Encoding of Speech Spectral Parameters Using Adaptive Vector-Scalar Quantization Methods for Mobile Communication Systems

*In-Sung Lee and *Jong-Hark Kim

※This paper was supported in part by Non-Directed Research Fund, National Special Support Project (No.96F-6) of Chungbuk National University.

Abstract

In this paper, an efficient quantization method of line spectrum pairs(LSP) with cascaded structure of vector quantizer and scalar quantizer is proposed. First, input LSP parameters is vector-quantized using a codebook with a moderate number of entries. In the second stage of quantization, the components of residual vector are individually quantized by the scalar quantizer. The utilization of ordering property of LSP parameters and the inclusion of interframe prediction improve the quantizer performance and remove the stability check routine after quantization procedure. The new vector-scalar hybrid quantizer using 26 bits/frame shows a transparent quality of speech that an average spectral distortion is 1 dB and the frame proportion with above 2 dB spectral distortion is less than 2 %. The performances of proposed quantization method is evaluated in the transmission errors.

I. Introduction

Most of speech coders including code excited linear prediction(CELP) coders use linear predictive coding(LPC) parameters for transmitting the short-time spectral envelope information of speech. The LPC coefficients can be transformed into mathematically equivalent representation of line spectrum pairs(LSP) that have the favorable properties for quantization and interpolation [1]. As the scalar quantization algorithm, the adaptive quantization with the backward sequence (AQBW) using the ordering property of the LSP parameters was designed by Sugamura and Farvardin [2]. The variable rate QCELP coder which was adopted as the standard vocoder(IS-96) in the North American CDMA digital cellular system quantizes the residuals of LSP parameters in the differential pulse code modulation(DPCM) system with a uniform scalar quantizer [3].

Several vector quantization (VQ) methods [4, 5, 6] were recently developed in order to overcome the performance limit of scalar quantization method. A bench-marked VQ algorithm is the split VQ approach designed by Paliwal and Atal [4]. It can obtain the 1 dB average spectral distortion (SD) by spending 24 bits/frame in which the size 10 vector is split into 4-dimensional vector and 6-dimensional vector. Another approach to exploit advantages offered by vector quantization while reducing the computational complexity and memory usage is to use a vector-scalar hybrid quantization algorithm. Grass and Kabal proposed several methods for vector-scalar quantization that achieved 1 dB spectral distortion by spending 30 bits/frame [6]. Recently, an improved 4.8 kbps CELP coder using a vector-scalar LSP quantization method was developed [7].

In this paper, an efficient vector-scalar hybrid 0quantization methods which has good performance, low complexity and memory are proposed. The maximum quantization range of each LSP parameter is varied adaptively on the quantized value of the previous order's LSP parameter. The performances of proposed methods were evaluated in the noisy channel as well as in noise-free channel. This paper is organized as follows. In section 2, properties of LSP parameters are reviewed briefly. In section 3, the vector-scalar hybrid quantizer is explained. In section 4, adaptive vector-scalar quantizer using the ordering property is described. In section 5, the performances of proposed algorithms are evaluated. Their performances in the noisy channel are presented.

II. LSP Analysis

In a speech coder, the parameters for the short-term prediction are generally updated every 20 miliseconds (ms) to 30 ms at the 8 kHz sampling rate. The coefficients of the short-term prediction error filter (STP) are extracted by the autocorrelation method from the

* Department of Radio Engineering, Chungbuk National University
Manuscript Received : October 30, 1998.

bias-removed input speech signal. The short-term synthesis filter is given as follows:

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

The a_i, $i = 1, \cdots, p$, are the LPC coefficients and p is the order of the filter. In the low bit rate CELP coder, the order of STP filter is usually given by 10. These coefficients are then transformed into the LSP parameters which have the excellent properties for quantization such as the boundness of the parameters, easy stability chec-king condition of synthesis filter, and the ordering relation among the parameters. In order to extract the LSP parameters from the STP coefficients, two polynomial $P(z)$ and $Q(z)$ are introduced as follows:

$$\begin{aligned} P(z) &= A(z) - z^{-(p+1)} A(z^{-1}) \\ &= (1 - z^{-1}) \prod_{i=2,4,\cdots,p} (1 - 2z^{-1}\cos\omega_i + z^{-2}) \end{aligned} \tag{2}$$

$$\begin{aligned} Q(z) &= A(z) + z^{-(p+1)} A(z^{-1}) \\ &= (1 + z^{-1}) \prod_{i=1,3,\cdots,p-1} (1 - 2z^{-1}\cos\omega_i + z^{-2}) \end{aligned} \tag{3}$$

Note that w_i are the LSP parameters and e^{jw_i} are the roots of the $P(z)$ and $Q(z)$. Moreover, the roots of $P(z)$ and $Q(z)$ have very important properties [2]. Firstly, all roots of $P(z)$ and $Q(z)$ are located on the unit circle. Secondly, the roots of $P(z)$ and $Q(z)$ are interlaced with each other on the unit circle. From the second property the following specific relationship among the LSP parameters is obtained;

$$0 = \omega_0 < \omega_1 < \cdots < \omega_p < \omega_{p+1} = \pi \tag{4}$$

If the ordering property in (4) is satisfied, the stability of the short-term synthesis filter is guaranteed. The orde-ring property is also very useful for effectively quantizing the LSP parameters [2,11].

III. Vector-Scalar Hybrid Quantization

The advantage of vector-scalar hybrid quantization (VQ-SQ) is that a codebook size in vector quantization stage is much smaller than that in a single stage vector quantization with the same number of bits. Thus, the complexity for codebook search, and memory requirement can be drastically reduced by using the vector-scalar hybrid quantizer of LSP parameters. The block diagram is shown in Fig. 1. At first, input LSP parameters are vector quantized using a codebook with the moderate number of entries. Multiple candidate codevectors are selected in the VQ stage. In the second stage of quanti-zation, the components of residual vector are individually quantized by the scalar quantizer.

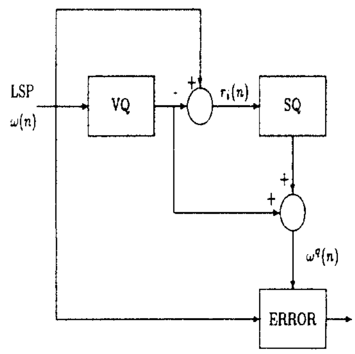

Figure 1. The Vector-Scalar Quantizer of LSP parameters.

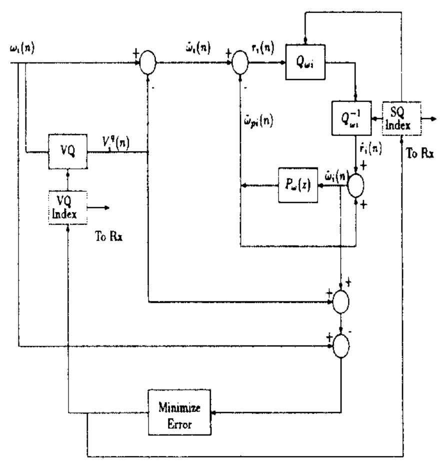

Figure 2. The Encoding Process of Vector-Scalar Hybrid Quantizer with Prediction.

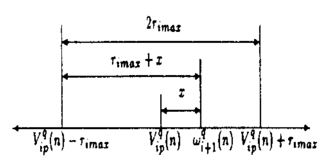

Figure 3. Maximum Quantization Range for VQ-SQ-BW.

Table I. Maximum LSP Quantization Level in Vector-Scalar Hybrid Quantizer.

LSP Frequency	Max. Range (r_{imax})
$r_{1\max}$	0.0125
$r_{2\max}$	0.0175
$r_{3\max}$	0.025
$r_{4\max}$	0.025
$r_{5\max}$	0.025
$r_{6\max}$	0.025
$r_{7\max}$	0.0225
$r_{8\max}$	0.0225
$r_{9\max}$	0.0175
$r_{10\max}$	0.0175

Even though the interframe correlation of residual LSP parameters after vector-quantization of LSP parameters is reduced, the interframe prediction in the second stage is utilized in the vector-scalar hybrid quantizer. The block diagram of encoding part of vector-scalar hybrid quantizater with interframe prediction is shown in Fig. 2. The following equations describe the encoding process of vector-scalar hybrid quantizer.

$$V_i^q(n) = V_i(index_{VQ}(n)) \tag{5}$$

$$\widetilde{\omega}_i(n) = \omega_i(n) - V_i^q(n) \tag{6}$$

$$r_i(n) = \widetilde{\omega}_i(n) - \widetilde{\omega}_\pi(n) \tag{7}$$

$$\widehat{r}_i(n) = Q_{\omega i}^{-1}[Q_{\omega i}[r_i(n)]] \tag{8}$$

$$\widehat{\omega}_i(n) = \widetilde{\omega}_\pi(n) + \widehat{r}_i(n) \tag{9}$$

$$\widetilde{\omega}_\pi(n) = 0.3625\,\widehat{\omega}_i(n-1) \tag{10}$$

$$where\ i = 1,\ 2, \cdots,\ p$$

The quantizer, $Q_{\omega i}$, for the ith LSP frequency is a linear quantizer with uniform step size. Each residual LSP frequency is quantized as follows:

$$Q_{\omega i}(x) = \max[0, \min(2^{N-1}, Q_{ti}(x))] \tag{11}$$

where $Q_{ti}(x) = round\left(\frac{2^{N-1}}{2}\frac{x + r_{imax}}{r_{imax}}\right)$, N is the number of quantization bits, r_{imax} is the maximum quantization level, and $round(x)$ is the function rounding to the closest integer. The maximum quantization range r_{imax} is given in Table 1. The final optimal VQ-SQ codevector which has a minimum distortion is selected among multiple VQ-SQ quantized value combinations.

IV. Adaptive Vector-Scalar Hybrid Quantization

The adaptive quantization range method utilizing the ordering property of LSP parameters is applied in the vector-scalar hybrid quantization of LSP parameters so that the maximum quantization range of LSP parameters in the second stage scalar quantizer can be adaptively varied. The scalar quantization of residual LSP parameters starts from $r_{10}(n)$ and then proceed to the quantization of lower order residual LSP parameters. The improved vector-scalar hybrid quantization method is called as vector-scalar hybrid quantization with backward sequence-(VQ-SQ-BW). In this algorithm, the 10th order residual LSP parameter, $r_{10}(n)$ is quantized with the usual maximum quantization range. From $r_9(n)$, decide if the maximum quantization range could be shrinkable by the examination that the quantized LSP parameters satisfy the ordering property of LSP parameters. The checking variable in VQ-SQ-BW that decides the maximum quantization range in the scalar quantizer is given by

$$x = \omega^q_{i+1}(n) - V^q_{ip}(n) \tag{12}$$

where, $\omega^q_{i+1}(n)$ is the quantized value of the $(i+1)$th order LSP parameter and $V^q_{ip}(n)$ is the sum of the ith order predicted value of scalar quantizer, $\omega_\pi(n)$, and the ith order quantized output of the first stage vector quantizer, $V_i^{q(n)}$. The maximum quantization range of each scalar quantizer in the VQ-SQ-BW may be shrunk by checking the relation of the $(i+1)$th order quantized value and the sum of ith element of vector quantizer output, $V_i^q(n)$ and the predicted value $\omega_\pi(n)$ of the ith LSP parameter. The quantized value of $r_i(n)$ can not greater than x because $\omega_i^q(n)$ should be less than $\omega^q_{i+1}(n)$ to gaurantee the stability of short-term synthesis filter. Therefore, if $|x| < r_{imax}$, it is not necessary to assign the normal quantization range of $r_i(n)$ that covers from $-r_{imax}$ to $+r_{imax}$. The maximum quantization range of $r_i(n)$ can be reduced to $[-r_{imax}, x]$. The reduced maximum quantization range is shown in Fig. 3. The following quantizatiom method is possible. If $|x| < r_{imax}$, quantize $r_i(n)$ with the shrunk quantization range, $i.e.$, $-r_{imax} \sim x$ and otherwise, quantize $r_i(n)$ with the nominal quantization range, $i.e.$, $-r_{imax} \sim +r_{imax}$. The final optimal VQ-SQ codevector which has a minimum distortion is selected among multiple VQ-SQ quantized value combinations. The idea used in vector-scalar hybrid quantization with backward sequence can be also applied in forward sequence. This method is called by vector-scalar hybrid quantization with forward sequence (VQ-SQ-FW).

V. Simulation Results

The performances of new algorithms are evaluated in terms of the average spectral distortion(SD). The SD is defined as follows:

$$SD(dB) = \frac{1}{NF} \sum_{n=1}^{NF} [(\frac{1}{\pi} \int_0^{\pi} [10 \log_{10} |A_n(e^{j\omega})|^2 - 10 \log_{10} |\widehat{A}_n(e^{j\omega})|^2]^2 d\omega)]^{\frac{1}{2}} \quad (13)$$

where NF is the number of total frames, and $A_n(e^{j\omega})$ and $\widehat{A}_n(e^{j\omega})$ are the spectra of the nth speech frame without quantization and with quantization. Two kinds of Korean speech database are used for the experiment. One consists of 6 male and 5 female speech data which are recorded from FM radio station. Another consists of 3 male and 3 female speech data recorded in two anechoic rooms with different recording conditions. About 15000 frames with 20 msec duration are used for testing.

In the vector-scalar hybrid quantization, the codebook size of the first stage vector quantizer is selected as 256(8 bits). The codebook of vector quantizer is designed by using the LBG algorithm [12] on the training speech data. The performance of vector-scalar quantizer is improved as the number of candidate vectors in the first stage increase. The performance of the quantizer according to the number of candidate vectors is shown in Table 3. Sixteen candidate vectors in the first stage are adequate. The ratio of speech frame that is quantized with the shrunk quantization range in adaptive VQ-SQ quantization process is shown in Table 2. The simulation result shows that the shrunk quantization range is used in over 40 % of total speech frame.

Table 2. The Ratio of Frame Used the Shrunk Quantizarion Range in VQ-SQ-BW.

LSP Parameter	1	2	3	4	5
Percent	61.8	50.7	72.0	63.0	69.6
LSP Parameter	6	7	8	9	10
Percent	56.7	53.6	47.5	42.0	0.

Table 3. Performance of 27 bit/frame Vector-scalar Quantization According to the Number of Candidate Vectors.

Can Vec	1	4	8	12	16	20	24
SD	1.51	1.14	1.00	0.97	0.95	0.94	0.94

Table 4. The SD (dB) Values in Vector-Scalar Quantization Schemes.

bit/frame	VQ-SQ	VQ-SQ-BW
	SD > 2dB	SD > 2dB
	dB %	dB %
24	1.25 3.0	1.15 1.0
25	1.20 2.8	1.10 0.5
26	1.10 1.5	1.01 0.4
27	1.05 1.2	0.95 0.3
28	0.98 1.0	0.88 0.2

Table 5. Channel Error Performance of VQ-SQ-BW at 27 bit/frame.

BER	VQ-SQ-BW-AR	VQ-SQ-BW-NP
	SD > 2dB	SD > 2dB
	dB %	dB %
0	0.95 0.3	1.00 0.7
10-4	0.96 0.4	1.01 0.8
5 × 10-4	0.99 1.0	1.04 1.4
10-3	1.03 1.8	1.07 2.2
5 × 10-3	1.34 8.1	1.37 8.4
10-2	1.69 15.4	1.70 15.4

Table 6. Performance Comparision of VQ-SQ-BW and IS-96A in the noisy channel.

BER	VQ-SQ-BW	IS-96A
	SD > 2dB	SD > 2dB
	dB %	dB %
0	0.95 0.3	0.92 0.4
10-4	0.96 0.4	0.93 0.6
5 × 10-4	0.99 1.0	0.96 1.5
10-3	1.03 1.8	1.01 2.8
5 × 10-3	1.34 8.1	1.33 10.9
10-2	1.69 15.4	1.75 21.0

The performance comparisons between the conventional VQ-SQ and the VQ-SQ using adaptive quantization range method are shown in Table 4. The simulation results show that the adaptive VQ-SQ provides 1-2 bits/frame saving over the conventional VQ-SQ. The VQ-SQ-BW quantization algorithm needs 26 bits/frame to achieve 1 dB average spectral distortion. Note that the proportion of outlier frame with distortion greater than 2 dB is significantly low. As the reference quantization method, the 40 bits/frame LSP scalar quantization method of

QCELP(IS-96A) which was adopted as the standard vocoder in code division multiple access(CDMA) digital cellular system was evaluated in same speech data. The 40 bits/frame scalar quantization method of QCELP(IS-96A) shows 0.92 dB average spectral distortion and 0.4 % outlier frame above 2 dB spectral distortion.

The performances of proposed quantization methods are evaluated in the presence of channel errors and the structure of quantizers may be modified to have the robustness to the channel errors. The scalar quantizer with interframe prediction may be very sensitive on the effect of channel errors due to autoregressive(AR) structure of predictor. Therefore, the VQ-SQ-BW with no predictor is simulated in the noisy channel. The performances of VQ-SQ-BW at 27 bits/frame in the presence of channel errors are shown in Table 5. In the adaptive vector-scalar quantizer, the choice of predictor structure did not affect the performance of quantizer in the noisy channel. The simulation results show that the adaptive quantizers using the ordering property of LSP parameters perform as well as other robust LSP quantizer [13] in the presence of channel errors. The proposed adaptive quantization methods did not show the severe performance degradation at the BER below 10^{-3}. The performance comparison between VQ-SQ-BW at 27 bits/frame and LSP quantizer of IS-96A at 40 bits/frame is shown in Table 6. The proposed VQ-SQ-BW method showed better performance than LSP quantizer of QCELP(IS-96A) in the noisy channel

VI. Conclusions

The vector-scalar hybrid quantization algorithm of the LSP parameters is improved by investigating the important ordering property of the LSP parameters. The simulation results show that the adaptive vector-scalar hybrid quantization method provides about 1-2 bits/frame saving over the vector-scalar hybrid quantization scheme without the extra increase of complexity. The VQ-SQ-BW quantization algorithm needs 26 bits/frame to achieve the transparent quality of speech. The simulation results show that the adaptive quantizers using the ordering property of LSP parameter maintains the good quality of speech in the presence of channel errors.

References

1. F. K. Soong and B. H. Juang, "Line Spectrum and Speech Data Compression," in Proc. IEEE Int. Conf. Acous., Speech, Signal processing, pp.1.10.1-1.10.4, 1984.
2. N. Sugamura and N. Farvardin, "Quantizer design in LSP speech analysis-synthesis," *IEEE J. on Areas in Commun.*, vol. 6, no. 2, pp.432-440, Feb. 1988.
3. QUALCOMM Inc., "TIA/EIA IS-96 Interim Standard; Speech service option standard for wideband spread spectrum digital cellular system," 1993.
4. K. Paliwal and B. Atal, "Efficient vector quantization of LPC parameters at 24 bits/frame," *IEEE Trans. Speech and Audio Processing.*, vol.11, no.1, pp.3-14, Jan. 1993.
5. W. LeBlanc, B. Bhattacharya, S. 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. Speech and Audio Procs.*, vol.1, no.4, pp.373-385, 1993.
6. J. Grass and P. Kabal, "Methods of improving vector-scalar quantization of LPC coefficients", in Proc. *IEEE Int. Conf. Acous., Speech, Signal processing*, Toronto, pp.657-660, 1991.
7. T. Miyano,M. Serizawa, J. Takizawa, S. Ikeda, and K. Ozawa, "Improved 4.8 kb/s CELP coding using two-stage vector quantization with multiple candidates (LCELP)," in Proc. *IEEE Int. Conf. Acous., Speech, Signal Processing*, pp. I.312-I.324, 1992.
8. Soong, F. K. and Juang, B. H.: 'Optimal quantization of LSP parameters,' *IEEE Trans. Speech and Audio Processing*, 1993, vol. 1, no. 1, pp. 15-24
9. N. Farvardin and R. Laroia, "Efficient coding of speech LSP paramters using discrete cosine transform," in Proc. *IEEE Int. Conf. Acous., Speech, Signal* Processing, pp. 801-804, 1989.
10. C. C. Kuo. F. R. Jean, and H. C. Wang, "Low bit-rate quantization of LSP parameters using two-dimensional differential coding," in Proc. *IEEE Int. Conf. Acous., Speech, Signal Processing*, pp. I.97-I.100, 1992.
11. K. T. Malone and T. R. Fischer, "Enumeration and trellis-searched coding schemes for speech LSP parameters," *IEEE Trans. on Speech and Audio Processing*, vol. 1, no. 3, pp. 304-314, Jul. 1993.
12. Y. Linde, A. Buzo, and R. Gray, "An algorithm for vector quantizer design," *IEEE Trans. on Commun.*, vol. COM-28, pp.84-95, Jan. 1980.
13. K. Paliwal and B. Atal, "Vector quantization of LPC parameters in the presence of channel errors," *Speech and Audio Coding for Wireless and Network Applications*, pp.191-201, 1993.

▲In Sung Lee

Insung Lee received the B.S and M.S degrees in electronic engineering from Yonsei University, Korea, in 1983 and 1985, respectively, and the PH.D degree in electrical engineering from Texas A&M Unversity, U.S.A, in 1992.

From 1986 to 1987, he was a research engineer at the Korea Telecom Research Center, Korea. From 1989 to 1992, he was a graduate research assistant in the Department of Electrical Engineering, Texas A&M University, U.S.A. From 1993 to 1995, he was with the Signal Processing Section of the Mobile Communication Division at the Electronics and Telecommunications Research Institute (ETRI), Korea, as a senior member of technical staff. Since 1995, he has been with the Department of Radio Engineering, Chungbuk National University, as an Assistant Professor.

His current research interests include speech and image compression, mobile communication, adaptive filters.

▲Jong Hark Kim

Jong Hark Kim received the B.S. degrees in electrical engineering and he was a graduate student, M.S., in radio engineering from the Cungbuk National University, Korea, in 1997, 1998, respectively.

His current interests include speech and audio coding, image data compression and digital signal processing.