

LSF 参数的模拟退火法连接分裂矢量量化

鲍长春, 卓 力, 王永会

(北京工业大学电子信息与控制工程学院, 北京 100022)

摘 要: 本文基于模拟退火算法, 提出了一种全局最优的模拟退火法连接分裂矢量量化 (SA-LSVQ) 方案. 实验表明, 该算法与 LBG 法相比, 有效地改善了 LSVQ 的性能.

关键词: 语音编码; 矢量量化; 线谱频率; 模拟退火

中图分类号: TN912.3 文献标识码: A 文章编号: 0372-2112 (2001) 01-0127-03

Simulated Annealing Linked Split Vector Quantizer of LSF Parameters

BAO Chang-chun, ZHUO Li, WANG Yong-hui

(School of Electronic Information and Control Engineering, Beijing Polytechnic University, Beijing 100022, China)

Abstract: Based on Simulated Annealing Algorithm, a Simulated Annealing Linked Split Vector Quantizer (SA-LSVQ) optimized globally is proposed in this paper. Experimental results show that the performance of LSVQ is improved effectively in comparison with LBG method.

Key words: speech coding; vector quantization; line spectrum frequency; simulated annealing

1 引言

在中低速率语音编码中, 以尽可能少的比特对 LSF 参数编码是人们研究的热点问题之一. 由于矢量量化在相同的编码比特数下能比标量量化给出更低的量化失真, 因此, 对 LSF 参数量化的研究主要集中在矢量量化上. 但将 10 维 LSF 参数作为一个矢量进行整体量化的方法并不可行, 这是因为这种方法存储量大, 计算复杂, 现有硬件条件无法承受, 因此, 人们大多采用次优矢量量化算法, 如多级矢量量化、分裂矢量量化和预测矢量量化等^[3-7].

在低比特率 LSF 分裂矢量量化中, 连接分裂矢量量化 (LSVQ)^[3] 的性能要优于传统的分裂式矢量量化器^[4,5]. 为了进一步改善这种量化器的性能, 本文借助了模拟退火算法提出了一种 LSF 参数的模拟退火法连接分裂矢量量化 (SA-LSVQ) 方案, 避免了 LBG 迭代算法容易陷入局部最优点的缺陷, 使 LSVQ 量化器的性能得到了改善.

2 LSVQ 的基本原理

LSVQ 将 LSF 矢量 $\{f_i\}, i = 1, 2, \dots, 10$ 分裂为低、中和高三个子矢量 $\{f_1, f_2, f_3\}$ 、 $\{f_4, f_5, f_6\}$ 和 $\{f_7, f_8, f_9, f_{10}\}$. 其码书训练算法如下:

(1) 用所有的 $\{f_4, f_5, f_6\}$ 训练矢量, 采用 LBG 算法训练中间码书 (COM);

(2) 根据 f_4 的大小, 将训练矢量 $\{f_1, f_2, f_3\}$ 分成四块, 分别采用 LBG 算法, 训练相应的低端码书 (COL1, COL2, COL3

和 COL4). 块划分的原则是, 在每块中 f_4 是等概率的, 即在每块区间中矢量个数是相同的:

$$P(0 \leq f_4 < x_1) = P(x_1 \leq f_4 < x_2) = P(x_2 \leq f_4 < x_3) = P(x_3 \leq f_4 < 4000) \quad (1)$$

式中, $P(a \leq f_4 < b)$ 表示 f_4 大于或等于 a 而小于 b 的概率. 本文由训练序列得到 $x_1 = 959, x_2 = 1163$ 和 $x_3 = 1318$.

(3) 根据 f_6 的值将训练矢量 $\{f_7, f_8, f_9, f_{10}\}$ 的集合划分为四块, 分别采用 LBG 算法, 训练相应的高端码书 (COU1, COU2, COU3 和 COU4); 块划分的原则是在每块中 f_6 是等概率的, 即:

$$P(0 \leq f_6 < x_1) = P(x_1 \leq f_6 < x_2) = P(x_2 \leq f_6 < x_3) = P(x_3 \leq f_6 < 4000) \quad (2)$$

本文由训练序列得到 $x_1 = 1837, x_2 = 2036$ 和 $x_3 = 2120$.

编码时, 先对中间矢量 $\{f_4, f_5, f_6\}$ 量化, 应用码书 COM 求取最小失真码矢, 得到量化矢量 $\{\bar{f}_4, \bar{f}_5, \bar{f}_6\}$ 的相应的码矢下标 index1. 然后, 根据已量化的 \bar{f}_4 , 低端 LSF 分矢量被划分到四个码书中的某一个. 例如, 如果 \bar{f}_4 为 900, 则用码书 COL1, 从中搜索出量化矢量 $\{\bar{f}_1, \bar{f}_2, \bar{f}_3\}$ 和相应的码矢下标 index2. 类似地, 根据 \bar{f}_6 , 可得到量化矢量 $\{\bar{f}_7, \bar{f}_8, \bar{f}_9, \bar{f}_{10}\}$ 和相应的码矢下标 index3.

译码时, 由 index1 在码书 COM 中直接得到量化中间矢量 $\{\bar{f}_4, \bar{f}_5, \bar{f}_6\}$, 再由公式 (1) 和 (2) 以及 index2 和 index3 得到响应的低端矢量 $\{\bar{f}_1, \bar{f}_2, \bar{f}_3\}$ 和高端矢量 $\{\bar{f}_7, \bar{f}_8, \bar{f}_9, \bar{f}_{10}\}$.

值得注意的是, 在信号传输中没有必要增加比特数来区

分用到了那一个 COL 和 COU 码书,因为在译码器中利用 index1 就可以获得有关的信息。

3 SA-LSVQ 码书训练算法

尽管 LSVQ 算法有效地提高了分裂矢量量化中高、低两个码书空间的利用率,但因 VQ 码书是由 LBG 迭代生成,所以,易陷入局部最优解。为得到全局最优码书,提高量化质量,我们尝试采用 SA 技术与 LSVQ 算法结合的方式,以获得性能方面的改善。SA 算法对初始码书不敏感,因此直接采用 LBG LSVQ 码书作为 SA 初始码书。

设训练序列为 $\{x_i^{(n)}\}$, 其中 $i=1, 2, \dots, 10$ 为 LSF 矢量的分量, $n=1, 2, \dots, N$ 为训练序列序号 (N 为训练矢量个数), 将 LSF 矢量 $\{x_i\}, i=1, 2, \dots, 10$ 分裂为三个子矢量 $\{x_1, 2, 3\}, \{x_4, 5, 6\}$ 和 $\{x_7, 8, 9, 10\}$, 分别称为低、中和高子矢量。与 LSVQ 相同, 根据 x_4 的值, 将训练序列低子矢量分为四块, 根据 x_6 的值, 将训练序列高子矢量分成四块。

如果中间码书的大小为 N , 设 R_i 是中间码书索引号为 i 的所有输入训练矢量, 则 $G_M = \{R_1, R_2, \dots, R_N\}$ 为其相应的分区。对此分区进行扰动, 扰动的方法是随机选择一个分区 R_i 中的一个训练矢量, 使其移动到分区 R_d 中, 形成两个新的分区 R_i^* 和 R_d^* , 这个扰动是否被接受, 取决于系统能量增量

$$D_M = D(X_M, Q_M^*, G_M^*) - D(X_M, Q_M, G_M) \quad (3)$$

式中, Q_M^* 与 G_M^* 分别为扰动后的码书和分区, Q_M 与 G_M 分别为扰动前的码书和分区。如果 $D_M \leq 0$, 则接受这次扰动, 否则应按 SA 接受规则判断。这时首先计算一个温度值 $T_m = T_0 \cdot e^{-D_M/T_m}$, 如果 $e^{-D_M/T_m} > r$ (r 为 $0 \sim 1$ 之间的随机数) 成立, 则接受扰动, 否则不接受扰动。这样就形成了中间码书 COM 新的分区所对应的形心组成的中间码书 $\{y_1, y_2, \dots, y_N\}$ 。在每个 $T_m = T_0 \cdot e^{-D_M/T_m}$ 下, 都要进行一定次数的扰动, 当扰动接受率降到给定值时, 就认为达到了热平衡, 才能进入下一个温度 $T_{m+1} = T_m \cdot e^{-D_M/T_m}$ 进行新的扰动。对中间码书进行扰动, 这个循环一直持续到冰点温度值 (即几乎不再有扰动被接受)。同理, 对高、低子码书施加扰动, 方法与中间码书扰动相同。由于在码书训练中采用了 SA 技术, 最终获得的码书 COM, COL 和 COU 应是全局最优码书的逼近。SA-LSVQ 码书的具体训练步骤如下

(1) 确定初始条件: 包括初始码书 COM, $r, T_0, \alpha = 0.9$, 最低扰动接受率 $P_0 = 0.4$, 最大扰动循环次数 L_{M0} ;

(2) 把输入训练矢量 $X = \{X_1, X_2, \dots, X_M\}$, 按码书 COM 映射成子集合 $G_M = \{R_1, R_2, \dots, R_N\}$, 计算形心

$$y_i = \frac{1}{R_i} \sum_{X \in R_i} X \quad (4)$$

(3) 计算方差 $D = d(X, Q_M, G_M)$ 作为系统能量;

(4) 扰动码书 COM, 随机选取一个训练矢量, 使其从 R_i 分区移动到 R_d 分区中, 获得新的分配 G_M^* , 计算新的分区 R_i^* 和 R_d^* , 以及新的形心 y_i^* 和 y_d^* , 形成新的码书 COM * ;

(5) 计算能量增量 $D_M = D(X_M, Q_M^*, G_M^*) - D(X_M, Q_M, G_M)$ 。若 $D_M \leq 0$ 时, 接受此次扰动; 若 $e^{-D_M/T_m} > r$, 则接受扰动; 若 $e^{-D_M/T_m} \leq r$, 则不接受扰动;

(6) 扰动固定次数后, 计算扰动接受率 P , 并判断 $P > P_0$ 是否成立, 如成立, 则计算 $m = m + 1$ 和 $T_m = T_0 \cdot e^{-m}$, 再返回到第 (4) 步; 如不成立, 则直接返回到第 (4) 步;

(7) 判断总循环次数, 当 $L = L_{M0}$ 时结束算法, 否则返回第 (4) 步继续;

(8) 分别对高端码书 COU 与低端码书 COL 进行扰动, 方法同 (1) ~ (7)。

4 实验与结论

以 dB 为单位的平均谱失真定义为^[4]:

$$SD = \frac{1}{N} \sum_{n=1}^{N_f} \left[10 \lg \frac{S_n(\cdot)}{S_n(\cdot)} \right]^2 \cdot d \quad (5)$$

式中, $S_n(\cdot)$ 和 $S_n(\cdot)$ 分别为第 n 帧的原始功率谱和量化功率谱, N_f 为测试数据的总帧数。以谱失真为客观量化性能评价指标。

语音以 8kHz 采样, 每个样值用 16 位线性码表示。分析帧长为 160 个样点, LP 分析窗长为 240 个样点。用于训练的语音约 16 万帧, 用于测试的语音约 25000 帧。

表 1 给出了 20bits (低、中和高分矢量各占 6, 7, 7bits) LSVQ 与 SA-LSVQ 量化器谱失真的对比结果。SA-LSVQ1 是每一温度下扰动的最大次数为 30 次的结果, SA-LSVQ2 是每一温度下扰动的最大次数为 250 次的结果。从表中可以看到, 当每一温度下扰动次数增加到 250 时, 谱失真在 2 ~ 4dB 之间比例由 6.204% 减少为 2.182%, 谱失真大于 4dB 的比例降为零, 平均谱失真由 1.26740dB 降为 1.19802dB, 量化器性能得到了有效改善。

实验发现, 扰动次数是影响 SA-LSVQ 性能的一个重要因素。由于参加训练的数据量庞大, 增加扰动次数势必增加算法的复杂度。在增大扰动次数的前提下, 如何减小 SA-LSVQ 训练码书的复杂度是今后研究的目标。

表 1 20bits LSVQ 与 SA-LSVQ 谱失真的比较

量化器	平均谱失真 (dB)	2 ~ 4dB	> 4dB
LSVQ	1.26740	6.204 %	0.0479 %
SA-LSVQ1	1.25784	5.677 %	0.0294 %
SA-LSVQ2	1.19802	2.182 %	0.0000 %

参考文献:

- [1] K. K. Paliwal and W. Bastiaan Kleijn. Speed coding and synthesis [J]. Elsevier Science B. V, 1995: 435 - 462.
- [2] 胡征, 杨有为. 矢量量化原理与应用 [M]. 西安电子科技大学出版社. 1988.
- [3] M. Y. Kim, N. K. Ha, S. R. Kim. Linked Split-Vector Quantizer of LPC Parameters [A]. Proceedings of IEEE ICASSP, 1995: 741 - 744.
- [4] K. K. Paliwal and B. S. Atal. Efficient vector quantization of LPC parameters at 24bits/frame [J]. IEEE Trans. Speech and Audio Processing, 1993, 1(1): 3 - 14.
- [5] 鲍长春等. 线谱频率参数的分裂矢量量化 [J]. 电子科学学刊, 1998, 20(4): 508 - 514.
- [6] Bao Changchun, and Dai Yisong. One-step interpolation predictive vec-

tor quantization of LSP parameters [J]. The Journal of China Universities of Posts and Telecommunications, 1996, 3(1): 21 - 26.

- [7] 卓力, 鲍长春. 一种高效、透明的线谱频率参数矢量量化器 [J]. 信号处理, 1999, 15(增刊): 154 - 158.
- [8] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi. Optimization by Simulated Annealing [J]. Science, 1983, 220: 671 - 680.
- [9] Kenneth Zeger, Jacques Vaisey, and Allen Gersho. Globally Optimal Vector Quantizer Design by Stochastic Relaxation [J]. IEEE Trans. on Signal Processing, Feb. 1992, 40(2): 310 - 322.

作者简介:



鲍长春 博士, 现为北京工业大学电子信息与控制工程学院教授, 副院长, 硕士生导师, 北京市高等学校青年学科带头人, 中国电子学会高级会员, 中国电子学会青年工作委员会委员, 通信学报编委, 长期从事语音压缩编码的研究工作, 在该领域发表学术论文 25 篇, 出版专著 1 部。



卓力 硕士, 现为北京工业大学电子信息与控制工程学院讲师, 计算机应用专业在职博士生, 研究兴趣为语音编码, 智能信息处理, 已发表多篇论文。

王永会 硕士, 现为美国密西西比洲立大学电子与计算机工程系博士生, 研究方向为语音与图像处理。

第 11 届三束(电子束、离子束、光子束)学术年会征文通知

由中国电子学会半导体与集成技术分会和生产技术分会联合主办“第 11 届三束(电子束、离子束、光子束)学术年会”, 将于 2001 年 10 月下旬在四川省成都市(暂定)召开, 会议由中国科学院光电技术研究所承办, 现将征文通知如下:

一、征文范围

- 1、电子束、离子束光刻及微细加工工艺
- 2、电子束、离子束辐照材料改性
- 3、电子束曝光及相关技术
- 4、半导体离子注入与离子束合成(包括器件和集成电路)
- 5、离子注入及相关技术
- 6、光子束的薄膜淀积、掺杂及材料加工工艺
- 7、X-射线光刻(包括掩模)
- 8、分子束外延(MBE)、CBE、MOCVD 及其它新外延技术
- 9、等离子体加工技术(包括 PECVD、溅射等技术)
- 10、借助电子束、离子束、光子束的测试分析技术
- 11、电子束、离子束、光子束与材料相互作用以及新效应的研究与计算机模拟
- 12、光学光刻(包括相移光刻、远紫外光刻、激光光刻)
- 13、抗蚀剂及干法刻蚀技术
- 14、半导体纳米技术, 纳米器件及应用
- 15、电子束、离子束、光子束在半导体材料器件电路及其领域中的应用
- 16、微机械、微系统研制及应用
- 17、生物芯片研制及应用
- 18、激光技术及其应用
- 19、微电子器件封装技术

二、截稿日期: 2001 年 5 月 15 日

录用通知发出日期: 2001 年 6 月 15 日

三、论文请寄: 四川成都市双流 350 信箱科技信息中心

联系人: 温尚明

联系电话: (028) 5100502 或 5100569

传 真: 028-5100070, 5865125

E-mail: infocenter @ioe. ac. cn.