



(12) 发明专利

(10) 授权公告号 CN 110992965 B

(45) 授权公告日 2024. 09. 03

(21) 申请号 201911345336.0

夫

(22) 申请日 2015.02.24

(74) 专利代理机构 北京铭硕知识产权代理有限公司 11286

(65) 同一申请的已公布的文献号  
申请公布号 CN 110992965 A

专利代理师 朱志玲 田方

(43) 申请公布日 2020.04.10

(51) Int.Cl.

(30) 优先权数据

G10L 19/02 (2013.01)

61/943,638 2014.02.24 US

G10L 19/04 (2013.01)

62/029,672 2014.07.28 US

G10L 19/18 (2013.01)

(62) 分案原申请数据

(56) 对比文件

201580021378.2 2015.02.24

CN 102089803 A, 2011.06.08

(73) 专利权人 三星电子株式会社

审查员 叶萌萌

地址 韩国京畿道水原市

(72) 发明人 朱基岬

安东·维克托罗维奇·普罗夫

康斯坦丁·谢尔盖耶维奇·奥斯波

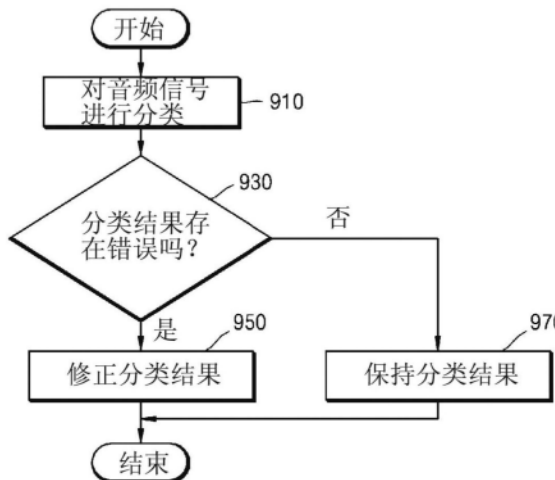
权利要求书2页 说明书10页 附图5页

(54) 发明名称

信号分类方法和装置以及使用其的音频编  
码方法和装置

(57) 摘要

公开了一种信号分类方法和装置以及使用其的音频编码方法和装置。本发明涉及了一种音频编码,更具体地,涉及一种能够减小由于编码模式切换而导致的延迟同时提高重建的声音的质量的信号分类方法和装置以及使用所述信号分类方法和装置的音频编码方法和装置。所述信号分类方法可包括以下操作:将当前帧分类为语音信号和音乐信号中的一个;基于从多个帧获得的特征参数,确定对当前帧进行分类的结果是否包含错误;并且依据确定的结果,修正对当前帧进行分类的结果。通过基于修正参数修正音频信号的初始分类结果,本发明能够确定针对音频信号的特征的最优编码模式并且能够防止帧之间的频繁的编码模式切换。



1. 一种信号分类方法,包括:  
基于多个第一信号特征,将当前帧分类为语音信号和音乐信号中的一个;  
基于多个第二信号特征估计条件,其中,所述多个第二信号特征是从多个帧获得的;  
首先确定所述条件的估计出的结果是否与第一阈值相应;  
其次确定迟滞参数是否与第二阈值相应;并且  
基于首先确定的结果和其次确定的结果,修正当前帧的分类结果,  
其中,所述多个第二信号特征包括1-2KHz范围的第一音调和2-4KHz范围的第二音调。
2. 如权利要求1所述的信号分类方法,其中,修正当前帧的分类结果的步骤是基于多个独立的状态机执行的。
3. 如权利要求2所述的信号分类方法,其中,所述多个独立的状态机包括音乐状态机和语音状态机。
4. 如权利要求1所述的信号分类方法,其中,修正当前帧的分类结果的步骤包括:当确定当前帧的分类结果指示音乐信号并且当前帧具有语音特征时,将所述分类结果修正为语音信号。
5. 如权利要求1所述的信号分类方法,其中,修正当前帧的分类结果的步骤包括:当确定当前帧的分类结果指示语音信号并且当前帧具有音乐特征时,将所述分类结果修正为音乐信号。
6. 一种音频编码方法,包括:  
基于多个第一信号特征,将当前帧分类为语音信号和音乐信号中的一个;  
基于多个第二信号特征估计条件,其中,所述多个第二信号特征是从多个帧获得的;  
首先确定所述条件的估计出的结果是否与第一阈值相应;  
其次确定迟滞参数是否与第二阈值相应;  
基于首先确定的结果和其次确定的结果,修正当前帧的分类结果;并且  
基于当前帧的分类结果或修正后的分类结果,对当前帧进行编码,  
其中,所述多个第二信号特征包括1-2KHz范围的第一音调和2-4KHz范围的第二音调。
7. 如权利要求6所述的音频编码方法,其中,对当前帧进行编码的步骤是使用CELP型编码器和变换编码器中的一个执行的。
8. 如权利要求6所述的音频编码方法,其中,对当前帧进行编码的步骤是使用CELP型编码器、变换编码器和CELP/变换混合编码器中的一个执行的。
9. 一种信号分类设备,包括至少一个处理器,其中,所述至少一个处理器被配置为:  
基于多个第一信号特征,将当前帧分类为语音信号和音乐信号中的一个;  
基于多个第二信号特征估计条件,其中,所述多个第二信号特征是从多个帧获得的;  
首先确定所述条件的估计出的结果是否与第一阈值相应;  
其次确定迟滞参数是否与第二阈值相应;并且  
基于首先确定的结果和其次确定的结果,修正当前帧的分类结果,  
其中,所述多个第二信号特征包括1-2KHz范围的第一音调和2-4KHz范围的第二音调。
10. 一种音频编码设备,包括至少一个处理器,其中,所述至少一个处理器被配置为:  
基于多个第一信号特征,将当前帧分类为语音信号和音乐信号中的一个;  
基于多个第二信号特征估计条件,其中,所述多个第二信号特征是从多个帧获得的;

首先确定所述条件的估计出的结果是否与第一阈值相应；  
其次确定迟滞参数是否与第二阈值相应；  
基于首先确定的结果和其次确定的结果,修正当前帧的分类结果;并且  
基于当前帧的分类结果或修正后的分类结果,对当前帧进行编码,  
其中,所述多个第二信号特征包括1-2KHz范围的第一音调和2-4KHz范围的第二音调。

## 信号分类方法和装置以及使用其的音频编码方法和装置

[0001] 本申请是向中国国家知识产权局递交的申请日为2015年2月24日,申请号为201580021378.2,发明名称为“信号分类方法和装置以及使用其的音频编码方法和装置”的发明专利申请的分案申请。

### 技术领域

[0002] 一个或多个示例性实施例涉及音频编码,更具体地,涉及一种能够提高恢复的声音的质量并且减少由于编码模式切换而导致的延迟的信号分类方法和设备以及一种采用所述信号分类方法和装置的音频编码方法和设备。

### 背景技术

[0003] 众所周知,音乐信号在频域中被有效地编码并且语音信号在时域中被有效地编码。因此,已经提出了对混合了音乐信号和语音信号的音频信号与音乐信号相应还是与语音信号相应进行分类并且响应于分类的结果确定编码模式的各种技术。

[0004] 然而,编码模式的频繁切换导致发生延迟和恢复的声音的质量的劣化,并且修正初始分类结果的技术还未提出,因此,当初始信号分类中存在错误时,会发生恢复的声音质量的劣化。

### 发明内容

[0005] 技术问题

[0006] 一个或多个示例性实施例包括一种能够通过确定编码模式以便适合于音频信号的特征来提高恢复的声音质量的信号分类方法和设备、以及一种采用所述信号分类方法和设备的音频编码方法和设备。

[0007] 一个或多个示例性实施例包括一种能够在确定编码模式以便适合于音频信号的特征的同时减少由于编码模式切换而导致的延迟的信号分类方法和设备、以及一种采用所述信号分类方法和设备的音频编码方法和设备。

[0008] 技术方案

[0009] 根据一个或多个示例性实施例,一种信号分类方法包括:将当前帧分类为语音信号和音乐信号中的一个;基于从多个帧获得的特征参数,确定在当前帧的分类结果中是否存在错误;并且响应于确定的结果,修正当前帧的分类结果。

[0010] 根据一个或多个示例性实施例,一种信号分类设备包括至少一个处理器,其中,所述至少一个处理器被配置为:将当前帧分类为语音信号和音乐信号中的一个,基于从多个帧获得的特征参数确定在当前帧的分类结果中是否存在错误,并且响应于确定的结果修正当前帧的分类结果。

[0011] 根据一个或多个示例性实施例,一种音频编码方法包括:将当前帧分类为语音信号和音乐信号中的一个;基于从多个帧获得的特征参数,确定在当前帧的分类结果中是否存在错误;响应于确定的结果,修正当前帧的分类结果;并且基于当前帧的分类结果或修

正后的分类结果,对当前帧进行编码。

[0012] 根据一个或更多个示例性实施例,一种音频编码装置包括至少一个处理器,其中,所述至少一个处理器被配置为:将当前帧分类为语音信号和音乐信号中的一个,基于从多个帧获得的特征参数确定在当前帧的分类结果中是否存在错误,响应于确定的结果修正当前帧的分类结果,并且基于当前帧的分类结果或修正后的分类结果对当前帧进行编码。

[0013] 发明的有益效果

[0014] 通过基于修正参数来修正音频信号的初始分类结果,可在确定对于音频信号的特征最优的编码模式的同时防止编码模式的频繁切换。

## 附图说明

[0015] 图1是根据示例性实施例的音频信号分类设备的框图;

[0016] 图2是根据另一示例性实施例的音频信号分类设备的框图;

[0017] 图3是根据示例性实施例的音频编码设备的框图;

[0018] 图4是用于描述根据示例性实施例的在CELP核中修正信号分类的方法的流程图;

[0019] 图5是用于描述根据示例性实施例的在HQ核中修正信号分类的方法的流程图;

[0020] 图6示出根据示例性实施例的用于在CELP核中修正基于上下文的信号分类的状态机;

[0021] 图7示出根据示例性实施例的用于在HQ核中修正基于上下文的信号分类的状态机;

[0022] 图8是根据示例性实施例的编码模式确定设备的框图;

[0023] 图9是用于描述根据示例性实施例的音频信号分类方法的流程图;

[0024] 图10是根据示例性实施例的多媒体装置的框图;

[0025] 图11是根据另一示例性实施例的多媒体装置的框图;

## 具体实施方式

[0026] 以下,将参照附图详细地描述本发明的一方面。在下面的描述中,当确定相关的公知的公式或功能的详细描述可能模糊要点时,省略详细的描述。

[0027] 当描述特定元件“连接”或“链接”到另一元件时,应理解,特定元件可直接地或经由中间元件连接或链接到另一元件。

[0028] 尽管诸如‘第一’和‘第二’的术语可被用于描述各种元件,但是元件不能被所述术语限制。所述术语可被用于将特定元件与另一元件进行分类。

[0029] 在实施例中出现的组件被单独地示出以表示不同特征的功能,但是并不表示每一组件在分离的硬件或单个软件配置单元中被形成。为了方便描述,组件被示出为单独的组件,并且一个组件可通过组合两个组件被形成或者一个组件可被分离成多个组件以执行功能。

[0030] 图1是示出根据示例性实施例的音频信号分类设备的配置的框图。

[0031] 图1所示的音频信号分类设备100可包括信号分类器110和修正器130。这里,除了需要被实现为单独的硬件部件的情况外,组件可被整合成至少一个模块并且被实现为至少一个处理器(未示出)。此外,音频信号可指示音乐信号、语音信号或音乐和语音的混合信

号。

[0032] 参照图1,基于各种初始分类参数,信号分类器110可对音频信号与音乐信号相应还是与语音信号相应进行分类。音频信号分类处理可包括至少一个操作。根据实施例,音频信号可基于当前帧和多个先前帧的信号特征被分类为音乐信号或语音信号。信号特征可包括短期特征和长期特征中的至少一个。此外,信号特征可包括时域特征和频域特征中的至少一个。这里,如果音频信号被分类为语音信号,则音频信号可使用码激励线性预测(CELP)型编码器被编码。如果音频信号被分类为音乐信号,则音频信号可使用变换编码器被编码。例如,变换编码器可以是例如改进的离散余弦变换(MDCT)编码器,但不限制于此。

[0033] 根据另一示例性实施例,音频信号分类处理可包括根据音频信号是否具有语音特征将音频信号分类为语音信号和通用音频信号(即,音乐信号)的第一操作以及确定通用音频信号是否适用于通用信号音频编码器(GSC)的第二操作。可通过结合第一操作的分类结果和第二操作的分类结果来确定音频信号可被分类为语音信号还是音乐信号。当音频信号被分类为语音信号时,可由CELP型编码器对音频信号进行编码。根据比特率或信号特征,CELP型编码器可包括清音编码(UC)模式、浊音编码(VC)模式、瞬态编码(TC)模式和通用编码(GC)模式中的多个模式。通用信号音频编码(GSC)模式可由单独的编码器实现或者被包括作为CELP型编码器的一个模式。当音频信号被分类为音乐信号时,可使用变换编码器或CELP/变换混合编码器对音频信号进行编码。具体地,变换编码器可被应用于音乐信号,CELP/变换混合编码器可被应用于非音乐信号(其中,非音乐信号不是语音信号)或者混合了音乐和语音的信号。根据实施例,根据带宽,可使用CELP型编码器、CELP/变换混合编码器和变换编码器中的所有,或者可使用CELP型编码器和变换编码器。例如,CELP型编码器和变换编码器可被用于窄带(NB),CELP型编码器、CELP/变换混合编码器和变换编码器可被用于宽带(WB)、超宽带(SWB)和全频带(FB)。CELP/变换混合编码器通过将时域中操作的基于LP的编码器和变换域编码器组合被获得,并且也可被称为通用信号音频编码器(GSC)。

[0034] 第一操作的信号分类可基于高斯混合模型(GMM)。各种信号特征可被用于GMM。信号特征的示例可包括开环基音、归一化相关、频谱包络、音调稳定性、信号非平稳性、LP残留误差、频谱差值和频谱平稳性,但不限制于此。用于第二操作的信号分类的信号特征的示例可包括频谱能量变化特征、LP分析残留能量的倾斜特征、高频带频谱峰度特征、相关特征、浊化特征和音调特征,但不限制于此。用于第一操作的特征可被用于确定音频信号具有语音特征还是非语音特征,以便确定CELP型编码器是否适用于进行编码,用于第二操作的特征可被用于确定音频信号具有音乐特征还是非音乐特征,以便确定GSC是否适用于进行编码。例如,在第一操作中被分类为音乐信号的一组帧可在第二操作中被改变为语音信号并随后通过CELP模式中的一个模式被编码。也就是说,当音频信号是大相关性的信号或攻击信号同时具有大基音周期和高稳定性时,音频信号可在第二操作中被从音乐信号改变到语音信号。编码模式可根据上述信号分类的结果被改变。

[0035] 修正器130可基于至少一个修正参数修正或保持信号分类器110的分类结果。修正器130可基于上下文修正或保持信号分类器110的分类结果。例如,当当前帧被分类为语音信号时,当前帧可被修正为音乐信号或被保持为语音信号,当当前帧被分类为音乐信号时,当前帧可被修正为语音信号或被保持为音乐信号。为了确定当前帧的分类结果中是否存在错误,包括当前帧的多个帧的特征被使用。例如,八个帧被使用,但实施例不限制于此。

[0036] 修正参数可包括音调、线性预测误差、浊化和相关性中的至少一个的结合。这里，音调可包括1-2KHz范围的音调 $ton_2$ 和2-4KHz范围的音调 $ton_3$ ，其中，音调 $ton_2$ 和音调 $ton_3$ 可分别由公式1和2确定。

$$[0037] \quad ton_2 = 0.2 * \log_{10} \left[ \sqrt{\frac{1}{8} \sum_{i=0}^7 \{tonality2^{[-i]}\}^2} \right] \quad (1)$$

$$[0038] \quad ton_3 = 0.2 * \log_{10} \left[ \sqrt{\frac{1}{8} \sum_{i=0}^7 \{tonality3^{[-i]}\}^2} \right] \quad (2)$$

[0039] 其中，上标 $[-j]$ 表示先前帧。例如， $tonality2^{[-1]}$ 表示一帧先前帧的1-2KHz范围的音调。

[0040] 低频带长期音调 $ton_{LT}$ 可被定义为 $ton_{LT} = 0.2 * \log_{10} [lt\_tonality]$ 。这里， $lt\_tonality$ 可表示全频带长期音调。

[0041] 在第 $n$ 帧的1-2KHz范围内的音调 $ton_2$ 和2-4KHz范围的音调 $ton_3$ 之间的差值 $d_{ft}$ 可被定义为 $d_{ft} = 0.2 * \{\log_{10}(tonality2(n)) - \log_{10}(tonality3(n))\}$ 。

[0042] 下面，线性预测误差 $LP_{err}$ 可由公式3定义。

$$[0043] \quad LP_{err} = \sqrt{\frac{1}{8} \sum_{i=0}^7 [FV_s^{[-i]}(9)]^2} \quad (3)$$

[0044] 其中， $FV_s(9)$ 被定义为 $FV_s(i) = sfa_i FV_i + sfb_i$  ( $i=0, \dots, 11$ ) 并且对应于通过缩放用于信号分类器110或210的特征参数中的由公式4定义的LP残留对数能量比特征参数而获得的值。此外， $sfa_i$ 和 $sfb_i$ 可根据特征参数的类型和带宽而改变，并且用于将每一特征参数近似到 $[0;1]$ 的范围。

$$[0045] \quad FV_9 = \log\left(\frac{E(13)}{E(1)}\right) + \log\left(\frac{E^{[-1]}(13)}{E^{[-1]}(1)}\right) \quad (4)$$

[0046] 其中， $E(1)$ 表示第一LP系数的能量，并且 $E(13)$ 表示第十三LP系数的能量。

[0047] 下面，值 $FV_s(1)$ 与值 $FV_s(7)$ 之间的差值 $d_{vcor}$ 可被定义为 $d_{vcor} = \max(FV_s(1) - FV_s(7), 0)$ ，其中，值 $FV_s(1)$ 是通过基于 $FV_s(i) = sfa_i FV_i + sfb_i$  ( $i=0, \dots, 11$ ) 缩放用于信号分类器110或210的特征参数中的归一化相关特征或由公式5定义的浊化特征 $FV_1$ 而获得的，值 $FV_s(7)$ 是基于 $FV_s(i) = sfa_i FV_i + sfb_i$  ( $i=0, \dots, 11$ ) 通过缩放由公式6定义的相关图特征 $FV(7)$ 而获得的。

$$[0048] \quad FV_1 = C_{norm}^{[.]} \quad (5)$$

[0049] 其中， $C_{norm}^{[.]}$ 表示在第一或第二半帧中的归一化相关。

$$[0050] \quad FV_7 = \sum_{j=0}^{127} M_{cor}(j) + \sum_{j=0}^{127} M_{cor}^{[-1]}(j) \quad (6)$$

[0051] 其中,  $M_{cor}$  表示一帧的相关图。

[0052] 包括条件1至4中的至少一个的修正参数可单独地或组合地使用多个特征参数来生成。这里, 条件1和条件2可指示语音状态SPEECH\_STATE可被改变所依据的条件, 条件3和4可指示音乐状态MUSIC\_STATE可被改变所依据的条件。具体地, 条件1使语音状态SPEECH\_STATE能够被从0改变到1, 条件2使语音状态SPEECH\_STATE能够被从1改变到0。此外, 条件3使音乐状态MUSIC\_STATE能够被从0改变到1, 条件4使音乐状态能够被从1改变到0。1的语音状态SPEECH\_STATE可指示语音概率高(也就是说, CELP型编码是适合的), 0的语音状态SPEECH\_STATE可指示非语音概率高。1的音乐状态MUSIC\_STATE可指示变换编码是适合的, 0的音乐状态MUSIC\_STATE可指示CELP/变换混合编码(即, GSC)是适合的。作为另一示例, 1的音乐状态MUSIC\_STATE可指示变换编码是适合的, 0的音乐状态MUSIC\_STATE可指示CELP型编码是适合的。

[0053] 例如, 条件1( $f_A$ )可被定义如下。也就是说, 当 $d_{vcor} > 0.4$ 且 $d_{ft} < 0.1$ 且 $FV_s(1) > (2 * FV_s(7) + 0.12)$ 且 $ton_2 < d_{vcor}$ 且 $ton_3 < d_{vcor}$ AND $ton_{LT} < d_{vcor}$ 且 $FV_s(7) < d_{vcor}$ 且 $FV_s(1) > d_{vcor}$ 且 $FV_s(1) > 0.76$ 时,  $f_A$ 可被设置为1。

[0054] 例如, 条件2( $f_B$ )可被定义如下。也就是说, 当 $d_{vcor} < 0.4$ 时,  $f_B$ 可被设置为1。

[0055] 例如, 条件3( $f_C$ )可被定义如下。也就是说, 当 $0.26 < ton_2 < 0.54$ 且 $ton_3 > 0.22$ 且 $0.26 < ton_{LT} < 0.54$ 且 $LP_{err} > 0.5$ 时,  $f_C$ 可被设置为1。

[0056] 例如, 条件4( $f_D$ )可被定义如下。也就是说, 当 $ton_2 < 0.34$ 且 $ton_3 < 0.26$ 且 $0.26 < ton_{LT} < 0.45$ 时,  $f_D$ 可被设置为1。

[0057] 用于生成每个条件的特征或特征集不限于此。此外, 每个常数值仅是示例性的并且可根据实现方法被设置为最优值。

[0058] 具体地, 修正器130可通过使用两个独立的状态机(例如, 语音状态机和音乐状态机)修正初始分类结果中的错误。每个状态机具有两个状态, 并且可在每一状态中使用迟滞来防止频繁切换。例如, 迟滞可包括六帧。当语音状态机中的迟滞变量由 $hang_{sp}$ 表示并且音乐状态机中的迟滞变量由 $hang_{mus}$ 表示时, 如果分类结果在给定状态下被改变, 则每个变量被初始化为6, 并且此后, 针对后续每个帧, 迟滞减少1。在每一个状态机中, 可使用通过组合从音频信号提取的至少一个特征而产生的修正参数。

[0059] 图2是示出根据另一实施例的音频信号分类设备的配置的框图。

[0060] 图2所示的音频信号分类设备200可包括信号分类器210、修正器230和精细分类器250。图2的音频信号分类设备200与图1的音频信号分类设备100的不同在于图2的音频信号分类设备200还包括精细分类器250, 而信号分类器210和修正器230的功能与参照图1所描述的相同, 因此信号分类器210和修正器230的功能的详细描述被省略。

[0061] 参照图2, 精细分类器250可基于精细分类参数对由修正器230修正或保持的分类结果进行精细分类。根据实施例, 精细分类器250将通过确定是否音频信号适合于由CELP/变换混合编码器(例如, GSC)进行编码, 修正被分类为音乐信号的音频信号。在这种情况下, 作为修正方法, 特定参数或标志被改变以不选择变换编码器。当从修正器230输出的分类结

果指示音乐信号时,精细分类器250可再次执行精细分类来对音频信号是音乐信号还是语音信号进行分类。当精细分类器250的分类结果指示音乐信号时,变换编码器也可被用于在第二编码模式下对音频信号进行编码,当精细分类器250的分类结果指示语音信号时,可使用CELP/变换混合编码器在第三编码模式下对音频信号进行编码。当从修正器230输出的分类结果指示语音信号时,可使用CELP型编码器在第一编码模式下对音频信号进行编码。例如,精细分类参数可包括诸如音调、浊化、相关性、基音增益和基音差的特征但不限制于此。

[0062] 图3是示出根据实施例的音频编码设备的配置的框图。

[0063] 图3所示的音频编码设备300可包括编码模式确定器310和编码模块330。编码模式确定器310可包括图1的音频信号分类设备100或图2的音频信号分类设备200的组件。编码模块330可包括第一至第三编码器331、333和335。这里,第一编码器331可对应于CELP型编码器,第二编码器333可对应于CELP/变换混合编码器,第三编码器335可对应于变换编码器。当GSC被实现为CELP型编码器的一个模式时,编码模块330可包括第一编码器331和第三编码器335。根据比特率或带宽,编码模块330和第一编码器331可具有各种配置。

[0064] 参照图3,编码模式确定器310可基于信号特征对音频信号是音乐信号还是语音信号进行分类,并且响应于分类结果确定编码模式。编码模式可在超帧单元、帧单元或带单元中被执行。可选地,编码模式可在多个超帧组、多个帧组或多个带组的单元中被执行。这里,编码模式的示例可包括变化域模式和线性预测域模式这两种类型,但不限于此。线性预测域模式可包括UC、VC、TC和GC模式。GSC模式可被分类为单独的编码模式或被包括在线性预测域模式的子模式中。当处理器的性能、处理速度等被支持并且由于编码模式切换而导致的延迟可被解决时,编码模式还可被细分,并且编码方案也可响应于编码模式被细分。具体地,编码模式确定器310可基于初始分类参数将音频信号分类为音乐信号和语音信号中的一个。基于修正参数,编码模式确定器310可将作为音乐信号的分类结果修正为语音信号或保持音乐信号,或者将作为语音信号的分类结果修正为音乐信号或保持语音信号。编码模式确定器310可基于精细分类参数将修正的或保持的分类结果(例如,作为音乐信号的分类结果)分类为音乐信号和语音信号中的一个。编码模式确定器310可通过使用最终分类结果确定编码模式。根据实施例,编码模式确定器310可基于比特率和带宽中的至少一个确定编码模式。

[0065] 在编码模块330中,当修正器130或230的分类结果对应于语音信号时,第一编码器331可进行操作。当修正器130的分类结果对应于音乐信号时,或者当精细分类器350的分类结果对应于语音信号时,第二编码器333可进行操作。当修正器130的分类结果对应于音乐信号时,或者当精细分类器350的分类结果对应于音乐信号时,第三编码器335可进行操作。

[0066] 图4是用于描述根据实施例的在CELP核中修正信号分类的方法的流程图,并且可由图1的修正器130或图2的修正器230执行。

[0067] 参数图4,在操作410中,可接收修正参数,例如,条件1和条件2。此外,在操作410中,可接收语音状态机的迟滞信息。在操作410中,还可接收初始分类结果。初始分类结果可从图1的信号分类器110或图2的信号分类器210提供。

[0068] 在操作420中,可确定是否初始分类结果(即,语音状态)是0、条件1( $f_A$ )是1并且语音状态机的迟滞 $hang_{sp}$ 是0。如果在操作420中确定初始分类结果(即,语音状态)是0、条件1是1并且语音状态机的迟滞 $hang_{sp}$ 是0,则在操作430中,语音状态可被改变为1并且迟滞可被

初始化为6。初始化后的迟滞值可被提供给操作460。否则,如果在操作420中语音状态不是0、条件1不是1或者语音状态机的迟滞 $hang_{sp}$ 不是0,则所述方法可继续到操作440。

[0069] 在操作440中,可确定是否初始分类结果(即,语音状态)是1、条件2( $f_b$ )是1并且语音状态机的迟滞 $hang_{sp}$ 是0。如果在操作440中确定语音状态是1、条件2是1并且语音状态机的迟滞 $hang_{sp}$ 是0,则在操作450中,语音状态可被改变为0并且 $hangover_{sp}$ 可被初始化为6。初始化后的迟滞值可被提供给操作460。否则,如果在操作440中语音状态不是1、条件2不是1或者语音状态机的迟滞 $hang_{sp}$ 不是0,则所述方法可继续到操作460,以执行用于将迟滞减小1的迟滞更新。

[0070] 图5是用于描述根据实施例的在高质量(HQ)核中修正信号分类的方法的流程图,并且可由图1的修正器130或图2的修正器230执行。

[0071] 参照图5,在操作510中,可接收修正参数,例如,条件3和条件4。此外,在操作510中,可接收音乐状态机的迟滞信息。在操作510中,还可接收初始分类结果。初始分类结果可从图1的信号分类器110或图2的信号分类器210提供。

[0072] 在操作520中,可确定是否初始分类结果(即,音乐状态)是1、条件3( $f_c$ )是1并且音乐状态机的迟滞 $hang_{mus}$ 是0。如果在操作520中确定初始分类结果(即,音乐状态)是1、条件3是1并且音乐状态机的迟滞 $hang_{mus}$ 是0,则在操作530中,音乐状态可被改变为0并且迟滞可被初始化为6。初始化后的迟滞值可被提供给操作560。否则,如果在操作520中音乐状态不是1、条件3不是1或者音乐状态机的迟滞 $hang_{mus}$ 不是0,则所述方法可继续到操作540。

[0073] 在操作540中,可确定是否初始分类结果(即,音乐状态)是0、条件4( $f_d$ )是1并且音乐状态机的迟滞 $hang_{sp}$ 是0。如果在操作540中确定音乐状态是0、条件4是1并且音乐状态机的迟滞 $hang_{mus}$ 是0,则在操作550中,音乐状态可被改变为1并且迟滞 $hang_{mus}$ 可被初始化为6。初始化后的迟滞值可被提供到操作560。否则,如果在操作540中语音状态不是0、条件4不是1或者音乐状态机的迟滞 $hang_{mus}$ 不是0,则所述方法可继续到操作560,以执行用于将迟滞减少1的迟滞更新。

[0074] 图6示出根据实施例的用于在适用于CELP核的状态(即,语音状态下)下修正基于上下文的信号分类的状态机,并且可与图4对应。

[0075] 参照图6,在修正器(图1的130或230)中,可根据由音乐状态机确定的音乐状态和由语音状态机确定的语音状态来应用对分类结果的修正。例如,当初始分类结果被设置为音乐信号时,音乐信号可基于修正参数被改变为语音信号。具体地,当初始分类结果的第一操作的分类结果指示音乐信号并且语音状态是1时,第一操作的分类结果和第二操作的分类结果两者可被改变为语音信号。在这种情况下,可以确定在初始分类结果中存在错误,从而修正分类结果。

[0076] 图7示出根据实施例的在针对高质量(HQ)核的状态(例如,音乐状态)下修正基于上下文的信号分类的状态机,并且可与图5对应。

[0077] 参照图7,在修正器(图1的130或230)中,可根据由音乐状态机确定的音乐状态和由语音状态机确定的语音状态应用对分类结果的修正。例如,当初始分类结果被设置为语音信号时,语音信号可基于修正参数被改变为音乐信号。具体地,当初始分类结果的第一操作的分类结果指示语音信号并且音乐状态是1时,第一操作的分类结果和第二操作的分类结果两者可被改变为音乐信号。当初始分类结果被设置为音乐信号时,音乐信号可基于修

正参数被改变为语音信号。在这种情况下,可以确定在初始分类结果中存在错误,从而修正分类结果。

[0078] 图8是示出根据实施例的编码模式确定设备的配置的框图。

[0079] 图8所示的编码模式确定设备可包括初始编码模式确定器810和修正器830。

[0080] 参照图8,初始编码模式确定器810可确定音频信号是否具有语音特征,并且当音频信号具有语音特征时可将第一编码模式确定为初始编码模式。在第一编码模式下,可由CELP型编码器对音频信号进行编码。当音频信号具有非语音特征时,初始编码模式确定器810可将第二编码模式确定为初始编码模式。在第二编码模式下,可由变换编码器对音频信号进行编码。可选地,当音频信号具有非语音特征时,初始编码模式确定器810可根据比特率将第二编码模式和第三编码模式中的一个确定为初始编码模式。在第三编码模式下,可由CELP/变换混合编码器对音频信号进行编码。根据实施例,初始编码模式确定器810可使用三种方式的方案。

[0081] 当初始编码模式被确定为第一编码模式时,修正器830可基于修正参数将初始编码模式修正为第二编码模式。例如,当初始分类结果指示语音信号但是具有音乐特征时,初始分类结果可被修正为音乐信号。当初始编码模式被确定为第二编码模式时,修正器830可基于修正参数将初始编码模式修正为第一编码模式或第三编码模式。例如,当初始分类结果指示音乐信号但是具有语音特征时,初始分类结果可被修正为语音信号。

[0082] 图9是用于描述根据实施例的音频信号分类方法的流程图。

[0083] 参照图9,在操作910中,音频信号可被分类为音乐信号和语音信号中的一个。具体地,在操作910中,可基于信号特征对当前帧对应于音乐信号还是语音信号进行分类。操作910可由图1的信号分类器110或图2的信号分类器210执行。

[0084] 在操作930中,可基于修正参数确定操作910的分类结果中是否存在错误。如果在操作930中确定分类结果中存在错误,则可在操作950中修正分类结果。如果在操作930中确定分类结果中不存在错误,则在操作970中保持分类结果不变。操作930至操作970可由图1的修正器130或图2的修正器230执行。

[0085] 图10是示出根据实施例的多媒体装置的配置的框图。

[0086] 图10所示的多媒体装置1000可包括通信单元1010和编码模块1030。此外,根据音频比特流的使用,还可包括用于存储获得的作为编码结果的音频比特流的存储单元1050。此外,多媒体装置1000还可包括麦克风1070。也就是说,存储单元1050和麦克风1070可选择性地被提供。根据示例性实施例,图10所示的多媒体装置1000还可包括任意解码装置(未示出),例如,用于执行通用解码功能的解码模块或根据示例性实施例的解码模块。这里,编码模块1030可与提供给多媒体装置1000的其他组件(未示出)集成并且被实现为至少一个处理器(未示出)。

[0087] 参照图10,通信单元1010可接收从外部提供的音频和编码比特流中的至少一个,或者发送重建的音频和获得的作为编码模块1030的编码结果的音频比特流中的至少一个。

[0088] 通信单元1010被配置为能够通过无线网络(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网络、Wi-Fi直连(WFD)网络、第三代(3G)网络、4G网络、蓝牙网络、红外数据协会(IrDA)网络、射频识别(RFID)网络、超宽带(UWB)网络、ZigBee网络和近场通信(NFC)网络)或者有线网络(诸如有线电话网络或有线互联网)向外部多媒体装

置或服务器发送数据或从外部多媒体装置或服务器接收数据。

[0089] 根据实施例,编码模块1030可对时域的音频信号进行编码,其中,时域的音频信号通过通信单元1010或麦克风1070被提供。可使用图1至图9所示的设备或方法实现编码处理。

[0090] 存储单元1050可存储操作多媒体装置1000所需的各种程序。

[0091] 麦克风1070可向编码模块1030提供用户或外部的音频信号。

[0092] 图11是示出根据另一实施例的多媒体装置的配置的框图。

[0093] 图11所示的多媒体装置1100可包括通信单元1110、编码模块1120和解码模块1130。此外,根据音频比特流或重建的音频信号的使用,还可包括用于存储获得的作为编码结果的音频比特流或获得的作为解码结果的重建的音频信号的存储单元1140。此外,多媒体装置1100还可包括麦克风1150或扬声器1160。这里,编码模块1120和解码模块1130可与提供给多媒体装置1100的其他组件(未示出)集成,并被实现为至少一个处理器(未示出)。

[0094] 省略了图11中示出的组件之中的与图10中所示的多媒体装置1000中的组件相同的组件的详细描述。

[0095] 解码模块1130可接收通过通信单元1110提供的比特流并且对包括在比特流中的音频频谱进行解码。解码模块1130可与图3的编码模块330相应地被实现。

[0096] 扬声器1170可向外部输出由解码模块1130生成的重建的音频信号。

[0097] 图10所示的多媒体装置1000和图11所示的多媒体装置1100可包括语音通信专用终端(包括电话或移动电话)、广播或音乐专用装置(包括TV或MP3播放器)或语音通信专用终端和广播或音乐专用装置的混合终端装置,但不限于此。此外,多媒体装置1000或1100可被用作布置在客户端中、服务器中或者在客户端和服务器之间的换能器。

[0098] 例如,当多媒体装置1000或1100是移动电话时,尽管未示出,但还可包括用户输入单元(诸如,键盘)、用于显示用户界面或由移动电话处理的信息的显示单元以及用于控制移动电话的通用功能的处理器。此外,移动电话还可包括具有图像拾取功能的相机单元以及用于执行由移动电话所需的功能的至少一个组件。

[0099] 例如,当多媒体装置1000或1100是TV时,尽管未示出,但还可包括用户输入单元(诸如键盘)、用于显示接收到的广播信息的显示单元和用于控制TV的通用功能的处理器。此外,TV还可包括用于执行由TV所需要的功能的至少一个组件。

[0100] 根据实施例的所述方法可通过计算机可执行的程序被编辑并且通过使用计算机可读记录介质在用于执行程序通用数字计算机中被实现。此外,可在本发明的实施例中使用的数据结构、程序命令或数据文件可通过各种方式被记录在计算机可读记录介质中。计算机可读记录介质可包括用于存储可由计算机系统读取的数据的所有类型的存储装置。计算机可读记录介质的示例包括磁介质(诸如硬盘、软盘或磁带)、光学介质(诸如压缩盘只读存储器(CD-ROM)或数字通用盘(DVD))、磁光介质(诸如软光盘)和被专门配置为用于存储并实施程序命令的硬件装置(诸如ROM、RAM或闪存)。此外,计算机可读记录介质可以是用于传输用于指定程序命令、数据结构等的信号的传输介质。程序命令的示例包括可由计算机使用解释器执行的高级语言代码和由编译器生成的机器语言代码。

[0101] 尽管已参照有限的实施例和附图描述了本发明的实施例,但是本发明的实施例不限于以上描述的实施例,并且它们的更新和修改可从本公开由本领域的普通技术人员进行

各种实施。因此,本发明的范围不是由以上描述定义的而是由权利要求定义,并且它们所有的相同或等同的修改将属于本发明的技术构思的范围。

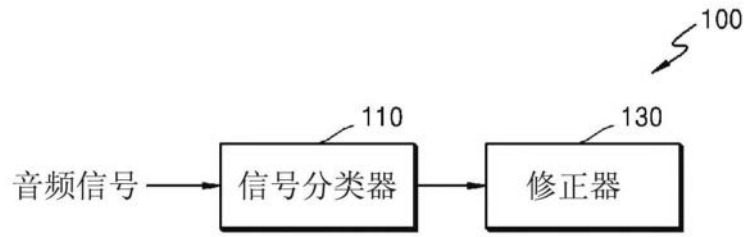


图1

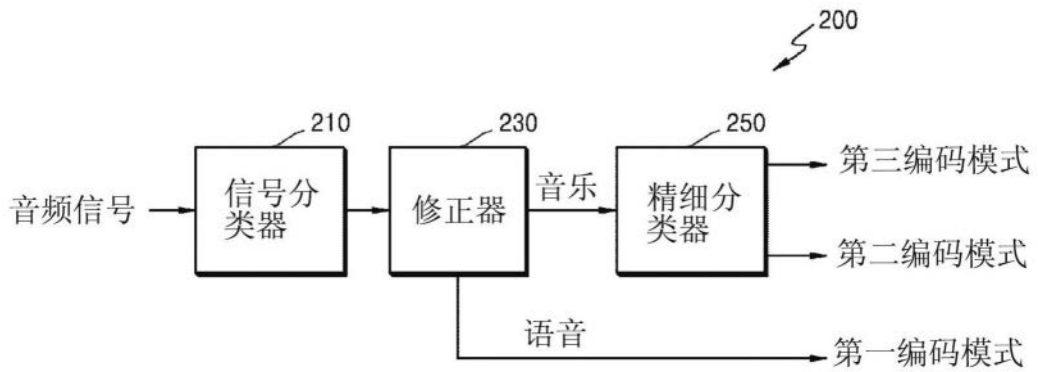


图2

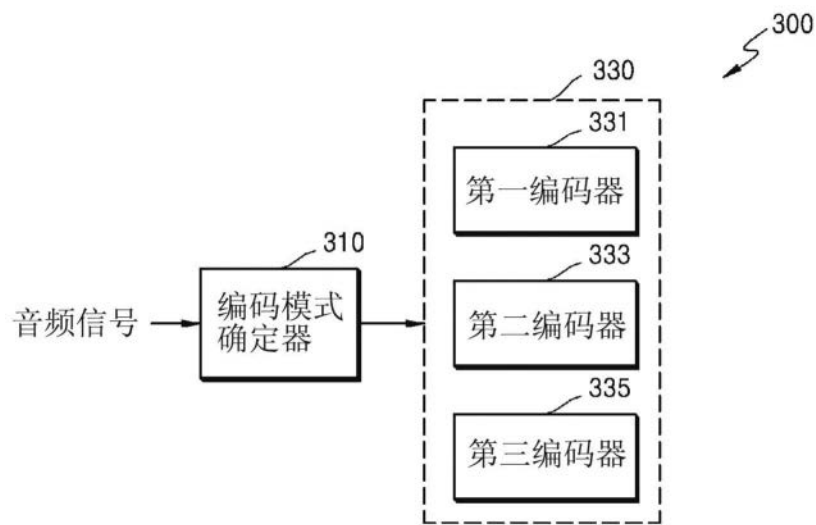


图3

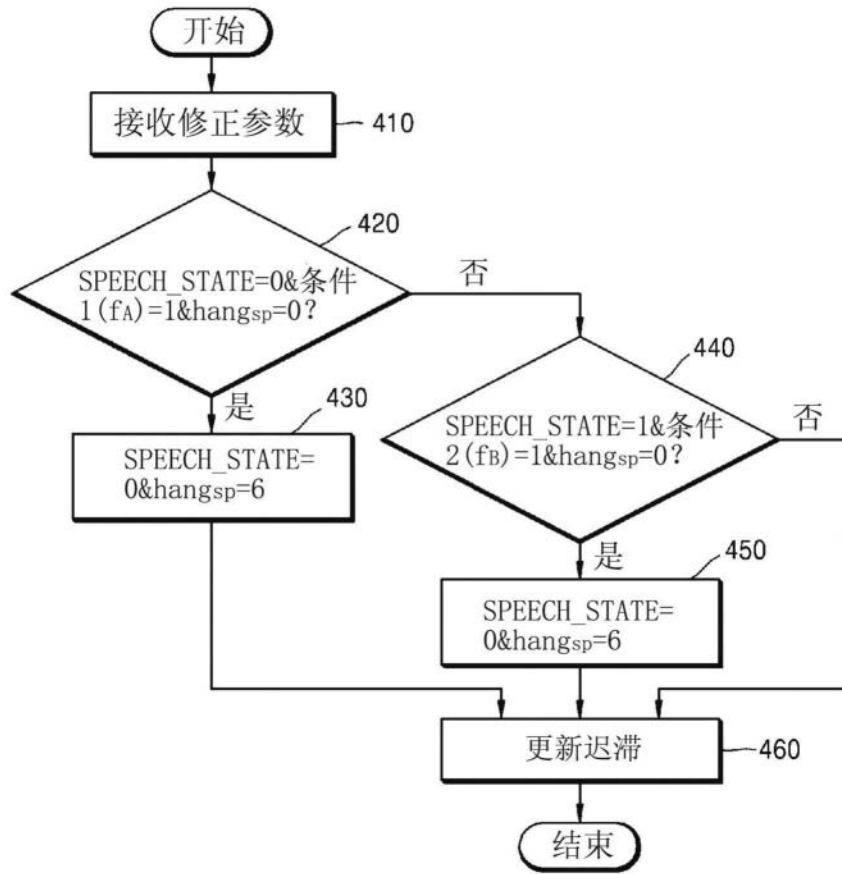


图4

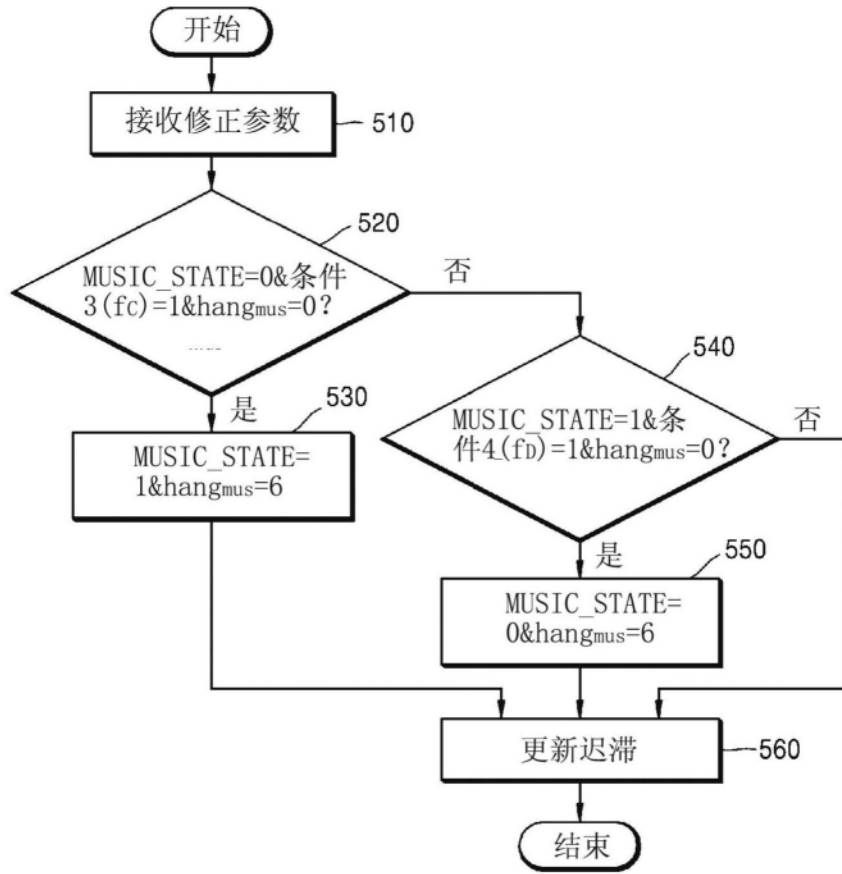


图5

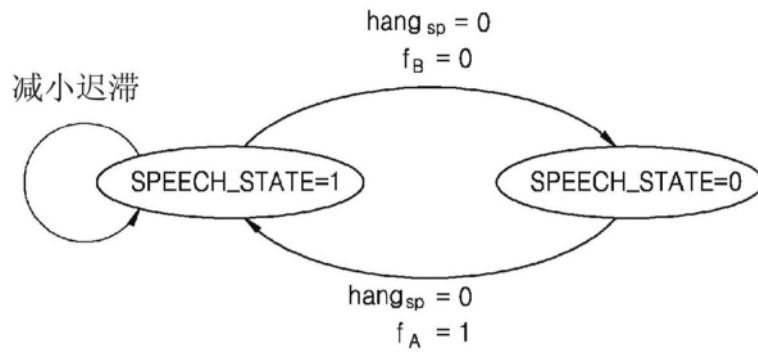


图6

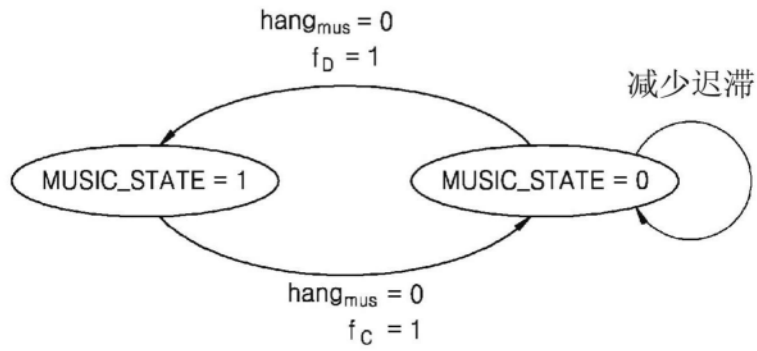


图7

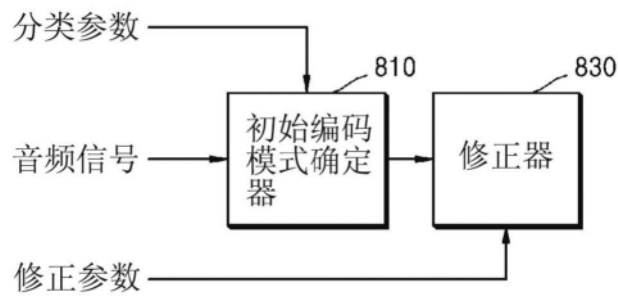


图8

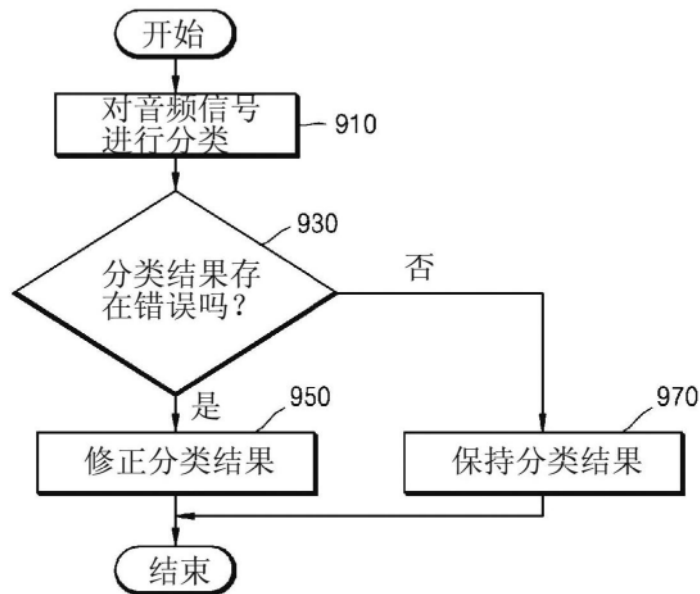


图9

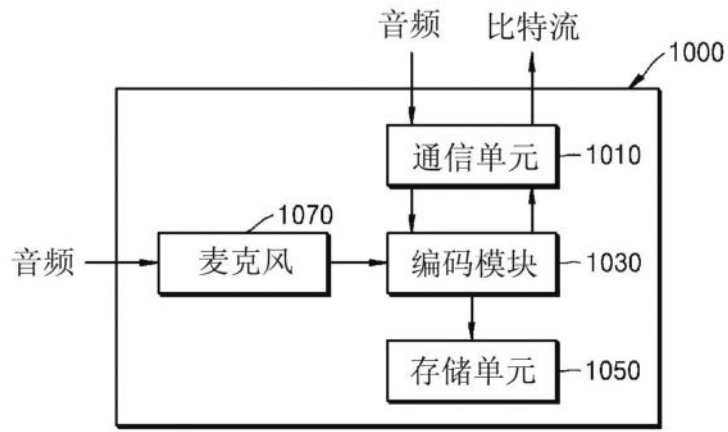


图10

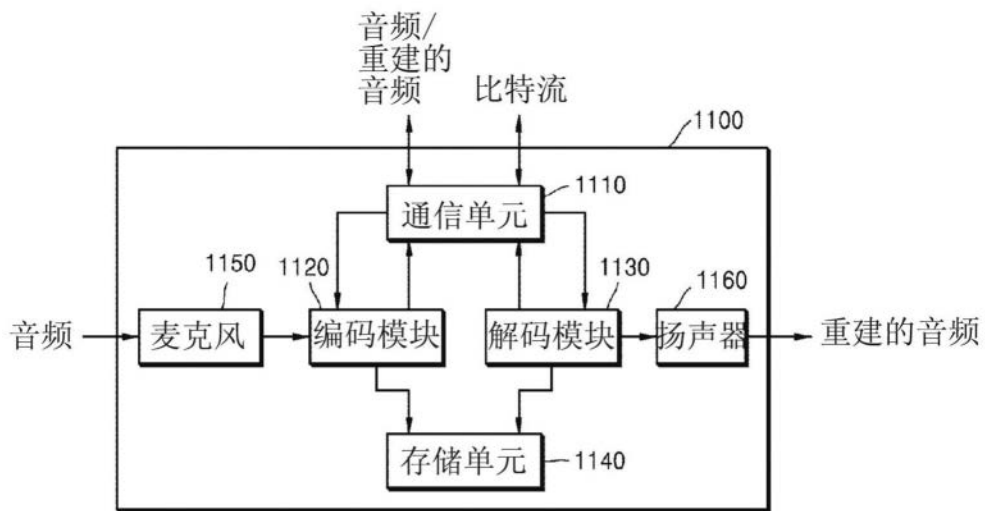


图11