

論文2001-38SP-5-13

주파수대역 정보를 이용한 가변률 IMBE-LP 음성부호화 알고리즘 (Variable Rate IMBE-LP Coding Algorithm Using Band Information)

朴萬鎬*, 裴建星**

(Man Ho Park and Keun Sung Bae)

요약

MBE(Multi-Band Excitation) 음성부호화 방식은 프레임 단위로 유/무성음을 구분하는 기존의 분석-합성 방식과는 달리 한 프레임 내에서의 주파수 영역을 여러 대역으로 나누고, 각 대역별로 유/무성음 구간을 판정하여 그에 맞는 여기신호를 이용하여 음성을 합성한다. 이러한 MBE 방식은 프레임 단위로 유/무성음을 구분하는 기존의 방식들이 갖는 합성음의 buzziness 영향이나 잡음이 섞인 음성을 분석할 때 생길 수 있는 유/무성음 판정 오류의 영향을 최소화함으로써 음질 향상을 이룰 수 있다. IMBE-LP 방식은 MBE 방식을 이용하여 2.4 kbps의 저전송률을 얻기 위한 음성부호화 알고리즘으로 MBE 모델에서 사용되는 각 대역별 스펙트럼 정보를 LP(Linear Prediction) 계수로 모델링 한다. 본 연구에서는 2.4 kbps IMBE-LP 알고리즘을 구현하고, 주파수대역 정보를 이용하여 분석프레임의 음성특성에 따라 LP차수를 달리 함으로써 전송률을 줄일 수 있는 방법을 제안하고 실험하였다.

Abstract

The Multi-Band Excitation(MBE) speech coder uses a different approach for the representation of the excitation signal. It replaces the frame-based single voiced/unvoiced classification of a classical speech coder with a set of such decision over harmonic intervals in the frequency domain. This enables each speech segment to be a mixture of voiced and unvoiced, and improves the synthetic speech quality by reducing decision errors that might occur on the frame-based single voiced and unvoiced decision process when input speech is degraded with noise. The IMBE-LP, improved version of MBE with linear prediction, represents the spectral information of MBE model with linear prediction coefficients to obtain low bit rate of 2.4 kbps. In this paper, we proposed a variable rate IMBE-LP vocoder that has lower bit rate than IMBE-LP without degrading the synthetic speech quality. To determine the LP order, it uses the spectral band information of the MBE model that has something to do with the input speech's characteristics. Experimental results are given with our findings and discussions.

* 正會員, 韓國電子通信研究員

(Electronics and Telecommunications Research Institute)

** 正會員, 慶北大學校 電子·電氣工學部

(Dept. of Electronic and Electrical Engineering
Kyungpook National University)

※ 본 논문은 1999년 정보통신부 대학기초연구지원 사업의 지원으로 수행되었으며, 지원에 감사드립니다.

接受日字:2001年5月16日, 수정완료일:2001年5月17日

I. 서론

MBE 음성부호화 방식은 기본적으로 유성음의 경우 기본주파수 및 고조파 성분을 정현파의 가중적 합으로 표현하는 sinusoidal coding에 기초한다. 전체적인 과정은 입력음성을 기본주파수의 고조파를 기본으로 하는 주파수 대역으로 나누어서 각 대역에 대한 유/무성음

여부를 판단하여 그 정보를 전송하며, 각 대역 내에 포함되는 고조파에 대한 스펙트럼 크기를 부호화한다. 이 정보에 기인하여 복호화기에서는 여기신호를 유성음인 경우 정현파 신호로, 무성음인 경우에는 잡음신호를 이용하여 각 대역별로 음성을 합성하게 된다^[1]. 그러므로 MBE 음성부호화기는 분석 프레임단위로 단순하게 유성음 혹은 무성음으로 판단하기 어려운 혼합된 부분을 적절히 표현할 수 있으므로 기존의 CELP계열의 음성 부호화기가 저전송속도에서 명료하고 자연스러운 음성을 합성하지 못한다는 단점을 극복할 수 있다. MBE 음성부호화기의 계산적 복잡도를 줄인 IMBE(Improved MBE)는 음성모델에 있어서는 MBE와 동일한 형태를 갖는데, IMBE 2.4 kbps의 경우 MBE 모델의 전체 전송률에서 2/3 가량을 차지하는 고조파의 스펙트럼 크기 정보를 LP 계수로 나타낸다. 일반적으로 가변 전송률을 가지는 음성부호화 방식은 각 음성 프레임을 음성의 특징에 따라 서로 다른 전송률로 전송하는 방식이다. 이러한 가변 전송률 방식의 장점은 적절한 전송률 선택 모델 적용을 통하여 합성음의 음질은 고정률 방식과 유사하게 유지하면서 평균적으로 낮은 전송률을 가진다는데 있다. 본 논문에서는 주파수대역 정보에 따라 실질적인 주파수 영역의 스펙트럼 크기 정보의 해상도가 달라진다는 점을 이용하여 고조파의 크기를 추정하기 위한 LP 모델의 차수를 가변함으로써 음질의 저하 없이 전송률을 낮출 수 있는 알고리즘을 제안하였고, 실험을 통해 제안된 알고리즘의 타당성을 확인하였다.

II. IMBE 음성 부호화기

1. MBE 음성모델

IMBE는 계산의 복잡도를 감소시키고 양자화 기법을 달리한 점을 제외하고는 MBE와 동일한 음성모델을 사용하므로 본 절에서는 MBE 음성 모델의 분석/합성 과정에 대하여 설명한다. 먼저, 음성을 발성학적 모델에서 관찰해보면 성대에서의 여기신호가 성도를 거치면서 발생하는 것으로 여기신호 성분과 성도의 전달특성으로 특징지어진다. 음성신호는 크게 나누어 여기신호가 성도의 특정 부분에서 공기의 흐름에 대한 장애가 일어나고 이로 인해 공기의 불규칙적인 흐름으로 인해 발생하는 무성음과 성대의 주기적인 진동으로 발생하는 유성음으로 구분된다. 이때 무성음의 경우 잡음과

같은 성질을 갖게 되고 유성음의 경우 신호의 주기성으로 인한 고조파 성분이 나타난다. 이와 같은 유/무성음의 여기신호가 성도를 지나게 되는데 이때 성도는 공진기와 같은 역할을 함으로써 여기신호의 스펙트럼을 변형시키게 된다. 유성음의 경우 대략 4개 정도의 공진점을 형성하게 되고 이 공진점들을 포먼트(formant)라 부른다.

유성음의 스펙트럼을 관찰하면 포먼트 정보를 가진 스펙트럼 상에 여기신호의 주기적인 고조파 성분이 실려있는 것을 볼 수 있다. 그러나 유성음의 스펙트럼에서도 주기적인 고조파만 나타나는 것이 아니라 진폭 스펙트럼의 일부분은 잡음 특성을 갖는 에너지가 존재하는 것을 볼 수 있다. 특히 유/무성음이 천이하는 혼합 구간일 경우에는 기본주파수에 의한 주기적인 고조파 성분들과 잡음 에너지에 의한 스펙트럼이 동시에 존재한다는 것을 알 수 있다. 이러한 음성신호 특성을 보다 더 잘 반영하기 위하여 MBE 방식에서는 프레임 단위의 유/무성음을 구분하는 기존의 LPC(Linear Predictive Coding) 방식과는 달리 음성신호의 스펙트럼을 주파수대역에서 기본주파수의 고조파 성분으로 구분한 후 각 고조파 성분에서의 에너지를 비교하여 각 주파수 대역에 대해 유/무성음 구간을 판별한다. 그림 1은 음성신호의 스펙트럼을 MBE 모델의 분석-합성 과정으로 나타낸 예를 보인 것이다.

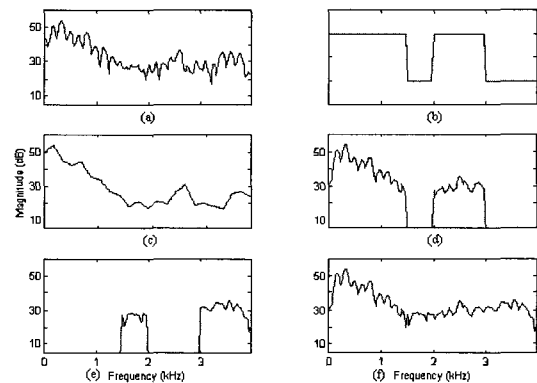


그림 1. MBE 음성모델의 예

- (a) 원음성의 스펙트럼 (b) 유/무성음 정보
- (c) 스펙트럼 포락선 (d) 유성음 스펙트럼
- (e) 무성음 스펙트럼 (f) 합성음 스펙트럼

Fig. 1. Illustration of MBE model.

- (a) Original spectrum (b) V/UV information
- (c) Spectral envelope (d) Voiced spectrum
- (e) Unvoiced spectrum (f) Synthetic spectrum

그림 1의 (a)는 원음성의 스펙트럼을, (b)는 MBE 모델에서 기본주파수 및 고조파 성분에 대해 추정된 스펙트럼 포락선을, (c)는 분석 프레임 내에서 주파수 대역별로 유/무성음을 판정한 결과를 나타낸 것이다. (d)는 유성음 구간으로 판정된 대역에 대해 정현파의 가중적 합으로 합성된 신호의 스펙트럼을, (e)는 무성음 구간으로 판정된 대역에 대해 잡음 여기신호를 이용하여 합성된 신호의 스펙트럼을, (f)는 유성음 및 무성음 대역을 합한 전체 합성음의 스펙트럼을 보인 것이다. 이러한 MBE 모델은 인간이 청각적으로 기본주파수 및 고조파가 지배적인 주파수 영역과 잡음과 같은 성질의 스펙트럼이 지배적인 주파수영역을 구별해 낼 수 있다는 사실에 기인한 것이다^[4]. 따라서 MBE 모델에서의 분석 방식은 단순한 프레임 단위로 하나의 유/무성음을 구분하여 사용하는 기존의 모델들이 갖는 합성 음질의 buzziness 영향이나 잡음이 섞인 음성을 분석할 경우 생길 수 있는 유/무성음 판단 오류를 최소화할 수 있다^[1,3].

2. IMBE의 주파수대역 정보 구조

초기 피치 추정 및 refinement 과정을 통하여 정밀한 기본주파수, ω_0 를 추정한 후 이 피치 정보에 의하여 식 (1)을 이용하여 현재 프레임의 고조파의 수를 결정하게 된다. 이때 추정되는 고조파의 수는 피치의 범위에 의하여 $9 \leq \hat{L} \leq 56$ 사이의 값이 된다^[4,5]. 식 (1)에 의하여 고조파의 수가 결정되면 각 대역의 하위와 상위 경계점을 나타내는 $\hat{\omega}_l$ 과 $\hat{\omega}_r$ 을 식 (2), (3)에 따라 얻게 된다. 한 프레임 내에서의 유/무성음 판정은 전체 주파수 대역을 식 (4)에 따라 \hat{K} 개의 주파수 대역으로 나눈 뒤 각 대역에 대하여 주기적 특성을 갖는 에너지와 잡음과 같은 특성을 갖는 에너지의 비를 이용하여 결정하게 된다.

$$\hat{L} = \lfloor 0.9254 * \lfloor \frac{\pi}{\omega_0} + 0.25 \rfloor \rfloor \quad (1)$$

$$\hat{\omega}_l = \frac{256}{2\pi} (l - 0.5)\hat{\omega}_0 \quad (2)$$

$$\hat{\omega}_r = \frac{256}{2\pi} (l + 0.5)\hat{\omega}_0 \quad (3)$$

$$\hat{K} = \begin{cases} \lfloor \frac{(\hat{L} + 2)}{3} \rfloor & \text{if } \hat{L} \leq 36 \\ 12 & \text{otherwise} \end{cases} \quad (4)$$

여기서 $\lfloor x \rfloor$ 는 x 를 넘지 않는 최대의 정수를 나타낸다. 이와 같은 구성을 통하여 표현되는 MBE 모델의

주파수 대역구조는 그림 2와 같이 표현 할 수 있다.

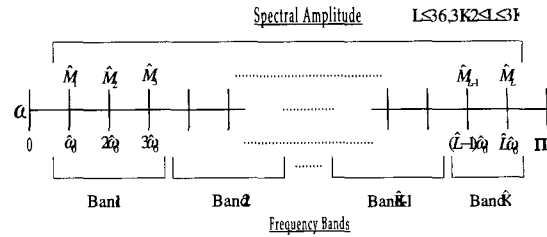


그림 2. MBE 모델의 주파수 대역 구조
Fig. 2. MBE frequency band structure.

3. IMBE 음성부호화기

MBE 음성모델에서 피치, 유/무성음 정보, 그리고 스펙트럼 크기 정보등의 인자가 양자화 되어서 전송되게 되는데 MBE와 IMBE는 이러한 인자의 양자화 방법에서 차이를 보인다. MBE 음성부호화기에서 고조파 크기정보를 양자화 할 때, 현재 프레임의 스펙트럼 크기는 이전 프레임과 현재 프레임의 정보를 이용하여 예측한 값과의 차이를 이용한다. 이렇게 구해진 스펙트럼 크기 잔차신호는 6개의 부대역으로 나누어지고 각 대역내의 스펙트럼 크기 잔차신호간의 자기상관관계는 DCT(Discrete Cosine Transform)를 이용하여 제거된다^[6]. MBE 음성부호화기에서는 이 계수들이 비적응적 비트 할당방법에 의하여 양자화 되게 되는데, 이러한 MBE의 비적응적 비트 할당방법을 개선하여 적응적 비트 할당방법을 도입하고 계산적 복잡도를 줄여서 성능을 향상시킨 모델이 IMBE(Improved MBE) 음성부호화기이다^[6-8].

표 1. 2.4 kbps IMBE-LP 음성부호화기의 비트 할당

Table 1. Bit allocation form of the 2.4 kbps IMBE-LP vocoder.

| Parameters | Bits/Frame | Bit Rate (kbits/s) |
|------------|------------|--------------------|
| Pitch | 8 | 0.4 |
| 10 LSFs | 26 | 1.3 |
| Energy | 5 | 0.25 |
| V/UV | 9 | 0.45 |
| Total | 48 | 2.4 |

IMBE 4.4 kbps 모델의 경우 전체 88 비트중 67~76 비트를 스펙트럼 크기정보에 할당하는데 이는 전체의

76~86%에 해당하는 양이 된다. IMBE-LP는 이러한 전체 비트중 스펙트럼 크기정보의 비중을 줄이기 위하여 스펙트럼 크기정보를 LPC모델을 통하여 나타낸 형태이며 2.4 kbps의 전송률을 가진다. 표 1은 2.4 kbps IMBE-LP의 비트 할당을 나타낸 것이다^[4].

Ⅲ. 가변률 IMBE-LP 알고리즘

1. 가변률 IMBE-LP

음성신호의 특징은 유성음, 무성음 그리고 잠음 혹은 묵음 구간으로 분류할 수 있다. 유성음의 경우 여러 개의 포먼트를 가지고 있으며 이러한 포먼트를 적절히 표현하기 위하여 높은 차수의 LP 모델이 필요하다. 반면에 무성음이나 묵음의 경우 스펙트럼 포락선이 평탄한 형태를 보이므로 적은 차수의 LP 모델로도 이러한 형태를 충분히 나타낼 수 있다^[9]. 그러므로 음성신호의 특징에 따라 적절한 LP 모델의 차수를 선택해 줌으로써 합성음의 음질은 고정적인 차수를 사용할 때와 같이 비슷하게 유지하면서 스펙트럼 크기 정보의 표현을 위한 평균적인 전송률을 줄이는 이득을 가져올 수 있다.

IMBE의 경우 피치정보에 의해서 주파수 대역의 수가 정해지며 각 대역은 3개의 스펙트럼 크기 정보를 포함하게 된다. 유성음인 경우 변화가 심한 포락선 구조를 가지며 구해진 피치에 의하여 전체 주파수 대역을 3~12개의 대역으로 구분하여 전체 대역에 대한 포락선 정보를 최소 9개, 최대 57개의 스펙트럼 크기 정보로 표현한다. 무성음 혹은 묵음의 경우 평탄한 포락선 모양을 나타내며 최소 대역수인 3개로 구분되는데 전체 주파수 대역을 9개의 스펙트럼 크기 정보로 표현한다. 이것은 나누어지는 대역의 수가 전체 주파수 대역에 대한 스펙트럼의 해상도를 결정하는 것으로 볼 수 있으며, 이러한 대역 정보를 이용하여 LP 차수를 결정할 경우 수신단에서 전송된 LP 차수를 피치 정보에 의하여 역추정할 수 있으므로 일반적으로 QCELP^[10]와 같은 가변률 음성부호화기에서 발생하는 부가적인 전송모드 정보에 대한 비트 할당부분을 제거 할 수 있다는 장점이 있다.

그림 3은 유성음과 무성음 구간의 음성 한 프레임에 대하여 유성음인 경우 10차의 LP 모델을 이용하여, 무성음과 묵음인 경우 6차의 LP 모델을 이용하여 스펙트럼

포락선을 추정한 예를 보인 것이다.

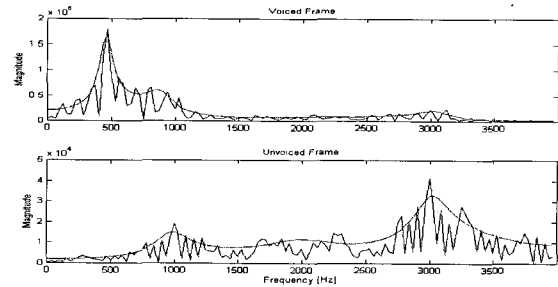


그림 3. 유성음/무성음 한 프레임에 대한 가변적 LP Model의 적용 예

Fig. 3. Example of variable order LP model to a V/UV frame.

2. LP 모델 차수의 선택

본 논문에서 제안한 가변률 IMBE-LP 알고리즘은 4.4 kbps의 전송률을 가지는 IMBE 알고리즘을 이용하여 추정된 주파수대역 정보에 따라 LP 모델의 차수를 10, 8, 그리고 6으로 가변하는 방법을 사용하였다. 주파수대역 정보에 따른 LP 모델의 차수는 분석/합성 실험을 통하여 실험적으로 결정하였다. 고조파의 수가 9~15개인 경우 상대적인 스펙트럼 해상도는 떨어지게 되고 이때 LPC 모델의 차수는 6차를 할당하였다. 그리고 고조파의 수가 16~27개인 경우는 LP 모델의 차수를 8차로 하였으며 가장 스펙트럼의 해상도가 높은 경우인 고조파의 수가 28~57개인 때는 LP 모델의 차수를 10차로 설정하였다. 그림 4는 본 연구에서 제안한 가변률을 적용한 IMBE-LP 부호화기의 블럭도이다.

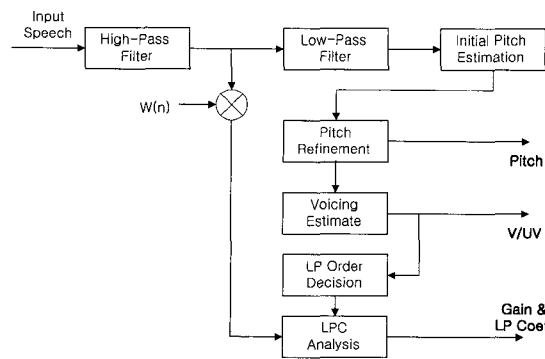


그림 4. 가변률 IMBE-LP 부호화기의 블럭도

Fig. 4. Block diagram of variable order IMBE-LP encoder.

LP 모델의 추정은 Levinson 알고리즘을 이용하였는

데 ill-condition의 발생 확률을 줄이기 위하여 자기상관 계수에 대하여 식 (5)와 같이 bandwidth expansion을 적용하였다.

$$w_{lag}(k) = \exp\left[-\frac{1}{2}\left(\frac{2\pi f_0 k}{f_s}\right)^2\right] \quad k = 1, \dots, 10 \quad (5)$$

여기서 f_0 는 기본주파수를, f_s 는 표본화 주파수를 나타내며 k 는 LPC 계수의 색인을 말한다. 이때 적용되는 bandwidth expansion의 크기는 LP 모델의 차수에 따라 10차의 경우에는 120Hz, 8차의 경우는 60Hz bandwidth expansion을 적용하였으며 6차의 경우 무성음/목음과 같이 평탄한 스펙트럼의 경우이므로 bandwidth expansion을 생략하였다.

IV. 실험 및 고찰

본 논문에서 제안한 방법을 남성화자가 발성한 2개의 한국어 문장과 3개의 영어 문장에 대해 실험하였다. 실험에 사용된 16비트로 양자화된 음성데이터를 표 2에 나타내었다.

표 2. 실험에 사용된 음성 데이터
Table 2. Processing speech data for simulation.

| 문장 번호 | 내 용 |
|-------|----------------------------------|
| 1 | She said that she adores the man |
| 2 | Card games are fun to play |
| 3 | We saw the ten fink fish |
| 4 | 한국의 가을하늘은 참으로 맑고 푸르롭니다. |
| 5 | 하늘을 날고자 하는 인간의 욕망은 끝이 없습니다. |

그림 5는 가변률 IMBE-LP에서 실험 문장에 대한 LP 모델의 차수 선택 결과를 보인 것이다. 그림에서 보는 바와 같이 음성의 성질에 따라 적절한 LP 차수 배분이 나타나며 명확한 유성음 구간의 경우 대체로 높은 차수가, 그리고 무성음 구간과 목음 구간에서는 낮은 차수가 선택됨을 볼 수 있다. 문장의 시작부분에서 LP 차수를 높게 배정한 것은 잘못된 피치검출 때문임을 확인하였다.

표 3은 실험 음성에 대해 일반적인 2.4 kbps

IMBE-LP와 가변률을 적용한 IMBE-LP사이의 비트 할당률을 보이고 있다. IMBE-LP 방식에서 스펙트럼 크기 정보를 나타내는데 사용되는 비트수는 1.3 kbps인데 비해 가변률 방식에서는 평균 1.092 kbps로 한 프레임당 평균적인 비트 할당량의 감소는 4.16 비트이고 부가적인 정보는 필요로 하지 않는다. 또한 실제적인 통화 상황에서 통화시간의 60% 정도가 목음인 것을 감안하면 평균적인 비트 할당량은 더욱 줄어 들 수 있음을 알 수 있다.

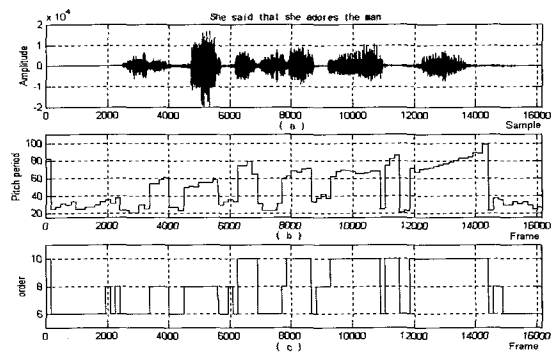


그림 5. 실험 문장 1
(a) 원 음성신호 (b) 피치 궤적
(c) LPC 차수선택 결과

Fig. 5. Example of test speech No. 1.
(a) original speech (b) Pitch contour
(c) Results of LPC model order selection

표 3. 가변률과 고정률 IMBE-LP의 스펙트럼 정보의 전송률 비교

Table 3. Bit-rate comparison of spectral information between variable rate and fixed rate IMBE-LP model

| Model 차수 | 빈도 (%) | LPC bits | 전송률 (Kbps) | 평균 전송률 (Kbps) |
|----------|--------|----------|------------|---------------|
| 가변 | 10 | 38 | 0.494 | 1.092 |
| | 8 | 20 | 0.220 | |
| | 6 | 42 | 0.378 | |
| 10 (고정) | 100 | 26 | 1.3 | 1.3 |

표 4는 실험 음성들에 대한 4.4 kbps IMBE와 2.4 kbps IMBE-LP, 가변률 IMBE-LP 방식에 대하여 주관적인 성능평가와 높은 상관 관계를 갖는 cepstral distortion 측정 결과를 나타낸 것이다. 이때 cepstral

distortion 측정은 후처리 이전의 각 합성음을 대상으로 14차의 LPC-cepstrum을 이용한 방법으로 측정하였다. 표 4에서 보는 바와 같이, IMBE-LP 방식과 본 논문에서 제안한 가변률 IMBE-LP 방식에 대하여 cepstral distortion이 유사한 값을 보이는 것을 알 수 있다.

MOS 실험은 10여명의 청취자를 대상으로 표 1에 제시된 실험 음성을 사용하여 비공식적으로 시행하였다. 표 5는 각 방식에 의한 후처리 이전의 합성음에 대한 MOS 테스트 결과를 보인 것이다. 적절한 후처리를 통한 음질을 향상한 음성에 대한 INMARSAT-M 4.15 kbps의 공식적인 MOS 점수가 3.3이고 2.4 kbps MBE-LPC의 경우 3.0인 것을 고려하면 MOS 점수가 신뢰할 만한 영역에 포함되어 있으며^[4], 기존의 방식에 대하여 가변률 방식의 음질이 크게 차이를 보이지 않음을 알 수 있다. 따라서, 위의 결과들을 종합하여 보면 고정률 방식과 가변률 방식에 있어서 음질에 그다지 큰 차이를 나타내지 않음을 알 수 있다.

표 4. 4.4 kbps IMBE, IMBE-LP, 가변률 IMBE-LP 방식의 Cepstral Distortion 비교

Table 4. Cepstral distortion comparison between 4.4 kbps IMBE, IMBE-LP, variable IMBE-LP within the corresponding frames.

| 구 분 | 4.4 Kbps | LP(고정) | LP(가변) |
|-----|----------|--------|--------|
| 평 균 | 0.30 | 0.31 | 0.33 |
| 분 산 | 0.35 | 0.29 | 0.34 |

표 5. 4.4 kbps IMBE, IMBE-LP, 가변률 IMBE-LP 방식의 MOS 결과 비교

Table 5. The MOS results of 4.4 kbps IMBE, IMBE-LP, Variable IMBE-LP.

| 구 분 | 4.4 Kbps | LP(고정) | LP(가변) |
|-----|----------|--------|--------|
| MOS | 3.0 | 2.6 | 2.4 |

V. 결 론

본 연구에서는 IMBE-LP 음성부호화 방식에서 주파수대역 정보를 이용하여 가변적 전송률을 갖는 알고리즘을 제안하였다. 고정 전송률의 IMBE-LP 방식에 비하여 가변 전송률 방식은 IMBE 알고리즘에서 전송률

의 대부분을 차지하는 스펙트럼 크기 정보 전송의 부담을 감소시키는 것이다. 제안한 방법은 고정 전송률 방식에서는 프레임당 26 비트의 정보전송이 필요한데 비하여 프레임당 평균 21.84 비트의 전송률을 나타내어 프레임당 평균 4.16 비트의 정보 전송량을 감소시킬 수 있었다. 또한 cepstral distortion 및 MOS 실험 결과로써 기존 고정률 방식에 대한 가변률 방식의 음질 저하가 크지 않음을 확인할 수 있었다. 앞으로 보다 개선된 합성음의 음질을 보장받기 위해서는 적절한 후처리 과정의 도입과 음성 특징에 따른 LP 모델 차수의 배정에 관한 연구가 계속 되어야 한다.

참 고 문 헌

- [1] D. W. Griffin and J. S. Lim, "Multiband Excitation Vocoder," IEEE Transactions on ASSP, Vol.36, No.8, August 1988.
- [2] M. S. Brandstein et. al., "A Real-Time Implementation of The Improved MBE Speech Coder," Proc. of ICASSP, pp.5~8, April 1990.
- [3] INMARSAT-M Voice Coding System Description (DRAFT version 1.3). February 1991.
- [4] A. M. Kondoz, DIGITAL SPEECH coding for low bit rate communication systems, John Wiley and Sons.
- [5] Digital Voice System, Inc. "APCO Project 25 Vocoder Description," July 1993.
- [6] C.V. Pavlovic, M. Rossi, and R. Espesser, "Use of the Magnitude Estimation Technique for Assessing the Performance of Text-to-Speech Synthesis System," J. Acoust. Soc. Am., Vol. 87, pp. 373-381, 1990.
- [7] D.B. Pisoni, B.G. Greene, and H.C. Nusbaum, "Some Human Factors Issues in the Perception of Synthetic Speech," Proc. Speech Tech '85, New York, pp. 57-61, 1985.
- [8] D.B. Pisoni, H.C. Nusbaum, and B.G. Greene, "Perception of Synthetic Speech Generated by Rule," Proceedings IEEE, Vol. 73, pp. 1665-1676, 1985.

[9] Pasi Ojala and Ari Lakaniemi, "Variable Model Order LPC Quantization," Proc. of the IEEE ICASSP, Vol.1, pp.49-52, May 1998.

[10] QUALCOMM Inc. TIA/EIA IS-96 Interim

Standard. "Speech service option standard for wideband spread spectrum digital cellular system," April 1993.

저 자 소 개



朴萬鎬(正會員)
 1998년 2월 경북대학교 전자공학과 학사. 2000년 2월 경북대학교 전자공학과 석사. 2000년 5월~현재, 한국전자통신연구원 무선방송 기술연구소 ※ 주관심분야 : 디지털 신호처리, 음성부호화 등



裴建星(正會員)
 1977년 2월 서울대학교 전자공학과 학사. 1979년 2월 한국과학기술원 전기 및 전자공학과 석사. 1989년 5월 University of Florida 공학박사. 1979년 3월~현재, 경북대학교 전자·전기공학부 교수. ※ 주관심분야 : 음성분석 및 인식, 디지털 신호처리, 디지털 통신, 음성부호화, 웨이브렛 이론 등