

(12) 发明专利

(10) 授权公告号 CN 101335000 B

(45) 授权公告日 2010.04.21

(21) 申请号 200810084077.6

US 5559832 A, 1996.09.24, 全文.

(22) 申请日 2008.03.26

审查员 隋欣

(73) 专利权人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为
总部办公楼

(72) 发明人 艾雅·舒默特 张立斌 代金良

(51) Int. Cl.

G10L 19/00 (2006.01)

G10L 19/14 (2006.01)

(56) 对比文件

JP 3090842 B2, 2000.09.25, 全文.

CN 1656817 A, 2005.08.17, 全文.

US 5189669 A, 1993.02.23, 全文.

US 5577087 A, 1996.11.19, 全文.

JP 2595372 B2, 1997.04.02, 全文.

CN 1512487 A, 2004.07.14, 全文.

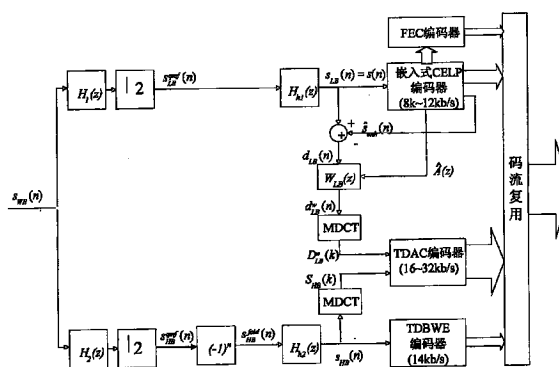
权利要求书 3 页 说明书 20 页 附图 7 页

(54) 发明名称

编码的方法及装置

(57) 摘要

本发明实施例公开了一种编码方法,包括:提取拖尾时间内的背景噪声特征参数;对所述拖尾时间后的第一个超帧,根据提取的所述背景噪声特征参数进行背景噪声编码;对第一个超帧后的超帧,对每一帧都进行背景噪声特征参数提取和 DTX 判决;对第一个超帧后的超帧,根据提取的当前超帧的背景噪声特征参数和当前超帧之前若干帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码。还公开了一种编码的装置。利用本发明实施例,能够在保证信号质量的情况下,明显降低占用的带宽。



1. 一种编码方法,其特征在于,包括:

提取拖尾时间内的背景噪声特征参数;

对所述拖尾时间后的第一个超帧,根据提取的所述拖尾时间内的背景噪声特征参数和所述第一个超帧的背景噪声特征参数,进行背景噪声编码;

对第一个超帧后的超帧,对每一帧都进行背景噪声特征参数提取和 DTX 判决;

对第一个超帧后的超帧,根据提取的当前超帧的背景噪声特征参数和所述当前超帧之前若干超帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码。

2. 如权利要求 1 所述的方法,其特征在于,所述拖尾时间为 120 毫秒或 140 毫秒。

3. 如权利要求 1 所述的方法,其特征在于,所述提取拖尾时间内的背景噪声特征参数具体为:

在所述拖尾时间内,对每一个超帧的每一帧,保存每帧背景噪声的自相关系数。

4. 如权利要求 1 所述的方法,其特征在于,对所述拖尾时间后的第一个超帧,所述根据提取的所述拖尾时间内的所述背景噪声特征参数和所述第一个超帧的背景噪声特征参数,进行背景噪声编码包括:

在所述拖尾时间后的第一个超帧的第一帧和所述拖尾时间后的第一个超帧的第二帧保存每个帧背景噪声的自相关系数;

在所述拖尾时间后的第一个超帧的第二帧,根据所述提取出的所述两帧的自相关系数和所述拖尾时间内的背景噪声特征参数,提取出所述第一个超帧的 LPC 滤波器系数和残差能量,进行背景噪声编码。

5. 如权利要求 4 所述的方法,其特征在于,所述提取所述 LPC 滤波器系数具体为:

计算所述第一个超帧和所述第一个超帧之前的所述拖尾时间内的最后四个超帧的自相关系数的算术平均值;

由所述自相关系数的算术平均值,根据 Levinson-Durbin 算法计算出 LPC 滤波器系数;

所述提取所述残差能量 E_t 具体为:

根据 Levinson-Durbin 算法计算出残差能量;

所述在第二帧进行背景噪声编码具体为:

将所述 LPC 滤波器系数转化到 LSF 域,进行量化编码;

将所述残差能量在对数域进行线性量化编码。

6. 如权利要求 1 所述的方法,其特征在于,所述对第一个超帧后的超帧,对每一帧都进行背景噪声特征参数提取具体为:

根据最近四个相邻帧自相关系数的值,计算当前自相关系数的稳态平均值,所述自相关系数的稳态平均值是所述最近四个相邻帧中具有中间自相关系数范数值的两个帧的自相关系数的算术平均值;

对所述稳态平均值,根据 Levinson-durbin 算法计算背景噪声 LPC 滤波器系数和残差能量。

7. 如权利要求 1 所述的方法,其特征在于,所述对每一帧都进行 DTX 判决具体为:

如果当前超帧中有一个帧的 DTX 判决结果为 1,则当前超帧窄带部分的 DTX 判决结果为 1。

8. 如权利要求7所述的方法,其特征在于,如果所述当前超帧的所述最终 DTX 判决结果为 1,则所述:“对第一个超帧后的超帧,根据提取的当前超帧的背景噪声特征参数和所述当前超帧之前若干超帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码”过程包括:

对于所述当前超帧,确定平滑因子,包括:

如果当前超帧第一帧的 DTX 为零,第二帧的 DTX 为 1,则所述平滑因子为 0.1,否则所述平滑因子为 0.5;

对所述当前超帧的两个帧进行参数平滑,将参数平滑后的参数作为对所述当前超帧的进行背景噪声编码的特征参数,所述参数平滑包括:

计算所述两个帧自相关系数稳态平均值的滑动平均值 $R^t(j)$:

$R^t(j) = \text{smooth_rate}R^{t-1}(j) + (1-\text{smooth_rate})R^{t-2}(j)$,所述 smooth_rate 为所述平滑因子, $R^{t-1}(j)$ 为第一帧的自相关系数稳态平均值, $R^{t-2}(j)$ 为第二帧的自相关系数稳态平均值;

对所述两个帧自相关系数稳态平均值的滑动平均值 $R^t(j)$,根据 Levinson-Durbin 算法,得到 LPC 滤波器系数,

计算所述两个帧能量估计的滑动平均值 E_t :

$E_t = \text{smooth_rate}E_{t,1} + (1-\text{smooth_rate})E_{t,2}$,所述 $E_{t,1}$ 为第一帧的能量估计, $E_{t,2}$ 为第二帧的能量估计。

9. 如权利要求1所述的方法,其特征在于,所述“根据提取的当前超帧的背景噪声特征参数和所述当前超帧之前若干超帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码”具体为:

计算当前超帧之前若干超帧的自相关系数的算术平均值;

根据所述自相关系数的算术平均值,计算当前超帧之前若干个超帧的平均 LPC 滤波器系数;

如果所述平均 LPC 滤波器系数与当前超帧的 LPC 滤波器系数差值小于或等于预设值,则将所述平均 LPC 滤波器系数转化到 LSF 域,进行量化编码;

如果所述平均 LPC 滤波器系数与当前超帧的 LPC 滤波器系数差值大于预设值,则将所述当前超帧的 LPC 滤波器系数转化到 LSF 域,进行量化编码;

对能量参数,在对数域进行线性量化编码。

10. 如权利要求9所述的方法,其特征在于,所述若干超帧的数量为 5。

11. 如权利要求1所述的方法,其特征在于,在所述提取拖尾时间内的背景噪声特征参数步骤之前,还包括:

用语音编码速率对所述拖尾时间内的背景噪声进行编码。

12. 一种编码装置,其特征在于,包括:

第一提取单元,用于:提取拖尾时间内的背景噪声特征参数;

第二编码单元,用于:对所述拖尾时间后的第一个超帧,根据提取的所述拖尾时间内的背景噪声特征参数和所述第一个超帧的背景噪声特征参数,进行背景噪声编码;

第二提取单元,用于:对所述第一个超帧后的超帧,对每一帧都进行背景噪声特征参数提取;

DTX 判决单元,用于:对所述第一个超帧后的超帧,对每一帧都进行 DTX 判决;

第三编码单元,用于:对所述第一个超帧后的超帧,根据提取的当前超帧的背景噪声特征参数和所述当前超帧之前若干超帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码。

13. 如权利要求 12 所述的装置,其特征在于,所述拖尾时间为 120 毫秒或 140 毫秒。

14. 如权利要求 14 所述的装置,其特征在于,所述第一提取单元具体为:

缓存模块,用于:在所述拖尾时间内,对每一个超帧的每一帧,保存每帧背景噪声的自相关系数。

15. 如权利要求 12 所述的装置,其特征在于,所述第二编码单元具体为:

提取模块,用于:在所述拖尾时间后的第一个超帧的第一帧和所述拖尾时间后的第一个超帧的第二帧保存每个帧背景噪声的自相关系数;

编码模块,用于:在所述拖尾时间后的第一个超帧的第二帧,根据所述提取出的所述两帧的自相关系数和所述拖尾时间内的背景噪声特征参数,提取出所述第一个超帧的 LPC 滤波器系数和残差能量,进行背景噪声编码。

16. 如权利要求 12 所述的装置,其特征在于,所述第二提取单元具体为:

第一计算模块,用于:根据最近四个相邻帧自相关系数的值,计算当前自相关系数的稳态平均值,所述自相关系数的稳态平均值是所述最近四个相邻帧中具有中间自相关系数范数值的两帧的自相关系数的算术平均值;

第二计算模块,用于:对所述稳态平均值,根据 Levinson-durbin 算法计算背景噪声 LPC 滤波器系数和残差能量。

17. 如权利要求 12 所述的装置,其特征在于,所述第三编码单元具体为:

第三计算模块,用于:根据计算得到的当前超帧之前若干超帧的自相关系数的算术平均值,计算当前超帧之前若干个超帧的平均 LPC 滤波器系数;

第一编码模块,用于:如果所述平均 LPC 滤波器系数与当前超帧的 LPC 滤波器系数差值小于或等于预设值,则将所述平均 LPC 滤波器系数转化到 LSF 域,进行量化编码;

第二编码模块,用于:如果所述平均 LPC 滤波器系数与当前超帧的 LPC 滤波器系数差值大于预设值,则将所述当前超帧的 LPC 滤波器系数转化到 LSF 域,进行量化编码;

第三编码模块,用于:对能量参数,在对数域进行线性量化编码。

18. 如权利要求 12 所述的装置,其特征在于,还包括:

第一编码单元,用于:用语音编码速率对所述拖尾时间内的背景噪声进行编码。

编码的方法及装置

技术领域

[0001] 本发明涉及通信技术领域,特别是涉及编码、解码的方法及装置。

背景技术

[0002] 在语音通信中,对于背景噪声的编解码是按 ITU(Intemational TelecomUnion,国际电信联盟)制定的 G. 729B 中规定的噪声处理方案进行的。

[0003] 在语音编码器中引入了静音压缩技术,其信号处理原理框图如图 1 所示。

[0004] 静音压缩技术主要包括三大模块:VAD(Voice Activity Detection,语音激活检测)、DTX(Discontinuous Transmission,非连续传输)和 CNG(ComfortNoise Generator,舒适噪声生成),其中 VAD、DTX 是编码器中的模块,CNG 是解码端中的模块。图 1 为一个简单的静音压缩系统原理框图,其基本流程为:

[0005] 首先在发送端(编码端),对每一输入信号帧,VAD 模块对当前的输入信号进行分析和检测,检测当前信号中是否包含语音信号,如果包含,则将当前帧设为语音帧,否则设为非语音帧。

[0006] 其次,编码器根据 VAD 检测结果对当前信号进行编码,如果 VAD 检测结果为语音帧,则信号进入语音编码器进行语音编码,输出为语音帧;如果 VAD 检测结果为非语音帧,则信号进入 DTX 模块用非语音编码器进行背景噪声处理,并输出非语音帧。

[0007] 最后,在接收端(解码端)对接收到的信号帧(包括语音帧和非语音帧)进行解码。如果接收到的信号帧为语音帧,则用语音解码器对其解码,否则进入 CNG 模块,在 CNG 模块根据非语音帧传过来的参数对背景噪声进行解码,产生舒适背景噪声或静音,使解码后的信号听起来更为自然和连续。

[0008] 在编码器中引入这种变速率的编码方式,通过对静音阶段的信号进行适当的编码,静音压缩技术有效的解决了背景噪声不连续的问题,提高了信号合成质量,因此,解码端的背景噪声也可称为舒适噪声。另外,由于背景噪声的编码速率要远远小于语音编码速率,因此系统的平均编码速率也大大降低,从而有效的节省了带宽。

[0009] G. 729B 处理信号时对信号采用分帧处理,帧长为 10ms。为了节省带宽,G. 729. 1 还定义了静音压缩系统的需求,要求其在背景噪声的情况下在不降低信号整体编码质量的前提下采用低速率的编码方式对背景噪声进行编码传输,即定义了 DTX 和 CNG 的需求,更重要的一点是要求其 DTX/CNG 系统要能够兼容 G. 729B。虽然可以将 G. 729B 的 DTX/CNG 系统简单移植到 G. 729. 1 中,但有两个问题需要解决:一是这两个编码器的处理帧长不同,直接移植会带来一些问题,而且 729B 的 DTX/CNG 系统有些简单,尤其是参数提取部分,为了满足 G. 729. 1 DTX/CNG 系统的需求需要对 729B 的 DTX/CNG 系统进行扩展。二是 G. 729. 1 处理的信号带宽为宽带,而 G. 729B 处理的带宽为窄带,在 G. 729. 1 的 DTX/CNG 系统还要加入背景噪声信号高频带部分(4000Hz ~ 7000Hz)的处理方式,使其成为一个完整的系统。

[0010] 现有技术中至少存在如下问题:现有的 G. 729B 系统由于处理的带宽为窄带背景噪声,在移植到 G. 729. 1 系统中时,不能保证编码信号的质量。

发明内容

[0011] 有鉴于此,本发明一个或多个实施例的目的在于提供一种编码、解码的方法及装置,以实现在将 G. 729B 进行扩展后,能够符合 G. 729. 1 技术标准的要求,实现了在保证编码质量的情况下,显著降低信号的通信带宽。

[0012] 为解决上述问题,本发明实施例提供了一种编码的方法,包括:

[0013] 提取拖尾时间内的背景噪声特征参数;

[0014] 对所述拖尾时间后的第一个超帧,根据提取的所述拖尾时间内的背景噪声特征参数和所述第一个超帧的背景噪声特征参数,进行背景噪声编码;

[0015] 对第一个超帧后的超帧,对每一帧都进行背景噪声特征参数提取和 DTX 判决;

[0016] 对第一个超帧后的超帧,根据提取的当前超帧的背景噪声特征参数和所述当前超帧之前若干超帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码。

[0017] 还提供了一种编码装置,包括:

[0018] 第一提取单元,用于:提取拖尾时间内的背景噪声特征参数;

[0019] 第二编码单元,用于:对所述拖尾时间后的第一个超帧,根据提取的所述拖尾时间内的背景噪声特征参数和所述第一个超帧的背景噪声特征参数,进行背景噪声编码;

[0020] 第二提取单元,用于:对所述第一个超帧后的超帧,对每一帧都进行背景噪声特征参数提取;

[0021] DTX 判决单元,用于:对所述第一个超帧后的超帧,对每一帧都进行 DTX 判决;

[0022] 第三编码单元,用于:对所述第一个超帧后的超帧,根据提取的当前超帧的背景噪声特征参数和所述当前超帧之前若干超帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码。

[0023] 与现有技术相比,本发明实施例具有以下优点:

[0024] 本发明实施例通过提取拖尾时间内的背景噪声特征参数;对所述拖尾时间后的第一个超帧,根据提取的所述背景噪声特征参数和所述第一个超帧的背景噪声特征参数,进行背景噪声编码;对第一个超帧后的超帧,对每一帧都进行背景噪声特征参数提取和 DTX 判决;对第一个超帧后的超帧,根据提取的当前超帧的背景噪声特征参数和所述当前超帧之前若干超帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码。实现了:

[0025] 首先,在保证编码质量的情况下,显著降低信号的通信带宽。

[0026] 其次,通过对于 G. 729B 系统的扩展,符合了 G. 729. 1 系统指标的要求。

[0027] 再次,通过灵活准确的背景噪声特征参数的提取,使得背景噪声的编码更加精确。

附图说明

[0028] 图 1 所示,为一个简单的静音压缩系统原理框图;

[0029] 图 2 所示,是 G. 729. 1 编码器功能框图;

[0030] 图 3 所示,是 G. 729. 1 解码器系统框图;

[0031] 图 4 所示,是本发明的编码的方法的实施例一的流程图;

[0032] 图 5 所示,是对第一个超帧编码的流程示意图;

[0033] 图 6 所示,是窄带部分参数提取及 DTX 判决的流程图;

- [0034] 图 7 所示,是当前超帧中的窄带部分背景噪声参数提取及 DTX 判决流程图;
- [0035] 图 8 所示,是本发明的解码方法的实施例一的流程图;
- [0036] 图 9 所示,是本发明的编码装置的实施例一的框图;
- [0037] 图 10 所示,是本发明的解码装置的实施例一的框图。

具体实施方式

[0038] 下面结合附图对本发明具体实施方式做进一步的详细阐述。

[0039] 首先,介绍 G. 729B 系统的有关原理。

[0040] 1. 1. 2. 语音编码码流和背景噪声编码码流中编码参数的异同

[0041] 在当前的语音编码器中,背景噪声的合成原理与语音的合成原理相同,采用的模型均是 CELP (Code Excited Linear Prediction, 码激励线性预测) 模型。语音的合成原理是:语音 $s(n)$ 可以看成是一个激励信号 $e(n)$ 激励一个合成滤波器 $v(n)$ 所产生的输出,即 $s(n) = e(n)*v(n)$,这就是语音产生的数学模型。在合成背景噪声时用的也是这个模型,所以背景噪声编码码流中所传输的描述背景噪声和静音特性的特征参数内容与语音编码码流中的特征参数基本相同,为信号合成时的合成滤波器参数和激励参数。

[0042] 在语音编码码流中,合成滤波器参数主要为线谱频率 LSF 量化参数,而激励信号参数包括:基音延迟参数、基音增益参数、固定码本参数和固定码本增益参数。不同的语音编码器,这些参数的量化比特数和量化形式有所不同;相同的编码器,如果其包含多个速率,在不同速率下,由于描述信号特性的侧重点不同,编码参数的量化比特数和量化形式也有所不同。

[0043] 与语音编码参数不同,背景噪声编码参数描述的是背景噪声特性,由于背景噪声的激励信号可以认为是简单的噪声随机序列,而这些序列在编解码端均可以简单的用随机噪声产生模块产生,然后用能量参数控制这些序列的幅度,就可产生最终的激励信号,因此激励信号特征参数可以简单的用能量参数来表示,而不需要用其它的一些特征参数来进一步描述,所以在背景噪声编码码流中,其激励参数为当前背景噪声帧的能量参数,这与语音帧不同;与语音帧相同的是,背景噪声编码码流中的合成滤波器参数也为线谱频率 LSF 量化参数,只是具体的量化方法有所差别。通过以上分析,可以认为对背景噪声的编码方式本质上就是一种简单的“语音”编码方式。

[0044] G. 729B 的噪声处理方案(引用 729B 协议)

[0045] 1. 2. 1 DTX/CNG 总体技术介绍

[0046] G. 729B 的静音压缩方案是较早的一种静音压缩技术,其背景噪声编解码技术基于的算法模型是 CELP,因此其所传输的背景噪声参数也是基于 CELP 模型来提取的,是描述背景噪声的合成滤波器参数与激励参数,其中激励参数是描述背景噪声能量的能量参数,没有描述语音激励的自适应和固定码本参数,滤波器参数与语音编码参数基本一致,是 LSF 参数。在编码端,对每帧输入的语音信号,若 VAD 的判决结果为“0”,表示当前的信号为背景噪声,那么编码器将信号送入 DTX 模块,在 DTX 模块对背景噪声参数进行提取,然后根据每帧参数变化情况来对背景噪声编码:若当前帧提取的的滤波器参数和能量参数与前几帧的变化较大,那么表示当前的背景噪声特性与之前的背景噪声特性相比有较大的差别,则在噪声编码模块将当前帧提取的背景噪声参数进行编码,组装成 SID 帧 (Silence Insertion

Descriptor, 静音描述帧) 发送到解码端, 否则发送 NODATA 帧 (无数据) 到解码端。SID 帧和 NODATA 帧称为非语音帧。在解码端, 若进入背景噪声阶段, 则在 CNG 模块中根据接收的非语音帧合成出描述编码端背景噪声特性的舒适噪声。

[0047] G. 729B 处理信号时对信号采用分帧处理, 帧长为 10ms。下面分三节分别描述 729B 的 DTX、噪声编码和 CNG 模块。

[0048] 1. 2. 2DTX 模块

[0049] DTX 模块主要用来做背景噪声参数的估计与量化以及 SID 帧的发送。在非语音阶段, DTX 模块需要将背景噪声信息发送到解码端, 背景噪声信息被封装在 SID 帧中发送, 若当前的背景噪声不平稳那么发送 SID 帧, 否则不发送 SID 帧, 而发送没有任何数据的 NODATA 帧。另外两个相邻 SID 帧间的间隔是有限制的, 限制为两帧, 若背景噪声不平稳, 需要连续的发送 SID 帧, 那么后一个 SID 帧的发送会延迟。

[0050] 在编码端, DTX 模块会从编码器中接收 VAD 模块的输出, 自相关系数以及过去的激励样点, 在每一帧, DTX 模块会用三个值 0、1、2 来分别描述非发送帧, 语音帧和 SID 帧, 它们的帧类型分别为 $F_{typ} = 0$ 、 $F_{typ} = 1$ 和 $F_{typ} = 2$ 。

[0051] 背景噪声估计的内容是背景噪声的能量电平和频谱包络, 这个和语音编码参数在本质上是一致的, 因此频谱包络的计算和语音编码参数的计算基本一致, 用到的参数包括了前两帧的参数; 而能量参数也是前几帧能量的一个平均值。

[0052] DTX 模块的主要操作:

[0053] a、每帧自相关系数的存储

[0054] 对每一输入的信号帧, 包括语音帧的和非语音帧, 将当前帧 t 的自相关系数保留在缓存中, 这些自相关系数表示为: $r'_t(j)$, $j = 0 \dots 10$ 。其中 j 为每帧自相关函数的序号。

[0055] b、估计当前的帧类型

[0056] 如果当前是一个语音帧, 即 $VAD = 1$, 那么将当前的帧类型设为 1, 若为非语音帧, 则依据前一帧和本帧的自相关系数计算一个当前的 LPC 滤波器 $A_t(z)$, 在计算 $A_t(z)$ 之前会首先计算相邻两帧自相关系数的平均值:

$$[0057] \quad R^t(j) = \sum_{i=t-N_{cur}+1}^t r'_i(j), j = 0 \dots 10$$

[0058] 其中 $N_{cur} = 2$, 计算出 $R^t(j)$ 后依据 Levinson-Durbin 算法计算出 $A_t(z)$ 。另外 Levinson-Durbin 算法也会计算出残差能量 E_t , 并以此做帧激励能量的简单估计。

[0059] 当前帧的帧类型会用以下的方式进行估计:

[0060] (1)、如果当前的帧是第一个非活动帧, 那么将此帧设为 SID 帧, 并令表征信号能量的变量 E 等于 E_t , 表征帧数目的参数 k_E 会设成 1:

$$[0061] \quad (Vad_{t-1} = 1) \Rightarrow \begin{cases} F_{typ} = 2 \\ \bar{E} = E_t \\ k_E = 1 \end{cases}$$

[0062] (2)、对于其它的非语音帧, 算法将之前的 SID 帧参数与当前相应的参数进行对比, 如果当前的滤波器与之前的滤波器差别较大或者当前的激励能量与之前的激励能量较大, 那么令标志 $flag_change$ 等于 1, 否则标志的值不变。

[0063] (3)、当前的计数器 count_fr 表示了当前帧与上一个 SID 之间的帧的数目。如果其值大于 N_{\min} , 那么发送 SID 帧; 另外如果 flag_change 等于 1, SID 帧也会发送, 其它情况下, 不发送当前帧:

$$[0064] \left. \begin{array}{l} \text{count_fr} \geq N_{\min} \\ \text{flag_chang} = 1 \end{array} \right\} \Rightarrow \text{Ftyp}_t = 2$$

[0065] Otherwise: $\text{Ftyp}_t = 0$

[0066] 在 SID 帧的情况下, 计数器 count_fr 和标志 flag_change 重新初始化成 0。

[0067] c、LPC 滤波器系数:

[0068] 设上一个 SID 的 LPC 滤波器 $A_{\text{sid}}(z)$ 的系数为 $a_{\text{sid}}(j)$, $j = 0 \dots 10$, 如果当前帧和上一帧的 SID-LPC 滤波器的 Itakura 距离超过了一定的门限, 就认为二者有很大的不同:

$$[0069] \sum_{j=0}^{10} R_a(j) \times R'(j) \geq E_t \times \text{thr1}$$

[0070] 其中, $R_a(j)$, $j = 0 \dots 10$ 是 SID 滤波器系数的自相关系数:

$$[0071] \begin{cases} R_a(j) = 2 \sum_{k=0}^{10-j} a_{\text{sid}}(k) \times a_{\text{sid}}(k+j) & \text{if } (j \neq 0) \\ R_a(0) = \sum_{k=0}^{10} a_{\text{sid}}(k)^2 \end{cases}$$

[0072] d、帧能量:

[0073] 计算帧能量的和:

$$[0074] \bar{E} = \sum_{i=t-k_E+1}^t E_i$$

[0075] 然后对 E 用 5 比特的对数量化器量化。解码之后的对数能量 E_q 会与上一个解码后的 SID 对数能量 E_q^{sid} 进行比较, 如果二者的差超过了 2dB, 那么认为二者的能量差别较大。

[0076] 1. 2. 3 噪声编码及 SID 帧

[0077] SID 帧中的参数就是 LPC 滤波器系数 (频谱包络) 和能量的量化参数。

[0078] 在对 SID-LPC 滤波器计算时考虑了相邻噪声帧之间的稳定情况:

[0079] 首先, 计算当前 SID 帧之前 N_p 帧的平均 LPC 滤波器 $A_p(z)$, 这要用到自相关函数和 $R_p(j)$, 然后将 $R_p(j)$ 送入 Levinson-Durbin 算法中得到 $A_p(z)$, 而 $R_p(j)$ 表示为:

$$[0080] \bar{R}_p(j) = \sum_{k=t'-N_p}^t r_k(j), j = 0 \dots 10$$

[0081] 其中 N_p 的值定为 6。帧数目 t' 的范围是 $[t-1, t-N_{\text{cur}}]$ 。这样, SID-LPC 滤波器表示为:

$$[0082] A_{\text{sid}}(z) = \begin{cases} A_t(z) & \text{if } \text{distance}(A_t(z), \bar{A}_p(z)) \geq \text{thr3} \\ \bar{A}_p(z) & \text{otherwise} \end{cases}$$

[0083] 即算法会计算前几帧的平均 LPC 滤波器系数 $A_p(z)$, 然后用其与当前的 LPC 滤波器系数 $A_t(z)$ 进行比较, 若二者差值较小, 那么当前帧在量化 LPC 系数时选择的的就是前几帧的平均值 $A_p(z)$, 否则就是当前帧的 $A_t(z)$ 。选择好 LPC 滤波器系数之后, 算法将这些 LPC 滤波器系数转化到 LSF 域, 然后进行量化编码, 而量化编码选择的方式与语音编码的量化编码

方式是一样的。

[0084] 能量参数的量化在对数域完成,采用的是线性的量化,然后用 5bit 进行编码。这样对背景噪声的编码就已完成,然后将这些编码比特封装在 SID 帧中。如表 A 所示:

[0085] 表 A

[0086] TABLE B. 2/G. 729

[0087]

Parameterdescription	Bits
SwitchedpredictorindexofLSFquantizer	1
FirststagevectorofLSFquantizer	5
SecondstagevectorofLSFquantizer	4
Gain (Energy)	5

[0088] SID 帧中的参数由四个码本索引构成,其中一个用来指示能量量化索引 (5 比特),另三个指示频谱量化的索引 (10 比特)。

[0089] 1. 2. 4CNG 模块

[0090] 在解码端,算法用一个电平可控的伪白噪声激励一个经内插得到的 LPC 合成滤波器得到舒适的背景噪声,这在本质上与语音的合成方式一样。其中激励电平和 LPC 滤波器系数分别从上一个 SID 帧中得到。子帧的 LPC 滤波器系数通过 SID 帧中的 LSP 参数的内插来得到,而内插方法与语音编码器中的内插方式是一致的。

[0091] 伪白噪声激励 $ex(n)$ 是语音激励 $ex1(n)$ 和高斯白噪声激励 $ex2(n)$ 的一个混合。 $ex1(n)$ 的增益较小,而采用 $ex1(n)$ 的目的是为了让语音和非语音间的过渡更为自然。

[0092] 这样得到激励信号后用其激励合成滤波器即可得到舒适的背景噪声。

[0093] 由于编解码双方的非语音编解码要保持同步,所以在双方都要为 SID 帧和不发送帧产生激励信号。

[0094] 首先,定义目标激励增益 \tilde{G}_t ,令其作为当前帧激励平均能量的平方根, \tilde{G}_t 由下面的平滑算法得到,其中 \tilde{G}_{sid} 是解码后的 SID 帧的增益:

$$[0095] \quad \tilde{G}_t = \begin{cases} \tilde{G}_{sid} & \text{if } (Vad_{t-1} = 1) \\ \frac{7}{8}\tilde{G}_{t-1} + \frac{1}{8}\tilde{G}_{sid} & \text{otherwise} \end{cases}$$

[0096] 80 个采样点被分成两个子帧,对每个子帧,CNG 模块的激励信号会用如下的方式来合成:

[0097] (1)、在 [40, 103] 范围内随机选择基音延迟;

[0098] (2)、子帧的固定码本矢量中非零脉冲的位置和符号随机选择(这些非零脉冲的位置和符号的结构与 G. 729 是一致的);

[0099] (3)、选择一个带增益的自适应码本激励信号,将其标记为 $e_a(n)$, $n = 0 \dots 39$,而选

择的固定码本激励信号标记为 $e_f(n)$, $n = 0 \dots 39$ 。然后以子帧能量为依据计算自适应增益 G_a 和固定码本增益 G_f :

$$[0100] \quad \frac{1}{40} \sum_{n=0}^{39} (G_a \times e_a(n) + G_f \times e_f(n))^2 = \tilde{G}_t^2$$

[0101] 需要注意的是 G_f 可以选择负值。

[0102] 定义: $E_a = \left(\sum_{n=0}^{39} e_a(n)^2 \right)$, $I = \left(\sum_{n=0}^{119} e_a(n)e_f(n) \right)$, $K = 40 \times \tilde{G}_t^2$, 而由 ACELP 的激励结构可知 $\sum_{n=0}^{39} e_f(n)^2 = 4$ 。

[0103] 如果将自适应码本增益 G_a 固定, 那么表现 \tilde{G}_t 的方程就变成了一个关于 G_f 的二阶方程:

$$[0104] \quad G_f^2 + \frac{G_a \times I}{2} G_f + \frac{E_a \times G_a^2 - K}{4} = 0$$

[0105] G_a 的值会被限定以确保上面的方程有解, 更进一步, 可以对一些大的自适应码本增益值的应用进行限制, 这样, 自适应码本增益 G_a 可以在如下的范围内随机的选择:

$$[0106] \quad \left[0, \text{Max} \left\{ 0.5, \sqrt{\frac{K}{A}} \right\} \right], \text{ with } A = E_a - I^2/4$$

[0107] 将方程 $\frac{1}{40} \sum_{n=0}^{39} (G_a \times e_a(n) + G_f \times e_f(n))^2 = \tilde{G}_t^2$ 的根中绝对值最小的作为 G_f 的值。

[0108] 最后, 用下式构建 G. 729 的激励信号:

$$[0109] \quad ex_1(n) = G_a \times e_a(n) + G_f \times e_f(n), n = 0 \dots 39$$

[0110] 合成激励 $ex(n)$ 可由如下方法合成:

[0111] 设 E_1 是 $ex_1(n)$ 的能量, E_2 是 $ex_2(n)$ 的能量, E_3 是 $ex_1(n)$ 和 $ex_2(n)$ 的点积:

$$[0112] \quad E_1 = \sum ex_1^2(n)$$

$$[0113] \quad E_2 = \sum ex_2^2(n)$$

$$[0114] \quad E_3 = \sum ex_1(n) \cdot ex_2(n)$$

[0115] 而计算的点数超过自身的大小。

[0116] 令 α 和 β 分别是混合激励中 $ex_1(n)$ 和 $ex_2(n)$ 的比例系数, 其中 α 设为 0.6, 而

[0117] β 依照下面的二次方程确定:

$$[0118] \quad \beta^2 E_2 + 2\alpha\beta E_3 + (\alpha^2 - 1)E_1 = 0, \text{ with } \beta > 0$$

[0119] 如果 β 没有解, 那么 β 将被设成 0, 而 α 设成 1。最终的 CNG 模块的激励变为 $ex(n)$:

$$[0120] \quad ex(n) = \alpha ex_1(n) + \beta ex_2(n)$$

[0121] 以上即为 729. B 编码器的 DTX/CNG 模块的基本原理。

[0122] 1. 3G. 729. 1 编解码器的基本流程

[0123] G. 729. 1 是 ITU 最新发布的新一代语音编解码标准 (见参考文献 [1]), 其是 ITU-TG. 729 在 8-32kbit/s 可分级宽带 (50-7000Hz) 上的扩展。默认情况下, 编码器输入和解码器输出端采样频率为 16000Hz。编码器产生的码流具有可分级性, 包含 12 个嵌入式层,

分别被称作第 1 ~ 12 层。第 1 层为核心层,对应比特率为 8kbit/s。该层与 G. 729 码流一致,从而使得 G. 729EV 与 G. 729 具有互操作性。第 2 层为窄带增强层,增加了 4kbit/s,而第 3 ~ 12 层是宽带增强层,以每层 2kbit/s 速度共增加 20kbit/s。

[0124] G. 729. 1 编解码器基于三阶段结构:嵌入式码激励线性估计 (CELP) 编解码,时域带宽扩展 (TDBWE) 以及被称为时域混叠消除 (TDAC) 的估计转换编解码。嵌入式 CELP 阶段产生第 1 和第 2 层,生成 8kbit/s 和 12kbit/s 窄带合成信号 (50-4000Hz)。TDBWE 阶段产生第 3 层,生成 14kbit/s 宽带输出信号 (50-7000Hz)。TDAC 阶段工作在改进离散余弦变换 (MDCT) 域,生成第 4 ~ 12 层,将信号质量从 14kbit/s 提高到 32kbit/s。TDAC 编解码同时代表 50-4000Hz 频带加权 CELP 编解码误码信号和 4000-7000Hz 频带输入信号。

[0125] 参考图 2 所示,给出了 G. 729. 1 编码器功能框图。编码器工作于 20ms 输入超帧。默认情况下,输入信号 $s_{WB}(n)$ 在 16000Hz 进行采样。因此,输入超帧具有 320 个采样点长度。

[0126] 首先,输入信号 $s_{WB}(n)$ 经过 QMF 滤波 ($H_1(z)$, $H_2(z)$) 分成两个子带,低子带信号 $s_{LB}^{qmf}(n)$ 经过 50Hz 截止频率的高通滤波器进行预处理,输出信号 $s_{LB}(n)$ 使用 8kb/s ~ 12kb/s 的窄带嵌入式 CELP 编码器进行编码, $s_{LB}(n)$ 和 12Kb/s 码率下 CELP 编码器的本地合成信号 $\hat{s}_{enh}(n)$ 之间的差值信号为 $d_{LB}(n)$, 将其经过知觉加权滤波 ($W_{LB}(z)$) 后得到信号 $d_{LB}^w(n)$, 将 $d_{LB}^w(n)$ 通过 MDCT 变换到频域。加权滤波器 $W_{LB}(z)$ 包含了增益补偿,用来保持滤波器输出 $d_{LB}^w(n)$ 与高子带输入信号 $s_{HB}(n)$ 之间的谱连续性。

[0127] 高子带分量乘上 $(-1)^n$ 进行谱折叠之后获得信号 $s_{HB}^{fold}(n)$, 将 $s_{HB}^{fold}(n)$ 通过截止频率为 3000Hz 的低通滤波器进行预处理,滤波后的信号 $s_{HB}(n)$ 使用 TDBWE 编码器进行编码。信号 $s_{HB}(n)$ 也通过 MDCT 变换为频域信号。

[0128] 两组 MDCT 系数 $D_{LB}^w(k)$ 和 $S_{HB}(k)$ 最后使用 TDAC 编码器进行编码。

[0129] 另外,还有一些参数用 FEC (丢帧错误隐蔽) 编码器进行传输,用以改进在传输中出现丢帧时造成的错误。

[0130] 解码器系统框图如图 3 所示,解码器的实际工作模式由接收到的码流层数决定,也等价于由接收到的码率决定。

[0131] (1)、如果接收到的码率为 8kb/s 或 12kb/s (即仅接收到第一层或者前两层):第一层或者前两层的码流由嵌入式 CELP 解码器进行解码,得到解码后的信号 $\hat{s}_{LB}(n)$, 再进行后滤波得到 $\hat{s}_{LB}^{post}(n)$, 并经高通滤波获得 $\hat{s}_{LB}^{qmf}(n) = \hat{s}_{LB}^{hpf}(n)$ 。输出信号由 QMF 合成滤波器组产生,其中高频合成信号 $\hat{s}_{HB}^{qmf}(n)$ 被置零。

[0132] (2)、如果接收到的码率为 14kb/s (即接收到前三层):除了 CELP 解码器解码出窄带分量以外, TDBWE 解码器也解码出高带信号分量 $\hat{s}_{HB}^{bwe}(n)$ 。对 $\hat{s}_{HB}^{bwe}(n)$ 进行 MDCT 变换,把高子带分量谱中 3000Hz 以上 (对应于 16kHz 采样率中 7000Hz 以上) 频率分量置 0, 然后进行逆 MDCT 变换,迭加之后并进行谱翻转,然后在 QMF 滤波器组中将重建的高频带信号 $\hat{s}_{HB}^{qmf}(n)$ 与 CELP 解码器解出的低带分量 $\hat{s}_{LB}^{qmf}(n) = \hat{s}_{LB}^{post}(n)$ 一起合成 16kHz 的宽带信号 (不进行高通滤波)。

[0133] (3)、如果接收到 14kb/s 以上速率的码流 (对应于前四层或者更多层):除了 CELP

解码器解码出低子带分量 $\hat{s}_{LB}^{post}(n)$ 、TDBWE 解码器解码出高子带分量 $\hat{s}_{HB}^{bwe}(n)$ 以外,TDAC 解码器还负责重建 MDCT 系数 $\hat{D}_{LB}^w(k)$ 和 $\hat{S}_{HB}(k)$,二者分别对应于低频带(0-4000Hz)重建加权差值和 高频带(4000-7000Hz)重建信号(注意到在高频带中,非接收子带和 TDAC 零码分配子带被替换为电平调整子带信号 $\hat{S}_{HB}^{bwe}(k)$)。 $\hat{D}_{LB}^w(k)$ 及 $\hat{S}_{HB}(k)$ 通过反向 MDCT 和重叠相加变换为时域信号。然后,低频带信号 $\hat{d}_{LB}^w(n)$ 经由感知加权滤波器进行处理。为减少变化编码带来的影响,对低频带和高频带信号 $\hat{d}_{LB}^w(n)$ 和 $\hat{s}_{HB}(n)$ 进行前向/后向回声监测和压缩。低频带合成信号 $\hat{s}_{LB}(n)$ 经由后滤波处理,而高频带合成信号 $\hat{s}_{HB}^{fold}(n)$ 经由 $(-1)^n$ 频谱折叠处理。然后,QMF 合成滤波器组对信号 $\hat{s}_{LB}^{qmf}(n) = \hat{s}_{LB}^{post}(n)$ 和 $\hat{s}_{HB}^{qmf}(n)$ 进行组合和上采样,得到最终的 16kHz 的宽带信号。

[0134] 1. 4G. 729. 1DTX/CNG 系统的需求

[0135] 为了节省带宽,G. 729. 1 还定义了静音压缩系统的需求,要求其在背景噪声的情况下在不降低信号整体编码质量的前提下采用低速率的编码方式对背景噪声进行编码传输,即定义了 DTX 和 CNG 的需求,更重要的一点是要求其 DTX/CNG 系统要能够兼容 G. 729B。虽然可以将 G. 729B 的 DTX/CNG 系统简单移植到 G. 729. 1 中,但有两个问题需要解决:一是这两个编码器的处理帧长不同,直接移植会带来一些问题,而且 729B 的 DTX/CNG 系统有些简单,尤其是参数提取部分,为了满足 G. 729. 1DTX/CNG 系统的需求需要对 729B 的 DTX/CNG 系统进行扩展。二是 G. 729. 1 处理的信号带宽为宽带,而 G. 729B 处理的带宽为窄带,在 G. 729. 1 的 DTX/CNG 系统还要加入背景噪声信号高频带部分(4000Hz ~ 7000Hz)的处理方式,使其成为一个完整的系统。

[0136] 在 G. 729. 1 中,可以将背景噪声的高频带和低频带分别进行处理。其中高频带的处理方式比较简单,其背景噪声特征参数的编码方式可参考语音编码器的 TDBWE 编码方式,判决部分简单比较频域包络和时域包络的稳定性即可。本发明的技术方案及要解决的问题是在低频带,也即窄带。以下所指的 G. 729. 1DTX/CNG 系统,是指应用于对窄带 DTX/CNG 部分的有关处理过程。

[0137] 参考图 4 所示,是本发明的编码的方法的实施例一,包括步骤:

[0138] 步骤 401、提取拖尾时间内的背景噪声特征参数;

[0139] 步骤 402、对所述拖尾时间后的第一个超帧,根据提取的所述拖尾时间内的背景噪声特征参数和所述第一个超帧的背景噪声特征参数,进行背景噪声编码,得到第一个 SID 帧;

[0140] 步骤 403、对第一个超帧后的超帧,对每一帧都进行背景噪声特征参数提取和 DTX 判决;

[0141] 步骤 404、对第一个超帧后的超帧,根据提取的当前超帧的背景噪声特征参数和所述当前超帧之前若干超帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码。

[0142] 利用本发明实施例,通过提取拖尾时间内的背景噪声特征参数;对所述拖尾时间后的第一个超帧,根据提取的所述拖尾时间内的背景噪声特征参数和所述第一个超帧的背景噪声特征参数,进行背景噪声编码;

[0143] 对第一个超帧后的超帧,对每一帧都进行背景噪声特征参数提取和 DTX 判决;

[0144] 对第一个超帧后的超帧,根据提取的当前超帧的背景噪声特征参数和所述当前超帧之前若干超帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码。实现了:

[0145] 首先,在保证编码质量的情况下,显著降低信号的通信带宽。

[0146] 其次,通过对于 G. 729B 系统的扩展,符合了 G. 729. 1 系统指标的要求。

[0147] 再次,通过灵活准确的背景噪声特征参数的提取,使得背景噪声的编码更加精确。

[0148] 在本发明的各个实施例中,如果为了适应 G. 729. 1 相关技术标准的要求,可以通过将每个超帧设置为 20 毫秒,将每个超帧所包含的帧设置为 10 毫秒。利用本发明的各个实施例,均可以实现对于 G. 729B 的扩展,满足 G. 729. 1 的技术指标。同时,对于所属领域的技术人员而言,可以明白,对于非 G. 729. 1 系统,应用本发明的各个实施例所提供的技术方案,同样可以达到对背景噪声以较低的频带占用,带来较高的通信质量。即本发明的应用范围不仅仅局限于 G. 729. 1 系统之内。

[0149] 下面结合附图,详细介绍本发明的编码的方法的实施例二:

[0150] 由于 G729. 1 和 G729B 的编码帧长不同,其中,前者是 20ms 一帧,后者是 10ms 一帧。也就是说, G729. 1 的一帧与 G729B 的两帧的长度相对应。为描述方便,这里将 G729. 1 的一帧为超帧 (superframe), G729B 的一帧为帧 (frame), 本发明主要针对这种差别来描述 G729. 1 的 DTX/CNG 系统,即通过对 G729B DTX/CNG 系统进行升级和扩展,以适应 ITU729. 1 的系统特点。

[0151] 一、噪声的学习:

[0152] 首先,用语音编码速率对背景噪声的前 120ms 进行编码;

[0153] 为了准确的对背景噪声的特征参数进行提取,在语音帧结束(根据 VAD 结果的指示,表明当前帧已经从活动的语音变为非活动的背景噪声)之后一段时间内,并不马上进入背景噪声处理阶段,而是继续用语音编码速率对背景噪声进行编码。这一拖尾的时间一般为 6 个超帧,即 120ms(可参考 AMR 及 AMRWB)。

[0154] 其次,在这一拖尾时间内,对每一个超帧的每一 10ms 帧,均对背景噪声的自相关系数 $R'_{t,k}(j)$, $j = 0,10$ 进行缓存,其中 t 为超帧序号, $k = 1,2$ 为每个超帧中第 1 个和第 2 个 10ms 帧的序号。由于这些自相关系数表征了拖尾阶段背景噪声的特性,因此在对背景噪声进行编码时就可以依据这些自相关系数,准确的提取出背景噪声的特征参数,从而使对背景噪声的编码更加精确。在实际运用中,噪声学习所持续的时间长度,可以根据实际需要设定,不限于 120ms;可以根据需要将拖尾时间设置为其他的数值。

[0155] 二、对拖尾阶段后第一个超帧进行编码

[0156] 在拖尾阶段结束之后,采用背景噪声的处理方式对背景噪声进行处理。参考图 5 所示,是对第一个超帧编码的流程示意图,包括步骤:

[0157] 在拖尾阶段结束后的第一个超帧,对噪声学习阶段以及当前超帧提取出来的背景噪声特征参数进行编码,得到第一个 SID 超帧,由于在拖尾阶段之后的第一个超帧要进行背景噪声参数的编码传输,因此这一超帧一般称为第一个 SID 超帧;编码后生成的第一个 SID 超帧在发送到解码端后进行解码。由于一个超帧对应两个 10ms 帧,为了准确获得编码参数,会在第 2 个 10ms 帧提取出背景噪声的特征参数 $A_t(z)$ 和 E_t :

[0158] LPC 滤波器 $A_t(z)$ 和残差能量 E_t 的计算方式如下:

[0159] 步骤 501、计算缓存中所有自相关系数的平均值：

$$[0160] \quad R^t(j) = \frac{1}{2 * N_{cur}} \sum_{i=t-N_{cur}+1}^t \sum_{k=1}^2 r'_{i,k}(j), j = 0 \dots 10$$

[0161] 其中 $N_{cur} = 5$ ，也就是缓存的大小为 10 个 10ms 帧。

[0162] 步骤 502、由自相关系数的平均值 $R^t(j)$ ，根据 Levinson-Durbin 算法计算出 LPC 滤波器 $A_t(z)$ ，其系数为 $a_t(j)$ ， $j = 0, \dots, 10$ ，同时 Levinson-Durbin 算法也会计算出残差能量 E_t ，并以此作为当前超帧能量参数的简单估计。

[0163] 其中，在实际运用中，为了获得更加稳定的超帧能量参数估计，还可以对估计出的残差能量 E_t 进行了长时的平滑，并将平滑后的能量估计 E_{LT} 作为当前超帧能量参数的最终估计重新赋给 E_t ，平滑操作如下：

$$[0164] \quad E_{LT} = \alpha E_{LT} + (1 - \alpha) E_t$$

$$[0165] \quad E_t = E_{LT}$$

[0166] 其中， α 取值范围为： $0 < \alpha < 1$ ，作为优选实施例中， α 取值可以为 0.9。也可以根据需要设置为其他值。

[0167] 步骤 503、算法将 LPC 滤波器系数 $A_t(z)$ 转化到 LSF 域，然后进行量化编码；

[0168] 步骤 504、残差能量参数 E_t 的量化在对数域完成，采用的是线性的量化。

[0169] 在背景噪声窄带部分的编码完成之后，将这些编码比特封闭在 SID 帧中传送到解码端，这样就完成了第一个 SID 帧窄带部分的编码。

[0170] 在本发明的实施例中，对于第一个 SID 帧窄带部分的编码充分考虑了拖尾阶段背景噪声的特性，将背景噪声在拖尾阶段的特性反映在了编码参数中，从而使得这些编码参数最大限度地表征了当前背景噪声的特性。因此，本发明的实施例中的参数提取相对于 G. 729B 更加准确合理。

[0171] 三、DTX 的判决

[0172] 为了清楚描述的需要，设提取的参数的表现形式是 $PARA_{t,k}$ ，其中 t 为超帧序号，“ $k = 1, 2$ ”为每个超帧中第 1 个和第 2 个 10ms 帧的序号。那么对于除了第一个超帧之外的其他非语音超帧，需要对每一 10ms 帧的参数进行提取和 DTX 判决。

[0173] 参考图 6 所示，是窄带部分参数提取及 DTX 判决的流程图，包括步骤：

[0174] 首先，进行第一个超帧后的第一个 10 毫秒帧的背景噪声参数提取及 DTX 判决；

[0175] 对于所述第一个 10 毫秒帧，背景噪声的谱参数 $A_{t,1}(z)$ 和激励能量参数 $E_{t,1}$ 的计算方式如下：

[0176] 步骤 601、根据最近四个相邻 10ms 帧自相关系数 $r'_{t,1}(j)$ 、 $r'_{(t-1),2}(j)$ 、 $r'_{(t-1),1}(j)$ 和 $r'_{(t-2),2}(j)$ 的值，计算当前自相关系数的稳态平均值 $R^{t-1}(j)$ ：

$$[0177] \quad R^{t-1}(j) = 0.5 * r_{min 1}(j) + 0.5 * r_{min 2}(j), j = 0 \dots 10$$

[0178] 其中， $r_{min 1}(j)$ 和 $r_{min 2}(j)$ 表示 $r'_{t,1}(j)$ 、 $r'_{(t-1),2}(j)$ 、 $r'_{(t-1),1}(j)$ 和 $r'_{(t-2),2}(j)$ 中具有次最小和次次最小自相关系数范数值的自相关系数，也就是除去具有最大和最小自相关系数范数值所剩下的具有中间自相关系数范数值的两个 10ms 帧的自相关系数：

[0179] $r'_{t,1}(j)$ 、 $r'_{(t-1),2}(j)$ 、 $r'_{(t-1),1}(j)$ 和 $r'_{(t-2),2}(j)$ 的自相关系数范数分别为：

$$[0180] \quad norm_{t,1} = \sum_{j=0}^{10} r'_{t,1}(j)^2$$

$$[0181] \quad norm_{(t-1),2} = \sum_{j=0}^{10} r_{(t-1),2}^2(j)$$

$$[0182] \quad norm_{(t-1),1} = \sum_{j=0}^{10} r_{(t-1),1}^2(j)$$

$$[0183] \quad norm_{(t-2),2} = \sum_{j=0}^{10} r_{(t-2),2}^2(j)$$

[0184] 将这四个自相关系数范数值进行排序,则 $r_{\min 1}(j)$ 和 $r_{\min 2}(j)$ 对应于具有中间自相关系数范数值大小的两个 10ms 帧的自相关系数。

[0185] 步骤 602、由当前自相关系数的稳态平均值 $R^{t-1}(j)$, 根据 Levinson-Durbin 算法计算出背景噪声的 LPC 滤波器 $A_{t,1}(z)$, 其系数为 $a_t(j)$, $j = 0, \dots, 10$, 同时 Levinson-Durbin 算法也会计算出残差能量 $E_{t,1}$;

[0186] 其中, 在实际运用中, 算法为了获得更加稳定的帧能量估计, 对估计出的 $E_{t,1}$, 还可以进行长时的平滑, 并将平滑后的能量估计 E_{LT} 作为当前帧激励能量估计重新赋值给 $E_{t,1}$ 操作如下:

$$[0187] \quad E_{LT1} = \alpha E_{LT} + (1 - \alpha) E_{t,1}$$

$$[0188] \quad E_{t,1} = E_{LT1}$$

[0189] α 取值为 0.9。

[0190] 步骤 603、参数提取之后, 进行当前 10ms 帧的 DTX 判决; DTX 判决的具体内容是:

[0191] 算法将之前的 SID 超帧 (SID 超帧就是经 DTX 判决后最终要编码发送的背景噪声超帧, 如果 DTX 判决结果, 该超帧不发送, 则不称之为 SID 超帧) 中窄带部分编码参数与当前 10 毫秒帧相应的编码参数进行对比, 如果当前的 LPC 滤波器系数与之前 SID 超帧中的 LPC 滤波器系数差别较大, 或者当前的能量参数与之前 SID 超帧中的能量参数差别较大 (见下面的公式), 则将当前 10ms 帧的参数变化标志 $flag_change_first$ 置 1, 否则清零。本步骤中的具体的确定方法与 G. 729B 类似:

[0192] 首先, 设上一个 SID 超帧中的 LPC 滤波器 $A_{sid}(z)$ 的系数为 $a_{sid}(j)$, $j = 0 \dots 10$, 如果当前 10ms 帧和上一 SID 超帧的 LPC 滤波器的 Itakura 距离超过了一定的门限, 就令 $flag_change_first$ 置 1, 否则置零:

$$[0193] \quad \text{if} \quad \left(\sum_{j=0}^{10} R_a(j) \times R^{t,1}(j) > E_{t,1} \times thr \right)$$

$$[0194] \quad flag_change_first = 1$$

[0195] else

$$[0196] \quad flag_change_first = 0$$

[0197] 其中, thr 是具体的门限值, 一般在 1.0 到 1.5 之间, 本实施例中为 1.342676475, $R_a(j)$, $j = 0 \dots 10$ 是上一 SID 超帧 LPC 滤波器系数的自相关系数:

$$[0198] \quad \begin{cases} R_a(j) = 2 \sum_{k=0}^{10-j} a_{sid}(k) \times a_{sid}(k+j) & \text{if } (j \neq 0) \\ R_a(0) = \sum_{k=0}^{10} a_{sid}(k)^2 \end{cases}$$

[0199] 其次, 计算当前 10ms 帧和最近三个 10ms 帧共四个 10ms 帧残差能量的平均值:

[0200] $E_{t,1} = (E_{t,1} + E_{t-1,2} + E_{t-1,1} + E_{t-2,2}) / 4$

[0201] 需要注意的是,如果当前超帧是噪声编码阶段的第二个超帧(即前一个超帧是第一个超帧),那么 $E_{t-2,2}$ 的值为 0。对 $E_{t,1}$ 用对数量化器量化。将解码之后的对数能量 $E_{q,1}$ 与上一 SID 超帧解码后的对数能量 E_q^{sid} 进行比较,如果二者的差超过 3dB,就令 flag_change_first 置一,否则置零:

[0202] *if* $abs(E_q^{sid} - E_{q,1}) > 3$

[0203] flag_change_first = 1

[0204] *else*

[0205] flag_change_first = 0

[0206] 对于所属领域的技术人员而言,可以根据实际需要两个激励能量的差设置为其他的值,这没有超出本发明的保护范围。

[0207] 在进行了第一个 10ms 帧的背景噪声参数提取及 DTX 判决之后,就进行第二个 10ms 帧的背景噪声参数提取及 DTX 判决。

[0208] 第二个 10ms 帧的背景噪声参数提取及 DTX 判决流程与第一个 10ms 帧一致,其中第二个 10ms 帧的相关参数为:相邻四个 10ms 帧自相关系数的稳态平均值 $R^{t,2}(j)$,相邻四个 10ms 帧帧能量的平均值 $E_{t,2}$ 以及第二个 10ms 帧的 DTX 标志 flag_change_second。

[0209] 四、当前超帧中的窄带部分背景噪声参数提取及 DTX 判决。

[0210] 参考图 7 所示,是当前超帧中的窄带部分背景噪声参数提取及 DTX 判决流程图,包括步骤:

[0211] 步骤 701、确定当前超帧窄带部分的最终 DTX 标志 flag_change,其确定方式如下:

[0212] $flag_change = flag_change_first || flag_change_second$

[0213] 即只要有一个 10ms 帧的 DTX 判决结果为 1,则当前超帧窄带部分的最终判决结果为 1。

[0214] 步骤 702、确定当前超帧的最终 DTX 判决结果;包括当前超帧高频带部分在内的当前超帧的最终 DTX 判决结果,则还要考虑高频带部分的特性,由窄带部分和高频带部分综合决定当前超帧的最终 DTX 判决结果。如果当前超帧的最终 DTX 判决结果为 1,则进入步骤 703;如果当前超帧的 DTX 判决结果为 0,则不进行编码,只向解码端发送没有任何数据的 NODATA 帧。

[0215] 步骤 703、如果当前超帧的最终 DTX 判决结果为 1,则对当前超帧的背景噪声特征参数进行提取;提取当前超帧的背景噪声特征参数的来源是当前两个 10ms 帧的参数,即将当前两个 10ms 帧的参数进行平滑得到当前超帧的背景噪声编码参数。提取背景噪声特征参数并进行背景噪声特征参数平滑的过程如下:

[0216] 首先,确定平滑因子 smooth_rate:

[0217] *if* (flag_change_first == 0 && flag_change_second == 1)

[0218] smooth_rate = 0.1

[0219] *else*

[0220] smooth_rate = 0.5

[0221] 即:如果第一个 10ms 帧的 DTX 判决结果为 0,而第二个 10ms 帧的 DTX 判决结果为

1,则在平滑时,第一个 10ms 帧背景噪声特征参数的平滑权重为 0.1,第二个 10ms 帧背景噪声特征参数的平均权重为 0.9,否则两个 10ms 帧背景噪声特征参数的平滑权重都为 0.5。

[0222] 然后,对两个 10ms 帧的背景噪声特征参数进行平滑,得到当前超帧的 LPC 滤波器系数和计算两个 10ms 帧帧能量的平均值,其过程包括:

[0223] 首先,计算两个 10ms 帧自相关系数稳态平均值的滑动平均值 $R^t(j)$:

$$[0224] \quad R^t(j) = \text{smooth_rate}R^{t-1}(j) + (1-\text{smooth_rate})R^{t-2}(j)$$

[0225] 得到自相关系数的滑动平均值 $R^t(j)$ 之后,根据 Levinson-Durbin 算法,得到 LPC 滤波器 $A_t(z)$,其系数为 $a_t(j)$, $j = 0, \dots, 10$;

[0226] 其次,计算两个 10ms 帧帧能量的平均值 E_t :

$$[0227] \quad E_t = \text{smooth_rate}E_{t,1} + (1-\text{smooth_rate})E_{t,2}$$

[0228] 这样就得到当前超帧的窄带部分的编码参数:LPC 滤波器系数和帧能量平均值。背景噪声特征参数提取以及 DTX 控制充分靠了了当前超帧每一 10ms 帧的特性,因此算法较为严谨。

[0229] 五、SID 帧的编码

[0230] 与 G. 729B 一样,在对 SID 帧的谱参数最终编码时,考虑了相邻噪声帧之间的稳定情况,具体的操作与 G. 729B 一致:

[0231] 首先,计算当前超帧之前 N_p 个超帧的平均 LPC 滤波器 $A_p(z)$,这要用到自相关函数平均值 $R_p(j)$,然后将 $R_p(j)$ 送入 Levinson-Durbin 算法中得到 $A_p(z)$,而 $R_p(j)$ 表示为:

$$[0232] \quad \bar{R}_p(j) = \frac{1}{2 * N_p} \sum_{i=t-1-N_p}^{t-1} \sum_{k=1}^2 r_{i,k}(j), j = 0 \dots 10$$

[0233] 其中 N_p 的值定为 5。这样,SID-LPC 滤波器表示为:

$$[0234] \quad A_{sid}(z) = \begin{cases} A_t(z) & \text{if } \text{distance}(A_t(z), \bar{A}_p(z)) > \text{thr3} \\ \bar{A}_p(z) & \text{otherwise} \end{cases}$$

[0235] 即算法会计算前几个超帧的平均 LPC 滤波器系数 $A_p(z)$ 然后,用其与当前的 LPC 滤波器系数 $A_t(z)$ 进行比较,若二者差别较小,那么当前超帧在量化 LPC 系数时选择的就是前几个超帧的平均值 $A_p(z)$,否则就是当前超帧的 $A_t(z)$,具体的比较方法与步骤 602 中 10ms 帧 DTX 判决的方法一样,其中 thr3 是具体的门限值,一般在 1.0 到 1.5 之间,本实施例中为 1.0966466。所述领域的技术人员可以根据实际需要,取其他的值,这没有超出本发明的保护范围。

[0236] 选择好 LPC 滤波器系数之后,算法将这些 LPC 滤波器系数转化到 LSF 域,然后进行量化编码,而量化编码选择的方式与 G. 729B 的量化编码方式类似。

[0237] 能量参数的量化在对数域完成,采用的是线性的量化,然后进行编码。这样对背景噪声的编码就已完成,然后将这些编码比特封装在 SID 帧中。

[0238] 六、CNG 的方式

[0239] 在基于 CELP 模型的编码中,为了获得最佳的编码参数,在编码端也包含了解码的过程,对于 CNG 系统也不例外,即在 G. 729. 1 中编码端也要包含 CNG 模块。对于 G. 729. 1 中的 CNG 来说,其处理的流程基于 G. 729B,虽然其帧长为 20ms,但还是以 10ms 为基本的数据处理长度来处理背景噪声。但是,从上一节可知,第一个 SID 超帧的编码参数在第 2 个 10ms

帧才会对进行编码,但这时系统需要在第一个 SID 超帧的第一个 10ms 帧就产生 CNG 的参数。显然,第一个 SID 超帧的第一个 10ms 帧的 CNG 参数不能从 SID 超帧的编码参数中得到,而只能从之前语音编码超帧中得到。由于存在这种特殊情况,因此 G. 729. 1 的第一个 SID 超帧的第一个 10ms 帧的 CNG 方式与 G. 729B 有所不同,相对前述有关内容介绍的 G. 729B 的 CNG 方式,这种不同表现在:

[0240] (1) 目标激励增益 \tilde{G}_f 由长时平滑的语音编码超帧量化的固定码本增益 LT_G_f 定义:

$$[0241] \quad \tilde{G}_f = LT_G_f * \gamma$$

[0242] 其中, $0 < \gamma < 1$, 本实施例中可以选择 $\gamma = 0.4$ 。

[0243] (2) LPC 滤波器系数 $A_{sid}(z)$ 由长时平滑的语音编码超帧量化的 LPC 滤波器系数 $LT_A(z)$ 定义:

$$[0244] \quad A_{sid}(z) = LT_A(z)$$

[0245] 其他的操作与 729B 一致。

[0246] 设语音编码帧量化的固定码本增益和 LPC 滤波器系数分别为 $gain_code$ 和 $A_q(z)$, 则这些长时平滑的参数分别计算如下:

$$[0247] \quad LT_G_f = \beta LT_G_f + (1 - \beta) gain_code$$

$$[0248] \quad LT_A(z) = \beta LT_A(z) + (1 - \beta) A_q(z)$$

[0249] 以上操作在语音超帧的每一子帧都进行平滑,其中平滑因子 β 的取值范围为 $0 < \beta < 1$, 本实施例中为 0.5。

[0250] 另外,除了第一个 SID 超帧的第一个 10ms 帧与 729B 略有不同外,其他所有 10ms 帧的 CNG 方式均与 G. 729B 一致。

[0251] 其中,在上述实施例中,所述拖尾时间为 120 毫秒或 140 毫秒。

[0252] 其中,在上述实施例中,所述提取拖尾时间内的背景噪声特征参数具体为:

[0253] 在所述拖尾时间内,对每一个超帧的每一帧,保存每帧背景噪声的自相关系数。

[0254] 其中,在上述实施例中,对所述拖尾时间后的第一个超帧,所述根据提取的所述拖尾时间内的背景噪声特征参数和所述第一个超帧的背景噪声特征参数,进行背景噪声编码包括:

[0255] 在第一帧和第二帧保存每个帧背景噪声的自相关系数;

[0256] 在第二帧,根据所述提取出的所述两帧的自相关系数和所述拖尾时间内的背景噪声特征参数,提取出所述第一个超帧的 LPC 滤波器系数和残差能量,进行背景噪声编码。

[0257] 其中,在上述实施例中,所述提取所述 LPC 滤波器系数具体为:

[0258] 计算所述第一个超帧和所述第一个超帧之前的所述拖尾时间内的四个超帧的自相关系数的平均值;

[0259] 由所述自相关系数的平均值,根据 Levinson-Durbin 算法计算出 LPC 滤波器系数;

[0260] 所述提取所述残差能量 E_i 具体为:

[0261] 根据 Levinson-Durbin 算法计算出残差能量;

[0262] 所述在第二帧进行背景噪声编码具体为:

[0263] 将所述 LPC 滤波器系数转化到 LSF 域,进行量化编码;

- [0264] 将所述残差能量在对数域进行线性量化编码。
- [0265] 其中,在上述实施例中,在计算出所述残差能量后,进行量化编码之前,还包括:
- [0266] 对所述残差能量进行长时的平滑;
- [0267] 平滑公式为: $E_{LT} = \alpha E_{LT} + (1 - \alpha)E_t$, α 取值范围为: $0 < \alpha < 1$;
- [0268] 将平滑后的能量估计 E_{LT} 的值作为残差能量的值。
- [0269] 其中,在上述实施例中,所述对第一个超帧后的超帧,对每一帧都进行背景噪声特征参数提取具体为:
- [0270] 根据最近四个相邻帧自相关系数的值,计算当前自相关系数的稳态平均值,所述自相关系数的稳态平均值是所述最近四个相邻帧中具有中间自相关系数范数值的两帧的自相关系数的平均值;
- [0271] 对所述稳态平均值,根据 Levinson-durbin 算法计算背景噪声 LPC 滤波器系数和残差能量。
- [0272] 其中,在上述实施例中,在计算所述残差能量后,还包括:
- [0273] 对所述残差能量进行长时平滑,得到当前帧能量估计;平滑方式为:
- [0274] $E_{LT} = \alpha E_{LT1} + (1 - \alpha)E_{t,k}$;
- [0275] α 取值为: $0 < \alpha < 1$;
- [0276] 将平滑后的当前帧能量估计赋值给所述残差能量;赋值方式为:
- [0277] $E_{t,k} = E_{LT}$ 。
- [0278] 其中 $k = 1, 2$, 分别表示第一帧和第二帧。
- [0279] 其中,各个实施例中: $\alpha = 0.9$ 。
- [0280] 其中,在上述实施例中,所述对第一个超帧后的超帧,对每一帧都进行 DTX 判决具体为:
- [0281] 如果当前帧 LPC 滤波器系数和上一 SID 超帧 LPC 滤波器系数的值超过预设的门限值,或者当前帧的能量估计与上一 SID 超帧中的能量估计相比差别较大,则将当前帧的参数变化标志设为 1;
- [0282] 如果当前 10 毫秒帧 LPC 滤波器系数和上一 SID 超帧 LPC 滤波器系数的值没有超过预设的门限值,或者当前 10 毫秒帧的能量估计与上一 SID 超帧中的能量估计相比差别不大,则将当前 10 毫秒帧的参数变化标志设为 0。
- [0283] 其中,在上述实施例中,所述当前帧的能量估计与前一 SID 超帧中的能量估计相比差别较大具体为:
- [0284] 计算当前 10 毫秒帧和之前最近 3 个帧共 4 个帧的残差能量的平均值作为当前帧的能量估计;
- [0285] 将所述残差能量的平均值使用对数量化器量化;
- [0286] 如果解码后的对数能量与上一 SID 超帧解码后的对数能量的差超过预设值,则确定所述当前帧的能量估计与前一 SID 超帧中的能量估计相比差别较大。
- [0287] 其中,在上述实施例中,所述对每一个帧都进行 DTX 判决具体为:
- [0288] 如果当前超帧中有一个帧的 DTX 判决结果为 1,则当前超帧窄带部分的 DTX 判决结果为 1。
- [0289] 其中,在上述实施例中,述当前超帧的所述最终 DTX 判决结果为 1,则所述:“对第

一个超帧后的超帧,根据提取的当前超帧的背景噪声特征参数和所述当前超帧之前若干超帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码”过程包括:

[0290] 对于所述当前超帧,确定平滑因子,包括:

[0291] 如果当前超帧第一帧的 DTX 为零,第二帧的 DTX 为 1,则所述平滑因子为 0.1,否则所述平滑因子为 0.5;

[0292] 对所述当前超帧的两个帧进行参数平滑,将参数平滑后的参数作为对所述当前超帧的进行背景噪声编码的特征参数,所述参数平滑包括:

[0293] 计算所述两个帧自相关系数稳态平均值的滑动平均值 $R^t(j)$:

[0294] $R^t(j) = \text{smooth_rate}R^{t+1}(j) + (1 - \text{smooth_rate})R^{t+2}(j)$,所述 smooth_rate 为所述平滑因子, $R^{t+1}(j)$ 为第一帧的自相关系数稳态平均值, $R^{t+2}(j)$ 为第二帧的自相关系数稳态平均值;

[0295] 对所述两个帧自相关系数稳态平均值的滑动平均值 $R'(j)$,根据 Levinson-Durbin 算法,得到 LPC 滤波器系数,

[0296] 计算所述两个帧能量估计的滑动平均值 E_t :

[0297] $E_t = \text{smooth_rate}E_{t,1} + (1 - \text{smooth_rate})E_{t,2}$,所述 $E_{t,1}$ 为第一帧的能量估计, $E_{t,2}$ 为第二帧的能量估计。

[0298] 其中,在上述实施例中,所述“根据提取的当前超帧的背景噪声特征参数和所述当前超帧之前若干超帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码”具体为:

[0299] 计算当前超帧之前若干超帧的自相关系数的平均值;

[0300] 根据所述自相关系数的平均值,计算当前超帧之前若干个超帧的平均 LPC 滤波器系数;

[0301] 如果所述平均 LPC 滤波器系数与当前超帧的 LPC 滤波器系数差值小于或等于预设值,则将所述平均 LPC 滤波器系数转化到 LSF 域,进行量化编码;

[0302] 如果所述平均 LPC 滤波器系数与当前超帧的 LPC 滤波器系数差值大于预设值,则将所述当前超帧的 LPC 滤波器系数转化到 LSF 域,进行量化编码;

[0303] 对能量参数,在对数域进行线性量化编码。

[0304] 其中,在上述实施例中,所述若干帧的数量为 5。所属领域的技术人员也可以根据需要选择其他数量的帧数。

[0305] 其中,在上述实施例中,在所述提取拖尾时间内的背景噪声特征参数步骤之前,还包括:

[0306] 用语音编码速率对所述拖尾时间内的背景噪声进行编码。

[0307] 参考图 8 所示,是本发明的解码方法的实施例一,包括步骤:

[0308] 步骤 801、从第一个超帧的第一帧之前的语音编码帧,获得第一个超帧的第一帧的 CNG 参数;

[0309] 步骤 802、根据所述 CNG 参数,对第一个超帧的第一帧进行背景噪声解码,所述 CNG 参数包括:

[0310] 目标激励增益,所述目标激励增益由长时平滑的语音编码帧参数量化的固定码本增益确定;

[0311] 其中,在实际运用中,所述确定目标增益可以具体为:目标激励增益 = γ * 固定码本增益, $0 < \gamma < 1$;

[0312] 滤波器系数,所述滤波器系数由长时平滑的语音编码帧参数量化的滤波器系数定义;

[0313] 其中,在实际运用中,所述定义所述滤波器系数可以具体为:

[0314] 滤波器系数 = 长时平滑的语音编码帧量化的滤波器系数。

[0315] 其中,上述实施例中,所述长时平滑因子取值范围为:大于 0 且小于 1。

[0316] 其中,上述实施例中,所述长时平滑因子可以为 0.5。

[0317] 其中,上述实施例中,所述 $\gamma = 0.4$ 。

[0318] 其中,上述实施例中,所述对第一个超帧的第一帧进行背景噪声解码过程之后,还可以包括:

[0319] 对除所述第一个超帧的第一帧之外的所有帧,从上一步 SID 超帧获取 CNG 参数后,根据获取的所述 CNG 参数进行背景噪声解码。

[0320] 参考图 9 所示,是本发明的编码装置的实施例一,包括:

[0321] 第一提取单元 901,用于:提取拖尾时间内的背景噪声特征参数;

[0322] 第二编码单元 902,用于:对所述拖尾时间后的第一个超帧,根据提取的所述拖尾时间内的背景噪声特征参数和所述第一个超帧的背景噪声特征参数,进行背景噪声编码;

[0323] 第二提取单元 903,用于:对第一个超帧后的超帧,对每一帧都进行背景噪声特征参数提取;

[0324] DTX 判决单元 904,用于:对第一个超帧后的超帧,对每一帧都进行 DTX 判决;

[0325] 第三编码单元 905,用于:对第一个超帧后的超帧,根据提取的当前超帧的背景噪声特征参数和所述当前超帧之前若干超帧的背景噪声特征参数,以及最终 DTX 判决结果,进行背景噪声编码。

[0326] 其中,上述实施例中,所述拖尾时间为 120 毫秒或 140 毫秒。

[0327] 其中,上述实施例中,所述第一提取单元具体为:

[0328] 缓存模块,用于:在所述拖尾时间内,对每一个超帧的每一帧,保存每帧背景噪声的自相关系数。

[0329] 其中,上述实施例中,所述第二编码单元具体为:

[0330] 提取模块,用于:在第一帧和第二帧保存每帧背景噪声的自相关系数;

[0331] 编码模块,用于:在第二帧,根据所述提取出的所述两个帧的自相关系数和所述拖尾时间内的背景噪声特征参数,提取出所述第一个超帧的 LPC 滤波器系数和残差能量,进行背景噪声编码。

[0332] 其中,上述实施例中,所述第二编码单元还可以包括:

[0333] 残差能量平滑模块,用于:对所述残差能量进行长时的平滑;

[0334] 平滑公式为: $E_{LT} = \alpha E_{LT} + (1 - \alpha) E_t$, α 取值范围为: $0 < \alpha < 1$;

[0335] 将平滑后的能量估计 E_{LT} 的值作为所述残差能量的值。

[0336] 其中,上述实施例中,所述第二提取单元具体为:

[0337] 第一计算模块,用于:根据最近四个相邻帧自相关系数的值,计算当前自相关系数的稳态平均值,所述自相关系数的稳态平均值是所述最近四个相邻帧中具有中间自相关系

数范数值的一个帧的自相关系数的平均值；

[0338] 第二计算模块,用于:对所述稳态平均值,根据 Levinson-durbin 算法计算背景噪声 LPC 滤波器系数和残差能量。

[0339] 其中,上述实施例中,所述第二提取单元还可以包括:

[0340] 第二残差能量平滑模块,用于:对所述残差能量进行长时平滑,得到当前帧能量估计;平滑方式为:

[0341] $E_{LT} = \alpha E_{LT1} + (1-\alpha)E_{t,k}$;

[0342] α 取值为: $0 < \alpha < 1$;

[0343] 将平滑后的当前帧能量估计赋值给所述残差能量;赋值方式为:

[0344] $E_{t,k} = E_{LT}$ 。

[0345] 其中 $k = 1, 2$, 分别表示第一帧和第二帧。

[0346] 其中,上述实施例中,所述 DTX 判决单元具体为:

[0347] 门限比较模块,用于:如果当前帧 LPC 滤波器系数和上一 SID 超帧 LPC 滤波器系数的值超过预设的门限值,则生成判决指令;

[0348] 能量比较模块,用于:计算当前帧和之前最近 3 个帧共 4 个帧的残差能量的平均值作为当前帧的能量估计,将所述残差能量的平均值使用对数量化器量化,如果解码后的对数能量与上一 SID 超帧解码后的对数能量的差超过预设值,则生成判决指令;

[0349] 第一判决模块,用于:根据所述判决指令,将当前帧的参数变化标志设为 1。

[0350] 其中,上述实施例中,还可以包括:

[0351] 第二判决单元,用于:如果当前超帧中有一个帧的 DTX 判决结果为 1,则当前超帧窄带部分的 DTX 判决结果为 1;

[0352] 所述第三编码单元具体为:

[0353] 平滑指示模块,用于:如果当前超帧的所述最终 DTX 判决结果为 1,则生成平滑指令;;

[0354] 平滑因子确定模块,用于:收到所述平滑指令后,确定所述当前超帧的平滑因子:

[0355] 如果当前超帧第一帧的 DTX 为零,第二帧的 DTX 为 1,则所述平滑因子为 0.1,否则所述平滑因子为 0.5;

[0356] 参数平滑模块,用于:对所述当前超帧的两个帧进行参数平滑,将平滑后的参数作为对所述当前超帧的进行背景噪声编码的特征参数,包括:

[0357] 计算所述两个帧自相关系数稳态平均值的滑动平均值 $R^t(j)$:

[0358] $R^t(j) = \text{smooth_rate}R^{t-1}(j) + (1-\text{smooth_rate})R^{t-2}(j)$, 所述 smooth_rate 为所述平滑因子, $R^{t-1}(j)$ 为第一帧的自相关系数稳态平均值, $R^{t-2}(j)$ 为第二帧的自相关系数稳态平均值;

[0359] 对所述两个帧自相关系数稳态平均值的滑动平均值 $R^t(j)$, 根据 Levinson-Durbin 算法, 得到 LPC 滤波器系数,

[0360] 计算所述两个帧能量估计的滑动平均值 E_t :

[0361] $E_t = \text{smooth_rate}E_{t,1} + (1-\text{smooth_rate})E_{t,2}$, 所述 $E_{t,1}$ 为第一帧的能量估计, $E_{t,2}$ 为第二帧的能量估计。

[0362] 其中,上述实施例中,所述第三编码单元具体为:

[0363] 第三计算模块,用于:根据计算得到的当前超帧之前若干超帧的自相关系数的平均值,计算当前超帧之前若干个超帧的平均 LPC 滤波器系数;

[0364] 第一编码模块,用于:如果所述平均 LPC 滤波器系数与当前超帧的 LPC 滤波器系数差值小于或等于预设值,则将所述平均 LPC 滤波器系数转化到 LSF 域,进行量化编码;

[0365] 第二编码模块,用于:如果所述平均 LPC 滤波器系数与当前超帧的 LPC 滤波器系数差值大于预设值,则将所述当前超帧的 LPC 滤波器系数转化到 LSF 域,进行量化编码;

[0366] 第三编码模块,用于:对能量参数,在对数域进行线性量化编码。

[0367] 其中,上述实施例中, $\alpha = 0.9$ 。

[0368] 其中,上述实施例中,还可以包括:

[0369] 第一编码单元,用于:用语音编码速率对拖尾时间内的背景噪声进行编码;

[0370] 本发明的编码装置具体与本发明的编码方法相适应的工作过程,相应地,也具有与相应的方法实施例同样的技术效果。

[0371] 参考图 10 所示,是本发明的解码装置的实施例一,包括:

[0372] CNG 参数获得单元 1001,用于:从第一个超帧的第一帧之前的语音编码帧,获得第一个超帧的第一帧的 CNG 参数;

[0373] 第一解码单元 1002,用于:根据所述 CNG 参数,对第一个超帧的第一帧进行背景噪声解码,所述 CNG 参数包括:

[0374] 目标激励增益,所述目标激励增益由长时平滑的语音编码帧量化的固定码本增益确定,其中,在实际运用中,所述目标激励增益确定具体为:

[0375] 目标激励增益 = γ * 固定码本增益,所述 γ 的取值范围是: $0 < \gamma < 1$;

[0376] LPC 滤波器系数,所述 LPC 滤波器系数由长时平滑的语音编码帧量化的 LPC 滤波器系数定义,其中,在实际运用中,所述定义 LPC 滤波器系数可以具体为:

[0377] LPC 滤波器系数 = 长时平滑的语音编码帧量化的 LPC 滤波器系数。

[0378] 其中,在上述实施例中,所述长时平滑因子取值范围为:大于 0 且小于 1。

[0379] 在优选的情况下,所述长时平滑因子可以为 0.5。

[0380] 其中,在上述实施例中,还可以包括:

[0381] 第二解码单元,用于:对除所述第一个超帧之外的所有帧,从上一个 SID 超帧获取 CNG 参数后,根据获取的所述 CNG 进行背景噪声编码。

[0382] 其中,在上述实施例中,所述 $\gamma = 0.4$ 。

[0383] 本发明的解码装置具体与本发明的解码方法相适应的工作过程,相应地,也具有与相应的解码方法实施例同样的技术效果。

[0384] 以上所述的本发明实施方式,并不构成对本发明保护范围的限定。任何在本发明的精神和原则之内所作的修改、等同替换和改进等,均应包含在本发明的保护范围之内。

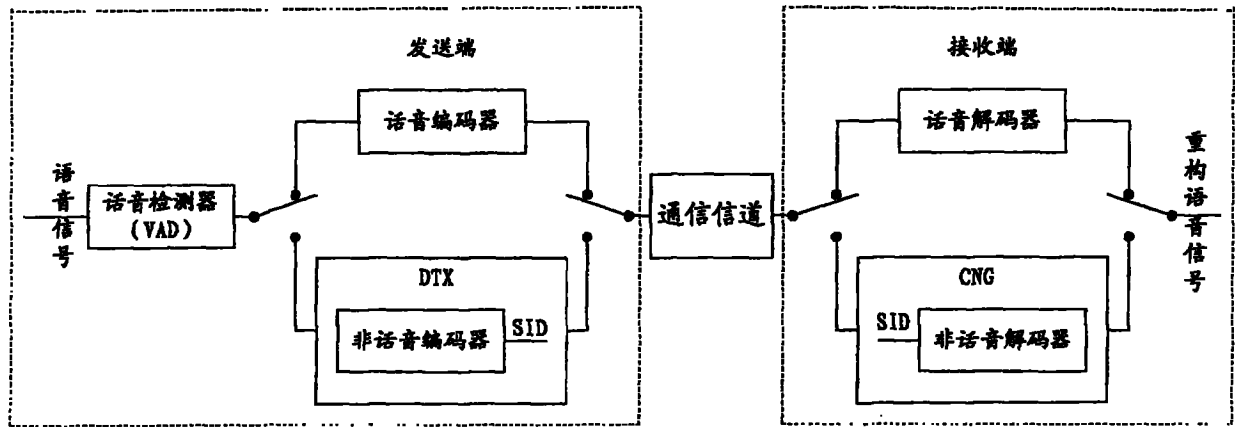


图 1

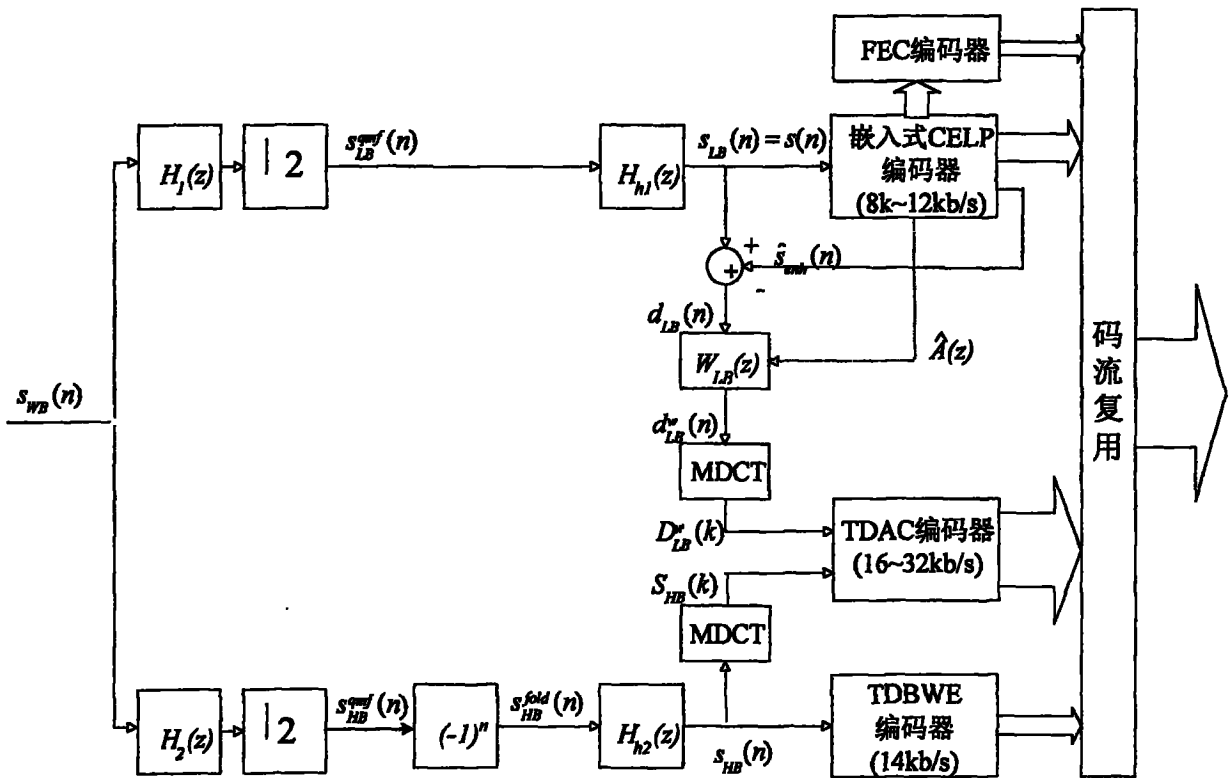


图 2

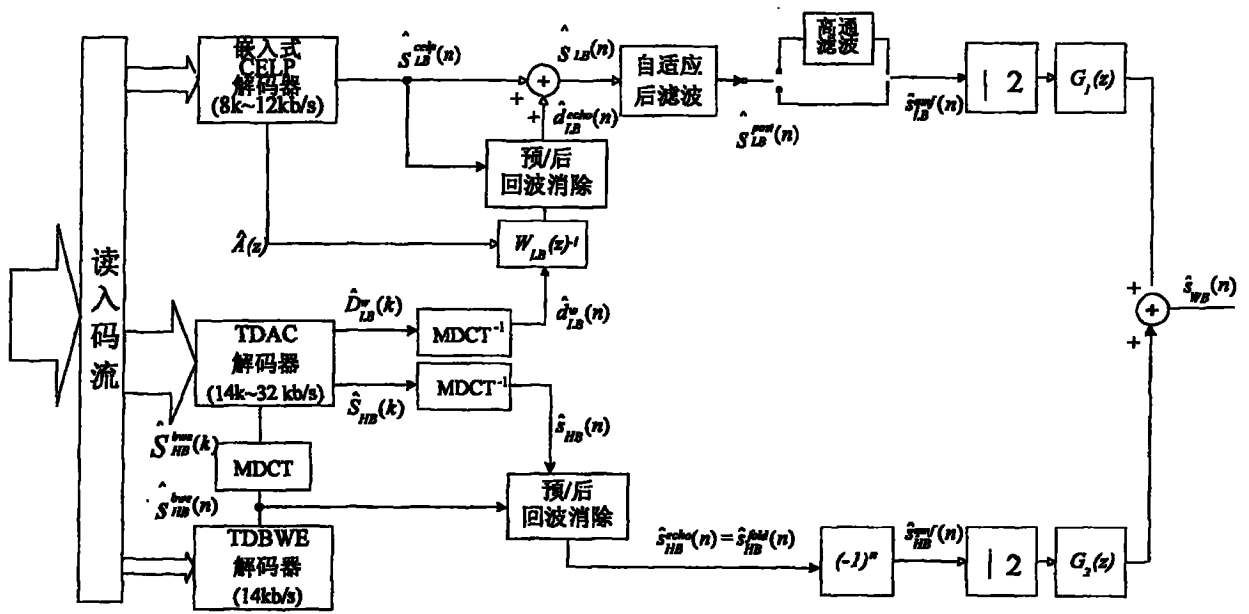


图 3

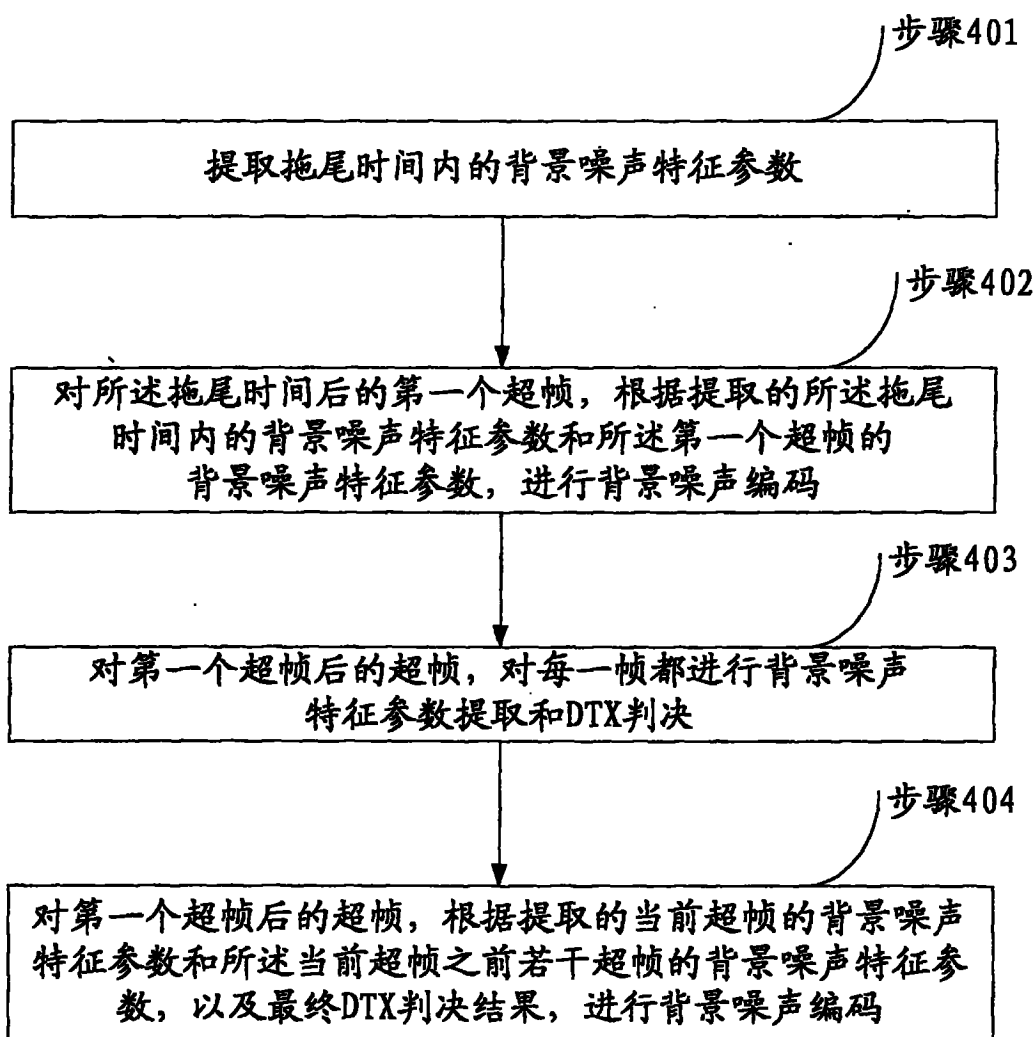


图 4

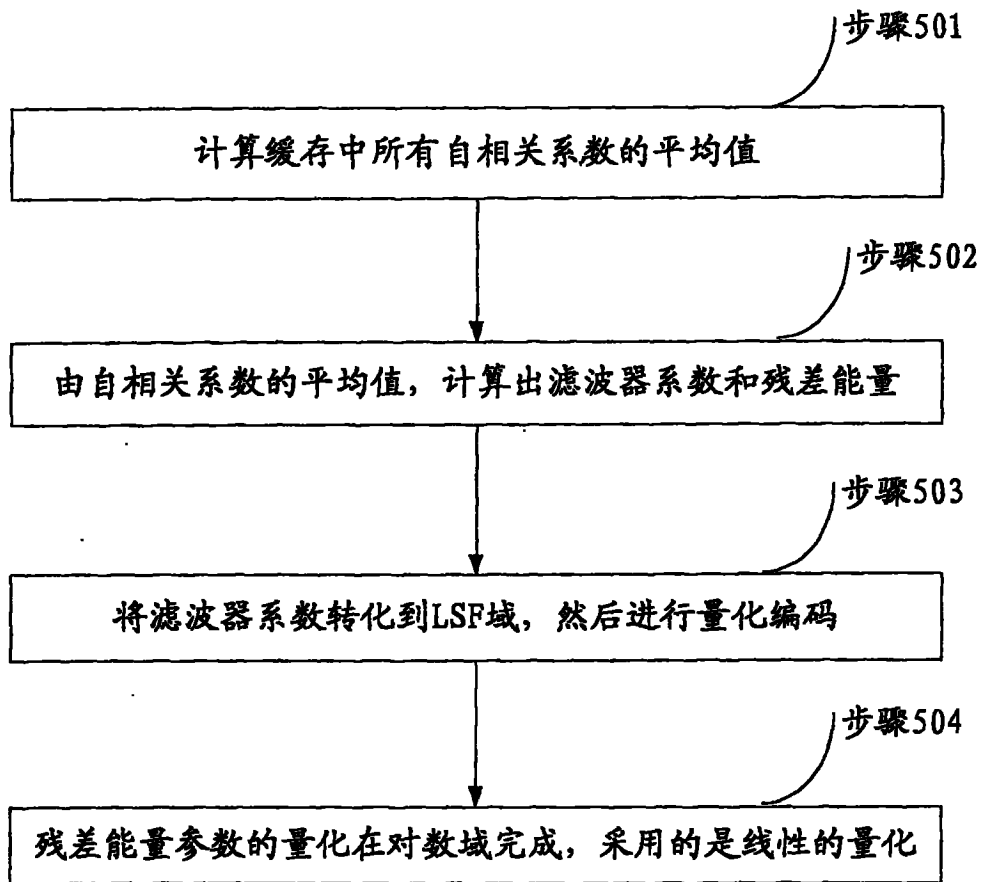


图 5

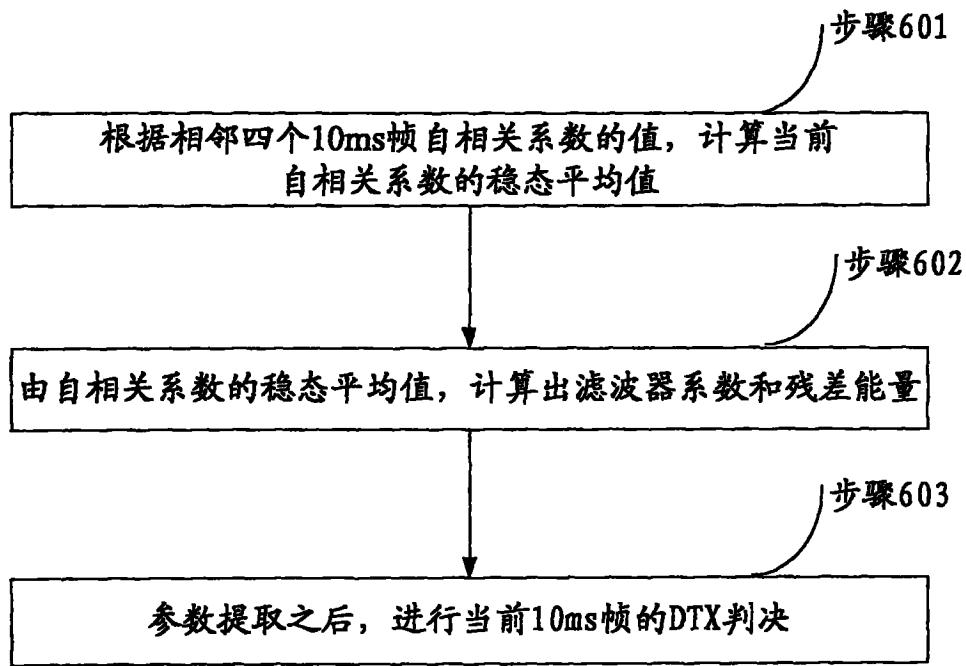


图 6

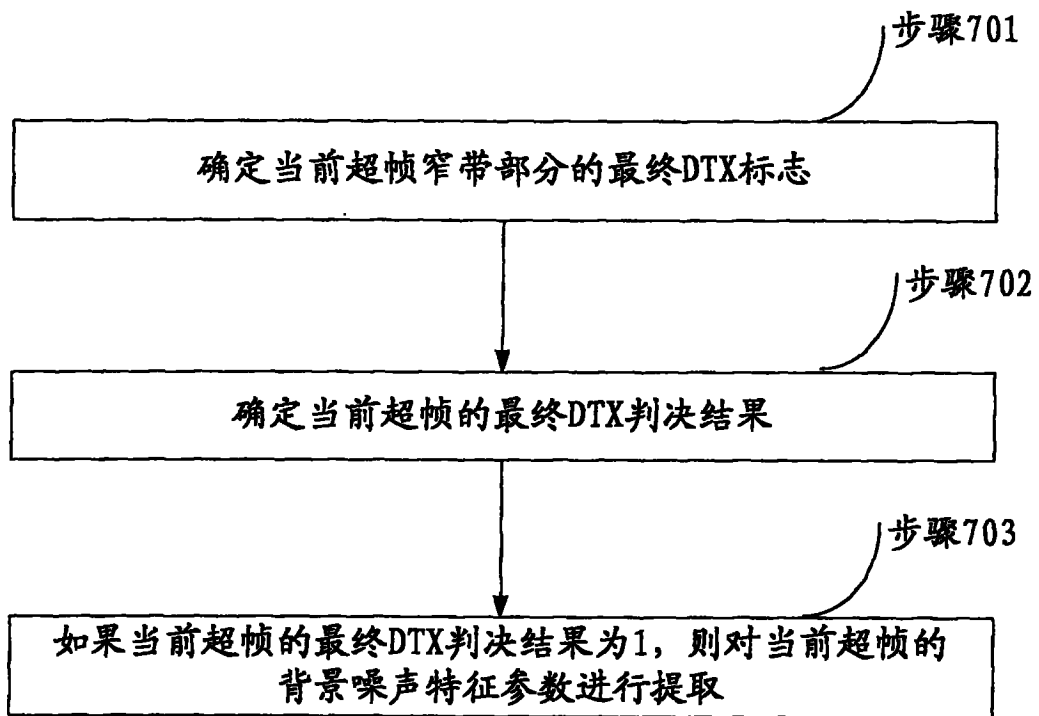


图 7

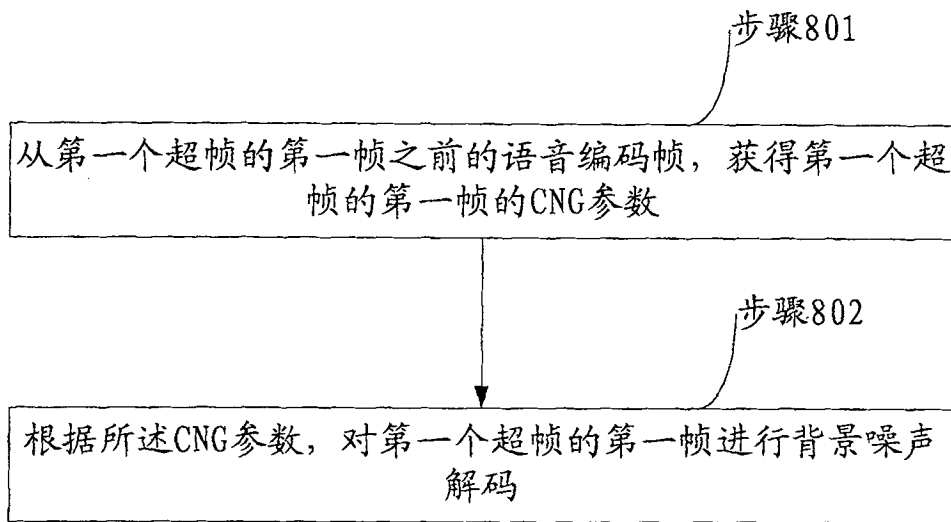


图 8

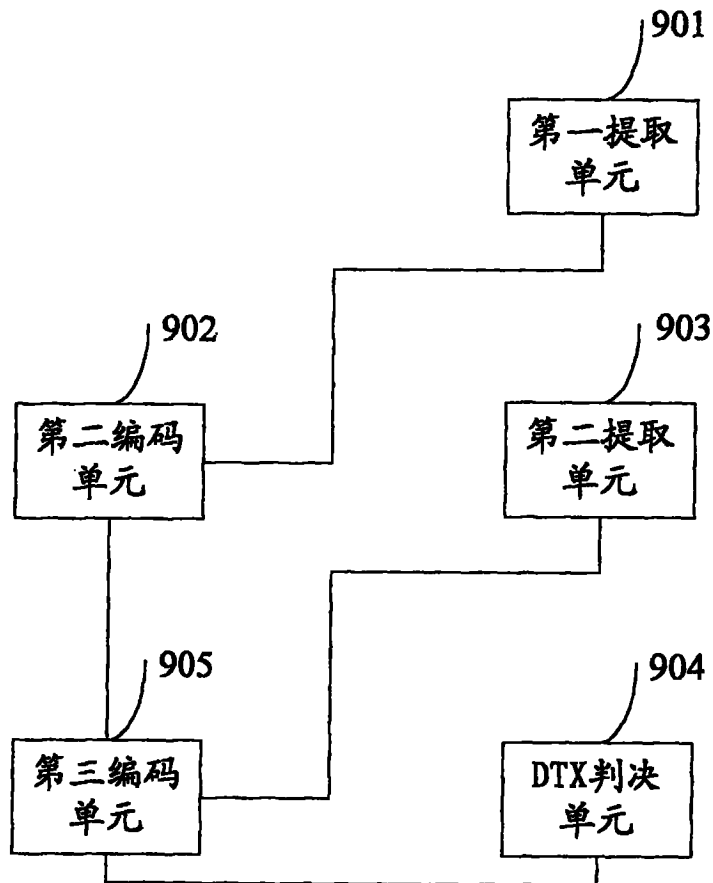


图 9

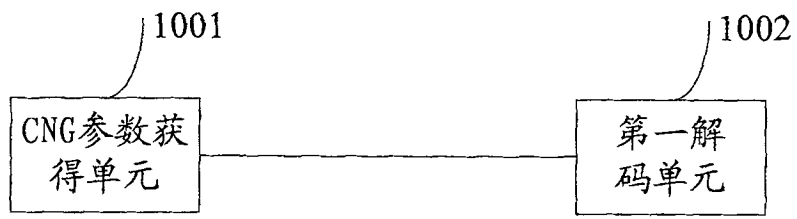


图 10