

Complexity Reduction Algorithm of Speech Coder(EVRC) for CDMA Digital Cellular System

So-Yeon Min[†]

ABSTRACT

The standard of evaluating function of speech coder for mobile telecommunication can be shown in channel capacity, noise immunity, encryption, complexity and encoding delay largely. This study is an algorithm to reduce complexity applying to CDMA(Code Division Multiple Access) mobile telecommunication system, which has a benefit of keeping the existing advantage of telecommunication quality and low transmission rate. This paper has an objective to reduce the computing complexity by controlling the frequency band nonuniform during the changing process of LSP(Line Spectrum Pairs) parameters from LPC(Line Predictive Coding) coefficients used for EVRC(Enhanced Variable-Rate Coder, IS-127) speech coders. Its experimental result showed that when comparing the speech coder applied by the proposed algorithm with the existing EVRC speech coder, it's decreased by 45% at average. Also, the values of LSP parameters, Synthetic speech signal and Spectrogram test result were obtained same as the existing method.

Keywords: EVRC Speech Coder, Complexity, LSP Parameter, CDMA System

1. INTRODUCTION

Voice compression means encoding analog voice signal into digital signal in comprehensive meaning, and as the feature of voice signal is not utilized in this case, it can not compress much amount[1,2]. The compression in detailed meaning is to encode analog voice signal into digital signal using the feature of voice signal, so it can compress much amount and if the input is not voice signal, its function is deteriorated[3].

CELP(Code-Excited Linear Prediction) voice compressor[4,5] used in mobile telecommunication system widely, is a method to use code book(vector table), where one vector is selected among many

entries of codebook and then it's used as $ex(n)$ signal after multiplying benefit to this, so this is a method to transmit only the index of selected vector as encoder and decoder has identical code book[6,7].

The standard voice compressor of domestic mobile telecommunication has 3 types as follows.

QCELP(Qualcomm CELP, IS-96A) is a method to enable us to obtain the output among nearby code-vectors by composing codebook in circular type, and this has a transmission rate of 8.55 kbps.

QCELP 13k(CDG-27) method is similar to QCELP but it has increased the voice quality just by raising bit rate and has 13.3kbps of transmission rate.

EVRC(Enhanced Variable Rate Codec) is the optimized technology of telephone voice quality that transmit only the desired sound with noise to be eliminated when the voice of caller is switched to digital. Also, this is the method to reduce the surrounding's noise with 8kbps speech coder standard stipulated in IS-127 and to transmit the

※ Corresponding Author : So-Yeon Min, Address : (131-702) 49-3, Myeonmok-8 Dong, Jungnang-Gu, Seoul, Korea, TEL : +82-2-490-7583, FAX : +82-2-490-7407, E-mail : symin@seoil.ac.kr

Receipt date : Oct. 8, 2007, Approval date : Nov. 19, 2007

[†] Dept. of Information and Communication

※ "This work was supported by the Korca Research Foundation Grant funded by the Korean Government (MOEHRD)"(KRF-2006-331-D00423)

speech coder. The method to be used mostly for this process is the real root method among the existing ones. Fig. 2 showed the root searching process of real root method. In other words, the root of odd number time is searched first and then, root of even number time is searched among the roots of odd number that were already searched in the process of seeking LSP root. When searching for the root of odd number time, it's searched with the frequency band to be divided into 800 equal parts one after the other. Therefore, the time to search for the root of odd number time comes to occupy the most of total searching time from LPC coefficients into LSP parameters.

In this study, when the LSP parameter is 10th, the frequency band is controlled non-uniformly according to the equation (1).

$$f_{n+1} = (10^{point(n)/(1000/\log 2)} - 1) * 1000 * 0.5 \quad (1)$$

for $0 \leq n < 399$

Equation (1) means,

$$f_0 = 0, \quad point(n) = index * (n+1), \quad \text{for } 0 \leq n < 399,$$

$$index = F_{mel} / 399.$$

Also, it is $F_{mel} = 1000 / \log 2 * \log [1 + FS / 1000]$,
 $FS = 8000.$

While the existing method has 800pcs of uniform interval of frequency band in order to seek for LSP parameters, the frequency band will be divided into non-uniformly 400pcs of searching interval from 0Hz to 4kHz when equation (1) is used in this study[8].

Equation (2) and equation (3) show the value of starting point f_n and value of ending point f_i about

$$Q_n'(f_n) = [A_0 \cos(2\pi f_n * \frac{p}{2}) + A_1 \cos(2\pi f_n * \frac{p-2}{2}) + \dots + \frac{1}{2} A_{p/2}] \quad \text{for } f_n : 0 \leq n < 399 \quad (2)$$

$$Q_i'(f_i) = [A_0 \cos(2\pi f_i * \frac{p}{2}) + A_1 \cos(2\pi f_i * \frac{p-2}{2}) + \dots + \frac{1}{2} A_{p/2}] \quad \text{for } f_i : i = n+1 \quad (3)$$

$$P_n'(f_n) = [A_0 \cos(2\pi f_n * \frac{p}{2}) + A_1 \cos(2\pi f_n * \frac{p-2}{2}) + \dots + \frac{1}{2} A_{p/2}] \quad \text{for } f_n = LSP(\psi_m), \quad m = 1, 3, 5, 7, 9 \quad (4)$$

$$P_i'(f_i) = [A_0 \cos(2\pi f_i * \frac{p}{2}) + A_1 \cos(2\pi f_i * \frac{p-2}{2}) + \dots + \frac{1}{2} A_{p/2}] \quad \text{for } f_i = LSP(\psi_{m+2}), \quad m = 1, 3, 5, 7, 9 \quad (5)$$

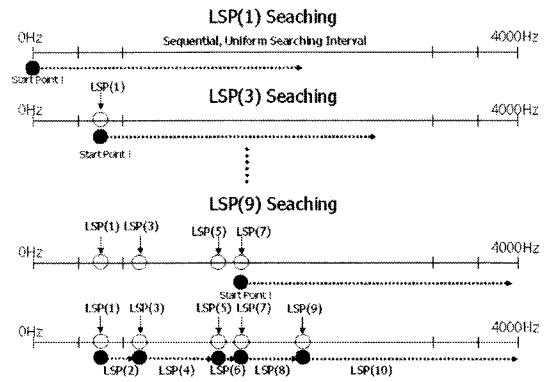


Fig. 2. LSP parameter searching process of real root method.

random searching interval after dividing the each searching interval into 400pcs[9,10]. Also, through the two equations above, 5pcs of LSP parameter of 1,3,5,7,9 number, the root of odd number time are obtained. Likewise, through the equation (4) and equation (5), 5 roots of even number time are searched among the roots of odd number time.

Through the equation (2),(3) and (4),(5), searching for LSP parameter passes through the same process. If we look into the process of searching LSP parameter of odd number time using equation (2),(3) in detail, it has the feature as follows;

In case of $Q_i(f_i) = 0$, f_i becomes the LSP parameter value to be searched. Also, in case of $Q_n(f_n) * Q_i(f_i) < 0$, it's determined that a root exist between f_n and f_i , and LSP parameter is searched passing through following process[11,12].

And, in case of $f_{mid} - f_n > intv_{thres}$, searching interval grows to be narrowed in order to search for LSP parameter. In short, if it is $Q(f_{mid}) = 0$ also in this case, f_{mid} becomes also the LSP parameter

value to be searched[13,14]. Also, in case of $Q(f_{mid}) * Q_n(f_n) < 0$, it's set like $f_i = f_{mid}, Q_i(f_i) = Q(f_{mid})$ and in case of $Q(f_{mid}) * Q_i(f_i) < 0$, it's set like $f_n = f_{mid}, Q_n(f_n) = Q(f_{mid})$. The next is to search for root passing through the equation (6). In short, if the searching band that was controlled un-equally, becomes bigger than the threshold value through the equation (6), it searches for root by increasing the resolution after adjusting the frequency interval again and adjusting into the range where the resolution does not become too low. Fig. 3 shows the whole block diagram of proposed algorithm as mentioned above. Also, fig. 4 represents the partial block diagram in case of odd order parameters and fig. 5 shows the partial block diagram in case of even order parameters as mentioned above.

$$f_{mid} = (f_n + f_i) / 2$$

$$\text{if } f_{mid} - f_n < \text{intv}_{thres}$$

$$LSP(\psi) = f_n + |Q_n(f_n)| / |Q_i(f_i) - Q_n(f_n)| * (f_i - f_n)$$

(6)

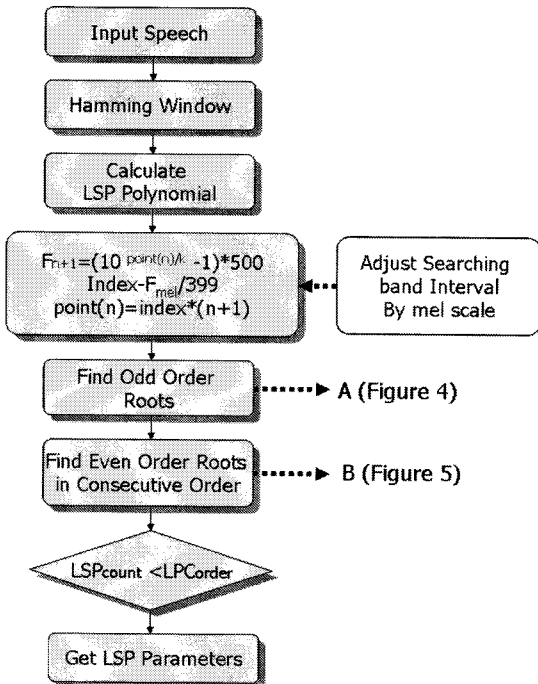


Fig. 3. Whole block diagram of the proposed algorithm.

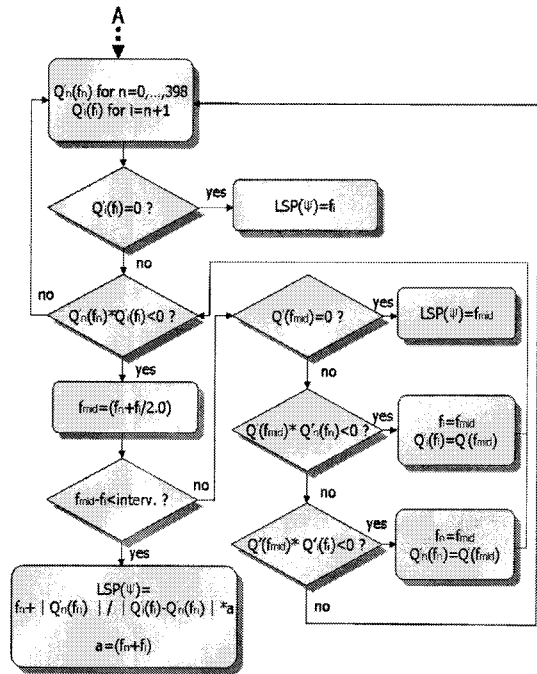


Fig. 4. Partial block diagram of the proposed algorithm. (In case of searching odd order parameter)

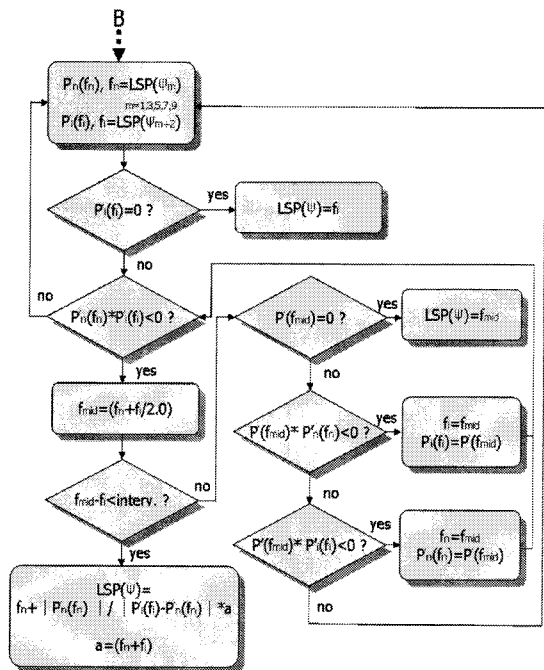


Fig. 5. Partial block diagram of the proposed algorithm. (In case of searching even order parameter)

4. EXPERIMENTAL RESULTS

We've interfaced A/D converter which enables the microphone input into PC for test. The proposed algorithm has the feature of a certain computing volume without being affected by feature change of the used voice sample. In order to prove this feature, we've used voice samples of 4 countries' languages. The simulation of the proposed algorithm was materialized in C language. The test was performed through 2 steps. In the first step, we've compared the computing volume and the distribution feature of LSP parameter in the changing process from LPC coefficients to LSP parameters between the real root method and the proposed method. 1 frame was processed as 256 sample in test process. Table 1 is the test result of comparing the real root algorithm extracted from EVRC vocoder with the computing volume per frame of the suggested algorithm in case that Korean voice sample and Japanese voice sample were used. In the test result about Korean voice sample, 41.3% was reduced in case of + and - computing, and 48.3% was decreased in case of * and / computing.

In the case Japanese voice sample was used, the computing of + and - was reduced by 41.3% and the computing of * and / was reduced by 47.8%. Table 2 is the result about computing volume per frame and computing volume reduction when using Chinese and English voice samples. In the test result about Chinese voice sample, the computing of + and - was decreased by 41.3% and the computing of * and / was reduced by 47.8%. In case of using English voice sample, the computing of + and - was decreased by 41.3% and the computing of * and / was reduced by 47.8%. As we find out from the test result, the proposed method compared to the existing algorithm showed toward various voice samples that the computing volume of '+,-' was decreased by about 41% and the computing volume of '*,/ ' was reduced by about 47%.

Table 1. Computing volume per frame and reducing rate. (In case of using Korean and Japanese)

unit: [times/frames]		Real Root Method	Proposed Method	Decreased Ratio [%]
Korean : 173 frames	+	15664	9217	41.5
	-			
	*	28671	14992	48.3
	/			
Japanese : 1835 frames	+	15665	9218	41.3
	-			
	*	28672	14993	47.8
	/			

Table 2. Computing volume per frame and reducing rate. (In case of using Chinese and English)

unit: [times/frames]		Real Root Method	Proposed Method	Decreased Ratio [%]
Chinese : 2428 frames	+	15665	9216	41.3
	-			
	*	28672	14991	47.8
	/			
English : 1835 frames	+	15667	9218	41.3
	-			
	*	28674	14993	47.8
	/			

Fig. 6 shows the distribution feature of LSP parameter about random frame.

In other words, (a) of fig. 6 shows the waveform of time domain toward voiced sound section. Index (b) represents the result of conducting frequency spectrum analysis about the same section. Index (c) and (d) show the LSP parameter obtained by the existing real root method and the LSP parameter obtained through the suggested method.

Through the result of fig. 6, we can learn that the parameter values obtained through the existing method and the suggested method are identical.

Table 3 corresponds to the location value of LSP parameter which indicated the result of (c) and (d) in number. Through the result of table 3, we can learn that the parameter values obtained through the existing method and the suggested method are identical.

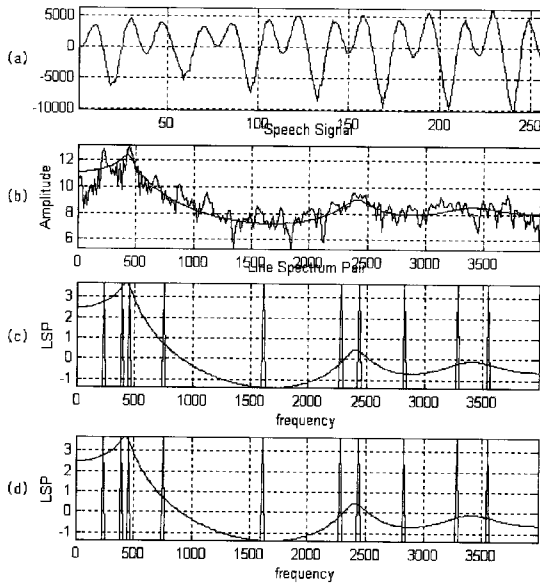


Fig. 6. Distribution feature of LSP parameter. (a) Waveform of time domain toward voiced sound section. (b) Result of conducting frequency spectrum analysis. (c) LSP parameters obtained by the existing real root method. (d) LSP parameters obtained by the proposed method.

In the second step, we've compared the type of synthetic speech of encoding device, spectrogram and SNR(Signal to Noise Ratio) analysis result in case of applying the proposed algorithm to EVRC speech coder.

Fig. 7 shows the comparison about synthetic speech obtained from output after inputting the male/female announcers selected from weather forecasting news into encoding device. Index (a) of fig. 7 represents the original voice signal and (b) shows the waveform of synthetic speech of

EVRC encoding device. Also, (c) of fig. 7 represents the waveform of synthetic speech of encoding device applied with the proposed method. At the result of analyzing the synthetic speech of encoding device applied with the suggest method, we've learned that there's no inferior part to the synthetic speech of EVRC encoding device.

Fig. 8 shows the result of performing spectrogram analysis of the entire voice. Index (a) of fig. 8 represents the original voice signal of time domain and (b) shows the result of spectrogram analysis of original voice signal. Index (c) and (d) of fig. 8 show the spectrogram result of synthetic speech of the existing EVRC speech coder and synthetic speech of the encoding device applied

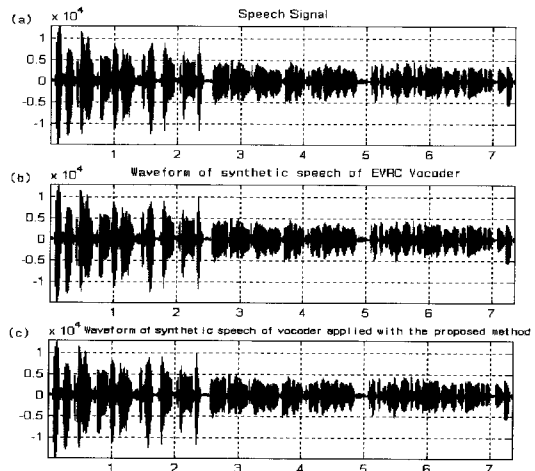


Fig. 7. Comparison of synthetic speech of encoding device. (a) Original voice signal. (b) Waveform of synthetic speech of EVRC speech coder. (c) Waveform of synthetic speech of encoding device applied with the proposed method.

Table 3. LSP parameter values obtained through the proposed method

	LSP Parameter values									
	LSP (1)	LSP (2)	LSP (3)	LSP (4)	LSP (5)	LSP (6)	LSP (7)	LSP (8)	LSP (9)	LSP (10)
Real Root Method	0.0278	0.0408	0.0565	0.0915	0.2053	0.2950	0.3245	0.3895	0.4279	0.4520
Proposed Method	0.0278	0.0408	0.0565	0.0915	0.2053	0.2950	0.3245	0.3895	0.4279	0.4520

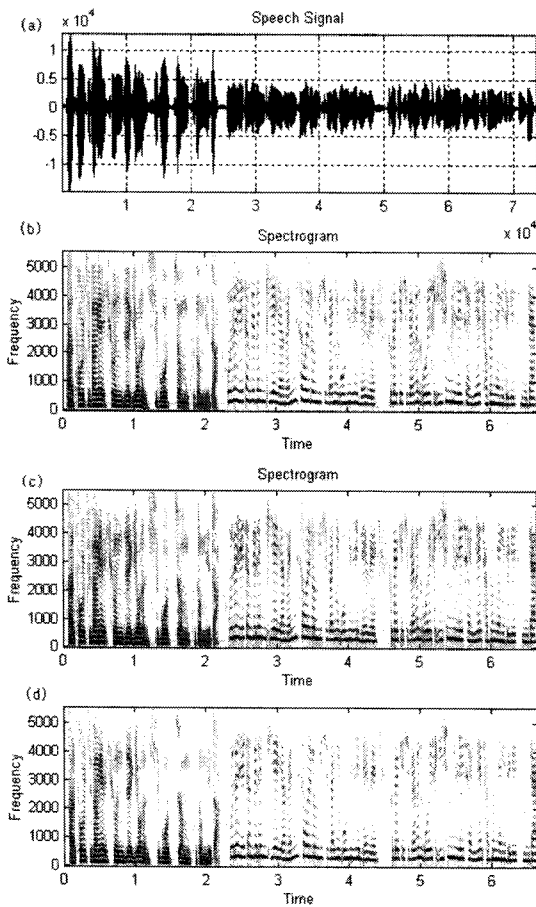


Fig. 8. Spectrogram comparison of entire voice
 (a) Original voice signal.
 (b) Spectrogram of original signal.
 (c) Spectrogram analysis result of synthetic speech passed through existing EVRC speech coder.
 (d) Spectrogram analysis result of synthetic speech of the encoding device applied with the suggested method.

with the suggest methods. From the result of spectrogram analysis, we learn that the case of using the suggested method brought about the same test result as the existing test result.

Table 4 represents the result of performing SNR(Signal to Noise Ratio) analysis of the entire speech signal. From the result of SNR analysis, we obtain that the case of using the suggested method brought about the same test result as the existing test result.

Table 4. SNR analysis of synthetic speech signal

	Real Root Method	Proposed Method
Korean: 173frames	13.66[dB]	13.67[dB]
Japanese: 1835 frames	14.61[dB]	14.61[dB]
Chinese: 2428frames	13.31[dB]	13.33[dB]
English: 1835 frames	14.56[dB]	14.58[dB]

5. CONCLUSION

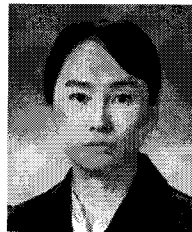
This paper is an algorithm to reduce complexity of speech coder for CDMA mobile telecommunication system(EVRC), which can be used for various speech coder and analyzer, etc that use LSP parameter because it's developed in independent form, not subordinate to a specific analyzer.

In short, this paper is a method to reduce the complexity by searching for frequency band divided unequally to search for the root, differentiated from the existing method of searching for the root dividing frequency band in equal interval. If searching interval becomes wider than the threshold value during passing through this process, inaccurate root will be searched, but because it's searching for the root by narrowing the searching frequency interval gradually in this paper, the same root as the existing methods could be obtained. At the result of test, we could achieve the same result as the existing methods in the aspect of sound quality while the computing volume was reduced by 45% at average.

REFERENCES

- [1] TIA/EIA IS-96, *Speech service option standard for wideband spread spectrum digital cellular system*, April 1994.
- [2] Qualcomm Inc., *High rate speech service option for wideband spread spectrum communications systems*, Feb. 22 1995.

- [3] G. Schroeder, "The standardization process for the proposed 8kbps speech coder standard," *IEEE Speech Coding Workshop*, Annapolis, 1995.
- [4] ITU-T Recommendation G.729, *Coding of Speech at 8kbit/s CS-ACELP Speech Coder*, 1996.
- [5] ITU-T Recommendation G.729 Annex A, *Reduced Complexity 8kbit/s CS-ACELP Speech Codec*, 1996.
- [6] ITU-T Recommendation G.723.1, March, 1996.
- [7] Stev Ammon and Jianxin Liu, "Implementation of the ITU-T G.723.1 Speech Coding Algorithm on the Hitachi SH-DSP," *ICSPAT* 1995.
- [8] Thomas Parson, *Voice and Speech Processing*, Mac Graw Hill, pp. 71-73, 1986.
- [9] F. Soong and B. H. Juang, "Line Spectrum pairs and speech data compression," *Proc. of ICASSP*, 1.10.1-1.10.4, 1984.
- [10] N. Jayant and P. Noll, *Digital Coding of Waveforms : Principles and Applications to Speech and Video, Signal Processing*, Prentice-Hall, pp. 221-220, 1984.
- [11] S. Saito and K. Nakata, *Fundamentals of Speech Signal*, Academic Press, pp. 126-132, 1985.
- [12] A. M. Kondoz, *Digital Speech*, John Wiley & Sons Ltd, pp. 84-92, 1994.
- [13] M. J. Bae, *Digital Speech Analysis*, Dongyoung Press, pp. 95-120, 1998.
- [14] P. Kabal and R. P. Ramachandran, "The computation of line spectral frequencies using Chebyshev polynomials," *IEEE Trans. on ASSP*, December 1986.



So-Yeon Min

So-Yeon Min received the B.S., M.S., and Ph.D. degree in electronics engineering from Soongsil University in 1993, 1995, and 2003 respectively. In 2006, she joined the Seoil College, where she is an assistant professor. Her research field has been in speech signal processing, data communication, communication system.