(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 8 月 3 日 (03.08.2017)

WIPO | PCT

(10) 国际公布号
WO 2017/128910 A1

(51) 国际专利分类号:
G10L 21/0216 (2013.01) *G10L 21/0224* (2013.01)

(21) 国际申请号: PCT/CN2016/112323

(22) 国际申请日: 2016 年 12 月 27 日 (27.12.2016)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201610049402.X 2016 年 1 月 25 日 (25.01.2016) 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: 汪法兵 (WANG, Fabing); 中国北京市海淀区学院路 40 号, Beijing 100191 (CN)。 梁民 (LIANG, Min); 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, DJ, 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, KH, KN, KP, KR, KW, 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, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

[见续页]

(54) Title: METHOD, APPARATUS AND ELECTRONIC DEVICE FOR DETERMINING SPEECH PRESENCE PROBABILITY

(54) 发明名称 : 一种语音出现概率的确定方法、装置及电子设备

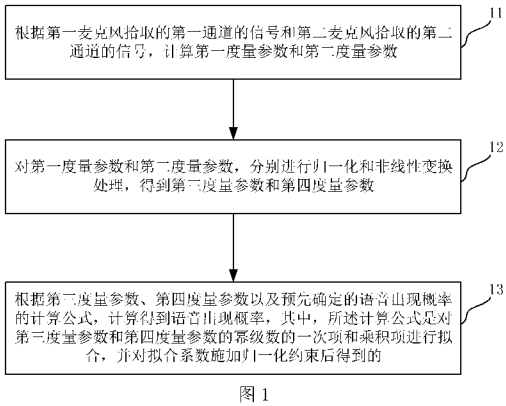


图 1

11 CALCULATING A FIRST MEASUREMENT PARAMETER AND A SECOND MEASUREMENT PARAMETER ACCORDING TO A FIRST CHANNEL SIGNAL PICKED UP BY THE FIRST MICROPHONE AND A SECOND CHANNEL SIGNAL PICKED UP BY THE SECOND MICROPHONE

12 PERFORMING NORMALIZATION AND NONLINEAR CONVERSION ON THE FIRST MEASUREMENT PARAMETER AND THE SECOND MEASUREMENT PARAMETER, RESPECTIVELY, TO OBTAIN A THIRD MEASUREMENT PARAMETER AND A FOURTH MEASUREMENT PARAMETER

13 CALCULATING TO OBTAIN A SPEECH PRESENCE PROBABILITY ACCORDING TO THE THIRD MEASUREMENT PARAMETER, THE FOURTH MEASUREMENT PARAMETER AND A PRE-DETERMINED CALCULATION EQUATION FOR SPEECH PRESENCE PROBABILITY, WHEREIN SAID CALCULATION EQUATION IS OBTAINED BY PERFORMING FITTING ON LINEAR TERMS AND PRODUCT TERMS OF THE TWO-VARIABLE POWER SERIES OF THE THIRD MEASUREMENT PARAMETER AND FOURTH MEASUREMENT PARAMETER, AND THEN APPLYING NORMALIZED CONSTRAINTS ON A FITTING COEFFICIENT

(57) Abstract: A method, an apparatus and an electronic device for determining speech presence probability, to be applied to a first microphone and a second microphone set up by using an end-fire structure, comprising: calculating a first measurement parameter and a second measurement parameter according to a first channel signal picked up by the first microphone and a second channel signal picked up by the second microphone (11), said first measurement parameter being a signal-noise ratio of signals in the first channel, and said second measurement parameter being a difference between signal power levels in the first channel and in the second channel; performing normalization and nonlinear conversion on the first measurement parameter and the second measurement parameter, respectively, to obtain a third measurement parameter and a fourth measurement parameter (12); calculating to obtain a speech presence probability according to the third measurement parameter, the fourth measurement parameter and a pre-determined calculation equation for speech presence probability, wherein said calculation equation is obtained by performing fitting on linear terms and product terms of the two-variable power series of the third measurement parameter and fourth measurement parameter, and then applying normalized constraints on a fitting coefficient (13).

(57) 摘要:

[见续页]

WO 2017/128910 A1

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种语音出现概率的确定方法、装置及电子设备，应用于采用端射 End-fire 结构设置的第一麦克风和第二麦克风，该包括：根据第一麦克风拾取的第一通道的信号和第二麦克风拾取的第二通道的信号，计算第一度量参数和第二度量参数（11），该第一度量参数为第一通道的信号信噪比，第二度量参数为第一通道与第二通道的信号功率电平差；对第一度量参数和第二度量参数，分别进行归一化和非线性变换处理，得到第三度量参数和第四度量参数（12）；根据第三度量参数、第四度量参数以及预先确定的语音出现概率的计算公式，计算得到语音出现概率，其中，该计算公式是通过对第三度量参数和第四度量参数的二元幂级数的一次项和乘积项进行拟合，并对拟合系数施加归一化约束后得到的（13）。

一种语音出现概率的确定方法、装置及电子设备

相关申请的交叉引用

本申请主张在 2016 年 1 月 25 日在中国提交的中国专利申请号 No. 201610049402.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处理技术领域，具体涉及一种语音出现概率的确定方法、装置及电子设备。

背景技术

在正常的语音通话中，用户大约有 50%的时间段是处于停顿/倾听等非发话状态。相关技术中的语音增强系统则通过语音激活检测（Voice Activity Detection，VAD）算法来鉴别出语音非激活段，并在该段进行环境噪声统计特性的估计与更新。目前的 VAD 技术大都通过计算语音信号时域波形的过零率或短时能量等参数、并将其与预定的阈值作相比，来做出语音激活与否的二元判决。但这种简单的二元判定方法时常发生误判（即把语音段判定为非语音段或者把非语音段判定为语音段），由此影响环境噪声统计参数估计的准确性，从而降低语音增强系统的质量。

为克服 VAD 的这种局限性，人们提出了 VAD 的软判决技术。VAD 软判决技术则首先计算语音出现概率（Speech Presence Probability，SPP）或者语音缺席概率（Speech Absence Probability，SAP），然后再利用 SPP 或 SAP 来估计噪声的统计信息。然而，对于双麦克风语音增强系统，相关技术中的计算语音出现概率的方法，大多存在计算量大，对参数波动敏感，以及在语音非激活段不趋近于零的缺点。

发明内容

本公开实施例要解决的技术问题是提供一种语音出现概率的确定方法、装置及电子设备，其计算复杂度较低，且对参数波动具有较好的健壮性，满足语

音非激活段语音出现概率趋近于零的约束条件，能够广泛应用于各种双麦克风语音增强系统。

为解决上述技术问题，本公开实施例提供的语音出现概率的确定方法，应用于采用端射 End-fire 结构设置的第一麦克风和第二麦克风，包括:

根据第一麦克风拾取的第一通道的信号和第二麦克风拾取的第二通道的信号，计算第一度量参数和第二度量参数，所述第一度量参数为第一通道的信号信噪比，第二度量参数为第一通道与第二通道的信号功率电平差;

对第一度量参数和第二度量参数，分别进行归一化和非线性变换处理，得到第三度量参数和第四度量参数;

根据第三度量参数、第四度量参数以及预先确定的语音出现概率的计算公式，计算得到语音出现概率，其中，所述计算公式是通过对第三度量参数和第四度量参数的二元幂级数的一次项和乘积项进行拟合，并对拟合系数施加归一化约束后得到的。

可选的，上述方案中，

所述第一度量参数的计算包括:

利用以下公式，计算第一度量参数：$M_{SNR}(n,k)=\frac{\xi_1(n,k)}{\xi_0(k)}$

其中，$M_{SNR}(n,k)$表示第一度量参数，$\xi_1(n,k)$表示第一通道的第 n 帧信号第 k 个频率分量上的先验信噪比，$\xi_0(k)$表示预先设定的第 k 个频率分量上的信噪比参考值。

可选的，上述方案中，

所述第二度量参数的计算包括:

利用以下公式，计算第二度量参数：$M_{PLD}(n,k)=\frac{\Phi_{y_1y_1}-\Phi_{y_2y_2}}{\Phi_{y_1y_1}+\Phi_{y_2y_2}}$

其中，$M_{PLD}(n,k)$表示第二度量参数，$\Phi_{y_1y_1}$表示第一通道的第 n 帧信号第 k 个频率分量上的信号功率谱密度，$\Phi_{y_2y_2}$表示第二通道的第 n 帧信号第 k 个频率分量上的信号功率谱密度。

可选的，上述方案中，

所述归一化和非线性变换处理包括:

对待处理参数进行数值更新，得到中间参数，其中，在数值超出区间[0，1]时，将数值更新为 1，否则保持数值不变，所述待处理参数为第一度量参数或第二度量参数;

对中间参数进行分段线性变换，得到最终参数，所述最终参数是所述中间参数的分段线性函数，且接近于所述中间参数取值范围中心的区段的斜率，大于远离所述中间参数取值范围中心的区段的斜率，所述最终参数为第三度量参数或第四度量参数。

可选的，上述方案中，

所述语音出现概率的计算公式为:

$$P_1 = \mathrm{c}(aM'_{SNR} + (1-a)M'_{PLD}) + (1-\mathrm{c})M'_{SNR}M'_{PLD}$$

其中，P_1表示第 n 帧信号第 k 个频率分量上的语音出现概率，M'_{SNR}表示第三度量参数，M'_{PLD}表示第四度量参数，a、c均为取值范围在[0，1]之内的拟合系数。

可选的，上述方案中，所述拟合系数a、c的取值是预先设定的固定值。

可选的，上述方案中，所述拟合系数a的取值是根据环境噪声的类型而预先设确定的;

所述拟合系数c的取值，随着M'_{SNR}与M'_{PLD}的差值的减小而增大。

其中，上述方案中，

拟合系数c的取值，按照以下任一公式计算得到:

$$\mathrm{c} = \frac{(M'_{PLD}+M'_{SNR}-1)^2}{(M'_{PLD}+M'_{SNR}-1)^2+(M'_{PLD}-M'_{SNR})^2};$$

$$\mathrm{c} = 1 - |M'_{PLD} - M'_{SNR}|$$

本公开实施例还提供了一种语音出现概率的确定装置，应用于采用端射End-fire 结构设置的第一麦克风和第二麦克风，包括:

采集单元，用于根据第一麦克风拾取的第一通道的信号和第二麦克风拾取的第二通道的信号，计算第一度量参数和第二度量参数，所述第一度量参数为

第一通道的信号信噪比，第二度量参数为第一通道与第二通道的信号功率电平差；

转换单元，用于对第一度量参数和第二度量参数，分别进行归一化和非线性变换处理，得到第三度量参数和第四度量参数；

计算单元，用于根据第三度量参数、第四度量参数以及预先确定的语音出现概率的计算公式，计算得到语音出现概率，其中，所述计算公式是通过对第三度量参数和第四度量参数的二元幂级数的一次项和乘积项进行拟合，并对拟合系数施加归一化约束后得到的。

可选的，上述方案中，

所述采集单元，具体用于：

利用以下公式，计算第一度量参数：$M_{SNR}(n,k)=\frac{\xi_1(n,k)}{\xi_0(k)}$

其中，$M_{SNR}(n,k)$表示第一度量参数，$\xi_1(n,k)$表示第一通道的第 n 帧信号第 k 个频率分量上的先验信噪比，$\xi_0(k)$表示预先设定的第 k 个频率分量上的信噪比参考值。

可选的，上述方案中，

所述采集单元，具体用于：

利用以下公式，计算第二度量参数：$M_{PLD}(n,k)=\frac{\Phi_{y_1y_1}-\Phi_{y_2y_2}}{\Phi_{y_1y_1}+\Phi_{y_2y_2}}$

其中，$M_{PLD}(n,k)$表示第二度量参数，$\Phi_{y_1y_1}$表示第一通道的第 n 帧信号第 k 个频率分量上的信号功率谱密度，$\Phi_{y_2y_2}$表示第二通道的第 n 帧信号第 k 个频率分量上的信号功率谱密度。

可选的，上述方案中，

所述转换单元，具体用于：对待处理参数进行数值更新，得到中间参数，其中，在数值超出区间[0，1]时，将数值更新为 1，否则保持数值不变，所述待处理参数为第一度量参数或第二度量参数；对中间参数进行分段线性变换，得到最终参数，所述最终参数是所述中间参数的分段线性函数，且接近于所述

中间参数取值范围中心的区段的斜率，大于远离所述中间参数取值范围中心的区段的斜率，所述最终参数为第三度量参数或第四度量参数。

可选的，上述方案中，

所述语音出现概率的计算公式为：

$$P_1 = \mathrm{c}(aM'_{SNR} + (1-a)M'_{PLD}) + (1-\mathrm{c})M'_{SNR}M'_{PLD}$$

其中，P_1表示第 n 帧信号第 k 个频率分量上的语音出现概率，M'_{SNR}表示第三度量参数，M'_{PLD}表示第四度量参数，a、c均为取值范围在[0，1]之内的拟合系数。

可选的，上述方案中，所述拟合系数a、c的取值是预先设定的固定值。

可选的，上述方案中，

所述拟合系数a的取值是根据环境噪声的类型而与确预先设定的；

所述拟合系数c的取值，随着M'_{SNR}与M'_{PLD}的差值的减小而增大。

其中，上述方案中，

拟合系数c的取值，按照以下任一公式计算得到：

$$\mathrm{c} = \frac{(M'_{PLD} + M'_{SNR} - 1)^2}{(M'_{PLD} + M'_{SNR} - 1)^2 + (M'_{PLD} - M'_{SNR})^2}$$

$$\mathrm{c} = 1 - |M'_{PLD} - M'_{SNR}|$$

本公开实施例还提供了一种电子设备，包括：

处理器；以及，通过总线接口与所述处理器相连接的存储器、第一麦克风和第二麦克风，所述第一麦克风和第二麦克风采用端射 End-fire 结构配置；所述存储器用于存储所述处理器在执行操作时所使用的程序和数据，当处理器调用并执行所述存储器中所存储的程序和数据时，实现如下的功能模块：

采集单元，用于分别采集第一麦克风对应的第一通道和第二麦克风对应的第二通道的声音信号，计算第一度量参数和第二度量参数，其中，所述第一度量参数为第一通道的信号信噪比，第二度量参数为第一通道与第二通道的信号功率电平差；

转换单元，用于对第一度量参数和第二度量参数，分别进行归一化和非线性变换处理，得到第三度量参数和第四度量参数；

计算单元，用于根据第三度量参数、第四度量参数以及预先确定的语音出现概率的计算公式，计算得到语音出现概率，其中，所述计算公式是通过对第三度量参数和第四度量参数的二元幂级数的一次项和乘积项进行拟合，并对拟合系数施加归一化约束后得到的。

与相关技术相比，本公开实施例提供的语音出现概率的确定方法、装置及电子设备，大大减少了语音出现概率计算的运算量，满足语音非激活段语音出现概率趋近于零的约束，且使得计算结果对参数波动具有较好的健壮性。另外，本公开实施例既能应用于稳态/准稳态噪声场中的情形，又能应用于瞬态噪声和第三方语音干扰的情形，能够广泛适用于各种双麦克风语音增强系统的应用场景。

附图说明

图 1 为本公开实施例提供的语音出现概率的确定方法的流程示意图；

图 2 为本公开实施例提供的语音出现概率的确定方法的又一流程示意图；

图 3 为本公开实施例中对第一度量参数进行分段线性变换的示意图；

图 4 为本公开实施例中对第二度量参数进行分段线性变换的示意图；

图 5 为本公开实施例中拟合系数的确定方式的举例示意图；

图 6 为本公开实施例提供的语音出现概率的确定装置的结构示意图；

图 7 为本公开实施例提供的电子设备的结构示意图。

具体实施方式

为使本公开要解决的技术问题、技术方案和优点更加清楚，下面将结合附图及具体实施例进行详细描述。

相关技术中的双麦克风语音增强系统的语音出现概率的确定方法，由于运算量非常大，且计算结果对参数波动敏感等缺点，以及语音非激活段不趋近于零，不能很好的适用于实际设备中。本公开实施例通过引入两种度量参数，并提出一种新的语音出现概率的确定模型，可以降低计算量，并使计算结果对参数波动具有更好的健壮性，并满足语音非激活段趋近于零的约束。

在介绍本公开实施例之前，为帮助更好的理解本公开，首先介绍一下相关

技术中的语音出现概率的计算原理。

假设麦克风拾取的信号为:

$$y(n) = x(n)+d(n) \qquad (1)$$

这里，$x(n)$是用户的语音信号，$d(n)$是噪声信号（包括环境噪声和其它声源干扰总和），$y(n)$为麦克风拾取的信号。

对上述公式（1）做短时傅里叶变换可以得到:

$$Y(n,k) = X(n,k)+ D(n,k) \qquad (2)$$

假设麦克风拾取信号存在如下表述的两种状态假设检验:

$$\begin{aligned} &H_0\text{（即无语音信号）:} \quad Y(n,k) = D(n,k) \\ &H_1\text{（即有语音信号）:} \quad Y(n,k) = X(n,k)+ D(n,k) \end{aligned} \qquad (3)$$

利用软判决方法 ,计算噪声功率谱:

$$E[|D|^2|Y]= E[|D|^2|Y,H_0]p(H_0|Y)+ E[|D|^2|Y,H_1]p(H_1|Y) \qquad (4)$$

上述公式（4）中， $p(H_1|Y)$是当前的时频单元的语音出现概率，$p(H_0|Y)$是当前的时频单元的语音缺席概率。

利用贝叶斯公式可以得到:

$$\begin{aligned} p(H_1|Y(n,k)) &= \frac{p(Y(n,k)\mid H_1)p(H_1)}{p(Y(n,k))} \\ &= \frac{p(Y(n,k)\mid H_1)p(H_1)}{p(Y(n,k)\mid H_1)p(H_1)+p(Y(n,k)\mid H_0)p(H_0)} \\ &= \frac{1}{1+\frac{p(H_0)}{p(H_1)}\frac{p(Y(n,k)\mid H_0)}{p(Y(n,k)\mid H_1)}} \\ &\triangleq \frac{1}{1+q\Lambda} \end{aligned} \qquad (5)$$

其中，$q=\frac{p(H_0)}{p(H_1)}$是语音缺席与语音出现的先验概率之比，$\Lambda=\frac{p(Y(n,k)\mid H_0)}{p(Y(n,k)\mid H_1)}$是麦克风拾取信号的第 n 帧信号第 k 个频点的条件概率之比，假设各频点振幅幅度是高斯分布，利用 MMSE-STSA 方法计算，则可以得到:

$$\Lambda = (1+\xi(n,k))\exp\left(-\frac{\gamma(n,k)\xi(n,k)}{1+\xi(n,k)}\right) \qquad (6)$$

上述公式（6）中，$\xi(n,k)$，$\gamma(n,k)$分别是麦克风拾取信号第 n 帧信号第 k 个频点的先验信噪比和后验信噪比。

上述公式（5）是相关技术中应用较广的单通道 SPP 计算方法。

近年来，双麦克风阵列已广泛地应用于移动终端提升语音增强的功能。双麦克风阵列通常包括采用端射 End-fire 结构设置的第一麦克风和第二麦克风，其中一个麦克风的部署位置通常更接近于用户嘴部。考虑到上述语音出现概率的计算方法则是基于单麦克风情况下推导的，它并不能完全适用于多麦克风系统。为此，相关技术中已经将上述方法拓展到多麦克风语音出现概率的计算上，通过基于高斯模型的语音出现概率之假设，推导出同公式（5）和（6）类似的理论公式：

$$P(H_1|\mathrm{Y}) = \frac{1}{1+q(1+\xi(n,k))\exp\left(-\frac{\beta(n,k)}{1+\xi(n,k)}\right)} \qquad (7)$$

上述公式（7）的参数$\xi(n,k)$，$\beta(n,k)$替换成如下多通道的计算公式：

$$\xi(n,k) \triangleq tr[\mathbf{\Phi}_{\mathbf{dd}}^{-1}(n,k)\mathbf{\Phi}_{\mathbf{xx}}(\mathrm{n,k})] \qquad (8)$$

$$\beta(n,k) \triangleq \mathbf{y}^{\mathrm{H}}(n,k)\mathbf{\Phi}_{\mathbf{dd}}^{-1}(n,k)\mathbf{\Phi}_{\mathbf{xx}}(n,k)\mathbf{\Phi}_{\mathbf{dd}}^{-1}(n,k)\mathrm{y}(n,k) \qquad (9)$$

其中，

$$\mathbf{y}(n,k) = [y_1(n,k)y_2(n,k)\ldots y_N(n,k)]^{\mathrm{T}},$$

$$\mathbf{X}(n,k) = [x_1(n,k)x_2(n,k)\ldots x_N(n,k)]^{T},$$

$$\mathbf{d}(n,k) = [d_1(n,k)d_2(n,k)\ldots d_N(n,k)]^{\mathrm{T}};$$

下标 N 为多麦克风阵列（如双麦克风阵列）的通道数，当用于双麦克风情形时，取 N=2；$\mathbf{\Phi}_{\mathbf{xx}}$，$\mathbf{\Phi}_{\mathbf{dd}}$分别是多通道语音信号和背景噪声的功率谱密度矩阵；$\mathbf{\Phi}_{\mathbf{xx}}(n,k) \triangleq \mathbf{E}\{\mathbf{x}(n,k)\mathbf{x}^{\mathrm{H}}(n,k)\} = \mathbf{\Phi}_{\mathbf{yy}}(n,k) - \mathbf{\Phi}_{\mathbf{dd}}(n,k)$，$\Phi_{\mathbf{dd}}(n,k) \triangleq \mathbf{E}\{\mathbf{d}(n,k)\mathbf{d}^{H}(n,k)\}$，期望值可通过递归计算逼近：

$$\mathbf{\Phi_{yy}}(n,k) = \big(1-\alpha_y\big)\mathbf{\Phi_{yy}}(n-1,k) + \alpha_y\,\mathbf{y}(n,k)\mathbf{y^H}\,(n,k) \tag{10}$$

$$\mathbf{\Phi_{dd}}(n,k) = (1-\alpha_d)\mathbf{\Phi_{dd}}(n-1,k) + \alpha_d\,\mathbf{d}(n,k)\mathbf{d^H}\,(n,k) \tag{11}$$

其中，$0 \le \alpha_y \le 1$，$0 \le \alpha_d \le 1$。

将上述公式（7）应用到双麦克风系统，便可得到双通道语音出现概率的计算公式。

然而，将上述的理论公式应用到移动终端时，存在计算量大，对参数敏感等问题。对于双麦克风语音增强系统，利用（7）式至（9）式计算 SPP，涉及大量的矩阵乘积与矩阵求逆运算，在实时处理的语音增强系统中，因为占用过多的计算资源而实用性低。其次，在现实应用环境中，语音和噪声信号大多是非稳态信号，经常性出现的第三方干扰源往往是瞬态信号，这时，参数$\xi(n,k)$，$\beta(n,k)$估计值与真实值存在较大误差，而由（7）式可知，SPP 对参数$\xi(n,k)$，$\beta(n,k)$的依赖关系是指数函数型的，对参数的变化非常敏感，$\xi(n,k)$，$\beta(n,k)$的微小计算误差，会导致 SPP 计算值的剧烈波动，进而影响语音增强系统的整体性能。

此外，单麦克风和多麦克风阵列的语音出现概率的理论公式（5）（6）（7）都是基于高斯统计模型推导得到的，它们存在一个缺陷，即当某个时频单元的先验信噪比$\xi(n,k)\to 0$时，$P(H_1|\mathrm{Y})\to\frac{1}{1+q}$。这与经验是相抵触的，当信噪比趋近于零时，语音是不存在的，即语音出现概率应当趋近于零。

另一方面，移动终端通话过程中经常遇到的瞬态噪声、第三方语音干扰等情况，这种噪声源和干扰源具有和语音相似或相同的时变特性，利用上述公式（7）计算语音出现概率，会将这种类型的噪声和干扰判定为语音，导致 SPP 的计算失效。

针对以上 SPP 估计方法的缺点，本公开实施例提出了一种计算复杂度小，对参数波动不敏感的 SPP 估计方法，使之满足如下条件：当$\xi(n,k)\to 0$时，$P(H_1|\mathrm{Y})\to 0$，应用于双麦克风阵列的语音出现概率计算，其中，双麦克风阵列包括采用端射（End-fire）结构设置的第一麦克风和第二麦克风，这里，假设第一麦克风与用户嘴部的距离小于第二麦克风与用户嘴部的距离，即第一麦

克风相比于第二麦克风，更接近于用户嘴部。

本公开实施例定义两个参数（后文中也称为第一度量参数和第二度量参数）: $M_{SNR}(n,k)$、$M_{PLD}(n,k)$（为简便起见，下文中也分别记为M_{SNR}和M_{PLD}）。M_{SNR}作为第一通道信号信噪比（SNR，Signal Noise Ratio）的度量参数，M_{PLD}作为第一、第二通道之间信号功率电平差（PLD，Power Level Difference）的度量参数，并用这两个参数计算 SPP。

具体的，请参照图 1 所示，本公开实施例提供的语音出现概率的确定方法，应用于采用 End-fire 结构设置的第一麦克风和第二麦克风，包括以下步骤:

步骤 11，根据第一麦克风拾取的第一通道的信号和第二麦克风拾取的第二通道的信号，计算第一度量参数和第二度量参数，所述第一度量参数为第一通道的信号信噪比，第二度量参数为第一通道与第二通道的信号功率电平差。

这里，将双通道信号间的功率电平差（第二度量参数）作为噪声干扰和目标语音之间的一个区分依据，结合信噪比度量参数（第一度量参数），计算双麦克风系统的语音出现概率，例如，在步骤 11 中提取跟 SNR、PLD 相关的两个参数M_{SNR}和M_{PLD}，用于后续 SPP 的计算。其中，M_{SNR}是利用信号的信噪比特征作为检测语音的判据，M_{PLD}是利用近场目标语音与远场噪声干扰的近远场特征不同，作为检测近场语音的判据。

步骤 12，对第一度量参数和第二度量参数，分别进行归一化和非线性变换处理，得到第三度量参数和第四度量参数。

这里，在步骤 12 中，可以通过分段线性变换，对M_{SNR}和M_{PLD}进行归一化和非线性变换处理，得到第三度量参数（可以记为M'_{SNR}）和第四度量参数（可以记为M'_{PLD}）。所述归一化和非线性变换处理具体包括:

对待处理参数进行数值更新，得到中间参数，其中，在数值超出区间[0，1]时，将数值更新为 1，否则保持数值不变，所述待处理参数为第一度量参数或第二度量参数;

对中间参数进行分段线性变换，得到最终参数，所述最终参数是所述中间参数的分段线性函数，且接近于所述中间参数取值范围中心的区段的斜率，大于远离所述中间参数取值范围中心的区段的斜率，所述最终参数为第三度量参数或第四度量参数。

步骤 13，根据第三度量参数、第四度量参数以及预先确定的语音出现概率的计算公式，计算得到语音出现概率，其中，所述计算公式是利用第三度量参数和第四度量参数幂级数的一次项和乘积项得到拟合公式，并对拟合系数施加归一化约束后得到的。

这里，所述语音出现概率的计算公式是利用归一化处理后的功率电平差度量参数（第四度量参数）和信噪比度量参数（第三度量参数）的二次函数，拟合出的语音出现概率。例如，可以利用M'_{SNR}、M'_{PLD} 的一次项和乘积项拟合 SPP 的计算公式。然后，在具体计算过程中，还可以利用功率电平差度量参数和信噪比度量参数的相关性强弱，自适应调整二次函数的各项的权重，即调整 SPP 计算公式的拟合系数，使计算结果更为准确。当然，所述拟合系数a、c的取值也可以是预先设定的固定值，例如根据当前应用场景中经常出现的噪声类型，预先设置好拟合参数的数值。

可以看出，本公开实施例提供的上述确定方法，具有较低的计算复杂度，并且对参数的波动具有更好的健壮性。另外，相关技术中的 SPP 计算方法大多是针对稳态和准稳态噪声，当受到瞬态噪声和第三方语音干扰时，其计算方法容易失效。本公开实施例提出的 SPP 计算方法，既能应用于稳态/准稳态噪声场中的情形，又能应用于瞬态噪声和第三方语音干扰的情形，能够广泛适用于各种双麦克风语音增强系统的应用场景。

为了更好的理解上述步骤，下面将进一步通过具体公式以及详细文字描述的方式，对本公开实施例作进一步的说明。

本公开实施例中，第一度量参数用于反映第一通道的信号信噪比，具体可以多种形式，可以直接采用第一通道的信号先验信噪比$\xi_1(n,k)$来表征，也可以采用第一通道的信号先验信噪比$\xi_1(n,k)$与一参考值的比值（如下述公式（12））来表征。第二度量参数用于反映两个通道之间的信号功率电平差，具体可以采用两个通道的信号功率电平的比值（如下述公式（13））来表征，也可以两个通道的功率谱密度矩阵的比值（如$\boldsymbol{\Phi}_{y_2y_2}/\boldsymbol{\Phi}_{y_1y_1}$）来表征，还可以采用两个通道的功率谱密度的差值与和值的比值来表征。

对双麦克风系统而言，目标语音表现为近场信号，环境噪声，第三方干扰等表现为远场信号。双麦克风系统第一通道和第二通道的信号功率电平差可以

作为区分近场信号和远场信号的一个重要判据，将近场目标语音检测出来。

不同于相关技术中多通道 SPP 估计方法，本公开实施例将双通道信号间的功率电平差作为噪声干扰和目标语音之间的一个区分依据，结合信噪比度量参数，计算双麦克风系统的 SPP。

在忽略两个麦克风信号间相位信息时，SPP 与变量M_{SNR}和M_{PLD}存在复杂的函数关系，可以用这两个变量的幂级数拟合。为了降低算法复杂度，本公开实施例首先对M_{SNR}和M_{PLD}做分段线性变换，再做幂级数展开，并取前几项，根据经验拟合其系数。可参考图 2 所示，首先提取M_{SNR}和M_{PLD}（步骤 21、23），然后对M_{SNR}和M_{PLD}归一化和分段线性变换处理得到M'_{SNR}、M'_{PLD}（步骤 22、24），然后，在利用计算公式加权计算 SPP 之前还可以自适应调整拟合系数（步骤 25），最后，利用 M'_{SNR}、M'_{PLD}的一次项及乘积项加权计算 SPP（步骤 26），得到 SPP 的计算结果（记为 p_1）。

下面介绍本公开实施例提取信噪比度量参数M_{SNR}和功率电平差度量参数M_{PLD}的一种实现方式。这里以下面公式（12）（13）作为第一、第二度量参数的表征方式来进行说明，其他表征方式的原理类似，为节约篇幅，不再一一赘述。

$$M_{SNR}(n,k)=\frac{\xi_1(n,k)}{\xi_0(k)} \qquad (12)$$

$$M_{PLD}(n,k)=\frac{\boldsymbol{\Phi}_{y_1y_1}-\boldsymbol{\Phi}_{y_2y_2}}{\boldsymbol{\Phi}_{y_1y_1}+\boldsymbol{\Phi}_{y_2y_2}} \qquad (13)$$

上述公式中，$M_{SNR}(n,k)$表示第一度量参数，$\xi_1(n,k)$表示第一通道的第 n 帧信号第 k 个频率分量上的先验信噪比，$\xi_0(k)$表示预先设定的第 k 个频率分量上的信噪比参考值。中，$M_{PLD}(n,k)$表示第二度量参数，$\Phi_{y_1y_1}$表示第一通道的第 n 帧信号第 k 个频率分量上的信号功率谱密度，$\Phi_{y_2y_2}$表示第二通道的第 n 帧信号第 k 个频率分量上的信号功率谱密度。

利用上述公式（12），提取 第一度量参数，即信噪比参数M_{SNR}。其中，$\xi_0(k)$可以根据频点分段预先设定。比如，本公开实施例将语音频率分为低频、中频、高频三个频带，每个频带预设一个信噪比参考值：

$$\xi_0(k)=\begin{cases}\xi_L & 0\le k<k_L\\ \xi_M & k_L\le k<k_H\\ \xi_H & k_H\le k<k_{FS}\end{cases} \quad (14)$$

其中，k_L是低频带和中频带的分界频点，k_H是中频带与高频带的分界频点，k_{FS}是频带上界对应的频点。ξ_L，ξ_M，ξ_H是这三个频带内参数值，可以根据经验确定，以下举例进行说明。

实例 1: 本公开实施例在应用于窄带语音信号时，$k_L \in [800,2000]$Hz，$k_H \in [1500,3000]$Hz，对应的ξ_L，ξ_M，ξ_H取值范围为（1,20）。

实例 2：本公开实施例应用于宽带语音信号，$k_L \in [800,3000]$Hz，$k_H \in [2500,6000]$Hz。对应的ξ_L，ξ_M，ξ_H取值范围为（1,20）。

然后，利用公式（14）计算各个频点的$M_{SNR}(n,k)$。

利用公式（13）即可以提取功率电平差度量参数M_{PLD}。

在提取得到对M_{SNR}和M_{PLD}后，可通过非线性变换处理得到M'_{SNR}、M'_{PLD}。下面将介绍本公开实施例非线性变换的一种处理方式，即归一化和分段线性变换。分段线性变换是指把非线性特性曲线分成若干个区段，在每个区段中用直线段近似地代替特性曲线，这种处理方式也称为分段线性化，可以减小后续的计算复杂度。

由上述公式(7)可知，当 $M_{SNR}\to -0$， $p_1\to 0$；当 $M_{SNR}\to +\infty$， $p_1\to 1$。本公开实施例利用归一化和分段线性函数处理 M_{SNR}得到M'_{SNR}，以拟合 SPP 对参数 M_{SNR}依赖的函数特征。如图 3 所示，M'_{SNR}的取值范围为[0,1]。

具体的，先将 M_{SNR}的取值范围公式 $M_{SNR}=\min(M_{SNR},1)$归一化到[0,1]区间，然后对 M_{SNR}做分段线性变换，下面公式（15）以划分为 3 个区段为例进行说明，当然本公开实施例可以划分成更多或更少的区段:

$$M'_{SNR}=\begin{cases}k_1*M_{SNR} & M_{SNR}<s_1\\ k_1*s_1+k_2*(M_{SNR}-s_1) & s_1\le M_{SNR}<s_2\\ k_1*s_1+k_2*(s_2-s_1)+k_3*(M_{SNR}-s_2) & M_{SNR}\ge s_2\end{cases} \quad (15)$$

可以看出，上述对第一度量参数 M_{SNR}进行归一化和非线性变换处理，得

到第三度量参数M'_{SNR}的步骤具体包括：根据第一度量参数的数值，对第一度量参数进行更新，其中在第一度量参数超出区间[0，1]时，将第一度量参数更新为 1，否则保持第一度量参数不变；然后，对更新后的第一度量参数，进行分段线性变换，转换为第三度量参数，所述第三度量参数是第一度量参数的分段线性函数。考虑到 SPP 对参数M_{SNR}依赖的函数特征，该分段线性函数的多个区段中，接近于第一度量参数取值范围中心的区段的斜率，大于远离第一度量参数取值范围中心的区段的斜率。例如，对于公式（15），k_2大于 1，而k_1、k_3均小于 1。而s_1、s_2、s_3的取值，则可以根据经验值设置。

对远场噪声与干扰，$M_{PLD}\to 0$，$p_1\to 0$；对于近场语音，$M_{PLD}\to 1$，$p_1\to 1$。本公开实施例利用图 4 所示分段线性函数归一化M_{PLD}，首先根据经验数据确定一个接近于 1 的参数 x_{max}，用公式$M_{PLD}=\min(M_{PLD}, x_{max})$将$M_{PLD}$的取值映射到区间$[0, x_{max}]$，然后利用公式（16）进行分段线性化，得到的$M'_{PLD}$的取值范围为[0,1]。下面公式（16）以划分为 3 个区段为例进行说明，当然本公开实施例可以划分成更多或更少的区段。

$$M'_{PLD}=\begin{cases} t_1 * M_{PLD} & M_{PLD}<x_1 \\ t_1 * x_1 + t_2 * (M_{PLD}-x_1) & x_1 \le M_{PLD} < x_2 \\ t_1 * x_1 + t_2 * (x_2-x_1) + t_3 * (M_{PLD}-x_2) & M_{PLD} \ge x_2 \end{cases} \quad (16)$$

可以看出，上述对第二度量参数M_{PLD}进行归一化和非线性变换处理，得到第四度量参数M'_{PLD}的步骤包括：根据第二度量参数的数值，对第二度量参数进行更新，其中在第二度量参数超出区间[0，1]时，将第二度量参数更新为 1，否则保持第二度量参数不变；对更新后的第二度量参数，进行分段线性变换，转换为第四度量参数，所述第四度量参数是第二度量参数的分段线性函数。考虑到 SPP 对参数M_{PLD}依赖的函数特征，接近于第二度量参数取值范围中心的区段的斜率，大于远离第二度量参数取值范围中心的区段的斜率。例如，对于公式（16），t_2大于 1，而t_1、t_3均小于 1。而x_1、x_2、x_3的取值，则可以根据经验值设置。

如前所述，用M'_{SNR}、M'_{PLD}的一次项及乘积项拟合得到 SPP，并对拟合系数施加归一化约束，可以得到如下所述的 SPP 的计算公式：

$$P_1 = \mathrm{c}(aM'_{SNR} + (1-a)M'_{PLD}) + (1-\mathrm{c})M'_{SNR}M'_{PLD} \qquad (17)$$

公式（17）中，存在a, c两个参数，a, c的取值范围都为[0,1]。本公开实施例根据M_{SNR} M_{PLD}的相关性，自适应调整c的大小，以及，根据麦克风的一致性特征，自适应调整a的大小。

理论上，M'_{SNR}、M'_{PLD}都可以独立作为 VAD 的判据或独立来计算 SPP。受各种因素影响，计算值与理论值有一定的偏离。特别的是，M'_{SNR}对平稳噪声，扩散场噪声有更好的适应性；M_{PLD}对远场的非平稳噪声，瞬态噪声及第三方讲话者的干扰语音有更好的适应性。

如图 5 所示，图 5 示出的是参数M'_{SNR}与M'_{PLD}的取值空间，M'_{SNR}、M'_{PLD}的取值空间可以分为示意性的四个区域，其中，图 5 中的 A1 区域，M'_{PLD}接近于 0，M'_{SNR}接近于 0；A2 区域M'_{PLD}接近于 1，且M'_{SNR}接近于 1；B1 区域，M'_{PLD}接近于 0，且M'_{SNR}接近于 1；B2 区域，M'_{PLD}接近于 1，且M'_{SNR}接近于 0。

在A_1, A_2区域，这两个参数具有较强的相关性，c取值较大，强调公式（17）的线性部分；在B_1，B_2区域，这两个参数相关性较弱，c取值较小，突出公式（17）的乘积项$M'_{SNR}M'_{PLD}$。本公开实施例可以根据M_{SNR} M_{PLD}分布的区域，自适应调整公式（17）中的参数c。具体的，拟合系数c的取值，随着M'_{SNR}与M'_{PLD}的差值的减小而增大。

下面利用两个例子来说明参数c的取值策略，需要指出的是，本公开实施例并不局限于这两种举例的实现方式。

实例 1：假设当前的参数M'_{SNR}与M'_{PLD}对应于图 5 中的参考点 R，即参考点 R 的坐标为（M'_{PLD}，M'_{SNR}）。假设第一线段与第二射线的夹角θ，可以用$\cos^2(\theta)$作为参数c的取值，如下述公式（18）所示 ，这里，第一线段以点（0.5,0.5）作为起点，R 为终点；第二射线以点（0.5,0.5）作为起点，且与M'_{PLD}轴呈 45 度夹角：

$$\mathrm{c} = \frac{(M'_{PLD} + M'_{SNR} - 1)^2}{(M'_{PLD} + M'_{SNR} - 1)^2 + (M'_{PLD} - M'_{SNR})^2} \qquad (18)$$

实例 2：可以根据下述公式（19）确定 c 的取值：

$$c = 1 - |M'_{PLD} - M'_{SNR}| \qquad (19)$$

本公开实施例中，参数a可以根据经验在$0 \leqslant a \leqslant 1$范围内取值，也可以根据噪声类型的预判而预先调整$a$的大小。例如，当预判噪声为稳态准稳态时，加大$M'_{SNR}$的权重，加大$a$的取值，当噪声为瞬态噪声或第三方语音干扰时，增加$M'_{PLD}$的权重，减小$a$的取值。例如，用户基于当前所处环境确定当前环境中可能的噪声类型，本公开实施例根据上述噪声类型来设置a的取值。

在确定了拟合系数a, c的取值之后，本公开实施例即可利用公式（17）确定语音出现概率。上述公式（17）大大减少了 SPP 计算的运算量，且语音出现概率不再是参数$\xi(n,k)$，$\beta(n,k)$的指数函数，使得计算结果对参数波动具有较好的健壮性。另外，相关技术中的 SPP 计算方法大多是针对稳态和准稳态噪声，当受到瞬态噪声和第三方语音干扰时，其计算方法容易失效。而本公开实施例提出的 SPP 计算方法，既能应用于稳态/准稳态噪声场中的情形，又能应用于瞬态噪声和第三方语音干扰的情形，能够广泛适用于各种双麦克风语音增强系统的应用场景。

基于以上所述的语音出现概率的确定方法，本公开实施例还提供了一种实现上述方法的确定装置和电子设备。请参照图 6 所示，本公开实施例提供的确定装置，应用于采用端射（End-fire）结构设置的第一麦克风和第二麦克风，该装置包括：

采集单元 61，用于分别采集第一麦克风对应的第一通道和第二麦克风对应的第二通道的声音信号，计算第一度量参数和第二度量参数，其中，所述第一度量参数为第一通道的信号信噪比，第二度量参数为第一通道与第二通道的信号功率电平差；

转换单元 62，用于对第一度量参数和第二度量参数，分别进行归一化和非线性变换处理，得到第三度量参数和第四度量参数；

计算单元 63，用于根据第三度量参数、第四度量参数以及预先确定的语

音出现概率的计算公式，计算得到语音出现概率，其中，所述计算公式是通过对第三度量参数和第四度量参数的幂级数的一次项和乘积项进行拟合，并对拟合系数施加归一化约束后得到的。

本公开实施例中所述采集单元 61，具体用于：

利用以下公式，计算第一度量参数：$M_{SNR}(n,k)=\frac{\xi_1(n,k)}{\xi_0(k)}$

其中，$M_{SNR}(n,k)$表示第一度量参数，$\xi_1(n,k)$表示第一通道的第 n 帧信号第 k 个频率分量上的先验信噪比，$\xi_0(k)$表示预先设定的第 k 个频率分量上的信噪比参考值。

所述采集单元 61，还可以用于：

利用以下公式，计算第二度量参数：$M_{PLD}(n,k)=\frac{\Phi_{y_1y_1}-\Phi_{y_2y_2}}{\Phi_{y_1y_1}+\Phi_{y_2y_2}}$

其中，$M_{PLD}(n,k)$表示第二度量参数，$\Phi_{y_1y_1}$表示第一通道的第 n 帧信号第 k 个频率分量上的信号功率谱密度，$\Phi_{y_2y_2}$表示第二通道的第 n 帧信号第 k 个频率分量上的信号功率谱密度。

本公开实施例中，所述转换单元 62，具体用于：对待处理参数进行数值更新，得到中间参数，其中，在数值超出区间[0，1]时，将数值更新为 1，否则保持数值不变，所述待处理参数为第一度量参数或第二度量参数；对中间参数进行分段线性变换，得到最终参数，所述最终参数是所述中间参数的分段线性函数，且接近于所述中间参数取值范围中心的区段的斜率，大于远离所述中间参数取值范围中心的区段的斜率，所述最终参数为第三度量参数或第四度量参数。

作为一种可选方式，本公开实施例中，所述语音出现概率的计算公式为：

$$P_1=\mathrm{c}(aM'_{SNR}+(1-a)M'_{PLD})+(1-\mathrm{c})M'_{SNR}M'_{PLD}$$

其中，P_1表示第 n 帧信号第 k 个频率分量上的语音出现概率，M'_{SNR}表示第三度量参数，M'_{PLD}表示第四度量参数，a、c均为取值范围在[0，1]之内的拟合系数。

作为一种可选方式，所述拟合系数a、c的取值是预先设定的固定值。

作为另一种可选方式，所述拟合系数a、c的取值是根据M'_{SNR}与M'_{PLD}确定的，其中，拟合系数a的取值是根据（M'_{PLD}，M'_{SNR}）所在区域确定的，不同区域对应于不同的值。

拟合系数c的取值，随着M'_{SNR}与M'_{PLD}的差值的减小而增大。

可选的，所述拟合系数c的取值，可以按照以下任一公式计算得到：

$$c = \frac{(M'_{PLD} + M'_{SNR} - 1)^2}{(M'_{PLD} + M'_{SNR} - 1)^2 + (M'_{PLD} - M'_{SNR})^2}$$

$$c = 1 - |M'_{PLD} - M'_{SNR}|$$

请参照图 7，本公开实施例提供的电子设备，包括：

处理器 71；以及，通过总线接口 72 与所述处理器相连接的存储器 73、第一麦克风 74 和第二麦克风 75，所述第一麦克风 74 和第二麦克风 75 采用端射 End-fire 结构配置，第一麦克风 74 通常与用户嘴部的距离小于第二麦克风 75 与用户嘴部的距离；所述存储器 73 用于存储所述处理器 71 在执行操作时所使用的程序和数据，当处理器 71 调用并执行所述存储器 73 中所存储的程序和数据时，实现如下的功能模块：

采集单元，用于分别采集第一麦克风对应的第一通道和第二麦克风对应的第二通道的声音信号，计算第一度量参数和第二度量参数，其中，所述第一度量参数为第一通道的信号信噪比，第二度量参数为第一通道与第二通道的信号功率电平差；

转换单元，用于对第一度量参数和第二度量参数，分别进行归一化和非线性变换处理，得到第三度量参数和第四度量参数；

计算单元，用于根据第三度量参数、第四度量参数以及预先确定的语音出现概率的计算公式，计算得到语音出现概率，其中，所述计算公式是通过对第三度量参数和第四度量参数的幂级数的一次项和乘积项进行拟合，并对拟合系数施加归一化约束后得到的。

以上所述是本公开的可选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本公开所述原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本公开的保护范围。

权 利 要 求 书

1. 一种语音出现概率的确定方法，应用于采用端射 End-fire 结构设置的第一麦克风和第二麦克风，包括:

根据第一麦克风拾取的第一通道的信号和第二麦克风拾取的第二通道的信号，计算第一度量参数和第二度量参数，所述第一度量参数为第一通道的信号信噪比，第二度量参数为第一通道与第二通道的信号功率电平差;

对第一度量参数和第二度量参数，分别进行归一化和非线性变换处理，得到第三度量参数和第四度量参数;

根据第三度量参数、第四度量参数以及预先确定的语音出现概率的计算公式，计算得到语音出现概率，其中，所述计算公式是通过对第三度量参数和第四度量参数的二元幂级数的一次项和乘积项进行拟合，并对拟合系数施加归一化约束后得到的。

2. 如权利要求 1 所述的确定方法，其中，

所述第一度量参数的计算包括:

利用以下公式，计算第一度量参数：$M_{SNR}(n,k)=\frac{\xi_1(n,k)}{\xi_0(k)}$

其中，$M_{SNR}(n,k)$表示第一度量参数，$\xi_1(n,k)$表示第一通道的第 n 帧信号第 k 个频率分量上的先验信噪比，$\xi_0(k)$表示预先设定的第 k 个频率分量上的信噪比参考值。

3. 如权利要求 2 所述的确定方法，其中，

所述第二度量参数的计算包括:

利用以下公式，计算第二度量参数：$M_{PLD}(n,k)=\frac{\Phi_{y_1y_1}-\Phi_{y_2y_2}}{\Phi_{y_1y_1}+\Phi_{y_2y_2}}$

其中，$M_{PLD}(n,k)$表示第二度量参数，$\Phi_{y_1y_1}$表示第一通道的第 n 帧信号第 k 个频率分量上的信号功率谱密度，$\Phi_{y_2y_2}$表示第二通道的第 n 帧信号第 k 个频率分量上的信号功率谱密度。

4. 如权利要求 3 所述的确定方法，其中，

所述归一化和非线性变换处理包括:

对待处理参数进行数值更新，得到中间参数，其中，在数值超出区间[0，1]时，将数值更新为 1，否则保持数值不变，所述待处理参数为第一度量参数或第二度量参数;

对中间参数进行分段线性变换，得到最终参数，所述最终参数是所述中间参数的分段线性函数，且接近于所述中间参数取值范围中心的区段的斜率，大于远离所述中间参数取值范围中心的区段的斜率，所述最终参数为第三度量参数或第四度量参数。

5. 如权利要求 4 所述的确定方法，其中，

所述语音出现概率的计算公式为:

$$P_1 = \mathrm{c}(aM'_{SNR} + (1-a)M'_{PLD}) + (1-\mathrm{c})M'_{SNR}M'_{PLD}$$

其中，P_1表示第 n 帧信号第 k 个频率分量上的语音出现概率，M'_{SNR}表示第三度量参数，M'_{PLD}表示第四度量参数，a、c均为取值范围在[0，1]之内的拟合系数。

6. 如权利要求 5 所述的确定方法，其中，所述拟合系数a、c的取值是预先设定的固定值。

7. 如权利要求 5 所述的确定方法，其中，

所述拟合系数a的取值是根据环境噪声的类型而预先设确定的;

所述拟合系数c的取值，随着M'_{SNR}与M'_{PLD}的差值的减小而增大。

8. 如权利要求 7 所述的确定方法，其中，

拟合系数c的取值，按照以下任一公式计算得到:

$$\mathrm{c} = \frac{(M'_{PLD}+M'_{SNR}-1)^2}{(M'_{PLD}+M'_{SNR}-1)^2+(M'_{PLD}-M'_{SNR})^2};$$

$$\mathrm{c} = 1 - |M'_{PLD} - M'_{SNR}|$$

9. 一种语音出现概率的确定装置，应用于采用端射 End-fire 结构设置的第一麦克风和第二麦克风，包括:

采集单元，用于根据第一麦克风拾取的第一通道的信号和第二麦克风拾取

的第二通道的信号，计算第一度量参数和第二度量参数，所述第一度量参数为第一通道的信号信噪比，第二度量参数为第一通道与第二通道的信号功率电平差；

转换单元，用于对第一度量参数和第二度量参数，分别进行归一化和非线性变换处理，得到第三度量参数和第四度量参数；

计算单元，用于根据第三度量参数、第四度量参数以及预先确定的语音出现概率的计算公式，计算得到语音出现概率，其中，所述计算公式是通过对第三度量参数和第四度量参数的二元幂级数的一次项和乘积项进行拟合，并对拟合系数施加归一化约束后得到的。

10. 如权利要求 9 所述的确定装置，其中，

所述采集单元，具体用于：

利用以下公式，计算第一度量参数：$M_{SNR}(n,k)=\frac{\xi_1(n,k)}{\xi_0(k)}$

其中，$M_{SNR}(n,k)$表示第一度量参数，$\xi_1(n,k)$表示第一通道的第 n 帧信号第 k 个频率分量上的先验信噪比，$\xi_0(k)$表示预先设定的第 k 个频率分量上的信噪比参考值。

11. 如权利要求 10 所述的确定装置，其中，

所述采集单元，具体用于：

利用以下公式，计算第二度量参数：$M_{PLD}(n,k)=\frac{\Phi_{y_1y_1}-\Phi_{y_2y_2}}{\Phi_{y_1y_1}+\Phi_{y_2y_2}}$

其中，$M_{PLD}(n,k)$表示第二度量参数，$\Phi_{y_1y_1}$表示第一通道的第 n 帧信号第 k 个频率分量上的信号功率谱密度，$\Phi_{y_2y_2}$表示第二通道的第 n 帧信号第 k 个频率分量上的信号功率谱密度。

12. 如权利要求 11 所述的确定装置，其中，

所述转换单元，具体用于：对待处理参数进行数值更新，得到中间参数，其中，在数值超出区间[0，1]时，将数值更新为 1，否则保持数值不变，所述待处理参数为第一度量参数或第二度量参数；对中间参数进行分段线性变换，

得到最终参数，所述最终参数是所述中间参数的分段线性函数，且接近于所述中间参数取值范围中心的区段的斜率，大于远离所述中间参数取值范围中心的区段的斜率，所述最终参数为第三度量参数或第四度量参数。

13. 如权利要求 12 所述的确定装置，其中，

所述语音出现概率的计算公式为：

$$P_1 = \mathrm{c}(aM'_{SNR} + (1-a)M'_{PLD}) + (1-\mathrm{c})M'_{SNR}M'_{PLD}$$

其中，P_1表示第 n 帧信号第 k 个频率分量上的语音出现概率，M'_{SNR}表示第三度量参数，M'_{PLD}表示第四度量参数，a、c均为取值范围在[0，1]之内的拟合系数。

14. 如权利要求 13 所述的确定装置，其中，所述拟合系数a、c的取值是预先设定的固定值。

15. 如权利要求 13 所述的确定装置，其中，

所述拟合系数a的取值是根据环境噪声的类型而与确预先设定的；

所述拟合系数c的取值，随着M'_{SNR}与M'_{PLD}的差值的减小而增大。

16. 如权利要求 15 所述的确定装置，其中，

拟合系数c的取值，按照以下任一公式计算得到：

$$\mathrm{c} = \frac{(M'_{PLD} + M'_{SNR} - 1)^2}{(M'_{PLD} + M'_{SNR} - 1)^2 + (M'_{PLD} - M'_{SNR})^2}$$

$$\mathrm{c} = 1 - |M'_{PLD} - M'_{SNR}|$$

17. 一种电子设备，包括：

处理器；以及，通过总线接口与所述处理器相连接的存储器、第一麦克风和第二麦克风，所述第一麦克风和第二麦克风采用端射 End-fire 结构配置；所述存储器用于存储所述处理器在执行操作时所使用的程序和数据，当处理器调用并执行所述存储器中所存储的程序和数据时，实现如下的功能模块：

采集单元，用于分别采集第一麦克风对应的第一通道和第二麦克风对应的第二通道的声音信号，计算第一度量参数和第二度量参数，其中，所述第一度量参数为第一通道的信号信噪比，第二度量参数为第一通道与第二通道的信号功率电平差；

转换单元，用于对第一度量参数和第二度量参数，分别进行归一化和非线性变换处理，得到第三度量参数和第四度量参数；

计算单元，用于根据第三度量参数、第四度量参数以及预先确定的语音出现概率的计算公式，计算得到语音出现概率，其中，所述计算公式是通过对第三度量参数和第四度量参数的二元幂级数的一次项和乘积项进行拟合，并对拟合系数施加归一化约束后得到的。

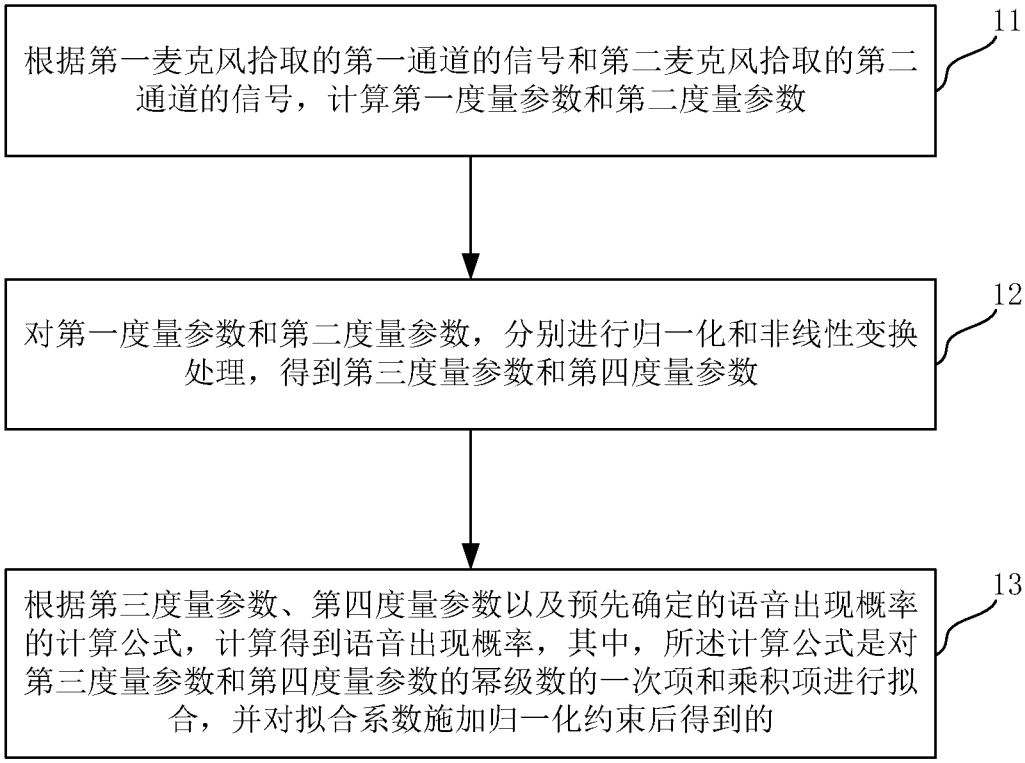

根据第一麦克风拾取的第一通道的信号和第二麦克风拾取的第二通道的信号，计算第一度量参数和第二度量参数
11
对第一度量参数和第二度量参数，分别进行归一化和非线性变换处理，得到第三度量参数和第四度量参数
12
根据第三度量参数、第四度量参数以及预先确定的语音出现概率的计算公式，计算得到语音出现概率，其中，所述计算公式是对第三度量参数和第四度量参数的幂级数的一次项和乘积项进行拟合，并对拟合系数施加归一化约束后得到的
13

图 1

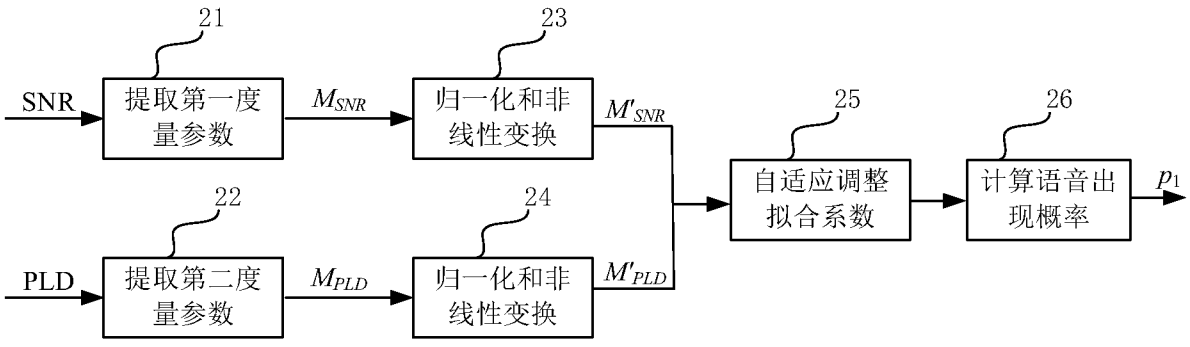

21
SNR
提取第一度量参数
M_{SNR}
23
归一化和非线性变换
M'_{SNR}
25
自适应调整拟合系数
26
计算语音出现概率
p_1
22
PLD
提取第二度量参数
M_{PLD}
24
归一化和非线性变换
M'_{PLD}

图 2

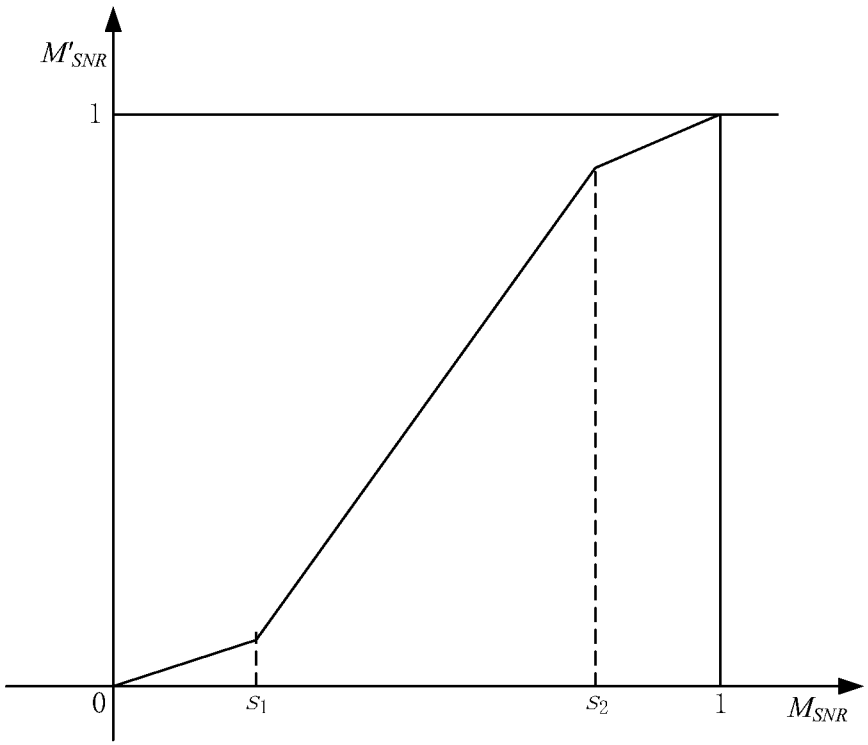


图 3

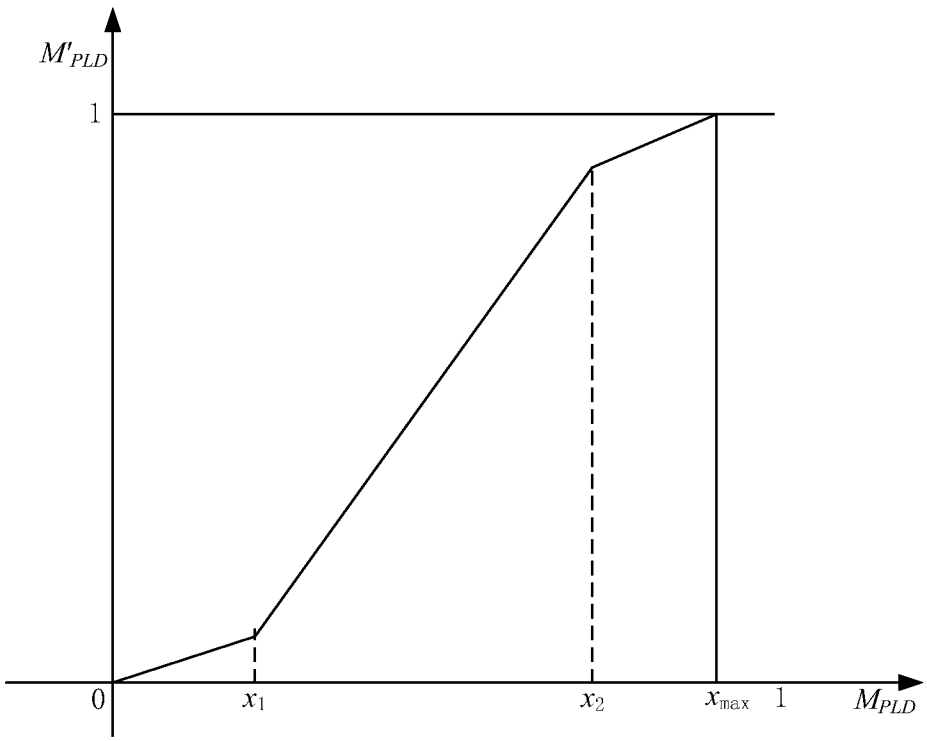


图 4

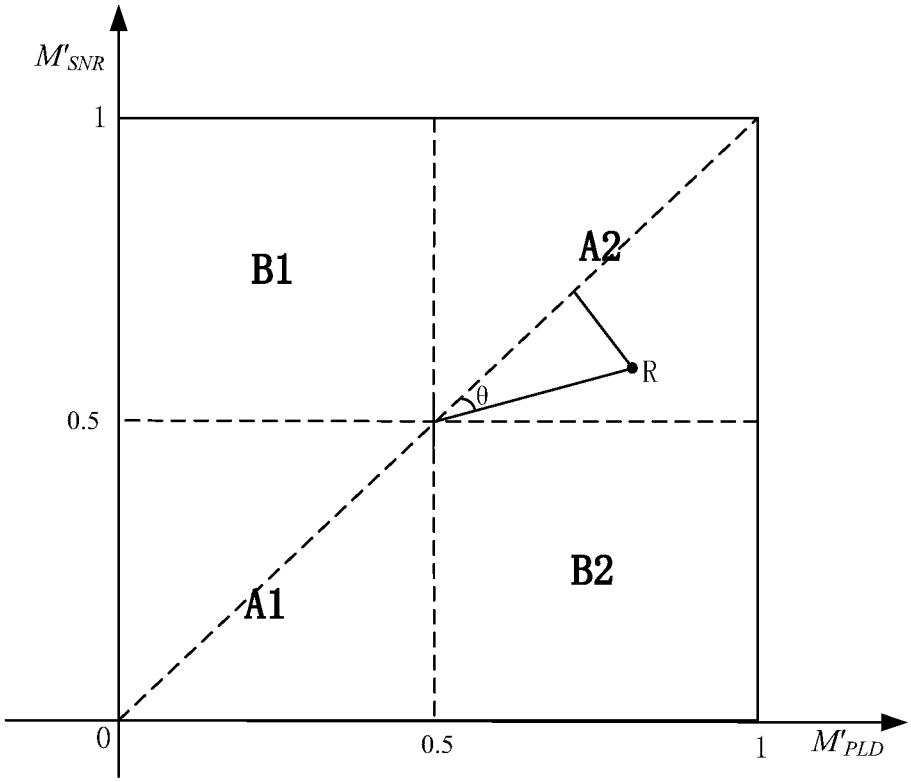


图 5

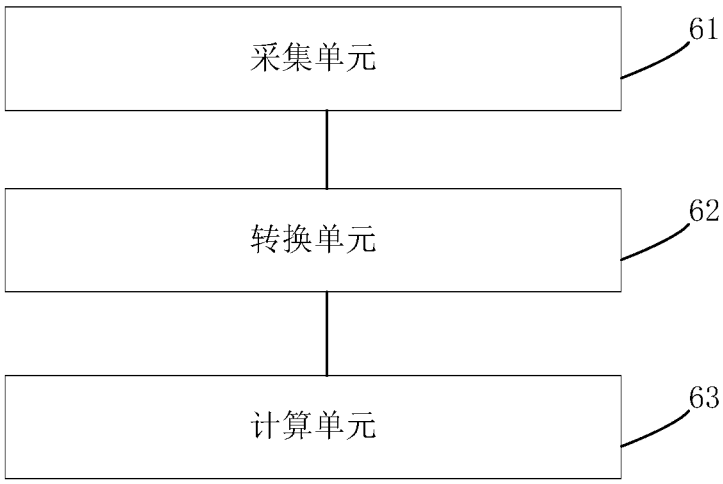


图 6

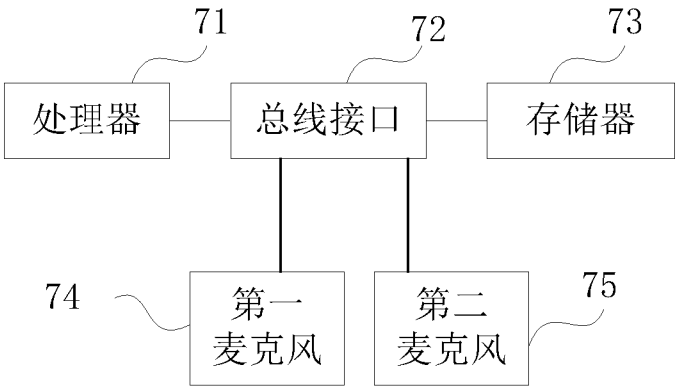


图 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/112323

A. CLASSIFICATION OF SUBJECT MATTER

G10L 21/0216 (2013.01) i; G10L 21/0224 (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)

VEN, CNABS, IEEE, CNKI, CPC: activation; CHINA ACADEMY OF TELECOMMUNICATIONS TECHNOLOGY; VAD, voice, activity, detection, SPP, speech, absence, probability, presence, SAP

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2015221322 A1 (APPLE INC.), 06 August 2015 (06.08.2015), description, paragraphs [0005]-[0057], and figures 1-7	1-17
A	CN 103646648 A (TSINGHUA UNIVERSITY et al.), 19 March 2014 (19.03.2014), description, paragraphs [0005]-[0085], and figures 1-3	1-17
A	CN 101790752 A (QUALCOMM INC.), 28 July 2010 (28.07.2010), the whole document	1-17
A	US 2012121100 A1 (ZHANG, X. et al.), 17 May 2012 (17.05.2012), the whole document	1-17
A	CN 101510426 A (BEIJING VIMICRO CORP.), 19 August 2009 (19.08.2009), the whole document	1-17

☐ 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
22 March 2017 (22.03.2017)	**29 March 2017 (29.03.2017)**
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 **ZHANG, Yaling** Telephone No.: (86-10) **62085722**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/112323

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US 2015221322 A1	06 August 2015	US 9524735 B2	20 December 2016
CN 103646648 A	19 March 2014	CN 103646648 B	23 March 2016
CN 101790752 A	28 July 2010	JP 2010541010 A	24 December 2010
		US 2009089053 A1	02 April 2009
		WO 2009042948 A1	02 April 2009
		TW 200926151 A	16 June 2009
		AT 531030 T	15 November 2011
		US 8954324 B2	10 February 2015
		CA 2695231 C	17 February 2015
		TW I398855 B	11 June 2013
		KR 101265111 B1	16 May 2013
		EP 2201563 A1	30 June 2010
		CA 2695231 A1	02 April 2009
		RU 2010116727 A	10 November 2011
		CN 101790752 B	04 September 2013
		RU 2450368 C2	10 May 2012
		JP 5102365 B2	19 December 2012
		EP 2201563 B1	26 October 2011
		KR 20100075976 A	05 July 2010
		ES 2373511 T3	06 February 2012
US 2012121100 A1	17 May 2012	US 2012123771 A1	17 May 2012
		US 8977545 B2	10 March 2015
		US 2012123772 A1	17 May 2012
		US 2012123773 A1	17 May 2012
		US 8965757 B2	24 February 2015
		US 8924204 B2	30 December 2014
		US 9330675 B2	03 May 2016
CN 101510426 A	19 August 2009	CN 101510426 B	27 March 2013
		US 2010241426 A1	23 September 2010
		US 9286908 B2	15 March 2016
		US 2014067386 A1	06 March 2014
		US 8612217 B2	17 December 2013

国际检索报告

国际申请号
PCT/CN2016/112323

A. 主题的分类

G10L 21/0216(2013.01)i；G10L 21/0224(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））

VEN, CNABS, IEEE, CNKI, CPC: 语音，出现，检测，概率，激活，电信科学研究院，VAD, voice, activity, detection, SPP, speech, absence, probability, presence, SAP

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
A	US 2015221322 A1 (APPLE INC.) 2015年 8月 6日 (2015-08-06) 说明书第[0005]-[0057]，图1-7	1-17
A	CN 103646648 A (清华大学 等) 2014年 3月 19日 (2014-03-19) 说明书第[0005]-[0085]段，图1-3	1-17
A	CN 101790752 A (高通股份有限公司) 2010年 7月 28日 (2010-07-28) 全文	1-17
A	US 2012121100 A1 (ZHANG XIANXIAN等) 2012年 5月 17日 (2012-05-17) 全文	1-17
A	CN 101510426 A (北京中星微电子有限公司) 2009年 8月 19日 (2009-08-19) 全文	1-17

☐ 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
2017年 3月 22日	2017年 3月 29日

ISA/CN的名称和邮寄地址	受权官员
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	张亚玲
传真号 (86-10)62019451	电话号码 (86-10)62085722

表 PCT/ISA/210 (第2页) (2009年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2016/112323

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)
US	2015221322	A1	2015年 8月 6日	US	9524735	B2	2016年 12月 20日
CN	103646648	A	2014年 3月 19日	CN	103646648	B	2016年 3月 23日
CN	101790752	A	2010年 7月 28日	JP	2010541010	A	2010年 12月 24日
				US	2009089053	A1	2009年 4月 2日
				WO	2009042948	A1	2009年 4月 2日
				TW	200926151	A	2009年 6月 16日
				AT	531030	T	2011年 11月 15日
				US	8954324	B2	2015年 2月 10日
				CA	2695231	C	2015年 2月 17日
				TW	I398855	B	2013年 6月 11日
				KR	101265111	B1	2013年 5月 16日
				EP	2201563	A1	2010年 6月 30日
				CA	2695231	A1	2009年 4月 2日
				RU	2010116727	A	2011年 11月 10日
				CN	101790752	B	2013年 9月 4日
				RU	2450368	C2	2012年 5月 10日
				JP	5102365	B2	2012年 12月 19日
				EP	2201563	B1	2011年 10月 26日
				KR	20100075976	A	2010年 7月 5日
				ES	2373511	T3	2012年 2月 6日
US	2012121100	A1	2012年 5月 17日	US	2012123771	A1	2012年 5月 17日
				US	8977545	B2	2015年 3月 10日
				US	2012123772	A1	2012年 5月 17日
				US	2012123773	A1	2012年 5月 17日
				US	8965757	B2	2015年 2月 24日
				US	8924204	B2	2014年 12月 30日
				US	9330675	B2	2016年 5月 3日
CN	101510426	A	2009年 8月 19日	CN	101510426	B	2013年 3月 27日
				US	2010241426	A1	2010年 9月 23日
				US	9286908	B2	2016年 3月 15日
				US	2014067386	A1	2014年 3月 6日
				US	8612217	B2	2013年 12月 17日