

주부호화기 대역폭에 따른 오디오 부호화의 성능비교

*정용석 **김인철

서울시립대학교 전자전기컴퓨터공학과

*yongseok82@uos.ac.kr **rin@uos.ac.kr

Performance Comparison of Audio Coders According to Core Codec Bandwidth

*Jeong, Yong-Seok **Kim, Rin-Chul

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

요약

본 논문에서는 음향신호의 부호화에 있어 주부호화기로 부호화 되는 주파수 대역폭이 음질에 미치는 영향에 대하여 고찰한다. 또한 비트율을 변화시켜 양자화 잡음 발생을 줄일 수 있는 비트율의 기준에 대하여 고찰한다. 마지막으로 주관적 음질평가 및 객관적 음질평가를 통하여 그에 따른 성능을 평가한다.

1. 서론

일반적인 음성신호의 경우 주파수영역에서 관찰하였을 때 약 6kHz이하의 저주파 영역에 많은 에너지가 존재한다. 실제로 사용되는 부호화 방식에서도 저주파 영역에 대해 충실한 부호화를 수행한다[1]. 음향신호의 경우 음성신호와 다른 주파수 특성을 갖지만 음성신호와 비슷한 방식으로 저주파 영역 중심의 부호화를 수행하는 방식을 사용하고 있다[2].

본 논문에서는 음향신호의 부호화에 있어 주부호화기로 부호화되는 대역폭을 변화시켜 그에 따른 음질을 비교하여 본다. 또한 전체적인 비트율(bit-rate)을 변화시켜 주부호화기의 부호화 대역폭에 따른 최상의 음질을 나타내는 비트율을 제시한다. 그리고 주관적 음질평가 및 객관적 음질평가를 통해서 그 성능을 평가한다.

2. 음향 신호 부호화 방법

일반적으로 오디오 파일의 부호화는 에너지가 많이 존재하고 있는 약 6kHz이하의 저주파 성분에 대해 충실한 부호화를 진행한다. 또한 비교적 적은 에너지가 존재하는 6kHz 이상의 고주파 성분에 대해서는 적은양의 비트를 할당하여 부호화를 실행한다. 예를들어 MPEG AAC+(Advanced Audio Coding)에서는 SBR(Spectral Band Replication)을 이용하여 고주파 영역의 신호 복원을 수행한다[2].

전체적인 비트율이 증가하는 경우에 있어 고주파 영역에 할당되는 비트수는 크게 증가하지 않으며 저주파 영역 위주로 비트할당량이 증가한다. 이러한 형태의 부호화는 낮은 비트율에서와 높은 비트율에서 고주파 영역에 존재하는 정보를 비슷한 수준으로 복원시켜 고주파 영역에 대한 부호화 충실도는 증가하지 못한다. 또한 저주파 영역의 신호에서도 일정 수준의 비트율을 초과하면 양자화 잡음의 감소에 따른 음질 향상의 정도 또한 감소한다.

본 논문에서는 주부호화기의 대역폭을 증가시켜 12kHz이하의 전

영역에 대해서 주부호화기를 이용한 부호화를 수행한다. 부호화 방식으로 MDCT를 이용한 TCX방식을 사용한다. 또한 저주파 영역에 할당되는 비트수의 감소에 따른 음질의 저하를 막기 위한 최소한의 비트율에 대해서도 고려하였다.

가. 주부호화기 부호화 대역폭

주부호화기(Core Codec)에 의해 부호화되는 저주파 영역의 대역폭을 조절하여 약 6kHz 또는 4kHz이하의 주파수 영역 위주의 부호화를 12kHz 이하의 전체 주파수 영역으로 확장시켜 고주파 영역에 대한 부호화를 충실히 수행한다.

최초에 입력된 시간영역에서의 신호는 표본화 주파수 48kHz의 모노신호이며, 이는 24kHz의 대역폭을 갖는다. 이 신호를 12kHz, 6kHz 또는 4kHz를 전대역으로 갖는 신호로 만들기 위해 2배, 4배, 또는 6배의 하향표본화(down sampling)를 각각 수행한다. 이러한 신호를 각각 주부호화기의 입력으로 인가시켜 부호화를 수행한다.

6kHz 또는 4kHz이하의 저주파 성분에 대해서는 주부호화기를 이용하여 충실한 부호화를 수행하는 경우, 주부호화기는 MDCT(Modified Discrete Cosine Transform)를 이용한 TCX 부호화기를 사용하여 부호화를 수행하며, 6kHz(또는 4kHz)~12kHz의 주파수 범위에 존재하는 성분에 대해서는 SBR을 이용하여 고주파 영역을 복원한다. 12kHz 이내의 전영역의 성분에 대하여 주부호화기를 이용한 부호화를 실행하는 경우에서도 위와 동일한 MDCT를 이용한 TCX 부호화기를 사용하였다.

주부호화기에의 MDCT를 이용한 TCX 부호화기는 각각의 대역폭(12kHz, 6kHz, 4kHz)을 갖는 신호 1024샘플에 대하여 AMR-WB+에서와 같은 방식으로 최대의 SEGSNR(SEGmental Signal to Noise Ratio) 값을 갖는 서프프레임(sub-frame) 길이로 각각 부호화를 수행한다.

	12kbps	14kbps	16kbps	18kbps
0~12kHz	505.064	591.653	669.938	756.247
0~6kHz	878.157	995.518	1115.215	1326.953
0~4kHz	1233.427	1483.664	1733.55	1981.748

표 1. 프레임당 주부호화기 부호화 대역폭 별 평균 할당 비트수.

본 연구는 지식경제부, 방송통신위원회, 및 한국산업기술평가관리원의 산업원천기술개발사업(정보통신)의 일환으로 수행하였음.[관저번호 KI001932, 차세대 DTV핵심기술개발]

표 1.에서는 비트율 변화에 따른 3가지 실험파일의 주부호화기에 대해 프레임별 할당 비트수의 평균값을 나타내었다.

본 논문에서는 주부호화기로 부호화되는 대역폭을 조절하여 각각의 대역폭에 따른 오디오 부호화의 성능을 비교하여 음질에 미치는 영향에 대해서 파악하고자 한다.

나. 양자화 잡음 감소를 위한 비트율 설정

주부호화기를 이용하여 전 주파수 영역을 부호화한 경우 저주파 영역만을 주부호화기로 부호화하는 경우에 비해 저주파 영역에 할당되는 비트수가 감소하여 저주파 영역에 대한 양자화 오차가 증가하게 된다. 이에 따라 발생하는 양자화 잡음(Quantization noise)이 발생하며, 이는 음질을 저하시킨다.

본 논문에서는 이러한 점을 고려하여 양자화 잡음의 영향을 줄일 수 있는 비트율 값을 실험적으로 고찰한다. 12kbps부터 18kbps까지 변화시켜가며 그에 따른 음질에 대해서 주관적 음질평가 및 PEAQ를 통한 객관적 음질평가를 수행한다.

3. 실험 및 결과

가. 실험 방법

본 논문에서 고찰한 주부호화기를 이용한 부호화 대역폭의 변화에 따른 성능을 평가하기 위해 주관적 음질 평가를 수행하였다. 본 논문에서는 음성신호처리를 전공하는 3명의 학생을 대상으로 주관적 평가를 시행하였다. 평가결과는 1부터 5까지의 값을 갖으며 5에 가까운 값일수록 좋은 음질을 보이고 있음을 의미한다.

또한 객관적 음질평가로 PEAQ (Perceptual Evaluation of Audio Quality)를 활용하였다[3]. 음질평가는 ODG(the Objective Difference Grade)를 기준으로 사용하였다. ODG값은 0부터 -5까지의 값을 갖으며 0에 가까운 값일수록 원본파일과의 오차가 작음을 의미한다.

평가에 사용된 실험 오디오는 표본화 주파수 48kHz의 mono 신호이다. 각각의 파일은 전자기타, 베이스, 드럼으로 이루어진 음악파일(14.5초, Music_1.wav)과 금관악기, 베이스 및 여성의 나레이션으로 이루어진 음악파일(14.5초, twinkle_ff51.wav), 현악기 및 금관악기로 이루어진 음악파일(14.5초, te15.wav)등 3곡이다.

나. 실험 결과

12kHz 전 영역에 대하여 MDCT를 이용한 TCX부호화기로 부호화를 실행한 결과 16kbps이하의 비트율을 갖는 경우 양자화 오차의 발생에 따라 음질이 저하되며, 6kHz 또는 4kHz 영역에 대한 TCX부호화기를 이용한 부호화의 결과보다 낮거나 비슷한 수준의 음질을 갖는다. 이는 저주파 영역에 할당되는 비트량의 감소에 따른 양자화 오차의 증가를 원인으로 볼 수 있다.

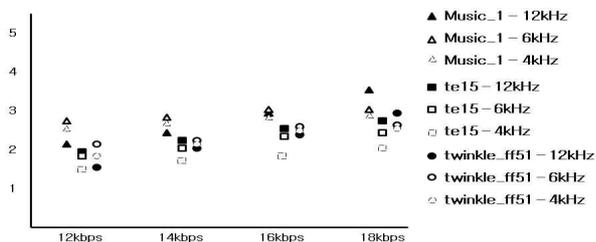


그림 1. 주관적 음질평가 결과 요약.

표 2. 객관적 음질평가 결과 요약.

		12kbps	14kbps	16kbps	18kbps
music_1	12kHz	-3.309	-3.300	-3.252	-3.220
	6kHz	-3.366	-3.342	-3.318	-3.298
	4kHz	-3.342	-3.332	-3.127	-3.043
te15	12kHz	-3.003	-3.040	-3.087	-3.105
	6kHz	-3.145	-3.130	-3.169	-3.174
	4kHz	-3.308	-3.290	-3.223	-3.177
twinkle_ff51	12kHz	-2.967	-2.974	-2.960	-2.942
	6kHz	-3.047	-3.039	-3.030	-3.030
	4kHz	-3.143	-3.106	-3.081	-3.037

18kbps이상의 비트율에서는 양자화 오차에 의한 잡음이 감소하며, 고주파 영역에 대한 충실한 부호화로 이전 방식보다 음질이 향상된다.

그림 1.에서는 주관적 음질평가의 결과를 요약하고 있다. 6kHz 또는 4kHz까지 주부호화기를 이용한 부호화를 수행하는 경우 비트율의 증가에 따른 음질향상의 정도가 작다. 잡음의 양은 모든 비트율에서 비슷하게 존재하며, 이는 음의 선명성을 저하시킨다.

전 영역에 대한 주부호화기를 이용한 부호화를 수행한 결과 비트율의 증가에 따라 양자화의 충실도가 증가하며 이는 전체적인 음질을 향상시키며 그 결과로 풍부한 음의 특성을 갖는다. 또한 6kHz 또는 4kHz의 대역폭에 대해 주부호화기를 이용한 부호화를 수행한 경우에서 비트율의 증가에 따른 음질향상의 정도가 작게 나타나는 결과에 대해 주부호화기의 대역폭을 증가시킴으로써 비트율의 증가에 따른 음질향상의 정도를 증가시킬 수 있다.

표 3.에서는 객관적 음질평가의 결과를 요약하고 있다. ODG값은 주관적 음질평가와 비슷한 양상을 보이고 있음을 확인할 수 있다.

4. 결론

본 논문에서는 부호화 하고자 하는 신호에 대해 주부호화기를 이용한 부호화방식의 주파수 대역폭에 대해 고찰하였다. 또한 비트율에 관한 기준을 제시하고자 하였다.

먼저, 부호화 하고자하는 전 영역의 대역폭에 대하여 주부호화기를 이용한 부호화를 수행하였다. 저주파 영역에 할당되는 비트량의 감소에 따른 전체적인 비트율의 기준을 제시하였다.

그리고 3가지 실험 오디오에 대하여 주관적 음질평가 및 객관적 음질평가를 시행한 결과 18kbps이상의 비트율을 갖는 상황에서 전체 주파수 영역에 대해 주부호화기를 이용한 부호화가 주관적 품질을 개선함을 확인했다. 또한 PEAQ를 통한 객관적 평가에서도 주관적 평가와 비슷한 양상을 보이고 있음을 확인할 수 있다.

본 논문을 통해 구현한 방법은 주부호화기의 성능 향상을 고려하면 18kbps 이상의 비트율이 보장되는 상황에서는 주부호화기의 대역폭을 12kHz으로 확장하여 부호화를 수행하는 방식이 6kHz 또는 4kHz를 주부호화기 대역폭으로 설정하는 경우보다 더욱 향상된 음질을 보일 수 있을 것이라 기대된다.

참고문헌

[1] M. Honda, F. Itakura, "Bit allocation in time and frequency domains for predictive coding of speech," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-32, no. 3, pp.465-473, JUNE 1984.
 [2] ISO/IEC 14496-3:2008 General Audio Coding
 [3] Rec. ITU-R BS.1387-1 Perceptual Evaluation of Audio Quality