EFFICIENT QUANTIZATION OF THE LPC COEFFICIENTS USING HYBRID CODEBOOK STRUCTURE

Kwok-Wah Law and Cheung-Fat Chan

Department of Electronic Engineering
City Polytechnic of Hong Kong

ABSTRACT An efficient k-parameters vector quantization algorithm using hybrid codebook structure is present. In the new scheme, the LPC vector is partitioned into 3 stages for vector quantization. In the first stage, full search codebook is used in order to maintain the spectral accuracy due to quantization. The binary-tree and quad-tree search codebook are used in the second and third stages for reducing the encoding complexity. By employing the hybrid codebooks in different stage, the proposed scheme can achieve 3.6 times complexity reduction when compared with the full-searched codebooks for all stages, and therefore very suitable for real-time implementation in low bit rate speech coding.

INTRODUCTION

Vector Quantization (VQ) of the LPC parameters has been shown to be the most successful method to encode the spectral envelop of the speech signal. However, this quantization scheme suffers many difficulties such as huge memory requirement, high computational complexity and large spread of distortion. In order to overcome these difficulties, Paliwal and Atal [1] proposed a Split Vector Quantization (SVQ) algorithm on the Line Spectrum Pairs parameters to quantize the LPC information with coding rate as efficient as 24 bits/frame. Although the spectral weighting was used to emphasize those important coefficients which are in the vicinity of the formant frequencies, quantizing the LPC parameters using the Euclidean distance of LSP parameters does not consist with classical LPC residual minimization concept. Therefore, minimizing the LSP parameters in the mean square sense, in fact, does not implicitly imply that the spectral distortion due to quantization is minimized.

Recently, we propose a split vector quantization algorithm of the k-parameters [2], which is consistent with the classical LPC residual minimization concept, for encoding the LPC coefficients. In order to reduce the encoding complexity but without sacrificing the performance of the quantizer severely, a hybrid codebook structure is thus applied to the newly developed split vector quantization system. For a 3 stages hybrid codebooks split VQ, the codebook in first stage, which consists of codewords having the highest spectral sensitivity, can be configured as full-searched codebook in order to minimized the spectral distortion due to quantization. The rest of the stages, in which the codewords have lower spectral sensitivity, are employed quad tree-searched or binary tree-searched codebooks for complexity reduction.

SPLIT VQ SYSTEM WITH HYBRID CODEBOOK STRUCTURE

The basic block diagram of the proposed split VQ system is depicted in Fig. 1. Like the other split vector quantization scheme, the input vector is partitioned into several small dimension subvectors for vector quantization. The major difference between our scheme and the others [1][3] is that the codewords are found inside the lattice analysis loop, based upon the residual minimization concept of the linear prediction. Therefore, the quantization errors in the first stage are forwarded to the second stage. Similarly, quantization errors in the subsequent stages are coupled to adjacent stage during the encoding process. Errors accumulated from previous stages would thus be compensated by the next stage codevector. Since the quantized k-parameters are used for the lattice analysis, traditional method such as Levinson Durbin algorithm for determined the optimal k-parameters are no longer valid in this case. A lattice analysis using quantized k-parameter [4] is adopted for the analysis.

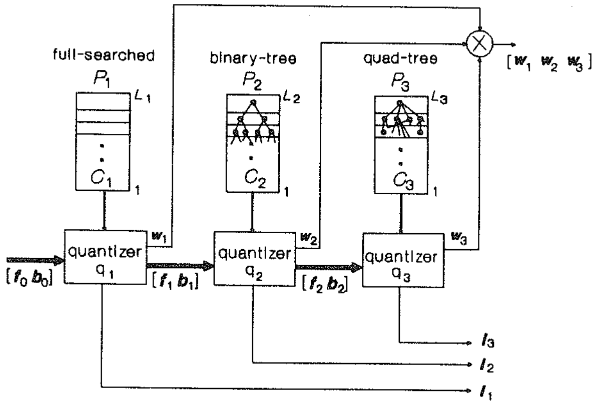

Figure 1 Block diagram of the k-parameters split vector quantization scheme using hybrid codebooks, where $P_1=3$, $P_2=3$ and $P_3=4$.

From [4], the lattice analysis recursive formulae using quantized k-parameters were derived as:

$$\begin{bmatrix} \alpha_t^{m+1} \\ \beta_t^{m+1} \end{bmatrix} = \begin{bmatrix} 1 & -\hat{k}_{m+1} \\ -\hat{k}_{m+1} & 1 \end{bmatrix} \cdot \begin{bmatrix} \alpha_t^m - \hat{k}_{m+1}\beta_{1-t}^m \\ \beta_{1+t}^m - \hat{k}_{m+1}\alpha_t^m \end{bmatrix} \qquad 0 \le |t| \le P-m \tag{1}$$

where $\hat{k}_{m+1}$ is the m^{th} stage quantization entity of the k-parameter and $\alpha_t^m = \sum_n e_f^m(n) e_f^m(n-t)$, $\beta_t^m = \sum_n e_f^m(n) e_b^m(n-t)$ with $e_f^m(n)$ and $e_b^m(n)$ being the m^{th} stage forward and backward prediction errors respectively. The structure of the split vector quantizer, shown in Fig. 1, can be derived from eq(1), the simplified computational structure of the 2 stages VQ is shown in Fig. 2. Since the main objective of eq (1) is to minimize the prediction residual using the quantized coefficients in the analysis. The prediction residual for a particular quantized k-parameter can be expressed as

$$\alpha_0^{m+1}(j) = \alpha_0^m(1+\hat{k}_{m+1}^2(j)) - 2\hat{k}_{m+1}(j)\beta_1^m \tag{2}$$

Therefore, for each stage, all codebook entries $w(j)=[k_i \,..\, k_{i+p}]$ are substituted into eq. (1) during analysis, the criterion for selecting the best codeword index is defined as:

$$I_{opt} = arg\{\ \min\ \alpha_0^{m+1}(j)\ \} \qquad w(j)\in C_{m+1} \tag{3}$$

It is clear from Fig. 1 that the input vector for the 1st stage is the autocorrelation of speech signal, whereas the input vectors for the other stages are α's and β's respectively. After determining codeword index for the 1st stage, the correlation vectors are updated by eq (1) using the best codeword. These two input correlation vectors are thus used as the input vectors for the 2nd stage analysis. Similarly, a binary tree-searched and quad tree-searched lattice analysis using eq(1) (2) and (3) are then performed sequentially, one stage at a time, in order to find the best codewords for quantization.

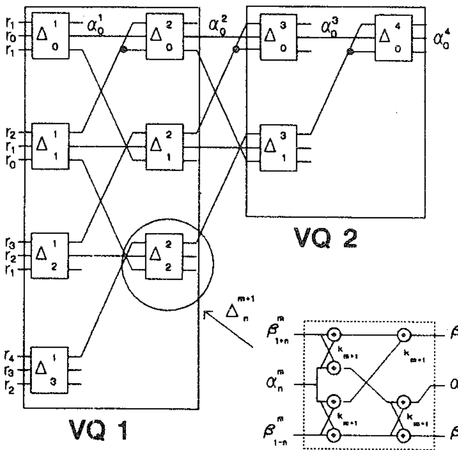

Figure 2 2 stages split vector quantizer using k-parameter, where $P=4$ and $P_1=P_2=2$

CODEBOOKS GENERATION ALGORITHM

A VQ codebook generation program for the new VQ structure was developed. The codebooks are trained stage by stage based on the LBG algorithm [5][6] with likelihood ratio as distortion measure

[7]. Note that, apart from having a lower LPC order, the codebook training procedure for the first stage VQ is the same as ordinary LPC-VQ design algorithm [5]. In the first stage, the training vectors are the autocorrelation sequences of speech signals. The training procedures for the subsequent stages are also similar to the first stage except that the input training vectors now consist of two sets of complementary vectors: the autocorrelation sequences of the forward prediction errors f_i and cross-correlation of the forward and backward prediction errors b_i. Because 2 correlation training vectors are used during the codebook training, it therefore should have two centroid corresponding to these input training sequences. The centroids at each stage is calculated from the input training vectors independently. They are firstly computed by averaging over all the gain-normalized correlation sequences (i,e. $\bar{\alpha}$ and $\bar{\beta}$). The lattice analysis is then performed to find the corresponding k-parameters codeword. The procedures of summation of the input training sequences and lattice analysis are iterated until the distortion ratio drops to a predefined threshold value. After the completion of one stage codebook generation, the entire input training sequence is then fed to the lattice quantizer for generating the input training sequences for next stage use.

SIMULATION RESULTS

A large speech database was set up for the codebook training and performance comparison among different quantization schemes. Speech data were recorded from 4 males and 2 females uttered in Cantonese sentences. The total number of frames for codebook training are about 12000 vectors , whereas 1200 speech frames are used for testing. These speech data are firstly band-limited to 3.4k Hz and digitized at 8k Hz sampling frequency. A 10^{th} order LPC analysis using autocorrelation method is then performed with 20ms frame size and 32ms Hamming window .

An average Likelihood Ratio (d_{LR}) is employed as the objective measurement to evaluate the quantization distortion for different quantizers at various bit rate. The average Likelihood ratio is defined as:

$$d_{LR} = \frac{\hat{\alpha}}{\alpha} - 1 \tag{4}$$

where α and $\hat{\alpha}$ are the prediction error energies associated with the unquantized and quantized LPC coefficients.

Firstly, we study the performance of proposed split VQ system with full-searched codebooks (FS-CB) for all stages and different codebook structure for different stage (i.e hybrid codebook, H-CB) based upon the Likelihood ratio distortion as well as the computational complexity. Table I summarizes the results of quantization distortion for these schemes. The complexity for the 2 schemes are given in table II. From Table I and Table II, we learn that the quantization performance is degraded by 1 bit/frame by using the hybrid codebooks. The complexity of the quantizer is, on the other hand, reduced about 3.6 times.

bit rate	FS-CB (LR)	H-CB (LR)
18	0.1348	0.1545
20	0.1076	0.1148
22	0.0826	0.0901
24	0.0604	0.0720

Table I Likelihood ratio(LR) of the split VQ (k-parameters) using the full-searched codebooks (FS-CB) and hybrid codebooks (H-CB).

	Stage 1		Stage 2		Stage 3		total no of operations
	Mul	Add	Mul	Add	Mul	Add	
FS-CB	10752	10752	10752	10752	17920	17920	78848/codeword
H-CB	10752	10752	336	336	1120	1120	24416/codeword

Table II Complexity comparison of the Full-searched and hybrid codebooks.

Next, the proposed scheme are compared with others split VQs algorithms, which are used the Euclidean Distance of the LSP [1] or LAR [3] parameters as distortion measure respectively. In order to have a fair comparison, all split VQ systems are partitioned into 3 stages with vector dimension of [3 3 4]. The stage with parameters having the highest spectral sensitivity is configured as full-searched codebook, and the rest are configured as binary or quad tree-searched codebooks. Fig. 2 illustrates plots of likelihood ratio vs bit rate for the three split VQ schemes. It is clear from Fig. 2 that the proposed split vector quantizer outperforms than other scheme having the same configuration.

CONCLUSIONS

A split VQ of the k-parameters using hybrid codebooks for different stages was proposed. By using different tree-searched codebooks in lower spectral sensitivity coefficients, 3.6 times complexity reduction has been obtained with only 1 bit/frame performance degradation trade-off. Since the codebooks are generated inside the lattice analysis loop, the new scheme is consistent with the residual minimization concept of traditional LPC analysis. The lattice filter is thus preserved after quantization. In addition, performance comparison with other split vector quantization algorithms has also been done. Simulation results showed that the new quantization algorithm outperformed than other schemes at all data rate.

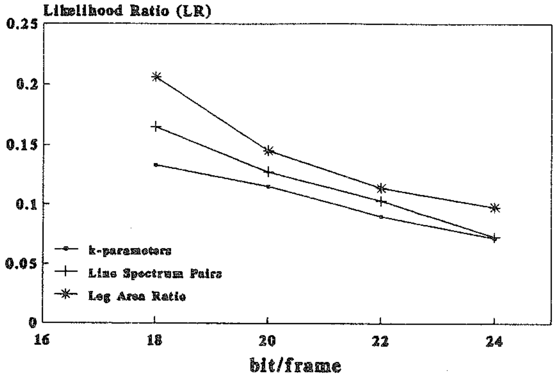

Figure 3 Quantization distortion vs bit rate for varios split vector quantization with hybrid codebooks structure

REFERENCES

[1] Paliwal, K. K. and Atal, B. S., "Efficient vector quantization of LPC parameters at 24 bits/frame," Proc. ICASSP-91, pp. 661-664, 1991.

[2] Chan, C. F. and Law, K. W., "Split-dimension piece-wise vector quantization of PARCOR coefficients", Proc. EUSIPCO-92, pp. 455-458, Belgium, Aug., 1992.

[3] Makhoul, J., Roucos, S. and Gish, H., " Vector quantization in speech coding," Proceeding of IEEE, Vol. 73, pp. 1551-1558, Nov., 1985.

[4] Chan, C. F. and Law, K. W., " In-loop quantization of reflection coefficients," Electronic Letter, Vol 27, pp. 2015-2016, Oct., 1991.

[5] Linde, Y., Buzo, A. and Gray, R. M., " An algorithm for vector quantization,", IEEE Tranc. on Comm., Vol. COM-28, Jan., 1980.

[6] A. Buzo, A. H. Gray, Jr., R. M. Gray and J. D. Markel, "Speech Coding Based Upon Vector Quantization," IEEE Trans. on Acoust., Speech, Signal Processing, Vol. ASSP-28, pp. 562-574, Oct., 1980.

[7] A. H. Gray, Jr. and J. D. Markel, "Distance Measures for Speech Processing," IEEE Trans. on Acoust., Speech, Signal Processing, Vol. ASSP-24, pp. 380-391, Oct., 1976.