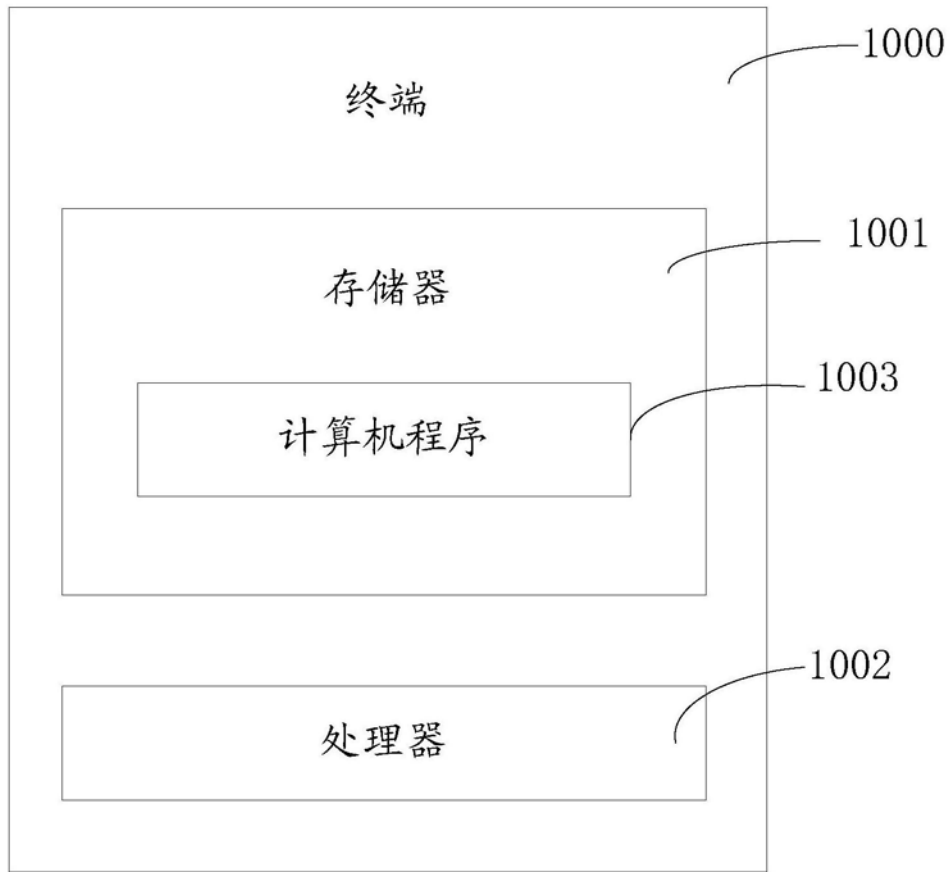


图10