

LSP 파라미터 분포특성을 이용한 주파수대역 조절법에 관한 연구

A Study on the Frequency Scaling Methods Using LSP Parameters Distribution Characteristics

민 소 연*, 배 명 진**
(So-Yeon MIN*, Myung-Jin BAE**)

* 숭실대학교 전자공학과, ** 숭실대학교 정보통신전자공학부
(접수일자: 2001년 11월 29일; 채택일자: 2002년 2월 28일)

LSP (Line Spectrum Pairs) 파라미터는 음성코덱 (codec)이나 인식기에서 음성신호를 분석하여 전송형이나 저장형 파라미터로 변환되어, 주로 저전송률 음성부호화에 사용된다. 그러나 LPC (Linear Predictive Coding) 계수를 LSP로 변환하는 방법이 복잡하여 계산시간이 많이 소요된다는 단점이 있다. 기존의 LSP 변환 방법 중 음성 부호화기에서 주로 사용하는 실근 (real root) 방법은 근을 구하기 위해 주파수 영역을 순차적으로 검색하기 때문에 계산시간이 많이 소요되는 단점을 갖는다. 본 논문에서 기존의 실근 방법과 비교 평가한 알고리즘은 첫 번째 검색 대역에 멜 스케일 (mel scale)을 사용하였고, 두 번째는 LSP 파라미터의 분포 특성을 조사하여 이를 토대로 검색구간의 순서와 검색간격을 달리하였다. 실험결과, 기존의 실근 방식에 비하여 두 가지 방식 모두가 변환시간의 47% 이상이 감소되는데 반하여 동일한 근을 찾음을 알 수가 있었다.

핵심용어: 실근방법, LSP 파라미터 분포특성, 멜 스케일

투고분야: 음성처리 분야 (2.2)

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 LSP. However, this method takes much time to compute, because the root searching is processed sequentially in frequency region. In this paper, to reduce the computation time of real root, we compare the real root method with two methods. In first method, we use the mel scale of searching frequency region that is linear below 1 kHz and logarithmic above. In second method, The searching frequency region and searching interval are ordered by each coefficient's distribution. In order to compare real root method with proposed methods, we measured the following two. First, we compared the position of transformed LSP (Line Spectrum Pairs) parameters in the proposed methods with these of real root method. Second, we measured how long computation time is reduced. The experimental results of both methods that the searching time was reduced by about 47% in average without the change of LSP parameters.

Keywords: Real root method, LSP (Line spectrum pairs) distribution characteristics, Mel scale

ASK subject classification: Speech signal processing (2.2)

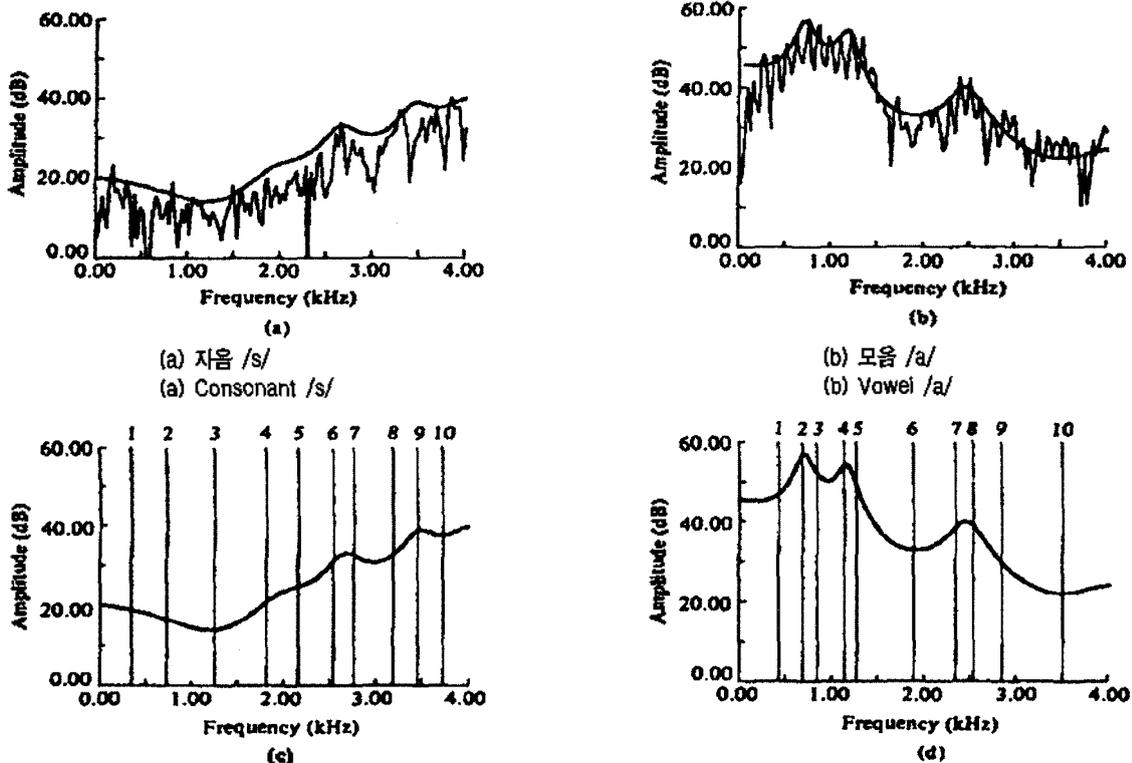
I. 서론

LSP 파라미터는 양자화 에러에 강하고 시스템의 안정성과 선형 보간성이 뛰어난 장점을 갖는다. 그러므로 음성 부호화기에서 음성신호를 분석하여 전송형이나 저장형 파라미터로 변환하는데 LSP파라미터가 사용된다. LPC (Linear Predictive Coding) 계수를 LSP 파라미터로 변환하기 위한 다항식의 근을 찾는 방법에는 복소근, 실근, 비울 필터, 체비셰프 급수, 적응적 순차형 최소제곱평균 (adaptive sequential LMS) 방법 등이 있다. 이 중에서 실근 방법이 간단하고 이해하기 쉽기 때문에 가장 많이 이용된다. 실근 방법의 특징은 홀수번째 근을 먼저 찾은 후에, 찾아진 홀수번째 근 사이에서 짝수번째 근을 찾게 된다. 그러므로 전체적인 검색시간은 홀수번째 근을 찾는 시간에 의해 좌우가 된다[1, 2]. LSP 홀수번째 계수들의 분포특성을 관찰한 결과 특정 주파수 영역에서 주로 나타나고 나머지 영역에서는 거의 나타나지 않았다. 본 논문에서는 LSP 근을 찾기 위한 몇 가지 방법들의 성능평가에 초점을 맞추어 보았다. 본 논문에서 고려한 첫 번째 방법은 분포특성을 이용하여 검색순서와 간격을 조절한

방법이고, 멜 스케일을 이용하여 주파수 대역을 검색한 방법이 두 번째 방법에 해당된다. 본 논문의 구성은 다음과 같다. 2장에서는 LPC와 LSP 파라미터의 일반적인 특징에 대하여 살펴본다. 3장에서는 실근 방법과 본 논문에서 고려한 두 가지 방법의 특징에 대하여 설명한다. 그리고 4장에서는 실험 결과를 분석하고, 5장에서는 결론을 맺는다.

II. LPC 파라미터의 특징

LSP 파라미터는 음성 신호의 포먼트 스펙트럼 정보를 반영한다. 유성음의 제 1 포먼트 (F_1)는 비교적 낮은 주파수 대역에서 좁은 대역폭을 갖는다. LSP 파라미터에서 선쌍 (line pair)의 간격이 좁으면, 강한 공명이 일어난 것을 나타낸다. 즉, LSP는 p 개의 불연속적인 주파수의 분포를 통하여 음성의 스펙트럼 포락선을 표현한다. 그림 1 (d)의 2, 3번째와 4, 5번째는 제 1포먼트와 제 2포먼트를 나타낸다. 따라서 포먼트의 위치 및 분포 특성은 LSP 파라미터에 직접적인 영향을 미치게 된다[1-3, 7].



(c), (d) /s/와 /a/에 대한 LPC 분석과 LSP
 (c), (d) The analysis of LPC and LSP for /s/ and /a/

그림 1. LSP 파라미터
 Fig. 1. The LSP parameter.

또한 PARCOR 구조에서 $k_{p+1} = \pm 1$ 인 전달함수를 $P_{p+1}(z)$ 와 $Q_{p+1}(z)$ 로 나타내면 다음과 같다.

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

$$k_{p+1} = -1 \text{ 일때, } Q_{p+1}(z) = A_p(z) + B_p(z)$$

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

두 개의 근($k_{p+1} = \pm 1$)을 알고 있으므로 $P_{p+1}(z)$ 의 $Q_{p+1}(z)$ 의 차수를 줄일 수 있다. 즉,

$$P(z) = \frac{P_{p+1}(z)}{(1-z)} = A_0 z^p + A_1 z^{(p-1)} + \dots + A_p \tag{3}$$

$$Q(z) = \frac{Q_{p+1}(z)}{(1-z)} = B_0 z^p + B_1 z^{(p-1)} + \dots + B_p \tag{4}$$

$$\text{조건 : } A_0 = 1, B_0 = 1 \tag{5}$$

$$A_k = (a_k - a_{p+1-k}) + A_{k-1} \tag{6}$$

$$B_k = (a_k - a_{p+1-k}) - A_{k-1} \text{ for } k = 1, \dots, p$$

LSP는 $0 \leq \omega_i \leq \pi$ 인 범위에서 $P(z)$ 와 $Q(z)$ 을 통해 얻어진 근의 각 (angular) 위치를 나타낸다. LSP는 다음과 같은 두가지 성질을 지닌다.

- 첫째, $P(z)$ 와 $Q(z)$ 는 단위원 상에 놓여 있다.
- 둘째, $P(z)$ 와 $Q(z)$ 의 근들이 단위원 상에 번갈아 나타난다.

III. LSP 파라미터의 피팅법

3.1. 실근 방법의 특징

$P(z)$ 와 $Q(z)$ 의 계수는 대칭적이기 때문에 식 (6)의 차수는 $p/2$ 로 줄어든다. 그리고 모든 근이 단위원 상에 있기 때문에, 아래와 같이 정의하면 단위원 상에서 값을 구할 수 있다.

$$\text{Let } z = e^{j\omega} \text{ then } z^1 + z^{-1} = 2 \cos(\omega) \tag{7}$$

$$P(z) = 2e^{j\omega/2} [A_0 \cos(\frac{\omega}{2}) + A_1 \cos(\frac{\omega-2}{2}) + \dots + \frac{1}{2} A_{p/2}] \tag{8}$$

$$Q(z) = 2e^{j\omega/2} [B_0 \cos(\frac{\omega}{2}) + B_1 \cos(\frac{\omega-2}{2}) + \dots + \frac{1}{2} B_{p/2}] \tag{9}$$

$x = \cos \omega$ 를 대입하면 식 (8)과 식 (9)를 x 에 대해서 풀 수 있다. $p=10$ 의 경우 다음 식이 얻어진다[4,7].

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

$$Q_{10}(x) = 16B_0 x^5 + 8B_1 x^4 + (4B_2 - 20B_3) x^3 + (2B_3 - 8B_4) x^2 + (5B_0 - 3B_2 + B_4) x + (B_1 - B_3 + 0.5B_5) \tag{11}$$

LSP는 식 (12)에 의해서 구해진다.

$$LSP(i) = \frac{\cos^{-1}(x_i)}{2\pi T}, \text{ for } 1 \leq i \leq p \tag{12}$$

3.2. 멜 스케일을 이용한 LSP 변환 알고리즘의 성능 개선

멜 스케일이란 인간의 청각특성을 고려한 주파수 척도이며 1000 Hz에 1000 Mel을 대응하고 이로부터 실험적으로 값을 결정한다. 일반적인 의미의 멜 스케일은 1000 Hz 이하에서는 주파수와 선형적으로 비례하며 1000 Hz 이상에서는 로그 스케일이다. 본 논문에서는 Fant에 의하여 근사된 식을 이용하였고 식 (13)에서 나타내었다. 식 (13)은 멜 스케일과 주파수 사이의 관계를 나타내고 있다[2,5,6,8,9].

$$F_{mel} = k * \log [1 + \frac{f}{1000}] \tag{13}$$

$$k = 1000 / \log 2$$

본 논문에서는 LSP 파라미터를 찾는 과정에 있어서 식 (13)을 이용하였다. 즉 1 kHz까지는 선형적으로 1 kHz 이상의 주파수 대역에서는 로그 스케일로 검색하였다.

3.3. LSP 파라미터의 분포특성을 이용하여 검색 간격을 조절한 방법

앞에서 설명한 실근 방법에서는 홀수번째 계수를 먼저 찾은 후 짝수번째 계수는 이미 찾아진 홀수번째 계수 사이에서 찾는다. 따라서 홀수번째 계수를 찾는데 걸리는 시간이 변환시간의 대부분을 차지하게 된다. 본 논문에서 LSP 파라미터의 피팅 (fitting)을 위해 고려한 방식 중에서 두 번째 방식은 다음과 같은 특징을 이용하였다. LSP 파라미터의 분포특성은 순차적이지 않다. 즉, 어느 특정 주파수 대역에 편중되어서 나타난다. 그러므로 분포특성을 이용하여 분포도가 높은 곳에서는 조밀한 간격으로 검색하고

표 1. 검색간격에 따른 LSP의 변화결과 (1)

Table 1. The LSP using adjustment search band interval (Utterance 1).

LSP	검색간격 (해상도)					
	5Hz	10Hz	20Hz	40Hz	80Hz	200Hz
LSP(1)	0.0278	0.0278	0.0279	0.0278	0.0278	0.0316
LSP(2)	0.0408	0.0408	0.0408	0.0408	0.0417	0.0413
LSP(3)	0.0565	0.0565	0.0565	0.0563	0.0553	0.0527
LSP(4)	0.0915	0.0915	0.0911	0.0903	0.0874	0.0594
LSP(5)	0.2053	0.2053	0.2053	0.2053	0.2052	0.2050
LSP(6)	0.2950	0.2950	0.2951	0.2951	0.2953	0.3054
LSP(7)	0.3245	0.3245	0.3245	0.3245	0.3245	0.3245
LSP(8)	0.3895	0.3895	0.3895	0.3896	0.3915	0.4026
LSP(9)	0.4279	0.4279	0.4279	0.4278	0.4278	0.4277
LSP(10)	0.4520	0.4520	0.4519	0.4518	0.4494	0.4531

낮은 분포특성을 보이는 주파수 대역에서는 해상도를 낮추어 검색하는 방법을 취한다면 LSP 파라미터의 근을 구하는데 있어서 계산시간을 줄이게 된다. 표 1, 2에서는 음성시료 1, 4에 대하여 임의의 프레임을 택하여 검색간격을 달리한 경우에 있어서 LSP의 변화를 나타내었다.

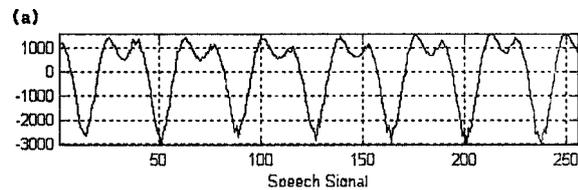
표 2에서 보여지듯이 음성시료 4를 사용하여 검색간격을 5 Hz, 10 Hz로 조절을 한 경우는 동일한 LSP 파라미터 값을 얻을 수가 있었고 검색간격을 20 Hz로 조정한 경우에 있어서는 미세한 변화를 보였다. 그리고 40 Hz 단위로 검색한 경우에 있어서는 근을 못 찾는 경우가 발생하였다. 그림 2에서는 음성시료 1을 사용하여 임의의 프레임에 대하여 해상도를 5 Hz와 200 Hz로 한 두가지 경우에 대하여 얻어진 LSP의 분포 특성을 나타내었다. 표 3에서는 주파수 대역별 검색순서와 검색간격을 나타내었다.

표 2. 검색간격에 따른 LSP의 변화결과 (2)

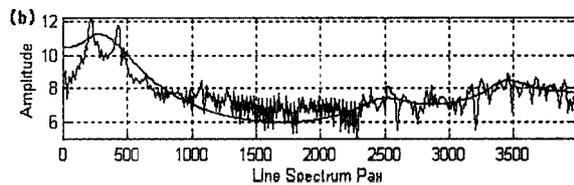
Table 2. The LSP using adjustment search band interval (Utterance 4).

LSP	검색간격 (해상도)					
	5Hz	10Hz	20Hz	40Hz	80Hz	200Hz
LSP(1)	0.0198	0.0198	0.0198	에러발생		
LSP(2)	0.0827	0.0827	0.0827			
LSP(3)	0.1273	0.1273	0.1273			
LSP(4)	0.1783	0.1783	0.1783			
LSP(5)	0.2647	0.2647	0.2647			
LSP(6)	0.2910	0.2910	0.2910			
LSP(7)	0.3362	0.3362	0.3362			
LSP(8)	0.3807	0.3807	0.3809			
LSP(9)	0.4138	0.4138	0.4138			
LSP(10)	0.4412	0.4412	0.4412			

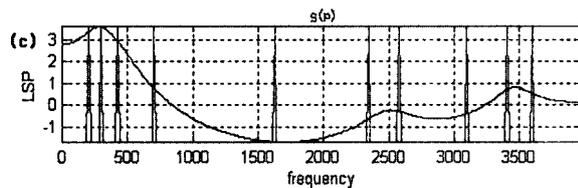
즉, 다섯번째 LSP 파라미터를 찾을 때 분포도가 가장 높은 가장 높은 주파수 대역인 1600~2000 Hz 대역은 5 Hz 단위로 검색하고 1200~1600 Hz에서는 10 Hz 단위로 검색한다. 그리고 분포도가 3, 4번째인 주파수 대역에서는 검색간격을 각각 20 Hz, 40 Hz로 조정한다. 그림 3은 음성신호에서 LSP 파라미터의 홀수번째 계수의 분포도를 나타낸 것으로, 8 kHz 표본화율에 10차의 선형예측계수를 사용하였다. 그림 3에서 LSP 홀수번째 계수들은 특정 주파수 영역에서 주로 나타나고 나머지 영역에서는 거의 나타나지 않음을 보인다. 실근 방법에서는 홀수 번째의 계수를 찾을 때 주파수대역을 순차적으로 검색한다. 그러나 그림 3의 (c), (d), (e)에서는 분포도가 순차적이지 않음을 보인다. 따라서 검색 순서를 파라미터가 많이 나타나는 주파수 대역부터 검색한다면 검색시간을 단축할 수 있다.



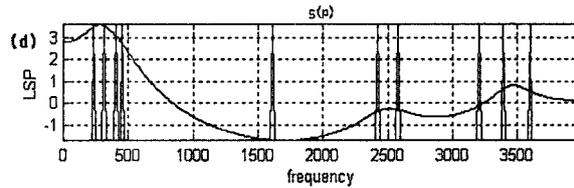
(a) 시간영역의 신호
(a) Time domain signal



(b) 주파수 스펙트럼
(b) Spectrum



(c) LSP 파라미터의 분포특성 (해상도: 5Hz)
(c) The distribution of LSPs (resolution: 5Hz)



(d) LSP 파라미터의 분포특성 (해상도: 200Hz)
(d) The distribution of LSPs (resolution: 200Hz)

그림 2. LSP 파라미터의 분포특성
Fig. 2. The distribution of LSPs.

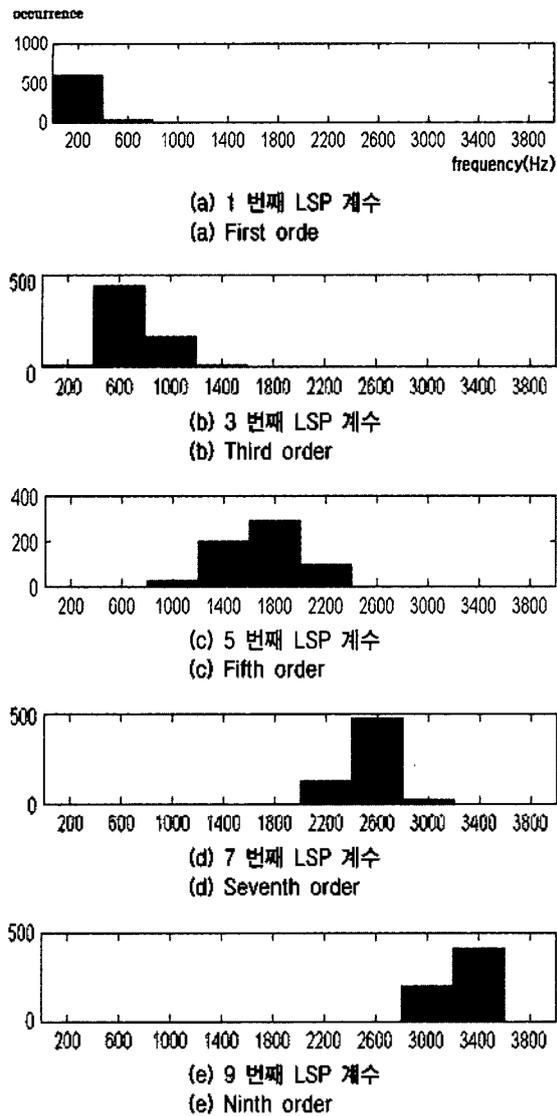


그림 3. LSP 홀수번째 계수의 분포도
Fig. 3. The distribution of odd order in LSPs.

IV. 실험 및 결과

실험에 사용된 음성시료는 연구실 환경 (30 dB의 SNR)에서 발성한 음성을 8 kHz로 표본화하고 16 bit로 양자화하여 사용하였다. 실험에 사용한 발성한 음성시료는 다음과 같다.

- 발성1: “인수네 꼬미는 천재소년을 좋아한다.”
- 발성2: “창공을 날으는 인간의 도전은 끝이 없다.”
- 발성3: “예수님께서 천지창조의 교훈을 말씀하셨다.”
- 발성4: “숭실대학교 음성통신 연구팀이다.”

본 논문에 사용한 실근 알고리즘은 C-언어로 구현된

표 3. 주파수 대역별 검색순서와 검색간격
Table 3. The searching order and Interval in searching frequency region.

검색순위 (Hz)	1	2	3	4	5
LSP 1	0-400 (5)	400-800 (5)	otherwise		
LSP 3	400-800 (5)	800-1200 (5)	otherwise		
LSP 5	1600-2000 (5)	1200-1600 (10)	800-1200 (20)	2000-2400 (40)	otherwise
LSP 7	2400-2800 (5)	2000-2400 (10)	1600-2000 (20)	otherwise	
LSP 9	2800-3200 (5)	3200-3600 (10)	otherwise		

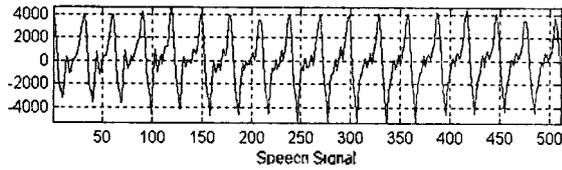
표 4. 실근 vs. 멜 스케일을 이용한 방법
Table 4. The real root method vs. The method using mel scale of searching frequency region.

비교 시료	Real Root 방법 (단위:sec)	비교평가 방법 2 (단위:sec)	감소율 (단위:%)
발성 1	80.09	41.56	48.11
발성 2	93.94	47.62	49.31
발성 3	104.27	55.84	46.45
발성 4	85.3	47.09	44.79

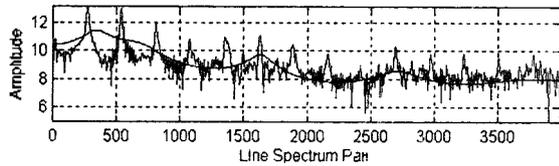
표 5. 실근 vs. 검색간격 조절 방법
Table 5. The real root method vs. The adjustment method of searching frequency region.

비교 시료	Real Root 방법 (단위:sec)	비교평가 방법 1 (단위:sec)	감소율 (단위:%)
발성 1	80.09	40.35	49.62
발성 2	93.94	46.83	50.15
발성 3	104.27	53.44	48.75
발성 4	85.3	45.82	46.28

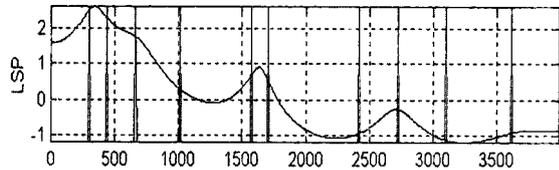
CELP 부호화기에서 발췌하여 사용하였다. 먼저 C-언어로 구현된 실근 방법의 LSP 변환함수 부분을 매트랩(Matlab) 언어로 구현하고 이것을 기반으로 하여 제안한 알고리즘을 구현하였다. 표 4는 실근 방법과 멜 스케일을 이용하여 계산시간을 단축한 방법의 비교결과이고, 표 5는 실근 방법과 LSP 분포특성을 이용하여 검색순서와 간격을 조절한 방법의 계산시간 비교결과이다. 표 4의 경우에 있어서 제안한 방법 1이 평균 47.17% 정도의 계산시간 단축을 가져옴을 나타내고 표 5의 경우에 있어서 제안한 방법 2가 약 48.7%의 계산시간이 단축되었고, 계산시간이 단축되었음을 보인다. 그림 4의 (c), (d), (e)는 각각 실근 방식으로 찾은 LSP 파라미터, 비교 평가한 첫 번째 방법, 두 번째 방법에 해당한다. 그림 4에서 보듯이 세 가지 방법에서 모두 동일한 근을 찾음을 알 수가 있다.



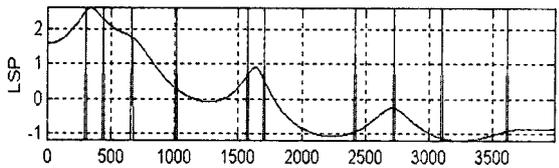
(a) 음성 파형
(a) Speech signal



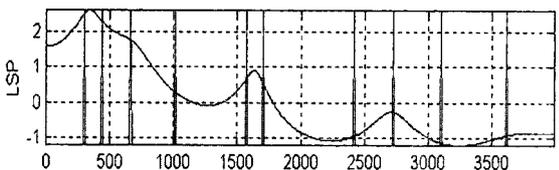
(b) 음성신호의 스펙트럼
(b) Spectrum characteristics



(c) Real root 방법
(c) The real root method



(d) Mel Scale을 이용한 방법
(d) The method using mel scale of searching frequency region



(e) 분포특성과 검색간격 조절을 이용한 방법
(e) The adjustment method of searching frequency region

그림 4. 실험에서 얻어진 LSP 파라미터의 분포
Fig. 4. The distribution of LSP parameters from experimental results.

V. 결론

저전송률 음성부호화기에서의 음성신호 선형예측 방법 중 LSP 파라미터를 이용하는 방법이 가장 많이 사용되어진다. 이것은 LSP 파라미터의 전송형 특징 중 낮은 전송률에서도 왜곡이 적고 선형보간 특성이 뛰어나기 때문이다. 하지만 LPC 계수를 LSP 파라미터로 변환하기 위해서는 많은 계산시간이 소요된다. 본 논문에서는 음성부호화기에서 주로 사용되는 실근 방법과 제한한 두가지

방법간에 계산시간 단축에 관하여 관찰하였다. 즉, 첫 번째 검색 주파수 대역에 멜 스케일을 적용한 경우에 있어서 평균 47.17%가 단축되었다. 두 번째, LSP 분포특성을 이용하여 검색순서와 간격 조절을 통하여 평균 48.7%의 계산시간을 단축하였다. 그러나 세 가지 방법 모두가 동일한 LSP 파라미터를 찾음을 알 수가 있었다. 향후 제안한 알고리즘을 음성부호화기에 적용하여 부호화기의 전체적 성능향상에 대한 연구가 이루어져야겠다.

참고 문헌

1. L. R. Rabiner, R. W. Schafer, "Digital Processing of Speech Signal", Prentice Hall, 433, 1978.
2. 배명진, "디지털 음성분석", 동영출판사, 95-120, 1998.
3. Oppenheim, Schafer, "Discrete Time Signal Processing", Prentice Hall, 318-323, 1989.
4. E. C. Ifeachor, "Discrete Time Signal Processing", Addison Wesley, 35-36, 1993.
5. T. Parson, "Voice and Speech Processing", McGraw-Hall, 71-73, 1986.
6. Douglas O. shaughnessy, "Speech Communication", IEEE Press, 127, 1996.
7. A. M. Kondoz, "Digital Speech", John Wiley & Sons Ltd, 84-88, 1994.
8. 민소연, 배명진, "A High Speed LSF Transformation Algorithm for CELP Vocoder", 한국음향학회지, 20 (1), 20-24, March 2001.
9. 민소연, 정찬중, 배명진, "주파수 대역의 적응적 조절을 통한 LSP 변환 알고리즘의 성능 개선에 관한 연구", 한국통신학회, 하계종합학술 발표회 논문집, 1316-1319, July 2001.

저자 약력

• 민 소 연 (So-Yeon MIN)



1993년 2월: 숭실대학교 전자공학과 졸업 (공학사)
1995년 2월: 숭실대학교 전자공학과 졸업 (공학석사)
2002년 2월: 숭실대학교 전자공학과 박사수료
※ 주관심분야: 음성코딩, 데이터통신

• 배 명 진 (Myung-Jin BAE)



1987년: 서울대학교 공과대학 전자공학과 졸업 (공학박사)
2002년 4월 현재: 숭실대학교 정보통신 전자공학부 교수
※ 주관심분야: 음성코딩, 음성인식, 음성합성, 데이터통신, 디지털 신호처리 등