



[12] 发 明 专 利 说 明 书

[21] ZL 专利号 95190717.4

[45] 授权公告日 2004 年 9 月 22 日

[11] 授权公告号 CN 1168071C

[22] 申请日 1995.8.1 [21] 申请号 95190717.4

[30] 优先权

[32] 1994.8.10 [33] US [31] 288,413

[86] 国际申请 PCT/US1995/009830 1995.8.1

[87] 国际公布 WO1996/005592 英 1996.2.2

[85] 进入国家阶段日期 1996.4.1

[71] 专利权人 夸尔柯姆股份有限公司

地址 美国加州

[72] 发明人 安德鲁·P·德雅克

威廉·R·加德纳

审查员 卞喜双

[74] 专利代理机构 上海专利商标事务所

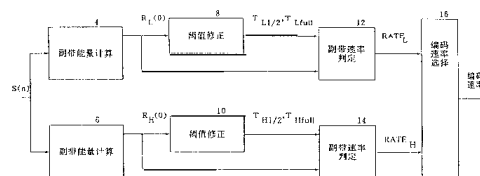
代理人 陈 亮

权利要求书 2 页 说明书 8 页 附图 1 页

[54] 发明名称 在速率可变的声码器中选择编码速率的方法和装置

[57] 摘要

本发明提供一种降低把低能量非语音语音作为背景噪声进行编码的概率的方法。用数字副带滤波器(4)和(6)把输入信号分成副带,在副带速率判定部件(12)和(14)中把这些副带中的能量与一组阈值比较,然后在编码速率选择器(16)内检查这些比较结果,通过这些步骤来确定编码速率。用这种方法,可以把非语音语音与背景噪声区别开。本发明还提供一种用输入信号的信噪比设置阈值电平的装置,本发明还提供一种用速率可变的声码器对音乐进行编码的方法,它通过检查输入信号的周期性以把音乐与背景噪声区别开来。



1.一种为速率可变声码器内的输入信号确定编码速率的装置，其特征在于，包含：

多个副带能量计算部件，用于接收所述输入信号并根据预定的副带能量计算公式确定多个副带能量值；

多个副带速率确定部件，每个副带速率确定部件用于接收所述多个副带能量值中相应的一个副带能量值，并根据所述多个副带能量值中所述相应的一个副带能量值确定一个建议的编码速率，从而提供多个建议的编码速率；以及

编码速率选择部件，用于接收所述多个建议的编码速率，并根据所述多个建议的编码速率选择所述输入信号的编码速率。

2.如权利要求1所述的装置，其特征在于，所述多个副带能量计算部件根据下面的公式确定所述多个副带能量值中的每个副带能量值：

$$\text{副带能量} = R_s(0) \cdot R_{hbp}(0) + 2 \cdot \sum_{i=1}^{L-1} R_s(i) \cdot R_{hbp}(i)$$

其中 L 为带通滤波器 $h_{bp}(n)$ 内的抽头数， $R_s(i)$ 为输入信号 $S(n)$ 的自相关函数， R_{hbp} 为带通滤波器 $h_{bp}(n)$ 的自相关函数。

3.如权利要求1所述的装置，其特征在于，进一步包含多个阈值修正部件，每个阈值修正部件设置在一个副带能量计算部件和一个副带速率确定部件之间，所述多个阈值修正部件接收所述多个副带能量值并根据所述多个副带能量值确定一组编码速率阈值。

4.如权利要求3所述的装置，其特征在于，所述多个阈值修正部件根据所述多个副带能量值确定信噪比。

5.如权利要求4所述的装置，其特征在于，每个阈值修正部件根据一个信噪比指数确定一个换算系数，其中所述信噪比指数是根据所述信噪比的值来确定的。

6.如权利要求5所述的装置，其特征在于，所述多个阈值修正部件通过把背景噪声估计值与所述换算系数相乘来确定至少一个阈值。

7.如权利要求1所述的装置，其特征在于，每个副带速率确定部件把所述相应的副带能量值与至少一个阈值比较以确定所述副带编码速率。

8.如权利要求6所述的装置，其特征在于，每个副带速率确定部件把所述相应的副带能量值与所述至少一个阈值比较以确定所述副带编码速率。

9. 如权利要求1所述的装置，其特征在于，所述多个副带能量值包括：

所述输入信号中的信息信号的能量的估计值；

所述输入信号中的背景噪声的能量的估计值；

其中所述多个副带能量计算部件根据所述信息信号的能量的估计值和所述背景噪声的能量的估计值来产生信噪比。

10.一种确定速率可变声码器中的输入信号的编码速率的方法，其特征在于，包含下列步骤：

接收输入信号；

根据预定的副带能量计算公式确定多个副带能量值；

对所述多个副带能量值中每一个副带能量值确定一个相应的建议的编码速率，以提供多个建议的编码速率；和

根据所述多个建议的编码速率选择所述输入信号的编码速率。

11. 如权利要求10所述的方法，其特征在于，所述根据预定的副带能量计算公式确定多个副带能量值的步骤包括：

对一帧的输入信号生成线性预测编码系数；

生成所述帧的背景噪声估计值；

产生相邻连续帧的平均归一化自相关值，其中所述背景噪声估计值已从预定的初始背景噪声估计值增加；以及

根据所述平均归一化自相关值和预定的阈值确定音乐的存在。

在速率可变的声码器中选择编码速率的方法和装置

技术领域

本发明涉及一种声码器。本发明尤其涉及在速率可变的声码器中确定话音编码速率的新颖的和经改进的方法。

背景技术

速率可变话音压缩系统一般在开始进行编码之前使用一些速率确定算法。这种速率确定算法把较高的比特率编码方法赋予了有话音出现的音频信号段，把较低的比特率编码方法赋予无声段。在这种方法中，可以实现较低的平均比特率，而重新构成的话音仍保持较高质量。因此，为了有效地进行工作，速率可变的话音声码器需要一种健全的速率确定算法，以能在各种背景噪声环境中区别话音和无声。

在1991年6月11日申请的，名称为“速率可变的声码器”的待批美国专利申请 No. 07/713/661 中揭示了这样一种速率可变的话音压缩系统或速率可变的声码器，该专利申请已转让给本发明的受让人，援引在此，以作参考。在这种速率可变的声码器的特定的实现方法中，用码激励线性预测编码技术(CELP)以根据话音活动性的程度确定的几种速率中的一种速率对输入话音进行编码。话音的活动性程度根据除了有声话音之外还可以包含背景噪声的输入音频样值内的能量来确定。为了使声码器在各种背景噪声下都提供高质量的声音编码，需要一种合适的调整阈值的技术来补偿背景噪声对速率判定算法的影响。

声码器一般用在诸如蜂窝电话等通信设备或个人通信设备中，以对转换成数字形式进行传输的模拟音频信号进行数字信号压缩。在可以使用蜂窝电话或个人通信设备的移动的环境中，高的背景噪声能量使得用基于信号能量的速率确定算法难以把低能量的非噪音声音从低背景噪声中区分开来。因此，经常对非噪音声音以较低的比特率进行编码，声音质量下降，诸如“s”、“x”、“ch”、“sh”、“t”等辅音在重新构成的话音中被丢失。

根据仅把背景噪声能量作为依据的速率判定的声码器在设定阈值时没有

考虑信号相对于背景噪声的强度。当背景噪声提高时，根据仅把背景噪声作为依据的声码器必然会一起压缩阈值。如果信号电平仍然维持不变，但设置阈值电平的校正方法是把信号电平与背景噪声电平一起提升，那么，压缩阈值电平不是最佳的解决方法。在速率可变的声码器内需要另一种考虑了信号强度的设置阈值电平的方法。

剩余的决定性的问题是在通过基于背景噪声能量的速率判定声码器来播放音乐时产生的。当人在说话时，他们必须暂停以便呼吸，这可以把阈值重新设置到适当的背景噪声电平上。然而，在通过声码器传输时，在音乐持续的情况下，没有暂停发生，并且阈值将持续提高，一直到开始对音乐以小于全速率的速率进行编码。在这种情况下，速率可变的编码器把音乐与背景噪声混为一谈。

发明内容

本发明是一种新颖的和经改进的在速率可变的声码器内确定编码速率的方法和装置。本发明的第一个目的是提供一种方法，用这种方法可降低把低能量的非语音话音作为背景噪声进行编码的概率。在本发明中，把输入信号滤波成高频分量和低频分量。然后单独地对输入信号的滤波信号进行分析，以检测是否有语音的存在。因为非语音话音有高频分量，所以相对于高频带来说其强度与背景噪声相比的区别比在整个频带上与背景噪声相比的区别来得更大。

本发明的第二个目的是提供一种装置，这种装置在设置阈值时考虑了信号能量以及背景噪声能量。在本发明中，根据输入信号的信噪比(SNR)的估计值来设定声音检测阈值。在一个典型的实施例中，把在存在语音期间的信号能量估计为最大信号能量，把在无声期间的背景噪声能量估计为最小信号能量。

本发明的第三个目的是提供一种通过速率可变的声码器对音乐进行编码的方法。在一个典型的实施例中，速率选择装置检测阈值电平上升的连续帧的数量，并检查帧数的周期。如果输入信号是有周期性的，这表示存在音乐。如果检测到有音乐存在，那么把阈值设置到以全速率对信号进行编码的电平上。

本发明提供一种为速率可变声码器内的输入信号确定编码速率的装置，包含：多个副带能量计算部件，用于接收所述输入信号并根据预定的副带能量计算公式确定多个副带能量值；多个副带速率确定部件，每个副带速率确定部件用于接收所述

多个副带能量值中相应的一个副带能量值,并根据所述多个副带能量值中所述相应的一个副带能量值确定一个建议的编码速率,从而提供多个建议的编码速率;以及编码速率选择部件,用于接收所述多个建议的编码速率,并根据所述多个建议的编码速率选择所述输入信号的编码速率。

本发明提供一种确定速率可变声码器中的输入信号的编码速率的方法,包含下列步骤:接收输入信号;根据预定的副带能量计算公式确定多个副带能量值;对所述多个副带能量值中每一个副带能量值确定一个相应的建议的编码速率,以提供多个建议的编码速率;和根据所述多个建议的编码速率选择所述输入信号的编码速率。通过下面结合附图的详细描述,本发明的特征、目的和优点将变得更明显,在整个描述中相同的参考字符表示相同的部件。

附图说明

图1是本发明的方框图。

具体实施方式

参见图1,把输入信号 $S(n)$ 提供给副带能量计算部件4和副带能量计算部件6。输入信号 $S(n)$ 包含音频信号和背景噪声。音频信号一般为话音,但也可以是音乐。在一个典型的实施例中,以每二十毫秒帧160样值的形式提供 $S(n)$ 。在一个典型的实施例中,输入信号 $S(n)$ 的频率分量从0kHz到4kHz,大约与人的话音信号的带宽相似。

在一个典型的实施例中,把4kHz的输入信号 $S(n)$ 滤波成两个分立的副带。这个分立的副带分别在0到2kHz和2kHz到4kHz之间。在一个典型的实施例中,可以用副带滤波器把输入信号分成副带,这种设计在已有技术中属于熟知的技术,并且在1994年2月1日提交的,名称为“频率选择自适应滤波”的美国专利申请 No. 08/189,819 中有详细的描述,该申请已转让给本发明的受让人,援引在此以作参考。

对于低通滤波器,副带滤波器的脉冲响应表示为 $h_L(n)$,对于高通滤波器,副带滤波器的脉冲响应表示为 $h_H(n)$ 。可以如现有技术中所熟知的那样,简单地取副带滤波器输出的样值平方之和计算得到的信号的所产生的副带分量的能量,给出 $R_L(0)$ 和 $R_H(0)$ 值。

在一个较佳实施例中,当把输入信号 $S(n)$ 提供给副带能量计算部件4时,

如下计算输入帧的低频分量的能量值 $R_L(0)$ ：

$$R_L(0) = R_s(0) \cdot R_{hL}(0) + 2 \cdot \sum_{i=1}^{L-1} R_s(i) \cdot R_{hL}(i) \quad (1)$$

其中， L 为具有脉冲响应 $h_L(n)$ 的低通滤波器内的抽头数， $R_s(i)$ 为输入信号 $S(n)$ 的自相关函数，它由下式给出：

$$R_s(i) = \sum_{n=1}^N S(n) \cdot S(n-i), \text{ 对 } i \in [0, L-1] \quad (2)$$

其中， N 为帧内的样值数， R_{hL} 是低通滤波器 $h_L(n)$ 的自相关函数，由下式给出：

$$R_{hL}(i) = \begin{cases} \sum_{n=0}^{L-1} h_L(n) \cdot h_L(n-i), & \text{对 } i \in [0, L-1] \\ 0 & \text{其它} \end{cases} \quad (3)$$

在副带能量计算部件 6 内用相似的方式计算高频能量 $R_H(0)$ 。

可以在减小计算负荷之前计算副带滤波器的自相关函数的值。另外，把计算得到的一些 $R_s(i)$ 值在对输入信号 $S(n)$ 进行编码时的另一些计算中使用，这进一步减轻了本发明的编码速率选择的方法的纯计算负荷。例如，运算 LPC 滤波器抽头值需要计算一组输入信号自相关系数。

对 LPC 滤波器抽头值的计算在现有技术中是众所周知的，并且在上面提到美国专利申请 08/004,484 中有详细的描述。如果一种是用需要十个抽头的 LPC 滤波器对话音进行编码，除了在对信号进行编码所用的之外，仅需要计算 i 值从 11 到 $L-1$ 的 $R_s(i)$ 值，因为， i 值从 0 到 10 的 $R_s(i)$ 在计算 LPC 滤波器抽头值时已经使用了。在一个典型的实施例中，副带滤波器具有 17 个抽头， $L=17$ 。

副带能量计算部件 4 向副带速率判定部件 12 提供计算得到的 $R_L(0)$ 值，副带能量计算部件 6 向副带速率判定部件 14 提供计算得到的 $R_H(0)$ 值。速率判定部件 12 把 $R_L(0)$ 值与两个预定的阈值 $TL1/2$ 和 $TLfull$ 作比较，把根据比较结果选定建议的编码速率 RATEL。速率的选定方式如下：

$$RATEL = \text{八分之一速率} \quad R_L(0) \leq TL1/2 \quad (4)$$

$$RATEL = \text{半速率} \quad TL1/2 < R_L(0) \leq TLfull \quad (5)$$

$$RATEL = \text{全速率} \quad R_L(0) > TLfull \quad (6)$$

副带速率判定部件 14 以相似的方式工作，并根据高频能量值 $R_H(0)$ 和一组

不同的阈值 $TH_{1/2}$ 和 TH_{full} 来选择一建议的编码速率。副带速率判定部件 12 把其建议的编码速率 $RATE_L$ 提供给编码速率选择部件 16, 副带速率判定部件 14 把其建议的编码速率 $RATE_H$ 提供给编码速率选择部件 16。在一个典型的实施例中, 编码速率选择部件 16 选择两个建议的速率中较高的一个速率, 并把较高的速率作为选出的编码速率(ENCODING RATE)提供。

副带能量计算部件 4 还把低频能量值 $R_L(0)$ 提供给阈值修正部件 8, 计算下一输入帧的阈值 $TL_{1/2}$ 和 TL_{full} 。相似地, 副带能量计算部件 6 把高频能量值 $R_H(0)$ 提供给阈值修正部件 10, 计算下一输入帧的阈值 $TL_{1/2}$ 和 TL_{full} 。

阈值修正部件 8 接收低频能量值 $R_L(0)$, 并确定 $S(n)$ 是否含有背景噪声或音频信号。在一个典型的实现方法中, 阈值修正部件 8 确定是否有音频信号存在的方法是检查归一化自相关函数 NACF, 它由下式给出:

$$NACF = \max_T \frac{\sum_{n=0}^{N-1} e(n) \cdot e(n-T)}{\frac{1}{2} \left[\sum_{n=0}^{N-1} e^2(n) + \sum_{n=0}^{N-1} e^2(n-T) \right]} \quad (7)$$

其中, $e(n)$ 为语音质量的特性分量残留信号, 它由 LPC 滤波器滤波输入信号 $S(n)$ 引起。

由 LPC 滤波器对信号滤波的设计在现有技术中是众所周知的, 并且在上面提及的美国专利申请 08/004, 484 中有详细的描述。LPC 滤波器对输入信号 $S(n)$ 进行滤波, 除去语音质量特性分量的相互影响。把 NACF 与阈值比较, 确定是否出现了音频信号。如果 NACF 大于预定的阈值, 它指示输入帧具有表示诸如语音或音乐的音频信号存在的周期性特征。请注意, 当一部分语音和音乐不是周期性时, 表现出 NACF 的值较小, 背景噪声一般决不会显示出周期性, 因此 NACF 几乎总是表现出较小的值。

如果确定 $S(n)$ 包含背景噪声, NACF 值小于阈值 TH_1 , 那末把值 $R_L(0)$ 用于更新当前背景噪声估计值 BGN_L 的值。在一个典型的实施例中, TH_1 为 0.35。把 $R_L(0)$ 与当前的背景噪声估计值 BGN_L 比较。如果 $R_L(0)$ 小于 BGN_L , 那末不管 NACF 的值如何, 总把背景噪声估计值 BGN_L 设置成等于 $R_L(0)$ 值。

背景噪声估计值只有在 NACF 小于阈值 TH_1 时才增加。如果 $R_L(0)$ 大于 BGN_L , 并且 NACF 小于 TH_1 , 那么把背景噪声能量 BGN_L 设置成 $\alpha_1 \cdot BGN_L$, 其中, α_1 为大

于 1 的数字。在一个典型的实施例中， α_1 等于 1.03。只要 NACF 小于阈值 TH1，并且 $R_L(0)$ 大于 BGN_L 的当前值，那末 BGN_L 就继续增加，直到 BGN_L 到达预定的最大值 BGN_{max} ，在该点上，背景估计值 BGN_L 被设置到 BGN_{max} 。

如果 NACF 值超过第二预定值 TH2 表示检测到音频信号，则更新信号能量估计值 S_L 。在一个典型的实施例中，TH2 被设置成 0.5。把 $R_L(0)$ 的值与当前低通信号能量估计值 S_L 比较。如果 $R_L(0)$ 大于当前 S_L 值，则把 S_L 设置成等于 $R_L(0)$ 。如果 $R_L(0)$ 小于当前 S_L 值，而且仅在 NACF 大于 TH2 时，把 S_L 设置成等于 $\alpha_2 * S_L$ 。在一个典型的实施例中， α_2 被设置为 0.96。

然后，阈值修正部件 8 根据下面的等式 8 计算信噪比估计值：

$$SNR_L = 10 \cdot \log \left[\frac{S_L}{BGN_L} \right] \quad (8)$$

然后，阈值修正部件 8 根据下面的等式 9-12 确定量化了的信噪比 $ISNRL$ 的指数：

$$I_{SNRL} = \text{nint} \left[\frac{SNR_L - 20}{5} \right], \quad \text{对 } 20 < SNR_L < 55 \quad (9)$$

$$\begin{aligned} &= 0, && \text{对 } SNR_L \leq 20, \\ &= 7, && \text{对 } SNR_L \geq 55. \end{aligned} \quad (10)$$

其中 nint 是把小数值四舍五入到最近的整数的函数。

然后阈值修正部件 8 根据信噪比指数 I_{SNRL} 选择或计算两个换算系数 $k_{L1/2}$ 和 k_{Lfull} 。下面的表 1 提供了一个典型的换算值查找表：

表 1

I_{SNRL}	$K_{L1/2}$	K_{Lfull}
0	7.0	9.0
1	7.0	12.6
2	8.0	17.0
3	8.6	18.5
4	8.9	19.4
5	9.4	20.9
6	11.0	25.5
7	15.8	39.8

这两个值用于根据下面式子计算选择速率的阈值：

$$TL1/2 = KL1/2 * BGN_L \quad (11)$$

和

$$TLfull = KLfull * BGN_L \quad (12)$$

其中， $TL1/2$ 为低频半速率阈值， $TLfull$ 为低频全速率阈值。

阈值修正部件 8 向速率判定部件 12 提供修正后的阈值 $TL1/2$ 和 $TLfull$ 。阈值修正部件 10 以相似的方式工作，并向副带速率判定部件 14 提供阈值 $TH1/2$ 和 $THfull$ 。

音频信号能量估计值 S 的初始值 (S 可以是 SL 或 SH) 如下进行设置。把初始信号能量估计值 $SINIT$ 设置到 $-18.0dBm0$ ，其中 $3.17dBm0$ 表示全正弦波的信号强度，在一个典型的实施例中，它是一个幅度范围从 -8031 到 8031 的数字正弦波。 $SINIT$ 一直被使用，直到确定出现了有声信号。

开始检测有声信号的方法是把 $NACF$ 值与一阈值比较，当 $NACF$ 在预定的连续数帧超过该阈值时，则确定出现了有声信号。在一个典型的实施例中， $NACF$ 必须连续 10 帧超过阈值。在这个条件得到满足后，在前 10 帧把信号能量估计值 S 设置到最大信号能量。

最初把背景噪声估计值 BGN_L 的初始值设置成 BGN_{max} 。只要接收到的副带帧能量小于 BGN_{max} ，就把背景噪声估计值复位到接收到的副带能量电平值上，并如上所述产生背景噪声 BGN_L 估计值。

在一个较佳实施例中，当跟了一串全速率话音帧时产生释放延迟情况，则检测低速率帧。在一个典型的实施例中，当在对四个连续的话音帧以全速率进行编码后跟一幅把编码速率设置到小于全速率的速率，并且计算得到的信噪比小于预定最小的 SNR 的帧时，把该帧的编码速率设置到全速率。在一个典型的实施例中，如在公式 8 中定义的那样，预定最小 SNR 为 $27.5dB$ 。

在一较佳实施例中，释放延迟的帧数是信噪比的函数。在一个典型的实施例中，释放延迟的帧数如下确定：

$$\text{释放延迟帧数} = 1 \quad 22.5 < SNR < 27.5 \quad (13)$$

$$\text{释放延迟帧数} = 2 \quad SNR \leq 22.5 \quad (14)$$

$$\text{释放延迟帧数} = 0 \quad SNR \geq 27.5 \quad (15)$$

本发明还提供一种检测是否有音乐存在的方法，如上所述音乐缺少可以测量背景噪声以进行复位的暂停。该检测音乐是否存在的方法假设在通话开始时没有出现音乐。这可以使本发明的编码速率选择装置适当地估计初始背景噪声能量 BGNinit。因为音乐不象背景噪声具有周期性的特征，本发明检查 NACF 的值来区别音乐和背景噪声。本发明的音乐检测方法根据下式计算平均 NACF：

$$NACF_{AVE} = \frac{1}{T} \sum_{i=1}^T NACF(i) \quad (16)$$

其中 NACF 由公式 7 定义，T 为连续帧数，在这些帧中估计的背景噪声值从初始背景噪声估计值 BGNINIT 已经并正在增加。

如果背景噪声 BGN 对预定的帧数 T 已经增加，并且 $NACF_{AVE}$ 超过了预定阈值，那么检测到了音乐，把背景噪声 BGN 复位到 BGNinit。应注意，为了使该方法可行，必须把值 T 设置得足够小，以使编码速率不低于全速率。因此，T 值应当设置成有声信号和 BGNinit 的函数。

提供了上面对较佳实施例的描述能使本技术领域的熟练人员实现或使用本发明。对于本技术领域的熟练人员来说对这些实施例的各种变化是容易的，此处限定的一般原理可以应用于其它实施例而无需创造性技能。因此，本发明并不限于此处所示的实施例，它被赋予与由此处的原理和新颖的特征相一致的最宽的范围。

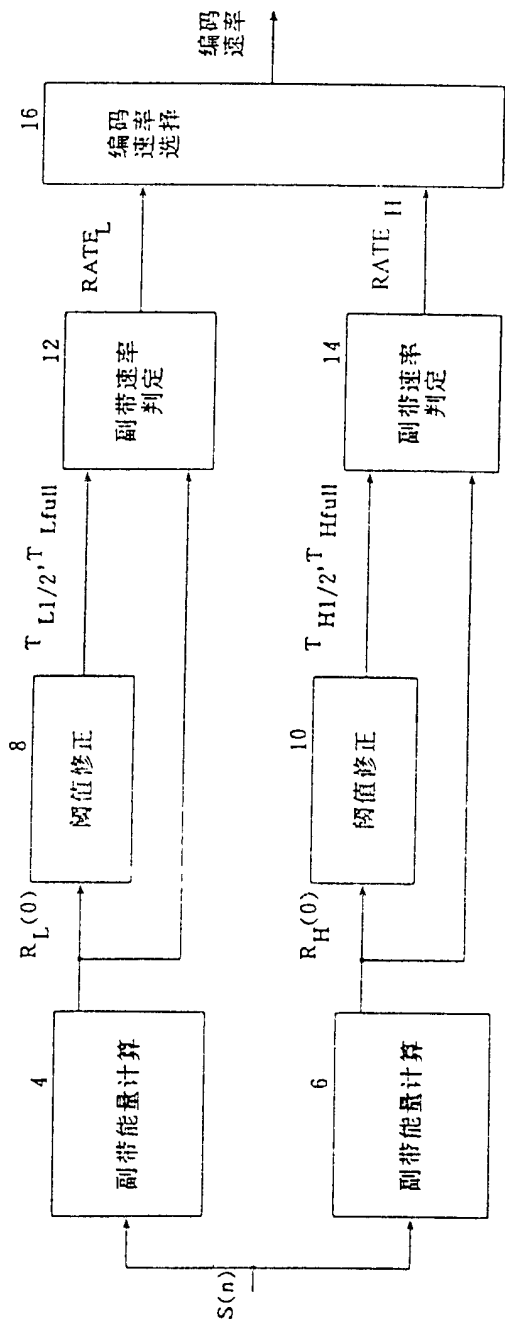


图 1