(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2017 年 5 月 11 日 (11.05.2017)

WIPO | PCT

(10) 国际公布号
WO 2017/075979 A1

(51) 国际专利分类号:
G10L 21/0364 (2013.01)

(21) 国际申请号: PCT/CN2016/083622

(22) 国际申请日: 2016 年 5 月 27 日 (27.05.2016)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201510741057.1 2015 年 11 月 4 日 (04.11.2015) CN

(71) 申请人: **腾讯科技（深圳）有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED)** [CN/CN]; 中国广东省深圳市福田区振兴路赛格科技园 2 栋东 403 室, Guangdong 518000 (CN)。

(72) 发明人: **袁豪磊 (YUAN, Haolei)**; 中国广东省深圳市福田区振兴路赛格科技园 2 栋东 403 室, Guangdong 518000 (CN)。

(74) 代理人: **深圳市深佳知识产权代理事务所（普通合伙） (SHENPAT INTELLECTUAL PROPERTY AGENCY)**; 中国广东省深圳市国贸大厦 15 楼西座 1521 室, Guangdong 518014 (CN)。

(81) 指定国 (除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明，要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH,

[见续页]

(54) **Title**: METHOD AND APPARATUS FOR PROCESSING VOICE SIGNAL

(54) **发明名称** : 语音信号的处理方法及装置

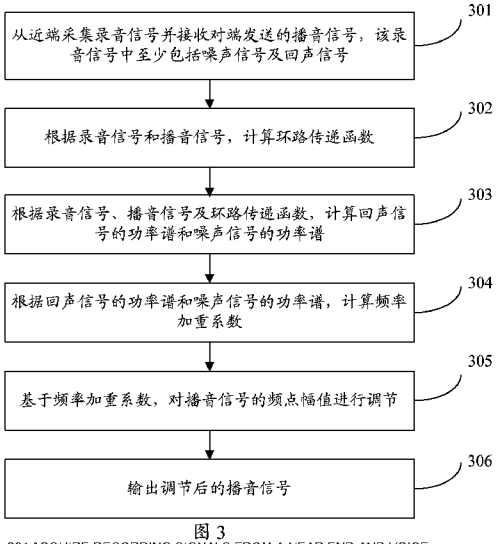


图 3

301 ACQUIRE RECORDING SIGNALS FROM A NEAR END AND VOICE SIGNALS SENT BY AN OPPOSITE END, THE RECORDING SIGNALS AT LEAST COMPRISING A NOISE SIGNAL AND AN ECHO SIGNAL
302 CALCULATE A LOOP TRANSFER FUNCTION ACCORDING TO THE RECORDING SIGNALS AND THE BROADCAST SIGNALS
303 CALCULATE A POWER SPECTRUM OF THE ECHO SIGNAL AND A POWER SPECTRUM OF THE NOISE SIGNAL ACCORDING TO THE RECORDING SIGNALS, THE BROADCAST SIGNALS AND THE LOOP TRANSFER FUNCTION
304 CALCULATE A FREQUENCY WEIGHING FACTOR ACCORDING TO THE POWER SPECTRUM OF THE ECHO SIGNAL AND THE POWER SPECTRUM OF THE NOISE SIGNAL
305 ADJUST A FREQUENCY CHANNEL NUMBER AMPLITUDE VALUE OF A BROADCAST SIGNAL ACCORDING TO THE FREQUENCY WEIGHING FACTOR
306 OUTPUT THE ADJUSTED BROADCAST SIGNAL

(57) **Abstract**: A method and apparatus for processing a voice signal, which relate to the technical field of terminals. The method for processing a voice signal comprises: acquiring recording signals and voice signals, the recording signals at least comprising a noise signal and an echo signal (301); calculating a loop transfer function according to the recording signals and the voice signals (302); calculating a power spectrum of the echo signal and a power spectrum of the noise signal according to the recording signals, the voice signals and the loop transfer function (303); calculating a frequency weighing factor according to the power spectrum of the echo signal and the power spectrum of the noise signal (304); adjusting a frequency channel number amplitude value of a voice signal according to the frequency weighing factor(S305); and outputting the adjusted voice signal (306). In the method and apparatus for processing a voice signal, on the premise of ensuring that a loudspeaker is not overloaded and the dynamic amplitude of an original broadcasting signal is not damaged, a frequency channel number amplitude value of a broadcast signal is automatically adjusted according to frequency distribution of the noise signal and the broadcast signal, thereby remarkably improving the voice intelligibility.

(57) **摘要**:

[见续页]

WO 2017/075979 A1

CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种语音信号的处理方法及装置，属于终端技术领域。语音信号处理方法包括：获取录音信号和语音信号，录音信号中至少包括噪声信号及回声信号(301)；根据录音信号和语音信号，计算环路传递函数(302)；根据录音信号、语音信号及环路传递函数，计算回声信号的功率谱和噪声信号的功率谱(303)；根据回声信号的功率谱和噪声信号的功率谱，计算频率加重系数(304)；基于频率加重系数，对语音信号的频点幅值进行调节(305)；输出调节后的语音信号(306)。该语音信号处理方法及装置在确保扬声器不过载，且不破坏原始播音信号的动态幅度的前提下，自动根据噪声信号与播音信号的频率分布，调整播音信号的频点幅值，明显提高了语音可懂度。

语音信号的处理方法及装置

本申请要求于2015年11月4日提交中国专利局，申请号为201510741057.1，发明名称为“语音信号的处理方法及装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本发明涉及终端技术领域，特别涉及一种语音信号的处理方法及装置。

背景技术

语音可懂度是指用户听懂声音系统所传递的语音信号的百分比，例如，如果用户听到声音系统传递了100个单词，但仅听懂了50个单词，则该系统的语音可懂度为 50%。随着便携式移动终端的外形尺寸逐渐向小型化方向发展，移动终端所能输出的最大声音功率逐渐减小，相应地用户使用移动终端进行通信时的语音可懂度也受到了影响。由于语音可懂度是衡量移动终端性能的一项重要指标，因此，移动终端如何处理语音信号，以改善语音可懂度，成为其发展的关键。

目前，在由移动终端、用户、噪声源所构成的典型声学应用场景下，采用自动增益控制算法检测待播放的播音信号，并对待播放的播音信号中的小信号进行放大，将放大后的播音信号转化为电信号，并将电信号传送到扬声器。通过上述放大处理，使得送到扬声器的电信号达到扬声器所允许的最大值，扬声器工作在最大输出功率的状态下，此时扬声器以最大的输出声压级输出语音信号。

在实现本发明的过程中，发明人发现相关技术至少存在以下问题:

由于通常播音信号的平均波动幅度远小于峰值波动幅度，对于一个最大额定输出功率为 1 瓦的扬声器来说，在正常语音信号的激励下，它正常工作时的平均输出功率一般仅达到最大额定输出功率的 10%左右（也就是 0.1W）。在正常工作状态下，如果继续加大输入到扬声器的电信号幅度，则播音信号中幅度较大的信号部分将导致扬声器过载，形成饱和失真，反而降低了语音可懂度与清晰度；另外，如果仅对播音信号中的小信号作放大处理，则将缩小播音信号

的有效动态范围，对应的语音可懂度同样也得不到明显提高。

发明内容

为了解决相关技术的问题，本发明实施例提供了一种语音信号的处理方法及装置。所述技术方案如下:

一方面，提供了一种语音信号的处理方法，所述方法包括:

获取录音信号和语音信号，所述录音信号中至少包括噪声信号及回声信号;

根据所述录音信号和所述语音信号，计算环路传递函数;

根据所述录音信号、所述语音信号及所述环路传递函数，计算所述回声信号的功率谱和所述噪声信号的功率谱;

根据所述回声信号的功率谱和所述噪声信号的功率谱，计算频率加重系数;

基于所述频率加重系数，对所述语音信号的频点幅值进行调节;

输出调节后的语音信号。

另一方面，提供了一种语音信号的处理装置，所述装置包括:

至少一个处理器；和

存储器，其中所述存储器存储有程序指令，所述指令当由所述处理器执行时，配置所述装置执行下述操作:

获取录音信号和语音信号，所述录音信号中至少包括噪声信号及回声信号;

根据所述录音信号和所述语音信号，计算环路传递函数;

根据所述录音信号、所述语音信号及所述环路传递函数，计算所述回声信号的功率谱和所述噪声信号的功率谱;

根据所述回声信号的功率谱和所述噪声信号的功率谱，计算频率加重系数;

基于所述频率加重系数，对所述语音信号的频点幅值进行调节;

输出调节后的语音信号。

本发明实施例提供的技术方案带来的有益效果是:

在确保扬声器不过载，且不破坏原始播音信号的动态幅度的前提下，自动根据噪声信号与播音信号的频率分布，调整播音信号的频点幅值，明显提高了语音可懂度。

附图说明

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

图 1 是本发明一实施例提供的语音信号的处理方法所涉及的实施环境的示意图；

图 2 是本发明另一实施例提供的语音信号的处理方法的系统架构图；

图 3 是本发明另一实施例提供的一种语音信号的处理方法流程图；

图 4 是本发明的另一实施例提供的一种语音信号的处理方法流程图；

图 5 是本发明的另一实施例提供的一种语音信号的处理方法对应的信号流的示意图；

图 6 是本发明的另一实施例提供的一种语音信号的处理方法的流程图；

图 7 是本发明另一实施例提供的一种语音信号的处理装置的结构示意图；

图 8 是本发明另一实施例提供的一种语音信号的处理终端的结构示意图。

具体实施方式

为使本发明的目的、技术方案和优点更加清楚，下面将结合附图对本发明实施方式作进一步地详细描述。显然，所描述的实施例仅仅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。

语音即时通讯应用是一种能够拨打网络电话或网络音频会议的应用，被广泛地安装在智能手机、平板电脑、笔记本电脑、可穿戴电子产品等移动终端设备上。随着这些移动终端的外形尺寸逐渐向小型化方向发展，移动终端设备中的微型扬声器所能够输出的最大声功率也遇到了瓶颈。

造成移动终端所输出的最大声功率遇到瓶颈的主要原因有以下两方面：

第一方面、现有的电声扩音技术主要依靠功放、扬声器、音腔三个部分共同作用实现声波的产生，当扬声器与音腔的物理尺寸与声波的波长成正比时，移动终端设备中的扬声器才能最大效率地实现电声转换。然而，随着便携式移动设备的外形尺寸向小型化发展，移动终端的外形尺寸往往比声波的波长要小

得多，以波长为 340Hz 的声波为例，若想实现最大的声电转换效率，移动终端的尺寸至少需要达到 1 米，扬声器尺寸的微型化导致移动终端输出的最大声音功率减小。另外，目前普遍使用的动圈式扬声器需要达到一定的尺寸厚度，以保证震膜有足够的运动空间，然而，随着移动终端的外形尺寸的减小、厚度变薄，移动终端内的整体声学设计受到了物理尺寸的限制，使得移动终端输出的最大声功率受到了限制。

第二方面、通常移动终端中所安装的语音即时通讯应用一般运行于操作系统之上，需要通过操作系统提供的应用程序接口才能实现对硬件的音量控制。对于音频输入输出而言，目前主流的实现方法是语音即时通讯应用向操作系统声明要求音频配置模式，由操作系统对相关硬件作出设置，完成配置之后，语音即时通讯应用只需要定时地将播音信号所对应的数据写入操作系统的录音 API，再从操作系统的录音 API 中读取数据即可。然而操作系统所支持的音频配置模式的类型是有限的，这些有限的音频配置模式是由移动终端生产商在硬件底层（固件 firmware）中实现的，应用程序对硬件输出音量的控制受到这一因素的制约，此外硬件厂商往往仅针对正常的使用场景做底层的音频优化，对于极端环境（比如存在很大环境噪声）下的使用场景，移动终端生产商一般不会对此作针对性的优化（比如移动终端生产商一般不会提供可以提高硬件输出音量的专用软件接口）。

在常见的移动终端中，输出音量从大到小排序依次是：笔记本电脑、平板电脑、智能手机（免提模式）、可穿戴设备等。在采用这些移动终端进行通信时，这几种移动终端面临的环境噪声问题却呈相反的变化趋势：通常笔记本电脑在室内使用的频率比较高，接触到的噪声也以室内低分贝的小噪声为主；平板电脑和智能手机在室外、公共场所使用的频率要更高，接触到的噪声以高分贝的大噪声为主；可穿戴设备由于长时间佩戴在人体上，接触到的噪声场景最多、最复杂。随着移动终端的外形尺寸向小型化发展，移动终端所面临的环境噪声问题越来越突出，严重影响了用户使用移动终端进行通信时的体验效果。

为了解决上述移动终端所输出的最大声功率遇到瓶颈的问题，本发明实施例提供了一种在不对移动终端在硬件方面作改动的前提下，通过对语音信号进行处理，来提高移动终端的语音可懂度的方法。采用本发明实施例提供的方法，移动终端的用户即便处于嘈杂的场景下，也能够听清通话对端的语音内容。

图 1 为本发明实施例提供的语音信号的处理方法和装置所涉及到的实施环

境示意图。参见图 1，该实施环境包括移动终端 P、用户 U 及噪声源 N 这 3 个声学主体，还包括声音输出和输入设备扬声器 S 和麦克风 M。该移动终端 P 可以是手机、平板电脑、笔记本电脑、可穿戴设备等，其中安装一个或多个语音即时通讯应用（App），基于这些语音即时通讯应用，用户可随时随地与其他用户进行通信。扬声器 S 和麦克风 M 既可以内置于移动终端内，也可以以外接设备如外接音响、外接扬声器、蓝牙音箱、蓝牙耳机的形式连接在移动终端上。麦克风 M 可以拾取到整个场景中的声音，包括：噪声源 N 发出的噪声、用户 U 说话时发出的语音、扬声器 S 播出的声音。当用户通过语音即时通讯软件与对端用户进行通信时，移动终端接收对端发送的要播放的语音信号（为了区分，下文简称播音信号），将该播音信号处理之后，由扬声器转换成声波，通过空气传播给用户 U 并被用户 U 所感知；与此同时噪声源 N 发出的声波也通过空气传播给用户 U，同时也被用户 U 感知，该噪声源 N 发出的声波会对用户 U 形成干扰，降低了移动终端的语音可懂度。

在声学领域，根据心理声学的掩蔽效应原理，当两个频率相近、幅值差别较大的信号同时出现时，幅值较大的信号会对幅值较小的信号形成掩蔽作用。也即是，当噪声源 N 发出的噪声强度很大时，用户 U 无法听清扬声器 S 中正在播放的语音内容。此时若想加大扬声器 S 的输出功率，则需要加大 S 的物理尺寸，而这又与移动终端小型化、轻薄化的设计相矛盾。鉴于此，本发明将利用心理声学的掩蔽效应解决噪声信号对播音信号的干扰问题。

通常播音信号、噪声信号都不是单频信号，它们各自占据不同的频带范围，并且它们在各个频点上的能量分布也不是均匀的。通过对比播音信号、噪声信号的功率谱分布，可找到噪声信号中能量最低的那些频点，记为 f_weak。本实施例在不超过扬声器输出功率的前提下，将播音信号能量集中到 f_weak 附近播放出去，与此同时衰减远离 f_weak 的频点上的播音信号的能量，以避免扬声器过载。通过这种处理方式，在临近 f_weak 的频点上，噪声信号被播音信号所掩蔽，用户所感知到的是播音信号的内容。在远离 f_weak 的频点上，播音信号仍旧被噪声信号所掩蔽。综合上述内容，增强后的播音信号在部分频点上将噪声信号掩蔽，使得噪声不再对播音信号形成整体掩蔽，此时用户可以听清播音信号的内容。

图 2 为本发明提供的语音信号的处理方法的系统架构图。参见图 2，该系统架构包括用户 U、扬声器 S、麦克风 M 以及各种功能模块。其中，功能模块包

括信号检测和分类模块、频谱估计模块、环路函数传递计算模块、语音可懂度估计模块等。频谱估计模块具体可以包括语音激活检测模块、噪声功率谱模块和回声功率谱模块。对于系统的各个模块的作用及模块间的相互关系如下:

其中，麦克风 M 用于拾取环境声音，在本实施例中将环境声音称为录音信号（记为 x），并将录音信号 x 送入信号检测与分类模块。

信号检测与分类模块用于对录音信号进行检测与区分，并输出三类信号:用户 U 讲话时的语音信号（记为近端信号 v）、噪声源 N 发出的噪声信号（记为噪声信号 n）、扬声器 S 播出的声音被 M 重新录回的信号（记为回声信号 e）。

频谱估计模块用于计算噪声信号功率谱、回声信号的功率谱及近端信号的功率特征值，其中，噪声信号的功率谱可用 P_n 表示、回声信号的功率可用 P_e 表示、近端信号的功率特征值可用 *VAD_v* 表示。 *VAD_v* 具有 true 和 false 两种状态，当 *VAD_v*=true 时，说明当前时刻有近端信号存在，也即是用户 U 正在说话，当 *VAD_v*=false 时，说明当前时刻近端信号不存在，也即是用户 U 未在说话，或者用户 U 的说话声音的音量明显小于噪声信号的音量或者回声信号的音量。

环路传递函数计算模块用于根据播音信号 y 和麦克风拾取到的录音信号 x 计算出 “加重滤波器--扬声器--声场--麦克风”这条路径上的传递函数，记为 *H_loop*。

语音可懂度估计模块用于根据 *H_loop*、*VAD_v*、P_n 和 P_e，确定语音可懂度（记为 SII），该语音可懂度还用于计算加重滤波器 W 的频率加重系数。

参见图 2，在实际应用中由于用户、移动终端、噪声源这三者在空间上的具体位置是无法确定的，而对播音信号和录音信号进行处理的目的，是希望将用户 U 耳朵所处位置上的 SII 调节到最大，而不是麦克风 M 所在的位置。为了解决这一问题，本实施例提供的方法采用了近似处理。为了便于后续叙述，在本发明实施例中，将声音在扬声器 S 与用户 U 耳朵之间的传播路径的长度用 h1 表示，将声音在噪声源 N 与用户耳朵之间的传播路径的长度用 h2 表示，将声音在噪声源 N 和麦克风 M 之间的传播路径的长度用 h3 表示，将声音在用户 U 的嘴与麦克风 M 之间的传播路径的长度用 h4 表示，将声音在麦克风 M 与扬声器 S 之间的传播路径的长度用 h5 表示。本发明实施例中所作的近似如下:

（1）、设定麦克风所拾取到的噪声与用户所感受到的噪声是近似相同，也即是 h2≈h3。

（2）、设定麦克风所拾取到的来自扬声器的回声与用户感受到的扬声器所

播放的声音近似相同，也即是h1≈h5。

在以上近似条件满足的前提下，可将计算用户U所在位置的最大语音可懂度问题，转化为计算麦克风M所在位置的最大语音可懂度问题。

上述所有可选技术方案，可以采用任意结合形成本发明的可选实施例，在此不再一一赘述。

图 3 示出根据本发明一个实施例提供的语音信号的处理方法的流程图。参见图3，本实施例提供的方法包括:

301、获取录音信号和语音信号，例如，从近端采集录音信号并接收对端发送的语音信号(即播音信号)。其中该录音信号中至少包括噪声信号及回声信号。

302、根据录音信号和播音信号，计算环路传递函数。

303、根据录音信号、播音信号及环路传递函数，计算回声信号的功率谱和噪声信号的功率谱。

304、根据回声信号的功率谱和噪声信号的功率谱，计算频率加重系数。

305、基于频率加重系数，对播音信号的频点幅值进行调节。

306、输出调节后的播音信号。

本发明实施例提供的方法，在确保扬声器不过载，且不破坏原始播音信号的动态幅度的前提下，自动根据噪声信号与播音信号的频率分布，调整播音信号的频点幅值，明显提高了语音可懂度。

在本发明的另一个实施例中，根据录音信号和播音信号，计算环路传递函数，包括:

计算录音信号与播音信号之间的频域互相关函数;

计算播音信号的频域自相关函数;

根据录音信号与播音信号之间的频域互相关函数以及播音信号的频域自相关函数，计算环路传递函数。

在本发明的另一个实施例中，对于录音信号，应用以下公式，计算录音信号的功率谱:

$$P_x = X(n).\hat{}2$$

其中，P_x为录音信号的功率谱，$X(n)$为将第n时刻采集到的录音信号进行傅立叶变换得到的向量，.^2用于将$X(n)$中的每个向量元素求平方。

在本发明的另一个实施例中，根据录音信号、播音信号及环路传递函数，

计算回声信号的功率谱和噪声信号的功率谱，包括:

计算所述录音信号的功率谱;

根据环路传递函数及播音信号，计算回声信号的频谱估计值;

计算回声信号的频谱估计值的平方，得到回声信号的功率谱;

将录音信号的功率谱减去回声信号的功率谱，得到噪声信号的功率谱。

在本发明的另一个实施例中，计算回声信号的频谱估计值的平方，得到回声信号的功率谱之前，还包括:

计算录音信号的功率特征值、播音信号的功率特征值及回声信号的功率特征值;

判断录音信号的功率特征值是否大于第一阈值、播音信号的功率特征值是否大于第二阈值、回声信号的功率特征值是否大于第三阈值;

当录音信号的功率特征值大于第一阈值、播音信号的功率值大于第二阈值且回声信号的功率特征值大于第三阈值时，计算回声信号的频谱估计值的平方，得到回声信号的功率谱。

在本发明的另一个实施例中，将录音信号的功率谱减去回声信号的功率谱，得到噪声信号的功率谱之前，还包括:

判断录音信号的功率特征值是否小于第一阈值、回声信号的功率特征值是否小于第三阈值;

当录音信号的功率特征值小于第一阈值且回声信号的功率特征值小于第三阈值时，将录音信号的功率谱减去回声信号的功率谱，得到噪声信号的功率谱。

在本发明的另一个实施例中，根据回声信号的功率谱、噪声信号的功率谱，计算频率加重系数，包括:

根据回声信号的功率谱及噪声信号的功率谱，构建语音可懂度函数;

在回声信号的功率谱保持不变的条件下，根据语音可懂度函数的极大值，得到频率加重系数。

图 4 示出根据本发明另一实施例的一种语音信号的处理方法的流程图。参见图 4，本实施例提供的方法包括:

401、移动终端从近端采集录音信号并接收对端发送的播音信号。

其中，近端为移动终端当前所处的环境，移动终端从近端采集录音信号方式，包括但不限于：开启麦克风，通过麦克风采集当前环境中的声音信号，并

将麦克风采集到的声音信号作为录音信号，该录音信号中包括噪声信号、回声信号及近端信号等。在本实施例中，录音信号可用 x 表示，噪声信号可用 n 表示，回声信号可用 e 表示，近端信号可用 v 表示。

对端通过麦克风采集到对端用户的语音信号，对采集到的语音信号处理后，通过网络发送至移动终端，移动终端上的即时通讯应用接收到对端发送来的语音信号，并将对端发送的语音信号作为播音信号。对端可以是与移动终端通过语音即时通讯应用进行通信的其它移动终端。在本实施例中，播音信号可用 y 表示。

可选地，为了提高采用语音即时通讯应用的时效性，移动终端侧的麦克风会每隔预设时长采集一次录音信号，对端侧麦克风也将每隔预设时长采集一次播音信号，并将采集到的播音信号发送给移动终端。其中，预设时长可以为 10ms(毫秒)、20ms、50ms 等等。

在本实施例中，移动终端从近端采集到的录音信号以及对端发送的播音信号实质上为时域信号，为了便于后续的计算，本实施例提供的方法还将采用傅里叶变换等方法分别对采集到的录音信号和接收到的播音信号进行处理，通过处理可将时域形式的录音信号转换为频域形式的录音信号，并将时域形式的播音信号转换为频域形式的播音信号，以用于后续计算。在本实施例中，频域形式的录音信号为一个列向量，向量长度等于所采用的傅立叶变换的点数，可用 X 表示；频域形式的播音信号也为一个列向量，向量长度也等于所采用的傅立叶变换的点数，可用 Y 表示。

可选地，在将时域形式的录音信号、播音信号经过傅里叶变换后，得到的频域形式的录音信号、频域形式的播音信号的维度相同。

402、移动终端根据录音信号和播音信号，计算环路传递函数。

在本实施例中，移动终端根据录音信号和播音信号，计算环路传递函数时，可采用如下步骤 4021~4023:

4021、移动终端获取录音信号与播音信号之间的频域互相关函数。

其中，互相关函数用于表示两个信号之间的相关程度。移动终端在获取录音信号与播音信号之间的频域互相关函数时，可采用如下公式<1>:

$$r_xy = \mathrm{E}\ [X .* Y] \qquad <1>$$

其中，r_xy 为录音信号与播音信号之间的互相关函数，E[.]为期望运算符，.*用于对向量按元素逐个相乘。例如，$X=\{a_1, a_2, a_3, a_4\}$，$Y=\{b_1, b_2, b_3, b_4\}$，

则 $X .* Y=\{a_1 b_1, a_2 b_2, a_3 b_3, a_4 b_4\}$。

4022、移动终端获取播音信号的频域自相关函数。

其中，自相关函数用于表示信号与该信号的延迟信号之间的相关程度。移动终端在获取播音信号的频域自相关函数时，可采用如下公式<2>:

$$R_yy = E[Y(n) * Y'(n-k)] \quad <2>$$

其中，R_yy 为播音信号的频域自相关函数，符号 * 表示矩阵乘积运算，符号 ' 表示共轭转置运算，$Y(n)$ 为将第 n 时刻采集到的播音信号进行傅里叶变换得到的向量，$Y(n-k)$ 为将第 n-k 时刻采集到的播音信号进行傅里叶变换得到的向量，$k=[0, K_{max}]$，$k \in Z$，即 k 是整数，K_{max} 的取值大小决定系统的阶数。

4023、基于上述步骤 4021 中所获取到的录音信号与播音信号之间的频域互相关函数，以及步骤 4022 中所获取到的播音信号的频域自相关函数，移动终端可应用以下公式<3>，计算环路传递函数:

$$H_loop = R_yy \text{ ^-1 } * r_xy \quad <3>$$

其中，H_loop 为环路传递函数，符号 ^-1 表示矩阵求逆运算。

403、移动终端获取录音信号的功率谱和播音信号的功率谱。

对于录音信号，移动终端可应用以下公式<4>，计算录音信号的功率谱:

$$P_x = X(n).\text{^}2 \quad <4>$$

其中，P_x 为录音信号的功率谱，$X(n)$ 为将第 n 时刻采集到的录音信号进行傅里叶变换得到的向量，.^2 用于将 $X(n)$ 中的每个向量元素求平方。

例如，第 n 时刻采集到的录音信号 $X(n)=\{a_1, a_2, a_3, \ldots, a_n\}$，移动终端通过应用公式 $P_x = X(n).\text{^}2$，可得到 $P_x=\{a_1^2, a_2^2, a_3^2, \ldots, a_n^2\}$。

对于播音信号，移动终端可应用以下公式<5>，计算播音信号的功率谱:

$$P_y = Y(n).\text{^}2 \quad <5>$$

其中，P_y 为播音信号的功率谱，$Y(n)$ 为将第 n 时刻采集到的播音信号进行傅里叶变换得到的向量，.^2 用于将 $Y(n)$ 中的每个向量元素求平方。

例如，第 n 时刻采集到的播音信号 $Y(n)=\{b_1, b_2, b_3, \ldots, b_n\}$，移动终端通过应用公式 $P_y = Y(n).\text{^}2$，可得到 $P_y=\{b_1^2, b_2^2, b_3^2, \ldots, b_n^2\}$。

404、移动终端根据环路传递函数及播音信号，计算回声信号的估计值。

移动终端根据环路传递函数及播音信号，可应用如下公式<6>，计算回声信号的估计值:

$$E(n)=\sum_{k=0}^{K_{max}} H_loop(k)\cdot Y(n-k) \quad <6>$$

其中，$E(n)$为回声信号的估计值。

405、移动终端获取录音信号的功率特征值、播音信号的功率特征值及回声信号的功率特征值。

其中，录音信号的功率特征值为对录音信号的功率谱进行衡量的一种量度，可通过对录音信号的功率谱进行处理得到，在本实施例中，录音信号的功率特征值可用VAD_x表示。VAD_x为一种二值状态，具有 true 和 flase 两种状态。当VAD_x=true 时，表示录音信号较强；当VAD_x=flase 时，表示录音信号较弱。

播音信号的功率特征值为对播音信号的功率谱进行衡量的一种量度，可通过对播音信号的功率谱进行处理得到，在本实施例中，播音信号的功率特征值可用VAD_y表示。VAD_y为一种二值状态，具有 true 和 flase 两种状态。当VAD_y=true 时，表示播音信号较强；当VAD_y=flase 时，表示播音信号较弱。

回声信号的功率特征值为对回声信号的功率谱进行衡量的一种量度，在本实施例中，回声信号的功率特征值可用VAD_e表示。VAD_e为一种二值状态，具有 true 和 flase 两种状态。当VAD_e=true 时，表示回声信号较强；当VAD_e=flase 时，表示回音信号较弱。此处需要说明的是，在获取回声信号的功率特征值时，可预先根据回声信号的频谱估计值，计算一个回声信号的功率谱，进而通过对回声信号的功率谱进行处理，得到回声信号的功率特征值。此处计算得到的回声信号的功率谱为对回声信号的功率谱的一种估计，对于回声信号的功率谱是否为此处计算得到的回声信号的功率谱，需要通过下述步骤 406 进一步判断。

406、移动终端判断录音信号的功率特征值是否大于第一阈值、播音信号的功率特征值是否大于第二阈值、回声信号的功率特征值是否大于第三阈值，如果是，执行步骤 407。

为了将噪声信号及近端信号进行区分，本实施例应用信号检测和分类模块以及语音激活检测机制，并根据录音信号的功率特征值、回声信号的功率特征值以及播音信号的功率特征值，按时间区分近端信号（叠加有背景噪声）和非近端信号，以获取噪声信号的功率谱。具体判断时，移动终端需要判断录音信号的功率特征值是否大于第一阈值，播音信号的功率特征值是否大于第二阈值、回声信号的功率特征值是否大于第三阈值。其中，第一阈值、第二阈值、第三阈值为预设门限值，在本实施例中，第一阈值可用Tx表示，第二阈值可用Ty表

示，第三阈值可用*Te*表示，第一阈值、第二阈值、第三阈值取值越小，移动终端对噪声的反应越敏感，反之，移动终端仅当噪声能量非常大时，才对噪声作出反应。

上述判断过程，可用如下公式<7>表示:

$$VAD_y=\begin{cases}true & Py>Ty\\ flase & Py\le Ty\end{cases}$$

$$VAD_x=\begin{cases}true & Px>Tx\\ flase & Px\le Tx\end{cases}$$

$$VAD_e=\begin{cases}true & Pe>Te\\ flase & Pe\le Te\end{cases}$$

<7>

一般情况下，移动终端通过麦克风所采集到的录音信号中可能并不存在近端信号，为了进一步判断录音信号中是否存在近端信号，可采用如下公式<8>进行判断:

当VAD_y=flase，并且VAD_e=flase 时，$VAD_v = VAD_x$ <8>

也即是，当移动终端的扬声器并没用播放声音（即VAD_y=flase）时，且未检测到回声信号（即VAD_e=flase），则此时麦克风所收集到的录音信号即为近端信号，此时用户正在说话，否则说明用户未在说话。

在判断过程中，如果判断出录音信号的功率特征值大于第一阈值、播音信号的功率特征值大于第二阈值、回声信号的功率特征值大于第三阈值，则执行下述步骤 407; 如果判断出录音信号的功率特征值大于第一阈值、播音信号的功率特征值大于第二阈值、回声信号的功率特征值小于或等于第三阈值，或者，录音信号的功率特征值大于第一阈值、播音信号的功率特征值小于或等于第二阈值，则忽略本次获取到的录音信号和播音信号。

407、移动终端计算回声信号的频谱估计值的平方，作为回声信号的功率谱。

当录音信号的功率特征值是大于第一阈值、播音信号的功率特征值大于第二阈值、回声信号的功率特征值大于第三阈值时，移动终端通过计算回声信号的频谱估计值的平方，获取回声信号的功率谱，具体计算时，可应用以下公式<9>:

$$P_e = E(n).\hat{}2$$ <9>

其中，P_e为回声信号的功率谱。

408、移动终端判断录音信号的功率特征值是否小于第一阈值、回声信号的功率特征值是否小于第三阈值，如果是，执行步骤 409。

基于上述步骤 407，移动终端还将继续判断录音信号的功率特征值是否小于第一阈值、回声信号的功率特征值是否小于第三阈值，以获取噪声信号的功率谱。

在判断过程中，如果判断出录音信号的功率特征值小于第一阈值且回声信号的功率特征值小于第三阈值，则执行下述步骤 409；如果判断出录音信号的功率特征值小于第一阈值、回声信号的功率特征值大于或等于第三阈值，则忽略本次获取到的录音信号和播音信号。

409、移动终端将录音信号的功率谱减去回声信号的功率谱，作为噪声信号的功率谱。

当判断出录音信号的功率特征值小于第一阈值且回声信号的功率特征值小于第三阈值，此时可认为未检测到近端信号，也即是用户此时并未讲话，此时移动终端通过将录音信号的功率谱减去回声信号的功率谱，作为噪声信号的功率谱。具体实施时，可参见下述公式<10>:

$$P_n = P_x - P_e \qquad <10>$$

其中，P_n 为噪声信号的功率谱。

410、移动终端根据回声信号的功率谱和噪声信号的功率谱，计算频率加重系数。

移动终端在根据回声信号的功率谱和噪声信号的功率谱，计算频率加重系数时，可采用如下步骤 4101~4102:

4101、移动终端根据回声信号的功率谱及噪声信号的功率谱，构建语音可懂度函数。

在声学领域，语音可懂度函数（SII）具有多套标准，本实施例中采用 ASNI-S3.5 中的标准[4]进行计算，在标准[4]中，语音可懂度函数可以表示为以回声信号的功率谱及噪声信号的功率谱为自变量的函数。因此，当移动终端计算出回声信号的功率谱和噪声信号的功率谱之后，即可构建出语音可懂度函数。构建的语音可懂度函数可参见如下公式<11>:

$$SII(Pe_i, Pn_i) = \sum_{i=1}^{i_{max}} I_i \cdot \min\left\{1 - \frac{Pe_i - Pu_i - 10}{160}, 1\right\} \cdot \min\left(1, \max\left(0, \frac{Pe_i - Pd_i + 15}{30}\right)\right) \qquad <11>$$

其中，i_{max}为所拆分的频带总数，i为i_{max}内的任一频带，SII为语音可懂度函数，Pe_i为回声信号在第i个频带内的功率谱，Pn_i为噪声信号在第i个频带内的功率谱，Pu_i为标准语音强度在第i个频带内功率谱，I_i为分频带加权权重，Pd_i为中间变量，可用如下公式<12>表示：

$$Pd_i = \begin{cases} 10\cdot\log_{10}\left(10^{\frac{Pn_i}{10}}+\sum_{k}^{i-1}10^{\frac{\max\{Pn_k,Pe_k-24\}+3.32C_k\log(0.89\frac{f_i}{f_k})}{10}}\right) & i\neq 0 \\ \max\{Pn_k,Pe_k-24\} & i=0 \end{cases} \quad <12>$$

其中，f_k表示第i个频带内的第k个频点，C_k为中间变量，可用如下公式<13>表示：

$$C_k = 0.6\left(\max\{Pn_k,Pe_k-24\}+10\log_{10}f_k-6.353\right)-80 \quad <13>$$

其中，Pe_k为回声信号在第k个频点上的功率谱，Pn_k为噪声信号在第k个频点上的功率谱。

需要说明的是，上述Pu_i和I_i的具体取值可以参考ANSI-S3.5标准[4]中规定的数值，也可以由设计人员根据需要自行确定。

4102、在回声信号的功率谱保持不变的条件下，移动终端计算语音可懂度函数的极大值，从而获得频率加重系数。

在本实施例中，频率加重系数即为移动终端中加重滤波器的系数，用于调节移动终端输出的播音信号的频点幅值。在不同时刻，移动终端所计算出的频率加重系数是不同的。

通过观察上述步骤4101中所构建的语音可懂度函数可以看出，语音可懂度函数为以回声信号的功率谱和噪声信号的功率谱为自变量的函数，也即是，语音可懂度函数中的变量有两个，此时很难计算计算语音可懂度函数的极大值。为此，本实施例提供的方法作了一个近似计算，设定第n时刻的噪声信号的功率谱近似等于n-1时刻的噪声信号的功率谱，这样在计算第n时刻的频率加重系数时，移动终端可直接使用第n-1时刻所计算出的噪声信号的功率谱。通过采用该种处理方式，移动终端将语音可懂度函数转换为以回声信号的功率谱为自变量的函数。

为了提高用户扬声器播放的语音信号的语音可懂度，移动终端在将播音信号通过扬声器播放之前，还将采用加重滤波器对播音信号进行处理，以提高播音信号在指定频点上的幅值，增加播音信号的能量。受限于移动终端的尺寸，

扬声器播放的最大声功率具极大值，为了避免扬声器不会过载，本实施例在基于所构建的语音可懂度函数，计算频率加重系数时，假设加重滤波器增强前后的回声信号功率谱保持不变，再计算语音可懂度函数的极大值，在数学上这一方法称为有约束条件下求极值问题。该极值问题，可用如下公式<14>表示:

$$\max \sum_i SII(Pe_i, Pn_i)$$
$$\sum_i Pe_i' = \sum_i Pe_i \qquad <14>$$

其中，Pe_i为增强前的回声信号在第 i 个频点上的功率谱，Pe_i'为增强后的回声信号在第 i 个频点上的功率谱，公式$\sum_i Pe_i' = \sum_i Pe_i$保证了增强前后的回声信号功率谱不变，从而确保扬声器不会过载。

需要注意的是，通过加重滤波器处理后的信号为电信号，电信号需通过扬声器转换后才变成声波。由于不同型号的移动终端的扬声器的输出频率响应是不同的，如果要获取不同移动终端的扬声器的输出频率响应，就需要分别测量每个移动终端的扬声器，并在运行时进行校正补偿，由此将产生硬件碎片化问题。为了避免该问题，本实施例提供的方法将采用如下方法，以避免对扬声器频响的直接测量。

通过对上述公式<6>的观察可以发现，$E(n)$与$Y(n)$可通过环路传递函数$H_{_loop}$建立起映射关系。本实施例将扬声器的频率响应记为H_{spk}，将麦克风的频率响应记为H_{mic}，根据公式<6>，则可得出:

$$Pe = Py \cdot \left|H_{_loop}\right|^2 = Pz \cdot \left|H_{spk}\right|^2 \cdot \left|H_{mic}\right|^2 = Pz \cdot \frac{\left|H_{_loop}\right|^2}{\left|W\right|^2} \qquad <15>$$

对上述公式<15>，可将公式<14>求极值问题转化为求偏导的问题，通过计算公式<15>的偏导数，可得到语音可懂度函数的拐点，具体过程可参见下述公式<16>:

$$\frac{\partial SII}{\partial Pe_i} = 0 \Rightarrow \frac{\partial SII}{\partial \left(Pz_i \cdot \frac{\left|H_{loop}\right|^2}{\left|W\right|^2} \right)} = 0 \Rightarrow \frac{\partial SII}{\partial \left(Py_i \cdot \left|W\right|^2 \cdot \frac{\left|H_{loop}\right|^2}{\left|W\right|^2} \right)} = 0 \Rightarrow$$
$$\frac{\partial SII}{\partial \left(Py_i \cdot \left|H_{loop}\right|^2 \right)} = 0 \Rightarrow \frac{\partial SII}{\partial \left(Py_i \cdot \left|H_{spk}\right|^2 \cdot \left|H_{mic}\right|^2 \cdot \left|W\right|^2 \right)} = 0 \qquad <16>$$

其中，$|W|^2$为频率加重系数，$|H_{_loop}|^2$可通过上述公式<3>得到，Py_i可通过上述公式<5>得到，SII 可通过公式<11>得到。

通过对上述公式<16>进行计算，可得到当前时刻的$|W|^2$。

411、基于频率加重系数，移动终端对播音信号的频点幅值进行调节。

基于所确定的频率加重系数，移动终端通过动态地跟踪并调整语音可懂度函数，以实现对噪声信号的功率谱P_n、回声信号的功率谱P_e的变化的自动适应。

412、移动终端输出调节后的播音信号。

为了提高移动终端当前时刻所输出的播音信号的准确性，移动终端将结合当前时刻之前的一段时间内所输出的播音信号及相应的频率加重系数，根据下述公式<17>，确定当前时刻要输出的播音信号。

$$z(n)=\sum_{k=0}^{K_{\max}} w(k)\cdot y(n-k) \qquad <17>$$

其中，$z(n)$为输出的播音信号，$w(k)$为第 n 时刻计算出的频率加重系数在时域上的对应值，$K_{\max}$等于加重滤波器 W 的阶数，$y(n-k)$为加重前的播音信号在第 n-k 时刻的值。

由于本步骤中移动终端所输出的调节后的播音信号能够掩蔽噪声信号，因此，当收听到待调节后的播音信号后，用户能够听清该播音信号的内容。

图 5 示出了本发明实施例提供的语音信号的处理方法对应的信号流，由图 5 可知，当基于所获取到的录音信号 X 和播音信号 Y，移动终端根据录音信号和播音信号之间的频域互相关函数r_xy及播音信号的频域自相关函数R_yy，可计算出环路传递函数 $H_1oop=R_yy\ \hat{}\ -1*\ r_xy$。移动终端根据播音信号及环路传递函数，可计算出回声信号的估计值$E(n)=H_loop\cdot Y(n)$，进一步地，移动终端根据录音信号的功率特征值、播音信号的功率特征值及回声信号的功率特征值，并采用语音激活检测机制，计算出回声信号的功率谱和噪声信号的功率谱，进而通过计算语音可懂度函数的极大值，获取频率加重系数，最后基于频率加重系数，采用加重滤波器对播音信号的频点幅值进行调节，并将调节后的播音信号输出。

图 6 示出了本发明另一实施例提供的语音信号的处理方法的流程图。该方法可以通过软件实现。当语音即时通讯应用启动后，移动终端会定时地获取麦克风从近端采集到的录音信号 x 与对端发送的播音信号 y，并计算出录音信号的

功率谱P_x、播音信号的功率谱P_y，进而基于公式<3>计算出环路传递函数H_loop。当确定了环路传递函数之后，移动终端可根据公式<6>，计算出回声信号的估计值$E(n)$。另外，由于回声信号、近端语音信号、噪声信号三者被同一个麦克风拾取，时间上存在重叠，因此，需要对录音信号进行分类，进而根据公式<9>计算出回声功率谱P_e，根据公式<10>计算出噪声功率谱P_n。之后，根据回声信号的功率谱和噪声信号的功率谱，构建语音可懂度函数SII，通过计算语音可懂度函数 SII 的极大值，可得到频谱加重系数 W。最后根据公式<17>计算输出增强后的播音信号送给扬声器，由扬声器转换成声音进行播放。

需要说明的是，上述方法可以在语音即时通讯应用层面中实现，也可以在操作系统层面实现，也可以固化在硬件芯片的固件(firmware)中实现。无论是在这三个层面的哪一层面上实现，本发明实施例提供的语音数据的处理方法均适用的，区别仅在于同一语音数据的处理方法具体是运行于移动终端系统中的哪个层面。

需要说明的是，上文以移动终端为例描述了本发明，本领域的技术人员可以理解，本发明还可以应用于其他终端设备，例如桌面计算机等。另外，上文的播音信号可以是从对端接收的，例如终端设备通过有线或无线网络从其他终端设备（即对端设备）接收的语音信号，上文的播音信号也可以是终端设备本地存储的语音信号。此外，上文以语音即时通讯应用为例进行了说明，本领域的技术人员可以理解，上文的语音即时通讯应用可以替换为任何其它语音播放应用。

需要说明的是，上述方法不仅可用于提高语音可懂度之外，还可以用于提高其他内容的音频信号。例如：可根据不同的环境噪声自动地对铃声、闹钟的提示音做增强，使得增强后的提示声音能更清楚地被用户听到，以达到克服环境噪声干扰的目的。

需要说明的是，上述方法除了用于对抗噪声场景之外，还可用于对抗非噪声的环境。例如：A 与 B 两个人同时在相近的距离内拨打电话，其中，A 与 a 通话、B 与 b 通话。由于 A 与 B 两个人距离很近，因而 A 的说话声将会对 B 的收听形成干扰，同时 B 的说话声也对 A 的收听形成干扰。本发明实施提供的方法同样可用于这种语音竞争场景，在该场景下， A 侧的移动终端将当把 B 的语音作为噪声信号，把 a 的语音作为需要做增强的信号；同理，B 侧的移动终端将把 A 作为噪声信号，把 b 的语音作为需要做增强的信号。

本发明实施例提供的方法，在确保扬声器不过载，且不破坏原始播音信号的动态幅度的前提下，自动根据噪声信号与播音信号的频率分布，调整播音信号的频点幅值，明显提高了语音可懂度。

参见图 7，本发明实施例提供了一种语音信号的处理装置的结构示意图，该装置包括:

采集模块 701，用于从近端采集录音信号，录音信号中至少包括噪声信号及回声信号;

接收模块 702，用于接收对端发送的播音信号;

第一计算模块 703，用于根据录音信号和播音信号，计算环路传递函数;

第二计算模块 704，用于计算录音信号的功率谱;

第三计算模块 705，用于根据录音信号的功率谱、播音信号及环路传递函数，计算回声信号的功率谱和噪声信号的功率谱;

第四计算模块 706，用于根据回声信号的功率谱和噪声信号的功率谱，计算频率加重系数;

调节模块 707，用于基于频率加重系数，对播音信号的频点幅值进行调节;

输出模块 708，用于输出调节后的播音信号。

在本发明的另一个实施例中，第一计算模块 703，用于计算录音信号与播音信号之间的频域互相关函数；计算播音信号的频域自相关函数；根据录音信号与播音信号之间的频域互相关函数以及播音信号的频域自相关函数，计算环路传递函数。

在本发明的另一个实施例中，第二计算模块 704，用于对于录音信号，应用以下公式，计算录音信号的功率谱:

$$P_x = X(n).\hat{}2$$

其中，P_x 为录音信号的功率谱，$X(n)$ 为将第 n 时刻采集到的录音信号进行傅里叶变换得到的向量，.^2 用于将 $X(n)$ 中的每个向量元素求平方。

在本发明的另一个实施例中，第三计算模块 705，用于根据环路传递函数及播音信号，计算回声信号的频谱估计值；计算回声信号的频谱估计值的平方，得到回声信号的功率谱；将录音信号的功率谱减去回声信号的功率谱，得到噪声信号的功率谱。

在本发明的另一个实施例中，该装置还包括:

第五计算模块，用于计算录音信号的功率特征值、播音信号的功率特征值及回声信号的功率特征值；

第一判断模块，用于判断录音信号的功率特征值是否大于第一阈值、播音信号的功率特征值是否大于第二阈值、回声信号的功率特征值是否大于第三阈值；

第三计算模块 705，用于当录音信号的功率特征值大于第一阈值、播音信号的功率值大于第二阈值且回声信号的功率特征值大于第三阈值时，计算回声信号的频谱估计值的平方，得到回声信号的功率谱。

在本发明的另一个实施例中，该装置还包括：

第二判断模块，用于判断录音信号的功率特征值是否小于第一阈值、回声信号的功率特征值是否小于第三阈值；

第三计算模块 705，用于当录音信号的功率特征值小于第一阈值且回声信号的功率特征值小于第三阈值时，将录音信号的功率谱减去回声信号的功率谱，得到噪声信号的功率谱。

在本发明的另一个实施例中，第四计算模块 706，用于根据回声信号的功率谱及噪声信号的功率谱，构建语音可懂度函数；在回声信号的功率谱保持不变的条件下，根据语音可懂度函数的极大值，得到频率加重系数。

综上，本发明实施例提供的装置，在确保扬声器不过载，且不破坏原始播音信号的动态幅度的前提下，自动根据噪声信号与播音信号的频率分布，调整播音信号的频点幅值，明显提高了语音可懂度。

参见图 8，其示出了本发明实施例所涉及的语音信号的处理终端的结构示意图，该终端可以用于实施上述实施例中提供的语音信号的处理方法。具体来讲：

终端 800 可以包括 RF（Radio Frequency，射频）电路 110、包括有一个或一个以上计算机可读存储介质的存储器 120、输入单元 130、显示单元 140、传感器 150、音频电路 160、WiFi（Wireless Fidelity，无线保真）模块 170、包括有一个或者一个以上处理核心的处理器 180、以及电源 190 等部件。本领域技术人员可以理解，图 8 中示出的终端结构并不构成对终端的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置。其中：

RF 电路 110 可用于收发信息或通话过程中，信号的接收和发送，特别地，将基站的下行信息接收后，交由一个或者一个以上处理器 180 处理；另外，将

涉及上行的数据发送给基站。通常，RF 电路 110 包括但不限于天线、至少一个放大器、调谐器、一个或多个振荡器、用户身份模块（SIM）卡、收发信机、耦合器、LNA（Low Noise Amplifier，低噪声放大器）、双工器等。此外，RF 电路 110 还可以通过无线通信与网络和其他设备通信。所述无线通信可以使用任一通信标准或协议，包括但不限于 GSM(Global System of Mobile communication，全球移动通讯系统)、GPRS(General Packet Radio Service，通用分组无线服务)、CDMA(Code Division Multiple Access，码分多址)、WCDMA(Wideband Code Division Multiple Access，宽带码分多址)、LTE(Long Term Evolution,长期演进)、电子邮件、SMS(Short Messaging Service，短消息服务)等。

存储器 120 可用于存储软件程序以及模块，处理器 180 通过运行存储在存储器 120 的软件程序以及模块，从而执行各种功能应用以及数据处理。存储器 120 可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序（比如声音播放功能、图像播放功能等）等；存储数据区可存储根据终端 800 的使用所创建的数据（比如音频数据、电话本等）等。此外，存储器 120 可以包括高速随机存取存储器，还可以包括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。相应地，存储器 120 还可以包括存储器控制器，以提供处理器 180 和输入单元 130 对存储器 120 的访问。

输入单元 130 可用于接收输入的数字或字符信息，以及产生与用户设置以及功能控制有关的键盘、鼠标、操作杆、光学或者轨迹球信号输入。具体地，输入单元 130 可包括触敏表面 131 以及其他输入设备 132。触敏表面 131，也称为触摸显示屏或者触控板，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敏表面 131 上或在触敏表面 131 附近的操作），并根据预先设定的程式驱动相应的连接装置。可选的，触敏表面 131 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，再送给处理器 180，并能接收处理器 180 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敏表面 131。除了触敏表面 131，输入单元 130 还可以包括其他输入设备 132。具体地，其他输入设备 132 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。

显示单元 140 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 800 的各种图形用户接口，这些图形用户接口可以由图形、文本、图标、视频和其任意组合来构成。显示单元 140 可包括显示面板 141，可选的，可以采用 LCD(Liquid Crystal Display，液晶显示器)、OLED(Organic Light-Emitting Diode, 有机发光二极管)等形式来配置显示面板 141。进一步的，触敏表面 131 可覆盖显示面板 141，当触敏表面 131 检测到在其上或附近的触摸操作后，传送给处理器 180 以确定触摸事件的类型，随后处理器 180 根据触摸事件的类型在显示面板 141 上提供相应的视觉输出。虽然在图 8 中，触敏表面 131 与显示面板 141 是作为两个独立的部件来实现输入和输入功能，但是在某些实施例中，可以将触敏表面 131 与显示面板 141 集成而实现输入和输出功能。

终端 800 还可包括至少一种传感器 150，比如光传感器、运动传感器以及其他传感器。具体地，光传感器可包括环境光传感器及接近传感器，其中，环境光传感器可根据环境光线的明暗来调节显示面板 141 的亮度，接近传感器可在终端 800 移动到耳边时，关闭显示面板 141 和/或背光。作为运动传感器的一种，重力加速度传感器可检测各个方向上（一般为三轴）加速度的大小，静止时可检测出重力的大小及方向，可用于识别手机姿态的应用（比如横竖屏切换、相关游戏、磁力计姿态校准）、振动识别相关功能（比如计步器、敲击）等；至于终端 800 还可配置的陀螺仪、气压计、湿度计、温度计、红外线传感器等其他传感器，在此不再赘述。

音频电路 160、扬声器 161，传声器 162 可提供用户与终端 800 之间的音频接口。音频电路 160 可将接收到的音频数据转换后的电信号，传输到扬声器 161，由扬声器 161 转换为声音信号输出；另一方面，传声器 162 将收集的声音信号转换为电信号，由音频电路 160 接收后转换为音频数据，再将音频数据输出处理器 180 处理后，经 RF 电路 110 以发送给比如另一终端，或者将音频数据输出至存储器 120 以便进一步处理。音频电路 160 还可能包括耳塞插孔，以提供外设耳机与终端 800 的通信。

WiFi 属于短距离无线传输技术，终端 800 通过 WiFi 模块 170 可以帮助用户收发电子邮件、浏览网页和访问流式媒体等，它为用户提供了无线的宽带互联网访问。虽然图 8 示出了 WiFi 模块 170，但是可以理解的是，其并不属于终端 800 的必须构成，完全可以根据需要在不改变发明的本质的范围内而省略。

处理器 180 是终端 800 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存储器 120 内的软件程序和/或模块，以及调用存储在存储器 120 内的数据，执行终端 800 的各种功能和处理数据，从而对手机进行整体监控。可选的，处理器 180 可包括一个或多个处理核心；可选的，处理器 180 可集成应用处理器和调制解调处理器，其中，应用处理器主要处理操作系统、用户界面和应用程序等，调制解调处理器主要处理无线通信。可以理解的是，上述调制解调处理器也可以不集成到处理器 180 中。

终端 800 还包括给各个部件供电的电源 190（比如电池），优选的，电源可以通过电源管理系统与处理器 180 逻辑相连，从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。电源 190 还可以包括一个或一个以上的直流或交流电源、再充电系统、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。

尽管未示出，终端 800 还可以包括摄像头、蓝牙模块等，在此不再赘述。具体在本实施例中，终端 800 的显示单元是触摸屏显示器，终端 800 还包括有存储器，以及一个或者一个以上的程序，其中一个或者一个以上程序存储于存储器中，且经配置以由一个或者一个以上处理器执行。

所述一个或者一个以上程序包含用于执行以下操作的指令:

获取录音信号和语音信号，所述录音信号中至少包括噪声信号及回声信号;

根据所述录音信号和所述语音信号，计算环路传递函数;

根据所述录音信号、所述语音信号及所述环路传递函数，计算所述回声信号的功率谱和所述噪声信号的功率谱;

根据所述回声信号的功率谱和所述噪声信号的功率谱，计算频率加重系数;

基于所述频率加重系数，对所述语音信号的频点幅值进行调节;

输出调节后的语音信号。

假设上述为第一种可能的实施方式，则在第一种可能的实施方式作为基础而提供的第二种可能的实施方式中，录音信号是使用终端设备的麦克风采集的声音信号。

在第一种可能的实施方式作为基础而提供的第三种可能的实施方式中，输出调节后的语音信号包括通过扬声器播放调节后的语音信号，其中所述语音信号是终端设备通过网络接收的或本地存储的要通过扬声器播放的播音信号。

在第三种可能的实施方式作为基础而提供的第四种可能的实施方式中，终

端的存储器中，还包含用于执行以下操作的指令:

根据录音信号和语音信号，计算环路传递函数，包括:

计算录音信号与播音信号之间的频域互相关函数;

计算播音信号的频域自相关函数;

根据录音信号与播音信号之间的频域互相关函数以及播音信号的频域自相关函数，计算环路传递函数;

或者，终端的存储器中，还包含用于执行以下操作的指令:

计算录音信号的功率谱，包括:

对于录音信号，应用以下公式，计算录音信号的功率谱:

$$P_x = X(n).\text{^}2$$

其中，P_x为录音信号的功率谱，$X(n)$为将第 n 时刻采集到的录音信号进行傅里叶变换得到的向量，.^2用于将$X(n)$中的每个向量元素求平方。

在第三种可能的实施方式作为基础而提供的第五种可能的实施方式中，终端的存储器中，还包含用于执行以下操作的指令:

根据录音信号、语音信号及环路传递函数，计算回声信号的功率谱和噪声信号的功率谱，包括:

计算所述录音信号的功率谱;

根据环路传递函数及播音信号，计算回声信号的频谱估计值;

计算回声信号的频谱估计值的平方，得到回声信号的功率谱;

将录音信号的功率谱减去回声信号的功率谱，得到噪声信号的功率谱。

在第五种可能的实施方式作为基础而提供的第六种可能的实施方式中，终端的存储器中，还包含用于执行以下操作的指令:

计算回声信号的频谱估计值的平方，得到回声信号的功率谱之前，还包括:

计算录音信号的功率特征值、播音信号的功率特征值及回声信号的功率特征值;

判断录音信号的功率特征值是否大于第一阈值、播音信号的功率特征值是否大于第二阈值、回声信号的功率特征值是否大于第三阈值;

当录音信号的功率特征值大于第一阈值、播音信号的功率值大于第二阈值且回声信号的功率特征值大于第三阈值时，执行计算回声信号的频谱估计值的平方，得到回声信号的功率谱的步骤。

在第六种可能的实施方式作为基础而提供的第七种可能的实施方式中，终

端的存储器中，还包含用于执行以下操作的指令:

将录音信号的功率谱减去回声信号的功率谱，得到噪声信号的功率谱之前，还包括:

判断录音信号的功率特征值是否小于第一阈值、回声信号的功率特征值是否小于第三阈值;

当录音信号的功率特征值小于第一阈值且回声信号的功率特征值小于第三阈值时，执行将录音信号的功率谱减去回声信号的功率谱，得到噪声信号的功率谱的步骤。

在第三种可能的实施方式作为基础而提供的第八种可能的实施方式中，终端的存储器中，还包含用于执行以下操作的指令:

根据回声信号的功率谱、噪声信号的功率谱，计算频率加重系数，包括:

根据回声信号的功率谱及噪声信号的功率谱，构建语音可懂度函数;

在回声信号的功率谱保持不变的条件下，根据语音可懂度函数的极大值，得到频率加重系数。

本发明实施例提供的终端，在确保扬声器不过载，且不破坏原始播音信号的动态幅度的前提下，自动根据噪声信号与播音信号的频率分布，调整播音信号的频点幅值，明显提高了语音可懂度。

本发明实施例还提供了一种计算机可读存储介质，该计算机可读存储介质可以是上述实施例中的存储器中所包含的计算机可读存储介质；也可以是单独存在，未装配入终端中的计算机可读存储介质。该计算机可读存储介质存储有一个或者一个以上程序，该一个或者一个以上程序被一个或者一个以上的处理器用来执行语音信号的处理方法，该方法包括:

获取录音信号和语音信号，所述录音信号中至少包括噪声信号及回声信号;

根据所述录音信号和所述语音信号，计算环路传递函数;

根据所述录音信号、所述语音信号及所述环路传递函数，计算所述回声信号的功率谱和所述噪声信号的功率谱;

根据所述回声信号的功率谱和所述噪声信号的功率谱，计算频率加重系数;

基于所述频率加重系数，对所述语音信号的频点幅值进行调节;

输出调节后的语音信号。

假设上述为第一种可能的实施方式，则在第一种可能的实施方式作为基础

而提供的第二种可能的实施方式中，录音信号是使用终端设备的麦克风采集的声音信号。

在第一种可能的实施方式作为基础而提供的第三种可能的实施方式中，输出调节后的语音信号包括通过扬声器播放调节后的语音信号，其中所述语音信号是终端设备通过网络接收的或本地存储的要通过扬声器播放的播音信号。

在第三种可能的实施方式作为基础而提供的第四种可能的实施方式中，终端的存储器中，还包含用于执行以下操作的指令:

根据录音信号和语音信号，计算环路传递函数，包括:

计算录音信号与播音信号之间的频域互相关函数;

计算播音信号的频域自相关函数;

根据录音信号与播音信号之间的频域互相关函数以及播音信号的频域自相关函数，计算环路传递函数;

或者，终端的存储器中，还包含用于执行以下操作的指令:

计算录音信号的功率谱，包括:

对于录音信号，应用以下公式，计算录音信号的功率谱:

$$P_x = X(n).\text{^}2$$

其中，P_x为录音信号的功率谱，$X(n)$为将第 n 时刻采集到的录音信号进行傅里叶变换得到的向量，.^2用于将$X(n)$中的每个向量元素求平方。

在第三种可能的实施方式作为基础而提供的第五种可能的实施方式中，终端的存储器中，还包含用于执行以下操作的指令:

根据录音信号、语音信号及环路传递函数，计算回声信号的功率谱和噪声信号的功率谱，包括:

计算所述录音信号的功率谱;

根据环路传递函数及播音信号，计算回声信号的频谱估计值;

计算回声信号的频谱估计值的平方，得到回声信号的功率谱;

将录音信号的功率谱减去回声信号的功率谱，得到噪声信号的功率谱。

在第五种可能的实施方式作为基础而提供的第六种可能的实施方式中，终端的存储器中，还包含用于执行以下操作的指令:

计算回声信号的频谱估计值的平方，得到回声信号的功率谱之前，还包括:

获取录音信号的功率特征值、播音信号的功率特征值及回声信号的功率特征值;

判断录音信号的功率特征值是否大于第一阈值、播音信号的功率特征值是否大于第二阈值、回声信号的功率特征值是否大于第三阈值;

当录音信号的功率特征值大于第一阈值、播音信号的功率值大于第二阈值且回声信号的功率特征值大于第三阈值时，执行计算回声信号的频谱估计值的平方，得到回声信号的功率谱的步骤。

在第六种可能的实施方式作为基础而提供的第七种可能的实施方式中，终端的存储器中，还包含用于执行以下操作的指令:

将录音信号的功率谱减去回声信号的功率谱，得到噪声信号的功率谱之前，还包括:

判断录音信号的功率特征值是否小于第一阈值、回声信号的功率特征值是否小于第三阈值;

当录音信号的功率特征值小于第一阈值且回声信号的功率特征值小于第三阈值时，执行将录音信号的功率谱减去回声信号的功率谱，得到噪声信号的功率谱的步骤。

在第三种可能的实施方式作为基础而提供的第八种可能的实施方式中，终端的存储器中，还包含用于执行以下操作的指令:

根据回声信号的功率谱、噪声信号的功率谱，计算频率加重系数，包括:

根据回声信号的功率谱及噪声信号的功率谱，构建语音可懂度函数;

在回声信号的功率谱保持不变的条件下，根据语音可懂度函数的极大值，得到频率加重系数。

本发明实施例提供的计算机可读存储介质，在确保扬声器不过载，且不破坏原始播音信号的动态幅度的前提下，自动根据噪声信号与播音信号的频率分布，调整播音信号的频点幅值，明显提高了语音可懂度。

本发明实施例中提供了一种图形用户接口，该图形用户接口用在语音信号的处理终端上，该执行语音信号的处理终端包括触摸屏显示器、存储器和用于执行一个或者一个以上的程序的一个或者一个以上的处理器；该图形用户接口包括:

获取录音信号和语音信号，录音信号中至少包括噪声信号及回声信号;

根据录音信号和语音信号，计算环路传递函数;

根据录音信号、语音信号及环路传递函数，计算回声信号的功率谱和噪声

信号的功率谱;

根据回声信号的功率谱和噪声信号的功率谱，计算频率加重系数;

基于频率加重系数，对语音信号的频点幅值进行调节;

输出调节后的语音信号。

本发明实施例提供的图形用户接口，在确保扬声器不过载，且不破坏原始播音信号的动态幅度的前提下，自动根据噪声信号与播音信号的频率分布，调整播音信号的频点幅值，明显提高了语音可懂度。

需要说明的是：上述实施例提供的语音信号的处理装置在处理语音信号时，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将语音信号的处理装置的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。另外，上述实施例提供的语音信号的处理装置与语音信号的处理方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成，也可以通过程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，上述提到的存储介质可以是只读存储器，磁盘或光盘等。

以上所述仅为本发明的较佳实施例，并不用以限制本发明，凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

权 利 要 求

1、一种语音信号的处理方法，包括:

获取录音信号和要输出的语音信号，所述录音信号中至少包括噪声信号及回声信号;

根据所述录音信号和所述语音信号，计算环路传递函数;

根据所述录音信号、所述语音信号及所述环路传递函数，计算所述回声信号的功率谱和所述噪声信号的功率谱;

根据所述回声信号的功率谱和所述噪声信号的功率谱，计算频率加重系数;

基于所述频率加重系数，对所述语音信号的频点幅值进行调节;

输出调节后的语音信号。

2，根据权利要求 1 所述的方法，其中，所述录音信号是使用终端设备的麦克风采集的声音信号。

3、根据权利要求 1 所述的方法，其中，输出调节后的语音信号包括通过终端设备的扬声器播放调节后的语音信号，其中所述语音信号是终端设备通过网络接收的或本地存储的要通过扬声器播放的播音信号。

4、根据权利要求 3 所述的方法，其中，所述根据所述录音信号和所述语音信号，计算环路传递函数，包括:

计算所述录音信号与所述播音信号之间的频域互相关函数;

计算所述播音信号的频域自相关函数;

根据所述录音信号与所述播音信号之间的频域互相关函数以及所述播音信号的频域自相关函数计算所述环路传递函数。

5、根据权利要求 3 所述的方法，其中，所述根据所述录音信号、所述语音信号及所述环路传递函数，计算所述回声信号的功率谱和所述噪声信号的功率谱，包括:

计算所述录音信号的功率谱;

根据所述环路传递函数及所述播音信号，计算所述回声信号的频谱估计值;

计算所述回声信号的频谱估计值的平方，得到所述回声信号的功率谱;

将所述录音信号的功率谱减去所述回声信号的功率谱，得到所述噪声信号的功率谱。

6、根据权利要求 5 所述的方法，还包括:

计算所述录音信号的功率特征值、所述播音信号的功率特征值及所述回声信号的功率特征值；和

判断所述录音信号的功率特征值是否大于第一阈值、所述播音信号的功率特征值是否大于第二阈值、所述回声信号的功率特征值是否大于第三阈值，

其中，所述计算所述回声信号的频谱估计值的平方，得到所述回声信号的功率谱包括:

当所述录音信号的功率特征值大于所述第一阈值、所述播音信号的功率值大于所述第二阈值且所述回声信号的功率特征值大于所述第三阈值时，计算所述回声信号的频谱估计值的平方，得到所述回声信号的功率谱。

7、根据权利要求 6 所述的方法，还包括:

判断所述录音信号的功率特征值是否小于所述第一阈值、所述回声信号的功率特征值是否小于所述第三阈值，

其中，所述将所述录音信号的功率谱减去所述回声信号的功率谱，得到所述噪声信号的功率谱包括:

当所述录音信号的功率特征值小于所述第一阈值且所述回声信号的功率特征值小于所述第三阈值时，将所述录音信号的功率谱减去所述回声信号的功率谱，得到所述噪声信号的功率谱。

8、根据权利要求 3 所述的方法，其中，所述根据所述回声信号的功率谱和所述噪声信号的功率谱，计算频率加重系数，包括:

根据所述回声信号的功率谱及所述噪声信号的功率谱，构建语音可懂度函数;

在所述回声信号的功率谱保持不变的条件下，根据所述语音可懂度函数的极大值，得到所述频率加重系数。

9、根据权利要求 1 所述的方法，其中所述终端设备包括加重滤波器、扬声器和麦克风，所述频率加重系数表示语音信号经过加重滤波器和扬声器后被麦克风拾取的比例。

10、一种语音信号的处理装置，包括:

至少一个处理器；和

存储器，其中所述存储器存储有程序指令，所述指令当由所述处理器执行时，配置所述装置执行下述操作:

获取录音信号和语音信号，所述录音信号中至少包括噪声信号及回声信号;

根据所述录音信号和所述语音信号，计算环路传递函数;

根据所述录音信号、所述语音信号及所述环路传递函数，计算所述回声信号的功率谱和所述噪声信号的功率谱;

根据所述回声信号的功率谱和所述噪声信号的功率谱，计算频率加重系数;

基于所述频率加重系数，对所述语音信号的频点幅值进行调节;

输出调节后的语音信号。

11、根据权利要求 10 所述的装置，其中，所述录音信号是使用终端设备的麦克风采集的声音信号。

12、根据权利要求 10 所述的装置，其中，输出调节后的语音信号包括通过扬声器播放调节后的语音信号，其中所述语音信号是终端设备通过网络接收的或本地存储的要通过扬声器播放的播音信号。

13、根据权利要求 12 所述的装置，其中，所述根据所述录音信号和所述语音信号，计算环路传递函数，包括:

计算所述录音信号与所述播音信号之间的频域互相关函数;

计算所述播音信号的频域自相关函数;

根据所述录音信号与所述播音信号之间的频域互相关函数以及所述播音信号的频域自相关函数计算所述环路传递函数。

14、根据权利要求 12 所述的装置，其中，所述根据所述录音信号、所述语音信号及所述环路传递函数，计算所述回声信号的功率谱和所述噪声信号的功率谱，包括:

计算所述录音信号的功率谱;

根据所述环路传递函数及所述播音信号，计算所述回声信号的频谱估计值;

计算所述回声信号的频谱估计值的平方，得到所述回声信号的功率谱;

将所述录音信号的功率谱减去所述回声信号的功率谱，得到所述噪声信号的功率谱。

15、根据权利要求 12 所述的装置，其中，所述装置还被配置为:

计算所述录音信号的功率特征值、所述播音信号的功率特征值及所述回声信号的功率特征值;

判断所述录音信号的功率特征值是否大于第一阈值、所述播音信号的功率特征值是否大于第二阈值、所述回声信号的功率特征值是否大于第三阈值;

当所述录音信号的功率特征值大于所述第一阈值、所述播音信号的功率值

大于所述第二阈值且所述回声信号的功率特征值大于所述第三阈值时，计算所述回声信号的频谱估计值的平方，得到所述回声信号的功率谱。

16、根据权利要求 12 所述的装置，其中，所述装置还被配置为：

判断所述录音信号的功率特征值是否小于所述第一阈值、所述回声信号的功率特征值是否小于所述第三阈值；

当所述录音信号的功率特征值小于所述第一阈值且所述回声信号的功率特征值小于所述第三阈值时，将所述录音信号的功率谱减去所述回声信号的功率谱，得到所述噪声信号的功率谱。

17、根据权利要求 12 所述的装置，其中，所述根据所述回声信号的功率谱和所述噪声信号的功率谱，计算频率加重系数，包括：

根据所述回声信号的功率谱及所述噪声信号的功率谱，构建语音可懂度函数；

在所述回声信号的功率谱保持不变的条件下，根据所述语音可懂度函数的极大值，得到所述频率加重系数。

18、一种计算机可读存储介质，所述存储介质存储有程序指令，所述指令当由计算装置的处理器执行时，配置所述装置执行根据权利要求 1-9 中任一项所述的方法。

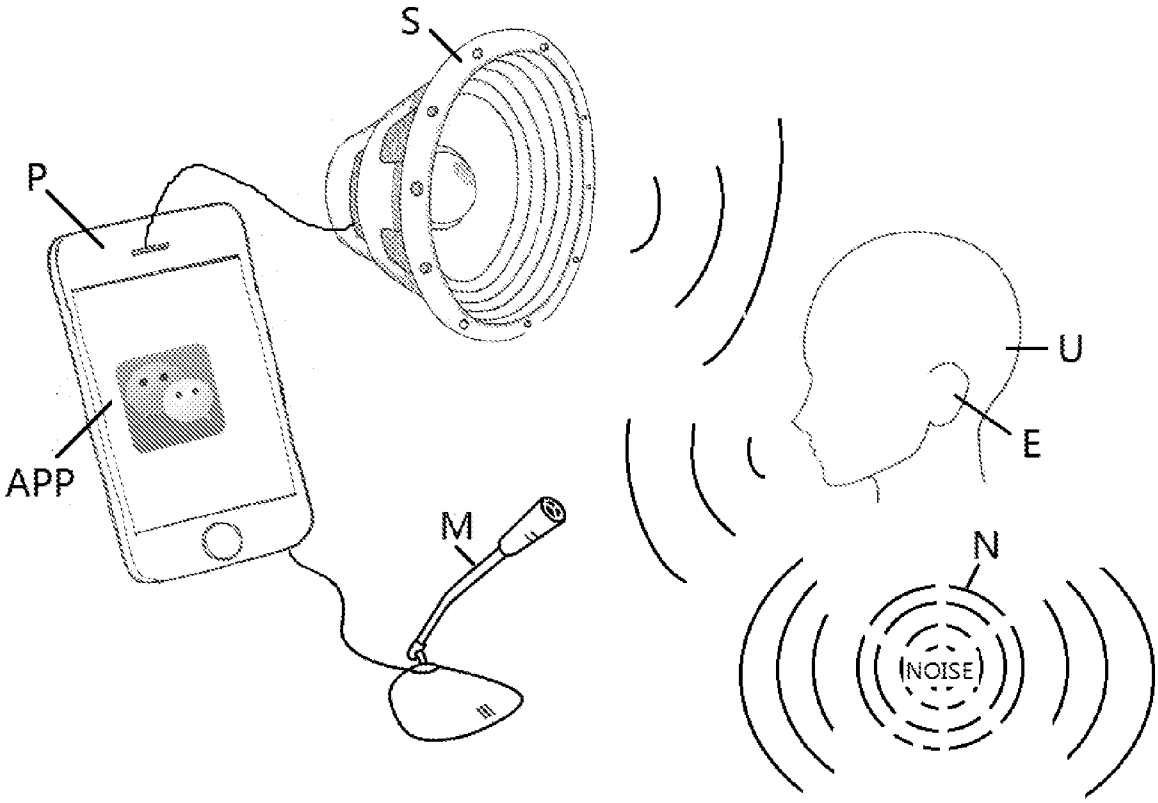


图 1

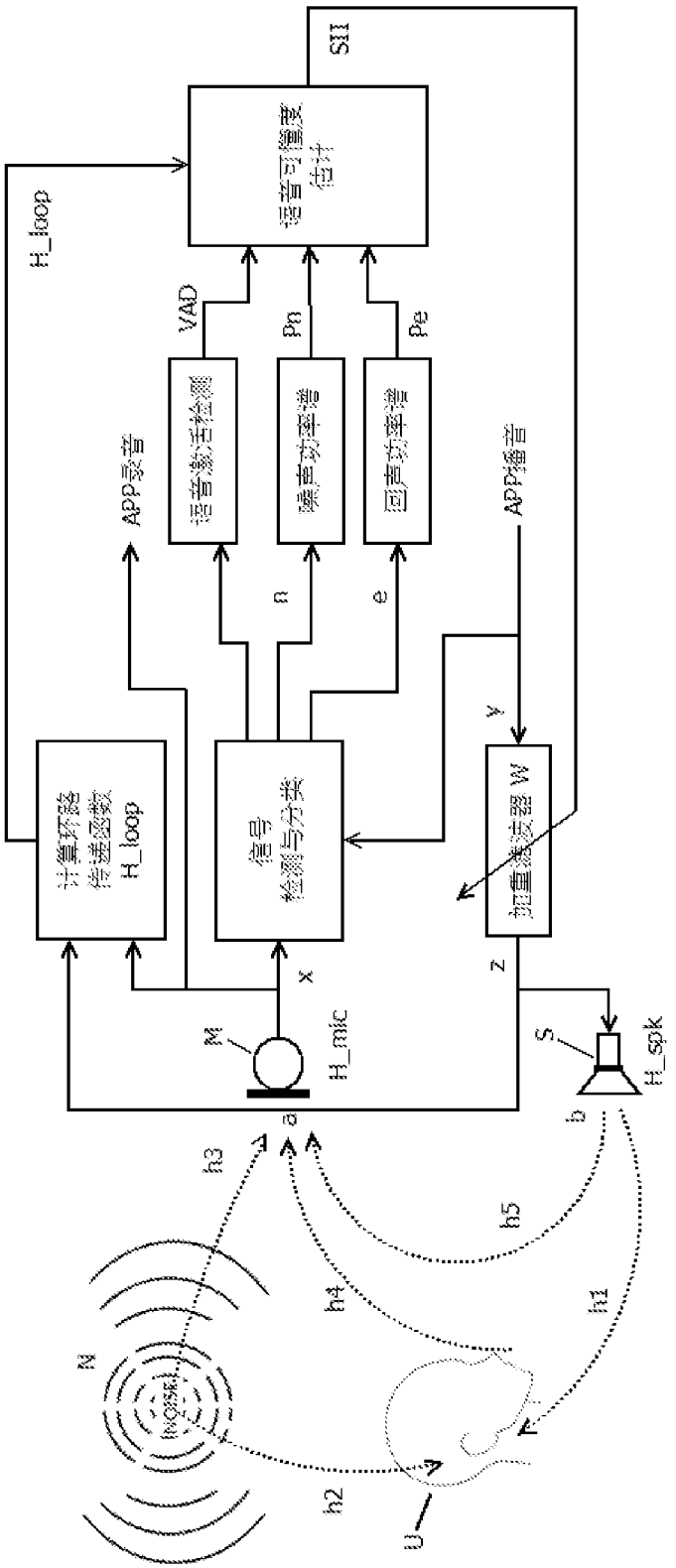

SII
语音可懂度估计
H_loop
VAD
Pn
Pe
语音激活检测
噪声功率谱
回声功率谱
APP录音
APP播音
n
e
y
计算环路传递函数 H_loop
信号检测与分类
加重滤波器 W
x
z
M
H_mic
S
H_spk
a
b
h3
h5
h4
h1
h2
N
U

图 2

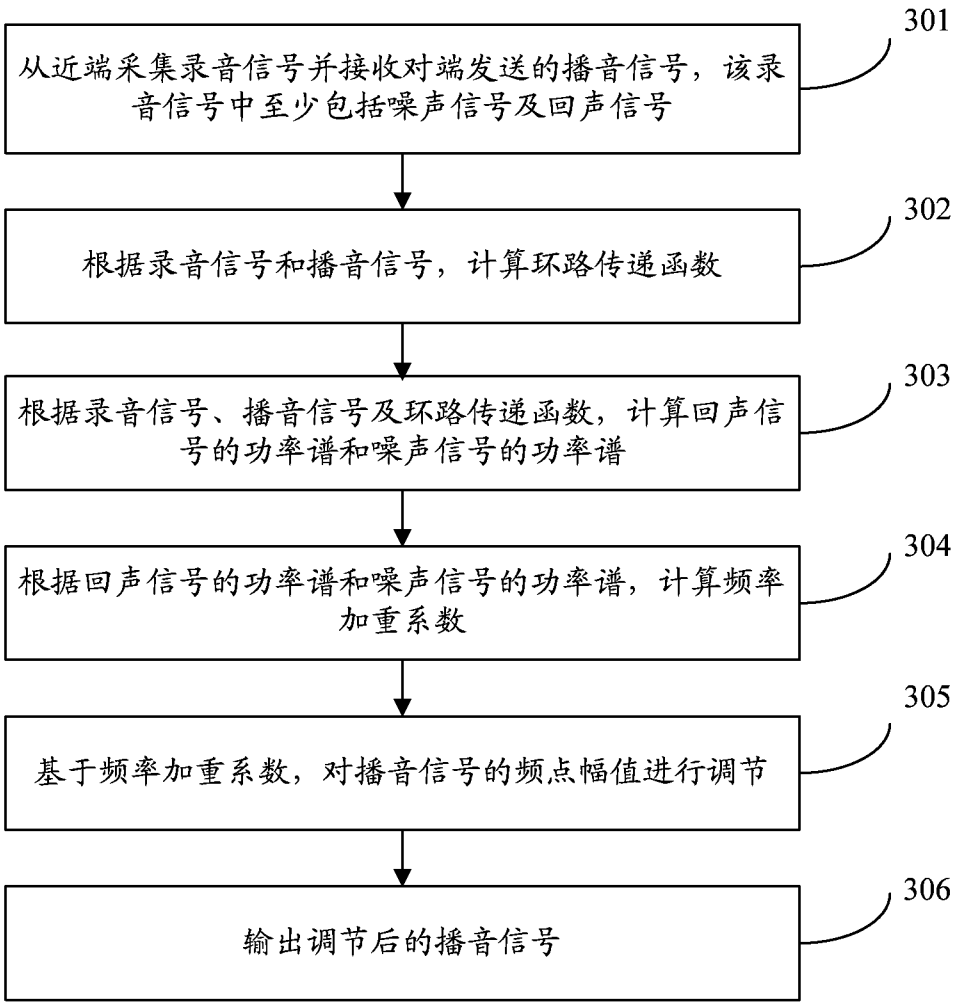

301
从近端采集录音信号并接收对端发送的播音信号，该录音信号中至少包括噪声信号及回声信号
302
根据录音信号和播音信号，计算环路传递函数
303
根据录音信号、播音信号及环路传递函数，计算回声信号的功率谱和噪声信号的功率谱
304
根据回声信号的功率谱和噪声信号的功率谱，计算频率加重系数
305
基于频率加重系数，对播音信号的频点幅值进行调节
306
输出调节后的播音信号

图 3

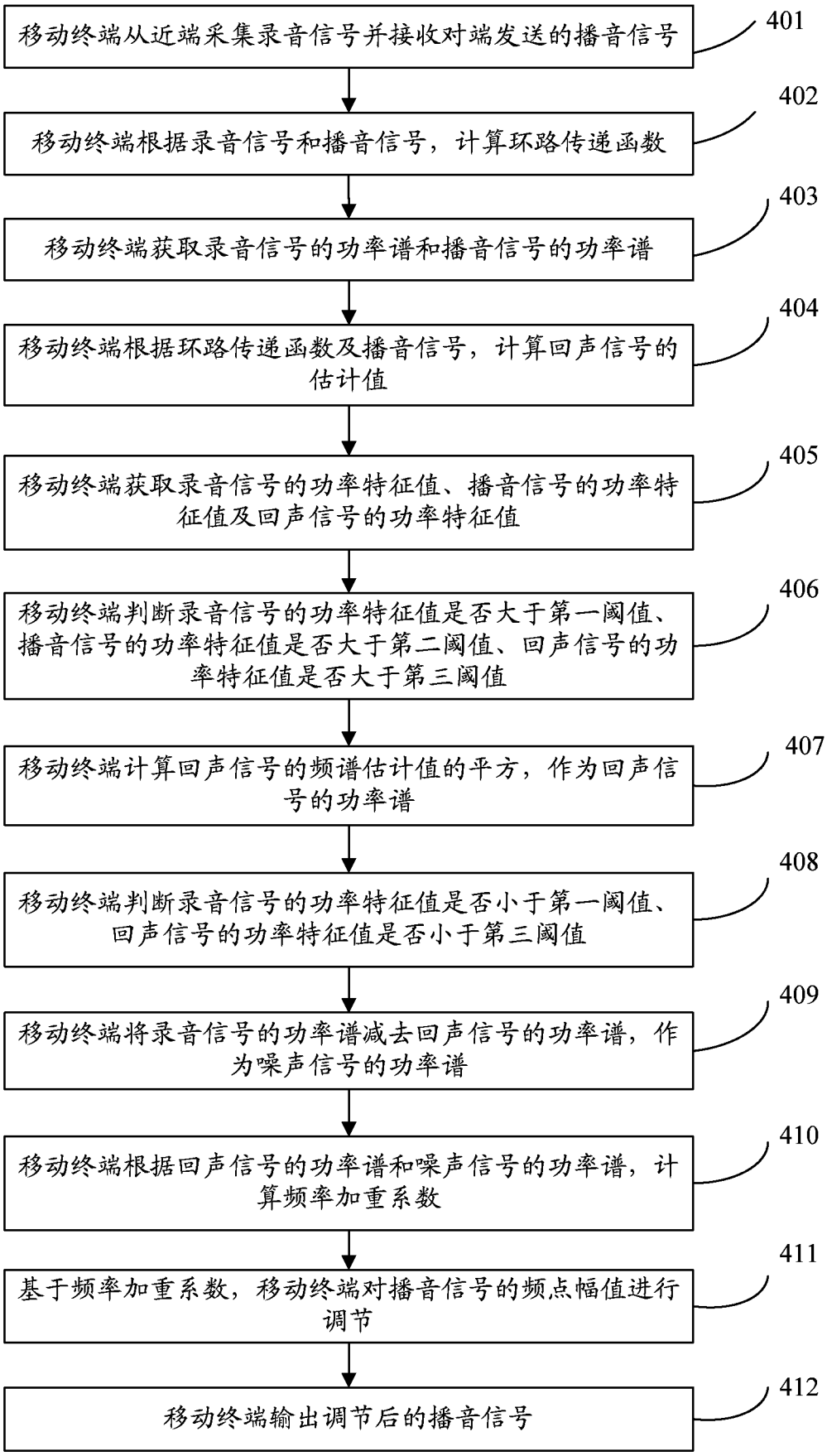

移动终端从近端采集录音信号并接收对端发送的播音信号
401
移动终端根据录音信号和播音信号，计算环路传递函数
402
移动终端获取录音信号的功率谱和播音信号的功率谱
403
移动终端根据环路传递函数及播音信号，计算回声信号的估计值
404
移动终端获取录音信号的功率特征值、播音信号的功率特征值及回声信号的功率特征值
405
移动终端判断录音信号的功率特征值是否大于第一阈值、播音信号的功率特征值是否大于第二阈值、回声信号的功率特征值是否大于第三阈值
406
移动终端计算回声信号的频谱估计值的平方，作为回声信号的功率谱
407
移动终端判断录音信号的功率特征值是否小于第一阈值、回声信号的功率特征值是否小于第三阈值
408
移动终端将录音信号的功率谱减去回声信号的功率谱，作为噪声信号的功率谱
409
移动终端根据回声信号的功率谱和噪声信号的功率谱，计算频率加重系数
410
基于频率加重系数，移动终端对播音信号的频点幅值进行调节
411
移动终端输出调节后的播音信号
412

图 4

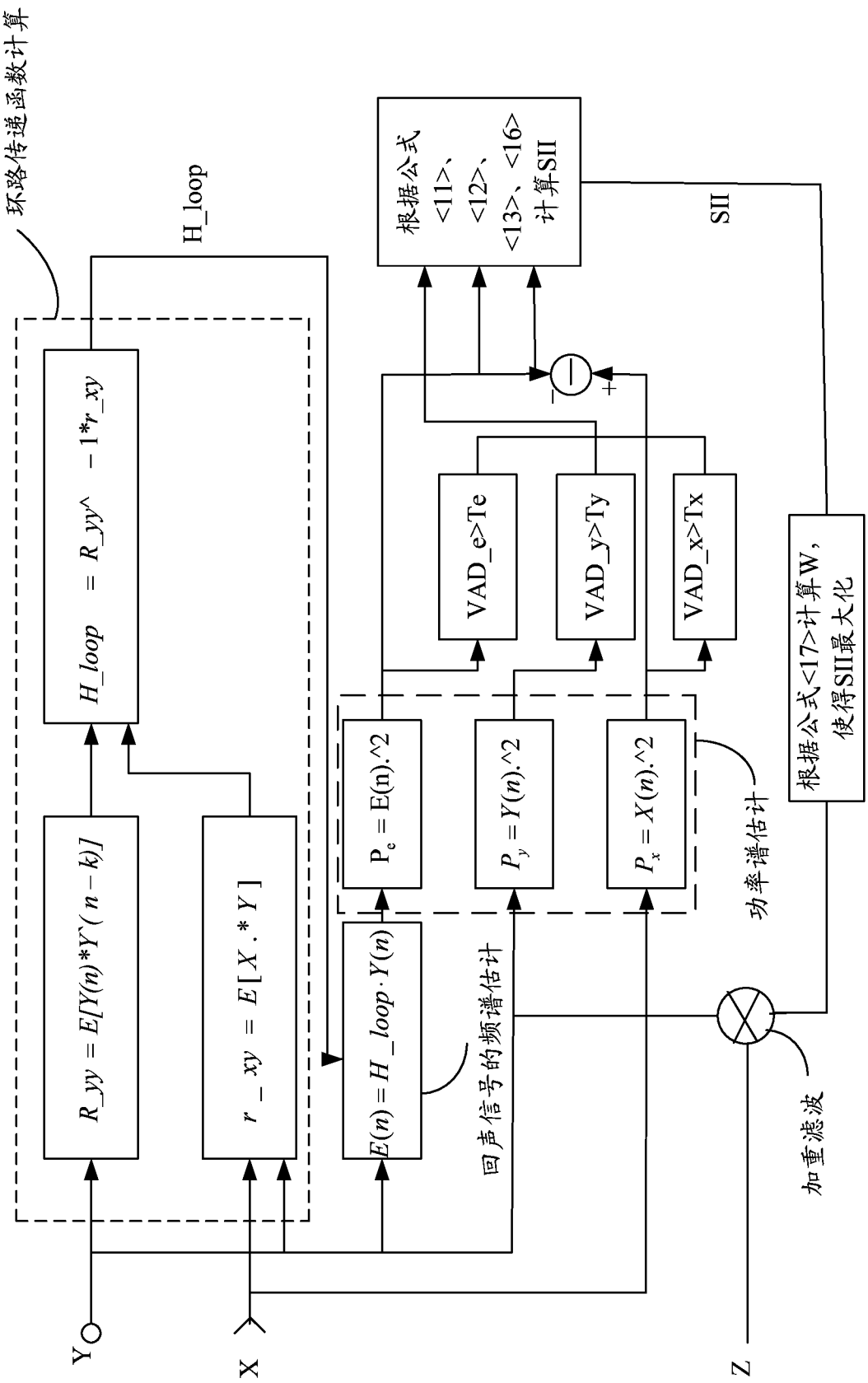

环路传递函数计算
H_loop
根据公式<11>、<12>、<13>、<16>计算SII
SII
H_loop = R_yy^ -1*r_xy
R_yy = E[Y(n)*Y'(n-k)]
r_xy = E[X.*Y]
E(n) = H_loop·Y(n)
回声信号的频谱估计
Pe = E(n).^2
Py = Y(n).^2
Px = X(n).^2
功率谱估计
VAD_e>Te
VAD_y>Ty
VAD_x>Tx
根据公式<17>计算W，使得SII最大化
加重滤波
Y
X
Z

图 5

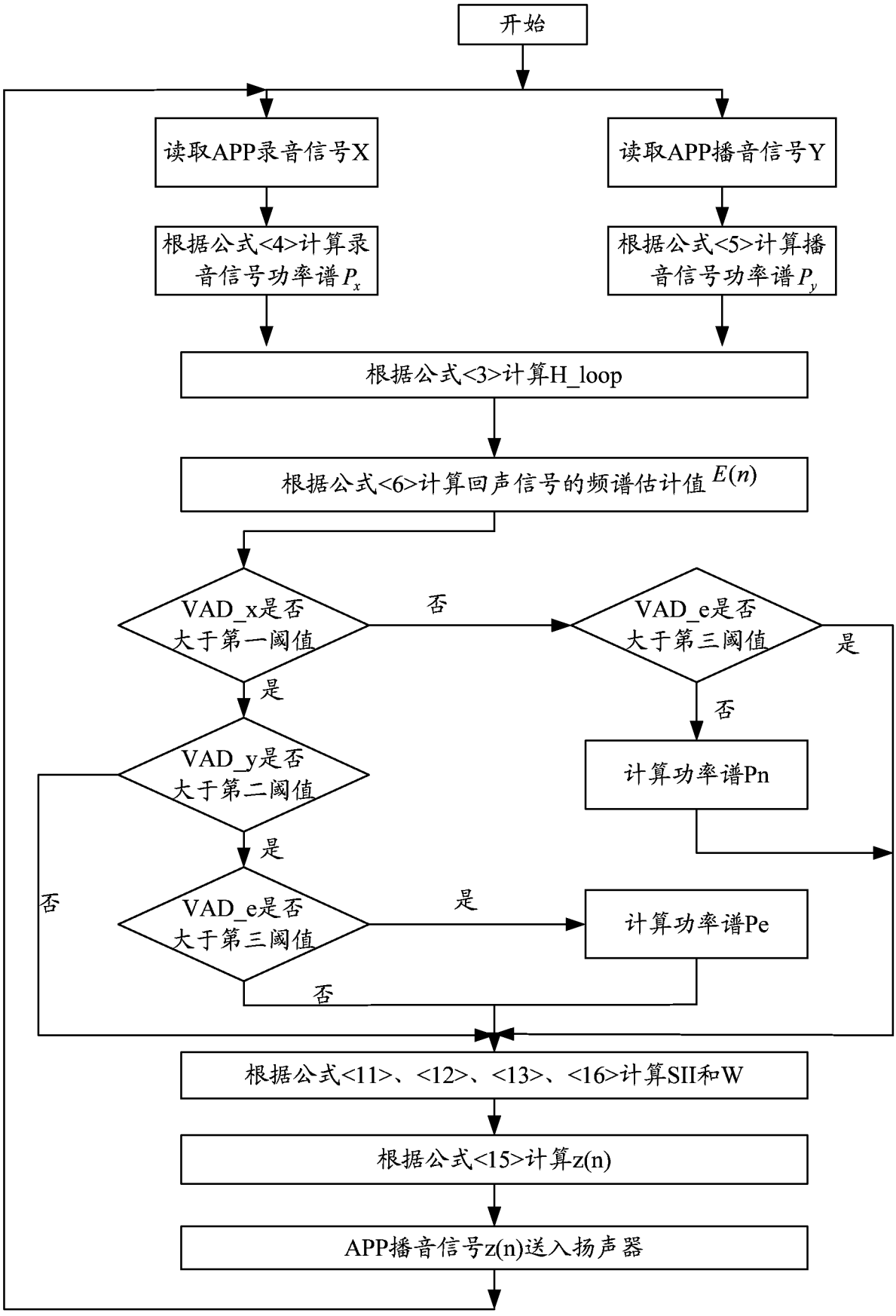

开始
读取APP录音信号X
读取APP播音信号Y
根据公式<4>计算录音信号功率谱P_x
根据公式<5>计算播音信号功率谱P_y
根据公式<3>计算H_loop
根据公式<6>计算回声信号的频谱估计值$E(n)$
VAD_x是否大于第一阈值
否
VAD_e是否大于第三阈值
是
是
否
VAD_y是否大于第二阈值
计算功率谱Pn
是
否
VAD_e是否大于第三阈值
是
计算功率谱Pe
否
根据公式<11>、<12>、<13>、<16>计算SII和W
根据公式<15>计算z(n)
APP播音信号z(n)送入扬声器

图 6

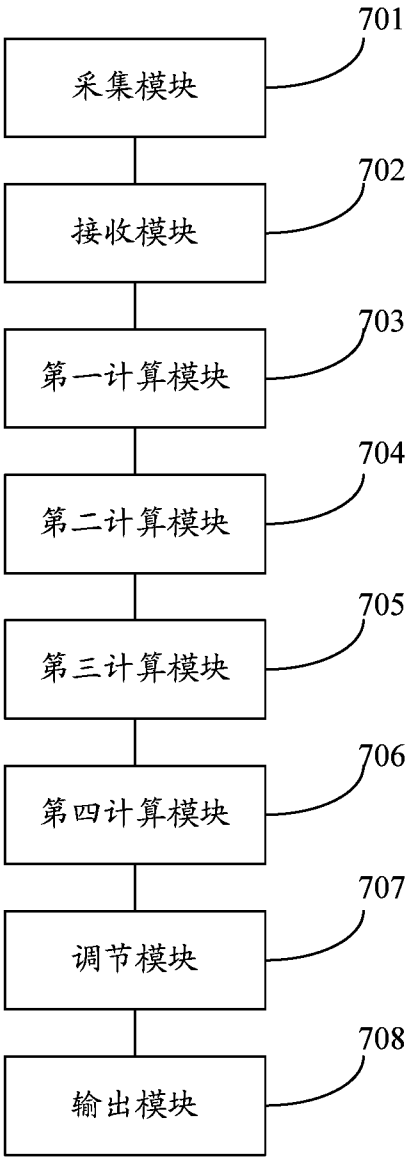

701
采集模块
702
接收模块
703
第一计算模块
704
第二计算模块
705
第三计算模块
706
第四计算模块
707
调节模块
708
输出模块

图 7

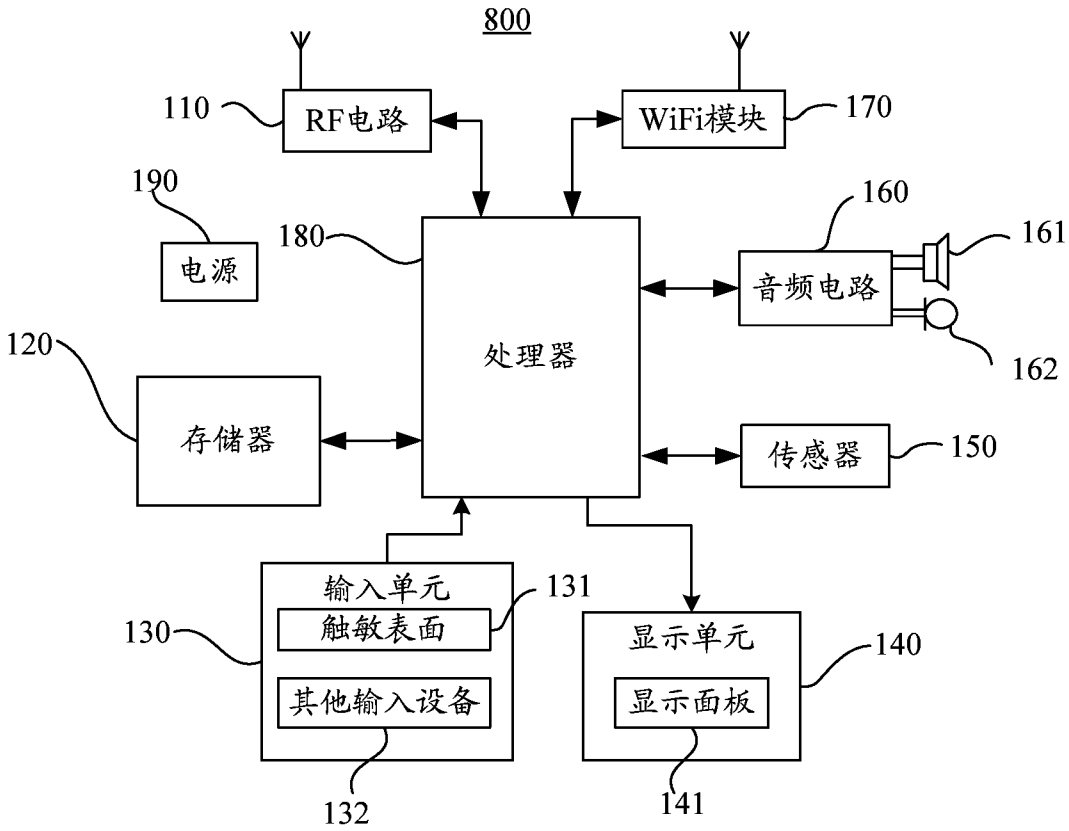

800
110
RF电路
WiFi模块
170
190
电源
180
处理器
160
音频电路
161
162
120
存储器
传感器
150
输入单元
131
触敏表面
130
其他输入设备
132
显示单元
140
显示面板
141

图 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/083622

A. CLASSIFICATION OF SUBJECT MATTER

G10L 21/0364 (2013.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNTXT; CNABS; CNKI; VEN; BAIDU ACADEMIC: record, play, broadcast, loop transfer, transfer function, energy, intelligibility, cross-correlation, self-correlation, micophone, loudspeaker, inteligibility, power, noise, echo, correlation

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 105280195 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.), 27 January 2016 (27.01.2016), the whole document	1-8
X	CN 104050971 A (TRUEPOSITION, INC.), 17 September 2014 (17.09.2014), description, paragraphs [0033]-[0090], and figure 2	1-3, 9-12, 18
A	CN 104050971 A (TRUEPOSITION, INC.), 17 September 2014 (17.09.2014), the whole document	4-8, 13-17
A	CN 103606374 A (STATE GRID CORPORATION OF CHINA et al.), 26 February 2014 (26.02.2014), the whole document	1-18
A	CN 101763858 A (AAC TECHNOLOGIES et al.), 30 June 2010 (30.06.2010), the whole document	1-18
A	CN 102893331 A (QUALCOMM INC.), 23 January 2013 (23.01.2013), the whole document	1-18

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
02 August 2016 (02.08.2016)	**17 August 2016 (17.08.2016)**
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer **YOU, Xiaomei** Telephone No.: (86-10) **010-62089539**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/083622

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105280195 A	27 January 2016	None	
CN 104050971 A	17 September 2014	WO 2014150211 A1	25 September 2014
		US 2016019909 A1	21 January 2016
		EP 2973557 A1	20 January 2016
CN 103606374 A	26 February 2014	None	
CN 101763858 A	30 June 2010	None	
CN 102893331 A	23 January 2013	US 2011288860 A1	24 November 2011
		WO 2011146903 A1	24 November 2011
		EP 2572353 A1	27 March 2013
		EP 2572353 B1	01 June 2016
		KR 20130042495 A	26 April 2013
		JP 5714700 B2	07 May 2015
		JP 2013531419 A	01 August 2013
		IN 201209560 P4	02 May 2014
		KR 20150080645 A	09 July 2015
		CN 102893331 B	09 March 2016

国际检索报告

国际申请号 PCT/CN2016/083622

A. 主题的分类

G10L 21/0364(2013.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G10L

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNTXT;CNABS;CNKI;VEN;百度学术：录音，麦克风，播放，播音，环路传递，传递函数，噪声，回声，噪音，回音，功率，能量，可懂度，互相关，自相关，micophone,loudspeaker,inteligibility,power,noise,echo,correlation

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN 105280195 A (腾讯科技深圳有限公司) 2016年 1月 27日 (2016-01-27) 全文	1-18
X	CN 104050971 A (杜比实验室特许公司) 2014年 9月 17日 (2014-09-17) 说明书第0033-0090段，附图2	1-3，9-12，18
A	CN 104050971 A (杜比实验室特许公司) 2014年 9月 17日 (2014-09-17) 全文	4-8，13-17
A	CN 103606374 A (国家电网公司 等) 2014年 2月 26日 (2014-02-26) 全文	1-18
A	CN 101763858 A (瑞声声学科技深圳有限公司 等) 2010年 6月 30日 (2010-06-30) 全文	1-18
A	CN 102893331 A (高通股份有限公司) 2013年 1月 23日 (2013-01-23) 全文	1-18

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2016年 8月 2日	2016年 8月 17日
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 游晓梅 电话号码 (86-10)010-62089539

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2016/083622

检索报告引用的专利文件			公布日（年/月/日）	同族专利			公布日（年/月/日）
CN	105280195	A	2016年 1月 27日	无			
CN	104050971	A	2014年 9月 17日	WO	2014150211	A1	2014年 9月 25日
				US	2016019909	A1	2016年 1月 21日
				EP	2973557	A1	2016年 1月 20日
CN	103606374	A	2014年 2月 26日	无			
CN	101763858	A	2010年 6月 30日	无			
CN	102893331	A	2013年 1月 23日	US	2011288860	A1	2011年 11月 24日
				WO	2011146903	A1	2011年 11月 24日
				EP	2572353	A1	2013年 3月 27日
				EP	2572353	B1	2016年 6月 1日
				KR	20130042495	A	2013年 4月 26日
				JP	5714700	B2	2015年 5月 7日
				JP	2013531419	A	2013年 8月 1日
				IN	201209560	P4	2014年 5月 2日
				KR	20150080645	A	2015年 7月 9日
				CN	102893331	B	2016年 3月 9日