

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

H04N 7/26

H04B 1/66 H03M 7/30

G10L 19/00



[12] 发明专利申请公开说明书

[21] 申请号 03154387.1

[43] 公开日 2004 年 4 月 7 日

[11] 公开号 CN 1487746A

[22] 申请日 2003.8.21 [21] 申请号 03154387.1

[30] 优先权

[32] 2002. 8. 28 [33] EP [31] 02090308.4

[71] 申请人 汤姆森许可贸易公司

地址 法国布洛里

[72] 发明人 瓦尔特·维尔森

[74] 专利代理机构 中科专利商标代理有限责任公司

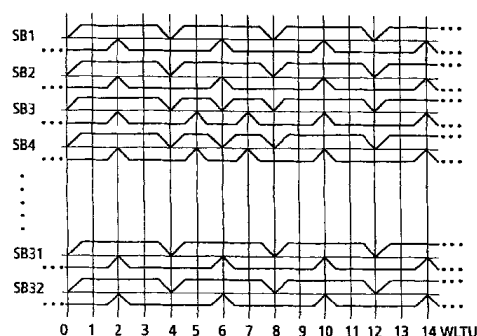
代理人 陈瑞丰

权利要求书 4 页 说明书 7 页 附图 2 页

[54] 发明名称 编码或解码音频信号的方法和设备

[57] 摘要

在 MPEG - 1 音频第三层中, 在参数 “window switching flag” 出现之后, 尽管解码器根据参数 “window_switching_flag” 的出现已经得知序列 “起始块”、“短窗”, “结束块” 和 “长窗” 会随之而来, 仍然会重复地发送对应的 2 比特数 “block_type”。所以, 多次传送可变参数 “block_type” 产生了冗余信息。根据本发明, 不再将多余的参数 “block_type” 标志出于标记块类型的目的而发送。而是利用两个对应的比特将不同的子带信号窗切换结构类型发给解码器。这些结构类型确定了在使用了窗切换的子带中, 哪个子带会被参数 “window_switching_flag” 影响, 而哪个不会。这些结构类型进一步确定了在整个子带中, 会被参数 “window_switching_flag” 影响的不同的固定子带组或可变子带组。



ISSN 1008-4274

1. 一种用于编码音频信号的方法，该音频信号经过了多个子带
5 (SB1... SB32)和交迭窗函数的处理(SAFW)，所述子带中的信号被所述交叠窗函数分割，

其中，在任何情况下，将数据块采样的结果变换(SAFW)为对应的
频谱域系数数据块并采用数据压缩对系数进行编码(ScFCod, NQCod)，

以及，使用不同的窗型，并将已使用窗型的信息(block_type)发送、
10 记录或存储于已编码系数的辅助信息中，

以及，在给定的时间周期里，当决定处理第一组子带(SB1, SB2)
的子带信号时，将有关这种窗型混合的附加信息(mixed_block_flag)发
送、记录或存储于辅助信息中，所述子带信号至少部分地带有给定的窗
型序列，这种窗型序列与对应的用于处理第二组子带(SB3... SB32)的
15 子带信号的窗型序列不同，

其特征在于以下步骤：

在决定处理第一组子带(SB1, SB2)的子带信号之后，不将给定序
列中使用的窗型信息(block_type)发送、记录或存储于辅助信息中，所
述子带信号至少部分地带有所述给定的窗型序列，这种窗型序列与对应
20 的用于处理第二组子带的子带信号的窗型序列不同；

而是将另外的子带信号窗切换结构类型的信息发送、记录或存储于
所述辅助信息中。

2. 根据权利要求1所述的方法，其特征在于所述音频信号是 MPEG-
1 音频第三层、MPEG-2 音频第三层或 AAC 音频信号。

25 3. 根据权利要求1或2所述的方法，其特征在于所述结构类型确定
了在整个使用了窗切换的子带中，哪一个受到有关窗型混合的附加信息
(mixed_block_flag)的影响，而哪一个没有受到影响。

4. 根据权利要求3所述的方法，其特征在于所述结构类型进一步确
定了所述附加信息(mixed_block_flag)会影响到的不同的子带组，这些
30 不同的子带组在整个子带中是固定的，

和/或所述结构类型进一步确定了附加信息 (mixed_block_flag) 在整个子带中会影响到的可变的子带组。

- 5 5. 一种解码音频信号的方法, 该音频信号经过了多个子带 (SB1... SB32) 和交迭窗函数的处理 (SAFW), 该子带中的信号所述交叠窗函数被分割,

在任何情况下, 将数据块采样的结果变换 (SAFW) 为对应的频谱域系数数据块并采用数据压缩对系数进行编码 (ScFCod, NQCod),

以及, 使用不同的窗型, 并将已使用窗型的信息 (block_type) 发送、记录或存储于已编码系数的辅助信息中,

- 10 以及, 在给定的时间周期里, 当决定处理第一组子带 (SB1, SB2) 的子带信号时, 将有关这种窗型混合的附加信息 (mixed_block_flag) 发送、记录或存储于辅助信息中, 所述子带信号至少部分地带有给定的窗型序列, 这种窗型序列与对应的用于处理第二组子带 (SB3... SB32) 的子带信号的窗型序列不同,

- 15 解码包括以下步骤:

将被接收、重放或读取信号的辅助信息解码 (SIDec),

利用所述已解码的辅助信息, 执行被接收、重放或读取码的数据压缩解码 (INQDec), 并在任何情况下, 将频谱域系数数据块反变换 (SSFW) 为对应的采样数据块,

- 20 利用所述交迭窗函数, 组合 (SSFW) 所述反变换得到的采样数据块, 并将所述的多个子带信号 (SB1... SB32) 组合 (SSFW) 为已解码的音频信号 (PCM),

其特征在于以下步骤:

- 25 当进行编码时, 在决定处理第一组子带 (SB1, SB2) 的子带信号之后, 不将所述给定序列中使用的窗型信息 (block_type) 发送、记录或存储于辅助信息中, 而是将另外的子带信号窗切换结构类型的信息发送、记录或存储于所述辅助信息中, 所述子带信号至少部分地带有给定的窗型序列, 这种窗型序列与对应的用于处理第二组子带的子带信号的窗型序列不同,

- 30 求出 (SIDec) 解码过程中进一步的子带信号窗切换类型信息, 并当

利用所述交迭窗函数组合 (SSFW) 所述反变换得到的采样数据块时, 以及将所述多个子带信号 (SB1... SB32) 组合 (SSFW) 为已解码的音频信号 (PCM) 时, 选择对应的窗型。

6. 根据权利要求 5 所述的方法, 其特征在于音频信号是 MPEG-1 音频第三层、MPEG-2 音频第三层或 AAC 音频信号。

7. 根据权利要求 5 或 6 所述的方法, 其特征在于所述结构类型确定了在整个使用窗切换的子带中, 哪一个受到了有关窗型混合的所述附加信息 (mixed_block_flag) 的影响, 而哪一个没有受到影响。

8. 根据权利要求 7 所述的方法, 其特征在于所述结构类型进一步确定了附加信息 (mixed_block_flag) 会影响到的不同的子带组, 这些不同的子带组在整个子带中是固定的,

和/或所述结构类型进一步确定了所述附加信息 (mixed_block_flag) 在整个子带中会影响到的可变的子带组。

9. 一种用于编码音频信号的设备, 包括:

15 装置 (SAFW), 用于利用多个子带 (SB1... SB32) 和交迭窗函数处理所述音频信号, 所述子带中的信号被所述交叠窗函数分割, 并用于在任何情况下, 将采样得到的数据块变换为对应的频谱域系数数据块;

装置 (ScFCal, ScFCod, NQCod, FFTA, ThCalSD, BALLCal, BRAdj), 用于利用数据压缩对所述系数进行编码, 其中, 使用了不同的窗型, 并将所使用的窗型的信息 (block_type) 附加在已编码信号的辅助信息里,

25 以及, 在给定的时间周期里, 当决定处理第一组子带 (SB1, SB2) 的子带信号时, 将有关这种窗型混合的附加信息 (mixed_block_flag) 添加到所述辅助信息中, 所述子带信号至少部分地带有给定窗型序列, 这种窗型序列与对应的用于处理第二组子带 (SB3... SB32) 的子带信号的窗型序列不同,

以及, 在决定处理第一组子带 (SB1, SB2) 的子带信号之后, 不将所述给定序列使用的窗型的信息 (block_type) 添加到所述辅助信息中, 而是将另外的子带信号窗切换结构类型的信息添加到所述辅助信息中, 所述子带信号至少部分地带有给定的窗型序列, 这种窗型序列与对应的用于处理第二组子带 (SB3... SB32) 的子带信号的窗型序列不同。

10. 一种用于解码音频信号的设备，所述音频信号经过了多个子带（SB1... SB32）和交迭窗函数的处理（SAFW），所述子带中的信号被所述交叠窗函数分割，

其中，在任何情况下，将数据块采样的结果变换为（SAFW）对应的
5 频谱域系数块并采用数据压缩对系数进行编码（ScFCod, NQCod），

以及，使用不同的窗型，并将已使用的窗型的信息（block_type）发送、记录或存储于已编码系数的辅助信息中，

以及，在编码过程中，在给定的时间周期里，当决定处理第一组子带（SB1，SB2）的子带信号时，将有关这种窗型混合的附加信息
10 （mixed_block_flag）添加到所述辅助信息中，所述子带信号至少部分地带有给定的窗型序列，这种窗型序列与对应的用于处理第二组子带（SB3... SB32）的子带信号的窗型序列不同，所述的设备包括：

用于解码被接收、重放或读取信号的辅助信息的装置（SIDec），，

装置（INQDec, SSFW），用于利用所述已解码的辅助信息，执行被
15 接收、重放或读取码的数据压缩解码，并在任何情况下，将频谱域系数数据块反变换为对应的采样数据块，利用交迭窗函数组合所述反变换得到的采样数据块，并将所述多个子带信号（SB1... SB32）组合为已解码的音频信号（PCM），

其中，在所述编码过程中，在决定处理第一组子带（SB1，SB2）的
20 子带信号后，不将所述给定序列中使用的窗型的信息发送、记录或存储于所述辅助信息中，而是将另外的子带信号窗切换结构类型的信息发送、记录或存储于所述辅助信息中，所述子带信号至少部分地带有给定的窗型序列，这种窗型序列与对应的用于处理第二组子带（SB3... SB32）的子带信号的窗型序列不同，

25 以及，所述用于解码辅助信息的装置（SIDec）估计出子带信号窗切换的类型信息，当利用所述交迭窗函数组合所述反变换得到采样数据块时，以及在所述用于压缩解码、反变换及组合的装置（INQDec, SSFW）中，将所述多个子带信号组合为已解码的音频信号（PCM）时，该类型信息用于选择对应的窗型。

编码或解码音频信号的方法和设备

5

技术领域

本发明涉及到一种用于编码或解码音频信号的方法和设备，其中所述音频信号经过了多个子带、交迭窗函数以及扩展的子带信号窗切换结构的处理。

10

背景技术

在音频编码中，短时余弦变换或短时傅立叶变换用于从时域输入采样产生谱系数。对这些系数进行编码，从而去除了信号冗余和不相关的部分。在接收机端，将已编码的系数解码，并反变换为时域采样。优选地，在输入信号幅度突然增加之前，可以由长到短地切换变换块的长度，利用或多或少的静音周期来掩蔽预回声，并减少数据块中出现的听得见的噪声，反之亦然，这取决于当前输入信号的性质。

例如，如 EP-A-0 998 051 所述的 ATRAC 压缩算法采用了子带内变换块长度独立切换的技术。

20 变换块长度切换还用于 ISO/IEC11172-3 (MPEG-1 音频第三层)、ISO/IEC13818-3 (MPEG-2 音频第三层) 以及 AAC (高级音频编码)。

发明内容

例如在 MPEG-1 音频第三层和 MPEG-2 音频第三层，利用所谓的
25 “window_switching_flag” 标志，将数据流中帧报头里的变换块长度切换信息或切换窗长度信息（在 “main_data_begin” 和 “main_data” 之间）发送给每一个系数的集合 “granule”。MPEG-1 音频和 MPEG-2 音频中的不同层以及其它的音频编/解码器，如 Minidisc 系统，均使用了子带编码/解码技术，这里将整个的频率带分解为预定数量的子带，如 32 个子
30 带，或者在 Minidisc 里为 3 个子带。图 2 所示为几个子带 SB1 ... SB32，

每一个子带都加了窗。窗的长度以及由此变换到频谱域的长度由“窗长度时间单位”WLTU给出。例如，当原始PCM的采样率为32kHz、44.1kHz或48kHz时，实际窗/变换块可能包含的采样数在12到2048之间。

如图2所示，例如各个窗相互之间的交迭率为50%。变换的类型可以是MDCT（改进的离散余弦变换），这种类型采用因子为2的二次采样，这样就不会增加输入系数的总量。

图2所示的窗函数只是示意性的，例如，实际窗函数的形状可以是正弦/余弦、Kaiser-Bessel或Fielder形状。

在MPEG-1音频第三层、MPEG-2音频第三层和Minidisc的编/解码过程中，在不同的子带，有可能在输入信号的给定时期里选择不同的变换块或窗口长度。这种情况下，需要数据流中包含有关哪一个子带或哪一组子带使用何种变换或窗长度的信息，用于解码器的估值。

例如在MPEG-1音频第三层、MPEG-2音频第三层中，这个参数被称为“mix_block_flag”，它决定了在两个最低的子带SB1和SB2中只使用长块，然而在同样方式下，上面的30个子带中使用的块长度不等，都介于长块和短块之间，这些块中包括分别被称之为起始块和结束块的过渡块。还可以利用2比特参数“block_type”通知数据块或窗口的类型。如果使用了短块，则在任何情况下，以图2中所示的子带3和子带4的情况为例，依次出现的块类型序列为：长块（码0）、起始块（码1，该窗函数波形的两个半区不对称）、3个短块（码2；一般来说至少有一个短块）、结束块（码3，该窗函数波形的两个半区不对称）、长块（码0）。

本发明要解决的问题是要对整个子带范围内容许的块、窗口长度或窗类型的适应性进行改进。通过本发明的权利要求1和权利要求5提供的方法，本发明解决了这一问题。在本发明的权利要求9和权利要求10提出了采用这些方法的设备。

例如在MPEG-1音频第三层和MPEG-2音频第三层中，在参数“window_switching_flag”出现之后，尽管解码器根据参数“window_switching_flag”的出现已经得知上述的序列“起始块”、“短窗”，“结束块”和“长窗”会随之而来，仍然会重复地发送对应的2比特数“block_type”。所以，多次发送可变参数‘block_type’产生了冗余信息。

根据本发明，不再将多余的参数‘block_type’标志为了块类型通知的目的而发送。而是利用两个对应的比特将不同的子带信号窗切换结构类型发给解码器。

5 这些结构类型确定了在整个使用了窗切换的子带中，哪个子带会被上述的参数“window_switching_flag”影响，或是在整个使用了窗切换的子带中，哪个子带不会被上述的参数“window_switching_flag”影响。

这些结构类型进一步确定了受参数“window_switching_flag”影响的不同子带组，这些子带组在整个子带中是固定的。

10 这些结构类型进一步确定了受参数“window_switching_flag”影响的整个子带中的可变子带组。

也可以将这两种选择相结合。

原则上，本发明的方法适用于音频信号的编码，该音频信号经过了多个子带和交迭窗函数的处理，而子带中的信号被该交叠窗函数分割，

15 其中在任何情况下，将数据块采样的结果变换为对应的频谱域系数数据块，并采用数据压缩对系数进行编码，

以及，使用不同的窗型，并将已使用的窗型信息发送、记录或存储于已编码系数的辅助信息中，

20 以及，在给定的时间周期里，当决定处理第一组子带的子带信号时，将有关这种窗型混和的附加信息发送、记录或存储于辅助信息中，所述子带信号至少部分地带有给定的窗型序列，这种窗型序列与对应的用于处理第二组子带的子带信号的窗型序列不同，

25 以及，在决定处理第一组子带的子带信号之后，不要将所述给定序列中使用的窗型信息发送、记录或存储于辅助信息中，所述子带信号至少部分地带有给定的窗型序列，这种窗型序列与对应的用于处理第二组子带的子带信号的窗型序列不同；

而是将另外的子带信号窗切换结构类型的信息发送、记录或存储于所述辅助信息中。

30 原则上，本发明的方法适用于音频信号的解码，该音频信号经过了多个子带和交迭窗函数处理，所述的子带中的信号被所述交叠窗函数分割，

其中在任何情况下，将数据块采样的结果变换为对应的频谱域系数数据块并使用数据压缩对系数进行编码，

以及，使用不同的窗型，并将已用窗型的信息发送、记录或存储于已编码系数的辅助信息中，

- 5 以及，在给定的时间周期里，当决定处理第一组子带的子带信号时，将有关这种窗型混合的附加信息发送、记录或存储于辅助信息中，所述子带信号至少部分地带有给定的窗型序列，这种窗型序列与对应的用于处理第二组子带的子带信号的窗型序列不同，解码包括步骤：

 将被接收、重放或读取信号的辅助信息解码，

- 10 利用所述已解码的辅助信息，执行被接收、重放或读取码的数据压缩解码，并在任何情况下，将频谱域系数数据块反变换为对应的采样数据块，

 利用所述的交迭窗函数，组合所述反变换得到的采样数据块，并将多个子带信号组合为已解码的音频信号，

- 15 其中，在进行所述编码时，在决定处理第一组子带的子带信号之后，不将所述给定序列中使用的窗型信息发送、记录或存储于辅助信息中，而是将另外的子带信号窗切换结构类型的信息发送、记录或存储于辅助信息中，所述子带信号至少部分地带有给定的窗型序列，这种窗型序列与对应的用于处理第二组子带的子带信号的窗型序列不同，

- 20 求出解码过程中进一步的子带信号窗切换类型信息，并且当利用交迭窗函数组合所述反变换得到的采样数据块时，以及将多个子带信号组合为已解码的音频信号时，选择对应的窗型。

 原则上，本发明用于编码音频信号的设备包括：

- 25 利用多个子带和交迭窗函数处理所述音频信号的装置，该子带中的信号被所述交叠窗函数分割，并在任何情况下，该装置将采样得到的数据块变换为对应的频谱域系数数据块；

 利用数据压缩对所述系数进行编码的装置，其中使用了不同的窗型并将已使用窗型的信息附加在已编码信号的辅助信息里，

- 30 以及，在给定的时间周期里，当决定处理第一组子带的子带信号时，将有关这种窗型混合的附加信息添加到辅助信息中，所述子带信号至少

部分地带有给定的窗型序列，这种窗型序列与对应的用于处理第二组子带的子带信号的窗型序列不同，

以及，在决定处理第一组子带的子带信号之后，不将所述给定序列使用的窗型信息添加到辅助信息中，而是将另外的子带信号窗切换结构类型的信息添加到辅助信息中，这种子带信号至少部分地带有给定的窗型序列，这种窗型序列与对应的用于处理第二组子带的子带信号的窗型序列不同。

原则上，本发明的设备用于音频信号的解码，这种音频信号经过了多个子带和交迭窗函数的处理，所述子带中的信号被所述交叠窗函数分割，

其中，在任何情况下，将数据块采样的结果变换为对应的频谱域系数块并采用数据压缩对系数进行编码，

以及，使用不同的窗型，并将已使用窗型的信息发送、记录或存储于辅助信息中，

以及，在编码过程中，在给定的时间周期里，当决定处理第一组子带的子带信号时，将有关这种窗型混合的附加信息添加到辅助信息中，这种子带信号至少部分地带有给定的窗型序列，这种窗型序列与对应的用于处理第二组子带的子带信号的窗型序列不同，所述的设备包括：

用于解码被接收、重放或读取信号的辅助信息的装置，

装置，用于利用所述已解码的辅助信息执行被接收、重放或读取码的数据压缩解码，并在任何情况下，将频谱域系数数据块反变换为对应的采样数据块，并利用交迭窗函数组合反变换得到的采样数据块，并将所述多个子带信号组合为已解码的音频信号，

其中，在所述编码过程中，在决定处理第一组子带的子带信号后，不将所述给定序列中使用的窗型信息发送、记录或存储于辅助信息中，而是将子带信号窗切换结构类型的信息发送、记录或存储于辅助信息中，所述子带信号至少部分地带有给定的窗型序列，这种窗型序列与对应的用于处理第二组子带的子带信号的窗型序列不同，

以及，所述用于解码辅助信息的装置求出子带信号窗切换类型的信息，当利用交迭窗函数组合所述反变换得到采样数据块时，以及在所述

述用于压缩解码、反变换及组合的装置中将所述多个子带信号组合为已解码的音频信号时，该类型信息用于选择对应的窗型。

各本发明的各个从属权利要求中公开了本发明其它具体的优势。

5 附图说明

本发明实施例的描述参考了这里给出的附图：

图 1 是实现本发明的编码器框图；

图 2 示出了窗在频率子带中的位置；

图 3 是实现本发明的解码器框图。

10

具体实施方式

在图 1 中，由模块 EINP 接收编码器输入信号。SAFW 级实现子带分析滤波（即产生 32 个子带信号）、加窗以及变换到频谱域。利用接收到的来自 BRAdj 级的辅助信息，ScFCal 级计算组成谱系数的尺度因子。同样利用来自 BRAdj 级的辅助信息，由 NQCod 级完成子带中系数的归一化、量化和编码。被发送、记录或存储的音频帧的格式化由 FrFo 级来执行。

FFTA 级并行地实现 EINP 输入信号的 FFT（快速傅立叶变换）分析，以便为心理声学信息提供了信源。根据心理声学信息，接下来的 ThCalSD 级计算掩蔽阈值、信号/掩蔽比并确定各子带所需的窗切换信息。在 SAFW 级，将这个窗切换信息施加给子带信号和对应的变换操作。BallCal 级计算所需的比特分配。随后 BRAdj 级通过向 ScFCod 级和 NQCod 级发送对应的控制信号来控制对期望固定比特率的调整。

图中只给出了两个（立体声）或多个信道中的一个信道的情形，一般地，在所有信道中，FFTA 级、ThCalSD 级、BallCal 级和 BRAdj 级是共有的。

图 3 中，由模块 DINP 接收、重放或读取解码器的输入信号，SIDec 级对解码器需要的辅助信息进行解码，辅助信息由编码器产生，包括如尺度因子信息、比特分配信息、窗切换信息、归一化信息、量化信息以及阈值信息。SIDec 级控制着后面的 INQDec 级和 SSFW 级。INQDec 级

执行来自子带的被接收或重放系数的反编码、反量化和反归一化。SSF_W级执行反变换、对应的窗切换以及子带合成滤波，并完成PCM采样的输出。

图中只给出了两个（立体声）或多个信道中的一个信道的情形。

- 5 如图2所示的子带1/2、3/4和31/32，本发明的窗切换分别被加到编码器的SAF_W级和解码器的SSF_W级，这种窗切换采用了不同的子带信号窗切换结构类型。结构类型信息的选择由ThCalSD级确定，并由解码器的SIDec级进行传送及估值。

- 10 例如，本发明可用于基于MPEG-1音频第三层、MPEG-2音频第三层或AAC的扩展系统。

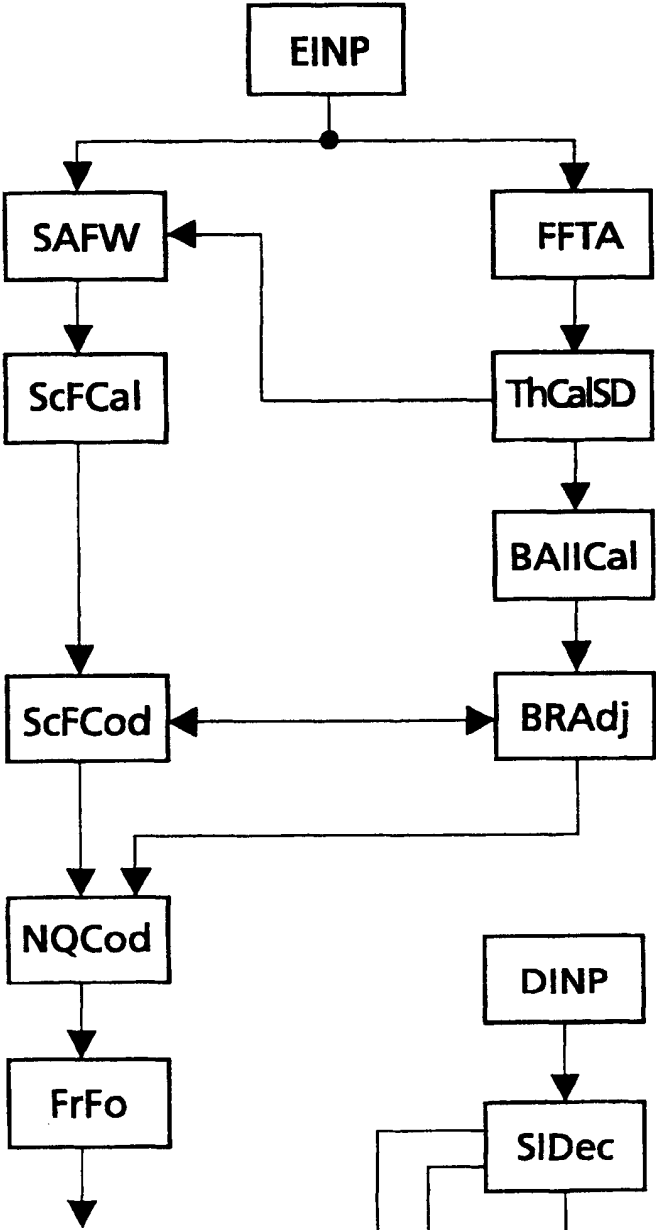


图 1

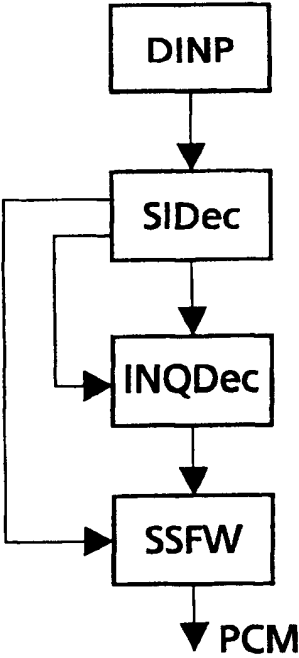


图 3

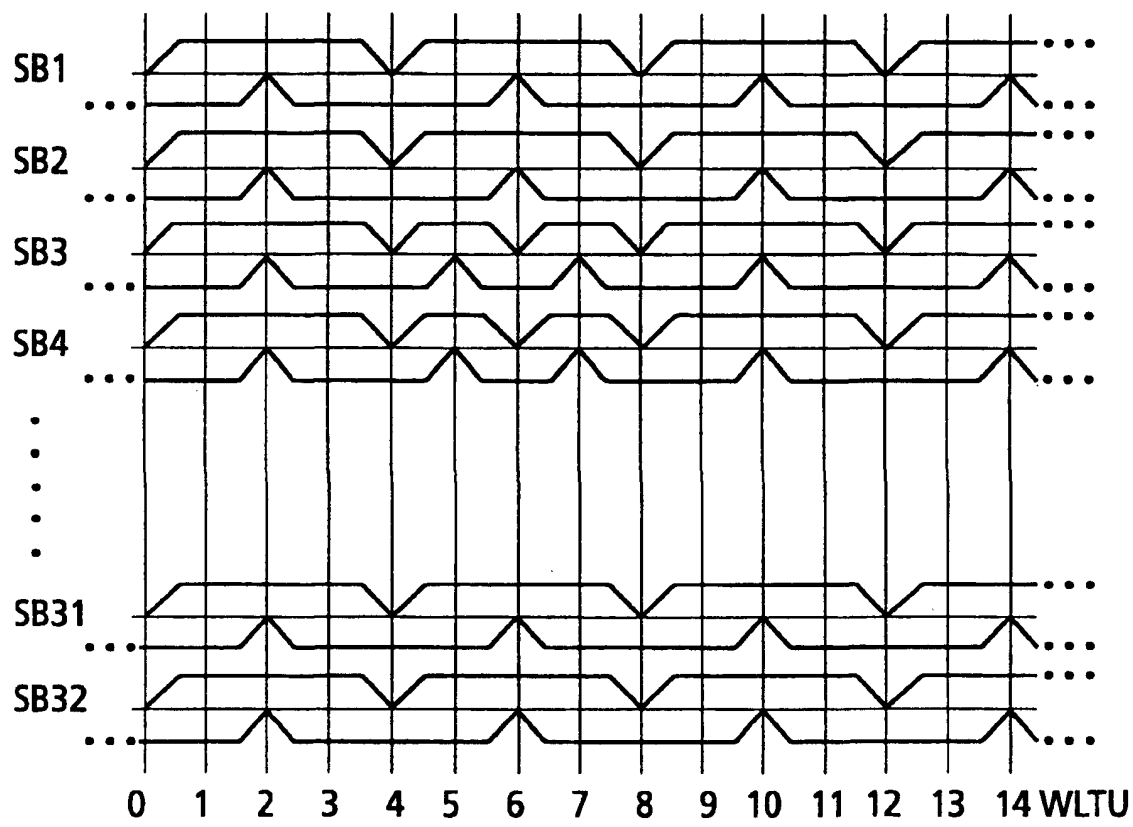


图 2