



(21) 申请号 202110466890.5

(22) 申请日 2015.12.08

(65) 同一申请的已公布的文献号
申请公布号 CN 112967727 A

(43) 申请公布日 2021.06.15

(30) 优先权数据
62/089,563 2014.12.09 US

(62) 分案原申请数据
201580067028.X 2015.12.08

(73) 专利权人 杜比国际公司
地址 荷兰阿姆斯特丹

(72) 发明人 A·比斯沃斯 T·弗瑞尔德里驰
K·佩秋

(74) 专利代理机构 中国贸促会专利商标事务所
有限公司 11038

专利代理师 刘前红

(51) Int.Cl.
G10L 19/005 (2013.01)
G10L 19/02 (2013.01)

(56) 对比文件
CN 107004417 A, 2017.08.01

审查员 李召卿

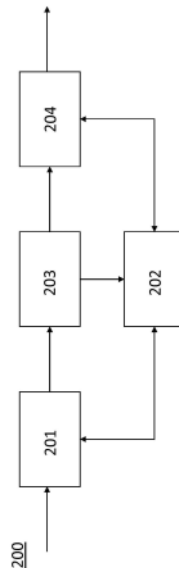
权利要求书3页 说明书16页 附图5页

(54) 发明名称

MDCT域错误掩盖

(57) 摘要

本发明涉及MDCT域错误掩盖。一种错误掩盖音频解码方法包括：接收包括对音频信号的时域样本的帧进行编码的MDCT系数集合的分组；将接收分组识别为错误的；基于与紧挨在错误分组前面的接收分组相关联的对应MDCT系数来产生替代错误分组的MDCT系数集合的估计MDCT系数；将估计MDCT系数中的MDCT系数的第一子集的符号指定为与所述前一个分组的对应MDCT系数的符号一致，其中，第一子集包括与音调状谱区间相关联的那些MDCT系数；随机地指定估计MDCT系数中的MDCT系数的第二子集的符号，其中，第二子集包括与噪声状谱区间相关联的MDCT系数；用包含估计MDCT系数和指定的符号的掩盖分组替代错误分组。



1. 一种用于掩盖在基于修正离散余弦变换MDCT的音频解码器中将被解码的数据分组中的错误的方法,所述基于MDCT的音频解码器被布置为将分组序列解码为解码帧序列,所述方法包括:

从被布置为对音频信号进行编码的基于MDCT的音频编码器接收包括与所述音频信号的 N 个加窗时域样本相关联的 $N/2$ 个MDCT系数的分组;

在所述分组包括一个或多个错误时,将所述分组识别为错误分组;

估计第一子集,所述第一子集包括中间帧的第一半的 $N/4$ 个加窗时域混叠样本,所述中间帧包括与所述错误分组相关联的 N 个加窗时域混叠样本,所述估计第一子集基于所述第一子集的加窗时域混叠样本和所述音频信号的 N 个加窗时域样本中的加窗时域样本之间的关系;

估计第二子集,所述第二子集包括所述中间帧的第一半的剩余 $N/4$ 个加窗时域混叠样本,所述估计第二子集基于所述第二子集的加窗时域混叠样本和所述第一子集的加窗时域混叠样本之间的对称关系;以及

通过将所述中间帧的第一半与前一个中间帧的第二半相加来产生与所述错误分组相关联的估计解码帧,所述前一个中间帧与在所述分组序列中紧挨在所述错误分组前面的接收分组相关联。

2. 根据权利要求1所述的方法,其中所述第一子集的估计基于与在所述分组序列中紧挨在所述错误分组前面的接收分组相关联的前一个解码帧。

3. 根据权利要求2所述的方法,进一步包括:

估计第三子集,所述第三子集包括与所述错误分组相关联的所述中间帧的第二半的 $N/4$ 个加窗时域混叠样本,所述估计第三子集基于与所述错误分组相关联的估计解码帧;以及

估计第四子集,所述第四子集包括所述中间帧的第二半的剩余 $N/4$ 个加窗时域混叠样本,所述估计第四子集基于所述第四子集的加窗时域混叠样本和所估计的第三子集的加窗时域混叠样本之间的对称关系。

4. 根据权利要求3所述的方法,进一步包括:

通过将所述中间帧的第二半与后一个中间帧的第一半相加来产生与在所述分组序列中紧挨在所述错误分组后面的接收分组相关联的后一个估计解码帧,所述后一个中间帧与在所述分组序列中紧挨在所述错误分组后面的接收分组相关联。

5. 根据权利要求2所述的方法,其中,包括 $N/4$ 个加窗时域混叠样本的第一子集是所述中间帧的第一半的第一半,并且其中,对于 n 等于 $0, 1, \dots, N/4-1$,所述第一子集的样本编号 n 被估计为前一个解码帧的样本编号 n 的加窗版本减去前一个解码帧的样本编号 $N/2-1-n$ 的加窗版本。

6. 根据权利要求4所述的方法,其中,包括 $N/4$ 个加窗时域混叠样本的第一子集是所述中间帧的第一半的第一半,包括 $N/4$ 个加窗时域混叠样本的第三子集是所述中间帧的第二半的第一半,并且其中,对于 n 等于 $0, 1, \dots, N/4-1$,所述第一子集的样本编号 n 被估计为前一个解码帧的样本编号 n 的加窗版本减去前一个解码帧的样本编号 $N/2-1-n$ 的加窗版本,并且其中,对于 n 等于 $0, 1, \dots, N/4-1$,所述第三子集的样本编号 n 被估计为所述估计解码帧的样本编号 n 的加窗版本加上所述估计解码帧的样本编号 $N/2-1-n$ 的加窗版本。

7. 根据权利要求1所述的方法,其中,所述第一子集的估计基于偏移集合,所述偏移集

合包括与在所述分组序列中紧挨在错误分组前面的接收分组相关联的前一个解码帧以及
与在所述分组序列中紧挨在与所述前一个解码帧相关联的分组前面的接收分组相关联的
再前一个解码帧的 $N/2$ 个样本,所述偏移集合包括所述再前一个解码帧的最后 k 个样本以及
所述前一个解码帧的除最后 k 个样本之外的所有样本,其中, $k < N/2$ 。

8. 根据权利要求7所述的方法,其中,基于要用前面的帧估计的帧的自相似性的最大化
来设置 k 。

9. 根据权利要求7或8所述的方法,其中, k 取决于 N 。

10. 根据权利要求3所述的方法,

其中,所述第一子集的估计进一步基于与在所述分组序列中紧挨在与所述前一个解码
帧相关联的分组前面的接收分组相关联的再前一个解码帧,

其中,包括 $N/4$ 个加窗时域混叠样本的第一子集是所述中间帧的第一半的第一半,包括
 $N/4$ 个加窗时域混叠样本的第三子集是所述中间帧的第二半的第一半,

其中,对于 n 等于 $0, 1, \dots, k$,所述第一子集的样本编号 n 被估计为所述再前一个解码帧
的样本编号 $N/2-1+n-k$ 的加窗版本减去所述前一个解码帧的样本编号 $N/2-1-n-k$ 的加窗版
本,并且对于 n 等于 $k+1, \dots, N/4-1$, $N/4-1$,所述第一子集的样本编号 n 被估计为所述前一个
解码帧的样本编号 $n-k-1$ 的加窗版本减去所述前一个解码帧的样本编号 $N/2-1-n-k$ 的加窗
版本,并且

其中,对于 n 等于 $0, 1, \dots, k$,第三子集的样本编号 n 被估计为所述前一个解码帧的样本
编号 $N/2-1+n-k$ 的加窗版本减去所述估计解码帧的样本编号 $N/2-1-n-k$ 的加窗版本,并且其
中,对于 n 等于 $k+1, \dots, N/4-1$,第三子集的样本编号 n 被估计为所述估计解码帧的样本编号
 $n-k-1$ 的加窗版本加上所述估计解码帧的样本编号 $N/2-1-n-k$ 的加窗版本,其中, $k \leq N/4-1$ 。

11. 一种用于掩盖在基于修正离散余弦变换MDCT的音频解码器中将被解码的数据分组
中的错误的解码系统,所述基于MDCT的音频解码器被布置为将分组序列解码为解码帧序
列,所述系统包括:

接收器部分,被配置为从被布置为对音频信号进行编码的基于MDCT的音频编码器接收
包括与所述音频信号的 N 个加窗时域样本相关联的 $N/2$ 个MDCT系数的分组;

错误检测部分,被配置为在所述分组包括一个或多个错误时,将所述分组识别为错误
分组;

错误掩盖部分,被配置为:

估计第一子集,所述第一子集包括中间帧的第一半的 $N/4$ 个加窗时域混叠样本,所述中
间帧包括与所述错误分组相关联的 N 个加窗时域混叠样本,所述估计第一子集基于所述第
一子集的加窗时域混叠样本和所述音频信号的所述 N 个加窗时域样本中的加窗时域样本之
间的关系;并且

估计第二子集,所述第二子集包括所述中间帧的第一半的剩余 $N/4$ 个加窗时域混叠样
本,所述估计第二子集基于所述第二子集的加窗时域混叠样本和所述第一子集的加窗时域
混叠样本之间的对称关系,以及

重叠相加部分,被配置为通过将所述中间帧的第一半与前一个中间帧的第二半相加来
产生与所述错误分组相关联的估计解码帧,所述前一个中间帧与在所述分组序列中紧挨在
所述错误分组前面的接收分组相关联。

12.一种包括计算机可读介质的计算机程序产品,所述计算机可读介质具有用于执行权利要求1-10中任一项的方法的指令。

13.一种用于掩盖在基于修正离散余弦变换MDCT的音频解码器中将被解码的数据分组中的错误的装置,所述基于MDCT的音频解码器被布置为将分组序列解码为解码帧序列,所述装置包括:

一个或多个处理器,

一个或多个非暂时性的计算机可读存储介质,存储有指令,所述指令在被所述一个或多个处理器执行时使得所述装置执行根据权利要求1-10中任一项所述的方法。

14.一种非暂时性的计算机可读存储介质,存储有指令,所述指令在被一个或多个处理器执行时使得实现根据权利要求1-10中任一项所述的方法。

MDCT域错误掩盖

[0001] 本申请是申请号为201580067028.X、申请日为2015年12月8日、发明名称为“MDCT域错误掩盖”的发明专利申请的分案申请。

技术领域

[0002] 本文公开的发明总体上涉及音频信号的编码和解码,并且特别地,涉及用于掩盖错误的方法和装置。

背景技术

[0003] 修正离散余弦变换(MDCT)和对应的逆修正离散变换(IMDCT)被用于例如音频译码和解码技术,诸如MPEG-2和MPEG-4音频层、高级音频编码、MPEG-4HE-AAC、MPEG-D USAC、杜比数字(Plus)以及其他专有格式。

[0004] 在这种技术的应用中,在分组在解码系统中被接收之前或之后,由于与音频信号的变换相关的分组丢失或分组错误,有时会出现错误。这样的错误例如包括分组丢失或分组失真,并且可能导致解码音频信号的听得见的失真。

[0005] 因此已经提供了用于在分组中出现错误的情况下掩盖错误的方法。错误掩盖法一般分为估计掩盖法和非估计掩盖法,在估计掩盖法中,错误帧被估计替代,非估计掩盖法例如使用错误帧的静默、帧重复或噪声代替。

[0006] 估计掩盖法包括在频域中使用估计的方法(诸如美国专利No.8620644中公开的方法)以及在时域中使用估计的方法(诸如国际专利公开No.WO/2014/052746中公开的方法)。

[0007] 用于错误掩盖的所有技术都面临与掩盖的质量和所需估计的复杂度之间的权衡相关的问题。因此,存在对于错误掩盖的其他方法的需求。

附图说明

[0008] 现在将参照附图来描述示例性实施例,其中:

[0009] 图1A和图1B分别以举例的方式描绘MDCT和IMDCT的概括框图,

[0010] 图2是第一解码系统的概括框图,

[0011] 图3是第二解码系统的概括框图,以及

[0012] 图4是第三解码系统的概括框图。

[0013] 所有图都是示意性的,并且一般仅描绘阐明本公开所必需的部分,而其他部分则可以被省略或仅被建议。除非另有指示,否则相似的附图标记指代不同图中的相同部分。

具体实施方式

[0014] 鉴于以上,一个目的是提供旨在以不显著的复杂度提供期望的错误掩盖的解码器系统和相关方法。

[0015] I.概述——第一方面

[0016] 根据第一方面,示例性实施例提出了用于解码的解码方法、解码系统以及计算机

程序产品。所提出的方法、解码系统以及计算机程序产品一般可以具有相同的特征和优点。

[0017] 根据示例性实施例,提供了一种用于掩盖在基于MDCT的音频解码器中将被解码的数据分组中的错误的方法,该音频解码器被布置为将分组序列解码为解码帧序列。该方法包括:从被布置为对音频信号进行编码的基于MDCT的音频编码器接收包括MDCT系数集合的分组,该MDCT系数集合与包括音频信号的时域样本的帧相关联;并且在接收分组(received packet)包括一个或多个错误时,将接收分组识别为错误分组。该方法还包括产生替代错误分组的MDCT系数集合的估计MDCT系数,估计MDCT系数基于与在分组序列中紧挨在错误分组前面的接收分组相关联的对应MDCT系数。该方法还包括:将估计MDCT系数中的MDCT系数的第一子集的符号指定为等于在分组序列中紧挨在错误分组前面的接收分组的对应MDCT系数的对应符号,其中,第一子集包括与分组的音调状谱区间(tonal-like spectral bin)相关联的这些MDCT系数;并且随机地指定估计MDCT系数中的MDCT系数的第二子集的符号,其中第二子集包括与分组的噪声状谱区间(noise-like spectral bin)相关联的这些MDCT系数;基于所述分组的选定符号和估计MDCT系数来产生掩盖分组;并且用掩盖分组来替代错误分组。

[0018] 如本文所使用的,“错误分组”表示包括以某种方式相对于音频信号的正确样本的正确MDCT的MDCT系数不同的MDCT系数。这可能意味着,整个分组的一部分在分组序列中丢失,或者整个分组的一部分包括失真。

[0019] 分组的音调状谱区间和噪声状谱区间的识别可以使用任何合适的方法来执行。音调状谱区间和噪声状谱区间的识别的次序是任意的,并且可以例如取决于所用方法。

[0020] 要注意,术语“第一子集”和“第二子集”仅用于在文本中将这两个子集相互区分开,而不是用于指示与这两个不同子集相关的处理次序。指定的执行次序是任意的。可以首先对用于第一子集的MDCT系数执行指定,最后对用于第二子集的MDCT系数执行指定,或者倒过来。此外,在一些示例性实施例中,可以不对MDCT系数执行指定,以使得与第一子集相关联的所有MDCT系数都被连续地指定,并且与第二子集相关联的所有MDCT系数都被连续地指定。在一些示例性实施例中,可以首先对其中一个子集的一个或多个MDCT系数进行指定,然后对其他子集的一个或多个MDCT系数进行指定。此外,分组不一定具有与噪声状谱区间和音调状谱区间这两者相关联的MDCT系数。在一些示例性实施例中,分组可以具有与噪声状谱区间相关联的所有MDCT系数或者全都与音调状谱区间相关联,以使得其中一个子集是空的。最后,MDCT系数通常被识别为要么属于第一子集、要么属于第二子集。

[0021] 要注意,使估计基于与在分组序列中紧挨在错误分组前面的接收分组相关联的MDCT系数和MDCT系数符号不排除估计可以附加地基于与在分组序列中比紧挨在错误分组前面的分组更早的接收分组相关联的MDCT系数和MDCT系数符号。

[0022] 如本文所使用的,“产生估计MDCT系数”涉及将如下值指定给MDCT系数:这些值不一定是在错误分组中尚未存在任何错误时MDCT系数将会具有的值的最佳近似(approximation),而是实现期望的错误掩盖性能以使得解码音频信号的不想要的失真被避免或减少。

[0023] 如本文所使用的,“估计MDCT系数”涉及所估计的MDCT系数的绝对值。

[0024] 根据示例性实施例,该方法还包括:对于每个估计MDCT系数,基于与错误分组相关联的功率谱的近似的谱峰检测来确定MDCT系数是与音调状谱区间相关联、还是与噪声状谱

区间相关联,其中,近似的功率谱基于与在分组序列中紧挨在错误分组前面的接收分组相关联的功率谱。

[0025] 根据一些实施例,该方法还包括:对于每个估计MDCT系数,基于与分组相关联的元数据来确定MDCT系数是与音调状谱区间相关联、还是与噪声状谱区间相关联,其中,元数据是在包括分组序列和元数据的比特流中接收的。

[0026] 如本文所使用的,“元数据”涉及用于控制音频解码器处理的比特流参数。

[0027] 元数据可以在分组序列的分组中以及在包括分组序列和元数据的比特流中的分组的外部被发送。

[0028] 可以用于确定MDCT系数是与音调状谱区间相关联还是与噪声状谱区间相关联的元数据是用于基于音频内容类型来控制某个音频解码器处理的元数据。这种元数据的一个例子是与AC-4中使用的压扩工具相关的元数据。在一些实施例中,对于音调信号可以关掉压扩(compand)工具,因此,如果压扩关闭,则信号被假定为是音调。作为另一个例子,如果最长的MDCT被使用,则音频内容最可能是音调信号。

[0029] 根据一些实施例,估计MDCT系数被选为等于在分组序列中紧挨在错误分组前面的接收分组的对应MDCT系数。

[0030] 根据一些实施例,估计MDCT系数被选为等于在分组序列中紧挨在错误分组前面的接收分组的以缩放因子带分辨率被用能量缩放因子进行能量调整的对应MDCT系数。关于缩放因子带分辨率的详细描述,参照ETSI TS 103 190V1.1.1“Digital Audio Compression (AC-4) Standard”,2014-04,其内容通过引用并入本文。

[0031] 根据一些实施例,接收分组包括与音频信号的N个加窗时域样本相关联的N/2个MDCT系数,进一步包括:通过IMDCT从掩盖帧产生包括N个加窗时域混叠样本的中间帧;基于中间帧的加窗时域混叠样本之间的对称关系来修改中间帧的加窗时域混叠样本。

[0032] 如本文所使用的,“N”是偶数。

[0033] 如本文所使用的,“包括N个加窗时域混叠样本的中间帧”表示由从编码器接收的MDCT系数在解码器系统中的IMDCT得到的样本帧。在一些示例性实施例中,中间帧是在解码系统中执行重叠相加以便生成解码帧序列中的解码帧之前的帧。

[0034] 根据一些实施例,修改使用包括N个加窗时域混叠样本的中间帧的第一半的第一半和包括N个加窗时域混叠样本的中间帧的第一半的第二半之间的对称关系、以及包括N个加窗时域混叠样本的中间帧的第二半的第一半和包括N个加窗时域混叠样本的中间帧的第二半的第二半之间的对称关系。

[0035] 如本文所使用的,“中间帧的第一半”表示中间帧的头N/2个样本。如果中间帧的样本被连续地编号为0至N-1,则第一半将是样本0至N/2-1。此外,“中间帧的第二半”表示中间帧的最后N/2个样本。如果中间帧的样本被连续地编号为0至N-1,则第二半将是样本N/2至N-1。

[0036] 如本文所使用的,“中间帧的第一半的第一半”表示包括中间帧的第一半的头N/4个样本的子集,“中间帧的第一半的第二半”表示包括中间帧的第一半的最后N/4个样本的子集,“中间帧的第二半的第一半”表示包括中间帧的第二半的头N/4个样本的子集,“中间帧的第二半的第二半”表示包括中间帧的第二半的最后N/4个样本的子集。

[0037] 根据一些实施例,接收分组包括与音频信号的N个加窗时域样本相关联的N/2个

MDCT系数,进一步包括:通过IMDCT从掩盖帧产生包括N个加窗时域混叠样本的中间帧;基于中间帧的加窗时域混叠样本和音频信号的N个时域样本的加窗时域样本之间的关系来修改中间帧的加窗时域混叠样本。

[0038] 示例性实施例提供与在分组序列中紧挨在错误分组前面的接收分组相关联的前一个解码帧可以被用作第一子集的加窗时域混叠样本和音频信号的N个加窗时域样本中的加窗时域样本之间的关系的近似。然后可以使用该关系来修改所产生的中间帧以便增强错误掩盖性能。

[0039] 根据示例性实施例,提供了一种用于掩盖在基于修正离散余弦变换MDCT的音频解码器中将被解码的数据分组中的错误的解码系统,所述基于MDCT的音频解码器被布置为将分组序列解码为解码帧序列,所述系统包括:接收器部分,被配置为从被布置为对音频信号进行编码的基于MDCT的音频编码器接收包括MDCT系数集合的分组,所述MDCT系数集合与包括所述音频信号的时域样本的帧相关联;错误检测部分,被配置为在所述接收分组包括一个或多个错误时,将所述接收分组识别为错误分组;以及错误掩盖部分,被配置为:产生用于替代所述错误分组的MDCT系数集合的估计MDCT系数,所述估计MDCT系数基于与在所述分组序列中紧挨在所述错误分组前面的接收分组相关联的对应MDCT系数;将所述估计MDCT系数中的MDCT系数的第一子集的符号指定为等于在所述分组序列中紧挨在所述错误分组前面的接收分组的对应MDCT系数的对应符号,其中,所述第一子集包括与所述分组的音调状谱区间相关联的MDCT系数;随机地指定所述估计MDCT系数中的MDCT系数的第二子集的符号,其中,所述第二子集包括与所述分组的噪声状谱区间相关联的MDCT系数;基于所述分组的估计MDCT系数和选定符号来产生掩盖分组;以及用所述掩盖分组来替代所述错误分组。

[0040] II.概述——第二方面

[0041] 根据第二方面,示例性实施例提出了用于解码的解码方法、解码系统以及计算机程序产品。所提出的方法、解码系统以及计算机程序产品一般可以具有相同的特征和优点。

[0042] 根据示例性实施例,提供了一种用于掩盖在基于MDCT的音频解码器中将被解码的数据分组中的错误的方法,该音频解码器被布置为将分组序列解码为解码帧序列。该方法包括:从被布置为对音频信号进行编码的基于MDCT的音频编码器接收包括与音频信号的N个加窗时域样本相关联的N/2个MDCT系数的分组;并且在该分组包括一个或多个错误时,将该分组识别为错误分组。该方法还包括:估计第一子集,第一子集包括中间帧的第一半的N/4个加窗时域混叠样本,中间帧包括与错误分组相关联的N个加窗时域混叠样本,该估计基于第一子集的加窗时域混叠样本和音频信号的N个加窗时域样本中的加窗时域样本之间的关系;并且估计第二子集,第二子集包括中间帧的第一半的剩余N/4个加窗时域混叠样本,该估计基于第二子集的加窗时域混叠样本和第一子集的加窗时域混叠样本之间的对称关系。

[0043] 如本文所使用的,“N”是偶数。

[0044] 如本文所使用的,“错误分组”表示包括以某种方式相对于音频信号的正确样本的正确MDCT的MDCT系数不同的MDCT系数。这可能意味着,整个分组的一部分在分组序列中丢失,或者整个分组的一部分包括失真。

[0045] 如本文所使用的,“包括N个加窗时域混叠样本的中间帧”表示由从编码器接收的MDCT系数在解码器系统中的逆IMDCT得到的样本帧。因此,中间帧是在解码系统中执行重叠

相加以便生成解码帧序列中的解码帧之前的加窗时域混叠样本的帧。

[0046] 如本文所使用的,“中间帧的第一半”表示中间帧的头 $N/2$ 个样本。如果中间帧的样本被连续地编号为0至 $N-1$,则第一半将是样本0至 $N/2-1$ 。

[0047] 如本文所使用的,“包括 $N/4$ 个加窗时域混叠样本的第一子集”表示包括中间帧的第一半的 $N/4$ 个样本的子集,这 $N/4$ 个样本无需是中间帧的第一半中的连续样本,但是应被选为使得不生成关于来自第二子集的样本和第一子集的样本之间的对称关系的信息的冗余信息。

[0048] 如本文所使用的,“估计第一子集”和“估计第二子集”涉及将如下值指定给第一子集和第二子集的加窗时域混叠样本:这些值不一定是在错误分组中尚未存在任何错误时它们将会具有的值的最佳近似,但是实现期望的错误掩盖性能以使得解码音频信号的不想要的失真被避免或减少。

[0049] 根据示例性实施例,第一子集的估计基于与在分组序列中紧挨在错误分组前面的接收分组相关联的前一个解码帧。

[0050] 要注意,使估计基于与在分组序列中紧挨在错误分组前面的接收分组相关联的前一个解码帧不排除估计可以附加地基于与在分组序列中比紧挨在错误分组前面的分组更早的接收分组相关联的更早的解码帧。

[0051] 基于前一个解码帧对第一子集的估计在示例性实施例中可以与包括 $N/4$ 个加窗时域混叠样本(这 $N/4$ 个加窗时域混叠样本是中间帧的第一半的第一半)的第一子集相组合,其中,对于 n 等于 $0, 1, \dots, N/4-1$,第一子集的样本编号 n 被估计为前一个解码帧的样本编号 n 的加窗版本减去前一个解码帧的样本编号 $N/2-1-n$ 的加窗版本。

[0052] 示例性实施例提供:第一子集的加窗时域混叠样本和音频信号的 N 个加窗时域样本中的加窗时域样本之间的关系可以通过使用与错误分组相关联的 N 个加窗时域样本和与在分组序列中紧挨在错误分组前面的接收分组相关联的先前的 N 个加窗时域样本的重叠性质来重新制定。因此,第一子集的加窗时域混叠样本和音频信号的先前的 N 个加窗时域样本之间的关系被导出。示例性实施例还提供:音频信号的先前的 N 个加窗时域样本中的加窗时域样本可以用前一个解码帧的样本的加窗版本来近似。

[0053] 基于前一个解码帧估计第一子集、产生估计解码帧、估计第三子集以及估计第四子集在示例性实施例中可以与包括 $N/4$ 个加窗时域混叠样本(这 $N/4$ 个加窗时域混叠样本是中间帧的第一半的第一半)的第一子集以及包括 $N/4$ 个加窗时域混叠样本(这 $N/4$ 个加窗时域混叠样本是中间帧的第二半的第一半)的第三子集相组合,并且其中,对于 n 等于 $0, 1, \dots, N/4-1$,第一子集的样本编号 n 被估计为前一个解码帧的样本编号 n 的加窗版本减去前一个解码帧的样本编号 $N/2-1-n$ 的加窗版本,并且其中,对于 n 等于 $0, 1, \dots, N/4-1$,第三子集的样本编号 n 被估计为估计解码帧的样本编号 n 的加窗版本加上估计解码帧的样本编号 $N/2-1-n$ 的加窗版本。

[0054] 要注意,使估计基于与错误分组相关联的估计解码帧不排除估计可以附加地基于与在分组序列中比错误分组更早的接收分组相关联的更早的解码帧。

[0055] 示例性实施例提供:音频信号的先前的 N 个加窗时域样本中的加窗时域样本可以用前一个解码帧和估计解码帧的样本的加窗版本来近似。

[0056] 在一些示例性实施例中,第一子集的估计基于偏移集合,该偏移集合包括与在分

组序列中紧挨在与前一个解码帧相关联的分组前面的接收分组相关联的再前一个解码帧以及在与在分组序列中紧挨在错误分组前面的接收分组相关联的前一个解码帧的 $N/2$ 个样本,该偏移集合包括所述再前一个解码帧的最后 k 个样本以及所述前一个解码帧的除了最后 k 个样本之外的所有样本,其中, $k < N/2$ 。在本示例性实施例中, k 可以基于要用前面的帧估计的帧的自相似性的最大化来设置,并且 k 可以例如取决于 N 。

[0057] 代替仅使用前一个解码帧的 $N/2$ 个样本,将前一个解码帧的 $N-k$ 个样本与来自所述再前一个解码帧的 k 个样本一起使用。更具体地说,使用所述再前一个解码帧的最后 k 个样本以及所述前一个解码帧的除了最后 k 个样本之外的所有样本。这要求 $k < N/2$ 。

[0058] 基于前一个解码帧估计第一子集、产生估计解码帧、估计第三子集以及估计第四子集在示例性实施例中可以与第一子集的估计相组合,第一子集的估计进一步基于与在分组序列中紧挨在与前一个解码帧相关联的分组前面的接收分组相关联的再前一个解码帧,第一子集包括 $N/4$ 个加窗时域混叠样本,这 $N/4$ 个加窗时域混叠样本是中间帧的第一半的第一半,第三子集包括 $N/4$ 加窗时域混叠样本,这 $N/4$ 个加窗时域混叠样本是中间帧的第二半的第一半,对于 n 等于 $0, 1, \dots, k$,第一子集的样本编号 n 被估计为所述再前一个解码帧的样本编号 $N/2-1+n-k$ 的加窗版本减去所述前一个解码帧的样本编号 $N/2-1-n-k$ 的加窗版本,并且对于 n 等于 $k+1, \dots, N/4-1$,第一子集的样本编号 n 被估计为所述前一个解码帧的样本编号 $n-k-1$ 的加窗版本减去所述前一个解码帧的样本编号 $N/2-1-n-k$ 的加窗版本,并且 n 等于 $0, 1, \dots, k$,第三子集的样本编号 n 被估计为所述前一个解码帧的样本编号 $N/2-1+n-k$ 的加窗版本减去估计解码帧的样本编号 $N/2-1-n-k$ 的加窗版本,并且其中,对于 n 等于 $k+1, \dots, N/4-1$,第三子集的样本编号 n 被估计为估计解码帧的样本编号 $n-k-1$ 的加窗版本加上估计解码帧的样本编号 $N/2-1-n-k$ 的加窗版本,其中, $k \leq N/4-1$ 。

[0059] 在示例性实施例中,提供了一种用于掩盖在基于MDCT的音频解码器中将被解码的数据分组中的错误的解码系统,该音频解码器被布置为将分组序列解码为解码帧序列,该系统包括:接收器部分,被配置为从被布置为对音频信号进行编码的基于MDCT的音频编码器接收包括与音频信号的 N 个加窗时域样本相关联的 $N/2$ 个MDCT系数的分组;错误检测部分,被配置为在该分组包括一个或多个错误时,将该分组识别为错误分组;以及错误掩盖部分,被配置为:估计第一子集,第一子集包括中间帧的第一半的 $N/4$ 个加窗时域混叠样本,中间帧包括与错误分组相关联的 N 个加窗时域混叠样本,该估计基于第一子集的加窗时域混叠样本和音频信号的 N 个加窗时域样本中的加窗时域样本之间的关系;并且估计第二子集,第二子集包括中间帧的第一半的剩余 $N/4$ 个加窗时域混叠样本,该估计基于第二子集的加窗时域混叠样本和第一子集的加窗时域混叠样本之间的对称关系。

[0060] III. 概述——第三方面

[0061] 根据第三方面,示例性实施例提出了用于解码的解码方法、解码系统以及计算机程序产品。所提出的方法、解码系统以及计算机程序产品一般可以具有相同的特征和优点。

[0062] 在一些示例性实施例中,提供了一种用于掩盖在基于MDCT的音频解码器中将被解码的数据分组中的错误的方法,该基于MDCT的音频解码器被布置为将分组序列解码为解码帧序列。该方法包括:从被布置为对音频信号进行编码的基于MDCT的音频编码器接收包括与音频信号的 N 个加窗时域样本相关联的 $N/2$ 个MDCT系数的分组;并且在该分组包括一个或多个错误时,将该分组识别为错误分组。该方法还包括将包括与错误分组相关联的 $N/2$ 个样

本的解码帧估计为等于前一个中间帧的第二半,该前一个中间帧包括与在分组序列中紧挨在错误分组前面的接收分组相关联的N个非加窗时域样本。

[0063] 如本文所使用的,“N”是偶数。

[0064] 如本文所使用的,“错误分组”表示包括以某种方式相对于音频信号的正确样本的正确MDCT的MDCT系数不同的MDCT系数。这可能意味着,整个分组的一部分在分组序列中缺失,或者整个分组的一部分包括失真。

[0065] 如本文所使用的,“估计解码帧”涉及将如下值指定给解码帧的样本:这些值不一定是在错误分组中尚未存在任何错误时这些样本将会具有的值的最佳近似,但是实现期望的错误掩盖性能,以使得解码音频信号的不想要的失真被避免或减少。

[0066] 如本文所使用的,“前一个中间帧的第二半”表示前一个中间帧的最后N/2个样本。如果中间帧的样本被连续地编号为0至N-1,则第二半将是样本N/2至N-1。

[0067] 在一些示例性实施例中,提供了将包括与在分组序列中紧挨在错误分组后面的接收分组相关联的N/2个样本的后一个解码帧估计为等于后一个中间帧的第一半,该后一个中间帧包括与在分组序列中紧挨在错误分组后面的接收分组相关联的非加窗时域样本。

[0068] 在一些示例性实施例中,提供了一种用于掩盖在基于MDCT的音频解码器中将被解码的数据分组中的错误的解码系统,该基于MDCT的音频解码器被布置为将分组序列解码为解码帧序列,该系统包括:接收器部分,被配置为从被布置为对音频信号进行编码的基于MDCT的音频编码器接收包括与音频信号的N个加窗时域样本相关联的N/2个MDCT系数的分组;错误检测部分,被配置为在该分组包括一个或多个错误时,将该分组识别为错误分组;错误掩盖部分,被配置为:将包括与错误分组相关联的N/2个样本的解码帧估计为等于前一个中间帧的第二半,前一个中间帧包括与在分组序列中紧挨在错误分组前面的接收分组相关联的非加窗时域样本。

[0069] 在一些示例性实施例中,所述方法还包括:确定可用的复杂度资源,并且基于可用的复杂度资源来确定适用于掩盖错误的方法。

[0070] IV. 示例性实施例

[0071] 图1A和图1B以举例的方式描绘示例性实施例可以分别与其一起实现的MDCT和逆变换。在音频编码/解码系统中,音频信号通常在编码器端被采样并且被划分为帧101-105的序列,其中,该序列中的每各帧对应于相应的时间间隔 $t-2$ 、 $t-1$ 、 t 、 $t+1$ 、 $t+2$ 。帧101-105各自包括N/2个样本,其中,取决于编码器类型和所选择的时间频率分辨率,N可以为2048、1920、1536等。代替将MDCT应用于帧101-105,将MDCT应用于两个相邻帧的组合。因此,MDCT使用重叠,并且是所谓的重叠变换的例子。从帧101-105的序列(每个帧包括音频信号的N/2个时域样本),按连续的次序将帧两个和两个地重叠地组合,以使得例如帧101-105序列的第一帧101和第二帧102被组合为第一组合帧110,第二帧102和第三帧103被组合为第二组合帧111,等等,这意味着第一组合帧110和第二组合帧111具有重叠,因为它们两个都包括第二帧102。为了使顺序帧之间的转变平滑,将窗口函数 $w[n]$ ($n=0, \dots, N-1$) 应用于帧序列的两个帧的每个组合以产生N个加窗时域样本的组合帧110-113。如图1A所描绘的,分别对应于时间间隔 $t-2$ 和 $t-1$ 的第一帧101和第二帧102被组合,并且加窗函数被应用于该组合以产生包括N个加窗时域样本 $x_n^{(t-2)}$ ($n=0, \dots, N-1$) 的第一组合帧110,对应于时间间隔 $t-$

1和t的第二帧102和第三帧103被组合,并且加窗函数被应用于该组合以产生包括N个加窗时域样本 $x_n^{(t-1)}$ ($n=0, \dots, N-1$) 的第二组合帧111,对应于时间间隔t和t+1的第三帧103和第四帧104被组合,并且加窗函数被应用于该组合以产生包括N个加窗时域样本 $x_n^{(t)}$ ($n=0, \dots, N-1$) 的第三组合帧112,并且对应于时间间隔t+1和t+2的第四帧104和第五帧105被组合,并且加窗函数被应用于该组合以产生包括N个加窗时域样本 $x_n^{(t+1)}$ ($n=0, \dots, N-1$) 的第四组合帧113。

[0072] MDCT然后被应用于组合帧110-113,得到分组120-123的序列,每个分组包括N/2个MDCT系数。如图1A所描绘的,MDCT被应用于第一组合帧110以产生包括N/2个MDCT系数 $c_k^{(t-2)}$ ($k=0, \dots, N/2-1$) 的第一分组120,MDCT被应用于第二组合帧111以产生包括N/2个MDCT系数 $c_k^{(t-1)}$ ($k=0, \dots, N/2-1$) 的第二分组121,MDCT被应用于第三组合帧112以产生包括N/2个MDCT系数 $c_k^{(t)}$ ($k=0, \dots, N/2-1$) 的第三分组122,MDCT被应用于第四组合帧113以产生包括N/2个MDCT系数 $c_k^{(t+1)}$ ($k=0, \dots, N/2-1$) 的第四分组123。

[0073] 在解码器端,IMDCT被应用于分组120-123,每个分组包括N/2个MDCT系数,以产生包括N个时域混叠样本的中间帧130-133。如图1B所描绘的,IMDCT被应用于第一分组120以产生包括N个时域混叠样本 $\hat{x}_n^{(t-2)}$ ($n=0, \dots, N-1$) 的第一中间帧130,IMDCT被应用于第二分组121以产生包括N个时域混叠样本 $\hat{x}_n^{(t-1)}$ ($n=0, \dots, N-1$) 的第二中间帧131,IMDCT被应用于第三分组122以产生包括N个时域混叠样本 $\hat{x}_n^{(t)}$ ($n=0, \dots, N-1$) 的第三中间帧132,IMDCT被应用于第四分组123以产生包括N个时域混叠样本 $\hat{x}_n^{(t+1)}$ ($n=0, \dots, N-1$) 的第四中间帧133。

[0074] 为了产生解码样本的解码帧150-152,在考虑窗口函数w[n]的情况下对中间帧130-133执行重叠相加操作140-142。如图1B所描绘的,在第二中间帧131的第一半和第一中间帧130的第二半之间执行第一重叠相加操作140以产生包括对应于时间间隔t-1的N/2个解码样本的第一解码帧150,在第三中间帧132的第一半和第二中间帧131的第二半之间执行第二重叠相加操作141以产生包括对应于时间间隔t的N/2个解码样本的第二解码帧151,在第四中间帧133的第一半和第三中间帧132的第二半之间执行第三重叠相加操作142以产生包括对应于时间间隔t+1的N/2个解码样本的第三解码帧152。

[0075] 在包括MDCT系数的分组中可能出现错误,或者分组或分组的一部分可能丢失。除非错误被纠正或者丢失的分组被重构,否则这种错误或丢失可能以如下方式影响解码帧:解码音频信号受损,而使得信息丢失或者不想要的伪像出现在解码音频信号中。例如并且参照图1B,如果在解码器端在第三分组122中检测到错误,则第三中间帧132通常将受到错误的第三分组122的影响。在本文件中,包括错误的分组将被称为错误分组,并且与和错误

分组相同的时间间隔相对应的中间帧将被称为与错误分组相关联的中间帧或包括与错误分组相关联的N个时域混叠样本的中间帧。此外,当第三中间帧132在重叠相加操作141中被使用以生成第二解码帧151时,第二解码帧151通常将受到错误分组的影响。在本文件中,与和错误分组相同的时间间隔相对应的解码帧将被称为与错误分组相关联的解码帧。此外,当第二中间帧132在重叠相加操作142中也被使用以生成第三解码帧152时,第三解码帧152通常也将受到错误分组的影响。

[0076] 由于组合帧的重叠性质,可以根据方程1导出组合帧的与时间间隔t相关联的头N/2个样本和组合帧的与时间间隔t-1相关联的后N/2个样本之间的关系:

$$[0077] \quad x_n^{(t)} = x_{\frac{N}{2}+n}^{(t-1)}, \text{ for } n = 0, 1, \dots, \frac{N}{2} - 1 \quad (1)$$

[0078] 此外,通过在中间帧的第一半和前一个中间帧的第二半之间使用重叠相加来产生解码帧。因此,根据以下方程产生与时间间隔t相关联的解码帧:

$$[0079] \quad x_n^{(t)} = \hat{x}_{\frac{N}{2}+n}^{(t-1)} + \hat{x}_n^{(t)}, \text{ for } n = 0, 1, \dots, \frac{N}{2} - 1 \quad (2)$$

[0080] 中间帧的加窗时域样本之间的特殊性质可以用于估计受错误分组影响的中间帧。更具体地说,可以证明,每个中间帧拥有第一半和第二半中的加窗时域样本之间的奇对称和偶对称。对于时间间隔t,以下关系可以被证明:

$$[0081] \quad \left. \begin{aligned} \hat{x}_n^{(t)} &= -\hat{x}_{\frac{N}{2}-1-n}^{(t)} \\ \hat{x}_{\frac{N}{2}+n}^{(t)} &= \hat{x}_{N-1-n}^{(t)} \end{aligned} \right\} \text{ 对于 } n = 0, 1, \dots, \frac{N}{4} - 1 \quad (3)$$

[0082] 此外,可以证明,加窗时域混叠样本可以根据以下方程(参见V.Britanak等人在Signal Processing第89卷第7期(2009年7月)第1379-1394页上的“Fast computational structures for an efficient implementation of the complete TDAC analysis/synthesis MDCT/MDST filter banks”,其内容通过引用并入本文)就音频信号的原始加窗样本显式地导出:

$$[0083] \quad \left. \begin{aligned} \hat{x}_n^{(t)} &= x_n^{(t)} - x_{\frac{N}{2}-1-n}^{(t)} \\ \hat{x}_{\frac{N}{2}+n}^{(t)} &= x_{\frac{N}{2}+n}^{(t)} + x_{N-1-n}^{(t)} \end{aligned} \right\} \text{ 对于 } n = 0, 1, \dots, \frac{N}{4} - 1 \quad (4)$$

[0084] 在方程(4)中使用方程(1),推出以下关系:

$$[0085] \quad \left. \begin{aligned} \hat{x}_n^{(t)} &= x_{\frac{N}{2}+n}^{(t-1)} - x_{N-1-n}^{(t-1)} \\ \hat{x}_{\frac{N}{2}+n}^{(t)} &= x_n^{(t+1)} + x_{\frac{N}{2}-1-n}^{(t+1)} \end{aligned} \right\} \text{ 对于 } n = 0, 1, \dots, \frac{N}{4} - 1 \quad (5)$$

[0086] 在另一个近似中,受错误分组影响的解码帧可以根据以下方程使用非加窗时域混叠信号 \tilde{x}_n 的帧估计:

$$[0087] \quad \left. \begin{array}{l} \tilde{x}_{\frac{N}{2}+n}^{(t-1)} \rightarrow x_n^{(t)} \\ \tilde{x}_{\frac{N}{2}-1-n}^{(t-1)} \rightarrow x_{\frac{N}{2}-1-n}^{(t)} \end{array} \right\} \text{对于 } n = 0, 1, \dots, \frac{N}{4} - 1 \quad (6)$$

$$[0088] \quad \left. \begin{array}{l} \tilde{x}_n^{(t+1)} \rightarrow x_n^{(t+1)} \\ \tilde{x}_{\frac{N}{2}-1-n}^{(t+1)} \rightarrow x_{\frac{N}{2}-1-n}^{(t+1)} \end{array} \right\} \text{对于 } n = 0, 1, \dots, \frac{N}{4} - 1 \quad (7)$$

[0089] 在方程 (6) 和 (7) 中, 记号 $a \rightarrow b$ 指示变量 b 被指定值 a 。

[0090] 图2以举例的方式描绘了第一解码系统200的概括框图。解码系统200被布置为掩盖在基于MDCT的音频解码器中将被解码的数据分组中的错误, 该基于MDCT的音频解码器被布置为将分组序列解码为解码帧序列。

[0091] 该系统包括接收器部分201, 接收器部分201被配置为接收分组序列, 其中, 每个分组包括与包括音频信号的时域样本的帧相关联的MDCT系数集合。分组序列通常如关于图1A所描述的那样通过将MDCT应用于 N 个加窗时域样本的组合帧而产生。分组序列的每个分组包括 $N/2$ 个MDCT系数。

[0092] 解码系统200还包括错误检测部分(未示出), 错误检测部分被配置为因为接收分组包括一个或多个错误而识别接收分组是否是错误分组。在错误检测部分中检测错误的方式是任意的, 并且错误检测部分的位置也是任意的, 只要检测到需要错误掩盖的错误分组并且检测到的错误分组可以在解码系统200的错误掩盖中被识别即可。

[0093] 解码系统200还包括错误掩盖部分202, 错误掩盖部分202被配置为估计错误分组的MDCT系数、将符号指定给估计MDCT系数、产生掩盖分组并且在分组序列中用掩盖分组替代错误分组。掩盖分组被产生为具有错误分组的对应选定符号的估计MDCT系数。

[0094] 解码系统200还包括IMDCT部分203, IMDCT部分203用于对分组序列的每个分组应用IMDCT, 该分组序列包括替代分组序列中的错误分组的掩盖分组。IMDCT部分203的输出是 N 个加窗时域混叠样本的中间帧的序列。

[0095] 解码系统200还包括重叠相加部分204, 重叠相加部分204用于执行中间帧序列中的连续的中间帧的重叠部分之间的重叠相加操作以便产生 $N/2$ 个样本的解码帧。

[0096] 在一个实施例中, 估计MDCT系数基于与在分组序列中紧挨在错误分组前面的接收分组相关联的对应MDCT系数。在其他实施例中, 估计MDCT系数被选为等于在分组序列中紧挨在错误分组前面的接收分组的对应MDCT系数。此外, 估计MDCT系数中的MDCT系数的第一子集的符号被指定为等于在分组序列中紧挨在错误分组前面的接收分组的对应MDCT系数的对应符号。第一子集包括与分组的音调状谱区间相关联的那些MDCT系数。估计MDCT系数中的MDCT系数的第二子集的符号是任意指定的。第二子集包括与分组的噪声状谱区间相关联的那些MDCT系数。连同每个MDCT系数的符号一起, 错误掩盖部分202从接收部分201连续地接收分组序列的每个分组的MDCT系数。错误掩盖部分202还从接收部分接收错误帧的标识。当错误帧被接收到时, 错误掩盖部分202可以提取在分组序列中的错误分组前面接收的前一个分组的MDCT系数和对应符号, 产生错误分组的估计MDCT系数, 并且使用来自前一个分组的MDCT系数和符号一起来对符号进行指定。当系数和符号已经被估计和指定时, 产生

基于估计分组的估计MDCT系数和选定符号的掩盖分组,在接收部分201中错误掩盖部分用掩盖分组替代错误分组,并且将掩盖分组从接收部分201转发给MDCT部分203。

[0097] 要注意,当关于估计以及将符号指定给每个估计MDCT系数而提到估计MDCT系数时,这隐含地指的是估计MDCT系数的绝对值。即使对于MDCT系数的符号指定被公开是首先针对第一子集、第二针对第二子集,符号的指定也可以按相反的次序执行。因此,在示例性实施例中,可以首先对第二子集执行指定,最后对第一子集执行指定。事实上,可以按任何次序对MDCT系数执行指定。在示例性实施例中,可能不一定连续地对与音调状谱区间相关联的所有MDCT系数以及连续地对与噪声状谱区间相关联的所有MDCT系数执行指定。例如,可以首先对与第一子集相关联的MDCT系数中的一个或多个进行指定,然后对与第二子集相关联的MDCT系数中的一个或多个进行指定,然后对与第一子集相关联的MDCT系数中的一个或多个进行指定,等等。此外,分组不一定具有与噪声状谱区间和音调状谱区间这两者相关联的MDCT系数。相反,分组可以具有与噪声状谱区间相关联的所有MDCT系数或与音调状谱区间相关联的所有MDCT系数,以使得第一子集和第二子集的一个是空的。最后,MDCT系数通常被要么识别为属于第一子集、要么属于第二子集。

[0098] 就错误掩盖性能而言,相比于仅使用随机指定的估计或仅基于分组序列中先前接收的分组的MDCT系数的符号的估计,基于内容类型来估计MDCT系数的符号可以提供改善的效果。与噪声状谱区间相关的MDCT系数如果通过随机指定估计的话,可能是足够精确的,而就错误掩盖性能而言,与音调状谱区间相关的MDCT系数可以通过基于在分组序列中紧挨在错误分组前面的接收分组的对应MDCT系数的指定来提供改善的结果。此外,当基于与在分组序列中紧挨在错误分组前面的接收分组相关联的对应MDCT系数估计MDCT系数时,可以仅使用来自先前接收的分组的数据来实现错误掩盖。

[0099] 在一些现有技术中,包括估计所有MDCT系数的符号而不使用随机指定的更复杂的方法已经被使用。在其他现有技术中,附加的元数据已经被提供用于估计符号(这给该方法添加了更多的复杂度),并且需要改变从译码器到解码器的数据流。此外,这种元数据必须在错误分组后面的分组中传送,这使得在解码系统中可执行符号估计的时间延迟。

[0100] 通过将估计MDCT系数选为等于前一个分组的对应MDCT系数,可以使复杂度保持低,同时如果这与根据示例性实施例的基于内容类型对MDCT系数的符号的估计相组合的话,可以实现提供期望的错误掩盖性能的掩盖分组。

[0101] 在其他实施例中,在前一个分组的MDCT系数被选为错误分组的MDCT系数的估计之前,通过能量缩放因子以缩放因子带分辨率对这些MDCT系数进行能量调整。

[0102] 通过将估计MDCT系数选为等于前一个分组的通过能量缩放因子以缩放因子带分辨率进行能量调整的对应MDCT系数,可以增强通过掩盖分组实现的错误掩盖性能,同时可以仅略微提高复杂度。

[0103] 存在几种替代方式来确定分组序列中的分组(例如错误分组)的MDCT系数是与音调状谱区间相关联、还是与噪声状谱区间相关联。在一个例子中,该确定基于与错误分组相关联的功率谱的近似的谱峰检测,其中,经近似的功率谱基于与在分组序列中紧挨在错误分组前面的接收分组相关联的功率谱。在另一个例子中,使用MDCT子带谱平坦度量。如果MDCT子带谱平坦度的值高于某个阈值,则子带谱是平坦的,这暗示它是有噪声的。否则,谱是有峰的,这暗示它是音调的。MDCT子带平坦度被估计为MDCT系数的幅值的几何平均值和

算术平均值之间的比率。它表示信号的功率谱相对于平坦形状的偏差。该度量是逐个带地计算的,其中,术语“带”与MDCT系数集合有关,并且这些带的宽度是根据感知相关的缩放因子带分辨率的。关于谱平坦度度量的描述,参照N.Jayant和P.Noll在Englewood Cliffs, NJ:Prentice-Hall(1984)上的“Digital Coding of Waveforms, Principles and Applications to Speech and Video”。在其他例子中,确定是基于在分组中或者在包括分组序列和元数据的比特流中接收的元数据的。将被使用的元数据可以例如是用于基于音频内容类型控制某个音频解码器处理的元数据。例如在AC-4中,存在对于音调信号必须被关掉的压扩工具。因此,如果接收到指示压扩被关掉的元数据,则信号可以假定是音调的。此外,如果例如最长MDCT被使用,则音频内容非常有可能是音调信号。

[0104] 在一个实施例中,方程(3)的与错误帧相关联的中间帧的加窗时域混叠样本之间的对称关系被用于修改与错误帧相关联的中间帧的加窗时域混叠样本。当与时间间隔 t 相关联的错误帧已经被识别出时,在错误掩盖部分202中产生掩盖分组,并且掩盖分组替代错误帧。在IMDCT部分203中,对掩盖分组应用IMDCT,这产生与错误分组相关联的中间帧。所产生的与错误分组相关联的中间帧被从IMDCT部分203转发给错误掩盖部分202。错误掩盖部分202然后修改所产生的中间帧的加窗时域混叠样本,以使得方程(3)的关系被更好地满足。

[0105] 能在中间帧的加窗时域混叠样本之间得到证明的对称关系可以被用于修改中间帧的加窗时域混叠样本,以便增强错误掩盖性能。于是,可以实现错误掩盖性能的增强,同时可以仅略微提高复杂度。

[0106] 在其他实施例中,方程(5)的与错误帧相关联的中间帧的加窗时域混叠样本和原始数据样本之间的关系被用于修改与错误帧相关联的中间帧的加窗时域混叠样本。当与时间间隔 t 相关联的错误帧已经被识别出时,在错误掩盖部分202中产生掩盖分组,并且掩盖分组替代错误帧。在IMDCT部分203中,对掩盖分组应用IMDCT,这产生与错误分组相关联的中间帧。所产生的与错误分组相关联的中间帧被从IMDCT部分203转发给错误掩盖部分202。错误掩盖部分202然后修改所产生的中间帧的加窗时域混叠样本,以使得方程(5)的关系被更好地满足。例如,用在错误估计部分202中从重叠相加部分204接收的与时间间隔 $t-1$ 相关联的过去的解码帧来近似方程(5)的与和错误分组相关联的中间帧的第一半相关的第一关系的右手边。结果是和错误分组相关联的中间帧的第一半的替代性估计,该替代性估计可以被用于修改通过将IMDCT应用于在掩盖部分202中产生的掩盖分组而产生的与错误分组相关联的中间帧的第一半。此外,用与时间间隔 t 相关联的解码帧(即,基于与错误分组相关联的中间帧的经修改的第一半的解码帧)来近似方程(5)的与和错误分组相关联的中间帧的第二半相关的第二关系的右手边。与时间间隔 t 相关联的解码帧是在错误估计部分202中从重叠相加部分204接收的。结果是和错误分组相关联的中间帧的第二半的替代性估计,该替代性估计可以被用于修改通过将IMDCT应用于在掩盖部分202中产生的掩盖分组而产生的与错误分组相关联的中间帧的第二半。

[0107] 图3以举例的方式描绘了第二解码系统300的概括框图。解码系统300被布置为掩盖在基于MDCT的音频解码器中将被解码的数据分组中的掩盖,该基于MDCT的音频解码器被布置为将分组序列解码为解码帧序列。

[0108] 该系统包括接收器部分301,接收器部分301被配置为接收分组序列,其中,每个分

组包括与包括音频信号的时域样本的帧相关联的MDCT系数集合。分组序列通常如关于图1A所描述的那样通过将MDCT应用于N个加窗时域样本的组合帧而产生。分组序列的每个分组包括N/2个MDCT系数。

[0109] 解码系统300还包括错误检测部分(未示出),错误检测部分被配置为识别是否由于接收分组包括一个或多个错误从而接收分组是错误分组。错误检测部分中检测错误的方式是任意的,并且错误检测部分的位置也是任意的,只要检测到需要错误掩盖的错误分组并且检测到的错误分组可以在解码系统200的错误掩盖中被识别即可。

[0110] 解码系统300还包括错误掩盖部分302,错误掩盖部分302被配置为估计包括与错误分组相关联的N个加窗时域混叠样本的中间帧的加窗时域混叠样本。

[0111] 解码系统300还包括IMDCT部分303,IMDCT部分303用于对分组序列的每个分组应用IMDCT。IMDCT部分303的输出是N个加窗时域混叠样本的中间帧的序列。

[0112] 错误掩盖部分302还被配置为用估计中间帧替代包括与错误分组相关联的N个加窗时域混叠样本的中间帧。

[0113] 解码系统300还包括重叠相加部分304,重叠相加部分304用于执行中间帧序列中的连续的中间帧的重叠部分之间的重叠相加操作以便产生N/2个样本的解码帧。

[0114] 在实施例中,当在时间间隔t内识别出错误分组时,可以对与错误分组相关联的中间帧进行估计。使用与时间间隔t相关联的中间帧的加窗时域混叠样本和方程(5)的音频信号的原始加窗样本的项之间的关系以及方程(3)的对称关系来执行估计。对包括中间帧的第一半的头N/4个加窗时域混叠样本的第一子集进行估计,该中间帧包括与错误分组相关联(即,与时间间隔t相关联)的N个加窗时域混叠样本。通过方程(5)的第一关系来进行估计,其中,用前一个解码帧的样本来近似右手边的样本,其中,前一个解码帧与时间间隔t-1相关联。在错误估计部分302中从重叠相加部分304接收与时间间隔t-1相关联的解码帧。更具体地说,对于 $n=0,1,\dots,N/4-1$,第一子集的样本编号n被估计为前一个解码帧的样本编号n的加窗版本减去前一个解码帧的样本编号 $N/2-1-n$ 的加窗版本。通过方程(3)的对称关系,估计第二子集,第二子集包括中间帧的第一半的剩余(即,最后的)N/4个加窗时域混叠样本。在重叠相加部分304中通过将估计中间帧的第一半与和在分组序列中紧挨在错误分组前面的接收分组相关联(即,与时间间隔t-1相关联)的前一个中间帧的第二半相加来产生与错误分组相关联(即,与时间间隔t相关联)的估计解码帧。

[0115] 通过使用第二子集的加窗时域混叠样本和第一子集的加窗时域混叠样本之间的对称关系来估计第二子集,可以在保持所实现的错误掩盖性能的同时实现估计的复杂度降低。

[0116] 通过使用前一个解码帧作为第一子集的加窗时域混叠样本和音频信号的N个加窗时域样本中的加窗时域样本之间的关系的近似以便产生第一子集的估计,可以在实现期望的错误掩盖性能的同时实现估计的低复杂度。

[0117] 对第三子集进行估计,第三子集包括与错误分组相关联的中间帧的第二半的头N/4个加窗时域混叠样本。通过方程(5)的第二关系来进行估计,其中,用估计解码帧的样本来近似右手边的样本,其中,估计解码帧与错误分组相关联,即,与时间间隔t相关联。与时间间隔t相关联的估计解码帧是在错误估计部分302中从重叠相加部分304接收的。更具体地说,对于 $n=0,1,\dots,N/4-1$,第三子集的样本编号n被估计为估计解码帧的样本编号n的加窗

版本加上估计解码帧的样本编号 $N/2-1-n$ 的加窗版本。通过方程(3)的对称关系来估计第四子集,第四子集包括中间帧的第二半的剩余(即,最后的) $N/4$ 个加窗时域混叠样本。要注意,对于 $n=0,1,\dots,N/4-1$,第三子集的样本编号 n 是中间帧的样本编号 $N/2+n$,因为第三子集是中间帧的第二半的第一半。在重叠相加部分304中通过将时间与时间间隔 t 相关联的估计中间帧的第二半与后一个估计中间帧的第一半相加来产生与紧跟在错误分组后面的接收分组相关联(即,与时间间隔 $t+1$ 相关联)的后一个估计解码帧。

[0118] 在替代性实施例,第一子集的估计基于如下偏移集合,该偏移集合包括与时间间隔 $t-2$ 相关联的再前一个解码帧(未示出)以及与时间间隔 $t-1$ 相关联的前一个解码帧的 $N/2$ 个样本,并且第三子集的估计基于如下偏移集合,该偏移集合包括与时间间隔 $t-1$ 相关联的前一个解码帧以及与时间间隔 t 相关联的估计解码帧的 $N/2$ 个样本。偏移集合包括所述再前一个解码帧的最后 k 个样本以及所述前一个解码帧的除了最后 k 个样本之外的所有样本,其中, $k < N/2$ 。更具体地说,对于 $k \leq N/4-1$,对于 $n=0,1,\dots,k$,第一子集的样本编号 n 被估计为所述再前一个解码帧(未示出)的样本编号 $N/2-1+n-k$ 的加窗版本减去所述前一个解码帧的样本编号 $N/2-1+n-k$ 的加窗版本。对于 n 等于 $k+1,\dots,N/4-1$,第一子集的样本编号 n 被估计为所述前一个解码帧的样本编号 $n-k-1$ 的加窗版本减去所述前一个解码帧的样本编号 $N/2-1-n-k$ 的加窗版本。对于 n 等于 $0,1,\dots,k$,第三子集的样本编号 n 被估计为所述前一个解码帧的样本编号 $N/2-1+n-k$ 的加窗版本减去估计解码帧的样本编号 $N/2-1-n-k$ 的加窗版本。对于 $n=k+1,\dots,N/4-1$,第三子集的样本编号 n 被估计为估计解码帧的样本编号 $n-k-1$ 的加窗版本加上估计解码帧的样本编号 $N/2-1-n-k$ 的加窗版本。

[0119] k 的值可以被计算为使将用前面的帧估计的帧的自相似性最大化,或者它可以被预先计算以节省复杂度。此外, k 通常取决于 N 。

[0120] 关于只有前一个解码帧的样本的加窗版本被用于估计第一子集的加窗时域混叠样本时,错误掩盖性能可以得到改善。更具体地说,增强的错误掩盖性能可以是由在第一子集的加窗时域混叠样本的估计中使用若干样本的偏移或时间的偏移而引起的。

[0121] 图4以举例的方式描绘了第三解码系统400的概括框图。解码系统400被布置为掩盖在基于MDCT的音频解码器中将被解码的数据分组中的错误,该基于MDCT的音频解码器被布置为将分组序列解码为解码帧序列。

[0122] 该系统包括接收器部分401,接收器部分401被配置为接收分组序列,其中,每个分组包括与包括音频信号的时域样本的帧相关联的MDCT系数集合。分组序列通常如关于图1A所描述的那样通过将MDCT应用于 N 个加窗时域样本的组合帧而产生。分组序列的每个分组包括 $N/2$ 个MDCT系数。

[0123] 解码系统400还包括错误检测部分(未示出),错误检测部分被配置为识别是否由于接收分组包括一个或多个错误从而接收分组是错误分组。错误检测部分中检测错误的方式是任意的,并且错误检测部分的位置也是任意的,只要检测到需要错误掩盖的错误分组并且检测到的错误分组可以在解码系统400的错误掩盖中被识别即可。

[0124] 解码系统400还包括错误掩盖部分402,错误掩盖部分402被配置为估计包括与错误分组相关联的 $N/2$ 个样本的解码帧以产生估计解码帧。解码帧被估计为等于包括与在分组序列中紧挨在错误分组前面的接收分组相关联的 N 个非加窗时域样本的前一个中间帧的第二半。

[0125] 解码系统400还包括IMDCT部分403,IMDCT部分403用于对分组序列的每个分组应用IMDCT。IMDCT部分403的输出是N个加窗时域混叠样本的中间帧的序列。

[0126] 解码系统400还包括重叠相加部分404,重叠相加部分404用于执行中间帧序列中的连续的中间帧的重叠部分之间的重叠相加操作以便产生N/2个样本的解码帧。

[0127] 错误掩盖部分402还被配置为将包括与在分组序列中紧挨在错误分组后面的接收分组相关联的N/2个样本的后一个解码帧估计为等于包括与在分组序列中紧挨在错误分组后面的接收分组相关联的非加窗时域样本的后一个中间帧的第一半。错误掩盖部分402还被配置为用估计解码分组来替代来自重叠相加部分404的与错误分组相关联的解码帧并且用估计解码分组来替代来自重叠相加部分404的与错误分组相关联的后一个解码帧。

[0128] 解码系统400使用方程(6)和(7)的近似。

[0129] 用前一个中间帧的非加窗时域样本估计与错误分组相关联的样本的解码帧的样本可以提供用于提供错误掩盖的低复杂度方法。

[0130] 此外,可以提供确定可用复杂度资源的适应性方法,例如该方法连续地确定对于错误掩盖所允许的复杂度水平。例如,当错误分组被识别到时,确定可用的复杂度资源,并且根据所确定的可用资源来选择用于错误掩盖的方法。

[0131] V.等同、扩展、替代及其他

[0132] 在研究以上描述之后,对于本领域技术人员来说,本公开的其他实施例将变得清楚。即使本描述和附图公开了实施例和例子,本公开也不限于这些具体例子。在不脱离所附权利要求书限定的本公开的范围的情况下,可以做出许多修改和变化。权利要求书中出现的任何引用符号不被解释为限制它们的范围。

[0133] 另外,通过研究附图、本公开以及所附权利要求书,技术人员在实施本公开时可以理解和实行所公开的实施例的变型。在权利要求书中,词语“包括”不排除任何元件或步骤,并且不定冠词“一个”不排除复数。只有某些措施被记载在相互不同的从属权利要求中的事实并不表明这些措施的组合不能被有利地使用。

[0134] 以上所公开的装置和方法可以实现为软件、固件、硬件或它们的组合。在硬件实现中,以上描述中提到的功能单元之间的任务划分不一定对应于划分为物理单元的划分;相反,一个物理组件可以具有多个功能,并且一个任务可以由几个物理组件合作执行。某些组件或所有组件可以实现为数字信号处理器或微处理器执行的软件,或者可以实现为硬件或专用集成电路。这样的软件可以分布在计算机可读介质上,计算机可读介质可以包括计算机存储介质(或非暂时性介质)和通信介质(或暂时性介质)。软件可以分布在专门编程的装置上,这些装置在本文中一般可以被称为“模块”。模块的软件组件部分可以用任何计算机语言编写,并且可以是庞大代码库的一部分,或者可以在更多的离散代码部分中开发,诸如通常用面向对象的计算机语言。另外,模块可以分布在多个计算机平台、服务器、终端、移动装置等上。给定模块甚至可以被实现为使得所描述的功能由分开的处理器和/或计算机硬件平台执行。如本领域技术人员众所周知的,计算机存储介质这一术语包括按照用于存储信息(诸如计算机可读指令、数据结构、程序模块或其他数据)的任何方法或技术实现的易失性和非易失性、可移动和不可移动介质。计算机存储介质包括但不限于RAM、ROM、EEPROM、闪存或其他存储器技术、CD-ROM、数字多功能盘(DVD)或其他光学盘存储器、磁盒、磁带、磁盘存储器或其他磁性存储装置、或可以用于存储期望的信息并且可以被计算机访

问的任何其他介质。如本申请中所使用的,术语“部分”是指以下所有方面:(a) 仅硬件的电路实现(诸如仅在模拟和/或数字电路中的实现);(b) 电路和软件(和/或固件)的组合,诸如(如果适用的话):(i) 处理器(一个或多个)的组合或(ii) 处理器(一个或多个)/软件的部分(包括一起工作以使设备(诸如移动电话或服务器)执行各种功能的数字信号处理器(一个或多个))、软件和存储器(一个或多个));以及(c) 需要用于操作的软件或固件的电路,诸如微处理器(一个或多个)或微处理器(一个或多个)的一部分,即使该软件或固件不是物理存在的。此外,技术人员众所周知的是,通信介质通常包含计算机可读指令、数据结构、程序模块、或模块化数据信号(诸如载波或其他传输介质)中的其他数据,并且包括任何信息递送介质。

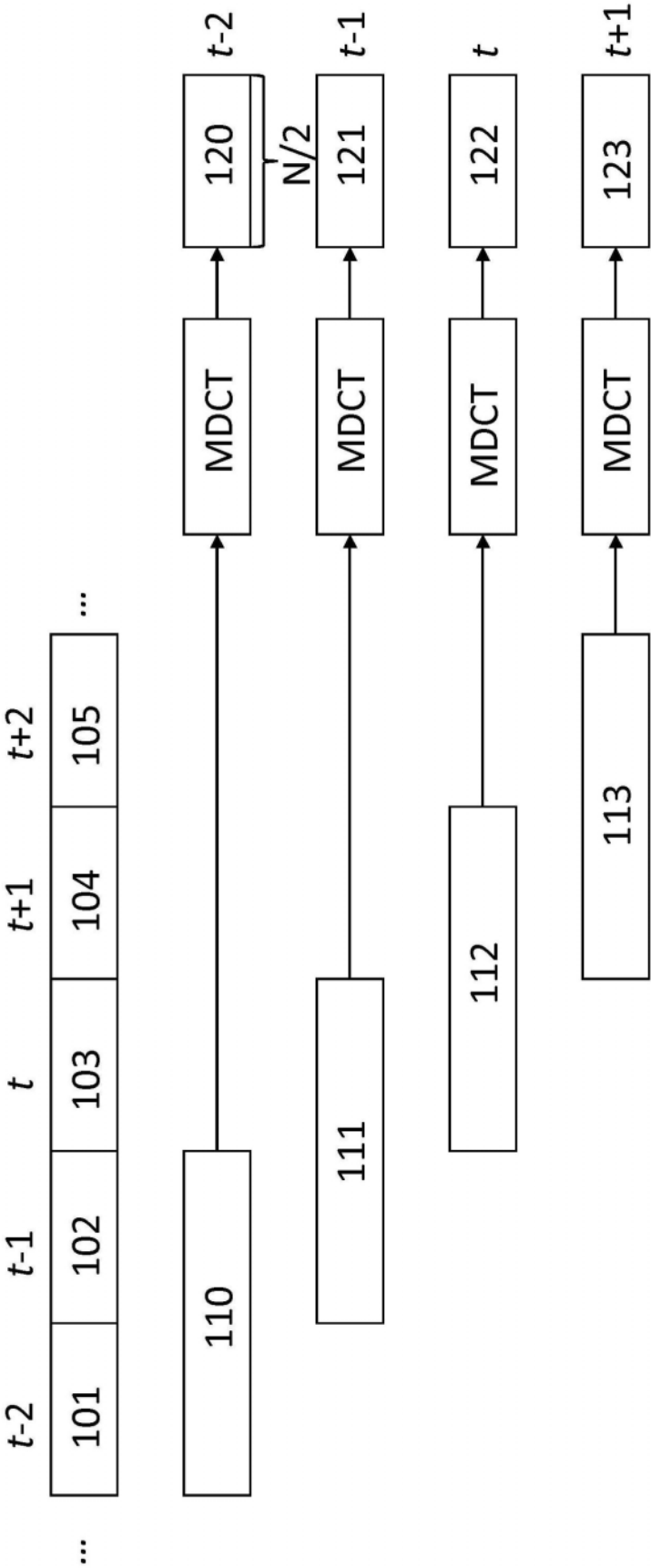


图1A

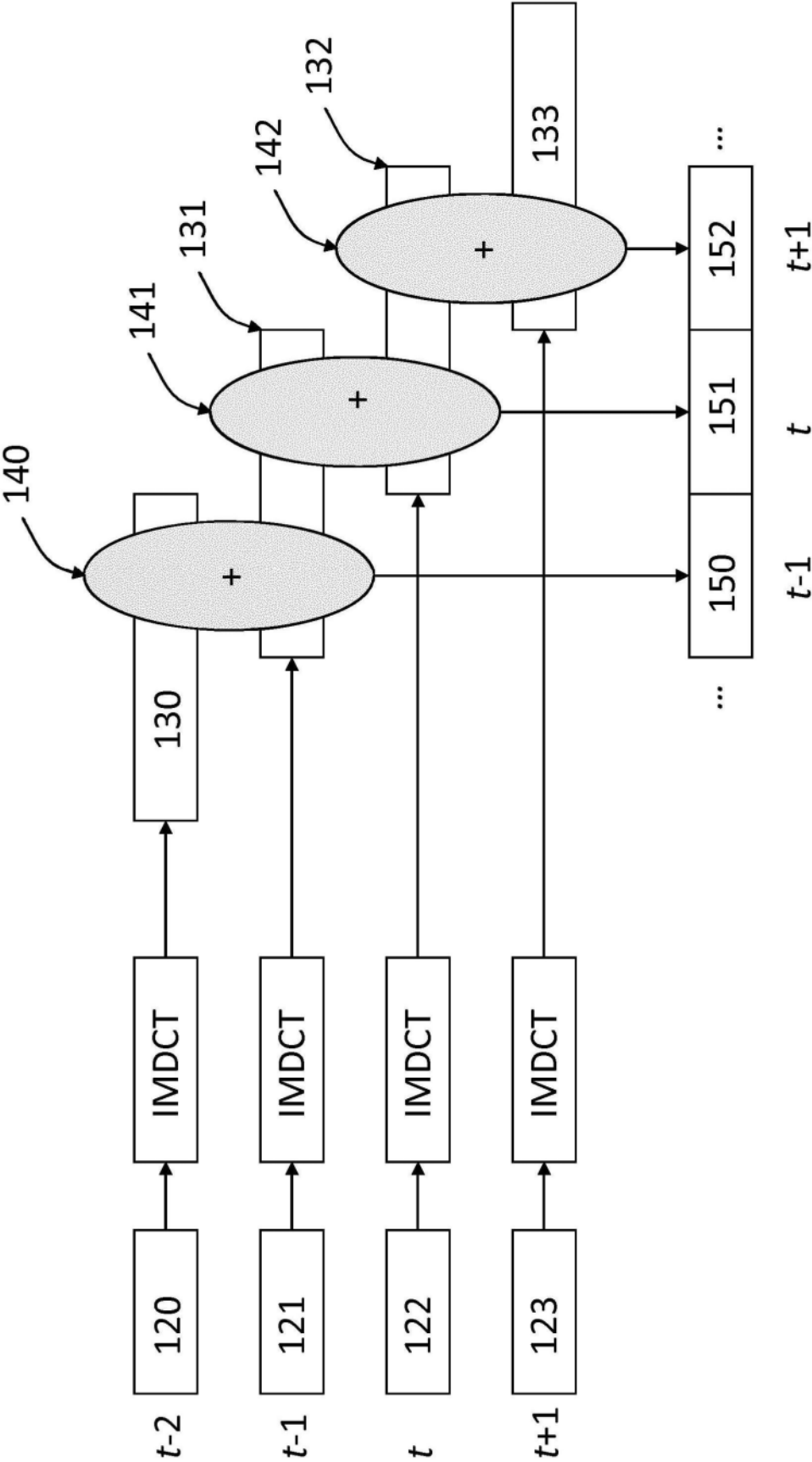


图1B

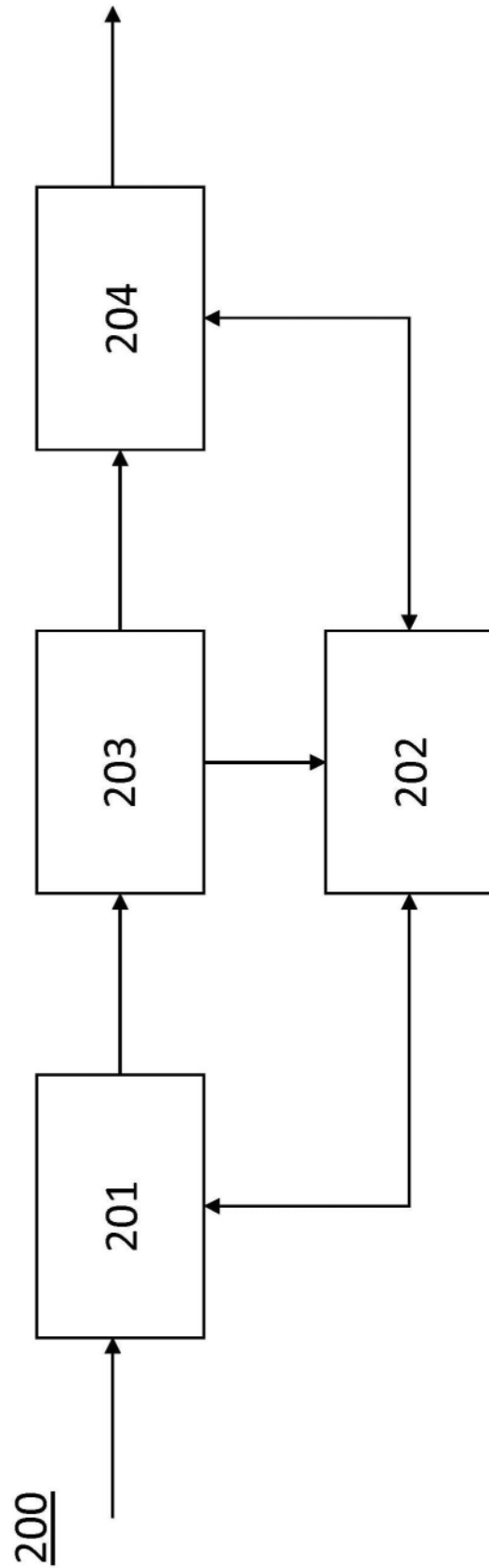


图2

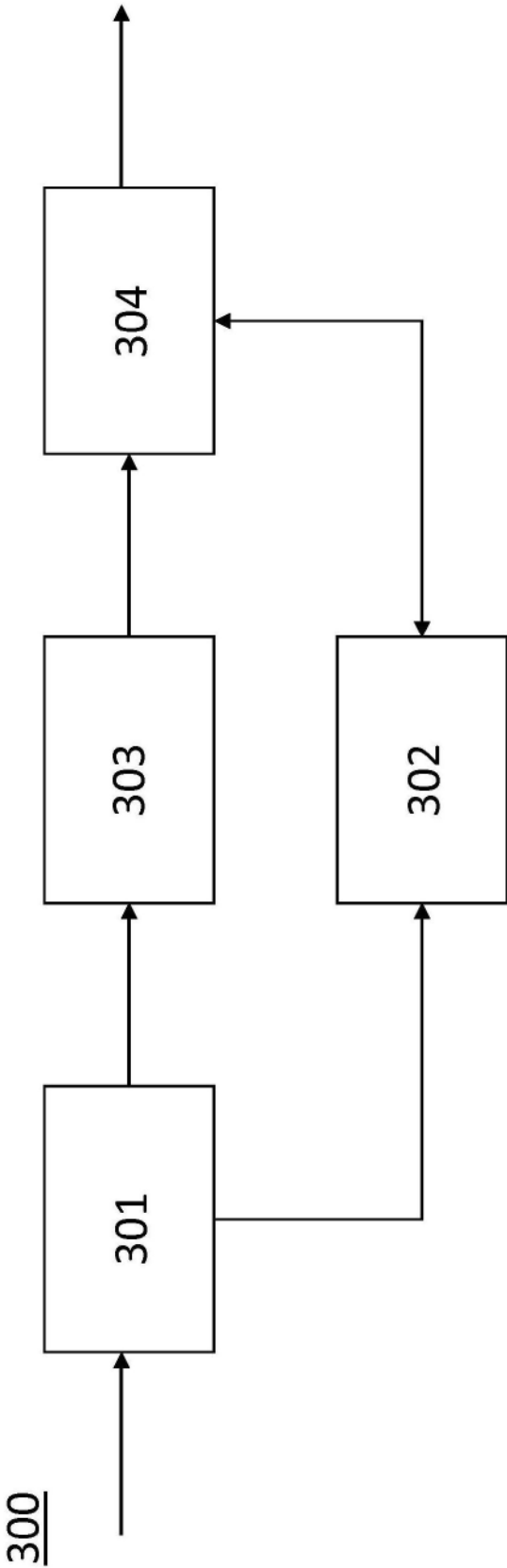


图3

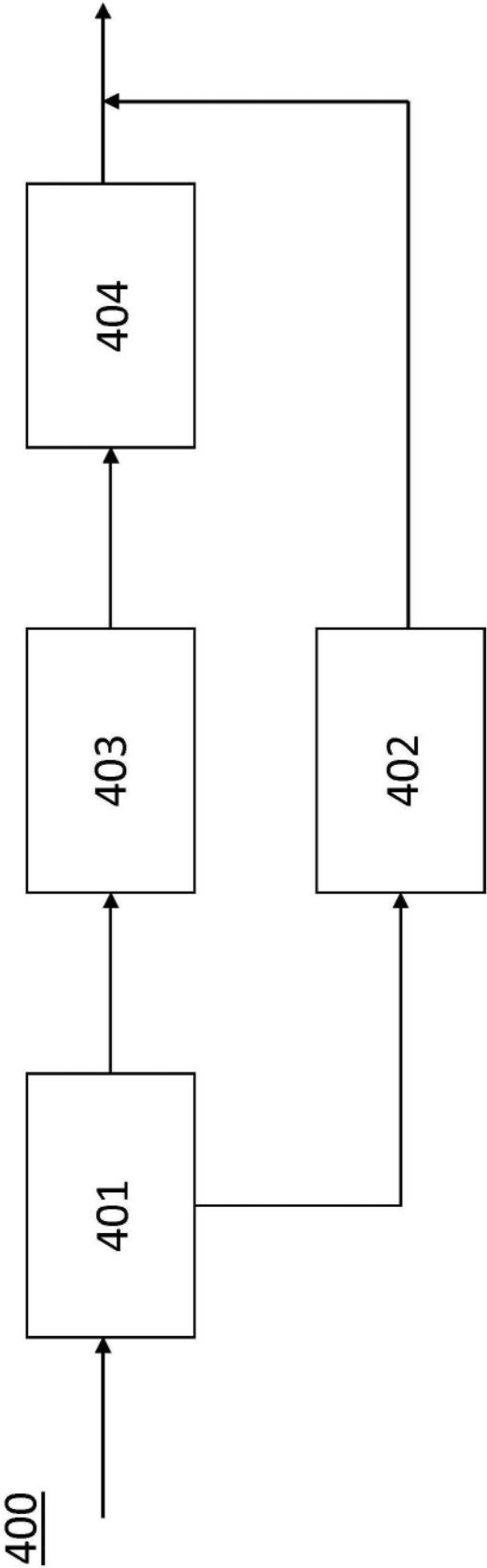


图4