ENCODING OF SPEECH SPECTRAL PARAMETERS USING ADAPTIVE QUANTIZATION METHODS

Insung Lee[1] *Hong Chae Woo*[2]

[1]School of Electrical and Electronic Eng., Chungbuk National University
San-42, Gaesin-dong, Cheongju, Korea
email: inslee@cbucc.chungbuk.ac.kr
[2]Dept. of Electronic Eng., Taegu Univ., Kyungbuk, Korea

ABSTRACT

Efficient quantization methods of the line spectrum pairs(LSP) which have good performances, low complexity and memory are proposed. The adaptive quantization method utilizing the ordering property of LSP parameters is used in a scalar quantizer and a vector-scalar hybrid quantizer. The maximum quantization range of each LSP parameter is varied adaptively on the quantized value of the previous order's LSP parameter. The proposed scalar quantization algorithm needs 31 bits/frame which is 3 bits less than in the conventional scalar quantization method with interframe prediction to maintain the transparent quality of speech. The improved vector-scalar quantizer achieves an average spectral distortion of 1 dB using 26 bits/frame. The performances of proposed quantization methods are evaluated in the transmission errors.

1. 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). For the quantization of LSP parameters, several vector quantization (VQ) methods [1, 2, 3] were recently developed in order to overcome the performance limit of scalar quantization method. A typical VQ algorithm is the split VQ approach designed by Paliwal and Atal [1]. It can obtain the 1 dB average spectral distortion (SD) by spending 24 bits/frame in which the size 10 vector is split into the small size 4 vector and the size 6 vector. Another approach to exploit advantages offered by vector quantization while reducing the computational complexity and memory is to use a vector-scalar quantization algorithm. Grass and Kabal proposed several methods for vector-scalar quantization that achieved 1 dB spectral distortion by spending 30 bits/frame [3]. Recently a vector-scalar quantizer was used in a 4.8 kbps CELP coder.[4].

Even though vector quantization techniques show good performances in LSP quantization, applications of them have been so far limited due to the increase of computational complexity and storage. 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 [5]. 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 [6].

In this paper, an efficient scalar quantization method which has a good performance and very low complexity and memory is proposed. A new scheme utilizing the ordering property in the DPCM algorithm is developed so that the quantization ranges of LSP parameters can be reduced significantly. Moreover, the adaptive quantization scheme utilizing the ordering property of LSP parameters is applied in the vector-scalar hybrid quantization method. We study the performance of proposed methods in the presence of transmission errors and modify them in order to maintain good performances in the noisy channel. This paper is organized as follows. In section 2 and 3, the DPCM scheme and vector-scalar quantizer utilizing the ordering property of LSP parameters are presented. In section 4, the performances of proposed algorithms are evaluated. In section 5, their performances in the noisy channel are presented.

2. ADAPTIVE SCALAR QUANTIZATION OF LSP PARAMETERS

In a speech coder the short-term synthesis filter is given as follows:

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

The a_i, $i = 1, \cdots, p$, are the LPC coefficients and the filter order p is given by 10. These coefficients are then transformed into the LSP parameters. The following specific relationship among the LSP parameters is obtained:

$$0 = \omega_0 < \omega_1 < \cdots < \omega_p < \omega_{p+1} = \pi \tag{2}$$

If the ordering property in (2) is satisfied, the stability of the short-term synthesis filter is guaranteed.

The encoding process of DPCM system utilizing the interframe correlation is shown in Fig. 1. The following equations describe the encoding process of LSP quantization.

$$B_i = \frac{0.5i}{p+1} \quad \text{where } p = 10 \tag{3}$$

*This work was funded by The Information and Communications Department of Korea, Contract No. U96-158

Table 1. Maximum LSP Quantization Level

LSP Frequency	Max. Range (e_{imax})
ω_1	0.03
ω_2	0.04
ω_3	0.07
ω_4	0.07
ω_5	0.06
ω_6	0.06
ω_7	0.05
ω_8	0.05
ω_9	0.04
ω_{10}	0.04

$$
\begin{aligned}
\tilde{\omega}_i(n) &= \omega_i(n) - B_i & (4)\\
e_i(n) &= \tilde{\omega}_i(n) - \tilde{\omega}_i^p(n) & (5)\\
\hat{e}_i(n) &= Q_{\omega i}^{-1}[Q_{\omega i}[e_i(n)]] & (6)\\
\hat{\omega}_i(n) &= \tilde{\omega}_i^p(n) + \hat{e}_i(n) & (7)\\
\tilde{\omega}_i^p(n) &= 0.90625\hat{\omega}_i(n-1) & (8)\\
& \text{where } i = 1,\ 2, \cdots,\ p &
\end{aligned}
$$

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

$$Q_{\omega i}(x) = max[0, min(2^N - 1, Q_{ti}(x))] \quad (9)$$

where $Q_{ti}(x) = round(\frac{2^N-1}{2}\frac{x+e_{imax}}{e_{imax}})$, N is the number of quantization bits, e_{imax} is the maximum quantization level, and $round(x)$ is the function rounding to the closest integer. The maximum quantization range e_{imax} is given in Table 1.

The adaptive quantization method considering the ordering property of LSP parameters is applied in the DPCM system so that the maximum quantization range of LSP parameters can be adaptively varied. The scalar quantization of a LSP parameter starts from the ω_{10} and then proceeds to the quantization of the lower order LSP parameters. This approach is called the DPCM with backward sequence (DPCM-BW). In this algorithm, $\omega_{10}(n)$ is quantized with the usual maximum quantization range. From $\omega_9(n)$, check if the maximum quantization range could be shrinkable by using the ordering property of the LSP parameters. Define a checking variable:

$$z = \omega_{i+1}^q(n) - \omega_i^p(n) \quad (10)$$

Here, $\omega_{i+1}^q(n)$ is the quantized value of the $(i+1)$th order LSP parameter, and the predicted value of the ith order LSP parameter adding the bias term, $\omega_i^p(n)$, is given by $\tilde{\omega}_i^p(n) + B_i$. The quantized value of $e_i(n)$ can not be greater than z because the $\omega_i^q(n)$ should be less than $\omega_{i+1}^q(n)$ to guarantee the stability of short-term synthesis filter. Therefore, if $|z| < e_{imax}$, it is not necessary to assign the quantization range of $e_i(n)$ that covers from $-e_{imax}$ to $+e_{imax}$. The maximum quantization range of $e_i(n)$ can be reduced to $[-e_{imax}, z]$. The reduced maximum quantization range is shown in Fig. 2. Another advantage of the proposed quantization method is that it does not need to have the stability

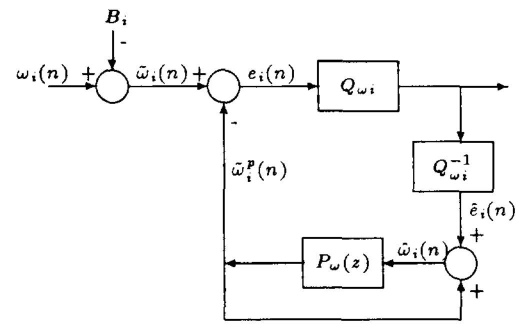

Figure 1. The DPCM Encoding Process for the LSP Parameters

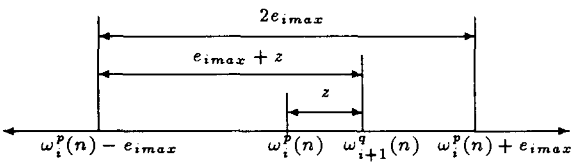

Figure 2. Maximum Quantization Range for DPCM-BW

checking routine in the decoding process because the new algorithm is already satisfied with the ordering condition of LSP parameters.

The quantization algorithm of DPCM-BW is as follows:
DPCM with Backward Sequence(DPCM-BW)

1. Quantize $\omega_{10}(n)$ with DPCM quantizer which have maximum quantizer range $[-e_{imax}, e_{imax}]$ and set i=9.
2. Compute a checking variable $z = \omega_{i+1}^q(n) - \omega_i^p(n)$.
3. (a) if $|z| < e_{imax}$, then quantize $e_i(n)$ using the uniform quantizer with a quantization range $[-e_{imax}, z]$ in DPCM system

   (b) if $|z| > e_{imax}$, then quantize $e_i(n)$ using the uniform quantizer with quantization range $[-e_{imax}, +e_{imax}]$ in DPCM system
4. If $i = 1$, stop; otherwise, set $i = i - 1$ and go back to step 2.

While the DPCM-BW method is adaptive quantization scheme, no additional information bits are needed. This is because of the fact that the quantization range used for encoding the ist LSP parameter is uniquely determined by the quantized value of ω_{i+1}, which is also available in the decoder.The idea used in the DPCM with backward sequence can be applied in the forward sequence. This method is called by the DPCM with forward sequence (DPCM-FW).

3. ADAPTIVE VECTOR-SCALAR HYBRID QUANTIZATION

The advantage of vector-scalar hybrid quantization(VQ-SQ) is that codebook size in vector quantization stage is much smaller than that for a single stage vector quantization with same number of bits. Thus, the complexity for codebook search, and memory size can be drastically reduced by using the vector-scalar hybrid quantizer of LSP parameters. The block diagram is shown in Fig. 3.

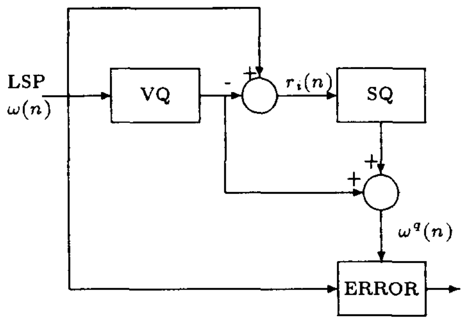

Figure 3. The Encoding Process in Vector-Scalar Quantizer of LSP parameters

First, input LSP parameters is 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 quantization, the components of residual vector are individually quantized by the scalar quantizer. Finally, the optimal VQ-SQ codevector which have a minimum distortion is selected among multiple VQ-SQ combinations.

The adaptive quantization method utilizing the ordering property of LSP parameters is applied in the vector-scalar hybrid quantization of LSP parameters. The improved vector-scalar hybrid quantization method is called as vector-scalar hybrid quantization with forward sequence(VQ-SQ-FW) or vector-scalar hybrid quantization with backward sequence(VQ-SQ-BW).

The checking variable in VQ-SQ-BW that decides the maximum quantization range in the scalar quantizer is given by

$$x = \omega^q_{i+1}(n) - [\omega_{pi}(n) + V_i(index_v, n)] \qquad (11)$$

where, $\omega^q_{i+1}(n)$ is the quantized value of the $(i+1)$th order LSP parameter, $\omega_{pi}(n)$ is the predicted value of the ith order LSP parameter in the scalar quantizer, and $V_i(index_v, n)$ is the ith order quantized output of first stage vector quantizer. 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 sum of ith element of vector quantizer output and predicted value of the ith LSP parameter. The nominal maximum quantization level of each parameter is given by r_{imax}. 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 have a minimum distortion is selected among multiple VQ-SQ combinations.

4. SIMULATION RESULTS

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

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

Table 2. The Performance of three Scalar Quantization Schemes

bit/ frm	AQBW SD dB	AQBW $>2dB$ %	DPCM SD dB	DPCM $>2dB$ %	DPCM-BW SD dB	DPCM-BW $>2dB$ %
27	1.45	4.6	1.59	12.6	1.32	2.6
29	1.26	2.4	1.42	6.3	1.17	1.7
31	1.18	2.0	1.21	2.0	1.02	1.0
33	1.06	1.8	1.07	1.4	0.92	0.8
36	0.86	1.6	0.89	1.1	0.77	0.7

Table 3. The SEGSNR (dB) Values of the QCELP Coder According to LSP Quantization Methods

	DPCM	DPCM-BW
8 kbps	14.84	15.11
Variable Rate	13.79	14.02

where NF is the number of total frames, and $A_n(e^{j\omega})$ and $\hat{A}_n(e^{j\omega})$ 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.

Table 2 shows the performance comparisons among AQBW, the DPCM scheme of the QCELP, and the proposed DPCM-BW at different bit rates in terms of the average spectral distortion and the number of outlier frames greater than 2 dB. The DPCM-BW algorithm which utilizes ordering property achieves 2-3 bits/frame saving over the DPCM and AQBW in terms of the SD. The DPCM-BW needs 31 bits/frame to maintain the transparent quality that the average distortion is less than 1 dB and the percent of frames with over 2 dB distortion is less than 2%. The proposed scalar quantizer was applied to the LSP quantization of the QCELP coder. The average SEGSNR value of the QCELP coder using the DPCM-BW at 40 bits/frame is about 0.3 dB higher than that of the QCELP coder using the DPCM at 40 bits/frame. In an informal listening test, the performance improvement was not evident because the QCELP coder assigned sufficient bits(40 bits/frame) for the LPC parameter transmission.

In the vector-scalar hybrid quantization, the codebook size of the first stage vector quantizer is given by 256(8 bits). The codebook of vector quantizer is designed by using the LBG algorithm [7] on the training data. The performance of vector-scalar quantizer improves as the number of candidate vectors in the first stage is increased. The performance of quantizer according to number of candidate vectors is shown in Table 4. Sixteen candidate vectors in the first stage are adequate. The performance comparisons between the conventional VQ-SQ and the adaptive VQ-SQ are shown in Table 5. The simulation result shows that the adaptive

Table 4. Performance of 27 bits/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 5. The SD (dB) Values in Vector-Scalar Quantization Schemes

bits/ frame	VQ-SQ		VQ-SQ-BW	
	SD	> 2 dB	SD	> 2 dB
	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

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 as low as 0.4 %.

5. EFFECT OF CHANNEL ERRORS

The performances of proposed quantization methods are evaluated in the presence of channel errors and the structure of quantizers is 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 predictor in DPCM structure of adaptive scalar quantizer is changed to moving average(MA) predictor from the AR predictor. The order of MA predictor is given by 3. The use of the MA predictor in proposed adaptive scalar quantizer showed the 0.02 dB performance degradation in spectral distortion in the error free channel. But the use of MA predictor provided 0.22 dB gain at bit error rate(BER) of 10^{-3} and 2.06 dB gain at BER of 10^{-2}.

In the adaptive vector-scalar quantizer, the choice of predictor structure did not affect the performance of quantizer in the noisy channel. The performances of DPCM-BW at 32 bits/frame and VQ-SQ-BW at 27 bits/frame in the presence of channel errors are shown in Table 6. The simulation result showed that the adaptive quantizers using the ordering property of LSP parameters performed as well as other robust LSP quantizer[8] in the presence of channel errors. The proposed adaptive quantization methods did not show the severe performance degradation at the BER below 10^{-3}

6. CONCLUSIONS

The DPCM quantization algorithm of the LSP parameters is improved by investigating the important ordering property of the LSP parameters. The simulation results show that the DPCM-BW provides about 3 bits/frame saving over the DPCM scheme without the extra increase of complexity. The DPCM-BW quantization algorithm needs 31 bits/frame to achieve the transparent quality of speech. Also, the new quantization method considering the ordering property of LSP parameters has been applied in the vector-scalar hybrid quantizer. The proposed vector-scalar quantizer achieves an average spectral distortion of 1 dB using 26 bits/frame.

Table 6. Channel Error Performance of DPCM-BW at 32 bits/frame and VQ-SQ-BW at 27 bits/frame

BER	DPCM-BW		VQ-SQ-BW	
	SD	> 2 dB	SD	> 2 dB
	dB	%	dB	%
0	0.98	0.8	0.95	0.3
10^{-4}	0.99	1.2	0.96	0.4
5×10^{-4}	1.05	3.0	0.99	1.0
10^{-3}	1.13	5.6	1.03	1.8
5×10^{-3}	1.66	22.0	1.34	8.1
10^{-2}	2.28	40.1	1.69	15.4

REFERENCES

[1] K. Paliwal and B. Atal,"Efficient vector quantization of LPC parameters at 24 bits/frame", *IEEE Trans. Speech and Audio Procs.*, vol.11, no.1, pp.3-14, Jan. 1993.

[2] 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.

[3] J. Grass and P. Kabal, "Methods of improving vector-scalar quantization of LPC coefficients", *in Proc. ICASSP-91,* Toronto, pp.657-660, 1991.

[4] 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. ICASSP-92,* San Francisco, pp. I.312-I.324, 1992.

[5] 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.

[6] QUALCOMM Inc., "TIA/EIA IS-96 Interim Standard; Speech service option standard for wideband spread spectrum digital cellular system", 1993.

[7] 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.

[8] 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.