



(12) 发明专利申请

(10) 申请公布号 CN 101925950 A

(43) 申请公布日 2010.12.22

(21) 申请号 200880125581.4

(74) 专利代理机构 中国国际贸易促进委员会专利商标事务所 11038

(22) 申请日 2008.12.30

代理人 李镇江

(30) 优先权数据

08009531.8 2008.05.24 EP

0800032.5 2008.01.04 SE

61/055,975 2008.05.24 US

(51) Int. Cl.

G10L 19/00(2006.01)

(85) PCT申请进入国家阶段日

2010.07.26

(86) PCT申请的申请数据

PCT/EP2008/011145 2008.12.30

(87) PCT申请的公布数据

W02009/086919 EN 2009.07.16

(71) 申请人 杜比国际公司

地址 荷兰阿姆斯特丹

(72) 发明人 A·比斯沃斯 H·普恩哈根

K·克约尔林 B·雷施

L·F·维勒莫斯 P·H·海德林

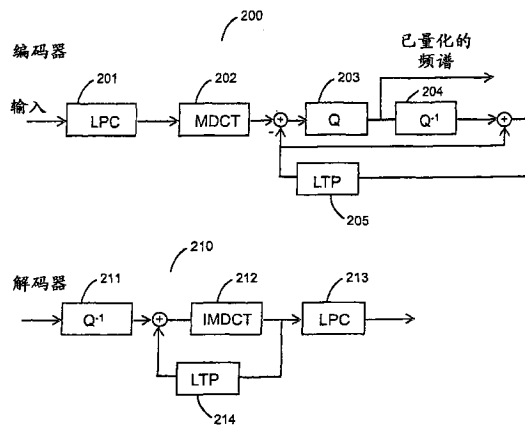
权利要求书 4 页 说明书 23 页 附图 22 页

(54) 发明名称

音频编码器和解码器

(57) 摘要

本发明涉及音频编码器和解码器。本发明教导一种新型音频编码系统,其可以在低比特率下良好地编码通用音频和语音信号。本方面提出的音频编码系统包括:线性预测单元,其用于基于自适应滤波器对输入信号进行滤波;变换单元,其用于将滤波后的输入信号帧变换成变换域;量化单元,其用于量化变换域信号;长期预测单元,其用于基于滤波输入信号的先前片段的重构,确定滤波输入信号帧的估计;及变换域信号组合单元,其用于在变换域中组合长期预测估计与变换输入信号,以生成所述变换域信号。



1. 一种音频编码系统,包括:
线性预测单元,其用于基于自适应性滤波器对输入信号进行滤波;
变换单元,其用于将滤波输入信号帧变换成变换域;
量化单元,其用于量化变换域信号;
长期预测单元,其用于基于所述滤波输入信号的先前片段的重构,确定所述滤波输入信号帧的估计;及
变换域信号组合单元,其用于在所述变换域中组合所述长期预测估计与所述变换输入信号,以生成所述变换域信号。
2. 如权利要求 1 所述的音频编码系统,包括:
反量化和逆变换单元,其用于生成所述滤波输入信号帧的时间域重构;及
长期预测缓冲器,其用于存储所述滤波输入信号的先前帧的时间域重构。
3. 如权利要求 1 或 2 所述的音频编码系统,其特征在于,
用于对所述输入信号进行滤波的所述自适应性滤波器是基于以第一帧长进行操作并产生白化输入信号的线性预测编码(LPC)分析,且
应用于所述滤波输入信号帧的变换是以可变第二帧长进行操作的修正离散余弦变换(MDCT)。
4. 如权利要求 3 所述的音频编码系统,包括:
窗体序列控制单元,其用于通过最小化输入信号块的编码成本函数,所述编码成本函数优选为简化感知熵,以便为所述输入信号块确定用于重合 MDCT 窗体的第二帧长。
5. 如权利要求 4 所述的音频编码系统,其特征在于,接连的 MDCT 窗体的长度至多以因子 2 改变。
6. 如权利要求 4 或 5 所述的音频编码系统,其特征在于,所述 MDCT 窗体长度是所述输入信号块的二元分割。
7. 如权利要求 4 至 6 中任一项所述的音频编码系统,其特征在于,所述窗体序列控制单元配置为:当搜索最小化所述输入信号块的所述编码成本函数的 MDCT 窗体长度序列时,考虑所述长期预测单元针对窗体长度候选所生成的长期预测估计。
8. 如权利要求 4 至 7 中任一项所述的音频编码系统,包括用于对序列中的 MDCT 窗体长度和窗体形状进行联合编码的窗体序列编码器。
9. 如权利要求 8 所述的音频编码系统,其特征在于,所述窗体序列编码器在对序列的窗体长度和形状进行编码时考虑窗体尺寸的约束。
10. 如任一先前权利要求所述的音频编码系统,包括 LPC 编码器,其用于以可变速率对所述线性预测单元所生成的线谱频率进行递归编码。
11. 如任一先前权利要求所述的音频编码系统,包括线性预测内插单元,用于内插以对应于所述第一帧长的速率生成的线性预测参数,以便匹配于以对应于所述第二帧长的速率生成的所述变换域信号的帧。
12. 如任一先前权利要求所述的音频编码系统,包括感知建模单元,其通过对所述线性预测单元为 LPC 帧而生成的 LPC 多项式进行啁啾和/或倾斜来修改所述自适应性滤波器的特性。
13. 如任一先前权利要求所述的音频编码系统,包括时间扭曲单元,其用于通过根据时

间扭曲曲线对所述滤波输入信号进行再采样来均匀地对准所述滤波信号帧中的音高组分，其中所述变换单元和所述长期预测单元对时间扭曲信号进行操作。

14. 如任一先前权利要求所述的音频编码系统，其特征在于，所述线性预测滤波在扭曲频率轴上进行操作。

15. 如任一先前权利要求所述的音频编码系统，包括用于编码所述输入信号的高频带组分的高频带编码器，其中当量化所述变换域信号时，所述量化单元中用于编码所述变换域中信号组分属于所述高频带的组分的量化步骤不同于用于编码属于所述输入信号低频带的组分。

16. 如权利要求 1 至 14 中任一项所述的音频编码系统，包括：

分频单元，其用于将所述输入信号分成低频带组分和高频带组分；及

高频带编码器，其用于编码所述高频带组分，其中所述低频带组分输入到所述线性预测单元。

17. 如权利要求 16 所述的音频编码系统，其特征在于，所述分频单元包括配置为对所述输入信号进行下采样的正交镜像滤波器组及正交镜像滤波器合成单元。

18. 如权利要求 16 或 17 所述的音频编码系统，其特征在于，所述低频带与所述高频带之间的边界是可变的，且所述分频单元基于输入信号性质和 / 或编码器带宽要求确定交越频率。

19. 如权利要求 16 至 18 中任一项所述的音频编码系统，包括：

第二正交镜像滤波器合成单元，其将所述高频带组分转换成低通信号；及

第二基于变换的编码器，其用于编码所述低通信号。

20. 如权利要求 16 至 19 中任一项所述的音频编码系统，包括信号表达式组合单元，其用于组合覆盖同一频率范围的不同信号表达式并且生成指示所述信号表达式的组合方式的信令数据。

21. 如权利要求 15 至 20 中任一项所述的音频编码系统，其特征在于，所述高频带编码器是频谱频带复制编码器。

22. 如任一先前权利要求所述的音频编码系统，其特征在于，所述长期预测单元包括用于将能量引入所述长期预测估计的高频组分的频谱频带复制单元。

23. 如任一先前权利要求所述的音频编码系统，包括用于计算左输入信道和右输入信道的参数立体声表达式的参数立体声单元。

24. 如任一先前权利要求所述的音频编码系统，包括用于由较低频率线预测较高频率 MDCT 线的谐波预测分析单元。

25. 如任一先前权利要求所述的音频编码系统，其特征在于，所述量化单元基于输入信号特性决定是用基于模型的量化器还是用非基于模型的量化器来编码所述变换域信号。

26. 如任一先前权利要求所述的音频编码系统，包括量化步长尺寸控制单元，其用于基于线性预测和长期预测参数来确定所述变换域信号的组分的量化步长尺寸。

27. 如任一先前权利要求所述的音频编码系统，其特征在于，所述长期预测单元包括：

长期预测提取器，用于确定滞后值，指定所述滤波信号中的最佳适合于所述滤波信号当前帧的重构片段；及

长期预测增益估计器，其用于估计应用于所述滤波信号的所选片段的信号的增益值，

其中确定所述滞后值和所述增益值以最小化失真标准。

28. 如权利要求 27 所述的音频编码系统,其特征在於,所述失真标准与对感知域中的所述变换输入信号的所述长期预测估计的差有关,所述失真标准是通过搜索所述感知域内的所述滞后值和所述增益值来最小化。

29. 如权利要求 27 或 28 所述的音频编码系统,其特征在於,当最小化所述失真标准时,将所述感知建模单元生成的所述修正线性预测多项式应用为 MDCT 域均衡增益曲线。

30. 如权利要求 27 至 29 中任一项所述的音频编码系统,其特征在於,所述长期预测单元包括将所述所选片段的重构信号变换成所述变换域,所述变换优选地是 IV 型离散余弦变换。

31. 如权利要求 27 至 30 中任一项所述的音频编码系统,其特征在於,所述长期预测单元包括虚拟矢量生成器,用于在所述滞后值小于所述 MDCT 帧长时生成所述重构信号的扩展片段。

32. 如权利要求 31 所述的音频编码系统,其特征在於,所述虚拟矢量生成器应用迭代的叠入-展开程序来细化所述重构信号的所述生成片段。

33. 如权利要求 27 至 32 中任一项所述的音频编码系统,其特征在於,当所述变换单元对时间扭曲信号进行操作时,所述长期预测单元基于从所述时间扭曲单元接收的时间扭曲曲线,对所述重构滤波输入信号进行再采样。

34. 如任一先前权利要求所述的音频编码系统,其特征在於,所述长期预测单元包括用于编码所述长期预测滞后和增益值的可变速率编码器。

35. 如任一先前权利要求所述的音频编码系统,其特征在於,所述长期预测单元包括噪声矢量缓冲器和 / 或脉冲矢量缓冲器。

36. 如任一先前权利要求所述的音频编码系统,包括联合编码单元,其用于对与音高有关的信息进行联合编码,例如,长期预测参数、谐波预测参数及时间扭曲参数。

37. 一种音频解码器,包括:

去量化单元,其用于对输入比特流的帧进行去量化;

逆变换单元,其用于对变换域信号进行逆变换;

长期预测单元,其用于确定所述去量化帧的估计;

变换域信号组合单元,其用于在所述变换域中组合所述长期预测估计和所述去量化帧,以生成所述变换域信号;及

线性预测单元,其用于对所述逆变换的变换域信号进行滤波。

38. 一种音频编码方法,包括如下步骤:

基于自适应性滤波器对输入信号进行滤波;

将所述滤波输入信号的帧变换成变换域;

对变换域信号进行量化;

基于所述滤波输入信号的先前片段的重构,估计所述滤波输入信号的帧;及

在所述变换域中,组合所述长期预测估计和所述变换输入信号,以生成所述变换域信号。

39. 一种音频解码方法,包括如下步骤:

对输入比特流的帧进行去量化;

对变换域信号进行逆变换；
确定所述去量化帧的估计；
在所述变换域中，组合所述长期预测估计和所述去量化帧，以生成所述变换域信号；
对所述逆变换的变换域信号进行滤波；及
输出重构的音频信号。

40. 一种用于致使可编程设备实施如权利要求 38 或 39 所述的音频编码方法的电脑程序。

音频编码器和解码器

技术领域

[0001] 本发明涉及音频信号的编码,具体而言,涉及对不仅限于语音、音乐或者其组合的任何音频信号的编码。

背景技术

[0002] 在现有技术中,有专门设计为通过将编码基于信号的源模型,即,人的发音系统,来对语音信号进行编码的语音编码器。这些编码器不能处理诸如音乐或任何其他非语音信号之类的任意音频信号。另外,在现有技术中还有通常被称为音频编码器的音乐编码器,它们将编码基于人的听觉系统的构想,而不是基于信号的源模型。这些编码器可以很好地处理任意信号,但是对于低比特率的语音信号,专用的语音编码器提供了优良的音频质量。因此,到目前为止,还不存在用于编码任意音频信号的通用的编码结构,并且当以低比特率操作时,其既能作为针对语音的语音编码器又能作为针对音乐的音乐编码器。

[0003] 因此,需要一种能改善音频质量和 / 或降低比特率的增强型音频编码器和解码器。

发明内容

[0004] 本发明涉及以等同或优于为具体信号而特别定制的系统的质量水平,有效地编码任意音频信号。

[0005] 本发明涉及音频编码 - 解码器算法,其包含线性预测编码 (LPC) 以及经 LPC 处理的信号进行操作的变换编码器部分两者。

[0006] 本发明还涉及通过可变的帧尺寸 (大小) 来有效地利用音频编码器中的比特储存器 (reservoir)。

[0007] 本发明还涉及结合具有可变帧尺寸 (size) 的变换编码器进行长期预测的操作。

[0008] 本发明还涉及一种用于编码音频信号并生成比特流的编码器;以及一种用于解码比特流并生成与输入音频信号在感知上无法区分的重构音频信号的解码器。

[0009] 本发明提供一种音频编码系统,其基于变换编码器并包括来自语音编码器的基础预测模块和整形模块。本发明系统包括:线性预测单元,其用于基于自适应滤波器对输入信号进行滤波;变换单元,其用于将滤波后的输入信号帧变换成变换域;量化单元,其用于量化变换域信号;长期预测单元,其用于基于对滤波输入信号先前段的重构来估计滤波后输入信号的帧;以及变换域信号组合单元,其用于在变换域中组合长期预测估计与变换后的输入信号,以生成输入到量化单元的变换域信号。

[0010] 该音频编码系统还可包括用于对滤波后输入信号的帧生成时间域重构的反量化和逆变换单元。此外,还可提供长期预测缓冲器,用来存储滤波后输入信号的先前帧的时间域重构。这些单元可布置在自量化单元至长期预测提取单元的反馈环路中,长期预测提取单元在长期预测缓冲器中搜索最佳匹配于滤波后输入信号的当前帧的重构片段 (segment)。另外,还可提供长期预测增益估计单元,其调节从长期预测缓冲器选择的片段

(段)的增益,以使得该片段最佳地匹配于当前帧。优选地,从变换域内的变换后输入信号中减去该长期预测估计。因此,可提供第二变换单元,用来将所选片段变换成变换域。该长期预测环路还可包括:在反量化之后和逆变换成时间域之前,在变换域中将长期预测估计添加到反馈信号。因此,可使用反向自适应长期预测方案,其在变换域中基于先前帧来预测滤波后输入信号的当前帧。为更加具有效率,如下文针对一些实例所提出的,可以不同的方式来进一步自适应该长期预测方案。

[0011] 对输入信号进行滤波的自适应滤波器优选地是基于线性预测编码(LPC)分析,其包括产生白化输入信号的LPC滤波器。输入数据的当前帧的LPC参数可由现有技术中已知的算法来确定。LPC参数估计单元可为输入数据帧计算任何适合的LPC参数表达式,例如,多项式、转移函数、反射系数、线谱频率等。用于编码或其它处理的LPC参数表达式的具体类型取决于相应的要求。如本领域的技术人员所知,一些表达式相比于其它操作更加适于某些操作,因此优选用于实施这些操作。线性预测单元可以以第一固定帧长(例如,20毫秒)进行操作。线性预测滤波还可以以扭曲频率轴进行操作,以相对于其它频率选择性地加强某些频率范围,例如,低频。

[0012] 应用于滤波后输入信号帧的变换优选地是以可变第二帧长进行操作的修正离散余弦变换(MDCT)。音频编码系统可包括窗体(window)序列控制单元,窗体(窗口)序列控制单元通过最小化整个输入信号块(包括数个帧)的编码成本函数(优选地是简化感知熵)的方式,为输入信号块确定用于重合MDCT窗体的帧长。因此,获得了将输入信号块分成具有相应第二帧长的MDCT窗体的最佳分段方式。由此,提出了一种变换域编码结构,其包括语音编码器元件,具有自适应长度的MDCT帧作为除LPC以外的所有处理的唯一基本单元。由于MDCT帧长可取很多不同的值,所以可找到最佳序列并且可避免帧尺寸突变,而这在仅应用小窗体尺寸和大窗体尺寸的现有技术中是普遍的。另外,一些在现有技术方法中用于在小窗体尺寸与大窗体尺寸之间过渡的具有陡沿(锋利的边)的过渡变换窗体没有必要。

[0013] 优选地,接连MDCT窗体的长度至多以因子2变化,和/或MDCT窗体长度是二元值。更具体地,MDCT窗体长度可以是输入信号块的二元分割。因此,MDCT窗体序列就局限于通过少量比特就容易编码的预定序列。另外,窗体序列具有帧尺寸的平滑过渡,由此排除了帧尺寸的突变。

[0014] 可提供用于对窗体序列内的MDCT窗体长度和窗体形状进行联合编码的窗体序列编码器。联合编码可去除冗余,并且需要较少的比特。当编码窗体序列的窗体长度和形状时,窗体序列编码器可考虑窗体尺寸限制,以省略可在解码器中重构的不必要信息(比特)。

[0015] 窗体序列控制单元还可配置为:当搜索使得输入信号块的编码成本函数最小化的MDCT窗体长度序列时,可为窗体长度候选考虑由长期预测单元生成的长期预测估计。在该实施例中,当确定MDCT窗体长度从而使应用于编码的MDCT窗体序列得到改善时,长期预测环路关闭。此外,还可提供时间扭曲单元,用于通过根据时间扭曲曲线对滤波后输入信号进行再采样,从而均匀地对准滤波后信号帧中的音高组分(分量)。优选地,可确定时间扭曲曲线来均匀地对准帧内的音高组分(component)。因此,变换单元和/或长期预测单元可以以具有恒定音高的时间扭曲信号进行操作,从而改善信号分析的准确性。

[0016] 该音频编码系统还可包括以可变速率、以递归方式编码线谱频率或由线性预测单元生成的其它适当 LPC 参数表达式,以存储和 / 或传输至解码器。根据一个实施例,提供线性预测内插单元,内插以对应于第一帧长的速率生成的线性预测参数,从而匹配于变换域信号的可变帧长。

[0017] 根据本发明一个方面,音频编码系统可包括感知建模单元,该感知建模单元通过对为 LPC 帧由线性预测单元生成的 LPC 多项式进行啁啾和倾斜处理,从而修改自适应滤波器的特性。通过修改自适应滤波器特性而收到的感知模型可在本系统中用于多种目的。例如,可将其应用为量化或长期预测中的感知权重函数。

[0018] 本发明的另一个独立方面涉及通过提供各独立的装置来编码输入信号的高频带组分,由此来扩展音频编码器的带宽。根据一个实施例,提供用于编码输入信号高频带组分的高频带编码器。优选地,高频带编码器是频谱频带复制 (SBR) 编码器。通过高频带编码器对高频带进行的独立编码允许在量化变换域信号时在量化单元中使用不同的量化步长,以便相比于输入信号中属于低频带的组分来编码变换域信号中属于高频带的组分。更具体地,量化器可应用高频带信号组分的较粗量化,其中高频带信号组分同样也是由高频带编码器进行编码,但这可降低比特率。

[0019] 根据另一实施例,提供用于将输入信号分成低频带组分和高频带组分的分频单元。然后,通过高频带编码器编码高频带组分,并且将低频带组分输入至线性预测单元并由上文提出的变换编码器来进行编码。优选地,分频单元包括正交镜像滤波器组以及正交镜像滤波器合成单元,其配置为用于对将输入到线性预测单元的输入信号进行下采样。来自正交镜像滤波器组的信号可直接输入到高频带编码器。当高频带编码器是可由正交镜像滤波器组直接馈送信号的频谱频带复制编码器时,这尤其有用。另外,正交镜像滤波器组与正交镜像滤波器合成单元的组合可用作下频带组分的额外下采样器。

[0020] 低频带与高频带之间的界限可以是可变的,且分频单元可动态地确定低频带与高频带之间的交越 (crossover) 频率。这允许 (例如) 基于输入信号性质和 / 或编码器带宽要求进行自适应频率分配。

[0021] 根据另一个方面,音频编码系统可包括将高频带组分转换成低通信号的第二正交镜像滤波器合成单元。可通过基于变换的第二编码器 (可能地,具有低解析度,即较大量化步长) 来编码该经过下调制的高频范围。当高频带还由其它装置 (例如,频谱频带复制编码器) 进一步编码时,这尤其有用。那么,组合这两种方式来编码高频带可更为有效。

[0022] 涵盖相同频率范围的不同信号表达式可通过信号表达式组合单元来组合,该表达式组合单元利用信号表达式中的相关性来降低必要的比特率。表达式组合单元还可生成指示信号表达式如何被组合的信令数据。该信令数据可存储或传输至解码器,以重构来自不同信号表达式的编码音频信号。

[0023] 还可在长期预测单元中提供频谱带复制单元以将能量引入长期预测估计的高频组分中。这可用于改善长期预测的效率。

[0024] 根据一个实施例,将具有左右输入信道的立体声信号输入至参数立体声单元,以计算包括输入信号单音表达式的立体声信号参数立体声表达式。然后,可将单音表达式输入至上文提出的 LPC 分析单元及后续的变换编码器。因此,获得了一种编码立体声信号的有效手段,其中基本上仅对单音表达式进行波形编码并且通过低比特率参数立体声表达式

实现了立体声效果。

[0025] 编码信号质量的进一步增强涉及使用谐波预测分析单元来预测频域 / MDCT 域中的谐波信号组分。

[0026] 本发明另一独立的编码器特殊方面涉及处理可变帧尺寸的比特储存器。在能编码可变长度帧的音频编码系统中,是通过在各个帧之间分配可用比特的方式来控制比特储存器。鉴于个别帧以及限定尺寸的比特储存器存在合理的难度度量 (measure),因此与所需恒定比特率的某一偏差可实现更好的总体质量,而不违背比特储存器尺寸所施加的缓冲器要求。本发明将使用比特储存器的概念扩展到以可变帧尺寸对通用音频编码解码器进行比特储存器控制。因此,音频编码系统可包括比特储存器控制单元,用于以基于帧长及帧的难度度量 (测度) 来确定许可用于编码滤波信号帧的比特数。优选地,比特储存器控制单元具有用于不同帧难度度量和 / 或不同帧尺寸的单控制方程式。可归一化不同帧尺寸的难度度量,以使得可以更加容易地对它们进行比对。为控制可变速率编码器的比特分配,比特储存器控制单元优选地将许可的比特控制算法的允许下限设定为最大允许帧尺寸的平均比特数。

[0027] 本发明还涉及这个方面:在变换编码器中量化 MDCT 线。不管编码器使用 LPC 分析还是使用长期预测,这个方面都适用。这种提出的量化策略是根据输入信号的特性 (例如,变换帧尺寸) 进行调整的。这表明,量化单元可基于变换单元所应用的帧尺寸来决定是通过基于模型的量化器,还是通过非基于模型的量化器来编码变换域信号。优选地,量化单元配置为通过基于模型的熵约束量化的方式、以小于阈值的帧尺寸来编码帧的变换域信号。该基于模型的量化可根据分类的参数进行调整。大尺寸帧可由 (例如) 标量量化器、通过 (例如) AAC 编码解码器中使用的 (例如) 基于 Huffman 的熵编码来量化。

[0028] 本发明优选实施例的另一方面是在 MDCT 线的不同量化方法之间进行切换。通过对不同变换尺寸使用不同的量化策略,编码解码器可在 MDCT 域中进行所有的量化和编码,而无需具有与变换域编码解码器并联或串联运行的专用时间域语音编码器。本发明教导了,对于其中存在 LTP 增益的语音类信号,优选地使用短变换和基于模型的量化器对该信号进行编码。基于模型的量化器尤其适用于短变换,并且如稍后概述的,其给出时间域语音专用矢量量化器 (VQ) 的优势,而仍操作于 MDCT 域中,并且不要求输入信号是语音信号。换句话说,当基于模型的量化器与 LTP 结合用于短变换时,可保持专用时间域语音编码器 VQ 的效率,而同时又不丧失通用性并且不离开 MDCT 域。

[0029] 另外,对于较为静止的音乐信号,优选地是使用相对较大尺寸的变换 (这在音频编码解码器中普遍使用),以及能利用由大尺寸变换所区分稀疏谱线的优势的量化方案。因此,本发明教导了使用这类量化方案来进行长变换。

[0030] 因此,根据帧尺寸来切换量化策略使得编码解码器通过简单地选择变换尺寸就能既保持专用语音编码解码器的性质,又能保持专用音频编码解码器的性质。这避免了现有技术系统中在低速率下力图将语音信号和音频信号处理得同样好的全部问题,这是因为这类系统都会不可避免地遇到有效组合时间域编码 (语音编码器) 与频域编码 (音频编码器) 的问题和困难。

[0031] 根据本发明另一方面,量化使用自适应步长尺寸。优选地,变换域信号组分的量化步长尺寸是基于线性预测和 / 或长期预测参数来进行自适应的。量化步长尺寸还可配置为

依据于频率。在本发明实施例中，量化步长尺寸是基于如下中的至少一个来确定：自适应滤波器的多项式、编码速率控制参数、长期预测增益值以及输入信号方差。

[0032] 本发明另一方面涉及长期预测 (LTP)，更具体来说，涉及 MDCT 域中的长期预测、MDCT 帧自适应 LTP 及 MDCT 加权的 LTP 搜索。无论变换编码器上游是否存在 LPC 分析，这些方面都适用。

[0033] 根据一个实施例，长期预测单元包括用于确定滞后值的长期预测提取器，该滞后值指定了滤波信号中的最佳适合于滤波信号当前帧的重构片段。长期预测增益估计器可估计应用于滤波信号的所选片段信号的增益值。优选地，确定滞后值和增益值以便最小化与长期预测估计和变换输入信号在感知域内的差有关的失真准则。该失真准则可与长期预测估计和变换输入信号在感知域内的差有关。优选地，通过搜索感知域内的滞后值和增益值来最小化该失真准则。当最小化失真准则时，可将修改后的线性预测多项式应用为 MDCT 域均衡增益曲线。

[0034] 长期预测单元可包括变换单元，用于将来自 LTP 缓冲器的片段的重构信号变换成变换域。为有效地实施 MDCT 变换，该变换优选地是 IV 型离散余弦变换。

[0035] 当滞后值小于 MDCT 帧长时，可使用虚拟矢量来生成重构信号的扩展片段。虚拟矢量优选地由叠入 - 展开的迭代过程来生成，以精化 (细化) 重构信号的生成片段。因此，在长期预测的滞后搜索过程期间，就生成了重构信号的尚不存在的片段。

[0036] 当变换单元以时间扭曲信号进行操作时，可基于时间扭曲曲线对长期预测缓冲器中的重构信号进行再采样。这使得时间扭曲 LPT 提取与时间扭曲 MDCT 相匹配。

[0037] 根据一个实施例，可提供用于编码长期预测滞后值和增益值的可变速率编码器，以实现低比特率。此外，长期预测单元还可包括噪声矢量缓冲器和 / 或脉冲矢量缓冲器，以增强对于 (例如) 噪音信号或瞬态信号的预测准确性。

[0038] 可提供用于对音高相关的信息 (例如，长期预测参数、谐波预测参数及时间扭曲参数) 进行联合编码的联合编码单元。联合编码还可通过利用这些参数内的相关性来降低必要的比特率。

[0039] 本发明的另一方面涉及用于解码由上述编码器实施例所生成的比特流的音频解码器。该音频解码器包括：用于对输入比特流的帧进行反量化的反量化单元；用于对变换域信号进行逆变换的逆变换单元；用于确定反量化帧的估计的长期预测单元；用于在变换域内进行组合的变换域信号组合单元；用于生成变换域信号的长期预测估计和反量化帧；以及用于对逆变换的变换域信号进行滤波的线性预测单元。

[0040] 另外，该解码器可包括上文关于编码器公开的许多方面。一般，解码器将镜像编码器的操作，尽管一些操作仅在编码器中实施，并且在解码器中没有对应的组件。因此，除非另有规定，否则关于编码器公开的内容被视为同样适用于解码器。

[0041] 本发明的上述方面可实施为装置、设备、方法或在可编程器件上操作的计算机程序。本发明的方面还可实施为信号、数据结构和比特流。

[0042] 因此，本申请会进一步公开音频编码方法和音频解码方法。示例性音频编码方法包括如下步骤：基于自适应滤波器对输入信号进行滤波；将滤波输入信号帧变换成变换域；量化变换域信号；基于滤波输入信号的先前片段的重构，估计滤波输入信号的帧；以及在变换域中，组合长期预测估计和变换输入信号，以生成变换域信号。

[0043] 示例性音频解码方法包括如下步骤：去量化输入比特流的帧；对变换域信号进行逆变换；确定对去量化帧的估计；在变换域中，组合长期预测估计和去量化帧，以生成变换域信号；对逆变换的变换域信号进行滤波；以及输出重构后的音频信号。

[0044] 这些只是本申请教导的优选的音频编码 / 解码方法和计算机程序的示例，并且本领域的技术人员可从下文关于示例性实施例的描述获得这些实例。

附图说明

[0045] 现将参照附图，以说明性实例的方式，而非限定本发明范围或精神的方式描述本发明，其中：

[0046] 图 1 显示根据本发明的编码器和解码器的一个优选实施例；

[0047] 图 2 显示根据本发明的编码器和解码器的更详细的图示；

[0048] 图 3 显示根据本发明的编码器的另一实施例；

[0049] 图 4 显示根据本发明的编码器的一个优选实施例；

[0050] 图 5 显示根据本发明的解码器的一个优选实施例；

[0051] 图 6 显示根据本发明的 MDCT 线编码和解码的一个优选实施例；

[0052] 图 7 显示本发明的结合 SBR 编码器的一个优选实施例；

[0053] 图 8 显示立体声系统的一个优选实施例；

[0054] 图 9 显示根据本发明的核心编码器与高频重构编码的更为复杂整合的一个优选实施例；

[0055] 图 10 显示根据本发明的 SBR 编码与核心编码器的组合的一个优选实施例；

[0056] 图 11 显示根据本发明的编码器和解码器的一个优选实施例、以及从一方传输到另一方的相关控制数据的一个实例；

[0057] 图 11a 是根据本发明一个实施例的编码器的方面的另一图示；

[0058] 图 12 显示根据本发明一个实施例的窗体序列以及 LPC 数据与 MDCT 数据之间相关性的实例；

[0059] 图 13 显示根据本发明的缩放因子数据与 LPC 数据的组合；

[0060] 图 14 显示根据本发明的将 LPC 多项式转换 (translate) 成 MDCT 增益曲线的一个优选实施例；

[0061] 图 15 显示根据本发明的将恒定更新速率的 LPC 参数映射成自适应 MDCT 窗体序列数据的一个优选实施例；

[0062] 图 16 显示根据本发明的基于变换尺寸和量化器类型对感知加权滤波器计算进行自适应的一个优选实施例。

[0063] 图 17 显示根据本发明的取决于帧尺寸对量化器进行自适应的一个优选实施例；

[0064] 图 18 显示根据本发明的取决于帧尺寸对量化器进行自适应的一个优选实施例；

[0065] 图 19 显示根据本发明的根据 LPC 和 LTP 数据对量化步长尺寸进行自适应的一个优选实施例；

[0066] 图 19a 显示如何借助 Δ - 自适应模块从 LPC 参数和 LTP 参数导出 Δ - 曲线；

[0067] 图 20 显示根据本发明的利用随机偏移 (offset) 的基于模型量化器的一个优选实施例；

- [0068] 图 21 显示根据本发明的基于模型量化器的一个优选实施例；
- [0069] 图 21a 显示根据本发明的基于模型量化器的另一优选实施例；
- [0070] 图 22 显示根据本发明的在 LTP 环路中使用 SBR 模块的一个优选实施例；
- [0071] 图 23a 示意地显示根据本发明一个实施例的 MDCT 变换的相邻窗体。
- [0072] 图 23b 描述本发明的使用四个不同 MDCT 窗体形状的一个实施例；
- [0073] 图 23c 描述根据本发明一个实施例的窗体序列编码方法的一个实例；
- [0074] 图 24 显示根据本发明的 MDCT 域内的谐波预测的一个优选实施例；
- [0075] 图 25 显示根据本发明的 LTP 提取精化过程；
- [0076] 图 25a 显示 MDCT 自适应的 LTP 提取过程；
- [0077] 图 25b 显示初始的 LTP 提取信号的迭代精化；
- [0078] 图 25c 显示精化单元的一个可选实施方案；
- [0079] 图 25d 显示精化单元的另一可选实施方案；
- [0080] 图 26 显示根据本发明的组合用于谐波预测、LTP 和时间扭曲的控制数据的一个优选实施例；
- [0081] 图 27 显示根据本发明的通过噪声和脉冲缓冲器扩展 LTP 搜索的一个优选实施例；
- [0082] 图 28a 显示比特储存器控制的基本概念；
- [0083] 图 28b 显示根据本发明的可变帧尺寸的比特储存器控制的概念；
- [0084] 图 29 显示根据本发明的以时间扭曲 MDCT 为背景的 LTP 搜索和应用；
- [0085] 图 29a 显示时间扭曲 MDCT 分析的效果；
- [0086] 图 30 显示根据本发明的在 MDCT 域和 QMF 域中的组合 SBR。

具体实施方式

[0087] 下文所述实施例仅为举例说明本发明关于音频编码器和解码器的原则。应理解的是，本文所述布置和细节的修改和变化对于本领域的技术人员很明显。因此，本发明意图在于仅由权利要求书的范围来限定，而不由通过描述和解释本文实施例而提出的具体细节来限定。相似的实施例组件由相似的附图标记来表示。

[0088] 在图 1 中，显示了编码器 101 和解码器 102。编码器 101 获取时间域输入信号并产生随后发送至解码器 102 的比特流 103。解码器 102 基于接收的比特流 103 产生输出波形。该输出信号在心理声学上与原始输入信号相似。

[0089] 图 2 中显示了编码器 200 和解码器 210 的一个优选实施例。编码器 200 中的输入信号通过 LPC(线性预测编码)模块 201，LPC 模块 201 为具有第一帧长和对应线性预测参数的 LPC 帧生成白化残余信号。另外，增益归一化可包含在 LPC 模块 201 中。来自 LPC 的残余信号由以第二可变帧长进行操作的 MDCT(修正离散余弦变换)模块 202 变换成频域。图 2 所示的编码器 200 中包括了 LTP(长期预测)模块 205。LTP 将在本发明的另一实施例中进一步详述。MDCT 线经过量化 203 并且也经过去量化 204，以便给 LTP 缓冲器馈送供解码器 210 使用的解码输出的副本。由于量化失真，所以该副本被称为是相应输入信号的重构。图 2 下半部分中显示了解码器 210。解码器 210 获取已量化的 MDCT 线，对其进行去量化 211，添加来自 LTP 模块 214 的贡献，并进行逆 MDCT 变换 212，之后是 LPC 合成滤波器

213。

[0090] 上述实施例的一个重要方面是：尽管 LPC 具有其自身的帧尺寸（在一个实施例中恒定）且 LPC 参数也经过编码，但 MDCT 帧是用于编码的唯一基础单元。该实施例由变换编码器开始，并引入来自语音编码器的基本预测和整形模块。如将在稍后讨论的，MDCT 的帧尺寸是可变的，并且通过使得简化的感知熵成本函数最小化来确定整个块的最佳 MDCT 窗体序列，由此使可变帧尺寸与输入信号块自适应。这允许缩放，以保持最佳的时间 / 频率控制。此外，这种提出的统一结构避免了不同编码范例的切换组合或分层组合。

[0091] 图 3 中更加详细地示意描述了编码器 300 的部分。从图 2 编码器中的 LPC 模块 201 输出的白化信号被输入至 MDCT 滤波器组 302。任选地，MDCT 分析可以是确保信号音高（如果该信号是周期性的，具有很好定义的音高）在 MDCT 变换窗体上保持恒定的时间扭曲 MDCT 分析。

[0092] 图 3 中更加详细地描述了 LTP 模块 310。LTP 模块 310 包括保存着先前输出信号片段的重构时间域样本的 LTP 缓冲器 311。LTP 提取器 312 根据当前输入片段在 LTP 缓冲器 311 中找出最佳匹配的片段。在从当前正输入到量化器 303 的片段减去该增益值之前，增益单元 313 将适合的增益值应用到该片段。明显地，为了在量化之前进行减法运算，LTP 提取器 312 同样将所选的信号片段变换到 MDCT 域。当组合先前重构的输入信号片段与变换的 MDCT 域输入帧时，LTP 提取器 312 会搜索使感知域内的误差函数最小化的最佳增益值和滞后值。例如，来自 LTP 模块 310 的变换后的重构片段与变换输入帧（即，减法之后的残余信号）之间的均方误差（MSE）函数被优化。该优化可在感知域内实施，其中频率组分（即，MDCT 线）根据其感知重要性进行加权。LTP 模块 310 在 MDCT 帧单元中进行操作，并且，例如对于在量化模块 303 中的量化，编码器 300 一次考虑一个 MDCT 帧残余。可在感知域内实施滞后和增益搜索。任选地，LTP 10 可以具有频率选择性，即根据频率对增益和 / 或滞后进行适应。图中显示了反量化单元 304 和逆 MDCT 单元 306。稍后将解释 MDCT 可以时间扭曲。

[0093] 图 4 中显示编码器 400 的另一实施例。除了图 3 以外，为清楚起见，图 4 中包含 LPC 分析 401。图中显示用于将所选信号片段变换到 MDCT 域的 DCT-IV 变换 414。另外，还显示了为 LTP 片段选择计算最小误差的数种方式。除了图 4 中所示的最小化残余信号（图 4 中表示为 LTP2）以外，图 4 中还显示了在逆变换到存储在 LTP 缓冲器 411 中的重构时间域信号之前，最小化变换输入信号与反量化 MDCT 域信号之间的差（指示为 LTP3）。该 MSE 函数的最小化会将 LTP 贡献引向变换输入信号与存储在 LTP 缓冲器 411 中的重构输入信号的最佳相似度（尽可能）。另一可选误差函数（指示为 LTP 1）是基于这些信号在时间域内的差。在该情况下，最小化 LPC 滤波输入帧与在 LTP 缓冲器 411 中的对应时间域重构之间的 MSE。有利地，基于 MDCT 帧尺寸来计算 MSE，MDCT 帧尺寸可能不同于 LPC 帧尺寸。另外，量化器和去量化器块由频谱编码块 403 和频谱解码块 404（“Spec enc”和“Spec dec”）替代，如图 6 中所示，频谱编码块 403 和频谱解码块 404 可包含除量化以外的额外模块。同样，MDCT 和逆 MDCT 可时间扭曲（WMDCT、IWMDC）。

[0094] 图 5 中显示了本发明所提出的解码器 500。来自所接收比特流的频谱数据经过反量化 511，并且添加由 LTP 提取器从 LTP 缓冲器 515 提供的 LTP 贡献。图中也显示了解码器 500 中的 LTP 提取器 516 和 LTP 增益单元 517。这些合计的 MDCT 线由 MDCT 合成模块合成

到时间域,且时间域信号由 LPC 合成滤波器 513 进行频谱整形。任选地,MDCT 合成可以是时间扭曲的 MDCT,和 / 或 LPC 合成滤波可以是频率扭曲。

[0095] 频率扭曲的 LPC 是基于频率轴的非均匀采样,以便在确定 LPC 滤波器参数时允许对 LPC 误差贡献进行频率选择控制。虽然平常的 LPC 是基于在线性频率轴上最小化 MSE 以使得 LPC 多项式在频谱波峰区域内基本上准确,然而在确定 LPC 滤波器参数时,频率扭曲的 LPC 允许频率选择集中。例如,当以较高带宽(例如,16 或 24kHz 的采样速率)进行操作时,扭曲频率轴允许将较低频带上的 LPC 多项式准确性集中直至 4kHz。

[0096] 图 6 中将更加详细地描述图 4 的频谱解码块 403 和频谱编码块 404。图中显示在右边的频谱编码块 603 在一个实施例包括谐波预测分析模块 610、TNS 分析(时间噪声整形)模块 611、之后是 MDCT 线的比例因子缩放模块 612、以及最后是 Enc 线模块 613 中的线的量化和编码。显示于图中左边的解码器的频谱解码块 604 进行该逆过程,即,所接收的 MDCT 线在 Dec 线模块 620 进行去量化,且由比例因子(SCF)缩放模块 621 来取消缩放。下文将解释应用 TNS 合成 622 和谐波预测合成 623。

[0097] 图 7 中描述本发明的另一优选实施例。除了已描述的 LPC 701、MDCT 量化 704 和 LTP 705 以外,增加了 QMF 分析模块 710 和 QMF 合成模块 711 以及 SBR(频谱频带复制)模块 712。QMF(正交镜像滤波器)滤波器组具有某一数量的子带,在该具体实例中为 64。复杂的 QMF 滤波器组允许对子带进行独立操控,并不会引入频域假像而高于在使用原型滤波器的情况下的假像拒斥水平。在该具体实例 32 中,然后将某一数量的较低(频率)子带合成到时间域,从而通过因子 2 形成下采样信号。如前所述,这就是输入到编码器模块的输入信号。使用 QMF 分析和合成模块作为再采样器,确保了 LPC 仅在减小后的带宽上操作,后续的变换编码器也在该带宽上进行编码。将较高 32 的子带发送至从高频带原始信号提取相关 SBR 参数的 SBR 编码器模块 712。可选地,将输入信号供应至 QMF 分析模块,其进而连接至 SBR 编码器和下采样模块,如前所述下采样模块产生用于变换编码器模块的下采样信号。

[0098] SBR(频谱频带复制)提供了一种编码频谱中高频部分的有效方式。SBR 通过音频信号的低频来再生成音频信号的高频、以及少量的附加控制信息。由于 SBR 方法能够减小核心编码器的带宽,并且 SBR 技术对频率范围进行编码所需的比特率显著少于波形编码器,所以通过降低分配给波形核心编码器的比特率可实现编码增益,同时保持全音频带宽。当然,通过降低核心编码器与 SBR 部分之间的交越(crossover)频率,这给出了几乎连续地降低总数据率的可能性。

[0099] 通过对量化噪声进行整形,使其始终被信号所掩蔽,感知音频编码器可降低比特率。这会导致很低的信噪比,但只要保持量化噪声低于掩蔽曲线,这不会有太大影响。该量化所代表的失真是听不见的。然而,当以低比特率进行操作时,将超出掩蔽阈值,并且失真变得听得见。感知音频编码器能采用的一种方法是对该信号进行低通滤波,即,仅编码该频谱的部分,只因为没有足够的比特来编码信号的整个频率范围。SBR 算法对于这种情况是非常有益的,因为 SBR 能够以低比特率实现全音频带宽。

[0100] SBR 解码的概念包括如下几个方面:

[0101] • 通过从低频带复制低通信号排除低频,来完成高频带再生成。

[0102] • 将频谱包络信息从编码器发送到解码器,确保重构的高频带的粗频谱包络正确。

[0103] • 还可将设计为补偿高频重构缺点的附加信息从编码器传输至解码器。

[0104] • 另外的手段（例如，反向滤波、噪声和正弦添加，所有这些手段都由所传输的信息来引导）可补偿任何频带扩展方法的源自低频带与高频带之间存在的偶尔基本差异的缺点。

[0105] 在图 8 中，通过如下方式，本发明的一个实施例扩展到立体声：添加两个用于左和右信道的 QMF 分析滤波器组 820、821、以及称为参数立体声 (PS) 模块的旋转模块 830，旋转模块 830 通过 QMF 域内的两个输入信号再生成了两个新的信号以及对应的旋转参数。这两个新的信号代表单缩混信号以及残余信号。可将它们视为左 / 右立体声信号的中 / 侧变换，其中中 / 侧立体声空间经过旋转，从而中信号（即，缩混信号）内的能量最大，且侧信号（即，残余信号）内的能量被最小化。作为一个具体实例，向左或向右展开 45 度的单音源将（以不同的电平）存在于左信道和右信道两个信道内。现有技术的波形音频编码器通常会选择独立地或作为中 / 侧表达式地编码左信道，或者会选择独立地或作为中 / 侧表达式地编码右信道。对于该具体示例，由于无论表达式如何展开的单音源都将存在于两个信道中，所以左 / 右表达式以及中 / 侧表达式都将不会是有意义的。然而，如果中 / 侧表达式旋转 45 度，则展开的单音源将完全结束于旋转后的中信号（在此称其为缩混信号）中，且旋转后的侧信号将为零（在此称为残余信号）。这就提供了相对于普通左 / 右编码或中 / 侧编码的编码优势。

[0106] 随后，可将代表结合所提取的参数的立体声信号的这两个新信号输入到（例如）图 7 中所示的 QMF 合成模块和 SBR 模块。对于低比特率，可低通滤波掉或完全略去该残余信号。该参数立体声解码器将用去相关版本的缩混信号取代略去的残余信号。当然，在此提出的立体声信号处理也可以与本发明其它实施例结合。

[0107] 更加详细地，PS 模块会比较对应时间 / 频率倾斜 (tilt) 的两个输入信号（左和右）。倾斜频带被设计为接近心理声学激励缩放，而同时片段的长度很好地匹配于双耳听觉系统的已知限制。基本上，每个时间 / 频率倾斜，提取代表在感知上最为重要的空间性质的三个参数：

[0108] (i) 信道间级别差，其代表信道之间的级别差，这类似于混合控制台上的“声象移动器 (pan pot)”。

[0109] (ii) 信道间相位差 (IPD)，其代表信道之间的相位差。在频域中，该特征几乎可与信道间时间差 (ITD) 互换。该 IPD 可由附加整体相位差 (OPD) 扩大，描述左和右相位调节的分布。

[0110] (iii) 信道间相关性 (IC)，其代表信道间的相关性或互关性。尽管第一两个参数与音源方向耦合，但第三个参数更多地与音源的空间扩散性相关联。

[0111] 在参数提取之后，对输入信号进行缩混以形成单信号。可通过求和过程的普通方法来完成缩混，但优选地可通过包含时间对准和能量保存技术的更为先进的方法来避免缩混中潜在的相位消减。在解码器侧，提供有 PS 解码模块，该 PS 解码模块基本上包括对应编码器的逆过程，并基于 PS 参数来重构立体声输出信号。

[0112] 图 9 中显示本发明的另一实施例。此处，输入信号再次由 64-子带信道 QMF 模块 920 分析。此外，与图 7 中所示的系统相反，核心编码器覆盖范围与 SBR 编码器覆盖范围之间的界限是可变的。因此，该系统在模块 911 中合成为覆盖时间域信号的带宽所需同样多

的子带,时间域信号随后由 LPC、MDCT 和 LTP 模块 901 进行编码。其余(频率较高)子带样本被输入至 SBR 编码器 912。

[0113] 除了早先的示例以外,还可将高频子带样本输入至 QMF 合成模块 920,该 QMF 合成模块 920 将较高频率范围合成到低通信号,由此包含下调制高频范围。随后,附加的基于 MDCT 的编码器 930 来编码该信号。来自附加的基于 MDCT 的编码器 930 可在任意的组合单元 940 中与 SBR 编码器输出相组合。产生信令并将信令发送至解码器,该解码器指示哪一部分由 SBR 编码,哪一部分由基于 MDCT 的波形编码器编码。这使得能够从 SBR 编码到波形编码平滑过渡。此外,由于用单独的 MDCT 变换对信号进行编码,所以能够关于 MDCT 编码中用于较低频和较高频的变换尺寸进行自由地选择。

[0114] 图 10 中显示另一实施例。输入信号被输入至 QMF 分析模块 1010。对应于 SBR 范围的输出子带输入至 SBR 编码器 1012。LPC 分析和滤波是通过覆盖信号的全频范围的方式完成,并且是直接使用输入信号或者使用 QMF 合成模块 1011 产生的合成版本的 QMF 子带信号来完成。当与图 8 的立体声实施结合时,后者更为有用。LPC 滤波信号输入至 MDCT 分析模块 1002,来提供待编码的频谱线。在该本发明的这个实施例中,量化 1003 布置为使得 SBR 区域中发生显著较粗的量化(即,频率区域同样由 SBR 编码器覆盖),从而仅仅覆盖最强的频谱线。该信息被输入至组合单元 1040,该组合单元 1040 在给定量化频谱和 SBR 编码数据的情况下向解码器提供信令,指示在 SBR 范围内针对不同的频率范围使用何种信号,即, SBR 数据或波形编码数据。

[0115] 图 11 中显示本发明编码系统的非常概括的图示。该示例性编码器获取输入信号并产生除其它数据外还包含以下数据的比特流:

- [0116] ● 量化 MDCT 线;
- [0117] ● 比例因子;
- [0118] ● LPC 多项表达式;
- [0119] ● 信号片段能量(例如,信号方差);
- [0120] ● 窗体序列
- [0121] ● LTP 数据;

[0122] 根据该实施例的解码器读取所提供的比特流,并产生在心理声学上相似于原始信号的音频输出信号。

[0123] 图 11a 是根据本发明一个实施例的编码器 1100 的各方面的另一图示。编码器 1100 包括 LPC 模块 1101、MDCT 模块 1104、LTP 模块 1105(仅简化显示)、量化模块 1103 及用于将重构信号馈送回到 LTP 模块 1105 的反量化模块 1104。此外,还提供了:音高(pitch)估计模块 1150,其用于估计确定输入信号的音高;以及窗体序列确定模块 1151,其用于确定较大输入信号块(例如,1 秒)的最佳 MDCT 窗体序列。在该实施例中,是基于开路的方法来确定 MDCT 窗体序列,其中确定最小化编码成本函数(例如,简化感知熵)的 MDCT 窗体尺寸候选值序列。当搜索最佳 MDCT 窗体序列时,可选地,可考虑被窗体序列确定模块 1151 最小化的 LTP 模块 1105 对编码成本函数的贡献。优选地,为每一评估的窗体尺寸候选值,确定对对应于窗体尺寸候选值的 MDCT 帧的长期预测贡献,并且确定相应的编码成本。一般,短的 MDCT 帧尺寸更加适合于语音输入,而对于音频信号,优选为具有细化频谱解析度的长变换窗体。

[0124] 感知权重或感知加权函数是基于 LPC 模块 1101 所计算的 LPC 参数来确定的,这在下文将更加详细地解释。将感知权重供应至两者都操作于 MDCT 域的 LTP 模块 1105 和量化模块 1103,以根据其相应的感知重要性对频率组分(分量)的误差贡献或失真贡献进行加权。图 11a 还显示,哪些编码参数会优选地由稍后论述的适当编码方案传输至解码器。

[0125] 接下来,将论述 LPC 和 MDCT 数据的共存、以及 LPC 在 MDCT 中的效果的仿真,两者都用于反作用及实际滤波省略。

[0126] 根据一个实施例,LP 模块对输入信号进行滤波,以去除信号的频谱形状,且 LP 模块随后的输出是频谱平信号。这对(例如)LTP 的操作是有利的。然而,知晓原始信号在 LP 滤波之前是何种频谱形状对于对频谱平信号进行操作的编码解码器的其它部分是有益处的。由于编码器模块在滤波之后对频谱平信号的 MDCT 变换进行操作,所以本发明教导了通过将所使用的 LP 滤波器的转移函数(即,原始信号的频谱包络)映射成施加于频谱平信号 MDCT 表达式的频率点(bin)的增益曲线或均衡曲线,可以将原始信号在 LP 滤波以前的频谱形状重新施加于频谱平信号的 MDCT 表达式。相反地,LP 模块可略去实际的滤波,并且仅对随后映射到可施加于信号的 MDCT 表达式的增益曲线的转移函数进行估计,因此不再需要输入信号的时间域滤波。

[0127] 本发明实施例的一个突出方面是使用对 LPC 白化后的信号进行灵活窗体分段的方式来操作基于 MDCT 的变换编码器。这概括于图 12 中,其中给出了示例性的 MDCT 窗体序列以及 LPC 的窗体化。因此,如图清楚显示,LPC 以恒定帧尺寸(例如,20ms)进行操作,而 MDCT 以可变窗体序列(例如,4 至 128ms)进行操作。这允许独立地选择 LPC 的最佳窗体长度以及 MDCT 的最佳窗体序列。

[0128] 图 12 进一步显示以第一帧率生成的 LPC 数据(具体来说,LPC 参数)与以第二可变速率生成的 MDCT 数据(具体来说,MDCT 线)之间的关系。图中向下的箭头标志着内插于 LPC 帧(循环)之间以匹配于对应 MDCT 帧的 LPC 数据。例如,为 MDCT 窗体序列所确定的时间点内插 LPC 生成的感知加权函数。向上的箭头标志着用于 MDCT 线编码的精细化数据(即,控制数据)。对于 AAC 帧,该数据通常是比例因子,且对于 ECQ 帧,该数据通常是方差修正数据等。虚线对实线代表在给定量化器的情况下对于 MDCT 线编码最重要的数据。向下的双箭头标志着编码谱线。

[0129] 可利用共存于编码器中的 LPC 数据和 MDCT 数据,例如,以便通过将根据 LPC 参数所估计的感知掩蔽曲线进行考虑的方式来降低编码 MDCT 比例因子的比特要求。此外,可在确定量化失真时使用由 LPC 导出的感知加权。如图所示且如下文论述的,取决于所接收数据的帧尺寸(即,对应于 MDCT 帧尺寸或窗体尺寸),量化器以两种模式进行操作,并生成两种类型的帧(ECQ 帧和 AAC 帧)。

[0130] 图 15 显示将恒定速率的 LPC 数据映射成自适应 MDCT 窗体序列数据的一个优选实施例。LPC 映射模块 1500 根据 LPC 更新率来接收 LPC 参数。另外,LPC 映射模块 1500 接收 MDCT 窗体序列上的信息。然后,LPC 映射模块 1500 生成 LPC 到 MDCT 的映射,例如,将基于 LPC 的心理声学数据映射成以可变 MDCT 帧率生成的相应 MDCT 帧。例如,LPC 映射模块针对应于 MDCT 帧的时间点内插 LPC 多项式或相关数据,以将其用作(例如)LTP 模块或量化器中的感知权重。

[0131] 现将参照图 13 论述基于 LPC 的感知模型的具体细节。在本发明一个实施例中,LPC

模块 1301 自适应于通过使用 16 阶 -16kHz 采样率的信号的线性预测来产生白输出信号。例如,来自图 2 的 LPC 模块 201 的输出是 LPC 参数估计及滤波之后的残余。如图 13 左下部示意显示的,经估计的 LPC 多项式可通过带宽扩展因子进行啁啾,并且此外,在本发明的一个实施方案中,经估计的 LPC 多项式通过修改对应 LPC 多项式的第一反射系数而发生倾斜。啁啾通过将多项式的极点向内移动到单位圆内来扩展 LPC 转移函数中波峰的带宽,从而导致较“软”的波峰。倾斜使 LPC 转移函数较平坦,以平衡较低和较高频率的影响。这些修改是力图根据可供系统的编码器和解码器边使用的估计 LPC 参数来生成感知掩蔽曲线 $A'(z)$ 。下文在图 16 中展示调整 PC 多项式的操控细节。

[0132] 在本发明一个实施方案中,针对 LPC 残余的 MDCT 编码操作具有用于控制量化器解析度或量化步长尺寸(以及因此量化所引入的噪声)的比例因子。比例因子估计模块 1360 在原始输入信号上对这些比例因子进行估计。例如,这些比例因子是从根据原始信号估计的感知掩蔽阈值曲线推导出的。在一个实施例中,可使用单独的频率变换(可能具有不同的频率解析度)来确定掩蔽阈值曲线,但这并不总是必要的。可选地,可从变换模块所生成的 MDCT 线来估计掩蔽阈值曲线。图 13 的右下部分示意地显示由比例因子估计模块 1360 生成的比例因子,来用于对量化进行控制,以使得所引入的量化噪声局限于听不见的失真。

[0133] 如果 LPC 滤波器连接在 MDCT 变换模块的上游,则白化信号被变换到 MDCT 域。由于该信号具有白频谱,所以由该信号来推导感知掩蔽曲线就不是很适合了。因此,当估计掩蔽阈值曲线和 / 或比例因子时,可使用生成用于补偿频谱白化的 MDCT 域均衡增益曲线。这是因为需要对具有原始信号绝对频谱性质的信号来估计比例因子,以正确地对掩蔽进行感知估计。

[0134] 下文参照图 14 更加详细地论述由 LPC 多项式来计算 MDCT 域均衡增益曲线。

[0135] 使用上文概述的方法,在编码器与解码器之间传输的数据包含当使用基于模型的量化器时可由它推导出相关感知信息和信号模型的 LP 多项式以及在变换编码解码器中普遍使用的比例因子。

[0136] 更详细地,返回到图 13,图中的 LPC 模块 1301 由输入信号来估计该信号的频谱包络 $A(z)$,并推导出该感知表达式 $A'(z)$ 。另外,通过输入信号来估计通常在基于变换的感知音频编码解码器中使用的比例因子,或者如果在比例因子估计中考虑了 LP 滤波器的转移函数(如下文图 14 上下文中描述的),则可以对 LP 滤波器所产生的白信号来进行估计。然后,如下文概述的,在给定 LP 多项式的情况下,可在比例因子自适应模块 1361 中进行比例因子的自适应,以降低传输比例因子所需的比特率。

[0137] 通常地,将比例因子传输至解码器,且 LP 多项式也是如此。现在,在比例因子和 LP 多项式两者都是由原始输入信号估计出的并且其与原始输入信号的绝对频谱性质都有一些相关性的情况下,所以本文提出对两者之间的 Δ 表达式进行编码,以去除在对它们进行单独传输情况下可能发生的冗余。根据一个实施例,如下地利用该相关性。由于 LPC 多项式在经过正确啁啾和倾斜时力图表达掩蔽阈值曲线,所以可组合这两个表达式,以使得变换编码器的传输比例因子表达所期望的比例因子与可从传输的 LPC 多项式推导出的比例因子之间的差。因此,图 13 中所示的比例因子自适应模块 1361 计算出由原始输入信号生成的期望比例因子与由 LPC 导出的比例因子之间的差。这个方面保留了既具备基于 MDCT 量化器的同时,仍有可能切换到基于模型的量化器的能力;基于 MDCT 的量化器具有如下的

概念：在 LPC 结构内，普遍使用于变换编码器中的比例因子对 LPC 残余进行操作；而基于模型的量化器仅由线性预测数据导出量化步长尺寸。

[0138] 图 14 显示将 LPC 多项式转换成 MDCT 增益曲线的一个优选实施例。如图 2 所显示的，MDCT 对白化信号进行操作，该白化信号是由 LPC 滤波器 1401 白化的。为保留原始输入信号的频谱包络，通过 MDCT 增益曲线模块 1470 来计算 MDCT 增益曲线。对于 MDCT 变换中点所代表的频率，可通过估计 LPC 滤波器描述的频谱包络的幅值响应来获得 MDCT 域均衡增益曲线。然后，例如当计算图 3 所示的最小均方差信号时，或当为比例因子确定（如上参照图 13 所概述的）估计感知掩蔽曲线时，可将该增益曲线应用于 MDCT 数据。

[0139] 图 16 显示基于变换尺寸和 / 或量化器的类型来对感知加权滤波器计算进行自适应的一个优选实施例。在图 16 中，LP 多项式 $A(z)$ 由 LPC 模块 1601 来估计。LPC 参数修改模块 1671 接收 LPC 参数（例如，LPC 多项式 $A(z)$ ），并通过修改 LPC 参数来生成感知加权滤波器 $A'(z)$ 。例如，扩展 LPC 多项式 $A(z)$ 的带宽并 / 或使多项式倾斜。自适应啁啾和倾斜模块 1672 的输入参数是默认的啁啾值和倾斜值 ρ 和 γ 。 ρ 和 γ 在给出预定规则的情况下是基于所使用的变换尺寸和 / 或所使用的量化策略 Q 来修改的。修改后的啁啾参数和倾斜参数 ρ' 和 γ' 输入至 LPC 参数修改模块 1671，从而将由 $A(z)$ 表达的输入信号频谱包络转换成由 $A'(z)$ 表达的感知掩蔽曲线。

[0140] 下文中将解释以帧尺寸进行调整的量化策略、及以根据本发明一个实施例的分类参数进行调整的基于模型的量化。本发明的一个方面是针对不同的变换尺寸或帧尺寸，使用不同的量化策略。这显示于图 17 中，其中帧尺寸是作为使用基于模型的量化器还是使用非基于模型的量化器的选择参数来使用。须注意到，该量化方面独立于本发明公开的编码器 / 解码器的其它方面，并且同样可在其它编码解码器中应用。非基于模型的量化器的一个实例是在 AAC 音频编码标准中使用的基于 Huffman 表格的量化器。基于模型的量化器可以是使用算术编码的熵约束量化器 (ECQ)。然而，也可在本发明实施例中使用其它的量化器。此外，在本发明的当前概述的实施例中，所选择的量化器会按照变换尺寸向解码器隐性发送信号。应清楚的是，也可使用其它的信令方式，例如，向解码器显性地发送关于针对具体帧尺寸使用何种量化策略的信息。

[0141] 根据本发明的一个独立方面，提出了根据帧尺寸在不同量化策略之间进行切换，以在给出的特定帧尺寸的情况下能使用最佳的量化策略。作为一个实例，窗体序列可命令 (dictate) 针对信号的非常静止的单音音乐片段使用长变换。对于这种使用长变换的特定信号类型，使用能利用信号频谱中的“稀疏”特性（即，很好定义的离散音）的量化策略是非常有益的。

[0142] 在 AAC 中与 Huffman 表格结合使用的量化方法以及同样在 AAC 中使用的频谱线的分组是非常有益的。然而，反过来，对于语音片段，窗体序列可在给出 LTP 编码增益的情况下命令（指示）使用短变换。对于该信号类型和变换尺寸，采用如下策略是有益的：不试图寻找或引入频谱中的稀疏性，而是保持宽带能量，其在给定 LTP 情况下保留原始输入信号的脉冲类特征。

[0143] 图 18 给出了所述概念更为一般化的显示，其中将输入信号变换成 MDCT 域，随后由受到针对 MDCT 变换所使用的变换尺寸或帧尺寸控制的量化器进行量化。

[0144] 根据本发明另一方面，量化步长尺寸是根据 LPC 和 / 或 LTP 数据进行自适应的。这

允许取决于帧的难度来确定步长尺寸,并且允许控制分配用于对帧进行编码的比特数。图 19 中给出 LPC 和 LTP 数据可以如何控制基于模型的量化的示例。图 19 的上部分给出 MDCT 线的示意。下部分绘示了作为频率的函数的量化步长尺寸 Δ 。由该特定实例清楚的看出,量化步长尺寸随频率增大而增大,即,较高频率引起更多的量化失真。通过图 19a 中绘示的 Δ 自适应模块,由 LPC 和 LTP 参数来导出 Δ 曲线。此外,还可通过参照 13 解释的啁啾和/或倾斜,由预测多项式 $A(z)$ 来导出 Δ 曲线。

[0145] 如下等式给出了由 LPC 数据导出的优选感知加权函数:

$$[0146] \quad P(z) = \frac{1 - (1 - \tau)r_1 z^{-1}}{A(z/\rho)}$$

[0147] 其中 $A(z)$ 是 LPC 多项式, τ 是倾斜参数, ρ 控制啁啾,且 r_1 是由 $A(z)$ 多项式计算出的第一反射系数。应注意的是,可对不同的表达式的分类重新计算 $A(z)$ 多项式,以从多项式提取相关的信息。如果关注频谱的斜率,从而想应用倾斜来抵消频谱的斜率,则优选地可重新计算反射系数的多项式,这是因为第一反射系数代表频谱的斜率。

[0148] 另外, Δ 值可作为输入信号方差 σ 、LTP 增益 g 以及由预测多项式导出的第一反射系数 r_1 的函数进行自适应。例如,自适应可以基于如下等式:

$$[0149] \quad \Delta' = \Delta (1 + r_1 (1 - g^2))$$

[0150] 下文将概述根据本发明一个实施例的基于模型量化器的各方面。图 20 显示基于模型的量化器的其中一个方面。MDCT 线输入至一个使用均匀标量量化器的量化器。另外,随机偏移输入至量化器,并且将随机偏移用作使间隔边界移位的量化间隔的偏移值。该提出的量化器提供矢量量化的优势的同时,还保持了标量量化器的可搜索性。量化器对一组不同的偏移值进行迭代操作,并计算这些偏移值的量化误差。将使得正处于量化的特定 MDCT 线的量化失真最小化的偏移值(或偏移值矢量)被用于量化。然后,将该偏移值与量化后的 MDCT 线一同传输至解码器。随机偏移的使用会在去量化后的解码信号中引入噪声填充,并且通过这样做,避免了量化频谱中的频谱缺陷(hole)。这对于低比特率来说尤其重要,在低比特率时,很多 MDCT 线都会被另外地量化成 0 值,其会在重构信号的频谱中导致听得见的缺陷。

[0151] 图 21 示意性地显示根据本发明一个实施例的基于模型的 MDCT 线量化器(MBMLQ)。图 21 的上部分绘示 MBMLQ 编码器 2100。MBMLQ 编码器 2100 获取 MDCT 帧中的 MDCT 线或 LTP 残余的 MDCT 线(如果该系统中存在 LTP)作为输入。MBMLQ 使用 MDCT 线的统计模型,并且源码以逐 MDCT 帧为基础的方式被自适应为信号性质,产生到比特流的有效压缩。

[0152] 可以将 MDCT 线的局部增益估计为 MDCT 线的 RMS 值,且可以在输入至 MBMLQ 编码器 2100 之前在增益归一化模块 2120 中归一化 MDCT 线。局部增益对 MDCT 线进行归一化,并且局部增益是对 LP 增益归一化的补充。而 LP 增益在较大的时间标度内自适应于信号水平的变化,局部增益在较小的时间标度内自适应于信号水平的变化,会改善瞬态声的品质和语音中的开始(on-sets)。局部增益通过固定速率或可变速率编码进行编码并传输至解码器。

[0153] 可使用速率控制模块 2110 来控制编码 MDCT 帧所用的比特数。速率控制指数控制所使用的比特数。速率控制指数指向标称量化器步长尺寸的列表。该表格可以按照步长尺寸以降序来分类。

[0154] MBMLQ 编码器以一组不同的速率控制指数运行,并且针对所述帧,使用产生的比特计数低于比特储存器控制所给出的许可比特数的速率控制指数。速率控制指数缓慢地改变,并且这可用来降低搜索的复杂性并用来有效地编码指数。如果是在先前 MDCT 帧的指数的附近开始测试,则可减少被测的指数组。同样地,如果概率在该指数(例如,由 32 个步长尺寸 of 列表)的先前值的附近到达峰值,则可获得指数的有效熵编码,并且平均每 MDCT 帧使用 2 个比特来编码速率控制指数。

[0155] 图 21 进一步示意地显示 MBMLQ 解码器 2150,其中如果是在编码器 2100 中估计局部增益,则对 MDCT 帧进行增益再归一化。

[0156] 图 21a 更加详细地示意显示基于模型的熵约束编码器 2140。通过输入的 MDCT 线除以感知掩蔽曲线(优选地由 LPC 多项式导出)的值,对输入 MDCT 线进行感知加权,导致加权的 MDCT 线矢量 $y = (y_1, \dots, y_N)$ 。后续编码的目的是向感知域内的 MDCT 线引入白量化噪声。在解码器中,应用感知加权的逆,导致遵循感知掩蔽曲线的量化噪声。

[0157] 现在在量化器的背景下,将随机偏移论述为避免因粗量化而引起的频谱缺陷的方法。如图 22 中所概述的,避免频谱缺陷的另一方法是在 LTP 环路中并入 SBR 模块 2212。

[0158] 在图 22 中,SBR 模块 2212 在 MDCT 域中操作,并由较低频再生成高频。与完整的编码器/解码器的 SBR 系统相反,LTP 环路中的 SBR 模块不需要任何包络调节,这是因为整个操作都在频谱平坦的 MDCT 域中实施。在 LTP 环路中放置高频重构模块的优势是:在量化之前减去由高频再生成的信号且在量化之后再将其加上。因此,如果具备用于编码整个频率范围的比特,则量化器将编码该信号以便保留原始高频(因为 SBR 贡献在量化之前减去且在量化之后加上),且如果比特约束太过于严苛,则量化器将不能产生高频中的能量,且在输出时加上 SBR 再生成的高频作为“后退”,从而确保高频范围内的能量。

[0159] 在本发明一个实施例中,LTP 环路中的 SBR 模块是简单的复制(copy up)机制(即,将低频线复制到高频线)。在另一实施例中,使用了谐波高频再生成模块。应注意的是,对于谐波信号,优选的是产生与低带频谱谐波相关的高频频谱的 SBR 模块,因为量化之前从输入信号减去的高频可能与原始高频很一致并且因此降低了进入量化器的信号的能量,从而使得在给定的某一比特率要求的情况下的量化更容易进行。在第三实施例中,LTP 环路中的 SBR 模块可取决于变换尺寸以及隐性地取决于信号的特性,自适应其再形成高频的方式。

[0160] 本发明还包含新窗体序列编码格式。根据本发明一个实施例,如图 23a、b、c 所示,用于 MDCT 变换的窗体为二元尺寸,且各窗体的尺寸仅可以以因子 2 来变化。二元变换尺寸在 16kHz 采样率下为(例如)对应于 4、8...、128ms 的 64、128...、2048 样本。一般,我们提出了可变尺寸的窗体,其可呈现在最小窗体尺寸与最大尺寸之间的多个窗体尺寸。在一个序列中,接连的窗体尺寸仅可以以因子 2 来变化,使得窗体尺寸的序列平滑而不会发生突变变化。该实施例所限定的窗体序列(即,限于二元尺寸并且各窗体的尺寸仅允许以因子 2 来变化)具有数种优势。首先,不需要特定的开始窗体或停止窗体,即,不需要锐缘(锋利的边)窗体。这保持良好的时间解析度/频率解析度。第二,窗体序列编码变得很有效率,即,向解码器发送关于使用何种特定窗体序列的信号。根据一个实施例,发送关于序列中的下一窗体尺寸增大因子 2 还是减小因子 2 的信号仅需要一个比特。当然,在上述约束下,也可能存在其它的有效编码整个窗体尺寸序列的编码模式。最后,窗体序列将始终良好

地适合于超帧结构。

[0161] 当在实界系统中操作编码器时,超帧结构是有利的,其中需要传输某些解码器配置参数,以便能够起动解码器。该数据通常存储在比特流中的描述编码音频信号的标头字段中。为使比特率最小化,并非为每个编码数据帧都传输标头,尤其在本发明提出的系统中,其中 MDCT 帧尺寸可从非常小变化到非常大。因此,本发明提出将某一数量的 MDCT 帧一起分组到一个超帧内,其中是在超帧的开头传输标头数据。超帧通常在时间上限定为特定长度。因此,需注意的是要使得 MDCT 帧尺寸的变化适合于恒定长度、预限定的超帧长度。上述的本发明窗体序列确保了所选的窗体序列始终适合于超帧结构。

[0162] 图 23a 显示由 MDCT 理论给出的对于 MDCT 变换的相邻窗体的优选兼容性要求。左边的窗体容纳变换尺寸 L_1 ,且右边的窗体容纳变换尺寸 L_2 。窗体间的重合支撑于直径的时间间隔上或时间段(期间)D 上。对于本发明一个实施例教导的 MDCT 变换,变换尺寸或可相等 $L_1 = L_2$,或变换尺寸相差因子 2, $L_1 = 2L_2$ 或 $L_2 = 2L_1$ 。图示绘示的是后种情形。此外,作为另一优选的约束,变换尺寸间隔的位置必须通过规则的等距超帧序列的二元分割来获得。即,变换间隔位置必须由从一个超帧间隔开始的一系列分离间隔一半来产生。即使当给出变换尺寸间隔时,仍留有某些选择重合直径 D 的自由度。根据本发明一个实施例,避免了直径 D 与邻近变换尺寸 L_1 、 L_2 相比小得非常多的情况,因为锐缘会导致所获得的 MDCT 变换的频率解析度很差。

[0163] 图 23b 示意地显示本发明的使用 4 个不同 MDCT 窗体形状的一个实施例。这 4 个形状表示为如下:

[0164] LL:左长-右长重合;

[0165] LS:左长-右短重合;

[0166] SL:左短-右长重合;

[0167] SS:左短-右短重合。

[0168] 所使用的 MDCT 窗体是这 4 个窗体类型的再缩放版本,其中是以等于 2 的幂为因子再缩放。图 23b 中时间轴上的刻度标记表示变换尺寸间隔,可以看出,长重合的直径等于变换尺寸,而短重合的直径为变换尺寸的一半。在实际的实施方案中,最大变换尺寸为最小变换尺寸的 2^N 倍,其中 N 通常等于小于 6 的整数。此外,对于最小的变换尺寸,仅可考虑 LL 窗体。

[0169] 图 23c 以实例方式描述根据本发明一个实施例的窗体序列编码方法。时间轴的标度被归一化到最小变换尺寸的单元。超帧尺寸为 $H = 16$ 个所述单元,并且超帧的左边缘限定时间标度的起始 $t = 0$ 。此外,为简单起见,假设所允许的最大变换尺寸为 $4 = 2^N$,其中 $N = 2$ 。变换尺寸间隔形成超帧间隔 $[0, 16]$ 的二元部分,间隔由如下 7 个间隔组成: $[0, 4]$ 、 $[4, 6]$ 、 $[6, 8]$ 、 $[8, 9]$ 、 $[9, 10]$ 、 $[10, 12]$ 、 $[12, 16]$,它们分别具有长度 4、2、2、1、1、2、4。可以看出,这些长度遵循两个邻近之间的尺寸最多以因子 2 变化的条件。所有 7 个窗体都是再缩放图 23b 中 4 个基本形状中的一个而获得。

[0170] 由于变换尺寸保持不变、增倍或减半,所以对那些变换尺寸进行递归编码的第一种方法是通过第三符号沿着窗体序列追踪这种选择。然而,这将导致变换尺寸的过度编码以及窗体形状的含糊描述。有时前者由于使用二元分割的要求而无法增倍变换尺寸。例如,在间隔 $[4, 6]$ 之后,在间隔 $[6, 10]$ (其并不是 $[0, 16]$ 的二元子间隔)中会造成增倍。后

者一窗体形状的含糊描述在图 23b 的实例中仍然存在,这是因为相等尺寸的相邻间隔可共享长重合或共享短重合。这些重合要求可从 MDCT 理论中得知,且这些重合要求使得能够去假频 (alias) 滤波器组的消除性质。

[0171] 另外,根据一个实施例的编码原则如下:对于每个窗体,最大化 2 个比特被如下定义:

[0172] 如果变换尺寸大于左重合, $b_1 = 1$, 否则 $b_1 = 0$;

[0173] 如果右重合小于变换尺寸, $b_2 = 1$, 否则 $b_2 = 0$

[0174] 换个说法,比特矢量 (b_1, b_2) 与图 23b 的窗体类型的映射如下给出:

	b_2	0	1
b_1			
0		LL	LS
1		SL	SS

[0175]

[0176] 然而,如果可从二元变换间隔的约束或从变换尺寸的限制来推出的一个比特,则不传输这个比特。

[0177] 返回图 23c 的特定实例,4 个单元的最左边的重合尺寸是通过先前超帧的最终状态或通过独立超帧情况下的绝对传输获得的当前超帧的初始状态。要考虑的第一个比特是最左边窗体的 b_1 。由于间隔 $[0, 4]$ 的长度不大于 4,所以该比特的值是 0。然而,由于 4 是该实例考虑的最大变换尺寸,所以略去第一比特。这在第一窗体上方绘示为划叉的 0。由于右边重合小于变换尺寸,所以该窗体的第二比特 $b_2 = 1$,如上所述,重合点 $t = 4$ 。接下来,间隔 $[4, 6]$ 在 $t = 4$ 附近具有等于重合的尺寸,因此第二窗体的第一比特是 $b_1 = 0$ 。 $t = 6$ 附近的重合不小于 2,因此下一个比特是 0。第三窗体的变换尺寸比特 b_1 具有 0 值,但在这里较长变换的选择与二元结构不相一致,因此可由所述情形推出这个比特,因此不传输该比特并在图中划叉。该过程继续,直至在 $t = 16$ 处到达超帧的末端,其中短重合的比特是 1。在中间处,由于没有使用最短变换尺寸的重合并由于放大的错误位置,所以间隔上方 $[9, 10]$ 的三个比特被划叉。因此,全部未划叉的比特序列是 01000100001011,但在使用编码器和解码器处获得的信息之后,其简化为用于编码 7 个窗体的 9 个比特 100101011。

[0178] 对于本领域的技术人员很明显,通过对这些纯描述性比特进行熵编码,可实现比特率的进一步降低。

[0179] 图 24 中提出本发明编码器/解码器系统的附加特征。输入信号输入至 MDCT 分析模块,信号的 MDCT 表达式输入到谐波预测模块 2400 中。谐波预测是在给定参数滤波器的情况下沿着频率轴的滤波。在给定音高信息、增益信息以及相位信息的情况下,如果输入信号包含谐波系,则接着可由较低的线预测出较高(频率上)的 MDCT 线。谐波预测模块的控制参数是音高信息、增益和相位信息。

[0180] 根据一个实施例,如图 25 中所概述,使用了 MDCT 域中的虚拟 LTP 矢量。图 25 绘示了包含的两个模块:LTP 提取模块 2512 及 LTP 精化(refinement)模块 2518。LTP 的概念是将输出信号的先前片段用于当前片段或帧的解码。在给定最小化编码信号失真的迭代的情况下,使用哪个先前片段是由 LTP 提取模块 2512 决定。当在 MDCT 域中实施 LTP 时,本发明提供一种将 MDCT 帧的重合纳入考虑的新方法,即当选择 LTP 滞后时,使得将经过 MDCT

分析并且将在当前输出片段的解码过程中使用的先前输出信号的片段包括（由于重合）尚未产生的先前输出片段的部分。

[0181] 该迭代过程例示如下：从 LTP 缓冲器开始，通过 LTP 提取模块 2512 实施信号的第一提取。第一提取的结果由精化模块 2518 精化，其目的是当所选的滞后 T 小于待编码帧的 MDCT 窗体的时间段时改善 LTP 信号的质量。首先参照图 25a 简短地概述对于时间滞后小于分析的帧来精化 LTP 贡献的迭代过程。在第一图表中，显示了 LTP 缓冲器中的所选片段，其中 MDCT 分析窗体超叠加。重合窗体的右边部分不包含可用数据：时间信号的虚线部分。迭代过程通过如下步骤：

[0182] 1) 如 MDCT 分析所通常进行的叠入重合部分；

[0183] 2) 展开重合部分（请注意，右边的部分起初不包含任何数据，而现在具有展开数据）；

[0184] 3) 使窗体向右移位所选的 LTP 滞后；

[0185] 4) 叠入重合的部分并计算 Δ ；

[0186] 5) 求 Δ 与上部图表中的原始 LTP 片段的和。

[0187] 该迭代过程优选地进行 2 至 4 次。

[0188] 图 25b 中更加详细地绘示经 MDCT 自适应的 LTP 提取过程，其显示了由 LTP 提取模块执行的步骤：

[0189] a) 对格式化的输入信号 $x(t)$ 进行描述；仅在有限的时间间隔（该时间间隔或为 LTP 缓冲器的范围、或为当前 MDCT 帧窗体的范围或由系统约束给出的某一其它间隔）内，输入信号是已知的。然而，对于操作的定义，假设输入信号是始终已知的。这是通过在间隔以外（信号是已知的）将信号设为 0 来实现的。

[0190] b) 对输入信号执行的第一操作是使输入信号移位 LTP 滞后 T ；即，

[0191] $x_1(t) = x(t-T)$ 。

[0192] c) 下一个步骤是应用 MDCT 窗体 $w(t)$ 。该窗体由时间段 $2r_1$ 的上升部分和时间段 $2r_2$ 的下降部分以及可能的二者间的恒定部分组成。示例性窗体由虚线图表绘示。窗体的上升部分和下降部分的支撑分别集中在镜像点 t_1 和 t_2 的附近。信号 $x_1(t)$ 逐个点地与窗体相乘得到：

[0193] $x_2(t) = w(t) \cdot x_1(t)$

[0194] 再次假设在已知范围 $[t_1-r_1, t_2+r_2]$ 以外，窗体 $w(t)$ 为 0。

[0195] 关于从 $x(t)$ 到 $x_2(t)$ 的操作的另一等效观点是执行如下步骤：

[0196] **(i)** $\tilde{x}_2(t) = w(t+T) \cdot x(t)$

[0197] **(ii)** $x_2(t) = \tilde{x}_2(t-T)$ ；

[0198] 其中步骤 (i) 相当于与支撑于 (t_1-r_1-T, t_2+r_2-T) 上窗体的窗体化，且步骤 (ii) 是使结果移位 LTP 滞后 T 。

[0199] d) 将现在窗体化后的信号 $x_2(t)$ 叠入支撑于 $[t_1, t_2]$ 上的信号，定义如下：

[0200]

$$\begin{aligned}
 & \mathbf{x}_2(t) + \varepsilon_1 \mathbf{x}_2(2t_1 - t), \quad \text{其中 } t_1 \leq t \leq t_1 + r_1; \\
 \mathbf{x}_3(t) = & \mathbf{x}_2(t), \quad \text{其中 } t_1 + r_1 < t < t_2 - r_2; \\
 & \mathbf{x}_2(t) + \varepsilon_2 \mathbf{x}_2(2t_2 - t), \quad \text{其中 } t_2 - r_2 \leq t \leq t_2.
 \end{aligned}$$

[0201] 对于该所示的实例,符号的值是对应于 MDCT 变换的既定实施方案的 $[\varepsilon_1, \varepsilon_2] = (-1, 1)$, 其它可能性为 $(1, -1)$ 、 $(1, 1)$ 或 $(-1, -1)$ 。

[0202] e) 随后,将叠入的信号 $x_3(t)$ 展开成如下给出的支撑于间隔 $[t_1 - r_1, t_2 + r_2]$ 上的信号:

[0203]

$$\begin{aligned}
 & \varepsilon_1 \mathbf{x}_3(2t_1 - t), \quad \text{其中 } t_1 - r_1 \leq t \leq t_1; \\
 \mathbf{x}_4(t) = & \mathbf{x}_3(t), \quad \text{其中 } t_1 < t < t_2; \\
 & \varepsilon_2 \mathbf{x}_3(2t_2 - t), \quad \text{其中 } t_2 \leq t \leq t_2 + r_2.
 \end{aligned}$$

[0204] 也可将从 $x_2(t)$ 到 $x_4(t)$ 的操作组合成一种操作:加上或减去间隔 $[t_1 - r_1, t_1 + r_1]$ 和 $[t_2 - r_2, t_2 + r_2]$ 上的信号部分的镜像图像。

[0205] f) 最后,将信号 $x_4(t)$ 与 MDCT 窗体进行窗体化,以产生 LTP 提取操作的结果:

[0206] $y(t) = w(t) \cdot x_4(t)$

[0207] 对于本技术领域的技术人员很明显,从 $x_1(t)$ 到 $y(t)$ 的组合操作等效于跟随有 MDCT 合成的 MDCT 分析,且这实现了当前 MDCT 帧子空间的正交投影。

[0208] 重要的是要注意,在没有重合的情况下,即 $r_1 = r_2 = 0$,由于 d) 至 f) 中操作的缘故,所以 $x_2(t)$ 什么也不会发生。那么,窗体化就由简单地提取间隔 $[t_1, t_2]$ 内的信号 $x_1(t)$ 组成。在该情况下,LTP 提取模块 2512 执行与现有技术的 LTP 提取器所完全的。

[0209] 图 25c 显示初始的 LTP 提取的信号 $y_1(t)$ 的迭代精化。其由如下步骤组成:应用 LTP 提取操作达 $N-1$ 次;及将结果添加到初始信号。如果 S 表示 LTP 提取操作,则迭代由如下公式定义:

[0210] $\Delta_0 = y_1;$

[0211] $\Delta_k = S(\Delta_{k-1}), k = 1, \dots, N-1;$

[0212] $y_k = y_{k-1} + \Delta_{k-1}, k = 2, \dots, N-1$

[0213] 如果 LTP 滞后 $T > \max(2r_1, 2r_2)$,则从图 25b 可看出,有使得 $\Delta_N = 0$ 的 N 。如果 $T > (r_1 + r_2 + t_2 - t_1)$,则已经有 $\Delta_1 = 0$,且可略去精化(细化)。在实际中, N 的适合选择是介于 2 至 4 的范围。

[0214] 在没有重合的情况下,即 $r_1 = r_2 = 0$,所述方法与现有技术方法中的虚拟矢量产生相一致。

[0215] 图 25d 显示精化单元的可选实施方案,该精化单元实施如下的迭代:

[0216] $y_k = y_1 + S(y_{k-1}), k = 2 \dots N$

[0217] 在两种实施方案中,来自迭代的最终输出都可写为:

[0218]
$$y_k = \sum_{k=0}^{N-1} S^k y_1 = \sum_{k=1}^N S^k x$$

[0219] 其中 x 是 LTP 缓冲器信号。

[0220] 根据本发明的一个实施例,以可变速率的方式编码 LTP 滞后和 LTP 增益。这是有利的,因为:由于对于静止周期信号的 LTP 效率,所以 LTP 滞后在有些长的片段上往往是相同的。因此,通过算术编码可对此加以利用,导致可变速率的 LTP 滞后和 LTP 增益编码。

[0221] 类似地,本发明的一个实施例也利用了比特储存器和可变速率编码来进行 LP 参数的编码。另外,本发明还教示了递归 LP 编码。

[0222] 如前文概述的,可使用设计用于改善谐波信号编码的技术。这些技术(例如)为谐波预测、LTP 和时间扭曲。所有上述的工具都隐性地或显性地依赖于某种音高或音高相关的信息。在本发明一个实施例中,在给定存在独立性或相关性的情况下,可有效地编码这些不同技术需要的这种不同信息。这显示于图 26 中,其示意地显示组合单元 2600,其用于组合音高以及音高相关的参数(例如,来自时间扭曲的 LTP 滞后和 Δ 音高),并且产生组合的音高信号。

[0223] 如上概述,根据一个实施例的编码解码器可使用 MDCT 域中的 LTP。为改善 MDCT 域中 LTP 的性能,可引入两个额外的 LTP 缓冲器 2512、2513。如图 27 所示,当 LTP 提取器在 LTP 缓冲器 2511 中搜索最佳滞后时,该搜索中也同样包括噪声矢量和脉冲矢量。例如,在存储于 LTP 缓冲器中的先前片段的信号不适合的瞬变中,可将噪声和脉冲用作预测信号。因此,本发明提出了一种带有脉冲和噪声码本条目的增强型 LTP。

[0224] 本发明的另一方面是针对编码器中的可变帧尺寸来处理比特储存器。本发明教示了比特储存器控制单元。除了作为输入提供的难度度量以外,比特储存器控制单元还接收关于当前帧长度的信息。在比特储存器控制单元内使用的难度度量的一个实例是感知熵,或者是能谱的对数。比特储存器控制在其中帧长度可在一组不同帧长度上变化的系统中是重要的。如下文所概述,当计算用于待编码帧的许可比特数时,本发明提出的比特储存器控制单元考虑帧长度。

[0225] 其中比特储存器限定为:缓冲器中的某一固定量的比特必须大于帧所允许用于给定比特率的平均比特数。如果帧为相同的尺寸,则帧的比特数可能没有变化。在按照实际帧所允许的比特数取出许可用于编码算法的比特之前,比特储存器控制始终观察着比特储存器的水平。因此,满比特储存器意味着比特储存器中可用的比特数等于比特储存器尺寸。在帧编码之后,将从缓冲器减去使用过的比特数,并且比特储存器通过添加表示恒定比特率的比特数而得到更新。因此,如果在编码帧之前比特储存器中的比特数等于每个帧的平均比特数,则存储器是空的。

[0226] 图 28a 中绘示比特储存器控制的基本概念。编码器提供用于计算相比于先前帧编码实际帧的难度的手段。对于 1.0 的平均难度,许可的比特数取决于比特储存器中可用的比特数。根据给定的控制线,如果比特储存器很满,则将从比特储存器中取出多于对应于平均比特率的比特。在空比特储存器的情况下,与平均比特相比更少的比特将被用于编码帧。对于平均难度的较长帧序列,这种行为会让步于平均的比特储存器水平。对于较高难度的帧,控制线可向上移位,具有的效果是允许编码有难度的帧在相同比特储存器水平下使用更多的比特。相应地,对于编码容易的帧,只将图 28a 中的控制线从平均难度情况移位到容易的难度情况,用于帧的允许的比特数会更少。也可能存在除了简单地移位控制线的其它修改。例如,如图 28a 中所示,可取决于帧难度改变控制曲线的斜率。

[0227] 当计算许可比特数时,必须遵守对于比特储存器下端的限制,以便不从缓冲器中取出比允许的比特更多的比特。包括计算图 28a 中所示控制线的许可比特的比特储存器控制方案仅仅是可能的比特储存器水平和难度度量与许可比特的关系的一个实例。此外,其它控制算法在比特储存器水平的下端普遍存在硬性限制,其防止比特储存器违反空比特储存器的制约,并且在上端也有限制,其中如果编码器要消耗过少数量的比特,则编码器将被迫写入填充比特。

[0228] 对于能处理一组可变帧尺寸的控制机制,必须对该简单的控制算法进行自适应。必须归一化所使用的难度度量,以使得不同帧尺寸的难度值相兼容。对于每个帧尺寸,会有着许可比特的不同的允许范围,并且由于每个帧的平均比特数对于可变的帧尺寸是不同的,所以每个帧尺寸都具有带有其自身限制的其自身的控制等式。图 28b 中显示一个实例。对于固定帧尺寸情况的一种重要修改是控制算法的较低允许边界。在取出用于实际帧的比特之前,现在最大允许的帧尺寸的平均比特数是比特储存器水平的最低允许值,而不是对应于固定比特率情况的实际帧尺寸的平均比特数。这是固定帧尺寸的比特储存器控制的主要差别中的一个。该制约确保了随后具有最大可能帧尺寸的帧至少能使用该帧尺寸的平均比特数。

[0229] 难度度量可按照 AAC 中的做法,基于(例如)由心理生理学模型的掩蔽阈值导出的感知熵(PE)计算,或者可选地,按照根据本发明一个实施例的编码器的 ECQ 部分中的做法,基于具有固定步长尺寸的量化的比特计数。可关于可变帧的尺寸来归一化这些值,这可通过简单地除以帧长度来实现,并且结果将是分别每个样本一个比特计数的 PE。另一归一化步骤可相对于平均难度发生。出于该目的,可使用过去的帧上的移动平均值,导致对于难度大的帧难度值大于 1.0,或对于难度容易的帧难度值小于 1.0。在两通编码器的情况下或在大超前性的情况下,也可将未来帧的难度值用于考虑难度度量归一化。

[0230] 图 29 概述在本发明提出的编码器和解码器的一个实施例中使用的扭曲 MDCT 域。如图所示,时间扭曲意指对时间标度进行再采样以实现恒定的音高。该图示的 x 轴显示具有变化音高的输入信号,且图示的 y 轴显示再采样后的恒定音高信号。可通过对当前片段使用音高检测算法并且估计片段内的音高演变来确定时间扭曲曲线。然后,使用音高演变信号,对片段中的信号进行再采样,由此生成扭曲曲线。由于确定音高演变仅需要音高差别而不需要绝对的音高信息,所以建立扭曲曲线的算法具有对于音高检测误差是鲁棒性的。

[0231] 根据本发明一个方面,时间扭曲 MDCT 与 LTP 结合使用。在该情况下,LTP 搜索是在编码器中的恒定音高片段域中完成。这对于包括数个音高脉冲(其由于音高变化而在 MDCT 帧中非等距地排列)的长 MDCT 帧尤其有用。因此,来自 LTP 缓冲器的恒定音高片段将不会适当地适合于多个音高脉冲上。根据一个实施例,基于当前 MDCT 帧的扭曲曲线,对 LTP 缓冲器中的所有片段进行再采样。此外在解码器中,在给出扭曲数据信息的情况下,将 LTP 缓冲器中的所选片段再采样到当前帧的扭曲数据。可将扭曲信息作为比特流的一部分传输到解码器。

[0232] 图 29 的上部显示了窗体(即,LTP 缓冲器中的片段)以及当前虚线帧的窗体。图 29a 中可看到扭曲 MDCT 分析的效果。展现在左边的是未扭曲分析的频率曲线。由于音高在窗体上变化,所以频率较高的谐波未得到适当地解析。图示的右边是同一信号的频率曲线,尽管利用时间扭曲 MDCT 分析进行分析。由于音高在分析窗体上恒定,所以较高的谐波得到

了更好的解析。

[0233] 图 30 中显示根据本发明一个实施例的另一分层式 SBR 重构方法。根据图 7, 编码器和解码器可实施为双率系统, 其中以一半的采样率对核心编码器进行采样, 且以原始采样率采样的高频由高频重构模块处理。假设原始采样率为 32kHz, LPC 滤波器以 16kHz 的采样频率进行操作, 提供 8kHz 的白化信号。然而, 在给定施加的比特率约束下, 随后的核心编码器可能不能编码 8kHz 的带宽。本发明提供数种手段来处理这种情况。本发明的一个实施例在 LPC 下 (即, 基于 LPC 滤波信号) 于 MDCT 域中应用高频重构, 以提供 8kHz 的带宽。这概述于图 30 中, 其中 LPC 覆盖了从 0 至 8kHz 的频率范围, 且 0 到 5kHz 的范围由 MDCT 波形量化器来处理。从 5 到 8kHz 的频率范围由 MDCT SBR 算法来处理, 且最后从 8 到 16kHz 的范围由 QMFSBR 算法处理。MDCT SBR 是基于如上所述的基于 QMF 的 SBR 中使用的复制机制。然而, 有利地还可使用其它方法, 例如, 根据变换尺寸自适应 MDCT 的 SBR 方法。

[0234] 在本发明另一实施例中, 取决于帧尺寸和信号性质, 对 LP 频谱的上频率范围进行量化和编码。对于某些帧尺寸和信号, 根据上述来编码频率范围, 而对于其它变换尺寸, 采用稀疏量化和噪声填充技术。

[0235] 尽管已经参照本发明的特定实施例公开了上述内容, 但应理解的是本发明概念并非限于所述的实施例。另一方面, 本申请中提出的公开内容将使得本领域的技术人员能够理解并实施本发明。本领域的技术人员应理解, 可作出各种修改, 而不违背权利要求书排他性提出的本发明的精神和范围。

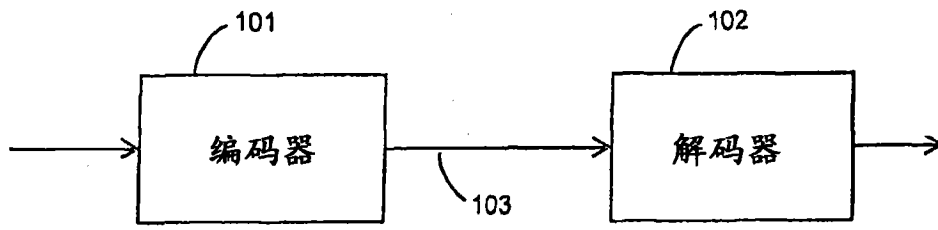


图 1

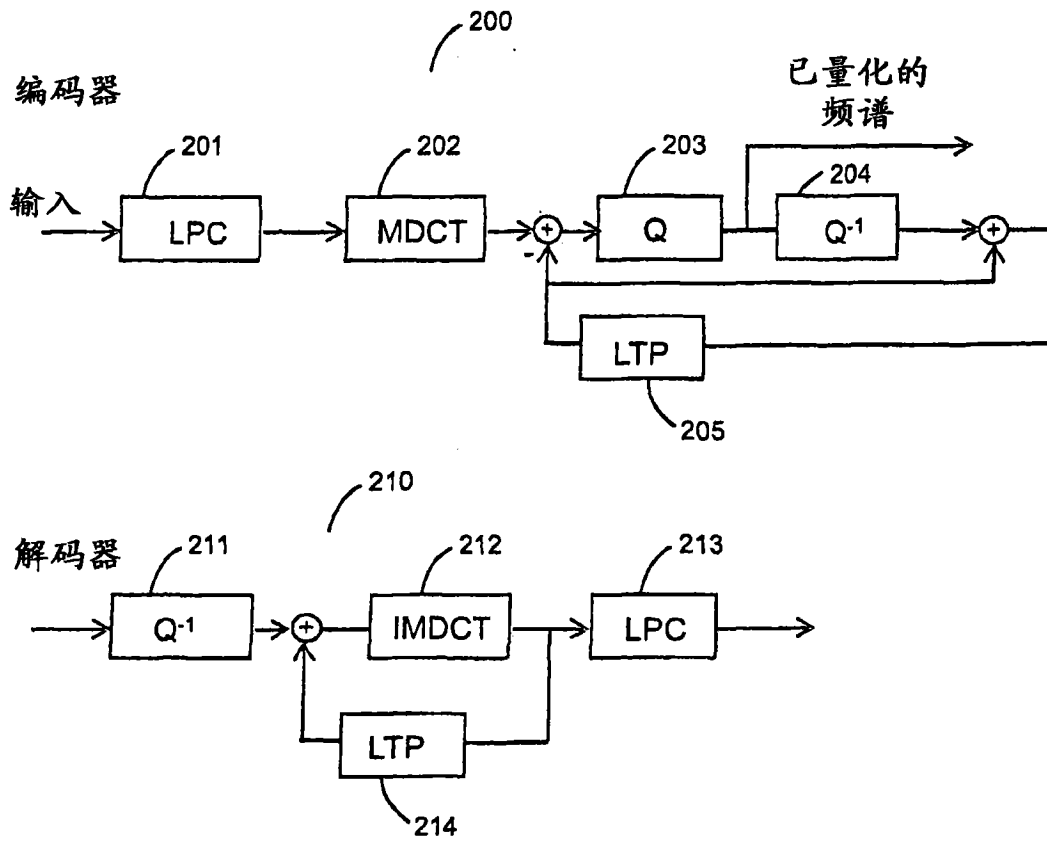


图 2

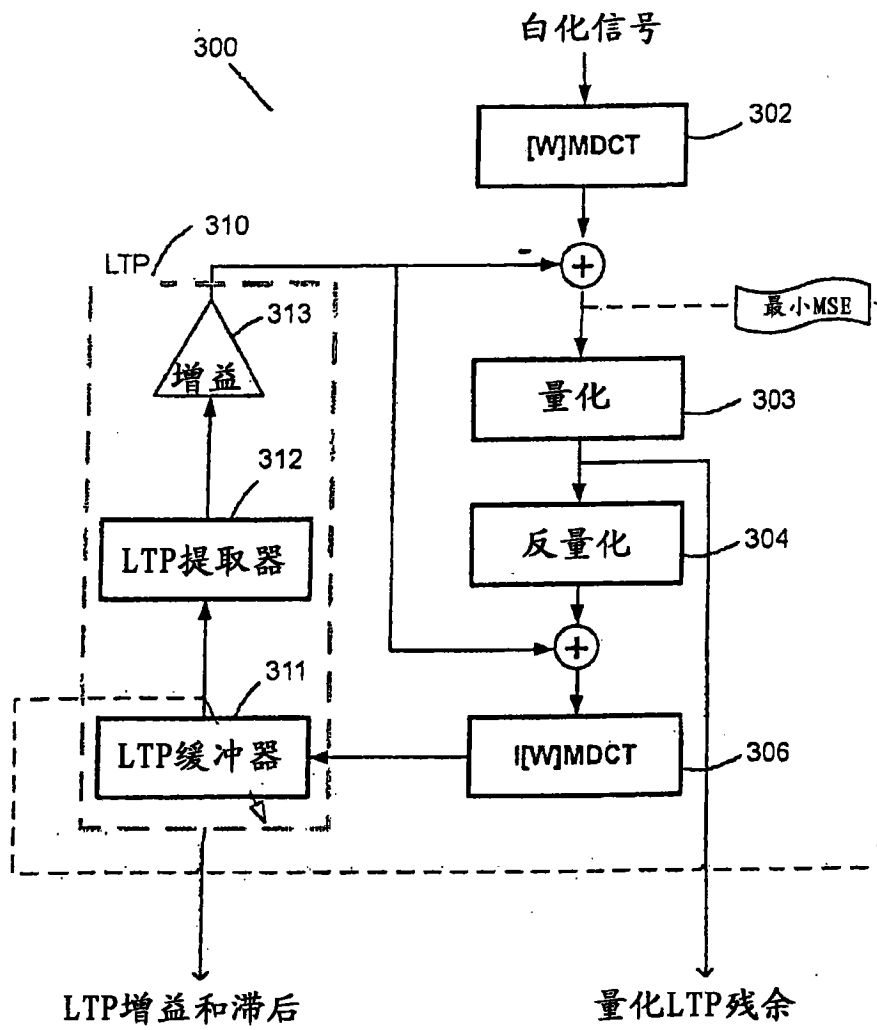


图 3

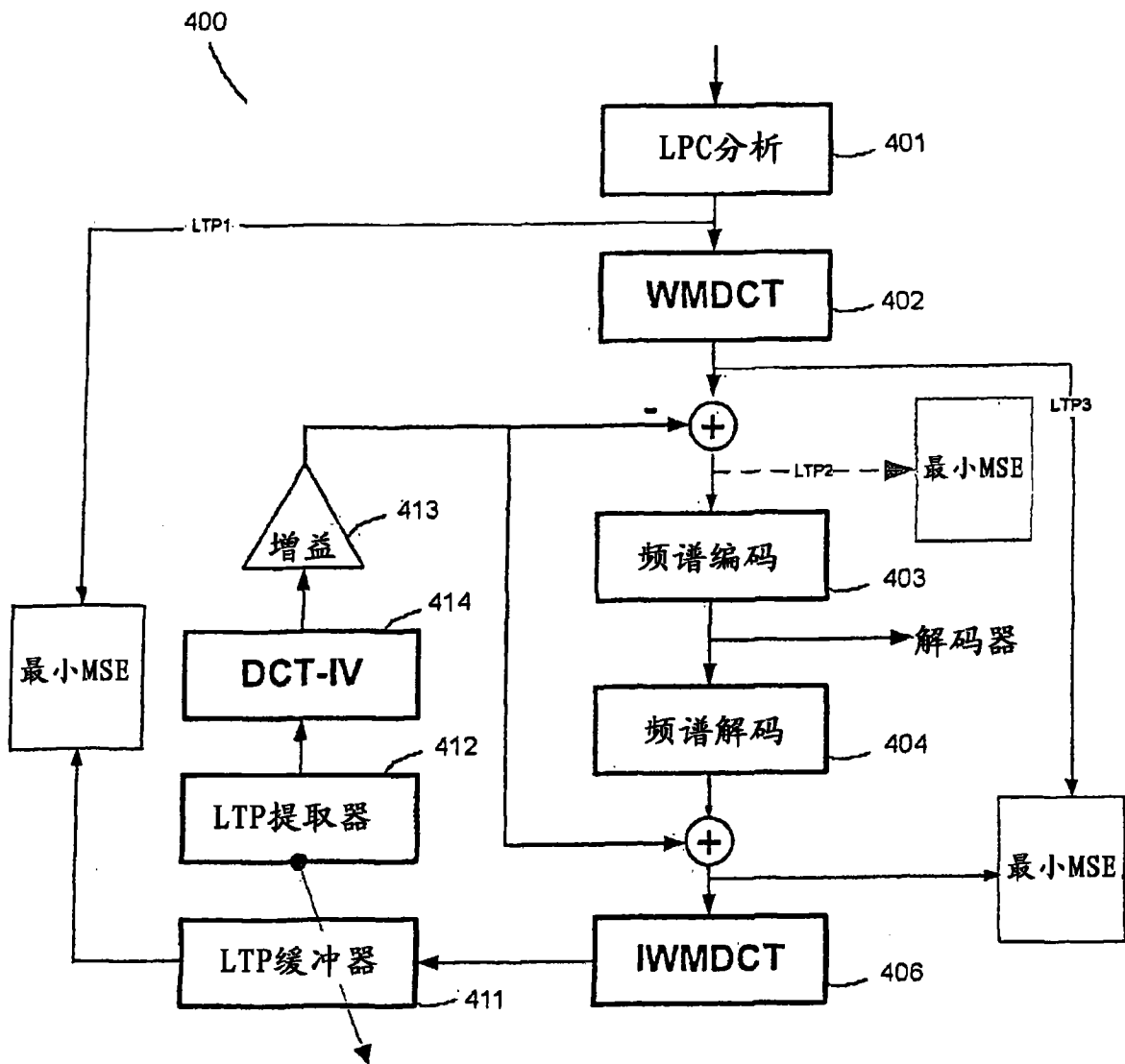


图 4

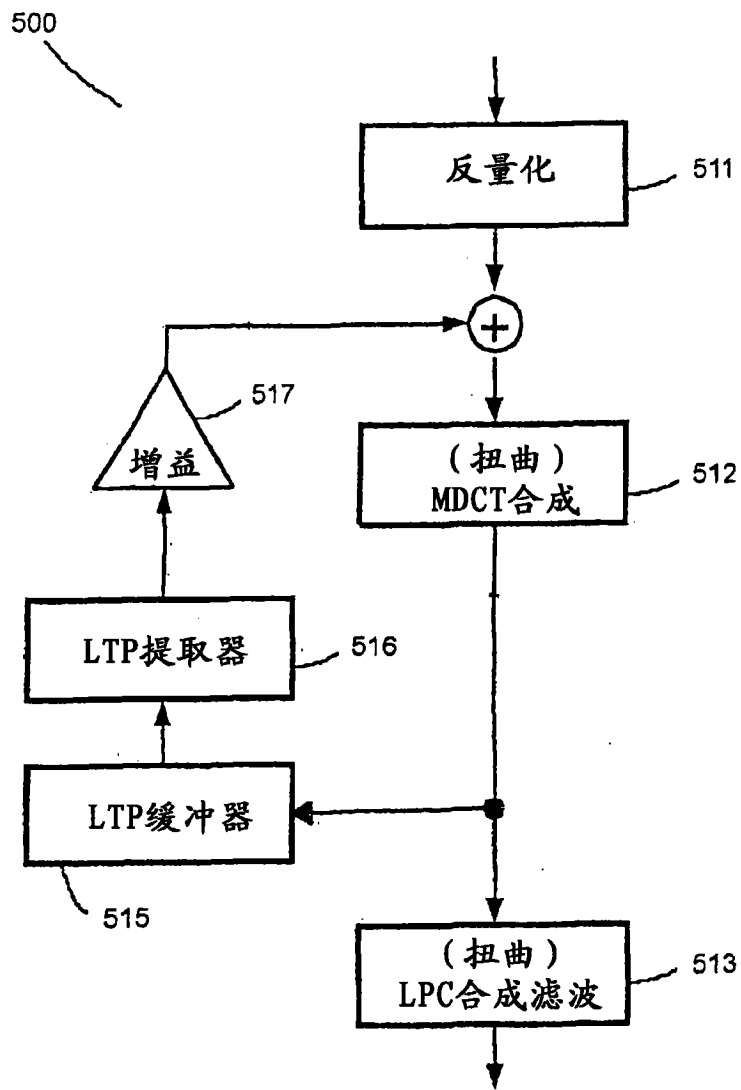


图 5

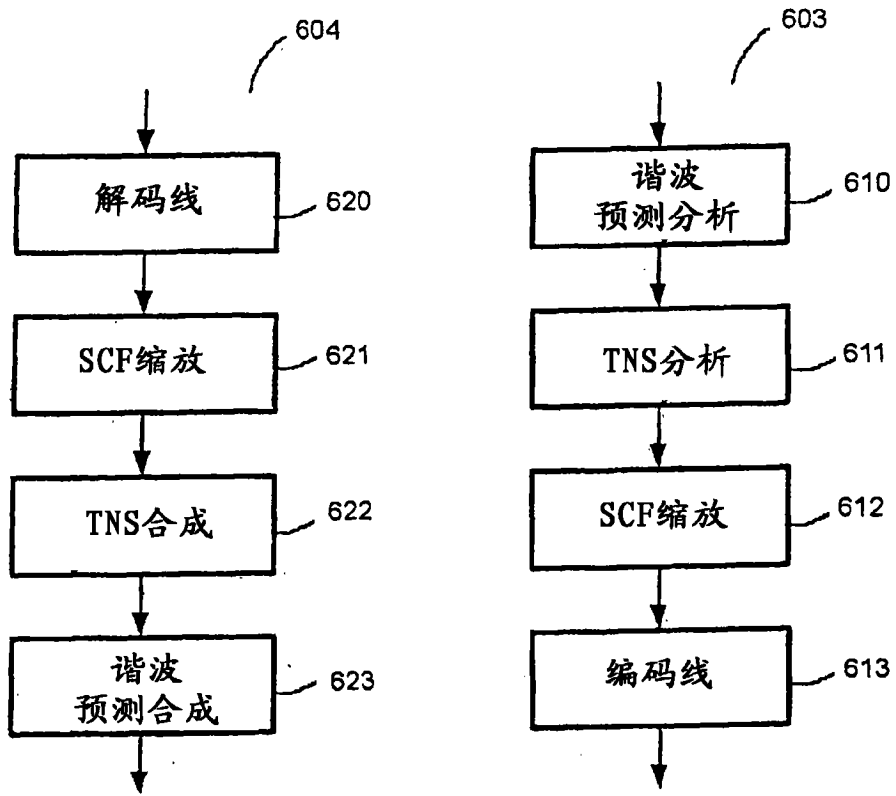


图 6

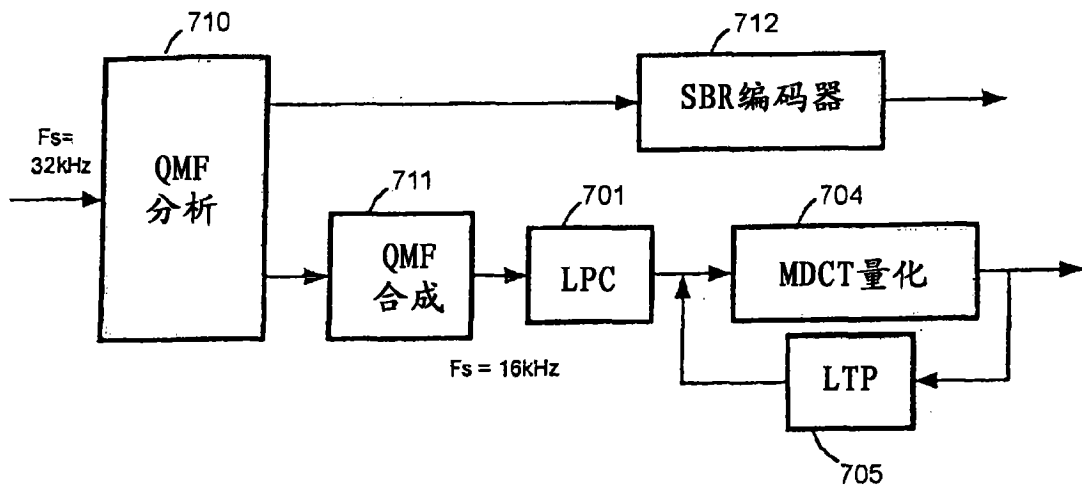


图 7

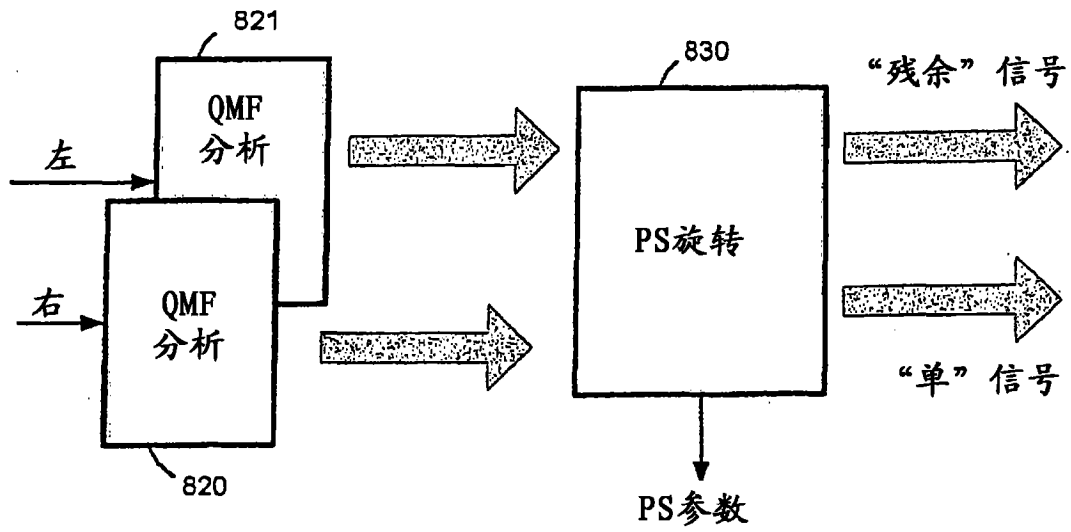


图 8

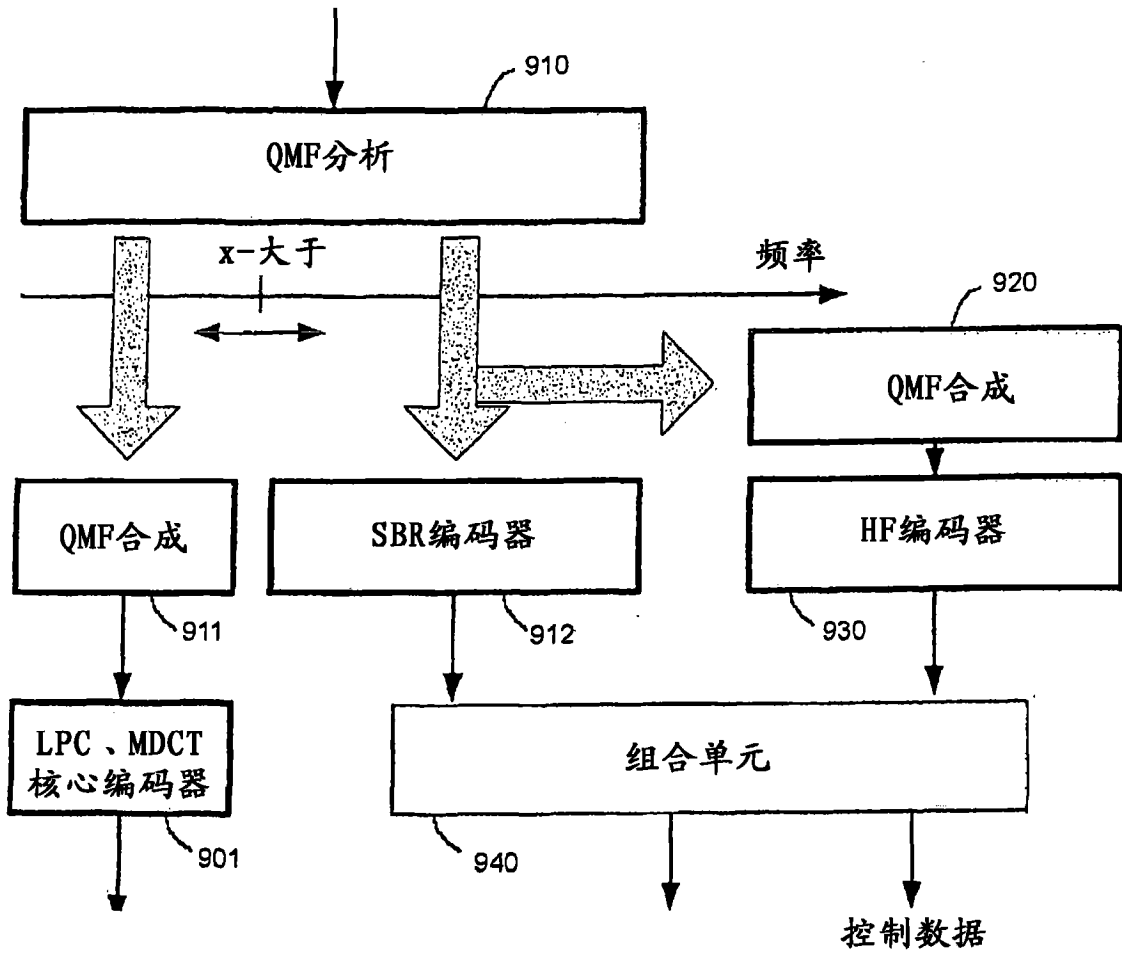


图 9

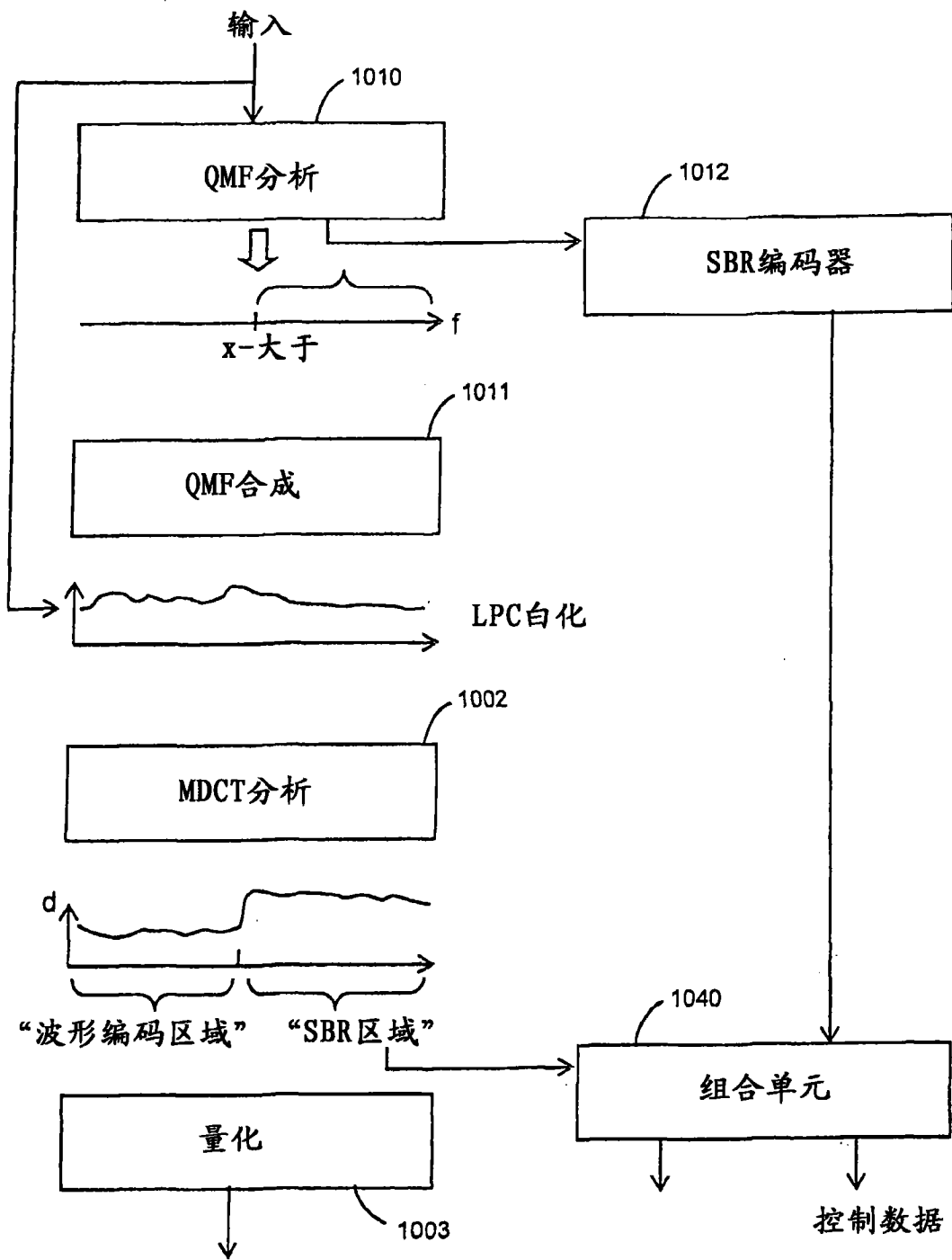


图 10

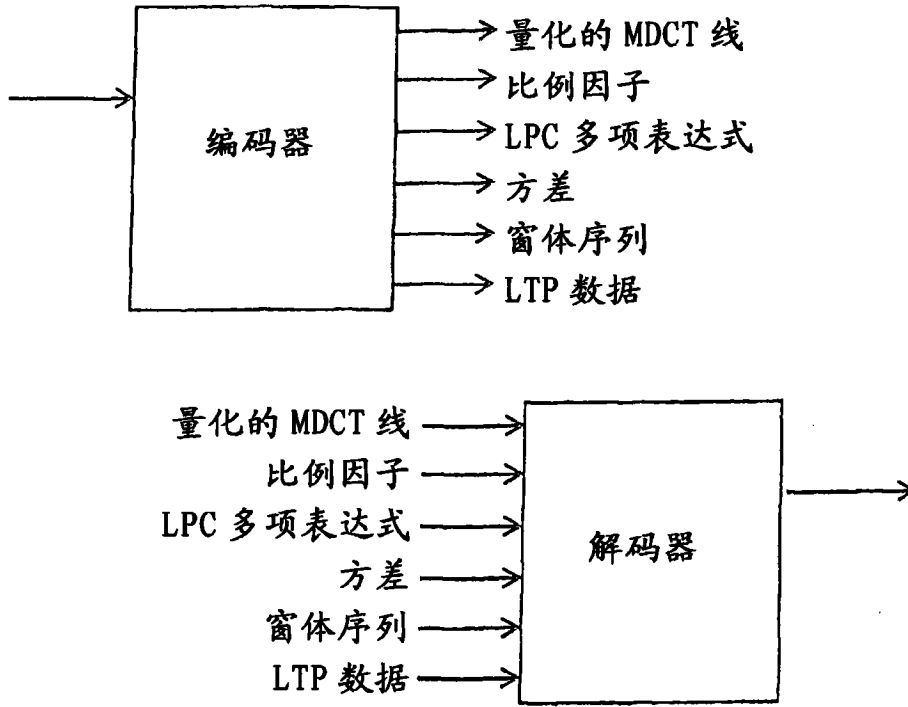


图 11

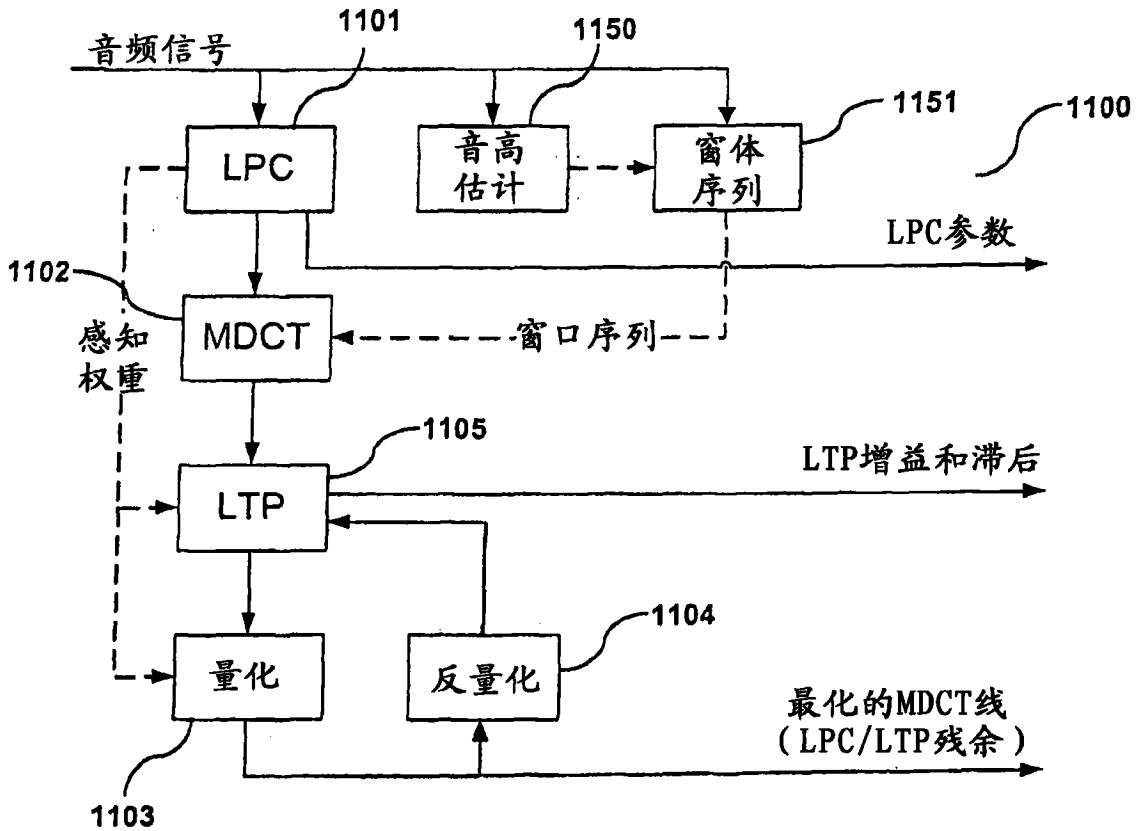


图 11a

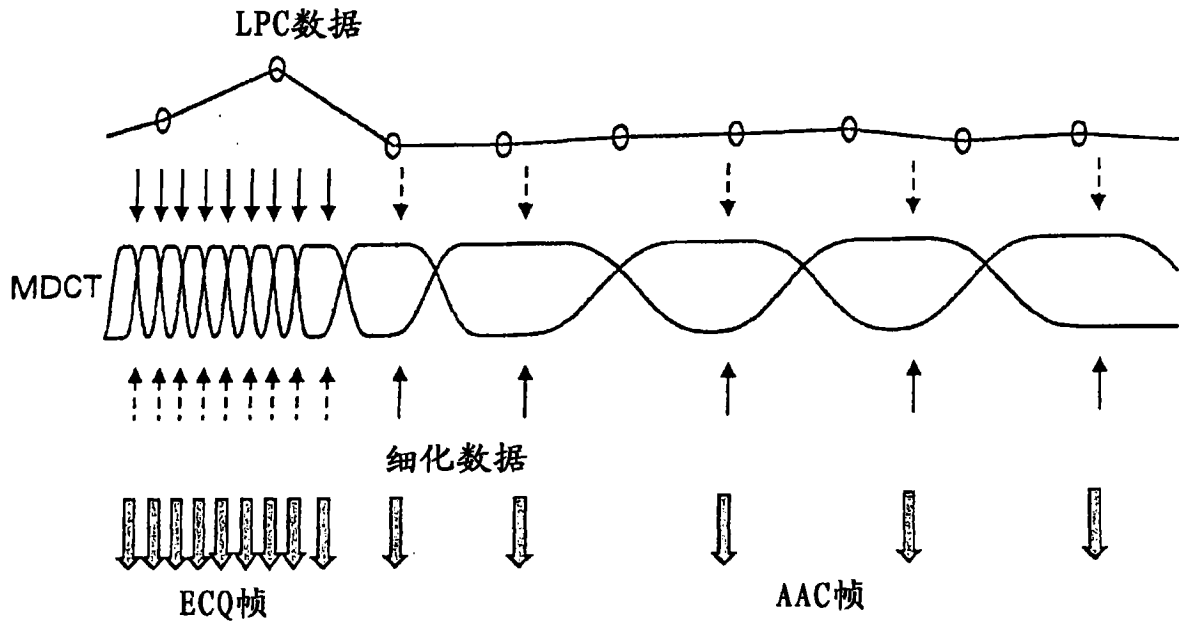


图 12

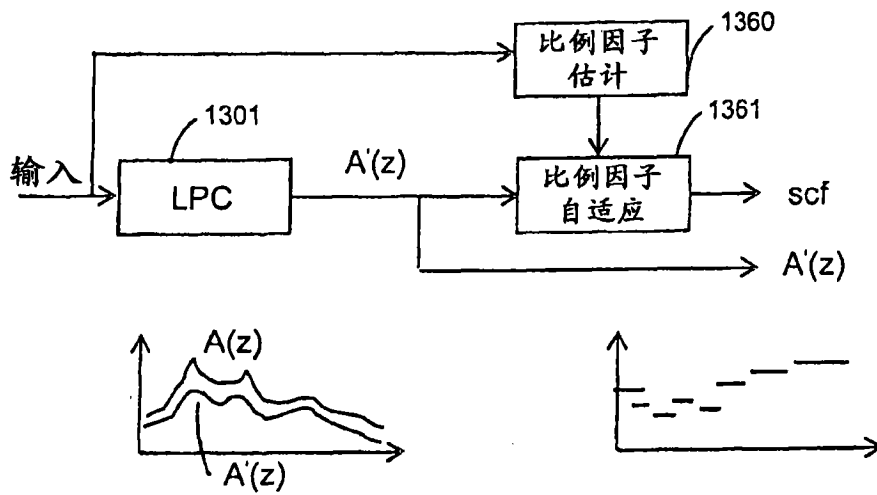


图 13

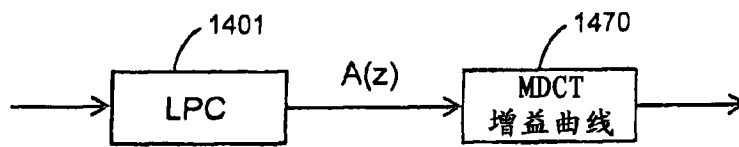


图 14

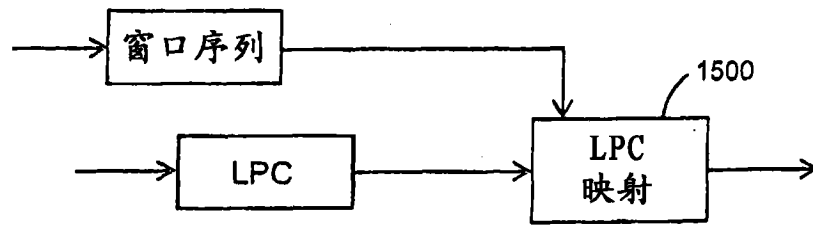


图 15

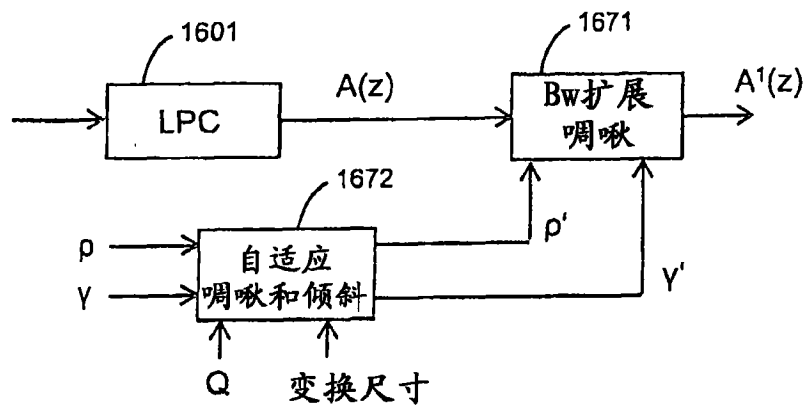


图 16

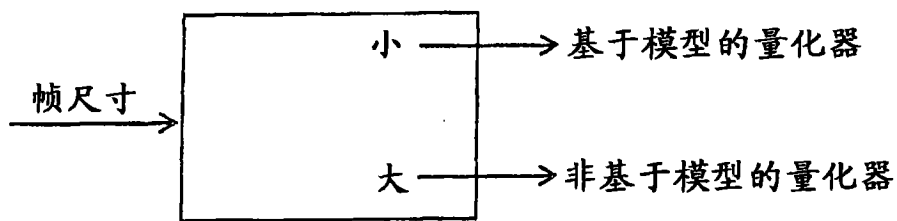


图 17

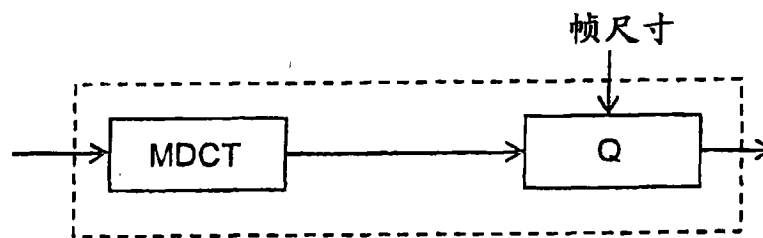


图 18

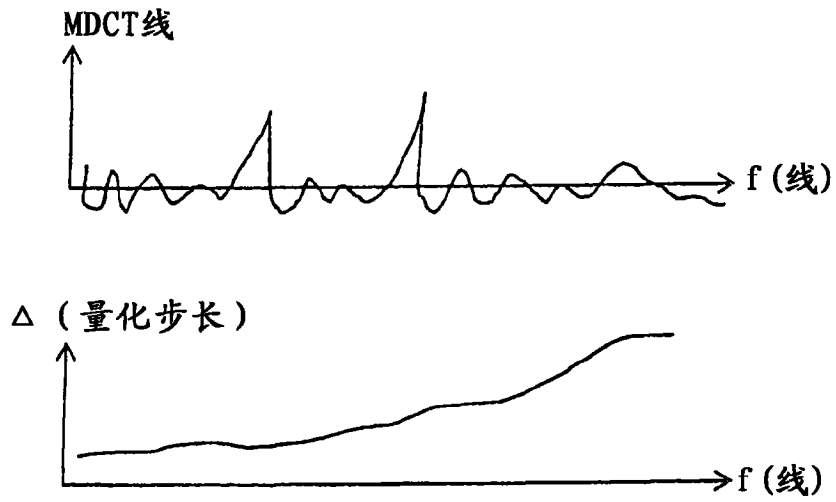


图 19

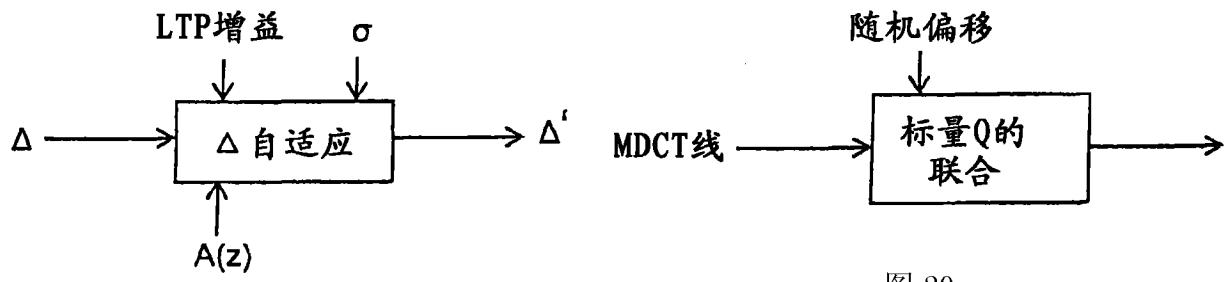


图 20

图 19a

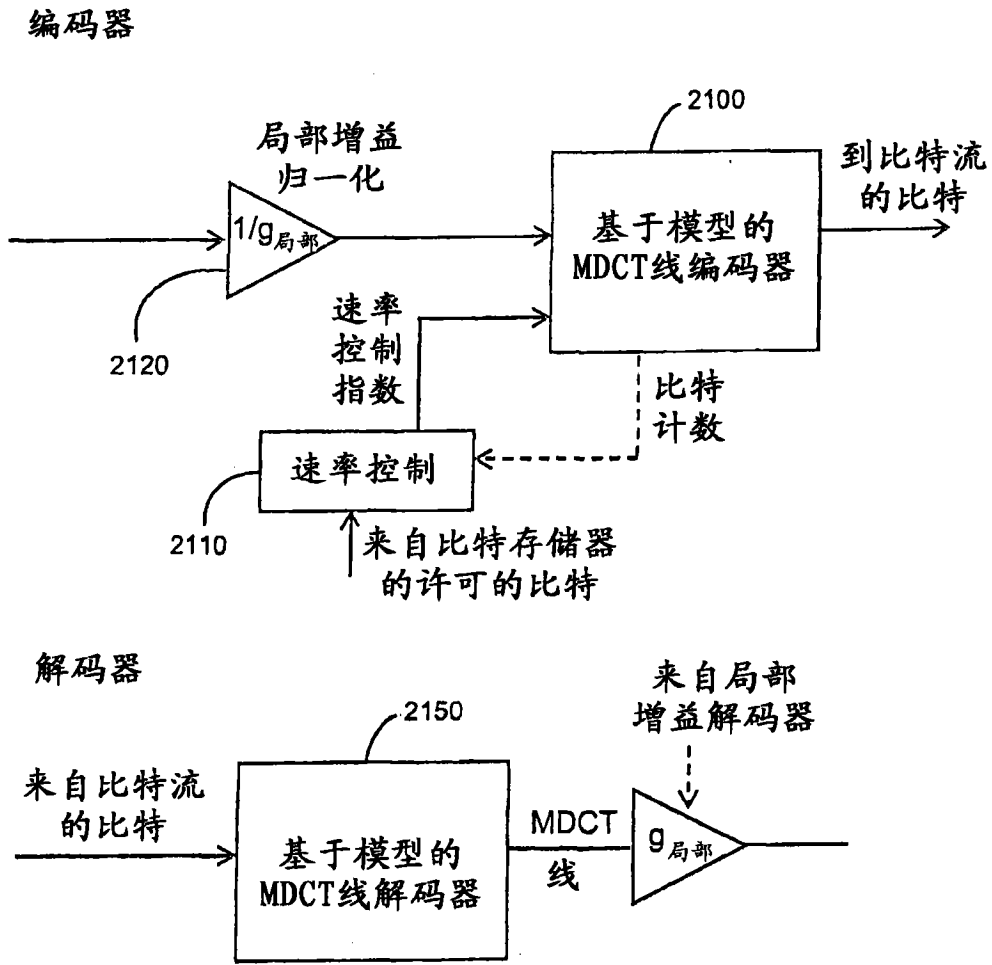


图 21

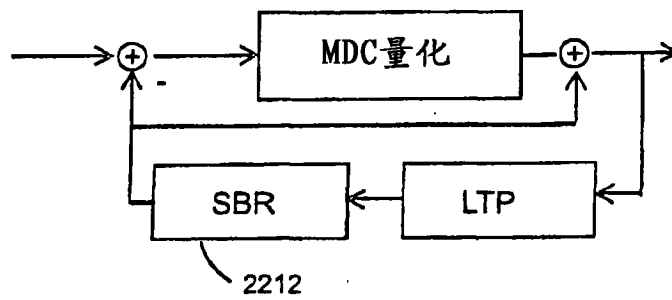


图 22

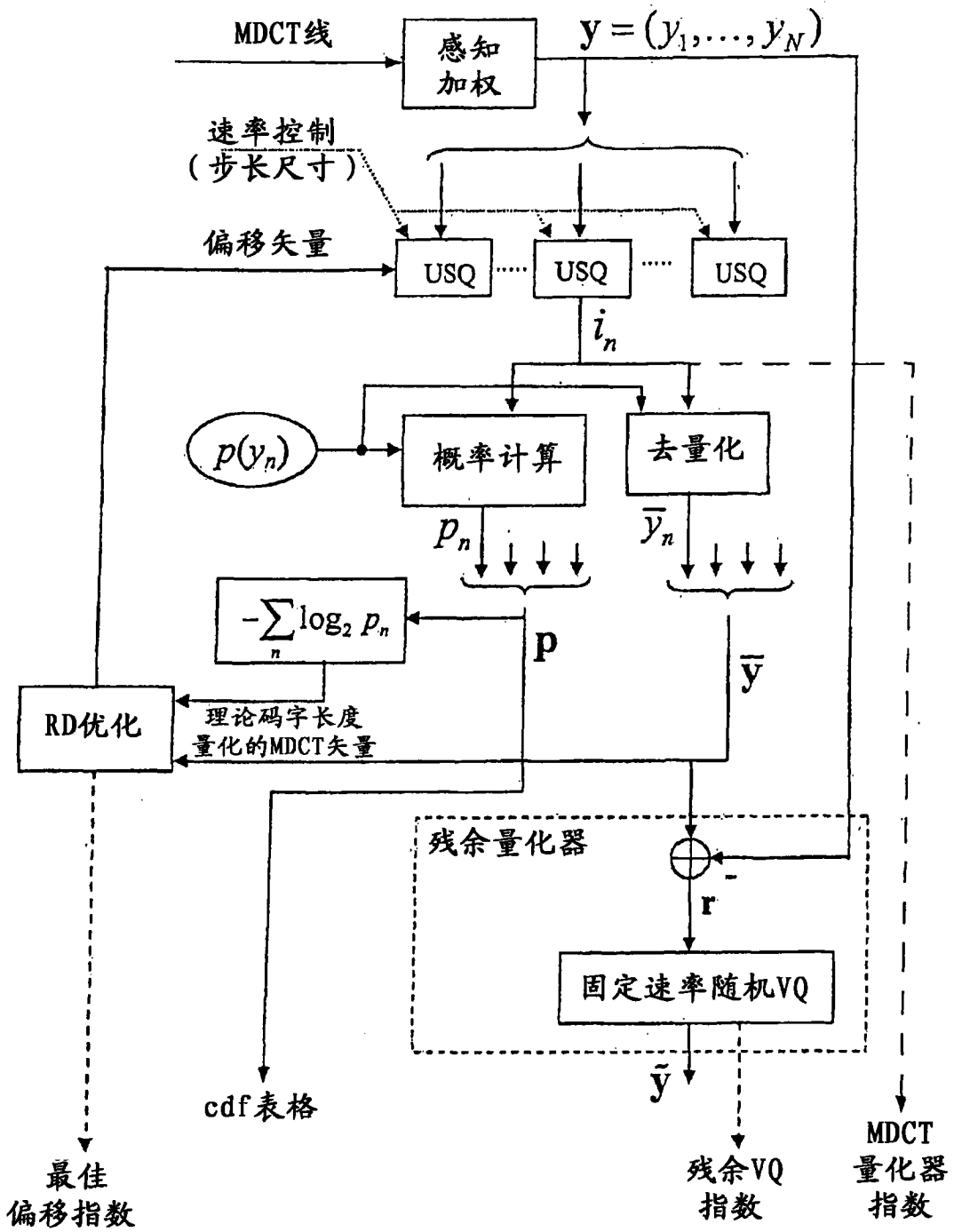


图 21a

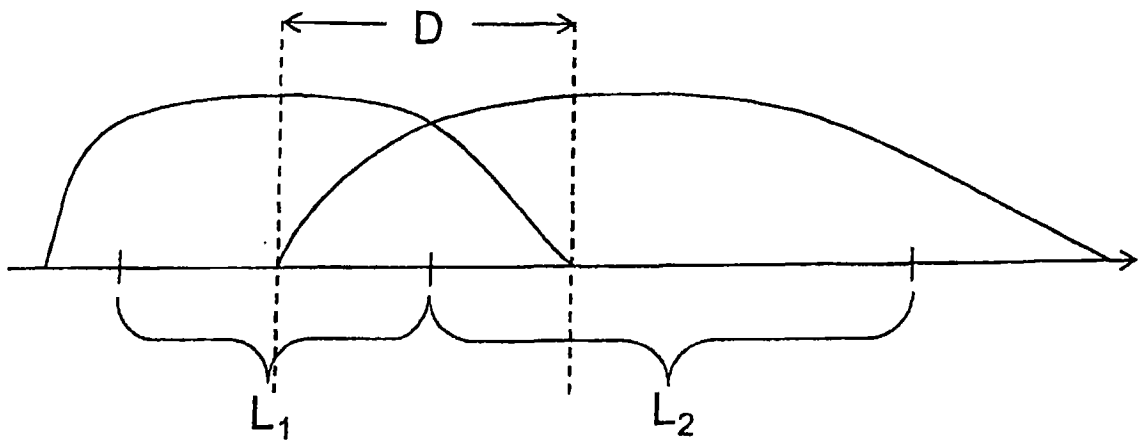


图 23a

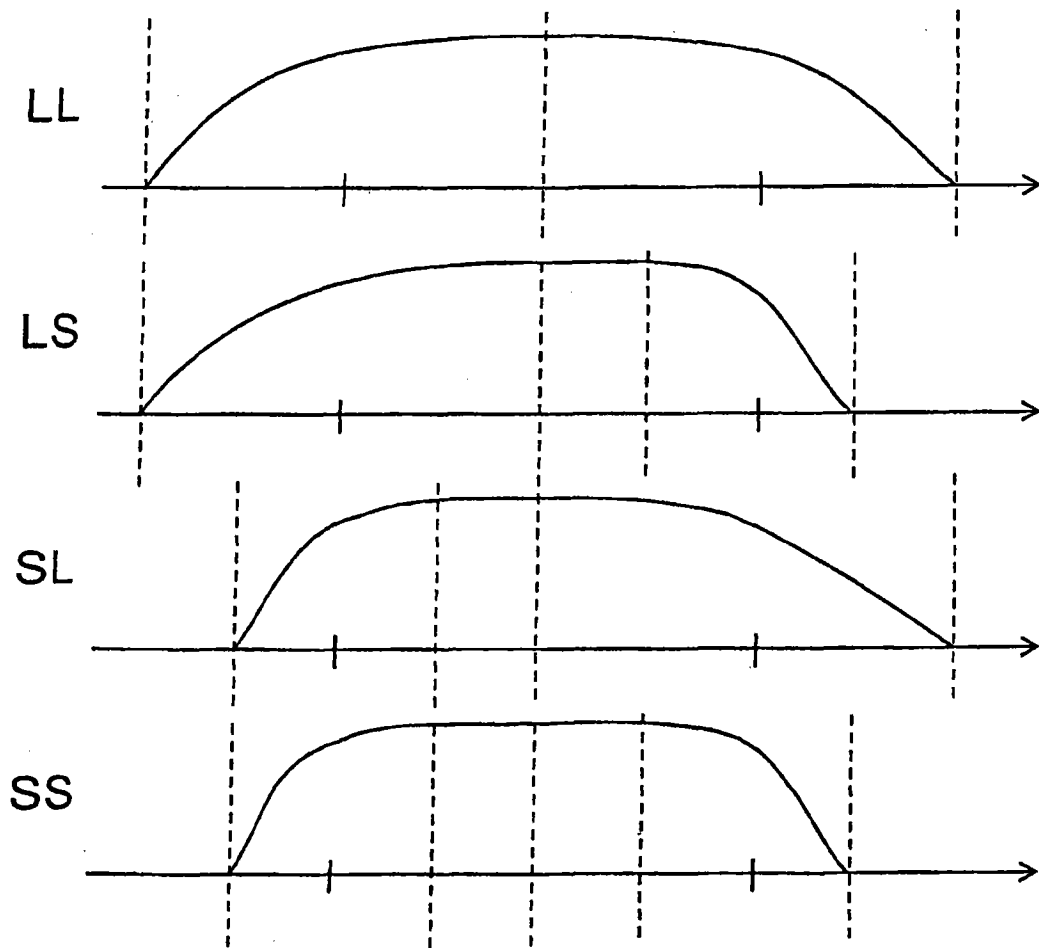


图 23b

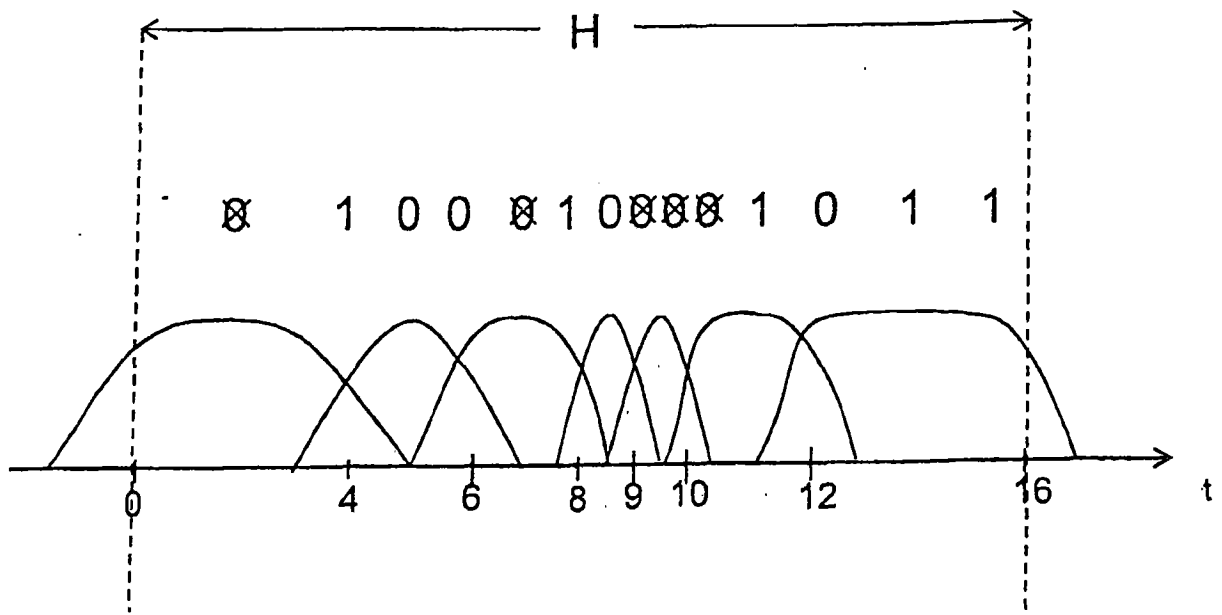


图 23c

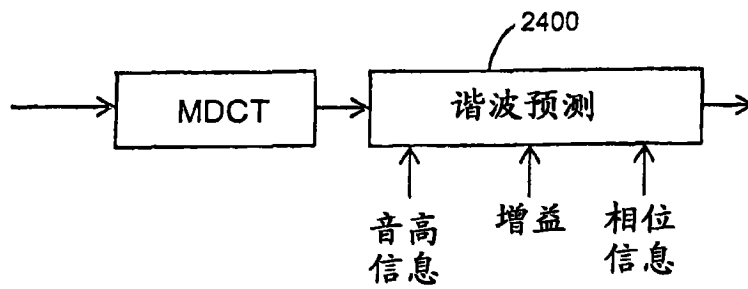


图 24

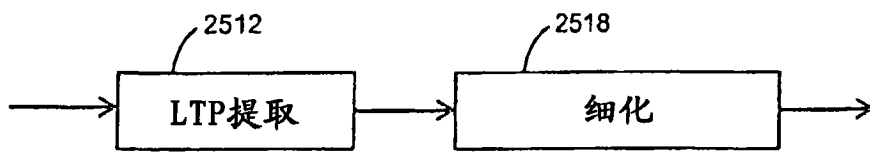


图 25

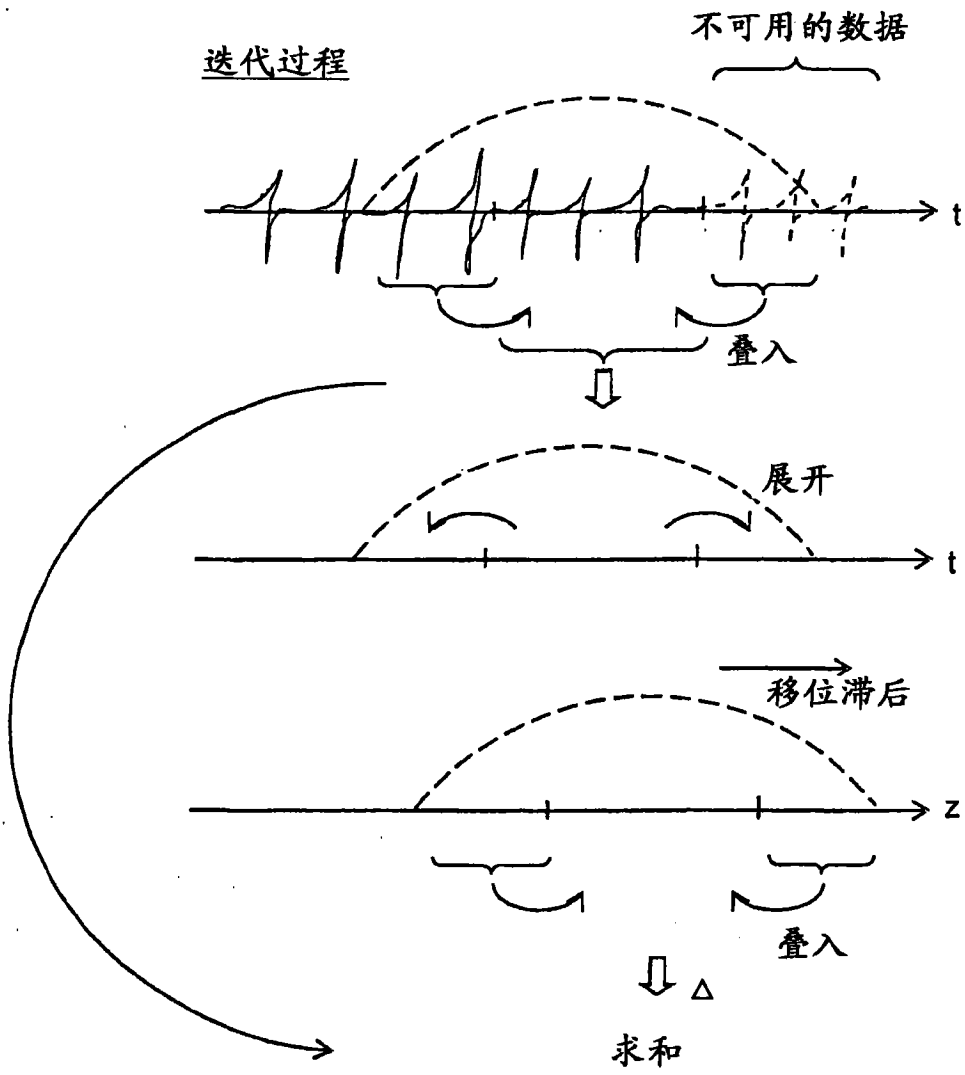


图 25a

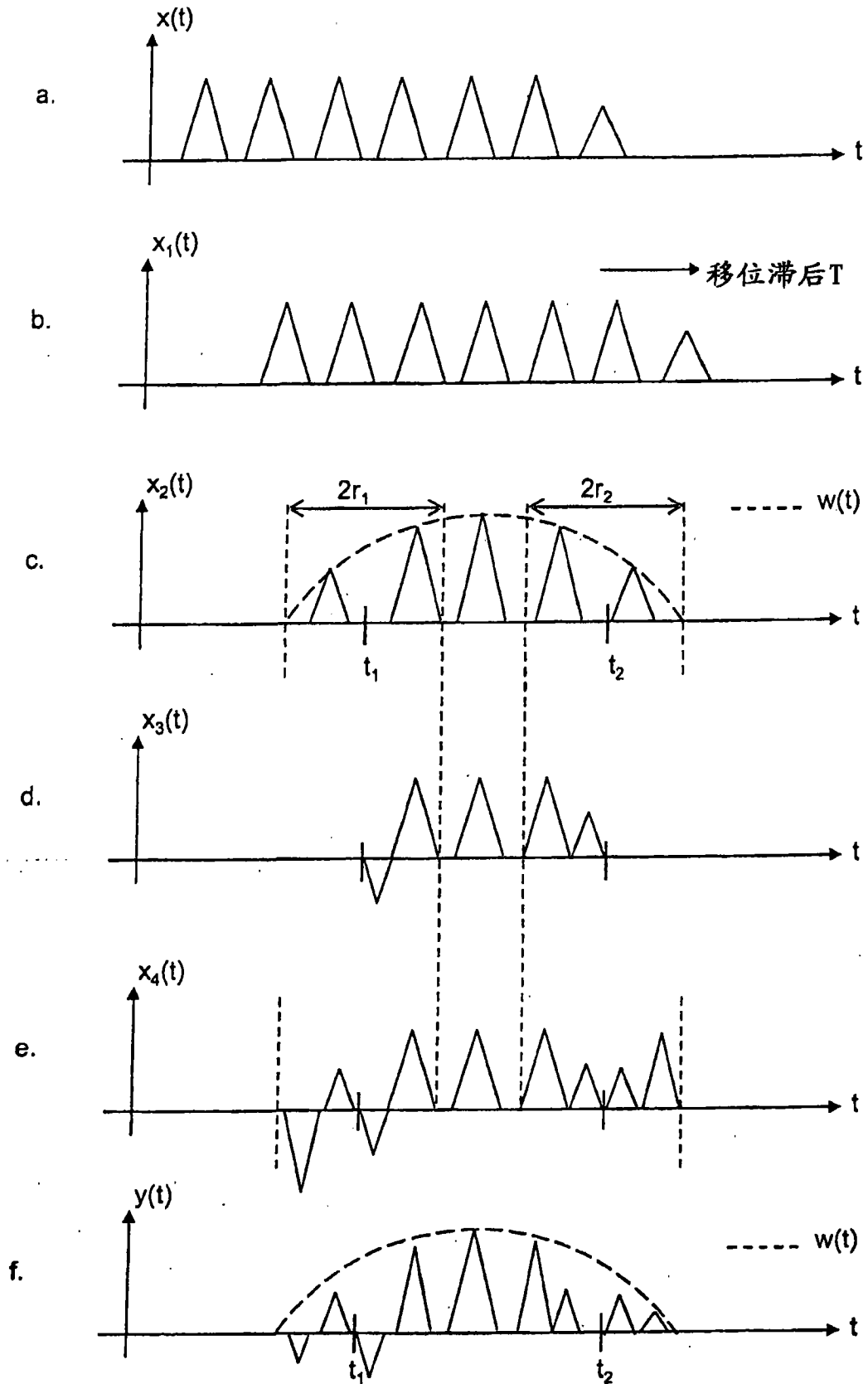


图 25b

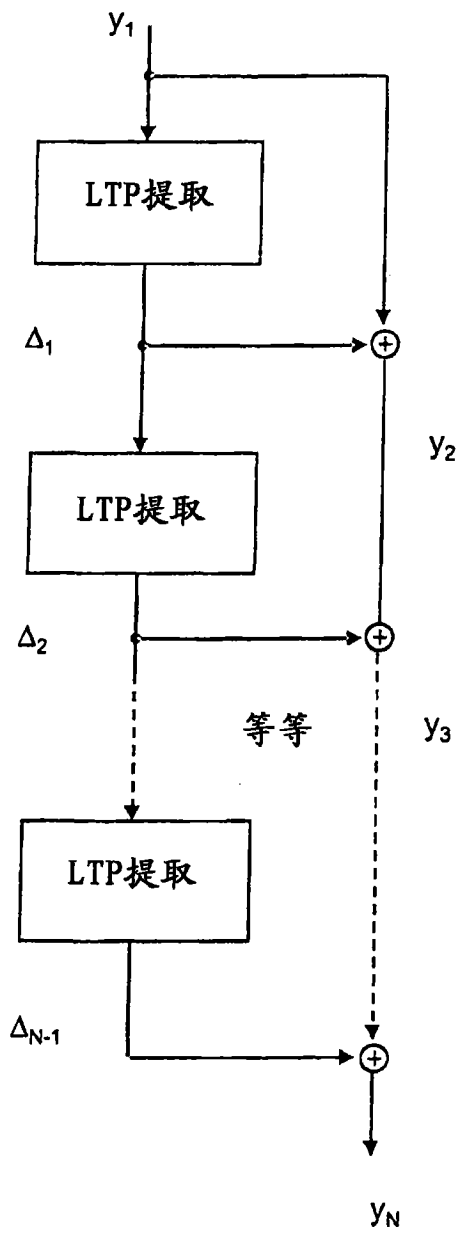


图 25c

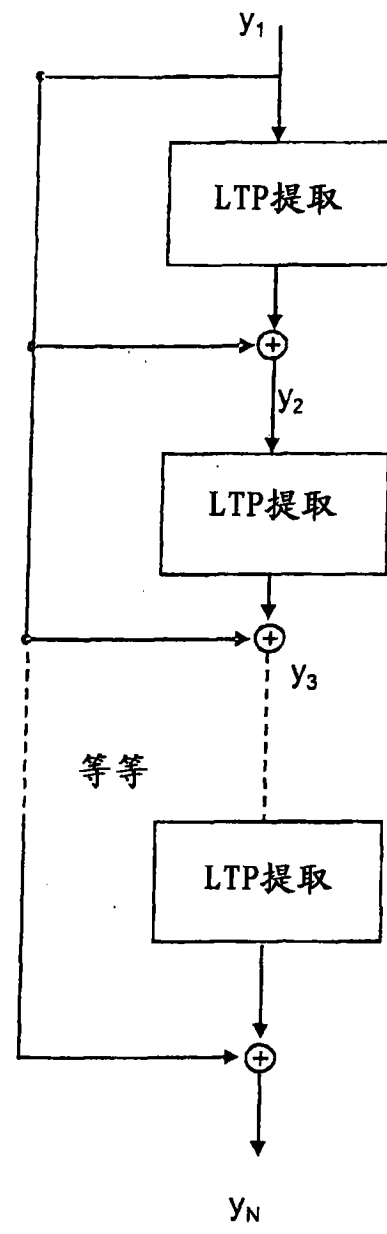


图 25d

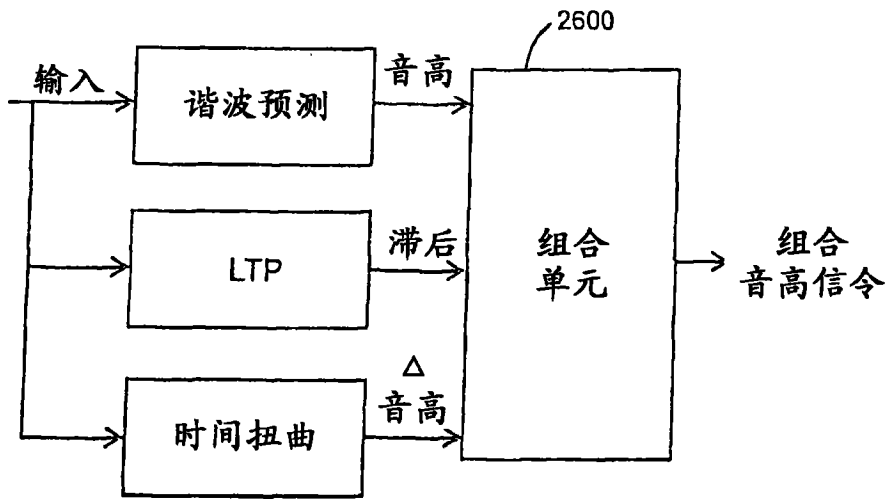


图 26

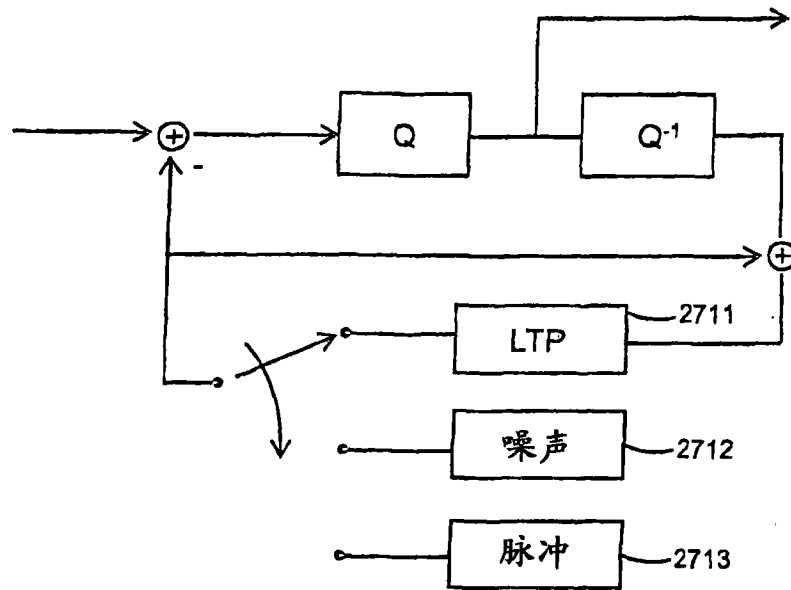


图 27

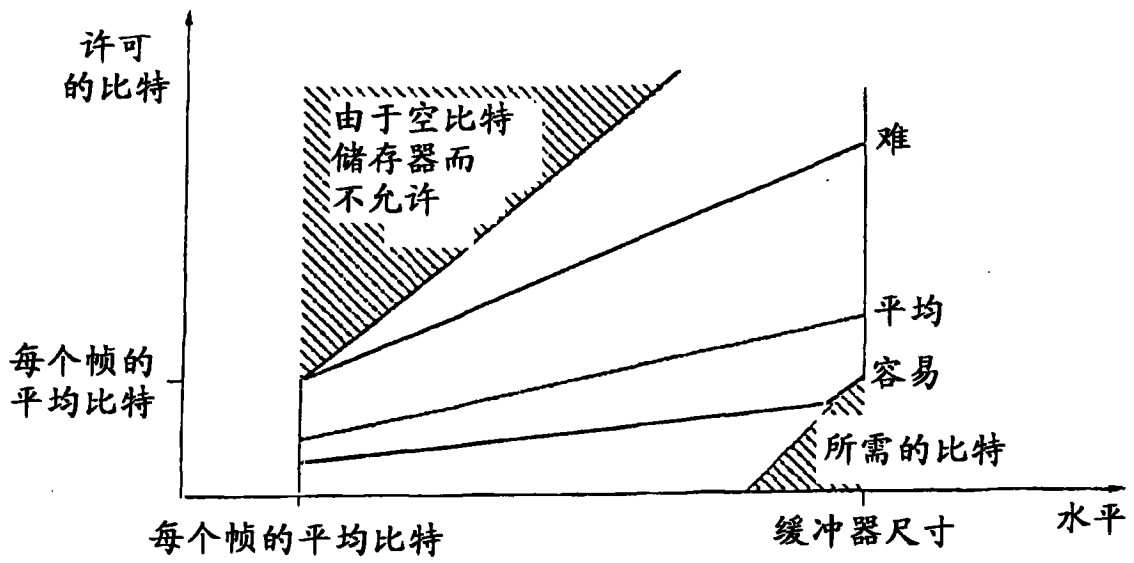


图 28a

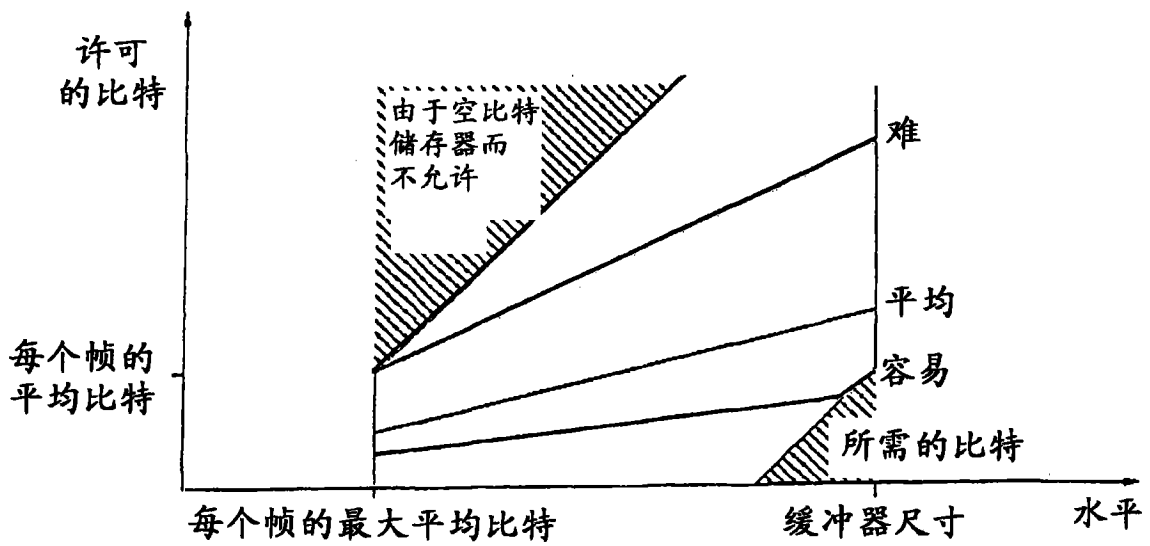


图 28b

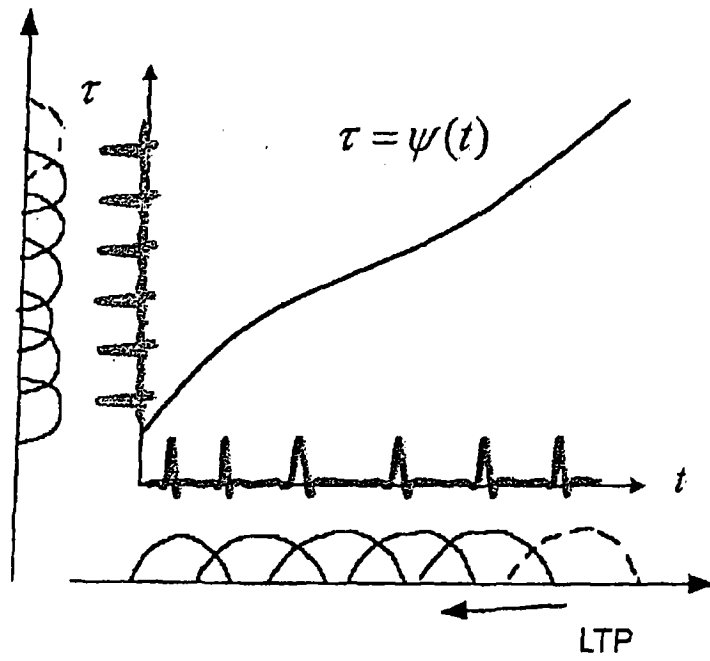


图 29

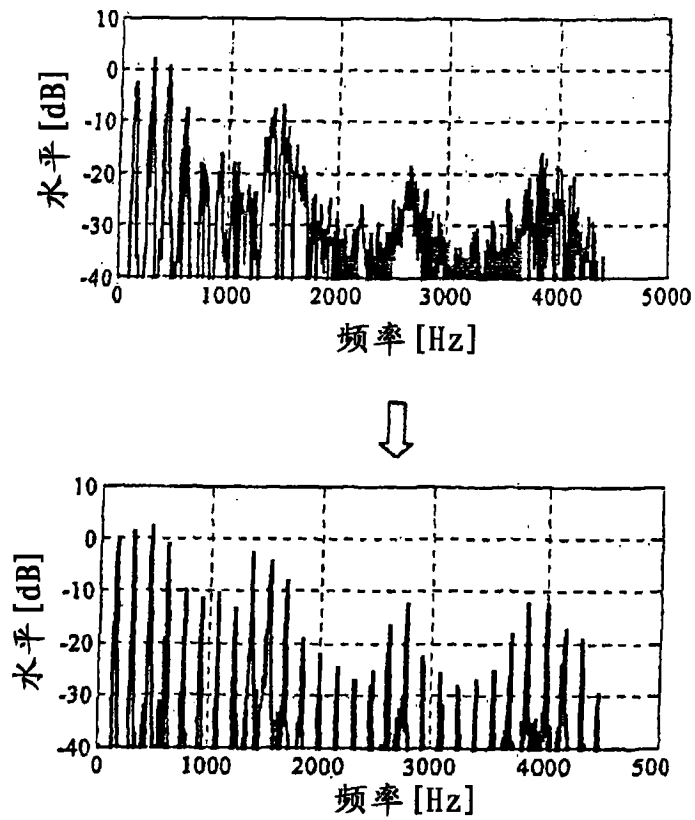


图 29a

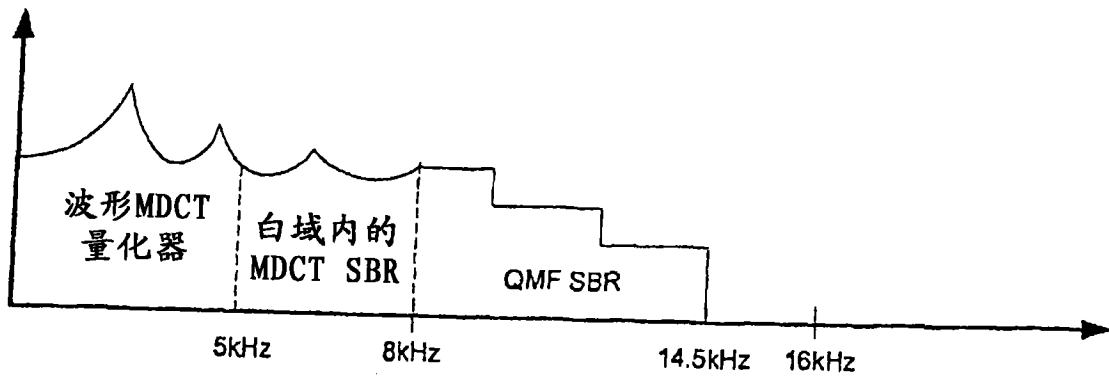


图 30