



(10)授权公告号 CN 104221082 B

(45)授权公告日 2017.03.08

(21)申请号 201280071983.7

(22)申请日 2012.12.21

(65)同一申请的已公布的文献号

申请公布号 CN 104221082 A

(43)申请公布日 2014.12.17

(30)优先权数据

61/617,175 2012.03.29 US

(85)PCT国际申请进入国家阶段日

2014.09.28

(86)PCT国际申请的申请数据

PCT/SE2012/051470 2012.12.21

(87)PCT国际申请的公布数据

W02013/147668 EN 2013.10.03

(73)专利权人 瑞典爱立信有限公司

地址 瑞典斯德哥尔摩

(72)发明人 塞巴斯蒂安·内斯隆德

沃洛佳·格兰恰诺夫

托马斯·詹森·托夫特戈德

(74)专利代理机构 中科专利商标代理有限责任

公司 11021

代理人 赵伟

(51)Int.Cl.

G10L 21/038(2006.01)

(56)对比文件

CN 1408109 A, 2003.04.02,

CN 1897467 A, 2007.01.17,

CN 1960351 A, 2007.05.09,

US 2010/0063827 A1, 2010.03.11,

WO 2012/017621 A1, 2012.02.09,

CN 102044250 A, 2011.05.04,

审查员 李春雨

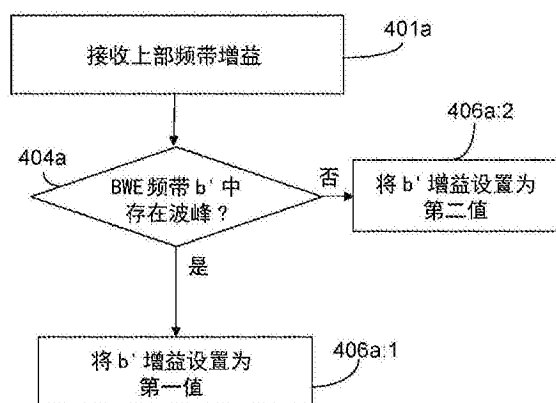
权利要求书2页 说明书10页 附图7页

(54)发明名称

谐波音频信号的带宽扩展

(57)摘要

一种在编解码器中支持谐波音频信号的带宽扩展(BWE)的方法和装置。编解码器的解码器部分中的方法包括:接收与频带b及频带b的多个相邻频带相关联的多个增益值。所述方法还包括:确定重建的对应频带b'是否包括谱峰。当频带b'包括谱峰时,基于接收到的多个增益值将与频带b'相关的增益值设置为第一值;否则基于接收到的多个增益值将所述增益值设置为第二值。本发明使得增益值与带宽扩展频域中的波峰位置一致。



1. 一种由变换音频解码器执行的用于支持谐波音频信号的带宽扩展“BWE”的方法,所述方法包括:

-接收 (401a) 与频带b及频带b的多个相邻频带相关联的多个增益值;

-确定 (404a) 带宽扩展频率区域的重建的对应频带b' 是否包括谱峰,以及:

当重建的频带b' 包括至少一个谱峰时:

-基于接收到的多个增益值,将与重建的频带b' 相关联的增益值设置 (406a:1) 为第一值,其中所述第一值是所述接收到的多个增益值的加权和;以及

当重建的频带b' 不包括任何谱峰时:

-基于接收到的多个增益值,将与重建的频带b' 相关联的增益值设置 (406a:2) 为第二值,其中,所述第二值小于或等于所述第一值;

从而使得增益值与所述带宽扩展频率区域中的波峰位置一致。

2. 根据权利要求1所述的方法,其中,所述加权和是接收到的多个增益值的平均值。

3. 根据前述权利要求的任意一项所述的方法,其中,所述第二值是接收到的多个增益值中最小的增益值之一。

4. 根据权利要求1或2所述的方法,其中,所述第二值是接收到的多个增益值之中的最小增益值。

5. 根据权利要求1或2所述的方法,还包括:

-接收 (402b) 反映原始信号的高频部分的至少一段的波峰能量与噪声底能量之间的关系系数 α ;

-基于接收到的系数 α ,将对应的重建的高频段的变换系数与噪声混合 (403b),

从而能够实现所述原始信号的所述高频部分的噪声特性的重建。

6. 一种用于支持谐波音频信号的带宽扩展“BWE”的音频解码器 (501),所述音频解码器包括:

-接收单元,适于接收与频带b及频带b的多个相邻频带相关联的多个增益值;

-确定单元 (504),适于确定带宽扩展频率区域的重建的对应频带b' 是否包括谱峰;

-增益应用单元 (508),适于:

-当重建的频带b' 包括至少一个谱峰时,基于接收到的多个增益值将与重建的频带b' 相关联的增益值设置为第一值,使得所述第一值是所述接收到的多个增益值的加权和;以及

-当重建的频带b' 不包括任何谱峰时,基于接收到的多个增益值将与重建的频带b' 相关联的增益值设置为第二值,其中,所述第二值小于或等于所述第一值;

从而使得增益值与所述带宽扩展频率区域中的波峰位置一致。

7. 根据权利要求6所述的音频解码器,其中,所述加权和是接收到的多个增益值的平均值。

8. 根据权利要求6-7中任一项所述的音频解码器,其中,所述第二值是接收到的多个增益值之中最小的增益值之一。

9. 根据权利要求6-7中任一项所述的音频解码器,其中,所述第二值是接收到的多个增益值之中的最小增益值。

10. 根据权利要求6-7中任一项所述的音频解码器,还适于接收反映原始信号的高频部

分的至少一段的波峰能量和噪声底能量之间的关系的系数 α ;以及还包括:

-噪声混合单元 (510),适于基于接收到的系数 α 将对应的重建的高频段的变换系数与噪声混合,

从而能够实现所述原始信号的所述高频部分的噪声特性的重建。

11.一种包括根据权利要求6-10中任一项所述的音频解码器的用户设备。

谐波音频信号的带宽扩展

技术领域

[0001] 本发明涉及音频信号的编码和解码,更具体地,涉及支持谐波音频信号的带宽扩展(BWE)。

背景技术

[0002] 基于变换的编码是当今的音频压缩/传输系统中最常用的方案。这种方案的主要步骤是首先通过合适的变换(例如DFT(离散傅立叶变换)、DCT(离散余弦变换)或MDCT(修正离散余弦变换))将信号波形的短块转换到频域。然后变换系数被量化、发送或存储且随后用于重建音频信号。这种方案对于一般的音频信号有效,但是需要足够高的比特速率以创建变换系数的足够好的表示。下面将给出这种变换域编码方案的高层概述。

[0003] 将要编码的波形逐块地变换到频域。用于此目的一个常用变换是所谓的修正离散余弦变换(MDCT)。将由此得到的频域变换矢量分成频谱包络(缓慢变化的能量)和频谱残留。频谱残留是使用所述频谱包络对所获得的频域矢量进行归一化而获得的。对频谱包络进行量化,将量化指标发送到解码器。接下来,使用量化的频谱包络作为比特分配算法的输入,并基于频谱包络的特性分配用于编码残留矢量的比特。作为此步骤的结果,某个数量的比特被分配给残留的不同部分(残留矢量或“子矢量”)。一些残留矢量不接收任何比特且必定是充满噪声的或带宽扩展的。通常,残留矢量的编码是两个步骤的过程;首先对矢量项的幅值进行编码,接着对非零项的符号(不要与“相位”混淆,相位是与例如傅立叶变换相关的)进行编码。将针对于残留幅值和符号的量化指标发送给解码器,在解码器中将残留和频谱包络进行组合,并最终变换回时域。

[0004] 电信网络的容量在持续增长。然而,尽管有增长的容量,仍然存在限制每一通信信道所需带宽的强驱动力。在移动网络中,用于每个呼叫的较小的传输带宽在移动设备和服务于该设备的基站两者中都产生较低的功率消耗。这转化成了移动运营商能量和成本的节约,同时,最终用户将体验到延长的电池寿命和增长的通话时间。此外,每一用户消耗的带宽越少,移动网络能够(并行地)服务的用户就越多。

[0005] 提高将被以低的或中等的比特速率传送的音频信号质量的一个方法是集中可用的比特以精确地表示音频信号中的较低频率。因此,使用BWE技术来基于仅需要少量比特的较低频率塑造较高频率。这些技术的背景是人类听觉系统的灵敏度取决于频率。具体而言,人类听觉系统(例如我们的听力)对于较高频率不太精确。

[0006] 在典型的频域BWE方案中,高频变换系数按照频带进行分组。针对每个频带,对增益(能量)进行计算、量化并发送(至信号的解码器)。在解码器端,用高频增益对所接收到的低频系数的翻转或平移(translate)及能量归一化版本进行缩放(scale)。这样,BWE并不完全是“瞎的”,因为至少频谱能量与目标信号的高频频带的频谱能量相似。

[0007] 然而,某些音频信号的BWE可能会导致音频信号包含瑕疵,这对听众带来烦恼。

发明内容

[0008] 本文提出了支持和改善谐波音频信号的BWE的技术。

[0009] 根据本发明的第一方面,提出了一种变换音频解码器中的方法。该方法用于支持谐波音频信号的带宽扩展(BWE)。所提出的方法可以包括与频带b及频带b的多个相邻频带相关的多个增益值的接收。所提出的方法还包括确定带宽扩展频率区域的重建的对应频带b'是否包括谱峰。此外,如果该频带包括至少一个谱峰,该方法包括基于所接收到的多个增益值将与频带b'相关的增益值 G_b 设置为第一值。如果该频带不包括任何谱峰,该方法包括基于所接收到的多个增益值将与频带b'相关的增益值 G_b 设置为第二值。因此使得增益值与频谱的带宽扩展部分中的波峰位置一致。

[0010] 此外,该方法可能包括:接收反映原始信号的高频部分的至少一段的波峰能量与噪声底能量之间关系的参数或系数 α 。该方法还可以包括:基于接收到的系数 α ,将对应的重建的高频段的变换系数与噪声混合。从而使得可以重建/仿真原始信号的高频部分的噪声特性。

[0011] 根据本发明的第二方面,提出了一种支持谐波音频信号的带宽扩展(BWE)的变换音频解码器或编解码器。该变换音频编解码器包括适于执行上面描述的动作的功能单元。此外,提出了一种变换音频编码器或编解码器,包括适于导出或提供一个或多个参数的功能单元,当将所述一个或多个参数提供给变换音频解码器时,能够进行本文描述的噪声混合。

[0012] 根据本发明的第三方面,提出了一种用户终端,其包括根据本发明第二方面的变换音频编解码器。用户终端可能是例如移动终端、平板设备、计算机、智能电话等诸如此类的设备。

附图说明

[0013] 现在将通过示例性的实施例并参照附图的方式更详细地描述本发明,在附图中:

[0014] 图1示出了谐波音频频谱,即谐波音频信号的频谱。这种类型的频谱通常针对于例如单一乐器声音、嗓音等。

[0015] 图2示出了谐波音频频谱的带宽扩展。

[0016] 图3a示出了用解码器接收到的对应BWE频带增益 $\hat{G}_b$ 来进行缩放的BWE频谱(也在图2中示出)。频谱的BWE部分严重失真。

[0017] 图3b示出了本文提出的利用修正的BWE频带增益 $\hat{G}_b^{\text{mod}}$ 来进行缩放的BWE频谱。在这种情况下,频谱的BWE部分得到了所希望的形状。

[0018] 图4a和图4b是示出了根据示例性实施例的变换音频解码器中的过程中的动作的流程图。

[0019] 图5是示出了根据示例性实施例的变换音频解码器的框图。

[0020] 图6是示出了根据示例性实施例的变换音频编码器中的过程中的动作的流程图。

[0021] 图7是示出了根据示例性实施例的变换音频编码器的框图。

[0022] 图8是示出了根据示例性实施例的变换音频解码器中的装置的框图。

具体实施方式

[0023] 如上所述,音频信号的带宽扩展与一些问题相关联。在解码器中,当翻转或平移低频带(即经编码、传输和解码的频带部分)以形成高频带时,不能确定谱峰将会与原始信号或“真实的”高频带中的谱峰结束于(end up)相同的频带中。来自低频带中的谱峰也许会结束于在原始信号没有波峰的频带中。还可能是另一种方式,即,低频信号的没有波峰的部分(在翻转或平移之后)结束于原始信号具有波峰的频带中。图1提供了谐波频谱的示例,图2提供了BWE原理的图示,接下来将对此作进一步的描述。

[0024] 如上描述的效果可能导致具有主要谐波内容的信号产生严重质量下降。原因是波峰和增益位置间的这种不匹配将导致不必要的波峰衰减,或者两个谱峰之间低能量频谱系数的放大。

[0025] 本文描述的方案涉及一种基于与波峰位置有关的信息来控制带宽扩展区域的带宽增益的新方法。此外,本文提出的BWE算法能够通过所发送的噪声混合等级来控制“谱峰噪声底比”。这导致了在扩展的高频中保留大量结构的BWE。

[0026] 本文描述的方案适用于谐波音频信号。图1示出了谐波音频信号的频谱(还可以表示为谐波频谱)。从图中可以看出,频谱包括波峰。这种类型的频谱典型地适用于例如单一乐器(如长笛)的声音或者噪音等。

[0027] 本文将讨论谐波音频信号频谱的两个部分。一个包括较低频率的下部,其中“下”表示低于将要进行带宽扩展的部分;一个包括较高频率的上部,例如比下部高。本文使用的像“下部”或“低/较低频率”之类的表达指的是谐波音频频谱的低于BWE交叉频率的部分(参见图2)。类似地,像“上部”或者“高/较高频率”之类的表达指的是谐波音频频谱的高于BWE交叉频率的部分(参见图2)。

[0028] 图2示出了谐波音频信号的频谱。本文接下来讨论的两个部分,在BWE交叉频率左边的可看作下部,在BWE交叉频率右边的可看作上部。在图2中,用浅灰色示出了原始频谱,即(在编码器端看到的)原始音频信号的频谱。用深/较深灰色示出了频谱的带宽扩展部分。频谱的带宽扩展部分不由编码器进行编码,而是通过使用先前描述的接收到的频谱的下部在解码器端进行重建。在图2中,出于比较的原因,对于较高频率能看到原始(浅灰色)频谱和BWE(深灰色)频谱两者。较高频率的原始频谱对于解码器是未知的,针对于每一个BWE频带(或者高频带)的增益值例外。在图2中,BWE频带用虚线分隔。

[0029] 为了更好地理解在频谱的带宽扩展部分中的增益值和波峰位置之间的不匹配的问题,可以研究图3a。在频带302a,原始频谱包括波峰,但重建的BWE频谱不包括波峰。这在图2的频带202中可以看到。因此,当将针对包括波峰的原始频带计算的增益应用于不包括波峰的BWE频带时,如在频带302a中看到的那样,将放大BWE频带的低能量频谱系数。

[0030] 图3a中的频带304a表示了相反的情况,即,原始频谱的对应频带不包括波峰,但重建的BWE频谱的对应频带包括波峰。因此,针对(从编码器接收到的)频带所获得的增益是针对低能量频带计算的。当将该增益应用于包括波峰的对应频带时,结果变成了衰减的波峰,如在图3a的频带304a中所见。从感知或心理声学的观点来看,因为多种原因,频带302a中示出的情况比频带304a中的情况对听众来说要更糟糕。简单地描述,也就是,与声音分量的异常缺失相比,体验到声音分量的异常出现通常令听众更加不愉快。

[0031] 接下来将描述新的BWE算法的一个示例,以阐述本文所描述的概念。

[0032] 设 $Y(k)$ 表示BWE区域(高频变换系数)中变换系数的集合。将这些变换系数分组到B

个频带 $\{\mathbf{Y}_b\}_{b=1}^B$ 中。频带大小 M_b 可以是恒定的,或者朝着高频增加。例如,如果频带是8次元的且均匀的(即所有的 $M_b=8$),我们得到: $\mathbf{Y}_1 = \{Y(1) \dots Y(8)\}$, $\mathbf{Y}_2 = \{Y(9) \dots Y(16)\}$ 等等。

[0033] BWE算法中的第一步骤是计算所有频带的增益:

$$[0034] \quad G_b = \sqrt{\frac{\mathbf{Y}_b^T \mathbf{Y}_b}{M_b}} \quad (1)$$

[0035] 将这些增益进行量化 $\hat{G}_b = Q(G_b)$, 并发送到解码器。

[0036] BWE算法中的第二步骤(该步骤是可选的)是计算噪声混合参数或系数 α , α 例如是 BWE 频谱的平均波峰能量 $\bar{E}_p$ 和平均噪声底能量 $\bar{E}_{nf}$ 的函数,如:

$$[0037] \quad \alpha = f\left(\frac{\bar{E}_{nf}}{\bar{E}_p}\right) \quad (2)$$

[0038] 在这里,已根据下面的 (3) 导出参数 α 。然而,可以通过不同的方式(如取决于什么适用于所使用的编解码器或量化器的类型等等)来选择所使用的确切表达式。

$$[0039] \quad \alpha = \left(10 \frac{\bar{E}_{nf}}{\bar{E}_p}\right)^3 \quad (3)$$

[0040] 可以例如通过追踪相应的最大和最小频谱能量来计算波峰和噪声底能量。

[0041] 可以使用少量的比特来对噪声混合参数 α 进行量化。在这里,作为示例,用两个比特来量化 α 。当对噪声混合参数 α 进行量化时,得到参数 $\hat{\alpha}$, 例如, $\hat{\alpha} = Q(\alpha)$, 将参数 $\hat{\alpha}$ 发送到解码器。将 BWE 区域分成两个或更多个段 's', 并在这些段的每一个中独立地计算噪声混合参数 α_s 。在这种情况下,编码器将向解码器发送噪声混合参数的集合,例如每个段一个。

[0042] 解码器操作

[0043] 解码器从比特流中提取所计算的量化增益 $\hat{G}_b$ (每个频带一个) 的集合以及一个或更多个量化噪声混合参数或因子 $\hat{\alpha}$ 。解码器还接收针对于频谱的低频部分(即,(谐波音频信号的)频谱的已编码部分,该低频部分与要被带宽扩展的高频部分相对)的量化变换系数。

[0044] 设 $\hat{\mathbf{X}}_b$ 为能量归一化的、量化的低频系数的集合。然后将这些系数与噪声(例如存储在例如噪声码本中的预先生成的噪声 $\mathbf{N}_b$) 进行混合。使用预先生成、预先存储的噪声有机会确保噪声的质量,也就是噪声不包括任何无意的差异和偏差。然而,可以备选地在需要时“抽空”生成噪声。例如将系数 $\hat{\mathbf{X}}_b$ 与噪声码本中的噪声 $\mathbf{N}_b$ 混合如下:

$$[0045] \quad \hat{\mathbf{X}}_b^{\text{mod}} = (1 - \hat{\alpha}) \hat{\mathbf{X}}_b + \hat{\alpha} \mathbf{N}_b \quad (4)$$

[0046] 可以通过不同的方式设置噪声混合参数或因子的范围。例如,这里将噪声混合因子的范围设置为 $\alpha \in [0, 0.4)$ 。这个范围意味着例如在某些情况下噪声贡献被完全忽略 ($\alpha = 0$), 在某些情况下噪声码本在混合矢量中贡献 40% ($\alpha = 0.4$), 这是使用该范围时的最大贡

献。引入这种类型的噪声混合(所得矢量包含例如60%至100%之间的原始低频带结构)的原因是频谱的高频部分通常比频谱的低频部分有更多噪声。因此,与由翻转或平移的低频频谱区域构成的BWE高频频谱区域相比,上面描述的噪声混合操作创建了能够更好地拟合原始信号频谱的高频部分的统计特征的矢量。例如,如果提供并接收到多个噪声混合因子(α),可以在BWE区域的不同部分上独立地执行噪声混合操作。

[0047] 在现有技术方案中,将接收到的量化的增益 $\hat{G}_b$ 的集合直接用于BWE区域的对应频带。然而,根据本文所描述的方案,例如在适当的时候首先基于有关BWE频谱波峰位置的信息来修正这些接收到的量化的增益 $\hat{G}_b$ 。所需要的关于波峰位置的信息可以从比特流中的低频区域信息中提取,或者由波峰抓取算法基于低频带的量化变换系数(或者所导出的BWE频带的系数)进行估计。然后将低频区域中与波峰相关的信息转移到高频(BWE)区域中。也就是说,当从低频带信号推导高频带(BWE)信号时,该算法能够登记谱峰位于(BWE区域的)哪些频带。

[0048] 例如,可以使用标志 $f_p(b)$ 来指示移动(翻转或平移)到BWE区域的频带b的低频系数是否包含波峰。例如, $f_p(b) = 1$ 指示频带b至少包含一个波峰, $f_p(b) = 0$ 指示频带b不包含任何波峰。如前所述,BWE区域中的每一个频带b都与增益 $\hat{G}_b$ 相关联, $\hat{G}_b$ 取决于原始信号的对应频带中包括的波峰的数目和大小。为了使增益与BWE区域中每个频带的实际波峰内容相匹配,需使增益相适应。例如根据下面的表达式对每个频带作增益修正:

$$[0049] \quad \hat{G}_b^{\text{mod}} = \begin{cases} \frac{1}{3}(\hat{G}_{b-1} + \hat{G}_b + \hat{G}_{b+1}) & \text{如果 } f_p(b) = 1 \\ \min\{\hat{G}_{b-1}, \hat{G}_b, \hat{G}_{b+1}\} & \text{如果 } f_p(b) = 0 \end{cases} \quad (5a)$$

[0050] 作此增益修正的动机如下:在(BWE)频带包含波峰($f_p(b) = 1$)的情况下,为了避免对应的增益来自(原始信号的)没有任何波峰的频带时的波峰衰减,将此频带的增益修正为当前频带及两个相邻频带的增益的加权和。在上面的示例性公式(5a)中,权重是相等的(即1/3),这导致修正的增益为当前频带的增益及两个相邻频带的增益的均值。例如可根据下式实现备选增益修改:

$$[0051] \quad \hat{G}_b^{\text{mod}} = \begin{cases} (0.1\hat{G}_{b-1} + 0.8\hat{G}_b + 0.1\hat{G}_{b+1}) & \text{如果 } f_p(b) = 1 \\ \min\{\hat{G}_{b-1}, \hat{G}_b, \hat{G}_{b+1}\} & \text{如果 } f_p(b) = 0 \end{cases} \quad (5b)$$

[0052] 在频带不包含波峰($f_p(b) = 0$)的情况下,我们不希望通过应用从包含一个或更多个波峰的原始信号中计算出的强增益来放大该频带中的类似噪声的结构。为了避免这种情况,选择例如当前频带增益及两个相邻频带增益的最小值作为该频带的增益。备选地,可以将包括波峰的频带的增益选择或计算为3个以上频带(如5或7个频带)的加权和(例如,均值),或者选择为如3、5或7个频带的中间值。与使用“真实的”增益相比,通过使用加权和(例如均值或中间值),波峰也许会轻微衰减。然而,与相反的情况相比,与“真实的”增益相比的

衰减或许是有益的,如前所述,因为从感知的观点来看,与导致过大音频分量的放大相比,适度的衰减更好。

[0053] 波峰不匹配的原因且因此也是增益修正的原因是频谱带位于预定的网格上,但是(翻转和平移低频系数之后的)波峰位置和波峰是随时间变化的。这可能导致波峰以不可控的方式进入或移出频带。因此因此,频谱的BWE部分的波峰位置不一定与原始信号中的波峰位置相匹配,因此与频带相关联的增益和频带的波峰内容之间可能会存在不匹配。图3a示出了用未修正的增益进行缩放的示例,图3b示出了用修正的增益进行缩放的示例。

[0054] 在图3b中可以看出使用本文提出的修正的增益的结果。在频带302b中,低能量频谱系数不再像在图3a的频带302a中那样被放大,而是用更合适的带宽增益进行缩放。此外,频带304b中的波峰不再像在图3a的频带304a中那样被衰减。与对应于图3a的频谱的音频信号相比,图3b中示出的频谱很可能对应于对于听众来说更令人愉快的音频信号。

[0055] 因此,BWE算法可创建频谱的高频部分。因为(例如为了节省带宽的原因)在解码器处高频系数 Y_b 不可用,所以作为替代,通过使用修正的量化增益缩放经翻转(或平移)的低频系数(可能在噪声混合之后)来重建或形成高频变换系数 $\tilde{Y}_b$:

$$[0056] \quad \tilde{Y}_b = \hat{G}_b^{\text{mod}} \hat{X}_b^{\text{mod}} \quad (6)$$

[0057] 该变换系数 $\tilde{Y}_b$ 的集合用于重建音频信号波形的高频部分。

[0058] 本文描述的方案是BWE原理的改进,通常用于变换域音频编码。所提出的算法保留了BWE区域中的多峰结构(波峰噪声底比),因此提供了重建信号的改进的音频质量。

[0059] 术语“变换音频编解码器”或“变换编解码器”包含编码器-解码器对,是该领域中的常用术语。在本发明的公开中,为了分别描述变换编解码器的功能/部件,使用了术语“变换音频编码器”或“编码器”和“变换音频解码器”或“解码器”。因此,可以把术语“变换音频编码器”/“编码器”和“变换音频解码器”/“解码器”与术语“变换音频编解码器”或“变换编解码器”相交换。

[0060] 解码器中的示例性过程,图4a和图4b

[0061] 下面将参照图4a描述在解码器中支持谐波音频信号的带宽扩展(BWE)的示例性过程。该过程适用于变换音频编码器(例如MDCT编码器)或其它编码器。音频信号主要包括音乐,还可以或者备选地包括例如语音。

[0062] 在动作401a中,接收到与频带b(原始频带)有关的增益值以及和与频带b相邻的多个其他频带有关的增益值。然后在动作404a中确定BWE区域的重建的对应频带b'是否包括谱峰。当重建的频带b'至少包括一个谱峰时,在动作406a:1中,基于所接收到的多个增益值,将与重建的频带b'相关的增益值设置为第一值。当重建的频带b'不包括任何谱峰时,在动作406a:2中,基于所接收到的多个增益值,将与重建的频带b'相关的增益值设置为第二值。第二值小于等于第一值。

[0063] 在图4b中,以略有不同及更加扩展的方式(例如具有与先前描述的噪声混合有关的附加可选动作),示出了图4a中所示的过程。下面将描述图4b。

[0064] 在动作401b中接收到与频谱的上部的频带有关的增益值。假定还在某一时间点接收到涉及频谱的下部的信息(在图4a或图4b中未示出),即变换系数和增益值等等。此外,假

定在某个时间点执行带宽扩展,通过如前所述的翻转或平移低频带频谱来创建高频带频谱。

[0065] 可以在可选动作402b中接收一个或更多个噪声混合系数。所接收的一个或更多个噪声混合系数已在编码器中基于原始高频带频谱中的能量分布计算出。在(同样可选的)动作403b中,参见上面的公式(4),使用噪声混合系数将高频带区域中的系数与噪声混合。因此,对于“噪声特性”或“噪声分量”,带宽扩展区域的频谱将更好地对应原始高频带频谱。

[0066] 此外,在动作404b中,确定所创建的BWE区域的频带是否包括谱峰。例如,如果频带包括谱峰,可以将与该频带相关联的指示符设置为1。如果另一个频带不包括谱峰,可以将与那个频带相关联的指示符设置为0。基于频带是否包括谱峰的信息,在动作405b中修正与所述频带有关的增益。如前所述,当修正频带的增益时,为了达到想要的结果,也要考虑相邻频带的增益。通过这种方式修正增益,能够得到改进的BWE频谱。如在动作406b所示的,然后将修正的增益应用于BWE频谱的各个频带。

[0067] 示例性解码器

[0068] 下面,将参照图5描述适于执行支持谐波音频信号的带宽扩展(BWE)的上述过程的示例性变换音频解码器。变换音频解码器可以是例如MDCT解码器,或其它解码器。

[0069] 将变换音频解码器501示为经由通信单元502与其它实体进行通信。将虚线包围的变换音频解码器的部分示为装置500,该部分适于实现上述过程的性能。变换音频解码器还可以包括其它功能单元516,例如提供常规解码器和BWE功能的功能单元,并还可以包括一个或更多个存储单元514。

[0070] 可以通过例如以下各项中的一个或更多个来实现变换音频解码器501和/或装置500:处理器或微处理器及带有合适存储设备的适当软件、可编程逻辑设备(PLD)或其它电子元件。

[0071] 假定变换音频解码器包括用于获得从编码实体提供的适当参数的功能单元。与现有技术相比,噪声混合系数是要获得的新的参数。因此,解码器应当适于使得需要一个或更多个噪声混合系数时能够获取该一个或更多个噪声混合系数。将音频解码器描述和实现为包括接收单元,该接收单元适于接收与频带b及频带b的多个相邻频带相关联的多个增益值;或许还接收噪声混合系数。然而,图5中并未显式地示出这种接收单元。

[0072] 变换音频解码器包括确定单元504,也可以称作波峰检测单元,该单元适于确定和指示BWE频谱区域的哪些频带包括波峰及哪些频带不包括波峰。也就是说确定单元适于确定带宽扩展频率区域的重建的对应频带b'是否包括谱峰。此外,变换音频解码器包括增益修正单元506,该单元适于根据频带是否包括波峰来修正与频带相关联的增益。如果频带包括波峰,将修正增益计算为加权和,例如与所讨论的频带相邻的多个频带的(原始)增益(包括所讨论的频带的增益)的均值或中间值。

[0073] 变换音频解码器还包括适于将修正增益应用或设置到BWE频谱的适当频带的增益应用单元508。也就是说,增益应用单元适于:当重建的频带b'包括至少一个谱峰时,基于接收到的多个增益值将与重建的频带b'相关联的增益值设置为第一值;以及当重建的频带b'不包括任何谱峰时,基于接收到的多个增益值将与重建的频带b'相关联的增益值设置为第二值,其中第二值小于等于第一值。因此使得增益值与带宽扩展频率区域中的波峰位置一致。

[0074] 备选地,如果可能不作修改,可以由(常规的)其他功能516提供应用功能,只是所应用的增益不是原始增益,而是修正的增益。此外,变换音频解码器包括噪声混合单元510,该噪声混合单元510适于基于音频信号的编码器提供的一个或更多个噪声系数或参数来将频谱的BWE部分的系数与(例如来自码本的)噪声进行混合。

[0075] 示例性过程编码器

[0076] 下面将参照图6描述在编码器中支持谐波音频信号的带宽扩展(BWE)的示例性过程。该过程适用于变换音频编码器(例如MDCT编码器)或其它编码器。如上文所述,音频信号被主要认为包括音乐,也可以或者备选地包括例如语音。

[0077] 下面描述的过程涉及背离使用变换编码器的谐波音频信号的传统编码方法的编码过程部分。因此,下面描述的动作对于用于频谱下部的变换系数和增益等的获取和频谱上部的频带(该部分将在解码器端由BWE构建)的增益的获取是可选的附加动作。

[0078] 在动作602中,确定涉及频谱上部的波峰能量。此外,在动作603中,确定涉及频谱上部的噪声底能量。例如,如上所述,计算BWE频谱的一个或更多个段的平均波峰能量 $\overline{E}_p$ 和平均噪声底能量 $\overline{E}_{nf}$ 。此外,在动作604中,根据一些适当的公式,如上述的公式(3)计算噪声混合系数,使得与BWE频谱某些段有关的噪声系数反映所述段的噪声量或“噪声特性”。在动作606中,向解码实体或存储器提供一个或更多个噪声混合系数以及由编码器提供的常规信息。所述提供包括例如仅仅将计算的噪声混合系数输出到输出端,和/或例如将该系数发送到解码器。如前所述,在提供噪声混合系数之前,可先对其进行量化。

[0079] 示例性编码器

[0080] 下面将参照图7描述适合执行支持谐波音频信号的带宽扩展(BWE)的上述过程的示例性变换音频解码器。变换音频解码器可以是例如MDCT解码器或其它解码器。

[0081] 将变换音频解码器701示为经由通信单元702与其它实体进行通信。将虚线包围的变换音频解码器的部分示为装置700,该部分适于实现上述过程中的性能。变换音频解码器还可以包括其它功能单元712,例如提供常规编码功能的功能单元,并且还包括一个或更多个存储单元710。

[0082] 变换音频编码器701和/或装置700可以通过例如以下各项中的一项或更多项来实现:处理器或微处理器及带有合适存储设备的适当软件、可编程逻辑设备(PLD)或其它电子元件。

[0083] 变换音频编码器可以包括确定单元704,该确定单元704适于确定频谱上部的波峰能量和噪声底能量。此外,变换音频编码器可以包括噪声系数单元706,该噪声系数单元706适于对于频谱的整个上部或其中的段计算一个或更多个噪声混合系数。变换音频编码器还包括提供单元708,适于提供由编码器使用的所计算的噪声混合系数。所述提供可以包括例如仅将计算的噪声混合系数输出到输出端,和/或将该系数发送到解码器。

[0084] 示例性装置

[0085] 图8示意性地示出了适于在变换音频解码器中使用的装置800的实施例,其还可以是公开图5示出的变换音频解码器中使用的装置的实施例的备选方法。在此,在装置800中包括例如带有DSP(数字信号处理器)的处理单元806。处理单元806可以是单一的单元或者是执行本文描述的过程的不同步骤的多个单元。装置800还包括用于接收信号(例如经编码

的频谱的下部、整个频谱的增益和噪声混合系数(参见,如果是编码器:谐波频谱的上部))的输入单元802和用于输出信号(例如修正的增益和/或整个频谱(参见,如果是编码器:噪声混合系数))的输出单元804。可以在装置的硬件中将输入单元802和输出单元804安排为同一个。

[0086] 此外,装置800包括以诸如EEPROM、闪存和硬盘等非易失性或易失性存储器形式的至少一个计算机程序产品808。计算机程序产品808包括计算机程序810,计算机程序810包括代码,当该代码在装置800的处理单元806中运行时,使装置和/或变换音频编码器执行先前结合图4描述的过程的动作。

[0087] 因此,在描述的示例性中,装置800的计算机程序810中的代码可以包括用于获得涉及音频频谱下部的信息和涉及整个音频频谱的增益的获得模块810a。此外,还可以获得涉及音频频谱上部的噪声系数。计算机程序包括检测模块810b,检测模块810b用于检测和指示带宽扩展频率区域的重建频带b的频带是否包括谱峰。计算机程序810还可以包括增益修正模块810c,用于修正与频谱的重建的上部的频带相关联的增益。计算机程序810还可以包括用于将修正的增益应用于频谱上部的对应频带的增益应用模块810d。此外,计算机程序810可以包括用于基于所接收到的噪声混合系数将频谱的上部与噪声进行混合的噪声混合模块810d。

[0088] 计算机程序810是以计算机程序模块的方式构成的计算机程序代码的形式。模块810a-d实质上执行图4a或图4b中示出的过程的动作以模仿图5中示出的装置500。换言之,当在处理单元806中运行不同的模块810a-d时,它们至少对应于图5的单元504-510。

[0089] 尽管上述结合图8公开的实施例的代码被实现为计算机程序模块,当程序在处理单元中运行时,使装置和/或变换音频编码器执行结合上述附图描述的上述步骤,在备选实施例中,代码中的至少一个可以至少部分地实现为硬件电路。

[0090] 类似地,将包括计算机程序模块的示例性实施例描述为图7中示出的变换音频编码器中的对应装置。

[0091] 虽然本发明是参考某些的示例实施例来描述的,但是这里的描述一般仅旨在说明本发明的概念,而不应视为限制本发明的范围。可以根据需要、需求或偏好以不同方式组合上面的示例性实施例的不同特征。

[0092] 任何应用音频编解码器的地方均可使用上面描述的方案,例如用在像移动终端、平板设备、计算机、智能电话等设备中。

[0093] 应当理解的是,交互单元或模块的选择以及单元的命名仅仅用于示例目的,可以以多种可替代方式配置适于执行上述任一方法的节点,以执行所提出的过程动作。

[0094] 还应注意:本公开中描述的单元或模块应被视为逻辑实体,而不一定作为单独的物理实体。虽然上面的说明包含许多某些术语,但这些术语不应视为限制本公开的范围,而应视为在此只是提供本发明的一些现有优先实施例的说明。将理解的是,本文本发明的范围完全涵盖本领域技术人员显而易见的其它实施例,并且本公开的范围对应地未受到限制。除非明确说明,否则,对单数形式的单元的引用并不是要表示“一个和只有一个”,而是“一个或更多个”。通过引用将本领域技术人员已知的上述实施例的单元的所有结构和功能等效物明确并入本文,并且由此涵盖它们。另外,本发明所涵盖的设备或方法不一定要解决本发明在此寻求解决的每个问题。

[0095] 在前面的说明中,为便于解释而不是限制,陈述了某些的细节,如某些的架构、接口、技术等,以便提供详尽理解。然而,本领域的技术人员将明白,本发明可以脱离这些某些细节的方式实践。也就是说,本领域技术人员将能够设计各种布置,这些布置虽然在本文中未明确描述或示出,但采用本发明的原理。一些情况下,忽略了熟知的设备、电路和方法的详细描述以免不必要的细节混淆本发明的描述。本文中叙述原理、方法和本发明实施例及其某些示例的所有声明旨在涵盖其结构和功能等效物,另外,此类等效物当前已知等效物和将来形成的等效物,即,开发的执行相同功能的任何单元,而无论结构如何。

[0096] 因此,例如,本领域技术人员将理解的是,本文中的框图可表示采用技术的原理的说明性电路或其它功能单元的概念视图。类似地,将理解的是任何流程图、状态转换图、伪代码及诸如此类表示各过程,这些过程实质上可在计算机可读媒体中表示并因此由计算机或处理器执行,而无论是否明确示出了这类计算机或处理器。

[0097] 包括功能块在内的各种单元的功能可通过使用诸如电路硬件等硬件和/或能够执行在计算机可读媒体上存储的编码指令形式的软件的硬件提供,功能块包括但不限于标记或描述为“功能单元”、“处理器”或“控制器”。因此,此类功能和所示功能块要理解为是硬件实现和/或计算机实现,并且因此是机器实现。

[0098] 就硬件实现而言,功能可包括或涵盖但不限于数字信号处理器 (DSP) 硬件、精简指令集处理器、包括但不限于专用集成电路 (ASIC) 的硬件 (例如,数字的或模拟的) 电路、以及 (在适当之处) 能够执行此类功能的状态机。

[0099] 缩略语

[0100] BWE 带宽扩展

[0101] DFT 离散傅里叶变换

[0102] DCT 离散余弦变换

[0103] MDCT 修正离散余弦变换

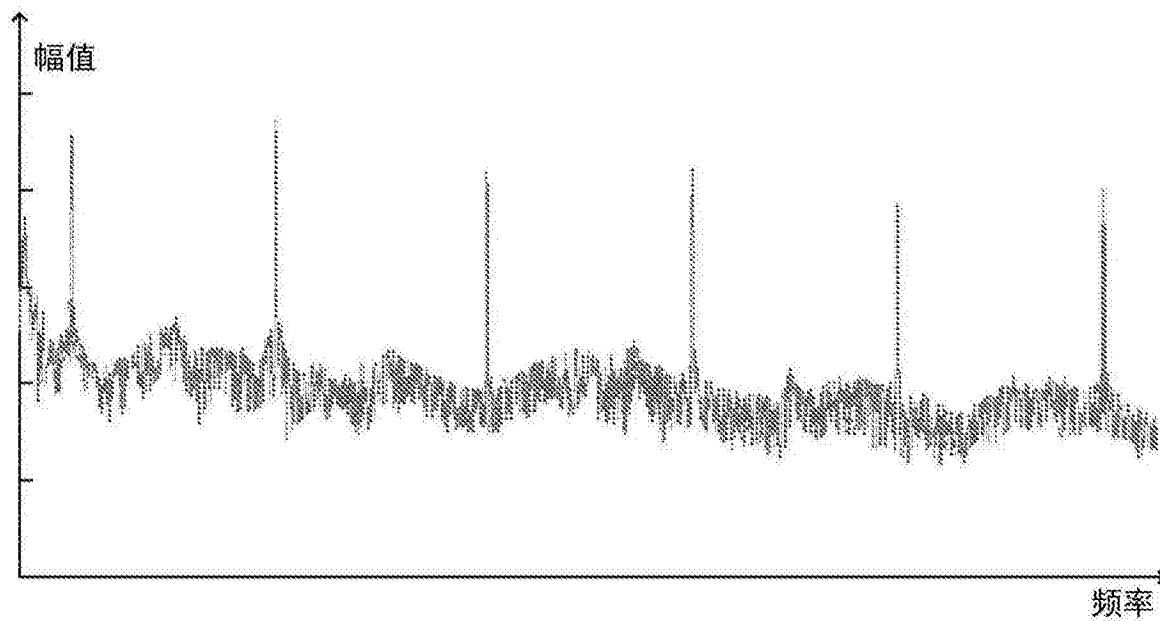


图1

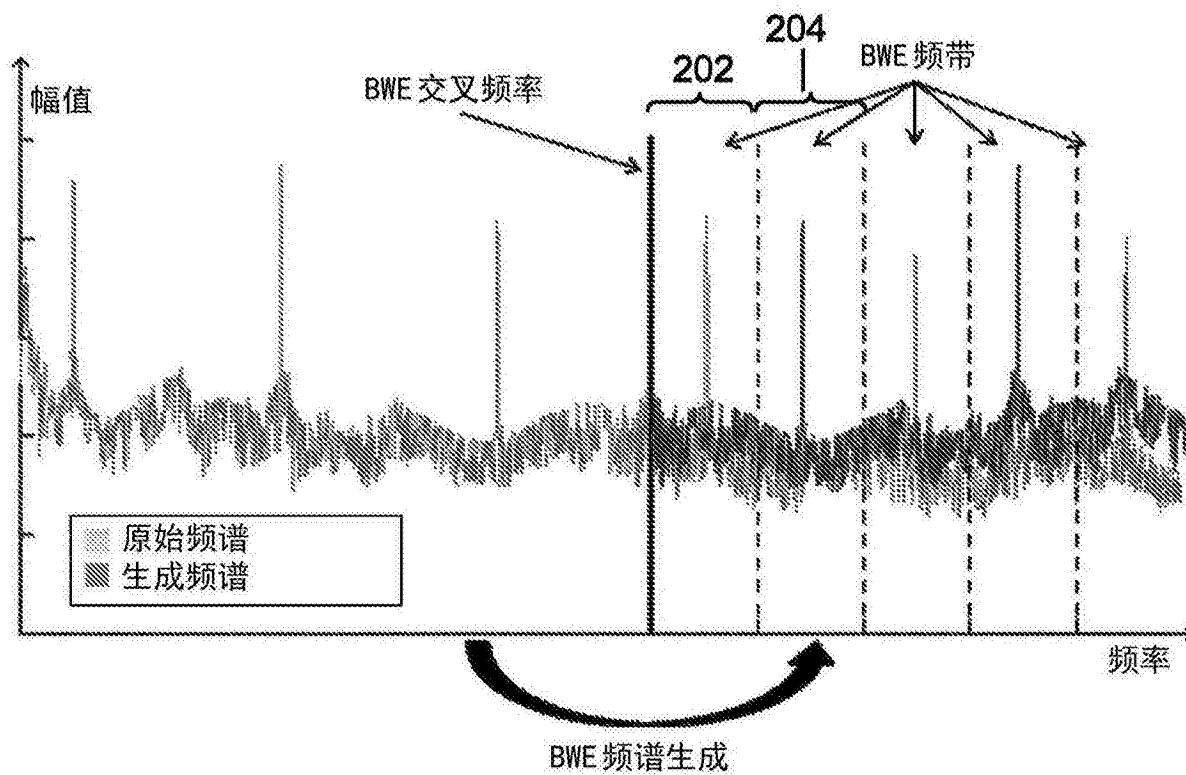


图2

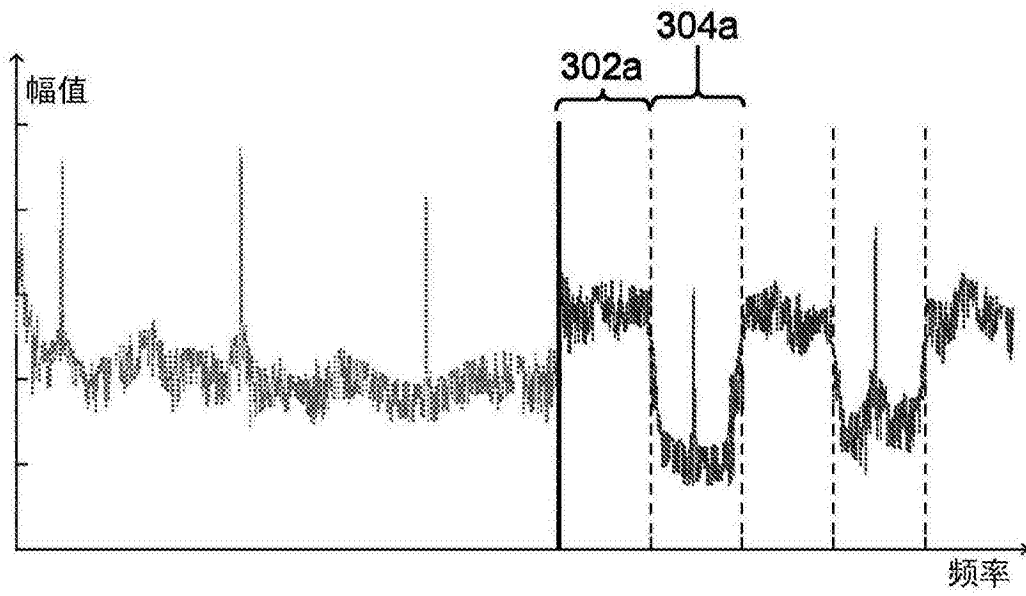


图3a

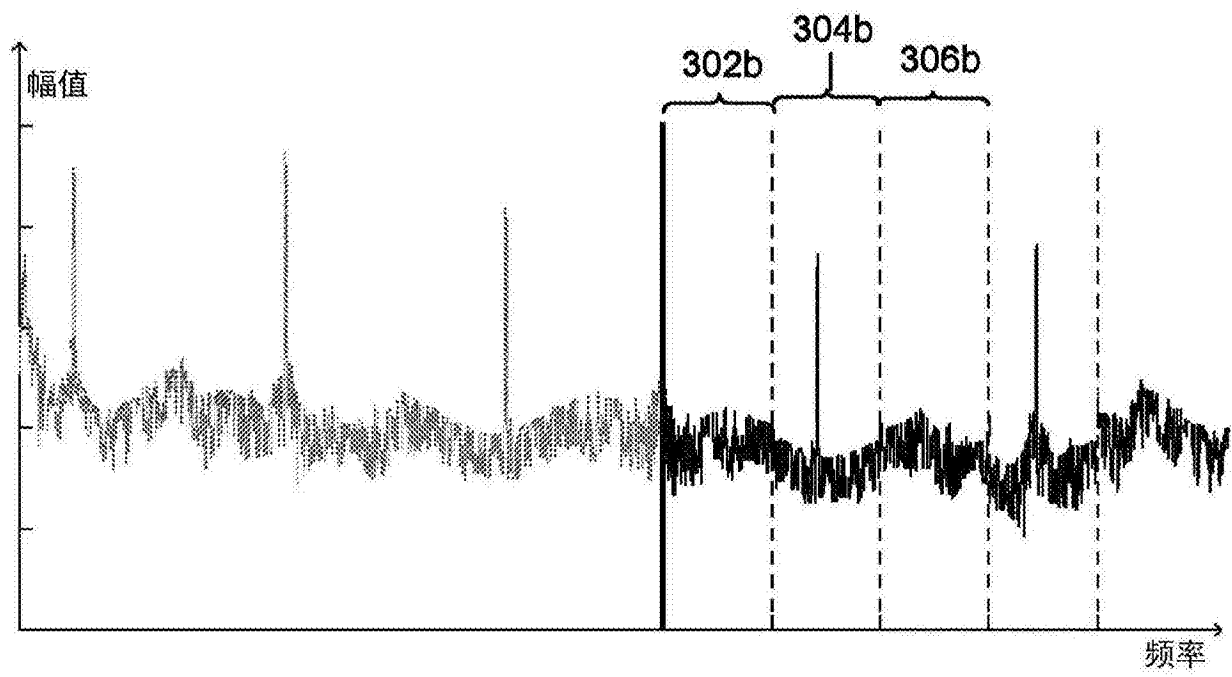


图3b

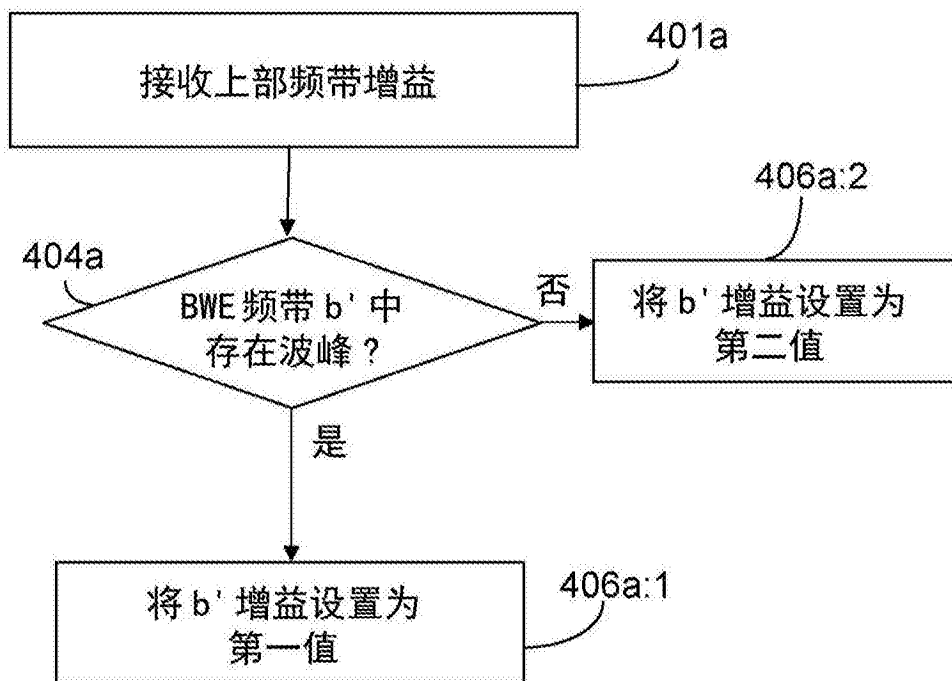


图4a

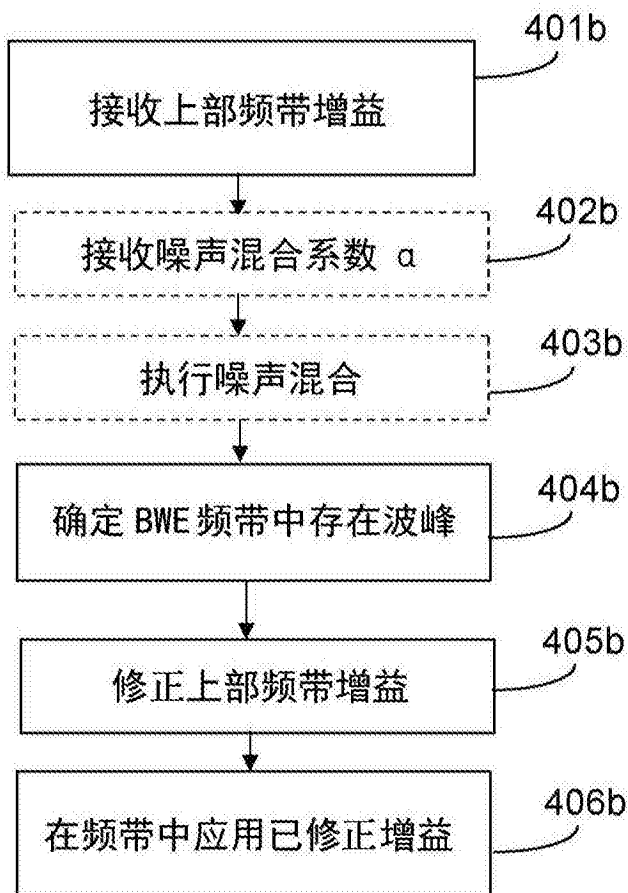


图4b

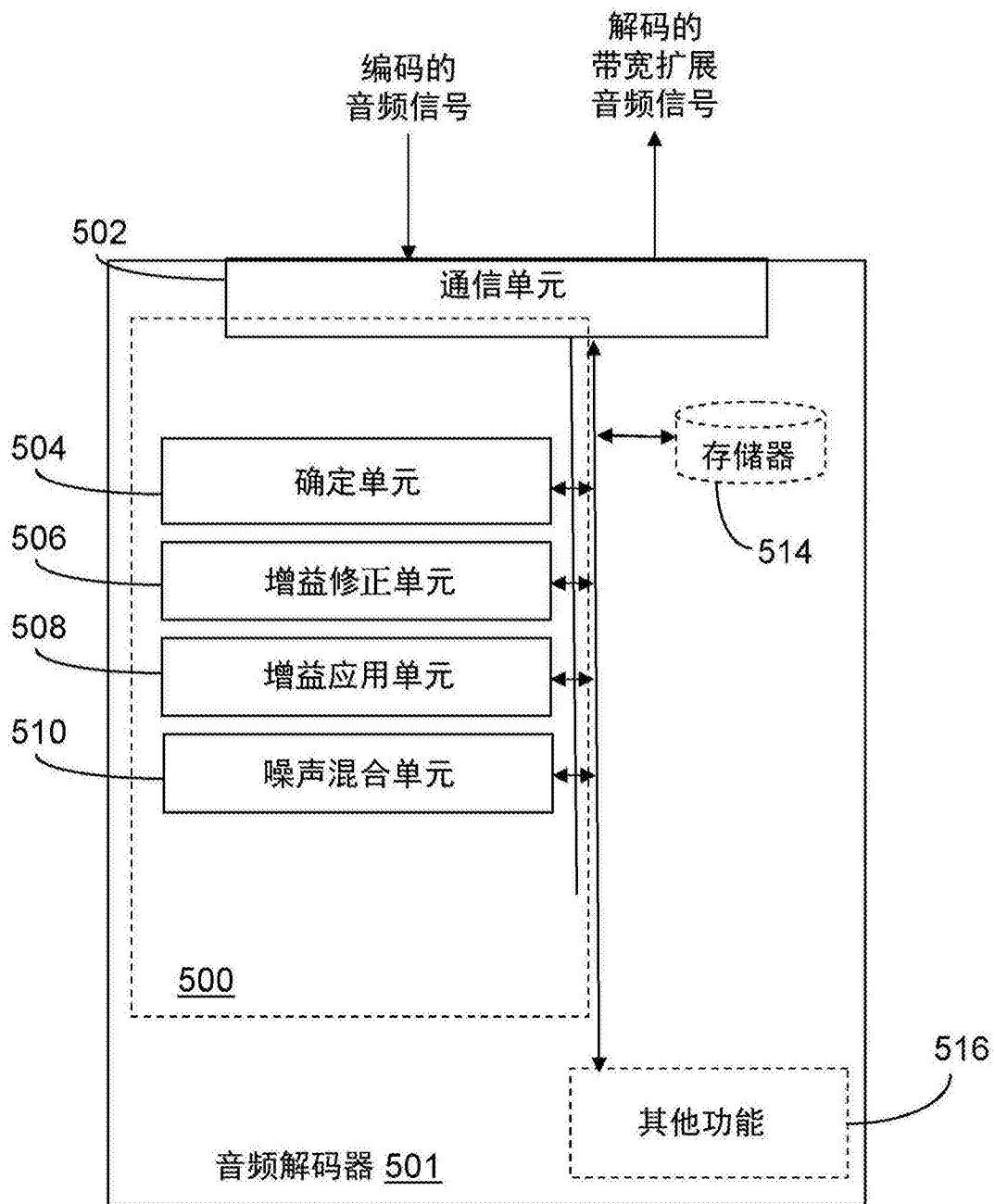


图5

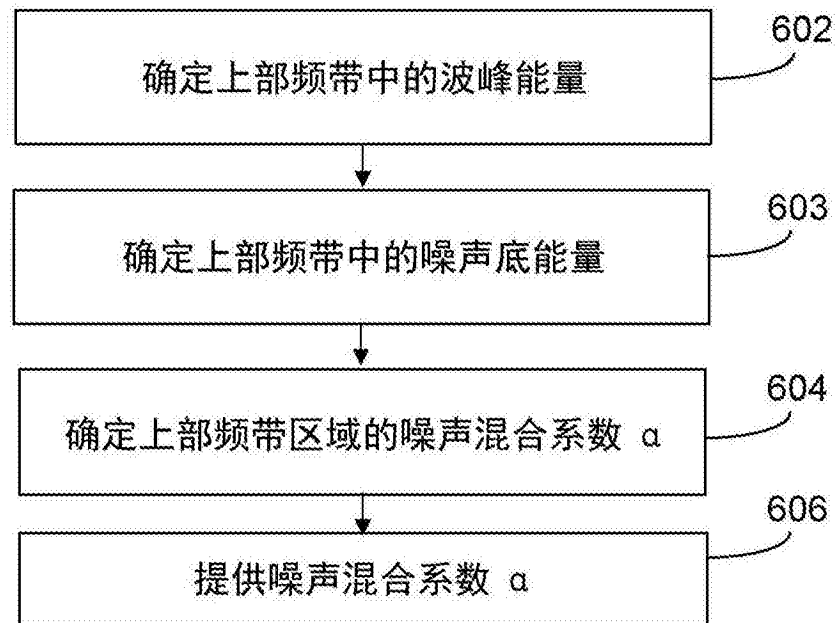


图6

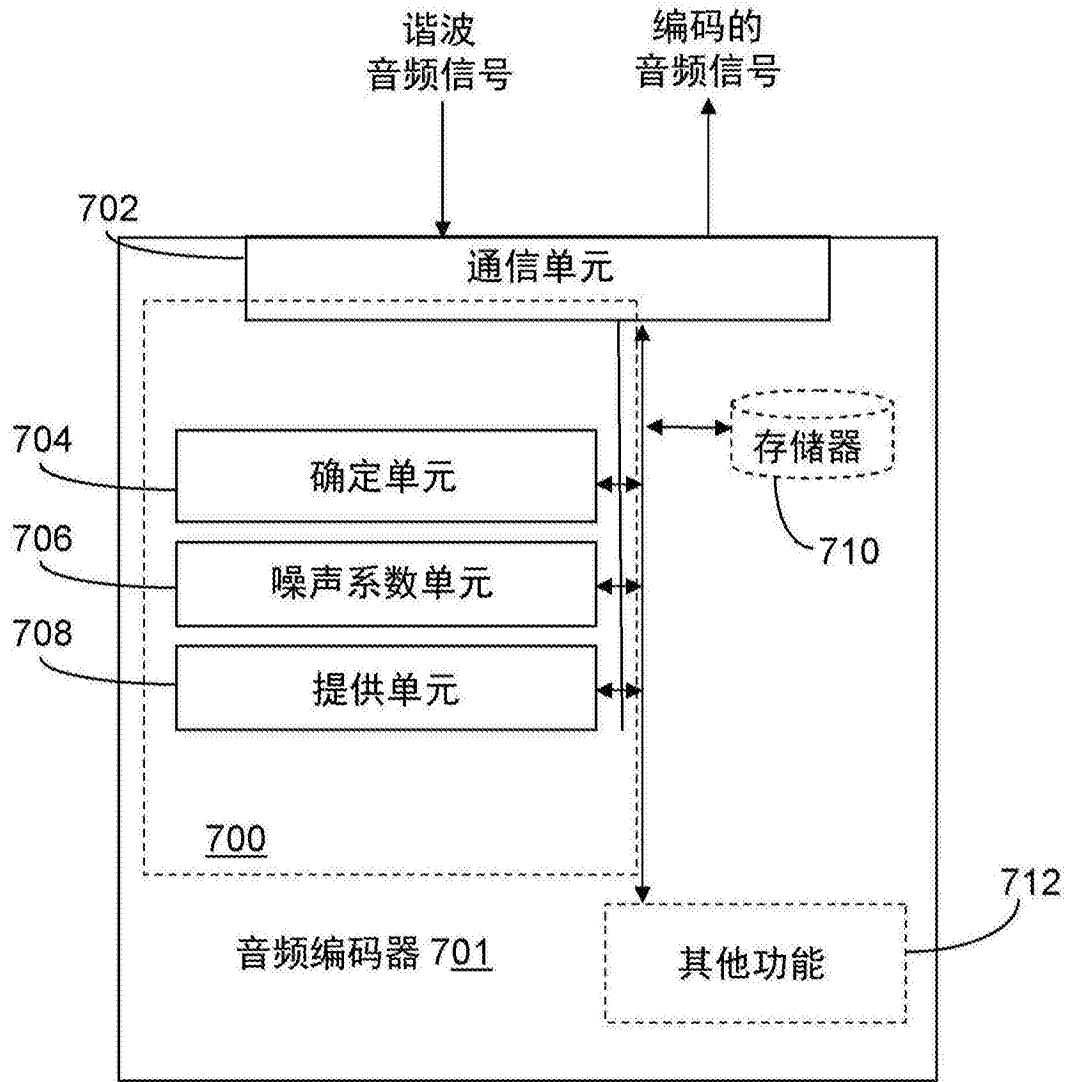


图7

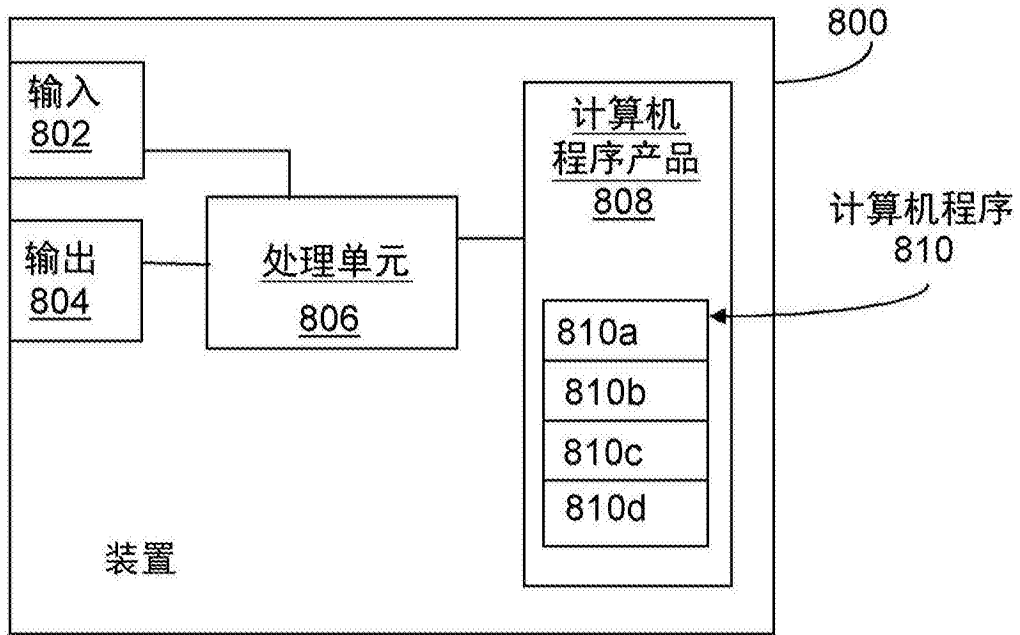


图8