

Spectral Band Replication 대역폭 변화에 따른 오디오 부호화 성능 평가

*안경준 **김인철

서울시립대학교 전자전기컴퓨터공학과

*akj0628@paran.com **rin@uos.ac.kr

Performance Evaluation of Audio Coders According to Spectral Band Replication Bandwidth

*An, Kyung-Jun **Kim, Rin-Chul

Dept. of Electrical & Computer Eng., University of Seoul

요약

본 논문에서는 오디오 신호의 부호화에 있어 고대역 신호의 부호화를 담당하는 대역 확장 기법인 SBR 대역폭의 변화에 따른 부호화 성능 변화에 대해 고찰한다. 주부호화기(Core coder)의 주파수 대역에 따라 적절한 SBR의 고대역 적용 범위를 선택하여 음질 평가를 통해 부호화 성능을 평가한다.

1. 서론

오디오 부호화기에서 SBR(Spectral Band Replication)은 적은 양의 비트(bit)로 고대역 신호를 부호화 하는 대표적인 대역 확장 기법이다. SBR은 MPEG HE-A¹AC 나 AAC+부호화기에서 사용된다.

본 논문에서는 음향과 음성 오디오를 부호화 함에 있어 그림 1에서와 같이 SBR을 적용하는 시작 주파수(Start Frequency)와 끝 주파수(Stop Frequency)의 변화를 통해서 음질의 차이를 비교하여 적합한 SBR 대역을 제안하고 평가하고자 한다.

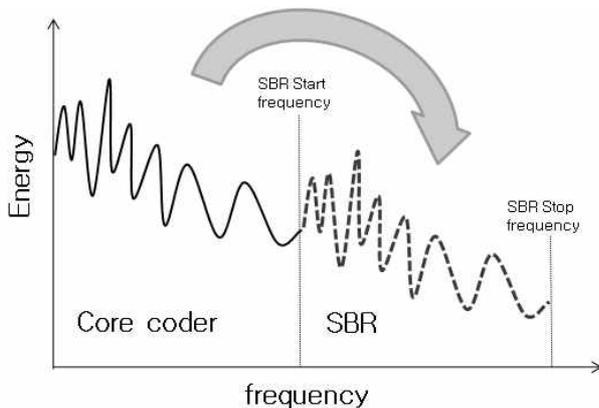


그림 1. SBR

2. 대역확장기법 SBR(Spectral Band Replication)

대역 확장 기법이란 오디오 신호의 저주파 성분과 고주파 성분이 높은 연관성을 갖고 있음에서 시작된다.

SBR의 기본 원리는 주부호화기를 통해 부호화한 저대역 신호를 근간으로 하여 적은 비트수의 파라미터를 이용하여 고대역 신호를 복

원하는 것이다. 이때 고대역 신호 복원에 사용되는 파라미터를 SBR 파라미터라고 한다.

SBR 부호화기로 입력된 신호는 QMF(Quadrature Mirror Filterbanks) 분석을 통해 64 대역 신호로 분해된 다음 SBR 부호화기에서 SBR 파라미터를 추출하여 전송하게 되고 복호기에 파라미터를 이용하여 저대역 신호를 복사하고 수정하여 고대역 신호를 복원한다.

가. SBR 대역 결정

SBR 부호화기에서의 대역 결정 방법은 주부호화기에 따라 결정된다. SBR 시작 주파수와 끝 주파수는 비트전송률(bit rate)과 채널 수, 그리고 표본화율(sampling rate)에 따라서 결정되어 진다. 저대역 신호는 주부호화기에서 담당하게 되고 이를 토대로 고대역 신호를 복원하게 된다.

SBR은 전대역(full band)을 64-채널의 QMF를 사용하여 주파수 성분으로 분석하기 때문에 시작 주파수와 끝 주파수는 QMF 인덱스(Index)로 결정 된다.

SBR의 시작 주파수의 결정 방법은 주부호화기의 대역 이후의 값으로 정의 내릴 수 있다. 하지만 음성, 음향 등의 오디오의 성격과 주부호화기의 특성에 따라 시작 주파수에 의한 음질의 차이를 보일 수 있다.

오디오가 음성 신호만 가지고 있을 경우 낮은 주파수 대역에서 큰 에너지를 갖기 때문에 끝 주파수는 큰 영향을 주지 않는다. 반면 고주파수를 가지고 있는 음향신호의 경우에는 고대역 신호의 복원이 중요하게 작용하기 때문에 끝 주파수에 따라 큰 영향을 받는다고 할 수 있다.

본 논문에서는 주부호화기의 대역폭이 결정된 이후에 SBR이 적용될 시작 주파수와 끝 주파수에 변화를 주어 대역폭에 따른 음질 평가를 수행한다.

본 연구는 지식경제부, 방송통신위원회, 및 한국산업기술평가관리원의 산업원천기술개발사업(정보통신)의 일환으로 수행하였음.[과제번호 KI001932, 차세대 DTV 핵심 기술 개발]

3. 실험 및 결과

가. 실험 방법

본 논문에서는 SBR의 시작 주파수와 끝 주파수에 따른 음질 평가 방법으로 주관적 평가인 청취 테스트와 객관적 평가인 PEAQ 테스트를 사용하였다.

주관적 평가자는 4명의 음악 작곡가와 2명의 학생으로 평가하였다. 평가 방법은 1점(듣기 힘들 정도로 거북하게 들린다)에서 5점(음질이 매우 자연스럽다)으로 평가 하여 평균을 구하였다.

객관적 평가 기준인 ODG(Objective Difference Grade)값은 0에서 부터 -5까지 평가되며 0에 가까울수록 음질 차이가 없음을 뜻한다.

실험에 사용한 오디오 신호는 48kHz의 표본화율을 갖는 Arirang_speech(남성의 나레이션), twinkle_ff5(여성의 나레이션과 배경음악), tel5(오케스트라)의 오디오 신호를 12kbps의 비트전송률로 부호화 하여 실험을 진행하였다. 부호화 방법은 MDCT를 이용한 TCX 또는 ACELP 부호화기를 사용하였으며 주부호화기의 대역폭은 6kHz 이다. 시작 주파수에 따른 음질을 평가하기 위해 끝 주파수를 15.375kHz로 고정하고 시작 주파수를 5.25, 5.625, 6kHz로 변화를 주어 테스트를 하였고, 끝 주파수에 따른 음질을 평가하기 위해 시작 주파수를 5.625kHz로 고정하고, 끝 주파수를 10.875, 13.125, 15.375, 18.375kHz 로 변화를 주어 실험 한다.

나. 실험 결과

ㄱ. SBR 시작 주파수에 따른 음질 평가

SBR 시작 주파수에 따른 주관적 음질 평가에서 음성신호의 경우 표 1에서와 같이 시작 주파수에 따라 음질 차이를 구분하기가 힘들 정도로 비슷하다고 할 수 있다. 음향 신호의 경우에는 시작 주파수가 5.25kHz일 경우 살짝 건조함이 느껴지고 시작 주파수가 6kHz일 경우에는 잡음이 다른 것들에 비해 심하게 들렸다. 시작 주파수가 5.625kHz일 때의 특성은 위의 두 가지 특성을 조금씩 가지고 있었다. 위와 같은 결과는 주부호화기의 특성상 6kHz 주변 신호의 에너지가 감소하기 때문에 SBR 시작 주파수가 6kHz일 경우 에너지 차이가 발생하게 되고 5.35kHz인 경우 정확하게 표현된 주부호화기 신호가 줄어들기 때문에 손실을 가져오게 된다.

시작 주파수에 따른 객관적 음질 평가는 표 2 에서와 같이 음성신호와 음향신호에 관계없이 시작 주파수를 최대화 하였을 때 좋은 결과를 보여준다. 위와 같은 결과는 시작 주파수가 5.25kHz일 경우 주부호화기의 신호를 SBR로 표현하게 되고 이때 왜곡이 발생하기 때문에 객관적 평가에 영향을 미치게 된다.

start frequency	5.25 kHz	5.625 kHz	6 kHz
Arirang_speech	3	3	3
twinkle_ff51	3.5	3.17	3.17
tel5	3.33	3.5	3

표 1. 시작 주파수에 따른 주관적 음질평가 결과 요약.

start frequency	5.25 kHz	5.625 kHz	6 kHz
Arirang_speech	-3.389	-3.342	-3.295
twinkle_ff51	-3.098	-3.078	-3.054
tel5	-3.117	-3.083	-3.021

표 2. 시작 주파수에 따른 객관적 음질평가 결과 요약.

ㄴ. SBR 끝 주파수에 따른 음질 평가

SBR 끝 주파수에 따른 주관적 음질 평가에서는 표 3 에서와 같이

음성신호만 존재 할 경우 음성신호는 저대역에서 강한 신호를 갖기 때문에 대역폭을 확장한 결과와 큰 차이를 보이지 않고 있다. 그러나 음성과 음향 신호가 같이 존재하거나 음향신호만 있을 경우 대역폭이 넓어짐에 따라 좋은 평가를 받는 것을 확인 할 수 있다. 특히 심벌 소리와 같은 고대역의 신호의 경우 더 선명하게 들을 수 있다.

끝 주파수에 따른 객관적 음질 평가에서는 표 4 에서와 같이 끝 주파수가 18.375kHz인 경우 모두 가장 좋은 평가를 받았다. 18.375kHz를 제외한 특성은 음성신호만 존재 할 경우 끝 주파수가 커짐에 따라 음질이 저 평가 되는 특성을 나타내고 있고, 음성과 음향이 같이 존재 할 경우에는 큰 차이를 보이지 않았다. 음성이 없고 음향 신호만 존재 할 경우 대역폭이 넓어짐에 따라 좋은 평가를 보여 주고 있다.

stop frequency	10.875 kHz	13.125 kHz	15.375 kHz	18.375 kHz
arirang_speech	2.5	2.67	3	2.83
twinkle_ff51	1	2.5	3	3.5
tel5	1	2	2.83	3.17

표 3. 끝 주파수에 따른 주관적 음질평가 결과 요약.

stop frequency	10.875 kHz	13.125 kHz	15.375 kHz	18.375 kHz
arirang_speech	-3.179	-3.325	-3.344	-3.059
twinkle_ff51	-3.035	-3.106	-3.078	-2.709
tel5	-3.231	-3.208	-3.083	-2.824

표 4. 끝 주파수에 따른 객관적 음질평가 결과 요약.

4. 결론

본 논문에서는 음질 향상을 위해 SBR의 적용범위인 시작 주파수와 끝 주파수 결정 범위를 제안하고자 하였다. 실험 결과를 통하여 SBR의 적용 범위는 오디오 파일을 특성에 따라 달라짐을 볼 수 있었다. 특히 음성신호만 있는 경우에는 저대역의 특성이 강하기 때문에 SBR 대역에 따른 음질 변화는 작았다. 그러나 음향신호의 경우에는 표현하고자 하는 주파수 대역이 넓을 때 SBR의 대역폭에 따라 큰 차이를 보이고 있다.

앞에서의 실험은 SBR 대역에 달라짐에 따라 SBR에 적용될 파라미터 값들도 달라지기 때문에 위의 실험이 절대적인 결과라고 보기는 힘들다. 하지만 위의 결과를 통해 SBR의 대역을 결정함에 있어서 현재 제안하고 있는 전송률과 표본화율을 가지고 SBR 대역을 결정 하기에는 문제점들이 있다고 생각하며 오디오의 특성을 파악하여 SBR의 대역을 결정할 수 있다면 향상된 음질을 표현하는 부호화기를 만들 수 있다고 본다.

참고문헌

- [1] ISO/IEC 14496-3:2008 General Audio Coding.
- [2] ITU-R BS.1387-1 Perceptual Evaluation of Audio Quality (PEAQ).
- [3] E. Laresen and R. M. Aarts, Audio Bandwidth Extension, *John Wiley & Sons*, 2004.
- [4] Thomas Ziegler, Andreas Ehret, Per Ekstrand, and Manfred Lutzky. "Enhancing MP3 with SBR: Features and Capabilities of the New MP3PRO Algorithm." *112th AES Convention*, Munich, May, 2002. Preprint 5560.