

위상 일치와 가변 지수 감쇄 가중치 방법이 적용된 가상 저음 시스템

문현기, *박영철, **윤대희
연세대학교, *연세대학교, **연세대학교
orexis@dsp.yonsei.ac.kr

Phase-matched variable slope exponential weighting virtual bass system

Hyeonggi Moon *Yeong-cheol Park and **Dae Hee Youn
Yonsei University *Yonsei University **Yonsei University

요 약

본 논문은 스피커의 저역 재생 대역을 확장하기 위한 가상 저음 시스템의 성능을 향상시키기 위한 방법을 제안하였다. 제안 방법은 생성하는 신호의 위상을 기존 신호의 위상과 맞추어 강화 과정에서 발생하는 상쇄 간섭의 가능성을 제거한다. 또한 생성한 배음의 상대적 크기를 기본 주파수에 맞추어 가변적으로 적용하여 저역감을 효과적으로 강화하면서 음색의 열화를 줄이는 방법을 제시한다. 제안한 알고리즘은 기존 가상 저음 시스템보다 더 높은 성능을 가지는 것을 확인하였으며, 스피커 환경에서 저역 재생 성능의 향상을 확인하였다.

1. 서론

최근 멀티미디어 장치의 소형화 및 슬림화로 인하여 소형, 평면형 스피커 사용이 늘어나고 있다. 이러한 스피커는 물리적인 한계로 인하여 저역 재생능력이 제한받는다. 위 문제점을 해결하기 위하여 EQ 를 사용할 경우 클리핑이 발생하거나 스피커 드라이버에 손상을 가할 수 있다.

가상 저음 시스템은 스피커가 재생 못하는 저역 성분을 직접 재생하는 대신 심리음향적인 방법인 “missing fundamental [1]” 을 사용하여 저역감을 생성한다. “missing fundamental” 이란 기본 주파수 성분 없이 그 배음만 존재하더라도 이를 기본 주파수로 인지하는 현상이다. 이는 음고의 인식에 음과 음 사이의 간격을 사용하는 청각 특성 때문이다. 따라서 스피커가 재생하지 못하는 저역 성분의 배음을 생성하여 스피커의 저역 재생 대역을 확장한다.

가상 저음 시스템은 배음 생성 방법으로 Non-Linear Device (NLD) [2]나 페이즈 보코더 (Phase vocoder) [3]를 사용한다. NLD 는 시간 영역의 배음 생성 방법으로 연산량이 작고 시간지연이 적지만, 서로 다른 주파수 성분이 입력될 경우 Inter-Modulation Distortion (IMD) 이 발생하는데, 이는 생성한 저역의 음질을 손상시킨다.

페이즈 보코더는 저역 강화 성능과 관련이 높은 배음의 가중치를 정교하게 조작 가능하며, IMD 에서 자유롭다. 하지만 프레임 기반 처리과정으로 인하여 과도 신호의 스미어링 현상이 발생한다. 따라서 과도 성분과 정상상태 성분을 분리하여 과도 성분에는 NLD[1]나 EQ[4]를, 정상상태 성분에는 페이즈 보코더를 사용하는 하이브리드 방법이 제안되었다[1].

생성한 배음의 스펙트럼 포락선은 음색에 중요한 영향을 미친다[5]. 기존 배음 생성 방법은 원 신호 스펙트럼의 포락선 정보를 추정하여 이와 동일한 포락선을 가지는 배음을 생성하거나[6] 배음의 차수에 따라 지수 감소하도록 가중치를 부여하는[5] 방법이 있다. 전자의 경우 다른 악기의 스펙트럼으로 인하여 과도한 중저음 영역의 배음을 생성할 가능성이 있다. 후자의 경우 고정된 계수의 감쇠율을 사용할 경우 높은 기본주파수에 대하여 과도한 중저음 영역의 배음을 생성하게 된다. 과도한 중저음 영역의 배음들은 음색을 날카롭게 만들어 주관적인 성능을 저하시킨다.

또한 가상 저음 시스템이 생성한 배음은 기존 악기의 배음과 더해지게 되는데, 기존 가상 저음 시스템은 이를 고려하지 않아 상쇄간섭이 발생할 수 있다. 이는 저역 강화 효과를 약화시키고 음색에 부정적인 영향을 미치게 된다.

본 논문에서 제안하는 방법은 이상의 문제점을 해결하기 위하여 기본 주파수에 대하여 가변하는 감쇠율을 가지는 배음을 생성하여 저역 강화 효과를 높임과 동시에 음색의 열화를 최소화 하였다. 그리고 기존 배음의 위상을 고려한 배음 생성 방법을 사용하여 상쇄 간섭이 발생하지 않도록 하였다.

본 논문의 구성은 다음과 같다. 2 절에서는 기존의 배음 가중치 부여 방법의 문제점과 이를 개선한 방법을 제안한다. 3 절에서는 위상 정보를 고려한 배음 생성 방법을 제안한다. 4 절에서는 본 논문에서 제안한 기법이 적용된 가상 저음 시스템을 제안하고, 5 절에서는 객관 평가와 주관 평가를 사용하여 제안한 기법의 성능을 확인한다. 마지막 6 절에서는 본 논문에 대한 결론을 맺는다.

2. 배음 가중치 부여 방법

본 절에서는 기존 가중치 부여 방법과 그 한계점 및 개선안을 제안하였다.

2.1. 음색 매칭 방법

음색 매칭 방법은 원 신호의 스펙트럼 포락선과 동일하게 배음의 가중치를 부여하는 방법이다[6]. 먼저 스펙트럼의 추정을 위해 식(1)을 구한다.

$$B_j = \sqrt{\frac{\sum_{k \in j} (|F_j(k)| |X(k)|)^2}{\sum_{k \in j} (|F_j(k)|)^2}} \quad (1)$$

k 는 FFT 빈의 인덱스이며 $X(k)$ 는 원 신호의 스펙트럼, B_j 는 이를 Bark-scale로 그룹화한 것이다. 이를 다시 인접한 프레임에 대하여 평균을 내어 큐빅 보간을 적용한 스펙트럼이 $Y(k)$ 이며, i 번째 배음의 가중치 W_i 는 식 (2)와 같다.

$$W_i^{Timbre} = \frac{Y(i \cdot f_0)}{Y(f_0)} \quad (2)$$

본 방법은 생성한 배음이 기존 신호와 유사한 음색을 가지는 장점이 있지만, 실제 음악 신호에 적용할 경우 음성 혹은 다른 악기로 인하여 스펙트럼 추정에 오차가 발생한다. 그림 1은 음성 신호가 없을 때 저역 악기(베이스 기타)의 스펙트럼 추정 결과와 200Hz, 280Hz 대역에 음성 신호가 함께 존재할 때의 스펙트럼 추정 결과이다. 잘못 추정된 스펙트럼으로 생성한 배음은 음색의 열화를 가져온다.

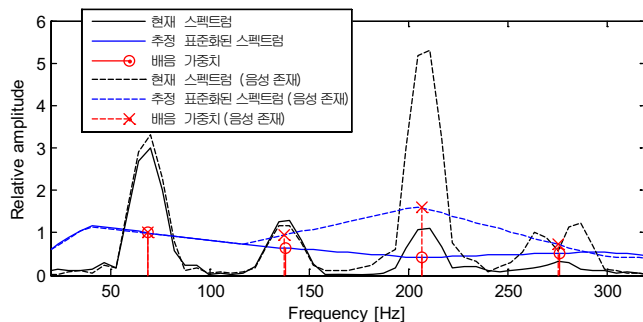


그림 1. 음성 신호로 인한 스펙트럼 추정 오차 (추정된 스펙트럼은 10개 프레임 평균)

2.2. 지수 감쇄 가중치 방법

지수 감쇄 가중치 방법은 i 번째 배음에 아래 식 (3)과 같이 지수 감소하는 가중치를 부여한다. α 는 감쇄 계수이다.

$$W_i^{Exp} = 10^{-\frac{\alpha}{20}i} \quad (3)$$

위 방법은 배음 생성에 기본 주파수 성분만을 사용하기 때문에 음색 매칭 방법에서 발생하는 스펙트럼 추정에서의 문제점이 발생하지 않지만, 감쇄 계수에 따라서 저역 강화의

성능과 음색의 변화 간 trade-off 관계가 발생한다. 낮은 감쇄 계수를 사용할 경우 높은 저역 강화 성능을 가지지만 높은 기본 주파수에서 과도한 중저음 영역의 배음을 생성하여 음색을 날카롭게 만든다. 높은 감쇄 계수를 사용할 경우 저역 강화 효과가 줄어들게 된다.

2.3. 가변 지수 감쇄 가중치 방법

제안 방법은 기본 주파수에 따라 감쇄 계수를 가변적으로 적용하였다. 식 (5)의 감쇄 계수 $g(f)$ 는 기본 주파수 f 에 종속되어 있으며 중저역 대역에 과도한 배음의 생성을 막기 위하여 특정 주파수 f_{max} 에 대하여 배음의 이득을 $W_{f_{max}}$ 로 제한하였다. i 번째 배음에 대한 가중치는 식 (4)와 같다.

$$W_i^{Var-Exp} = 10^{-\frac{g(\bar{f}_0 \cdot i)}{20}} \quad (4)$$

$$g(f) = -20 \log_{10} \left(w_{f_{max}} \right) \frac{f}{f_{max} - f} \quad (5)$$

제안 방법은 낮은 기본 주파수일 경우 낮은 감쇄 계수로 저역 강화 성능을 높이고 높은 기본 주파수일 경우 높은 감쇄 계수로 과도한 중저음 대역의 배음 생성을 막는다. 그림 1은 $f_{max} = 300\text{Hz}$, $W_{f_{max}} = -30\text{dB}$ 일 때 기본 주파수에 대한 배음의 이득이다.

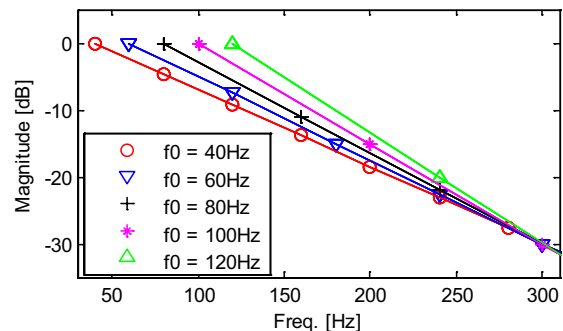


그림 2. 기본 주파수에 따른 제안 배음 가중치 부여 방법

3. 위상 매칭

현악기의 정상상태 성분의 시간 파형 $x(n)$ 은 식 (6), (7)과 같이 모델링 된다[7]. f_0 는 기본 주파수이며 A_i , φ_i 는 i 번째 배음의 진폭과 위상이다. 연주자의 연주 방식, 악기 특성으로 인하여 배음의 차수마다 $\varphi_i(n)$ 가 다른 값을 가지기 때문에[7] 기본 주파수와 배음 성분은 non-coherent 하다.

$$x(n) = \sum_{i=0}^L A_i(n) \cos(\theta_i(n)) \quad (6)$$

$$\theta_i(n) = i \cdot f_0 \cdot n + \varphi_i(n) \quad (7)$$

대표적인 저역 악기인 일렉 및 어쿠스틱 베이스 기타, 첼로 등 에서도 동일한 현상이 발생한다. 기존 가상 저음 시스템에서

사용된 페이즈 보코더 기반의 배음 생성 방법들은 모두 기본 주파수와 coherent 한 배음들을 생성하기 때문에 원 신호와 상쇄 간섭을 일으켜 저역 강화 성능을 감소시킨다.

제안 방법은 “Peak shift method[8]” 를 사용하여 배음을 생성하였다. 추정된 기본 주파수에 해당하는 주파수 빈과 인접한 주파수 빈의 진폭 스펙트럼을 배음에 해당하는 주파수 빈 위치로 옮기고, 식 (4)의 가중치를 부여한다. 생성한 배음의 위상은 기존 신호의 위상과 일치시키기 위하여 기존 신호의 위상 스펙트럼을 그대로 사용하였다. 기본 주파수의 추정은 주파수 영역에서 스펙트럼의 피크를 검출하는 방법을 사용하였다.

그림 3은 원신호와 위상 매칭이 적용된 저역강화신호 그리고 위상 매칭이 비 적용된 저역강화신호 스펙트럼이다. 기본 주파수 성분은 두 저역강화방법 모두 위상이 동일하다. 위상 매칭이 없는 경우 상쇄 간섭이 일어남을 확인할 수 있다.

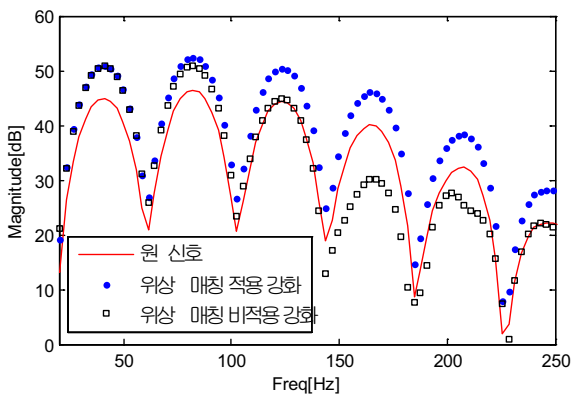


그림 3. 위상 매칭 적용 여부에 따른 저역 강화 효과

4. 제안 저역 강화 시스템

제안 저역 강화 시스템은 크로스오버 필터 H_{lpf} , H_{hpf} 를 사용하여 저역 성분을 추출한다. 연산량을 줄이기 위하여 다운샘플링을 한 다음, Short-time Fourier Transform (STFT)을 적용한다. 정상상태 성분과 과도 성분의 분리를 위하여 Harmonic Percussive Sound Separation (HPSS) [9]을 적용한다. 정상상태 성분 $X_n(k)$ 는 기본 주파수를 추정하여 배음을 생성하고 식 (4)의 가중치 부여 방법을 적용하였으며, 과도 성분 $X_p(k)$ 는 EQ를 적용하였다. 제안 시스템의 블록 다이어그램은 그림 4와 같다.

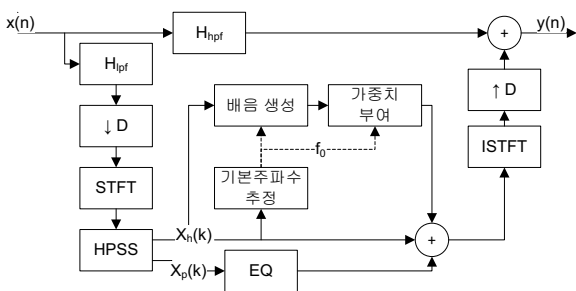


그림 4. 제안 저역 강화 시스템 블록 다이어그램

5. 객관 및 주관 평가

5.1. 배음 가중치 부여 방법 객관 평가

음색 매칭 방법과 지수 감쇄 가중치 방법 그리고 제안한 가변 지수 감쇄 가중치 방법의 성능 비교를 수행하였다. 원 신호를 레퍼런스 스피커 출력으로 가정하였다. 소형 스피커를 모델링 하기 위하여 120Hz의 저지 대역을 가지고 기울기가 12dB/Oct. 인 고역 통과 필터를 사용하였으며, 과도 성분은 강화하지 않고 정상상태 신호만 강화하였으며, 위 세가지 가중치 부여 방법을 비교하였다.

생성한 배음의 개수는 4개이다. 음색 매칭 방법은 10개의 인접한 프레임의 평균을 스펙트럼을 추정하였으며, 지수 감쇄 가중치 부여 방법의 감쇄 계수는 6dB이다. 제안한 가중치 부여 방법은 $f_{max} = 300\text{Hz}$, $W_{fmax} = -30\text{dB}$ 와 $f_{max} = 400\text{Hz}$, $W_{fmax} = -40\text{dB}$ 를 사용하였다.

소형 스피커로 재생되는 저역 강화 신호와 레퍼런스 신호의 주관적인 차이를 예측하기 위하여 Perceptual Evaluation of Audio Quality (PEAQ)의 Objective Difference Grade (ODG) [10]를 측정하였다. 테스트 신호는 총 31개의 베이스 신호로 41~110Hz의 기본 주파수를 가지며, 일렉 베이스, 어쿠스틱 베이스 기타, 키보드 베이스를 포함한다. 각 가중치 부여 방법에 따른 저역 강화 신호와 레퍼런스 신호간의 ODG 값은 그림 5와 같다.

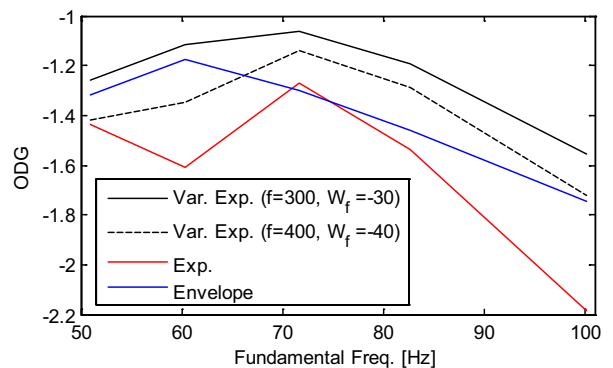


그림 5. 배음 가중치 부여 방법에 따른 ODG

지수 감쇄 가중치 방법(적색)은 음색 매칭 방법(청색)보다 낮은 성능을 보이며 [5]의 결과와 일치한다. 점선의 가변 지수 감쇄 가중치 방법은 낮은 기본 주파수(약 50Hz)에서 지수 감쇄 가중치 방법과 동일한 감쇄율을 가지도록 하였다. 제안 방법은 낮은 주파수에서는 지수 감쇄 가중치 방법과 ODG 값이 동일하며, 높은 주파수에서는 높은 감쇄율로 과도한 중저음 영역의 배음 생성을 줄여 높은 ODG 값을 가지게 한다. 제안 방법에 적합한 파라미터를 선정한다면(흑색 실선) 음색 매칭 방법보다 주관적인 성능을 더 높일 수 있다.

5.2. 제안 저역 강화 시스템 성능 주관 평가

제안 방법의 효과를 확인하기 위하여 약 100Hz의 저역 재생 한계 주파수를 가지는 소형 스피커에 제안한 가상 저음 시스템을 적용한 다음 주관적인 저역의 성능을 비교하였다.

실험 방법은 ITU-T BS.1284[11]의 5 점 척도 비교를

사용하였으며, 레퍼런스 신호는 49Hz 의 저역 재생 재생 한계 주파수를 가지는 모니터 스피커에서 재생된 신호이다. 비교 항목은 저역의 양과 전반적인 소리의 품질이며 실험음원으로 30 초 내외의 팝 음원 2 개와 락 음원이 사용되었다. 각 음원은 청취지점에서 평균 음압이 $70 \pm 0.5\text{dB}$ 이 되도록 보정하였다.

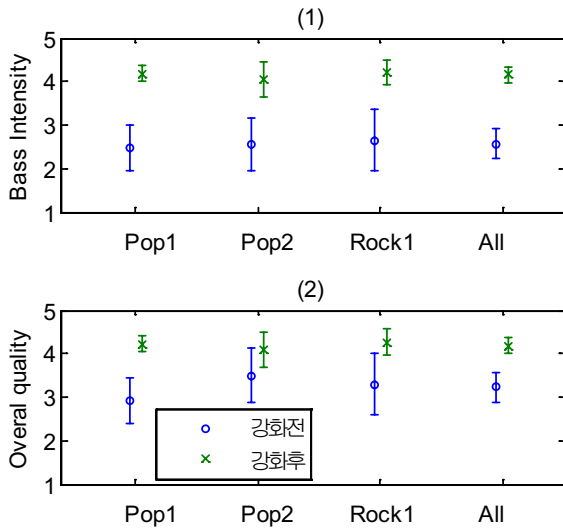


그림 6. 저역 강화 시스템 주관 성능 평가 결과
(1) 저역의 양, (2) 소리의 전반적인 품질

레퍼런스 시스템을 5 점 만점 기준으로, 제안한 방법의 적용 전과 후를 비교하면 저역의 양은 1.6 점, 전반적인 소리의 품질은 1 점 향상되었다.

6. 결론

본 논문으로 정확한 배음의 가중치를 부여하기 위하여 위상 매칭이 필수적임을 확인하였으며, 가변 가중치 배음 생성 방법이 기존 배음 생성 방법보다 높은 성능을 가지고 있음을 확인하였으며, 실제 소형 스피커를 사용하여 제안 알고리즘의 저역 강화 효과를 확인하였다.

7. 참고문헌

[1] A. J. Hill and M. O. J. Hawksford, "A Hybrid Virtual Bass System for Optimized Steady-State and Transient Performance," in Proc. 2nd Computer Sci. and Electronic Eng. Conf. (CEEC), Colchester, UK, pp. 1, 6, 8-9., 2010.

[2] E. Larsen and R. M. Aarts, "Perceiving Low Pitch Through Small Loudspeakers," presented at the AES 108th Convention, Paris, France, 2000.

[3] M. R. Bai and W. Lin, "Synthesis and Implementation of Virtual Bass System with a

Phase-Vocoder Approach," J. Audio Eng. Soc., vol. 54, pp. 1077-1091, 2006.

[4] T. Lee, S. Lee, Y. C. Park and D. H. Youn, "Virtual Bass Systems Based on a Multiband Harmonic Generation," IEEE International Conference on Consumer Electronics (ICCE), 2013

[5] H. Mu, W. S. Gan, and E. L. Tan, "An Objective Analysis Method for Perceptual Quality of a Virtual Bass System," Audio Speech Lang. Process. IEEEACM Trans. On, vol. 23, no. 5, pp. 840-850, 2015

[6] H. Mu, W. S. Gan, and E. L. Tan, "A Timbre Matching Approach to Enhance Audio Quality of Psychoacoustic Bass Enhancement System," in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), pp. 36-40, 2013.

[7] S. Dubnov and X. Rodet, "Investigation of Phase Coupling Phenomena in Sustained Portion of Musical Instrument Sound," J. Acoust. Soc. Am., Vol. 113, pp. 348-359, 2002.

[8] J. Laroche and M. Dolson, "New Phase-Vocoder Techniques for Pitch-Shifting, Harmonizing and Other Exotic Effects," Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 1999.

[9] D. Fitzgerald, "Harmonic/Percussive Separation Using Median Filtering," in 13th International Conference on Digital Audio Effects (DAFX10), 2010.

[10] "Method for objective measurements of perceived audio quality," ITU Recommendation BS.1387-1, 2001.

[11] "General methods for the subjective assessment of sound quality," ITU Recommendation BS.1284-1, 2014.