



(12)发明专利

(10)授权公告号 CN 103733257 B

(45)授权公告日 2017.02.15

(21)申请号 201280037719.1

(22)申请日 2012.06.01

(65)同一申请的已公布的文献号
申请公布号 CN 103733257 A

(43)申请公布日 2014.04.16

(30)优先权数据
2011-121982 2011.06.01 RU

(85)PCT国际申请进入国家阶段日
2014.01.28

(86)PCT国际申请的申请数据
PCT/KR2012/004362 2012.06.01

(87)PCT国际申请的公布数据
W02012/165910 KO 2012.12.06

(73)专利权人 三星电子株式会社
地址 韩国京畿道水原市

(72)发明人 安顿·普波沃

康斯坦丁·奥斯波夫 朱基峴

(74)专利代理机构 北京铭硕知识产权代理有限公司 11286

代理人 曾世骁 胡江海

(51)Int.Cl.
G10L 19/02(2013.01)
G10L 19/00(2013.01)

(56)对比文件
CN 1784020 A, 2006.06.07,
CN 101317217 A, 2008.12.03,
JP 特开2000-132193 A, 2000.05.12,
ITU.Low-complexity, full-band audio
coding for high-quality, conversational
applications.《ITU-T G.719》.2008,

审查员 陈红红

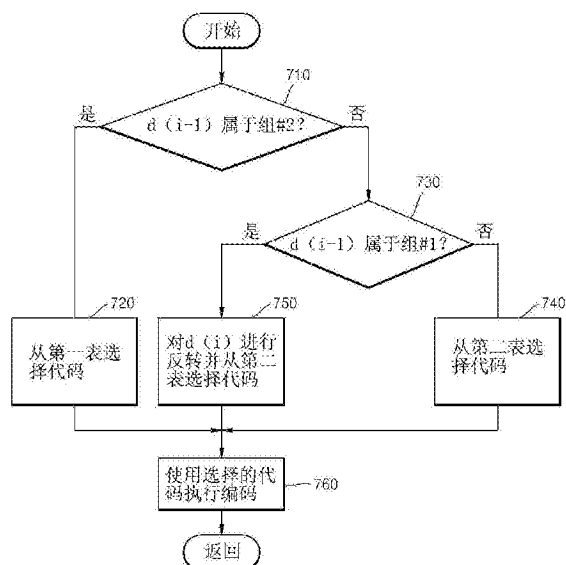
权利要求书1页 说明书14页 附图6页

(54)发明名称

音频编码方法和设备、音频解码方法和设备和采用音频编码方法和设备、音频解码方法和设备的多媒体装置

(57)摘要

提供一种音频编码方法。所述音频编码方法包括：针对音频频谱基于预定的子频带获取包络；基于预定的子频带对包络进行量化；获得相邻子频带的量化的包络之间的差值并通过将先前子频带的差值用作上下文来对当前子频带的差值进行无损编码。因此，对音频频谱的包络信息进行编码所需的比特的数量可被减少到有限的比特范围内，从而增加对实际频谱分量进行编码所需的比特的数量。



1. 一种音频编码方法,包括:

通过将音频信号从时域变换到频域来获得音频频谱的包络,其中,音频频谱包括多个子频带;

对包络进行量化以基于子频带获得量化索引;

获得相对于先前子频带的量化索引的当前子频带的差分量化索引;

对当前子频带的差分量化索引进行无损编码,

其中,无损编码的步骤包括:

通过使用先前子频带的差分量化索引来获得当前子频带的上下文;

基于与所述上下文相应的霍夫曼表对当前子频带的差分量化索引进行霍夫曼编码。

2. 如权利要求1所述的音频编码方法,其中,包络是相应的子频带的平均能量、平均幅度、功率和范数中的一个。

3. 如权利要求1所述的音频编码方法,其中,霍夫曼编码的步骤包括:将与所述上下文相应的差分量化索引划分为多个组中的一个组,并通过使用针对所述一个组定义的霍夫曼表来对当前子频带的差分量化索引执行霍夫曼编码。

4. 如权利要求3所述的音频编码方法,其中,通过使用针对所述多个组中的一些组共享的霍夫曼表来执行霍夫曼编码。

5. 如权利要求3所述的音频编码方法,其中,霍夫曼编码的步骤包括:当针对所述多个组中的一些组共享的霍夫曼表被使用时,通过反转处理过程来对当前子频带的差分量化索引进行处理。

6. 如权利要求1所述的音频编码方法,其中,霍夫曼编码的步骤包括:针对不存在先前子频带的第一子频带,按原样地对量化索引进行霍夫曼编码,并通过将第一子频带的量化索引与预定参考值之间的差用作上下文来对接着第一子频带的第二子频带的差分量化索引进行霍夫曼编码。

音频编码方法和设备、音频解码方法和设备和采用音频编码方法和设备、音频解码方法和设备的多媒体装置

技术领域

[0001] 与示例性实施例一致的设备和方法涉及音频编码/解码,更具体地讲,涉及一种能够在不增加复杂度和恢复的声音质量的恶化的情况下,通过减少在有限比特范围内对音频频谱的包络信息进行编码所需的比特的数量来增加对实际频谱分量进行编码所需的比特的数量的音频编码方法和设备、音频解码方法和设备、记录介质和采用该音频编码方法和设备、音频解码方法和设备的多媒体装置。

背景技术

[0002] 当对音频信号进行编码时,除实际频谱分量以外的附加信息(诸如,包络)可包括在比特流中。在这种情况下,通过在使损失最小化的同时减少分配用于附加信息的编码的比特的数量,分配用于实际频谱分量的编码的比特的数量可被增加。

[0003] 也就是说,当对音频信号进行编码或解码时,需要通过以特别低的比特率有效地使用有限数量的比特来构建在相应比特范围内具有最优声音质量的音频信号。

发明内容

[0004] 技术问题

[0005] 一个或更多个示例性实施例的多个方面提供一种音频编码方法和设备、音频解码方法和设备、记录介质和采用其的多媒体装置,其中,所述音频编码方法和设备能够在不增加复杂度和恢复的声音质量的恶化的情况下,在将对音频频谱的包络信息进行编码所需的比特的数量减少到有限的比特范围内的同时,增加对实际频谱分量进行编码所需的比特的数量。

[0006] 解决方案

[0007] 根据一个或更多个示例性实施例的一方面,提供一种音频编码方法,包括:针对音频频谱基于预定的子频带获取包络;基于预定的子频带对包络进行量化;获得相邻子频带的量化的包络之间的差值并通过将先前子频带的差值用作上下文来对当前子频带的差值进行无损编码。

[0008] 根据一个或更多个示例性实施例的一方面,提供一种音频编码设备,包括:包络获取单元,用于针对音频频谱基于预定的子频带获取包络;包络量化器,用于基于预定的子频带对包络进行量化;包络编码器,用于获得相邻子频带的量化的包络之间的差值,并通过将先前子频带的差值用作上下文来对当前子频带的差值进行无损编码;频谱编码器,用于对音频频谱进行量化和无损编码。

[0009] 根据一个或更多个示例性实施例的一方面,提供一种音频解码方法,包括:从比特流获得相邻子频带的量化的包络之间的差值,并通过将先前子频带的差值用作上下文来对当前子频带的差值进行无损解码;通过从作为无损解码的结果而被重构的当前子频带的差值获得基于子频带的量化的包络来执行反量化。

[0010] 根据一个或更多个示例性实施例的一方面,提供一种音频解码设备,包括:包络解码器,用于从比特流获得相邻子频带的量化的包络之间的差值,并通过将先前子频带的差值用作上下文来对当前子频带的差值进行无损解码;包络反量化器,用于通过从作为无损解码的结果而被重构的当前子频带的差值获得基于子频带的量化的包络来执行反量化;频谱解码器,用于对包括在比特流中的频谱分量进行无损解码和反量化。

[0011] 根据一个或更多个示例性实施例的一方面,提供一种包括编码模块的多媒体装置,其中,编码模块用于针对音频频谱基于预定的子频带获取包络,基于预定的子频带对包络进行量化,获得相邻子频带的量化的包络之间的差值并通过将先前子频带的差值用作上下文来对当前子频带的差值进行无损编码。

[0012] 多媒体装置还可包括解码模块,其中,解码模块用于从比特流获得相邻子频带的量化的包络之间的差值,并通过将先前子频带的差值用作上下文来对当前子频带的差值进行无损解码,通过从作为无损解码的结果而被重构的当前子频带的差值获得基于子频带的量化的包络来执行反量化。

[0013] 效果

[0014] 可在不增加复杂度和恢复的声音质量的恶化的情况下,通过减少在有限的比特范围内对音频频谱的包络信息进行编码所需的比特的数量,来增加对实际频谱分量进行编码所需的比特的数量。

附图说明

[0015] 通过下面结合附图的示例性实施例的描述,这些和其他方面将变得清楚和更容易理解,其中:

[0016] 图1是根据示例性实施例的数字信号处理设备的框图;

[0017] 图2是根据另一示例性实施例的数字信号处理设备的框图;

[0018] 图3A和图3B分别示出在量化分辨率是0.5且量化步长是3.01时的相互比较的非优化的对数尺度和优化的对数尺度;

[0019] 图4A和图4B分别示出在量化分辨率是1且量化步长是6.02时的相互比较的非优化的对数尺度和优化的对数尺度;

[0020] 图5A和图5B是分别示出相互比较的非优化的对数尺度的量化结果和优化的对数尺度的量化结果的曲线图;

[0021] 图6是示出在先前子频带的量化差(delta)值被用作上下文时选择的三个组的概率分布的曲线图;

[0022] 图7是示出根据示例性实施例的图1的数字信号处理设备的包络编码器中的基于上下文的编码处理的流程图;

[0023] 图8是示出根据示例性实施例的图2的数字信号处理设备的包络解码器中的基于上下文的解码处理的流程图;

[0024] 图9是根据示例性实施例的包括编码模块的多媒体装置的框图;

[0025] 图10是根据示例性实施例的包括解码模块的多媒体装置的框图;

[0026] 图11是根据示例性实施例的包括编码模块和解码模块的多媒体装置的框图。

具体实施方式

[0027] 示例性实施例可允许各种改变或修改以及形式上的各种改变,特定实施例将被示出在附图中并在说明书中被详细描述。然而,应理解特定实施例不将本发明构思限制为特定公开形式而是包括本发明构思的精神和技术范围内的所有修改的、等同的或替代的实施例。在以下描述中,由于公知的功能或构造将在不必要的细节上使本发明构思不清楚,因此不被详细描述。

[0028] 尽管诸如“第一”和“第二”的术语可用于描述各种元件,但所述元件可不被所述术语限制。所述术语可用于将特定元件与另一元件分开。

[0029] 在本申请中使用的术语仅用于描述特定实施例,而不具有任何限制本发明构思的意图。尽管在考虑在本发明构思中的功能的同时将当前尽可能广泛使用的通用术语选作在本发明构思中使用的术语,但它们可根据本领域的普通技术人员的意图、先前使用或新技术的出现改变。另外,在特定情况下,可使用由申请人有意地选择的术语,在这种情况下,将在本发明构思的相应描述中公开这些术语的意义。因此,在本发明构思中使用的术语不应仅由术语的名称而应由术语的含义和整个本发明构思中的内容来定义。

[0030] 除非在上下文中单数的表达与复数的表达明显彼此不同,否则单数的表达包括复数的表达。在本申请中,应理解诸如“包括”和“具有”的术语用于指示实施的特征、数字、步骤、操作、元件、部件或它们的组合的存在,而不是预先排除一个或多个其他特征、数字、步骤、操作、元件、部件或它们的组合的存在或添加的可能性。

[0031] 在下文中,将参照示出本发明构思的示例性实施例的附图来更全面地描述本发明构思。附图中的相同的标号表示相同的元件,因此将省略它们的重复的描述。

[0032] 当诸如“…中的至少一个”的表述在一列元件之后时,其修饰整列元件而不修饰列出的单个元件。

[0033] 图1是根据示例性实施例的数字信号处理设备100的框图。

[0034] 图1中示出的数字信号处理设备100可包括变换器110、包络获取单元120、包络量化器130、包络编码器140、频谱归一化器150和频谱编码器160。数字信号处理设备100的组件可被集成到至少一个模块中,并可由至少一个处理器实现。这里,数字信号可指示媒体信号,诸如,视频、图像、音频或语音、或指示通过对音频和语音进行合成来获得的信号的声音,但在下文中,为便于描述,数字信号通常指示音频信号。

[0035] 参照图1,变换器110可通过将音频信号从时域变换到频域来产生音频频谱。可通过使用诸如改进的离散余弦变换(MDCT)的各种公知方法来执行时域到频域的变换。例如,用于时域的音频信号的MDCT可使用等式1来执行。

$$[0036] \quad x_i = \sum_{j=0}^{2N-1} h_j s_j \cos[\pi(j+(N+1)/2)(i+1/2)/N], \quad i = 0, \dots, N-1 \quad (1)$$

[0037] 在等式1中,N表示包括在单个帧中的采样的数量(即,帧大小), h_j 表示应用的窗口, s_j 表示时域的音频信号, x_i 表示MDCT系数。可选择地,可使用正弦窗口(例如, $h_j = \sin[\pi(j+1/2)/2N]$)替代等式1的余弦窗口。

[0038] 由变换器110获得的音频频谱的变换系数(例如,MDCT系数 x_i)被提供给包络获取单元120。

[0039] 包络获取单元120可从自变换器110提供的变换系数基于预定的子频带获取包络值。子频带是对音频频谱的采样进行分组的单位,并可通过反映临界频带具有均匀或不均匀的长度。当子频带具有不均匀的长度时,子频带可被设置为使得针对一帧包括在每个子频带中的从开始采样到最后采样的采样的数量渐渐增加。另外,当支持多比特率时,可进行设置使得包括在不同比特率的相应子频带的每个子频带中的采样的数量相同。可预先确定包括在一个帧中的子频带的数量或包括在每个子频带中的采样数量。包络值可指示包括在每个子频带中的变换系数的平均幅度、平均能量、功率或标准(norm)值。

[0040] 每个子频带的包络值可使用等式2来计算,但不限于此。

$$[0041] \quad n = \sqrt{\frac{1}{w} \sum_{i=1}^w x_i^2} \quad (2)$$

[0042] 在等式2中,w表示包括在子频带中的变换系数的数量(即,子频带大小), x_i 表示变换系数,n表示子频带的包络值。

[0043] 包络量化器130可以以优化的对数尺度对每个子频带的包络值n进行量化。由包络量化器获得的每个子频带的包络值n的量化索引 n_q 可使用例如等式3来获得。

$$[0044] \quad n_q = \left\lfloor \frac{1}{r} \log_c n + \frac{b}{r} \right\rfloor \quad (3)$$

[0045] 在等式3中,b表示取整系数,在优化之前的其初始值是 $r/2$ 。另外,c表示对数尺度的底数,r表示量化分辨率。

[0046] 根据实施例,包络量化器130可可变地改变与每个量化索引相应的量化区域的左边界和右边界,从而与每个量化索引相应的量化区域中的总量化误差被最小化。为此,取整系数b可被调整,使得在量化索引与和每个量化索引相应的量化区域的左边界和右边界之间获得的左量化误差和右量化误差彼此相同。以下,描述包络量化器130的详细操作。

[0047] 可通过等式4执行每个子频带的包络值n的量化索引 n_q 的反量化。

$$[0048] \quad \tilde{n} = c^{r n_q} \quad (4)$$

[0049] 在等式4中, $\tilde{n}$ 表示每个子频带的反量化的包络值,r表示量化分辨率,c表示对数尺度的底数。

[0050] 由包络量化器130获得的每个子频带的包络值n的量化索引 n_q 可被提供给包络编码器140,每个子频带的反量化的包络值 $\tilde{n}$ 可被提供到频谱归一化器150。

[0051] 尽管未示出,但基于子频带获得的包络值可用于对归一化的频谱(即,归一化的系数)进行编码所需的比特分配。在这种情况下,基于子频带量化和无损编码的包络值可包括在比特流中,并被提供给解码设备。与使用基于子频带获得的包络值的比特分配关联地,可应用反量化的包络值以在编码设备和相应的解码设备中使用相同的处理。

[0052] 例如,当包络值是标准值时,可使用基于子频带的标准值计算掩蔽(masking)阈值,并可使用掩蔽阈值预测感知上所需的比特数量。也就是说,掩蔽阈值是与恰可察觉失真(JND)相应的值,当量化噪声小于掩蔽阈值时,感知噪声可能不被感测到。因此,可使用掩蔽阈值计算感测不到感知噪声所需的比特的最小数量。例如,可使用标准值相对于基于子频带的掩蔽阈值的比例来计算信号掩蔽比(SMR),可使用对于SMR的6.025dB $\approx$ 1比特的关系预测满足掩蔽阈值的比特数量。尽管预测的比特数量是感测不到感知噪声的所需的比特的最小数量,但在压缩方面不需要使用多于预测的比特数量的比特,所以预测的比特的数量可

被认为是基于子频带允许的最大数量的比特(在下文中,被称为可允许的比特数量)。可以以小数点单位表示每个子频带的可允许的比特数量,但不限于此。

[0053] 另外,可使用以小数点为单位的标准值来执行基于子频带的比特分配,但不限于此。从具有较大标准值的子频带顺序地分配比特,可调整分配的比特,从而通过基于每个子频带的感知上的重要性对每个子频带的标准值进行加权来将更多的比特分配到感知上更重要的子频带。可通过例如在ITU-T G.719中定义的心理声学加权(psycho-acoustic weighting)来确定感知上的重要性。

[0054] 包络编码器140可获得用于从包络量化器130提供的每个子频带的包络值 n 的量化索引 n_q 的量化差值,可基于用于量化差值的上下文执行无损编码,可将无损编码结果包括在比特流中,并可发送和存储比特流。先前子频带的量化差值可用作上下文。以下描述包络编码器140的详细操作。

[0055] 频谱归一化器150通过使用每个子频带的反量化包络值 $\tilde{n} = e^{-m}$ 将变换系数归一化为 $y_i = x_i / \tilde{n}$,从而使频谱平均能量为1。

[0056] 频谱编码器160可执行归一化的变换系数的量化和无损编码,可将量化和无损编码结果包括在比特流中,并可发送和存储比特流。这里,频谱编码器160可通过使用根据基于子频带的包络值最终确定的可允许的比特数量,来执行归一化的变换系数的量化和无损编码。

[0057] 归一化的变换系数的无损编码可使用例如阶乘脉冲编码(FPC)。FPC是通过使用单位幅度脉冲来有效地对信息信号进行编码的方法。根据FPC,可使用四个分量(即,非零脉冲位置的数量、非零脉冲的位置、非零脉冲的幅度和非零脉冲的符号)来表示信息内容。详细地,FPC可基于均方误差(MSE)标准确定 $\tilde{y} = \{\tilde{y}_1, \tilde{y}_2, \tilde{y}_3, \dots, \tilde{y}_{k-1}\}$ 的最优解,其中,在MSE

中在满足 $m = \sum_{i=0}^{k-1} |\tilde{y}_i|$ (m 表示单位幅度脉冲的总数量)的同时使子频带的原始矢量 y 和FPC矢量 $\tilde{y}$ 之间的差最小。

[0058] 可通过使用如等式5中的拉格朗日函数找到条件极值来获得最优解。

$$[0059] \quad L = \sum (y_i - \tilde{y}_i)^2 + \lambda (\sum \tilde{y}_i - m)$$

$$[0060] \quad \begin{cases} \frac{\partial L}{\partial \tilde{y}_i} = 2\tilde{y}_i - 2y_i + \lambda \tilde{y}_i = 0 \\ \frac{\partial L}{\partial \lambda} = \sum \tilde{y}_i - m = 0 \end{cases}$$

$$[0061] \quad \tilde{y}_i = \text{Round}\left(\frac{y_i m}{\sum y_i}\right) \quad (5)$$

[0062] 在等式5中, L 表示拉格朗日函数, m 表示子频带中的单位幅度脉冲的总数量, λ 表示作为拉格朗日乘子(即,优化系数)的用于寻找给定函数的最小值的控制参数, y_i 表示归一化的变换系数, $\tilde{y}_i$ 表示在位置 i 处所需的脉冲的最优数量。

[0063] 当无损编码使用FPC被执行时,基于子频带获得的总集的 $\tilde{y}$ 可被包括在比特流中

并被发送。另外,还可将最优乘子包括在比特流中并进行发送,其中,最优乘子用于使每个子频带中的量化误差最小化并执行平均能量的对齐(alignment)。可通过等式6获得最优乘子。

$$[0064] \quad D = \frac{\sum (y_i - G \tilde{y}_i)^2}{\sum y_i^2} \rightarrow 0$$

$$[0065] \quad \frac{\partial D}{\partial G} = 0$$

$$[0066] \quad G = \frac{\sum y_i \tilde{y}_i}{\sum \tilde{y}_i^2} \quad (6)$$

[0067] 在等式6中,D表示量化误差,G表示最优乘子。

[0068] 图2是根据示例性实施例的数字信号解码设备200的框图。

[0069] 图2中示出的数字信号解码设备可包括包络解码器210、包络反量化器220、频谱解码器230、频谱反归一化器240和反变换器250。数字信号解码设备200的组件可集成到至少一个模块中并由至少一个处理器实现。这里,数字信号可指示媒体信号,诸如,视频、图像、音频或语音、或指示通过对音频和语音进行合成而获得的信号的声音,但在下文中,数字信号通常指示音频信号以与图1的编码设备相应。

[0070] 参照图2,包络解码器210可经由通信信道或网络接收比特流,对包括在比特流中的每个子频带的量化差值进行无损解码,并重构每个子频带的包络值的量化索引 n_q 。

[0071] 包络反量化器220可通过对每个子频带的包络值的量化索引 n_q 进行反量化来获得反量化的包络值 $\tilde{n} = c^{m_q}$ 。

[0072] 频谱解码器230可通过对接收的比特流进行无损解码和反量化来重构归一化的变换系数。例如,当编码设备已使用FPC时,包络反量化器可对关于每个子频带的总集的 $\tilde{y}_i$ 进行无损解码和反量化。可按照等式7使用最优乘子G来执行每个子频带的平均能量对齐。

$$[0073] \quad \tilde{y}_i = \tilde{y}_i G \quad (7)$$

[0074] 如同在图1的频谱编码器160中,频谱解码器230可通过使用基于根据子频带的包络值最终确定的可允许的比特数量来执行无损解码和反量化。

[0075] 频谱反归一化器240可通过使用从包络反量化器220提供的反量化的包络值,对从包络解码器210提供的归一化的变换系数进行反量化。例如,当编码设备已使用FPC时,按照 $\tilde{x}_i = \tilde{y}_i \tilde{n}$ 使用反量化的包络值 $\tilde{n}$ 将执行了能量对齐的 $\tilde{y}_i$ 反归一化。通过执行反归一化,重构每个子频带的原始频谱平均能量。

[0076] 反变换器250通过对从频谱反归一化器240提供的变换系数进行反变换来重构时域的音频信号。例如,可通过使用与等式1相应的等式8来对频谱分量 $\tilde{x}_i$ 进行反变换来获得时域的音频信号 s_j 。

$$[0077] \quad s_j = \frac{1}{N} h_j \sum_{i=0}^{N-1} x_i \cos[\pi(j+(N+1)/2)(i+1/2)/N], \quad j = 0, \dots, 2N-1 \quad (8)$$

[0078] 在下文中,将更详细地描述图1的包络量化器130的操作。

[0079] 当包络量化器130以底数为c的对数尺度对每个子频带的包络值进行量化时,与量化索引相应的量化区域的边界 B_i 可由 $B_i = c^{(S_i+S_{i+1})/2}$ 表示,近似点(即,量化索引) A_i 可由 $A_i = c^{S_i}$ 表示,量化分辨率r可由 $r = S_i - S_{i-1}$ 表示,量化步长可由 $20\lg A_i - 20\lg A_{i-1} = 20r\lg c$ 表示。可通过等式3获得每个子频带的包络值n的量化索引 n_q 。

[0080] 在非优化的线性尺度的情况下,与量化索引 n_q 相应的量化区域的左边界和右边界从近似点远离不同的距离。由于该差异,如图3A和图4A所示,用于量化的信噪比(SNR)测量(即,量化误差)从近似点对于左边界和右边界具有不同的值。图3A示出非优化的对数尺度(底数为2)的量化,其中,量化分辨率为0.5和量化步长为3.01。如图3A所示,量化区域中的在左边界处和右边界处距离近似点的量化误差 SNR_L 和 SNR_R 分别是14.46dB和15.96dB。图4A示出非优化的对数尺度的量化(底数是2),其中,量化分辨率是1并且量化步长是6.02dB。如图4A所示,在量化区域中在左边界处和右边界处距离近似点的量化误差 SNR_L 和 SNR_R 分别是7.65dB和10.66dB。

[0081] 根据实施例,通过可变地改变与量化索引相应的量化区域的边界,与每个量化索引相应的量化区域中的总量化误差可被最小化。当在量化区域中的左边界处和右边界处获得的距离近似点的量化误差相同时,量化区域中的总量化误差可被最小化。可通过可变地改变取整系数b来获得量化区域的边界位移。

[0082] 可通过等式9来表示在与量化索引i相应的量化区域中的左边界处和右边界处获得的距离近似点的量化误差 SNR_L 和 SNR_R 。

$$[0083] \quad SNR_L = -20\lg((c^{S_i} - c^{(S_i+S_{i-1})/2})/c^{(S_i+S_{i-1})/2})$$

$$[0084] \quad SNR_R = -20\lg((c^{(S_i+S_{i+1})/2} - c^{S_i})/c^{(S_i+S_{i+1})/2}) \quad (9)$$

[0085] 在等式9中,c表示对数尺度的底数, S_i 表示与量化索引i相应的量化区域中的边界的指数。

[0086] 可使用由等式10定义的参数 b_L 和 b_R 来表示与量化索引相应的量化区域中的左边界和右边界的指数位移。

$$[0087] \quad b_L = S_i - (S_i + S_{i-1})/2$$

$$[0088] \quad b_R = (S_i + S_{i+1})/2 - S_i \quad (10)$$

[0089] 在等式10中, S_i 表示与量化索引i相应的量化区域中的边界处的指数, b_L 和 b_R 表示量化区域中的左边界和右边界距离近似点的指数位移。

[0090] 量化区域中的左边界处和右边界处的距离近似点的指数位移之和与量化分辨率相同,因此,可通过等式11来表示。

$$[0091] \quad b_L + b_R = r \quad (11)$$

[0092] 基于量化的一般特性,取整系数与和量化索引相应的量化区域中的左边界处的距离近似点的指数位移相同。因此,等式9可由等式12表示。

$$[0093] \quad SNR_L = -20\lg((c^{S_i} - c^{S_i+b_L})/c^{S_i+b_L}) = -20\lg(c^{b_L} - 1)$$

$$[0094] \quad SNR_R = -20\lg((e^{S_i+b_R} - e^{S_i})/e^{S_i+b_R}) = -20\lg(1 - e^{-r+b_L}) \quad (12)$$

[0095] 通过使在与量化索引相应的量化区域中的左边界处和右边界处的距离近似点的量化误差SNRL和SNRR相同,参数bL可通过等式13来确定。

$$[0096] \quad -20\lg(e^{b_L} - 1) = -20\lg(1 - e^{-r+b_L})$$

$$[0097] \quad e = e^{b_L} + e^{-r+b_L} = e^{b_L}(1 + e^{-r}) \quad (13)$$

[0098] 因此,可通过等式14表示取整系数bL。

$$[0099] \quad b_L = 1 - \log_e(1 + e^{-r}) \quad (14)$$

[0100] 图3B示出优化的对数尺度(底数为2)的量化,其中,量化分辨率为0.5并且量化步长为3.01。如图3B所示,量化区域中的左边界处和右边界处的距离近似点的量化误差SNRL和SNRR两者都是15.31dB。图4B示出优化的对数尺度(底数为2)的量化,其中,量化分辨率为1并且量化步长为6.02。如图4B所示,量化区域中的左边界处和右边界处的距离近似点的量化误差SNRL和SNRR两者都是9.54dB。

[0101] 取整系数b=bL确定从与量化索引i相应的量化区域中的左边界和右边界中的每个到近似点的指数距离。因此,可通过等式15执行根据实施例的量化。

$$[0102] \quad n_q = \lfloor \frac{1}{r} \log_e n + \frac{b_L}{r} \rfloor \quad (15)$$

[0103] 在图5A和图5B中示出通过以底数为2的对数尺度执行量化而获得的测试结果。根据信息论,比特率失真函数H(D)可用作参考,其中,通过所述参考可比较和分析各种量化方法。量化索引集合的熵可看作比特率并具有维数b/s,dB尺度的SNR可看作失真测量。

[0104] 图5A是在正态分布下执行的量化的比较曲线图。在图5A中,实线指示非优化的对数尺度的量化的比特率失真函数,点划线指示优化的对数尺度的量化的比特率失真函数。图5B是在均匀分布下执行的量化的比较曲线图。在图5B中,实线指示非优化的对数尺度的量化的比特率失真函数,点划线指示优化的对数尺度的量化的比特率失真函数。根据相应的分布规律、零期望值和单个方差使用随机数量的传感器产生正态分布和均匀分布的采样。可针对各种量化分辨率计算比特率失真函数H(D)。如图5A和图5B所示,点划线位于实线下面,这表示优化的对数尺度的量化的性能优于非优化的对数尺度的量化的性能。

[0105] 也就是说,根据优化的对数尺度的量化,可以以相同的比特率使用较少的量化误差执行量化,或以相同的比特率按照相同的量化误差,使用较少的比特执行量化。表1和表2中示出测试结果,其中,表1示出非优化的对数尺度的量化,表2示出优化的对数尺度的量化。

[0106] 表1

[0107]

量化分辨率(r)	2.0	1.0	0.5
取整系数(b/r)	0.5	0.5	0.5
正态分布			
比特率(H), b/s	1.6179	2.5440	3.5059
量化误差(D), dB	6.6442	13.8439	19.9534
均匀分布			
比特率(H), b/s	1.6080	2.3227	3.0830
量化误差, dB	6.6470	12.5018	19.3640

[0108] 表2

[0109]

量化分辨率(r)	2.0	1.0	0.5
取整系数(b/r)	0.3390	0.4150	0.4569
正态分布			
比特率(H), b/s	1.6069	2.5446	3.5059
量化误差(D), dB	8.2404	14.2284	20.0495
均匀分布			
比特率(H), b/s	1.6345	2.3016	3.0449
量化误差(D), dB	7.9208	12.8954	19.4922

[0110] 根据表1和表2,在量化分辨率0.5特性值SNR提高0.1dB,在量化分辨率1.0特性值SNR提高0.45dB,在量化分辨率2.0特性值SNR提高1.5dB。

[0111] 由于根据实施例的量化方法仅更新基于取整系数的量化索引的搜索表,因此复杂度不增加。

[0112] 现将更详细地描述图1的包络解码器140的操作。

[0113] 使用差值编码(delta coding)执行包络值的基于上下文的编码。可通过等式16表示当前子频带和先前子频带的包络值之间的量化差值。

$$[0114] \quad d(i) = n_q(i+1) - n_q(i) \quad (16)$$

[0115] 在等式16中, $d(i)$ 表示子频带 $(i+1)$ 的量化差值, $n_q(i)$ 表示子频带 (i) 的包络值的量化索引, $n_q(i+1)$ 表示子频带 $(i+1)$ 的包络值的量化索引。

[0116] 每个子频带的量化差值 $d(i)$ 限制在 $[-15, 16]$ 的范围内, 如下所述, 首先调整负量化差值, 随后调整正量化差值。

[0117] 首先, 通过使用等式16以从高频子频带到低频子频带的顺序获得量化差值 $d(i)$ 。在这种情况下, 如果 $d(i) < -15$, 则通过 $n_q(i) = n_q(i+1) + 15$ ($i = 42, \dots, 0$) 来执行调整。

[0118] 接下来, 通过使用等式16以从低频子频带到高频子频带的顺序获得量化差值 $d(i)$ 。在这种情况下, 如果 $d(i) > 16$, 则通过 $d(i) = 16, n_q(i+1) = n_q(i) + 16$ ($i = 0, \dots, 42$) 来执行调整。

[0119] 最后, 通过将偏移15添加到所有获得的量化差值 $d(i)$ 来产生范围为 $[0, 31]$ 的量化差值。

[0120] 根据等式16, 当 N 个子频带存在于单个帧中时, 获得 $n_q(0), d(0), d(1), d(2), \dots, d(N-2)$ 。使用上下文模型对当前子频带的量化差值进行编码, 根据实施例, 可将先前子频带的量化差值用作上下文。由于第一子频带的 $n_q(0)$ 存在于 $[0, 31]$ 的范围内, 因此可通过使用5比特按照其原样对量化差值 $n_q(0)$ 进行无损编码。当第一子频带的 $n_q(0)$ 用作 $d(0)$ 的上下文时, 通过使用预定参考值从 $n_q(0)$ 获得的值可被使用。也就是说, 当 $d(i)$ 的霍夫曼编码被执行时, $d(i-1)$ 可用作上下文, 当 $d(0)$ 的霍夫曼编码被执行时, 通过从 $n_q(0)$ 减去预定参考值而获得的值可用作上下文。预定参考值可以是例如通过仿真或实验预先设置为最优值的预定常数值。参考值可被包括在比特流中并被发送, 或在编码设备或解码设备中预先被提供。

[0121] 根据实施例, 包络编码器140可将用作上下文的先前子频带的量化差值的范围划分为多个组, 并基于针对所述多个组预定义的霍夫曼表对当前子频带的量化差值执行霍夫曼编码。可例如通过使用大数据库的训练处理产生霍夫曼表。也就是说, 基于预定标准收集数据, 并基于收集的数据产生霍夫曼表。根据实施例, 在先前子频带的量化差值的范围内收集当前子频带的量化差值的频率的数据, 并且可针对所述多个组产生霍夫曼表。

[0122] 可使用当前子频带的量化差值的概率分布的分析结果选择各种分布模型, 并因此可执行对具有类似分布模型的量化级别的分组, 其中, 通过将先前子频带的量化差值用作上下文来获得所述分析结果。在表3中示出三个组的参数。

[0123] 表3

[0124]

组编号	量化差值的下限	量化差值的上限
#1	0	12
#2	13	17
#3	18	31

[0125] 图6中示出了三个组的概率分布。组#1的概率分布于组#3的概率分布类似, 它们基于 x 轴基本被反转(reverse)(或翻转(flip))。这表示在没有任何编码效率的损失的情况

下,相同的概率模型可用于两个组#1和#3。也就是说,两个组#1和#3可使用相同的霍夫曼表。因此,可使用关于组#2的第一霍夫曼表和由组#1和#3共享的第二霍夫曼表。在这种情况下,组#1中的代码的索引可相对于组#3相反地表示。也就是说,当由于作为上下文的先前子频带的量化差值而将当前子频带的量化差值 $d(i)$ 的霍夫曼表确定为组#1时,可通过对编码端的处理的反向处理来将当前子频带的量化差值 $d(i)$ 改变为 $d'(i)=A-d(i)$,从而通过参考关于组#3的霍夫曼表来执行霍夫曼解码。在解码端,通过参考关于组#3的霍夫曼表来执行霍夫曼解码,并且通过 $d(i)=A-d'(i)$ 的转换处理来从 $d'(i)$ 提取最终值 $d(i)$ 。这里,值A可被设置为使得组#1和#3的概率分布彼此对称。可预先将值A设置为最优值而不是在编码和解码处理中提取。可选择地,可使用关于组#1的霍夫曼表而不是关于组#3的霍夫曼表,并可改变组#3中的量化差值。根据实施例,当 $d(i)$ 具有范围在 $[0,31]$ 中的值时,值A可以是31。

[0126] 图7是示出根据示例性实施例的图1的数字信号处理设备100的包络编码器140中的基于上下文的霍夫曼编码处理的流程图。在图7中,使用根据三个组中的量化差值的概率分布确定的两个霍夫曼表。另外,当对当前子频带的量化差值 $d(i)$ 执行霍夫曼编码时,先前子频带的量化差值 $d(i-1)$ 用作上下文,例如,使用关于组#2的第一霍夫曼表和关于组#3的第二霍夫曼表。

[0127] 参照图7,在操作710,确定先前子频带的量化差值 $d(i-1)$ 是否属于组#2。

[0128] 如果在操作710确定先前子频带的量化差值 $d(i-1)$ 属于组#2,则在操作720,从第一霍夫曼表选择当前子频带的量化差值 $d(i)$ 的代码。

[0129] 如果在操作710另外确定先前子频带的量化差值 $d(i-1)$ 不属于组#2,则在操作730,确定先前子频带的量化差值 $d(i-1)$ 是否属于组#1。

[0130] 如果在操作730确定先前子频带的量化差值 $d(i-1)$ 不属于组#1,即,如果先前子频带的量化差值 $d(i-1)$ 属于组#3,则在操作740,从第二霍夫曼表选择当前子频带的量化差值 $d(i)$ 的代码。

[0131] 如果在操作730另外地确定先前子频带的量化差值 $d(i-1)$ 属于组#1,则在操作750,对当前子频带的量化差值 $d(i)$ 进行反转并且从第二霍夫曼表选择当前子频带的反转的量化差值 $d'(i)$ 的代码。

[0132] 在操作760,使用在操作720、740或750选择的代码执行当前子频带的量化差值 $d(i)$ 的霍夫曼编码。

[0133] 图8是示出根据示例性实施例的图2的数字信号解码设备200的包络解码器210中的基于上下文的霍夫曼解码处理的流程图。如在图7中,在图8中,使用根据在三个组中的量化差值的概率分布确定的两个霍夫曼表。另外,当对当前子频带的量化差值 $d(i)$ 执行霍夫曼编码时,先前子频带的量化差值 $d(i-1)$ 用作上下文,例如,使用关于组#2的第一霍夫曼表和关于组#3的第二霍夫曼表。

[0134] 参照图8,在操作810,确定先前子频带的量化差值 $d(i-1)$ 是否属于组#2。

[0135] 如果在操作810确定先前子频带的量化差值 $d(i-1)$ 属于组#2,则在操作820,从第一霍夫曼表选择当前子频带的量化差值 $d(i)$ 的代码。

[0136] 如果在操作810另外地确定先前子频带的量化差值 $d(i-1)$ 不属于组#2,则在操作830,确定先前子频带的量化差值 $d(i-1)$ 是否属于组#1。

[0137] 如果在操作830确定先前子频带的量化差值 $d(i-1)$ 不属于组#1,即,如果先前子频

带的量化差值 $d(i-1)$ 属于组#3,则在操作840,从第二霍夫曼表选择当前子频带的量化差值 $d(i)$ 的代码。

[0138] 如果在操作830另外地确定先前子频带的量化差值 $d(i-1)$ 属于组#1,则在操作850,对当前子频带的量化差值 $d(i)$ 进行反转,并且从第二霍夫曼表选择当前子频带的反转的量化差值 $d'(i)$ 的代码。

[0139] 在操作860,使用在操作820、840或850选择的代码执行当前子频带的量化差值 $d(i)$ 的霍夫曼解码。

[0140] 在表4中示出了按照帧的比特代价差分析。如在表4中所示,根据图7的实施例的编码效率较原始霍夫曼编码算法平均增加9%。

[0141] 表4

[0142]

算法	比特率, kbps	增益, %
霍夫曼编码	6.25	-
上下文+霍夫曼编码	5.7	9

[0143] 图9是根据示例性实施例的包括编码模块930的多媒体装置900的框图。

[0144] 图9的多媒体装置900可包括通信单元910和编码模块930。另外,根据被获得作为编码结果的音频比特流的用途,图9的多媒体装置900还可包括存储单元950以存储音频比特流。另外,图9的多媒体装置900还可包括麦克风970。也就是说,存储单元950和麦克风970是可选的。图9的多媒体装置900还可包括解码模块(未示出),例如,用于执行一般解码功能的解码模块或根据示例性实施例的解码模块。编码模块930可与包括在多媒体装置900中的其他组件(未示出)集成并由至少一个处理器实现。

[0145] 参照图9,通信单元910可接收从外部提供的音频信号和编码的比特流中的至少一个,或可发送作为编码模块930的编码的结果而获得的重构的音频信号和音频比特流中的至少一个。

[0146] 通信单元910被配置为通过如下的无线网络将数据发送到外部多媒体装置并从外部多媒体装置接收数据:诸如,无线互联网、无线内联网、无线电话网、无线局域网(LAN)、Wi-Fi、Wi-Fi直连(WFD)、第三代(3G)、第四代(4G)、蓝牙、红外数据协会(IrDA)、射频识别(RFID)、超宽带(UWB)、Zigbee或近场通信(NFC)或有线网络(诸如,有线电话网络或有线互联网)。

[0147] 根据实施例,编码模块930可通过以下操作来产生比特流:将通过通信单元910或麦克风970提供的时域的音频信号变换到频谱的音频频谱,基于关于音频频谱的预定的子频带获取包络,基于预定的子频带对包络进行量化,获得相邻的子频带的量化的包络之间的差,并通过将先前子频带的差值用作上下文来对当前子频带的差值进行无损编码。

[0148] 根据另一实施例,当包络被量化时,编码模块930可对与预定的量化索引相应的量化区域的边界进行调整,使得量化区域中的总量化误差被最小化并可使用通过调整更新的量化表来执行量化。

[0149] 存储单元950可存储由编码模块930产生的编码的比特流。另外,存储单元950可存储操作多媒体装置900所需的各种程序。

[0150] 麦克风970可将来自用户或外部的音频信号提供给编码模块930。

[0151] 图10是根据示例性实施例的包括解码模块1030的多媒体装置1000的框图。

[0152] 图10的多媒体装置1000可包括通信模块1010和解码模块1030。另外,根据作为解码结果而获得的重构的音频信号的用途,图10的多媒体装置1000还可包括存储单元1050以存储重构的音频信号。另外,图10的多媒体装置1000还可包括扬声器1070。也就是说,存储单元1050和扬声器1070是可选的。图10的多媒体装置1000还可包括编码模块(未示出),例如,用于执行一般编码功能的编码模块或根据示例性实施例的编码模块。解码模块1030可与包括在多媒体装置1000中的其他组件(未示出)集成并由至少一个处理来实现。

[0153] 参照图10,通信单元1010可接收从外部提供的音频信号和编码的比特流中的至少一个,或可发送作为解码模块1030的解码的结果而获得的重构的音频信号以及作为编码的结果而获得的音频比特流中的至少一个。通信单元1010可被实现为基本与图9的通信单元910相同。

[0154] 根据实施例,解码模块1030可通过以下操作执行反量化:接收通过通信单元1010提供的比特流,通过将先前子频带的差值用作上下文来对当前子频带的差值进行无损解码,从作为无损解码的结果而重构的当前子频带的差值基于子频带获得量化的包络。

[0155] 存储单元1050可存储由解码模块1030产生的重构的音频信号。另外,存储单元1050可存储操作多媒体装置1000所需的各种程序。

[0156] 扬声器1070可将由解码模块1030产生的重构的音频信号输出到外部。

[0157] 图11是根据示例性实施例的包括编码模块1120和解码模块1130的多媒体装置1100的框图。

[0158] 图11的多媒体装置1100可包括通信单元1110、编码模块1120和解码模块1130。另外,根据作为编码结果而获得的音频比特流或作为解码结果而获得的重构的音频信号的用途,图11的多媒体装置1100还可包括存储单元1140,用于存储音频比特流或重构音频信号。另外,图11的多媒体装置1100还可包括麦克风1150或扬声器1160。编码模块1120和解码模块1130可与包括在多媒体装置1100中的其他组件(未示出)集成,并由至少一个处理器来实现。

[0159] 由于图11的多媒体装置1100中的组件与图9的多媒体装置900中的组件或图10的多媒体装置1000中的组件相同,因此省略其详细描述。

[0160] 图9、图10或图11的多媒体装置900、1000或1100可包括:包括电话或移动电话的仅语音通信的终端、包括TV或MP3播放器的仅广播或音乐的装置或仅语音通信终端和仅广播或音乐的装置的混合型终端装置,但不限于此。另外,图9、图10或图11的多媒体装置900、1000或1100可用作客户机、服务器或布置在客户机或服务器之间的变换器。

[0161] 例如,如果多媒体装置900、1000或1100是移动电话,则尽管未示出,但移动电话还可包括诸如键区的用户输入单元、用户界面或用于显示由移动电话处理的信息的显示单元和用于控制移动电话的一般功能的处理器。另外,移动电话还可包括具有图像拾取功能的相机单元和用于执行移动电话所需的功能的至少一个组件。

[0162] 作为另一示例,如果多媒体装置900、1000或1100是TV,则尽管未示出,但TV还可包括诸如键区的用户输入单元、用于显示接收的广播信息的显示单元和用于控制TV的一般功能的处理器。另外,TV还可包括用于执行TV所需的功能的至少一个组件。

[0163] 根据示例性实施例的方法可被编写为计算机可执行程序,并可被实现在通用数字

计算机中,其中,通用数字计算机通过使用非暂时性计算机可读记录介质执行程序。另外,可在实施例中使用的数据结构、程序指令或数据文件可以以各种方式记录在计算机可读记录介质中。非暂时性计算机可读记录介质是可存储随后可由计算机系统读取的数据的任何数据存储装置。非暂时性计算机可读记录介质的示例包括:磁存储介质(诸如,硬盘、软盘和磁带)、光学记录介质(诸如,CD-ROM、DVD)、磁光介质(诸如,光盘)和专门被配置为存储和执行程序指令的硬件装置(诸如,ROM、RAM和闪存)。另外,非暂时性计算机可读记录介质可以是用于传输指定程序指令、数据结构等的信号的传输介质。程序指令的示例可不仅包括由编译器创建的机器语言代码,还包括由计算机使用解释器等可执行的高级语言代码。

[0164] 虽然以上已具体地示出和描述了示例性实施例,但本领域的普通技术人员将理解,在不脱离由权利要求限定的本发明构思的精神和范围的情况下,可在形式和细节上进行各种改变。示例性实施例应被理解为描述性意义,而不是限制的目的。因此,本发明构思的范围不由示例性实施例的详细描述而由权利要求限定,在范围内的所有不同将被解释为包括在本发明构思中。

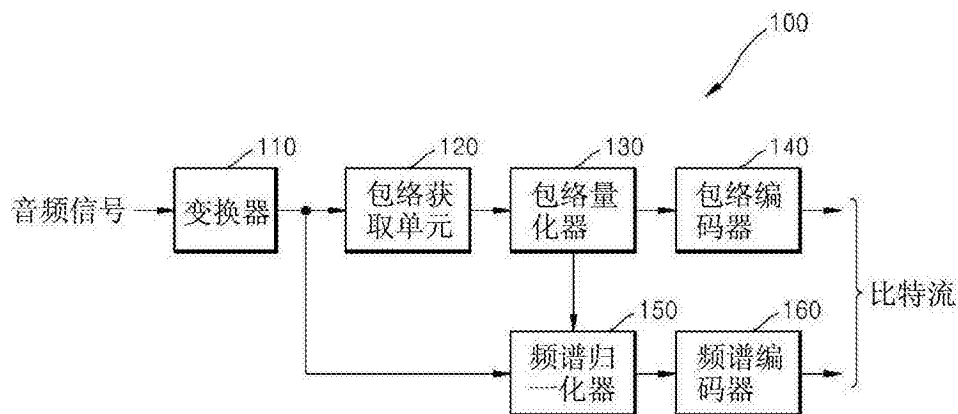


图1

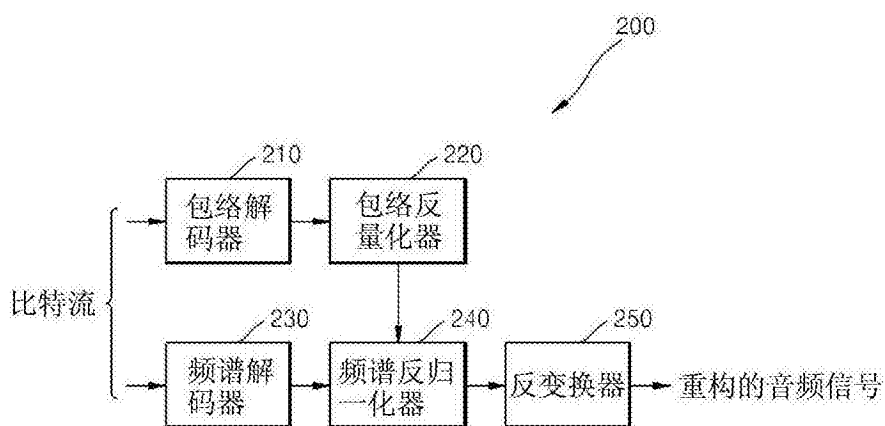


图2

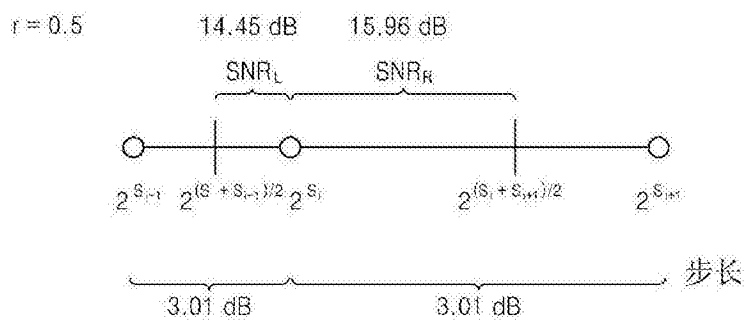


图3A

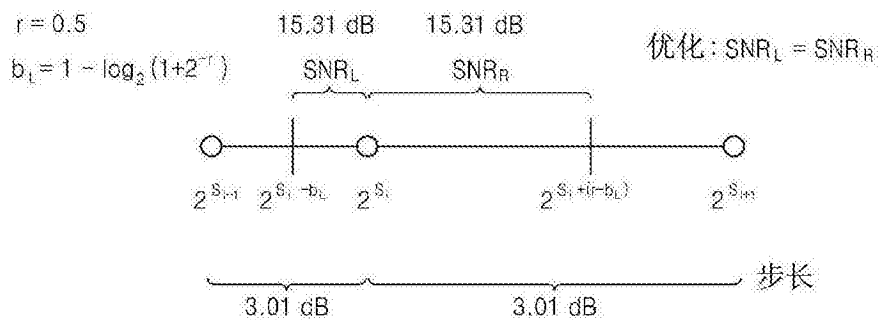


图3B

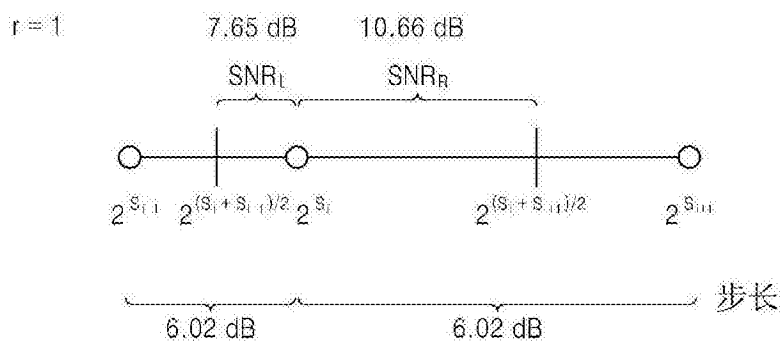


图4A

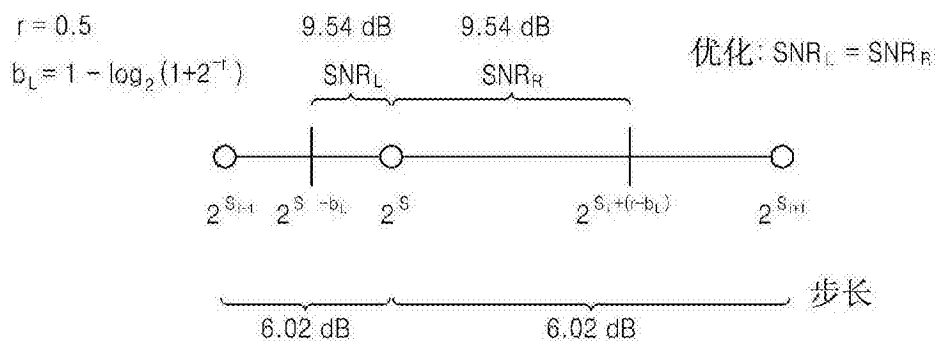


图4B

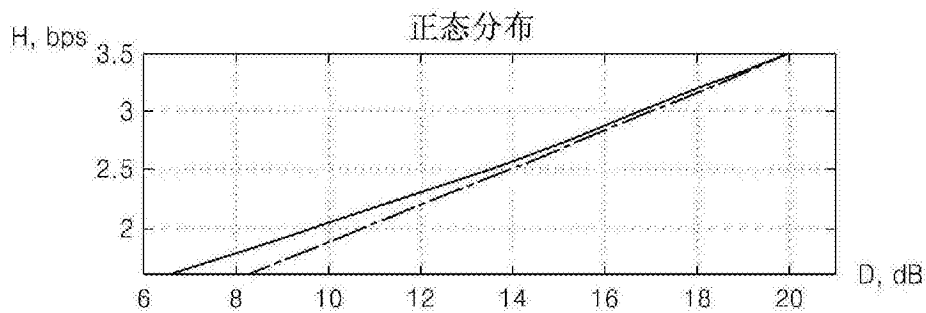


图5A

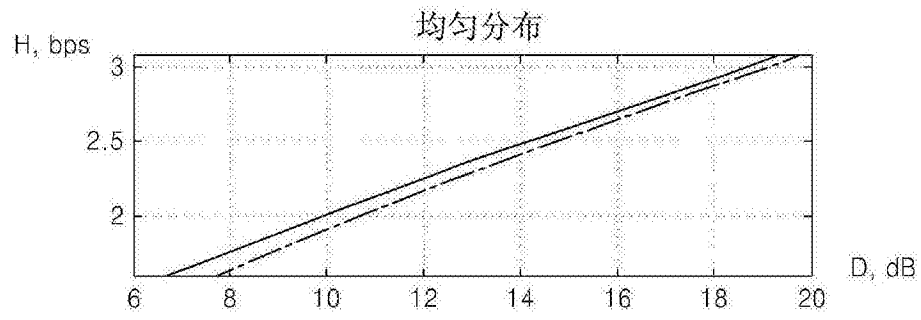


图5B

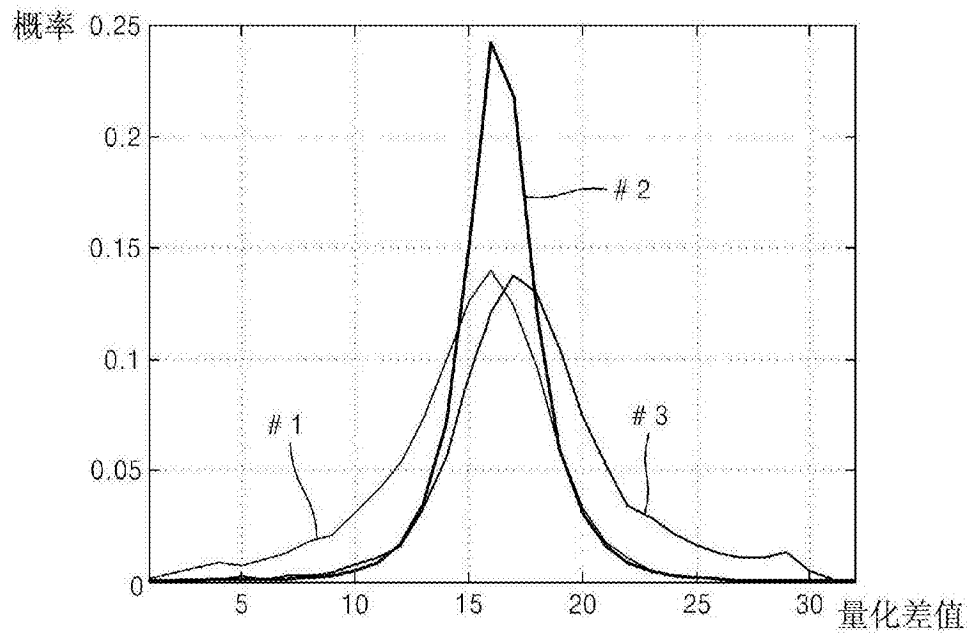


图6

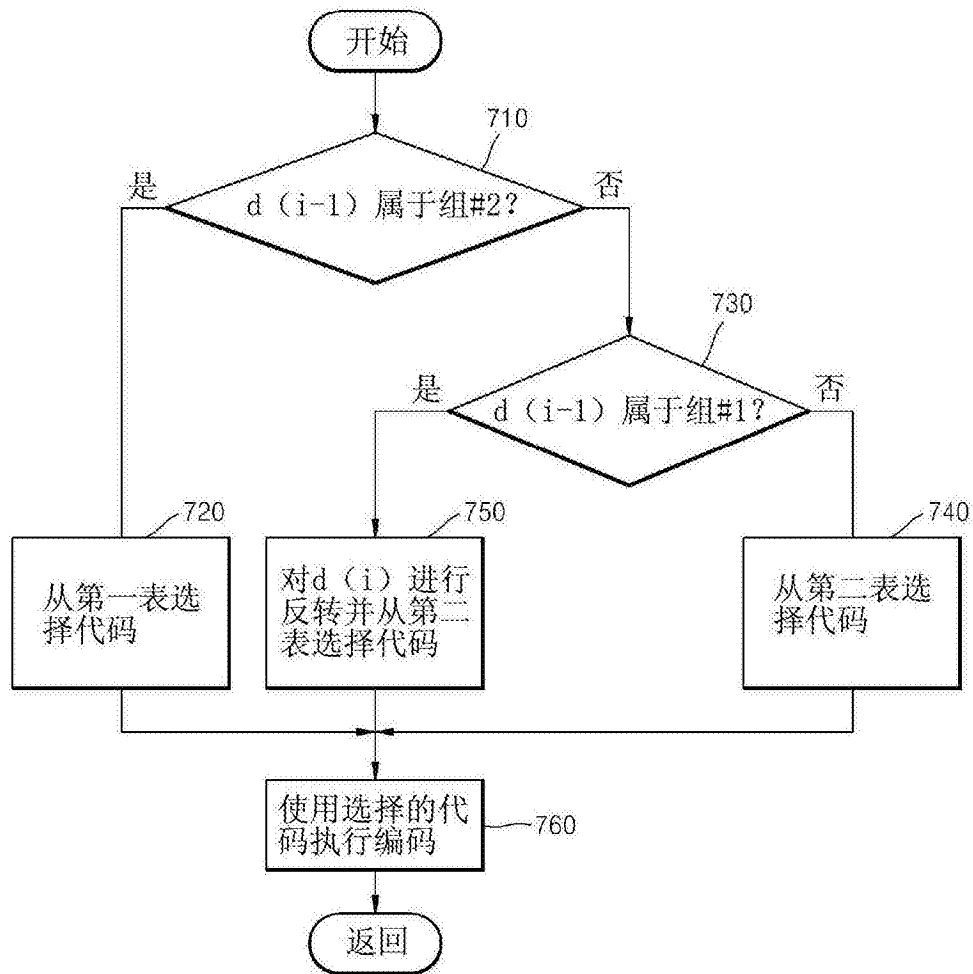


图7

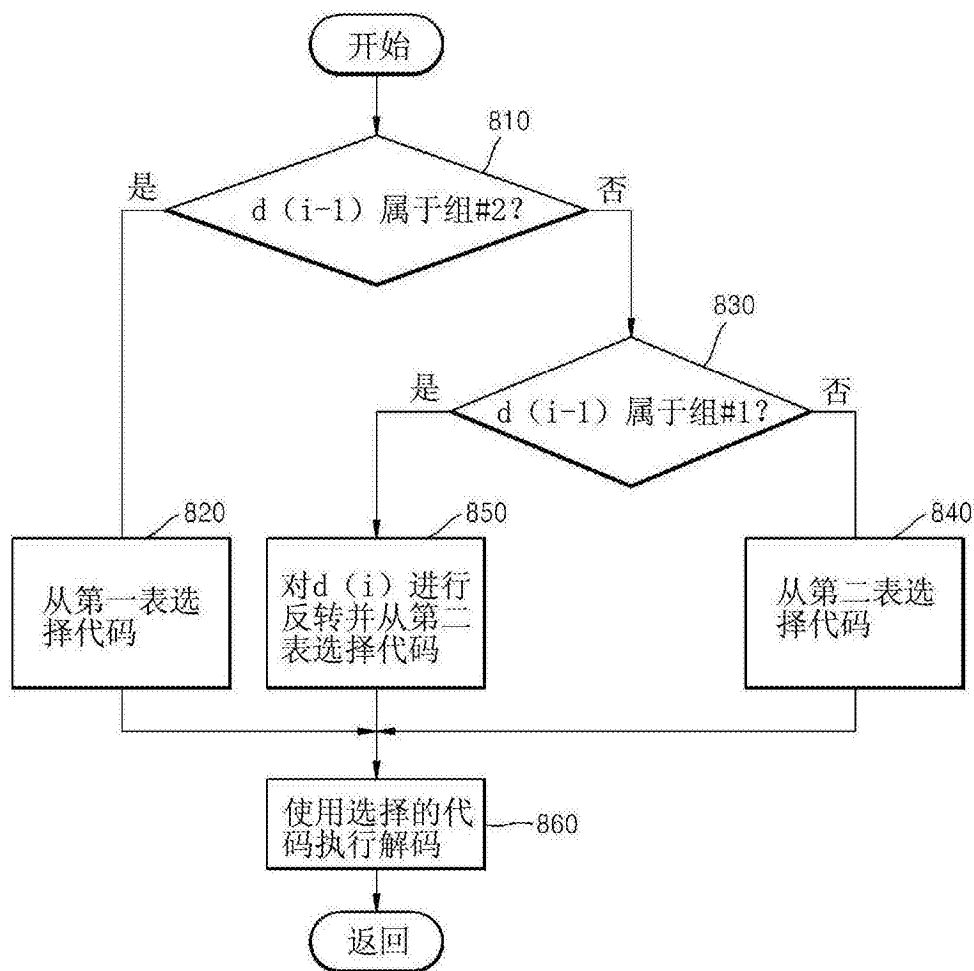


图8

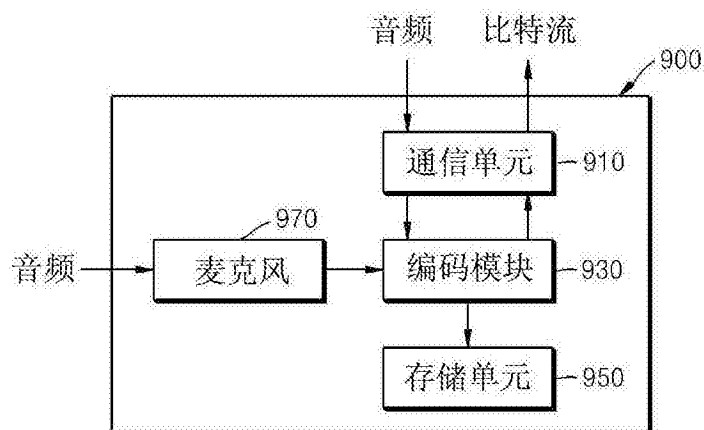


图9

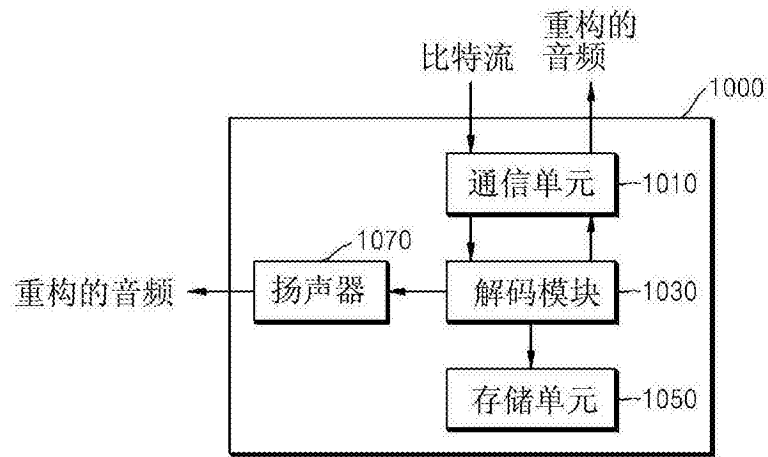


图10

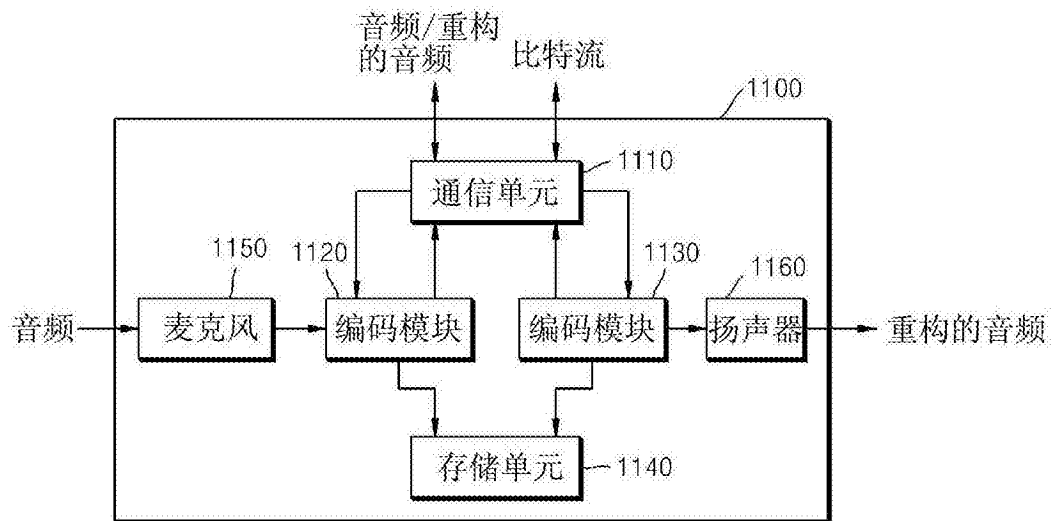


图11