

# 정현파 크기로 가중치 된 위상 오류 함수를 사용한 음성의 중첩합산 정현파 합성 방법

준회원 박종배\*, 김규진\*, 정규혁\*, 정회원 김종학\*, 이인성\*

## Overlap and Add Sinusoidal Synthesis Method of Speech Signal using Amplitude-weighted Phase Error Function

Jong-bae Park\*, Gyu-Jin Kim\*, Jeong-gyu Hyeok\* *Associate Members*  
Jong-Hark Kim\*, In-sung Lee\* *Regular Members*

### 요 약

본 논문에서는 음성신호의 정현파 합성방법 중 하나인 선형위상을 사용한 중첩합산방법에 대하여 합성음성의 연속성을 개선시킨 새로운 방법을 제안한다. 기존의 중첩합산 정현파 합성방법은 프레임의 중간 지점에 대한 위상 값을 예측하기 위해 합성 파형의 경계면에 대한 위상 차이를 최소화하는 방법을 사용하였으나, 제안된 방법은 중간지점의 위상만이 아닌 최종 합성된 신호들 간의 차이를 최소화하는 방법을 사용한다. 이러한 합성신호의 오차 함수를 최소화하는 과정은 정현파 크기 값을 가중치로 사용한 위상오차 함수로 근사화 될 수 있다. 제안된 합성 방법의 연속성에 대한 성능을 검증하기 위해 합성신호의 전 구간에 대한 합성 신호들 간의 평균 오차 값 및 청취자 테스트를 하였다. 기존의 위상 오차함수와 비교하여 개선된 연속성 성능을 얻을 수 있었다.

**Key Words** : Sinusoidal model, Phase Interpellation, Matching Pursuit, Overlap and Add

### ABSTRACT

In this paper, we propose a new overlap and add speech synthesis method which demonstrates improved continuity performance. The proposed method uses a weighted phase error function and minimizes the wave discontinuity of the synthesis signal, rather than the phase discontinuity, to estimate the mid-point phase. Experimental results show that the proposed method improves the continuity between the synthesized signals relative to the existing method.

### 1. 서 론

일반적인 정현파 또는 하모닉 모델은 시간에 따라 변화하는 주파수, 진폭, 그리고 위상을 가진 정현파 성분의 선형 합으로 정의 한다<sup>[1]</sup>.

정현파 모델은 낮은 비트 전송률로 음성신호를

부호화하는 효율적인 기술로 알려져 있고, 최근에는 음성 변환이나 음질 개선, 그리고 저 전송률의 음성 부호화의 디지털 신호처리 분야에 이용되고 있다<sup>[2][3][4]</sup>. 이러한 음성 부호화기의 처리 단계는 파라미터의 보간법이나 파형의 중첩합산(Overlap and Add) 방법을 통해 수행된다. 파라미터의 예측은 일

※ "이 논문은 2007년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임" (지방연구중심대학육성사업/충북BIT연구중심대학육성사업단

\* 충북대학교 전자공학과(Department of Radio Engineering, Chungbuk National University) (inslee@chungbuk.ac.kr)

논문번호 : KICS2007-06-248, 접수일자 : 2007년 6월 8일, 최종논문접수일자 : 2007년 12월 5일

반적으로 MP(Matching Pursuit)방법이 정확한 것으로 알려져 있다<sup>5)</sup>.

정현파 모델은 한 프레임 안에서 정현파 파라미터가 일정한 값을 가진다는 기본 가정 때문에 프레임간의 불연속성이 생기게 된다. 일반적으로 이러한 프레임 불연속을 줄이기 위해 스펙트럼 크기의 선형 보간(linear interpolation)과 위상 보간(phase interpolation)이 합성 시 함께 사용된다<sup>1)</sup>.

본 논문에서는 합성신호를 보간 할 때 위상오차에 대해서 최소화하도록 프레임 중간지점의 정현파 파라미터를 추정하지 않고 합성신호들 간의 오차를 최소화하는 방법으로 정현파 파라미터를 추정함으로써 합성과형에 대한 연속성을 개선한다.

본 논문의 II장에서는 기존의 선형위상을 사용한 중첩합산 정현파 방법<sup>6)</sup>에 대해 설명하고, III장에서는 정현파 합성 함수들 간의 오차를 최소화 하는 과정이 정현파 크기(amplitude)로 가중치 된 위상 오차 식으로 유도되는 것을 설명하고, 이 정현파 크기의 가중치 값이 첨가된 새로운 선형위상을 사용한 중첩합산 정현파 합성 방법을 설명한다. IV장에서는 기존의 합성 방법과 제안된 합성 방법 간의 성능을 평가하고 V장에서 결론을 맺는다.

## II. 선형 위상 중첩 합산 정현파 합성법

프레임간의 연속성을 유지하기 위해 정현파 파라미터들의 연결은 필수적이다. McAulay와 Quatieri는 3차 위상 보간<sup>6)</sup>에 비해 적은 복잡도를 가지면서 프레임간의 연속성을 유지하기 위한 방법으로 선형위상을 가지는 중첩합산(OLA-LP: overlap and add-linear phase) 정현파 합성법을 제안하였다<sup>6)</sup>. 그 합성 방법의 주된 요점은 중첩길이 N/2 (overlap-length N/2)에 대하여 그림 1에서의 같이 k

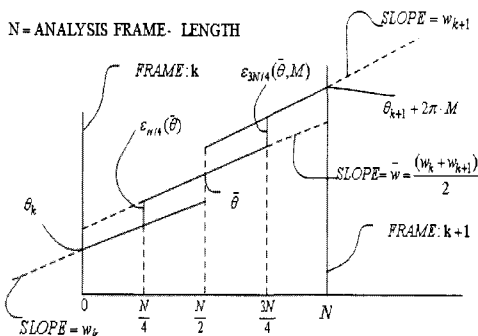


그림 1. N/2 지점의 프레임이 첨가된 위상 계획

번째 프레임과 k+1번째 프레임 사이의 N/2 지점을 중심으로 파라미터들의 연속성을 유지하도록 N/2 지점에서의 정현파 파라미터를 찾는 것이다. 즉, k번째와 k+1번째 프레임의 중간에 위치한 정현파의 크기, 주파수, 위상을 찾는 것이다. k번째 프레임에서의 크기, 주파수, 위상은 각각  $A_k, w_k, \theta_k$ 라 정의하고 N/2 지점인 중간에 위치한 크기, 주파수, 위상을  $\bar{A}, \bar{w}, \bar{\theta}$ 라 한다면 중간에 위치한 크기 값은 간단히 다음과 같이 과거와 현재 크기 파라미터의 평균값으로 표현 할 수 있다<sup>6)</sup>.

$$\bar{A} = \frac{A_k + A_{k+1}}{2} \quad (1)$$

유사하게, 중간지점의 주파수 또한 과거와 현재 주파수 파라미터의 평균값으로 다음과 같이 표현할 수 있다<sup>6)</sup>.

$$\bar{w} = \frac{w_k + w_{k+1}}{2} \quad (2)$$

하지만 중간지점의 위상 파라미터의 선택은 분명하지 않다<sup>6)</sup>. 주파수는 중첩합산 윈도우의 범위 내에서 일정(linear phase)하기 때문에 정현파 파라미터의 위상은 선형적으로 증가한다. 예로, k번째 프레임에서 k+1번째 프레임에 대해 위상을 예측하면 k+1번째 프레임의 위상은  $2\pi \cdot M$ 이 더해진다<sup>6)</sup>. 여기서 M은 위상 비 구속 정수 인자(phase unwrapping integer factor)이다<sup>6)</sup>. 중첩길이 N/2(overlap-length N/2)에 대한 중첩합산(OLA: overlap and add) 과정에서 각 프레임마다 삼각 윈도우(triangle window)가 적용되기 때문에 정현파의 최대 중첩(overlap)은 윈도우가 서로 만나는 지점에서 발생한다. 그러므로 최대 중첩 지점에서의 위상 불연속성을 최소화하는 방법으로 중간지점의 위상을 얻는 것이 적절하다. 이 두 점은 N/4, 3N/4이며 각각의 위상 에러를  $\epsilon_{N/4}(\bar{\theta})$ 와  $\epsilon_{3N/4}(\bar{\theta}, M)$ 이라 한다면 다음과 같이 정의 될 수 있다<sup>6)</sup>.

$$\begin{aligned} \epsilon_{N/4}(\bar{\theta}) &= (\theta_k + w_k \cdot \frac{N}{4}) + \bar{w} \cdot \frac{N}{4} - \bar{\theta} \\ \epsilon_{3N/4}(\bar{\theta}, M) &= (\theta_{k+1} + 2\pi M - w_{k+1} \cdot \frac{N}{4}) - \bar{w} \cdot \frac{N}{4} - \bar{\theta} \end{aligned} \quad (3)$$

여기서 위상의 불연속성을 줄이기 위해 첨가된 위상  $\bar{\theta}$ 와 정수 M을 추정하는 것이 중요하다. 두

지점에서 에러의 제곱의 합을 위상오차함수  $l(\bar{\theta}, M)$  로 정의 하면 다음과 같다<sup>6)</sup>.

$$l(\bar{\theta}, M) = \epsilon_{N/4}^2(\bar{\theta}) + \epsilon_{3N/4}^2(\bar{\theta}, M) \quad (4)$$

위상 불연속을 최소로 하는 위상  $\bar{\theta}$ 와 정수  $M$ 은 식(4)에서  $\bar{\theta}$ 와  $M$  각각에 대하여 편미분 값이 0 이 되는 값으로 다음과 같이 구해진다<sup>6)</sup>.

$$\bar{\theta} = \frac{\theta_k + \theta_{k+1}}{2} + \pi M + \frac{w_k - w_{k+1}}{2} \cdot \frac{N}{4} \quad (5)$$

$$M = \text{round}\left(\frac{\theta_k - \theta_{k+1} + \bar{w} \cdot N}{2\pi}\right) \quad M = 0, \pm 1, \pm 2, \dots \quad (6)$$

### III. 정현파 크기로 가중치 된 선형위상 중첩합산 정현파 합성법

기존의 선형위상을 가지는 중첩합산(OLA-LP: overlap and add-linear phase) 정현파 합성법은 최대 중첩 윈도우가 만나는 지점에 대해서 단순히 위상 오차에 대해서만 오차를 최소화 하는 방법으로 정현파 파라미터의 위상 값을 예측하였다. 하지만 기존의 OLA-LP 정현파 합성법은 합성 파형에 대해서 불연속성을 최소화 하는 것이 아니기 때문에 정현파 파라미터의 연속성을 향상시키기에는 충분하지 않다. 그러므로 합성파형에 대한 오차를 최소화 하는 과정에서 정현파 파라미터를 찾는 것이 필요하다.

중간 지점의 정현파 파라미터(크기  $\bar{A}$ , 주파수  $\bar{w}$ , 위상  $\bar{\theta}$ )에 대해서 정현파 파라미터의 크기는 식 (1) 과 똑같이 정의하고 주파수는 식 (2)와 다르게 크기 값에 비례하여 다음과 같이 정의한다.

$$\bar{w} = \frac{A_k \cdot w_k + A_{k+1} \cdot w_{k+1}}{A_k + A_{k+1}} \quad (7)$$

중간 지점의 정현파 파라미터의 위상  $\bar{\theta}$  을 얻기 위해서 우리는 최대 중첩지점  $N/4, 3N/4$  에 대해서 합성파형에 대한 MSE(mean-squared error) 값을  $\epsilon_{N/4}^s(\bar{\theta})$ 와  $\epsilon_{3N/4}^s(\bar{\theta}, M)$ 로 각각 정의 하면 다음과 같다.

$$\begin{aligned} \epsilon_{N/4}^s{}^2(\bar{\theta}) &= |W_{1/4} \cdot e^{j\phi_k(N/4)} - W_{1/4} \cdot e^{j\bar{\phi}(-N/4)}|^2 \\ \epsilon_{3N/4}^s{}^2(\bar{\theta}, M) &= |W_{3/4} \cdot e^{j\phi_{k+1}(-N/4)} - W_{3/4} \cdot e^{j\bar{\phi}(N/4)}|^2 \end{aligned}$$

$$\begin{aligned} \text{where } W_{1/4} &= \frac{3A_k + A_{k+1}}{4}, \quad W_{3/4} = \frac{A_k + 3A_{k+1}}{4} \quad (8) \\ \phi_k(n) &= w_k \cdot n + \theta_k, \quad \bar{\phi}(n) = \bar{w} \cdot n + \bar{\theta} \end{aligned}$$

여기서  $W_{1/4}$ 와  $W_{3/4}$ 는  $N/4$  지점과  $3N/4$  지점에서의  $k$ 번째 프레임과  $k+1$ 번째 프레임으로부터 각 지점의 평균 크기 값이다.

식 (8)은 MSE(mean-squared error)를 최소로 하는 중간 지점의 위상을 얻기 위해 편미분시 0이 되는 해를 구하게 된다. 그러나 위상 항에 대한 삼각 함수를 포함 하고 있고 많은 국소해(local solutions)를 가지고 있어 최적 해를 구하기 힘들다. 그러므로 이 문제를 해결하기 위해서 식 (8)은 중간 지점의 위상  $\bar{\theta}$ 에 대해서 삼각 함수의 공식을 사용하여 보다 간단한 식으로 유도하면 다음과 같다.

$$\begin{aligned} \epsilon_{N/4}^s{}^2(\bar{\theta}) &= (W_{1/4})^2 |e^{j\phi_k(N/4)} - e^{j\bar{\phi}(-N/4)}|^2 \\ &= (W_{1/4})^2 \left( \sqrt{2 - 2 \left( \cos \phi_k \left( \frac{N}{4} \right) \cdot \cos \bar{\phi} \left( -\frac{N}{4} \right) + \sin \phi_k \left( \frac{N}{4} \right) \cdot \sin \bar{\phi} \left( -\frac{N}{4} \right) \right)} \right)^2 \\ &= (W_{1/4})^2 \left( \sqrt{2 \left( 1 - \cos \left( \phi_k \left( \frac{N}{4} \right) - \bar{\phi} \left( -\frac{N}{4} \right) \right) \right)} \right)^2 \\ &= 2 \cdot (W_{1/4})^2 \left( 1 - \cos \left( \phi_k \left( \frac{N}{4} \right) - \bar{\phi} \left( -\frac{N}{4} \right) \right) \right) \\ \epsilon_{3N/4}^s{}^2(\bar{\theta}, M) &= (W_{3/4})^2 |e^{j\phi_{k+1}(-N/4)} - e^{j\bar{\phi}(N/4)}|^2 \quad (9) \\ &= 2 \cdot (W_{3/4})^2 \left( 1 - \cos \left( \phi_{k+1} \left( -\frac{N}{4} \right) - \bar{\phi} \left( \frac{N}{4} \right) \right) \right) \end{aligned}$$

위의 식(9)의 유사한 결과가 <sup>7)</sup>에서도 관찰된다. 만약 변수  $x$ 가 작은 값이라면, 1차 테일러 급수(1st Taylor Series)에 의해서  $\cos \cong 1 - \frac{x^2}{2}$  이므로 식(9)는 다음과 같이 근사화 된다.

$$\begin{aligned} \epsilon_{N/4}^s{}^2 &= (W_{1/4})^2 \left( \phi_k \left( \frac{N}{4} \right) - \bar{\phi} \left( -\frac{N}{4} \right) \right)^2 \quad (10) \\ &= (W_{1/4})^2 \left( \theta_k + w_k \cdot \frac{N}{4} + \bar{w} \cdot \frac{N}{4} - \bar{\theta} \right)^2 \end{aligned}$$

$$\begin{aligned} \epsilon_{3N/4}^s{}^2 &= (W_{3/4})^2 \left( \phi_{k+1} \left( -\frac{N}{4} \right) - \bar{\phi} \left( \frac{N}{4} \right) \right)^2 \\ &= (W_{3/4})^2 \left( \theta_{k+1} + 2\pi M - w_{k+1} \cdot \frac{N}{4} - \bar{w} \cdot \frac{N}{4} - \bar{\theta} \right)^2 \end{aligned}$$

실제로 우리는 유성음에 대한 많은 실험(8kHz NTT 데이터 남녀 64문장샘플)으로부터  $\phi_k \left( \frac{N}{4} \right) - \bar{\phi} \left( -\frac{N}{4} \right)$ ,  $\phi_{k+1} \left( -\frac{N}{4} \right) - \bar{\phi} \left( \frac{N}{4} \right)$  값의 차이가 대부분 약 0.32 라디안(radian)보다 작은 값을 가짐

을 확인할 수 있었으며, 이것은 식 (10)이 충분히 근사값으로 적용 가능함을 의미한다. 위의 식(10)으로부터 최종 합성된 신호의 MSE가 기존의 위상 오차함수에서 정현파 크기(amplitude)가 가중치 된 위상오차 함수식으로 간략하게 유도됨을 알 수 있다. 우리는 이것을 정현파 크기에 가중치 된 선형 위상 중첩 합산(WE-OLA-LP: weighted-overlap and add-linear phase) 합성법으로 정의한다.

식 (4)와 마찬가지로 식 (10)에서 유도된 가중치 된 위상오차 함수의 개념을 사용하면  $N/4$ 과  $3N/4$  지점에 대한 에러 제곱의 합을  $r(\bar{\theta}, M)$ 로 다음과 같이 정의 할 수 있다.

$$r(\bar{\theta}, M) = \varepsilon_{N/4}^2(\bar{\theta}) + \varepsilon_{3N/4}^2(\bar{\theta}, M) \quad (11)$$

위의 식 (11)에서 합성된 신호를 최소로 하는 위상  $\bar{\theta}$ 와 위상 비 구속 정수 인자  $M$ 을 구할 수 있는데, 이것은 각각에 대해서 편미분 값이 0 이 되는 값을 구하면 된다.

$$\bar{\theta} = \frac{W_{1/4}X + W_{3/4}Y}{W_{1/4} + W_{3/4}} \quad (12)$$

$$M = \partial \left[ \frac{\frac{W_{1/4}X + W_{3/4}Z}{W_{1/4} + W_{3/4}} - \theta_{k+1} + w_{k+1} \cdot \frac{N}{4} + \bar{w} \cdot \frac{N}{4}}{\left( 2\pi - \frac{W_{3/4} \cdot 2\pi}{W_{1/4} + W_{3/4}} \right)} \right]$$

where  $X = \theta_k + \frac{3w_k + w_{k+1}}{2} \cdot \frac{N}{4}$   
 $Y = \theta_{k+1} + 2\pi M - \frac{w_k + 3w_{k+1}}{2} \cdot \frac{N}{4}$   
 $Z = \theta_{k+1} - \frac{w_k + 3w_{k+1}}{2} \cdot \frac{N}{4}$  (13)

#### IV. 시뮬레이션 및 성능 평가

기존의 OLA-LP 정현파 합성법이 단순히 위상에 대한 불연속성을 최소화하기 위해 프레임의 중간지점의 정현파 파라미터를 찾았다면 제안된 WE-OLA-LP 정현파 합성법은 위상 차이만이 아닌 합성된 신호들 간의 차이를 최소화 하면서 프레임의 중간지점의 정현파 파라미터를 찾는다. 이러한 정현파 모델은 프레임간의 연속성을 유지하기 위해 정현파 파라미터의 연결은 필수적이다. 기존의

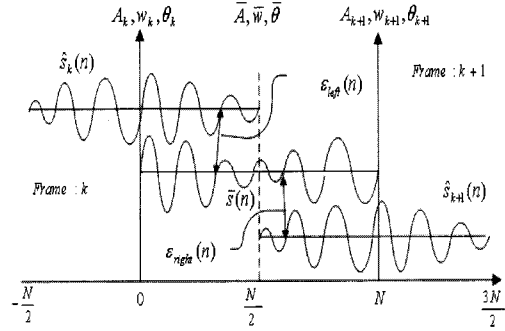


그림 2. k와 k+1번째 프레임의 하모닉 성분과 N/2 지점에서의 하모닉 성분으로 합성된 파형

모양OLA-LP 정현파 합성법과 제안된 WE-OLA-LP 정현파 합성법에 대한 연속성 성능을 평가하기 위해서 그림 2에서와 같이 k번째 프레임의 정현파 파라미터  $A_k, w_k, \theta_k$ 으로부터 합성된 신호를  $\hat{s}_k(n)$ , k+1번째 프레임의 정현파 파라미터  $A_{k+1}, w_{k+1}, \theta_{k+1}$ 으로부터 합성된 신호는  $\hat{s}_{k+1}(n)$ 으로 정의한다. 여기서, 정현파 파라미터는 MP를 통해서 측정된다<sup>5)</sup>.

그리고 정현파 파라미터  $\bar{A}, \bar{w}, \bar{\theta}$ 으로부터 합성된 신호를  $\bar{s}(n)$ 이라 하면, 그 정현파 파라미터는 2장에서 설명된 기존의 OLA-LP 정현파 합성법과 3장에서 설명된 새롭게 제안된 WE-OLA-LP 정현파 합성법으로부터 각각 구해진다.

$N/4$  지점과  $3N/4$  지점뿐만이 아니라, 각 샘플에 대한 연속성을 관찰하기 위하여  $\hat{s}_k(n)$ 와  $\bar{s}(n)$ 의 차이 값의 전체 프레임에 대한 평균값을 식 (14)으로 정의의 하고  $\bar{s}(n)$ 와  $\hat{s}_{k+1}(n)$ 의 차이 값에 대한 전체 프레임에 대한 평균값을 식 (15)으로 정의하여 성능을 관찰한다.

$$\varepsilon_{\text{left}}(n) = \frac{1}{P} \sqrt{\sum_{n=0}^{N/2-1} \sum_{k=1}^P (\hat{s}_k(n) - \bar{s}(n))^2} \quad (14)$$

$$\varepsilon_{\text{right}}(n) = \frac{1}{P} \sqrt{\sum_{n=N/2k=1}^{N-1} \sum_{k=1}^P (\bar{s}(n) - \hat{s}_{k+1}(n))^2} \quad (15)$$

여기서 P는 전체 프레임 개수이다.

그림 3은 OLA-LP 정현파 합성법과 제안된 WE-OLA-LP 정현파 합성법에 대한 식 (14), (15)의 결과 값을 나타낸 것이다. 여기서 실험 샘플은 8kHz NTT 음성 데이터 한국어 남자, 여자 각각 32

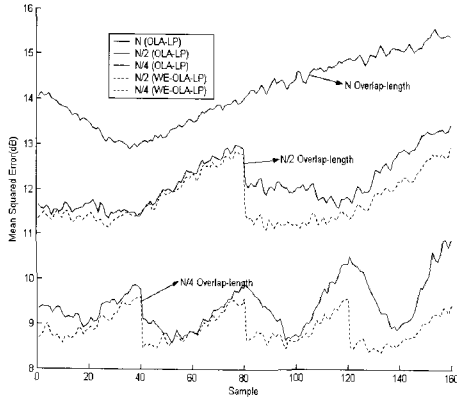


그림 3. 합성 프레임틀에 따른 각 샘플마다 평균MSE값 (dB) (실선: OLA-LP, 점선: WE-OLA-LP)

문장이 사용 되었고 전이구간을 포함한 유성음 구간에 대해서 측정되었다. 전체적으로 기존의 OLA-LP 정현파 합성법과 제안한 WE-OLA-LP 정현파 합성법 모두 중첩길이(overlap-length)를 N/2, N/4으로 짧게 하였을 경우 MSE(mean-squared error) 값이 줄어드는 것을 볼 수 있고, 제안한 WE-OLA-LP 정현파 합성법이 기존의 OLA-LP 정현파 합성법보다 MSE 값이 작음을 볼 수 있다.

프레임 전 구간에 대한 연속성에 대한 평균 MSE(mean-squared error) 값을 측정하기 위해 표 1에서는 다음과 같은 식의 값을 나타내었다.

$$E = \frac{1}{2} \sum_{n=0}^{N/2-1} \epsilon_{\text{left}}(n) + \frac{1}{2} \sum_{n=0}^{N/2-1} \epsilon_{\text{right}}(n) \quad (16)$$

표1에 나타난 것처럼 제안된 WE-OLA-LP 방법이 기존의 OLA-LP 에 비해 MSE 값의 평균값이 중첩길이 N/2 (overlap-length N/2)에서는 0.45dB, 중첩길이 N/4 (overlap-length N/4)에서는 0.58dB 작아짐을 볼 수 있었고, 이것은 정현파의 크기 값을 고려하여 합성할 경우 좀 더 연속성이 개선된 부드러운 합성을 할 수 있다는 것을 나타낸다.

표 1. 각 샘플과 합성 프레임틀에 따른 MSE의 평균값(dB)

Overlap-length	OLA-LP(dB)	WE-OLA-LP(dB)
N	14.15	14.15
N/2	12.14	11.69
N/4	9.41	8.83

그림 4는 실제로 제안된 방법의 연속성이 더욱 부드러운 합성을 한다는 예를 보이기 위해 기존의 방법과 제안된 방법으로 합성한 파형의 스펙트로그램을 비교해 보았다. 원음과 비교해 봤을 때 그림 4의 표시한 부분에서 기존의 합성방법은 정현파 파라미터의 연결이 제안한 합성방법보다 끊어져 있음을 볼 수 있다. 즉 제안한 방법이 정현파 모델의 프레임간의 연결성을 유지하는데 좋음을 볼 수 있다.

표 2는 기존의 OLA-LP 합성법과 제안한 WE-OLA-LP 합성법에서 보간 하려는 지점에서 정현파 파라미터를 구하는 방법에 따른 연산량을 비교하여 측정하였다. 즉, 식 (1),(2),(5),(6)과 식 (1),(7),(12),(13)의 연산량을 DSP에 사용되는 기본 연산인 덧셈, 곱셈의 연산 복잡도 가중치를 1로 하여 구한 연산량이다. 여기서 L은 하모닉 개수, A는 덧셈, M은 곱셈이다. 여기서 제안된 방법의 연산량이 기존의 방법보다 약간 많은 정도이나 실제로 매트랩에서 시간 함수를 사용하여 연산량 측정결과 제안한 방법이 시간이 평균 1.25배 길며 그리 큰 연산량의 차이는 아니었다.

제안된 모델의 성능을 확인하기 위해 주관적 음질테스트를 하였다. 실험은 NTT 음성 여자 24문장

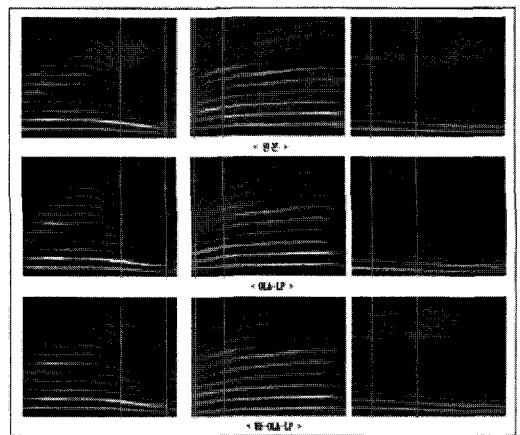


그림 4. 합성 방법(N/2 overlap-length)에 따른 스펙트로그램 비교

표 2. 합성 방법에 따른 추정 연산량

Synthesis Method	Operation
Conventional (OLA-LP)	L(8A+10M)
Proposed (WE-OLA-LP)	L(18A+24M)

표 3. 합성 방법에 따른 선호도

	Conventional (OLA-LP)	Proposed (WE-OLA-LP)
Man	44.63%	55.37%
Woman	43.98%	56.02%
Total	44.30%	55.70%

과 남자 24문장을 사용하였고, 테스트에 참여한 인원은 15명이다. 테스트 방법은 잡음이 없는 동일한 음성을 각각의 합성법으로 합성하여 약 2초간의 간격을 두고 들려준 후 더 듣기 좋은 음성을 고르는 A/B 테스트 방법을 사용하였다. 표 3의 음질 테스트 결과에서 보여주듯이 제안된 합성법은 기존의 합성법보다 나은 선호도를 보였다.

### V. 결 론

본 논문에서는 음성신호의 정현과 합성방법 중 하나인 선형위상을 사용한 중첩합산방법에 대한 합성음성의 연속성을 개선시킨 새로운 방법을 제안하였다. 이렇게 제안한 합성신호들 간의 차이를 최소화 하는 과정에서 기존에 OLA-LP 정현과 합성법에서 단순히 위상 에러(phase errors)를 최소화하기 위한 나온 식 앞에 가중치 값이 첨가 되는 식을 유도하였고 기존의 방법과 제안한 방법의 연산량 및 성능평가를 보았을 때 연산량에는 많은 시간 차이가 나지 않는 반면 제안된 WE-OLA-LP 정현과 합성법으로 합성 하였을 때 기존의 OLA-LP 정현과 합성법보다 하모닉 성분들의 연속성이 좋아짐을 확인할 수 있었다.

### 참 고 문 헌

[1] R. J. McAulay and T. F. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," *IEEE Trans. on ASSP*, vol. 34, no. 4, pp. 744-754, Aug. 1986.

[2] W. B. Kleijin and K. K. Paliwal, "Speech coding and synthesis," *Elevier Science Publishers*, Amsterdam, 1995

[3] J. Nieuwenhuijse, R. Heusdens, and E.F. Deprettere, "Robust exponential modeling of audio signals," *IEEE International Conference*

on Acoustics, Speech, and Signal Processing, *ICASSP '98*, Seattle, Washington, USA, vol. 6, pp. 3581-3584, May 1998.

[4] T. S. Verma and T. H. Y. Meng, "Sinusoidal modeling using frame-based perceptually weighted matching pursuits," *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '99*, Phoenix, Arizona, USA, vol. 2, pp. 981984, May 1999.

[5] M. Goodwin, "Matching Pursuit with Damped Sinusoids", *Proc. IEEE ICASSP 1997*, vol.3, pp.2037-2040.

[6] R. J. McAulay and T. F. Quatieri, "Computationally efficient sine-wave synthesis and its application to sinusoidal Transform coding," *Proc. IEEE ICASSP 1998*, pp.370-373.

[7] David L. Thomson, "Parametric Models of the Magnitude/Phase Spectrum for Harmonic Speech Coding," *ICASSP 1988*, pp.378-381, 198.

박 종 배 (Jong-bae Park)

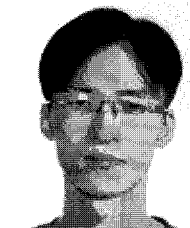
준회원



2007년 2월 충북대학교 정보통신공학과(공학사)  
2007년 3월~현재 충북대학교 전파공학과(석사)  
<관심분야> 음성/오디오 부호화, 통신 신호 처리, 적응 필터

김 종 학 (Jong-hark Kim)

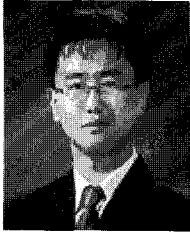
정회원



1999년 2월 충북대학교 전자공학과(공학사)  
2000년 2월 충북대학교 전파공학과(공학석사)  
2000년 3월~현재 충북대학교 전파공학과(공학박사)  
<관심분야> 음성/오디오 부호화, 영상 압축, 적응 필터

김 규 진 (Gyu-jin Kim)

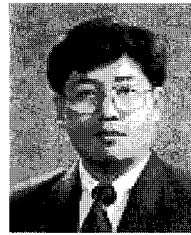
준회원



2006년 2월 충북대학교 전자공학과(공학사)  
2007년 3월~현재 충북대학교 전자공학과(공학석사)  
<관심분야> 음성/오디오 부호화, 통신 신호 처리, 적응 필터

이 인 성 (In-sung Lee)

정회원



1983년 2월 연세대학교 전자공학과(공학사)  
1985년 2월 연세대학교 전자공학과(공학석사)  
1992년 2월 Texas A&M University 전기공학과(공학박사)  
1993년 2월~1995년 9월 한국전자통신연구원, 이동통신 기술연구단 선임연구원  
1995년 10월~현재 충북대학교 전기전자공학부 정교수  
<관심분야> 음성/영상 신호 압축, 이동통신, 적응필터.

정 규 혁 (Jeong-gyu Hyeok)

준회원



2004년 2월 충북대학교 전기전자공학(공학사)  
2006년 2월 충북대학교 전자공학과(공학석사)  
2006년 3월~현재 충북대학교 전자공학과(공학박사)  
<관심분야> 음성/오디오 부호화,

통신신호처리, VoIP