

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

G10L 19/02

H03M 7/30



[12] 发明专利说明书

专利号 ZL 01818972.5

[45] 授权公告日 2005 年 12 月 21 日

[11] 授权公告号 CN 1232950C

[22] 申请日 2001.11.14 [21] 申请号 01818972.5

[30] 优先权

[32] 2000.11.15 [33] SE [31] 0004187-1

[86] 国际申请 PCT/SE2001/002533 2001.11.14

[87] 国际公布 WO2002/041302 英 2002.5.23

[85] 进入国家阶段日期 2003.5.15

[71] 专利权人 编码技术股份公司

地址 瑞典斯德哥尔摩

[72] 发明人 弗雷德里克·翰

安德烈亚斯·埃雷特 迈克·舒格

审查员 刘亚斌

[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

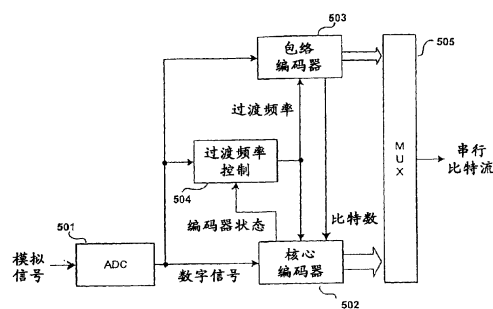
代理人 陆丽英

权利要求书 3 页 说明书 10 页 附图 4 页

[54] 发明名称 性能增强的使用高频重建方法的编码系统和方法

[57] 摘要

本发明涉及使用高频重建(HFR)方法的数字音频编码系统。它教导了如何通过核心编解码器编码的低频带和HFR系统编码的高频带之间的过渡频率的时间上的一个匹配来改进此类系统的总体性能。建立过渡频率瞬时最佳选择的各种方法被介绍。



ISSN 1008-4274

1. 一种用于编码音频信号以获得一种编码音频信号的装置，该编码音频信号由一个解码器使用，所述的解码器具有一个高频重建模块，用以对高于一个过渡频率的一个频率范围执行高频重建，其特征在于，所述装置包括：

一个核心编码器(502)，用于对直到所述过渡频率的所述音频信号的一个较低频带进行编码，其中所述的过渡频率是可变的，且其中所述的核心编码器以帧为基础在一帧上操作；和

一个过渡频率控制模块(504)，根据由所述核心编码器(502)编码所述音频信号的难度和/或根据所述音频信号的音调的和噪声的频率范围之间的边界，来估计由所述核心编码器(502)为一个当前帧所选择的过渡频率，因此所述的过渡频率在时间上是自适应地被选择的。

2. 根据权利要求1所述的装置，其特征在于，所述的难度是以所述音频信号的知觉熵为基础的。

3. 根据权利要求1所述的装置，其特征在于，所述的难度是以在用所述核心编码器编码之后的一个失真能量为基础的。

4. 根据权利要求1所述的装置，其特征在于，所述的难度是以在与所述核心编码器相关的一个比特储藏状态为基础的。

5. 根据权利要求1所述的装置，其特征在于，知觉熵、核心编码器失真能量和核心编码器比特储藏状态的任意组合被用来获得由所述核心编码器(502)为一个帧所选择的过渡频率。

6. 一种用于编码音频信号以获得一种编码音频信号的方法，该编码音频信号由一个解码器使用，所述的解码器具有一个高频重建模块，用于对高于一个过渡频率的频率范围执行高频重建，其特征在于，所述的方法包括以下步骤：

对于直到一个过渡频率的所述音频信号的一个较低频带进行核心编码，其中所述的过渡频率是可变的，并且其中所述的核心编码是以帧的基础在一帧上发生；和

根据在所述的核心编码步骤中对于音频信号编码的难度和/或根据所述音频信号的音调的和噪声的频率范围之间的边界,来估计在所述核心编码步骤中为一个当前帧所选择的一个过渡频率,因此所述的过渡频率在时间上是自适应地被选择的。

7.一种用于解码一种编码音频信号的装置,所述的编码音频信号已经利用一个可变过渡频率而被编码,所述的编码音频信号包括:关于在时间上自适应地变化的一个过渡频率的信息,其特征在于,所述的用于解码的装置包括:

一个比特流去复用器(701),用于将一个比特流分离成为核心编码解码器数据和包络数据;

一个核心解码器(702),用于接收来自所述比特流去复用器的所述核心编码解码器数据和用于输出低频带数据;

一个高频再生包络检波器,用于接收来自所述比特流去复用器(701)中的所述包络数据和用于产生一个频谱包络输出;

一个置换模块(705),用于根据关于所述可变过渡频率的信息来从所述低频带数据中产生一个复制的高频带信号;

一个增益控制模块(704),响应所述高频再生包络检波器,用于调整所述复制的高频带信号到所述高频再生包络检波器输出的一个频谱包络以获得一个包络调整了的高频带信号;和

一个加法器,用于把经过延迟了的所述低频带数据和所述包络调整了的高频带信号相加以获得一种数字宽带信号。

8.一种用于解码一种编码音频信号所述的方法,所述的编码音频信号已经利用一个可变过渡频率而被编码,所述的编码音频信号包括:关于在时间上自适应变化的一个过渡频率的信息,其特征在于,所述的用于解码的方法包括以下步骤:

将一个比特流分离成为核心编码解码器数据和包络数据(701);

利用一个核心解码器(702)接收来自所述比特流去复用器的核心编码解码器数据和输出低频带数据;

利用一个高频再生包络检波器接收所述包络数据和产生一个频

谱包络输出;

利用一个置换模块(705), 根据关于所述可变过渡频率的信息, 从所述的低频带数据中产生一个复制的高频带信号;

响应所述高频再生包络解码器, 利用一个增益控制模块(704), 调整所述复制的高频带信号到所述高频率再生包络解码器输出的一个频谱包络, 以获得一个包络调整了的高频带信号; 和

将经过延迟了的所述低频带数据和所述包络调整了的高频带信号相加, 以获得一个数字宽带信号。

性能增强的使用高频重建方法的编码系统和方法

技术领域

本发明涉及使用高频重建(HFR)方法的数字音频编码系统。它实现一个更相容的核心编解码性能，并且获得组合的核心编解码器和HFR系统的改良音频质量。

背景技术

音频信源编码技术可以被分成两类：正常的音频编码和语音编码。正常的音频编码通常以介质比特率来用于音乐或随机信号。语音编解码器基本上被局限于语音再现，但是另一方面能够以非常低的比特率而被使用。在两类中，信号通常被分开成两个主要信号成分，一个频谱包络和一个相应的残余信号。使用这种划分的编解码器利用了频谱包络可以比残余信号更加有效地被编码的这一事实。在使用高频重建方法的那些系统中，没有与高频带相应的残余信号被发射。替代为：一个高频带在解码器处从被核心编解码器覆盖的低频带中产生并且被整形以便获得期望的高频带频谱包络。在双端 HFR系统中，与较高频率范围相应的包络数据被发射，而在单端 HFR系统中高频带包络从低频带中得到。在任一情况中，现有技术的音频编解码器在核心编解码器频率范围和 HFR频率范围之间应用一个时间恒定的过渡频率。因此，在一个给定比特率处，过渡频率被选择如此以使对于典型程序材料获得在核心编解码器引入的人工产物和 HFR系统引入的人工产物之间的一个优良交换。清楚地，这样一个静态设置可能远离一个特定信号的最佳值：核心编解码器或者被过载，导致比必需的低频带人工产物更高，HFR方法固有的此现象也衰落了高频带质量；或者不被用于它的整个位势，即，比必需的 HFR频率范围更大的一个被使用。因此，只是偶而通过现有技术系统达到联合编码系统的最大性能。此

外，对准具有不同频谱性质的区域（例如声音的和类似噪声的区域）之间的转换交叉的可能性未被利用。

发明内容

本发明提供一种新方法和一种设备用于改善使用高频重建方法(HFR)的编码系统。通过在由低频带编解码器和 HFR系统分别引入的人工产物之间产生最佳值折衷的过渡频率的继续估计和应用，本发明从使用传统编码方案(例如 MPEG层 3或 AAC)的低频带和使用 HFR编码方案的高频带之间的一个固定过渡频率的传统使用中分离出来。根据本发明，该选择可以以如下为基础：利用核心编解码器编码一个信号的难度测量、一个短时比特需求检测以及一个频谱音调分析或者它们的任意组合。难点测量可以从知觉熵或者心理学相关的核心编解码器失真中得到。由于最佳选择经常在时间上改变，所以可变化过渡频率的应用导致一个基本上改良的音频质量，这也较少依靠程序材料特性。本发明适用于单端以及双端 HFR系统。

根据本发明的一个方面，这里提供一种用于编码音频信号以获得一种编码音频信号的装置，该编码音频信号由一个解码器使用，所述的解码器具有一个高频重建模块，用以对高于一个过渡频率的一个频率范围执行高频重建，其特征在于，所述装置包括：一个核心编码器，用于对直到所述过渡频率的所述音频信号的一个较低频带进行编码，其中所述的过渡频率是可变的，且其中所述的核心编码器以帧为基础在一帧上操作；和一个过渡频率控制模块，根据由所述核心编码器编码所述音频信号的难度和/或根据所述音频信号的音调的和噪声的频率范围之间的边界，来估计由所述核心编码器为一个当前帧所选择的过渡频率，因此所述的过渡频率在时间上是自适应地被选择的。

根据本发明的另一方面，这里提供一种用于编码音频信号以获得一种编码音频信号的方法，该编码音频信号由一个解码器使用，所述的解码器具有一个高频重建模块，用于对高于一个过渡频率的频率范

围执行高频重建，其特征在于，所述的方法包括以下步骤：对于直到一个过渡频率的所述音频信号的一个较低频带进行核心编码，其中所述的过渡频率是可变的，并且其中所述的核心编码是以帧的基础在一帧上发生；和根据在所述的核心编码步骤中对于音频信号编码的难度和/或根据所述音频信号的音调的和噪声的频率范围之间的边界，来估计在所述核心编码步骤中为一个当前帧所选择的一个过渡频率，因此所述的过渡频率在时间上是自适应地被选择的。

根据本发明的再一方面，这里提供一种用于解码一种编码音频信号的装置，所述的编码音频信号已经利用一个可变过渡频率而被编码，所述的编码音频信号包括：关于在时间上自适应地变化的一个过渡频率的信息，其特征在于，所述的用于解码的装置包括：一个比特流去复用器，用于将一个比特流分离成为核心编码解码器数据和包络数据；一个核心解码器，用于接收来自所述比特流去复用器的所述核心编码解码器数据和用于输出低频带数据；一个高频再生包络检波器，用于接收来自所述比特流去复用器中的所述包络数据和用于产生一个频谱包络输出；一个置换模块，用于根据关于所述可变过渡频率的信息来从所述低频带数据中产生一个复制的高频带信号；一个增益控制模块，响应所述高频再生包络检波器，用于调整所述复制的高频带信号到所述高频再生包络检波器输出的一个频谱包络以获得一个包络调整了的高频带信号；和一个加法器，用于把经过延迟了的所述低频带数据和所述包络调整了的高频带信号相加以获得一种数字宽带信号。

根据本发明的又一方面，这里提供一种用于解码一种编码音频信号所述的方法，所述的编码音频信号已经利用一个可变过渡频率而被编码，所述的编码音频信号包括：关于在时间上自适应变化的一个过渡频率的信息，其特征在于，所述的用于解码的方法包括以下步骤：将一个比特流分离成为核心编码解码器数据和包络数据；利用一个核心解码器接收来自所述比特流去复用器的核心编码解码器数据和输出低频带数据；利用一个高频再生包络检波器接收所述包络数据和产生一个频谱包络输出；利用一个置换模块，根据关于所述可变过渡

频率的信息，从所述的低频带数据中产生一个复制的高频带信号；响应所述高频再生包络解码器，利用一个增益控制模块调整所述复制的高频带信号到所述高频再生包络解码器输出的一个频谱包络，以获得一个包络调整了的高频带信号；和将经过延迟了的所述低频带数据和所述包络调整了的高频带信号相加，以获得一个数字宽带信号。

附图说明

现在将参考附图通过不限制本发明范围或精神的说明示例来描述本发明，附图中：

图 1是说明各项低频带、高频带以及过渡频率的曲线图。

图 2是说明核心编解码器工作负载测量的曲线图。

图 3是说明恒定比特率编解码器的短时比特需求变化的曲线图。

图 4是说明信号划分成为声音和类似噪声的频率范围的曲线图。

图 5是被一个过渡频率控制模块增强的一个基于 HFR的编码器框图。

图 6是详细地说明过渡频率控制模块的一个框图。

图 7是相应的基于 HFR的解码器框图。

具体实施方式

下面描述的实施例只是说明本发明的原理。不言而喻，在此描述的配置和细节的改进与变化对本领域技术人员来说是显而易见的。因此，想要仅仅通过随后的专利权利要求范围来限制而非通过在此的实施例说明和解释呈现的具体细节来限制。

在如图 1 给出的低频带或低频范围 101 被核心编解码器编码、并且高频带或高频范围 102 被一个适当的 HFR 方法适用的一个系统中，两个范围之间的边界可以被定义为过渡频率 103。因为编码方案按块规一帧帧的基础进行操作，所以人们可对于每个被处理帧自由地改变过渡频率。根据本发明，可设置一种适应过渡频率的检测算法如此以使获得组合编码系统的最佳质量。它的配置在下文中被称为过渡频率控制模块。

考虑核心编解码器的音频质量也是重建高频带质量的根据，很明显，在低频带范围中的一个优良且恒定的音频质量被期望。通过减低过渡频率，核心编解码器不得不对付的频率范围变小，并因此容易编码。因此，通过测量编码一个帧的难度并相应地调整过渡频率，则可以获得核心编解码器的一个更恒定的音频质量。

作为如何测量难度的一个示例，知觉熵 [ISO/TEC 13818-7 附件 B.2.1] 可以被使用：在这里，一个基于频谱分析的心理学模型被应用。通常，分析滤波器组的谱线被归组到频带中，在此一个频带内的线数取决于频带中心频率并且根据熟知的 bark scale 来被选择，目的是全频带的一个知觉固定频率解决方案。通过使用一个利用诸如频谱或暂时屏蔽之类的效果的心理学模式，则获得对于每个频带的听觉阈值。一个频带内的知觉熵然后如下给出：

$$e(b) = \frac{1}{2} \sum_{i=0}^{L(b)-1} \log_2(r(i)) + 1 \quad (\text{公式 1})$$

其中

$$r(i) = s(i)^2 \frac{L(b)}{l(b)}$$

并且

i-当前频带内的谱线指标

s(i)=线路 i 的频谱值

$L(b)$ =当前频带中的线路数

$t(b)$ =当前频带的心理学阈值

b =频带指标

l = 当前频带中的线数如此以致 $r(i) > 1.0$

并且只是项如此以致 $r(i) > 1.0$ 被用于累加中。

通过总计在低频带范围中不得被编码的所有频带的知觉熵，则获得当前帧的编码难点的一个测量。

根据等式2，则一种类似的方法通过总计每个频带的失真能量来计算在核心编解码器编码处理最后的失真能量。

$$n_{tot} = \sum_{b=0}^{B-1} n(b) \quad (\text{公式2})$$

在此

$$n(b) = \begin{cases} n_q(b) - t(b) & \text{for } n_q(b)/t(b) > 1.0 \\ 0 & \text{otherwise} \end{cases}$$

并且

$n_q(b)$ =量化噪声能量

$t(b)$ =心理学阈值

b =频带指标

B =频带数

此外，失真能量可以被一个音量曲线加权，以便把实际的失真加权到它的心理学相关性上。作为一个示例，等式 2 中的累加可以被修改为

$$n'_{tot} = \sum_{b=0}^{B-1} (n(b))^{0.23} \quad (\text{公式3})$$

在此按照 Zwicker 的音量函数的一个简化形式得到了使用 [“ Psychoacoustics”, Eberhard Zwicker 和 Hugo Fastl, Springer-Verlag, Berlin 1990]。

一个编码难点或工作负载测量因此可以被定义为总失真的一个函数。图 2 给出了一个知觉音频编解码器的失真能量以及一个相应工作负载测量的例子，在此，一个非线性递归已被用来计算工作负载。可以看到工作负载示出了时间上的高漂移并且取决于输入材料特性。

高知觉熵或高度失真能量表示一个信号心理声学上难以以一个有限比特率来编码，并且低频带中的可听人工产物可能出现。在这种情况下，过渡频率控制模块将指令使用一个较低的过渡频率以便对于知觉音频编码器更容易对付给定信号。同时，低知觉熵或低失真能量指示一个易编码信号。因此，过渡频率将被选择更高以便对于低频带允许一个更宽频率范围，从而减少由于任何现有 HFR方法有限性能而可能在高频带中引入的人工产物。如果过渡频率的调整已经在分析阶段被指令过的话，则两种方法都还允许通过重新编码当前帧来使用一种分析合成方法。可是，因为重叠转换被使用于最新科技的音频编解码器中，所以系统性能可以通过在时间上应用分析输入参数的一个平滑而被改良，以免过渡频率的太频繁切换，其可能引起抑制效应。如果实际的设备不需要就处理延迟方面被最佳化，则通过使用时间上的一个更大预见、提供查找在时间上的点（在该时间点处可以进行具有切换人工产物最小值的偏移）可以进一步改良该检测算法。非实时应用表示这种特定情况：如果期望的话，则要被编码的整个文件可以被分析。

在恒定比特率(CBR)音频编解码器的情况下，一个短时比特需求变化分析可以被使用作为交叉判断中的一个附加输入参数：诸如 MPEG层 3或 MPEG 2 AAC之类的最新科技的音频编码器使用一个比特储藏技术以便补偿每一帧从可用比特平均数中的短时峰值比特需求偏移。此类比特储藏的满度表示核心编码器是否能够解决好即将来临的难以编码的帧。每一帧被使用的比特数以及时间上的比特储藏满度的实际例子在图 3中给出。因此，如果比特储藏满度为高，则核心编码器将能够处理一个困难的帧并且不需要选择一个较低的过渡频率。同时，如果比特储藏满度为低，则通过减低过渡频率在随后的帧中结果的音频质量可以基本上被改良，如此以致由于不得被编码的频率范围较小而可以填满比特储藏。再一次，由于可以很好地预先预知比特储藏满度的特性，所以一个大的预见能够改善检测方法。

除当前帧的编码难点之外，基于过渡频率选择的另外一个重要的

参数被描述如下：诸如语音或一些乐器之类的大量音频信号示出这样一个性质：即，频谱范围可以被分成一个音调或音频范围和一个类似噪音范围。图 4 示出了此性质被清楚显示的一个音频输入信号的频谱。在频谱域中使用音调和/或噪声分析方法，可以检测两个范围，其可以被分别分类为声音的和类似噪声的。例如在 AAC 标准中给出的，音调可以被计算出 [ISO/IEC 13818-7: 1997(E), pp. 96-98, section B.2.1.4 "Steps in threshold calculation"]。诸如频谱均匀性测量之类的其它熟知的音调或噪声检测算法也适合于本目的。因此，在这些范围之间的过渡频率被使用作为本发明环境中的过渡频率以便更好地分开声音的和类似噪声的频谱范围并且把它们分别馈送给核心编码器，分别的 HFR 方法。因此组合编解码器系统的整个音频质量在这些情况下能基本上被改良。

显然，上面的方法同样适用于双端和单端 HFR 系统。在后一种情况中，被核心编解码器编码的变化带宽的仅仅一个低频带被发射。HFR 解码器然后推断来自低频带截止频率以上的一个包络。此外，本发明可应用到通过与被用于低频带编码的方法不同的任意方法来产生高频带的那些系统中。

当应用诸如频率转换之类的传统置换方法时，把 HFR 开始频率适应到低频带信号的变化带宽将是一个非常冗长乏味的任务。那些方法通常包括引起频移的低频带信号的滤波以便提取一个低通或带通信号，其随后在时域中被调制。因此，一个适配将合并低通或带通滤波器的切换并且在调制频率中变化。此外，滤波器的一个改变引起输出信号中的不连续性，这推动窗口技术的使用。可是，在一个基于滤波器组的系统中，通过从一组连续的滤波频带中提取次能带信号来自动地获得滤波。然后通过滤波器组内修补提取的次能带信号获得时域调制的一个等价物。修补很容易适合变化的过渡频率，并且前述的窗口为次能带域所固有，因此转换参数的改变以很少的附加复杂性而被实现。

图 5 示出了一个根据本发明增强的基于 HFR 的编解码器的编码

器侧的示例。模拟输入信号被馈送给一个 A/D转换器501，形成一个数字信号。数字音频信号被馈送给一个核心编码器 502，在此源编码被执行。另外，数字信号被馈送给一个 HFR包络编码器 503。HFR包络编码器的输出表示覆盖起始于图 1中说明的过渡频率 103的高频带 102的包络数据。在包络编码器中包络数据所需要的比特数被传递到核心编码器以便对于一个给定帧从总有效比特中减去之。核心编码器然后将编码剩余低频带频率范围直到过渡频率。正如本发明所教导的，一个过渡频率控制模块 504被加到编码器。输入信号以及核心编解码器状态信号的一个时域和/或频域表示被馈送给过渡频率控制模块。模块 504的输出，以过渡频率的最佳选择的形式，被馈送给核心和包络编码器以便指令将被编码的频率范围。两个编码方案每一个的频率范围例如也通过一个有效表格查寻方案进行编码。如果在两个后续帧之间的频率范围未变化，则这可以被一个单个比特发送信号以便保持尽可能小的比特率开销。因此频率范围不必在每一帧中都被明确发射。两个编码器的编码数据然后被馈送给多路复用器，形成被发射或储存的一个串行比特流。

图 6分别给出了在过渡频率控制模块 504和 601内的子系统示例。一个编码器工作负载测量分析模块 602例如使用如上所述的知觉熵或失真能量方法来探测对于核心编码器编码当前帧多么难。倘若核心编解码器使用一个比特储藏，则一个缓存器满度分析模块可以被包括，603。当可应用时，一个音调分析模块 604指令一个目标过渡频率与声音的/噪声的过渡频率相应。当计算为了获得最大总体性能而使用的过渡频率时，根据所使用的核心和 HFR编解码器的实际设备合并并均衡联合判断模块 606的所有输入参数。

相应的解码器侧如图 7所示。去复用器 701把比特流信号分离成为核心编解码器数据、包络数据，核心编解码器数据被馈送给核心解码器 702，包络数据被馈送给 HFR包络解码器 703。核心解码器产生一个覆盖低频带频率范围的信号。同样地，HFR包络解码器把数据解码成为高频带频率范围的频谱包络的一个表示。解码的包络数据然后

被馈送给增益控制模块 704。来自核心解码器中的低频带信号被路由到置换模块 705，它根据过渡频率产生来自低频带中的一个重复的高频带信号。该高频带信号被馈送给增益控制模块以便把高频带频谱包络调整到发射包络的上面。输出因此是一个包络调整了的高频带音频信号。此信号被加到来自延迟单元 706中的输出上，其用低频带音频信号馈送然而该延迟补偿了高频带信号的处理时间。最后，所获得的数字宽带信号在 D/A转换器 707中被转换成一个模拟音频信号。

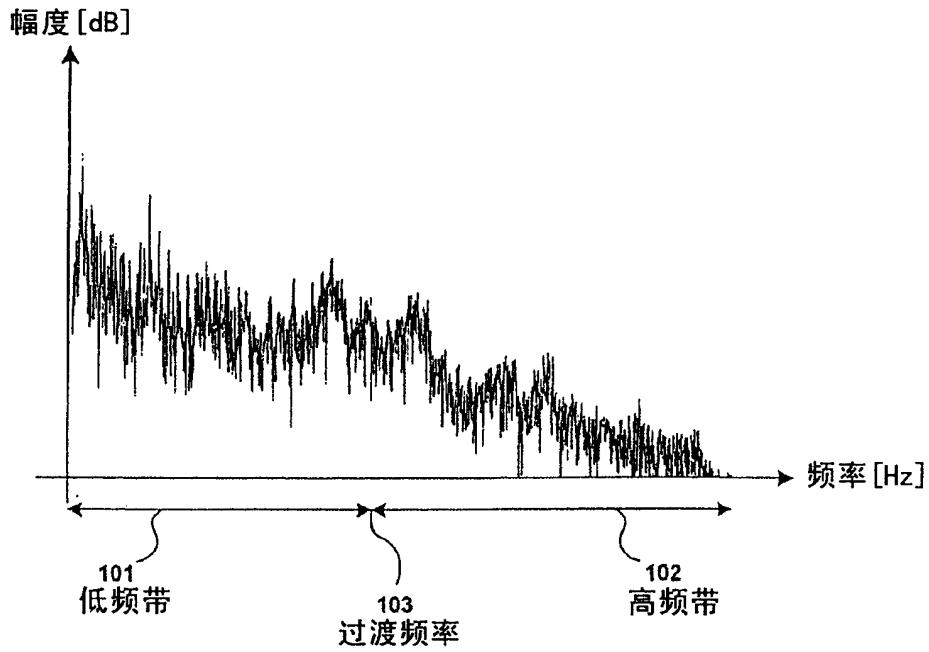


图1

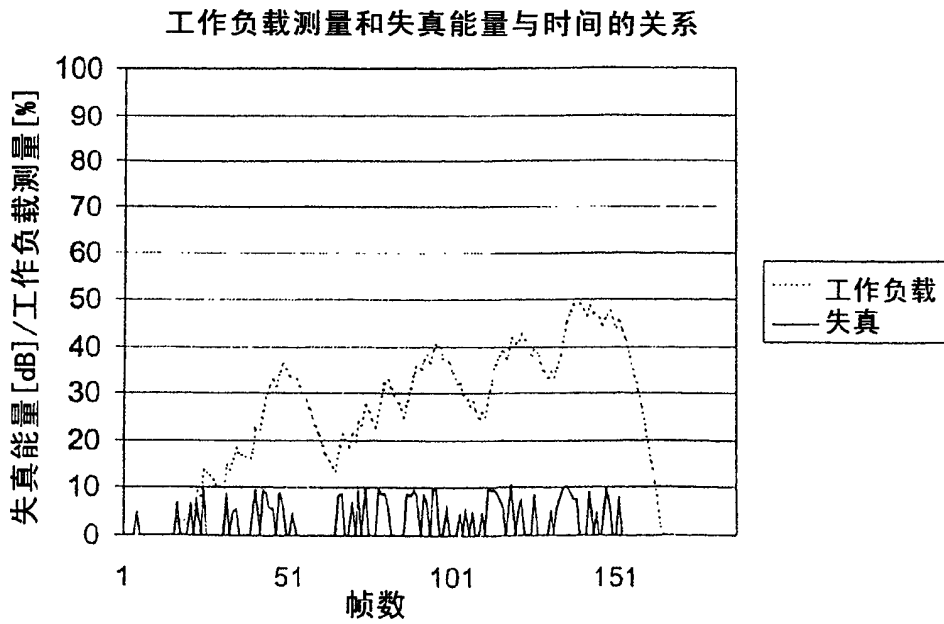


图2

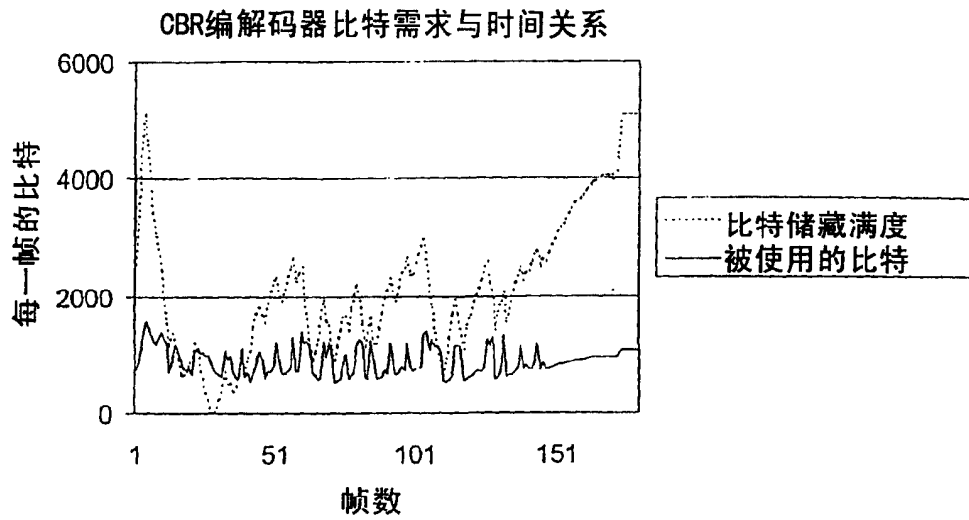


图 3

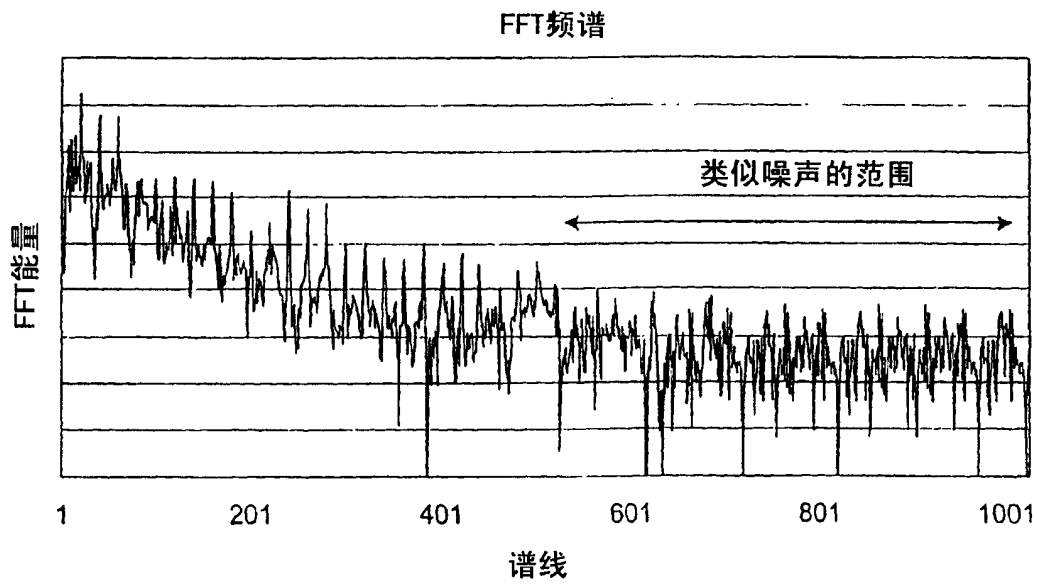
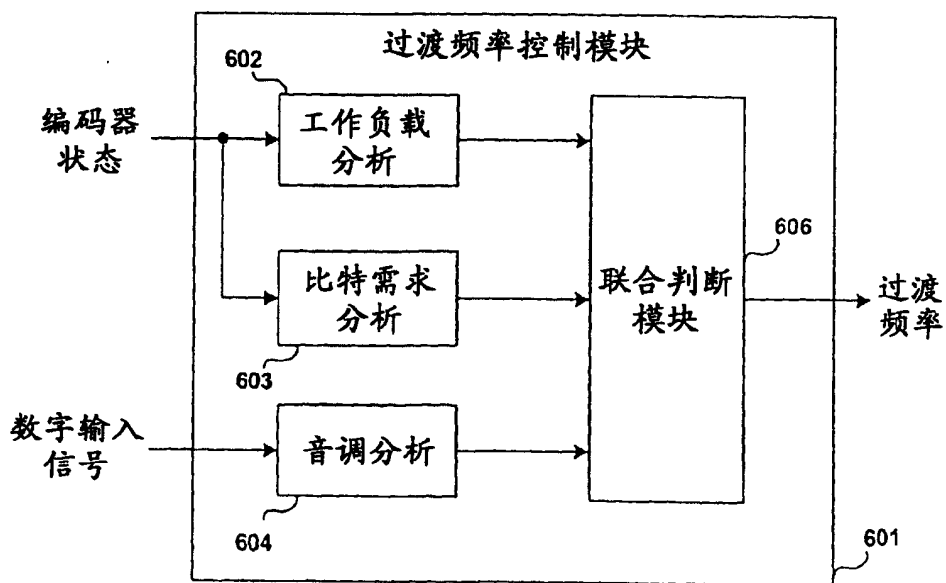
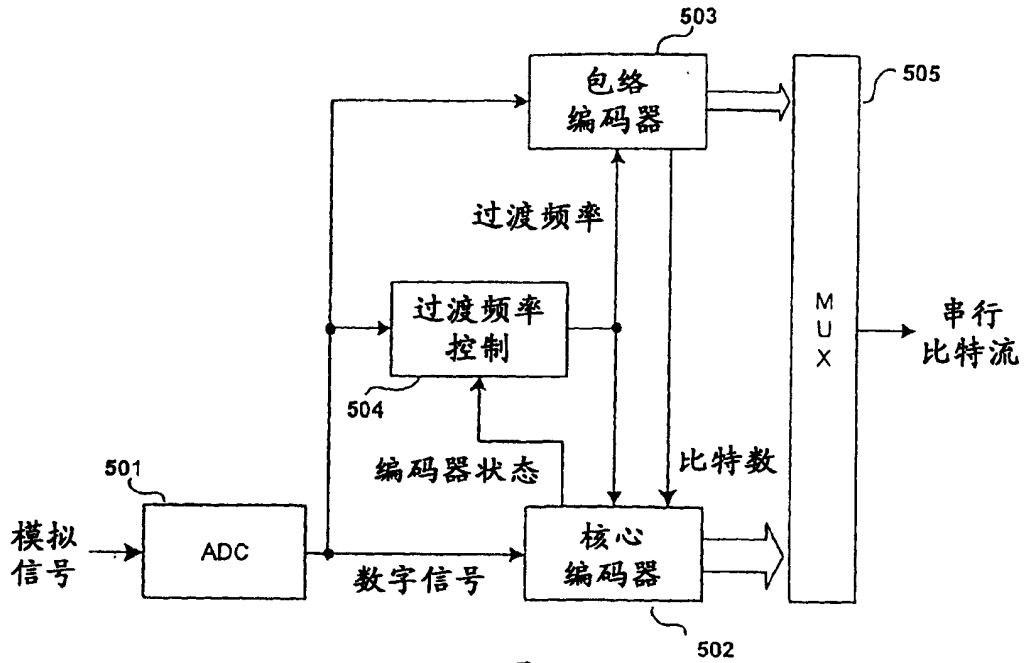


图 4



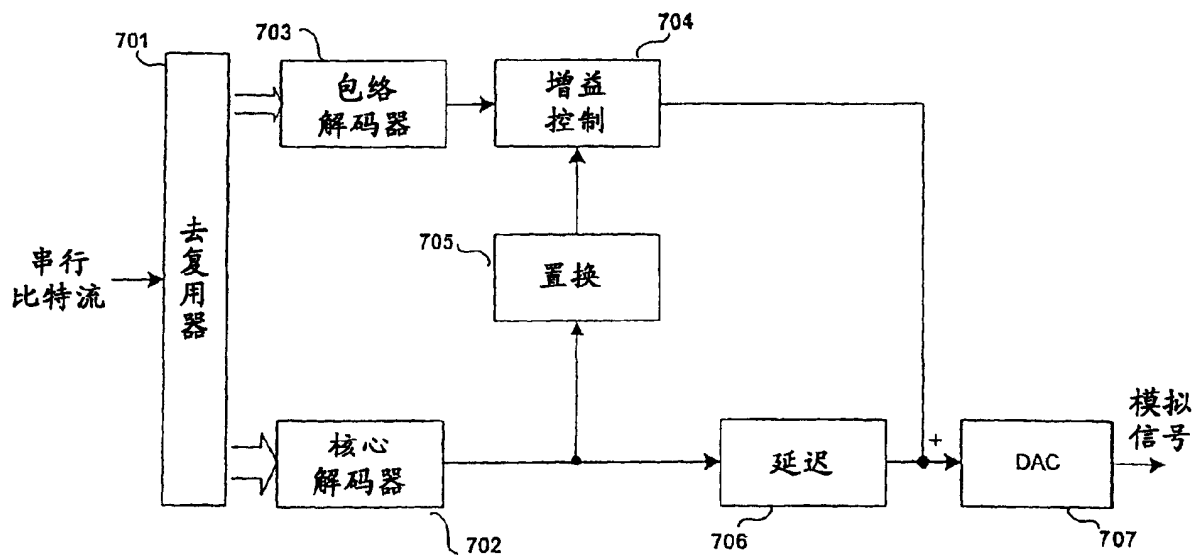


图7