



[12] 发明专利说明书

[21] ZL 专利号 99803763.X

[45] 授权公告日 2003 年 9 月 17 日

[11] 授权公告号 CN 1121683C

[22] 申请日 1999.2.12 [21] 申请号 99803763.X
 [30] 优先权
 [32] 1998.3.9 [33] FI [31] 980532
 [86] 国际申请 PCT/FI99/00112 1999.2.12
 [87] 国际公布 WO99/46764 英 1999.9.16
 [85] 进入国家阶段日期 2000.9.7
 [71] 专利权人 诺基亚移动电话有限公司
 地址 芬兰埃斯波
 [72] 发明人 P·奥亚拉
 审查员 杨 叁

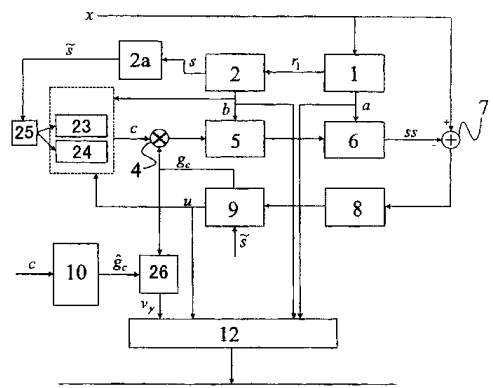
[74] 专利代理机构 中国专利代理(香港)有限公司
 代理人 陈 霁 张志醒

权利要求书 5 页 说明书 12 页 附图 2 页

[54] 发明名称 语音编码

[57] 摘要

可变比特率语音编码方法为每个子帧确定一个量化矢量 $d(i)$ ，该矢量包括数量可变的脉冲。用于激励 LTP 和 LPC 合成滤波器的激励矢量 $c(i)$ 是通过量化矢量 $d(i)$ 进行滤波得到的，增益值 g_c 被确定以用于缩放脉冲幅度激励矢量 $c(i)$ ，使得缩放后的激励矢量表示加权后的残留信号 \tilde{s} ，该信号在通过 LPC 和 LTP 分析除去子帧语音信号中的冗余信息之后仍然保留在该子帧语音信号中。预测的增益值 \hat{g}_c 是根据以前处理的子帧确定的，并且当依据量化矢量 $d(i)$ 中的脉冲数 m 对矢量残留信号 $c(i)$ 的幅度进行缩放时， \hat{g}_c 为包含于激励矢量 $c(i)$ 中的能量 E_c 的函数。量化后的增益校正因子 γ_{g_c} 可以利用增益值 g_c 和预测增益值 \hat{g}_c 确定。



ISSN 1008-4274

1. 一种编码语音信号的方法，其中的信号包括含有数字化语音样本的子帧序列，对于每个子帧，该方法包括：

5 (a) 选出一个包括至少一个脉冲的量化矢量 $d(i)$ ，其中矢量 $d(i)$ 中的脉冲数 m 和脉冲位置可能在子帧之间变化；

(b) 确定增益值 g_c 用于缩放量化矢量 $d(i)$ 的幅度或用于缩放从量化矢量 $d(i)$ 得到的另一个矢量 $c(i)$ 的幅度，其中缩放后的矢量与加权后的残留信号 \tilde{s} 同步；

10 (c) 确定缩放因子 k ，该因子为预定能量值与量化矢量 $d(i)$ 中能量之比的函数；

(d) 在一个或多个以前处理过的子帧基础上确定预测的增益值 \hat{g}_c ，该增益值为量化矢量 $d(i)$ 的能量 E_c 的函数或当另一个矢量 $c(i)$ 的幅度由所述的缩放因子 k 缩放时该矢量 $c(i)$ 的能量 E_c 的函数；

15 (e) 利用所述的增益值 g_c 和所述的预测增益值 \hat{g}_c 确定量化的增益校正因子 $\hat{\rho}_{g_c}$ 。

2. 根据权利要求 1 的方法，该方法是变比特率编码方法，该方法包括：

20 通过从语音信号子帧中基本上除去长时和短时冗余而产生所述的加权残留信号 \tilde{s} ；

根据包含于加权残留信号 \tilde{s} 中的能量将语音信号子帧分类，并利用该分类确定量化矢量 $d(i)$ 中的脉冲数 m 。

3. 根据权利要求 1 或 2 的方法，包括：

25 为每个帧产生一组线性预测编码 LPC 系数 a 和为每个子帧产生一组长时预测 LTP 参数 b ，其中一帧包括多个语音子帧；

在 LPC 系数，LTP 参数，量化矢量 $d(i)$ 以及量化增益校正因子 $\hat{\rho}_{g_c}$ 的基础上产生编码语音信号。

4. 根据权利要求 1 的方法，包括通过代数码 u 在编码信号中定义量化矢量 $d(i)$ 。

30 5. 根据权利要求 1 的方法，其中预测增益值是根据下面等式确定的：

$$\hat{g}_c = 10^{0.05(\hat{E}(n) + \bar{E} - E_c)}$$

其中 \bar{E} 是常数, $\hat{E}(n)$ 是在所述以前处理过的子帧基础上确定的当前子帧中能量的预测值。

- 5 6. 根据权利要求 1 的方法, 其中所述预测增益值 \hat{g}_c 是量化矢量 $d(i)$ 的除去均值后的能量 $E(n)$ 的函数, 或者当每个以前处理的子帧的所述另一个矢量 $c(i)$ 的幅度被所述缩放因子 k 缩放时, 为该矢量 $c(i)$ 的能量 E_c 的函数。

- 10 7. 根据权利要求 1 的方法, 其中增益值 g_c 被用于对所述另一个矢量 $c(i)$ 进行缩放, 该另一个矢量是通过对量化矢量 $d(i)$ 滤波得到的。

8. 根据权利要求 5 的方法, 其中:

- 15 所述预测增益值 \hat{g}_c 是量化矢量 $d(i)$ 的除去均值后的激励能量 $E(n)$ 的函数, 或者当每个以前处理过的子帧的所述另一个矢量 $c(i)$ 的幅度被所述缩放因子 k 缩放时, 为该矢量 $c(i)$ 的能量 E_c 的函数;

增益值 g_c 被用于对所述另一个矢量 $c(i)$ 进行缩放, 该另一个矢量是通过对量化矢量 $d(i)$ 滤波得到的;

预测能量是利用下面等式得到的:

$$\hat{E}(n) = \sum_{i=1}^P b_i \hat{R}(n-i)$$

20

其中 b_i 是移动平均预测系数, P 是预测阶数, $\hat{R}(j)$ 是以前子帧 j 中预测能量 $\hat{E}(j)$ 中的误差, 由下式给出:

$$\hat{R}(n) = E(n) - \hat{E}(n)$$

25 其中

$$E(n) = 10 \log \left(\frac{1}{N} g_c^2 \sum_{i=0}^{N-1} (kc(i))^2 \right) - \bar{E}.$$

9. 根据权利要求 5 的方法，其中项 E_c 由下面等式确定：

$$E_c = 10 \log \left(\frac{1}{N} \sum_{i=0}^{N-1} (kc(i))^2 \right)$$

5

其中 N 是子帧中的样本数。

10. 根据权利要求 1 的方法，其中，如果量化矢量 $d(i)$ 包括两个或多个脉冲，则所有脉冲具有相同的幅度。

11. 根据权利要求 1 的方法，其中缩放因子由下式给出：

10

$$k = \sqrt{\frac{M}{m}}$$

其中 M 是量化矢量 $d(i)$ 中最大允许的脉冲数。

12. 根据权利要求 1 的方法，该方法包括搜索一个增益校正因子码书来确定量化增益校正因子 $\hat{\gamma}_{gc}$ ，该因子使误差最小化：

15

$$e_Q = (g_c - \hat{\gamma}_{gc} \hat{g}_c)^2$$

并对所识别出的量化增益校正因子进行码书索引编码。

13. 对数字化采样语音信号子帧序列解码的方法，对于每个子帧，该方法包括：

20

(a) 从编码信号恢复包括至少一个脉冲的量化矢量 $d(i)$ ，其中矢量 $d(i)$ 中的脉冲数 m 和脉冲位置可能在子帧之间变化；

(b) 从编码信号恢复量化增益校正因子 $\hat{\gamma}_{gc}$ ；

(c) 确定缩放因子 k ，该因子为预定能量值与量化矢量 $d(i)$

中能量之比的函数;

(d) 在一个或多个以前处理过的子帧基础上确定预测的增益值 \hat{g}_c , 该增益值是量化矢量 $d(i)$ 的能量 E_c 的函数或当得自该量化矢量的另一个矢量 $c(i)$ 的幅度由所述的缩放因子 k 缩放时, 该矢量 $c(i)$ 的能量 E_c 的函数;

(e) 利用 量化增益校正因子 $\hat{\gamma}_{gc}$ 来校正预测增益值 \hat{g}_c 以给出校正后的增益值 g_c ;

(f) 利用 增益值 g_c 对量化矢量 $d(i)$ 或所述另一个矢量 $c(i)$ 进行缩放以产生与残留信号 \tilde{s} 同步的激励矢量, 其中的残留信号 \tilde{s} 在从原始子帧语音信号中除去冗余信息之后仍然保留在该子帧中。

14. 根据权利要求 13 的方法, 其中每个所接收信号的编码子帧包括一个定义量化矢量 $d(i)$ 的代数码 μ 和对获得量化增益校正因子 $\hat{\gamma}_{gc}$ 的量化增益校正因子码书编址的索引。

15. 用于编码语音信号的装置, 其中的信号包括含有数字化语音样本的子帧序列, 该装置具有依次对所述每个子帧编码的装置, 这些装置包括:

用于选出包括至少一个脉冲的量化矢量 $d(i)$ 的矢量选择装置, 其中矢量 $d(i)$ 中脉冲数 m 和脉冲位置可能在子帧之间变化;

用于确定增益值 g_c 的第一信号处理装置, 该增益值用于缩放量化矢量 $d(i)$ 的幅度或从量化矢量 $d(i)$ 得到的另一个矢量 $c(i)$ 的幅度, 其中缩放后的矢量与加权后的残留信号 \tilde{s} 同步;

用于确定缩放因子 k 的第二信号处理装置, 其中 k 为预定能量值与量化矢量 $d(i)$ 中能量之比的函数;

在一个或多个以前处理过的子帧基础上确定预测增益值 \hat{g}_c 的第三信号处理装置, 该增益值为量化矢量 $d(i)$ 的能量 E_c 的函数或当另一个矢量 $c(i)$ 的幅度由所述的缩放因子 k 缩放时, 为该矢量 $c(i)$ 的能量 E_c 的函数;

用于利用所述的增益值 g_c 和所述的预测增益值 \hat{g}_c 确定量化增益校正因子 $\hat{\gamma}_{gc}$ 的第四信号处理装置。

16. 用于对数字化采样语音信号的编码子帧序列进行解码的装置, 该装置具有用于对每个所述子帧依次解码的装置, 所说依次解码装置包括:

从编码信号恢复包括至少一个脉冲的量化矢量 $d(i)$ 的第一信号处理装置，其中矢量 $d(i)$ 中脉冲数 m 和脉冲位置可能在子帧之间变化；

从编码信号恢复量化增益校正因子 $\hat{\gamma}_{gc}$ 的第二信号处理装置；

- 5 确定缩放因子 k 的第三信号处理装置，该因子为预定能量值与量化矢量 $d(i)$ 中能量之比的函数；

- 在一个或多个以前处理过的子帧基础上确定预测增益值 \hat{g}_c 的第四信号处理装置，该增益值为量化矢量 $d(i)$ 的能量 E_c 的函数或当另一个得自该量化矢量的矢量 $c(i)$ 的幅度由所述的缩放因子 k 缩放时，
10 为矢量 $c(i)$ 的能量 E_c 的函数；

利用 量化增益校正因子 $\hat{\gamma}_{gc}$ 来校正预测增益值 \hat{g}_c 以给出校正后的增益值 g_c 的校正装置；

- 利用 增益值 g_c 对量化矢量 $d(i)$ 或所述另一个矢量 $c(i)$ 进行缩放以产生与残留信号 \tilde{s} 同步的激励矢量的缩放装置，其中的残留信号 \tilde{s} 在从原始子帧语音信号中除去冗余信息之后仍然保留在该子帧中。
15

语音编码

技术领域

- 5 本发明涉及语音编码，更具体地，涉及在包含数字化语音样本的离散时间帧中对语音信号编码，但是本发明特别适用于，尽管是不必要的，变长比特语音编码。

背景技术

- 10 在欧洲，被接受的数字蜂窝电话的标准以字首 GSM 而闻名（用于移动通讯的全球系统），最近版本的 GSM 标准（GSM2；06.60）导致已知为增强全速率（EFR）的新语音编码算法（或编解码器）的细则。如传统的语音编解码器那样，EFR 被设计为降低个体声音或数据通讯所需的比特率。通过最小化该比特率，可以复用到给定信号带宽的独立呼叫数可以增加。

- 15 类似于 EFR 中使用的语音编码器结构的通常的图解说明在图 1 中给出。采样后的语音信号被分成 20 毫秒的帧 x ，每个包含 160 个样本。每个样本由 16 个比特表示。通过首先将样本帧应用于线性预测编码器（LPC1），这些样本帧被编码，其中的线性预测编码器为每个帧产生一组 LPC 系数 a 。这些系数代表帧中的短时冗余。

- 20 来自 LPC1 的输出包括 LPC 系数 a 和残余信号 γ_1 ，该信号是通过 LPC 分析滤波器从输入语音帧中除去短时冗余而产生的。然后，残留信号被提供给长时预测器（LPT）2，它产生一组表示残留信号 γ_1 中长时冗余的 LTP 参数 b ，并且还产生长时冗余被除去的残留信号 s 。实际上，长时预测分两个阶段，（1）首先为整个帧进行开环估计得到一组 LTP 参数；（2）其次对估计所得的参数进行闭环精确化以便为该帧的每个 40 样本子帧产生一组 LTP 参数。LTP2 提供的残留信号 s 依次通过滤波器 $1/A(z)$ 和 $W(z)$ 而被滤波（以图 1 中的方框 2a 给出）以给出加权后的残留信号 \tilde{s} 。这些滤波器中的第一个是 LPC 合成滤波器，而第二个是强调谱中的共振峰结构的感知加权滤波器。所有滤波器的参数是由 LPC 分析阶段给出的（块 1）。
- 25
- 30

代数激励码书 3 被用于产生激励矢量 c 。对于每个 40 样本子帧（每帧有 4 个子帧），通过缩放单元 4，一些不同的“候选”激励矢量依次

被施加给 LTP 合成滤波器 5。滤波器 5 接受当前子帧的 LTP 参数，并且在激励矢量中引入 LTP 参数预测的长时冗余。所产生的信号然后被提供给 LPC 合成滤波器 6，该滤波器接收连续帧的 LPC 系数。对于给定的子帧，利用帧到帧的内插会产生一组 LPC 系数，所产生的系数依次应用于产生合成信号 ss 。

图 1 的编码器不同于以前的码激励线性预测 (CELP) 编码器，后者用到了包含预定的激励矢量组的码书。前者类型的编码器却依赖于激励矢量的代数产生和确定 (例如，见 W09624925)，并且常常被称为代数 CELP 或 ACELP。更具体的，量化矢量 $d(i)$ 被定义为包含 10 个非零脉冲。所有的脉冲幅度可以为 +1 或 -1。子帧中 40 个样本位置 ($i=0$ 到 39) 被划分成 5 个“轨道”，每个轨道包括两个脉冲 (即 8 个可能位置中的 2 个)。如在下面表中给出的。

表 1: 代数码书中各个脉冲的可能位置

轨道	脉冲	位置
1	i_0, i_5	0, 5, 10, 15, 20, 25, 30, 35
2	i_1, i_6	1, 6, 11, 16, 21, 26, 31, 36
3	i_2, i_7	2, 7, 12, 17, 22, 27, 32, 37
4	i_3, i_8	3, 8, 13, 18, 23, 28, 33, 38
5	i_4, i_9	4, 9, 14, 19, 24, 29, 34, 39

15

给定轨道中每对脉冲的位置以 6 比特编码 (即，总共 30 比特，每个脉冲 3 比特)，而轨道中第一脉冲的符号以 1 比特编码 (总共 5 比特)。第二脉冲的符号并不会被特别编码，而是根据其相对于第一脉冲的位置来获得，如果第二脉冲的采样位置先于第一脉冲，那么第二脉冲被定义为与第一脉冲符号相反，否则，两个脉冲被定义具有相同的符号。所有的 3 比特脉冲位置被进行格雷编码，以便提高针对信道误差的强度，使得量化矢量可以用 35 比特代数码 u 来编码。

为了产生激励矢量 $c(i)$ ，由代数码 u 定义的量化矢量 $d(i)$ 被预滤波器 $F_E(z)$ 滤波，其中的预滤波器增强了特殊的谱分量以便提高合成语音的质量。预滤波器 (常常称为色彩滤波器) 用于该子帧产生

25

的某些 LTP 参数来定义。

如传统的 CELP 编码器那样，差值单元 7 在逐个样本（逐个子帧）的基础上确定合成信号和输入信号之间的差值。加权滤波器 8 被用于对误差信号加权以考虑人类音频感知。对于给定的子帧，搜索单元 9 从代数码书 3 产生的候选矢量中选出适当的激励矢量 $\{c(i), \text{其中 } i=0 \text{ 到 } 39\}$ ，其方式是识别出最小化加权均方误差的矢量。该过程通常称为“矢量量化”。

如已经注意到的，在缩放单元 4 激励矢量被乘以增益 g_c 。导致缩放后的激励矢量的能量等于加权残留信号 \tilde{s} 能量的增益值被选出，其中的残留信号由 LTP2 给出。该增益由下式给出：

$$g_c = \frac{\tilde{s}^T H c(i)}{c(i)^T H^T H c(i)} \quad (1)$$

其中 H 是线性预测模型（LTP 和 LPC）脉冲响应矩阵。

有必要将增益信息连同定义激励矢量的代数码一起引入编码后的语音子帧，以使得子帧能够被正确重构。然而，与其直接引入增益 g_c ，不如在处理单元 10 中根据以前的语音子帧产生预测增益 \hat{g}_c ，并且在单元 11 中确定校正因子，即：

$$\gamma_{g_c} = g_c / \hat{g}_c \quad (2)$$

然后，在包括 5 比特码矢量的校正因子码书情况下，相关因子被进行矢量量化。索引矢量 v_γ 表明量化后的增益相关因子 $\hat{\gamma}_{g_c}$ ，该因子被引入编码后的帧。假定增益 g_c 在帧与帧之间略有不同，那么 $\gamma_{g_c} \cong 1$ ，并且可以用相对较短的码书来正确量化。

实际上，预测增益 \hat{g}_c 是利用具有固定系数的移动平均（MA）预测得到的，如下所示，对激励能量进行了 4 阶 MA 预测。使得子帧 n 中除去平均激励能量（以 dB）后得到 $E(n)$ ，由下式给出：

$$E(n) = 10 \log \left(\frac{1}{N} g_c^2 \sum_{i=0}^{N-1} c^2(i) \right) - \bar{E} \quad (3)$$

其中 $N = 40$ 是子帧的大小, $c(i)$ 是激励矢量 (包括预滤波)。 $\bar{E} = 36\text{dB}$ 是典型激励能量的预定均值。子帧 n 的能量可以由下式预测:

$$\hat{E}(n) = \sum_{i=1}^4 b_i \hat{R}(n-i) \quad (4)$$

5

其中 $[b_1 b_2 b_3 b_4] = [0.68 \ 0.58 \ 0.34 \ 0.19]$ 是 MA 预测系数, $\hat{R}(j)$ 是子帧 j 的预测能量 $\hat{E}(j)$ 中的误差。根据下面等式, 当前子帧的误差被计算出来, 用在处理后续子帧中:

$$\hat{R}(n) = E(n) - \hat{E}(n) \quad (5)$$

10

通过以 $\hat{E}(n)$ 代替等式 (3) 中的 $E(n)$, 预测能量可以用于计算预测增益 \hat{g}_c , 如下式:

$$\hat{g}_c = 10^{0.05(\hat{E}(n) + \bar{E} - E_c)} \quad (6)$$

15 其中

$$E_c = 10 \log \left(\frac{1}{N} \sum_{i=0}^{N-1} c^2(i) \right) \quad (7)$$

是激励矢量 $c(i)$ 的能量。

增益校正因子码书搜索被执行以识别量化后的增益校正因子 $\hat{\gamma}_{gc}$, 它使得误差最小化:

$$e_Q = (g_c - \hat{\gamma}_{gc} \hat{g}_c)^2 \quad (8)$$

编码帧包括 LPC 系数, LTP 参数, 定义激励矢量的代数码, 以及量

化后的增益校正因子码书索引。在发送之前，在编码和复用单元 12 中会对某些编码参数进行进一步编码。实际上，LPC 系数被转换成相应数量的线性谱对 (LSP) 系数，如在“Efficient Vector Quantisation of LPC Parameters at 24Bits/Frame”Kuldip K. P 和 Bishnu S. A, IEEE Trans Speech and Audio Processing, 卷 1, 第 1 期, January 1993 中描述的那样，整个的编码帧也被编码以用于误差检测和校正。为 GSM2 制定的编解码器以完全相同的比特数，即 244 对每个语音帧进行编码。在引入卷积编码和附加了循环冗余检验比特后增加到 456 比特。

图 2 给出 ACELP 解码器的通常结构，适于对被图 1 的编码器编码的信号解码。解复用器 13 将所接收的编码信号分离为各个分量。相同于编码器处的码书 3 的代数码书 14 确定编码矢量并对该矢量进行预滤波（利用 LTP 参数）以产生激励矢量，其中的编码矢量由所接收的编码信号中的 35 比特代数码确定。增益校正因子是利用所接收的量化增益校正因子并根据增益校正因子码书确定的，并且该因子在块 15 中用于校正在块 16 确定的、根据以前解码的子帧得到的预测增益。在块 17 中，激励矢量被乘以校正后的增益，然后该乘积被传送给 LTP 合成滤波器 18 和 LPC 合成滤波器 19。LTP 和 LPC 滤波器分别接收由编码信号传送的 LTP 参数和 LPC 系数，并在激励矢量中再次引入长时和短时冗余。

语音在其本质上变化性很强，包括强活动期和弱活动期，并且常常包括相对的无声段。因此使用固定比特率编码会浪费带宽资源。一些语音编解码器被推荐，这些编解码器的帧与帧之间，子帧与子帧之间的编码比特率是变化的。例如，US5, 657, 420 推荐了一种语音编解码器用于 US CDMA 系统中，在该系统中，数据帧的编码比特率是根据数据帧中的语音活动等级而从一些可能的比特率中选出的。

至于 ACELP 编解码器，建议将语音信号子帧划分成两类或多类，并用不同的代数码书对不同的类别进行编码。更具体的，加权信号 \tilde{s} 随时间变化很慢的子帧可以利用具有相对较少脉冲（如 2）的码矢量 $d(i)$ 来编码，而加权残留信号变化相对较快的子帧可以用具有相对较多脉冲（例如 10）的码矢量 $d(i)$ 来编码。

参考上面的等式 (7)，码矢量 $d(i)$ 中激励脉冲数量的变化，例如从 10 变为 2 将导致激励矢量 $c(i)$ 中能量的相应降低。因为等式 (4)

的能量预测是基于以前子帧的，在激励脉冲数量大量减少的情况下，该预测值可能会很差。这样会导致预测增益 \hat{g}_c 中相对较大的误差，造成增益校正因子在整个语音信号上变化很大。为了能够正确地对这种变化范围很大的增益校正因子量化，增益校正因子量化表必须相对很大，需要相应较长的码书索引 v_γ ，例如 5 比特。这样会在编码子帧数据中加入额外的比特。

要理解的是预测增益中的较大误差也会产生于 CELP 编码器中，在该编码器中码矢量 $d(i)$ 的能量在帧与帧之间变化很大，需要类似的较大的码书用于量化增益校正因子。

10 发明内容

本发明的目的是克服或至少减轻上面提到的现存可变速率编解码器的不利之处。

根据本发明的第一方面，这里给出了一种对语音信号编码的方法，其中的信号包括含有数字化语音样本的子帧序列，对于每个子帧，该方法包括：

15 (a) 选出一个至少包括一个脉冲的量化矢量 $d(i)$ ，其中矢量 $d(i)$ 中的脉冲数 m 和脉冲位置可能在子帧之间变化。

(b) 确定增益值 g_c 用于缩放量化矢量 $d(i)$ 的幅度或用于缩放从量化矢量 $d(i)$ 得到的另一个矢量 $c(i)$ 的幅度，其中缩放后的矢量与加权后的残留信号 \tilde{s} 同步。

20 (c) 确定缩放因子 k ，该因子为预定能量值与量化矢量 $d(i)$ 中能量之比的函数；

(d) 在一个或多个以前处理过的子帧基础上确定预测的增益值 \hat{g}_c ，该因子为量化矢量 $d(i)$ 的能量 E_c 的函数或者当另一个矢量 $c(i)$ 的幅度由所述的缩放因子 k 缩放时，为该矢量 $c(i)$ 的能量 E_c 的函数。

(e) 利用所述的增益值 g_c 和所述的预测增益值 \hat{g}_c 确定量化的增益校正因子 $\hat{\gamma}_{g_c}$ 。

30 通过如上述的那样缩放激励矢量的能量，当量化矢量 $d(i)$ 中的脉冲数（或能量）在子帧之间变化时，本发明会提高预测增益值 \hat{g}_c 的准确性。这样会减小增益校正因子 γ_{g_c} 的范围，并且在与前文相比更小的量化码书的情况下，能够进行正确量化。使用较小的码书降低了用

于索引该码书的矢量的比特长度。此外，可以用与以前所用码书大小相同的码书来提高量化准确性。

在本发明的一个实施方案中，矢量 $d(i)$ 中的脉冲数 m 取决于子帧语音信号的本质。在另一个可选实施方案中，脉冲数 m 是由系统需求或特性确定的。例如在编码信号通过传输信道传输的情况下，当信道干涉较高时，脉冲数可以很小，这样可以允许更多的保护比特加入信号中。当信道干涉较低时，信号需要较少的保护比特，矢量中的脉冲数可以增加。

最好的是，本发明的方法是一种可变比特率的编码方法，该方法包括通过从语音信号子帧中基本除去长时和短时冗余而产生所述加权残留信号 \tilde{s} ，根据包括在加权残留信号 \tilde{s} 中的能量而将语音信号子帧分类，并利用该分类来确定量化矢量 $d(i)$ 中的脉冲数 m 。

最好的是，该方法包括为每个帧产生一组线性预测编码 (LPC) 系数 a ，并为每个子帧产生一组长时预测 (LTP) 参数 b ，其中的数据帧包括多个语音子帧，并在 LPC 系数，LTP 参数，量化矢量 $d(i)$ 和量化增益校正因子 \hat{g}_c 的基础上产生编码的语音信号。

最好的是，量化矢量 $d(i)$ 由代数码 μ 定义，该码被引入编码语音信号中。

最好的是增益值 g_c 被用于缩放所述矢量 $c(i)$ ，该矢量是通过对量化矢量 $d(i)$ 滤波得到的。

最好的是，预测增益值根据下面等式确定。

$$\hat{g}_c = 10^{0.05(\hat{E}(n) + \bar{E} - E_c)}$$

其中 \bar{E} 是常数， $\hat{E}(n)$ 是在以前子帧基础上确定的当前子帧中能量的预测值。该预测能量可以用下面等式确定：

$$\hat{E}(n) = \sum_{i=1}^p b_i \hat{R}(n-i)$$

其中 b_i 是移动平均预测系数， p 是预测阶数， $\hat{R}(j)$ 是以前子帧 j 的预测能量 $\hat{E}(j)$ 的误差，误差由下式给出：

$$\hat{R}(n) = E(n) - \hat{E}(n)$$

项 E_c 是由下面等式确定的:

$$E_c = 10 \log \left(\frac{1}{N} \sum_{i=0}^{N-1} (kc(i))^2 \right)$$

5

其中 N 是子帧中的样本数, 最好的是:

$$k = \sqrt{\frac{M}{m}}$$

其中 M 是量化矢量 $d(i)$ 中最大允许的脉冲数。

10 最好的是, 量化矢量 $d(i)$ 包括两个或多个脉冲, 其中所有的脉冲具有相同的幅度。

最好的是, 步骤 (d) 包括搜索一个增益校正因子码书来确定最小化误差的量化增益校正因子 $\hat{\gamma}_{gc}$ 。

$$e_Q = (g_c - \hat{\gamma}_{gc} \hat{g}_c)^2$$

15

并对识别出的量化增益校正因子进行码书索引编码。

根据本发明的第二方面, 这里给出一种方法, 对数字化采样语音信号的编码子帧序列进行解码, 对于每个子帧, 该方法包括:

20 (a) 从编码信号恢复至少包括一个脉冲的量化矢量 $d(i)$, 其中矢量 $d(i)$ 中的脉冲数 m 和脉冲位置可能在子帧之间变化。

(b) 从编码信号恢复量化增益校正因子 $\hat{\gamma}_{gc}$ 。

(c) 确定缩放因子 k , 该因子为预定能量值与量化矢量 $d(i)$ 中能量之比的函数;

(d) 在一个或多个以前处理过的子帧基础上确定预测的增益值

\hat{g}_c ，该增益值为量化矢量 $d(i)$ 的能量 E_c 的函数，或者当另一个得自 $d(i)$ 的矢量 $c(i)$ 的幅度由所述缩放因子 k 缩放时，为该矢量 $c(i)$ 的能量 E_c 的函数。

5 (e) 利用 量化增益校正因子 $\hat{\gamma}_{gc}$ 来校正预测增益值 \hat{g}_c 以给出校正后的增益值 g_c 。

(f) 利用 增益值 g_c 对量化矢量 $d(i)$ 或所述另一个矢量 $c(i)$ 进行缩放以产生与残留信号 \tilde{s} 同步的激励矢量，其中的残留信号 \tilde{s} 在从原始子帧语音信号中基本上除去冗余信息之后仍然保留在该子帧中。

10 最好的是，每个所接收信号的编码子帧包括一个代数码 u ，该码定义了量化矢量 $d(i)$ ，每个编码子帧还包括一个索引，该索引定义了获得量化增益校正因子 $\hat{\gamma}_{gc}$ 的量化增益校正因子码书的地址。

根据本发明的第三方面，这里给出一种装置用于编码语音信号，该信号包括含有数字语音样本的子帧序列，该装置具有依次编码所述
15 每个子帧的装置，这些该装置包括：

用于选出包括至少一个脉冲的量化矢量 $d(i)$ 的矢量选择装置，其中矢量 $d(i)$ 中的脉冲数 m 和脉冲位置可能在子帧之间变化。

20 用于确定增益值 g_c 的第一信号处理装置，该增益值用于缩放量化矢量 $d(i)$ 的幅度或用于缩放得自量化矢量 $d(i)$ 的另一个矢量 $c(i)$ 的幅度，其中缩放后的矢量与加权后的残留信号 \tilde{s} 同步。

用于确定缩放因子 k 的第二信号处理装置，其中 k 为预定能量值与量化矢量 $d(i)$ 中能量之比的函数；

25 在一个或多个以前处理过的子帧基础上确定预测增益值 \hat{g}_c 的第三信号处理装置，该增益值为量化矢量 $d(i)$ 的能量 E_c 的函数或当另一个矢量 $c(i)$ 的幅度由所述缩放因子 k 缩放时，为该矢量 $c(i)$ 的能量 E_c 的函数。

用于利用所述的增益值 g_c 和所述的预测增益值 \hat{g}_c 确定量化增益校正因子 $\hat{\gamma}_{gc}$ 的第四信号处理装置。

30 根据本发明的第四方面，这里给出一种装置，用于对数字化采样语音信号的编码子帧序列解码，该装置具有对所述每个子帧依次解码的装置，这些装置包括：

用于从编码信号恢复包括至少一个脉冲的量化矢量 $d(i)$ 的第一

信号处理装置，其中矢量 $d(i)$ 中的脉冲数 m 和脉冲位置可能在子帧之间变化。

用于从编码信号恢复量化增益校正因子 $\hat{\gamma}_{gc}$ 的第二信号处理装置。

5 用于确定缩放因子 k 的第三信号处理装置，该因子为预定能量值与量化矢量 $d(i)$ 中能量之比的函数；

用于在一个或多个以前处理过的子帧基础上确定预测增益值 \hat{g}_c 的第四信号处理装置，该因子为量化矢量 $d(i)$ 的能量 E_c 的函数或者当另一个矢量 $c(i)$ 的幅度由所述的缩放因子 k 缩放时，为该矢量 $c(i)$ 的能量 E_c 的函数。

10 用于利用 量化增益校正因子 $\hat{\gamma}_{gc}$ 来校正预测增益值 \hat{g}_c 以给出校正后的增益值 g_c 的校正装置。

用于利用增益值 g_c 对量化矢量 $d(i)$ 或所述另一个矢量 $c(i)$ 进行缩放以产生与残留信号 \tilde{s} 同步的激励矢量的缩放装置，其中的残留信号在从原始子帧语音信号中除去冗余信息之后仍然保留在该子帧中。

附图说明

为了更好地理解本发明以及本发明是如何实现的，下面通过例子参考附图描述，其中：

图 1 给出 ACELP 语音编码器的方框图。

20 图 2 给出 ACELP 语音解码器的方框图。

图 3 给出修正后的能够进行可变比特率编码的 ACELP 语音编码器的方框图。

图 4 给出修正后的能够进行可变比特率解码的 ACELP 语音解码器的方框图。

25 具体实施方式

上面已经参考图 1 和 2 简要描述了类似于为 GSM2 推荐的 ACELP 语音编解码器。图 3 说明了适于对数字化采样语音信号进行变比特率编码的经修正的 ACELP 语音编码器，其中的功能块已经参考图 1 描述了，这些功能块被标以类似的参考标号。

30 在图 3 的编码器中，图 1 的单个代数码书 3 被一对代数码书 23, 24 代替。第一码书 23 被用于基于包含两个脉冲的码矢量 $d(i)$ 来产生激励矢量 $c(i)$ ，而第二码书 24 被用于基于包含 10 个脉冲的码书矢量 d

(i) 来产生激励矢量 $c(i)$ 。对于给定的子帧，码书选择单元 25 根据 LTP2 给出的加权残留信号 \tilde{s} 中的能量选出码书 23, 24。如果加权残留信号中的能量超过了某个预定的 (或自适应的) 阈值 -- 表明变化很大的加权残留信号, 那么 10 个脉冲码书 24 被选出。另一方面, 如果

5 加权残留信号中的能量低于定义的阈值, 那么 2 脉冲码书 23 被选出。在使用 3 个或多个码书的情况下, 建议定义两个或多个阈值。为了更详细地描述适当的码书选择过程, 应该参考文献 "Toll Quality Variable-Rate Speech Codec"; Ojala P; Proc. Of IEEE International Conference on Acoustics, Speech and Signal Processing, Munich,

10 Germany, Apr. 21-24 1997.

用于缩放单元 4 的增益 g_c 的推导是如上面参考等式 (1) 描述的那样实现的。然而, 在获得预测增益 \hat{g}_c 的过程中, 通过如下所示对激励矢量施加一个幅度缩放因子 k , 等式 (7) 被修正 (在修正处理单元 26 中) 为下式:

15

$$E_c = 10 \log \left(\frac{1}{N} \sum_{i=0}^{N-1} (kc(i))^2 \right) \quad (9)$$

在选择 10 个脉冲码书的情况下, $k=1$, 在选择 2 个脉冲码书的情况下, $k=\sqrt{5}$ 。更通用的表达是, 缩放因子由下式给出:

$$k = \sqrt{\frac{10}{m}} \quad (10)$$

20

其中 m 是对应码书矢量 $d(i)$ 中的脉冲数。

对于给定子帧在计算除去均值后的激励能量 $E(n)$ 的过程中, 为了能够以等式 (4) 预测能量, 还需要引入缩放因子 k 。这样等式 (3) 被修正为:

25

$$E(n) = 10 \log \left(\frac{1}{N} g_c^2 \sum_{i=0}^{N-1} (kc(i))^2 \right) - \bar{E} \quad (11)$$

然后通过等式 (6)、等式 (9) 给出的修正后的激励矢量能量和等式 (11) 给出的修正后的除去均值的激励能量来计算预测增益。

将缩放因子 k 引入等式 (9) 和 (11) 明显改善了增益预测使得一般来说 $\hat{g}_c \cong g_c$, $\gamma_{gc} \cong 1$ 。当与以前技术相比增益校正因子的范围缩小时, 可以使用较小的增益校正因子码书, 使用较短长度的码书索引 v_y , 例如 3 或 4 比特。

图 4 说明了适于对图 3 的 ACELP 编码器编码的语音信号解码的解码器, 其中在图 3 中语音子帧以变比特率被编码。图 4 中解码器的大部分功能与图 3 的解码器相同, 并且这些功能块已经参考图 2 描述了, 并且这些功能块在图 2 和图 4 中被标以相同的参考标号。主要的差别在于两个代数码书 20, 21 的给出, 它们对应于图 3 编码器中的 2 脉冲码书和 10 脉冲码书。所接收代数码 u 的本质确定了适当码书 20, 21 的选择, 此后解码过程以前面描述的同样方式进行。然而, 如编码器那样, 在块 22 中利用等式 (6)、等式 (9) 给出的缩放后的激励矢量能量 E_c 和等式 (11) 给出的缩放后的除去均值的激励能量 $E(n)$ 来计算预测增益 \hat{g}_c 。

技术人员将会理解在不偏离本发明范围的情况下可以对上面描述的实施方案进行各种修改。特别是, 图 3 和 4 中的编码器和解码器可以用软件或硬件实现, 也可以软硬结合来实现。尽管上面的描述集中于 GSM 蜂窝电话系统, 本发明也能够很好地应用于其它的蜂窝无线电系统以及非无线电通讯系统如互联网。本发明还可以应用于数据存储中对语音数据的编码和解码过程。

本发明可以应用于 CELP 编码器, 以及 ACELP 编码器。然而, 因为 CELP 编码器有一个固定码书用于产生量化矢量 $d(i)$, 并且给定量化矢量中脉冲的幅度可以变化, 用于缩放激励矢量 $c(i)$ 幅度的缩放因子 k 并不是 (如等式 (10) 那样) 脉冲数 m 的简单函数。而且, 每个固定码书的每个量化矢量 $d(i)$ 的能量必须被计算出来并且该能量相对于例如最大量化矢量能量的比例要确定。该比例的平方根给出缩放因子 k 。

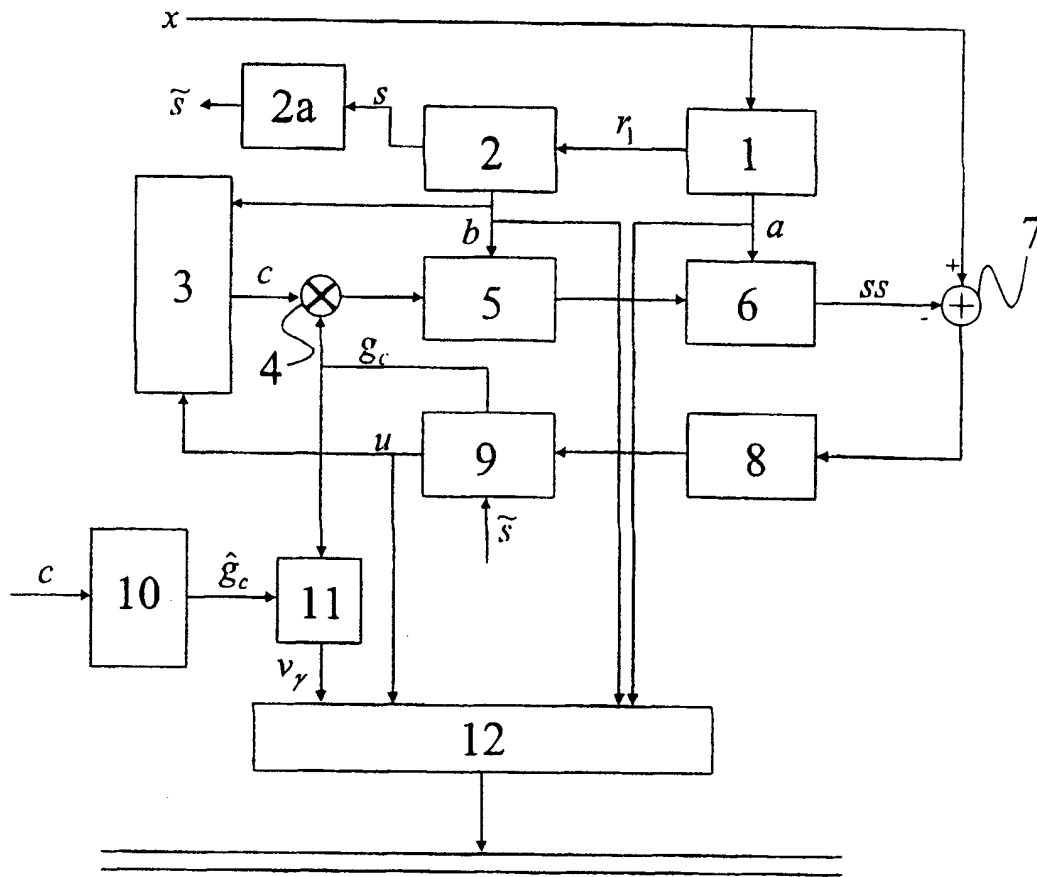


图 1

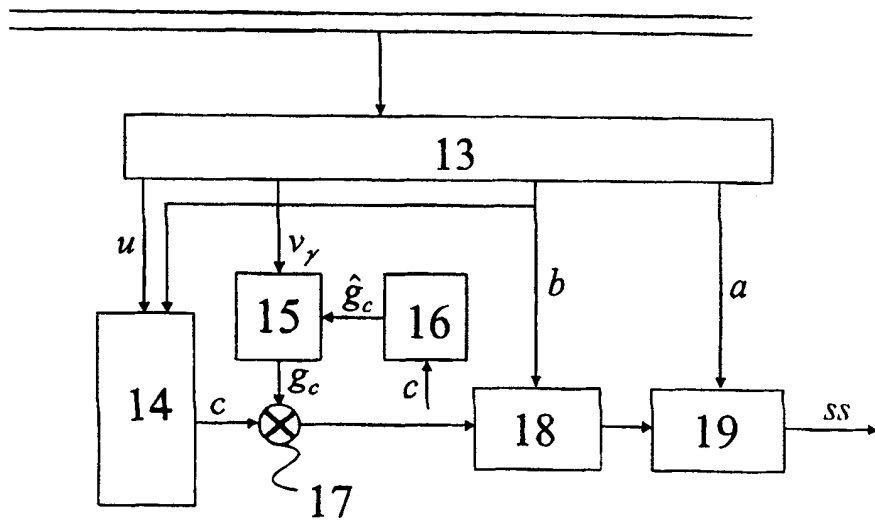


图 2

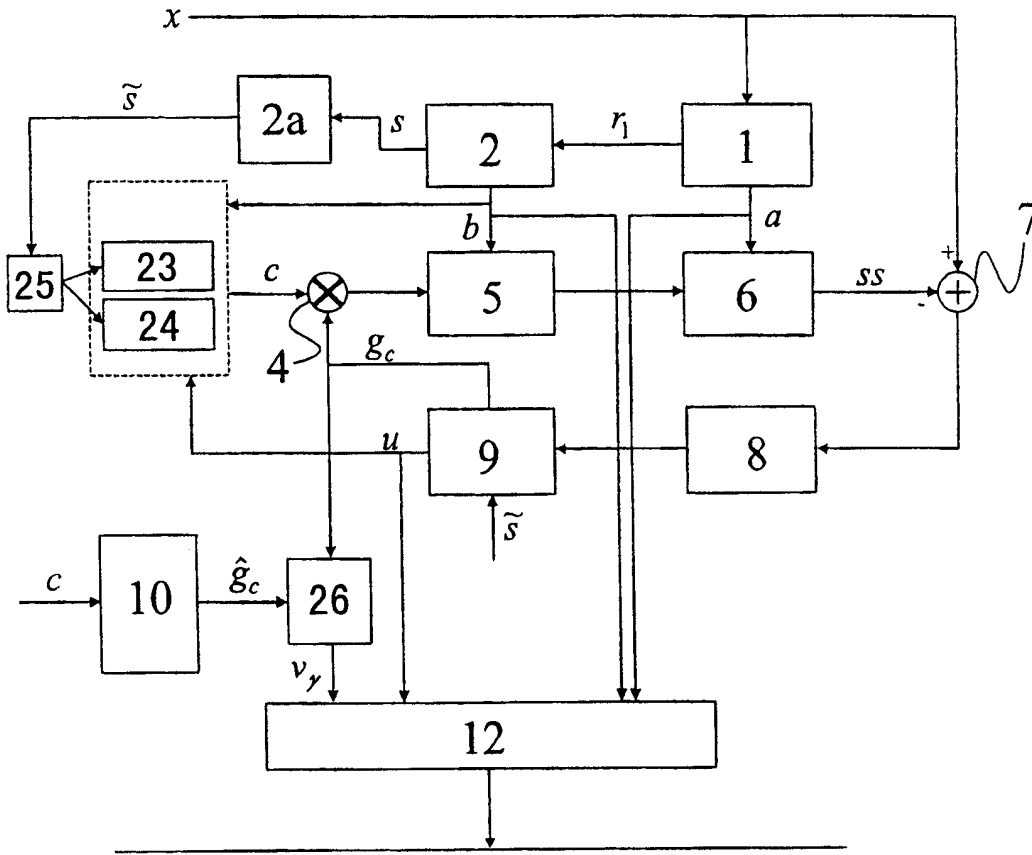


图 3

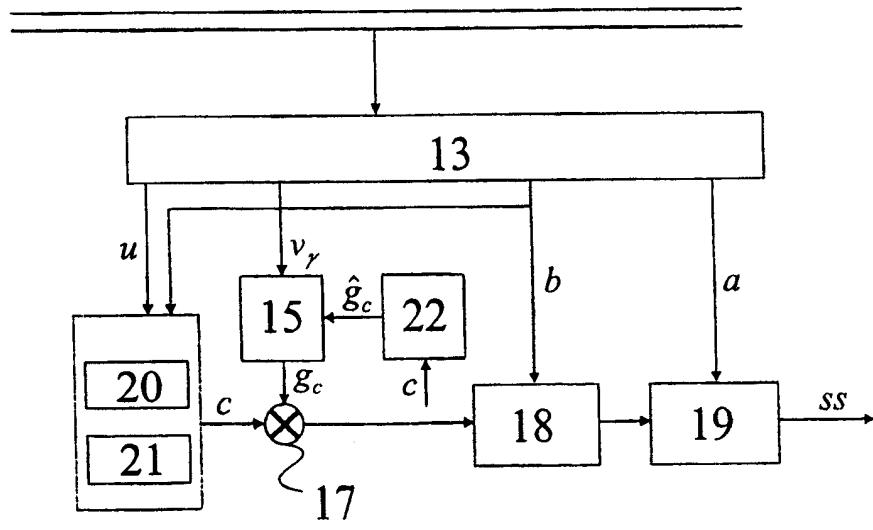


图 4