

대역 분할 장구간 예측을 이용한 CELP 부호화기

A CELP Coder using the Band-Divided Long Term Prediction

최 용 수*, 강 홍 구*, 임 명 섭**, 안 동 순***, 윤 대 회*

(Young Soo Choi*, Hong Goo Kang*, Myoung Seob Lim**,
Dong Soon Ahn***, Dae Hee Youn*)

본 연구는 한국전자통신연구소의 연구비 지원에 의한 것입니다.

요 약

본 논문에서는 CELP(Code Excited Linear Prediction) 부호화기의 장구간 예측단에 MBE(Multi-Band Excitation) 방법을 도입하여 4.8 kbps 이하의 낮은 전송률에서 성능을 향상시키는 방법을 제안한다. 제안한 방법에서는 기존의 CELP 방법으로 장구간 예측을 한 후에도 여전히 남아 있는 주기적 성분들을 다시 한번 다중대역(multiband)으로 장구간 예측을 한다. 이때 전 스펙트럼을 기본 주파수의 하모닉 간격으로 대역 분할하고, 주기적 다중대역 여기 신호는 각 대역 내의 하모닉들에서 여기 신호 스펙트럼과 근사한 크기를 갖는 정현파(sine wave)의 합으로 표현함으로써 실제 여기 신호의 특성을 잘 반영하도록 한다. 제안된 방법의 성능 평가를 위해서 4.8 kbps의 전송률에서 컴퓨터 모의 실험을 하였다. 음질 평가의 비교 기준 대상으로 4.8 kbps DoD CELP와 4.4 kbps IMBE를 선정하여 주관적인 음질 평가를 실시한 결과 4.8 kbps DoD CELP 보다 우수하였고, 4.4 kbps IMBE와는 비슷하였다.

ABSTRACT

In this paper a way to improve the performance of the long term prediction is proposed, which adopts the Multi-Band Excitation (MBE) method in addition to the Code-Excited Linear Prediction (CELP) method at low bit rates below 4.8 kbps.

In the proposed method, the multiband long term prediction is performed on the periodic components which still remain after the long term prediction of the conventional CELP method. At this point, the whole frequency region is divided into subbands whose size is equal to the spacing between the harmonics of the fundamental frequency, and the periodic multiband excitation signals are represented as the sum of sine waves approximately as large as the spectrum of the excitation signals, so that the

*연세대학교 전자공학과
Dept. of Electronics Eng. Yonsei Univ.

**한국전자통신연구소

***목포대학교 컴퓨터공학과

접수일자: 1995년 3월 17일

actual characteristics of the excitation signals can be better taken into account.

To evaluate the performance of the proposed method, computer simulation is performed at 4.8 kbps. The 4.8 kbps DoD CELP and the 4.4 kbps IMBE were chosen as the reference vocoders for the speech quality measure. The result of the perceptual speech quality measure showed that the performance of the proposed method is better than that of the 4.8 kbps DoD CELP vocoder, and similar to that of the 4.4 kbps IMBE vocoder.

I. 서 론

이동 통신에 대한 수요가 급증함에 따라 채널을 효율적으로 사용하기 위해서는 낮은 전송률에서 고품질을 유지하는 음성 부호화 알고리즘 연구는 필수적이며, 다른 한편으로 음성 우편과 응답 전화기와 같은 시스템에서도 효율적인 음성 저장을 위하여 낮은 비트율 음성 부호화에 관한 연구는 매우 중요하다[1].

CELP[2]는 4.8 kbps 이하의 낮은 전송률에서도 비교적 양호한 음질을 얻을 수 있어서 많은 연구가 이루어지고 있는 부호화 방법이다.

일반적인 CELP 알고리즘에서 전송률을 낮추기 위한 손쉬운 방법은 프레임의 길이를 늘리는 것이다. 그러나 이 경우에는 장시간 예측 후에도 피치 성분들이 많이 남아 있게 되어 합성음의 음질은 크게 저하된다. 이러한 문제점들을 해결하기 위하여 여러 방법들이 제안되었다[3]-[5]. 이러한 방법들로부터 합성음의 주기성을 향상시키는 것이 주관적(perceptual) 음질을 높일 수 있다는 사실을 알 수 있다. 이 중에서 [5]에서 제안된 MBCELP (Multi Band CELP)는 기존의 CELP 처럼 첫번째 적응 코드북으로 부터 장구간 예측을 한 후, 다중대역으로 필터링된 임펄스열로 구성된 두번째 적응 코드북으로 부터 한번 더 장시간 예측을 하여 잔차 신호에 남아 있는 주기적 성분들을 제거함으로써 합성음의 음질을 크게 향상시켰다. 그러나 MBCELP는 4.8 kbps 이하의 전송률로 구현될 경우에 제한된 비트로 인해 다중 대역의 수가 제한되고, 실제 이상적 여기 신호 스펙트럼은 각 대역 내의 하모닉들에서 크기가 다르게 분포되어 있음에도 불구하고, MBCELP에서의 주기적 다중대역 여기 신호 스펙트럼은 각 대역 내의 하모닉들에서 일정한 크기를 갖도록 구성된다.

본 논문에서는 위의 MBCELP와는 다른 접근 방식의 Sinusoidal MBCELP(SMBCELP)를 제안한다. 제안된 SMBCELP에서는 장구간 예측 후 잔차 신호의 스펙트럼을 기본 주파수의 하모닉 간격으로

대역 분할하고, 주기적 다중대역 여기 신호의 스펙트럼을 각 대역 내의 하모닉들에서 여기 신호 스펙트럼과 근사한 크기를 갖는 정현파(sine wave)의 합으로 모델링함으로써 LPC 합성 여기 신호의 특성을 잘 반영하도록 하였다.

본 논문은 다음과 같이 구성된다. 2장에서는 SMBCELP 알고리즘을 설명하며, 3장에서는 SMBCELP 모델 파라메타를 추정하는 방법을 보여준다. 4장에서는 4.8 kbps SMBCELP에 대한 컴퓨터 모의 실험 결과를 고찰하고, 5장에서 결론을 맺는다.

II. SMBCELP 모델

SMBCELP의 전체적인 구조는 그림 1과 같으며, MBCELP[5]와는 근본적으로 주파수 대역 분할 방법과 주기적 다중대역 여기 신호의 생성 방법이 다르다. 그림 1에서 MBAC(MultiBand Adaptive Codebook), MBESEFC(MultiBand Excitation Spectral Envelope Fixed Codebook)는 각각 다중대역 적응 코드북, 다중대역 여기 신호의 스펙트럼 포락선 고정 코드북으로서 구성 및 사용 방법은 다음 장에서 자세히 설명한다.

단구간 예측기로는 인접 샘플 간의 상관 관계는 제거할 수 있지만 유성을 구간에서는 준주기성(quasi periodicity)을 제거하지는 못한다[2]. 이러한 준주기성을 제거하기 위하여 장구간 예측기를 사용한다. 그러나, 장구간 예측기가 주기적 여기 신호를 나타내는데 커다란 기여를 했음에도 불구하고 예측 후에도 여전히 주기적 성분들이 상당히 남아 있다. 그림 2(b)에 예시된 장구간 예측 후 여기 신호의 스펙트럼에서 알 수 있듯이 주기적 여기 신호의 스펙트럼에 너지는 주파수 영역에 따라 크기가 다르게 나타난다. 여기서 주목할 것은 주기적 성분들은 그림 2(a)에서 관찰할 수 있는 것과 같이 피치 주기를 갖고 있고, 그 스펙트럼은 그림 2(b)에서 볼 수 있듯이 기본 주파수의 정수배에 해당하는 하모닉들에 에너지가 집중되

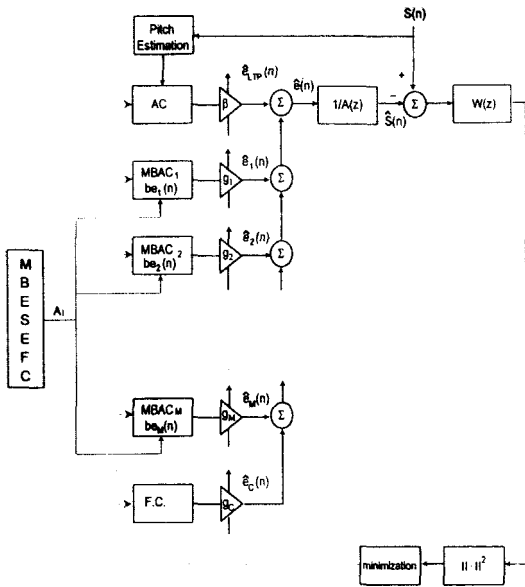
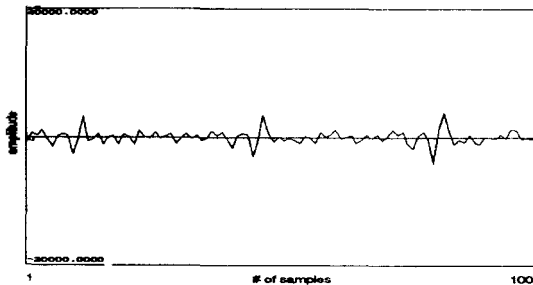
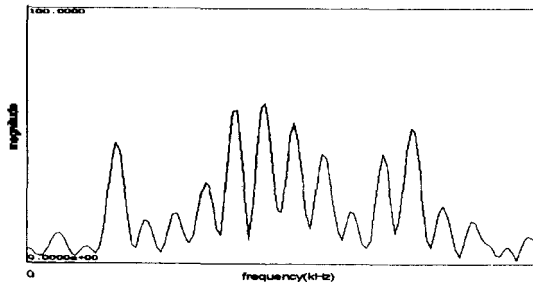


그림 1. SMBCELP 부호화기



(a) 신호 파형



(b) 스펙트럼

그림 2. 장구간 예측 후의 여기 신호

어 있으며 그 스펙트럼 포락선은 평탄하지 않다는 것이다.

MBCELP[5]에서는 그림 2(b)에서 보듯이 장구간 예측 후 여기 신호의 스펙트럼이 각 대역 내의 하모닉들에서 크기가 다르게 분포되어 있음에도 불구하고 비트할당의 제한 때문에 일정한 크기를 갖도록 구성하므로 정확한 모델링이 어렵고, 또한, 대역필터를 사용하여 대역 분할된 여기 신호를 얻음으로써 분할 대역의 수가 많아지면 필터링에 소요되는 계산량이 증가하고, 필터링에 의한 리플등의 좋지 않은 영향이 커지는 등 단점이 있다.

이를 해결하기 위해 제안된 SMBCELP에서는 주기적 다중대역 여기 신호 스펙트럼을 각 대역 내의 하모닉들에서 실제 여기 신호 스펙트럼과 유사한 크기를 갖는 정현파의 합으로 표현함으로써 균일하지 않은 여기 신호 스펙트럼 포락선의 실제적인 특성을 반영하고, 필터링에 의한 계산량과 리플의 영향 등을 줄일 수 있다.

다중대역의 갯수 M 은 기본주파수 ω_0 의 하모닉 갯수 L 로 하는 것이 이상적이지만, SMBCELP를 4.8 kbps 이하의 낮은 전송률로 구현하기 위해서 IMBE [6]에서와 유사한 방법으로 다중대역들을 다시 몇개씩 묶어서 그 수를 줄일 수 있다. 그림 3은 다중대역의 갯수 M 이 2인 경우의 스펙트럼 포락선과 분할 대역 구조의 예를 보여 준다.

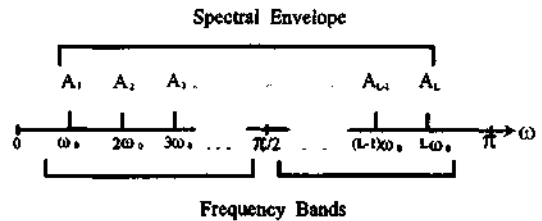


그림 3. 다중대역의 갯수 M 이 2인 경우의 스펙트럼 포락선과 분할 대역 구조

다중대역 적용 코드북(MBAC)의 여기 신호 $be_m(n)$ 은 분할된 주파수 대역 B_m 내의 기본 주파수 ω_0 의 하모닉들 $l\omega_0$ 를 주파수로 하고 크기 A_l 을 갖는 정현파의 합으로 표현될 수 있다.

$$be_m(n) = \sum_l A_l \sin(l\omega_0 n + \phi_l(-1)),$$

$$\omega_0 \in B_m, 1 \leq l \leq L, 1 \leq m \leq M \quad (1)$$

여기서 M 은 다중대역의 갯수이고, L 은 기본주파수의

ω_0 의 하모닉 갯수이며, $\phi_l(-1)$ 은 초기 위상으로 각 하모닉 성분들의 상대적 위상을 보상해주는 역할을 하는데 다음과 같이 표현될 수 있다[6].

$$\phi_l(0) = \phi_l(-1) + \frac{l(\omega_0(0) + \omega_0(-1))}{2} N \quad (2)$$

여기서 N , $\phi_l(-1)$, $\phi_l(0)$, $\omega_0(0)$, $\omega_0(-1)$ 는 각각 부프레임 길이, 과거 프레임에서의 위상, 현재 프레임에서의 위상, 현재 프레임에서의 기본 주파수, 과거 프레임에서의 기본 주파수이다.

III. SMBCELP 모델 파라메타의 추정

SMBCELP 모델에서 전송될 파라메타들은 선형 예측 계수(LSP), 피치(Pitch1, Pitch2), 장구간 예측 이득(β), 다중대역 여기 신호의 스펙트럼 포락선 코드북 인덱스(I_m), 다중대역 적응 코드북에서의 주기적 여기 신호의 최적 위치(Pos_m), 다중대역 주기 여기 신호들의 이득(g_m), 통계적 코드북에서 여기 벡터의 최적 인덱스(I_c)와 이득(g_c)이다.

두번째 부프레임에서의 피치 파라메타 Pitch2는 첫번째 프레임에서의 피치 Pitch1과의 차이이고, 첫번째 다중대역 적응 코드북 MBAC1에서 부프레임 길이의 주기적 여기 신호 $be_1(n)$ 의 최적 위치 Pos_1 이 선택되면 다른 다중대역 적응 코드북 MBACm에 대해서는 여기 신호 벡터의 위치 Pos_m 은 첫번째 주기적 여기 신호 $be_1(n)$ 의 최적 위치 Pos_1 주위의 작은 범위 내에서 찾거나, 또는 비트 할당의 제한을 고려하여 같은 위치로 하여 이득을 다르게 찾을 수도 있다.

3.1 장구간 예측 파라메타 추정

장구간 예측 알고리즘으로 얻어진 장구간 예측 피치는 음성의 실제 피치와 항상 일치하지는 않는다. 그 값은 실제 피치 주기의 정수배일 수도 있고 실수배일 때도 있다. 따라서 피치가 서서히 변한다 할지라도 인접한 부프레임에서의 장구간 예측 피치의 변화의 차이를 그대로 부호화 하면 음질 저하가 나타날 수 있다. 특히, 부프레임 길이가 길어지게 되는 경우에는 합성음의 음질에 더욱 나쁜 영향을 미치게 된다. 이러한 영향을 피하기 위해서 매 프레임마다 장구간 예측 피치의 초기값을 추정하는 방법을 사용한다. 원 신호로부터 피치를 추정하는 많은 기법들이 있지만 비교적 신뢰성 있는 IMBE[6]에서의 피치 알

고리즘을 사용한다. [6]에서는 동적 프로그램을 사용하여 현재 프레임을 중심으로 과거 및 미래의 몇 프레임부터 부터의 전후방 피치 추적을 하여 프레임간 피치의 연속성을 유지하는 최적의 피치를 구한다.

각 프레임에서 장구간 예측 피치는 피치 추적에 의해 찾아진 초기 피치를 주위도 아주 작은 범위 내에서 추정된다. 장구간 예측 피치는 천천히 변하므로 두번째 부프레임에 대해서는 장구간 예측 피치 변화의 차이만을 부호화함으로써 비트수를 줄일 수 있다.

3.2 주기적 다중대역 여기 신호 파라메타 추정

장구간 예측 후의 여기 신호는 주기성을 갖는 다중대역 신호들로 구성되어 있는 MBAC로 부터 예측을 하게 되는데 이 방법은 장구간 예측 방법과 비슷하다.

다중대역 여기 신호 파라메타들은 각 주파수 대역에 대해서 독립적으로 추정될 수 있으며, 분할된 다중대역의 갯수에 따라 두가지 접근 방식이 있을 수 있다.

3.2.1 다중대역 갯수가 기본 주파수의 하모닉 갯수와 같을 경우

다중대역의 갯수 M 이 기본 주파수 ω_0 의 하모닉 갯수 L 과 같을 때는 매우 이상적인 경우로서 각 대역에서의 파라메타 추정은 MBCELP[5]에서와 같은 추정 방법을 사용한다.

각 다중대역 내의 기본 주파수의 하모닉에서의 크기 A_k 은 장구간 예측 후 여기 신호의 스펙트럼 포락선을 기본 주파수의 하모닉 간격으로 샘플링하여 구할 수 있다. 여기서 장구간 예측 후 여기 신호의 스펙트럼 포락선을 추정하는 방법은 세가지가 있다.

첫째, 스펙트럼을 기본주파수의 하모닉에서 샘플링하여 피크를 검출하는 방법[7], IMBE[6]에서 유성음 대역에 적용되는 방법인 다중대역 스펙트럼과 창함수가 곱해진 임펄스열의 스펙트럼과의 상대적 비로 하는 방법, 셋째, IMBE에서 무성음 대역에 적용되는 대역 스펙트럼의 평균 크기로 하는 방법 등이 있다. 본 논문에서는 이 중에서 주기적 스펙트럼과 잡음 성분이 섞여 있는 스펙트럼 모두에 대해서 스펙트럼 포락선을 비교적 잘 나타낼 수 있는 세번째 방법을 사용하였다.

3.2.2 다중대역 갯수가 기본 주파수의 하모닉

갯수 보다 적을 경우

3.2.1의 경우에서 기본 주파수의 하모닉에서 추정된 다중대역 여기 신호의 스펙트럼 포락선 사각을 직접 사용하기에는 너무 많은 비트가 필요하므로 낮은 비트율로 구현하기에는 문제가 있다. 그러므로 실제로 4.8 kbps이하의 낮은 비트율로 SMBCELP를 구현하기 위해서는 2장에서 언급했듯이 전체 L 개의 하모닉들을 다시 몇개의 묶어서 다중대역의 갯수 M 을 줄이고, 묶여진 대역 내에서 하모닉들의 크기 A_i 을 벡터 양자화 방법을 사용하여 다중대역 여기 신호의 스펙트럼 포락선을 표현한다.

벡터 양자화에 사용되는 다중대역 여기 신호의 스펙트럼 포락선 고정 코드북(MBESEFC)은 구간 별로 기본 주파수가 다르고 에너지도 다르므로 각 하모닉에서 추정된 스펙트럼 포락선을 선형 보간(linear interpolation)하고, 이를 다시 에너지에 대해서 정규화(normalization)한 후, 최종적으로 LBG 알고리즘[10]을 이용하여 훈련(training) 과정을 거쳐 구성한다.

코드북 검색시에는 스펙트럼 포락선으로 구성된 코드 벡터를 각 대역에 대해 평균 크기를 구한 후, 하모닉의 크기 A_i 로 사용하여 다중대역 적용 코드북(MBAC)의 여기 신호 벡터 $be_m(n)$ 을 구성한다.

3.3 통계적 코드북 파라메타 추정

다중대역 장구간 예측을 거친 잔차 신호는 다시 통계적 코드북을 사용하여 예측을 하게된다. 통계적 코드북은 백색 가우시안 잡음을 센터 클리핑(center clipping)한 sparse 코드 벡터로 구성하였다.

정현파를 이용한 다중대역 장구간 예측을 한 후, 피치 성분이 상당히 제거되어 백색 잡음에 가까워진 잔차 신호를 MBCELP에 비해서 그 크기가 상당히 줄어든 통계적 코드북으로도 잘 예측할 수 있다.

IV. 실험 및 성능 평가

4.1 4.8 kbps SMBCELP의 컴퓨터 모의 실험

제안된 방법의 성능 평가를 위해서, 각 파라메타들에 표 1과 같은 비트를 할당 하여 4.8 kbps의 전송률을 갖는 SMBCELP 부호화기를 구현하였다.

LPC(Linear Predictive Coding) 분석을 위해 240 샘플(30ms)의 Hamming 창함수를 40샘플씩 중첩하여 사용하였다. 한 프레임의 길이는 200샘플(25ms)로 하였으며, 각 프레임은 100샘플(12.5ms)씩 두개

표 1. 4.8 kbps SMBCELP 파라메타와 비트 할당

Parameters	Bits/frame
LPC(LSP)	34
Pitch1 + 1Pitch2	7 + 3
β (LTP gain)	2 × 4
I_m (MB excit. spec env. codebook index)	2 × 6
Pos_1 (MB1 excit. position)	2 × 7
Pos_2 (MB2 excit. position)	0
g_1 (MB excit. gain1)	2 × 4
g_2 (MB excit. gain2)	2 × 4
I_c (codebook index)	2 × 8
g_c (codebook gain)	2 × 5
Total	120

의 부프레임으로 구성된다.

4.8 kbps라는 낮은 비트율의 제한으로 인해 다중대역의 갯수 M 은 두개로 하고, 두번째 다중대역 여기 신호 $be_2(n)$ 의 위치 Pos_2 는 첫번째 다중대역 여기 신호 $be_1(n)$ 의 위치 Pos_1 와 같게 놓고 이득만을 다르게 사용하였다.

LPC는 전송을 위해서 LSP(Line Spectrum Pair)로 변환되고 변환된 LSP는 프레임 간 LSP 값의 차이를 34 비트 스칼라(scalar) 양자화를 사용하여 부호화하였다[9]. 첫번째 부프레임에서 장구간 예측 피치 $Pitch_1$ 은 7 비트로 부호화되고, 두번째 부프레임에서는 첫번째 부프레임에서의 장구간 예측 피치와의 상대적 차 $Pitch_2$ 를 3비트로 부호화하였다. 장구간 예측 이득 β 는 4비트 대수(logarithmic) 양자화를 사용하여 부호화하였다. 다중대역 여기 벡터의 스펙트럼 포락선은 통계적 특성을 잘 반영하도록 훈련된(trained) 코드북으로 하여 6비트 벡터 양자화하였다. 다중대역 여기 벡터 이득 g_1 과 통계적 코드북 이득 g_c 는 각각 4비트, 5비트 대수 양자화를 사용하여 부호화하였다.

4.2 성능 평가

4.8 kbps SMCELP의 성능을 평가하기 위해 비교 기준 부호화기로 4.8 kbps DoD CELP[8]와 4.4 kbps IMBE[6]를 선정하였다.

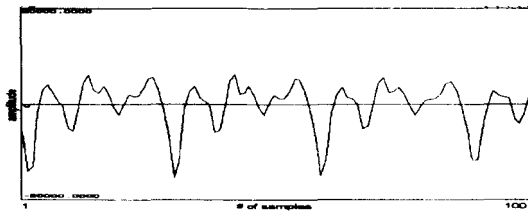
그림 4는 음성 신호의 한 구간에 대한 4.8 kbps

SMBCELP 관련 신호들이다. 장구간 예측 후의 잔차 신호 (b)와 다중대역 장구간 예측 후의 잔차 신호 (c)를 비교해보면 주기적 성분들이 상당히 제거되었음을 관찰할 수 있다.

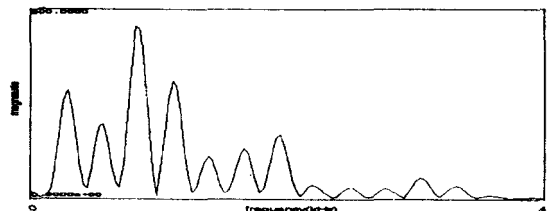
그림 5는 그림 4의 스펙트럼이다. 그림 6의 약 6초 길이의 한 문장에 대한 4.8 kbps SMBCELP 관련 잔차 신호들의 스펙트로그램(spectrogram)이다. 다중대역 장구간 예측 후 잔차 신호의 스펙트로그램 (b)는 장구간 예측 후 잔차 신호의 스펙트로그램 (a)에 비해서 하모닉 성분들의 에너지가 상당히 줄어들었다.

음을 알 수 있다.

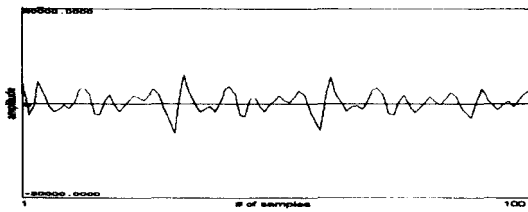
주관적 음질 평가에는 21세 전후의 남녀 2명이 각각 두 분장씩 발음한 4개의 음성 시료를 가지고 10명에 대해서 평가 시험을 행하였다. 평가 방법인 MOS (Mean Opinion Score)[5]를 구하기 위해 등급은 1에서 5까지 다섯 등급으로 나누고, 원음을 등급 '5'라고 가정하였다. 그 결과는 표 2에 나타나 있다. 그 결과를 보면 남녀화자 모두에 비해 4.8 kbps DoD CELP 보다 더 높은 점수를 얻었고, 4.4 kbps IMBE 보다는 약간 낮은 점수를 얻었다.



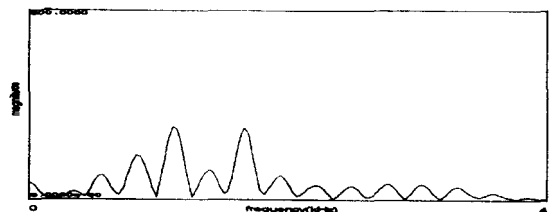
(a) 원 음성 신호



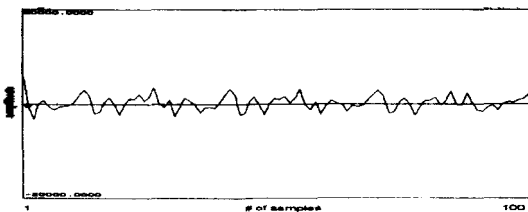
(a) 원 음성 신호의 스펙트럼



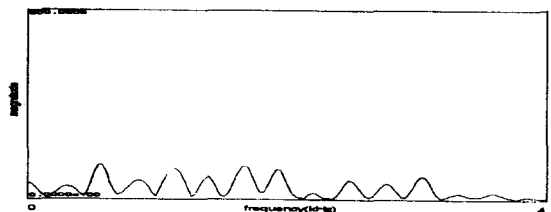
(b) 장구간 예측 후 잔차 신호



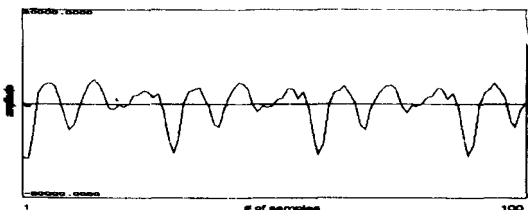
(b) 장구간 예측 후 잔차 신호의 스펙트럼



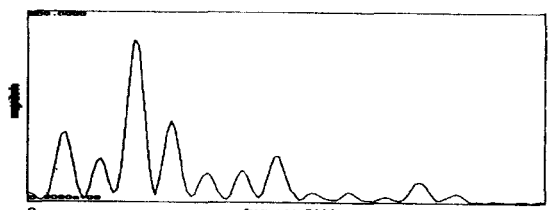
(c) 다중대역 장구간 예측 후 잔차 신호



(c) 다중대역 장구간 예측 후 잔차 신호의 스펙트럼



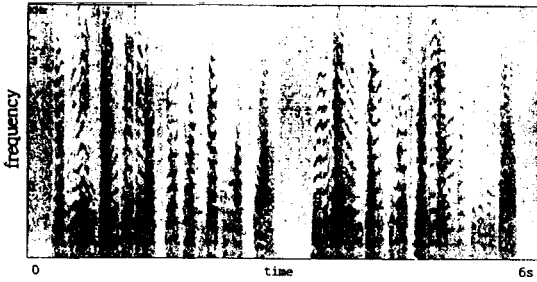
(d) 합성된 음성 신호



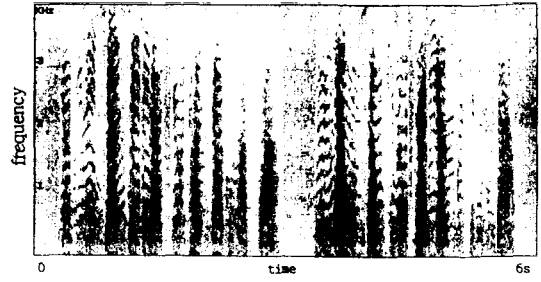
(d) 합성된 음성 신호의 스펙트럼

그림 4. 4.8 kbps SMBCELP 관련 신호

그림 5. 4.8 kbps SMBCELP 관련 신호의 스펙트럼



(a) 장구간 예측 후 잔차 신호의 스펙트로그램



(b) 다중대역 장구간 예측 후 잔차 신호의 스펙트로그램

그림 6. 4.8 kbps SMBCELP 관련 잔차 신호의 스펙트로그램

표 2 주관적 음질 평가(MOS) 결과

	문장1(여)	문장2(여)	문장1(남)	문장2(남)	평균
4.8 kbps DoD CELP	2.85	2.76	3.11	3.04	2.94
4.4 kbps IMBE	3.18	3.21	3.37	3.36	3.22
4.8 kbps SMBCELP	3.13	3.02	3.29	3.21	3.16

시료로 사용한 문장은 다음과 같다.

(문장1) “여기는 연세대학교 전자공학과 음향 음성 및 신호 처리 연구실입니다.”

(문장2) “대학가라는 말을 붙일 수 있는 거리는 어느 시중의 거리와는 차별될 수 있는 풍속도를 가지고 있어야 한다.”

V. 결 론

본 논문에서는 4.8 kbps 이하의 낮은 전송률에서 MBE 방법을 도입하여 CELP 부호화기의 장구간 예측 성능을 향상시키는 알고리즘과 파라메타의 추정 방법을 제안하였다.

제안된 방법에서는 기존의 CELP 방법으로 장구간 예측을 한 후, 그 잔차 신호를 에너지가 집중되어 있는 기본 주파수의 하모닉 간격으로 주파수 대역 분할하여 대역에 따라 일정하지 않은 스펙트럼 포락선을 추정하였다. 4.8 kbps 이하의 낮은 전송률로 구현하기 위해 다시 몇개의 대역들을 묶어서 대역수를 줄였다. 기본 주파수의 하모닉 간격에서 추정된 스펙트럼 포락선을 크기로 하고 기본 주파수의 하모닉을 주파수로 하는 정현파의 합으로 대역 적용 코드북 여기 신호를 만들어 이것으로 대역 장구간 예측을 하였다.

4.8 kbps의 전송률을 갖는 SMBCELP의 실험 결과, 장구간 예측 후의 잔차 신호에 남아 있던 주기적 성분들이 다중대역 장구간 예측으로 줄어들었음을 확인하였다. 다중대역 장구간 예측 후의 잔차 신호는 하모닉 성분들이 줄어들어서 통계적 코드북으로 여기 신호를 잘 모델링할 수 있다. 그러므로 제안된 방법에서는 CELP 방법에 MBE 방법을 결합하여 실제 여기 신호의 주파수 분포 특성을 잘 반영하였다.

제안된 방법의 한 예로서 4.8 kbps SMBCELP를 구현하였으며, 비교 기준 부호화기로 4.8 kbps DoD CELP와 4.4 kbps IMBE를 선정하여 음질을 비교 평가하였다. 주관적 음질 평가에서는 4.8 kbps DoD CELP 보다 우수하였고, 4.4 kbps IMBE와는 비슷하였다.

제안된 방법은 위상 모델링 함수와 실제 위상과의 오차가 큰 구간에서는 다중대역 장구간 예측 이득이 떨어지는 단점이 있으므로 주기적 다중대역 여기 신호를 정현파의 합으로 표현할 때, 위상을 더 잘 모델링하여 다중대역 장구간 예측 이득을 높이고, 프레임간 기본 주파수의 변화를 반영할 수 있도록 하는 것이 추후 과제가 될 것이다. 현재, 이러한 위상 모델링의 개선 방법의 한 가지로서 초기 위상 코드북을 통한 성능 향상 연구가 진행되고 있다.

참 고 문 헌

1. B.S. Atal, et al., eds., *Speech and Audio Coding for Wireless and Network Application*, Kluwer, 1993.
2. M.R. Schroeder, B.S. Atal, "Code-Excited Linear Prediction(CELP) High Quality Speech at Very Low Bit Rates," *Proc. ICASSP*, pp. 937-940, 1985.
3. J.S. Marques and Tribolet, "Pitch prediction with fractional delays in CELP coding," *Proc. EURO-SPEECH*, pp. 509-512, 1989.
4. S. Miki, K. Mano, H. Ohmuro, and T. Moriya, "A Pitch Synchronous Innovation CELP (PSI-CELP) Coder for 2.4kbit/s," *IEEE ICASSP*, II.113-II.116, 1994.
5. G. Yang, H. Leich and R. Boite, "Multiband Code-Excited Linear Prediction (MBCELP) for Speech Coding," *Signal Processing*, Vol. 31, pp. 215-227, 1993.
6. IMBE Vcoder description, DVSI, July 1993.
7. R.J. McAulay and T.F. Quatieri, "Speech Analysis-Synthesis Based on a Sinusoidal Representation," *IEEE Trans. ASSP*, Vol. ASSP-34, No. 4, August 1986.
8. J.P. Campbell, T.E. Tremain, V.C. Welch, "An Expandable Error-protected 4800 bps CELP Coders(U. S. Federal Standard 4800 bps Voice Coder," *IEEE ICASSP*, pp. 735-738, 1989.
9. Proposed EIA/TIA Wideband Speech Spectrum Standard, Qualcomm, 1992.
10. Y.Linde, A. Buzo, R.M. Gray, "An Algorithm for Vector Quantizer Design," *IEEE Trans. Com.*, Vol. Com-28, No. 1, pp. 84-95, 1980.

▲최 응 수(Yong-Soo Choi)



1993년 2월 : 연세대학교 공과대학 전자공학과 졸업 (공학사)
 1995년 2월 : 연세대학교 대학원 전자공학과 졸업(공학석사)
 1995년 3월~현재 : 연세대학교 대학원 전자공학과 박사과정

※주관심 분야 : 음성 부호화 및 응용, 실시간 신호처리 시스템

▲강 흥 구(Hong Goo Kang)



1980년 2월 : 연세대학교 공과대학 전자공학과 졸업 (공학사)
 1991년 2월 : 연세대학교 대학원 전자공학과 졸업(공학석사)
 1991년 3월~현재 : 연세대학교 대학원 전자공학과 박사과정

※주관심 분야 : 음성 부호화, 오디오 부호화, 실시간 신호처리 시스템

▲임 명 섭(Myoung-Soeb Lim)



1980년 2월 : 연세대학교 공과대학 전자공학과 졸업 (공학사)
 1982년 2월 : 연세대학교 대학원 전자공학과 졸업(공학석사)
 1990년 2월 : 연세대학교 대학원 전자공학과 졸업(공학박사)

1984년 1월~1985년 9월 : 대우통신 종합연구소 광통신 시스템 개발

1985년 9월~현재 : 한국전자통신연구소 이동통신 기술연구단 신호처리 연구실장

※주관심 분야 : 음성 부호화, 대역 확산 통신, 디지털 모뎀 및 신호처리

▲안 등 순(Dong-Soon Ahn)

한국음향학회지 Vol. 8, No. 5, Oct., 1989. 참조.

▲윤 대 희(Dae-Hee Youn)

한국음향학회지 Vol. 13, No. 1E, Jan., 1994. 참조.