



(12) 发明专利

(10) 授权公告号 CN 107430863 B

(45) 授权公告日 2021.01.26

(21) 申请号 201680014669.3

本杰明·舒伯特 拉尔夫·盖革

(22) 申请日 2016.03.07

(74) 专利代理机构 北京清亦华知识产权代理事

(65) 同一申请的已公布的文献号
申请公布号 CN 107430863 A

务所(普通合伙) 11201

代理人 宋融冰

(43) 申请公布日 2017.12.01

(51) Int.Cl.

(30) 优先权数据

G10L 19/008 (2013.01)

15158233.5 2015.03.09 EP

G10L 19/02 (2013.01)

15172594.2 2015.06.17 EP

G10L 19/04 (2013.01)

G10L 19/18 (2013.01)

G10L 21/038 (2013.01)

(85) PCT国际申请进入国家阶段日
2017.09.08

(56) 对比文件

(86) PCT国际申请的申请数据

CN 102388417 A, 2012.03.21

PCT/EP2016/054776 2016.03.07

CN 102741831 A, 2012.10.17

(87) PCT国际申请的公布数据

W02016/142337 EN 2016.09.15

CN 101925950 A, 2010.12.22

CN 1437747 A, 2003.08.20

CN 102859589 A, 2013.01.02

(73) 专利权人 弗劳恩霍夫应用研究促进协会
地址 德国慕尼黑

CN 101253557 A, 2008.08.27

CN 102177426 A, 2011.09.07

EP 0797324 A2, 1997.09.24

WO 2013156814 A1, 2013.10.24

3GPP. "3rd Generation Partnership Project". 《3GPP standard》. 2014,

(72) 发明人 萨沙·迪施 纪尧姆·福克斯
伊曼纽尔·拉韦利
克里斯蒂安·诺伊卡姆
康斯坦丁·施密特
康拉德·本多尔夫
安德烈·尼德迈尔

审查员 董小东

权利要求书6页 说明书25页 附图21页

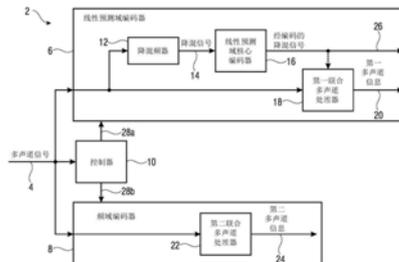
(54) 发明名称

用于编码的音频编码器及用于解码的音频解码器

用于编码降混信号的线性预测域核心编码器(16),此外,线性预测域编码器包括用于从多声道信号(4)生成第一多声道信息(20)的第一联合多声道编码器(18)。

(57) 摘要

展示用于编码多声道音频信号(4)的音频编码器(2)的示意性框图。音频编码器包括线性预测域编码器(6)、频域编码器(8)以及用于在线性预测域编码器(6)与频域编码器(8)之间切换的控制器(10)。控制器被配置为使得多声道信号的部分由线性预测域编码器的编码帧表示或由频域编码器的编码帧表示。线性预测域编码器包括用于对多声道信号(4)进行降混以获得降混信号(14)的降混滤波器(12)。线性预测域编码器还包括



CN 107430863 B

1. 一种用于编码多声道信号(4)的音频编码器(2),包括:

线性预测域编码器(6);

频域编码器(8);

用于在所述线性预测域编码器(6)与所述频域编码器(8)之间切换的控制器(10),

其中所述线性预测域编码器(6)包括用于对所述多声道信号(4)进行降混以获得降混信号(14)的降混频器(12)、用于编码所述降混信号(14)的线性预测域核心编码器(16)以及用于从所述多声道信号(4)生成第一多声道信息(20)的第一联合多声道编码器(18),

其中所述频域编码器(8)包括用于从所述多声道信号(4)生成第二多声道信息(24)的第二联合多声道编码器(22),其中所述第二联合多声道编码器(22)不同于所述第一联合多声道编码器(18),以及

其中所述控制器(10)被配置为执行切换使得所述多声道信号(4)的部分由所述线性预测域编码器(6)的编码帧表示或由所述频域编码器(8)的编码帧表示;

其中所述线性预测域核心编码器(16)包括ACELP处理器(30)、TCX处理器(32)以及时域带宽扩展处理器(36),其中所述ACELP处理器用于对通过降取样获得的经降取样的降混信号(34)进行操作,且其中所述时域带宽扩展处理器(36)用于对通过降取样从ACELP输入信号中移除的所述降混信号的部分的频带进行参数化编码,以及其中所述TCX处理器(32)用于对未被降取样或以小于由所述ACELP处理器(30)操作的用于生成经降取样的降混信号(34)的降取样的程度被降取样的降混信号(14)进行操作,所述TCX处理器包括第一时间-频率转换器(40)、用于生成第一频带集合的参数化表示(46)的第一参数生成器(42)以及用于生成用于第二频带集合的经量化的编码的频谱线的集合(48)的第一量化器编码器(44);

或

其中所述音频编码器还包括用于对由所述线性预测域核心编码器(16)输出的经编码的降混信号(14)进行解码以获得经编码且经解码的降混信号(54)的线性预测域解码器(50);以及用于使用经编码且经解码的降混信号(54)计算和编码多声道残余信号(58)的多声道残余编码器(56),所述多声道残余信号(58)表示使用第一多声道信息(20)获得的经解码的多声道表示与降混之前的多声道信号(4)之间的误差;

或

其中所述控制器(10)用于在所述多声道信号的当前帧(204)内从使用所述频域编码器(8)对先前帧进行编码切换至使用所述线性预测域编码器(6)对即将到来的帧(206)进行编码;其中第一联合多声道编码器(18)用于从用于所述当前帧的多声道信号计算合成多声道参数(210a,210b,212a,212b)作为所述第一多声道信息(20);以及其中所述第二联合多声道编码器(22)用于使用停止窗口(200a,200b)对所述多声道信号(4)进行加权以获得所述当前帧(204)中的经加权的多声道信号,使得其中执行从所述频域编码器(8)至所述线性预测域编码器(6)的转换的所述当前帧(204)包括所述经加权的多声道信号、关于所述降混信号(14)的信息(208)以及所述合成多声道参数(210a,210b,212a,212b)。

2. 根据权利要求1所述的音频编码器(2),其中所述第一联合多声道编码器(18)包括第一时间-频率转换器(82),其中所述第二联合多声道编码器(22)包括第二时间-频率转换器(66),且其中所述第一时间-频率转换器(82)和所述第二时间-频率转换器(66)彼此不同。

3. 根据权利要求1所述的音频编码器(2),其中所述第一联合多声道编码器(18)为参数

化联合多声道编码器;或

其中所述第二联合多声道编码器(22)为波形保持联合多声道编码器。

4.根据权利要求3所述的音频编码器,

其中所述参数化联合多声道编码器包括立体声产生编码器、参数化立体声编码器或基于旋转的参数化立体声编码器,或

其中所述波形保持联合多声道编码器包括频带选择性切换中间/侧或左/右立体声编码器。

5.根据权利要求1所述的音频编码器(2),其中所述频域编码器(8)包括用于将所述多声道信号(4)的第一声道(4a)及所述多声道信号(4)的第二声道(4b)转换成频谱表示(72a,b)的第二时间-频率转换器(66)、用于生成第二频带集合的参数化表示的第二参数生成器(68)以及用于生成第一频带集合的经量化且经编码的表示(80)的第二量化器编码器(70)。

6.根据权利要求1所述的音频编码器(2),

其中所述TCX处理器包括MDCT操作和智能间隙填充功能,或

其中所述频域编码器(8)包括用于所述多声道信号(4)的第一声道(4a)和第二声道(4b)的MDCT操作和AAC操作以及智能间隙填充功能,或

其中所述第一联合多声道编码器(18)用于以得出用于多声道信号(4)的全带宽的多声道信息的方式进行操作。

7.根据权利要求1所述的音频编码器(2),

其中所述降混信号具有低频带和高频带,其中所述线性预测域编码器(6)用于施加用于对高频带进行参数化编码的带宽扩展处理,其中所述线性预测域解码器用于仅获得表示所述降混信号的低频带的低频带信号以作为经编码且经解码的降混信号(54),且其中经编码的多声道残余信号(58)仅具有在降混之前的多声道信号(4)的低频带内的频率成分。

8.根据权利要求1所述的音频编码器(2),

其中所述多声道残余编码器(56)包括:

用于使用所述第一多声道信息(20)及所述经编码且经解码的降混信号(54)生成经解码的多声道信号(64)的联合多声道解码器(60);以及

用于形成所述经解码的多声道信号(64)与降混之前的多声道信号(4)之间的差异以获得多声道残余信号(58)的差异处理器(62)。

9.根据权利要求1所述的音频编码器(2),

其中所述降混频器(12)用于将所述多声道信号(4)转换成频谱表示,且其中使用频谱表示或使用时域表示来执行降混,且

其中所述第一联合多声道编码器(18)用于使用频谱表示生成用于频谱表示的各个频带的单独第一多声道信息。

10.一种用于对经编码的音频信号(103)进行解码的音频解码器(102),包括:

线性预测域解码器(104);

频域解码器(106);

用于使用所述线性预测域解码器(104)的输出和使用第一多声道信息(20)生成第一多声道表示(114)的第一联合多声道解码器(108);

用于使用所述频域解码器(106)的输出和第二多声道信息(22,24)生成第二多声道表

示(116)的第二联合多声道解码器(110);以及

用于组合所述第一多声道表示(114)和所述第二多声道表示(116)以获得经解码的音频信号(118)的第一组合器(112),

其中所述第二联合多声道解码器(110)不同于所述第一联合多声道解码器(108);

其中所述线性预测域解码器(104)包括:ACELP解码器(120)、低频带合成器(122)、用于对由所述低频带合成器(122)生成的信号进行升取样的升取样器(124)、时域带宽扩展处理器(126)以及用于组合由所述升取样器(124)生成的经升取样的信号和由所述时域带宽扩展处理器(126)生成的经带宽扩展的信号的信号的第二组合器(128);

或

其中所述经编码的音频信号(3)包括用于由所述线性预测域解码器(104)输出的经解码的降混信号(142)的多声道残余信号(58),以及其中所述第一联合多声道解码器(108)用于除所述经解码的降混信号(142)和所述第一多声道信息(20)之外还使用所述多声道残余信号(58)来生成所述第一多声道表示(114);

或

其中所述音频解码器(102)用于在多声道音频信号的当前帧(204)内从使用所述频域解码器(106)对先前帧进行解码切换至使用所述线性预测域解码器(104)对即将到来的帧(206)进行解码;其中所述组合器(112)用于从当前帧(204)的第二多声道表示(116)计算合成中间信号(226);其中所述第一联合多声道解码器(108)用于使用所述合成中间信号(226)和所述第一多声道信息(20)生成所述第一多声道表示(114);以及其中所述组合器(112)用于组合所述第一多声道表示(114)和所述第二多声道表示(116)以获得所述经解码的音频信号(118)的当前帧(204);

或

其中所述音频解码器(102)用于在多声道音频信号的当前帧(232)内从使用所述线性预测域解码器(104)对先前帧进行解码切换至使用所述频域解码器(106)对即将到来的帧进行解码;其中所述第一联合多声道解码器(108)包括立体声解码器(146),所述立体声解码器(146)用于使用先前帧的第一多声道信息(20)从用于当前帧的所述线性预测域解码器(104)生成的经解码的单声道信号计算合成多声道音频信号(250a,250b);其中所述第二联合多声道解码器(110)用于计算用于所述当前帧(232)的第二多声道表示(116)并使用开始窗口对所述第二多声道表示(116)进行加权以获得经加权的第二多声道表示;以及其中所述组合器(112)用于组合所述合成多声道音频信号(250a,250b)和所述经加权的第二多声道表示以获得所述经解码的音频信号的当前帧(232)。

11.根据权利要求10所述的音频解码器(102),

其中所述第一联合多声道解码器(108)为参数化联合多声道解码器,且其中所述第二联合多声道解码器(110)为波形保持联合多声道解码器,

其中所述第一联合多声道解码器(108)用于基于复数预测、参数化立体声操作或旋转操作而进行操作,以及

其中所述第二联合多声道解码器(110)用于将频带选择性切换应用于中间/侧或左/右立体声解码算法。

12.根据权利要求10所述的音频解码器(102),其中所述线性预测域解码器(104)包括:

TCX解码器(130)和智能间隙填充IGF处理器(132)；

用于组合所述TCX解码器(130)的第二组合器(128)和IGF处理器(132)的输出的全频带合成处理器(134)，以及

用以使用通过低频带频谱-时间转换从所述TCX解码器(130)及所述IGF处理器(132)生成的信号得出的信息来初始化所述低频带合成器(122)的交叉路径(136)。

13. 根据权利要求10所述的音频解码器(102)，

其中所述第一联合多声道解码器(108)包括：用于将所述线性预测域解码器(104)的输出转换成频谱表示(145)的时间-频率转换器(144)；

受所述第一多声道信息(20)控制而对所述频谱表示(145)进行操作的升混频器；以及

用于将升混结果转换成与所述第一多声道表示(114)对应的时间表示的频率-时间转换器(148)。

14. 根据权利要求10所述的音频解码器(102)，

其中所述第二联合多声道解码器(110)用于：

使用由所述频域解码器(106)获得的频谱表示作为输入，所述频谱表示包括至少针对多个频带的第一声道信号和第二声道信号；以及

对所述第一声道信号和所述第二声道信号的多个频带施加联合多声道操作并将所述联合多声道操作的结果转换成时间表示以获得所述第二多声道表示(116)。

15. 根据权利要求14所述的音频解码器(102)，其中所述第二多声道信息(24)是为各个频带指示左/右或中间/侧联合多声道编码的掩码，且其中联合多声道操作是用于将由所述掩码所指示的频带从中间/侧表示转换为左/右表示的中间/侧至左/右转换操作。

16. 根据权利要求13所述的音频解码器(102)，其中所述多声道残余信号(58)具有低于所述第一多声道表示(114)的带宽，且其中所述第一联合多声道解码器(108)用于使用第一多声道信息(20)和所述经解码的降混信号(142)重建中间第一多声道表示(114)并将所述多声道残余信号(58)添加至所述中间第一多声道表示。

17. 根据权利要求13所述的音频解码器(102)，

其中所述时间-频率转换器(144)包括复数运算或过取样的操作，且

其中所述频域解码器(106)包括IMDCT操作(152)或临界取样的操作。

18. 根据权利要求13所述的音频解码器(102)或根据权利要求1所述的音频编码器(2)，其中所述多声道指的是两个或更多个声道。

19. 一种对多声道信号(4)进行编码的方法(800)，包括：

执行线性预测域编码；

执行频域编码；

在所述线性预测域编码与所述频域编码之间切换，

其中所述线性预测域编码包括：对所述多声道信号进行降混以获得降混信号、对所述降混信号进行线性预测域核心编码以及从所述多声道信号生成第一多声道信息的第一联合多声道编码，

其中所述频域编码包括从所述多声道信号生成第二多声道信息的第二联合多声道编码，其中所述第二联合多声道编码不同于所述第一多声道编码，且

其中切换被执行以使得所述多声道信号的部分由所述线性预测域编码的编码帧表示

或由所述频域编码的编码帧表示；

其中所述线性预测域核心编码包括ACELP处理、TCX处理以及时域带宽扩展处理，其中所述ACELP处理用于对通过降取样获得的经降取样的降混信号(34)进行操作，且其中所述时域带宽扩展处理用于对通过降取样从ACELP输入信号中移除的所述降混信号的部分的频带进行参数化编码，以及其中所述TCX处理用于对未被降取样或以小于由所述ACELP处理操作的用于生成经降取样的降混信号(34)的降取样的程度被降取样的降混信号(14)进行操作，所述TCX处理包括第一时间-频率转换、生成第一频带集合的参数化表示(46)以及生成用于第二频带集合的经量化的编码的频谱线的集合(48)；

或

其中所述音频编码的方法还包括线性预测域解码，包括对由所述线性预测域核心编码输出的经编码的降混信号(14)进行解码以获得经编码且经解码的降混信号(54)；以及多声道残余编码，包括使用经编码且经解码的降混信号(54)计算和编码多声道残余信号(58)，所述多声道残余信号(58)表示使用第一多声道信息(20)获得的经解码的多声道表示与降混之前的多声道信号(4)之间的误差；

或

其中所述切换包括在所述多声道信号的当前帧(204)内从使用所述频域编码(8)对先前帧进行编码切换至使用所述线性预测域编码(6)对即将到来的帧(206)进行编码；其中第一联合多声道编码包括从用于所述当前帧的多声道信号计算合成多声道参数(210a,210b,212a,212b)作为所述第一多声道信息(20)；以及其中所述第二联合多声道编码包括使用停止窗口(200a,200b)对所述多声道信号(4)进行加权以获得所述当前帧(204)中的经加权的多声道信号，使得其中执行从所述频域编码(8)至所述线性预测域编码(6)的转换的所述当前帧(204)包括所述经加权的多声道信号、关于所述降混信号(14)的信息(208)以及所述合成多声道参数(210a,210b,212a,212b)。

20.一种对经编码的音频信号进行解码的方法(900)，包括：

线性预测域解码；

频域解码；

使用所述线性预测域解码的输出和使用第一多声道信息生成第一多声道表示的第一联合多声道解码；

使用所述频域解码的输出和第二多声道信息生成第二多声道表示的第二多声道解码；
以及

组合所述第一多声道表示和所述第二多声道表示以获得经解码的音频信号，

其中所述第二多声道解码不同于所述第一多声道解码；

其中所述线性预测域解码包括：ACELP解码、低频带合成、包括对由所述低频带合成生成的信号进行升取样的升取样、时域带宽扩展处理以及包括组合通过所述升取样生成的经升取样的信号和通过所述时域带宽扩展处理生成的经带宽扩展的信号的第二组合；

或

其中所述经编码的音频信号(103)包括用于通过所述线性预测域解码输出的经解码的降混信号(142)的多声道残余信号，以及其中所述第一联合多声道解码用于除所述经解码的降混信号(142)和所述第一多声道信息(20)之外还使用所述多声道残余信号(58)来生成

所述第一多声道表示 (114) ;

或

其中所述解码的方法包括在多声道音频信号的当前帧 (204) 内从使用所述频域解码对先前帧进行解码切换至使用所述线性预测域解码对即将到来的帧 (206) 进行解码;其中所述组合包括从当前帧 (204) 的第二多声道表示 (116) 计算合成中间信号 (226);其中所述第一联合多声道解码包括使用所述合成中间信号 (226) 和所述第一多声道信息 (20) 生成所述第一多声道表示 (114);以及其中所述组合包括组合所述第一多声道表示 (114) 和所述第二多声道表示 (116) 以获得所述经解码的音频信号的当前帧 (204);

或

其中所述解码的方法包括在多声道音频信号的当前帧 (232) 内从使用所述线性预测域解码对先前帧进行解码切换至使用所述频域解码对即将到来的帧进行解码;其中所述第一联合多声道解码包括立体声解码,所述立体声解码包括使用先前帧的第一多声道信息 (20) 从用于当前帧的所述线性预测域解码生成的经解码的单声道信号计算合成多声道音频信号 (250a, 250b);其中所述第二联合多声道解码包括计算用于所述当前帧 (232) 的第二多声道表示 (116) 并使用开始窗口对所述第二多声道表示 (116) 进行加权以获得经加权的第二多声道表示;以及其中所述组合包括组合所述合成多声道音频信号 (250a, 250b) 和所述经加权的第二多声道表示以获得所述经解码的音频信号的当前帧 (232)。

21. 一种存储介质,具有存储于其上的计算机程序,当计算机程序在计算机或处理器上运行时用于执行根据权利要求19或权利要求20所述的方法。

用于编码的音频编码器及用于解码的音频解码器

技术领域

[0001] 本发明涉及一种用于编码多声道音频信号的音频编码器及用于解码经编码的音频信号的音频解码器。实施例涉及包括波形保持及参数化立体声编码的切换式感知音频编解码器。

背景技术

[0002] 音频信号的感知编码出于用于此等信号的高效存储或传输的数据缩减的目的而被广泛实际应用。特别地,当将达到最高效率时,使用紧密适合于信号输入特性的编解码器。一个示例为MPEG-D USAC核心编解码器,其可用于主要对语音信号使用代数码本激励线性预测(ACELP, Algebraic Code-Excited Linear Prediction)编码、对背景噪声及混合信号使用变换编码激励(TCX, Transform Coded Excitation)以及对音乐内容使用高级音频编码(AAC, Advanced Audio Coding)。所有的三个内部编解码器配置可响应于信号内容以信号自适应方式被立即切换。

[0003] 此外,使用联合多声道编码技术(中间/侧编码等)或为了最高效率而使用参数化编码技术。参数化编码技术基本上以感知等同音频信号的再造而非给定波形的忠实重建为目标。示例包括噪声填充、带宽扩展以及空间音频编码。

[0004] 在现有技术水平的编解码器中,当将信号自适应核心编码器与联合多声道编码或参数化编码技术进行组合时,核心编解码器被切换以匹配信号特性,但多声道编码技术(如,M/S立体声、空间音频编码或参数化立体声)的选择保持固定且独立于信号特性。这些技术通常被用于核心编解码器以作为核心编码器的预处理器及核心解码器的后处理器,这两种处理器不知道核心编解码器的实际选择。

[0005] 另一方面,用于带宽扩展的参数化编码技术的选择有时是信号相依地做出的。举例而言,应用于时域中的技术对于语音信号更有效率,而频域处理对于其他信号更相关。在此情况下,所采用的多声道编码技术必须与两种带宽扩展技术兼容。

[0006] 现有技术水平中的相关话题包括:

[0007] 作为MPEG-D USAC核心编解码器的预处理器/后处理器的PS及MPS

[0008] MPEG-D USAC标准

[0009] MPEG-H 3D音频标准

[0010] 在MPEG-D USAC中,描述了可切换核心编码器。然而,在USAC中,多声道编码技术被定义为整个核心编码器常见的固定选择,与其编码原理的内部切换为ACELP或TCX(“LPD”)或AAC(“FD”)无关。因此,若期望切换式核心编解码器配置,编解码器被限制为针对整个信号始终使用参数化多声道编码(parametric multichannel coding,PS)。然而,为了编码(例如)音乐信号,使用联合立体声编码将更恰当,其可每频带及每帧地在L/R(左/右)与M/S(中间/侧)方案之间动态地切换。

[0011] 因此,需要经改良的方法。

发明内容

[0012] 本发明的目标为提供用于处理音频信号的经改良的概念。通过独立权利要求的主题实现此目标。

[0013] 本发明基于如下发现：使用多声道编码器的(时域)参数化编码器对参数化多声道音频编码是有利的。多声道编码器可以是多声道残余编码器，其与用于每个声道的单独编码相比可减小用于编码参数的传输的带宽。(例如)结合频域联合多声道音频编码器此可被有利地使用。时域及频域联合多声道编码技术可被组合，以使得(例如)基于帧的决策可将当前帧引导至基于时间或基于频率的编码周期。换言之，实施例展示用于将使用联合多声道编码及参数化空间音频编码的可切换核心编解码器组合成完全可切换感知编解码器的经改良的概念，完全可切换感知编解码器允许依据核心编码器的选择而使用不同的多声道编码技术。此概念是有利的，因为与已经存在的方法相比，实施例展示可与核心编码器一起被立即切换且因此紧密匹配于且适合于核心编码器的选择的多声道编码技术。因此，可避免由于多声道编码技术的固定选择而出现的所描绘的问题。此外，实现给定核心编码器与其相关联且经调适的多声道编码技术的完全可切换组合。举例而言，此种编码器(例如，使用L/R或M/S立体声编码的AAC(高级音频编码))能够使用专用联合立体声或多声道编码(例如，M/S立体声)在频域(FD)核心编码器中对音乐信号进行编码。此决策可单独地应用于每个音频帧中的每个频带。在(例如)语音信号的情况下，核心编码器可立即切换至线性预测性解码(linear predictive decoding, LPD)核心编码器及其相关联的不同技术(例如，参数化立体声编码技术)。

[0014] 实施例展示对于单声道LPD路径而言是唯一的立体声处理，及将立体声FD路径的输出与来自LPD核心编码器及其专用立体声编码的输出进行组合的基于立体声信号的无缝切换方案。此情况是有利的，因为实现了无伪声(artifact)的无缝编解码器切换。

[0015] 实施例涉及一种用于编码多声道信号的编码器。编码器包括线性预测域编码器及频域编码器。此外，编码器包括用于在线性预测域编码器与频域编码器之间进行切换的控制器。此外，线性预测域编码器可包括：用于对多声道信号进行降混以获得降混信号的降混频率器；用于编码降混信号的线性预测域核心编码器；以及用于从多声道信号生成第一多声道信息的第一多声道编码器。频域编码器包括用于从多声道信号生成第二多声道信息的第二联合多声道编码器，其中第二多声道编码器不同于第一多声道编码器。控制器被配置为使得多声道信号的部分由线性预测域编码器的编码帧表示或由频域编码器的编码帧表示。线性预测域编码器可包括ACELP核心编码器及(例如)作为第一联合多声道编码器的参数化立体声编码算法。频域编码器可包括(例如)作为第二联合多声道编码器的使用(例如)L/R或M/S处理的AAC核心编码器作为第二联合多声道编码器。控制器可关于(例如)帧特性而分析多声道信号(例如，语音或音乐)，且用于针对每帧或帧的序列或多声道音频信号的部分，决定是线性预测域编码器还是频域编码器应被用于编码多声道音频信号的此部分。

[0016] 实施例进一步展示一种用于解码经编码的音频信号的音频解码器。音频解码器包括线性预测域解码器及频域解码器。此外，音频解码器包括：用于使用线性预测域解码器的输出及使用多声道信息生成第一多声道表示的第一联合多声道解码器；以及用于使用频域解码器的输出及第二多声道信息生成第二多声道表示的第二多声道解码器。此外，音频解码器包括用于将第一多声道表示及第二多声道表示进行组合以获得经解码的音频信号的

第一组合器。组合器可在作为(例如)经线性预测的多声道音频信号的第一多声道表示与作为(例如)频域经解码的多声道音频信号的第二多声道表示之间执行无缝无伪声切换。

[0017] 实施例展示可切换音频编码器内的LPD路径中的ACELP/TCX编码与频域路径中的专用立体声编码及独立AAC立体声编码的组合。此外,实施例展示LPD与FD立体声之间的无缝瞬时切换,其中其他实施例涉及用于不同信号内容类型的联合多声道编码的独立选择。举例而言,针对主要使用LPD路径编码的语音,使用参数化立体声,而对于在FD路径中被编码的音乐,使用更加自适应的立体声编码,其可每频带及每帧地在L/R方案与M/S方案之间动态地切换。

[0018] 根据实施例,针对主要使用LPD路径来编码且通常位于立体声影像的中心的语音,简单的参数化立体声是恰当的,而在FD路径中被编码的音乐通常具有更复杂的空间分布且可利用更加自适应的立体声编码,该更加自适应的立体声编码可每频带及每帧地在L/R方案与M/S方案之间动态地切换。

[0019] 其他实施例展示音频编码器,该音频编码器包括:用于对多声道信号进行降混以获得降混信号的降混频器(12);用于编码降混信号的线性预测域核心编码器;用于生成多声道信号的频谱表示的滤波器组;以及用于从多声道信号生成多声道信息的联合多声道编码器。降混信号具有低频带和高频带,其中线性预测域核心编码器用于施加带宽扩展处理以用于对高频带进行参数化编码。此外,多声道编码器用于处理包括多声道信号的低频带及高频带的频谱表示。此是有利的,因为每个参数化编码可使用其最佳时间-频率分解以得到其参数。此可(例如)使用代数码本激励线性预测(ACELP)加上时域带宽扩展(TDBWE)以及利用外部滤波器组的参数化多声道编码(例如DFT)的组合来实施,其中ACELP可编码音频信号的低频带且TDBWE可编码音频信号的高频带。此组合特别有效率,因为已知用于语音的最佳带宽扩展应在时域中且多声道处理在频域中。由于ACELP+TDBWE不具有任何时间-频率转换器,因此外部滤波器组或如DFT的变换是有利的。此外,多声道处理器的成帧可与ACELP中所使用的成帧相同。即使多声道处理是在频域中进行,用于计算其参数或进行降混的时间分辨率应理想地接近于或甚至等于ACELP的成帧。

[0020] 所描述的实施例是有益的,因为可应用于不同信号内容类型的联合多声道编码的独立选择。

附图说明

[0021] 随后将参照附图论述本发明的实施例,其中:

[0022] 图1展示用于编码多声道音频信号的编码器的示意性框图;

[0023] 图2展示根据实施例的线性预测域编码器的示意性框图;

[0024] 图3展示根据实施例的频域编码器的示意性框图;

[0025] 图4展示根据实施例的音频编码器的示意性框图;

[0026] 图5a展示根据实施例的有源降混频器的示意性框图;

[0027] 图5b展示根据实施例的无源降混频器的示意性框图;

[0028] 图6展示用于解码经编码的音频信号的解码器的示意性框图;

[0029] 图7展示根据实施例的解码器的示意性框图;

[0030] 图8展示编码多声道信号的方法的示意性框图;

- [0031] 图9展示解码经编码的音频信号的方法的示意性框图；
- [0032] 图10展示根据另一方面的用于编码多声道信号的编码器的示意性框图；
- [0033] 图11展示根据另一方面的用于解码经编码的音频信号的解码器的示意性框图；
- [0034] 图12展示根据另一方面的用于编码多声道信号的音频编码方法的示意性框图；
- [0035] 图13展示根据另一方面的解码经编码的音频信号的方法的示意性框图；
- [0036] 图14展示从频域编码至LPD编码的无缝切换的示意性时序图；
- [0037] 图15展示从频域解码至LPD域解码的无缝切换的示意性时序图；
- [0038] 图16展示从LPD编码至频域编码的无缝切换的示意性时序图；
- [0039] 图17展示从LPD解码至频域解码的无缝切换的示意性时序图；
- [0040] 图18展示根据另一方面的用于编码多声道信号的编码器的示意性框图；
- [0041] 图19展示根据另一方面的用于解码经编码的音频信号的解码器的示意性框图；
- [0042] 图20展示根据另一方面的用于编码多声道信号的音频编码方法的示意性框图；
- [0043] 图21展示根据另一方面的解码经编码的音频信号的方法的示意性框图。

具体实施方式

[0044] 在下文中,将更详细地描述本发明的实施例。各个附图中所示的具有相同或相似功能的元件将与相同的附图标记相关联。

[0045] 图1展示用于编码多声道音频信号4的音频编码器2的示意性框图。音频编码器包括线性预测域编码器6、频域编码器8以及用于在线性预测域编码器6与频域编码器8之间切换的控制器10。控制器可分析多声道信号且针对多声道信号的部分决定是线性预测域编码还是频域编码有利。换言之,控制器被配置为使得多声道信号的部分由线性预测域编码器的编码帧表示或由频域编码器的编码帧表示。线性预测域编码器包括用于对多声道信号4进行降混以获得降混信号14的降混频器12。线性预测域编码器还包括用于编码降混信号的线性预测域核心编码器16,此外,线性预测域编码器包括用于从多声道信号4生成第一多声道信息20的第一联合多声道编码器18,第一多声道信息包括(例如)双耳声级差(interaural level difference,ILD)和/或双耳相位差(interaural phase difference,IPD)参数。多声道信号可以是(例如)立体声信号,其中降混频器将立体声信号转换为单声道信号。线性预测域核心编码器可编码单声道信号,其中第一联合多声道编码器可生成经编码的单声道信号的立体声信息以作为第一多声道信息。当与关于图10及图11所描述的另一方面相比时,频域编码器及控制器是可选择的。然而,为了时域编码与频域编码之间的信号自适应切换,使用频域编码器及控制器是有利的。

[0046] 此外,频域编码器8包括第二联合多声道编码器22,其用于从多声道信号4生成第二多声道信息24,其中第二联合多声道编码器22不同于第一多声道编码器18。然而,对于被第二编码器更好地编码的信号,第二联合多声道处理器22获得允许比由第一多声道编码器获得的第一多声道信息的第一再现质量高的第二再现质量的第二多声道信息。

[0047] 换言之,根据实施例,第一联合多声道编码器18用于生成允许第一再现质量的第一多声道信息20,其中第二联合多声道编码器22用于生成允许第二再现质量的第二多声道信息24,其中第二再现质量高于第一再现质量。此情况至少与被第二多声道编码器更好地编码的信号(如,语音信号)相关。

[0048] 因此,第一多声道编码器可以是包括(例如)立体声预测编码器、参数化立体声编码器或基于旋转的参数化立体声编码器的参数化联合多声道编码器。此外,第二联合多声道编码器可以是波形保持的,如(例如)频带选择性切换至中间/侧或左/右立体声编码器。如图1中所描绘,经编码的降混信号26可被传输至音频解码器且选择性地伺服第一联合多声道处理器,在第一联合多声道处理器处,例如,经编码的降混信号可被解码,且可计算来自编码之前及解码经编码的信号之后的多声道信号的残余信号以改良解码器侧的经编码的音频信号的解码质量。此外,在确定用于多声道信号的当前部分的合适编码方案之后,控制器10可分别使用控制信号28a、28b来控制线性预测域编码器及频域编码器。

[0049] 图2展示根据实施例的线性预测域编码器6的框图。至线性预测域编码器6的输入是由降混滤波器12降混的降混信号14。此外,线性预测域编码器包括ACELP处理器30及TCX处理器32。ACELP处理器30用于对经降取样的降混信号34进行操作,降混信号可由被降取样器35降取样。此外,时域带宽扩展处理器36可参数化编码降混信号14的部分的频带,其被从输入至ACELP处理器30的经降取样的降混信号34中移除。时域带宽扩展处理器36可输出降混信号14的部分的经参数化编码的频带38。换言之,时域带宽扩展处理器36可计算可包括相比降取样器35的截止频率较高的频率的降混信号14的频带的参数化表示。因此,降取样器35可具有其他属性以将高于降取样器的截止频率的那些频带提供至时域带宽扩展处理器36,或将截止频率提供至时域带宽扩展(TD-BWE)处理器以使TD-BWE处理器36能够计算用于降混信号14的正确部分的参数38。

[0050] 此外,TCX处理器用于对降混信号进行操作,降混信号(例如)未被降取样或以小于用于ACELP处理器的降取样的程度被降取样。当与输入至ACELP处理器30的经降取样的降混信号35相比时,小于ACELP处理器的降取样的程度的降取样可以是使用较高的截止频率的降取样,其中大量的降混信号的频带被提供至TCX处理器。TCX处理器还可包括第一时间-频率转换器40,如MDCT、DFT或DCT。TCX处理器32还可包括第一参数生成器42及第一量化器编码器44。第一参数生成器42(例如,智能间隙填充(intelligent gap filling,IGF)算法)可计算第一频带集合的第一参数化表示46,其中第一量化器编码器44(例如)使用TCX算法来计算用于第二频带集合的经量化编码的频谱线的第一集合48。换言之,第一量化器编码器可参数化编码入站信号(inbound signal)的相关频带(如,音调频带),其中第一参数生成器将例如IGF算法应用于入站信号的剩余频带以进一步减小经编码的音频信号的带宽。

[0051] 线性预测域编码器6还可包括线性预测域解码器50,其用于解码降混信号14(例如,由经ACELP处理的经降取样的降混信号52来表示)和/或第一频带集合的第一参数化表示46和/或用于第二频带集合的经量化编码的频谱线的第一集合48来表示的降混信号14。线性预测域解码器50的输出可以是经编码且经解码的降混信号54。此信号54可被输入至多声道残余编码器56,其可使用经编码且经解码的降混信号54来计算并编码多声道残余信号58,其中经编码的多声道残余信号表示使用第一多声道信息的经解码的多声道表示与降混之前的多声道信号之间的误差。因此,多声道残余编码器56可包括联合编码器侧多声道解码器60和差异处理器62。联合编码器侧多声道解码器60可使用第一多声道信息20及经编码且经解码的降混信号54而生成经解码的多声道信号,其中差异处理器可形成经解码的多声道信号64与降混之前的多声道信号4之间的差异以获得多声道残余信号58。换言之,音频编码器内的联合编码器侧多声道解码器可执行解码操作,这是有利的,相同的解码操作在解

码器侧上执行。因此,在用于解码经编码的降混信号的联合编码器侧多声道解码器中使用可在传输之后由音频解码器得出的第一联合多声道信息。差异处理器62可计算经解码的联合多声道信号与原始多声道信号4之间的差异。经编码的多声道残余信号58可改进音频解码器的解码质量,因为经解码的信号与原始信号之间的由于(例如)参数化编码所致的差异可通过了解这两个信号之间的差异来减小。此使得第一联合多声道编码器能够以得出用于多声道音频信号的全带宽的多声道信息的方式进行操作。

[0052] 此外,降混信号14可包括低频带及高频带,其中线性预测域编码器6用于使用(例如)时域带宽扩展处理器36来施加带宽扩展处理以用于参数化编码高频带,其中线性预测域解码器6用于仅获得表示降混信号14的低频带的低频带信号作为经编码且经解码的降混信号54,且其中经编码的多声道残余信号仅具有在降混之前的多声道信号的低频带内的频率。换言之,带宽扩展处理器可计算用于高于截止频率的频带的带宽扩展参数,其中ACELP处理器对低于截止频率的频率进行编码。解码器因此用于基于经编码的低频带信号及带宽参数38来重建较高频率。

[0053] 根据其他实施例,多声道残余编码器56可计算侧信号,且其中降混信号为M/S多声道音频信号的对应中间信号。因此,多声道残余编码器可计算并编码经计算的侧信号(其可计算自由滤波器组82获得的多声道音频信号的全频带频谱表示)与经编码且经解码的降混信号54的倍数的经预测的侧信号的差异,其中倍数可由成为多声道信息的部分的预测信息来表示。然而,降混信号仅包括低频带信号。因此,残余编码器还可计算用于高频带的残余(或侧)信号。此计算可(例如)通过模拟时域带宽扩展(如在线性预测域核心编码器中所进行的)或通过预测作为经计算的(全频带)侧信号与经计算的(全频带)中间信号之间的差异的侧信号来执行,其中预测因子用于将两个信号之间的差异最小化。

[0054] 图3展示根据实施例的频域编码器8的示意性框图。频域编码器包括第二时间-频率转换器66、第二参数生成器68以及第二量化器编码器70。第二时间-频率转换器66可将多声道信号的第一声道4a及多声道信号的第二声道4b转换成频谱表示72a、72b。第一声道及第二声道的频谱表示72a、72b可被分析且各自分裂成第一频带集合74及第二频带集合76。因此,第二参数生成器68可生成第二频带集合76的第二参数化表示78,其中第二量化器编码器可生成第一频带集合74的经量化且经编码的表示80。频域编码器或更特别地,第二时间-频率转换器66可针对第一声道4a及第二声道4b执行(例如)MDCT操作,其中第二参数生成器68可执行智能间隙填充算法且第二量化器编码器70可执行(例如)AAC操作。因此,如关于线性预测域编码器已描述的,频域编码器也能够以得出用于多声道音频信号的全带宽的多声道信息的方式来操作。

[0055] 图4展示根据优选实施例的音频编码器2的示意性框图。LPD路径16由含有“有源或无源DMX”降混计算12的联合立体声或多声道编码组成,降混计算指示LPD降混可为有源(“频率选择性”)或无源(“恒定混合因子”),如图5中所描绘。降混还可由TD-BWE模块或IGF模块支持的可切换单声道ACELP/TCX核心来编码。应注意的是,ACELP对经降取样的输入音频数据34进行操作。可对经降取样的TCX/IGF输出执行由于切换所致的任何ACELP初始化。

[0056] 由于ACELP不含有任何内部时间-频率分解,LPD立体声编码借助于LP编码之前的分析滤波器组82及LPD解码之后的合成滤波器组来添加额外的复调制滤波器组。在优选的实施例中,使用具有低重叠区域的过取样的DFT。然而,在其他实施例中,可使用具有类似时

间分辨率的任何过取样的时间-频率分解。然后,可在频域中计算立体声参数。

[0057] 参数化立体声编码通过“LPD立体声参数编码”区块18执行,该区块18将LPD立体声参数20输出至比特流。可选择地,后续区块“LPD立体声残余编码”将向量量化的低通降混残余58添加至比特流。

[0058] FD路径8被配置为具有其自身的内部联合立体声或多声道编码。对于联合立体声编码,路径再次使用其自身的临界取样及实值的滤波器组66,即(例如)MDCT。

[0059] 提供至解码器的信号可(例如)多路复用至单一比特流。比特流可包括经编码的降混信号26,经编码的降混信号还可包括以下各者中的至少一个:经参数化编码的经时域带宽扩展的频带38、经ACELP处理的经降取样的降混信号52、第一多声道信息20、经编码的多声道残余信号58、第一频带集合的第一参数化表示46、第二频带集合的经量化编码的频谱线的第一集合48以及包括第一频带集合的经量化且经编码的表示80及第一频带集合的第二参数化表示78的第二多声道信息24。

[0060] 实施例展示用于将可切换核心编解码器、联合多声道编码以及参数化空间音频编码组合成完全可切换感知编解码器的经改良的方法,完全可切换感知编解码器允许依据核心编码器的选择而使用不同的多声道编码技术。特别地,在可切换音频编码器内,组合本地(native)频域立体声编码与基于ACELP/TCX的线性预测性编码(其具有其自身的专用独立参数化立体声编码)。

[0061] 图5a及图5b分别展示根据实施例的有源降混频器及无源降混频器。有源降混频器使用(例如)用于将时域信号4变换成频域信号的时间频率转换器82而在频域中操作。在降混之后,频率-时间转换(例如,IDFT)可将来自频域的降混信号转换成时域中的降混信号14。

[0062] 图5b展示根据实施例的无源降混频器12。无源降混频器12包括加法器,其中第一声道4a及第一声道4b在分别使用权重a 84a及权重b 84b加权之后被组合。此外,第一声道4a及第二声道4b在传输至LPD立体声参数化编码之前可被输入至时间-频率转换器82。

[0063] 换言之,降混频器用于将多声道信号转换成频谱表示,且其中使用频谱表示或使用时域表示执行降混,且其中第一多声道编码器用于使用频谱表示来生成用于频谱表示的各个频带的单独第一多声道信息。

[0064] 图6展示根据实施例的用于对经编码的音频信号103进行解码的音频解码器102的示意性框图。音频解码器102包括线性预测域解码器104、频域解码器106、第一联合多声道解码器108、第二多声道解码器110以及第一组合器112。经编码的音频信号103(其可以是先前所描述的编码器部分的经多路复用的比特流,例如音频信号的帧)可被联合多声道解码器108使用第一多声道信息20来解码或被频域解码器106解码,且被第二联合多声道解码器110使用第二多声道信息24进行多声道解码。第一联合多声道解码器可输出第一多声道表示114,且第二联合多声道解码器110的输出可以是第二多声道表示116。

[0065] 换言之,第一联合多声道解码器108使用线性预测域编码器的输出及使用第一多声道信息20生成第一多声道表示114。第二多声道解码器110使用频域解码器的输出及第二多声道信息24生成第二多声道表示116。此外,第一组合器组合第一多声道表示114及第二多声道表示116(例如,基于帧)以获得经解码的音频信号118。此外,第一联合多声道解码器108可以是使用(例如)复数预测(complex prediction)、参数化立体声操作或旋转操作的

参数化联合多声道解码器。第二联合多声道解码器110可以是使用(例如)频带选择性切换至中间/侧或左/右立体声解码算法的波形保持联合多声道解码器。

[0066] 图7展示根据另一实施例的解码器102的示意性框图。本文中,线性预测域解码器102包括ACELP解码器120、低频带合成器122、升取样器124、时域带宽扩展处理器126或用于组合经升取样的信号和经带宽扩展的信号的第二组合器128。此外,线性预测域解码器可包括TCX解码器132及智能间隙填充处理器132,其在图7中被描绘为一个区块。此外,线性预测域解码器102可包括用于组合第二组合器128及TCX解码器130及IGF处理器132的输出的全频带合成处理器134。如关于编码器已展示的,时域带宽扩展处理器126、ACELP解码器120以及TCX解码器130并行地工作以解码各个经传输的音频信息。

[0067] 可提供交叉路径136,其用于使用从低频带频谱-时间转换(使用例如频率-时间转换器138)自TCX解码器130及IGF处理器132得出的信息来初始化低频带合成器。参照声道的模型,ACELP数据可模型化声道的形状,其中TCX数据可模型化声道的激励。由低频带频率-时间转换器(如IMDCT解码器)表示的交叉路径136使低频带合成器122能够使用声道的形状及当前激励来重新计算或解码经编码的低频带信号。此外,经合成的低频带被升取样器124升取样,且使用例如第二组合器128而被与经时域带宽扩展的高频带140组合,以(例如)整形经升取样的频率以恢复(例如)每个经升取样的频带的能量。

[0068] 全频带合成器134可使用第二组合器128的全频带信号及来自TCX处理器130的激励来形成经解码的降混信号142。第一联合多声道解码器108可包括时间-频率转换器144,其用于将线性预测域解码器的输出(例如,经解码的降混信号142)转换成频谱表示145。此外,升混频器(例如,实施于立体声解码器146中)可由第一多声道信息20控制以将频谱表示升混成多声道信号。此外,频率-时间转换器148可将升混结果转换成时间表示114。时间-频率和/或频率-时间转换器可包括复数运算(complex operation)或过取样的操作,诸如DFT或IDFT。

[0069] 此外,第一联合多声道解码器或更特别地立体声解码器146可使用(例如)由多声道经编码的音频信号103提供的多声道残余信号58来生成第一多声道表示。此外,多声道残余信号可包括比第一多声道表示低的带宽,其中第一联合多声道解码器用于使用第一多声道信息重建中间第一多声道表示且将多声道残余信号添加至中间第一多声道表示。换言之,立体声解码器146可包括使用第一多声道信息20的多声道解码,且在经解码的降混信号的频谱表示已被升混成多声道信号之后,可选择地包括通过将多声道残余信号添加至经重建的多声道信号的经重建的多声道信号的改良。因此,第一多声道信息及残余信号可能已对多声道信号起作用。

[0070] 第二联合多声道解码器110可使用由频域解码器获得的频谱表示作为输入。频谱表示包括至少针对多个频带的第一声道信号150a及第二声道信号150b。此外,第二联合多声道处理器110可适用于第一声道信号150a及第二声道信号150b的多个频带。联合多声道操作,(如掩码(mask),)为各个频带指示左/右或中间/侧联合多声道编码,且其中联合多声道操作是用于将由掩码指示的频带从中间/侧表示转换为左/右表示的中间/侧或左/右转换操作,其为联合多声道操作的结果至时间表示以获得第二多声道表示的转换。此外,频域解码器可包括频率-时间转换器152,其为(例如)IMDCT操作或特定取样的操作。换言之,掩码可包括指示(例如)L/R或M/S立体声编码的旗标,其中第二联合多声道编码器将对应立体

声编码算法应用于各个音频帧。可选择地,智能间隙填充可应用于经编码的音频信号以进一步减小经编码的音频信号的带宽。因此,例如,音调频带可使用前面提及的立体声编码算法以高分辨率被编码,其中其他频带可使用(例如)IGF算法被参数化编码。

[0071] 换言之,在LPD路径104中,经传输的单声道信号是被(例如)由TD-BWE 126或IGF模块132支持的可切换ACELP/TCX 120/130解码器重建的。将对经降取样的TCX/IGF输出执行因切换所致的任何ACELP初始化。ACELP的输出使用(例如)升取样器124被升取样至全采样率。所有信号,使用(例如)混频器128以高采样率在时域中被混合,且被LPD立体声解码器146进一步处理以提供LPD立体声。

[0072] LPD“立体声解码”是由受经传输的立体声参数20的应用操控的经传输的降混的升混组成。可选择地,降混残余58也包含于比特流中。在此情况下,残余被“立体声解码”146解码且被包括于升混计算中。

[0073] FD路径106被配置为具有其自身的独立内部联合立体声或多声道解码。对于联合立体声解码,路径再次使用其自身的临界取样及实值的滤波器组152,例如(即)IMDCT。

[0074] LPD立体声输出及FD立体声输出使用(例如)第一组合器112而在时域中被混频,以提供全切换式编码器的最终输出118。

[0075] 尽管关于相关附图中的立体声解码描述了多声道,但相同原理通常也可应用于利用两个或更多个声道的多声道处理。

[0076] 图8展示用于编码多声道信号的方法800的示意性框图。方法800包括:执行线性预测域编码的步骤805;执行频域编码的步骤810;在线性预测域编码与频域编码之间切换的步骤815,其中线性预测域编码包括降混多声道信号以获得降混信号、对降混信号进行线性预测域核心编码以及从多声道信号生成第一多声道信息的第一联合多声道编码,其中频域编码包括从多声道信号生成第二多声道信息的第二联合多声道编码,其中第二联合多声道编码不同于第一多声道编码,且其中切换被执行以使得多声道信号的部分由线性预测域编码的编码帧或由频域编码的编码帧表示。

[0077] 图9展示对经编码的音频信号进行解码的方法900的示意性框图。方法900包括:线性预测域解码的步骤905;频域解码的步骤910;使用线性预测域解码的输出及使用第一多声道信息来生成第一多声道表示的第一联合多声道解码的步骤915;使用频域解码的输出及第二多声道信息来生成第二多声道表示的第二多声道解码的步骤920;以及组合第一多声道表示及第二多声道表示以获得经解码的音频信号的步骤925,其中第二多声道信息解码不同于第一多声道解码。

[0078] 图10展示根据另一方面的用于编码多声道信号的音频编码器的示意性框图。音频编码器2'包括线性预测域编码器6及多声道残余编码器56。线性预测域编码器包括用于对多声道信号4进行降混以获得降混信号14的降混频器12、用于编码降混信号14的线性预测域核心编码器16。线性预测域编码器6还包括用于从多声道信号4生成多声道信息20的联合多声道编码器18。此外,线性预测域编码器包括用于对经编码的降混信号26进行解码以获得经编码且经解码的降混信号54的线性预测域解码器50。多声道残余编码器56可使用经编码且经解码的降混信号54来计算并编码多声道残余信号。多声道残余信号可表示使用多声道信息20的经解码的多声道表示54与降混之前的多声道信号4之间的误差。

[0079] 根据实施例,降混信号14包括低频带及高频带,其中线性预测域编码器可使用带

宽扩展处理器来施加带宽扩展处理以用于参数化编码高频带,其中线性预测域解码器用于仅获得表示降混信号的低频带的低频带信号作为经编码且经解码的降混信号54,且其中经编码的多声道残余信号仅具有对应于降混之前的多声道信号的低频带的频带。此外,关于音频编码器2的相同描述可应用于音频编码器2'。然而,省略编码器2的其他频率编码。此省略简化了编码器配置,且因此在以下情况下是有利的:编码器仅用于仅包括可在时域中被参数化编码而无明显质量损失的信号的音频信号,或经解码的音频信号的质量仍在规范内。然而,专用残余立体声编码对于增加经解码的音频信号的再现质量是有利的。更特别地,编码之前的音频信号与经编码且经解码的音频信号之间的差异被得出且被传输至解码器以增加经解码的音频信号的再现质量,因为经解码的音频信号与经编码的音频信号的差异已被解码器知晓。

[0080] 图11展示根据另一方面的用于解码经编码的音频信号103的音频解码器102'。音频解码器102'包括线性预测域解码器104,以及用于使用线性预测域解码器104的输出及联合多声道信息20来生成多声道表示114的联合多声道解码器108。此外,经编码的音频信号103可包括多声道残余信号58,其可被多声道解码器用来生成多声道表示114。此外,与音频解码器102相关的相同解释可应用于音频解码器102'。本文中,从原始音频信号至经解码的音频信号的残余信号被使用并被应用至经解码的音频信号以至少几乎达成与原始音频信号相比的相同质量的经解码的音频信号,即使在使用了参数化且因此有损的编码的情况下。然而,在音频解码器102'中省略关于音频解码器102所展示的频率解码部分。

[0081] 图12展示用于编码多声道信号的音频编码方法1200的示意性框图。方法1200包括:线性预测域编码的步骤1205,其包括对多声道信号进行降混以获得降混多声道信号,以及线性预测域核心编码器从多声道信号生成多声道信息,其中方法还包括对降混信号进行线性预测域解码以获得经编码且经解码的降混信号;以及多声道残余编码的步骤1210,其使用经编码且经解码的降混信号来计算经编码的多声道残余信号,多声道残余信号表示使用第一多声道信息的经解码的多声道表示与降混之前的多声道信号之间的误差。

[0082] 图13展示对经编码的音频信号进行解码的方法1300的示意性框图。方法1300包括线性预测域解码的步骤1305,以及联合多声道解码的步骤1310,其使用线性预测域解码的输出及联合多声道信息来生成多声道表示,其中经编码的多声道音频信号包括声道残余信号,其中联合多声道解码使用多声道残余信号以生成多声道表示。

[0083] 所描述的实施例可得以用在所有类型的立体声或多声道音频内容(在给定低比特率下具有恒定感知质量的语音及相似音乐)的广播的分布如关于数字无线电、因特网串流及音频通信应用中。

[0084] 图14至图17描述如何在LPD编码与频域编码之间及相反情况应用所提出的无缝切换的实施例。通常,之前的窗口化或处理使用细线来指示,粗线指示切换被施加的当前的窗口化或处理,且虚线指示只为过渡或切换进行的当前处理。从LPD编码至频率编码的切换或过渡。

[0085] 图14展示指示频域编码至时域编码之间的无缝切换的实施例的示意性时序图。若(例如)控制器10指示使用LPD编码而不是用于先前帧的FD编码来更好地编码当前帧,则此图可能相关。在频域编码期间,停止窗口200a及200b可应用于每个立体声信号(其可选择性地扩展至两个以上声道)。停止窗口不同于在第一帧204的开始202处衰落的标准MDCT重叠

相加。停止窗口的左边部分可为用于使用(例如)MDCT时间-频率变换来编码先前帧的经典重叠相加。因此,切换之前的帧仍被合适地编码。对于施加切换的当前帧204,计算额外立体声参数,即便用于时域编码的中间信号的第一参数化表示被计算用于随后的帧206。进行这两个额外的立体声分析以用于能够生成用于LPD预看的中间信号208。不过,在两个第一LPD立体声窗口中(另外地)传输立体声参数。正常情况下,立体声参数以两个LPD立体声帧的延迟被发送。为了更新ACELP内存(如为了LPC分析或前向混叠消除(forward aliasing cancellation,FAC)),中间信号也变得可用于过去。因此,在(例如)应用使用DFT的时间-频率转换之前,可在分析滤波器组82中应用用于第一立体声信号的LPD立体声窗口210a至210d及用于第二立体声信号的LPD立体声窗口212a至212d。中间信号在使用TCX编码时可包括典型淡入淡出渐变(crossfade ramp),从而导致例示性LPD分析窗口214。若将ACELP用于编码音频信号(诸如单声道低频带信号),则简单地选择应用LPC分析的多个频带,通过矩形LPD分析窗口216来指示。

[0086] 此外,由垂直线218指示的时序展示:施加有过渡的当前帧包括来自频域分析窗口200a、200b的信息以及经计算的中间信号208及对应立体声信息。在线202与线218之间的频率分析窗口的水平部分期间,帧204使用频域编码而被完美地编码。从线218至频率分析窗口在线220处的结束,帧204包括来自频域编码及LPD编码二者的信息,且从线220至帧204在垂直线222处的结束,仅LPD编码有助于帧的编码。进一步地注意编码的中间部分,因为第一及最后(第三)部分仅从一个编码技术得出而不具有混叠。然而,对于中间部分,其应在ACELP与TCX单声道信号编码之间进行区分。由于TCX编码使用淡入淡出,如关于频域编码已应用,频率经编码的信号的简单淡出及经TCX编码的中间信号的淡入提供用于编码当前帧204的完整信息。若将ACELP用于单声道信号编码,则可应用更复杂的处理,因为区域224可能不包括用于编码音频信号的完整信息。所提出的方法为前向混叠校正(forward aliasing correction,FAC),例如,在USAC规范中在章节7.16中所描述。

[0087] 根据实施例,控制器10用于在多声道音频信号的当前帧204内从使用频域编码器8对先前帧进行编码切换至使用线性预测域编码器对即将到来的帧(upcoming frame)进行解码。第一联合多声道编码器18可从当前帧的多声道音频信号计算合成多声道参数210a、210b、212a、212b,其中第二联合多声道编码器22用于使用停止窗口对第二多声道信号进行加权。

[0088] 图15展示对应于图14的编码器操作的解码器的示意性时序图。本文中,根据实施例来描述当前帧204的重建。如在图14的编码器时序图中所见,从应用停止窗口200a及200b的先前帧提供频域立体声声道。如在单声道情况下,首先对经解码的中间信号进行从FD至LPD模式的过渡。这通过从以FD模式解码的时域信号116人工建立中间信号226来达成,其中ccfl为核心码帧长度且L_fac表示频率混叠消除窗口或帧或区块或变换的长度。

[0089]
$$x[n-ccfl/2]=0.5 \cdot l_{i-1}[n]+0.5 \cdot r_{i-1}[n], \text{ 对于 } ccfl \leq n < \frac{ccfl}{2} + L_{fac}$$

[0090] 此信号接着被传输至LPD解码器120以用于更新内存及应用FAC解码,如在单声道情况下针对FD模式至ACELP的过渡所进行的。在USAC规范[ISO/IEC DIS 23003-3,Usac]中在章节7.16中描述了处理。在FD模式至TCX的情况下,执行传统的重叠相加。例如,通过将所传输的立体声参数210及212用于立体声处理,其中过渡已经完成,LPD立体声解码器146接

收经解码的(在频域中,在应用时间-频率转换器144的时间-频率转换之后)中间信号作为输入信号。然后,立体声解码器输出与以FD模式解码的先前帧重叠的左声道信号228及右声道信号230。然后,信号(即,用于施加过渡的帧的经FD解码的时域信号及经LPD解码的时域信号)在每个声道上淡入淡出(在组合器112中)以用于平滑左声道及右声道中的过渡:

$$[0091] \quad \begin{aligned} & l \left[n - \frac{ccfl}{2} + L_{fac} \right] \\ & = \begin{cases} l_{i-1}[ccfl + n] & , \text{ 对于 } 0 \leq n < \frac{ccfl}{2} - L_{fac} - L \\ l_{i-1} \left[ccfl + \frac{ccfl}{2} - L_{fac} - L + n \right] \cdot w[L-1-n] + l_i[n] \cdot w[n] & , \text{ 对于 } 0 \leq n < L \\ l_i[n] & , \text{ 对于 } L \leq n < M \end{cases} \end{aligned}$$

$$[0092] \quad \begin{aligned} & r \left[n - \frac{ccfl}{2} + L_{fac} \right] \\ & = \begin{cases} r_{i-1}[ccfl + n] & , \text{ 对于 } 0 \leq n < \frac{ccfl}{2} - L_{fac} - L \\ r_{i-1} \left[ccfl + \frac{ccfl}{2} - L_{fac} - L + n \right] \cdot w[L-1-n] + r_i[n] \cdot w[n] & , \text{ 对于 } 0 \leq n < L \\ r_i[n] & , \text{ 对于 } L \leq n < M \end{cases} \end{aligned}$$

[0093] 在图15中,使用 $M=ccfl/2$ 来示意性地说明过渡。此外,组合器可在仅使用FD或LPD解码来解码而没有这些模式之间的过渡的连续帧处的执行淡入淡出。

[0094] 换言之,FD解码的重叠相加过程(尤其当将MDCT/IMDCT用于时间-频率/频率-时间转换时)被替换为经FD解码的音频信号及经LPD解码的音频信号的淡入淡出。因此,解码器应计算用于经FD解码的音频信号的淡出部分至经LPD解码的音频信号的淡入部分的LPD信号。根据实施例,音频解码器102用于在多声道音频信号的当前帧204内从使用频域解码器106对先前帧进行解码切换至使用线性预测域解码器104对即将到来的帧进行解码。组合器112可从当前帧的第二多声道表示116来计算合成中间信号226。第一联合多声道解码器108可使用合成中间信号226及第一多声道信息20来生成第一多声道表示114。此外,组合器112用于组合第一多声道表示及第二多声道表示以获得多声道音频信号的经解码的当前帧。

[0095] 图16展示用于在当前帧232中执行使用LPD编码至使用FD解码的过渡的编码器中的示意性时序图。为了从LPD切换至FD编码,可对FD多声道编码应用开始窗口300a、300b。当与停止窗口200a、200b相比时,开始窗口具有类似功能。在垂直线234与236之间的LPD编码器的经TCX编码的单声道信号的淡出期间,开始窗口300a、300b执行淡入。当使用ACELP替代TCX时,单声道信号不执行平滑淡出。尽管如此,可使用(例如)FAC在解码器中重建正确音频信号。LPD立体声窗口238及240被默认地计算且参考经ACELP或TCX编码的单声道信号(由LPD分析窗口241指示)。

[0096] 图17展示对应于关于图16所描述的编码器的时序图的解码器中的示意性时序图。

[0097] 对于从LPD模式至FD模式的过渡,由立体声解码器146解码额外帧。来自LPD模式解码器的中间信号以零进行扩展以用于帧索引 $i=ccfl/M$ 。

$$[0098] \quad x[i \cdot M + n - L] = \begin{cases} x[i \cdot M + n - L] & , \text{ 对于 } 0 \leq n < L + 2 \cdot L_{fac} \\ 0 & , \text{ 对于 } L + 2 \cdot L_{fac} \leq n < M \end{cases}$$

[0099] 如先前所描述的立体声解码可通过保留上一个立体声参数及通过切断侧信号反量化(即,将code_mode设定为0)来执行。此外,反向DFT之后的右侧窗口化不被应用,此导致额外LPD立体声窗口244a、244b的陡峭边缘242a、242b。可清晰地看到,形状边缘位于平面区

段246a、246b处,其中帧的对应部分的整个信息可从经FD编码的音频信号得出。因此,右侧窗口化(无陡峭边缘)可导致LPD信息对FD信息的不想要的干扰且因此不被应用。

[0100] 然后,通过在TCX至FD模式的情况下使用重叠相加处理或通过ACELP至FD模式的情况下对每个声道使用FAC将所得的左和右(经LPD解码的)声道250a、250b(使用由LPD分析窗口248指示的经LPD解码的中间信号及立体声参数)组合至下个帧的经FD模式解码的声道。在图17中描绘过渡的示意性说明,其中 $M = ccf1/2$ 。

[0101] 根据实施例,音频解码器102可在多声道音频信号的当前帧232内从使用线性预测域解码器104对先前帧进行解码切换至使用频域解码器106对即将到来的帧进行解码。立体声解码器146可使用先前帧的多声道信息从用于当前帧的线性预测域解码器的经解码的单声道信号计算合成多声道音频信号,其中第二联合多声道解码器110可计算用于当前帧的第二多声道表示及使用开始窗口对第二多声道表示加权。组合器112可组合合成多声道音频信号及经加权的第二多声道表示以获得多声道音频信号的经解码的当前帧。

[0102] 图18展示用于编码多声道信号4的编码器2”的示意性框图。音频编码器2”包括降混频器12、线性预测域核心编码器16、滤波器组82以及联合多声道编码器18。降混频器12用于对多声道信号4进行降混以获得降混信号14。降混信号可以是单声道信号,例如M/S多声道音频信号的中间信号。线性预测域核心编码器16可编码降混信号14,其中降混信号14具有低频带及高频带,其中线性预测域核心编码器16用于施加带宽扩展处理以用于对高频带进行参数化编码。此外,滤波器组82可生成多声道信号4的频谱表示,且联合多声道编码器18可用于处理包括多声道信号的低频带及高频带的频谱表示以生成多声道信息20。多声道信息可包括ILD和/或IPD和/或双耳强度差异(IID, Interaural Intensity Difference)参数,从而使解码器能够从单声道信号重新计算多声道音频信号。根据此方面的实施例的其他方面的更详细的附图可在先前图中、尤其在图4中找到。

[0103] 根据实施例,线性预测域核心编码器16还可包括用于对经编码的降混信号26进行解码以获得经编码且经解码的降混信号54的线性预测域解码器。本文中,线性预测域核心编码器可形成被编码的M/S音频信号的中间信号用于传输至解码器。此外,音频编码器还包括用于使用经编码且经解码的降混信号54来计算经编码的多声道残余信号58的多声道残余编码器56。多声道残余信号表示使用多声道信息20的经解码的多声道表示与降混之前的多声道信号4之间的误差。换言之,多声道残余信号58可以是M/S音频信号的侧信号,其对应于使用线性预测域核心编码器计算的中间信号。

[0104] 根据其他实施例,线性预测域核心编码器16用于施加带宽扩展处理以用于对高频带进行参数化编码以及仅获得表示降混信号的低频带的低频带信号以作为经编码且经解码的降混信号,且其中经编码的多声道残余信号58仅具有对应于降混之前的多声道信号的低频带的频带。额外地或可选地,多声道残余编码器可模拟在线性预测域核心编码器中应用于多声道信号的高频带的时域带宽扩展,并计算用于高频带的残余或侧信号以使得能够更准确地解码单声道或中间信号从而得出经解码的多声道音频信号。模拟可包括相同或类似计算,其在解码器中执行以解码经带宽扩展的高频带。作为模拟带宽扩展的替代或补充的方法可以是预测侧信号。因此,多声道残余编码器可从滤波器组82中的时间-频率转换之后的多声道音频信号4的参数化表示83来计算全频带残余信号。可比较此全频带侧信号与从参数化表示83类似地得出的全频带中间信号的频率表示。全频带中间信号可(例如)被计

算为参数化表示83的左声道及右声道的总和,且全频带侧信号可被计算为左声道及右声道的差。此外,预测因此可计算全频带中间信号的预测因子,最小化预测因子及全频带中间信号的乘积与全频带侧信号的绝对差。

[0105] 换言之,线性预测域编码器可用于计算降混信号14以作为M/S多声道音频信号的中间信号的参数化表示,其中多声道残余编码器可用于计算对应于M/S多声道音频信号的中间信号的侧信号,其中残余编码器可使用模拟时域带宽扩展来计算中间信号的高频带,或其中残余编码器可使用发现预测信息来预测中间信号的高频带,预测信息最小化来自先前帧的经计算的侧信号与经计算的全频带中间信号之间的差异。

[0106] 其他实施例展示包括ACELP处理器30的线性预测域核心编码器16。ACELP处理器可对经降取样的降混信号34进行操作。此外,时域带宽扩展处理器36用于对降混信号的通过第三降取样从ACELP输入信号中移除的部分的频带进行参数化编码。额外地或可选地,线性预测域核心编码器16可包括TCX处理器32。TCX处理器32可对降混信号14进行操作,该降混信号未被降取样或以小于用于ACELP处理器的降取样的程度而被降取样。此外,TCX处理器可包括第一时间-频率转换器40、用于生成第一频带集合的参数化表示46的第一参数生成器42以及用于生成第二频带集合的经量化编码的频谱线的集合48的第一量化器编码器44。ACELP处理器及TCX处理器可分别地执行(例如,使用ACELP编码第一数目的帧,且使用TCX编码第二数目的帧),或以ACELP及TCX均贡献信息以解码一个帧的联合方式执行。

[0107] 其他实施例展示不同于滤波器组82的时间-频率转换器40。滤波器组82可包括经优化以生成多声道信号4的频谱表示83的滤波器参数,其中时间-频率转换器40可包括经优化以生成第一频带集合的参数化表示46的滤波器参数。在另一步骤中,必须注意,线性预测域编码器在带宽扩展和/或ACELP的情况下使用不同滤波器组或甚至不使用滤波器组。此外,滤波器组82可不依赖于线性预测域编码器的先前参数选择而计算单独的滤波器参数以生成频谱表示83。换言之,LPD模式中的多声道编码可使用用于多声道处理的滤波器组(DFT),其并非是在带宽扩展(时域用于ACELP且MDCT用于TCX)中所使用的滤波器组。此情况的优点为每个参数化编码可使用其最佳时间-频率分解以得到其参数。例如,ACELP+TDBWE与利用外部滤波器组(例如,DFT)的参数化多声道编码的组合是有利的。此组合特别有效率,因为已知用于语音的最佳带宽扩展应在时域中且多声道处理应在频域中。由于ACELP+TDBWE不具有任何时间-频率转换器,因此如DFT的外部滤波器组或变换是优选的或甚至可能是必需的。其他概念始终使用相同滤波器组且因此不使用不同的滤波器组,例如:

[0108] -在MDCT中用于AAC的IGF及联合立体声编码

[0109] -在QMF中用于HeAACv2的SBR+PS

[0110] -在QMF中用于USAC的SBR+MPS212

[0111] 根据其他实施例,多声道编码器包括第一帧生成器且线性预测域核心编码器包括第二帧生成器,其中第一和第二帧生成器用于从多声道信号4形成帧,其中第一和第二帧生成器用于形成具有类似长度的帧。换言之,多声道处理器的成帧可与ACELP中所使用的成帧相同。即使多声道处理是在频域中进行,用于计算其参数或降混的时间分辨率应理想地接近于或甚至等于ACELP的成帧。此情况下的类似长度可指ACELP的成帧,其可等于或接近于用于计算用于多声道处理或降混的参数的时间分辨率。

[0112] 根据其他实施例,音频编码器还包括线性预测域编码器6(其包括线性预测域核心

编码器16及多声道编码器18)、频域编码器8以及用于在线性预测域编码器6与频域编码器8之间切换的控制器10。频域编码器8可包括用于对于来自多声道信号的第二多声道信息24进行编码的第二联合多声道编码器22,其中第二联合多声道编码器22不同于第一联合多声道编码器18。此外,控制器10被配置为使得多声道信号的部分由线性预测域编码器的编码帧表示或由频域编码器的编码帧表示。

[0113] 图19展示根据另一方面的用于对经编码的音频信号103进行解码的解码器102”的示意性框图,经编码的音频信号包括经核心编码的信号、带宽扩展参数以及多声道信息。音频解码器包括线性预测域核心解码器104、分析滤波器组144、多声道解码器146以及合成滤波器组处理器148。线性预测域核心解码器104可对经核心编码的信号进行解码以生成单声道信号。此信号可为M/S经编码的音频信号的全频带)中间信号。分析滤波器组144可将单声道信号转换成频谱表示145,其中多声道解码器146可从单声道信号的频谱表示及多声道信息20生成第一声道频谱及第二声道频谱。因此,多声道解码器可使用多声道信息,多声道信息(例如)包括对应于经解码的中间信号的侧信号。合成滤波器组处理器148用于对第一声道频谱进行合成滤波以获得第一声道信号且用于对第二声道频谱进行合成滤波以获得第二声道信号。因此,优选地,与分析滤波器组144相比的反向操作可应用于第一声道信号及第二声道信号,若分析滤波器组使用DFT,则反向操作可为IDFT。然而,滤波器组处理器可使用(例如)相同的滤波器组并行地或以连续次序来(例如)处理两个声道频谱。关于此另一方面的其他详细附图可在先前图中、尤其关于图7看出。

[0114] 根据其他实施例,线性预测域核心解码器包括:用于从带宽扩展参数及低频带单声道信号或经核心编码的信号生成高频带部分140以获得音频信号的经解码的高频带140的带宽扩展处理器126;用于解码低频带单声道信号的低频带信号处理器;以及用于使用经解码的低频带单声道信号及音频信号的经解码的高频带来计算全频带单声道信号的组合器128。低频带单声道信号可以是(例如)M/S多声道音频信号的中间信号的基带表示,其中带宽扩展参数可被应用以(在组合器128中)从低频带单声道信号计算全频带单声道信号。

[0115] 根据其他实施例,线性预测域解码器包括ACELP解码器120、低频带合成器122、升取样器124、时域带宽扩展处理器126或第二组合器128,其中第二组合器128用于组合经升取样的低频带信号及经带宽扩展的高频带信号140以获得全频带经ACELP解码的单声道信号。线性预测域解码器还可包括TCX解码器130及智能间隙填充处理器132以获得全频带经TCX解码的单声道信号。因此,全频带合成处理器134可组合全频带经ACELP解码的单声道信号及全频带经TCX解码的单声道信号。另外,可提供交叉路径136以用于使用通过低频带频谱-时间转换从TCX解码器及IGF处理器得出的信息来初始化低频带合成器。

[0116] 根据其他实施例,音频解码器包括:频域解码器106;用于使用频域解码器106的输出及第二多声道信息22、24生成第二多声道表示116的第二联合多声道解码器110;以及用于将第一声道信号及第二声道信号与第二多声道表示116进行组合以获得经解码的音频信号118的第一组合器112,其中第二联合多声道解码器不同于第一联合多声道解码器。因此,音频解码器可在使用LPD的参数化多声道解码或频域解码之间切换。已关于先前附图详细地描述了此方法。

[0117] 根据其他实施例,分析滤波器组144包括DFT以将单声道信号转换成频谱表示145,且其中全频带合成处理器148包括IDFT以将频谱表示145转换成第一声道信号及第二声道

信号。此外,分析滤波器组可对经DFT转换的频谱表示145应用窗口,以使得先前帧的频谱表示的右边部分及当前帧的频谱表示的左边部分重叠,其中先前帧和当前帧是连续的。换言之,淡入淡出可从一个DFT区块应用至另一区块以执行连续的DFT区块之间的平滑过渡和/或减少区块伪声。

[0118] 根据其他实施例,多声道解码器146用于从单声道信号获得第一声道信号及第二声道信号,其中单声道信号为多声道信号的中间信号,且其中多声道解码器146用于获得M/S多声道经解码的音频信号,其中多声道解码器用于从多声道信息计算侧信号。此外,多声道解码器146可用于自M/S多声道经解码的音频信号计算L/R多声道经解码的音频信号,其中多声道解码器146可使用多声道信息及侧信号来计算用于低频带的L/R多声道经解码的音频信号。额外地或可选地,多声道解码器146可从中间信号计算经预测的侧信号,且其中多声道解码器还用于使用经预测的侧信号及多声道信息的ILD值来计算用于高频带的L/R多声道经解码的音频信号。

[0119] 此外,多声道解码器146还可用于对L/R经解码的多声道音频信号执行复数运算,其中多声道解码器可使用经编码的中间信号的能量及经解码的L/R多声道音频信号的能量来计算复数运算的幅度以获得能量补偿。此外,多声道解码器用于使用多声道信息的IPD值计算复数运算的相位。在解码之后,经解码的多声道信号的能量、水平或相位可不同于经解码的单声道信号。因此,可确定复数运算,以使得多声道信号的能量、水平或相位被调整至经解码的单声道信号的值。此外,可使用(例如)来自在编码器侧所计算的多声道信息的经计算的IPD参数将相位调整至编码之前的多声道信号的相位的值。此外,经解码的多声道信号的人类感知可适于编码之前的原始多声道信号的人类感知。

[0120] 图20展示用于编码多声道信号的方法2000的流程图的示意性说明。该方法包括:对多声道信号进行降混以获得降混信号的步骤2050;编码降混信号的步骤2100,其中降混信号具有低频带和高频带,其中线性预测域核心编码器用于施加带宽扩展处理以用于对高频带进行参数化编码;生成多声道信号的频谱表示的步骤2150;以及处理包括多声道信号的低频带及高频带的频谱表示以生成多声道信息的步骤2200。

[0121] 图21展示对经编码的音频信号进行解码的方法2100的流程图的示意性说明,经编码的音频信号包括经核心编码的信号、带宽扩展参数以及多声道信息。该方法包括:对经核心编码的信号进行解码以生成单声道信号的步骤2105;将单声道信号转换成频谱表示的步骤2110;从单声道信号的频谱表示及多声道信息生成第一声道频谱及第二声道频谱的步骤2115;以及对第一声道频谱进行合成滤波以获得第一声道信号并对第二声道频谱进行合成滤波以获得第二声道信号的步骤2120。

[0122] 其他实施例描述如下。

[0123] 比特流语法变化

[0124] USAC规范[1]在章节5.3.2辅助有效载荷中的表23应修改如下:

[0125] 表1-UsacCoreCoderData()的语法

语法	比特数	助记符
<pre> UsacCoreCoderData (nrChannels, indepFlag) { for (ch=0; ch < nrChannels; ch++) { core_mode[ch]; } if (nrChannels == 2) { StereoCoreToolInfo(core_mode); } for (ch=0; ch<nrChannels; ch++) { if (core_mode[ch] == 1) { if(ch==1 && core_mode[1] == core_mode[0]){ lpd_stereo_stream(); }else{ lpd_channel_stream(indepFlag); } } else { if((nrChannels==1) ((core_mode[0]!= core_mode[1]))){ tns_data_present[ch]; } fd_channel_stream(common_window, common_tw, tns_data_present[ch], noiseFilling, indepFlag); } } } </pre>	<p>1</p> <p>1</p>	<p>uimsbf</p> <p>uimsbf</p>

[0127] 应添加下表:

[0128] 表1-lpd_stereo_stream() 的语法

语法	比特数	助记符
lpd_stereo_stream (indepFlag)		
{		
for (l=0,n=0;l<ccfl;l+=M,n++) {		
res_mode	1	uimsbf
q_mode	1	uimsbf ,
ipd_mode	2	uimsbf
pred_mode	1	uimsbf
cod_mode	2	uimsbf
nbands=band_config (N, res_mode)		
ipd_band_max=max_band[res_mode][ipd_mode]		
cod_band_max=max_band[res_mode][cod_mode]		
cod_L=2* (band_limits[cod_band_max]-1)		
for (k=1;k>=0;k--) {		
if (q_mode==0 k == 1) {		
for (b=0;b< nbands;b++) {		
ild_idx[2n+k][b]	5	
}		
for (b=0;b< ipd_band_max;b++) {		
ipd_idx[2n+k][b]	3	
}		
if (pred_mode==1) {		
for (b=cod_band_max;b< nbands;b++) {		
pred_gain_idx[2n+k][b]	3	
}		
}		
}		
if (cod_mode==1) {		
cod_gain_idx[2n+k]		
for (i=0;i< cod_L/8;i++) {		
code_book_indices (i, 1, 1)		
}		
}		
}	7	
}		

[0130] 应在章节6.2 USAC有效载荷中添加以下有效载荷描述。

[0131] 6.2.x lpd_stereo_stream()

[0132] 详细解码程序在7.x LPD立体声解码章节中描述。

[0133] 术语及定义

[0134] lpd_stereo_stream() 用以关于LPD模式解码立体声数据的数据元素

[0135] res_mode 指示参数频带的频率分辨率的旗标。

[0136] q_mode 指示参数频带的时间分辨率的旗标。

[0137] ipd_mode 定义用于IPD参数的参数频带的最大值的比特字段。

[0138] pred_mode 指示是否使用预测的旗标。

[0139] cod_mode 定义侧信号被量化的参数频带的最大值的比特字段。

[0140] Ild_idx[k][b] 帧k及频带b的ILD参数索引。

[0141] Ipd_idx[k][b] 帧k及频带b的IPD参数索引。

- [0142] pred_gain_idx[k][b] 帧k及频带b的预测增益索引。
- [0143] cod_gain_idx 经量化的侧信号的全局增益索引。
- [0144] 协助元素
- [0145] ccf1 核心码帧长度。
- [0146] M如表7.x.1中所定义的立体声LPD帧长度。
- [0147] band_config() 传回经编码的参数频带的数目的函数。函数定义于7.x中
- [0148] band_limits() 传回经编码的参数频带的数目的函数。函数定义于7.x中
- [0149] max_band() 传回经编码的参数频带的数目的函数。函数定义于7.x中
- [0150] ipd_max_band() 传回经编码的参数频带的数目的函数。函数
- [0151] cod_max_band() 传回经编码的参数频带的数目的函数。函数
- [0152] cod_L 用于经解码的侧信号的DFT线的数目。
- [0153] 解码过程
- [0154] LPD立体声编码
- [0155] 工具描述
- [0156] LPD立体声为离散M/S立体声编码,其中由单声道LPD核心编码器对中间声道进行编码且在DFT域中对侧信号进行编码。经解码的中间信号从LPD单声道解码器输出且然后由LPD立体声模块来处理。立体声解码在DFT域中进行,L及R声道在DFT域中被解码。两个经解码的声道变换回到时域且可然后在此域中与来自FD模式的经解码的声道组合。FD编码模式使用其自身立体声工具,即具有或不具有复数预测的离散立体声。
- [0157] 数据元素
- [0158] res_mode 指示参数频带的频率分辨率的旗标。
- [0159] q_mode 指示参数频带的时间分辨率的旗标。
- [0160] ipd_mode 定义用于IPD参数的参数频带的最大值的比特字段。
- [0161] pred_mode 指示是否使用预测的旗标。
- [0162] cod_mode 定义侧信号被量化的参数频带的最大值的比特字段。
- [0163] Ild_idx[k][b] 帧k及频带b的ILD参数索引。
- [0164] Ipd_idx[k][b] 帧k及频带b的IPD参数索引。
- [0165] pred_gain_idx[k][b] 帧k及频带b的预测增益索引。
- [0166] cod_gain_idx 经量化的侧信号的全局增益索引。
- [0167] 协助元素
- [0168] ccf1 核心码帧长度。
- [0169] M如表7.x.1中所定义的立体声LPD帧长度。
- [0170] band_config() 传回经编码的参数频带的数目的函数。函数定义于7.x中
- [0171] band_limits() 传回经编码的参数频带的数目的函数。函数定义于7.x中
- [0172] max_band() 传回经编码的参数频带的数目的函数。函数定义于7.x中
- [0173] ipd_max_band() 传回经编码的参数频带的数目的函数。函数
- [0174] cod_max_band() 传回经编码的参数频带的数目的函数。函数
- [0175] cod_L 用于经解码的侧信号的DFT线的数目。
- [0176] 解码过程

[0177] 在频域中执行立体声解码。立体声解码充当LPD解码器的后处理。其从LPD解码器接收单声道中间信号的合成。然后，在频域中解码或预测侧信号。然后在时域中重新合成之前在频域中重建声道频谱。独立于LPD模式中所使用的编码模式，立体声LPD对等于ACELP帧的大小的固定帧大小起作用。

[0178] 频率分析

[0179] 从长度为M的经解码的帧x来计算帧索引i的DFT频谱。

$$[0180] \quad X_i[k] = \sum_{n=0}^{N-1} w[n] \cdot x[i \cdot M + n - L] \cdot e^{-2\pi jkn/N}$$

[0181] 其中N为信号分析的大小，w为分析窗口且x为来自LPD解码器的经延迟DFT的重叠大小L的帧索引i处的经解码的时间信号。M等于FD模式中所使用的采样率下的ACELP帧的大小。N等于立体声LPD帧大小加DFT的重叠大小。大小视所使用LPD版本而定，如表7.x.1中所报告。

[0182] 表7.x.1-立体声LPD的DFT及帧大小

[0183] LPD版本	DFT大小N	帧大小M	重叠大小L
0	336	256	80
1	672	512	160

[0184] 窗口w为正弦窗口，其被定义为：

$$[0185] \quad w[n] = \begin{cases} \sin\left(\frac{\pi}{2L}\left(n + \frac{1}{2}\right)\right) & \text{对于 } 0 \leq n < L \\ 1 & \text{对于 } L \leq n < M \\ \sin\left(\frac{\pi}{2L}\left(L + n + \frac{1}{2}\right)\right) & \text{对于 } M \leq n < M + L \end{cases}$$

[0186] 参数频带的配置

[0187] DFT频谱被划分为所谓的参数频带的非重叠频带。频谱的分割是不均匀的并模仿听觉频率分解。频谱的两个不同划分可能具有遵照大致两倍或四倍的等效矩形带宽 (ERB) 的带宽。

[0188] 频谱分割是通过数据元素res_mod来选择的且由以下伪码定义：

```

function nbands=band_config(N,res_mod)
band_limits[0]=1;
nbands=0;
while(band_limits[nbands++]<(N/2)){
    if(stereo_lpd_res==0)
        band_limits[nbands]=band_limits_erb2[nbands];
[0189]     else
        band_limits[nbands]=band_limits_erb4[nbands];
}
nbands--;
band_limits[nbands]=N/2;
return nbands

```

[0190] 其中nbands为参数频带的总数目且N为DFT分析窗口大小。表band_limits_erb2及band_limits_erb4在表7.x.2中定义。解码器可每两个立体声LPD帧自适应地改变频谱的参数频带的分辨率。

[0191] 表7.x.2-关于DFT索引k的参数频带极限

参数频带索引 <i>b</i>	<i>band_limits_erb2</i>	<i>band_limits_erb4</i>
0	1	1
1	3	3
2	5	7
3	7	13
4	9	21
[0192] 5	13	33
6	17	49
7	21	73
8	25	105
9	33	177
10	41	241
11	49	337
12	57	
13	73	
14	89	
[0193] 15	105	
16	137	
17	177	
18	241	
19	337	

[0194] 用于IPD的参数频带的最大数目在2比特字段ipd_mod数据元素内发送：

[0195] ipd_max_band=max_band[res_mod][ipd_mod]

[0196] 用于侧信号的编码的参数频带的最大数目在2比特字段cod_mod数据元素内发送：

[0197] $\text{cod_max_band} = \text{max_band}[\text{res_mod}][\text{cod_mod}]$

[0198] 表 $\text{max_band}[\][\]$ 定义于表7.x.3中。

[0199] 然后,计算期望用于侧信号的经解码的线的数目:

[0200] $\text{cod_L} = 2 \cdot (\text{band_limits}[\text{cod_max_band}] - 1)$

[0201] 表7.x.3-用于不同码模式的频带的最大数目

[0202]

模式索引	$\text{max_band}[0]$	$\text{max_band}[1]$
0	0	0
1	7	4
2	9	5
3	11	6

[0203] 立体声参数的反量化

[0204] 立体声参数声道间等级差 (Interchannel Level Differences, ILD)、声道间相位差 (Interchannel Phase Differences, IPD) 以及预测增益将依据旗标 q_mode 而每一帧或每两帧地发送。若 q_mode 等于0,则在每一帧地更新参数。否则,仅针对USAC帧内的立体声LPD帧的奇数索引 i 更新参数值。USAC帧内的立体声LPD帧的索引 i 在LPD版本0中可在0与3之间而在LPD版本1中可在0与1之间。

[0205] ILD被解码如下:

[0206] $\text{ILD}_i[b] = \text{ild_q}[\text{ild_idx}[i][b]]$, 对于 $0 \leq b < \text{nbands}$

[0207] 针对前 ipd_max_band 个频带解码IPD:

[0208] $\text{IPD}_i[b] = \frac{\pi}{4} \cdot \text{ipd_idx}[i][b] - \pi$, 对于 $0 \leq b < \text{ipd_max_band}$

[0209] 预测增益仅在 pred_mode 旗标设定为一时被解码。经解码的增益因而为:

[0210]
$$\text{pred_gain}_i[b] = \begin{cases} 0 & \text{, 对于 } 0 \leq b < \text{cod_max_band} \\ \text{res_pred_gain_q}[\text{pred_gain_idx}[i][b]] & \text{, 对于 } \text{cod_max_band} \leq b < \text{nbands} \end{cases}$$

[0211] 若 pred_mode 等于零,则所有增益被设定为零。

[0212] 不依赖于 q_mode 的值,若 code_mode 为非零值,则侧信号的解码每一帧地执行。其首先解码全局增益:

[0213] $\text{cod_gain}_i = 10^{\text{cod_gain_idx}[i]} \cdot 20 \cdot 127/90$

[0214] 侧信号的经解码的形状为USAC规范[1]中在章节中所描述的AVQ的输出。

[0215] $S_i[1+8k+n] = kv[k][0][n]$, 对于 $0 \leq n < 8$ 和 $0 \leq k < \frac{\text{cod_L}}{8}$

[0216] 表7.x.4-反量化表 $\text{ild_q}[\]$

[0217]

索引	输出	索引	输出
0	-50	16	2
1	-45	17	4
2	-40	18	6
3	-35	19	8
4	-30	20	10
5	-25	21	13

6	-22	22	16
7	-19	23	19
8	-16	24	22
9	-13	25	25
10	-10	26	30
11	-8	27	35
12	-6	28	40
13	-4	29	45
14	-2	30	50
15	0	31	预留

[0218] 表7.x.5-反量化表res_pres_gain_q[]

索引	输出
0	0
1	0.1170
2	0.2270
3	0.3407
4	0.4645
5	0.6051
6	0.7763
7	1

[0220] 反声道映射

[0221] 首先,中间信号X及侧信号S被如下地转换至左声道L及右声道R:

[0222] $L_i[k] = X_i[k] + gX_i[k]$, 对于 $\text{band_limits}[b] \leq k < \text{band_limits}[b+1]$,

[0223] $R_i[k] = X_i[k] - gX_i[k]$, 对于 $\text{band_limits}[b] \leq k < \text{band_limits}[b+1]$,

[0224] 其中每个参数频带的增益g是从ILD参数得出的:

[0225] $g = \frac{c-1}{c+1}$, 其中 $c = 10^{ILD_i[b]/20}$

[0226] 对于低于cod_max_band的参数频带,用经解码的侧信号来更新两个声道:

[0227] $L_i[k] = L_i[k] + \text{cod_gain}_i \cdot S_i[k]$, 对于 $0 \leq k < \text{band_limits}[\text{cod_max_band}]$,

[0228] $R_i[k] = R_i[k] - \text{cod_gain}_i \cdot S_i[k]$, 对于 $0 \leq k < \text{band_limits}[\text{cod_max_band}]$,

[0229] 对于较高参数频带,对侧信号进行预测且声道更新如下:

[0230] $L_i[k] = L_i[k] + \text{cod_pred}_i[b] \cdot X_{i-1}[k]$, 对于 $\text{band_limits}[b] \leq k < \text{band_limits}[b+1]$,

[0231] $R_i[k] = R_i[k] - \text{cod_pred}_i[b] \cdot X_{i-1}[k]$, 对于 $\text{band_limits}[b] \leq k < \text{band_limits}[b+1]$,

[0232] 最终,声道以复值倍增,其目标为恢复信号的原始能量及声道间相位:

[0233] $L_i[k] = a \cdot e^{j2\pi\beta} \cdot L_i[k]$

[0234] $R_i[k] = a \cdot e^{j2\pi\beta} \cdot R_i[k]$

[0235] 其中

$$[0236] \quad \alpha = \sqrt{2 \cdot \frac{\sum_{k=\text{band_limits}[b]}^{\text{band_limits}[b+1]} X_i^2[k]}{\sum_{k=\text{band_limits}[b]}^{\text{band_limits}[b+1]-1} L_i^2[k] + \sum_{k=\text{band_limits}[b]}^{\text{band_limits}[b+1]-1} R_i^2[k]}}$$

[0237] 其中c被约束于-12dB及12dB。

[0238] 且其中

[0239] $\beta = \text{atan2}(\sin(\text{IPD}_i[b]), \cos(\text{IPD}_i[b]) + c)$

[0240] 其中 $\text{atan2}(x, y)$ 为x相对于y的四象限反正切。

[0241] 时域合成

[0242] 从两个经解码的频谱L及R,通过反DFT来合成两个时域信号l及r:

$$[0243] \quad l_i[n] = \sum_{k=0}^{N-1} L_i[k] \cdot e^{\frac{2\pi jkn}{N}}, \text{ 对于 } 0 \leq n < N$$

$$[0244] \quad r_i[n] = \sum_{k=0}^{N-1} R_i[k] \cdot e^{\frac{2\pi jkn}{N}}, \text{ 对于 } 0 \leq n < N$$

[0245] 最终,重叠相加法操作允许重建M个样本的帧:

$$[0246] \quad l[i \cdot M + n - L] = \begin{cases} l_{i-1}[M+n] \cdot w[L-1-n] + l_i[n] \cdot w[n], & \text{对于 } 0 \leq n < L \\ l_i[n] & \text{, 对于 } L \leq n < M \end{cases}$$

$$[0247] \quad r[i \cdot M + n - L] = \begin{cases} r_{i-1}[M+n] \cdot w[L-1-n] + r_i[n] \cdot w[n], & \text{对于 } 0 \leq n < L \\ r_i[n] & \text{, 对于 } L \leq n < M \end{cases}$$

[0248] 后处理

[0249] 巴斯后处理被分别地应用于两个声道。处理用于两个声道,与[1]的章节7.17中所描述的相同。

[0250] 应理解,在本说明书中,线上的信号有时以用于线的附图标记来命名或有时以已经属于线的附图标记本身来指示。因此,标记为使得具有某一信号的线指示信号本身。线在固线式实施中可为实体线。然而,在计算机化实施中,实体线并不存在,但由线表示的信号将从一个计算模块传输至另一计算模块。

[0251] 尽管已在区块表示实际或逻辑硬件组件的框图的上下文中描述本发明,但本发明亦可由计算机实施方法来实施。在后一情况下,区块表示对应方法步骤,其中这些步骤代表由对应逻辑或实体硬件区块执行的功能性。

[0252] 尽管已在设备的上下文中描述一些方面,但显然,这些方面也表示对应方法的描述,其中区块或装置对应于方法步骤或方法步骤的特征。类似地,方法步骤的上下文中所描述的方面也表示对应区块或项目或对应设备的特征的描述。可由(或使用)硬件设备(类似于(例如)微处理器、可程序化计算机或电子电路)来执行方法步骤中的一些或全部。在一些实施例中,可由此设备来执行最重要方法步骤中的某一个或多个。

[0253] 本发明的经传输或经编码的信号可被存储于数字存储媒体上或可在如无线传输媒体的传输媒体或如因特网的有线传输媒体上传输。

[0254] 依据某些实施要求,本发明的实施例可在硬件中或在软件中实施。可使用上面存

储有电子可读控制信号、与可程序化计算机系统协作(或能够协作)以使得执行各个方法的数字存储媒体(例如,软盘、DVD、Blu-Ray、CD、ROM、PROM以及EPROM、EEPROM或闪存)来执行实施。因此,数字存储媒体可以是计算机可读的。

[0255] 根据本发明的一些实施例包括具有电子可读控制信号的数据载体,其能够与可程序化计算机系统协作,从而执行本文中所描述的方法中的一个。

[0256] 通常,本发明的实施例可实施为具有程序代码的计算机程序产品,当计算机程序产品在计算机上执行时,程序代码操作性地用于执行方法中的一个。程序代码可(例如)存储于机器可读载体上。

[0257] 其他实施例包括存储于机器可读载体上的用于执行本文中所描述的方法中的一个的计算机程序。

[0258] 换言之,因此,本发明方法的实施例为计算机程序,其具有用于在计算机程序运行于计算机上时执行本文中所描述的方法中的一个的程序代码。

[0259] 因此,本发明方法的另一实施例为包括数据载体(或诸如数字存储媒体的非暂时性存储媒体,或计算机可读媒体),其包括记录于其上的用于执行本文中所描述的方法中的一个的计算机程序。数据载体、数字存储媒体或记录媒体通常为有形的和/或非暂时性的。

[0260] 因此,本发明方法的另一实施例为表示用于执行本文中所描述的方法中的一个的计算机程序的数据串流或信号的序列。数据串流或信号的序列可(例如)用于经由数据通信连接(例如,经由因特网)来传送。

[0261] 另一实施例包括处理构件,例如,经配置或经调适以执行本文中所描述的方法中的一个的计算机或可规划逻辑设备。

[0262] 另一实施例包括上面安装有用于执行本文中所描述的方法中的一个的计算机程序的计算机。

[0263] 根据本发明的另一实施例包括用于将用于执行本文中所描述的方法中的一个的计算机程序进行传送(例如,用电子地或光学地)至接收器的设备或系统。接收器可(例如)为计算机、行动装置、内存装置等。设备或系统可(例如)包括用于将计算机程序传送至接收器的文件服务器。

[0264] 在一些实施例中,可程序化逻辑设备(例如,现场可程序化门阵列)可用以执行本文中所描述的方法的功能性中的一些或全部。在一些实施例中,现场可规划门阵列可与微处理器合作以便执行本文中所描述的方法中的一个。通常,优选地由任何硬设备来执行方法。

[0265] 上述实施例仅说明本发明的原理。应理解,本文中所描述的配置及细节的修改及变化对于本领域技术人员是显而易见的。因此,其仅意在由所附专利权利要求的范畴限制,而不是由借助于本文中的实施例的描述及解释所呈现的特定细节限制。

[0266] 参考文献

[0267] [1] ISO/IEC DIS 23003-3, Usac

[0268] [2] ISO/IEC DIS 23008-3, 3D音频

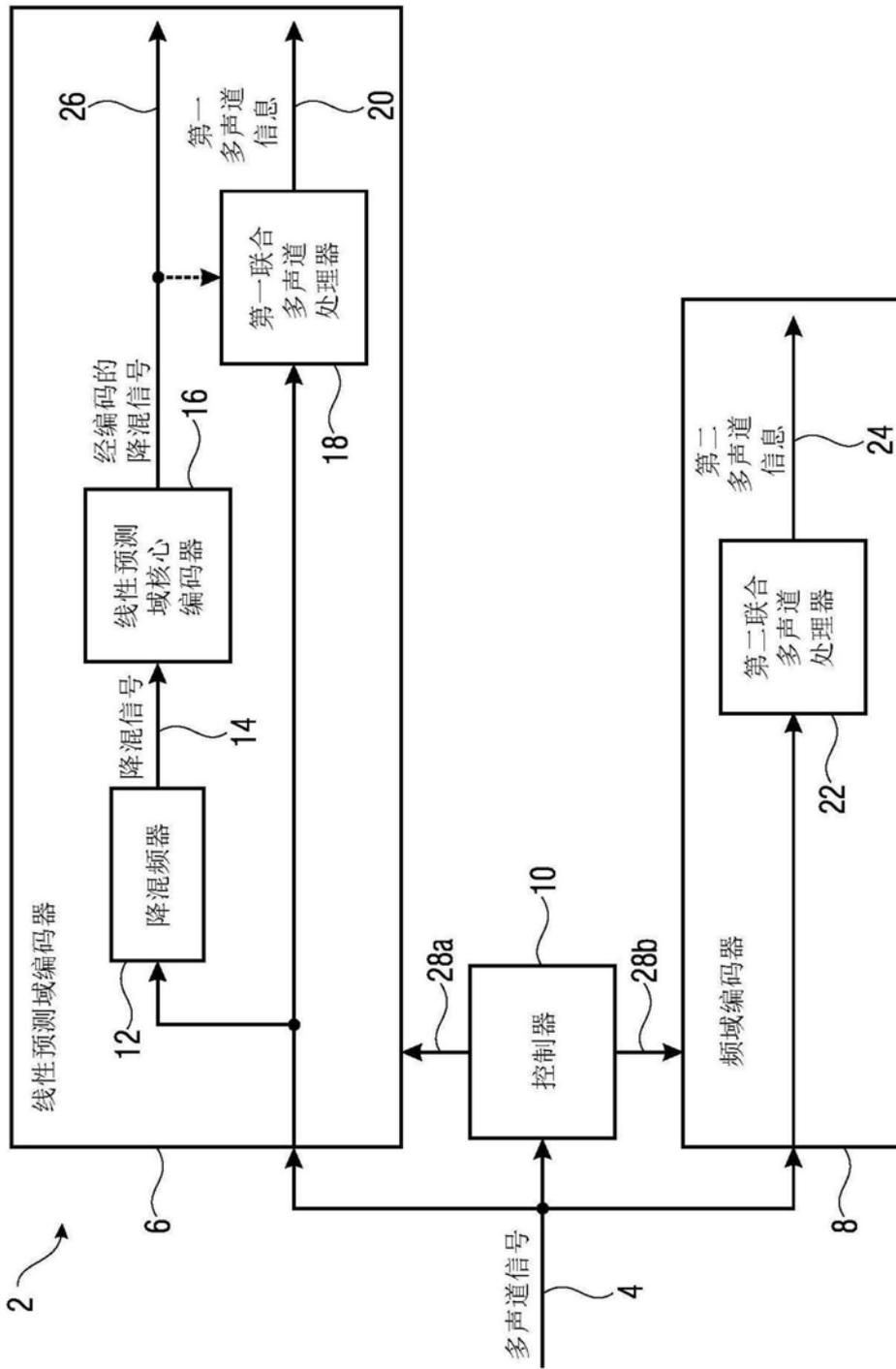


图1

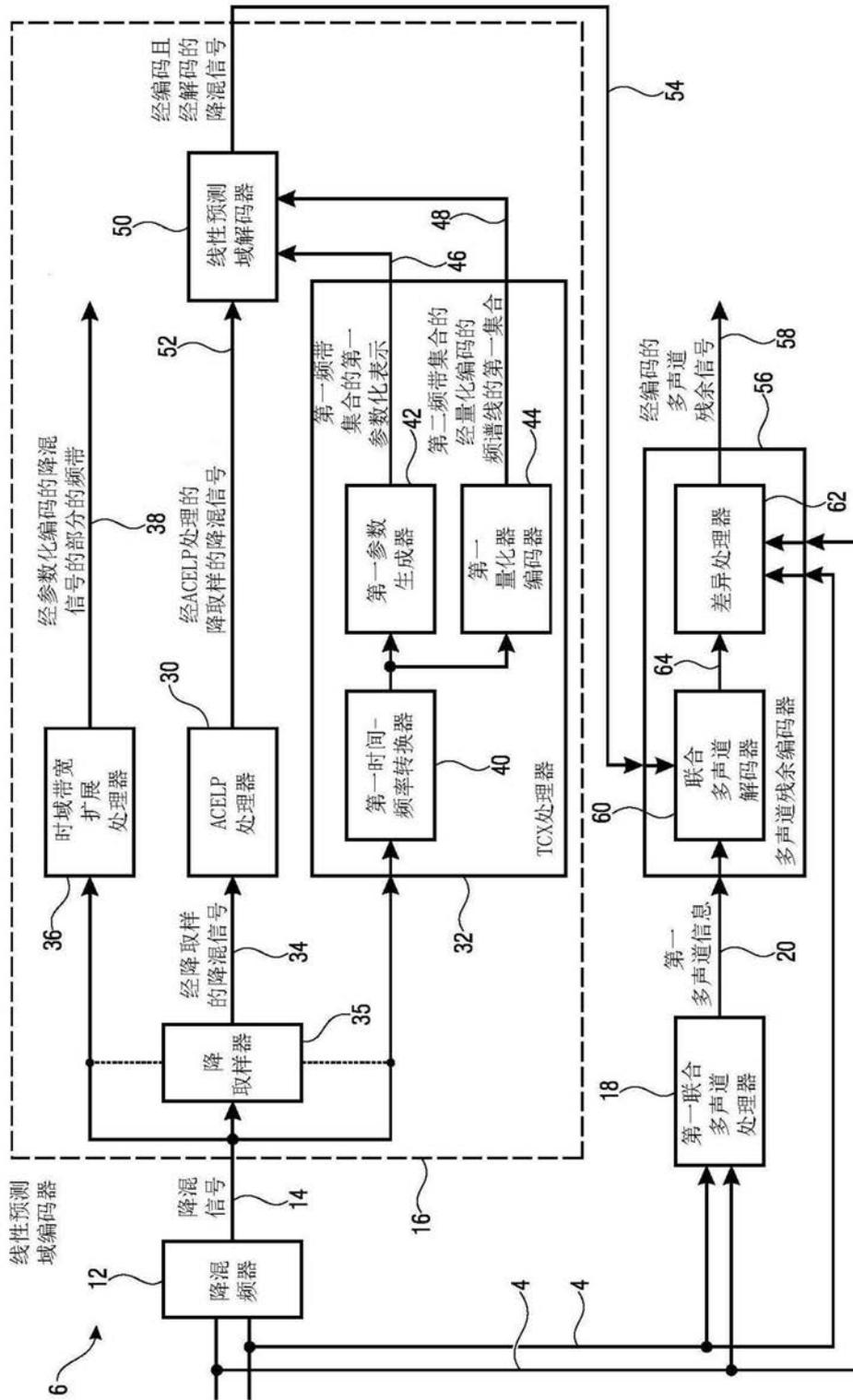


图2

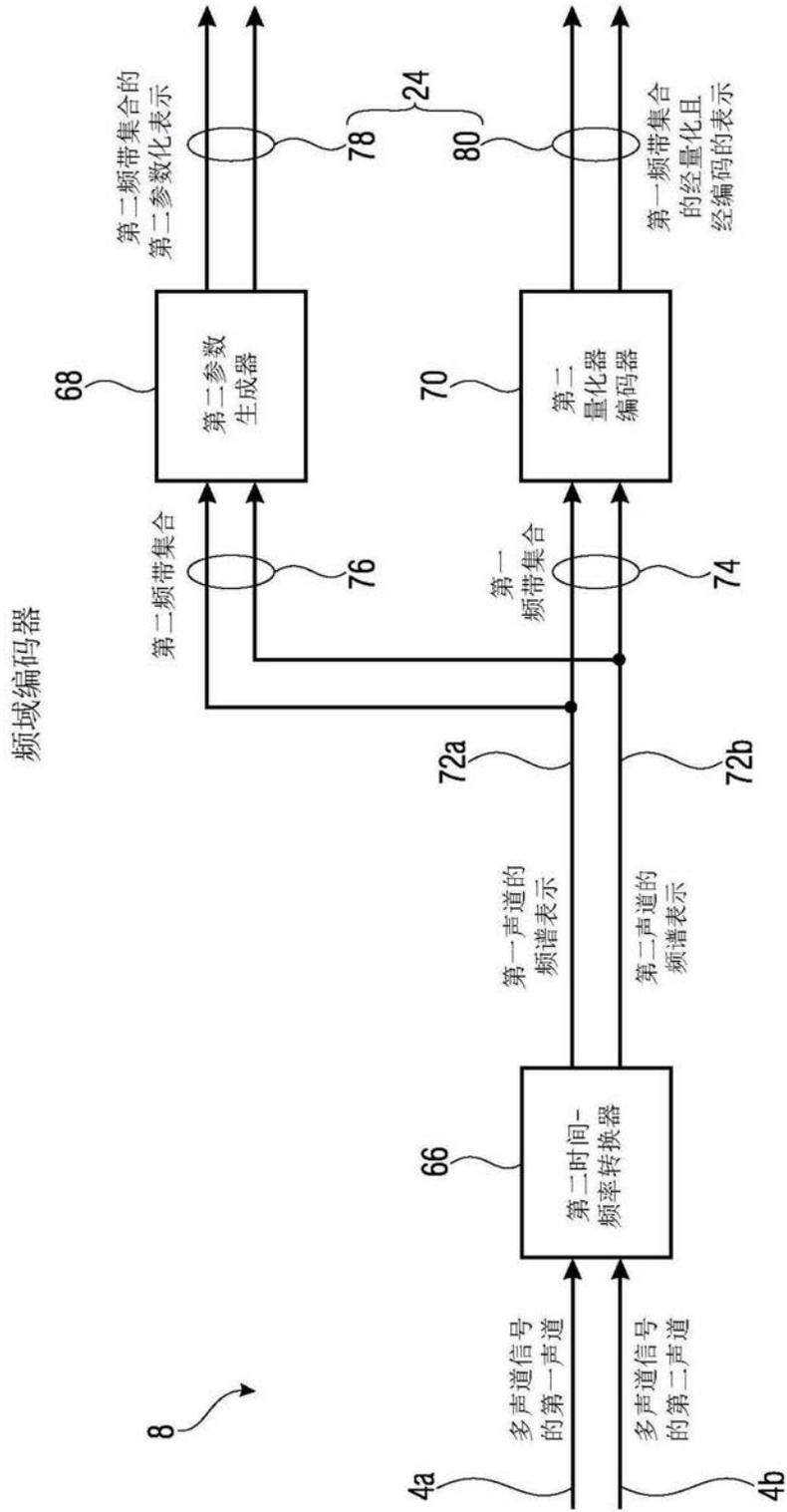


图3

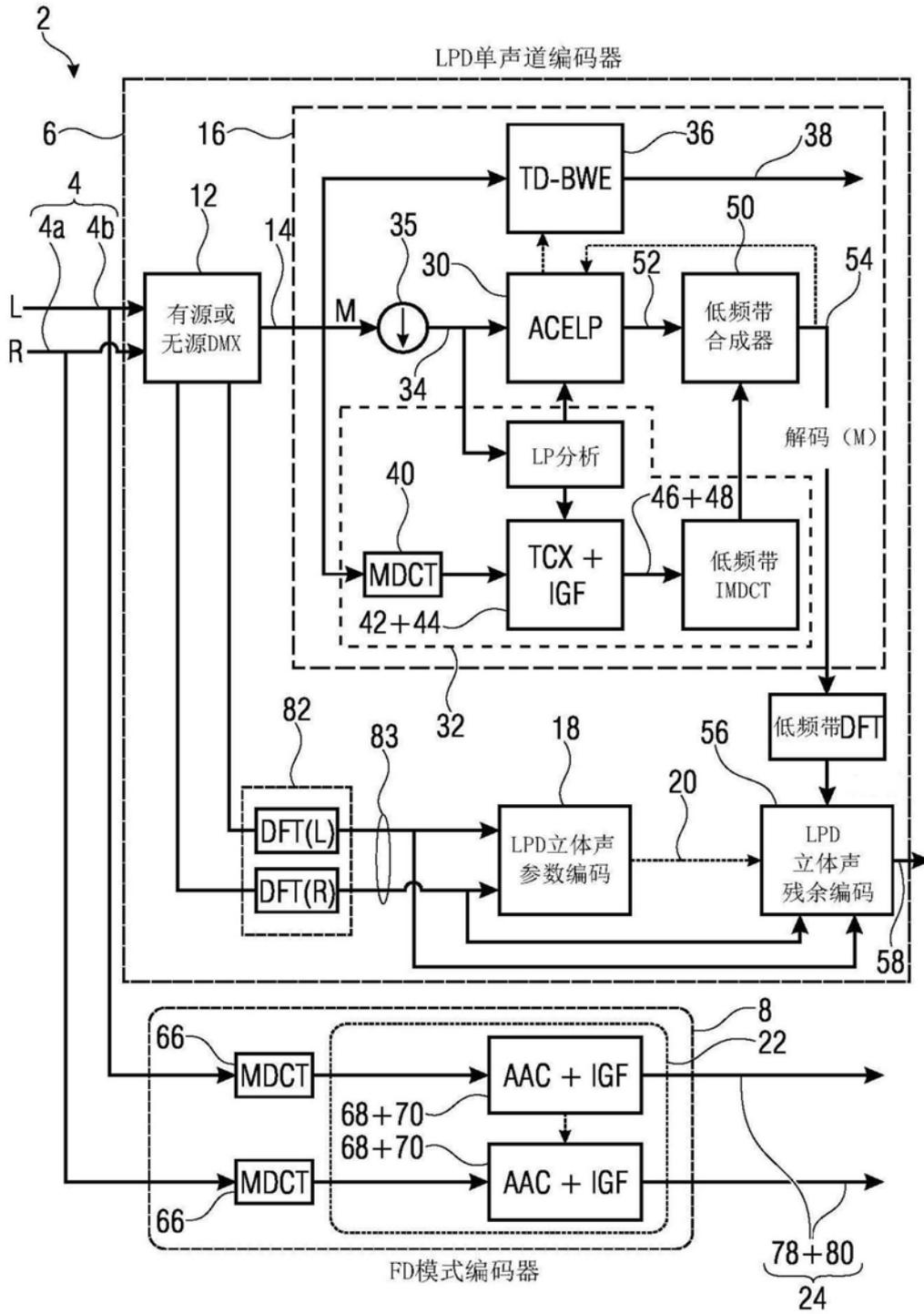


图4

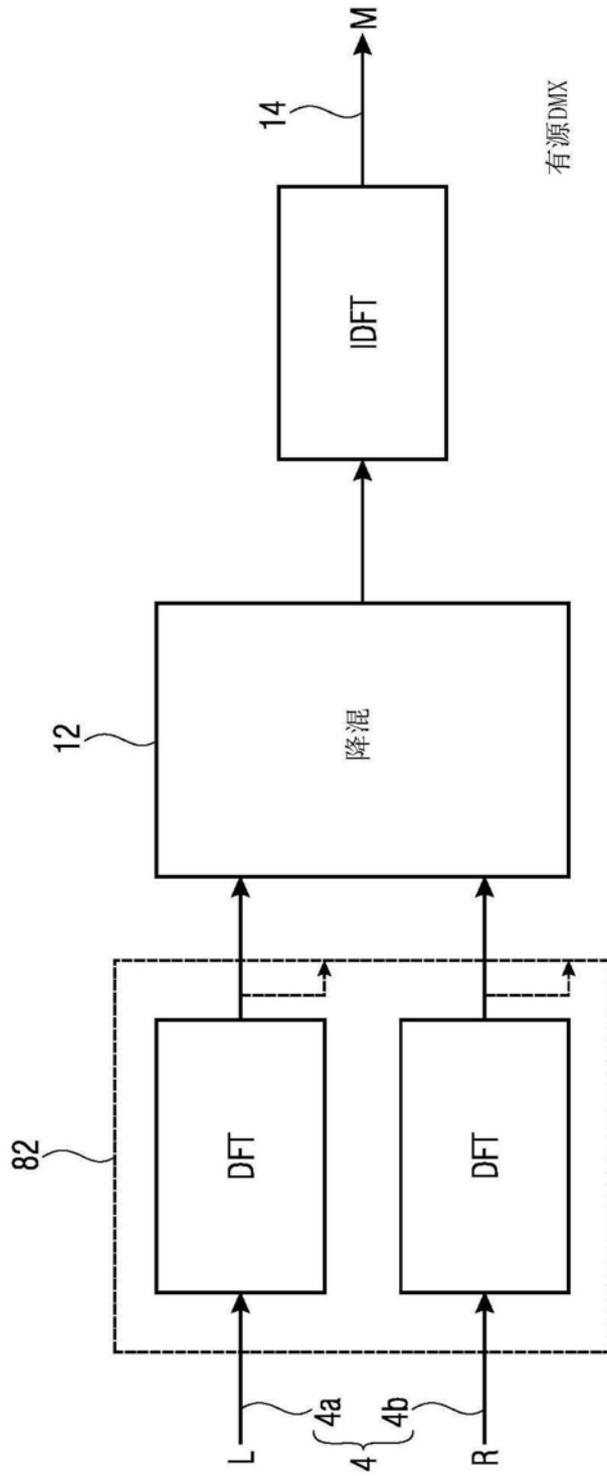


图5A

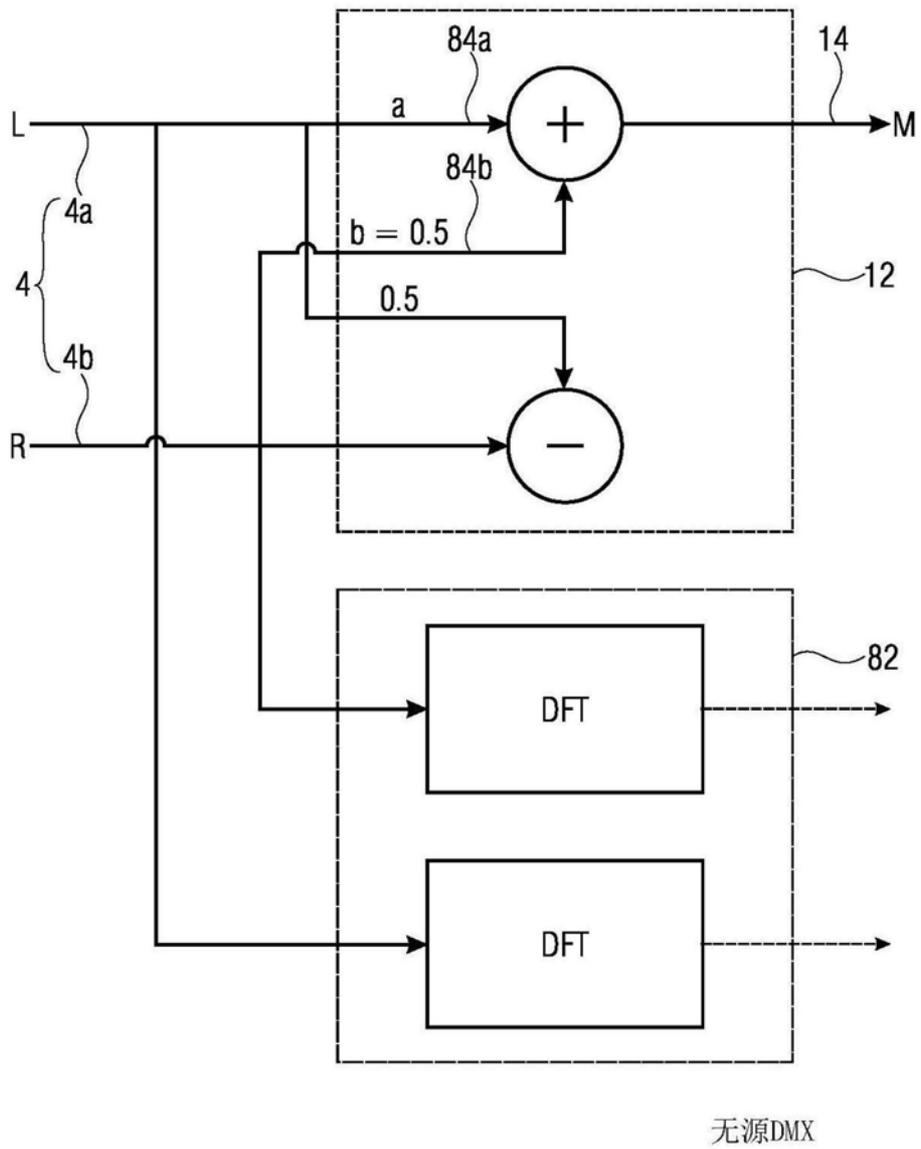


图5B

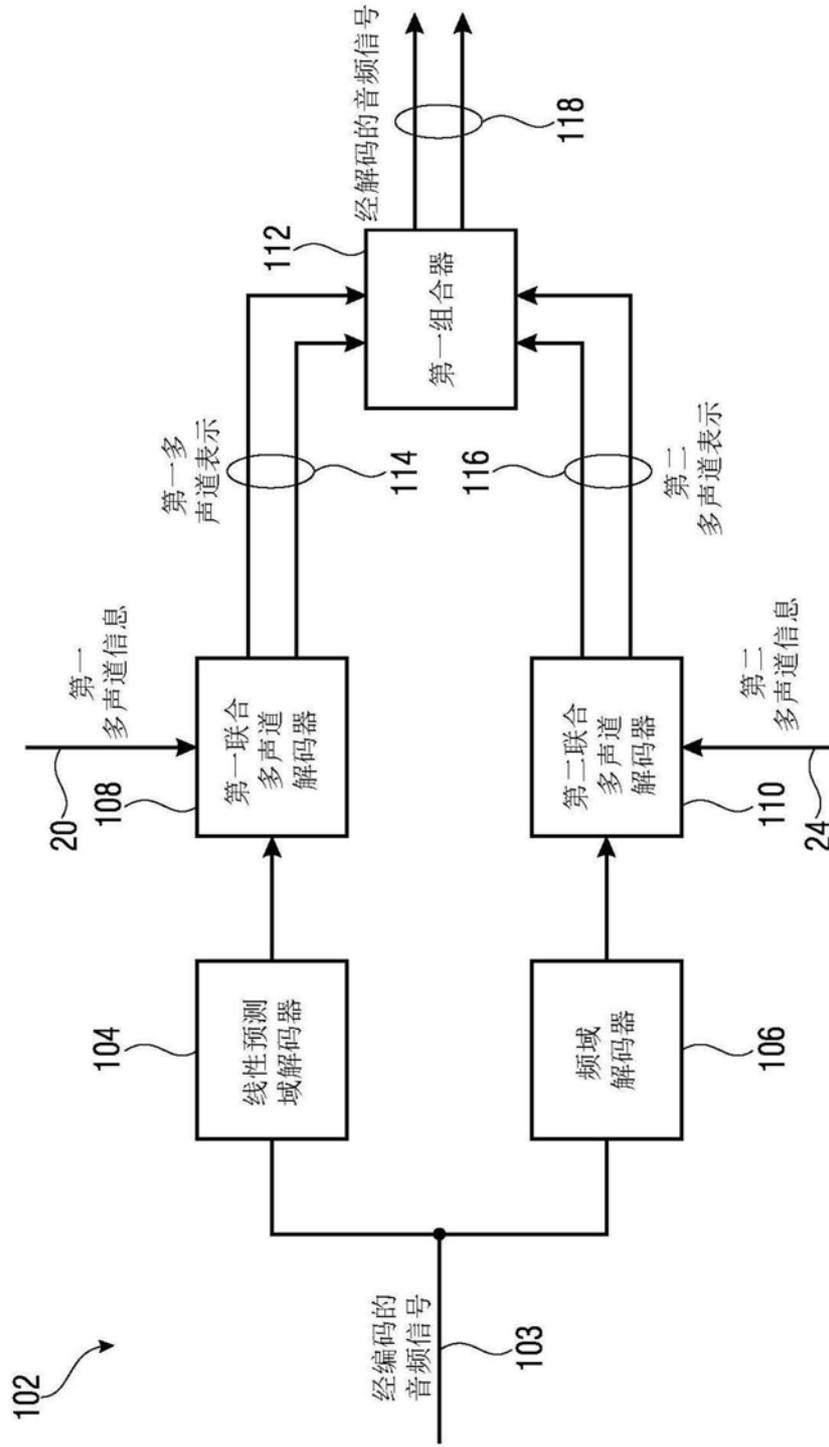


图6

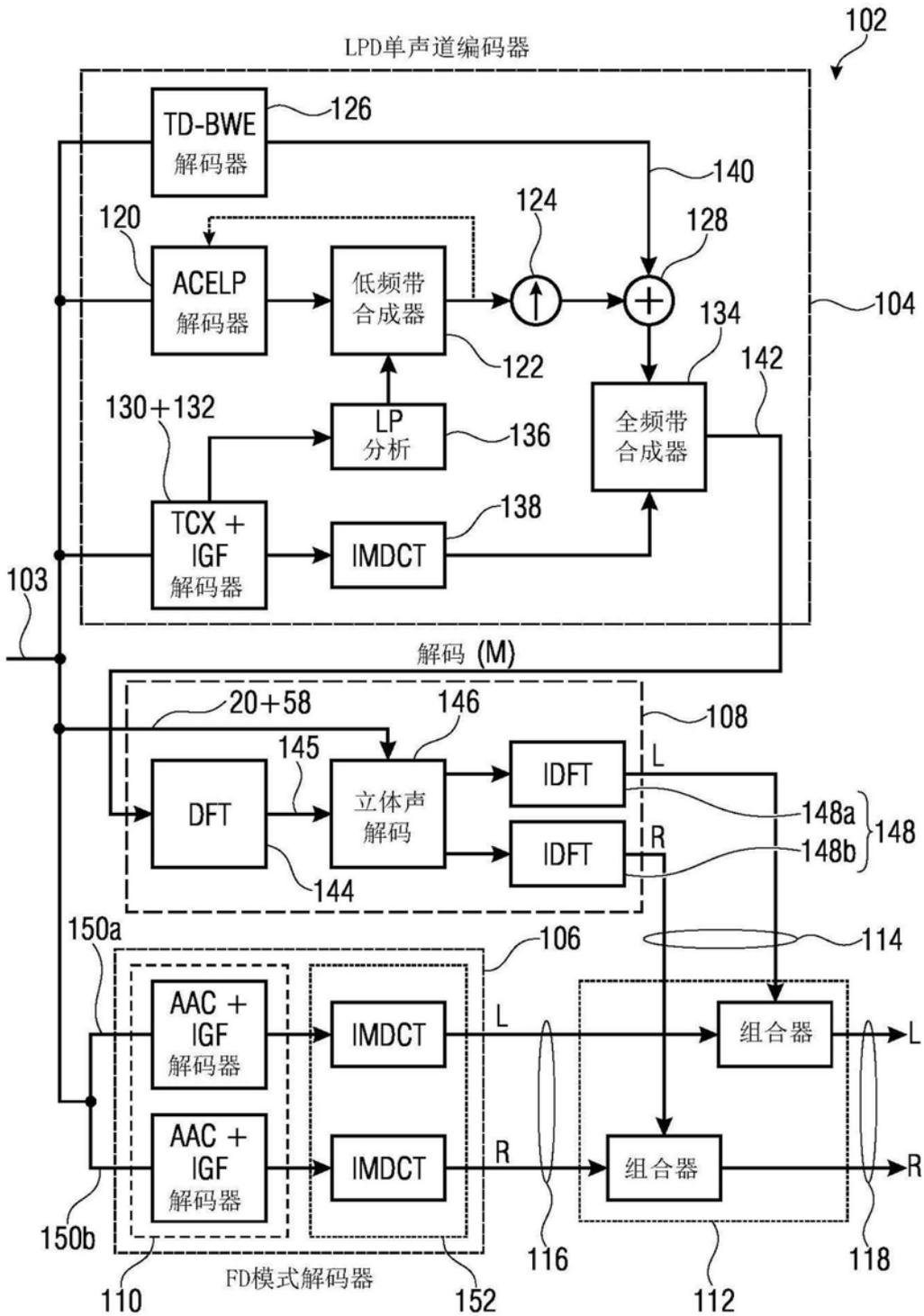


图7

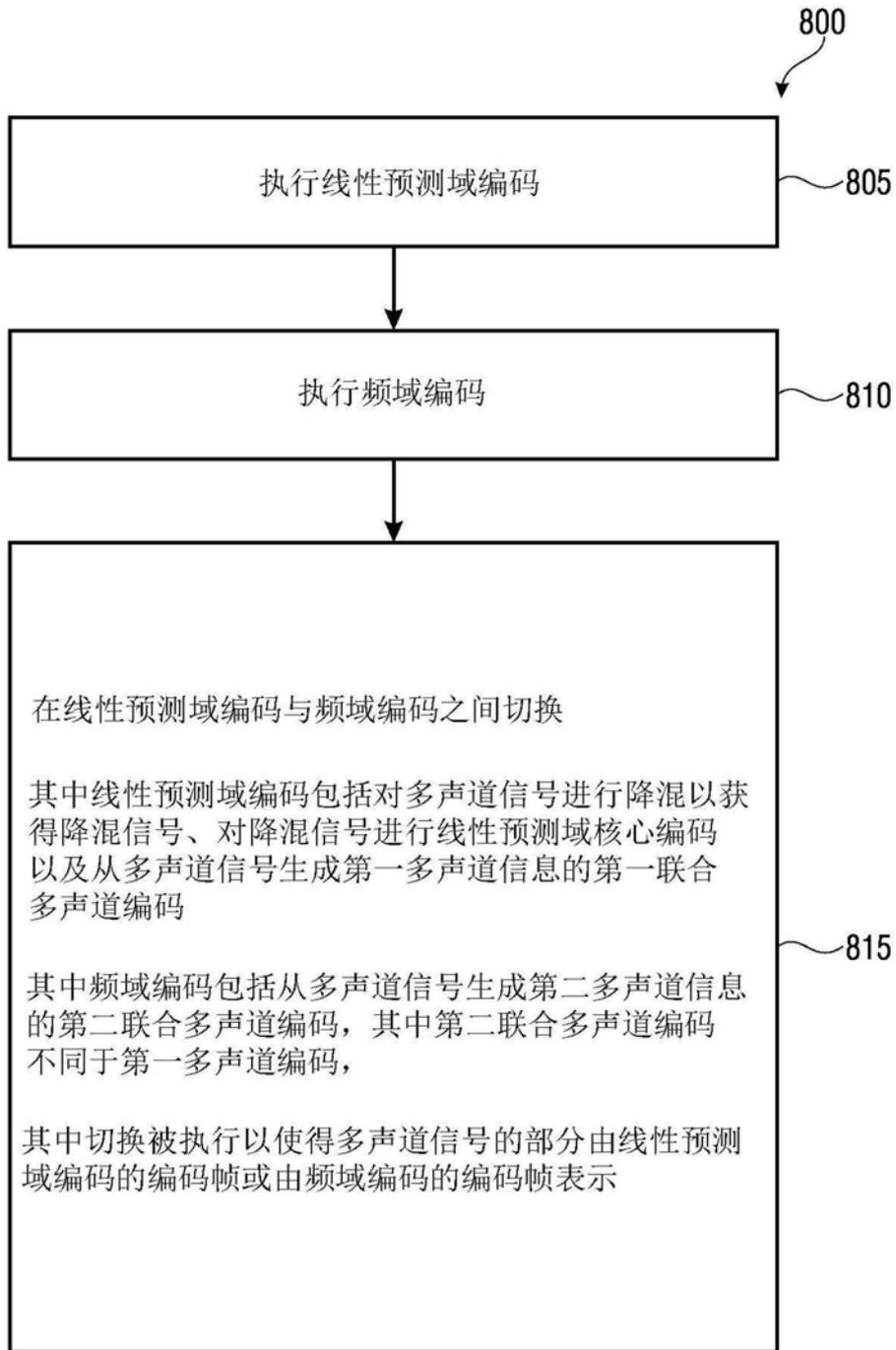


图8

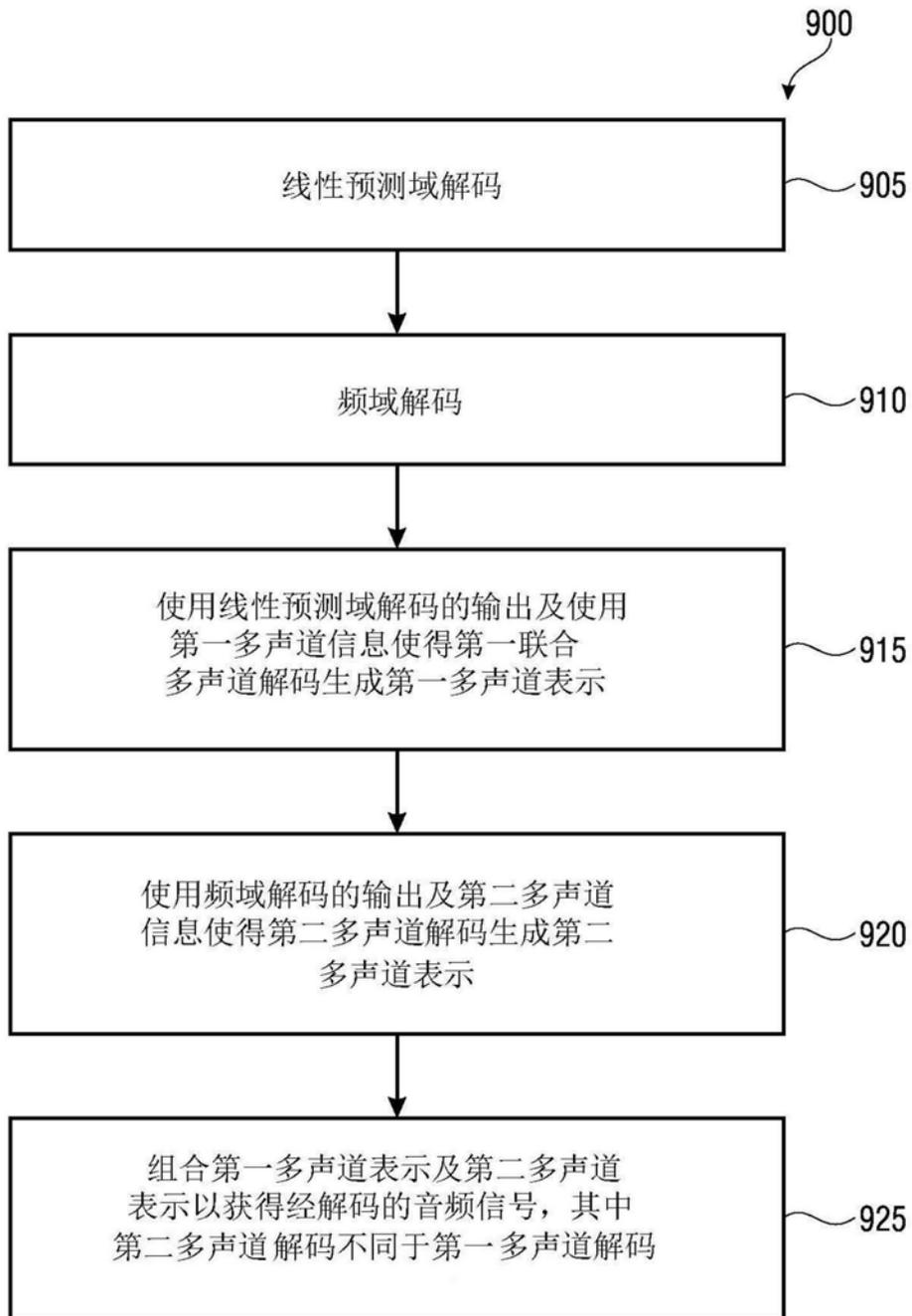


图9

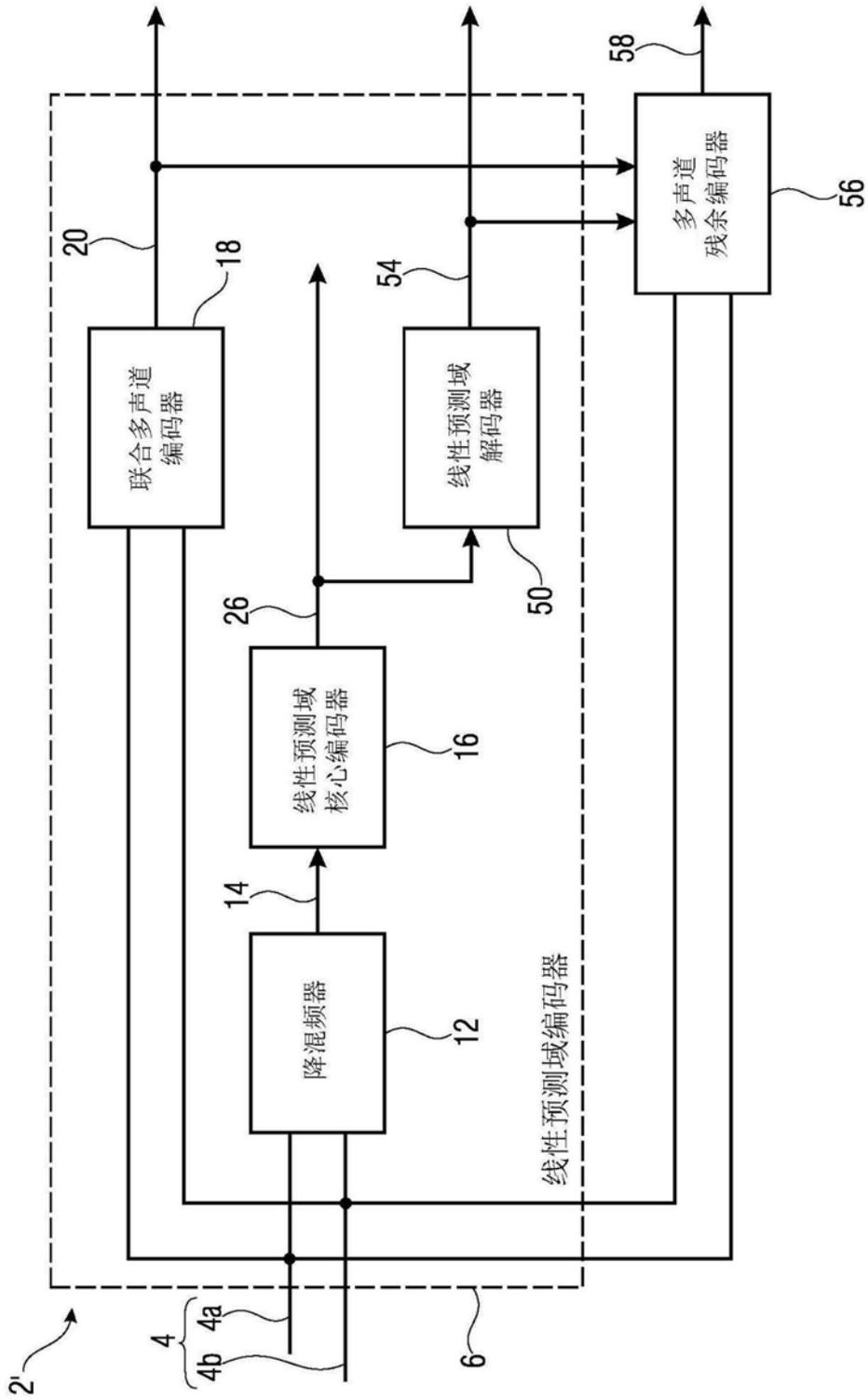


图10

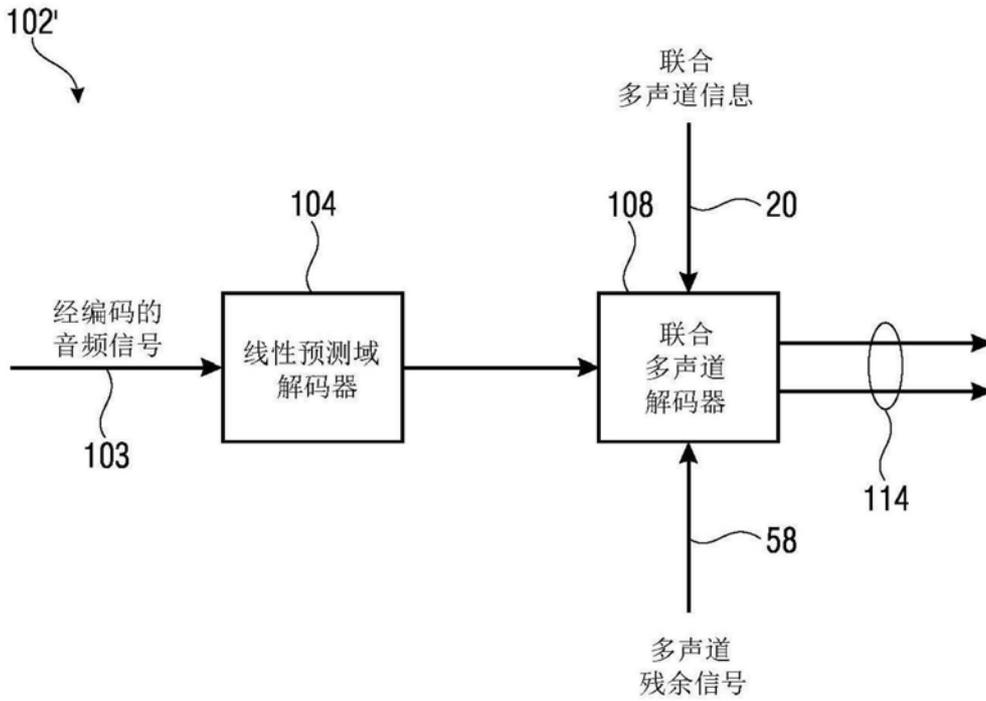


图11

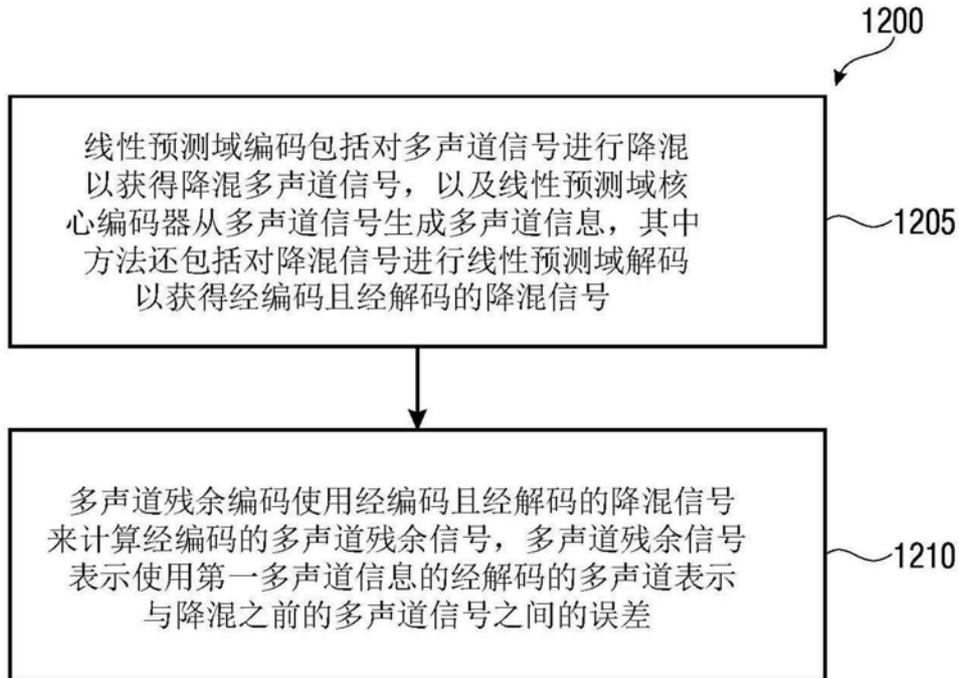


图12

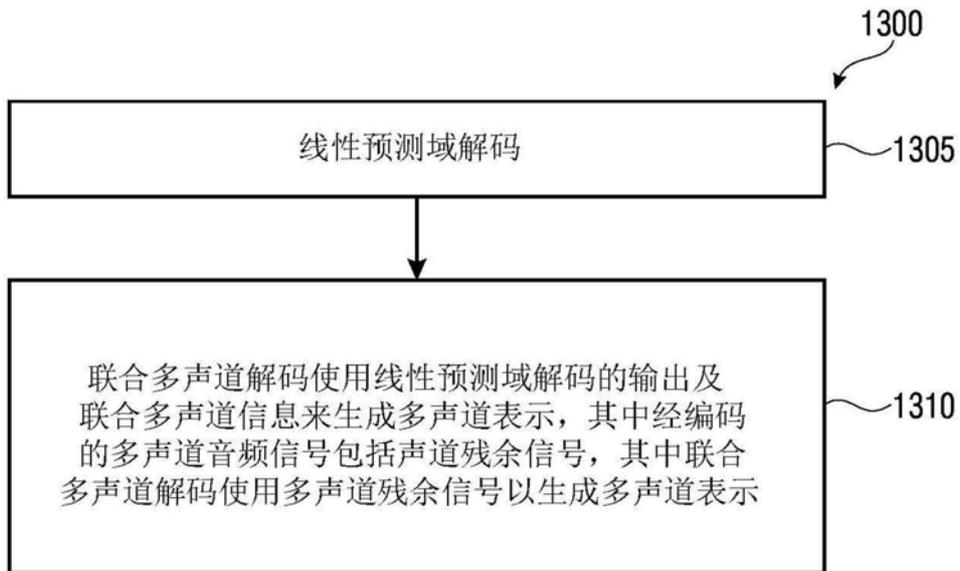


图13

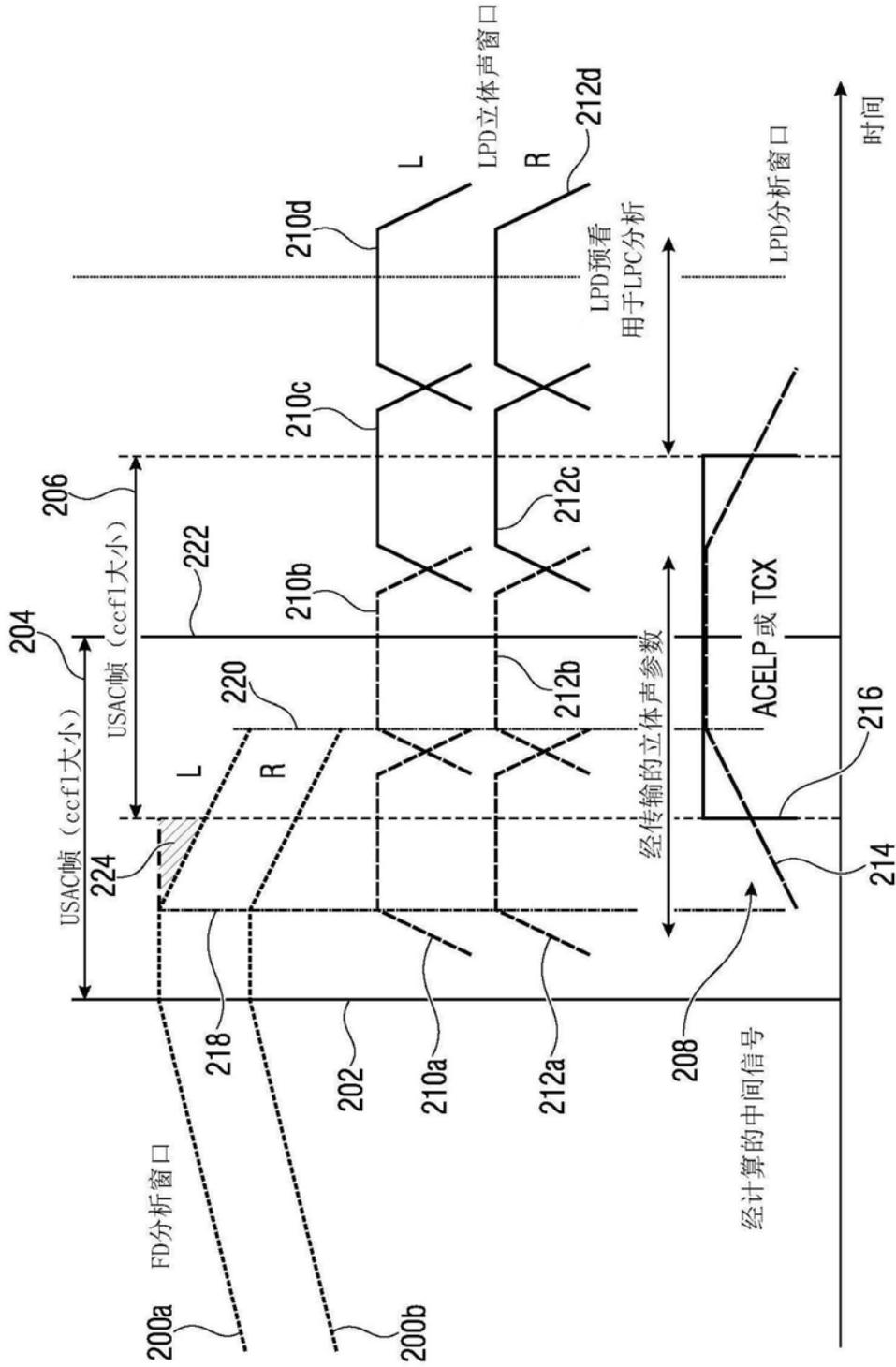


图14

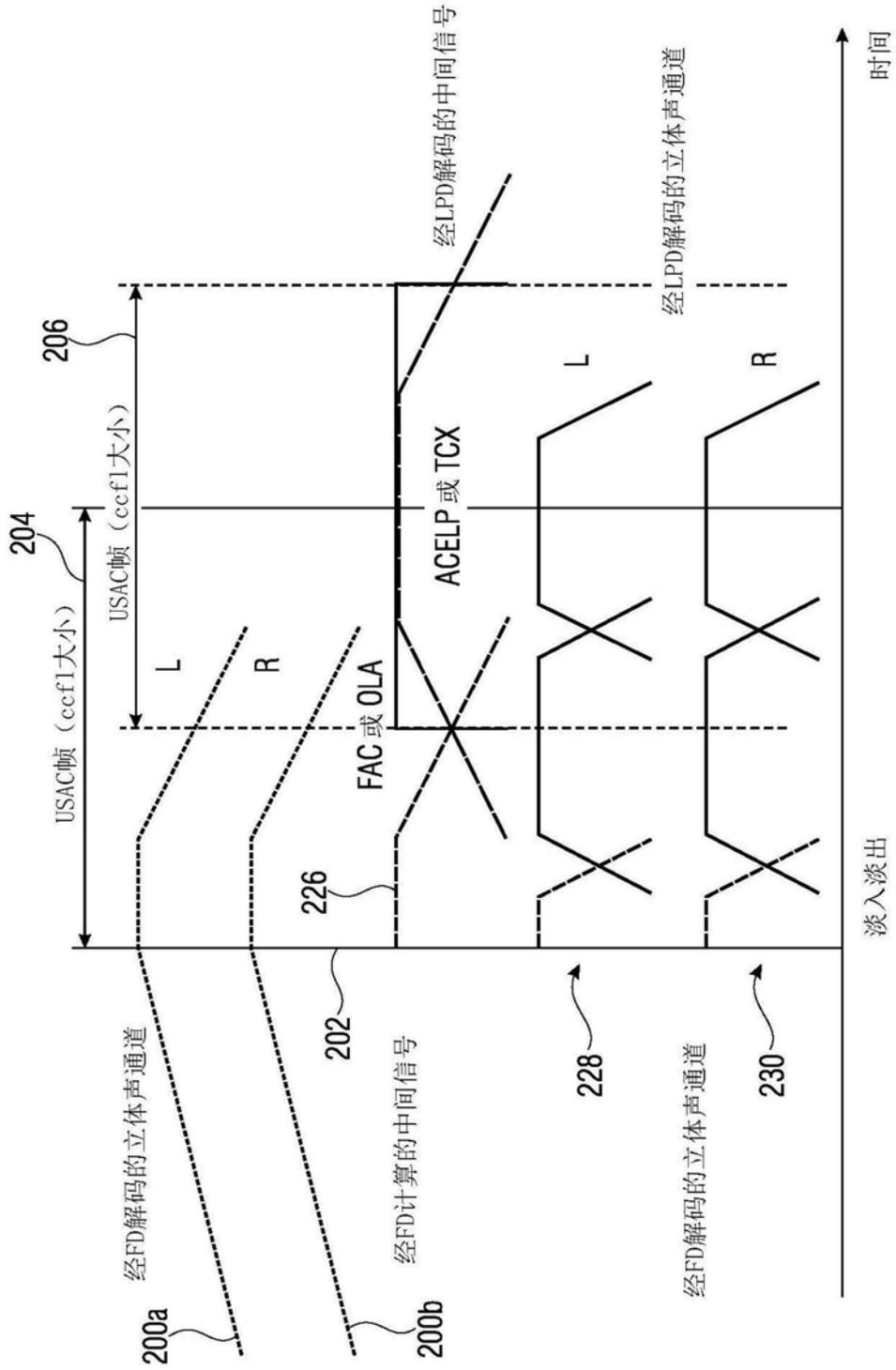


图15

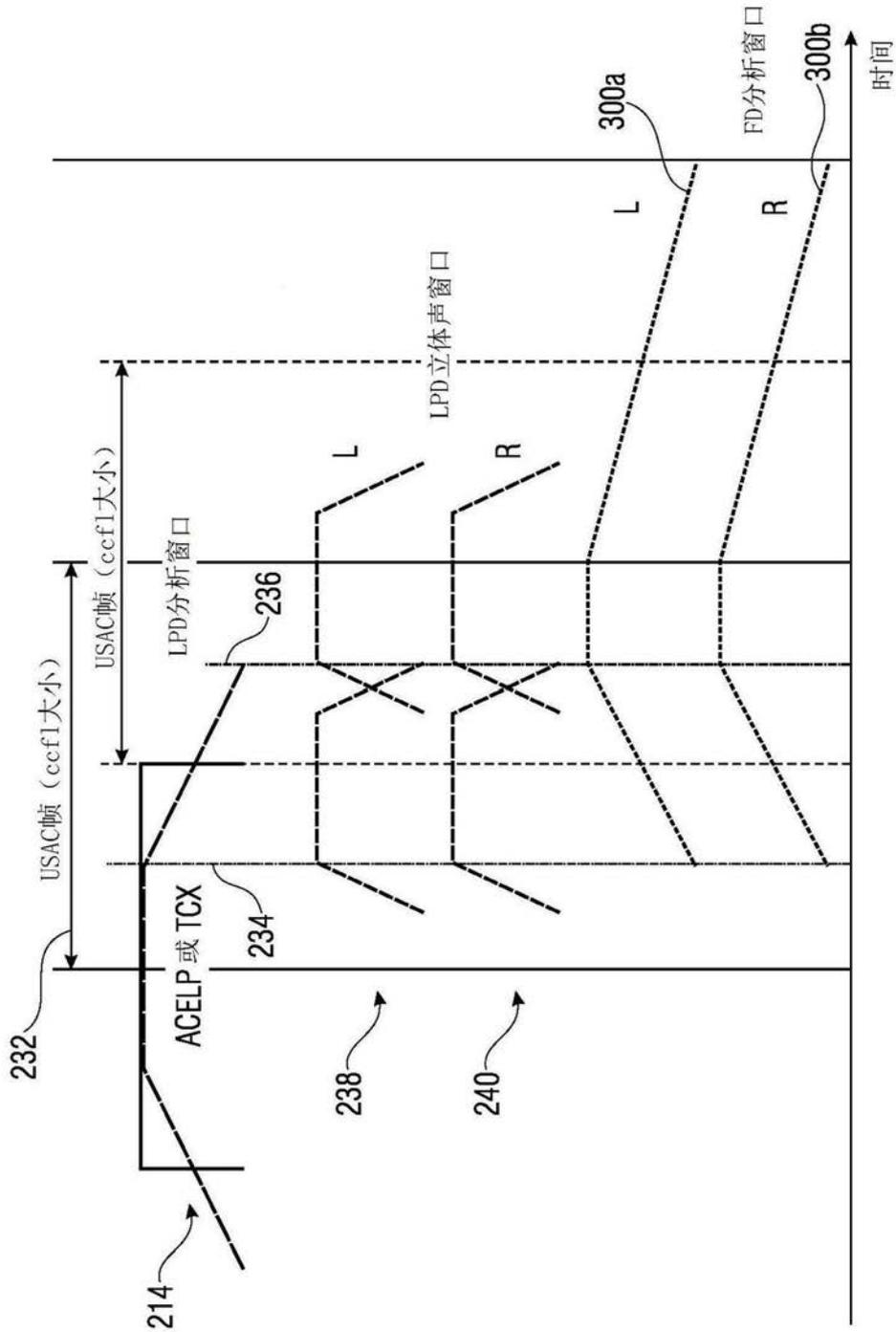


图16

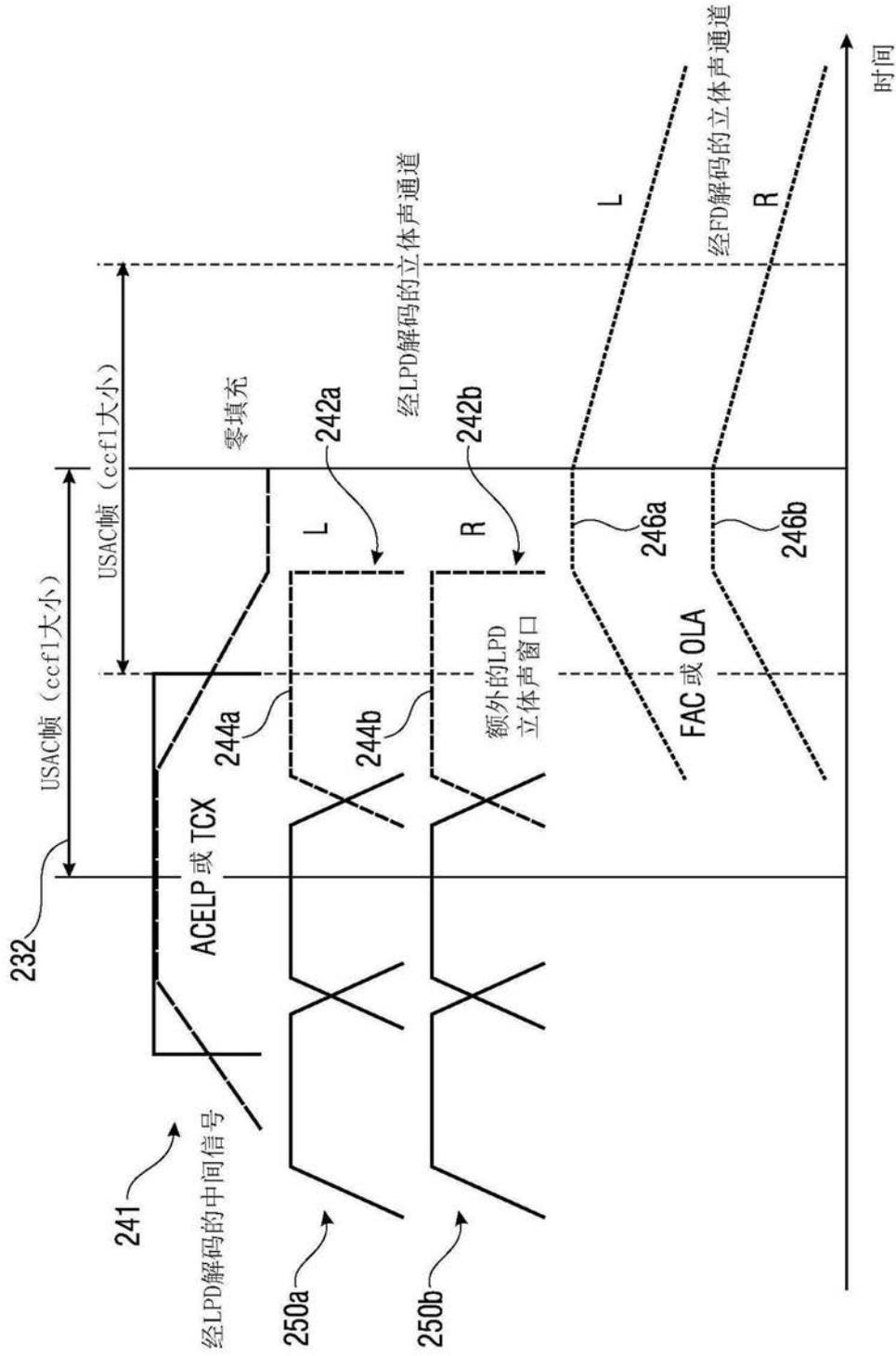


图17

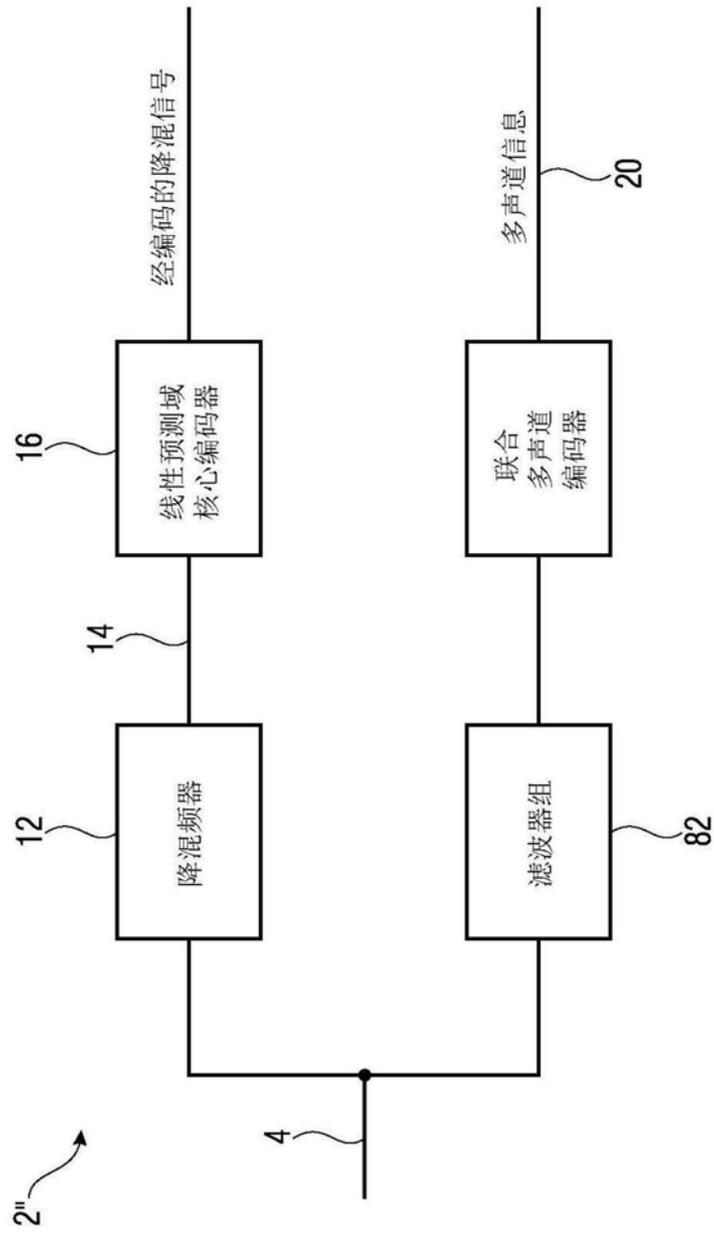


图18

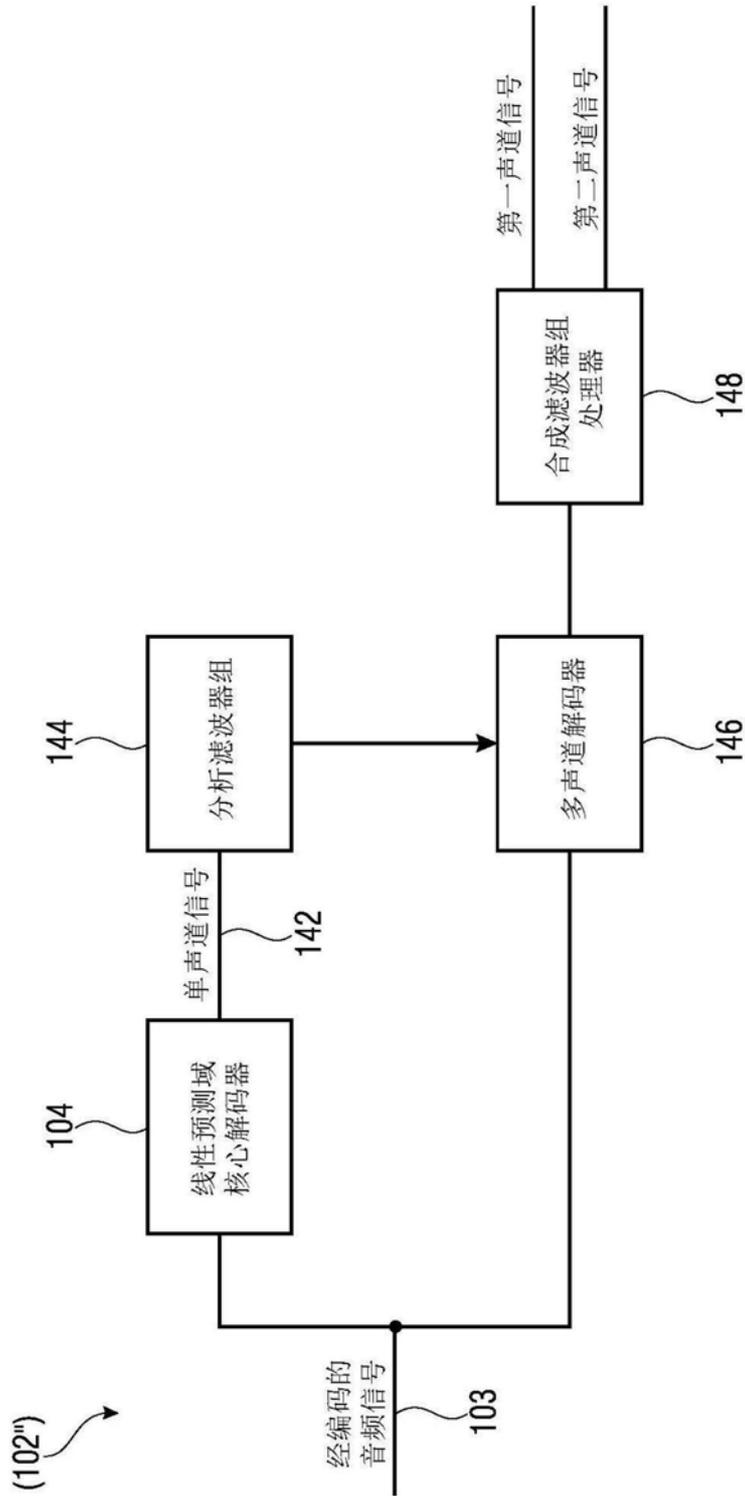


图19

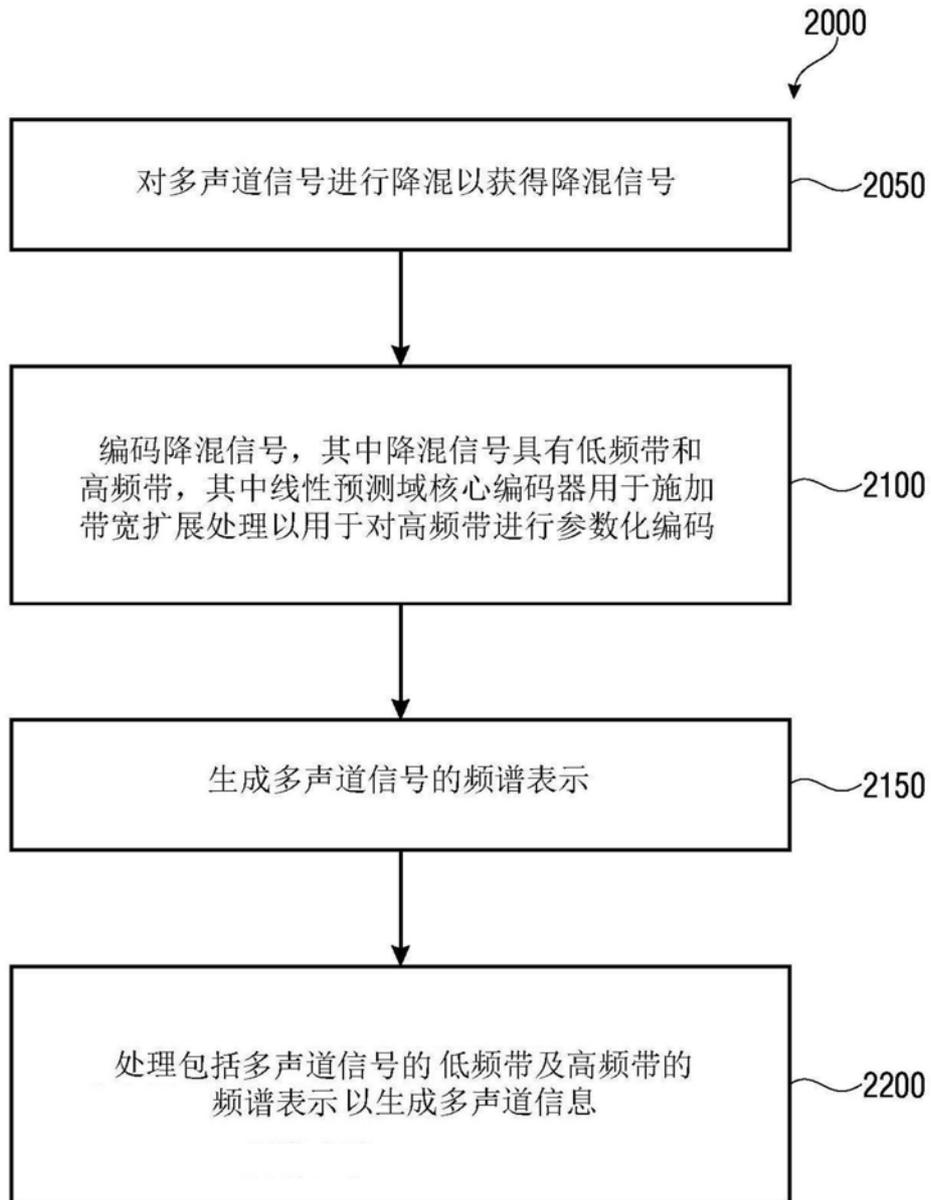


图20

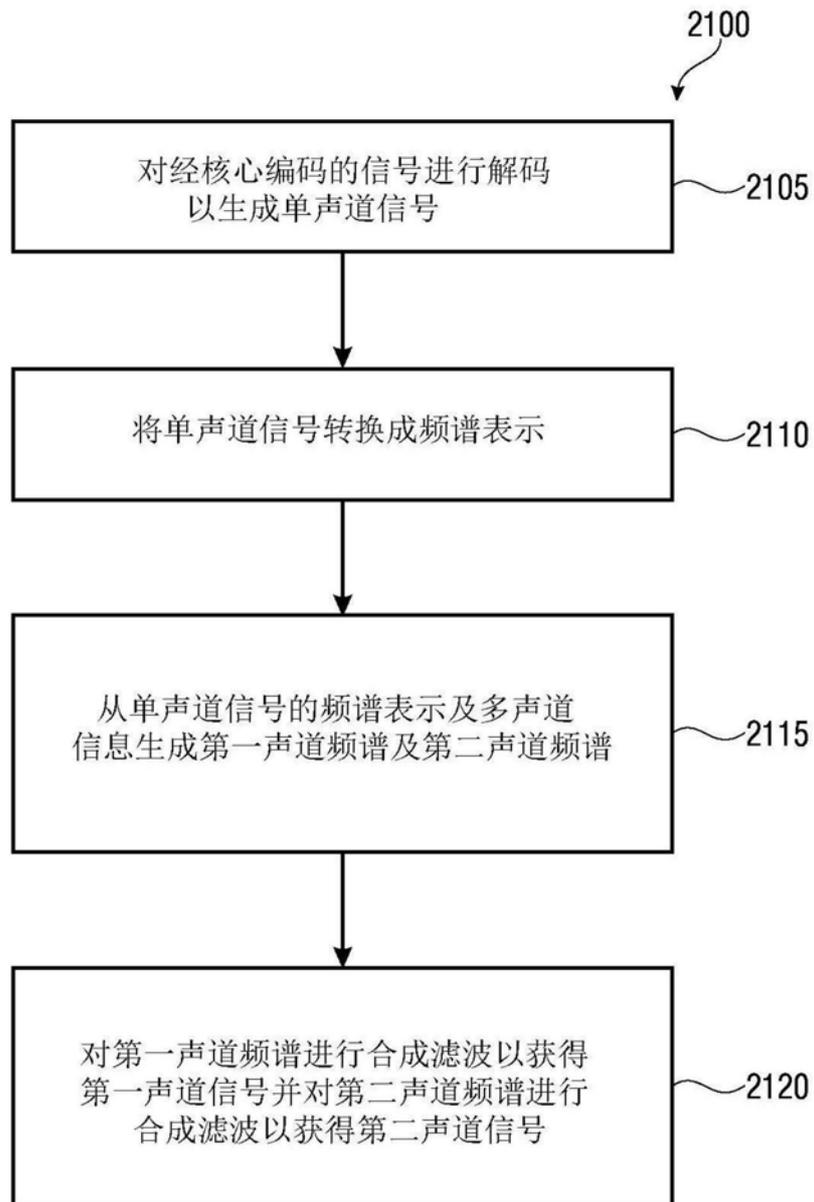


图21