

(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
2016 年 9 月 15 日 (15.09.2016)



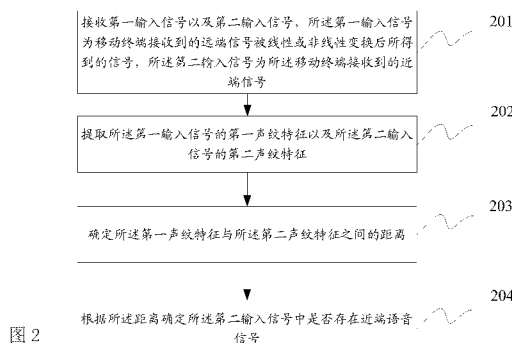
(10) 国际公布号
WO 2016/141773 A1

- (51) 国际专利分类号:
G10L 25/78 (2013.01)
- (21) 国际申请号: PCT/CN2016/070253
- (22) 国际申请日: 2016 年 1 月 6 日 (06.01.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510102968.X 2015 年 3 月 9 日 (09.03.2015) CN
- (71) 申请人: 电信科学技术研究院 (CHINA ACADEMY OF TELECOMMUNICATIONS TECHNOLOGY) [CN/CN]; 中国北京市海淀区学院路 40 号, Beijing 100191 (CN)。
- (72) 发明人: 梁民 (LIANG, Min); 中国北京市海淀区学院路 40 号, Beijing 100191 (CN)。 韩波 (HAN, Bo); 中国北京市海淀区学院路 40 号, Beijing 100191 (CN)。
- (74) 代理人: 北京银龙知识产权代理有限公司 (DRAGON INTELLECTUAL PROPERTY LAW FIRM); 中国北京市海淀区西直门北大街 32 号院枫蓝国际中心 2 号楼 10 层, Beijing 100082 (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, 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,

[见续页]

(54) Title: NEAR-END VOICE SIGNAL DETECTION METHOD AND APPARATUS

(54) 发明名称: 一种近端语音信号检测方法及装置



201 RECEIVE A FIRST INPUT SIGNAL AND A SECOND INPUT SIGNAL, WHEREIN THE FIRST INPUT SIGNAL IS A SIGNAL OBTAINED BY LINEARLY OR NON-LINEARLY CHANGING A FAR-END SIGNAL RECEIVED BY A MOBILE TERMINAL, AND THE SECOND INPUT SIGNAL IS A NEAR-END SIGNAL RECEIVED BY THE MOBILE TERMINAL

202 EXTRACT A FIRST VOCAL PRINT CHARACTERISTIC OF THE FIRST INPUT SIGNAL AND A SECOND VOCAL PRINT CHARACTERISTIC OF THE SECOND INPUT SIGNAL

203 DETERMINE A DISTANCE BETWEEN THE FIRST VOCAL PRINT CHARACTERISTIC AND THE SECOND VOCAL PRINT CHARACTERISTIC

204 DETERMINE, ACCORDING TO THE DISTANCE, WHETHER THERE IS A NEAR-END VOICE SIGNAL IN THE SECOND INPUT SIGNAL

(57) Abstract: A near-end voice signal detection method and apparatus. The method comprises: receiving a first input signal and a second input signal (201), wherein the first input signal is obtained by linearly changing or non-linearly changing a far-end voice signal, and the second input signal is an output signal of an audio receiving sensor; determining, according to the first input signal, a first vocal print characteristic vector of the first input signal, and determining, according to the second input signal, a second vocal print characteristic vector of the second input signal (202); and determining a distance between a first vocal print characteristic and a second vocal print characteristic (203), and determining, according to the distance, whether a double-talk occurs or not.

(57) 摘要: 一种近端语音信号检测方法及装置, 该方法包括: 接收第一输入信号以及第二输入信号 (201), 所述第一输入信号通过对远端语音信号进行线性变换或非线性变换获得, 所述第二输入信号为音频接收传感器的输出信号; 根据所述第一输入信号确定出所述第一输入信号的第一声纹特征矢量, 根据所述第二输入信号确定出所述第二输入信号的第二声纹特征矢量 (202); 确定出所述第一声纹特征与所述第二声纹特征之间的距离 (203), 并根据所述距离确定是否发生双讲。



WO 2016/141773 A1



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))。

一种近端语音信号检测方法及装置

相关申请的交叉引用

本申请主张在 2015 年 3 月 9 日在中国提交的中国专利申请号 No. 201510102968.X 的优先权，其全部内容通过引用包含于此。

技术领域

本公开涉及语音信号检测技术领域，尤其涉及一种近端语音信号检测方法及装置。

背景技术

声学回波抵消器 (Acoustic Echo canceller, AEC) 是电话会议系统、免提通信终端等设备的一个重要模块，用来抵消由扬声器到麦克风的声学耦合反馈效应，即扬声器到麦克风之间的声学回波。

在声学回波抵消器中，用一个滤波器对回波路径进行自适应地数学建模，并由此合成一个声学回波的有效估计，然后在麦克风的接收信号的输出信号中减去该声学回波的有效估计，从而实现声学回波抵消的目的。当麦克风的接收信号中出现近端语音信号时，即发生双讲 (Double-Talk, DT) 情形，由于它与远端语音信号统计上不相关，因而其犹如一个突发的噪声，使得滤波器的系数将偏离实际声学回波路径所对应的真值而发生发散现象。这便相应地增大了回波残留量，使声学回波抵消器的性能恶化。为使声学回波抵消器的工作性能稳定可靠，准确而及时地检测出麦克风接收信号中是否发生双讲，是一项非常重要和必要的任务。在 DT 发生的条件下，滤波器系数的自适应学习必须停止进行，以避免在该情况下滤波器系数持续学习所致的发散现象。

为克服这一问题，一种自然的处理方法是：滤波器的滤波器系数矢量的学习算法应该在发生双讲的情况下被停止执行，而在未发生双讲时将持续进行。由此，双讲检测器 (DTD) 便应运而生。目前，双讲检测器主要是基于互相关

(Cross-Correlation) 准则实现的。在基于互相关准则的 DTD 中, 较典型的技术方案有以下两种:

第一种方案, 利用声学回波抵消器中的误差信号 $e(n)$ 和远端语音信号矢量 $\vec{x}(n)$ 之间互相关来进行双讲检测, 误差信号 $e(n)$ 和远端语音信号矢量 $\vec{x}(n)$ 之间互相关系数如下:

$$\vec{r}_{e\vec{x}} = E\{e(n) \cdot \vec{x}^T(n)\} = E\{[y(n) - \widehat{x_2}(n)] \cdot \vec{x}^T(n)\} \quad (1)$$

在由放大器过载和编码解码器引入的非线性失真可以忽略不计, 以及环境噪声是平稳的假设条件下 (在无特别注明的情况下, 以下均假设该条件成立), 式 (1) 变为:

$$\vec{r}_{e\vec{x}} = E\{[\vec{h}^T \vec{x}(n) + u(n) + v(n) - \vec{h}^T \vec{x}(n)] \vec{x}^T(n)\} = (\vec{h} - \vec{h})^T R_{XX} \quad (2)$$

其中, $\vec{h} = [h(0), h(1), \dots, h(L-1)]^T$, 为声学回波的回波路径中线性部分的冲击响应, L 为回波路径的长度; $\vec{h} = [\hat{h}(0), \hat{h}(1), \dots, \hat{h}(L-1)]^T$ 为滤波器的冲击响应; $R_{XX} = E\{\vec{x}(n) \cdot \vec{x}^T(n)\}$ 为远端语音信号的自相关矩阵。

式 (2) 中的 $\vec{r}_{e\vec{x}}$ 高度依赖于回波路径的变化, 因而适合用于检测声学回波路径是否发生变化, 而不是用来检测双讲是否发生。

第二种方案, 利用远端语音信号矢量 $\vec{x}(n)$ 和麦克风输出信号 $y(n)$ 之间的互相关来构造一个决策统计量用于双讲检测。矢量 $\vec{x}(n)$ 和 $y(n)$ 之间的互相关 $\vec{r}_{xy}$ 可表达为:

$$\vec{r}_{xy} = E\{\vec{x}(n) \cdot y(n)\} = E\{\vec{x}(n) [\vec{h}^T \vec{x}(n) + u(n) + v(n)]\} = R_{XX} \vec{h} \quad (3)$$

考虑到麦克风输出信号 $y(n)$ 的方差 σ_y^2 可表示成下式:

$$\sigma_y^2 = E\{y^2(n)\} = \vec{r}_{xy}^T R_{XX}^{-1} \vec{r}_{xy} + \sigma_v^2 + \sigma_u^2 \quad (4)$$

其中 σ_v^2 和 σ_u^2 分别为环境噪声和近端语音信号的功率。

在无 DT, 即 $u(n) = 0$ 时, 式 (4) 即为:

$$\sigma_y^2 = \vec{r}_{xy}^T R_{XX}^{-1} \vec{r}_{xy} + \sigma_v^2 \quad (5)$$

将决策统计量 ξ_{Benesty} 定义为 用式 (5) 除以式 (4) 后再开方, 即:

$$\xi_{\text{Benesty}} = \sqrt{\frac{\vec{r}_{xy}^T \mathbf{R}_{XX}^{-1} \vec{r}_{xy} + \sigma_v^2}{\vec{r}_{xy}^T \mathbf{R}_{XX}^{-1} \vec{r}_{xy} + \sigma_v^2 + \sigma_u^2}} = \sqrt{\frac{\vec{h}^T \mathbf{R}_{XX} \vec{h} + \sigma_v^2}{\vec{h}^T \mathbf{R}_{XX} \vec{h} + \sigma_v^2 + \sigma_u^2}} \quad (6)$$

根据式(6)可以确定,在无双讲时,决策统计量 ξ_{Benesty} 取值为 1;在无双讲时,决策统计量 ξ_{Benesty} 取值小于 1。因此可定义一个门限值参数 T_{Benesty} ,当 $\xi_{\text{Benesty}} < T_{\text{Benesty}}$,则确定发生双讲;否则,确定无双讲发生。

还可以利用误差信号 $e(n)$ 和麦克风输出信号 $y(n)$ 之间的互相关来构造 DTD 的决策统计量,具体的,将误差信号 $e(n)$ 和麦克风输出信号 $y(n)$ 之间的互相关定义为:

$$r_{ey} = E[e(n) \cdot y(n)] = E\{[y(n) - \widehat{x2}(n)]y(n)\} = [\vec{h} - \vec{\hat{h}}]^T \mathbf{R}_{XX} \vec{h} + \sigma_v^2 + \sigma_u^2 \quad (7)$$

构造的决策统计量 ξ_{Iqbal} 如下:

$$\xi_{\text{Iqbal}} = 1 - \frac{r_{ey}}{\sigma_y^2} = 1 - \frac{[\vec{h} - \vec{\hat{h}}]^T \mathbf{R}_{XX} \vec{h} + \sigma_v^2 + \sigma_u^2}{\vec{h}^T \mathbf{R}_{XX} \vec{h} + \sigma_v^2 + \sigma_u^2} = \frac{\vec{\hat{h}}^T \mathbf{R}_{XX} \vec{h} + \sigma_v^2}{\vec{h}^T \mathbf{R}_{XX} \vec{h} + \sigma_v^2 + \sigma_u^2} \quad (8)$$

在滤波器收敛时,滤波器的冲击响应趋于回波路径的冲击响应,即 $\vec{\hat{h}} \rightarrow \vec{h}$ 。那么在无双讲的情况下 $\xi_{\text{Iqbal}} \approx 1$,而在有双讲时 $\xi_{\text{Iqbal}} < 1$ 。因此可定义一个门限值参数 T_{Iqbal} ,当 $\xi_{\text{Iqbal}} < T_{\text{Iqbal}}$,则确定发生双讲;否则,就确定未发生双讲。

上述介绍的双讲检测技术都是基于以下两个假设:1、声学回波路径中非线性失真很小而忽略不计;2、环境噪声是平稳的。然而实际系统中,由于放大器过载和编码解码器所引发的非线性失真不可忽略,使得相关技术中的基于互相关技术的双讲检测技术的性能较差。此外,实际环境中的噪声也并非平稳的,这一非平稳性也将进一步加剧该类双讲检测技术性能的恶化程度,乃至有时无法正常检测出是否发生双讲。

发明内容

本公开的一些实施例提供了一种近端语音信号检测方法及装置,用以提高双讲检测性能。

本公开的一些实施例提供了一种近端语音信号检测方法,包括:

接收第一输入信号以及第二输入信号，其中，所述第一输入信号为移动终端接收到的远端信号被线性或非线性变换后所得到的信号，所述第二输入信号为所述移动终端接收到的近端信号；

提取所述第一输入信号的第一声纹特征以及所述第二输入信号的第二声纹特征；

确定所述第一声纹特征与所述第二声纹特征之间的距离；以及

根据所述距离确定所述第二输入信号中是否存在近端语音信号。

可选地，所述第一输入信号为所述移动终端的自适应滤波器输出的回波估计信号，所述回波估计信号是所述自适应滤波器对所述远端信号进行线性或非线性滤波得到的。

可选地，所述第一输入信号为所述远端信号被线性延时后得到的信号。

可选地，所述根据所述距离确定所述第二输入信号中是否存在近端语音信号，包括：

判断所述距离是否小于第一门限值，若是，则确定所述第二输入信号中不存在近端语音信号，否则，确定所述第二输入信号中存在近端语音信号。

可选地，所述确定所述第二输入信号中存在近端语音信号之后，还包括：

向所述移动终端的自适应滤波器发送指示信息，其中，所述指示信息用于指示所述自适应滤波器暂停更新滤波器系数。

本公开的一些实施例提供了一种近端语音信号检测方法，包括：

接收第一输入信号以及第二输入信号，其中，所述第一输入信号为移动终端接收到的远端信号，所述第二输入信号为所述移动终端接收到的近端信号；以及

检测所述第一输入信号是否大于第二门限值，以及检测所述第二输入信号是否大于第三门限值；

若所述第一输入信号大于所述第二门限值，且所述第二输入信号大于所述第三门限值，则提取所述第一输入信号的第一声纹特征，以及提取所述第二输入信号的第二声纹特征，确定所述第一声纹特征与所述第二声纹特征之间的距

离，并根据所述距离确定所述第二输入信号中是否存在近端语音信号。

可选地，所述检测所述第一输入信号是否大于第二门限值，以及检测所述第二输入信号是否大于第三门限值，包括：

检测所述第一输入信号在第一时间点是否大于所述第二门限值，以及检测所述第二输入信号在第二时间点是否大于所述第三门限值，其中，所述第二时间点为所述第一时间点经过延时后的时间点。

可选地，所述根据所述距离确定所述第二输入信号中是否存在近端语音信号，包括：

若所述距离小于第四门限值，则确定所述第二输入信号在所述第二时间点不存在近端语音信号，否则，确定所述第二输入信号在所述第二时间点存在近端语音信号。

可选地，还包括：

若所述第二输入信号小于所述第三门限值，则确定所述第二输入信号在所述第二时间点不存在近端语音信号；或者，

若所述第一输入信号小于所述第二门限值，且所述第二输入信号大于所述第三门限值，则确定所述第二输入信号在所述第二时间点存在近端语音信号。

可选地，确定所述第二输入信号中存在近端语音信号之后，还包括：

向所述移动终端的自适应滤波器发送指示信息，其中，所述指示信息用于指示所述自适应滤波器暂停更新滤波器系数。

本公开的一些实施例提供了一种近端语音信号检测装置，包括：

接收单元，用于接收第一输入信号以及第二输入信号，其中，所述第一输入信号为移动终端接收到的远端信号被线性或非线性变换后所得到的信号，所述第二输入信号为所述移动终端接收到的近端信号；

提取单元，用于提取所述第一输入信号的第一声纹特征以及所述第二输入信号的第二声纹特征；以及

确定单元，用于确定所述第一声纹特征与所述第二声纹特征之间的距离，并根据所述距离确定所述第二输入信号中是否存在近端语音信号。

可选地，所述第一输入信号为所述移动终端的自适应滤波器输出的回波估计信号，所述回波估计信号是所述自适应滤波器对所述远端信号进行线性或非线性滤波得到的。

可选地，所述第一输入信号为所述远端信号被线性延时后得到的信号。

可选地，所述确定单元具体用于：

判断所述距离是否小于第一门限值，若是，则确定所述第二输入信号中不存在近端语音信号，否则，确定所述第二输入信号中存在近端语音信号。

可选地，所述确定单元还用于：

向所述移动终端的自适应滤波器发送指示信息，其中，所述指示信息用于指示所述自适应滤波器暂停更新滤波器系数。

本公开的一些实施例提供了一种近端语音信号检测装置，包括：

接收单元，用于接收第一输入信号以及第二输入信号，其中，所述第一输入信号为移动终端接收到的远端信号，所述第二输入信号为所述移动终端接收到的近端信号；

检测单元，用于检测所述第一输入信号是否大于第二门限值以及检测所述第二输入信号是否大于第三门限值；以及

确定单元，用于在确定所述第一输入信号大于所述第二门限值且所述第二输入信号大于所述第三门限值时，提取所述第一输入信号的第一声纹特征，以及提取所述第二输入信号的第二声纹特征，确定所述第一声纹特征与所述第二声纹特征之间的距离，并根据所述距离确定所述第二输入信号中是否存在近端语音信号。

可选地，所述检测单元用于：

检测所述第一输入信号在第一时间点是否大于所述第二门限值，以及检测所述第二输入信号在第二时间点是否大于所述第三门限值，其中，所述第二时间点为所述第一时间点经过延时后的时间点。

可选地，所述确定单元具体用于：

若所述距离小于第四门限值，则确定所述第二输入信号在所述第二时间点

不存在近端语音信号，否则，确定所述第二输入信号在所述第二时间点存在近端语音信号。

可选地，所述确定单元还用于：

若所述第二输入信号小于所述第三门限值，则确定所述第二输入信号在所述第二时间点不存在近端语音信号；或者，

若所述第一输入信号小于所述第二门限值，且所述第二输入信号大于所述第三门限值，则确定所述第二输入信号在所述第二时间点存在近端语音信号。

可选地，所述确定单元还用于：

向所述移动终端的自适应滤波器发送指示信息，其中，所述指示信息用于指示所述自适应滤波器暂停更新滤波器系数。

根据本公开的一些实施例提供的方法及装置，提取远端信号的第一声纹特征，以及近端信号中的第二声纹特征之后，通过对比第一声纹特征与第二声纹特征确定是否发生双讲。由于本公开的一些实施例是根据远端信号和近端信号的声纹特征来进行近端语音信号的判决，即判断是否发生双讲，没有像现有技术一样基于互相关技术以及两个假设（1、声学回波路径中非线性失真很小而忽略不计；2、环境噪声是平稳的）来进行双讲检测，因此一定程度上避免了现有技术中以假设声学回波路径中非线性失真很小而忽略不计以及假设环境噪声平稳为前提条件去检测是否发生双讲时产生的误判等情况，从而更准确的实现双讲检测。

附图说明

图 1 为现有技术中声学回波抵消器的结构示意图；

图 2 为本公开的一些实施例提供的一种近端语音信号检测方法流程示意图；

图 3 为本公开的一些实施例提供的声纹特征提取流程示意图；

图 4 为本公开的一些实施例提供的一种近端语音信号检测方法流程示意图；

图 5 为本公开的一些实施例提供的一种近端语音信号检测装置结构图；

图 6 为本公开的一些实施例提供的一种近端语音信号检测方法流程示意图；

图7为本公开的一些实施例提供的第二种近端语音信号检测装置应用场景示意图；

图8为本公开的一些实施例提供的一种近端语音信号检测装置结构图；

图9为本公开的一些实施例提供的一种近端语音信号检测装置结构图。

具体实施方式

如图1所示,为现有技术中声学回波抵消器的结构示意图,包括扬声器101,自适应滤波器102,双讲检测器103,麦克风104。远端语音信号 $x(n)$ 从扬声器101输出时,扬声器101中放大器过载和编码解码器会导致远端语音信号 $x(n)$ 非线性失真;远端语音信号 $x(n)$ 从扬声器101传输到麦克风104的过程中,扬声器101传输到麦克风104之间的声学回波路径也会对远端语音信号 $x(n)$ 产生影响。

现假设导致远端语音信号 $x(n)$ 非线性失真的非线性冲击响应很小,可以忽略不计,那么有:

$$\begin{cases} x1(n) = x(n) \\ y(n) = x2(n) + u(n) + v(n) \end{cases} \quad (9)$$

其中, $y(n)$ 为麦克风104的接收信号, $u(n)$ 为近端语音信号, $v(n)$ 为系统噪声, $x1(n)$ 为远端语音信号 $x(n)$ 经过非线性冲击响应后的语音信号,它们均为零均值; $x2(n)$ 为回波信号,由下式确定:

$$x2(n) = \vec{h}^T \cdot \vec{x1}(n) = \vec{h}^T \cdot \vec{x}(n) \quad (10)$$

其中, $\vec{h} = [h(0), h(1), \dots, h(L-1)]^T$, 为扬声器101到麦克风104之间的声学回波路径中线性部分的冲击响应, L 为回波路径的长度;

$$\vec{x}(n) = [x(n), x(n-1), \dots, x(n-L+1)]^T$$

这时用自适应滤波器102对扬声器101馈入麦克风104的回波信号 $x2(n)$ 进行估计,得估计信号 $\widehat{x2}(n)$ 如下:

$$\widehat{x2}(n) = \vec{\hat{h}}^T \cdot \vec{x}(n) \quad (11)$$

其中, $\vec{\hat{h}} = [\hat{h}(0), \hat{h}(1), \dots, \hat{h}(L-1)]^T$, 为自适应滤波器102的系数矢量。

将 $\widehat{x_2}(n)$ 从麦克风 104 的输出信号 $y(n)$ 中减去, 获得相应的误差信号 $e(n)$ 为:

$$e(n) = y(n) - \widehat{x_2} \quad (12)$$

自适应滤波器 102 的系数矢量 $\vec{h}$ 是通过自适应算法学习获得的, 在 $\vec{h}$ 收敛于 $\vec{h}$ 的条件下, 误差信号 $e(n)$ 中的回波信号 $x_2(n)$ 会被抵消, 从而达到消除回波信号的目的。当近端语音信号 $u(n)$ 出现, 即发生双讲时, 由于近端语音信号 $u(n)$ 与远端语音信号 $x(n)$ 之间统计上不相关, 因此近端语音信号 $u(n)$ 对于远端语音信号 $x(n)$ 来说犹如一个突发干扰信号, 致使自适应滤波器 102 的系数矢量 $\vec{h}$ 的自适应学习算法发散, 由此导致误差信号 $e(n)$ 中将出现较大的残留回波。

目前通过检测双讲是否发生, 并在检测到双讲发生时停止自适应滤波器 102 的系数矢量 $\vec{h}$ 的更新, 从而避免导致误差信号 $e(n)$ 中将出现较大的残留回波。

现有技术中, 在检测双讲是否发生时, 都是基于以下两个假设: 1、声学回波路径中非线性失真很小而忽略不计; 2、环境噪声是平稳的。然而, 实际情况中, 声学回波路径中非线性失真往往很大, 或者环境噪声非常不平稳, 导致基于这两个假设条件的双讲检测技术的性能很不稳定, 有时无法正常检测出是否发生双讲。

本公开的一些实施例中将摒弃这两个假设条件, 从另外一个角度去实现双讲检测, 下面详细描述本公开的一些实施例提供的双讲检测方法是如何检测双讲是否发生。需要说明的是, 本公开的一些实施例提供的双讲检测方法并不仅仅是应用于带声学回波抵消器的电话会议系统、免提通信终端等设备, 还可以应用于其他设备和系统, 在此并不限定其应用场景。

如图 2 所示, 本公开的一些实施例提供的一种近端语音信号检测方法, 该方法包括:

步骤 201: 接收第一输入信号以及第二输入信号, 其中, 所述第一输入信号为移动终端接收到的远端信号被线性或非线性变换后所得到的信号, 所述第

二输入信号为所述移动终端接收到的近端信号；

步骤 202：提取所述第一输入信号的第一声纹特征以及所述第二输入信号的第二声纹特征；

步骤 203：确定所述第一声纹特征与所述第二声纹特征之间的距离；以及

步骤 204：根据所述距离确定所述第二输入信号中是否存在近端语音信号。

本公开的一些实施例中的移动终端可以为手机、平板电脑、会议电话等设备。

在步骤 201 中，第一输入信号为移动终端接收到的远端信号被线性或非线性的变换后所得到的信号。远端信号是经过编码、调制，并需要被扬声器等设备播放的信号。

第二输入信号，即近端信号，是由麦克风等音频接收传感器接收到的信号，可能包括远端信号经过回声路径形成的声学回波信号、环境噪声信号以及近端语音信号中的一种或多种组合，第二输入信号中的声学回波信号是需要消除的信号。第二输入信号中包含由远端信号经过回声路径形成的声学回波信号时，会产生一定的延时，导致与远端信号之间不同步，如果不对远端信号进行延时处理，直接采用远端信号与第二输入信号进行双讲检测，会降低检测的准确性。因此需要将远端信号进行线性变换或非线性的变换，形成与第二输入信号中声学回波信号同步的第一输入信号。

实现将远端信号进行线性变换或非线性的变换形成第一输入信号的方法有多种。第一输入信号可以为移动终端的自适应滤波器输出的回波估计信号，所述回波估计信号是所述自适应滤波器对所述远端信号进行线性或非线性的滤波得到的；也可以通过延时单元对远端语音信号延时，将延时后的远端语音信号作为第一输入信号。需要说明的是，该延时单元对信号的延时与回波路径的延时相匹配，可以通过声学回声路径延时估计算法确定出延时单元，也可以通过其他方法确定出延时单元，本公开对此并不限定。

在步骤 201 中，获得第一输入信号以及第二输入信号之前，还可以检测输入的第一输入信号和\或第二输入信号中是否有语音信号，在未获得第一输入

信号或者获得的第一输入信号中不包含语音信号时，移动终端中的自适应滤波器的滤波器系数可以停止系数的更新，以便节省功耗；获得的第一输入信号中包含语音信号时，若第二输入信号中存在近端语音信号时，移动终端中的自适应滤波器的滤波器系数可以停止系数的更新，若第二输入信号中不存在近端语音信号时，可以直接确定未发生双讲，此时移动终端中的自适应滤波器需要根据残差信号进行滤波器系数的更新。

检测输入的第一输入信号和\或第二输入信号中是否有语音信号的方法有多种，例如可以通过语音活动检测（Voice activity detection, VAD）来检测输入的信号是否包含语音信号。

步骤 202 中，在获得第一输入信号以及第二输入信号之后，分别提取第一输入信号的第一声纹特征以及第二输入信号第二声纹特征。

声纹（Voiceprint）是携带语音信息的声波频谱，由于人在讲话时使用的发生器在尺寸和形态方面各自有差异，所以任何两个人的声纹都存在差异；另一方面，人耳能在嘈杂的背景噪声中及各种变异的情况下听到语音信号，该特性是得益于这样一个事实：耳蜗实质上相当于一个滤波器组，其滤波作用是在对数频率尺度上进行的，从而使得人耳对低频信号比对高频信号更敏感。综合考虑人耳的听觉感知和人的语音产生的机理，在本公开的一些实施例中选择 Mel 频率的倒谱系数（Mel-Frequency Cepstral Coefficient, MFCC）作为语音信号的声纹特征参数，用来进行双讲检测。其基本原理是：首先分别提取第一输入信号和第二输入信号的 MFCC 特征参数矢量，然后计算它们之间的距离，根据距离判断有无发生双讲。在未发生双讲的情况下，第二输入信号中仅含回波信号，因而第一输入信号和第二输入信号的 MFCC 特征参数矢量间的距离较小；在发生双讲的情况下，第二输入信号中不仅含近端语音信号 $u(n)$ ，而且还可能包含回波信号（在有远端语音信号的前提下），此时第一输入信号和第二输入信号的 MFCC 特征参数矢量间的距离较大。由于声纹特征参数对声学回波路径中的非线性失真和噪声干扰具有较强的不敏感特性，因而本公开提出的基于声纹特征参数之 DTD 对环境噪声和声学回波路径中的非线性退变，具有较好

的鲁棒性。

需要说明的是，本公开的一些实施例中从音频信号中提取的声纹特征包括但不限于 MFCC，可以是能有效表征和鉴别信号的任何特征参数，并且该类参数对信号的噪声污染和非线性畸变具有较好的抵免性。

针对一个输入信号，根据预加重函数对所述输入信号进行预加重，获得预加重后的输入信号；通过窗函数对所述预加重后的输入信号进行加窗，并计算所述加窗后的输入信号的频谱；通过 Mel 滤波器组对所述加窗后的输入信号的频谱进行滤波，并对滤波后的所述加窗后的输入信号的频谱进行离散余弦变换，获得所述输入信号的声纹特征。

具体地，如图 3 所示，本公开的一些实施例提供的提取声纹特征流程图。

步骤 301：预加重处理；

将输入信号通过预加重函数进行预加重处理，预加重函数为：

$$z(n) = x(n) - \alpha \cdot x(n-1) \quad (13)$$

其中， $0.9 < \alpha < 1.0$ 为预加重系数， α 一般取 0.95， $x(n)$ 为输入信号，可以为第一输入信号或者第二输入信号， $z(n)$ 为预加重后的输入信号。对输入信号进行预加重可以提升信号的高频分量进而补偿声门脉冲形状和口唇辐射对语音信号产生的影响，从而提高检测的准确性。

步骤 302：加窗；

通过窗函数对预加重后的输入信号进行加窗，获得加窗后的输入信号 $z(n)w(n)$ ；其中 $w(n)$ 为长度 N 的窗函数，可以为汉明窗函数、高斯窗函数、矩形窗函数等。

步骤 303：计算频谱；

对加窗后的输入信号进行离散傅立叶变换，获得第 t 帧输入信号的频谱 $Z(t, k)$ ：

$$Z(t, k) = \sum_{n=0}^{N-1} z(n)w(t-n)\exp\left(-j\frac{2\pi nk}{N}\right), k=0, 1, \dots, N-1 \quad (14)$$

步骤 304：Mel 滤波器组滤波；

采用 M 组 Mel 滤波器 $\{H_m(k), m=0, 1, 2, \dots, M-1\}$ 对 $Z(t, k)$ 进行处理，每个 Mel 滤波器的输出能量 $E(t, m)$ 为：

$$E(t, m) = \sum_{k=0}^{N-1} |Z(t, k)|^2 H_m(k), m=0, 1, 2, \dots, M-1 \quad (15)$$

这里 $H_m(k)$ 为 Mel 滤波器组第 m 个滤波器的频响应函数，它定义为：

$$H_m(k) = \begin{cases} 0, & k < f_{m-1} \\ \frac{k-f_{m-1}}{f_m-f_{m-1}}, & f_{m-1} \leq k < f_m \\ \frac{f_{m+1}-k}{f_{m+1}-f_m}, & f_m \leq k < f_{m+1} \\ 0, & k > f_{m+1} \end{cases} \quad (16)$$

其中 f_m 为第 m 个 Mel 滤波器的中心频率，它由下式定义：

$$f_m = \frac{N}{F_s} B^{-1} \left[B(f_{low}) + m \frac{B(f_{high}) - B(f_{low})}{M+1} \right] \quad (17)$$

式 (17) 中 f_{low} 和 f_{high} 分别为 Mel 滤波器组的最低和最高频率， F_s 为采样率， M 为滤波器组的数目，函数 $B^{-1}(b) = 700 \left\{ \exp \left(\frac{b}{1125} \right) - 1 \right\}$ 。

步骤 305：取对数：

首先对式 (17) 取对数，获得 Mel 滤波器组中每个滤波器输出的对数能量 $S(t, m)$ ：

$$S(t, m) = \log_e E(t, m), m=0, 1, \dots, M-1 \quad (18)$$

步骤 306：离散余弦变换：

然后对 (18) 式经离散余弦变换 (DCT) 变换得 MFCC 的系数如下：

$$C_n(t) = \sum_{m=0}^{M-1} S(t, m) \cdot \cos\left(\frac{\pi n(m+0.5)}{M}\right), n=0, 1, \dots, M-1 \quad (19)$$

由此提取到输入信号的声纹特征矢量 $\overrightarrow{VP}_x(t)$ 为：

$$\overrightarrow{VP}_x(t) = [C_0(t), C_1(t), \dots, C_{M-1}(t)]^T$$

根据上述提取信号声纹特征的流程，可以提取第一输入信号的第一声纹特征 $\overrightarrow{VP}_1$ ，以及第二输入信号第二声纹特征 $\overrightarrow{VP}_2$ 。

在步骤 203 中，根据式 (20) 计算第一声纹特征 $\overrightarrow{VP}_1$ 与第二声纹特征 $\overrightarrow{VP}_2$ 之间的距离 D ：

$$D \triangleq \|\overrightarrow{VP}_x - \overrightarrow{VP}_y\| \quad (20)$$

其中， $\|\cdot\|$ 为矢量的范数，可为 1-范数、2-范数或者 ∞ -范数。

最后，在步骤 204 中，当第一声纹特征 $\overrightarrow{VP}_1$ 与第二声纹特征 $\overrightarrow{VP}_2$ 之间的距离 D 大于或等于门限值 T 时（为了与其他门限值相区别，此处可称该门限值为第一门限值），确定第二输入信号中包含近端语音信号，即发生双讲，否则确定未发生双讲，即处于单讲状态，具体如式 (21) 所示：

$$\left. \begin{array}{l} \text{如果 } D \geq T, \quad \text{则确定发生双讲} \\ \text{如果 } D < T, \quad \text{则确定未发生双讲} \end{array} \right\} \quad (21)$$

在确定发生双讲之后，向移动终端的自适应滤波器发送指示信息，所述指示信息用于指示所述自适应滤波器暂停更新滤波器系数。

根据以上图 2 所示流程的描述，图 4 和图 5 分别示出了两种具体应用场景的示意图。

图 4 示出了利用麦克风输出信号 $y(n)$ 和自适应滤波器输出信号 $\hat{x}_2(n)$ 来进行双讲检测的实施例。如图 4 所示，远端输入信号 $x(n)$ 经过自适应滤波器滤波后形成 $\hat{x}_2(n)$ ， $y(n)$ 为麦克风输出信号。对麦克风输出信号 $y(n)$ 和 $\hat{x}_2(n)$ 分别进行声纹特征提取，对所提取的声纹特征矢量进行匹配处理，若两路信号的声纹特征矢量是模式匹配的，则判为单讲状态；否则，判为双讲状态。这里所提取的声纹特征矢量可以是 MFCC 型特征参数，也可以是能有效表征和鉴别输入信号的任何其它类型的特征参数。所采用的“模式匹配”技术可以是特征矢量间的距离匹配技术，也可以是特征矢量间的其它“相似度”匹配技术。

图 5 给出了利用麦克风输出信号 $y(n)$ 和远端输入信号 $x(n)$ 来进行双讲检测的实施例。如图所示，对 $x(n)$ 通过延时单元进行延时处理后进行特征提取，所延时的长度由声学回声路径延时估计算法决定，并对 $y(n)$ 进行特征提取；然后，对所提取的声纹特征矢量进行匹配处理，若两路信号的声纹特征矢量是模式匹配的，则判为单讲状态；否则，判为双讲状态。这里所提取的声纹特征矢量可以是 MFCC 型特征参数，也可以是能有效表征和鉴别输入信号的任何其它类型的特征参数。所采用的“模式匹配”技术可以是特征矢量间的距离匹配技术，也可以是特征矢量间的其它“相似度”匹配技术。

上述实施例中，通过将第一输入信号的第一声纹特征与第二输入信号的第二声纹特征进行比较，在第一声纹特征与第二声纹特征相近时，认为第一输入信号与第二输入信号中均包含远端信号，且第二输入信号中不包含近端语音信号，因此可以认为并未发生双讲，否则认为发生双讲。

由于语音信号是非平稳信号，表现在时域或频域上为非连续信号。因此并

不需要一直检测第一输入信号的第一声纹特征，或第二输入信号的第二声纹特征，可以先检测第一输入信号或第二输入信号中是否有语音信号，如果存在语音信号，则提取第一输入信号或第二输入信号的声纹特征。下面通过具体的实施例来详细描述。

如图 6 所示，本公开的一些实施例提供的一种近端语音信号检测方法，包括：

步骤 601：接收第一输入信号以及第二输入信号，其中，所述第一输入信号为移动终端接收到的远端信号，所述第二输入信号为所述移动终端接收到的近端信号；

步骤 602：检测所述第一输入信号是否大于第二门限值，以及检测所述第二输入信号是否大于第三门限值；以及

步骤 603：若所述第一输入信号大于所述第二门限值，且所述第二输入信号大于所述第三门限值，则提取所述第一输入信号的第一声纹特征，以及提取所述第二输入信号的第二声纹特征，确定所述第一声纹特征与所述第二声纹特征之间的距离，并根据所述距离确定所述第二输入信号中是否存在近端语音信号。

本公开的一些实施例中的移动终端可以为手机、平板电脑、会议电话等设备。

在步骤 601 中接收到的第一输入信号为远端信号。远端信号是经过编码、调制，并需要被扬声器等设备播放的信号。

第二输入信号，即近端信号，是由麦克风等音频接收传感器接收到的信号，可能包括远端信号经过回声路径形成的声学回波信号、环境噪声信号以及近端语音信号中的一种或多种组合，第二输入信号中的声学回波信号是需要消除的信号。

在步骤 602 中，分别检测第一输入信号以及第二输入信号中是否具有语音信号特征的信号，检测的方法有多种，可以通过语音活性检测算法进行检测，也可以通过其他方法进行检测，本公开的一些实施例对此并不限定。

在步骤 602 中,第二门限值可以是预设的信号能量与噪声能量比的短时能量差,当检测到第一输入信号的短时能量差高于第二门限值时,确定第一输入信号为语音信号。对应的,第三门限值可以是预设的信号能量与噪声能量比的短时能量差

在步骤 602 中,由于第二输入信号中包含由远端信号经过回声路径形成的声学回波信号时,会产生一定的延时,导致与第一输入信号之间不同步,第二输入信号相对于第一输入信号有一定的滞后。如果不对第一输入信号进行延时处理,直接检测第一输入信号是否大于所述第二门限值,那么需要将第一输入信号的检测结果与第二输入信号在经过延时后的检测结果相比较;如果对第一输入信号进行延时处理,那么需要将第一输入信号的检测结果与同一时间点第二输入信号的检测结果相比较。

综上所述,检测第一输入信号在第一时间点是否大于所述第二门限值,以及检测第二输入信号在第二时间点是否大于第三门限值,其中,第二时间点为第一时间点经过延时后的时间点,延时的时间长度可以根据实际情况确定。由上面的描述可知,延时的时间长度的取值可以分为下面两种情况:

第一种,不对第一输入信号进行延时处理,此时延时的时间长度大于 0,即第二时间点为第一时间点之后的时间点;延时的时间长度的具体取值可以根据远端信号在回波路径中的延时确定;

第二种,对第一输入信号进行延时处理,此时延时的时间长度等于 0,即第二时间点与第一时间点重合。

最后,在步骤 603 中,对第一输入信号和第二输入信号的检测结果可以分为以下三种情况:

一、若第二输入信号小于第三门限值,则确定第二输入信号在第二时间点不存在近端语音信号;

二、若第一输入信号小于第二门限值,且第二输入信号大于第三门限值,则确定第二输入信号在第二时间点存在近端语音信号。

三、若第一输入信号大于第二门限值,且第二输入信号大于第三门限值,

则提取第一输入信号的第一声纹特征，以及提取第二输入信号的第二声纹特征，确定第一声纹特征与第二声纹特征的距离，根据距离确定所述第二输入信号中是否存在近端语音信号。

第三种情况中，若第一声纹特征与第二声纹特征的距离小于第四门限值，则确定第二输入信号在第二时间点不存在近端语音信号，否则，确定第二输入信号在所述第二时间点存在近端语音信号。其中，这里的“第四门限值”与图 2 所示流程中的“第一门限值”取值可以相同也可以不同。

具体如何提取第一输入信号的第一声纹特征，以及提取第二输入信号的第二声纹特征，可以参考前一实施例的描述，在此不再赘述。

当确定第二输入信号中存在近端语音信号之后，向移动终端的自适应滤波器发送指示信息，所述指示信息用于指示所述自适应滤波器暂停更新滤波器系数。

根据以上图 6 所示流程的描述，图 7 示出了两种具体应用场景的示意图。

图 7 给出了基于 VAD 并利用麦克风输出信号 $y(n)$ 和远端输入信号 $x(n)$ 来进行双讲检测的实施例。如图所示，对远端输入信号 $x(n)$ 进行 VAD 监测，如果有语音信号，则对信号 $x(n)$ 提取声纹特征矢量 VP_x ，否则，不作处理。对下行链路中的麦克风输出信号 $y(n)$ 进行 VAD 监测，如果有语音信号，则对信号 $y(n)$ 提取声纹特征矢量 VP_y ，否则，不作处理。在声纹特征矢量 VP_x 可使用时开始等待直到声纹特征矢量 VP_y 可使用时即刻进行模式匹配处理。具体如下：

为了方便，标记下行链路的 VAD 在第 t 个时刻的值为 $DL_VAD(t)$ ，上行链路的 VAD 在第 t 个时刻的值为 $UL_VAD(t)$ ，如果 $DL_VAD(t)=0$ 并且 $UL_VAD(t)=1$ 时，则判定为双讲；如果 $DL_VAD(t)=0$ 并且 $UL_VAD(t)=0$ 时，则判定为单讲；如果 $DL_VAD(t)=1$ 并且 $UL_VAD(t+t_0)=1$ （这里 $t_0>0$ ）时，则按声纹识别技术判决是否为双讲。这里所提取的声纹特征矢量可以是 MFCC 型特征参数，也可以是能有效表征和鉴别输入信号的任何其它类型的特征参数。所采用的“模式匹配”技术可以是特征矢量间的距离匹配技术，也可以是特征矢量间的其它“相似度”匹配技术。

针对上述方法流程，本公开的一些实施例还提供一种近端语音信号检测装置，该装置的具体内容可以参照上述方法实施，在此不再赘述。

如图 8 所示，本公开的一些实施例提供了一种近端语音信号检测装置，包括：

接收单元 801，用于接收第一输入信号以及第二输入信号，其中，所述第一输入信号为移动终端接收到的远端信号被线性或非线性变换后所得到的信号，所述第二输入信号为所述移动终端接收到的近端信号；

提取单元 802，用于提取所述第一输入信号的第一声纹特征以及所述第二输入信号的第二声纹特征；以及

确定单元 803，用于确定所述第一声纹特征与所述第二声纹特征之间的距离，并根据所述距离确定所述第二输入信号中是否存在近端语音信号。

可选地，所述第一输入信号为所述移动终端的自适应滤波器输出的回波估计信号，所述回波估计信号是所述自适应滤波器对所述远端信号进行线性或非线性滤波得到的。

可选地，所述第一输入信号为所述远端信号被线性延时后得到的信号。

可选地，所述确定单元 803 具体用于：

判断所述距离是否小于第一门限值，若是，则确定所述第二输入信号中不存在近端语音信号，否则，确定所述第二输入信号中存在近端语音信号。

可选地，所述确定单元 803 还用于：

向所述移动终端的自适应滤波器发送指示信息，所述指示信息用于指示所述自适应滤波器暂停更新滤波器系数。

如图 9 所示，本公开的一些实施例提供了一种近端语音信号检测装置，包括：

接收单元 901，用于接收第一输入信号以及第二输入信号，其中，所述第一输入信号为移动终端接收到的远端信号，所述第二输入信号为所述移动终端接收到的近端信号；

检测单元 902，用于检测所述第一输入信号是否大于第二门限值，以及检

测所述第二输入信号是否大于第三门限值；以及

确定单元 903，用于在确定所述第一输入信号大于所述第二门限值，且所述第二输入信号大于所述第三门限值时，提取所述第一输入信号的第一声纹特征，以及提取所述第二输入信号的第二声纹特征，确定所述第一声纹特征与所述第二声纹特征之间的距离，并根据所述距离确定所述第二输入信号中是否存在近端语音信号。

可选地，所述检测单元 902 用于：

检测所述第一输入信号在第一时间点是否大于所述第二门限值，以及检测所述第二输入信号在第二时间点是否大于所述第三门限值，其中，所述第二时间点为所述第一时间点经过延时之后的时间点。

可选地，所述确定单元 903 具体用于：

若所述距离小于第四门限值，则确定所述第二输入信号在所述第二时间点不存在近端语音信号，否则，确定所述第二输入信号在所述第二时间点存在近端语音信号。

可选地，所述确定单元 903 还用于：

若所述第二输入信号小于所述第三门限值，则确定所述第二输入信号在所述第二时间点不存在近端语音信号；或者，

若所述第一输入信号小于所述第二门限值，且所述第二输入信号大于所述第三门限值，则确定所述第二输入信号在所述第二时间点存在近端语音信号。

可选地，所述确定单元 903 还用于：

向所述移动终端的自适应滤波器发送指示信息，其中，所述指示信息用于指示所述自适应滤波器暂停更新滤波器系数。

综上所述，根据本公开的一些实施例提供的方法及装置，本公开的一些实施例中通过提取远端语音信号的第一声纹特征，以及音频接收设备的输出信号中的第二声纹特征，通过对比第一声纹特征与第二声纹特征确定是否发生双讲。通过本公开的一些实施例提供的方法，避免了现有技术中以假设声学回波路径中非线性失真很小而忽略不计以及假设环境噪声平稳为前提条件去检测是否

发生双讲时产生的误判等情况，从而更准确的实现双讲检测。

本领域内的技术人员应明白，本公开的一些实施例可提供为方法、系统、或计算机程序产品。因此，本公开可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本公开可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器和光学存储器等）上实施的计算机程序产品的形式。

本公开是参照根据本公开的一些实施例的方法、设备（系统）、和计算机程序产品的流程图和 / 或方框图来描述的。应理解可由计算机程序指令实现流程图和 / 或方框图中的每一流程和 / 或方框、以及流程图和 / 或方框图中的流程和 / 或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的步骤。

显然，本领域的技术人员可以对本公开进行各种改动和变型而不脱离本公开的精神和范围。这样，倘若本公开的这些修改和变型属于本公开的权利要求及其等同技术的范围之内，则本公开也意图包含这些改动和变型在内。

权利要求书

1、一种近端语音信号检测方法，包括：

接收第一输入信号以及第二输入信号，其中，所述第一输入信号为移动终端接收到的远端信号被线性或非线性变换后所得到的信号，所述第二输入信号为所述移动终端接收到的近端信号；

提取所述第一输入信号的第一声纹特征以及所述第二输入信号的第二声纹特征；

确定所述第一声纹特征与所述第二声纹特征之间的距离；以及

根据所述距离确定所述第二输入信号中是否存在近端语音信号。

2、根据权利要求 1 所述的方法，其中，所述第一输入信号为所述移动终端的自适应滤波器输出的回波估计信号，其中，所述回波估计信号是所述自适应滤波器对所述远端信号进行线性或非线性滤波得到的。

3、根据权利要求 1 所述的方法，其中，所述第一输入信号为所述远端信号被线性延时后得到的信号。

4、根据权利要求 1 所述的方法，其中，所述根据所述距离确定所述第二输入信号中是否存在近端语音信号，包括：

判断所述距离是否小于第一门限值，若是，则确定所述第二输入信号中不存在近端语音信号，否则，确定所述第二输入信号中存在近端语音信号。

5、根据权利要求 1 至 4 中任一项所述的方法，其中，所述确定所述第二输入信号中存在近端语音信号之后，还包括：

向所述移动终端的自适应滤波器发送指示信息，其中，所述指示信息用于指示所述自适应滤波器暂停更新滤波器系数。

6、一种近端语音信号检测方法，包括：

接收第一输入信号以及第二输入信号，其中，所述第一输入信号为移动终端接收到的远端信号，所述第二输入信号为所述移动终端接收到的近端信号；以及

检测所述第一输入信号是否大于第二门限值，以及检测所述第二输入信号是否大于第三门限值；

若所述第一输入信号大于所述第二门限值，且所述第二输入信号大于所述第三门限值，则提取所述第一输入信号的第一声纹特征，以及提取所述第二输入信号的第二声纹特征，确定所述第一声纹特征与所述第二声纹特征之间的距离，并根据所述距离确定所述第二输入信号中是否存在近端语音信号。

7、根据权利要求 6 所述的方法，其中，所述检测所述第一输入信号是否大于第二门限值，以及检测所述第二输入信号是否大于第三门限值，包括：

检测所述第一输入信号在第一时间点是否大于所述第二门限值，以及检测所述第二输入信号在第二时间点是否大于所述第三门限值，其中，所述第二时间点为所述第一时间点经过延时后的时间点。

8、根据权利要求 7 所述的方法，其中，所述根据所述距离确定所述第二输入信号中是否存在近端语音信号，包括：

若所述距离小于第四门限值，则确定所述第二输入信号在所述第二时间点不存在近端语音信号，否则，确定所述第二输入信号在所述第二时间点存在近端语音信号。

9、根据权利要求 7 所述的方法，还包括：

若所述第二输入信号小于所述第三门限值，则确定所述第二输入信号在所述第二时间点不存在近端语音信号；或者，

若所述第一输入信号小于所述第二门限值，且所述第二输入信号大于所述第三门限值，则确定所述第二输入信号在所述第二时间点存在近端语音信号。

10、根据权利要求 6 至 9 任一项所述的方法，其中，确定所述第二输入信号中存在近端语音信号之后，还包括：

向所述移动终端的自适应滤波器发送指示信息，其中，所述指示信息用于指示所述自适应滤波器暂停更新滤波器系数。

11、一种近端语音信号检测装置，包括：

接收单元，用于接收第一输入信号以及第二输入信号，其中，所述第一输

入信号为移动终端接收到的远端信号被线性或非线性变换后所得到的信号，所述第二输入信号为所述移动终端接收到的近端信号；

提取单元，用于提取所述第一输入信号的第一声纹特征以及所述第二输入信号的第二声纹特征；以及

确定单元，用于确定所述第一声纹特征与所述第二声纹特征之间的距离，并根据所述距离确定所述第二输入信号中是否存在近端语音信号。

12、根据权利要求 11 所述的装置，其中，所述第一输入信号为所述移动终端的自适应滤波器输出的回波估计信号，所述回波估计信号是所述自适应滤波器对所述远端信号进行线性或非线性滤波得到的。

13、根据权利要求 11 所述的装置，其中，所述第一输入信号为所述远端信号被线性延时后得到的信号。

14、根据权利要求 11 所述的装置，其中，所述确定单元具体用于：

判断所述距离是否小于第一门限值，若是，则确定所述第二输入信号中不存在近端语音信号，否则，确定所述第二输入信号中存在近端语音信号。

15、根据权利要求 11 至 14 中任一项所述的装置，其中，所述确定单元还用于：

向所述移动终端的自适应滤波器发送指示信息，其中，所述指示信息用于指示所述自适应滤波器暂停更新滤波器系数。

16、一种近端语音信号检测装置，包括：

接收单元，用于接收第一输入信号以及第二输入信号，其中，所述第一输入信号为移动终端接收到的远端信号，所述第二输入信号为所述移动终端接收到的近端信号；

检测单元，用于检测所述第一输入信号是否大于第二门限值以及检测所述第二输入信号是否大于第三门限值；以及

确定单元，用于在确定所述第一输入信号大于所述第二门限值且所述第二输入信号大于所述第三门限值时，提取所述第一输入信号的第一声纹特征，以及提取所述第二输入信号的第二声纹特征，确定所述第一声纹特征与所述第二

声纹特征之间的距离，并根据所述距离确定所述第二输入信号中是否存在近端语音信号。

17、根据权利要求 16 所述的装置，其中，所述检测单元用于：

检测所述第一输入信号在第一时间点是否大于所述第二门限值，以及检测所述第二输入信号在第二时间点是否大于所述第三门限值，其中，所述第二时间点为所述第一时间点经过延时后的时间点。

18、根据权利要求 17 所述的装置，其中，所述确定单元具体用于：

若所述距离小于第四门限值，则确定所述第二输入信号在所述第二时间点不存在近端语音信号，否则，确定所述第二输入信号在所述第二时间点存在近端语音信号。

19、根据权利要求 17 所述的装置，其中，所述确定单元还用于：

若所述第二输入信号小于所述第三门限值，则确定所述第二输入信号在所述第二时间点不存在近端语音信号；或者，

若所述第一输入信号小于所述第二门限值，且所述第二输入信号大于所述第三门限值，则确定所述第二输入信号在所述第二时间点存在近端语音信号。

20、根据权利要求 16 至 19 任一项所述的装置，其中，所述确定单元还用于：

向所述移动终端的自适应滤波器发送指示信息，其中，所述指示信息用于指示所述自适应滤波器暂停更新滤波器系数。

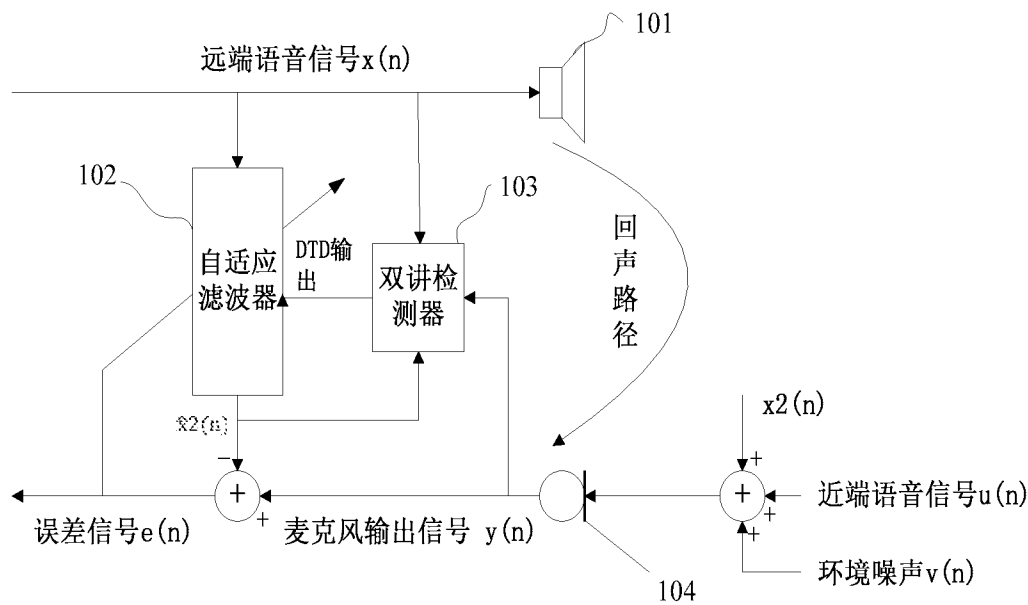


图 1

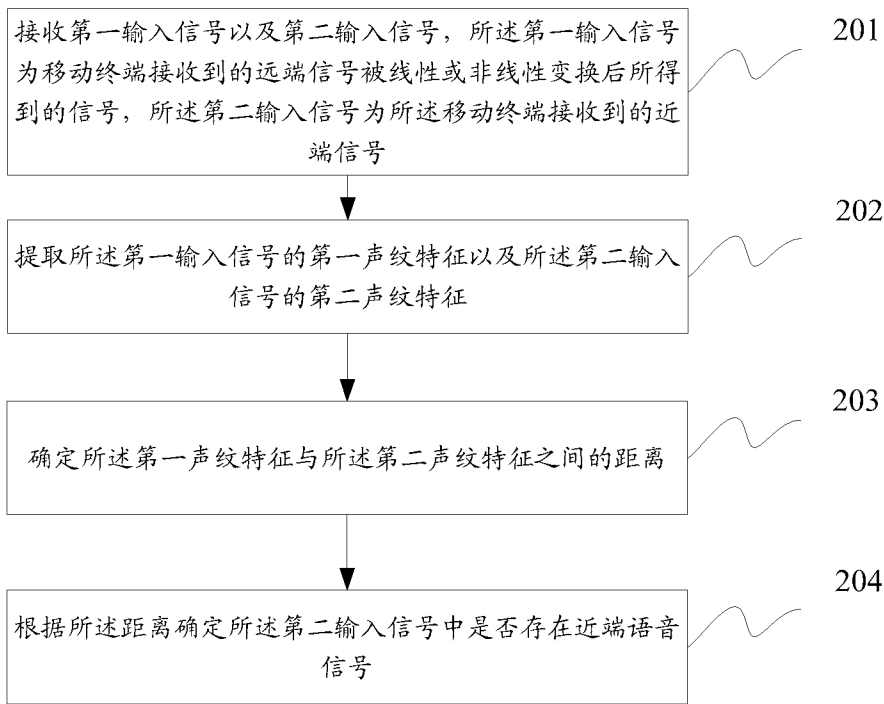


图 2

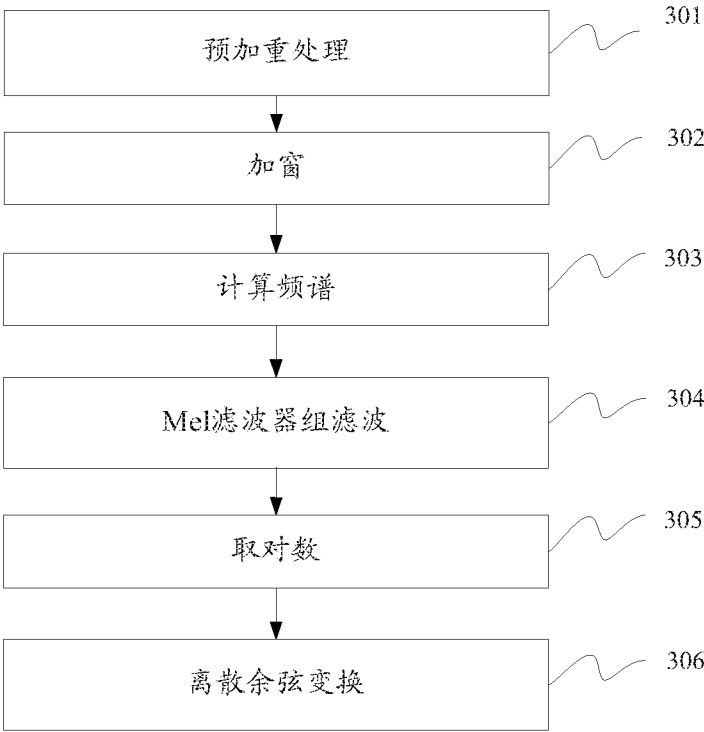


图 3

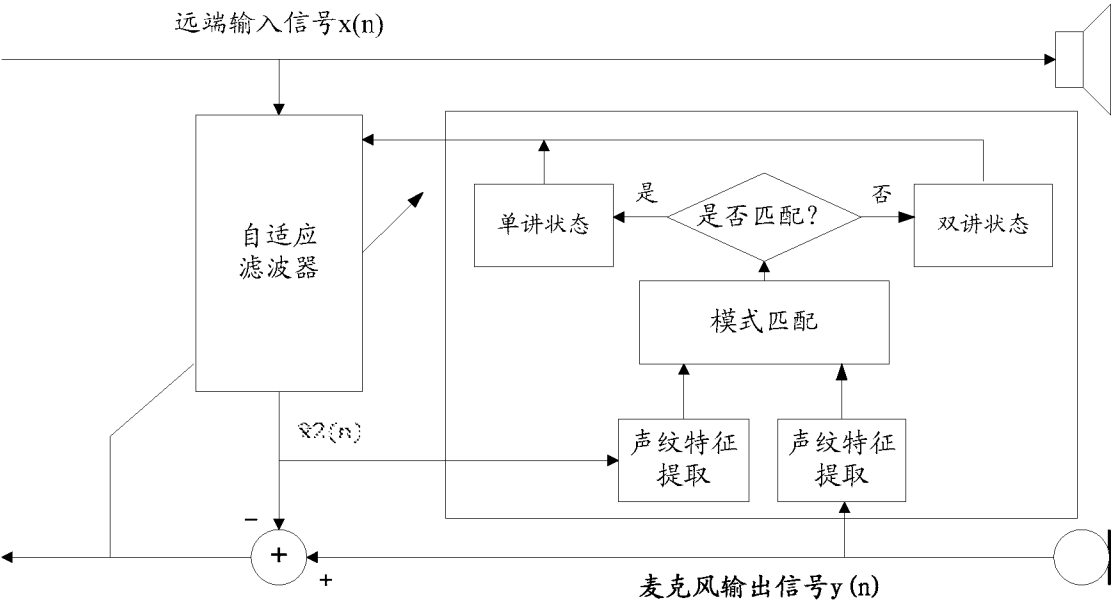


图 4

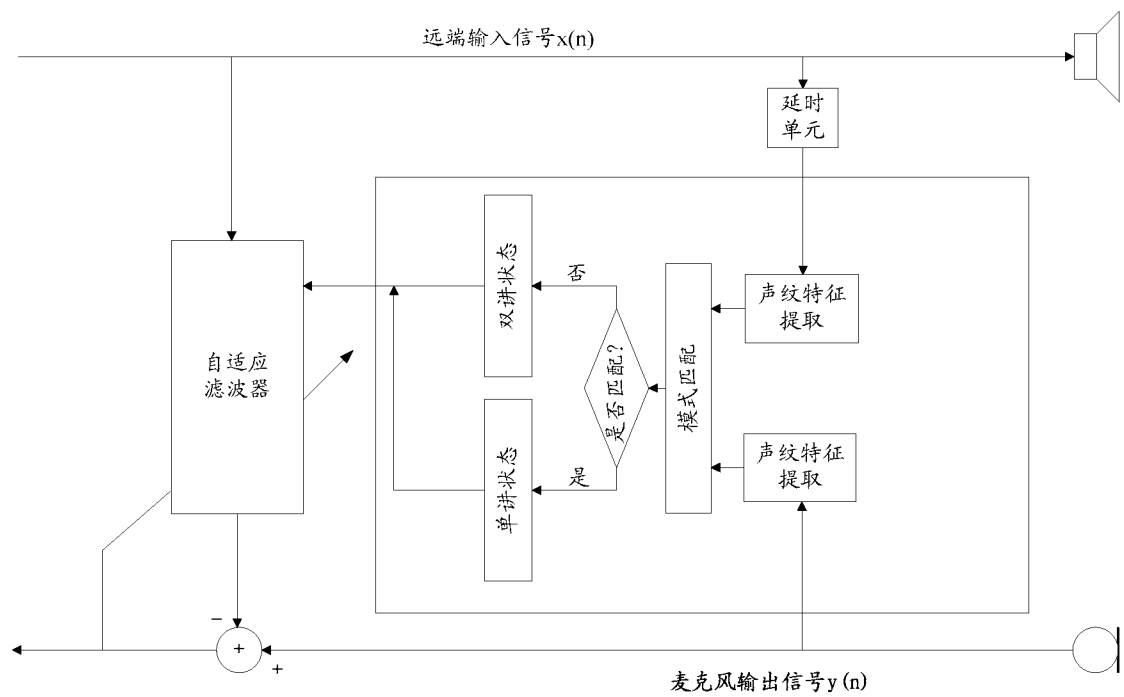


图 5

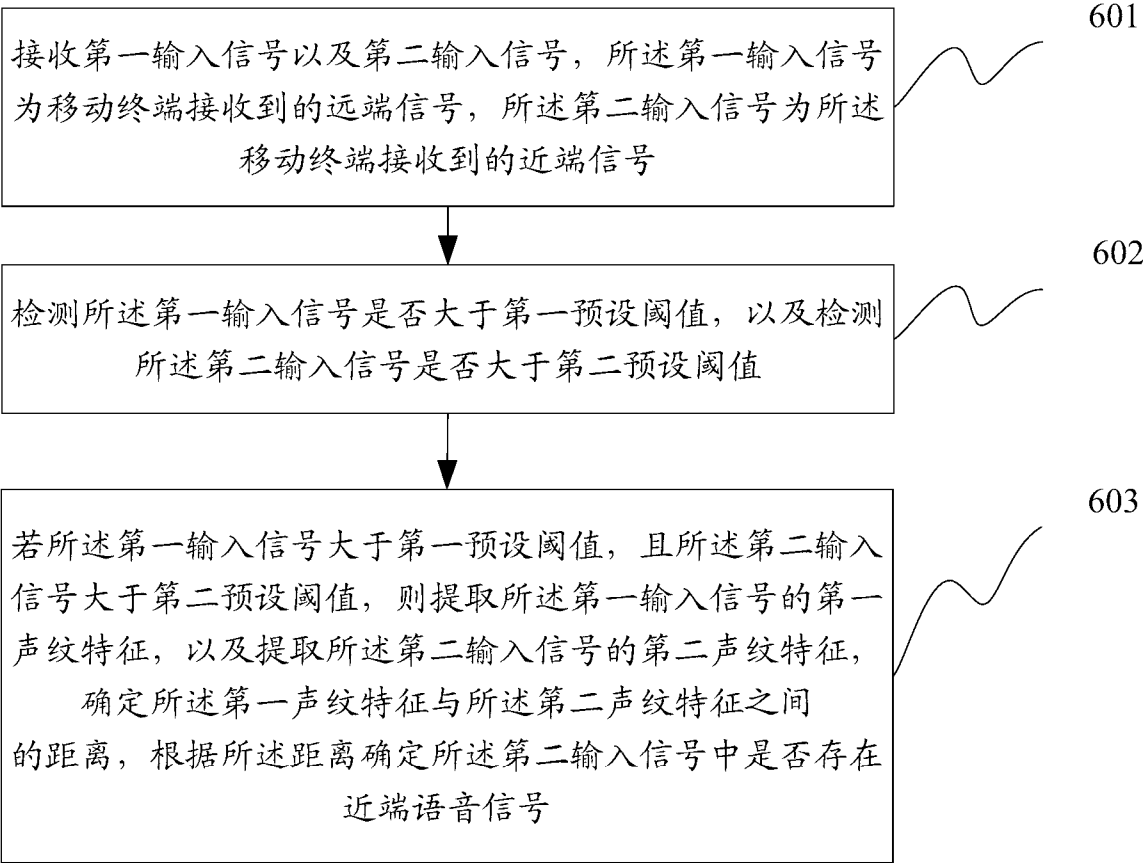


图 6

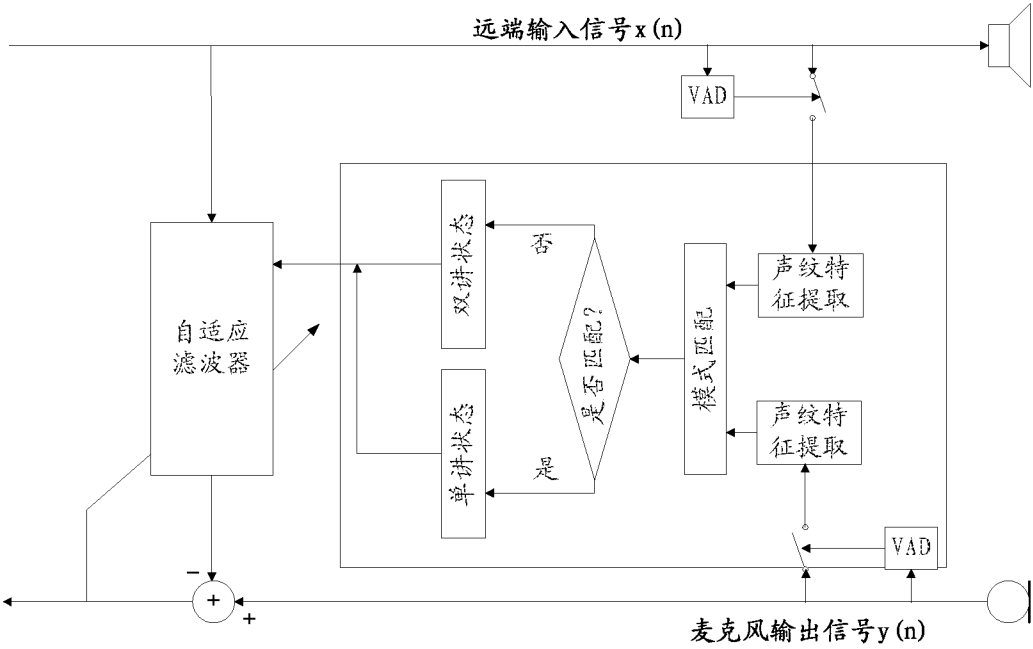


图 7

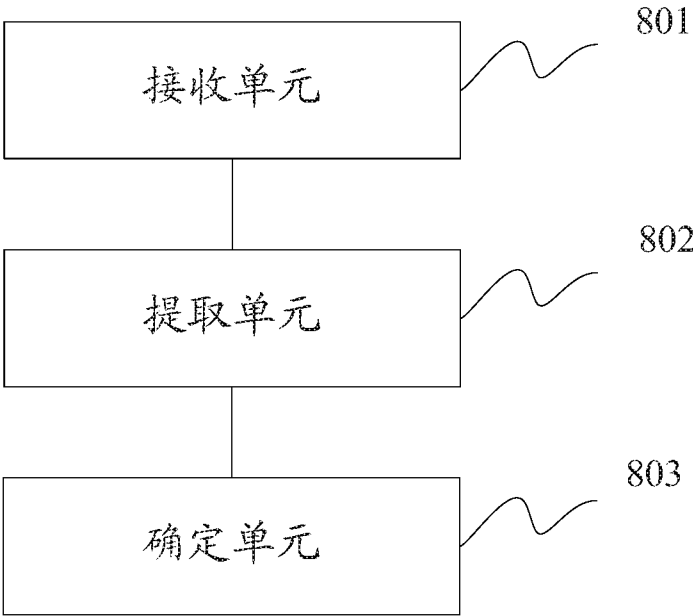


图 8

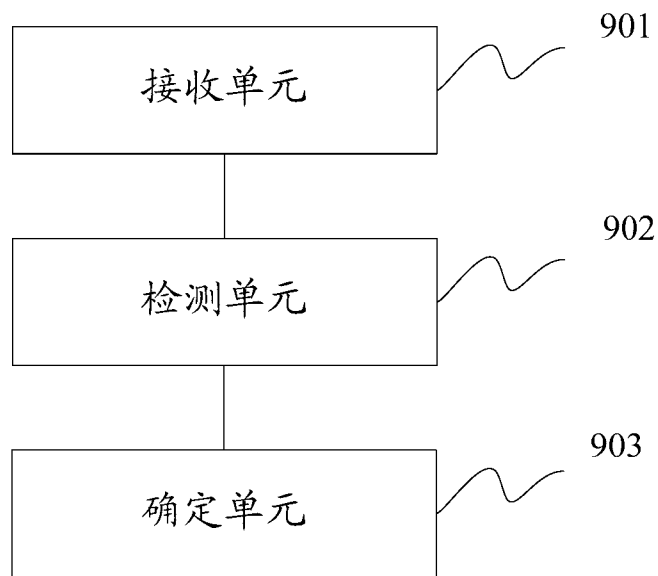


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/070253

A. CLASSIFICATION OF SUBJECT MATTER

G10L 25/78 (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 11/-; H04M 3/-; G10L 15/-

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, TWABS, TWTXT, CNKI, WPI, EPODOC: CHINA ACADEMY OF TELECOMMUNICATIONS TECHNOLOGY; LIANG, Min; HAN, Bo; distance, voice print, characteristic parameter, far-end, near-end, double-talk, estimation signal, echo signal, parameter, sound, outlying, dipped, linearity, non-linear, estimate

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 102137194 A (HUAWEI DEVICE CO., LTD.), 27 July 2011 (27.07.2011), claims 1 and 2, and description, paragraphs 28- 63	1-20
A	EP 0640953 A1 (CANON KABUSHIKI KAISHA), 01 March 1995 (01.03.1995), the whole document	1-20
A	CN 103337242 A (HUAWEI TECHNOLOGIES CO., LTD.), 02 October 2013 (02.10.2013), the whole document	1-20
A	CN 103905656 A (LEADCORE TECHNOLOGY CO., LTD.), 02 July 2014 (02.07.2014), the whole document	1-20
A	CN 1584977 A (ZTE CORP.), 23 February 2005 (23.02.2005), the whole document	1-20

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 10 March 2016 (10.03.2016)	Date of mailing of the international search report 24 March 2016 (24.03.2016)
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 LIU, Zihan Telephone No.: (86-10) 010-62414489

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/070253

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102137194 A	27 July 2011	CN 102137194 B	01 January 2014
EP 0640953 A1	01 March 1995	EP 0640953 B1	04 July 2001
		US 5764779 A	09 June 1998
		DE 69427621 T2	08 May 2002
		DE 69427621 D1	09 August 2001
		JP H0764578 A	10 March 1995
		JP H07334192 A	22 December 1995
CN 103337242 A	02 October 2013	None	
CN 103905656 A	02 July 2014	None	
CN 1584977 A	23 February 2005	CN 1299254 C	07 February 2007

A. 主题的分类 G10L 25/78(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) G10L11/-;H04M3/-;G10L15/- 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, TWABS, TWTXT, CNKI, WPI, EPODOC:电信科学技术研究院, 梁民, 韩波, 距离, 远近, 参数, 声纹, 特征参数, 远端, 近端, 双讲, 线性, 非线性, 估计信号, 回波信号, parameter, sound, outlying, dipped, linearity, non-linear, estimate		
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
X	CN 102137194 A (华为终端有限公司) 2011年 7月 27日 (2011 - 07 - 27) 权利要求1-2、说明书第28-63段	1-20
A	EP 0640953 A1 (CANON KABUSHIKI KAISHA) 1995年 3月 1日 (1995 - 03 - 01) 全文	1-20
A	CN 103337242 A (华为技术有限公司) 2013年 10月 2日 (2013 - 10 - 02) 全文	1-20
A	CN 103905656 A (联芯科技有限公司) 2014年 7月 2日 (2014 - 07 - 02) 全文	1-20
A	CN 1584977 A (中兴通讯股份有限公司) 2005年 2月 23日 (2005 - 02 - 23) 全文	1-20
☐ 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 2016年 3月 10日		国际检索报告邮寄日期 2016年 3月 24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		授权官员 刘子菡 电话号码 (86-10)010-62414489

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/070253

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	102137194	A	2011年 7月 27日	CN 102137194 B	2014年 1月 1日
EP	0640953	A1	1995年 3月 1日	EP 0640953 B1	2001年 7月 4日
				US 5764779 A	1998年 6月 9日
				DE 69427621 T2	2002年 5月 8日
				DE 69427621 D1	2001年 8月 9日
				JP H0764578 A	1995年 3月 10日
				JP H07334192 A	1995年 12月 22日
CN	103337242	A	2013年 10月 2日	无	
CN	103905656	A	2014年 7月 2日	无	
CN	1584977	A	2005年 2月 23日	CN 1299254 C	2007年 2月 7日