

동적 세그멘테이션을 이용한 폴리포닉 오디오 신호의 정현파 모델링

Sinusoidal Modeling of Polyphonic Audio Signals Using Dynamic Segmentation Method

장 호 근*, 박 주 성**

(Ho Keun Jang*, Ju Sung Park**)

요 약

본 논문에서는 폴리포닉 오디오 신호에 대한 정현파 모델링 방법을 제안한다. 정현파 모델링을 폴리포닉 오디오 신호에 적용하는데 있어서 가장 큰 문제점은 스펙트럼 분석을 위한 분석 윈도우의 크기를 결정할 수 없다는 것이다. 또한 고품질의 합성음을 위해서는 악기음의 특성을 결정짓는 어택이 잘 보존되어야 한다. 본 논문에서는 입력 신호를 6개의 옥타브 밴드 구조의 다중 해상도 필터 뱅크를 통과시키고, 각 서브밴드 신호에 대해 서로 다른 크기의 분석 윈도우를 적용시킴으로써 폴리포닉 오디오 신호에 대한 분석 윈도우 크기 결정 문제를 해결한다. 정현파 모델링에서 발생하는 어택과 같은 천이 구간에 서의 퍼짐 현상을 개선하기 위해 각 서브밴드 신호에 동적 세그멘테이션 방법을 적용하여 천이 구간 근처에서는 분석과 합성 프레임 크기를 작게 하는 방법을 사용한다. 이 방법을 통해 서브밴드 신호의 구간별 시간-주파수 특성에 따라 적절한 크기의 윈도우를 선택할 수 있다. 동적 세그멘테이션 방법으로는 기존의 방법보다 계산량과 성능 면에서 더 나은 특성을 보이는 방법을 제안한다. 여러가지 폴리포닉 오디오 신호에 대한 시뮬레이션 결과 제안한 정현파 모델링 방법이 음질의 손상 없이 원래 신호를 잘 복원할 수 있음을 확인하였다.

핵심용어: 오디오 신호의 정현파 모델링, 다중해상도 필터 뱅크, 동적 세그멘테이션

투고분야: 전기음향 분야(3.5)

ABSTRACT

This paper proposes a sinusoidal modeling of polyphonic audio signals. Sinusoidal modeling which has been applied well to speech and monophonic signals cannot be applied directly to polyphonic signals because a window size for sinusoidal analysis cannot be determined over the entire signal. In addition, for high quality synthesized signal transient parts like attacks should be preserved which determines timbre of musical instrument. In this paper, a multiresolution filter bank is designed which splits the input signal into six octave-spaced subbands without aliasing and sinusoidal modeling is applied to each subband signal. To alleviate smearing of transients in sinusoidal modeling a dynamic segmentation method is applied to subbands which determines the analysis-synthesis frame size adaptively to fit time-frequency characteristics of the subband signal. The improved dynamic segmentation is proposed which shows better performance about transients and reduced computation. For various polyphonic audio signals the result of simulation shows the suggested sinusoidal modeling can model polyphonic audio signals without loss of perceptual quality.

Key words: Sinusoidal modeling, Multiresolution filter bank, Dynamic segmentation.

I. 서 론

1980년대 중반에 소개된 정현파 모델은 신호의 분석을 통한 합성 모델로, 오디오 신호를 시간에 따라 변하는 주파수와 진폭을 갖는 정현파의 합으로 나타내는 것이다. 이 모델은 원래 신호의 음질을 그대로 복원할 수 있을 뿐만 아니라, 신호의 음정이나 음색을 변화시키는 주파수 변화와 음정이나 음색을 그대로 유지하면서 재생

*현대전자 시스템 IC 선임연구원

**부산대학교 전자공학과 부교수

접수일자 : 2000년 1월 5일

속도를 변화시키는 시간축 변화에 특히 뛰어난 모델로 알려져 있으며, 음의 분석/변형/합성 모델로서 가장 널리 사용되고 활발히 연구되고 있는 분야의 하나이다[1-4].

지금까지 정현파 모델링에 대해 연구되어 온 것은 주로 음성 신호나 단일 악기음에 대한 것이었으므로 이것을 여러가지 악기음이 복잡하게 섞여 있는 폴리포닉 음에는 적용하기 어렵다. 폴리포닉 음을 정현파 모델링 할 때 가장 문제가 되는 것은 정현파 분석 과정에서 필요한 윈도우의 크기, 윈도우의 모양, FFT 크기, 그리고 윈도우 적용 간격 등의 분석 파라미터를 일률적으로 정할 수 없다는 것이다. 음성 신호나 단일 악기음의 경우 이것은 신호의 피치 주파수로부터 대부분 결정된다[1-4]. 그러나 폴리포닉 음은 여러 개의 복잡한 음이 섞여 있으므로 신호의 피치 주파수를 찾아내기 어렵고, 설정 찾아낼 수 있다 하더라도 그것은 전체 음을 구성하는 하나의 음에만 적용 가능한 것이다.

정현파 모델링의 또다른 문제점은 분석/합성 과정에서 일정한 크기의 파형 구간에서 분석/합성이 이루어짐에 따라 짧은 시간 내에 급격하게 신호의 특성이 변화하는 어택(attack)과 같은 천이 구간(transients)을 잘 표현할 수 없다는 것이다. 분석시에 얻어지는 정현파의 파라미터는 분석 구간내의 주파수 성분들의 평균화된 파라미터 값이므로, 실제 천이 구간에 포함된 정확한 파라미터 변화를 얻기 어렵다. 더구나 합성 과정에서 파라미터에 대한 인터폴레이션이 일어나므로, 천이 구간의 파형은 시간상으로 퍼져 버리는 효과를 가져온다. 심리 음향의 실험 결과에 따르면 음성의 명료도나 악기음의 음색을 결정하는데 있어 음의 초기 천이 구간이 매우 결정적인 역할을 하는 것으로 알려져 있다[5]. 따라서 오디오 신호를 모델링할 때는 신호의 천이 구간을 잘 표현할 수 있도록 모델링하는 것이 필요하다.

본 논문은 폴리포닉 음과 같은 복잡한 스펙트럼을 갖는 오디오 신호의 정현파 모델링을 위한 분석 방법을 제안한다. 앞에서 언급한 것처럼 폴리포닉 음에 일정한 크기의 분석 윈도우를 적용할 수 없는 문제를 해결하기 위해 본 논문에서는 입력 신호를 옥타브 밴드 구조의 다중 해상도(multiresolution) 필터 बैं크를 통과시킨 서브밴드 신호에 대해 정현파 분석을 행한다. 이렇게 하면 서브밴드 신호는 대역폭이 제한되어 있으므로, 서브밴드 신호별로 시간-주파수 해상도에 따라 정현파 모델링에 필요한 적절한 분석/합성 구간 크기를 결정할 수 있다. 한편, 정현파 모델로 합성된 신호의 음질에 가장 큰 영향을 미치는 천이 구간의 특성을 보존하기 위해, 각 서브밴드 신호에서 시간에 따라 변화하는 신호의 특성에 맞게 정현파 분석/합성 구간의 크기를 다르게 해준다. 이렇게 하면 단일 피치를 갖지 않는 폴리포닉 신호의 국지적인 특성에 맞춰 가장 최적의 분석/합성 구간 크기를 결정할 수 있을 뿐만 아니라, 천이 구간까지도 정현파 성분으로 효과적으로 모델링할 수 있다.

본 논문은 다음과 같은 내용으로 구성되어 있다. 2장에서는 오디오 신호에 대한 기존의 정현파 모델링 방법에 대해서 간략하게 설명한다. 3장에서는 복잡한 오디오 신호를 분석하기 위한 다중 해상도 필터 बैं크의 설계에 대해서 설명한다. 4장에서는 서브밴드 신호에 대한 동적 세그멘테이션 방법을 설명하고, 계산량과 성능면에서 기존의 방법보다 나은 특성을 보이는 방법을 제안한다. 5장에서는 본 논문에서 제안된 알고리즘을 이용하여 여러가지 폴리포닉 오디오 신호를 정현파 모델로 합성한 시뮬레이션 결과를 보여준다. 마지막으로 6장에서는 결론을 제시한다.

II. 정현파 모델링

정현파 모델은 신호를 시간에 따라 변하는 주파수와 진폭을 갖는 정현파의 합으로 나타낸다. 이것은 식 (1)과 같이 표현될 수 있다.

$$x(t) \approx \hat{x}(t) = \sum_{i=1}^{Q(t)} A_i(t) \cos[\theta_i(t)] \tag{1}$$

여기서 $A_i(t)$ 는 시간에 따라 변하는 정현파의 진폭을 나타내고, $\theta_i(t)$ 는 정현파의 순시 위상(instantaneous phase), 그리고 $Q(t)$ 는 정현파의 개수를 나타낸다. 순시 위상은 다음과 같이 정현파 성분의 순시 주파수(instantaneous frequency)로 나타낼 수 있다.

$$\theta_i(t) = \int_0^t \omega_i(\tau) d\tau + \phi_i \tag{2}$$

여기서 ϕ_i 은 각 정현파 성분의 위상이 같지 않음을 나타낸다.

정현파 모델의 분석 과정은 정현파 성분의 진폭, 주파수, 위상 등의 파라미터들을 찾아내는 것이다. 정현파 모델의 분석/합성 과정은 그림 1로 나타낼 수 있다.

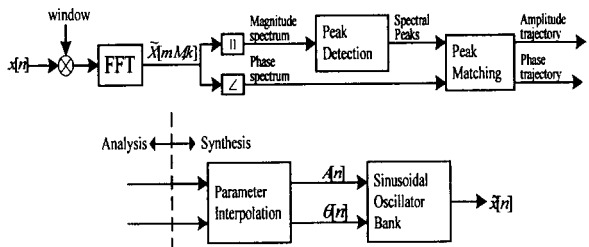


그림 1. 정현파 모델의 분석/합성 과정
Fig. 1. Sinusoidal analysis/synthesis process.

정현파 분석의 첫번째 단계는 단구간 푸리에 변환(short time Fourier transform:STFT)에 의해 신호의 일정한 구간에서의 진폭 스펙트럼과 위상 스펙트럼을 구하는 것

이다. STFT는 신호의 주파수 성분이 시간에 따라 어떻게 변화하는지를 나타내는 시간-주파수 분석 방법의 하나이다. STFT는 신호에 고정된 크기의 윈도우를 일정한 간격으로 이동시켜 가면서 적용시켜 푸리에 변환을 통해 국지적인 스펙트럼의 변화를 분석한다. 이것은 수식으로 다음과 같이 나타낼 수 있다.

$$\tilde{X}[mM, k] = \sum_{n=0}^{N-1} h[n]x[mM+n]e^{-j\frac{2\pi}{N}kn}, k=0, 1, \dots, N-1 \quad (3)$$

여기서 $h[n]$ 은 유한한 크기를 가지는 윈도우 함수를 나타내고, M 은 윈도우를 적용시키는 간격, 그리고 N 은 DFT(Discrete Fourier Transform) 크기를 나타낸다.

STFT의 출력값으로 각 주파수 성분에서의 진폭과 위상은 다음과 같이 계산된다.

$$|\tilde{X}[mM, k]| = \sqrt{\text{Re}(\tilde{X}[mM, k])^2 + \text{Im}(\tilde{X}[mM, k])^2} \quad (4)$$

$$\angle \tilde{X}[mM, k] = \tan^{-1} \frac{\text{Im}(\tilde{X}[mM, k])}{\text{Re}(\tilde{X}[mM, k])} \quad (5)$$

여기서 $\text{Re}(\cdot)$, $\text{Im}(\cdot)$ 은 각각 실수 부분과 허수 부분을 의미한다.

다음 단계는 STFT의 스펙트럼으로부터 신호에 포함된 정현파 성분을 찾아내는 피크 선택 과정이다. 이론적으로 단일 주파수와 고정된 진폭을 갖는 정현파는 스펙트럼 상에서 뚜렷한 피크로 나타난다. 음성 신호나 단일 악기음의 파형이 대부분 가장 안정적인 부분에 있어서도 완벽한 주기성을 나타낸다고 볼 수 없지만, STFT를 수행하는 구간이 어느 정도 안정적이고 주기성을 나타낸다면 스펙트럼 상에 나타나는 피크들은 신호에 포함된 정현파 성분으로 추정할 수 있다[1]. 정현파 분석에서의 피크는 진폭 스펙트럼 상에서 나타나는 국지적인 최대값으로 정의된다. 따라서 스펙트럼 분석에 사용되는 윈도우의 모양이 어떤 것이든 주파수 성분간의 간섭과 사이드로브(sidelobe) 피크의 영향을 무시할 수 없다. 이 때문에 피치 주파수를 정의할 수 있는 신호의 경우에는 피크를 찾는 주파수 범위와 피크의 최소 진폭 크기 등을 지정해서 실제 정현파 성분에 해당하는 피크만을 찾는 데 이용할 수 있다.

피크 선택 과정에서 나오는 것은 일정한 신호 구간에 포함된 각 정현파 성분의 주파수, 진폭, 위상에 대한 파라미터 값이다. 이것은 분석 프레임(일정 간격을 두고 윈도우를 적용시키는 구간)의 파형에서 나타나는 정현파 성분의 고정된 값이다. 이들 파라미터 값의 시간적인 변화는 각 프레임 간에 관계있는 정현파 성분들을 연결해 줌으로써 얻을 수 있다. 이 연결 과정이 피크 매칭 혹은 라인 트래킹(line tracking)이라 부르는 과정이다.

정현파 모델링에서 피크 매칭의 가장 간단한 방법은 각 프레임의 피크 주파수들 중 가장 가까운 것끼리 연결시키는 것이다. 이 방법에서는 최대 주파수 차이(maximum frequency difference)가 정의되고, 피크 매칭은 최대 주파수 차이 내에 있는 주파수 중에서 가장 가까운 것끼리 연결된다. 이것은 식 (6)으로 나타낼 수 있다.

$$|\omega_k^{i+1} - \omega_n^i| < \beta \omega_n^i \quad (6)$$

여기서 i 는 프레임 번호, k 와 n 은 주파수 번호를 나타내고, 최대 주파수 차이를 나타내는 β 값은 신호의 특성에 따라 실험을 통해 설정되는 것이 보통이다.

피크 매칭 과정을 통해 프레임간에 매치된 정현파 성분의 주파수, 진폭, 위상의 파라미터 세트는 정현파 합성에 필요한 샘플 단위의 파라미터 값으로 인터폴레이션된다. 프레임 i 의 n 번째 정현파 성분의 파라미터 세트 $(\omega_n^i, A_n^i, \theta_n^i)$ 와 이와 매치된 프레임 $i+1$ 의 k 번째 정현파 성분의 파라미터 세트 $(\omega_k^{i+1}, A_k^{i+1}, \theta_k^{i+1})$ 은 다음과 같이 샘플 단위의 값으로 인터폴레이션된다.

먼저, 진폭 파라미터의 경우는 식 (7)에서와 같이 단순히 선형 인터폴레이션을 사용하여 순시 진폭을 얻게 된다.

$$A[m] = A^i + \frac{A^{i+1} - A^i}{S} m, m=0, 1, \dots, S-1 \quad (7)$$

여기서 S 는 합성 프레임의 크기를 나타내고, m 은 합성 프레임 내의 샘플 번호를 나타낸다. 각 프레임의 주파수 번호는 편의상 생략하였다.

주파수와 위상에 관련된 파라미터 값은 간단한 선형 인터폴레이션을 통해 얻을 수 없다. 왜냐하면 순시 주파수는 순시 위상의 미분으로 나타나고, 실제로 정현파 합성에 필요한 것은 순시 위상 $\theta[m]$ 이기 때문이다. 일반적으로 정현파 모델에서는 순시 위상을 얻기 위해 3차 인터폴레이션을 사용한다[1].

$$\theta[m] = \xi + \gamma m + \alpha m^2 + \beta m^3 \quad (8)$$

여기서 $m=0$, 즉 합성 프레임의 시작 위치에서 주파수, 위상이 각각 ω^i, θ^i 이고, $m=S$, 즉 합성 프레임의 끝 위치에서의 주파수, 위상이 각각 $\omega^{i+1}, \theta^{i+1}$ 인 경계 조건을 이용하면, 식 (8)은 다음과 같이 표시된다.

$$\theta[m] = \theta^i + \omega^i m + \alpha m^2 + \beta m^3 \quad (9)$$

$$\alpha = \frac{3}{S^2} (\theta^{i+1} - \theta^i - \omega^i S + 2\pi P) - \frac{1}{S} (\omega^{i+1} - \omega^i), \quad (10)$$

$$\beta = -\frac{3}{S^3} (\theta^{i+1} - \theta^i - \omega^i S + 2\pi P) - \frac{1}{S^2} (\omega^{i+1} - \omega^i) \quad (11)$$

식 (10)과 식 (11)에서 α 와 β 값은 위상의 unwrapping factor P 에 따라 다르다. McAulay와 Quatieri[1]에 따르면 프레임간에 변하는 위상과 주파수의 값이 최대로 천천히 변하도록 다음과 같이 P 를 결정할 수 있다.

$$P^* = \frac{1}{2\pi} [(\theta' + \omega' S - \theta^{i+1}) + \frac{S}{2}(\omega^{i+1} - \omega^i)] \quad (12)$$

식 (12)로부터 P 는 P^* 에 가장 가까운 정수가 된다.

샘플 단위의 순시 진폭과 순시 위상이 구해지면 한 프레임 신호에 해당하는 정현파 합성은 식 (13)을 통해 합성된다.

$$x_i[n] \approx \hat{x}_i[n] = \sum_{l=1}^{Q_i} A_l[n] \cos(\theta_l[n]), \quad n=0,1,\dots,S-1 \quad (13)$$

III. 다중 해상도 필터 बैं크의 설계

단일 피치를 가지는 음성 신호나 단일 악기음과는 달리 폴리포닉 음은 피치를 정의할 수 없으므로, 스펙트럼 분석을 위한 윈도우 크기를 결정하기 어렵다. 이것을 해결하기 위한 하나의 방법이 다중 해상도 필터 बैं크를 사용하는 것이다. 다중 해상도 필터 बैं크는 constant Q나 웨이브릿에서처럼 서브밴드의 대역폭이 주파수에 따라 다른 것이다. 심리 음향 모델에 따르면 사람의 귀는 저주파수 대역에서는 주파수 해상도에 민감하고, 고주파수 대역에서는 시간 해상도에 민감하다고 알려져 있다. 이런 원리를 바탕으로 다중 해상도 필터 बैं크는 그림 2에 나타낸 것처럼 저주파수로 갈수록 서브밴드의 대역폭이 좁고, 고주파수로 갈수록 서브밴드의 대역폭이 넓다. STFT는 고정된 크기의 윈도우를 사용하므로 모든 주파수 영역에서의 대역폭이 일정하고, 저주파수 성분과 고주파수 성분의 시간-주파수 해상도가 동일하다고 볼 수 있다.

폴리포닉 음의 경우 정현파 분석을 위해 다중 해상도 필터 बैं크를 적용하면, 각 서브밴드 신호는 대역폭이 제한되므로 스펙트럼 분석을 위한 윈도우 크기를 어느 정도 결정할 수 있게 된다. 이 때 저주파수 대역의 서브밴드 신호에 대해서는 분석 윈도우 크기를 크게 함으로써 주파수 해상도를 높이고, 고주파수 대역의 서브밴드 신호에 대해서는 윈도우 크기를 작게 함으로써 시간 해상도를 높일 수 있다.

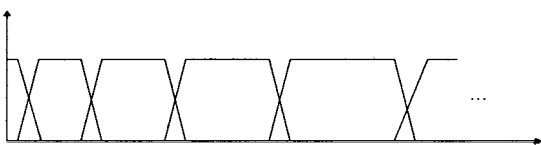


그림 2. 다중 해상도 필터 बैं크
Fig. 2. Multiresolution filter bank.

정현파 모델링을 위한 다중 해상도 필터 बैं크에 대한 연구는 Levine을 비롯한 몇 사람에게 의해 행해졌다[6-9]. 그 중에 Levine이 제안한 aliasing이 거의 없는 다중 해상도 필터 बैं크는 계산량과 음질면에서 폴리포닉 음의 정현파 분석에 적합한 것으로 알려져 있다. 이것을 바탕으로 본 논문에서는 그림 3과 같이 신호의 주파수 영역을 서로 다른 대역폭을 가지는 6개의 서브밴드로 나누었다. 각 서브밴드는 저주파수 대역으로 갈수록 상위 서브밴드의 대역폭보다 크기가 1/2인 옥타브 밴드 구조를 갖는다. 서브밴드 신호를 발생시키는 complementary 필터 बैं크의 구조는 그림 4와 같다. 여기서 데시메이션 필터의 컷오프(cutoff) 주파수는 0.4π 이고, 정지대역(stopband) 주파수는 0.45π , 그리고 필터 탭의 개수는 256이다. 이 필터 बैं크의 고주파 출력 신호는 2배로 오버샘플링 되어 있고, 저주파 출력 신호는 데시메이션 필터에 의해 aliasing이 방지되므로 aliasing이 거의 없는 서브밴드 신호를 얻을 수 있다. 물론 고주파 출력 신호에 대한 계산량은 증가되지만 고음질의 정현파 모델링을 위해서는 aliasing이 먼저 제거되어야 한다. 그림 3의 다중 해상도 필터 बैं크에서 complementary 필터 बैं크의 저주파 출력이 다시 다음 complementary 필터 बैं크의 입력이 되므로, 각 밴드에서 나타나는 신호는 원래 신호에 비해 샘플링 주파수가 각각 1배, 2배, 4배, 8배, 16배, 32배로 작아진 값이다. 따라서 이 구조에서 각 서브밴드 신호에 일정한 크기의 윈도우를 적용할 때, 실제로 원래 신호에 적용되는 윈도우 크기에 따른 각 서브밴드 신호의 시간-주파수 해상도는 그림 5와 같이 나타낼 수 있다.

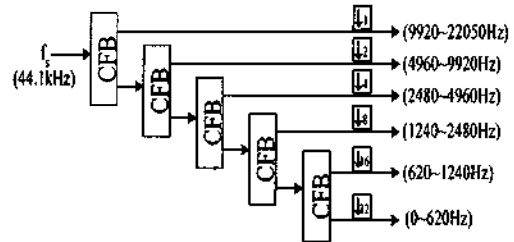


그림 3. 정현파 분석을 위한 다중 해상도 필터 बैं크
Fig. 3. Multiresolution filter bank for sinusoidal analysis.

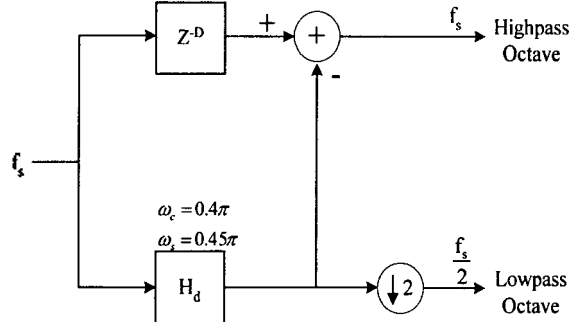


그림 4. Complementary 필터 बैं크
Fig. 4. Complementary filter bank.

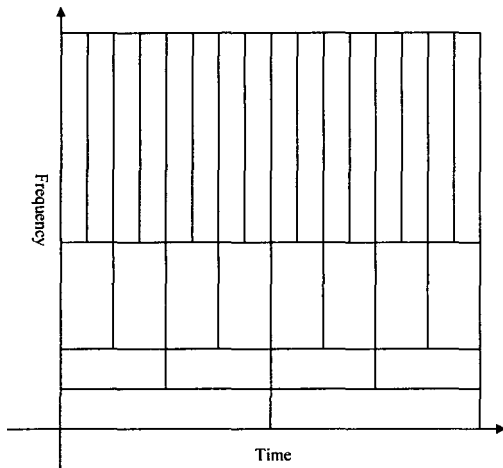


그림 5. 서브밴드 신호에 고정된 크기의 윈도우 적용시 시간-주파수 해상도

Fig. 5. Time-frequency resolution of multiresolution filter bank for a fixed window size.

그림 6은 서브밴드 신호에 대한 스펙트럼을 보여준다. 편의상 0~620Hz와 620~1240Hz의 두 개의 서브밴드 신호에 대해서만 나타내었다. 그림에서 보면 전체 주파수 대역의 스펙트럼에서 각 주파수 내역에 해당되는 부분이 서브밴드 신호의 스펙트럼에서 확대되어 나타난다. 이것은 서브밴드 신호가 complementary 필터 뱅크를 통과하면서 데시메이션되기 때문이다.

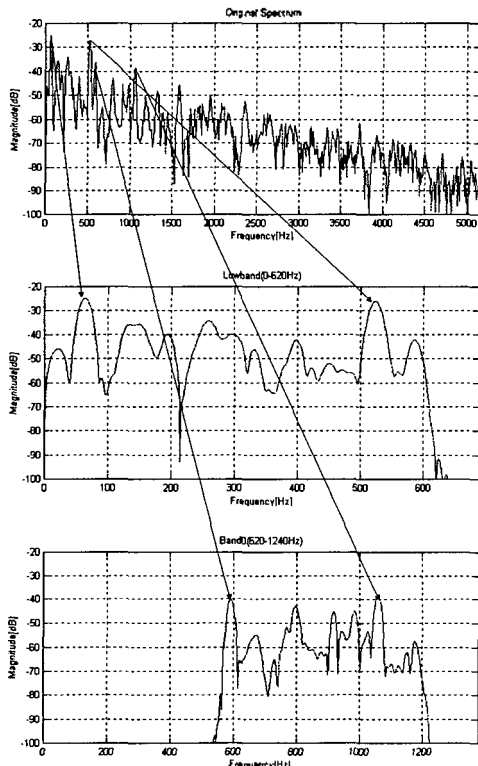


그림 6. 서브밴드 스펙트럼

Fig. 6. Subband spectrum.

IV. 천이 구간을 고려한 동적 세그멘테이션 방법

폴리포닉 음의 정현파 모델링에서 원래 음의 음질을 유지하기 위해서는 천이 구간에 대한 처리가 필요하다. 이것을 위한 하나의 방법은 Levine[6]이 제안한 것처럼 천이 구간의 파형을 그대로 보존하고, 천이 구간을 제외한 나머지 부분에 대해서만 정현파 모델을 적용하는 것이다. 천이 구간은 원래 파형을 그대로 복원할 수 있도록 DCT(Discrete Cosine Transform)를 이용한 코딩 방법을 사용하고, 합성시에는 단지 시간축으로 이동시켜 정현파 성분으로 합성된 부분과 연결한다. 비록 이러한 방법이 음의 천이 구간에 대한 특성을 그대로 유지시킬 수 있는 장점이 있으나, 분석과 합성 과정의 복잡도를 증가시키고, 어택이 많은 오디오 신호의 경우에는 천이 구간의 수에 따라 시간축 변화와 같은 신호 변형시에 음의 속도 조절에 문제를 가져올 수 있다. 또한 천이 구간 근처에서의 음의 속도가 정현파 합성 구간의 속도와 다르게 됨으로써 전체적으로 부자연스럽게 들릴 수도 있다.

본 논문에서는 천이 구간을 따로 처리하지 않고, 정현파 모델링을 그대로 적용한다. 천이 구간에 정현파 모델링을 적용하는데 있어서의 문제점은 일정한 크기의 분석/합성 프레임 크기로부터 발생된다. 정현파 분석을 통해 얻어지는 파라미터 값은 분석 프레임 내의 주파수 성분에 대한 평균값이다. 또한 정현파 합성시에는 합성 프레임과 인접 프레임 사이에서 인터플레이션이 일어나므로 신호 성분은 시간상으로 퍼져 버리게 된다. 이것이 어택과 같은 천이 구간을 무디게 만들고, 시간축 확장시에 음질을 저하시키는 가장 큰 원인이 된다. 이것을 해결하는 하나의 방법은 천이 구간 근처에서 분석/합성 프레임 크기를 작게 하는 것이다. 프레임 크기를 작게 하는 것은 신호의 시간 해상도를 높이는 것이므로 천이 구간과 같이 신호 성분이 짧은 시간 내에 급격하게 변하는 부분을 효과적으로 모델링할 수 있다.

신호의 구간별 특성에 따라 서로 다른 프레임 크기를 할당하는 것을 동적 세그멘테이션이라고 한다. Goodwin[10]은 정현파 모델링을 위한 몇 가지 동적 세그멘테이션 방법을 제안하였다. 여기서는 Goodwin이 기술한 동적 세그멘테이션 방법 중 동적 프로그램(dynamic program)을 이용한 방법과 포워드(forward) 세그멘테이션 방법을 소개하고, 계산량과 성능 면에서 좀 더 향상된 세그멘테이션 방법을 제안한다.

4.1. 동적 프로그램을 이용한 동적 세그멘테이션

동적 프로그램을 이용한 동적 세그멘테이션은 그림 7로 설명될 수 있다. 신호는 미리 설정된 합성 프레임 크기 중 가장 작은 프레임 크기의 셀이라는 영역으로 구분된다. 셀의 경계 부분을 노드라고 부르며, 그림 7에 각 노드의 일련 번호가 표시되어 있다. 세그멘테이션에서 허용되는 합성 프레임 크기는 셀 크기의 정수배로 최대

합성 프레임 크기까지이다. 각 노드에서는 하위 노드로부터 그 노드에 이르는 합성 구간에 대한 MSE(Mean Square Error)가 계산된다. 임의의 노드에서 MSE는 세그멘테이션에 허용되는 모든 합성 프레임 크기에서 계산되어진다. 동적 프로그램 방법에서 MSE는 식 (14)와 같이 정의될 수 있다.

$$D_{n_1, n_2} = \sum_{n=n_1, M}^{n_2, M+1} w[n](x[n] - \hat{x}[n])^2 \quad (14)$$

여기서 n_1 은 시작 노드 번호, n_2 는 끝 노드 번호, M 은 셀의 크기, 그리고 $w[n]$ 은 임의의 윈도우를 나타낸다.

동적 프로그램은 첫번째 노드부터 임의의 노드에 이를 때까지 계산된 MSE를 모두 더했을 때 가장 작은 값을 가지는 프레임 구성을 그 노드에 이를 때까지의 최적의 프레임 크기 구성으로 선택하게 된다. 그림 7을 참조하여 최대 프레임 크기는 셀 크기의 3배라고 가정하고, 6번 노드에 이를 때까지의 최적 프레임 구성을 구한다고 생각하자. 먼저, 노드 1에 이르는 방법은 노드 0에서 오는 한 가지 방법 밖에 없다. 이것으로부터 이 구간 동안의 MSE D_{01} 을 계산하고 이것을 $D(n_1)$ 으로 저장한다. 노드 2에 이르는 방법은 노드 1에서 오는 방법과 노드 0에서 오는 두 가지 방법이 있다. 먼저 D_{12} 와 D_{02} 가 계산된다. 이 두 개의 경로 중 노드 2에 이르는 최적의 경로는 이 노드에 이를 때까지 최소의 MSE 합, $\min(D(n_1) + D_{12}, D_{02})$ 을 나타내는 경로가 선택되고, 이것의 MSE 합을 $D(n_2)$ 로 저장한다. 노드 3에서는 노드 2, 노드 1, 그리고 노드 0에서 오는 것의 MSE에 해당하는 D_{23} , D_{13} , D_{03} 이 계산된다. 이 중 최소의 MSE 합, $\min(D(n_2) + D_{23}, D(n_1) + D_{13}, D_{03})$ 을 나타내는 경로가 선택되고, 이것의 MSE 합을 $D(n_3)$ 으로 저장한다.

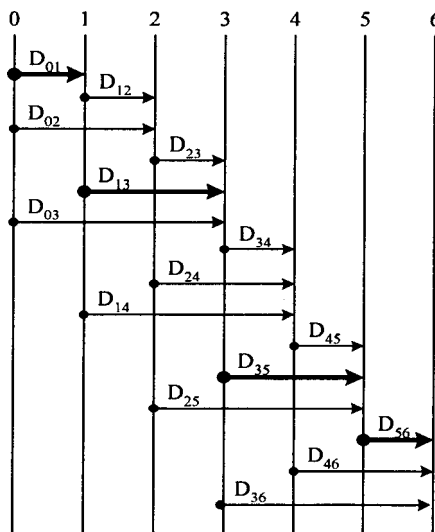


그림 7. 동적 프로그램을 이용한 동적 세그멘테이션 방법
Fig. 7. Dynamic segmentation using dynamic program.

이와 같은 과정을 노드 4, 5, 6까지 반복하면 노드 0부터 노드 6까지 이르는 동안 최소의 MSE를 가지는 프레임 구성을 찾을 수 있다. 그림 7에 노드 6에 이를 때까지의 프레임 구성의 한 예를 굵은 선으로 표시하였다.

4.2. 포워드 세그멘테이션 방법

포워드 세그멘테이션은 동적 프로그램 방법보다 계산량은 작은 반면 신호 전체에 대한 최적의 프레임 구성을 찾지 못한다. 포워드 방법은 신호의 국지 구간에 대한 MSE를 최소화하는 방향으로 프레임 크기를 설정하게 된다. 포워드 방법에서는 모든 노드에서의 MSE를 구하지 않으며, 후보 프레임 크기가 셀 크기의 정수배일 필요도 없다.

그림 8을 참조하여 포워드 세그멘테이션 방법에서 프레임 크기를 설정하는 방법을 살펴보자. 먼저 노드 0에서 세그멘테이션에 사용되는 각각 다른 크기의 후보 프레임들에 대해서 정현파 분석, 합성 과정을 통해 원래 신호와 합성 신호에 대한 MSE를 구하게 된다. 그림 8과 같이 후보 프레임 크기를 셀 크기의 1배, 2배, 3배로 했을 때는 D_{01} , D_{02} , D_{03} 이 구해지는 것이 된다. 이 중 최소의 MSE, $\min(D_{01}, D_{02}, D_{03})$ 를 나타내는 프레임 크기가 노드 0에서의 합성 프레임 크기로 선택된다. 다음으로는 노드 0에서 선택된 프레임의 끝 노드로 이동하여 그 노드에서 역시 후보 프레임들에 대해 최소의 MSE를 가지는 프레임 크기를 찾는다. 예를 들어 노드 0에서 D_{01} 에 해당하는 프레임 크기가 선택되었다고 하면, 다음으로는 노드 1으로 이동하여 이 노드에서 프레임 크기를 설정하는 것이다. 노드 1에서도 역시 모든 후보 프레임 크기에 대한 MSE, D_{12} , D_{13} , D_{14} 를 구하고 최소의 MSE를 갖는 프레임 크기를 선택한다. 여기서 D_{13} 이 선택되었다고 하면 이번에는 노드 3으로 이동하여 동일한 계산이 행해진다. 이러한 과정이 신호의 끝까지 반복되어 전체 프레임 구성이 이루어지게 된다.

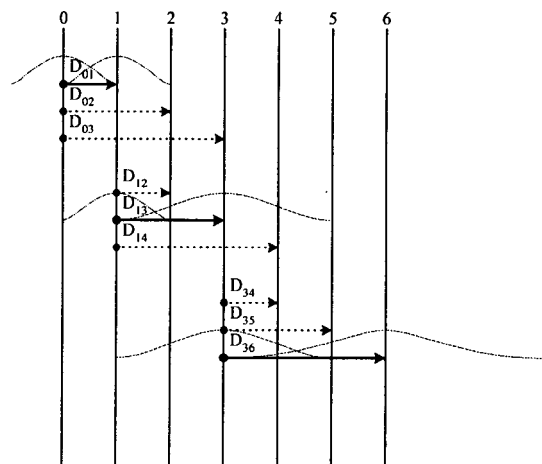


그림 8. 포워드 세그멘테이션 방법
Fig. 8. Forward segmentation

포워드 방법에서는 각 노드에서 합성되는 프레임 크기가 서로 다르므로 동적 프로그램 방법에서와 같이 합성 프레임에서 발생하는 절대적인 오차 에너지를 MSE로 사용할 수 없다. 포워드 세그멘테이션 방법에서는 식 (15)와 같이 MSE를 정의할 수 있다.

$$D_{n_1, n_2} = \frac{\sum_{n=n_1, M}^{n_2, M-1} w[n](x[n] - \hat{x}[n])^2}{\sum_{n=n_1, M}^{n_2, M-1} w[n](x[n])^2} \quad (15)$$

식 (15)는 원래 신호의 에너지에 대한 오차 신호의 에너지 비를 나타내는 것으로 이 값이 0에 가까울수록 정현파 모델링이 잘 되었음을 의미한다.

4.3. 제안하는 동적 세그멘테이션 방법

기존의 세그멘테이션 방법에서는 각 후보 프레임에 대한 MSE를 비교하기 위해 오버랩되는 두 개의 분석 프레임을 통해 정현파 합성을 하게 되므로 정현파 모델링의 분석/합성 과정이 모두 필요하다(그림 8에 분석 윈도우가 표시되어 있다). 이것은 스펙트럼 분석, 주파수 매칭과 파라미터 인터플레이션 등 많은 계산량을 필요로 하는 알고리즘이 후보 프레임의 개수만큼 반복되어야 하는 것을 의미하므로 계산량이 매우 많다. 본 논문에서 제안하는 방법은 두 개의 분석 프레임으로부터 합성 구간에 대한 파형을 합성하는 것이 아니라, 하나의 분석 프레임에서 찾아낸 정현파 성분으로 그 분석 프레임 구간 동안의 파형을 합성하고 MSE를 구하는 것이다. 이것을 그림 9를 참조하여 설명하면 다음과 같다.

먼저 노드 0에서 각 후보 합성 프레임 크기에 해당하는 분석 프레임에서 정현파 성분의 추출과 합성이 이루어진다. 그림 9에서는 노드 0-1 구간의 합성 프레임에 해당하는 분석 프레임이 노드 0-2가 되고, 노드 0-2 구간의 합성 프레임에 해당하는 분석 프레임이 노드 0-4, 그리고 노드 0-3 구간의 합성 프레임에 해당하는 분석 프레임이 노드 0-6이 된다. 이들 분석 프레임에서 정현파

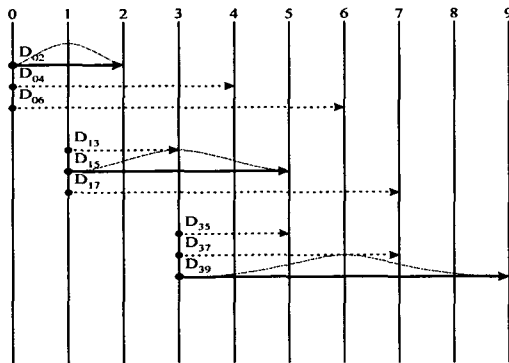


그림 9. 제안하는 동적 세그멘테이션 방법
Fig. 9. Proposed dynamic segmentation method.

분석/합성 과정을 통해 분석 프레임 구간에 해당하는 파형을 합성하고, 이것으로부터 각각의 MSE D_{02} , D_{04} , D_{06} 이 계산된다. 이들 중 최소의 MSE를 나타내는 분석 프레임 크기가 노드 0에서의 프레임 크기가 된다. 예를 들어 D_{02} 가 최소의 값을 나타낸다면 노드 0에서의 분석 프레임은 노드 0-2 구간이 되고, 합성 프레임은 노드 0-1 구간이 된다. 다음으로 프레임 크기가 설정되어야 할 노드는 노드 0에서 설정된 합성 프레임 구간의 끝 노드가 된다. 즉 선택된 분석 프레임의 중앙으로 이동하여 위의 과정이 반복된다. 그림 10에서는 노드 0에서 합성 프레임이 노드 0-1 구간이 설정되었으므로 노드 1로 이동하여 D_{13} , D_{15} , D_{17} 을 계산한다. 노드 1에서는 D_{13} 가 최소의 MSE를 나타내므로 합성 프레임은 노드 1-3 구간이 되고, 노드 3으로 이동한다. 이와 같은 과정이 신호의 끝까지 반복되어 전체 신호에 대한 프레임 구성을 찾아낸다.

개선된 방법에서는 두 분석 프레임 간의 정현파 성분에 대한 피크 매칭, 파라미터 인터플레이션 과정이 필요 없으므로 기존의 방법보다 훨씬 계산량이 줄어든다. 또한 정현파 합성 과정도 역푸리에 변환(IFFT)을 이용해 간단하게 할 수 있다. IFFT를 이용한 합성 방법은 프레임 신호에 포함된 정현파 성분의 일정한 주파수, 진폭, 위상으로 신호를 추정하는 것을 의미하므로, 주기적이고 안정된 신호일수록 합성 결과는 원래 신호의 파형에 가깝게 나오게 된다. 그림 10은 분석 프레임 크기가 32, 64, 96일 때 어택이 나타나는 근처에서의 프레임 신호와 합성 신호를 나타내고 있다. 그림에서 실선으로 표시된 것이 프레임 신호이고 점선으로 표시된 것이 합성 신호이다. 그림에서 보면 분석 프레임이 32일 때는 천이 구간이 포함되어 있지 않으므로 합성 신호는 원래 신호에 가깝다. 반면 64, 96의 경우에는 천이 구간이 많이 포함될수록 오차 신호의 값이 증가되는 것을 알 수 있다. 따라서 천이 구간에 가까워질수록 프레임 크기는 작게 설정되어 천이 구간 파형이 시간상으로 퍼지는 효과를 줄일 수 있다.

한편, 개선된 방법에서는 식 (16)과 같이 MSE가 정의된다.

$$D_{n_1, n_2} = \frac{\sum_{n=n_1, M}^{n_2, M-1} w[n](x[n] - \hat{x}[n])^2}{N_w \left(\sum_{n=0}^{N_w-1} w[n] \right)} \quad (16)$$

여기서 $x[n]$ 이 분석 프레임 구간의 원래 신호, $\hat{x}[n]$ 이 IFFT를 통해 합성한 신호, $w[n]$ 은 임의의 윈도우를 나타내고, n_1 은 분석 프레임의 시작 노드 번호, n_2 는 끝 노드 번호, M 은 셀의 크기, 그리고 $N_w (= n_2 M - n_1 M)$ 는 분석 프레임의 크기를 나타낸다. 이것이 식 (15)와 다른 것은 먼저 오차 에너지를 원래 신호의 에너지로 나누지 않고, 프레임의 크기로 나누어 주는 것이다. 이것은 샘플

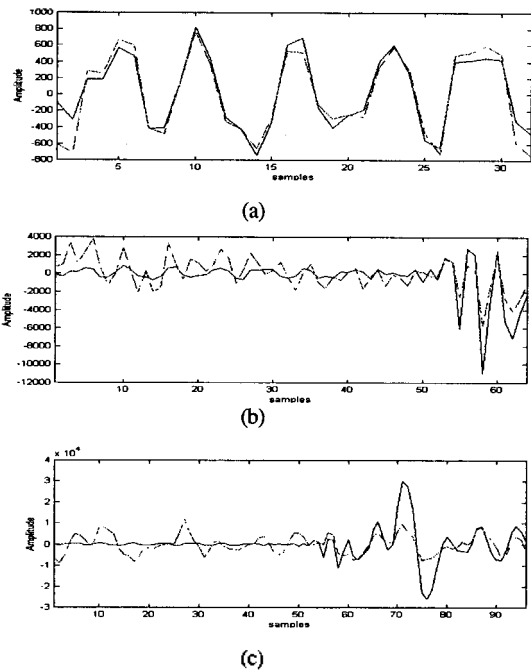


그림 10. 천이 구간 근처에서 각 프레임 크기에 대한 합성 신호 비교
 (a) 32 (b) 64 (c) 96 samples
 Fig. 10. Comparison of synthesized signals for each frame size about transient;
 (a) 32, (b) 64, (c) 96 samples.

당 평균 오차 에너지를 구하는 것이라 볼 수 있다. 오차 에너지를 원래 신호의 에너지로 나누어 주면, 비록 서로 다른 크기의 프레임에서 오차 에너지의 값이 비슷하게 나온다 해도 원래 신호의 에너지가 클수록 MSE는 작게 나오게 된다. 이것은 어택과 같이 에너지가 큰 천이 구간이 포함되는 프레임의 경우 프레임 크기가 클수록 원래 신호의 에너지가 증가하므로 MSE가 더 작게 나오는 것을 의미한다. 따라서 천이 구간 근처에서는 작은 프레임이 설정되어야 함에도 불구하고 오히려 프레임 크기가 큰 것이 선택될 가능성이 있다. 한편, 식 (16)과 같이 오차 에너지를 프레임 크기로 나누어주면 어느 정도 절대적인 오차 에너지를 알 수 있으므로 천이 구간의 프레임 크기를 올바르게 설정할 수 있다. 그러나 오차 에너지를 단순히 프레임 크기로만 나누어 주면, 천이 구간이 아닌 다른 정적인 구간에서도 프레임 크기가 작게 선택될 가능성이 많다. 천이 구간이 아닌 영역에서는 주파수 해상도를 높이기 위해 프레임 크기를 되도록이면 크게 하는 것이 유리하다. 식 (16)에서 분모에 나타나는 윈도우의 샘플값의 합은 천이 구간이 아닌 영역에서 프레임 크기가 크게 설정될 수 있도록 작용한다. 윈도우의 크기가 클수록 샘플값의 합도 증가하므로 MSE는 작게 나오기 때문이다. 본 논문에서는 일반적인 hanning 타입의 윈도우를 사용하였다.

그림 11은 지금까지 설명된 동적 세그멘테이션 방법들을 비교한 결과를 보여주고 있다. 동적 프로그램을 이용한 방법과 포워드 방법에서는 정적인 구간에서도 프레임 크기가 작게 설정되어 있는 것을 알 수 있다. 그러나 제안하는 방법에서는 신호의 천이 구간에서만 프레임 크기가 작게 설정되고 나머지 구간에서는 큰 프레임 크기가 설정됨을 확인할 수 있다.

그림 12는 서브밴드 신호에 동적 세그멘테이션을 적용한 경우의 각 서브밴드 신호에 대한 시간-주파수 해상도를 나타낸다. 그림 5와 비교해 볼 때 고정된 필터뱅크 구조보다 좀 더 신호의 구간별 특성을 잘 반영할 수 있음을 알 수 있다.

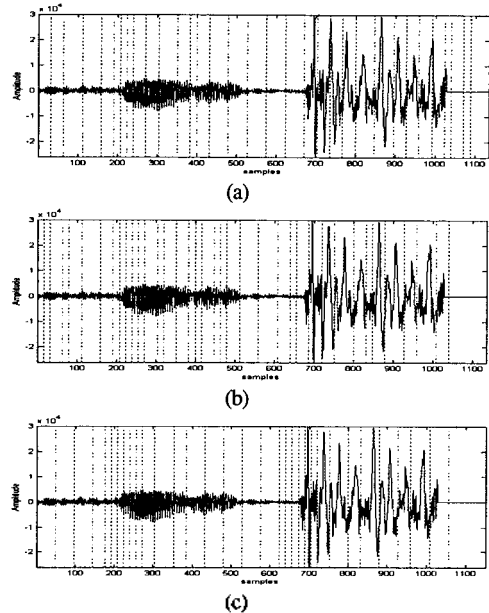


그림 11. 동적 세그멘테이션 결과 비교
 (a) 동적 프로그램을 이용한 방법
 (b) 포워드 세그멘테이션
 (c) 제안하는 동적 세그멘테이션 방법

Fig. 11. The results of dynamic segmentation methods;
 (a) Dynamic program,
 (b) Forward segmentation,
 (c) Proposed method.

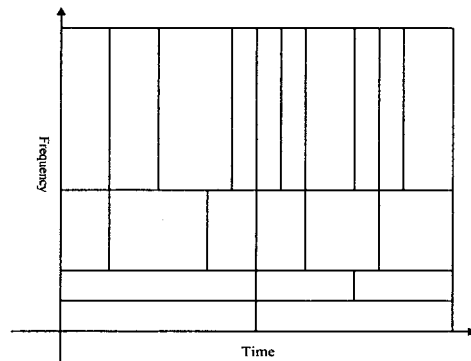


그림 12. 서브밴드 신호에 동적 세그멘테이션 적용시 시간-주파수 해상도

Fig. 12. Time-frequency resolution of multiresolution filter bank for dynamic segmentation.

V. 실험 및 고찰

본 논문에서 제안된 알고리즘을 검증하기 위해 실험에 사용된 오디오 데이터는 표 1에 나타낸 것과 같이 음성 신호에서 오케스트라 연주 음악에 이르기까지 다양한 종류의 음악 신호를 포함한다.

표 1. 실험에 사용된 오디오 데이터 목록

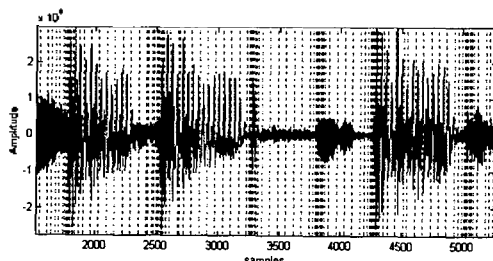
Table 1. List of audio data for simulation.

곡 이름	종류
Orig1	팝 발라드
Orig2	연주음악
Figaro	클래식 성악곡
Pop	팝 랩음악
Titasub	영화음악
Speech1	CNN 뉴스
Speech2	여성 인사말

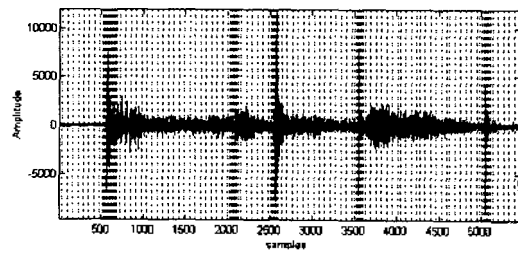
그림 13에서 그림 14는 표 1의 오디오 데이터 중 편의상 'orig1' 음과 'orig2' 음에 대해서만 실험 결과를 나타낸 것이다.

그림 13은 'orig1' 음의 low band(0~620Hz)와 band0(620~1240Hz) 신호의 일부 구간에 대한 동적 세그멘테이션 결과를 보여준다. 그림에서 보면 신호의 어택 근처에서 프레임 크기가 작게 나오는 것을 확인할 수 있다.

그림 14는 'orig2' 음의 일부 천이 구간 근처에서 본 논문에서 제안된 동적 세그멘테이션 알고리즘으로 합성한 신호와 서브밴드 신호에 고정된 윈도우 크기를 적용해서 합성한 신호를 비교한 것이다. 그림 14의 (a)는 원래 신호를 나타내고, (b)와 (c)는 각각 동적 세그멘테이션을 이용해 합성한 신호와 원래 신호와의 오차 신호, 그리고 (d)와 (e)는 각각 고정된 윈도우 크기를 적용해 합성한 신호와 원래 신호와의 오차 신호를 나타낸다. 그림에서 보면 본 논문에서 제안한 방식



(a)



(b)

그림 13. 'orig1' 음에 대한 동적 세그멘테이션 결과

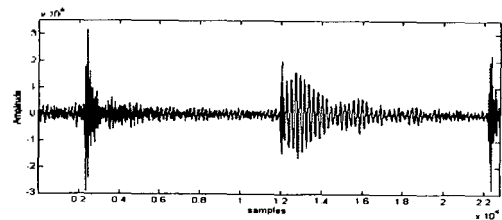
(a) Lowband(0~620Hz)

(b) Band0(620~1240Hz)

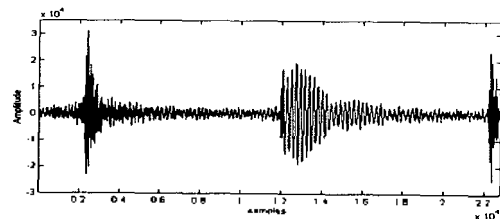
Fig. 13. The result of dynamic segmentation for 'orig1';

(a) Lowband(0~620Hz),

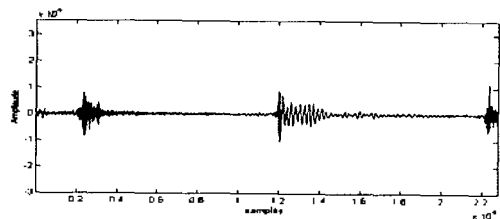
(b) Band0(620~1240Hz).



(a)



(b)



(c)

그림 14. 서브밴드 신호에 대한 동적 세그멘테이션 방법과 고정

세그멘테이션 방법의 합성 신호와 오차 신호 비교

(a) 원래 신호('orig2' 음)

(b) 동적 세그멘테이션 방법으로 합성한 신호

(c) 동적 세그멘테이션 방법에서의 오차 신호

Fig. 14. Comparison of synthesized signals using dynamic segmentation and fixed segmentation method;

(a) Original signal('orig2'),

(b) Synthesized signal using dynamic segmentation method,

(c) Residual signal of dynamic segmentation.

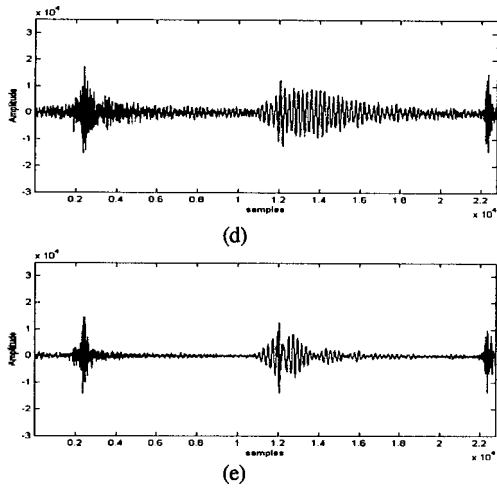


그림 14. 계속

- (d) 고정 세그멘테이션 방법으로 합성한 신호
- (e) 고정 세그멘테이션 방법에서의 오차 신호

Fig. 14. Continued;

- (d) Synthesized signal using fixed segmentation method,
- (e) Residual signal of fixed segmentation method.

으로 정현파 합성한 신호가 천이 구간 근처에서 천이 구간의 특성을 보다 잘 표현하고 있음을 확인할 수 있다.

본 논문에서 제안된 알고리즘이 실제로 얼마나 우수한 음질을 나타내는가를 살펴보기 위해 주관적인 음질 평가를 수행하였다. 음질 평가는 표 1에 나타난 테스트 음으로 부산대학교 전자공학과 대학원생 10명을 상대로 수행하였다. 음질 평가 방법은 제안된 정현파 모델링 방법으로 원래 신호를 합성했을 때 얼마나 원래 신호와 가까운가 하는 것이다. 이것은 원래 음과 합성한 음을 차례로 들려준 다음 원래 음에 가까운 정도를 4점(구별 안됨)에서 0점(음질 매우 떨어짐)까지의 다섯 단계로 나누어 점수를 매기도록 하였다. 표 2에 나타난 결과를 보면 테스트에 사용된 모든 음에 대해 비교적 우수한 음질을 얻을 수 있음을 확인할 수 있다.

표 2. 합성음의 음질에 대한 평가 결과

Table 2. The result of quality for the synthesized signals.

시료	점수
Orig1	3.8
Orig2	3.7
Figaro	3.3
Pop	3.7
Titasub	3.4
Speech1	3.6
Speech2	3.9

VI. 결 론

본 논문은 폴리포닉 음과 같은 복잡한 스펙트럼을 갖는 오디오 신호에 대해 천이 구간을 고려해서 정현파 모델링하는 방법을 제안하였다.

정현파 모델링은 지금까지 주로 음성 신호나 단일 악기음에 대해 연구되어 왔으므로 이것을 그대로 폴리포닉 음에 적용시키기 어렵다. 가장 큰 문제점은 폴리포닉 음의 경우 여러 개의 악기음이 섞여 있으므로 단일 피치 주파수에 의한 분석 윈도우의 크기를 결정할 수 없다는 것이다. 또한 고음질의 합성음을 얻기 위해서는 악기음의 음색을 결정짓는 어택과 같은 천이 구간의 특성을 보존하는 것이 필요하다.

본 논문에서는 폴리포닉 음의 정현파 분석을 위해 옥타브 밴드 구조의 다중 해상도 필터 뱅크를 설계하였고, 이 필터 뱅크의 출력인 서브밴드 신호에 대해 정현파 모델링을 행하였다. 서브밴드 신호는 대역폭이 제한된 신호이므로, 각 서브밴드 신호별로 시간-주파수 해상도에 따라 적절한 분석 윈도우 크기를 설정할 수 있다. 천이 구간에 대해서는 정현파 모델링 방법을 그대로 적용함으로써 기존 알고리즘에 비해 복잡도를 줄였다. 이를 위해 각 서브밴드 신호에 동적 세그멘테이션을 적용하여 천이 구간에서의 분석 합성 프레임 크기를 작게 함으로써 기존 정현파 모델링에서의 퍼짐 현상을 개선하고, 시간축 변화시에도 천이 구간의 특성을 그대로 보존함으로써 우수한 음질을 얻을 수 있었다. 동적 세그멘테이션 방법으로는 기존의 방법보다 계산량과 성능 면에서 향상된 방법을 제안하였다.

본 논문의 알고리즘을 적용해 다양한 폴리포닉 음에 대해 실험한 결과 제안한 정현파 모델링 방법이 원래 신호의 음질을 잘 복원할 수 있고, 신호의 천이 구간을 잘 표현할 수 있음을 확인하였다. 주관적인 음질 평가를 통해 여러가지 시간축 알고리즘을 비교한 결과 폴리포닉 음에 대해 본 논문의 알고리즘의 정현파 모델링 방법이 우수한 음질을 내는 것을 확인하였다.

참 고 문 헌

1. R. J. McAulay and T. F. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," *IEEE Trans. on Acoust., Speech, and Signal Processing*, vol. 34(4), pp. 744-754, 1986.
2. T. F. Quatieri and R. J. McAulay, "Speech transformations based on a sinusoidal representation," *IEEE Trans. on Acoust., Speech, and Signal Processing*, vol. 34(6), pp. 1449-1464, 1986.
3. E. B. George and M. J. T. Smith, "Analysis-by-synthesis/Overlap-add sinusoidal modeling applied to the analysis and synthesis of musical tones," *J. Audio*

- Eng. Soc., vol. 40(6), pp. 497-516, 1992.
4. X. Serra, A system for sound analysis/transformation/synthesis based on a deterministic plus stochastic decomposition, Ph. D. thesis, Stanford University, 1989.
 5. S. Furui, "On the role of spectral transition for speech perception," J. Acoust. Soc. Amer., vol. 80, pp. 1016-1025, 1979.
 6. S. N. Levine, Audio representations for data compression and compressed domain processing, Ph. D. thesis, Stanford University, 1998.
 7. D. V. Anderson, "Speech analysis and coding using a multiresolution sinusoidal transform," Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, 1996.
 8. M. Rodriguez-Hernandez and F. Casajus-Quiros, "Improving time-scale modification of audio signals using wavelets." ICSPAT-1994, vol. 2, pp. 1573-1577, 1994.
 9. B. Tang, A. Shen, G. Pottie, and A. Alwon, "Spectral analysis of subband filtered signals," Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, 1995.
 10. M. Goodwin, "Multiresolution sinusoidal modeling using adaptive segmentation," Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, pp. 1525-1528, 1998.

▲장 호 근(Ho Keun Jang)

1968년 1월 15일생

1993년 2월 : 부산대학교 전자공학과

1995년 2월 : 부산대학교 전자공학과 공학석사

2000년 2월 : 부산대학교 전자공학과 공학박사

2000년 2월 ~ 현재 : 현대전자 시스템 IC 선임 연구원

※ 주관심분야 : 오디오 신호 처리, DSP 설계

▲박 주 성(Ju Sung Park)

1953년 12월 19일생

1976년 2월 : 부산대학교 전자공학과

1978년 2월 : 한국과학기술원 전기 및 전자공학과 공학석사

1978년 3월~1985년 7월 : 한국전자기술연구소

1985년 8월~1989년 7월 : University of Florida, Ph.D.

1989년 8월~1991년 3월 : 한국전자통신연구소 책임연구원

1991년 3월~현재 : 부산대학교 전자공학과 부교수

※ 주관심분야 : DSP 설계 및 응용, 사운드 합성, 반도체 소자 모델링