

특집논문 (Special Paper)

방송공학회논문지 제21권 제6호, 2016년 11월 (JBE Vol. 21, No. 6, November 2016)

<http://dx.doi.org/10.5909/JBE.2016.21.6.889>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

위상 일치와 가변 지수 감쇠 가중치 부여 방법이 적용된 가상 저음 시스템

문현기^{a)}, 박영철^{b)}, 황영수^{c)}†

Phase-matched Harmonic Generation and Variable Slope Exponential Weighting for Virtual Bass System

Hyeongi Moon^{a)}, Young-cheol Park^{b)}, and Young-soo Whang^{c)}†

요약

가상 저음 시스템은 기본 주파수 성분의 배음을 생성하여 스피커의 저역 재생 대역을 확장하는 방법으로 소형 스피커에 널리 사용된다. 가상 저음 시스템의 주관적인 성능은 배음의 가중치 부여 방법과 관련이 높기 때문에, 기존 연구에서는 지수 감쇠 가중치 부여 방법과 음색 매칭 방법 등 다양한 가중치 부여 방법이 제안되었다. 그러나 생성한 배음과 기존 신호간의 위상을 맞추지 않을 경우 정확한 가중치 부여가 불가능하다. 본 논문에서는 기존 가중치 부여 방법의 한계점을 분석하고 이를 개선한 가중치 부여 방법을 제안하였다. 제안한 방법은 생성한 배음의 위상을 기존신호의 위상과 일치시키고, 기본 주파수에 따라 배음 가중치를 가변적으로 부여하는 방법이다. 기존 가상 저음 시스템과 객관 및 주관 비교 평가를 수행한 결과, 위상 일치 방법은 자연스럽고 효과적인 저역강화에 필수적임을 확인하였으며, 제안한 배음 가중치 부여 방법은 제한된 상황에서 기존 가중치 부여 방법보다 효과적임을 확인하였다.

Abstract

Virtual Bass System (VBS) is widely used to extend the lower frequency limit of small loudspeakers, which generates harmonics of a fundamental frequency. The perceptual quality of the VBS is highly dependent on the harmonic weighting strategy. There have been several weighting methods, including exponential attenuation and timbre matching. However, it is essential to match phases between harmonics in the original signal and generate harmonics to precisely convey the weighting strategy. This paper shows the limitations of the previous harmonic weighting schemes and proposes a new harmonic weighting scheme. The proposed weighting scheme proposes phase matching between the original and generated harmonics and varies the slope of the attenuation weighting dynamically according to the missing fundamental frequency. Objective and subjective tests show that the proposed harmonic weighting scheme provides more natural and effective bass perception in a limited situation than the conventional schemes, which implies that the phase matching is essential for the high quality bass enhancement.

Keyword : Virtual bass system, Bass enhancement, Psycho-acoustic signal processing, Audio quality evaluation, Phase vocoder

Copyright © 2016 Korean Institute of Broadcast and Media Engineers. All rights reserved.

"This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered."

I. 서 론

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 EQ를 사용할 경우 클리핑이나 스피커 드라이버의 손상이 발생할 수 있다.

가상 저음 시스템(Virtual Bass System ; VBS)은 스피커가 재생 못하는 저역 성분을 직접 재생하는 대신 "Missing fundamental^[1]"을 사용하여 저역감을 생성한다. "Missing fundamental" 이란 기본 주파수 성분 없이 그 배음들만이 존재하더라도 이를 기본 주파수로 인지하는 심리음향적인 현상이다. 이는 음고의 인식에 음과 음 사이의 간격을 사용하는 청각 특성 때문에 발생한다. VBS는 스피커가 재생하지 못하는 저역 성분의 배음을 생성하여 스피커의 저역 재생 대역을 확장한다.

VBS는 배음 생성 방법으로 Non-Linear Device^[2] (NLD)나 페이즈 보코더^[3] (Phase vocoder)를 사용한다. NLD는 시간 영역의 배음 생성 방법으로 연산량이 작고 시간지연이 적지만, 서로 다른 주파수 성분이 입력될 경우 Inter-Modulation Distortion (IMD)이 발생하며, 이는 생성한 저역의 음질을 손상시킨다. 페이즈 보코더는 저역 강화 성능과 관련이 높은 배음의 가중치를 정교하게 조작 가능하며, IMD에서 자유롭지만 프레임 기반 처리과정으로 인하여 과도상태 신호의 스미어링(Smeearing) 현상이 발생한다. 따라서 과도상태 신호와 정상상태 신호를 분리하여 과도상태 신호에는 NLD^[1]나 EQ^[4]를 사용하고 정상상태 신호에는 페이즈 보코더를 사용하는 하이브리드 방법이

제안되었다^[1].

한편 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 방법^[5]이 있다. 그러나 전자의 경우 다른 악기의 스펙트럼으로 인하여 저음이 아닌 중음 대역에 배음을 생성할 가능성이 있으며, 후자의 경우 고정된 감쇠 계수를 사용할 경우 높은 기본주파수에 대하여 과도한 중음 영역의 배음을 생성하게 되어 음색을 날카롭게 만들어 주관적인 성능을 저하시킨다.

과거 제안되었던 여러 가지 VBS 알고리즘들이 가지는 공통적인 단점은 음질이 일관적이지 않다는 것이다. 이는 기존의 VBS가 생성한 배음과 기존의 배음 간의 위상 차이를 전혀 고려하지 않고 있어서 발생하는 문제이다. 이러한 위상 차이로 인한 상쇄간섭은 VBS 자체의 성능을 떨어뜨릴 뿐만 아니라 음질을 열화 시킬 수 있다.

따라서 본 논문에서는 이상의 문제점을 해결하고 원하는 배음 가중치 부여방법을 적용할 수 있도록 기존의 페이즈 보코더를 수정하여 생성한 배음에 기존 배음의 위상을 사용하도록 하였다. 그리고 넓은 기본 주파수대역에서 효과적인 저역 강화를 위하여 기본 주파수에 따라 생성한 배음의 감쇠 계수를 가변적으로 부여하는 가변 지수 감쇠 가중치 부여 방법을 제안한다.

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

II. 배음 가중치 부여 방법

배음의 가중치 부여 방법은 저역의 주관적인 성능에 중요한 요소이며, 지수 감쇠 가중치 부여 방법^[5]과 음색 매칭^[6], 등청감곡선을 사용하는 방법^[3]이 대표적이다. 이를 중

a) 연세대학교 전기전자공학부(School of Electrical and Electronic Engineering, Yonsei University)

b) 연세대학교 컴퓨터정보통신공학부(Computer and Telecommunications Engineering Division, Yonsei University)

c) 가톨릭관동대학교 전자공학과(Departmnent of Electronic and Communication Engineering, Kwandong University)

‡ Corresponding Author : 황영수(Young-soo Whang)

E-mail: hysoo@cku.ac.kr

Tel: +82-033-649-7554

ORCID: <http://orcid.org/0000-0002-7188-3678>

※ 이 논문의 연구결과 중 일부는 “2016년 한국방송·미디어공학회 학계학술대회”에서 발표한 바 있음.

· Manuscript received September 12, 2016; Revised November 17, 2016; Accepted November 17, 2016.

지수 감쇠 가중치 부여 방법과 음색 매칭 방법이 상대적으로 높은 주관 성능을 가진다^{[5][6]}. 2절에서는 두 가중치 부여 방법의 한계점을 분석하고 이를 개선한 가변 지수 감쇠 가중치 부여 방법을 제안하였다.

1. 음색 매칭 방법

음색 매칭 방법은 배음의 가중치를 원 신호의 스펙트럼 포락선과 동일하게 부여하는 방법이다^[6]. 스펙트럼의 추정을 위해 먼저 원 신호의 스펙트럼 $X(k)$ 를 Bark-scale로 그룹화한 B_j 를 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 주파수 빈, j 는 Bark-scale의 인덱스이다. $F_j(k)$ 는 j -번째 Bark-scale critical band에 대한 Triangular filter이다^[6]. 식 (2)의 $Y(f)$ 는 B_j 를 인접한 프레임에 대하여 평균을 낸 다음, 주파수 f 에 대하여 큐빅 보간을 적용한 스펙트럼이다. 기본 주파수 f_0 성분에 대하여 i -번째 배음의 가중치 W_i 는 식 (2)와 같다.

$$W_i^{\text{Timbre}} = \frac{Y(f_0i)}{Y(f_0)} \quad (2)$$

본 방법은 음성 혹은 다른 악기로 인하여 저역 악기의 스펙트럼 추정에 오차가 발생한다. 그림 1의 좌측 그림은

음성 신호가 없을 때 저역 악기(베이스 기타)의 스펙트럼 추정 결과와 200Hz, 280Hz 대역에 음성 신호가 함께 존재할 때의 스펙트럼 추정 결과이다.

그림 1의 베이스 기타 스펙트럼에 음색 매칭 방법으로 배음을 생성할 경우, 130Hz 대역에 존재하는 첫 번째 배음의 가중치가 가장 커야 한다. 하지만 음성 신호가 존재하면 그림 1의 우측 그림과 같이 200 Hz와 280 Hz에 더욱 큰 가중치를 부여하여 원래 강화해야 할 대역보다 높은 대역을 강화하게 된다. Harmonic Spectral Centroid (HSC)는 배음의 가중 평균으로 음색의 날카로움(Sharpness)과 관련이 높으며^[7], 이는 VBS의 주요한 성능 열화 요인이다. 여러 악기가 섞인 음원에서 음색 매칭 방법을 적용할 경우 HSC를 높이게 되어 의도한 저역강화 효과를 이루지 못하고 음색의 열화가 발생하거나 높은 주파수 대역의 다른 악기를 강화할 수 있다.

2. 지수 감쇠 가중치 부여 방법

지수 감쇠 가중치 부여 방법은 i -번째 배음에 식 (3)과 같이 지수 감소하는 가중치 W_i^{Exp} 를 부여한다. α 는 감쇠 계수이다.

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

위 방법은 배음 생성에 기본 주파수 성분만을 사용하기 때문에 음색 매칭 방법에서 발생하는 스펙트럼 추정에서의 문제점이 발생하지 않지만, 감쇠 계수에 따라서 저역 강화의 성능과 음색의 열화 간 trade-off 관계가 발생한다. 항상 일정한 감쇠 계수를 사용하기 때문에 낮은 기본 주파수에

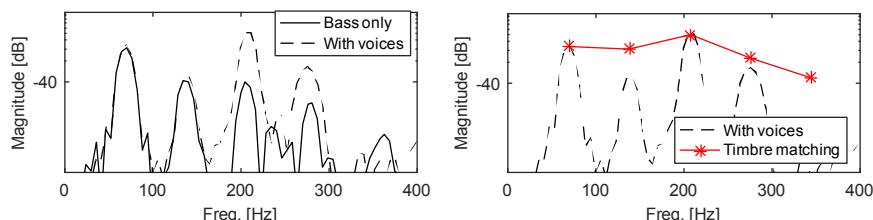


그림 1. 음성신호 유무에 따른 스펙트럼(좌), 음성신호가 있을 때 음색 매칭 방법의 생성 배음(우)

Fig. 1. Magnitude spectrum of bass guitar with and without voices (left), Magnitude spectrum of the bass guitar with voices and harmonics generated by timbre matching (right)

적합한 낮은 감쇠 계수를 사용할 경우, 높은 기본 주파수에서 과도한 중음 영역의 배음을 생성하여 자연스럽지 못한 저역 강화 효과를 제공한다.

3. 가변 지수 감쇠 가중치 부여 방법

제안 방법은 저역 신호의 기본 주파수 f_0 에 따라 가변적인 감쇠 계수를 적용한다. 이는 낮은 기본 주파수일 경우 낮은 감쇠 계수를 사용하여 저역 강화 성능을 높이고, 높은 기본 주파수일 경우 높은 감쇠 계수를 사용하여 음색을 날카롭게 만드는 중음 대역의 배음 생성을 막는다.

식 (4)는 제안 방법의 감쇠 계수이다. 기본 주파수 성분의 크기가 0 dB일 때, 모든 f_0 에 대하여 생성한 배음 성분의 포락선은 f_w Hz 일 때 W_w dB의 크기를 가진다. 적절한 f_w 와 W_w 를 사용하여 원하는 저역 대역으로 제한된 배음을 생성할 수 있다. i -번째 배음에 대한 가변 지수 감쇠 가중치 W_i^{VarExp} 는 식 (5)와 같다.

$$g(f_0) = -20 \log_{10}(W_w) \frac{f_0}{f_w - f_0} \quad (4)$$

$$W_i^{VarExp} = 10^{-\frac{g(f_0)}{20} i} \quad (5)$$

그림 2는 $f_w = 300$ Hz, $W_w = -70$ dB 일 때 기본 주파수에 대한 감쇠 계수 $g(f_0)$ 와 가중치 W_i 를 나타낸 그림이다. 생성한 배음의 스펙트럼 포락선이 f_w Hz 일 때 W_w dB의 크기를 가지는 것을 확인할 수 있다.

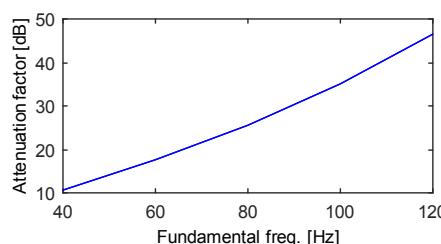


그림 2. 기본 주파수에 대한 감쇠 계수(좌), 기본 주파수에 대한 생성 배음의 가중치 (우)

Fig. 2. Attenuation factor according to fundamental frequency (left), Weighting ratio according to fundamental frequency (right)

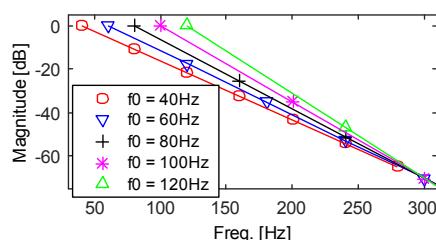
III. 위상 일치 방법

페이즈 보코더 기반의 배음 생성방법은 크게 두 가지로 나뉜다. 첫 번째 방법은 푸리에 도메인 혹은 필터뱅크 (Filter-bank) 도메인에 적용되는 방법이다. 기본 주파수 성분의 순시 주파수(Instantaneous frequency)를 구하고, 여기에 생성할 배음의 차수를 곱하여 배음의 순시 주파수를 구한다. 각 배음에 대하여 이를 적분하여 초기 위상에 더한 값이 각 배음의 위상이다. 초기 위상은 일반적으로 0이 사용된다.

두 번째 방법은 "Peak shift method^[8]"로 푸리에 도메인에 적용되는 방법이다. 기본 주파수에 해당하는 크기 스펙트럼의 주엽을 배음의 위치에 복사하여 배음의 크기 스펙트럼을 생성한다. 배음의 순시 주파수는 첫 번째 방법과 같이 구한다. 각 배음의 위상은 각 배음의 순시 주파수를 적분하여 기본 주파수의 위상과 더한 값이다.

페이즈 보코더 기반의 VBS는 위 방법 중 하나를 사용하여 일렉트릭 베이스와 어쿠스틱 베이스 기타, 첼로 등의 저역을 담당하는 악기의 기본 주파수 성분의 배음을 생성한다. 앞서 언급한 악기들은 현악기로, 연주자의 연주 방식과 악기 특성으로 인하여 정상상태 성분의 기본주파수와 배음 성분이 서로 결맞지 않는(Non-coherent) 현상이 발생한다^[9]. 기존의 배음 생성 방법은 모두 기본 주파수와 결맞은 (Coherent) 배음들을 생성하기 때문에, 현악기에 기존 배음 생성 방법을 적용할 경우 원 신호와 상쇄간섭이 발생하여 저역 강화 성능이 감소되거나 음질의 열화가 발생할 수 있다.

그림 3은 베이스 기타에 제안한 위상 일치 방법을 적용한 저역 강화 결과와 위상 일치 방법을 비적용한 저역 강화



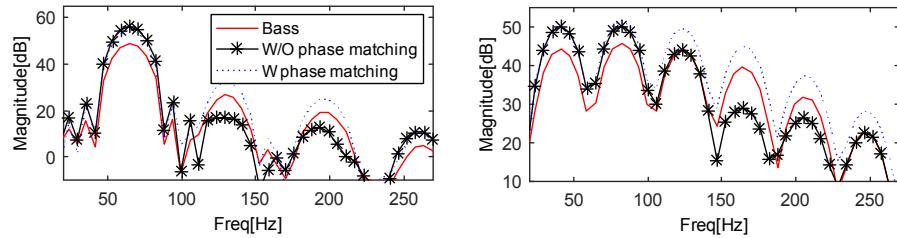


그림 3. 베이스 기타의 크기 스펙트럼과 저역 강화 후 크기 스펙트럼
 Fig. 3. Magnitude spectra of the bass guitar and the reinforced harmonics

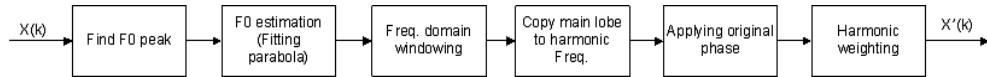


그림 4. 위상 일치 방법이 적용된 배음 생성 방법 블록도
 Fig. 4. Block diagram of the phase-matched harmonic generation

결과의 스펙트럼이다. 오른쪽과 왼쪽은 각각 서로 다른 베이스 기타의 스펙트럼과 그 강화 결과를 나타낸다. 적색 선이 베이스 기타의 스펙트럼이며 위상 일치 방법을 적용한 저역 강화 결과가 점선, 적용하지 않은 저역 강화 결과가 흑색 선이다. 위상 일치 방법을 적용하지 않을 경우 상쇄간섭이 발생하여 강화된 신호가 원 신호보다 작아지는 경우가 발생한다. 오른쪽 그림의 경우 첫 번째 배음과 두 번째 배음에서, 왼쪽 그림의 경우 세 번째 이상의 배음에서 상쇄간섭이 발생한다.

제안한 위상 일치 방법을 적용한 배음 신호 생성 방법은 "Peak shift method"^[8]를 기반으로 한다. 기본 주파수의 크기 스펙트럼을 복사하여 배음의 크기 스펙트럼으로 사용한 것은 기존 방법과 동일하지만, 악기의 스펙트럼이 배음 구조를 가지는 점을 이용하여 배음의 위치에 존재하는 위상 스펙트럼을 합성하지 않고 기존 배음의 위상 스펙트럼을 그대로 사용한 점이 차이점이다.

제안한 배음 신호 생성 방법의 상세 블록도는 그림 4와 같다. 기본주파수의 추정을 위하여 크기 스펙트럼 $X(k)$ 의 극대 값을 검출한다. 검출된 극대 값과 인접한 주파수 빈에 포물선 근사법을 적용하여 정확한 기본 주파수를 추정한다. 제안한 방법을 사용하여 배음의 크기 및 위상 스펙트럼을 생성한 다음, 기본 주파수에 따라 식 (4)의 가중치를 부여하여 최종적으로 원 신호와 위상이 일치된 가변 지수 감쇠 가중치가 적용된 배음 신호 $X'(k)$ 를 생성한다.

IV. 제안하는 저역 강화 시스템

VBS는 일반적으로 스피커의 재생 한계 대역을 최대한 옥타브 반 아래로 확장한다^[10]. 소형 스피커의 재생 한계 대역이 f_L 일 때, 제안 저역 강화 시스템은 $2^{-1.5}f_L \sim f_L$ Hz의 기본 주파수 성분에 대하여 배음을 생성하여 저역을 강화한다. 제안 시스템의 블록도는 그림 5와 같다.

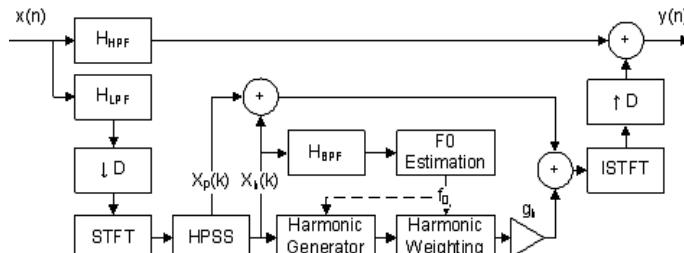


그림 5. 가상 저음 시스템 테스트 플랫폼
 Fig. 5. Block diagram of the test platform of the VBS

입력 신호 $x(n)$ 에 크로스오버 필터 H_{LPF} , H_{HPF} 를 적용하여 생성할 배음을 주파수대역을 포함한 저역 신호를 추출한다. 연산량을 줄이기 위하여 다운샘플링을 수행한 다음 단구간 푸리에 변환(Short-time Fourier Transform ; STFT)을 수행한다. VBS는 정상상태 신호 $X_h(k)$ 의 저역 강화에 효과적이며 과도상태 신호 $X_p(k)$ 에는 대역 제한된 이득을 적용하는 것이 효과적이다^[4]. 따라서 Harmonic Percussive Sound Separation^[11] (HPSS)을 이용하여 $X_h(k)$ 와 $X_p(k)$ 를 분리한 다음 $X_h(k)$ 에만 VBS를 적용한다.

$X_h(k)$ 에 통과 대역이 $2^{-1.5}f_L \sim f_L$ Hz인 대역 통과 필터 H_{BPF} 를 적용하여 배음을 생성할 저역 신호를 추출하고 f_0 를 추정한다. 기본 주파수 성분에 대하여 위상 일치 방법을 사용한 배음을 생성한 뒤 가변 지수 감쇠 가중치 부여 방법을 적용한다. f_w 와 W_w 값은 스피커의 f_L 에 의하여 결정되며 f_L 이 100~120 Hz일 경우 $f_w=700$ Hz, $W_w=-100$ dB 이 적절하였다. 생성한 배음 성분의 이득 g_h 는 0~2 dB를 사용하였다.

V. 실험결과

VBS는 생성한 배음 신호를 원음에 더해주는 과정에서 저역 강화와 함께 음질의 저하가 발생한다. 본 절에서는 제안한 가중치 부여 방법과 위상 매칭 방법의 저역 강화 효과 및 저역의 품질을 확인하기 위한 객관 평가와 주관 평가를 수행하였다. 기준 가중치 부여 방법과 제안 가중치 부여 방법에 대한 객관비교를 수행하여 주관적인 품질을 예측하였으며, 청취평가를 통하여 이를 검증하였다. 최종적으로 제안 VBS를 적용한 소형 스피커와 저역 재생 성능이 뛰어난 레퍼런스 스피커 간의 주관 성능 비교 평가를 수행하였다.

1. 배음 가중치 부여 방법 성능 객관 평가

Perceptual Evaluation of Audio Quality^[12] (PEAQ)를 사용하여 배음 가중치 부여 방법의 주관적인 품질을 예측하는 객관 평가를 수행하였다. 평가에 사용된 PEAQ의 Model Output Variables (MOV)는 Total noise-to-mask ratio (NMR), Harmonic structure error (EHS), Objective Differ-

ence Grade (ODG)이다.

NMR과 EHS는 VBS를 사용한 정상상태 신호의 저역 강화에서 주관적인 품질과 높은 관련이 있는 변수이며^[5] 주관적인 품질과 음의 상관관계가 있다. ODG는 원 신호를 기준으로 열화된 신호의 주관적인 품질 저하를 5점 척도로 예측하는 변수이다.

120 Hz의 저지대역과 기울기가 12 dB/Oct인 고역 통과 필터로 소형 스피커의 저역 감쇠를 모델링 하였다. 정상상태 신호의 저역 강화 방법으로 음색 매칭 방법과 지수 감쇠 가중치 부여 방법, 제안한 가변 지수 감쇠 가중치 부여 방법을 사용하였다. 과도상태 신호는 아무런 처리를 하지 않았다.

모든 VBS는 5개의 배음을 생성하였으며 위상 일치 방법을 적용하였다. 음색 매칭 방법은 10개의 인접 프레임의 평균 스펙트럼을 사용하였으며 생성된 배음에 500 Hz의 저지대역을 가지는 저역 통과필터를 적용하였다. 지수 감쇠 가중치 부여 방법과 가변 지수 감쇠 가중치 부여 방법의 파라미터는 예비 청취 실험을 통해 테스트할 음원에서 가장 높은 성능을 보이는 파라미터를 선정하였다. 지수 감쇠 가중치 부여 방법은 6, 9 dB의 감쇠 계수를 사용하였다. 가변 지수 감쇠 가중치 부여방법은 $f_w = 700$ Hz, $W_w = -100$ dB를 사용하였으며, 이는 40~120 Hz의 기본 주파수에 대하여 6~20 dB의 감쇠 계수를 가진다. 생성한 배음 성분의 이득 g_h 은 0 dB이다.

실험에 사용된 음원은 31개의 저역 악기 신호로 41~110Hz의 기본 주파수를 가지며, 일렉트릭 베이스, 어쿠스틱 베이스, 신디사이저 등의 다양한 악기를 포함한다. 원음과 저역을 강화한 음원의 NMR과 EHS, ODG를 각 음원의 기본 주파수에 대하여 측정하였다. 그림 6은 측정 결과로 Var Exp (흑색)는 가변 지수 감쇠 가중치 방법, Exp (적색)는 지수 감쇠 가중치 방법, Envelope (청색)는 음색 매칭 방법의 결과이다.

고정된 감쇠 계수를 사용한 Exp는 기본 주파수가 높아질수록 NMR, EHS가 상승하고 ODG가 감소하는 경향을 보인다. 따라서 의도한 기본주파수 대역보다 높은 기본주파수에 대하여 저역의 주관적인 품질이 감소할 것으로 예측된다. 하지만 그림 6에서 보이는 것처럼 제안한 방법은 다른 가중치 방법보다 NHR, EHS가 낮으며 ODG값이 높게

설계할 수 있어 다른 가중치 방법보다 높은 품질을 가질 것으로 예측된다.

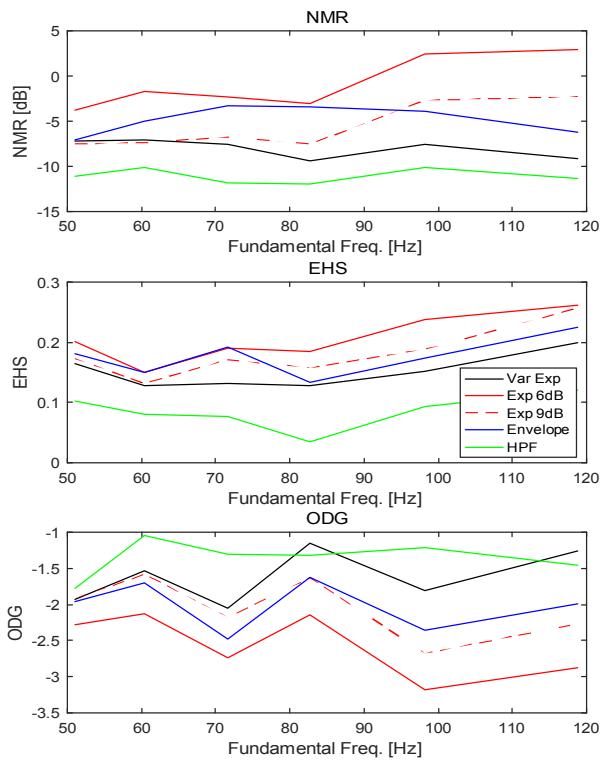


그림 6. 기본 주파수와 배음 가중치 부여 방법에 따른 NMR, EHS, ODG값
Fig. 6. NMR, EHS and ODG values according to different fundamental frequencies and harmonic weighting methods

6 dB의 감쇠 계수를 사용한 Exp는 음색 매칭 방법보다 낮은 품질을 보일 것으로 예측하며, 이는 청취평가 결과[5]와 일치한다. 저역이 제거된 신호의 ODG값이 가장 높게 측정되었는데, 이는 VBS에서 추가한 배음들이 PEAQ측정에서는 왜곡으로 작용하였기 때문이다. 다음 절에서는 위상 일치방법과 제안한 가중치 부여방법의 주관적인 저역강화 성능을 확인하기 위한 청취 실험을 진행하였다.

2. 배음 가중치 부여 방법 및 위상 일치 방법 성능 주관 평가

본 절에서는 저역의 품질 뿐만 아니라 저역의 양 측면에서 제안 가중치 부여방법과 위상 일치방법의 효과를 확인

하기 위한 청취실험을 진행하였다. 11명의(남 10명, 여 1명) 잘 훈련된 피 실험자를 대상으로 ITU-T BS.1534^[13]의 MUSHRA (MULTI Stimulus test with Hidden Reference and Anchor)실험을 수행하였다.

다양한 저역악기에 대하여 VBS의 성능 비교를 위하여 어쿠스틱 베이스, 일렉트릭 베이스, 신디사이저, 첼로가 포함된 7개 장르의 음원을 사용하였다. 각 음원의 길이는 10~15초이다. 외부 소음이 없는 청취실에서 HD600 헤드폰과 DAC를 사용하여 실험을 수행하였다.

원 음원을 레퍼런스로 하였으며, 120 Hz의 저지대역과 12 dB/Oct.의 기울기의 고역 통과 필터를 사용하여 저역을 제거한 음원이 앵커(Anchor)이다. 저역을 제거한 음원에 [표 1]의 저역 강화 방법을 적용하였다.

표 1. 배음 가중치 부여 방법 및 위상 일치 방법 성능 주관 평가에 사용된 레퍼런스, 앵커 및 저역 강화 신호

Table 1. Reference, anchor and bass enhanced signal for the subjective test of the harmonic weighting and phase matching scheme

Test condition	Parameter	Phase matching	LPF
Reference (Ref.)	N/A	N/A	X
Anchor (Anc.)	N/A	N/A	
Timbre matching (T)	N=10	O	
Exponential weighting 1 (E1)	$\alpha = 9\text{dB}$	O	
Exponential weighting 2 (E2)	$\alpha = 9\text{dB}$	X	
Variable slope Exp. weighting (V)	$f_w = 700\text{Hz}$, $W_w = -100\text{dB}$	O	
Shelving filter EQ (EQ)	$f_C = 150\text{Hz}$, $G = 6\text{dB}$	N/A	

$f_c = 120\text{Hz}$
 -12dB/Oct.

지수 감쇠 가중치 방법(E1, E2)의 감쇠 계수는 실험할 음원을 이용한 예비 실험에서 가장 높은 저역 강화 효과를 보이는 9 dB 값을 사용하였다. 이는 100 Hz의 저지대역을 가지는 고역 통과 필터를 사용한 실험[5]에서 사용한 6, 12 dB의 중간 값이다. 음색 매칭 방법(T)은 위 논문[5]과 동일한 파라미터를 사용하였다. 가변 지수 감쇠 가중치 방법(V)은 $f_w = 700\text{Hz}$, $W_w = -100\text{dB}$ 를 사용하였으며, 이 경우 60 Hz에서 9 dB의 감쇠 계수를 가진다. 모든 VBS의 g_h 는 2 dB이며, E2를 제외한 모든 VBS는 위상 매칭을 적용하였

다. Shelving filter EQ는 성능 비교를 위해 사용되었다.

평가 항목은 저역의 품질(Overall quality)과 저역의 양(Bass intensity)이다. 저역의 품질은 저역 강화 방법을 적용한 음원에서 발생한 주관적인 음질 열화의 정도를 원 음원과 비교하여 평가한 항목이다. 저역의 양은 저역 강화 방법을 적용한 음원의 주관적인 저역의 양을 원음과 비교하여 평가하는 항목이다.

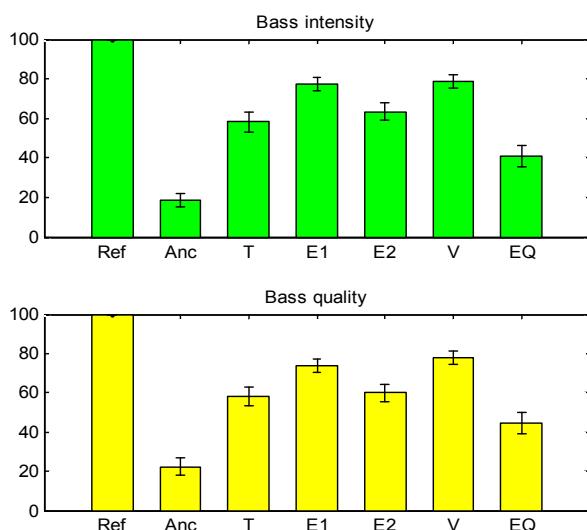


그림 7. 모든 음원에 대한 배음 가중치 부여 방법 및 위상 일치 방법 성능 주관 평가 결과, 저역의 양(위), 저역의 품질(아래)

Fig. 7. Subjective test result of the harmonic weighting and phase matching scheme, bass intensity (up) and bass quality (down) for all music clips

그림 7은 전체 음원에 대한 배음 가중치 부여 방법 및 위상 일치 방법 성능 주관 평가 결과이다. 위상 일치 방법을 적용한 지수 감쇠 저역 강화 방법인 E1은 위상 일치 방법을 적용하지 않은 E2에 비해 저역의 양과 저역의 품질에 대해 각각 15점 높은 점수를 획득하였으며, 95% 신뢰구간이 겹치지 않았다. E1과 V의 경우, 전체 음원을 평균을 낸 값에서는 제안한 방법이 지수 감쇠 저역 강화 방법에 비해 저역의 양이 1점, 저역의 품질이 4점 높았지만 통계적으로 유의한 차이가 없었다.

각 음원에 대한 실험 결과는 그림 8과 같다. 위상 일치 방법을 적용한 E2는 위상 일치 방법을 적용하지 않은 E1과 비교하여 다양한 음원에서 강인한 저역 강화 성능을 가지

는 것을 확인할 수 있다. Jazz 음원은 다른 저역 악기와 비교하여 과도상태 소리를 상대적으로 많이 생성하는 어쿠스틱 베이스 기타를 사용하기 때문에 저역의 양 측면에서는 상쇄간섭에 의한 영향이 적지만 품질 측면에서는 큰 차이가 나타난다.

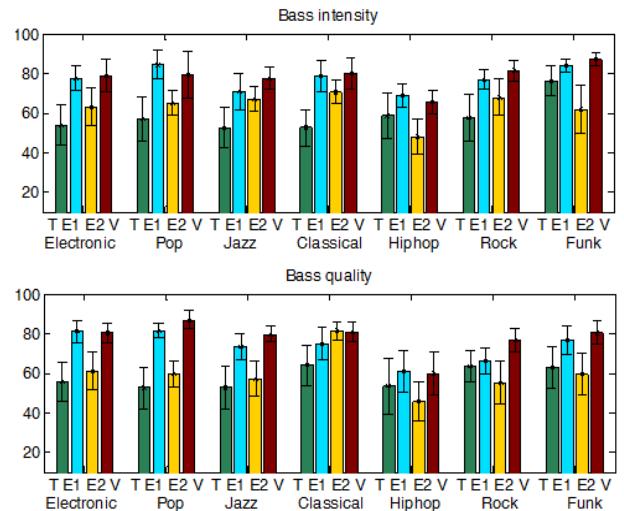


그림 8. 각 음원에 대한 배음 가중치 부여 방법 및 위상 일치 방법 성능 주관 평가 결과, 저역의 양(위), 저역의 품질(아래)

Fig. 8. Subjective test result of the harmonic weighting and phase matching scheme, bass intensity (up) and bass quality (down) for each music clip

저역 신호만 있는 경우 가장 높은 성능을 얻었던 음색 매칭 방법^[5]은 일반적인 음원을 사용한 실험에서 가장 낮은 성능을 보였는데 그 원인은 2장 1절에서 살펴본 이유로, 다른 악기로 인하여 실제 생성해야 할 대역보다 높은 대역에 배음을 생성하였기 때문이다.

일반적으로 저역의 양을 증가시키기 위하여 배음을 추가 할수록, 높은 주파수 성분이 증가하여 저역의 품질은 낮아지는 Trade-off 관계가 성립한다. 하지만 60 Hz 이상의 기본 주파수를 많이 포함하는 Rock과 Jazz, Pop 음원의 경우, V가 E1보다 더 높은 감쇠 계수를 가지고 있지만 저역의 양과 저역의 품질 모두에서 다소 높은 점수를 얻었다. V는 생성한 배음의 양이 E1보다 작지만 생성된 배음의 대역이 저역 대역으로 제한되어서 나타난 효과로 예측되지만 통계적으로 큰 차이는 없었다.

베이스 신호만을 사용한 예비 청취 실험과 객관 실험에서는 높은 기본주파수(약 80 Hz이상)에서 E1보다 V가 저역의 품질 측면에서 효과적이었지만, 일반적인 음악신호에서는 성능차이가 크지 않았다. 그 이유로는 일반적인 음악에서 저역 악기의 기본주파수가 주로 낮은 값을 가지기 때문에 E1과 V간에 차이가 적다. 또한 저역 악기가 높은 기본주파수를 가지는 경우에도 생성된 배음들의 크기가 다른 악기 혹은 음성보다 상대적으로 작기 때문에 마스킹(Masking) 현상으로 인하여 음색의 차이가 크지 않다.

3. 제안 저역 강화 시스템 성능 주관 평가

약 100 Hz의 저역 재생 한계 주파수를 가지는 소형 스피커에 위상 일치 방법과 제안 가중치 부여 방법의 적용 여부에 대한 효과를 확인하기 위하여 ITU-T BS.1284^[14]의 5점 척도 비교 실험을 수행하였다. 평가 항목은 저역의 양과 저역의 품질이다. 레퍼런스 신호는 49 Hz의 저역 재생 한계 주파수를 가지는 모니터 스피커에서 재생된 신호이다. 실험 음원으로 30초 내외의 Pop음원 2개와 Rock음원을 사용하였다. 각 음원은 청취자점에서 평균 음압이 70 ± 0.5 dB이 되도록 보정하였다.

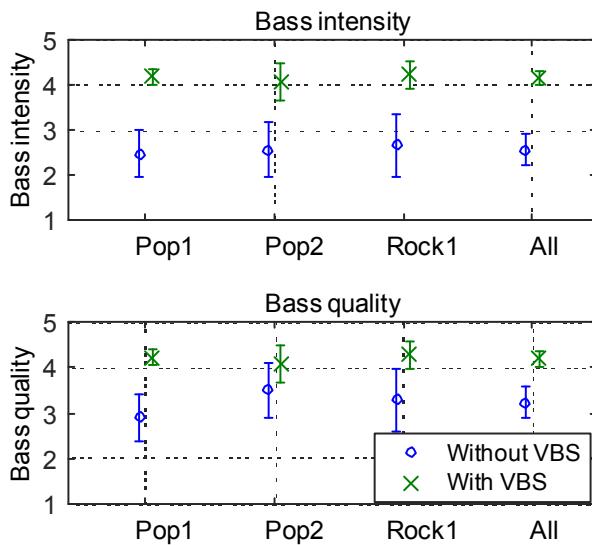


그림 9. 레퍼런스 스피커와 비교한 제안 저역 강화 시스템 성능 주관 평가 결과, 저역의 양(위), 저역의 품질(아래)

Fig. 9. Bass intensity (up) and bass quality (down) of the proposed VBS compared with the reference speaker

실험 결과는 그림 9와 같다. 레퍼런스 신호를 5점 만점 기준으로, 제안한 저역 강화 방법은 저역의 양을 1.6점, 저역의 품질을 1점 향상시켰다.

VI. 결 론

본 논문은 기존 페이즈 보코더 기반 VBS의 성능을 향상하기 위한 위상 일치 방법과 가변 지수 감쇠 가중치 부여 방법을 제안하였다. 기존 VBS는 기본 주파수 성분의 위상을 이용하여 배음 성분의 위상을 계산하는데, 이와 달리 위상 일치 방법은 기존 저역 성분의 배음의 위상을 그대로 사용하는 방법이다. 더 간단할 뿐만 아니라, 기본주파수와 배음성분이 서로 결맞지 않을 때에도 기존 방법과 비교하여 효과적인 저역 강화 성능을 보여준다.

가변 지수 가중치 부여 방법은 지수 가중치 부여 방법에서 발생하는 저역 강화의 성능과 음색의 열화간의 Trade-off 관계를 개선하기 위하여 제안되었다. 그러나 예비 테스트를 통하여 각 저역 강화 방법마다 적절한 파라미터를 선정한 다음 일반적인 음악신호에 저역 강화를 적용한 실험에서는 통계적으로 의미 있는 성능향상을 얻지 못하였다. 높은 기본 주파수의 저역 신호만을 사용한 예비 테스트에서만 더 높은 주관적인 성능을 보여주었다.

본 논문에서는 생성한 배음의 이득을 고정된 값으로 사용하였다. 그러나 라우드니스(Loudness)는 음질에 중요한 영향을 미치기 때문에 콘텐츠 제작자가 의도한 라우드니스를 고려하는 것이 주관적인 선호도와 높은 관련이 있을 것으로 예상된다. 후속 연구에서는 재생 불가능한 대역으로 인하여 감소한 라우드니스를 고려한 저역 강화 전략을 찾아낼 필요가 있다.

참 고 문 헌 (References)

- [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] H. Kim, N. Moreau and T. Sikora, "MPEG-7 Audio and Beyond ; Audio Content Indexing and Retrieval," Wiley, England, pp. 45-46., 2005.
- [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] 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.
- [10] D. Ben-Tzur and M. Colloms, "The Effect of MaxxBass Psychoacoustic Bass Enhancement on Loudspeaker Design," in 106th AES convention, Munich, Germany, 1999.
- [11] D. Fitzgerald, "Harmonic/Percussive Separation Using Median Filtering," in 13th International Conference on Digital Audio Effects (DAFX10), 2010.
- [12] "Method for objective measurements of perceived audio quality," ITU Recommendation BS.1387-1, 2001.
- [13] "Method for the Subjective Assessment of Intermediate Quality Level of Audio Systems," ITU Recommendation BS.1534-2, 2014.
- [14] "General methods for the subjective assessment of sound quality," ITU Recommendation BS.1284-1, 2014.

저자소개

문현기



- 2013년 : 연세대학교 전기전자공학부 학사 졸업
- 2013년 ~ 현재 : 연세대학교 전기전자공학부 석박 통합 과정
- ORCID : <http://orcid.org/0000-0001-8806-6335>
- 주관심분야 : 오디오 신호처리, 3D 오디오, 오디오 부호화

박영철



- 1986년 : 연세대학교 전자공학과 학사 졸업
- 1988년 : 연세대학교 전자공학과 석사 졸업
- 1993년 : 연세대학교 전자공학과 박사 졸업
- 현재 : 연세대학교 컴퓨터정보통신공학부 교수
- ORCID : <http://orcid.org/0000-0003-3274-076X>
- 주관심분야 : 음성 신호처리, 적응 신호처리, 오디오 신호처리

황영수



- 1982년 : 연세대학교 전자공학과 학사 졸업
- 1994년 : 연세대학교 전자공학과 석사 졸업
- 1990년 : 연세대학교 전자공학과 박사 졸업
- 현재 : 가톨릭관동대학교 전자공학과 교수
- ORCID : <http://orcid.org/0000-0002-7188-3678>
- 주관심분야 : 음향 및 음성 신호처리, 음향학, 멀티미디어