

저전송률 코드여기 선형 예측 부호화기를 위한 선택적 대역 하모닉 모델 기반 여기신호 개선 알고리즘

Excitation Enhancement Based on a Selective-Band Harmonic Model for Low-Bit-Rate
Code-Excited Linear Prediction Coders

이 미 숙* · 김 홍 국** · 최 승 호*** · 김 도 영*
Mi Suk Lee · Hong Kook Kim · Seung Ho Choi · Do Young Kim

ABSTRACT

In this paper, we propose a new excitation enhancement technique to improve the speech quality of low bit-rate code-excited linear prediction (CELP) coders. The proposed technique is based on a harmonic model and it is employed only in the decoding process of speech coders without any additional bits. We develop the procedure of harmonic model parameter estimation and harmonic generation, and apply this technique to a current state-of-the-art low bit rate speech coder, ITU-T G.729 Annex D. Also, its performance is measured by using the ITU-T P.862 PESQ score and compared to those of the phase dispersion filter and the long-term postfilter applied to the decoded excitation. It is shown that the proposed excitation enhancement technique can improve the quality of decoded speech and provide better quality for male speech than other techniques.

Keywords: excitation enhancement(여기신호 향상), postfilter(후처리 필터), CELP coder(CELP 부호화기)

1. 서 론

음성신호를 압축 또는 부호화할 때는 필연적으로 부호화 잡음이 발생하며, 특히 낮은 비트율(bit-rate)로 부호화할 때는 부호화 잡음이 증가한다. 후처리 필터(postfilter)는 이러한 부호화 잡음이 존재하는 경우에, 수신 측에서 인간의 청각특성을 이용하여 음질을 개선하는 방법이다. 후처리 필터는 주로 스펙트럼 포락선이나 여기신호를 개선하는 알고리즘들의 형태로 개발되어 사용되고 있다 [1]. 스펙트럼 포락선을 개선시키는 방법 중 하나는 단구간 후처리 필터(short-term postfilter)로 스펙트럼의 피크(peak) 부분은 강조하고 널(null) 부분은 더욱 깊게 하여 스펙트럼의 널 부분에서 발생하는 잡음에 의한 머플링 효과를 감소시키는 것이다 [2]. 또한 8 kbit/s 미만의 전송률을 갖

* 한국전자통신 연구원 광대역통합망 연구단

** 광주과학기술원 정보통신공학과

*** 서울산업대학교 전자정보공학과

는 코드여기 선형 예측 (code-excited linear prediction: CELP) 유형의 음성 부호화기에서는 적은 수의 펄스로 여기신호를 모델링함으로써 거친 소리가 발생하게 되는 데, 이는 phase dispersion 필터를 사용하여 완화시킬 수 있다 [1][3]. 그럼에도 불구하고 원래의 음성신호에 비해 하모닉 구조가 제대로 복원되지 않기 때문에 재생된 음성신호에 왜곡이 발생한다. 일반적으로, 복호화된 음성신호의 하모닉 구조를 개선하기 위해서 하모닉 후처리 필터 또는 피치 후처리 필터를 사용하기도 한다 [2]. 그러나 이들 방법은 전 주파수 대역에 걸쳐 하모닉 구조를 변경하기 때문에 제대로 복원된 낮은 주파수 대역에 존재하는 하모닉 구조도 변경시킬 수 있는 단점이 있다.

본 논문에서는 하모닉 모델링 방법에 기반을 둔 여기신호 개선 알고리즘을 제안한다. 이 알고리즘은 낮은 주파수 대역에 있는 하모닉 구조는 그대로 유지하면서 높은 대역에 존재하는 하모닉 구조를 개선시킨다.

본 논문의 구성은 다음과 같다. 먼저 제 2 장에서는 여기신호 개선을 위한 기존 알고리즘들에 대해 간단히 살펴본다. 제 3 장에서는 제안된 하모닉 모델 기반의 여기신호 개선 알고리즘에 대해 자세히 설명한다. 그리고 제 4 장에서는 본 논문에서 제안하고자 하는 하모닉 모델 기반 여기신호 개선 알고리즘과 기존 여기신호 개선 알고리즘을 G.729 Annex D에 적용한 결과에 대해 논의하고, 제 5 장에서 결론을 맺도록 한다.

2. 기존의 여기신호 개선 알고리즘

저 전송률 음성 부호화기에서는 고정 코드북의 펄스 수가 작기 때문에 발생하는 문제점을 완화시키기 위해서 phase dispersion 알고리즘이나 피치 필터와 같은 여기신호 개선 알고리즘을 사용할 수 있다. 그림 1은 일반적인 CELP 구조의 복호화기에 대한 블록도이다. 합성 필터의 입력신호는 적응 코드북 (adaptive codebook)과 고정 코드북 (fixed codebook)에 의해 만들어진다. 즉, 복호화된 여기신호 $e(n)$ 은

$$e(n) = g_p x(n) + g_c c(n) \quad (1)$$

으로 표현되며, $x(n)$ 과 $c(n)$ 은 각각 적응 코드북과 고정 코드북에 의해 복원된 신호이고 g_p 와 g_c 는 각 코드북의 이득이다. 일반적으로는 $e(n)$ 을 합성필터의 입력으로 사용하지만 저 전송률 음성 부호화기에서는 $e(n)$ 을 개선시킨 $e_h(n)$ 을 합성 필터의 입력으로 사용한다. 본 절에서는 $e(n)$ 을 개선하여 $e_h(n)$ 을 구하는데 사용되는 phase dispersion 알고리즘과 pitch filtering 알고리즘에 대해 간단히 살펴보기로 한다.

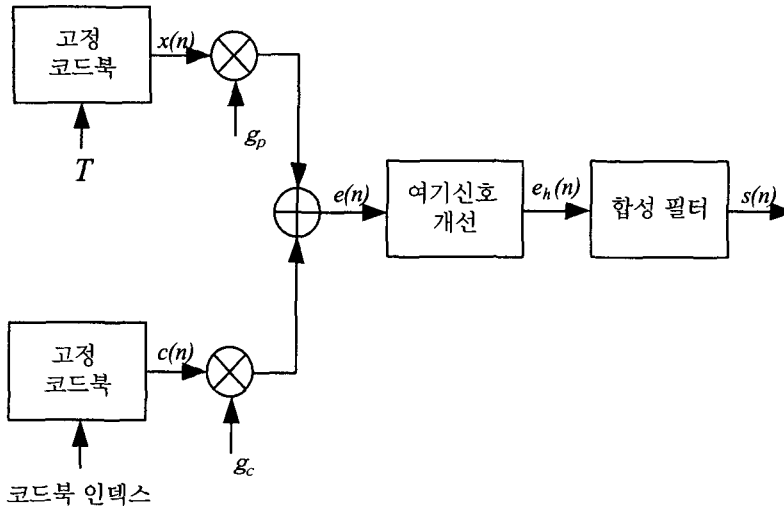


그림 1. 여기신호 개선 알고리즘을 갖는 CELP 복호화기

2.1 Phase Dispersion

CELP 구조의 음성 부호화기에서 전송률을 줄일 수 있는 가장 간단한 방법은 고정 코드북의 펄스 수를 줄이는 것이다. 그러나 고정 코드북의 펄스 수가 적어지면, 준 주기적인 성분이 복호화된 음성신호에 더해지는 것과 같은 효과를 주어 음질을 저하시키는 원인이 된다. 이러한 영향은 특히 3,000 Hz 이상에서 두드러지며 유성음보다는 무성음이나 배경잡음과 같은 신호에서 크게 나타난다. 고정 코드북의 펄스 수가 적기 때문에 나타나는 이러한 왜곡은 위상을 dispersion함으로써 완화시킬 수 있다. 다음 식은 phase dispersion 알고리즘을 이용하여 $e(n)$ 으로부터 $e_h(n)$ 을 구하는 과정을 나타낸다.

$$e_h(n) = g_p x(n) + g_c c(n) * h_p(n) \quad (2)$$

여기서 $h_p(n)$ 은 phase dispersion 필터의 임펄스 응답이다. 식 (2)에서와 같이 $h_p(n)$ 은 고정 코드북에만 적용되며 음성 신호의 종류에 따라 특성이 달라진다. 유성음의 경우에는 위상을 너무 많이 수정하면 잡음이 섞인 듯한 소리가 발생할 수 있다. 따라서 분석하고자 하는 프레임을 유성음과 무성음으로 분류한 후 그 결과에 따라 phase dispersion의 정도를 다르게 한다. CELP형 부호화기에서는 적용 코드북의 이득 값에 따라 음성신호를 세 가지 유형으로 분류하고, 각 유형에 따라 필터의 cut-off 주파수와 위상 수정을 위한 최대 각도가 선택된다 [1]. 예를 들면, 피치 이득이 0.9 이상이면 phase dispersion을 하지 않고, 피치 이득이 0.5 이하이면 2,000 Hz에서부터 π 만큼, 0.5에서 0.9 사이에서는 3,000Hz부터 $\pi/2$ 만큼 위상을 dispersion시킨다 [1][3].

2.2 Pitch Filtering

피치 필터는 피치 고조파들 사이에 있는 주파수 성분을 감쇄시켜 피치의 주기성을 강조하기 위해 사용하는 필터이다. ITU-T G.729 Annex D에서는 합성된 음성신호의 음질을 향상시키기 위해 식 (3)과 같은 형태의 피치 필터를 장구간 후처리 필터 (long-term postfilter)와 함께 사용한다. 식 (3)의 필터는 기본적으로 comb filter의 확장된 형태에 해당하며, 여기신호의 주기성 향상에도 사용할 수 있다 [4]. 즉,

$$H_p(z) = \frac{1 + r_p g_l z^{-T}}{1 + r_p g_l} \quad (3)$$

여기서 g_l 은 적응 코드북 이득에 비례하는 값이고 r_p 는 피치 필터링의 정도를 나타내는 값으로 G.729 Annex D에서는 0.5를 사용하고 있다. T 는 수신된 피치 값을 기반으로 새롭게 구한 피치 값으로 두 단계로 구해진다. 먼저, 첫 번째 프레임의 피치 값으로부터 ± 1 에 해당하는 구간에서 복호화된 여기신호로부터 정수형 피치를 구한다. 그리고 이렇게 구해진 피치 근처에서 1/8의 분해능을 갖는 피치를 다시 구한다.

3. 하모닉 모델 기반 여기신호 개선 알고리즘

본 논문에서 제안하고 있는 하모닉 모델 기반 여기신호 개선 기술은 그림 2와 같은 과정에 의해 여기신호의 고주파 대역에 있는 하모닉 성분을 개선한다. 복호화된 여기신호 $r(n)$ 과 피치 값 T_0 를 이용하여 개선된 여기신호 $r_p(n)$ 을 발생시키며, 주파수 영역에서 개선된 여기신호의 스펙트럼 $R_p(\omega)$ 는 식 (4)와 같이 표현된다.

$$R_p(\omega) = \begin{cases} R(\omega), & \omega < \omega_l \text{ or } \omega > \omega_h \\ (1 - G_p(\omega))R(\omega) + G_p(\omega)R_h(\omega), & \omega_l \leq \omega \leq \omega_h \end{cases} \quad (4)$$

여기서 $G_p(\omega)$ 는 $r(n)$ 과 $r_p(n)$ 의 이득을 정규화하기 위한 것이다. 위의 식을 보면 ω_l 부터 ω_h 사이의 주파수 대역에 있는 하모닉 성분만을 개선시키고 그 외의 주파수 대역에서는 원래 여기신호의 하모닉을 그대로 사용한다. 또한 계산량을 줄이기 위해 식 (4)는 시간영역에서 부프레임 단위로 다음 식 (5)와 같이 실행된다. 우선, $0 \leq n < N_s$ 구간에서 $r_h(n)$ 은 다음과 같이 표현된다.

$$r_h(n) = \sum_{m=M_l(T_0)}^{M_h(T_0)} A_m C_m(n; T_0, n_m), \quad (5)$$

$$C_m(n; T_0, n_m) = \cos(n + n_m) \frac{2\pi m}{T_0} \quad (6)$$

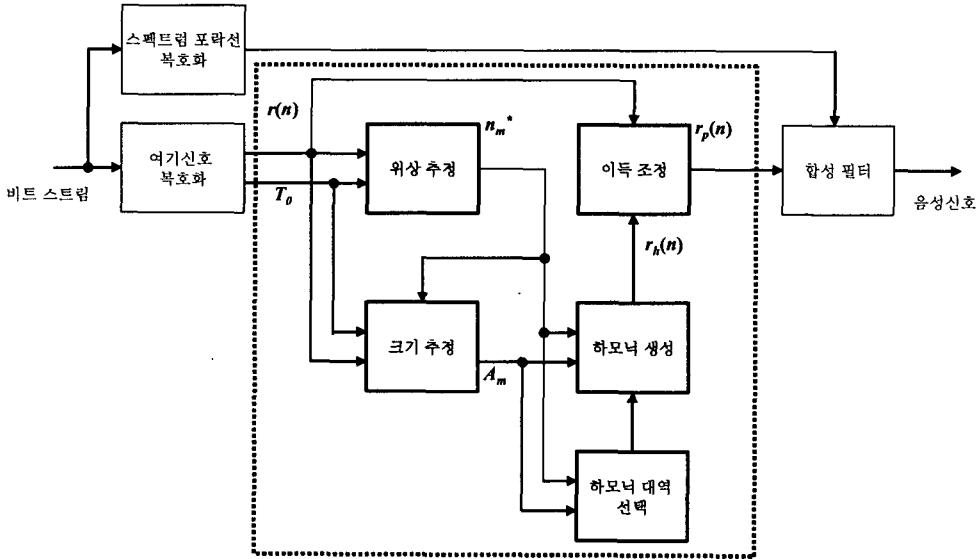


그림 2. 하모닉 모델 기반 여기신호 개선 과정을 갖는 음성 부호화기의 복호화부

여기서 N_s 는 부프레임의 크기이고 T_0 는 적용 코드북의 피치주기이며, n_m 과 A_m 은 각각 현재 부프레임에 대한 m 번째 사인파의 추정된 첫 번째 위상과 크기이다. 그리고 $M_l(T_0)$ 와 $M_h(T_0)$ 는 주어진 T_0 에서 ω_l 과 ω_h 에 해당하는 하모닉 수이다. 결국 시간영역에서의 하모닉 개선은 다음과 같이 수행된다.

$$r_p(n) = r(n) + g_p r_h(n) \quad (7)$$

여기서 g_p 는 $G_p(\omega)$ 에 해당하는 하모닉 개선을 위한 시간영역 이득이다. 식 (5)와 (6)에서의 위상 (n_m)과 크기 (A_m)는 다음 세부 절에서 구한다.

3.1 위상 추정

m 번째 사인파의 위상 파라미터인 n_m 은 입력 여기신호와 식 (6)에서 정의된 사인파들 간의 상관도를 기반으로 구한다. 즉, 주어진 m 에 대해서 n_m 는 $[0, \lfloor T_0/m \rfloor]$ 영역에 포함되며 추정된 시간 지연 n_m^* 은 다음의 수식에 의해서 얻어진다. 여기서 $\lfloor x \rfloor$ 는 x 보다 작거나 같은 정수를 의미한다.

$$n_m^* = \operatorname{argmax}_{0 \leq k \leq \lfloor T_0/m \rfloor} \sum_{n=0}^{N_T-1} r(n) C_m(n; T_0, k), \quad M_l(T_0) \leq m \leq M_h(T_0). \quad (8)$$

식 (8)은 모든 m 에 대해서 n_m^* 를 추정하기 위해서 모든 부프레임마다 $(M_h(T_0) + M_l(T_0) + 1)$ 회 반복된다.

3.2 크기 추정

각 하모닉 주파수의 크기는 $r(n)$ 과 $r_h(n)$ 사이의 최소평균자승오차(minimum mean-square error)를 기반으로 수행되며 이렇게 함으로써 스펙트럼이 피크를 가지는 주파수를 보전할 수 있다. A_m 은 다음 식의 오차를 최소화함으로써 얻어진다.

$$E = \sum_{n=0}^{N_T-1} (r(n) - r_h(n))^2 = \sum_{n=0}^{N_T-1} (r(n) - \sum_{m=M_l(T_0)}^{M_h(T_0)} A_m C_m(n; T_0, n_m))^2. \quad (9)$$

위 수식을 A_m 에 대해 미분함으로써 A_m 의 최적 값인 A_m^* 는 다음과 같다.

$$A_m^* = \frac{\sum_{n=0}^{N_T-1} r(n) C_m(n; T_0, n_m^*)}{\sum_{n=0}^{N_T-1} C_m^2(n; T_0, n_m^*)} \quad (10)$$

여기서 $N_T (\leq N_s)$ 는 T_0 의 정수배가 N_s 에 근접한 가장 큰 정수이다. 식 (10)에서 사인파들의 직교성(orthogonality)을 다음과 같이 가정된다. 즉,

$$\sum_{n=0}^{N_T-1} C_m(n; T_0, n_m^*) C_k(n; T_0, n_m^*) = \delta(m-k), \quad 1 \leq m, k \leq M(T_0). \quad (11)$$

여기서 $M(T_0) = T_0/2$ 이다. 그리고 $M_h(T_0)$ 부근에서 스펙트럼이 지나치게 개선되는 것을 방지하기 위해 식 (12)와 같은 방법을 취한다.

$$\overline{A}_m^* = (1 - C_p \frac{m - (M_l(T_0) - 1)}{M(T_0)}) A_m^*, \quad M_l(T_0) \leq m \leq M_h(T_0) \quad (12)$$

여기서 $M(T_0) = M_h(T_0) - M_l(T_0) + 1$ 이고 C_p 는 고주파수에서의 신호 크기를 조절하기 위한 파라미터이며, 실험적으로 0.5로 고정되었다.

3.3 하모닉 대역 선택

본 논문에서는 여기신호의 하모닉 성분을 개선하기 위한 상위 주파수 ω_h 는 $\lfloor T_0/2 \rfloor - 1$ 로 고정하였다. 그리고 하위 주파수 ω_l 은 매 부프레임마다 다음과 같이 적응적으로 결정하였다. 앞의 3.1절에서 구한 위상 추정 값과 앞의 3.2절에서 구한 크기 추정 결과를 이용하여 m 번째 하모닉 성분을 $h_m(n)$ 이라 할 때, 이것과 원래의 여기신호 $r(n)$ 과의 정규화된 상관도 ρ_m 을 식 (13)과 같이 구할 수 있다.

$$\rho_m = \frac{\sum_{n=0}^{N_i-1} r(n)h_m(n)}{\sqrt{\sum_{n=0}^{N_i-1} r^2(n) \sum_{n=0}^{N_i-1} h_m^2(n)}} \quad (13)$$

이 식으로부터 $\sum_{k=1}^{m^*} (1 - \rho_k)$ 이 임계치를 넘기 시작하는 m^* 을 최저 주파수의 하모닉 인덱스로 결정한다. 다시 말해서, 식 (14)에 의해 구해진 $D(m^*)$ 값이 임계치 D_{thres} 를 넘기 시작하는 m^* 을 최저 주파수의 하모닉 인덱스로 결정한다.

$$D(m^*) = \sum_{k=1}^{m^*} (1 - \rho_k) > D_{thres} \left(= \lambda \sum_{m=1}^{M_i(T_0)} (1 - \rho_m) \right) \quad (14)$$

여기서 λ 는 객관적 음질척도인 ITU-T 표준 P.862 “perceptual evaluation of speech quality” (PESQ) [6] 값을 최대로 하는 파라미터로, 이 값은 실험에 의해 0.85로 정했다. 그림 3은 남성화자가 발성한 유성음 신호에 대해 음성신호, 여기신호, 여기신호의 크기 스펙트럼, 여기신호간의 스펙트럼 차, 그리고 $D(m^*)$ 값을 보여주고 있다. 그림에서와 같이 $D(m^*)$ 값은 각각의 주파수별 여기신호에 대한 주파수 값 차이가 축적되는 형태로 증가하며, $D(m^*)$ 이 정해진 임계치를 초과하는 시점을 ω_l 로 선택하여 ω_l 로부터 ω_h 사이에 존재하는 하모닉 성분만을 개선시킨다.

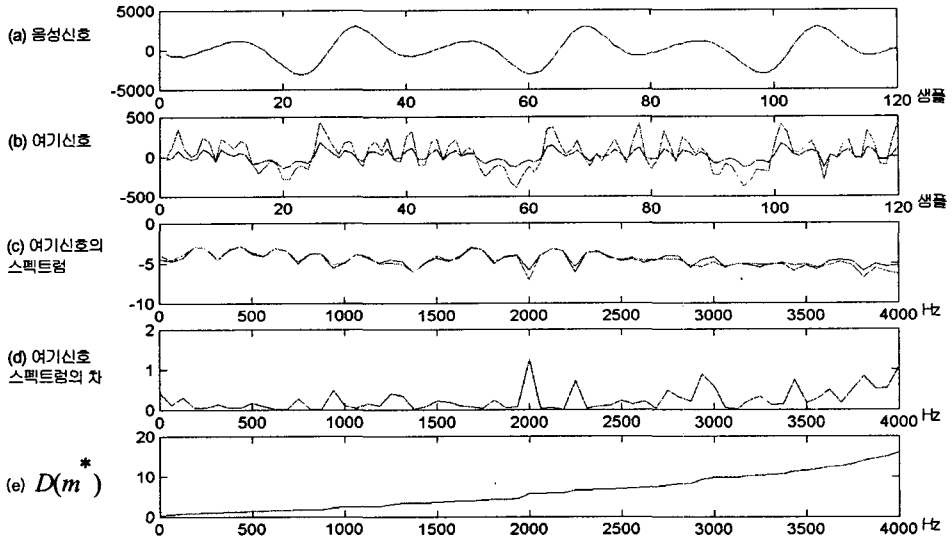


그림 3. 여기신호 개선을 위한 주파수 대역 선택에 대한 일례

3.4 하모닉 생성

부프레임 단위로 구해진 각 하모닉 주파수에 대해 추정된 위상과 크기는 이전 부프레임에서 구해진 위상과 크기와 보간(interpolation)되며, 이는 부프레임 경계에서 스펙트럼의 변화를 부드럽게 하는 역할을 한다. 그러므로 개선된 여기신호 $r_h(n)$ 은 다음과 같이 표현된다.

$$r_h(n) = \sum_{m=M_l(T_0)}^{M_h(T_0)} A_m(n) \cos(\theta_m(n)), \quad 0 \leq n < N_s \tag{15}$$

여기서 크기 $A_m(n)$ 과 위상 $\theta_m(n)$ 의 보간은 [5]에 있는 스펙트럼 크기 및 위상의 보간 과정을 따랐다. 즉, 전 부프레임의 피치 값을 $T_0(-1)$, 전 부프레임의 m 번째 스펙트럼의 크기를 $A_m(-1)$ 이라 하면, 시간 n 에서의 m 번째 스펙트럼 크기, $A_m(n)$ 은

$$A_m(n) = A_m(-1) + \frac{n+1}{N_s} (\overline{A}_m^* - A_m(-1)), \quad M_l(T_0) \leq m \leq M_h(T_0) \tag{16}$$

으로 얻어지며, 위상, $\theta_m(n)$ 은

$$\begin{aligned} \theta_m(n) = & \frac{2\pi mn_m^*}{T_0} + \frac{2\pi mn}{T_0} \frac{n}{2N_s} \\ & + \frac{2\pi mn}{T_0(-1)} \left(1 - \frac{n}{2N_s}\right), \quad M_l(T_0) \leq m \leq M_h(T_0) \end{aligned} \tag{17}$$

으로 계산되며, 결과적으로 여기신호의 개선된 부분 $r_h(n)$ 은 식 (16)과 식 (17)을 식 (15)에 대입하므로써 얻어진다. 또한, 이득 정규화 과정은 식 (18)을 이용하여 식 (19)과 같은 수식에 의해서 얻어진다.

$$\sum_{n=0}^{N_i-1} r_p^2(n) = \sum_{n=0}^{N_i-1} (r^2(n) + g_p r_h(n))^2 = \sum_{n=0}^{N_i-1} r^2(n), \quad (18)$$

$$g_p = \frac{-2 \sum_{n=0}^{N_i-1} r(n)r_h(n)}{\sum_{n=0}^{N_i-1} r_h^2(n)}. \quad (19)$$

마지막으로 식 (7)에 의해 $r_p(n)$ 을 구하고, 이를 그림 2에서 합성필터의 입력으로 사용한다.

4. 성능 평가

본 논문에서 제안하고 있는 하모닉 모델 기반 여기신호 개선 알고리즘을 G.729 Annex D에 적용한 후에 기존의 여기신호 개선 알고리즘과 성능을 비교하였다. 성능 평가를 위해서 NTT-AT 데이터베이스 [7]에 있는 한국어 음성 중에서 남·여 각각 4 명의 화자가 발성한 약 6~8 초 정도가 되는 음성 데이터를 이용하였다. NTT-AT 데이터베이스에 있는 음성 데이터는 16 kHz로 샘플링된 신호이며, ITU-T 표준 G.191 소프트웨어 툴 [8]을 이용하여 8 kHz로 다운 샘플링하였다. 그리고 성능 측정은 PESQ를 이용하였다. PESQ를 구하기 위해 G.729 Annex D에서 사용하고 있는 여기신호 개선 알고리즘인 phase dispersion을 pitch filter와 본 논문에서 제안하고 있는 하모닉 모델 기반 여기신호 개선 알고리즘으로 대체하였다.

'Phase dispersion'은 주로 무성음의 하모닉 성분을 개선하는 특징을 가지고 있고 'Pitch Filter'와 본 논문에서 제안하고 있는 하모닉 모델 기반 여기신호 개선 알고리즘은 유성음의 하모닉을 개선하는 특징이 있기 때문에 이들을 결합하는 방식을 고려할 수 있다. 즉, 'Phase Dispersion+Pitch Filter'는 적용 코드북의 이득이 큰 경우에는 'Pitch Filter'를 적용하고 그 외의 경우에는 'Phase Dispersion'을 적용한 경우를 의미한다. 마찬가지로, phase dispersion과 하모닉 모델 기반 여기신호 개선 알고리즘도 결합할 수 있다. 이를 'Harmonic+Phase Dispersion'이라 한다. 표 1은 남·여 화자에 대해 각 방식을 적용한 후 얻은 PESQ 점수를 보여 준다. PESQ 점수는 직접적인 청취 실험 대신에 이에 상응하는 mean opinion score (MOS)를 나타내는 것으로 점수가 높을수록 좋은 음질을 의미한다. 표 1을 보면 여성화자의 경우에는 'Pitch Filter'나 'Harmonic Filter'를 사용하는 경우에 좀더 나은 성능을 보임을 알 수 있다. 여성화자의 음성신호에 대해서는 'Pitch Filter'가, 남성화자의 음성신호에 대해서는 본 논문에서 제안하고 있는 하모닉 모델 기반의 여기신호 개선 알고리즘이 좋은 성능을 보이고 있음을 알 수 있다.

표 1. 각 방식에 대한 PESQ 점수 비교

여기신호 개선 방식	화 자	
	여성	남성
Phase Dispersion	3.371	3.660
Pitch Filter	3.414	3.656
Pitch Filter + Phase Dispersion	3.409	3.651
Harmonic	3.378	3.677
Harmonic + Phase Dispersion	3.376	3.672

5. 결 론

본 논문에서는 저 전송률 음성 부호화기에서 나타나는 음질 저하를 줄이기 위해 하모닉 모델에 기반을 둔 여기신호 개선 알고리즘을 제안하였다. 제안된 알고리즘은 여성화자보다는 남성화자에서 좋은 성능을 보였으며, 기존의 G.729 Annex D에서 사용하고 있는 phase dispersion 알고리즘에 비해 성능개선을 보여주고 있다. 본 논문에서 제안된 알고리즘은 수신된 피치 주기를 이용하여 하모닉을 발생시키므로 피치 검출의 정확성에 영향을 받을 수 있다. 따라서 피치 필터처럼 복호화단에서 좀더 정확한 피치를 사용한다면 보다 더 향상된 성능을 가져올 수 있다.

참 고 문 헌

- [1] ITU-T Recommendation G.729 Annex, D. Jan., 1998. *6.4 kbit/s CS-ACELP speech coding algorithm*.
- [2] Chen, J. H., Gersho, A. Jan., 1995. "Adaptive postfiltering for quality enhancement of coded speech." *IEEE Trans. Speech Audio Process.*, vol.3, no. 1, pp. 59-71.
- [3] Hagen, R., Ekudden, E., Johansson, B., Kleijin, W. B., May, 1998. "Removal of sparse-excitation artifacts in CELP." in *Proc. ICASSP*, Seattle, WA, pp. 145-148.
- [4] Lim, J., Oppenheim, A. V., & Braida, L., Aug., 1978. "Evaluation of an adaptive comb filtering method for enhancing speech degraded by white noise addition," *IEEE Trans. Acoustic. Speech Signal Process.*, vol. 26, no. 4, pp. 354-358.
- [5] Griffin, D. W., Lim, J. S., Aug., 1988. "Multiband excitation vocoder." *IEEE Trans. Acoustic. Speech Signal Process.*, vol. 36, no. 8, pp. 1223-1235.
- [6] ITU-T Recommendation P.862, Feb., 2001. *Perceptual evaluation of speech quality (PESQ); An objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs*.
- [7] NTT-AT. 1994. *Multi-lingual speech database for telephony*.
- [8] ITU-T Recommendation G.191. Nov., 2000. *Software tools for speech and audio coding standardization*.

접수일자: 2004. 4. 29

게재결정: 2004. 6. 15

▲ 이미숙

대전시 유성구 가정동 161번지 (우: 305-350)
한국전자통신 연구원 광대역 통합망 연구단
Tel: +82-42-860-6148 (O) Fax: +82-42-860-6342
E-mail: lms@etri.re.kr

▲ 김홍국

광주광역시 북구 오룡동 1번지 (우: 500-712)
광주과학기술원 정보통신공학과
Tel: +82-62-970-2228 (O) Fax: +82-62-970-2204
E-mail: hongkook@gist.ac.kr

▲ 최승호

서울시 노원구 공릉 2동 172번지 (우: 139-743)
서울산업대학교 전자정보공학과
Tel: +82-2-970-6461 (O) Fax: +82-2-970-7903
E-mail: shchoi@snut.ac.kr

▲ 김도영

대전시 유성구 가정동 161번지 (우: 305-350)
한국전자통신 연구원 광대역 통합망 연구단
Tel: +82-42-860-5180 (O) Fax: +82-42-860-6342
E-mail: dyk@etri.re.kr