



## [12] 发明专利说明书

[21] ZL 专利号 98810093.2

[45] 授权公告日 2005 年 4 月 27 日

[11] 授权公告号 CN 1199179C

[22] 申请日 1998.10.15 [21] 申请号 98810093.2

[30] 优先权

[32] 1997.10.17 [33] US [31] 08/953,106

[86] 国际申请 PCT/US1998/021700 1998.10.15

[87] 国际公布 WO1999/021187 英 1999.4.29

[85] 进入国家阶段日期 2000.4.12

[71] 专利权人 多尔拜实验特许公司

地址 美国加利福尼亚

[72] 发明人 路易斯·杜恩·菲尔德尔

审查员 王永真

[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 罗亚川

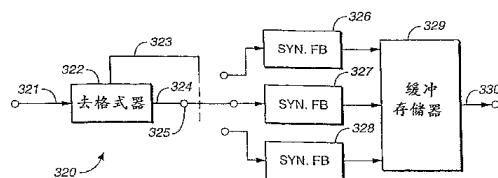
权利要求书 9 页 说明书 27 页 附图 8 页

[54] 发明名称 在帧边界处衰减频谱邻频干扰的音频编码

[57] 摘要

可以按照不同的组合方式来使用几种音频信号处理技术，以改进通过拼接编辑两个以上的其他信息流而形成的一个信息流所代表的音质。在将音频信息跟视频信息捆绑在一起的各种应用中，这种技术是特别有用的。在一种技术中，在音频信息流中所载运的各增益控制字被用来对跨越一个接合部的回放声音电平进行内插计算。在另一种技术中，特定的各种滤波器阵列或者各种 TDAC 变换形式被用来抑制出现在一个接合部两侧的各种混叠伪差。在又一种技术中，特定的滤波器阵列或者交叉衰落窗口函数被用来优化对产生于接合部的频谱邻频干扰的衰减。在再一种技术中，根据各种帧长度和频率来转换音频采样率，使得音频信息得以跟，例如，视频信息捆绑在一起。在还一种技术中，各音频块被动态地对准，使得在跨越一个接合

部时，能保持良好的同步关系。讨论了一个采样率为 48kHz 的音频跟 NTSC 视频相组合的实例。



1. 一种信号处理方法，其特征在于，该方法包括以下步骤：

接收一组输入信号，包括诸输入帧的一个序列，一个各自的输入信号帧包括多个用以运载已编码的音频信息的已编码的诸块，其中一个第1输入帧具有跟随在第1输入帧里面的一个第1临时的已编码块的一个终了的已编码块，并且一个紧跟在第1输入帧后面的第2输入帧具有领先于在第2输入帧里面的一个第2临时的已编码块的一个起始的已编码块，其中该起始的已编码块运载着由一个第1分析滤波器阵列进行滤波的音频信息，终了的已编码块运载着由一个第3分析滤波器阵列进行滤波的音频信息，并且第1和第2临时的已编码块运载着由一个第2分析滤波器阵列进行滤波的音频信息；

接收一组控制信号（323），它对介于所述第1输入帧和所述第2输入信号帧之间的一段边界加以标识；

紧挨着一个由诸合成样本块组成的第2合成帧的前面，产生一个由诸合成样本块组成的第1合成帧，该第1合成帧具有跟随在第1合成帧里面的一个第1临时的合成样本块的一个终了的合成样本块，并且该第2合成帧具有领先于在第2合成帧里面的一个第2临时的合成样本块的一个起始的合成样本块，其中，通过向起始的已编码块施加一个第1合成滤波器阵列（326），产生起始的合成样本块，通过向终了的已编码块施加一个第3合成滤波器阵列（328），产生终了的合成样本块，通过向第1临时的已编码块施加一个第2合成滤波器阵列（327），产生第1临时的合成样本块，并且通过向第2临时已编码块施加第2合成滤波器阵列，产生第2临时的合成样本块，并且其中；

在以 Hz 表示的、开始于大约 50,000 除以在起始的合成样本块中的样本数目、结束于大约 150,000 除以在起始的合成样本块中的样本数目的频率范围内，第1合成滤波器阵列与第1分析滤波器阵列相结合的频率响应（341）对频谱能量的衰减，大于第2合成滤波器阵列

与第2分析滤波器相结合的频率响应(342)对频谱能量的衰减;

在以 Hz 表示的、开始于大约 50,000 除以在终了的合成样本块中的样本数目、结束于大约 150,000 除以在终了的合成样本块中的样本数目的频率范围内,第3合成滤波器阵列与第3分析滤波器阵列相结合的频率响应(341)对频谱能量的衰减,大于第2合成滤波器阵列与第2分析滤波器相结合的频率响应对频谱能量的衰减;和

通过使相邻的合成样本块互相重叠以及将对应的互相重叠的诸样本相加来产生一组输出信号(330)。

2. 根据权利要求1中所述的信号处理方法,其特征在于,还包括:丢弃在上述终了的合成样本块中的一个或多个样本和在起始的合成样本块中的一个或多个样本。

3. 一种信号处理方法,其特征在于,包括以下步骤:

接收一组输入信号,其中包括:一个输入帧序列,一个各自的输入信号帧包括:多个用以运载已编码的音频信息的已编码块,其中一个第1输入帧具有:跟随在第1输入帧里面的一个第1临时的已编码块的一个终了的已编码块,而一个紧跟在第1输入帧后面的第2输入帧具有:领先于在第2输入帧里面的一个第2临时的已编码块的一个起始的已编码块,其中所述的起始的已编码块运载着由一个第1分析滤波器阵列进行滤波的音频信息,而所述的终了的已编码块运载着由一个第3分析滤波器阵列进行滤波的音频信息,以及所述的第1和第2临时的已编码块运载着由一个第2分析滤波器阵列进行滤波的音频信息;

接收一组控制信号(323),它对介于所述第1输入帧和所述第2输入信号帧之间的一段边界加以标识;

紧挨着一个由合成样本块组成的第2合成帧的前面,产生一个由合成样本块组成的第1合成帧,该第1合成帧具有跟随在第1合成帧里面的一个第1临时的合成样本块的一个终了的合成样本块,并且该第2合成帧具有领先于在第2合成帧里面的一个第2临时的合成样本块的一个起始的合成样本块,其中,通过向起始的已编码块施加

一个第 1 合成滤波器阵列 (326)，产生起始的合成样本块，通过向终了的已编码块施加一个第 3 合成滤波器阵列 (328)，产生终了的合成样本块，通过向第 1 临时的已编码块施加一个第 2 合成滤波器阵列 (327)，产生第 1 临时的合成样本块，并且通过向第 2 临时的已编码块施加第 2 合成滤波器阵列，产生第 2 临时的合成样本块，并且其中，

相对于一个参考频率响应 (343) 来说，通过第 1 合成滤波器阵列与第 1 分析滤波器阵列相结合而形成的一个第 1 频率响应 (341)，使频谱能量的衰减得以优化，以便与从参考频率响应所获得的衰减相比，在第 1 频率响应的阻带内获得较小的衰减，而在第 1 频率响应的阻带以外，对频谱邻频干扰提供更大的衰减，其中，参考频率响应对应于在一个大约 5 毫秒的时间间隔内，跟一个线性递减的斜坡相一致的脉冲响应，

相对于一个参考频率响应来说，通过第 3 合成滤波器阵列与第 3 分析滤波器相结合而形成的一个第 3 频率响应 (341)，使频谱能量的衰减得以优化，以便与从参考频率响应所获得的衰减相比，在第 3 频率响应的阻带内获得较小的衰减，而在第 3 频率响应的阻带以外，对频谱邻频干扰提供更大的衰减；以及

第 2 合成滤波器阵列与第 2 分析滤波器阵列相结合而形成的一个第 2 的频率响应 (342)，不同于第 1 频率响应以及第 3 频率响应；以及

通过使相邻的合成样本块互相重叠以及在互相重叠的合成样本块中，将对应的互相重叠的样本相加来产生一组输出信号 (330)。

4. 根据权利要求 3 所述的方法，其特征在于，还包括以下步骤：在所述终了的合成样本块中舍弃一个或多个样本以及在所述起始的合成样本块中舍弃一个或多个样本，

5. 根据权利要求 1 至 4 中任何一项所述的方法，其特征在于，还包括以下步骤：在所述第 1 输入帧以及所述第 2 输入帧之间标识一个拼接编辑点并响应于此而产生所述控制信号。

6. 根据权利要求 1 至 4 所述的方法, 其特征在于, 还包括以下步骤: 在所述输入信号中标识介于连续的诸帧之间的边界, 并且响应于此, 产生所述控制信号, 以便将领先于一个各自的边界的每一个输入帧标识为一个各自的第 1 输入帧, 并且将跟随在一个各自的边界后面的每一个输入帧标识为一个各自的第 2 输入帧。

7. 根据权利要求 1 至 4 中任何一项所述的方法, 其特征在于, 所述第 1 合成滤波器阵列包括: 从一个 $\alpha$ 值小于 2 的凯撒-贝塞尔窗口函数导出的一个或多个反变换以及一个合成窗口函数, 所述第 2 合成滤波器阵列包括从一个 $\alpha$ 值大于 2 的凯撒-贝塞尔窗口函数导出的一个或多个反变换以及一个合成窗口函数, 以及所述第 3 合成滤波器阵列包括从一个 $\alpha$ 值小于 2 的凯撒-贝塞尔窗口函数导出的一个或多个反变换以及一个合成窗口函数。

8. 根据权利要求 1 至 4 中任何一项所述的方法, 其特征在于, 所述第 1 合成滤波器阵列包括一个反变换以及一个第 1 合成窗口函数, 所述第 2 合成滤波器阵列包括一个反变换以及一个第 2 合成窗口函数, 所述第 3 合成滤波器阵列包括一个反变换以及一个第 3 合成窗口函数。

9. 根据权利要求 1 至 4 中任何一项所述的方法, 其特征在于, 所述第 1 分析滤波器阵列、所述第 2 分析滤波器阵列和所述第 3 分析滤波器阵列是相同的。

10. 根据权利要求 1 至 4 中任何一项所述的方法, 其特征在于, 包括以下步骤:

从所述第 1 输入帧获得多个增益控制字 (145), 它们表示一个第 1 起始增益水平以及一个第 1 終了增益水平, 并从所述第 2 输入帧获得多个增益控制字 (145), 它们表示一个第 2 起始增益水平以及一个第 2 終了增益水平;

产生一个第 1 增益控制信号 (149a, 149b), 表示第 1 起始和第 1 終了增益水平的一个内插值, 以及产生一个第 2 增益控制信号 (149a, 149b), 表示第 2 起始和第 2 終了增益水平的一个内插值;

以及

根据所述第 1 增益控制信号，调制所述终了的合成样本块，以实现所述第 3 合成滤波器阵列（328），并且根据所述第 2 增益控制信号，调制所述起始的合成样本块，以实现所述第 1 合成滤波器阵列（326）。

11. 一种信号处理装置，其特征在于，包括：

用于接收一组输入信号的装置，上述输入信号包括诸输入帧的一个序列，一个各自的输入帧包括多个用以运载已编码的音频信息的多个已编码的块，其中一个第 1 输入帧具有跟随在第 1 输入帧里面的一个第 1 临时的已编码块的一个终了的已编码块，并且一个紧跟在第 1 输入帧后面的第 2 输入帧具有领先于在第 2 输入帧里面的一个第 2 临时的已编码块的一个起始的已编码块，其中该起始的已编码块运载着由一个第 1 滤波器阵列进行滤波的音频信息，该终了的已编码块运载着由一个第 3 滤波器阵列进行滤波的音频信息，并且第 1 和第 2 临时的已编码块运载着由一个第 2 分析滤波器阵列进行滤波的音频信息；

用于接收一组控制信号（323）的装置，该控制信号对介于所述第 1 输入帧和所述第 2 输入信号帧之间的一段边界加以标识；

紧挨着一个由已合成的诸样本块组成的第 2 合成帧的前面，用以产生一个由已合成的诸样本块组成的第 1 合成帧的装置，该第 1 合成帧具有跟随在第 1 合成帧里面的一个第 1 临时的合成样本块的一个终了的合成样本块，并且该第 2 合成帧具有领先于在第 2 合成帧里面的一个第 2 临时的合成样本块的一个起始的合成样本块，其中，通过向起始的已编码块施加一个第 1 合成滤波器阵列（326），产生起始的合成样本块，通过向终了的已编码块施加一个第 3 合成滤波器阵列（328），产生终了的合成样本块，通过向第 1 临时的已编码块施加一个第 2 合成滤波器阵列（327），产生第 1 临时的合成样本块，并且通过向第 2 临时已编码块施加第 2 合成滤波器阵列，产生第 2 临时的合成样本块，并且其中，

在以 Hz 表示的、开始于大约 50,000 除以在起始的合成样本块中的样本数目、结束于大约 150,000 除以在起始的合成样本块中的样本数目的频率范围内,第 1 合成滤波器阵列与第 1 分析滤波器阵列相结合的频率响应 (341) 对频谱能量的衰减,大于第 2 合成滤波器阵列与第 2 分析滤波器相结合的频率响应 (342) 对频谱能量的衰减,以及

在以 Hz 表示的、开始于大约 50,000 除以在终了的合成样本块中的样本数目、结束于大约 150,000 除以在终了的合成样本块中的样本数目的频率范围内,第 3 合成滤波器阵列与第 3 分析滤波器阵列相结合的频率响应 (341) 对频谱能量的衰减,大于第 2 合成滤波器阵列与第 2 分析滤波器阵列相结合的频率响应对频谱能量的衰减; 以及

通过使相邻的诸合成样本块互相重叠以及在互相重叠的诸合成样本块中,将对应的互相重叠的诸样本相加来产生一组输出信号 (330) 的装置。

12. 根据权利要求 11 中所述的装置,其特征在于,还包括:丢弃在上述终了的合成样本块中的一个或多个样本和在起始的合成样本块中的一个或多个样品的装置。

13. 一种信号处理装置,其特征在于,包括:

用于接收一组输入信号的装置,上述输入信号包括诸输入帧的一个序列,一个各自的输入信号帧包括多个用以运载已编码的音频信息的已编码的诸块,其中一个第 1 输入帧具有跟随在第 1 输入帧里面的一个第 1 临时的已编码块的一个终了的已编码块,并且一个紧跟在第 1 输入帧后面的第 2 输入帧具有领先于在第 2 输入帧里面的一个第 2 临时的已编码块的一个起始的已编码块,其中该起始的已编码块运载着由一个第 1 滤波器阵列进行滤波的音频信息,该终了的已编码块运载着由一个第 3 滤波器阵列进行滤波的音频信息,并且第 1 和第 2 临时的已编码块运载着由一个第 2 分析滤波器阵列进行滤波的音频信息;

用于接收一组控制信号 (323) 的装置,该控制信号对介于所述

第 1 输入帧和所述第 2 输入信号帧之间的一段边界加以标识;

具有下列功能的装置:紧挨着一个由已合成的诸样本块组成的第 2 合成帧的前面,产生一个由已合成的诸样本块组成的第 1 合成帧,该第 1 合成帧具有跟随在第 1 合成帧里面的一个第 1 临时的已合成的样本块的一个终了的已合成样本块,并且该第 2 合成帧具有领先于在第 2 合成帧里面的一个第 2 临时的已合成的样本块的一个起始的已合成样本块,其中,通过向起始的已编码块施加一个第 1 合成滤波器阵列 (326),产生起始的合成样本块,通过向终了的已编码块施加一个第 3 合成滤波器阵列 (328),产生终了的合成样本块,通过向第 1 临时的已编码块施加一个第 2 合成滤波器阵列 (327),产生第 1 临时的合成样本块,并且通过向第 2 临时的已编码块施加第 2 合成滤波器阵列,产生第 2 临时的合成样本块,并且其中,

相对于一个参考频率响应 (343) 来说,通过第 1 合成滤波器阵列与第 1 分析滤波器阵列相结合而形成的一个第 1 的频率响应(341),使频谱能量的衰减得以优化,以便与从参考频率响应所获得的衰减相比,在第 1 频率响应的阻带内获得较小的衰减,而在第 1 频率响应的阻带以外,对频谱邻频干扰提供更大的衰减,其中,参考频率响应对应于在一个大约 5 毫秒的时间间隔内,跟一个线性递减的斜坡相一致的脉冲响应,

相对于一个参考频率响应来说,通过第 3 合成滤波器阵列与第 3 分析滤波器阵列相结合而形成的一个第 3 频率响应 (341),使频谱能量的衰减得以优化,以便与从参考频率响应所获得的衰减相比,在第 3 频率响应的阻带内获得较小的衰减,而在第 3 频率响应的阻带以外,对频谱邻频干扰提供更大的衰减;以及

第 2 合成滤波器阵列与第 2 分析滤波器阵列相结合而形成的一个第 2 频率响应 (342),不同于第 1 频率响应以及第 3 频率响应;以及

通过使相邻的诸合成样本块互相重叠以及在互相重叠的诸合成样本块中,将对应的互相重叠的诸样本相加来产生一组输出信号

(330) 的装置。

14. 根据权利要求 11 所述的装置, 其特征在于, 还包括: 在所述第 1 临时信号块中舍弃一个或多个样本、以及在所述第 2 临时信号块中舍弃一个或多个样本的装置,

15. 根据权利要求 11 至 14 中任何一项所述的装置, 其特征在于, 还包括: 在所述第 1 输入帧以及所述第 2 输入帧之间标识一个拼接编辑点, 并响应于此, 产生所述控制信号的装置。

16. 根据权利要求 11 至 14 所述的装置, 其特征在于, 还包括: 用于标识介于所述输入信号的连续的各帧之间的边界, 以及响应于此产生控制信号的装置, 以便将前置于一个各自的边界的每一个输入信号帧标识为一个各自的第 1 输入帧, 以及将跟随于一个各自的边界后面的每一个输入信号帧标识为一个各自的第 2 输入帧。

17. 根据权利要求 11 至 14 中任何一项所述的装置, 其特征在于, 所述第 1 合成滤波器阵列包括从一个  $\alpha$  值小于 2 的凯撒-贝塞尔窗口函数导出的一个或多个反变换以及一个合成窗口函数, 所述第 2 合成滤波器阵列包括从一个  $\alpha$  值大于 2 的凯撒-贝塞尔窗口函数导出的一个或多个反变换以及一个合成窗口函数, 以及所述第 3 合成滤波器阵列包括从一个  $\alpha$  值小于 2 的凯撒-贝塞尔窗口函数导出的一个或多个反变换以及一个合成窗口函数。

18. 根据权利要求 11 至 14 中任何一项所述的装置, 其特征在于, 所述第 1 合成滤波器阵列包括: 一个反变换以及一个第 1 合成窗口函数, 所述第 2 合成滤波器阵列包括: 所述反变换以及一个第 2 合成窗口函数, 并且所述第 3 合成滤波器阵列包括: 所述反变换以及一个第 3 合成窗口函数。

19. 根据权利要求 11 至 14 中任何一项所述的装置, 其特征在于, 所述第 1 分析滤波器阵列、所述第 2 分析滤波器阵列和所述第 3 分析滤波器阵列是相同的。

20. 根据权利要求 11 至 14 中任何一项所述的装置, 其特征在于, 包括:

用于从所述第 1 输入帧获得多个增益控制字(145)的装置(142), 上述增益控制字表示一个第 1 起始增益水平以及一个第 1 终止增益水平, 以及用于从所述第 2 输入帧获得多个增益控制字(145)的装置(142), 上述增益控制字表示一个第 2 起始增益水平以及一个第 2 终止增益水平;

用于产生一个第 1 增益控制信号(149a, 149b)的装置, 该增益控制信号表示第 1 起始和第 1 终止增益水平的一个内插值, 以及用于产生一个第 2 增益控制信号(149a, 149b)的装置, 该增益控制信号表示第 2 起始和第 2 终止增益水平的一个内插值; 以及

用于根据所述第 1 增益控制信号, 调制所述终止的合成样本块, 以实现所述第 3 合成滤波器阵列(328)的装置, 以及用于根据所述第 2 增益控制信号, 调制所述起始的合成样本块, 以实现所述第 1 合成滤波器阵列(326)的装置。

## 在帧边界处衰减 频谱邻频干扰的音频编码

### 技术领域:

本发明涉及音频信号处理,在其中,音频信息流被排列成信息帧的形式。特别是,本发明涉及改进音频信息流的音质,上述音频信息流是通过将基于帧的诸音频信息流加以拼接而形成的。

### 背景技术

编辑音频或视频素材的过程实质上是将两段素材拼接或对接(其中之一)在一起的过程。一个简单的编辑范例就是剪接动画电影胶片的过程。待拼接的两段素材可以取自不同的来源,例如,音频信息的不同声道,或者它们可以取自相同的来源。无论在哪一种情况下,拼接过程通常会在音频或视频素材中产生可觉察的或不可觉察的不连续性。

### 音频编码

#### 块处理

数字音频的日益增长的应用使得它难以在不产生可听见的各种伪差的条件下编辑音频素材。这种情况的出现,部分地是由于必须以块的形式频繁地对诸数字样本块进行处理或编码。许多感知的或基于心理声学的音频编码系统利用滤波器阵列或者各种变换,将各信号样本块转换为已编码的子带信号样本的诸块或者各种变换系数,它们必须经过合成滤波或者被逆变换为各块,以便恢复原始信号的一个复制品。起码,必须在一个块的边界上来完成已处理的音频信号;否则,由剩余的部分块所代表的音频信息不能完全地被恢复。

在本文的以下部分中,诸如“编码”以及“编码器”这样的名词指的是用于信号处理的各种方法和装置,诸如“已编码的”这样的其他名词指的是这样的处理的诸结果。在这些名词中,没有一条隐含着任何特定的处理方式,例如在一组信号中降低信息的无关性或冗余度。例如,编码包括产生代表一组信号的脉冲编码调制(PCM)诸样本,并按照某种规格将信息排列成样板或格式。在本文中使用的诸如“块”和“帧”这样的名词指的是跟这些名词在别处(例如在 ANSI S4.40 - 1992 标准,有时也称为 AES - 3/EBU 数字音频标准中)所指的对象不同的那些信息的诸分组或诸间隔,本文所使用的名词“滤波器”和“滤波器阵列”基本上包括任何形式的递归的和非递归的

滤波方法，例如正交镜像滤波器（QMF）和变换，并且“已滤波”的信息是使用这样的滤波器所得到的结果。下面将对通过各种变换来实现的滤波器阵列作出更专门的说明。

使用重叠的块结构来处理 and 编码节目素材的各种编码系统对编辑工作产生了附加的限制。由于已编码的各块的重叠的性质，所以即使从已编码的诸样本或诸系数的一个完整的块，也无法完全地恢复原始信号。

借助于一种常用的重叠块变换，就能清楚地说明这种限制，这种修正的离散余弦变换(DCT)，在 Princen, Johnson 和 Bradley 合写的题为《使用基于时域混叠抵消的滤波器阵列设计的子带 / 变换编码》的论文中对此作了叙述，该文被收入 1987 年国际声学、语音和信号处理会议论文集，1987 年 5 月，第 2161-2164 页。这种变换是一种奇数堆栈临界采样单边带分析-合成系统的时域等价物，并且在本文中被称为奇数堆栈时域混叠抵消（O-TDAC）方法。对以半个块长互相重叠的诸样本块实施正变换，并且通过将诸变换系数除以 2 来获得临界采样；然而，由于这种缩减所带来的信息损失将在已恢复的信号中产生时域混叠。通过对诸变换系数的诸块实施逆变换以产生合成样本的诸块，对已合成样本的诸块施加一个形状合适的合成窗口函数，以及对窗口内的诸块进行重叠和相加，该合成过程就能抵消这种混叠。例如，若一个 TDAC 编码系统产生诸块 B1 - B2 的一个序列，则处于块 B1 的后半部和块 B2 的前半部之中的混叠伪差将互相抵消。

若来自一个 TDAC 编码系统的两组已编码的信息流在一个块的边界处被拼接，则所得到的诸块的序列将不能抵消互相之间的混叠伪差。例如，假设一段已编码的信息流被剪断，使得它结束于介于诸块 B1 - B2 之间的一个块边界上，并且另一个已编码的信息流被剪断，使得它开始于介于诸块 A1 - A2 之间的一个块边界上。若这两段已编码的信息流被拼接，并使得块 B1 紧挨着块 A2 的前面，则处于块 B1 的后半部以及块 A2 的前半部的各种混叠伪差通常也将无法互相抵消。

现有技术的方法和装置不是忽略了这个问题就是提出了不能令人满意的解决方案。一种解决方案通过从每一组已编码的音频流中恢复或解码原始音频信号，来降低未被抵消的混叠伪差的可听度。将一组音频流交叉衰落（平滑过渡）到另一组，并且将所得到的交叉衰落流再编码为一组新的已编码音频流。不幸的是，解码 / 再编码过程使所得到的结果信号恶化，该过程的开销很大，使得它没有吸引力，并且由于交叉衰落无法取消，所以紧挨着接合部两侧原始信号不能独立地被恢复。

### 频谱邻频干扰

拼接编辑产生了现有技术无法解决的另一个问题。在分割频带的感知编码技术中，这个问题带来特别的麻烦。感知分割频带编码方法将一个滤波器阵列施加到输入信号上，以产生具有与人的听觉系统的临界带宽相当的带宽的诸子带信号或者诸变换系数组。理想地，用刚好足够的位数来对每一组子带信号或者变换系数组进行量化或编码，并且通过让噪声被原始信号中的频谱成分所掩盖，来使所得到的量化噪声变为听不见。编码性能显著地受到施加于输入信号以产生子带信号或诸变换系数的滤波器阵列的频率响应特性的影响。一般来说，通过在滤波器阻带的频率上增加衰减来换取较宽的滤波器通带，使这些特性得以优化。例如，见美国专利第 5,109,417 号。

拼接编辑倾向于在通常处于滤波器通带或介于通带与阻带之间的过渡区域内的频率范围内（并且不在滤波器阻带范围内），产生显著的假频谱成分或者“频谱邻频干扰”，因此，为了优化总的编码性能而设计的滤波器阵列对在拼接编辑中产生的频谱邻频干扰并不提供足够的衰减。由于这些伪差通常是如此之大，以致不能被原始信号所掩盖，所以它们通常是听得见的。

### 发明内容：

本发明的一个目的是，通过减少在接合部的频谱邻频干扰来改进由拼接两个或多个基于帧的音频信息流而形成的一个音频信息流所代表的音质。

根据本发明的一个方面，这里提供一种信号处理方法，其特征在于，该方法包括以下步骤：接收一组输入信号，包括诸输入帧的一个序列，一个各自的输入信号帧包括多个用以运载已编码的音频信息的已编码的诸块，其中一个第 1 输入帧具有跟随在第 1 输入帧里面的一个第 1 临时的已编码块的一个终了的已编码块，并且一个紧跟在第 1 输入帧后面的第 2 输入帧具有领先于在第 2 输入帧里面的一个第 2 临时的已编码块的一个起始的已编码块，其中该起始的已编码块运载着由一个第 1 分析滤波器阵列进行滤波的音频信息，终了的已编码块运载着由一个第 3 分析滤波器阵列进行滤波的音频信息，并且第 1 和第 2 临时的已编码块运载着由一个第 2 分析滤波器阵列进行滤波的音频信息；接收一组控制信号，它对介于所述第 1 输入帧和所述第 2 输入信号帧之间的一段边界加以标识；紧挨着一个由诸合成样本块组成的第 2 合成帧的前面，产生一个由诸合成样本块组成的第 1 合成帧，该第 1 合成帧具有跟随在第 1 合成帧里面的一个第 1 临时的合成样本块的一个终了的合成样本块，并且该第 2 合成帧具有领先于在第 2 合成

帧里面的一个第 2 临时的合成样本块的一个起始的合成样本块，其中，通过向起始的已编码块施加一个第 1 合成滤波器阵列，产生起始的合成样本块，通过向终了的已编码块施加一个第 3 合成滤波器阵列，产生终了的合成样本块，通过向第 1 临时的已编码块施加一个第 2 合成滤波器阵列，产生第 1 临时的合成样本块，并且通过向第 2 临时已编码块施加第 2 合成滤波器阵列，产生第 2 临时的合成样本块，并且其中；在以 Hz 表示的、开始于大约 50,000 除以在起始的合成样本块中的样本数目、结束于大约 150,000 除以在起始的合成样本块中的样本数目的频率范围内，第 1 合成滤波器阵列与第 1 分析滤波器阵列相结合的频率响应对频谱能量的衰减，大于第 2 合成滤波器阵列与第 2 分析滤波器相结合的频率响应对频谱能量的衰减；在以 Hz 表示的、开始于大约 50,000 除以在终了的合成样本块中的样本数目、结束于大约 150,000 除以在终了的合成样本块中的样本数目的频率范围内，第 3 合成滤波器阵列与第 3 分析滤波器阵列相结合的频率响应对频谱能量的衰减，大于第 2 合成滤波器阵列与第 2 分析滤波器相结合的频率响应对频谱能量的衰减；和通过使相邻的合成样本块互相重叠以及将对应的互相重叠的诸样本相加来产生一组输出信号。

根据本发明的另一个方面，这里提供一种信号处理方法，其特征在于，包括以下步骤：接收一组输入信号，其中包括：一个输入帧序列，一个各自的输入信号帧包括：多个用以运载已编码的音频信息的已编码块，其中一个第 1 输入帧具有：跟随在第 1 输入帧里面的一个第 1 临时的已编码块的一个终了的已编码块，而一个紧跟在第 1 输入帧后面的第 2 输入帧具有：领先于在第 2 输入帧里面的一个第 2 临时的已编码块的一个起始的已编码块，其中所述的起始的已编码块运载着由一个第 1 分析滤波器阵列进行滤波的音频信息，而所述的终了的已编码块运载着由一个第 3 分析滤波器阵列进行滤波的音频信息，以及所述的第 1 和第 2 临时的已编码块运载着由一个第 2 分析滤波器阵列进行滤波的音频信息；接收一组控制信号，它对介于所述第 1 输入帧和所述第 2 输入信号帧之间的一段边界加以标识；紧挨着一个由合成样本块组成的第 2 合成帧的前面，产生一个由合成样本块组成的第 1 合成帧，该第 1 合成帧具有跟随在第 1 合成帧里面的一个第 1 临时的合成样本块的一个终了的合成样本块，并且该第 2 合成帧具有领先于在第 2 合成帧里面的一个第 2 临时的合成样本块的一个起始的合成样本块，其中，通过向起始的已编码块施加一个第 1 合成滤波器阵列，产生起始的合成样本块，通过向终了的已编码块施加一个第 3 合成滤波器阵列，产生终了的合成样本块，通过向第 1 临时的已编码块施加

一个第2合成滤波器阵列，产生第1临时的合成样本块，并且通过向第2临时的已编码块施加第2合成滤波器阵列，产生第2临时的合成样本块，并且其中，相对于一个参考频率响应来说，通过第1合成滤波器阵列与第1分析滤波器阵列相结合而形成的一个第1频率响应，使频谱能量的衰减得以优化，以便与从参考频率响应所获得的衰减相比，在第1频率响应的阻带内获得较小的衰减，而在第1频率响应的阻带以外，对频谱邻频干扰提供更大的衰减，其中，参考频率响应对应于在一个大约5毫秒的时间间隔内，跟一个线性递减的斜坡相一致的脉冲响应，相对于一个参考频率响应来说，通过第3合成滤波器阵列与第3分析滤波器相结合而形成的一个第3频率响应，使频谱能量的衰减得以优化，以便与从参考频率响应所获得的衰减相比，在第3频率响应的阻带内获得较小的衰减，而在第3频率响应的阻带以外，对频谱邻频干扰提供更大的衰减；以及第2合成滤波器阵列与第2分析滤波器阵列相结合而形成的一个第2的频率响应，不同于第1频率响应以及第3频率响应；以及通过使相邻的合成样本块互相重叠以及在互相重叠的合成样本块中，将对应的互相重叠的样本相加来产生一组输出信号。

根据本发明的再一个方面，这里提供一种信号处理装置，其特征在于，包括：用于接收一组输入信号的装置，上述输入信号包括诸输入帧的一个序列，一个各自的输入帧包括多个用以运载已编码的音频信息的多个已编码的块，其中一个第1输入帧具有跟随在第1输入帧里面的一个第1临时的已编码块的一个终了的已编码块，并且一个紧跟在第1输入帧后面的第2输入帧具有领先于在第2输入帧里面的一个第2临时的已编码块的一个起始的已编码块，其中该起始的已编码块运载着由一个第1滤波器阵列进行滤波的音频信息，该终了的已编码块运载着由一个第3滤波器阵列进行滤波的音频信息，并且第1和第2临时的已编码块运载着由一个第2分析滤波器阵列进行滤波的音频信息；用于接收一组控制信号的装置，该控制信号对介于所述第1输入帧和所述第2输入信号帧之间的一段边界加以标识；紧挨着一个由已合成的诸样本块组成的第2合成帧的前面，用以产生一个由已合成的诸样本块组成的第1合成帧的装置，该第1合成帧具有跟随在第1合成帧里面的一个第1临时的合成样本块的一个终了的合成样本块，并且该第2合成帧具有领先于在第2合成帧里面的一个第2临时的合成样本块的一个起始的合成样本块，其中，通过向起始的已编码块施加一个第1合成滤波器阵列，产生起始的合成样本块，通过向终了的已编码块施加一个第3合成滤波器阵列，产生终了的合成样本块，通过向第1临时的已编码块施加一个第2合成滤波器阵列，产生第1临

时的合成样本块，并且通过向第 2 临时已编码块施加第 2 合成滤波器阵列，产生第 2 临时的合成样本块，并且其中，在以 Hz 表示的、开始于大约 50,000 除以在起始的合成样本块中的样本数目、结束于大约 150,000 除以在起始的合成样本块中的样本数目的频率范围内，第 1 合成滤波器阵列与第 1 分析滤波器阵列相结合的频率响应对频谱能量的衰减，大于第 2 合成滤波器阵列与第 2 分析滤波器相结合的频率响应对频谱能量的衰减，以及在以 Hz 表示的、开始于大约 50,000 除以在终了的合成样本块中的样本数目、结束于大约 150,000 除以在终了的合成样本块中的样本数目的频率范围内，第 3 合成滤波器阵列与第 3 分析滤波器阵列相结合的频率响应对频谱能量的衰减，大于第 2 合成滤波器阵列与第 2 分析滤波器阵列相结合的频率响应对频谱能量的衰减；以及通过使相邻的诸合成样本块互相重叠以及在互相重叠的诸合成样本块中，将对应的互相重叠的诸样本相加来产生一组输出信号的装置。

根据本发明的又一个方面，这里提供一种信号处理装置，其特征在于，包括：用于接收一组输入信号的装置，上述输入信号包括诸输入帧的一个序列，一个各自的输入信号帧包括多个用以运载已编码的音频信息的已编码的诸块，其中一个第 1 输入帧具有跟随在第 1 输入帧里面的一个第 1 临时的已编码块的一个终了的已编码块，并且一个紧跟在第 1 输入帧后面的第 2 输入帧具有领先于在第 2 输入帧里面的一个第 2 临时的已编码块的一个起始的已编码块，其中该起始的已编码块运载着由一个第 1 滤波器阵列进行滤波的音频信息，该终了的已编码块运载着由一个第 3 滤波器阵列进行滤波的音频信息，并且第 1 和第 2 临时的已编码块运载着由一个第 2 分析滤波器阵列进行滤波的音频信息；用于接收一组控制信号的装置，该控制信号对介于所述第 1 输入帧和所述第 2 输入信号帧之间的一段边界加以标识；具有下列功能的装置：紧挨着一个由已合成的诸样本块组成的第 2 合成帧的前面，产生一个由已合成的诸样本块组成的第 1 合成帧，该第 1 合成帧具有跟随在第 1 合成帧里面的一个第 1 临时的已合成的样本块的一个终了的已合成样本块，并且该第 2 合成帧具有领先于在第 2 合成帧里面的一个第 2 临时的已合成的样本块的一个起始的已合成样本块，其中，通过向起始的已编码块施加一个第 1 合成滤波器阵列，产生起始的合成样本块，通过向终了的已编码块施加一个第 3 合成滤波器阵列，产生终了的合成样本块，通过向第 1 临时的已编码块施加一个第 2 合成滤波器阵列，产生第 1 临时的合成样本块，并且通过向第 2 临时的已编码块施加第 2 合成滤波器阵列，产生第 2 临时的合成样本块，并且其中，相对于一个参考频率响应来说，通过第 1 合成滤波器阵列与

第1分析滤波器阵列相结合而形成的一个第1的频率响应，使频谱能量的衰减得以优化，以便与从参考频率响应所获得的衰减相比，在第1频率响应的阻带内获得较小的衰减，而在第1频率响应的阻带以外，对频谱邻频干扰提供更大的衰减，其中，参考频率响应对应于在一个大约5毫秒的时间间隔内，跟一个线性递减的斜坡相一致的脉冲响应，相对于一个参考频率响应来说，通过第3合成滤波器阵列与第3分析滤波器阵列相结合而形成的一个第3频率响应，使频谱能量的衰减得以优化，以便与从参考频率响应所获得的衰减相比，在第3频率响应的阻带内获得较小的衰减，而在第3频率响应的阻带以外，对频谱邻频干扰提供更大的衰减；以及第2合成滤波器阵列与第2分析滤波器阵列相结合而形成的一个第2频率响应，不同于第1频率响应以及第3频率响应；以及通过使相邻的诸合成样本块互相重叠以及在互相重叠的诸合成样本块中，将对应的互相重叠的诸样本相加来产生一组输出信号的装置。

通过参照以下的讨论和诸附图，将能更好地理解本发明的各种特征及其优选的诸实施例，在若干附图中，相同的参考数字表示相同的元件。说明各种装置的诸附图表示出主要的部件，这对理解本发明是有帮助的。为了明确起见，这些附图省略了在实际的实施例中可能是重要的、但对理解本发明的概念来说是不重要的其他各种特征。为实践本发明所需的信号处理可以通过多种途径来完成，包括由微处理器、数字信号处理器、逻辑阵列以及其他形式的计算电路来执行的程序。实质上可以用任何方法来实现诸信号滤波器，包括递归的、非递归的以及格子数字滤波器。根据该项应用的各种需求和特性，可以按照不同的组合方式来使用数字的和模拟的技术。

关于处理音频和视频信息流的条件将作出更专门的说明，然而，本发明的诸方面可以在不包括视频信息处理的应用场合中来实行。以下的讨论和诸附图的内容仅作为实例而被说明，并且不应当被理解为对本发明的范围施加各种限制。

附图说明：

图1a和1b是被排列成各块、各帧和各超帧的视频和音频信息的概略的表示。

图2a到2c是被各窗口函数调制的各重叠块以及含有各窗口块的诸帧的所得到的增益包络的概略表示。

图3表示由一种混叠抵消变换产生的信号与诸混叠成分。

图4a到4c是表示在一个已编码的信息流中，生成、改变和响应于各增益控制字的诸装置的功能框图。

图5a和5b是采用交替的滤波器阵列来抑制在帧边界处的各种混

叠伪差的功能框图。

图 6a 到 6d 是可以被用来抑制在帧边界处的各种混叠伪差的诸窗口函数的概略表示。

图 7 是在各帧边界处使用各种窗口函数所得到的频率响应特性。

图 8 表示采用交替的滤波器阵列以增加在接合部处的频谱邻频干扰衰减的一个装置的功能框图。

图 9、10a 和 11a 是图 8 的装置的几个窗口函数的概略表示。

图 10b 和 11b 是在图 8 的装置中，通过使用各种窗口函数所得到的诸频率响应特性。

具体实施方式：

诸信号及其处理

各信号块和各帧

图 1a 说明被排列在各音频（数据）块 10 到 18 的一个序列之中的已编码的音频信息流，以及被排列在一个视频帧（例如视频帧 1）序列之中的视频信息流。在某些格式（制式）中，例如 NTSC 视频，每一个视频帧包括两个视频场，它们合起来定义一幅单独的画面或图像。各音频块 11 到 17 跟视频帧 1 组合为一个已编码的信号帧 21。

某些应用具有这样的视频帧，它们不能把已编码的音频整除为整数个样本或诸变换系数之类。通过将已编码的诸信号帧的诸小组排到各自的超帧中，就能适应这种情况。在图 1b 中示出了将 5 个已编码的信号帧 21 到 25 组合为超帧 31 的一种安排。这种特定的安排可以用于使用 NTSC 视频和每秒 48k 个样本的 PCM 音频的应用中。

已处理的诸信号块

已编码的音频信息块的序列可以表示一组音频信号的互相重叠的诸间隔。例如，某些分割频带的感知编码系统处理以半个块长互相重叠的诸音频样本块。典型地，在这些互相重叠的诸块中的诸样本被一个分析窗口函数所调制。

图 2a 表示将一个分析窗口函数施加到一个互相重叠的诸音频块的一个序列中的每一块之后所得到的诸调制包络 61 到 67。重叠的长度等于块长的一半。长度间隔通常被某些信号分析-合成系统，例如上述的 O-TDAC 变换，所使用。

图 2b 表示针对一个已编码的信号帧而施加到互相重叠的各块的一个序列中去的一个窗口函数的所得到的调制包络。如图 2b 所示，这种调制的净效果或增益包络 81 就是在重叠的诸间隔中相邻各块的

调制包络之和。最好是，跨越每一个重叠（块）的净效果应当是单位增益（译者注：即增益 = 1）。

图 2c 表示跨越相邻的已编码信号帧的窗口函数调制的总的效果。如图所示，增益包络 80 到 82 互相重叠并相加，使得净效果为单位增益。

在仅使用分析窗口函数的那些系统中，所有窗口函数调制的净效果等效于单独由分析窗口函数所产生的诸调制效果。通过保证分析窗口函数的调制包络互相重叠和相加为一个常数，就能得到理想的增益包络。

在使用分析和合成诸窗口函数的那些系统中，所有窗口函数调制的净效果等效于由分析窗口函数以及合成窗口函数的一个乘积所形成的一个“乘积”窗口函数。在这样的系统中，通过让在重叠间隔中的乘积窗口函数的调制包络相加为一个常数，就能得到理想的增益包络。

贯穿这里的公开的是，针对使用分析与合成这两种窗口函数的各种编码系统和各种方法来进行叙述。在本文中，从互相重叠的诸分析窗口函数所得到的增益包络有时可以被说成是等于一个常数。类似地，从互相重叠的诸合成窗口函数所得到的增益轮廓有时可以被说成是等于一个常数。应当理解，这样的叙述旨在把问题归结为在该系统中所有窗口的净调制效果。

### 窗口函数

分析窗口函数的形状不仅影响到信号的增益包络，而且它还影响到一个对应的滤波器阵列的频率响应特性。

### 频谱邻频干扰

如上所述，在许多分割频带的感知编码系统中，通过增加在滤波器中的阻带诸频率上的衰减，以取代一个较宽的滤波器通带，从而使用具有为感知编码而优化的各种频率响应特性。遗憾的是，在不属于滤波器阻带范围以内的一个频率范围内，拼接编辑倾向于产生显著的频谱伪差或者“频谱邻频干扰”。被设计用来优化总的感知编码性能的诸滤波器阵列不会提供足够的衰减，使得在拼接编辑过程中所产生的这些频谱伪差变为听不见。

### TDAC 变换混叠抵消

就 O - TDAC 而言，分析窗口函数，连同在应用合成变换之后所施加的一个合成窗口函数一起，还应当满足许多约束条件，以便抵消各种时域混叠伪差。

通过合成变换所恢复的信号可以被概念化为原始信号以及由分析变换所产生的时域混叠成分之和。在图3中，诸曲线91、93和95表示从逆变换或合成变换所恢复的、并且受到分析与合成窗口函数调制的一组输入信号的振幅包络的一些区段。诸曲线92、94和96表示从逆变换或合成变换所恢复的、并且受到分析与合成窗口函数调制的时域混叠成分。正如在图中可以看到的以及将在下面说明的那样，时域混叠成分是由于受到分析与合成窗口函数的调制而反射的原始输入信号的复制品。

分析与合成 O-TDAC 变换的诸核心函数被设计成用以产生各种时域混叠成分，它们是在一个数据块的每一个半块中用窗口截取的信号的端对端的反射物。如 Princen 等所公开的那样，O-TDAC 变换在两个不同的区域中产生时域混叠成分。在区域2，时域混叠成分是在该区域中原始信号的一个用窗口截取的端对端反射物。在区域1，时域混叠成分是在该区域中输入信号的一个用窗口截取的端对端反射物。但该反射物的振幅被倒置。

例如，混叠成分94a是信号成分93a的一个用窗口截取的端对端反射物。除了反射成分的振幅被倒置以外，混叠成分92b也是信号成分91b的一个用窗口截取的端对端反射物。

通过让相邻的各块互相重叠和相加，使原始信号得以恢复，并且各混叠成分得以抵消。例如，诸信号成分91b和93a被叠加，以便恢复没有窗口函数调制影响的信号，并且诸混叠成分92b和94a被叠加，以便抵消混叠。类似地，诸信号成分93b和95a被相加，以便恢复信号，并且诸混叠成分94b和96a被叠加，以便抵消混叠。

由于紧挨着接合部前面的已合成的诸音频样本的半块中的诸混叠伪差不等于紧挨着接合部后面的已合成的诸音频块的半块的诸混叠伪差的倒置产物，所以在—个接合部边界两侧的诸时域混叠伪差通常不能互相抵消。

例如作者 Princen 和 bradley 在《基于时域混叠抵消的分析/合成滤波器阵列设计》一文中所述，类似的考虑也应用于其他的混叠抵消滤波器阵列(见《IEEE 声学、语音和信号处理汇刊》，ASSP-34卷，1986年，第1153-1161页)。这种滤波器阵列系统是一个偶堆栈临界采样单边带分析-合成系统的时域等价物，并且在本文中被称为偶堆栈时域混叠抵消(E-TDAC)。

### 各窗口函数的推导

针对一个优选实施例，根据下列技术来推导分析与合成诸窗口函数。为了便于讨论起见，将这些窗口函数表示为

$W_2(n)$ 。

使用在下列各段中所描述一项技术，就能从一个基本窗口函数推导出窗口函数  $W_2(n)$ 。虽然具有适当的重叠 - 相加特性的任何窗口函数都可以被用来作为基础窗口函数，但是在一个优选实施例中所使用的基础窗口函数是凯撒 - 贝塞尔窗口函数。

$$W_{KB}(n) = \frac{I_0 \left[ \pi \alpha \sqrt{1 - \left( \frac{n}{N/2} \right)^2} \right]}{I_0[\pi \alpha]} \quad (1)$$

上式在  $0 \leq n < N$  的条件下成立

式中， $\alpha$  = 凯撒 - 贝塞尔窗口函数的  $\alpha$  因子，

$n$  = 窗口内的样本数目，

$N$  = 以样本数目来表示的窗口长度，以及

$$I_0[x] = \sum_{k=0}^{\infty} \frac{(x/2)^k}{k!}$$

通过将凯撒 - 贝塞尔窗口函数  $W_{KB}(n)$  跟一个长度等于块长  $N$  减去重叠间隔  $v$  的矩形窗口函数  $s(k)$  进行卷积运算，这样的推导就产生一个分析 - 合成乘积窗口函数  $WP(n)$ ，或：

$$WP(n) = \frac{\sum_{k=0}^{N-1} s(k) W_{KB}(n-k)}{\sum_{k=0}^v W_{KB}(k)}$$

上式在  $0 \leq n < N$  的条件下成立，此式可以简化为：

$$WP(n) = \frac{\sum_{k=0}^{N-v-1} W_{KB}(n-k)}{\sum_{k=0}^v W_{KB}(k)}$$

上式在  $0 \leq n < N$  的条件下成立

式中,  $n$  = 乘积 - 窗口内的样本数目,  
 $v$  = 在窗口重叠间隔内的样本数目,  
 $N$  = 所希望的乘积 - 窗口长度,  
 $W_{KB}(n)$  = 长度为  $v+1$  的基本窗口函数,  
 $WP(n)$  = 推导出来的长度为  $N$  的乘积-窗口, 以及  
 $s(k) = 1$  对条件  $0 \leq k < N-v$  成立  
 $= 0$  对其他条件成立。

对于 O-TDAC 变换来说, 重叠间隔  $v = N/2$ , 并且分析窗口函数以及合成窗口函数都是如此, 因此, 可以从下式得到其中任何一个窗口函数:

$$W_2(n) = \sqrt{\frac{\sum_{k=0}^{N/2-1} W_{KB}(n-k)}{\sum_{k=0}^{N/2} W_{KB}(k)}} \quad (2)$$

上式在  $0 \leq n < N$  的条件下成立

在本文中, 以这种方式推导出来的分析与合成窗口函数被称为一个凯撒-贝塞尔-导出 (KBD) 窗口函数。乘积窗口函数被称为一个 KBD 乘积窗口函数。可以选择基本凯撒-贝塞尔窗口函数的  $\alpha$  因子以优化编码性能。在许多应用中, 用于编码的最佳  $\alpha$  因子处于 2 到 6 的范围内。

用以降低接合部的频谱邻频干扰的诸滤波器阵列

从优化感知编码这个意义上来说, 处于上述范围内的  $\alpha$  因子对许多编码应用来说是最佳的。如上所述, 通常用增加在滤波器阻带内诸频率上的衰减, 来取代一个较宽的滤波器通带, 从而使编码得以优化。图 7 中的曲线 342 表示针对一个为感知编码而优化的一种典型的频率响应的一个例子。这条曲线表示一个使用 KBD 窗口函数的 O

-TDAC 分析 -合成系统的帧增益包络的频率响应，上述 KBD 窗口函数的  $\alpha = 6$ ，并且具有等于 256 个样本的帧重叠间隔。虽然介于通带和阻带之间的边界没有明确地规定，但在本例中通带覆盖着高达 200 Hz 的诸频率，并且阻带覆盖着大约 1kHz 以上的诸频率。一个过渡区域在这两个频带之间延伸。

在使用施加于 256 个样本块的变换的各种应用中，在约为 200 Hz 到 1 kHz 的滤波器的中心频率范围内，拼接编辑倾向于产生显著的假频谱成分或“频谱邻频干扰”。对于使用其他长度的（数据）块的应用来说，这个频率范围可以表示为两个常数除以块的长度；因此，显著的频谱邻频干扰出现在从大约 50,000 到 256,000 Hz（每一个都除以块长度）的频率范围内。

在图 7 所示的例子中，这些频率处于被认为是滤波器阻带的范围以外。被设计用于优化感知编码性能的诸滤波器阵列对在拼接编辑过程中所生成的频谱邻频干扰未能提供足够的衰减。这些伪差通常是可听的，因为它们通常是如此之大，以致于无法被信号所遮盖。

图 7 中的曲线 341 和曲线 343 表示两个其他的分析-合成系统的频率响应，该系统在阻带内提供显著地小的衰减，但在一个受到在接合部生成的频谱邻频干扰的影响的频率范围内，提供更大的衰减。通过牺牲在感知编码中的某些性能来增加对频谱邻频干扰的衰减。最好是，在一个对 256 个样本块进行滤波的系统中，在包括 200 Hz 和 600 Hz 的一个频率范围内，或者在大约 50,000 到 150,000 Hz 的频率范围内，每一个都除以块长度，该频率响应使频谱能量的衰减得以优化。

有时，在满足对一般编码以及对在接合部处的交叉衰落各帧的频率响应要求之间达到一种折衷。在不能达到这样一种折衷的各种应用中，要检出一个接合部，并且改变分析-合成系统的频率响应。由于分析滤波器阵列一般地不能预测拼接操作，所以这种改变应当结合合成滤波过程来完成。

图 8 表示通过改变一个分析-合成系统的端到端频率响应，装置

320 可以被用来减少频谱邻频干扰。在这个装置中，去格式器 322 从通路 321 接收一组输入信号，从中获得沿着通路 324 传送的已编码的音频信息，并沿着通路 323 产生一组控制信号，用以指示在一帧的开始或终止处，是否出现一个接合部。一个接合部的出现可以在输入信号中明确地表达，或者可以根据在信号中所表达的其他信息来推断。

例如，根据 AES-3/EBU 标准，连续的音频信息块所含有的块号码从 0 增加到 255，然后反过来减少到 0。两个邻接的而不按顺序的块号码表示一个接合部，然而，由于某些装置在处理 AES/EBU 数据流时并不增加这个号码，所以这种检验方法是不可靠的。若该音频流已被编码，则编码方案可能提供顺序的编号或者某些其他形式的可预测的信息。若该信息跟所预期的不相符，则可能产生一组信号以表明一个接合部的存在。

响应于从通路 323 接收的控制信号，开关 325 将已编码的音频流引导到 3 个合成滤波器阵列中的一个。开关 325 将跟随在一个接合部后面的一帧中的对应于第 1 块的已编码的音频信息引导到第 1 合成滤波器阵列 326，将前置于一个接合部前面的一帧中的对应于最后一块的已编码的音频信息引导到第 3 合成滤波器阵列 328，以及将对应于其他各块的已编码音频信息引导到第 2 合成滤波器阵列 327。另一方面，也可以根据以下结合图 5b 所讨论的技术，将对应于这些其他各块的已编码的音频信息引导到 3 个滤波器阵列中的一个。响应于从这 3 个合成滤波器阵列所接收的已合成的各音频块，缓冲存储器 329 沿着通路 330 产生一组输出信号。

第 1 和第 3 合成滤波器阵列被设计成这样，使其在与某些分析滤波器相结合的条件下，能获得所需的频率响应。在许多应用中，这个分析滤波器被设计成这样，使其在与第 2 合成滤波器相结合的条件下，能优化总的编码性能。基本上可以通过能提供所需的总的频率响应的任何方式来实现第 1 和第 3 合成滤波器阵列。一般来说，这两个滤波器阵列将具有相同的频率响应，但将具有在时序上互为倒置的复制品那样的脉冲响应。在使用诸变换与诸窗口函数来实现诸滤波器

阵列的那些应用中，通过使用在一个接合部两侧相邻的各帧之间增加重叠间隔的合成窗口函数，就能实现适当的滤波器阵列。

### 已合成音频的调制

可以通过几种方法来完成这一步。一种方法就是对从合成滤波器阵列中恢复的已合成音频信号进行调制，使得在一个接合部两侧的各帧通过交叉衰落融入对方。可以在一个装置（例如将在下面讨论并示于图 4c 的装置 140）中完成这一步。解码器 146 降低在前置于接合部并跨越所需的接合部重叠间隔的帧中的已合成信号的幅度。实际上，在跨越这个间隔时，前置于接合部的帧的增益包络从 1 减少到某个较低的水平。解码器 146 还增加在跟随着接合部并跨越所需的接合部重叠间隔的帧中的已合成信号的幅度。实际上，在跨越这个间隔时，跟随着接合部的帧的增益包络从较低的水平增加到 1。若在各增益包络的有效变化中考虑到分析-合成窗口的调制效应，则重叠的各帧的总的增益得以保留。

增益包络的有效变化可以是直线性的。图 7 中的曲线 343 表示一个宽度约为 5 毫秒的线性递减的帧增益包络的频率响应特性。在每秒 48 k 个样本的采样率下，这个间隔对应于大约 256 个样本。在许多编码应用中，对具有 256 个样本的样本块实施变换；因此，在这些特定的应用中，一个含有 256 个样本的斜坡或线性递减的增益包络在延伸，它越过处于帧边界处的一个“終了”块，并且跨越一个跟这个終了块重叠的相邻块的一部分。这等效于将一个滤波器阵列施加于終了块，经另一个滤波器阵列施加于紧挨着的相邻块，以及将又一个滤波器阵列施加于该帧内部的其他各块。参看图 8 所示的装置 320，它需要两个附加的合成滤波器阵列，用以处理邻接于并重叠于“終了”块的各块。

这个线性递减的斜坡的频率响应表示一个参照的响应，其他各种频率响应都可以相对于它而作出评估。一般来说，在减少产生于接合部的频谱邻频干扰方面，能优化针对这种参考响应的频谱能量衰减的

各种滤波器阵列都是有效的。

### 修正的合成窗口函数

用以改变一个分析-合成系统的总的频率响应特性的另一种方法就是修改合成窗口函数,使得分析-合成窗口的净效应获得所需的响应。实际上,总的频率响应根据所得到的分析-合成乘积窗口函数而发生改变。

图 7 中的曲线 341 表示一种频率响应,它跟曲线 343 所表示的 5 毫秒的线性递减增益包络的频率响应相比,能在更大程度上衰减在各接合部的频谱邻频干扰。借助于 O-TDAC 分析-合成系统,使用 256 点变换以及  $\alpha = 1$  的 KBD 窗口函数来获得曲线 341 的响应。如上所述,曲线 342 对应于  $\alpha = 6$  的 KBD 窗口函数。

这些分析-合成系统的端到端频率响应等效于由分析窗口函数与合成窗口函数的乘积所形成的窗口的频率响应。用代数方法可以表示为:

$$WP_6(n) = WA_6(n) WS_6(n) \quad (3a)$$

$$WP_1(n) = WA_1(n) WS_1(n) \quad (3b)$$

式中,  $WA_6(n)$  为  $\alpha = 6$  时的分析 KBD 窗口函数,

$WS_6(n)$  为  $\alpha = 6$  时的合成 KBD 窗口函数,

$WP_6(n)$  为  $\alpha = 6$  时的 KBD 乘积窗口函数,

$WA_1(n)$  为  $\alpha = 1$  时的分析 KBD 窗口函数,

$WS_1(n)$  为  $\alpha = 1$  时的合成 KBD 窗口函数,以及

$WP_1(n)$  为  $\alpha = 1$  时的 KBD 乘积窗口函数。

若一个合成窗口函数被修改成将端到端频率响应修改为所需的某种其他响应,则它应当被这样修改,使得它本身与分析窗口函数的乘积等于具有所需响应的乘积窗口。若希望得到对应于  $WP_1$  的一个频率响应,并且分析窗口函数  $WA_6$  被用于信号分析,则这种关系可以用代数方法表示为:

$$WP_1(n) = WA_6(n) WX(n) \quad (3c)$$

式中， $WX(n)$  = 为转换频率响应所需的合成窗口函数。

上式可以写成：

$$WX(n) = \frac{WP_1(n)}{WA_6(n)} \quad (3d)$$

若接合部重叠间隔延伸到在帧中跟“终了”块相重叠的一个相邻的音频块之上时，则窗口函数  $WX$  的形状跟表达式 3d 所表示的形状相比，显得更复杂一些。下面将对此进行更充分的讨论。在任何情况下，表达式 3d 精确地表示了对在终了块中不跟该帧中任何其他块重叠的那一部分中的窗口函数  $WX$  的要求。对于使用 O-TDAC 的诸系统来说，该部分等于块长的一半，或对应于  $0 \leq n < N/2$ 。

若合成窗口函数  $WX$  被用来将端到端频率响应从较高的  $\alpha$  轮廓转换到一个较低的  $\alpha$  轮廓，则在接近帧边界处，它必须具有非常大的数值。图 9 示出了一个实例，在其中曲线 351 表示一个  $\alpha = 1$  的 KBD 分析或合成窗口函数，曲线 352 表示一个  $\alpha = 1$  的 KBD 乘积窗口（函数）。曲线 356 表示一个  $\alpha = 6$  的 KBD 分析或合成窗口函数，以及曲线 359 表示一个根据表达式 3d 的合成窗口函数。随着曲线 356 接近帧边界，它变得比曲线 352 小很多，因此，曲线 359 变得非常大。不幸的是，一个具有类似于曲线 359（它在窗口函数  $WX$  的边缘处具有大的增加）的形狀的合成窗口函数具有很差的频率响应特性，并将使被恢复信号的音质恶化。有两种技术可以被用来解决这个问题，这将在下面加以讨论。

### 舍弃样本

用于修正一个合成窗口函数的第 1 种技术通过在分析窗口函数具有最小值的帧边界处舍弃某一数目的样本来避免在窗口函数  $WX$  中的大量增加。通过改变被舍弃的样本的数目，就能调整为在帧重叠间隔中传送诸样本所需的带宽，从而抵消由于解码器的不良的频率响应特性而引起的系统编码性能的下降。

例如，通过修正对应于一帧中的前3块的各合成窗口函数，以获得一种对应于乘积窗口函数  $WP_1$  的所需的频率响应，以及用于信号分析的窗口函数为  $WA_6$ ，则所需的修正合成窗口函数如下式所示：

$$WX1(n) = \begin{cases} 0 & \text{在 } 0 \leq n < x \text{ 条件下成立} \\ \frac{WP_1(n-x)}{WA_6(n)} & \text{在 } x \leq n < N/2 \text{ 条件下成立} \\ WP_1(n-x) WA_6(n) & \text{在 } N/2 \leq n < N \text{ 条件下成立} \end{cases} \quad (4a)$$

$$WX2(n) = \begin{cases} WP_1(n-x + \frac{N}{2}) WA_6(n) & \text{在 } 0 \leq n < N/2 + x \text{ 条件下成立} \\ WA_6(n) & \text{在 } N/2 + x \leq n < N \text{ 条件下成立} \end{cases} \quad (4b)$$

$$WX3(n) = \begin{cases} WP_1(n-x+N) WA_6(n) & \text{在 } 0 \leq n < x \text{ 条件下成立} \\ WA_6(n) & \text{在 } x \leq n < N \text{ 条件下成立} \end{cases} \quad (4c)$$

式中，

$WX1(n)$  = 对应于第1块的修正的合成窗口函数，

$WX2(n)$  = 对应于第2块的修正的合成窗口函数，

$WX3(n)$  = 对应于第3块的修正的合成窗口函数，以及

$x$  = 在帧边界处被舍弃的样本数目。

图 10a 表示，对于  $x$  的若干数值来说，修正的合成窗口函数的形状要求使用一个  $KBD \alpha = 6$  的分析窗口函数将一个 256 点 O-TDAC 分析 - 合成系统转换为一个这样的分析 - 合成系统，后者的频率响应等效于一个使用  $KBD \alpha = 1$  的、帧重叠间隔等于 256 个样本的分析

与合成窗口函数。曲线 361, 362, 363 和 364 分别是对应于  $x = 8, 16, 24$  和  $32$  个样本的修正的各合成窗口函数。

使用这些修正的窗口函数的各合成滤波器阵列的频率响应示于图 10 b。诸曲线 372, 373 和 374 分别是对应于  $x = 8, 16$  和  $24$  个样本的频率响应。曲线 371 是使用一个  $\alpha = 1$  的 KBD 窗口函数的一个合成滤波器阵列的频率响应。如图所示, 一个  $x = 16$  的修正的合成窗口函数将大约  $200\text{Hz}$  以上的各频率衰减到跟使用  $\alpha = 1$  的 KBD 窗口函数的合成滤波器阵列所获得的衰减相同的程度。换句话说, 一个舍弃了  $16$  个样本的合成滤波器阵列, 当跟一个分析滤波器阵列以及一个  $\alpha = 6$  的分析窗口函数配合使用时, 它所获得的端到端分析-合成系统频率响应等效于一个使用  $\alpha = 1$  的分析与合成窗口函数的系统的端到端频率响应, 并且, 与此同时, 提供这样一种合成滤波器阵列频率响应, 它对大约  $200\text{Hz}$  以上的各频率的衰减程度接近于一个使用  $\alpha = 1$  的合成窗口函数的合成滤波器阵列的衰减程度。

一般来说, 将具有较低  $\alpha$  数值的 KBD 窗口函数用于正常编码的各种系统将要求对合成窗口函数进行较小的修正, 并且在帧的末尾舍弃较少的样本。在一帧的末尾处所需要的修正的合成窗口函数类似于在表达式 4a 到 4c 所示的各窗口函数, 所不同的是发生了时序倒置。

#### 对帧增益包络进行调制

用于修正一个合成窗口函数的第 2 种技术通过在一个接合部的两侧允许帧增益包络稍为偏离理想水平来避免在窗口函数  $WX$  中的大量增加。通过改变增益包络的偏离量, 就能用该偏离量的可听度来抵消频谱邻频干扰的可听度。

这种技术使修正的合成窗口函数变得平滑, 因此它在帧边界处或靠近贴边界处具有小的数值。当适当地完成这一步时, 所得到的合成窗口函数将具有可接受的频率响应, 并且在帧边界或靠近帧边界处的增益相当低, 帧增益包络将偏离理想的 KBD 乘积窗口函数。跟由一种理想的交叉衰落的增益形状所提供的衰减相比, 频谱邻频干扰的

衰减将仅有轻微地降低。

例如，若必须修正对应于一帧中前3块的合成窗口函数以获得所需的频率响应，则为第2和第3块所需的修正的合成窗口函数  $WX$  一般地相同于在  $x=0$  条件下由上面的表达式 4b 和 4c 所表示的情形。通过在平滑窗口函数长度的前半部逐点地跟一个平滑窗口函数相乘，就能使在上面的表达式 4a 中所表示的修正的合成窗口函数  $WX1$  变得平滑。所得到的对应于第1块的修正的合成窗口函数为：

$$WX1(n) = \begin{cases} \frac{WP_1(n)WM(n)}{WA_6(n)} & \text{在 } 0 \leq n < P/2 \text{ 条件下成立} \\ \frac{WP_1(n)}{WA_6(n)} & \text{在 } P/2 \leq n < N/2 \text{ 条件下成立} \\ WP_1(n)WA_6(n) & \text{在 } N/2 \leq n < N \text{ 条件下成立} \end{cases} \quad (5)$$

式中， $WM(n)$  = 平滑窗口函数，以及

$P$  = 平滑窗口函数的长度，假设小于  $N$ 。

在一帧终了处所需的修正的合成窗口函数也相同于这个窗口函数，所不同的是时序发生倒置。

平滑窗口函数  $WM$  基本上可以基于任何窗口函数，但是，一个  $KBD$  平滑窗口函数似乎工作得更好一些。在本例中，平滑窗口函数是一个长度为 128、 $\alpha = 6$  的  $KBD$  窗口函数。在图 11a 中，曲线 381 表示未经平滑的修正的合成窗口函数的形状，而曲线 382 则表示经平滑后的修正的合成窗口函数的形状。

图 11b 表示对应于一个使用平滑后的修正的窗口函数的分析-合成系统的频率响应。曲线 391 表示使用平滑后的修正的窗口函数的频率响应。曲线 341 表示使用  $\alpha = 1$  的  $KBD$  窗口函数的一个分析-合成系统的频率响应，曲线 393 则表示使用上面所讨论的、并示于曲线 343 的、宽度约为 5 毫秒的线性递减帧交叉衰落窗口函数所得到的频率响应的诸峰值的一个包络。从这份图中可以看出，一个经过平滑的

修正的合成窗口函数获得一个类似于通过一个使用  $\alpha = 1$  的 KBD 窗口函数的分析-合成系统所获得的频率响应。

### 混合分析-合成窗口函数的修正

在以上所讨论的各项技术中，帧增益包络的所有变化都在信号合成过程中发生。作为另一种方法，分析过程也可以对处于帧边界处的各块使用具有一种频率响应的滤波器阵列，并且对内部各块使用另一种滤波器阵列。用于在帧边界处各块的各滤波器阵列可以被设计成减少在合成过程中所需的修正量，以便获得对产生于接合部的频谱邻频干扰的足够的衰减。

### 用以在帧边界处抑制混叠的诸滤波器阵列

在使用一种混叠抵消（例如一种由 TDAC 变换所提供的）形式的编码系统中，由于上面所讨论的原因，诸拼接编辑妨碍了在接合部每一侧的诸混叠伪差的互相抵消。通过在每一帧的开始和结尾对各音频块施加交替的滤波器阵列，就能避免这些未被抵消的混叠伪差。例如，参照图 1a 的帧 21，一个第 1 滤波器阵列被施加于块 11，一个第 2 滤波器阵列被施加于块 12 到 16，以及一个第 3 滤波器阵列被施加于块 17。这些滤波器阵列具有这样的特性，使得从每一帧所恢复的音频基本上不含有未被抵消的诸混叠伪差。在这里所描述的用于在帧边界处抑制混叠的各项技术是一项共同未决的专利申请的主题；然而，它们可以有利地跟上面所讨论的用于减少频谱邻频干扰的各项技术组合在一起使用。

参照图 5a，装置 200 包括缓冲存储器 202，它接收各音频信息块，并且沿着通路 203 产生一个控制信号，以表明一个音频块是一帧中的第 1 或起始块，一帧中的最后或终了块，或者是一帧中的一个临时块。响应于从通路 203 接收的控制信号，开关 204 将每一帧中的第 1 或起始块引导到第 1 滤波器阵列 205，将每一帧中的所有临时块引导到第 2 滤波器阵列 206，将每一帧中的最后或终了块引导到第 3

滤波器阵列 207。格式器 208 将从这 3 个滤波器阵列中的每一个所接收的已滤波的音频信息组合成一组输出信号，并沿着通路 209 传送。

图 5b 表示装置 220，在其中，去格式器 222 从通路 221 接收一组输入信号，从中获得已编码的音频信息，它沿着通路 224 被传送。并沿着通路 223 产生一组控制信号，以表明已编码的音频信息是一帧中的第 1 或起始块，一帧中的最后或终了块，或者是一帧中的一个临时块。响应于从通路 223 接收的控制信号，开关 225 将已编码的音频信息引导到 3 个合成滤波器阵列中的一个。开关 225 将对应于第 1 块的已编码的音频信息引导到第 1 合成滤波器阵列 226，将对应于各临时块的已编码音频信息引导到第 2 合成滤波器阵列 227，将对应于最后块的已编码的音频信息引导到第 3 合成滤波器阵列 228。响应于从 3 个合成滤波器阵列接收的已合成的各音频块，缓冲存储器 229 沿着通路 230 产生一组输出信号。

## 第 2 滤波器阵列

在一个编码器的一个实施例中，根据上面所引用的 Princen 等所公开的 O-TDAC 变换，通过一种 N 点修正的离散余弦变换 (DCT) 以及一个 N 点分析窗口函数来实现第 2 滤波器阵列。在一个互补的解码器中，根据 O-TDAC 变换，通过一种 N 点修正的反向 DCT 以及一个 N 点合成窗口函数来实现第 2 滤波器阵列。正向和反向的 O-TDAC 变换分别示于表达式 6 和 7:

$$X(k) = \sum_{n=0}^{M-1} x(n) \cos \left[ \frac{2\pi}{M} \left( k + \frac{1}{2} \right) \left( n + \frac{m+1}{2} \right) \right] \quad \text{for } 0 \leq k < M \quad (6)$$

$$x(n) = \frac{1}{M} \sum_{k=0}^{M-1} X(k) \cos \left[ \frac{2\pi}{M} \left( k + \frac{1}{2} \right) \left( n + \frac{m+1}{2} \right) \right] \quad \text{for } 0 \leq n < M \quad (7)$$

式中，k = 频率指标，

n = 信号样本数，

M = 样本块长度，

$m$  = 用于 O - TDAC 的相位项,

$x(n)$  = 被窗口截取的输入信号样本数  $n$ , 以及

$X(k)$  = 变换系数  $k$ 。

如图 3 所示, 第 2 滤波器阵列的长度  $M = N$ , 并且生成两个混叠反射区域, 在一个块的中点处有一段介于这两个区域之间的边界。为生成这两个区域所需的 TDAC 相位项为  $m = N/2$ 。

在一个优选实施例中, 根据前面所述的技术来导出分析与合成诸窗口函数。在图 6a 中, 通过曲线 242 来说明这些窗口函数的形状。为了便于讨论起见, 这些窗口函数被称为  $W_2(n)$ 。

### 第 1 滤波器阵列

在同一实施例中, 通过上面所示的修正的 DCT 以及窗口函数  $W_2(n)$  的一种修正形式来实现在编码器和互补的解码器中的第 1 滤波器阵列。正向和逆向的 O - TDAC 变换分别示于表达式 6 和 7。第 1 滤波器阵列的长度为  $M = 3N/2$ , 并且生成一个单独的混叠反射区域 1。诸混叠伪差是在该块内信号的倒置的端到端反射。实际上, 反射区域 2 的长度为 0, 并且介于这两个区域之间的边界位于该块的右边缘的前沿。为生成这个单独区域所需的 TDAC 相位项为  $m = 0$ 。

对应于诸第 1 滤波器阵列的分析与合成窗口函数  $W_1(n)$  与此相同。在图 6b 中, 通过曲线 241 来说明这些窗口函数的形状。它由 3 部分组成。第 1 和第 2 部分表示为区段 1 和 2, 跟上面所描述的并示于图 6a 的窗口函数  $W_2(n)$  相同。第 3 部分表示为区段 3, 它等于 0。

第 1 分析窗口函数  $W_1(n)$  保证在区段 3 中的信号为 0。其结果是, 从区段 3 反射到区段 1 的混叠伪差也是 0。从区段 1 反射到区段 3 的混叠伪差一般地将不是 0; 然而, 当向合成的音频块施加第 1 合成窗口函数  $W_1(n)$  时, 反射到区段 3 的任何伪差将被消除。其结果是, 混叠伪差仅存在于区段 2。

### 第3滤波器阵列

在同一实施例中，通过上面所示的修正的 DCT 以及窗口函数  $W_2(n)$  的一种修正形式来实现在编码器和互补的解码器中的第3滤波器阵列。正变换和诸反变换分别示于表达式 6 和 7。第3滤波器阵列的长度为  $M = 3N/2$ ，并且生成一个单独的混叠反射区域 2。混叠伪差是在该块内信号的一种端到端反射。实际上，反射区域 1 的长度为 0，并且介于这两个区域之间的边界位于该块的左边缘的后沿。为生成这个单独区域所需的 TDAC 相位项为  $m = 3N/2$ 。

对应于第3滤波器阵列的分析与合成窗口函数  $W_3(n)$  与此相同。在图 6c 中，通过曲线 243 来说明这些窗口函数的形状。它由 3 部分组成。表示为区段 1 的第 1 部分等于 0。第 2 和第 3 部分，表示为区段 2 和 3，跟上面所描述的并示于图 6a 的窗口函数  $W_2(n)$  相同。

第3分析窗口函数  $W_3(n)$  保证在区段 1 中的信号为 0。其结果是，从区段 1 反射到区段 3 的混叠伪差也是 0。从区段 3 反射到区段 1 的混叠伪差一般地将不是 0；然而，当向合成的音频块施加第3合成窗口函数  $W_3(n)$  时，反射到区段 1 的任何伪差将被消除。其结果是，混叠伪差仅存在于区段 2。

图 6d 说明诸窗口函数  $W_1(n)$ ， $W_2(n)$ ， $W_3(n)$  241 到 243 是如何互相重叠的。增益包络 240 表示以端到端方式使用窗口函数的净效果，对 TDAC 来说，它是由对应的分析与合成诸窗口函数的乘积所形成的重叠的乘积窗口函数的一个序列。由分析-合成窗口函数  $W_1(n)$  所加权的、在块 11 的区段 2 中的混叠伪差被由分析-合成窗口函数  $W_2(n)$  所加权的、在块 12 的前半部中的混叠伪差所抵消。由分析-合成窗口函数  $W_3(n)$  所加权的、在块 17 的区段 2 中的混叠伪差被由分析-合成窗口函数  $W_2(n)$  所加权的、在块 16 的后半部中的混叠伪差所抵消。根据常规的 TDAC 来完成在临时块对（例

如块 12 和 13 或者块 15 和 16) 中的信号恢复与混叠抵消。

通过使用此项技术, 就可以在任何帧边界处进行拼接编辑, 并且不会留下未被抵消的各种混叠伪差。

在全帧范围内都没有未被抵消的混叠伪差这一事实使得任何窗口函数基本上都可以用于一个接合部。一般来说, 这些窗口函数都具有这样一种形状, 使得它在跨越重叠的间隔时能保持一个恒定的增益轮廓。在接合部, 重叠间隔可以延伸到跨越许多帧; 然而, 在许多应用中, 希望使用一种处于 5 到 30 毫秒范围内的“接合部-重叠间隔”。由于将在下面讨论的理由, 重要的是, 跨越一个接合部的重叠间隔还可以延长。

#### 用以衰减在接合部的各种伪差的增益控制

一种可以用来降低由一个接合部产生的各种伪差的可听度的技术就是将多个增益控制字纳入到一组已编码的音频信号之中, 上述增益控制字指示一个解码器或回放系统去改变回放信号的幅度。在以下各段中将讨论使用这些控制字的诸装置的简单的诸实施例。在这里所描述的关于使用增益控制字的各项技术是一项共同未决的专利申请的主题, 然而, 它们可以有利地跟上面所讨论的用于减少频谱邻频干扰的各项技术组合在一起来使用。

图 4a 表示装置 100 的一份功能框图, 在其中, 格式器 111 沿着通路 112 产生一组被排列为诸帧的输出信号, 该信号包括视频信息、代表多条音频信道的已编码的音频信息, 以及各增益控制字。响应于从通路 108 接收的一组信号 (该信号被排列为运送视频信息以及针对多条音频信道的已编码的音频信息的诸帧的形式), 并且响应于从运送各增益控制字的通路 110 接收的一组信号, 格式器 111 产生输出信号。过程 109 从诸通路 103a 和 103b 接收多组控制信号, 其中的每一组都跟多条音频信道中的一条有关, 并且响应于每一组控制信号, 沿着通路 110 为一条有关的音频信道产生一对增益控制字, 它们表示在一个各自的帧里面的一个起始增益以及一个终了增

益。为了简明起见，在图中仅示出了两组控制信号 103 和两条有关的音频信道 102。必要时，这种增益控制技术可以应用于两条以上的声道。

在所示的实施例中，响应于从通路 102a 和 102b 接收的多组音频信道信号，编码器 105 沿着通路 106a 和 106b 为多条音频信道产生已编码的音频信息，并且通过以帧的形式来排列从通路 101 接收的视频信息以及从通路 106a 和 106b 接收的已编码的音频信息，成帧器 107 沿着通路 108 产生该信号。

增益控制技术可以用于类似于沿着通路 108 通过的信号的诸输入信号，因此，用不着编码器 105，也用不着成帧器 107。在含有编码器 105 的各实施例中，可以独立地对每一条音频信道进行编码，也可以联合地对多条音频信道进行编码。例如，可以联合地对两条或多条声道使用 AC-3 编码技术，通过取消或降低介于各声道之间的冗余度来降低总的带宽要求。

图 4c 表示装置 140 的一个功能框图，装置 140 根据在一组输入信号中的增益控制字来产生诸输出信号，以便再现或回放多条音频信道的信息。去格式器 142 从通路 141 接收一组被排列为帧的形式的输入信号，其中包括视频信息、已编码的音频信息，以及各增益控制字。去格式器 142 从每一帧输入信号中获得代表多条音频信道的已编码的音频信息，以及获得跟每一条音频信道有关的一对增益控制字。过程 148 从通路 145 接收各增益控制字，并且作为响应，沿着通路 149a 和 149b 产生各增益控制字。解码器 146 从通路 144a 和 144b 接收多条声道的已编码的音频信息，并且作为响应，为每一条音频信道产生一组输出信号，因此，响应于一组有关的增益控制信号，每一组输出信号的幅度或电平将发生改变。

一对增益控制字表示在一个特定帧里面针对一条有关的音频信道的一个起始增益和一个终了增益。步骤 148 产生表示这对增益控制字的一个内插值的增益控制诸信号。可以遵循任何所希望的轨迹，例如直线、二次曲线、对数或指数曲线来进行内插。例如，在直线

性内插的情况下，增益控制信号将代表在一个特定帧的范围内按线性规律发生变化的一个增益。

可以独立地对每一条音频信道进行解码，也可以联合地对多条音频信道进行解码。例如，解码过程对那些取消或降低介于各声道之间的冗余度的编码过程的各种形式来说，起到互补的作用。在使用一个合成滤波器阵列以及一个合成窗口函数的分割频带编码应用中，通过在应用合成滤波器阵列之前修改已编码的音频，通过在应用合成窗口函数之前修改从合成滤波器阵列中获得的合成音频，或者通过修改从应用合成窗口函数而获得的音频信息，就能根据一个增益控制信号来有效地调制输出信号。

图 4b 表示用以修改现有的各增益控制字的装置 120 的一份功能框图。去格式器 123 从通路 121 接收一组排列成诸帧的输入信号，其中包括视频信息、代表多条音频信道的已编码的音频信息，以及各输入控制字。去格式器 123 从输入信号中获得跟针对一条或多条音频信道的已编码音频信息有关的一个或多个输入增益控制字，并沿着通路 124a 和 124b 传送诸输入增益控制字。响应于从通路 122 接收的一组控制信号，步骤 126 通过修改一个或多个输入增益控制字，沿着通路 127 产生一个或多个输出增益控制字。格式器 128 沿着通路 129 产生一组被排列成诸帧的输出信号，其中包括视频信息、针对多条音频信道的已编码的音频信息，输出增益控制字，以及跟各输出增益控制字不相符合的各输入控制字。

在一种编辑应用中，控制信号 122 指示在输入信号中的一个接合部。作为响应，步骤 126 产生一个或多个输出增益控制字，它(们)将使得一个装置，例如装置 140，在紧挨着接合部的前面去衰减一组回放信号，并且紧挨着接合部的后面，让衰减量按相反方向变回来。增益的改变可以延伸到跨越若干帧，然而，在许多应用中，这种改变被限制在接合部任何一侧的 1 帧上。通过平衡由增益改变乘以增益改变本身的可听度所产生的调制产物的可听度，就能确定增益改变的间隔。增益控制字技术不局限于编辑应用。

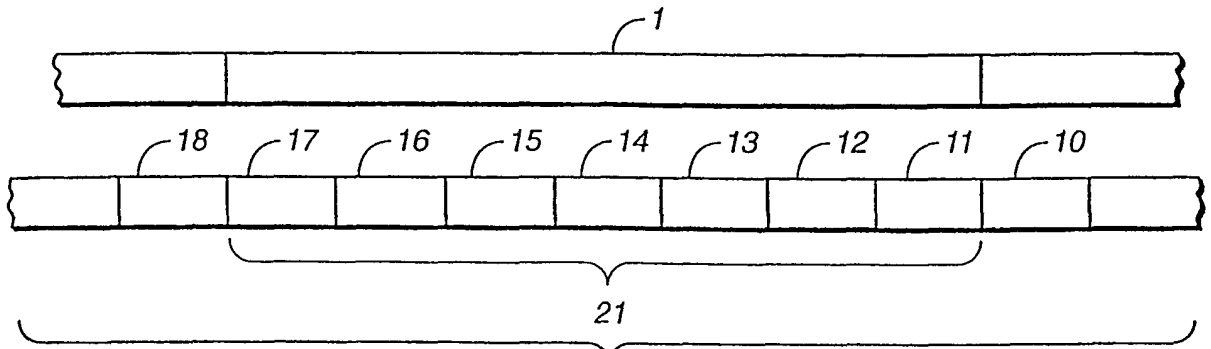


图1a

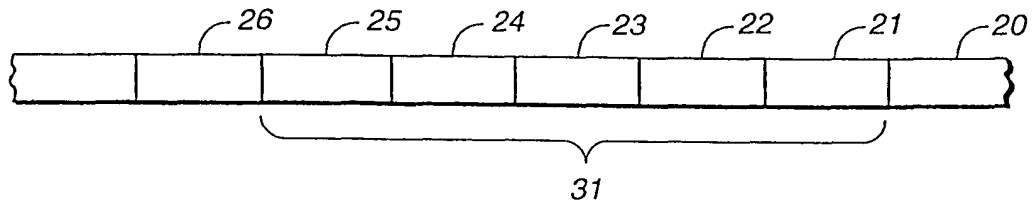


图1b

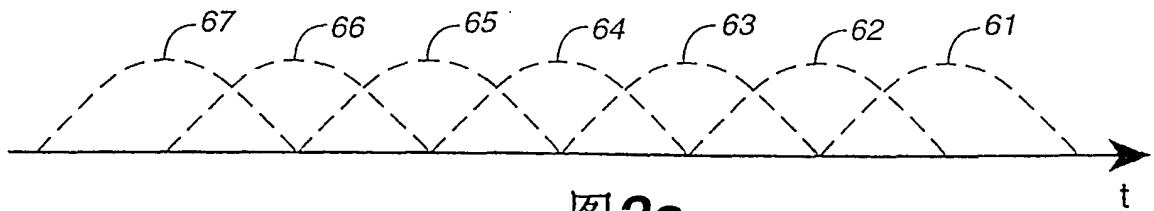


图2a

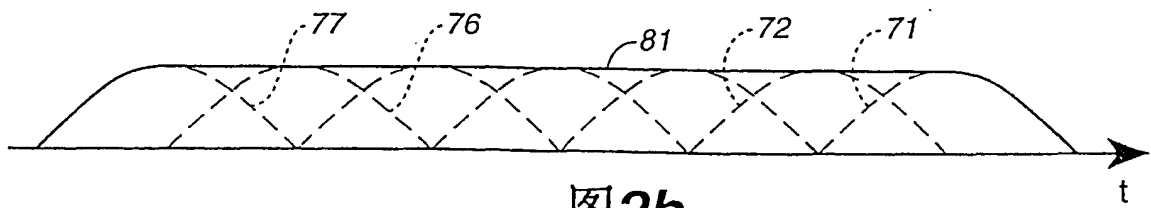


图2b

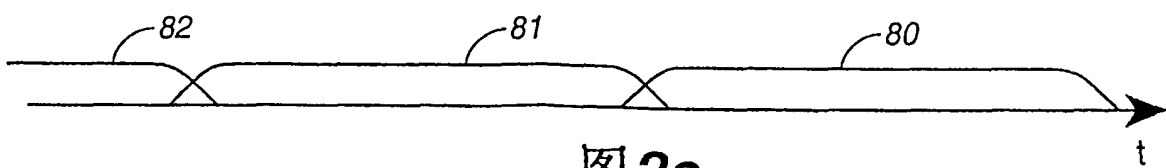


图2c

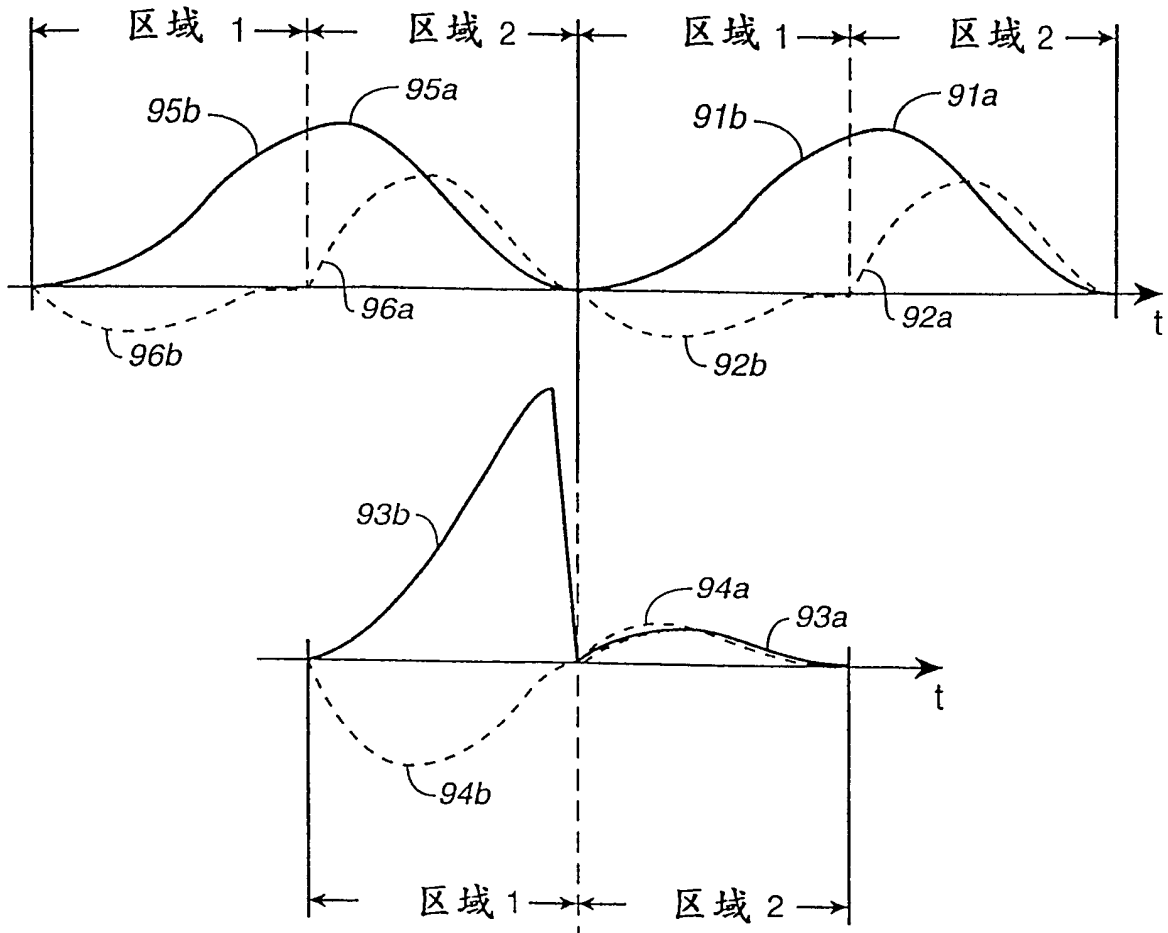


图3

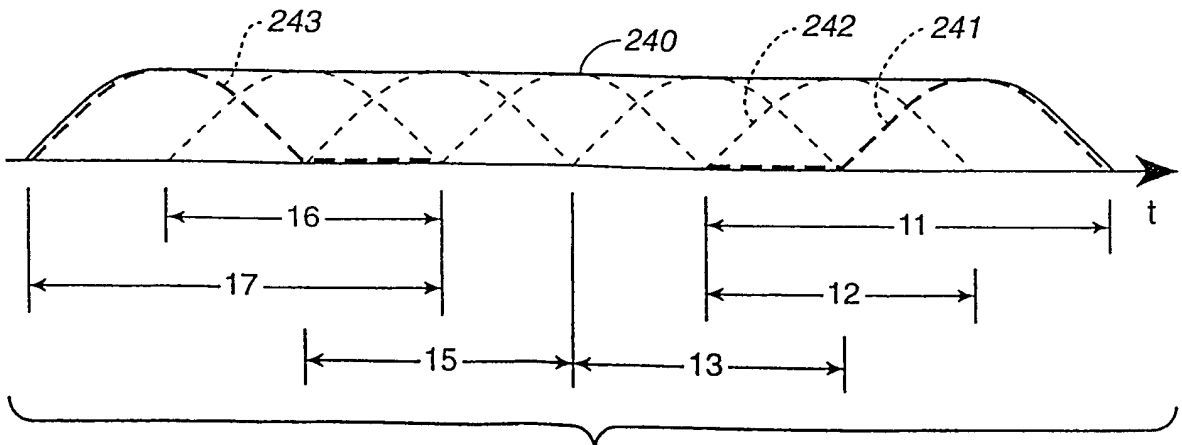
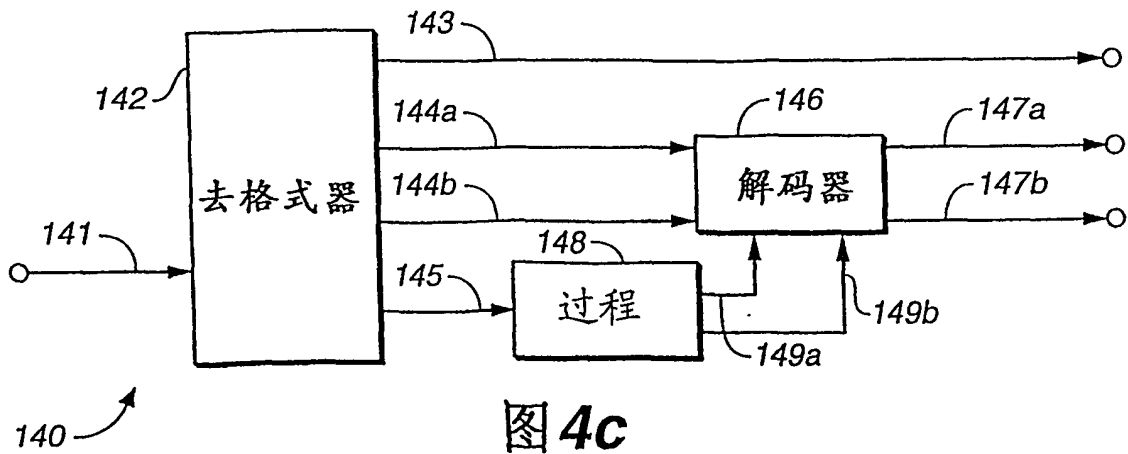
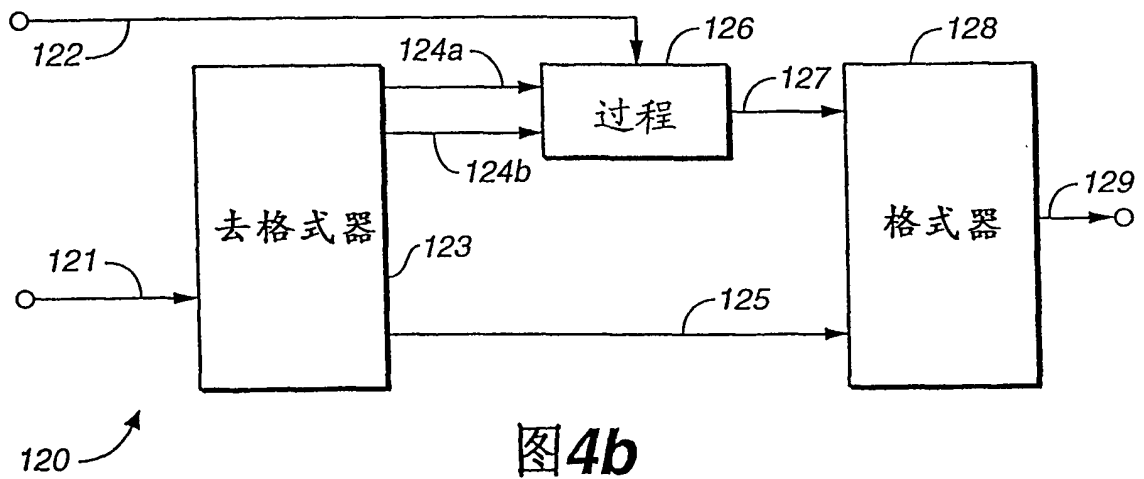
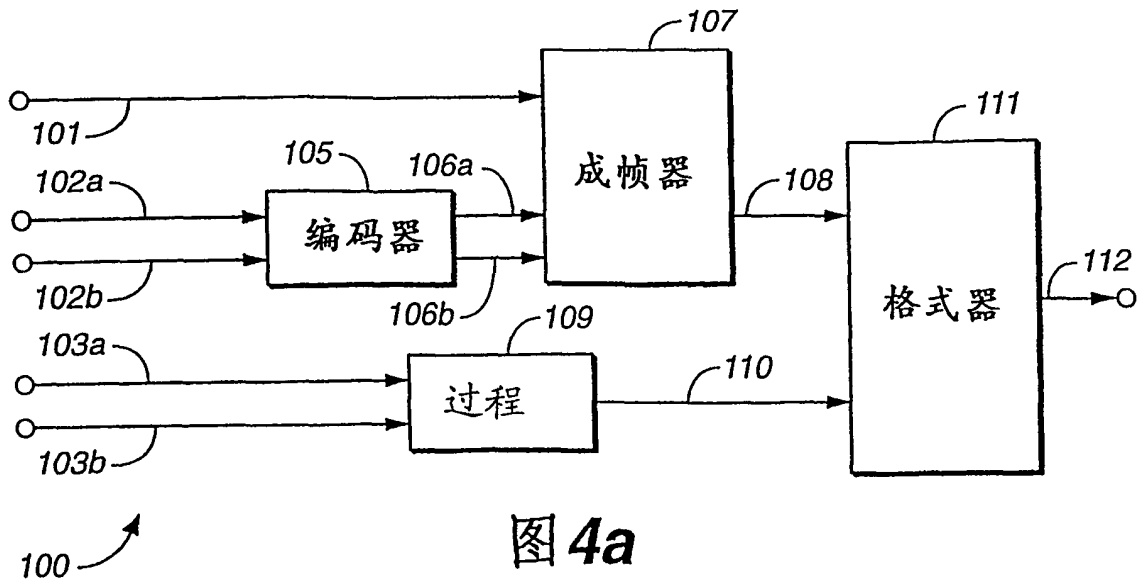


图6d



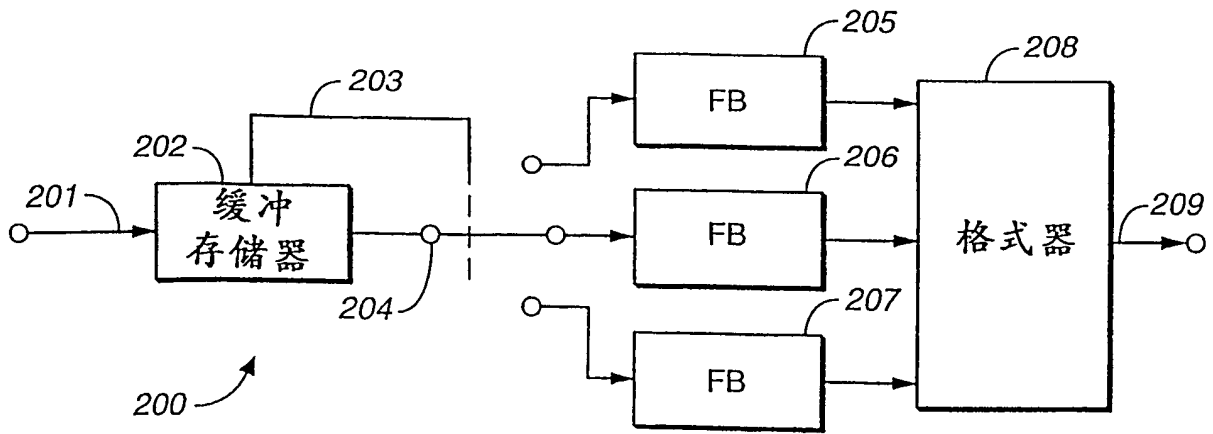


图5a

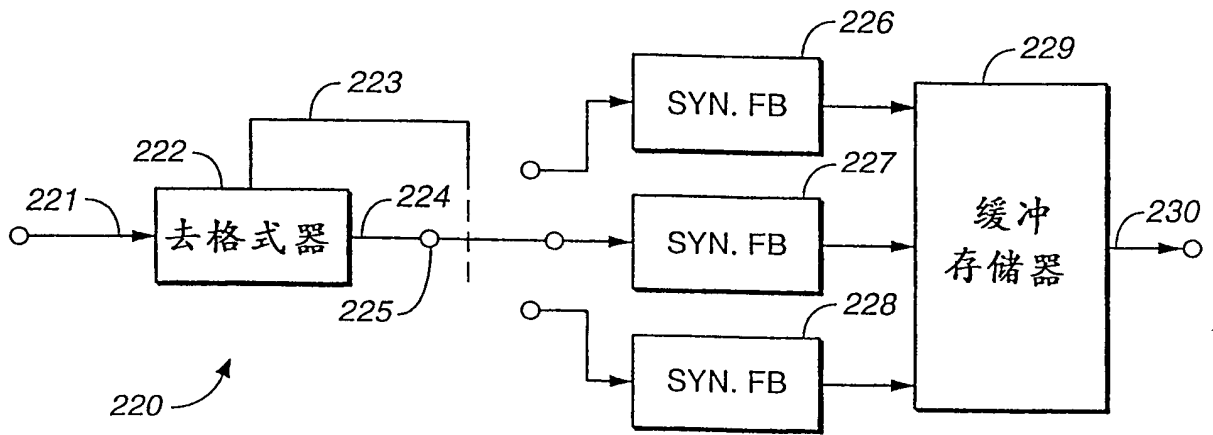


图5b

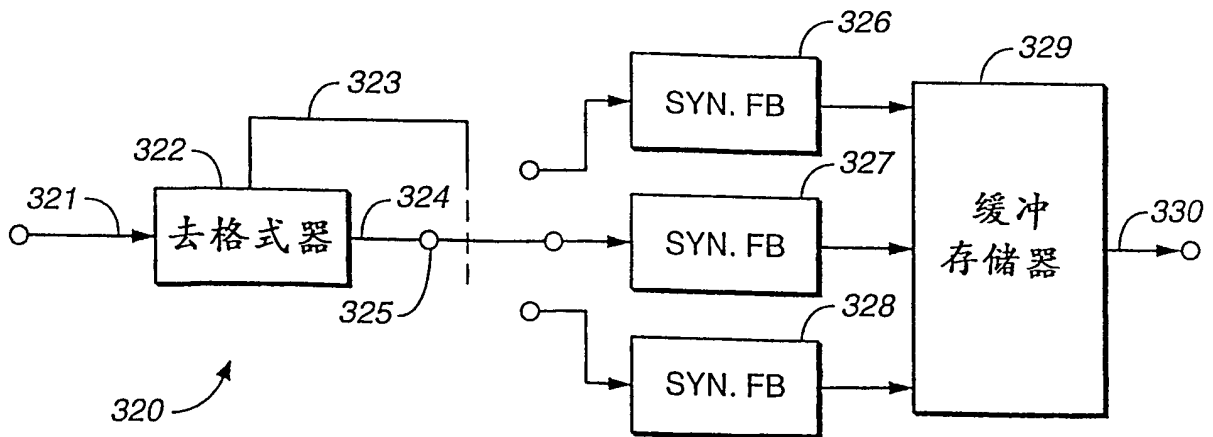


图8

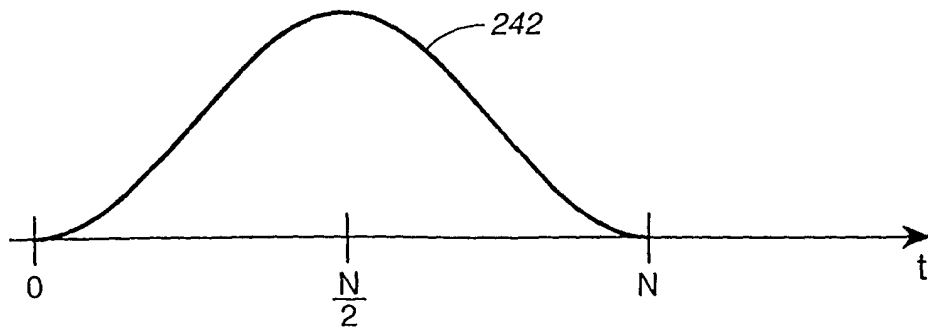


图6a

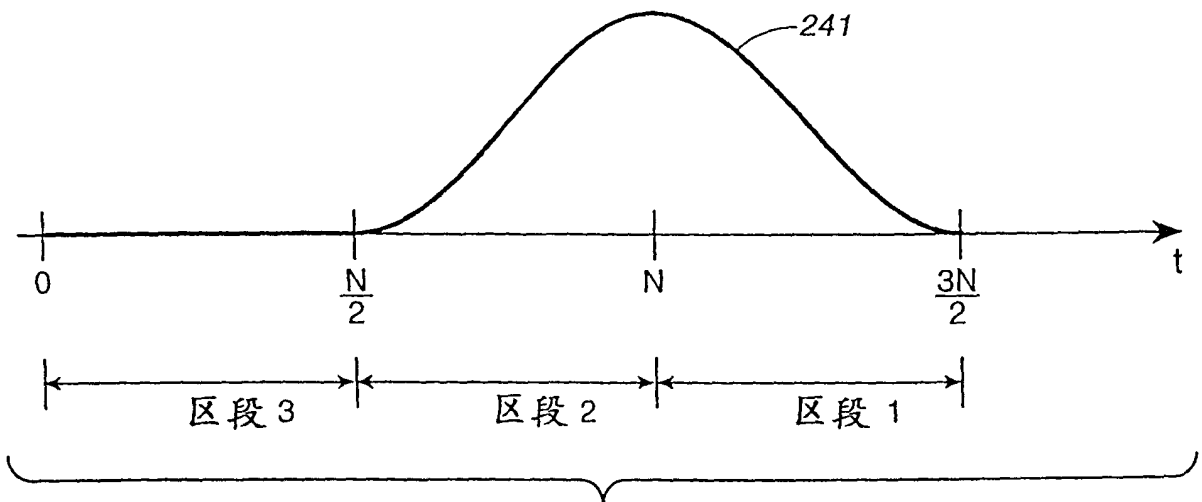


图6b

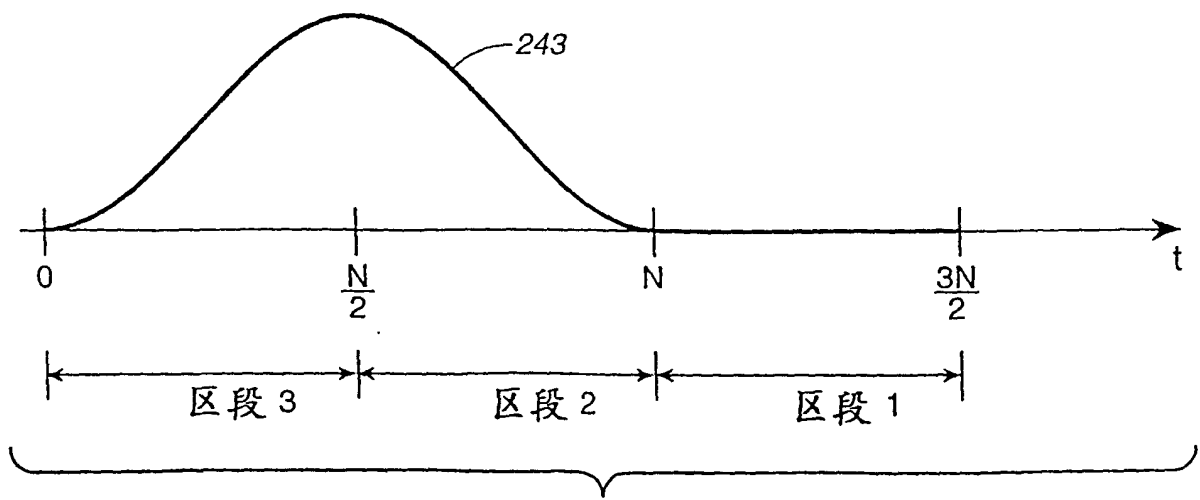
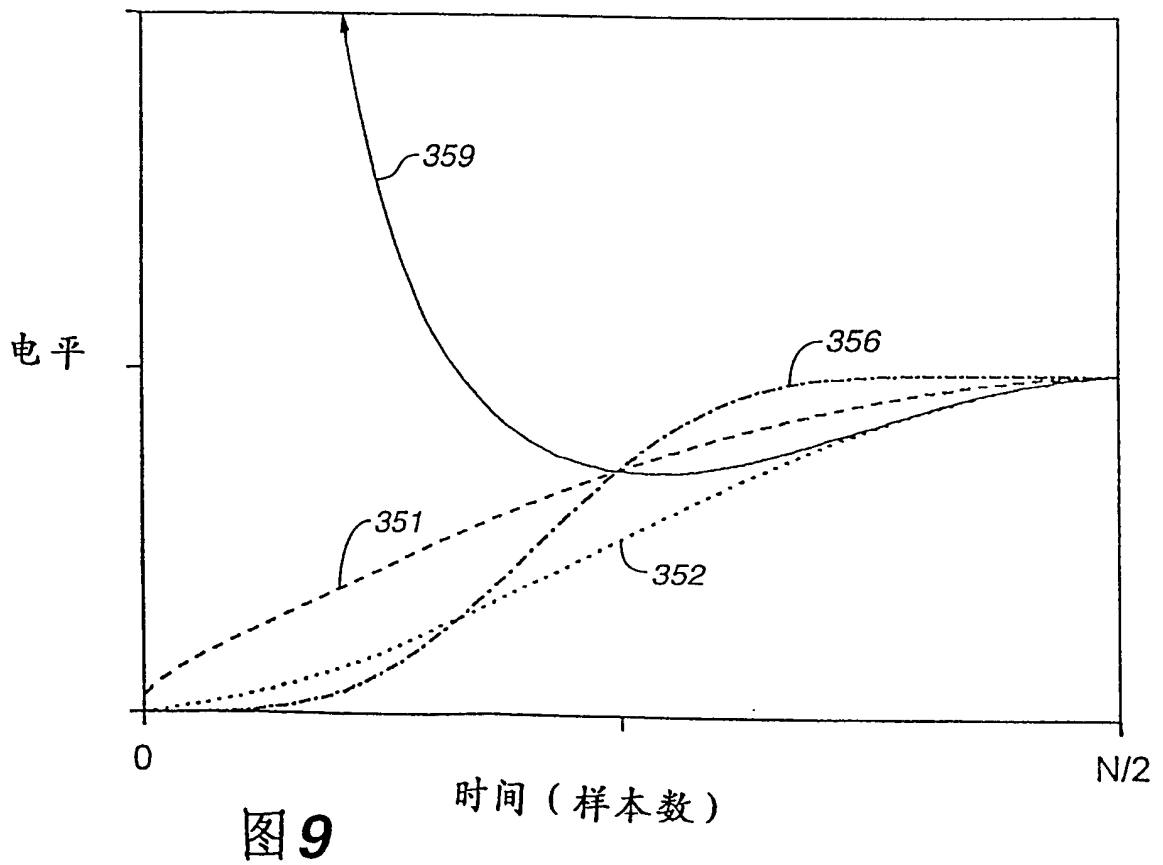
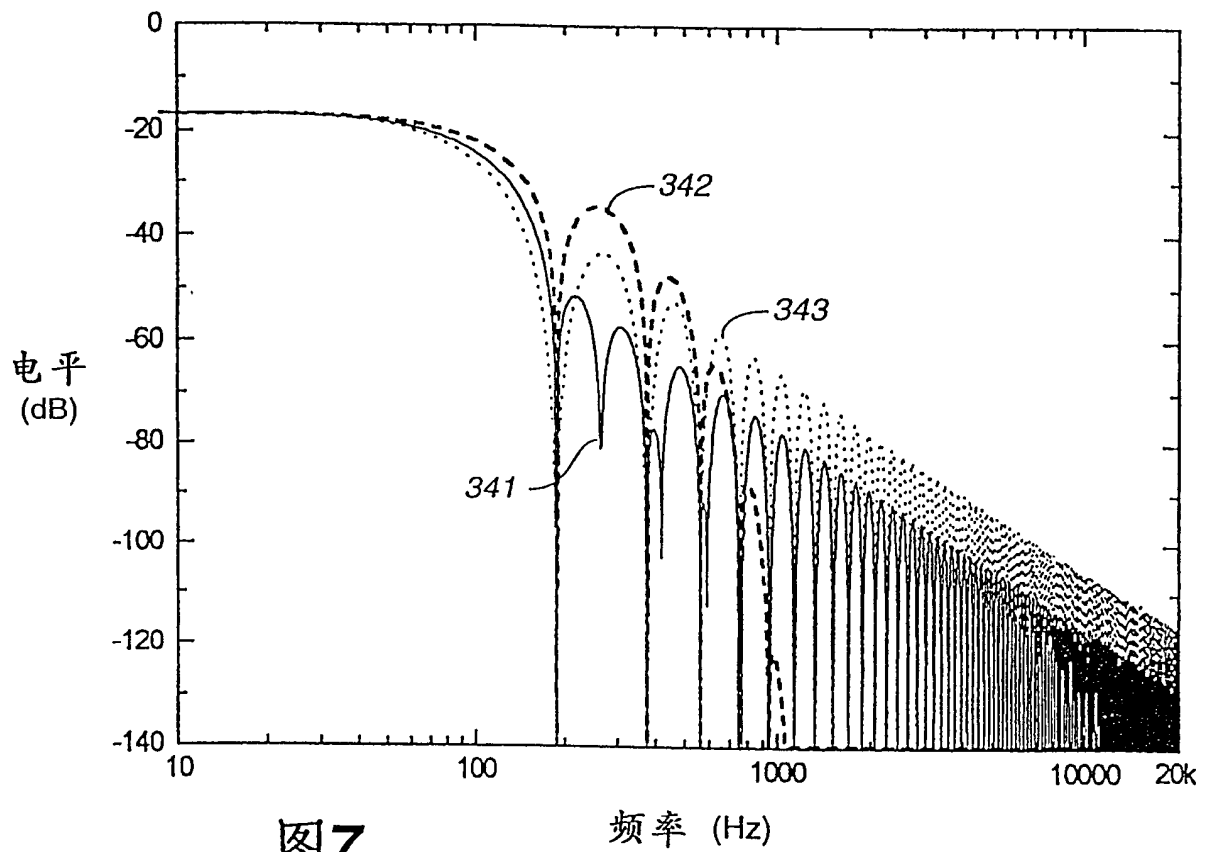


图6c



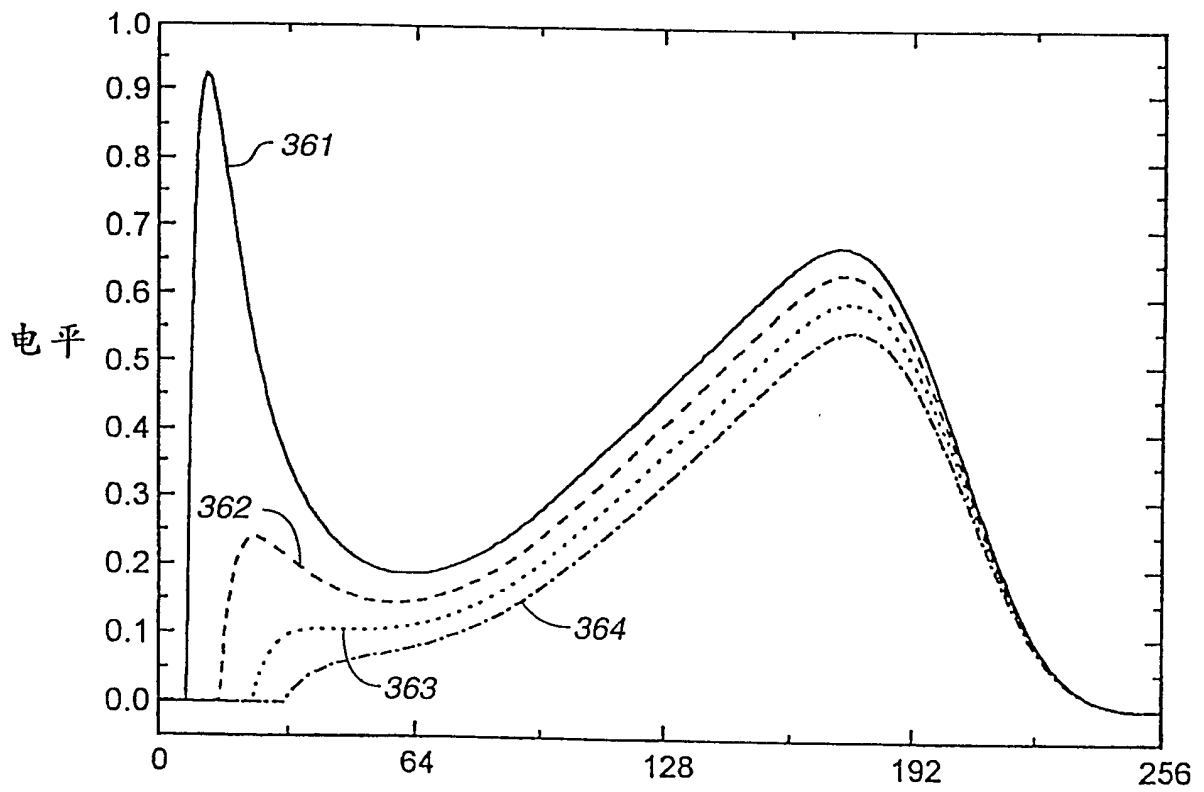


图10a 时间 (样本数)

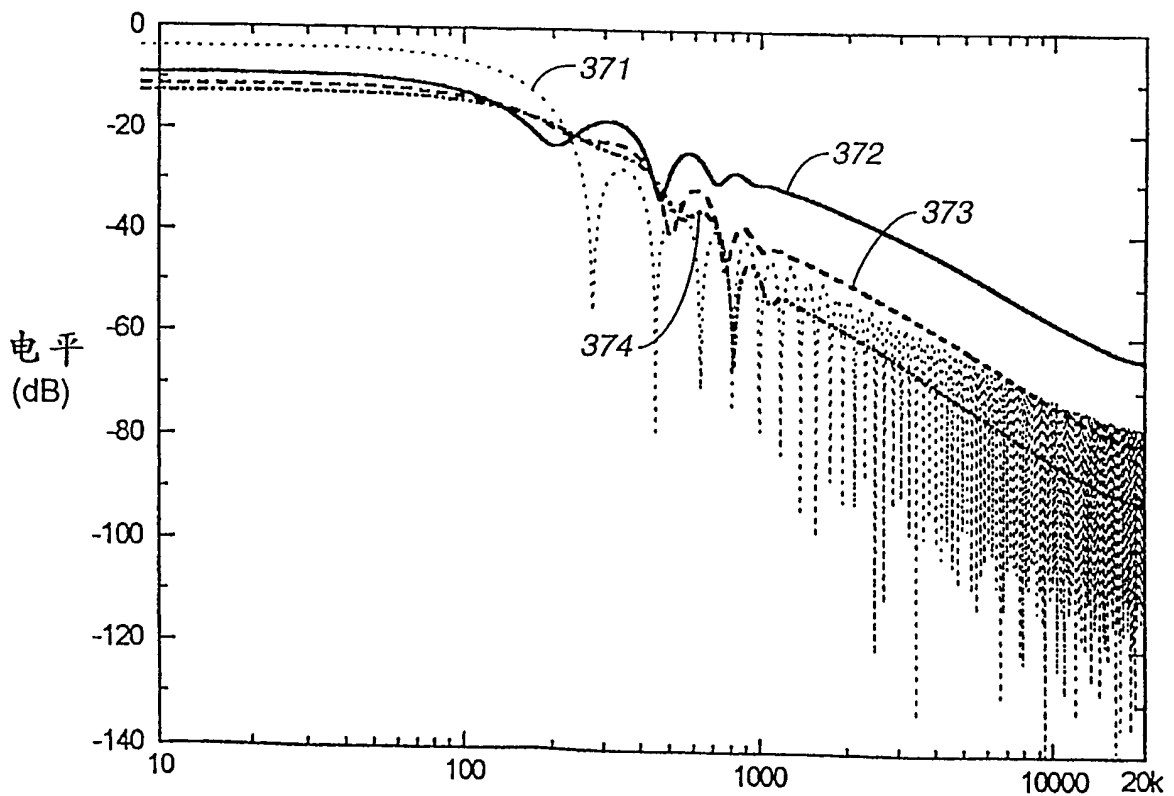


图10b 频率 (Hz)

