

(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
2020 年 6 月 4 日 (04.06.2020)



(10) 国际公布号
WO 2020/107269 A1

(51) 国际专利分类号:
G10L 21/0316 (2013.01)

(21) 国际申请号: PCT/CN2018/117972

(22) 国际申请日: 2018 年 11 月 28 日 (28.11.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 深圳市汇顶科技股份有限公司(SHENZHEN GOODIX TECHNOLOGY CO.,LTD.)
[CN/CN]; 中国广东省深圳市福田保税区腾飞工业大厦B座13层, Guangdong 518045 (CN)。

(72) 发明人: 朱虎(ZHU, Hu); 中国广东省深圳市福田保税区腾飞工业大厦B座13层, Guangdong 518045 (CN)。 王鑫山(WANG, Xinshan); 中国广东省深圳市福田保税区腾飞工业大厦B

座13层, Guangdong 518045 (CN)。 李国梁(LI, Guoliang); 中国广东省深圳市福田保税区腾飞工业大厦B座13层, Guangdong 518045 (CN)。 曾端(ZENG, Duan); 中国广东省深圳市福田保税区腾飞工业大厦B座13层, Guangdong 518045 (CN)。 郭红敬(GUO, Hongjing); 中国广东省深圳市福田保税区腾飞工业大厦B座13层, Guangdong 518045 (CN)。

(74) 代理人: 上海晨皓知识产权代理事务所(普通合伙)(SHANGHAI CHENHAO INTELLECTUAL PROPERTY LAW FIRM GENERAL PARTNERSHIP); 中国上海市黄浦区制造局路787号二幢202B室, Shanghai 200011 (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,

(54) **Title:** SELF-ADAPTIVE SPEECH ENHANCEMENT METHOD, AND ELECTRONIC DEVICE

(54) 发明名称: 自适应语音增强方法和电子设备

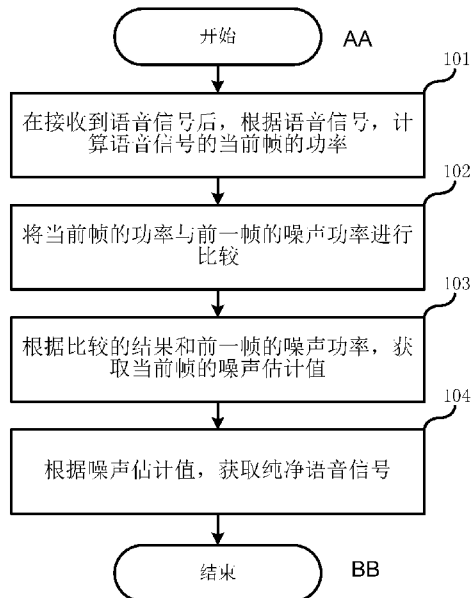


图 1

- 101 After receiving a speech signal, calculate, according to the speech signal, the power of a current frame of the speech signal
102 Compare the power of the current frame with the noise power of a previous frame
103 Acquire, according to a comparison result and the noise power of the previous frame, a noise estimation value of the current frame
104 Acquire, according to the noise estimation value, a pure speech signal
AA Start
BB End

(57) **Abstract:** A self-adaptive speech enhancement method and an electronic device. The self-adaptive speech enhancement method comprises: after receiving a speech signal, calculating, according to the speech signal, the power of a current frame of the speech signal (101); comparing the power of the current frame with the noise power of a previous frame (102); acquiring, according to a comparison result and the noise power of the previous frame, a noise estimation value of the current frame (103); and acquiring, according to the noise estimation value, a pure speech signal (104). The use of this method makes the estimation of noise more accurate and reduces

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, JO, 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)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

the complexity of an algorithm, thereby facilitating the enhancement of a speech signal and improving the quality of human auditory perception.

(57) 摘要: 一种自适应语音增强方法和电子设备。自适应语音增强方法, 包括: 在接收到语音信号后, 根据语音信号, 计算语音信号的当前帧的功率(101); 将当前帧的功率与前一帧的噪声功率进行比较(102); 根据比较的结果和前一帧的噪声功率, 获取当前帧的噪声估计值(103); 根据噪声估计值, 获取纯净语音信号(104)。采用该方法, 使得对于噪声的估计更加准确, 且降低了算法的复杂度, 从而有利于对语音信号进行增强, 改善人的听觉感知质量。

自适应语音增强方法和电子设备

技术领域

[0001] 本申请涉及信息处理技术领域，特别涉及一种自适应语音增强方法和电子设备。

背景技术

[0002] 现实生活中，由于说话人经常要处于各种嘈杂的环境中，语音信号不可避免的会受到背景噪声的污染，背景噪声使许多语音处理系统的性能急剧下降。语音增强作为一种信号处理方法，是解决噪声污染的一种高效途径。一方面，通过进行语音增强，可以提高噪声环境下语音的清晰度，可懂度和舒适度，改善人的听觉感知质量；另一方面，语音增强也是语音处理系统中必不可少的环节，在进行各种语音信号处理操作前首先要进行语音增强，以减小噪声对语音处理系统的影响，提高系统的工作技能。

[0003] 语音增强最主要包括噪声估计和滤波器系数求解两部分。代表性的语音增强方法包括谱减法、维纳滤波法、最小均方误差估计法、子空间法、基于小波变换的增强方法等。这些方法大都基于频率中语音和噪声分量的统计模型，并结合各种估计理论来设计具有针对性的噪声消除技术。

[0004] 现有技术中的语音增强算法中，存在噪声估计不准确，算法复杂的问题。

发明内容

[0005] 本申请部分实施例的目的在于提供一种自适应语音增强方法，使得对于噪声的估计更加准确，且降低了算法的复杂度，从而有利于对语音信号进行增强，改善人的听觉感知质量。

[0006] 本申请实施例提供了一种自适应语音增强方法，包括：在接收到语音信号后，根据语音信号，计算语音信号的当前帧的功率；将当前帧的功率与前一帧的噪声功率进行比较；根据比较的结果和前一帧的噪声功率，获取当前帧的噪声估计值；根据噪声估计值，获取纯净语音信号。

[0007] 本申请实施例还提供了一种电子设备，包括：至少一个处理器；以及，与所述至少一个处理器通信连接的存储器；其中，所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行上述的自适应语音增强方法。

[0008] 本申请实施例相对于现有技术而言，根据接收的语音信号，计算语音信号的当前帧的功率，将当前帧的功率与前一帧的噪声功率进行比较、根据比较的结果和前一帧的噪声功率，获取当前帧的噪声估计值。对于噪声的估计无需利用 VAD 算法以检测区分当前帧是语音帧还是噪声帧，从而可以避免由于 VAD 算法检测的不准确将导致噪声估计的偏差很大，有利于快速估计语音信号中的噪声成分。本申请采用了迭代的估计方法，每一帧的噪声功率都会进行自适应更新，将当前帧的功率与前一帧的噪声功率进行比较，估计当前帧的噪声值，在不断的迭代过程中，也使得估计的噪声值越来越精准。而且，本申请中，对每一帧都重新计算其功率，能够实现噪声的连续估计和不断更新，只需要将当前帧的功率与前一帧的噪声功率对比，不需要存储前 D 帧数据和按照功率大

小进行排序，从而减小了算法资源开销，降低了算法的复杂度。根据噪声估计值，获取纯净语音信号，有利于对语音信号进行增强，改善人的听觉感知质量。

[0009] 例如，当前帧的功率具体为：当前帧的对数功率谱；前一帧的噪声功率具体为：前一帧的对数分位数。对数坐标能够放大细节，可以提取出一般坐标刻度下提取不出的信号，有利于压缩数值的动态范围，使得在对数坐标系下，当前帧的对数功率谱和前一帧的对数分位数之间的比较更加精确，从而有利于后续的精准处理。

[0010] 例如，根据比较结果和前一帧的噪声功率，获取当前帧的噪声估计值，具体包括：根据当前帧的对数功率谱与前一帧的对数分位数的比较结果，获取当前帧的增量步长；根据前一帧的对数分位数和当前帧的增量步长获取当前帧的对数分位数；根据所述当前帧的对数分位数，获取当前帧的噪声估计值。当前帧的增量步长为获取当前的帧的对数分位数提供有针对意义的参考，有利于准确的获取当前帧的对数分位数，从而对前帧的噪声值进行准确的估计。

[0011] 例如，根据前一帧的对数分位数和当前帧的增量步长获取当前帧的对数分位数，具体包括：若当前帧的对数功率谱大于或等于前一帧的对数分位数，则根据增量步长将前一帧的对数分位数自适应增加，得到当前帧的对数分位数；若当前帧的对数功率谱小于前一帧的对数分位数，则根据增量步长将前一帧的对数分位数自适应减小，得到当前帧的对数分位数。通过将前一帧的对数分位数根据增量步长进行自适应的增加或减小，有利于准确的获取当前帧的对数分位数。

[0012] 例如，根据当前帧的对数功率谱与前一帧的对数分位数的比较结果，获取当前帧的增量步长，具体包括：根据当前帧的对数功率谱与前一帧的对数分

位数的比较结果，获取密度函数；根据密度函数获取当前帧的增量步长，提供了一种获取当前帧的增量步长的方式。

[0013] 例如，获取密度函数，具体用过以下公式获取密度函数 *density*:

$$density(\lambda, k) = \begin{cases} \frac{(\beta - 1) \cdot density(\lambda - 1, k) + 1 / (2\xi)}{\beta} & |\log(|Y_w(\lambda)|^2) - lq(\lambda - 1, k)| \geq \xi \\ density(\lambda - 1, k) & others \end{cases}$$

其中，所述 λ 为当前帧的帧号，所述 k 为频点数，所述 β 为实验的经验值，所述 ξ 为预设的门限阈值，所述 $\log(|Y_w(\lambda)|^2)$ 为当前帧的对数功率谱，所述 $lq(\lambda - 1, k)$ 为前一帧的对数分位数；提供了一种获取密度函数的具体计算公式，有利于快速准确地获取密度函数。

[0014] 例如，根据密度函数获取当前帧的增量步长，具体通过以下公式获取增量步长 *delta*：

$$delta(\lambda, k) = \begin{cases} \frac{K}{density(\lambda - 1, k)} & density(\lambda - 1, k) \geq 1 \\ K & density(\lambda - 1, k) < 1 \end{cases}$$

其中， λ 为当前帧的帧号， K 为增量步长控制因子， $density(\lambda - 1, k)$ 为前一帧的密度函数，提供了一种获取增量步长的具体计算公式，有利于快速准确地获取增量步长。

[0015] 例如，根据增量步长将前一帧的对数分位数自适应增加，得到当前帧的对数分位数，具体通过以下公式得到当前帧的对数分位数：

$lq(\lambda, k) = lq(\lambda - 1, k) + \alpha \cdot delta(\lambda, k) / \beta$ ；根据增量步长将前一帧的对数分位数自适应减小，得到当前帧的对数分位数，具体包括：具体通过以下公式得到当前帧的对数分位数： $lq(\lambda, k) = lq(\lambda - 1, k) - (1 - \alpha) \cdot delta(\lambda, k) / \beta$ ；其中， λ 为当前帧的帧号， k 为频点数， α 为实验的经验值， $delta(\lambda, k)$ 为增量步长。提供了自适应增

加和减小对数分位数的计算公式，有利于直接、快速并准确的获取当前帧的对数分位数。

[0016] 例如，根据噪声估计值，获取纯净语音信号，具体包括：获取语音信号的当前帧的功率谱；根据噪声估计值，获取谱增益系数；根据谱增益系数获取当前帧的纯净语音信号，有利于自适应的跟踪每一帧噪声的变化，对每一帧均进行语音增强，提高噪声环境下语音的清晰度，可懂度和舒适度，减小噪声对语音处理系统的影响，提高系统的工作技能。

[0017] 例如，根据噪声估计值，获取谱增益系数，具体包括：根据前一帧的噪声估计值和前一帧的纯净语音信号，计算先验信噪比；根据当前帧的噪声估计值和当前帧的功率，计算后验信噪比；根据先验信噪比和后验信噪比获取谱增益系数，提供了一种获取谱增益系数的方式。

[0018] 例如，根据先验信噪比和后验信噪比获取谱增益系数，具体包括：根据以下公式获取谱增益系数：

$$G = \frac{\sqrt{\gamma_k}}{\gamma_k} \left[\frac{\Gamma\left(\frac{p+\beta}{2}+1\right) \Phi\left(-\frac{p+\beta}{2}, 1; -\nu_k\right)}{\Gamma\left(\frac{p}{2}+1\right) \Phi\left(-\frac{p}{2}, 1; -\nu_k\right)} \right]^{1/\beta}$$

其中， γ_k 为后验信噪比， ξ_k 为先验信噪比， $\nu_k = \frac{\xi_k}{1+\xi_k} \gamma_k$ ， p 为感知加权阶数，

β 为高阶幅度谱的阶数。提供了一种获取谱增益系数的具体计算公式，有利于准确快速的获取谱增益系数。

[0019] 例如，计算若干个子带的信噪比，具体包括：通过以下公式计算所述若干个子带的信噪比：

$$SNR(b, k) = 10 \log 10 \frac{\sum_{k=B_{low}(b)}^{B_{up}(b)} |Y_k(b, k) - \sqrt{\lambda_d(b, k)}|^2}{\sum_{k=B_{low}(b)}^{B_{up}(b)} \lambda_d(b, k)}$$

其中, b 为子带的序号, k 为频点数, $B_{low}(b)$ 为 Bark 域第 b 个子带的频点起点, $B_{up}(b)$ 为 Bark 域第 b 子带的频点终点, 考虑到了人耳对语音在 Bark 域更加的敏感和人耳掩蔽机制, 有利于改善人的听觉感知质量。

[0020] 例如, 根据若干个子带的信噪比计算感知加权阶数, 具体为: 通过以下公式计算所述感知加权阶数 p :

$$p(b, k) = \max\{\min[\alpha_1 SNR(b, k) + \alpha_2, p_{\max}], p_{\min}\}$$

其中, α_1 , α_2 , $p_{\min}$ 和 $p_{\max}$ 均为实验经验值。提供了一种获取感知加权阶数的具体计算公式, 有利于准确快速的获取感知加权阶数。

[0021] 例如, $\Gamma\left(\frac{p+\beta}{2}+1\right)$ 和 $\Gamma\left(\frac{p}{2}+1\right)$ 具体通过以下方式得到: 根据预存的 Γ 函数的输入输出对应关系查询 $\Gamma\left(\frac{p+\beta}{2}+1\right)$ 和 $\Gamma\left(\frac{p}{2}+1\right)$; $\Phi\left(-\frac{p+\beta}{2}, 1; -\nu_k\right)$ 和 $\Phi\left(-\frac{p}{2}, 1; -\nu_k\right)$ 具体通过以下方式得到: 根据预存的 Φ 函数的输入输出对应关系查询 $\Phi\left(-\frac{p+\beta}{2}, 1; -\nu_k\right)$ 和 $\Phi\left(-\frac{p}{2}, 1; -\nu_k\right)$ 。采用根据对应关系查询的方式, 使得该方法的计算复杂度大大降低, 减小了运算量, 更加具有工程适用性。

[0022] 例如, 根据谱增益系数获取纯净语音信号, 具体通过以下公式获取:

$$\hat{X}_k = \frac{\sqrt{\nu_k}}{\gamma_k} \left[\frac{\Gamma\left(\frac{p+\beta}{2}+1\right) \Phi\left(-\frac{p+\beta}{2}, 1; -\nu_k\right)}{\Gamma\left(\frac{p}{2}+1\right) \Phi\left(-\frac{p}{2}, 1; -\nu_k\right)} \right]^{1/\beta} Y_w(k)$$

其中, 所述 $Y_w(k)$ 为当前帧的信号幅值, 提供了获取纯净语音信号的具体公式,

有利于快速准确的得到当前帧的纯净语音信号。

附图说明

[0023] 一个或多个实施例通过与之对应的附图中的图片进行示例性说明，这些示例性说明并不构成对实施例的限定，附图中具有相同参考数字标号的元件表示为类似的元件，除非有特别申明，附图中的图不构成比例限制。

[0024] 图 1 是根据本申请第一实施例中的自适应语音增强方法的流程图；

[0025] 图 2 是根据本申请第一实施例中的 Kaiser 窗函数示意图；

[0026] 图 3 是根据本申请第一实施例中的步骤 104 的子步骤的示意图；

[0027] 图 4 是根据本申请第二实施例中的自适应语音增强方法的流程图；

[0028] 图 5 是根据本申请第二实施例中的实现自适应语音增强方法的模块示意图；

[0029] 图 6 是根据本申请第三实施例中的自适应语音增强方法的流程图；

[0030] 图 7 是根据本申请第四实施例中的电子设备的结构示意图。

具体实施方式

[0031] 为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请部分实施例进行进一步详细说明。应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。以下各个实施例的划分是为了描述方便，不应对本发明的具体实现方式构成任何限定，各个实施例在不矛盾的前提下可以相互结合相互引用。

[0032] 本申请第一实施例涉及一种自适应语音增强方法，包括：在接收到语音信号后，根据语音信号，计算语音信号的当前帧的功率；将当前帧的功率与自适应更新的噪声功率进行比较；其中，自适应更新的噪声功率为所述语音信号的前一帧的噪声功率；根据比较的结果，获取当前帧的噪声估计值；根据噪声估计值，获取纯净语音信号，使得对于噪声的估计更加准确，且降低了算法的复杂度，从而有利于对语音信号进行增强，改善人的听觉感知质量。下面对本实施方式的自适应语音增强方法的实现细节进行具体的说明，以下内容仅为方便理解提供的实现细节，并非实施本方案的必须。

[0033] 本实施例的自适应语音增强方法能够运用于语音信号处理技术领域并适用于低功耗语音增强、语音识别、语音交互产品中，包括但不限于耳机、音响、手机、电视、汽车、可穿戴设备、智能家居等电子设备。

[0034] 本实施例中自适应语音增强方法的具体流程如图 1 所示，包括：

[0035] 步骤 101：在接收到语音信号后，根据语音信号，计算语音信号的当前帧的功率。

[0036] 具体的说，在接收到语音信号后可以对语音信号进行时域和频域的变换，得到频域语音，频域语音是描述语音信号在频率方面特性时用到的一种坐标系。语音信号从时间域变换到频率域主要通过傅立叶级数和傅立叶变换实现。周期信号靠傅立叶级数，非周期信号靠傅立叶变换。通常，语音信号的时域越宽，其频域越短。根据频域语音信号的当前帧的幅值得到当前帧的功率。

[0037] 在一个例子中，假设对语音信号的采样速率为 $F_s=8000\text{Hz}$ ，数据长度一般处理数据为 $8\text{ms}\sim 30\text{ms}$ 之间，对语音信号进行处理可以为 64 点并且重叠上一帧 64 个点，则实际一次处理 128 个点，即当前帧与上一帧的重叠率为 50%，但

在实际应用中，并不以此为限。对接收的语音信号进行预加重处理，提高语音信号的高频成分，具体的运算可以为： $\hat{y}(n) = y(n) - \alpha y(n-1)$ ，其中， α 为平滑因子，本实施例中 α 可取值为0.98，但在实际应用中可根据实际需求进行不同的设置。 $y(n)$ 为采样的当前帧的语音信号， $y(n-1)$ 为采样得到的上一帧的语音信号。

[0038] 进一步的，在预加重处理后，为了减少频谱能量泄漏，可采用截取函数对信号进行截断，截断函数称为窗函数，即对语音信号进行加窗处理，窗函数设计可以根据不同运用场景可选择矩形窗、汉明窗、汉宁窗和高斯窗函数等，实际设计中可灵活选择，本实施例采用如图2所示的Kaiser窗函数，重叠50%。

[0039] 另外，由于计算语音信号的当前帧的功率通常在频域进行处理，所以可以将加窗数据通过以下公式，进行快速傅里叶变换FFT，得到频域信号。

$$y_w(n, m) = w(n) \hat{y}(n, m)$$

$$Y_k(m) = \sum_{n=0}^{N-1} y_w(n, m) e^{-j2\pi \frac{k}{N} n}$$

其中， k 代表频点数， $w(n)$ 为Kaiser窗函数， N 为128，即实际一次处理128个点，本实施例只是以 $N=128$ 为例，但在实际应用中并不以此为限。 m 为帧数， n 的取值可以为1到128。对于当前帧的功率的计算，可以对变换后的128个频点的频域信号求幅值，并分别对128个频点的幅值求平方得到。

[0040] 步骤102：将当前帧的功率与前一帧的噪声功率进行比较。

[0041] 具体的说，前一帧的噪声功率为自适应更新的噪声功率。在实际应用中，首先可以根据实验值初始化噪声功率，如果当前帧为第一帧，那么可以将当前帧的功率与初始化的噪声功率进行比较。自适应更新的噪声功率是指不同帧的

噪声功率是不同的，设置噪声功率的初始值后，在迭代的过程中可自适应的更新当前帧的噪声功率。比如说，分别将当前帧的 128 个频点的功率和前一帧的 128 个频点的功率进行比较，自适应的更新当前帧中每一个频点对应的噪声功率。

[0042] 步骤 103：根据比较的结果和前一帧的噪声功率，获取当前帧的噪声估计值。

[0043] 具体的说，如果当前帧的功率大于前一帧的噪声功率，则可以自适应的增大前一帧的噪声功率，以此作为当前帧的噪声估计值，比如说，可以预设增量步长，根据增量步长进行自适应的增加。较佳的，增量步长也可以在迭代的过程中自适应的更新。如果当前帧的功率小于前一帧的噪声功率，则可以自适应的减小前一帧的噪声功率，将减小后的噪声功率作为当前帧的噪声估计值。

[0044] 步骤 104：根据噪声估计值，获取纯净语音信号。

[0045] 具体的说，步骤 104 可以包括如图 3 所示的如下子步骤：

[0046] 步骤 1041：根据前一帧的噪声估计值和前一帧的纯净语音信号，计算先验信噪比。

[0047] 具体的说，先验信噪比的计算可以采用经典的改进判决引导方法，可以根据以下公式计算先验信噪比 $\hat{\xi}_k(\lambda)$ ：

$$\hat{\xi}_k(\lambda) = \max \left[a \frac{\hat{X}_k^2(\lambda-1)}{\lambda_d(\lambda-1)} + (1-a) \max[\gamma_k(\lambda)-1, 0], \xi_{\min} \right]$$

其中， a 为平滑因子， $\xi_{\min}$ 为预设的经验值， $\hat{X}_k^2(\lambda-1)$ 为所述前一帧的纯净语音信号功率， λ 为当前帧的帧号。在一个例子中， a 的取值可以为 0.98， $\xi_{\min}$ 可根据经验取值为-15dB，但在实际应用中并不以此为限。

[0048] 需要说明的是，本实施方式以通过上述公式计算先验信噪比为例，但在实际应用中并不以此为限。

[0049] 步骤 1042：根据当前帧的噪声估计值和当前帧的功率，计算后验信噪比。

[0050] 具体的说，可以根据以下公式计算后验信噪比：

$$\gamma_k = \frac{Y_w^2(k)}{\lambda_d(k)}$$

其中， $Y_w^2(k)$ 为当前帧的功率， $\lambda_d(k)$ 为当前帧的噪声估计值。

[0051] 需要说明的是，本实施方式以通过上述公式计算后验信噪比为例，但在实际应用中并不以此为限。而且本实施例中也并不限定步骤 1041 与步骤 1042 的执行顺序，在实际应用中，也可以先执行步骤 1042 再执行步骤 1041，或者同时执行步骤 1041 和步骤 1042。

[0052] 步骤 1043：计算感知加权阶数 p 。

[0053] 具体的说，可以根据子带信噪比和 Bark 域的特性自适应的计算参数 p 。

具体的，在语音信号的频率谱中，可以将 Bark 域划分为若干个子带，比如说：

可以将 Bark 域分为 18 个子带，每个子带的上限频率为：

[100, 200, 300, 400, 510, 630, 770, 920, 1080, 1270, 1480, 1720, 2000, 2320, 2700,

3150, 3700, 4400]，根据人耳对语音在 Bark 域更加的敏感，计算子带的信噪比，

通过以下公式计算若干个子带的信噪比：

$$SNR(b, k) = 10 \log_{10} \frac{\sum_{k=B_{low}(b)}^{B_{up}(b)} \left| Y_k(b, k) - \sqrt{\lambda_d(b, k)} \right|^2}{\sum_{k=B_{low}(b)}^{B_{up}(b)} \lambda_d(b, k)}$$

其中， b 为子带的序号，子带的序号 $1 \leq b \leq 18$ ， k 为频点数， $B_{low}(b)$ 为 Bark 域

第 b 子带的频点起点, $B_{up}(b)$ 为 Bark 域第 b 子带的频点终点。进一步的, 可以通过以下公式计算参数 p :

$$p(b, k) = \max \{ \min [\alpha_1 SNR(b, k) + \alpha_2, p_{\max}], p_{\min} \}$$

其中, α_1 , α_2 , $p_{\min}$ 和 $p_{\max}$ 均为实验经验值, 在本实施例中, 比如说实验经验值的取值可以如下所示: $\alpha_1 = 0.251$, $\alpha_2 = -1.542$, $p_{\max} = 4$, $p_{\min} = -1$, 但在实际应用中并不以此为限。

[0054] 步骤 1044: 计算高阶幅度谱的阶数 β 。

[0055] 具体的说, 通过以下公式计算高阶幅度谱的阶数 β :

$$\beta(b, k) = \beta_{\max} - \frac{(\beta_{\max} - \beta_{\min}) \sum_{k=B_{low}(b)}^{B_{up}(b)} \frac{20}{3} \log_{10} \left(\frac{f(k)}{A} + 1 \right)}{(B_{up}(b) - B_{low}(b)) \log_{10} \left(\frac{F_s}{2A} + 1 \right)}$$

其中, F_s 为采样频率, $f(k) = kF_s / N$, 代表 FFT 以后每个频点代表的频率数, $\beta_{\max}$ 和 A 为实验经验值。比如说, 在本实施例中, 上述经验值的取值可以分别如下: $\beta_{\max} = 0.8$, $\beta_{\min} = 0.2$, $A = 165.4 \text{ Hz}$, 但在实际应用中并不以此为限。

[0056] 需要说明的是, 本实施方式并不限定步骤 1043 和步骤 1044 的执行顺序, 在实际应用中, 也可以先执行步骤 1044 再执行步骤 1043, 或者同时执行步骤 1043 和步骤 1044。

[0057] 步骤 1045: 根据先验信噪比、后验信噪比、感知加权阶数和高阶幅度谱的阶数获取谱增益系数。

[0058] 具体的说, 得到谱增益系数的核心思想可以为贝叶斯短时幅度谱估计, 其代价函数为:

$$d(X_k, \hat{X}_k) = X_k^p \left(X_k^\beta - \hat{X}_k^\beta \right)^2$$

类似经典的 MMSE 估计器的推导过程, 可以得到:

$$\hat{X}_k^\beta = \frac{\int_0^\infty X_k^{p+\beta} p(X_k | Y_w(k)) dX_k}{\int_0^\infty X_k^p p(X_k | Y_w(k)) dX_k}$$

假设 X_k 和 D_k 都是复高斯随机分布, 可以得到:

$$\hat{X}_k = \frac{\sqrt{\nu_k}}{\gamma_k} \left[\frac{\Gamma\left(\frac{p+\beta}{2}+1\right) \Phi\left(-\frac{p+\beta}{2}, 1; -\nu_k\right)}{\Gamma\left(\frac{p}{2}+1\right) \Phi\left(-\frac{p}{2}, 1; -\nu_k\right)} \right]^{1/\beta} Y_w(k)$$

$$\gamma_k = \frac{Y_w^2(k)}{\lambda_d(k)}, \xi_k = \frac{\lambda_x(k)}{\lambda_d(k)}, \nu_k = \frac{\xi_k}{1+\xi_k} \gamma_k$$

其中, $\xi_k = \frac{\lambda_x(k)}{\lambda_d(k)}$ 为先验信噪比的理论公式, 由于实际中很难得到当前帧的纯净

语音功率 $\lambda_x(k)$, 所以通常可以采用以下公式来估计和近似求解先验信噪比 ξ_k :

$$\hat{\xi}_k(\lambda) = \max \left[a \frac{\hat{X}_k^2(\lambda-1)}{\lambda_d(\lambda-1)} + (1-a) \max[\gamma_k(\lambda)-1, 0], \xi_{\min} \right].$$

由以上推导公式可以得到谱增益系数 G 的计算公式如下:

$$G = \frac{\sqrt{\nu_k}}{\gamma_k} \left[\frac{\Gamma\left(\frac{p+\beta}{2}+1\right) \Phi\left(-\frac{p+\beta}{2}, 1; -\nu_k\right)}{\Gamma\left(\frac{p}{2}+1\right) \Phi\left(-\frac{p}{2}, 1; -\nu_k\right)} \right]^{1/\beta}$$

由上述谱增益系数 G 的表达式可知, 根据先验信噪比 ξ_k , 后验信噪比 γ_k , 参数 β 和 p 可以计算谱增益系数。

[0059] 进一步的, 考虑到 Γ 函数和 Φ 函数的复杂性, 可以采用查表的形式计算谱增益系数, 具体的可以预存的 Γ 函数和 Φ 函数的输入输出对应关系, 比如说:

根据预存的 Γ 函数的输入输出对应关系表进行查询, 当输入为 $\frac{p+\beta}{2}+1$ 时, 对应

的输出值 $\Gamma\left(\frac{p+\beta}{2}+1\right)$ ；当输入为 $\frac{p}{2}+1$ 时，对应的输出值 $\Gamma\left(\frac{p}{2}+1\right)$ 。在预存的 Φ 函数的输入输出对应关系表中进行查询：当输入为 $-\frac{p+\beta}{2}, 1; -\nu_k$ 时，对应的输出值 $\Phi\left(-\frac{p+\beta}{2}, 1; -\nu_k\right)$ ；当输入为 $-\frac{p}{2}, 1; -\nu_k$ 对应的输出值 $\Phi\left(-\frac{p}{2}, 1; -\nu_k\right)$ 。最后将查找到的输出值带入谱增益系数的计算表达式中可以得到谱增益系数，这样使得该方法的计算复杂度大大降低。

[0060] 需要说明的是，本实施例中只是以通过谱增益系数 G 的表达式得到谱增益系数为例，但在实际应用中并不以此为限。

[0061] 步骤 1046：根据谱增益系数获取当前帧的纯净语音信号。

[0062] 具体的说，在得到谱增益系数后，可以根据以下公式计算得到当前帧的纯净语音信号 $\hat{X}$ ：

$$\hat{X}_k = \frac{\sqrt{\nu_k}}{\gamma_k} \left[\frac{\Gamma\left(\frac{p+\beta}{2}+1\right) \Phi\left(-\frac{p+\beta}{2}, 1; -\nu_k\right)}{\Gamma\left(\frac{p}{2}+1\right) \Phi\left(-\frac{p}{2}, 1; -\nu_k\right)} \right]^{1/\beta} Y_w(k)$$

其中， $Y_w(k)$ 为当前帧的信号幅值。

[0063] 需要说明的是，本实施例中只是以通过纯净语音信号 $\hat{X}$ 的上述计算公式得到的到纯净语音信号为例，在实际应用中任何通过谱增益系数获取当前帧的纯净语音信号的方法均在本实施例保护范围之内。

[0064] 本实施例相对于现有技术而言，具有以下技术效果：第一、与传统噪声估计相比，不需要进行语音有声无声检测，在噪声帧和语音帧同时更新噪声，能够自适应的跟踪噪声的变化。第二、与传统的分位数噪声估计相比，不需要存储前 D 帧数据和按照功率大小进行排序，降低了算法资源开销。第三、在计算谱增益系数时同时考虑人耳掩蔽机制和对噪声与频谱振幅的敏感程度，自适

应的更新参数 p 和 β ，与传统的广义加权高阶谱估计量的语音增强相比，减小了运算量，更加具有工程适用性。

[0065] 本申请第二实施例涉及一种自适应语音增强方法，本实施例中的当前帧的功率，具体为：当前帧的对数功率谱；本实施例中的噪声功率具体为对数分位数。在对数坐标系下，当前帧的对数功率谱和前一帧的对数分位数之间的比较更加精确，从而有利于后续的精处理。

[0066] 本实施例中自适应语音增强方法的具体流程如图 4 所示，包括：

[0067] 步骤 201：在接收到语音信号后，根据语音信号，计算语音信号的当前帧的对数功率谱。

[0068] 具体的说，步骤 201 与步骤 101 大致相同，不同之处在于，步骤 101 计算的为当前帧的功率，而本步骤中计算的为当前帧的对数功率谱，即还需对计算得到的当前帧的功率取对数。比如说，对当前帧的语音信号进行处理可为 64 个点并且重叠上一帧 64 个点，则实际一次处理 128 个点，即可以得到 128 个点的功率值，对 128 个点的功率值分别取对数可以得到 128 个频点对应的对数功率，128 个对数功率组成当前帧的对数功率谱。

[0069] 步骤 202：根据当前帧的对数功率谱与前一帧的对数分位数的比较结果，获取密度函数。

[0070] 具体的说，本实施例中，可以预先设置初始对数分位数和初始密度函数。

即首先可以根据实验值初始化密度函数和对数分位数，比如说，根据实验值初始化后的对数分位数可以为： $lq(1,k)=8$ 。如果当前帧为第一帧，则可以将第一帧的对数功率谱与初始对数分位数进行比较。在之后的处理中可以根据当前帧的对数功率谱和前一帧的对数分位数，对当前帧的密度函数进行更新，具体可

以根据以下公式进行更新：

$$density(\lambda, k) = \begin{cases} \frac{(\beta - 1) \cdot density(\lambda - 1, k) + 1 / (2\xi)}{\beta} & |\log(|Y_w(\lambda)|^2) - lq(\lambda - 1, k)| > \xi \\ density(\lambda - 1, k) & others \end{cases}$$

其中， λ 为当前帧的帧号， k 为频点数， β 为实验的经验值， ξ 为预设的门限阈值， $\log(|Y_w(\lambda)|^2)$ 为当前帧的对数功率谱， $lq(\lambda - 1, k)$ 为前一帧的对数分位数。

[0071] 需要说明的是，本实施例只是以通过上述密度函数的计算公式得到当前帧的密度函数为例，但在实际应用中并不以此为限。

[0072] 步骤 203：根据密度函数获取当前帧的增量步长。

[0073] 具体的说，可以预先设置初始增量步长。比如说，根据实验值初始化后得到的初始增量步长可以为： $\delta(1, k) = 40$ 。在之后的处理中，根据前一帧的密度函数对当前帧的增量步长进行更新，具体可以根据以下公式进行更新：

$$\delta(\lambda, k) = \begin{cases} \frac{K}{density(\lambda - 1, k)} & density(\lambda - 1, k) \geq 1 \\ K & density(\lambda - 1, k) < 1 \end{cases}$$

其中， K 为增量步长控制因子。如果当前帧为第一帧，则增量步长控制因子 K 即为初始增量步长。

[0074] 需要说明的是，本实施例只是以通过上述增量步长的计算公式得到当前帧的增量步长为例，在实际应用中任何根据密度函数获取当前帧的增量步长的方法均在本实施例保护范围之内。

[0075] 步骤 204：根据前一帧的对数分位数和当前帧的增量步长获取当前帧的对数分位数。

[0076] 具体的说，若当前帧的对数功率谱大于或等于前一帧的对数分位数，则可以根据增量步长将前一帧的对数分位数自适应增加，得到当前帧的对数分位

数；若当前帧的对数功率谱小于前一帧的对数分位数，则可以根据增量步长将前一帧的对数分位数自适应减小，得到当前帧的对数分位数。

[0077] 步骤 205：根据当前帧的对数分位数，获取当前帧的噪声估计值。

[0078] 具体的说，在得到当前帧的对数分位数 $lq(\lambda, k)$ 后，可以通过以下公式计算噪声估计值：

$$\hat{\lambda}_d(\lambda, k) = e^{lq(\lambda, k)}$$

[0079] 步骤 206：根据噪声估计值，获取纯净语音信号。

[0080] 步骤 206 与第一实施例中步骤 104 大致相同，为避免重复在此不再赘述。

[0081] 为方便说明，本实施例提供如图 5 所示的框图对本实施例中的自适应语音增强方法进行说明：

[0082] 预加重模块 301，主要实现高通滤波器的功能，滤除低频成分，增强高频语音成分，即对接收的带噪语音信号 $y(n) = x(n) + d(n)$ ，滤除低频成分，其中， $x(n)$ 为纯净语音信号， $d(n)$ 为噪声信号。去预加重模块 310，主要是低通滤波器，去预加重模块 310 和预加重模块 301 为互逆过程，两者结合可以达到去混响的效果。

[0083] 加窗模块 302，主要是避免重叠信号出现突变的情况。窗口合成模块 309 主要是去除窗函数对输出纯净语音信号的影响。本实施例中加窗模块 302 和窗口合成模块 309 在实现的过程中，采用同一窗函数，因此，窗函数必须为保幂映射，即语音信号重叠部分的窗口平方和必须为 1，如下公式所示：

$$w^2(N) + w^2(N + M) = 1$$

其中，N 是 FFT 处理的点数，取值为 128，M 是帧长度取值为 64。

[0084] 快速傅里叶变换 FFT 模块 303，主要是进行时域信号与频域信号之间的

互相转换。FFT 模块 303 与逆 FFT 模块 308 互为逆过程，FFT 模块 303 将时域信号转换为频域信号，转换为频域信号后可以根据频域信号得到信号幅值 Y_w 。逆 FFT 模块 308 将频域信号转换为时域信号。

[0085] 功率谱计算模块 304，用于通过对频域信号求得的幅值进行平方运算得到当前帧的功率 P 。对数功率谱计算模块 305，用于对当前帧的功率取对数得到当前帧的对数功率谱。功率谱计算模块 304 和对数计算模块 305 主要是噪声估计之前的预处理过程。

[0086] 噪声值估计模块 306，主要是对带噪语音信号进行噪声估计，尽可能估计准确的噪声信号，主要根据自适应分位数噪声估计原理进行噪声估计得到噪声估计值 $\hat{\lambda}_d$ 。

[0087] 计算谱增益系数模块 307，主要是完成根据噪声估计值和带噪语音信号的功率进行谱增益系数的计算得到谱增益系数 G 。具体的，谱增益系数的计算主要根据广义加权高阶短时频谱振幅估计器原理。

[0088] 进一步的，根据谱增益系数 G 和信号幅值 Y_w 得到频域纯净语音信号 $\hat{X}$ ，接着再经过逆 FFT 模块 308 将频域信号转换为时域信号，通过窗口合成模块 309 和去预加重模块 310 的处理，输出时域纯净语音信号 $\hat{x}(n)$ ，从而完成对语音信号的增强。

[0089] 本实施例相对于现有技术而言，将带噪语音的当前帧的对数功率谱与前一帧的对数分位数比较来修改对数分位数以得到噪声估计值，这种做法可以避免现有技术中语音信号有无检测、大量数据存储和功率谱排序操作，减小算法资源开销。而且对数坐标能够放大细节，可以提取出一般坐标刻度下提取不出来的信号，有利于压缩数值的动态范围，使得在对数坐标系下，当前帧的对数功

率谱和前一帧的对数分位数之间的比较更加精确,从而有利于后续的精准处理。

[0090] 本申请第三实施例涉及一种自适应语音增强方法,在本实施例中,提供了一种具体公式,用以根据增量步长将前一帧的对数分位数自适应增加,得到当前帧的对数分位数,有利于直接、快速并准确的获取当前帧的对数分位数。

[0091] 本实施例中自适应语音增强方法的具体流程如图 6 所示,包括:

[0092] 步骤 401: 在接收到语音信号后,根据语音信号,计算语音信号的当前帧的对数功率谱。

[0093] 步骤 402: 根据当前帧的对数功率谱与前一帧的对数分位数的比较结果,获取密度函数。

[0094] 步骤 403: 根据密度函数获取当前帧的增量步长。

[0095] 步骤 401 至步骤 403 与第二实施例中步骤 201 至步骤 203 大致相同,为避免重复在此不再一一赘述。

[0096] 步骤 404: 判断当前帧的对数功率谱是否大于或等于前一帧的对数分位数,如果是则执行步骤 405,否则执行步骤 406。

[0097] 步骤 405: 根据公式 $lq(\lambda, k) = lq(\lambda - 1, k) + \alpha \cdot \text{delta}(\lambda, k) / \beta$ 计算当前帧的对数分位数。

[0098] 也就是说,当 $\log(|Y_w(\lambda)|^2) \geq lq(\lambda - 1, k)$ 时,当前帧的对数分位数为:将前一帧的对数分位数根据增量步长自适应增加,具体通过以下公式自适应增加,计算得到当前帧的对数分位数: $lq(\lambda, k) = lq(\lambda - 1, k) + \alpha \cdot \text{delta}(\lambda, k) / \beta$ 。其中, λ 为当前帧数, k 为频点数, α 和 β 都是实验的经验值,本实施例中实验的经验值可以为: $\alpha = 0.25$, $\beta = 67$,但在实际应用中并不以此为限。

[0099] 步骤 406: 根据公式 $lq(\lambda, k) = lq(\lambda - 1, k) - (1 - \alpha) \cdot \text{delta}(\lambda, k) / \beta$ 计算当前帧的

对数分位数。

[00100] 也就是说，当 $\log(|Y_w(\lambda)|^2) < lq(\lambda-1, k)$ 时，当前帧的对数分位数为：将前一帧的对数分位数根据增量步长自适应减小，具体通过以下公式自适应减小，计算得到当前帧的对数分位数： $lq(\lambda, k) = lq(\lambda-1, k) - (1-\alpha) \cdot \text{delta}(\lambda, k) / \beta$ 。

[00101] 步骤 407：根据公式 $\hat{\lambda}_d(\lambda, k) = e^{lq(\lambda, k)}$ ，获取当前帧的噪声估计值。

[00102] 步骤 408：根据噪声估计值，获取纯净语音信号。

[00103] 步骤 407 至步骤 408 与第二实施例中步骤 205 至步骤 206 大致相同，为避免重复此处不再赘述。

[00104] 本实施例相对于现有技术而言，提供了根据增量步长将前一帧的对数分位数自适应增加，得到当前帧的对数分位数的具体公式，有利于根据当前帧的增量步长直接、快速并准确的获取当前帧的对数分位数，从而有利于根据当前帧的对数分位数进行噪声估计。

[00105] 本申请第四实施例涉及一种电子设备，如图 7 所示，包括至少一个处理器 501；以及，与至少一个处理器 501 通信连接的存储器 502；其中，存储器 502 存储有可被所述至少一个处理器 501 执行的指令，指令被至少一个处理器 501 执行，以使至少一个处理器 501 能够执行上述的自适应语音增强方法。

[00106] 其中，存储器 502 和处理器 501 采用总线方式连接，总线可以包括任意数量的互联的总线和桥，总线将一个或多个处理器 501 和存储器 502 的各种电路连接在一起。总线还可以将诸如外围设备、稳压器和功率管理电路等之类的各种其他电路连接在一起，这些都是本领域所公知的，因此，本文不再对其进行进一步描述。总线接口在总线和收发机之间提供接口。收发机可以是一个元件，也可以是多个元件，比如多个接收器和发送器，提供用于在传输介质上

与各种其他装置通信的单元。经处理器 501 处理的数据通过天线在无线介质上进行传输，进一步，天线还接收数据并将数据传送给处理器 501。

[00107] 处理器 501 负责管理总线和通常的处理，还可以提供各种功能，包括定时，外围接口，电压调节、电源管理以及其他控制功能。而存储器 502 可以被用于存储处理器 501 在执行操作时所使用的数据。

[00108] 本领域技术人员可以理解，实现上述实施例方法中的全部或部分步骤是可以通程序来指令相关的硬件来完成，该程序存储在一个存储介质中，包括若干指令用以使得一个设备（可以是单片机，芯片等）或处理器（processor）执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（ROM, Read-Only Memory）、随机存取存储器（RAM, Random Access Memory）、磁碟或者光盘等各种可以存储程序代码的介质。

[00109] 本领域的普通技术人员可以理解，上述各实施例是实现本申请的具体实施例，而在实际应用中，可以在形式上和细节上对其作各种改变，而不偏离本申请的精神和范围。

权利要求书

1. 一种自适应语音增强方法，其特征在于，包括：

在接收到语音信号后，根据所述语音信号，计算所述语音信号的当前帧的

5 功率；

将所述当前帧的功率与前一帧的噪声功率进行比较；

根据比较的结果和所述前一帧的噪声功率，获取所述当前帧的噪声估计值；

根据所述噪声估计值，获取纯净语音信号。

2. 根据权利要求 1 所述的自适应语音增强方法，其特征在于，

10 所述当前帧的功率为：当前帧的对数功率谱；

所述前一帧的噪声功率为：前一帧的对数分位数。

3. 根据权利要求 2 所述的自适应语音增强方法，其特征在于，所述根据比较结果和所述前一帧的噪声功率，获取所述当前帧的噪声估计值，包括：

根据所述当前帧的对数功率谱与所述前一帧的对数分位数的比较结果，获

15 取所述当前帧的增量步长；

根据所述前一帧的对数分位数和所述当前帧的增量步长获取所述当前帧的对数分位数；

根据所述当前帧的对数分位数，获取所述当前帧的噪声估计值。

4. 根据权利要求 3 所述的自适应语音增强方法，其特征在于，所述根据所述前一帧的对数分位数和所述当前帧的增量步长获取所述当前帧的对数分位数，
20 包括：

若所述当前帧的对数功率谱大于或等于前一帧的对数分位数，则根据所述

增量步长将所述前一帧的对数分位数自适应增加, 得到当前帧的对数分位数;

若所述当前帧的对数功率谱小于前一帧的对数分位数, 则根据所述增量步长将所述前一帧的对数分位数自适应减小, 得到当前帧的对数分位数。

5. 根据权利要求 3 所述的自适应语音增强方法, 其特征在于, 还包括:

5 预先设置初始对数分位数和初始增量步长。

6. 根据权利要求 3 所述的自适应语音增强方法, 其特征在于, 所述根据所述当前帧的对数功率谱与所述前一帧的对数分位数的比较结果, 获取所述当前帧的增量步长, 包括:

根据所述当前帧的对数功率谱与所述前一帧的对数分位数的比较结果, 获

10 取密度函数;

根据所述密度函数获取所述当前帧的增量步长。

7. 根据权利要求 6 所述的自适应语音增强方法, 其特征在于, 所述获取概率密度, 包括:

通过以下公式获取密度函数 *density*:

$$15 \quad density(\lambda, k) = \begin{cases} \frac{(\beta - 1) \cdot density(\lambda - 1, k) + 1 / (2\xi)}{\beta} & |\log(|Y_w(\lambda)|^2) - lq(\lambda - 1, k)| > \xi \\ density(\lambda - 1, k) & others \end{cases}$$

其中, 所述 λ 为当前帧的帧号, 所述 k 为频点数, 所述 β 为实验值, 所述 ξ 为预设的门限阈值, 所述 $\log(|Y_w(\lambda)|^2)$ 为当前帧的对数功率谱, 所述 $lq(\lambda - 1, k)$ 为前一帧的对数分位数。

8. 根据权利要求 6 所述的自适应语音增强方法, 其特征在于, 所述根据所述密度函数获取所述当前帧的增量步长, 包括:

通过以下公式获取增量步长 *delta*:

$$\text{delta}(\lambda, k) = \begin{cases} \frac{K}{\text{density}(\lambda-1, k)} & \text{density}(\lambda-1, k) \geq 1 \\ K & \text{density}(\lambda-1, k) < 1 \end{cases}$$

其中，所述 λ 为当前帧的帧号，所述 K 为增量步长控制因子，所述 $\text{density}(\lambda-1, k)$ 为所述前一帧的密度函数。

9. 根据权利要求4所述的自适应语音增强方法，其特征在于，所述根据所述增量步长将所述前一帧的对数分位数自适应增加，得到当前帧的对数分位数，包括：

通过以下公式得到所述当前帧的对数分位数：

$$lq(\lambda, k) = lq(\lambda-1, k) + \alpha \cdot \text{delta}(\lambda, k) / \beta$$

所述根据所述增量步长将所述前一帧的对数分位数自适应减小，得到当前帧的对数分位数，包括：

通过以下公式得到所述当前帧的对数分位数：

$$lq(\lambda, k) = lq(\lambda-1, k) - (1 - \alpha) \cdot \text{delta}(\lambda, k) / \beta$$

其中，所述 λ 为当前帧的帧号，所述 k 为频点数，所述 α 为实验的经验值，所述 $\text{delta}(\lambda, k)$ 为增量步长。

10. 根据权利要求3所述的自适应语音增强方法，其特征在于，所述根据所述当前帧的对数分位数，获取所述当前帧的噪声估计值，包括：

通过以下公式获取所述当前帧的噪声估计值：

$$\hat{\lambda}_d(\lambda, k) = e^{lq(\lambda, k)}$$

其中，所述 $\hat{\lambda}_d$ 为噪声估计值，所述 $lq(\lambda, k)$ 为所述当前帧的对数分位数，所述 λ 为所述当前帧的帧号，所述 k 为频点数。

11. 根据权利要求1所述的自适应语音增强方法，其特征在于，所述根据

所述噪声估计值, 获取纯净语音信号, 包括:

根据所述噪声估计值, 获取谱增益系数;

根据所述谱增益系数获取所述当前帧的纯净语音信号。

12. 根据权利要求 11 所述的自适应语音增强方法, 其特征在于, 所述根据

5 所述噪声估计值, 获取谱增益系数, 包括:

根据所述前一帧的噪声估计值和所述前一帧的纯净语音信号, 计算先验信噪比;

根据所述当前帧的噪声估计值和所述当前帧的功率, 计算后验信噪比;

根据所述先验信噪比和所述后验信噪比获取谱增益系数。

10 13. 根据权利要求 12 所述的自适应语音增强方法, 其特征在于, 所述根据所述先验信噪比和所述后验信噪比获取谱增益系数, 包括:

根据以下公式获取谱增益系数 G :

$$G = \frac{\sqrt{\gamma_k}}{\gamma_k} \left[\frac{\Gamma\left(\frac{p+\beta}{2}+1\right) \Phi\left(-\frac{p+\beta}{2}, 1; -\gamma_k\right)}{\Gamma\left(\frac{p}{2}+1\right) \Phi\left(-\frac{p}{2}, 1; -\gamma_k\right)} \right]^{\frac{1}{\beta}}$$

其中, 所述 γ_k 为后验信噪比, 所述 ξ_k 为先验信噪比, 所述 $\gamma_k = \frac{\xi_k}{1+\xi_k}$, 所

15 述 p 为感知加权阶数, 所述 β 为高阶幅度谱的阶数。

14. 根据权利要求 13 所述的自适应语音增强方法, 其特征在于, 所述感知加权阶数通过以下方式获取:

在所述语音信号的频率谱中, 将 Bark 域的频带划分为若干个子带;

计算所述若干个子带的信噪比:

20 根据所述若干个子带的信噪比计算所述感知加权阶数。

15. 根据权利要求 14 所述的自适应语音增强方法, 其特征在于, 所述计算所述若干个子带的信噪比, 包括:

通过以下公式计算所述若干个子带的信噪比 SNR :

$$SNR(b, k) = 10 \log 10 \frac{\sum_{k=B_{low}(b)}^{B_{up}(b)} |Y_k(b, k) - \sqrt{\lambda_d(b, k)}|^2}{\sum_{k=B_{low}(b)}^{B_{up}(b)} \lambda_d(b, k)}$$

5 其中, 所述 b 为子带的序号, 所述 k 为频点数, 所述 $B_{low}(b)$ 为所述 Bark 域第 b 个子带的频点起点, 所述 $B_{up}(b)$ 为所述 Bark 域第 b 子带的频点终点。

16. 根据权利要求 15 所述的自适应语音增强方法, 其特征在于, 所述根据所述若干个子带的信噪比计算感知加权阶数, 为:

通过以下公式计算所述感知加权阶数 p :

$$10 \quad p(b, k) = \max \{ \min [\alpha_1 SNR(b, k) + \alpha_2, p_{\max}], p_{\min} \}$$

其中, 所述 α_1 , 所述 α_2 , 所述 $p_{\min}$ 和所述 $p_{\max}$ 均为实验经验值。

17. 根据权利要求 13 所述的自适应语音增强方法, 其特征在于, 所述高阶幅度谱的阶数 通过以下方式获取:

在所述语音信号的频率谱中, 将 Bark 域划分为若干个子带;

15 通过以下公式计算所述高阶幅度谱的阶数 β :

$$\beta(b, k) = \beta_{\max} - \frac{(\beta_{\max} - \beta_{\min}) \sum_{k=B_{low}(b)}^{B_{up}(b)} \frac{20}{3} \log 10 \left(\frac{f(k)}{A} + 1 \right)}{(B_{up}(b) - B_{low}(b)) \log 10 \left(\frac{F_s}{2A} + 1 \right)}$$

其中, 所述 F_s 为采样频率, 所述 $\beta_{\min}$, 所述 $\beta_{\max}$, 所述 $p_{\min}$, 所述 $p_{\max}$ 和所述 A 均为实验经验值, 所述 b 为子带的序号, 所述 k 为频点数, 所述 $B_{low}(b)$ 为所述 Bark 域第 b 个子带的频点起点, 所述 $B_{up}(b)$ 为所述 Bark 域第 b 个子带的

频点终点, 所述 $f(k) = kFs / N$ 为对所述接收的语音信号进行快速傅里叶变换后第 k 个频点的频率。

18. 根据权利要求 13 所述的自适应语音增强方法, 其特征在于,

所述 $\Gamma\left(\frac{p+\beta}{2}+1\right)$ 和 $\Gamma\left(\frac{p}{2}+1\right)$ 通过查询预存的 Γ 函数的输入输出对应关系得

5 到:

所述 $\Phi\left(-\frac{p+\beta}{2}, 1; -v_k\right)$ 和 $\Phi\left(-\frac{p}{2}, 1; -v_k\right)$ 通过查询预存的 Φ 函数的输入输出对应关系得到。

19. 根据权利要求 13 所述的自适应语音增强方法, 其特征在于, 所述根据所述谱增益系数获取纯净语音信号, 包括:

10 通过以下公式获取纯净语音信号 $\hat{x}_k$:

$$\hat{x}_k = \frac{\sqrt{v_k}}{\gamma_k} \left[\frac{\Gamma\left(\frac{p+\beta}{2}+1\right) \Phi\left(-\frac{p+\beta}{2}, 1; -v_k\right)}{\Gamma\left(\frac{p}{2}+1\right) \Phi\left(-\frac{p}{2}, 1; -v_k\right)} \right]^{1/\beta} Y_w(k)$$

其中, 所述 $Y_w(k)$ 为所述当前帧的信号幅值。

20. 一种电子设备, 其特征在于, 包括:

至少一个处理器; 以及,

15 与所述至少一个处理器通信连接的存储器; 其中,

所述存储器存储有可被所述至少一个处理器执行的指令, 所述指令被所述至少一个处理器执行, 以使所述至少一个处理器能够执行如权利要求 1 至 19 中任一所述的自适应语音增强方法。

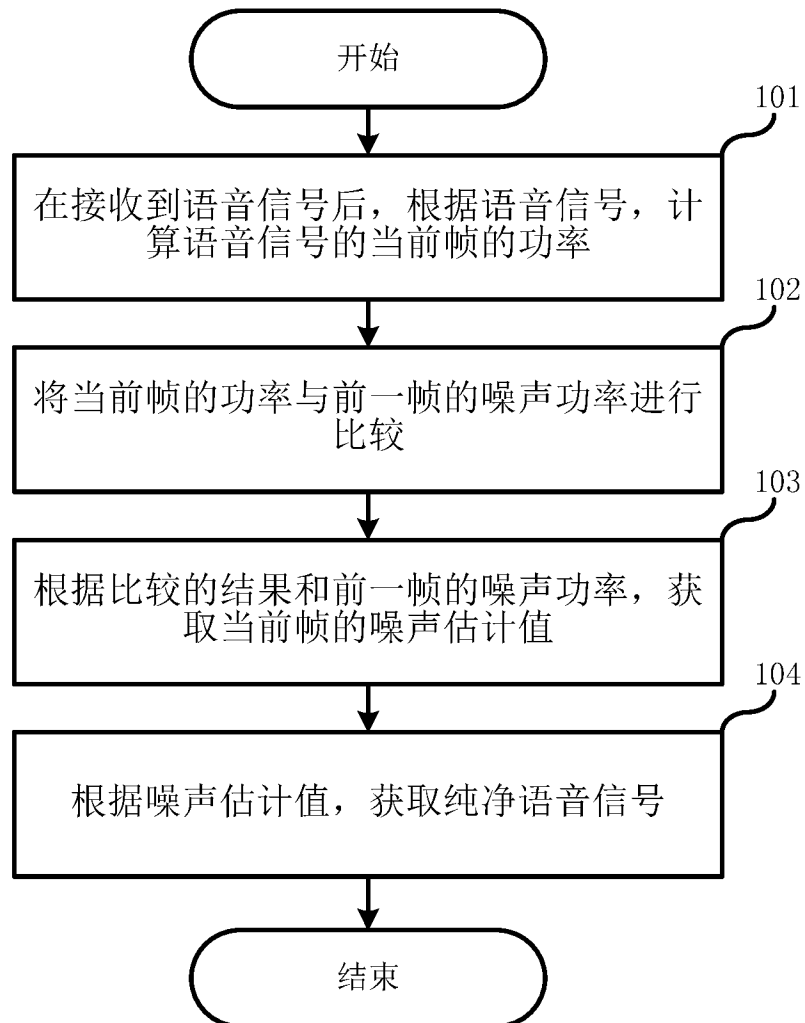


图 1

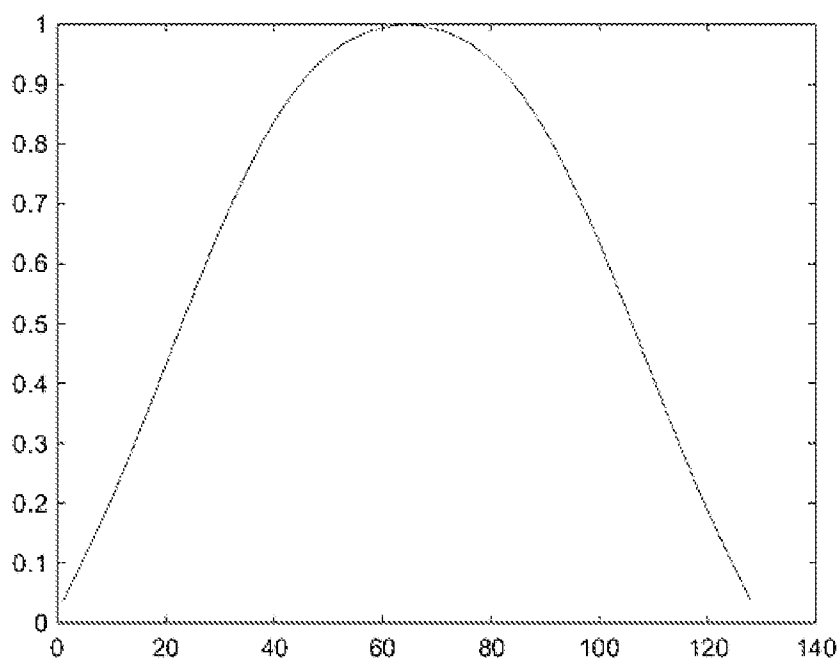


图 2

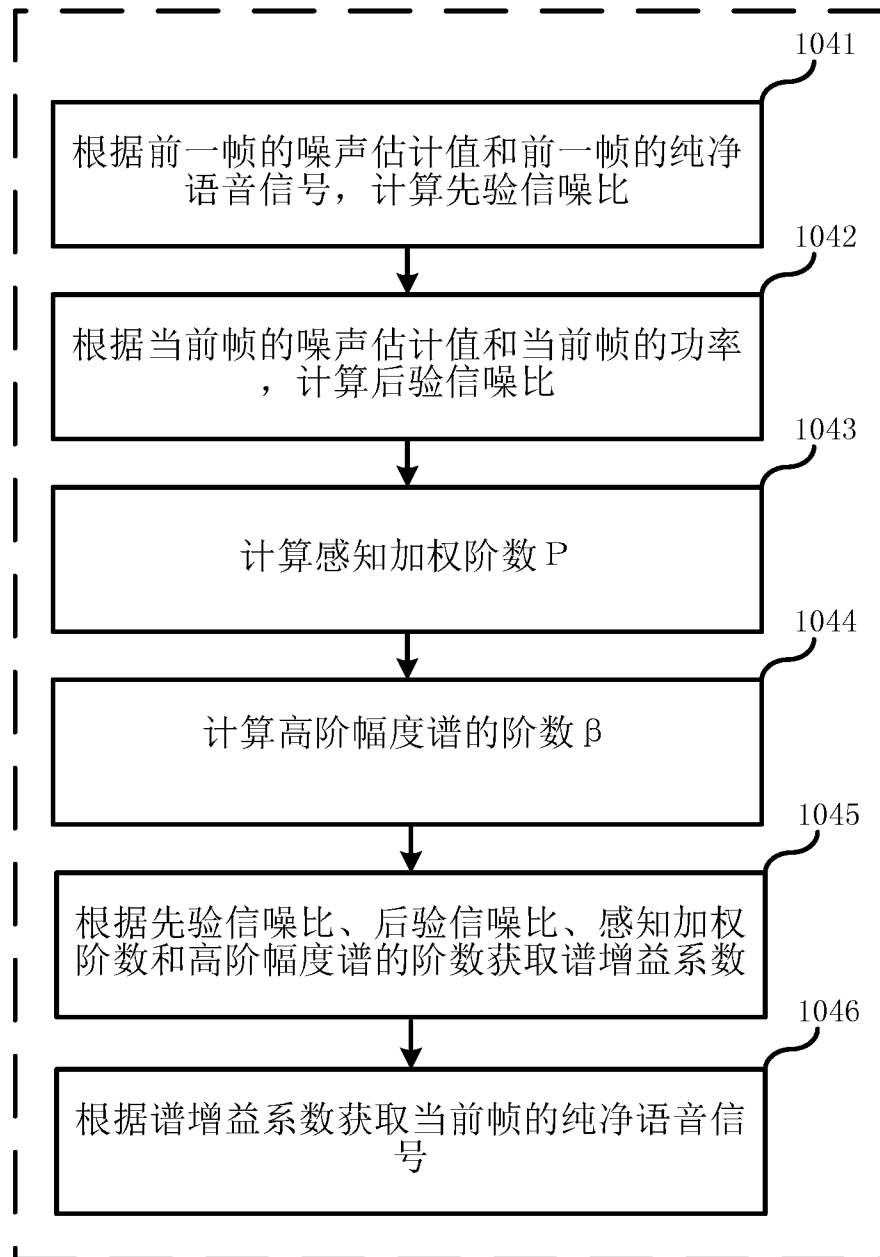


图 3

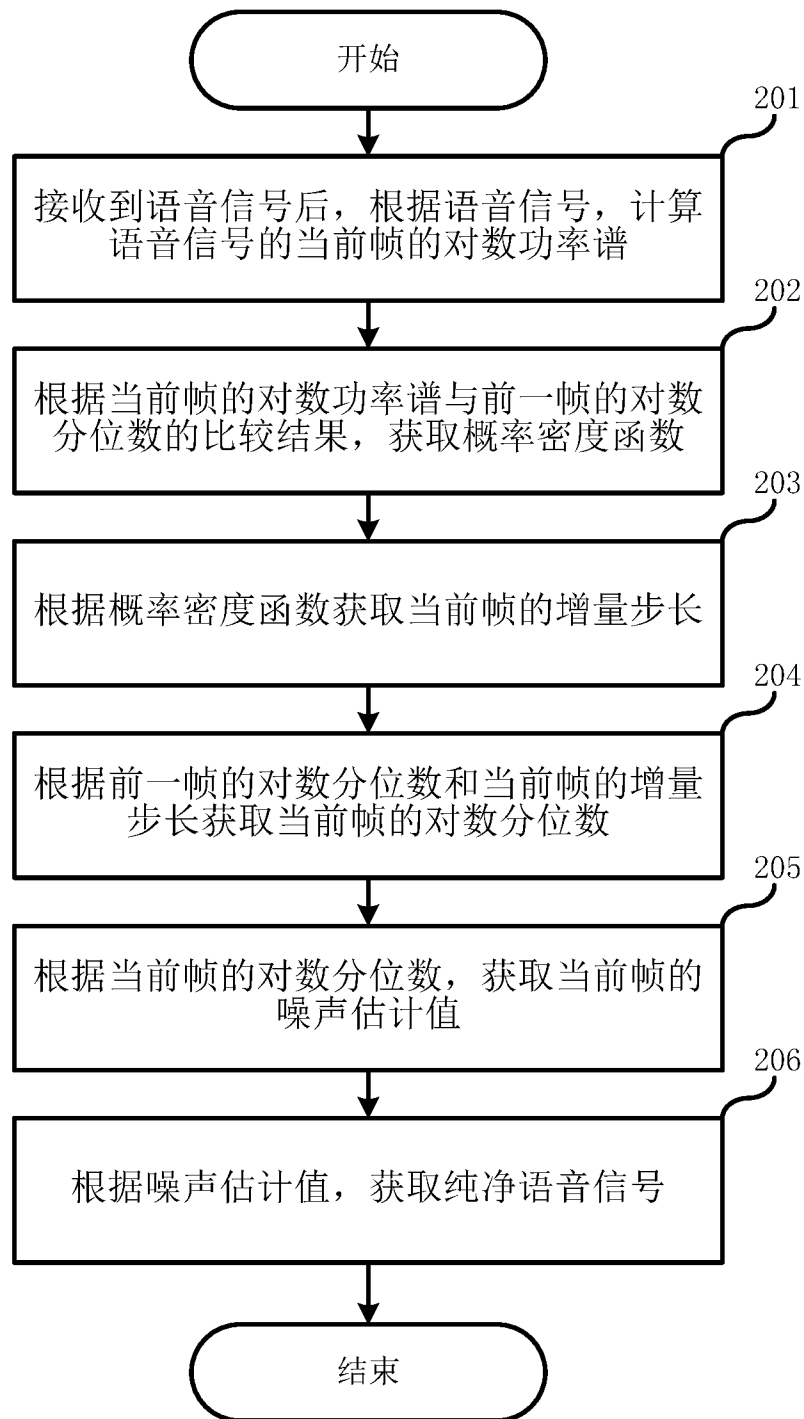


图 4

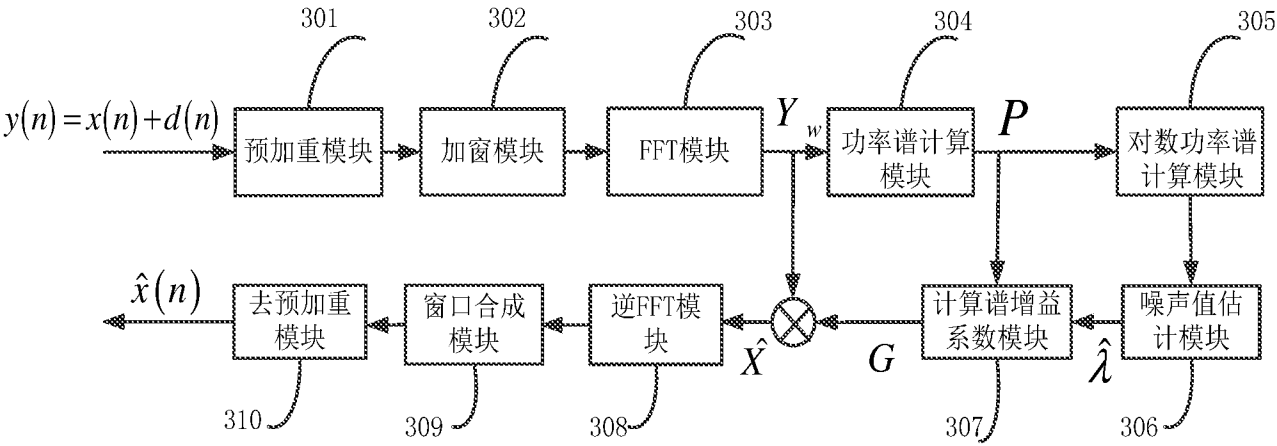


图 5

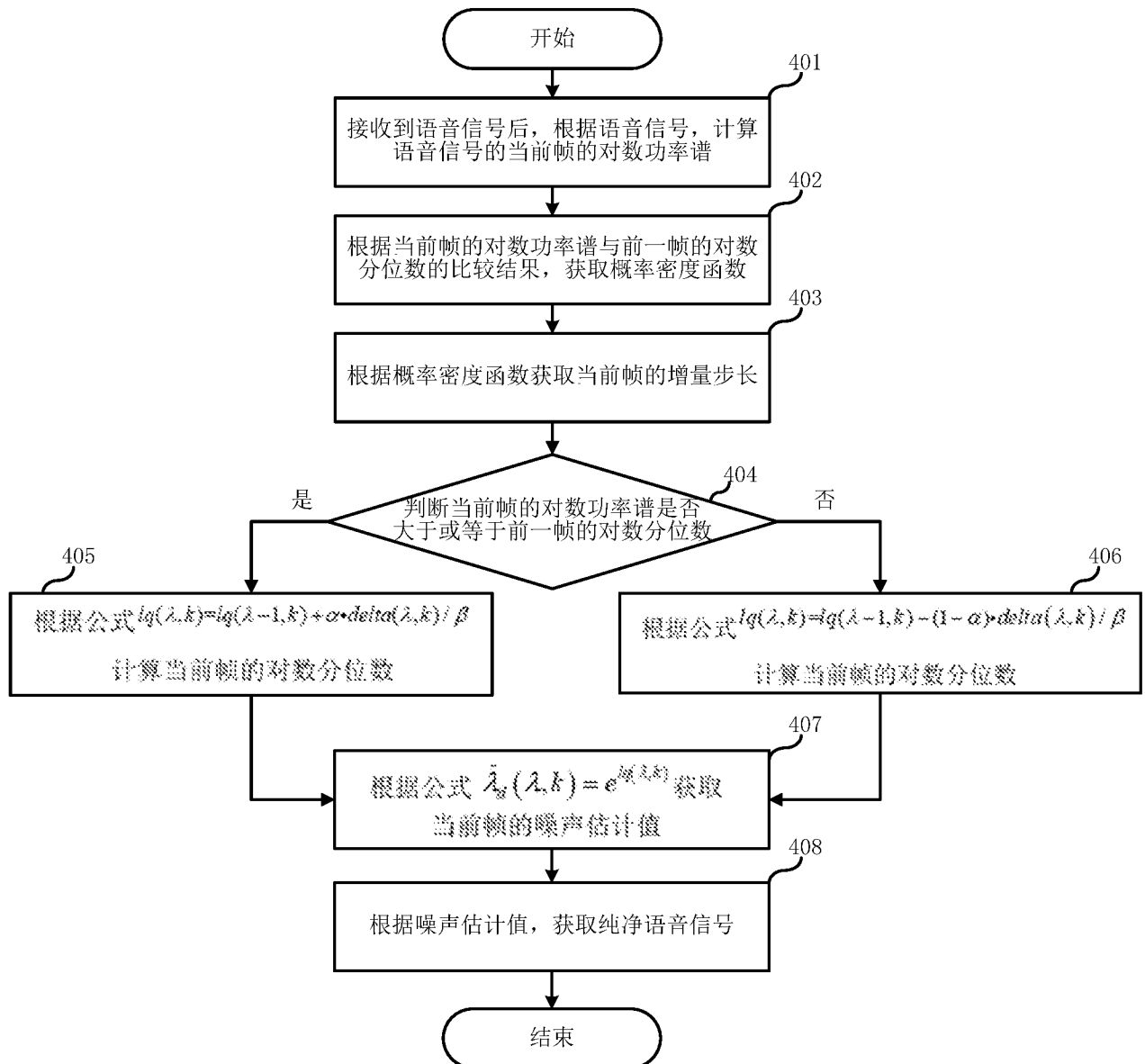


图 6

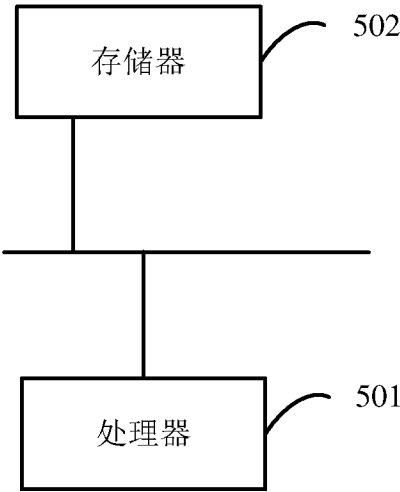


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/117972

A. CLASSIFICATION OF SUBJECT MATTER

G10L 21/0316(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)

WPI; EPODOC; CNPAT; CNKI; IEEE: 语音, 增强, 当前帧, 功率, 前一帧, 噪声功率, 噪声估计值, speech, enhance, noise, estimate, value, compare, previous, frame, current

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103646648 A (TSINGHUA UNIVERSITY et al.) 19 March 2014 (2014-03-19) description, paragraphs [0041]-[0095]	1, 20
A	CN 103021420 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES) 03 April 2013 (2013-04-03) entire document	1-20
A	CN 103730124 A (SHANGHAI JIAO TONG UNIVERSITY WUXI RESEARCH INSTITUTE) 16 April 2014 (2014-04-16) entire document	1-20
A	US 2006271362 A1 (NEC CORPORATION) 30 November 2006 (2006-11-30) entire document	1-20



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

14 August 2019

Date of mailing of the international search report

03 September 2019

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/117972

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	103646648	A	19 March 2014	None			
CN	103021420	A	03 April 2013	None			
CN	103730124	A	16 April 2014	None			
US	2006271362	A1	30 November 2006	KR	20060125572	A	06 December 2006
				CN	1892822	A	10 January 2007
				JP	2006337415	A	14 December 2006
				EP	1729286	A2	06 December 2006

国际检索报告

国际申请号

PCT/CN2018/117972

A. 主题的分类 G10L 21/0316(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)) WPI;EPDOC;CNPAT;CNKI;IEEE:语音, 增强, 当前帧, 功率, 前一帧, 噪声功率, 噪声估计值, speech, enhance, noise, estimate, value, compare, previous, frame, current																	
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 103646648 A (清华大学 等) 2014年 3月 19日 (2014 - 03 - 19) 说明书第[0041]-[0095]段</td> <td>1、20</td> </tr> <tr> <td>A</td> <td>CN 103021420 A (中国科学院自动化研究所) 2013年 4月 3日 (2013 - 04 - 03) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 103730124 A (上海交通大学无锡研究院) 2014年 4月 16日 (2014 - 04 - 16) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 2006271362 A1 (NEC CORPORATION) 2006年 11月 30日 (2006 - 11 - 30) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 103646648 A (清华大学 等) 2014年 3月 19日 (2014 - 03 - 19) 说明书第[0041]-[0095]段	1、20	A	CN 103021420 A (中国科学院自动化研究所) 2013年 4月 3日 (2013 - 04 - 03) 全文	1-20	A	CN 103730124 A (上海交通大学无锡研究院) 2014年 4月 16日 (2014 - 04 - 16) 全文	1-20	A	US 2006271362 A1 (NEC CORPORATION) 2006年 11月 30日 (2006 - 11 - 30) 全文	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
X	CN 103646648 A (清华大学 等) 2014年 3月 19日 (2014 - 03 - 19) 说明书第[0041]-[0095]段	1、20															
A	CN 103021420 A (中国科学院自动化研究所) 2013年 4月 3日 (2013 - 04 - 03) 全文	1-20															
A	CN 103730124 A (上海交通大学无锡研究院) 2014年 4月 16日 (2014 - 04 - 16) 全文	1-20															
A	US 2006271362 A1 (NEC CORPORATION) 2006年 11月 30日 (2006 - 11 - 30) 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 2019年 8月 14日		国际检索报告邮寄日期 2019年 9月 3日															
ISA/CN的名称和邮寄地址 中国国家知识产权局(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 李萌 电话号码 86-(10)-53961536															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/117972

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	103646648	A	2014年 3月 19日	无			
CN	103021420	A	2013年 4月 3日	无			
CN	103730124	A	2014年 4月 16日	无			
US	2006271362	A1	2006年 11月 30日	KR	20060125572	A	2006年 12月 6日
				CN	1892822	A	2007年 1月 10日
				JP	2006337415	A	2006年 12月 14日
				EP	1729286	A2	2006年 12月 6日