



(12)发明专利

(10)授权公告号 CN 105556599 B

(45)授权公告日 2019.12.10

(21)申请号 201480041809.7

(22)申请日 2014.07.15

(65)同一申请的已公布的文献号
申请公布号 CN 105556599 A

(43)申请公布日 2016.05.04

(30)优先权数据
13177351.7 2013.07.22 EP
13189336.4 2013.10.18 EP

(85)PCT国际申请进入国家阶段日
2016.01.22

(86)PCT国际申请的申请数据
PCT/EP2014/065173 2014.07.15

(87)PCT国际申请的公布数据
W02015/010966 EN 2015.01.29

(73)专利权人 弗朗霍夫应用科学研究促进协会
地址 德国慕尼黑

(72)发明人 弗洛林·基多 安德烈·尼德迈尔

(74)专利代理机构 北京清亦华知识产权代理事务
所(普通合伙) 11201

代理人 宋融冰

(51)Int.Cl.

G10L 19/02(2013.01)

G10L 19/032(2013.01)

G10L 21/038(2013.01)

G10L 19/038(2013.01)

(56)对比文件

CN 1194749 A,1998.09.30,

WO 00/45379 A2,2000.08.03,

CN 101185126 A,2008.05.21,

CN 102568484 A,2012.07.11,

US 6978236 B1,2005.12.20,

WO 2010/003479 A1,2010.01.14,

CN 1272259 A,2000.11.01,

CN 101180677 A,2008.05.14,

CN 102089811 A,2011.06.08,

Jing Wang ET AL.“Context-based

adaptive arithmetic coding in time and
frequency domain for the lossless

compression of audio coding parameters at
variable rate”.《EURASIP Journal on Audio
speech and Music Processing2013》.2013,

审查员 董小东

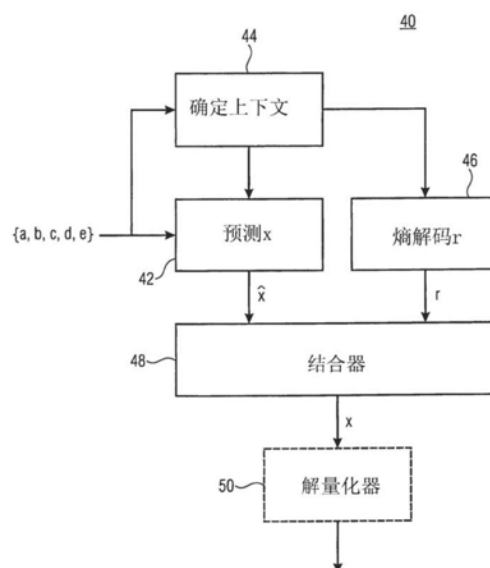
权利要求书3页 说明书16页 附图13页

(54)发明名称

频谱包络的样本值的基于上下文的熵编码、
解码的装置及方法

(57)摘要

通过结合一方面的频谱时间预测与另一方面的对残差进行基于上下文的熵编码,而特别地根据对当前样本值的频谱时间邻近区域中的频谱包络的成对的已编码/解码的样本值之间的偏差的测量,确定用于当前样本值的上下文,获得用于对频谱包络的样本值进行编码的改进概念。一方面的频谱时间预测与另一方面的随着根据偏差测量的上下文的选择来对预测残差进行基于上下文的熵编码的结合与频谱包络的本质相协调。



1. 一种用于对音频信号的频谱包络 (10) 的样本值 (12) 进行解码的基于上下文的熵解码器, 用于:

在频谱时间上预测 (42) 所述频谱包络的当前样本值, 以获得所述当前样本值的估计值;

根据对所述当前样本值的频谱时间邻近区域中的频谱包络的成对的已解码的样本值之间的偏差的测量, 确定 (44) 用于所述当前样本值的上下文;

使用确定的上下文, 熵解码 (46) 所述当前样本值的预测残差值; 以及

结合 (48) 所述估计值与所述预测残差值, 以获得所述当前样本值。

2. 根据权利要求1所述的基于上下文的熵解码器, 还用于通过线性预测以执行频谱时间预测。

3. 根据权利要求1所述的基于上下文的熵解码器, 还用于使用所述当前样本值的所述频谱时间邻近区域中的所述频谱包络的所述成对的已解码的样本值之间的带符号的差, 以测量所述偏差。

4. 根据权利要求1所述的基于上下文的熵解码器, 还用于根据对所述当前样本值的所述频谱时间邻近区域中的所述频谱包络的第一对已解码的样本值之间的偏差的第一测量, 以及对所述当前样本值的所述频谱时间邻近区域中的所述频谱包络的第二对已解码的样本值之间的偏差的第二测量, 确定用于所述当前样本值的上下文, 其中所述第一对在频谱上彼此邻近, 且所述第二对在时间上彼此邻近。

5. 根据权利要求4所述的基于上下文的熵解码器, 还用于通过线性地结合所述第一对和第二对的已解码的样本值, 在频谱时间上预测所述频谱包络的所述当前样本值。

6. 根据权利要求5所述的基于上下文的熵解码器, 还用于设定线性结合的因子, 以使得在所述音频信号被编码的比特率大于预设阈值的情况下, 对于不同的上下文, 所述因子是相同的; 而在所述比特率低于所述预设阈值的情况下, 对于不同的上下文, 单独地设定所述因子。

7. 根据权利要求1所述的基于上下文的熵解码器, 还用于在解码所述频谱包络的样本值中, 使用解码顺序 (30), 顺序地解码所述样本值, 所述解码顺序 (30) 按照时刻遍历所述样本值, 其中在每个时刻中从最低频率至最高频率。

8. 根据权利要求1所述的基于上下文的熵解码器, 还用于在确定所述上下文中, 量化对所述偏差的测量, 并使用量化的测量确定所述上下文。

9. 根据权利要求8所述的基于上下文的熵解码器, 还用于在对所述偏差的测量的量化中使用量化函数 (32), 所述量化函数 (32) 对于在预设区间 (34) 之外的对所述偏差的测量的值为恒定的, 所述预设区间包括零。

10. 根据权利要求9所述的基于上下文的熵解码器, 其中所述频谱包络的值被表示为整数, 且所述预设区间 (34) 的长度小于或等于所述频谱包络的值的整数表示的可表示状态的数目的 $1/16$ 。

11. 根据权利要求1所述的基于上下文的熵解码器, 还用于将由结合推导的所述当前样本值从对数域转换 (50) 至线性域。

12. 根据权利要求1所述的基于上下文的熵解码器, 所述基于上下文的熵解码器管理多个上下文, 每个上下文具有与其相关联的概率分布, 所述概率分布为所述预测残差值的每

个可能值分配各个概率,其中所述基于上下文的熵解码器还用于在熵解码所述预测残差值中,沿着解码顺序,顺序地解码所述样本值,并使用各个的上下文概率分布的集合,所述各个的上下文概率分布在顺序地对频谱包络的样本值进行解码的期间为恒定的。

13.根据权利要求1所述的基于上下文的熵解码器,还用于在熵解码所述预测残差值中,在所述预测残差值在预设值域(68)之外的情况下,使用逸出编码机制。

14.根据权利要求13所述的基于上下文的熵解码器,其中所述频谱包络的所述样本值被表示为整数,且所述预测残差值被表示为整数,且所述预设值域的区间边界(70,72)的绝对值低于或等于所述预测残差值的可表示状态的数目的1/8。

15.参数化解码器,包括:

根据前述权利要求中任一项的用于对音频信号的频谱包络的样本值进行解码的基于上下文的熵解码器(40);

精细结构确定器(82),用于从以频谱线间距而频谱地布置的数据流接收频谱线值(160),以确定所述音频信号的频谱图的精细结构;以及

频谱成形器(84),用于根据所述频谱包络成形所述精细结构。

16.根据权利要求15所述的参数化解码器,其中所述精细结构确定器用于使用人工随机噪声生成、频谱再生以及使用频谱预测和/或频谱熵上下文推导的频谱线状解码中的至少一个,确定所述频谱图的所述精细结构。

17.根据权利要求15所述的参数化解码器,还包括用于解码所述音频信号的频谱图的低频区间(98)的低频区间解码器(94),其中配置所述基于上下文的熵解码器、所述精细结构确定器以及所述频谱成形器,以使得根据所述频谱包络的所述精细结构的成形在所述低频区间的频谱高频延伸(18)内执行。

18.根据权利要求17所述的参数化解码器,其中所述低频区间解码器(94)用于使用频谱线状解码或使用解码的时域低频带音频信号的频谱分解,确定所述频谱图的所述精细结构,所述频谱线状解码使用频谱预测和/或频谱熵上下文推导。

19.根据权利要求15所述的参数化解码器,其中所述精细结构确定器用于使用频谱线状解码以推导第一频率区间(130)内的所述音频信号的所述频谱图的所述精细结构,并定位与所述第一频率区间交叠的第二频率区间(18)内的所述精细结构的零量化部分(142),以及将人工随机噪声生成和/或频谱再生应用至所述零量化部分(142),其中所述频谱线状解码使用频谱预测和/或频谱熵上下文推导,其中所述频谱成形器(84)用于根据所述零量化部分(142)处的频谱包络执行所述精细结构的成形。

20.一种用于对音频信号的频谱包络的样本值进行编码的基于上下文的熵编码器,用于:

在频谱时间上预测所述频谱包络的当前样本值,以获得所述当前样本值的估计值;

根据对所述当前样本值的频谱时间邻近区域中的频谱包络的成对的已编码的样本值之间的偏差的测量,确定用于所述当前样本值的上下文;

基于所述估计值与所述当前样本值之间的偏差确定预测残差值;以及

使用确定的上下文,熵编码所述当前样本值的所述预测残差值。

21.一种用于使用基于上下文的熵解码对音频信号的频谱包络的样本值进行解码的方法,包括:

在频谱时间上预测所述频谱包络的当前样本值,以获得所述当前样本值的估计值;
根据对所述当前样本值的频谱时间邻近区域中的频谱包络的成对的已解码的样本值之间的偏差的测量,确定用于所述当前样本值的上下文;
使用确定的上下文,熵解码所述当前样本值的预测残差值;以及
结合所述估计值与所述预测残差值,以获得所述当前样本值。

22.一种用于使用基于上下文的熵编码,对音频信号的频谱包络的样本值进行编码的方法,包括:

在频谱时间上预测所述频谱包络的当前样本值,以获得所述当前样本值的估计值;
根据对所述当前样本值的频谱时间邻近区域中的频谱包络的成对的已编码的样本值之间的偏差的测量,确定用于所述当前样本值的上下文;
基于所述估计值与所述当前样本值之间的偏差,确定预测残差值;以及
使用确定的上下文,熵编码所述当前样本值的所述预测残差值。

23.一种计算机可读存储介质,存储有计算机程序,所述计算机程序具有程序代码,当在计算机上执行时,用于执行根据权利要求21或22的方法。

频谱包络的样本值的基于上下文的熵编码、解码的装置及方法

技术领域

[0001] 本申请涉及频谱包络的样本值的基于上下文的熵编码及其在音频编码/压缩中的使用。

背景技术

[0002] 如在[1]和[2]中所述的许多现有技术的有损(lossy)音频编码器,基于MDCT变换,并使用不相关性减小和冗余度减小,以对于给定的感知品质最小化所需的比特率。为了减小表示精确度或移除非感知相关的频率信息,不相关性减小通常利用人类听觉系统的感知限制。为了实现剩余数据的最简洁的表示,通常使用结合熵编码的统计建模,应用冗余度减少以利用统计结构或相关性。

[0003] 除了别的之外,参数化编码概念用于高效地对音频内容进行编码。使用参数化编码,使用参数而非使用实际时域音频样本或类似来描述音频信号的部分,如其频谱图的部分。例如,音频信号的频谱图的部分可在解码器侧与仅包括参数(如频谱包络和可选的其他控制合成的参数)的数据流合成,以使合成的频谱图部分适于所传输的频谱包络。此种的新技术为频谱频带复制(SBR),根据此,核心编解码器用于编码并传输音频信号的低频分量,然而所传输的频谱包络用在解码侧,以频谱地成形/形成音频信号的低频带分量的重建的频谱复制,从而在解码侧合成音频信号的高频带分量。

[0004] 以上所述的编码技术的架构内的频谱包络,在数据流内以某种合适的频谱时间分辨率传输。在某种程度上类似于频谱包络样本值的传输,用于缩放频谱线系数或频域系数(如 MDCT系数)的比例因子,同样地以合适的频谱时间分辨率传输,该合适的频谱时间分辨率比原始的频谱线分辨率更粗糙,例如在频谱意义上。

[0005] 为了传递描述频谱包络或比例因子或频域系数的样本的信息,可使用固定的哈夫曼(Huffman)编码表。改进的方法使用上下文编码,如在[2]和[3]中所述,其中用于选择概率分布以用于对值进行编码的上下文跨时间与频率延伸。个别的频谱线(如MDCT系数值)为复频谱线的实数投射,且其可能在本质上显现某些随机性,即便复频谱线的振幅为跨时间恒定的,但相位因帧而异。为了有如[3]所描述的良好结果,这需要上下文选择、量化以及映射的相当复杂的机制。

[0006] 在图像编码中,所使用的上下文通常为图像的跨x轴和y轴的二维的,如在[4]中所述。在图像编码中,例如利用伽马(gamma)调节,值处于线性域或幂律域。此外,在每个上下文中,单个固定的线性预测可被用作平面拟合及基本的边缘检测机制,而且预测误差可被编码。参数化Golomb或Golomb-Rice编码可用于对预测误差进行编码。此外,例如使用基于比特的编码器,运行长度编码(run length coding)用于补偿对非常低的熵信号(每样本低于1比特)进行直接编码的困难度。

[0007] 然而,尽管有关于比例因子和/或频谱包络的编码的改进,仍需要用于对频谱包络的样本值进行编码的改进概念。因此,本发明的目的在于提供用于对频谱包络的频谱值进

行编码的概念。

发明内容

[0008] 通过待决的独立权利要求的主旨实现本发明的目的。

[0009] 本文中所述的实施例基于此发现：通过结合一方面的频谱时间预测与另一方面的对残差进行基于上下文的熵编码，而特别地根据对当前样本值的频谱时间邻近区域中的频谱包络的成对的已编码/解码的样本值之间的偏差的测量，确定用于当前样本值的上下文，可以获得对频谱包络的样本值进行编码的改进概念。一方面的频谱时间预测与另一方面的随着根据偏差测量的上下文的选择来对预测残差进行基于上下文的熵编码的结合与频谱包络的本质相协调：频谱包络的平滑度导致简洁的预测残差分布，以使得关于预测结果的熵编码，频谱时间的交互相关在预测之后几乎被完全地移除，且在上下文选择中可被无视。这反过来降低了用于管理上下文的开销。然而，当前样本值的频谱时间邻近区域中的已编码/解码的样本值之间的偏差测量的使用，仍使得上下文自适应性的规定可行，其由此造成的额外开销是合理的方式改进熵编码效率。

[0010] 根据以下描述的实施例，线性预测与使用差值作为偏差测量相结合，借此保持用于编码的开销低下。

[0011] 根据实施例，选择（用于确定差值最终用于选择/确定上下文的）已编码/解码的样本值的位置，以使得它们以与当前样本值相对准的方式在频谱上或时间上彼此邻近，即它们沿着与时间轴或频谱轴相平行的线分布，而当确定/选择上下文时，额外地考虑差值的符号（sign）。通过此测量，当确定/选择用于当前样本值的上下文时，可考虑预测残差中的一种“趋势”，而仅合理地增加上下文开销。

附图说明

[0012] 以下，关于附图描述本申请的优选实施例，其中：

[0013] 图1显示频谱包络的示意图并示出源于样本值的其组成，和其中定义的可能解码顺序及用于频谱包络的当前编码/解码的样本值的可能频谱时间邻近区域。

[0014] 图2显示根据实施例的用于对频谱包络的样本值进行编码的基于上下文的熵编码器的方块图。

[0015] 图3显示示出在对推导测量进行量化中可使用的量化函数的示意图。

[0016] 图4显示与图2的编码器相配的基于上下文的熵解码器的方块图。

[0017] 图5显示根据另一实施例的用于对频谱包络的样本值进行编码的基于上下文的熵编码器的方块图。

[0018] 图6显示示出根据使用逸出编码（escape coding）的实施例的预测残差的熵编码的可能值的区间相对于预测残差可能值的整体区间的放置的示意图。

[0019] 图7显示与图5的编码器相配的基于上下文的熵解码器的方块图。

[0020] 图8显示使用特定记号的频谱时间邻近区域的可能定义。

[0021] 图9显示根据实施例的参数化音频解码器的方块图。

[0022] 图10通过显示一方面的由频谱包络覆盖的频率区间与另一方面的覆盖整体音频信号的频率范围的另一区间的精细结构之间的关系，显示示出图9的参数化解码器的可能

的实施变型的示意图。

[0023] 图11显示根据图10的变型的与图9的参数化音频解码器相配的音频编码器的方块图。

[0024] 图12显示示出当支持智能间隙填充 (Intelligent Gap Filling, IGF) 时的图9的参数化音频解码器的变型的示意图。

[0025] 图13显示示出根据实施例的频谱包络的源于精细结构频谱图的频谱 (即频谱片), 频谱的IGF填充及其成形的示意图。

[0026] 图14显示支持IGF的音频编码器的方块图, 其与根据图12的图9的参数化解码器的变型相配。

具体实施方式

[0027] 如以下所概述的实施例 (其通常适于频谱包络的编码) 的动机, 使用智能间隙填充 (IGF) 作为示例, 在此呈现引致以下概述的有利实施例的一些想法。IGF为一种新的方法, 用于显著地改进即使以非常低的比特率编码的信号的品质。详细内容请参考以下描述。在任何情况下, IGF处理这样的事实: 由于通常不足的比特预算, 高频区域中的频谱的重要部分被量化成零。为了尽可能保持上频率区域的精细结构, 在IGF信息中低频区域用作自适应地替代高频区域中的大部分被量化成零的目的区域的来源。为了实现良好的感知品质, 重要的需求是, 频谱系数的解码的能量包络与原始信号的解码的能量包络相匹配。为了实现此, 从一个或多个连续的AAC比例因子频带在频谱系数上计算平均频谱能量。使用由比例因子频带所定义的边界来计算平均能量是由已存在的将这些边界小心调整至临界频带的片段所激发, 该临界频带对于人类听觉是特有的。平均能量被使用公式 (类似于用于AAC比例因子的公式) 而转换成dB比例表示, 然后被均匀量化。在IGF中, 根据所请求的总比特率, 可选择地使用不同的量化准确性。平均能量构成由IGF所产生的信息的重要部分, 因此其高效的表示对于IGF的整体性能是高度重要的。

[0028] 因此, 在IGF中, 比例因子能量描述频谱包络。比例因子能量 (SFE) 表示描述频谱包络的频谱值。当SFE进行解码时, 可以利用SFE的特定属性。特别地, 相比于[2]和[3], 已实现SFE表示MDCT频谱线的平均值, 因此其值更“平滑”并与对应复频谱线的平均振幅线性相关。利用此情景, 以下的实施例使用一方面的频谱包络样本值预测与另一方面的使用取决于此频谱包络的成对的邻近的已编码/解码的样本值的偏差的测量的上下文的预测残差的基于上下文的熵编码的结合。此结合的使用特别地适于此类待被编码的数据, 即频谱包络。

[0029] 为了容易理解以下所描述的实施例, 图1显示频谱包络10以及源于样本值12的其组成, 样本值12以特定频谱时间分辨率采样音频信号的频谱包络10。在图1中, 样本值12 示意性地沿着时间轴14和频谱轴16布置。每个样本值12描述或定义对应时空平铺 (tile) 之内的频谱包络10的高度, 此对应时空平铺覆盖, 例如音频信号的频谱图的时空域的特定矩形。因此, 样本值为已通过整合其相关联的频谱时间平铺上的频谱图而获得的整合值。依据能量或一些其他物理测量, 样本值12可测量频谱包络10的高度或强度, 且可被定义在非对数或线性域中, 或对数域中。其中由于其额外地分别沿着轴14和16对样本值进行平滑的特性, 对数域可提供额外的优点。

[0030] 应注意的是, 就以下描述而言, 仅为了示出目的而假设样本值12在频谱上和时间

上规律地布置,即对应于样本值12的对应时空平铺规律地覆盖源于音频信号的频谱图的频带 18,但是此规律并非强制性的。相反地,也可使用样本值12对频谱包络10的不规则采样,每个样本值12表示在其对应时空平铺之内的频谱包络10的高度的平均值。以下概述的邻近区域的定义仍可用于频谱包络10的不规则采样的可选实施例。以下呈现此可能性的简要描述。

[0031] 然而,之前,应注意的是,上述频谱包络可受到为了各种理由的用于从编码器到解码器的传输的编码和解码。例如,为了可扩展性目的可以使用频谱包络,以延伸音频信号的低频带的核心编码,即将此低频带朝更高的频率延伸,即延伸至频谱包络所涉及的高频带。在此情况下,例如,以下描述的基于上下文的熵解码器/编码器可为SBR解码器/编码器的部分。可选地,可以为使用如上已提及的IGF的音频编码器/解码器的部分。在IGF中,此外,使用描述频谱图的高频部分频谱包络的频谱值来描述音频信号频谱图的高频部分,以能够使用频谱包络填充高频部分内的频谱图的零量化区。以下描述相关细节。

[0032] 图2显示根据本申请的实施例的用于对音频信号的频谱包络10的样本值12进行编码的基于上下文的熵编码器。

[0033] 图2的基于上下文的熵编码器大体使用附图标记20来标示,且包括预测器22、上下文确定器24、熵编码器26以及残差确定器28。上下文确定器24和预测器22具有输入,并在输入处访问频谱包络(图1)的样本值12。熵编码器26具有连接至上下文确定器24的输出的控制输入,以及连接至残差确定器28的输出的数据输入。残差确定器28具有两个输入,其中一个连接至预测器22的输出,而另一个为残差确定器28提供对频谱包络10的样本值12的访问。特别地,残差确定器28在其输入处接收当前待被编码的样本值 x ,而上下文确定器24和预测器22在其输入处接收已被编码且位于此当前样本值 x 的频谱时间邻近区域内的样本值12。

[0034] 预测器22用于在频谱时间上预测频谱包络10的当前样本值 x ,以获得估计值 \hat{x} 。如关于以下概述的更详细实施例将被示出的,预测器22可使用线性预测。特别地,在执行频谱时间预测中,预测器22检查在当前样本值 x 的频谱时间邻近区域中已编码的样本值。例如,参见图1。使用粗体连续绘制的轮廓线示出此当前样本值 x 。当前样本值 x 的频谱时间邻近区域中的样本值以细线显示,根据实施例,其形成预测器22的频谱时间预测的基础。例如,“a”表示直接邻近于当前样本值 x 的样本值12,其在频谱上共置至当前样本值 x ,但在时间上先于当前样本值 x 。同样地,邻近样本值“b”表示直接邻近于当前样本值 x 的样本值,其在时间上共置至当前样本值 x ,但相比于当前样本值 x ,其涉及更低的频率。在当前样本值 x 的频谱时间邻近区域中的样本值“c”为当前样本值 x 的最接近的邻近样本值,其在时间上先于当前样本值 x ,且涉及更低的频率。频谱时间邻近区域甚至可以围绕表示当前样本值 x 的邻近值中的下一个的样本值。例如,样本值“a”分隔当前样本值 x 与样本值“d”,即样本值“d”在时间上被共置至当前样本值 x ,且先于样本值 x 而仅有样本值“a”位于两者之间。同样地,样本值“e”邻近于样本值 x 而在时间上被共置至当前样本值 x ,且沿着频谱轴16邻近于样本值 x 而仅有样本值“b”位于两者之间。

[0035] 如以上已概述的,虽然假设样本值12沿着时间轴14和频谱轴16规律地布置,但此规则并非强制性,而邻近区域的定义和邻近样本值的识别可被延伸至不规则的情况。例如,邻近样本值“a”可被定义为沿着时间轴邻近于当前样本的频谱时间平铺的左上角,且在时

间上先于左上角的样本值。类似定义可被用于定义其他邻近值,例如邻近值b至e。

[0036] 如以下将更详细概述的,预测器22可以根据当前样本值x的频谱时间位置,使用频谱时间邻近区域内的所有样本值的不同子集,即 $\{a, b, c, d, e\}$ 的子集。实际上使用哪个子集可以,例如取决于频谱时间邻近区域(由集合 $\{a, b, c, d, e\}$ 所定义)内的邻近样本值的可用性。例如,由于当前样本值x直接随后于随机访问点(即使得解码器开始解码的时间点)从而对频谱包络10的在先部分的依赖被禁止/阻止,邻近样本值a、d及c可能是不可用的。可选地,由于当前样本值x表示区间18的低频边缘从而各个邻近样本值的位置落在区间18之外,邻近样本值b、c及e可能是不可用的。在任何情况下,预测器22可通过线性结合频谱时间邻近区域内的已编码的样本值,在频谱时间上预测当前样本值x。

[0037] 上下文确定器24的任务为选择用于对预测残差进行熵编码的一些支持的上下文中的一个,即 $r = x - \hat{x}$ 。至此,上下文确定器24根据对频谱时间邻近区域中的a至e之间的成对的已编码的样本值的偏差的测量,确定用于当前样本值x的上下文。在以下概述的特定实施例中,频谱时间邻近区域内的成对的样本值的差被用作对其间的偏差的测量,例如a-c、b-c、b-e、a-d或类似,但可选地,可以使用其他偏差测量,例如,商数(即 a/c 、 b/c 、 a/d),差的幂(幂的值不等于1,例如为不等于1的非偶数n)(即 $(a-c)^n$ 、 $(b-c)^n$ 、 $(a-d)^n$),或一些其他类型的偏差测量,如 $a^n - c^n$ 、 $b^n - c^n$ 、 $a^n - d^n$ 或 $(a/c)^n$ 、 $(b/c)^n$ 、 $(a/d)^n$,其中 $n \neq 1$ 。在此,例如,n也能是大于1的任意值。

[0038] 如以下将更详细显示的,上下文确定器24可用于根据对频谱时间邻近区域中的第一对已编码的样本值之间的偏差的第一测量和频谱时间邻近区域内的第二对已编码的样本值之间的偏差的第二测量,确定用于当前样本值x的上下文,其中第一对在频谱上彼此邻近,第二对在时间上彼此邻近。例如,当a与c在频谱上彼此邻近且b与c在时间上彼此邻近,可使用差值b-c和a-c。预测器22可使用相同集合的邻近样本值,即 $\{a, c, b\}$,以获得估计值 \hat{x} ,即例如通过样本值的线性结合。在样本值a、c和/或b中的任一个不可用的情况下,不同集合的邻近样本值可被用于上下文确定和/或预测。如以下所设定的,可以设定线性结合的因子,以使得在音频信号被编码的比特率大于预设阈值的情况下,对于不同的上下文,因子是相同的,而在音频信号被编码的比特率低于预设阈值的情况下,对于不同的上下文因子被单独地设定。

[0039] 作为中间注释,应提及的是,频谱时间邻近区域的定义可适于编码/解码顺序,而基于上下文的熵编码器20顺序地沿着此顺序对样本值12进行编码。例如,如图1所示,基于上下文的熵编码器可用于使用解码顺序30顺序地对样本值12进行编码,其按照时刻(time instant)遍历样本值12,其中在每个时刻中,从最低频率至最高频率。以下,“时刻”被表示为“帧”,但是时刻可选地被称为时隙、时间单元或类似。在任何情况下,在时间前馈之前使用此频谱遍历中,被延伸至在先时间且向更低的频率延伸的频谱时间邻近区域的定义提供对应样本值已被编码/解码且为可用的最高可行概率。在本申请中,邻近区域内的值总是编码/解码的,并假设它们存在,但是对于其他邻近区域和解码顺序对,这可能是不同的。自然地,此解码器使用相同的解码顺序30。

[0040] 如上已表示的,样本值12可表示对数域中的频谱包络10。特别地,已使用对数量化函数将频谱值12量化成整数值。因此,由于量化,由上下文确定器24确定的偏差测量已经是

本质上的整数。例如,此为使用差作为偏差测量的情况。不考虑由上下文确定器24确定的偏差测量的本质整数的本质,上下文确定器24可使偏差测量受到量化并使用量化的测量确定上下文。特别地,例如,如以下所概述的,对于预设区间之外的偏差测量的值,上下文确定器24使用的量化函数可以是恒定的,此预设区间包含零。

[0041] 图3示意性地显示将未量化的偏差测量量化至量化的偏差测量的此量化函数32,其中,在此示例中,刚刚提及的预设区间34从-2.5延伸至2.5,其中在此区间之上的未量化的偏差测量值被持续地映射至量化的偏差测量值3,而在此区间34之下的未量化的偏差测量值被持续地映射至量化的偏差测量值-3。因此,仅七个上下文被区分出且必须由基于上下文的熵编码器所支持。在以下概述的实施方式示例中,区间34的长度为5,其仅为示例,其中频谱包络的样本值的该集合的可能值的基数为 2^n (例如=128),即大于区间长度的16倍。在如稍后所示出的使用逸出编码的情况下,此频谱包络的样本值的可能值的范围可被定义成 $[0; 2^n]$,其中n为选择的整数,以使得 2^{n+1} 低于预测残差值的可编码的可能值的基数,根据以下描述的特定实施方式示例,其为311。

[0042] 熵编码器26使用由上下文确定器24确定的上下文以高效地对预测残差r进行熵编码,该预测残差r反过来由残差确定器28基于实际的当前样本值x和估计值 \hat{x} 来确定,例如以减法的方式。优选地,可使用算术编码。上下文可具有与其相关联的恒定的概率分布。对于每个上下文,与其相关联的概率分布将特定的概率值分配给源于熵编码器26的符号字母表的每个可能符号。例如,熵编码器26的符号字母表与预测残差r的可能值的范围相一致,或符号字母表覆盖此范围。在可选的实施例中,以下其被更详细地概述,可使用特定的逸出编码机制,以保证待被熵编码器26熵编码的值r在熵编码器26的符号字母表内。当使用算术编码时,熵编码器26使用由上下文确定器24所确定的确定的上下文的概率分布,以将当前概率区间(其表示熵编码器26的内部状态)细分成对每个字母表值有一个子区间,其中根据r的实际值来选择子区间中的一个,并利用,例如重新归一化处理,输出通知解码侧更新概率区间偏移及宽度的算术编码的比特流。然而,可选地,对于每个上下文,熵编码器26可使用个别的变长编码表,其将各个上下文的概率分布转译成r的可能值至与各个可能值r的各个频率相对应的的长度的码上的对应映射。也可使用其他熵编解码器。

[0043] 为了完整性起见,图2显示量化器36可连接在残差确定器28的输入之前,利用例如被应用至未量化的样本值x的对数量化函数,当前样本值x在此输入处回传,以获得此当前样本值 \hat{x} ,如以上已概述的。

[0044] 图4显示根据实施例的基于上下文的熵解码器,其与图2的基于上下文的熵编码器相配。

[0045] 使用附图标记40来标示并类似于图2的编码器来解释图4的基于上下文的熵解码器。因此,基于上下文的熵解码器40包括预测器42、上下文确定器44、熵解码器46以及结合器48。上下文确定器44和预测器42像图2的编码器20的预测器22和上下文确定器24一样操作。即,预测器42在频谱时间上预测当前样本值x,即当前待解码的样本值,以获得估计值 \hat{x} ,并将其输出至结合器48,且上下文确定器44确定用于根据样本值x的频谱时间邻近区域内的成对的已解码的样本值之间的偏差测量对当前样本值x的预测残差r进行熵解码的上下文,通过熵解码器46的控制输入向熵解码器46通知确定的上下文。因此,上下文确定器44和

预测器42均访问频谱时间邻近区域中的样本值。结合器48具有分别连接至预测器42和熵解码器46的输出的两个输入以及用于输出当前样本值的输出。特别地,熵编码器46使用上下文确定器44确定的上下文对用于当前样本值 x 的残差值 r 进行熵解码,且结合器48结合估计值 \hat{x} 和对应残差值 r 以获得当前样本值 x ,例如通过相加。仅为了完整性起见,图4显示解量化器50,其可后接于结合器48的输出,以对结合器48输出的样本值解量化,例如通过使用例如指数函数使样本值受到从对数域至线性域的转换。

[0046] 熵解码器46逆转熵编码器26执行的熵编码。即,熵解码器也管理上下文的数量,并对于当前样本值 x 使用由上下文确定器44选择的上下文,其中每个上下文具有与其相关联的对应概率分布,其为 r 的每个可能值分配特定的概率,而该特定的概率与上下文确定器24为熵编码器26所选择的相同。

[0047] 当使用算术编码,熵解码器46逆转,例如熵编码器26的区间细分顺序。例如,在当前概率区间内,熵解码器46的内部状态由当前区间的概率区间宽度和指向子区间的偏移值所定义,该子区间得自当前样本值 x 的 r 的实际值所对应的当前概率区间。熵解码器46使用熵编码器26所输出的回传的算术编码的比特流来更新概率区间及偏移值(例如以重新归一化处理的方式),并通过检查偏移值以及识别其落入的子区间以获得 r 的实际值。

[0048] 如以上已提及的,将残差值的熵编码限制在预测残差 r 的可能值的一些小区间上是有利的。图5显示图2的基于上下文的熵编码器的变型以实现此。除了图2所显示的元件,图5的上下文熵编码器包括连接在残差确定器28与熵编码器26之间的控制器,即控制器60,以及通过控制器60控制的逸出编码处理器62。

[0049] 图5中以粗略的方式示出控制器60的功能。如图5所示,控制器60检查残差确定器28基于实际样本值 x 与其估计值 \hat{x} 的比较所确定的初始确定的残差值 r 。特别地,如图5所示的在64处,控制器60检查 r 是否在预设值区间之内或之外,参见例如图6。图6显示初始预测残差 r 的沿 x 轴的可能值,而 y 轴显示实际熵编码的 r 。此外,图6显示初始预测残差 r 的可能值的范围,即66,以及在检查64中涉及的刚刚提及的预设区间68。例如,设想样本值12为0到 $2^n - 1$ 之间的整数值,也包含两边的数值。然后,用于预测残差 r 的可能值的范围66可从 $-(2^n - 1)$ 延伸到 $2^n - 1$,也包含两边的数值,且区间68的区间边界70和72的绝对值可小于或等于 2^{n-2} ,即,区间边界的绝对值可小于范围66内的该集合的可能值的基数的 $1/8$ 。在关于xHE-AAC的以下设定的实施方式示例中的一个中,区间68为从-12到+12(包括两边的数值),区间边界70和72为-13和+13,且逸出编码通过对VLC编码的绝对值进行编码而延伸区间68,即使用4个比特将区间68延伸至 $-/+ (13+15)$,以及如果在先4个比特为15,则使用另外的7个比特将区间68延伸至 $-/+ (13+15+127)$ 。因此,为了充分地覆盖用于预测残差的可能值的范围66(其从-127延伸至127),预测残差可被编码在 $-/+155$ 的范围内(包含两边的数值)。正如可看出的, $[127;127]$ 的基数为255,且13(即内部边界70和72的绝对值)小于 $32 \approx 255/8$ 。当对比区间68的长度与使用逸出编码的可编码的可能值的基数(即 $[-155;155]$)时,然后发现有利地选择内部边界70和72的绝对值小于所述基数(在此为311)的 $1/8$ 或甚至 $1/16$ 。

[0050] 在初始预测残差 r 位于区间68内的情况下,控制器60使熵编码器26直接地对此初始预测残差 r 进行熵编码。没有采用特定测量。然而,如果如残差确定器28提供的 r 在区间68之外,控制器60启动逸出编码过程。特别地,根据一个实施例,直接邻近于区间68的区间边界70和72的直接邻近值可以属于熵编码器26的符号字母表且用作逸出码本身。即,熵编

码器26的符号字母表将围绕区间68的所有值加上低于及高于区间68的直接邻近值（如以大括号74所标示的），而在残差值 r 大于区间68的上边界72的情况下，控制器60 将简单地将待被熵编码的值减小至直接邻近于区间68的上边界72的最高字母表值76，以及在初始预测残差 r 小于区间68的下边界70的情况下，控制器60将直接邻近于区间68 的下边界70的最低字母表值78传送至熵编码器26。

[0051] 通过使用刚概述的实施例，在熵编码的值 r 位于区间68内的情况下，熵编码的值 r 对应于，即等于，实际预测残差。然而，如果熵编码的值 r 等于值76，则清楚的是，当前样本值 x 的实际预测残差 r 等于76或高于76的某值，而如果熵编码的残差值 r 等于值78，则实际预测残差 r 等于此值78或低于78的某值。即，在此情况下实际上有两个逸出码76 及78。在初始值 r 位于区间68之外的情况下，控制器60触发逸出编码处理器62以将使得解码器以独立于等于逸出码76或78的熵编码的值 r 的自足性方式或取决于等于逸出码76 或78的熵编码的值 r 而复原实际预测残差的编码插入数据流内，熵编码器26将其熵编码的数据流输出至该数据流。例如，逸出编码处理器62可将使用充分比特长度（例如长度 2^{n+1} ）的二进制表示的包含实际预测残差 r 的符号的实际预测残差 r 直接写入数据流，或仅将使用比特长度 2^n 的二进制表示的实际预测残差 r 的绝对值（其使用逸出码76用以信号化符号“+”，及使用逸出码78用以信号化符号“-”）写入数据流。可选地，在初始预测残差超过上边界72的情况下，仅初始预测残差值 r 与逸出码76的值之间的差的绝对值被编码，而在初始预测残差低于下边界70的情况下，仅初始预测残差 r 与逸出码78的值之间的差的绝对值被编码。即，根据一个实施方式示例，使用条件编码来完成：首先在逸出编码的情况下，使用四个比特对 $\min(|x-\hat{x}|-13; 15)$ 编码，以及如果 $\min(|x-\hat{x}|-13; 15)$ 等于15，则使用另外的七个比特对 $|x-\hat{x}|-13-15$ 编码。

[0052] 显然地，逸出编码比区间68内的常见预测残差的编码较不复杂。例如，没有使用上下文自适应性。相反地，在逸出情况下编码的值的编码可通过简单地写用于值的二进制表示（如 $|r|$ 或甚至 x ）来直接地执行。然而，优选地选择区间68，以使得统计地很少发生逸出过程，且仅在样本值 x 的统计下表示“outliers”。

[0053] 图7显示图4的基于上下文的熵解码器的变型，与图5的熵编码器相对应或相配。类似于图5的熵编码器，图7的基于上下文的熵解码器与图4所显示的不同在于，一方面控制器71连接在熵解码器46与结合器48之间，另一方面其中图7的熵解码器额外地包括逸出码处理器73。类似于图5，控制器71执行检查74以检查熵解码器46输出的熵解码的值 r 是否在区间68之内或与某逸出码相对应。如果应用后者环境，控制器71触发逸出码处理器73以从也携带由熵解码器46熵解码的熵编码数据流的数据流中抽取由逸出码处理器62 插入的前述码，例如充分比特长度的二进制表示，其可以以独立于熵解码的值 r 所指示的逸出码的自足方式或以取决于熵解码的值 r 所假定的实际逸出码的方式来指示实际预测残差 r ，正如关于图6已说明的。例如，逸出码处理器73从数据流读取值的二进制表示，并将其加到逸出码的绝对值（即分别加到上或下边界的绝对值），并使用值的符号以读取各个边界的符号（即用于上边界的“+”符号，用于下边界的“-”符号）。可使用条件编码。即，如果熵解码器46输出的熵解码的值 r 位于区间68之外，则逸出码处理器73首先从数据流读取 p -比特绝对值，并检测其是否为 2^p-1 。如果不是，如果逸出码为上边界72，则通过将 p -比特绝对值加到熵解码的值 r 以更新熵解码的值 r ；如果逸出码为下边界70，则通过从熵解码的值 r 减去 p -比特绝对值

以更新熵解码的值 r 。然而,如果 p -比特绝对值为 2^p-1 ,则从比特流读取另一 q -比特绝对值,且如果逸出码为上边界72,则通过将 q -比特绝对值加 2^p-1 加到熵解码的值 r 以更新熵解码的值 r ;如果逸出码为下边界70,则通过从熵解码的值 r 减去 p -比特绝对值及 2^p-1 以更新熵解码的值 r 。

[0054] 然而,图7显示另一实施例。根据此实施例,由逸出码处理器62和72实现的逸出码过程对完整的样本值 x 直接编码,以使得在逸出码的情况下,估计值 \hat{x} 为多余的。例如,在此情况, 2^n 比特表示可能足够并指示 x 的值。

[0055] 仅作为预防措施,应注意的是,通过不对频谱值(其预测残差超过区间68或位于区间 68之外)进行任何熵解码,利用这些可选的实施例,实现逸出编码的另一方式也是可行的。例如,对于每个语法元素,可传输指示是否使用熵编码或使用逸出编码对其进行编码的旗标。在此情况下,对于每个样本值,旗标将指示编码的选择方式。

[0056] 以下,描述用于实现以上实施例的具体示例。特别地,以下设定的此明确示例举例说明如何处理前述的频谱时间邻近区域中的特定的之前编码/解码的样本值的不可用。此外,呈现特定示例用于设定可能值域66、区间68、量化函数32及范围34等等。之后将描述可关于IGF使用的具体示例。然而,应注意的是,以下设定的描述可容易地转用到其他情况,例如,其布置有频谱包络的样本值的时间网格由其他时间单元而非帧来定义(例如多组 QMF 隙),且同样地通过将子频带的分组至频谱时间平铺的子分组来定义频谱分辨率。

[0057] 使用 t (时间)表示跨时间的帧数量, f (频率)表示跨比例因子(或比例因子群组)的频谱包络的各个样本值的位置。以下,样本值被称为SFE值。期望使用已从位置 $(t-1)$ 、 $(t-2) \dots$ 处的之前解码的帧以及从在频率 $(f-1)$ 、 $(f-2) \dots$ 处的位置 (t) 处的当前帧可得到的信息,对 x 的值进行编码。图8再次绘示此状况。

[0058] 对于单独帧,设定 $t=0$ 。单独帧为使其本身能够作为用于解码实体的随机访问点。其因此表示随机访问解码的时刻在解码侧是可行的。就频谱轴16而言,与最低频率相关联的第一SFE 12有 $f=0$ 。在图8中,用于计算上下文的时间和频率中的邻近值(在编码器和解码器均可用),为如图1中的情况下的 a 、 b 、 c 、 d 及 e 。

[0059] 根据是否 $t=0$ 或 $f=0$,有几种情况。在每种情况下,在每个上下文中,可以基于邻近值计算值 x 的自适应估计 \hat{x} ,如下所示:

[0060]

$t = 0$	频谱时间预测 $\hat{x} = 0$,
$f = 0$	使用 7 比特未经加工的二进制数, 上下文自适应地编码 $r = x - \hat{x}$;
$t = 0$	频谱时间预测 $\hat{x} = b$,
$f = 1$	使用上下文 se01 , 上下文自适应地编码 $r = x - \hat{x}$;
$t = 0$	频谱时间预测 $\hat{x} = b$,
$f \geq 2$	使用上下文 se02 $[Q(b - e)]$, 上下文自适应地编码 $r = x - \hat{x}$;
$t = 1$	频谱时间预测 $\hat{x} = a$,
$f = 0$	使用上下文 se10 , 上下文自适应地编码 $r = x - \hat{x}$;
$t \geq 2$	频谱时间预测 $\hat{x} = a$,
$f = 0$	使用上下文 se20 $[Q(a - d)]$, 上下文自适应地编码 $r = x - \hat{x}$;
$t \geq 1$	频谱时间预测
$f \geq 1$	$\hat{x} = r \text{INT}(\alpha_{[Q(b-c)][Q(a-c)]}a + \beta_{[Q(b-c)][Q(a-c)]}b + \gamma_{[Q(b-c)][Q(a-c)]}c + \delta_{[Q(b-c)][Q(a-c)]})$, 使用上下文 se11 $[Q(b - c)][Q(a - c)]$, 上下文自适应地编码 $x - \hat{x}$ 。

[0061] 如以上已表示的, 数值 $b-e$ 和 $a-c$ 表示偏差测量。其表示靠近待被解码/编码的值的跨频率变化性的噪声的期望数量, 即 x 。值 $b-c$ 和 $a-d$ 表示靠近 x 的跨时间变化性的噪声的期望数量。为了显著地减少上下文的总数, 在它们被用于选择上下文之前, 可将其非线性地量化, 例如关于图3所设定的。上下文指示估计值 \hat{x} 的置信, 或等效地, 编码分布的峰值。例如, 可如图3中所示的量化函数。其可被定义为 $Q(x) = x$, 对于 $|x| \leq 3$; 以及 $Q(x) = 3 \sin(x)$, 对于 $|x| > 3$ 。此量化函数将所有整数值映射至七个值 $\{-3, -2, -1, 0, 1, 2, 3\}$ 。请注意以下。在写 $Q(x) = x$ 时, 已利用了两个整数之差本身为整数。为了分别匹配前述的更多一般描述及图3中的函数, 此公式可被写成 $Q(x) = r \text{Int}(x)$ 。然而, 如果仅用于偏差测量的整数输入, $Q(x) = x$ 在功能上等效 $Q(x) = r \text{Int}(x)$, 对于整数 x , 其中 $|x| \leq 3$ 。

[0062] 上述表格中的术语 **se02**[], **se20**[] 及 **se11**[][] 为上下文向量/矩阵。即, 这些向量/矩阵的实体中的每个为/表示上下文索引, 其索引可用上下文中的一个。此三个向量/矩阵中的每个可索引不相交集的上下文中的上下文。即, 不同集合的上下文可由以上概述的上下文确定器根据可用条件来选择。以上表格示意性地区分六个不同的可用条件。与 **se01** 及 **se10** 相对应的上下文也可与不同于由 **se02**、**se20** 与 **se11** 索引的上下文群组中的任

何上下文的上下文相对应。 x 的估计值被计算为 $\hat{x} = \text{rINT}(\alpha a + \beta b + \gamma c + \delta)$ 。对于较高的比特率,可使用 $\alpha=1$ 、 $\beta=-1$ 、 $\gamma=1$ 及 $\delta=0$,而对于较低的比特率,基于来自训练数据集合的信息,独立的系数集合可用于每个上下文。

[0063] 可使用用于每个上下文的独立分布来对预测误差或预测残差 $r = x - \hat{x}$ 编码,使用从表示性训练数据集合中抽取的信息推导该独立分布。两个特殊符号可用在编码分布74的两侧,即76及78,以指示范围外的大的负值或正值,然后使用逸出编码技术对其进行编码,如以上已概述的。例如,根据实施方式示例,在逸出编码的情况下使用四个比特对 $\min(|x - \hat{x}| - 13; 15)$ 编码,如果 $\min(|x - \hat{x}| - 13; 15)$ 等于15,则使用另外的七个比特对 $|x - \hat{x}| - 13 - 15$ 编码。

[0064] 关于以下附图,描述关于以上提及的基于上下文的熵编码器/解码器如何建造于各个音频解码器/编码器的各种可能性。例如,图9显示参数化解码器80,有利地根据以上概述的实施例中的任一个的基于上下文的熵解码器40建造其中。除了基于上下文的熵解码器40,参数化解码器80包括精细结构确定器82及频谱成形器84。可选择地,参数化解码器80包括逆变换器86。如以上概述的,基于上下文的熵解码器40接收根据基于上下文的熵编码器的以上概述的实施例中的任一个而编码的熵编码的数据流88。因此,数据流88具有编码至其的频谱包络。以以上概述的方式,基于上下文的熵解码器40对参数化解码器80 寻求重建的音频信号的频谱包络的样本值进行解码。精细结构确定器82用于确定此音频信号的频谱图的精细结构。至此,精细结构确定器82可从外部接收信息,例如数据流的另一部分也包括数据流88。以下描述另一实施例。然而,在另一实施例中,精细结构确定器82 可通过本身使用随机或伪随机处理来确定此精细结构。频谱成形器84反过来用于根据如由基于上下文的熵解码器40解码的频谱值所定义的频谱包络成形此精细结构。换句话说,频谱成形器84的输入分别地连接至基于上下文的熵解码器40和精细结构确定器82的输出,以从其中一方面接收频谱包络另一方面接收音频信号的频谱图的精细结构。频谱成形器84 在其输出处输出根据频谱包络成形的频谱图的精细结构。逆变换器86可对成形的精细结构执行逆变换,以在其输出处输出音频信号的重建。

[0065] 特别地,精细结构确定器82可用于使用人工随机噪声产生、频谱再生以及使用频谱预测和/或频谱熵上下文推导的频谱线状解码中的至少一个,确定频谱图的精细结构。关于图 10,描述第一个两种可能性。图10示出由基于上下文的熵解码器40解码的频谱包络10属于频率区间18的可能性,该区间18形成低频率区间90的高频延伸,即区间18将较低的低频率区间90朝向较高的频率延伸,即区间18在区间19的高频侧的界定区间19。因此,图10显示待被参数化解码器80重现的音频信号实际上覆盖频率区间92的可能性,其中区间18仅表示整体频率区间92的高频部分。如图9所示,参数化解码器80可以,例如额外地包括低频解码器94,其用于对伴随数据流88的低频数据流96进行解码,以在其输出处获得音频信号的低频带版本。图10中使用附图标记98绘示的此低频版本的频谱图。音频信号的频率版本98和区间18内的成形的精细结构导致完整的频率区间92(即跨完整的频率区间92)的其频谱图的音频信号重建。如图9中的虚线所指示的,逆变换器86可对完整的区间92执行逆变换。在此框架中,精细结构确定器82可在时域或频域中从解码器94 接收低频版本98。在第一种情况下,精细结构确定器82可使所接收的低频版本受到变换至频谱域以获得频谱图98,并使

用如用箭头100所示出的频谱再生根据基于上下文的熵解码器40提供的频谱包络来获得待被频谱成形器84成形的精细结构。然而,如以上已概述的,精细结构确定器82甚至可不从LF解码器94接收音频信号的低频版本,而仅使用随机或伪随机处理来产生精细结构。

[0066] 图11中绘示与根据图9和10的参数化解码器相对应的对应参数化编码器。图11的参数化编码器包括频率交越110,其接收待被编码的音频信号112、高频带编码器114及低频带编码器116。频率交越110将回传音频信号112分解成两个分量,即与回传音频信号112的高通滤波版本相对应的第一信号118,以及与回传音频信号112的低通滤波版本相对应的低频信号120。其中被高频信号118和低频信号120覆盖的频带在一些交越频率处交界(图10中的比较122)。低频带编码器116接收低频信号120,并将其编码至低频数据流,即96,且高频带编码器114计算样本值,其描述高频区间18内的高频信号118的频谱包络。高频带编码器114也包括上述的基于上下文的熵编码器,其用于对频谱包络的这些样本值进行编码。例如,低频带编码器116可以为变换编码器,低频带编码器116对低频信号120的变换或频谱图进行编码的频谱时间分辨率可大于样本值12决定高频信号118的频谱包络的频谱时间分辨率。因此,高频带编码器114输出除了别的以外的数据流88。如图11中的虚线124所显示的,例如,关于描述频谱包络的样本值的生成或至少关于在样本值采样频谱包络的频谱时间分辨率的选择,为了控制高频带编码器114,低频带编码器116可向高频带编码器114输出信息。

[0067] 图12显示实现图9的参数化解码器80以及特别地,实现精细结构确定器82的另一可能性。特别地,根据图12的示例,精细结构确定器82本身接收数据流,并基于其,使用应用频谱预测和/或频谱熵上下文推导的频谱线状解码,确定音频信号频谱图的精细结构。即,例如,精细结构确定器82本身从数据流以频谱图的形式复原精细结构,其由重叠变换的频谱的时序所构成。然而,在图12的情况下,由精细结构确定器82确定的精细结构涉及第一频率区间130,且与音频信号的完整的频率区间(即,92)相一致。

[0068] 在图12的示例中,频谱包络10所涉及的频率区间18与区间130完全地重叠。特别地,区间18形成区间130的高频部分。例如,由精细结构确定器82复原的并覆盖频率区间130的频谱图132内的许多频谱线,将被量化成零,尤其是在高频部分18内的。然而,为了以高品质重建音频信号,即使以合理的比特率在高频部分18内,参数化解码器80利用频谱包络10。频谱包络10的频谱值12以比由精细结构确定器82解码的频谱图132的频谱时间分辨率粗糙的频谱时间分辨率描述高频部分18内的音频信号的频谱包络。例如,频谱包络10的频谱时间分辨率在频谱上较粗糙,即频谱分辨率比精细结构132的频谱线粒度较为粗糙。如上所述,例如,在频谱上,频谱包络10的样本值12可描述频带134中的频谱包络10,频谱图132的频谱线被分组至该频带134用于频谱线系数的比例因子频带状缩放。

[0069] 然后,使用类似频谱再生或是人工噪声生成的机制,频谱成形器84可使用样本值12填充频谱线群组或对应于各个样本值12的频谱时间平铺内的频谱线,根据描述频谱包络的对应样本值来调整产生的精细结构水平或各个频谱时间平铺/比例因子群组内的能量。例如,参见图13。图13示意性地显示频谱图132中的频谱,其对应于一个帧或其时刻,如图12中的时刻136。使用附图标记140示意性地指示频谱。如图13所示出的,其一些部分142被量化成零。图13显示高频部分18以及频谱的140频谱线细分成由大括号指示的比例因子频带。使用“x”、“b”以及“e”,图13示意性地示出三个样本值12描述时刻136中的高频部分

18内的频谱包络,一个针对每个比例因子频带。对应于这些样本值e、b及x 的每个比例因子频带内,例如通过从完整的频率区间130的低频部分146的频谱再生,然后通过根据或使用样本值e、b及x缩放人工精细结构144以调整所产生的频谱的能量,精细结构确定器82生成频谱140的至少零量化部分142内的精细结构,如阴影区域144所示出的。有趣的是,或高频部分18的比例因子频带之内或之间,存在频谱140的非零量化部分148,因此,根据图12使用智能间隙填充,定位频谱140内的峰值是可行的,即使以频谱线分辨率在完整的频率区间130的高频部分18中及在任何频谱线位置处,有机会使用样本值x、b及e填充零量化部分142用于成形插入这些零量化部分142内的精细结构。

[0070] 最后,图14显示当根据图12和13的描述而具现时的用于馈送图9的参数化解码器的可能的参数化编码器。特别地,在此情况下,参数化编码器可包括变换器150,其用于在频谱上将回传音频信号152分解成覆盖完整的频率区间130的完整的频谱图。可使用能变化变换长度的重叠变换。频谱线编码器154以频谱线分辨率对此频谱图进行编码。至此,频谱线编码器154从变换器150接收高频部分18以及剩余的低频部分,两部分无间隙且没有覆盖完整的频率区间130的重叠。参数化高频编码器156仅从变换器150接收频谱图132 的高频部分18,并至少生成数据流88,即描述高频部分18内的频谱包络的样本值。

[0071] 即,根据图12-14的实施例,音频信号的频谱图132被频谱线编码器154编码至数据流158。因此,频谱线编码器154可对每完整的区间130的频谱线、每个时刻或帧136,编码一个频谱线值。图12中的小箱子160显示这些频谱线值。沿着频谱轴16,频谱线可被分组至比例因子频带。换句话说,频率区间16可被细分成由多组频谱线组成的比例因子频带。频谱线编码器154可选择用于每时刻内的每个比例因子频带的比例因子,以缩放通过数据流158编码的量化的频谱线值160。以比由时刻和规律地布置有频谱线值160的频谱线所定义的频谱时间网格至少较粗糙的且可与由比例因子分辨率所定义的栅相一致的频谱时间分辨率,参数化高频编码器156描述高频部分18内的频谱包络。有趣的是,根据其落入的比例因子频带的比例因子缩放的非零量化的频谱线值160,可以以频谱线分辨率散置在高频部分18内的任何位置,因此,使用描述高频部分内的频谱包络的样本值,在频谱成形器84内,它们幸免于在解码侧的高频率合成,正如精细结构确定器82和频谱成形器84 将其精细结构合成和成形限制在频谱图132的高频部分18内的零量化部分142。因此,在一方面在比特率消耗以及另一方面能获得品质之间产生非常有效的折中方案。

[0072] 如图14的虚线箭头所标示的,在164处所指示的,例如,频谱线编码器154可向参数化高频编码器156通知如可从数据流158重建的频谱图132的可重建版本,其中参数化高频编码器156使用此信息控制样本值12的生成和/或由样本值12的频谱包络10的表现的频谱时间分辨率。

[0073] 总结上述内容,相比于[2]及[3]以样本值表示频谱线的平均值,以上实施例的优点在于,频谱包络的样本值的特定属性。以上所概述的所有实施例中,变换可使用MDCT,且因此逆MDCT可用于所有的逆变换。在任何情况下,频谱包络的样本值更“平滑”许多,且与对应复频谱线的平均振幅线性相关。此外,根据至少一些的以上实施例,以下被称为SFE 值的频谱包络的样本值,确实为dB域或更通常地为对数域,其为对数表示。相比于用于频谱线的线性域或幂律域内的值,此进一步改进“平滑度”。例如,在AAC中幂律指数为0.75。相比于[4],在至少一些实施例中,频谱包络样本值在对数域中,且编码分布的属性及结构显著地

不同(取决于其振幅,一个对数域的值通常映射至指数增加的数目的线性域的值)因此,上述实施例中的至少一些的优点在于,在上下文(通常出现少量的上下文)的量化中及在对每个上下文中的分布的尾部(每个分布的尾部较广)进行编码中的对数表示。相比于[2],一些以上实施例基于与在计算量化的上下文中使用的相同的数据,额外地在每个上下文中使用固定或自适应的线性预测。此方法在上下文的数量剧烈减少中是有用的且仍获得最优性能。例如,相比于[4],在至少一些实施例中,对数域中的线性预测具有显著不同的用法及意义。例如,允许完美地预测不变的能量频谱区及信号的淡入及淡出频谱区。相比于[4],一些上述实施例使用算术编码,其使用从表示性训练数据集合抽取出的信息允许对任意分布进行最优编码。相比于[2],其也使用算术编码,但根据以上实施例,预测误差值而非原始值被编码。而且,在以上实施例中,无需使用比特平面编码。然而,对于每个整数值,比特平面编码需要几个算术编码步骤。相较之下,根据以上实施例,频谱包络的每个样本值可在一个步骤内编码/解码,如以上所概述的,其包含对于全部样本值分布的中心之外的值,可选择使用逸出编码,这将更快。

[0074] 再次简要地总结参数解码器支持IGF的实施例,如以上关于图9、12及13所述的,根据此实施例,精细结构确定器82用于使用频谱线状编码(其使用频谱预测和/或频谱熵上下文推导),以推导第一频率区间130(即完整的频率区间)内的音频信号的频谱图的精细结构132。频率线状解码表示这样的事实:精细结构确定器82从以频谱线间距而频谱地布置的数据流接收频谱线值160,借此在对应于各个时间部分的每时刻形成频谱136。例如,频谱预测的使用可涉及沿着频谱轴16的此频谱线值的差分编码,即仅频谱上直接先于频谱线值的差值被从数据流中解码,然后被加至此前值。频谱熵上下文推导可表示这样的事实:用于熵解码各个频谱线值160的上下文可取决于(即可被额外地选择,基于)当前解码的频谱线值160的频谱时间邻近区域中或至少频谱邻近区域中的已解码的频谱线值。为了填充精细结构的零量化的部分142,精细结构确定器82可使用人工随机噪声生成和/或频谱再生。精细结构确定器82仅在第二频率区间18(例如,其可限制在整体频率区间130的高频部分)内执行此。例如,部分频谱再生可从剩余频率部分146取得。然后,频谱成形器执行因此根据由零量化的部分处的样本值12所描述的频谱包络而获得的精细结构的成形。显著地,区间18内的精细结构的非零量化的部分对于成形后的精细结构的结果的贡献与实际频谱包络10不相关。这意味着:任一人工随机噪声生成和/或频谱再生,即填充,完全地限制于零量化的部分142,以使得在最终的精细结构频谱中,仅部分142已由人工随机噪声生成和/或使用频谱包络成形的频谱再生所填充,而非零贡献148维持,因为它们散置于部分142之间,或可选地所有人工随机噪声生成和/或频谱再生结果,即各个合成的精细结构也以加法方式放置在部分148,然后根据频谱包络10来成形产生的合成的精细结构。然而,即使在此情况下,维持初始解码的精细结构的非零量化的部分148的贡献。

[0075] 关于图12-14的实施例,最后应注意的是,关于这些附图所描述的智能间隙填充(IGF)过程或概念,显著地改进即使以非常低的比特率编码的信号的品质,由于通常不足的比特预算,高频区域18中的频谱的重要部分被量化成零。为了尽可能保持上频率区域18的精细结构,IGF信息、低频区域用作自适应地替代高频区域中的大部分被量化成零的目的区域(即区域142)的来源。为了实现良好的感知品质,重要的需求是,频谱系数的解码的能量包络与原始信号的解码的能量包络相匹配。为了实现此,从一个或多个连续的AAC比例因

子频带在频谱系数上计算平均频谱能量。产生的值为描述频谱包络的样本值12。使用由比例因子频带所定义的边界来计算平均是由已存在的将那些边界小心调整成临界频带的片段所激发,该临界频带对于人类听觉是特有的。如上所述,平均能量可被使用公式(该公式可以,例如类似于已知的用于AAC比例因子的公式)而变换成对数(如dB比例表示),然后被均匀地量化。在IGF中,根据请求的总比特率,可选择地使用不同的量化准确性。平均能量构成由IGF所生成的信息的重要部分,因此数据流88内的其高效的表示对于IGF 概念的整体性能是非常重要的。

[0076] 尽管已在装置的上下文中描述一些方面,但显然,这些方面亦表示对应方法的描述,其中区块或装置对应于方法步骤或方法步骤的特征。类似地,在方法步骤的上下文中所描述的方面亦表示对应装置的对应区块或项目或特征的描述。一些或所有的方法步骤可由(或使用)硬件装置执行,如像是微处理器、可编程的计算机或是电子电路。在一些实施例中,最重要的方法步骤可由此种装置执行。

[0077] 取决于特定的实施要求,本发明的实施例可以硬件或软件实施。可使用具有存储于其上的电子可读控制信号的数字存储介质,例如软性磁盘、硬盘、DVD、蓝光、CD、ROM、PROM、EPROM、EEPROM或闪存,执行实施方案,这些电子可读控制信号与可编程计算机系统协作(或能够协作)以使得执行各个方法。因此,数字存储介质可是计算机可读的。

[0078] 根据本发明的一些实施例包含具有电子可读控制信号的非暂时性数据载体,这些电子可读控制信号能够与可编程计算机系统协作,使得执行本文中所描述的方法中的一个。

[0079] 通常,本发明的实施例可被实施为具有程序代码的计算机程序产品,当计算机程序产品执行于计算机上时,程序代码操作性地用于执行这些方法中的一个。程序代码可(例如)储存于机器可读载体上。

[0080] 其他实施例包含储存于机器可读载体上的用于执行本文中所描述的方法中的一个的计算机程序。

[0081] 换言之,因此,本发明方法的实施例为具有程序代码的计算机程序,当计算机程序执行于计算机上时,该程序代码用于执行本文中所描述的方法中的一个。

[0082] 因此,本发明方法的另一实施例为包含记录于其上的,用于执行本文中所描述的方法中的一个的计算机程序的数据载体(或数字存储介质,或计算机可读介质)。数据载体、数字存储介质或记录的介质通常为有形的和/或非暂时性的。

[0083] 因此,本发明方法的另一实施例为表示用于执行本文中所描述的方法中的一个的计算机程序的数据流或信号序列。数据流或信号序列可例如用于经由数据通信连接(例如,经由因特网)而传送。

[0084] 另一实施例包含用于或经调适以执行本文中所描述的方法中的一个的处理构件,例如,计算机或可编程逻辑器件。

[0085] 另一实施例包含安装有用于执行本文中所描述的方法中的一个的计算机程序的计算机。

[0086] 根据本发明的另一实施例包含用以将用于执行本文中所描述的方法中的一个的计算机程序传输(例如电性或光学)到接收器的装置或是系统。例如,接收器可为计算机、移动装置、存储器装置或类似。例如,此装置或系统可包含用于将计算机程序传输至接收器的

文件服务器。

[0087] 在一些实施例中,可编程逻辑器件(例如,场可编程门阵列)可用于执行本文中所描述的方法的功能性中的一些或所有。在一些实施例中,场可编程门阵列可与微处理器协作,以便执行本文中所描述的方法中的一个。大体而言,优选地由任何硬件装置执行这些方法。

[0088] 上文所描述的实施例仅仅说明本发明的原理。应理解,对本文中所描述的配置及细节的修改及变型对本领域技术人员而言将是显而易见。因此,仅意欲由待决专利的权利要求的范围限制,而不由通过本文的实施例的描述及解释而提出的特定细节限制。

[0089] 参考文献

[0090] [1]International Standard ISO/IEC 14496-3:2005,Information technology-Coding of audio-visual objects-Part 3:Audio,2005.

[0091] [2]International Standard ISO/IEC 23003-3:2012,Information technology-MPEG audio technologies-Part 3:Unified Speech and Audio Coding, 2012.

[0092] [3]B.Edler and N.Meine:Improved Quantization and Lossless Coding for Subband Audio Coding,AES 118th Convention,May 2005.

[0093] [4]M.J.Weinberger and G.Seroussi:The LOCO-I Lossless Image Compression Algorithm: Principles and Standardization into JPEG-LS, 1999.Available online at http://www.hpl.hp.com/research/info_theory/loco/HPL-98-193R1.pdf

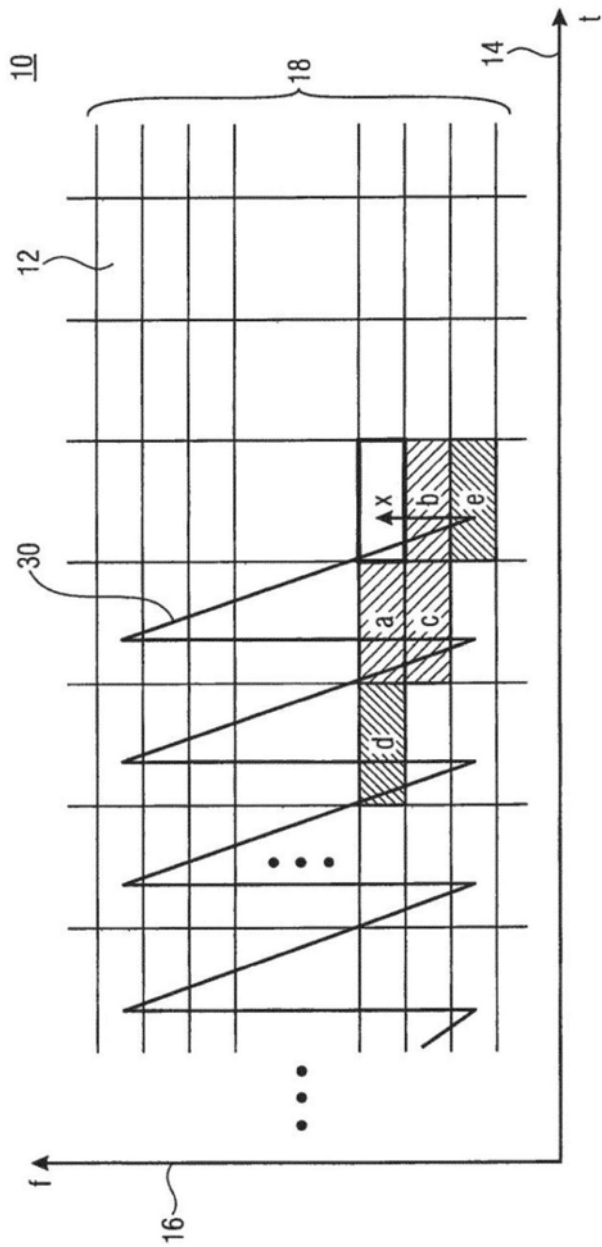


图1

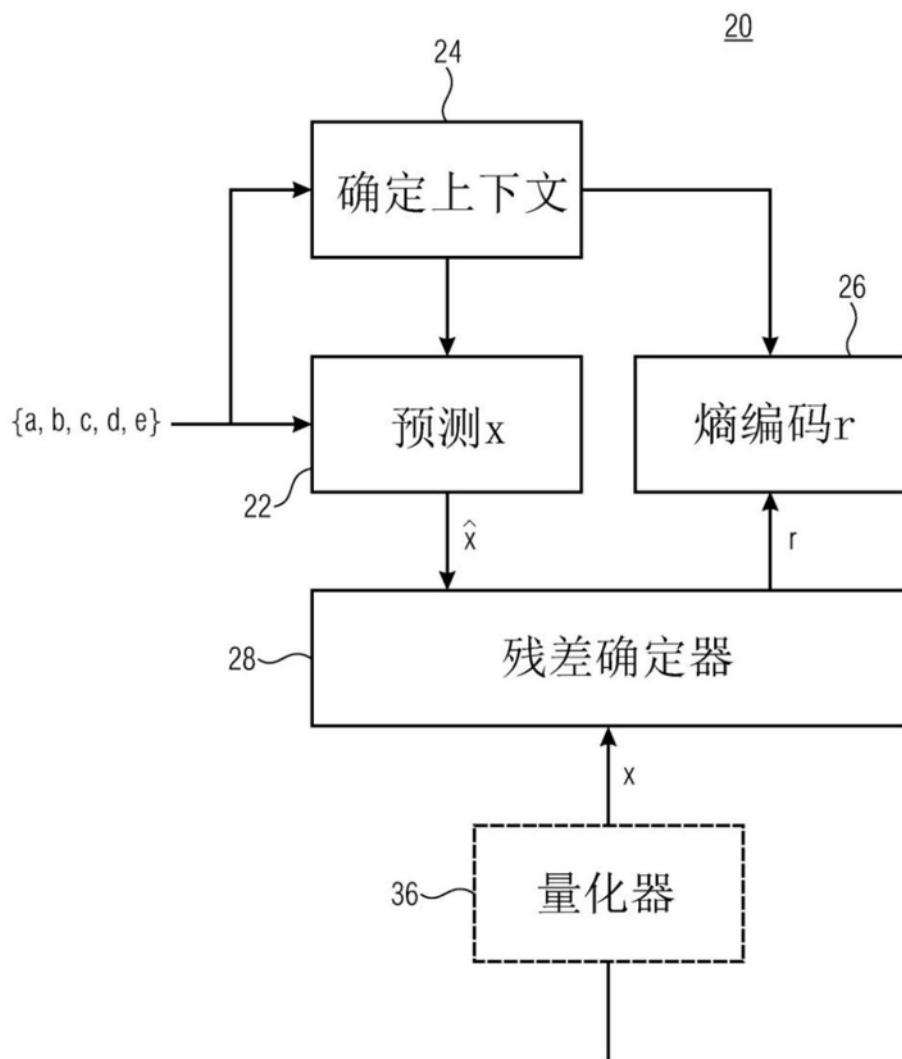


图2

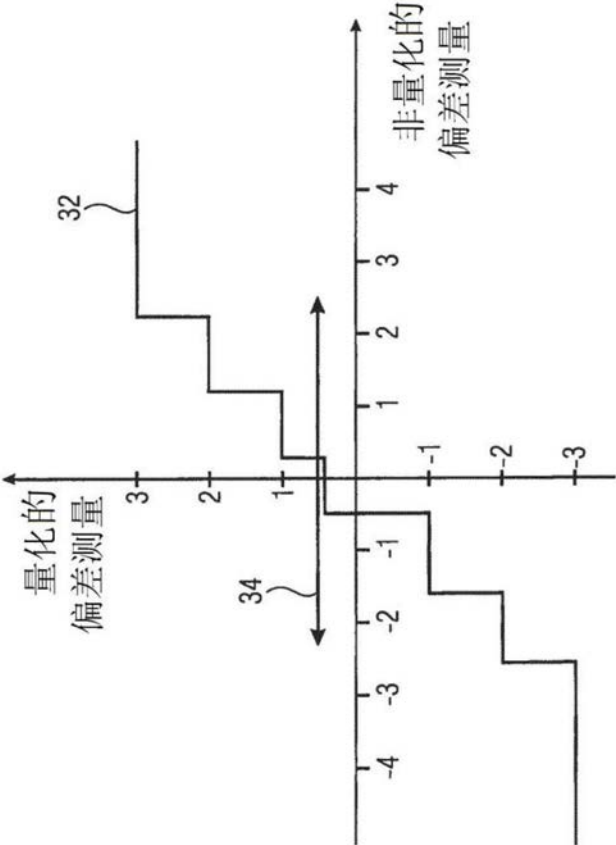


图3

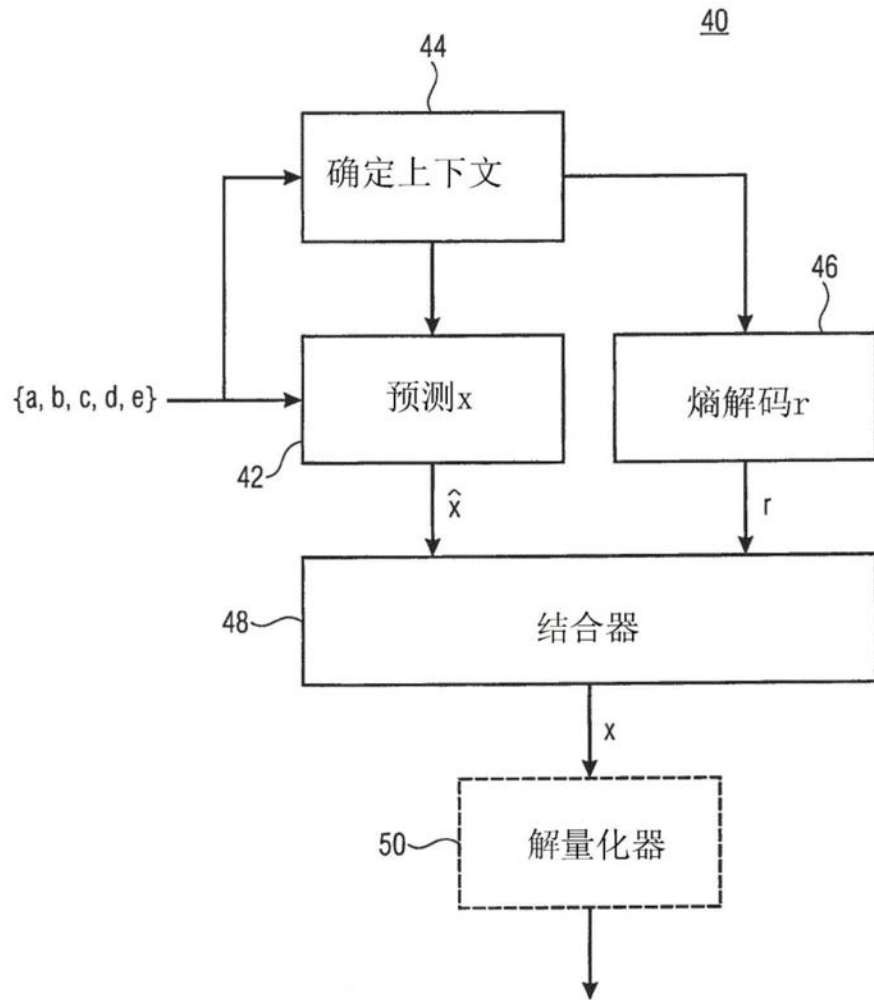


图4

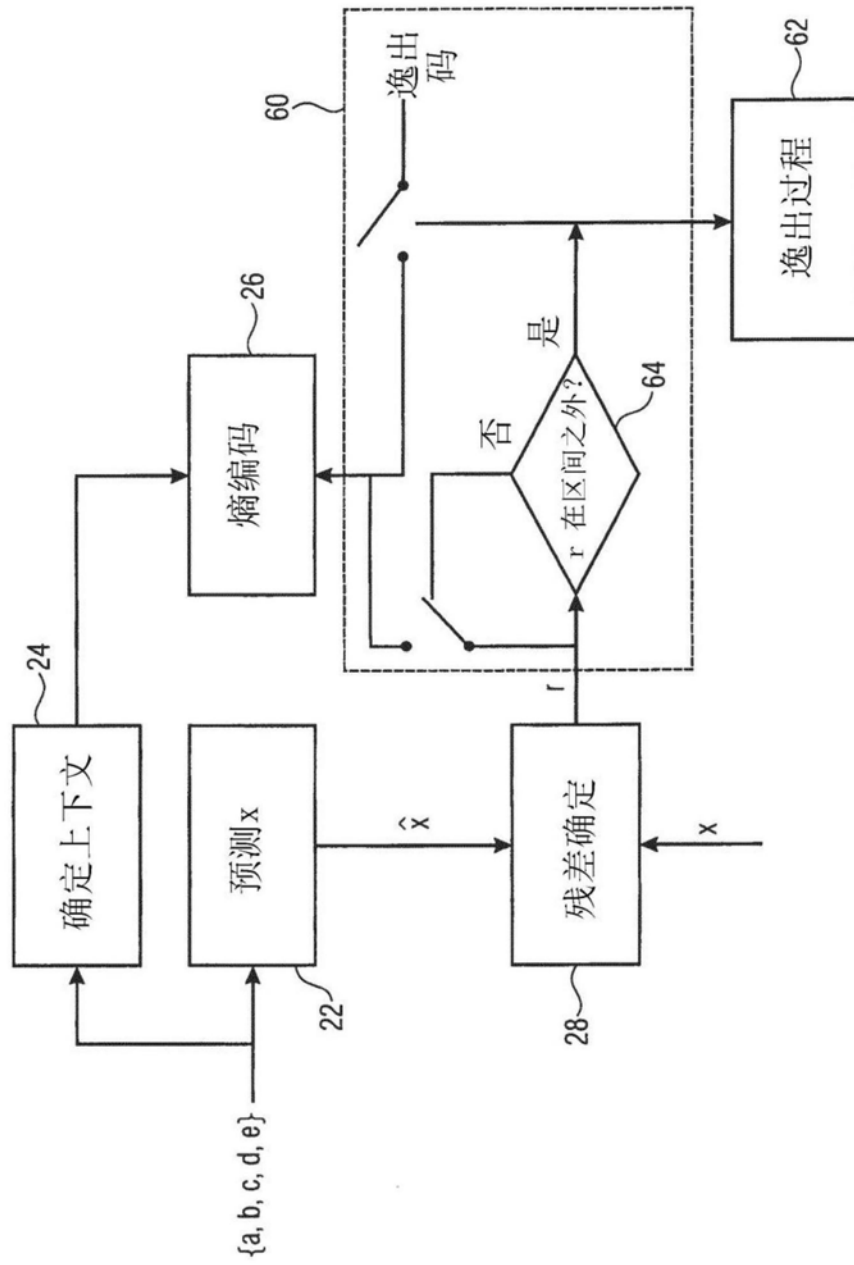


图5

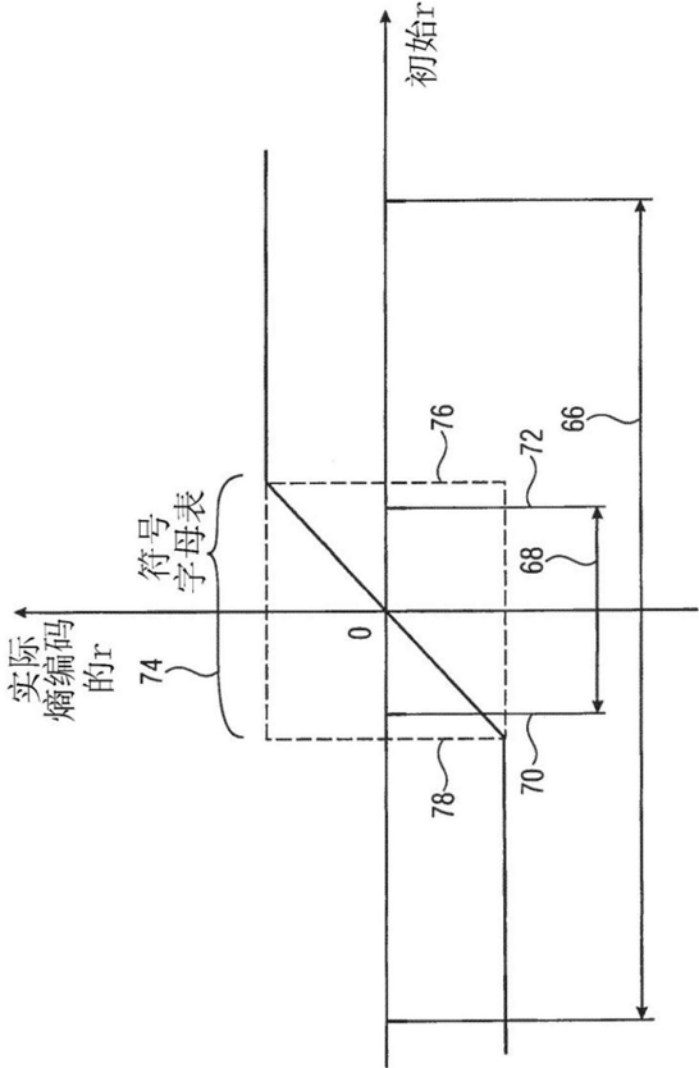


图6

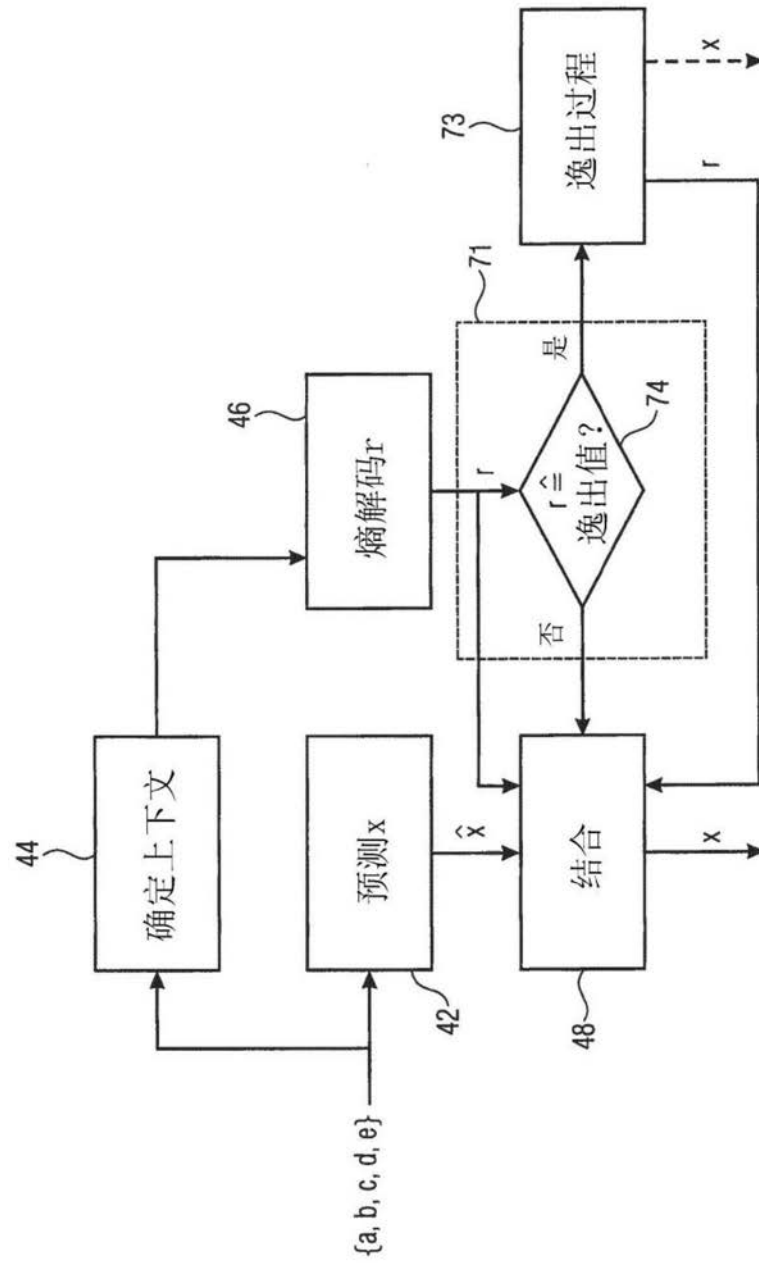


图7

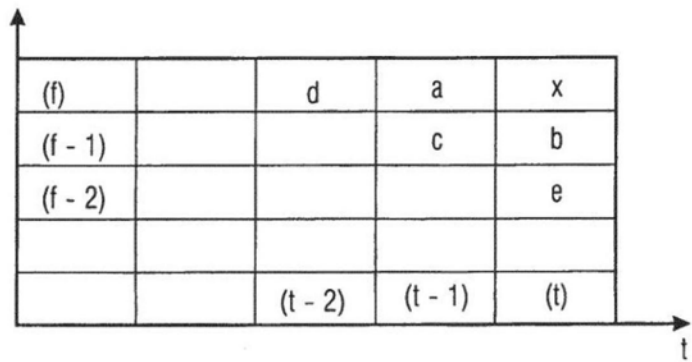


图8

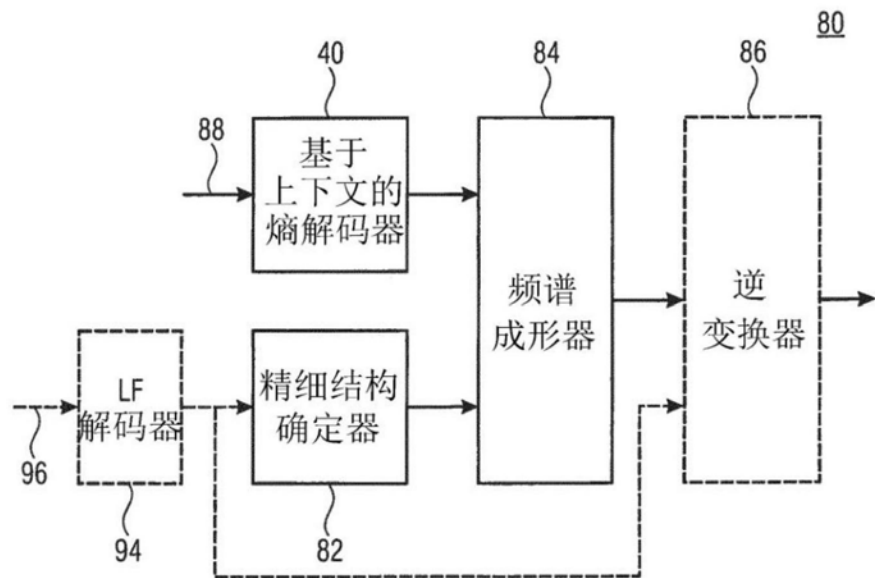


图9

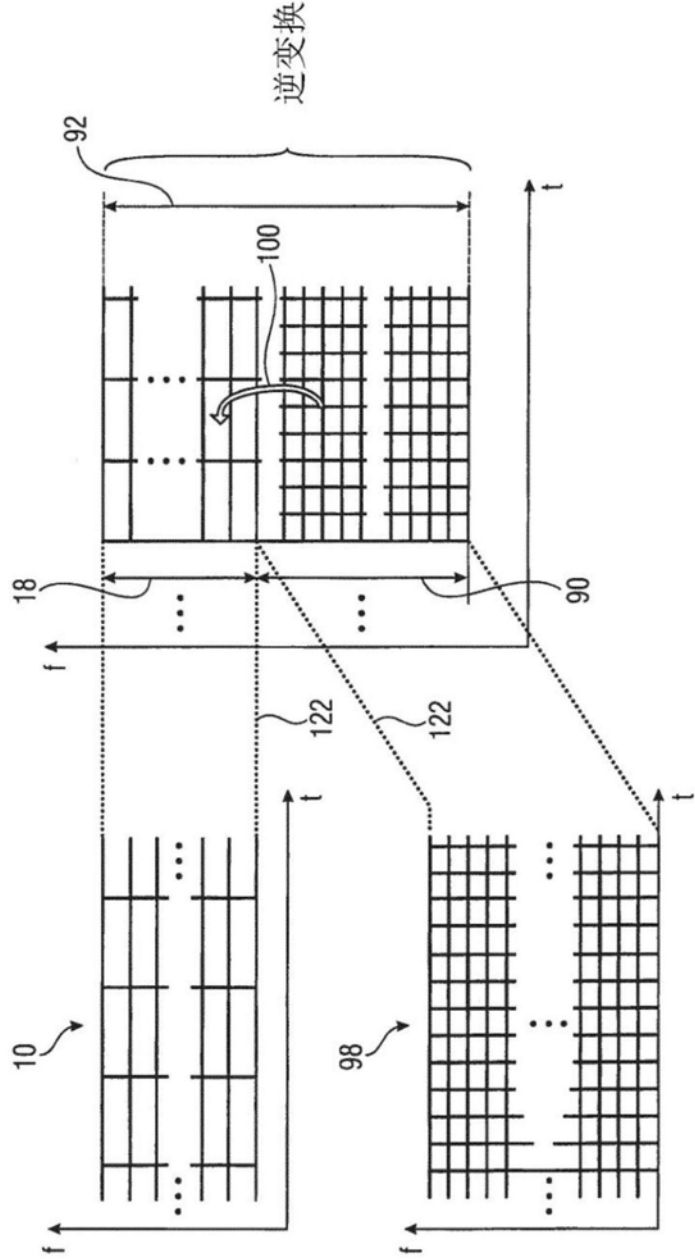


图10

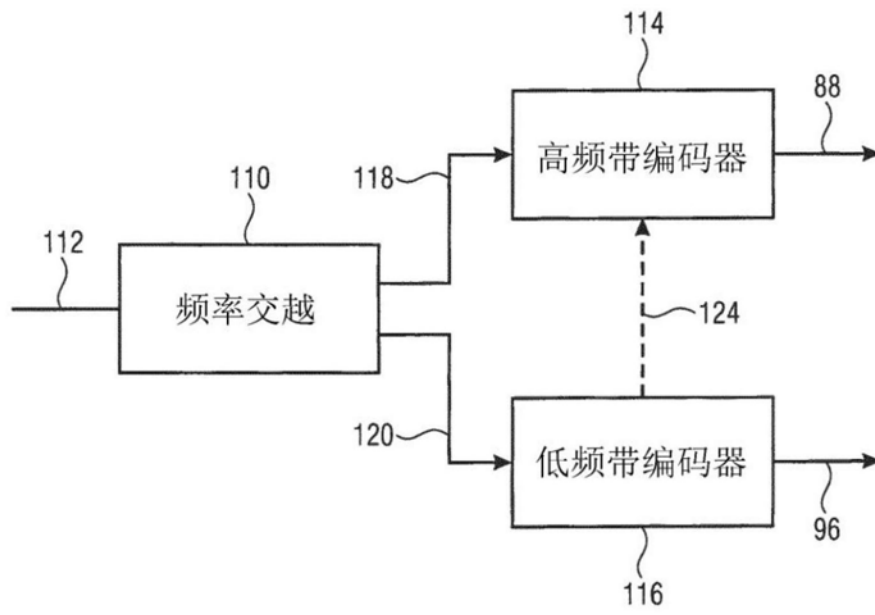


图11

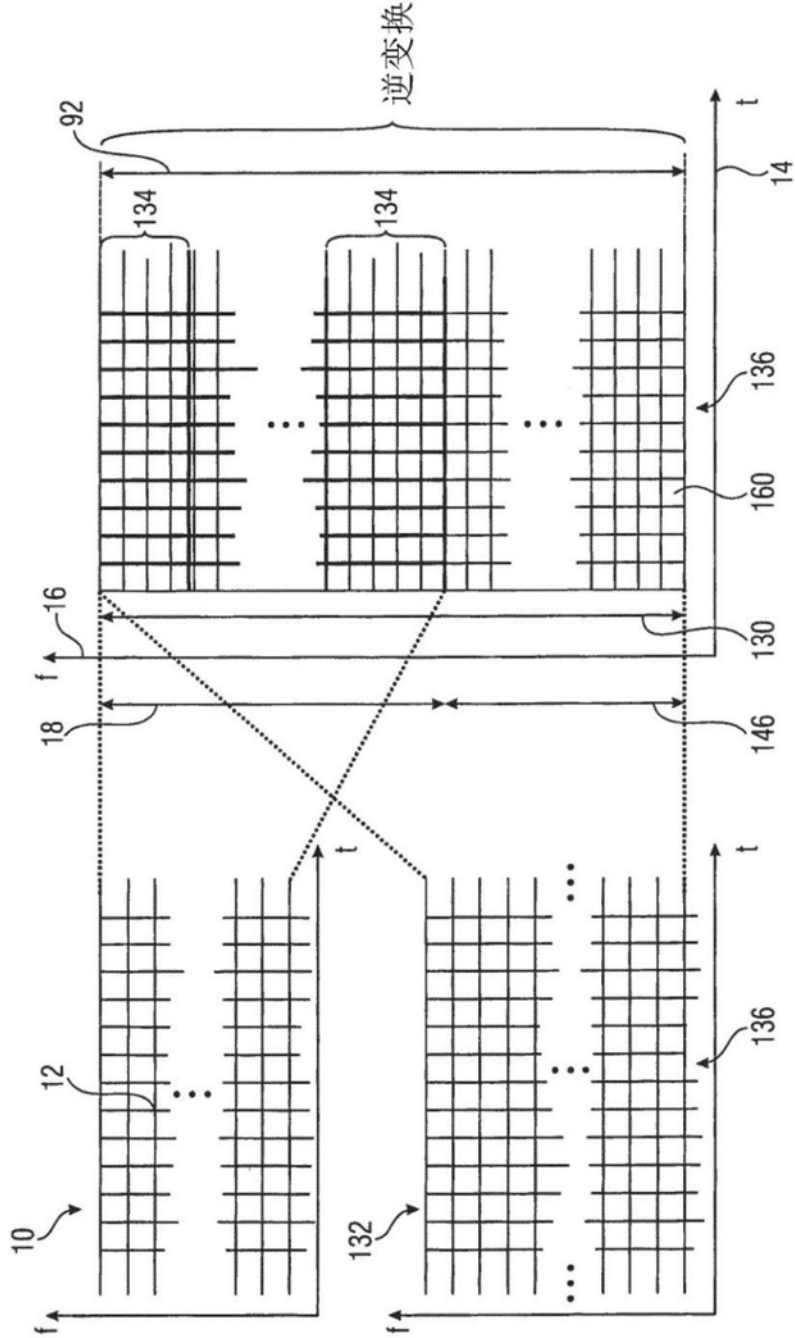


图12

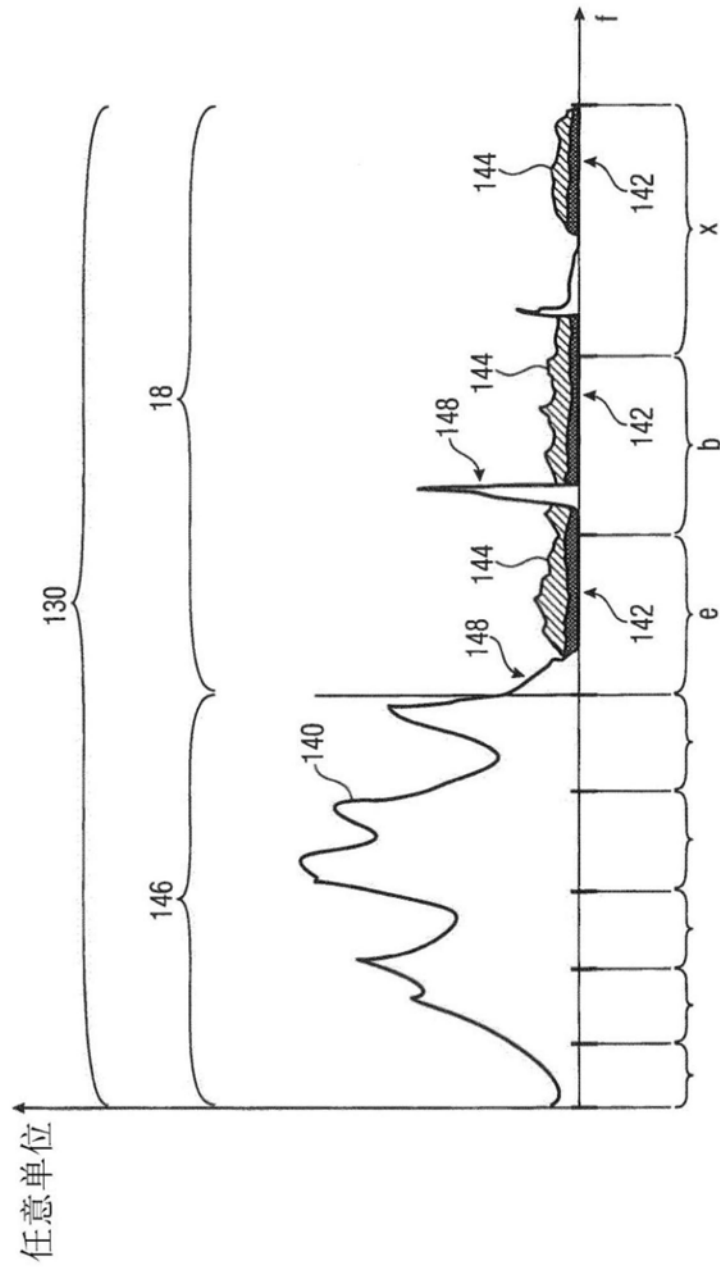


图13

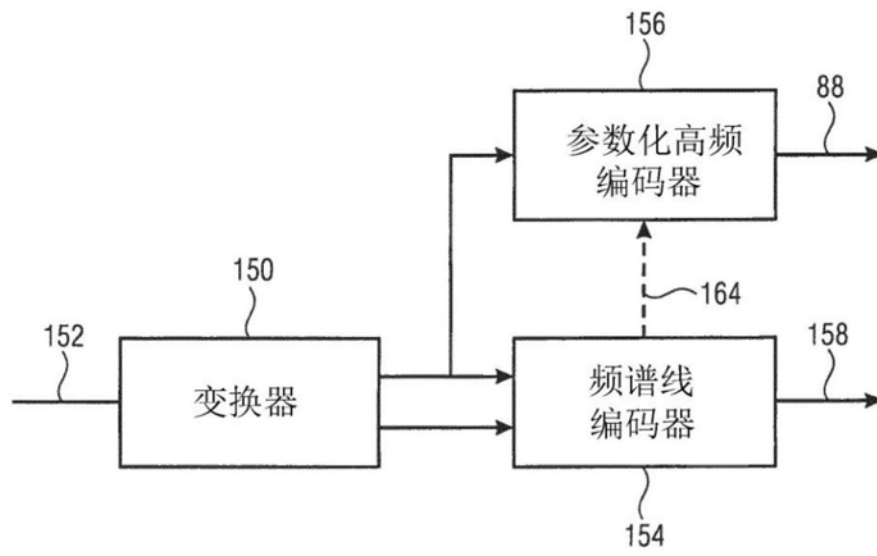


图14