Evaluation Performance of Speech Coder in Speech Signal Processing

Kwang-Seok Lee, *Member, KIMICS*

Abstract—We compared CS-ACELP with QCELP speech coder in CDMA cellular under channel error environment and experimented performance with its measured value under channel error environment. Also, we specified the effective coding scheme to overcome. CS-ACELP speech coder using a LSP vector quantizer shows transparent speech quality from the results that SD is 0.92dB and outlier frames over 2dB is 2.9% in the BER 0.10% condition. CS-ACELP speech coder which is utilizing MA predictor shows better results on SVR and SEGSNR than QCELP speech coder(IS-96) adopting DPCM type predictor when bit error occurs from BER 0.01% to 0.50%.

Index Terms—CDMA, QCELP, CS-ACELP, Speech Coder, Coding Scheme.

I. INTRODUCTION

The CDMA system uses variable rate Qualcomm Code Excited Prediction(QCELP) speech coder, the GSM system uses 13Kbps Regular Pulse Excited-Long Term Prediction (RPE-LTP) speech code and the TDMA system uses 8Kbps Vector Sum Excited Linear Prediction(VSELP) speech coder. It has been chosen as an interim standard of a speech coder(IS-96A) for the CDMA digital cellular system.[1][5]

In this research, we presented performance evaluation of a speech coder in channel error environments. As known well, fading, noise, and interference is inherent in the wireless communication system. This eventually degrade speech quality and annoy mobile subscribers. The robustness of a speech coder on channel error environment is known to one of the most important requirement for wireless communication system. Line Spectrum Pairs(LSP) quantizer gives a great contribution to the speech quality under channel error environment.

Therefore, the performance evaluation of a LSP quantizer used to be a standard measure to estimate speech quality. LSP quantization scheme can be classified into scalar quantization and Vector Quantization(VQ) method. QCELP uses typical scalar quantization scheme, comparing with that CS-ACELP utilizes VQ method. This method shows better results, however needs more memory for codebook and computational complexity to calculate distortion between input vector and codebook value. Performance compareison between OCELP and CS-ACELP speech coder is accomplished considering various channel error environment.

II. SPEECH CODING ALGORITHM

A. QCELP Algorithm

The QCELP speech coder is a variable rate speech coder applicable to CDMA digital cellular system. This variable rate encoding operates at rates of 8, 4, 2, or 1Kbps, depending on the level of voice activity. The QCELP algorithm is based upon the general CELP structure proposed by AT&T. That is, pitch and codebook parameters are selected by analysis-by-synthesis structure and codebook is utilized to quantize excitation signal. QCELP uses stochastic codebook which has 128 entries to minimize computational complexity and LSP which is suitable for interpolation and stability check.

The overall block diagram of encoder is shown in the figure 1. The input speech is sampled at 8KHz and is broken down 20ms speech frames consisting of 160 samples. The Linear Predictive Coding(LPC) with 10th order are calculated regardless of data rate selected. The LPC parameters are transformed into line LSP for more finely quantization, interpolation, and ease of stability check.

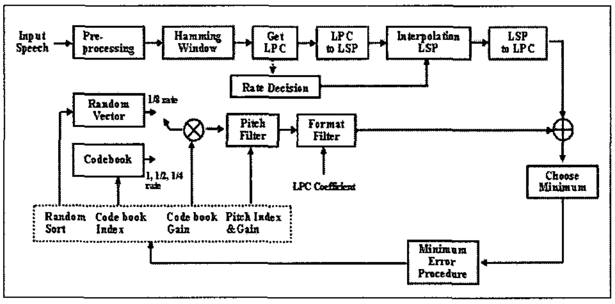

Fig. 1 Encoding procedure of QCELP algorithm

The LSPs are updated in each subframe through interpolation of LSPs in the neighbor frames. QCELP algorithm of IS-96A does not use Auto Regressive(AR) predictor which propagates previous frame's error to the future frames when channel error occurs. This algorithm has same bit assignment scheme with the IS-96 which depends on transmission rate however, does not utilize

Manuscript received April 13, 2007.

Kwang-seok Lee (phone: +82-55-751-3333, email: kslee@ jinju.ac.kr) is with the Department of Electronic Engineering, Jinju National University, Jinju, Korea.

predictor in the case of Rate 1. In the pitch search procedure, pitch gain and pitch lag are determined by standard analysis-by-synthesis error minimization procedures. The pitch lag is quantized from 17 to 143 sampled using 7 bits for each pitch update. The codebook gain and codebook index is determined once for each codebook update. The codebook is organized as an overlapping codebook such that each code vector differs from the adjacent code vector by one sample. The use of overlapping codebook provides several advantages such that the storage requirements for codebook of various rates can be reduced significantly and fast codebook search can be applied by using the dependency of the neighbor candidate code vector. Table 1 shows subframe size and bit assignment on every transmission rate.

B. CS-ACELP Algorithm

The 8Kbps CS-ACELP algorithm has recommended as G.729 for IMT-2000 at 1995 by ITU-T SG15(International Telecommunication Union-Telecommunication Standardization Sector).[2-3]

Table 1 Subframe size and bit assignment of QCELP

Parameter	Rate 1	Rate 1-2	Rate 1-4	Rate 1-8
frames for LPC bits/LPC update	160 40	160 20	160 10	160 10
pitch subframe bits/pitch update	40 10	80 10	160 10	- -
codebook subframe bits/codebook update	20 10	40 10	80 10	160 10

Figure 2 shows encoding procedure of the CS-ACELP algorithm. The frame size needed LPC analysis is 10ms consisting of 80 samples. It uses asymmetric hamming window that minimizes algorithmic delay and then determines LPC parameters. Once LPC coefficients selected, they are transformed into LSP parameters. The difference between LSP value of current frame and average LSP value of previous 4 frame's is quantized by using a 2-stage vector quantizer. In the first stage, vector quantization is accomplished on 10-dimensional 128 entries and split vector quantization is done on the 5-dimensional 32 entries in the second stage. Quantized LSP values are interpolated for pitch and codebook search.

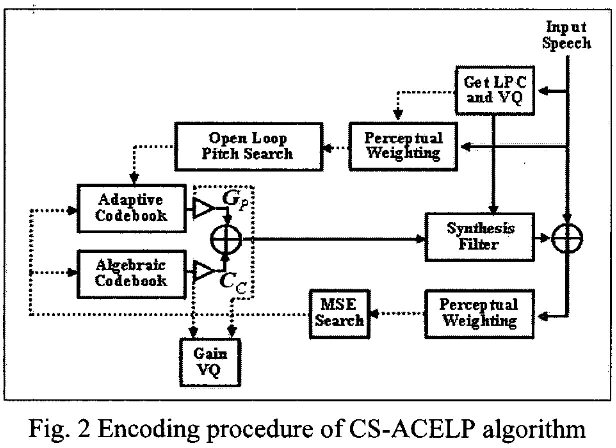

Fig. 2 Encoding procedure of CS-ACELP algorithm

After open-loop pitch analysis, fractional pitch search is executed sequentially on the smaller range around the selected pitch delay.

This pitch parameter search (adaptive-codebook search) is done every subframe(5ms). Codebook is composed with 4 pulses which values are +1 or -1. With the nature of algebraic structure of the codebook, search procedure can be done easily comparing with the conventional analysis-by-synthesis method. And this structure produces more candidate codebook, this eventually support better speech quality.

Pitch and codebook gain is quantized with the 2-stage conjugate structure vector quantizer.

Table 2 Bit assignment of the CS-ACELP

Parameter	Subframe1	Subframe2	Total
LSP	-	-	18
pitch delay	8	5	13
pitch parity	1		1
codebook index	13	13	26
codebook sign	4	4	8
codebook gain 1	3	3	6
codebook gain 2	4	4	8
total			80

III. PERFORMANCE COMPARISON UNDER CHANNEL ERROR ENVIRONMENT

As known well, fading, noise, and interference is inherent in the wireless communication system. This eventually degrade speech quality and annoy mobile subscribers. The robustness of a speech coder under channel error environment is known to be one of the most important requirement for wireless communication system. The performance of a speech coder can be evaluated through the quantization scheme of a transmission parameters. It is possible for speech coder to quantize LSPs through comparably low bit assignment by using correlation property between frames. However, this might propagates error and make speech quality worse because it utilizes quantization value of previous frames. To overcome this problem IS-96A does not use the AR predictor when quantize LSP values.

To evaluate the performance of the LSP quantizer average Spectral Distortion(SD) and its outlier frames over 2dB is generally used. It can be accepted as a "transparent speech quality", when the performance of the LSP quantizer shows average SD value is lower than 1dB and outlier frames over 2dB are around 2%. Table 3 shows performance comparison quantizer between QCELP and CS-ACELP on various bit error rates. QCELP and CS-ACELP shows approximately same results under channel environment in the LSP quantizer. CS-ACELP speech coder needs extra computational complexity, however, uses little bit assignment for LSP quantization comparing with the QCELP and maintains speech quality in BER 0.1%. For this simulation we used

50.30 second long speech data composed of 5 English sentences and 9 Korean sentences.

Table 4 shows simulation on the performance comparison between IS-96 and IS-96A by using Signal to Noise Ratio(SNR) and Segmental SNR(SEGSNR) which is known to be one of the objective measure. QCELP has I bit error recovery capability with CRC 11 bit on the transmission parameter. Without errors, IS-96 shows better results than IS-96A about 0.2dB in SEGSNR, however performance degradation is occurred under channel error environment.

Table 3 Performance comparison of the LSP quantizer between QCELP and CS-ACELP

(a) QCELP

BER(%)	IS-96		IS-96A	
	SD[dB]	Qutliers>(%) 2[dB]	SD[dB]	Qutliers>(%) 2[dB]
0.00	0.74	0.5	0.92	0.4
0.01	0.79	1.6	0.93	0.6
0.05	1.00	7.2	0.96	1.5
0.10	1.27	14.5	1.01	2.8
0.50	2.69	57.7	1.33	10.9

(b) CS-ACELP and Difference with IS-96

BER(%)	CS-ACELP		Difference
	SD[dB]	Qutliers>(%) 2[dB]	SD[dB]
0.00	0.85	1.3	0.11
0.01	0.86	1.5	0.07
0.05	0.89	1.8	-0.11
0.10	0.92	2.9	-0.35
0.50	1.27	9.1	-0.42

CS-ACELP speech coding algorithm uses quantization predictor adopting Moving Average(MA) model instead of DPCM type predictor in the quantization process of the LSP parameter, pitch, and codebook gain. That is, with quantized LSP values of previous 4 frames, current LSP value is estimated and then difference between calculated and estimated LSP value is quantized by VQ scheme. And also, codebook gain is quantized with the same method such as the LSP quantizer, current codebook gain is quantized using codebook gain of previous 4 subframes with MA model. This type of the quantization predictor minimizes errors effectively comparing with the DPCM predictor type which is used in the QCELP speech coder.

Table 4 Objective performance comparison between QCELP and CS-ACELP

(a) QCELP [dB]

BER(%)	IS-96		IS-96A	
	SNR	SEGSNR	SNR	SEGSNR
0.00	13.96	11.66	13.97	11.44
0.01	12.93	10.91	13.07	11.10
0.05	9.89	9.55	11.22	10.18
0.10	8.00	8.32	9.13	9.17
0.50	2.08	3.32	3.94	5.27
1.00	0.15	1.17	1.00	2.73

(b) CS-ACELP and Difference with IS-96 [dB]

BER(%)	CS-ACELP		Difference	
	SNR	SEGSNR	SNR	SEGSNR
0.00	16.38	15.49	2.42	3.83
0.01	15.85	15.20	3.12	4.29
0.05	13.71	14.35	3.82	4.80
0.10	11.81	13.33	3.81	5.01
0.50	5.74	8.76	3.66	5.44
1.00	3.13	5.44	2.98	4.27

In this simulation, CS-ACELP speech coder does not use error recovery function. Simulation results shows that CS-ACELP without recovery capability gives better results than QCELP(IS-96) which has that capability from BER 0.01% to 0.1%. That is, performance of the CS-ACELP speech coder becomes better in the case of error occurrence increases. However, in the case of BER 1.00%, it shows bad results, that means MA predictor is not effective any more over BER 1.00% occurrence environment.

IV. CONCLUSIONS

In this paper, we presented Objective performance comparison between QCELP and CS-ACELP speech coder under channel error environment. LSP quantizer scheme of the CS-ACELP speech coder adopts VQ method. Though it needs more memory for codebook and computational complexity to calculate distortions between input and codebook vectors, assigns less information bit for LSP quantizer and shows better results. CS-ACELP speech coder using a LSP vector quantizer shows transparent speech quality from the results that SD is 0.92dB and outlier frames over 2dB is 2.9% in the BER 0.10% condition.

CS-ACELP speech coder which is utilizing MA predictor shows better results on SVR and SEGSNR than QCELP speech coder(IS-96) adopting DPCM type predictor when bit error occurs from BER 0.01% to 0.50%. This clarifies the fact that speech coder which uses MA predictor gives more good results under channel environment when quantizes transmission parameters. However, it is impossible to expect performance improvement over BER 1.00% even though with the MA predictor. In further studies, channel error masking model considering real propagation condition should be generated instead of simple BER environment.

REFERENCES

[1] Noboru Sugamura and Nariman Farvardin, 2002, "Quantizer Design in LSP Speech Analysis-Synthesis", IEEE Journals on Selected Areas in Communications, Vol.6, No.2, pp.432 ~ 440.

[2] TIA/EIA IS-96, 2001, Speech Service Option Standard for Wideband Spread Spectrum Digital Cellular System.

[3] Redwan Salami, Claude Laflamme, Jean-Pierre Adoul and Dominique Massaloux, 1999, "A Toll Quality 8Kb/s Speech Codec for the Personal Communications System", IEEE Trans. on Vehicular Technology, Vol.43, No.3, pp.808 ~ 816.

[4] D. Lin, 1998, "Speech Coding using Pseudo Stochastic Block Codes", Proc. Int. Conf. Acoustics. Speech Signal Processing, pp. 1354 ~ 1357.

[5] M.R. Schroeder and B.S. Atal, 2000, "Code Exited Linear Prediction(CELP): High Quality Speech at Very Low Bit Rates", Proc. Int. Conf. Acoustics, Speech Signal Processing, pp. 937 ~ 940.

[6] Bishnu S. Atal, Vladmior Cuperman, Allen Gersho, 2001, Speech and Audio Coding for Wireless and Network Applications. Kluwer Academic Publishers.

[7] Shuzo Saito and Kazuo Kakata, 2001, Fundamentals of Speech Signal Processing. Academic Press.

Kwang-Seok Lee
Received B.S. and M.S. degrees of Electronic Engineering in 1983 and 1985 respectively, from Dong-A University. And Ph.D. from Dong-A University, in 1992. In 1995, he joined the Jinju National University in Jinju, Korea, where is currently a professor and his research interests are in Intelligent System, Neural Network, Speech Processing, Speech Recognition and Synthesis and Biometrics.