



(12) 发明专利申请

(10) 申请公布号 CN 102396024 A

(43) 申请公布日 2012. 03. 28

(21) 申请号 201080016829. 0

代理人 李芳华

(22) 申请日 2010. 02. 16

(51) Int. Cl.

(30) 优先权数据

G10L 19/00 (2006. 01)

10-2009-0012356 2009. 02. 16 KR

10-2009-0092717 2009. 09. 29 KR

(85) PCT申请进入国家阶段日

2011. 10. 17

(86) PCT申请的申请数据

PCT/KR2010/000955 2010. 02. 16

(87) PCT申请的公布数据

W02010/093224 KO 2010. 08. 19

(71) 申请人 韩国电子通信研究院

地址 韩国大田市

(72) 发明人 李美淑 裴贤珠 李炳培

(74) 专利代理机构 北京市柳沈律师事务所
11105

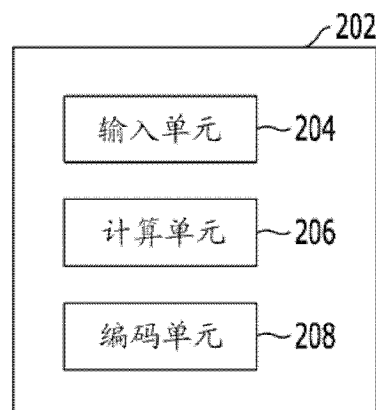
权利要求书 3 页 说明书 10 页 附图 7 页

(54) 发明名称

使用自适应正弦波脉冲编码的用于音频信号的
编码 / 解码方法及其设备

(57) 摘要

本发明涉及一种使用自适应正弦波形脉冲编码的用于音频信号的编码 / 解码方法及其设备。根据本发明实施例的用于音频信号的编码方法包括 : 将音频信号划分为多个子带的步骤 ; 单独地操作每个子带的能量的步骤 ; 按照预定数目来从所有子带之中选择具有相对较大能量的子带的步骤 ; 以及针对所选择的子带来执行整形波形脉冲编码的步骤。本发明具有以下效果, 即通过在考虑到合成信号的每个子带的能量数量的情况下应用正弦波形脉冲编码来有效地增加合成信号的质量。



1. 一种用于对音频信号进行编码的方法,包括:
接收所合成的音频信号;
将所合成的音频信号划分为多个子带;
计算所述子带中的每个子带的能量;
从所述子带中选择具有相对大量能量的预定数目的子带;以及
对于所选择的子带执行正弦脉冲编码。
2. 根据权利要求1的方法,其中通过要在一个子带中进行编码的正弦脉冲的数目和用于对一个正弦脉冲进行编码所必需的比特数目,来确定该预定数目。
3. 根据权利要求1的方法,其中所述对于所选择的子带执行正弦脉冲编码的步骤包括:
确认在所选择的子带之中是否存在相邻的子带;
将所述相邻的子带合并为一个子带;以及
对于所合并的子带执行正弦脉冲编码。
4. 根据权利要求1的方法,其中在所述对于所选择的子带执行正弦脉冲编码的步骤中,
按照能量的数量的顺序,对于所选择的子带执行正弦脉冲编码。
5. 一种用于对音频信号进行编码的设备,包括:
输入单元,被配置为接收所转换的音频信号;
计算单元,被配置为将所转换的音频信号划分为多个子带,计算所述子带中的每个子带的能量,并且从所述子带中选择具有相对大量能量的预定数目的子带;以及
编码单元,被配置为对于所选择的子带执行正弦脉冲编码。
6. 根据权利要求5的设备,其中该计算单元被配置为基于要在一个子带中进行编码的正弦脉冲的数目和用于对一个正弦脉冲进行编码所必需的比特数目,来确定该预定数目。
7. 根据权利要求5的设备,其中该计算单元被配置为确认在所选择的子带之中是否存在相邻的子带,并且将所述相邻的子带合并为一个子带,并且
该编码单元被配置为对于所合并的子带执行正弦脉冲编码。
8. 根据权利要求5的设备,其中该编码单元被配置为按照能量的数量的顺序,对于所选择的子带执行正弦脉冲编码。
9. 一种用于对音频信号进行解码的方法,包括:
接收所转换的音频信号;
将所转换的音频信号划分为多个子带;
计算所述子带中的每个子带的能量;
从所述子带中选择具有相对大量能量的预定数目的子带;以及
对于所选择的子带执行正弦脉冲解码。
10. 根据权利要求9的方法,其中通过要在一个子带中进行编码的正弦脉冲的数目和用于对一个正弦脉冲进行编码所必需的比特数目,来确定该预定数目。
11. 根据权利要求9的方法,其中在所述对于所选择的子带执行正弦脉冲解码的步骤中,
按照能量的数量的顺序,对于所选择的子带执行正弦脉冲解码。

12. 一种用于对音频信号进行解码的设备,包括:

输入单元,被配置为接收所转换的音频信号;

计算单元,被配置为将所转换的音频信号划分为多个子带,计算所述子带中的每个子带的能量,并且从所述子带中选择具有相对大量能量的预定数目的子带;以及

解码单元,被配置为对于所选择的子带执行正弦脉冲解码。

13. 根据权利要求 12 的设备,其中该计算单元被配置为基于要在一个子带中进行编码的正弦脉冲的数目和用于对一个正弦脉冲进行编码所必需的比特数目来确定该预定数目。

14. 根据权利要求 12 的设备,其中该解码单元被配置为按照能量的数量的顺序,对于所选择的子带执行正弦脉冲解码。

15. 一种用于对音频信号进行编码的方法,包括:

接收音频信号;

对于该音频信号执行改进的离散余弦变换(MDCT),以输出 MDCT 系数;

使用该 MDCT 系数来对高频音频信号进行合成;以及

对于该高频音频信号执行正弦脉冲编码。

16. 根据权利要求 15 的方法,其中所述对于高频音频信号执行正弦脉冲编码的步骤包括:

将该高频音频信号划分为多个子带;

计算所述子带中的每个子带的能量;

从所述子带中选择具有相对大量能量的预定数目的子带;以及

对于所选择的子带执行正弦脉冲编码。

17. 一种用于对音频信号进行编码的设备,包括:

输入单元,被配置为接收音频信号;

MDCT 单元,被配置为对于该音频信号执行 MDCT,以输出 MDCT 系数;

合成单元,被配置为使用该 MDCT 系数来对高频音频信号进行合成;以及

正弦脉冲编码单元,被配置为对于该高频音频信号执行正弦脉冲编码。

18. 根据权利要求 17 的设备,其中该正弦脉冲编码单元被配置为将该高频信号划分为多个子带,计算所述子带中的每个子带的能量,从所述子带中选择具有相对大量能量的预定数目的子带,并且对于所选择的子带执行正弦脉冲编码。

19. 一种用于对音频信号进行解码的方法,包括:

接收音频信号;

对于该音频信号执行 MDCT,以输出 MDCT 系数;

使用该 MDCT 系数来对高频音频信号进行合成;以及

对于该高频音频信号执行正弦脉冲解码。

20. 根据权利要求 19 的方法,其中所述对于高频音频信号执行正弦脉冲解码的步骤包括:

将该高频信号划分为多个子带;

计算所述子带中的每个子带的能量;

从所述子带中选择具有相对大量能量的预定数目的子带;以及

对于所选择的子带执行正弦脉冲解码。

21. 一种用于对音频信号进行解码的设备,包括:

输入单元,被配置为接收音频信号;

MDCT 单元,被配置为对于该音频信号执行 MDCT,以输出 MDCT 系数;

合成单元,被配置为使用该 MDCT 系数来对高频音频信号进行合成;以及

正弦脉冲解码单元,被配置为对于该高频音频信号执行正弦脉冲解码。

22. 根据权利要求 21 的设备,其中该正弦脉冲解码单元被配置为将该高频信号划分为多个子带,计算所述子带中的每个子带的能量,从所述子带中选择具有相对大量能量的预定数目的子带,并且对于所选择的子带执行正弦脉冲解码。

使用自适应正弦波脉冲编码的用于音频信号的编码 / 解码 方法及其设备

技术领域

[0001] 本发明的示范实施例涉及一种用于对音频信号进行编码和解码的方法和设备 ; 且更具体地, 涉及一种用于使用自适应正弦脉冲编码来对音频信号进行编码和解码的方法和设备。

背景技术

[0002] 由于用于数据传送的带宽随着通信技术的发展而增加, 所以用户对于使用多声道语音和音频的高质量服务的需求正在增加。首先, 高质量语音和音频服务的提供需要能够有效地对立体声语音和音频信号进行压缩和解压缩的编码技术。

[0003] 因此, 对于用于编码窄带 (NB : 300-3, 400 赫兹 (Hz))、宽带 (WB : 50-7, 000Hz)、和超宽带 (SWB : 50-14, 000Hz) 信号的编解码器的扩展研究正在进行中。例如, ITU-T G. 729. 1 是代表性扩展编解码器, 其是基于 G. 729 的 WB 扩展编解码器 (NB 编解码器)。此编解码器按照 8 千比特 / 秒 (kbit/s) 来提供与 G. 729 的比特流级别的兼容性, 并且按照 12 千比特 / 秒来提供较好质量的 NB 信号。在 14-32 千比特 / 秒的范围中, 编解码器可以对具有 2 千比特 / 秒的比特率可伸缩性的 WB 信号进行编码, 并且输出信号的质量随着比特率增加而改善。

[0004] 近来, 正在开发能够基于 G. 729. 1 提供 SWB 信号的扩展编解码器。此扩展编解码器可以对 NB、WB 和 SWB 信号进行编码和解码。

[0005] 在这种扩展编解码器中, 正弦脉冲编码可以用于改善所合成信号的质量。当使用正弦脉冲编码时, 需要考虑输入信号的能量, 以增加编码效率。具体地, 当可用于正弦脉冲编码的比特的数目不足时, 有效的是, 优先编码对于所合成信号的质量具有较大影响的频带 (即, 具有相对大量能量的频带)。

发明内容

[0006] 技术问题

[0007] 本发明的实施例针对一种用于对音频信号进行编码和解码的方法和设备, 其可以使用正弦脉冲编码来改善合成信号的质量。

[0008] 本发明的另一实施例针对一种用于对音频信号进行编码和解码的方法和设备, 其可以通过基于对于合成信号的每个子带的能量数量的考虑、而应用正弦脉冲编码, 来更加有效地改善合成信号的质量。

[0009] 本发明的目的不限于上述的目的, 并且本发明的其他目的和优点可以通过以下描述来理解, 并可以参考本发明的实施例而变得明显。同样, 对于本发明所属技术领域的技术人员显而易见的是, 可以通过所要求保护的手段及其组合来实现本发明的目的和优点。

[0010] 技术解决方案

[0011] 根据本发明的实施例, 一种用于对音频信号进行编码的方法包括 : 将所转换的音

频信号划分为多个子带；计算所述子带中的每个子带的能量；从所述子带中选择具有相对大量能量的预定数目的子带；以及对于所选择的子带执行正弦脉冲编码。

[0012] 根据本发明的另一实施例，一种用于对音频信号进行编码的设备包括：输入单元，被配置为接收所转换的音频信号；计算单元，被配置为将所合成的音频信号划分为多个子带，计算所述子带中的每个子带的能量，并且从所述子带中选择具有相对大量能量的预定数目的子带；以及编码单元，被配置为对于所选择的子带执行正弦脉冲编码。

[0013] 根据本发明的另一实施例，一种用于对音频信号进行解码的方法包括：接收所转换的音频信号；将所编码的音频信号划分为多个子带；计算所述子带中的每个子带的能量；从所述子带中选择具有相对大量能量的预定数目的子带；以及对于所选择的子带执行正弦脉冲解码。

[0014] 根据本发明的另一实施例，一种用于对音频信号进行解码的设备包括：输入单元，被配置为接收所转换的音频信号；计算单元，被配置为将所编码的音频信号划分为多个子带，计算所述子带中的每个子带的能量，并且从所述子带中选择具有相对大量能量的预定数目的子带；以及解码单元，被配置为对于所选择的子带执行正弦脉冲解码。

[0015] 根据本发明的另一实施例，一种用于对音频信号进行编码的方法包括：接收音频信号；对于该音频信号执行改进的离散余弦变换 (MDCT)，以输出 MDCT 系数；使用该 MDCT 系数来对高频音频信号进行合成；以及对于该高频音频信号执行正弦脉冲编码。

[0016] 根据本发明的另一实施例，一种用于对音频信号进行编码的设备包括：输入单元，被配置为接收音频信号；MDCT 单元，被配置为对于该音频信号执行 MDCT，以输出 MDCT 系数；合成单元，被配置为使用该 MDCT 系数来对高频音频信号进行合成；以及正弦脉冲编码单元，被配置为对于该高频音频信号执行正弦脉冲编码。

[0017] 根据本发明的另一实施例，一种用于对音频信号进行解码的方法包括：接收音频信号；对于该音频信号执行 MDCT，以输出 MDCT 系数；使用该 MDCT 系数来对高频音频信号进行合成；以及对于该高频音频信号执行正弦脉冲解码。

[0018] 根据本发明的另一实施例，一种用于对音频信号进行解码的设备包括：输入单元，被配置为接收音频信号；MDCT 单元，被配置为对于该音频信号执行 MDCT，以输出 MDCT 系数；合成单元，被配置为使用该 MDCT 系数来对高频音频信号进行合成；以及正弦脉冲解码单元，被配置为对于该高频音频信号执行正弦脉冲解码。

[0019] 有益效果

[0020] 根据本发明的示范实施例，使用正弦脉冲编码来改善合成信号的质量。

[0021] 另外，基于对于合成信号的每个子带的能量数量的考虑的正弦脉冲编码的应用更加有效地改善了合成信号的质量。

附图说明

[0022] 图 1 示出了用于提供与 NB 编解码器的兼容性的 SWB 扩展编解码器的结构。

[0023] 图 2 示出了根据本发明实施例的音频信号编码设备的构造。

[0024] 图 3 示出了根据本发明实施例的音频信号解码设备的构造。

[0025] 图 4 是示出了根据本发明实施例的音频信号编码方法的流程图。

[0026] 图 5 是示出了根据本发明实施例的执行正弦脉冲编码的步骤 (图 4 中的 S410) 的

流程图。

[0027] 图 6 是示出了根据本发明实施例的音频信号解码方法的流程图。

[0028] 图 7 示出了传统的正弦脉冲编码与根据本发明的自适应正弦脉冲编码的结果之间的比较。

[0029] 图 8 示出了根据本发明另一实施例的音频信号编码设备的构造。

[0030] 图 9 示出了根据本发明另一实施例的音频信号解码设备的构造。

具体实施方式

[0031] 下面,将参考附图来更加详细地描述本发明的示范实施例。然而,本发明可以按照不同的形式来实施,并且不应被诠释为限于在这里提出的实施例。相反地,提供这些实施例,使得本公开将是彻底和完全的,并将向本领域技术人员充分地传达本发明的范围。贯穿整个公开中,贯穿本发明的各个图和实施例中,同样的附图标记表示同样的部分。

[0032] 图 1 示出了用于提供与 NB 编解码器的兼容性的 SWB 扩展编解码器的结构。

[0033] 一般地,扩展编解码器具有以下构造,其中将输入信号划分为多个频带,并且对相应频带中的信号进行编码或解码。参考图 1,将输入信号输入到初级低通滤波器 102 和初级高通滤波器 104。将初级低通滤波器 102 配置为执行滤波和下采样,使得输出输入信号的低频带信号 A(0-8 千赫兹 (kHz))。将初级高通滤波器 104 配置为执行滤波和下采样,使得输出输入信号的高频带信号 B(8-16kHz)。

[0034] 将从初级低通滤波器 102 输出的低频带信号 A 输入到次级低通滤波器 106 和次级高通滤波器 108。将次级低通滤波器 106 配置为执行滤波和下采样,使得输出低低频带信号 A1(0-4kHz)。将次级高通滤波器 108 配置为执行滤波和下采样,使得输出低高频带信号 A2(4-8kHz)。

[0035] 结果,将低低频带信号 A1 输入到 NB 编码模块 110,将低高频带信号 A2 输入到 WB 扩展编码模块 112,并且将高频带信号 B 输入到 SWB 扩展编码模块 114。当 NB 编码模块 110 单独地进行操作时,仅仅重新生成 NB 信号,并且当 NB 编码模块 110 和 WB 扩展编码模块 112 两者进行操作时,重新生成 WB 信号。当 NB 编码模块 110、WB 扩展编码模块 112 和 SWB 扩展编码模块 114 全部都进行操作时,重新生成 SWB 信号。

[0036] 图 1 所示的扩展编解码器的代表性示例可以是 ITU-T G. 729. 1,其是基于 G. 729 的 WB 扩展编解码器 (NB 编解码器)。此编解码器按照 8kbit/s 来提供与 G. 729 的比特流级别的兼容性,并且按照 12kbit/s 来提供质量被大大改善的 NB 信号。在 14-32kbit/s 的范围中,编解码器可以对具有 2kbit/s 的比特率可伸缩性的 WB 信号进行编码,并且输出信号的质量由于比特率增加而改善。

[0037] 近来,正在开发能够基于 G. 729. 1 来提供 SWB 质量的扩展编解码器。此扩展编解码器可以对 NB、WB 和 SWB 信号进行编码和解码。

[0038] 在这种扩展编解码器中,可以将不同的编码方案应用于如图 1 所示的相应频带。例如,G. 729. 1 和 G. 711. 1 编解码器采用其中使用传统的 NB 编解码器 (即,G. 729 和 G. 711) 来对 NB 信号进行编码的编码方案,并且对于剩余信号执行改进的离散余弦变换 (MDCT),使得对所输出的 MDCT 系数进行编码。

[0039] 在 MDCT 域编码的情况下,将 MDCT 系数划分为多个子带,对每个子带的增益和形状

进行编码,并且使用代数码激励线性预测(ACELP)或正弦脉冲来对MDCT系数进行编码。扩展编解码器一般具有以下结构,其中首先对用于带宽扩展的信息进行编码,并然后对用于质量改善的信息进行编码。例如,使用每个子带的增益和形状来对7-14kHz频带中的信号进行合成,并且使用ACELP或正弦脉冲编码来改善合成信号的质量。

[0040] 具体地,在用于提供SWB质量的第一层中,使用诸如增益和形状之类的信息来对与7-14kHz波段对应的信号进行合成。然后,使用附加比特来应用例如正弦脉冲编码,以改善合成信号的质量。由于比特率增加,所以此结构可以改善合成信号的质量。

[0041] 一般地,在正弦脉冲编码的情况下,对与在给定间隔中具有最大幅度的脉冲(即,对于质量具有最大影响的脉冲)的位置、幅度和符号有关的信息进行编码。计算的数量与这种脉冲搜索间隔成比例地增加。因此,取代了将正弦脉冲编码应用于整个帧(在时域的情况下)或整个频带,优选地将正弦脉冲编码应用于每个子帧或子带。正弦脉冲编码的有利之处在于,尽管需要相对大数目的比特来传送一个脉冲,但是可以准确地表达影响信号质量的信号。

[0042] 输入到编解码器的信号的能量分布取决于频率而变化。具体地,在音乐信号的情况下,频率方面的能量变化比在语音信号的情况下更加严重。具有大量能量的子带中的信号对于合成信号的质量具有更大的影响。如果存在足够的比特来对整个子带进行编码,则将不会存在问题,但是如果并非如此,则有效的是,优先编码对于合成信号的质量具有大影响的子带中的(即,具有大量能量的)信号。

[0043] 本发明针对音频信号的编码和解码,其可以在如图1所示的扩展编解码器的情况下,通过基于对于有限比特数目的考虑执行更有效的正弦脉冲编码,来改善合成信号的质量。在下文中,在本发明的以下描述中将简单地把语音和音频信号称作音频信号。

[0044] 图2示出了根据本发明实施例的音频信号编码设备的构造。

[0045] 参考图2,音频信号编码设备202包括输入单元204、计算单元206和编码单元208。将输入单元204配置为接收所转换的音频信号,例如作为通过MDCT进行的音频信号的转换的结果的MDCT系数。

[0046] 将计算单元206配置为将已经通过输入单元204所输入的、转换后的音频信号划分为多个子带,并且计算每个子带的能量。将计算单元206配置为从子带中选择具有相对大量能量的、预定数目的子带。该预定数目通过要在一个子带中进行编码的正弦脉冲的数目和用于对一个正弦脉冲进行编码所必需的比特数目来确定。

[0047] 将编码单元208配置为对于计算单元206所选择的子带执行正弦脉冲编码。编码单元208可以按照能量数量的顺序,对于具有相对大量能量的、预定数目的子带来执行正弦脉冲编码。根据本发明的另一实施例,编码单元208可以按照除了能量数量的顺序之外的顺序(例如,按照带宽或索引的顺序),对于具有相对大量能量的、预定数目的子带来执行正弦脉冲编码。

[0048] 计算单元206可以确认在所选择的子带之中是否存在相邻的子带,并且将相邻的子带合并为一个子带。然后,编码单元208可以然后对于这样所合并的子带来执行正弦脉冲编码。

[0049] 图3示出了根据本发明实施例的音频信号解码设备的构造。

[0050] 参考图3,音频信号解码设备302包括输入单元304、计算单元306和解码单元

308。将输入单元 204 配置为接收所转换的音频信号,例如,MDCT 系数。

[0051] 将计算单元 306 配置为将已经通过输入单元 304 所输入的、转换后的音频信号划分为多个子带,并且计算每个子带的能量。将计算单元 306 配置为从子带中选择具有相对大量能量的、预定数目的子带。该预定数目通过要在一个子带中进行编码的正弦脉冲的数目和用于对一个正弦脉冲进行编码所必需的比特数目来确定。

[0052] 将解码单元 308 配置为对于计算单元 306 所选择的子带来执行正弦脉冲解码。解码单元 308 可以按照能量数量的顺序,对于具有相对大量能量的、预定数目的子带来执行正弦脉冲编码。根据本发明的另一实施例,解码单元 308 可以按照除了能量数量的顺序之外的顺序(例如,按照带宽或索引的顺序),对于具有相对大量能量的、预定数目的子带来执行正弦脉冲编码。

[0053] 图 2 和 3 所示的音频信号编码设备 202 和音频信号解码设备 302 可以包括在图 1 所示的 NB 编码模块 110、WB 扩展编码模块 112 或 SWB 扩展编码模块 114 中。

[0054] 在下文中,将结合图 1 所示的 SWB 扩展编码模块 114 进行的音频信号的示范编码或解码,参考图 4 到 6 来描述根据本发明实施例的用于对音频信号进行编码和解码的方法。

[0055] SWB 扩展编码单元 114 将与 7-14kHz 对应的 MDCT 系数划分为多个子带,并且对每个子带的增益和形状进行编码或解码,以获得误差信号。然后,SWB 扩展编码模块 114 对于误差信号来执行正弦脉冲编码或解码。如果存在充足数目的比特要用于正弦脉冲编码,则正弦脉冲编码可以应用于每个子带。然而,由于比特数目在大多数的情况下都很难是充足的,所以仅仅对于有限数目的子带来应用正弦脉冲编码。因此,将正弦脉冲编码应用于对于合成信号的质量具有较大影响的子带保证了在给定相同比特率的情况下获得较好的信号质量。

[0056] 图 4 是示出了根据本发明实施例的音频信号编码方法的流程图。

[0057] 参考图 4,在步骤 S402 中,在 SWB 扩展编码模块 114 中包括的音频信号编码设备接收所转换的音频信号,例如,与 7-14kHz 对应的 MDCT 系数。该设备在步骤 S404 中将所接收到的转换后的音频信号划分为多个子带,并且在步骤 S406 中对所述多个子带中的每一个子带的能量进行计算。图 7 示出了被划分为九个子带的 MDCT 系数、以及每个子带的能量的相对数量。根据图 7 所清楚的是,子带 1、4、5、6 和 7 的能量的数量大于其他子带的能量数量。

[0058] 下面,表 1 列举了已经被划分为八个子带的 MDCT 系数的索引和能量。

[0059] 表 1

[0060]

索引	1	2	3	4	5	6	7	8
能量	350	278	657	245	1500	780	200	190

[0061] 在步骤 S408 中,音频信号编码设备从所述子带中选择具有大量能量的、预定数目的子带。例如,如下面的表 2 所示,按照能量的顺序来对表 1 的 MDCT 系数进行归类,并且从它们中选择具有最大能量数量的五个子带(其用阴影表示)。

[0062] 表 2

[0063]

索引	5	6	3	1	2	4	7	8
能量	1500	780	657	350	278	245	200	190

[0064] 根据本发明,如表 2 所示地选择预定数目的(例如,五个)子带。该预定数目通过要在一个子带中进行编码的正弦脉冲的数目和用于对一个正弦脉冲进行编码所必需的比特的数目来确定。

[0065] 如下地确定用于对一个正弦脉冲进行编码所必需的比特的数目:需要一个比特来对一个正弦脉冲的符号(+,-)进行编码。通过正弦脉冲搜索间隔的尺寸(例如,一个子带的尺寸)来确定用于对正弦脉冲的位置进行编码所需的比特的数目。如果子带的尺寸是 32,则需要五个比特来对正弦脉冲的位置进行编码($2^5 = 32$)。通过量化器的结构和和代码本的尺寸来确定用于对正弦脉冲的幅度(增益)进行编码所需的比特的数目。综上,用于对一个正弦脉冲进行编码所必需的比特的数目是用于对正弦脉冲的符号、位置和幅度进行编码所需的比特的总数。

[0066] 将假设:在已经确认了针对正弦脉冲编码所给定的比特的数目和用于对一个正弦脉冲进行编码所必需的比特的数目的情况下,可以传送十个正弦脉冲。当针对每个子带来对两个正弦脉冲进行编码时,可以将正弦脉冲编码应用于总共五个子带。因此,音频信号编码设备如表 2 所示地选择具有最大数量能量的五个子带,并且在步骤 S410 中对于所选择的子带 5、6、3、1 和 2 来执行正弦脉冲编码。

[0067] 图 5 是示出了根据本发明实施例的执行正弦脉冲编码的步骤(图 4 中的 S410)的流程图。

[0068] 根据本发明的另一实施例,在步骤 S502 中确认在图 4 的步骤 S408 中所选择的子带之中是否存在相邻的子带。在步骤 S504 中将相邻的子带合并为一个子带,并且在步骤 S506 中对于所合并的子带执行正弦脉冲编码。

[0069] 例如,假设已经如表 2 所示地选择了五个子带 5、6、3、1 和 2,则确认子带 5 是否具有在所选择的子带之中的相邻的子带,即子带 4 或 6。确认出这五个子带中包括与子带 5 相邻的子带 6。因此,取代了对用于子带 5 和 6 中的每一个的两个正弦脉冲进行编码,音频信号编码设备将两个子带合并为单一子带,并且对于该单一子带来对四个正弦脉冲进行编码。例如,如果子带 5 比子带 6 具有更大数量的能量,则所有的四个正弦脉冲都可位于所合并的子带中的子带 5 中。照这样,对相邻的子带进行合并并且将正弦脉冲编码应用于所合并的子带保证了更加有效的正弦脉冲编码。

[0070] 其间,取决于编解码器的特性,编码器和解码器所合成的 7-14kHz 波段中的信号可能彼此不一致。为了减少由于编码器和解码器所分别计算的子带的能量的差异所产生的误差,音频信号编码设备可以重新安排子带,如下面表 3 所示,并且执行正弦脉冲编码。

[0071] 表 3

[0072]

索引	1	2	3	5	6	4	7	8
----	---	---	---	---	---	---	---	---

能量	350	278	657	1500	780	245	200	190
----	-----	-----	-----	------	-----	-----	-----	-----

[0073] 即,取代了按照能量数量的顺序而对于五个子带执行正弦脉冲编码,音频信号编码设备可以按照带宽或索引的顺序来执行正弦脉冲编码。照这样,不考虑所选择的子带的能量数量的顺序减少了由于可能在编码器和解码器中出现的较高频带合成信号的差异所产生的误差。

[0074] 图 6 是示出了根据本发明实施例的音频信号解码方法的流程图。

[0075] 首先,在步骤 S602 中接收所转换的音频信号。在步骤 S604 中,将所转换的音频信号划分为多个子带,并且在步骤 S606 中计算每个子带的能量。

[0076] 在步骤 S608 中,从子带中选择具有大量能量的、预定数目的子带,并且在步骤 S610 中对于所选择的子带执行正弦脉冲解码。图 6 的步骤 S602 到 S610 与上述的根据本发明实施例的音频信号编码方法的相应步骤相似,并且在这里将省略其详细描述。

[0077] 图 7 示出了传统的正弦脉冲编码与根据本发明实施例的自适应正弦脉冲编码的结果之间的比较。

[0078] 在图 7 中,(a) 对应于传统正弦脉冲编码的结果。根据图 7 所示的每个子带的能量的相对数量的比较所清楚的是,子带 1、4、5、6 和 7 的能量的数量比其他子带的能量数量更大。然而,传统正弦脉冲编码按照带宽或索引的顺序来应用脉冲编码,而与子带的能量数量无关,使得相对于子带 1、2、3、4 和 5 来对正弦脉冲进行编码,如图 (a) 所示。

[0079] 在图 7 中,(b) 对应于根据本发明的自适应正弦脉冲编码的结果。根据 (b) 所清楚的是,根据本发明,将正弦脉冲编码应用于具有相对大量能量的子带,即,子带 1、4、5、6 和 7。

[0080] 如上所述,本发明可应用于包括语音的音频信号。语音信号的能量分布如下:浊音(voiced)声音具有大多数都位于低频带中的能量,而清音(unvoiced)和爆破音(plosive)声音具有位于相对高频带中的能量。相反,音乐信号的能量取决于频率而极大地变化。这意味着,与语音信号不同地,难以在频带的方面对音乐信号的能量分布的特性进行定义。合成信号的质量受到具有大量能量的频带中的信号的更大影响。因此,取代了对要应用正弦脉冲编码的子带进行固定,如本发明所提出的,根据输入信号的特性来选择子带并且相应地应用正弦脉冲编码可以改善按照相同的比特率所合成的信号的质量。

[0081] 现在,将参考图 8 和 9 来描述根据本发明另一实施例的用于对音频信号进行编码和解码的方法和设备。

[0082] 图 8 示出了根据本发明另一实施例的音频信号编码设备的构造。

[0083] 将图 8 所示的音频信号编码设备配置为接收 32kHz 的输入信号,并且合成和输出 WB 和 SWB 信号。音频信号编码设备包括 WB 扩展编码模块 802、808 和 822 以及 SWB 扩展编码模块 804、806、810 和 812。WB 扩展模块(具体地,G.729.1 核心编解码器)使用 16kHz 信号来进行操作,而 SWB 扩展编码模块使用 32kHz 信号。在 MDCT 域中执行 SWB 扩展编码。使用两个模式(即,一般(generic)模式 814 和正弦模式 816)来对 SWB 扩展编码模块的第一层进行编码。基于输入信号的所测量的音调(tonality),来做出关于要使用一般模式和正弦模式 814 和 816 中的哪一个的确定。通过用于改善高频内容的质量的正弦编码单元 818 和 820,或者通过用于改善 WB 内容的感知质量的 WB 信号改善单元 822,来对较高 SWB 波段

进行编码。

[0084] 首先,将 32kHz 的输入信号输入到下采样单元 802 中,并且将它下采样到 16kHz。将下采样的 16kHz 信号输入到 G. 729. 1 编解码器 808。G. 729. 1 编解码器 808 对于所输入的 16kHz 信号执行 WB 编码。将从 G. 729. 1 编解码器 808 所输出的合成后的 32kbit/s 信号输入到 WB 信号改善单元 822,并且 WB 信号改善单元 822 改善输入信号的质量。

[0085] 另一方面,将 32kHz 输入信号输入到 MDCT 单元 806,并且将它转换到 MDCT 域中。在步骤 S810 中,将转换到 MDCT 域中的输入信号输入到音调测量单元 804,以确定输入信号是否是音调的。换言之,基于音调测量来定义第一 SWB 层中的编码模式,该音调测量是通过在 MDCT 域中对输入信号的当前帧和先前帧的对数域能量进行比较来执行的。音调测量基于输入信号的当前帧和先前帧的谱峰值之间的相关性分析。

[0086] 在步骤 S810 中,基于音调测量单元 804 所输出的音调信息,确定输入信号是否是音调的。例如,如果音调信息高于给定的阈值,则确认该输入信号是音调的,并且如果并非如此,则确认该输入信号不是音调的。还将该音调信息包括在传输到解码器的比特流中。如果输入信号是音调的,则使用正弦模式 816,并且如果并非如此,则使用一般模式 814。

[0087] 当输入信号的帧不是音调的(音调(tonal) = 0)时,使用一般模式 814。一般模式 814 利用 G. 729. 1WB 编解码器 808 的编码 MDCT 域表达式,来对高频进行编码。将高频频带(7-14kHz)划分为四个子带,并且从编码后的、包络的归一化 WB 内容中搜索与每个子带有关的所选择的相似性准则。通过两个缩放(scaling)因子(具体地,线性域的第一缩放因子和对数域的第二缩放因子)来对最相似的匹配进行缩放,以获取所合成的高频内容。还通过正弦编码单元 818 和一般模式 814 内的附加正弦脉冲来改善此内容。

[0088] 在一般模式 814 中,可以通过根据本发明的音频编码方法来改善所编码信号的质量。例如,比特预算(budget)允许将两个正弦脉冲添加到 4kbit/s 的第一 SWB 层。基于所合成的高频信号的子带能量来选择轨道的开始位置,使用该开始位置来搜索要添加的正弦脉冲的位置。可以根据下面的等式 1 来计算合成子带的能量。

$$[0089] \quad SbE(k) = \sum_{n=0}^{n=31} \ddot{M}_{32}(k \times 32 + n)^2 \quad k = 0, \dots, 7 \quad (\text{等式 1})$$

[0090] 其中,k 是指子带索引,SbE(k) 是指第 k 子带的能量,而 $\ddot{M}_{32}(k)$ 是指所合成的高频信号。每个子带由 32 个 MDCT 系数组成。作为用于正弦编码的搜索轨道来选择具有相对大量能量的子带。例如,搜索轨道可以包括具有单位尺寸 1 的 32 个位置。在此情况下,搜索轨道与子带相符。

[0091] 分别通过 4 比特、一维代码本来对两个正弦脉冲的幅度进行量化。

[0092] 当输入信号是音调的时,使用正弦模式 816。在正弦模式 816 中,通过将一组有限数目的正弦分量添加到高频频谱,来创建高频信号。例如,假设添加总共十个正弦脉冲,则四个正弦脉冲可以位于 7000-8600Hz 的频率范围中,四个正弦脉冲可以位于 8600-10200Hz 的频率范围中,一个正弦脉冲可以位于 10200-11800Hz 的频率范围中,并且一个正弦脉冲可以位于 11800-12600Hz 的频率范围中。将正弦编码单元 818 和 820 配置为改善一般模式 814 或正弦模式 816 所输出的信号的质量。正弦编码单元 818 和 820 所添加的正弦脉冲的数目(Nsin)取决于比特预算而变化。基于高频内容的子带能量来选择用于正弦编码单元 818 和 820 进行正弦编码的轨道。

[0093] 例如,将 7000–13400Hz 的频率范围中的所合成的高频内容划分为八个子带。每个子带由 32 个 MDCT 系数组成,并且可以根据等式 1 来计算每个子带的能量。

[0094] 通过找到与 $N_{\sin}/N_{\sin_track}$ 一样多的具有相对大量能量的子带,来选择用于正弦编码的轨道。对此, $N_{\sin_track}$ 是指每个轨道的正弦脉冲的数目,并且被设置为 2。所选择的 $(N_{\sin}/N_{\sin_track})$ 个子带分别对应于用于正弦编码所使用的轨道。例如,假设 $N_{\sin}$ 是 4,则最初两个正弦脉冲位于具有最大数量子带能量的子带中,而剩余的两个正弦脉冲位于具有次大数量能量的子带中。用于正弦编码的轨道位置取决于可用的比特预算和高频信号能量特性而逐帧地变化。

[0095] 图 9 示出了根据本发明另一实施例的音频信号解码设备的构造。

[0096] 将图 9 所示的音频信号解码设备配置为接收该编码设备已经编码的 WB 和 SWB 信号,并且输出对应的 32kHz 信号。音频信号解码设备包括 WB 扩展解码模块 902、914、916 和 918 以及 SWB 扩展解码模块 902、920 和 922。将 WB 扩展解码模块配置为对所输入的 16kHz 信号进行解码,并且将 SWB 扩展解码模块配置为对高频进行解码,以提供 32kHz 输出。使用两个模式(具体地,一般模式 906 和正弦模式 908)来对扩展的第一层进行解码,并且此操作取决于首先解码的音调指示符。第二层使用与编码器相同的比特分配,以改善 WB 信号,并且在附加的正弦脉冲之间分布比特。第三 SWB 层由正弦解码单元 910 和 912 组成,并且这改善了高频内容的质量。第四和第五扩展层提供 WB 信号改善。为了改善所合成的 SWB 内容,在时域中使用后处理。

[0097] 将编码设备所编码的信号输入到 G. 729. 1 编解码器 902。G. 729. 1 编解码器 902 输出被输入到 WB 信号改善单元 914 的、16kHz 的合成信号。WB 信号改善单元 914 对所输入的信号的质量进行改善。从 WB 信号改善单元 914 输出的信号经历后处理单元 916 进行的后处理和上采样单元 918 所进行的上采样。

[0098] 其间,在开始高频解码以前,需要对 WB 信号进行合成。通过 G. 729. 1 编解码器 902 来执行这种合成。在高频信号解码的情况下,在应用一般的后处理功能以前使用 32kbit/s WB 合成。

[0099] 高频信号的解码通过从 G. 729. 1WB 解码中获取所合成的 MDCT 域表达式而开始。需要 MDCT 域 WB 内容来对一般编码帧的高频信号进行解码,并且通过从 WB 频率范围自适应地复制所编码的子带,来构造此情况下的高频信号。

[0100] 一般模式 906 通过自适应子带复制来构造高频信号。此外,将两个正弦分量添加到最初 4kbit/s SWB 扩展层的频谱。一般模式 906 和正弦模式 908 利用基于正弦模式解码技术的相似增强层。

[0101] 在一般模式 906 中,可以通过根据本发明的音频解码方法来改善所解码的信号的质量。一般模式 906 将两个正弦分量添加到所重构的整个高频频谱。在位置、符号和幅度方面表达这些正弦脉冲。如上所述,从具有相对大量能量的子带的索引中获取用于添加正弦脉冲所使用的轨道的开始位置。

[0102] 在正弦模式 908 中,通过一组有限数目的正弦分量来创建高频信号。例如,假设添加总共十个正弦脉冲,则四个正弦脉冲可以位于 7000–8600Hz 的频率范围中,四个正弦脉冲可以位于 8600–10200Hz 的频率范围中,一个正弦脉冲可以位于 10200–11800Hz 的频率范围中,并且一个正弦脉冲可以位于 11800–12600Hz 的频率范围中。

[0103] 将正弦解码单元 902 和 912 配置为改善通过一般模式 906 或正弦模式 908 所输出的信号的质量。第一 SWB 改善层将十个正弦分量添加到正弦模式帧的高频信号频谱。在一般模式帧中,根据低频与高频改善之间的自适应比特分配,来设置所添加的正弦分量的数目。

[0104] 正弦解码单元 910 和 912 进行的解码的处理如下:首先,从比特流中获取正弦脉冲的位置。然后,对该比特流进行解码,以获得所传送的符号索引和幅度代码本索引。

[0105] 通过找到与 $N_{\sin}/N_{\sin_track}$ 一样多的具有相对大量能量的子带,来选择用于正弦解码的轨道。对此, $N_{\sin_track}$ 是指每个轨道的正弦脉冲的数目,并且被设置为 2。所选择的 $(N_{\sin}/N_{\sin_track})$ 个子带分别对应于用于正弦解码的轨道。

[0106] 初始地从该比特流中获得与相应的对应轨道相关的十个正弦脉冲的位置索引。然后,对这十个正弦脉冲的符号进行解码。最终,对正弦脉冲的幅度(三个 8 比特代码本索引)进行解码。

[0107] 正弦解码单元 910 和 912 已经这样改善其质量的信号经历 IMDCT 920 进行的逆 MDCT 和后处理单元 922 进行的后处理。添加从上采样单元 918 和后处理单元 922 所输出的信号,使得输出 32kHz 输出信号。

[0108] 尽管已经相对于特定实施例而描述了本发明,但是对于本领域技术人员明显的是,可以做出各种改变和修改,而不脱离由以下权利要求所限定的本发明的精神和范围。

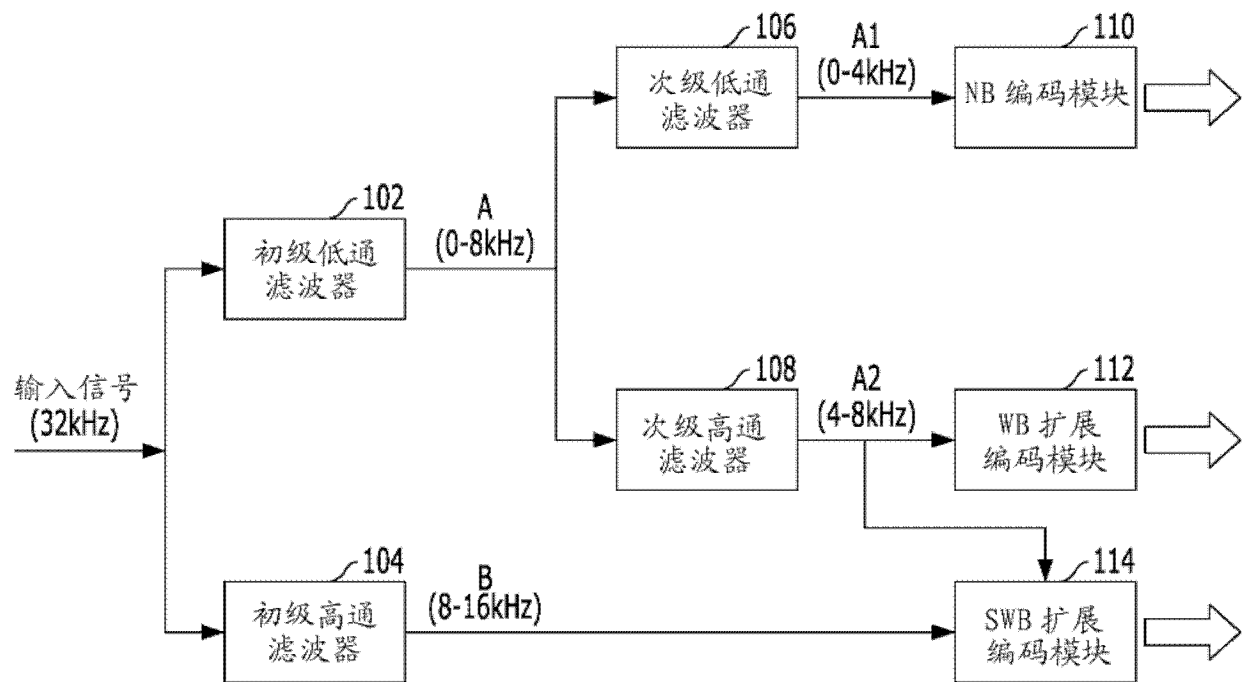


图 1

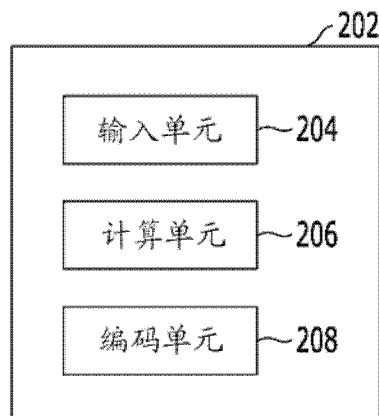


图 2

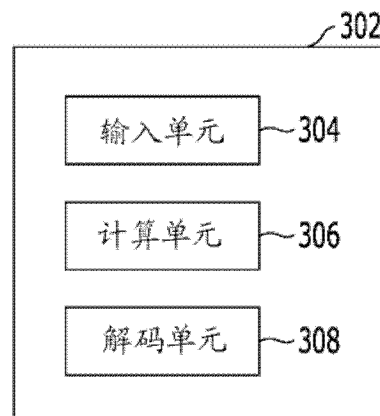


图 3

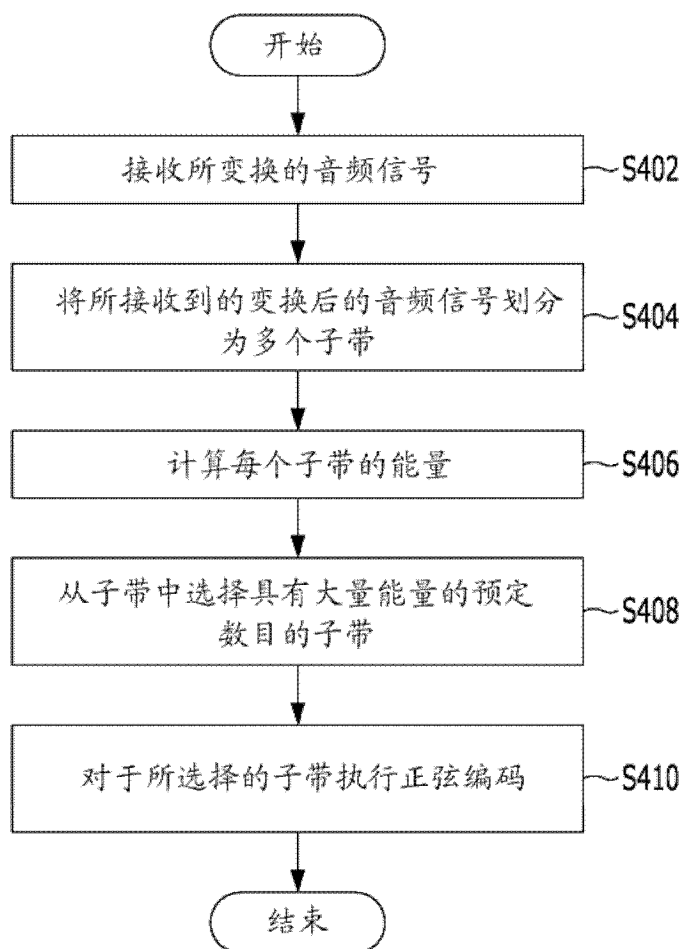


图 4

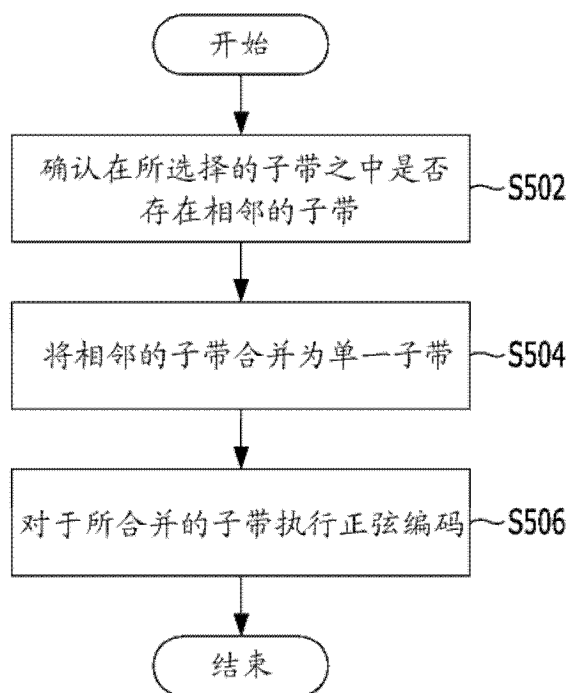


图 5

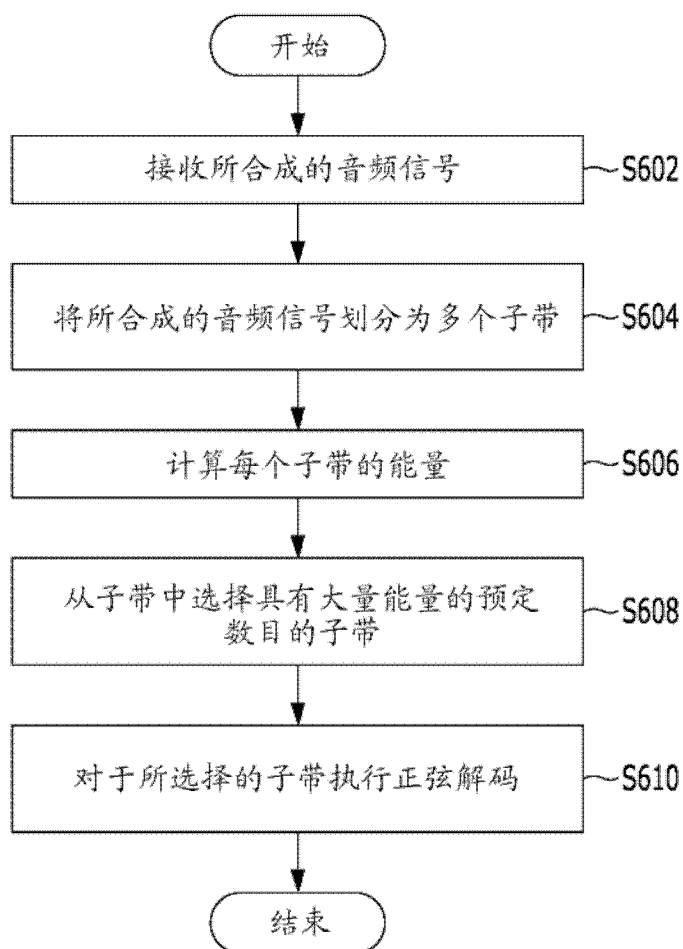


图 6

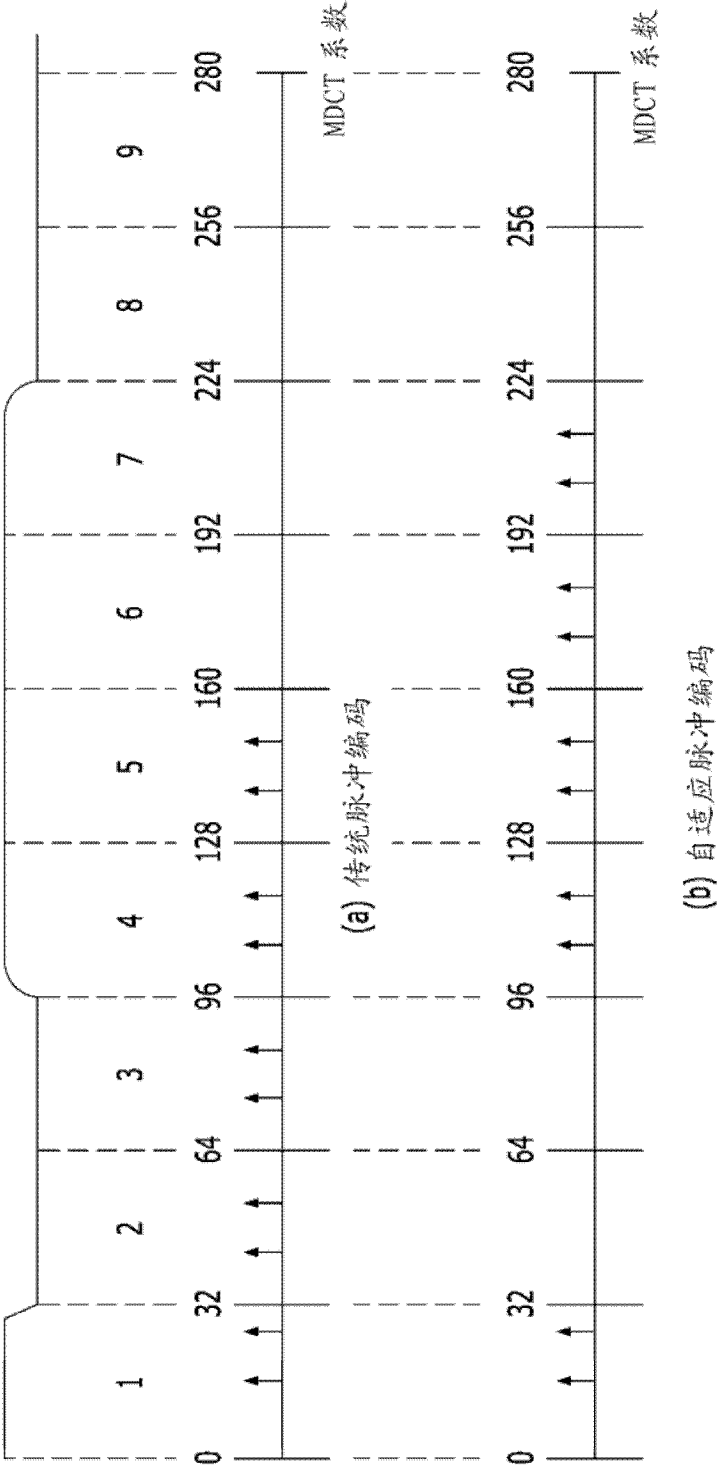


图 7

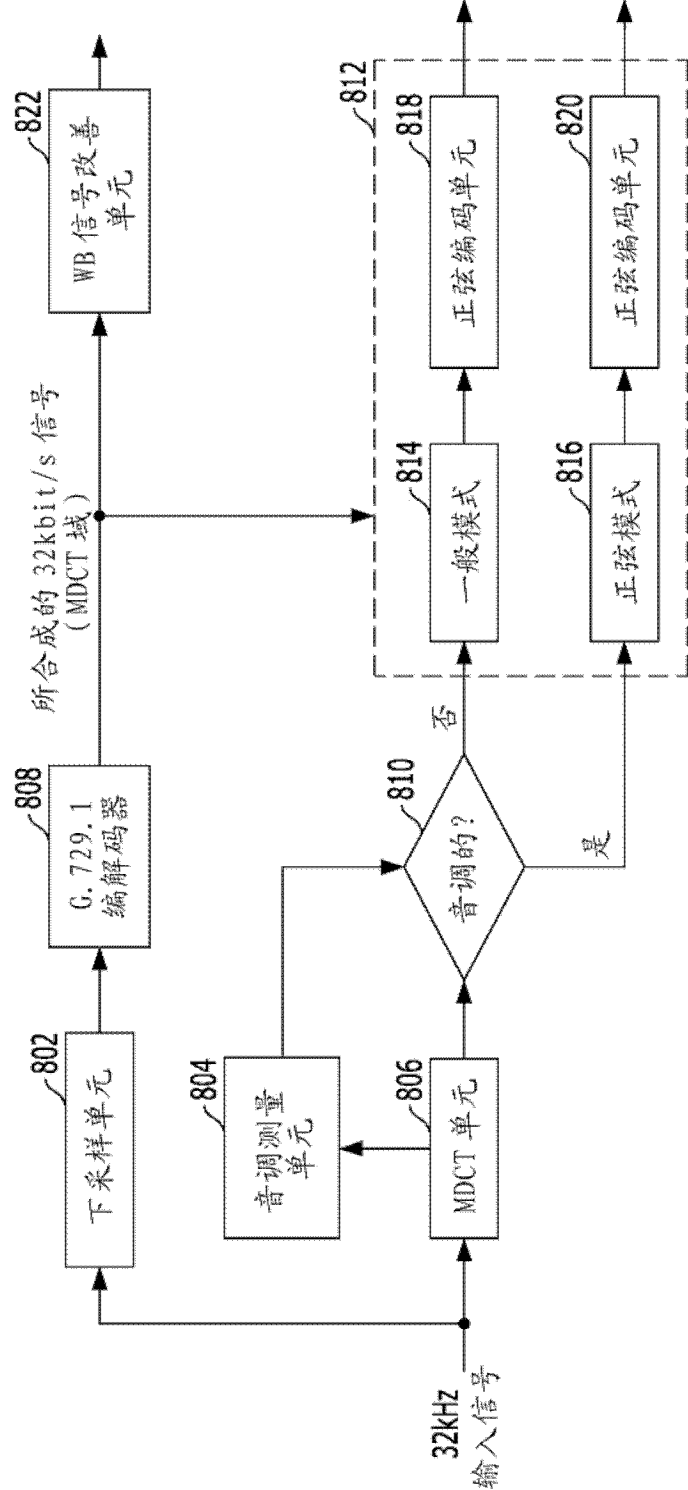


图 8

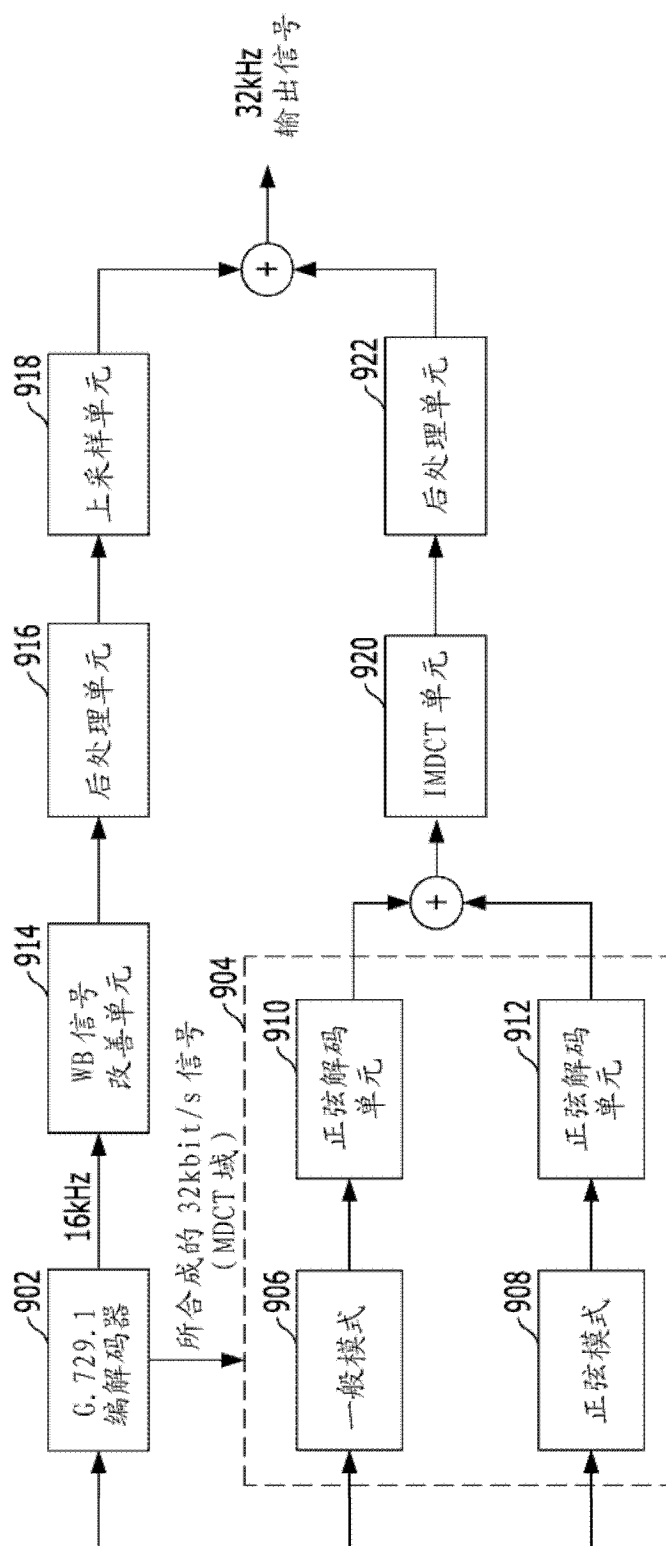


图 9