



(12) 发明专利申请

(10) 申请公布号 CN 117542365 A

(43) 申请公布日 2024. 02. 09

(21) 申请号 202311493628.5

(22) 申请日 2017.01.20

(30) 优先权数据

16152457.4 2016.01.22 EP

16152454.1 2016.01.22 EP

16199895.0 2016.11.21 EP

(62) 分案原申请数据

201780012788.X 2017.01.20

(71) 申请人 弗劳恩霍夫应用研究促进协会

地址 德国慕尼黑

(72) 发明人 以马利·拉韦利 马库斯·施内尔

斯蒂芬·朵拉 乌尔夫冈·雅吉斯

马丁·迪茨 克里斯汀·赫姆瑞希

戈兰·马尔科维奇

埃伦尼·福托普楼

马库斯·马特拉斯 斯特凡·拜尔

纪尧姆·福克斯 于尔根·赫勒

(74) 专利代理机构 中科专利商标代理有限责任

公司 11021

专利代理师 罗松梅 潘剑颖

(51) Int.Cl.

G10L 19/008 (2013.01)

G10L 19/02 (2013.01)

G10L 19/03 (2013.01)

G10L 19/032 (2013.01)

G10L 19/22 (2013.01)

权利要求书9页 说明书39页 附图23页

(54) 发明名称

用于具有全局ILD和改进的中/侧决策的MDCT M/S立体声的装置和方法

(57) 摘要

提出了一种用于对包括两个或更多个声道的音频输入信号的第一声道和第二声道进行编码以获得编码音频信号的装置。该装置包括归一化器(110),归一化器(110)被配置为根据音频输入信号的第一声道并且根据音频输入信号的第二声道来确定音频输入信号的归一化值。此外,该装置包括编码单元(120),编码单元(120)被配置为产生具有第一声道和第二声道的处理后的音频信号,使得处理后的音频信号的第一声道的一个或多个频谱带是归一化音频信号的第一声道的一个或多个频谱带,使得处理后的音频信号的第二声道的一个或多个频谱带是归一化音频信号的第二声道的一个或多个频谱带。



1. 一种用于对包括两个或更多个声道的音频输入信号的第一声道和第二声道进行编码以获得编码音频信号的装置,其中,所述装置包括:

归一化器(110),所述归一化器(110)被配置为根据所述音频输入信号的第一声道并且根据所述音频输入信号的第二声道来确定所述音频输入信号的归一化值,其中所述归一化器(110)被配置为通过根据所述归一化值修正所述音频输入信号的第一声道和第二声道中的至少一个声道,来确定归一化音频信号的第一声道和第二声道;

编码单元(120),所述编码单元(120)被配置为产生具有第一声道和第二声道的处理后的音频信号,使得所述处理后的音频信号的第一声道的一个或多个频谱带是所述归一化音频信号的第一声道的一个或多个频谱带,使得所述处理后的音频信号的第二声道的一个或多个频谱带是所述归一化音频信号的第二声道的一个或多个频谱带,使得所述处理后的音频信号的第一声道的至少一个频谱带是根据所述归一化音频信号的第一声道的频谱带并且根据所述归一化音频信号的第二声道的频谱带的中央信号的频谱带,以及使得所述处理后的音频信号的第二声道的至少一个频谱带是根据所述归一化音频信号的第一声道的频谱带并且根据所述归一化音频信号的第二声道的频谱带的侧信号的频谱带,其中,所述编码单元(120)被配置为对所述处理后的音频信号进行编码以获得所述编码音频信号。

2. 根据权利要求1所述的装置,

其中,所述编码单元(120)被配置为根据所述归一化音频信号的第一声道的多个频谱带并且根据所述归一化音频信号的第二声道的多个频谱带,在全-中-侧编码模式、全-双-单声道编码模式和逐频带编码模式之间选择,

其中,所述编码单元(120)被配置为:如果选择所述全-中-侧编码模式,则根据所述归一化音频信号的第一声道并且根据所述归一化音频信号的第二声道产生中央信号作为中-侧信号的第一声道,根据所述归一化音频信号的第一声道和根据所述归一化音频信号的第二声道产生侧信号作为所述中-侧信号的第二声道,以及对所述中-侧信号进行编码以获得所述编码音频信号,

其中,所述编码单元(120)被配置为:如果选择所述全-双-单声道编码模式,则对所述归一化音频信号进行编码以获得所述编码音频信号,以及

其中,所述编码单元(120)被配置为:如果选择所述逐频带编码模式,则产生所述处理后的音频信号,使得所述处理后的音频信号的第一声道的一个或多个频谱带是所述归一化音频信号的第一声道的一个或多个频谱带,使得所述处理后的音频信号的第二声道的一个或多个频谱带是所述归一化音频信号的第二声道的一个或多个频谱带,使得所述处理后的音频信号的第一声道的至少一个频谱带是根据所述归一化音频信号的第一声道的频谱带并且根据所述归一化音频信号的第二声道的频谱带的中央信号的频谱带,以及使得所述处理后的音频信号的第二声道的至少一个频谱带是根据所述归一化音频信号的第一声道的频谱带并且根据所述归一化音频信号的第二声道的频谱带的侧信号的频谱带,其中所述编码单元(120)被配置为对所述处理后的音频信号进行编码以获得所述编码音频信号。

3. 根据权利要求2所述的装置,

其中,所述编码单元(120)被配置为:如果选择所述逐频带编码模式,则针对所述处理后的音频信号的多个频谱带中的每个频谱带,决定是采用中-侧编码还是采用双-单声道编码,

其中,如果针对所述频谱带采用所述中-侧编码,则所述编码单元(120)被配置为:基于所述归一化音频信号的第一声道的所述频谱带并且基于所述归一化音频信号的第二声道的所述频谱带,产生所述处理后的音频信号的第一声道的所述频谱带作为中央信号的频谱带,并且所述编码单元(120)被配置为:基于所述归一化音频信号的第一声道的所述频谱带并且基于所述归一化音频信号的第二声道的所述频谱带,产生所述处理后的音频信号的第二声道的所述频谱带作为侧信号的频谱带,以及

其中,如果针对所述频谱带采用所述双-单声道编码,则

所述编码单元(120)被配置为:使用所述归一化音频信号的第一声道的所述频谱带作为所述处理后的音频信号的第一声道的所述频谱带,并且被配置为使用所述归一化音频信号的第二声道的所述频谱带作为所述处理后的音频信号的第二声道的所述频谱带,或者

所述编码单元(120)被配置为:使用所述归一化音频信号的第二声道的所述频谱带作为所述处理后的音频信号的第一声道的所述频谱带,并且被配置为使用所述归一化音频信号的第一声道的所述频谱带作为所述处理后的音频信号的第二声道的所述频谱带。

4.根据权利要求2所述的装置,其中,所述编码单元(120)被配置为:通过确定估计在采用所述全-中-侧编码模式时编码所需的第一比特数的第一估计,通过确定估计在采用所述全-双-单声道编码模式时编码所需的第二比特数的第二估计,通过确定估计在采用所述逐频带编码模式时编码所需的第三比特数的第三估计,以及通过在所述全-中-侧编码模式、所述全-双-单声道编码模式和所述逐频带编码模式之中选择具有所述第一估计、所述第二估计和所述第三估计之中的最小比特数的编码模式,在所述全-中-侧编码模式、所述全-双-单声道编码模式和所述逐频带编码模式之间进行选择。

5.根据权利要求4所述的装置,

其中,所述编码单元(120)被配置为根据以下公式估计所述第三估计 b_{BW} ,所述第三估计估计在采用所述逐频带编码模式时编码所需的第三比特数:

$$b_{BW} = nBands + \sum_{i=0}^{nBands-1} \min(b_{bwLR}^i, b_{bwMS}^i)$$

其中, $nBands$ 是所述归一化音频信号的频谱带的个数,

其中, b_{bwMS}^i 是用于编码所述中央信号的第*i*个频谱带和用于编码所述侧信号的第*i*个频谱带所需的比特数的估计,以及

其中, b_{bwLR}^i 是用于编码第一信号的第*i*个频谱带和用于编码第二信号的第*i*个频谱带所需的比特数的估计。

6.根据权利要求2所述的装置,其中,所述编码单元(120)被配置为:通过确定估计在以所述全-中-侧编码模式进行编码时所保存的第一比特数的第一估计,通过确定估计在以所述全-双-单声道编码模式进行编码时所保存的第二比特数的第二估计,通过确定估计在以所述逐频带编码模式进行编码时所保存的第三比特数的第三估计,以及通过在所述全-中-侧编码模式、所述全-双-单声道编码模式和所述逐频带编码模式之中选择具有所述第一估计、所述第二估计和所述第三估计之中的所保存的最大比特数的编码模式,在所述全-中-侧编码模式、所述全-双-单声道编码模式和所述逐频带编码模式之间进行选择。

7. 根据权利要求2所述的装置,其中,所述编码单元(120)被配置为:通过估计在采用所述全-中-侧编码模式时发生的第一信噪比,通过估计在采用所述全-双-单声道编码模式时发生的第二信噪比,通过估计在采用所述逐频带编码模式时发生的第三信噪比,并且通过在所述全-中-侧编码模式、所述全-双-单声道编码模式和所述逐频带编码模式之中选择具有所述第一信噪比、所述第二信噪比和所述第三信噪比之中的最大信噪比的编码模式,在所述全-中-侧编码模式、所述全-双-单声道编码模式和所述逐频带编码模式之间进行选择。

8. 根据权利要求1所述的装置,

其中,所述编码单元(120)被配置为:产生所述处理后的音频信号,使得所述处理后的音频信号的第一声道的所述至少一个频谱带是所述中央信号的所述频谱带,并且使得所述处理后的音频信号的第二声道的所述至少一个频谱带是所述侧信号的所述频谱带,

其中,为了获得所述编码音频信号,所述编码单元(120)被配置为通过确定所述侧信号的所述频谱带的校正因子来编码所述侧信号的所述频谱带,

其中,所述编码单元(120)被配置为根据残差并且根据与所述中央信号的所述频谱带相对应的先前中央信号的频谱带,确定所述侧信号的所述频谱带的所述校正因子,其中所述先前中央信号在时间上在所述中央信号之前,

其中,所述编码单元(120)被配置为根据所述侧信号的所述频谱带、并且根据所述中央信号的所述频谱带来确定所述残差。

9. 根据权利要求8所述的装置,

其中,所述编码单元(120)被配置为根据以下公式确定所述侧信号的所述频谱带的所述校正因子:

$$\text{correction_factor}_{fb} = ERes_{fb} / (EprevDmx_{fb} + \epsilon)$$

其中, $\text{correction_factor}_{fb}$ 指示所述侧信号的所述频谱带的所述校正因子,

其中, $ERes_{fb}$ 指示根据与所述中央信号的所述频谱带相对应的所述残差的频谱带的能量的残差能量,

其中, $EprevDmx_{fb}$ 指示根据先前中央信号的频谱带的能量的先前能量,以及

其中, $\epsilon = 0$, 或其中 $0.1 > \epsilon > 0$ 。

10. 根据权利要求8所述的装置,

其中,根据以下公式定义所述残差:

$$Res_R = S_R - a_R Dmx_R,$$

其中, Res_R 是所述残差,其中 S_R 是所述侧信号,其中 a_R 是系数,其中 Dmx_R 是所述中央信号,

其中,所述编码单元(120)被配置为根据以下公式来确定所述残差能量:

$$ERes_{fb} = \sum_{fb} Res_R^2.$$

11. 根据权利要求8所述的装置,

其中,根据以下公式定义所述残差:

$$Res_R = S_R - a_R Dmx_R - a_I Dmx_I,$$

其中, Res_R 是所述残差,其中 S_R 是所述侧信号,其中 a_I 是复数系数的实部,并且其中 a_I 是所述复数系数的虚部,其中 Dmx_R 是所述中央信号,其中 Dmx_I 是根据所述归一化音频信号的第

一声道并且根据所述归一化音频信号的第二声道的另一中央信号，

其中，根据以下公式定义根据所述归一化音频信号的第一声道并且根据所述归一化音频信号的第二声道的另一侧信号 S_1 的另一残差：

$$Res_{s_1} = S_1 - a_R Dmx_R - a_L Dmx_L,$$

其中，所述编码单元 (120) 被配置为根据以下公式来确定所述残差能量：

$$ERes_{fb} = \sum_{fb} Res_R^2 + \sum_{fb} Res_L^2$$

其中，所述编码单元 (120) 被配置为根据与所述中央信号的所述频谱带相对应的所述残差的频谱带的能量、以及根据与所述中央信号的所述频谱带相对应的所述另一残差的频谱带的能量，确定先前能量。

12. 根据权利要求1所述的装置，

其中，所述归一化器 (110) 被配置为根据所述音频输入信号的第一声道的能量并且根据所述音频输入信号的第二声道的能量来确定所述音频输入信号的归一化值。

13. 根据权利要求1所述的装置，

其中，所述音频输入信号在频谱域中表示，

其中，所述归一化器 (110) 被配置为根据所述音频输入信号的第一声道的多个频谱带并且根据所述音频输入信号的第二声道的多个频谱带来确定所述音频输入信号的归一化值，以及

其中，所述归一化器 (110) 被配置为通过根据所述归一化值修正所述音频输入信号的第一声道和第二声道中的至少一个声道的多个频谱带来确定所述归一化音频信号。

14. 根据权利要求13所述的装置，

其中，所述归一化器 (110) 被配置为基于以下公式确定所述归一化值：

$$NRG_L = \sqrt{\sum MDCT_{L,k}^2}$$

$$NRG_R = \sqrt{\sum MDCT_{R,k}^2}$$

$$ILD = \frac{NRG_L}{NRG_L + NRG_R}$$

其中， $MDCT_{L,k}$ 是所述音频输入信号的第一声道的MDCT频谱的第k个系数，并且 $MDCT_{R,k}$ 是所述音频输入信号的第二声道的MDCT频谱的第k个系数，以及

其中，所述归一化器 (110) 被配置为通过量化ILD来确定所述归一化值。

15. 根据权利要求13所述的装置，

其中，所述用于编码的装置还包括变换单元 (102) 和预处理单元 (105)，

其中，所述变换单元 (102) 被配置为将时域音频信号从时域变换到频域以获得变换后的音频信号，

其中，所述预处理单元 (105) 被配置为通过对所述变换后的音频信号应用编码器侧频

域噪声整形操作来产生所述音频输入信号的第一声道和第二声道。

16. 根据权利要求15所述的装置,

其中,所述预处理单元(105)被配置为通过在对所述变换后的音频信号应用编码器侧频域噪声整形操作之前对所述变换后的音频信号应用编码器侧时间噪声整形操作,来产生所述音频输入信号的第一声道和第二声道。

17. 根据权利要求1所述的装置,

其中,所述归一化器(110)被配置为根据在时域中表示的所述音频输入信号的第一声道并且根据在时域中表示的所述音频输入信号的第二声道来确定所述音频输入信号的归一化值,

其中,所述归一化器(110)被配置为通过根据所述归一化值修正在时域中表示的所述音频输入信号的第一声道和第二声道中的至少一个声道来确定所述归一化音频信号的第一声道和第二声道,

其中,所述装置还包括变换单元(115),所述变换单元(115)被配置为将所述归一化音频信号从时域变换到频谱域,使得所述归一化音频信号在频谱域中表示,以及

其中,所述变换单元被配置为将在频谱域中表示的所述归一化音频信号馈送到所述编码单元(120)中。

18. 根据权利要求17所述的装置,

其中,所述装置还包括被配置为接收包括第一声道和第二声道的时域音频信号的预处理单元(106),

其中,所述预处理单元(106)被配置为对时域音频信号中的、产生第一感知白化频谱的第一声道应用滤波器,以获得在时域中表示的音频输入信号的第一声道,以及

其中,所述预处理单元(106)被配置为对时域音频信号中的、产生第二感知白化频谱的第二声道应用所述滤波器,以获得在时域中表示的音频输入信号的第二声道。

19. 根据权利要求17所述的装置,

其中,所述变换单元(115)被配置为将归一化音频信号从时域变换到频谱域以获得变换后的音频信号,

其中,所述装置还包括频谱域预处理器(118),所述频谱域预处理器(118)被配置为对所述变换后的音频信号执行编码器侧时间噪声整形,以获得在频谱域中表示的归一化音频信号。

20. 根据权利要求1所述的装置,

其中,所述编码单元(120)被配置为通过对所述归一化音频信号或所述处理后的音频信号应用编码器侧立体声智能间隙填充来获得所述编码音频信号。

21. 根据权利要求1所述的装置,其中,所述音频输入信号是恰好包括两个声道的音频立体声信号。

22. 一种用于对包括四个或更多个声道的音频输入信号的四个声道进行编码以获得编码音频信号的系统,其中,所述系统包括:

根据权利要求1所述的第一装置(170),用于对所述音频输入信号的四个或更多个声道中的第一声道和第二声道进行编码,以获得所述编码音频信号的第一声道和第二声道,以及

根据权利要求1所述的第二装置(180),用于对所述音频输入信号的四个或更多个声道中的第三声道和第四声道进行编码,以获得所述编码音频信号的第三声道和第四声道。

23.一种用于对包括第一声道和第二声道的编码音频信号进行解码以获得包括两个或更多个声道的解码音频信号的第一声道和第二声道的装置,

其中,所述装置包括解码单元(210),所述解码单元(210)被配置为针对多个频谱带中的每个频谱带,确定所述编码音频信号的第一声道的所述频谱带和所述编码音频信号的第二声道的所述频谱带是使用双-单声道编码来编码的还是使用中-侧编码来编码的,

其中,如果使用了所述双-单声道编码,则所述解码单元(210)被配置为使用所述编码音频信号的第一声道的所述频谱带作为中间音频信号的第一声道的频谱带,并且被配置为使用所述编码音频信号的第二声道的所述频谱带作为所述中间音频信号的第二声道的频谱带,

其中,如果使用了所述中-侧编码,则所述解码单元(210)被配置为基于所述编码音频信号的第一声道的所述频谱带并且基于所述编码音频信号的第二声道的所述频谱带来产生所述中间音频信号的第一声道的频谱带,以及基于所述编码音频信号的第一声道的所述频谱带并且基于所述编码音频信号的第二声道的所述频谱带,来产生所述中间音频信号的第二声道的频谱带,以及

其中,所述装置包括去归一化器(220),所述去归一化器(220)被配置为根据去归一化值来修正所述中间音频信号的第一声道和第二声道中的至少一个声道,以获得所述解码音频信号的第一声道和第二声道。

24.根据权利要求23所述的装置,

其中,所述解码单元(210)被配置为确定所述编码音频信号是以全-中-侧编码模式、以全-双-单声道编码模式、还是以逐频带编码模式来编码的,

其中,所述解码单元(210)被配置为:如果确定所述编码音频信号是以所述全-中-侧编码模式编码的,则根据所述编码音频信号的第一声道并且根据所述编码音频信号的第二声道来产生所述中间音频信号的第一声道,以及根据所述编码音频信号的第一声道并且根据所述编码音频信号的第二声道来产生所述中间音频信号的第二声道,

其中,所述解码单元(210)被配置为:如果确定所述编码音频信号是以所述全-双-单声道编码模式编码的,则使用所述编码音频信号的第一声道作为所述中间音频信号的第一声道,以及使用所述编码音频信号的第二声道作为所述中间音频信号的第二声道,以及

其中,所述解码单元(210)被配置为:如果确定所述编码音频信号是以所述逐频带编码模式编码的,则

针对多个频谱带中的每个频谱带,确定所述编码音频信号的第一声道的所述频谱带和所述编码音频信号的第二声道的所述频谱带是使用所述双-单声道编码来编码的还是使用所述中-侧编码来编码的,

如果使用了所述双-单声道编码,则使用所述编码音频信号的第一声道的所述频谱带作为所述中间音频信号的第一声道的频谱带,并且使用所述编码音频信号的第二声道的所述频谱带作为所述中间音频信号的第二声道的频谱带,以及

如果使用了所述中-侧编码,则基于所述编码音频信号的第一声道的所述频谱带并且基于所述编码音频信号的第二声道的所述频谱带,产生所述中间音频信号的第一声道的频

谱带,以及基于所述编码音频信号的第一声道的所述频谱带并且基于所述编码音频信号的第二声道的所述频谱带,产生所述中间音频信号的第二声道的频谱带。

25. 根据权利要求23所述的装置,

其中,所述解码单元(210)被配置为针对所述多个频谱带中的每个频谱带来确定所述编码音频信号的第一声道的所述频谱带和所述编码音频信号的第二声道的所述频谱带是使用双-单声道编码来编码的还是使用中-侧编码来编码的,

其中,所述解码单元(210)被配置为通过重新构建所述第二声道的所述频谱带来获得所述编码音频信号的第二声道的所述频谱带,

其中,如果使用中-侧编码,则所述编码音频信号的第一声道的所述频谱带是中央信号的频谱带,并且所述编码音频信号的第二声道的所述频谱带是侧信号的频谱带,

其中,如果使用中-侧编码,则所述解码单元(210)被配置为根据所述侧信号的所述频谱带的校正因子并且根据与所述中央信号的所述频谱带相对应的先前中央信号的频谱带,重新构建所述侧信号的所述频谱带,其中所述先前中央信号在时间上在所述中央信号之前。

26. 根据权利要求25所述的装置,

其中,如果使用了中-侧编码,则所述解码单元(210)被配置为通过根据以下公式重新构建所述侧信号的所述频谱带的频谱值来重新构建所述侧信号的所述频谱带,

$$S_i = N_i + \text{facDmx}_{fb} \cdot \text{prevDmx}_i$$

其中, S_i 指示所述侧信号的所述频谱带的频谱值,

其中, prevDmx_i 指示所述先前中央信号的频谱带的频谱值,

其中, N_i 指示噪声填充频谱的频谱值,

其中,根据如下公式来定义 facDmx_{fb} :

$$\text{facDmx}_{fb} = \sqrt{\text{correction_factor}_{fb} - \frac{EN_{fb}}{(E\text{prevDmx}_{fb} + \varepsilon)}}$$

其中, $\text{correction_factor}_{fb}$ 是所述侧信号的所述频谱带的所述校正因子,

其中, EN_{fb} 是所述噪声填充频谱的能量,

其中, $E\text{prevDmx}_{fb}$ 是所述先前中央信号的所述频谱带的能量,以及

其中, $\varepsilon = 0$,或其中 $0.1 > \varepsilon > 0$ 。

27. 根据权利要求23所述的装置,

其中,所述去归一化器(220)被配置为根据所述去归一化值来修正所述中间音频信号的第一声道和第二声道中的至少一个声道的多个频谱带,以获得所述解码音频信号的第一声道和第二声道。

28. 根据权利要求23所述的装置,

其中,所述去归一化器(220)被配置为根据所述去归一化值来修正所述中间音频信号的第一声道和第二声道中的至少一个声道的多个频谱带,以获得去归一化音频信号,

其中,所述装置还包括后处理单元(230)和变换单元(235),以及

其中,所述后处理单元(230)被配置为对所述去归一化音频信号执行解码器侧时间噪声整形和解码器侧频域噪声整形中的至少一个,以获得后处理音频信号,

其中,所述变换单元(235)被配置为将所述后处理音频信号从频谱域变换到时域,以获得所述解码音频信号的第一声道和第二声道。

29.根据权利要求23所述的装置,

其中,所述装置还包括被配置为将所述中间音频信号从频谱域变换到时域的变换单元(215),

其中,所述去归一化器(220)被配置为根据所述去归一化值来修正在时域中表示的中间音频信号的第一声道和第二声道中的至少一个声道,以获得所述解码音频信号的第一声道和第二声道。

30.根据权利要求23所述的装置,

其中,所述装置还包括被配置为将所述中间音频信号从频谱域变换到时域的变换单元(215),

其中,所述去归一化器(220)被配置为根据所述去归一化值来修正在时域中表示的中间音频信号的第一声道和第二声道中的至少一个声道,以获得去归一化音频信号,

其中,所述装置还包括后处理单元(235),所述后处理单元(235)被配置为处理作为感知白化音频信号的所述去归一化音频信号,以获得所述解码音频信号的第一声道和第二声道。

31.根据权利要求29所述的装置,

其中,所述装置还包括被配置为对所述中间音频信号执行解码器侧时间噪声整形的频谱域后处理器(212),

其中,所述变换单元(215)被配置为在已经对所述中间音频信号执行了解码器侧时间噪声整形之后,将所述中间音频信号从频谱域变换到时域。

32.根据权利要求23所述的装置,

其中,所述解码单元(210)被配置为对所述编码音频信号应用解码器侧立体声智能间隙填充。

33.根据权利要求23所述的装置,其中,所述解码音频信号是恰好包括两个声道的音频立体声信号。

34.一种用于对包括四个或更多个声道的编码音频信号进行解码以获得包括四个或更多个声道的解码音频信号的四个声道的系统,其中,所述系统包括:

根据权利要求23所述的第一装置(270),用于对所述编码音频信号的四个或更多个声道中的第一声道和第二声道进行解码,以获得所述解码音频信号的第一声道和第二声道,以及

根据权利要求23所述的第二装置(280),用于对所述编码音频信号的四个或更多个声道中的第三声道和第四声道进行解码,以获得所述解码音频信号的第三声道和第四声道。

35.一种用于根据音频输入信号来产生编码音频信号以及根据编码音频信号来产生解码音频信号的系统,包括:

根据权利要求1所述的装置(310),其中根据权利要求1所述的装置(310)被配置为根据所述音频输入信号产生所述编码音频信号,以及

根据权利要求23所述的装置(320),其中根据权利要求23所述的装置(320)被配置为根据所述编码音频信号产生所述解码音频信号。

36.一种用于根据音频输入信号来产生编码音频信号以及根据编码音频信号来产生解码音频信号的系统,包括:

根据权利要求22所述的系统,其中,根据权利要求22所述的系统被配置为根据所述音频输入信号产生所述编码音频信号,以及

根据权利要求34所述的系统,其中,根据权利要求34所述的系统被配置为根据所述编码音频信号产生所述解码音频信号。

37.一种用于对包括两个或更多个声道的音频输入信号的第一声道和第二声道进行编码以获得编码音频信号的方法,其中,所述方法包括:

根据所述音频输入信号的第一声道并且根据所述音频输入信号的第二声道来确定所述音频输入信号的归一化值,

通过根据所述归一化值修正所述音频输入信号的第一声道和第二声道中的至少一个声道来确定归一化音频信号的第一声道和第二声道,

产生具有第一声道和第二声道的处理后的音频信号,使得所述处理后的音频信号的第一声道的一个或多个频谱带是所述归一化音频信号的第一声道的一个或多个频谱带,使得所述处理后的音频信号的第二声道的一个或多个频谱带是所述归一化音频信号的第二声道的一个或多个频谱带,使得所述处理后的音频信号的第一声道的至少一个频谱带是根据所述归一化音频信号的第一声道的频谱带并且根据所述归一化音频信号的第二声道的频谱带的中央信号的频谱带,以及使得所述处理后的音频信号的第二声道的至少一个频谱带是根据所述归一化音频信号的第一声道的频谱带并且根据所述归一化音频信号的第二声道的频谱带的侧信号的频谱带,以及编码所述处理后的音频信号以获得所述编码音频信号。

38.一种用于对包括第一声道和第二声道的编码音频信号进行解码以获得包括两个或更多个声道的解码音频信号的第一声道和第二声道的方法,其中所述方法包括:

针对多个频谱带中的每个频谱带,确定所述编码音频信号的第一声道的所述频谱带和所述编码音频信号的第二声道的所述频谱带是使用双-单声道编码来编码的还是使用中-侧编码来编码的,

如果使用了所述双-单声道编码,则使用所述编码音频信号的第一声道的所述频谱带作为中间音频信号的第一声道的频谱带,并且使用所述编码音频信号的第二声道的所述频谱带作为所述中间音频信号的第二声道的频谱带,

如果使用了所述中-侧编码,则基于所述编码音频信号的第一声道的所述频谱带并且基于所述编码音频信号的第二声道的所述频谱带来产生所述中间音频信号的第一声道的频谱带,以及基于所述编码音频信号的第一声道的所述频谱带并且基于所述编码音频信号的第二声道的所述频谱带来产生所述中间音频信号的第二声道的频谱带,以及

根据去归一化值,修正所述中间音频信号的第一声道和第二声道中的至少一个声道,以获得解码音频信号的第一声道和第二声道。

39.一种计算机可读存储介质,存储有计算机程序,所述计算机程序用于在计算机或信号处理器上执行时实施根据权利要求37或38所述的方法。

用于具有全局ILD和改进的中/侧决策的MDCT M/S立体声的装置和方法

[0001] 本申请是申请日为2017年01月20日的PCT国际申请PCT/EP2017/051177在2018年08月22日向中国专利局递交并进入中国国家阶段后对应的题为“用于具有全局ILD和改进的中/侧决策的MDCT M/S立体声的装置和方法”的发明专利申请No.201780012788.X的分案申请。

技术领域

[0002] 本发明涉及音频信号编码和音频信号解码,并且更具体地涉及用于具有全局ILD和改进的中/侧决策的MDCT M/S立体声的装置和方法。

背景技术

[0003] 基于MDCT (MDCT=修正的离散余弦变换)的编码器中的逐频带 (Band-wise) M/S (M/S=中/侧) 处理是用于立体声处理的已知且有效的方法。然而,对于平移 (panned) 信号这种方法不足够,还需要附加处理 (例如,复数预测、或中央声道和侧声道之间的角度编码)。

[0004] 在[1]、[2]、[3]和[4]中,描述了对加窗和变换的非归一化 (非白化) 信号的M/S处理。

[0005] 在[7]中,描述了中央声道和侧声道之间的预测。在[7]中,公开了一种编码器,其基于两个音频声道的组合对音频信号进行编码。该音频编码器获得作为中央信号的组合信号,并且还获得预测残差信号,该预测残差信号是从中央信号导出的预测侧信号。第一组合信号和预测残差信号被编码并与预测信息一起写入数据流。此外,[7]公开了一种解码器,其使用预测残差信号、第一组合信号和预测信息来产生解码的第一音频声道和第二音频声道。

[0006] 在[5]中,描述了在分别对每个频带进行归一化后的M/S立体声耦合的应用。特别地,[5]指代Opus编解码器。Opus将中央信号和侧信号编码为归一化信号 $m = M / ||M||$ 和 $s = S / ||S||$ 。为了从 m 和 s 恢复 M 和 S ,对角度 $\theta_s = \arctan (||S|| / ||M||)$ 进行编码。当 N 是频带的大小并且 a 是 m 和 s 可用的总比特数时, m 的最优分配是 $a_{mid} = (a - (N-1) \log_2 \tan \theta_s) / 2$ 。

[0007] 在已知的方法中 (例如在[2]和[4]中),复杂的速率/失真回路与其中将 (例如,使用M/S,也可以跟随来自[7]的M到S预测残差计算) 变换频带声道的决策相组合,以减少声道之间的相关性。这种复杂的结构具有高计算成本。将感知模型与速率回路分离 (如[6a]、[6b]和[13]中那样) 显著简化了系统。

[0008] 此外,对每个频带中的预测系数或角度进行编码需要大量的比特 (例如,如在[5]和[7]中的那样)。

[0009] 在[1]、[3]和[5]中,仅对整个频谱执行单一决策,以决定整个频谱是应该被M/S编码还是被L/R编码。

[0010] 如果存在ILD (耳间水平差),即如果声道被平移,则M/S编码效率不高。

[0011] 如上所述,已知基于MDCT的编码器中的逐频带M/S处理是用于立体声处理的有效

方法。M/S处理编码增益从针对不相关声道的0%变化到针对单声道或针对声道之间的 $\pi/2$ 相位差的50%。由于立体声解屏蔽和逆解屏蔽(参见[1]),因此有鲁棒的M/S决策是很重要的。

[0012] 在[2]中,在每个频带中,左右之间的掩蔽阈值变化小于2dB,选择M/S编码作为编码方法。

[0013] 在[1]中,M/S决策基于针对M/S编码的和针对声道的L/R(L/R=左/右)编码的估计比特消耗。使用感知熵(PE)根据频谱和根据掩蔽阈值来估计针对M/S编码和针对L/R编码的比特率需求。针对左和右声道计算掩蔽阈值。假设针对中央声道的掩蔽阈值和针对侧声道的掩蔽阈值是左阈值和右阈值的最小值。

[0014] 此外,[1]描述了如何导出要被编码的各个声道的编码阈值。具体地,左声道和右声道的编码阈值是通过针对这些声道的相应感知模型来计算的。在[1]中,M声道和S声道的编码阈值被相等地选择,并且被导出为左编码阈值和右编码阈值的最小值。

[0015] 此外,[1]描述了在L/R编码和M/S编码之间做决定,从而实现了良好的编码性能。具体地,使用阈值来估计针对L/R编码和针对M/S编码的感知熵。

[0016] 在[1]和[2]以及[3]和[4]中,对加窗和变换的非归一化(非白化)信号进行M/S处理,M/S决策基于掩蔽阈值和感知熵估计。

[0017] 在[5]中,左声道和右声道的能量被明确地编码,并且编码的角度保留差信号的能量。在[5]中,假设即使L/R编码更有效,M/S编码也是安全的。根据[5],仅当声道之间的相关性不够强时才选择L/R编码。

[0018] 此外,对每个频带中的预测系数或角度进行编码需要大量的比特(例如,参见[5]和[7])。

[0019] 因此,如果将提供针对音频编码和音频解码的改进构思,将会高度赞赏。

发明内容

[0020] 本发明的目的是提供用于音频信号编码、音频信号处理和音频信号解码的改进构思。通过根据权利要求1所述的音频解码器、通过根据权利要求23所述的装置、通过根据权利要求37所述的方法、通过根据权利要求38所述的方法以及通过根据权利要求39所述的计算机程序来实现本发明的目的。

[0021] 根据实施例,提供了用于对包括两个或更多个声道的音频输入信号的第一声道和第二声道进行编码以获得编码音频信号的装置。

[0022] 该用于编码的装置包括归一化器,归一化器被配置为根据音频输入信号的第一声道并且根据音频输入信号的第二声道来确定音频输入信号的归一化值,其中归一化器被配置为通过根据归一化值修正音频输入信号的第一声道和第二声道中的至少一个声道,来确定归一化音频信号的第一声道和第二声道。

[0023] 此外,该用于编码的装置包括编码单元,编码单元被配置为产生具有第一声道和第二声道的处理后的音频信号,使得处理后的音频信号的第一声道的一个或多个频谱带是归一化音频信号的第一声道的一个或多个频谱带,使得处理后的音频信号的第二声道的一个或多个频谱带是归一化音频信号的第二声道的一个或多个频谱带,使得处理后的音频信号的第一声道的至少一个频谱带是根据归一化音频信号的第一声道的频谱带并且根据归

一化音频信号的第二声道的频谱带的中央信号的频谱带,以及使得处理后的音频信号的第二声道的至少一个频谱带是根据归一化音频信号的第一声道的频谱带并且根据归一化音频信号的第二声道的频谱带的侧信号的频谱带。编码单元被配置为对处理后的音频信号进行编码以获得编码音频信号。

[0024] 此外,提供了一种用于对包括第一声道和第二声道的编码音频信号进行解码以获得包括两个或更多个声道的解码音频信号的第一声道和第二声道的装置。

[0025] 该用于解码的装置包括解码单元,解码单元被配置为针对多个频谱带中的每个频谱带,来确定编码音频信号的第一声道的所述频谱带和编码音频信号的第二声道的所述频谱带是使用双-单声道编码来编码的还是使用中-侧编码来编码的。

[0026] 如果使用了双-单声道编码,则解码单元被配置为使用编码音频信号的第一声道的所述频谱带作为中间音频信号的第一声道的频谱带,并且被配置为使用编码音频信号的第二声道的所述频谱带作为中间音频信号的第二声道的频谱带。

[0027] 此外,如果使用了中-侧编码,则解码单元被配置为基于编码音频信号的第一声道的所述频谱带并且基于编码音频信号的第二声道的所述频谱带来产生中间音频信号的第一声道的频谱带,以及基于编码音频信号的第一声道的所述频谱带并且基于编码音频信号的第二声道的所述频谱带,来产生中间音频信号的第二声道的频谱带。

[0028] 此外,该用于解码的装置包括去归一化器,去归一化器被配置为根据去归一化值来修正中间音频信号的第一声道和第二声道中的至少一个声道,以获得解码音频信号的第一声道和第二声道。

[0029] 此外,提供了用于对包括两个或更多个声道的音频输入信号的第一声道和第二声道进行编码以获得编码音频信号的方法。所述方法包括:

[0030] -根据音频输入信号的第一声道并且根据音频输入信号的第二声道来确定音频输入信号的归一化值。

[0031] -通过根据归一化值修正音频输入信号的第一声道和第二声道中的至少一个声道来确定归一化音频信号的第一声道和第二声道。

[0032] -产生具有第一声道和第二声道的处理后的音频信号,使得处理后的音频信号的第一声道的一个或多个频谱带是归一化音频信号的第一声道的一个或多个频谱带,使得处理后的音频信号的第二声道的一个或多个频谱带是归一化音频信号的第二声道的一个或多个频谱带,使得处理后的音频信号的第一声道的至少一个频谱带是根据归一化音频信号的第一声道的频谱带并且根据归一化音频信号的第二声道的频谱带的中央信号的频谱带,以及使得处理后的音频信号的第二声道的至少一个频谱带是根据归一化音频信号的第一声道的频谱带并且根据归一化音频信号的第二声道的频谱带的侧信号的频谱带,以及编码处理后的音频信号以获得编码音频信号。

[0033] 此外,提供了一种用于对包括第一声道和第二声道的编码音频信号进行解码以获得包括两个或更多个声道的解码音频信号的第一声道和第二声道的方法。所述方法包括:

[0034] -针对多个频谱带中的每个频谱带,确定编码音频信号的第一声道的所述频谱带和编码音频信号的第二声道的所述频谱带是使用双-单声道编码来编码的还是使用中-侧编码来编码的。

[0035] -如果使用了双-单声道编码,则使用编码音频信号的第一声道的所述频谱带作为

中间音频信号的第一声道的频谱带,并且使用编码音频信号的第二声道的所述频谱带作为中间音频信号的第二声道的频谱带。

[0036] -如果使用了中-侧编码,则基于编码音频信号的第一声道的所述频谱带并且基于编码音频信号的第二声道的所述频谱带,来产生中间音频信号的第一声道的频谱带,以及基于编码音频信号的第一声道的所述频谱带并且基于编码音频信号的第二声道的所述频谱带,来产生中间音频信号的第二声道的频谱带。以及:

[0037] -根据去归一化值,修正中间音频信号的第一声道和第二声道中的至少一个声道,以获得解码音频信号的第一声道和第二声道。

[0038] 此外,提供了计算机程序,其中每个计算机程序被配置为当在计算机或信号处理器上执行时实现上述方法之一。

[0039] 根据实施例,提供了能够使用最小侧信息处理平移信号的新构思。

[0040] 根据一些实施例,如在[6a]和[6b]中结合如[8]中所述的频谱包络翘曲描述的那样来使用具有速率回路的FDNS(FDNS=频域噪声整形)。在一些实施例中,对FDNS白化频谱使用单个ILD参数,然后使用逐频带决策,无论使用M/S编码还是L/R编码来编码。在一些实施例中,M/S决策基于估计的比特节省。在一些实施例中,逐频带M/S处理声道之间的比特率分配可以例如取决于能量。

[0041] 一些实施例提供了对白化频谱应用单个全局ILD、之后是具有有效M/S决策机制以及具有控制单个全局增益的速率回路的逐频带M/S处理的组合。

[0042] 一些实施例尤其结合频谱包络翘曲(例如,基于[8])来采用具有速率回路的FDNS(例如,基于[6a]或[6b])。这些实施例提供了用于分离量化噪声的感知整形和速率回路的有效率且非常有作用的方式。对FDNS白化频谱使用单个ILD参数允许简单且有效的方式来决定是否在如上所述的M/S处理的优点。使频谱白化并去除ILD允许有效的M/S处理。对于所描述的系统来说编码单个全局ILD就足够了,因此与已知方法相比实现了比特节省。

[0043] 根据实施例,M/S处理基于感知白化信号完成。实施例确定编码阈值并以最优方式确定在处理感知白化和ILD补偿信号时是否采用L/R编码或M/S编码的决策。

[0044] 此外,根据实施例,提供了新的比特率估计。

[0045] 与[1]至[5]相反,在实施例中,感知模型与速率回路分离(如[6a]、[6b]和[13])。

[0046] 尽管如[1]中提出的那样M/S决策基于估计比特率,但与[1]相反,M/S和L/R编码的比特率需求的差异不依赖于通过感知模型确定的掩蔽阈值。相反,比特率需求是通过所使用的无损熵编码器来确定的。换言之:替代根据原始信号的感知熵导出比特率需求,比特率需求是根据感知白化信号的熵导出的。

[0047] 与[1]至[5]相反,在实施例中,M/S决策是基于感知白化信号来确定的,并且获得所需比特率的更好估计。为此,可以应用如[6a]或[6b]中所述的算术编码器比特消耗估计。不必明确考虑掩蔽阈值。

[0048] 在[1]中,假设中央声道和侧声道的掩蔽阈值是左掩蔽阈值和右掩蔽阈值中的最小值。频谱噪声整形在中央声道和侧声道上完成,并且可以例如基于这些掩蔽阈值。

[0049] 根据实施例,频谱噪声整形可以例如在左和右声道上进行,并且在这样的实施例中,感知包络可以在估计的地方精确地应用。

[0050] 此外,实施例基于以下发现:如果ILD存在(即,如果声道被平移),则M/S编码不是

有效的。为了避免这种情况,实施例对感知白化频谱使用单个ILD参数。

[0051] 根据一些实施例,提供了处理感知白化信号的M/S决策的新构思。

[0052] 根据一些实施例,编解码器使用不是经典音频编解码器(例如,如[1]中所述)的一部分的新构思。

[0053] 根据一些实施例,感知白化信号用于进一步编码,例如,类似于感知白化信号在语音编码器中使用的方式。

[0054] 这种方法具有若干优点,例如,简化了编解码器架构、实现了噪声整形特性和掩蔽阈值的复数表示(例如,作为LPC系数)。此外,变换和语音编解码器架构是统一的,因此能够实现组合的音频/语音编码。

[0055] 一些实施例采用全局ILD参数来有效地编码平移源。

[0056] 在实施例中,编解码器采用频域噪声整形(FDNS)以利用速率回路感知白化信号(例如,如在[6a]或[6b]中结合如[8]中所述的频谱包络翘曲描述的那样)。在这样的实施例中,编解码器可以例如对FDNS白化频谱进一步使用单个ILD参数,之后是逐频带M/S与L/R决策。逐频带M/S决策可以例如基于在以L/R和M/S模式编码时每个频带中的估计比特率。选择具有最少所需比特的模式。逐频带M/S处理声道之间的比特率分配基于能量。

[0057] 一些实施例使用熵编码器的每频带估计比特数对感知白化和ILD补偿频谱应用逐频带M/S决策。

[0058] 在一些实施例中,采用具有速率回路的FDNS(例如,如[6a]或[6b]中结合如[8]中描述的频谱包络翘曲描述的)。这提供了分离量化噪声的感知整形和速率回路的有效率的、非常起作用的方式。对FDNS白化频谱使用单个ILD参数允许简单且有效的方式来决定是否存在所述的M/S处理的优点。使频谱白化并去除ILD允许有效的M/S处理。对于所描述的系统来说编码单个全局ILD就足够了,因此与已知方法相比实现了比特节省。

[0059] 实施例修改了[1]中提供的在处理感知白化和ILD补偿信号时的构思。特别地,实施例对L、R、M和S采用相等的全局增益,该全局增益与FDNS一起形成编码阈值。全局增益可以根据SNR估计或根据一些其它构思导出。

[0060] 所提出的逐频带M/S决策精确地估计用算术编码器对每个频带进行编码所需的比特数。这是可能的,因为M/S决策是对白化频谱进行的,之后直接进行量化。不需要实验性搜索阈值。

附图说明

[0061] 以下,参考附图更详细地描述本发明的实施例,其中:

[0062] 图1a示出了根据实施例的用于编码的装置,

[0063] 图1b示出了根据另一实施例的用于编码的装置,其中该装置还包括变换单元和预处理单元,

[0064] 图1c示出了根据另一实施例的用于编码的装置,其中该装置还包括变换单元,

[0065] 图1d示出了根据另一实施例的用于编码的装置,其中该装置包括预处理单元和变换单元,

[0066] 图1e示出了根据另一实施例的用于编码的装置,其中该装置还包括频谱域预处理器,

- [0067] 图1f示出了根据实施例的用于对包括四个或更多个声道的音频输入信号中的四个声道进行编码以获得编码音频信号的四个声道的系统，
- [0068] 图2a示出了根据实施例的用于解码的装置，
- [0069] 图2b示出了根据实施例的用于解码的装置，其还包括变换单元和后处理单元，
- [0070] 图2c示出了根据实施例的用于解码的装置，其中用于解码的装置还包括变换单元，
- [0071] 图2d示出了根据实施例的用于解码的装置，其中用于解码的装置还包括后处理单元，
- [0072] 图2e示出了根据实施例的用于解码的装置，其中该装置还包括频谱域后处理器，
- [0073] 图2f示出了根据实施例的用于对包括四个或更多个声道的编码音频信号进行解码以获得包括四个或更多个声道的解码音频信号的四个声道的系统，
- [0074] 图3示出了根据实施例的系统，
- [0075] 图4示出了根据另一实施例的用于编码的装置，
- [0076] 图5示出了根据实施例的用于编码的装置中的立体声处理模块，
- [0077] 图6示出了根据另一实施例的用于解码的装置，
- [0078] 图7示出了根据实施例的针对逐频带M/S决策的比特率的计算，
- [0079] 图8示出了根据实施例的立体声模式决策，
- [0080] 图9示出了根据实施例的编码器侧的采用立体声填充的立体声处理，
- [0081] 图10示出了根据实施例的解码器侧的采用立体声填充的立体声处理，
- [0082] 图11示出了根据一些特定实施例的解码器侧的侧信号的立体声填充，
- [0083] 图12示出了根据实施例的编码器侧的不采用立体声填充的立体声处理，以及
- [0084] 图13示出了根据实施例的解码器侧的不采用立体声填充的立体声处理。

具体实施方式

[0085] 图1a示出了根据实施例的用于对包括两个或更多个声道的音频输入信号的第一声道和第二声道进行编码以获得编码音频信号的装置。

[0086] 该装置包括归一化器110，归一化器110被配置为根据音频输入信号的第一声道并且根据音频输入信号的第二声道来确定音频输入信号的归一化值。归一化器110被配置为通过根据归一化值修正音频输入信号的第一声道和第二声道中的至少一个声道来确定归一化音频信号的第一声道和第二声道。

[0087] 例如，在实施例中，归一化器110可以被配置为根据音频输入信号的第一声道和第二声道的多个频谱带确定音频输入信号的归一化值，归一化器110例如可以被配置为通过根据归一化值修正音频输入信号的第一声道和第二声道中的至少一个声道的多个频谱带来确定归一化音频信号的第一声道和第二声道。

[0088] 或者，例如，归一化器110可以例如被配置为根据在时域中表示的音频输入信号的第一声道并且根据在时域中表示的音频输入信号的第二声道来确定音频输入信号的归一化值。此外，归一化器110被配置为通过根据归一化值修正在时域中表示的音频输入信号的第一声道和第二声道中的至少一个声道来确定归一化音频信号的第一声道和第二声道。该装置还包括变换单元(图1a中未示出)，变换单元被配置为将归一化音频信号从时域变换到

频谱域,使得归一化音频信号在频谱域中表示。变换单元被配置为将在频谱域中表示的归一化音频信号馈送到编码单元120中。例如,音频输入信号可以是例如时域残差信号,其由LPC(LPC=线性预测编码)滤波时域音频信号的两个声道产生。

[0089] 此外,该装置包括编码单元120,编码单元120被配置为产生具有第一声道和第二声道的处理后的音频信号,使得处理后的音频信号的第一声道的一个或多个频谱带是归一化音频信号的第一声道的一个或多个频谱带,使得处理后的音频信号的第二声道的一个或多个频谱带是归一化音频信号的第二声道的一个或多个频谱带,使得处理后的音频信号的第一声道的至少一个频谱带是根据归一化音频信号的第一声道的频谱带并且根据归一化音频信号的第二声道的频谱带的中央信号的频谱带,以及使得处理后的音频信号的第二声道的至少一个频谱带是根据归一化音频信号的第一声道的频谱带并且根据归一化音频信号的第二声道的频谱带的侧信号的频谱带。编码单元120被配置为对处理后的音频信号进行编码以获得编码音频信号。

[0090] 在实施例中,编码单元120可以例如被配置为根据归一化音频信号的第一声道的多个频谱带并且根据归一化音频信号的第二声道的多个频谱带,在全-中-侧编码模式、全-双-单声道编码模式和逐频带编码模式之间选择。

[0091] 在这样的实施例中,编码单元120可以例如被配置为:如果选择全-中-侧编码模式,则根据归一化音频信号的第一声道并且根据归一化音频信号的第二声道产生中央信号作为中-侧信号的第一声道,根据归一化音频信号的第一声道和根据归一化音频信号的第二声道产生侧信号作为中-侧信号的第二声道,以及编码中-侧信号以获得编码音频信号。

[0092] 根据这样的实施例,编码单元120可以例如被配置为如果选择全-双-单声道编码模式,则对归一化音频信号进行编码以获得编码音频信号。

[0093] 此外,在这样的实施例中,编码单元120可以例如被配置为:如果选择逐频带编码模式,则产生处理后的音频信号,使得处理后的音频信号的第一声道的一个或多个频谱带是归一化音频信号的第一声道的一个或多个频谱带,使得处理后的音频信号的第二声道的一个或多个频谱带是归一化音频信号的第二声道的一个或多个频谱带,使得处理后的音频信号的第一声道的至少一个频谱带是根据归一化音频信号的第一声道的频谱带并且根据归一化音频信号的第二声道的频谱带的中央信号的频谱带,以及使得处理后的音频信号的第二声道的至少一个频谱带是根据归一化音频信号的第一声道的频谱带并且根据归一化音频信号的第二声道的频谱带的侧信号的频谱带,其中编码单元120可以例如被配置为对处理后的音频信号进行编码以获得编码音频信号。

[0094] 根据实施例,音频输入信号可以是例如恰好包括两个声道的音频立体声信号。例如,音频输入信号的第一声道可以例如是音频立体声信号的左声道,并且音频输入信号的第二声道可以例如是音频立体声信号的右声道。

[0095] 在实施例中,编码单元120可以例如被配置为:如果选择逐频带编码模式,则针对处理后的音频信号的多个频谱带中的每个频谱带,决定是采用中-侧编码还是采用双-单声道编码。

[0096] 如果针对所述频谱带采用中-侧编码,则编码单元120可以例如被配置为基于归一化音频信号的第一声道的所述频谱带并且基于归一化音频信号的第二声道的所述频谱带,来产生处理后的音频信号的第一声道的所述频谱带作为中央信号的频谱带。编码单元120

可以例如被配置为基于归一化音频信号的第一声道的所述频谱带并且基于归一化音频信号的第二声道的所述频谱带,来产生处理后的音频信号的第二声道的所述频谱带作为侧信号的频谱带。

[0097] 如果针对所述频谱带采用双-单声道编码,则编码单元120可以例如被配置为使用归一化音频信号的第一声道的所述频谱带作为处理后的音频信号的第一声道的所述频谱带,并且可以例如被配置为使用归一化音频信号的第二声道的所述频谱带作为处理后的音频信号的第二声道的所述频谱带。或者,编码单元120被配置为使用归一化音频信号的第二声道的所述频谱带作为处理后的音频信号的第一声道的所述频谱带,并且可以例如被配置为使用归一化音频信号的第一声道的所述频谱带作为处理后的音频信号的第二声道的所述频谱带。

[0098] 根据实施例,编码单元120可以例如被配置为:通过确定估计在采用全-中-侧编码模式时编码所需的第一比特数的第一估计,通过确定估计在采用全-双-单声道编码模式时编码所需的第二比特数的第二估计,通过确定估计在可以例如采用逐频带编码模式时编码所需的第三比特数的第三估计,以及通过在全-中-侧编码模式、全-双-单声道编码模式和逐频带编码模式之中选择具有第一估计、第二估计和第三估计之中的最小比特数的编码模式,来在全-中-侧编码模式、全-双-单声道编码模式、和逐频带编码模式之间进行选择。

[0099] 在实施例中,编码单元120可以例如被配置为根据以下公式估计第三估计 b_{BW} ,从而估计在采用逐频带编码模式时编码所需的第三比特数:

$$[0100] \quad b_{BW} = nBands + \sum_{i=0}^{nBands-1} \min(b_{BWLR}^i, b_{BWS}^i)$$

[0101] 其中, $nBands$ 是归一化音频信号的频谱带数,其中 b_{BWS}^i 是对中央信号的第*i*个频谱带进行编码和对侧信号的第*i*个频谱带进行编码所需的比特数的估计,并且其中 b_{BWLR}^i 是对第一信号的第*i*个频谱带进行编辑和对第二信号的第*i*个频谱带进行编辑所需的比特数的估计。

[0102] 在实施例中,可以例如采用用于在全-中-侧编码模式、全-双-单声道编码模式以及逐频带编码模式之间进行选择的客观质量测量。

[0103] 根据实施例,编码单元120可以例如被配置为:通过确定估计在以全-中-侧编码模式进行编码时所保存的第一比特数的第一估计,通过确定估计在以全-双-单声道编码模式进行编码时所保存的第二比特数的第二估计,通过确定估计在以逐频带编码模式进行编码时所保存的第三比特数的第三估计,以及通过在全-中-侧编码模式、全-双-单声道编码模式和逐频带编码模式之中选择具有第一估计、第二估计和第三估计之中的所保存的最大比特数的编码模式,来在全-中-侧编码模式、全-双-单声道编码模式、和逐频带编码模式之间进行选择。

[0104] 在另一实施例中,编码单元120可以例如被配置为:通过估计在采用全-中-侧编码模式时发生的第一信噪比,通过估计在采用全-双-单声道编码模式时发生的第二信噪比,通过估计在采用逐频带编码模式时发生的第三信噪比,并且通过在全-中-侧编码模式、全-双-单声道编码模式和逐频带编码模式之中选择具有第一信噪比、第二信噪比和第三信噪

比之中的最大信噪比的编码模式,来在全-中-侧编码模式、全-双-单声道编码模式和逐频带编码模式之间进行选择。

[0105] 在实施例中,归一化器110可以例如被配置为根据音频输入信号的第一声道的能量并且根据音频输入信号的第二声道的能量来确定音频输入信号的归一化值。

[0106] 根据实施例,音频输入信号可以例如在频谱域中表示。归一化器110可以例如被配置为根据音频输入信号的第一声道的多个频谱带并且根据音频输入信号的第二声道的多个频谱带来确定音频输入信号的归一化值。此外,归一化器110可以例如被配置为通过根据归一化值修正音频输入信号的第一声道和第二声道中的至少一个声道的多个频谱带来确定归一化音频信号。

[0107] 在实施例中,归一化器110可以例如被配置为基于以下公式确定归一化值:

$$[0108] \quad NRG_L = \sqrt{\sum MDCT_{L,k}^2}$$

$$[0109] \quad NRG_R = \sqrt{\sum MDCT_{R,k}^2}$$

$$[0110] \quad ILD = \frac{NRG_L}{NRG_L + NRG_R}$$

[0111] 其中, $MDCT_{L,k}$ 是音频输入信号的第一声道的MDCT频谱的第k个系数,并且 $MDCT_{R,k}$ 是音频输入信号的第二声道的MDCT频谱的第k个系数。归一化器110可以例如被配置为通过量化ILD来确定归一化值。

[0112] 根据图1b所示的实施例,用于编码的装置可以例如还包括变换单元102和预处理单元105。变换单元102可以例如被配置为将时域音频信号从时域变换到频域以获得变换后的音频信号。预处理单元105可以例如被配置为通过对变换后的音频信号应用编码器侧频域噪声整形操作来产生音频输入信号的第一声道和第二声道。

[0113] 在特定实施例中,预处理单元105可以例如被配置为通过在对变换后的音频信号应用编码器侧频域噪声整形操作之前,对变换后的音频信号应用编码器侧时间噪声整形操作,来产生音频输入信号的第一声道和第二声道。

[0114] 图1c示出了根据另一实施例的用于编码的装置还包括变换单元115。归一化器110可以例如被配置为根据在时域中表示的音频输入信号的第一声道并且根据在时域中表示的音频输入信号的第二声道来确定音频输入信号的归一化值。此外,归一化器110可以例如被配置为通过根据归一化值修正在时域中表示的音频输入信号的第一声道和第二声道中的至少一个声道来确定归一化音频信号的第一声道和第二声道。变换单元115可以例如被配置为将归一化音频信号从时域变换到频谱域,使得归一化音频信号在频谱域中表示。此外,变换单元115可以例如被配置为将在频谱域中表示的归一化音频信号馈送到编码单元120中。

[0115] 图1d示出了根据另一实施例的用于编码的装置,其中该装置还包括被配置为接收包括第一声道和第二声道的时域音频信号的预处理单元106。预处理单元106可以例如被配

置为对时域音频信号中的、产生第一感知白化频谱的第一声道应用滤波器,以获得在时域中表示的音频输入信号的第一声道。预处理单元106可以例如被配置为对时域音频信号中的、产生第二感知白化频谱的第二声道应用滤波器,以获得在时域中表示的音频输入信号的第二声道。

[0116] 在实施例中,如图1e所示,变换单元115可以例如被配置为将归一化音频信号从时域变换到频谱域,以获得变换后的音频信号。在图1e的实施例中,该装置还包括频谱域预处理器118,频谱域预处理器118被配置为对变换后的音频信号执行编码器侧时间噪声整形,以获得在频谱域中表示的归一化音频信号。

[0117] 根据实施例,编码单元120可以例如被配置为通过对归一化音频信号或处理后的音频信号应用编码器侧立体声智能间隙填充来获得编码音频信号。

[0118] 在另一实施例中,如图1f所示,提供了一种用于对包括四个或更多个声道的音频输入信号的四个声道地进行编码以获得编码音频信号的系统。该系统包括根据上述实施例之一的第一装置170,第一装置170用于对音频输入信号的四个或更多个声道中的第一声道和第二声道进行编码,以获得编码音频信号的第一声道和第二声道。此外,该系统包括根据上述实施例之一的第二装置180,第二装置180用于对音频输入信号的四个或更多个声道中的第三声道和第四声道进行编码,以获得编码音频信号的第三声道和第四声道。

[0119] 图2a示出了根据实施例的用于对包括第一声道和第二声道的编码音频信号进行解码以获得解码音频信号的装置。

[0120] 用于解码的装置包括解码单元210,解码单元210被配置为针对多个频谱带中的每个频谱带,来确定编码音频信号的第一声道的所述频谱带和编码音频信号的第二声道的所述频谱带是使用双-单声道编码来编码的还是使用中-侧编码来编码的。

[0121] 如果使用了双-单声道编码,则解码单元210被配置为使用编码音频信号的第一声道的所述频谱带作为中间音频信号的第一声道的频谱带,并且被配置为使用编码音频信号的第二声道的所述频谱带作为中间音频信号的第二声道的频谱带。

[0122] 此外,如果使用了中-侧编码,则解码单元210被配置为基于编码音频信号的第一声道的所述频谱带并且基于编码音频信号的第二声道的所述频谱带来产生中间音频信号的第一声道的频谱带,以及基于编码音频信号的第一声道的所述频谱带并且基于编码音频信号的第二声道的所述频谱带,来产生中间音频信号的第二声道的频谱带。

[0123] 此外,用于解码的装置包括去归一化器220,去归一化器220被配置为根据去归一化值来修正中间音频信号的第一声道和第二声道中的至少一个声道,以获得解码音频信号的第一声道和第二声道。

[0124] 在实施例中,解码单元210可以例如被配置为确定编码音频信号是以全-中-侧编码模式、以全-双-单声道编码模式、还是以逐频带编码模式来编码的。

[0125] 此外,在这样的实施例中,解码单元210可以例如被配置为:如果确定编码音频信号是以全-中-侧编码模式编码的,则根据编码音频信号的第一声道并且根据编码音频信号的第二声道来产生中间音频信号的第一声道,以及根据编码音频信号的第一声道并且根据编码音频信号的第二声道来产生中间音频信号的第二声道。

[0126] 根据这样的实施例,解码单元210可以例如被配置为:如果确定编码音频信号是以全-双-单声道编码模式编码的,则使用编码音频信号的第一声道作为中间音频信号的第一

声道,以及使用编码音频信号的第二声道作为中间音频信号的第二声道。

[0127] 此外,在这样的实施例中,解码单元210可以例如被配置为如果确定编码音频信号是以逐频带编码模式编码的,则:

[0128] -针对多个频谱带中的每个频谱带,确定编码音频信号的第一声道的所述频谱带和编码音频信号的第二声道的所述频谱带是使用双-单声道编码来编码的还是使用中-侧编码来编码的,

[0129] -如果使用了双-单声道编码,则使用编码音频信号的第一声道的所述频谱带作为中间音频信号的第一声道的频谱带,并且使用编码音频信号的第二声道的所述频谱带作为中间音频信号的第二声道的频谱带,以及

[0130] -如果使用了中-侧编码,则基于编码音频信号的第一声道的所述频谱带并且基于编码音频信号的第二声道的所述频谱带,来产生中间音频信号的第一声道的频谱带,以及基于编码音频信号的第一声道的所述频谱带并且基于编码音频信号的第二声道的所述频谱带,来产生中间音频信号的第二声道的频谱带。

[0131] 例如,在全-中-侧编码模式下,例如可以应用以下公式:

[0132] $L = (M+S) / \sqrt{2}$, 以及

[0133] $R = (M-S) / \sqrt{2}$

[0134] 来获得中间音频信号的第一声道L并获得中间音频信号的第二声道R,其中M是编码音频信号的第一声道,S是编码音频信号的第二声道。

[0135] 根据实施例,解码输入信号可以是例如恰好包括两个声道的音频立体声信号。例如,解码音频信号的第一声道可以例如是音频立体声信号的左声道,并且解码音频信号的第二声道可以例如是音频立体声信号的右声道。

[0136] 根据实施例,去归一化器220可以例如被配置为根据去归一化值来修正中间音频信号的第一声道和第二声道中的至少一个声道的多个频谱带,获得解码音频信号的第一声道和第二声道。

[0137] 在图2b中所示的另一实施例中,去归一化器220可以例如被配置为根据去归一化值来修正中间音频信号的第一声道和第二声道中的至少一个声道的多个频谱带,以获得去归一化音频信号。在这样的实施例中,该装置可以例如还包括后处理单元230和变换单元235。后处理单元230可以例如被配置为对去归一化音频信号执行解码器侧时间噪声整形和解码器侧频域噪声整形中的至少一个,以获得后处理音频信号。变换单元(235)可以例如被配置为将后处理音频信号从频谱域变换到时域,以获得解码音频信号的第一声道和第二声道。

[0138] 根据图2c所示的实施例,该装置还包括被配置为将中间音频信号从频谱域变换到时域的变换单元215。去归一化器220可以例如被配置为根据去归一化值来修正在时域中表示的中间音频信号的第一声道和第二声道中的至少一个声道,以获得解码音频信号的第一声道和第二声道。

[0139] 在图2d所示的类似实施例中,变换单元215可以例如被配置为将中间音频信号从频谱域变换到时域。去归一化器220可以例如被配置为根据去归一化值来修正在时域中表示的中间音频信号的第一声道和第二声道中的至少一个声道,以获得去归一化音频信号。该装置还包括后处理单元235,后处理单元235可以例如被配置为处理去归一化音频信号

(作为感知白化音频信号),以获得解码音频信号的第一声道和第二声道。

[0140] 根据如图2e所示的另一实施例,该装置还包括被配置为对中间音频信号执行解码器侧时间噪声整形的频谱域后处理器212。在这样的实施例中,变换单元215被配置为在已经对中间音频信号执行了解码器侧时间噪声整形之后,将中间音频信号从频谱域变换到时域。

[0141] 在另一实施例中,解码单元210可以例如被配置为对编码音频信号应用解码器侧立体声智能间隙填充。

[0142] 此外,如图2f所示,提供了一种用于对包括四个或更多个声道的编码音频信号进行解码以获得包括四个或更多个声道的解码音频信号的四个声道的系统。该系统包括根据上述实施例之一的第一装置270,第一装置270用于对具有四个或更多个声道的编码音频信号中的第一声道和第二声道进行解码,以获得解码音频信号的第一声道和第二声道。该系统包括根据上述实施例之一的第二装置280,第二装置280用于对编码音频信号的四个或更多个声道中的第三声道和第四声道进行解码,以获得解码音频信号的第三声道和第四声道。

[0143] 图3示出了根据实施例的用于根据音频输入信号来产生编码音频信号以及用于根据编码音频信号来产生解码音频信号的系统。

[0144] 该系统包括根据上述实施例之一的用于编码的装置310,其中用于编码的装置310被配置为根据音频输入信号来产生编码音频信号。

[0145] 此外,该系统包括如上所述的用于解码的装置320。用于解码的装置320被配置为根据编码音频信号来产生解码音频信号。

[0146] 类似地,提供了一种用于根据音频输入信号来产生编码音频信号以及根据编码音频信号来产生解码音频信号的系统。该系统包括根据图1f的实施例的系统以及根据图2f的实施例的系统,其中根据图1f的实施例的系统被配置为根据音频输入信号来产生编码音频信号,其中图2f的实施例的系统被配置为根据编码音频信号来产生解码音频信号。

[0147] 在下文中,描述了优选实施例。

[0148] 图4示出了根据另一实施例的用于编码的装置。尤其是,示出了根据特定实施例的预处理单元105和变换单元102。变换单元102尤其被配置为将音频输入信号从时域变换到频谱域,并且变换单元被配置为对音频输入信号执行编码器侧时间噪声整形和编码器侧频域噪声整形。

[0149] 此外,图5示出了根据实施例的用于编码的装置中的立体声处理模块。图5示出了归一化器110和编码单元120。

[0150] 此外,图6示出了根据另一实施例的用于解码的装置。尤其是,图6示出了根据特定实施例的后处理单元230。后处理单元230尤其被配置为从去归一化器220获得处理后的音频信号,并且后处理单元230被配置为对处理后的音频信号执行解码器侧时间噪声整形和解码器侧频域噪声整形中的至少一个。

[0151] 时域瞬态检测器(TD TD)、加窗、MDCT、MDST和OLA可以例如如[6a]或[6b]中所述的那样进行。MDCT和MDST形成复数调制重叠变换(MCLT);单独地执行MDCT和MDST相当于执行MCLT;“MCLT到MDCT”表示仅采用MCLT的MDCT部分并丢弃MDST(参见[12])。

[0152] 在左声道和右声道中选择不同的窗口长度可以例如在该帧中强制执行双-单声道

编码。

[0153] 时间噪声整形 (TNS) 可以例如与 [6a] 或 [6b] 中描述的那样类似地进行。

[0154] 频域噪声整形 (FDNS) 和对 FDNS 参数的计算可以例如类似于 [8] 中描述的处理。例如, 一个差异可以是根据 MCLT 频谱计算针对 TNS 不活跃的帧的 FDNS 参数。在 TNS 是活跃的帧中, 可以例如根据 MDCT 来估计 MDST。

[0155] FDNS 也可以用时域中的感知频谱白化替代 (例如, 如 [13] 中所述)。

[0156] 立体声处理由全局 ILD 处理、逐频带 M/S 处理、声道间的比特率分配组成。

[0157] 单个全局 ILD 被计算为:

$$[0158] \quad NRG_L = \sqrt{\sum MDCT_{L,k}^2}$$

$$[0159] \quad NRG_R = \sqrt{\sum MDCT_{R,k}^2}$$

$$[0160] \quad ILD = \frac{NRG_L}{NRG_L + NRG_R}$$

[0161] 其中, $MDCT_{L,k}$ 是左声道中的 MDCT 频谱的第 k 个系数, $MDCT_{R,k}$ 是右声道中的 MDCT 频谱的第 k 个系数。全局 ILD 被均匀量化为:

$$[0162] \quad \widehat{ILD} = \max\left(1, \min(ILD_{range} - 1, \lfloor ILD_{range} \cdot ILD + 0.5 \rfloor)\right)$$

$$[0163] \quad ILD_{range} = 1 \ll ILD_{bits}$$

[0164] 其中, ILD_{bits} 是用于编码全局 ILD 的比特数。 \widehat{ILD} 存储在比特流中。

[0165] 《是比特移位操作, 通过插入 0 比特将比特向左移位 ILD_{bits} 。

[0166] 换言之: $ILD_{range} = 2^{ILD_{bits}}$ 。

[0167] 则, 声道的能量比是:

$$[0168] \quad ratio_{ILD} = \frac{ILD_{range}}{\widehat{ILD}} - 1 \approx \frac{NRG_R}{NRG_L}$$

[0169] 如果 $ratio_{ILD} > 1$, 则右声道以 $\frac{1}{ratio_{ILD}}$ 来缩放, 否则左声道以 $ratio_{ILD}$ 来缩放。这实际上意味着更大声的声道被缩放了。

[0170] 如果使用时域中的感知频谱白化 (例如, 如 [13] 中所述), 则在时域到频域的变换之前 (即, 在 MDCT 之前), 也可以在时域中计算和应用单个全局 ILD。或者, 备选地, 感知频谱白化之后可以是时域到频域变换, 之后是在频域中的单个全局 ILD。备选地, 可以在到时域到频域变换之前在时域中计算单个全局 ILD, 并且在时域到频域变换之后在频域中应用所计算出的单个全局 ILD。

[0171] 中央声道 $MDCT_{M,k}$ 和侧声道 $MDCT_{S,k}$ 是通过使用左声道 $MDCT_{L,k}$ 和右声道 $MDCT_{R,k}$ 、依据

$MDCT_{M,k} = \frac{1}{\sqrt{2}}(MDCT_{L,k} + MDCT_{R,k})$ 和 $MDCT_{S,k} = \frac{1}{\sqrt{2}}(MDCT_{L,k} - MDCT_{R,k})$ 而形成的。频谱被

划分为频带,并且针对每个频带,决定是使用左声道、右声道、中央声道还是侧声道。

[0172] 对包括级联的左声道和右声道的信号估计全局增益 G_{est} 。因此不同于[6b]和[6a]。例如,假设来自标量量化的每比特每样本的SNR增益为6dB,可以使用如[6b]或[6a]的第5.3.3.2.8.1.1节“Global gain estimator”中描述的增益的第一估计。

[0173] 所估计的增益可以乘以常数以得到低估或高估的最终 G_{est} 。然后,使用 G_{est} 来量化左声道、右声道、中央声道和侧声道中的信号,即,量化步长为 $1/G_{est}$ 。

[0174] 然后使用算术编码器、霍夫曼编码器或任何其它熵编码器对量化后的信号进行编码,以便获得所需比特数。例如,可以使用在[6b]或[6a]的节5.3.3.2.8.1.3至节5.3.3.2.8.1.7中描述的基于上下文的算术编码器。由于将在立体声编码之后运行速率回路(例如,[6b]中或[6a]中的5.3.3.2.8.1.2),因此所需比特的估计是足够的。

[0175] 例如,对于每个量化声道,如[6b]的或[6a]的节5.3.3.2.8.1.3至节5.3.3.2.8.1.7中所述的那样来估计基于上下文的算数编码所需的比特数。

[0176] 根据实施例,基于以下示例代码来确定每个量化声道(左、右、中或侧)的比特估计:

```
int context_based_arithmetic_coder_estimate (
    int spectrum[],
    int start_line,
    int end_line,
    int lastnz, // lastnz = last non-zero spectrum line
    int & ctx, // ctx = context
    int & probability, // 14 bit fixed point probability
    const unsigned int cum_freq[N_CONTEXTS][]
[0177] // cum_freq = cumulative frequency tables, 14 bit fixed point
)
{
    int nBits = 0;

    for (int k = start_line; k < min(lastnz, end_line); k+=2)
    {
        int a1 = abs(spectrum[k]);
```

```
int b1 = abs(spectrum[k+1]);

/* Signs Bits */
nBits += min(a1, 1);
nBits += min(b1, 1);

while (max(a1, b1) >= 4)
{
    probability *= cum_freq[ctx][VAL_ESC];

    int nlz = Number_of_leading_zeros(probability);
    nBits += 2 + nlz;
    probability >>= 14 - nlz;

    a1 >>= 1;
    b1 >>= 1;

    ctx = update_context(ctx, VAL_ESC);
}

int symbol = a1 + 4*b1;
probability *= (cum_freq[ctx][symbol] -
               cum_freq[ctx][symbol+1]);

int nlz = Number_of_leading_zeros(probability);
nBits += nlz;
hContextMem->proba >>= 14 - nlz;

ctx = update_context(ctx, a1+b1);
}
```

[0178]

```

    return nBits;
[0179] }

```

[0180] 其中,spectrum被设置为指向要被编码的量化频谱,start_line被设置为0,end_line被设置为频谱的长度,lastnz被设置为频谱的最后一个非零元素的索引,ctx被设置为0,并且probability被设置为在14比特定点数表示法下的1(16384=1<<14)。

[0181] 如所概述的,例如,可以采用上述示例代码来获得针对左声道、右声道、中央声道和侧声道中的至少一个声道的比特估计。

[0182] 一些实施例采用如[6b]和[6a]中所述的算术编码器。进一步的细节可以在例如[6b]的节5.3.3.2.8“Arithmetic coder”中找到。

[0183] 然后,针对“全-双-单声道”的估计比特数(b_{LR})等于左和右声道所需的比特之和。

[0184] 然后,针对“全M/S”的估计比特数(b_{MS})等于中央声道和侧声道所需的比特之和。

[0185] 在为上述示例代码的备选项的备选实施例中,可以采用例如以下公式来计算针对“全-双-单声道”的估计比特数(b_{LR}):

$$[0186] \quad b_{LR} = \sum_{i=0}^{nBands-1} b_{bwLR}^i$$

[0187] 此外,在为上述示例代码的备选项的备选实施例中,可以采用例如以下公式来计算针对“全M/S”的估计比特数(b_{MS}):

$$[0188] \quad b_{MS} = \sum_{i=0}^{nBands-1} b_{bwMS}^i$$

[0189] 对于具有边界 $[lb_i, ub_i]$ 的每个频带*i*,检查在L/R模式下将有多少比特(b_{bwLR}^i)

用于编码频带中的量化信号和在M/S模式下将有多少比特(b_{bwMS}^i)用于编码频带中的量

化信号。换句话说,对于每个频带*i*针对L/R模式执行逐频带比特估计: b_{bwLR}^i 由此产生针对

频带*i*的L/R模式频带比特估计,并且对于每个频带*i*针对M/S模式执行逐频带比特估计,由此产生针对频带*i*的M/S模式逐频带比特估计: b_{bwMS}^i 。

[0190] 为频带选择利用较少比特的模式。如[6b]的或[6a]的节5.3.3.2.8.1.3至节5.3.3.2.8.1.7中所述的那样来估算数编码所需的比特数。在“逐频带M/S”模式下编码频谱所需的总比特数(b_{BW})等于 $\min(b_{bwLR}^i, b_{bwMS}^i)$ 之和:

$$[0191] \quad b_{BW} = nBands + \sum_{i=0}^{nBands-1} \min(b_{bwLR}^i, b_{bwMS}^i)$$

[0192] 无论是使用L/R还是M/S编码,“逐频带M/S”模式都需要用于在每个频带中发信号的附加比特nBands。在“逐频带M/S”、“全-双-单声道”和“全M/S”之间的选择可以例如作为

立体声模式被编码到比特流中,然后与“逐频带M/S”相比,“全-双-单声道”和“全M/S”无需用于发信号的附加比特。

[0193] 对于基于上下文的算术编码器,用于计算bLR的 b_{bwLR}^i 不等于用于计算bBW的 b_{bwLR}^i ,用于计算bMS的 b_{bwMS}^i 也不等于用于计算bBW的 b_{bwMS}^i ,因为 b_{bwLR}^i 和 b_{bwMS}^i 取决于针对先前的 b_{bwLR}^j 和 b_{bwMS}^j 的上下文的选择,其中 $j < i$ 。bLR可以被计算为针对左声道和针对右声道的比特的总和,并且bMS可以被计算为针对中央声道和针对侧声道的比特的总和,其中可以使用如下示例代码来计算针对每个声道的比特:context_based_arithmetic_coder_estimate_bandwise,其中start_line设置为0,并且end_line设置为lastnz。

[0194] 在为上述示例代码的备选项的备选实施例中,可以采用例如以下公式来计算针对“全-双-单声道”的估计比特数(b_{LR}),并且在每个频带中发信号时可以使用L/R编码:

$$[0195] \quad b_{LR} = nBands + \sum_{i=0}^{nBands-1} b_{bwLR}^i$$

[0196] 此外,在为上述示例代码的备选项的备选实施例中,可以采用例如以下公式来计算针对“全M/S”的估计比特数(b_{MS}),并且在每个频带中发信号时可以使用M/S编码:

$$[0197] \quad b_{MS} = nBands + \sum_{i=0}^{nBands-1} b_{bwMS}^i$$

[0198] 在一些实施例中,首先,可以例如估计增益G,并且可以例如估计量化步长,预期有足够的比特来编码L/R中的声道。

[0199] 在下文中,提供了描述如何确定逐频带比特估计的不同方式的实施例,例如,根据特定实施例,描述了如何确定 b_{bwLR}^i 和 b_{bwMS}^i 。

[0200] 如已经概述的,根据特定实施例,对于每个量化声道,例如如[6b]的节5.3.3.2.8.1.7“Bit consumption estimation”或者[6a]的类似节中描述的那样来估计算术编码所需的比特数。

[0201] 根据实施例,使用用于计算针对每个i的 b_{bwLR}^i 和 b_{bwMS}^i 中的每一个的context_based_arithmetic_coder_estimate,通过将start_line设置为 lb_i 、将end_line设置为 ub_i 、将lastnz设置为频谱的最后非零元素的索引来确定逐频带比特估计。

[0202] 初始化四个上下文($ctx_L, ctx_R, ctx_M, ctx_M$)和四个概率(p_L, p_R, p_M, p_M),然后对其重复更新。

[0203] 在估计开始时(对于 $i=0$),将每个上下文($ctx_L, ctx_R, ctx_M, ctx_M$)设置为0,并且将每个概率(p_L, p_R, p_M, p_M)设置为14比特定点数表示法下的1($16384=1 < 14$)。

[0204] b_{bwLR}^i 被计算为 b_{bwL}^i 和 b_{bwR}^i 之和,其中 b_{bwL}^i 是使用context_based_arithmetic_coder_estimate、通过将spectrum设置为指向要被编码的量化左频谱、将ctx设置为 ctx_L 、并且将probability设置为 p_L 来确定的,并且 b_{bwR}^i 是使用context_based_

arithmic_coder_estimate、通过将spectrum设置为指向要被编码的量化右频谱、将ctx设置为ctx_R、并且将probability设置为p_R来确定的。

[0205] b_{bwMS}^i 被计算为 b_{bwM}^i 和 b_{bwS}^i 之和,其中 b_{bwM}^i 是使用context_based_arithmetic_coder_estimate、通过将spectrum设置为指向要被编码的量化中央频谱、将ctx设置为ctx_M、并且将probability设置为p_M来确定的,并且 b_{bwS}^i 是使用context_based_arithmetic_coder_estimate、通过将spectrum设置为指向要被编码的量化侧频谱、将ctx设置为ctx_S、并且将probability设置为p_S来确定的。

[0206] 如果 $b_{bwLR}^i < b_{bwMS}^i$, 则将ctx_L设置为ctx_M,将ctx_R设置为ctx_S,将p_L设置为p_M,将p_R设置为p_S。

[0207] 如果 $b_{bwLR}^i \geq b_{bwMS}^i$, 则将ctx_M设置为ctx_L,将ctx_S设置为ctx_R,将p_M设置为p_L,将p_S设置为p_R。

[0208] 在备选实施例中,如下获得逐频带比特估计:

[0209] 频谱被划分为频带,并且对于每个频带,决定是否应该进行M/S处理。对于使用M/S的所有频带,MDCT_{L,k}和MDCT_{R,k}被替代为MDCT_{M,k} = 0.5 (MDCT_{L,k} + MDCT_{R,k})和MDCT_{S,k} = 0.5 (MDCT_{L,k} - MDCT_{R,k})。

[0210] 逐频带M/S与L/R决策可以例如基于M/S处理情况下保存的估计比特:

$$[0211] \quad bitsSaved_i = nlines_i \cdot \log_2 \sqrt{\frac{NRG_{R,i} NRG_{L,i}}{NRG_{M,i} NRG_{S,i}}}$$

[0212] 其中,NRG_{R,i}是右声道的第i个频带中的能量,NRG_{L,i}是左声道的第i个频带中的能量,NRG_{M,i}是中央声道的第i个频带中的能量,NRG_{S,i}是侧声道的第i个频带中的能量,并且nlines_i是第i个频带中的频谱系数的数量。中央声道是左和右声道之和,侧声道是左和右声道之差。

[0213] bitsSaved_i受限于将用于第i个频带的估计比特数:

$$[0214] \quad maxBits_{LR} = \left(\frac{NRG_{R,i}}{NRG_R} + \frac{NRG_{L,i}}{NRG_L} \right) \cdot bitsAvailable$$

$$[0215] \quad maxBits_{MS} = \left(\frac{NRG_{M,i}}{NRG_M} + \frac{NRG_{S,i}}{NRG_S} \right) \cdot bitsAvailable$$

$$[0216] \quad \overline{bitsSaved}_i = \max \left(maxBits_{LR}, \min(-maxBits_{MS}, bitsSaved_i) \right)$$

[0217] 图7示出了根据实施例的计算针对逐频带M/S决策的比特率。

[0218] 特别地,在图7中,描绘了用于计算b_{BW}的处理。为了降低复杂度,保存直至频带i-1

的用于编码频谱的算术编码器上下文,并且在频带*i*中重新使用所保存的算术编码器上下文。

[0219] 应当注意,对于基于上下文的算术编码器, b_{bwLR}^i 和 b_{bwMS}^i 取决于算术编码器上下文,而该算术编码器上下文取决于在小于*i*的所有频带*j*中的M/S与L/R选择(例如如上所述的那样)。

[0220] 图8示出了根据实施例的立体声模式决策。

[0221] 如果选择“全-双-单声道”,则完整频谱由 $MDCT_{L,k}$ 和 $MDCT_{R,k}$ 组成。如果选择“全M/S”,则完整频谱由 $MDCT_{M,k}$ 和 $MDCT_{S,k}$ 组成。如果选择“逐频带M/S”,则频谱的一些频带由 $MDCT_{L,k}$ 和 $MDCT_{R,k}$ 组成,并且其它频带由 $MDCT_{M,k}$ 和 $MDCT_{S,k}$ 组成。

[0222] 立体声模式被编码到比特流中。在“逐频带M/S”模式中,还将逐频带M/S决策编码到比特流中。

[0223] 在立体声处理后两个声道中的频谱的系数表示为 $MDCT_{LM,k}$ 和 $MDCT_{RS,k}$ 。 $MDCT_{LM,k}$ 根据立体声模式和逐频带M/S决策,等于M/S频带中的 $MDCT_{M,k}$ 或者L/R频带中的 $MDCT_{L,k}$,并且 $MDCT_{RS,k}$ 等于M/S频带中的 $MDCT_{S,k}$ 或者L/R频带中的 $MDCT_{R,k}$ 。由 $MDCT_{LM,k}$ 组成的频谱可以例如称为联合编码声道0(联合Chn 0),或者可以例如称为第一声道,并且由 $MDCT_{RS,k}$ 组成的频谱可以例如称为联合编码声道1(联合Chn 1)或者可以例如被称为第二声道。

[0224] 使用立体声处理声道的能量来计算比特率拆分比:

$$[0225] \quad NRG_{LM} = \sqrt{\sum MDCT_{LM,k}^2}$$

$$[0226] \quad NRG_{RS} = \sqrt{\sum MDCT_{RS,k}^2}$$

$$[0227] \quad r_{split} = \frac{NRG_{LM}}{NRG_{LM} + NRG_{RS}}$$

[0228] 比特率拆分比被均匀量化为:

$$[0229] \quad \widehat{r}_{split} = \max\left(1, \min\left(rsplit_{range} - 1, \lceil rsplit_{range} \cdot r_{split} + 0.5 \rceil\right)\right)$$

$$[0230] \quad rsplit_{range} = 1 \ll rsplit_{bits}$$

[0231] 其中, $rsplit_{bits}$ 是用于编码比特率拆分比的比特数。如果 $r_{split} < \frac{8}{9}$ 并且

$$\widehat{r}_{split} > \frac{9rsplit_{range}}{16}, \text{ 则 } \widehat{r}_{split} \text{ 减少 } \frac{rsplit_{range}}{8}。 \text{ 如果 } r_{split} > \frac{1}{9} \text{ 并且 } \widehat{r}_{split} < \frac{7rsplit_{range}}{16},$$

则 \widehat{r}_{split} 增加 $\frac{rsplit_{range}}{8}$ 。 \widehat{r}_{split} 存储在比特流中。

[0232] 声道间的比特率分配为:

$$[0233] \quad bits_{LM} = \left\lfloor \frac{\widehat{r}_{split}}{rsplit_{range}} (totalBitsAvailable - stereoBits) \right\rfloor$$

$$[0234] \quad bits_{RS} = (totalBitsAvailable - stereoBits) - bits_{LM}$$

[0235] 此外,通过检查 $bits_{LM} - sideBits_{LM} > minBits$ 和 $bits_{RS} - sideBits_{RS} > minBits$,来确保每个声道中用于熵编码器的比特是足够的,其中 $minBits$ 熵编码器所需的最小比特数。如果用于熵编码器的比特不足够,则将 \widehat{r}_{split} 增加/减少1,直到满足 $bits_{LM} - sideBits_{LM} > minBits$ 和 $bits_{RS} - sideBits_{RS} > minBits$ 。

[0236] 量化、噪声填充和熵编码,包括速率回路,如[6b]中或[6a]中5.3.3“MDCT based TCX”的5.3.3.2“General encoding procedure”中所述的那样。可以使用估计的 G_{est} 来优化速率回路。功率谱P(MCLT的幅度)用于量化和智能间隙填充(IGF)中的音调/噪声测量,如[6a]或[6b]中所述。由于白化和逐频带M/S处理的MDCT频谱用于功率谱,因此将对MDST频谱进行相同的FDNS和M/S处理。将如同针对MDCT所做的那样,针对MDST进行基于更大声的声道的全局ILD的相同缩放。对于TNS是活跃的帧,用于功率谱计算的MDST频谱是根据白化和M/S处理的MDCT频谱估计的: $P_k = MDCT_k^2 + (MDCT_{k+1} - MDCT_{k-1})^2$ 。

[0237] 解码处理开始于联合编码声道的频谱的解码和逆量化,之后为如[6b]或[6a]中的6.2.2“MDCT based TCX”中所述的噪声填充。分配给每个声道的比特数是基于被编码到比特流中的窗口长度、立体声模式和比特率拆分比来确定的。在完全解码比特流之前,必须知道分配给每个声道的比特数。

[0238] 在智能间隙填充(IGF)块中,在某一范围的频谱(称为目标区块(tile))中被量化为零的谱线(line)填充有来自不同频谱范围(称为源区块)的处理内容。由于逐频带立体声处理,立体声表示(即L/R或M/S)对于源区块和目标区块来说可以不同。为了确保良好的质量,如果源区块的表示与目标区块的表示不同,则在解码器中间隙填充之前,对源区块进行处理以将其变换为目标区块的表示。[9]中已经描绘了该过程。与[6a]和[6b]相反,IGF本身应用于白化频谱域而不是原始频谱域。与已知的立体声编解码器(例如[9])相反,IGF应用于白化的ILD补偿频谱域。

[0239] 基于立体声模式和逐频带M/S决策,根据联合编码声道来构建左和右声道::

$$MDCT_{L,k} = \frac{1}{\sqrt{2}}(MDCT_{LM,k} + MDCT_{RS,k}) \text{ 和 } MDCT_{R,k} = \frac{1}{\sqrt{2}}(MDCT_{LM,k} - MDCT_{RS,k})。$$

[0240] 如果 $ratio_{ILD} > 1$,则右声道以 $ratio_{ILD}$ 缩放,否则左声道以 $\frac{1}{ratio_{ILD}}$ 缩放。

[0241] 对于可能发生除以0的每种情况,向分母添加小的正数。

[0242] 对于中间比特率(例如,48kbps),基于MDCT的编码可以很粗略地对频谱进行量化,以匹配比特消耗目标。这提出了对参数编码的需求,参数编码与相同频谱区域中的离散编码相结合、在帧到帧的基础上进行适配,从而提高了保真度。

[0243] 在下文中,描述了采用立体声填充的那些实施例中的一些实施例的方面。应注意,对于上述实施例,不必采用立体声填充。因此,仅上述实施例中的一些实施例采用立体声填

充。上述实施例的其它实施例根本不采用立体声填充。

[0244] MPEG-H频域立体声中的立体声频率填充例如在[11]中被描述。在[11]中,通过以缩放因子形式从编码器发送的频带能量(例如,在AAC中)来实现针对每个频带的目标能量。如果应用频域噪声(FDNS)整形并且通过使用LSF(线谱频率)对频谱包络进行编码(参见[6a]、[6b]、[8]),则无法如[11]中所述的立体声填充算法所要求的那样仅针对一些频带(频谱带)改变缩放。

[0245] 首先提供一些背景信息。

[0246] 当采用中/侧编码时,可以以不同方式来编码侧信号。

[0247] 根据第一组实施例,以与中央信号M相同的方式来编码侧信号S。执行量化,但不执行进一步的步骤以降低必要的比特率。通常,这种方法旨在允许在解码器侧非常精确地重新构建侧信号S,但另一方面需要大量的比特用于编码。

[0248] 根据第二组实施例,基于M信号根据原始侧信号S来产生残差侧信号S。在实施例中,可以例如根据以下公式计算残差侧信号:

$$[0249] \quad S_{\text{res}} = S - g \cdot M。$$

[0250] 其它实施例可以例如采用针对残差侧信号的其它定义。

[0251] 残差信号 S_{res} 被量化并与参数g一起发送到解码器。通过量化残差信号 S_{res} 而不是原始侧信号S,通常,更多的频谱值被量化为0。也就是说,通常,与量化原始侧信号S相比,这节省了编码和发送所必须的比特量。

[0252] 在第二组实施例的这些实施例的一些中,针对完整频谱确定单个参数g,并且将单个参数g发送到解码器。在第二组实施例的其它实施例中,频率频谱的多个频带/频谱带中的每一个可以例如包括两个或更多个频谱值,并且针对每个频带/频谱带确定参数g,并且将参数g发送到解码器。

[0253] 图12示出了根据第一组实施例或第二组实施例的编码器侧的不采用立体声填充的立体声处理。

[0254] 图13示出了根据第一组实施例或第二组实施例的解码器侧的不采用立体声填充的立体声处理。

[0255] 根据第三组实施例,采用立体声填充。在这些实施例的一些实施例中,在解码器侧,针对某一时间点t的侧信号S是根据紧接在前的时间点t-1的中央信号来产生的。

[0256] 例如,针对某一时间点t的侧信号S根据紧接在前的时间点t-1的中央信号来产生可以根据以下公式来执行:

$$[0257] \quad S(t) = h_b \cdot M(t-1)。$$

[0258] 在编码器侧,针对频谱的多个频带的每个频带确定参数 h_b 。在确定参数 h_b 之后,编码器向解码器发送参数 h_b 。在一些实施例中,侧信号S本身或其残差的频谱值不被发送到解码器。这种方法旨在节省所需比特数。

[0259] 在第三组实施例的一些其它实施例中,至少对于侧信号比中央信号更大声的那些频带,那些频带的侧信号的频谱值被明确地编码并被发送到解码器。

[0260] 根据第四组实施例,通过明确地编码原始侧信号S(参见第一组实施例)或残差侧信号 S_{res} 来编码侧信号S的一些频带,而对于其它频带,采用立体声填充。这种方法将第一组实施例或第二组实施例与采用立体声填充的第三组实施例组合。例如,可以例如通过量化

原始侧信号S或残差侧信号 S_{res} 来编码较低频带,而对于其它较高频带,可以例如采用立体声填充。

[0261] 图9示出了根据第三组实施例或第四组实施例的编码器侧的采用立体声填充的立体声处理。

[0262] 图10示出了根据第三组实施例或第四组实施例的解码器侧的采用立体声填充的立体声处理。

[0263] 上述实施例中的不采用立体声填充的那些实施例可以例如采用如MPEG-H中所述的立体声填充(参见MPEG-H频域立体声(参见,例如[11]))。

[0264] 采用立体声填充的一些实施例可以例如将[11]中描述的立体声填充算法应用于其中频谱包络被编码为LSF与噪声填充相组合的系统。对频谱包络进行编码可以例如如[6a]、[6b]、[8]中所描述的那样来实现。噪声填充可以例如如[6a]和[6b]中所述的那样来实现。

[0265] 在一些特定实施例中,可以例如在频域内的M/S频带中(例如,从诸如 $0.08F_s$ (F_s = 采样频率)之类的较低频率至诸如IGF交叉频率之类的较高频率)执行包括立体声填充参数计算的立体声填充处理。

[0266] 例如,对于低于较低频率(例如, $0.08F_s$)的频率部分,原始侧信号S或根据原始侧信号S导出的残差侧信号可以例如被量化并被发送到解码器。对于大于较高频率(例如,IGF交叉频率)的频率部分,可以例如执行智能间隙填充(IGF)。

[0267] 更具体地,在一些实施例中,对于立体声填充范围内的、完全被量化为0的那些频带(例如,采样频率的0.08倍直到IGF交叉频率),例如可以使用来自先前帧的白化MDCT频谱缩混(IGF=智能间隙填充)的“复制”来填充侧信道(第二信道)。例如,“复制”可以与噪声填充互补地应用,并且相应地根据从编码器发送的校正因子进行缩放。在其它实施例中,较低频率可以呈现为除 $0.08F_s$ 之外的其它值。

[0268] 在一些实施例中,替代 $0.08F_s$,较低频率可以是例如0至 $0.50F_s$ 范围内的值。具体地,在实施例中,较低频率可以是 $0.01F_s$ 至 $0.50F_s$ 范围内的值。例如,较低频率可以是例如 $0.12F_s$ 或 $0.20F_s$ 或 $0.25F_s$ 。

[0269] 在其它实施例中,除了采用智能间隙填充之外或替代采用智能间隙填充,对于大于较高频率的频率,可以例如执行噪声填充。

[0270] 在其它实施例中,没有较高频率,并且对于大于较低频率的每个频率部分执行立体声填充。

[0271] 在其它实施例中,没有较低频率,并且对于从最低频带到较高频率的频率部分执行立体声填充。

[0272] 在其它实施例中,没有较低频率且没有较高频率,并且对整个频率谱执行立体声填充。

[0273] 在下文中,描述了采用立体声填充的特定实施例。

[0274] 特别地,描述了根据特定实施例的具有校正因子的立体声填充。在图9(编码器侧)和图10(解码器侧)的立体声填充处理块的实施例中,可以采用具有校正因子的立体声填充。

[0275] 在下文中,

- [0276] $-Dmx_R$ 可以例如表示白化的MDCT频谱的中央信号,
- [0277] $-S_R$ 可以例如表示白化的MDCT频谱的侧信号,
- [0278] $-Dmx_I$ 可以例如表示白化的MDCT频谱的中央信号,
- [0279] $-S_I$ 可以表示白化的MDST频谱的侧信号,
- [0280] $-prevDmx_R$ 可以例如表示延迟一帧的白化的MDCT频谱的中央信号,以及
- [0281] $-prevDmx_I$ 可以例如表示延迟一帧的白化的MDST频谱的中央信号。
- [0282] 当立体声决策是针对所有频带的M/S(全M/S)或针对所有立体声填充频带的M/S(逐频带M/S)时,可以应用立体声填充编码。
- [0283] 当确定应用全-双-单声道处理时,绕过立体声填充。此外,当针对某些频谱带(频带)选择L/R编码时,针对这些频谱带也绕过立体声填充。
- [0284] 现在,考虑采用立体声填充的特定实施例。在这样的特定实施例中,块内的处理可以例如如下执行:
- [0285] 对于落在从较低频率(例如, $0.08F_s$ (F_s = 采样频率))开始到较高频率(例如,IGF交叉频率)的频率区域内的频带(fb):
- [0286] -例如,根据以下公式来计算侧信号 S_R 的残差 Res_R :
- [0287] $Res_R = S_R - a_R Dmx_R - a_I Dmx_I$.
- [0288] 其中, a_R 是复数预测系数的实部, a_I 是复数预测系数的虚部(参见[10])。
- [0289] 根据以下公式来计算侧信号 S_I 的残差 Res_I :
- [0290] $Res_I = S_I - a_R Dmx_R - a_I Dmx_I$.
- [0291] -计算残差 Res 的以及先前帧缩混(中央信号) $prevDmx$ 的能量(例如,复值能量):
- [0292]
$$ERes_{fb} = \sum_{fb} Res_R^2 + \sum_{fb} Res_I^2,$$
- [0293]
$$EprevDmx_{fb} = \sum_{fb} prevDmx_R^2 + \sum_{fb} prevDmx_I^2$$
- [0294] 在以上公式中:
- [0295] $\sum_{fb} Res_R^2$ Res_R 的频带fb内的所有频谱值的平方之和。
- [0296] $\sum_{fb} Res_I^2$ Res_I 的频带fb内的所有频谱值的平方之和。
- [0297] $\sum_{fb} prevDmx_R^2$ $prevDmx_R$ 的频带fb内的所有频谱值的平方之和。
- [0298] $\sum_{fb} prevDmx_I^2$ $prevDmx_I$ 的频带fb内的所有频谱值的平方之和。
- [0299] -根据这些计算的能量($ERes_{fb}$ 、 $EprevDmx_{fb}$),计算立体声填充校正因子,并且将其作为侧信息发送到解码器:
- [0300] $correction_factor_{fb} = ERes_{fb} / (EprevDmx_{fb} + \epsilon)$
- [0301] 在实施例中, $\epsilon = 0$ 。在其它实施例中,例如, $0.1 > \epsilon > 0$,例如以避免除以0。
- [0302] -可以例如根据例如针对采用立体声填充的每个频谱带计算的立体声填充校正因

子来计算逐频带缩放因子。为了补偿能量损失,引入按照缩放因子将输出中央和侧(残差)信号进行逐频带缩放,因为没有用于根据解码器侧的残差重新构建侧信号的逆复数预测操作($a_R = a_I = 0$)。

[0303] 在特定实施例中,可以例如根据以下公式计算逐频带缩放因子:

$$[0304] \quad \text{scaling_factor}_{fb} = \sqrt{\frac{\sum_{fb} (S_R - a_R Dmx_R)^2 + \sum_{fb} (S_I - a_I Dmx_I)^2 + EDmx_{fb}}{ERes_{fb} + EDmx_{fb} + \varepsilon}}$$

[0305] 其中, $EDmx_{fb}$ 是当前帧缩混的(例如复数)能量(其可以例如如上所述地计算)。

[0306] 在一些实施例中,在立体声处理块中的立体声填充处理之后且在量化之前,如果对于等同频带,缩混(中央)比残差(侧)大声,则可以例如将落入立体声填充频率范围内的残差的仓(bin)设置为0:

$$[0307] \quad \frac{E_{fb}^M}{E_{fb}^S} > \text{阈值}$$

$$[0308] \quad E_{fb}^M = \sum_{fb} Dmx_R^2$$

$$[0309] \quad E_{fb}^S = \sum_{fb} Res_R^2$$

[0310] 因此,在编码缩混和残差的较低频率仓时花费更多比特,从而提高了整体质量。

[0311] 在备选实施例中,可以例如将残差(侧)的所有比特设置为0。这样的备选实施例可以例如基于缩混在大多数情况下比残差更大声的假设。

[0312] 图11示出了解码器侧的根据特定实施例的侧信号的立体声填充。

[0313] 在解码、逆量化和噪声填充之后,对侧声道应用立体声填充。对于立体声填充范围内的、被量化为0的频带,如果噪声填充后的频带能量不能达到目标能量,则可以例如应用来自最后帧的白化MDCT频谱缩混的“复制”(如图11所示)。例如,根据以下公式,根据作为参数从编码器发送的立体声校正因子来计算每个频带的目标能量。

$$[0314] \quad ET_{fb} = \text{correction_factor}_{fb} \cdot E_{prevDmx_{fb}}$$

[0315] 例如根据以下公式实现在解码器侧产生侧信号(例如,可以称为先前缩混“复制”):

$$[0316] \quad S_i = N_i + \text{facDmx}_{fb} \cdot \text{prevDmx}_i, i \in [fb, fb+1],$$

[0317] 其中*i*表示频带*fb*内的频率仓(频谱值),*N*是噪声填充频谱,并且 facDmx_{fb} 是应用于先前缩混的因子,其取决于从编码器发送的立体声填充校正因子。

[0318] 在特定实施例中,例如,可以针对每个频率带*fb*将 facDmx_{fb} 计算为:

$$[0319] \quad facDmx_{fb} = \sqrt{\frac{correction_factor_{fb} - EN_{fb}}{(EprevDmx_{fb} + \epsilon)}}$$

[0320] 其中, EN_{fb} 是频带 fb 中的噪声填充频谱的能量, 并且 $EprevDmx_{fb}$ 是相应先前帧缩混能量。

[0321] 在编码器侧, 备选实施例不考虑 MDST 频谱 (或 MDCT 频谱)。在那些实施例中, 如下地适配编码器侧的进程:

[0322] 对于落在从较低频率 (例如, $0.08F_s$ ($F_s =$ 采样频率)) 开始到较高频率 (例如, IGF 交叉频率) 的频率区域内的频带 (fb):

[0323] - 例如, 根据以下公式来计算侧信号 S_R 的残差 Res :

$$[0324] \quad Res = S_R - a_R Dmx_R,$$

[0325] 其中, a_R 是 (例如, 实数的) 预测系数。

[0326] - 计算残差 Res 的以及先前帧缩混 (中央信号) $prevDmx$ 的能量:

$$[0327] \quad ERes_{fb} = \sum_{fb} Res_R^2,$$

$$[0328] \quad EprevDmx_{fb} = \sum_{fb} prevDmx_R^2$$

[0329] - 根据这些计算的能量 ($ERes_{fb}$ 、 $EprevDmx_{fb}$), 计算立体声填充校正因子, 并且将其作为侧信息发送到解码器:

$$[0330] \quad correction_factor_{fb} = ERes_{fb} / (EprevDmx_{fb} + \epsilon)$$

[0331] 在实施例中, $\epsilon = 0$ 。在其它实施例中, 例如, $0.1 > \epsilon > 0$, 例如以避免除以 0。

[0332] - 可以例如根据例如针对采用立体声填充的每个频谱带计算的立体声填充校正因子来计算逐频带缩放因子。

[0333] 在特定实施例中, 可以例如根据以下公式计算逐频带缩放因子:

$$[0334] \quad scaling_factor_{fb} = \sqrt{\frac{\sum_{fb} (S_R - a_R Dmx_R)^2 + EDmx_{fb}}{ERes_{fb} + EDmx_{fb} + \epsilon}}$$

[0335] 其中, $EDmx_{fb}$ 是当前帧缩混的能量 (其可以例如如上所述地计算)。

[0336] - 在一些实施例中, 在立体声处理块中的立体声填充处理之后且在量化之前, 如果对于等同频带, 缩混 (中央) 比残差 (侧) 大声, 则可以例如将落入立体声填充频率范围内的残差的仓 (bin) 设置为 0:

$$[0337] \quad \frac{E_{fb}^M}{E_{fb}^S} > \text{阈值}$$

$$[0338] \quad E_{fb}^M = \sum_{fb} Dmx_R^2$$

$$[0339] \quad E_{fb}^S = \sum_{fb} Res_R^2$$

[0340] 因此,在编码缩混和残差的较低频率仓时花费更多比特,从而提高了整体质量。

[0341] 在备选实施例中,可以例如将残差(侧)的所有比特设置为0。这样的备选实施例可以例如基于缩混在大多数情况下比残差更大声的假设。

[0342] 根据一些实施例,可以例如提供用于在具有FDNS的系统中应用立体声填充的装置,其中使用LSF(或者不可能在单个频带中独立地改变缩放的类似编码)对频谱包络进行编码。

[0343] 根据一些实施例,可以例如提供用于在没有复数/实数预测的系统中应用立体声填充的装置。

[0344] 在从编码器向解码器发送明确参数(立体声填充校正因子)的意义上,一些实施例可以例如采用参数立体声填充,以控制白化的左和右MDCT频谱的立体声填充(例如,利用先前帧的缩混)。

[0345] 更一般地:

[0346] 在一些实施例中,图1a至图1e的编码单元120可以例如被配置为产生处理后的音频信号,使得处理后的音频信号的第一声道的所述至少一个频谱带是所述中央信号的所述频谱带,并且使得处理后的音频信号的第二声道的所述至少一个频谱带是所述侧信号的所述频谱带。为了获得编码音频信号,编码单元120可以例如被配置为通过确定所述侧信号的所述频谱带的校正因子来编码所述侧信号的所述频谱带。编码单元120可以例如被配置为根据残差并且根据与所述中央信号的所述频谱带相对应的先前中央信号的频谱带,确定所述侧信号的所述频谱带的所述校正因子,其中在时间上先前中央信号在所述中央信号之前。此外,编码单元120可以例如被配置为根据所述侧信号的所述频谱带、并且根据所述中央信号的所述频谱带来确定残差。

[0347] 根据一些实施例,编码单元120可以例如被配置为根据以下公式确定所述侧信号的所述频谱带的所述校正因子。

$$[0348] \quad \text{correction_factor}_{fb} = ERes_{fb} / (EprevDmx_{fb} + \epsilon)$$

[0349] 其中,correction_factor_{fb}指示所述侧信号的所述频谱带的所述校正因子,其中ERes_{fb}指示根据与所述中央信号的所述频谱带相对应的所述残差的频谱带的能量的残差能量,其中EprevDmx_{fb}指示根据先前中央信号的频谱带中能量的先前能量,并且其中 $\epsilon=0$,或者其中 $0.1 > \epsilon > 0$ 。

[0350] 在一些实施例中,可以根据以下公式来定义所述残差:

$$[0351] \quad Res_R = S_R - a_R Dmx_R,$$

[0352] 其中,Res_R是所述残差,其中S_R是所述侧信号,其中a_R是(例如,实数)系数(例如,预测系数),其中Dmx_R是所述中央信号,其中编码单元(120)被配置为根据以下公式来确定所述残差能量:

$$[0353] \quad ERes_{fb} = \sum_{fb} Res_R^2$$

[0354] 根据一些实施例,根据以下公式来定义所述残差:

$$[0355] \quad Res_R = S_R - a_R Dmx_R - a_I Dmx_I,$$

[0356] 其中, Res_R 是所述残差,其中 S_R 是所述侧信号,其中 a_R 是复数(预测)系数的实部,并且其中 a_I 是所述复数(预测)系数的虚部,其中 Dmx_R 是所述中央信号,其中 Dmx_I 是根据归一化音频信号的第一声道和根据归一化音频信号的第二声道的另一中央信号,其中根据以下公式定义根据归一化音频信号的第一声道和根据归一化音频信号的第二声道的另一侧信号 S_I 的另一残差:

$$[0357] \quad ReS_I = S_I - a_R Dmx_R - a_I Dmx_I,$$

[0358] 其中,编码单元120可以例如被配置为根据以下公式来确定所述残差能量:

$$[0359] \quad ERes_{fb} = \sum_{fb} Res_R^2 + \sum_{fb} Res_I^2$$

[0360] 其中编码单元120可以例如被配置为根据与所述中央信号的所述频谱带相对应的所述残差的频谱带的能量、以及根据与所述中央信号的所述频谱带相对应的所述另一残差的频谱带的能量,来确定先前的能量。

[0361] 在一些实施例中,图2a至图2e的解码单元210可以例如被配置为针对所述多个频谱带的每个频谱带,来确定编码音频信号的第一声道的所述频谱带和编码音频信号的第二声道的所述频谱带是使用双-单声道编码来编码的还是使用中-侧编码来编码的。此外,解码单元210可以例如被配置为通过重新构建第二声道的所述频谱带来获得编码音频信号的第二声道的所述频谱带。如果使用中-侧编码,则编码音频信号的第一声道的所述频谱带是中央信号的频谱带,并且编码音频信号的第二声道的所述频谱带是侧信号的频谱带。此外,如果使用中-侧编码,则解码单元210可以例如被配置为根据侧信号的所述频谱带的校正因子、并且根据与所述中央信号的所述频谱带相对应的先前中央信号的频谱带,来重新构建侧信号的所述频谱带,其中在时间上先前中央信号在所述中央信号之前。

[0362] 根据一些实施例,如果使用中-侧编码,则解码单元210可以例如被配置为通过根据以下公式重新构建侧信号的所述频谱带的频谱值来重新构建侧信号的所述频谱带。

$$[0363] \quad S_i = N_i + facDmx_{fb} \cdot prevDmx_i$$

[0364] 其中, S_i 指示侧信号的所述频谱带的频谱值,其中 $prevDmx_i$ 指示所述先前中央信号的频谱带的频谱值,其中 N_i 指示噪声填充频谱的频谱值,其中根据以下公式来定义 $facDmx_{fb}$:

$$[0365] \quad facDmx_{fb} = \sqrt{correction_factor_{fb} - \frac{EN_{fb}}{(EprevDmx_{fb} + \epsilon)}}$$

[0366] 其中, $correction_factor_{fb}$ 是所述侧信号的所述频谱带的校正因子,其中, EN_{fb} 是噪声填充频谱的能量,其中 $EprevDmx_{fb}$ 是所述前述中央信号的所述频谱带的能量,并且其中 $\epsilon = 0$, 或其中 $0.1 > \epsilon > 0$ 。

[0367] 在一些实施例中,残差可以例如根据编码器处的复数立体声预测算法导出,而在解码器侧不存在立体声预测(实数或复数)。

[0368] 根据一些实施例,编码器侧处的对频谱进行能量校正缩放可以例如用于补偿解码器侧没有逆预测处理的事实。

[0369] 尽管已经在装置的上下文下描述了一些方面,但是将清楚的是,这些方面还表示对应方法的描述,其中,块或设备与方法步骤或方法步骤的特征相对应。类似地,在方法步骤的上下文下描述的方面也表示对对应块或者对应装置的项或特征的描述。可以由(或使用)硬件装置(诸如,微处理器、可编程计算机或电子电路)来执行一些或全部方法步骤。在一些实施例中,可以由这种装置来执行最重要方法步骤中的一个或多个方法步骤。

[0370] 根据某些实现要求,本发明的实施例可以用硬件或软件实现,或者至少部分地用硬件、或至少部分地用软件实现。可以使用其上存储有电子可读控制信号的数字存储介质(例如,软盘、DVD、蓝光、CD、ROM、PROM、EPROM、EEPROM或闪存)来执行实现,该电子可读控制信号与可编程计算机系统协作(或者能够与之协作)从而执行相应方法。因此,数字存储介质可以是计算机可读的。

[0371] 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,该电子可读控制信号能够与可编程计算机系统协作从而执行本文所述的方法之一。

[0372] 通常,本发明的实施例可以被实现为具有程序代码的计算机程序产品,程序代码可操作以在计算机程序产品在计算机上运行时执行方法之。程序代码可以例如存储在机器可读载体上。

[0373] 其它实施例包括存储在机器可读载体上的计算机程序,该计算机程序用于执行本文所述的方法之一。

[0374] 换言之,本发明方法的实施例因此是具有程序代码的计算机程序,该程序代码用于在计算机程序在计算机上运行时执行本文所述的方法之一。

[0375] 因此,本发明方法的另一实施例是其上记录有计算机程序的数据载体(或者数字存储介质或计算机可读介质),该计算机程序用于执行本文所述的方法之一。数据载体、数字存储介质或记录的介质通常是有形的和/或非暂时性的。

[0376] 因此,本发明方法的另一实施例是表示计算机程序的数据流或信号序列,所述计算机程序用于执行本文所述的方法之一。数据流或信号序列可以例如被配置为经由数据通信连接(例如,经由互联网)传送。

[0377] 另一实施例包括处理装置,例如,计算机或可编程逻辑器件,所述处理装置被配置为或适于执行本文所述的方法之一。

[0378] 另一实施例包括其上安装有计算机程序的计算机,该计算机程序用于执行本文所述的方法之一。

[0379] 根据本发明的另一实施例包括被配置为向接收机(例如,以电子方式或以光学方式)传送计算机程序的装置或系统,该计算机程序用于执行本文所述的方法之一。接收机可以是例如计算机、移动设备、存储设备等。装置或系统可以例如包括用于向接收机传送计算机程序的文件服务器。

[0380] 在一些实施例中,可编程逻辑器件(例如,现场可编程门阵列)可以用于执行本文所述的方法的功能中的一些或全部。在一些实施例中,现场可编程门阵列可以与微处理器

协作以执行本文所述的方法之一。通常,方法优选地由任意硬件装置来执行。

[0381] 本文描述的装置可以使用硬件装置、或者使用计算机、或者使用硬件装置和计算机的组合来实现。

[0382] 本文描述的方法可以使用硬件装置、或者使用计算机、或者使用硬件装置和计算机的组合来执行。

[0383] 上述实施例对于本发明的原理仅是说明性的。应当理解的是:本文所述的布置和细节的修改和变形对于本领域其他技术人员将是显而易见的。因此,旨在仅由所附专利权利要求的范围来限制而不是由借助对本文实施例的描述和解释所给出的具体细节来限制。

[0384] 示例实施例1,一种用于对包括两个或更多个声道的音频输入信号的第一声道和第二声道进行编码以获得编码音频信号的装置,其中,所述装置包括:

[0385] 归一化器(110),所述归一化器(110)被配置为根据所述音频输入信号的第一声道并且根据所述音频输入信号的第二声道来确定所述音频输入信号的归一化值,其中所述归一化器(110)被配置为通过根据所述归一化值修正所述音频输入信号的第一声道和第二声道中的至少一个声道,来确定归一化音频信号的第一声道和第二声道;

[0386] 编码单元(120),所述编码单元(120)被配置为产生具有第一声道和第二声道的处理后的音频信号,使得所述处理后的音频信号的第一声道的一个或多个频谱带是所述归一化音频信号的第一声道的一个或多个频谱带,使得所述处理后的音频信号的第二声道的一个或多个频谱带是所述归一化音频信号的第二声道的一个或多个频谱带,使得所述处理后的音频信号的第一声道的至少一个频谱带是根据所述归一化音频信号的第一声道的频谱带并且根据所述归一化音频信号的第二声道的频谱带的中央信号的频谱带,以及使得所述处理后的音频信号的第二声道的至少一个频谱带是根据所述归一化音频信号的第一声道的频谱带并且根据所述归一化音频信号的第二声道的频谱带的侧信号的频谱带,其中,所述编码单元(120)被配置为对所述处理后的音频信号进行编码以获得所述编码音频信号。

[0387] 示例实施例2,根据示例实施例1所述的装置,

[0388] 其中,所述编码单元(120)被配置为根据所述归一化音频信号的第一声道的多个频谱带并且根据所述归一化音频信号的第二声道的多个频谱带,在全-中-侧编码模式、全-双-单声道编码模式和逐频带编码模式之间选择,

[0389] 其中,所述编码单元(120)被配置为:如果选择所述全-中-侧编码模式,则根据所述归一化音频信号的第一声道并且根据所述归一化音频信号的第二声道产生中央信号作为中-侧信号的第一声道,根据所述归一化音频信号的第一声道和根据所述归一化音频信号的第二声道产生侧信号作为所述中-侧信号的第二声道,以及对所述中-侧信号进行编码以获得所述编码音频信号,

[0390] 其中,所述编码单元(120)被配置为:如果选择所述全-双-单声道编码模式,则对所述归一化音频信号进行编码以获得所述编码音频信号,以及

[0391] 其中,所述编码单元(120)被配置为:如果选择所述逐频带编码模式,则产生所述处理后的音频信号,使得所述处理后的音频信号的第一声道的一个或多个频谱带是所述归一化音频信号的第一声道的一个或多个频谱带,使得所述处理后的音频信号的第二声道的一个或多个频谱带是所述归一化音频信号的第二声道的一个或多个频谱带,使得所述处理后的音频信号的第一声道的至少一个频谱带是根据所述归一化音频信号的第一声道的频

谱带并且根据所述归一化音频信号的第二声道的频谱带的中央信号的频谱带,以及使得所述处理后的音频信号的第二声道的至少一个频谱带是根据所述归一化音频信号的第一声道的频谱带并且根据所述归一化音频信号的第二声道的频谱带的侧信号的频谱带,其中所述编码单元(120)被配置为对所述处理后的音频信号进行编码以获得所述编码音频信号。

[0392] 示例实施例3,根据示例实施例2所述的装置,

[0393] 其中,所述编码单元(120)被配置为:如果选择所述逐频带编码模式,则针对所述处理后的音频信号的多个频谱带中的每个频谱带,决定是采用中-侧编码还是采用双-单声道编码,

[0394] 其中,如果针对所述频谱带采用所述中-侧编码,则所述编码单元(120)被配置为:基于所述归一化音频信号的第一声道的所述频谱带并且基于所述归一化音频信号的第二声道的所述频谱带,产生所述处理后的音频信号的第一声道的所述频谱带作为中央信号的频谱带,并且所述编码单元(120)被配置为:基于所述归一化音频信号的第一声道的所述频谱带并且基于所述归一化音频信号的第二声道的所述频谱带,产生所述处理后的音频信号的第二声道的所述频谱带作为侧信号的频谱带,以及

[0395] 其中,如果针对所述频谱带采用所述双-单声道编码,则

[0396] 所述编码单元(120)被配置为:使用所述归一化音频信号的第一声道的所述频谱带作为所述处理后的音频信号的第一声道的所述频谱带,并且被配置为使用所述归一化音频信号的第二声道的所述频谱带作为所述处理后的音频信号的第二声道的所述频谱带,或者

[0397] 所述编码单元(120)被配置为:使用所述归一化音频信号的第二声道的所述频谱带作为所述处理后的音频信号的第一声道的所述频谱带,并且被配置为使用所述归一化音频信号的第一声道的所述频谱带作为所述处理后的音频信号的第二声道的所述频谱带。

[0398] 示例实施例4,根据示例实施例2或3所述的装置,其中,所述编码单元(120)被配置为:通过确定估计在采用所述全-中-侧编码模式时编码所需的第一比特数的第一估计,通过确定估计在采用所述全-双-单声道编码模式时编码所需的第二比特数的第二估计,通过确定估计在采用所述逐频带编码模式时编码所需的第三比特数的第三估计,以及通过在所述全-中-侧编码模式、所述全-双-单声道编码模式和所述逐频带编码模式之中选择具有所述第一估计、所述第二估计和所述第三估计之中的最小比特数的编码模式,在所述全-中-侧编码模式、所述全-双-单声道编码模式和所述逐频带编码模式之间进行选择。

[0399] 示例实施例5,根据示例实施例4所述的装置,

[0400] 其中,所述编码单元(120)被配置为根据以下公式估计所述第三估计 b_{BW} ,所述第三估计估计在采用所述逐频带编码模式时编码所需的第三比特数:

$$[0401] \quad b_{BW} = nBands + \sum_{i=0}^{nBands-1} \min(b_{bwLR}^i, b_{bwMS}^i)$$

[0402] 其中, $nBands$ 是所述归一化音频信号的频谱带的个数,

[0403] 其中, b_{bwMS}^i 是用于编码所述中央信号的第*i*个频谱带和用于编码所述侧信号的第*i*个频谱带所需的比特数的估计,以及

[0404] 其中, b_{bwLR}^i 是用于编码第一信号的第i个频谱带和用于编码第二信号的第i个频谱带所需的比特数的估计。

[0405] 示例实施例6,根据示例实施例2或3所述的装置,其中,所述编码单元(120)被配置为:通过确定估计在以所述全-中-侧编码模式进行编码时所保存的第一比特数的第一估计,通过确定估计在以所述全-双-单声道编码模式进行编码时所保存的第二比特数的第二估计,通过确定估计在以所述逐频带编码模式进行编码时所保存的第三比特数的第三估计,以及通过在所述全-中-侧编码模式、所述全-双-单声道编码模式和所述逐频带编码模式之中选择具有所述第一估计、所述第二估计和所述第三估计之中的所保存的最大比特数的编码模式,在所述全-中-侧编码模式、所述全-双-单声道编码模式和所述逐频带编码模式之间进行选择。

[0406] 示例实施例7,根据示例实施例2或3所述的装置,其中,所述编码单元(120)被配置为:通过估计在采用所述全-中-侧编码模式时发生的第一信噪比,通过估计在采用所述全-双-单声道编码模式时发生的第二信噪比,通过估计在采用所述逐频带编码模式时发生的第三信噪比,并且通过在所述全-中-侧编码模式、所述全-双-单声道编码模式和所述逐频带编码模式之中选择具有所述第一信噪比、所述第二信噪比和所述第三信噪比之中的最大信噪比的编码模式,在所述全-中-侧编码模式、所述全-双-单声道编码模式和所述逐频带编码模式之间进行选择。

[0407] 示例实施例8,根据示例实施例1所述的装置,

[0408] 其中,所述编码单元(120)被配置为:产生所述处理后的音频信号,使得所述处理后的音频信号的第一声道的所述至少一个频谱带是所述中央信号的所述频谱带,并且使得所述处理后的音频信号的第二声道的所述至少一个频谱带是所述侧信号的所述频谱带,

[0409] 其中,为了获得所述编码音频信号,所述编码单元(120)被配置为通过确定所述侧信号的所述频谱带的校正因子来编码所述侧信号的所述频谱带,

[0410] 其中,所述编码单元(120)被配置为根据残差并且根据与所述中央信号的所述频谱带相对应的先前中央信号的频谱带,确定所述侧信号的所述频谱带的所述校正因子,其中所述先前中央信号在时间上在所述中央信号之前,

[0411] 其中,所述编码单元(120)被配置为根据所述侧信号的所述频谱带、并且根据所述中央信号的所述频谱带来确定所述残差。

[0412] 示例实施例9,根据示例实施例8所述的装置,

[0413] 其中,所述编码单元(120)被配置为根据以下公式确定所述侧信号的所述频谱带的所述校正因子:

$$[0414] \text{correction_factor}_{fb} = ERes_{fb} / (EprevDmx_{fb} + \epsilon)$$

[0415] 其中, $\text{correction_factor}_{fb}$ 指示所述侧信号的所述频谱带的所述校正因子,

[0416] 其中, $ERes_{fb}$ 指示根据与所述中央信号的所述频谱带相对应的所述残差的频谱带的能量的残差能量,

[0417] 其中, $EprevDmx_{fb}$ 指示根据先前中央信号的频谱带的能量的先前能量,以及

[0418] 其中, $\epsilon = 0$, 或其中 $0.1 > \epsilon > 0$ 。

[0419] 示例实施例10,根据示例实施例8或9所述的装置,

[0420] 其中,根据以下公式定义所述残差:

[0421] $Res_R = S_R - a_R Dmx_R$,

[0422] 其中, Res_R 是所述残差, 其中 S_R 是所述侧信号, 其中 a_R 是系数, 其中 Dmx_R 是所述中央信号,

[0423] 其中, 所述编码单元 (120) 被配置为根据以下公式来确定所述残差能量:

$$[0424] \quad ERes_{fb} = \sum_{fb} Res_R^2$$

[0425] 示例实施例11, 根据示例实施例8或9所述的装置,

[0426] 其中, 根据以下公式定义所述残差:

[0427] $Res_R = S_R - a_R Dmx_R - a_I Dmx_I$,

[0428] 其中, Res_R 是所述残差, 其中 S_R 是所述侧信号, 其中 a_R 是复数系数的实部, 并且其中 a_I 是所述复数系数的虚部, 其中 Dmx_R 是所述中央信号, 其中 Dmx_I 是根据所述归一化音频信号的第一声道并且根据所述归一化音频信号的第二声道的另一中央信号,

[0429] 其中, 根据以下公式定义根据所述归一化音频信号的第一声道并且根据所述归一化音频信号的第二声道的另一侧信号 S_i 的另一残差:

[0430] $Res_i = S_i - a_R Dmx_R - a_I Dmx_i$,

[0431] 其中, 所述编码单元 (120) 被配置为根据以下公式来确定所述残差能量:

$$[0432] \quad ERes_{fb} = \sum_{fb} Res_R^2 + \sum_{fb} Res_i^2$$

[0433] 其中, 所述编码单元 (120) 被配置为根据与所述中央信号的所述频谱带相对应的所述残差的频谱带的能量、以及根据与所述中央信号的所述频谱带相对应的所述另一残差的频谱带的能量, 确定先前能量。

[0434] 示例实施例12, 根据前述示例实施例中的任一项所述的装置,

[0435] 其中, 所述归一化器 (110) 被配置为根据所述音频输入信号的第一声道的能量并且根据所述音频输入信号的第二声道的能量来确定所述音频输入信号的归一化值。

[0436] 示例实施例13, 根据前述示例实施例中的任一项所述的装置,

[0437] 其中, 所述音频输入信号在频谱域中表示,

[0438] 其中, 所述归一化器 (110) 被配置为根据所述音频输入信号的第一声道的多个频谱带并且根据所述音频输入信号的第二声道的多个频谱带来确定所述音频输入信号的归一化值, 以及

[0439] 其中, 所述归一化器 (110) 被配置为通过根据所述归一化值修正所述音频输入信号的第一声道和第二声道中的至少一个声道的多个频谱带来确定所述归一化音频信号。

[0440] 示例实施例14, 根据示例实施例13所述的装置,

[0441] 其中, 所述归一化器 (110) 被配置为基于以下公式确定所述归一化值:

$$[0442] \quad NRG_L = \sqrt{\sum MDCT_{L,k}^2}$$

$$[0443] \quad NRG_R = \sqrt{\sum MDCT_{R,k}^2}$$

$$[0444] \quad ILD = \frac{NRG_L}{NRG_L + NRG_R}$$

[0445] 其中, $MDCT_{L,k}$ 是所述音频输入信号的第一声道的MDCT频谱的第k个系数, 并且 $MDCT_{R,k}$ 是所述音频输入信号的第二声道的MDCT频谱的第k个系数, 以及

[0446] 其中, 所述归一化器 (110) 被配置为通过量化ILD来确定所述归一化值。

[0447] 示例实施例15, 根据示例实施例13或14所述的装置,

[0448] 其中, 所述用于编码的装置还包括变换单元 (102) 和预处理单元 (105),

[0449] 其中, 所述变换单元 (102) 被配置为将时域音频信号从时域变换到频域以获得变换后的音频信号,

[0450] 其中, 所述预处理单元 (105) 被配置为通过对所述变换后的音频信号应用编码器侧频域噪声整形操作来产生所述音频输入信号的第一声道和第二声道。

[0451] 示例实施例16, 根据示例实施例15所述的装置,

[0452] 其中, 所述预处理单元 (105) 被配置为通过在对所述变换后的音频信号应用编码器侧频域噪声整形操作之前对所述变换后的音频信号应用编码器侧时间噪声整形操作, 来产生所述音频输入信号的第一声道和第二声道。

[0453] 示例实施例17, 根据示例实施例1至12中任一项所述的装置,

[0454] 其中, 所述归一化器 (110) 被配置为根据在时域中表示的所述音频输入信号的第一声道并且根据在时域中表示的所述音频输入信号的第二声道来确定所述音频输入信号的归一化值,

[0455] 其中, 所述归一化器 (110) 被配置为通过根据所述归一化值修正在时域中表示的所述音频输入信号的第一声道和第二声道中的至少一个声道来确定所述归一化音频信号的第一声道和第二声道,

[0456] 其中, 所述装置还包括变换单元 (115), 所述变换单元 (115) 被配置为将所述归一化音频信号从时域变换到频谱域, 使得所述归一化音频信号在频谱域中表示, 以及

[0457] 其中, 所述变换单元被配置为将在频谱域中表示的所述归一化音频信号馈送到所述编码单元 (120) 中。

[0458] 示例实施例18, 根据示例实施例17所述的装置,

[0459] 其中, 所述装置还包括被配置为接收包括第一声道和第二声道的时域音频信号的预处理单元 (106),

[0460] 其中, 所述预处理单元 (106) 被配置为对时域音频信号中的、产生第一感知白化频谱的第一声道应用滤波器, 以获得在时域中表示的音频输入信号的第一声道, 以及

[0461] 其中, 所述预处理单元 (106) 被配置为对时域音频信号中的、产生第二感知白化频谱的第二声道应用所述滤波器, 以获得在时域中表示的音频输入信号的第二声道。

[0462] 示例实施例19, 根据示例实施例17或18所述的装置,

[0463] 其中, 所述变换单元 (115) 被配置为将归一化音频信号从时域变换到频谱域以获得变换后的音频信号,

[0464] 其中, 所述装置还包括频谱域预处理器 (118), 所述频谱域预处理器 (118) 被配置为对所述变换后的音频信号执行编码器侧时间噪声整形, 以获得在频谱域中表示的归一化

音频信号。

[0465] 示例实施例20,根据前述示例实施例中的任一项所述的装置,

[0466] 其中,所述编码单元(120)被配置为通过对所述归一化音频信号或所述处理后的音频信号应用编码器侧立体声智能间隙填充来获得所述编码音频信号。

[0467] 示例实施例21,根据前述示例实施例中任一项所述的装置,其中,所述音频输入信号是恰好包括两个声道的音频立体声信号。

[0468] 示例实施例22,一种用于对包括四个或更多个声道的音频输入信号的四个声道进行编码以获得编码音频信号的系统,其中,所述系统包括:

[0469] 根据示例实施例1至20中任一项所述的第一装置(170),用于对所述音频输入信号的四个或更多个声道中的第一声道和第二声道进行编码,以获得所述编码音频信号的第一声道和第二声道,以及

[0470] 根据示例实施例1至20中任一项所述的第二装置(180),用于对所述音频输入信号的四个或更多个声道中的第三声道和第四声道进行编码,以获得所述编码音频信号的第三声道和第四声道。

[0471] 示例实施例23,一种用于对包括第一声道和第二声道的编码音频信号进行解码以获得包括两个或更多个声道的解码音频信号的第一声道和第二声道的装置,

[0472] 其中,所述装置包括解码单元(210),所述解码单元(210)被配置为针对多个频谱带中的每个频谱带,确定所述编码音频信号的第一声道的所述频谱带和所述编码音频信号的第二声道的所述频谱带是使用双-单声道编码来编码的还是使用中-侧编码来编码的,

[0473] 其中,如果使用了所述双-单声道编码,则所述解码单元(210)被配置为使用所述编码音频信号的第一声道的所述频谱带作为中间音频信号的第一声道的频谱带,并且被配置为使用所述编码音频信号的第二声道的所述频谱带作为所述中间音频信号的第二声道的频谱带,

[0474] 其中,如果使用了所述中-侧编码,则所述解码单元(210)被配置为基于所述编码音频信号的第一声道的所述频谱带并且基于所述编码音频信号的第二声道的所述频谱带来产生所述中间音频信号的第一声道的频谱带,以及基于所述编码音频信号的第一声道的所述频谱带并且基于所述编码音频信号的第二声道的所述频谱带,来产生所述中间音频信号的第二声道的频谱带,以及

[0475] 其中,所述装置包括去归一化器(220),所述去归一化器(220)被配置为根据去归一化值来修正所述中间音频信号的第一声道和第二声道中的至少一个声道,以获得所述解码音频信号的第一声道和第二声道。

[0476] 示例实施例24,根据示例实施例23所述的装置,

[0477] 其中,所述解码单元(210)被配置为确定所述编码音频信号是以全-中-侧编码模式、以全-双-单声道编码模式、还是以逐频带编码模式来编码的,

[0478] 其中,所述解码单元(210)被配置为:如果确定所述编码音频信号是以所述全-中-侧编码模式编码的,则根据所述编码音频信号的第一声道并且根据所述编码音频信号的第二声道来产生所述中间音频信号的第一声道,以及根据所述编码音频信号的第一声道并且根据所述编码音频信号的第二声道来产生所述中间音频信号的第二声道,

[0479] 其中,所述解码单元(210)被配置为:如果确定所述编码音频信号是以所述全-双-

单声道编码模式编码的,则使用所述编码音频信号的第一声道作为所述中间音频信号的第一声道,以及使用所述编码音频信号的第二声道作为所述中间音频信号的第二声道,以及

[0480] 其中,所述解码单元(210)被配置为:如果确定所述编码音频信号是以所述逐频带编码模式编码的,则

[0481] 针对多个频谱带中的每个频谱带,确定所述编码音频信号的第一声道的所述频谱带和所述编码音频信号的第二声道的所述频谱带是使用所述双-单声道编码来编码的还是使用所述中-侧编码来编码的,

[0482] 如果使用了所述双-单声道编码,则使用所述编码音频信号的第一声道的所述频谱带作为所述中间音频信号的第一声道的频谱带,并且使用所述编码音频信号的第二声道的所述频谱带作为所述中间音频信号的第二声道的频谱带,以及

[0483] 如果使用了所述中-侧编码,则基于所述编码音频信号的第一声道的所述频谱带并且基于所述编码音频信号的第二声道的所述频谱带,产生所述中间音频信号的第一声道的频谱带,以及基于所述编码音频信号的第一声道的所述频谱带并且基于所述编码音频信号的第二声道的所述频谱带,产生所述中间音频信号的第二声道的频谱带。

[0484] 示例实施例25,根据示例实施例23所述的装置,

[0485] 其中,所述解码单元(210)被配置为针对所述多个频谱带中的每个频谱带来确定所述编码音频信号的第一声道的所述频谱带和所述编码音频信号的第二声道的所述频谱带是使用双-单声道编码来编码的还是使用中-侧编码来编码的,

[0486] 其中,所述解码单元(210)被配置为通过重新构建所述第二声道的所述频谱带来获得所述编码音频信号的第二声道的所述频谱带,

[0487] 其中,如果使用中-侧编码,则所述编码音频信号的第一声道的所述频谱带是中央信号的频谱带,并且所述编码音频信号的第二声道的所述频谱带是侧信号的频谱带,

[0488] 其中,如果使用中-侧编码,则所述解码单元(210)被配置为根据所述侧信号的所述频谱带的校正因子并且根据与所述中央信号的所述频谱带相对应的先前中央信号的频谱带,重新构建所述侧信号的所述频谱带,其中所述先前中央信号在时间上在所述中央信号之前。

[0489] 示例实施例26,根据示例实施例25所述的装置,

[0490] 其中,如果使用了中-侧编码,则所述解码单元(210)被配置为通过根据以下公式重新构建所述侧信号的所述频谱带的频谱值来重新构建所述侧信号的所述频谱带,

[0491] $S_i = N_i + \text{facDmx}_{fb} \cdot \text{prevDmx}_i$

[0492] 其中, S_i 指示所述侧信号的所述频谱带的频谱值,

[0493] 其中, prevDmx_i 指示所述先前中央信号的频谱带的频谱值,

[0494] 其中, N_i 指示噪声填充频谱的频谱值,

[0495] 其中,根据如下公式来定义 facDmx_{fb} :

$$[0496] \quad \text{facDmx}_{fb} = \sqrt{\text{correction_factor}_{fb} - \frac{EN_{fb}}{(E\text{prevDmx}_{fb} + \epsilon)}}$$

- [0497] 其中, $\text{correction_factor}_{fb}$ 是所述侧信号的所述频谱带的所述校正因子,
- [0498] 其中, EN_{fb} 是所述噪声填充频谱的能量,
- [0499] 其中, EprevDmx_{fb} 是所述先前中央信号的所述频谱带的能量, 以及
- [0500] 其中, $\varepsilon = 0$, 或其中 $0.1 > \varepsilon > 0$ 。
- [0501] 示例实施例27, 根据示例实施例23至26中任一项所述的装置,
- [0502] 其中, 所述去归一化器 (220) 被配置为根据所述去归一化值来修正所述中间音频信号的第一声道和第二声道中的至少一个声道的多个频谱带, 以获得所述解码音频信号的第一声道和第二声道。
- [0503] 示例实施例28, 根据示例实施例23至26中任一项所述的装置,
- [0504] 其中, 所述去归一化器 (220) 被配置为根据所述去归一化值来修正所述中间音频信号的第一声道和第二声道中的至少一个声道的多个频谱带, 以获得去归一化音频信号,
- [0505] 其中, 所述装置还包括后处理单元 (230) 和变换单元 (235), 以及
- [0506] 其中, 所述后处理单元 (230) 被配置为对所述去归一化音频信号执行解码器侧时间噪声整形和解码器侧频域噪声整形中的至少一个, 以获得后处理音频信号,
- [0507] 其中, 所述变换单元 (235) 被配置为将所述后处理音频信号从频谱域变换到时域, 以获得所述解码音频信号的第一声道和第二声道。
- [0508] 示例实施例29, 根据示例实施例23至26中任一项所述的装置,
- [0509] 其中, 所述装置还包括被配置为将所述中间音频信号从频谱域变换到时域的变换单元 (215),
- [0510] 其中, 所述去归一化器 (220) 被配置为根据所述去归一化值来修正在时域中表示的中间音频信号的第一声道和第二声道中的至少一个声道, 以获得所述解码音频信号的第一声道和第二声道。
- [0511] 示例实施例30, 根据示例实施例23至26中任一项所述的装置,
- [0512] 其中, 所述装置还包括被配置为将所述中间音频信号从频谱域变换到时域的变换单元 (215),
- [0513] 其中, 所述去归一化器 (220) 被配置为根据所述去归一化值来修正在时域中表示的中间音频信号的第一声道和第二声道中的至少一个声道, 以获得去归一化音频信号,
- [0514] 其中, 所述装置还包括后处理单元 (235), 所述后处理单元 (235) 被配置为处理作为感知白化音频信号的所述去归一化音频信号, 以获得所述解码音频信号的第一声道和第二声道。
- [0515] 示例实施例31, 根据示例实施例29或30所述的装置,
- [0516] 其中, 所述装置还包括被配置为对所述中间音频信号执行解码器侧时间噪声整形的频谱域后处理器 (212),
- [0517] 其中, 所述变换单元 (215) 被配置为在已经对所述中间音频信号执行了解码器侧时间噪声整形之后, 将所述中间音频信号从频谱域变换到时域。
- [0518] 示例实施例32, 根据示例实施例23至31中任一项所述的装置,
- [0519] 其中, 所述解码单元 (210) 被配置为对所述编码音频信号应用解码器侧立体声智能间隙填充。
- [0520] 示例实施例33, 根据示例实施例23至32中任一项所述的装置, 其中, 所述解码音频

信号是恰好包括两个声道的音频立体声信号。

[0521] 示例实施例34,一种用于对包括四个或更多个声道的编码音频信号进行解码以获得包括四个或更多个声道的解码音频信号的四个声道的系统,其中,所述系统包括:

[0522] 根据示例实施例23至32中任一项所述的第一装置(270),用于对所述编码音频信号的四个或更多个声道中的第一声道和第二声道进行解码,以获得所述解码音频信号的第一声道和第二声道,以及

[0523] 根据示例实施例23至32中任一项所述的第二装置(280),用于对所述编码音频信号的四个或更多个声道中的第三声道和第四声道进行解码,以获得所述解码音频信号的第三声道和第四声道。

[0524] 示例实施例35,一种用于根据音频输入信号来产生编码音频信号以及根据编码音频信号来产生解码音频信号的系统,包括:

[0525] 根据示例实施例1至21中任一项所述的装置(310),其中根据示例实施例1至21中任一项所述的装置(310)被配置为根据所述音频输入信号产生所述编码音频信号,以及

[0526] 根据示例实施例23至33中任一项所述的装置(320),其中根据示例实施例23至33中任一项所述的装置(320)被配置为根据所述编码音频信号产生所述解码音频信号。

[0527] 示例实施例36,一种用于根据音频输入信号来产生编码音频信号以及根据编码音频信号来产生解码音频信号的系统,包括:

[0528] 根据示例实施例22所述的系统,其中,根据示例实施例22所述的系统被配置为根据所述音频输入信号产生所述编码音频信号,以及

[0529] 根据示例实施例34所述的系统,其中,根据示例实施例34所述的系统被配置为根据所述编码音频信号产生所述解码音频信号。

[0530] 示例实施例37,一种用于对包括两个或更多个声道的音频输入信号的第一声道和第二声道进行编码以获得编码音频信号的方法,其中,所述方法包括:

[0531] 根据所述音频输入信号的第一声道并且根据所述音频输入信号的第二声道来确定所述音频输入信号的归一化值,

[0532] 通过根据所述归一化值修正所述音频输入信号的第一声道和第二声道中的至少一个声道来确定归一化音频信号的第一声道和第二声道,

[0533] 产生具有第一声道和第二声道的处理后的音频信号,使得所述处理后的音频信号的第一声道的一个或多个频谱带是所述归一化音频信号的第一声道的一个或多个频谱带,使得所述处理后的音频信号的第二声道的一个或多个频谱带是所述归一化音频信号的第二声道的一个或多个频谱带,使得所述处理后的音频信号的第一声道的至少一个频谱带是根据所述归一化音频信号的第一声道的频谱带并且根据所述归一化音频信号的第二声道的频谱带的中央信号的频谱带,以及使得所述处理后的音频信号的第二声道的至少一个频谱带是根据所述归一化音频信号的第一声道的频谱带并且根据所述归一化音频信号的第二声道的频谱带的侧信号的频谱带,以及编码所述处理后的音频信号以获得所述编码音频信号。

[0534] 示例实施例38,一种用于对包括第一声道和第二声道的编码音频信号进行解码以获得包括两个或更多个声道的解码音频信号的第一声道和第二声道的方法,其中所述方法包括:

[0535] 针对多个频谱带中的每个频谱带,确定所述编码音频信号的第一声道的所述频谱带和所述编码音频信号的第二声道的所述频谱带是使用双-单声道编码来编码的还是使用中-侧编码来编码的,

[0536] 如果使用了所述双-单声道编码,则使用所述编码音频信号的第一声道的所述频谱带作为中间音频信号的第一声道的频谱带,并且使用所述编码音频信号的第二声道的所述频谱带作为所述中间音频信号的第二声道的频谱带,

[0537] 如果使用了所述中-侧编码,则基于所述编码音频信号的第一声道的所述频谱带并且基于所述编码音频信号的第二声道的所述频谱带来产生所述中间音频信号的第一声道的频谱带,以及基于所述编码音频信号的第一声道的所述频谱带并且基于所述编码音频信号的第二声道的所述频谱带来产生所述中间音频信号的第二声道的频谱带,以及

[0538] 根据去归一化值,修正所述中间音频信号的第一声道和第二声道中的至少一个声道,以获得解码音频信号的第一声道和第二声道。

[0539] 示例实施例39,一种计算机程序,用于在计算机或信号处理器上执行时实施根据示例实施例37或38所述的方法。

[0540] 文献

[0541] [1]J.Herre,E.Eberlein and K.Brandenburg,"Combined Stereo Coding,"in 93rd AES Convention, San Francisco,1992.

[0542] [2]J.D.Johnston and A.J.Ferreira,"Sum-difference stereo transform coding,"in Proc.ICASSP,1992.

[0543] [3]ISO/IEC 11172-3,Information technology-Coding of moving pictures and associated audio for digital storage media at up to about 1,5Mbit/s-Part 3:Audio,1993.

[0544] [4]ISO/IEC 13818-7,Information technology-Generic coding of moving pictures and associated audio information-Part 7:Advanced Audio Coding(AAC), 2003.

[0545] [5]J.-M.Valin,G.Maxwell,T.B.Terriberry and K.Vos,"High-Quality,Low-Delay Music Coding in the Opus Codec,"in Proc.AES 135th Convention,New York, 2013.

[0546] [6a]3GPP TS 26.445,Codec for Enhanced Voice Services(EVS);Detailed algorithmic description,V 12.5.0,Dezember 2015.

[0547] [6b]3GPP TS 26.445,Codec for Enhanced Voice Services(EVS);Detailed algorithmic description,V 13.3.0,September 2016.

[0548] [7]H.Purnhagen,P.Carlsson,L.Villemoes,J.Robilliard,M.Neusinger,C.Helmrich,J.Hilpert,N.Rettelbach,S.Disch and B.Edler,"Audio encoder,audio decoder and related methods for processing multi-channel audio signals using complex prediction".US Patent 8,655,670 B2,18February 2014.

[0549] [8]G.Markovic,F.Guillaume,N.Rettelbach,C.Helmrich and B.Schubert,"Linear prediction based coding scheme using spectral domain noise shaping".European Patent 2676266 B1,14February 2011.

[0550] [9] S. Disch, F. Nagel, R. Geiger, B. N. Thoshkahna, K. Schmidt, S. Bayer, C. Neukam, B. Edler and C. Helmrich, "Audio Encoder, Audio Decoder and Related Methods Using Two-Channel Processing Within an Intelligent Gap Filling Framework". International Patent PCT/EP2014/065106, 15072014.

[0551] [10] C. Helmrich, P. Carlsson, S. Disch, B. Edler, J. Hilpert, M. Neusinger, H. Purnhagen, N. Rettelbach, J. Robilliard and L. Villemoes, "Efficient Transform Coding Of Two-channel Audio Signals By Means Of Complex-valued Stereo Prediction," in Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, Prague, 2011.

[0552] [11] C. R. Helmrich, A. Niedermeier, S. Bayer and B. Edler, "Low-complexity semi-parametric joint-stereo audio transform coding," in Signal Processing Conference (EUSIPCO), 2015 23rd European, 2015.

[0553] [12] H. Malvar, "A Modulated Complex Lapped Transform and its Applications to Audio Processing" in Acoustics, Speech, and Signal Processing (ICASSP), 1999. Proceedings., 1999 IEEE International Conference on, Phoenix, AZ, 1999.

[0554] [13] B. Edler and G. Schuller, "Audio coding using a psychoacoustic pre- and post-filter," Acoustics, Speech, and Signal Processing, 2000. ICASSP'00.

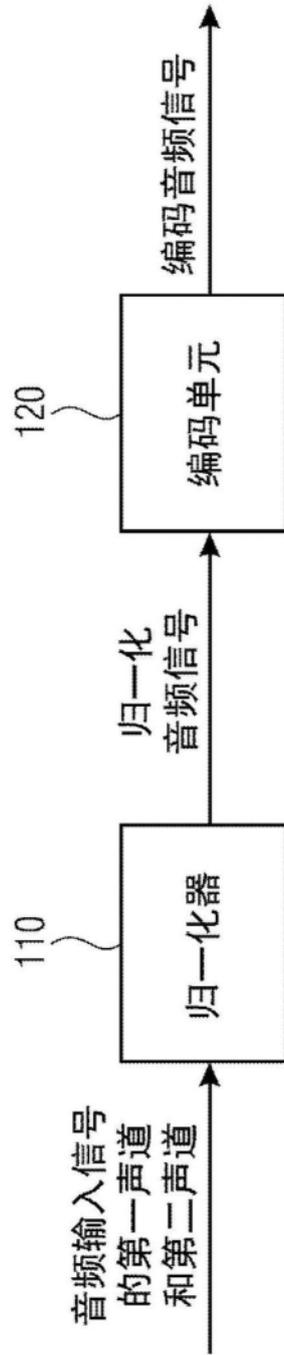


图1a

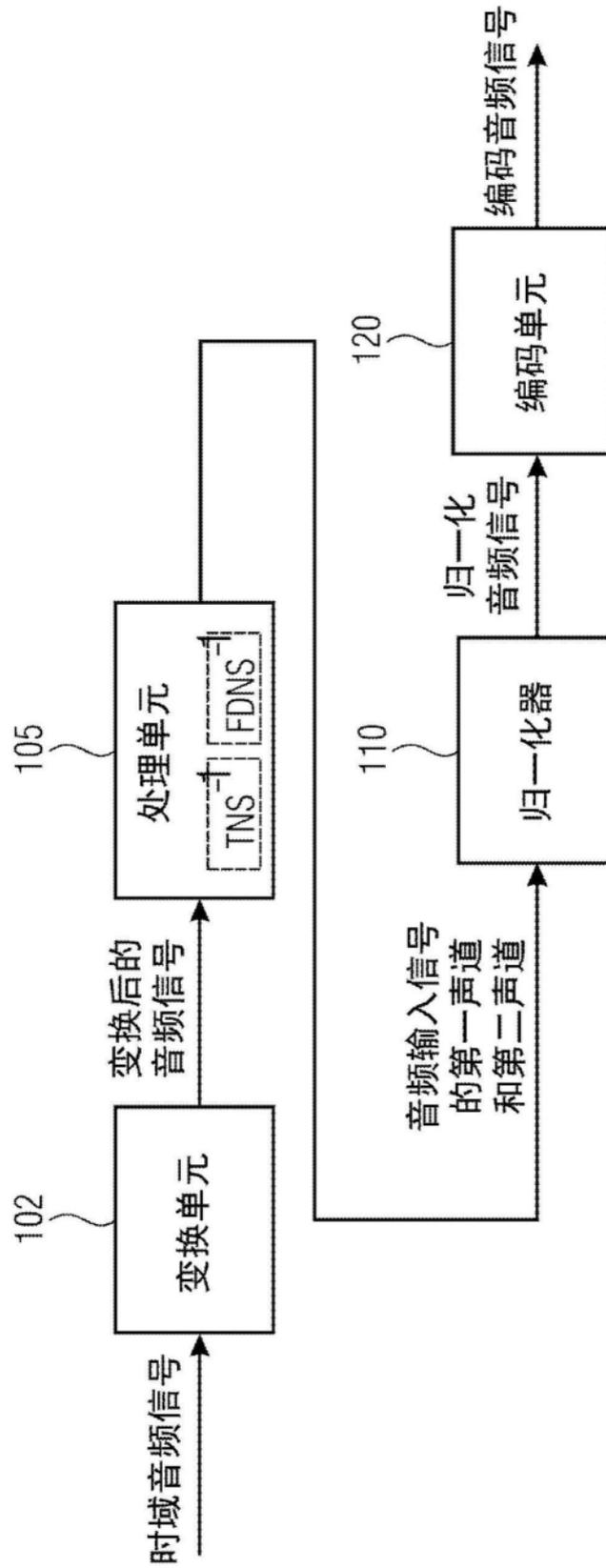


图1b

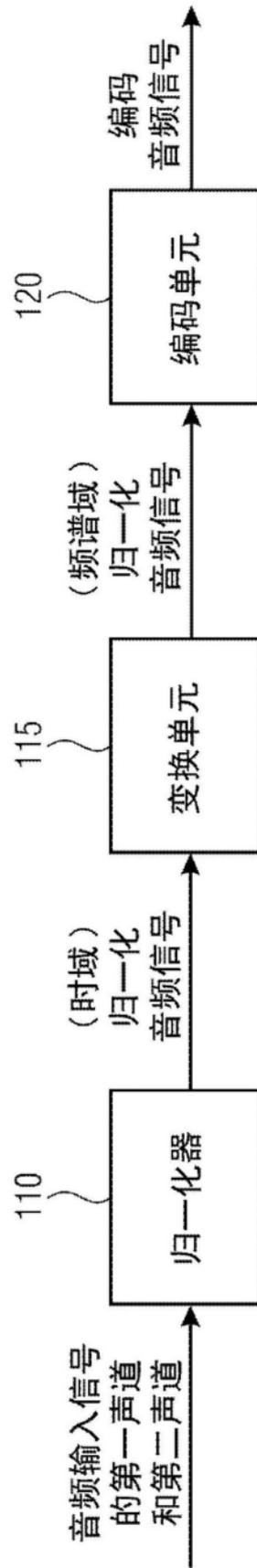


图1c

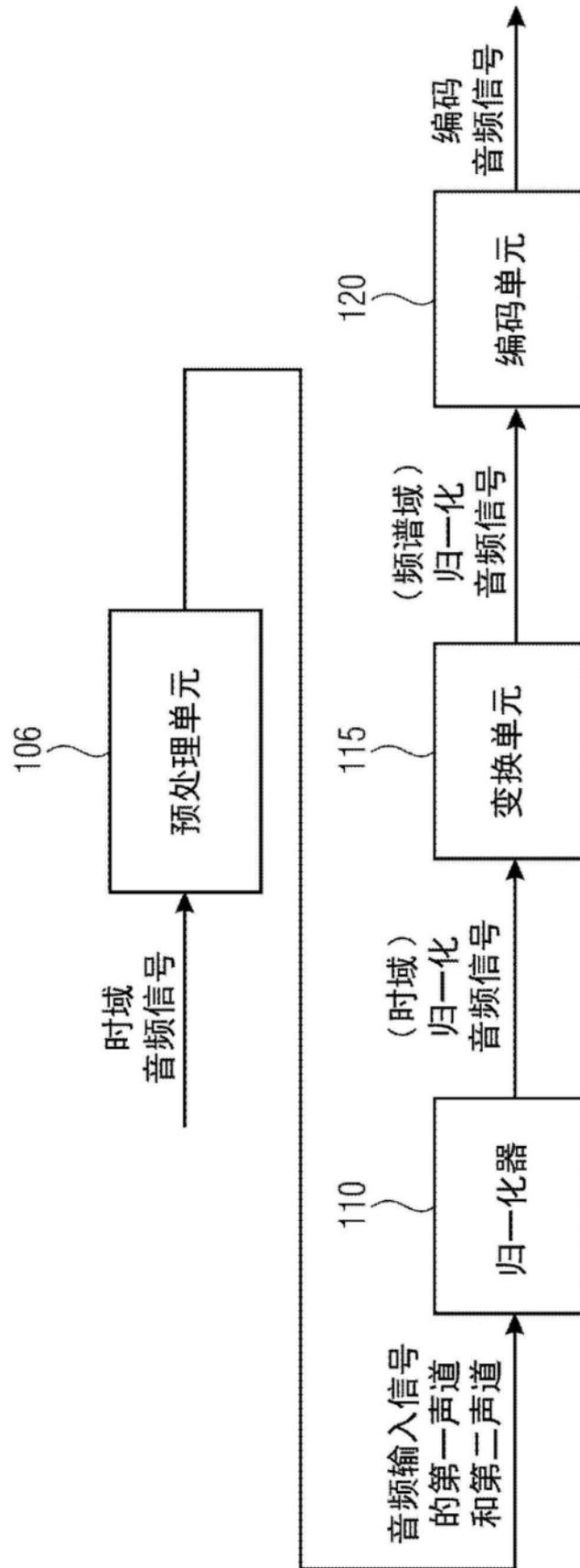


图1d

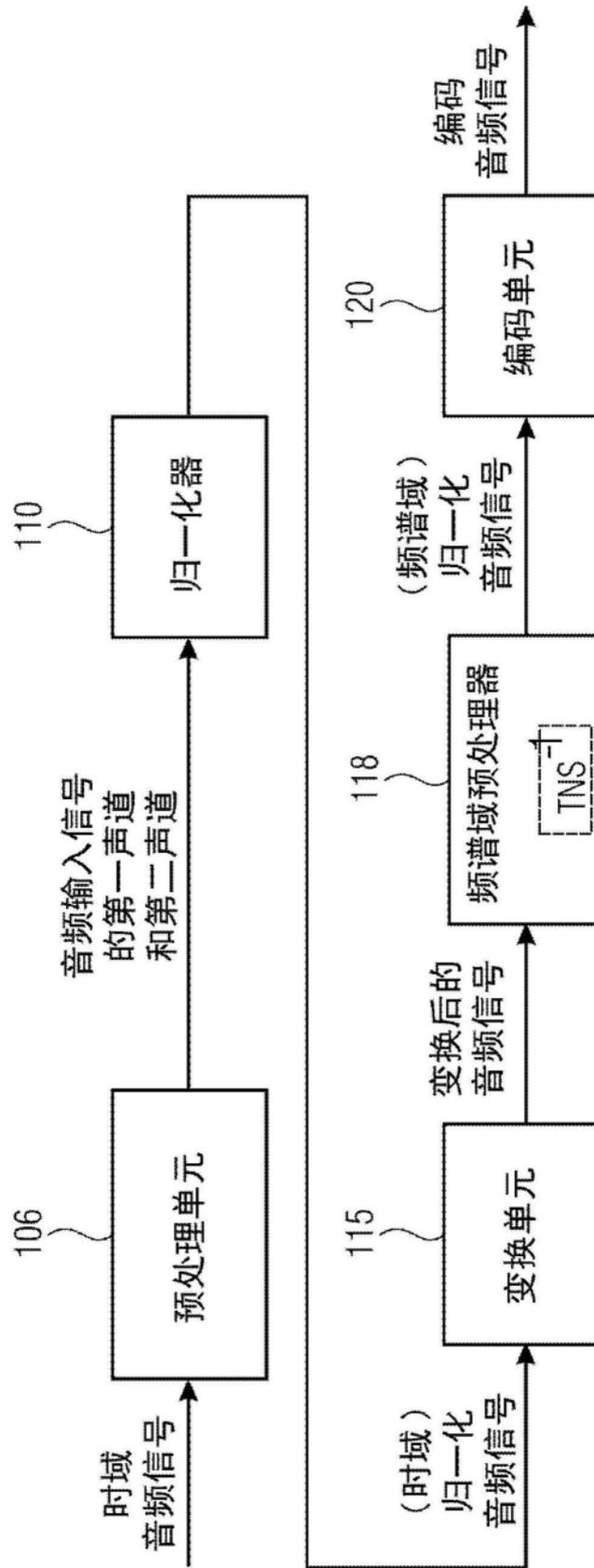


图1e

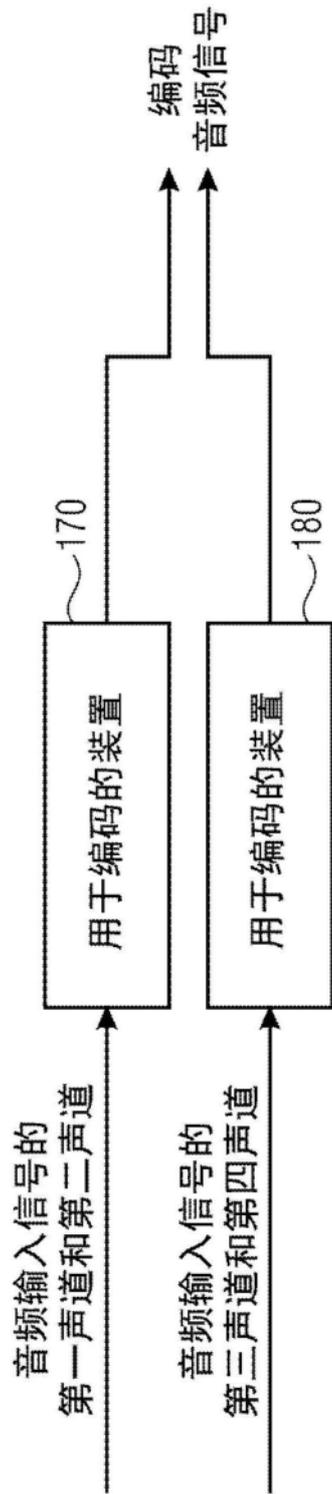


图1f

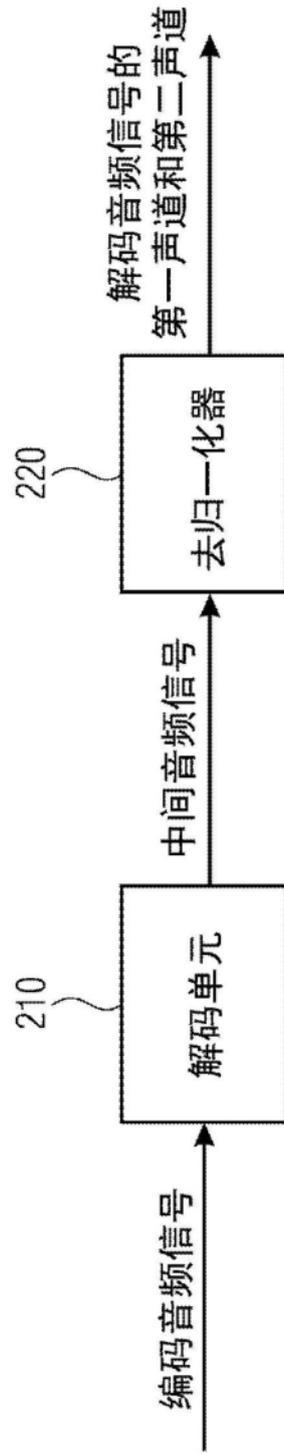


图2a

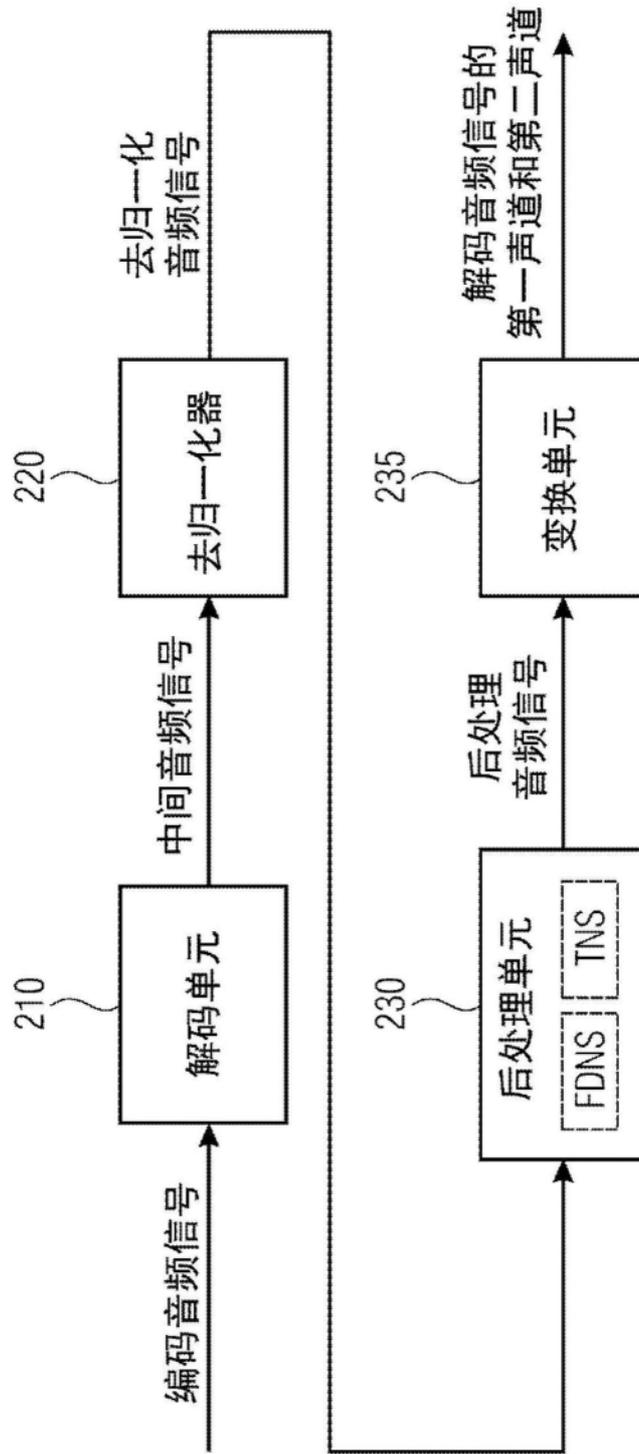


图2b

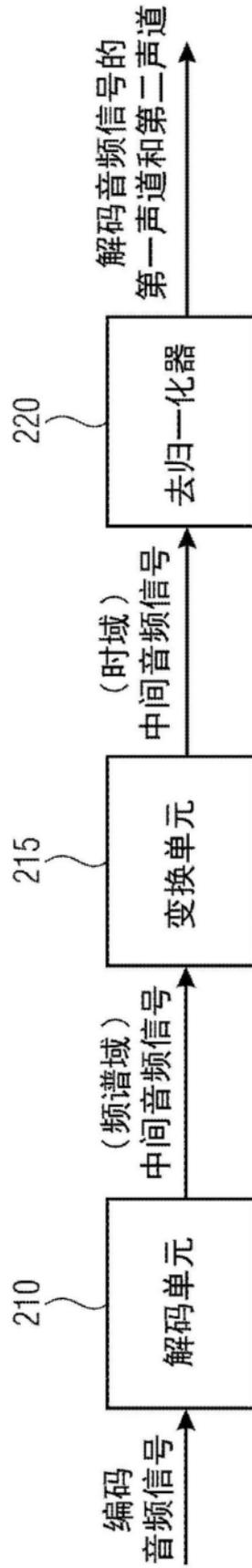


图2c

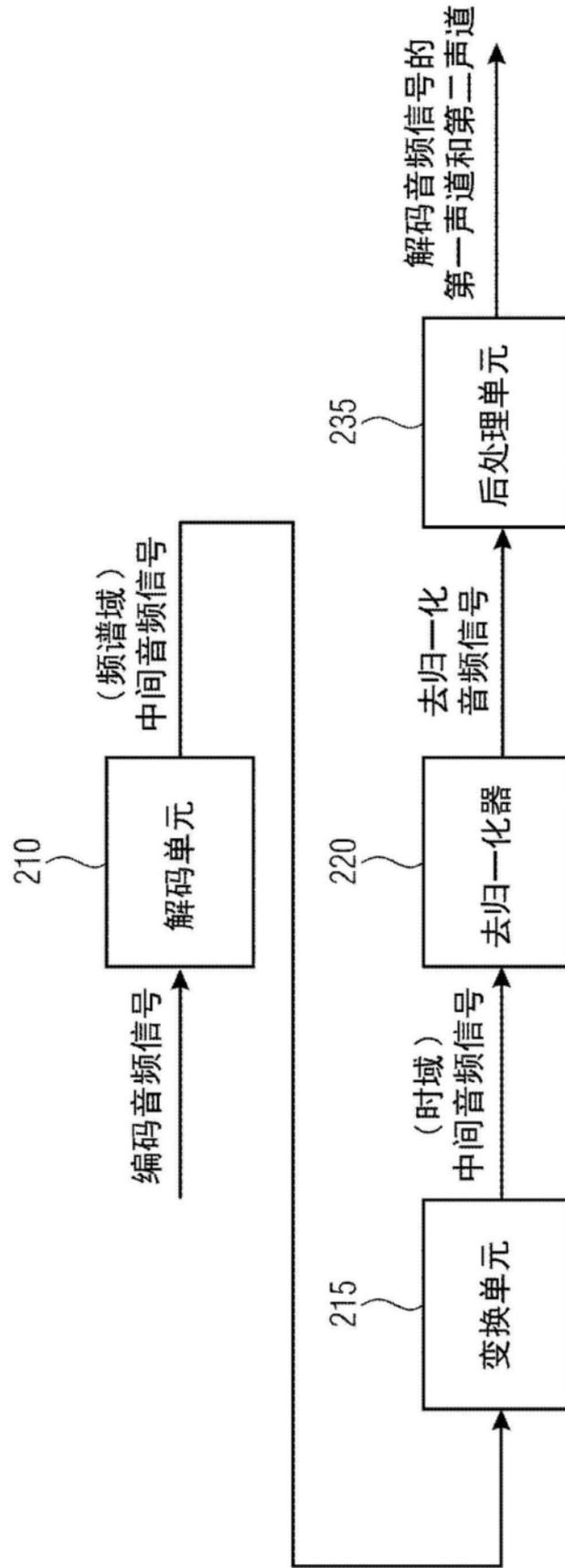


图2d

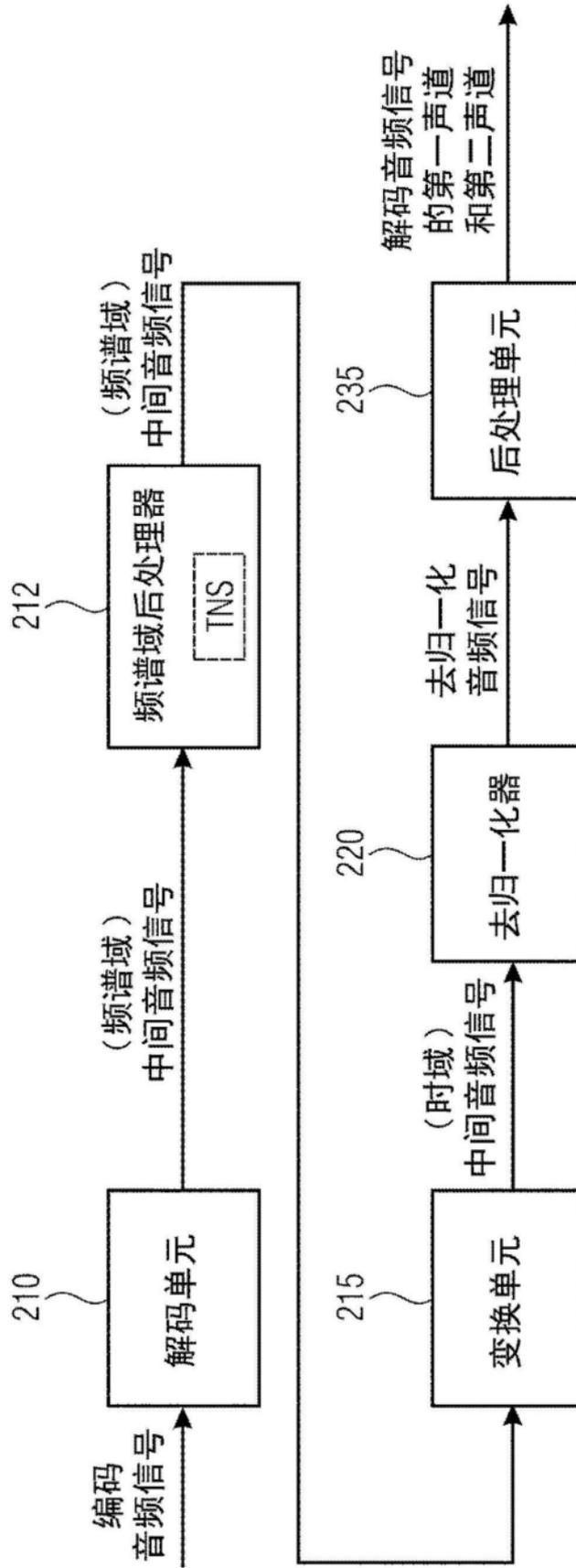


图2e

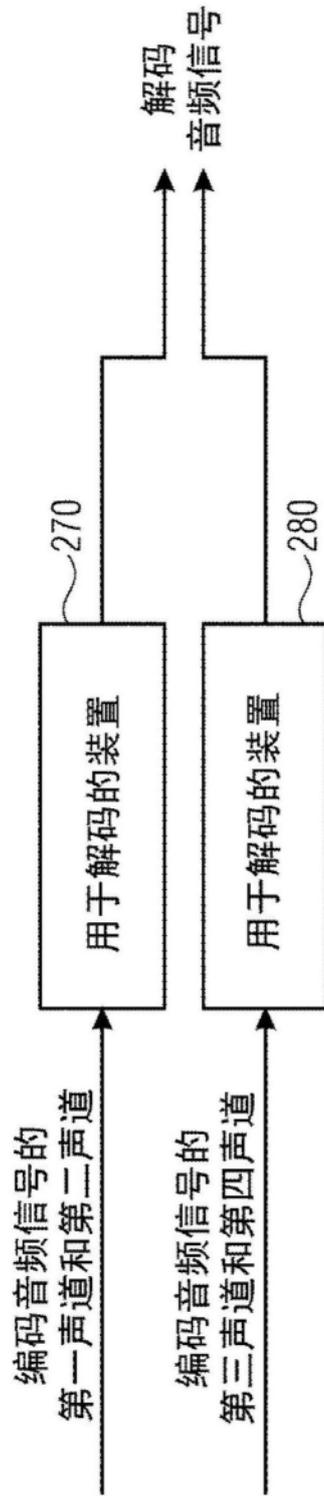


图2f

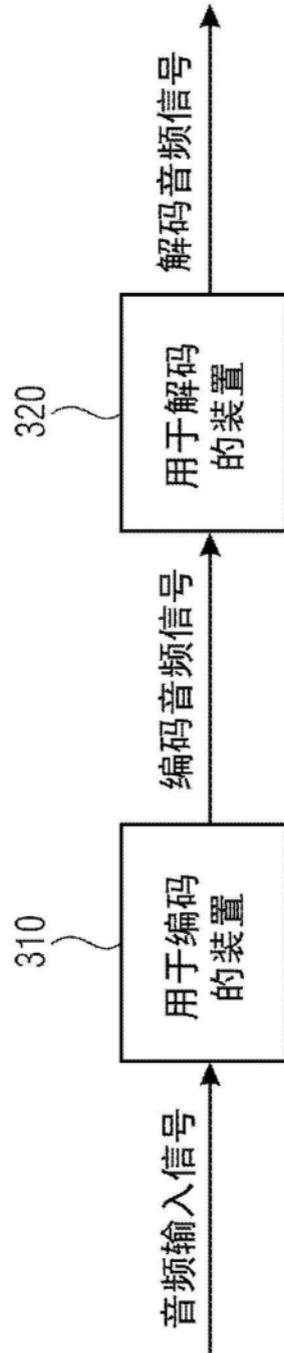


图3

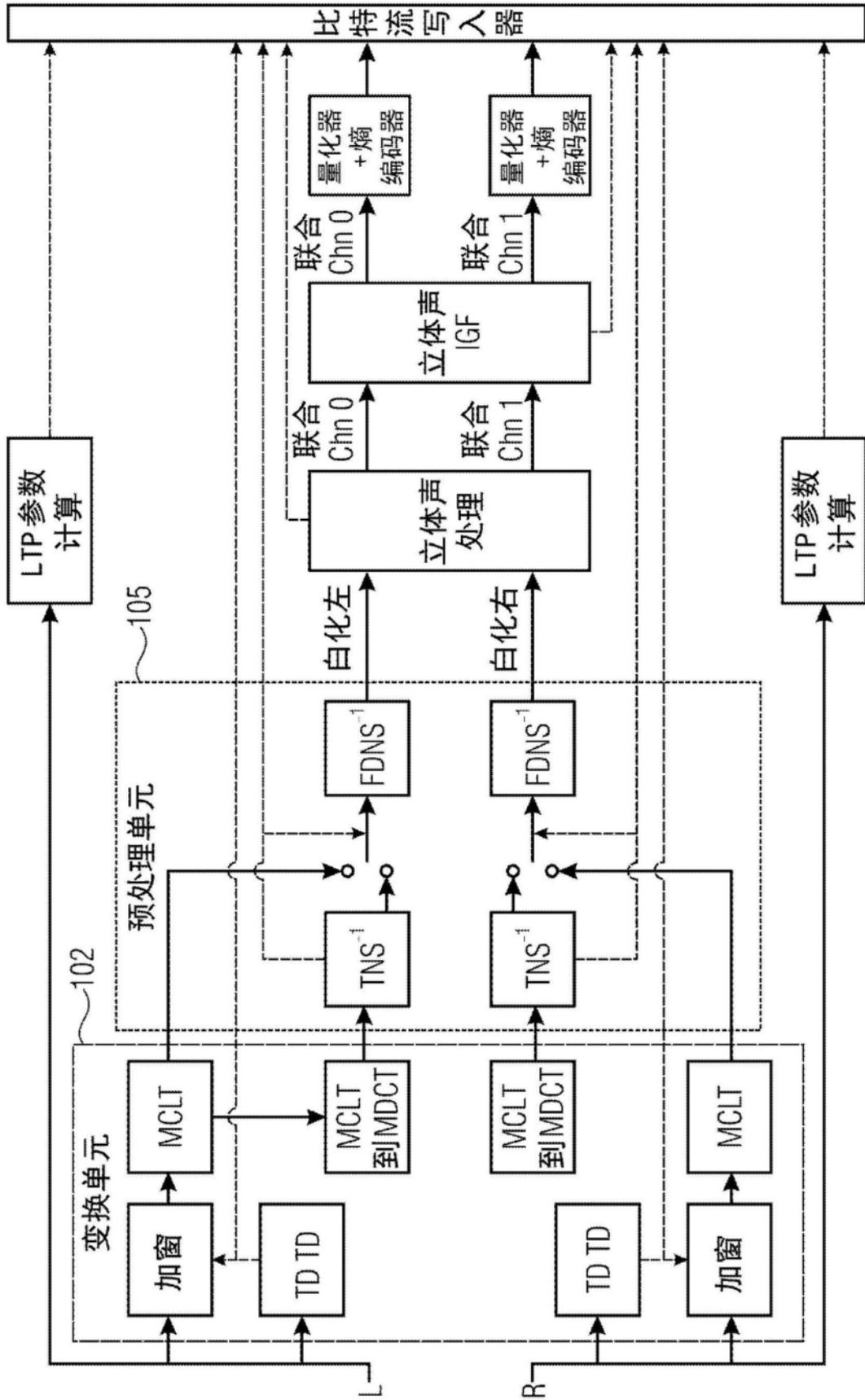


图4

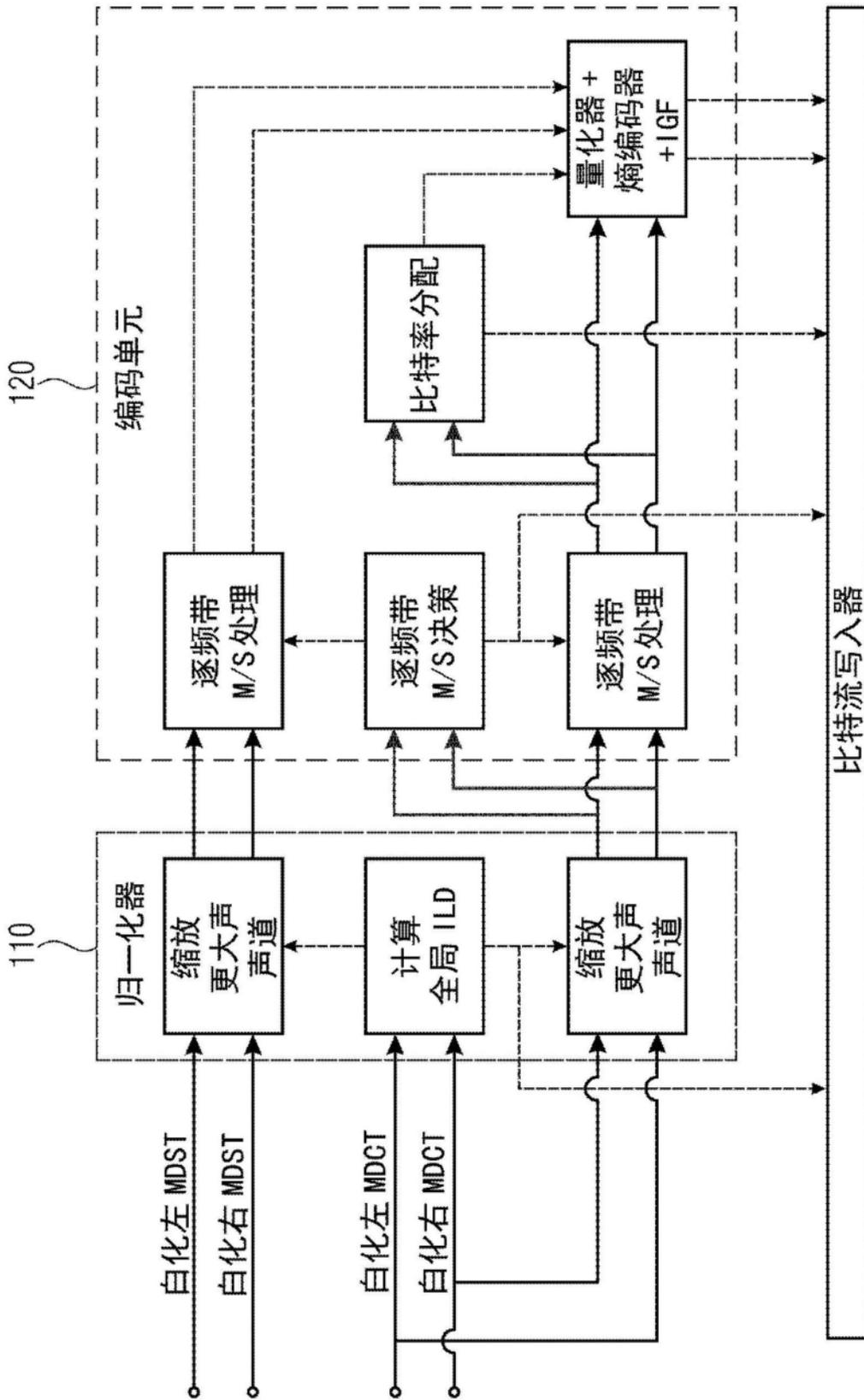


图5

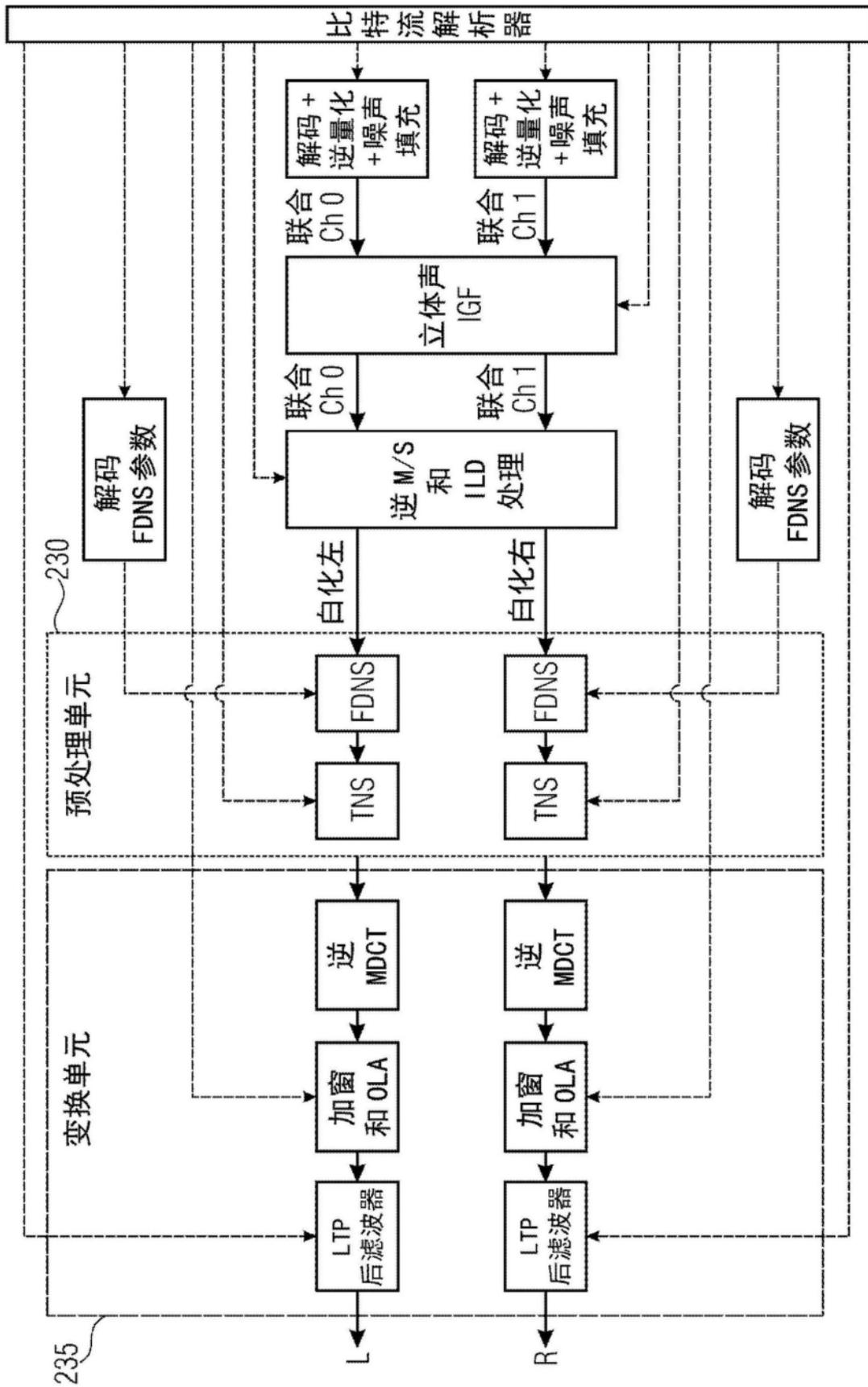


图6

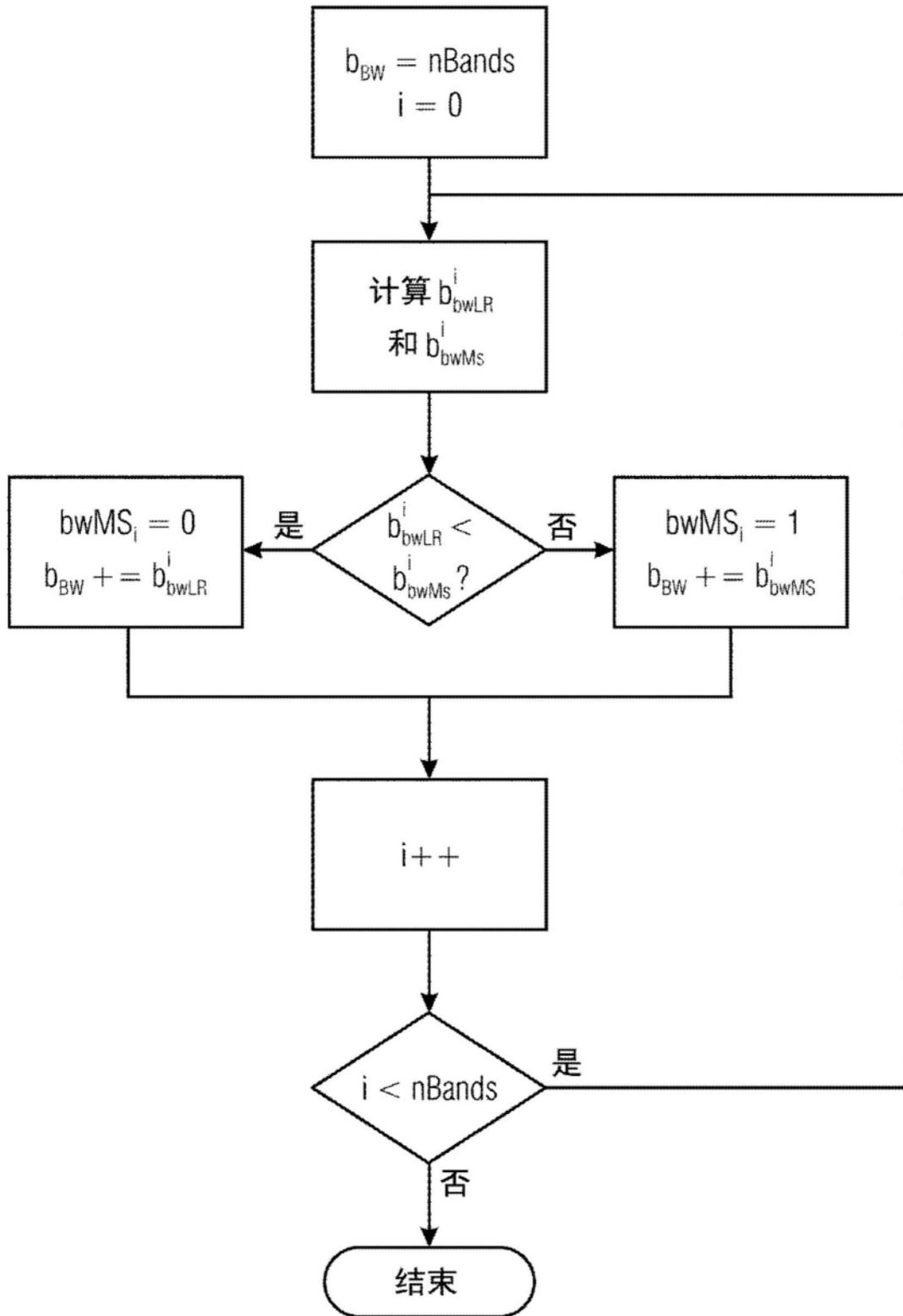


图7

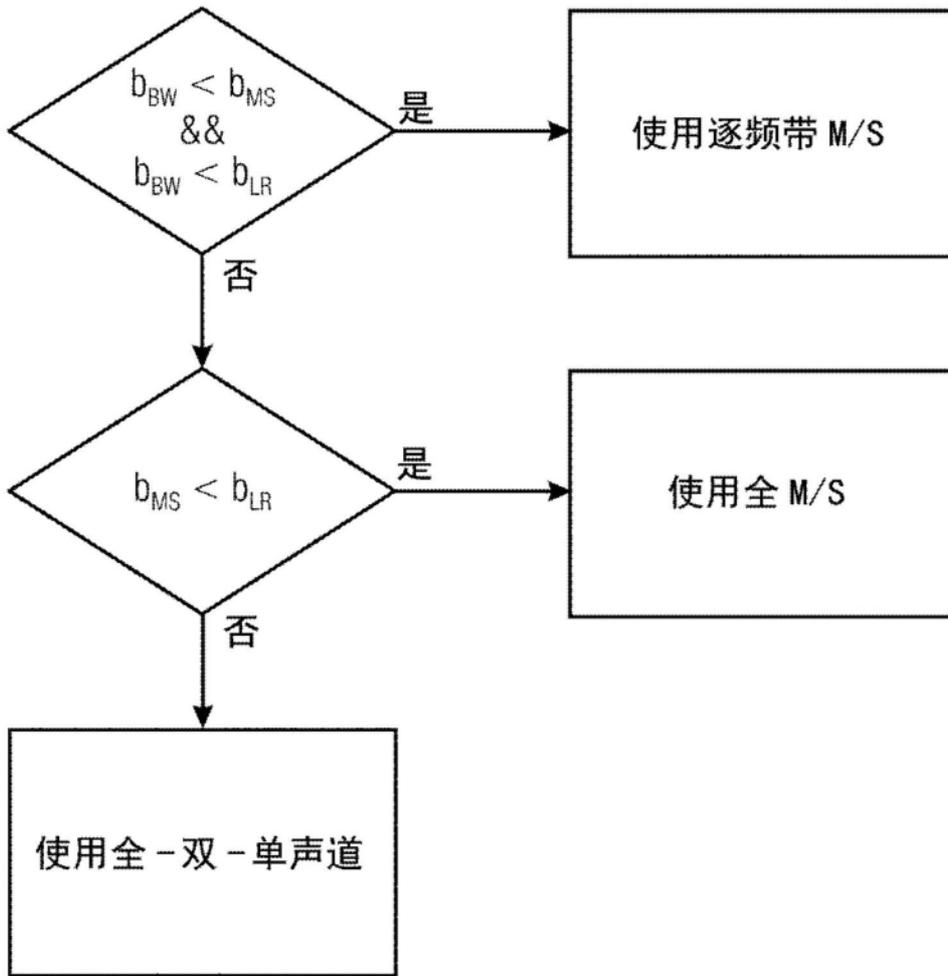


图8

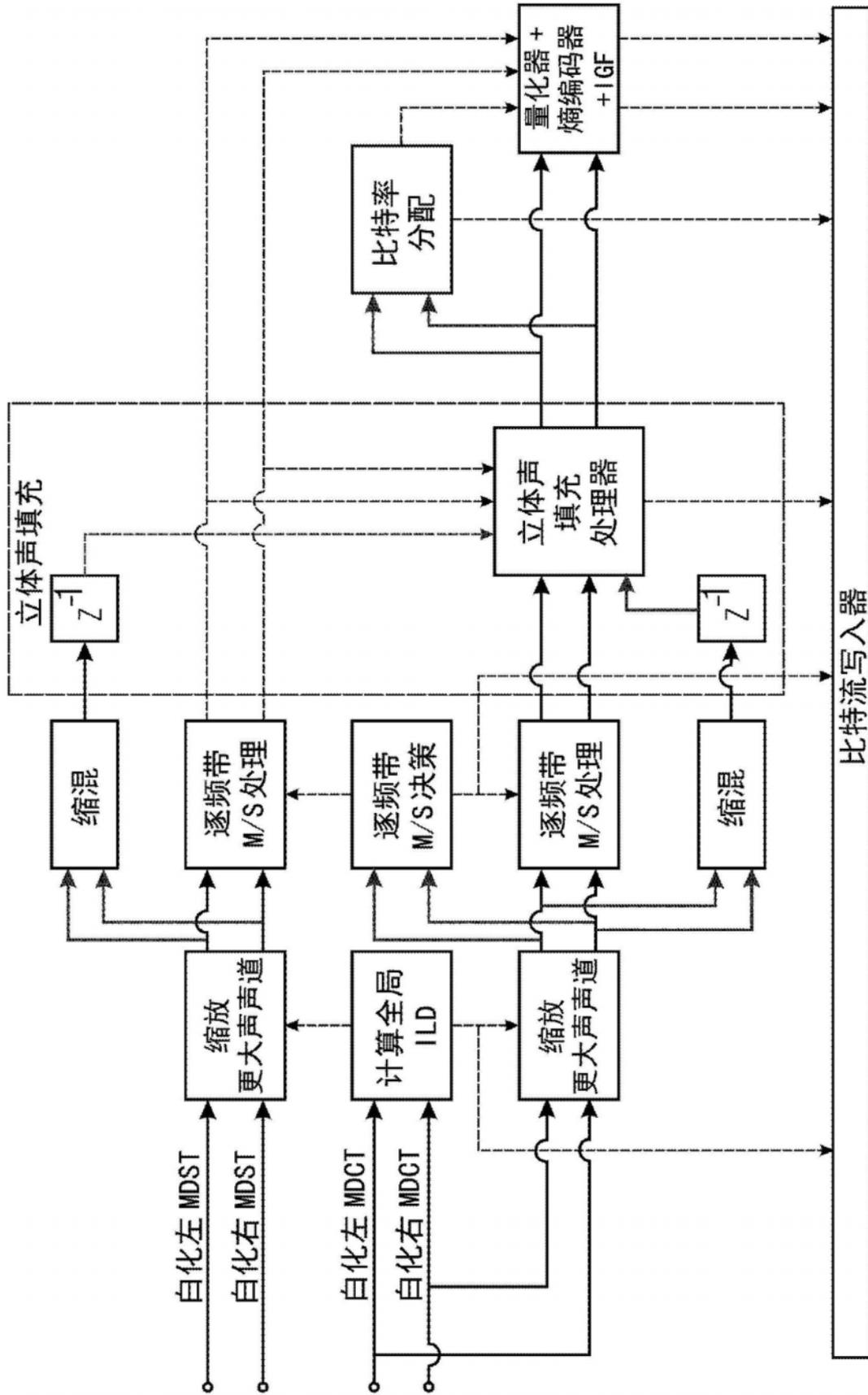


图9

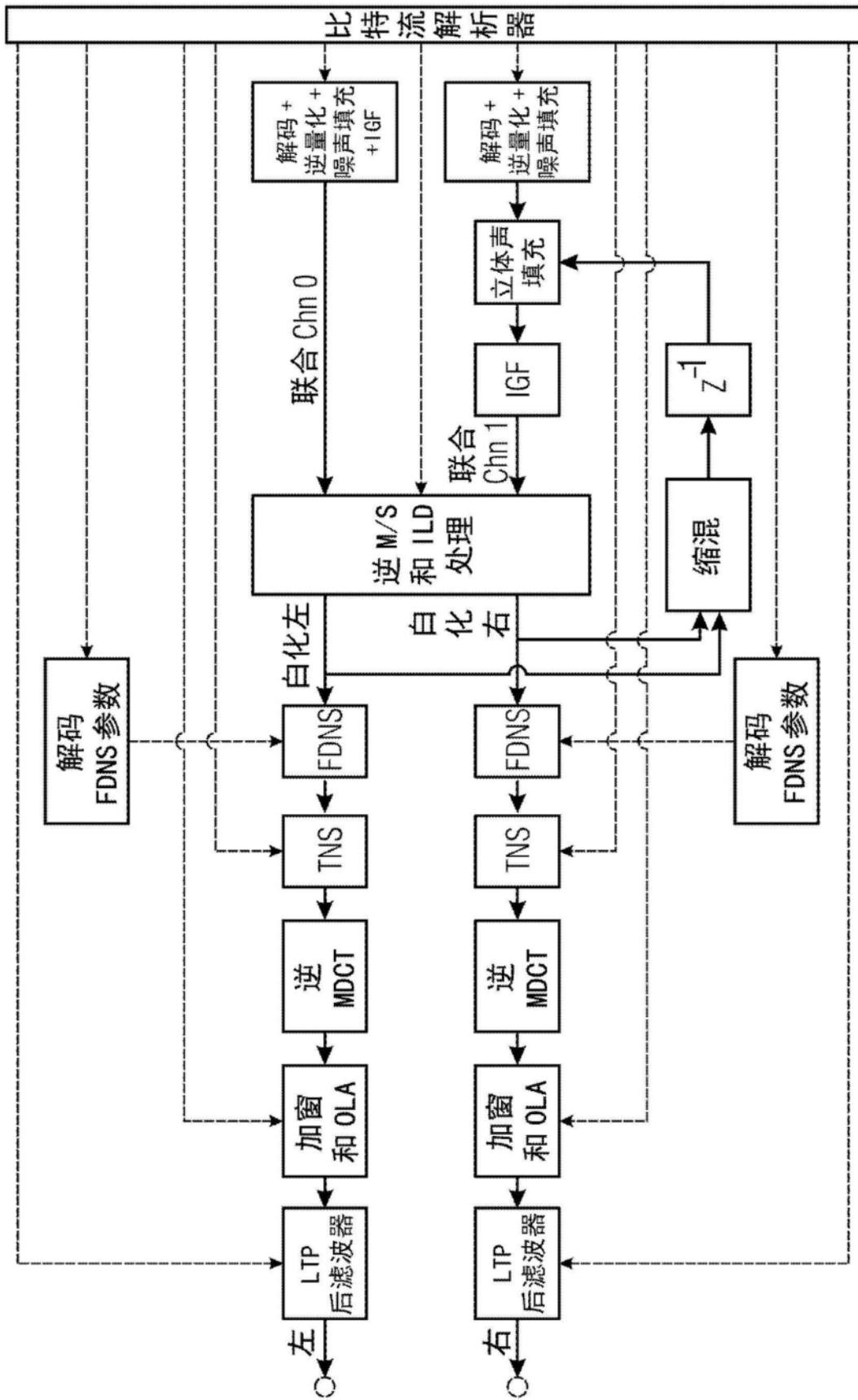


图10

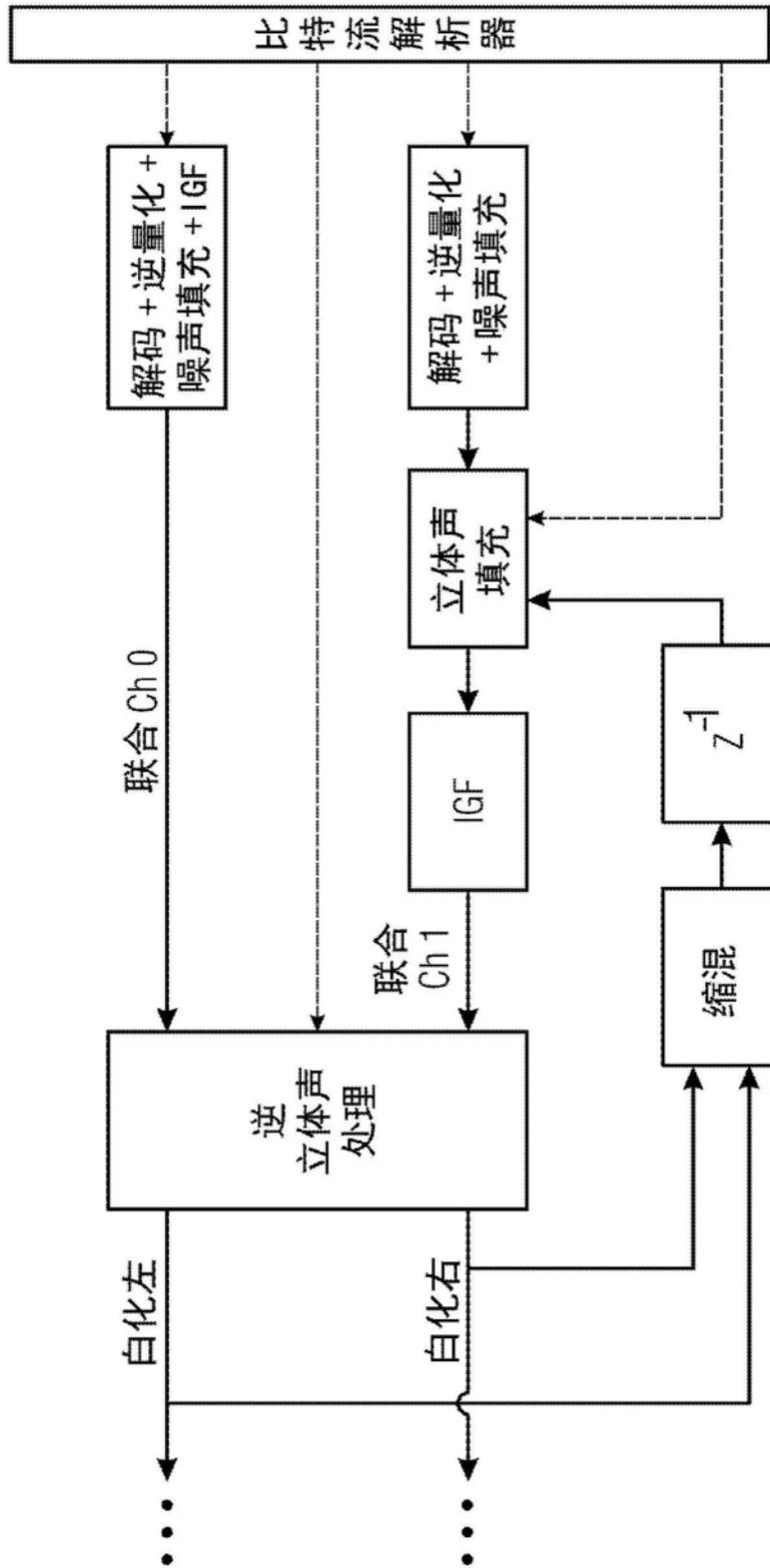


图11

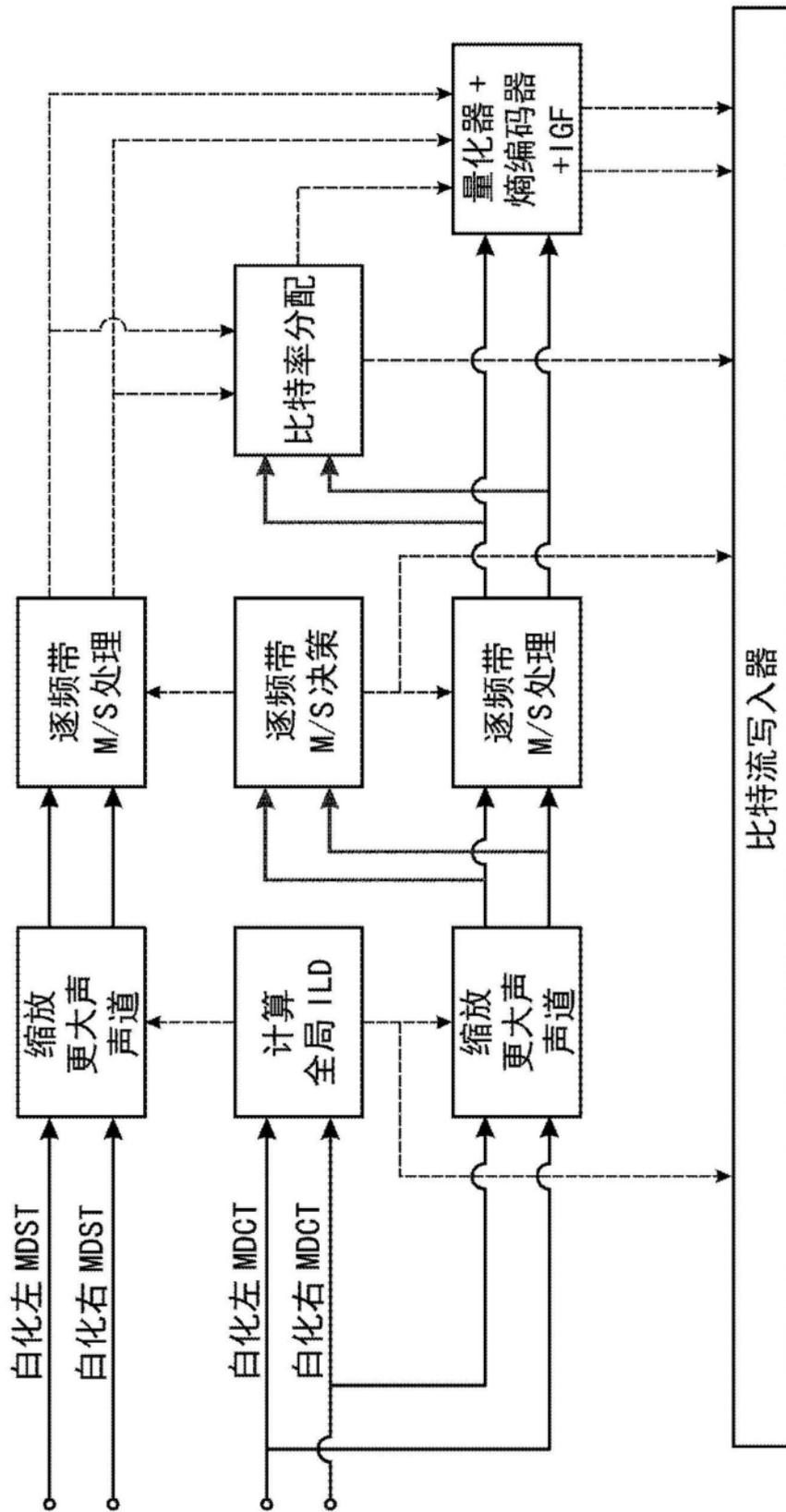


图12

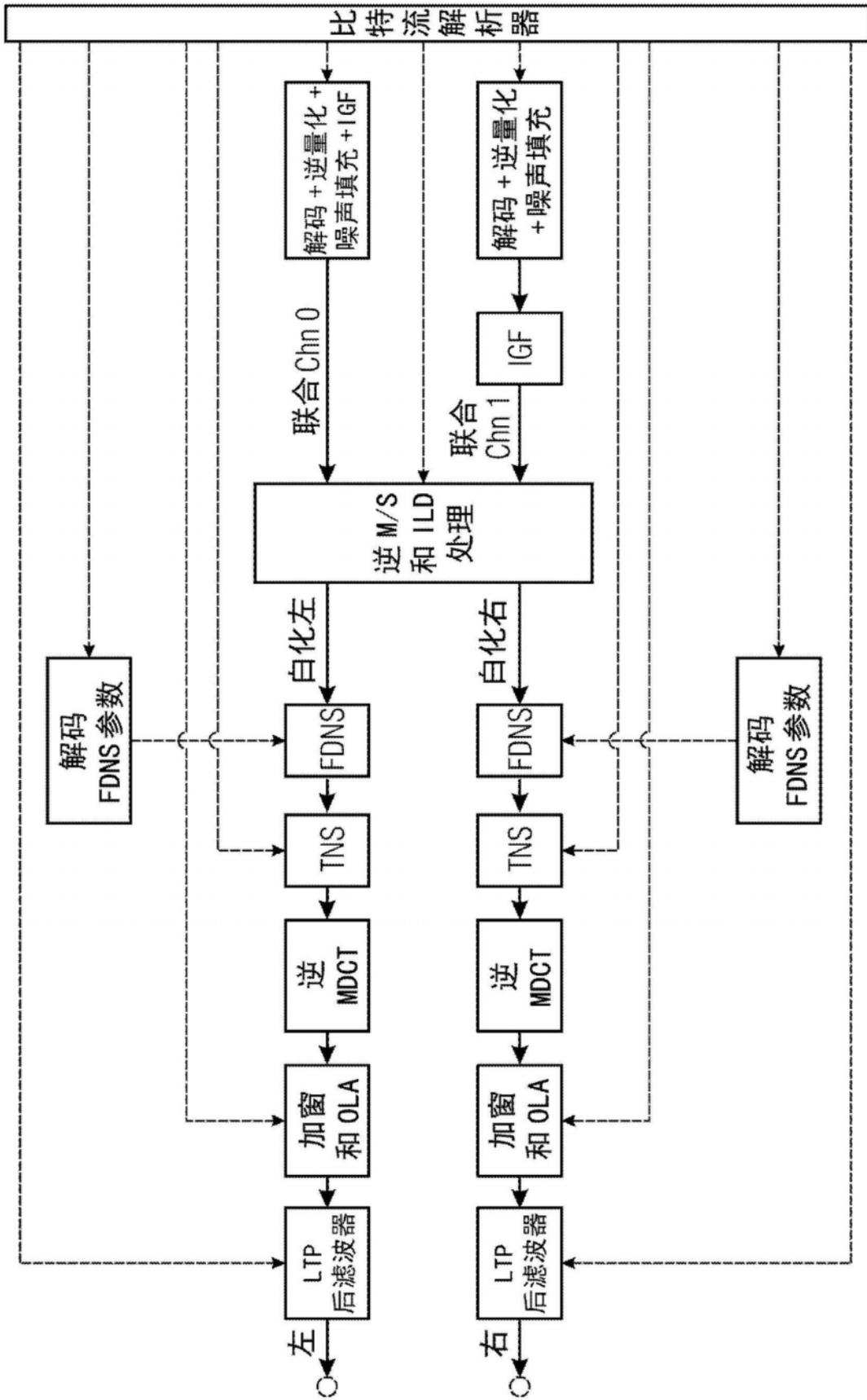


图13