



(12) 发明专利

(10) 授权公告号 CN 110619891 B

(45) 授权公告日 2023. 01. 17

(21) 申请号 201910918149.0
(22) 申请日 2015.05.07
(65) 同一申请的已公布的文献号
申请公布号 CN 110619891 A

(43) 申请公布日 2019.12.27
(30) 优先权数据
61/990,354 2014.05.08 US

(62) 分案原申请数据
201580023968.9 2015.05.07

(73) 专利权人 瑞典爱立信有限公司
地址 瑞典斯德哥尔摩

(72) 发明人 艾力克·诺维尔
沃洛佳·格兰恰诺夫

(74) 专利代理机构 中科专利商标代理有限责任
公司 11021
专利代理师 陈慧

(51) Int.Cl.
G10L 25/51 (2013.01)
G10L 19/20 (2013.01)
G10L 25/18 (2013.01)
G10L 25/81 (2013.01)

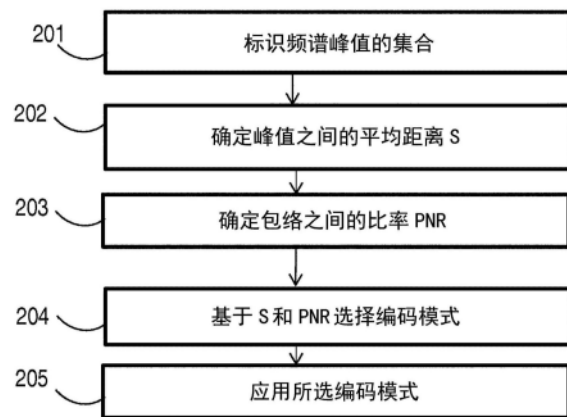
(56) 对比文件
US 2003101050 A1, 2003.05.29
US 2012158401 A1, 2012.06.21
US 2011035213 A1, 2011.02.10
CN 101399039 A, 2009.04.01
CN 102044246 A, 2011.05.04
CN 101145345 A, 2008.03.19
US 2013185063 A1, 2013.07.18

审查员 姚黎为

权利要求书1页 说明书8页 附图6页

(54) 发明名称
音频信号区分器和编码器

(57) 摘要
本发明涉及编解码器和区分器以及其中用于音频信号区分和编码的方法。编码器执行的方法的实施例包括, 针对音频信号的分段: 识别频谱峰值的集合; 确定所述集合中峰值之间的平均距离S; 以及确定峰值包络与噪声基底包络之间的比率PNR。所述方法还包括: 至少基于所述平均距离S和比率PNR, 从多个编码模式中选择编码模式; 以及应用所选编码模式, 用于所述音频信号的分段的编码。



1. 一种音频信号分类方法,所述方法包括:
针对音频信号的分段:
 - 识别频谱峰值的集合;
 - 确定所述集合中的峰值之间的平均距离S;
 - 确定峰值包络能量与噪声基底包络能量之间的比率PNR;
 - 比较所述平均距离S与第一阈值;
 - 比较所述比率PNR与第二阈值;以及
 - 基于所述平均距离S与第一阈值的比较以及所述比率PNR与第二阈值的比较,将音频信号分段分类为多个音频信号类别中的一个类别。
2. 根据权利要求1所述的方法,其中,当确定S时,每个峰值由一个频谱系数表示,所述一个频谱系数是与所述峰值相关联的频谱系数中具有最大平方幅度的频谱系数。
3. 根据权利要求1所述的方法,其中,基于频谱系数的绝对值和与低能量系数相比突出高能量系数的贡献的加权因子来估计峰值包络。
4. 根据权利要求1所述的方法,其中,基于频谱系数的绝对值和与高能量系数相比突出低能量系数的贡献的加权因子来估计噪声基底包络。
5. 一种音频信号分类器,所述音频信号分类器被配置为:
针对所述音频信号的分段:
 - 识别频谱峰值的集合;
 - 确定所述集合中的峰值之间的平均距离S;
 - 确定峰值包络能量与噪声基底包络能量之间的比率PNR;
 - 比较所述平均距离S与第一阈值;
 - 比较所述比率PNR与第二阈值;以及
 - 基于所述平均距离S与第一阈值的比较以及所述比率PNR与第二阈值的比较,将音频信号分段分类为多个音频信号类别中的一个类别。
6. 根据权利要求5所述的音频信号分类器,其中,当确定所述平均距离S时,每个峰值由一个频谱系数表示,所述一个频谱系数是与所述峰值相关联的频谱系数中具有最大平方幅度的频谱系数。
7. 根据权利要求5所述的音频信号分类器,被配置为:基于频谱系数的绝对值和与低能量系数相比突出高能量系数的贡献的加权因子来估计峰值包络。
8. 根据权利要求5所述的音频信号分类器,被配置为:基于频谱系数的绝对值和与高能量系数相比突出低能量系数的贡献的加权因子来估计噪声基底包络。
9. 一种音频编码器,包括根据权利要求5-8中任一项所述的音频信号分类器。
10. 一种通信设备,包括根据权利要求5-8中任一项所述的音频信号分类器。
11. 一种存储计算机程序的计算机可读存储介质,所述计算机程序包括指令,所述指令当在至少一个处理器上执行时使所述至少一个处理器执行根据权利要求1-4中任一项所述的方法。

音频信号区分器和编码器

[0001] 本申请是2015年5月7日提交的、申请号为201580023968.9、发明名称为“音频信号区分器和编码器”的专利申请的分案申请。

技术领域

[0002] 所提出的技术大体上涉及用于音频编码的编解码器和方法。

背景技术

[0003] 现代音频编解码器由针对具有不同属性的信号优化的多个压缩方案组成。实际上没有例外,用时域编解码器处理类语音信号,而用变换域编解码器处理音乐信号。应该处理语音和音乐信号二者的编码方案需要一种识别输入信号包括语音还是音乐的机制,并且在适当的编解码器模式之间切换。这种机制可以被称为语音音乐分类器或区分器。在图1a中示出了使用基于输入信号的模式判决逻辑的多模音频编解码器的概述说明。

[0004] 用类似的方式,在音乐信号类别中,可以从谐波音乐信号中区分出更多的类噪声音乐信号,并且为这些组中的每一个构建分类器和最佳编码方案。图1b中示出了创建分类器以确定信号的类别、该分类器然后控制模式判决的概要。

[0005] 在音频编码领域中存在各种语音音乐分类器。然而,这些分类器不能在音乐信号的空间中的不同类别之间区分。事实上,许多已知的分类器不能提供足够的分辨率以能够在复杂多模编解码器中应用所需的方式来区分音乐的类别。

发明内容

[0006] 本文中通过使用直接在频域系数上计算的新颖度量来解决例如谐波和类噪声音乐分段之间区分的问题。该度量基于预选频谱峰值候选的分布和平均峰值噪声基底比率。

[0007] 所提出的解决方案允许识别谐波和类噪声音乐分段,所提出的解决方案进而允许这些信号类型的最佳编码。该编码概念提供优于传统编码方案的优良的质量。本文描述的实施例涉及找到用于区分谐波和类噪声音乐信号的更好的分类器。

[0008] 根据第一方案,提供了一种可由音频信号编码器执行的、用于对音频信号进行编码的方法。所述方法包括,针对音频信号的分段:识别频谱峰值的集合并确定所述集合中峰值之间的平均距离 S 。所述方法还包括:确定峰值包络与噪声基底包络之间的比率 PNR ;至少基于所述平均距离 S 和所述比率 PNR ,从多个编码模式中选择编码模式;以及应用所选编码模式。

[0009] 根据第二方案,提供了一种用于对音频信号进行编码的编码器。所述编码器被配置为,针对所述音频信号的分段:识别频谱峰值的集合并确定所述集合中峰值之间的平均距离 S 。所述编码器还被配置为:确定峰值包络与噪声基底包络之间的比率 PNR ;基于所述平均距离 S 和所述比率 PNR ,从多个编码模式中选择编码模式;以及还应用所选编码模式。

[0010] 根据第三方案,提供了一种由音频信号区分器执行的、用于音频信号区分的方法。所述方法包括,针对音频信号的分段:识别频谱峰值的集合并确定所述集合中峰值之间的

平均距离 S 。所述方法还包括：确定峰值包络与噪声基底包络之间的比率PNR。所述方法还包括：至少基于所述平均距离 S 和比值PNR，从多个音频信号类别中确定所述分段属于的音频信号的类别。

[0011] 根据第四方案，提供了一种音频信号区分器。所述区分器被配置为，针对音频信号的分段：识别频谱峰值的集合；以及确定所述集合中峰值之间的平均距离 S 。所述区分器还被配置为确定峰值包络与噪声基底包络之间的比率PNR，以及还至少基于所述平均距离 S 和比值PNR，从多个音频信号类别中确定所述分段属于的音频信号的类别。

[0012] 根据第五方案，提供了一种通信设备，包括根据第二方案的编码器。

[0013] 根据第六方案，提供了一种通信设备，包括根据第四方案的音频信号区分器。

[0014] 根据第七方案，提供了一种计算机程序，包括指令，所述指令当在至少一个处理器上执行时使所述至少一个处理器执行根据第一和/或第三方案的方法。

[0015] 根据第八方案，提供了一种包含前述权利要求所述的计算机程序的载体，其中，所述载体是电信号、光信号、无线电信号或计算机可读存储介质中的一种。

附图说明

[0016] 根据以下对附图中示出的实施例更具体的描述，本文公开的技术的以上或其他目的、特征、优点将显而易见。附图不必须按比例绘制，相反重点在于说明本文公开的技术的原理。

[0017] 图1a是可以应用本发明的实施例的音频编解码器的示意图。图1b是明确示出信号分类器的音频编解码器的示意图。

[0018] 图2是示出根据示例实施例的方法的流程图。

[0019] 图3a是示出根据示例实施例的峰值选择算法和瞬时峰值和噪声基底值的图；

[0020] 图3b是示出根据示例实施例的峰值距离 d_i 的图；

[0021] 图4示出了根据示例实施例的判决的维恩图。

[0022] 图5a-c示出根据示例实施例的编码器的实施方式。

[0023] 图5d示出了根据示例实施例的区分器的实施方式。

[0024] 图6示出了编码器的实施例。

具体实施方式

[0025] 所提出的技术可以应用于例如可以是有线或无线设备的用户终端或用户设备的编码器和/或解码器。本文描述的所有备选设备和节点被概括为术语“通信设备”，在“通信设备”中可以应用本文描述的解决方案。

[0026] 如本文中所使用的，非限制性术语“用户设备”和“无线设备”可以指移动电话、蜂窝电话、配备有无线通信能力的个人数字助理PDA、智能电话、膝上型电脑或配备有内部或外部的移动宽带调制解调器的个人计算机PC，具有无线通信能力的平板PC、目标设备、设备到设备UE、机器类型的UE或支持机器到机器通信的UE、iPAD、客户住宅设备CPE、膝上型嵌入式设备LEE、膝上安装的设备LME、USB加密狗、便携式电子无线通信设备、配备有无线通信能力的传感器设备等。具体地，术语“UE”和术语“无线设备”应当理解为非限制性的，包括在蜂窝或移动通信系统中与无线网络节点通信的任意类型无线设备，或配备有用于根据蜂窝

或移动通信系统内的任意相关标准进行无线通信的无线电电路。

[0027] 如本文所使用的,术语“有线设备”可以指被配置为或准备与网络有线连接的任意设备。具体地,有线设备可以是当配置为用于有线连接时具有或不具有无线电通信能力的上述装置的至少一些。

[0028] 所提出的技术还可以应用于无线电网络节点的编码器和/或解码器。如本文所使用的,非限制性术语“无线电网络节点”可以指基站、网络控制节点,例如网络控制器、无线电网络控制器、基站控制器等。具体地,术语“基站”可以涵盖不同类型的无线电基站,其中包括标准基站(例如,节点B或演进节点B,eNB),还可以包括宏/微/微微无线基站、家庭基站(也称为毫微微基站)、中继节点、中继器、无线电接入点、基本收发机站、BTS、甚至控制一个或多个远程无线单元RRU的无线控制节点等。

[0029] 本文描述的解决方案的实施例适用于音频编解码器。因此,将在输入波形的短块(例如20ms)上进行操作的示例音频编解码器的上下文中描述该实施例。应注意,本文所描述的解决方案还可以适用于在其它块大小上操作的其它音频编解码器。此外,所提出的实施例示出针对目前的实施例优选的示例性数值。应理解,这些数值仅作为示例给出,并可以适用于目前的音频编解码器。

[0030] 示例实施例

[0031] 下面,将参照图2描述与用于对音频信号进行编码的方法相关的示例实施例。由编码器执行该方法。编码器可以被配置为符合音频编码的一个或多个标准。所述方法包括,针对音频信号的分段:识别201频谱峰值的集合;确定202所述集合中峰值之间的平均距离S;以及确定203峰值包络与噪声基底包络之间的比率PNR。该方法还包括:至少基于所述平均距离S和比率PNR,从多个编码模式中选择204编码模式;以及应用205所选编码模式。

[0032] 可以用不同的方式来识别频谱峰值,这也将下面更详细地描述。例如,其幅度超过定义阈值的频谱系数可以被识别为属于峰值。当确定峰值之间的平均距离S时,每个峰值可以由单个频谱系数表示。该单个系数将优选地是与峰值相关联的频谱系数(如果多于一个)中具有最大平方幅度的频谱系数。也就是说,当多于一个频谱系数被识别为与一个频谱峰值相关联时,则在确定平均距离S时,可以选择与峰值相关联的多个系数之一以表示峰值。这可以在图3b中看到,并且将在下面进一步描述。平均距离S也可以称为例如“峰值稀疏度”。

[0033] 为确定峰值包络和噪声基底包络之间的比率,需要估计这些包络。可以基于频谱系数的绝对值和突出(emphasize)低能量系数的贡献的加权因子来估计噪声基底包络。相应地,可以基于频谱系数的绝对值和突出高能量系数的贡献的加权因子来估计峰值包络。图3a和3b示出了估计的噪声基底包络(短虚线)和峰值包络(长虚线)的示例。“低能量”和“高能量”系数应被理解为具有与阈值有某一关系的幅度的系数,其中低能量系数将通常是具有低于(或可能等于)某一阈值的幅度的系数,并且高能量系数通常是具有高于(或可能等于)某一阈值的幅度的系数。

[0034] 根据示例实施例,在执行频谱分析之前,例如使用一阶高通滤波器 $H(z) = 1 - 0.68z^{-1}$ 对输入波形(即音频信号)进行预加重(pre-emphasized)。可以例如完成这一点,以便增加高频区域的建模精度,但应当注意,目前这对于本发明来说不是必要的。

[0035] 离散傅里叶变换(DFT)可以用于将滤波的音频信号转换至变换域或频域。在具体

示例中,使用256点快速傅里叶变换(FFT)每帧执行一次频谱分析。

[0036] 对预加重的、加窗的输入信号(即在音频信号的分段上)执行FFT,以获得一组频谱参数:

$$[0037] \quad X(k) = \sum_{n=0}^{255} x(n) e^{-j2\pi \frac{kn}{256}}$$

[0038] 其中 $k=0, \dots, 255$,是频率系数或频谱系数的指数,并且 n 是波形采样的指数。应当注意,可以使用任意长度 N 的变换。该系数也可以被称为变换系数。

[0039] 本文描述的解决方案的目的是实现其不仅可以区分语音与音乐、还可以区分不同类型的音乐的分类器或区分器。下面将更详细地描述如何根据区分器的示例实施例实现该目的:

[0040] 示例性区分器需要知道输入音频信号的分段的频谱峰值的(例如在频率上的)位置。这里,将频谱峰值定义为具有高于自适应阈值(例如基于峰值和噪声基底包络的比率)的绝对值的系数。

[0041] 可以使用对变换系数 $|X(k)|$ 的绝对值进行操作的噪声基底估计算法。可以根据递归估计瞬时噪声基底能量 $E_{nf}(k)$:

$$[0042] \quad E_{nf}(k) = \alpha E_{nf}(k-1) + (1-\alpha) |X(k)|^2$$

$$[0043] \quad \alpha = \begin{cases} 0.9578 & \text{如果 } |X(k)|^2 > E_{nf}(k-1) \\ 0.6472 & \text{如果 } |X(k)|^2 \leq E_{nf}(k-1) \end{cases}$$

[0044] 特定形式的加权因子 α 使高能量变换系数的影响最小,并突出低能量系数的贡献。最后,通过简单地对瞬时能量 E_{nf} 求平均值来估计噪声基底电平 \bar{E}_{nf} 。

$$[0045] \quad \bar{E}_{nf} = (\sum_{k=0}^{255} E_{nf}(k)) / 256$$

[0046] 本文提出的“峰值拾取”算法的一个实施例需要知道噪声基底能量电平和频谱峰值的平均能量电平。本文使用的峰值能量估计算法与上述噪声基底估计算法类似,但是替代于低能量,它跟踪如下高频谱能量:

$$[0047] \quad E_p(k) = \beta E_p(k-1) + (1-\beta) |X(k)|^2$$

$$[0048] \quad \beta = \begin{cases} 0.4223 & \text{如果 } |X(k)|^2 > E_p(k-1) \\ 0.8029 & \text{如果 } |X(k)|^2 \leq E_p(k-1) \end{cases}$$

[0049] 在这种情况下,加权因子 β 使低能量变换系数的影响最小,并突出高能量系数的贡献。这里通过对瞬时能量求平均值将总峰值能量 \bar{E}_p 估计为:

$$[0050] \quad \bar{E}_p = (\sum_{k=0}^{255} E_p(k)) / 256$$

[0051] 当计算峰值和噪声基底电平时,阈值电平 τ 可以形成为:

$$[0052] \quad \tau = \left(\frac{\bar{E}_p}{\bar{E}_{nf}} \right)^\gamma \bar{E}_{nf}$$

[0053] 将 γ 设置为示例值 $\gamma = 0.88579$ 。然后将输入音频信号的分段的变换系数与阈值进行比较,并且具有超过阈值的幅度的变换系数形成峰值候选的矢量。即,包括被假定为属于频谱峰值的系数的矢量。

[0054] 可以使用可能需要与计算 τ 相比更小的计算复杂度的备选阈值 $\theta(k)$, 以检测峰值。在一个实施例中, $\theta(k)$ 被建立为具有固定缩放因子的瞬时峰值包络水平 $E_p(k)$,。这里, 缩放因子 0.64 被用作示例, 使得:

$$[0055] \quad \theta(k) = E_p(k) \cdot 0.64$$

[0056] 当使用备选阈值 θ 时, 峰值候选被定义为具有高于瞬时阈值电平的平方幅度的所有系数, 如下:

$$[0057] \quad \begin{cases} |X(k)|^2 > \theta(k), k \in P \\ |X(k)|^2 \leq \theta(k), k \notin P \end{cases}$$

[0058] 其中 P 标记峰值候选的位置的频率有序集。考虑 FFT 频谱, 一些峰将是宽的并由若干变换系数组成, 而其他峰是窄的并由单个系数表示。为获得各个系数的峰值表示, 即每峰值一个系数, 假定连续位置中的峰值候选系数是更宽峰值的一部分。通过在连续峰候选位置 $\dots k-1, k, k+1, \dots$ 的范围中找到变换系数的最大平方幅度 $|X(k)|^2$, 创建细分集合 P , 其中宽峰由每个范围中的最大位置 (即由具有在范围中的最高值 $|X(k)|^2$ 、其也可以被标记为范围中具有最大谱幅度的系数) 表示。图 3a 示出了峰值包络和噪声基底包络的导出和峰值选择算法。

[0059] 上述计算用于生成用于形成分类器判决的两个特征: 即峰值稀疏度 S 的估计和峰值噪声基底比率 PNR 。峰值稀疏度 S 可以使用峰值之间的平均距离 d_i 表示或定义为:

$$[0060] \quad S = \begin{cases} \frac{1}{N_d} \sum_{i=1}^N d_i, N_d > 1 \\ 0, N_d \leq 1 \end{cases}$$

[0061] 其中 N_d 是集合 \hat{P} 中细分峰值的数量。 PNR 可以被计算为:

$$[0062] \quad PNR = \frac{\bar{E}_p}{\bar{E}_{nf}}$$

[0063] 可以使用这些特征结合判决阈值来形成分类器判决。可以将这些判决命名为“issparse”和“isclean”, 如:

$$[0064] \quad \text{issparse} = S > S_{\text{THR}}$$

$$[0065] \quad \text{isclean} = PNR > PNR_{\text{THR}}$$

[0066] 这些判决的结果可以用于形成不同类别的信号。图 4 中示出这些类别的说明。当分类基于两个二进制判决时, 类别的总数可以至多为 4。作为下一步骤, 如表 1 所示, 可以使用

类别信息来形成编解码器判决。

[0067] 表1:使用两个特征判决形成的可能类别。

	isclean	Issparse
类别A	假	假
类别B	真	假
类别C	真	真
类别D	假	真

[0069] 在音频编解码器中的以下步骤中,做出将哪个处理步骤应用于哪个类别的判决。即,至少基于S和PNR选择编码模式。该选择或映射将取决于可用的不同编码模式或处理步骤的特征和能力。作为示例,编解码器模式1将处理A类和C类,而编解码器模式2将处理B类和D类。编码模式判决可以是分类器的最终输出,以指导编码过程。编码模式判决通常将与来自所选编码模式的编解码器参数一起在比特流中传送。

[0070] 应当理解,上述类别可以进一步与其他分类器判决组合。组合可以导致更大数量的类别,或者它们可以使用优先级顺序组合,使得所呈现的分类器可以被另一个分类器支配,或者相反,所呈现的分类器可以支配另一分类器。

[0071] 本文所描述的解决方案提供了高分辨率音乐类型区分器,其可以有利地应用于音频编码中。区分器的判决逻辑基于具有显著能量的频率系数的位置分布的统计。

[0072] 实施方式

[0073] 上述方法和技术可以在编码器和/或解码器中实现,编码器和/或解码器可以是例如通信设备的一部分。

[0074] 编码器,图5a-5c

[0075] 在图5a中以一般方式示出了编码器的示例实施例。编码器指被配置用于对音频信号进行编码的编码器。编码器可以进一步被配置为对其他类型的信号进行编码。编码器500被配置为执行上述例如参照图2方法实施例中的至少一个。编码器500关联于与前述方法实施例相同的技术特征、目的和优点。编码器可以被配置为符合音频编码的一个或多个标准。为了避免不必要的重复,将简要描述编码器。

[0076] 可以如下实现和/或描述编码器:

[0077] 编码器500被配置为对音频信号进行编码。编码器500包括处理电路或处理装置501和通信接口502。处理电路501被配置为:针对音频信号的分段,使所述编码器500:识别频谱峰值的集合;确定所述集合中峰值之间的平均距离S;以及确定峰值包络与噪声基底包络之间的比值PNR。处理电路501还被配置为:使所述编码器至少基于所述平均距离S和比率PNR从多个编码模式中选择编码模式;并应用所选编码模式。通信接口502,也可以被标记为例如输入/输出(I/O)接口,包括用于向其它实体或模块发送数据和从其他实体或模块接收数据的接口。

[0078] 如图5b所示,处理电路501可以包括处理装置,例如处理器503(例如CPU)和用于存储或保持指令的存储器504。然后,存储器将包括例如计算机程序505的形式的指令,所述指令当由处理装置503执行时,使编码器500执行上述动作。

[0079] 在图5c中示出了处理电路501的备选实施方式。本文的处理电路包括识别单元506,其被配置为针对音频信号的分段识别一组频谱峰值。处理电路还包括第一确定单元

507,被配置为使所述编码器500确定所述集合中峰值之间的平均距离S。处理电路还包括第二确定单元508,被配置为使所述编码器确定峰值包络与噪声基底包络之间的比率PNR。处理电路还包括选择单元509,被配置为使所述编码器至少基于所述平均距离S和比率PNR从多个编码模式中选择编码模式。处理电路还包括编码单元510,被配置为使所述编码器应用所选编码模式。处理电路501可以包括更多单元,例如被配置为使编码器对输入信号进行滤波的滤波器单元。该任务在执行时可以备选地由一个或多个其他单元执行。

[0080] 上述编码器或编解码器可以被配置用于本文所述的不同方法实施例,例如使用不同的阈值来检测峰值。可以假设编码器500包括用于执行常规编码器功能的附加功能。

[0081] 处理电路的示例包括但不限于,一个或多个微处理器、一个或多个数字信号处理器(DSP)、一个或多个中央处理单元(CPU)、视频加速硬件、和/或任意合适的可编程逻辑电路,例如一个或多个现场可编程门阵列(FPGA)或者一个或多个可编程逻辑控制器(PLC)。

[0082] 也应当理解,可以重用其中实现了所提出的技术的任意传统设备或单元的通用处理能力。还可以例如通过重新编程现有的软件或者通过添加新的软件组件来重新使用现有的软件。

[0083] 区分器,图5d

[0084] 图5d示出了可以应用于编码器或解码器中的区分器或分类器的示例性实施方式。如图5d中所示,本文所述的区分器可以由例如处理器、适当软件以及合适存储装置或存储器中的一个或多个来实现,以根据本文中描述的实施例执行输入信号矢量的区分动作。在图5d中所示的实施例中,输入(IN)接收到来的信号,处理器和存储器与输入(IN)连接,并且从输出(OUT)输出从软件获得的音频信号(参数)的区分表示。

[0085] 区分器可以通过针对音频信号的分段识别频谱峰值的集合并确定所述集合中峰值之间的平均距离S而在不同语音信号类型之间区分。此外,区分器可以确定峰值包络和噪声基底包络之间的比率PNR,并然后至少基于平均距离S和比率PNR,从多个音频信号类别段中确定出分段所属的音频信号的类别。通过执行该方法,区分器使能例如适当地选择用于音频信号的编码方法或其它信号处理相关方法。

[0086] 如先前提及的,上述技术可以例如在发送器中使用,该发送器可以在移动设备(例如,移动电话、膝上型计算机)或固定设备(例如,个人计算机)中使用。

[0087] 可以在图6中看到示例性音频信号区分器的概述。图6示出根据示例实施例的具有区分器的编码器的示意框图。区分器包括被配置为接收表示要处理的音频信号的输入信号的输入单元、成帧单元、可选的预加重单元、频率变换单元、峰值/噪声包络分析单元、峰值候选选择单元、峰值候选细分单元、特征计算单元、类别判决单元、编码模式判决单元、多模编码器单元、音频信号的比特流/存储器和输出单元。所有这些单元可以用硬件实现。存在可使用并组合以实现编码器的单元功能的各种电路元件的变型。这些变体由实施例涵盖。区分器的硬件实现的特定示例是在数字信号处理器(DSP)硬件和集成电路技术中的实现,包括通用电子电路和专用电路。

[0088] 如先前所描述的,根据本文描述的实施例的区分器可以是编码器的一部分,并且根据本文描述的实施例的编码器可以是设备或节点的一部分。如先前提及的,本文的技术可以例如在发送器中使用,该发送器可以在移动设备(例如,移动电话或膝上型计算机)或固定设备(例如,个人计算机)中使用。

[0089] 应当理解,对互动单元或模块的选择以及单元的命名仅用于示例目的,并且可通过多个备选方式来配置,从而能够执行所公开的处理动作。

[0090] 还应当注意,本公开中描述的单元或模块应被认为是逻辑实体,并且不必是分离的物理实体。可以理解,本文公开技术的范围完全覆盖对本领域技术人员来说显而易见的其他实施例,相应地,本公开的范围不限于此。

[0091] 除非明确地阐述,单数形式的元件的参考不意图表示“一个且仅一个”,而是“一个或多个”。上述优选元素实施例的元素的对于本领域的普通技术人员已知的所有结构和功能等同物明确通过引用并入本文,并旨在由本权利要求所涵盖。此外,设备或方法不必须解决本文公开的技术所要解决的每个问题,其用于被包含于此。

[0092] 在前述说明中,出于描述而非限制的目的,阐述例如特定结构、接口、技术等特定细节,以提供对所公开的技术的透彻理解。然而,本领域技术人员将显而易见,所公开技术可以在偏离这些特定细节的其他实施例或实施例的组合中实践。即,本领域技术人员将能够想出体现所公开技术的原理的各种配置,尽管并未这里明确地描述或示出。在一些实例中,省略众所周知的装置、电路和方法的详细说明,以免不必要的细节模糊所公开技术的说明。本文中列出所公开技术的原理、方面和实施例,以及其具体实例的所有陈述旨在包括其结构和功能等同物。此外,不考虑结构,希望这种等价形式既包括当前已知的等价形式,也包括未来发展的等价形式,例如执行相同功能的发展的单元。

[0093] 因此,例如本领域技术人员将理解本文的附图可以代表体现技术的原理的说明性电路或其他功能单元的概念性视图,和/或可以大体上在计算机可读介质中表示和利用计算机或处理器执行的各种过程,即便不能在附图中明确示出这种计算机或处理器。

[0094] 可以通过诸如电路硬件和/或能够执行在计算机可读介质上存储的编码指令形式的软件的硬件的使用来提供包括功能模块的各种单元的功能。因此,这种功能和所示出的功能模块被理解为或者是硬件实现的和/或计算机实现,并因此是机器实现的。

[0095] 上述实施例被理解为本发明的几个说明性示例。本领域技术人员将理解,在不脱离本发明的范围的前提下,可以对实施例作出各种修改、合并和改变。尤其是,不同实施例中的不同部分的方案可在其他技术上可行配置中进行组合。

[0096] 缩写

[0097] DFT 离散傅里叶变换

[0098] FFT 快速傅里叶变换

[0099] MDCT 修正离散余弦变换

[0100] PNR 峰值噪声基底比率。

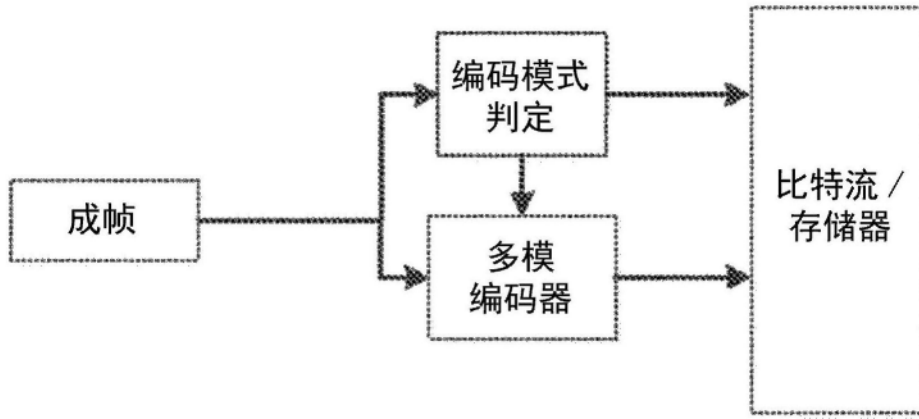


图1a

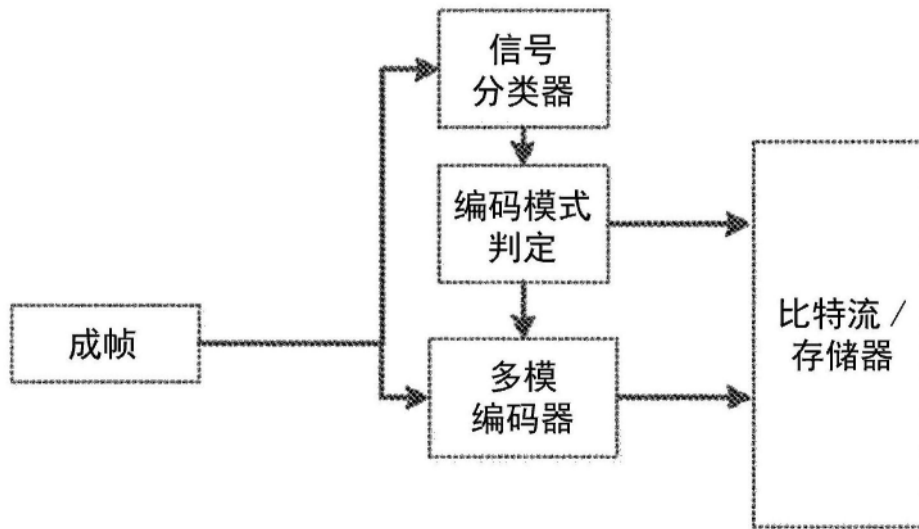


图1b

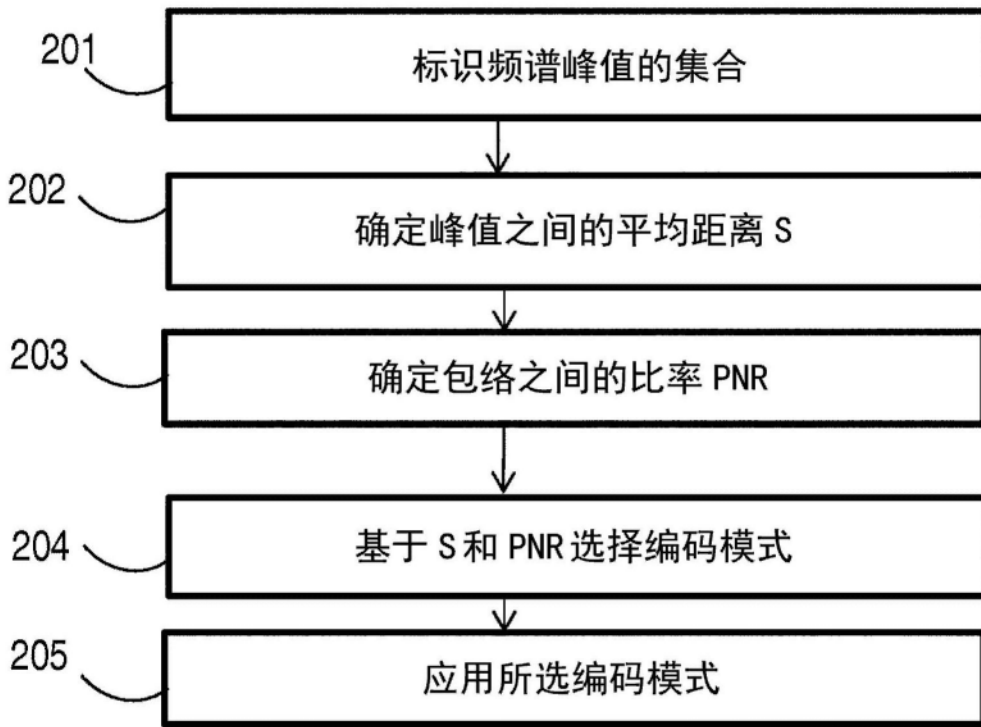


图2

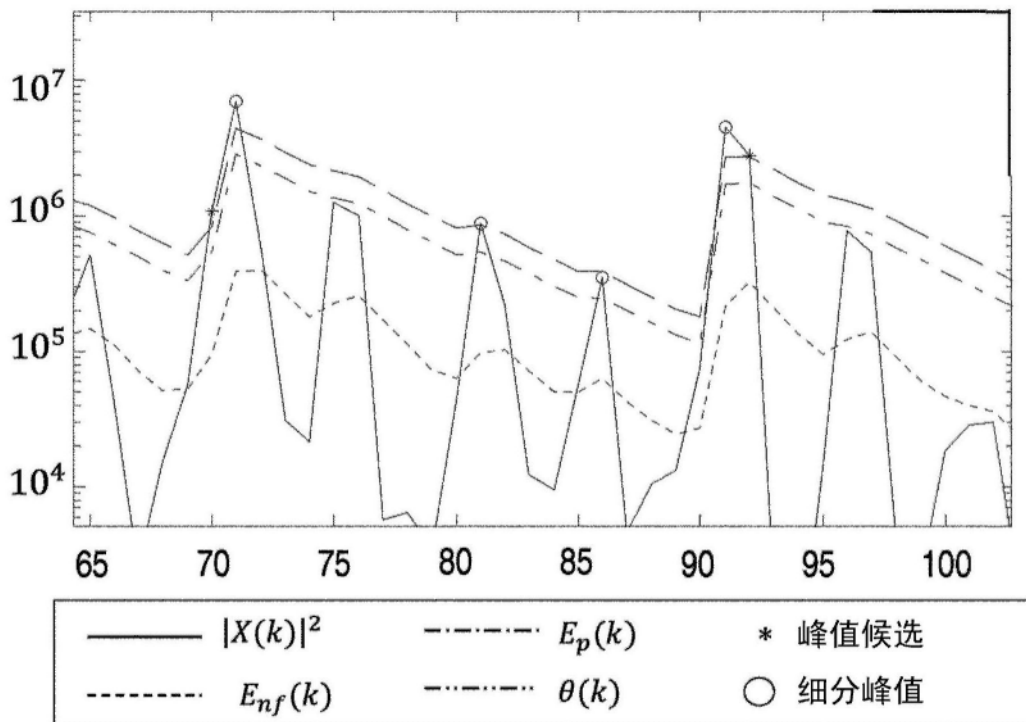


图3a

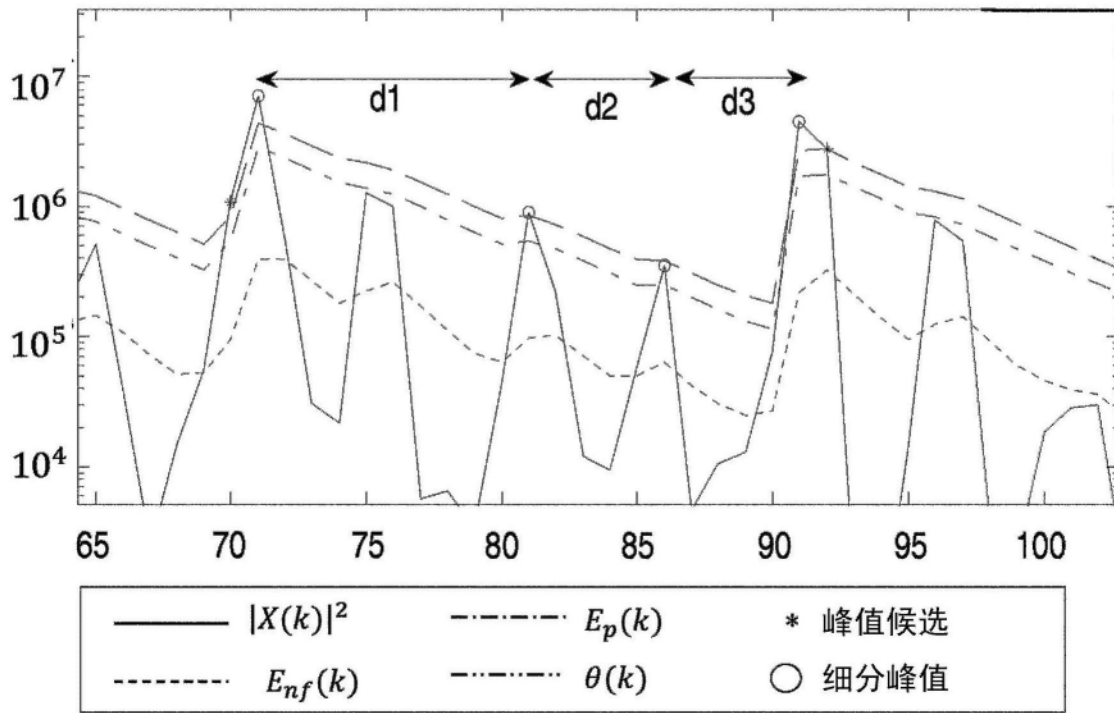


图3b

$$issparse = S > S_{THR}$$

$$isclean = PNR > PNR_{THR}$$

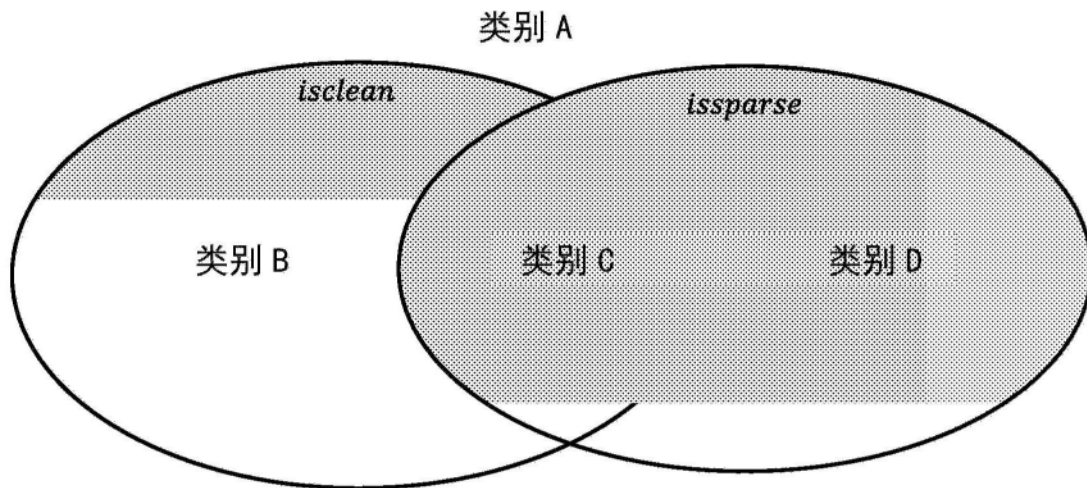


图4

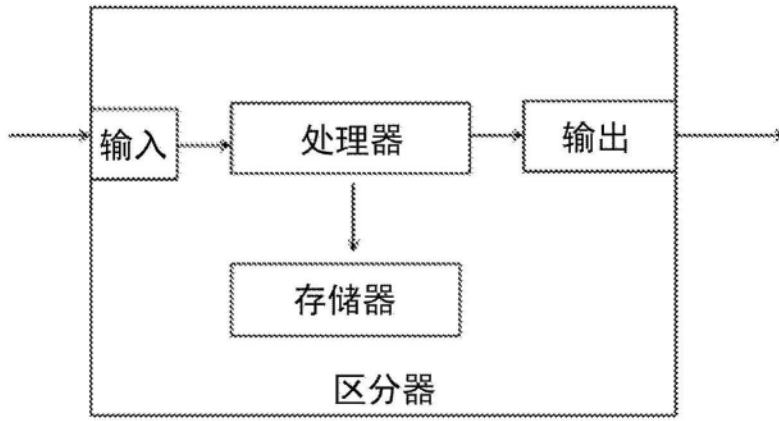


图5d

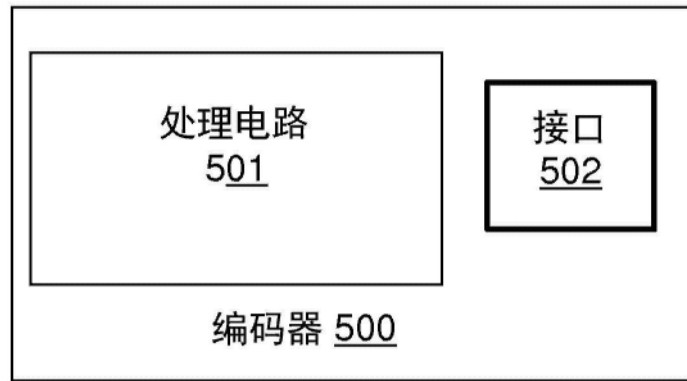


图5a

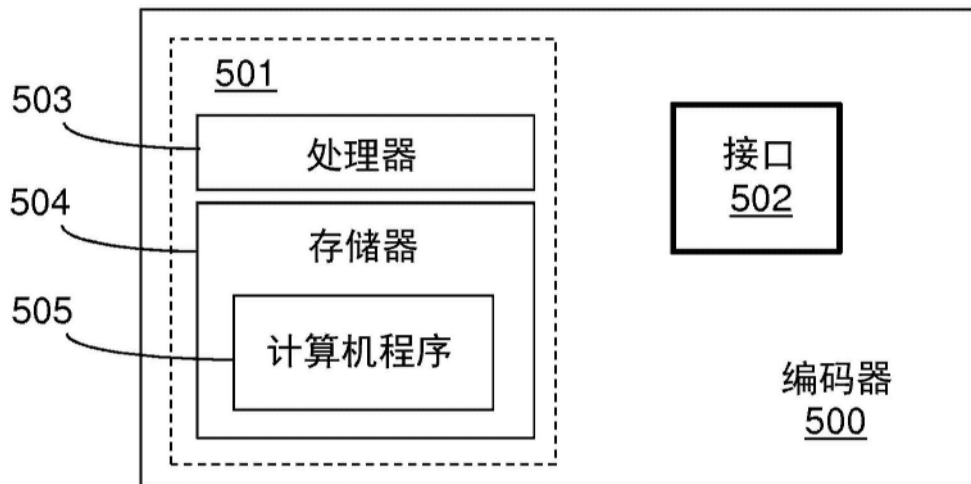


图5b

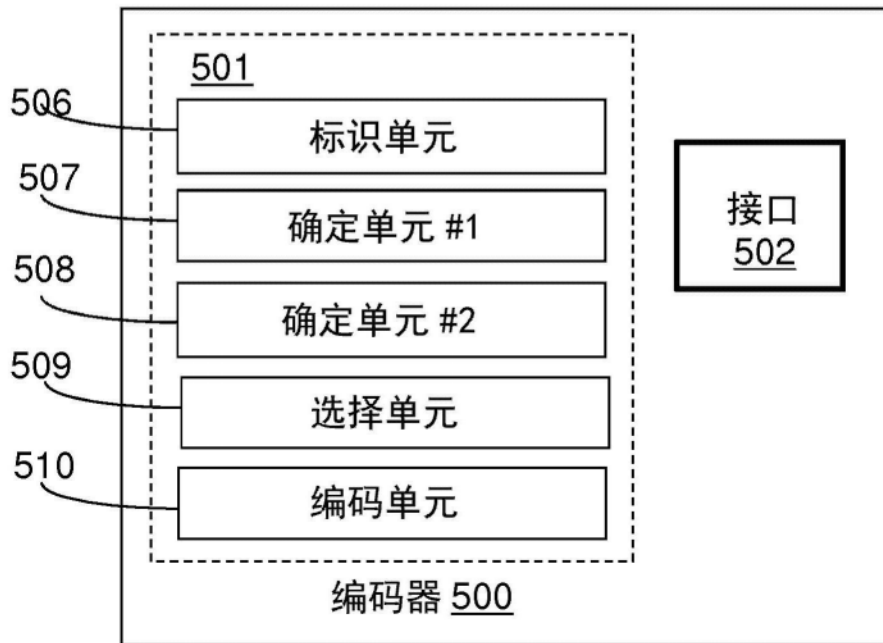


图5c

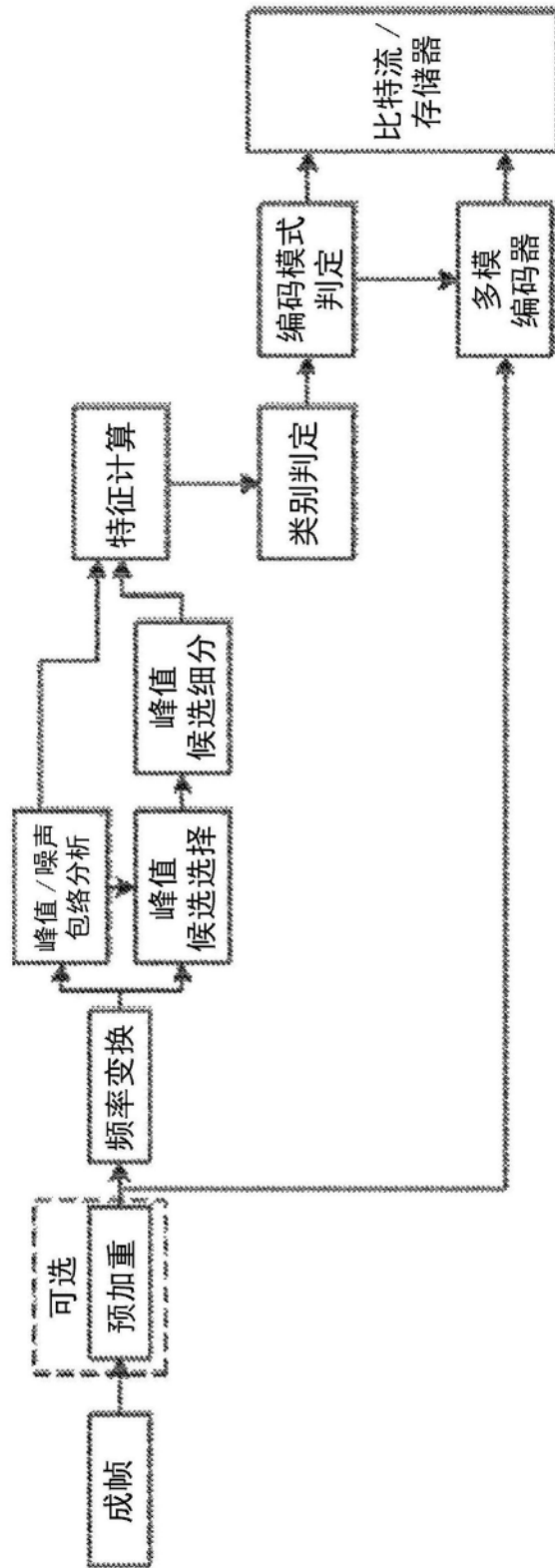


图6