

(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
2015 年 12 月 17 日 (17.12.2015)



(10) 国际公布号
WO 2015/188627 A1

- (51) 国际专利分类号:
G10L 19/022 (2013.01) *G10L 19/20* (2013.01)
G10L 19/135 (2013.01) *G10L 25/45* (2013.01)
- (21) 国际申请号: PCT/CN2015/071727
- (22) 国际申请日: 2015 年 1 月 28 日 (28.01.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201410260730.5 2014 年 6 月 12 日 (12.06.2014) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 刘泽新 (LIU, Zexin); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 苗磊 (MIAO, Lei); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (74) 代理人: 北京同立钧成知识产权代理有限公司 (LEADER PATENT & TRADEMARK FIRM); 中国

北京市海淀区西直门北大街 32 号枫蓝国际 A 座 8F-6, Beijing 100082 (CN)。

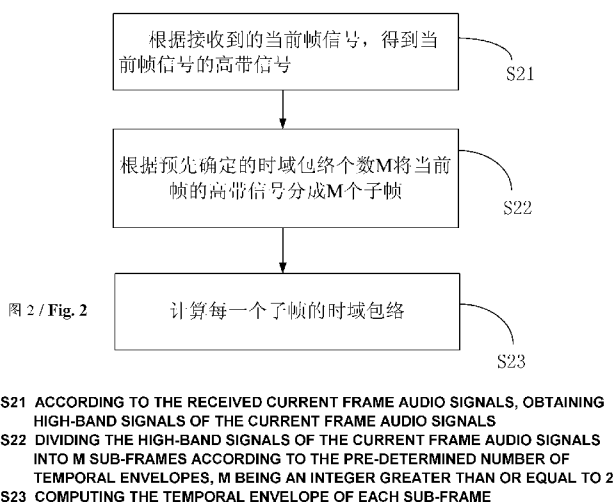
(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,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

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, 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, DEVICE AND ENCODER OF PROCESSING TEMPORAL ENVELOPE OF AUDIO SIGNAL

(54) 发明名称: 一种音频信号的时域包络处理方法及装置、编码器



(57) Abstract: A method, device and encoder of processing temporal envelopes of audio signals, capable of better maintaining the continuity of signal energy when solving a plurality of temporal envelopes, while reducing complexity of computing temporal envelopes. The method comprises: according to the received current frame audio signals, obtaining high-band signals of the current frame audio signals (S21); dividing the high-band signals of the current frame audio signals into M sub-frames according to the pre-determined number of temporal envelopes, M being an integer greater than or equal to 2 (S22); computing the temporal envelope of each sub-frame (S23); windowing the sub-frames at the most significant end in M sub-frames and of the frames at the least significant end in M sub-frames by using asymmetrical windows; windowing sub-frames other than the sub-frames at the most significant end sub-frame and at the least significant end in M sub-frames.

(57) 摘要:

[见续页]



WO 2015/188627 A1

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种音频信号的时域包络处理方法及装置、编码器，在求解多个时域包络能够很好的保持信号能量的连续，同时降低了计算时域包络的复杂度。该方法包括：根据接收到的当前帧音频信号，得到当前帧音频信号的高带信号（S21）；根据预先确定的时域包络个数 M 将当前帧音频信号的高带信号分成 M 个子帧，其中， M 为大于等于 2 的整数（S22）；计算每一个子帧的时域包络（S23）；采用非对称窗对 M 个子帧中的最前端的子帧和 M 个子帧中的最末端的子帧进行加窗；对 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧进行加窗。

一种音频信号的时域包络处理方法及装置、编码器

技术领域

- 5 本发明实施例涉及通信技术领域，尤其涉及一种音频信号的时域包络处理方法及装置、编码器。

背景技术

- 随着语音频压缩技术的高速发展，各种语音频编码算法也相继出现。在
10 语音频编码算法的处理过程中，需要计算时域包络，现有的计算并量化时域包络的过程为：根据事先设定好的计算时域包络的个数 M ， M 为正整数，将预处理后的原始高带信号和预测的高带信号分别分成 M 个子帧，对子帧进行加窗，然后计算各个子帧内预处理后的原始高带信号和预测的高带信号的能量或幅度比。其中，事先设定好的计算时域包络的个数 M 是根据前向缓存
15 (lookahead buffer) 的长度来确定。前向缓存是当前帧为了计算一些参数的需要，将输入信号的最后某些样点缓存不用，在下一帧计算参数时使用，当前帧使用的是前一帧缓存的样点。缓存的这些样点即为前向缓存，缓存的样点的个数即为前向缓存的长度。

- 上述对时域包络的处理过程存在的问题是：在求解时域包络时，利用
20 的都是对称窗，同时为了保证子帧间和帧间的混叠，根据前向缓存 (lookahead) 的长度计算了多个时域包络。但在计算时域包络时，如果信号的时域分辨率太高，会造成帧内能量的不连续，从而引入很差的听觉感受。

25 发明内容

本发明实施例提供一种音频信号的时域包络处理方法及装置、编码器，可解决在计算时域包络时造成的帧内能量的不连续的问题。

第一方面，本发明实施例提供一种音频信号的时域包络处理方法，包括：

根据接收到的当前帧信号，得到所述当前帧信号的高带信号；

根据预先确定的时域包络个数 M 将所述当前帧的高带信号分成 M 个子帧，其中， M 为大于等于 2 的整数；

计算每一个所述子帧的时域包络；

5 其中，所述计算每一个所述子帧的时域包络包括：

采用非对称窗对所述 M 个子帧中的最前端的子帧和所述 M 个子帧中的最末端的子帧进行加窗；

对所述 M 个子帧中除所述最前端的子帧和所述最末端的子帧之外的子帧进行加窗。

10 根据本发明实施例提供的音频信号的时域包络的处理方法，在不同的条件下采用不同的窗长度和/或窗形状求解时域包络，减少因为时域包络差别太大引入的能量不连续的影响，能够提升输出信号的性能。

在第一方面的第一种可能的实施方式中，在采用非对称窗对所述 M 个子帧中的最前端的子帧和所述 M 个子帧中的最末端的子帧进行加窗之前，所述方法还包括：

15 根据所述当前帧信号的高带信号的前向缓存的长度确定所述非对称窗；或者，

根据所述当前帧信号的高带信号的前向缓存的长度和所述时域包络个数 M 确定所述非对称窗。

20 结合第一方面或第一方面的第一种可能的实施方式，在第一方面的第二种可能的实施方式中，所述对所述 M 个子帧中除所述最前端的子帧和所述最末端的子帧之外的子帧进行加窗，包括：

对所述 M 个子帧中除最前端的子帧和所述最末端的子帧之外的子帧采用对称窗进行加窗；或者，

25 对所述 M 个子帧中除最前端的子帧和所述最末端的子帧之外的子帧采用非对称窗进行加窗。

结合第一方面，在第一方面的第三种可能的实施方式中，所述非对称窗的窗长与对所述 M 个子帧中除所述最前端的子帧和所述最末端的子帧之外的子帧进行加窗所采用的窗的窗长相同。

30 结合第一方面的第一种可能的实施方式至第一方面的第三种可能的

实施方式任意之一所述的方法，在第一方面的第四种可能的实施方式中，所述根据所述当前帧音频信号的高带信号的前向缓存的长度确定非对称窗，包括：

当所述当前帧信号的高带信号的前向缓存的长度小于第一阈值时，
5 根据当前帧的前一帧信号的高带信号和所述当前帧信号的高带信号的前向缓存的长度确定所述非对称窗，其中，所述当前帧的前一帧信号的高带信号的最末端子帧采用的非对称窗和所述当前帧信号的高带信号的最前端子帧采用的非对称窗的混叠部分等于所述当前帧信号的高带信号的前向缓存的长度，所述第一阈值等于所述当前帧的高带信号的帧长除以
10 M 。

结合第一方面的第一种可能的实施方式至第一方面的第三种可能的实施方式任意之一所述的方法，在第一方面的第五种可能的实施方式中，所述根据所述当前帧信号的高带信号的前向缓存的长度确定非对称窗，包括：

15 当所述当前帧信号的高带信号的前向缓存的长度大于第一阈值时，根据所述当前帧的前一帧信号的高带信号和所述当前帧信号的高带信号的前向缓存的长度确定所述非对称窗，其中，所述当前帧的前一帧信号的高带信号的最末端子帧采用的非对称窗和所述当前帧信号的高带信号的最前端子帧采用的非对称窗的混叠部分等于所述第一阈值，所述第一
20 阈值等于所述当前帧的高带信号的帧长除以 M 。

结合第一方面至第一方面的第五种可能的实施方式任意之一所述的方法，在第一方面的第六种可能的实施方式中，根据下列之一方式确定所述时域包络个数 M ：

根据所述当前帧信号得到所述当前帧信号的低带信号，当所述当前
25 帧信号的低带信号的基音周期大于第二阈值时， $M=M1$ ；或者，

根据所述当前帧信号得到所述当前帧信号的低带信号，当所述当前帧信号的基音周期不大于第二阈值时， $M=M2$ ；

其中， $M1$ ， $M2$ 均为正整数，且 $M2>M1$ 。

结合第一方面至第一方面的第五种可能的实施方式任意之一所述的方法，在第一方面的第七种可能的实施方式中，所述方法还包括：
30

根据所述当前帧信号得到所述当前帧信号的低带信号的基音周期；
当所述当前帧信号的类型与所述当前帧的前一帧信号的类型相同，
且所述当前帧的低带信号的基音周期大于第三阈值时，对每一个所述子
帧的时域包络进行平滑处理。

5 第二方面，本发明实施例提供一种音频信号的时域包络处理装置，
包括：

高带信号获取模块，用于根据接收到的当前帧信号，得到所述当前
帧信号的高带信号；

子帧获取模块，用于根据预先确定的时域包络个数 M 将所述当前帧
10 的高带信号分成 M 个子帧，其中， M 为大于等于 2 的整数；

时域包络获取模块，用于计算每一个所述子帧的时域包络；

其中，所述时域包络获取模块具体用于：

采用非对称窗对所述 M 个子帧中的最前端的子帧和所述 M 个子帧中
的最末端的子帧进行加窗；

15 对所述 M 个子帧中除所述最前端的子帧和所述最末端的子帧之外的
子帧进行加窗。

根据本发明实施例提供的音频信号的时域包络的处理装置，在不同的
条件下采用不同的窗长度和/或窗形状求解时域包络，减少因为时域包
络差别太大引入的能量不连续的影响，能够提升输出信号的性能。

20 在第二方面的第一种可能的实施方式中，所述时域包络获取模块还
用于：

根据所述当前帧信号的高带信号的前向缓存的长度确定所述非对称
窗；或者，

25 根据所述当前帧信号的高带信号的前向缓存的长度和所述时域包络
个数 M 确定所述非对称窗。

结合第二方面的实施方式，在第二方面的第二种可能的实施方式
中，所述时域包络获取模块具体用于：

30 采用非对称窗对所述 M 个子帧中的最前端的子帧和所述 M 个子帧中
的最末端的子帧进行加窗，对所述 M 个子帧中除最前端的子帧和所述最
末端的子帧之外的子帧采用对称窗进行加窗；或者，

采用非对称窗对所述 M 个子帧中的最前端的子帧和所述 M 个子帧中的最末端的子帧进行加窗，对所述 M 个子帧中除最前端的子帧和所述最末端的子帧之外的子帧采用非对称窗进行加窗。

结合第二方面的实施方式，在第二方面的第三种可能的实施方式中，所述非对称窗的窗长与对所述 M 个子帧中除所述最前端的子帧和所述最末端的子帧之外的子帧进行加窗所采用的窗的窗长相同。

结合第二方面至第二方面的第三种可能的实施方式任意之一所述的装置，在第二方面的第四种可能的实施方式中，还包括：确定模块，用于根据下列之一方式确定所述时域包络个数 M ：

10 根据所述当前帧信号得到所述当前帧信号的低带信号，当所述当前帧信号的低带信号的基音周期大于第二阈值时， $M=M1$ ；或者，

根据所述当前帧信号得到所述当前帧信号的低带信号，当所述当前帧信号的低带信号的基音周期不大于第二阈值时， $M=M2$ ；

其中， $M1$ ， $M2$ 均为正整数，且 $M2>M1$ 。

15 本发明第三方面的实施例公开了一种编码器，所述编码器具体用于：

用于根据接收到的当前帧信号，得到所述当前帧信号的低带信号和所述当前帧信号的高带信号；

20 对所述当前帧信号的低带信号进行编码，得到低带编码的激励信号；

对所述当前帧信号的高带信号进行线性预测，得到线性预测系数；

量化所述线性预测系数，得到量化后的线性预测系数；

根据所述低带编码的激励信号和所述量化后的线性预测系数得到预测的高带信号；

25 计算及量化所述预测的高带信号的时域包络；

其中，所述计算所述预测的高带信号的时域包络包括：

根据预先确定的时域包络个数 M 将所述预测的高带信号分成 M 个子帧，其中， M 为大于等于 2 的整数，

30 采用非对称窗对所述 M 个子帧中的最前端的子帧和所述 M 个子帧中的最末端的子帧进行加窗，

对所述 M 个子帧中除所述最前端的子帧和所述最末端的子帧之外的子帧进行加窗；

对量化后的时域包络进行编码。

根据本发明实施例提供的编码器，在不同的条件下采用不同的窗长度和/或窗形状求解时域包络，减少因为时域包络差别太大引入的能量不连续的影响，能够提升输出信号的性能。

附图说明

为了更清楚地说明本发明实施例中的技术方案，下面将对实施例描述中所需要使用的附图作一简单地介绍，显而易见地，下面描述中的附图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性的前提下，还可以根据这些附图获得其他的附图。

图 1 为一种对音频信号进行编码的过程示意图；

图 2 为本发明音频信号的时域包络处理方法实施例一的流程图；

图 3 为本发明实施例中对音频信号进行处理的示意图；

图 4 为本发明另一实施例的对音频信号进行处理的示意图；

图 5 为本发明另一实施例的对音频信号进行处理的示意图；

图 6 为本发明音频信号的时域包络处理方法实施例二的流程图；

图 7 为本发明实施例的时域包络处理装置的结构示意图；

图 8 为本发明实施例的编码器的结构示意图。

具体实施方式

为使本发明实施例的目的、技术方案和优点更加清楚，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

图 1 为一种对语音信号进行编码的过程示意图，如图 1 所示，在编码端，在获得原始音频信号后，首先对原始音频信号进行信号分解，得到原始音频信号的低带信号和高带信号，接着对低带信号通过已有算法进行编码得

到低带的码流，已有算法（例如代数码本激励线性预测编码（Algebraic Code Excited Linear Prediction，简称：ACELP），或码本激励线性预测编码（Code Excited Linear Prediction，简称：CELP 等算法），同时，在进行低带编码过程中，得到低带的激励信号，并对低带激励信号进行预处理；对于原始音频

5 信号的高带信号，首先进行预处理，然后做线性预测（Linear prediction，以下简称：LP）分析得到 LP 系数，量化该 LP 系数。接着将预处理后的低带激励信号通过 LP 合成滤波器（滤波器系数为量化后的 LP 系数）得到预测的高带信号。根据预处理后的高带信号和预测的高带信号，计算及量化高带信号的时域包络，最后输出编码码流（MUX）。计算并量化高带信号的时域包络

10 的过程为：根据事先设定好的时域包络的个数 N ，将预处理后的高带信号和预测的高带信号分别分成 N 个子帧，对每一个子帧进行加窗，然后计算预处理后的原始高带信号每一个子帧和预测的高带信号的相对应的每一个子帧的时域能量或子帧内每个样点幅度的平均值。其中，事先设定好的时域包络的个数 N 是根据前向缓存（lookahead）的长度来确定的， N 为正整数。

15 本发明实施例提供一种音频信号的时域包络处理方法，主要用于图 1 中所示的计算及量化时域包络的步骤，还可以用于其它采用同样原理的求解时域包络的处理流程中。下面结合附图详细说明本发明实施例提供的音频信号的时域包络处理方法。

图 2 为本发明音频信号的时域包络处理方法实施例一的流程图，如图

20 2 所示，本实施例的方法包括：

S21、根据接收到的当前帧信号，得到当前帧信号的高带信号。

当前帧信号即可以是语音信号，也可以是音乐信号，还可能是噪音信号，在此不做具体的限制。

S22、根据预先确定的时域包络个数 M 将当前帧的高带信号分成 M 个

25 子帧，其中， M 为大于等于 2 的整数。

其中，具体来说，要预先确定的时域包络个数 M 可以根据整体算法要求和经验值确定。时域包络个数 M 例如是编码器事先根据整体算法或经验值确定，确定后不会改变。例如一般对 20ms 一帧的输入信号，如果输入信号相对平稳，求解 4 个或者 2 个时域包络，但对一些非平稳信号，需

30 要求解更多如 8 个时域包络。

S23、计算每一个子帧的时域包络。

其中，计算每一个子帧的时域包络包括：

采用非对称窗对 M 个子帧中的最前端的子帧和 M 个子帧中的最末端的子帧进行加窗。

5 对 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧进行加窗。

进一步地，在采用非对称窗对 M 个子帧中的最前端的子帧和 M 个子帧中的最末端的子帧进行加窗之前，本实施例的方法还可以包括：

10 根据当前帧信号的高带信号的前向缓存的长度确定非对称窗；或者，

根据当前帧信号的高带信号的前向缓存的长度和时域包络个数 M 确定非对称窗。

其中，对 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧进行加窗，具体可以包括：

15 对 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧采用对称窗进行加窗；或者，

对 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧采用非对称窗进行加窗。

20 其中，在一种可能的实施方式中，对最前端子帧和最末端子帧加窗使用的非对称窗的窗长与对 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧进行加窗所采用的窗的窗长相同。

在上述实施例中，作为一种可实施的方式，根据当前帧音频信号的高带信号的前向缓存的长度确定非对称窗，包括：

25 当当前帧信号的高带信号的前向缓存的长度小于第一阈值时，根据当前帧的前一帧信号的高带信号和当前帧信号的高带信号的前向缓存的长度确定非对称窗，其中，当前帧的前一帧信号的高带信号的最末端子帧采用的非对称窗和当前帧信号的高带信号的最前端子帧采用的非对称窗的混叠部分等于当前帧信号的高带信号的前向缓存的长度，第一阈值等于当前帧的高带信号的帧长除以 M 。

30 在一种可能的实施方式中，根据当前帧信号的高带信号的前向缓存的

长度确定非对称窗，包括：

当当前帧信号的高带信号的前向缓存的长度大于第一阈值时，根据当前帧的前一帧信号的高带信号和当前帧信号的高带信号的前向缓存的长度确定非对称窗，其中，当前帧的前一帧信号的高带信号的最末端子帧采用的非对称窗和当前帧信号的高带信号的最前端子帧采用的非对称窗的混叠部分等于第一阈值，第一阈值等于当前帧的高带信号的帧长除以 M 。

在本发明的一种实施例中，根据下列之一方式确定时域包络个数 M ：

根据当前帧信号得到当前帧信号的基音周期，当当前帧信号的基音周期大于第二阈值时， $M=M1$ ；或者，

根据当前帧信号得到当前帧信号的基音周期，当当前帧信号的基音周期不大于第二阈值时， $M=M2$ ；

其中， $M1$ ， $M2$ 均为正整数，且 $M2>M1$ 。在一种可能的方式中， $M1=4$ ， $M2=8$ 。

在上述实施例中，进一步地，本实施例的方法还可以包括：

根据当前帧信号得到当前帧信号的基音周期；

当当前帧信号的类型与当前帧的前一帧信号的类型相同，且当前帧的低带信号的基音周期大于第三阈值时，对每一个子帧的时域包络进行平滑处理。

对时域包络做平滑处理，具体可以是：将相邻的两个子帧的时域包络加权，加权后的时域包络作为这两个子帧的时域包络。例如，当解码端连续两帧信号都是浊音信号，或者一帧是浊音信号一帧是普通信号，且低带信号的基音周期大于给定阈值（大于 70 个样点，此时低带信号的采样率为 12.8kHz 采样）时，则对解码的高带信号时域包络做平滑处理，否则保持时域包络不变。平滑处理可以为：

$$\text{env}[0] = 0.5 * (\text{env}[0] + \text{env}[1]);$$

$$\text{env}[1] = 0.5 * (\text{env}[0] + \text{env}[1]);$$

...

$$\text{env}[N-1] = 0.5 * (\text{env}[N-1] + \text{env}[N]);$$

$$\text{env}[N] = 0.5 * (\text{env}[N-1] + \text{env}[N]).$$

其中， $env[]$ 为时域包络。

可以理解的是，上述步骤序号只是为了帮助理解本发明实施例而做出的一种示例，而不是对本发明实施例的具体限制。在实际的处理过程中，并不需要严格的按照上述顺序的限制。例如，可以先对除最前端和最末端的子帧之外的子帧进行加窗，再对最前端和最末端的子帧进行加窗。

图3为本发明实施例中对音频信号进行处理的示意图。

如图3所示，在编码端，在获得原始音频信号后，首先对原始音频信号进行信号分解，得到原始音频信号的低带信号和高带信号，接着对低带信号通过已有算法进行编码得到低带的码流，同时，在进行低带编码过程中，得到低带的激励信号，并对低带激励信号进行预处理；对于原始音频信号的高带信号，首先进行预处理，然后做LP分析得到LP系数，量化该LP系数。接着将预处理后的低带激励信号通过LP合成滤波器（滤波器系数为量化后的LP系数）得到预测的高带信号。根据预处理后的高带信号和预测的高带信号，计算及量化高带信号的时域包络，最后输出编码码流。

除了计算及量化高带信号的时域包络的步骤之外，对于音频信号的其它步骤的处理可以参考现有技术中所采用的方法，在此不再赘述。

下面以具体对图3中所示的 $N+1$ 帧的处理来描述本发明实施例中计算及量化时域包络的步骤。

如图3所示，将第 $N+1$ 帧按照需要计算的时域包络的个数划分为 M 个子帧， M 为正整数。在一种可能的实施方式中， M 的值可以是3、4、5、8等。在此不做限制。

对 M 个子帧中的最前端的子帧和 M 个子帧中最末端的子帧采用非对称窗进行加窗。 $N+1$ 帧的 M 个子帧中最前端的子帧为与前一帧（ N 帧）的信号有重叠部分的子帧；最末端的子帧为与后一帧（ $N+2$ 帧，图中未示出）的信号有重叠部分的子帧。在一种可能的方式中，如图3所示，最前端的子帧即为 $N+1$ 帧中最左端的子帧，最末端的子帧即为 $N+1$ 帧中最右端的子帧。可以理解的是，最左和最右只是结合图3的一种具体示例，而不是对本发明实施例的限制。实际中子帧的划分是不存在最左、最右这种方向性限制的。

对于最前端的子帧和最末端的子帧加窗所使用的非对称窗可以完全相

同，也可以不同。在此不做限制。在一种可能的实现方式中，最前端子帧使用的非对称窗的窗长和最末端子帧所使用的非对称窗的窗长相同。

在本发明的一个实施例中，如图3所示，对 $N+1$ 帧的 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧采用对称窗进行加窗。

- 5 在本发明的一个实施例中，对于最前端的子帧和最末端的子帧加窗所采用的非对称窗的窗长与对其它子帧采用的对称窗的窗长相等。可以理解的是，在另一种可能的方式中，非对称窗的窗长和对称窗的窗长也可以不等。

在本发明的一个实施例中，当第 $N+1$ 帧的帧长为80个样点，采样率为4kHz时，可以求解8个时域包络。

- 10 在一种可能的实现方式中，当第 $N+1$ 帧的帧长为80个样点，采样率为4kHz时，也可以求解4个时域包络。

在本发明的一个实施例中，除了预先设定之外，还可以根据 $N+1$ 帧的其它信息预先确定时域包络的个数 N 。下面是确定时域包络的个数 N 的实现方式的示例：

- 15 在一种可能实现的方式中，当第 $N+1$ 帧的低带信号的基音周期大于第二阈值时， $N=4$ ；或者，当第 $N+1$ 帧的低带信号的基音周期不大于第二阈值时， $N=8$ 。对于采用率为12.8kHz的低带信号，第二阈值可以为70个样点。可以理解的是，上述数值只是为了帮助理解本发明实施例而做出的一种具体举例，而不是对本发明实施例的具体限制。如图3所示，在对第 $N+1$ 帧的信号进行信号分解时可以得到第 $N+1$ 帧的低带信号，信号分解所采用的方法和求解低带信号的基音周期的方式可以采用现有技术中的任意一种方式，在此不做具体的限制。
- 20

可以理解的是，除了利用低带信号的基音周期以外，还可以利用信号的能量等其它参数。

- 25 在本发明的一个实施例中，在利用非对称窗对最前端的子帧和最末端的子帧进行加窗时，根据前向缓存的长度确定非对称窗。

- 在一种可能的实现方式中，当第 $N+1$ 帧的帧长为80个样点，采样率为4kHz，求解8个时域包络时，加窗所采用的非对称窗的窗长和对称窗的窗长可以都为20个样点。利用帧长除以包络个数得到第一阈值，此示例中第一阈值
- 30 等于10。则当前向缓存的长度小于10个样点时，第8个子帧（即，最末端的子

帧)采用的窗和第1个子帧(即,最前端的子帧)采用的窗的混叠部分等于前向缓存的长度。当前向缓存的长度大于等于10个样点时,第8个子帧采用的窗的右侧和第1个子帧采用的窗的左侧的长度可以等于另一侧(例如第一个子帧采用的窗的右侧或第八个子帧采用的窗的左侧)的窗长(10个样点),也可以根据经验设定一个长度(如,保持和前向缓存小于10个样点时相同的长度)。

在一种可能的实现方式中,当第 $N+1$ 帧的帧长为80个样点,采样率为4kHz,求解4个时域包络时,加窗所采用的非对称窗的窗长和对称窗的窗长可以都为40个样点。利用帧长除以包络个数得到第一阈值,此示例中第一阈值等于20。

在加窗后,计算各个子帧内预处理后的原始高带信号和预测的高带信号的时域能量或子帧内每个样点幅度的平均值。具体的计算方式可参考现有技术中提供的方式,本发明实施例提供的信号处理的方法在加窗时所采用的窗的形状和所需要加窗的个数的确定方式与现有技术不同。其它的计算方式均可参考现有技术中提供的方式。

根据本发明实施例提供的音频信号的时域包络的处理方法,在不同的条件下采用不同的窗长度和/或窗形状求解时域包络,减少因为时域包络差别太大引入的能量不连续的影响,能够提升输出信号的性能。

下面以具体对图4中所示的 $N+1$ 帧的处理来描述本发明另一实施例中计算及量化时域包络的步骤。

图4为本发明另一实施例的对音频信号进行处理的示意图,如图4所示,和图3所示类似,将第 $N+1$ 帧按照需要计算的时域包络的个数划分为 M 个子帧, M 为正整数。在一种可能的实施方式中, M 的值可以是3、4、5、8等。在此不做限制。

对 M 个子帧中的最前端的子帧和 M 个子帧中最末端的子帧采用非对称窗进行加窗。如图4所示,对于最前端的子帧和最末端的子帧加窗所使用的非对称窗不同。在一种可能的实现方式中,最前端子帧使用的非对称窗的窗长和最末端子帧所使用的非对称窗的窗长相同,也可以不同。

在本发明的一个实施例中,如图4所示,对 $N+1$ 帧的 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧采用形状相同的非对称窗进行加窗。

在本发明的一个实施例中，当第N+1帧的帧长为80个样点，采样率为4kHz时，可以求解8个时域包络。

在一种可能的实现方式中，当第N+1帧的帧长为80个样点，采样率为4kHz时，也可以求解4个时域包络。

- 5 在本发明的一个实施例中，除了预先设定之外，还可以根据N+1帧的其它信息预先确定时域包络的个数N。下面是确定时域包络的个数N的实现方式的示例：

 在一种可能实现的方式中，当第N+1帧的低带信号的基音周期大于第二阈值时， $N=4$ ；或者，当第N+1帧的低带信号的基音周期不大于第二阈值时， $N=8$ 。对于采用率为12.8kHz的低带信号，第二阈值可以为70个样点。可以理解的是，上述数值只是为了帮助理解本发明实施例而做出的一种具体举例，而不是对本发明实施例的具体限制。如图4所示，在对第N+1帧的信号进行信号分解时可以得到第N+1帧的低带信号，信号分解所采用的方法和求解低带信号的基音周期的方式可以采用现有技术中的任意一种方式，在此不做具体的限制。

 可以理解的是，除了利用低带信号的基音周期以外，还可以利用信号的能量等其它参数。

 在本发明的一个实施例中，在利用非对称窗对最前端的子帧和最末端的子帧进行加窗时，根据前向缓存的长度确定非对称窗。

- 20 在一种可能的实现方式中，当第N+1帧的帧长为80个样点，采样率为4kHz，求解8个时域包络时，加窗所采用的非对称窗的窗长和对称窗的窗长可以都为20个样点。利用帧长除以包络个数得到第一阈值，此示例中第一阈值等于10。则当前向缓存的长度小于10个样点时，第8个子帧采用的窗（即，最末端的子帧）和第1个子帧（即，最前端的子帧）采用的窗的混叠部分等于前向缓存的长度。当前向缓存的长度大于等于10个样点时，第8个子帧采用的窗的右侧和第1个子帧采用的窗的左侧的长度可以等于另一侧（例如第1个子帧采用的窗的右侧或第8个子帧采用的窗的左侧）的窗长（10个样点），也可以根据经验设定一个长度（如，保持和前向缓存小于10个样点时相同的长度）。

- 25 在一种可能的实现方式中，当第N+1帧的帧长为80个样点，采样率为4kHz，求解4个时域包络时，加窗所采用的非对称窗的窗长和对称窗的窗长可

以都为40个样点。利用帧长除以包络个数得到第一阈值，此示例中第一阈值等于20。

在加窗后，计算各个子帧内预处理后的原始高带信号和预测的高带信号的时域能量或子帧内每个样点幅度的平均值。具体的计算方式可参考现有技术中提供的方式，本发明实施例提供的信号处理的方法在加窗时所采用的窗的形状和所需要加窗的个数的确定方式与现有技术不同。其它的计算方式均可参考现有技术中提供的方式。

下面以具体对图5中所示的N+1帧的处理来描述本发明另一实施例中计算及量化时域包络的步骤。

图5为本发明另一实施例的对音频信号进行处理的示意图，如图5所示，在编码端，在获得原始音频信号后，首先对原始音频信号进行信号分解，得到原始音频信号的低带信号和高带信号，接着对低带信号通过已有算法进行编码得到低带的码流，同时，在进行低带编码过程中，得到低带的激励信号，并对低带激励信号进行预处理；对于原始音频信号的高带信号，首先进行预处理，然后做LP分析得到LP系数，量化该LP系数。接着将预处理后的低带激励信号通过LP合成滤波器（滤波器系数为量化后的LP系数）得到预测的高带信号。根据预处理后的高带信号和预测的高带信号，计算及量化高带信号的时域包络，最后输出编码码流。

除了计算及量化高带信号的时域包络的步骤之外，对于音频信号的其它步骤的处理可以参考现有技术中所采用的方法，在此不再赘述。

下面以具体对图5中所示的N+1帧的处理来描述本发明实施例中计算及量化时域包络的步骤。

如图5所示，将第N+1帧按照需要计算的时域包络的个数划分为M个子帧，M为正整数。在一种可能的实施方式中，M的值可以是3、4、5、8等。在此不做限制。

对M个子帧中的最前端的子帧和M个子帧中最末端的子帧采用非对称窗进行加窗。N+1帧的M个子帧中最前端的子帧为与前一帧（N帧）的信号有重叠部分的子帧；最末端的子帧为与后一帧（N+2帧，图中未示出）的信号有重叠部分的子帧。在一种可能的方式中，如图3所示，最前端的子帧即为N+1帧中最左端的子帧，最末端的子帧即为N+1帧中最右端的子帧。可以理解的

是,最左和最右只是结合图3的一种具体示例,而不是对本发明实施例的限制。实际中子帧的划分是不存在最左、最右这种方向性限制的。

对于最前端的子帧和最末端的子帧加窗所使用的非对称窗可以完全相同,也可以不同。在此不做限制。在一种可能的实现方式中,最前端子帧使用的非对称窗的窗长和最末端子帧所使用的非对称窗的窗长相同。

在本发明的一种可能实现的方式中,对 M 个子帧中的最前端的子帧和 M 个子帧中最末端的子帧采用非对称窗进行加窗,其中对 M 个子帧中的最前端的子帧采用的非对称窗与对 M 个子帧中最末端的子帧采用的非对称窗的形状不同,其中一个非对称窗以水平方向旋转 180 度可以与另一个非对称窗重合。在一种可能的实现方式中,最前端子帧使用的非对称窗的窗长和最末端子帧所使用的非对称窗的窗长相同。在本发明的一个实施例中,如图 5 所示,对 $N+1$ 帧的 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧采用对称窗进行加窗。对称窗的窗长与非对称窗的窗长不同。例如,对帧长为 20ms (80 个样点)采样率为 4kHz 的信号:如果前向缓存为 5 个样点,求解 4 个时域包络,采用本实施例的窗,两端的窗长为 30 个样点,连续两帧混叠时的样点数为 5 个样点,中间的两个窗长为 50 个样点,混叠 25 个样点。

在本发明的一个实施例中,如图5所示,对 $N+1$ 帧的 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧采用对称窗进行加窗。

在本发明的一个实施例中,对于最前端的子帧和最末端的子帧加窗所采用的非对称窗的窗长与对其它子帧采用的对称窗的窗长相等。可以理解的是,在另一种可能的方式中,非对称窗的窗长和对称窗的窗长也可以不等。

在本发明的一个实施例中,当第 $N+1$ 帧的帧长为 80 个样点,采样率为 4kHz 时,可以求解 8 个时域包络。

在一种可能的实现方式中,当第 $N+1$ 帧的帧长为 80 个样点,采样率为 4kHz 时,也可以求解 4 个时域包络。

在本发明的一个实施例中,除了预先设定之外,还可以根据 $N+1$ 帧的其它信息预先确定时域包络的个数 N 。下面是确定时域包络的个数 N 的实现方式的示例:

在一种可能实现的方式中,当第 $N+1$ 帧的低带信号的基音周期大于第二

阈值时, $N=4$; 或者, 当第 $N+1$ 帧的低带信号的基音周期不大于第二阈值时, $N=8$ 。对于采用率为 12.8kHz 的低带信号, 第二阈值可以为70个样点。可以理解的是, 上述数值只是为了帮助理解本发明实施例而做出的一种具体举例, 而不是对本发明实施例的具体限制。如图3所示, 在对第 $N+1$ 帧的信号进行信号分解时可以得到第 $N+1$ 帧的低带信号, 信号分解所采用的方法和求解低带信号的基音周期的方式可以采用现有技术中的任意一种方式, 在此不做具体的限制。

可以理解的是, 除了利用低带信号的基音周期以外, 还可以利用信号的能量等其它参数。

在本发明的一个实施例中, 在利用非对称窗对最前端的子帧和最末端的子帧进行加窗时, 根据前向缓存的长度确定非对称窗。

在一种可能的实现方式中, 当第 $N+1$ 帧的帧长为80个样点, 采样率为 4kHz , 求解8个时域包络时, 加窗所采用的非对称窗的窗长和对称窗的窗长可以都为20个样点。利用帧长除以包络个数得到第一阈值, 此示例中第一阈值等于10。则当前向缓存的长度小于10个样点时, 第8个子帧(即, 最末端的子帧)采用的窗和第1个子帧(即, 最前端的子帧)采用的窗的混叠部分等于前向缓存的长度。当前向缓存的长度大于等于10个样点时, 第8个子帧采用的窗的右侧和第1个子帧采用的窗的左侧的长度可以等于另一侧(例如第一个子帧采用的窗的右侧或第八个子帧采用的窗的左侧)的窗长(10个样点), 也可以根据经验设定一个长度(如, 保持和前向缓存小于10个样点时相同的长度)。

在一种可能的实现方式中, 当第 $N+1$ 帧的帧长为80个样点, 采样率为 4kHz , 求解4个时域包络时, 加窗所采用的非对称窗的窗长和对称窗的窗长可以都为40个样点。利用帧长除以包络个数得到第一阈值, 此示例中第一阈值等于20。

在加窗后, 计算各个子帧内预处理后的原始高带信号和预测的高带信号的时域能量或子帧内每个样点幅度的平均值。具体的计算方式可参考现有技术中提供的方式, 本发明实施例提供的信号处理的方法在加窗时所采用的窗的形状和所需要加窗的个数的确定方式与现有技术不同。其它的计算方式均可参考现有技术中提供的方式。

根据本发明实施例提供的音频信号的时域包络的处理方法, 在不同的条

件下采用不同的窗长度和/或窗形状求解时域包络，减少因为时域包络差别太大引入的能量不连续的影响，能够提升输出信号的性能。

本实施例提供的音频信号的时域包络处理方法，通过根据接收到的音频帧信号得到音频帧的高带信号，然后根据预先确定的时域包络个数 M 将音频帧的高带信号分成 M 个子帧，最后计算每一个子帧的时域包络。从而有效避免了在 *lookahead* 很短，同时要保证子帧间很好的混叠引起的求解过多时域包络的问题，进而避免了对一些信号，因过多求解时域包络而引入的能量不连续的问题，同时降低了计算复杂度。

图 6 为本发明音频信号的时域包络处理方法实施例二的流程图，如图 6 所示，本实施例的方法可以包括：

S60、接收到待处理信号后，根据第一频带内时域信号的平稳状态或第二频带信号的基音周期大小，确定对待处理信号计算的时域包络个数 M ，第一频带为待处理信号的时域信号的频带或整个输入信号的频带，第二频带为低于给定阈值的频带或整个输入信号的频带。

其中，确定对待处理信号计算的时域包络个数 M ，具体包括：

当第一频带内时域信号处于平稳状态或第二频带信号的基音周期大于预设阈值时， M 等于 M_1 ，否则 M 等于 M_2 ， M_1 大于 M_2 ， M_1 、 M_2 都为正整数，预设阈值根据采样率确定。

平稳状态是指时域信号在一定时间内的能量或幅度的均值变化不大，或时域信号在一定时间内的偏差小于给定阈值。

例如，对帧长为 20ms（80 个样点）采样率为 4kHz 的高带信号，如果高带时域信号子帧间的能量的比值小于给定阈值（小于 0.5），或低带信号的基音周期大于给定阈值（大于 70 个样点，此时低带信号的采样率为 12.8kHz 采样），则在对该高带信号求解时域包络时，求解 4 个时域包络；否则，求解 8 个时域包络。

例如，对帧长为 20ms（320 个样点）采样率为 16kHz 的高带信号，如果高带时域信号子帧间的能量的比值小于给定阈值（小于 0.5），或低带信号的基音周期大于给定阈值（大于 70 个样点，此时低带信号的采样率为 12.8kHz 采样），则在对该高带信号求解时域包络时，求解 2 个时域包络；否则，求解 4 个时域包络。

S61、将待处理信号分成 M 个子帧，计算每一个子帧的时域包络。

其中，本实施例对每一个子帧进行加窗处理时，不限定采用何种加窗方式进行加窗处理。

本实施例提供的音频信号的时域包络处理方法，通过根据不同的条件求解不同个数的时域包络，有效避免了对一定条件下的信号求解过多的时域包络造成的能量不连续，进而引起的听觉质量下降，同时，可以有效降低算法的平均复杂度。

本发明实施例还提供一种音频信号的时域包络处理装置，可以用于执行图 1-图 5 中所示部分方法，还可以用于其它采用同样原理的求解时域包络的处理流程中。下面结合附图详细说明本发明实施例提供的音频信号的时域包络处理装置的结构。

图 7 为本发明实施例的时域包络处理装置的结构示意图，如图 7 所示，本实施例的时域包络处理装置 70 包括：高带信号获取模块 71，用于根据接收到的当前帧信号，得到当前帧信号的高带信号；子帧获取模块 72，用于根据预先确定的时域包络个数 M 将当前帧的高带信号分成 M 个子帧，其中， M 为大于等于 2 的整数；时域包络获取模块 73，用于计算每一个子帧的时域包络；其中，时域包络获取模块 73 具体用于：采用非对称窗对 M 个子帧中的最前端的子帧和 M 个子帧中的最末端的子帧进行加窗；对 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧进行加窗。

在本发明实施例一种可能的方式中，时域包络获取模块 73 还用于：根据当前帧信号的高带信号的前向缓存的长度确定非对称窗；或者，

根据当前帧信号的高带信号的前向缓存的长度和时域包络个数 M 确定非对称窗。

在本发明一个实施例中，时域包络获取模块 73 具体用于：

采用非对称窗对 M 个子帧中的最前端的子帧和 M 个子帧中的最末端的子帧进行加窗，对 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧采用对称窗进行加窗；或者，

采用非对称窗对 M 个子帧中的最前端的子帧和 M 个子帧中的最末端

的子帧进行加窗，对 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧采用非对称窗进行加窗。

在本发明实施例一种可能的实现方式中，非对称窗的窗长与对 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧进行加窗所采用的窗的窗长相同。在本发明的一个实施例中，时域包络获取模块 73 还用于：根据当前帧信号得到当前帧信号的低带信号的基音周期；

当当前帧信号的类型与当前帧的前一帧信号的类型相同，且当前帧的低带信号的基音周期大于第三阈值时，对每一个子帧的时域包络进行平滑处理。

对时域包络做平滑处理，具体可以是：将相邻的两个子帧的时域包络加权，加权后的时域包络作为这两个子帧的时域包络。例如，当解码端连续两帧信号都是浊音信号，或者一帧是浊音信号一帧是普通信号，且低带信号的基音周期大于给定阈值（大于 70 个样点，此时低带信号的采样率为 12.8kHz 采样）时，则对解码的高带信号时域包络做平滑处理，否则保持时域包络不变。平滑处理可以为：

$$\text{env}[0] = 0.5 * (\text{env}[0] + \text{env}[1]);$$

$$\text{env}[1] = 0.5 * (\text{env}[0] + \text{env}[1]);$$

...

$$\text{env}[N-1] = 0.5 * (\text{env}[N-1] + \text{env}[N]);$$

$$\text{env}[N] = 0.5 * (\text{env}[N-1] + \text{env}[N]).$$

其中， $\text{env}[]$ 为时域包络。

在本发明的一个实施例中，时域包络处理装置 70 还包括：确定模块 74，用于根据下列之一方式确定时域包络个数 M ：

根据当前帧信号得到当前帧信号的低带信号，当当前帧信号的低带信号的基音周期大于第二阈值时， $M=M1$ ；或者，

根据当前帧信号得到当前帧信号的低带信号，当当前帧信号的低带信号的基音周期不大于第二阈值时， $M=M2$ ；

其中， $M1$ ， $M2$ 均为正整数，且 $M2 > M1$ 。

在本发明的实施例中，要预先确定的时域包络个数 M 可以是根据整体算法要求和经验值确定。时域包络个数 M 例如是编码器事先根据整体

算法或经验值确定，确定后不会改变。例如一般对 20ms 一帧的输入信号，如果输入信号相对平稳，求解 4 个或者 2 个时域包络，但对一些非平稳信号，需要求解更多如 8 个时域包络。

具体来说，首先，在编码端，在获得原始音频信号后，首先对原始音频信号进行信号分解，得到原始音频信号的低带信号和高带信号，接着对低带信号通过已有算法进行编码得到低带的码流，同时，在进行低带编码过程中，得到低带的激励信号，并对低带激励信号进行预处理；对于原始音频信号的高带信号，首先进行预处理，然后做 LP 分析得到 LP 系数，量化该 LP 系数。接着将预处理后的低带激励信号通过 LP 合成滤波器（滤波器系数为量化后的 LP 系数）得到预测的高带信号。根据预处理后的高带信号和预测的高带信号，计算及量化高带信号的时域包络，最后输出编码码流。

除了计算及量化高带信号的时域包络的步骤之外，对于音频信号的其它步骤的处理可以参考现有技术中所采用的方法，在此不再赘述。

本实施例的装置，可以用于执行图 2-图 5 所示方法实施例的技术方案，其实现原理类似。

在一个具体的示例中，在编码端，在获得原始音频信号后，首先对原始音频信号进行信号分解，得到原始音频信号的低带信号和高带信号，接着对低带信号通过已有算法进行编码得到低带的码流，同时，在进行低带编码过程中，得到低带的激励信号，并对低带激励信号进行预处理；对于原始音频信号的高带信号，首先进行预处理，然后做 LP 分析得到 LP 系数，量化该 LP 系数。接着将预处理后的低带激励信号通过 LP 合成滤波器（滤波器系数为量化后的 LP 系数）得到预测的高带信号。根据预处理后的高带信号和预测的高带信号，计算及量化高带信号的时域包络，最后输出编码码流。

除了计算及量化高带信号的时域包络的步骤之外，对于音频信号的其它步骤的处理可以参考现有技术中所采用的方法，在此不再赘述。

将第 $N+1$ 帧按照需要计算的时域包络的个数划分为 M 个子帧， M 为正整数。在一种可能的实施方式中， M 的值可以是 3、4、5、8 等。在此不做限制。

对 M 个子帧中的最前端的子帧和 M 个子帧中最末端的子帧采用非对称窗进行加窗。 $N+1$ 帧的 M 个子帧中最前端的子帧为与前一帧（ N 帧）的信号有重

叠部分的子帧；最末端的子帧为与后一帧（ $N+2$ 帧，图中未示出）的信号有重叠部分的子帧。在一种可能的方式中，最前端的子帧即为 $N+1$ 帧中最左端的子帧，最末端的子帧即为 $N+1$ 帧中最右端的子帧。可以理解的是，最左和最右只是一种具体示例，而不是对本发明实施例的限制。实际中子帧的划分是不存在最左、最右这种方向性限制的。

对于最前端的子帧和最末端的子帧加窗所使用的非对称窗可以完全相同，也可以不同。在此不做限制。在一种可能的实现方式中，最前端子帧使用的非对称窗的窗长和最末端子帧所使用的非对称窗的窗长相同。

在本发明的一个实施例中，对 $N+1$ 帧的 M 个子帧中除最前端的子帧和最末端的子帧之外的子帧采用对称窗进行加窗。

在本发明的一个实施例中，对于最前端的子帧和最末端的子帧加窗所采用的非对称窗的窗长与对其它子帧采用的对称窗的窗长相等。可以理解的是，在另一种可能的方式中，非对称窗的窗长和对称窗的窗长也可以不等。

在本发明的一个实施例中，当第 $N+1$ 帧的帧长为80个样点，采样率为4kHz时，可以求解8个时域包络。

在一种可能的实现方式中，当第 $N+1$ 帧的帧长为80个样点，采样率为4kHz时，也可以求解4个时域包络。

在本发明的一个实施例中，除了预先设定之外，还可以根据 $N+1$ 帧的其它信息预先确定时域包络的个数 N 。下面是确定时域包络的个数 N 的实现方式的示例：

在一种可能实现的方式中，当第 $N+1$ 帧的低带信号的基音周期大于第二阈值时， $N=4$ ；或者，当第 $N+1$ 帧的低带信号的基音周期不大于第二阈值时， $N=8$ 。对于采用率为12.8kHz的低带信号，第二阈值可以为70个样点。可以理解的是，上述数值只是为了帮助理解本发明实施例而做出的一种具体举例，而不是对本发明实施例的具体限制。在对第 $N+1$ 帧的信号进行信号分解时可以得到第 $N+1$ 帧的低带信号，信号分解所采用的方法和求解低带信号的基音周期的方式可以采用现有技术中的任意一种方式，在此不做具体的限制。

可以理解的是，除了利用低带信号的基音周期以外，还可以利用信号的能量等其它参数。

在本发明的一个实施例中，在利用非对称窗对最前端的子帧和最末端的

子帧进行加窗时，根据前向缓存的长度确定非对称窗。

在一种可能的实现方式中，当第N+1帧的帧长为80个样点，采样率为4kHz，求解8个时域包络时，加窗所采用的非对称窗的窗长和对称窗的窗长可以都为20个样点。利用帧长除以包络个数得到第一阈值，此示例中第一阈值
5 等于10。则当前向缓存的长度小于10个样点时，第8个子帧（即，最末端的子帧）采用的窗和第1个子帧（即，最前端的子帧）采用的窗的混叠部分等于前向缓存的长度。当前向缓存的长度大于等于10个样点时，第8个子帧采用的窗的右侧和第1个子帧采用的窗的左侧的长度可以等于另一侧（例如第一个子帧采用的窗的右侧或第八个子帧采用的窗的左侧）的窗长（10个样点），也可以
10 根据经验设定一个长度（如，保持和前向缓存小于10个样点时相同的长度）。

在一种可能的实现方式中，当第N+1帧的帧长为80个样点，采样率为4kHz，求解4个时域包络时，加窗所采用的非对称窗的窗长和对称窗的窗长可以都为40个样点。利用帧长除以包络个数得到第一阈值，此示例中第一阈值
15 等于20。

在加窗后，计算各个子帧内预处理后的原始高带信号和预测的高带信号的时域能量或子帧内每个样点幅度的平均值。具体的计算方式可参考现有技术中提供的方式，本发明实施例提供的信号处理的方法在加窗时所采用的窗的形状和所需要加窗的个数的确定方式与现有技术不同。其它的计算方式均可参考现有技术中提供的方式。
20

本实施例提供的音频信号的时域包络处理装置，通过根据不同的条件求解不同个数的时域包络，有效避免了对一定条件下的信号求解过多的时域包络造成的能量不连续，进而引起的听觉质量下降，同时，可以有效降低算法的平均复杂度。

下面结合图8描述本发明实施例的一种编码器80，图8为本发明实施例的编码器的结构示意图，如图8所示，编码器80具体用于：
25

用于根据接收到的当前帧信号，得到当前帧信号的低带信号和当前帧信号的高带信号；

对当前帧信号的低带信号进行编码，得到低带编码的激励信号；

对当前帧信号的高带信号进行线性预测，得到线性预测系数；

30 量化线性预测系数，得到量化后的线性预测系数；

根据低带编码的激励信号和量化后的线性预测系数得到预测的高带信号；

计算及量化预测的高带信号的时域包络；

其中，计算所述预测的高带信号的时域包络包括：

- 5 根据预先确定的时域包络个数 M 将预测的高带信号分成 M 个子帧，其中， M 为大于等于 2 的整数，

采用非对称窗对 M 个子帧中的最前端的子帧和 M 个子帧中的最末端的子帧进行加窗，

- 10 对 M 个子帧中除所述最前端的子帧和最末端的子帧之外的子帧进行加窗；

对量化后的时域包络进行编码。

可以理解的是，编码器 80 可以用于执行上述任意的方法实施例。也可以包括任意实施例的时域包络处理装置 70。具体的编码器 80 所执行的功能可参考前述方法和装置实施例，在此不再赘述。

- 15 本领域普通技术人员可以理解：实现上述各方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成。前述的程序可以存储于一计算机可读取存储介质中。该程序在执行时，执行包括上述各方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

- 20 最后应说明的是：以上各实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述各实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分或者全部技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

25

权 利 要 求 书

1、一种音频信号的时域包络处理方法，其特征在于，包括：

根据接收到的当前帧信号，得到所述当前帧信号的高带信号；

根据预先确定的时域包络个数 M 将所述当前帧的高带信号分成 M 个
5 子帧，其中， M 为大于等于 2 的整数；

计算每一个所述子帧的时域包络；

其中，所述计算每一个所述子帧的时域包络包括：

采用非对称窗对所述 M 个子帧中的最前端的子帧和所述 M 个子帧中的最末端的子帧进行加窗；

10 对所述 M 个子帧中除所述最前端的子帧和所述最末端的子帧之外的子帧进行加窗。

2、根据权利要求 1 所述的方法，其特征在于，在采用非对称窗对所述 M 个子帧中的最前端的子帧和所述 M 个子帧中的最末端的子帧进行加窗之前，所述方法还包括：

15 根据所述当前帧信号的高带信号的前向缓存的长度确定所述非对称窗；或者，

根据所述当前帧信号的高带信号的前向缓存的长度和所述时域包络个数 M 确定所述非对称窗。

3、根据权利要求 1 或 2 所述的方法，其特征在于，所述对所述 M 个
20 子帧中除所述最前端的子帧和所述最末端的子帧之外的子帧进行加窗，包括：

对所述 M 个子帧中除最前端的子帧和所述最末端的子帧之外的子帧采用对称窗进行加窗；或者，

25 对所述 M 个子帧中除最前端的子帧和所述最末端的子帧之外的子帧采用非对称窗进行加窗。

4、根据权利要求 1 所述的方法，其特征在于，所述非对称窗的窗长与对所述 M 个子帧中除所述最前端的子帧和所述最末端的子帧之外的子帧进行加窗所采用的窗的窗长相同。

5、根据权利要求 2-4 任意之一所述的方法，其特征在于，所述根据
30 所述当前帧音频信号的高带信号的前向缓存的长度确定非对称窗，包

括：

当所述当前帧信号的高带信号的前向缓存的长度小于第一阈值时，根据当前帧的前一帧信号的高带信号和所述当前帧信号的高带信号的前向缓存的长度确定所述非对称窗，其中，所述当前帧的前一帧信号的高带信号的最末端子帧采用的非对称窗和所述当前帧信号的高带信号的最前端子帧采用的非对称窗的混叠部分等于所述当前帧信号的高带信号的前向缓存的长度，所述第一阈值等于所述当前帧的高带信号的帧长除以 M 。

6、根据权利要求 2-4 任意之一所述的方法，其特征在于，所述根据所述当前帧信号的高带信号的前向缓存的长度确定非对称窗，包括：

当所述当前帧信号的高带信号的前向缓存的长度大于第一阈值时，根据所述当前帧的前一帧信号的高带信号和所述当前帧信号的高带信号的前向缓存的长度确定所述非对称窗，其中，所述当前帧的前一帧信号的高带信号的最末端子帧采用的非对称窗和所述当前帧信号的高带信号的最前端子帧采用的非对称窗的混叠部分等于所述第一阈值，所述第一阈值等于所述当前帧的高带信号的帧长除以 M 。

7、根据权利要求 1-6 任意之一所述的方法，其特征在于，根据下列之一方式确定所述时域包络个数 M ：

根据所述当前帧信号得到所述当前帧信号的基音周期，当所述当前帧信号的基音周期大于第二阈值时， $M=M1$ ；或者，

根据所述当前帧信号得到所述当前帧信号的基音周期，当所述当前帧信号的基音周期不大于第二阈值时， $M=M2$ ；

其中， $M1$ ， $M2$ 均为正整数，且 $M2>M1$ 。

8、根据权利要求 1-6 任意之一所述的方法，其特征在于，所述方法还包括：

根据所述当前帧信号得到所述当前帧信号的基音周期；

当所述当前帧信号的类型与所述当前帧的前一帧信号的类型相同，且所述当前帧的基音周期大于第三阈值时，对每一个所述子帧的时域包络进行平滑处理。

9、一种音频信号的时域包络处理装置，其特征在于，包括：

高带信号获取模块，用于根据接收到的当前帧信号，得到所述当前帧信号的高带信号；

子帧获取模块，用于根据预先确定的时域包络个数 M 将所述当前帧的高带信号分成 M 个子帧，其中， M 为大于等于 2 的整数；

5 时域包络获取模块，用于计算每一个所述子帧的时域包络；

其中，所述时域包络获取模块具体用于：

采用非对称窗对所述 M 个子帧中的最前端的子帧和所述 M 个子帧中的最末端的子帧进行加窗；

10 对所述 M 个子帧中除所述最前端的子帧和所述最末端的子帧之外的子帧进行加窗。

10、根据权利要求 9 所述的装置，其特征在于，所述时域包络获取模块还用于：

根据所述当前帧信号的高带信号的前向缓存的长度确定所述非对称窗；或者，

15 根据所述当前帧信号的高带信号的前向缓存的长度和所述时域包络个数 M 确定所述非对称窗。

11、根据权利要求 9 所述的装置，其特征在于，所述时域包络获取模块具体用于：

20 采用非对称窗对所述 M 个子帧中的最前端的子帧和所述 M 个子帧中的最末端的子帧进行加窗，对所述 M 个子帧中除最前端的子帧和所述最末端的子帧之外的子帧采用对称窗进行加窗；或者，

采用非对称窗对所述 M 个子帧中的最前端的子帧和所述 M 个子帧中的最末端的子帧进行加窗，对所述 M 个子帧中除最前端的子帧和所述最末端的子帧之外的子帧采用非对称窗进行加窗。

25 12、根据权利要求 9 所述的装置，其特征在于，所述非对称窗的窗长与对所述 M 个子帧中除所述最前端的子帧和所述最末端的子帧之外的子帧进行加窗所采用的窗的窗长相同。

13、根据权利要求 9-12 任意之一所述的装置，其特征在于，还包括：确定模块，用于根据下列之一方式确定所述时域包络个数 M ：

30 根据所述当前帧信号得到所述当前帧信号的低带信号，当所述当前帧

信号的低带信号的基音周期大于第二阈值时， $M=M1$ ；或者，

根据所述当前帧信号得到所述当前帧信号的低带信号，当所述当前帧信号的低带信号的基音周期不大于第二阈值时， $M=M2$ ；

其中， $M1$ ， $M2$ 均为正整数，且 $M2>M1$ 。

- 5 14、根据权利要求 9-13 任意之一所述的装置，其特征在于，所述时域包络获取模块还用于：

根据所述当前帧信号得到所述当前帧信号的低带信号的基音周期；

当所述当前帧信号的类型与所述当前帧的前一帧信号的类型相同，且所述当前帧的低带信号的基音周期大于第三阈值时，对每一个所述子
10 帧的时域包络进行平滑处理。

15、一种编码器，其特征在于，所述编码器具体用于：

用于根据接收到的当前帧信号，得到所述当前帧信号的低带信号和所述当前帧信号的高带信号；

- 15 对所述当前帧信号的低带信号进行编码，得到低带编码的激励信号；

对所述当前帧信号的高带信号进行线性预测，得到线性预测系数；

量化所述线性预测系数，得到量化后的线性预测系数；

根据所述低带编码的激励信号和所述量化后的线性预测系数得到预测的高带信号；

- 20 计算及量化所述预测的高带信号的时域包络；

其中，所述计算所述预测的高带信号的时域包络包括：

根据预先确定的时域包络个数 M 将所述预测的高带信号分成 M 个子帧，其中， M 为大于等于 2 的整数，

- 25 采用非对称窗对所述 M 个子帧中的最前端的子帧和所述 M 个子帧中的最末端的子帧进行加窗，

对所述 M 个子帧中除所述最前端的子帧和所述最末端的子帧之外的子帧进行加窗；

对量化后的时域包络进行编码。

1/4

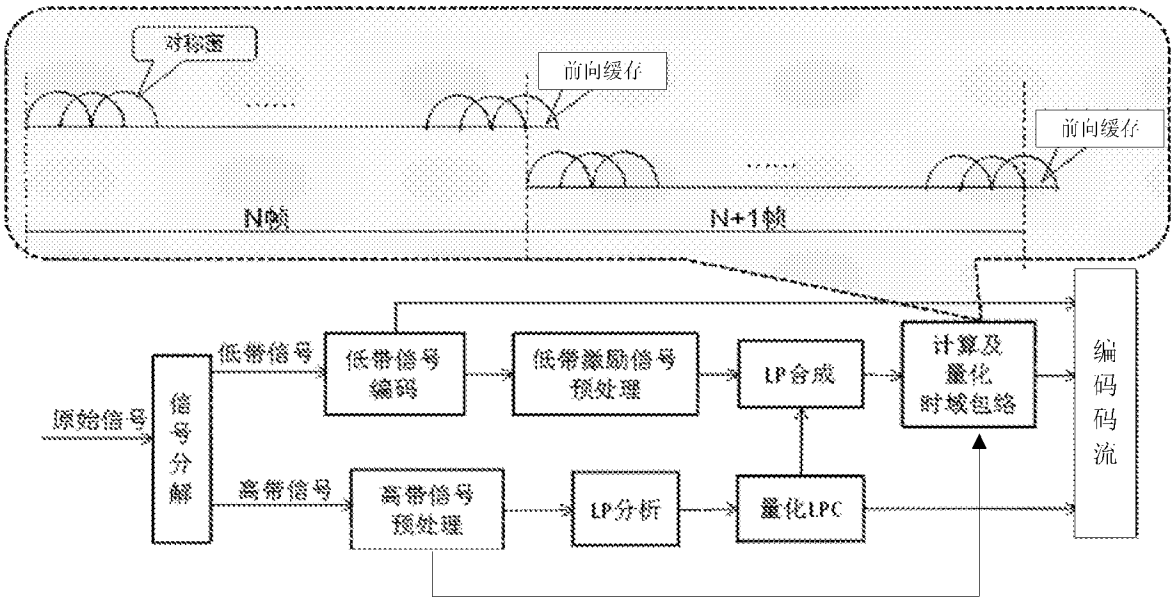


图 1

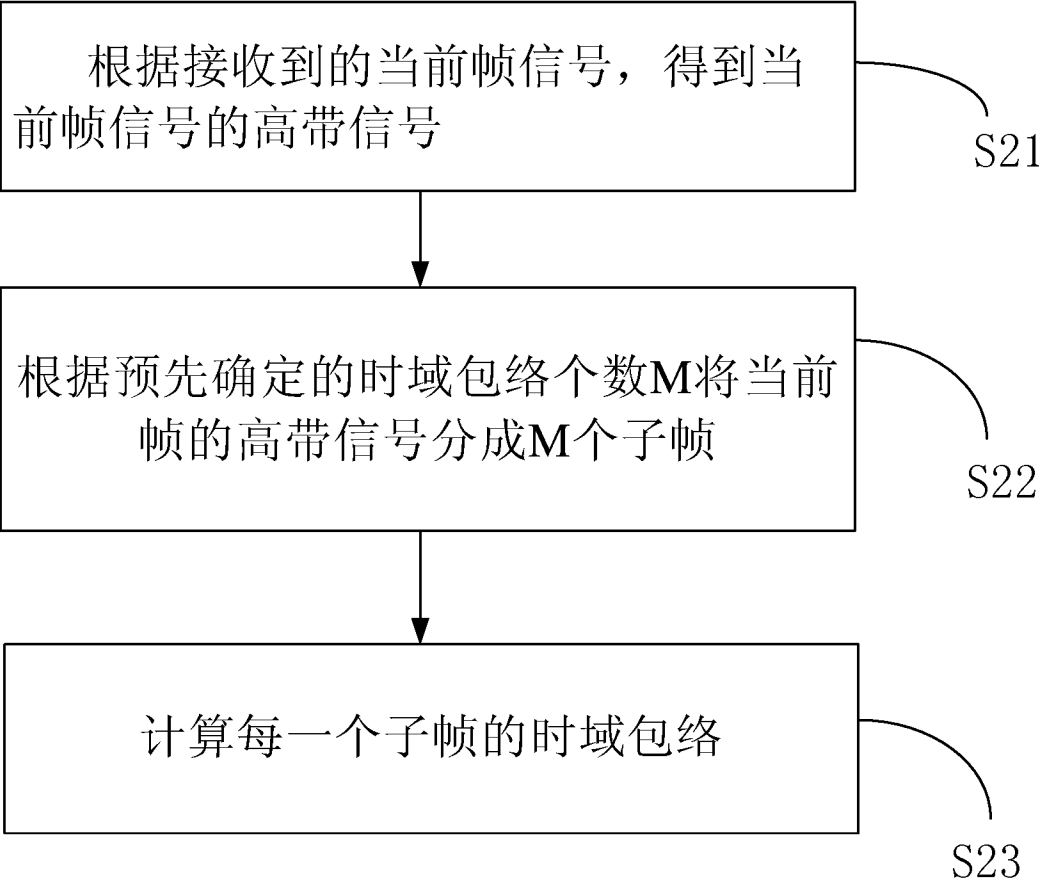


图 2

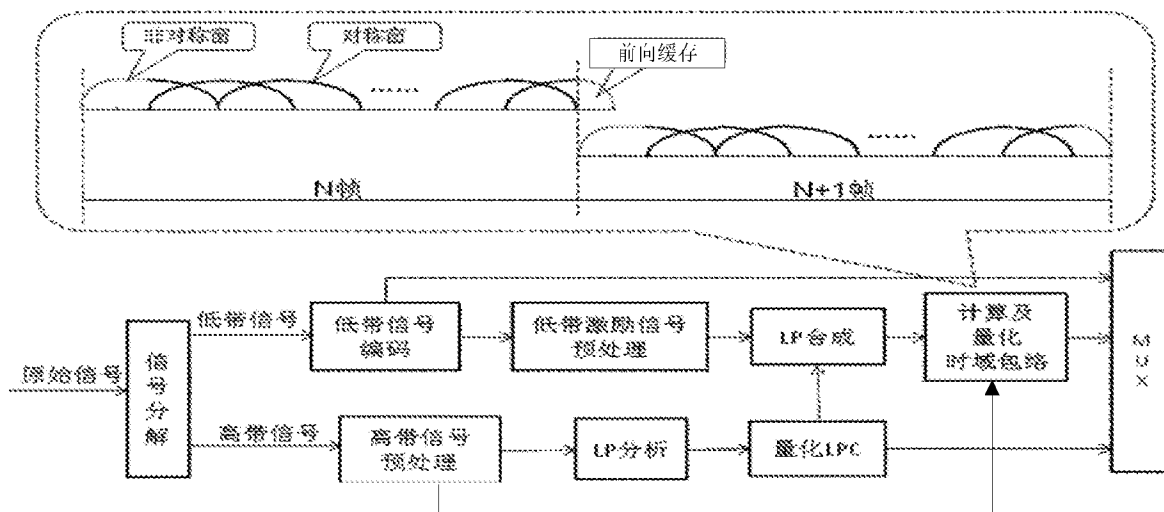


图 3

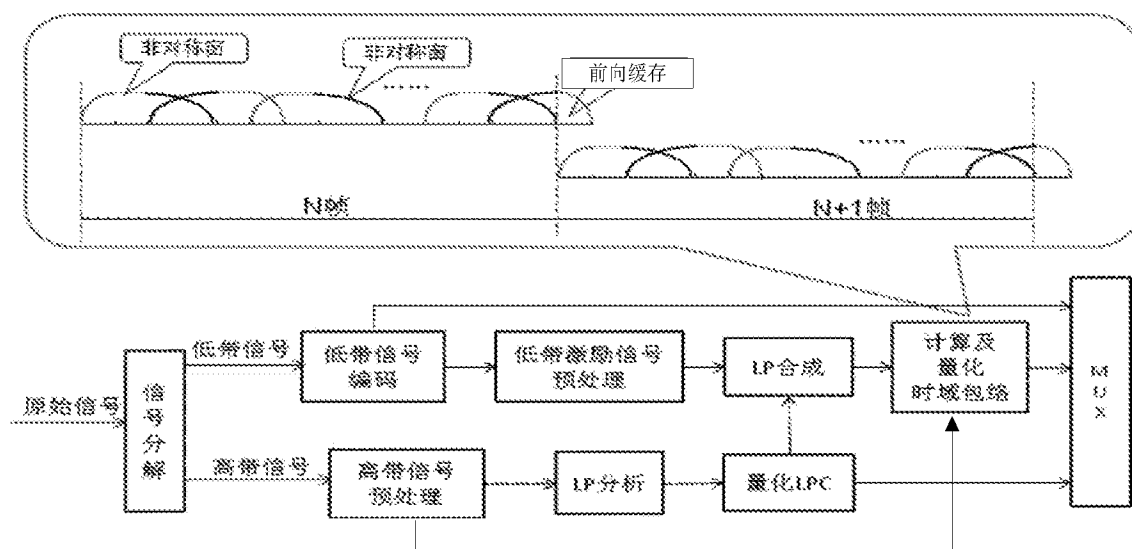


图 4

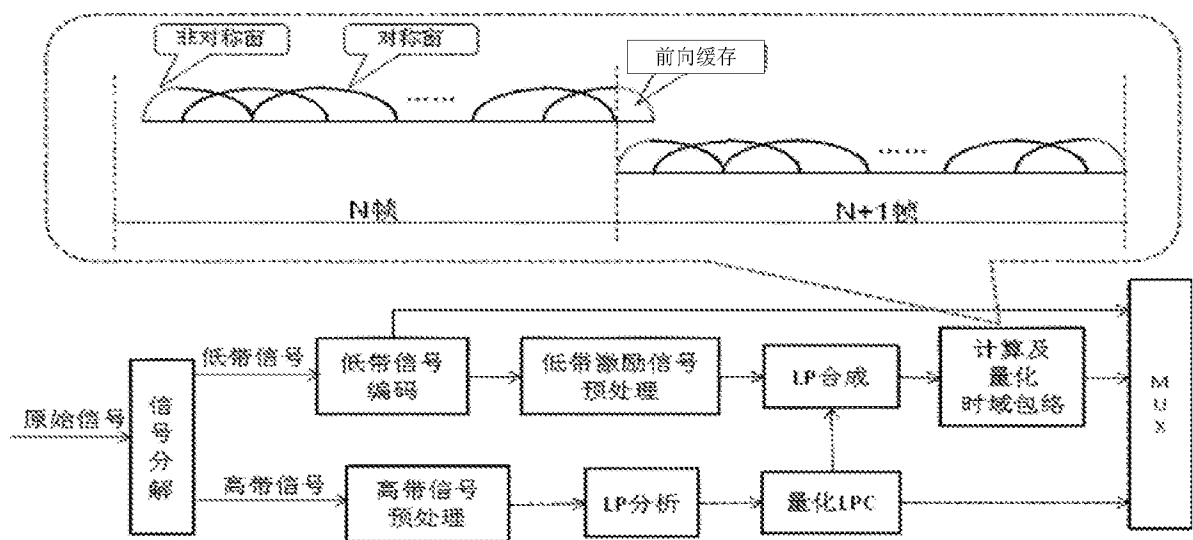


图 5

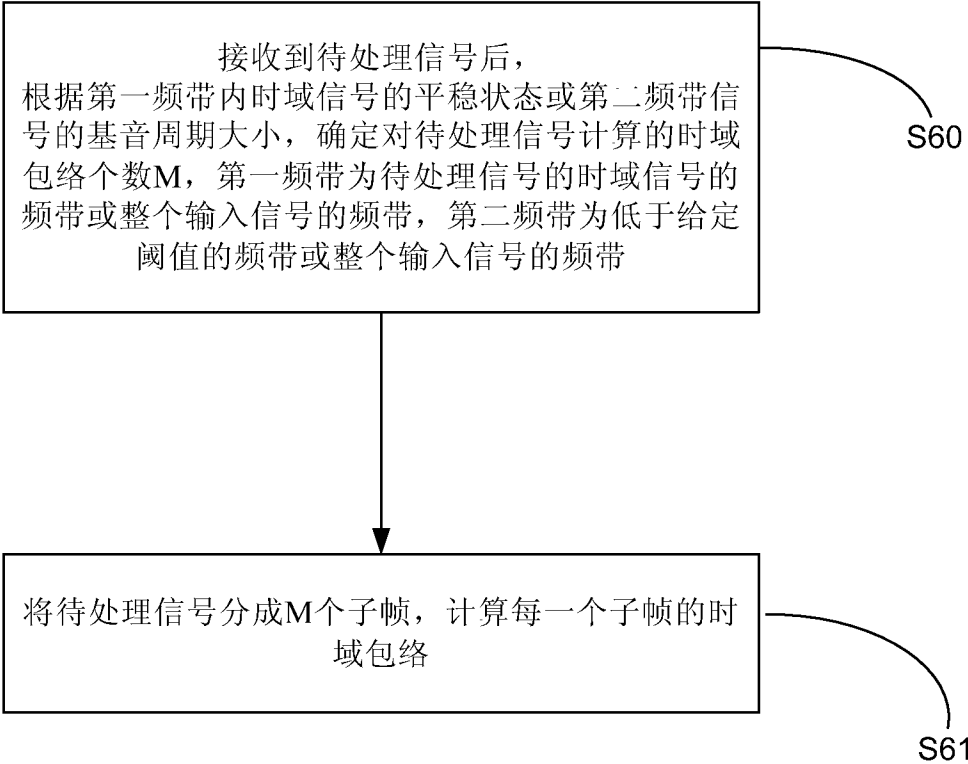


图 6

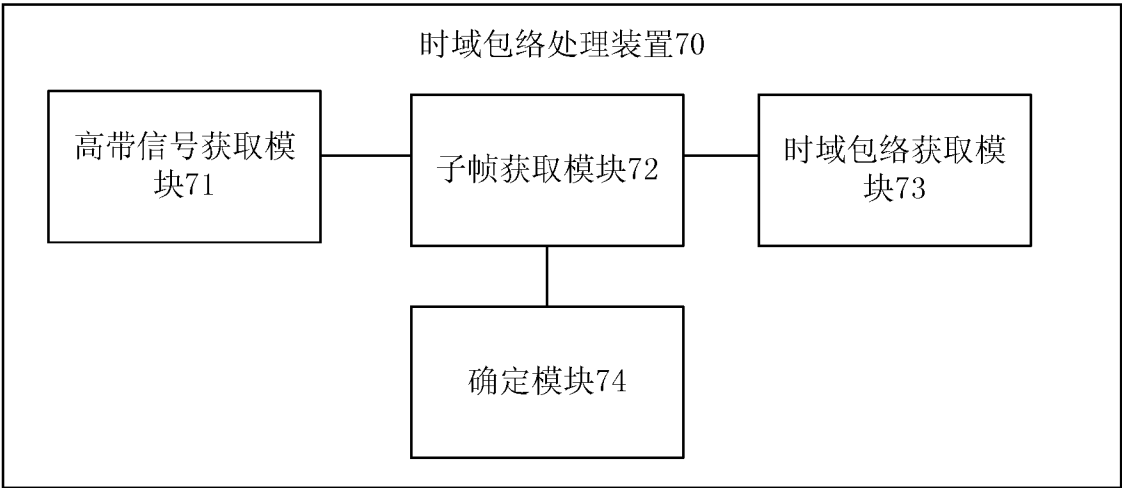


图 7

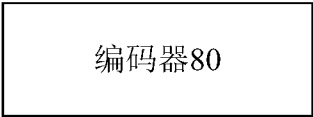


图 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2015/071727

A. CLASSIFICATION OF SUBJECT MATTER

G10L 19/022 (2013.01) i; G10L 19/135 (2013.01) i; G10L 19/20 (2013.01) i; G10L 25/45 (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)

CATXT; USTXT; CPRSABS; GBTXT; EPTXT; CNTXT; CNABS; WOTXT; SGTXT; TWTXT; VEN: sub, terminal?, end?, highband?,
high 1w band?, high 1w frequency 1w band?, start+, envelop+, end, frequency, high, G10L, front, band?, tail?, envelope?, frame?,
subframe?, window+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2012016668 A1 (FUTUREWEI TECHNOLOGIES INC.) 19 January 2012 (19.01.2012) description, paragraphs [0035]-[0076] and figures 1-11	1-15
A	CN 1424712 A (UNIV. BEIJING TECHNOLOGY) 18 June 2003 (18.06.2003) the whole document	1-15
A	WO 2013066238 A3 (ERICSSON CO., LTD. et al.) 01 August 2013 (01.08.2013) the whole document	1-15
A	JP 10222194 A (HETAI SEMICONDUCTOR CO., LTD.) 21 August 1998 (21.08.1998) the whole document	1-15
A	JP 2001166800 A (NIPPON TELEGRAPH & TELEPHONE CORP.) 22 June 2001 (22.06.2001) the whole document	1-15

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

☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 20 April 2015	Date of mailing of the international search report 27 April 2015
Name and mailing address of the ISA 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 SUI, Xin Telephone No. (86-10) 62085740

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/CN2015/071727

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US 2012016668 A1	19 January 2012	US 8560330 B2	15 October 2013
CN 1424712 A	18 June 2003	CN 1186765 C	26 January 2005
WO 2013066238 A3	01 August 2013	MX 2014004670 A	28 May 2014
		EP 2791937 A2	22 October 2014
		US 2014257827 A1	11 September 2014
		WO 2013066238 A2	10 May 2013
		CN 104221081 A	17 December 2014
JP 10222194 A	21 August 1998	None	
JP 2001166800 A	22 June 2001	JP 3510168 B2	22 March 2004

--

Form PCT/ISA /210 (patent family annex) (July 2009)

A. 主题的分类		
G10L 19/022(2013.01)i; G10L 19/135(2013.01)i; G10L 19/20(2013.01)i; G10L 25/45(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))		
CATXT;USTXT;CPRSABS;GBTXT;EPTXT;CNTXT;CNABS;WOTXT;SGTXT;TWTXT;VEN:高带, 高频带, 窗, 子帧, 加窗, 包络, sub, terminal?, end?, highband?, start+, envelop+, end, frequency, high, G10L, front, band?, tail?, envelope?, frame?, subframe?, window+		
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 2012016668 A1 (FUTUREWEI TECHNOLOGIES INC.) 2012年 1月 19日 (2012 - 01 - 19) 说明书[0035]到[0076]段, 附图1-11	1-15
A	CN 1424712 A (北京工业大学) 2003年 6月 18日 (2003 - 06 - 18) 全文	1-15
A	WO 2013066238 A3 (ERICSSON CO., LTD. 等) 2013年 8月 1日 (2013 - 08 - 01) 全文	1-15
A	JP 10222194 A (HETAI SEMICONDUCTOR CO., LTD.) 1998年 8月 21日 (1998 - 08 - 21) 全文	1-15
A	JP 2001166800 A (NIPPON TELEGRAPH & TELEPHONE CORP.) 2001年 6月 22日 (2001 - 06 - 22) 全文	1-15
☐ 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
2015年 4月 20日		2015年 4月 27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) 北京市海淀区蓟门桥西土城路6号 100088 中国		隋欣
传真号 (86-10)62019451		电话号码 (86-10)62085740

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2015/071727

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	2012016668	A1	2012年 1月 19日	US	8560330	B2	2013年 10月 15日
CN	1424712	A	2003年 6月 18日	CN	1186765	C	2005年 1月 26日
WO	2013066238	A3	2013年 8月 1日	MX	2014004670	A	2014年 5月 28日
				EP	2791937	A2	2014年 10月 22日
				US	2014257827	A1	2014年 9月 11日
				WO	2013066238	A2	2013年 5月 10日
				CN	104221081	A	2014年 12月 17日
JP	10222194	A	1998年 8月 21日	无			
JP	2001166800	A	2001年 6月 22日	JP	3510168	B2	2004年 3月 22日