



(12) 发明专利申请

(10) 申请公布号 CN 114005454 A

(43) 申请公布日 2022. 02. 01

(21) 申请号 202111026302.2

(51) Int. Cl.

(22) 申请日 2016.06.17

G10L 19/008 (2013.01)

(30) 优先权数据

G10L 19/002 (2013.01)

62/181,096 2015.06.17 US

G10L 19/00 (2013.01)

62/241,082 2015.10.13 US

G10L 19/16 (2013.01)

62/241,098 2015.10.13 US

H04S 3/00 (2006.01)

62/245,191 2015.10.22 US

(62) 分案原申请数据

201680035415.X 2016.06.17

(71) 申请人 三星电子株式会社

地址 韩国京畿道

(72) 发明人 金善民 孙尚模

(74) 专利代理机构 北京市立方律师事务所

11330

代理人 谢玉斌 赵莎

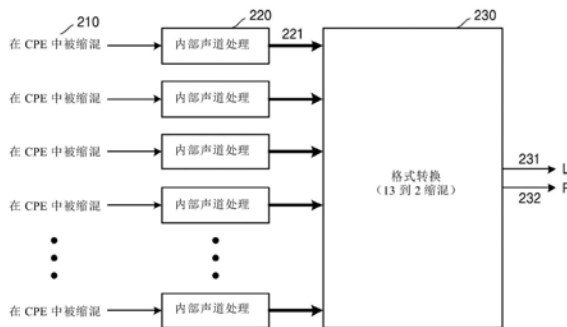
权利要求书2页 说明书30页 附图8页

(54) 发明名称

实现低复杂度格式转换的内部声道处理方法和装置

(57) 摘要

根据本发明实施例的处理音频信号的方法包括如下步骤:接收通过MPEG环绕212 (MPS212) 编码的音频比特流;基于接收到的音频比特流以及在格式转换器中定义的用于MPS212输出声道的渲染参数中的均衡(EQ)值和增益值,生成用于一个双声道单元(CPE)的内部声道信号;以及,基于所产生的内部声道信号生成立体声输出信号。



1. 一种对包括第一双声道单元 (CPE) 和第二CPE的四声道单元 (QCE) 进行解码的方法, 所述方法包括:

通过对所述第一CPE进行解码获得用于立体声输出布局的第一频带复制 (SBR) 有效载荷和第一MPS212有效载荷;

通过对所述第二CPE进行解码获得用于立体声输出布局的第二SBR有效载荷和第二MPS212有效载荷;

基于所述第一MPS212有效载荷、所述第二MPS212有效载荷和一对内部声道增益 (ICG) 生成一对带限内部声道 (IC) 信号;

使用格式转换器的渲染参数将所述第一SBR有效载荷和所述第二SBR有效载荷缩混为缩混SBR参数; 和

基于所生成的一对带限内部声道信号和所述缩混SBR参数生成一对全频带内部声道信号。

2. 如权利要求1所述的方法, 其中, 所述带限IC信号的生成包括判定对所述第一CPE和所述第二CPE的IC处理是否可能。

3. 如权利要求1所述的方法, 所述方法还包括基于所生成的一对全频带内部声道信号来生成立体声输出声道信号。

4. 如权利要求1所述的方法, 其中, 所述一对带限IC信号的生成包括:

基于所述格式转换器的所述渲染参数计算关于所述QCE的所述一对ICG。

5. 如权利要求1所述的方法, 其中, 所述一对ICG的内部声道增益 $G_{ICH}^{l,m}$ 是通过使用

$$G_{ICH}^{l,m} = \sqrt{(c_{left}^{l,m} \times G_{left} \times G_{EQ,left}^m)^2 + (c_{right}^{l,m} \times G_{right} \times G_{EQ,right}^m)^2}$$

来计算的, 其中l表示时隙索引, m表示频带索引,

$c_{left}^{l,m}$ 和 $c_{right}^{l,m}$ 是所述第一CPE的声道电平差 (CLD) 值,

G_{left} 和 G_{right} 是根据MPS212输出声道在所述格式转换器中定义的增益值, 并且

$G_{EQ,left}^m$ 和 $G_{EQ,right}^m$ 是根据所述MPS212输出声道在所述格式转换器中定义的均衡 (EQ) 增益值。

6. 一种用于处理音频信号的装置, 所述装置包括:

接收器, 所述接收器被配置为接收包括第一双声道单元 (CPE) 和第二CPE的四声道单元 (QCE); 和

内部声道 (IC) 信号发生器, 所述IC信号发生器被配置为:

通过对所述第一CPE进行解码获得用于立体声输出布局的第一频带复制 (SBR) 有效载荷和第一MPS212有效载荷;

通过对所述第二CPE进行解码获得用于立体声输出布局的第二SBR有效载荷和第二MPS212有效载荷;

基于所述第一MPS212有效载荷、所述第二MPS212有效载荷和一对内部声道增益 (ICG) 生成一对带限内部声道 (IC) 信号;

使用格式转换器的渲染参数将所述第一SBR有效载荷和所述第二SBR有效载荷缩混为

缩混SBR参数;和

基于所生成的一对带限内部声道信号和所述缩混SBR参数生成一对全频带内部声道信号。

7. 如权利要求6所述的装置,其中,所述IC信号发生器还被配置为判定对所述第一CPE和所述第二CPE的IC处理是否可能。

8. 如权利要求6所述的装置,所述装置还包括立体声输出信号发生器,所述立体声输出信号发生器被配置为基于所生成的一对全频带内部声道信号生成立体声输出声道信号。

9. 如权利要求6所述的装置,其中,所述IC信号发生器还被配置为基于所述格式转换器的所述渲染参数计算关于所述QCE的所述一对ICG。

10. 如权利要求6所述的装置,其中,所述一对ICG的内部声道增益 $G_{ICH}^{l,m}$ 是通过使用

$$G_{ICH}^{l,m} = \sqrt{(c_{left}^{l,m} \times G_{left} \times G_{EQ,left}^m)^2 + (c_{right}^{l,m} \times G_{right} \times G_{EQ,right}^m)^2}$$

来计算的,其中1表示时隙索引,m表示频带索引,

$c_{left}^{l,m}$ 和 $c_{right}^{l,m}$ 是所述第一CPE的声道电平差(CLD)值,

G_{left} 和 G_{right} 是根据MPS212输出声道在所述格式转换器中定义的增益值,并且

$G_{EQ,left}^m$ 和 $G_{EQ,right}^m$ 是根据所述MPS212输出声道在所述格式转换器中定义的均衡(EQ)增益值。

实现低复杂度格式转换的内部声道处理方法和装置

[0001] 本申请是申请日为2016年6月17日、申请号为201680035415X、发明名称为“实现低复杂度格式转换的内部声道处理方法和装置”的中国专利申请的分案申请。

技术领域

[0002] 本发明涉及实现低复杂度格式转换的内部声道(IC)处理方法和装置,更具体地说,涉及一种通过在立体声输出布局环境中对输入声道执行IC处理以减少格式转换器中的IC数量进而减少在格式转换器中进行的协方差运算次数的方法和装置。

背景技术

[0003] 根据MPEG-H 3D音频,可以处理各种类型的信号并且可以容易地控制输入/输出的类型。因此,MPEG-H 3D音频可以用作下一代音频信号处理的解决方案。此外,根据装置小型化的趋势,在立体声再现环境中借助于移动设备进行音频再现的百分比已经增加。

[0004] 当通过诸如22.2声道的多声道实现的沉浸式音频信号被传送到立体声再现系统时,将要解码所有输入声道,并且沉浸式音频信号要被缩混(downmix)以转换成立体声格式。

[0005] 随着输入声道数量的增加和输出声道数量的减少,在上述过程中,协方差分析和相对对齐所需的解码器的复杂度增加。这种复杂度的增加不仅影响移动设备的运行速度,还影响移动设备的电池消耗。

发明内容

[0006] 技术问题

[0007] 如上所述,增加输入声道的数量以提供沉浸式音频,然而减少输出声道的数量以实现便携性。在这种环境下,解码过程中格式转换的复杂度成为问题。

[0008] 为了解决这个问题,本发明提供了解码器中格式转换的复杂度的降低。

[0009] 技术方案

[0010] 为了实现上述目标,本发明的代表性特征如下。

[0011] 根据本发明的一个方面,提供了一种处理音频信号的方法,所述方法包括:接收通过MPEG环绕212(MPS212)编码的音频比特流;基于接收到的音频比特流、格式转换器中定义的MPS212输出声道的均衡(EQ)值和MPS212输出声道的增益值,生成用于单个双声道单元(CPE)的内部声道(IC)信号;以及基于生成的IC信号生成立体声输出信号。

[0012] 所述IC信号的生成可以包括:基于包含在MPS212有效载荷中的声道电平差(CLD),将所述接收到的音频比特流上混(upmix)为用于包含在所述单个CPE中的声道对的信号;基于所述EQ值和所述增益值,缩放上混后的比特流;以及混合缩放后的比特流。

[0013] 所述IC信号的生成可以进一步包括判定是否生成了所述用于单个CPE的IC信号。

[0014] 可以基于包含在单个CPE中的声道对是否属于同一个IC组,来判定是否生成所述用于单个CPE的IC信号。

[0015] 当包含在单个CPE中的声道对均包含在左IC组中时,所述IC信号可以只通过立体声输出声道中的左输出声道输出。当包含在单个CPE中的声道对均包含在右IC组中时,所述IC信号可以只通过立体声输出声道中的右输出声道输出。

[0016] 当包含在单个CPE中的声道对均包含在中央IC组中或者包含在单个CPE中的声道对均包含在低频音效(LFE) IC组中时,所述IC信号可以通过立体声输出声道中的左输出声道和右输出声道均匀地输出。

[0017] 所述音频信号可以是沉浸式音频信号。

[0018] 所述IC信号的生成还可以包括:计算IC增益(ICG);以及应用所述ICG。

[0019] 根据本发明的另一方面,提供了一种用于处理音频信号的装置,所述装置包括:接收器,被配置为接收通过MPEG环绕212(MPS212)编码的音频比特流;内部声道(IC)信号发生器,被配置为基于接收到的音频比特流、格式转换器中定义的MPS212输出声道的均衡(EQ)值和MPS212输出声道的增益值,生成用于单个双声道单元(CPE)的IC信号;以及立体声输出信号发生器,被配置为基于生成的IC信号生成立体声输出信号。

[0020] 所述IC信号发生器可以被配置为:基于包含在MPS212有效载荷中的声道电平差(CLD),将所述接收到的音频比特流上混为用于包含在所述单个CPE中的声道对的信号;基于所述EQ值和所述增益值,缩放上混后的比特流;以及混合缩放后的比特流。

[0021] 所述IC信号发生器可以被配置为判定是否生成了所述的用于单个CPE的IC信号。

[0022] 可以基于包含在单个CPE中的声道对是否属于同一个IC组,来判定是否生成所述IC信号。

[0023] 当包含在单个CPE中的声道对均包含在左IC组中时,所述IC信号可以只通过立体声输出声道中的左输出声道输出。当包含在单个CPE中的声道对均包含在右IC组中时,所述IC信号可以只通过立体声输出声道中的右输出声道输出。

[0024] 当包含在单个CPE中的声道对均包含在中央IC组中或者包含在单个CPE中的声道对均包含在低频音效(LFE) IC组中时,所述IC信号可以通过立体声输出声道中的左输出声道和右输出声道均匀地输出。

[0025] 所述音频信号可以是沉浸式音频信号。

[0026] 所述IC信号发生器可以被配置为计算IC增益(ICG)和应用所述ICG。

[0027] 根据本发明的另一个方面,提供了一种计算机可读记录介质,所述计算机可读记录介质上记录有用于执行前述的方法的计算机程序。

[0028] 根据本发明的其他实施例,提供了其它方法、其它系统和上面记录有用于执行所述方法的计算机程序的计算机可读记录介质。

[0029] 有益效果

[0030] 根据本发明,通过使用内部声道(IC)减少了输入到格式转换器的声道的数量,因此可以降低格式转换器的复杂度。更详细地说,由于减少了输入到格式转换器的声道数量,所以简化了要在格式转换器中执行的协方差分析,因此,降低了格式转换器的复杂度。

附图说明

[0031] 图1是根据一个实施例的用于将24个输入声道格式转换为立体声输出声道的解码结构的框图。

[0032] 图2是根据一个实施例的解码结构的框图,该解码结构利用13个内部声道(IC)将22.2声道沉浸式音频信号格式转换为立体声输出声道。

[0033] 图3图示了从单个双声道单元(CPE)生成单个IC的实施例。

[0034] 图4是根据本发明实施例的解码器的IC增益(ICG)应用单元将ICG应用于IC信号的详细框图。

[0035] 图5是示出了根据本发明实施例的编码器预处理ICG时的解码的框图。

[0036] 图6的流程图示出了根据本发明实施例的当通过立体声再现布局输出CPE时执行单声道频带复制(SBR)解码然后执行MPEG环绕(MPS)解码的结构中的IC处理方法。

[0037] 图7的流程图示出了根据本发明实施例的当通过立体声再现布局输出CPE时执行MPS解码并随后执行立体声SBR解码的结构中的IC处理方法。

[0038] 图8是根据本发明实施例的当通过立体声再现布局输出四声道单元(QCE)时使用立体声SBR的结构中的IC处理方法的框图。

[0039] 图9是根据本发明另一实施例的在通过立体声再现布局输出QCE时使用立体声SBR的结构中的IC处理方法的框图。

[0040] 图10a示出了当第一包络的开始边界相同并且最后包络的停止边界相同时确定时间包络格子的实施例。

[0041] 图10b示出了当第一包络的开始边界不同并且最后包络的停止边界相同时确定时间包络格子的实施例。

[0042] 图10c示出了当第一包络的开始边界相同并且最后包络的停止边界不同时确定时间包络格子的实施例。

[0043] 图10d示出了当第一包络的开始边界不同并且最后包络的停止边界不同时确定时间包络格子的实施例。

[0044] 表1示出了将22.2声道沉浸式音频信号渲染为立体声信号的格式转换器的混合矩阵的实施例。

[0045] 表2示出了使用IC将22.2声道沉浸式音频信号渲染为立体声信号的格式转换器的混合矩阵的实施例。

[0046] 表3示出了根据本发明实施例的使用IC来配置22.2声道的CPE结构。

[0047] 表4示出了根据本发明实施例的与解码器输入声道对应的IC的类型。

[0048] 表5示出根据本发明实施例的根据IC类型另外定义的声道的位置。

[0049] 表6示出根据本发明的实施例的与IC类型对应的格式转换器输出声道,以及将被应用于每个格式转换器输出声道的增益和EQ索引。

[0050] 表7示出了根据本发明实施例的ICGConfig的句法。

[0051] 表8示出了根据本发明实施例的mpegh3daExtElementConfig()的句法。

[0052] 表9示出了根据本发明实施例的usacExtElementType的句法。

[0053] 表10示出了根据本发明实施例的speakerLayoutType的句法。

[0054] 表11示出了根据本发明实施例的SpeakerConfig3d()的句法。

[0055] 表12示出了根据本发明实施例的immersiveDownmixFlag的句法。

[0056] 表13示出了根据本发明实施例的SAOC3DgetNumChannels()的句法。

[0057] 表14示出了根据本发明实施例的声道分配顺序的句法。

[0058] 表15示出了根据本发明实施例的mpegh3daChannelPairElementConfig() 的句法。

[0059] 表16示出了根据本发明实施例的基于声道单元和再现布局确定的MPS 和SBR的解码场景。

[0060] 最佳实施方式

[0061] 为了实现上述目标,本发明的代表性特征如下。

[0062] 处理音频信号的方法包括:接收通过MPEG环绕212 (MPS212) 编码的音频比特流;基于接收到的音频比特流、格式转换器中定义的MPS212输出声道的均衡(EQ)值以及MPS212输出声道的增益值,生成用于单个双声道单元(CPE)的内部声道(IC)信号;并基于生成的IC信号生成立体声输出信号。

具体实施方式

[0063] 现在将参考示出本发明的特定实施例的附图对本发明进行详细描述。提供这些实施例是为了使本公开内容透彻和完整,并且将本发明的构思充分地传达给本领域的普通技术人员。应该理解的是,本发明的各种实施例彼此不同,但并不彼此排斥。

[0064] 例如,在不脱离本发明的精神和范围的情况下,说明书中描述的特定形状、特定结构和特定特征在一个实施例中可以与另一个实施例中不同。还可以理解的是,在不脱离本发明的精神和范围的情况下,每个实施例中的每个元件的位置或布局可以改变。因此,下面的详细描述应该被认为只是描述性的,而不是为了限制的目的,并且本发明的范围应当由所附权利要求及其等同物定义。

[0065] 在整个说明书中,附图中相似的附图标记表示相同或相似的元件。在附图中,为了简化说明,省略了与说明无关的部分,并且相同的标号始终表示相同的元件。

[0066] 在下文中,将通过参考附图解释本发明的示例性实施例来详细描述本发明。然而,本发明可以以许多不同的形式来实施并且不应该被解释为限于在此阐述的实施例。

[0067] 在整个说明书中,当元件被称为“连接”或“耦合”到另一元件时,其可以直接连接或耦合到另一元件,或者可以利用插入其间的插入元件而电连接或电耦合到另一元件。另外,当在本说明书中使用术语“包括”或者“包含”指明所述元件的存在,但不排除存在或添加一种或多种其他元件。

[0068] 这里使用的术语定义如下。

[0069] 内部声道(IC)是在格式转换中使用的虚拟中间声道,并且考虑了立体声输出以便去除在MPS212 (MPEG环绕立体声)上混期间和格式转换器(FC)缩混期间产生的不必要的运算,将进来。

[0070] IC信号是为了提供立体声信号而在格式转换器中混合的单声道信号,并且是利用IC增益(ICG)生成的。

[0071] IC处理表示利用MPS212解码块生成IC信号并且在IC处理块中执行的处 理。

[0072] ICG表示根据声道电平差(CLD)值和格式转换参数计算出的并应用于IC信号的增益。

[0073] IC组表示基于核心编解码器输出声道位置确定的IC的类型,核心编解码器输出声道位置和IC组在稍后将描述的表4中定义。

[0074] 现在将参照示出了本发明的示例性实施例的附图更充分地描述本发明。

[0075] 图1是根据一个实施例的用于将24个输入声道格式转换成立体声输出声道的解码结构的框图。

[0076] 当多声道输入的比特流被传送到解码器时,解码器根据再现系统的输出声道布局来缩混输入声道布局。例如,当遵循MPEG标准的22.2声道输入信号由如图1所示的立体声声道输出系统再现时,包含在解码器中的格式转换器130根据格式转换器130内规定的格式转换器规则,将24输入声道布局缩混成2输出声道布局。

[0077] 输入到解码器的22.2声道输入信号包含双声道单元(CPE)比特流110,该CPE比特流110是通过针对包含在单个CPE中的两个声道来缩混信号而获得的。由于CPE比特流是通过MPS212(基于MPEG的立体声)编码的,所以通过MPS212 120解码CPE比特流。在这种情况下,CPE比特流中不包含LFE声道,即低音扬声器声道。因此,输入到解码器的22.2声道输入信号包含用于11个CPE的比特流和用于两个低音扬声器声道的比特流。

[0078] 当对构成22.2声道输入信号的CPE比特流执行MPS212解码时,为每个CPE生成两个MPS212输出声道,即121和122,并且这两个MPS212输出声道121和122成为格式转换器130的输入声道。在如图1的情况中,包括两个低音扬声器声道的格式转换器130的输入声道的数量 N_{in} 是24。因此,格式转换器130应当执行 $24*2$ 缩混。

[0079] 格式转换器130根据协方差分析执行相位对齐,以防止由于多声道信号的相位之间的差异而导致发生音色失真。在这种情况下,因为协方差矩阵具有 $N_{in} \times N_{in}$ 维度,理论上应当执行 $(N_{in} \times (N_{in} - 1) / 2 + N_{in}) \times 71$ 个频带 $\times 2 \times 16 \times (48000 / 2048)$ 次复数乘法以分析协方差矩阵。

[0080] 当输入声道的数量 N_{in} 是24时,对于一次复数乘法应执行四次运算,并且需要大约每秒六千四百万次运算(MOPS)的性能。

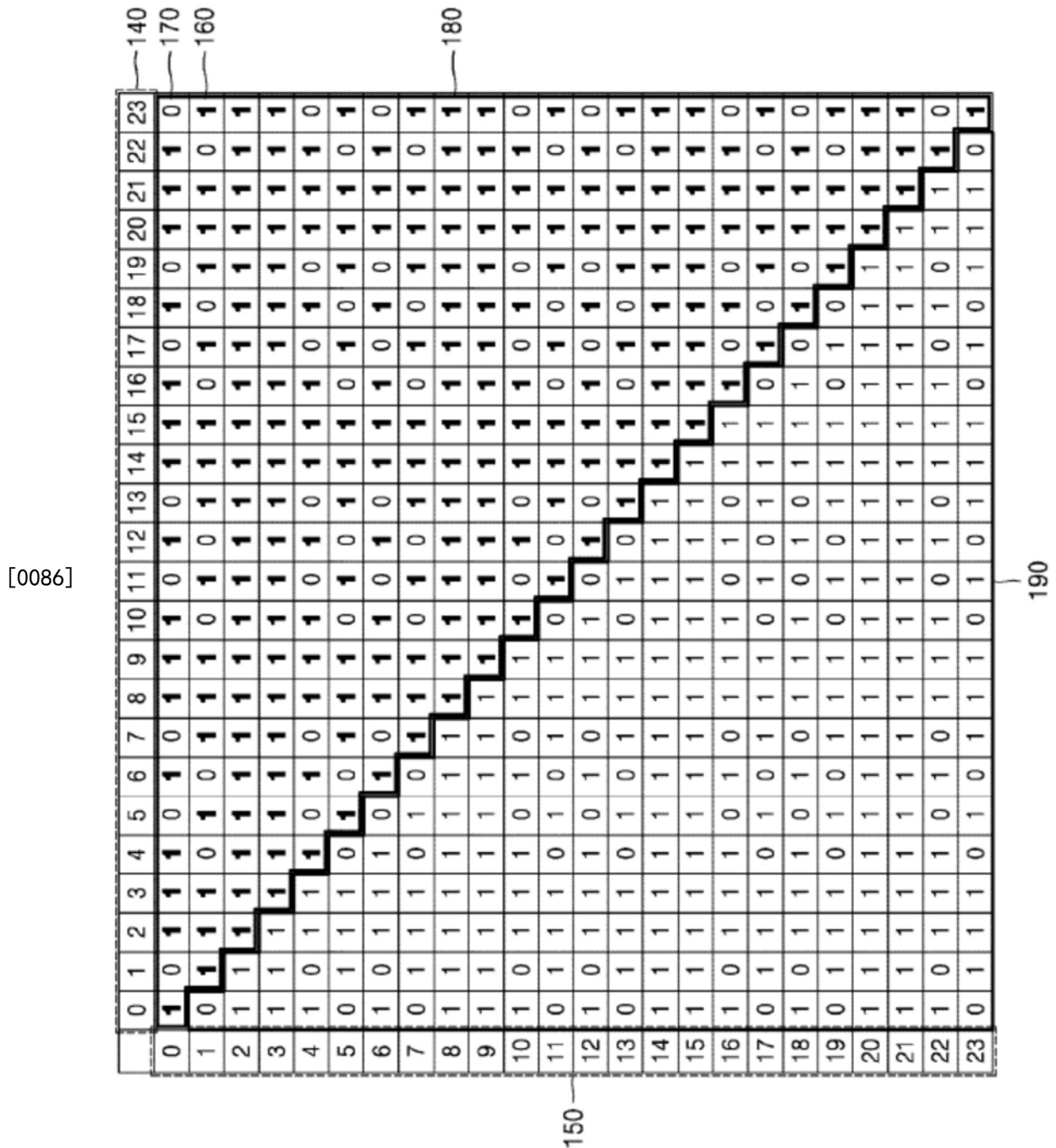
[0081] 表1示出了将22.2声道沉浸式音频信号渲染成立体声信号的格式转换器的混合矩阵的实施例。

[0082] 在表1的混合矩阵中,编了号的24个输入声道在横轴140和纵轴150上表示。编了号的24个输入声道的顺序在协方差分析中没有任何特别的相关性。在表1所示的实施例中,当混合矩阵的每个元素的值都为1时(如参考数字160所示),协方差分析是必要的,但是当混合矩阵的每个元素的值为0时(如参考数字170所示),可以省略协方差分析。

[0083] 例如,在格式转换成立体声输出布局期间没有彼此混合的输入声道(例如声道CM_M_L030和CH_M_R030)的情况下,与未混合的输入声道对应的混合矩阵中的元素的值为0,因此可以省略未混合的声道CM_M_L030和CH_M_R030之间的协方差分析。

[0084] 因此,可以从 $24*24$ 个协方差分析中排除彼此未混合的输入声道的128个协方差分析。

[0085] 另外,因为混合矩阵被配置为按照输入声道对称,所以表1的混合矩阵关于对角线被分成下部190和上部180,并且可以省略对与下部190相对应的区域的协方差分析,在【表1】中:



[0087] 此外,因为仅对与上部180相对应的区域的粗体部分执行了协方差分析,所以最终执行了236个协方差分析。

[0088] 在混合矩阵的值为0(声道未彼此混合)并且基于混合矩阵的对称性去除了不必要的协方差分析的情况下,为了协方差分析应该进行 236×71 个频带 $\times 2 \times 16 \times (48000/2048)$ 次复数乘法。

[0089] 因此,在这种情况下,需要执行50MOPS的性能,因此与对混合矩阵的整个部分执行协方差分析的情况相比,减小了由于协方差分析而导致的系统 负荷。

[0090] 图2是根据一个实施例的解码结构的框图,该解码结构利用13个IC将22.2 声道沉浸式音频信号格式转换为立体声输出声道。

[0091] MPEG-H 3D音频使用CPE以便在受限传输环境中更有效地传送多声道音 频信号。

当对应于单个声道对的两个声道混合成立体声布局时,IC相关性 (ICC) 被设置为1,从而不应用解相关器。因此,这两个声道具有相同的相位信息。

[0092] 换句话说,当通过考虑立体声输出来确定包含在每个CPE中的声道对时,上混后声道对具有相同的平移系数(panning coefficient),这将在稍后进行描述。

[0093] 通过混合包含在单个CPE中的两个同相声道来产生单个IC。当包含在IC 中的两个输入声道被转换为立体声输出声道时,根据基于格式转换器转换规则的混合增益和均衡(EQ)值对单个IC信号进行缩混。在这种情况下,因为包含在单个CPE中的两个声道是同相声道,所以不需要在缩混之后的声道间对齐相位的过程。

[0094] MPS212上混器的立体声输出信号彼此之间没有相位差。然而,在图1的实施例中并没有考虑到这一点,因此复杂度不必要地增加了。当再现布局是立体声布局时,可以通过使用单个IC而不是被上混的CPE声道对作为格式转换器的输入,来减少格式转换器的输入声道的数量。

[0095] 根据图2所示的实施例,不是每个CPE比特流210经历MPS 212上混以产生两个声道,而是每个CPE比特流210经历IC处理220以生成单个IC 221。在这种情况下,由于低音扬声器声道不形成CPE,所以每个低音扬声器声道信号变成IC信号。

[0096] 根据图2的实施例,在22.2声道的情况下,包含用于一般声道的11个CPE 的IC和用于2个低音扬声器声道的IC的13个IC(即, $N_{in}=13$),理论上变成格式转换器230的输入声道。因此,格式转换器230执行13*2缩混。

[0097] 在这样的立体声再现布局情况下,利用IC进一步消除了在通过MPS212上混并且然后通过格式转换进行缩混的过程中产生的不必要的处理,从而进一步降低了解码器的复杂度。

[0098] 当用于单个CPE的两个输出声道i和j的混合矩阵 $M_{mix}(i, j)$ 的值为1时,可以将声道间关联(ICC) $ICC^{1:m}$ 设置为1,并且可以省略解相关和残差处理。

[0099] IC被定义为与格式转换器的输入相对应的虚拟中间声道。如图2所示,每个IC处理块220通过使用MPS212有效载荷(payload)(例如,CLD)和渲染参数(例如,EQ值和增益值)来生成IC信号。EQ值和增益值表示在格式转换器的转换规则表中定义的用于MPS212块的输出声道的渲染参数。

[0100] 表2示出了使用IC将22.2声道沉浸式音频信号渲染为立体声信号的格式转换器的混合矩阵的一个实施例。

[0101] 【表2】

[0102]

	A	B	C	D	E	F	G	H	I	J	K	L	M
A	1	1	1	1	1	1	1	1	1	1	1	1	1
B	1	1	1	1	1	1	1	1	1	1	1	1	1
C	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1
E	1	1	1	1	1	1	1	1	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	0	0
G	1	1	1	1	1	1	1	1	1	0	0	0	0
H	1	1	1	1	1	1	1	1	1	0	0	0	0

I	1	1	1	1	1	1	1	1	1	0	0	0	0
J	1	1	1	1	1	0	0	0	0	1	1	1	1
K	1	1	1	1	1	0	0	0	0	1	1	1	1
L	1	1	1	1	1	0	0	0	0	1	1	1	1
M	1	1	1	1	1	0	0	0	0	1	1	1	1

[0103] 与表1类似,表2的混合矩阵的水平轴和垂直轴表示输入声道的索引,并且索引的顺序在协方差分析中并不是很重要。

[0104] 如上所述,因为一般混合矩阵具有基于对角线的对称性,所以表2的混合矩阵也基于对角线被划分为上部和下部,并且因此对于这两个部分中选定的部分的协方差分析可以省略。在格式转换为立体声输出声道布局期间未混合的输入声道的协方差分析也可以省略。

[0105] 然而,与表1的实施例相比,根据表2的实施例,包括由一般声道组成的11个IC和2个低音扬声器声道的13个声道被缩混成立体声输出声道,并且格式转换器的输入声道数 N_{in} 是13。

[0106] 因此,根据如表2的使用IC的实施例,执行75次协方差分析,理论上需要19MOPS的性能。因此,与不使用IC的情况相比,由协方差分析而导致的格式转换器的负荷可以大大降低。

[0107] 在格式转换器中定义了用于缩混的缩混矩阵 M_{Dmx} ,并且利用 M_{Dmx} 计算的混合矩阵 M_{Mix} ,如下:

```

MMix = zero Nin × Nin Matrix
for i = 1 to Nout
  for j = 1 to Nin
    set_l = 0
    if MDmx(i, j) > 0.0
      set_l = 1
    end
    for k = 1 to Nin
      set_k = 0
      if MDmx(i, j) > 0.0
        set_k = 1
      end
      if set_l == 1 and set_k == 1
        MMix(j, k) = 1
      end
    end
  end
end
end
end

```

[0108]

[0109] 每个OTT解码块输出对应于声道编号i和j的两个声道,并且将混合矩阵 M_{Mix} 为1的

情况设置为 $ICC^{1,m}=1$ ，由此计算上混矩阵 $R_2^{l,m}$ 的 $H11_{OTT}^{l,m}$ 和 $H21_{OTT}^{l,m}$ 。这样，每个OTT解码块都不使用解相关器。

[0110] 表3示出了根据本发明实施例的使用IC来配置22.2声道的CPE结构。

[0111] 【表3】

输入声道	单元	要应用于左输出声道的混合增益	要应用于右输出声道的混合增益	内部声道
CH_M_000	CPE	0.707	0.707	ICH_A
CH_L_000				
CH_U_000	CPE	0.707	0.707	ICH_B
CH_T_000				
CH_M_180	CPE	0.707	0.707	ICH_C
CH_U_180				
CH_LFE2	LFE	0.707	0.707	ICH_D
CH_LFE3	LFE	0.707	0.707	ICH_E
CH_M_L135	CPE	1	0	ICH_F
CH_U_L135				
CH_M_L030	CPE	1	0	ICH_G
CH_L_L045				
CH_M_L090	CPE	1	0	ICH_H
CH_U_L090				
CH_M_L060	CPE	1	0	ICH_I
CH_U_L045				
CH_M_R135	CPE	0	1	ICH_J
CH_U_R135				
CH_M_R030	CPE	0	1	ICH_K
CH_L_R045				
CH_M_R090	CPE	0	1	ICH_L
CH_U_R090				
CH_M_R060	CPE	0	1	ICH_M
CH_U_R045				

[0112] 当22.2声道比特流具有如表3所示的结构时，可以将13个IC定义为ICH_A 至ICH_M，并且可以如表2中那样确定13个IC的混合矩阵。

[0114] 表3的第一列指示输入声道的索引，并且其第一行指示输入声道是否构成 CPE、应用于立体声声道的混合增益以及IC的索引。

[0115] 例如，当CM_M_000和CM_L_000是包含在单个CPE中的ICH_A IC时，分别要应用到左输出声道和右输出声道以便将CPE上混到立体声输出声道的混合增益的值为0.707。换句话说，上混到左输出声道和右输出声道的信号以相同的大小被再现。

[0116] 作为另一示例，当CH_M_L135和CH_U_L135是包括在单个CPE中的 ICH_F IC时，要应用到左输出声道的混合增益的值为1，并且要应用到右输出声道的混合增益的值为0，以便将CPE上混到立体声输出声道。换句话说，所有信号仅通过左输出声道而不通过右输出声道来再现。

[0117] 另一方面，当CH_M_R135和CH_U_R135是包括在单个CPE中的ICH_F IC 时，要应用到左输出声道的混合增益的值为0且要应用到右输出声道的混合增益的值为1，以将CPE上

混到立体声输出声道。换句话说,所有信号只通过右 输出声道而不通过左输出声道来再现。

[0118] 图3是根据一个实施例的用于从单个CPE生成单个IC的装置的框图。

[0119] 通过将正交镜像滤波器(QMF)域的格式转换参数(例如,CLD、增益 和EQ)应用于经缩混的单声道信号,可以产生单个CPE的IC。

[0120] 图3的IC生成装置包括:上混器310、缩放器320和混合器330。

[0121] 在输入通过缩混用于CH_M_000和CH_L_000这对声道的信号而获得的 CPE信号340的情况下,上混器310通过使用CLD参数上混CPE信号340。CPE 信号340可以通过上混器310被上混成用于CH_M_000的信号351和用于 CH_L_000的信号352,并且上混后的信号351和352可以保持相同的相位并且 可以在格式转换器中被混合在一起。

[0122] 分别使用缩放器320和缩放器321,以子带为单位,用与在格式转换器中 定义的转换规则相对应的增益和EQ值,对作为上混结果的CH_M_000声道信 号351和CH_L_000声道信号352进行缩放。

[0123] 当缩放后的信号361和362作为针对CH_M_000和CH_L_000这对声道进 行缩放的结果而生成时,混合器330混合缩放后的信号361和362,并对混合的 结果进行功率归一化,以生成IC信号ICH_A 370,IC信号ICH_A 370是用于格 式转换的中间声道信号。

[0124] 在这种情况下,用于没有利用CLD上混的单声道单元(SCE)和低音扬声 器声道的IC与原始输入声道相同。

[0125] 由于使用IC的核心编解码器输出是在混合QMF域中执行的,所以不执行 ISO IEC23308-3 10.3.5.2的处理。为了分配核心编码器的每个声道,定义如表 4-6所示的附加声道分配规则和缩混规则。

[0126] 表4示出了根据本发明实施例的与解码器输入声道相对应的IC的类型。

[0127] **【表4】**

[0128]

类型	声道	平移 (L,R)
低频效果	CH_LFE1, CH_LFE2, CH_LFE3	(0.707, 0.707)
中央	CH_M_000, CH_L_000, CH_U_000, CH_T_000, CH_M_180, CH_U_180	(0.707, 0.707)
左	CH_M_L022, CH_M_L030, CH_M_L045, CH_M_L060, CH_M_L090, CH_M_L110, CH_M_L135, CH_M_L150, CH_L_L045, CH_U_L045, CH_U_L030, CH_U_L045, CH_U_L090, CH_U_L110, CH_U_L135, CH_M_LSCR, CH_M_LSCH	(1, 0)
右	CH_M_R022, CH_M_R030, CH_M_R045, CH_M_R060, CH_M_R090, CH_M_R110, CH_M_R135, CH_M_R150, CH_L_R045, CH_U_R045, CH_U_R030, CH_U_R045, CH_U_R090, CH_U_R110, CH_U_R135, CH_M_RSCR, CH_M_RSCH	(0,1)

[0129] IC对应于格式转换器与核心编码器的输入声道之间的中间声道,并且包 括四种类型的IC,即,低音扬声器声道、中央声道、左声道和右声道。

[0130] 当以CPE表示的不同类型的声道具有相同的IC类型时,格式转换器具有 相同的平移系数和相同的混合矩阵,因此可以使用IC。换句话说,当包括在 CPE中的两个声道具有相同的IC类型时,IC处理是可能的,并且因此需要用 具有相同IC类型的声道来配置CPE。

[0131] 当解码器输入声道对应于低音扬声器声道,即CH_LFE1、CH_LFE2或 CH_LFE3时,解码器输入声道的IC类型被确定为低音扬声器声道CH_I_LFE。

[0132] 当解码器输入声道对应于中央声道,即CH_M_000、CH_L_000、CH_U_000、CH_T_000、CH_M_180或CH_U_180时,解码器输入声道的IC类型被确定为中央声道CH_I_CNTR。

[0133] 当解码器输入声道对应于左声道,即CH_M_L022、CH_M_L030、CH_M_L045、CH_M_L060、CH_M_L090、CH_M_L110、CH_M_L135、CH_M_L150、CH_L_L045、CH_U_L045、CH_U_L030、CH_U_L045、CH_U_L090、CH_U_L110、CH_U_L135、CH_M_LSCR或CH_M_LSCH,解码器输入声道的IC类型被确定为左声道CH_I_LEFT。

[0134] 当解码器输入声道对应于右声道,即,CH_M_R022、CH_M_R030、CH_M_R045、CH_M_R060、CH_M_R090、CH_M_R110、CH_M_R135、CH_M_R150、CH_L_R045、CH_U_R045、CH_U_R030、CH_U_R045、CH_U_R090、CH_U_R110、CH_U_R135、CH_M_RSCR或CH_M_RSCH,解码器输入声道的IC类型被确定为右声道CH_I_RIGHT。

[0135] 表5示出根据本发明实施例的根据IC类型另外定义的声道的位置。

[0136] 【表5】

[0137]

在 ISO/IEC 23001-8 中定义的扬声器几何 (Loudspeaker Geometry)	声道	方位角 (度)	仰角 (度)	扇区的方位角开始角度 (度)	扇区的方位角结束角度 (度)	扇区的仰角开始角度 (度)	扇区的仰角结束角度 (度)	声道是 LFE	位置是相对的
43	CH_I_CNTR	0	0	0	0	0	0	0	0
44	CH_I_LFE	0	n/a	n/a	n/a	n/a	n/a	1	0
45	CH_I_LEFT	30	0	30	30	0	0	0	0
46	CH_I_RIGHT	-30	0	-30	-30	0	0	0	0

[0138] CH_I_LFE是低音扬声器声道,并且位于0度的仰角处,并且CH_I_CNTR 对应于仰角和方位角均为0度的声道。CH_I_LFET对应于仰角为0度且方位角 在左侧的30度至60度之间的扇区的声道,CH_I_RIGHT对应于仰角为0度且方 位角在右侧30度至60度之间的扇区的声道。

[0139] 在这种情况下,重新定义的IC的位置不是声道之间的相对位置,而是相 对于参考

点的绝对位置。

[0140] 甚至可以将IC应用于由CPE对组成的四声道单元(QCE, Quadruple Channel Element),稍后将对其进行描述。

[0141] 可以使用两种方法来产生IC。

[0142] 第一种方法是在MPEG-H 3D音频编码器中进行预处理,而第二种方法是在MPEG-H 3D音频解码器中进行后处理。

[0143] 当在MPEG中使用IC时,表5可作为新行添加到ISO/IEC 23008-3表90中。

[0144] 表6示出根据本发明的实施例的对应于IC类型的格式转换器输出声道,以及将被应用于每个格式转换器输出声道的增益和EQ指标。

[0145] **【表6】**

源	目的地	增益	EQ指标
CH_I_CNTR	CH_M_L030, CH_M_R030	1.0	0(关)
CH_I_LFE	CH_M_L030, CH_M_R030	1.0	0(关)
CH_I_LEFT	CH_M_L030	1.0	0(关)
CH_I_RIGHT	CH_M_L030	1.0	0(关)

[0147] 为了使用IC,应在格式转换器中添加诸如表6的附加规则。

[0148] IC信号是考虑了格式转换器的增益值和EQ值而产生的。因此,可以使用如表6所示的增益值为1并且EQ指标为0的附加转换规则来产生IC信号。

[0149] 当IC类型是对应于中央声道的CH_I_CNTR或对应于低音扬声器声道的CH_I_LFE时,输出声道是CH_M_L030和CH_M_R030。此时,因为增益值被确定为1, EQ指标被确定为0,并且两个立体声输出声道被全部使用,所以每个输出声道信号应该被乘以 $1/\sqrt{2}$ 以便保持输出信号的功率。

[0150] 当IC类型是对应于左声道的CH_I_LEFT时,输出声道是CH_M_L030。此时,因为增益值被确定为1, EQ指标被确定为0,并且仅使用左输出声道,所以增益1被应用到CH_M_L030,增益0被应用到CH_M_R030。

[0151] 当IC类型是对应于右声道的CH_I_RIGHT时,输出声道是CH_M_R030。此时,因为增益值被确定为1, EQ指标被确定为0,并且仅使用右输出声道,所以增益1被应用到CH_M_R030,并且增益0被应用到CH_M_L030。

[0152] 在这种情况下,通用格式转换规则被应用于其IC和输入声道相同的SCE声道。

[0153] 当在MPEG中使用IC时,表6可作为新行添加到ISO/IEC 23008-3表96中。

[0154] 表7-表15示出了为了在MPEG中利用IC而要更改的现有标准的一部分。

[0155] 表7示出了根据本发明实施例的ICGConfig的句法。

[0156] **【表7】**

[0157]

句法	比特数	助记符
<pre> ICGConfig () { if (ICGDisabledPresent) { for (elemIdx=0; elemCPE=0; elemIdx<numElements; ++elemIdx) { If (usacElementType[elemIdx] == ID_USAC_CPE) { ICGDisabledCPE[elemCPE]; elemCPE++; } } } if (ICGPreAppliedPresent) { for (elemIdx=0; elemCPE=0; elemIdx<numElements; ++elemIdx) { If (usacElementType[elemIdx] == ID_USAC_CPE) { ICGPreAppliedCPE[elemCPE]; elemCPE++; } } } } </pre>	1	Uimsbf

[0158] 表7中所示的ICGconfig定义了将在IC处理块中执行的处理的类型。

[0159] ICGDisabledPresent指示针对CPE的至少一个IC处理是否由于声道分配而被禁用。换句话说,ICGDisabledPresent是表示至少一个ICGDisabledCPE是否具有值1的指示符。

[0160] ICGDisabledCPE指示每个针对CPE的IC处理是否由于声道分配而被禁用。换句话说,ICGDisabledCPE是表示每个CPE是否使用IC的指示符。

[0161] ICGPreAppliedPresent指示是否已经考虑了ICG来编码至少一个CPE。

[0162] ICGPreAppliedCPE是表示是否已经考虑了ICG来编码每个CPE(即,ICG是否已经在编码器中被预处理)的指示符。

[0163] 当针对每个CPE将ICGAppliedPresent设置为1时,ICGPreAppliedCPE(其为ICGPreAppliedCPE的1-位标志)被读出。换句话说,确定是否应该将ICG应用于每个CPE,并且当确定应该将ICG应用于每个CPE时,就确定了ICG是否已经在编码器中被预处理。如果确定ICG已经在编码器中被预处理,则解码器不应用ICG。另一方面,如果确定ICG没有在编码器中被预处理,则解码器应用ICG。

[0164] 当沉浸式音频输入信号是利用CPE或QCE由MPS212编码的并且输出布局是立体声布局时,为了减少格式转换器的输入声道的数量,核心编解码器 解码器生成IC信号。在这种情况下,对于其ICGDisabledCPE被设置为1的CPE, 省略了IC信号生成。IC处理对应于将解码的单声道信号乘以ICG的处理,并且 ICG是根据CLD和格式转换参数计算出来的。

[0165] ICGDisabledCPE[n]指示第n个CPE是否可能经历IC处理。当包括在第n个 CPE中的两个声道属于表4中定义的不同声道组时,第n个CPE能够经历IC处 理,并且 ICGDisabledCPE[n]被设置为0。

[0166] 例如,当输入声道中的CH_M_L060和CH_T_L045构成单个CPE时,由于 两个声道属于同一个声道组,所以可以将ICGDisabledCPE[n]设置为0,并且可 以生成CH_I_LEFT的IC。另一方面,当输入声道中的CH_M_L060和 CH_M_000构成单个CPE时,由于两个声道属于不同的声道组,所以 ICGDisabledCPE[n]被设置为1,并且不执行IC处理。

[0167] 对于包括CPE对的QCE,在用属于一个组的四个声道配置QCE的情况(1) 下,或者在用属于一个组的两个声道和属于另一个组的两个声道配置QCE的 情况(2)下,IC处理是可能的,并且ICGDisableCPE[n]和ICGDisableCPE[n+1] 都被设置为0。

[0168] 作为情况(1)中的示例,当用CH_M_000、CH_L_000、CH_U_000和 CH_T_000四个声道配置QCE时,IC处理是可能的,并且QCE的IC类型是 CH_I_CNTR。作为的情况(2)的例子,在用 CH_M_L060,CH_U_L045, CH_M_R060和CH_U_R045这四个声道配置QCE的情况下,IC处理是可能的, 并且QCE的IC类型为CH_I_LEFT和CH_I_RIGHT。

[0169] 在除了情况(1)和情况(2)的情况下,构成对应的QCE的CPE对的 ICGDisableCPE [n]和ICGDisableCPE[n+1]均应被设置为1。

[0170] 当编码器应用ICG时,与解码器应用ICG时相比,解码器所需的复杂度可 以降低。

[0171] ICGConfig的ICGPreAppliedCPE[n]指示是否已经在编码器中将ICG应用 于第n个 CPE。如果ICGPreAppliedCPE[n]为真,则解码器的IC处理块绕开用于 立体声再现第n个CPE的缩混信号。另一方面,如果ICGPreAppliedCPE[n]为假, 则解码器的IC处理块将ICG应用于缩混信号。

[0172] 如果ICGDisableCPE[n]是1,则不可能计算用于对应的QCE或CPE的ICG, 并且因此 将ICGPreApplied[n]设置为0。对于包含CPE对的QCE,包含在QCE中 的两个CPE的指标 ICGPreApplied[n]和ICGPreApplied[n+1]应具有相同的值。

[0173] 现在将使用表8-表16描述对于IC处理要被更改或添加的比特流结构和比 特流句 法。

[0174] 表8示出了根据本发明实施例的mpeg3daExtElementConfig()的句法。

[0175] **【表8】**

[0180]

usacExtElementType	值
ID_EXT_ELE_FILL	0
ID_EXT_ELE_MPEGS	1
ID_EXT_ELE_SAOC	2
ID_EXT_ELE_AUDIOPREROLL	3
ID_EXT_ELE_UNIDRC	4
ID_EXT_ELE_OBJ_METADATA	5
ID_EXT_ELE_SAOC_3D	6
ID_EXT_ELE_HOA	7
ID_EXT_ELE_FMT_CNVTR	8
ID_EXT_ELE_ICG	9
/*为 ISO 使用保留*/	10-127
/*为在 ISO 范围之外使用保留*/	128 以及更高

注意：授权特定应用的 usacExtElementType 值位于为 ISO 范围之外使用而保留的空间中。因为解码器需要最小结构以跳过这些扩展，所以解码器跳过这些特定应用的 usacExtElementType 值。

[0181] 如表9所示,在usacExtElementType中,可以为IC处理添加 ID_EXT_ELE_ICG,并且 ID_EXT_ELE_ICG的值可以是9。

[0182] 表10示出了根据本发明实施例的speakerLayoutType的句法。

[0183] 【表10】

[0184]

值	含义
0	如在 ISO/IEC 23001-8 中定义的, 通过 ChannelConfiguration 索引来告知扬声器布局
1	如在 ISO/IEC 23001-8 中定义的, 通过 LoudspeakerGeometry 索引列表来告知扬声器布局
2	通过明确的几何位置信息列表来告知扬声器布局
3	通过 LCChannelConfiguration 索引来告知扬声器布局。注意 LCChannelConfiguration 与 ChannelConfiguration 具有相同的布局但是不同的声道顺序, 从而能够实现利用 CPE 的最佳内部声道结构

[0185] 对于IC处理,应定义用于IC的扬声器布局类型speakerLayoutType。表10 显示了 speakerLayoutType的每个值的含义。

[0186] 当speakerLayoutType是3时,通过索引LCChannelConfiguration来表示扬声器布局。索引LCChannelConfiguration与ChannelConfiguration具有相同的布局, 但 LCChannelConfiguration所具有声道分配顺序能够实现使用CPE的最佳IC结构。

[0187] 表11示出了根据本发明实施例的SpeakerConfig3d()的句法。

[0188] 【表11】

句法	比特数	助记符
SpeakerConfig3d() { speakerLayoutType; if (speakerLayoutType == 0 speakerLayoutType == 3) { CICPspeakerLayoutIdx; } else { numSpeakers = escapedValue(5, 8, 16) + 1; if (speakerLayoutType == 1) { for (i = 0; i < numSpeakers; i++) { CICPspeakerIdx; } } if (speakerLayoutType == 2) { mpeg3daFlexibleSpeakerConfig(numSpeakers); } } }	2 6 7	uimsbf uimsbf uimsbf

[0190] 当如上所述speakerLayoutType为3时,实施例使用与 CICPspeakerLayoutIdx相同的布局,但就最佳声道分配排序而言实施例与 CICPspeakerLayoutIdx不同。

[0191] 当speakerLayoutType是3并且输出布局是立体声布局时,输入声道数 N_{in} 变为核心编解码器之后的IC的数量。

[0192] 表12示出了根据本发明实施例的immersiveDownmixFlag的句法。

[0193] 【表12】

immersiveDownmixFlag	含义
0	应该如在子句 10 中定义的那样应用一般格式转换器
1	如果由 LoudspeakerRendering()通知的本地扬声器设置被通知为 (speakerLayoutType==0 或 3,CICPspeakerLayoutIdx==5) 或 (speakerLayoutType==0 或 3,CICPspeakerLayoutIdx==6)而与可能通知的扬声器位移角无关,则应该如子句 11 中定义的那样应用沉浸式渲染格式转换器。 在其他所有情况下,应该如在子句 10 中定义的那样应用一般格式转换器

[0195] 通过新定义用于IC的扬声器布局类型,immersiveDownmixFlag也应该被校正。当immersiveDownmixFlag为1时,应如表12所示添加用于处理 speakerLayoutType为3的情况的句子。

[0196] 目标展开应该满足以下要求:

[0197] -本地云扬声器设置由LoudspeakerRendering()通知,

[0198] -speakerLayoutType应该是0或3,

[0199] -CICPspeakerLayoutIdx具有值4、5、6、7、9、10、11、12、13、14、15、16、17或18。

[0200] 表13示出了根据本发明实施例的SAOC3DgetNumChannels()的句法。

[0201] 如表13所示,SAOC3DgetNumChannels应该被校正以包括 speakerLayoutType为3的情况。

[0202] 【表13】

句法	比特数	助记符
<pre>SAOC3DgetNumChannels(Layout) { numChannels = numSpeakers; for (i = 0; i < numSpeakers; i++){ if (Layout.isLFE[i] == 1){ numChannels = numChannels - 1; } } return numChannels; }</pre>		<p>Note 1</p> <p>Note 2</p>
<p>注释 1: 函数 SAOC3DgetNumChannels()返回可用的非-LFE 声道的数量 numChannels。</p> <p>注释 2: numSpeakers 在 SpeakerConfig3d()的句法中定义。如在 ISO/IEC 23001-8 中定义的, 如果 speakerLayoutType==0 或 speakerLayoutType==3, 则 numSpeakers 表示对应于 ChannelConfiguration 值的扬声器个数, CICPSpeakerLayoutIdx。</p>		

[0203]

[0204] 表14示出了根据本发明实施例的声道分配顺序的句法。

[0205] 表14按照扬声器布局或LCChannelConfiguration,将声道数、声道顺序以及可能的IC类型指示为针对IC新定义的声道分配顺序。

[0206] 【表14】

[0207]

扬声器布局索引或 LCChannelConfiguration	声道数	声道 (有顺序)	可能的内部声道类型
1	1	CH_M_000	中央
2	2	CH_M_L030, CH_M_R030	左 右
3	3	CH_M_000, CH_M_L030, CH_M_R030	中央 左 右
4	4	CH_M_000, CH_M180, CH_M_L030, CH_M_R030,	中央 左 右
5	5	CH_M_000, CH_M_L030, CH_M_L110, CH_M_R030, CH_M_R110	中央 左 右
6	6	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_M_R030, CH_M_R110	中央 Lfe 左 右
7	8	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_M_L060, CH_M_R030, CH_M_R110, CH_M_R060	中央 Lfe 左 右
8		n.a.	
9	3	CH_M_180, CH_M_L030, CH_M_R030	中央 左 右
10	4	CH_M_L030, CH_M_L110, CH_M_R030, CH_M_R110	左 右
11	7	CH_M_000, CH_M_180 CH_LFE1, CH_M_L030, CH_M_L110, CH_M_R030, CH_M_R110	中央 Lfe 左 右
12	8	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_M_L135, CH_M_R030, CH_M_R110, CH_M_R135	中央 Lfe 左 右
13	24	CH_M_000, CH_L_000, CH_U_000, CH_T_000, CH_M_180, CH_T_180, CH_LFE2, CH_LFE3, CH_M_L135, CH_U_L135, CH_M_L030, CH_L_L045, CH_M_L090, CH_U_L090, CH_M_L060, CH_U_L045, CH_M_R135, CH_U_R135, CH_M_R030, CH_L_R045, CH_M_R090, CH_U_R090, CH_M_R060, CH_U_R045	中央 Lfe 左 右

[0208]

14	8	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_U_L030, CH_M_R030, CH_M_R110, CH_U_R030	中央 Lfe 左 右
15	12	CH_M_000, CH_U_180, CH_LFE2, CH_LFE3, CH_M_L030, CH_M_L135, CH_M_L090, CH_U_L045, CH_M_R030, CH_M_R135, CH_M_R090, CH_U_R045	中央 Lfe 左 右
16	10	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_U_L030, CH_U_L110, CH_M_R030, CH_M_R110, CH_U_R030, CH_U_R110	中央 Lfe 左 右
17	12	CH_M_000, CH_U_000, CH_T_000 CH_LFE1, CH_M_L030, CH_M_L110, CH_U_L030, CH_U_L110, CH_M_R030, CH_M_R110, CH_U_R030, CH_U_R110,	中央 Lfe 左 右
18	14	CH_M_000, CH_U_000, CH_T_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_M_L150, CH_U_L030, CH_U_L110, CH_M_R030, CH_M_R110, CH_M_R150, CH_U_R030, CH_U_R110	中央 Lfe 左 右
19	12	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L135, CH_M_L090, CH_U_L030, CH_U_L135, CH_M_R030, CH_M_R135, CH_M_R090, CH_U_R030, CH_U_R135	中央 Lfe 左 右
20	14	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L135, CH_M_L090, CH_U_L045, CH_U_L135, CH_M_LSCR, CH_M_R030, CH_M_R135, CH_M_R090, CH_U_R045, CH_U_R135, CH_M_RSCR	中央 Lfe 左 右

[0209] 表15示出了根据本发明实施例的mpeg3daChannelPairElementConfig()的句法。

[0210] 对于IC处理,如表15所示,当stereoConfigIndex大于0时,应校正 mpeg3daChannelPairElementConfig()使得Mps212Config()处理之后是 isInternalChangeProcessed()。

[0211] 【表15】

句法	比特数	助记符
<pre> mpeg3daChannelPairElementConfig(sbrRatiIndex) { mpeg3daCoreConfig(); if (enhancedNoiseFilling) { igfIndependentTiling; } if (sbrRatiIndex > 0) { SbrConfig(); stereoConfigIndex; } else { stereoConfigIndex = 0; } if (stereoConfigIndex > 0) { Mps212Config(stereoConfigIndex); } } </pre>		
[0212]	<pre> isInternalChannelProcessed } qcelIndex; if (qcelIndex > 0) { shiftIndex0; if (shiftIndex0 > 0) { shiftChannel0; } } shiftIndex1; if (shiftIndex1 > 0) { shiftChannel1; } } </pre>	<p>1 uimsbf</p> <p>2 uimsbf</p> <p>1 uimsbf</p> <p>nBits¹⁾</p> <p>1 uimsbf</p> <p>nBits¹⁾</p>
<p>¹⁾ nBits = floor(log2(numAudioChannels + numAudioObjects + numHOATransportChannels + numSAOCTransportChannels - 1)) + 1</p>		

[0213] 图4是根据本发明实施例的解码器的ICG应用单元将ICG应用于IC信号的详细框图。

[0214] 当speakerLayout是3、isInternalProcessed是0并且再现布局是立体声布局的条件被满足并且因此解码器应用ICG时，执行如图4中的IC处理。

[0215] 图4中所示的ICG应用单元包括ICG获取器410和乘法器420。

[0216] 假设输入CPE包括CH_M_000和CH_L_000这对声道，则当输入CPE的单声道QMF子带样本430被输入时，ICG获取器410通过使用CLD来获取ICG。乘法器420通过将所接收的单声道QMF子带样本430乘以所获取的ICG来获取IC信号ICH_A 440。

[0217] 可以通过将CPE的单声道QMF子带样本与ICG $G_{ICH}^{l,m}$ 相乘来简单地重组IC信号，其中l指示时间索引且m指示频率索引。

[0218] ICG $G_{ICH}^{l,m}$ 在等式1中定义：

[0219] 【等式1】

[0220]
$$G_{ICH}^{l,m} = \frac{\sqrt{(c_{left}^{l,m} \times G_{left} \times G_{EQ, left}^m)^2 + (c_{right}^{l,m} \times G_{right} \times G_{EQ, right}^m)^2}}{\sqrt{(c_{left}^{l,m} \times G_{left} \times G_{EQ, left}^m + c_{right}^{l,m} \times G_{right} \times G_{EQ, right}^m)}} \times c_{left}^{l,m} \times G_{left} \times G_{EQ, left}^m + c_{right}^{l,m} \times G_{right} \times G_{EQ, right}^m$$

[0221] 其中， $c_{left}^{l,m}$ 和 $c_{right}^{l,m}$ 指示CLD的平移系数， G_{left} 和 G_{right} 指示在格式转换规则中定义的增益，并且 $G_{EQ, left}^m$ 和 $G_{EQ, right}^m$ 指示在格式转换规则中定义的EQ值的第m个频带的增益。

[0222] 图5是示出根据本发明实施例的当编码器预处理ICG时的解码过程的框图。

[0223] 当speakerLayout是3、isInternalProcessed是1并且再现布局是立体声布局 的条件被满足且因此编码器应用ICG并发送ICG时,执行如图5中的IC处理。

[0224] 当输出布局是立体声布局时,MPEG-H 3D音频编码器预处理对应于CPE 的ICG,使得解码器绕开MPS212,因此可以减少解码器的复杂度。

[0225] 然而,当输出布局不是立体声布局时,MPEG-H 3D音频编码器不执行IC 处理,因此解码器需要执行乘ICG的逆 $1/G_{ICH}^{l,m}$ 并运行MPS212的处理以实现解 码,如图5所示。

[0226] 类似于图3和图4,假定输入CPE包括CH_M_000和CH_L_000这对声道。当输入具有在编码器中预处理的ICG的单声道QMF子带样本540时,解码器确 定输出布局是否是立体声布局,如附图标记510所示。

[0227] 当输出布局是立体声布局时,使用IC,并且因此解码器将所接收的单声 道QMF子带样本540作为IC ICH_A 550的IC信号输出。另一方面,当输出布局 不是立体声布局时,在IC处理期间不使用IC,因此解码器执行逆ICG处理520 以恢复如附图标记560所指示的经IC处理过的信号,以及通过如附图标记530 所示的MPS212上混所恢复的信号,从而输出CH_M_000的信号571和 CH_L_000的信号572。

[0228] 因为在输入声道的数量大并且输出声道的数量小时由格式转换器中协方 差分析导致的负荷成为问题,所以当输出布局是立体布局时,MPEG-H音频 具有最大的解码复杂度。

[0229] 另一方面,当输出布局不是立体布局时,在每帧有两组CLD的情况下,为了乘以ICG的逆而添加的运算的数目是(5次乘法、2次加法、1次除法、取 1次平方根 \approx 55次运算) \times (71条频带) \times (2个参数集合) \times (48000/2048) \times (13个 IC),因此变成大约2.4MOPS且在系统上并没有充当大负荷。

[0230] 在生成IC之后,IC的QMF子带样本、IC数量以及IC的类型被发送到格式 转换器,并且格式转换器中的协方差矩阵的大小取决于IC数量。

[0231] 表16示出了根据本发明实施例的基于声道单元和再现布局确定的MPEG 环绕(MPS)和频带复制(SBR)的解码场景。

[0232] 【表16】

再现布局	单元	MPS和SBR的顺序
立体声	CPE	MPS在单声道SBR之后
立体声	CPE	MPS在立体声SBR之前
立体声	QCE	两个MPS在两个立体声SBR之前
非立体声	CPE/QCE	与顺序无关

[0234] MPS是通过使用包含空间线索参数和人对多声道音频信号的感知特性的 辅助数据来编码多声道音频信号的技术,所述空间线索参数表示混合到最小 声道(单声道或立体声)的缩混。

[0235] MPS编码器接收N个多声道音频信号并且提取空间参数作为辅助数据, 例如,该空间参数可以表示成基于双耳效应和声道之间的相关性的两个耳朵 的音量之间的差异。由于所提取的空间参数的信息量非常少(每个声道不超 过4kbps),因此即使在仅能够提供单声道或立体声音频服务的带宽中也可以 提供高质量的多声道音频。

[0236] MPS编码器还从所接收的N个多声道音频信号生成缩混信号,并且所生成的缩混信号通过例如MPEG USAC(这是一种音频压缩技术)进行编码,并且与空间参数一起发送。

[0237] 此时,由MPS编码器接收的N个多声道音频信号被分析滤波器组分成多个频带。将频域分成子带的代表性方法包括离散傅里叶变换(DFT)或使用QMF。在MPEG环绕中,使用QMF以低复杂度将频域分成子带。当使用QMF时,可以确保其与SBR的兼容性,并且因此可以执行更有效的编码。

[0238] SBR这样一种技术:将低频带复制并粘贴到人类相对不易觉察到的高频带,并且参数化和发送关于高频带信号的信息。因此,根据SBR,可以在低比特率下实现宽带宽。SBR主要用于压缩率高、比特率低的编解码器,并且由于一些高频段信息的丢失难以表达谐波。然而,SBR在听得见的频率内提供了高还原率。

[0239] 除了处理的域不同之外,在IC处理中使用的SBR与ISO/IEC 23003-3:2012相同。在QMF域中定义ISO/IEC 23003-3:2012的SBR,但在混合QMF域中处理IC。因此,当QMF域的索引号是k时,针对IC的整体SBR处理的频率索引号是k+7。

[0240] 在通过立体声再现布局输出CPE时执行单声道SBR解码然后执行MPS解码的解码场景的实施例在图6中示出。

[0241] 在将CPE输出到立体声再现布局时执行MPS解码然后执行立体声SBR解码的解码场景的实施例在图7中示出。

[0242] 图8和图9示出了在通过立体声再现布局输出QCE时对CPE对执行MPS解码然后对每个已解码的信号执行立体声SBR解码的解码场景的实施例。

[0243] 当用以输出CPE或QCE的再现布局不是立体声布局时,执行MPS解码和SBR解码的顺序就不重要了。

[0244] 如下定义由解码器处理的通过MPS212编码的CPE信号:

[0245] `cplx_out_dmx[]`是通过复合预测立体声解码获得的CPE缩混信号。

[0246] `cplx_out_dmx_preICG[]`是已经在编码器中被应用了ICG的单声道信号,其要通过复合预测立体声解码和混合QMF域中的混合QMF分析滤波器组解码被解码。

[0247] `cplx_out_dmx_postICG[]`是已经经历了复合预测立体声解码且在混合QMF域中经历了IC处理并且在解码器中要被应用ICG的单声道信号。

[0248] `cplx_out_dmx_ICG[]`是混合QMF域中的全频带IC信号。

[0249] 如下定义由解码器处理的通过MPS212编码的QCE信号:

[0250] `cplx_out_dmx_L[]`是已经经历复合预测立体声解码的第一CPE的第一声道信号。

[0251] `cplx_out_dmx_R[]`是已经经历了复合预测立体声解码的第一CPE的第二声道信号。

[0252] `cplx_out_dmx_L_preICG[]`是混合QMF域中的第一预先应用ICG的IC信号。

[0253] `cplx_out_dmx_R_preICG[]`是混合QMF域中的第二预先应用ICG的IC信号。

[0254] `cplx_out_dmx_L_postICG[]`是混合QMF域中的第一后应用ICG的IC信号。

[0255] `cplx_out_dmx_R_postICG[]`是混合QMF域中的第二后应用ICG的IC信号。

[0256] `cplx_out_dmx_L_ICG_SBR`是包括用于22.2到2格式转换的缩混参数和由SBR生成的高频分量的第一全频带解码的IC信号。

[0257] `cplx_out_dmx_R_ICG_SBR`是包括用于22.2到2格式转换的缩混参数和由SBR生成

的高频分量的第二全频带解码的IC信号。

[0258] 图6的流程图示出了根据本发明实施例的当通过立体声再现布局输出 CPE时用于执行单声道SBR解码然后执行MPS解码的结构中的IC处理方法。

[0259] 当接收到CPE比特流时,在操作610中首先通过ICGDisabledCPE[n]标志 来判定使用或不使用CPE。

[0260] 当ICGDisabledCPE[n]为真时,在操作620中,如在ISO/IEC 23008-3中所 定义的那样解码CPE比特流。另一方面,当ICGDisabledCPE[n]为假时,在操 作630中,如果SBR是必要的,对CPE比特流执行单声道SBR,并且对其执行 立体声解码以生成缩混信号cplx_out_dmx。

[0261] 在操作640中,通过ICGPreAppliedCPE判定在编码器端是否已经应用ICG。

[0262] 当ICGPreAppliedCPE[n]为假时,在操作650中,缩混信号cplx_out_dmx在 混合QMF域中受到IC处理,从而生成后应用ICG的缩混信号 cplx_out_dmx_postICG。在操作650中,使用MPS参数来计算ICG。通过ISO/IEC 23008-3计算针对CPE去量化的线性CLD值,并且使用等式2计算ICG。

[0263] 通过将缩混信号cplx_out_dmx乘以使用等式2计算出的ICG,来生成后应 用ICG的缩混信号cplx_out_dmx_postICG:

[0264] 【等式2】

$$[0265] \quad G_{ICH}^{l,m} = \sqrt{\left(c_{left}^{l,m} \times G_{left} \times G_{EQ,left}^m\right)^2 + \left(c_{right}^{l,m} \times G_{right} \times G_{EQ,right}^m\right)^2}$$

[0266] 其中 $c_{left}^{l,m}$ 和 $c_{right}^{l,m}$ 指示对于CPE信号的第1个时隙和第m个混合QMF频带 的去量化的线性CLD值, G_{left} 和 G_{right} 指示在ISO/IEC 23008-3表96中(即,在 格式转换规则表中)定义 的输出声道的增益列的值,并且 $G_{EQ,left}^m$ 和 $G_{EQ,right}^m$ 指 示在格式转换规则表中定义的输 出声道的EQ值的第m个频带的增益。

[0267] 当ICGPreAppliedCPE[n]为真时,在操作660中,分析缩混信号 cplx_out_dmx以获 取预先应用ICG的缩混信号cplx_out_dmx_preICG。

[0268] 根据ICGPreAppliedCPE[n]的设置,信号cplx_out_dmx_preICG或 cplx_out_dmx_postICG变成最终的经过IC处理的输出信号cplx_out_dmx_ICG。

[0269] 图7的流程图示出了根据本发明的实施例的当通过立体声再现布局输出 CPE时执 行MPS解码然后执行立体声SBR解码的IC处理方法。

[0270] 根据图7的实施例,与图6的实施例相比,由于MPS解码之后是SBR解码, 所以当IC 未被使用时执行立体声SBR解码。另一方面,当使用IC时,执行单 声道SBR,并且为此将用于 立体声SBR的参数缩混。

[0271] 因此,与图6相比,图7的方法进一步包括:缩混用于两个声道的SBR参 数来生成用 于一个声道的SBR参数的操作780;以及通过使用所生成的SBR参 数来执行单声道SBR的操 作770,并且已经经过单声道SBR的 cplx_out_dmx_ICG成为最终的经过IC处理的输出信号 cplx_out_dmx_ICG。

[0272] 在如图7所示的操作布局中,由于在IC处理之后SBR的执行导致高频分量 扩展,所 以信号cplx_out_dmx_preICG或信号cplx_out_dmx_postICG对应于带 限信号。应当在参数

域中缩混用于上混立体声信号的SBR参数对,以便扩展带限IC信号cplx_out_dmx_preICG或cplx_out_dmx_postICG的带宽。

[0273] SBR参数缩混器应当包括将由于SBR而扩展的高频带乘以格式转换器的EQ值和增益参数的过程。下面将详细描述缩混SBR参数的方法。

[0274] 图8是根据本发明实施例的当通过立体声再现布局输出QCE时使用立体声SBR的结构中的IC处理方法的框图。

[0275] 图8的实施例是ICGPreApplied[n]和ICGPreApplied[n+1]均为0的情况,即在解码器中应用ICG的方法的实施例。

[0276] 参考图8,整体解码过程按照比特流解码810、立体声解码820、混合QMF分析830、IC处理840和立体声SBR850的顺序进行。

[0277] 当包括在QCE中的两个CPE的比特流分别经历比特流解码811和比特流解码812时,从与比特流解码的结果对应的解码信号中提取SBR有效载荷、MPS212有效载荷和CplxPred有效载荷。

[0278] 使用CplxPred有效载荷执行立体声解码821,并且将分别经历混合QMF分析831和混合QMF分析832的立体声解码信号cplx_dmx_L和cplx_dmx_R,分别作为IC处理单元841和IC处理单元842的输入信号发送。

[0279] 此时,生成的IC信号cplx_dmx_L_PostICG和cplx_dmx_R_PostICG是带限信号。因此,这两个IC信号通过使用缩混SBR参数经历立体声SBR 851,所述缩混SBR参数是通过对从两个CPE的比特流提取的SBR有效载荷进行缩混而获得的。带限IC信号的高频率通过立体声SBR 851被扩展,从而生成全频带的经IC处理的输出信号cplx_dmx_L_ICG和cplx_dmx_R_ICG。

[0280] 缩混SBR参数用于扩展带限IC信号的频带以生成全频带IC信号。

[0281] 因此,当对QCE使用IC时,仅使用一个立体声解码块并且仅使用一个立体声SBR块,因此可省略立体声解码块822和立体声SBR块852。换句话说,与每个CPE都被处理的情况相比,图7的情况通过使用QCE实现了简单的解码结构。

[0282] 图9是根据本发明另一实施例的在通过立体声再现布局输出QCE时使用立体声SBR的结构中的IC处理方法的框图。

[0283] 图9的实施例是ICGPreApplied[n]和ICGPreApplied[n+1]都是1的情况,即在编码器中应用ICG的方法的实施例。

[0284] 参考图9,整个解码过程按照比特流解码910、立体声解码920、混合QMF分析930和立体声SBR 950的顺序进行。

[0285] 当编码器已经应用了ICG时,解码器不执行IC处理,并且因此图9的方法省略了图8的IC处理块841和IC处理块842。图9的其他过程与图8的相似,在此省略对这些过程的重复说明。

[0286] 立体声解码信号cplx_dmx_L和cplx_dmx_R分别经过混合QMF分析931和混合QMF分析932,然后作为立体声SBR块951的输入信号被发送。在立体声解码信号cplx_dmx_L和cplx_dmx_R通过立体声SBR块951后,生成全频带的经IC处理的输出信号cplx_dmx_L_ICG和cplx_dmx_R_ICG。

[0287] 当输出声道不是立体声声道时,使用IC可能不合适。因此,当编码器已经应用了

ICG时,如果输出声道不是立体声声道,则解码器应当应用ICG的逆。

[0288] 在这种情况下,如表8所示,MPS和SBR的解码顺序不重要,但是为了便于解释,将描述执行单声道SBR解码然后执行MPS212解码的场景。

[0289] 如等式3所示,使用MPS参数和格式转换参数来计算ICG的逆IG:

[0290] 【等式3】

$$[0291] \quad IG_{ICH}^{l,m} = \frac{1}{\sqrt{(c_{left}^{l,m} \times G_{left} \times G_{EQ,left}^m)^2 + (c_{right}^{l,m} \times G_{right} \times G_{EQ,right}^m)^2}}$$

[0292] $c_{left}^{l,m}$ 和 $c_{right}^{l,m}$ 指示CPE信号的第1个时隙和第m个混合QMF频带的去量化线性CLD值, G_{left} 和 G_{right} 指示在ISO/IEC 23008-3表96中(即在格式转换规则表中)定义的输出声道的增益列的值,并且 $G_{EQ,left}^m$ 和 $G_{EQ,right}^m$ 指示在格式转换规则表中定义的输出声道的EQ值的第m个频带的增益。

[0293] 如果ICGPreAppliedCPE[n]为真,则在通过MPS块之前,第n个cplx_dmz应与ICG的逆相乘,并且剩余解码过程应遵循ISO/IEC 23008-3。

[0294] 当解码器使用IC处理块或编码器预处理ICG并且输出布局是立体布局时,对于CPE/QCE,不是经MPS上混的立体声/四声道信号,而是带限IC信号在SBR块之前的端生成。

[0295] 由于对于经MPS上混的立体声/四声道信号,SBR有效载荷已经通过立体声SBR被编码,所以立体声SBR有效载荷应当通过与参数域中的格式转换器的增益和EQ值相乘而被缩混,以便实现IC处理。

[0296] 现在将详细描述参数缩混立体声SBR的方法。

[0297] (1) 逆滤波

[0298] 通过允许立体声SBR参数在每个本底噪声带中具有最大值来选择逆滤波模式。

[0299] 其可以利用等式4获得:

[0300] 【等式4】

[0301] for (i=0; i<N_q; i++)

[0302] bs_invf_mode_{Downmixed}(i) = MAX (bs_invf_mode_{ch1}(i), bs_invf_mode_{ch2}(i))

$$[0303] \quad \begin{pmatrix} \text{ch1} \\ \text{ch2} \end{pmatrix} = \begin{cases} \text{在Cplx_out_dmz_L的情况下} & \begin{pmatrix} \text{CPE1的左声道} \\ \text{CPE2的左声道} \end{pmatrix} \\ \text{在Cplx_out_dmz_R的情况下} & \begin{pmatrix} \text{CPE1的右声道} \\ \text{CPE2的右声道} \end{pmatrix} \end{cases}$$

[0304] 2) 附加谐波

[0305] 包括基频f和基频f的奇次谐波3f、5f、7f、...的声波具有半波对称性。但是,包含基频f的偶次谐波0f、2f、...的声波不具有对称性。相反,引起声源波形变化而不是简单的缩放或移动的非线性系统产生额外的谐波,从而发生谐波失真。

[0306] 附加谐波是额外的正弦波的组合,并且可以如等式5表示:

[0307] 【等式5】

[0308] for (i=0; i<N_{High}; i++)

[0309] bs_add_harmonic_{Downmixed}(i) = OR (bs_add_harmonic_{ch1}(i), bs_add_harmonic_{ch2}(i))

[0310] (3) 包络时间边界

[0311] 图10a、图10b、图10c和图10d示出了根据本发明实施例的确定时间边界的方法，该时间边界是SBR参数。

[0312] 图10a示出了当第一包络的开始边界相同并且最后包络的停止边界相同时的时间包络网格。

[0313] 图10b示出了当第一包络的开始边界不同并且最后包络的停止边界相同时的时间包络格子。

[0314] 图10c示出了当第一包络的开始边界相同并且最后包络的停止边界不同时的时间包络格子。

[0315] 图10d示出了当第一包络的开始边界不同并且最后包络的停止边界不同时的时间包络格子。

[0316] 用于IC的SBR的时间包络格子 t_{E_Merged} 是通过将立体声SBR时间格分割成具有最高分辨率的最小片段而生成的。

[0317] 将 t_{E_Merged} 的开始边界值设置为立体声道的最大开始边界值。时间格子0和开始边界之间的包络已经在前一帧中处理过了。将两个声道的最后包络的停止边界中的具有最大值的停止边界，作为最后包络的停止边界。

[0318] 如在图10a-10d中所示，通过获得两个声道的时间边界之间的交点，确定第一个包络和最后一个包络的开始/停止边界具有最多分段分辨率。如果至少有5个包络，则反向搜索从 t_{E_Merged} 的停止点到 t_{E_Merged} 的开始点的点，找到不到4个包络，从而去除这不到4个包络的开始边界，以便减少包络的数量。这个过程一直持续到剩下5个包络。

[0319] (4) 噪声时间边界

[0320] 通过采用两个声道的噪声时间边界之中具有较大值的噪声时间边界，来确定缩混的噪声时间边界 L_{Q_Merged} 的数量。通过采用包络时间边界 t_{E_Merged} 的第一格和最后一格来确定第一格和合并的噪声时间边界 t_{Q_Merged} 。

[0321] 如果缩混的噪声时间边界 L_{Q_Merged} 大于1，则选择 $t_{Q_Merged}(1)$ ，作为其中噪声时间边界 L_Q 大于1的声道的 $t_Q(1)$ 。如果两个声道都具有大于1的噪声时间边界 L_Q ，则 $t_Q(1)$ 的最小值被选为 $t_{Q_Merged}(1)$ 。

[0322] (5) 包络数据

[0323] 选择合并的包络时间边界的频率分辨率 r_{Merged} 。频率分辨率 r_{Merged} 的每一部分的频率分辨率 r_{ch1} 和频率分辨率 r_{ch2} 之间的最大值被选择为 r_{Merged} 。

[0324]

[0325] 利用等式6，考虑格式转换参数来根据包络数据 E_{Orig} 计算所有包络的包络数据 E_{Orig_Merged} ：

[0326] 【等式6】

[0327]
$$E_{Orig_Merged}(k, l) = E_{ch1Orig}(g_{ch1}(k), h_{ch1}(l)) \times (EQ_{ch1}(k, h_{ch1}(l)))^2 + E_{ch2Orig}(g_{ch2}(k), h_{ch2}(l)) \times (EQ_{ch2}(k, h_{ch2}(l)))^2$$

[0328] 其中, $EQ_{ch1}(k, l) = \frac{\sum_m (G_{ch1}^m \times G_{EQ, ch1}^m)}{F(k+1, r_{Merged}(l)) - F(k, r_{Merged}(l))}$, $F(k, r_{Merged}(l)) \leq m < F(k+1, r_{Merged}(l))$,

[0329] $EQ_{ch2}(k, l) = \frac{\sum_m (G_{ch2}^m \times G_{EQ, ch2}^m)}{F(k+1, r_{Merged}(l)) - F(k, r_{Merged}(l))}$, $F(k, r_{Merged}(l)) \leq m < F(k+1, r_{Merged}(l))$

$, 0 \leq k < n(r_{Merged}(l)), 0 \leq l < L_{E_Merged}$,

[0330] $h_{ch1}(l)$ 定义为:

[0331] $t_{E_ch1}(h_{ch1}(l)) \leq t_{E_Merged}(l) < t_{E_ch1}(h_{ch1}(l)+1)$,

[0332] $h_{ch2}(l)$ 定义为:

[0333] $t_{E_ch2}(h_{ch2}(l)) \leq t_{E_Merged}(l) < t_{E_ch2}(h_{ch2}(l)+1)$,

[0334] $g_{ch1}(k)$ 定义为:

[0335] $F(g_{ch1}(k), r_{ch1}(h_{ch1}(l))) \leq F(k, r_{Merged}(l)) < F(g_{ch1}(k)+1, r_{ch1}(h_{ch1}(l)))$,

[0336] 并且 $g_{ch2}(k)$ 定义为:

[0337] $F(g_{ch2}(k), r_{ch2}(h_{ch2}(l))) \leq F(k, r_{Merged}(l)) < F(g_{ch2}(k)+1, r_{ch2}(h_{ch2}(l)))$ 。

[0338] (6) 本底噪声数据

[0339] 根据等式7, 合并的本底噪声数据被确定为两个声道数据的总和:

[0340] 【等式7】

[0341] $Q_{OrigMerged}(k, l) = Q_{Origch1}(k, h_{ch1}(l)) + Q_{Origch2}(k, h_{ch2}(l))$,

[0342] $0 \leq k < N_Q, 0 \leq l < L_{Q_Merged}$

[0343] 其中 $h_{ch1}(l)$ 定义为 $t_{Q_ch1}(h_{ch1}(l)) \leq t_{Q_Merged}(l) < t_{Q_ch1}(h_{ch1}(l)+1)$, 并且 $h_{ch2}(l)$ 定义为 $t_{Q_ch2}(h_{ch2}(l)) \leq t_{Q_Merged}(l) < t_{Q_ch2}(h_{ch2}(l)+1)$ 。

[0344] 本发明的上述实施例可以实现为可由各种计算机配置元件执行并且可以记录在计算机可读记录介质上的程序命令。计算机可读记录介质可以包括单独或组合的程序命令、数据文件、数据结构等等。要被记录在计算机可读记录介质上的程序命令可以为了本发明的实施例而被专门设计和配置, 或者可以是计算机软件领域的普通技术人员熟知和可用的。计算机可读记录介质的示例包括磁性介质(例如, 硬盘、软磁盘或磁带)、光学介质(例如, 光盘只读存储器(CD-ROM)或者数字多功能盘(DVD)、磁光介质(例如, 软光盘)以及专门配置为存储和执行程序命令的硬件设备(例如, ROM、随机存取存储器(RAM)或闪存存储器等)。计算机程序的示例包括可以由计算机使用解释器等执行的高级语言代码, 以及由编译器制作的机器语言代码。硬件设备可以被配置用作一个或多个软件模块以执行本发明的操作, 或者反之亦然。

[0345] 虽然已经参照本发明的示例性实施例具体示出和描述了本发明, 但是可以理解, 在不脱离所附权利要求的精神和范围的情况下, 可以进行各种形式和细节上的改变。

[0346] 因此, 本发明的范围不是由详细描述而是由所附权利要求限定的, 并且该范围内的所有差异将被解释为包括在本发明中。

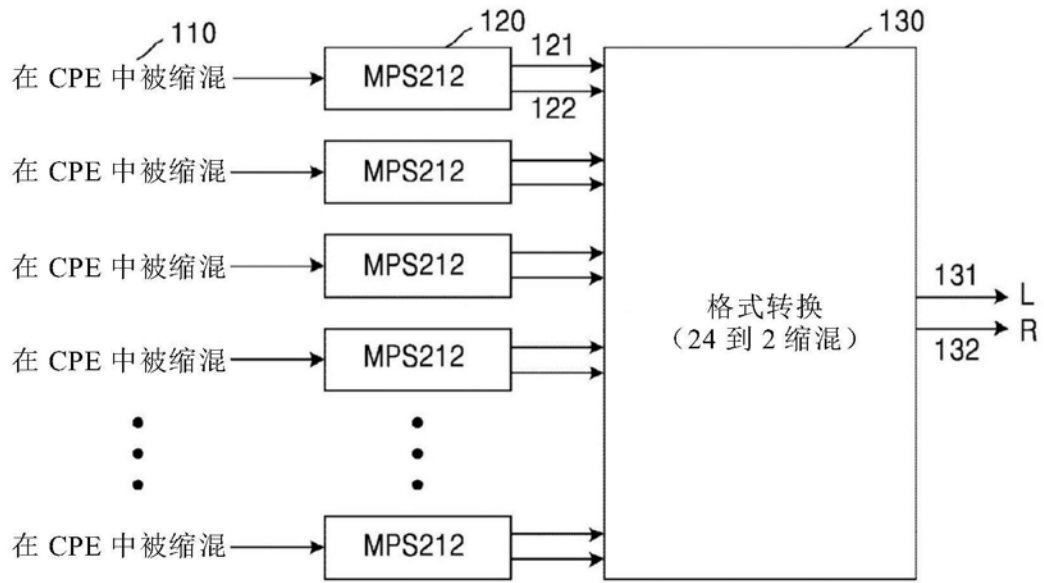


图1

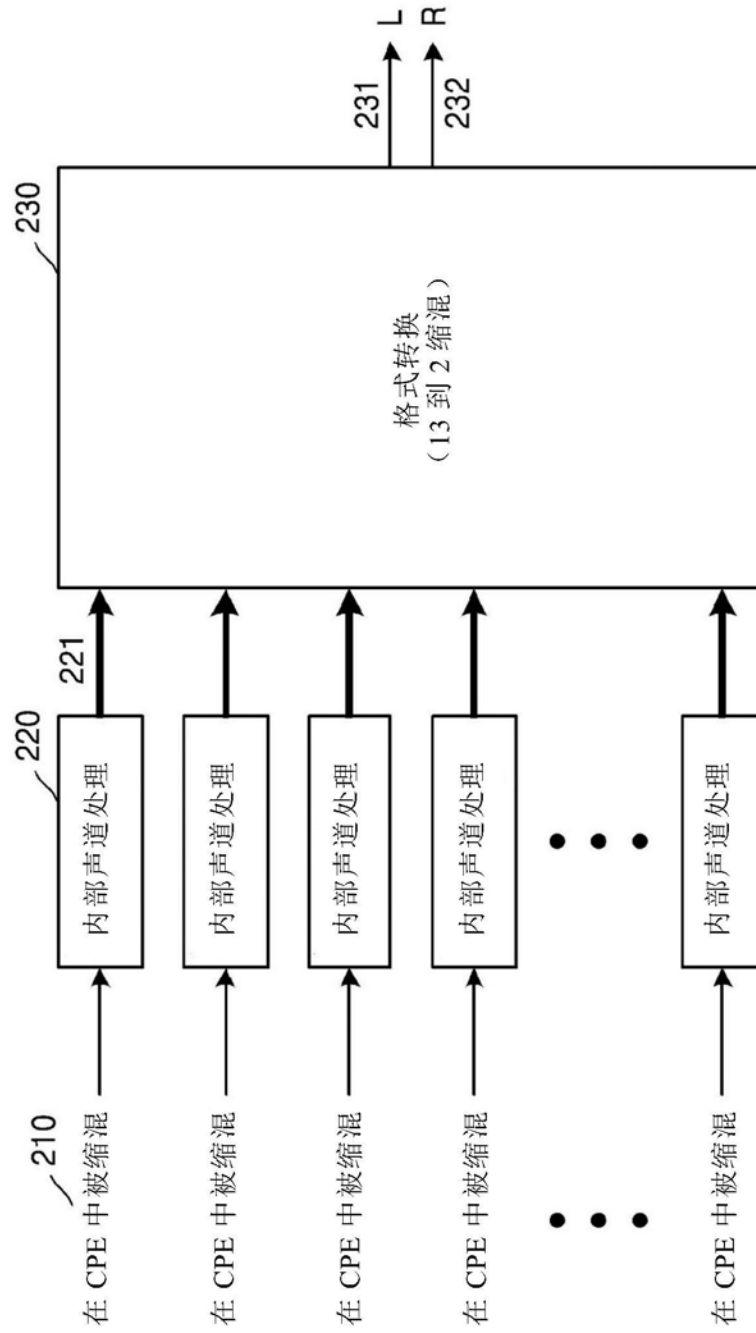


图2

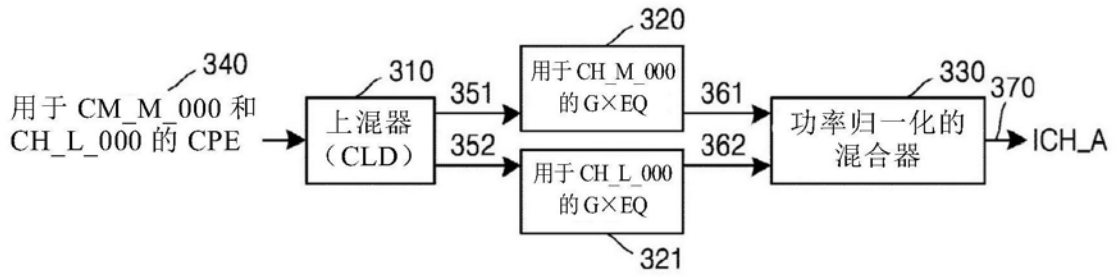


图3

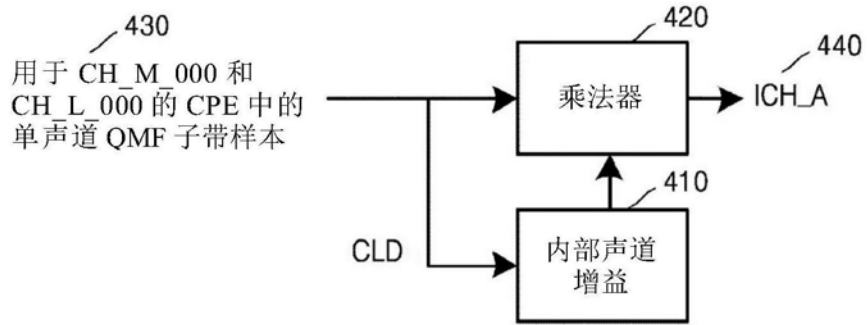


图4

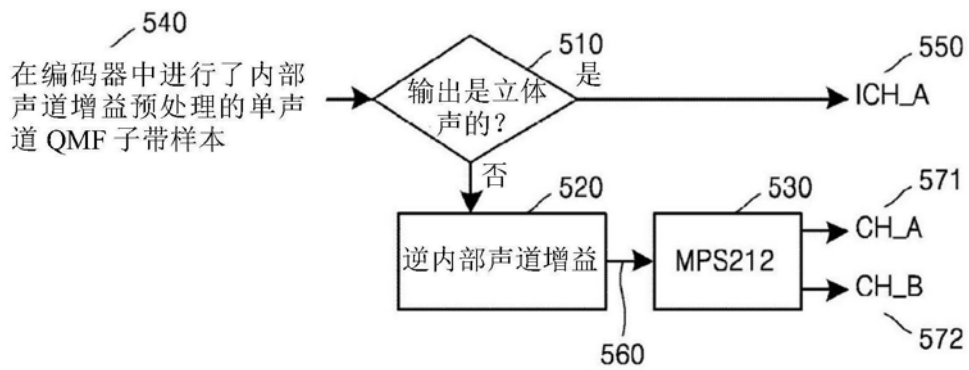


图5

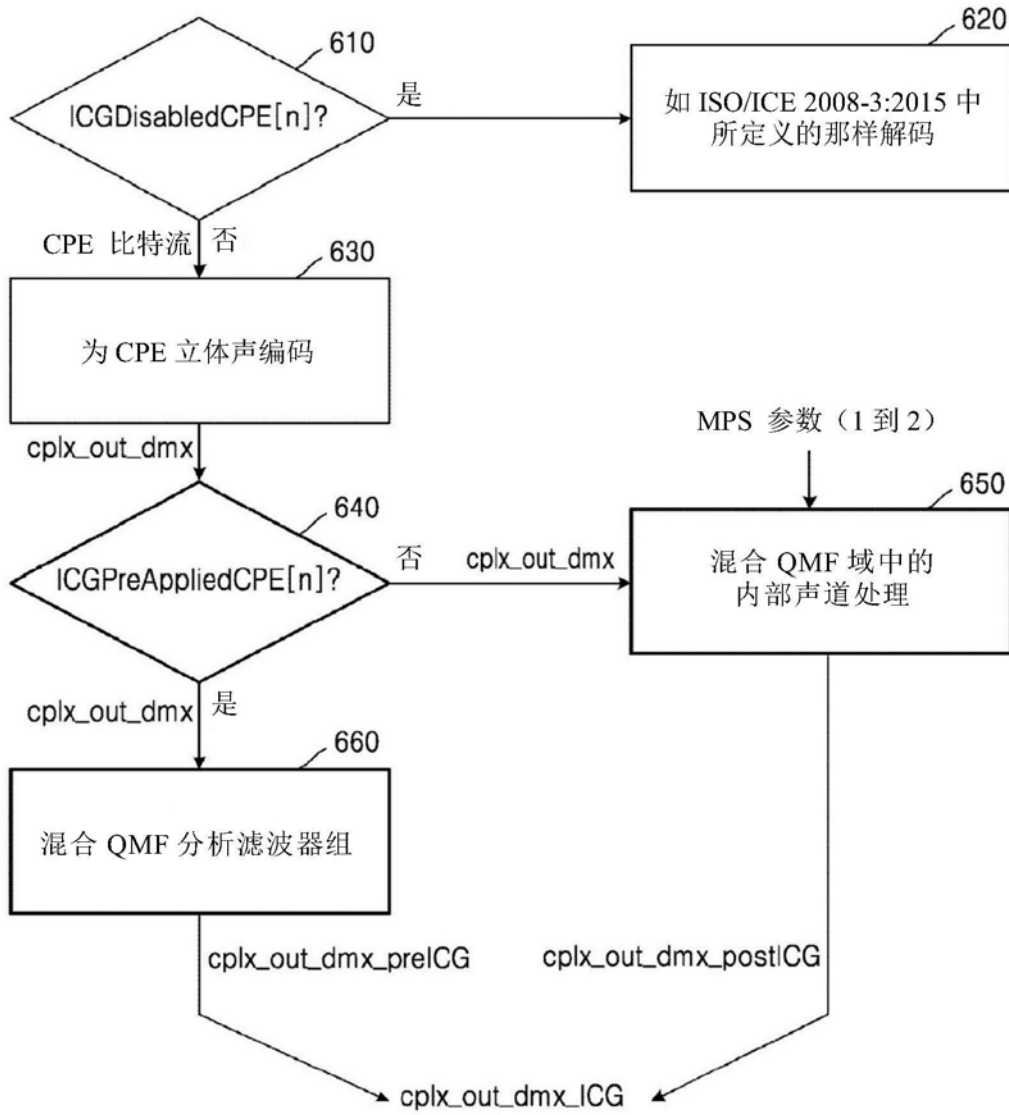


图6

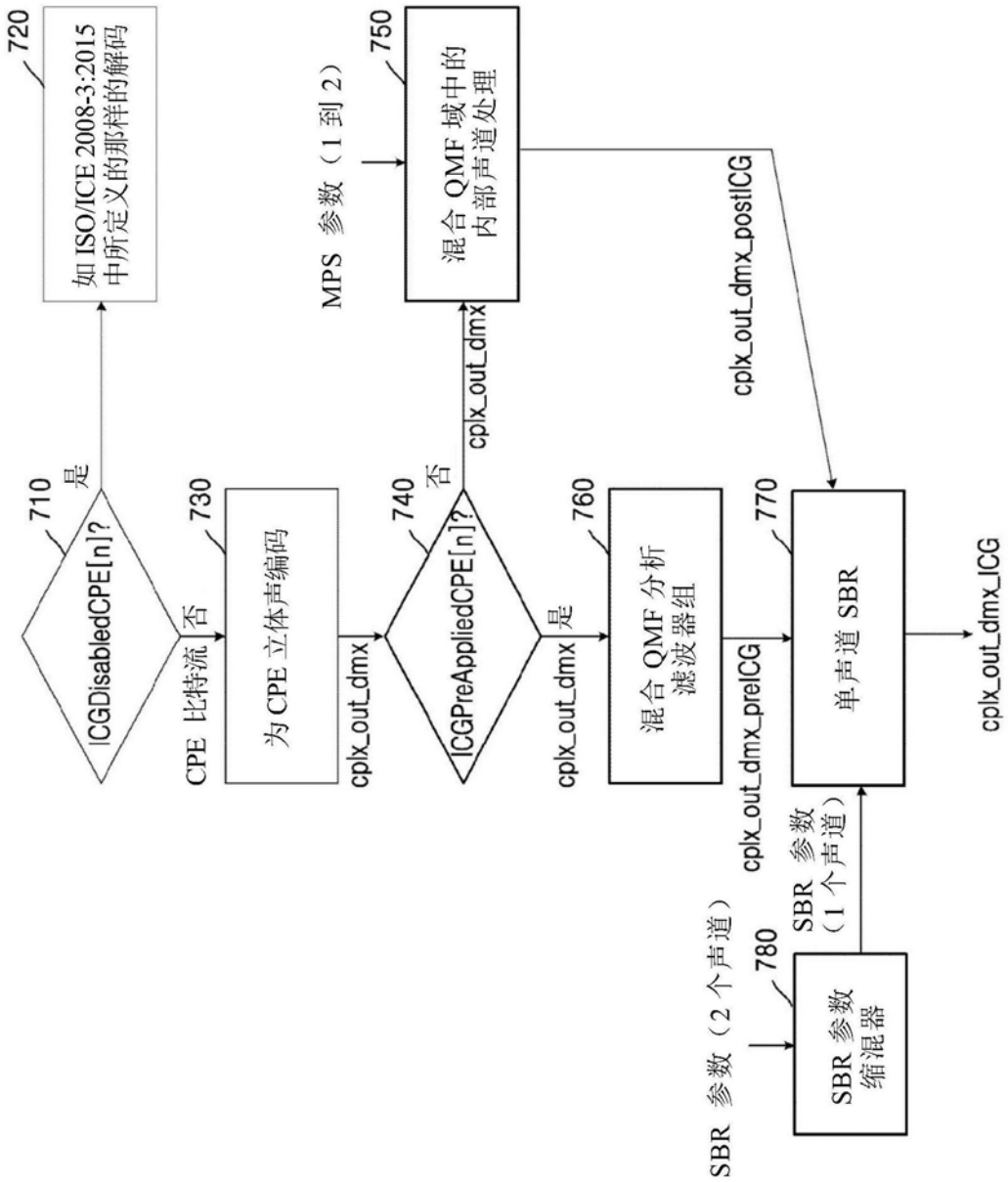


图7

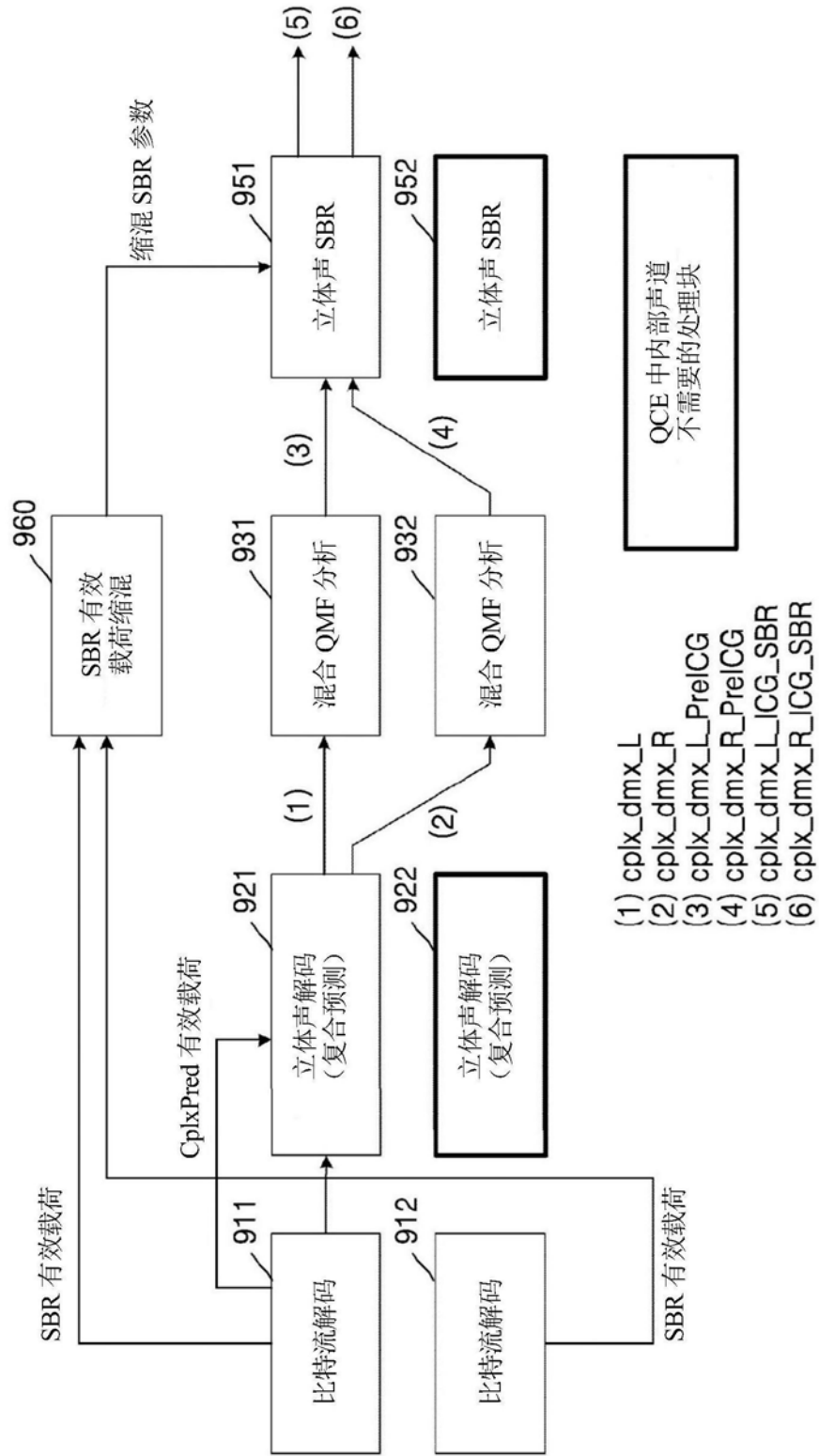


图9

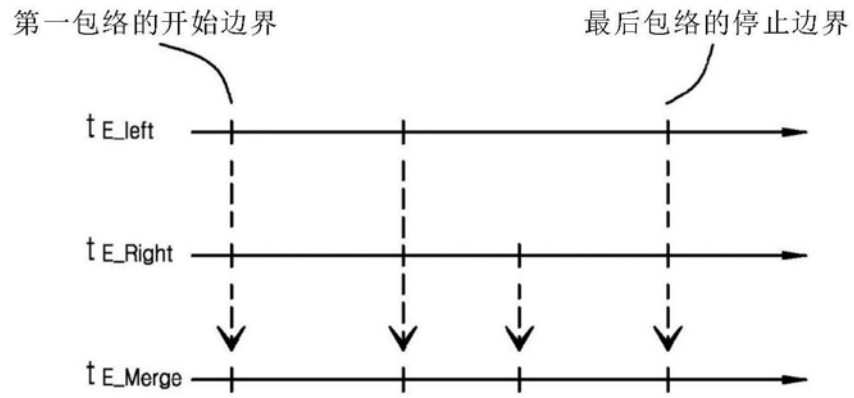


图10a

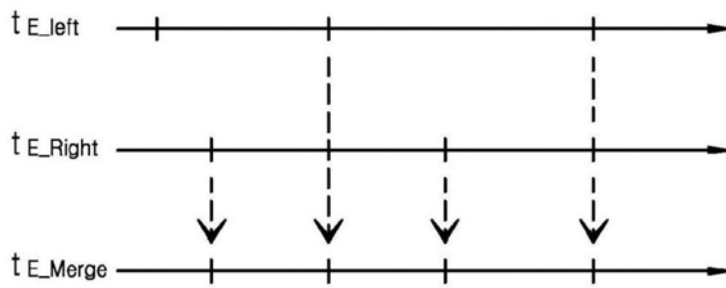


图10b

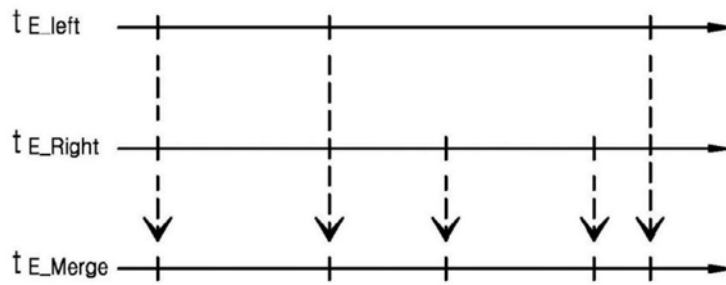


图10c

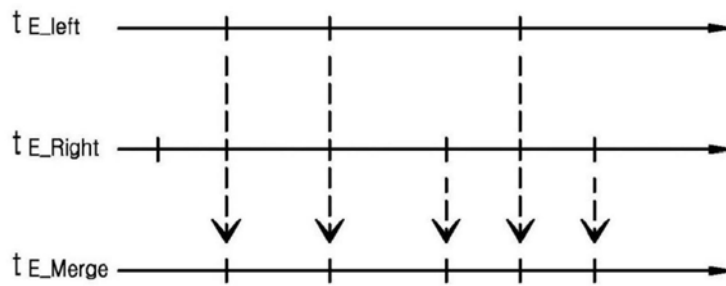


图10d