



(12)发明专利

(10)授权公告号 CN 105451842 B

(45)授权公告日 2019.06.11

(21)申请号 201580000798.2

马库斯·穆赖特鲁斯

(22)申请日 2015.07.21

斯特凡·多赫拉

(65)同一申请的已公布的文献号

伯恩哈德·格里尔

申请公布号 CN 105451842 A

曼努埃尔·扬德尔

(43)申请公布日 2016.03.30

(74)专利代理机构 北京康信知识产权代理有限公司 11240

(30)优先权数据

代理人 梁丽超 陈鹏

14178809.1 2014.07.28 EP

(51)Int.CI.

G10L 19/22(2013.01)

(85)PCT国际申请进入国家阶段日

2015.11.20

(56)对比文件

(86)PCT国际申请的申请数据

CN 103620672 A, 2014.03.05,

PCT/EP2015/066677 2015.07.21

US 2008312914 A1, 2008.12.18,

(87)PCT国际申请的公布数据

CN 1957398 A, 2007.05.02,

W02016/016053 EN 2016.02.04

CN 1708997 A, 2005.12.14,

(73)专利权人 弗劳恩霍夫应用研究促进协会

CN 103000178 A, 2013.03.27,

地址 德国慕尼黑

审查员 薛彦杰

(72)发明人 埃曼努埃尔·拉维利

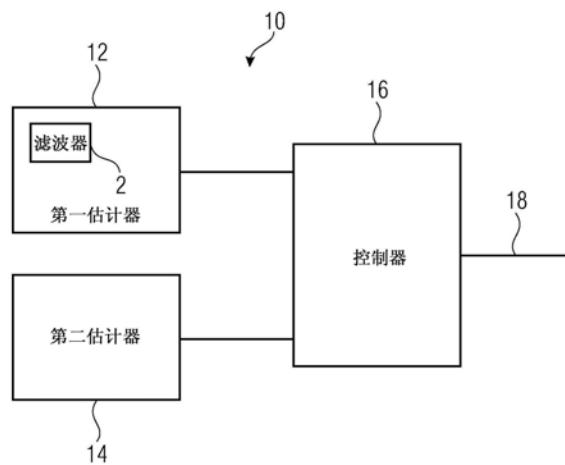
权利要求书3页 说明书19页 附图4页

(54)发明名称

选择第一编码演算法和第二编码演算法之一的装置与方法

(57)摘要

一种装置用于选择第一编码演算法以及第二编码演算法的其中之一以编码一音信,以得到音信的一编码版,一滤波器接收音信、减少音信的谐波的振幅、并输出音信的一滤波版。第一估计器使用音信的滤波版来估计音信的一信噪比或一分段式信噪比,以作为音信的第一品质测量,非实际上使用第一编码演算法来编码与解码音信。第二估计器估计一信噪比或一分段式信噪比以作为音信的第二品质测量,非实际上使用第二编码演算法来编码与解码音信。控制器依据在第一品质测量与第二品质测量之间的比较而选择第一编码演算法或第二编码演算法。



1. 一种用于选择具有一第一特征的一第一编码演算法以及具有一第二特征的一第二编码演算法的其中之一以编码一音信(40)的一部分从而得到该音信的该部分的一编码版的装置(10),该装置包含:

一长期预测滤波器,接收该音信、减少该音信的谐波的振幅、并输出该音信的一滤波版;

一第一估计器(12),使用该音信的该滤波版来估计该音信的该部分的一信噪比或一分段式信噪比,以作为该音信的该部分的一第一品质测量,该第一品质测量关于第一编码演算法,其中估计该第一品质测量包含执行该第一编码演算法的一近似,以得到该第一编码演算法的一变形估计并基于该音信的该部分与该第一编码演算法的该变形估计来估计该第一品质测量,且非实际上使用该第一编码演算法来编码与解码该音信的该部分;

一第二估计器(14),估计一信噪比或一分段式信噪比以作为该音信的该部分的一第二品质测量,该第二品质测量关于该第二编码演算法,其中估计该第二品质测量包含执行该第二编码演算法的一近似,以得到该第二编码演算法的一变形估计并藉由使用该音信的该部分与该第二编码演算法的该变形估计来估计该第二品质测量,且非实际上使用第二编码演算法来编码与解码该音信的该部分;

一控制器(16),依据在第一品质测量与第二品质测量之间的一比较而选择该第一编码演算法或该第二编码演算法,

其中,该第一编码演算法为一转换编码演算法,一改进离散余弦变换MDCT的编码演算法或一转换编码激励TCX编码演算法,并且该第二编码演算法为一码激励线性预测CELP编码演算法或一代数码激励线性预测ACELP编码演算法。

2. 根据权利要求1所述的装置(10),其中该长期预测滤波器的一转移函数包含一音高延迟的一整数部分以及依据该音高延迟的一分数部分的一多阀门滤波器。

3. 根据权利要求1所述的装置(10),其中该长期预测滤波器具有如下转移函数:

$$P(z) = 1 - \beta g B(z, T_{fr}) z^{-T_{int}}$$

其中,Tint与Tfr分别为一音高延迟的整数与分数部分,g为一增益,β为一权重,B(z,Tfr)为一有限脉冲响应低通滤波器,其系数依据该音高迟的分数部分。

4. 根据权利要求1所述的装置,还包含一失能单元,其基于至少一谐波性测量及/或至少一时间结构测量的一组合而失能该滤波器。

5. 根据权利要求4所述的装置,其中该谐波性测量包含正规化关联性与预测增益的至少其中之一,该至少一时间结构测量包含时间平整性测量与能量改变的至少其中之一。

6. 根据权利要求1所述的装置,其中该滤波器以一画面接画面基准而应用到该音信,该装置还包含一单元,其移除该音信中由该滤波器所造成的不连续。

7. 根据权利要求1所述的装置(10),其中该第一估计器与该第二估计器被配置为估计该音信的一加权重版的一部分的一信噪比或一分段式信噪比。

8. 根据权利要求1所述的装置(10),其中当量化该音信的该部分时,该第一估计器(12)被配置为决定由用于该第一编码演算法的一量化器所引入的一被估计量化器变形,并且基于该音信的一加权重版的一部分的一能量以及该被估计量化器变形而估计该第一品质测量,其中当藉由用于该第一编码演算法的一量化器与一熵变码器而被编码时,该第一估计

器(12)被配置为估计该音信的该部分的一全局增益,以致该音信的该部分产生一给定目标比特率,其中该第一估计器(12)还被配置为基于该被估计全局增益来决定该被估计量化器变形。

9.根据权利要求1所述的装置(10),其中该第二估计器(14)被配置为决定一被估计调适码本变形,当使用一调适本来编码该音信的该部分时,该被估计调适码本变形由用于该第二编码演算法的该调适码本所引入,其中该第二估计器(14)被配置为基于该音信的一加权重版的一部分的一能量以及该被估计调适码本变形来估计该第二品质测量,其中对于该音信的该部分的多个子部分,该第二估计器(14)被配置为基于藉由在一预处理阶段中所决定的一音高延迟而切换到过去的该加权重音信的该子部分的一版本而近似该调适码本,并且估计一调适码本增益,以致在该加权重音信的该部分的该子部分与该被近似调适码本之间的一误差能最小化,并且基于在该加权重音信的该部分的该子部分与藉由该调适码本增益而缩放的该被近似调适码本之间的一误差的能量而决定该被估计调适码本变形。

10.根据权利要求9所述的装置(10),其中该第二估计器(14)还被配置为减少该被估计调适码本变形,其为了该音信的该部分的各子部分而藉由一固定因子所决定。

11.根据权利要求1所述的装置(10),其中该第二估计器(14)被配置为决定一被估计调适码本变形,当使用一调适本来编码该音信的该部分时,该被估计调适码本变形由用于该第二编码演算法的该调适码本所引入,其中该第二估计器(14)被配置为基于该音信的一加权重版的一部分的一能量以及该被估计调适码本变形来估计该第二品质测量,其中该第二估计器(14)被配置为基于藉由在一预处理阶段中所决定的一音高延迟而切换到过去的该加权重音信的该部分的一版本而近似该调适码本,并且估计一调适码本增益,以致在该加权重音信的该部分与该被近似调适码本之间的一误差能最小化,并且基于在该加权重音信的该部分与藉由该调适码本增益而缩放的该被近似调适码本之间的一误差的能量而决定该被估计调适码本变形。

12.一种用以编码一音信的一部分的装置(20),其包含根据权利要求1至11中任一项所述的装置(10)、用以执行该第一编码演算法的一第一编码器阶段(26)以及用以执行该第二编码演算法的一第二编码器阶段(28),其中用于编码的该装置(20)被配置为依据藉由控制器(16)的选择而使用该第一编码演算法或该第二编码演算法以编码该音信的该部分。

13.一种用于编码与解码的系统,其包含根据权利要求12所述的用于编码的一装置(20)以及一解码器,该解码器被配置为接收该音信的该部分的该编码版与该演算法的一指示,该演算法用来编码该音信的该部分并解码该音信的该部分的该编码版。

14.一种用于选择具有一第一特征的一第一编码演算法以及具有一第二特征的一第二编码演算法的其中之一以编码一音信的一部分从而得到该音信的该部分的一编码版的方法,该方法包含:

使用一长期预测滤波器来滤波该音信以减少该音信的谐波的振幅并输出该音信的一滤波版;

使用该音信的滤波版来估计音信的该部分的一信噪比或一分段式信噪比,以作为该音信的该部分的一第一品质测量,该第一品质测量关于该第一编码演算法,其中估计该第一品质测量包含执行该第一编码演算法的一近似,以得到该第一编码演算法的一变形估计并基于第一音信的该部分与该第一编码演算法的该变形估计来估计该第一品质测量,且非实

际上使用该第一编码演算法来编码与解码该音信的该部分；

估计一信噪比或一分段式信噪比作为该音信的该部分的一第二品质测量，该第二品质测量关于该第二编码演算法，其中估计该第二品质测量包含执行该第二编码演算法的一近似，以得到该第二编码演算法的一变形估计并藉由使用该音信的该部分与该第二编码演算法的该变形估计来估计该第二品质测量，且非实际上使用该第二编码演算法来编码与解码该音信的该部分；以及

基于在该第一品质测量与该第二品质测量之间的一比较而选择该第一编码演算法或该第二编码演算法，

其中该第一编码演算法为一转换编码演算法，一改进离散余弦变换MDCT的编码演算法或一转换编码激励TCX编码演算法，并且该第二编码演算法为一码激励线性预测CELP编码演算法或一代数码激励线性预测ACELP编码演算法。

15. 一种存储介质，包括具有一程序代码的计算机程序，当该程序代码执行于一包括所述存储介质的计算机上时，使所述计算机执行根据权利要求14所述的方法。

选择第一编码演算法和第二编码演算法之一的装置与方法

技术领域

[0001] 本发明关于一种音信编码,更关于一种被切换的音信编码,其中对于一声音信号的不同部分,被编码的信号系藉由使用不同编码演算法而产生。

背景技术

[0002] 在习知技术中,已有一些被切换的音信编码器可以为一音信的不同部分决定不同的编码演算法。一般而言,被切换的音信编码器为了两个不同模式之间的切换而提供,也就是演算法,例如代数码激励线性预测 (Algebraic Code Excited Linear Prediction, ACELP) 与转换编码激励 (Transform Coded Excitation, TCX)。

[0003] 动态影像专家组USAC (MPEG Unified Speech Audio Coding) 的线性预测域 (linear prediction domain, LPD) 基于两个不同的模式ACELP、TCX。ACELP系提供较佳的品质给类语音 (speech-like) 以及类瞬时 (transient-like) 的信号。TCX提供较佳的品质给类音乐 (music-like) 以及类噪音 (noise-like) 的信号。编码器决定在一画面接画面 (frame-by-frame) 基准上使用哪一种模式。编码器所作的决定对于编解码品质是相当重要的。单个错误决定能产生强大的人为影响,特别是在低比特率的情况。

[0004] 决定使用哪一种模式最直接的方式系为一种闭回路模式选择,也就是先执行两种模式的一完整的编码/解码,然后基于音信以及编码/解码音信来计算两个模式的一选择标准(例如分段式信噪比segmental SNR),最后基于选择标准来选择一模式。这种方式一般皆能产生一个稳定且强健的决定。然而,它也需要大量的复杂度,因为两种模式必需在每一画面中来执行。

[0005] 为减少复杂度,另一种方式系为开回路模式选择。开回路选择不是对两种模式执行完整的编码/解码,而是藉由使用由低复杂度所计算出来的一选择标准而选择其中一模式。然后,最差情况复杂度可藉由最少复杂度模式 (通常是TCX) 而被减少,亦即减少计算选择标准所需的复杂度。在复杂度上的节省通常是具大的,以致当编解码最差情况复杂度处于受限时,这种方式能具有吸引力。

[0006] AMR-WB+标准 (定于国际标准3GPP TS 26.290V6.1.0 2004-12) 包含一种开回路模式选择,其使用于在一个80毫秒的画面中,并在 ACELP/TCX20/TCX40/TCX80的所有组合之间进行决定。它被描述于 3GPP TS 26.290中的5.2.4章节。它亦被描述于研讨会论文“Low Complex Audio Encoding for Mobile, Multimedia, VTC 2006, Makinen et al.”以及美国专利(US 7,747,430B2and US 7,739,120B2),且上述文献皆相同作者。

[0007] 美国专利US 7,747,430 B2揭露一种基于长期预测参数之一分析之一开回路模式选择,美国专利US 7,739,120 B2揭露一种基于指出一音信的各部分的音信内容之类型之信号特征之开回路模式选择。其中,假使这样一个选择系为不可行,则该选择系更基于一种统计估价 (为了音信的相邻部分而实行)。

[0008] AMR-WB+之开回路模式选择可描述于两个主要的步骤。在第一主要步骤中,多个特征系计算于音信上,例如能量等级之标准偏差、低频/高频能量关系、总能量、导抗频谱对

(immittance spectral pair, ISP) 距离、音高延迟 (pitch lag) 与增益以及频谱倾斜。然后, 藉由使用一简单的阀值应用分类器 (threshold-based classifier), 这些特征可使用来在 ACELP 与 TCX 之间作出一选择。假使 TCX 在第一主要步骤中被选择, 则第二主要步骤系在 TCX20/TCX40/TCX80 之可能组合之间并以一闭回路方式来选择。

[0009] 专利 WO 2012/110448 A1 系揭露基于一音信之一瞬时检测结果以及一品质结果而在两个编码演算法 (具有不同特征) 之间进行选择之一方法。此外, 其系揭露一种滞后 (hysteresis), 其中该滞后系依赖过去的选择模式, 亦即音信之早期部分。

[0010] 在研讨会论文 “Low Complex Audio Encoding for Mobile, Multimedia, VTC 2006, Makinen et al.” 中, AMR-WB+ 之闭回路以及开回路模式选择系被比较。主观的听力测验系指出开回路模式选择相较于闭回路模式选择有较差的执行。但它也指出开回路模式选择可减少最差情况复杂度约 40%。

发明内容

[0011] 本发明之一目的在于提供一种改良的方法, 其系能在一第一编码演算法以及一第二编码演算法之间进行选择, 且可达到较佳的性能并减少复杂度。

[0012] 本发明之实施例系提供可选择具有一第一特征的一第一编码演算法以及具有一第二特征的一第二编码演算法的其中之一的一装置, 以编码一音信的一部分, 以得到该音信的该部分的一编码版, 其包含:

[0013] 一滤波器, 接收该音信、减少音信的谐波的振幅、并输出该音信的一滤波版;

[0014] 一第一估计器, 使用该音信的滤波版来估计音信的该部分的一信噪比或一分段式信噪比, 以作为该音信的该部分的一第一品质测量, 其关于第一编码演算法, 但非实际上使用第一编码演算法来编码与解码该音信的该部分;

[0015] 一第二估计器, 估计一信噪比或一分段式信噪比以作为该音信的该部分的一第二品质测量, 其关于第二编码演算法, 但非实际上使用第二编码演算法来编码与解码该音信的该部分; 以及

[0016] 一控制器, 依据在第一品质测量与第二品质测量之间的一比较而选择第一编码演算法或第二编码演算法。

[0017] 本发明之实施例系提供可选择具有一第一特征的一第一编码演算法以及具有一第二特征的一第二编码演算法的其中之一之一方法, 以编码一音信的一部分, 以得到该音信的该部分的一编码版, 其包含:

[0018] 滤波该音信以减少音信的谐波的振幅并输出该音信的一滤波版;

[0019] 使用该音信的滤波版来估计音信的该部分的一信噪比或一分段式信噪比, 以作为该音信的该部分的一第一品质测量, 其关于第一编码演算法, 但非实际上使用第一编码演算法来编码与解码该音信的该部分;

[0020] 估计该音信的该部分的一第二品质测量, 其关于第二编码演算法, 但非实际上使用第二编码演算法来编码与解码该音信的该部分; 以及

[0021] 依据在第一品质测量与第二品质测量之间的一比较而选择第一编码演算法或第二编码演算法。

[0022] 本发明之实施例基于下面识别, 即藉由第一及第二编码演算法的各品质测量以及

依据在第一与第二品质测量之间的一比较而选择该等编码演算法的其中之一,具有较佳性能之一开回路选择可被实现。品质测量可被估计,亦即音信并非实际地被编码与解码以得到品质测量。因此,品质测量可在复杂度减少的情况下被得到。然后,与一闭回路模式选择相比,模式选择可藉由使用被估计之品质测量而执行。此外,本发明基于下面识别,即假使第一品质测量之估计使用音信的该部分的一滤波版,则可得到改良的模式选择,其中相比于音信之非滤波版,谐波可被减少。

[0023] 在本发明的实施例中,先实现一开回路模式选择,在其中ACELP与 TCX之分段式信噪比先被估计并且具有低复杂度。然后,藉由使用这些被估计之分段式信噪比值而执行模式选择,就如在一闭回路模式选择中一样。

[0024] 本发明的实施例并非使用一种习知特征加上分类器的方法,就如在 AMR-WB+之开回路模式选择所作的一样。相反地,本发明的实施例先估计各模式之一品质测量,再选择能给出最佳品质之模式。

附图说明

[0025] 图1为本发明一实施例的一装置的示意图,该装置选择一第一编码演算法以及一第二编码演算法的其中之一。

[0026] 图2系为本发明一实施例的编码一音信的一装置的示意图。

[0027] 图3为本发明一实施例的一装置的示意图,该装置选择一第一编码演算法以及一第二编码演算法的其中之一。

[0028] 图4a与图4b为信噪比与分段式信噪比之可能态样。

具体实施方式

[0029] 以下将参照相关图式,说明依本发明较佳实施例之一种藉由使用谐波抑制而从一第一编码演算法以及一第二编码演算法中选择其中之一的装置与方法,其中相同的组件将以相同的参照符号加以说明。

[0030] 在下面的叙述中,不同图式之相似的组件/步骤由相同的标号所表示。需注意者,在图式中,一些属于了解本发明之非必要的东西,例如信号连接或类似的,在本发明中系省略。

[0031] 图1系显示一装置10,其选择一第一编码演算法(例如TCX演算法)以及一第二编码演算法(例如ACELP演算法)的其中之一,并作为编码器以编码一音信的一部分。装置10包含一第一估计器12,其估计音信的该部分的一信噪比或一分段式信噪比,以作为该音信部分的一第一品质测量。第一品质测量关于第一编码演算法。装置10包含一滤波器2,其接收该音信、减少音信的谐波的振幅、并输出该音信的一滤波版。滤波器2可在第一估计器12内,就如图1所示,或者在第一估计器12之外。第一估计器12使用音信的滤波版来估计第一品质测量。换言之,第一估计器12 估计音信的该部分所应有的第一品质测量,但并非实际上使用第一编码演算法来编码与解码该音信的该部分。装置10包含一第二估计器14,其估计音信部分的一第二品质测量。第二品质测量关于第二编码演算法。换言之,第二估计器14估计音信的该部分所应有的第二品质测量,但并非实际上使用第二编码演算法来编码与解码该音信的该部分。另外,装置10 包含一控制器16,其依据在第一品质测量与第二品质测量之

间的一比较而选择第一编码演算法或第二编码演算法。控制器可包含一输出18,其系指出被选择的编码演算法。

[0032] 在下面叙述中,假使滤波器2被提供来减少谐波振幅并且没有失能,则第一估计器使用音信的滤波版,也就是音信的该部分的滤波版,以估计第一品质测量,即使在没有明确指出的情况下。

[0033] 在一实施例中,第一编码演算法之第一特征系较适合应用于类音乐与类噪音之信号,而第二编码演算法之第二特征系较适合应用于类语音与类瞬时之信号。在本发明之实施例中,第一编码演算法系为一音信编码演算法,就如一变换编码演算法(transform coding algorithm),例如一改进离散余弦变换(modified discrete cosine transform, MDCT),就如TCX编码演算法。其他变换编码演算法可基于一快速傅立叶变换、或任何其他变换或滤波器组(filterbank)。在本发明之实施例中,第二编码演算法系为一语音编码演算法,就如码激励线性预测(code excited linear prediction, CELP)编码演算法,就如ACELP编码演算法。

[0034] 在实施例中,品质测量系代表一种感知品质测量。在实施例中,系计算一种作为第一编码演算法之主观品质之一估计之单一值以及作为第二编码演算法之主观品质之一估计之单一值。基于这两个值的比较而可选择能给出最佳估计主观品质的编码演算法。这与在AMR-WB+标准中所作的不同,在其中代表信号的不同特征的许多特征先被计算,然后一分类器被应用来决定选择哪一个演算法。

[0035] 在实施例中,各别的品质测量基于加权重(weighted)的音信的一部分而估计,也就是音信的一加权重版。在实施例中,加权重音信可定义为被一权重函数所滤波之一音信,其中该权重函数系为一加权重线性预测编码(LPC)滤波A(z/g),其中A(z)为一LPC滤波器,g为一介于0与1之间的权重,如0.68。藉此方法就可得到良好的感知品质之测量。需注意者,LPC滤波器A(z)与加权重LPC滤波器A(z/g)是在一前处理阶段所决定的,并且它们也使用于两个编码演算法中。在其他实施例中,权重函数可为一线性滤波器、一有限脉冲响应(FIR)滤波器或一线性预测滤波器。

[0036] 在实施例中,品质测量系为在加权重信号域中的分段式信噪比。如此,在加权重信号域中之分段式信噪比系代表一个良好的感知品质之测量,并且因此可以一个有利的方式作为品质测量。这也是使用于ACELP与TCX 编码演算法中的品质测量,以估计编码参数。

[0037] 另一个品质测量可为在加权重信号域中的信噪比。其他的品质测量可为分段式的信噪比,也就是在非加权重信号域中的音信的对应部分之信噪比,也就是没有被(加权重)之线性预测编码参数所滤波。

[0038] 一般而言,信噪比系一个取样接一个取样的比较原有以及处理过的音信(例如语音信号)。它的目的在于测量重现输入波形之波形编码器之变形。信噪比可如图4a所计算,其中x(i)与y(i)分别为由i所表示之原有以及处理过之取样,并且N为取样的总数。分段式的信噪比在没有以整个信号来进行的情况下,系计算多个短分段(例如1到10毫秒,就如5毫秒)的信噪比值之平均。信噪比可如图4b所计算,其中N与M分别为分段长度与分段的总量。

[0039] 在本发明的实施例中,音信的该部分系代表藉由窗口化(windowing) 音信所得到的音信之一画面(frame),并且一合适的编码演算法的选择为了多个由窗口化一音信所得到之连续的画面而执行。在下面的叙述中,在与音信的关连下,「部分」与「画面」这两个词是

可交换的。在实施例中,各画面系分为多个子画面,并且分段式的信噪比为了各画面并藉由计算各子画面的信噪比而被估计的,并且被转换为dB的单位且计算出子画面信噪比的平均值(dB)。

[0040] 因此,在实施例中,并非估计出输入音信与被解码音信之间的(分段式)信噪比,而是估计出加权重输入音信与加权重被解码音信之间的(分段式)信噪比。在关注此(分段式)信噪比的情况下,可参考AMR-WB+ 标准之5.2.3章节 (International Standard 3GPP TS 26.290 V6.1.0 2004-12)。

[0041] 在本发明的实施例中,各别的品质测量基于加权重音信的一部分的能量并基于在使用各演算法来编码该信号部分的情况下所导入的一被估计变形(estimated distortion)而被估计的,其中第一与第二估计器系可依据一加权重音信的能量而决定被估计变形。

[0042] 在本发明之实施例中,当需量化(quantizing)音信的该部分时,系决定藉由使用于第一编码演算法中的一量化器而被引入的一被估计量化器变形,并且第一品质测量基于加权重音信的该部分的能量与被估计之量化器变形而被决定。在这样的实施例中,当需藉由用于第一编码演算法的一量化器以及一熵编码器来进行编码时,为了音信部分的一全局增益可被估计,以致音信部分可产生一给定的目标比特率,其中该被估计量化器变形基于被估计全局增益而决定。在这样的实施例中,被估计量化器变形可基于被估计增益之一能力(power)而决定。当用于第一编码演算法之量化器为一均匀标量量化器(uniform scalar quantizer)时,第一估计器可藉由使用公式 $D=G*G/12$ 而决定被估计量化器变形,其中D为被估计量化器变形,G为被估计全局增益。在第一编码演算法使用另一种量化器的例子中,量化器变形可以另一种方式而从全局增益来决定。

[0043] 发明人认可一品质测量(例如一分段式信噪比)可藉由使用上述特征之任何组合并以一合适的方式而被估计。其中当使用第一编码演算法(例如TCX演算法)来编码与解码该音信的该部分时,可得到该分段式信噪比。

[0044] 在本发明之实施例中,第一品质测量系为一分段式信噪比。分段式信噪比系藉由计算音信部分的各子部分的一被估计信噪比(其基于加权重音信之对应子部分的一能量以及被估计量化器变形)并藉由计算加权重音信的该等子部分之信噪比之一平均值而被估计,以得到加权重音信的该部分之被估计分段式信噪比。

[0045] 在本发明之实施例中,当使用调适码本(adaptive codebook)来编码音信部分时,系决定一被估计调适码本变形,其系藉由使用于第二编码演算法中之一调适码本而被引入。并且第二品质测量基于加权重音信的该部分的一能量以及被估计调适码本变形而被估计。

[0046] 在这样的实施例中,对音信的该部分的各子部分而言,调适码本可基于被转移至过去之加权重音信之子部分的一版本并藉由在一前处理阶段中决定的一音高延迟而被近似,并且可估计一调适码本增益以致能最小化加权重音信的该部分之子部分与被近似之调适码本之间的一误差,并且可基于在加权重音信的该部分之子部分与藉由调适码本增益而缩放之被近似调适码本之间的一误差的能量而决定一被估计调适码本变形。

[0047] 在本发明之实施例中,为了音信部分的各子部分所决定之被估计调适码本变形可被一固定因子减少,以将变形之一减量考虑在内,这是藉由在第二编码演算法中之一创新

码本所达到的。

[0048] 在本发明之实施例中,第二品质测量系为一分段式信噪比。分段式信噪比系藉由计算各子部分的一被估计信噪比(其基于加权重音信之对应子部分的能量以及被估计调适码本变形)并藉由计算该等子部分之信噪比之一平均值而被估计,以得到被估计分段式信噪比。

[0049] 在本发明之实施例中,调适码本可基于被转移至过去之加权重音信之部分的一版本并藉由在一前处理阶段中决定的一音高延迟而被近似,并且可估计一调适码本增益以致能最小化加权重音信的该部分与被近似之调适码本之间的一误差,并且可基于在加权重音信的该部分与藉由调适码本增益而缩放之被近似调适码本之间的能量而决定一被估计调适码本变形。如此,被估计调适码本变形可在低复杂度的情况下被决定。

[0050] 发明人认可品质测量(例如一分段式信噪比)可藉由使用上述特征之任何组合并以一合适的方式而被估计。其中当使用第二编码演算法(例如ACELP演算法)来编码与解码该音信的该部分时,可得到该分段式信噪比。

[0051] 在本发明之实施例中,一滞后机制用来比较被估计之品质测量。这可用来让使用哪一种演算法的决定更稳定。该滞后机制可依据被估计之品质测量(例如它们之间的差异)以及其他参数,例如关于先前决定的统计、时间静止画面的数量与画面中的瞬时。当考虑到这些滞后机制时,可例如参考申请专利WO 2012/110448 A1。

[0052] 在本发明之实施例中,编码一音信的一编码器包含装置10、一阶段以执行第一编码演算法以及一阶段以执行第二编码演算法。其中,该编码器依据控制器16的选择而使用第一编码演算法或第二编码演算法来编码音信的该部分。在本发明之实施例中,可编码与解码之一系统包含该编码器以及一解码器,其可接收音信部分的编码版以及用来编码音信部分之演算法的一指示,并可使用被指出之演算法来解码音信部分的编码版。

[0053] 如图1所示以及上述之一开回路模式选择演算法系在一早先申请案PCT/EP2014/051557中被描述。该演算法用以在两种模式(例如ACELP 与 TCX)之间并以一画面接画面的基准上(frame-by-frame basis)作出一选择。该选择可基于ACELP与TCX二者之分段式信噪比之一估计。带有最高被估计分段式信噪比之模式系被选择。非必要地,一滞后机制可使用来提供更强健的选择。ACELP之分段式信噪比可藉由使用调适码本变形的一近似以及创新码本变形的一近似而被估计。调适码本可藉由使用由一音高分析演算法所估计的一音高延迟而在加权重信号域中被近似。该变形可在加权重信号域中来计算并作为一最佳增益。然后,该变形可藉由一固定因子而减少,以近似于该创新码本变形。TCX之分段式信噪比可藉由使用真实TCX编码器之一简化版而被估计。输入信号可先由一改良离散余弦转换(MDCT)来转换,再藉由使用一加权重线性预测编码滤波器而被塑形。最后,该变形可藉由使用一全局增益与一全局增益估计器而在加权重MDCT域中来估计。

[0054] 结果,在早先申请案中所描述之开回路模式选择演算法系大部分时间皆提供期望中的决定,即在类语音与类瞬时信号上选择ACELP并且在类音乐与类噪音信号上选择TCX。然而,发明人认可下面事项可能会发生,就是在一些谐波音乐信号上,一些时候选择ACELP。在这样的信号上,调适码本一般由于谐波信号的高可预测性而具有一高预测增益,并能产生低变形以及较TCX更高的分段式信噪比。然而,TCX在大部分的谐波音乐信号上显为更好听,所以TCX在这些情况中应是更好的选择。

[0055] 如此,本发明建议使用输入信号的一版本而执行信噪比或分段式信噪比之估计,而作为第一品质测量,其系被滤波以减少其谐波。如此,就可得到谐波音乐信号上之一改良模式选择。

[0056] 一般而言,任何能减少谐波之合适的滤波器皆可使用。在本发明之实施例中,该滤波器为一长期预测滤波器。一长期预测滤波器之一简化例子系为 $F(z) = 1 - g \cdot z^{-T}$

[0057] 其中,滤波器参数为增益g以及音高延迟T,其系从音信来决定。

[0058] 本发明之实施例基于一长期预测滤波器,其系应用于音信并在TCX 分段式信噪比估计中之MDCT分析之前。长期预测滤波器系在MDCT分析之前减少在输入信号中的谐波振幅。结果就是在加权重MDCT域中的变形被减少、TCX之被估计分段式信噪比被增加、以及TCX在谐波音乐信号上更常被选择。

[0059] 在本发明的实施例中,长期预测滤波器的一转移函数包含一音高延迟的一整数部分以及依据该音高延迟的一分数部分的一多阀门滤波器(multi tap filter)。由于整数部分 $[z]^{T_{int}}$ 只被使用于正常取样率架构中,所以可得到有效率的实行。同时,由于分数部分使用于多阀门滤波器中,故可达到高精确度。藉由将多阀门滤波器中的分数部分考虑进来,可达到谐波能量的去除,同时在谐波附近的该等部分的能量亦被去除。

[0060] 在本发明的实施例中,长期预测滤波器系描述如下:

$$P(z) = 1 - \beta g B(z, T_{fr}) z^{-T_{int}}$$

[0062] 其中 T_{int} 与 T_{fr} 分别为一音高延迟的整数与分数部分, g 为一增益, β 为一权重, $B(z, T_{fr})$ 为一有限脉冲响应(FIR) 低通滤波器,其系数依据该音高延迟的分数部分。上述长期预测滤波器之实施例的叙述会在下面提出。

[0063] 音高延迟与增益可在一画面接画面基准上来估计。

[0064] 预测滤波器可基于至少一谐波性测量(harmonicity measure) (例如正规化(normalized) 关联性或预测增益) 及/或至少一时间结构测量(例如时间平整性(flatness) 测量或能量改变)的一组合而失能(即增益等于0)。

[0065] 滤波器可在一画面接画面的基准上应用于输入音信。假使滤波器参数从一画面到下一画面的过程中改变,则在两画面之间的边界上会产生一不连续。在实施例中,装置还包含一单元以移除由滤波器造成在音信中的不连续。为移除可能的不连续,可使用任何技术,例如可与那些描述于专利 US5012517、EP0732687A2、US5999899A 或 US7353168B2 内相比之技术。移除可能不连续之另一技术系在下面描述。

[0066] 在详细描述图3所示之第一估计器12与第二估计器14之一实施例之前,系描述如图2所示的一编码器20之一实施例。

[0067] 编码器20包含第一估计器12、第二估计器14、控制器16、一前处理单元22、一切换器24、一第一编码阶段26以执行一TCX演算法、一第二编码阶段28以执行一ACELP演算法、以及一输出接口30。前处理单元 22 可为一常用的语音/音信统合编码(Unified speech and audio coding, USAC) 编码器的一部分并且可输出线性预测编码系数、加权重线性预测编码系数、加权重音信以及一组音高延迟。需注意者,这些参数皆使用于两个编码演算法,即TCX演算法与ACELP演算法。如此,这些参数就不必为了开回路模式决定而被额外的计算出来。使用在开回路模式决定中之已经被计算出来的参数的好处就是能降低复杂度。

[0068] 如图2所示,装置包含谐波减少滤波器2。装置还包含一非必要失能单元4以基于至

少一谐波性测量 (harmonicity measure) (例如正规化 (normalized) 关联性或预测增益) 及/或至少一时间结构测量 (例如时间平整性 (flatness) 测量或能量改变) 的一组合而失能该谐波减少滤波器2。装置包含一非必要的不连续移除单元6以移除音信的滤波版本之不连续。此外,装置非必要地包含一单元8以估计谐波减少滤波器2的滤波器参数。在图2中,这些组件(2、4、6、8)系显示而作为第一估计器12的一部分。不用说,这些组件可实现于第一估计器之外面或与第一估计器分离,并且能提供音信的滤波版本给第一估计器。

[0069] 一输入音信40系提供于一输入线。输入音信40系应用于第一估计器 12、前处理单元22以及两编码阶段26、28。在第一编码器12中,输入音信40系应用于滤波器2,并且输入音信的滤波版本用于估计第一品质测量。在滤波器被失能单元4所失能的情况下,输入音信40被使用来估计第一品质测量,而非使用输入音信的滤波版本。前处理单元22以一习知方式处理输入音信以得到线性预测编码系数以及加权重线性预测编码系数42并滤波带有加权重线性预测编码 (LPC) 系数42的音信40,以得到加权重音信44。前处理单元22系输出加权重LPC系数42、加权重音信44 以及一组音高延迟48。就如习知技艺者所知道的,加权重LPC系数42与加权重音信44可被分段为多个画面或子画面。该分段可藉由以一合适的方式来窗口化音信而得到。

[0070] 在其他实施例中,可提供一前处理器,其基于音信的滤波版而产生加权重LPC系数与一加权重音信。然后,基于音信的滤波版之加权重LPC 系数与加权重音信系应用于第一估计器以估计第一品质测量,而非使用加权重LPC系数42与加权重音信44。

[0071] 在本发明之实施例中,可使用被量化之LPC系数或被量化之加权重 LPC系数。如此,下面叙述应被了解,就是LPC系数亦涵盖被量化之LPC 系数,并且加权重LPC系数亦涵盖加权重被量化之LPC系数。就这一点而言,值得注意的是,语音/音信统合编码 (USAC) 之TCX 演算法使用被量化加权重LPC系数以塑形改良离散余弦转换 (MCDT) 频谱。

[0072] 第一估计器12接收音信40、加权重LPC系数42与加权重音信44,并基于此而估计第一品质测量46,且输出第一品质测量至控制器16。第二估计器16接收加权重音信44与该组音高延迟48,并基于此而估计第二品质测量50,且输出第二品质测量50至控制器16。就如习知技艺者所知,加权重LPC系数42、加权重音信44与该组音高延迟48已经在前一模块 (亦即前处理单元22) 被计算出来,所以现在可不用其他花费而被使用。

[0073] 控制器基于所接收之品质测量的一比较而选择TCX演算法与ACELP 演算法的其中之一。就如上所指出的,控制器可使用一滞后机制来决定使用哪一演算法。第一编码阶段26 或第二编码阶段28的选择由图2之切换器24所示意,切换器24系藉由控制器16所输出之一控制信号52而控制。控制信号52系指出第一编码阶段26或第二编码阶段28将被使用。基于控制信号52,由图2之箭号54所表示之所需要的信号 (其系至少包含LPC 系睥、加权重LPC系数、音信、加权重音信与该组音高延迟) 系被应用至第一编码阶段26或第二编码阶段28。被选择的编码阶段系应用相关连的编码演算法并输出被编码的表述56或58至输出接口30。输出接口30 可输出一被编码音信60,其可包含被编码的表述56或58、LPC系数或加权重LPC系数、被选择的编码演算法之参数以及关于被选择的编码演算法之信息。

[0074] 图3系描述特定的实施例,其估计第一与第二品质测量,其中第一与第二品质测量系为在加权重信号域中的分段式信噪比。图3以流程图的形式来显示第一估计器12、第二估计器14与其功能,并显示各别的估计之步骤。

[0075] TCX分段式信噪比的估计

[0076] 第一(TCX)估计器接收音信40(输入信号)、加权重LPC系数42 与加权重音信44作为输入。音信40的滤波版系在步骤98中产生。在音信40的滤波版中,谐波系被减少或抑制。

[0077] 音信40可被分析以决定至少一谐波性测量(例如正规化关联性或预测增益)及/或至少一时间结构测量(例如时间平整性测量或能量改变)。基于这些测量的其中之一或这些测量的一组合,滤波器2连同滤波98可被失能。假使滤波98失能,则使用音信40来进行第一品质测量之估计,而非使用其滤波版本。

[0078] 在本发明之实施例中,移除不连续(未显示于图3)之一步骤可跟在滤波98之后,以移除音信中的不连续,其由滤波98所导致。

[0079] 在步骤100,音信40的滤波版系被窗口化。窗口化可由一个10毫秒之低重迭正弦窗口 (low-overlap sine window)来进行。当该过去画面 (past-frame) 为ACELP时,遮挡尺寸 (block-size) 可增加5毫秒、窗口的左边可为四方形、并且ACELP合成滤波器之窗口化零脉冲反应 (windowed zero impulse response) 可从窗口化输入信号中被移除。这跟在TCX演算法中所作的相像。音信40的滤波版之一画面(其系代表音信的一部分)系从步骤100输出。

[0080] 在步骤102中,窗口化音信,亦即产生的画面,系被一MDCT而转换。在步骤104中,频谱塑形系藉由塑形带有加权重LPC系数之MDCT 频谱而进行。

[0081] 在步骤106中,当用一熵编码器(例如一算术编码器)进行编码时,一全局增益G系被估计,以致用增益G所量化之加权重频谱可产生一给定目标R。由于一增益为了整个画面而决定,故使用全局增益这个词。

[0082] 以下要解释全局增益估计之一实现的例子。需注意者,该全局增益估计系合适于特定的实施例,即TCX编码演算法使用带有一算术编码器之一标量量化器 (scalar quantizer)。这样带有一算术编码器之一缩放量化器系在MPEG USAC标准中被假定。

[0083] 起始化

[0084] 首先,用于增益估计之变量系藉由下面来起始化:

[0085] 1. Set $en[i] = 9.0 + 10.0 * \log_{10}(c[4*i+0] + c[4*i+1] + c[4*i+2] + c[4*i+3])$, 其中, $0 \leq i < L/4$, $c[]$ 系为要量化之系数的向量, L 为 $c[]$ 的长度。

[0086] 2. Set $fac = 128$, $offset = fac$ and $target = \text{any value (e.g. 1000)}$

[0087] 迭代

[0088] 然后,下列的操作方块执行NITER次(例如NITER=10)。

[0089] 1. $fac = fac/2$

[0090] 2. $offset = offset - fac$

[0091] 3. $ener = 0$

[0092] 4. for every i where $0 \leq i < L/4$ do the following:

[0093] if $en[i] - offset > 3.0$, then $ener = ener + en[i] - offset$

[0094] 5. if $ener > target$, then $offset = offset + fac$

[0095] 上述迭代的结果为补偿值 (offset value)。在迭代之后,全局增益估计为 $G = 10^{\lceil (offset/20) \rceil}$ 。

[0096] 估计全局增益的方式可依据所使用的量化器与熵编码器而变化。在MPEG USAC标准中,带有一算术编码器之一标量量化器系被假定。其他 TCX方式可使用一个不同的量化

器,并且习知技术者系知道如何估计对应这样不同的量化器之全局增益。举例来说,AMR-WB+系假定使用一RE8 格状量化器(lattice quantizer)。对于这样一个量化器,全局增益的估计可如同在3GPP TS 26.290 V6.1.0 2004-12之第34页的章节5.3.5.7所描述的来进行,其中系假定一固定目标比特率。

[0097] 在步骤106之估计全局增益之后,步骤108系进行变形估计。特定地说,量化器基于被估计之全局增益而被近似。在本实施例中,其系假定使用一均匀标量量化器。如此,量化器变形由简单公式 $D=G*G/12$ 来决定,其中D代表被决定之量化器变形,G代表被估计之全局增益。这对应至一均匀标量量化器变形之高比例近似 (high-rate approximation)。

[0098] 基于被决定之量化器变形,分段式信噪比计算系在步骤110中来进行。该画面的各子画面之信噪比系被计算并作为加权重音信能量与变形D之比例,变形D系假定为在该等子画面中为定值。举例而言,该画面系分为连续的四个子画面(参照图4)。然后,分段式信噪比系为四个子画面之信噪比之平均值并且可以dB来表示。

[0099] 该方式可准许第一分段式信噪比之估计,当使用TCX演算法而实际地编码与解码该目标画面时,可得到第一分段式信噪比,然而却不需要实际地编码与解码音信,因此可大幅降低复杂度并减少计算时间。

[0100] ACELP分段式信噪比之估计

[0101] 第二估计器14接收加权重音信44与该组音高延迟48,其系已经在前处理单元22中被计算出来。

[0102] 如步骤112所示,在各子画面中,调适码本系藉由简单地使用加权重音信与音高延迟T而被近似。调适码本系藉由下面而被近似:

[0103] $x_w(n-T), n=0, \dots, N$

[0104] 其中 x_w 为加权重音信,T为对应子画面的音高延迟,N为子画面长度。据此,调适码本系藉由使用被T转移到过去之子画面的一版本而被近似。因此,在本发明之实施例中,调适码本以一非常简单的方式而被近似。

[0105] 在步骤114中,系决定各子画面之一调适码本增益。特定来说,在各子画面中,码本增益G系被估计,以致其最小化在加权重音信与被近似之调适码本之间之误差。这可藉由简单地比较各取样之两种信号之间的差异以及找到使这些差异之总和最小化之增益而达到。

[0106] 在步骤116中,系决定各子画面之调适码本变形。在各子画面中,由调适码本所引入之变形D就是在加权重音信与由增益G所缩放之被近似之调适码本之间的误差的能量。

[0107] 在步骤116中所决定的变形可在一非必要的步骤118中来调整,以考虑到创新的码本。用于ACELP演算法中之创新码本的变形可被估计为一定值。在本发明已描述的实施例中,其系假定创新码本藉由一固定因子而减少变形D。如此,在步骤116所得到的各子画面之变形可在步骤118中乘以一固定因子,例如为0到1次方的固定因子,例如为0.055。

[0108] 步骤120系进行分段式信噪比之计算。在各子画面中,信噪比系计算而作为加权重音信能量与变形D之比值。然后,分段式信噪比系为四个子画面之信噪比之平均值并可以dB表示。

[0109] 该方式系准许第二信噪比之估计,当使用ACELP演算法而实际地编码与解码该目标画面时,可得到第二信噪比,然而却不需要实际地编码与解码音信,因此可大幅降低复杂度并减少计算时间。

[0110] 第一与第二估计器12、14系输出被估计之分段式信噪比46、50至控制器16,并且控制器16基于被估计之分段式信噪比46、50而决定哪一演算法要用于音信之相关部分。控制器可非必要地使用一滞后机制,以使该决定更稳定。举例而言,在闭回路决定中之滞后机制可被使用,但带着些许不同的调音参数。这样的滞后机制可计算出一值dsnr,其依据被估计之分段式信噪比(例如在它们之间的差异)以及其他参数,例如关于先前决定之统计、时间静止画面的数量与画面中的瞬时。

[0111] 在没有滞后机制时,控制器可选择具有较高被估计之信噪比的编码演算法,亦即,假使第二被估计信噪比高于第一被估计信噪比,则选择 ACELP,假使第一被估计信噪比高于第二被估计信噪比,则选择TCX。在具有滞后机制的情况下,控制器可依据下面的决定规则来选择编码演算法,其中acelp_snr系为第二被估计信噪比而tcx_snr为第一被估计信噪比:

[0112] if acelp_snr+dsnr>tcx_snr then select ACELP, otherwise select TCX.

[0113] 为了减少谐波振幅的滤波器之参数的确定

[0114] 以下系描述为减少谐波振幅而确定滤波器参数之一实施例。滤波器参数可在编码器侧被估计,就如在单元8中。

[0115] 音高估计

[0116] 每一画面(画面大小例如20毫秒)的一音高延迟系被估计。这是在三个步骤中来进行的,以减少复杂度并提升估计精确度。

[0117] (a) 音高延迟的整数部分的第一估计

[0118] 产生一平滑音高进展曲线(smooth pitch evolution contour)的一音高分析演算法系被使用(例如在Rec.ITU-T G.718, sec.6.6中所描述之开回路音高分析)。该分析一般系在一子画面基准(子画面大小例如10毫秒)上来进行,并产生每一子画面的一音高延迟。需注意者,这些音高延迟估计并没有任何分数部分且一般系在一缩减取样(downsampling)信号(取样率例如6400Hz)上来估计。使用的信号可为任何音信,例如为LPC加权重音信,就如在Rec.ITU-T G.718, sec.6.5中所描述的。

[0119] (b) 音高的整数部分Tint之提炼

[0120] 最后的音高的整数部分系在一音信x[n]上并以核心编码器取样率(core encoder sampling rate)而被估计,核心编码器取样率一般系高于用于(a)之缩减音信的取样率(例如12.8kHz、16kHz、32kHz)。该信号x[n]可为任何音信,例如LPC加权重音信。

[0121] 然后,音高延迟的整数部分Tint系为能最大化自相关函数(autocorrelation function)之延迟。

$$[0122] C(d) = \sum_{n=0}^N x[n]x[n-d]$$

[0123] 其中d系在(a)中所估计的一音高T之周围。

[0124] $T - \delta_1 \leq d \leq T + \delta_2$

[0125] (c) 音高延迟的分数部分Tfr之估计

[0126] 分数部分Tfr系藉由插入在步骤(b)中所计算出来之自相关函数C(d)并藉由选择能最大化被插入之自相关函数的分数音高而被找到。该插入可藉由使用在例如Rec.ITU-T

G.718, sec.6.6.7中所描述之一低通有限脉冲响应(FIR)滤波器而进行。

[0127] 增益估计与量化

[0128] 增益一般系在输入音信上并以核心编码器取样率来估计,但其亦可为任何音信,例如LPC加权重音信。该信号系标注为y[n]并可与x[n]相同或不同。

[0129] y[n]的预测yP[n]系藉由使下面滤波器滤波y[n]而先被找到。

$$P(z) = B(z, T_{fr}) z^{-T_{int}}$$

[0131] 其中T_{int}为音高的整数部分(在步骤(b)中被估计),B(z, T_{fr})为一低通FIR滤波器,其系数依据音高T_{fr}的分数部分(在步骤(c)中被估计)。

[0132] 以下为当音高分辨率为1/4时之B(z)之一例子:

$$[0133] T_{fr} = \frac{0}{4} \quad B(z) = 0.0000z^{-2} + 0.2325z^{-1} + 0.5349z^0 + 0.2325z^1$$

$$[0134] T_{fr} = \frac{1}{4} \quad B(z) = 0.0152z^{-2} + 0.3400z^{-1} + 0.5094z^0 + 0.1353z^1$$

$$[0135] T_{fr} = \frac{2}{4} \quad B(z) = 0.0609z^{-2} + 0.4391z^{-1} + 0.4391z^0 + 0.0609z^1$$

$$[0136] T_{fr} = \frac{3}{4} \quad B(z) = 0.1353z^{-2} + 0.5094z^{-1} + 0.3400z^0 + 0.0152z^1$$

[0137] 则,增益g系计算如下:

$$[0138] g = \frac{\sum_{n=0}^{N-1} y[n] y_p[n]}{\sum_{n=0}^{N-1} y_p[n] y_p[n]}$$

[0139] 最后,增益g系藉由使用例如均匀量化而量化于2位上。

[0140] β用来控制滤波器之强度。当β系等于1,其系产生全部的功效;当β等于0时,其系失能滤波器。如此,在本发明之实施例中,滤波器可藉由将β设为0而失能。在本发明之实施例中,假使滤波器被致能,则β可被设为介于0.5到0.75之间的值。在本发明之实施例中,假使滤波器被致能,则β可设为0.625。B(z, T_{fr})的一个例子系提供于上。B(z, T_{fr})的阶数与系数亦可依据比特率与输出取样率。一个不同的频率响应可被设计并调整以对应比特率与输出取样率的各组合。

[0141] 失能滤波器

[0142] 滤波器可基于至少一谐波性测量及/或至少一时间结构测量的一组合而失能。这样的测量之例子系如下所述。

[0143] (i) 谐波性测量如同在步骤(b)所估计的整数音高延迟之正规化关联性。

$$[0144] \text{norm.corr.} = \frac{\sum_{n=0}^{N-1} x[n] x[n - T_{int}]}{\sqrt{\sum_{n=0}^{N-1} x[n] x[n]} \sqrt{\sum_{n=0}^{N-1} x[n - T_{int}] x[n - T_{int}]}}$$

[0145] 假使输入信号藉由整数音高延迟而完美地可预测,则正规化关联性为1;假使并非可预测,则正规化关联性为0。再者,一高值(逼近1)可指出一谐波信号。为达到更强健的决定,过去画面的正规化关联性亦可使用于该决定,例如:

[0146] If (norm.corr(curr.) * norm.corr(prev.)) > 0.25, then the filter is not disabled

[0147] (ii) 例如在能量取样基准上被计算之时间结构测量亦可藉由为了瞬时检测(例如时间平整性测量、能量改变)之一瞬时检测器而被使用,例如:

[0148] if (temporal flatness measure>3.5 or energy change>3.5) then the filter is disabled.

[0149] 更多关于至少一谐波性测量之确定的细节系于下描述。

[0150] 谐波性的测量系例如藉由音信之一正规化关联性或其一前调整 (pre-modified) 版本并在音高延迟或在音高延迟附近而计算。音高延迟可甚至在包含一第一阶段与一第二阶段之阶段中而确定,其中,在第一阶段中,音高延迟之一初步估计系在一第一取样率之一缩减取样域中确定,而在第二阶段中,该音高延迟之初步估计系在一第二取样率上被精炼,其高于第一取样率。音高延迟例如系藉由使用自相关而确定。该至少一时间结构测量系例如在一时间范围内确定,该时间范围依据音高信息。该时间范围之一时间上的过去航向 (past-heading) 端系例如依据音高信息而置放。时间范围之时间上的过去航向端可被置放,以致时间范围之时间上的过去航向端系藉由带着音高信息之一增加之单调的增加之一时间量而调换到过去方向。时间范围之时间上的未来航向端可在一时间的候选范围(其系从该时间范围之时间上的过去航向端或从对时间结构测量之确定有较高影响的范围之时间上的过去航向端到一现行画面之一时间上的未来航向端)并依据音信之时间结构而被定位。在时间候选范围内之最大与最小能量取样之间的振幅或比率可被使用于这目的。举例而言,该至少一时间结构测量可测量在时间范围内的音信之一平均或最大能量变化,并且假使该至少一时间结构测量系小于一默认第一阀值并且谐波性测量对于一现行画面及/或一先前画面系为在一第二阀值之上,则失能之一条件可被满足。假使谐波性测量对于一现行画面系在一第三阀值之上并且谐波性测量对于一现行画面及/或一先前画面系在一第四阀值(其系随着音高延迟之增加而减少)之上,则该条件亦可被满足。

[0151] 现在要给予的是确定该等测量之一具体实施例之步骤接步骤的描述。

[0152] 步骤1:瞬时检测与时间测量

[0153] 输入信号 $s_{HP}(n)$ 系输入至时间域瞬时检测器。输入信号 $s_{HP}(n)$ 系经过高通滤波。瞬时检测之高通 (HP) 滤波的转移函数系如下所示:

[0154] $H_{TD}(z) = 0.375 - 0.5z^{-1} + 0.125z^{-2}$ (1)

[0155] 由瞬时检测之HP滤波器所滤波之信号系标注为 $s_{TD}(n)$ 。高通滤波之信号 $s_{TD}(n)$ 系被分段为8个连续且相同长度的分段。每一分段之高通滤波信号 $s_{TD}(n)$ 的能量系计算如下:

$$[0156] E_{TD}(i) = \sum_{n=0}^{L_{segment}-1} (s_{TD}(iL_{segment} + n))^2, \quad i = 0, \dots, 7 \quad (2)$$

[0157] 一累积能量系如下所计算:

[0158] $E_{Acc} = \max(E_{TD}(i-1), 0.8125E_{Acc})$ (3)

[0159] 假使一分段 $E_{TD}(i)$ 的能量超过累积能量,则藉由一固定因子 $attackRatio = 8.5$ 检测到一攻击,并且攻击指标 (attackIndex) 系设为 i :

[0160] $E_{TD}(i) > attackRatio \cdot E_{Acc}$ (4)

[0161] 假使没有攻击基于上述标准而被检测到,但有检测到分段 i 具有一大幅能量增加,则攻击指标系设为 i 且未指出有攻击的出现。基本上,攻击指标系设为在一画面中上个攻击

的位置,且带有一些额外的限制。

[0162] 每一分段的能量改变系计算如下:

$$[0163] E_{chng}(i) = \begin{cases} \frac{E_{TD}(i)}{E_{TD}(i-1)}, & E_{TD}(i) > E_{TD}(i-1) \\ \frac{E_{TD}(i-1)}{E_{TD}(i)}, & E_{TD}(i-1) > E_{TD}(i) \end{cases} \quad (5)$$

[0164] 时间平整性测量系计算如下:

$$[0165] TFM(N_{past}) = \frac{1}{8 + N_{past}} \sum_{i=-N_{past}}^7 E_{chng}(i) \quad (6)$$

[0166] 最大能量改变系计算如下:

$$[0167] MEC(N_{past}, N_{new}) = \max(E_{chng}(-N_{past}), E_{chng}(-N_{past}+1), \dots, E_{chng}(N_{new}-1)) \quad (7)$$

[0168] 假使 $E_{chng}(i)$ 或 $E_{TD}(i)$ 的指标为负,则其指出从带有与现行画面相关之分段指标之先前分段之一值。

[0169] N_{past} 系为从过去画面之分段的数量。假使时间平整性测量为了使用于 ACELP/TCX 中而计算,则 N_{past} 等于0。假使时间平整性测量为了TCX LTP 决定而计算,则 N_{past} 等于:

$$[0170] N_{past} = 1 + \min\left(8, \left\lceil 8 \frac{pitch}{L} + 0.5 \right\rceil\right) \quad (8)$$

[0171] N_{new} 系为从现行画面之分段的数量。对于非瞬时画面,其系等于8。对于瞬时画面,具有最大与最小能量之分段的位置系如下:

$$[0172] i_{\max} = \arg \max_{i \in \{-N_{past}, \dots, 7\}} E_{TD}(i) \quad (9)$$

$$[0173] i_{\min} = \arg \min_{i \in \{-N_{past}, \dots, 7\}} E_{TD}(i) \quad (10)$$

[0174] If $E_{TD}(i_{\min}) > 0.375 E_{TD}(i_{\max})$ then N_{new} is set to $i_{\max} - 3$, otherwise N_{new} is set to 8.

[0175] 步骤2:转换方块长度切换

[0176] TCX之重迭长度与转换方块长度依据一瞬时之存在以及其位置。

[0177] 表1:基于瞬时位置之重迭与转换长度的编码

[0178]	攻击指标 Attack-Index	带有下一画面之第一窗口之重迭 Overlap with the first window of the following frame	短/长转换决定 (二元编码) Short/Long Transform decision (binary coded) 0 - Long, 1 - Short	重迭宽度之二元编码 Binary code for the overlap width	重迭编码 Overlap code
	none	ALDO	0	0	00
	-2	FULL	1	0	10
	-1	FULL	1	0	10
[0179]	0	FULL	1	0	10
	1	FULL	1	0	10
	2	MINIMAL	1	10	110
	3	HALF	1	11	111
	4	HALF	1	11	111
	5	MINIMAL	1	10	110
	6	MINIMAL	0	10	010
	7	HALF	0	11	011

[0180] 如上所描述之瞬时检测器基本上系回报带有限制之上个攻击之指标,假使有多个瞬时,则MINIMAL重迭是更好于HALF重迭,HALF重迭系更好于FULL重迭。假使在位置2或6的攻击不够强,则选择HALF重迭,而非MINIMAL重迭。

[0181] 步骤3:音高估计

[0182] 每一画面的一音高延迟(整数部分加上分数部分)系被估计(画面大小例如20毫秒),就如上述3个步骤(a)到(c)所述,以减少复杂度并提升估计精确度。

[0183] 步骤4:决定定位

[0184] 假使输入音信未包含任何谐波内容,或者假使技术性的预测可将变形导入时间结构(例如一短瞬时的重复),则采取让滤波器失能之一决定。

[0185] 该决定基于多个参数而作出,参数例如是在整数音高延迟之正规化关联性以及时间结构测量。

[0186] 在整数音高之正规化关联性norm_corr系被估计,就如上所述。假使输入信号系可藉由整数音高延迟而完美的预测,则正规化关联性为1,若无法可预测,则正规化关联性为0。然后,一高值(逼近于1)可指出一谐波信号。对于更强健的决定,除了对于现行画面之正规化关联性 (norm_corr(curr)) 可被使用之外,过去画面之正规化关联性 (norm_corr

(prev)) 亦可使用于该决定,例如:

[0187] If (norm_corr (curr) * norm_corr (prev)) > 0.25

[0188] 或

[0189] If max (norm_corr (curr) , norm_corr (prev)) > 0.5,

[0190] 则现行画面包含一些谐波内容。

[0191] 时间结构测量可藉由一瞬时检测器而计算(例如时间平整性测量(方程式(6))以及最大能量改变方程式(7)),以避免激活在包含一强瞬时或大时间改变之一信号上的滤波器。时间特征系在包含现行画面(N_{new} 分段)与直到音高之过去画面(N_{past} 分段)之信号上而被计算。对于像慢慢衰退之瞬时的步骤,由于在由LTP滤波所导入之频谱之非谐波部分中之变形可藉由强壮且长持续的瞬时(例如碎音钹)之屏蔽(masking)而被抑制,所以全部或一些的特征仅可被计算到瞬时的位置($i_{max}-3$)。

[0192] 对于低音高信号之脉冲串可藉由一暂检测器而被检测而作为一瞬时。对于具有低音高之信号,从瞬时检测器来之特征可被忽略,并且反而有对应依据音高之正规关联性之额外的阀值,就如:

[0193] If norm_corr <= 1.2-T_{int}/L, then disable the filter.

[0194] 决定的一个例如如下所述,其中,b1为某一比特率,例如48kbps, TCX_20系指出画面藉由使用信号长方块而编码,TCX_10系指出画面系藉由使用2、3、4或更多的短方块而编码,TCX_20/TCX_10之决定基于如上所述之瞬时检测器之输出。tempFlatness系为在方程式(6)中所定义之时间平整性测量。maxEnergyChange系为在方程式(7)中所定义之最大能量改变。条件式norm_corr (curr) > 1.2-T_{int}/L亦可被写为 (1.2-norm_corr (curr)) * L < T_{int}。

[0195]

```
enableLTP =
(bitrate < b1 && tcxMode == TCX_20 && (norm_corr(curr) * norm_corr(prev)) > 0.25 && tempFlatness < 3.5) ||
(bitrate >= b1 && tcxMode == TCX_10 && max(norm_corr(curr),norm_corr(prev)) > 0.5 && maxEnergyChange < 3.5) ||
(bitrate >= b1 && norm_corr(curr) > 0.44 && norm_corr(curr) > 1.2-Tint/L) ||
(bitrate >= b1 && tcxMode == TCX_20 && norm_corr(curr) > 0.44 &&
  (tempFlatness < 6.0 || (tempFlatness < 7.0 && maxEnergyChange < 22.0)));
  (bitrate >= b1 && tcxMode == TCX_20 && norm_corr > 0.44 &&
```

[0196] 从上述例子可明显地看见,一瞬时之检测系影响哪一长期预测之决定机制会被使用以及信号的哪一部分会被使用在用于决定中之测量,并且不是它直接触发长期预测滤波器之失能。

[0197] 用于转换长度决定之时间测量可完全与用于LTP滤波器决定之时间测量不同,或者它们可相互重迭或完全相同但被计算于不同范围。对于低音高信号,假使达到依据音高延迟之正规化关联性之阀值,则瞬时之检测可完全被忽略。

[0198] 移除可能不连续之技术

[0199] 现在系描述藉由以画面接画面的方式应用一线性滤波器H(z)而移除不连续之一可能技术。线性滤波器可为已描述之LTP滤波器。线性滤波器可为一FIR滤波器或一无限脉冲反应(infinite impulse response, IIR)滤波器。所提的方法并非用过去画面的滤波器

参数来滤波现行画面的一部分,因而避免已知方法之可能的问题。所提的方法使用一LPC滤波器以移除不连续。该LPC滤波器系在音信上(由一线性时间不变滤波器 $H(z)$ 来滤波或没有滤波)被估计,并因此成为音信(由 $H(z)$ 所滤波或没有滤波)之频谱形状之一良好模型。然后,LPC滤波器系被使用以致音信的频谱形状屏蔽不连续。

[0200] LPC滤波器可用不同方式来估计。它可例如使用音信(现行及/或过去画面)与 Levinson-Durbin演算法而被估计。它亦可藉由使用 Levinson-Durbin演算法而在过去滤波画面信号上被计算。

[0201] 假使 $H(z)$ 被使用于一音信编解码器并且该音信编解码器已使用一LPC滤波器(量化或没有量化),以例如塑形在一运用转换(transform-based)音信编解码中之量化噪音,则该LPC滤波器可被直接使用于平滑化不连续,且不需要额外的复杂度来估计一个新的LPC滤波器。

[0202] 以下系描述FIR滤波器例子以及IIR滤波器例子中的现行画面之处理。过去画面系假定已经被处理。

[0203] FIR滤波器例子:

[0204] 1、用现行画面的滤波器参数来滤波现行画面,以产生一被滤波现行画面。

[0205] 2、考虑具有M次之LPC滤波器(量化与否),且在音信(滤波与否)上被估计。

[0206] 3、过去画面的前M个取样用滤波器 $H(z)$ 与现行画面之系数来滤波,以产生被滤波信号的第一部分。

[0207] 4、被滤波之过去画面的前M个取样系从被滤波信号的第一部分中减去,以产生被滤波信号的第二部分。

[0208] 5、LPC滤波器之一零脉冲反应(ZIR)系藉由用LPC滤波器以及等于被滤波信号之第二部分之起始状态来滤波零取样之一画面而产生。

[0209] 6、ZIR可非必要地被窗口化,以致其振幅更快的达到零。

[0210] 7、ZIR的一起始部分系从被滤波现行画面之一对应起始部分中减去。

[0211] IIR滤波器例子:

[0212] 1、考虑具有M次之一LPC滤波器(量化与否),并在音信(滤波与否)上被估计。

[0213] 2、过去画面的前M个取样用滤波器 $H(z)$ 与现行画面之系数来滤波,以产生被滤波信号的第一部分。

[0214] 3、被滤波之过去画面的前M个取样系从被滤波信号的第一部分中减去,以产生被滤波信号的第二部分。

[0215] 4、LPC滤波器之一零脉冲反应(ZIR)系藉由用LPC滤波器以及等于被滤波信号之第二部分之起始状态来滤波零取样之一画面而产生。

[0216] 5、ZIR可非必要地被窗口化,以致其振幅更快的达到零。

[0217] 6、现行画面之一起始部分以取样接取样的方式并由现行画面之第一取样开始而被处理。

[0218] 7、取样用滤波器 $H(z)$ 与现行画面参数来滤波,以产生一第一被滤波取样。

[0219] 8、ZIR的对应取样系从第一被滤波取样中减去,以产生被滤波之现行画面之对应取样。

[0220] 9、移动至下一取样。

[0221] 10、重复上述9到12,直到现行画面之起始部分之前一个取样被处理。

[0222] 11、用现行画面的滤波器参数来滤波现行画面之剩余取样。

[0223] 据此,本发明之实施例可让估计分段式信噪比与一合适编码演算法的选择变得更简单及精确。特别地,本发明之实施例可允许一合适编码演算法之一开回路选择,其中,在音信具有谐波的情况下,编码演算法之不合适的选择可被避免。

[0224] 在上述实施例中,藉由计算各子画面所估计之信噪比之一平均而估计分段式信噪比。在另一实施例中,在不用将画面分为子画面的情况下,可估计一整个画面之信噪比。

[0225] 由于闭回路选择所需要的许多步骤都可省略,因此与闭回路选择相比,本发明的实施例可大幅减少计算时间。

[0226] 据此,藉由创新方法,可大幅节省许多步骤及其计算时间,同时合适编码演算法的选择仍可维持良好的效能。

[0227] 虽然一些方面已描述于装置的叙述中,但清楚地,这些方面亦可代表对应之方法的叙述,其中一功能方块或装置系对应一方法步骤或一方法步骤的特征。类似地,在方法步骤中所描述的方面亦可代表对应功能方块或一对应装置之项目或特征之叙述。

[0228] 这里所描述的装置的实施例以及其特征可藉由一计算机、至少一处理器、至少一微处理器、现场可编程门阵列 (FPGA)、专用集成电路 (ASIC)、类似装置或上述之任一组合来实现,上述组件可配置或编程以提供所叙述之功能性。

[0229] 一些或全部的方法步骤可由(或使用)一硬件装置,例如一微处理器、一可编程计算机或一电子电路来执行。在一些实施例中,至少一最重要的方法步骤可由这样的装置来执行。

[0230] 依据某些实施需求,本发明的实施例可以硬件或软件来实施。实施态样可使用一非瞬时储存媒介来实行,例如一数字储存媒介,如一软盘、一 DVD、一蓝光光盘、一CD、一只读存储器 (ROM)、一可编程只读存储器 (PROM)、一可擦除可编程只读存储器 (EPROM)、一电子可擦除可编程只读存储器 (EEPROM) 或一闪存,其具有电子可读控制信号储存于上并与一可编程计算机系统相互合作(或能合作),以致执行各方法。因此,数字储存媒介可为计算机可读。

[0231] 本发明之实施例包含一数据载体,其具有电子可读控制信号,其系能举一可编程计算机系统合作,以致本发明之方法的其中之一可被执行。

[0232] 一般而言,本发明之实施例可实施为带有一程序代码之一计算机程序产品,当计算机程序产生执行于一计算机上时,程序代码系能执行该等方法之一。程序代码可例如储存于一机械可读载体。

[0233] 其他实施例包含可执行本发明之方法的其中之一的计算机程序,其系储存于一机械可读载体。

[0234] 换言之,本发明方法之一实施例系为具有一程序代码之一计算机程序,当计算机程序执行于一计算机上时,可执行该等方法的其中之一。

[0235] 本发明方法之另一实施例系为一数据载体(或一数字储存媒介、或一计算机可读媒介),其包含,即记录于其上,可执行本发明之方法的其中之一之计算机程序。数据载体、数字储存媒介或记录媒介系为实体及/或非瞬时的。

[0236] 本发明方法之另一实施例系为一数据流或一信号串以代表执行本发明的其中一

方法之计算机程序。数据流或信号串可例如经由一数据通讯连结(例如因特网)而转移。

[0237] 另一实施例包含一处理手段,例如一计算机或一可编程逻辑设备,其被配置或被编程而执行本发明的其中一方法。

[0238] 另一实施例包含一计算机,其已安装可执行本发明的其中一方法之计算机程序。

[0239] 本发明另一实施例包含一装置或一系统,其系能转移(例如电子地或光学地)能执行本发明其中一方法之一计算机程序至一接收器。接收器可例如为一计算机、一行动装置、一内存装置或类似装置。该装置或系统可例如包含一文件服务器以将计算机程序转移至接收器。

[0240] 在一些实施例中,一可编程逻辑设备(例如一现场可编程门阵列)可被使用来执行本发明之方法之些或全部的功能性。在一些实施例中,一现场可编程门阵列可与一微处理器合作以执行本发明的其中一方法。一般而言,该等方法系较佳为藉由任何硬件装置来执行。

[0241] 以上所述仅为举例性,而非为限制性者。任何未脱离本发明之精神与范畴,而对其进行之等效修改或变更,均应包含于后附之申请专利范围内。

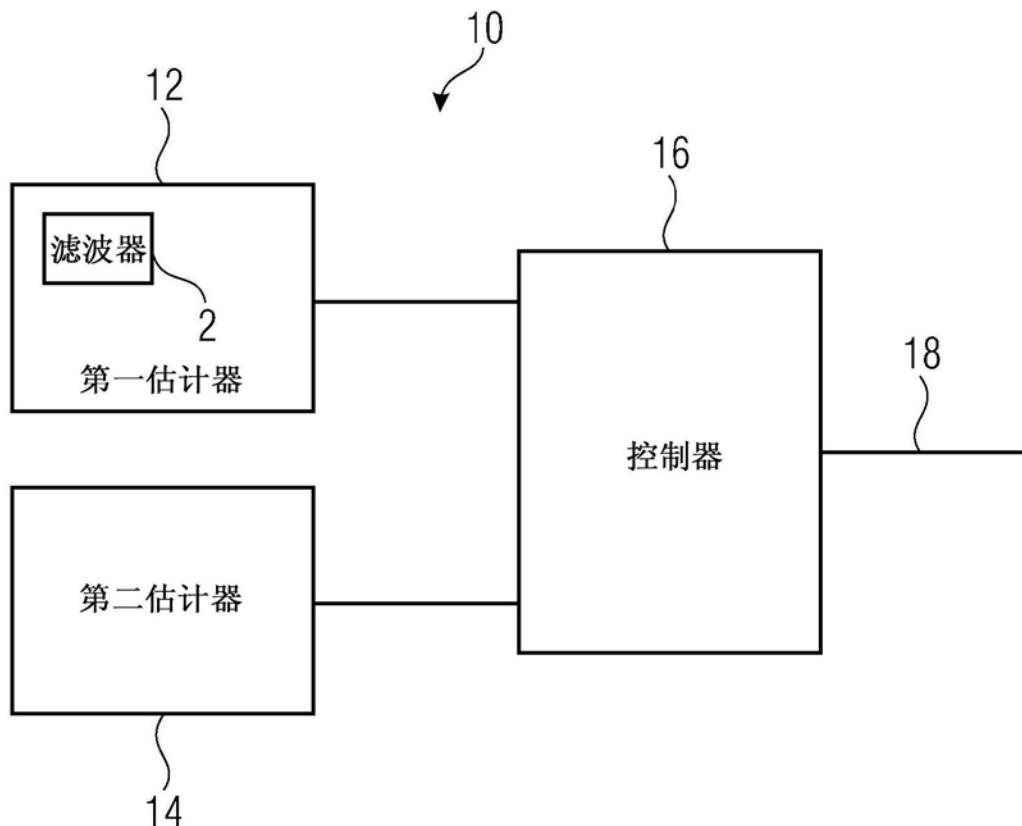


图1

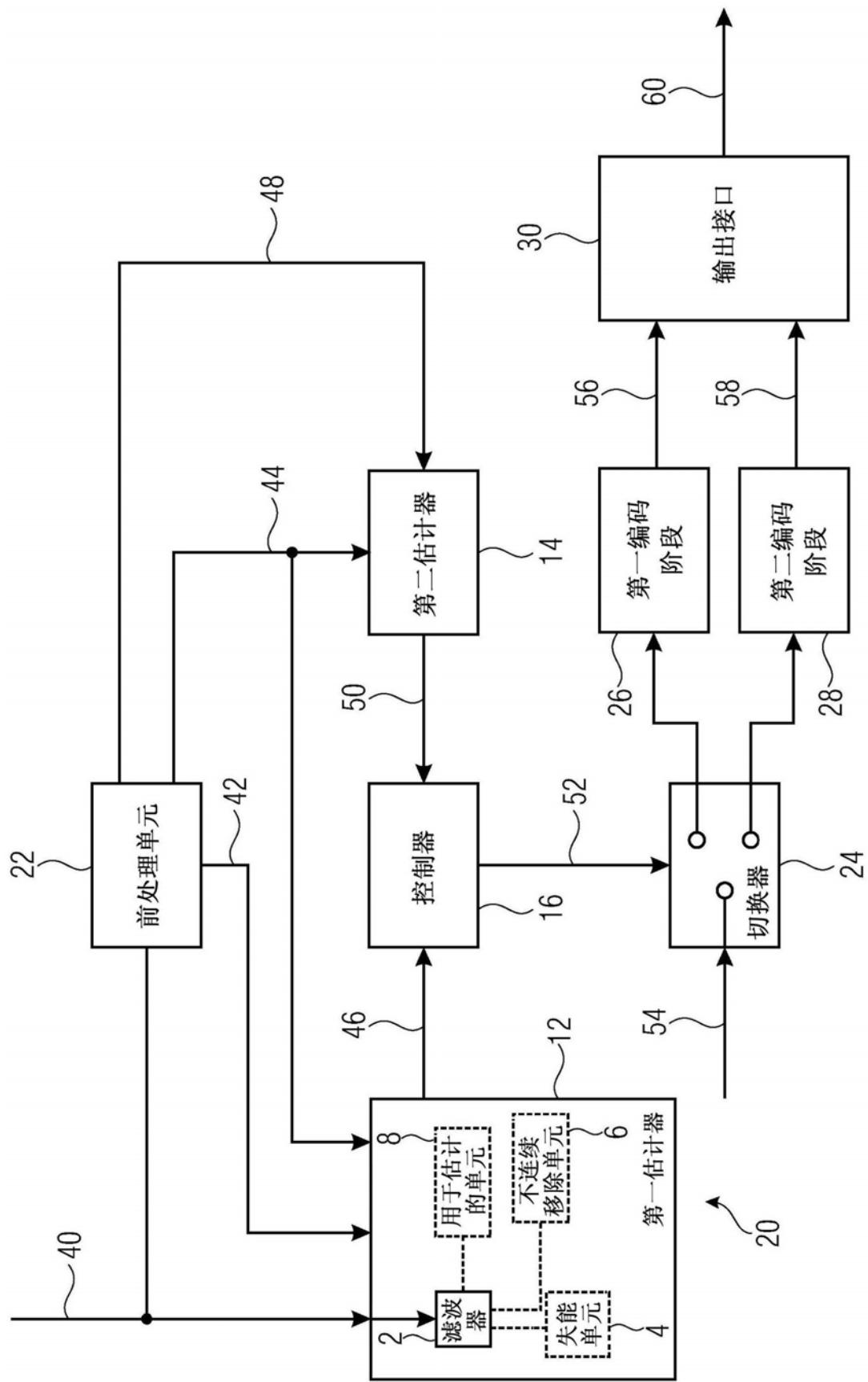


图2

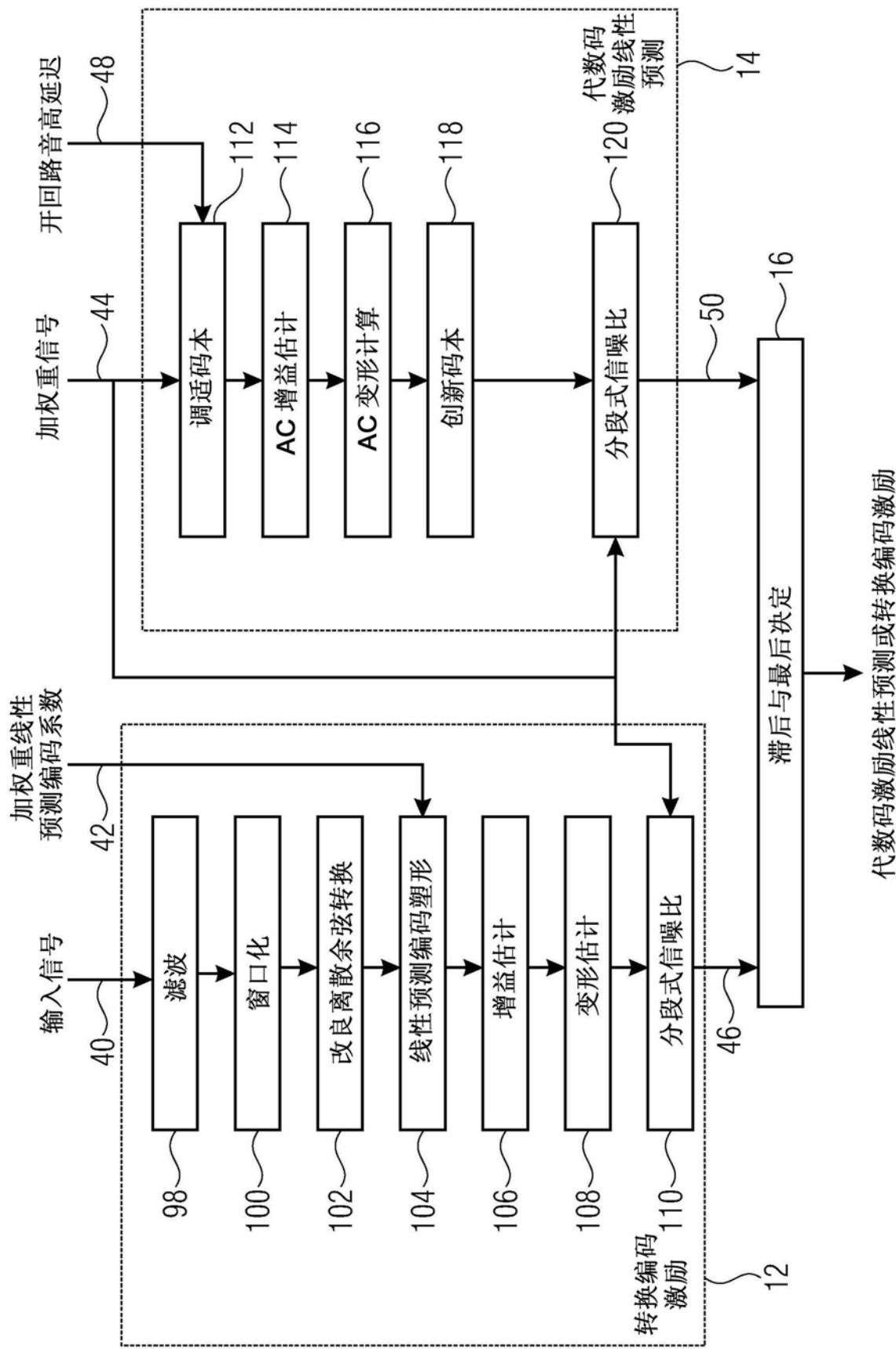


图3

$$SNR = 10 \log_{10} \frac{\sum_{i=1}^N x^2(i)}{\sum_{i=1}^N (x(i) - y(i))^2}$$

图4A

$$SNR_{seg} = \frac{10}{M} \sum_{m=0}^{M-1} \log_{10} \sum_{i=Nm}^{Nm+N-1} \left(\frac{\sum_{i=1}^N x^2(i)}{\sum_{i=1}^N (x(i) - y(i))^2} \right)$$

图4B