THE JOURNAL OF THE ACOUSTICAL SOCIETY OF KOREA Vol.20, No.1E 2001. 3. pp. 20~24

A High-Speed LSF Transformation Algorithm for CELP Vocoders

SoYeon Min*, MyungJin Bae**

*Dept. of Electronics Engr, Soongsil University
**Dept. of Information & Telecommunication Engr, Soongsil University
(Received 30 October 2000 ; accepted 29 November 2000)

Abstract

We propose the computation reduction method of real root method that is mainly used in the CELP(Code Excited Linear Prediction) vocoder. The real root method is that if polynomial equations have the real roots, we are able to find those and transform them into LSF[1]. However, this method takes much time to compute, because the root searching is processed sequentially in frequency region. But, the important characteristic of LSF is that most of coefficients are occurred in specific frequency region. So, the searching frequency region is ordered by each coefficient's distribution. And coefficients are searched in ordered frequency region.

Transformation time can be reduced by this method than the sequential searching method in frequency region. When we compare this proposed method with the conventional real root method, the experimental result is that the searching time was reduced about 46% in average.

Keywords *: LSP, High speed LSF transformation, CELP vocoders*

I. Introduction

Linear predictive coding(LPC) is very powerful analysis technique, and is used in many speech processing system. For speech coding and synthesis system, apart from the different analysis techniques available to obtain the LPC parameters, auto correlation, covariance, lattice, etc., the quantization of the LPC parameters is also a very important aspect of the LPC analysis, as minimization of coding capacity is the ultimate aim in these applications. The main objective of the quantization procedure is to code the LPC parameters with as few bits as possible without introducing additional spectral distortion. Whilst perfect reconstruction is not possible, subjective transparency is achievable[3].

Considerable amounts of work on scalar and vector LPC quantizers have already been reported in the past, but these have been predominantly directed at coding schemes operating above 9.6kb/s (APC, RELP, etc.) or at very low rate vocoders, less than 4.8kb/s. Thus these have tended to be good quality but high capacity schemes, e.g. 40-50 bits scalar quantization, or low capacity but only reasonable quality vector quantization schemes, e.g. 10-bit codebook vector quantization. Therefore, for medium to low bit rates, i.e. 9.6-4.8kb/s, the previously reported LPC quantization schemes are not directly applicable, hence further investigations are required. A promising and popular method is the use of the line spectrum pairs(LSP)(related

Corresponding author: Soyeon Min(pasternak@hanmail.net)
Soongsil University, Seoul 156-743, Korea

to line spectral frequency(LSF)) representation of the LPC parameters[2][8]. LSF(Line Spectrum Frequency) is used for speech analysis in vocoders or recognizers since it has advantages of constant spectrum sensitivity, low spectrum distortion and easy linear interpolation[1]. However the method of transforming LPC(Linear Predictive Coding) coefficients into LSF is so complex that it takes much time to compute[2][3].

In order to acquire LSP, the process of finding the roots of polynomial equations is implemented. The conventional methods are complex root, real root, ratio filter, Chebyshev series, and adaptive sequential LMS(Least Mean Square) methods[3]. Among these methods, the real root method is considerably simpler than others, but nevertheless, it still suffer from its indeterministic computation time. In this paper, we propose the computation reduction method of real root method using the distribution of LSP and formant characteristics.

II. LPC to LSP Transformation

An all-pole digital filter for speech synthesis, H(z), can be derived form linear predictive analysis, and is given by

$$H(z) = 1/A_p(z) \tag{2.1}$$

where

$$A_p(z) = 1 + \sum_{k=1}^{p} \alpha_k z^{-k} \tag{2.2}$$

The PARCOR system is an equivalent representation, and its digital form is as shown in Figure 1,

where

$$A_{p-1}(z) = A_p(z) + k_p B_{p-1}(z) \tag{2.3}$$

$$B_p(z) = z^{-1}[B_{p-1}(z) - k_p A_{p-1}(z)] \tag{2.4}$$

where

$A_0(z) = 1$ and $B_0(z) = z^{-1}$, and

$$B_p(z) = z^{-(p+1)} A_p(z) \tag{2.5}$$

The PARCOR system as shown in Figure 1, is stable for $|k_i| < 1$ for all i. The PARCOR synthesis process can be viewed as sound wave propagation through a lossless acoustic tube, consisting of p sections of equal length but

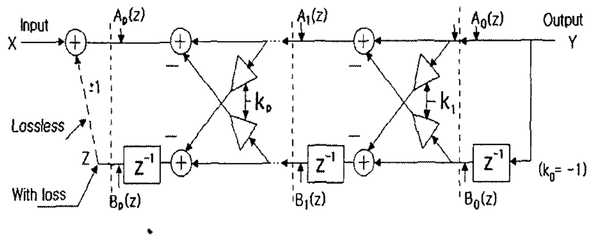

Figure 1. PARCOR structure of LPC synthesis.

non-uniform cross sections. The acoustic tube is open at the terminal corresponding to the lips, and each section is numbered from the lips. Mismatching between the adjacent sections p and $(p+1)$ causes wave propagation reflection. The reflection coefficients are equal to the $P_{\_th}$ PARCOR coefficient k_p section $p+1$, which corresponds to the glottis, is terminated by a matched impedance. The excitation signal applied to the glottis drives the acoustic tube.

In PARCOR analysis, the boundary condition at the glottis is impedance matched. Now consider a pair of artificial boundary conditions where the acoustic tube is completely closed or open at the glottis. These conditions correspond to $k_{p+1} = 1$ and $k_{p+1} = -1$, a pair of extreme values for the artificially extended PARCOR coefficients, which corresponds to perfectly lossless tubes. The value Q of each resonance becomes infinite and the spectrum of distributed energy is concentrated in several line spectra. The feedback conditions $k_{p+1} = -1$ correspond to a perfect closure at the input (glottis) and for $k_{p+1} = -1$ correspond to an opening to infinite free space[3].

To derive the line spectra or line spectrum frequencies (LSF), we proceed as follows, where it is assumed that the PAECOR filter is stable and the order is even. $A_p(z)$may be decomposed to a set of two transfer functions, one having an even symmetry, and the other having an odd symmetry. This can be accomplished by taking a difference and sum between $A_p(z)$ and its conjugate functions.Hence the transfer functions with $k_{p+1} = +1$:and denoted by $P_{p+1}(z)$ and $Q_{p+1}(z)$:

$$For\ \ k_{p+1} = 1,\ P_{p+1}(z) = A_p(z) - B_p(z) \tag{2.6}$$

$$For\ \ k_{p+1} = -1,\ Q_{p+1}(z) = A_p(z) + B_p(z) \tag{2.7}$$

$$\rightarrow A_p(z) = \frac{1}{2}[P_{p-1}(z) + Q_{p+1}(z)] \tag{2.8}$$

Substituting equation (2.5) into (2.6)

$$P_{p+1}(z) = A_p(z) + z^{-(p+1)} A_p(z^{-1}) \tag{2.9}$$
$$= 1 + (a_1 - a_p) z^{-1} + \cdots + (a_p - a_1) z^{-p} - z^{-(p+1)}$$
$$= z^{-(p+1)} \prod_{i=0}^{p=1} (z + a_i)$$

Similarly,

$$Q_{p+1}(z) = z^{-(p+1)} \prod_{i=0}^{p=1} (z + b_i) \tag{2.10}$$

As we know that two roots exits ($k_{p+1} = \pm 1$), the order of $P_{p+1}(z)$ and $Q_{p+1}(z)$ can be reduced,

$$P'(z) = \frac{P_{p+1}(z)}{(1-z)} = A_0 z^p + A_1 z^{(p-1)} + \cdots A_p \tag{2.11}$$

and

$$Q'(z) = \frac{Q_{p+1}(z)}{(1+z)} = B_0 z^p + B_1 z^{(p-1)} + \cdots B_p \tag{2.12}$$

where $A_0 = 1, B_0 = 1$

$$A_k = (a_k - a_{p+1-k}) + A_{k-1} ,$$
$$B_k = (a_k + a_{p+1-k}) + B_{k-1}$$
$$\text{for } k = 1, \cdots, p \tag{2.13}$$

The LSF's are the a ngular of the roots of $P'(z)$ and $Q'(z)$ with $0 \leq \omega_i \leq \pi$.

Real root method

As the coefficients of $P'(z)$ and $Q'(z)$ are symmetrical, the order of equation (2.11) can be reduced to *p/2*.

$$P'(z) = A_0 z^p + A_1 z^{(p-1)} + \cdots + A_1 z^1 + A_0 \tag{2.14}$$
$$= z^{p/2} [A_0 (z^{p/2} + z^{-p/2}) + A_1 (z^{(p/2-1)} + z^{-(p/2-1)})$$
$$+ \cdots + A_{p/2}]$$

Similarly,

$$Q'(z) = B_0 z^p + B_1 z^{(p-1)} + \cdots + B_1 z^1 + B_0 \tag{2.15}$$
$$= z^{p/2} [B_0 (z^{p/2} + z^{-p/2}) + B_1 (z^{(p/2-1)} + z^{-(p/2-1)})$$
$$+ \cdots + B_{p/2}]$$

As all roots are on the unit circle, we can evaluate equation (2.14) on the unit circle only

$$\text{Let} \quad Z = e^{w} \quad \text{then} \quad z^1 + z^{-1} = 2\cos(\omega) \tag{2.16}$$
$$P'(z) = 2e^{jp\omega/2} [A_0 \cos(\frac{p}{2}\omega) + A_1 \cos(\frac{p-2}{2}\omega)$$
$$+ \ldots + \frac{1}{2} A_{p/2}] \tag{2.17}$$
$$Q'(z) = 2e^{jp\omega/2} [B_0 \cos(\frac{p}{2}\omega) + B_1 \cos(\frac{p-2}{2}\omega)$$
$$+ \ldots + \frac{1}{2} B_{p/2}] \tag{2.18}$$

By making the substitution $x = \cos(\omega)$, equation (2.16) and (2.17) can be solved for x. For example, with p=10, the following is obtained:

$$P'_{10}(x) = 16A_0 x^5 + 8A_1 x^4 + (4A_2 - 20A_0)x^3 + (2A_3 - 8A_1)x^2$$
$$+ (4A_0 - 3A_2 + A_4)x + (A_1 - A_3 + 0.5A_5) \tag{2.19}$$

and similarly,

$$Q'_{10}(x) = 16B_0 x^5 + 8B_1 x^4 + (4B_2 - 20B_0)x^3 + (2B_3 - 8B_1)x^2$$
$$+ (4B_0 - 3B_2 + B_4)x + (B_1 - B_3 + 0.5B_5) \tag{2.20}$$

The LSFs are then given by

$$LSF(i) = \frac{\cos^{-1}(x_i)}{2\pi T}, \quad \text{for } 1 \leq i \leq p \tag{2.21}$$

This method is obviously considerably simpler than others, but nevertheless, it still suffer from its indeterministic computation time[5][6][7].

III. The High-Speed LSF Transformation Algorithm

In the real root method, odd order LSF parameters are searched first and then even order parameters are searched between odd order parameters. The searching time of odd order parameters take up most transformation time because the searching is processed sequentially in the whole of frequency region. But the important characteristic of LSF is that most LSF parameters are occurred in specific frequency region. So, to reduce the computation time in real root method, the searching frequency region is put in order by each coefficient's distribution. And, odd order coefficients are searched in order of the searching frequency order.

Figure 2 shows the distribution of odd order LSF parameters that is calculated by using the 8kHz sampled speech signal and tenth linear predictive analysis. In Figure

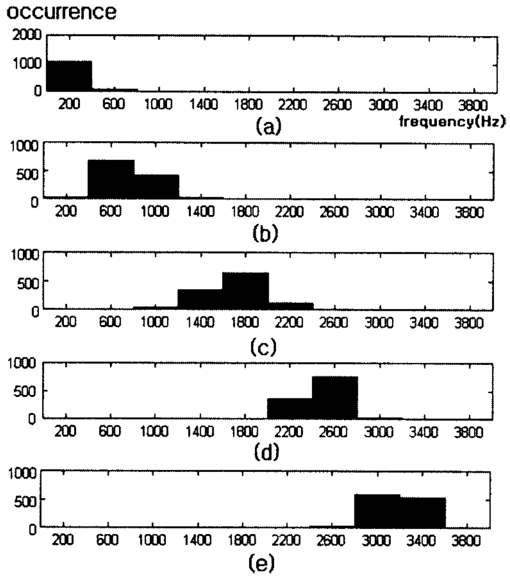

Figure 2. The Distribution of Odd Order in LSFs (p=10).
(a) First Order, (b) Third Order, (c) Fifth Order, (d) Seventh Order, (e) Ninth Order.

2, odd order LSF parameters are occurred in specific frequency region and not occurred in the others region. In the real root method, odd order LSF parameters are searched in frequency region sequentially. But the distribution of LSF in (c), (d) and (e) in Figure 2, is not sequential. So, if we will search in many occurred frequency region first, we will be able to reduce the computation time of LPC to LSF parameters transformation.

IV. Experimental Result

Computer simulation was performed to evaluate the proposed algorithm using an IBM PC(300MHz) interfaced with the 16-bit AD/DA converter. To measure the performance of the proposed algorithm, we used the following speech data. Speech data was sampled at 8kHz and was quantized at 16bits. Following sentences were uttered five times by 5 male and female speakers five times who are in the middle or later twenties. The data were recorded in a quiet room, with the SNR(Signal to Noise Ratio)greater than 30dB.

Table 2. The Computation Time of LSF Transformation.

	real root method (sec)	proposed method (sec)	decreased ratio (%)
sentence(1)	1.59	0.84	47.17
sentence(2)	2.35	1.21	48.5
sentence(3)	1.95	1.07	45.13
sentence(4)	2.2	1.22	44.55

1. /Insune komaneun cheonjaesonyuneul joahanda/
2. /Yesunimkeoseo cheonjichangjoeu gyohuneul malseumhasyuda/
3. /Soongsildae jeongbotongshingonghakkwa eumseiongtongshin yeungusilida/
4. /Changgongeul hechye naganeun inganeu dojeoneun keuchieobda/

The real root algorithm used in this paper is extracted from G.723.1 ACELP Vocoder(G.723.1 Annex A) implemented in C-language[4].

Table 2 shows the computation times of the conventional real root method and the proposed method. The proposed LPC to LSF transformation algorithm is faster than conventional real root algorithm and the computation time of the proposed method is reduced by 46% in average. But the transformed LSF parameters of the proposed method was the same as those of conventional real root method. That is, computation time is reduced by 46% in average with the same results.

Table 1. The Searching Frequency Order.

LSF \ Searching Order	1	2	3	4	5
LSF(1)	0-400(Hz)	400-800(Hz)	Otherwise		
LSF(3)	400-800(Hz)	800-1200(Hz)	Otherwise		
LSF(5)	1600-2000(Hz)	1200-1600(Hz)	2000-2400(Hz)	800-1200(Hz)	Otherwise
LSF(7)	2400-2800(Hz)	2000-2400(Hz)	2800-3200(Hz)	Otherwise	
LSF(9)	2800-3200(Hz)	3200-3600(Hz)	Otherwise		

V. Conclusion

LSF parameter is used for speech analysis in low-bit rate speech vocoders or recognizers since it has advantages of constant spectrum sensitivity, low spectrum distortion and easy linear interpolation. However the method of transforming LPC(Linear Predictive Coding) coefficients into LSF is so complex that it takes much time to compute. In this paper, we proposed the new transformation algorithm based on the real root algorithm that which widely used in the CELP vocoder. The real root method is simpler than other transformation methods but this takes much time to compute, because the root searching is processed sequentially in frequency region. So, to reduce the computation time of real root, we used the characteristic of LSF that most coefficients occurrs in specific frequency region. Firstly, we found the distribution of LSF parameters and put in order the LSF searching frequency region. Transformation time can be reduced by this method than the sequential searching of real root method. The experimental result is that the searching time was reduced by about 46% in average without the change of LSF parameters.

References

1. M. J. Bae, "Digital Speech Analysis",pp.95-120,Dongyoun, 4, 1998.
2. S. Saito and K. Nakata," Fundamentals of Speech Signal", pp.126-132, Academic Pressl, 1985.
3. A. M. Kondoz, "Digital Speech", pp. 84-92, John Wiley & sons Ltd, 1994.
4. ITU-T Recommendation G.723.1, March, 1996.
5. N. Jayant and P. Noll, "Digital Coding of Waveforms:Principles and Applications to Speech and Video, pp 221-220, Prentice-Hall Signal Processing, Prentice-Hall, 1984.
6. G. Kang and C. J. Francen, "Low Bit Rate Speech Encoder based on Line Spectrum Frequencies(LSF)", Technical Report8857, U.S. Naval Lab. Report, November 1984.
7. P. Kabal and R. P. Ramachandran, "The computation of line spectral frequencies using Chebyshev polynomials", IEEE Trans. on ASSP, 1419-1425, December 1986.
8. F. Soong and B. H. Juang, "Line Spectrum pairs and speech data compression", Proc. of ICASSP,1.10.1-1.10.4, 1984.

[Profile]

• **Soyeon Min**

Soyeon Min received the BS and MS degrees in Electronic Dept., from Soongsil University in 1994. Now she is working in the ph. D course on Electronic Dept., at Soongsil University. Her current research interests are Speech Coding and Synthesis.

• **Myungjin Bae**

Prof. Myungjin Bae, ph. D., finished his academic discipline at the Seoul National University in 1987.
He has served as a professor of Dept., of Information and Telecommunication Engineering in Soongsil University. He is a member of Acoustical Society of Korea. His research fields are Speech Coding and Synthesis, Speaker Recognition, Digital Signal Processing and Digital Communication.