



(12) 发明专利

(10) 授权公告号 CN 102947880 B

(45) 授权公告日 2014. 10. 08

(21) 申请号 201180028342. 9

(51) Int. Cl.

(22) 申请日 2011. 04. 06

G10L 19/008 (2013. 01)

(30) 优先权数据

61/322, 458 2010. 04. 09 US

(56) 对比文件

CN 101067931 A, 2007. 11. 07,

CN 101253557 A, 2008. 08. 27,

CN 1677490 A, 2005. 10. 05,

US 2009/0210234 A1, 2009. 08. 20,

(85) PCT国际申请进入国家阶段日

2012. 12. 07

(86) PCT国际申请的申请数据

PCT/EP2011/055369 2011. 04. 06

审查员 张飞弦

(87) PCT国际申请的公布数据

W02011/124616 EN 2011. 10. 13

(73) 专利权人 杜比国际公司

地址 荷兰祖伊杜斯特

(72) 发明人 海科·普尔哈根 蓬图斯·卡尔森

拉尔斯·维尔默斯

(74) 专利代理机构 北京集佳知识产权代理有限公司

公司 11227

代理人 王萍 李春晖

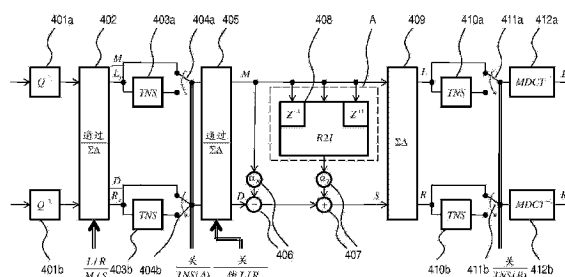
权利要求书4页 说明书21页 附图16页

(54) 发明名称

解码器系统和解码方法

(57) 摘要

本发明提供了用于在频域中使用复合预测来进行立体声编码与解码的方法及装置。在一种实施方式中,一种用于从通过复合预测编码编码的输入立体声信号来获得输出立体声信号并且包括有两个输入声道的第一频域表示的方法包括以下上混合步骤:(i) 计算第一输入声道的第二频域表示;以及(ii) 基于第一输入声道的第一频域表示和第二频域表示、第二输入声道的第一频域表示和复合预测系数来计算输出声道。该方法包括在上混合前后选择性地进行频域修改。



1. 一种用于通过复合预测立体声编码来提供立体声信号的解码器系统,所述解码器系统包括:

上混合级 (406、407、408、409 ;1433),所述上混合级适于基于下混合信号 (M) 和残差信号 (D) 的第一频域表示来生成所述立体声信号,所述第一频域表示中的每一个包括在多维空间的第一子空间中表达的表示对应信号的频谱含量的第一频谱分量,所述上混合级包括:

模块 (408),所述模块基于所述下混合信号的所述第一频域表示来计算所述下混合信号的第二频域表示,所述第二频域表示包括在所述多维空间的第二子空间中表达的表示所述信号的频谱含量的第二频谱分量,所述第二子空间包括在所述第一子空间中不包括的所述多维空间的一部分;

加权加法器 (406、407),所述加权加法器基于所述下混合信号的所述第一频域表示和所述第二频域表示、所述残差信号的所述第一频域表示和在比特流信号中编码的复合预测系数 (α) 来计算侧信号 (S);以及

和与差级 (409),所述和与差级基于所述侧信号和所述下混合信号的所述第一频域表示来计算所述立体声信号;

第一频域修改器级 (403 ;1431),所述第一频域修改器级布置在所述上混合级的上游并且能够在主动模式和被动模式下进行操作,在所述主动模式中,所述第一频域修改器级处理至少一个信号的频域表示,在所述被动模式中,所述第一频域修改器级用作直通;以及

第二频域修改器级 (410 ;1435),所述第二频域修改器级布置在所述上混合级的下游并且能够在主动模式和被动模式下进行操作,在所述主动模式中,所述第二频域修改器级处理至少一个信号的频域表示,在所述被动模式中,所述第二频域修改器级用作直通。

2. 根据权利要求 1 所述的解码器系统,其中,所述频域修改器级中的至少一个是时域噪声整形 TNS 级。

3. 根据权利要求 2 所述的解码器系统,所述解码器系统还适于针对每个时间帧接收与所述帧相关联的数据域,并且适于响应所述数据域的值来在主动模式或直通模式下操作所述第一频域修改器级以及在主动模式或直通模式下操作所述第二频域修改器级。

4. 根据前述权利要求中任一项所述的解码器系统,还包括:

去量化级 (401),所述去量化级布置在所述上混合级的上游,以基于比特流信号来提供所述下混合信号 (M) 和所述残差信号 (D) 的所述第一频域表示。

5. 根据权利要求 1-3 中任一项所述的解码器系统,其中:

所述第一频谱分量具有在所述第一子空间中表达的实数值;

所述第二频谱分量具有在所述第二子空间中表达的虚数值;

选择性地,所述第一频谱分量能够通过以下中的一个来获得:

离散余弦变换 DCT,或

改进的离散余弦变换 MDCT,以及

选择性地,所述第二频谱分量能够通过以下中的一个来获得:

离散正弦变换 DST,或

改进的离散正弦变换 MDST。

6. 根据权利要求 5 所述的解码器系统,其中:

所述下混合信号被分割成连续的时间帧,所述连续的时间帧中的每一个与所述复合预测系数的值相关联;以及

用于计算所述下混合信号的第二频域表示的所述模块适于响应于所述复合预测系数的虚部的绝对值小于用于时间帧的预定容差而使所述模块自身无效,以使得所述模块没有针对该时间帧生成输出。

7. 根据权利要求 1-3 中任一项所述的解码器系统,其中:

所述第一频谱分量是布置在变换系数的一个或更多个时间块中的变换系数,每个时间块是通过时对域信号的时间片段施加变换来生成的;以及

用于计算所述下混合信号的第二频域表示的所述模块适于:

从所述第一频谱分量中的至少一些中导出一个或更多个第一中间分量;

根据一个或更多个脉冲响应中的至少一部分来形成所述一个或更多个第一频谱分量的组合,以获得一个或更多个第二中间分量;以及

从所述一个或更多个第二中间分量来导出所述一个或更多个第二频谱分量。

8. 根据权利要求 7 所述的解码器系统,其中,所述一个或更多个脉冲响应的所述部分是基于所述一个或更多个变换的频率响应特征,其中,选择性地,所述变换的所述频率响应特征取决于一个或更多个分析窗函数的特征,所述一个或更多个分析窗函数与变换一起应用到所述信号的所述时间片段上。

9. 根据权利要求 7 所述的解码器系统,其中,用于计算所述下混合信号的第二频域表示的所述模块适于基于下列中的一个来获得所述第二频谱分量的每个时间帧:

(a) 所述第一频谱分量的同时期的时间帧;

(b) 所述第一频谱分量的同时期的和先前的时间帧;以及

(c) 所述第一频谱分量的同时期的、先前的和随后的时间帧。

10. 根据权利要求 1-3 中任一项所述的解码器系统,其中,用于计算所述下混合信号的第二频域表示的所述模块适于计算近似第二频谱表示,所述近似第二频谱表示包括通过至少两个时间上相邻和/或频率上相邻的所述第一频谱分量的组合确定的近似第二频谱分量。

11. 根据权利要求 1-3 中任一项所述的解码器系统,所述立体声信号被表示在时域中,并且所述解码器系统还包括:

切换组件 (203),所述切换组件布置在去量化级与所述上混合级之间,所述切换组件能操作以用作以下任何一种:

(a) 直通级,或

(b) 和与差级,

从而,能够在直接地编码的立体声输入信号与联合地编码的立体声输入信号之间切换;

逆变换级 (209),所述逆变换级适于计算所述立体声信号的时域表示;以及

选择器装置 (208),所述选择器装置布置在所述逆变换级的上游并且适于选择性地将所述逆变换级连接到以下任何一种:

(a) 在所述上混合级下游的点,从而将通过复合预测获得的立体声信号提供给所述逆变换级;或

(b) 在所述上混合级上游并且在所述切换组件 (203) 下游的点,从而将通过直接立体声编码获得的立体声信号提供给所述逆变换级。

12. 根据权利要求 1-3 中任一项所述的解码器系统,其中,用于计算所述下混合信号的第二频域表示的所述模块包括:

逆变换级 (306),所述逆变换级用于基于各个信号在所述多维空间中的所述第一子空间中的所述第一频域表示,计算所述下混合信号和 / 或所述侧信号的时域表示;以及

变换级 (307),所述变换级用于基于所述信号的所述时域表示来计算所述各个信号的所述第二频域表示,

其中,优选地,所述逆变换级 (306) 执行逆的改进的离散余弦变换 (MDCT),所述变换级执行改进的离散正弦变换 (MDST)。

13. 根据权利要求 12 所述的解码器系统,所述立体声信号被表示在时域中,并且所述解码器系统还包括:

切换组件 (302),所述切换组件布置在去量化级与所述上混合级之间,所述切换组件能够操作以用作以下任何一种:

(a) 直通级,以在联合立体声编码中使用;或

(b) 和与差级,以在直接立体声编码中使用;

另外的逆变换级 (311),所述另外的逆变换级布置在所述上混合级中,用于计算所述侧信号的时域表示;

选择器装置 (305、310),所述选择器装置布置在所述逆变换级 (306、301) 的上游,适于选择性地所述逆变换级 (306、301) 连接到下列中的任何一个:

(a) 另外的和与差级 (304),所述另外的和与差级又连接到在所述切换组件 (302) 的下游并且在所述上混合级的上游的点;或

(b) 从所述切换组件 (302) 获得的下混合信号和从所述加权加法器 (308、309) 获得的侧信号。

14. 一种解码方法,所述解码方法用于通过复合预测立体声编码将输入立体声信号上混合成输出立体声信号,其中:

所述输入立体声信号包括下混合声道 (M) 和残差声道 (D) 的第一频域表示以及复合预测系数 (α);以及

所述第一频域表示中的每一个包括在多维空间的第一子空间中表达的表示对应信号的频谱含量的第一频谱分量,

所述方法由上混合级来执行并且包括以下步骤:

基于所述下混合声道的所述第一频域表示来计算所述下混合信号的第二频域表示,所述第二频域表示包括在所述多维空间的第二子空间中表达的表示所述信号的频谱含量的第二频谱分量,所述第二子空间包括在所述第一子空间中不包括的所述多维空间的一部分;

基于所述下混合信号的所述第一频域表示和所述第二频域表示、所述残差信号的所述第一频域表示和所述复合预测系数来计算侧声道;

以及还包括以下步骤中的任何一个:

在所述上混合步骤前执行的、对所述下混合信号的所述第一频域表示和 / 或所述残差

信号的所述第一频域表示应用时域噪声整形 TNS 的步骤；

或者，在所述上混合步骤后执行的、对所述立体声信号的至少一个声道应用 TNS 的步骤。

解码器系统和解码方法

技术领域

[0001] 本文所公开的发明一般涉及立体声 (stereo) 音频编码,更准确地,涉及用于使用频域中的复合预测的立体声编码的技术。

背景技术

[0002] 立体声信号的左 (L) 声道和右 (R) 声道的联合编码相比独立的 L 和 R 编码能够更加有效地进行编码。用于联合立体声编码的通常的方法是中 / 侧 (M/S) 编码。此处,中 (M) 信号是通过将 L 信号与 R 信号相加而形成的,例如,M 信号可以具有以下形式:

$$[0003] \quad M = (L+R)/2$$

[0004] 此外,侧 (S) 信号是通过将两个声道 L 和 R 的信号相减来形成的,例如,S 信号可以具有以下形式:

$$[0005] \quad S = (L-R)/2$$

[0006] 在 M/S 编码的情形下,对 M 信号与 S 信号进行编码来替代对 L 信号与 R 信号进行编码。

[0007] 在 MPEG (运动图像专家组) ACC (高级音频编码) 标准 (参见标准文献 ISO/IEC 13818-7) 中,能够以时变和频变的方式来选择 L/R 立体声编码和 M/S 立体声编码。因此,立体声编码器可以针对一些频带的立体声信号应用 L/R 编码,而对其他频带的立体声信号使用 M/S 编码 (频变)。此外,编码器能够随着时间在 L/R 编码与 M/S 编码之间切换 (时变)。在 MPEG AAC 中,立体声编码是在频域中执行的,更具体地,是在 MDCT (改进的离散余弦变换) 域中执行的。这使得能够自适应地以频变方式以及时变方式选择 L/R 或 M/S 编码中的任一种。

[0008] 参数立体声编码是一种用于对作为单声道信号的立体音频信号加上少量的关于立体声参数的边信息进行有效编码的技术。它是 MPEG-4 音频标准 (参见标准文献 ISO/IEC 14496-3) 的一部分。能够使用任何音频编码器对单声道信号进行编码。立体声参数能够嵌入在单声道比特流的辅助部分中,因此,实现了全面的前向与后向兼容性。在解码器中,首先对单声道信号进行解码,之后,借助于立体声参数来重构立体声信号。与单声道信号具有零互相关的已解码单声道信号的去相关版本是借助于去相关器来生成的 (例如,可以包括一个或更多个延迟线的适合的全通滤波器)。基本上,去相关信号与单声道信号具有相同的频谱和时间能量分布。将单声道信号与去相关信号一起输入到由立体声参数控制并且重构立体声信号的上混合 (upmix) 处理中。更多的信息,参见论文 "Low Complexity Parametric Stereo Coding in MPEG-4", H. Purnhagen, Proc. of the 7th Int. Conference on Digital Audio Effects (DAFx' 04), Naples, Italy, October 5-8, 2004, pages 163-168。

[0009] 环绕 MPEG (MPEG Surround) (MPS; 参见 ISO/IEC 23003-1 以及论文 "MPEG Surround-The ISO/MPEG Standard for Efficient and Compatible Multi-Channel Audio Coding", J. Herre et al., Audio Engineering Convention Paper 7084, 122nd Convention, May 5-8, 2007) 允许将参数立体声编码与残差编码的原理组合,以及使用所传

输的残差来替换去相关信号,因此提高了感观质量。可以通过对多声道信号进行下混合以及可选地,通过提取空间线索来实现残差编码。在下混合处理期间,对表示误差信号的残差信号进行计算,然后对其进行编码与传输。在解码器中,它们可以取代去相关信号。在混合方法中,它们可以在某些频带中取代去相关信号,优选地,在相对地低的频带。

[0010] 根据现有的 MPEG 统一语音 / 音频编码 (USAC) 系统 (在图 1 中示出了它的两个示例), 解码器包括位于核心解码器的下游的复值正交镜像滤波器 (QMF) 组。作为该滤波器组的输出而获得的 QMF 表示是复合 (因此, 以因子 2 来进行过采样) 并且可以被配置为下混合信号 (或, 等同地, 中信号) M 和残差信号 D, 对其应用具有复合输入的上混合矩阵。如下来获得 L 信号和 R 信号 (在 QMF 域):

$$[0011] \quad \begin{bmatrix} L \\ R \end{bmatrix} = g \begin{bmatrix} 1 - \alpha & 1 \\ 1 + \alpha & -1 \end{bmatrix} \begin{bmatrix} M \\ D \end{bmatrix}$$

[0012] 其中, g 是实值增益因子以及 α 是复值预测系数。优选地, 选择 α 以使得残差信号 D 的能量最小化。该增益因子可由标准化来确定, 即, 确保和信号的功率等于左信号和右信号的功率的和。L 信号和 R 信号中的每一个信号的实部和虚部是互冗余的 (原理上, 它们中的每一个都可以基于另一个来计算), 但是对于能够在不发生音频失真伪影的情况下进行频带复制 (SBR) 解码器的后续应用来说是有益的。还可以以防止与其他时间或频率自适应信号处理 (未示出) (如单声道 - 立体声上混合) 有联系的伪影为目的, 针对类似的理由选择过采样信号表示的用途。逆 QMF 滤波是解码器中的最后一个处理步骤。应当注意, 信号的限带 QMF 表示考虑到了限带残差技术和“残差填充”技术, 它们可以被结合到该类型的解码器中。

[0013] 上面的编码结构非常适合低的比特率 (通常小于 80kb/s), 但是从计算的复杂度来将, 它对于较高的比特率不是最佳的。更准确地, 在较高的比特率下, 通常不使用 SBR 工具 (由于它不会提高编码效率)。然后, 在没有 SBR 级的解码器中, 仅复值上混合矩阵使 QMF 滤波器组存在, 其在计算上需要并引入延迟 (在 1024 个样本的帧长度中, QMF 分析 / 综合滤波器组引入 961 个样本的延迟)。这很清楚地表明了对更加有效的编码结构的需要。

发明内容

[0014] 本发明的目的是提供用于立体声编码的方法及装置, 该方法和装置在高的比特率范围中也有计算上的效率。

[0015] 本发明通过分别提供如在独立权利要求中限定的编码器和解码器、编码和解码方法以及用于编码和解码的计算机程序来实现该目的。从属权利要求限定了本发明的实施方式。

[0016] 在第一方面, 本发明提供了一种用于通过复合预测立体声编码来提供立体声信号的解码器系统, 所述解码器系统包括:

[0017] 上混合级, 所述上混合级基于下混合信号 (M) 和残差信号 (D) 的第一频域表示来生成立体声信号, 第一频域表示中的每一个包括在多维空间的第一子空间中表达的表示对应信号的频谱含量的第一频谱分量, 上混合级包括:

[0018] 模块, 所述模块用于基于下混合信号的第一频域表示来计算下混合信号的第二频

域表示,第二频域表示包括在多维空间的第二子空间中表达的表示信号的频谱含量的第二频谱分量,第二子空间包括在在第一子空间中没有包括的多维空间的一部分;

[0019] 加权加法器,所述加权加法器基于下混合信号的第一频域表示和第二频域表示、残差信号的第一频域表示和在比特流信号中编码的复合预测系数(α)来计算侧信号(S);以及

[0020] 和与差级,所述和与差级用于基于下混合信号和侧信号的第一频域表示来计算立体声信号,

[0021] 其中,上混合级还以直通模式来操作,其中,下混合信号和残差信号被直接提供给和与差级。

[0022] 在第二方面,本发明提供了用于通过复合预测立体声编码由比特流信号来对立体声信号进行编码的编码器系统,包括:

[0023] 估计器,所述估计器用于对复合预测系数进行估计;

[0024] 编码级,所述编码级能够操作用于:

[0025] (a) 以由复合预测系统的值确定的关系来将立体声信号变换为下混合信号和残差信号的频域表示;以及

[0026] 复用器,所述复用器从编码级和估计器接收输出并且使用比特流信号对该输出进行编码。

[0027] 在本发明的第三与第四方面中,提供了用于将立体声信号编码为比特流并且将比特流解码为至少一种立体声信号的方法。每种方法的技术特征分别与编码器系统和解码器系统的技术特征相似。在第五与第六方面,本发明还提供了包括用于在计算机上执行所述方法中的每一种的指令的计算机程序产品。

[0028] 本发明获益于 MPEG USAC 系统中的统一立体声编码的优点。在较高比特率下也保留了这些优点,在高比特率下,在没有伴随基于 QMF 方法的计算复杂度的明显增加的情况下通常不使用 SBR。这是可能的,因为至少在下混合声道与残差声道的编码音频带宽相同并且上混合处理不包括去相关的情况下,作为 MPEG USAC 变换编码系统的基础的临界采样的 MDCT 变换可以用于如由本发明提供的复合预测立体声编码。这意味着不再需要额外的 QMF 变换。QMF 域中的复合预测立体声编码的代表性实施相比传统的 L/R 或 M/S 立体声可以实际明显地增加每单位时间的操作数量。因此,根据本发明的编码装置在该种比特率上显现出竞争力,从而以适度的计算花销提供了高的音质。

[0029] 如本领域普通技术人员所意识到的,上混合级的还可以以直通模式来操作的事实使得解码器能够根据常规的直接或联合编码以及如在编码器侧确定的复合预测编码来自适应地进行解码。因此,在解码器不能够肯定地增加质量水平以超过常规的直接 L/R 立体声编码或联合的 M/S 立体声编码的质量水平的情况下,可以至少保证维持相同的水平。因此,从功能观点来看,可以认为根据本发明的此方面的解码器是与现有技术有关的超集。

[0030] 作为基于 QMF 的预测编码立体声的优点,可以对信号进行理想的重构(除了可以为任意小的量化误差)。

[0031] 因此,本发明提供了用于通过复合预测进行基于变换的立体声编码的编码装置。优选地,根据本发明的装置不限于复合预测立体声编码,而且也可以根据现有技术的直接 L/R 立体声编码或联合 M/S 立体声编码方式操作,以使得可以针对具体的应用或在具体时

间间隔期间选择最适合的编码方法。

[0032] 信号（包括第一频谱分量与第二频谱分量）的过采样（如，复合）表示用作对于根据本发明的复合预测的基础，因此，将用于计算这种过采样表示的模块布置在根据本发明的编码器系统和解码器系统中。频谱分量是指多维空间的第一子空间和第二子空间，其可以是关于以有限采样频率采样的给定长度（如，预定义的时间帧长度）的间隔的时变函数的集合。以下是公知的：在此具体的多维空间中的函数可以通过基函数的有限加权和来逼近。

[0033] 如本领域普通技术人员应当理解的，适于与解码器协作的编码器配备有用于提供过采样表示的等同模块（预测编码基于该过采样表示），以使得能够对已编码信号进行可靠的再现。这些等同模块可以是相同或相似的模块或是具有相同或相似的传递特性的模块。具体地，分别位于编码器和解码器中的模块可以是执行用于进行等同的数学运算集合的各个计算机程序的类似或不同的处理单元。

[0034] 在解码器系统或编码器系统的一些实施方式中，第一频谱分量具有表达在第一子空间的实数值，第二频谱分量具有表达在第二子空间的虚数值。第一频谱分量和第二频谱分量一起形成了信号的复合频谱表示。第一子空间可以是第一组基函数的线性生成，而第二子空间可以是第二组基函数的线性生成，第二组基函数中的一些基函数线性地独立于第一组基函数。

[0035] 在一种实施方式中，用于计算复合表示的模块是实到虚变换，即，用于基于信号的实谱（real spectral）表示来计算离散时间信号的频谱的虚部的模块。该变换可以基于精确或近似的数学关系，例如来自谐波分析或启发式关系的公式。

[0036] 在解码器系统或编码器系统的一些实施方式中，第一频谱分量可通过离散时域信号的时域到频域的变换（优选地，为傅里叶变换）来获得，例如，通过离散余弦变换（DCT）、改进的离散余弦变换（MDCT）、离散正弦变换（DST）、改进的离散正弦变换（MDST）、快速傅里叶变换（FFT）、基于素因子的傅里叶算法等来获得。在前四种情况下，第二频谱分量可然后分别通过 DST、MDST、DCT 和 MDCT 来获得。如公知的，在单位间隔上是周期性的余弦的线性生成形成了没有完全包含在相同间隔的正弦周期的线性生成的子空间。优选地，第一频谱分量可由 MDCT 获得以及第二频谱分量可由 MDST 获得。

[0037] 在一种实施方式中，解码器系统包括至少一个时域噪声整形模块（TNS 模块，或 TNS 滤波器），其布置在上混合级的上游。一般来讲，TNS 的使用增加了对于具有类瞬态分量的感知音质，并且这也适用于以 TNS 为特征的本发明的解码器系统的实施方式。在常规的 L/R 和 M/S 立体声编码中，TNS 滤波器可以被应用作为紧在逆变换前的频域中的最后一个处理步骤。但是，在复合预测立体声编码的情况下，对下混合信号和残差信号应用 TNS 滤波器通常是有利的（即，在上混合矩阵之前）。换言之，将 TNS 应用到左声道和右声道的线性组合，这具有若干优点。首先，可以在给定的情况下证明：TNS 仅对例如下混合信号是有益的。然后，可以针对残差信号来抑制或省略 TNS 滤波，其可能意味对可用带宽的更加经济的使用，TNS 滤波器系数仅需要被传递用于下混合信号。其次，复合预测编码所需的对下混合信号的过采样表示的计算（如，从 MDCT 数据中导出 MDST 数据以形成复合频域表示）可能需要下混合信号的时域表示是可计算的。而这又意味着：优选地，可获得下混合信号作为以统一的方式获得的 MDCT 频谱的时间序列。如果在将下混合表示 / 残差表示转换为左 / 右

表示的上混合矩阵后在解码器中应用 TNS 滤波器,则仅仅下混合信号的 TNS 残差 MDCT 频谱的序列会是可用的。这将使得对应的 MDST 频谱的有效计算非常有挑战性,尤其是如果左/右声道使用了具有不同特性的 TNS 滤波器。

[0038] 应当强调的是,MDCT 频谱的时间序列的可用性不是为了获得适于用作复合预测编码的基础的 MDST 表示的绝对准则。除了实验证据以外,该事实可以用 TNS 一般仅适用于较高频率(例如大于几千赫兹)来解释,以使得经 TNS 滤波的残差信号近似与针对较低频率的非滤波残差信号对应。因此,本发明可以实施为针对复合预测立体声编码的解码器,其中,如下面所指出的,TNS 滤波器与上混合级的上游具有不同的布置。

[0039] 在一种实施方式中,解码器系统包括位于上混合级的下游的至少一个另外的 TNS 模块。借助于选择器布置,可以选择上混合级上游的 TNS 模块或上混合级下游的 TNS 模块。在某些环境下,复合频域表示的计算不需要下混合信号的时域表示是可计算的。此外,如上面所提出的,解码器可以选择性地操作在直接或联合编码模式下而不应用复合预测编码,则它可以更加适合于应用常规的 TNS 的定位,即,作为在频域中的最后处理步骤之一。

[0040] 在一种实施方式中,解码器系统适于当不需要下混合信号时通过停用用于计算下混合信号的第二频域表示的模块来节省处理资源和可能的能量。假设下混合信号被分割成连续的时间块,其中每个时间块与复合预测系数的值相关联。该值可以通过与解码器协作的编码器针对每个时间块采取的决定来确定。此外,在此实施方式中,用于计算下混合信号的第二频域表示的模块适于在对于给定的时间块,复合预测系数的虚部的绝对值为 0 或小于预定的容差时停用自己。模块的停用可以暗示着没有下混合信号的第二频域表示是针对该时间块计算的。如果停用没有发生,则第二频域表示(如,一组 MDST 系数)可以乘以 0 或与解码器的机器精度(舍入单位)基本上是同一数量级的数或某个其他适合的阈值。

[0041] 在前一实施方式的进一步开发中,在从下混合信号分割成的时间块的子水平上实现处理资源的节省。例如,这种时间块内的子水平可以是频带,其中,编码器确定对于时间块内的每个频带的复合预测系数的值。类似地,用于产生第二频域表示的模块适于抑制其对于复合预测系数为 0 或具有小于容差的大小的时间块中的频带的运算。

[0042] 在一种实施方式中,第一频谱分量是布置在变换系数的一个或更多个时间块中的变换系数,每个块是通过应用对时域信号的时间片段的变换来生成的。此外,用于计算下混合信号的第二频域表示的模块适于:

[0043] • 从第一频谱分量中的至少一些来导出一个或更多个第一中间分量;

[0044] • 根据一个或更多个脉冲响应来形成上述一个或更多个第一中间分量的组合以获得一个或更多第二中间分量;以及

[0045] • 从上述一个或更多个第二中间分量导出一个或更多个第二频谱分量。

[0046] 此步骤实现了直接从第一频域表示来对第二频域表示进行计算,如在 US6,980,933 B2 中所更加详细地描述的,特别是第 8 栏到第 28 栏,具体是其中的公式 41。如本领域普通技术人员应当意识到的,该计算不是通过时域来进行的,与如跟随有不同变换的逆变换形成对照。

[0047] 对于根据本发明的复合预测立体声编码的示例性实施,估计相比传统的 L/R 或 M/S 立体声,计算复杂度仅稍微地增加(明显地小于由在 QMF 域中的复合预测立体声编码引起的增加)。包括对第二频谱分量的精确计算的该类型的实施方式引入了通常仅比由基于

QMF 的实施所引入的延迟长了几个百分比的延迟（假设时间块长度为 1024 个样本并且与是 961 个样本的混合 QMF 分析 / 综合滤波器组的延迟相比）。

[0048] 适合地,在先前实施方式中的至少一些中,脉冲响应适于可获得第一频域表示的变换（更精确地,根据其频率响应特性来获得第一频域表示）。

[0049] 在一些实施方式中,下混合信号的第一频域表示是通过结合一种或更多种分析窗函数（或截断函数,例如,矩形窗、正弦窗、Kaiser-Bessel-derived 窗等）应用的变换来获得的,其一个目的是在不引入有害量的噪声或以不期望的方式改变频谱的情况下实现时间分割。可能地,这些窗函数是部分地交叠。然后,优选地,变换的频率响应特性取决于上述一个或更多个分析窗函数的特性。

[0050] 仍参考以在频域内对第二频域表示进行计算为特征的实施方式,可以通过使用近似的第二频域表示来减小所涉及的计算负载。这种近似可以不需要计算所基于的完整信息而实现。通过 US 6,980,933 B2 的教导,例如,对一个块（即,与输出块、前一块和后一块同时存在的块）中的下混合信号的第二频域表示的精确计算需要来自三个时间块的第一频域数据。出于根据本发明的复合预测编码的目的,可以通过省略或用 0 来代替从后一块（从而,模块的运算可能变成因果性的,即,没有贡献延迟）和 / 或从前一块发出的数据来获得适合的近似,以使得第二频域表示的计算仅基于来自以一个或两个时间块的数据。应当注意,如上面所提到的,即使输入数据的省略可能暗示着第二频域表示的尺度改变（从如不再表示相等功率的意义上来讲）,只要他在编码器端和解码器端两者中是以等同的方式来计算的,则它还能够用作复合预测编码的基础。实际上,可能的该类型的尺度改变可以由预测系数值的对应变化来补偿。

[0051] 用于计算形成下混合的第二频域表示的一部分的频谱分量的又一种近似的方法可以包括对来自第一频域表示的至少两个分量进行组合。来自第一频域表示的至少两个分量可以关于时间和 / 或频率相邻。替代地,可以以相对地少的抽头 (tap) 通过有限脉冲响应 (FIR) 滤波来对它们进行组合。例如,在应用 1024 时间块大小的系统中,这些 FIR 滤波器可以包括 2 个、3 个、4 个等抽头。此特性的近似计算方法的描述例如可以在 US 2005/0197831 A1 中找到。如果使用给每个时间块边界的附近相对地较小的权重的窗函数（例如,非矩形函数）,则可以便利地使时间块中的第二频谱分量仅基于同一时间块中的第一频谱分量的组合,从而暗示着对最外面的分量可用的信息量不是相同的信息量。窗函数的形状在一定程度上抑制或消除了由这种实践可能地引入的近似误差。

[0052] 在设计成输出时域立体声信号的解码器的一种实施方式中,包括了在直接或联合立体声编码与复合预测编码之间切换的可能性。这是通过提供以下来实现的：

[0053] • 可以选择性地作为直通级（不对信号进行修改）或作为和与差变换来操作的开关；

[0054] • 用于进行频率到时间的变换的逆变换级；以及

[0055] • 用于使用直接地（或联合地）编码的信号或由复合预测编码的信号来供给逆变换级的选择器装置。

[0056] 如本领域普通技术人员应当意识到的,关于解码器的该部分的灵活性给予编码器在常规的直接或联合编码与复合预测编码之间选择的余地。因此,在不能够超越常规的直接 L/R 立体声编码或联合 M/S 立体声编码的质量水平的情况下,该实施方式能够至少保证

维持相同的水平。因此,根据此实施方式的解码器可以认为是关于相关领域的超集。

[0057] 解码器系统的另一组实施方式实现通过时域来对第二频域表示中的第二频谱分量进行的计算。更精确地,应用由其获得了(可获得)第一频谱分量的变换的逆过程,并且变换的逆过程跟随具有将第二频谱分量作为输出的不同的变换。具体地,可以在 MDCT 之后跟随有 MDST。为了减小变换及逆变换的数量,在这种实施方式中,可以将逆 MDCT 的输出提供给 MDST 和解码系统的输出终端(可能有之前的其它处理步骤)两者。

[0058] 对于根据本发明的复合预测立体声编码的示例性实施,估计到:相比传统的 L/R 或 M/S 立体声而言,计算复杂度仅轻微地增加(但仍明显地小于由 QMF 域中的复合预测立体声编码所引起的增加)。

[0059] 作为在前一段所提及的实施方式的进一步开发,上混合级可以包括用于处理侧信号的另外的逆变换级。然后,给和与差级提供由上述另外的逆变换级生成的侧信号的时域表示和由已经提及的逆变换级生成的下混合信号的时域表示。回想起,有利地从计算复杂度的观点来看,将后一信号提供给和与差级和上面所提及的上述不同的变换级。

[0060] 在一种实施方式中,设计成输出时域立体声信号的解码器包括在直接 L/R 立体声编码或联合 M/S 立体声与复合预测立体声编码之间切换的可能性。这通过提供以下来实现:

[0061] • 可作为直通级或和与差级操作的开关;

[0062] • 用于计算侧信号的时域表示的另外的逆变换级;

[0063] • 选择器装置,该选择器装置用于将逆变换级连接到连接至上混合级的上游与开关的下游的点的另外的和与差级(优选地,当开关被驱动用作通带滤波器时,其可以是对由复合预测编码生成的立体声信号进行解码的情况),或连接到来自开关的下混合信号与来自加权加法器的侧信号的组合(优选地,当开关被驱动用作和与差级时,其可以是对直接编码立体声信号进行解码的情况)。

[0064] 如本领域普通技术人员所意识到的,这给出了编码器在常规的直接或联合编码与复合预测编码之间进行选择的余地,其意味着能够保证至少与直接或联合编码等同的质量水平。

[0065] 在一种实施方式中,根据本发明的第二方面的编码器系统可以包括为了使残差信号的信号功率或平均信号功率减小或最小化而对复合预测系数进行估计的估计器。最小化可以发生在时间间隔上,优选地,在待编码的信号的时间片段或时间块或时间帧上。幅度的平方可以作为瞬时信号功率的测量值,并且在平方后的幅度(波形)的时间间隔上的积分可以作为在该间隔上的平均信号功率的测量值。适当地,复合预测系数是基于时间块和频带来确定的,即,它的值以以下方式来设置:它减小了在该时间块和频带中的残差信号的平均功率(即,总能量)。具体地,用于对如 IID、ICC 和 IPD 等参数立体声编码参数进行估计的模块可以提供基于其能够根据普通技术人员已知的数学关系来计算复合预测系数的输出。

[0066] 在一种实施方式中,编码器系统的编码阶段是可操作进一步地用作直通级,以使得能够进行直接的立体声编码。在期望提供较高质量的情况下通过选择直接的立体声编码,编码器系统能够保证编码后的立体声信号至少具有与直接编码的质量相同的质量。类似地,在由复合预测编码引起的较大的计算努力不是以明显的质量增加为动机的情况下,

编码系统因此可获得节省计算资源的选项。编码器中的联合、直接、实数预测和复合预测之间的决定一般是基于速率 / 失真最佳化原理。

[0067] 在一种实施方式中,编码器系统可以包括用于基于第一频谱分量来直接(即,没有将逆变换应用到时域中以及没有使用信号的时域数据)计算第二频域表示的模块。关于上述解码器系统的对应实施方式,该模块可以具有类似的结构,即,包括类似的处理运算但顺序不同,以使得编码器适于输出适于作为解码器侧的输入的数据。处于说明此实施方式的目的,假设待编码的立体声信号包括中和侧声道,或已经被变换为此结构,并且编码级适于接收第一频域表示。编码级包括用于计算中声道的第二频域表示的模块。(此处所指的第一频域表示和第二频域表示是如上面所定义的;具体地,第一频域表示可以是 MDCT 表示以及第二频域表示可以是 MDST 表示。)编码级还包括用于计算作为由侧信号和分别由复合预测系数的实部和虚部加权的中信号的两个频域表示形成的线性组合的残差信号的加权加法器。中信号,或恰当地为其第一频域表示,可以直接用作下混合信号。在此实施方式中,另外,估计器为了将残差信号的功率或平均功率最小化来确定复合预测系数的值。最终的运算(最佳化)可以通过反馈控制来实现,其中,估计器可以接收通过待进一步调节的(如果需要)当前预测系数值,或以前馈的方式通过作用于原始立体声信号的左/右声道或中/侧声道的计算,来获得的残差信号。优选前馈方法,通过前馈方法并基于中信号的第一频域表示和第二频域表示以及侧信号的第一频域表示来直接地(具体地,以非迭代或非反馈的方式)确定复合预测系数。应当注意,复合预测系数的确定可以跟随在是否应用直接、联合、实数预测或复合预测编码的决定之后,其中,考虑了每个可用选项的作为结果的质量(优选地,将感知质量考虑在内,例如,信掩比效果);因此,上面所陈述的不应当解释为在编码器中不存在反馈机制的结果。

[0068] 在一种实施方式中,编码器系统包括用于通过时域来计算中(或下混合)信号的第二频域表示的模块。应当理解,至少就所关注的第二频域表示的计算而言,与此实施方式有关的实施细节与对应的解码器实施方式类似或能够做得与对应的解码器类似。在此实施方式中,编码级包括:

[0069] • 用于将立体声信号转换为包括有中声道和侧声道形式的和与差级;

[0070] • 用于提供侧声道的频域表示和中声道的复合值的(因此,过采样的)频域表示的变换级;以及

[0071] • 用于计算残差信号的加权加法器,其中,复合预测系数用作权重。此处,估计器可以接收残差信号并且可能以反馈控制方式来确定复合预测系数,以使残差信号的功率或平均功率减小或最小化。但是,优选地,估计器接收待编码的立体声信号并且基于该信号来确定预测系数。从计算经济性的观点来看有利的是使用侧声道的临界采样的频域表示,因为在本实施方式中后者不会与复数相乘。适当地,变换级可以包括并行布置的 MDCT 级和 MDST 级,两者皆具有作为输入的中声道的时域表示。因此,产生了中声道的过采样频域表示和侧声道的临界采样的频域表示。

[0072] 注意,经过在包括有例行实验的能力的技术人员能力内的适当修改后,本部分所公开的方法及装置可以应用到具有多于两个声道的信号的编码。对这种对于多声道可操作性的修改可以按照上述 J. Herre 等人的论文中的第 4 部分和第 5 部分中的各行来进行。

[0073] 在另外的实施方式中,除非来自上面所概述的两种或更多种实施方式的特征是互

余的,否则可以将它们进行组合。在不同的权利要求中列出两个特征的这一事实不排除能够将它们有利地组合。同样地,另外的实施方式还能够忽略对于期望的目的是不需要的或不是必需的某些特征。作为一个示例,根据本发明的解码系统可以在没有去量化级情况下来实施,在这些情况下,待处理的已编码信号没有被量化或已经是适于由上混合级进行处理的可用形式。

附图说明

[0074] 将参考附图通过在接下来的部分中描述的实施方式来对本发明做进一步说明,在附图中:

[0075] 图 1 示出了根据现有技术的基于 QMF 的解码器的两个概括框图组成;

[0076] 图 2 是根据本发明实施方式的利用复合预测的基于 MDCT 立体声解码器系统的概括框图,其中,待解码的信号的声道的复合表示是在频域中计算出来的;

[0077] 图 3 是根据本发明实施方式的利用复合预测的基于 MDCT 立体声解码器系统的概括框图,其中,待解码的信号的声道复合表示是在时域中计算出来的;

[0078] 图 4 示出了图 2 中的解码器系统的替代实施方式,其中,主动 TNS 级的位置是可选的;

[0079] 图 5 包括示出了根据本发明的另一个方面的实施方式的利用复合预测的基于 MDCT 的立体声编码器系统的概括框图;

[0080] 图 6 是根据本发明的实施方式的利用复合预测的基于 MDCT 的立体声编码器的概括框图,其中,待编码的信号的声道的复合表示是基于其时域表示计算出来的;

[0081] 图 7 示出了图 6 中的编码器系统的替代实施方式,其也可操作在直接 L/R 编码模式;

[0082] 图 8 示出了根据本发明的实施方式的利用复合预测的基于 MDCT 的立体声编码器系统的概括框图,其中,待编码的信号的声道的复合表示是基于其第一频域表示计算出来的,其解码器系统也可操作在直接 L/R 编码模式下;

[0083] 图 9 示出了图 7 中的编码器系统的替代实施方式,其还包括有布置在编码级的下游的 TNS 级;

[0084] 图 10 示出了图 2 和图 8 中的标记为 A 的部分的替代实施方式;

[0085] 图 11 示出了图 8 中的编码器系统的替代实施方式,其还包括分别布置在编码级的下游和上游的两个频域修改器件;

[0086] 图 12 来自 6 个对象的在 96kb/s 处的收听测试结果的图示,其显示了针对 MDST 频谱的计算或近似的不同的复杂度-质量权衡选项,其中,标记为“+”的数据点是指隐藏参考,“×”是指 3.5kHz 限带锚,“✱”是指 USAC 传统的立体声 (M/S 或 L/R),“□”是指通过使用禁用了预测系数的虚部的复合预测来进行的 MDCT 域统一立体声编码(即,实值预测,不需要 MDST)。“■”是指通过使用当前 MDCT 帧以计算 MDST 的近似来进行的 MDCT 域统一立体声编码,“○”是指通过使用当前的和前一 MDCT 帧以计算 MDST 近似来进行的 MDCT 域统一立体声编码以及“●”是指通过使用当前帧、前一个和下一个 MDCT 帧计算 MDST 的复合预测来进行的 MDCT 域统一立体声编码;

[0087] 图 13 呈现了图 12 中的数据,但是,是作为与通过使用当前 MDCT 帧计算 MDST 的近

似的复合预测来进行的 MDCT 域统一立体声编码有关的差分分数；

[0088] 图 14 包括示出了根据本发明实施方式的解码器系统的三种实施方式的概括框图；

[0089] 图 15 是示出了根据本发明实施方式的解码方法的流程图；以及

[0090] 图 16 是示出了根据本发明实施方式的编码方法的流程图。

具体实施方式

[0091] I. 解码器系统

[0092] 图 2 以一般框图的形式示出了一种解码系统，该解码系统用于对包括复合预测系数 $\alpha = \alpha_R + i\alpha_I$ 和具有下混合声道 M 和与残差声道 D 的立体声信号的 MDCT 表示中的至少一个值的比特流进行解码。可以对预测系数的实部 α_R 和虚部 α_I 进行量化和 / 或联合地编码。但是，优选地，通常以 0.1（无量纲数）的步长来独立与统一地对实部和虚部进行量化。用于复合预测系数的频带分辨率不是必需与用于根据 MPEG 标准的比例因子频带（sfb；即，使用相同 MDCT 量化步长和量化范围的一组 MDCT 线）的分辨率相同。具体地，用于预测系数的频带分辨率可以是在心里声学中证明了的频带分辨率（例如吠声规模）。去复用器 201 适于从对其提供的比特流中提取这些 MDCT 表示和预测系数（如图中所指出的控制信息的一部分）。事实上，在比特流中还可以对不仅仅是复合预测系数的更多的控制信息进行编码（例如，将比特流在预测模式还是在非预测模式下解码的指令、TNS 信息等）。TNS 信息可以包括要由解码器系统的 TNS（综合）滤波器应用的 TNS 参数的值。如果 TNS 参数的相同集合要用于若干 TNS 滤波器（例如，用于两个声道），则与以独立地接收两组参数相比，以接收指示该种等同的参数组的比特的形式的信息是经济的。还可以包括是否在上混合级前后应用 TNS 的信息（如适当地基于如两个可用选项的心理声学评估）。此外，控制信息可以单独地指示下混信号和残差信号的限定带宽。对于每个声道，不会对大于带宽极限的频带进行解码而是将其设置为 0。在某些情况下，最高的频带具有如此小的能量含量以至于它们已经被量化为 0。普通实践（参照 MPEG 标准中的参数 max_sfb）已经对下混合信号和残差信号两者使用相同的带宽限制。但是，残差信号相比下混合信号在较大程度上使其能量含量集中到较低的频带。因此，通过对残差信号布置专用的较高的带宽极限，在没有明显的质量损失的情况下，比特率的减小是可能的。例如，这可以由在比特流中编码的两个独立的 max_sfb 参数来控制，一个用于下混合信号，一个用于残差信号。

[0093] 在本实施方式中，立体声信号的 MDCT 表示被分割成包括有固定数量的数据点（如，1024 点）、若干固定数量的数据点（如，128 或 1024 点）之一和变换数量的点的连续的时间帧（或时间块）。如本领域普通技术人员所已知的，MDCT 是临界采样。在图的右部指出的解码器系统的输出是具有左 L 声道和右 R 声道的时域立体声信号。去量化模块 202 适于处理输入到解码器系统中的比特流，或者如果需要，处理在对原始的比特流进行去复用之后获得的、与下混合声道和残差声道中的每一个对应的两个比特流。将经去量化声道信号提供给可操作在与下面的各个变换矩阵对应的直通模式或和与差模式的切换组件 203：

[0094]
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

[0095] 以及

$$[0096] \quad \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

[0097] 如在下一段进一步讨论的,解码器系统包括第二切换组件 205。与在此实施方式和待描述的其他实施方式中的大多数其他开关以及切换组件一样,切换组件 203、205 可以频率选择性方式来操作。这使得在相关领域已知的各种解码模式的解码有效,例如,取决于解码频率的 L/R 或 M/S 解码。因此,可以将根据本发明的解码器认为是与相关领域有关的超集。

[0098] 现假设切换组件 203 处于直通模式,在本实施方式中,去量化声道信号通过各个 TNS 滤波器 204。TNS 滤波器 204 对于解码系统的操作不是必需的而是可以由直通元件来取代。此后,将信号提供给与位于上游的切换组件 203 具有相同功能的第二切换组件 205。利用如先前描述的输入信号和设置在直通模式的第二切换组件 205,第二切换组件 205 的输出是下混合信号和残差信号。将仍由其时间连续的 MDCT 频谱表示的下混合信号提供给适于基于下混合信号的时间连续的 MDCT 频谱表示来计算下混合信号的 MDST 频谱的实到虚变换 206。在此实施方式中,一个 MDST 帧是基于三个 MDCT 帧、一个先前帧、一个当前(或同时)帧和一个后续帧。 (Z^{-1}, Z) 象征性地表示实到虚变换 206 的输入侧包括延迟分量。

[0099] 从实到虚变换 206 获得的下混合信号的 MDST 表示是由预测系数的虚部 α_I 来加权,并且被增加到由预测系数的实部 α_R 加权的下混合信号的 MDCT 表示和残差信号的 MDCT 表示。两个相加和相乘是通过一起形成(功能性地)加权加法器的乘法器与加法器 210、211 来执行的,加权加法器是由在解码器系统所初始接收的比特流中编码的复合预测系数 α 的值来提供的。可以为每个时间帧确定一次复合预测系数。还可以更加频繁地确定,例如,对一个帧内的每个频带确定一次,该频带是根据心理声学来划分的。还可以较不频繁地确定,如下面将联系根据本发明的编码系统来讨论的。实到虚变换 206 与加权加法器以以下方式来同步:将下混合声道信号的当前 MDST 帧与各个下混合声道信号和残差声道混合信号中的一个同时 MDCT 帧进行组合。这三个信号的和是侧信号 $S = \text{Re}\{\alpha M\} + D$ 。在该表达式中, M 包括下混合信号的 MDCT 和 MDST 表示,即, $M = M_{\text{MDCT}} - iM_{\text{MDST}}$,但是, $D = D_{\text{MDCT}}$ 是实值。因此,获得了具有下混合声道和侧声道的立体声信号,其中和与差变换 207 如下存储左声道和右声道:

$$[0100] \quad \begin{cases} L = M + S \\ R = M - S \end{cases}$$

[0101] 这些信号被表示在 MDCT 域。解码系统的最后一步是将逆 MDCT 209 应用到每个声道,由此获得左/右立体声信号的时域表示。

[0102] 如上面所提到的,在申请人的专利 US 6,980,933 B2 中对实到虚变换 206 的可能的实现做了进一步的描述。通过其中的公式 41,变换能够表达为有限脉冲响应滤波器,例如,对于偶数点,

$$S(2v) = \frac{2}{N} \sum_{p=0}^{N-1} [(-1)^{p+1} X_I(p) + X_{III}(p)] h_{I,III}(2v-p) \\ + \frac{4}{N} \sum_{l=0}^{N-1} X_{II}(2l+1) h_{II}(2v-2l-1),$$

[0103] 其中, $S(2v)$ 是 $2v^{\text{th}}$ MDST 数据点, X_I , X_{II} , X_{III} 是来自每个帧的 MDCT 数据, 以及 N 是帧长。此外, $h_{I,III}$, h_{II} 是取决于所应用的窗函数的脉冲响应, 因此, 是针对窗函数 (如矩形窗、正弦窗和凯萨尔-贝塞尔导出窗等) 的每个选择以及针对每个帧长来确定的。可以通过省略具有相对地较小的能量含量和对 MDST 数据贡献相对地较少的那些脉冲响应来减小该计算的复杂度。作为该简化的替代或扩展, 可以缩短脉冲响应自身, 如, 从整个帧长到较小数量的点。例如, 脉冲响应长度可以从 1024 点 (抽头) 减少到 10 点。仍被认为有意义的最极端的截取是:

$$[0105] \quad S(v) = X_{II}(v+1) - X_{II}(v-1)$$

[0106] 其他直接的方法能够在 US 2005/0197831 A1 中找到。

[0107] 还可以减小计算所基于的输入数据的量。为了说明, 用在图 10 中示出的简化的变量 A' 和 A'' 取代在图中的“ A ”来表示实到虚变换 206 和其上游联系。变量 A' 提供了信号的近似的虚数表示。此处, MDST 计算仅将当前帧和前一帧考虑在内。参考此段中的上面的公式, 可以通过设置 $X_{III}(p) = 0$, $p = 0, \dots, N-1$ (索引 III 表示后一时间帧) 来实现。因为变量 A' 不需要作为输入的后一帧的 MDCT 频谱, 所以 MDST 计算不会引起任何时间延迟。明显地, 该近似一定程度上减小了所获得的 MDST 信号的精确度, 但是还可以暗示着该信号的能量变小了; 信号的能量变小这一事实能够通过由预测编码的性质导致的 α_I 的增加来补偿。

[0108] 在图 10 中还示出了变量 A'' , 其仅将当前时间帧的 MDCT 数据用作输入。相比变量 A' , 变量 A'' 可论证地产生了较为不精确的 MDST 表示。另一方面, 它可操作在 0 延迟, 就像变量 A' 一样, 并且具有较低的计算复杂度。如已经提到的, 只要在编码器系统和解码器系统中使用相同的近似, 则波形编码属性是不受影响的。

[0109] 应当注意, 无论是否使用变量 A 、 A' 或 A'' 或任何另外的推导, 仅有复合预测系数的虚部非 0 ($\alpha_I \neq 0$) 的 MDST 频谱的那些部分需要计算。在实际的环境中, 这会被采用以意味着系数的虚部的绝对值 $|\alpha_I|$ 大于 $|\alpha_I|$ 能与所使用的硬件的单位舍入有关的预定的阈值。在系数的虚部对于时间帧内的所有频带都为 0 的情况, 不需要计算针对该帧的任何 MDST 数据。因此, 适当地, 实到虚变换 206 适于通过不生成 MDST 输出来对非常小的 $|\alpha_I|$ 值的出现做出响应, 因此, 能够节省计算资源。但是, 在使用不止当前帧的多个帧来产生一帧 MDST 数据的情况下, 即使不需要 MDST 频谱, 变换 206 的任何上游单元也应该适当地继续操作, 具体地, 第二切换组件 205 应该保留前向 MDCT 频谱以使得当与非 0 预测系数相关联的下一个时间帧出现时对于实到虚变换 206 已经有充分的输入数据是可用的; 这当然可以是下一个时间块。

[0110] 又参照图 2, 解码系统的功能在假设切换组件 203、205 两者被设置在它们相应的

直通模式的情况下已经做了描述。如现将要讨论的,解码器系统也能够对非预测编码的信号进行解码。对于此用途,第二切换组件 205 可以设置在和与差模式,并且适当地,如在图 1 中所指出的,在其下游位置中可以设置有选择器装置 208,从而,保证了将信号从 TNS 滤波器 204 与第二切换组件 205 之间的源点直接输送到逆变换 209。为了保证正确的解码,信号合适地具有在源点处的 L/R 形式。因此,为了确保一直(而不是,比如通过左信号间断地)给实到虚变换提供正确的中(即,下混合)信号,优选地,在非预测编码立体声信号的解码期间将第二切换组件 205 设置在和与差模式。如上所提到的,基于例如数据速率对音频质量判断,对于某些帧可以由常规的直接或联合编码取代预测编码。这种判断的结果可以以各种方式从编码器对解码器通信,例如,通过每个帧中的专用指示符比特的值,或通过预测系数的值的缺失或存在。已经建立了这些事实,所能够容易地实现第一切换组件 203 的作用,实际上,在非预测编码模式下,解码器系统能够根据直接(L/R)立体声编码或联合(M/S)编码以及通过在直通或和与差模式下操作第一切换组件 203 来处理两种信号,可以确保总是给原点提供已直接地编码信号。明显地,切换组件 203 当用作和与差级时,将 M/S 形式的输入信号转换成 L/R 形式的输出信号(提供给可选的 TNS 滤波器 204)。

[0111] 解码器系统接收表示具体的时间帧是由解码器系统以预测编码模式或是以非预测编码模式来解码的信号。非预测模式可以由每个帧中的专用指示符比特的值或通过预测系数的缺失(或 0 值)来发信号通知。预测模式可以类似地来表达。具体地有利的实施(其在没有任何开销的情况下使得能够回退)使用 2 位域 *ms_mask_present* (见, MPEG-2AAC, 文献 ISO/IEC 13818-7) 的保留的第四个值,每时间帧对其进行传输并且将其定义如下:

[0112]

表 1: USAC 中 <i>ms_mask_present</i> 的定义	
值	含义
00	对于所有频带的 L/R 编码
01	每频带的一个信令比特用于指示
10	对于所有频带的 M/S 编码
11	保留

[0113] 通过对值 11 重新定义以表示“复合预测编码”,解码器能够操作在所有传统模式(具体地, M/S 和 L/R 编码)下而没有任何比特损失,并且还能够接收指示用于相关帧的复合预测编码模式的信号。

[0114] 图 4 示出了与图 2 中示出的解码器系统具有相同的总体结构的解码器系统,但包括至少两个不同的结构。首先,图 4 的系统包括开关 404、411,其使得能够在上混合级的上游和/或下游应用一些涉及频域修改的处理步骤。另一方面,这是通过与第一开关 404 一起提供的第一组频域修改器 403(在该图中画为 TNS 综合滤波器)来实现的,第一开关 404 位于去量化模块 401 和第一切换组件 402 的下游并且位于紧布置在上混合级 406、407、408、409 的上游的第二切换组件 405 的上游。另一方面,解码器系统包括与第二开关 411 一起提供的第二组频域修改器 410,第二开关 411 位于上混合级 406、407、408、409 的下游并且位于逆变换级 412 的上游。有利地,如在图上所指示的,与在上游连接到频域修改器的输入侧的直通线并行地布置每个频域修改器,并且每个频域修改器在下游连接至相关联的开

关。借助于此结构,给频域修改器一直提供信号数据,从而使得能够基于不止仅仅当前帧的多个时间帧来在频域中进行处理。是应用第一组频域修改器 403 还是应用第二组频域修改器 410 的决定可以由编码器来做出(并且在比特流中传送该决定),或可以基于是否应用预测编码,或可以基于所找到的适合于具体环境的一些其他准则来做出该决定。例如,如果频域修改器是 TNS 滤波器,则第一组频域修改器 403 对于一些类型的信号可能是有利的,而第二组频域修改器 410 对于其他类型的信号可能是有利的。如果将该选择的结果编码在比特流中,则解码器系统可以相应地激活各个 TNS 滤波器组。

[0115] 为了便于理解图 4 中示出的解码器系统,应当清楚地指出,当 $\alpha = 0$ (暗示着伪 L/R 和 L/R 是相同的并且侧声道与残差声道不同) 时对直接 (L/R) 编码信号进行解码,第一切换组件 402 处于直通模式,第二切换组件处于和与差模式,从而,使得信号在上混合级的第二切换组件 405 与和与差级 409 之间具有 M/S 形式。因为上混合级然后将有效地是直通步骤,所以是触发第一组频域修改器还是第二组频域修改器(使用各自的开关 404、411) 是无关紧要的。

[0116] 图 3 示出了根据本发明的实施方式解码器系统,其与图 2 和图 4 所示有关并且表示与提供上混合所需的 MDST 数据的方法不同的方法。与已经描述的解码器系统一样,图 3 的系统包括从解码器系统的输入端开始全部串行地布置的去量化模块 301、可操作在直通或和与差模式任何一种模式下的第一开关组件 302,和 TNS(综合)滤波器 303。借助于两个开关 305、310 选择性地利用该点下游的模块,优选地,如图中所指出的,开关 305、310 联合地操作从而两者处于它们的上位置或下位置。在解码器系统的输出端,存在和与差级 312,并且紧在其上游存在用于将每个声道的 MDCT 域表示变换成时域表示的两个逆 MDCT 模块 306、311。

[0117] 在复合预测解码中(其中,给解码器系统提供对下混合/残差信号和复合预测系数的值进行编码的比特流),第一切换组件 302 设置在其直通模式并且第二开关 305、310 设置在上位置。在 TNS 滤波器的下游,以不同的方式处理(去量化、TNS 滤波、MDCT) 立体声信号的两个声道。一方面,将下混合声道提供给乘法器与加法器 308(其将下混合声道的由实部 α_R 加权的 MDCT 表示相加到残差信号的 MDCT 表示),并且,另一方面,将下混合声道提供给逆 MDCT 变换模块中的一个 306。将下混合声道 M 的从逆 MDCT 变换模块 306 输出的时域表示提供给最终的和与差级 312 和 MDST 变换模块 307 两者。从计算复杂度的观点来看,下混合声道的时域表示的这种双用途是有利的。将由此所获得的下混合声道的 MDST 表示提供给另外的乘法器与加法器 309,其在预测系数的虚部 α_I 的加权之后,将该信号增加到从加法器 308 输出的线性组合;因此,加法器 309 的输出是侧声道信号, $S = \text{Re}\{\alpha M\} + D$ 。与图 2 所述的解码器系统类似,乘法器与加法器 308、309 可以容易地组合以形成具有将下混合信号的 MDCT 和 MDST 表示、残差信号的 MDCT 表示和复合预测系数的值作为输入的加权多信号加法器。在本实施方式中的该点的下游处,在将侧声道信号提供给最后的和与差级 312 之前,只剩下通过逆 MDCT 变换模块 311 的通道。

[0118] 可以通过在逆 MDCT 变换模块 306 和 311 两者处应用相同的变换长度和窗形状来获得在解码器系统中的必需的同步性,如已经在频率选择性 M/S 和 L/R 编码中所实践的。逆 MDST 模块 307 的某些实施方式与逆 MDCT 模块 306 某些实施方式的组合引入了一帧延迟。因此,设置有五个可选的延迟块 313(或用计算机实施该效果的软件指令),以使得当必需

时系统的位于虚线的右边的这部分能够延迟与左部分有关的一帧。明显地,除了逆 MDCT 模块 306 与 MDST 变换模块 307 之间的连接线(其是延迟出现需要补偿的地方),虚线与连接线之间的所有交叉处均设置有延迟块。

[0119] 针对一个时间帧的 MDST 数据的计算需要来自时域表示的一个帧的数据。但是,逆 MDCT 变换是基于一个(当前)、两个(优选地,前一个和当前)或三个(优选地,前一个、当前和后一个)连续帧。借助于公知的与 MDCT 相关联的时域混叠抵消(TDAC),三个帧的选项实现了输入帧的完整交叠,因此,至少在包括有时域混叠的帧中提供了最高(可能是完美的)的精确度。明显地,三帧逆 MDCT 以一帧操作延迟操作。通过接受将近似的时域表示用作到 MDST 变换的输入,可以避免该延迟以及因此对解码器系统的不同部分之间的延迟进行补偿的需要。在两帧选项中,使得 TDAC 的交叠/添加发生在该帧的前一半,并且混叠可以仅出现在后一半。在一帧选项中,TDAC 的缺失暗示着混叠可能发生在整个帧上;但是,以此方式实现的并且在复合预测编码中用作中间信号的 MDST 表示仍然可以提供满足的质量。

[0120] 图 3 中使出的解码系统还可以操作在两种非预测解码模式下。对于将直接 L/R 编码立体声信号进行解码,第二开关 305、310 设置在较低的位置并且第一切换组件 302 设置在直通模式下。因此,信号具有和与差级 304 上游的 L/R 形式,当逆 MDCT 变换和最终的和与差操作发生时,和与差级 304 将信号转换为 M/S 形式。对于将设置成联合 M/S 编码形式的立体声信号进行解码,替代地,将第一切换组件 302 设置在其和与差模式,以使得信号在第一切换组件 302 与和与差级 304 之间具有 L/R 形式,从 TNS 滤波的观点来看,L/R 形式比 M/S 会更加适合。和与差级 304 下游的处理与直接 L/R 解码的情况下的处理相同。

[0121] 图 14 由根据本发明实施方式的解码器的三个概括框图组成。与本发明若干其他框图相比,图 14 中的连接线可以将多声道信号符号化。具体地,这种连接线可以布置成传输包括有左/右、中/侧、下混合/残差、伪左/伪右声道以及其他组合的立体声信号。

[0122] 图 14A 示出了用于对输入信号的频域表示(出于该图的目的而被指示为 MDCT 表示)进行解码的解码系统。该解码器系统适于作为其输出来提供基于输入信号生成的立体声信号的时域表示。为了能够对通过复合预测编码来编码的输入信号进行解码,该解码器系统设置有上混合级 1410。但是,它还能够处理以可能随着时间在若干编码格式之间交替的其他格式进行编码的输入信号,例如,通过复合预测编码编码的时间帧序列后面可以跟随有通过直接左/右编码进行编码的时间部分。解码器系统处理不同编码格式的能力是通过提供与上混合级 1410 并行设置的连接线(直通)来实现的。借助于开关 1411,可以选择将来自上混合级 1410(图中的下开关位置)的输出还是将在连接线(图中的上开关位置)上可获得的未处理的信号提供到进一步布置成在下游的解码器模块。在本实施方式中,逆 MDCT 模块 1412 布置在开关的下游,其将信号的 MDCT 表示变换成时域表示。例如,提供给上混合级 1410 的信号可以是下混合/残差形式的立体声信号。上混合级 1410 则适于导出侧信号并且进行和与差操作以使得左/右立体声信号(在 MDCT 域中)被输出。

[0123] 图 14B 示出了与图 14A 的解码器系统类似的解码器系统。该系统适于接收器输入信号处的比特流。由组合的去复用器和去量化模块 1420 来初始地对该比特流进行处理。如由具有与图 14A 的开关 1411 的功能类似的功能的开关 1422 的位置所确定的,去复用器和去量化模块 1420 提供用于进一步处理的多声道立体声信号的 MDCT 表示作为第一输出信

号。更精确地,开关 1422 确定来自去复用器和去量化器的第一输出信号是由上混合级 1421 与逆 MDCT 模块 1423(下位置)来处理还是仅由逆 MDCT 模块 1423(上位置)来处理。组合的去复用器和去量化模块 1420 还输出控制信息。在该情况下,与立体声信号相关联的控制信息可以包括指示开关 1422 的上位置还是下位置适合对信号进行解码,或更加抽象地,要根据什么编码格式对立体声信号解码。控制信息还可以包括用于调节上混合级 1421 的属性的参数,例如,如上面已经描述的在复合预测编码中使用的复合预测系数 α 的值。

[0124] 图 14C 示出了除了包括与图 14B 中类似的实体外还包括有分别布置在上混合级 1433 的上游和下游的第一频域修改器件 1431 和第二频域修改器件 1432 的解码器系统。出于该图的目的,每个时域修改器件由 TNS 滤波器示出。但是,通过该术语,时域修改器件还可以被理解为除了应用在上混合级前还是上混合级后的敏感的 TNS 滤波的其他处理。频域修改的示例包括预测、噪声添加、带宽扩展和非线性处理。可能包括待处理的信号的属性和/或这种频域修改器件的设置或配置的心理声学考虑及类似的理由有时表明:将该频域修改应用在上混合级 1433 的上游而不是下游是有利的。在其他情况下,可以通过类似的考虑来确立:频域修改的下游位置相对于上游位置是优选的。借助于开关 1432、1436,频域修改器件 1431、1435 可以被选择性地激活,以使得解码器系统能够响应于控制信息来选择所期望的配置。例如,图 14C 示出了以下配置:来自组合去复用器和去量化模块 1430 的立体声信号由第一频域修改器件 1431 初始地进行处理,然后提供给上混合级 1433 并且最后直接地转发给逆 MDCT 模块 1437 而不通过第二频域修改器件 1435。如在发明内容部分所说明的,相比在复合预测编码中的上混合后进行 TNS 的选择,此配置是优选的。

[0125] II. 编码器系统

[0126] 现将参考图 5 来描述根据本发明的编码器系统,其是用于通过复合预测编码将左/右(L/R)立体声信号编码为输出比特流的编码器系统的概括框图。该编码器系统接收信号的时域或频域表示并且将其提供给下混合级和预测系数估计器两者。将预测系数的实部和虚部提供给下混合级以对左与右声道到下混合与残差声道的转换进行控制,然后将其提供给最终的复用器 MUX。如果没有将信号作为时域表示提供给编码器,则将其在下混合级或复用器中变换为这种表示。

[0127] 在预测编码中的一个原理是将左/右信号转换为中/侧形式,即

$$[0128] \quad \begin{cases} M = (L + R)/2 \\ S = (L - R)/2 \end{cases}$$

[0129] 然后利用这些声道的剩余的相关性,即,通过设置:

$$[0130] \quad S = \text{Re}\{\alpha M\} + D$$

[0131] 其中, α 是待确定的复合预测系数并且 D 是残差信号。可以选择 α 以使得残差信号的能量 $D = S - \text{Re}\{\alpha M\}$ 最小化。能量最小化可以关于瞬时功率、对于均值意义下的最佳化的离散信号量的情况下的较短期或较长期能量(能量平均)来实现。

[0132] 预测系数的实部 α_R 和虚部 α_I 可以被联合地量化和/或编码。但是,优选地,实部和虚部通常被以 0.1(无量纲数字)的步长独立地并且统一地进行量化。用于复合预测系数的频带分辨率不是必需地与用于根据 MPEG 标准的尺度因子带(sfb;即,使用相同的 MDCT 量化步长和量化范围的一组 MDCT 线)的分辨率相同。具体此,用于预测系数的频带分辨率可以是心理声学地证明了的频带分辨率(例如吠声规模)。应当注意,频带分辨率在变换长

度变化的情况下可以变化。

[0133] 如已经提到的,根据本发明的编码器系统可以具有是否应用预测立体声编码的余地,后一种情况暗示着回退到 L/R 或 M/S 编码。这种决定可以基于时间帧来做出,或者更精细地,基于时间帧内的频带来做出。如上面所提到的,该决定的否定结果可以以各种方式传达到解码实体,例如,通过每个帧中专用指示器的值,或通过预测系数的值的缺失(或 0 值)。可以类似地传达肯定的决定。具体地有利的实施(其在没有任何开销的情况下使得能够回退)使用 2 比特域 *ms_mask_present*(见 MPEG-2AAC,文献 ISO/IEC 131818-7)的保留的第四个值,每时间帧对其进行传输并且将其定义如下:

[0134]

表 1: USAC 中的 <i>ms_mask_present</i> 的定义	
值	含义
00	对于所有频带的 L/R 编码
01	每频带一个信令比特用于指示 L/R
10	对于所有频带的 M/S 编码
11	保留

[0135] 通过重定义值 11 来表示“复合预测编码”,编码器能够操作在所有的传统模式(具体地,M/S 和 L/R 编码)下而没有任何比特损失,并且还能够用信号通知对于一些帧(对于这些帧它是有利的)的复合预测编码。

[0136] 实际的决定可以基于数据速率对音频质量的基本原理。作为质量测量,可以使用利用包括在编码器中的心理声学模型获得的数据(如通常可获得的基于 MDCT 音频编码器的情况下)。具体地,编码器的一些实施方式提供了预测系数的速率失真最佳化选择。从而,在这些实施方式中,如果预测增益的增加没有节省用于对残差信号进行编码的足够比特以证明花费的对预测系数进行编码所需的比特,则将预测系数的虚部并且有可能还有实部设置为 0。

[0137] 编码器的实施方式可以对与比特流中的 TNS 有关的信息进行编码。这种信息可以包括待由解码器侧的 TNS(综合)滤波器应用的 TNS 参数的值。如果 TNS 参数的相同集合要用于两个声道,则包括指示等通参数集合的信令比特比独立地传输两组参数节约。还可以包括是在上混合级前还是后应用 TNS 的信息(如适当地基于如两个可用选项的心理声学评估)。

[0138] 作为又一个可选特征(从复杂度和比特率的观点来看,该特征是潜在地有利的),编码器可以适用于使用对于残差信号的编码的各个有限的带宽。大于该极限的频带不会传输到解码器但是会设置为 0。在某些情况下,最高频带具有如此小的能量含量以至于它们已经被向量化为小到 0。普通的实践(参考 MPEG 标准中的参数 *max_sfb*)已经使用用于下混合信号和残差信号两者的相同的带宽极限来承担。现在,发明人已经经验地发现一定程度上大于下混合信号的残差信号具有集中到下频带的能量含量。因此,通过在残差信号上设置专用的上带宽极限,以不明显的质量的损失来减小比特率是可以的。例如,这可以通过传输两个独立的 *max_sfb* 参数来实现,一个是针对下混合信号,一个是针对残差信号。

[0139] 需要指出的是,尽管预测系数的最佳确定、量化及其编码的问题回退到 M/S 或 L/R 模式,但是仍然参考图 5 所示的解码器系统来讨论 TNS 滤波和上带宽限制,相同的事实等同地可应用于随后公开在参考后续图的后续内容中的实施方式。

[0140] 图 6 示出了根据本发明的适于进行复合预测立体声编码的另一种编码器系统。该系统作为输入接收分割成连续的、可能交叠的时间帧并且包括左声道和右声道的立体信号的时域表示。和与差级 601 将信号转换成中声道和侧声道。将中声道提供给 MDCT 模块 602 和 MDST 模块 603 两者,而将侧声道仅提供给 MDCT 模块 604。预测系数估计器 605 针对每个时间帧(可能地,针对帧内的各个频带)对如上面说明的复合预测系数 α 的值进行估计。系数 α 的值作为权重提供给加权加法器 606、607,其形成作为中信号的 MDCT 和 MDST 表示与侧信号的 MDCT 表示的线性组合的残差信号 D。优选地,将复合预测系数提供给加权加法器 606、607,加权加法器 606、607 由将在它被编码为比特流时使用的相同量化方案表示;由于编码器和解码器应用了相同的预测系数的值,所以这明显地提供了更加可靠的重构。将残差信号、中信号(当其与残差组合在一起出现时,可以更加恰当称为下混合信号)和预测系数提供给组合的量化和复用级 608,其将这些和可能的附加信息编码作为输出比特流。

[0141] 图 7 示出了图 6 的编码器系统的变形例。如从图中的符号的类似性很明显的,它具有类似的结构,但是还具有在直接 L/R 编码回退模式下的操作的附加功能。借助于紧设置在组合的量化与复用级 709 上游的开关 710 在复合预测编码模式与回退模式之间驱动编码器系统。在开关 710 的上位置,如图所示,开关 710 可以使得编码器操作在回退模式。将中/侧信号从紧位于 MDCT 模块 702、704 的下游的点提供到和与差级 705,和与差级 705 将该信号转换成左/右形式后将它传递到开关 710,开关 710 然后将其连接到组合的量化与复用器级 709。

[0142] 图 8 示出了根据本发明的编码器系统。相比图 6 和图 7 的编码器系统,此实施方式直接从 MDCT 数据(即,通过频域中的实到虚变换)导出了复合预测编码所需的 MDST 数据。实到虚变换应用联系图 2 和图 4 的解码器系统讨论的方法中的任何一种。将解码器的计算方法与编码器的计算方法进行匹配是重要的,以使得能够实现可靠的解码;优选地,同样的实到虚变换方法用在编码器侧和解码器侧。关于解码器的实施方式,由虚线包围的并且包括实到虚变换 804 的 A 部分能够由近似的变体或使用较少的输入时间帧作为输入来代替。同样地,可以使用上面描述的其他近似方法中的任意一种来简化编码。

[0143] 在较高水平上,图 8 的编码器系统具有与可能跟随有通过(适当地连接)实到虚模块来取代图 7 中的 MDST 模块的直接动作的结构不同的结构。本构造是清晰的并且实现了以鲁棒性和计算节约性的方式在预测编码与直接 L/R 编码之间进行切换的功能。将输入的立体声信号馈送给输出每个声道的频域表示的 MDCT 变换模块 801。将其馈送给最后的开关 808 以用于将编码器系统在其预测编码器模式与直接编码模式之间进行驱动并且馈送给和与差级 802。在直接 L/R 编码或联合 M/S 编码(在预测系数 α 被设置为 0 的时间帧中对其进行执行)中,此实施方式仅对输入信号进行 MDCT 变换、量化和复用,后两个步骤是通过布置在提供比特流的系统输出端的组合的量化与复用器级 807 来实现的。在预测编码中,声道中的每一个在和与差级 802 与开关 808 之间受到进一步的处理。实到虚变换 804 从中信号的 MDCT 表示导出 MDST 数据并且将这些 MDST 数据转发给预测系数估计器 803 和加权加法器 806 两者。与在图 6 和图 7 中示出的编码器系统中的相同,将另外的加权加法器 805

用于将具有加权 MDCT 的侧信号与中信号的 MDST 表示进行组合以形成残差声道信号,通过组合的量化和复用器模块 807 来将残差声道信号与中(即,下混合)声道信号和预测系数编码在一起。

[0144] 又参考图 9,将说明,编码器系统的实施方式中的每一个可以与一个或多个 TNS(分析)滤波器进行组合。根据前述讨论,将 TNS 滤波应用到下混合形式的信号通常是有利的。因此,如在图 9 中示出的,调节图 7 的编码器系统以包括 TNS 是通过紧在组合的量化与复用器模块 909 的上游增加 TNS 滤波器 911 来实现的。

[0145] 除了右/残差 TNS 滤波器 911b,可以紧在适于处理右声道或残差声道的开关 910 部分的上游设置两个独立的 TNS 滤波器(未示出)。因此,可以对两个 TNS 滤波器中的每一个始终提供相应的声道信号,从而使得能够基于不止仅当前帧的更多的时间帧来进行 TNS 滤波。如已经提到的,TNS 滤波器是并且仅是一个频域修改器件的示例,特别是将其处理基于不止当前帧的多个帧的器件,从这种布置可以获得与在 TNS 滤波器获得的益处一样多或更多的益处。

[0146] 作为图 9 所述的实施方式的另一种可能的替代方案,用于选择性动作的 TNS 滤波器能够布置在用于每个声道的多于一个点处。这与图 4 所示的能够借助开关连接不同 TNS 滤波器组的解码器系统的结构类似。这使得能够选择针对对于每个时间帧的 TNS 滤波的最合适的可用级。具体地,以下是有利的:联系复合预测立体声编码与其他编码模式之间的切换来在不同的 TNS 位置之间进行切换。

[0147] 图 11 示出了基于图 8 的编码器系统的变形例,其中,下混合信号的第二频域表示是借助于实到虚变换 1105 来导出的。类似于图 4 中的解码器系统,此编码器系统也包括可选择性地驱动的频域修改器模块,一个 1102 设置在下混合级的上游以及一个 1109 设置在下混合级的下游。能够使用 4 个开关 1103a、1103b、1109a 以及 1109b 来将在此图中通过 TNS 滤波器例示的频域模块 1102、1109 连接到信号路径中的每一个中。

[0148] III. 非装置实施方式

[0149] 在图 15 和图 16 中示出了本发明的第三方面与第四方面。图 15 示出了用于将比特流解码成立体声信号的方法,包括以下步骤:

[0150] 1. 输入比特流。

[0151] 2. 量化比特流,从而获得立体声信号的下混合声道和残差声道的第一频域表示。

[0152] 3. 计算下混合声道的第二频域表示。

[0153] 4. 基于声道的 3 个频域表示来计算侧声道信号。

[0154] 5. 基于侧声道和下混合声道来计算优选地是左/右形式的立体声信号。

[0155] 6. 输出所获得的立体声信号。

[0156] 可以认为步骤 3 至 5 是上混合处理。步骤 1 至 6 中的每一个与在本文中的前述部分公开的任意一种解码器系统中的对应功能类似,并且与其实现有关的进一步的细节能够在相同的部分中找到。

[0157] 图 16 示出了用于将立体声信号编码成比特流信号的方法,包括以下步骤:

[0158] 1. 输入立体声信号

[0159] 2. 将立体声信号变换成第一频域表示。

[0160] 3. 确定复合预测系数。

[0161] 4. 对频域表示进行下混合。

[0162] 5. 将下混合声道和残差声道与复合预测系数一起编码为比特流。

[0163] 6. 输出比特流。

[0164] 步骤 1 至 5 中的每一个步骤与在本文的前述部分中公开的任意一种编码器系统中的对应功能类似,并且与其实现的进一步的细节能够在相同的部分中找到。

[0165] 两种方法都可以表达为软件程序形式的计算机可读指令,并且可以通过计算机来执行。本发明的保护范围延伸到这种软件以及用于分发这种软件的计算机程序产品。

[0166] IV. 经验评估

[0167] 已经经验地对本文所公开的若干实施方式进行了评价。将在此子部分中对在此过程中所获得的经验材料的最重要的部分进行总结。

[0168] 用于实验的实施方式具有以下特征:

[0169] (i) 通过来自当前、前一个以及下一个 MDCT 频谱的二维有限脉冲响应滤波来计算每个 MDST 频谱(用于时间帧)。

[0170] (ii) 使用了来自 USAC 立体声编码器的心理声学模型。

[0171] (iii) 替代 PS 参数 ICC、CLD 以及 IPD,传输复合预测系数 α 的实部和虚部。实部和虚部被独立地处理,限于范围 $[-3.0, 3.0]$,以及使用步长 0.1 对其进行量化。然后对它们进行时间微分编码并且最终使用 USAC 的尺度因子码本对它们进行霍夫曼编码。每第二尺度因子带更新预测系数,其产生与 MPEG Surround(参见,例如 ISO/IEC23003-1)的频率分辨率类似的频率分辨率。此量化与编码方案导致针对具有 96kb/s 的目标比特率的典型配置中的立体声侧信息的大约 2kb/s 的平均比特率。

[0172] (iv) 由于 2 比特 ms_mask_present 比特流元素当前仅具有三个可能值,因此在不打断当前 USAC 比特流的情况下修改比特流格式。通过使用第四个值来指示复合预测使得在不浪费任何比特的情况下允许基本的中/侧编码的回退模式(对于该主题的进一步的细节,见本公开内容的前一子部分)。

[0173] 听力测试是根据 MUSHRA 方法学来完成的,具体地使得通过耳机的回放和具有采样速率为 48kHz 的 8 个测试项的使用成为必需的。在每个测试中有 3 个、5 个或 6 个测试对象参与。

[0174] 评估不同的 MDST 近似的影响以说明存在于这些选项之间的实际的复杂度与质量的权衡。在图 12 和图 13 中发现这些结果,前者示出了所获得的绝对分值并且后者示出了与 96s USAC cp1f 有关的差分分值,即,通过使用当前 MDCT 帧的复合预测进行 MDCT 域统一立体声编码以计算 MDST 的近似值。可以看出,当应用用于计算 MDST 频谱的更加计算地复杂的方法时,由基于 MDCT 的统一立体声编码实现的质量增益增加了。考虑到所有测试上的平均值,基于单个帧的系统 96s USAC cp1f 与常规的立体声编码相比提供了编码效率上的明显的增加。反过来,甚至获得了对于 96s USAC cp1f 的明显较好的结果,即,通过使用当前、前一个和下一个 MDCT 帧计算 MDST 的复合预测来进行 MDCT 域统一立体声编码。

[0175] V. 结束语

[0176] 在阅读了上述描述之后,对于本领域普通技术人员,本发明的另外实施方式将变得很明显。尽管本描述和附图公开了实施方式和示例,但是本发明不限于这些具体的示例。在不背离由所附权利要求定义的范围的前提下,可以做出多种修改和变化。

[0177] 应当注意,在经过包括例程实验能力的技术人员能力内的适当修改后,在本申请中所公开的方法及装置可以应用到对具有不止两个声道的信号进行的编码。尤其要强调,联系所述实施方式提到的任何信号、参数以及矩阵可以是频率变化的或频率不变的和 / 或时间变化的或时间不变的。所描述的计算步骤可以每次以频率方式执行或针对所有频带执行,并且可以被嵌入所有的实体以具有频率选择作用。出于应用的目的,可根据心理声学模型来调节任何量化方案。此外应当注意,各种和与差转换,即,从下混合 / 残差形式到伪 L/R 形式的转换以及 L/R 到 M/S 的转换和 M/S 到 L/R 的转换都具有以下形式:

[0178]
$$g \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

[0179] 其中,仅仅增益因子 g 可以变化。因此,通过单独地调节增益因子,可以通过适当选择解码增益来补偿一定的编码增益。此外,如本领域普通技术人员意识到的,偶数个串联布置的和与差变换具有直通级的效果,可能具有不统一的增益。

[0180] 上文中公开的系统和方法可以以软件、固件、硬件或其组合来实施。某些部件或所有部件可以以实现为由数字信号处理器或微处理器执行的软件,或实现为硬件或专用集成电路。这种软件可以分布在可以包括有存储媒介和通信媒介的计算可读媒介上。如本领域普通技术人员所公知的,计算机存储媒介包括以任何方法或技术实施的用于如计算机可读指令、数据结构、程序模块或其他数据等信息的存储的易失和非易失媒介、可移除及不可移除媒介。计算存储媒介包括但不限于 RAM、ROM、EEPROM、闪存或其他存储技术、CD-ROM、数字多功能光盘 (DVD) 或其他光盘存储、盒式磁带、磁带、磁盘存储或其他磁性存储器件或其他能够用于存储期望的信息并且能够由计算机访问的任何其他介质。此外,对于技术人员已知的是,通信介质通常具体化计算机可读指令、数据结构、程序模块或在调制数据信号如载波或其他传输机构中的其他数据,并且包括任何信息传输媒介。

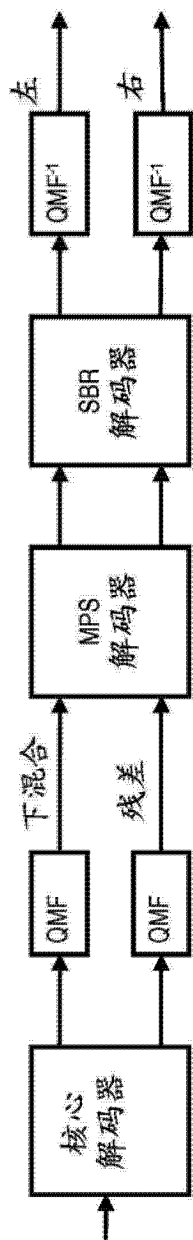


图 1A

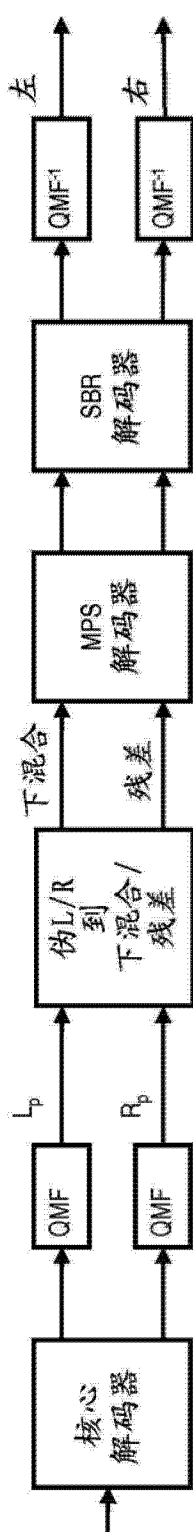


图 1B

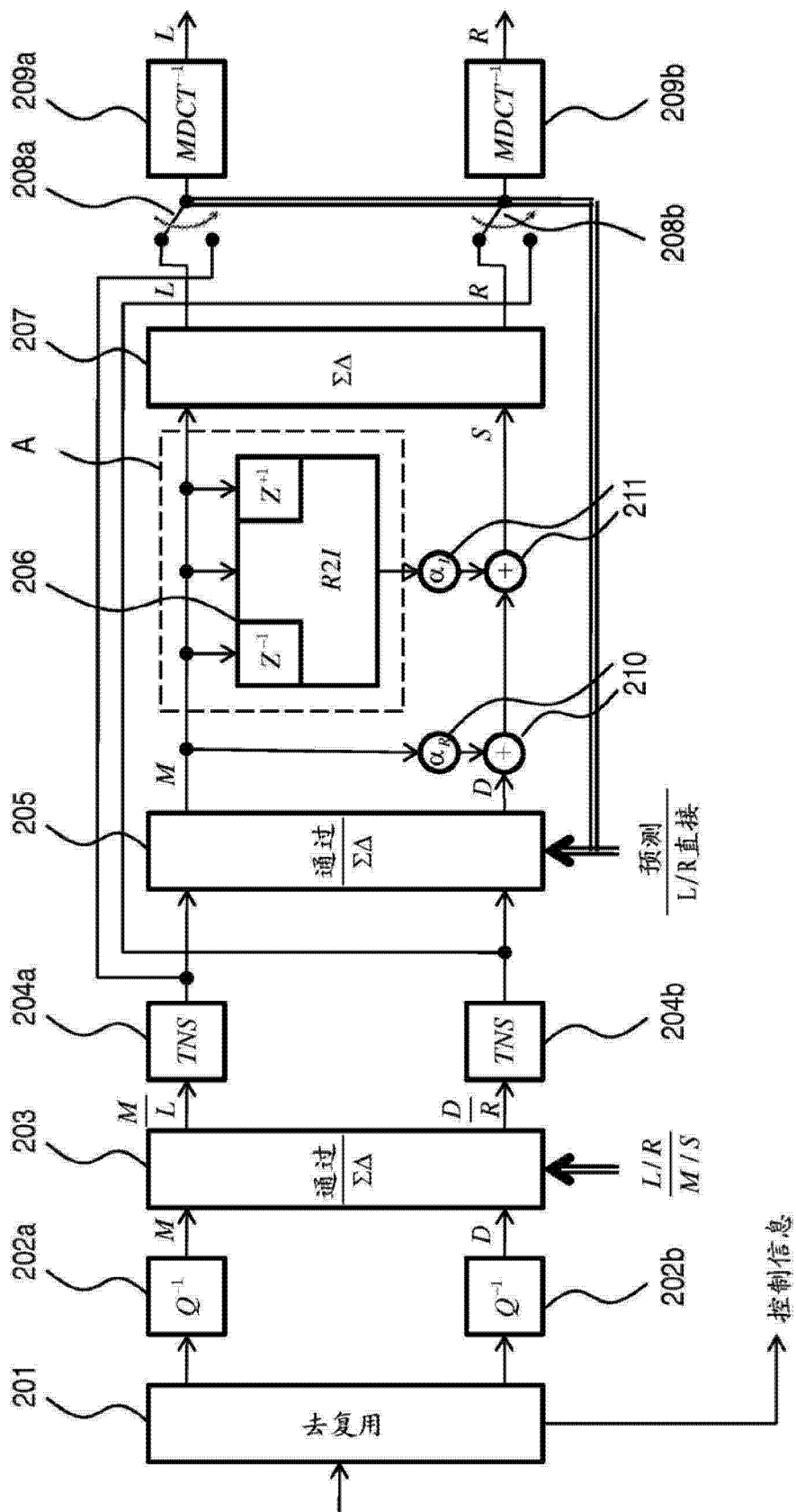


图 2

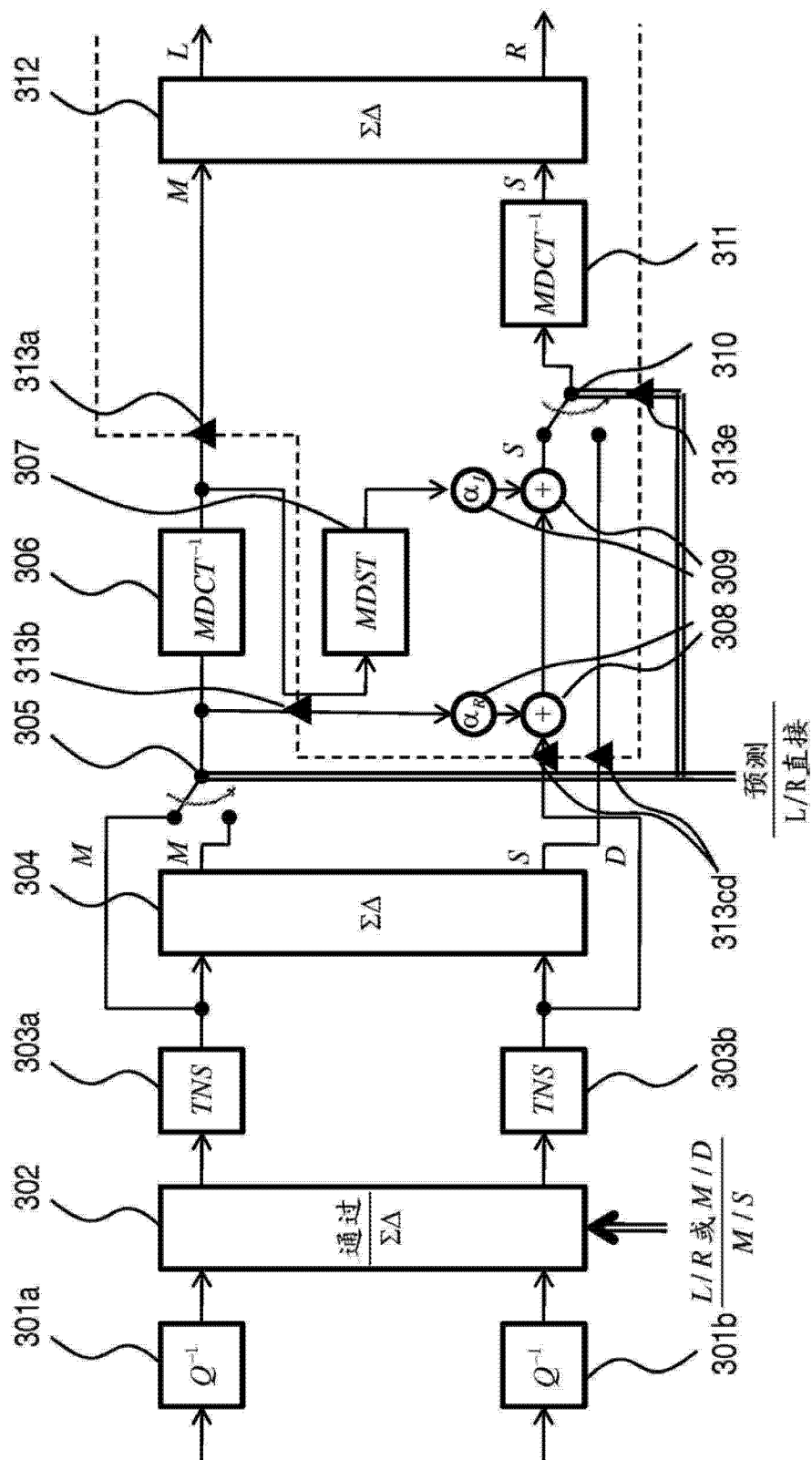


图 3

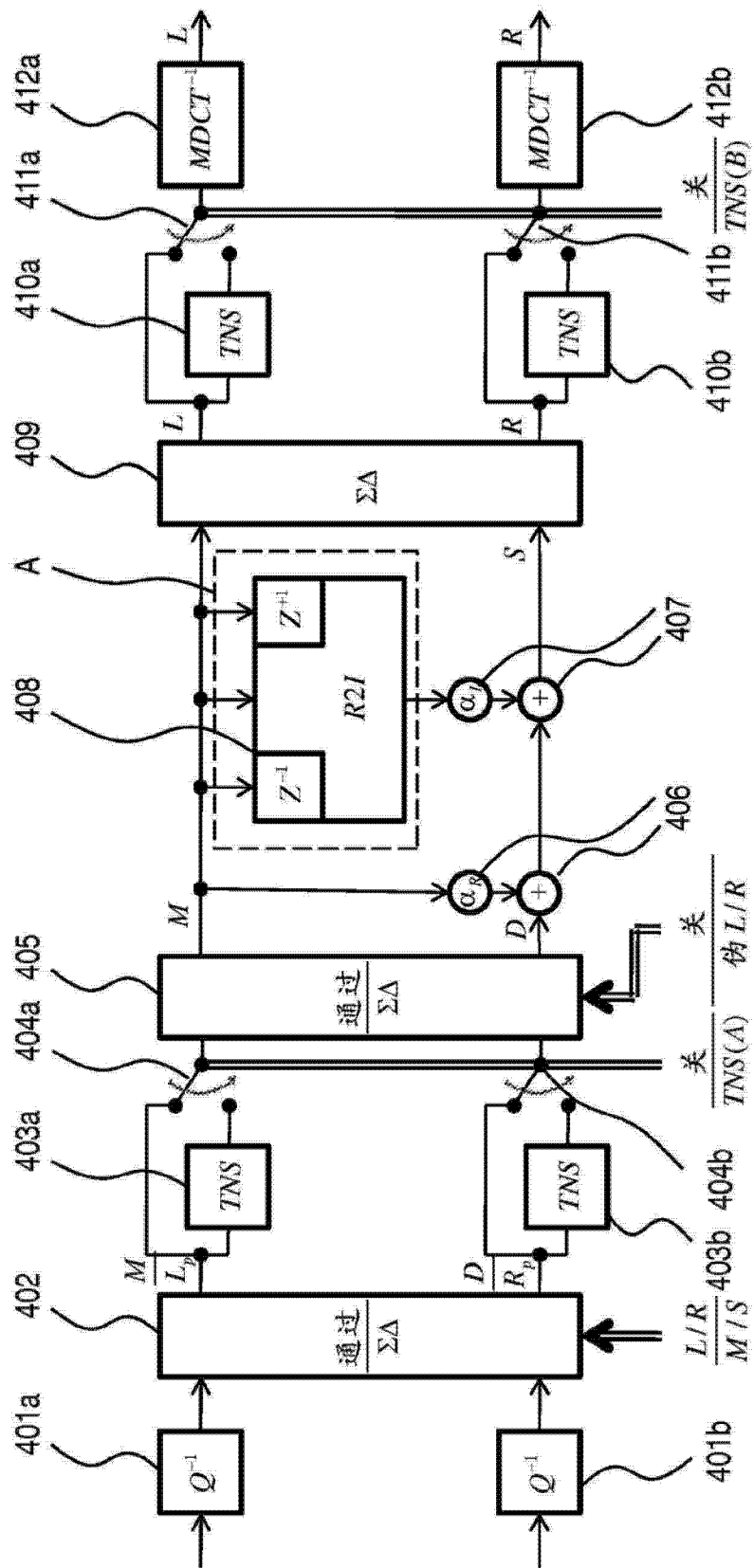


图 4

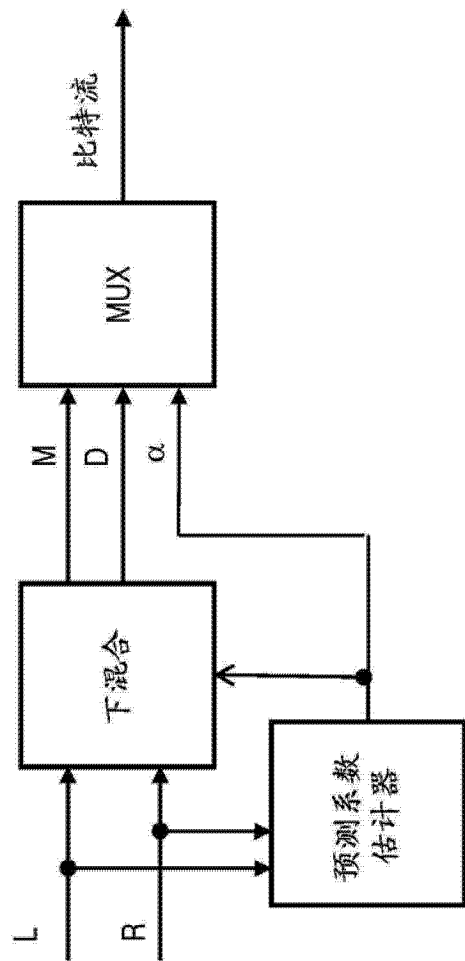


图 5

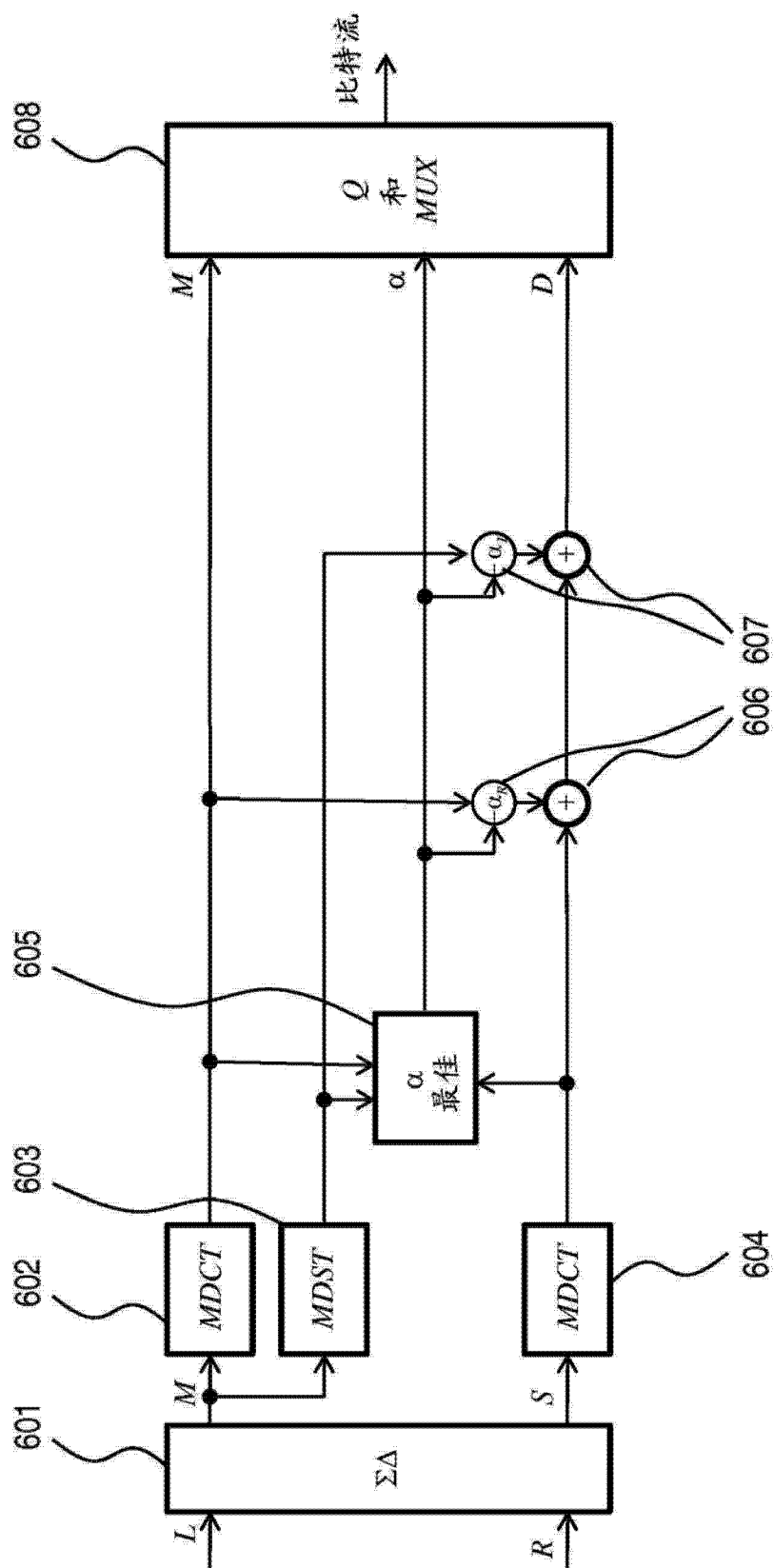


图 6

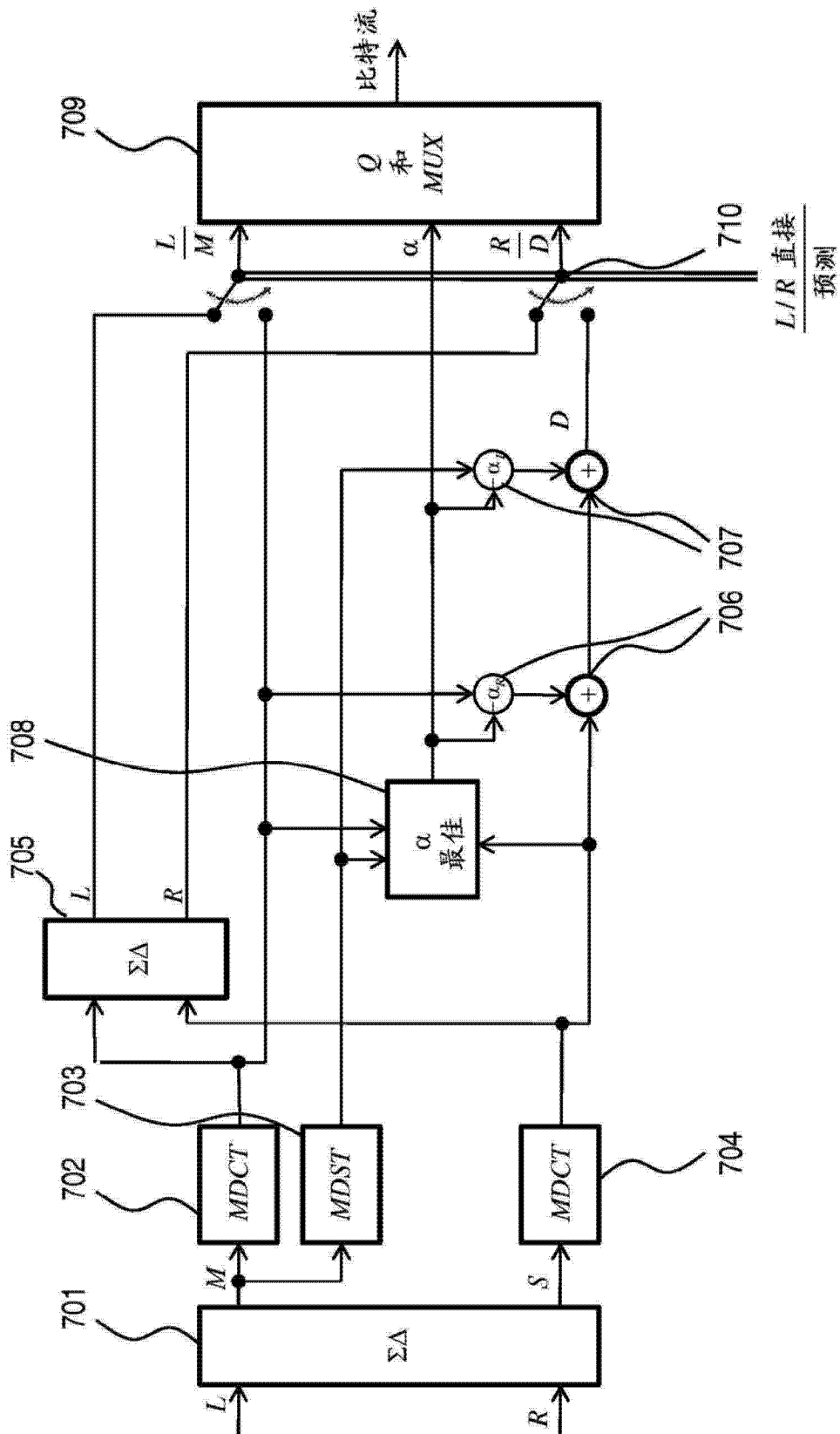


图 7

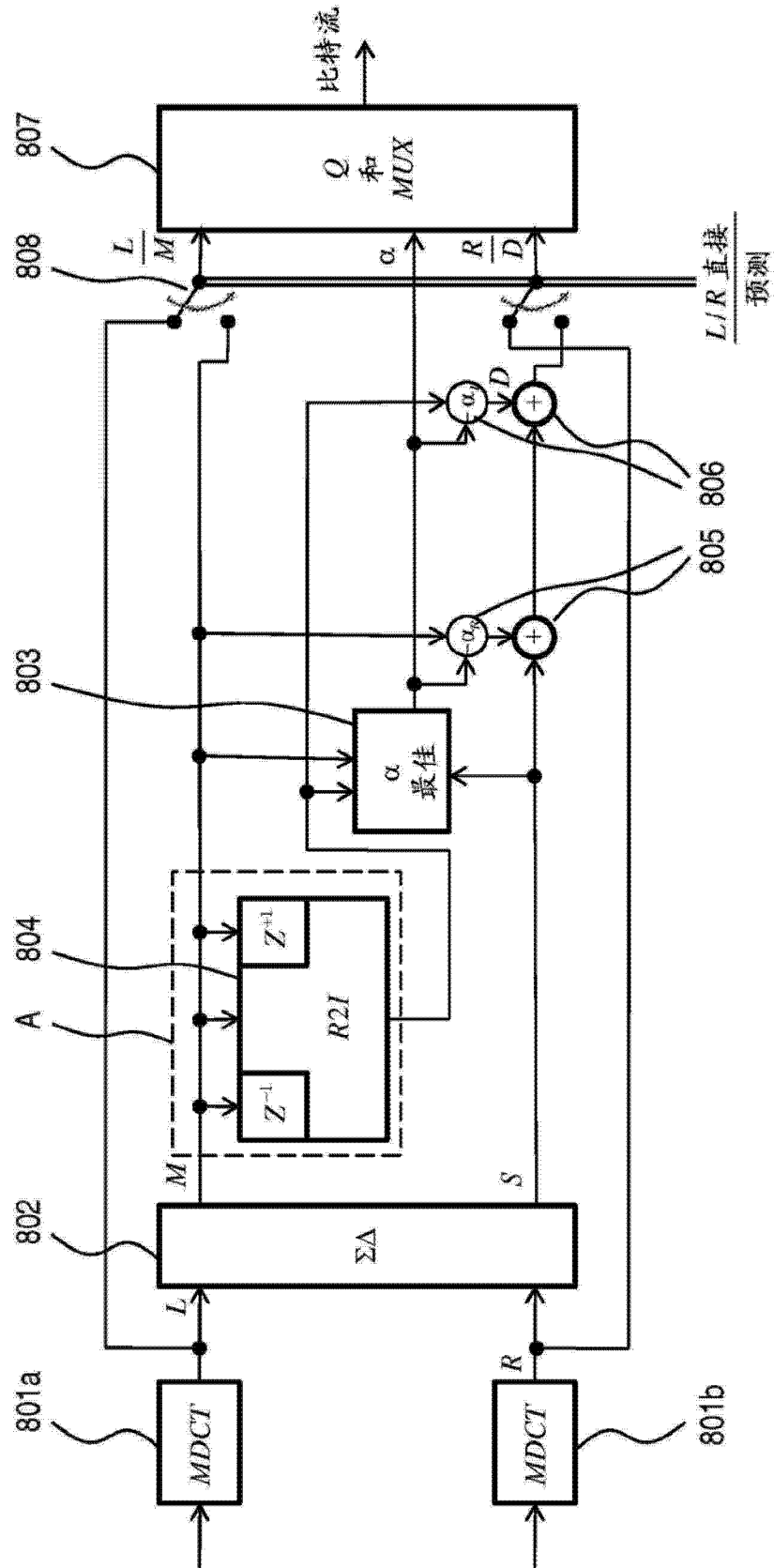


图 8

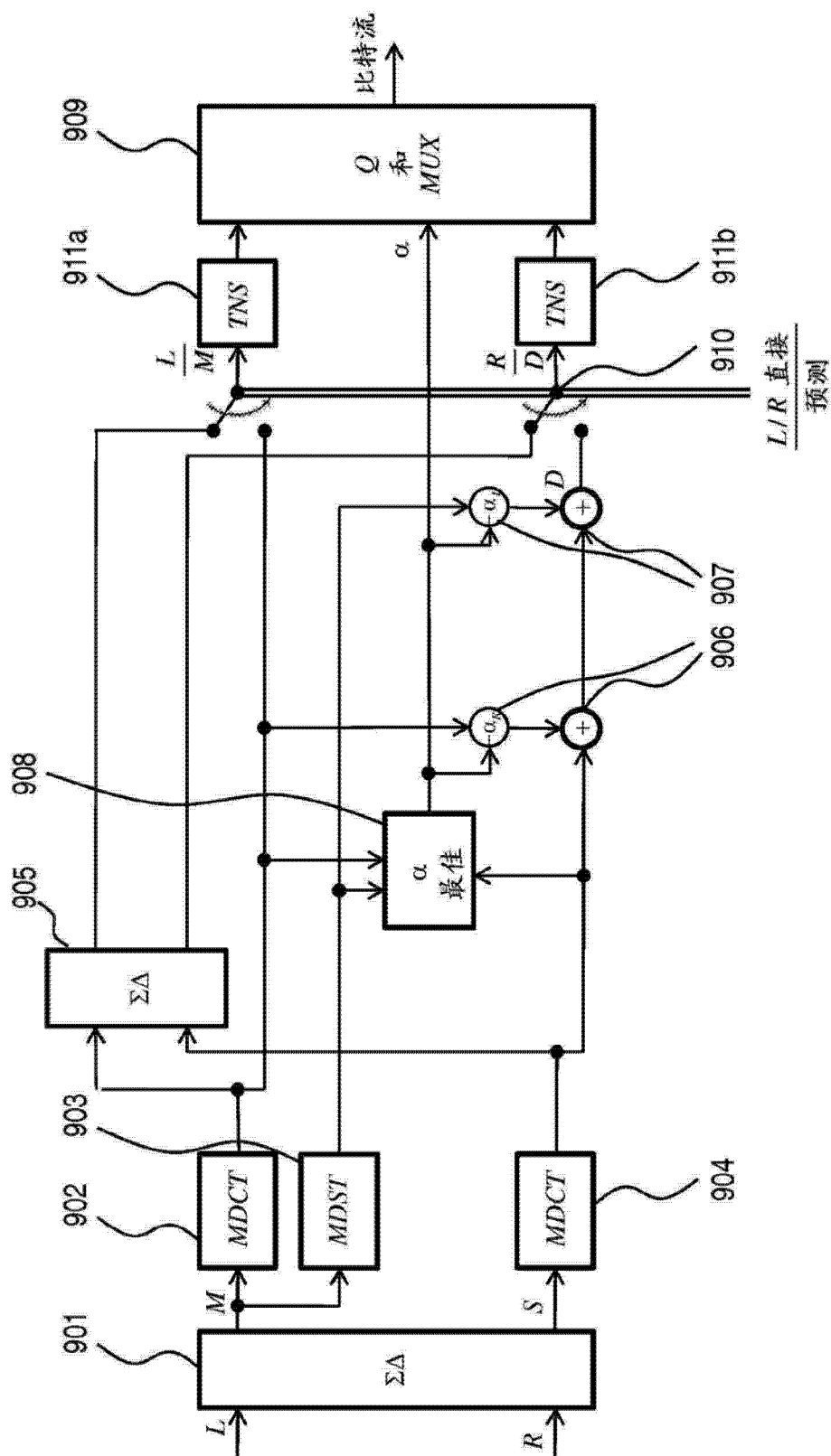


图 9

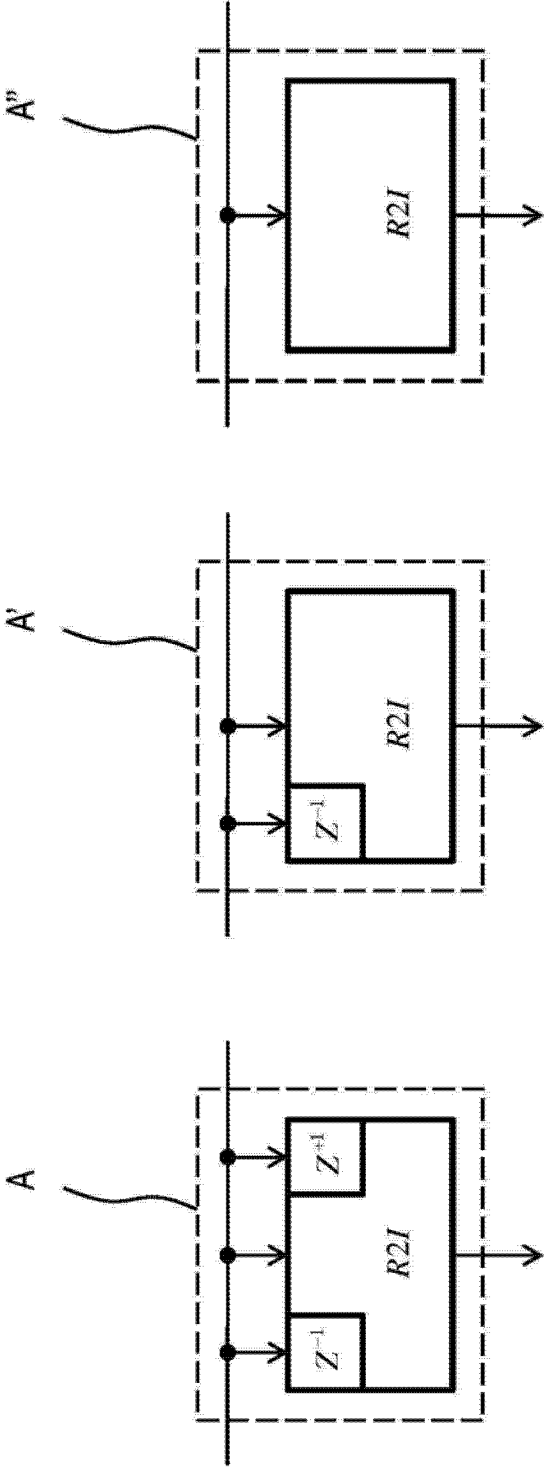


图 10

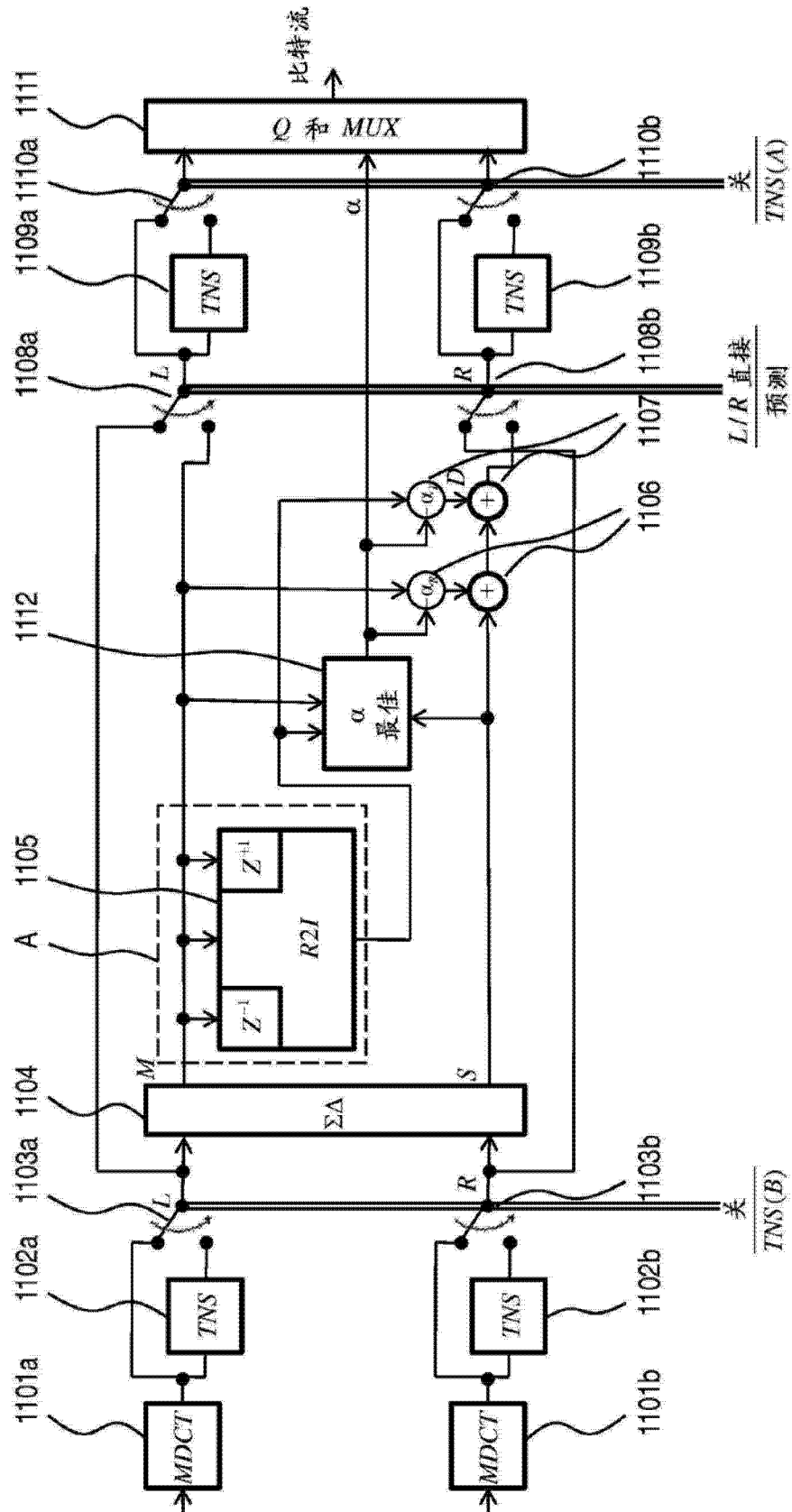


图 11

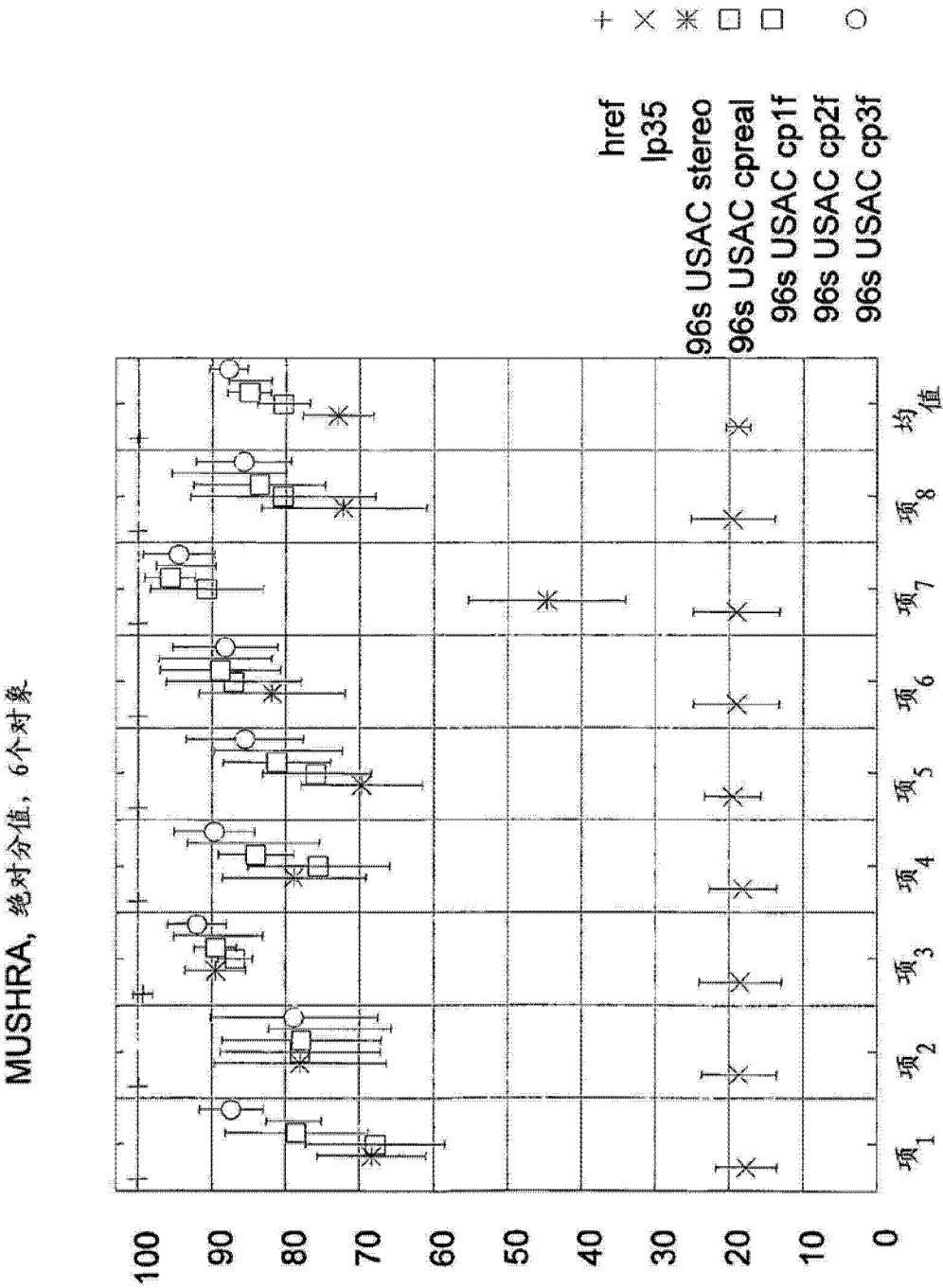


图 12

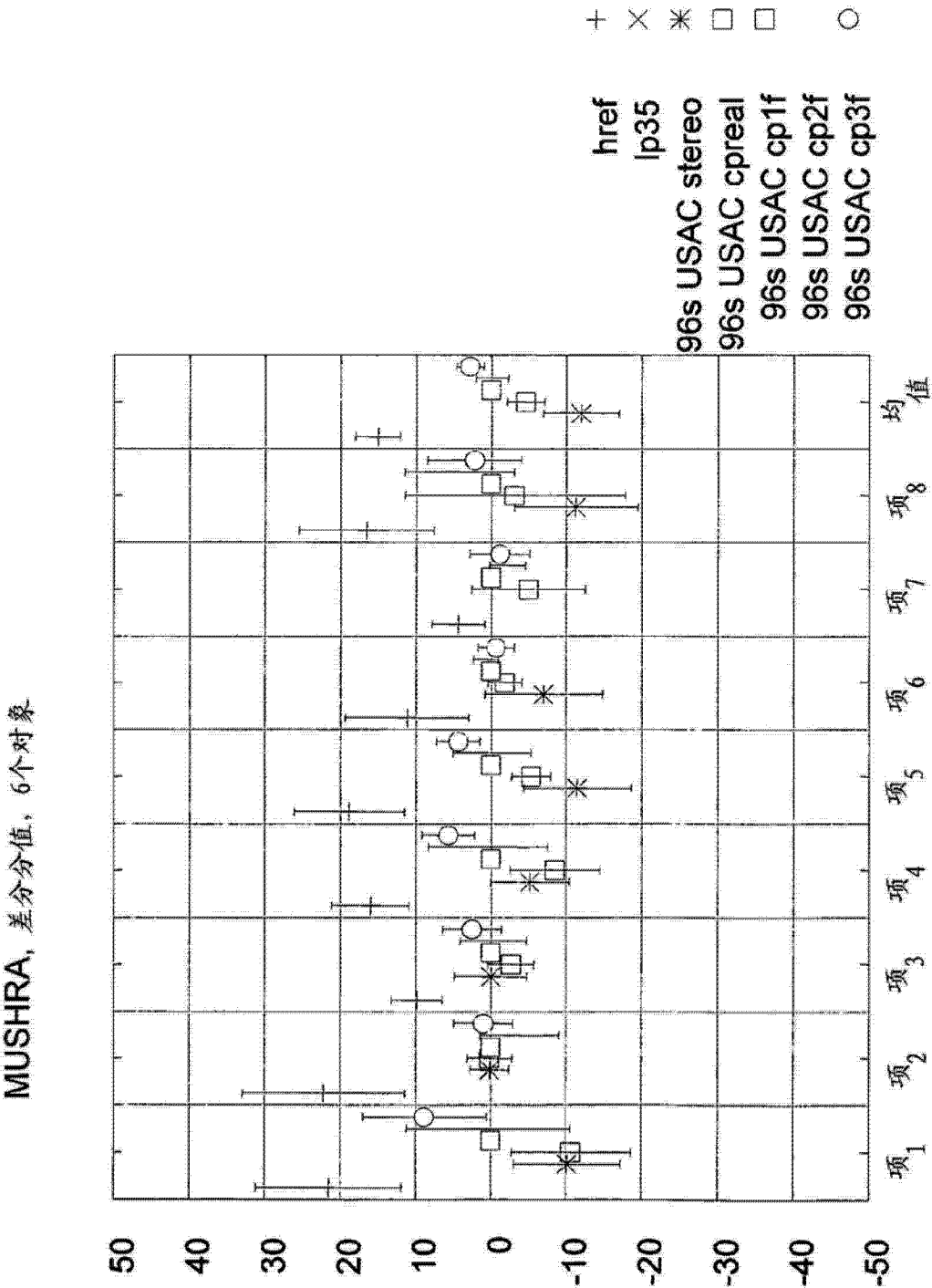


图 13

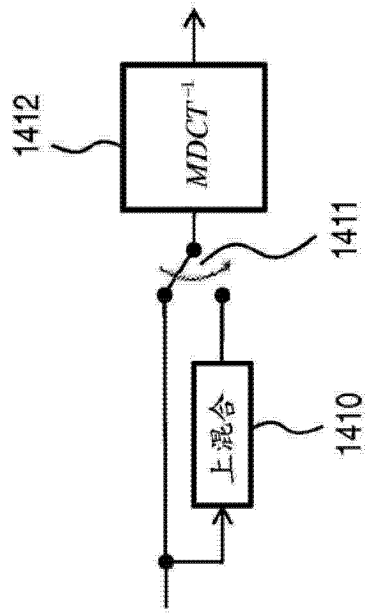


图 14A

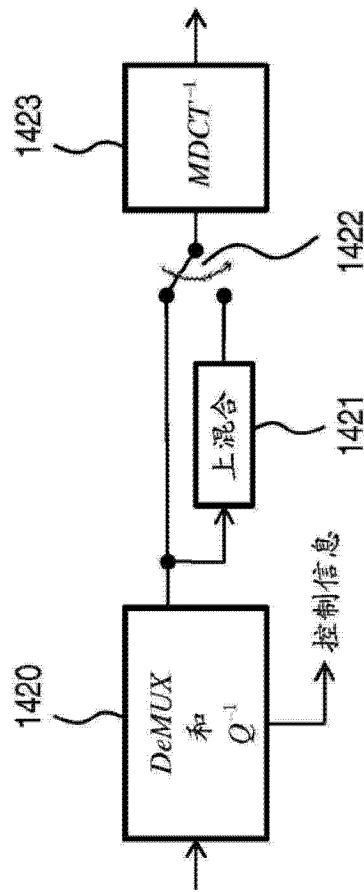


图 14B

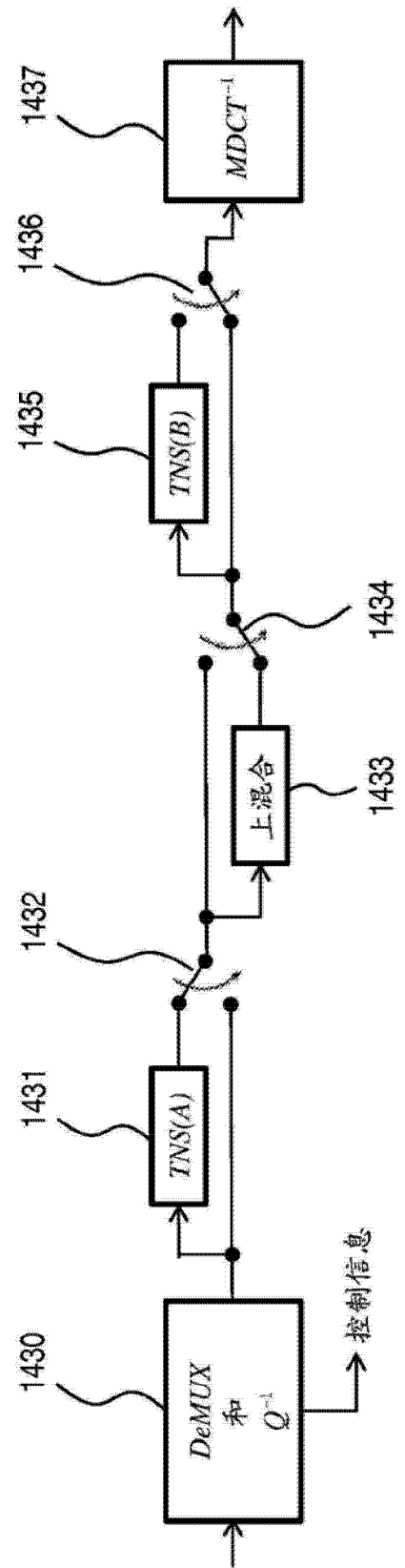


图 14C

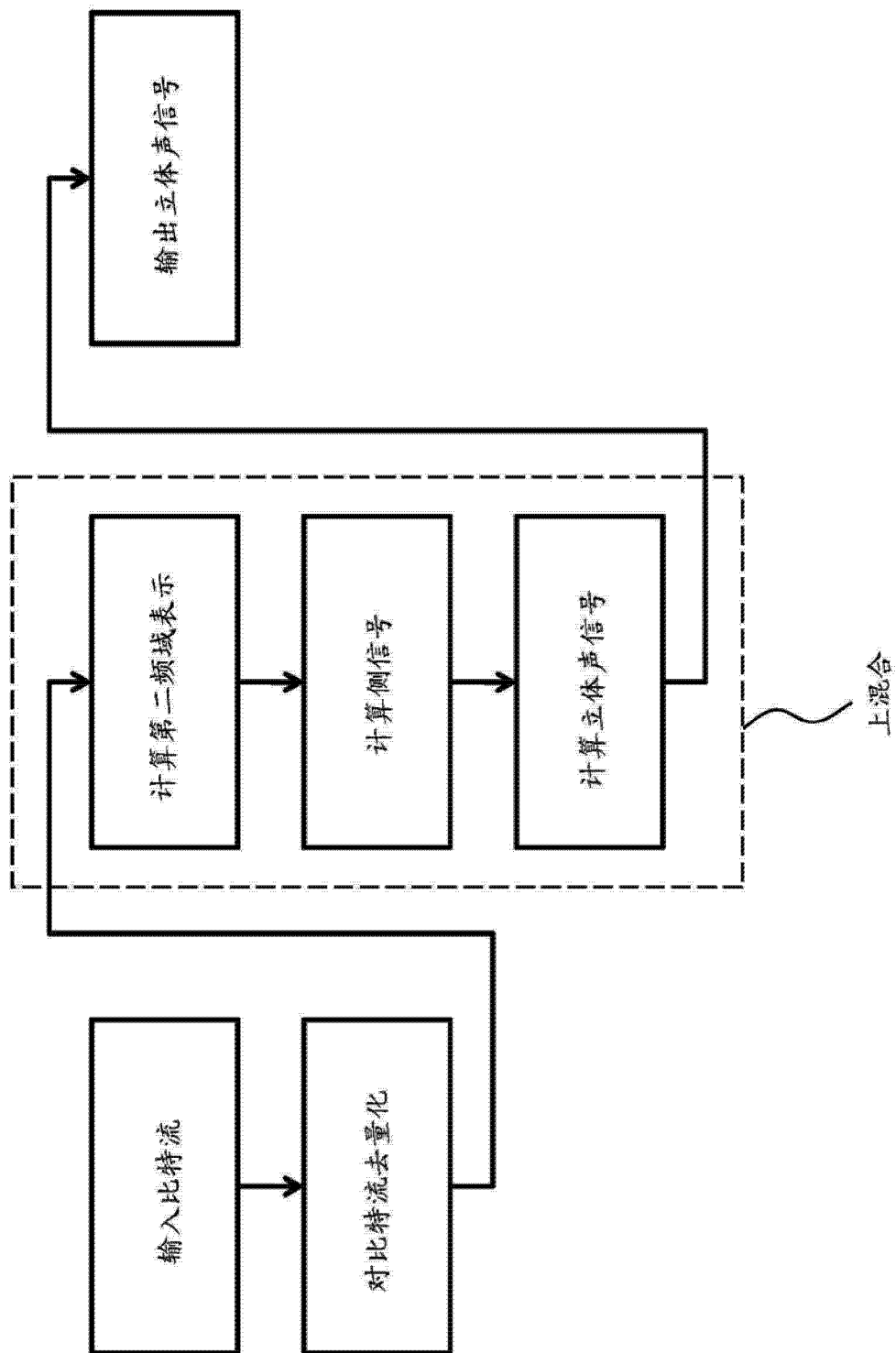


图 15

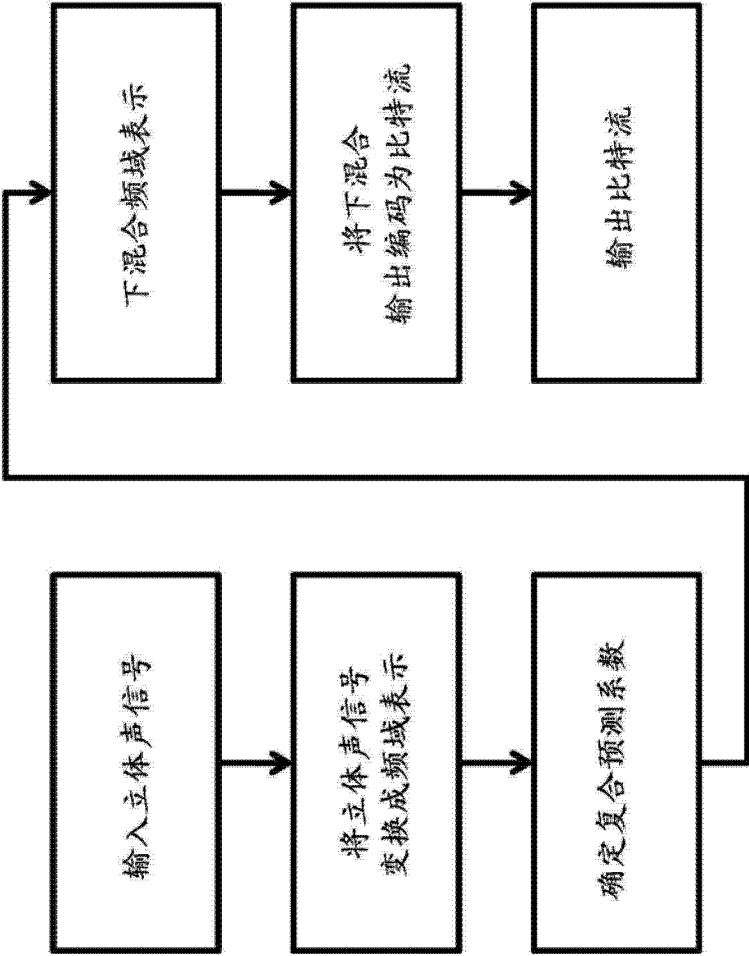


图 16