

LSP 분포 특성을 이용한 G.723.1 보코더의 계산량 감소

이 희 원, 배 명 진
승실대학교 정보통신공학과
전화 : (02) 824-0906 / 팩스 : (02) 820-0018

The Reduction of the Computation Speed using LSP Distribution in G.723.1 Vocoder

HeeWon LEE, MyungJin BAE
Dept. Information and Telecommunication Engr., Soongsil University
E-mail : mjbae@saint.soongsil.ac.kr

Abstract

현재까지 발표된 음성 부호화기 중에서 저전송률에서 양호한 음질을 제공하는 CELP 계열 보코더에 대한 연구가 가장 많이 이루어지고 있다. 그 중에서 G.723.1 부호화기는 인터넷 폰이나 화상회의 등 상용서비스로 개발되었다. G.723.1 부호화기에서는 음성신호의 선형 예측 방법 중 LSP 파라미터를 이용하는 방법이 많이 사용된다. 이것은 LSP 파라미터의 전송형 특징 중 낮은 전송률에서도 왜곡이 적고 선형보간 특성이 뛰어나기 때문이다. 하지만 LPC 계수를 LSP 파라미터로 변환하기 위해서는 많은 계산시간이 소요된다[1].

본 논문에서는 G.723.1 보코더에서 LSP 변환 시 다항식의 근을 찾는 순서를 음성신호의 LSP 분포 특성에 맞게 조정함으로써 전체 계산시간을 평균 2% 단축하였다.

I. 서론

현재까지 발표된 음성부호화기 중 가장 많은 연구가 이루어지고 있는 방식은 그림 1-1의 구조를 가지고 있는 CELP(Code Excited Linear Prediction)구조이다. 이 방식은 4.8kbps내외의 전송률에서 양호한 음질을 얻을 수 있으며 ITU-T, TIA/EIA 등 여러 국제표준화 기구를 통해 다양한 응용분야에서 표준화가 이루어지고 있

다. 특히 한국에서는 PCS 및 전화기 라인 상에서의 인터넷을 통한 화상회의를 위하여 낮은 전송률에서 고품질을 가지는 코덱이 많은 주목을 받고 있다[2]. 이러한 CELP 계열 보코더들 중에서 G.723.1은 멀티미디어 통신 환경하의 음성 전송 표준 보코더로 개발되었다[7].

G.723.1은 5.3/6.3kbps의 이중 전송률을 갖는 구조로 현재 별정 통신으로 상용화되는 인터넷폰과 그 외의 이동 통신용 보코더로 사용되고 있으며 낮은 전송률에 비해서 우수한 음질을 제공하고 있다. 더불어 최적의 전송 환경을 위하여 두 개의 전송률을 사용하기 때문에 다른 보코더 표준안들에 비해서 더욱 응용성이 높다[8]-[12].

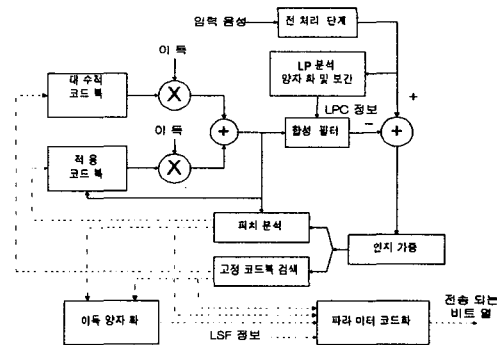


그림 1-1. 일반적인 CELP 보코더 블록도

현재 사용되는 음성 코덱(codec)이나 인식기에서 음

성신호를 분석하여 전송형이나 저장형 파라미터로 변환하는데 LSP 파라미터를 사용한다. LSP 파라미터는 양자화 에러에 강하고 시스템의 안정성이 보장된다 [5]-[6].

LSP 파라미터를 추출하기 위해서 다항식의 근을 찾는 과정을 수행한다. 본 논문에서는 음성신호의 LSP의 분포도가 LSP 파라미터의 순서에 따라 다르다는 특징을 이용하여 계산 시간을 단축하는 방법을 제안한다.

먼저 II장에서는 LSP 파라미터의 추출 방법과 근을 찾는 방법에 대해 살펴보고, III장에서는 음성신호의 LSP 분포 특성에 따른 제안한 변환 방법을 설명한다. 그리고 IV장에서는 제안한 방법의 실험 및 결과를 설명하고, V장에서는 결론을 맺겠다.

II. LSP 파라미터 추출

LSP 파라미터를 추출하기 위해서 먼저 LPC(Linear Predictive Coding)분석이 이루어져야 한다. LPC 분석은 고역 여파된 신호에 대해서 실행되어지며 10차의 선형 예측 분석(Linear predictive analysis)이 수행된다 [2]. 180표본의 해밍 윈도우(Hamming window)가 각 부프레임의 중앙에 위치하며 11개의 자기 상관계수가 윈도우 처리된 신호로부터 계산된다. 상관계수 $R[0]$ 는 1025/1024의 백색 잡음 상관 인자 처리가 되며 다른 10개의 상관계수는 이항 윈도우(binomial window) 계수 테이블에 의해 곱해진다. 선형 예측 계수(LPC)는 Levinson-Durbin recursion을 사용하여 계산되며, 모든 입력 프레임의 각 부프레임 마다 하나씩 계산되어 네 집합의 LPC 계수가 계산된다. 이 LPC 계수는 단구간 인지가중화 필터를 만들기 위해 사용되는데, 이때 LPC 합성 필터는 다음과 같이 정의된다.

$$A_i(z) = \frac{1}{1 - \sum_{j=1}^M a_{ij}z^{-j}}, \quad 0 \leq i \leq 3 \quad (2-1)$$

여기서 i 는 0과 3사이에서 정의되는 부프레임 인덱스이며 j 는 차수를 나타낸다.

LPC(Linear Predictive Coefficients) 양자화를 하기 위하여, 먼저 7.5Hz의 대역폭 확장(bandwidth expansion)이 수행된다. 그 결과로 $A_3(z)$ LP 필터는 예측 분할벡터 양자화기를 이용하여 양자화 되어진다. 양자화 과정 중에 LP 계수 $\{a_j\}_{j=1..10}$ 는 단위원과 영

교차에 의한 인터폴레이션(interpolation) 과정동안 LSP 계수 $\{b'_j\}_{j=1..10}$ 로 변환된다[4].

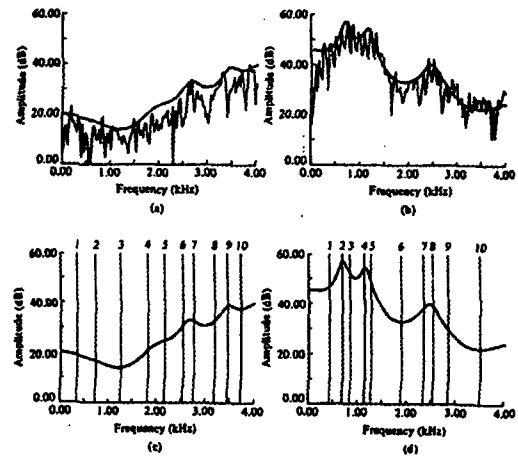


그림 2-1. 선형스펙트럼쌍의 예
(a) 자음 /s/ (b) 모음 /a/
(c),(d) /s/와 /a/에 대한 LPC 분석과 LSP

LSP에서 선쌍(line pair)의 간격이 좁으면, 강한 공명이 일어난 것을 나타낸다. 즉, LSP는 p개의 불연속적인 주파수의 분포를 통하여 음성의 스펙트럼 포락선을 표현하는 방법이다. 그림 2-1은 자음과 모음의 선형스펙트럼쌍의 예이다[3].

III. 음성 신호의 LSP 분포

음성 신호의 LSP 분포도를 확률적으로 조사하기 위해 20분 정도의 긴 음성의 분포도를 조사하였다. 음성 시료는 8kHz 표본화율에 10차의 선형예측계수를 사용하였다. 가로축은 주파수 대역이고 세로축은 발생빈도를 나타낸다. 256 포인트를 16개의 bin으로 나누어 분포도를 조사하였다. 그림 3-1은 음성신호에서 LSP 파라미터의 홀수 번째 계수의 분포도를 나타낸 것이다.

LSP 홀수 번째 계수들은 특정 주파수 영역에서 주로 나타나고 나머지 영역에서는 거의 나타나지 않는 것을 볼 수 있다. G.723.1 보코더에서 LSP 다항식의 근을 찾기 위해 순차적으로 영교차점을 체크하여 근을 구한다. 그러나 그림 3-1을 살펴보면 분포도가 순차적이지 않다는 것을 알 수 있다. 따라서 검색 순서를 파라미터가

많이 나타나는 주파수 대역부터 검색한다면 검색시간을 단축할 수 있다. 짝수 번째 LSP 계수는 순차적으로 검색한다.

표 3-1은 홀수 번째 LSP 파라미터의 주파수 대역별 검색 순서로써 분포도에 의해 제일 많이 나타나는 구간을 제일 먼저 검색하게 된다.

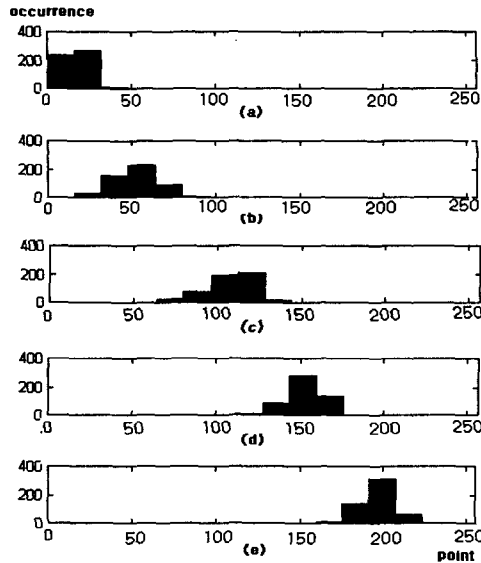


그림 3-1. LSP 홀수 번째 계수의 분포도

- (a) 1 번째 LSP 계수 (b) 3 번째 LSP 계수
(c) 5 번째 LSP 계수 (d) 7 번째 LSP 계수
(e) 9 번째 LSP 계수

표 3-1. 주파수 대역별 검색순서(단위 : point)

검색순위	1	2	3	4
LSP(1)	0~31	otherwise		
LSP(3)	48~63	32~47	64~79	otherwise
LSP(5)	96~127	otherwise		
LSP(7)	144~175	128~143	otherwise	
LSP(9)	192~207	176~191	208~223	otherwise

IV. 실험 및 결과

제안한 방법을 실험하기 위해서 먼저 IBM PC(233 MHz)에 마이크 입력이 가능한 A/D 변환기를 인터페이스 하였다. 음성시료는 남자와 여자가 연구실 환경(30dB의 SNR)에서 발성한 음성을 8kHz로 표본화하고 16bit로 양자화하여 사용하였다. 발성한 문장은 다음과 같다.

- 발성1) “인수네 꼬마는 천재소년을 좋아한다.”
 발성2) “창공을 날으는 인간의 도전은 끝이 없다.”
 발성3) “예수님께서 천지창조의 교훈을 말씀하셨다.”
 발성4) “숭실대학교 정보통신공학과 음성통신 연구팀이다.”
 발성5) “아나운서의 뉴스”

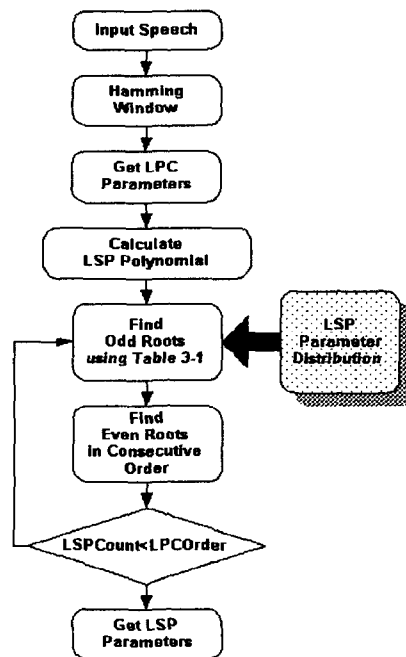


그림 4-1. 제안한 알고리즘의 블록도

제안한 알고리즘의 블록도는 그림 4-1과 같다. 음성 신호가 들어오면 해밍 윈도우를 취하고 LPC 계수를 추출한 후 이것으로 다차 방정식을 만든다[3]. 그 후 LSP 파라미터의 분포도를 적용한 표 3-1의 검색순서에 따라 홀수 번째 근을 찾고 홀수 번째 근을 찾은 지점부터

순차적으로 짝수 번째 근을 찾은 후 다시 표 3-1에 따라 홀수 번째 근을 찾는다. 이러한 과정은 10개의 근을 찾을 때까지 계속된다.

표 4-1.은 제안한 알고리즘과 기존의 알고리즘을 사용하여 입력음성을 완전히 코딩하는데 걸리는 시간을 나타낸다. 제안한 알고리즘을 이용하여 LSP 다항식의 근을 찾았을 때 기존의 방법을 이용한 것과 비교하여 평균 2% 빠르다는 것을 보여주고 있다. 검색 순서만 다르게 했을 뿐 근이 모두 찾아지지 않으면 모든 구간을 검색했기 때문에 변환된 LSP 계수를 비교해 보면 기존의 것과 동일하였다.

표 4-1. 기존의 방법과 제안한 방법의 계산시간 비교

	기존의 방법 (단위:초)	제안한 방법 (단위:초)	감소율 (단위:%)
발성(1)	6.530	6.310	3.37
발성(2)	7.910	7.470	2.78
발성(3)	7.580	7.470	1.45
발성(4)	11.860	11.700	1.35
발성(5)	32.960	32.620	1.03

V. 결 론

현재까지 발표된 음성부호화기 중에서 저전송률에서 양호한 음질을 제공하는 CELP 계열 보코더에 대한 연구가 가장 많이 이루어지고 있다. 이런 CELP 계열 보코더 중에서 G.723.1 보코더는 인터넷 폰이나 화상회의 등 상용서비스로 개발되었다. G.723.1 부호화기에서는 음성신호의 선형예측 방법 중 LSP 파라미터를 이용하는 방법이 많이 사용되어 지고 있다. 이것은 LSP 파라미터의 전송형 특징 중 낮은 전송률에서도 왜곡이 적고 선형 보간 특성이 뛰어나기 때문이다. 하지만 LPC 계수를 LSP 파라미터로 변환하기 위해서는 많은 계산시간이 소요된다.

본 논문에서는 G.723.1 보코더에서 LSP 변환 시 다항식의 근을 찾는 순서를 음성신호의 LSP 분포 특성에 맞게 조정함으로써 계산시간을 단축하였다. 기존의 알고리즘에서는 LSP 파라미터를 구하기 위해 주파수 대역을 순차적으로 검색한다. 그러나 제안한 알고리즘에서는 주파수 대역을 LSP 파라미터의 분포도에 따라 검색 구간의

순서를 정하여 검색함으로써 평균 2%의 변환 시간을 단축할 수 있었다.

참고 문헌

- [1] 배명진, "디지털 음성분석", pp.95-120, 동영출판사, 1998. 4
- [2] L. R. Rabiner, R.W. Schafer, "Digital Processing of Speech Signal", pp.38-115, Prentice Hall, 1978
- [3] A. M. Kondoz, "Digital Speech", pp. 84-92, John Wiley & Sons Ltd, 1994.
- [4] ITU-T Recommendation G.723.1, March, 1996
- [5] John R. Deller, Jr., John G. Proakis, John H.L. Hansen, "Discrete-Time Processing of Speech Signals", pp.124-125, Maxwell Macmillan International, 1993.
- [6] Sadaoki Furui, "Digital Speech Processing, Synthesis, and Recognition", pp.129, MARCEL DEKKER, INC. 1991.
- [7] YoungHo Park, MyungJin Bae, "On Improving the Performance of Low-Bit Rate Speech Coder", Proc. of KSCSP98, 1998. 8, pp.131-135
- [8] B. Atal. "Efficient coding of LPC parameter by temporal decomposition", Proc. of ICASSP, pp.81-85, 1983.
- [9] J. P. Campbell, Jr., V. C. Welch and T. E. Tremain, "Error-protected 4800 bps CELP Coder(U.S. Fedral Standard 4800 bps Voice Coder)", Proc. Int. Conf. on Acoustics, Speech and Signal Processing, 1989.
- [10] W. B. Kleijn et al., "Fast Methods for the CELP Speech Coding Algorithm", IEEE Trans., Acoustics, Speech and Signal Processing, Vol.38, No.8, pp.1330-1341, August 1990.
- [11] J. H. LEE, H. Y. JEON, M. BAE, and S. ANN, "A Fast Pitch Searching Algorithm using Correlation Characteristics in the CELP Vocoder," IEEE Communication Society, Proceeding of MILCOM '94, pp.699-702, Oct.2-5 1994.
- [12] D. KIM, M. BAE, J. KIM, K. BYUN, K. HAN, H.YOO, "On a Reduction of Pitch Searching Time by Preprocessing in the CELP Vocoder," J. Acoust. Soc, Korea, Vol.13, No.3, pp. 33-40, June 1994.