Split Vector Quantization of LSF Parameters with Minimum of dLSF Constraint

Sung-Joo Kim and Yung-Hwan Oh, *Member, IEEE*

***Abstract*—In this letter, the authors present two improved split vector quantization (SVQ) methods for line spectral frequency (LSF) parameters. By using these methods jointly, the codewords and quantization results conserve a given minimum difference LSF (dLSF), although they are trained and quantized with a weighted distance measure. Experimental results show that the proposed methods are more effective than conventional SVQ methods, because the total training error and number of outliers due to quantization are all reduced.**

***Index Terms*— Line spectral frequency, split vector quantization.**

I. Introduction

In speech coding, the spectral envelope of an analysis frame is often represented by line spectral frequencies (LSF's). LSF's are estimated from given linear predictive coefficients (LPC's) and can be transformed back to corresponding LPC's without loss of information. An important property of LSF's $\{\omega_i\}$ is that they are ordered in $(0, \pi)$ as

$$0 < \omega_1 < \omega_2 < \cdots < \omega_p < \pi. \tag{1}$$

Also (1) means that difference LSF's (dLSF's) $\{d_i = \omega_i - \omega_{i-1}\}$ with $d_1 = \omega_1$ are always larger than zero. This ordering property is a necessary and sufficient condition for the stability of the corresponding LPC synthesis filter.

Spectral distortion (SD) is defined by the root mean square difference between the original log-power spectrum and the quantized log-power spectrum [1]. To measure the performance of LSF quantization, the SD measure is commonly used, but during quantization, a weighted distance substitutes for SD because it is more computationally tractable [3]–[5]. However, none of the previously suggested weighted distance measures can deal with the dependency of LSF's accurately, since all the weights are evaluated from the original LSF's and do not consider the quantized LSF's. In Fig. 1, the squared SD caused by quantizing only one LSF, ω_k as $\hat{\omega}_k$, and the corresponding weighted distance $D = \alpha(\omega_k - \hat{\omega}_k)^2$ are plotted with varying $\hat{\omega}_k$ from ω_{k-1} to ω_{k+1}.[1] As shown in Fig. 1, the SD curve is not symmetric, but steeper on the smaller dLSF side. However, D is symmetric with respect to $\hat{\omega}_k = \omega_k$ and gives a reasonable value even when $\hat{\omega}_k$ goes out of the valid range $(\omega_{k-1}, \omega_{k+1})$. This mismatch causes some trained codewords and quantization results using a weighted distance to have too small or even negative dLSF values.

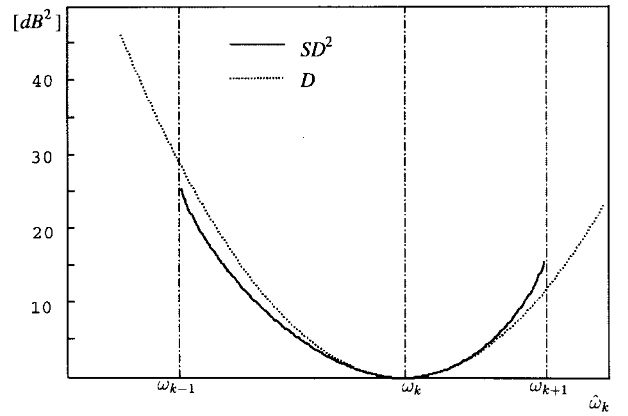

Fig. 1. Example of the mismatch between SD and a weighted distance.

To compensate for this weighted distance measure deficiency, we propose annexing a minimum dLSF constraint to the error criterion for quantizing LSF's. Thus, we develop a split vector quantization (SVQ) method, which gives a locally optimal quantizer [7] with respect to the given weighted distance measure and a minimum dLSF constraint. We also experimentally determine the proper minimum dLSF value.

II. SVQ of LSF's Conserving a Given Minumum dLSF

A. Training Phase of SVQ

Let the original and the quantized LSF vectors be $\boldsymbol{\omega} = [\omega_1, \cdots, \omega_p]^T$ and $\hat{\boldsymbol{\omega}} = [\hat{\omega}_1, \cdots, \hat{\omega}_p]^T$, respectively. The weighted distance measure is defined as

$$D(\boldsymbol{\omega}, \hat{\boldsymbol{\omega}}) = \sum_{i=1}^{p} \alpha_i(\omega_i - \hat{\omega}_i)^2. \tag{2}$$

In this paper, we use a modified inverse harmonic mean [4] as

$$\alpha_i = c_i\left(\frac{1}{\omega_{i+1} - \omega_i} + \frac{1}{\omega_i - \omega_{i-1}}\right) \tag{3}$$

assuming $\omega_0 = 0$, $\omega_{p+1} = \pi$ and c_i are constant weights. We use $c_i = 1.0$ except $c_9 = 0.8$, $c_{10} = 0.4$ [1]. In addition to the weighted distance measure, we restrict the minimum dLSF δ of the quantized LSF vector as

$$\hat{d}_i = (\hat{\omega}_i - \hat{\omega}_{i-1}) \geq \delta > 0, \qquad i = 1, \cdots, p. \tag{4}$$

Manuscript received March 24, 1999. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. K. K. Paliwal.

The authors are with the Department of Computer Science, Korea Advanced Institute of Science and Technology, Taejon 305-701, Korea (e-mail: sung@bulsai.kaist.ac.kr).

Publisher Item Identifier S 1070-9908(99)06638-9.

[1] The original ten LSF's were estimated from a vowel /a/ and ω_2 was changed. α was adjusted to match the squared SD curve near the original ω_2.

Now the total quantization error E for a certain cell $\{\boldsymbol{\omega}(n);\ n = 1, \cdots, N\}$ with quantization result $\hat{\boldsymbol{\omega}}$ can be defined by

$$E = \sum_{n=1}^{N} D(\boldsymbol{\omega}(n), \hat{\boldsymbol{\omega}}) = \sum_{n}\sum_{i} \alpha_i(n)(\omega_i(n) - \hat{\omega}_i)^2. \quad (5)$$

If $\hat{\omega}_i$ minimizes the error E, the following equation is always true for all i:

$$\frac{\partial E}{\partial \hat{\omega}_i} = -2\sum_{n} \alpha_i(n)(\omega_i(n) - \hat{\omega}_i) = 0. \quad (6)$$

Consequently, $\hat{\omega}_i$ should be set as

$$\hat{\omega}_i = \frac{\sum_n \alpha_i(n)\omega_i(n)}{\sum_n \alpha_i(n)}, \qquad i = 1, \cdots, p. \quad (7)$$

However, the $\hat{\boldsymbol{\omega}}$ estimated by (7) does not guarantee the minimum dLSF δ. Therefore, the average centroid was used in previous research [1], [6], but this is not consistent with the given weighted distance measure. Furthermore, it does not guarantee the convergence of the error theoretically.

Hence, we propose a novel solution for $\hat{\boldsymbol{\omega}}$, when $\hat{\boldsymbol{\omega}}$ violates the minimum dLSF constraint. For instance, if $\hat{\omega}_k - \hat{\omega}_{k-1} < \delta$, then we replace $\hat{\omega}_{k-1}$ and $\hat{\omega}_k$ with $\tilde{\omega}_{k-1}$ and $\tilde{\omega}_k = \tilde{\omega}_{k-1} + \delta$, respectively. In this case, the increment of the error E is calculated as

$$\Delta = (\hat{\omega}_{k-1} - \tilde{\omega}_{k-1})^2 \sum_n \alpha_{k-1}(n) + (\hat{\omega}_k - \tilde{\omega}_k)^2 \sum_n \alpha_k(n). \quad (8)$$

To minimize (8), we should set $\tilde{\omega}_{k-1}$ as

$$\tilde{\omega}_{k-1} = \frac{\hat{\omega}_{k-1}\sum_n \alpha_{k-1}(n) + (\hat{\omega}_k - \delta)\sum_n \alpha_k(n)}{\sum_n \alpha_{k-1}(n) + \sum_n \alpha_k(n)}. \quad (9)$$

As a result, we can train a codebook that minimizes the total weighted error E while conserving the minimum dLSF δ by using (7) and (9). Furthermore, this modification takes neither intensive computation nor extensive memory. It only requires the summation of weights for each dimension of the LSF vectors, which is used as the denominator in (7).

B. Quantization Phase of SVQ

In the quantization phase, the SVQ method independently quantizes each subvector of the original LSF vector, so it may produce a nonpositive dLSF at a boundary between two subvectors although all the quantized subvectors are ordered. As a solution to this problem, we can cancel the codeword combinations producing a nonpositive dLSF from the search space during the quantization. However, the quantization error of this method is always larger than that of the best codeword combination in point of the weighted distance measure.

We, however, propose a novel solution that enforces a minimum dLSF constraint on the best codeword combination, thus reducing the quantization error. If the SVQ result of $\boldsymbol{\omega}$ violates the minimum dLSF constraint as $\hat{\omega}_k - \hat{\omega}_{k-1} < \delta$, we then replace $\hat{\omega}_{k-1}$ and $\hat{\omega}_k$ with

$$\tilde{\omega}_{k-1} = (\hat{\omega}_{k-1} + \hat{\omega}_k - \delta)/2 \quad (10)$$

and

$$\tilde{\omega}_k = \tilde{\omega}_{k-1} + \delta = (\hat{\omega}_{k-1} + \hat{\omega}_k + \delta)/2. \quad (11)$$

This modification increases the quantization error by

$$\begin{aligned}\Delta =& \alpha_{k-1}\left\{(\omega_{k-1} - \tilde{\omega}_{k-1})^2 - (\omega_{k-1} - \hat{\omega}_{k-1})^2\right\} \\ &+ \alpha_k\left\{(\omega_k - \tilde{\omega}_k)^2 - (\omega_k - \hat{\omega}_k)^2\right\} \\ =& -0.5(\delta + \hat{\omega}_{k-1} - \hat{\omega}_k)[\alpha_k\{2(\omega_k - \tilde{\omega}_k) + (\tilde{\omega}_k - \hat{\omega}_k)\} \\ &- \alpha_{k-1}\{2(\omega_{k-1} - \tilde{\omega}_{k-1}) + (\tilde{\omega}_{k-1} - \hat{\omega}_{k-1})\}]. \quad (12)\end{aligned}$$

In (12), Δ becomes negative when $\omega_k > \tilde{\omega}_k$ and $\omega_{k-1} < \tilde{\omega}_{k-1}$, so that the quantization error is reduced, and experiments show that ω_k and ω_{k-1} satisfy this condition in most cases. Consequently, we can further reduce the quantization error of the best codeword combination by replacing an LSF pair, which violates the minimum dLSF constraint, with (10) and (11).

III. Experiments and Results

So far, we described how to enforce the minimum dLSF constraint on codewords and the quantization results of SVQ. We also quantitatively described the effect of the enforcement on the quantization error. In this section, we describe experiments with the proposed methods and their results.

We used the TIMIT database as the speech corpus which was downsampled at 8 kHz after being lowpass filtered to 3.4 kHz. 3696 spoken sentences from 462 speakers were used for the training, and 1344 spoken sentences from another 168 speakers, for the testing. A tenth-order LPC analysis was performed based on the autocorrelation method with a 30 ms Hamming window at a rate of 50 Hz. The LPC's were first 15-Hz bandwidth expanded and then converted to LSF's. As a result, we collected 566 117 LSF vectors as a training set and 205 804 as a test set.

First, to compare the proposed codeword updating method using (7) and (9) with others, we trained several codebooks. We split each LSF vector in the training set into three-, three-, and four-dimensional subvectors, $\boldsymbol{\omega}_1$, $\boldsymbol{\omega}_2$, and $\boldsymbol{\omega}_3$, and separately trained three 9-b codebooks for each subvector. We complied with the LBG algorithm [7] of an initial guess by splitting and set the distortion threshold ϵ to 0.0001 for the sake of fast convergence. When the reduction ratio of the total distortion error was equal or less than the distortion threshold ϵ, we stopped training and set the centroids as the codewords for the level at that time.

Table I shows the average weighted distance error in training each subvector codebook. Here, CB1 was trained with Euclidean distance, CB2 with the weighted distance and the average centroid as described in [1], [6], and CB3-δ with the weighted distance and the proposed centroid conserving the minimum dLSF δ. We used the same weights defined as (3) for all the weighted distances mentioned here and below. As shown in Table I, all of codebook CB3 with $\delta \leq 0.06$

TABLE I
TRAINING RESULTS OF EACH CODEBOOK FOR SVQ

type of codebook	average weighted distance error (enforcement occurred)		
	codebook for $\boldsymbol{\omega}_1$	codebook for $\boldsymbol{\omega}_2$	codebook for $\boldsymbol{\omega}_3$
CB1	0.005515	0.011327	0.018948
CB2	0.005392	0.010845	0.017030
CB3-0.03	0.005104 (0)	0.010520 (0)	0.016285 (0)
CB3-0.04	0.005108 (117)	0.010520 (0)	0.016285 (0)
CB3-0.05	0.005111 (3023)	0.010520 (0)	0.016285 (0)
CB3-0.06	0.005223 (11702)	0.010520 (0)	0.016282 (32)
CB3-0.07	0.005635 (17797)	0.010518 (21)	0.016287 (125)

TABLE II
TEST RESULTS OF EACH CODEBOOK FOR SVQ

type of codebook	SD [dB]	outliers [% (times)]		negative / modified
		2 - 4 dB	>4 dB	
CB1	1.121	2.47 (5095)	0.003 (7)	.
CB2	1.083	1.64 (3373)	0.002 (5)	.
CB3-0.03	1.082	1.22 (2502)	0.001 (2)	408/410
CB3-0.04	1.081	1.21 (2481)	0.001 (2)	795/826
CB3-0.05	1.081	1.19 (2455)	0.001 (2)	1400/1541
CB3-0.06	1.081	1.20 (2467)	0.001 (2)	2264/2652
CB3-0.07	1.085	1.22 (2507)	0.001 (2)	3331/4306
CB4	1.082	1.24 (2545)	0.001 (2)	.
CB5	1.082	1.58 (3258)	0.002 (5)	1521/1642

gives better performance than either CB1 or CB2, in that their average weighted distance errors are smaller than those of CB1 and CB2. For example, the average weighted distance errors of CB3-0.05 for the three subvectors are reduced 5.2, 3.0, and 4.4% more than those of CB2, respectively. It should be noted that only CB3 used a centroid consistent with the weighted distance measure, so that the average weighted distance error for CB3 was nonincreasing during the training phase, but not for CB1 nor for CB2. The enforcement as (9) occurred 3023 times during the training of the CB3-0.05 for $\boldsymbol{\omega}_1$, but did not occur with $\delta \leq 0.03$ for $\boldsymbol{\omega}_1$, $\delta \leq 0.06$ for $\boldsymbol{\omega}_2$, nor with $\delta \leq 0.05$ for $\boldsymbol{\omega}_3$.

Finally, we measured the SD of the quantized LSF vectors in the test set over the 0–3 kHz frequency band [1]. During the quantization, the codeword combinations producing a nonpositive dLSF were canceled for CB1 and CB2, but modified with (10) and (11) for CB3-δ. We additionally tested CB4 and CB5 to evaluate the isolated contributions of the two proposed methods to performance improvement. CB4 is identical with CB3-0.05, but the codeword combinations producing a nonpositive dLSF were canceled, as they were for CB2. In contrast, CB5 is identical with CB2, but the quantized LSF vectors were modified to guarantee a minimum dLSF of 0.05, as they were for CB3-0.05.

According to Table II, CB1 performs worst, as expected. This confirms that our weighted distance measure is effective. CB3-0.05 performs best and practically satisfies the condition for transparent coding, i.e., with an average SD of about 1 dB, less than 2% of outliers in the range 2–4 dB, and no outlier with a SD greater than 4 dB [1]. Especially, CB3 is effective in reducing outliers with the result that 27.3% of the outliers of CB2 can be reduced with $\delta = 0.05$. In addition, as shown in the last column of Table II, we checked the increment Δ in (12), which is caused by the modification of quantized LSF vectors. As a result, we found that, of the total 1541 modifications of quantized LSF vectors for CB3-0.05, 1400 modifications gave negative Δ. In other words, 90.9% of the quantized LSF vector modifications gave a smaller error than all possible codeword combinations of the CB3-0.05 SVQ codebook.

IV. CONCLUSION

Two improved SVQ methods for LSF parameters were presented. The proposed methods conserve a given minimum dLSF while training the codebook and quantizing the LSF's. The proposed centroid minimizes the total weighted distance error with minimum dLSF constraint, so that the codebook gives a smaller error than that of the conventional average centroid. During the SVQ of the LSF's, we demonstrate experimentally that modifying a pair of LSF's, which violates the minimum dLSF constraint, gives better quantization performance than does a canceling method. When we set the minimum dLSF to 0.05, the proposed methods give the best quantization performance. As a result, we can reduce 27 b/frame three-part SVQ outliers by 27.3%, using these methods jointly. The proposed methods do not adjust the weighted distance measure itself, but only annex a minimum dLSF constraint to the error criterion. Hence, these methods can be applied to any weighted distance measure as long as the weights are only dependent on the original LSF's.

REFERENCES

[1] K. K. Paliwal and B. S. Atal, "Efficient vector quantization of LPC parameters at 24 bits/frame," *IEEE Trans. Speech Audio Processing,* vol. 1, pp. 3–14, Jan. 1993.
[2] J. Pan and T. R. Fischer, "Vector quantization of speech line spectrum pair parameters and reflection coefficients," *IEEE Trans. Speech Audio Processing,* vol. 6, pp. 106–115, Mar. 1998.
[3] F. K. Soong and B.-H. Juang, "Optimal quantization of LSP parameters," *IEEE Trans. Speech Audio Processing,* vol. 1, pp. 15–24, Jan. 1993.
[4] R. Laroia, N. Phamdo, and N. Farvardin, "Robust and efficient quantization of speech LSP parameters using structured vector quantizers," in *Int. Conf. Acoustics, Speech, Signal Processing,* 1991, pp. 641–644.
[5] P. A. Laurent, "Expression of spectral distortion using line spectrum frequencies," *IEEE Trans. Speech Audio Processing,* vol. 5, pp. 481–484, Sept. 1997.
[6] R. P. Ramachandran, M. M. Sondhi, and N. Seshadri, "A two codebook format for robust quantization of line spectral frequencies," *IEEE Trans. Speech Audio Processing,* vol. 3, pp. 157–167, May 1995.
[7] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.,* vol. COMM-28, pp. 84–95, Jan. 1980.