Low Complexity Vector Quantizer Design for LSP Parameters

*Hong-Chae Woo

※The author wishes to acknowledge the financial support of the Korea Research Foundation made in the program year of 1997.

Abstract

Spectral information at a speech coder should be quantized with sufficient accuracy to keep perceptually transparent output speech. Spectral information at a low bit rate speech coder is usually transformed into corresponding line spectrum pair parameters and is often quantized with a vector quantization algorithm. As the vector quantization algorithm generally has high complexity in the optimal code vector searching routine, the complexity reduction in that routine is investigated using the ordering property of the line spectrum pair. When the proposed complexity reduction algorithm is applied to the well-known split vector quantization algorithm, the 46% complexity reduction is achieved in the distortion measure compu- tation.

I. Introduction

Most of recent speech coders are based upon linear prediction(LP) in which a current speech sample can be represented as an weighted sum of previous speech samples. As it is known, speech signals, especially voice signals, are correlated in the neighbor samples and in the pitch intervals. These correlations are usually investigated with an all-pole short-term prediction filter which models the spectral envelope of the speech signal, and a long-term prediction filter which models the spectral fine structure of it [1].

At low bit rate speech coders, the code excited linear prediction(CELP) scheme in Fig. 1 that is based on an analysis-by-synthesis technique is mostly adopted as a standard structure. The basic CELP algorithm has advanced as many standard speech coders such as 16Kbits/s low delay CELP in the public switched telephone network, 8Kbits/s vector sum excited linear prediction(VSELP) in the north America digital cellular system, the U. S. federal standard 4.8Kbits/s CELP, G. 929 8Kbits/s conjugate-structure algebraic CELP(CS-ACELP) in the digital cellular telephony, and IS-95 8Kbits/s Qualcomm CELP in the code division multiple access(CDMA) digital cellular system[2][3]. As shown in Fig. 1, the coefficients of both the long-term synthesis filter $1/P(z)$ and the short-term synthesis filter $1/A(z)$ are extracted from a block of input speech samples and after quantization routines they are sent to the other side as side information.

For quantizing the coefficients of the short-term synthesis filter, there are needed about 20-40bits/frame which

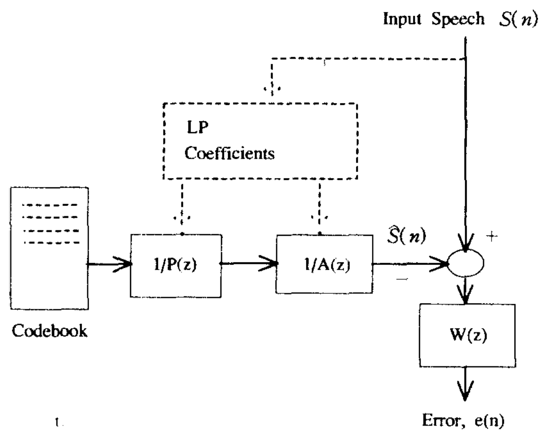

Figure 1. Block Diagram of Code Excited Linear Prediction Coder.

take a large portion of overall bit rates in a low bit rate speech coder [2]-[6]. The effective quantization of them are quite important since the spectrum envelopes of speech signals should be maintained for high quality speech. As the coefficients of a LP filter themselves do not have proper properties for quantization such as their wide dynamic range, these coefficients are often transformed into line spectrum pair(LSP) parameters that are suitable for quantization [5]. In section II, LSP quantization will be discussed focusing on vector quantization(VQ). In section III, a low complexity vector quantization algorithm will be proposed. The performances of the proposed algorithm will be evaluated at section IV. Section V summarizes the conclusions.

II. LSP Parameter Quantization

Recently the cellular phone is very popular and there

* School of Computer & Commun. Taegu University
Manuscript Received : September 28, 1998.

are more ten million users in Korea. In the wireless mobile communications, the available spectrum bandwidth is very limited so that the available total bits for a speech coder is usually at around 8Kbits/s [2][3]. But the speech quality even at low bit rates should be maintained to be at least toll quality. Since the spectral envelope of the speech signal is important for high perceptual speech, the quantization on the coefficients of the short-term LP filter should be done accurately under the restriction of the small amount of bit usage. For quantization the LP coefficients can be transformed into several types of parameters such as log-area ratio(LAR), arcsine coefficient (ASRC), and LSP. The LSP parameters are often selected because of the good properties for quantization [3][5][6].

Now let the sampling frequency of speech signals be 8KHz and the frame size for extracting the coefficients of the short-term LP filter be 160 samples/frame. The LP coefficients are obtained with the autocorrelation method. Let the all-pole short-term filter be expressed as $H(z) = 1/A(z)$, where $A(z)$ is given by

$$A(z) = 1 + a_1 z^{-1} + \cdots + a_p z^{-p} \tag{1}$$

Here the p is the order of the inverse filter $A(z)$ and is assumed as 10. The LSP parameters can be computed with defining two polynomials, $P(z)$ and $Q(z)$;

$$P(z) = A(z) + z^{-(p+1)} A(z^{-1}) \tag{2}$$

$$Q(z) = A(z) - z^{-(p+1)} A(z^{-1}) \tag{3}$$

The roots of $P(z)$ and $Q(z)$ are called the LSP parameters, which have special properties. The first property is that all the zeros of $P(z)$ and $Q(z)$ are located on the unit circle. The second is that the zeros of $P(z)$ and $Q(z)$ are interlaced on the unit circle and the magnitudes of the zeros are in the ascending order, which is called as the ordering property. It is also known that if the ordering property is satisfied, the stability of H(z) is guaranteed. Then, let zeros from $P(z)$ and $Q(z)$ be $\omega = \{\omega_1, \omega_2, \cdots, \omega_p\}$. The zeros of $P(z)$ are $\omega_1, \omega_3, \cdots, \omega_{p-1}$ and the zeros of $Q(z)$ are $\omega_2, \omega_4, \cdots, \omega_p$. Considering the ordering property, it can be said that $0 < \omega_1 < \omega_2 < \cdots < \omega_p < \pi$.

As the values of LSP parameters are bounded and the special ordering property exits among the parameters, they are suitable for quantization. LSP parameters are often quantized in several different scalar quantization methods because of simplicity in scalar quantization. In Sugamura and Farvardin's works [4], uniform, nonuniform, differential, and adaptive scalar quantization methods were investigated in 32-38 bits/frame. The U.S. federal standard 4.8Kbits/s CELP coder uses the 34bits/frame scalar quantizer [5].

However, vector quantization can provide more efficient compression than scalar quantization at the same bit usage. But vector quantization generally demands high computational complexity in the optimal codevector search routine and large memory in stbring a codebook [7]. Many modified algorithms on vector quantization have researched to avoid these difficulties. Split VQ splits a 10-dimensional LSP vector into 2 or 3 small dimensional vectors in which 1dB spectral distortion is obtained at 24 bits/frame [5]. The split VQ algorithm is known as a bench mark on the LSP VQ. Multistage VQ is based upon multi-level quantization in which 3-6 stages are used and 1dB spectral distortion is obtained at 22 bits/frame [6]. The residual values of one stage are quantized in the next stage, and this step is repeated again. The CS-ACELP coder uses both the split VQ and the multistage VQ method for the LSP quantization [3]. In the first, a 10-dimensional vector is quantized with the usual vector quantizer. Secondly, the difference vector between the original LSP and the quantized values is quantized with the split VQ method in which a 10-dimensional vector is splitted into two 5-dimensional vectors. The CS-ACELP with the 4th order moving average interframe prediction spends 18bits/frame for the LSP quantization. Thus, the complexity on VQ algorithms of the LSP needs to be reviewed and be investigated further.

III. Complexity reduction in split VQ

When the VQ method is considered for the LSP, it cannot be implemented with regarding all the LSP parameters as a vector since it requires a huge storage and very high complexity. Split VQ suggested by Paliwal and Atal [5] can reduce the storage requirement and high complexity by splitting a 10-dimensional LSP parameter vector into 2 or 3 small dimensional vectors. However, the complexity of two-part split VQ with 24 bits/frame is still high, in which about 40kilo of memory locations and about 4 million multiplications/sec are required in case of a weighted distortion measure [5]. As seen in split VQ, a large portion of the computational complexity comes from the distortion measure calculation between an input vector and each codevector in a codebook to search an optimal codevector. Thus, a complexity reduction algorithm in the optimal codevector search routine is further investigated using the ordering property in the LSP parameters.

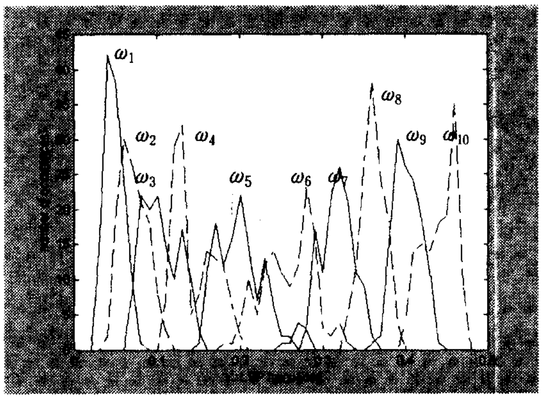

Figure 2. The histogram of LSP parameters in a test sentence.

Examining the histogram of LSP parameters in the normalized frequency in Fig. 2, it can be found that the distribution of each LSP parameter has overlapped with those of neighbored parameters. This implies that given an input LSP vector to be quantized, there can be many codevectors which do not agree with the ordering property of the input vector. Note that the violating codevectors cannot be candidates for vector quantization because of the stability condition in the linear filter H(z). These code vectors can be then exempted from the distortion computation in the optimal codevector search routine. Suppose ω_i is an element of an input LSP vector ω. Let $c^1{}_{jk}$ be an element of a codevector of the first part codebook in (4, 6) split VQ in which j indicates the index of a codevector and k does the index of an element in a codevector. As each codebook in the (4, 6) split VQ scheme with the 24 bits/frame is allocated with 12 bits, j ranges from 1 to 4096 and k does from 1 to 4 in $c^1{}_{jk}$. For the second part of (4, 6) split VQ, the indexes of the $c^2{}_{jk}$ codebook are defined in the same way but k goes to 6. During the complexity reduced optimal codevector search routine in the split VQ shown in Fig. 3, the violation of the ordering property at each codevector is checked before the distortion computation. Check if the ordering property, $\omega_i < c^1{}_{j(i+1)}$ is satisfied for $i = 1, 2, 3$ at each $j = 1, \cdots, 4096$. The distortion computation of the current codevector $c^1{}_j$ is exempted if there is any violation. For the $c^2{}_j$, check if $\omega_i < c^2{}_{j(i-3)}$ for $i = 5, \cdots, 9$ at each $j = 1, \cdots, 4096$ by the same way. And the ordering property is also considered between the vector c^1 and c^2. Then, check if $\widehat{\omega}_4 < c^2{}_{j1}$ in which $\widehat{\omega}_4$ is the quantized value of ω_4 and is known at the $c^1{}_j$ routine.

1. Initialize: set $j = 1$.
2. Ordering property:
Check if $\begin{cases} \omega_i < c^1{}_{j(i+1)} & \text{for } i = 1,2,3 \\ \omega_i < c^2{}_{j(i-3)} & \text{for } i = 5,\cdots,9 \end{cases}$
3. Decision
If inequality not satisfied, $j \leftarrow j + 1$.
Otherwise, compute distortion & $j \leftarrow j + 1$.
4. Repeat: go to step 2.
5. Stop when $j = 4096$.

Figure 3. Low complexity optimal codevector search routine.

IV. Performances

Split VQ is actually the same as general VQ. Thus, using the well-known LBG algorithm, the codebooks for (4, 6) split VQ at 24bits/frame are searched. Each codebook has 4096(12bits) codevectors. For the codebook generation, eight Korean test speech files are recorded at the FM radio stations. Each file has one million samples(125 second recording time) at the sampling rate 8KHz. The half of them are female voices. The short-term LP coefficients at every 160 samples are computed using the autocorrelation method with Hamming window. The frame rate is 50 frames/sec and the total frames from test speech files are about 50000. All the LSP parameters are then calculated from LP coefficients. Each of c^1 and c^2 codebooks are searched from 50000 LSP training vectors.

The spectral distortion(SD) is usually used to evaluate the performances of LSP parameter quantization. It is known that transparently reproduced speech quality can be obtained at average 1 dB SD, and a small number of high distortion frames which are about 1-2% outlier frames on the range of 2-4 dB SD and no outlier frames above 4 dB SD [5]. The average SD is defined as

$$SD = \frac{1}{N_f} \sum_{n=1}^{N_f} \left(\frac{100}{\pi} \int_0^{\pi} [\log|A_n(e^{j\omega})|^2 - \log|\widehat{A}_n(e^{j\omega})|^2]^2 d\omega \right) \quad (4)$$

where $1/|A_n(e^{j\omega})|^2$ and $1/|\widehat{A}_n(e^{j\omega})|^2$ are the spectras of the all-pole synthesis filters with respectively the unquantized and the quantized coefficients at the nth frame. And N_f is the total number of frames on database speech files.

Four Korean test sentences with each 20000 samples are again recorded from the FM radio stations and LSP parameters are also computed. The mean squared error (MSE) distortion measure is used in split VQ. Table 1 shows the spectral distortion performances. According to Paliwal and Atal's results [5], the average SD was 1.19 dB, outliers between 2-4 dB were 4.30%, and outliers above 4 dB were 0.03%. We have almost the same performances as theirs.

Table 1. SD performances of (4, 6) split VQ at 24 bits/frame.

test sentence	average SD(dB)	outliers(in %)	
		2-4dB	> 4dB
male 1	1.20	3.90	0.0
male 2	1.17	1.56	0.0
female 1	1.11	1.56	0.0
female 2	1.25	3.12	0.0

The complexity reduction algorithm is evaluated by counting the number of codevectors that are exempted from the MSE distortion computation during the optimal codevector search. Table 2 shows the MSE distortion computation exemption rate in percentage for four Korean test sentences.

Table 2. The MSE computation exemption rate for Korean test sentences.

test sentence	c^1 (%)	c^2 (%)
male 1	34.76	49.74
male 2	40.95	53.58
female 1	49.25	46.53
female 2	46.87	48.11

The exemption rate for c^2 is higher than that for c^1 since a codevector in c^2 has more elements to be checked than a codevector in c^1. The average exemption rate is 46.22%. For reliability in Table 2, the complexity reduction algorithm in split VQ was tested for two long sentences with each one million samples. The female long sentence showed the computation exemption rate 49.59% for c^1 and 42.63% for c^2. In the male long se- ntence, the rate was 38.98% for c^1 and 53.21% for c^2. In these cases, there were no performance degradations when the complexity reduction routine was applied.

In split VQ with the MSE distortion measure, two million multiplications/sec complexity in the optimal codevector search are needed because of 10 multiplications for each codevector, 4096 codevectors/codebook, and 50 frames/sec. It can said that, about one million multiplications/sec can be saved here. As Paliwal and Atal have computed, four million multiplications/sec. are necessary in case of the weighted distortion measure usage [5]. Even if the weighted distortion measure is used, the exemption rate will not be changed since the ordering property is independent of the type of a distortion measure. Then, about two million multiplications/sec. can be saved in a weighted distortion measure. In the proposed complexity reduction algorithm, the comparison routine for the ordering property check will be additionally required but the routine does not take much of computational power. It can be thus said that the ordering property checking algorithm in split VQ is very simple and the algorithm can also be included in any other VQ algorithms for the LSP such as multistage VQ and the CS-ACELP coder. Moreover, the higher the complexity of a distortion measure, the more the proposed algorithm is beneficial for the complexity reduction.

V. Conclusions

The vector quantization algorithms on the LSP parameters are very effective but they need high computational complexity in the optimal codevector search. Thus, using the ordering property of LSP parameters the simple complexity reduction algorithm is developed for LSP VQ. When the algorithm is applied to split VQ, the 46 percentage of the computational reduction are achieved in case of the MSE distortion measure computation. It corresponds to about one million multiplication reduction every second.

References

1. L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall, 1978.
2. A. M. Kondoz, *Digital Speech: Coding for Low Bit Rate Communication Systems*, Wiley, England, 1995.
3. International Telecommunication Union, *G.729 CS-ACELP Recommendation*, 1996.
4. Sugamura, N. and Farvardin, N., "Quantizer design in LSP speech analysis-synthesis," *IEEE J. on Sel. Areas in Commun.*, Vol. 6, No. 2, pp. 432-440, Feb. 1988.
5. Paliwal, K. K. and Atal, B. S., "Efficient vector quantization of LPC parameters at 24 bits/frame," *IEEE Trans. Speech and Audio Proces.*, Vol. 1, No. 1, pp. 3-14, Jan. 1993.
6. LeBlanc, W. P., Bhattacharya, B., Mahmoud, S. A. and Cuperman, V., "Efficient search and design procedures for robust multi-stage VQ of LPC parameters for 4kb/s speech coding," *IEEE Trans. on Speech and Audio Proces.*, Vol. 1, No. 4, pp. 373-385, Oct. 1993.
7. A. Gersho and R. M. Gray, Vector *Quantization and Signal Compression*, Kluwer Academic Press, Boston, 1992.

▲Hong Chae Woo

Hong Chae Woo was born in Kyungpook, Korea, on Jan. 20, 1957. He received the B.S. degree in the department of electronics, Kyungpook national university, in 1980 and the M.S. in 1988 and the Ph.D. degree in 1991 from the department of electrical engr., Texas A&M university. He has been with School of Computer and Communication Engr., Taegu University since 1992 where he is now an associate professor. His research interests are on speech coding, adaptive filters, and vector quantization.