



(12) 发明专利

(10) 授权公告号 CN 103106902 B

(45) 授权公告日 2015. 12. 16

(21) 申请号 201210441382. 2

US 6324505 B1, 2001. 11. 27, 全文.

(22) 申请日 2006. 07. 14

Heiko Purnhagen et al. Sinusoidal coding using loudness-based component selection.

(30) 优先权数据

10-2005-0064507 2005. 07. 15 KR

《Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International

(62) 分案原申请数据

200680025920. 2 2006. 07. 14

Conference on》. 2002, 第 2 卷 1817-1818.

(73) 专利权人 三星电子株式会社

Jesper Jensen et al. Speech Enhancement Using a Constrained Iterative Sinusoidal Model. 《IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING》. 2001, 第 9 卷 (第 7 期), 733.

地址 韩国京畿道水原市

(72) 发明人 金重会 吴殷美

康斯坦丁·奥斯波夫

波利斯·库德里亚索夫

Renat Vafin, et al.. EXPLOITING TIME AND FREQUENCY MASKING IN CONSISTENT SINUSOIDAL ANALYSIS-SYNTHESIS. 《Acoustics, Speech, and Signal Processing, IEEE International Conference on》. 2000, 第 2 卷 901-903.

(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 王艳娇

马延飞. MPEG 音频第 III 层编码的算法研究与硬件实现. 《CNKI 中国优秀硕士学位论文全文数据库》. 2002, 全文.

(51) Int. Cl.

G10L 19/02(2013. 01)

G10L 19/035(2013. 01)

审查员 刘红梅

(56) 对比文件

US 2004/0088160 A1, 2004. 05. 06, 全文.

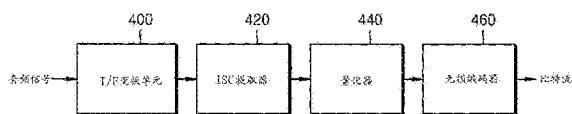
权利要求书 2 页 说明书 9 页 附图 6 页

(54) 发明名称

低比特率音频信号解码方法

(57) 摘要

一种提取具有重要频谱分量 (ISC) 的音频信号的方法和装置、以及使用该提取 ISC 的方法和装置的低比特率音频信号编码 / 解码方法。提取 ISC 的方法包括：通过使用心理声学模型计算包括变换的频谱音频信号的 SMR (信号掩蔽比) 值的感知重要性，使用 SMR 值将掩蔽阈值小于所述频谱音频信号的掩蔽阈值的频谱音频信号选作为第一 ISC；根据预定权重因数从选作为 ISC 的频谱音频信号提取频谱峰值以选择第二 ISC。因此，可有效地对感知重要频谱分量编码，从而获得低比特率的高声音质量。此外，通过使用心理声学模型能够提取感知重要频谱分量，能够无需相位信息执行编码，并且能够有效地表示低比特率的频谱信号。



1. 一种低比特率音频信号解码方法,包括:
提取音频信号的帧的概率模型信息;
针对音频信号产生指示感知重要频谱分量的存在的索引信息、量化器信息、量化步长、频谱分量的分组信息和音频信号量化值;
根据量化器信息、量化步长和所述分组信息对音频信号执行逆量化;
将逆量化的信号变换为时域信号。
2. 如权利要求 1 所述的低比特率音频信号解码方法,其中,通过考虑使用的比特量和量化误差对音频信号的频谱分量进行分组以形成多个组来获得分组信息,以便最小化附加信息,其中,附加信息包括缩放因数。
3. 如权利要求 1 所述的低比特率音频信号解码方法,还包括:
通过使用提取的概率模型信息对指示感知重要频谱分量的存在的索引信息、量化步长和所述分组信息执行无损解码。
4. 如权利要求 1 所述的低比特率音频信号解码方法,其中,产生感知重要频谱分量的步骤包括:
对感知重要频谱分量进行解码;
通过使用指示感知重要频谱分量的存在的索引信息将解码的感知重要频谱分量映射到频谱轴。
5. 一种低比特率音频信号解码设备,包括:
无损解码器,提取用于音频信号的帧的概率模型信息,并且通过使用该概率模型信息产生指示感知重要频谱分量的存在的索引信息、量化器信息、量化步长、频谱分量的分组信息和音频信号量化值;
逆量化器,根据量化器信息、量化步长和所述分组信息对音频信号执行逆量化;和
F/T 变换单元,将逆量化的信号变换为时域信号。
6. 如权利要求 5 所述的低比特率音频信号解码设备,其中,通过考虑使用的比特量和量化误差对音频信号的频谱分量进行分组以形成多个组来获得所述分组信息,以便最小化附加信息,其中,附加信息包括缩放因数。
7. 如权利要求 5 所述的低比特率音频信号解码设备,其中,无损解码器通过使用提取的概率模型对指示感知重要频谱分量的存在的索引信息、量化步长和所述分组信息执行无损解码。
8. 如权利要求 5 所述的低比特率音频信号解码设备,其中,无损解码器对感知重要频谱分量进行解码,并且通过使用指示感知重要频谱分量的存在的索引信息将解码的感知重要频谱分量映射到频谱轴。
9. 一种音频信号编码和 / 或解码系统,包括:
编码器,选择具有根据频带的信号掩蔽比 SMR 值、权重因数和信噪比 SNR 中的至少一个提取的一个或多个感知重要频谱分量的频谱音频信号,并且根据关于选择的感知重要频谱分量的信息对频谱音频信号编码,所述信息包括频谱分量的分组信息;和
解码器,根据所述信息对编码的频谱音频信号解码。
10. 如权利要求 9 所述的音频信号编码和 / 或解码系统,其中,通过考虑使用的比特量和量化误差对音频信号的频谱分量进行分组以形成多个组来获得所述分组信息,以便最小

化附加信息,其中,附加信息包括缩放因数。

11. 一种音频信号编码系统,包括:

编码器,选择具有根据频谱音频信号的频带的信号掩蔽比 SMR 值、权重因数和信噪比 SNR 中的至少一个提取的一个或多个感知重要频谱分量的频谱音频信号,并且根据关于选择的感知重要频谱分量的信息对频谱音频信号编码,所述信息包括频谱分量的分组信息。

12. 如权利要求 11 所述的音频信号编码系统,其中,通过考虑使用的比特量和量化误差对音频信号的频谱分量进行分组以形成多个组来获得所述分组信息,以便最小化附加信息,其中,附加信息包括缩放因数。

13. 一种音频信号解码系统,包括:

解码器,对编码的频谱音频信号解码,

其中,通过选择具有根据频带的信号掩蔽比 SMR 值、权重因数和信噪比 SNR 中的至少一个提取的一个或多个感知重要频谱分量的频谱音频信号,并且根据关于选择的感知重要频谱分量的信息对频谱音频信号编码,已经产生了编码的频谱音频信号,所述信息包括频谱分量的分组信息。

14. 如权利要求 13 所述的音频信号解码系统,其中,通过考虑使用的比特量和量化误差对音频信号的频谱分量进行分组以形成多个组来获得所述分组信息,以便最小化附加信息,其中,附加信息包括缩放因数。

低比特率音频信号解码方法

[0001] 本申请是申请日为 2006 年 7 月 14 日、申请号为 200680025920.2、题为“从音频信号提取重要频谱分量的方法和设备以及使用其的低比特率音频信号编码和 / 或解码方法和设备”的专利申请的分案申请。

技术领域

[0002] 本发明总体发明构思涉及一种音频信号编码和 / 或解码系统,更具体地讲,涉及一种提取音频信号的重要频谱分量的方法和设备以及使用其的对低比特率音频信号编码和解码的方法和设备。

背景技术

[0003] “MPEG (运动图像专家组) 音频”是用于高质量高性能立体声编码的 ISO/IEC 标准。MPEG 音频与运动图像编码根据 MPEG 的 ISO/IEC SC29/WG11 一起被标准化。对于 MPEG 音频,基于 32 个频带的子带编码 (频带分解编码) 和改进离散余弦变换 (MDCT) 用于压缩,具体地讲,通过使用心理特征执行高性能压缩。与传统压缩编码方案相比, MPEG 音频可实现高质量的声音。

[0004] 为了高性能地压缩音频信号, MPEG 音频利用“感知编码”压缩方案以减小音频信号的压缩量,在该“感知编码”压缩方案中,通过使用感测音频信号的人类的敏感特性来去除详细的低敏感信息。

[0005] 此外,在 MPEG 音频中,无声阶段的最小可听限制和掩蔽特性主要用于使用听觉心理特征的感知编码。无声阶段的最小可听限制是听觉可感知的声音的最小级别。最小可听限制与在无声阶段听觉可感知的噪声的限制有关。最小可听限制根据声音的频率改变。在一些频率,可听到比最小可听限制高的声音,但是在另一些频率,可能不会听到比最小可听限制低的声音。此外,特定声音的感测限制可根据与该特定声音一起听到的其他声音大大改变。这被称为“掩蔽效应”。发生掩蔽效应的频率的宽度被称为临界带。为了有效地利用听觉心理特征 (例如,临界带),将声音信号分解为频谱分量很重要。为此,频带被分为 32 个子带,随后执行子带编码。另外,在 MPEG 音频中,滤波器组用于消除 32 个子带的混叠噪声。

发明内容

[0006] 技术问题

[0007] MPEG 音频包括使用滤波器组和心理声学模型的比特分配和量化。通过 MDCT 产生的系数分配有最佳量化比特,并且通过使用心理声学模型 2 被压缩。用于分配最佳比特的心理声学模型 2 通过使用扩散函数基于 FFT 来估计掩蔽效应。因此,需要相对大量的复杂度。

[0008] 通常,对于低比特率 (32kbps 或更少) 音频信号的压缩,可分配给信号的比特数不足以量化音频信号的所有频谱分量及其无损编码。因此,需要提取感知的重要频谱分量

(ISC) 和量化及其无损编码。

[0009] 技术方案

[0010] 本发明总体发明构思提供一种从音频信号提取重要频谱分量以低比特率压缩音频信号的方法和设备。

[0011] 本发明总体发明构思还提供一种使用从音频信号提取重要频谱分量的方法和设备的低比特率音频信号编码方法和设备。

[0012] 本发明总体发明构思还提供一种对通过低比特率音频信号编码方法和设备编码的低比特率音频信号解码的低比特率音频信号解码方法和设备。

[0013] 将在接下来的描述中部分阐述本发明另外的方面和优点,还有一部分通过描述将是清楚的,或者可以经过本发明总体发明构思的实施而得知。

[0014] 可通过提供一种提取音频信号的重要频谱分量 (ISC) 的方法来实现本发明总体发明构思的前述和 / 或其他方面和优点,该方法包括:通过使用心理声学模型计算包括变换的频谱音频信号的信号掩蔽比 (SMR) 值的感知重要性,使用 SMR 值将掩蔽阈值小于所述频谱音频信号的掩蔽阈值的频谱音频信号选作为第一 ISC;根据预定权重因数从选作为第一 ISC 的频谱音频信号提取频谱峰值以选择第二 ISC。可通过使用权重因数将被获得的当前信号的频率附近的预定数量的频谱值获得权重因数。

[0015] 该方法还可包括获得频带的 SNR(信噪比);和将具有低 SNR 的频带中峰值大于预定值的频谱分量选作为 ISC。

[0016] 还可通过提供一种提取音频信号的重要频谱分量 (ISC) 的方法来实现本发明总体发明构思的前述和 / 或其他方面和优点,该方法包括:通过使用心理声学模型计算包括变换的频谱音频信号的 SMR(信号掩蔽比)值的感知重要性;使用 SMR 将掩蔽阈值小于所述频谱音频信号的掩蔽阈值的频谱音频信号选作为第一 ISC;和获得选作为第一 ISC 的频谱音频信号中的频带的 SNR 以将具有低 SNR 的频带中峰值大于预定值的频谱分量的频谱音频信号选作为另一 ISC。

[0017] 还可通过提供一种低比特率音频信号编码方法来实现本发明总体发明构思的前述和 / 或其他方面和优点,该方法包括:通过使用心理声学模型计算包括频谱音频信号的 SMR(信号掩蔽比)值的感知重要性;使用 SMR 值将掩蔽阈值小于所述频谱音频信号的掩蔽阈值的频谱音频信号选作为第一 ISC;和根据预定权重因数从选作为第一 ISC 的频谱音频信号提取频谱峰值,并将具有该频谱峰值的频率的频谱音频信号选作为第二 ISC;和对具有第二 ISC 的频谱音频信号执行量化和无损编码。提取频谱峰值的步骤可包括:获得频带的 SNR(信噪比),并且通过使用 SNR 将具有低 SNR 的频带中峰值大于预定值的频谱分量选作为第三 ISC。低比特率音频信号编码方法还可包括:通过使用 MDCT(改进离散余弦变换)和 MDST(改进离散正弦变换)来将时域音频信号变换为频谱音频信号以产生频谱音频信号。对 ISC 音频信号执行量化的步骤可包括:根据使用的比特量和量化误差将音频信号分成多个组以最小化附加信息;根据 SMR(信号掩蔽比)和所述多组的动态范围的数据分布确定量化步长;和通过使用所述多组的一个或多个预定量化器对音频信号量化。可通过使用采用组的最大值规格化的值和量化步长确定量化器。量化可以是 Max-Lloyd 量化。

[0018] 对量化的信号执行无损编码的步骤可包括:上下文算术编码。执行上下文算术编码的步骤可包括:采用指示 ISC 的存在的频谱索引表示组成帧的频谱分量;和根据与先

前帧的相关性和相邻 ISC 的分布选择随机模型,以对音频信号的量化值以及包括量化器信息、量化步骤、分组信息和频谱索引值的附加信息执行无损编码。

[0019] 还可通过提供一种低比特率音频信号编码方法来实现本发明总体发明构思的前述和/或其他方面和优点,该方法包括:通过使用心理声学模型计算包括频谱音频信号的 SMR(信号掩蔽比)值的感知重要性;使用 SMR 值将掩蔽阈值小于所述频谱音频信号的掩蔽阈值的频谱信号选作为第一 ISC;获得选作为第一 ISC 的频谱音频信号中的频带的 SNR,并且使用 SNR 将具有低 SNR 的频带中峰值大于预定值的频谱分量选作为另一 ISC;和对于具有另一 ISC 的频谱音频信号执行量化和无损编码。

[0020] 还可通过提供一种提取音频信号 ISC(重要频谱分量)的设备来实现本发明总体发明构思的前述和/或其他方面和优点,该设备包括:心理建模单元,通过使用心理声学模型计算包括变换的频谱音频信号的 SMR(信号掩蔽比)值的感知重要性;第一 ISC 选择单元,使用 SMR 将掩蔽阈值小于所述频谱音频信号的掩蔽阈值的频谱音频信号选作为第一 ISC;和第二 ISC 选择单元,根据预定权重因数从选作为第一 ISC 的频谱音频信号提取频谱峰值并选择第二 ISC。可通过使用权重因数将被获得的当前信号的频率附近的预定数量的频谱值获得第二 ISC 选择单元的权重因数。该设备还可包括:第三 ISC 选择单元,获得频带的 SNR(信噪比),并通过使用 SNR 将具有低 SNR 的频带中峰值大于预定值的频谱分量选作为第三 ISC。

[0021] 还可通过提供一种提取音频信号 ISC(重要频谱分量)的设备来实现本发明总体发明构思的前述和/或其他方面和优点,该设备包括:心理建模单元,通过使用心理声学模型计算包括变换的频谱音频信号的 SMR(信号掩蔽比)值的感知重要性;第一 ISC 选择单元,使用 SMR 将掩蔽阈值小于所述频谱音频信号的掩蔽阈值的频谱音频信号选作为第一 ISC;和另一 ISC 选择单元,获得选作为第一 ISC 的频谱音频信号中的频带的 SNR,并且使用 SNR 将具有低 SNR 的频带中峰值大于预定值的频谱分量选作为另一 ISC。

[0022] 还可通过提供一种低比特率音频信号编码提取设备来实现本发明总体发明构思的前述和/或其他方面和优点,该设备包括:心理建模单元,通过使用心理声学模型计算包括变换的频谱音频信号的 SMR(信号掩蔽比)值的感知重要性;第一 ISC(重要频谱分量)选择单元,使用 SMR 值将掩蔽阈值小于所述频谱音频信号的掩蔽阈值的频谱音频信号选作为第一 ISC;第二 ISC 选择单元,根据预定权重因数从选作为第一 ISC 的频谱音频信号提取频谱峰值并且选择第二 ISC;量化器,对具有第二 ISC 的频谱音频信号量化;和无损编码器,对量化的信号执行无损编码。

[0023] 低比特率音频信号编码设备还可包括:第三 ISC 选择单元,获得频带的 SNR(信噪比),并且使用 SNR 将具有低 SNR 的频带中峰值大于预定值的频谱分量选作为第三 ISC。

[0024] 低比特率音频信号编码设备还可包括:T/F 变换单元,通过使用 MDCT(改进离散余弦变换)和 MDST(改进离散正弦变换)来将时域音频信号变换为频谱音频信号。

[0025] 量化器可包括:分组单元,根据使用的比特量和量化误差将频谱音频信号分为多个组以最小化附加信息;量化步长确定单元,根据 SMR(信号掩蔽比)和所述多个组的数据分布(动态范围)确定量化步长;和组量化器,通过使用所述多组的预定量化器对频谱音频信号量化。组量化器的量化可以是 Max-Lloyd 量化,无损编码器的无损编码可以是上下文算术编码。

[0026] 无损编码器可包括：索引单元，采用指示 ISC 的存在的频谱索引表示组成帧的频谱分量；随机模型无损编码器，根据与先前帧的相关性和相邻 ISC 的分布选择随机模型，并且对频谱音频信号的量化值以及包括量化器信息、量化步长、分组信息和频谱索引值的附加信息执行无损编码。

[0027] 还可通过提供一种低比特音频信号编码设备来实现本发明总体发明构思的前述和 / 或其他方面和优点，该设备包括：心理建模单元，通过使用心理声学模型计算包括变换的频谱音频信号的 SMR（信号掩蔽比）值的感知重要性；第一 ISC（重要频谱分量）选择单元，使用感知重要性将掩蔽阈值小于所述频谱音频信号的掩蔽阈值的频谱音频信号选作为第一 ISC；另一 ISC 选择单元，获得选作为第一 ISC 的频谱音频信号中的频带的 SNR，并且通过使用 SNR 将具有低 SNR 的频带中峰值大于预定值的频谱分量选作为另一 ISC；和量化器，对具有所述另一 ISC 的频谱音频信号量化；和无损编码器，对量化的信号执行无损编码。

[0028] 还可通过提供一种低比特音频信号解码方法来实现本发明总体发明构思的前述和 / 或其他方面和优点，该方法包括：恢复指示 ISC（重要频谱分量）的存在的索引信息、量化器信息、量化步长、ISC 分组信息和音频信号量化值；参照恢复的量化器信息、量化步长和分组信息对音频信号执行逆量化；和将逆量化的值变换为时域信号。

[0029] 还可通过提供一种低比特音频信号解码设备来实现本发明总体发明构思的前述和 / 或其他方面和优点，该设备包括：无损解码器，提取用于帧的随机模型信息，并且通过使用该随机模型信息恢复指示 ISC（重要频谱分量）的存在的索引信息、量化器信息、量化步长、ISC 分组信息和音频信号量化值；逆量化器，参照恢复的量化器信息、量化步长和分组信息执行逆量化；和 F/T 变换单元，将逆量化的值变换为时域信号。

[0030] 还可通过提供一种实现用于执行以下方法的计算机程序的计算机可读介质来实现本发明总体发明构思的前述和 / 或其他方面和优点，该方法包括：根据心理声学模型计算包括变换的频谱音频信号的信号掩蔽比（SMR）值的感知重要性，使用感知重要性将掩蔽阈值小于所述频谱音频信号的掩蔽阈值的频谱音频信号选作为一个或多个第一重要频谱分量（ISC）；根据预定权重因数从选作为一个或多个第一 ISC 的频谱音频信号提取频谱峰值以选择将被用于对频谱音频信号编码的一个或多个第二 ISC。

[0031] 还可通过提供一种实现用于执行以下方法的计算机程序的计算机可读介质来实现本发明总体发明构思的前述和 / 或其他方面和优点，该方法包括：对音频信号恢复指示重要频谱分量（ISC）的存在的索引信息、量化器信息、量化步长、ISC 分组信息和音频信号量化值；根据恢复的量化器信息、量化步长和分组信息对音频信号执行逆量化；和将逆量化的信号变换为时域信号。

[0032] 还可通过提供一种音频信号编码和 / 或解码系统来实现本发明总体发明构思的前述和 / 或其他方面和优点，该系统包括：编码器，根据频带的信号掩蔽比（SMR）值、以及权重因数和信噪比（SNR）中的一个选择具有一个或多个重要频谱分量（ISC）的频谱音频信号，并且根据关于选择的 ISC 的信息对频谱音频信号编码；和解码器，根据所述信息对编码频谱音频信号解码。

[0033] 还可通过提供一种音频信号编码和 / 或解码系统来实现本发明总体发明构思的前述和 / 或其他方面和优点，该系统包括：编码器，根据频带的信号掩蔽比（SMR）值、以及权重因数和信噪比（SNR）中的一个选择具有一个或多个重要频谱分量（ISC）的频谱音频信

号,并且根据关于选择的 ISC 的信息对频谱音频信号编码。

[0034] 还可通过提供一种音频信号编码和 / 或解码系统来实现本发明总体发明构思的前述和 / 或其他方面和优点,该系统包括:解码器,根据关于 ISC 的信息对编码的音频信号解码。可根据频谱音频信号的频带的信号掩蔽比 (SMR) 值、以及权重因数和信噪比 (SNR) 中的一个获得 ISC。

附图说明

[0035] 通过下面结合附图对实施例进行的详细描述,本发明总体发明构思的这些和其他方面和优点将会变得更加清楚和更易于理解,其中:

[0036] 图 1 是示出根据本发明总体发明构思的实施例的从输入的音频信号提取重要频谱分量以按低比特率压缩音频信号的设备的框图;

[0037] 图 2 是示出根据本发明总体发明构思的实施例的从输入的音频信号提取重要频谱分量以按低比特率压缩音频信号的方法的流程图;

[0038] 图 3 是示出根据本发明总体发明构思的实施例的从输入的音频信号提取重要频谱分量以按低比特率压缩音频信号的方法的示意图;

[0039] 图 4 是示出根据本发明总体发明构思的实施例的使用从输入的音频信号提取重要频谱分量的设备按低比特率压缩音频信号的低比特率音频信号编码设备的构造的框图;

[0040] 图 5 是示出图 4 的设备的量化器的框图;

[0041] 图 6 是示出图 4 的设备的无损编码单元的框图;

[0042] 图 7 是示出根据本发明总体发明构思的实施例的使用从音频信号提取重要频谱分量的方法的低比特率音频信号编码方法的流程图;

[0043] 图 8 是示出图 7 的方法的 ISC 量化的详细流程图;

[0044] 图 9 是示出根据本发明总体发明构思的实施例的对通过使用从音频信号提取重要频谱分量的设备编码的低比特率音频信号进行解码的低比特率音频信号解码设备的框图;和

[0045] 图 10 是示出根据本发明总体发明构思的实施例的对通过使用提取音频信号的重要频谱分量的设备编码的低比特率音频信号进行解码的低比特率音频信号解码方法的流程图。

具体实施方式

[0046] 现在将对本发明总体发明构思的实施例进行详细参照,其示例在附图中表示,在整个附图中,相同的标号始终表示相同的部件。以下通过参考附图描述实施例以解释本发明总体发明构思。

[0047] 图 1 是示出根据本发明总体发明构思的实施例的从输入的音频信号提取重要频谱分量 (ISC) 以按低比特率压缩音频信号的设备的框图。音频信号 ISC 提取设备包括心理建模单元 100 和 ISC 选择单元 150。

[0048] 心理建模单元 100 对根据心理特征变换的频谱音频信号计算信号掩蔽比 (SMR) 值。通过使用改进离散余弦变换 (MDCT) 和改进离散正弦变换 (MDST) (而不是离散傅立叶

变换 (DFT)) 产生输入到心理建模单元 100 的频谱音频信号。由于 MDCT 和 MDST 分别代表音频信号的实部和虚部, 因此可表示音频信号的相位信息。因此, 可解决 DFT 和 MDCT 之间不匹配的问题。当通过使用经受了 DFT 的时域音频信号量化 MDCT 的系数时发生不匹配的问题。

[0049] ISC 选择单元 150 通过使用 SMR 值从音频信号选择 ISC。ISC 选择单元 150 包括第一 ISC 选择器 152、第二 ISC 选择器 154 和第三 ISC 选择器 156 以分别选择一个或多个第一 ISC、第二 ISC 和第三 ISC。一个或多个第一 ISC、第二 ISC 和 / 或第三 ISC 可被称为 ISC。

[0050] 第一 ISC 选择器 152 通过使用由心理建模单元 100 计算的 SMR 值选择掩蔽阈值小于频谱音频信号的掩蔽阈值的一个或多个频谱信号作为一个或多个第一重要频谱分量 (ISC)。

[0051] 第二 ISC 选择器 154 根据预定权重因数通过从在第一 ISC 选择器 152 中选作为一个或多个第一 ISC 的音频信号提取频谱峰值来选择一个或多个第二 ISC。

[0052] 在一个或多个第一 ISC 中搜索频谱峰值。基于信号的大小确定频谱峰值。由经过 MDCT 和 MDST 变换的信号实部平方加上虚部平方的根来定义信号的大小。通过使用该信号附近的频谱值获得该信号的权重因数。通过使用当前信号 (当前信号的权重因数将被获得) 的频率附近的预定数量的频谱值来获得第二 ISC 选择器 154 中的权重因数。可通过使用等式 1 获得该权重因数。

[0053] 等式 1

$$W_k = \frac{|SC_k|}{\sum_{i=k-len}^{k-1} |SC_i| + \sum_{j=k+1}^{k+len} |SC_j|}$$

[0054]

[0055] 这里, $|SC_k|$ 表示权重因数将被获得的当前信号的大小, $|SC_i|$ 和 $|SC_j|$ 表示当前信号附近的信号的大小。此外, len 表示当前信号附近的信号的数量。

[0056] 基于该信号的峰值和权重因数选择第二 ISC。例如, 峰值和权重因数的乘积与预定阈值进行比较以仅选择大于该阈值的值作为第二 ISC。

[0057] 第三 ISC 选择器 156 对音频信号执行信噪比 (SNR) 均衡。也就是, 该音频信号的频谱分量被分为频带, 并且获得这些频带的 SNR, 在具有低 SNR 的频带中, 峰值大于预定值的频谱分量被选作为一个或多个第三 ISC。执行这种操作来防止 ISC 集中在特定频带上。换句话说, 在具有低 SNR 的频带中选择主要峰值, 从而在整个频带中这些频带的 SNR 近似相等。其结果是, 具有低 SNR 的频带的 SNR 值增加, 从而整个频带的 SNR 值近似相等。

[0058] 组成 ISC 选择单元 150 的第一 ISC 选择器 152、第二 ISC 选择器 154 和第三 ISC 选择器 156 可选择性地用于提取具有感知的重要频谱分量 (ISC) 的音频信号。例如, 仅第一 ISC 选择器 152 和第二 ISC 选择器 154 可被使用。然而, 仅第一 ISC 选择器 152 和第三 ISC 选择器 156 可被使用。否则, 所有的第一 ISC 选择器 152、第二 ISC 选择器 154 和第三 ISC 选择器 156 都可被使用。因此, 可从音频信号提取第一 ISC、第二 ISC 和 / 或第三 ISC 以被用作 ISC, 从而在音频信号的所有频谱分量的量化和 / 或其无损编码中使用提取的 ISC 压缩音频信号。

[0059] 图 2 是示出根据本发明总体发明构思的实施例的提取音频信号的重要频谱分量以按低比特率压缩音频信号的方法的流程图。参照图 1 和图 2, 通过使用心理声学模型计算

变换到频域的音频信号的 SMR 值（操作 200）。接下来，通过使用 SMR 值，在掩蔽阈值低于频域中的音频信号的掩蔽阈值的频谱信号被选作为第一 ISC（操作 220）。

[0060] 根据预定权重因数从选作为第一 ISC 的音频信号提取频谱峰值并将该频谱峰值选作为第二 ISC（操作 240）。可通过使用当前信号（当前信号的权重因数将被获得）的频率附近的预定频率的频谱值来获得权重因数。操作 240 可以与前述图 1 的第二 ISC 选择器 154 的操作相同。因此，省略对其的描述。

[0061] 通过执行 SNR 均衡选择频率（或频带）的第三 ISC（操作 260）。也就是，音频信号的频谱分量被分为频带，获得频带的 SNR，并且在具有低 SNR 的频带中，峰值大于预定值的频谱分量被选作为第三 ISC。第一 ISC、第二 ISC 和第三 ISC 可被统称为 ISC。如上所述，执行这种操作来防止 ISC 集中在特定频带上。换句话说，在具有低 SNR 的频带中选择主要峰值，从而在整个频带中，具有低 SNR 的频带的 SNR 近似相等。其结果是，具有低 SNR 的频带的 SNR 值增加，从而整个频带的 SNR 值近似相等。

[0062] 另一方面，可选择地使用操作 220 至 260 中的 ISC 提取。例如，仅操作 200 和 200 可被用于提取 ISC。然而，仅操作 200 和 260 可用于提取 ISC。否则，所有的操作 200、240 和 260 可用于提取 ISC。

[0063] 图 3 是示出根据本发明总体发明构思的实施例的从输入的音频信号提取重要频谱分量以按低比特率压缩音频信号的方法的示意图。参照图 2 和图 3，例如使用 MDCT 和 MDST 将输入的音频信号变换为频谱音频信号，并且根据与可听见信号和听不见信号相应的心理声学模型的心理特征计算与变换的频谱音频信号相应的信号掩蔽比（SMR）值。可根据 SNR 值、权重因数（或权重最大值）和 / 或 SNR 均衡获得具有第一 ISC、第二 ISC 和 / 或第三 ISC 的频谱音频信号。

[0064] 图 4 是示出根据本发明总体发明构思的实施例的使用提取音频信号的重要频谱分量的设备的低比特率音频信号编码设备的构造的框图。低比特率音频信号编码设备包括 ISC 提取器 420、量化器 440 和无损编码器 460。低比特率音频信号编码设备还可包括 T/F 变换单元 400。

[0065] 参照图 1 和图 4，T/F 变换单元 400 通过使用改进离散余弦变换（MDCT）和改进离散正弦变换（MDST）将时域音频信号变换为频谱信号（频谱音频信号）。通过使用 MDCT 和 MDST（而不是离散傅立叶变换（DFT））产生输入给 ISC 提取器 420 的心理声学模型的频谱音频信号。通过这样做，MDCT 和 MDST 代表实部和虚部，从而可另外表示音频信号的相位分量。因此，可解决 DFT 和 MDST 不匹配的问题。当通过使用经过 DFT 的时域音频信号量化 MDCT 的系数时发生不匹配问题。

[0066] ISC 提取器 420 从频谱音频信号提取具有 ISC 的音频信号。ISC 提取器 420 可以与图 1 的音频信号 ISC 提取设备相同，因此省略对其的描述。也就是，ISC 提取器 420 包括心理建模单元 100 和 ISC 选择单元 150 来选择具有 ISC 的音频信号。

[0067] 量化器 440 量化 ISC 的音频信号。如图 5 所示，量化器 440 包括分组单元 442、量化步长确定单元 444 和量化器 446。

[0068] 分组单元 442 根据使用的比特量和量化误差执行分组以最小化附加信息。下面执行对选择的 ISC 的量化。首先，根据比率失真对选择的 ISC 执行分组以最小化附加信息。比率失真表示使用的比特量和量化误差之间的关系。使用的比特量和量化误差可交替换位。

也就是,如果使用的比特量增加,则量化误差减少。

[0069] 相反,如果使用的比特量减少,则量化误差增加。选择的 ISC 被分组,并且分组的成本被计算。执行分组从而降低成本。

[0070] 各组可以形成相同,并且可以合并,从而降低频带的成本。此外,如等式 2 所示,通过将各组所需的比特数和关于比特数的附加信息相加来获得 成本。

[0071] 等式 2

[0072] 成本 = q_{bit} + 附加信息 [比特数]

[0073] 这里, q_{bit} 表示每一组所需的比特数,附加信息包括缩放因数、量化信息等。

[0074] 当完成分组时,量化步长确定单元 444 根据 SMR 和各组的数据分布 (动态范围) 确定量化步长。此外,采用组成该组的 ISC 的最大值将该 ISC 规格化。

[0075] 量化器 446 量化组的音频信号。通过使用采用组的 ISC 的最大值规格化的值和量化步长来确定量化器 446。

[0076] 量化可以是 Max-Lloyd 量化。

[0077] 无损编码器 460 对量化的信号执行无损编码。如图 6 所示,无损编码器 460 包括索引单元 462 和随机模型无损编码器 464。无损编码可以是上下文算术编码。

[0078] 索引单元 462 产生一个或多个频谱索引以代表构成每一帧的频谱分量。频谱索引指示 ISC 的存在。通过使用上下文算术编码对 ISC 的频谱信息编码。更具体地讲,通过代表 ISC 的选择的频谱索引设置构成每一帧的频谱分量。频谱索引可以是具有代表 ISC 的存在或不存在的 0 或 1 的信号。

[0079] 随机模型无损编码器 464 根据与先前帧的相关性和相邻 ISC 的分布选择随机模型,并且对音频信号的量化值和附加信息 (包括量化器信息、量化步长、分组信息和频谱索引信息) 执行无损编码。

[0080] 图 7 是示出根据本发明总体发明构思的实施例的使用音频信号 ISC 提取方法的低比特率音频信号编码方法的流程图。

[0081] 参照图 4 和图 7,通过使用改进离散余弦变换 (MDCT) 和改进离散正弦变换 (MDST) 来将时域音频信号变换为频谱信号 (操作 700)。变换的频谱音频信号被输入到心理声学模型。在心理声学模型中,计算信号掩蔽比 (SMR) 以预测频谱音频信号的重要性 (操作 720)。通过使用 SMR 值提取 ISC (操作 740)。该 ISC 提取可以与图 2 的 ISC 提取方法相同,因此省略对其的描述。

[0082] 在提取 ISC 之后,执行 ISC 量化 (操作 760)。在图 8 中示出 ISC 量化的详细操作。参照图 8,根据使用的比特量和量化误差之间的关系执行分组以最小化附加信息 (操作 762)。该分组可以与图 5 的分组单元 442 的分组相同,因此省略对其的描述。

[0083] 在分组之后,根据 SMR 和各组的数据分布 (动态范围) 确定量化步长 (操作 764)。此外,采用 ISC 的最大值将组成组的 ISC 规格化。

[0084] 接下来,通过使用采用组的最大值规格化的值和量化步长确定量化器。

[0085] 量化可以是 Max-Lloyd 量化。

[0086] 参照图 7,在量化之后,执行无损编码 (操作 780)。通过上下文算术编码对 ISC 的量化值和频谱信息编码。此外,通过代表 ISC 的选择的频谱索引设置组成每一帧的频谱分量。频谱索引分别采用 0 和 1 代表 ISC 的存在和不存在。接下来,对频谱索引的值编码。

根据与先前帧的相关性和相邻 ISC 的分布选择随机模型,并且执行无损编码。接下来,对编码值执行比特打包。

[0087] 图 9 是示出对使用提取音频信号的重要频谱分量的设备编码的低比特率音频信号进行解码的低比特率音频信号解码设备的框图。低比特率音频信号解码设备包括无损解码器 900、逆量化器 920 和 F/T 变换单元 940。

[0088] 无损解码器 900 提取各组的随机模型信息,并且通过使用随机模型信息恢复各组的指示 ISC 的存在的索引信息、量化器信息、量化步长、ISC 分组信息和音频信号量化值。

[0089] 逆量化器 920 参照恢复的量化器信息、量化步长和分组信息执行逆量化。

[0090] F/T 变换单元 940 将逆量化的值变换为时域信号。

[0091] 图 10 是示出根据本发明总体发明构思的实施例的对使用提取具有 ISC 的音频信号的设备编码的低比特率音频信号进行解码的低比特率音频信号解码方法的流程图。将参看图 9 和图 10 描述低比特率音频信号解码方法和设备的操作。

[0092] 首先,通过无损解码器 900 提取帧的随机模型信息(操作 1000)。接下来,通过使用随机模型信息恢复指示 ISC 的存在的索引信息、量化器信息、量化步长、ISC 分组信息和音频信号量化值(操作 1020)。接下来,由逆量化器 920 根据恢复的量化器信息、量化步长和分组信息对量化值逆量化(操作 1040)。在逆量化之后,通过 F/T 变换单元 940 将逆量化的值变换为时域信号(操作 1060)。

[0093] 根据提取具有 ISC 的音频信号的方法和设备以及使用该方法和设备的低比特率音频信号编码/解码方法和设备,能够有效地对感知重要频谱分量编码以获得低比特率的高声音质量。此外,能够通过使用心理声学模型提取感知重要分量,无需相位信息执行编码,并且有效地代表低比特率频谱信号。此外,可在需要低比特率音频编码方案的所有应用中和下一代音频方案中应用本发明。

[0094] 本发明总体发明构思也可实现为计算机可读记录介质上的计算机可读代码。计算机可读记录介质是可存储其后由计算机系统读取的数据的任何数据存储装置。计算机可读记录介质的例子包括只读存储器(ROM)、随机存取存储器(RAM)、CD-ROM、磁带、软盘、光学数据存储装置和载波(例如,通过互联网的数据传输)。计算机可读记录介质也可分布在网络连接的计算机系统,从而以分布方式存储和执行计算机可读代码。此外,本发明所属领域的编程人员容易解释实现本发明的功能性程序、代码和代码段。

[0095] 尽管已经显示和描述了本发明总体发明构思的一些实施例,但是本领域的技术人员应该理解,在不脱离本发明总体发明构思的原理和精神的情况下,可以对这些实施例进行改变,在权利要求及其等同物中限定本发明总体发明构思的范围。

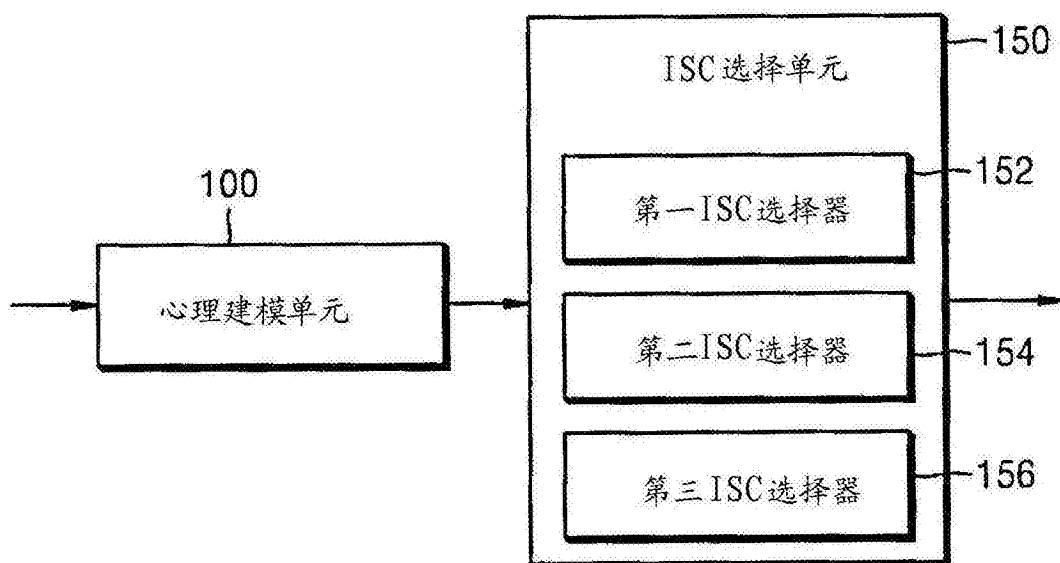


图 1

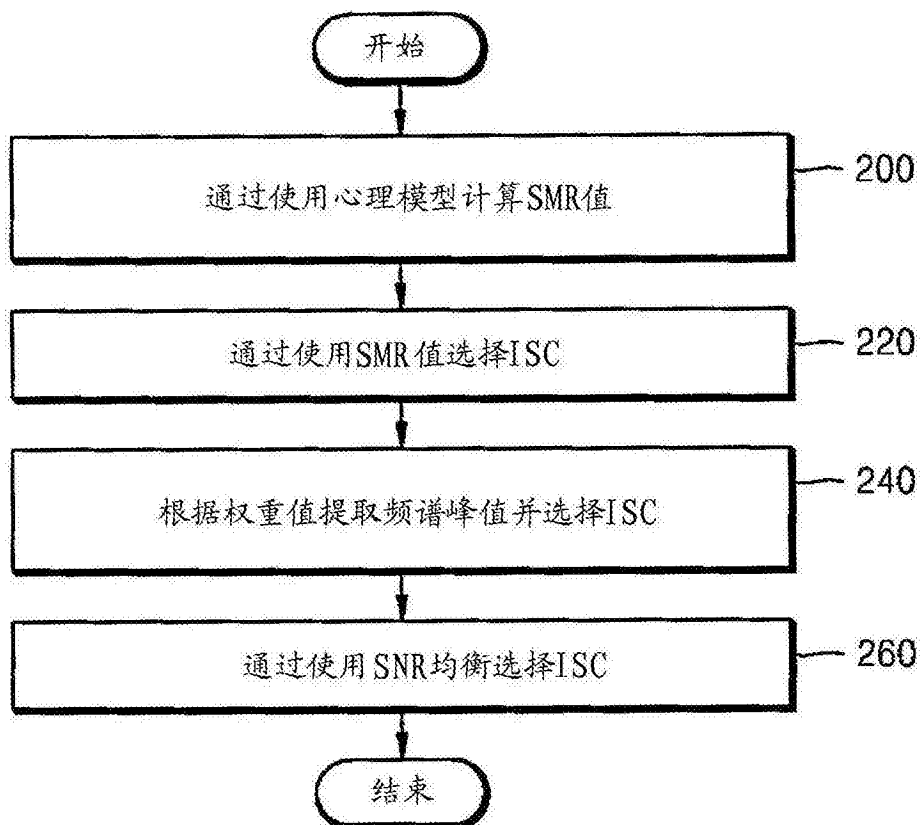


图 2

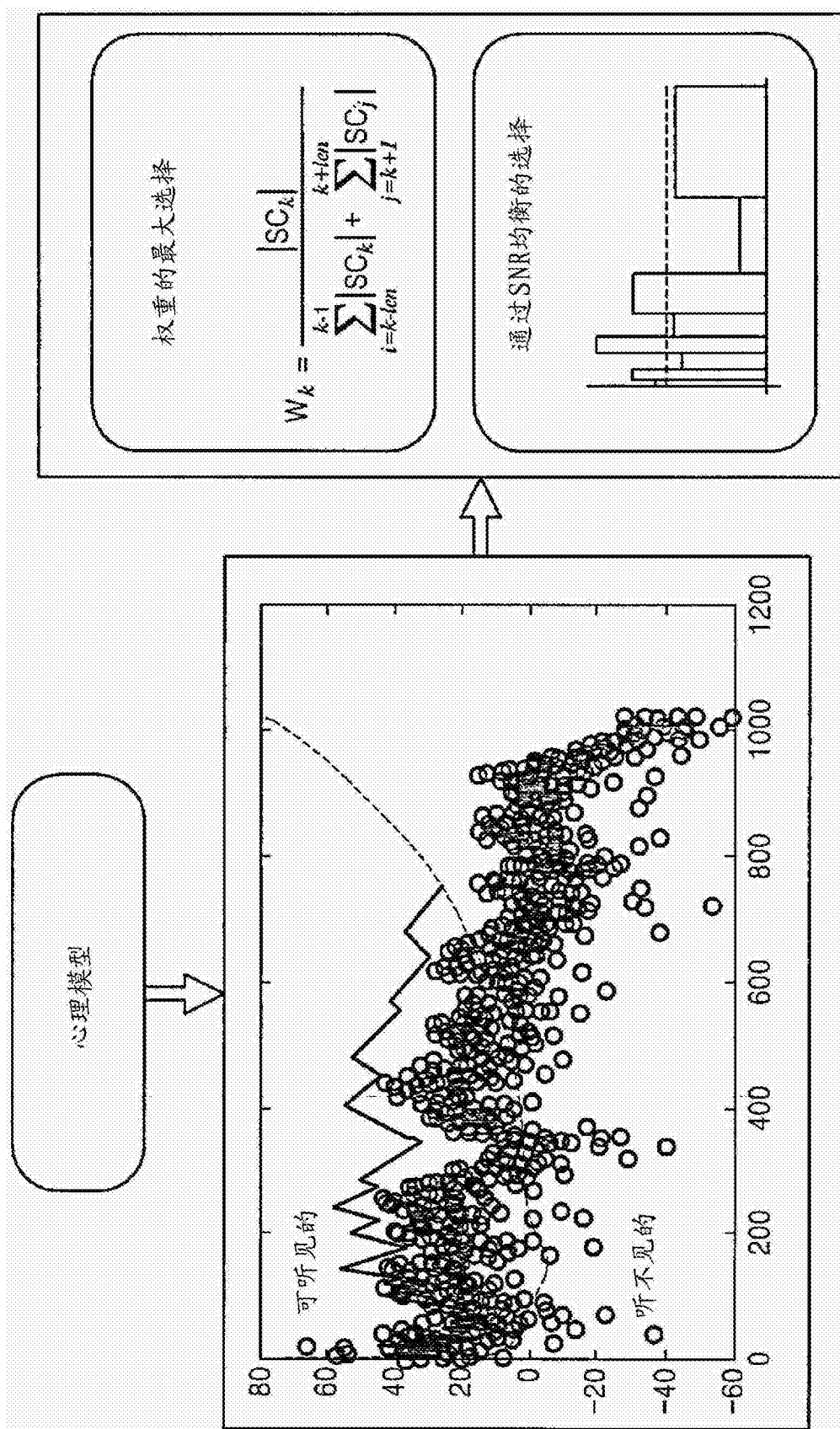


图 3

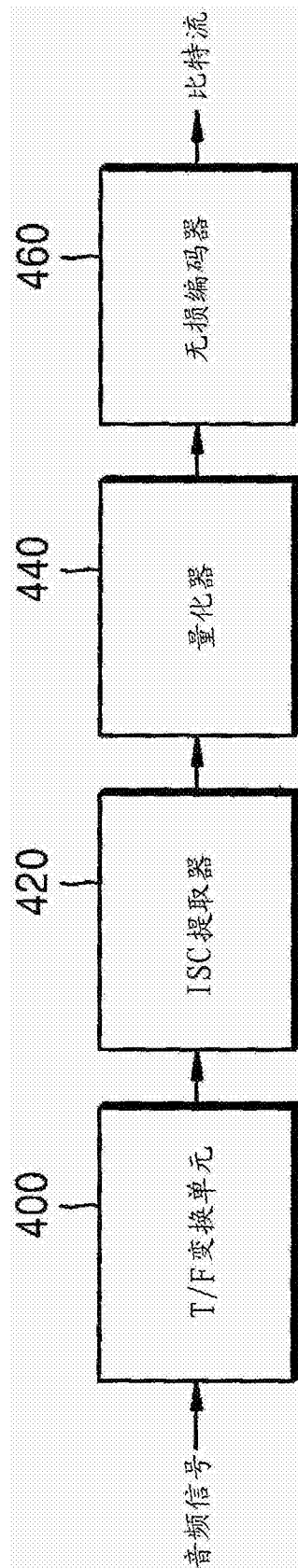


图 4

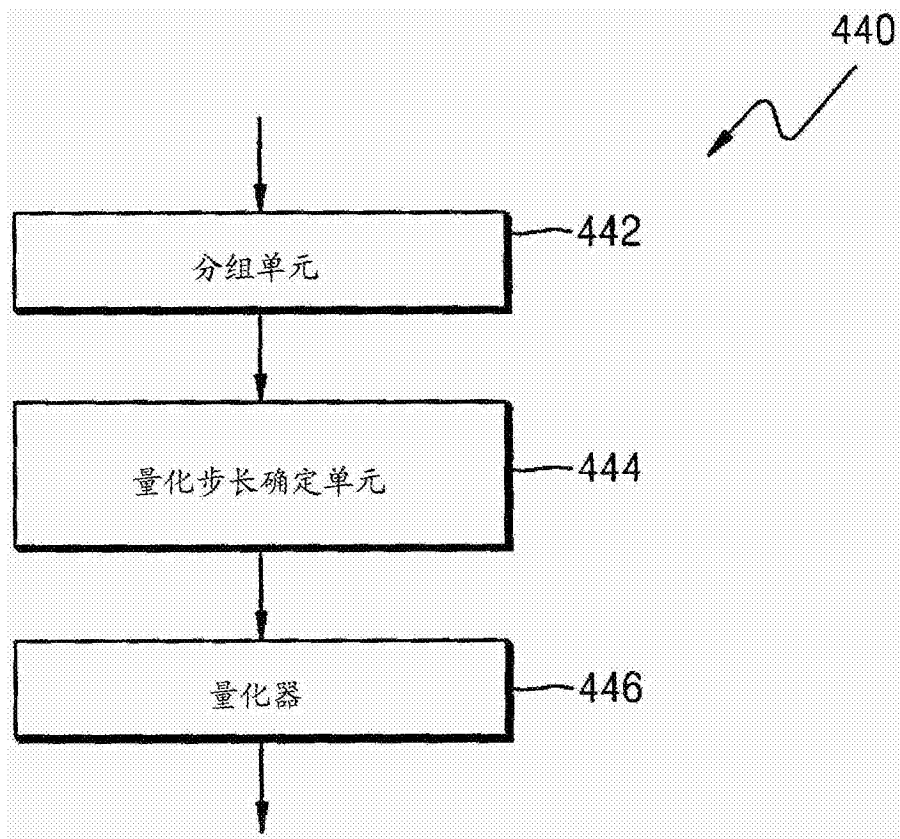


图 5

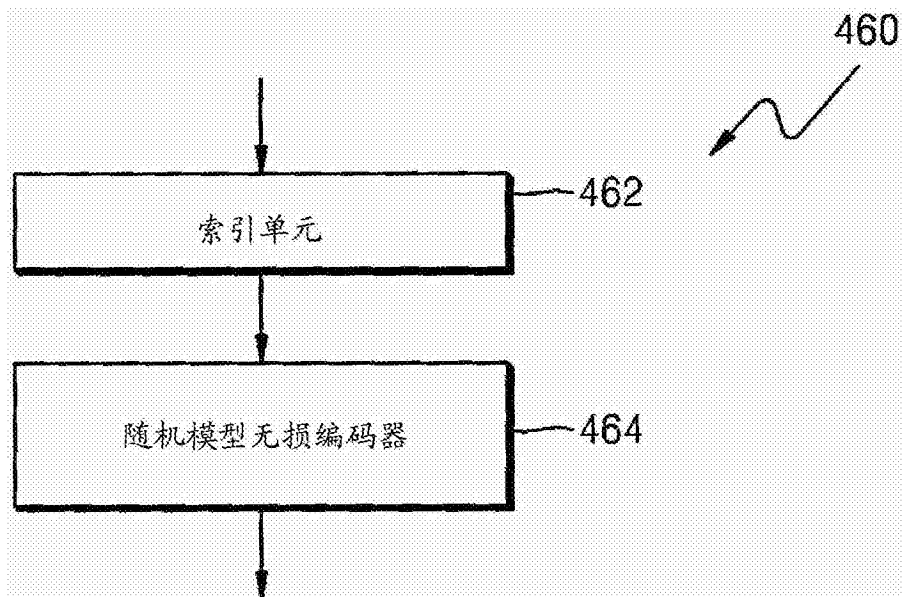


图 6

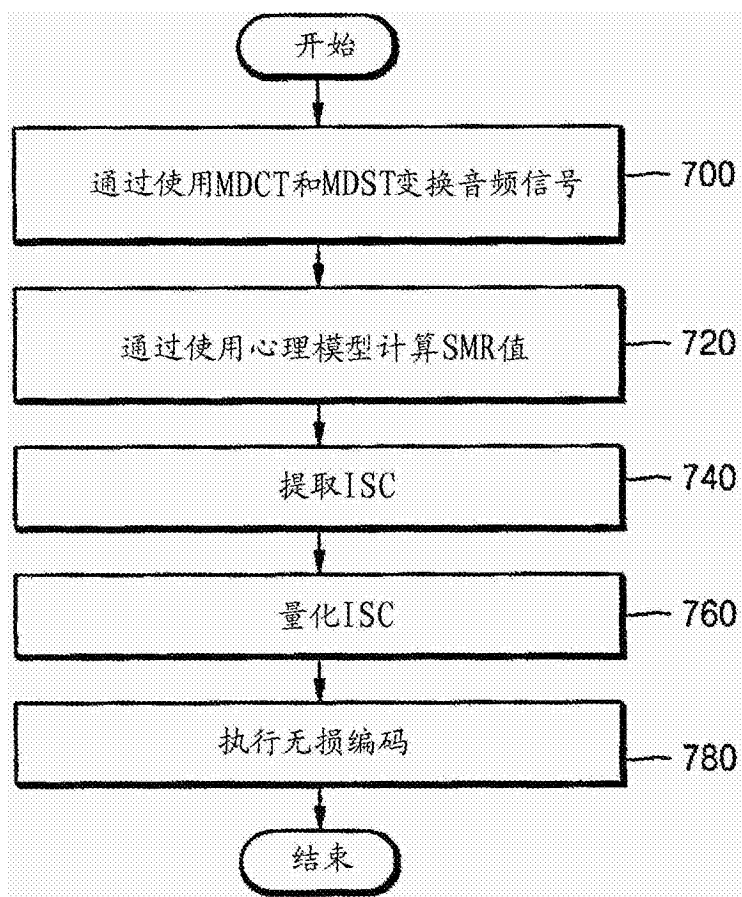


图 7

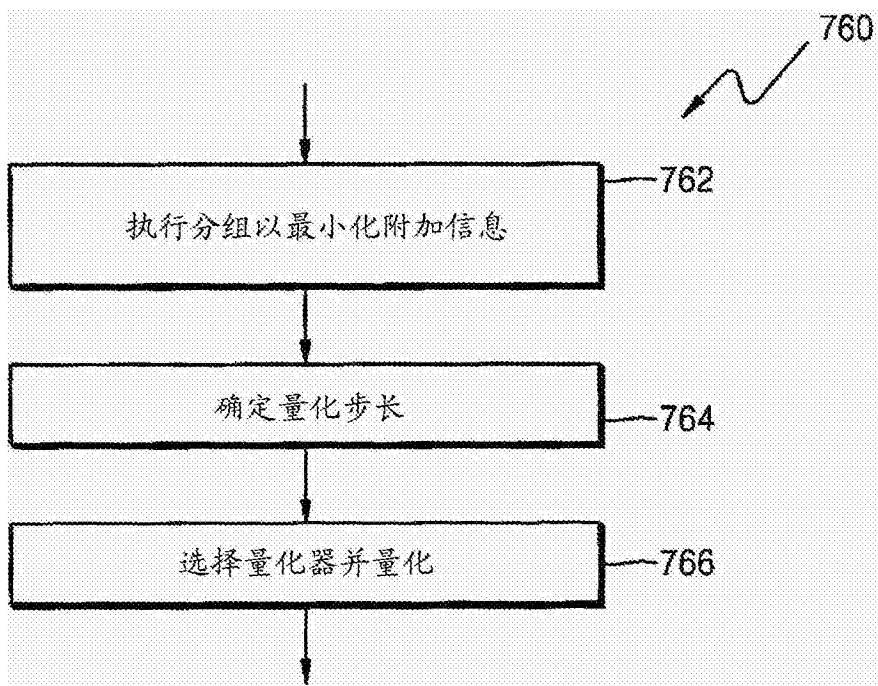


图 8

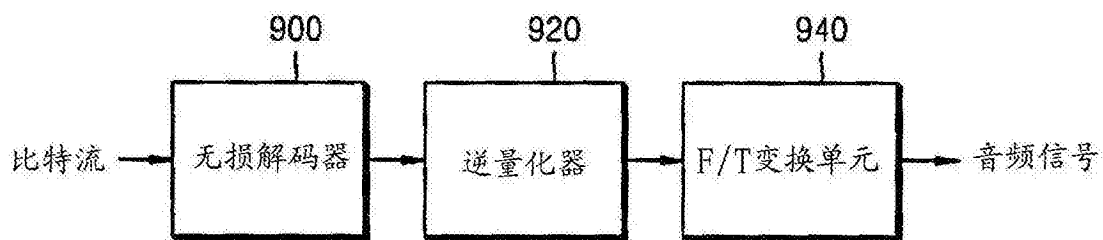


图 9

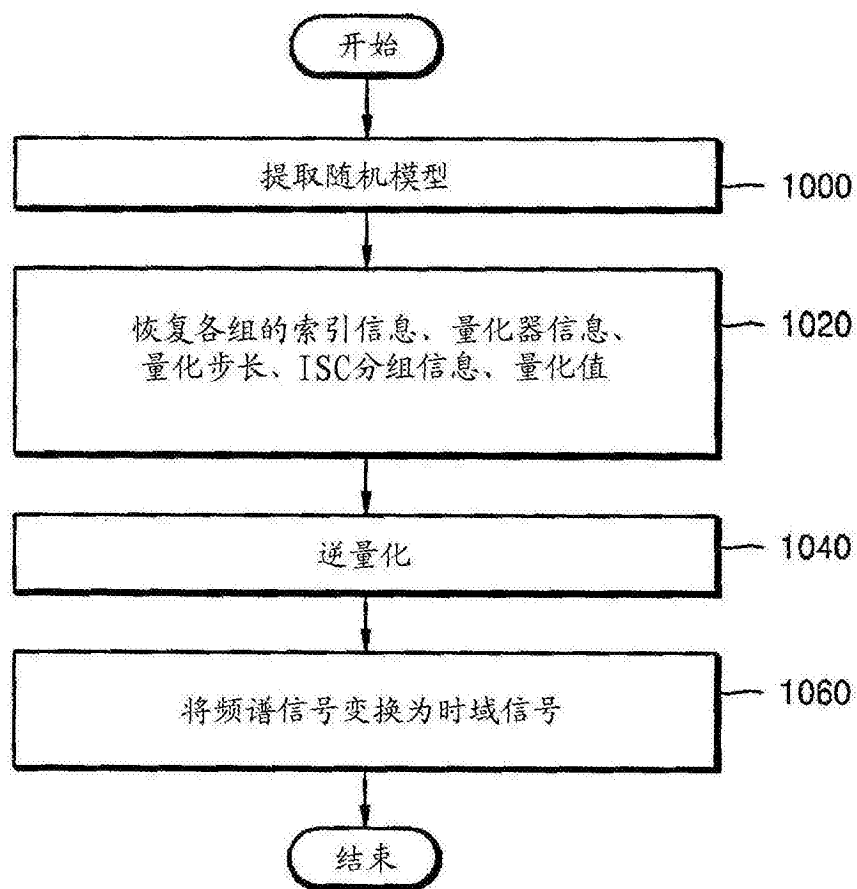


图 10