

# LD-CELP 语音编码算法中矢量量化过程的改进

黄德智, 马尽文

(汕头大学数学研究所, 广东汕头 515063)

摘要: 本文介绍了 LD-CELP 算法的基本原理, 在分析其编码的矢量量化过程的基础上提出了一种旨在提高编码速度的改进方案. 模拟实验表明, 改进算法的编码速度平均提高了一倍. 虽然信噪比有所下降, 但下降幅度仅为 1.2dB, 依然能够保证编码质量.

关键词: LD-CELP 算法; 矢量量化; 语音编码

中图分类号: TN912.3 文献标识码: A 文章编号: 0372-2112(2001)10-1415-03

## An Improved Scheme of Vector Quantization in LD-CELP Speech Coding Algorithm

HUANG De zhi, MA Jir wen

(Institute of Mathematics, Shantou University, Shantou, Guangdong 515063, China)

Abstract: In this paper, the basic principle of LD-CELP (Low-Delay Code Excited Linear Prediction) Algorithm is introduced. Based on a detailed analysis of the vector quantization of the algorithm, an improved scheme is proposed to increase the coding speed of LD-CELP algorithm. It is shown by simulation experiments that with the improved scheme, the coding speed has been doubled, while the SNR has decreased by about 1.2dB which is small enough to guarantee the quality of the coding.

Key words: LD-CELP algorithm; vector quantization; speech coding

### 1 引言

国际电报电话咨询委员会(CCITT)于1992年9月公布了16k bps LD-CELP(Low Delay Code Excited Linear Prediction)语音编码算法<sup>[1]</sup>. 该算法全面满足CCITT对于16Kbps算法的性能要求. 目前, LD-CELP算法已被广泛应用, 特别是在会议电视系统、IP电话等领域<sup>[2]</sup>. 其特点是不直接从语音信号中提取语音短时谱与长时谱预测系数、增益因子等参数, 而利用50阶后向预测的方法得到. 然而算法的编码复杂度(单位时间内的运算量)相当大, 仅编码过程就需要20MIPS. 人们一直希望能够减小算法复杂度, 以满足更低延时的要求. 而减小算法复杂度的主要途径是减少50阶LPC的计算量和加速码本搜索过程<sup>[3]</sup>. 本文在分析该算法的矢量量化过程的基础上, 提出了一种降低码本搜索复杂度的改进方案, 并且通过模拟实验验证了这种方案的有效性.

### 2 LD-CELP 编码过程的基本原理

对语音信号按8kHz的采样率进行采样, 并转换成PCM信号. 每5个采样点组成一个语音矢量, 每4个矢量连成一帧. PCM信号经过均匀化处理后送入缓冲区. 通过感觉加权滤波处理, 根据最小均方误差原则得到激励矢量. 对激励矢量进行量化得到编码的结果: 一个10bit索引. 再通过对前时间段上量化过的增益进行线性预测得到激励矢量的增益. 在每帧的处理过程中, 10阶感觉加权滤波器、50阶综合滤波器和50

阶后向增益适配器的系数均更新一次. 算法编码器有5个样点的输入延迟, 即0.625ms, 总的一路编码延迟不超过2ms<sup>[4]</sup>.

### 3 矢量量化的改进方案与分析

LD-CELP采用的矢量量化方法是闭合码本搜索, 即根据最小均方误差原则对码本进行全局搜索. 这种全局最优搜索保证了编码质量, 但运算量相当大, 编码复杂度较高, 不利于语音实时处理. 虽然适当减小码本个数可以提高码本搜索的速度, 但也会降低编码质量. 本文保持码本个数不变, 根据LD-CELP编码中H矩阵的性质建立了一种快速搜索的方案.

考虑LD-CELP算法中矢量量化过程. 假设 $\sigma(n)$ 是后向增益适配器的输出,  $n$ 是时间.  $x(n)$ 是感觉加权滤波器处理语音矢量后得到的加权矢量. 定义距离为 $D_{i,j} = \|x(n) - H^0(n)g_{y_j}\|$ , 其中 $\|\cdot\|$ 是欧氏距离,  $y_j$ 是第 $j$ 个码本向量,  $g_i$ 是第 $i$ 个增益常数.  $H$ 由综合滤波器的传输函数的系数构成, 形式如下:

$$H = \begin{bmatrix} h(0) & 0 & 0 & 0 & 0 \\ h(1) & h(0) & 0 & 0 & 0 \\ h(2) & h(1) & h(0) & 0 & 0 \\ h(3) & h(2) & h(1) & h(0) & 0 \\ h(4) & h(3) & h(2) & h(1) & h(0) \end{bmatrix} \quad (1)$$

其中 $h(0)=1$ . 矢量量化过程中, 每帧第二个矢量量化后更新H. 矢量量化的目标是从128个码本矢量和8个增益常数中

找到与  $x(n)$  距离最短的码本矢量和增益常数, 也就是搜索  $\{i, j\}$ , 使  $D_{i,j}$  达到最小. 若令  $\hat{x}(n) = x(n)/\sigma(n)$ , 则量化过程等价于求  $\{i, j\}$  使  $D_{i,j} = \|\hat{x}(n) - g_i \mathbf{H}y_j\|$  达到最小.

由于  $\hat{x}(n) - g_i \mathbf{H}y_j = \mathbf{H}(\mathbf{H}^{-1}\hat{x}(n) - g_i y_j)$ , 考虑用  $\tilde{D}_{i,j} = \|\mathbf{H}^{-1}\hat{x}(n) - g_i y_j\|$  近似  $D_{i,j}$ . 根据范数的性质, 有

$$\frac{1}{\|\mathbf{H}^{-1}\|} \tilde{D}_{i,j} \leq D_{i,j} \leq \|\mathbf{H}\| \tilde{D}_{i,j} \quad (2)$$

当  $\mathbf{H}$  是正交阵时,  $\|\mathbf{H}\| = \|\mathbf{H}^{-1}\| = 1$ ,  $D_{i,j} = \tilde{D}_{i,j}$ , 两种搜索过程等价. 然而这里  $\mathbf{H}$  是非对称下三角阵,  $D_{i,j}$  和  $\tilde{D}_{i,j}$  不一定相等. 记  $D_{i_0,j_0} = \min D_{i,j}$ ,  $\tilde{D}_{i_1,j_1} = \min \tilde{D}_{i,j}$ . 根据  $D_{i,j}$  的定义,  $D_{i_1,j_1} \geq D_{i_0,j_0}$ . 因此, 求  $\tilde{D}_{i,j}$  的最小值问题得到的解  $\{i_1, j_1\}$  替代原问题的解  $\{i_0, j_0\}$  会产生一定的误差. 但编码过程中,  $\|\mathbf{H}\|$  和  $1/\|\mathbf{H}^{-1}\|$  都接近于 1, 根据不等式(2),  $D_{i,j}$  和  $\tilde{D}_{i,j}$  近似相等. 这样,  $\tilde{D}_{i,j}$  的最小化问题近似等价于  $D_{i,j}$  的最小化问题. 替代产生的误差是微小的, 后面的误差分析将进一步证明这一点.

改进方案归结为如何求使  $\tilde{D}_{i,j}$  达到最小值的下标  $\{i_1, j_1\}$ . 首先令  $v(n) = \mathbf{H}^{-1}\hat{x}(n)$ ,  $p_j(n) = v^T y_j$ , 若  $y_j$ ,  $p_j(n)$  和  $g_i$  无关, 则

$$\tilde{D}_{i,j}^2 = \|v(n)\|^2 - 2g_i v^T(n) y_j + g_i^2 \|y_j\|^2 \quad (3)$$

又令  $c_j = p_j(n)/\|y_j\|^2$ ,  $R_j = \|v(n)\|^2 - c_j^2$ , 则有:

$$\tilde{D}_{i,j}^2 = (g_i - c_j)^2 + R_j \quad (4)$$

从方程(4)可以看出, 当  $j$  固定时, 使  $\tilde{D}_{i,j}$  达到最小值的  $i$  正是使  $|g_i - c_j|$  最小的  $i$ . 依此, 我们给出矢量量化改进方案的计算程序:

**Step 0:** 在矢量量化之前, 首先计算  $2g_i, g_i^2$  和  $\|y_j\|^2$ .

**Step 1:** 对于输入矢量  $\hat{x}(n) = x(n)/\sigma(n)$ , 令  $\hat{x}(n) = (x_0, x_1, x_2, x_3, x_4)^T$ ,

$v(n) = (v_0, v_1, v_2, v_3, v_4)^T$ ,  $v(n)$  按下列等式计算, 并计算  $\|v(n)\|^2$ .

$$v_0 = x_0, v_1 = x_1 - h(1)v_0, v_2 = x_2 - h(2)v_0 - h(1)v_1; v_3 = x_3 - h(3)v_0 - h(2)v_1 - h(1)v_2,$$

$$v_4 = x_4 - h(4)v_0 - h(3)v_1 - h(2)v_2 - h(1)v_3$$

**Step 2:** 对  $j = 1, 2, \dots, 128$ , 求出  $i_1$ , 使  $\tilde{D}_{i_1,j}$  在  $\{\tilde{D}_{i,j}, i = 1, 2, \dots, 8\}$  中达到最小.

首先计算  $p_j(n) = v^T(n) y_j$ , 然后计算  $c_j = p_j(n)/\|y_j\|^2$ . 对  $i = 1, 2, \dots, 8$  求解  $|g_i - c_j|$  的最小值问题, 得到  $i_1$ , 并按式(4)计算  $\tilde{D}_{i_1,j}$ .

**Step 3:** 求出 128 个  $\tilde{D}_{i_1,j}$  的最小值  $\tilde{D}_{i_1,j_1}$ , 得到量化结果  $\{i_1, j_1\}$ .

LD-CELP 算法每帧处理过程中计算  $\mathbf{H}y_j$  和  $\|\mathbf{H}y_j\|^2$  需要执行  $20 \times 128$  次乘法, 4 个语音矢量量化需要  $20 + 48 \times 128$  次乘法, 共需 8724 次乘法. 改进方案对每一个矢量量化执行  $20 + 8 \times 128$  次乘法, 4 个矢量量化需要 4176 次乘法.

改进方案导致了一定的量化误差. 我们现在根据实验结果分析这种误差的大小及随时间的变化情况. 考虑  $m$  个语音矢量, 对第  $n$  个矢量, 记  $D_{i_0,j_0}(n) = \min D_{i,j}(n)$ ,  $\tilde{D}_{i_1,j_1}(n) =$

$\min \tilde{D}_{i,j}(n)$ , 平均误差定义如下:

$$E(m) = \frac{\sum_{n=1}^m [D_{i_0,j_0}(n) - \tilde{D}_{i_0,j_0}(n)]}{m} \quad (5)$$

其中  $\tilde{D}_{i_0,j_0} = \frac{1}{m} \sum_{n=1}^m \tilde{D}_{i_0,j_0}(n)$ , 并记  $R(m) = E(m)/\tilde{D}_{i_0,j_0}$ . 显然,  $R(m)$  反映了改进方案产生的误差与原方案固有误差的比. 为了观察两种方案下误差的散布情况, 令  $\bar{D}_{i_1,j_1} = \frac{1}{m} \sum_{n=1}^m \tilde{D}_{i_1,j_1}(n)$ , 定义如下两个样本方差:

$$S_1^2 = \frac{1}{m} \sum_{n=1}^m [D_{i_0,j_0}(n) - \bar{D}_{i_0,j_0}(n)]^2,$$

$$S_2^2 = \frac{1}{m} \sum_{n=1}^m [\tilde{D}_{i_1,j_1}(n) - \bar{D}_{i_1,j_1}(n)]^2.$$

我们采用女声测试数据 Test 1(见图 3(a)) 和有大量背景噪音的语音数据 Test 4(见图 3(d)) 进行连续编码实验. 按不同时间段计算出  $E(m)$ ,  $R(m)$ ,  $S_1^2$  和  $S_2^2$ , 如表 1 和 2 所示.

表 1 女声数据的处理结果

$m$	200	300	500	800	1000	1200	1400	1600
$E(m)$	0.040	0.072	0.094	0.091	0.085	0.078	0.086	0.085
$R(m)$	0.053	0.077	0.098	0.086	0.078	0.074	0.081	0.080
$S_1$	0.148	1.419	1.000	0.940	0.933	0.820	0.827	0.760
$S_2$	0.167	1.534	1.131	1.076	1.081	0.953	1.001	0.927

表 2 含噪数据的处理结果

$m$	200	300	500	800	1000	1200	1400	1600
$E(m)$	0.101	0.110	0.113	0.107	0.151	0.147	0.133	0.132
$R(m)$	0.124	0.135	0.141	0.134	0.158	0.131	0.108	0.098
$S_1$	0.132	0.135	0.136	0.124	0.117	3.815	3.225	3.029
$S_2$	0.191	0.231	0.217	0.204	4.366	4.042	3.661	3.438

从以上两表可以看出, 平均误差  $E(m)$  比较小, 并且对于  $m$  是稳定的. 相对于原方案的固有平均误差  $R(m)$ , 改进方案的误差增加了 10% 左右. 当无噪声时, 误差的增长率不超过 10%. 并且从长时段来看, 误差的增长率更低. 利用表中的数据可以得到以下两条  $R(m)$  的拟合曲线.

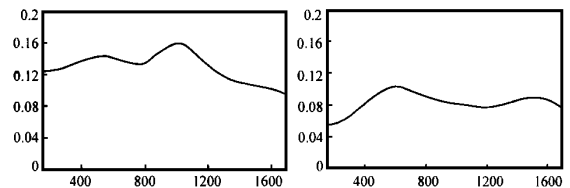


图 1 女声量化误差拟合曲线

图 2: A line graph showing the average error E(m) versus the number of vectors m for noisy speech data. The x-axis ranges from 0 to 1600, and the y-axis ranges from 0 to 0.2. The curve starts at approximately 0.10 for m=200, rises to a peak of about 0.15 at m=500, and then fluctuates between 0.10 and 0.15 for larger m values.

#### 4 模拟结果和性能比较

将 LD-CELP 算法中矢量量化部分切换成本文提出的改进方案, 就得到一种改进的 LD-CELP 算法, 记为 M-LDCELP, 并

将 ITU-T 公布的 LD-CELP 算法记为 ITU-LDCELP. 采用这两个算法, 分别对一组测试数据进行语音压缩编码实验. 待测试的语音数据如图 4 所示. 它们是 4 段语音的原始记录数据. Test1 和 Test2 是 ITU 提供的测试数据. Test3 和 Test4 是作者录制的

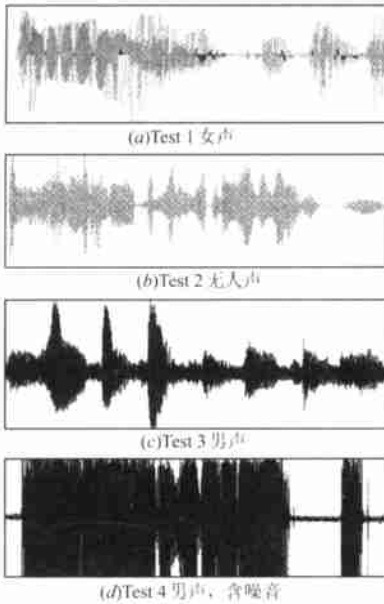


图 3 测试数据的波形图

采用 C 语言对 M-LDCELP 和 ITU-LDCELP 在 Pentium 200 微机上编程实现. 使用这两个算法分别对上述 4 段语音数据进行压缩编码, 记录其编码时间. 根据编码、解码的结果计算出每段语音上的信噪比 (SNR). 实验结果如表 3 所示, 其中 SNR 的单位为 dB, 编码时间的单位为 ms.

表 3 M-LDCELP 和 ITU-LDCELP 算法对于语音数据 Test1-4 的实验结果

数据	ITU-LDCELP SNR	M-LDCELP SNR	ITU-LDCELP 编码时间	M-LDCELP 编码时间
Test1	17.61	16.67	436	275
Test2	17.56	15.85	565	415
Test3	17.51	14.26	1182	522
Test4	10.23	9.81	1505	761

从表 3 可以看出, 改进的 LD-CELP 算法的编码速度平均提高了一倍, 而 SNR 仅平均下降 1.2dB.

## 5 总结

本文对 LD-CELP 算法的矢量量化过程进行分析, 提出了一种改进方案. 该方案可以大大减小算法的编码复杂度. 实验表明, 算法的编码速度平均提高了一倍. 虽然改进方案会产生一定的误差, 但这种误差相对较小, 并且不积累. 改进方案的信噪比仅下降了 1.2dB. 因此, 改进方案仍能保证很好的语音编码质量及语音信号的合成质量.

### 参考文献:

- [1] CCITT. Recommendation G. 728, Coding of speech at 16Kb/s using low-delay code excited linear prediction [S]. Geneva, 1992.
- [2] 金罗军, 等. 会议电视系统中 LD-CELP 算法的研究及其实现 [J]. 小型微型计算机系统, 1997, 18(12).
- [3] 张雪英. 减小 G. 728 语音编码计算量的三种方法 [J]. 太原理工大学学报, 1999, 30(3).
- [4] Juir-Hwey Chen. High quality 16kb/s speech coding with a one-way delay less than 2ms [J]. CH2847-2/90, 1990 IEEE ICASSP: 453-456.

### 作者简介:



黄德智 男, 1976 年 11 月出生于江西万年. 汕头大学数学所研究生. 研究方向为神经网络、信息处理与编码. 现已考入清华大学计算机科学与技术系, 攻读博士学位.



马尽文 男, 1962 年 7 月出生于陕西西安. 1992 年于南开大学概率论与数理统计专业博士生毕业, 现为汕头大学数学研究所教授. 1995~1996 年在香港中文大学计算机科学与工程学系访问. 目前已发表学术论文 40 余篇, 出版专著两部. 主要研究方向为: 神经网络理论及应用, 信息处理与编码, EM 算法, 模式识别与聚类分析等. 现已调入北京大学数学科学学院信息科学系.