EFFICIENT QUANTIZATION OF LSF PARAMETERS USING CLASSIFIED SVQ COMBINED WITH CONDITIONAL SPLITTING

Dong-il Chang, Young-kwon Cho and Souguil Ann

Department of Electronics Engineering, Seoul National University
San 56-1, Shinlim-dong, Kwanak-gu
Seoul 151-742, Korea

ABSTRACT

In this paper, we propose a classified SVQ of line spectral frequency (LSF) parameters combined with conditional splitting. The proposed algorithm adopts an independent conditional splitting scheme instead of the conventional fixed splitting scheme for each class. Considering the perceptual and spectral sensitivity characteristics of LSF's, we define an LSF perceptual importance index (LPII) to represent the relative perceptual importance of each one. Experimental results have shown that the proposed algorithm, conditional split VQ (CONSVQ), can achieve reduction of 37.5 % in searching complexity while maintaining the performance of quantization. From these results, we have found that the performance of VQ can be enhanced by considering and using the difference in relative importance of LSF's.

I. INTRODUCTION

Vector quantization (VQ) of LSF parameters may provide better performance than scalar quantization at any given rate. In spite of the large computational and storage complexity, the superior performance makes VQ of LSF's attractive for low rate coding of spectral information. However, an unstructured full-search VQ of LSF's requires impracticably high complexity in addition to the complexity in design procedure. Therefore, several structured VQ methods of LSF's have been proposed to overcome these problems [1], [2].

Among these methods, the split VQ (SVQ) can provide nearly transparent coding of LSF at a rate between 24-26 bits/frame [1]. In the conventional SVQ, several lower indexed LSF's are grouped and quantized with higher precision than the remaining upper LSF's. This fixed splitting scheme is based on the fact that, generally, lower indexed LSF's are perceptually more important than higher indexed ones. In terms of spectral sensitivity and perceptual quality, however, we should reduce the errors near spectral formants, because they may cause more severe degradation of performance than the errors in low frequency components. This means that we can improve the performance of SVQ by using a variable splitting scheme dependent on the characteristics of a given LPC spectrum.

In this paper, we propose a conditional split VQ (CONSVQ) of LSF's that adopts a classified SVQ structure combined with conditional splitting instead of fixed splitting. An LSF perceptual importance index (LPII) is also defined to denote the relative importance of an LSF. In the CONSVQ, LSF's are ordered and split according to the values of LPII. By introducing a classified SVQ structure, we can restrict efficiently the number of different splitting conditions.

This paper is organized as follows. We summarize the definition and properties of LSF's in Section II. In this section, we also discuss the relationship between LPC spectrum and LSF's and derive a localized approximation of LPC spectrum. In Section III, LPII is defined and a simple example of LPII is provided. In Section IV, the proposed CONSVQ algorithm is described. In Section V, we present the results of performance comparison and conclusions are followed in Section VI.

II. LSF: DEFINITION AND PROPERTIES.

The p-th order LPC prediction filter is given by

$$A(z) = 1 + \sum_{k=1}^{p} a_k z^{-k}, \tag{1}$$

where a_k's are LPC coefficients. Two LSF polynomials, $P(z)$ and $Q(z)$, are defined from $A(z)$ as follows:

$$\begin{aligned} P(z) &= A(z) + z^{-(p+1)} A(z^{-1}), \\ Q(z) &= A(z) - z^{-(p+1)} A(z^{-1}). \end{aligned} \tag{2}$$

LSF parameters, $\{\omega_i : i = 1, \ldots, p\}$, are the roots of $P(e^{-j\omega})$ and $Q(e^{-j\omega})$ in $0 < \omega < \pi$. Odd LSF's are the roots of $P(e^{-j\omega})$ and even LSF's are the roots of $Q(e^{-j\omega})$. If the roots of $P(e^{-j\omega})$ and $Q(e^{-j\omega})$ are interlaced and ordered in ascending manner, then the resulting LPC synthesis filter, $1/A(z)$, is always stable [3]. Therefore,

0-7803-2431-5/95 $4.00

these two properties must be also satisfied to ensure the stability of synthesis filter after LSF's are quantized.

The LPC power spectrum of $S(z)=1/A(z)$ can be decomposed by

$$\left|S(e^{-j\omega})\right|^2 = 4\Big/\left(\left|P(e^{-j\omega})\right|^2+\left|Q(e^{-j\omega})\right|^2\right). \quad (3)$$

From (3), we can find that the magnitude of power spectrum, $\left|S(e^{-j\omega})\right|^2$, becomes large when two or more LSF parameters are closely located together. This means that the LPC power spectrum are determined by the distribution of LSF's [4].

To show this relationship more clearly, we derive a localized approximation of LPC power spectrum as a function of difference between two adjacent LSF's. First of all, assume that the filter order p is even. Then two LSF polynomials $P(e^{-j\omega})$ and $Q(e^{-j\omega})$ can be decomposed, respectively, as follows:

$$P(e^{-j\omega}) = \left(e^{-j\omega}+1\right)\prod_{k=1,(odd)}^{p-1}\left(e^{-j\omega}-e^{-j\omega_k}\right)\left(e^{-j\omega}-e^{j\omega_k}\right),,$$

$$Q(e^{-j\omega}) = \left(e^{-j\omega}-1\right)\prod_{k=2,(even)}^{p}\left(e^{-j\omega}-e^{-j\omega_k}\right)\left(e^{-j\omega}-e^{j\omega_k}\right). \quad (4)$$

From (4), the power spectrums, $\left|P(e^{-j\omega})\right|^2$ and $\left|Q(e^{-j\omega})\right|^2$, can be also decomposed as, respectively,

$$\left|P(e^{-j\omega})\right|^2 = (2+2\cos\omega)\prod_{k=1,(odd)}^{p-1}\{2-2\cos(\omega-\omega_k)\}\cdot\{2-2\cos(\omega+\omega_k)\},$$

$$\left|Q(e^{-j\omega})\right|^2 = (2-2\cos\omega)\prod_{k=2,(even)}^{p}\{2-2\cos(\omega-\omega_k)\}\cdot\{2-2\cos(\omega+\omega_k)\}. \quad (5)$$

Let's consider the case in which only two LSF's, ω_i and ω_{i+1}, are closely located and the others are distributed apart from them. In this case, we may regard the other factors in (5) as nearly invariant. Therefore, we can derive an approximation of $\left|S(e^{-j\omega})\right|^2$, $A(e^{-j\omega})$, in this region as

$$A(e^{-j\omega}) = \left[K_1\{1-\cos(\omega-\omega_i)\}+K_2\{1-\cos(\omega-\omega_{i+1})\}\right]^{-1}, \quad (6)$$

where $\omega_i \le \omega \le \omega_{i+1}$. Assume $K_1 \approx K_2 = 2\cdot K^{-1}$, then

$$A(e^{-j\omega}) = K\left\{1-\frac{1}{2}\left[\cos(\omega-\omega_i)+\cos(\omega-\omega_{i+1})\right]\right\}^{-1},$$

$$= K\left\{1-\cos\left(\frac{\omega_{i+1}-\omega_i}{2}\right)\cos\left(\omega-\frac{\omega_{i+1}+\omega_i}{2}\right)\right\}^{-1},$$

$$= K\{1-D\cos(\omega-M)\}^{-1}, \quad \omega_i \le \omega \le \omega_{i+1}. \quad (7)$$

where $D=\cos\{(\omega_{i+1}-\omega_i)/2\}$, $M=(\omega_{i+1}+\omega_i)/2$. The localized approximation given by (7) has a peak at M with a magnitude of $K[1-D]^{-1}$. As D approaches to unity, that is, the difference between two LSF's approaches to zero, the magnitude of the peak is increased inversely. The magnitudes and frequency of the peak may have different values from those of approximation, because two constants, K_1 and K_2, are different in practice. Nevertheless, the local structure of LPC spectrum can be efficiently approximated by (7) when two LSF's are closely located.

In the cases that two LSF's are separated from each other, the localized approximation in (7) are not valid. This invalidity is resulted from the negligence of the effects of the other terms which cannot be neglected in these cases. However, the approximation has relatively small magnitude of peak and large bandwidth in these cases. Accordingly, it can roughly represents the envelope of LPC spectrum combined with the approximations of the other regions.

So far, we have derived the localized approximation of LPC spectrum as a function of LSF difference. Though this approximation is not valid in general, it can efficiently represent the relationship between local LPC spectrum and local distribution of LSF's. Since the purpose of the approximation is not to estimate the LPC spectrum itself , but to estimate the relationship, it still has special meanings to our application where LSF's are divided into subsets according to the relative importance.

III. LSF PERCEPTUAL IMPORTANCE INDEX

For more efficient quantization, the spitting scheme should be determined according to the relative importance of LSF. In terms of perception and spectral sensitivity, spectral formants are much more important than the spectral valleys. In additions, the relatively higher importance of low frequency components should be also considered for denoting the relative importance more accurately. Based on these, the LPII is defined by the following product form for representing the relative importance of each LSF:

$$L(\omega_i) = f(\omega_i)g(\omega_i), \quad i=1,2,\ldots,p. \quad (8)$$

In (8), $f(\omega_i)$ represents the relatively higher perceptual importance of lower frequency components, while $g(\omega_i)$ represents the higher perceptual importance of spectral formants.

The first component, $f(\omega_i)$, has nothing to do with the LPC spectrum at a given frame, whereas $g(\omega_i)$ is so closely related to it. This means that $f(\omega_i)$ may be regarded as a time-invariant function, whereas $g(\omega_i)$ should be a time varying function dependent on the LPC spectrum at a given frame. We may represent $g(\omega_i)$ by a function of differences between three adjacent LSF's as

$$g(\omega_i) = \sum_{k=1,\, k\neq i}^{p} d(|\omega_i - \omega_k|) \approx d(|\omega_i - \omega_{i-1}|) + d(|\omega_{i+1} - \omega_i|). \quad (9)$$

The approximation in (9) is based on the property that an LSF has restricted effects on the region of LPC spectrum close to it [5]. The approximation given by (7) may be used as $d(\cdot)$ in (9). In this paper, however, we choose a simple step function for simplicity which is given by,

$$d(\Delta\omega) = s(\Delta\omega) = \begin{cases} C, & \text{if } \Delta\omega \le T_1 \\ 0, & \text{if } T_1 < \Delta\omega \le T_2 \\ -C, & \text{if } T_2 < \Delta\omega \end{cases}, \quad (10)$$

where $\Delta\omega = \omega_{i+1} - \omega_i$, $C > 0$ and $T_1 < T_2$. In the conventional SVQ, the second term, $g(\omega_i)$, is neglected and a fixed splitting scheme based on $f(\omega_i)$ is used. On the contrary, in this paper, we neglect the effects of $f(\omega_i)$ in computing LPII and only consider the $g(\omega_i)$ to show the effects on the performance more directly. Therefore, we choose an LPII as the following form:

$$L(\omega_i) = s(|\omega_i - \omega_{i-1}|) + s(|\omega_i - \omega_{i+1}|). \quad (11)$$

In Fig. 1, we illustrate an example of LPII which is compared with the corresponding LPC spectrum. In this case, two thresholds, T_1 and T_2, are fixed as $\pi/22$ and $\pi/11$, respectively and the constant C is chosen by 10. The LPII plot in Fig. 1 is constructed by giving the same value as the LPII value of a given LSF within the region between its previous LSF. From Fig. 1, we can see that the LPII given by (11) can represent efficiently the relatively higher importance of LSF's corresponding to LPC formants and distinguish exactly them from the others.

IV. CONDITIONAL SPLIT VQ (CONSVQ) OF LSF PARAMETERS

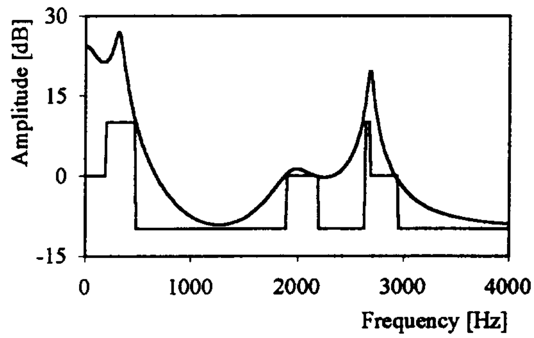

Fig. 1. A comparison of LPII with its corresponding LPC power spectrum.

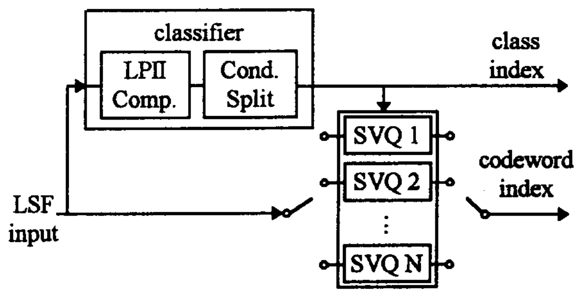

Fig. 2. The overall structure of the proposed CONSVQ of LSF parameters.

Using the LPII given by (11), we can group several LSF components with the highest relative importance more exactly for a given LSF vector. Because of the improvements in splitting scheme, we may enhance the performance of the SVQ of LSF's by reducing the errors at the selected important LSF's instead of increasing the errors at the remaining LSF's.

In this paper, we propose a CONSVQ of LSF's which is based on the classified SVQ combined with conditional splitting. In classified VQ, an input vector is classified into a class and quantized by a specific way for the class [6]. In CONSVQ, therefore, an input LSF vector is classified into a class among L classes. The classified LSF vector is split into subvectors and they are quantized independently according to the predetermined way. The overall structure of CONSVQ is illustrated in Fig. 2.

In our experiments, we design the classifier in classified SVQ as a VQ of LPII. For each class, we determine the condition of splitting and quantization scheme by considering the codeword of classifier for the class, which corresponds to the average LPII for the class. The dimension of each subvector and bit allocation should be determined carefully, so that the performance of CONSVQ is maximized while minimizing the storage and computational complexity.

V. EXPERIMENTS AND RESULTS

The speech database used in our experiments consists of about 27 minutes of speech recorded from 8 different FM radio station. First 25 minutes of speech is used for training VQ and the remaining about 2 minutes of speech is used for evaluating the performance. Input speech is sampled at 8 kHz and bandpass filtered from 200 to 3400 Hz. The 10-th order LPC analysis using the autocorrelation method is performed every 30 ms with a Hamming window. Bandwidth expansion of 15 Hz is followed and pre-emphasis is not used.

To implement CONSVQ, we choose the number of classes L and the number of subvectors N by 3 and 2, respectively. We divide the input database into two classes at first, and only one class is divided into two subclasses, because the sub-dividision of the other class have caused a highly sensitive VQ to the given data. In the undivided class, an input vector is divided into two subvectors with equal dimensions. In this case, a single bit is used for denoting the class information, and the remaining bits are divided unequally and allocated to subvectors, respectively. In this paper, we allocate only one more bit to the subvector of important LSF's. In the two sub-classes, an input vector is divided into two subvectors with unequal dimension of 4 and 6, respectively, and we allocate the same bits to each of them.

This resulting CONSVQ has the 37.5 % reduction in searching complexity compared with the conventional SVQ, while the required storage is 1.75 times more. We may reduce the required storage by varying the number of classes, bit allocation or dimensions of subvectors, but it may also cause the degradation in performance. This means that there is the trade-off between storage and performance in CONSVQ.

The evaluation of the performance is performed by using the average log spectral distortion,

$$D=\frac{1}{N_1-N_0}\sqrt{\sum_{k=N_0}^{N_1-1}\left(10\log_{10}\left|A_q(k)\right|^2-10\log_{10}\left|A(k)\right|^2\right)^2}, \quad (12)$$

where $A_q(k)$ is FFT of LPC filter constructed by the quantized LSF's and $A(k)$ is FFT of original LPC filter. The two constant N_0 and N_1 represent the FFT indexes corresponding to the cut-off frequencies of the input band-pass filter. Results of performance comparison between SVQ and CONSVQ are summarized in Table 1. We use the Euclidean mean square error (MSE) without weighting in searching for the best codeword.

In Table 1, we can find that proposed CONSVQ has slightly larger average distortion than the conventional SVQ at any given rate whereas the percentage of outlier of above 4 dB is significantly reduced.

Table 1. Average spectral distortion from the proposed CONSVQ and the conventional SVQ.

Quant. Method	Bits	SD (dB)	Outliers (%)	
			2-4 dB	> 4 dB
CONSVQ	22	1.49	17.2	0.12
SVQ	22	1.45	12.7	0.21
CONSVQ	24	1.34	11.2	0.00
SVQ	24	1.31	8.19	0.08
CONSVQ	26	1.20	7.80	0.00
SVQ	26	1.18	5.14	0.04

VI. CONCLUSION

In this paper, we have proposed the CONSVQ scheme of LSF parameters which are a classified SVQ scheme combined with conditional splitting. To overcome the suboptimality of the conventional fixed splitting scheme, we have introduced a conditional splitting scheme based on the relative importance of LSF. The LPII has been defined to represent the perceptual importance of LSF's. Experimental results have shown that an example of the proposed CONSVQ can achieve reduction of 37.5 % in the searching complexity while maintaining the performance of quantization.

REFERENCES

[1] K.K. Paliwal and B.S. Atal, "Efficient Vector Quantization of LPC Parameters at 24 Bits/Frame," *IEEE Trans. on Speech and Audio Processing*, vol. 1, pp. 3-14, 1993.

[2] W.P. LeBlanc, B. Bhattacharya, S.A. Mahmoud and V. Cuperman, "Efficient Search and Design Procedures for Robust Multi-Stage VQ of LPC Parameters for 4 kb/s Speech Coding," *IEEE Trans. on Speech and Audio Processing*, vol. 1, pp. 373-385, 1993.

[3] F.K. Soong and B.H. Juang, "Line Spectum Pair (LSP) and Speech Data Compression," *Proc. ICASSP-84*, pp. 10.1-10.4, 1984.

[4] S. Furui, *Digital Speech Processing, Synthesis, and Recognition*, Marcel Dekker, 1991.

[5] F.K. Soong and B.H. Juang, "Optimal Quantization of LSF Parameters," *IEEE Trans. on Speech and Audio Processing*, vol. 1, pp. 15-24, 1993.

[6] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Press, 1992.