

一种降低 ACELP 码书搜索复杂度的高效算法^{*}

金 静[†], 张天骐, 万义龙, 邓 灵

(重庆邮电大学 信号与信息处理重庆市重点实验室, 重庆 400065)

摘 要: 为了降低代数码激励线性预测(algebraic code-excited linear prediction, ACELP)语音编码算法的复杂度,以便更好地实时实现,提出了一种有效的改进算法。在自适应码书搜索上提出了不连续的开环基音搜索算法,利用时间抽取因子对不同延迟段语音样点进行不连续抽取;在代数码书的搜索上提出了一致脉冲替换法,采用脉冲位置预选和循环判断机制控制码书搜索的次数。以 G. 729A 为实验平台进行仿真,仿真结果表明,改进的算法在保证语音质量的情况下,有效降低了 ACELP 码书搜索的复杂度。

关键词: 代数码激励线性预测; 码书搜索; G. 729A; 搜索复杂度

中图分类号: TN912.32 **文献标志码:** A **文章编号:** 1001-3695(2013)12-3698-04

doi:10.3969/j.issn.1001-3695.2013.12.046

Efficient complexity reduction in codebook search for ACELP

JIN Jing[†], ZHANG Tian-qi, WAN Yi-long, DENG Ling

(Chongqing Key Laboratory of Signal & Information Processing, Chongqing University of Posts & Telecommunications, Chongqing 400065, China)

Abstract: In order to reduce the complexity of ACELP (algebraic code excited linear prediction), and achieve a better operation in real-time system, this paper proposed an effective algorithm. First, it proposed the discontinuous open-loop pitch estimation in adaptive codebook search, used the time decimation factor in different delays, and then proposed the unified pulse-replacement search algorithm in stochastic codebook search, used the prediction technique and judging mechanism to control the number of codebook search. Finally, it simulated the improved algorithm in G. 729A. The result shows that the proposed algorithm not only acquires complexity reduction but also get a high speech quality.

Key words: ACELP; codebook search; G. 729A; search complexity

0 引言

码激励线性预测(code-excited linear prediction, CELP)是一种有效的中低速率语音压缩编码技术,具有编码速率低,合成语音质量高等特点^[1]。代数码激励线性预测(ACELP)不仅继承了 CELP 的优点,而且更具优越性,其采用的代数码本具有稀疏性,不需要存储,能够进行高效搜索,它在 CELP 的基础上将语音压缩效率提高了几倍,所以被广泛应用于数字通信领域。近年来,许多基于 ACELP 的语音编码器已被标准化,如 G. 723.1、G. 729、AMR 等。其中 G. 723.1、G. 729 是 VoIP 系统中常用的语音编码标准^[2],AMR 是 3GPP 制定的应用于第三代移动通信 WCDMA 中的语音压缩编码^[3]。

G. 729 是 ITU-T 于 1996 年制定的基于共轭结构代数码激励线性预测(conjugate structure-algebraic code excited linear prediction, CS-ACELP)算法的语音压缩编码标准^[4],G. 729A 为其简化版本。文献[5]中分析了 CS-ACELP 的复杂度,其中自适应码书的搜索占到整个运算量的 27.3%,代数码书的搜索占

到 30.4%。由此可见,码书搜索是 ACELP 编码方式的主要运算负荷所在,因此,必须降低 ACELP 码书搜索的复杂度以达到更好的实时通信^[6,7]。

自适应码本搜索也叫做基音周期估计,自相关法是其最经典的方法,即利用语音信号的短时自相关函数来估计基音周期。目前国际语音编码标准中推荐的代数码书搜索的方法主要有聚焦搜索法、深度优先树搜索法^[8]和全域脉冲替换法^[9],此外还有一些改进的算法,如最不重要脉冲替换法、非迭代脉冲替换法^[10,11]。为降低 ACELP 码书搜索的复杂度,本文提出了一种有效的改进算法,在自适应码书搜索的开环部分提出了不连续的基音搜索算法,在代数码书的搜索部分提出了一致脉冲替换法,最后以 G. 729A 为平台进行仿真,验证算法的有效性。

1 ACELP 码书搜索

1.1 自适应码书的搜索

自适应码书的搜索采用开环搜索和闭环搜索相结合的方法

收稿日期: 2013-04-24; **修回日期:** 2013-05-27 **基金项目:** 国家自然科学基金资助项目(61071196,61102131);国家教育部新世纪优秀人才支持计划项目(NCET-10-0927);信号与信息处理重庆市市级重点实验室建设项目(CSTC2009CA2003);重庆市杰出青年基金资助项目(CSTC2011jjjq40002);重庆市自然科学基金资助项目(CSTC2010BB2398,CSTC2010BB2409,CSTC2010BB2411,CSTC2012JJ40008);重庆市教育委员会科研资助项目(KJ120525)

作者简介: 金静(1988-),女(通信作者),陕西西安人,硕士研究生,主要研究方向为变速率语音编码、中低速率语音编码(jinjing19888464@163.com);张天骐(1971-),男,四川眉山人,教授,主要研究方向为扩频通信、盲信号和微弱信号处理等;万义龙(1989-),男,江西南昌人,硕士研究生,主要研究方向为语音信号处理、语音识别;邓灵(1988-),女,湖北宜昌人,硕士研究生,主要研究方向为跳频信号的细微特征提取、识别及网台分选。

法。首先进行开环搜索,在基音周期所有可能的范围内找到一个粗略的估计值,然后通过闭环搜索确定基音周期值。

1.1.1 开环基音搜索

以 G. 729A 为例(本文所述算法均以 G. 729A 为例,其他 ACELP 算法类似),开环基音搜索是每帧进行一次,将加权的语音信号采用自相关法估算基音:

$$R(k) = \sum_{n=0}^{79} s_w(n) s_w(n-k) \quad (1)$$

为避免基音整数倍数的影响,将基音延迟 k 取值范围 $[20 \sim 143]$ 分成以下三段:

$$i = 1:20, \dots, 39$$

$$i = 2:40, \dots, 79$$

$$i = 3:80, \dots, 143$$

分别在这三段内求取 $R(k)$ 最大值 $R(t_i)$ ($i=1,2,3$),再对三个最大值分别进行归一化:

$$R'(t_i) = \frac{R(t_i)}{\sqrt{\sum_{n=0}^{79} s_w^2(n-t_i)}} \quad (2)$$

注意:在三个归一化的相关值中,要选取较低范围的延迟值作为最佳基音估计值 T_{op} 。

1.1.2 闭环基音搜索

闭环基音搜索是在开环基音搜索的基础上基于子帧进行的,要保证原始语音与合成语音之间的均方误差最小。G. 729A 自适应码书搜索,每一子帧做一次,第一子帧在延迟范围 $\left[19 \frac{1}{3}, 84 \frac{1}{3}\right]$,使用分辨率为 $1/3$ 的分数延迟 T_1 ,对于延迟范围在 $[85, 143]$,只使用整数基音延迟。对于后续子帧,只在范围 $\left[\text{int}(T_1) - 5 \frac{2}{3}, \text{int}(T_1) + 4 \frac{2}{3}\right]$ 中采用分辨率为 $1/3$ 的分数延迟 T_2 ,这里 $\text{int}(T_1)$ 是第一子帧分数基音延迟的整数部分。

每一子帧的最佳延迟值是由闭环分析得到。对于 G. 729A,第一子帧,延迟值 T_1 的搜索范围为 $[T_{op} - 3, T_{op} + 3]$,第二子帧,最佳延迟值 T_2 的搜索范围为 $[T_{op} - 5, T_{op} + 4]$ 。

1.2 代数码书的搜索

1.2.1 代数码本的结构

代数码字被设计成一个稀疏向量,其非零脉冲的个数和幅值是固定的,只有它的位置是变的,它的维数一般被设计成子帧的长度,所以,代数码书的搜索实际上就是搜索代数码字中非零值的最佳位置。G. 729A 中,代数码字是由四个非零脉冲组成的向量,脉冲的位置和幅度见表 1。

表 1 G. 729A 固定码本结构

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

代数码书的搜索采用加权输入语音与加权重构语音的均方误差 E_w 最小为原则,找出最佳码矢量:

$$E_w = \|x - gHc_k\|^2 \quad (3)$$

式中: c_k 是索引为 k 的码矢量; x 是代数码书搜索的目标矢量,由加权原始语音减去加权合成滤波器的零输入响应和自适应

码本的贡献而得到的; g 是固定码本增益; H 是下三角 Toeplitz 矩阵,由感知加权合成滤波器的脉冲响应 $h(n)$ 截断而成的。

令 $\partial E_k / \partial g = 0$, 可求得增益 g , 再代回式 (3), 则最佳码矢量应使式 (4) 最大化:

$$A_k = \frac{(c_k)^2}{E_k} = \frac{(x^T H c_k)^2}{c_k^T H^T H c_k} = \frac{(d^T c_k)^2}{c_k^T \Phi c_k} \quad (4)$$

其中: $d = H^T x$ 是代数码书搜索的目标信号 $x(n)$ 和脉冲响应 $h(n)$ 的互相关信号,则

$$d(n) = \sum_{i=n}^{39} x(i) h(i-n) \quad n=0, \dots, 39 \quad (5)$$

$\Phi = H^T H$ 是 $h(n)$ 的自相关矩阵,其元素为

$$\varphi(i, j) = \sum_{n=j}^{39} h(n-i) h(n-j) \quad i=0, \dots, 39; j=i, \dots, 39 \quad (6)$$

d 和 Φ 可在码书搜索之前计算。由于 c_k 仅有有限个少数非零脉冲,式 (4) 的分子和分母可相应地简化为式 (7) (8):

$$C_k = \sum_{i=0}^3 s_i d(m_i) \quad (7)$$

$$E_k = \sum_{i=0}^3 \varphi(m_i, m_i) + 2 \sum_{i=0}^2 \sum_{j=i+1}^3 s_i s_j \varphi(m_i, m_j) \quad (8)$$

为了简化码书搜索过程,通过参考信号 $b(n)$ 作脉冲幅度预判决,即设置某位置脉冲的符号与 $b(n)$ 在此位置的符号相同。对于 G. 729A, $b(n) = d(n)$, 则有

$$s_i = \text{sign}(b(m_i)) = \text{sign}(d(m_i)) \quad (9)$$

$$C_k = |d(m_0)| + |d(m_1)| + |d(m_2)| + |d(m_3)| \quad (10)$$

$$E_k/2 = \varphi'(m_0, m_0) + \varphi'(m_1, m_1) + \varphi'(m_0, m_1) + \varphi'(m_2, m_2) + \varphi'(m_0, m_2) + \varphi'(m_1, m_2) + \varphi'(m_3, m_3) + \varphi'(m_0, m_3) + \varphi'(m_1, m_3) + \varphi'(m_2, m_3) \quad (11)$$

式中:

$$\varphi'(i, j) = \begin{cases} 0.5\varphi(i, j) & i=j \\ \text{sign}(d(i)) \text{sign}(d(j)) \varphi(i, j) & i \neq j \end{cases} \quad (12)$$

1.2.2 代数码本的搜索

为简化码书搜索过程, G. 729A 采用 DFTS(深度优先树搜索)算法。此算法将每个子帧中可能的位置分成了五个脉轨(表 2),沿着有较优值的路径进行搜索。

表 2 G. 729A 深度优先树搜索的轨道分布

脉冲	轨道	位置
i_0	T_0	0 5 10 15 20 25 30 35
i_1	T_1	1 6 11 16 21 26 31 36
i_2	T_2	2 7 12 17 22 27 32 37
T_3	3	8 13 18 23 28 33 38
i_3	T_4	4 9 14 19 24 29 34 39

如图 1 所示,搜索过程分为两部分:

$$\left\{ \begin{array}{l} \text{part 1} \left\{ \begin{array}{l} T_2 \rightarrow T_3 \rightarrow T_0, T_1 \\ T_3 \rightarrow T_0 \rightarrow T_1, T_2 \end{array} \right. \\ \text{part 2} \left\{ \begin{array}{l} T_2 \rightarrow T_4 \rightarrow T_0, T_1 \\ T_4 \rightarrow T_0 \rightarrow T_1, T_2 \end{array} \right. \end{array} \right.$$

图 1 G. 729A 深度优先树搜索示意图

先从轨道 T_2 中找到使 $|d(n_i)|$ 最大的两个值,其对应位置与轨道 T_3 的八个位置分别进行组合,最终确定最佳位置组合,再对轨道 T_0 和 T_1 进行全搜索。之后,改变脉冲轨道的搜索顺序重复上述步骤。

2 改进的码书搜索算法

2.1 改进的自适应码书搜索

本文只对自适应码书搜索的开环搜索部分进行改进。对

式(1)进行了简化:

$$R_1(k) = \sum_{n=0}^{79} s_w(T_c n) s_w(T_c n - k) \quad T_c = 1, k \in [20, 39]$$

$$R_2(k) = \sum_{n=0}^{39} s_w(T_c n) s_w(T_c n - k) \quad T_c = 2, k \in [40, 79]$$

$$R_3(k) = \sum_{n=0}^{25} s_w(T_c n) s_w(T_c n - k) \quad T_c = 3, k \in [80, 142]$$

其中: T_c 为时间抽取因子, 即对每一帧语音的样点进行抽取, 这样每个时延的相关性将成倍地降低, 语音的质量也将有所下降, 因此本文将 T_c 对语音质量的影响作了以下研究。

本文任意选取 20 段语音信号 (10 段男声, 10 段女声), 以 G. 729A 为实验平台, 采用 TIU-T P. 862. 1^[12] 建议“窄带电话网络和语音编/解码器的端到端语音质量的客观评价方法” PESQ_LQO (感知语音质量评价) 进行测评。图 2 是其中任意三段语音在 $T_c = 1, 2, \dots, 10$ 时的 PESQ 值。

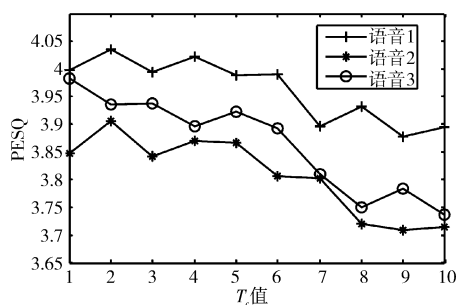


图2 T_c 值对语音质量的影响

从图 2 可以看出, T_c 值越大, 语音的质量下降越快, 但当 $T_c \leq 5$ 时, 语音的质量只有略微的下降, 因此本文建议 $T_c = 1, 2, \dots, 5$ 。

图 3 是改进的开环搜索部分的流程。

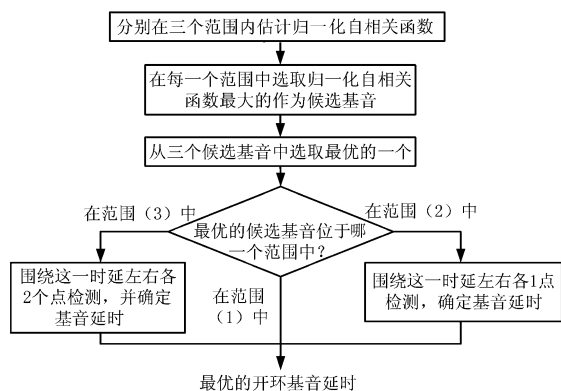


图3 改进开环基音搜索流程

算法步骤:

a) 分别在 $k \in [20, 39]$ 、 $[40, 79]$ 和 $[80, 142]$ 三个范围内, 采用时间抽取因子 T_c 计算归一化自相关函数, 并在每一个范围内选取归一化自相关函数值最大者作为候选值, 将得到三个候选值。

b) 在三个候选值中选取最优值。

c) 判定最优值所对应的延时 k 值在哪个范围内。如果在范围 2 (即 $[40, 79]$) 中, 再围绕此延时值进行 ± 1 检测; 如果在范围 3 (即 $[80, 142]$) 中, 则围绕此时延值进行 ± 2 检测。

d) 确定最优的开环基音延时。

2.2 改进的代数码书搜索

本文给出的一致脉冲替换法的基本流程如图 4 所示。下

面以 G. 729A 为例对算法进行详细说明。

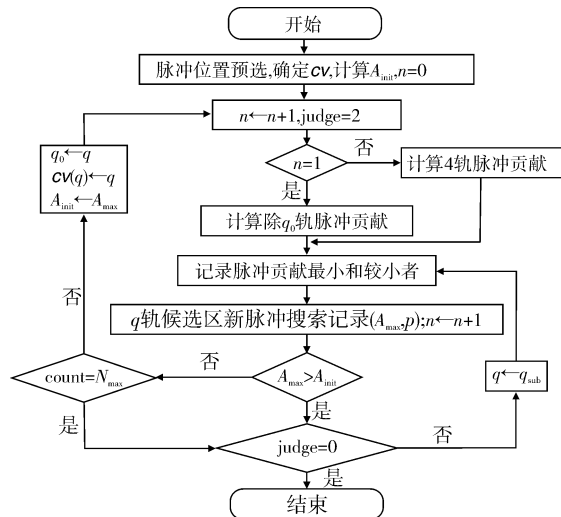


图4 一致脉冲替换法流程

a) 根据式(4)计算单个脉冲的重要性 (即去掉此脉冲而仅保留其他三个脉冲的 A_k 值), 以递减顺序排列脉冲位置, 用前 M 区的脉冲位置作为候选替换脉冲, 取每轨中最重要脉冲位置初始化码向量 cv 。根据式(4)计算 cv 的 A_{init} , 初始化 $n \leftarrow 0$ (n 控制循环的次数, 即搜索的轮数)。

b) 脉冲替换按以下步骤进行:

(a) 令 $n \leftarrow n + 1$, $judge = 2$ 。

(b) 若 n 为 1, 计算 cv 中所有脉冲的贡献值; 否则, 只计算 cv 中去掉 q_0 轨脉冲的贡献值。

(c) 选择贡献值最小和较小者, 记为 (Q, p) (Q_{sub}, q_{sub})。其中, q 和 q_{sub} 表示待替换脉冲所在的轨道, 取值为 0, 1, 2, 3。

(d) 候选区内属于 q 轨的新脉冲与保留的三个脉冲分别组合, 根据式(4)计算 A_k 值, 将最大值记为 (A_{max}, p) , p 表示替换脉冲的位置。judge 减 1。

(e) 若 $A_{max} > A_{init}$, 跳至步骤(f); 否则跳至步骤(g)。

(f) 若 $n = N_{max}$ (N_{max} 为最大迭代次数), 则结束; 否则, $q_0 \leftarrow q$, $cv(q) \leftarrow p$, $A_{init} \leftarrow A_{max}$, 跳至步骤(a)。

(g) 若 $judge = 0$, 结束; 否则, $q \leftarrow q_{sub}$, 跳至步骤(d)。

3 结果分析

3.1 改进开环基音搜索复杂度分析

如表 3 中所列, G. 729A 在开环阶段所需的搜索量为 124, 而本文提出的改进算法只需要 65, 即第一步搜索量为 62, 第三步大约需要 $3 \times (4 \times 64/124 + 2 \times 40/124)$, 总计为 65。也就是说, 在开环搜索阶段, 改进算法所需的搜索量只有 G. 729A 所需搜索量的 52.42%, 大大降低了搜索复杂度。

表3 改进开环基音搜索复杂度分析

算法	搜索量	百分比/%
G. 729A	124	100
改进算法	65	52.42

3.2 改进代数码书搜索复杂度分析

ACELP 代数码书的搜索需要不断计算和比较 A_k 值, 所以其复杂度可以以计算 A_k 的次数来衡量^[13, 14]。对 G. 729A, 第一步需搜索 $2 \times 80 = 160$ 次, 第二步与第一步类似, 只是用 T_4

替换掉 T_3 。因此, G. 729A 深度优先树搜索的搜索量为 $\text{load} = 2 \times 160 = 320$, 如表 4 所示。

表 4 改进的代数码书搜索复杂度分析

算法	搜索量	百分比/%
G. 729A	320	100
改进算法	$18 \leq \text{load} \leq 35$	5.6 ~ 10.9

本文改进算法的算法复杂度为

$$\text{load} = 4N + \sum_{i=1}^N [T_1(i) + T_2(i)] \quad (13)$$

其中: N 为实际执行的循环次数; $T_1(i)$ 、 $T_2(i)$ 分别表示当循环到第 i 次时, 第 1 级和第 2 级所搜索计算的 A_k 的次数。其所需的搜索量为 $18 \leq \text{load} \leq 35$, 且经本文反复测试表明, 当 $M = 4$, $N_{\max} = 2$ 时, 所需搜索量最低。

3.3 合成语音质量的主观评价

主观测试采用人耳的非正式倾听。算法解码输出的文件不能够直接播放, 可以用 MATLAB 中的 `wavwrite()` 函数, 将解码输出的文件写为 .wav 文件, 也可以采用 Cool Edit 软件将解码输出文件转换为 .wav 格式的文件, 这样就可以播放重建语音。经过试听, 本文算法合成的语音与原始语音在听觉上没有什么差别, 均可以清晰地听到语音内容。所以, 经主观测试, 本文算法在合成语音的清晰度上与原始语音几乎无差别。

3.4 合成语音质量的客观评价

本文采用 PESQ_LQO 进行测评, 得到的 MOS_LQO 范围为 1~4.5 分。采用十段男声和十段女声, 有中文也有英文, 得到的 PESQ 值如表 5 和 6 所示。

表 5 十段男声 PESQ 值比较

男声	PESQ		
	标准算法	改进算法	差值
第 1 段	3.943 1	3.864 6	0.075 8
第 2 段	3.967 4	3.905 4	0.062 0
第 3 段	3.903 6	3.876 0	0.027 6
第 4 段	3.911 8	3.846 4	0.065 4
第 5 段	3.904 7	3.861 3	0.043 4
第 6 段	3.919 2	3.823 5	0.095 7
第 7 段	3.912 4	3.843 1	0.069 3
第 8 段	4.036 5	3.934 0	0.012 5
第 9 段	3.952 2	3.896 4	0.055 8
第 10 段	3.974 1	3.955 9	0.018 2
平均值	3.942 5	3.880 7	0.052 5

表 6 十段女声 PESQ 值比较

女声	PESQ		
	标准算法	改进算法	差值
第 1 段	3.935 3	3.921 0	0.014 3
第 2 段	3.890 9	3.814 0	0.076 9
第 3 段	3.965 2	3.877 4	0.087 8
第 4 段	3.940 6	3.891 5	0.049 1
第 5 段	3.867 7	3.875 7	0.008 0
第 6 段	3.896 5	3.818 3	0.078 2
第 7 段	3.902 2	3.830 6	0.071 6
第 8 段	3.905 4	3.870 6	0.064 8
第 9 段	3.925 8	3.884 7	0.041 1
第 10 段	3.865 8	3.875 3	0.009 5
平均值	3.909 5	3.866 0	0.050 1

从表 5 和 6 中可以看出, 改进的算法在语音质量上只有略微的下降, 平均在 5% 左右, 不会影响正常的通信。

4 结束语

本文通过对自适应码书搜索的开环部分和代数码书搜索部分算法进行改进, 降低了 ACELP 码书搜索的复杂度。改进的算法在开环部分提出了不连续的基音搜索算法, 使复杂度降低了将近一半; 在代数码书的搜索部分提出了一致脉冲搜索算法, 使得代数码书搜索复杂度至少降低为原来的 1/10。最终实验结果表明, 在保证合成语音质量的条件下, 改进算法大大降低了码书搜索的复杂度, 有利于更好地实时实现。因此, 如何在 VoIP 上实现该算法是未来进一步研究的工作。

参考文献:

- [1] 王洪. 低速率语音编码[M]. 北京: 国防工业出版社, 2006: 49-52.
- [2] 屈丹, 王波. VoIP 语音处理和识别[M]. 北京: 国防工业出版社, 2010: 43-95.
- [3] 鲍长春. 数字语音编码原理[M]. 西安: 西安电子科技大学出版社, 2007.
- [4] ITU-T Recommendation G. 729, Coding of speech at 8 kbit/s using conjugate structure algebraic code excited linear prediction (CS-ACELP)[S]. 2007.
- [5] YE H C Y. Effective complexity reduction for G. 729 standard[C]//Proc of International Conference on Electrical Engineering/Electronics Computer Telecommunications and Information Technology. 2010.
- [6] CHEN Fu-kun, CHEN Guan-ming. Complexity scalability design in coding of the adaptive codebook for ITU-T G. 729 speech coder[C]//Proc of IEEE International Conference on Information and Communication Security. 2011.
- [7] YE H C Y, ZHOU Chang-zhi. An efficient complexity reduction algorithm for G. 729 speech codec[J]. Computers and Mathematics with Application, 2012, 64(5): 887-896.
- [8] SALAMI R, LAFLAMME C, BESSETTE B, et al. ITU-T G. 729 Annex A: reduced complexity 8 kb/s CS-ACELP codec for digital simultaneous voice and data[J]. IEEE Commutation Magazine, 1997, 35(9): 56-63.
- [9] PARK H, KIM Y, YOO J. Efficient codebook search method for AMR wideband speech codecs[J]. IEICE Trans on Information and Systems, 2004, E87-D(8): 2114-2120.
- [10] LEE E D, YUN S H, LEE S I, et al. Iteration-free pulse replacement method for algebraic codebook search[J]. Electronics Letters, 2007, 43(1): 59-60.
- [11] LEE E D, SUNG J M, SONG Y J, et al. Fixed codebook search method through iteration-free global pulse replacement and speech coder using the same method; world intellectual property organization, WO 2008/044817 A1[P]. 2008-04-17.
- [12] ITU-T Recommendation P. 862. 1, Telephone transmission quality methods for objective and subjective assessment of quality: mapping function for transforming P. 862 raw result scores to MOS-LQO[S]. 2003.
- [13] 张德民, 宋铁成, 张天骐, 等. 一种快速的代数码书搜索方法[J]. 重庆邮电大学学报: 自然科学版, 2010, 22(2): 183-187.
- [14] 刘晓明, 冯晓荣, 田雨, 等. 一种改进的 G. 729 标准固定码本快速搜索算法[J]. 计算机应用研究, 2011, 28(9): 3298-3300.