

캡스트럼 포락선을 이용한 해금 소리의 포먼트 합성

Formant Synthesis of Haegeum Sounds Using Cepstral Envelope

홍연우*, 조상진**, 김종면*, 정의필*
(Yeonwoo Hong*, Sangjin Cho**, Jongmyon Kim*, Uipil Chong*)

*울산대학교 컴퓨터정보통신공학부, **울산대학교 전기전자정보시스템공학부
(접수일자: 2009년 5월 25일; 채택일자: 2009년 6월 15일)

본 논문에서는 전통 현악기 해금의 스펙트럼 모델링을 위해 캡스트럼 포락선을 이용한 포먼트 합성법을 제안한다. 스펙트럼 모델링은 입력 신호를 정현파 성분과 노이즈 성분의 합으로 해석하여 음을 합성하는 기술로 주기성이 있는 현악기나 관악기의 음 합성에 효과적이다. 캡스트럼 포락선의 포먼트는 정현파 성분 합성을 위한 파라미터로 활용하였다. 정현파 성분을 합성하기 위해 기존의 가산합성 방식과는 달리 IIT (Impulse Invariant Transform)로 공명기를 설계하였으며 배음간 크기 보안을 위해 대역 통과 필터를 추가하였다. 원음과 합성된 정현파 성분의 차로 구해진 노이즈 성분에 포함된 일부 유효 배음을 제거하면 완전한 노이즈 성분을 구할 수 있으며 선형 보간법 (linear interpolation)에 기초하여 그 주파수 특성을 파라미터화 하였다. 최종적으로 합성된 노이즈 성분과 정현파 성분을 더하여 해금 단위음을 합성하였고 합성음은 원음과 매우 유사하였다.

핵심용어: 해금 소리 합성, 스펙트럼 모델링, 포먼트 합성, 캡스트럼 포락선, 스펙트럼 분석

투고분야: 음악음향 및 음향심리 분야 (8.6)

This paper proposes a formant synthesis method of Haegeum sounds using cepstral envelope for spectral modeling. Spectral modeling synthesis (SMS) is a technique that models time-varying spectra as a combination of sinusoids (the "deterministic" part), and a time-varying filtered noise component (the "stochastic" part). SMS is appropriate for synthesizing sounds of string and wind instruments whose harmonics are evenly distributed over whole frequency band. Formants extracted from cepstral envelope are parameterized for synthesis of sinusoids. A resonator by Impulse Invariant Transform (IIT) is applied to synthesize sinusoids and the results are bandpass filtered to adjust magnitude. The noise is calculated by first generating the sinusoids with formant synthesis, subtracting them from the original sound, and then removing some harmonics remained. Linear interpolation is used to model noise. The synthesized sounds are made by summing sinusoids, which are shown to be similar to the original Haegeum sounds.

Keywords: Sound synthesis of Haegeum, Spectral modeling, Formant synthesis, Cepstral envelope, Spectral analysis

ASK subject classification: Musical Acoustics and Psychoacoustics (8.6)

I. 서론

악기 음 합성에 대한 연구는 샘플링, 변조, 필터링, 모델링 등의 방식으로 꾸준히 연구되어왔다 [1]. 샘플링과 필터링은 가장 전통적인 방법으로 녹음된 악기 음을 사용하거나 필터를 이용하여 원하는 형태의 파형이나 스펙트럼을 만드는 방식이다. 샘플링 계통은 악기 본래의 음색이 가장 잘 표현되는 자연스러운 소리를 합성할 수 있으나 단일 형태의 연주 방식에 적합하며 연주 형태가 바

뀌면 다시 샘플링을 해야 하는 단점이 있다. 변조 방식은 자연의 소리가 아닌 전자음과 같은 새로운 소리를 창조한다는 장점이 있지만 인위적인 느낌이 강하여 다양한 악기의 자연스러운 소리를 합성하는 데는 제한이 있다. 반면에 모델링 방식은 악기의 구조적, 음향적 특성을 이용하여 설계한 디지털 필터를 통해 음을 합성하는 방식으로 필터의 파라미터를 바꿈으로써 음색을 조절할 수 있고 특정 필터를 추가, 제거함으로써 연주방식을 표현할 수 있지만 계산량이 많은 단점이 있다 [2].

악기 음을 사실적으로 재생하기 위해서는 각 악기의 특성에 맞는 적절한 모델링 방법을 선택하는 것이 중요한데, 모델링을 이용한 악기음 합성에는 물리적 모델링

책임저자: 정의필 (upchong@ulsan.ac.kr)
680-749 울산광역시 남구 대학로 102 울산대학교 컴퓨터정보통신공학부
(전화: 052-259-2220; 팩스: 052-259-1687)

(physical modeling)과 스펙트럼 모델링 (spectral modeling) 방식이 있다. 물리적 모델링은 악기의 소리 발생 원리를 분석하고 그에 적합한 모델을 만들어 음을 합성하는 것으로 디지털 도파관 (digital waveguide)을 이용한 방식이 음 합성 분야에서 많이 사용되고 있다. 스펙트럼 모델링은 대상 악기 음의 스펙트럼을 정현파 성분과 음의 특성에 영향을 미치는 다른 성분들의 합으로 해석하고 이를 통해 음을 합성하는 기술로 주기성 있는 현악기나 관악기의 음 합성에 효과적이다 [3]. 스펙트럼 모델링의 주요 장점은 음 합성을 위한 파라미터를 실제 사운드로부터 추출하기 때문에 사실적인 소리를 만들어 낼 수 있다는 것이다 [4].

스펙트럼 모델링은 최초의 음 합성 기술인 가산합성 방식에서 출발하였고 일반적으로 정현파 성분과 노이즈 성분의 합으로 악기 음을 모델링하는 방식을 말한다 [5]. 노이즈 성분은 음 높이와는 크게 상관이 없으므로 정현파 성분을 모델링 하는 방법이 더 중요하다. 정현파 성분의 모델링에는 가산합성 (additive synthesis), 감산합성 (subtractive synthesis), 포먼트 합성 (formant synthesis) 등이 쓰일 수 있다. 포먼트 합성은 주파수 영역에서 스펙트럼 포락선 (spectral envelope)을 통해 배음 성분들을 만들어 내며 각 배음의 크기와 주파수 정보를 모두 필요로 하지 않는다는 점과 배음 이외의 성분들도 합성해 낼 수 있다는 장점이 있다 [6]. 포먼트 합성에서 스펙트럼 포락선을 추출하는 방법으로는 보간법 (interpolation), LPC (Linear Prediction Coding), 캡스트럼 (cepstrum) 등이 있으며 본 논문에서는 캡스트럼 포락선을 이용하였다.

캡스트럼은 지진과 연구를 목적으로 1963년 Bogert 등에 의해 연구 되었으며 Oppenheim의 호모모픽 시스템 (homomorphic system)과 그 맥락을 같이한다 [7]. 입력 신호의 스펙트럼에 로그를 취하고 푸리에 변환하면 캡스트럼 계수를 얻을 수 있다. 시간 영역에서 컨볼루션으로 결합된 입력 신호는 주파수 영역에서 곱셈의 형태로 바뀌게 되고 로그를 취함으로써 덧셈의 형태로 다시 바뀌게 된다. 이를 다시 푸리에 변환하면 입력 신호에 포함된 성분들이 덧셈의 형태로 중첩되어 있다고 해석할 수 있으므로 입력 신호의 정현파 성분과 노이즈 성분은 캡스트럼 영역에서 쉽게 분리할 수 있다.

국악기 음 합성에 관한 연구로는 산조가야금의 물리적 모델링 [2,8], 대금, 피리의 음색 연구 [9-11], 국악기 샘플링 신디사이저 [12], 해금의 음향 특성 연구 [13] 등이 있으나 이는 국외의 활발한 연구에 비해 미비한 실정이다. 특히 해금의 경우 음향학적 특성 연구 정도에 그쳐

해금 음을 분석하여 합성하고자 한 연구는 아직 이루어지지 않았다. 해금은 비사비죽 (非絲非竹)의 악기라고 불릴 만큼 관악기의 특성을 많이 가진 악기로서 관악기와 유사하게 배음이 매우 높은 대역까지 규칙적으로 존재하므로 스펙트럼 모델링에 적합하다 [14].

본 논문에서는 해금의 스펙트럼 모델링을 위해 캡스트럼 포락선을 이용한 포먼트 합성법을 제안한다. 캡스트럼 포락선에서 포먼트를 추출하여 정현파 성분을 합성하기 위한 파라미터로 활용하였다. 정현파 성분을 합성하기 위해 IIT (Impulse Invariant Transform) 에 의한 필터를 공명기로 사용하였으며 크기 보인을 위해 공명기의 출력에 대역 통과 필터를 적용하였다. 원음에서 포먼트 합성을 통해 재현된 정현파 성분을 빼면 노이즈 성분을 구할 수 있으며 선형 보간법 (linear interpolation)에 기초하여 그 주파수 특성을 파라미터화 하였다. 합성된 노이즈 성분과 정현파 성분을 더하여 해금 단위를 합성하였다.

II. 해금

해금은 활로 현을 마찰해서 소리를 내는찰현악기로 6세기경 중국의 해 (奚) 부족 문화권에서 유래하여 고려시대부터 우리나라에서 연주되기 시작했다 [15]. 해금은 그림 1과 같이 통, 목판, 입죽, 활대로 이루어져 있으며 두 현은 각각 중현과 유현이라 하며 활대는 중현과 유현 사이에 위치한다. 해금을 연주할 때는 왼손을 입죽에 대고 줄을 쥐고 중현과 유현 사이의 활대를 잡아 활을

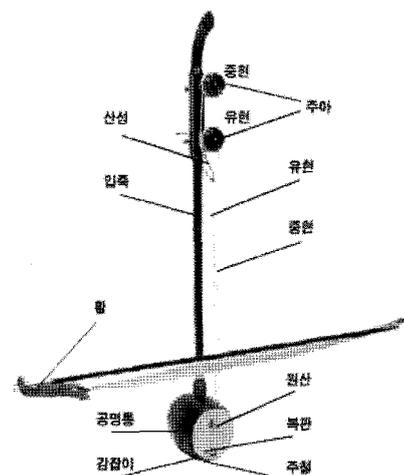


그림 1. 해금의 구조
Fig. 1. Structure of Haegeum.

밀거나 당기면서 소리를 낸다. 해금의 음역은 두 옥타브 반 정도이며 필요에 따라 더 넓어질 수 있다. 해금의 운지법은 크게 3가지로 이는 각각 주아 가까운 쪽에 왼손을 위치시키는 황종 운지법, 줄의 가운데 부분을 잡고 소리를 내는 중려 운지법, 울림통 가까운 곳에 위치시키고 소리를 내는 청황종 운지법이다. 왼손의 위치가 울림통에 가까워질수록 높은 소리를 내며 주아에 가까워질수록 낮은 소리를 낸다.

III. 캡스트럼

입력 신호를 $x(n)$ 이라 하면, $x(n)$ 은 식 (1)과 같이 여기 신호 $e(n)$ 과 이를 필터링하여 출력 신호를 만들어 내는 $s(n)$ 의 컨볼루션으로 표현된다.

$$x(n) = e(n) * s(n) \tag{1}$$

이를 푸리에 변환한 후 로그를 취하면 식 (2)가 된다.

$$\log [X(e^{j\omega})] = \log [E(e^{j\omega})] + \log [S(e^{j\omega})] \tag{2}$$

이 때, 여기 신호의 스펙트럼인 $E(e^{j\omega})$ 는 노이즈처럼 상대적으로 크기가 빠르게 변하는 ω 의 함수로, 필터의 스펙트럼인 $S(e^{j\omega})$ 는 악기음의 배음처럼 크기가 비교적 느리게 변하는 ω 의 함수로 볼 수 있으므로 입력 신호의 스펙트럼은 이 두 함수로 분리된다. 식 (2)의 좌변을 IDFT (Inverse Discrete Fourier Transform) 하면 캡스트럼 계수를 얻을 수 있으며 이는 식 (3)과 같다.

$$c(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log [X(e^{j\omega})] e^{j\omega n} d\omega \tag{3}$$

이 때, $c(n)$ 은 n 번째 캡스트럼 계수가 된다 [16]. 캡스트럼 계수가 큰 부분을 적절한 크기의 윈도우로 분리하여 푸리에 변환을 하면 주기성을 가지는 정현파 성분들에 대한 캡스트럼 포락선을 구할 수 있다. 나머지 캡스트럼 계수를 푸리에 변환하면 노이즈 성분을 얻을 수 있으므로 입력 신호에 포함된 정현파 성분과 노이즈 성분은 이론적으로 캡스트럼 영역에서 분리될 수 있다.

IV. 해금 합성 파라미터 추출

4.1. 해금 소리의 주파수 특성

해금 소리는 활과 현의 마찰에 의해 발생하므로 연주자가 활을 밀고 당기는 속도가 가장 큰 영향을 준다고 할 수 있다. 연주자가 활을 밀거나 당기는 동안은 계속해서 소리가 나고 속도를 줄이면 크기가 줄어든다. 해금 파형에서는 그림 2와 같이 피아노나 가야금과 같은 발현악기의 파형에서 나타나는 소리의 발생 (attack)과 감쇄 (decay) 구간이 명확히 나타나지 않는다. 해금 소리는 배음 분포가 상당히 규칙적이며 20 kHz의 고주파 대역까지 배음이 존재한다. 또한 그림 3에서 확인할 수 있듯이 1600 Hz ~ 1800 Hz 대역에서 크기가 가장 크게 나타나는데 이는 [13]에서 보여준 해금 몸통의 주파수 특성과 일치하며 악기의 몸통이 그 악기의 음색을 결정하는 중요한 요소라는 것을 증명하는 결과이다. 1600 Hz~1800 Hz 대역 이외에도 몇몇 주파수 대역에서 크기가 큰 패턴을 보이는데 이

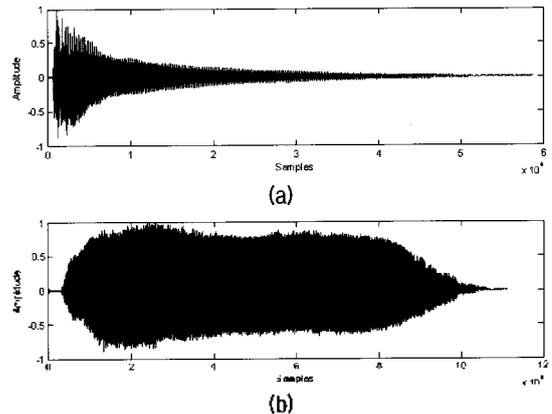


그림 2. 파형 비교 : (a) 가야금의 파형, (b) 해금의 파형
Fig. 2. Comparison of waveform: (a) Gayageum, (b) Haegeum.

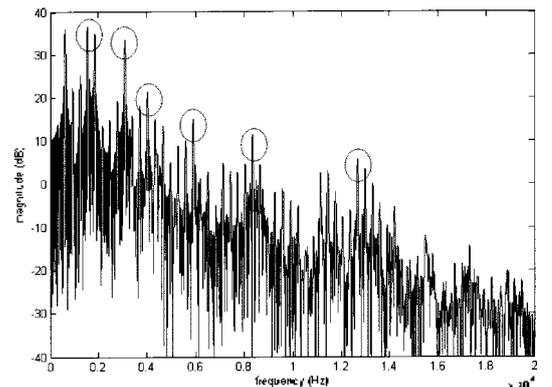


그림 3. 황종운자법 황음에 대한 포먼트
Fig. 3. Formants of note Hwang of Hwangjong position.

러한 포먼트들은 해금 소리의 특성에 영향을 주는 공명주파수가 된다.

4.2. 스펙트럼 파라미터 추출

본 논문에서는 본고 부향실에서 해금 연주 전문가가 해금의 세 가지 운지법에 따라 연주한 단위음을 녹음하여 사용하였다. 복판 바로 앞에 AKG C1000S 마이크를 설치하고 APOGEE Mini-Me A/D 컨버터를 통해 48000 Hz의 샘플링 주파수, 24 bit 양자화로 녹음하였다. 녹음된 각 단위음의 지속구간을 추출하여 이용하였고 분석도구로는 Mathworks 사의 MATLAB 7을 사용하였다. 세 가지 운지법의 단위음은 음높이가 같은 음이 많으므로 황종은 지법에 대한 결과만을 보이기로 한다.

해금은 소리의 발생과 감쇄 구간이 명확하지 않으므로 녹음된 해금 단위음에서 지속구간만을 파라미터 추출에 사용한다. 이러한 입력 샘플에서는 연주자가 활을 밀거나 당기는 속도가 일정하다고 가정할 수 있으므로 시간에 따른 변화를 고려할 필요가 없다. 따라서 STFT (Short-time Fourier Transform) 를 이용하여 시간에 따른 배음과 그 크기의 변화를 관찰하지 않고 입력 신호를 SIFT 분석에서의 한 프레임이라 보고 각 단위음에서 500~1000 샘플 정도를 추출하여 FFT 수행하였다. 본 논문에서 사용한 스펙트럼 파라미터 추출 과정을 그림 4에 나타내었다.

포먼트 합성을 위해 필요한 공명주파수와 그 크기와 대역폭 정보를 얻기 위해서는 스펙트럼 포락선을 구해야 한다. 이 때 캡스트럼 계수를 구하고 적절한 윈도우 크기를 결정하는 것이 중요하다. 그림 5는 황음에 대한 캡스트럼 계수를 보여주는데 기본주파수와 관계되는 두 번째 피크를 기준으로 윈도우 크기를 정하고 윈도우를 씌워 분리한다. 분리된 캡스트럼 계수를 푸리에 변환하면 캡스트럼 포락선을 구할 수 있다.

해금 단위음에서 추출된 캡스트럼 포락선은 많은 포먼트를 포함하고 있기 때문에 적절한 공명점을 찾아야 한다. 해금은 20 kHz 대역까지 배음이 분포하지만 그림 3에서 보듯이 15 kHz 이후에는 크기가 현저하게 작아지는 것을 확인할 수 있다. 그러므로 15 kHz 이내의 공명점을 대상으로 하였고 인접한 공명점 간의 거리가 기본주파수보다 크며 인접한 공명점보다 크기가 더 클 경우에 유효 공명점으로 판단하였다. 각 공명점의 대역폭은 에너지가 공명주파수의 절반이 되는 즉 -3 dB가 되는 지점을 기준으로 결정하였다. 그림 6은 임음에 대한 캡스트럼 포락선과 추출된 공명점을 보여준다.

추출한 공명점 정보를 이용하면 정현파 성분을 합성할 수 있고 원음과의 차로 노이즈 성분을 구할 수 있다. 노이즈 성분을 추출하는 과정을 통해 포먼트 합성이 배음 성분들을 얼마나 잘 묘사하고 있는지 확인할 수 있다. 만약 이전 단계에서 추출하지 못한 배음 성분들이 잔여신호(residual signal)에 포함되어 있다면 유효 배음 성분들을 찾아내어 정현파 성분에 추가한다. 유효 배음을 없앤 후의 잔여신호는 이론적으로 노이즈의 형태를 보이게 되므로 주파수 영역에서 선형모간법에 기초하여 파라미터화

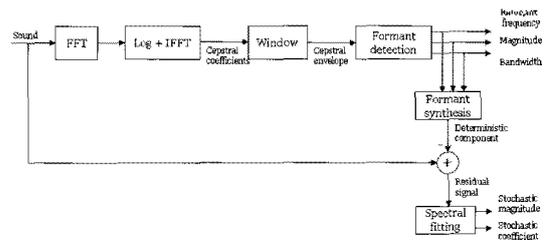


그림 4. 스펙트럼 파라미터 추출 과정
Fig. 4. Extraction of spectral parameters.

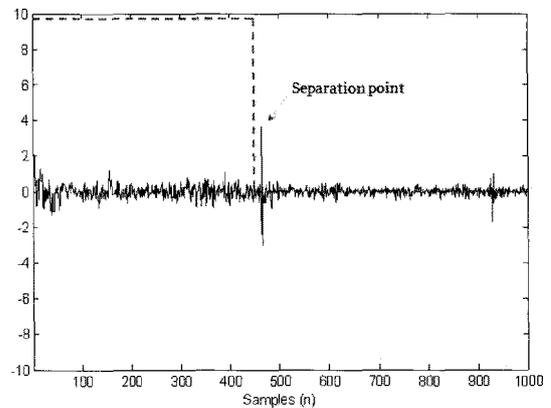


그림 5. 황음에 대한 캡스트럼 (실선)과 포락선 추출을 위한 윈도우 (점선)
Fig. 5. Cepstrum of note Hwang (solid) and window to extract cepstral envelope (dotted).

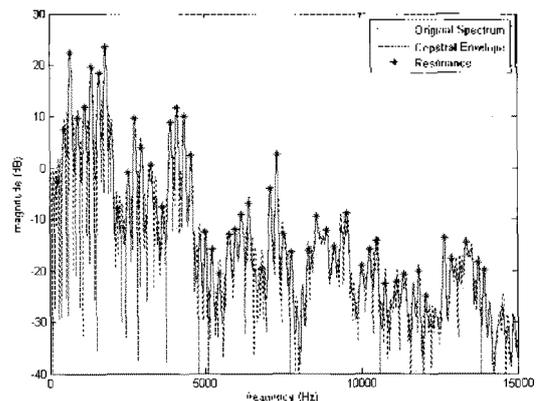


그림 6. 임음에 대한 캡스트럼 포락선과 공명점
Fig. 6. Cepstral envelope and resonances of note Im.

하였다. 본 논문에서는 노이즈 성분을 100샘플 단위로 나누어 각 구간에서 최대값을 구하였다.

V. 해금 단위음 합성

해금 단위음 합성 과정은 그림 7과 같다.

해금 단위음 합성을 위해 본 논문에서 사용한 공명기의 특성은 그림 8와 같으며 캡스트럼 포락선에서 추출한 공명주파수를 f_r 이라 하고 대역폭을 w_{BW} 라고 하면, 입력 신호 $x(n)$ 에 대한 공명기의 출력 $y(n)$ 은 식 (4)로 나타낼 수 있다.

$$y(n) = Ax(n) + By(n-1) + Cy(n-2) \quad (4)$$

여기서,

$$\begin{aligned} C &= -e^{-w_{BW}}, \\ B &= 2e^{-\frac{w_{BW}}{2}} \cos(w_r), \\ A &= 1 - B - C \end{aligned} \quad (5)$$

이고, 공명기의 전달함수는 $z = e^{jw}$ 일 때 식 (6)과 같다 [6].

$$T(z) = \frac{A}{1 - Bz^{-1} - Cz^{-2}} \quad (6)$$

공명기의 출력 신호를 단순히 더하게 되면 크기가 큰 공명점이 인접한 경우 그림 8의 감쇄부분에서 중첩이 일어나 공명점 부근의 감쇄영역을 제대로 묘사하지 못한다. 이를 보완하기 위해 공명기의 출력에 대역 통과 필터를 추가하였다. 이 때 필터의 중심주파수는 공명주파수로 하였고 대역폭은 인접 공명점 간의 주파수 차이로 정의하였다. 예를 들어 첫 번째 공명주파수가 250 Hz이고 두 번째 공명주파수가 550 Hz이면 첫 번째 공명주파수에 대한 대역 통과 필터의 중심주파수는 250 Hz이고 대역폭은 300 Hz이며 차단주파수는 각각 100 Hz와 400 Hz가 된다. 그림 9는 대역 통과된 공명기를 이용한 포먼트 합성 결과이다.

노이즈 성분은 각 구간에서 추출한 파라미터와 선형 보간법으로 스펙트럼을 근사화 하였다. 그림 10에 추출된 노이즈 성분과 선형 보간법을 통해 만든 스펙트럼을 비교하였다.

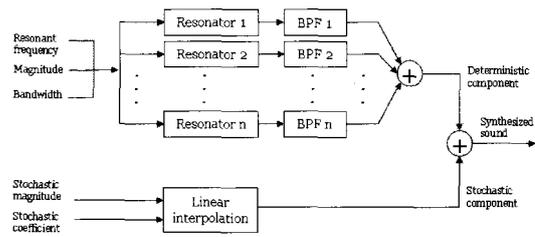


그림 7. 해금 단위음 합성 과정
Fig. 7. Block diagram of synthesis.

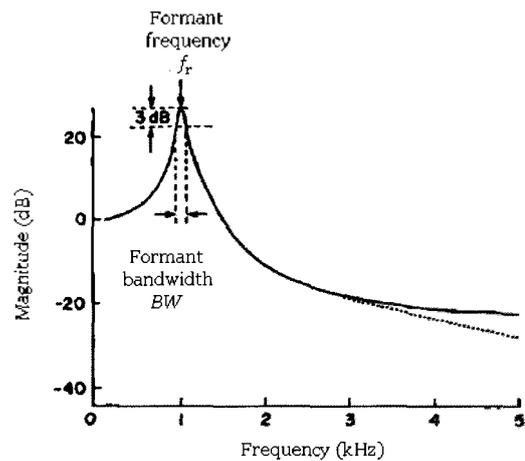
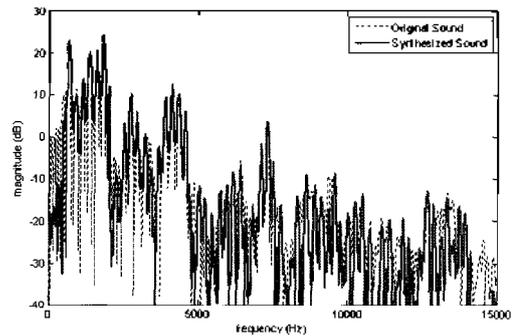
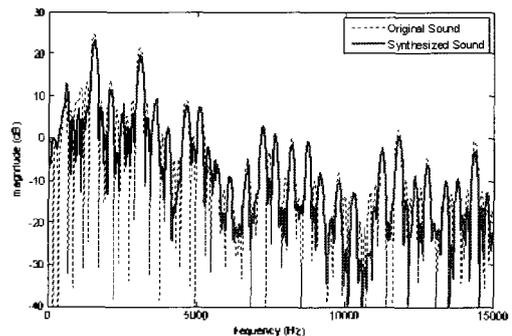


그림 8. 공명기의 주파수 특성 ($f_r=1000$ Hz, $BW=50$ Hz)
Fig. 8. Frequency response of resonator ($f_r=1000$ Hz, $BW=50$ Hz).



(a)



(b)

그림 9. 대역 통과된 공명기를 이용한 포먼트 합성 결과 : (a) 임음, (b) 남음
Fig. 9. Spectra of synthesized sounds by bandpass filtered formant synthesis: (a) note im, (b) note Nam.

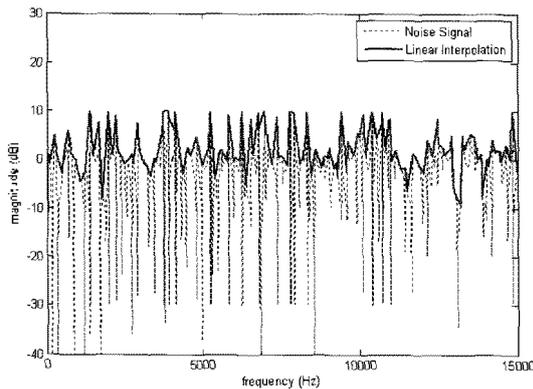


그림 10. 남음에 대한 노이즈 성분(점선)과 선형 보간법을 통한 합성 결과(실선)

Fig. 10. Spectrum of noise signal (dotted) and approximated spectrum of noise signal by linear interpolation (solid).

그림 11은 황종운지법의 단위음에 대한 합성 결과이다. 전체적으로 스펙트럼의 피크를 잘 표현하며 그림 10의 포먼트 합성 결과 보다 원음의 스펙트럼과 더 유사하다. 그림 11에서 (a)와 (b)의 경우 5 kHz 부근과 8 kHz 이후의 대역에서 합성음의 크기가 원음 보다 크고 (c)의 경우 전 대역에 걸쳐 대부분의 합성음 피크들의 크기가 좀 더 크며 (d)와 (e)에서는 몇몇 주요 피크에서 합성음의 크기가 더 크다. (f)는 10 kHz 이후의 대역에서 합성음의 크기가 조금 더 크며 (g)와 (h)에서는 합성음과 원음이 상당히 유사하다. 결과적으로 기본주파수가 비교적 높은 음에서는 대체로 합성음과 원음의 스펙트럼이 유사하고 기본주파수가 비교적 낮은 음에서는 합성음의 크기가 좀 더 커지는 경향을 보이며 특히 10 kHz 이후의 높은 주파수 대역에서 더 두드러진다. 이러한 현상의 원인은 두 가지로 요약될 수 있다. 첫째로 기본주파수가 낮은 음에서는 상대적으로 많은 공명점이 존재하므로 공명점간 거리가 가깝다. 본 논문에서는 인접 공명점과의 거리가 기본주파수 보다 크고 인접 공명점 보다 크기가 더 큰 경우 유효 공명점으로 판단하였기 때문에 고주파 대역에서 누락된 공명점이 더 많이 발생하게 되어 노이즈 성분에서 고주파 대역의 크기가 다소 크다. 이는 공명점 추출에 사용한 컨트를 파라미터를 조절하여 해결할 수 있다. 둘째로는 노이즈 스펙트럼을 선형 보간법으로 근사화하기 때문에 피크 주변의 감쇄 부분의 크기가 커지게 된다. 이러한 현상을 해결하기 위해서는 새로운 노이즈 모델의 적용이 필요할 것이다. 합성음의 데모 파일은 "http://signal.ulsan.ac.kr/duck16/haegeumsound.htm"의 웹페이지에서 확인할 수 있다.

VI. 토 의

스펙트럼은 주파수 영역에서 신호의 에너지 정보를 표현하므로 입력 신호의 샘플 수에 영향을 받는다. 캡스트럼은 스펙트럼 정보를 이용하기 때문에 샘플 양이 많아지면 시간에 따른 변화 보다는 전체적인 에너지의 평균을 나타내게 되며 피크 보다 피크 주변의 에너지 크기가 커지게 된다. 그러므로 해금 소리의 포먼트를 정확히 추출하기 위해서 각 단위음의 10~20 ms 정도의 구간을 추출하여 입력 신호로 사용하였다. 본 논문에서는 해금 소리의 지속 구간을 대상으로 하였지만 해금 연주를 사실적으로 재생하기 위해서는 음의 발생과 감쇄뿐만 아니라 음높이의 변화 또한 묘사해야하며 시간에 따른 변화도 고려해야 한다. 따라서 본 논문에서는 입력 신호를 STFT 분석에서 한 프레임이라 가정하였고 향후 해금 소리의 시간에 따른 변화를 묘사하고 보다 사실적인 음을 합성할 수 있는 STFT 기반의 모델로 개선해야 할 것이다.

스펙트럼 모델링에서 원음에 가까운 소리를 합성하기 위해서는 가능한 모든 배음 성분들을 묘사해야 한다. 정현파 성분 합성 부분에서 기존의 가산합성 방식은 각 배음 성분의 주파수, 위상, 크기 정보를 모두 필요로 하므로 유효 배음의 수가 많을수록 원음에 거의 흡사한 소리를 만들어 낼 수는 있지만 합성시 요구되는 파라미터 수는 비례적으로 많아질 수밖에 없다. 포락선을 이용한 포먼트 합성법은 적은 수의 파라미터로 배음 특성을 묘사할 수 있으며 각각의 피크만을 표현하는 가산합성과는 달리 스펙트럼에서 피크의 변화 곡선을 이용하므로 배음외의 불협화음도 표현할 수 있다는 장점이 있다. 포락선 추출에 FFT를 이용할 경우 가장 세밀하고 정확한 포락선을 구할 수 있지만 상대적으로 많은 데이터가 필요하며 LPC를 이용할 경우 완만한 포락선을 그리게 되어 해금의 배음 특성을 묘사하기 어려우며 차수를 높여 조밀한 포락선을 구하게 되면 계산량이 많아진다. 캡스트럼의 경우 캡스트럼 계수를 구하여 적절한 크기의 윈도우로 분리하면 계산량의 추가 없이 FFT를 이용한 방법보다 적은 데이터로 조밀한 포락선을 구할 수 있으므로 배음 성분이 많은 해금 소리에 적합하다.

원음에서 정현파 성분을 제외한 노이즈 성분에는 유효 피크들이 다소 포함되어있을 수 있는데 이러한 유효 피크들을 추출하여 정현파 성분에 추가하는 과정은 포함하여 정확도를 향상시켰다. 그러나 합성음의 스펙트럼에서 고주파 부분의 크기가 원음에 비해 커지는 문제가 발생하였

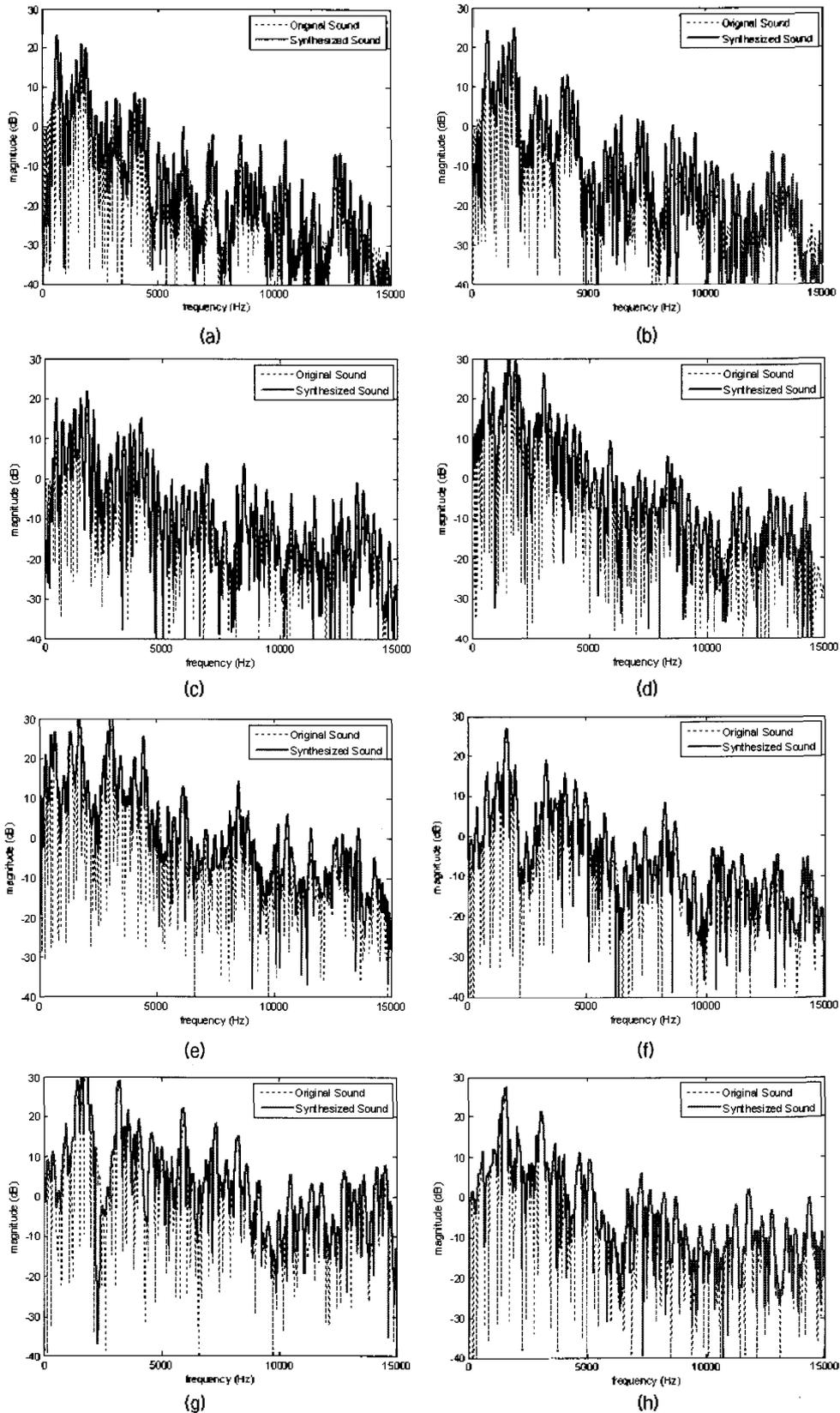


그림 11. 황종운지법 해금 단위를 합성 결과 : (a) 중, (b) 임, (c) 남, (d) 황, (e) 태, (f) 중, (g) 임, (h) 남
 Fig. 11. Spectra of synthesized sounds: (a) note Jung, (b) note Im, (c) note Nam, (d) note Hwang, (e) note Tae, (f) note Jung, (g) note Im, (h) note Nam of Hwangjong position

나. 이는 노이즈 성분 추가 후에 발생하는 현상이므로 노

이즈 모델의 수정이 필요하나 해금은 고주파 배음 성분을

많이 포함하고 있기 때문에 그 부분의 에너지가 노이즈 모델에 의해서 일부 증가한다 하더라도 합성음의 음색에는 거의 영향을 주지 못한다.

VII. 결론

본 논문에서는 전통 악기 해금의 스펙트럼 모델링을 위해 캡스트럼 포락선을 이용한 포먼트 합성을 수행하였다. 기존의 스펙트럼 모델링에서 정현파 성분을 위해 사용된 가산합성법 대신 포먼트 합성법을 적용함으로써 합성 과정에서 요구되는 파라미터의 양을 줄였다. 해금의 특징에 기초하여 캡스트럼 포락선에서 포먼트를 추출하고 정현파 합성에 활용하였다. 노이즈 성분은 원음과 포먼트 합성으로 재생된 정현파 성분과의 차이를 통해 구하였는데 노이즈 성분에 남아있는 유효 피크들을 정현파 성분에 추가함으로써 성능을 향상 시켰다. 정현파 성분의 합성을 위한 공병기로는 IIT에 의한 필터를 활용하였고 노이즈 성분의 합성에는 선형 보간법을 활용하였다. 정현파 성분과 노이즈 성분의 합으로 완성된 합성음은 원음과 매우 유사하였다. 본 연구를 바탕으로 향후 해금 연주의 시간적 변화를 다양하게 묘사할 수 있는 STFT 기반의 합성법에 대한 연구가 진행되어야 할 것이다.

감사의 글

이 논문은 2009년도 정부(교육과학기술부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구인(No. R01-2008-000-20493-0).

참고 문헌

1. Curtis Roads. *The Computer Music Tutorial*, The MIT press, London, 1996.
2. 조상진, 정의필, "개선된 선조 가야금의 물리적 모델링을 이용한 오른손 주법의 음 합성," *한국음향학회지*, 25권, 8호, 362-369쪽, 2006.
3. X. Serra and J. O. Smith, "Residual Minimization in a Musical Signal Model based on a Deterministic plus Stochastic Decomposition," *J. Acoust. Soc. Am.*, vol. 95, no. 5-2, pp. 2958-2959, 1994.
4. X. Serra and J. Smith, "Spectral Modeling Synthesis: A Sound Analysis/Synthesis System based on a Deterministic plus Stochastic Decomposition," *Computer Music Journal*, vol. 14, no. 4, pp. 12-24, 1990.

5. *Spectral Audio Signal Processing*, <http://ccrma.stanford.edu/~jos/sasp/>, Online Book, 2007.
6. Dennis H. Klatt, "Software for a cascade/parallel formant synthesizer," *Journal of Acoustical Society of America*, vol. 67, no. 3, pp. 971-995, 1980.
7. Alan V. Oppenheim, Ronald W. Schaler, *Discrete-Time Signal Processing*, Prentice Hall, 1989.
8. 조상진, 정의필, "선조가야금의 물리적 모델링," *한국음향학회지*, 23권, 7호, 521-531쪽, 2004.
9. 윤지원, 조희영, 조형제, 김준, "대금의 청 울림에 따른 음색변화 연구," *한국디지털아트미디어학회 학술발표대회 논문집*, 3권, 1호, 31-35쪽, 2005.
10. 변중배, 조상진, 정의필, "주파수 영역에서의 상관관수를 이용한 피리의 FM합성 파라미터 추출," *한국디지털아트미디어학회 학술발표대회 논문집*, 4권, 1호, 85-88쪽, 2006.
11. 변중배, 조상진, 정의필, "비대칭 FM합성방식을 이용한 피리 소리의 합성," *한국신호처리시스템학회 하계학술대회 논문집*, 7권, 1호, 37-40쪽, 2006.
12. 변중배, 채진욱, 조상진, 정의필, "샘플링 신디사이저를 이용한 가야금의 농현 사운드 프로그래밍," *한국디지털아트미디어학회 학술발표대회 논문집*, 2권, 1호, 128-131쪽, 2004.
13. 노정욱, 박상하, 성광모, "해금 몸체의 음향학적 특성에 관한 연구," *한국음향학회지*, 26권, 7호, 317-322쪽, 2007.
14. 홍연우, 조상진, 최명환, 정의필, "해금 소리 합성을 위한 스펙트럼 파라미터 추출," *한국디지털아트미디어학회 학술발표대회 논문집*, 5권, 1호, 7-10쪽, 2007.
15. 송혜진, *한국악기*, 열화당, 2001.
16. Ben Gold, Nelson Morgan, *Speech and Audio Signal Processing*, John Wiley & Sons, 1999.

저자 약력

•홍 연 우 (Yeonwoo Hong)



2006년 2월: 울산대학교 컴퓨터정보통신공학부 졸업
2006년 3월 ~ 현재: 울산대학교 컴퓨터정보통신공학과 석사박사통합과정
※주관심분야: 오디오 신호처리

•조 상 진 (Sangjin Cho)

한국음향학회지 제23권 7호 참조

•김 중 면 (Jongmyon Kim)



1995년: 명지대학교 공학사
2000년: (미)University of Florida, 전기컴퓨터공학과 석사
2005년: (미)Georgia Institute of Technology, 전기컴퓨터공학과 박사
2005년 5월~6월: Georgia Institute of Technology, 임시 교수
2005년 7월~2007년 2월: 삼성종합기술원, 전문연구원
2007년 3월 ~ 현재: 울산대학교 컴퓨터정보통신공학부 조교수
※주관심분야: Embedded Systems, Application Specific SoC Design, Computer Architectures, Parallel Processing, Signal Processing

•정 의 필 (Uipil Chong)

한국음향학회지 제23권 7호 참조