



[12] 发 明 专 利 说 明 书

[21] ZL 专利号 98800792.4

[45] 授权公告日 2005 年 6 月 15 日

[11] 授权公告号 CN 1206816C

[22] 申请日 1998.6.9 [21] 申请号 98800792.4

[30] 优先权

[32] 1997. 6. 10 [33] SE [31] 9702213 - 1

[32] 1997. 12. 12 [33] SE [31] 9704634 - 6

[32] 1998. 1. 30 [33] SE [31] 9800268 - 6

[86] 国际申请 PCT/IB1998/000893 1998.6.9

[87] 国际公布 WO1998/057436 英 1998.12.17

[85] 进入国家阶段日期 1999.2.10

[71] 专利权人 编码技术股份公司

地址 瑞典索尔纳

[72] 发明人 拉斯·古斯塔夫·里杰利德

泊·鲁那·阿尔宾·埃克斯兰德

拉斯·夫莱基克·汉恩

汉斯·玛格那斯·克里斯托弗·克乔

林

审查员 凌 林

[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

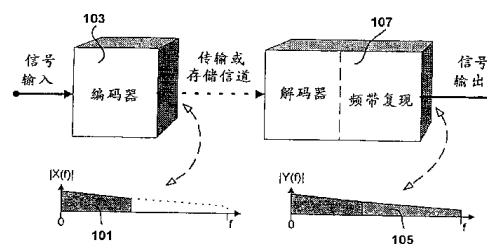
代理人 于 静

权利要求书 7 页 说明书 31 页 附图 13 页

[54] 发明名称 采用频带复现增强源编码

[57] 摘要

本发明提出增强源编码系统的新方法和装置。本发明采用在编码器(103)之前或之中减小带宽(101), 而后再在解码器(107)处复现频带(105)。这是通过使用新的搬移方法结合频谱包络调节实现的。从而在给定的感觉质量下降低了比特率, 或在给定的比特率下改进了感觉质量。本发明最好集成在硬件或软件编解码器中, 但是也可以作为与编解码器结合的独立的处理器实现。本发明提供了实际上与编解码器类型及技术进步的实质性的改进。



1. 一种对编码信号进行解码的方法，该编码信号源自原始信号并仅表示包含在原始信号中的频带的一部分，该方法包括：

提供多个子带的子带样本或多个频谱系数，子带样本或频谱系数表示编码信号的内容；

抽取带通信号，该带通信号具有来自预定数目的分析子带的子带样本，或具有预定数目的分析频谱系数，该预定数目小于表示编码信号的频率内容的子带或频谱系数的总数；

搬移包含在带通信号中的分析子带的子带样本或分析频谱系数到包含在原始信号中、而不包含在编码信号中的重构频带，该重构频带具有合成子带或合成频谱系数，搬移步骤包含将分析子带再连接到合成子带或将分析频谱系数再连接到合成频谱系数；

其中，在搬移步骤之前或之后，利用源自原始信号或编码信号的频谱包络信息，调节包含在带通信号中的子带样本或频谱系数的频谱包络，以获得经过频谱包络调节的、搬移的子带样本或经过频谱包络调节的、搬移的频谱系数，所述频谱包络信息控制频谱包络调节，使得重构频率的频谱包络由频谱包络信息来确定；和

将子带样本和被调节的、搬移的子带样本、或频谱系数和被调节的搬移的频谱系数相组合，以获得解码输出信号，该解码输出信号具有包含了编码信号的频率内容和重构频带的频率内容的频率内容。

2. 根据权利要求1的方法，其中设置被调节的、搬移的子带样本或被调节的、搬移的频谱系数的通带不与或只是部分地与表示编码信号的频率内容的子带样本或频谱系数的通带重叠。

3. 根据权利要求1的方法，还包含利用编码信号估计重构频带的频谱包络信息的步骤。

4. 根据权利要求1的方法，其中编码信号还包含有关重构频带的所传输的频谱包络信息，所述频带是当对原始信号编码时被丢弃的，该方法还包含抽取所传输的原始信号的丢弃频带的包络信息的步

骤。

5. 根据权利要求4的方法, 其中频谱包络信息作为编码信号的任意数目子带频道中的子带样本传输, 其中设置所述这些子带频道的增益为低电平。

6. 根据权利要求4的方法, 其中包络信息作为标度因子被传输, 不传输编码信号中对应的子带样本。

7. 根据权利要求4的方法, 其中包络信息作为标度因子传输, 且设置编码信号中对应的子带样本为零或常数值。

8. 根据权利要求1的方法, 其中解码输出信号为单声道音频信号, 该方法还包含下述步骤: 将解码输出信号分解为两个信号, 每一信号由所述输出信号和同一信号的延迟版本组成, 以获得准立体声信号。

9. 根据权利要求8的方法, 其中通过上边带(USB)调制获得所述移动步骤。

10. 根据权利要求1的方法,

其中所述提供步骤包含下述步骤:

通过一组具有分别包括频率 $[f_1, \dots, f_N]$ 的通带的 $N \geq 2$ 个带通滤波器对信号滤波;

其中所述搬移步骤包含下述步骤:

在频率上移动带通信号到包含频率 $M[f_1, \dots, f_N]$ 的区域, 其中 $M \neq 1$ 是搬移因子。

11. 根据权利要求1的方法,

其中所述提供步骤包含下述步骤:

使用分析滤波器组或产生低通型实或复值子带信号这种性质的变换, 对信号进行带通滤波;

其中所述搬移步骤包括下述步骤:

在合成滤波器组或变换中, 把所述分析滤波器组或变换的任意数目的频道 k 填补频道 Mk , 其中 M 是大于等于2的整数, 并且是搬移因子; 以及

其中其中在组合步骤中使用合成滤波器组或变换。

12. 根据权利要求 11 的方法, 其中滤波器组被最大抽取并根据以下关系进行填补步骤

$$v_{MK}(n) = (-1)^{(M-1)kn} v_K(n),$$

其中 $(-1)^{(M-1)kn}$ 是校正因子, $v_K(n)$ 是频道 k 的子带信号, 且 $v_{MK}(n)$ 是频道 Mk 的子带信号, 从而获得频谱反相子带信号的补偿。

13. 根据权利要求 11 的方法, 其中在填补步骤中, 采用所述搬移因子 M 的多个值。

14. 根据权利要求 1 的方法, 其中所述搬移步骤包括下述步骤:

填补来自一个分析滤波器组或变换的频道 k 的分析子带的子带样本或分析频率系数的相位为与合成频道 Mk 相关的子带样本的相位, 其中 M 是大于等于 2 的整数, 并且是搬移因子, k 是频道号; 以及

填补来自一个分析滤波器组或变换的连续频道 l 的分析子带的子带样本或分析频谱系数的量值为与连续合成频道 $l+S$ 相关的子带样本或频谱系数的量值, 其中 S 是 $\neq 1$ 的整数, l 是频道号。

15. 根据权利要求 14 的方法, 其中在搬移步骤中, 当在组合步骤中使用所述合成滤波器组或变换之前, 频道 k 的子带样本的相位乘以因子 M 。

16. 根据权利要求 14 的方法, 其中 $M=K^{\pm 1}$, 其中 K 是 >1 的整数。

17. 一种从至少其一个频带被丢弃的输入信号提供按搬移因子 M 搬移的搬移信号方法, 包含下述步骤:

通过一个并联 L 个具有如下的脉冲响应的滤波器的组对输入信号滤波

$$h_k(n) = K p_0(n) \exp \left[j \frac{\pi}{2L} (2k+1) \left(n - \frac{N-1}{2} \right) + j(-1)^k \frac{\pi}{4M} \right],$$

其中 $k=0,1,\dots,L-1$, K 是常数, 而 $p_0(n)$ 是长度为 N 的低通原型滤波器, 产生一组 L 个复值信号;

使用因子 L/M 对所述一组 L 个信号下降采样, 产生一组 L 个复值子带信号;

使所述一组复值子带信号的相位角乘以 M , 给出新的一组子带信号;

选择所述新的一组子带信号的实部, 得出一组 L 个实值子带信号;

使用因子 L' 对所述一组实值子带信号的子集上升采样, 产生一组实值信号;

通过一个并联 L' 个具有如下的脉冲响应的滤波器的组对所述一组实值信号滤波

$$f_k(n) = K' p'_0(n) \cos \left[\frac{\pi}{2L'} (2k+1) \left(n - \frac{N'-1}{2} \right) - (-1)^k \frac{\pi}{4} \right],$$

其中 $k=0,1,\dots,L'-1$, K' 是常数, 而 $p'_0(n)$ 是长度为 N' 的低通原型滤波器, 形成一组 L' 个经滤波的信号; 以及

把所述一组 L' 个经滤波的信号和输入信号相加而产生一个搬移的信号。

18. 根据权利要求 17 的方法, 其中所述相位角的相乘步骤和所述实部选择以得到一组 L 个实值子带信号的步骤按下述步骤计算:

把所述一组复值子带信号定义为

$$Z_k(n) = R_k(n) + jI_k(n),$$

其中 $R_k(n)$ 和 $I_k(n)$ 分别是 $Z_k(n)$ 的实部和虚部;

计算所述一组实值子带信号 $W_k(n)$ 如下

$$W_k(n) = |Z_k(n)| \cos \left\{ \text{Marctan} \left(\frac{I_k(n)}{R_k(n)} \right) \right\},$$

其中 $|Z_k(n)| = \sqrt{R_k(n)^2 + I_k(n)^2}$, 而 M 是大于等于 2 的整数, 并且是搬移因子, 使用下述三角恒等式

$$\cos(M\alpha) = \cos^M(\alpha) - \left(\frac{M}{2}\right) \sin^2(\alpha) \cos^{M-2}(\alpha) + \left(\frac{M}{4}\right) \sin^4(\alpha) \cos^{M-4}(\alpha) \dots,$$

其中 $\alpha = \arctan\{I_k(n)/R_k(n)\}$, 及下述关系式

$$\cos(\alpha) = \frac{R_k(n)}{|Z_k(n)|} \text{ 并且 } \sin(\alpha) = \frac{I_k(n)}{|Z_k(n)|}.$$

19. 一种对编码信号进行解码的解码器, 该编码信号源自原始信号并仅表示原始信号的包含的频带的一部分, 该解码器包括:

提供器, 用于提供多个子带的子带样本或多个频谱系数, 子带样本或频谱系数表示编码信号的内容;

抽取器, 用于抽取带通信号, 该带通信号具有来自预定数目的分析子带的子带样本, 或具有预定数目的分析频谱系数, 该预定数目小于表示编码信号的频率内容的子带或表示频谱系数的总数;

搬移器, 用于搬移包含在带通信号中的分析子带的子带样本或分析频谱系数到包含在原始信号中、而不包含在编码信号中的重构频带, 所述重构频带具有合成子带或合成频谱系数, 搬移步骤包含将分析子带再连接到合成子带或将分析频谱系数再连接到合成频谱系数;

其中, 在经过搬移器的处理之前或之后, 利用源自原始信号或编码信号的频谱包络信息调节包含在带通信号中的子带样本或频谱系数的频谱包络, 以获得经过频谱包络调节的、搬移的子带样本或经过频谱包络调节的搬移的频谱系数, 所述频谱包络信息控制频谱包络调节, 使得重构频率的频谱包络由频谱包络信息来确定; 和

组合器, 用于将子带样本和被调节的搬移的子带样本、或频谱系数和被调节的搬移的频谱系数相组合, 以获得解码输出信号, 该解码输出信号具有包含了编码信号的频率内容的, 和重构频带的频率内容的频率内容。

20. 根据权利要求 19 的解码器, 其中解码输出信号为单声道音频信号, 该解码器还包括:

用于由解码输出信号形成第一延迟信号的第一延迟器和第一衰减器;

用于由解码输出信号形成第二延迟信号的与第一延迟器不同的第二延迟器和第二衰减器;

用于把所述解码输出信号和所述第一延迟信号相加, 形成左声道输出信号的第一加法器; 以及

第二加法器，用于把所述解码输出信号和所述第二延迟信号相加，形成右声道输出信号，从而获得准立体声信号。

21. 一种用于从原始信号形成编码信号的编码器，该编码器包含：

处理器，其用于产生编码的原始信号，该编码的原始信号仅表示包含在原始信号中的频带的一部分，和用于对包含在原始信号中而不包含在编码的原始信号中的频带导出频谱包络信号；和

复用器，其用于形成编码信号，该编码信号包含编码的原始信号和频谱包络信息，其中处理器被设置成提供频谱包络信息作为任意数目的子带频率中的子带样本，其中子带频道的增益被设置为低电平，或者其中处理器被设置成产生频谱包络信息作为标度因子，对标度因子不传输对应的子带样本，或者其中处理器被设置成产生频谱包络信息作为标度因子，并且对应的子带样本被设置为零或常数值，从而降低子带样本的熵。

22. 一种从至少其一个频带被丢弃的输入信号提供按搬移因子 M 搬移的搬移信号的装置，包含：

通过一个并联 L 个具有如下的脉冲响应的滤波器的组对输入信号滤波的滤波器

$$h_k(n) = K p_0(n) \exp \left[\frac{\pi}{M} \left(n - \frac{n-1}{M} \right) + (-1)^k \frac{\pi}{M} \right],$$

其中 $k=0,1,\dots,L-1$, K 是常数，而 $p_0(n)$ 是长度为 N 的低通原型滤波器， M 是大于等于 2 的整数，并且是搬移因子，产生一组 L 个复值信号；

用于使用因子 L/M 对所述一组 L 信号下降采样，产生一组 L 个复值子带信号的下降采样器；

用于使所述一组复值子带信号的相位角乘以 M ，给出新的一组子带信号的乘法器；

用于选择所述新的一组子带信号的实部，得出一组 L 个实值子带信号的选择器；

使用因子 L' 对所述一组 L 个实值子带信号的子集上升采样，产

生一组实值信号的上升采样器;

通过一个并联 L' 个具有如下的脉冲响应的滤波器组对所述一组实值信号滤波的滤波器

$$f_k(n) = K' p'_0(n) \cos \left[\frac{\pi}{L'} \left(n - \frac{L'-1}{2} \right) - (-1)^k \frac{\pi}{2} \right],$$

其中 $k=0,1,\dots,L'-1$, K' 是常数, 而 $p'_0(n)$ 是长度为 N' 的低通原型滤波器, 形成一组 L' 个经滤波的信号; 以及

用于把所述一组 L' 个经滤波的信号和输入信号相加而产生一个搬移的信号的加法器。

23. 一种用于从原始信号形成编码信号的方法, 该方法包含:

产生编码的原始信号, 该编码的原始信号仅表示包含在原始信号中的频带的一部分, 和对包含在原始信号中而不包含在编码的原始信号中的频带导出频谱包络信号; 和

形成编码信号, 该编码信号包含编码的原始信号和频谱包络信息, 其中执行产生和导出步骤以提供作为任意数量的子带信道中的子带样本的频谱包络信息, 其中将子带信道的增益设置为低水平, 或者产生作为标度因子的频谱包络信息, 但不发送标度因子的对应子带样本, 或者产生作为标度因子的频谱包络信息并且设置到零或固定值对应子带样本, 使得降低子带样本的包络。

采用频带复现增强源编码

技术领域

在源编码系统中，数字数据在传输或存储之前被压缩，以便降低所需的比特率或存储量。本发明涉及利用频带复现(SBR)改进源编码系统的新方法和装置。在保持相同的感知质量之下达到了明显的比特率降低，或者反过来，在给定的比特率之下达到了感知质量的改进。这是利用在编码器侧减小带宽而以后在解码器处复现频带实现的，从而本发明开拓了频域中信号冗余的新概念。

背景技术

音频源编码技术可分为两类：自然音频编码和语音编码。自然音频编码通常用于音乐或中等比特率的任意信号，并一般提供很宽的音频带宽。语音编码器基本上限制在语音再生，但另一方面可在很低的比特率使用，虽然音频带宽低。宽带语音提供胜于窄带语音的显著主观质量的改进。增加带宽不仅改进了语音的可理解性和自然性，而且便于对说话者识别。于是，宽带语音编码在下一代电话系统中是很重要的项目，由于多媒体领域的急剧的发展，音乐和其它非语音信号通过电话系统高质量的传输是一种诱人的前景。

高保真线性 PCM 信号对感知熵就比特率来说是非常低效的。CD 标准规定采样频率为 44.1 kHz，每个样值为 16 位和立体声。这等于 1411 kbit/s 的比特率。为了有力地降低比特率，可以使用分裂频带感知音频编解码器进行源编码。这些自然音频编解码器利用了信号中的感知非相干性和统计冗余性。使用最佳编解码技术，对标准的 CD 格式信号能够达到大约 90% 数据压缩，而实际上感知不到质量有所下降。很高的立体声声音质量能够以 96 kbit/s，即大约为 15:1 的压缩因子获得。某些感知编解码器提供了更高的压缩比。为了达到这一点，通常要降低采样率，于是就要降低音频带宽。通常还要减少量化电平数，允许偶而有可听出的量化失真，并通过强

度编码,使用立体声场的变差。过分使用这种方法的结果是令人生厌的感觉质量下降。当前的编解码器技术近于极限,故不能指望进一步提高编码增益。为了进一步改进编码性能,需要新的方法。

人类噪音和许多乐器产生振动系统发出的准平稳信号。根据付立叶理论,任何周期信号能够表示为频率为 $f, 2f, 3f, 4f, 5f$ 等的正弦波之和。这里 f 是基频。这些频率形成一个调和级数。这种信号的带宽限制等价于调和级数的截尾。这种截尾改变了听到的乐器或噪音的音调,音色,并产生听起来“压抑”或“迟钝”的音频信号,并可能降低可理解性。这样,高频对于声音质量的主观印象是重要的。

先有的技术方法主要倾向于语音编解码器性能的改进,特别倾向于高频再生(HFR),这是语音编码中的一个问题。这种方法采用宽带线性频移,非线性性或折叠[U.S.Pat 5,127,054],这将产生互调或其它非谐波频率成分,当它们加到音乐信号时会引起严重不谐和音。在语音编码文献中这种不谐和音称为“刺耳的”或“难听的”声音。其它合成语音的 HFR 方法产生正弦谐波,该谐波基于基音调的估计,于是局限于音调平稳的声音[U.S.Pat.4,771,465]。这种先有的技术方法虽然对低质量的语音应用是有用的,但是对高质量的语音或音乐信号不适用。一些方法试图改进高质量音频源编解码器。一种方法使用在解码器产生的合成噪声信号以代替先前由编码器丢弃的语音或音乐中的类似噪声的信号[“Improving Audio Codecs by Noise Substitution” D. Schultz, JAES, Vol. 44, No. 7/8, 1996]。当噪声信号存在时,这是在基于间歇另一种正常传输的高频带内进行的。另一种方法则再生某些在编码过程中丢失的高频带谐波[“Audio Spectral Coder” A.J.S. Ferreira, AES Preprint 4201, 100th Convention, May 11-14 1996, Copenhagen],这也是依靠音调信号和音高检测。两种方法都基于低占空度工作,提供比较有限的编码或性能增益。

发明内容

本发明对数字源编码系统的实质改进或更具体来说对音频编解码器的改进,提供了新的方法和装置。目的包括降低比特率或改进感觉质量或这两者。本发明基于采用谐波冗余的新方法,提供在传输或存储之前丢弃信

号的一些通带的可能性。如果解码器根据本发明进行高质量的频谱复现，则不会觉察到感觉质量下降。被丢弃的这些比特表示在固定感觉质量下的编码增益。另外，如果比特率保持不变，那么就能够分配更多的比特用于对低频带信号的编码，因而达到较高的感觉质量。

本发明提供了一种对编码信号进行解码的方法，该编码信号源自原始信号并仅表示包含在原始信号中的频带的一部分，该方法包括：提供多个子带的子带样本或多个频谱系数，子带样本或频谱系数表示编码信号的内容；抽取带通信号，该带通信号具有来自预定数目的分析子带的子带样本，或具有预定数目的分析频谱系数，该预定数目小于表示编码信号的频率内容的子带或频谱系数的总数；搬移包含在带通信号中的分析子带的子带样本或分析频谱系数到包含在原始信号中、而不包含在编码信号中的重构频带，该重构频带具有合成子带或合成频谱系数，搬移步骤包含将分析子带再连接到合成子带或将分析频谱系数再连接到合成频谱系数；其中，在搬移步骤之前或之后，利用源自原始信号或编码信号的频谱包络信息，调节包含在带通信号中的子带样本或频谱系数的频谱包络，以获得经过频谱包络调节的、搬移的子带样本或经过频谱包络调节的、搬移的频谱系数，所述频谱包络信息控制频谱包络调节，使得重构频率的频谱包络由频谱包络信息来确定；和将子带样本和被调节的、搬移的子带样本、或频谱系数和被调节的搬移的频谱系数相组合，以获得解码输出信号，该解码输出信号具有包含了编码信号的频率内容和重构频带的频率内容的频率内容。

本发明还提供了一种从至少其一个频带被丢弃的输入信号提供按搬移因子 M 搬移的搬移信号方法，包含下述步骤：通过一个并联 L 个具有如下的脉冲响应的滤波器的组对输入信号滤波 $h_k(n) = Kp_0(n) \exp\left[j\frac{\pi}{2L}(2k+1)\left(n - \frac{N-1}{2}\right) + j(-1)^k \frac{\pi}{4M}\right]$ ，其中 $k=0,1,\dots,L-1$, K 是常数，而 $p_0(n)$ 是长度为 N 的低通原型滤波器，产生一组 L 个复值信号；使用因子 L/M 对所述一组 L 个信号下降采样，产生一组 L 个复值子带信号；使所述一组复值子带信号的相位角乘以 M ，给出新的一组子带信号；选择所述新的一组子带信号的实部，得出一组 L 个实值子带信号；使用因子 L' 对所述一组实值子带信号的子集上升采样，产生一组实值信号；通过一个并联 L' 个具有

如下的脉冲响应的滤波器的组对所述一组实值信号滤波 $f_k(n) = K'p'_0(n) \cos\left[\frac{\pi}{2L'}(2k+1)\left(n - \frac{N'-1}{2}\right) - (-1)^k \frac{\pi}{4}\right]$, 其中 $k=0,1,\dots,L'-1, K'$ 是常数, 而 $p'_0(n)$ 是长度为 N' 的低通原型滤波器, 形成一组 L' 个经滤波的信号; 以及把所述一组 L' 个经滤波的信号和输入信号相加而产生一个搬移的信号。

本发明还提供一种对编码信号进行解码的解码器, 该编码信号源自原始信号并仅表示原始信号的包含的频带的一部分, 该解码器包括: 提供器, 用于提供多个子带的子带样本或多个频谱系数, 子带样本或频谱系数表示编码信号的内容; 抽取器, 用于抽取带通信号, 该带通信号具有来自预定数目的分析子带的子带样本, 或具有预定数目的分析频谱系数, 该预定数目小于表示编码信号的频率内容的子带或表示频谱系数的总数; 搬移器, 用于搬移包含在带通信号中的分析子带的子带样本或分析频谱系数到包含在原始信号中、而不包含在编码信号中的重构频带, 所述重构频带具有合成子带或合成频谱系数, 搬移步骤包含将分析子带再连接到合成子带或将分析频谱系数再连接到合成频谱系数; 其中, 在经过搬移器的处理之前或之后, 利用源自原始信号或编码信号的频谱包络信息调节包含在带通信号中的子带样本或频谱系数的频谱包络, 以获得经过频谱包络调节的、搬移的子带样本或经过频谱包络调节的搬移的频谱系数, 所述频谱包络信息控制频谱包络调节, 使得重构频率的频谱包络由频谱包络信息来确定; 和组合器, 用于将子带样本和被调节的搬移的子带样本、或频谱系数和被调节的搬移的频谱系数相组合, 以获得解码输出信号, 该解码输出信号具有包含了编码信号的频率内容的, 和重构频带的频率内容的频率内容。

本发明还提供一种用于从原始信号形成编码信号的编码器, 该编码器包含: 处理器, 其用于产生编码的原始信号, 该编码的原始信号仅表示包含在原始信号中的频带的一部分, 和用于对包含在原始信号中而不包含在编码的原始信号中的频带导出频谱包络信号; 和复用器, 其用于形成编码信号, 该编码信号包含编码的原始信号和频谱包络信息, 其中处理器被设置成提供频谱包络信息作为任意数目子带频率中的子带样本, 其中子带频道的增益被设置为低电平, 或者其中处理器被设置成产生频谱包络信息作

为标度因子，对标度因子不传输对应的子带样本，或者其中处理器被设置成产生频谱包络信息作为标度因子，并且对应的子带样本被设置为零或常数值，从而降低子带样本的熵。

本发明还提供一种从至少其一个频带被丢弃的输入信号提供按搬移因子 M 搬移的搬移信号的装置，包含：通过一个并联 L 个具有如下的脉冲响应的滤波器的组对输入信号滤波的滤波器 $h_k(n) = K p_0(n) \exp\left[\frac{\pi}{2L}(2k+1)\left(n - \frac{N-1}{2}\right) + (-1)^k \frac{\pi}{4}\right]$ ，其中 $k=0,1,\dots,L-1, K$ 是常数，而 $p_0(n)$ 是长度为 N 的低通原型滤波器， M 是大于等于 2 的整数，并且是搬移因子，产生一组 L 个复值信号；用于使用因子 L/M 对所述一组 L 信号下降采样，产生一组 L 个复值子带信号的下降采样器；

用于使所述一组复值子带信号的相位角乘以 M ，给出新的一组子带信号的乘法器；用于选择所述新的一组子带信号的实部，得出一组 L 个实值子带信号的选择器；使用因子 L' 对所述一组 L 个实值子带信号的子集上升采样，产生一组实值信号的上升采样器；通过一个并联 L' 个具有如下的脉冲响应的滤波器组对所述一组实值信号滤波的滤波器 $f_k(n) = K' p'_0(n) \cos\left[\frac{\pi}{2L'}(2k+1)\left(n - \frac{N'-1}{2}\right) - (-1)^k \frac{\pi}{4}\right]$ ，其中 $k=0,1,\dots,L'-1, K'$ 是常数，而 $p'_0(n)$ 是长度为 N' 的低通原型滤波器，形成一组 L' 个经滤波的信号；以及用于把所述一组 L' 个经滤波的信号和输入信号相加而产生一个搬移的信号的加法器。

本发明还提供一种用于从原始信号形成编码信号的方法，该方法包含：产生编码的原始信号，该编码的原始信号仅表示包含在原始信号中的频带的一部分，和对包含在原始信号中而不包含在编码的原始信号中的频带导出频谱包络信号；和形成编码信号，该编码信号包含编码的原始信号和频谱包络信息，其中执行产生和导出步骤以提供作为任意数量的子带信道中的子带样本的频谱包络信息，其中将子带信道的增益设置为低水平，或者产生作为标度因子的频谱包络信息，但不发送标度因子的对应子带样本，或者产生作为标度因子的频谱包络信息并且设置到零或固定值对应子带样本，使得降低子带样本的包络。

本发明假设基于低频带和高频带频谱成分之间的直接关系能够扩展出被截去的谐波序列。如果遵从以下这些规则，这一扩展的序列就在感觉上与原始序列类似：首先，外推的频谱成分与截去的谐波序列必需调和相关，以避免关系不谐和的人工痕迹。本发明使用搬移作为频谱复现过程的手段，这保证了符合这一准则。然而对于成功的操作并不需要低频带频谱成分形成谐波序列，因为与低频带频谱成分谐和相关的新的被复现的成分将不会改变信号的噪声形或瞬变性质。搬移定义为在保持分音频率比的同时将分音从音乐音阶的一个位置向另一位置谐音的搬移。第二，被复现的高频带的频谱包络，即粗燥的频谱分布，必须与原始信号的频谱包络很好地合理相似。本发明提供了两种不同的调节频谱包络的操作模式，SBR-1 和 SBR-2。

用于改进中等质量编解码器应用的 SBR-1 是一种单端过程，它只依赖于包含在解码器接收的低频带或低通信号中的信息。例如使用多项式与规则集或代码簿一同确定并推导出这种信号的频谱包络。这种信息用来连续调节并均衡复现的高频带。本发明的 SBR-1 方法提供了后处理的优点，即不需要编码器侧的修改。播音机构将在频带利用上获得好处或能够提供改进的感觉质量或两者的组合。能够使用现有的比特流句法和标准而无需修改。

用于改进高质量编解码器应用的 SBR-2 是一种双端过程，其中除了根据 SBR-1 传输的低频带信号之外，还对高频带的频谱包络进行编码和传输。由于频谱包络的变化具有比高频带信号成分低得多的速率，故为了成功地表示频谱包络只需要传输有限的信息。SBR-2 能够用来改进当前编解码器技术的性能，对现有的句法或协议不需要或只需要很少的修改，可作为开发未来编解码器的有价值的工具。

当低频带中一些较小的通带在缺乏比特的条件下按心理声学模型规定的那样用解码器略去时，SBR-1 和 SBR-2 都能够用于复现这些通带。其结果是通过在低频带之外的频谱复现和在低频带内的频谱复现改进了感觉质量。而且，SBR-1 和 SBR-2 还能够用于采用比特率可变的编解码器，其中接收机的信号感觉质量依赖于传输频带的条件而变化。这通常意味着接收机音频带宽不良的变化。在这种条件下，SBR 方法能够成功地用来保持不

变的高频带宽，这又改进了感觉质量。

本发明基于连续方式操作，复现任何类型的信号内容，即音调的或非音调的(噪声形和瞬时信号)。本频谱复现方法根据解码器可得的频带生成在感觉上是被丢弃的频带的精确复现的拷贝。因而，SBR方法与先有技术相比提供了相当高的编码增益或改进的感觉质量。本发明能够与这种先有技术编解码器改进方法结合；然而，这种结合不会有性能上的提高。

SBR方法包括以下步骤：

- 对从原始信号推导得出的信号进行编码，形成第一信号，在这过程中信号的一些频带在编码之前或编码期间被丢弃，
- 在第一信号解码期间或解码之后，搬移第一信号的频带，形成第二信号，
- 进行频谱包络调节，以及
- 组合解码信号和第二信号，形成输出信号。

可以将第二信号的通带设置成与第一信号的通带不重叠或部分地重叠，并能够依照原始信号和/或第一信号的瞬时特征或传输信道的状态设置。基于来自所述第一信号的原始频谱包络的估计或被传输的原始信号包络信息，进行频谱包络调节。

本发明包括两种搬移器的基本类型：具有不同性质的多带搬移器和时变模式搜索预测搬移器。基本的多带搬移根据本发明可按以下进行：

- 通过通带分别包括频率 $[f_1, \dots, f_N]$ 的 $N \geq 2$ 个带通滤波器对要被搬移的信号进行滤波，形成 N 个带通信号，
- 将这些带通信号在频率上移到包含频率 $M[f_1, \dots, f_N]$ 的区域，其中 $M \neq 1$ 为搬移因子，以及
- 组合这些搬移了的带通信号，形成搬移的信号。

另外，可以根据本发明如下进行这一基本多带搬移：

- 使用分析滤波器组或产生低通型实或复值子带信号的变换对要被搬移信号的信号进行带通滤波，
- 将所述分析滤波器组或变换的任意个频道 k 连接到合成滤波器组或变换中的频道 Mk ， $M \neq 1$ ，以及

—使用合成的滤波器组或变换形成搬移的信号。

根据本发明一种改进的多带搬移结合了相位调节，故提高了基本多带搬移的性能。

根据本发明的时变模式搜索预测搬移可以如下进行：

- 对第一信号进行瞬时检测，
- 根据瞬时检测的结果确定在复现/丢弃第一信号的部分时要使用第一信号的哪一段，
- 根据瞬时检测的结果调节状态向量和码本的性质，以及
- 基于先前同步点搜索中找到的同步点，搜索第一信号的所选段中的同步点。

根据本发明的 SBR 方法和装置提供了以下特点：

- 1.该方法和装置采用了频域中信号冗余的新概念。
- 2.该方法和装置适用于任意信号。
- 3.每一谐波集是被分别生成和控制的。
- 4.所有被复现的谐波是以形成与现有谐波序列连续这样的方式产生的。
- 5.频谱复现过程基于搬移且不生成或可觉察的人工痕迹。
- 6.频谱复现能够覆盖多个较小的频带和/或一个宽的频率范围。
- 7.在 SBR-1 方法中，处理只在解码器侧进行，能够使用所有的标准和协议而无需修改。
- 8.SBR-2 方法能够根据大多数标准和协议实现，不需要或只要很少的修改。
- 9.SBR-2 方法向编解码器设计者提供了新的有力的压缩工具。
- 10.编码增益是显著的。

许多吸引人的应用涉及各种类型的低比特率编解码器的改进，诸如 MPEG1/2 Layer I/II/III [U.S. Pat. 5,040,217], MPEG2/4 AAC, Dolby AC-2/3, NTT TwinVQ [U.S. Pat. 5,684,920] AT&T/Lucent PAC 等等。本发明还可用于诸如宽带 CELP 和 SB-ADPCM G.722 等高质量语音编解码器，以改进感觉质量。以上的编解码器广泛用于多媒体，电话业，因特网以及专业应用。T-DAB(地面数字音频广播)系统使用低比特率协议，通过使用本发明的方法将增加频

道的利用，或者改进 FM 和 AM DAB 中的质量。卫星 S-DAB 将显著增加，由于所涉及的过高的系统成本，通过使用本发明的方法以增加 DAB 多路复用的节目频道数。此外，首先，使用低速率电话调制解调器，通过因特网能够实现全带宽音频实时流。

附图说明

以下将参照附图通过解释例的方式对本发明进行说明，但并不限制本发明的范围和精神，其中：

图 1 表示根据本发明装在编码系统中的 SBR；

图 2 表示根据本发明的上谐波频谱复现；

图 3 表示根据本发明的带内谐波的频谱复现；

图 4 是对于根据本发明的搬移器的时域实现的框图；

图 5 是表示根据本发明的模式搜索预测搬移器的一个操作周期的流程图；

图 6 是表示根据本发明对于同步点搜索的流程图；

图 7a-7b 表示根据本发明在瞬态期间的码本定位；

图 8 是表示对于根据本发明的 SBR 操作，与适当的滤波器组相连接的时域搬移器的实现的框图；

图 9a-9c 是表示根据本发明为产生 2 次谐波配置的 STFT 分析和合成装置的框图；

图 10a-10b 是根据本发明在 STFT 装置中带有线性频移子带的框图；

图 11 表示使用根据本发明的相位乘法器的一个子带；

图 12 表示根据本发明如何产生 3 次谐波；

图 13 表示根据本发明如何同时产生 3 次和 2 次谐波；

图 14 表示根据本发明几个谐波段非重叠组合的产生；

图 15 表示根据本发明几个谐波段交替组合的产生；

图 16 表示宽带线性频移的产生；

图 17 表示根据本发明如何产生子谐波；

图 18a-18b 是感觉编解码器的框图；

图 19 最大抽取滤波器组的基本结构；

图 20 表示根据本发明在最大抽取滤波器组中 2 次谐波的产生;

图 21 是根据本发明对于最大抽取滤波器组对子带信号操作的改进的多带搬移的框图;

图 22 是表示根据本发明在对子带信号操作的最大抽取滤波器组中改进的多带搬移的流程图;

图 23 表示典型的编解码器的子带采样和标度因子;

图 24 表示根据本发明的 SBR-2 的子带采样和包络信息;

图 25 表示根据本发明的 SBR-2 中隐藏包络信息的传输;

图 26 表示根据本发明的 SBR-2 中冗余编码;

图 27 表示使用根据本发明的 SBR-1 方法的编解码器的实现; 以及

图 28 表示使用根据本发明的 SBR-2 方法的编解码器的实现;

图 29 是根据本发明的“伪立体声”产生器的框图。

具体实施方式

在这里实施例的整个说明中, 强调自然音频源编码应用。然而应当明白, 本发明也可适用于编码和解码音频信号以外的源编码应用范围。

搬移基本原理

根据本发明所定义的搬移是对于频谱复现理想的方法, 与先有技术相比具有若干主要的优点, 诸如: 不需要音调检测, 对单音调和复调节目材料获得同等的高性能, 对于调性和非调性信号同等良好地作搬移。与其它方法不同, 根据本发明的搬移能够用于对任意信号类型进行任何音频源编码的系统。

带有时变振幅的余弦和的形式的离散时间信号 $x(n)$ 的因子为 M 的精确搬移, 由以下关系定义

$$x(n) = \sum_{i=0}^{N-1} e_i(n) \cos(2\pi f_i n / f_s + \alpha_i) \quad (1)$$

$$y(n) = \sum_{i=0}^{N-1} e_i(n) \cos(2\pi M f_i n / f_s + \beta_i) \quad (2)$$

其中 N 是以下称为分音的正弦波的个数, $f_i, e_i(n), \alpha_i$ 分别为各输入的频率, 时间包络和相位常数, β_i 为任意输出相位常数而 f_s 为采样频率, 且 $0 \leq M f_i \leq f_s/2$ 。

图 2 中, 示出 M 次谐波的产生, 其中 $M \geq 2$ 。为了简略起见使用术语 M 次谐波, 虽然该过程对一定的频率区所有的信号都产生 M 次谐波, 它们在许多情形下本身也是某次谐波。带有频域表示 $X(f)$ 的输入信号带宽限制在 0 到 $f_{\max}$ 的范围, 201。范围 $f_{\max}/M$ 到 $Qf_{\max}/M$ 中的信号内容是借助于带通滤波器抽取的, 形成带有频谱 $X_{BP}(f)$ 203 的带通信号, 其中 Q 是所需的带宽扩展因子 $1 < Q \leq M$ 。带通信号被搬移了一个因子 M , 形成带有覆盖范围为 $f_{\max}$ 到 $Qf_{\max}$ 的频谱 $X_T(f)$ 的第二带通信号, 205。这一信号的频谱包络借助于程控均衡器调节, 形成带有频谱 $X_E(f)$ 的信号 207。这一信号与输入信号的为补偿由带通滤波器和搬移器造成的延迟的延迟版本组合, 从而形成带有覆盖 0 到 $Qf_{\max}$ 的范围的频谱 $Y(f)$ 的输出信号 209。或者, 在搬移 M 之后可以使用截止频率 $f_{\max}$ 和 $Qf_{\max}$ 进行带通滤波。当然能够使用多个搬移器同时产生不同次谐波。如图 3 所示, 以上方案还可以用于“填充”输入信号中的阻带, 其中输入信号具有从 f_0 到 Qf_0 延伸的阻带。然后抽取通带 $[f_0/M, Qf_0/M]$ 303, 搬移一个因子 M , 成为 $[f_0, Qf_0]$ 305, 再调节包络 307, 并与延迟输入信号组合形成带有频谱 $Y(f)$ 的输出信号 309。

可以使用精确搬移的一种近似。根据本发明, 使用不谐和音理论确定这种近似的质量。不谐和音的一种标准由 Plomp[“Tonal Consonance and Critical Bandwidth” R.Plomp, W.J.M. Levelt JASA, Vol38,1965]提出, 并声称, 如果两个分音频率差在分音所在的临界带的带宽大约 5 到 50% 以内, 则认为它们是不谐和的。例如, 对于给定频率的临界带宽能够由以下近似表示

$$cb(f) = 25 + 75(1 + 1.4(\frac{f}{1000})^2)^{0.69} \quad (3)$$

其中 f 和 cb 的单位为 Hz。而且 Plomp 说, 如果两个分音频率差大约小于它们所在的临界带宽的百分之五, 则人类的听觉系统不能分辨这两个分音。方程式 2 中的精确搬移由以下公式近似表示

$$y_{approx}(n) = \sum_{i=0}^{N-1} e_i(n) \cos(2\pi(Mf_i \pm \Delta f_i)n / f_s + \beta_i) \quad (4)$$

其中 Δf_i 是对精确搬移的偏移。如果输入分音形成一个谐波序列，则本发明的假说称，对被搬移的分音的谐波序列的偏移不得超过它们所在的临界带宽的百分之五。这可以解释，为什么先有技术方法给出不满意的“刺耳的”或“难听的”结果，因为宽带线性频率移动产生了比可接受的偏移大得多的偏移。当先有技术方法对于只一个输入分音产生一个以上的分音时，这些分音必须仍然在上述偏移限度之内，感觉起来才如同一个分音。这又解释了使用非线性等的先有技术方法所获得的不良结果，因为它们产生了不在偏移限度之内的互调分音。

当使用根据本发明的基于以上搬移的频谱复现方法时，实现了以下重要性质：

- 通常，被复现的谐波和现有的分音之间不出现频域重叠。
- 被复现的分音与输入信号的谐分谐和相关，且不会引起任何令人讨厌的不谐和或人工痕迹。
- 被复现谐波的频谱包络形成输入信号频谱包络的平滑连续，感觉上与原始包络匹配。

基于时变模式搜索预测的搬移

有各种方式指定所需的搬移器。典型的时域实现通过基于音调周期复现各信号段在时间上扩展信号。然后这一信号以不同的速率读出。不幸的是这些方法对于信号段的精确的时间拼接严格依赖于音调检测。此外，强制处理基于音调周期的信号段使得它们对瞬态是敏感的。由于检测到的音调周期可能比实际的瞬态长得多，故显然有复现整个瞬态而不是在时间上扩展它的危险。另一种时域算法使用输出信号的模式搜索预测获得语音信号的时间扩展/压缩[“Pattern Search Prediction of speech” R.Bogner,T.Li, Proc.ICASSP'89,Vol.1, May 1989, “Time-Scale Modification of Speech based on nonlinear Oscillator Model” G.Kubin,W.B.Kleijn,IEEE,1994]。这是颗粒合成的形式，其中输入信号被分为一系列用来合成输出信号的小部分，即颗粒。这种合成通常通过对信号段进行相关确定最佳拼接点来实现。这意味着用来形成输出信号的段不依赖于音调周期，于是不需要音调检测的非平凡任务。然而，在

这些方法中仍然有快速改变的信号振幅带来的问题，并且高质量的搬移会引起高的计算需求。然而，现在提出改进的时域音调位移器/搬移器，其中使用瞬态检测和动态系统参数，以低的计算成本对平稳(调性或非调性)和瞬态声音期间都产生了更加精确的高搬移因子的搬移。

参照图示，其中相同的号码指示相同的元件，图4中示出九个分开的模块：瞬态检测器401，窗口位置调节器403，码本产生器405，同步信号选择器407，同步位置存储器409，最小差估计器411，输出段存储器413，混合单元415，以及下降采样器417。输入信号提供给码本产生器405和瞬态检测器401。如果检测到瞬态，则其位置送到窗口位置模块403。当生成码本时这一模块规定了乘以输入信号的窗口的大小和位置。码本产生器405从同步选择模块407接收同步位置，假设它已经连接到另一搬移器。如果这同步位置在码本内，则使用该同步位置，并产生输出段。否则码本被发送到最小差估计器411，该估计器返回一个新的同步位置。新的输出段在混合模块415中与先前的输出段一同起受到窗处理之后，在模块417中被下降采样。

为了使说明清晰，引入状态空间表示。这里，状态向量或颗粒表示输入和输出信号。输入信号由状态向量 $x(n)$ 表示：

$$x(n) = [x(n), x(n-D), x(n-2D), \dots, x(n-(N-1)D)] \quad (5)$$

这是从输入信号的 N 个延迟样本获得的，其中 N 是状态向量的维数，而 D 是用来构成向量的输入样本之间的延迟。颗粒映射产生在每一状态向量 $x(n-1)$ 后的样本 $x(n)$ 。这给出方程式6，其中 $a(*)$ 是映射：

$$x(n) = a(x(n-1)). \quad (6)$$

本方法中，颗粒映射使用状态搬移码本，根据先前的输出确定下一个输出。长度为 L 的码本被连续重构成包含状态向量和每一状态向量后的下一个样本。每一状态向量与其相邻的状态向量相隔 K 个样本；这使得系统能够依靠当前的被处理的信号的特性调节时间分辨率，这里 K 等于一表示最细分辨率。基于可能的瞬态的位置和先前码本中的同步位置选择用来构成码本的输入信号段。

这意味着在理论上对包含在码本中所有的搬移计算映射 $a(\cdot)$:

$$a\left(\begin{bmatrix} x(n-L) \\ x(n-L+K) \\ \vdots \\ x(n-1) \end{bmatrix}\right) = \begin{bmatrix} x(n-L+1) \\ x(n-L+K+1) \\ \vdots \\ x(n) \end{bmatrix}. \quad (7)$$

使用这个转移码本，通过搜索码本中与当前状态向量 $y(n-1)$ 最相似的状态向量，计算新的输出 $y(n)$ 。通过计算最小差进行这个最邻近的状态向量搜索；给出新的输出样本：

$$y(n) = a(y(n-1)). \quad (8)$$

然而系统不限于基于逐个样本的处理，而是最好基于逐段的处理。对新的输出段加窗，再与先前的输出段相加、混合被下降采样。通过由码本表示的输入段长度和从码本读出的输出段长度的比确定音调搬移因子。

返回图示，在图 5 和图 6 的中示出流程图，显示搬移器操作的周期。在 501 表示输入数据，对一段输入信号进行瞬态检测 503；对等于输出段长度的段长度进行瞬态搜索。如果找到瞬态 505，则存储瞬态的位置 507 并调节参数 L (表示码本长度)、 K (表示每一状态向量之间的样本数距离)，及 D (表示每一状态向量中的样本之间的延迟) 509。比较该瞬的位置与先前输出段的位置 511，以确定瞬态是否已经被处理。如果是 513，则调节码本(窗口 L)的位置、及参数 K 、 L 和 D ，515。在必须的参数调节之后，基于瞬态检测的结果，进行搜索新的同步或拼接点 517。这一过程显示在图 6。首先基于先前的同步点计算新的同步点 601，根据：

$$Sync_pos = Sync_pos_old + SM - S, \quad (9)$$

其中 $Sync_pos$ 和 $Sync_pos_old$ 分别是新的和老的同步位置， S 是正在被处理的输入段的长度，而 M 是搬移因子。这一同步点用来对新的拼接点的精确性与老的拼接点的精确性进行比较 603。如果匹配与先

前的相同或更好 605，则这一新的拼接点返回 607，假设它在码本内。如果不是，则在回路 609 搜索新的同步点。这是使用相似性测量进行的，这种情形下是最小差函数，然而也能够使用时域或频域的相关性。如果位置产生比先前找到的 613 位置的匹配更好的匹配，则存储同步点 615。当试验了所有位置 617 时，系统返回 619 图 5 的流程。所获得的新的同步点被存储 519，并从码本 521 在给定的同步点开始读出新的段。对这一段加窗并添加到先前的那一段上 523，通过搬移因子下降采样 525，并存储在输出缓冲器 527。

图 7 中示出在瞬态条件下关于码本位置的系统行为。在瞬态之前，表示输入段 1 的码本 1 配置在段的“左侧”。相关段 1 表示先前输出的部分并用来寻找码本 1 中的同步点 1。当检测到瞬态且对瞬态点处理时，根据图 7a 移动码本并静止到当前正在处理的输入段一旦再次到码本的“右侧”。由于不允许系统在瞬态之前搜索同步点，这使之不能复现瞬态。

许多音调搬移器，或时间扩展器，基于模式搜索预测对语音和单音调材料给出满意的结果。然而，它们的性能对于高复杂性信号（如音乐）迅速劣化，特别是在搬移因子较大时。本发明为改进性能提供了几种解决办法，从而对于任何类型的信号都产生极好的结果。与其它设计不同，该系统是时变的，且系统参数基于输入信号的性质及先前操作周期中所使用的参数。不仅控制码本大小和位置，而且控制所包含的状态向量的性质的瞬态检测器的使用是健壮的和计算上有效的方法，能够避免快速改变信号段期间可听到的质量下降。此外，不需要处理会引起较高计算需求的信号段长度的变化。而且本发明使用基于来自先前搜索的结果的加细的码本搜索。这意味着与普通两信号段相关性，基于模式搜索预测时域系统中通常所作的不同，不是顺序地试验所有位置而是首先试验最可能的同步位置。这新的用于减少码本搜索的方法大大降低了系统的计算复杂性。而且，当使用几个搬移器时，如以下所示的实施那样，能够在搬移器之间共享同步位置信息以进一步降低计算复杂性。

以上所述的时域搬移器，用来实现根据以下说明性的但非限制性例子构成的 SBR-1 和 SBR-2 系统。图 8 中使用了三个时间扩展模块，以

便产生二、三和四次谐波。在这一例子中由于每一时域扩展/搬移处理宽带信号，故不必添加分开的均衡器系统，最好在搬移之前调节源频率范围的频谱包络，这是考虑到在搬移之后这样作是没有意义的。频谱包络调节器 801、803 和 805 每一个处理几个滤波器组频道。必须这样设置包络调节器中每一频道的增益，使得在搬移之后输出的和 813、815 和 817 产生所需的频谱包络。搬移器 807、809 和 811 被互连以便共享同步位置信息。这是基于这样的事实，即在一定的条件下，在分开的搬移单元中相关期间码本中找到的同步位置之间将是高相关的。假如，作为也是不限制本发明范围的一个例子，四次谐波搬移器基于时间帧工作，二次谐波搬移器帧的一半，但是两倍于额定周期。进而假设，对于两个扩展器的码本是相同的，且两个时域扩展器的同步位置分别标以 $sync_pos4$ 和 $sync_pos2$ 。这产生以下的关系：

$$sync_pos2 = sync_pos4 - n \cdot 4 \cdot S - sync_offset, \text{ for } n=1,2,3,4..., \quad (10)$$

其中

$$sync_offset = sync_pos4 - sync_pos2, \text{ for } n=0, \quad (11)$$

且 S 是由码本表示的输入段的长度。只要任何同步位置指针都没有达到码本的结束，这就是有效的。在正常操作期间对每一由二次谐波搬移器处理的时间帧 n 增加一，并当不可避免地由指针之一达到码本末端时，计数器 n 被设置为 $n=0$ ，并分别计算 $sync_pos4$ 和 $sync_pos2$ 。对三次谐波搬移器当连接到四次谐波搬移器时，可获得类似的结果。

以上出现的几个互连的时域搬移器，显著减少了生成较高次谐波的计算。此外，提出的时域搬移器与适当的滤波器组连接使用则提供了调节所生成的频谱包络的机会，同时保持时域搬移器结构简单计算成本低，因为这些，或多或少可以实现使用定点算法和只使用加法/减法操作。

本发明的其它示例性但非限制性例子是：

- 在子带滤波器组中每一子带内使用一个时域搬移器，这样降低对每一搬移器的信号复杂性。

- 一个时域搬移器与一个频域搬移器组合使用, 这样能够使系统按照被处理的输入信号的特性使用不同的搬移方法。

- 在一个宽带语音编解码器中使用一个时域搬移器例如对线性预测后的获得的残留信号进行操作。

应当认识到以上概述的方法只要简单地省略采样速率变换就可以有利用于时标修改型。还可以理解, 虽然所述方法专注于音调搬移到较高音调, 即时间扩展, 但对业内专业人员明显的是, 在搬移到较低的音调即时间压缩时, 能够使用相同的原理。

基于搬移的滤波器组

现在说明基于搬移技术的各种新的和革新的滤波器组。被搬移的信号被分为一系列 BP 或子带信号。任何子带信号被精确或近似地搬移, 这最好通过以下称为“填补”的分析和合成子带的再连接实现。首先使用短时付立叶变换, STFT, 说明该方法。

离散时间信号 $x(n)$ 的 N -点 STFT 由以下公式定义

$$X_k(n) = \sum_{p=-\infty}^{\infty} x(p) h(n-p) e^{-j\omega_k p}, \quad (12)$$

其中 $k=0,1,\dots,N-1$, 及 $\omega_k=2\pi k/N$, 且 $h(n)$ 是窗函数。如果窗满足以下条件

$$\begin{cases} h(0)=1 \\ h(n)=0 \text{ for } n=\pm N, \pm 2N, \pm 3N, \dots \end{cases} \quad (13)$$

则逆变换存在并由以下公式给出

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X_k(n) e^{j\omega_k n}. \quad (14)$$

参见图 9a, 直接变换可以解释为一个分析器, 由一组脉冲响应为 $h(n)\exp(j\omega_k n)$ 的 N 个 BP 滤波器组成, 后跟一组乘以载波 $\exp(-j\omega_k n)$ 的 N 个乘法器, 乘法器把 BP 信号下移到 0Hz 区域附近, 形成 N 个分析信号 $X_k(n)$ 。窗口的作用是作为原型 LP-滤波器。 $X_k(n)$ 具有小的带宽并通常被下降采样。这样只需对方程式 12 在 $n=rR$ 估算, 其中 R 是抽取因子而 r 是新时间变量。参见图 9b, $X_k(n)$ 能够通过上升采

样从 $X_k(rR)$ 恢复, 即插零 907, 后接 LP 滤波 909。逆变换可解释为由一组乘以载波 $(1/N)\exp(j\omega_k n)$ 911 的 N 个乘法器组成的合成器, 乘法器把信号 $X_k(n)$ 上移到它们的原始频率, 后跟图 9c 的级 913, 把来自所有频道的贡献 $y_k(n)$ 相加。可以重新排布 STFT 和 ISTFT 以便使用 DFT 和 IDFT, 这使得能够使用 FFT 算法[“ Implementation of the Phase Vocoder using the Fast Fourier Transform ” M.R.Portnoff, IEEE ASSP, Vol.24, No.3, 1976]。

图 9c 示出用于产生二次谐波的填补 915, $M=2, N=32$ 。为简单起见, 只示出频道 0 到 16。BP 16 的中心频率等于奈奎斯特频率, 频道 17 到 31 对应于负频率。稍后将说明标以 P 的块 917 和增益块 919, 现在只简短考虑。在这一例子中对输入信号加了频带限制, 于是只有频道 0 到 7 包含信号。这样, 分析器 8 到 16 是空的, 不需要被映射到合成器。对应于输入信号延迟通路, 分析器频道 0 到 7 连接到合成器频道 0 到 7。分析频道 k , $4 \leq k \leq 7$, 还连接到合成频道 $Mk, M=2$, 它把信号移动到 BP 滤波器 k 的中心频率两倍的频率区域。这里, 信号被上移到它们的原始范围, 以及被向上搬移一个八度。为了探索以实值滤波器响应和调制器产生谐波, 还必须考虑负频率, 参见图 10a 的下面的分支。因而, 必须估计再映射 $k \rightarrow Mk$ 1001 和 $N-k \rightarrow N-Mk$ 1003 (其中 $4 \leq k \leq 7$) 的组合输出。

这产生了

$$y(n) = \frac{2}{N} [x(n) * h(n) \cos(\omega_k n)] \cos((M-1)\omega_k n) + \frac{2}{N} [x(n) * h(n) \sin(\omega_k n)] \sin((M-1)\omega_k n) \quad (15)$$

其中 $M=2$ 。方程式 15 可以解释为输入信号的 BP 滤波, 后跟线性频率移动或上边带(USB)调制, 即使用上边带的单边带调制, 参见图 10b, 其中 1005 和 1007 形成希尔伯特变换器, 1009 和 1011 为乘以余弦和正弦载波的乘法器, 而 1013 是选择上边带的差级。显然, 可以直接实现这种多带 BP 和 SSB 方法, 即无需在时域或频域中滤波器组填补, 允许任意选择各个通带和振荡器频率。

根据方程式 15, 分析频道 k 的通带内带有频率 ω_i 的正弦波产生频率为 $M\omega_k + (\omega_i - \omega_k)$ 的谐波。因而, 称为基本多带搬移的方法只对频率为

$\omega_i = \omega_k$ ($4 \leq k \leq 7$) 的输入信号产生精确的谐波。然而, 如果滤波器数足够大, 则对精确搬移的偏移可以忽略, 参见方程式 4。而且, 假设每一分析频道包含最大一个分音, 通过插入标以 P 的块 917(图 9c)对任意频率的准稳定调性信号进行精确搬移。这种情形下, $X_k(rR)$ 是频率等于分析滤波器的谐分音频率 ω_i 和中心频率 ω_k 之间的差的复指数。为了获得精确的搬移 M, 必须通过因子 M 使这些频率增加, 修改以上频率关系为 $\omega_i \rightarrow M\omega_k + M(\omega_i + \omega_k) = M\omega_i$ 。 $X_k(rR)$ 的频率等于它们各退缠相位角的时间导数, 并能够使用相继相位角的一阶差分估计。使频率估计乘以 M, 并使用那些新的频率计算合成相位角。然而, 除了相位常数之外, 通过使分析幅角直接乘以 M 的简化方式可获得相同的结果, 这免除了对频率估计的需要。这在表示块 917 的图 11 中描述。这样, $X_k(rR)$, 在本例中 $4 \leq k \leq 7$, 被从直角坐标转换为极坐标, 由块 $R \rightarrow P$, 1101 表示。幅角乘以 $M = 2$ (1103) 而量值不变。然后信号被转换回直角坐标 ($P \rightarrow R$) 1105, 形成信号 $Y_{MK}(rR)$ 并根据图 9c 供给分析频道。这样这一改进的多带搬移方法具有两级: 填补作为以基本方法提供了粗搬移, 而相位乘法器提供了细致的频率校正。以上多带搬移方法不同于使用 STFT 的传统的音调移动技术, 其中为了合成使用查找表振荡器, 或者当为了合成使用 ISTFT 时, 对信号进行时间伸展和抽取, 即不使用填补。

易于对非 2 的搬移因子修改图 9c 的谐波填补。图 12 示出为产生 3 次谐波的填补 1203, 其中 1201 是分析频道, 1205 是合成频道。如图 13 所示, 可以同时生成一些不同次的谐波, 其中使用了 2 和 3 次谐波。图 14 表示 2、3 和 4 次谐波的非重叠组合。使用频率尽可能高的最低次可能谐波。在谐波 M 的目标范围上限以上, 使用谐波 M+1。图 15 示出映射所有合成器频道的方法(所示 $N=64$, 频道 0-32)。根据源和目标频道号之间的以下关系映射带有非素数下标的所有高频带频道: $k_{\text{dest}} = Mk_{\text{source}}$, 其中 M 是满足 k_{source} 在低频带而 k_{dest} 在高频带这样的条件的最小整数 ≥ 2 。因而, 任何一个合成器频道不会从多于一个的分析频道接收信号。素数高频带频道可能被映射到 $k_{\text{source}} = 1$, 或者低频带频道 $k_{\text{source}} > 1$, 这产生以上关系的良好的近似(图 15 中只示出带有 $M=2, 3, 4, 5$ 的非素数连接)。

还能够从不同的分析器频道组合振幅和相位信息。振幅信号 $|X_k(rR)|$ 可根据图 16 连接, 而相位信号 $\arg\{X_k(rR)\}$ 根据图 16 的原理连接。这样, 低频带频率将仍然被搬移, 从而产生源区域包络的周期性重复, 而不是从根据图 2 的搬移得出的延伸包络。可以采用选通或其它措施, 以避免“空”源频道的放大。图 17 示出另一应用, 即使用从较高到较低子带的连接对高通滤波的或低音受限信号产生分谐波。当使用以上装置时, 可能有利于采用基于信号特性的修补自适应切换。

在以上说明中, 假设了包含在输入信号中的最高频率显著低于奈奎斯特频率。这样, 能够进行带宽扩展而无需增加采样率。然而, 这并不总是为什么先期上升采样也可能是必要的情形。当对搬移使用滤波器组方法时, 能够在过程中集成上升采样。

在时间到频率的映射中大多数感觉编解码器采用最大抽取滤波器组[“ Introduction to Perceptual Coding ” K. Brandenburg, AES, Collected Paper on Digital Audio Biterate, 1996]。图 18 示出感觉编解码器系统的基本结构。分析滤波器组 1801 把输入信号分为几个子带信号。使用减少的比特数对子带采样分别量化 1803, 其中从估计最小掩蔽阈值的感测模型 1807 确定量化电平数。子带采样被归一化后, 以可选冗余度编码方法编码, 并与由归一化因子、比特分配信息和其它编解码器特有的数据组成的副信息组合 1805, 以形成串行比特流。这个比特流然后被存储或传输。在解码器中, 图 18B, 编码的比特流被多路分解 1809, 解码, 且子带采样被重新量化为相等的比特数 1811。合成滤波器组组合子带采样, 以便重新生成原始信号 1813。使用最大抽取滤波器组的实现将大大降低计算成本。在以下说明, 专注于余弦调制滤波器组。然而应当理解, 本发明可以使用其它类型的滤波器组或变换实现, 包括小波变换的滤波器组解释程序, 其它不等带宽滤波器组或变换, 及多维滤波器组或变换。

在以下解释性的但非限制性的说明中, 假设 L 频道余弦调制的滤波器组把输入信号 $x(n)$ 分解为 L 子带信号。最大抽取滤波器组的一般结构示于图 19 中。分析滤波器 1901 标记为 $H_k(z)$, 其中 $k=0,1,\dots,L-1$ 。子带信号 $v_k(n)$ 被最大抽取, 每一采样频率 f_s/L , 其中 f_s 是 $x(n)$ 的采样频率。

在内插 1905 和滤波 1907 之后合成部分重新组合子带信号以产生 $\hat{x}(n)$ 。合成滤波器标记为 $F_k(z)$ 。此外，本发明对 $\hat{x}(n)$ 进行频谱复现，给出增强信号 $y(n)$ 。

以 QL 频道滤波器组合成子带信号结果将是带有采样频率 Qf_s 的输出比特流，其中只使用 L 个低频带频道，并选择带宽扩展因子 Q 使得 QL 为一整数值。因而，延伸的滤波器组的作用如同后跟一个上升采样器的 L 频道滤波器组。这种情形下，由于不使用 $L(Q-1)$ 个高频带滤波器（以零供给），故音频带宽将不变 - 滤波器组将只是重构 $\hat{x}(n)$ 的上升采样版本。然而如果对高频带滤波器填补 L 个子带信号，则 $\hat{x}(n)$ 的带宽将增加因子 Q，产生 $y(n)$ 。根据本发明，这是基本多带搬移器的最大抽取滤波器组版本。使用这一方案，如同先前的说明，上升采样过程被集成在合成滤波中。应当注意，可以使用任何大小的合成滤波器组，其结果是输出信号不同的采样率，从而不同的带宽扩展因子。根据带有整数搬移因子 M 的基本多带搬移方法的本发明，对 $\hat{x}(n)$ 进行频谱复现通过如下填补子带信号实现

$$v_{Mk}(n) = e_{Mk}(n) (-1)^{(M-1)kn} v_k(n), \quad (16)$$

其中 $k \in [0, L-1]$ 并选择成使 $Mk \in [L, QL-1]$ ， $e_{Mk}(n)$ 是包络校正，而 $(-1)^{(M-1)kn}$ 是对倒谱带的校正因子。频谱颠倒这是由于子带信号抽取造成的，且颠倒的信号可以通过在这些频道中每隔一次采样时改变符号而被重新颠倒回来。参见图 20，考虑对搬移因子 $M=2$ 以 $Q=2$ 填补 2009 的 16 频道合成滤波器组。块 2002 和 2003 分别标记图 19 的分析滤波器 $H_k(z)$ 和抽取器。类似地，2005 和 2007 是内插器及合成滤波器 $F_k(z)$ 。这时方程式 16 简化为把接收的数据的四个上频率子带信号填补为合成滤波器组中八个最高频道的每隔一个的频道。由于频谱颠倒，每隔一个填补的子带信号必须是合成之前被颠倒的频率。此外，填补的信号的量值必须根据 SBR-1 或 SBR-2 原理调节。

使用根据本发明的基本多频带搬移方法，所产生的谐波一般不是基频精确的倍数。每一子带中除了最低以外的所有频率都在某种程度上不同于精确的搬移。而且，由于目标区间覆盖了比源区间更宽的频率范

围，故复现的频谱包含一些零。此外，由于在目标区间中在频率上子带信号被分开，故余弦调制的滤波器组的混叠抵销性质消失。即相邻的子带信号在高频带区域不重叠。然而，业内专业人员所知道的混叠降低的方法可用来减少这种类型的人工痕迹。这种搬移方法的优点在于易于实现，以及很低的计算成本。

为了实现完美的正弦波搬移，现在提出改进的多带搬移方法的有效最大抽取滤波器组解决方案。该系统使用增加的修改的分析滤波器组，而合成滤波器组是受如 Vaidyanathan 所描述的余弦调制为 [“Multirate Systems and Filter Banks” p.p. Vaidyanathan, Prentice Hall, Engewood Cliffs, New Jersey, 1993, ISBN 0-13-605718-7]。使用根据本发明基于最大抽取滤波器组改进的多带搬移方法的操作步骤，简略示于图 21 及图 22 的流程图中如下：

1. 使用 QL 频道滤波器组 2101、2201、2203 合成 L 个收到的子带信号，其中对 $L(Q-1)$ 个上频道供给零，以便形成信号 $x_1(n)$ ，于是通过带宽扩展因子 Q 对其过采样。

2. 通过因子 Q 对 $x_1(n)$ 下降采样，以便形成信号 $x_2(n')$ 2103、2205，即 $x_2(n') = x_1(Qn')$ 。

3. 选择整数值 K 作为合成滤波器组的大小，限制其使得 $T=KM/Q$ 为一整数，其中 T 是修改的分析滤波器组的大小，而 M 是搬移因子 2207、2209、2211。最好应这样选择 K，对于平稳(调性)信号为大，对于动态(瞬态)信号为较小。

4. $x_2(n')$ 是通过 T 频道修改的分析滤波器组 2107、2213 滤波的，其中对 T 个分析滤波器作幂调制，产生一组复值子带信号。按因子 T/M 对子带信号进行下降采样，给出子带信号 $v_k^{(M)}(n'')$ ， $k=0,1,\dots,T-1$ 。从而对滤波器组将按因子 M 过采样。

5. 把样本 $v_k^{(M)}(n'')$ 转换为极坐标表示(幅值和相角)。相角乘以因子 M，并根据图 11 所示的方案样本被转换回直角坐标。提取复值样本的实部，给出信号 $s_k^{(M)}(n'')$ 2109、2215。在这一操作之后，对信号 $s_k^{(M)}(n'')$ 作临界采样。

6. 根据 SBR-1 或 SBR-2 的原理调节 $s_k^{(M)}(n'')$ 的增益。

7.使用普通余弦调制的 K 频道滤波器组合成子带信号 $s_k^{(M)}(n'')$, 这里 $k \in [T/M, \min(K, T)-1]$, 其中对频道 0 到 $T/M-1$ 供给零 2105、2221。这产生信号 $x_3^{(M)}(n)$ 。

8.最后把 $x_3^{(M)}(n)$ 加到 $x_1(n)$ 给出 $y(n)$ 2223, 这就是所需的频谱复现信号。

对不同的搬移因子 M 值可以重复步骤 3 到 6, 这样把多个谐波添加到 $x_1(n)$ 。这种操作模式由图 21 的虚线表示, 并在图 22 中通过重复框 2211 - 2219 上的回路表示。这种情形下, 选择 K 使 T 对于整数值 M : s 的 M 的所有选择为整数值; K 最好这样选择, 使 K/Q 为正整数。所有子带信号 $s_k^{(M_i)}(n'')$, $i=1, 2, \dots, m$, m 是搬移因子数, 根据

$$s_k(n'') = \sum_{i=1}^m s_k^{(M_i)}(n'') \quad (17)$$

对每一可用的 k 被相加。在图 22 的回路第一叠代中, 可以把 $s_k(n'')$ 看作只是零的子带样本, 这里 $k=0, 1, \dots, K-1$ 。在每次循环, 新的样本被添加 2219 到 $s_k(n'')$, 有

$$s_k(n'') = s_k(n'') + s_k^{(M_i)}(n''), \quad (18)$$

这里 $k=K/Q, K/Q+1, \dots, \min(K, T_i)-1$ 。使用 K 频道滤波器组根据步骤 7 一次合成子带信号 $s_k(n'')$ 。

步骤 4 的修改的分析滤波器组, 是通过余弦调制滤波器组的理论推导得出的, 其中调制叠搭变换[“Lapped Transform for Efficient Transform/Subband Coding” H.S Malvar, IEEE Trans ASSP, Vol.38, no.6, 1990]是一种特别的情形。 T 频道余弦调制滤波器组中的滤波器的脉冲响应 $h_k(n)$ 可以写为

$$h_k(n) = C p_0(n) \cos \left[\frac{\pi}{2T} (2k+1) \left(n - \frac{N-1}{2} \right) + \Phi_k \right], \quad (19)$$

这里 $k=0, 1, \dots, T-1$, N 是低通原型滤波器 $p_k(n)$ 的长度, C 是常数, 而 Φ_k 是保证相邻频道之间混叠消除的相位角。对 Φ_k 的限制是

$$\Phi_0 = \pm \frac{\pi}{4}, \Phi_{T-1} = \pm \frac{\pi}{4} \text{ and } \Phi_k = \Phi_{k-1} \pm \frac{\pi}{2} \quad (20a-c)$$

这可以简化为封闭形式的表达式

$$\Phi_k = \pm(-1)^k \frac{\pi}{4}. \quad (21)$$

由 Φ_k 的这一选择,使用带有如下脉冲响应的合成滤波器组,可以获得完美的重构系统或近似的重构系统(准 QMF 系统)

$$f_k(n) = C p_0(n) \cos \left[\frac{\pi}{2T} (2k+1)(n - \frac{N-1}{2}) - \Phi_k \right]. \quad (22)$$

考虑滤波器

$$h'_k(n) = C p_0(n) \sin \left[\frac{\pi}{2T} (2k+1)(n - \frac{N-1}{2}) + \Phi_k \right], \quad (23)$$

其中 $h'_k(n)$ 是原型滤波器 $p_0(n)$ 的正弦调制版本。滤波器 $h'_k(n)$ 和 $h_k(n)$ 具有等同的通带支持,但是相位响应不同。滤波器的通带实际上是彼此的希尔伯特变换(对于接近 $\omega = 0$ 和 $\omega = \pi$ 的频率这是无效的)。根据

$$h_k^a(n) = h_k(n) + j h'_k(n) = C p_0(n) \exp \left[\frac{j\pi}{2T} (2k+1)(n - \frac{N-1}{2}) + j\Phi_k \right] \quad (24)$$

组合方程式 19 和 23,产生具有对正频率与 $H_k(z)$ 相同的幅值响应,而对负频率为零的滤波器。使用带有如方程式 24 中的脉冲响应的滤波器组,给出可以解释为对应于从带有如方程式 19 中脉冲响应的滤波器组获得的子带信号的解析(复)信号的一组子带信号。解析信号适用于操作,因为复值样本可以写成极坐标的形式,即 $z(n) = r(n) + j i(n) = |z(n)| \exp \{j \arg(z(n))\}$ 。然而,当对于搬移使用复滤波器组时,对 Φ_k 的限制必须扩展到保持混叠消除性质。为了保证在与带有如方程式 22 中的脉冲响应的合成滤波器组的组合中的混叠消除,对 Φ_k 新的限制是

$$\Phi_k = \pm(-1)^k \frac{\pi}{4M} \quad (25)$$

当 $M=1$ 时这简化为方程式 21。使用这一选择,被搬移的分音将具

有与 $M=1$ (不搬移)时它们具有的相同的相对相位。

组合方程式 24 和方程式 25 结果为

$$h_k^a(n) = C p_0(n) \exp \left\{ j\pi \left[\frac{(2k+1)}{2T} \left(n - \frac{N-1}{2} \right) \pm \frac{(-1)^k}{4M} \right] \right\} \quad (26)$$

根据本发明这是在步骤 4 的修改的滤波器组中使用的滤波器。

一些涉及到步骤 5 的澄清: 按因子 Y/M 对复值子带信号的下降采样使它们被按 M 过采样, 当相位角相继乘以搬移因子 M 时, 这是一重要准则。在搬移到目标范围后, 过采样迫使每带宽的子带样本数等于源范围的子带样本数。由于相位乘法器, 被搬移的子带信号各带宽为 M 倍源范围的带宽。在使得子带信号在步骤 5 之后被临界采样, 并且, 当搬移调性信号时, 频谱中将没有零。

为了避免三角计算, 即必须计算新的子带信号如

$$\begin{aligned} s_k^{(M)}(n'') &= \text{real} \left\{ \left| v_k^{(M)}(n'') \right| \exp \left\{ jM \arctan \left(\frac{\text{imag}\{v_k^{(M)}(n'')\}}{\text{real}\{v_k^{(M)}(n'')\}} \right) \right\} \right\} = \\ &= \left| v_k^{(M)}(n'') \right| \cos \left\{ M \arctan \left(\frac{\text{imag}\{v_k^{(M)}(n'')\}}{\text{real}\{v_k^{(M)}(n'')\}} \right) \right\}. \end{aligned} \quad (27)$$

其中 $|v_k^{(M)}(n'')|$ 是 $v_k^{(M)}(n'')$ 的绝对值, 使用以下的三角关系式:

$$\cos(M\alpha) = \cos^M(\alpha) - \binom{M}{2} \sin^2(\alpha) \cos^{M-2}(\alpha) + \binom{M}{4} \sin^4(\alpha) \cos^{M-4}(\alpha) - \dots \quad (28)$$

设

$$\alpha = \arctan \left(\frac{\text{imag}\{v_k^{(M)}(n'')\}}{\text{real}\{v_k^{(M)}(n'')\}} \right), \quad (29)$$

并注意到

$$\cos(\alpha) = \cos \left(\arctan \left(\frac{\text{imag}\{v_k^{(M)}(n'')\}}{\text{real}\{v_k^{(M)}(n'')\}} \right) \right) = \frac{\text{real}\{v_k^{(M)}(n'')\}}{|v_k^{(M)}(n'')|}. \quad (30)$$

以及

$$\sin(\alpha) = \sin\left(\arctan\left(\frac{\text{imag}\{v_k^{(M)}(n^*)\}}{\text{real}\{v_k^{(M)}(n^*)\}}\right)\right) = \frac{\text{imag}\{v_k^{(M)}(n^*)\}}{|v_k^{(M)}(n^*)|}, \quad (31)$$

无需三角计算即可实现步骤 5 的计算，降低了降低复杂性。

当使用 M 为偶数的搬移时，取决于低通原型滤波器 $p_0(n)$ 的特性，可能引起相位乘法器的障碍。所有可使用的原型滤波器在 Z 平面中单位圆上有零点。单位圆上的一个零点迫使滤波器的相位响应出现 180° 的移动。对于 M 为偶数，相位乘法器把这些移动转变为 360° 的移动，即相移消失。位于这种相移消失的频率中的分音将引起合成信号中的混叠。最不好的情形是当分音在频率上位于对应于分析滤波器的第一旁瓣的顶部时。取决于幅值响应中这一旁瓣的抑制，混叠将或多或少能够听到。作为一例，ISO/MPEG 层 1 和 2 标准中使用的原型滤波器的第一旁瓣被抑制 96 分贝，而对于 ISO/MPEG 层 3 标准的 MDCT 方案中使用正弦窗口的第一旁瓣抑制仅为 23 分贝。很明显，使用正弦窗口的这种类型的混叠将是可听到的。将提出对这问题的解决方案，并被称为相对锁相。

滤波器 $h_k^a(n)$ 都具有线性相位响应。相位角 Φ_k 导致相邻频道之间的相对相位差，且单位圆上的零点导致频率中频道之间可能不同的位置的 180° 的相移。通过监视相邻子带信号之间的相位差，在相位乘法器被激活之前，易于检测到包含相位反相信息的频道。考虑调性信号，根据方程式 25 对于非反向信号相位差近似于 $\pi/2M$ ，因而对于两种信号之一反相的信号近似于 $\pi(1-1/2M)$ 。可以简单地通过相邻子带中样本的点积完成反相信号的检测，如

$$v_k^{(M)}(n^*) \circ v_{k+1}^{(M)}(n^*) = \text{real}\{v_k^{(M)}(n^*)\}\text{real}\{v_{k+1}^{(M)}(n^*)\} + \text{imag}\{v_k^{(M)}(n^*)\}\text{imag}\{v_{k+1}^{(M)}(n^*)\}. \quad (32)$$

如果方程式 32 中的乘积是负的，则相位差大于 90° ，于是出现相位反相状态。根据步骤 5 的方案，复值子带信号的相位角乘以 M ，并最后，标为反相的信号被反相。这种相对锁相方法迫使在相位乘法之后 180° 移动的子带信号保持这一移动，因而保持了混叠消除的性质。

谱包络调整

许多声音，象语音和音乐之类，被刻划为慢变包络和固定振幅的快

变载波的乘积, 如 Stockam 在[“ The Application of Generalized Linearity to Automatic Gain Control” T.G. Stockham, Jr, IEEE Trans. on Audio and Electroacoustics, Vol.AU-16, No.2, June 1968]及方程式 1 所述。

在分裂频带感觉音频编码器中, 把音频信号分段为帧, 并使用子带滤波器或时域到频域变换分为多个频带。在许多类型的编解码器中, 信号随之被分为用于传输或存储的两个主信号成分, 即频谱包络表示和归一化子带样本或系数。在以下的全部说明中, 术语“子带样本”或“系数”是指从子带滤波器获得的样本以及从时域到频域变换获得的系数。术语“频谱包络”或“标度因子”表示基于时间帧的子带的值, 诸如每一子带中平均或最大的幅值, 用于子带样本的归一化。然而, 频谱包络还可以使用线性预测 LPC 获得[U.S.Pat.5,684,920]。在典型的编码器中, 与可以用降低许多的速率编码(使用大约 10 % 的可用比特率)的慢变临时包络, 即频谱包络比较, 归一化子带样本需要以高比特率编码(使用大约 90 % 的可用比特率)。

如果要保持原始信号的音色的质量, 则被复现的带宽的精确的频谱包络是重要的。乐器或噪音的感觉音色主要由位于听觉最高八度中频率 f_{lim} 以下的频谱分布决定的。这样在 f_{lim} 以上的频谱细节不很重要, 于是通过以上搬移方法获得的高频带细结构不需要调节, 而粗结构一般是需要的。为了能够进行这种调节, 对信号的频谱表示滤波, 以便从细结构分离出包络粗结构, 则是有用的。

在根据本发明的 SBR-1 实现中, 是从解码器可得的低频带信息中估计高频带粗频谱包络。这种估计是通过连续监视低频带包络并根据特定的规则调节高频带频谱包络进行的。实现包络估计的一种新方法使用了对数频率幅值空间中的渐近线, 这等价于线性空间中以变阶多项式的曲线拟合。估计低频带频谱的上部的电平和倾斜, 并把估计用来定义表示新的高频带包络的一个或几个段的电平或倾斜。渐近线的交叉点在频率中固定, 其作用是作为枢点。然而并不总是需要的, 规定保持高频带包络在实际边界内的摆幅的限制是有益的。估计频谱包络的另一种方法是使用大量代表性的频谱包络的向量量化 VQ, 并在速查表或码本中存储

它们。向量量化是通过对大量的训练数据训练所需数量的向量进行的，这种情形下数据是音频频谱包络。训练通常使用推广的 Lloyd 算法进行 [“ Vector Quantization and Signal Compression ” A.Gersho, R.M.Gray Kluwer Academi Publisher, USA 1992 ISBN 0-7923-9181-0], 并产生最优覆盖训练数据内容的向量。考虑由 B 个包络训练的 A 个频谱包络组成的 VQ 码本($B \gg A$), 这时基于声音广泛变化的 B 个观察, A 个包络表示从低频带包络到高频带包络 A 个最可能的转变。这就是理论上基于 B 个观察预测包络的 A 个最优化规则。当估计新的高频带频谱包络时, 原来的低频带包络用来搜索码本, 且最佳匹配码本项的高频带部分用来生成新的高频带频谱。

图 23 中, 归一化子带样本由 2301 表示, 而频谱包络由标度因子 2305 表示。为了说明的目的, 到解码器 2303 的传输以并行的形式表示。在 SBR-2 的方法图 24 中, 根据图 23 产生并传输频谱包络信息, 从而只有低频带子带样本被传输。这样被传输的标度因子在全频率范围取值, 而子带样本只在排除高频带的限制的频率范围内取值。在解码器处, 低频带子带样本 2401 被搬移 2403, 并与收到的高频带频谱包络信息 2405 组合。这样, 在保持显著的比特率降低的同时, 合成的高频带频谱包络与原来的等同。

在某些编解码器中, 如图 24 所示, 能够在省略高频带子带样本的同时传输整个频谱包络的标度因子。其它的编解码器标准规定, 标度因子和子带样本必须覆盖相同的频率范围, 即如果子带样本被省略则标度因子不能传输。这种情形下, 有几种解决方案; 高频带频谱包络信息能够以独立的帧传输, 这些帧有自己的标题及可选的出错保护, 后跟数据。没有采纳本发明的优点的普通的解码器, 将不能识别标题, 从而摒弃这些额外的帧。在第二个解决方案中, 高频带频谱包络信息作为被编码的比特流内的辅助数据传输。然而, 可得的辅助数据字段必须足够大以保持包络信息。在头两个解决方案都不适合的情形下, 可以使用第三个解决方案, 将高频带频谱包络信息作为子带样本隐藏。子带标度因子覆盖了大的动态范围, 通常超过 100 分贝。这样能够把任意数目的子带标度因子(图 25 中的 2505)设置为非常低的值, “伪装”为子带样本,

传输这些高频带标度因子, 2501。这种把高频带标度因子传输到解码器 2503 的方式保证了与比特流句法的兼容性。因而, 任意的数据都能够以这种方式传输。有一种相关的方法, 把信息编码为子带样本流 [U.S.Pat.5,687,191]。在码本系统使用霍夫曼或其它冗余编码 2603 时, 可使用图 26 的第四个解决方案。这时对高频带的子带样本被设置为零 2601 或固定值, 以便达到高冗余度。

瞬态响应的改进

涉及人工痕迹的瞬态是音频编解码器中的共同问题, 且类似的人工痕迹出现在本发明中。一般来说, 填补产生对应于时域前或后回声的频谱“零点”或凹口, 即在“真实”的瞬态之前或之后的假的瞬态。尽管 P-块对慢变调性信号“填入零点”, 但仍有前和后回声。改进的多带方法是要处理离散正弦波的, 其中正弦波的数目限制为每子带一个。可以把子带中的瞬态或噪声看作是該子带中大量的离散正弦波。这产生互调失真。这些人工痕迹被认为是瞬态区间期间与复现的高频带频道连接的附加噪声源。于是为避免感觉音频编码器中前和后回声人工痕迹的传统方法, 例如自适应窗口切换, 在此可用来提高改进的多带方法的主观质量。通过使用由编解码器或独立的检测器提供的瞬态检测并降低瞬态条件下的频道数, 迫使“量化噪声”不超过与时间有关的掩蔽阈值。在瞬态通过期间使用较少数目的频道, 而在调性通过期间使用较大数目的频道。这种自适应窗口切换在编解码器中是通常使用的, 以便用频率分辨率换取时间分辨率。在滤波器组大小固定的应用中可以使用不同的方法。一个方法是通过频域中的线性预测使“量化噪声”在时间中成形。然后对作为线性预测滤波器输出的残余信号进行搬移。然后, 把逆向预测滤波器同时施加到原始和频谱复现的频道。另一方法采用压缩扩展器系统, 即在搬移或编码之前瞬态信号的动态振幅压缩, 以及在搬移之后的互补扩展。还能够以信号相关的方式在搬移方法之间切换, 例如对平稳信号使用高分辨率滤波器搬移方法, 而对瞬态信号使用时变模式搜索预测方法。

实际的实现

使用标准的信号处理器或高性能 PC, 可进行 SBR 强化的编解码器

的实时操作。还可以在通常的芯片上对 SBR 强化编解码器进行硬编码。还可以使用任意的编解码器在各种类系统中实现对模拟或数字信号的存储和传输,见图 27 和图 28。SBR-1 方法可以集成在解码器中或作为附加的硬件或软件后处理模块提供。SBR-2 方法需要对编码器进行另外的修改。图 27 中,模拟输入信号提供给 A/D 转换器 2701,形成供给进行源编码的任意编码器 2703 的数字信号。提供给系统的信号可以是那种低通类型的,其听觉范围内的一些频带已经被丢弃,或者在任意编码器中一些频带被丢弃。得到的低频带信号被提供给多路复用器 2705,形成被传输或存储的串行比特流。多路分解器 2709 恢复信号并把它们提供给任意解码器 2711。在解码器 2711 中,估计出频谱包络信息 2715 并提供给 SBR-1 单元 2713,该单元把低频带信号搬移为高频带信号并生成包络被调节的宽带信号。最后,数字宽带信号被转换 2717 为模拟输出信号。

SBR-2 方法需要对编码器作另外的修改。图 28 中,模拟输入信号提供给 A/D 转换器 2801,形成供给进行源编码的任意编码器 2803 的数字信号。抽取频谱包络信息 2805。所得信号,低频带子带样本或系数及宽带包络信息,提供给多路复用器 2807,形成被传输或存储的串行比特流 2809。多路分解器 2811 恢复这些信号,低频带子带样本或系数及宽带包络信息,并把它们提供给任意解码器 2815。频谱包络信息 2813 从多路分解器 2811 供给 SBR-2 单元 2817,该单元搬移低频带信号为高频带信号并生成包络经调节的宽带信号。最后,数字宽带信号被转换 2819 为模拟输出信号。

当只能使用很低的比特率时,(因特网及慢速电话调制解调器, AM-广播。)音频节目材料的单编码是不可避免的。为了改进收听质量并使节目声音更好听,通过引入抽头延迟线 2901 可获得简易的“准立体声”产生器,见图 29。除了原来的单信号这可以向每一输出频道提供大约-6 分贝 2903 的 10ms 和 15ms 延迟信号 2905。准立体声产生器以低的计算成本提供了有价值的可感知的改进。

以上的实施例只是例示本发明的原理用于音频源编码改进的情况。应当理解,这里所述的结构和细节的修改和改型对于业内其它专业

人员将是显然的。因此，本发明的专利保护范围仅由所附权利要求规定，而不局限于这里实施例的描述和解释方式所呈的特定细节。

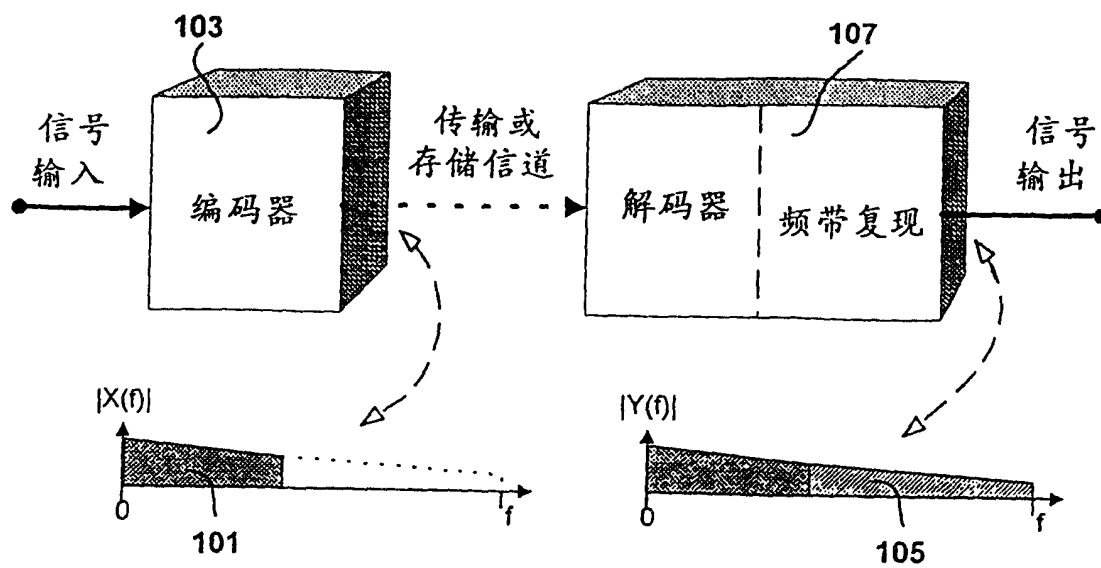


图 1

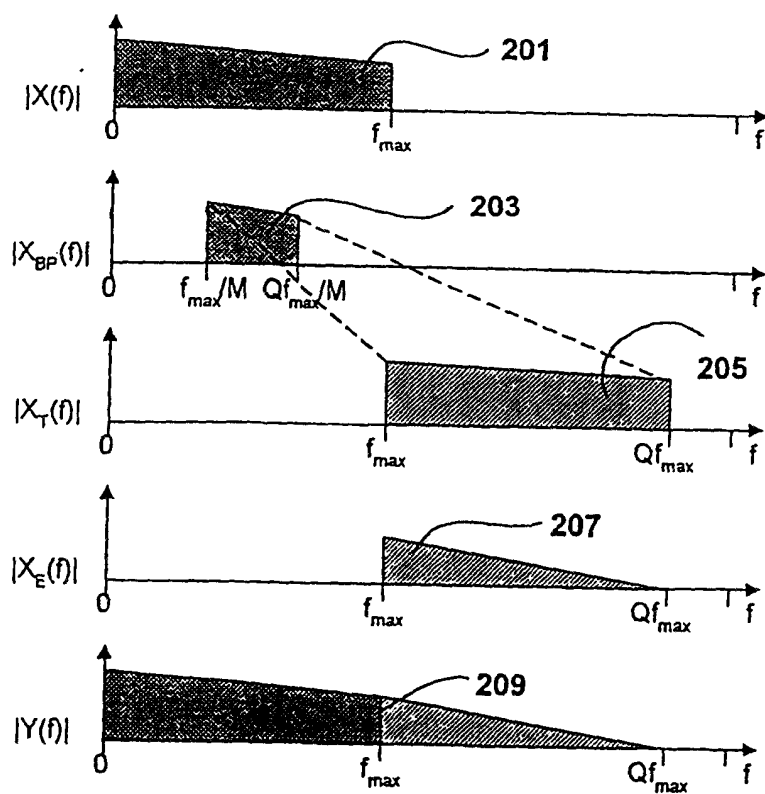


图 2

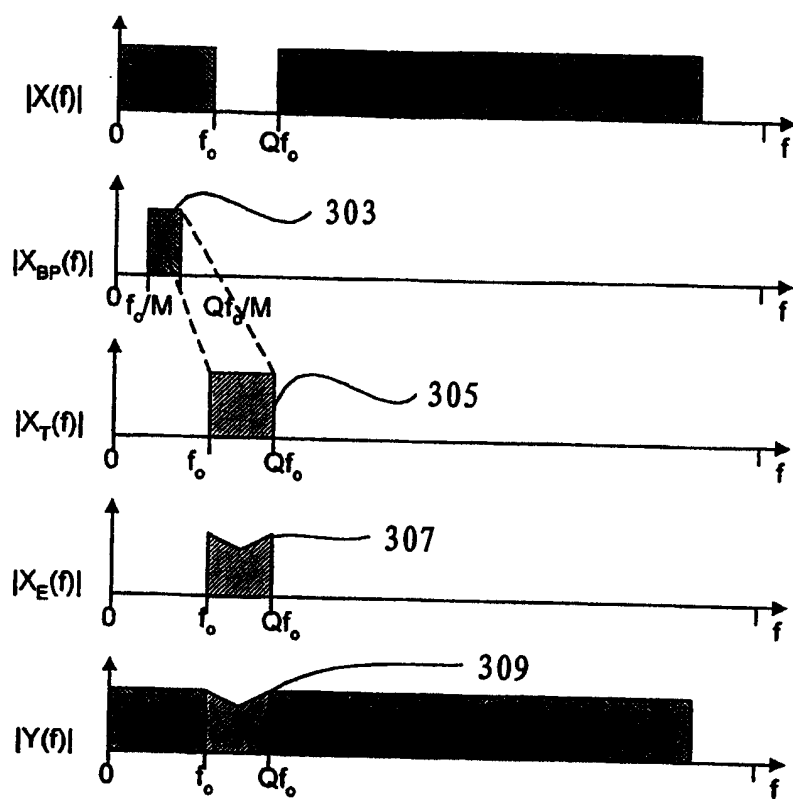


图 3

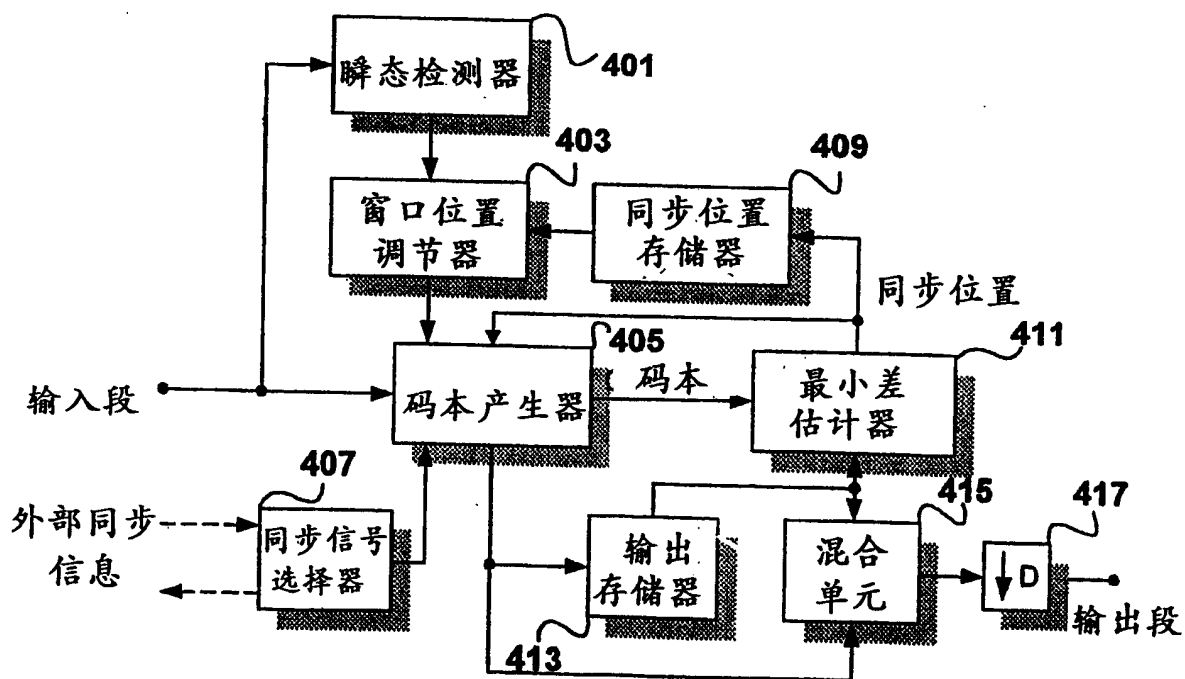


图 4

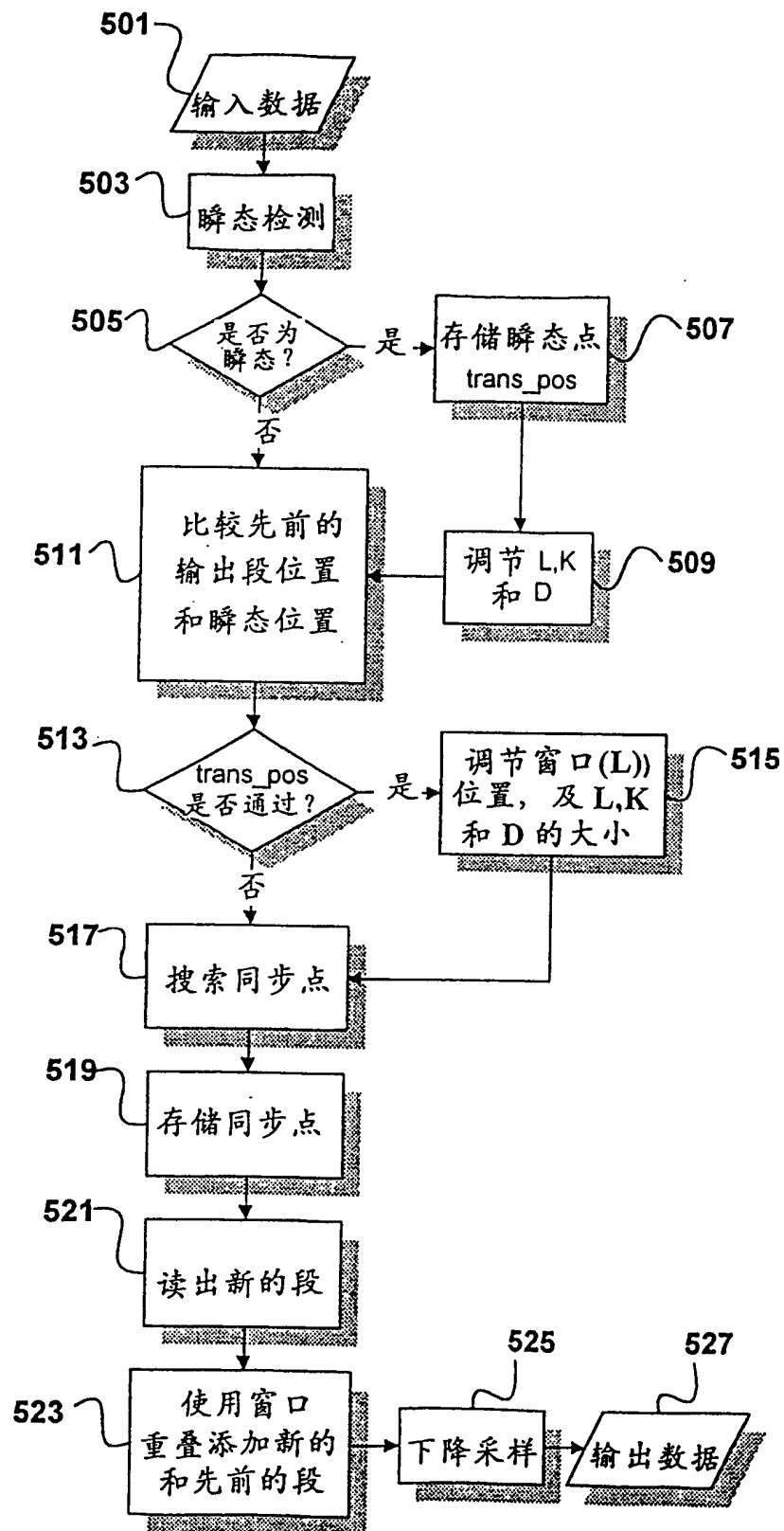


图 5

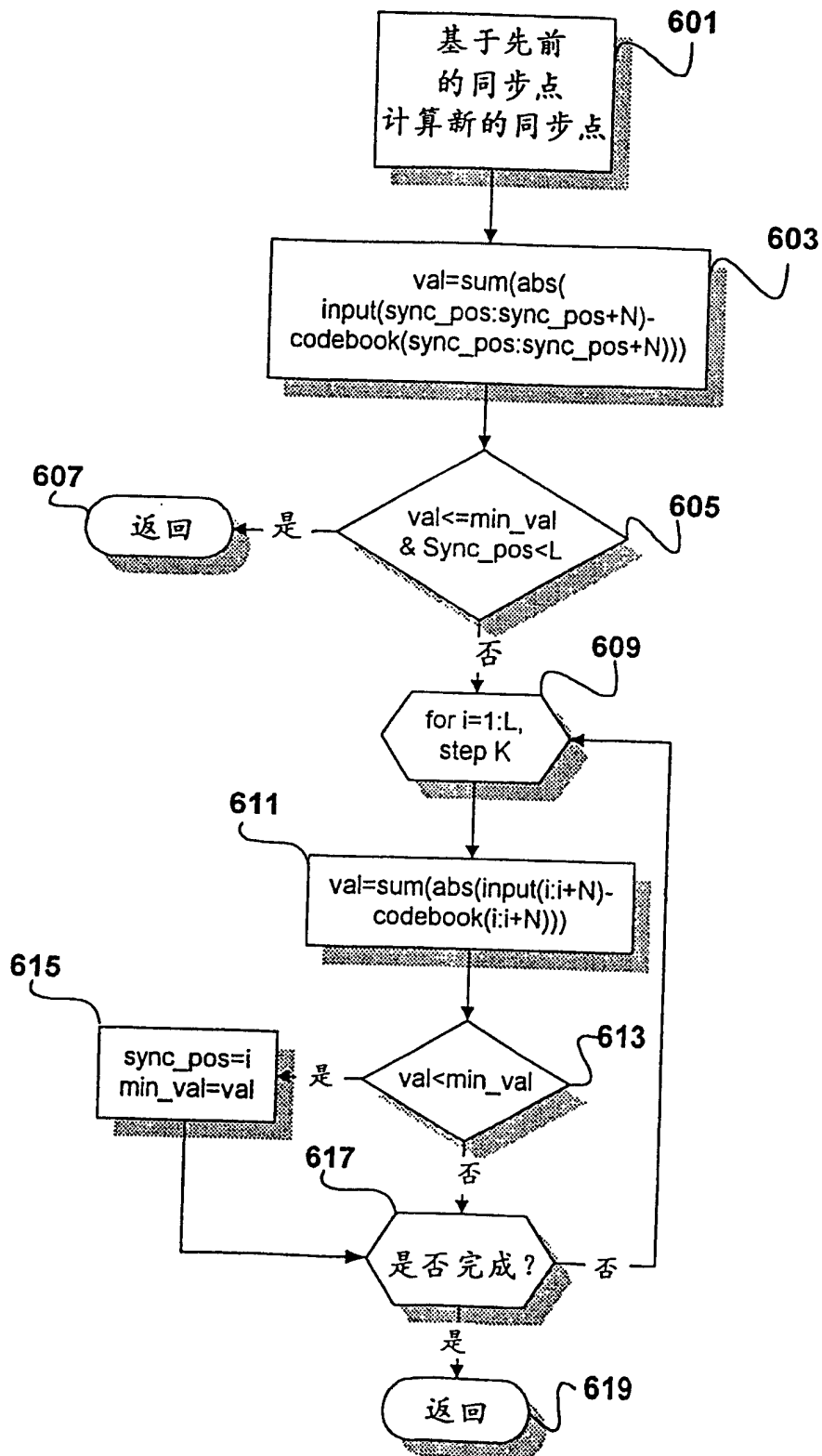


图 6

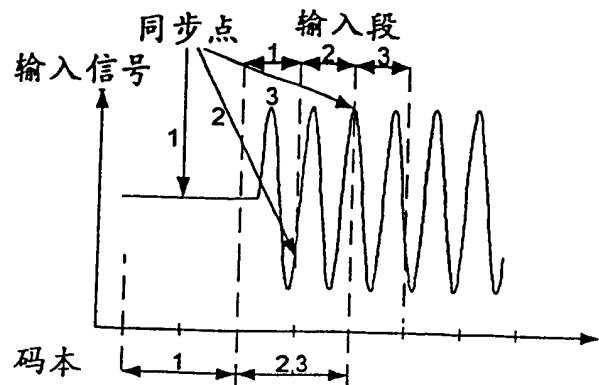


图 7a

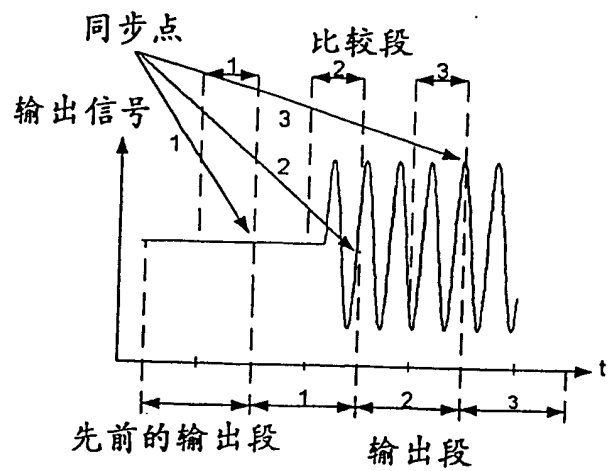


图 7b

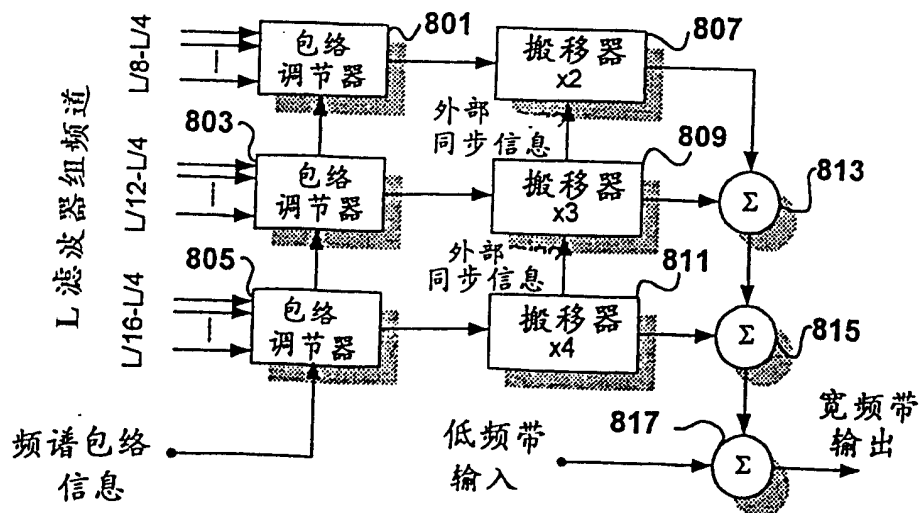


图 8

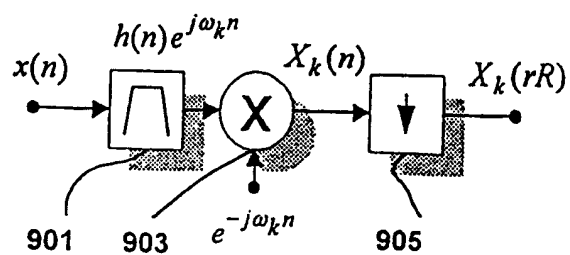


图 9a

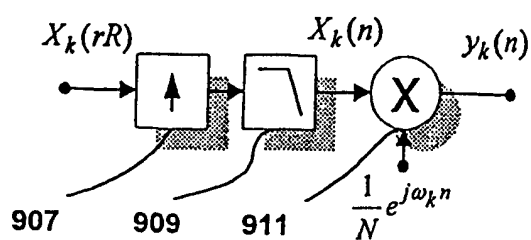


图 9b

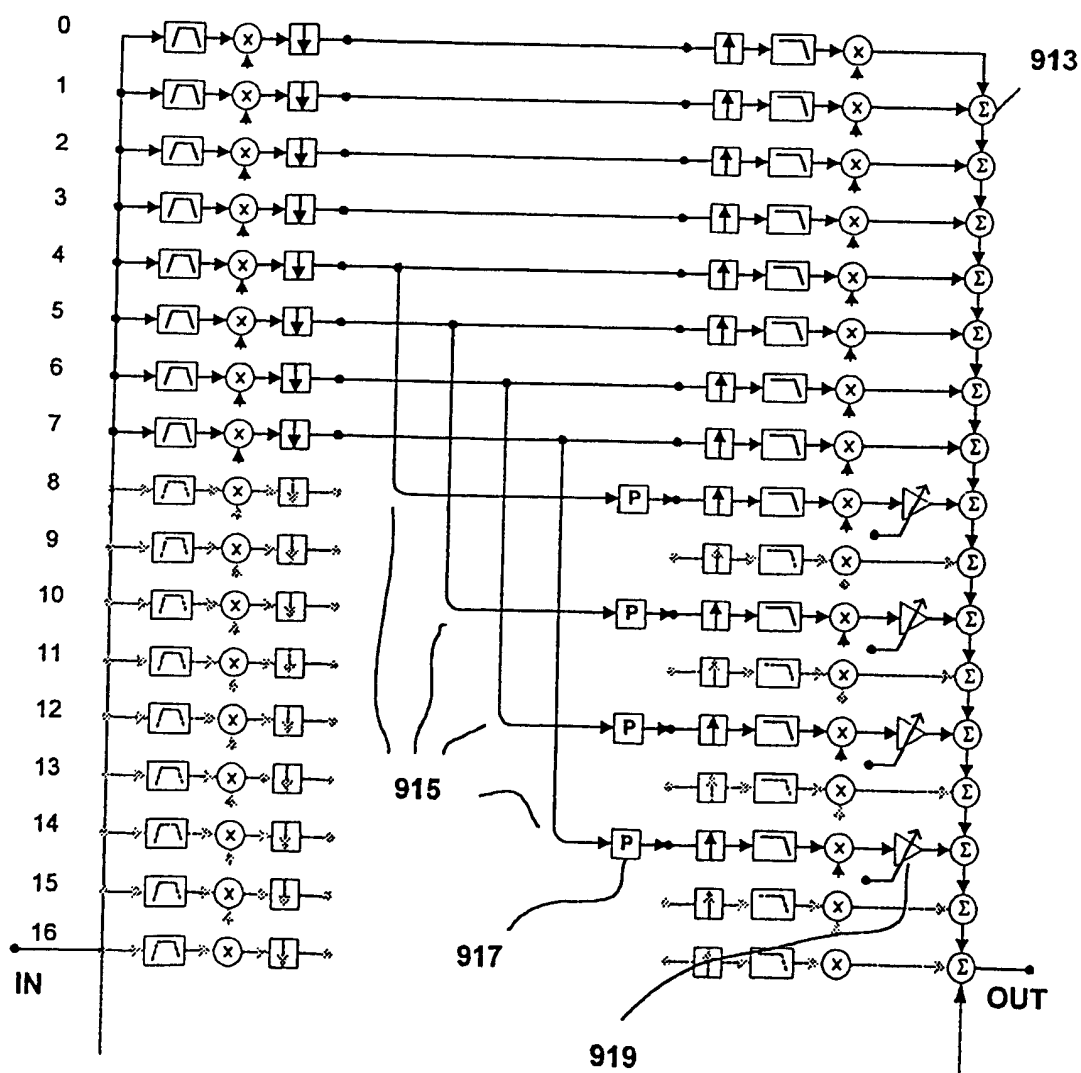


图 9c

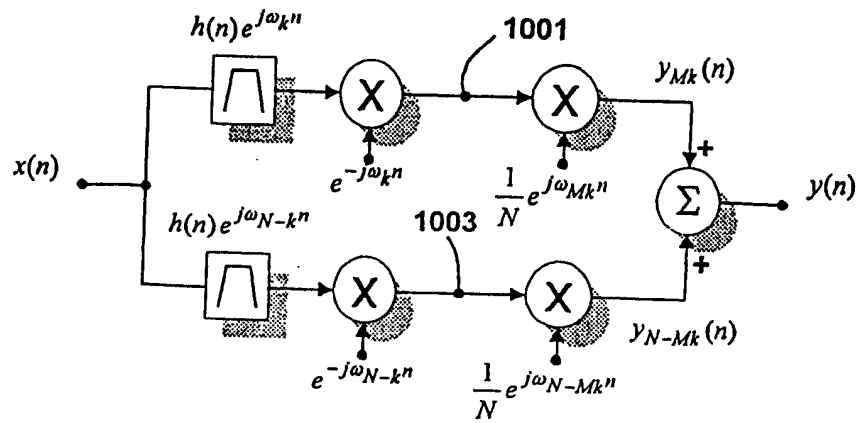


图 10a

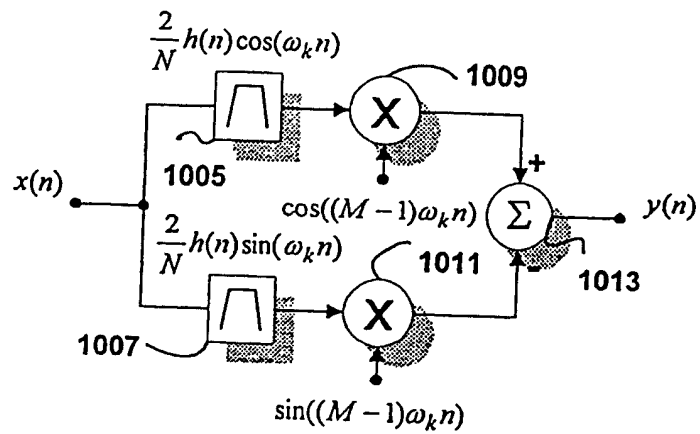


图 10b

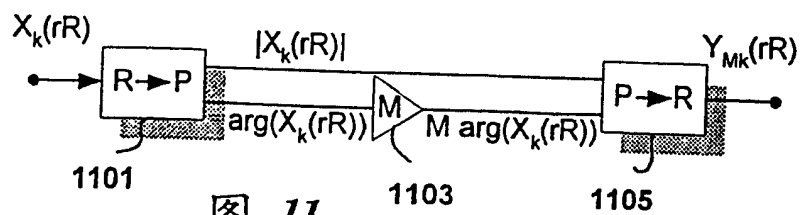


图 11

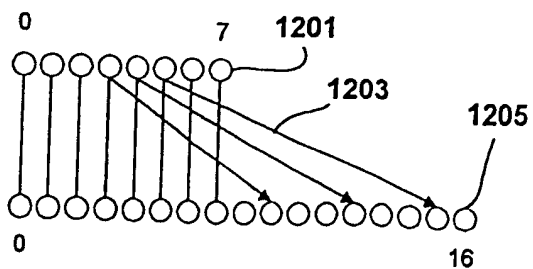


图 12

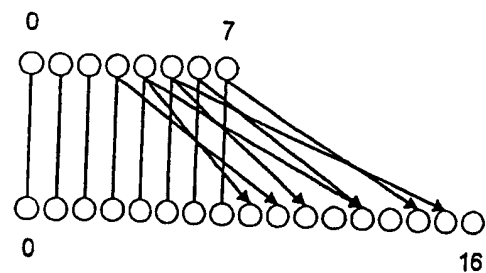


图 13

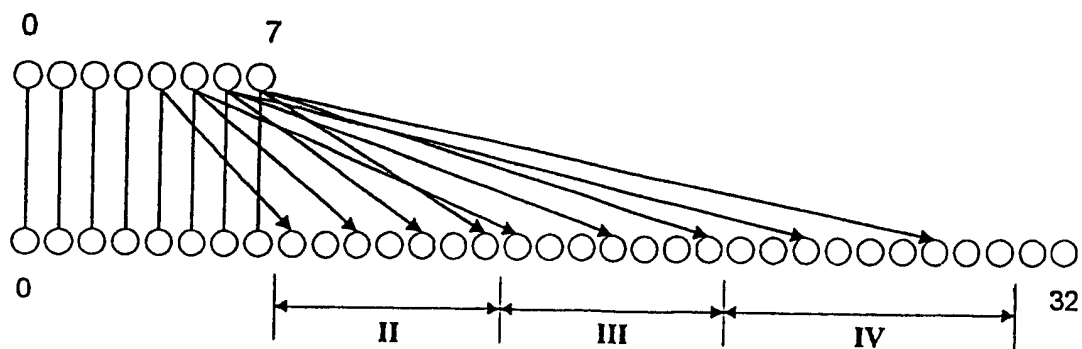


图 14

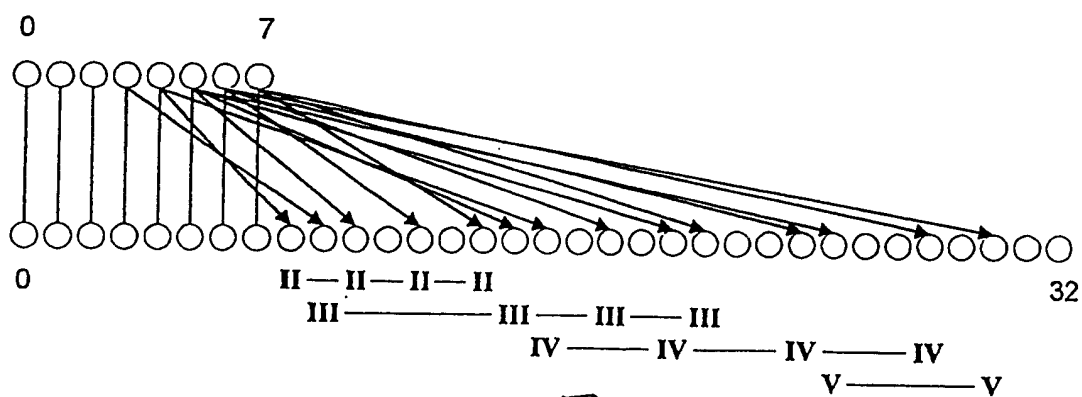


图 15

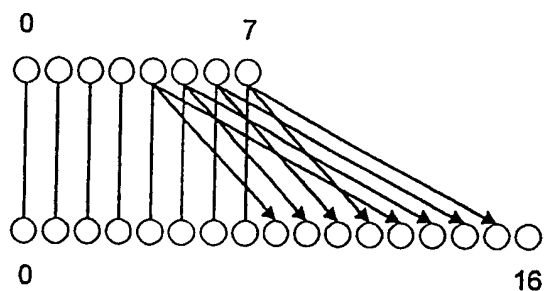


图 16

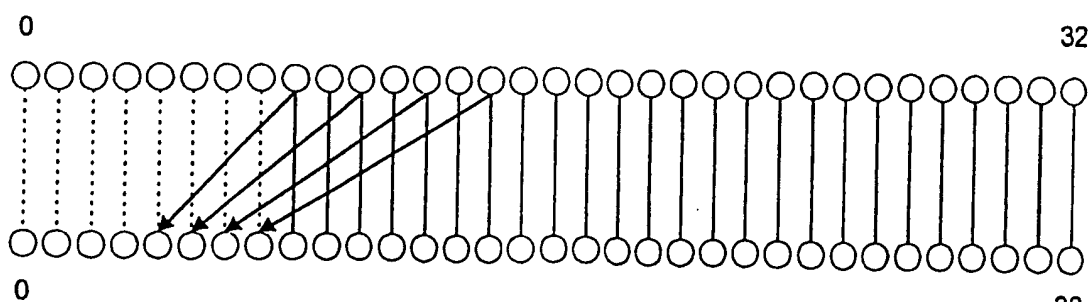


图 17

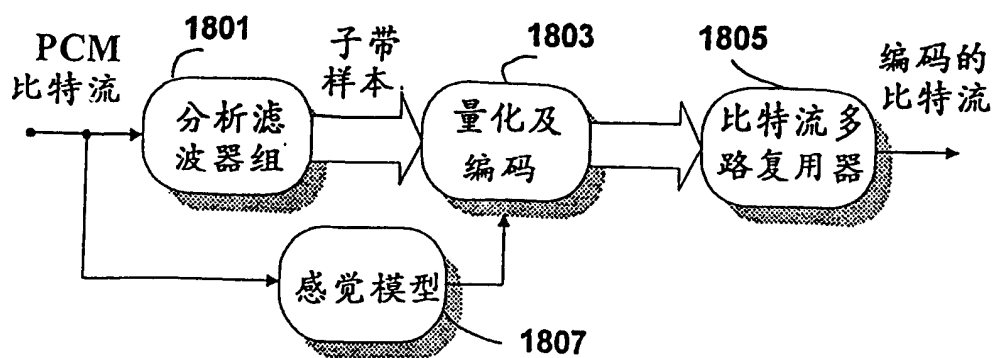


图 18a

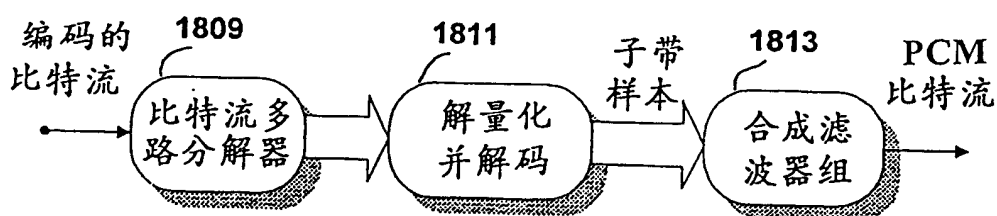


图 18b

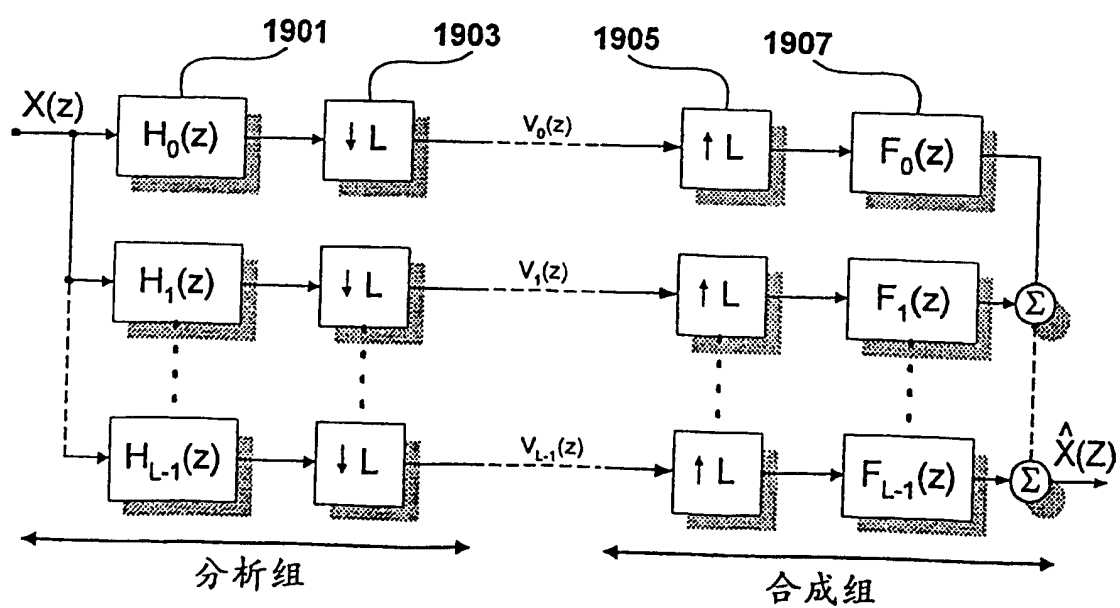


图 19

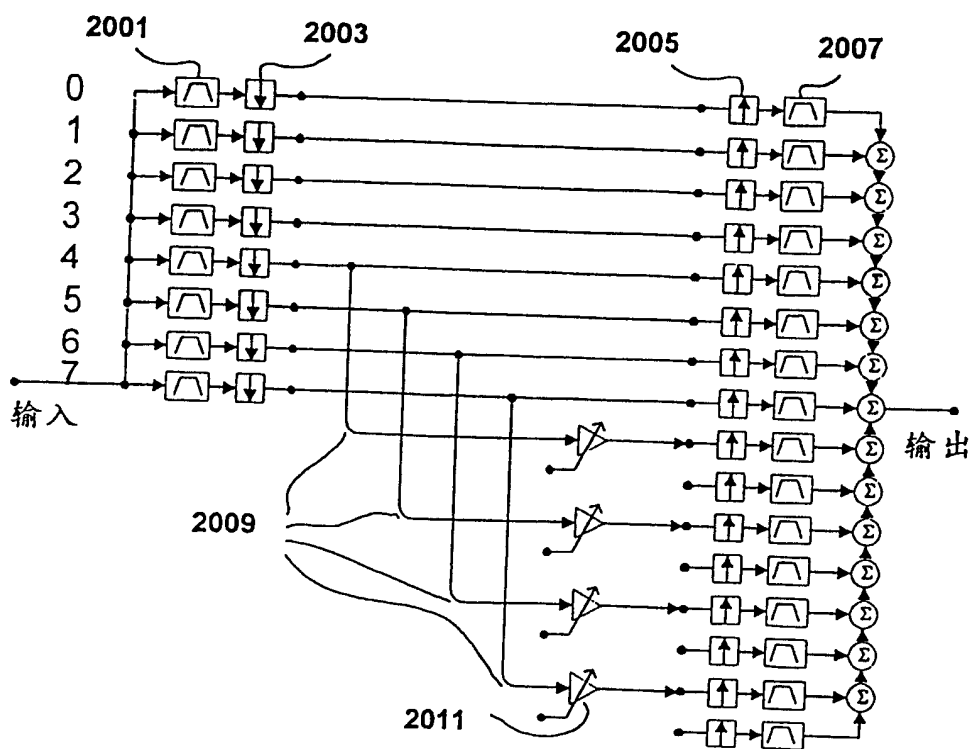


图 20

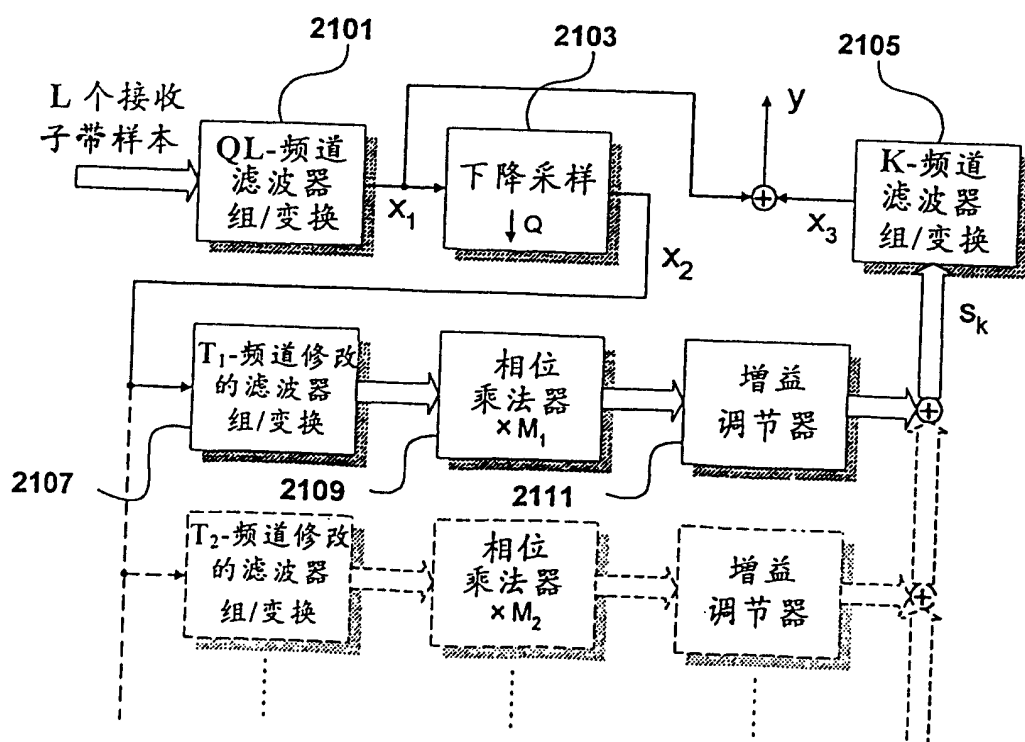


图 21

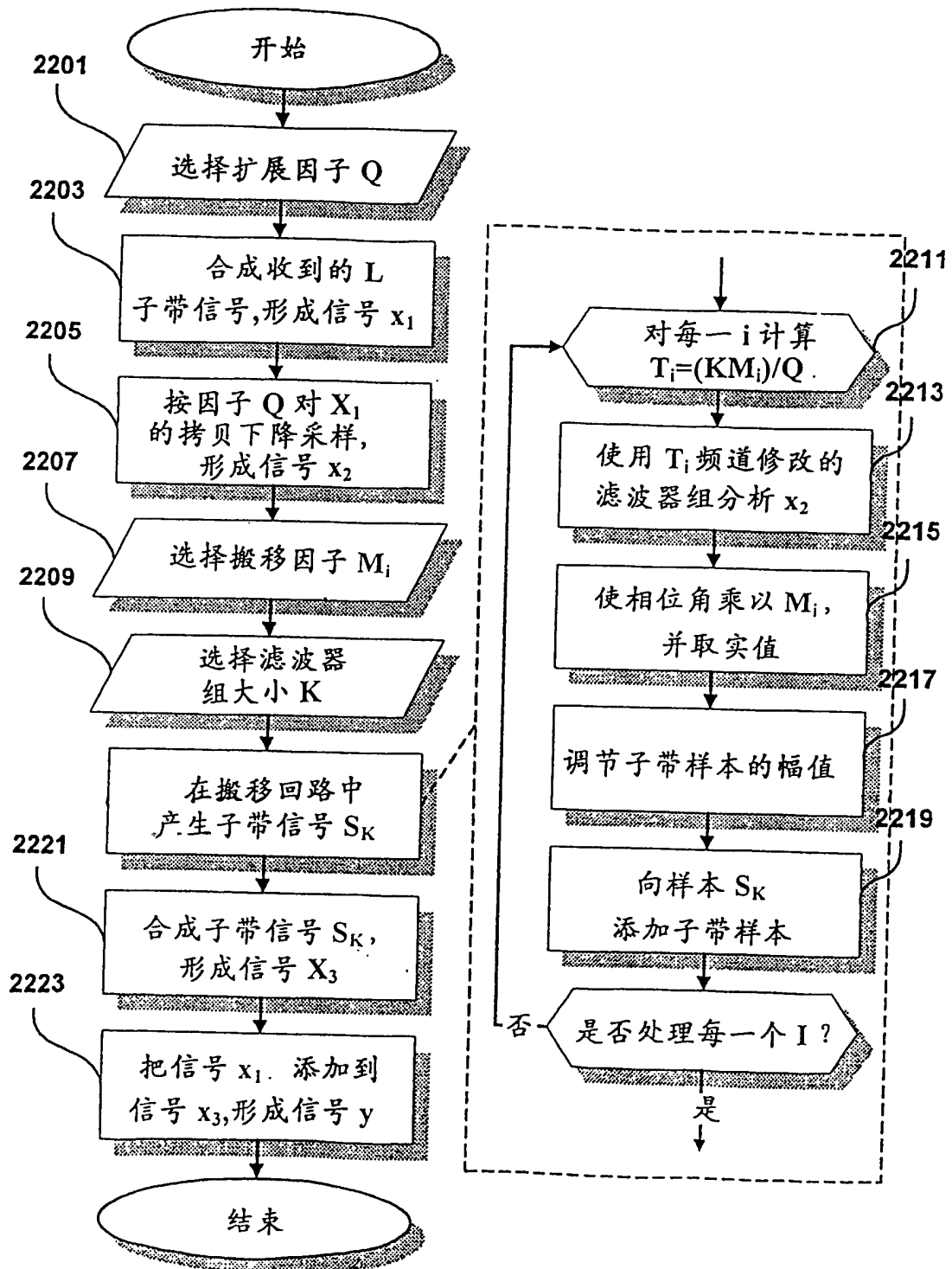
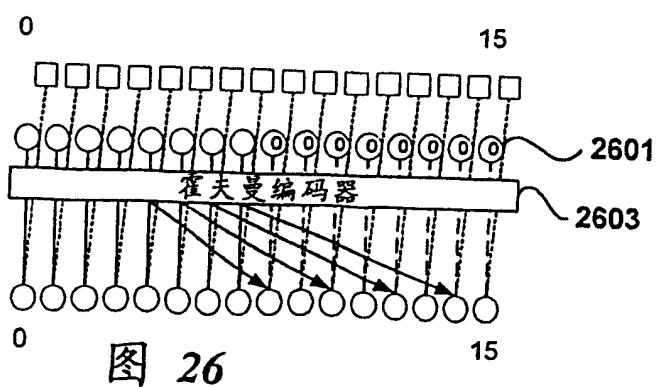
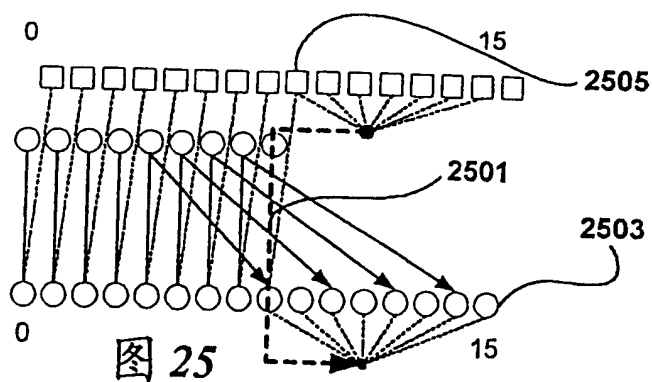
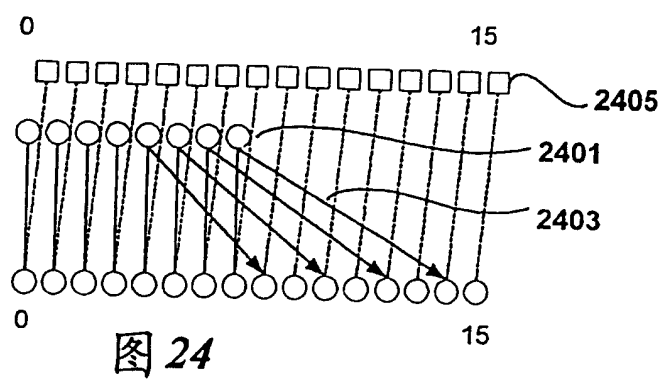
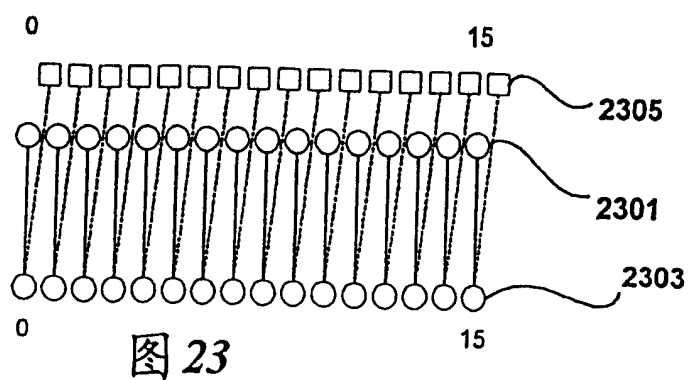


图 22



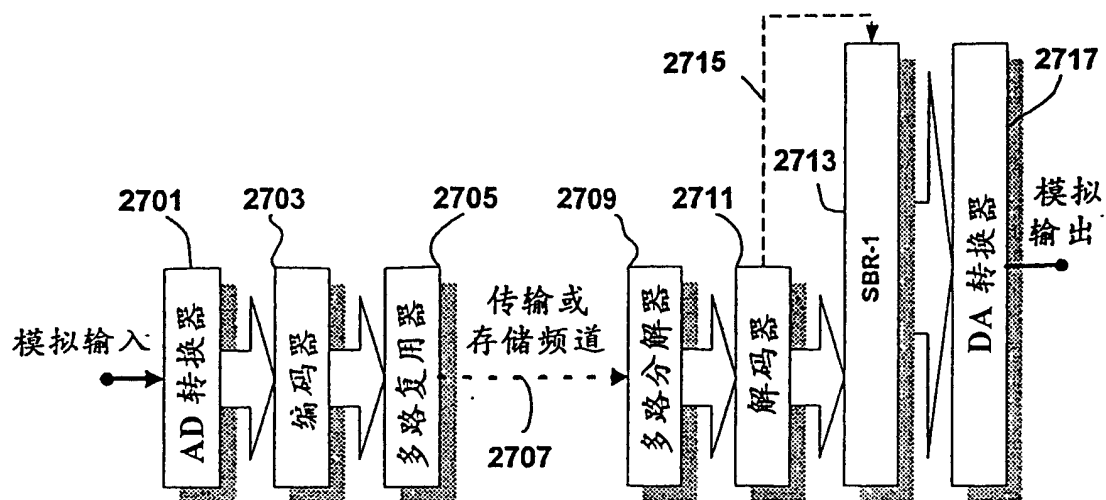


图 27

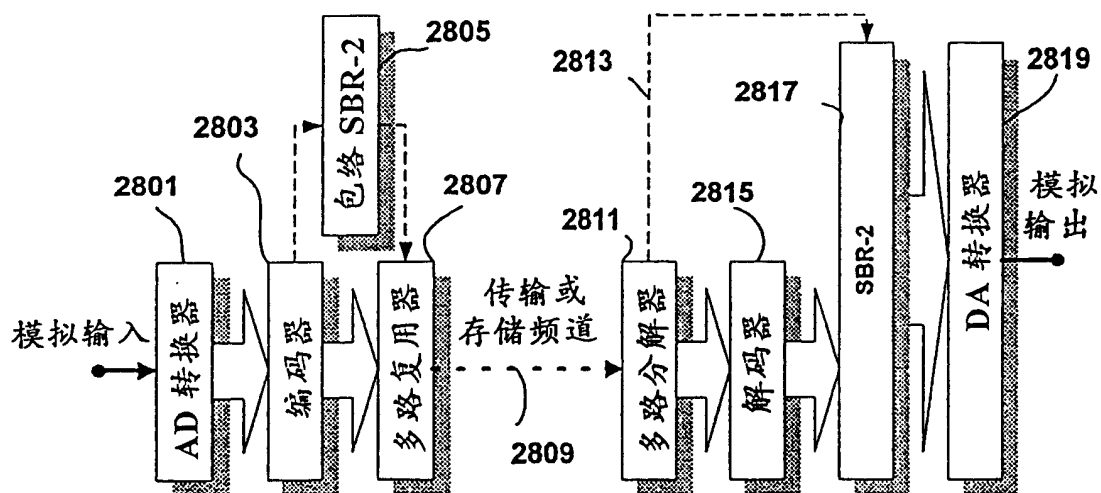


图 28

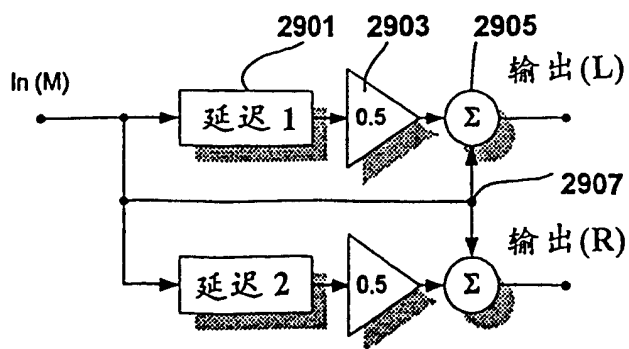


图 29