

(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



(10) 国际公布号
WO 2017/177782 A1

(43) 国际公布日
2017年10月19日 (19.10.2017)

- (51) 国际专利分类号:
G10L 21/0232 (2013.01) G10L 25/21 (2013.01)
G10L 21/0324 (2013.01)
- (21) 国际申请号: PCT/CN2017/076653
- (22) 国际申请日: 2017年3月14日 (14.03.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201610235392.9 2016年4月15日 (15.04.2016) CN
- (71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。
- (72) 发明人: 梁俊斌 (LIANG, Junbin); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。
- (74) 代理人: 广州华进联合专利商标代理有限公司 (ADVANCE CHINA IP LAW OFFICE); 中国广东省

广州市天河区花城大道85号3901房, Guangdong 510623 (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: VOICE SIGNAL CASCADE PROCESSING METHOD AND TERMINAL, AND COMPUTER READABLE STORAGE MEDIUM

(54) 发明名称: 语音信号级联处理方法、终端和计算机可读存储介质

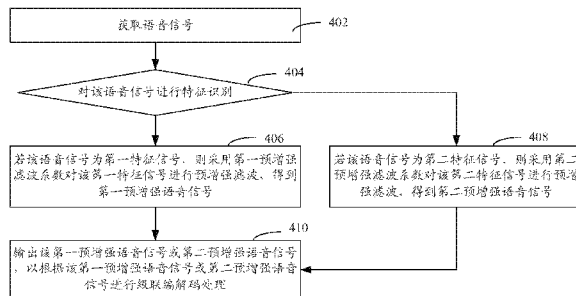


图 4

402 Acquire a voice signal
 404 Perform feature recognition on the voice signal
 406 If the voice signal is a first feature signal, perform pre-enhancement filtering on the first feature signal using a first pre-enhancement filtering coefficient to obtain a first pre-enhanced voice signal
 408 If the voice signal is a second feature signal, perform pre-enhancement filtering on the second feature signal using a second pre-enhancement filtering coefficient to obtain a second pre-enhanced voice signal
 410 Output the first pre-enhanced voice signal or the second pre-enhanced voice signal to perform cascade encoding/decoding processing according to the first pre-enhanced voice signal or the second pre-enhanced voice signal

(57) Abstract: A voice signal cascade processing method, comprising: acquiring a voice signal (402); performing feature recognition on the voice signal (404); if the voice signal is a first feature signal, performing pre-enhancement filtering on the first feature signal using a first pre-enhancement filtering coefficient to obtain a first pre-enhanced voice signal (406); if the voice signal is a second feature signal, performing pre-enhancement filtering on the second feature signal using a second pre-enhancement filtering coefficient to obtain a second pre-enhanced voice signal (408); outputting the first pre-enhanced voice signal or the second pre-enhanced voice signal to perform cascade encoding/decoding processing according to the first pre-enhanced voice signal or the second pre-enhanced voice signal (410).

(57) 摘要:

[见续页]



WO 2017/177782 A1



本国际公布:

- 包括国际检索报告(条约第 21 条(3))。

一种语音信号级联处理方法,包括:获取语音信号(402);对该语音信号进行特征识别(404);若该语音信号为第一特征信号,则采用第一预增强滤波系数对该第一特征信号进行预增强滤波,得到第一预增强语音信号(406);若该语音信号为第二特征信号,则采用第二预增强滤波系数对该第二特征信号进行预增强滤波,得到第二预增强语音信号(408);输出该第一预增强语音信号或第二预增强语音信号,以根据该第一预增强语音信号或第二预增强语音信号进行级联编解码处理(410)。

说明书

发明名称：语音信号级联处理方法、终端和计算机可读存储介质

本申请要求于 2016 年 04 月 15 日提交中国专利局、申请号为 201610235392.9、发明名称为“语音信号级联处理方法和装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5

技术领域

本发明涉及音频数据处理领域，特别是涉及一种语音信号级联处理方法、终端和非易失性计算机可读存储介质。

10 背景技术

随着 VOIP（Voice over Internet Protocol，网络电话）业务的推广，不同网络间互融应用日趋增多，如经互联网的 IP 电话与经 PSTN（Public Switched Telephone Network，公共交换电话网络）固定电话互通，IP 电话与无线网络的手机互通。不同网络的语音采用的是不同的语音编解码，如无线 GSM
15 （Global System for Mobile Communication，全球移动通信系统）网络采用 AMR-NB 编码，固定电话采用 G711 编码，IP 电话采用 G729 等编码，由于各网络终端支持的语音编码格式不一致，必然导致通话链路出现多次编解码过程，目的使级联编解码后不同网络终端可以互通语音对接，然而，当前所用语音编码绝大部分都是有损编码器，即每次编解码必然导致语音质量下降，
20 级联编解码的次数越多语音质量下降越剧烈，结果导致语音双方听不清对方的说话内容，即语音可懂度下降。

发明内容

根据本申请的各种实施例，提供一种语音信号级联处理方法、终端和非
25 易失性计算机可读存储介质。

一种语音信号级联处理方法，包括：

获取语音信号;

对所述语音信号进行特征识别;

若所述语音信号为第一特征信号, 则采用第一预增强滤波系数对所述第一特征信号进行预增强滤波, 得到第一预增强语音信号; 若所述语音信号为第二特征信号, 则采用第二预增强滤波系数对所述第二特征信号进行预增强滤波, 得到第二预增强语音信号; 以及

输出所述第一预增强语音信号或第二预增强语音信号, 以根据所述第一预增强语音信号或第二预增强语音信号进行级联编解码处理。

一种终端, 包括存储器及处理器, 所述存储器中储存有计算机可读指令, 所述指令被所述处理器执行时, 使得所述处理器执行以下步骤:

获取语音信号;

对所述语音信号进行特征识别;

若所述语音信号为第一特征信号, 则采用第一预增强滤波系数对所述第一特征信号进行预增强滤波, 得到第一预增强语音信号;

若所述语音信号为第二特征信号, 则采用第二预增强滤波系数对所述第二特征信号进行预增强滤波, 得到第二预增强语音信号; 以及

输出所述第一预增强语音信号或第二预增强语音信号, 以根据所述第一预增强语音信号或第二预增强语音信号进行级联编解码处理。

一个或多个包含计算机可执行指令的非易失性计算机可读存储介质, 当所述计算机可执行指令被一个或多个处理器执行时, 使得所述处理器执行以下步骤:

获取语音信号;

对所述语音信号进行特征识别;

若所述语音信号为第一特征信号, 则采用第一预增强滤波系数对所述第一特征信号进行预增强滤波, 得到第一预增强语音信号;

若所述语音信号为第二特征信号, 则采用第二预增强滤波系数对所述第二特征信号进行预增强滤波, 得到第二预增强语音信号; 以及

输出所述第一预增强语音信号或第二预增强语音信号，以根据所述第一预增强语音信号或第二预增强语音信号进行级联编解码处理。

本发明的一个或多个实施例的细节在下面的附图和描述中提出。本发明的其它特征、目的和优点将从说明书、附图以及权利要求书变得明显。

5

附图说明

为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图 1 为一个实施例中语音信号级联处理方法的应用环境示意图；

图 2 为一个实施例中终端的内部结构示意图；

图 3A 为一个实施例中经过级联编解码后第一特征信号的频率能量损伤示意图；

图 3B 为一个实施例中经过级联编解码后第二特征信号的频率能量损伤示意图；

图 4 为一个实施例中语音信号级联处理方法的流程图；

图 5 为根据音频训练集中的训练样本进行离线训练得到第一预增强滤波系数和第二预增强滤波系数的具体流程图；

图 6 为一个实施例中获取该语音信号的基音周期；

图 7 为三电平削波处理原理示意图；

图 8 为一段语音的基音周期计算结果示意图；

图 9 为一个实施例中离线训练的预增强滤波系数对在线通话的语音输入信号进行增强的示意图；

图 10 为级联编解码后信号经过预增强后的级联编解码信号的示意图；

图 11 为未做增强的级联编解码的信号频谱与增强后的级联编解码的信

号频谱的对比示意图；

图 12 为未做增强的级联编解码的信号频谱与增强后的级联编解码的信号频谱的中高频部分的对比示意图；

图 13 为一个实施例中语音信号级联处理装置的结构框图；

5 图 14 为另一个实施例中语音信号级联处理装置的结构框图；

图 15 为一个实施例中训练模块的内部结构示意图；

图 16 为另一个实施例中语音信号级联处理装置的结构框图。

具体实施方式

10 为了使本发明的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本发明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发明，并不用于限定本发明。

可以理解，本发明所使用的术语“第一”、“第二”等可在本文中用于描述各种元件，但这些元件不受这些术语限制。这些术语仅用于将第一个元件
15 与另一个元件区分。举例来说，在不脱离本发明的范围的情况下，可以将第一客户端称为第二客户端，且类似地，可将第二客户端称为第一客户端。第一客户端和第二客户端两者都是客户端，但其不是同一客户端。

图 1 为一个实施例中语音信号级联处理方法的应用环境示意图。如图 1 所示，该应用环境包括第一终端 110、第一网络 120、第二网络 130 和第二终端
20 140。第一终端 110 接收语音信号，将语音信号经过第一网络 120 和第二网络 130 的编解码处理后，被第二终端 140 接收。第一终端 110 将语音对所述语音信号进行特征识别；若所述语音信号为第一特征信号，则采用第一预增强滤波系数对所述第一特征信号进行预增强滤波，得到第一预增强语音信号；以及若所述语音信号为第二特征信号，则采用第二预增强滤波系数对所述第二特征信号进行预增强滤波，得到第二预增强语音信号；输出所述第一
25 预增强语音信号或第二预增强语音信号，经过第一网络 120 和第二网络 130 进行级联编解码处理后得到预增强后的级联编解码信号，第二终端 140 接收

到预增强后的级联编解码信号，接收到的信号可懂度高。第一终端 110 接收第二终端 140 发送的经过第二网络 130、第一网络 120 的语音信号，同样对接收到的语音信号进行预增强滤波处理。

图 2 为一个实施例中终端的内部结构示意图。如图 2 所示，该终端包括通过系统总线连接的处理器、存储介质、内存、网络接口、声音采集装置和扬声器。其中，终端的存储介质存储有操作系统和计算机可读指令，该计算机可读指令被执行时，使得处理器执行步骤以实现一种语音信号级联处理方法。该处理器用于提供计算和控制能力，支撑整个终端的运行，该处理器被用于执行一种语音信号级联处理方法，包括获取语音信号；对该语音信号进行特征识别；若该语音信号为第一特征信号，则采用第一预增强滤波系数对该第一特征信号进行预增强滤波，得到第一预增强语音信号；以及若该语音信号为第二特征信号，则采用第二预增强滤波系数对该第二特征信号进行预增强滤波，得到第二预增强语音信号；输出该第一预增强语音信号或第二预增强语音信号，以根据该第一预增强语音信号或第二预增强语音信号进行级联编解码处理。该终端可以是能进行网络通话的电话机、手机、平板电脑或者个人数字助理等。本领域技术人员可以理解，图 2 中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的终端的限定，具体的终端可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

20

因经过级联编解码后的语音信号，其中高频会明显的损伤，且第一特征信号和第二特征信号的语音可懂度在级联编解码后会有不同程度的而影响，因为影响语音可懂度的关键成分是语音信号的中高频能量信息，第一特征信号由于在其基频较低（一般在 125Hz（赫兹）以内），主要能量成分集中在中低频（1000Hz 以下），中高频（1000Hz 以上）成分较少，第二特征信号的基频较高（一般在 125Hz 以上），中高频成分比第一特征信号要多，如图 3A 和图 3B 所示，经过级联编解码后第一特征信号和第二特征信号的频率能量均有

25

损伤，由于第一特征信号中高频比例偏低，经过级联编解码后中高频能量更低，使第一特征信号的语音可懂度影响极大，导致收听方感觉听到的声音模糊难以听清楚说话内容，而第二特征信号虽然中高频也有损耗，但经过级联编解码后中高频还有足够能量以达到较好的语音可懂度。从语音编解码原理来说，以语音听觉失真最小为准则的一种编解码模型 CELP(Code Excited Linear Prediction, 码激励线性预测编码) 合成的语音为例，由于第一特征信号语音的频谱能量分布很不均衡，大部分能量集中在中低频，所以编码过程主要确保中低频失真最小，而占能量比例较小的中高频相对失真较大，相反，第二特征信号的频谱能量分布较为均衡，中高频也有较多成分，所以经过编解码后中高频成分能量损失相对不高。也就是，第一特征信号和第二特征信号在经过级联编解码后的可懂度下降表现是有明显差异的。图 3A 中曲实线为第一特征信号原始信号，虚线为经级联编解码后的信号。图 3B 中曲实线为第二特征信号原始信号，虚线为经级联编解码后的信号。图 3A 和图 3B 中横坐标为频率，纵坐标为能量，且为归一化后的能量值。归一化是基于第一特征信号或第二特征信号中最大峰值作为归一的。第一特征信号可为男声信号，第二特征信号可为女声信号。

图 4 为一个实施例中语音信号级联处理方法的流程图。如图 4 所示，一种语音信号级联处理方法，运行于图 1 的终端上，包括：

步骤 402，获取语音信号。

本实施例中，语音信号为识别输入的原始语音信号中的语音信号。终端获取到经过级联编解码处理后的原始语音信号，识别出原始语音信号中的语音信号。级联编解码与原始语音信号经过的实际链路环节相关，例如支持 G.729A 的 IP 电话与 GSM 手机互通，则级联编解码可为 G.729A 编码+G.729 解码+AMRNB 编码+AMRNB 解码。

语音可懂度是指听者听清楚和理解说话人的口头表述内容的程度。

步骤 404，对该语音信号进行特征识别。

本实施例中，对该语音信号进行特征识别包括：获取语音信号的基音周

期；判断该语音信号的基音周期是否大于预设周期值，若是，则该语音信号为第一特征信号，若否，则该语音信号为第二特征信号。

具体地，声带振动的频率称为基频，相应的周期称为基音周期。预设周期值可根据需要设定，如周期为 60 个样点。语音信号的基音周期大于 60 个样点，则该语音信号为第一特征信号，若小于或等于 60 个样点，则该语音信号为第二特征信号。

步骤 406，若该语音信号为第一特征信号，则采用第一预增强滤波系数对该第一特征信号进行预增强滤波，得到第一预增强语音信号。

步骤 408，若该语音信号为第二特征信号，则采用第二预增强滤波系数对该第二特征信号进行预增强滤波，得到第二预增强语音信号。

第一特征信号和第二特征信号可为不同频段范围内的语音信号。

步骤 410，输出该第一预增强语音信号或第二预增强语音信号，以根据该第一预增强语音信号或第二预增强语音信号进行级联编解码处理。

上述语音信号级联处理方法，通过对语音信号进行特征识别，对第一特征信号采用第一预增强滤波系数进行预增强滤波处理，对第二特征信号采用第二预增强滤波系数进行预增强滤波处理，将预增强语音进行级联编解码处理，接收方能更清楚听清语音信息，提高了经过级联编解码后的语音信号的可懂度，针对第一特征信号和第二特征信号分别采用对应的滤波系数进行增强滤波处理，针对性更强，滤波处理更加准确。

在一个实施例中，上述语音信号级联处理方法，在获取语音信号之前，还包括：获取输入的原始音频信号；检测该原始音频信号为语音信号或非语音信号；若该原始音频信号为语音信号，则获取语音信号；若该原始音频信号为非语音信号，则对该非语音信号进行高通滤波处理。

本实施例中，通过 VAD 判断样本语音信号为语音信号或非语音信号。

25 对非语音进行高通滤波处理，降低信号的噪声。

在一个实施例中，在该获取语音信号之前，该语音信号级联处理方法还包括：根据音频训练集中的训练样本进行离线训练得到第一预增强滤波系数

和第二预增强滤波系数。

本实施例中，男音频训练集中的训练样本可为录制或从网络上筛选得到的语音信号等。

如图 5 所示，在一个实施例中，该根据音频训练集中的训练样本进行离线训练得到第一预增强滤波系数和第二预增强滤波系数的步骤包括：

步骤 502，从音频训练集中获取样本语音信号，该样本语音信号为第一特征样本语音信号或第二特征样本语音信号。

本实施例中，预先建立音频训练集，音频训练集中包含多个第一特征样本语音信号和第二特征样本语音信号。音频训练集中的第一特征样本语音信号和第二特征样本语音信号独立存在。第一特征样本语音信号和第二特征样本语音信号为不同特征信号的样本语音信号。

在步骤 502 之后，还包括：判断该样本语音信号是否为语音信号，若是，则对样本语音信号进行模拟级联编解码处理，得到降级语音信号，若否，则重新从音频训练集中获取样本语音信号。

本实施例中，采用 VAD（Voice Activity Detection，语音活跃度检测）判断样本语音信号是否为语音信号。VAD 是一种语音检测算法，是基于能量、过零率和低噪估计等估算语音。

判断样本语音信号是否为语音信号的步骤包括（a1）至（a5）：

（a1）接收连续语音，并从此连续语音中获取音框；

（a2）计算音框的能量，并根据这些能量取得能量门限值；

（a3）分别计算获取的音框的越零率，并根据这些越零率取得越零率门限值；

（a4）使用线性回归演绎法，并以（a2）中获取的能量及（a3）中获取的越零率作为线性回归演绎法的输入参数，用来判断每一音框是否为活动语音或非活动语音；

（a5）根据能量门限值及越零率门限值，自（a4）中的活动语音及非活动语音中取得活动语音起点及活动语音终点。

VAD 检测方法可采用双门限检测法或基于自相关极大值的语音检测法。

双门限检测法的过程包括:

(b1) 在开始阶段做预加重和分帧处理, 将语音信号分成一帧一帧;

(b2) 设置初始化参数, 包括最大静音长度、短时能量的门限和短时过
5 零率的门限;

(b3) 判断当语音在静音段或过渡段时, 如果语音信号的短时能量值大
于短时能量的高门限, 或者语音信号的短时过零率大于短时过零率的高门限,
则确认进入语音段, 如果短时能量的值大于短时能量的低门限或者过零率的
值大于过零率的低门限, 则语音处于过渡段, 否则语音仍处于静音段;

10 (b4) 当语音信号在语音段时, 判断如果短时能量的低门限或短时过零
率的值大于短时过零率的低门限, 则语音信号仍然处于语音段;

(b5) 如果静音长度小于设置的最大静音长度, 则表明语音还尚未结束,
还在语音段, 如果语音的长度小于最小噪声长度, 则认为语音太短, 此时是
噪声, 同时判断语音处于静音段; 否则语音进入结束段。

15 步骤 504, 对该样本语音信号进行模拟级联编解码处理, 得到降级语音
信号。

模拟级联编解码是指模拟原始语音信号经过的实际链路环节, 例如支持
G.729A 的 IP 电话与 GSM 手机互通, 则模拟级联编解码可为 G.729A 编码
+G.729 解码+AMRNB 编码+AMRNB 解码。对样本语音信号经过离线的级联
20 编解码处理后得到降级语音信号。

步骤 506, 获取该降级语音信号与样本语音信号对应不同频点上的能量
衰减值, 将该能量衰减值作为频点能量补偿值。

具体地, 将每个频点的样本语音信号所对应的能量值减去降级语音信号
所对应的能量值得到对应频点的能量衰减值, 该能量衰减值即为后续需要的
25 该频点能量补偿值。

步骤 508, 对该音频训练集中的第一特征信号所对应的频点能量补偿值
求平均得到第一特征信号在不同频点上的能量平均补偿值, 以及对该音频训

练集中的第二特征信号所对应的频点能量补偿值求平均得到第二特征信号在不同频点上的能量平均补偿值。

具体地，对音频训练集中第一特征信号所有的能量补偿值求平均得到第一特征信号在不同频点上的能量平均补偿值，对音频训练集中第二特征信号所有的能量补偿值求平均得到第二特征信号在不同频点上的能量平均补偿值。

步骤 510，根据该第一特征信号在不同频点上的能量平均补偿值进行滤波拟合得到第一预增强滤波系数，以及根据该第二特征信号在不同频点上的能量平均补偿值进行滤波拟合得到第二预增强滤波系数。

本实施例中，基于第一特征信号在不同频点上的能量平均补偿值为目标，采用自适应滤波器拟合方式对第一特征信号的能量平均补偿值进行滤波拟合得到一组第一预增强滤波系数。基于第二特征信号在不同频点上的能量平均补偿值为目标，采用自适应滤波器拟合方式对第二特征信号的能量平均补偿值进行滤波拟合得到一组第二预增强滤波系数。

预增强滤波器可使用 FIR (Finite Impulse Response, 有限长单位冲激响应滤波器) 滤波器： $y[n] = a_0 * x[n] + a_1 * x[n-1] + \dots + a_m * x[n-m]$ 。

FIR 滤波器的预增强滤波系数 $a_0 \sim a_m$ 可通过 matlab 的 fir2 函数进行计算求得，函数 $b = \text{fir2}(n, f, m)$ 是用来设计多通带任意响应 FIR 滤波器，该滤波器的幅频特性由向量对 f 和 m 确定， f 为归一化频率向量， m 为对应频率点上的幅度， n 为滤波器的阶数。本实施例中，将各频点的能量补偿值作为 m ，输入 fir2 函数从而计算得到 b 。

上述离线训练得到第一预增强滤波系数和第二预增强滤波系数，通过离线训练可以准确得到第一预增强滤波系数和第二预增强滤波系数，方便后续进行在线滤波处理得到增强后的语音信号，有效提高级联编解码后的语音信号的可懂度。

如图 6 所示，在一个实施例中，该获取该语音信号的基音周期包括：
步骤 602，对该语音信号进行带通滤波。

本实施例中，对语音信号进行带通滤波可采用 80Hz~1500Hz 的滤波器进行滤波，也可采用 60~1000Hz 的带通滤波器进行滤波等，不限于此。也就是带通滤波的频率范围根据具体需求设置。

步骤 604，将该带通滤波后的语音信号进行预加重处理。

5 本实施例中，预加重是指发送端对输入信号高频分量的提升。

步骤 606，对该语音信号以矩形窗进行平移分帧，每帧窗长第一采样点数，每帧平移第二采样点数。

本实施例中，矩形窗的窗长为第一采样点数，第一采样点数可为 280 点，第二采样点可为 80 点，第一采样点数和第二采样点数不限于此。80 点对应的是 10ms（毫秒）数据，采用 80 点平移，则是每帧都会引入 10ms 的新数据进行计算。

步骤 608，对每帧信号进行三电平削波处理。

本实施例中，三电平削波处理，如设定正负阈值，如果样点值大于正阈值则输出 1，如果样点值小于负阈值则输出 -1，其余情况输出为 0。

15 如图 7 所示，正阈值为 C，负阈值为 -C，若样点值超过正阈值 C，则输出 1，样点值小于负阈值 -C，则输出 -1，其余输出为 0。

对每帧信号进行三电平削波处理得到 $t(i)$ ，其中， i 取值范围为 1~280。

步骤 610，对每帧内采样点计算自相关值。

本实施例中，每帧内采样点计算自相关值为两个因子的积除以各自的开方根的乘积。计算自相关值的公式为：

$$r(k) = \sum_{l=1}^{121} (t(k+l-1) * t(l)) / (\text{sqrt}(\sum_{l=1}^{121} (t(k+l-1) * t(k+l-1))) * \text{sqrt}(\sum_{l=1}^{121} (t(l) * t(l)))) , k = 20 \sim 160$$

其中， $r(k)$ 为自相关值， $t(k+l-1)$ 为对应的 $(k+l-1)$ 的三电平削波处理的结果， k 取值为 20 至 160 是常规的基音周期搜索范围，若对换为基频则为 8000/20~8000/160，即 50Hz~400Hz 范围，即人声正常基频范围， k 超出 20~160

25 可认为非人类正常声音基频范围，可不用计算，节省计算时间。

因 k 最大值为 160, l 的最大值为 121, 则 t 的最大范围为 $160+121-1=280$, 故三电平削波中 i 的最大值为 280。

步骤 612, 以每帧中自相关值最大者所对应的序号作为每帧的基音周期。

本实施例中, 通过计算每帧中自相关值, 可得到每帧中自相关值最大者
5 对应的序号, 将该自相关最大者对应的序号作为每帧的基音周期。

在其他实施例中, 步骤 602 和步骤 604 可以省略。

图 8 为一段语音的基音周期计算结果示意图。如图 8 所示, 第一幅图中的横坐标为采样点的序号, 纵坐标为采样点的样点值即采样点的幅值, 可知采样点的样点值呈现变化, 有的采样点的样点值大, 有的采样点的样点值小。
10 第二幅图中的横坐标为帧数, 纵坐标为基音周期值, 对于语音帧求取基音周期, 非语音帧的基音周期默认为 0。

下面结合具体的实施例描述上述语音信号级联处理方法。如图 9 所示, 以第一特征信号为男声, 第二特征信号为女声为例, 上述语音信号级联处理方法包括离线训练部分和在线处理部分。离线训练部分包括:

15 (c1) 从男女声训练集中获取样本语音信号。

(c2) vad 判决样本语音信号是否为语音信号, 若是, 则执行步骤 (c3), 若否, 则返回 (c2)。

(c3) 若为语音信号, 则对样本语音信号进行模拟级联编解码处理, 得到降级语音信号。

20 将样本语音信号经过实际链路环节所需要经过的多个编解码环节, 例如支持 G.729A 的 IP 电话与 GSM 手机互通, 则模拟级联编解码可为 G.729A 编码+G.729 解码+AMRNB 编码+AMRNB 解码。对样本语音信号经过离线的级联编解码处理后得到降级语音信号。

(c4) 计算各频点能量衰减值, 即为能量补偿值。

25 具体地, 将每个频点的样本语音信号所对应的能量值减去降级语音信号所对应的能量值得到对应频点的能量衰减值, 该能量衰减值即为后续需要的该频点能量补偿值。

(c5) 分别计算男声和女声的频点能量补偿值的平均值。

对该男女声训练集中的男声所对应的频点能量补偿值求平均得到男声在不同频点上的能量平均补偿值, 以及对该男女声训练集中的女声所对应的频点能量补偿值求平均得到女声在不同频点上的能量平均补偿值。

5 (c6) 计算男声预增强滤波系数和女声预增强滤波系数。

基于男声在不同频点上的能量平均补偿值为目标, 采用自适应滤波器拟合方式对男声的能量平均补偿值进行滤波拟合得到一组男声预增强滤波系数。基于女声在不同频点上的能量平均补偿值为目标, 采用自适应滤波器拟合方式对女声的能量平均补偿值进行滤波拟合得到一组女声预增强滤波系

10 数。

在线处理部分包括:

(d1) 语音信号输入。

(d2) vad 检测是否为语音信号, 若是, 则执行步骤 (d3), 若否执行步骤 (d4)。

15 (d3) 判断语音信号为男声或女声, 若为男声, 执行步骤 (d4), 若为女声, 执行步骤 (d5)。

(d4) 调用离线训练得到的男声预增强滤波系数对男声语音信号进行预增强滤波处理, 得到增强后的语音信号。

20 (d5) 调用离线训练得到的女声预增强滤波系数对女声语音信号进行预增强滤波处理, 得到增强后的语音信号。

(d6) 对非语音信号进行高通滤波处理, 得到增强后语音。

上述语音可懂度提升方法, 对非语音进行高通滤波处理, 降低信号的噪声, 通过识别出语音信号为男声信号或女声信号, 对男声信号采用离线训练得到的男声预增强滤波系数进行预增强滤波处理, 对女声信号采用离线训练

25 得到的女声预增强滤波系数进行预增强滤波处理, 针对男声信号和女声信号分别采用对应的滤波系数进行增强滤波处理, 提高了语音信号的可懂度, 因针对男声和女声分别处理, 针对性更强, 滤波处理更加准确。

图 10 为级联编解码后信号经过预增强后的级联编解码信号的示意图。如图 10 所示，第一幅图为原始信号，第二幅图为级联编解码后的信号，第三幅图 5 为经预增强滤波处理后的级联编解码信号。由此可知，经过预增强后的级联编解码信号相比于级联编解码后信号能量更强，听起来更清晰易懂，提高了语音的可懂度。

图 11 为未做增强的级联编解码的信号频谱与增强后的级联编解码的信号频谱的对比示意图。如图 11 所示，曲线为未做增强处理的级联编解码的信号频谱，各点为增强后的级联编解码的信号频谱，横坐标为频率，纵坐标为绝对能量，做增强处理后的信号频谱强度增强，可懂度提升。

10 图 12 为未做增强的级联编解码的信号频谱与增强后的级联编解码的信号频谱的中高频部分的对比示意图。曲线为未做增强处理的级联编解码的信号频谱，各点为增强后的级联编解码的信号频谱，横坐标为频率，纵坐标为绝对能量，做增强处理后的信号频谱强度增强，中高频部分做了预增强处理后信号能量更强，提高了可懂度。

15

图 13 为一个实施例中语音信号级联处理装置的结构框图。如图 13 所示，一种语音信号级联处理装置，包括语音信号获取模块 1302、识别模块 1304、第一信号增强模块 1306、第二信号增强模块 1308 和输出模块 1310。其中：

语音信号获取模块 1302 用于获取语音信号。

20 识别模块 1304 用于对该语音信号进行特征识别。

第一信号增强模块 1306 用于若该语音信号为第一特征信号，则采用第一预增强滤波系数对该第一特征信号进行预增强滤波，得到第一预增强语音信号。

25 第二信号增强模块 1308 用于若该语音信号为第二特征信号，则采用第二预增强滤波系数对该第二特征信号进行预增强滤波，得到第二预增强语音信号。

输出模块 1310 用于输出该第一预增强语音信号或第二预增强语音信号，

以根据该第一预增强语音信号或第二预增强语音信号进行级联编解码处理。

上述语音信号级联处理装置，通过对语音信号进行特征识别，对第一特征信号采用第一预增强滤波系数进行预增强滤波处理，对第二特征信号采用第二预增强滤波系数进行预增强滤波处理，将预增强语音进行级联编解码处理，接收方能更清楚听清语音信息，提高了经过级联编解码后的语音信号的可懂度，针对第一特征信号和第二特征信号分别采用对应的滤波系数进行增强滤波处理，针对性更强，滤波处理更加准确。

图 14 为另一个实施例中语音信号级联处理装置的结构框图。如图 14 所示，一种语音信号级联处理装置，除了包括语音信号获取模块 1302、识别模块 1304、第一信号增强模块 1306、第二信号增强模块 1308 和输出模块 1310，还包括训练模块 1312。

训练模块 1312 用于在该获取语音信号之前，根据音频训练集中的训练样本进行离线训练得到第一预增强滤波系数和第二预增强滤波系数。

图 15 为一个实施例中训练模块的内部结构示意图。如图 15 所示，该训练模块 1310 包括选取单元 1502、模拟级联编解码单元 1504、能量补偿值获取单元 1506、平均能量补偿值获取单元 1508 和滤波系数获取单元 1510。

选取单元 1502 用于从音频训练集中获取样本语音信号，该样本语音信号为第一特征样本语音信号或第二特征样本语音信号。

模拟级联编解码单元 1504 用于对该样本语音信号进行模拟级联编解码处理，得到降级语音信号。

能量补偿值获取单元 1506 用于获取该降级语音信号与样本语音信号对应不同频点上的能量衰减值，将该能量衰减值作为频点能量补偿值。

平均能量补偿值获取单元 1508 用于对该音频训练集中的第一特征信号所对应的频点能量补偿值求平均得到第一特征信号在不同频点上的能量平均补偿值，以及对该音频训练集中的第二特征信号所对应的频点能量补偿值求平均得到第二特征信号在不同频点上的能量平均补偿值。

滤波系数获取单元 1510 用于根据该第一特征信号在不同频点上的能量平均补偿值进行滤波拟合得到第一预增强滤波系数，以及根据该第二特征信号在不同频点上的能量平均补偿值进行滤波拟合得到第二预增强滤波系数。

上述离线训练得到第一预增强滤波系数和第二预增强滤波系数，通过离线训练可以准确得到第一预增强滤波系数和第二预增强滤波系数，方便后续进行在线滤波处理得到增强后的语音信号，有效提高级联编解码后的语音信号的可懂度。

在一个实施例中，识别模块 1304 还用于获取该语音信号的基音周期；以及判断该语音信号的基音周期是否大于预设周期值，若是，则该语音信号为第一特征信号，若否，则该语音信号为第二特征信号。

进一步的，识别模块 1304 还用于对该语音信号以矩形窗进行平移分帧，每帧窗长第一采样点数，每帧平移第二采样点数；对每帧信号进行三电平削波处理；对每帧内采样点计算自相关值；以及以每帧中自相关值最大者所对应的序号作为每帧的基音周期。

进一步的，识别模块 1304 还用于在该对该语音信号以矩形窗进行平移分帧，每帧窗长第一采样点数，每帧平移第二采样点数之前，对该语音信号进行带通滤波，以及将该带通滤波后的语音信号进行预加重处理。

图 16 为另一个实施例中语音信号级联处理装置的结构框图。如图 16 所示，一种语音信号级联处理装置，除了包括语音信号获取模块 1302、识别模块 1304、第一信号增强模块 1306、第二信号增强模块 1308 和输出模块 1310，还包括原始信号获取模块 1314、检测模块 1316、滤波模块 1318。

原始信号获取模块 1314 用于获取输入的原始音频信号。

检测模块 1316 用于检测该原始音频信号为语音信号或非语音信号。

该语音信号获取模块 1302 还用于若该原始音频信号为语音信号，获取语音信号。

滤波模块 1318 用于若该原始音频信号为非语音信号，则对该非语音信号

进行高通滤波处理。

上述语音信号级联处理装置，对非语音进行高通滤波处理，降低信号的噪声，通过对语音信号进行特征识别，对第一特征信号采用第一预增强滤波系数进行预增强滤波处理，对第二特征信号采用第二预增强滤波系数进行预增强滤波处理，将预增强语音进行级联编解码处理，接收方能更清楚听清语音信息，提高了经过级联编解码后的语音信号的可懂度，针对第一特征信号和

5 第二特征信号分别采用对应的滤波系数进行增强滤波处理，针对性更强，滤波处理更加准确。

在其他实施例中，一种语音信号级联处理装置，可包括语音信号获取模块 1302、识别模块 1304、第一信号增强模块 1306、第二信号增强模块 1308、输出模块 1310、训练模块 1312、原始信号获取模块 1314、检测模块 1316、滤波模块 1318 中所有可能的组合。

10

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，所述的程序可存储于一非易失性计算机可读取存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，所述的存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory，ROM）等。

15

以上所述实施例仅表达了本发明的几种实施方式，其描述较为具体和详细，但并不能因此而理解为对本发明专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本发明构思的前提下，还可以做出若干变形和改进，这些都属于本发明的保护范围。因此，本发明的保护范围应以所附权利要求为准。

权利要求书

1、一种语音信号级联处理方法，包括：

获取语音信号；

对所述语音信号进行特征识别；

若所述语音信号为第一特征信号，则采用第一预增强滤波系数对所述第
5 一特征信号进行预增强滤波，得到第一预增强语音信号；

若所述语音信号为第二特征信号，则采用第二预增强滤波系数对所述第
二特征信号进行预增强滤波，得到第二预增强语音信号；以及

输出所述第一预增强语音信号或第二预增强语音信号，以根据所述第一
预增强语音信号或第二预增强语音信号进行级联编解码处理。

10 2、根据权利要求 1 所述的方法，其特征在于，在所述获取语音信号之前，
所述方法还包括：

根据音频训练集中的训练样本进行离线训练得到第一预增强滤波系数和
第二预增强滤波系数，包括：

从音频训练集中获取样本语音信号，所述样本语音信号为第一特征样本
15 语音信号或第二特征样本语音信号；

对所述样本语音信号进行模拟级联编解码处理，得到降级语音信号；

获取所述降级语音信号与样本语音信号对应不同频点上的能量衰减值，
将所述能量衰减值作为频点能量补偿值；

对所述音频训练集中的第一特征信号所对应的频点能量补偿值求平均得
20 到第一特征信号在不同频点上的能量平均补偿值，以及对所述音频训练集中
的第二特征信号所对应的频点能量补偿值求平均得到第二特征信号在不同频
点上的能量平均补偿值；以及

根据所述第一特征信号在不同频点上的能量平均补偿值进行滤波拟合得
到第一预增强滤波系数，以及根据所述第二特征信号在不同频点上的能量平
25 均补偿值进行滤波拟合得到第二预增强滤波系数。

3、根据权利要求1所述的方法，其特征在于，所述对所述语音信号进行特征识别包括：

获取所述语音信号的基音周期；

判断所述语音信号的基音周期是否大于预设周期值，若是，则所述语音信号为第一特征信号，若否，则所述语音信号为第二特征信号。

4、根据权利要求3所述的方法，其特征在于，所述获取所述语音信号的基音周期包括：

对所述语音信号以矩形窗进行平移分帧，每帧窗长第一采样点数，每帧平移第二采样点数；

对每帧信号进行三电平削波处理；

对每帧内采样点计算自相关值；以及

以每帧中自相关值最大者所对应的序号作为每帧的基音周期。

5、根据权利要求4所述的方法，其特征在于，在所述对所述语音信号以矩形窗进行平移分帧，每帧窗长第一采样点数，每帧平移第二采样点数之前，所述获取所述语音信号的基音周期还包括：

对所述语音信号进行带通滤波；

将所述带通滤波后的语音信号进行预加重处理。

6、根据权利要求1所述的方法，其特征在于，在所述获取语音信号的步骤之前，所述方法还包括：

获取输入的原始音频信号；

检测所述原始音频信号为语音信号或非语音信号；

若所述原始音频信号为语音信号，则执行所述获取语音信号的步骤；

若所述原始音频信号为非语音信号，则对所述非语音信号进行高通滤波处理。

7、一种终端，包括存储器及处理器，所述存储器中储存有计算机可读指令，所述指令被所述处理器执行时，使得所述处理器执行以下步骤：

获取语音信号；

对所述语音信号进行特征识别；

若所述语音信号为第一特征信号，则采用第一预增强滤波系数对所述第一特征信号进行预增强滤波，得到第一预增强语音信号；

若所述语音信号为第二特征信号，则采用第二预增强滤波系数对所述第二特征信号进行预增强滤波，得到第二预增强语音信号；以及

输出所述第一预增强语音信号或第二预增强语音信号，以根据所述第一预增强语音信号或第二预增强语音信号进行级联编解码处理。

8、根据权利要求7所述的终端，其特征在于，在所述获取语音信号之前，所述处理器还用于执行以下步骤：

10 根据音频训练集中的训练样本进行离线训练得到第一预增强滤波系数和第二预增强滤波系数，包括：

从音频训练集中获取样本语音信号，所述样本语音信号为第一特征样本语音信号或第二特征样本语音信号；

对所述样本语音信号进行模拟级联编解码处理，得到降级语音信号；

15 获取所述降级语音信号与样本语音信号对应不同频点上的能量衰减值，将所述能量衰减值作为频点能量补偿值；

对所述音频训练集中的第一特征信号所对应的频点能量补偿值求平均得到第一特征信号在不同频点上的能量平均补偿值，以及对所述音频训练集中的第二特征信号所对应的频点能量补偿值求平均得到第二特征信号在不同频点上的能量平均补偿值；以及

20 根据所述第一特征信号在不同频点上的能量平均补偿值进行滤波拟合得到第一预增强滤波系数，以及根据所述第二特征信号在不同频点上的能量平均补偿值进行滤波拟合得到第二预增强滤波系数。

25 9、根据权利要求7所述的终端，其特征在于，所述对所述语音信号进行特征识别包括：

获取所述语音信号的基音周期；

判断所述语音信号的基音周期是否大于预设周期值，若是，则所述语音

信号为第一特征信号，若否，则所述语音信号为第二特征信号。

10、根据权利要求 9 所述的终端，其特征在于，所述获取所述语音信号的基音周期包括：

5 对所述语音信号以矩形窗进行平移分帧，每帧窗长第一采样点数，每帧
平移第二采样点数；

对每帧信号进行三电平削波处理；

对每帧内采样点计算自相关值；以及

以每帧中自相关值最大者所对应的序号作为每帧的基音周期。

11、根据权利要求 10 所述的终端，其特征在于，在所述对所述语音信号
10 以矩形窗进行平移分帧，每帧窗长第一采样点数，每帧平移第二采样点数之
前，所述获取所述语音信号的基音周期还包括：

对所述语音信号进行带通滤波；

将所述带通滤波后的语音信号进行预加重处理。

12、根据权利要求 7 所述的终端，其特征在于，在所述获取语音信号的
15 步骤之前，所述处理器还用于执行以下步骤：

获取输入的原始音频信号；

检测所述原始音频信号为语音信号或非语音信号；

若所述原始音频信号为语音信号，则执行所述获取语音信号的步骤；

20 若所述原始音频信号为非语音信号，则对所述非语音信号进行高通滤波
处理。

13、一个或多个包含计算机可执行指令的非易失性计算机可读存储介质，
当所述计算机可执行指令被一个或多个处理器执行时，使得所述处理器执行
以下步骤：

获取语音信号；

25 对所述语音信号进行特征识别；

若所述语音信号为第一特征信号，则采用第一预增强滤波系数对所述第
一特征信号进行预增强滤波，得到第一预增强语音信号；

若所述语音信号为第二特征信号，则采用第二预增强滤波系数对所述第二特征信号进行预增强滤波，得到第二预增强语音信号；以及

输出所述第一预增强语音信号或第二预增强语音信号，以根据所述第一预增强语音信号或第二预增强语音信号进行级联编解码处理。

5 14、根据权利要求 13 所述的非易失性计算机可读存储介质，其特征在于，在所述获取语音信号之前，所述处理器还用于执行以下步骤：

根据音频训练集中的训练样本进行离线训练得到第一预增强滤波系数和第二预增强滤波系数，包括：

10 从音频训练集中获取样本语音信号，所述样本语音信号为第一特征样本语音信号或第二特征样本语音信号；

对所述样本语音信号进行模拟级联编解码处理，得到降级语音信号；

获取所述降级语音信号与样本语音信号对应不同频点上的能量衰减值，将所述能量衰减值作为频点能量补偿值；

15 对所述音频训练集中的第一特征信号所对应的频点能量补偿值求平均得到第一特征信号在不同频点上的能量平均补偿值，以及对所述音频训练集中的第二特征信号所对应的频点能量补偿值求平均得到第二特征信号在不同频点上的能量平均补偿值；以及

20 根据所述第一特征信号在不同频点上的能量平均补偿值进行滤波拟合得到第一预增强滤波系数，以及根据所述第二特征信号在不同频点上的能量平均补偿值进行滤波拟合得到第二预增强滤波系数。

15、根据权利要求 13 所述的非易失性计算机可读存储介质，其特征在于，所述对所述语音信号进行特征识别包括：

获取所述语音信号的基音周期；

25 判断所述语音信号的基音周期是否大于预设周期值，若是，则所述语音信号为第一特征信号，若否，则所述语音信号为第二特征信号。

16、根据权利要求 15 所述的非易失性计算机可读存储介质，其特征在于，所述获取所述语音信号的基音周期包括：

对所述语音信号以矩形窗进行平移分帧，每帧窗长第一采样点数，每帧平移第二采样点数；

对每帧信号进行三电平削波处理；

对每帧内采样点计算自相关值；以及

5 以每帧中自相关值最大者所对应的序号作为每帧的基音周期。

17、根据权利要求 16 所述的非易失性计算机可读存储介质，其特征在于，在所述对所述语音信号以矩形窗进行平移分帧，每帧窗长第一采样点数，每帧平移第二采样点数之前，所述获取所述语音信号的基音周期还包括：

对所述语音信号进行带通滤波；

10 将所述带通滤波后的语音信号进行预加重处理。

18、根据权利要求 13 所述的非易失性计算机可读存储介质，其特征在于，在所述获取语音信号的步骤之前，所述处理器还用于执行以下步骤：

获取输入的原始音频信号；

检测所述原始音频信号为语音信号或非语音信号；

15 若所述原始音频信号为语音信号，则执行所述获取语音信号的步骤；

若所述原始音频信号为非语音信号，则对所述非语音信号进行高通滤波处理。

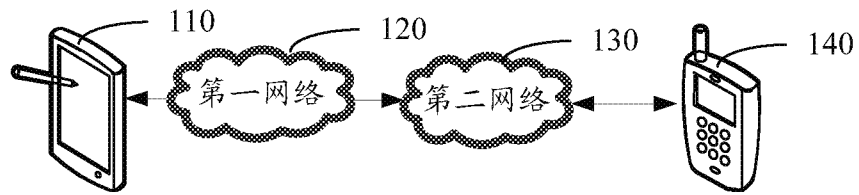


图 1

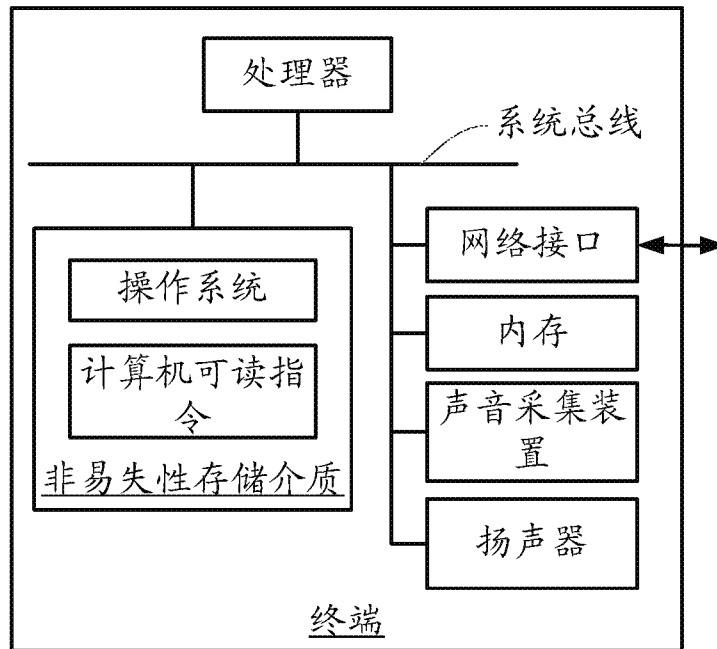


图 2

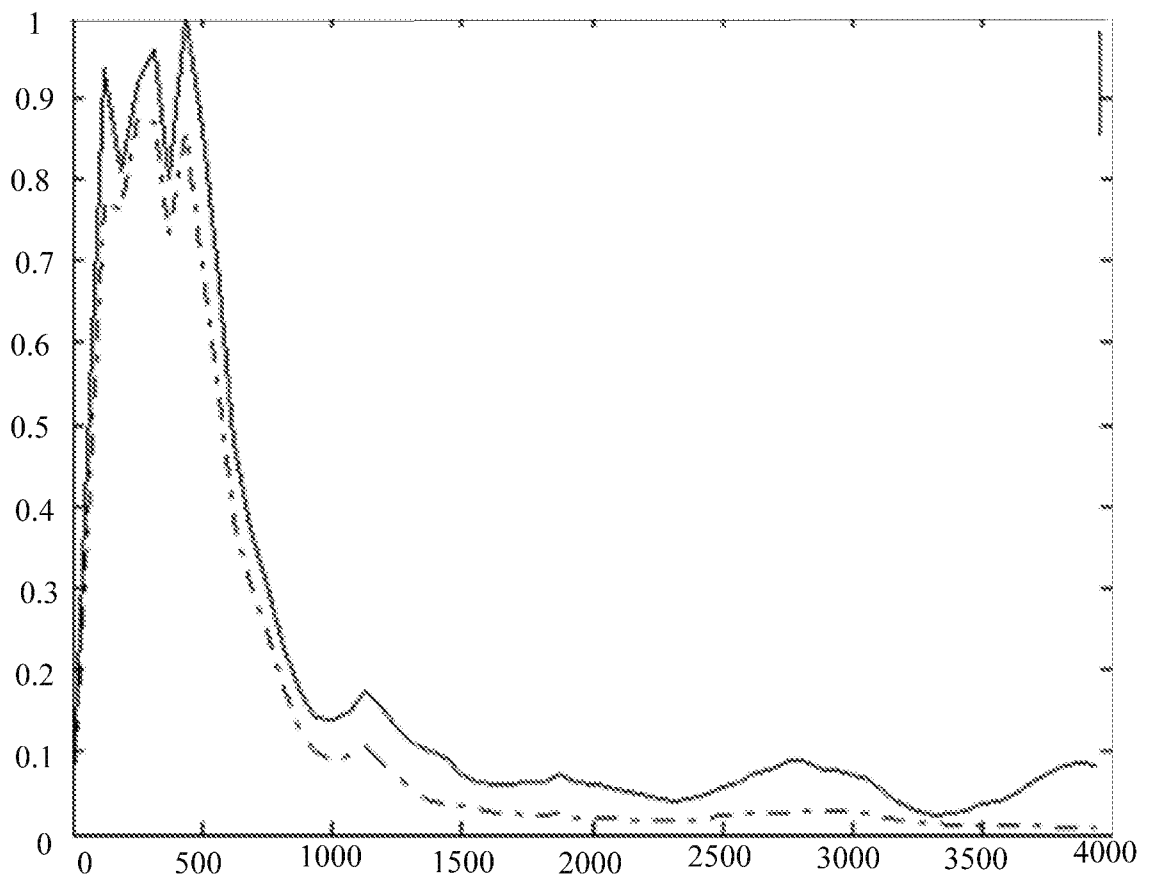


图 3A

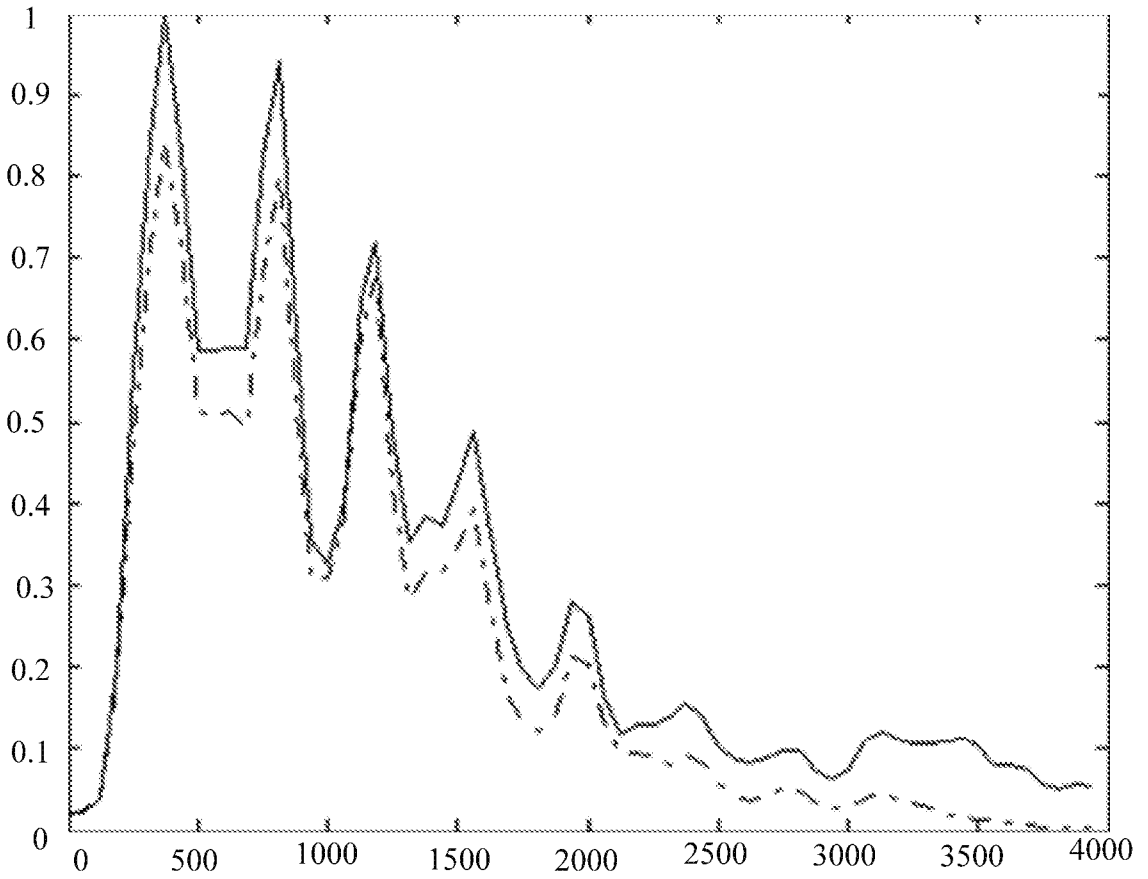


图 3B

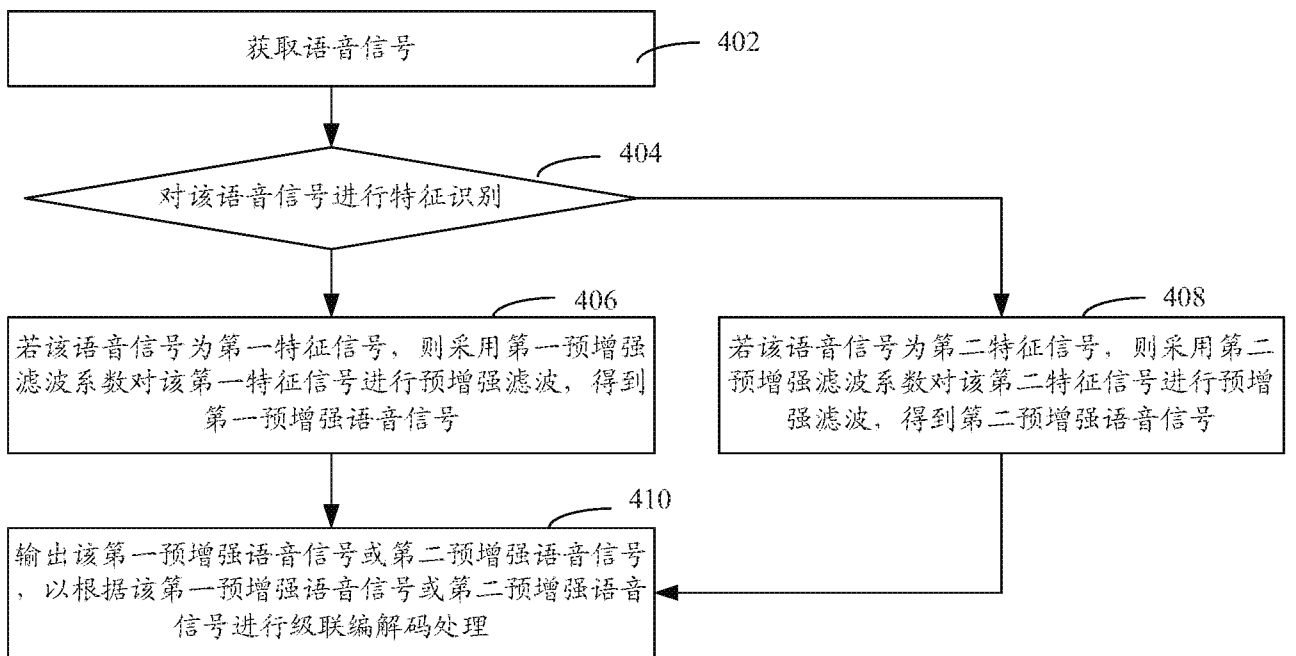


图 4

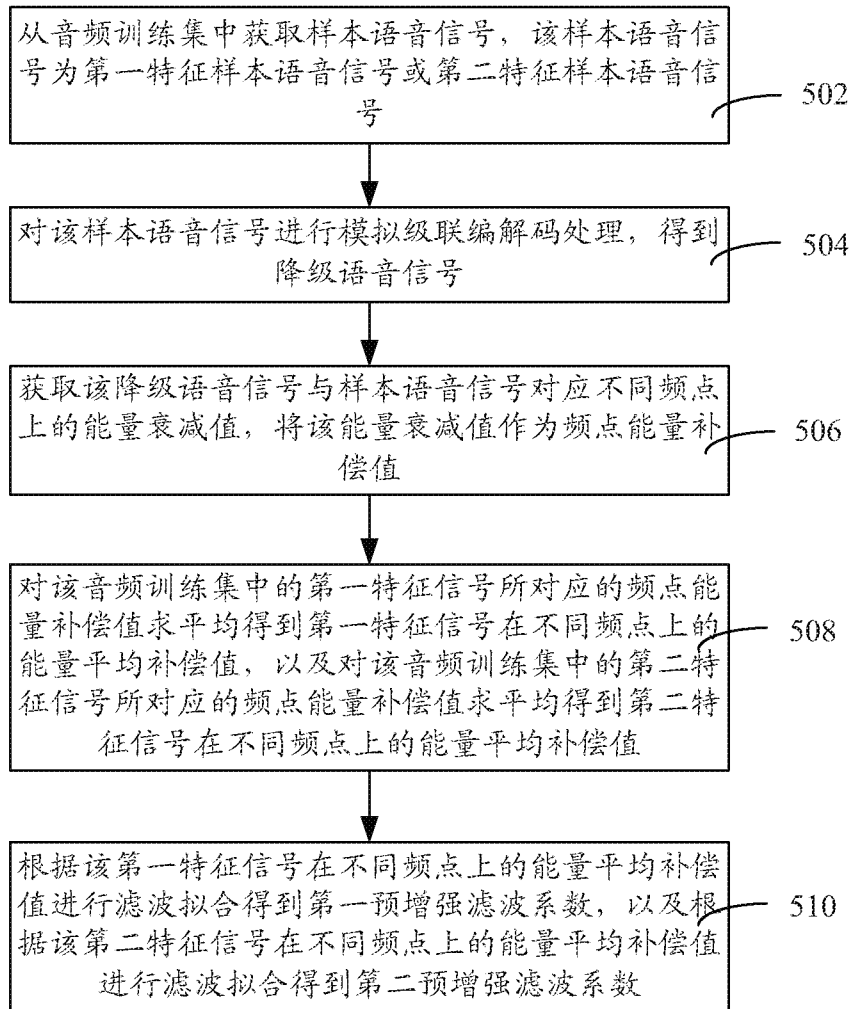


图 5

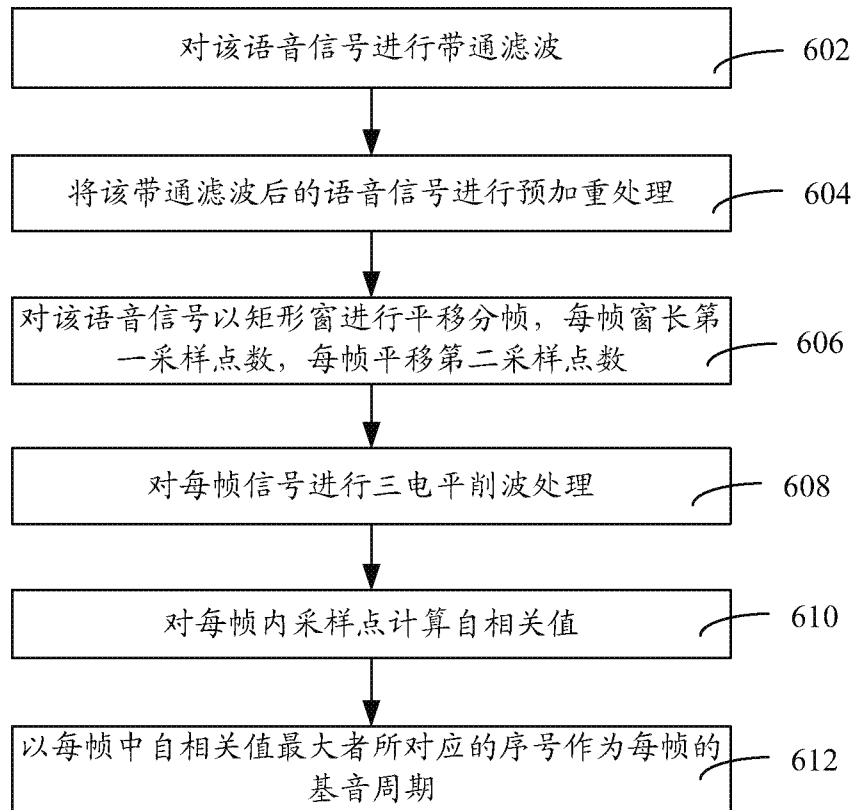


图 6

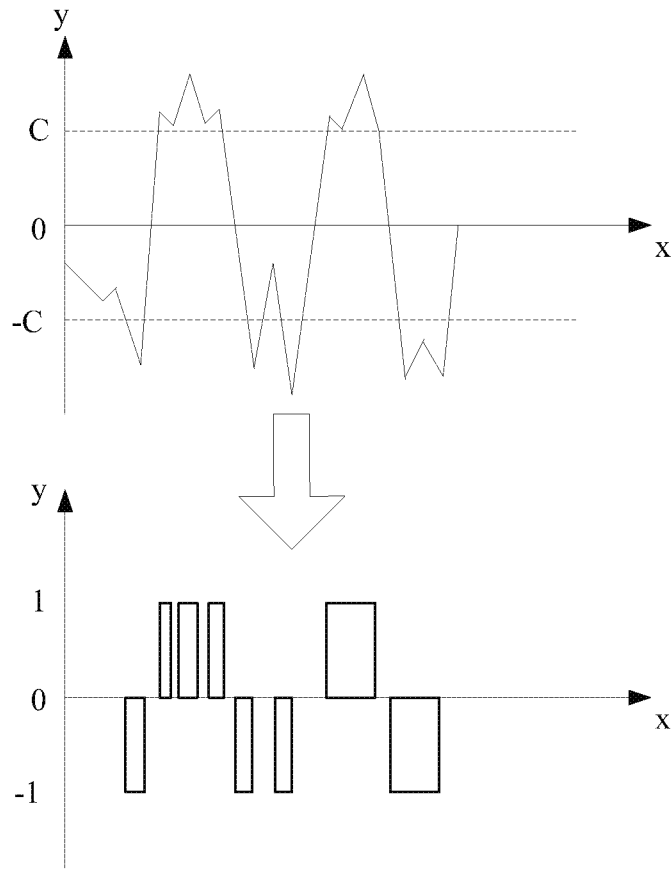


图 7

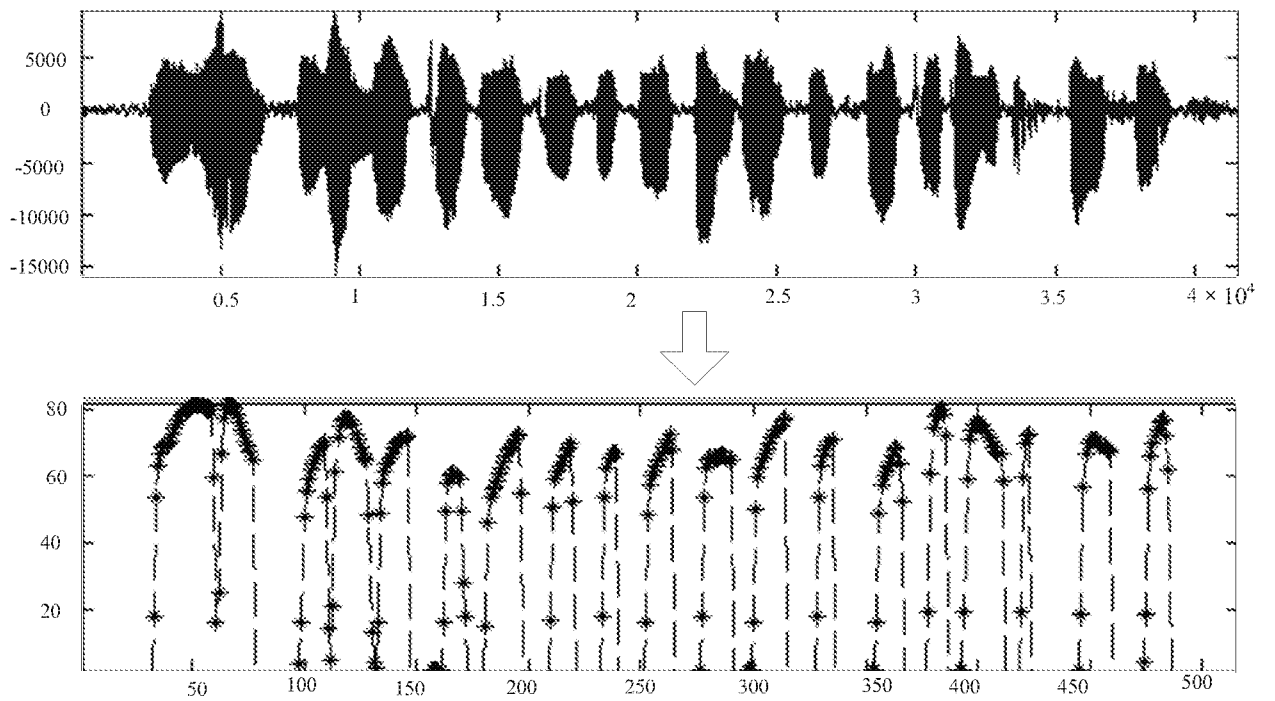


图 8

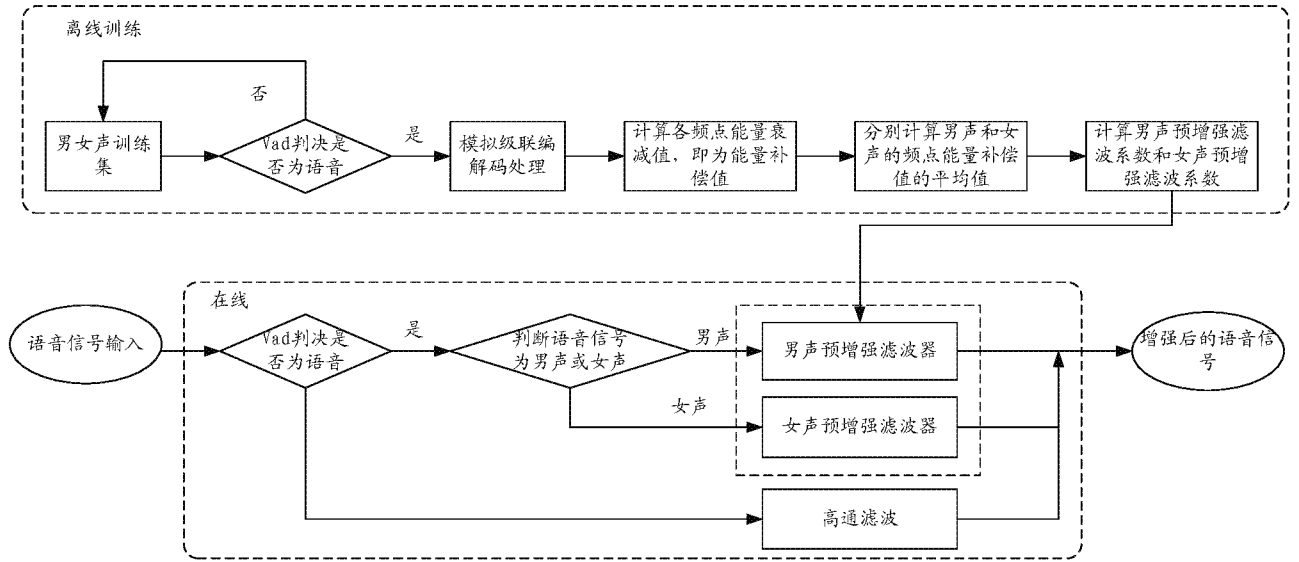


图 9

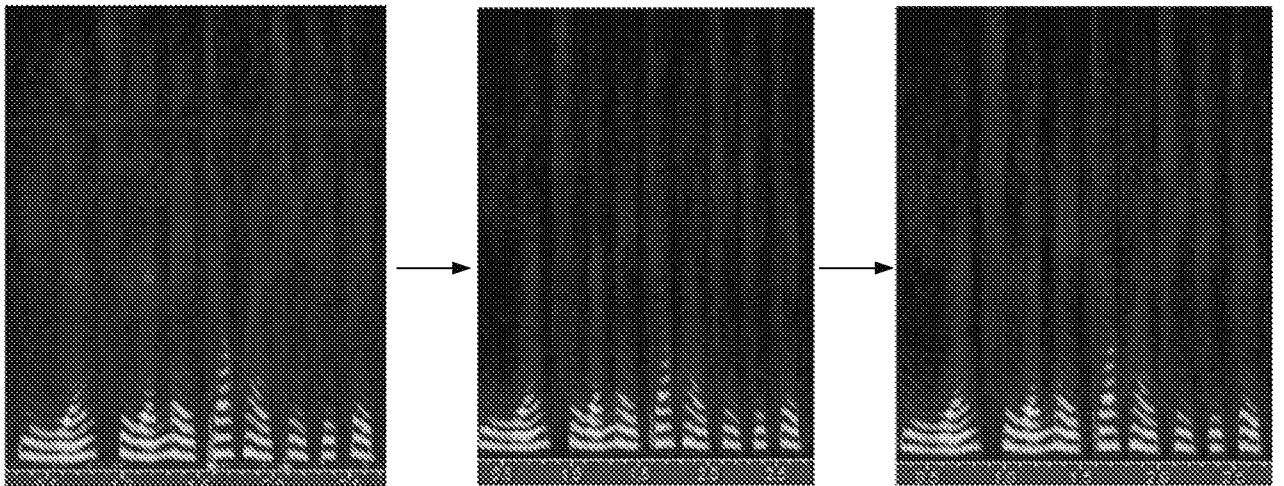


图 10

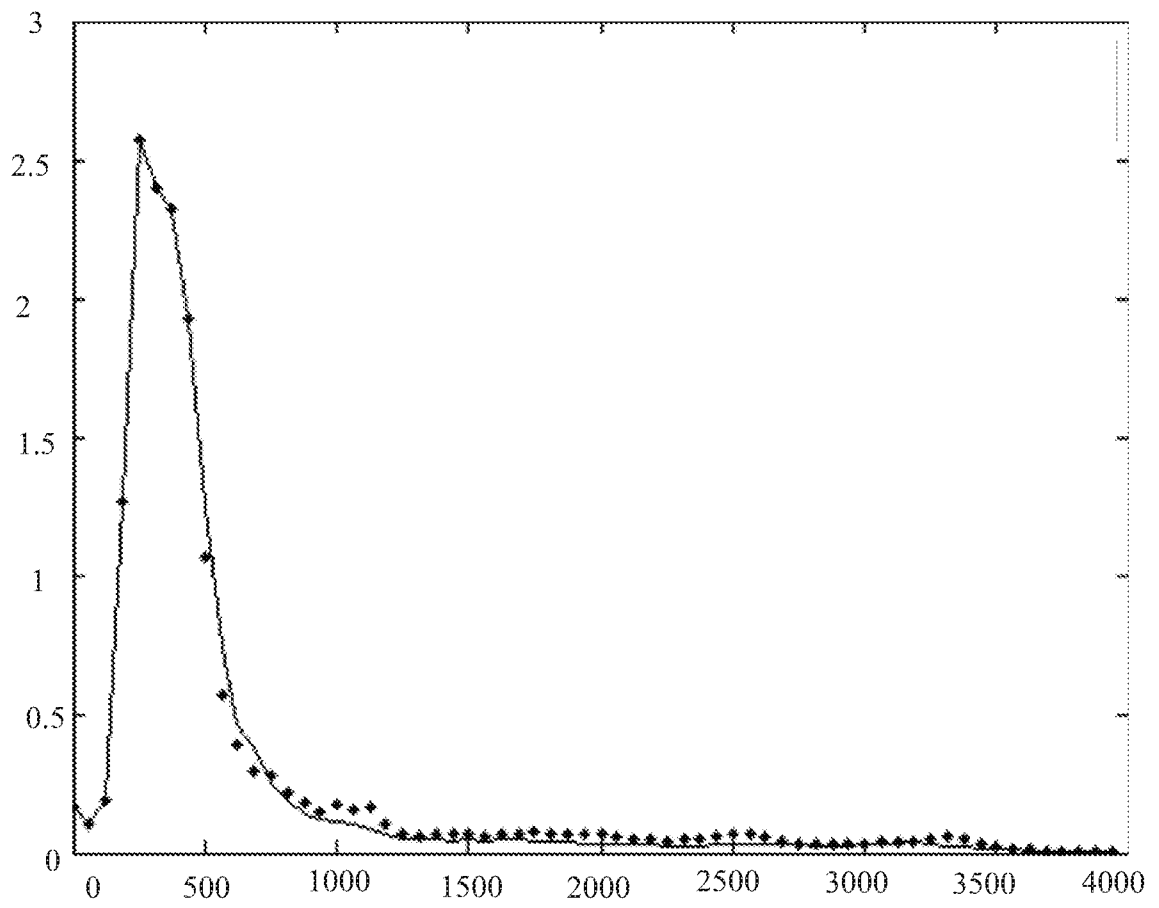


图 11

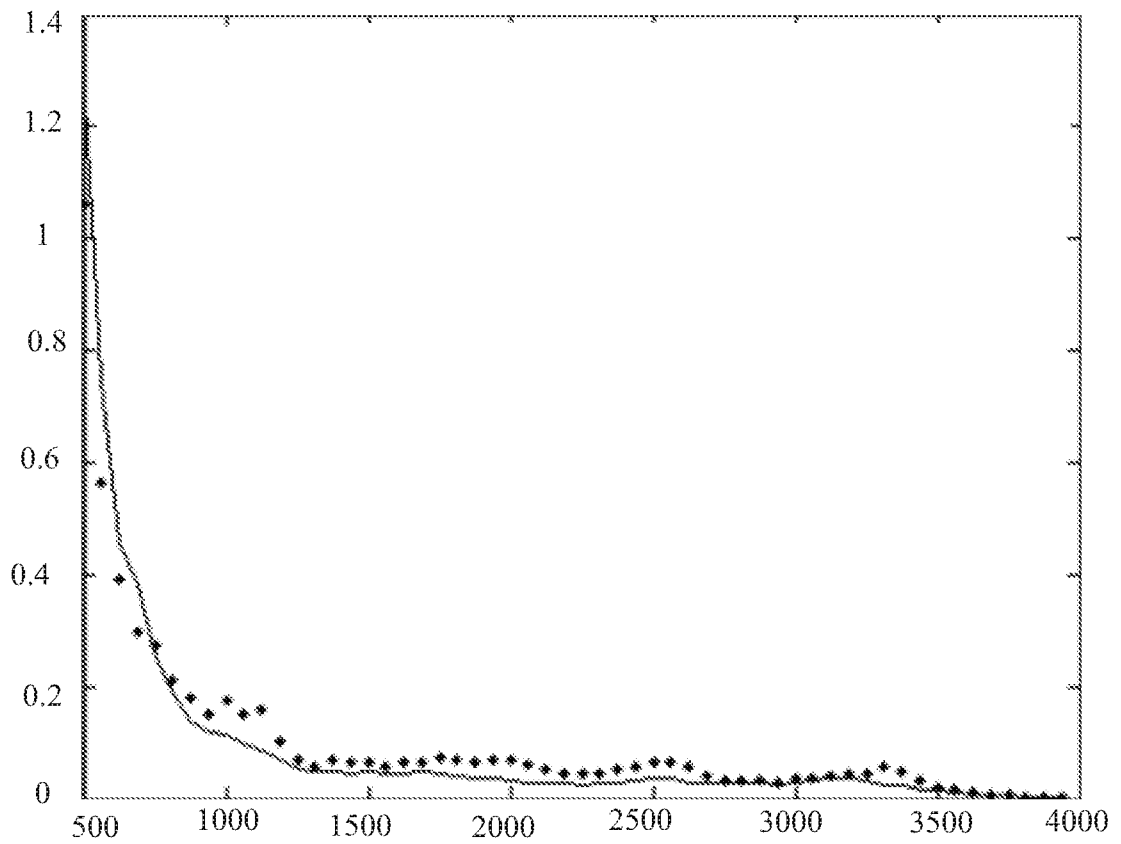


图 12

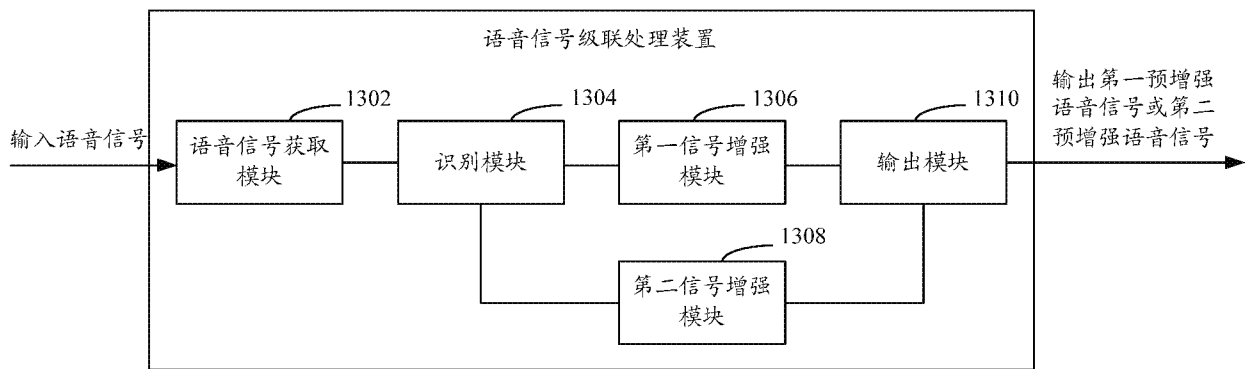


图 13

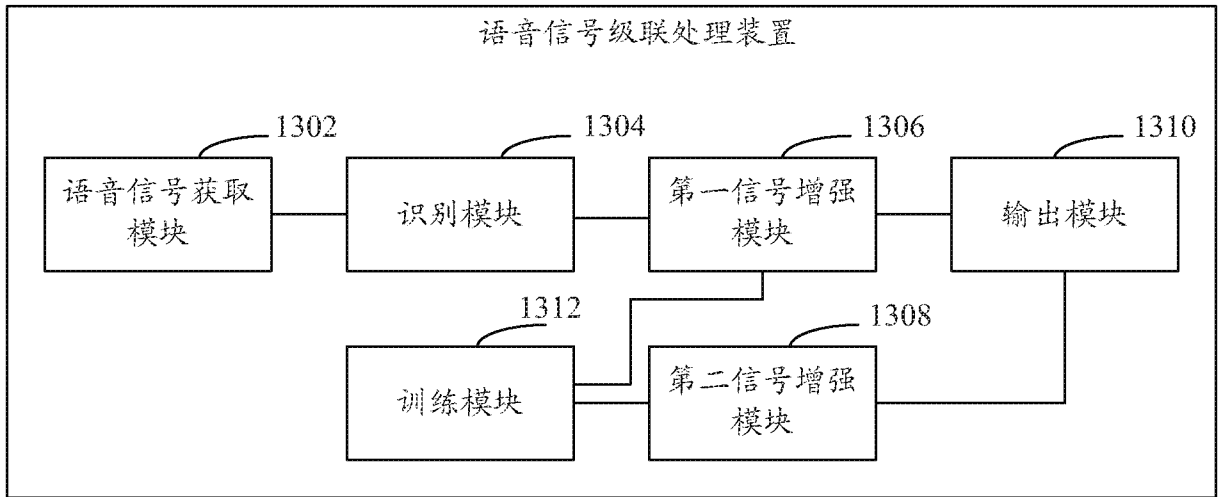


图 14

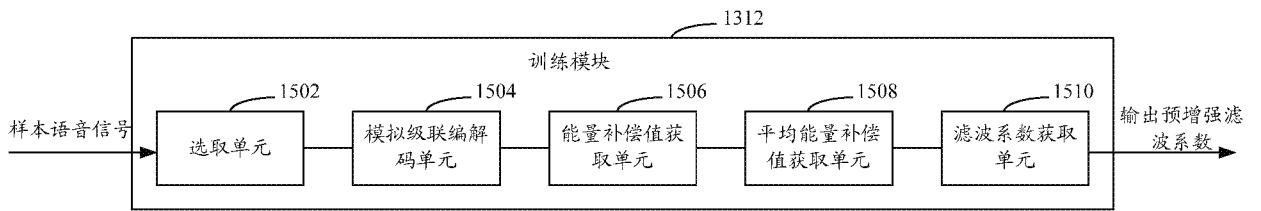


图 15

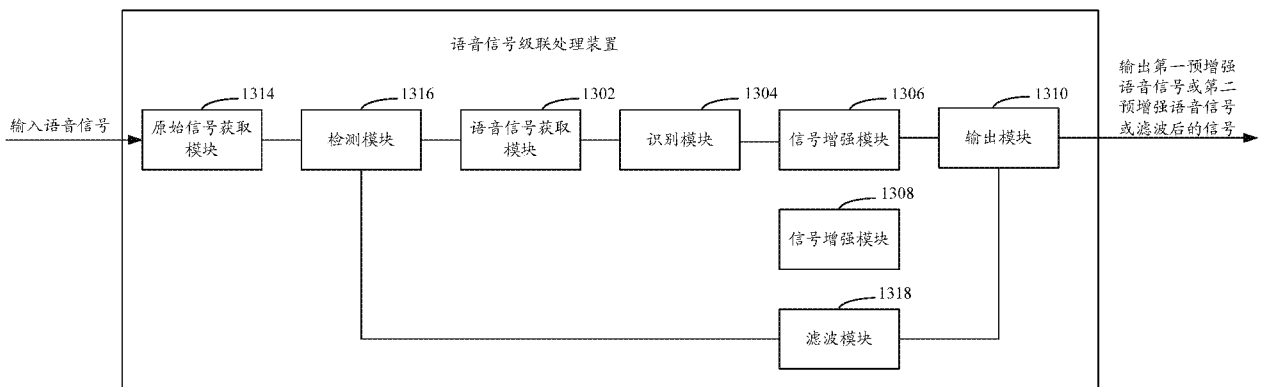


图 16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2017/076653

A. CLASSIFICATION OF SUBJECT MATTER

G10L 21/0232 (2013.01) i; G10L 21/0324 (2013.01) i; G10L 25/21 (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 21/-; G10L 25/-

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)

CNABS, CNTXT, CNKI, WPI, EPODOC: TENCENT, code/encode, fundamental tone, speech, cod+, decod+, encod+, enhanc+, filt+, coefficient+, character+, pitch, energy

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 105913854 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.), 31 August 2016 (31.08.2016), claims 1-12, and description, paragraph [0153]	1-18
X	CN 102779527 A (WUXI CHENGDIAN KEDA TECHNOLOGY DEVELOPMENT CO., LTD.), 14 November 2012 (14.11.2012), description, paragraphs [0004] and [0007]-[0010], and figures 1-3	1, 3-7, 9-13, 15-18
A	CN 103413553 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.), 27 November 2013 (27.11.2013), the whole document	1-18
A	CN 104269177 A (LENOVO (BEIJING) CO., LTD.), 07 January 2015 (07.01.2015), the whole document	1-18
A	CN 1285945 A (ERICSSON INC.), 28 February 2001 (28.02.2001), the whole document	1-18
A	EP 0929065 A2 (AT & T CORP.), 14 July 1999 (14.07.1999), the whole document	1-18

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>
---	---

Date of the actual completion of the international search
10 May 2017 (10.05.2017)

Date of mailing of the international search report
31 May 2017 (31.05.2017)

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
WANG, Xinning
Telephone No.: (86-10) **62413706**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2017/076653

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2004097799 A1 (MASSACHUSETTS INSTITUTE OF TECHNOLOGY), 11 November 2004 (11.11.2004), the whole document	1-18

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/076653

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105913854 A	31 August 2016	None	
CN 102779527 A	14 November 2012	CN 102779527 B	28 May 2014
CN 103413553 A	27 November 2013	US 2015127356 A1	07 May 2015
		CN 103413553 B	09 March 2016
		WO 2015024428 A1	26 February 2015
CN 104269177 A	07 January 2015	None	
CN 1285945 A	28 February 2001	EE 04070 B1	16 June 2003
		EE 200000414 A	17 December 2001
		US 6070137 A	30 May 2000
		BR 9813246 A	03 October 2000
		WO 9935638 A1	15 July 1999
		EP 1046153 B1	17 July 2002
		AU 1622699 A	26 July 1999
		EP 1046153 A1	25 October 2000
		DE 69806645 E	22 August 2002
EP 0929065 A2	14 July 1999	None	
WO 2004097799 A1	11 November 2004	US 2004252850 A1	16 December 2004
		US 7787640 B2	31 August 2010
		EP 1618559 A1	25 January 2006

<p>A. 主题的分类</p> <p>G10L 21/0232(2013.01)i; G10L 21/0324(2013.01)i; G10L 25/21(2013.01)i</p> <p>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</p>																							
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>G10L 21/-; G10L25/-</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CNABS, CNTXT, CNKI, WPI, EPODOC: 腾讯, 语音, 编码, 解码, 编解码, 增强, 滤波, 系数, 特征, 基音, 能量, speech, cod+, decod+, encod+, enhanc+, filt+, coefficient+, character+, pitch, energy</p>																							
<p>C. 相关文件</p> <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>PX</td> <td>CN 105913854 A (腾讯科技深圳有限公司) 2016年 8月 31日 (2016 - 08 - 31) 权利要求1-12、说明书第[0153]段</td> <td>1-18</td> </tr> <tr> <td>X</td> <td>CN 102779527 A (无锡成电科大科技发展有限公司) 2012年 11月 14日 (2012 - 11 - 14) 说明书第[0004], [0007]-[0010]段、附图1-3</td> <td>1, 3-7, 9-13, 15-18</td> </tr> <tr> <td>A</td> <td>CN 103413553 A (腾讯科技深圳有限公司) 2013年 11月 27日 (2013 - 11 - 27) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>CN 104269177 A (联想北京有限公司) 2015年 1月 7日 (2015 - 01 - 07) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>CN 1285945 A (艾利森公司) 2001年 2月 28日 (2001 - 02 - 28) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>EP 0929065 A2 (AT & T CORP.) 1999年 7月 14日 (1999 - 07 - 14) 全文</td> <td>1-18</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	PX	CN 105913854 A (腾讯科技深圳有限公司) 2016年 8月 31日 (2016 - 08 - 31) 权利要求1-12、说明书第[0153]段	1-18	X	CN 102779527 A (无锡成电科大科技发展有限公司) 2012年 11月 14日 (2012 - 11 - 14) 说明书第[0004], [0007]-[0010]段、附图1-3	1, 3-7, 9-13, 15-18	A	CN 103413553 A (腾讯科技深圳有限公司) 2013年 11月 27日 (2013 - 11 - 27) 全文	1-18	A	CN 104269177 A (联想北京有限公司) 2015年 1月 7日 (2015 - 01 - 07) 全文	1-18	A	CN 1285945 A (艾利森公司) 2001年 2月 28日 (2001 - 02 - 28) 全文	1-18	A	EP 0929065 A2 (AT & T CORP.) 1999年 7月 14日 (1999 - 07 - 14) 全文	1-18
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
PX	CN 105913854 A (腾讯科技深圳有限公司) 2016年 8月 31日 (2016 - 08 - 31) 权利要求1-12、说明书第[0153]段	1-18																					
X	CN 102779527 A (无锡成电科大科技发展有限公司) 2012年 11月 14日 (2012 - 11 - 14) 说明书第[0004], [0007]-[0010]段、附图1-3	1, 3-7, 9-13, 15-18																					
A	CN 103413553 A (腾讯科技深圳有限公司) 2013年 11月 27日 (2013 - 11 - 27) 全文	1-18																					
A	CN 104269177 A (联想北京有限公司) 2015年 1月 7日 (2015 - 01 - 07) 全文	1-18																					
A	CN 1285945 A (艾利森公司) 2001年 2月 28日 (2001 - 02 - 28) 全文	1-18																					
A	EP 0929065 A2 (AT & T CORP.) 1999年 7月 14日 (1999 - 07 - 14) 全文	1-18																					
<p><input checked="" type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p> <p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>																							
<p>国际检索实际完成的日期</p> <p>2017年 5月 10日</p>		<p>国际检索报告邮寄日期</p> <p>2017年 5月 31日</p>																					
<p>ISA/CN的名称和邮寄地址</p> <p>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</p> <p>传真号 (86-10)62019451</p>		<p>授权官员</p> <p>王馨宁</p> <p>电话号码 (86-10)62413706</p>																					

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	WO 2004097799 A1 (MASSACHUSETTS INSTITUTE OF TECHNOLOGY) 2004年 11月 11日 (2004 - 11 - 11) 全文	1-18

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/076653

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	105913854	A	2016年 8月 31日	无			
CN	102779527	A	2012年 11月 14日	CN	102779527	B	2014年 5月 28日
CN	103413553	A	2013年 11月 27日	US	2015127356	A1	2015年 5月 7日
				CN	103413553	B	2016年 3月 9日
				WO	2015024428	A1	2015年 2月 26日
CN	104269177	A	2015年 1月 7日	无			
CN	1285945	A	2001年 2月 28日	EE	04070	B1	2003年 6月 16日
				EE	200000414	A	2001年 12月 17日
				US	6070137	A	2000年 5月 30日
				BR	9813246	A	2000年 10月 3日
				WO	9935638	A1	1999年 7月 15日
				EP	1046153	B1	2002年 7月 17日
				AU	1622699	A	1999年 7月 26日
				EP	1046153	A1	2000年 10月 25日
				DE	69806645	E	2002年 8月 22日
EP	0929065	A2	1999年 7月 14日	无			
WO	2004097799	A1	2004年 11月 11日	US	2004252850	A1	2004年 12月 16日
				US	7787640	B2	2010年 8月 31日
				EP	1618559	A1	2006年 1月 25日

表 PCT/ISA/210 (同族专利附件) (2009年7月)