



(12)发明专利

(10)授权公告号 CN 105229736 B

(45)授权公告日 2019.07.19

(21)申请号 201480019093.0

(22)申请日 2014.01.28

(65)同一申请的已公布的文献号
申请公布号 CN 105229736 A

(43)申请公布日 2016.01.06

(30)优先权数据
61/758,100 2013.01.29 US

(85)PCT国际申请进入国家阶段日
2015.09.28

(86)PCT国际申请的申请数据
PCT/EP2014/051557 2014.01.28

(87)PCT国际申请的公布数据
W02014/118136 EN 2014.08.07

(73)专利权人 弗劳恩霍夫应用研究促进协会
地址 德国慕尼黑

(72)发明人 埃曼努埃尔·拉维利
斯特凡·多赫拉 纪尧姆·福奇斯
(续)

(74)专利代理机构 北京康信知识产权代理有限
责任公司 11240

代理人 梁丽超 陈鹏

(51)Int.Cl.
G10L 19/22(2013.01) (续)

(56)对比文件
WO 2012/110448 A1,2012.08.23,
CN 102113051 A,2011.06.29,
CN 102099856 B,2012.11.07, (续)
审查员 可杨

权利要求书4页 说明书10页 附图4页

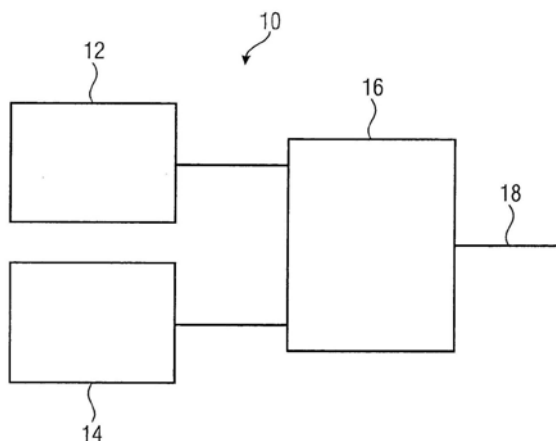
(54)发明名称

用于选择第一编码算法与第二编码算法中的一个的装置及方法

(57)摘要

一种用以选择具有一第一特性的一第一编码算法和具有一第二特性的一第二编码算法中的一个的装置,该算法用于编码一音频信号的一部分,以获得该音频信号的该部分的一经编码版本,该装置包含一第一估计器,其用于在实际上并不使用该第一编码算法编码及解码该音频信号的该部分的情况下,估计该音频信号的该部分的一第一质量测量,该第一质量测量与该第一编码算法相关联。提供用于在实际上并不使用该第二编码算法编码及解码该音频信号的该部分的情况下,估计该音频信号的该部分的一第二质量测量的一第二估计器,该第二质量测量与该第二编码算法相关联。该装置包含用于基于该第一质

量测量与该第二质量测量之间的一比较,选择该第一编码算法或该第二编码算法的一控制器。



CN 105229736 B

[接上页]

(72)发明人 埃莱尼·福托普洛
克里斯蒂安·赫尔姆里希

(51)Int.Cl.

G10L 19/08(2013.01)

G10L 19/02(2013.01)

(56)对比文件

CN 101261834 A,2008.09.10,
EUROPEAN TELECOMMUNICATIONS STANDARDS

INSTITUTE(ETSI).“DIGITAL CELLULAR
TELECOMMUNICATIONS SYSTEM(PHASE 2+);
UNIVERSAL MOBILE TELECOMMUNICATIONS
SYSTEM(UMTS);LTE;AUDIO CODEC PROCESSING
FUNCTIONS;EXTENDED ADAPTIVE MULTI-RATE-
WIDEBAND(AMR-WB+)CODEC;TRANSCODING
FUNCTIONS(3GPP TS 26.290 VERSION 11.0.0
RELEASE 11)”.《TECHNICAL SPECIFICATION》
.2012,第3GPP SA 4卷(第V11.0.0期),

1. 一种用以选择具有一第一特性的一第一编码算法及具有一第二特性的一第二编码算法中的一个的装置 (10), 该算法用于编码一音频信号 (40) 的一部分, 以获得该音频信号 (40) 的该部分的一经编码版本, 该装置包含:

一第一估计器 (12), 其用于在实际上并不使用该第一编码算法编码及解码该音频信号的该部分的情况下, 估计该音频信号的该部分的一第一质量测量, 该第一质量测量与该第一编码算法相关联;

一第二估计器 (14), 其用于在实际上并不使用该第二编码算法编码及解码该音频信号的该部分的情况下, 估计该音频信号的该部分的一第二质量测量, 该第二质量测量与该第二编码算法相关联;

以及

一控制器 (16), 其用于基于该第一质量测量与该第二质量测量之间的一比较, 选择该第一编码算法或该第二编码算法,

其中, 该第一质量测量和该第二质量测量是该音频信号的加权版本的对应部分的SNR (信噪比) 或区段性SNR,

其中该第一估计器 (12) 被配置为判定当量化该音频信号的该部分时, 用于该第一编码算法的一量化器将引入的一估计量化器失真, 并基于该音频信号的一加权版本的一部分能量及该估计量化器失真, 估计该第一质量测量。

2. 如权利要求1所述的装置 (10), 其中该第一编码算法为一变换编码算法、一基于编码算法的MDCT (修改离散余弦变换) 或一TCX (变换编码激励) 编码算法, 且其中该第二编码算法为一CELP (码激励线性预测) 编码算法, 或一ACELP (代数码激励线性预测) 编码算法。

3. 如权利要求1所述的装置 (10), 其中该第一估计器 (12) 被配置为估计该音频信号的该部分的一全局增益, 使得当使用该第一编码算法的一量化器及一熵编码器编码时, 该音频信号的该部分将产生一给定目标比特率, 其中该第一估计器 (12) 进一步被配置为基于所估计的全局增益的一功率, 判定该估计量化器失真, 其中用于该第一编码算法的该量化器为一均匀标量量化器, 且其中该第一估计器 (12) 被配置为使用公式 $D=G*G/12$ 判定该估计量化器失真, 其中D为该估计量化器失真且G为所估计的全局增益。

4. 如权利要求1所述的装置 (10), 其中该第一质量测量为加权音频信号的一部分的一区段性SNR, 且其中该第一估计器 (12) 被配置为基于该加权音频信号对应的子部分的一能量及该所估计量化器失真, 通过计算与该加权音频信号的该部分的多个子部分中的每一个相关联的一所估计的SNR来估计该区段性SNR, 并通过计算与该加权音频信号的该部分的该子部分相关联的该SNR的一平均, 以获得该加权音频信号的该部分的所估计的区段性SNR。

5. 如权利要求1所述的装置 (10), 其中该控制器 (16) 被配置为在比较该估计质量测量时利用一磁滞。

6. 一种用以选择具有一第一特性的一第一编码算法及具有一第二特性的一第二编码算法中的一个的装置 (10), 该算法用于编码一音频信号 (40) 的一部分, 以获得该音频信号 (40) 的该部分的一经编码版本, 该装置包含:

一第一估计器 (12), 其用于在实际上并不使用该第一编码算法编码及解码该音频信号的该部分的情况下, 估计该音频信号的该部分的一第一质量测量, 该第一质量测量与该第一编码算法相关联;

一第二估计器(14),其用于在实际上并不使用该第二编码算法编码及解码该音频信号的该部分的情况下,估计该音频信号的该部分的一第二质量测量,该第二质量测量与该第二编码算法相关联;

以及

一控制器(16),其用于基于该第一质量测量与该第二质量测量之间的一比较,选择该第一编码算法或该第二编码算法,

其中,该第一质量测量和该第二质量测量是该音频信号的加权版本的对应部分的SNR(信噪比)或区段性SNR,

其中该第二估计器(14)被配置为判定当使用一自适应码本以编码该音频信号的该部分时,用于该第二编码算法的该自适应码本将引入一估计自适应码本失真,且其中该第二估计器(14)被配置为基于该音频信号的一加权版本的一部分的一能量及该估计自适应码本失真,估计该第二质量测量,其中,对于该音频信号的该部分的多个子部分中的每一个,该第二估计器(14)被配置为:基于由一预处理平台中所判定的一音调滞后而转换到过去的该加权音频信号的该子部分的一版本,近似该自适应码本;估计一自适应码本增益,使得该加权音频信号的该部分的该子部分与该所近似的自适应码本之间的一误差最小化;并基于由该自适应码本增益按比例调整的,该加权音频信号的该部分的该子部分与所近似的自适应码本之间的一误差的能量,判定该估计自适应码本失真。

7.如权利要求6所述的装置(10),其中该第二估计器(14)进一步被配置为将该音频信号的该部分的每一子部分判定的该估计自适应码本失真减少一常数因子。

8.如权利要求6所述的装置(10),其中该第二质量测量为该加权音频信号的该部分的一区段性SNR,且其中该第二估计器(14)被配置为基于该加权音频信号的该对应的子部分的该能量及该估计自适应码本失真,通过计算与每个子部分相关联的一估计SNR来估计该区段性SNR,并通过计算与该子部分相关联的该SNR的一平均,以获得该加权音频信号的该部分的该估计区段性SNR。

9.如权利要求6所述的装置(10),其中该第二估计器(14)被配置为:基于由一预处理平台中所判定的一音调滞后而转换到过去的该加权音频信号的该部分的一版本,近似该自适应码本;估计一自适应码本增益,使得该加权音频信号的该部分与该所近似的自适应码本之间的一误差最小化;并基于由该自适应码本增益按比例调整的该加权音频信号的该部分与该所近似的自适应码本之间的一误差的能量,判定该估计自适应码本失真。

10.一种用于编码一音频信号的一部分的装置(20),包含如权利要求1的装置(10)、用于执行第一编码算法的一第一编码器平台(26),以及用于执行第二编码算法的一第二编码器平台(28),其中用于编码的该装置(20)被配置为取决于由该控制器(16)作出的选择使用该第一编码算法或该第二编码算法编码该音频信号的该部分。

11.一种用于编码及解码的系统,该系统包含如权利要求10的用于编码的一装置(20)及一解码器,该解码器被配置为接收音频信号的部分的经编码版本和用于编码该音频信号的该部分的算法的一指示,并使用所指示的算法解码该音频信号的该部分的该经编码版本。

12.一种用于编码一音频信号的一部分的装置(20),包含如权利要求6的装置(10)、用于执行第一编码算法的一第一编码器平台(26),以及用于执行第二编码算法的一第二编码

器平台(28),其中用于编码的该装置(20)被配置为取决于由该控制器(16)作出的选择使用该第一编码算法或该第二编码算法编码该音频信号的该部分。

13.一种用以选择具有一第一特性的一第一编码算法及具有一第二特性的一第二编码算法中的一个的方法,该算法用于编码一音频信号的一部分以获得该音频信号的该部分的一经编码版本,该方法包含:

在实际上并不使用该第一编码算法编码及解码该音频信号的该部分的情况下,估计该音频信号的该部分的一第一质量测量,该第一质量测量与该第一编码算法相关联;

在实际上并不使用该第二编码算法编码及解码该音频信号的该部分的情况下,估计该音频信号的该部分的一第二质量测量,该第二质量测量与该第二编码算法相关联;以及

基于该第一质量测量与该第二质量测量之间的一比较,选择该第一编码算法或该第二编码算法,

其中,该第一质量测量和该第二质量测量是该音频信号的加权版本的对应部分的SNR(信噪比)或区段性SNR,

所述方法进一步包含:判定(108)当量化该音频信号的该部分时,用于该第一编码算法的一量化器将引入的一估计量化器失真;

并基于该音频信号的一加权版本的一部分的一能量及该估计量化器失真判定该质量测量。

14.如权利要求13所述的方法,其中该第一编码算法为一变换编码算法、一基于编码算法的MDCT(修改离散余弦变换)或一TCX(变换编码激励)编码算法,且其中该第二编码算法为一CELP(码激励线性预测)编码算法,或一ACELP(代数码激励线性预测)编码算法。

15.如权利要求13所述的方法,其包含:估计(106)该音频信号的该部分的一全局增益,使得当通过用于该第一编码算法的一量化器及一熵编码器编码时,该音频信号的该部分将产生一给定目标比特率;

并基于所估计的全局增益的一功率判定(108)该估计量化器失真,其中该量化器为一均匀标量量化器,其中使用公式 $D=G*G/12$ 判定该估计量化器失真,其中D为该估计量化器失真且G为所估计的全局增益。

16.如权利要求13所述的方法,其中该第一质量测量为该加权音频信号的一部分的经LPC滤波版本的一区段性SNR,且该方法包含:基于该加权音频信号对应的子部分的一能量及该估计量化器失真,通过计算与该加权音频信号的该部分的多个子部分中的每一个相关联的一所估计的SNR来估计该区段性SNR;并通过计算与该加权音频信号的该部分的该子部分相关联的该SNR的一平均,以获得该加权音频信号的该部分的所估计的区段性SNR。

17.如权利要求13所述的方法,其包含在比较该估计质量测量时利用一磁滞。

18.一种用以选择具有一第一特性的一第一编码算法及具有一第二特性的一第二编码算法中的一个的方法,该算法用于编码一音频信号的一部分以获得该音频信号的该部分的一经编码版本,该方法包含:

在实际上并不使用该第一编码算法编码及解码该音频信号的该部分的情况下,估计该音频信号的该部分的一第一质量测量,该第一质量测量与该第一编码算法相关联;

在实际上并不使用该第二编码算法编码及解码该音频信号的该部分的情况下,估计该音频信号的该部分的一第二质量测量,该第二质量测量与该第二编码算法相关联;以及

基于该第一质量测量与该第二质量测量之间的一比较,选择该第一编码算法或该第二编码算法,

其中,该第一质量测量和该第二质量测量是该音频信号的加权版本的对应部分的SNR(信噪比)或区段性SNR,

所述方法进一步包含:判定当使用一自适应码本以编码该音频信号的该部分时,用于该第二编码算法的该自适应码本将引入一估计自适应码本失真(116);并基于该音频信号的一加权版本的一部分的一能量及该估计自适应码本失真,估计该第二质量测量,以及该方法包含:对于该音频信号的该部分的多个子部分中的每一个,基于由一预处理平台中所判定的一音调滞后而转换到过去的该加权音频信号的该子部分的一版本,近似(112)该自适应码本;估计(114)一自适应码本增益,使得该加权音频信号的该部分的该子部分与该所近似的自适应码本之间的一误差最小化;并基于由该自适应码本增益按比例调整的,该加权音频信号的该部分的该子部分与所近似的自适应码本之间的一误差的能量,判定(116)该估计自适应码本失真。

19.如权利要求18所述的方法,其包含将该音频信号的该部分的每一子部分判定的该估计自适应码本失真减少(118)一常数因子。

20.如权利要求18所述的方法,其中该第二质量测量为该加权音频信号的该部分的一区段性SNR,且该方法包含:基于该加权音频信号的该对应的子部分的该能量及该估计自适应码本失真,通过计算与每个子部分相关联的一估计SNR来估计该区段性SNR;并通过计算与该子部分相关联的该SNR的一平均,以获得该加权音频信号的该部分的该估计区段性SNR。

21.如权利要求18所述的方法,其包含:基于由一预处理平台中所判定的一音调滞后而转换到过去的该加权音频信号的该部分的一版本,近似该自适应码本;估计一自适应码本增益,使得该加权音频信号的该部分与该所近似的自适应码本之间的一误差最小化;并基于由该自适应码本增益按比例调整的该加权音频信号的该部分与该所近似自适应码本之间的一误差的能量,判定该估计自适应码本失真。

22.一种存储有计算机程序的非暂时性储存媒体,所述程序具有代码,当所述代码运行于一计算机上时,执行如权利要求13或18所述的方法。

用于选择第一编码算法与第二编码算法中的一个的装置及方法

技术领域

[0001] 本发明涉及音频编码,更具体地,涉及切换式音频编码,其中对于音频信号的不同部分,使用不同编码算法产生经编码信号。

背景技术

[0002] 已知判定用于音频信号的不同部分的不同编码算法的切换式音频编码器。大体而言,切换式音频编码器提供在两个不同模式(即,算法,诸如ACELP(代数码激励线性预测(Algebraic Code Excited Linear Prediction;ACELP))及TCX(变换编码激励(Transform Coded Excitation;TCX)))之间进行切换。

[0003] MPEG USAC(MPEG统一语音音频编码(Unified Speech Audio Coding;USAC))的LPD模式基于该两个不同模式ACELP及TCX。ACELP 为语音状及瞬态状信号提供较好质量。TCX为音乐状及噪声状信号提供较好质量。编码器逐帧地决定使用哪种模式。编码器所作出的决定对于编解码器质量至关重要。单一的错误决定可产生大量伪像,尤其在低比特率的情况下。

[0004] 用于决定使用哪种模式的最直接方法为封闭回路模式选择,亦即执行两种模式的完整编码/解码,接着基于音频信号及经编码/经解码音频信号计算用于两种模式的选择准则(例如,区段性SNR),且最后基于选择准则选择模式。此方法大体上产生稳定且稳固的决定。然而,其亦要求大量复杂性,因为必须在每一帧处运行两种模式。

[0005] 为减少复杂性,替代性方法为开放回路模式选择。开放回路选择由并不执行两种模式的完整编码/解码,但代替地使用通过低复杂性所计算的选择准则而选择一模式组成。接着,通过最不复杂模式(通常为TCX)的复杂性减去计算选择准则所需的复杂性而减少最差状况复杂性。通常节约大量复杂性,此情况使得当编解码器最差状况复杂性受到约束时,此种方法是有吸引力的。

[0006] AMR-WB+标准(国际标准3GPP TS 26.290 V6.1.0 2004-12中所定义)包括用于在80ms帧中,在ACELP/TCX20/TCX40/TCX80的所有组合之间决定的开放回路模式选择。其描述于3GPP TS 26.290的章节5.2.4中。其亦描述于「用于行动、多媒体的较不复杂音频编码(Low Complex Audio Encoding for Mobile,Multimedia),VTC 2006,Makinen等人」的会议文件中,及追溯至此会议文件的作者的US7,747,430 B2及US 7,739,120 B2。

[0007] US7,747,430 B2揭示基于长期预测参数的分析的开放回路模式选择。US 7,739,120 B2揭示基于信号特性的开放回路模式选择,该信号特性指示音频信号的各自区段中的音频内容的类型,其中,若此选择并不可行,则进一步基于统计评价进行用于各自相邻区段的选择。

[0008] 可以两个主要步骤描述AMR-WB+的开放回路模式选择。在第一主要步骤中,对音频信号进行若干特征计算,诸如能量位准的标准偏差、低频 /高频能量比例、总能量、ISP(导抗谱对(immittance spectral pair;ISP))距离、音调滞后及增益、频谱倾斜。接着,使用简

单的基于临界分类器,将此等特征用于在ACELP与TCX之间作出选择。若在第一主要步骤中选择TCX,则第二主要步骤以封闭回路方式在TCX20/TCX40/TCX80的可能组合之间决定。

[0009] WO 2012/110448 A1揭示用于基于音频信号的瞬态侦测结果及质量结果,在具有不同特性的两个编码算法之间作出决定的方法。另外,揭示应用磁滞,其中磁滞依赖于过去所作出的选择,亦即对音频信号的较早部分所作出的选择。

[0010] 在“用于行动、多媒体的较不复杂音频编码(Low Complex Audio Encoding for Mobile, Multimedia), VTC 2006, Makinen等人”的会议文件中,对AMR-WB+的封闭回路及开放回路模式选择进行比较。主观收听测试指示开放回路模式选择执行显著较差于封闭回路模式选择。但亦展示,开放回路模式选择减少40%的最差状况复杂性。

发明内容

[0011] 本发明的目标在于提供一种允许在一第一编码算法与具有良好效能及减少的复杂性的一第二编码算法之间作出选择的经改良方法。

[0012] 由装置、方法及计算机程序实现此目标。

[0013] 本发明的实施例提供一种用以选择具有一第一特性的一第一编码算法及具有一第二特性的一第二编码算法中的一个的装置,该算法用于编码一音频信号的一部分,以获得该音频信号的该部分的一经编码版本,该装置包含:

[0014] 一第一估计器,其用于在实际上并不使用该第一编码算法编码及解码该音频信号的该部分的情况下,估计该音频信号的该部分的一第一质量测量,该第一质量测量与该第一编码算法相关联;

[0015] 一第二估计器,其用于在实际上并不使用该第二编码算法编码及解码该音频信号的该部分的情况下,估计该音频信号的该部分的一第二质量测量,该第二质量测量与该第二编码算法相关联;以及

[0016] 一控制器,其用于基于该第一质量测量与该第二质量测量之间的一比较,选择该第一编码算法或该第二编码算法。

[0017] 本发明的实施例提供一种用以选择具有一第一特性的一第一编码算法及具有一第二特性的一第二编码算法中的一个的方法,该算法用于编码一音频信号的一部分,以获得该音频信号的该部分的一经编码版本,该方法包含:

[0018] 在实际上并不使用该第一编码算法编码及解码该音频信号的该部分的情况下,估计该音频信号的该部分的一第一质量测量,该第一质量测量与该第一编码算法相关联;

[0019] 在实际上并不使用该第二编码算法编码及解码该音频信号的该部分的情况下,估计该音频信号的该部分的一第二质量测量,该第二质量测量与该第二编码算法相关联;及

[0020] 基于该第一质量测量与该第二质量测量之间的一比较,选择该第一编码算法或该第二编码算法。

[0021] 本发明的实施例基于如下认识:可通过估计第一编码算法及第二编码算法中的每一个的一质量测量,并基于该第一质量测量与该第二质量测量之间的一比较选择该编码算法中的一个,而实施具有改良的效能的一开放回路选择。估计该质量测量,亦即实际上并不编码及解码该音频信号以获得该质量测量。因此,可通过减少的复杂性而获得该质量测量。接着,可使用该估计质量测量执行与一封闭回路模式选择类似的模式选择。

[0022] 在本发明的实施例中,实施首先通过较低的复杂性估计ACELP及 TCX的区段性SNR的一开放回路模式选择。且接着,类似于在一封闭回路模式选择中,使用此等估计区段性SNR值执行该模式选择。

[0023] 本发明的实施例并不类似于AMR-WB+的该开放回路模式选择中所进行地利用一经典的特征+分类器方法。但代替地,本发明的实施例试图估计每一模式的一质量测量,并选择给出最佳质量的该模式。

附图说明

[0024] 现将参考随附图进一步详细地描述本发明的实施例,其中:

[0025] 图1示出用以选择第一编码算法与第二编码算法中的一个的装置的实施例的示意图;

[0026] 图2示出用于编码音频信号的装置的实施例的示意图;

[0027] 图3示出用以选择第一编码算法与第二编码算法中的一个的装置的实施例的示意图;

[0028] 图4A和图4B可能表示SNR及区段性SNR。

具体实施方式

[0029] 在以下描述中,通过相同参考记号参考不同图式中的类似组件/步骤。应注意,在图式中,已省略理解本发明并不必要的特征(诸如,信号连接及类似者)。

[0030] 图1展示用以选择第一编码算法(诸如,TCX算法)与第二编码算法(诸如,ACELP算法)中的一个的装置10,如用于编码音频信号的一部分的编码器。装置10包含用于估计信号部分的第一质量测量的第一估计器12。第一质量测量与第一编码算法相关联。换言之,第一估计器12估计若使用第一编码算法编码并解码,音频信号的部分将具有的第一质量测量,而实际上并不使用第一编码算法编码及解码音频信号的部分。装置10包含用于估计信号部分的第二质量测量的第二估计器14。第二质量测量与第二编码算法相关联。换言之,第二估计器14估计若使用第二编码算法编码并解码,音频信号的部分将具有的第二质量测量,而实际上并不使用第二编码算法编码及解码音频信号的部分。此外,装置10包含用以基于第一质量测量与第二质量测量之间的比较,选择第一编码算法或第二编码算法的控制器16。控制器可包含指示所选择编码算法的输出18。

[0031] 在一实施例中,与第一编码算法相关联的第一特性较好地适于音乐状及噪声状信号,而与第二编码算法相关联的第二编码特性较好地适于语音状及瞬态状信号。在本发明的实施例中,第一编码算法为音频编码算法(诸如,变换编码算法),例如MDCT(修改离散余弦变换(modified discrete cosine transform;MDCT))编码算法,诸如TCX(变换编码激励)编码算法。其他变换编码算法可基于FFT变换或任何其他变换或滤波器组。在本发明的实施例中,第二编码算法为语音编码算法,诸如,CELP(码激励线性预测)编码算法,诸如ACELP(代数码激励线性预测)编码算法。

[0032] 在实施例中,质量测量表示感知质量测量。可计算为第一编码算法的主观质量的估计的单一值,及为第二编码算法的主观质量的估计的单一值。可仅基于此等两值的比较选择给出最佳经估计主观质量的编码算法。此情况不同于AMR-WB+标准中所进行的内容,在

AMR-WB+标准中,计算表示信号的不同特性的许多特征,且接着应用分类器以决定选择哪种算法。

[0033] 在实施例中,基于加权音频信号(亦即,音频信号的加权版本)的一部分估计各自品质测量。在实施例中,可将加权音频信号定义为由加权功能滤波的音频信号,其中加权功能为加权LPC滤波器 $A(z/g)$,其中 $A(z)$ 为LPC滤波器且 g 为介于0与1之间的重量(诸如,0.68)。结果为可以此方式获得感知质量的良好测量。应注意,在预处理平台中判定LPC滤波器 $A(z)$ 及加权LPC滤波器 $A(z/g)$,且其亦用于两种编码算法中。在其他实施例中,加权功能可为线性滤波器、FIR滤波器或线性预测滤波器。

[0034] 在实施例中,质量测量为加权信号域中的区段性SNR(信噪比(signal to noise ratio;SNR))。结果为,加权信号域中的区段性SNR表示感知质量的良好测量,且因此,其可以有有益方式用作品质测量。此区段性SNR亦为 ACELP及TCX编码算法两者中所使用以估计编码参数的质量测量。

[0035] 另一品质测量可为加权信号域中的SNR。其他质量测量可为区段性 SNR,非加权(亦即,未由(加权)LPC系数滤波)信号域中的音频信号的对应部分的SNR。其他质量测量可为倒频谱失真或杂遮比(noise-to-mask ratio;NMR)。

[0036] 大体而言,SNR逐样本地比较原始音频信号与经处理音频信号(诸如,语音信号)。其目的在于测量再生输入波形的波形编码器的失真。SNR可经计算为如图4A中所展示,其中 $x(i)$ 及 $y(i)$ 为由 i 索引的原始样本及经处理样本,且 N 为样本的总数目。区段性SNR计算较短区段(诸如,1ms 至10ms,诸如5ms)的SNR值的平均,而非操作全部信号。SNR可经计算为如图4B中所展示,其中 N 及 M 分别为区段长度及区段数目。

[0037] 在本发明的实施例中,音频信号的部分表示通过开窗口音频信号而获得的音频信号的帧,并针对通过开窗口音频信号而获得的多个连续帧执行适当编码算法的选择。在以下规范中,结合音频信号,以可交换方式使用术语「部分」及「帧」。在实施例中,将每一帧划分成子帧,并通过计算每一子帧的SNR(转变为dB),且以dB为单位计算子帧SNR的平均,而估计每一帧的区段性SNR。

[0038] 因此,在实施例中,并非估计输入音频信号与经解码音频信号之间的(区段性)SNR,而估计加权输入音频信号与加权经解码音频信号之间的(区段性)SNR。就相关此(区段性)SNR而言,可参考AMR-WB+标准(国际标准3GPP TS 26.290 V6.1.0 2004-12)的5.2.3章。

[0039] 在本发明的实施例中,基于加权音频信号的一部分的能量,并基于通过各自算法编码信号部分时所引入的估计失真估计各自质量测量,其中第一及第二估计器被配置为取决于加权音频信号的能量而判定估计失真。

[0040] 在本发明的实施例中,判定在量化音频信号的部分时由用于第一编码算法的量化器所引入的估计量化器失真,并基于加权音频信号的部分的能量及估计量化器失真判定第一质量测量。在此等实施例中,可估计音频信号的部分的全局增益,使得当通过用于第一编码算法的量化器及熵编码器编码时,音频信号的部分将产生给定目标比特率,其中基于估计全局增益判定估计的量化器失真。在此等实施例中,可基于估计增益的功率判定估计量化器失真。当用于第一编码算法的量化器为均匀标量量化器时,第一估计器可被配置为使用公式 $D=G*G/12$ 来判定估计量化器失真,其中 D 为估计量化器失真且 G 为估计全局增益。倘若第一编码算法使用另一量化器,则可自全局增益以不同方式判定量化器失真。

[0041] 本发明人认识到,可通过使用呈其任何组合的上文特征,来以适当方式估计将在使用第一编码算法(诸如,TCX算法)编码及解码音频信号的部分时获得的质量测量(诸如,区段性SNR)。

[0042] 在本发明的实施例中,第一质量测量为区段性SNR,且基于加权音频信号的对应子部分的能量及估计量化器失真,通过计算与音频信号的部分的多个子部分中的每一个相关联的估计SNR来估计区段性SNR,并通过计算与加权音频信号的部分的子部分相关联的SNR的平均,以获得加权音频信号的该部分的估计区段性SNR。

[0043] 在本发明的实施例中,判定在使用自适应码本编码音频信号的部分时由用于第二编码算法的自适应码本所引入的估计自适应码本失真,并基于加权音频信号的部分的能量及估计自适应码本失真,估计第二质量测量。

[0044] 在此等实施例中,对于音频信号的部分的多个子部分中的每一个,可基于由预处理平台中所判定的音调滞后转换至过去的加权音频信号的部分的版本,近似自适应码本;可估计自适应码本增益,使得最小化加权音频信号的部分的子部分与经近似自适应码本之间的误差;且可基于由自适应码本增益按比例调整的加权音频信号的部分的子部分与经近似自适应码本之间的误差的能量,判定估计自适应码本失真。

[0045] 在本发明的实施例中,可将所判定的音频信号的部分的每一子部分的估计自适应码本失真减少常数因子,以便考虑由第二编码算法中的创新码本所实现的失真减少。

[0046] 在本发明的实施例中,第二质量测量为区段性SNR,且基于加权音频信号的对应子部分的能量及估计自适应码本失真,通过计算与每一子部分相关联的估计SNR来估计区段性SNR,并通过计算与子部分相关联的SNR的平均以获得估计区段性SNR。

[0047] 在本发明的实施例中,基于由预处理平台中所判定的音调滞后转换至过去的加权音频信号的部分的版本,近似自适应码本;估计自适应码本增益,使得最小化加权音频信号的部分与经近似自适应码本之间的误差;并基于由自适应码本增益按比例调整的加权音频信号的部分与经近似自适应码本之间的能量,判定估计自适应码本失真。因此,可以较少的复杂性判定估计自适应码本失真。

[0048] 本发明人认识到,可通过使用呈其任何组合的上文特征,来以适当方式估计将在使用第二编码算法(诸如,ACELP算法)编码及解码音频信号的部分时获得的质量测量(诸如,区段性SNR)。

[0049] 在本发明的实施例中,磁滞机制用于比较估计质量测量。此操作可作出待较稳定地使用哪种算法的决定。磁滞机制可取决于估计质量测量(诸如,其间的差异)及其他参数,诸如,关于先前决定的统计、时间固定帧的数目、帧中的瞬态。就相关此等磁滞机制而言,可参考(例如)WO 2012/110448 A1。

[0050] 在本发明的实施例中,用于编码音频信号的编码器包含装置10,用于执行第一编码算法的平台及用于执行第二编码算法的平台,其中取决于由控制器16作出的选择,编码器被配置为使用第一编码算法或第二编码算法来编码音频信号的部分。在本发明的实施例中,用于编码及解码的系统包含被配置为接收音频信号的部分的经编码版本,及用于编码音频信号的部分的算法的指示,并使用所指示算法解码音频信号的部分的经编码版本的编码器及解码器。

[0051] 在参看图3详细描述第一估计器12及第二估计器14的实施例之前,参看图2描述编

码器20的实施例。

[0052] 编码器20包含第一估计器12、第二估计器14、控制器16、预处理单元22、切换器24、被配置为执行TCX算法的第一编码器平台26、被配置为执行ACELP算法的第二编码器平台28及输出接口30。预处理单元22可为共同USAC编码器的部分,且可被配置为输出LPC系数、加权LPC系数、加权音频信号及音调滞后的集合。应注意,所有此等参数皆用于两种编码算法,亦即TCX算法及ACELP算法。因此,不必另外针对开放回路模式决定计算此等参数。在开放回路模式决定中使用已计算出的参数的优势在于节约复杂性。

[0053] 在输入在线提供输入音频信号40。将输入音频信号40应用于第一估计器12、预处理单元22及编码器平台26、28两者。预处理单元22以习知方式处理输入音频信号,以导出LPC系数及加权LPC系数42,并通过加权LPC系数42滤波音频信号40以获得加权音频信号44。预处理单元22输出加权LPC系数42、加权音频信号44及音调滞后的集合48。如熟习此项技术者所理解,可将加权LPC系数42及加权音频信号44分段化为帧或子帧。可通过以适当方式开窗口音频信号而获得分段。

[0054] 在本发明的实施例中,可使用经量化LPC系数或经量化加权LPC系数。因此,应理解,术语「LPC系数」亦意欲涵盖「经量化LPC系数」,且术语「加权LPC系数」亦意欲涵盖「加权经量化系数」。就此而言,值得注意的是,USAC的TCX算法使用经量化加权LPC系数以塑形MCDT频谱。

[0055] 第一估计器12接收音频信号40、加权LPC系数42及加权音频信号44,基于上述各者估计第一品质测量46,并将第一质量测量输出至控制器16。第二估计器16接收加权音频信号44及音调滞后的集合48,基于上述两者估计第二质量测量50,并将第二质量测量50输出至控制器16。如熟习此项技术者所已知,已在先前模块(亦即,预处理单元22)中计算出加权LPC系数42、加权音频信号44及音调滞后的集合48,且因此,上述各者无需成本而可用。

[0056] 控制器基于所接收质量测量的比较,作出TCX算法抑或ACELP算法的选择决定。如上文所指示,控制器可在决定使用哪种算法时使用磁滞机制。图2中借助于由控制器16所输出的控制信号52控制的切换器24,示意性地展示选择第一编码器平台26抑或第二编码器平台28。控制信号52指示待使用第一编码器平台26抑或第二编码器平台28。基于控制信号52,由图2中的箭头54所示意性地指示,且至少包括LPC系数、加权LPC系数、音频信号、加权音频信号、音调滞后的集合的所要求信号应用于第一编码器平台26抑或第二编码器平台28。所选择的编码器平台应用相关联编码算法,并将经编码表示56或58输出至输出接口30。输出接口30可被配置为输出经编码音频信号,其可包含经编码表示56或58、LPC系数或加权LPC系数、用于所选择编码算法的参数及关于所选择编码算法的信息(以及其他数据)。

[0057] 现参看图3描述用于估计第一及第二质量测量的具体实施例,其中第一及第二质量测量为加权信号域中的区段性SNR。图3以逐步骤地展示各自估计的流程图的形式展示第一估计器12及第二估计器14及其功能性。

[0058] TCX区段性SNR的估计

[0059] 第一(TCX)估计器将音频信号40(输入信号)、加权LPC系数42及加权音频信号44接收为输入。

[0060] 在步骤100中,对音频信号40进行开窗口。可通过10ms低重迭正弦窗口发生开窗口。当过去帧为ACELP时,区块大小可增加5ms,窗口的左侧可以是矩形,且可自经开窗输入

信号移除ACELP合成滤波器的经开窗零脉冲回应。此情况类似于TCX算法中所进行的内容。自步骤100 输出表示音频信号的一部分的音频信号40的帧。

[0061] 在步骤102中,通过MDCT(修改离散余弦变换)变换经开窗音频信号(亦即,所得帧)。在步骤104中,通过用加权LPC系数塑形MDCT频谱而执行频谱塑形。

[0062] 在步骤106中,估计全局增益G,使得当通过熵编码器(例如,算术编码器)编码时,通过增益G所量化的加权频谱将产生给定目标R。使用术语「全局增益」,这是由于一增益是针对整个帧而判定。

[0063] 现解释全局增益估计的实施的实例。应注意,此全局增益估计适合于TCX编码算法使用具有算术编码器的标量量化器的实施例。在MPEG USAC标准中假定具有算术编码器的此标量量化器。

[0064] 初始化

[0065] 首先,通过如下各者初始化用于增益估计的变量:

[0066] 1. 将 $en[i]$ 设定为 $=9.0+10.0*\log_{10}(c[4*i+0]+c[4*i+1]+c[4*i+2]+c[4*i+3])$,

[0067] 其中 $0 \leq i < L/4$, $c[]$ 为用以量化的系数的向量,且L为 $c[]$ 的长度。

[0068] 2. 将 fac 设定为 $=128$, 将偏移设定为 $=fac$, 且将目标设定为 $=$ 任何值(例如, 1000)

[0069] 迭代

[0070] 接着,将操作的以下区块执行NITER次(例如,在此处,NITER=10)。

[0071] 1. $fac=fac/2$

[0072] 2. 偏移=偏移-fac

[0073] 3. $ener=0$

[0074] 4. 对于每一i(其中 $0 \leq i < L/4$),进行如下操作:

[0075] 若 $en[i]-\text{偏移} > 3.0$, 则 $ener=ener+en[i]-\text{偏移}$

[0076] 5. 若 $ener > \text{目标}$, 则偏移=偏移+fac

[0077] 迭代的结果为偏移值。迭代之后,将全局增益估计为 $G=10^{\wedge}(\text{偏移}/20)$ 。

[0078] 取决于所使用的量化器及熵编码器,估计全局增益的特定方式可发生变化。在MPEG USAC标准中假定具有算术编码器的标量量化器。其他TCX方法可使用不同的量化器,且熟习此项技术者应理解如何针对此等不同量化器而估计全局增益。举例来说,AMR-WB+标准假定使用RE8晶格量化器。对于此量化器,全局增益的估计可经估计为如3GPP TS 26.290 V6.1.0 2004-12的34页上的5.3.5.7章中所描述,其中假定固定的目标比特率。

[0079] 在于步骤106中估计全局增益之后,步骤108中发生失真估计。更特定而言,基于估计全局增益近似量化器失真。在本实施例中,假定使用均匀标量量化器。因此,通过单一公式 $D=G*G/12$ 判定量化器失真,其中D 表示所判定量化器失真且G表示估计全局增益。此情况对应于均匀标量量化器失真的高速率近似。

[0080] 基于所判定量化器失真,在步骤110中执行区段性SNR计算。将帧的每一子帧中的SNR计算为子帧中的加权音频信号能量与失真D(假定为常数)的比。举例来说,将帧分成四个连续子帧(参见图4A和图4B)。接着,区段性SNR为四个子帧的SNR的平均,且可以dB进行指示。

[0081] 此方法允许估计当实际上使用TCX算法编码及解码主题帧时将获得的第一区段性SNR,然而并不必实际上编码及解码音频信号,且因此,该方法具有大量减少的复杂性及减少的计算时间。

[0082] ACELP区段性SNR的估计

[0083] 第二估计器14接收预处理单元22中已计算出的加权音频信号44及音调滞后的集合48。

[0084] 如步骤112中所展示,在每一子帧中,通过简单地使用加权音频信号及音调滞后T而近似自适应码本。通过如下公式近似自适应码本

[0085] $x_w(n-T), n=0, \dots, N$

[0086] 其中 x_w 为加权音频信号,T为对应子帧的音调滞后,且N为子帧长度。因此,通过使用由T转换至过去的子帧的版本而近似自适应码本。因此,在本发明的实施例中,以极简单的方式近似自适应码本。

[0087] 在步骤114中,判定每一子帧的自适应码本增益。更特定而言,在每一子帧中,估计码本增益G,使得其最小化加权音频信号与所近似自适应码本之间的误差。可通过简单地比较每一样本的两信号之间的差异并发现增益而进行此操作,使得此等差异的和最小。

[0088] 在步骤116中,判定每一子帧的自适应码本失真。在每一子帧中,由自适应码本所引入的失真D仅为由增益G按比例调整的加权音频信号与所近似自适应码本之间的误差的能量。

[0089] 可在可选步骤118中调整步骤116中所判定的失真,以便考虑创新码本。可将用于ACELP算法的创新码本的失真简单地估计为常数值。在本发明的所描述实施例中,简单地假定创新码本将失真D减少常数因子。因此,可在步骤118中将步骤116中所获得的每一子帧的失真乘以常数因子,诸如0至1的级别的常数因子(诸如,0.055)。

[0090] 在步骤120中,发生区段性SNR的计算。在每一子帧中,将SNR计算为加权音频信号能量与失真D的比。接着,区段性SNR为四个子帧的SNR的平均值,且可以dB进行指示。

[0091] 此方法允许估计当实际上使用ACELP算法编码及解码主题帧时将获得的第二SNR,然而并不必实际上编码及解码音频信号,且因此,该方法具有大量减少的复杂性及减少的计算时间。

[0092] 第一及第二估计器12及14将估计区段性SNR 46、50输出至控制器 16,且控制器16基于估计区段性SNR 46、50,作出待对音频信号的相关联部分使用哪种算法的决定。控制器可视情况使用磁滞机制,以便作出较稳定的决定。举例来说,可通过略微不同的调谐参数使用相同于封闭回路决定中的磁滞机制的机制。此磁滞机制可计算值[dsnr],该值可取决于估计区段性SNR(诸如,其间的差异)及其他参数,诸如关于先前决定的统计、时间固定帧的数目及帧中的瞬态。

[0093] 在无磁滞机制的情况下,控制器可通过较高的估计SNR而选择编码算法,亦即若第二估计SNR高于第一估计SNR,则选择ACELP,及若第一估计SNR高于第二估计SNR,则选择TCX。在有磁滞机制的情况下,控制器可根据以下决定规则选择编码算法,其中acelp_snr为第二估计SNR且tcx_snr为第一估计SNR:

[0094] 若 $acelp_snr + dsnr > tcx_snr$,则选择ACELP,否则选择TCX。

[0095] 因此,本发明的实施例允许以简单且准确的方式估计区段性SNR并选择适当编码

算法。

[0096] 在上文的实施例中,通过计算估计的各自子帧的SNR的平均而估计区段性SNR。在替代实施例中,可无需将帧分成子帧而估计整个帧的SNR。

[0097] 当与封闭回路选择相比时,本发明的实施例允许大量减少计算时间,这是由于省略封闭回路选择中所要求的数个步骤。

[0098] 因此,可通过发明方法节省大量步骤及与此相关联的计算时间,同时又允许良好地执行适当编码算法的选择。

[0099] 尽管已在装置的上下文中描述一些方面,但显然,此方面亦表示对应方法的描述,其中区块或器件对应于方法步骤或方法步骤的特征。类似地,方法步骤的上下文中所描述的方面亦表示对应区块或项目或对应装置的特征的描述。

[0100] 可由被配置为或程序化,以便提供所描述的功能性的计算机、一或多个处理器、一或多个微处理器、场可程序化门阵列(FPGA)、特殊应用集成电路(ASIC)及其类似者或其组合来实施本文中所述的装置及其特征的实施例。

[0101] 可由(或使用)硬件装置(例如,微处理器、可程序化计算机或电子电路)执行方法步骤中的一些或所有。在一些实施例中,可由此装置执行最重要的方法步骤中的某一个或多者。

[0102] 取决于某些实施要求,本发明的实施例可以硬件或软件实施。可使用其上储存有电子可读控制信号,与可程序化计算机系统协作(或能够协作),使得执行各自方法的非暂时性储存媒体(诸如,数字储存媒体(例如,软盘、DVD、Blu-Ray、CD、ROM、PROM及EPROM、EEPROM或闪存))来执行实施。因此,数字储存媒体可以是计算机可读的。

[0103] 根据本发明的一些实施例包含具有电子可读控制信号的数据载体,其能够与可程序化计算机系统协作,使得执行本文中所描述的方法中的一个。

[0104] 大体而言,本发明的实施例可实施为具有程序代码的计算机程序产品,当计算机程序产品运行于计算机上时,程序代码操作性地用于执行该方法中的一个。程序代码可(例如)储存于机器可读载体上。

[0105] 其他实施例包含储存于机器可读载体上,用于执行本文中所描述的方法中的一个的计算机程序。

[0106] 换言之,因此,发明方法的实施例为具有用于当计算机程序运行于计算机上时,执行本文中所描述的方法中的一个的程序代码的计算机程序。

[0107] 因此,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(或数字储存媒体,或计算机可读媒体)。数据载体、数字储存媒体或记录媒体通常是有形的及/或非瞬变的。

[0108] 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)被配置为经由数据通信连接(例如,经由因特网)而传送。

[0109] 另一实施例包含被配置为或程序化以执行本文中所描述的方法中的一个的处理构件,例如,计算机或可程序化逻辑器件。

[0110] 另一实施例包含其上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。

[0111] 根据本发明的另一实施例包含被配置为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送(例如,用电子方式或光学方式)至接收器的装置或系统。接收器可(例如)为计算机、行动器件、内存器件或类似者。装置或系统可(例如)包含用于将计算机程序传送至接收器的文件服务器。

[0112] 在一些实施例中,可程序化逻辑器件(例如,场可程序化门阵列)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,场可程序化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,较佳地由任何硬件装置执行该方法。

[0113] 上文所描述的实施例仅仅说明本发明的原理。应理解,熟习此项技术者将显而易见对本文中所描述的配置及细节的修改及变化。因此,仅意欲受限于接下来的申请专利范围的范畴,而不受限于通过本文中的实施例的描述及解释所呈现的特定细节。

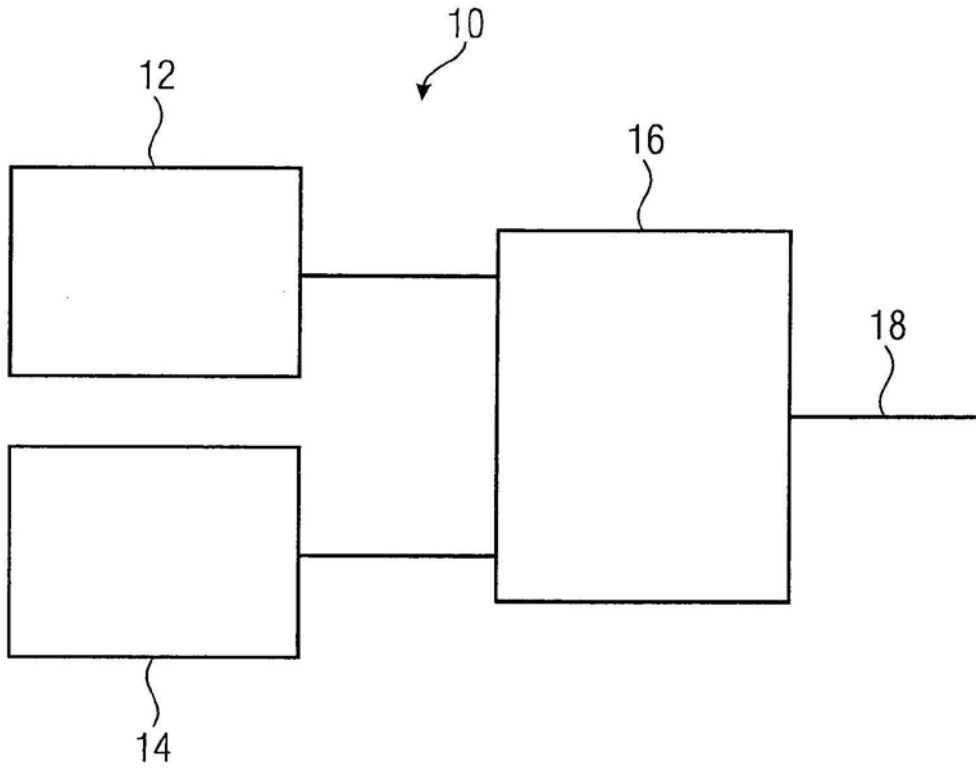


图1

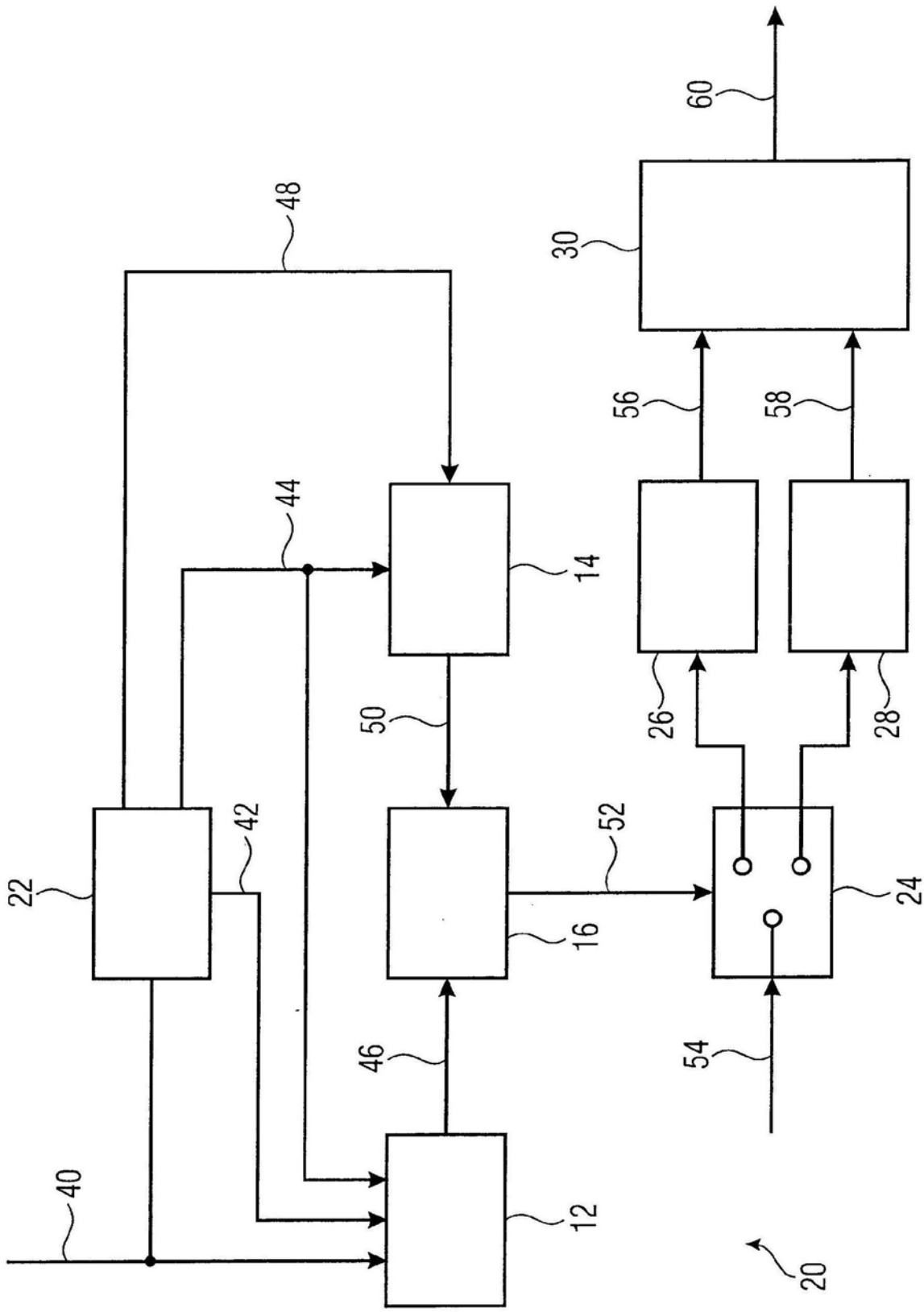


图2

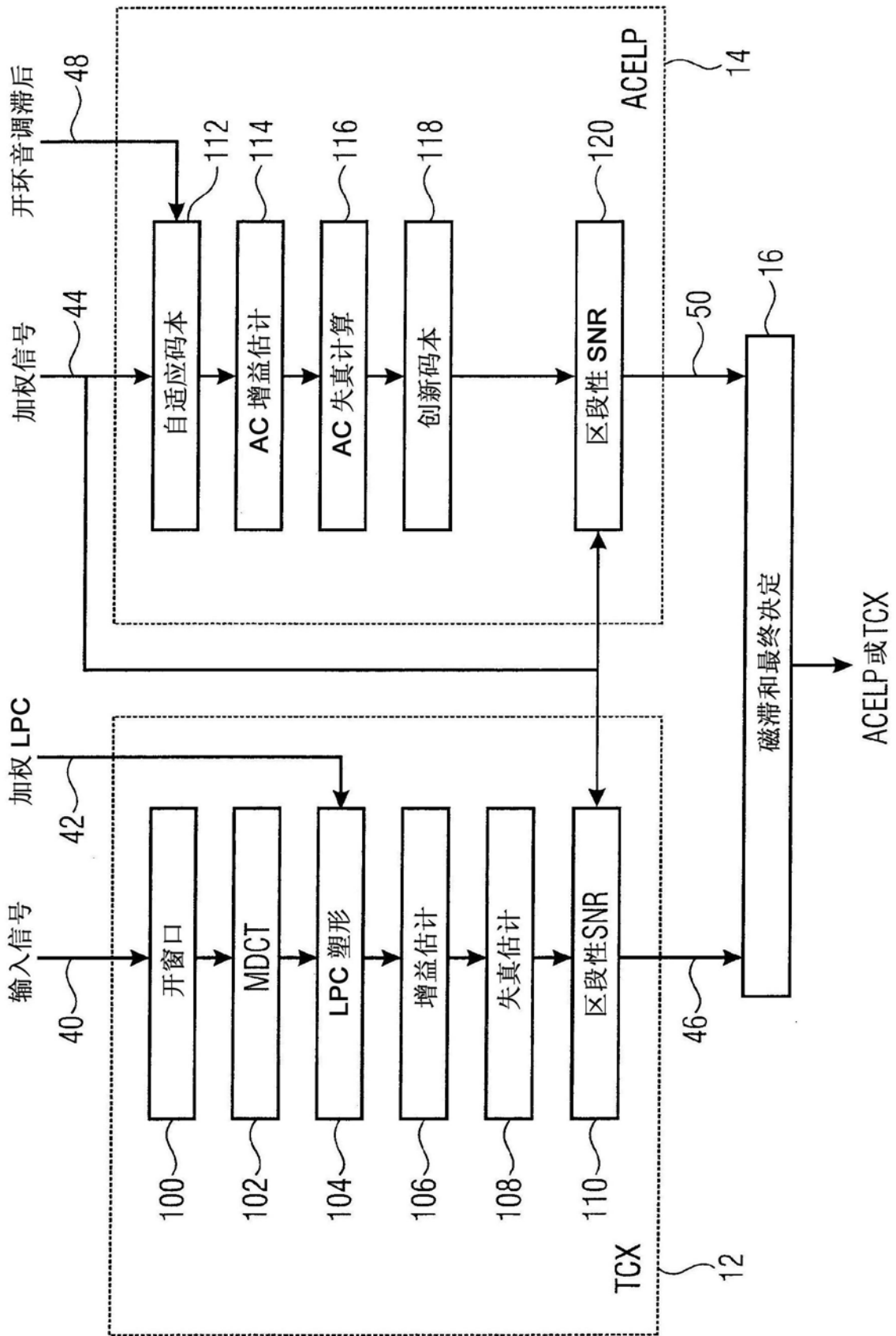


图3

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

图4A

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

图4B