AN EFFICIENT LSF QUANTIZATION USING DYNAMIC BIT ALLOCATION

Hosang Sung, Changyong Son, Kangeun Lee, and Eunmi Oh

Samsung Electronics Co., LTD.
San #14, Nongseo-Dong, Giheung-Gu, Yongin-City, Gyeonggi-Do, Korea

ABSTRACT

The current paper is concerned with an effective method to quantize a spectrum envelope of a speech signal without having an inter-frame prediction. In this paper, we proposed a method referred to as dynamic bit allocation-split vector quantization (DBA-SVQ). The main feature of this structure is that it makes use of the ordering property of line spectral frequencies (LSF) and exploits multiple codebooks, normalization and the DBA technique. As a result, we can limit the dynamic range of LSF sub-vectors and allocate different numbers of bits in accordance with the range sizes to maximize the overall efficiency of quantization. The performance is compared with delta line spectral pairs (LSP) VQ, which is used in EVRC-B, demonstrating reduction in spectral distortion (SD).

***Index Terms*—**LPC, quantization, speech coding

1. INTRODUCTION

Packet-based speech communication has become important and is in great demand. Quality of service (QoS) may not be guaranteed due to network congestion, maximum delay constraints, and buffer overflow occurred in heterogeneous network. Error robustness of a speech coder is desirable in packet-based network systems where packet loss occurs. Recently several frame erasure concealment (FEC) techniques were introduced, and error robustness of speech coder is enhancing gradually [1]. The current paper is concerned with an effective method to quantize the spectrum envelope without using inter-frame prediction. With this feature, this method can minimize the error propagation caused by the spectrum quantization and can be of help in maximizing the FEC performance. Additionally, it can enhance the coding efficiency of a spectrum quantizer for a signal region with low inter-frame correlation.

A typical speech coding system exploits linear prediction (LP) coefficients which can represent the spectrum envelope of speech signals. The LP coefficients are converted to LSP or LSF representation for efficient quantization and interpolation [2]. A moving average (MA) or autoregressive (AR) prediction method is frequently used to quantize LSF because this parameter has high inter-frame correlation [3]. These predictions can be an effective way to quantize the LSP in case of no-error condition, but tend to be fragile in the error-prone condition.

This paper is concerned with dynamic bit allocation-split vector quantization (DBA-SVQ) that is a non-predictive system described in section 2. The absence of the prediction is compensated by using a novel DBA technique as well as existing technologies: a linked split VQ (LSVQ) [4] and a normalization scheme [5]. In section 3, the performance of the DBA-SVQ is compared to delta LSP VQ method in EVRC-B.

2. DYNAMIC BIT ALLOCATION - SVQ

2.1. LSF Vector Split and LSVQ

An all-pole digital filter for speech synthesis, *H(z)*, is given by

$$H(z) = \frac{1}{1+\sum_{k=1}^{p} \alpha_k z^{-k}} \tag{1}$$

where p is LP filter order, which is generally 10 for narrowband speech, and α_k are filter coefficients (LP coefficients). Consequently, the LP coefficients are then converted to the LSP or LSF representation for quantization and interpolation purposes. The LSF is denoted by

$$\Omega = [\omega_0, \omega_1, \ldots, \omega_{p-1}] \tag{2}$$

By default, all LSF values can vary from 0 to π, and theoretically each element can exist at any location within the range. One of the important LSF properties is the natural ordering of its parameters [2]. This property can be represented by

$$0 \le \omega_0 < \omega_1 < \cdots < \omega_{p-1} \le \pi \tag{3}$$

This ordering property indicates that the filter is stable and the LSFs have a relatively high intra-frame correlation. The vector quantization (VQ) method efficiently quantizes the LSFs by using this correlation. However, the complexity and the memory size of the VQ are increased exponentially with respect to the size of dimension of the LSF vector. The split vector quantization (SVQ) and the multi-stage VQ (MSVQ) have been developed to reduce the complexity and memory size, even though the performances of these two VQ schemes are worse than that of the full dimensional VQ.

In order to increase the performance of the SVQ, an LSVQ scheme is introduced [4]. The basic concept of the LSVQ is a multi-codebook scheme using correlation among the adjacent sub-vectors. In the LSVQ, the input LSF vector is divided into three sub-vectors and the middle sub-vector is quantized first. The codebooks for lower and upper sub-vector quantization are selected according to the quantized values of the first and the last LSF ($\hat{\omega}_3$, and $\hat{\omega}_5$ in Fig. 1) of the middle sub-vector, respectively. Fig. 1 shows the basic structure of the LSVQ.

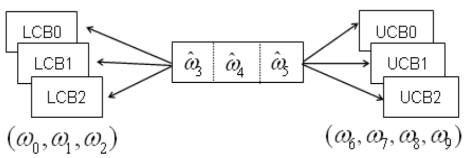

Fig. 1. The structure of LSVQ

2.2. Normalization and Weighting

When the values of the adjacent parameters of a given LSF parameter are known, the dynamic range of the parameter can be limited by the ordering property. In addition, if we normalize the distance by the adjacent values, then the dynamic range will be set from 0 to 1.

$$d_n = \frac{\omega_n - \omega_{n-1}}{\omega_{n+1} - \omega_{n-1}},$$
$$\omega_{n-1} < \omega_n < \omega_{n+1},\ and \quad 0 < d_n < 1 \tag{4}$$

A more intricate example of normalization is given in [6]. In the paper, the normalization serves the additional purpose of shaping the data into simple geometric regions in order to facilitate the design of lattice codebooks. With this idea, many design variations are possible [5]. In this paper, we merged this normalization idea with the LSVQ. The basic idea of the LSVQ is to divide the dynamic range and to design multi-codebooks for each divided range. On the other hand, the normalization idea is to limit and normalize the dynamic range. These two ideas commonly exploit the ordering property and pursue the same goal: minimizing the dynamic range for quantization. In order to realize the combination of these two ideas, we split the LSF vector with an anchor as shown in Fig. 2. According to [5], the performance of normalized SVQ is superior to that of conventional SVQ, so we selected the new anchor vector whose elements are ω_0, ω_4, and ω_9 in order to maximize the normalization effect. Except for the anchor vector, the two sub-vectors Ω_0 and Ω_1 can be represented by

$$\begin{aligned} \Omega_0 &= \{\omega_1, \omega_2, \omega_3\}, \\ \Omega_1 &= \{\omega_5, \omega_6, \omega_7, \omega_8\} \end{aligned} \tag{5}$$

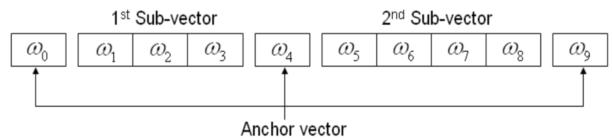

Fig. 2. LSF vector split method

In the above split structure, the range of Ω_n is limited by the adjacent anchor LSF and can be represented by

$$\omega_0 < \Omega_0 < \omega_4, \quad \omega_4 < \Omega_1 < \omega_9 \tag{6}$$

After quantization of the anchor vector, the normalized distances, D_n, of the remaining sub-vectors, Ω_n, are represented as

$$\begin{aligned} D_0 &= \{d_1, d_2, d_3\}, \\ D_1 &= \{d_5, d_6, d_7, d_8\} \end{aligned} \tag{7}$$

where

$$\begin{aligned} d_n &= \max\left\{0, \frac{\omega_n - \hat{\omega}_0}{\hat{\omega}_4 - \hat{\omega}_0}\right\}, n = 1,2,3 \\ d_n &= \max\left\{0, \frac{\omega_n - \hat{\omega}_4}{\hat{\omega}_9 - \hat{\omega}_4}\right\}, n = 5,6,7,8 \end{aligned} \tag{8}$$

Therefore, the codebook designed by normalized distance can vary depending on the anchor LSF.

When the normalization scheme is applied to vector quantization, weighted distance measure between the original normalized distance, D_n, of LSF sub-vector, Ω_n, and the approximated distance, $\hat{D}_n$, is defined

$$d\left(D_n, \hat{D}_n\right) = \left(D_n - \hat{D}_n\right) \cdot W^T \tag{9}$$

A dynamic weighting vector, W, which is the same vector as in EVRC-B [7], puts more weight to the closer elements among the LSF vector elements.

$$w(n) = \begin{cases} \dfrac{0.5}{2\pi\sqrt{(w_n - w_{n-1})(w_{n+1} - w_n)}}; \ 1 \le n \le 8 \\ \dfrac{0.5}{2\pi\sqrt{w_n (w_{n+1} - w_n)}}; \ n = 0 \\ \dfrac{0.5}{2\pi\sqrt{(w_n - w_{n-1})(0.5)}}; \ n = 9 \end{cases} \tag{10}$$

After applying weighting to the LSF vector, the importance among all elements of the vector becomes similar. This similar importance allows us to search a codeword that simply minimizes the MSE between the weighted distance and the approximated distance vectors.

2.3. Dynamic Bit Allocation

By introducing the normalization and LSVQ technologies which could minimize the dynamic range of sub-vectors, we can enhance the quantization efficiency. To combine the above-mentioned normalization and LSVQ techniques, we changed the anchor from ω_3 and ω_5 (Fig. 1) to ω_0, ω_4, and ω_9 (Fig. 2). Coding sequence is as follows. First, we quantize the anchor vector, whose elements are ω_0, ω_4, and ω_9, and then select the multi-codebook using the quantized anchor. Finally, we normalize the sub-vectors according to the quantized anchor vector, and then quantize the normalized sub-vectors.

In order to improve further upon the previously described quantizer, we propose the DBA technique to the structure. The basic idea of this technology is to allocate more bits where the dynamic range of sub-vectors is broad, and to allocate fewer bits where it is narrow. Due to the similar importance among all elements of the vector after weighting, we can allocate the number of bits in proportion to the width of sub-vector. Fig. 3 depicts the distributions of each element in LSF. The three solid lines represent the distribution for the anchors. It can be observed from the figure that the dynamic range of ω_4 is wide, which means ω_4 can be one of the most effective anchors for the DBA technique.

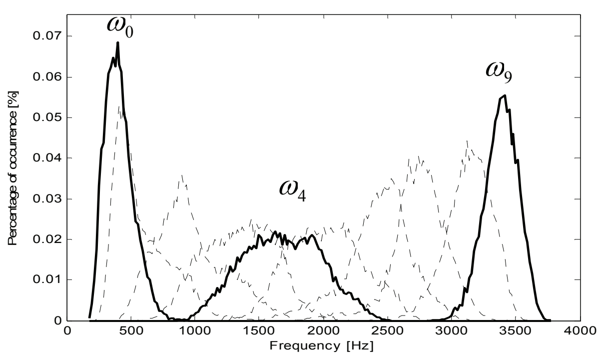

Fig. 3.LSF distributions

The concept of dynamic bit allocation in this split structure is depicted in Fig. 4, where the predefined thresholds, thr_0, thr_1 and thr_2, are reference points for determining the width of each sub-vector using anchor LSFs. The dynamic bit allocation sums up the same number of bits for the entire range regardless of different combinations of sub-vectors such as (a) - (d) in Fig. 4. Case (a) in Fig. 4 describes the bit allocation scheme where the quantized anchor LSF, $\hat{\omega}_4$, is smaller than thr_0. In case (a), the width of the lower sub-vector, Ω_0, is relatively narrow, while the width of the higher sub-vector, Ω_1 is relatively broad. Therefore the number of bits allocated to the upper sub-vector should be increased, while the number of bits allocated to the lower sub-vector should be reduced. Case (d) in Fig. 4 depicts the opposite of case (a). In case (d), the number of bits allocated to the upper sub-vector should be reduced, since the width of the higher sub-vector is smaller than that of case (a).

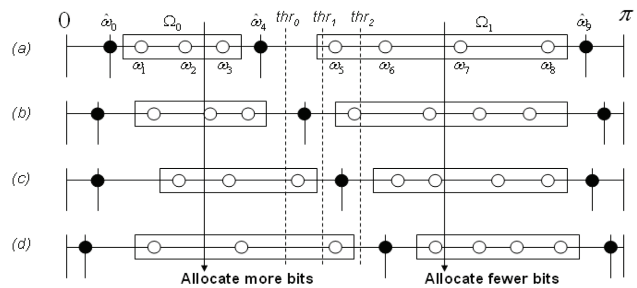

Fig. 4. Dynamic bit allocation scheme

In the quantization procedure, the sub-vector of anchor LSF is quantized by using an LBG VQ. For assigning the number of bits to each sub-vector, the anchor LSF is compared with a predefined threshold. Table 1 shows the used threshold and bit allocation examples in the 25-bit DBA-SVQ design, which has four multi-codebooks for upper and lower sub-vectors respectively. In order to calculate the threshold values, LSFs extracted from a training database (DB) are used. When the four multi-codebooks are used, the threshold values represent 1/4, 2/4, and 3/4 positions of anchor LSFs. In this 25-bit structure, 9 bits are used for anchor vector quantization and a total of 16 bits are distributed between sub-vectors 0 and 1.

Table 1. Threshold and Dynamic Bit Allocation examples

Condition with thresholds (Hz)	Bits used for Anchor	Bits used for Ω_0	Bits used for Ω_1
$\hat{\omega}_4 < 1,478$	9	6	10
$1478 \le \hat{\omega}_4 < 1700$		7	9
$1700 \le \hat{\omega}_4 < 1907$		8	8
$1907 \le \hat{\omega}_4$		8	8

3. TEST RESULTS OF THE DBA-SVQ

In order to objectively evaluate the distortion between a quantized and an original LPC vector, the spectral distortion (SD) is often used to evaluate quantizer performance [8]. The duration of the speech DB is approximately 80 minutes for training and 14 minutes for testing, both sampled at an 8 kHz sampling rate. The training DB consists of 45 minutes of NTT-AT multi-lingual speech DB and an additional 35 minutes of American English and Korean. The male and female sources comprise similar portions. The testing DB uses the clean and noisy samples with different levels. The performance of the DBA-SVQ is compared with the quantization schemes used in the full-rate EVRC which uses 4-split VQ and EVRC-B which employs a delta LSP VQ with 4-split structure. In order to have a fare comparison, the EVRC-B framework is used for the test of the proposed VQ algorithm.

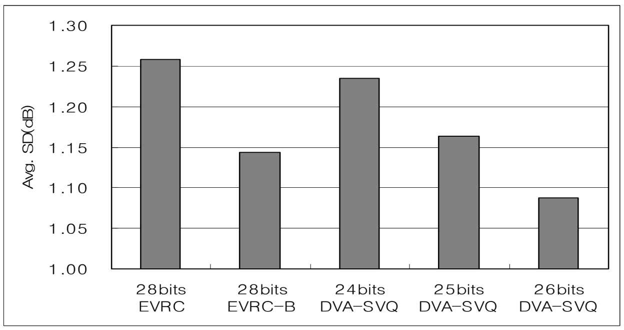

Fig. 5. SD performance at different bit rates

Table 2. SD performance of the 28-bit delta LSP VQ

28-bit delta LSP VQ	Avg. SD	Outlier (%)	
		2-4dB	>4dB
Clean speech (-22dBov)	1.1430	3.1663	0.0485
Car noise (SNR 26dB)	1.2152	5.5662	0.1114
Office noise (SNR 31dB)	1.1710	3.9033	0.0850
Street noise (SNR 15dB)	1.1912	4.8664	0.0632

Table 3. SD performance of the 25-bit DBA-SVQ

25-bit DVA-SVQ	Avg. SD	Outlier (%)	
		2-4dB	>4dB
Clean speech (-22dBov)	1.1633	1.2055	0
Car noise (SNR 26dB)	1.2256	2.5639	0.0074
Office noise (SNR 31dB)	1.1955	1.941	0
Street noise (SNR 15dB)	1.2248	2.6386	0

Table 4. Codebook size at different bit rates

	Used bits	Size (Bytes)
EVRC	28	4352
EVRC-B	28	2176
DBA-SVQ	25	13376

Fig. 5 describes the SD performance of the DBA-SVQ at different bit rates with clean samples. Table 2 and Table 3 show the spectral distortions of the DBA-SVQ and the delta LSP VQ method in EVRC-B. According to the test results, when we consider the average and outlying SDs, we can say that the performance of the 25-bit DVA-SVQ is comparable to that of the 28-bit delta LSP VQ in EVRC-B. As shown in Table 4 the codebook size of DVA-SVQ increases compared to LSP VQ in EVRC-B. However, the algorithmic complexity of the quantizer is similar, because only selected multi-codebook is searched.

4. CONCLUSION

The goal of this paper is to enhance the performance of quantizing LSF without inter-frame prediction. The main scheme of the DBA-SVQ is the dynamic bit allocation method based on the multi-codebook scheme of the LSVQ and normalized distance representation of codeword. Experimental results show lower spectral distortion and smaller number of bits than those of LSP VQ in EVRC-B. The proposed algorithm may enhance the performance in error-prone conditions and the coding efficiency of signal portions with low inter-frame correlation such as transition and onsets.

5. REFERENCES

[1] Tommy Vaillancourt, et al., "Efficient frame erasure concealment in predictive speech codecs using glottal pulse resynchronisation," in Proc. ICASSP-2007, vol. 4, pp. 1113-1116, 2007.

[2] F. Itakura, “Line spectral representation of linear predictive coefficients of speech signals,” J. Acoust. Soc. Amer., vol. 57, pp. S35, Supplement 1, 1975.

[3] J. Skoglund and J. Linden, “Predictive VQ for noisy channel spectrum coding: AR or MA?,” in Proc. ICASSP-97, vol. 2, pp. 1351–1354, 1997.

[4] M. Kim, N. Ha, and S. Kim, “Linked split-vector quantizer of LPC parameters,” ICASSP-96, vol. 2, pp. 741-744, 1996.

[5] Smith, A.M., Ashley, J.P., Jasiuk, M.A., et al., "Normalization and polygon error detection for split VQ of line spectral frequencies," 2000 IEEE Workshop on Speech Coding, pp. 123 – 125, 2000.

[6] M. Xie and J. Adoul, "Algebraic vector quantization of LSF parameters with low storage and computational complexity," IEEE Trans. Speech and Audio Proc., pp. 234 - 239, 1996.

[7] 3GPP2 C.P0014-B, “Enhanced variable rate codec, speech service option 3 and 68 for wideband spread spectrum digital systems,” version 1.0, Jan. 2007.

[8] K. K. Paliwal and B. S. Atal, “Efficient vector quantization of LPC parameter at 24 bits/frame,” IEEE Trans. Speech Audio Processing, vol. 1, no. 1, pp. 3-14, 1993.