한국음향학회지 제21권 제6호 pp. 566~575 (2002)

저비트율로 압축된 오디오의 음질 개선 방법

Audio Quality Enhancement at a Low-bit Rate Perceptual Audio Coding

서 정 일*, 서 진 수**, 홍 진 우*, 강 경 옥*
(Jeong-Il Seo*, Jin-Soo Seo**, Jin-Woo Hong*, Kyeong-Ok Kang*)

*한국전자통신연구원 방송미디어연구부, **한국과학기술원 전자전산학과 전자공학 전공
(접수일자: 2002년 5월 20일; 채택일자: 2002년 7월 22일)

이동통신망과 같이 제한된 대역폭에서 실시간 멀티미디어 스트리밍 서비스를 제공하기 위해서는 보다 낮은 비트율로 비디오와 오디오 데이터를 압축하여야 한다. 또한 대부분의 대역이 비디오 데이터를 위해 할당되어 있으므로 제한된 대역폭만이 오디오에 할당되게 된다. 오디오 데이터를 낮은 비트율로 압축하기 위해서는 압축율이 높은 알고리즘을 사용하거나, 표본화 주파수 (sampling frequency)를 낮춤으로써 데이터양을 줄여 낮은 비트율로 부호화하여야 한다. 본 논문에서는 이러한 문제점을 극복하기 위해서 낮은 표본화 주파수로 오디오 신호를 압축하고, 낮은 표본화 주파수를 사용함으로서 발생하는 대역폭의 손실은 소량의 부가정보를 이용하여 복원해 줌으로써 음질을 향상시키는 알고리즘을 제안한다. 높은 주파수의 스펙트럼을 복원하기 위하여 부호화단에서 낮은 주파수 대역과 다운 샘플링 과정 중에 손실되는 높은 주파수 대역간의 에너지비를 바크밴드에 구한 후 이를 부호화하여 복호화 단으로 전달하고 이를 이용하여 높은 주파수 성분을 복원하는 방법을 제안하였다. 제안된 방법을 이용하면 10%~ 20% 정도의 추가적인 비트를 사용하면서 기존의 방식보다 세그멘탈 신호대 잡음비는 1 dB ~ 3 dB의 성능 개선을 보였으며, 주관적인 MOS 듣기 평가를 수행한 결과 기존의 방식보다 음질이 향상됨을 확인하였다. 또한 본 논문에서 제안한 방법은 주파수 영역에서 압축을 수행하는 모든 오디오 부호화 방식에도 적용이 가능하다.

핵심용어: 지각오디오 부호화, 주파수 복원, 바크밴드, MPEG-2 AAC, 세그멘탈 신호대 잡음비
투고분야: 전기음향 분야 (3.2)

Low-bitrate audio coding enables a number of Internet and mobile multimedia streaming service more efficiently. For the help of next-generation mobile telephone technologies and digital audio/video compression algorithm, we can enjoy the real-time multimedia contents on our mobile devices (cellular phone, PDA, notebook, etc). But the limited available bandwidth of mobile communication network prohibits transmitting high-quality AV contents. In addition, most bandwidth is assigned to transmit video contents. In this paper, we design a novel and simple method for reproducing high frequency components. The spectrum of high frequency components, which are lost by down-sampling, are modeled by the energy rate with low frequency band in Bark scale, and these values are multiplexed with conventional coded bitstream. At the decoder side, the high frequency components are reconstructed by duplicating with low frequency band spectrum at a rate of decoded energy rates. As a result of segmental SNR and MOS test, we convinced that our proposed method enhances the subjective sound quality only 10%~20% additional

책임저자: 서정일 (seoji@etri.re.kr)
305-350 대전 광역시 유성구 가정동 161
한국전자통신연구원 방송미디어연구부
(전화: 042-860-1571; 팩스: 042-860-6465)

bits. In addition, this proposed method can apply all kinds of frequency domain audio compression algorithms, such as MPEG-1/2, AAC, AC-3, and etc.

Keywords: *Perceptual audio coding, Spectrum Reconstruction, Bark band, MPEG-2 AAC, segmental SNR*

ASK subject classification: *Electro acoustics (3.2)*

I. 서 론

지난 수년 동안 인터넷 사용이 급격하게 증가함에 따라 인터넷을 이용한 오디오, 비디오 등의 멀티미디어 데이터 이용이 폭팔적으로 증가하고 있다. 또한 스트리밍 기술을 이용하여 멀티미디어 데이터 전체를 다운로드 받지 않고 실시간으로 제공하는 서비스가 상용화되고 있다. 이러한 기술들에 힘입어 컴퓨터와 인터넷 접속이 가능한 곳에서는 언제, 어디서나 자신이 선택한 오디오, 비디오 데이터를 즐길 수 있게 되었다. 최근에는 CDMA2000-1x나 IMT-2000과 같은 차세대 이동망에서 멀티미디어 데이터를 스트리밍하기 위한 연구들이 활발히 이루어지고 있으며, 2001년에는 바로비젼 등에서 MPEG-4 기술을 이용한 이동통신용 멀티미디어 서비스가 상용화되었다. 그러나 현재의 이동통신망에서는 제한된 대역폭과 사용자 환경의 가변으로 인하여 고음질/고화질의 서비스를 동시에 제공하기는 아직도 해결해야 될 문제점이 남아있다. 보다 넓은 대역폭을 지원하는 차세대 이동통신 환경에서도 실시간으로 고음질/고화질의 멀티미디어 데이터 스트리밍은 여전히 해결해야 할 중요한 문제로 남아있을 것이다.

ISO/IEC 14496 (MPEG-4)는 오디오, 비쥬얼 데이터의 포괄적인 압축 표준안으로서 이동 멀티미디어 스트리밍 서비스에서도 널리 활용될 것으로 전망된다[1]. MPEG-4 오디오 규격에는 CELP, AAC, TwinVQ, BSAC 등의 다양한 압축 알고리즘을 지원한다[2]. CELP는 음성 코더에 적합한 알고리즘이고, TwinVQ와 BSAC는 상용화하기에 많은 제약조건을 가지고 있다. 따라서 MPEG-4에 기반한 이동 멀티미디어 서비스 개발자들은 오디오 압축을 위해서 AAC (Advanced Audio Coding)를 주로 이용하고 있다. AAC를 이용하여 데이터 비트율를 채널당 16 kbps 정도로 유지하면서 만족할만한 음질을 얻기 위해서는, 입력 오디오 신호를 16 kHz 이하의 샘플링 주파수로 샘플링하여 낮은 비트율로 압축하여야 한다. 낮은 샘플링 주파수를 사용할 때 발생하는 고주파수 성분의 손실을 보상하여 실감 음질을 향상시킬 수 있다면, 제한된 대역폭에서 고품질의 오디오 데이터를 서비스할 수 있다. 최근 Coding Technologies사에서는 자사의 Spectral Band Replication (SBR)을 이용하여 낮은 샘플링 주파수를 사용함으로써 발생하는 주파수 손실을 보상하는 알고리즘을 MP3에 적용한 mp3PRO를 발표하였다[3]. 그러나 SBR 알고리즘은 AAC 디코딩 과정과 거의 비슷한 정도의 연산량을 요구하므로 이동 단말기에서는 사용하기 적절하지 않다.

오디오 데이터가 낮은 샘플링 주파수와 비트율로 압축될 경우 대표적으로 다음의 두 가지 문제점이 발생한다[4]. 첫 번째는 고주파수 성분의 손실이다. 샘플링 이론에 따라서 샘플링 주파수의 절반 이상의 주파수 성분은 복원할 수 없다. 즉 16 kHz로 샘플링된 오디오 신호에서는 8 kHz 이상의 고주파 오디오 성분을 들을 수 없다. 이는 음질의 명료도에 심각한 영향을 끼치게 된다. 두 번째는 낮은 비트율로 압축되어 있을 경우 양자화 비트의 부족으로 양자화 (quantization) 오차가 피치 펄스들 사이에서 들리게 되어 압축된 오디오는 울림 (reverberation)을 가지게 된다. 이는 음성같이 피치들이 분명히 드러나는 음원에서는 귀에 거슬리는 잡음을 유발하게 된다.

본 논문에서는 낮은 샘플링 주파수로 샘플링하면서 발생하는 고주파수 성분의 손실을 복원할 수 있는 새로운 고주파 대역 복원 방법을 제안하였다. 또한 주파수 복원에 필요한 연산을 최소화 함으로써 이동통신 단말과 같이 저전력이 필요한 환경에서 효율적으로 적용될 수 있도록 하였다. 복원하고자 하는 고주파수 대역의 수를 줄이면서도 효과적으로 고주파수 성분을 복구하기 위하여 오디오 신호를 바크밴드로 변환한 후 스펙트럼 분석과 복원을 수행하였다[5]. 고주파수 성분의 복원에 필요한 비트스트림을 AAC 비트스트림의 프레임의 마지막 부분에 부가 데이터로 추가하여 다중화함으로써 기존의 방식으로 부호화된 AAC 비트스트림과도 호환성을 유지하였다.

II. MPEG-2 AAC와 SBR

제안된 알고리즘의 실험 및 검증을 위하여 MPEG-2 AAC (Advanced Audio Coding) 압축 알고리즘을 사용하

였다[6]. MPEG-2 AAC 압축 알고리즘은 현존하는 가장 효율적인 오디오 압축 방법의 하나이며 MPEG-2에서 ISO/IEC 13818-7로 국제 규격화되어 있으며, 몇 개의 기능만 추가하여 MPEG-4의 General Audio 코더로 국제 규격화되어 있다. AAC는 탁월한 압축 능력과 음질을 가지고 있으므로 인터넷 오디오 스트리밍, 디지털 오디오 방송, 휴대용 디지털 오디오 재생기 등의 영역에서 이미 상용화되었거나 상용화가 진행 중인 상태이다.

2.1. MPEG-2 AAC

1997년에 표준화된 MPEG-2 AAC는 지금까지 여러가지 검증 과정을 거쳤다. MPEG-2 AAC는 ITU에서 권고하는 채널당 64 kbps의 원본 오디오와 구별할 수 없는 음질을 (indistinguishable quality) 제공하는 것을 목표로 개발되었으며, ITU Recommendation BS.1116에 따라 BBC, NHK, CRC 등의 기관을 통하여 합성음질이 검증되었다[7,8]. 그림 1은 CRC에서 5.1채널 오디오 데이터를 AAC와 MP3, AC-3, PAC 등으로 압축하여 주관적인 테스트를 한 결과를 도표로 나타낸 것이다. 그림에서 나타나는 것과 같이 MPEG-2 AAC를 이용하여 5.1채널 오디오 신호를 320 kbps로 압축하였을 때의 음질이 480 kbps로 압축한 AC-3, MPEG-2 layer-II 보다도 우수한 결과를 나타냄을 확인하였다.

MPEG-2 AAC 알고리즘의 주요 특징은 다음과 같다.

- ▶ 매우 유연한 비트스트림 구조를 가진다. 오디오 데이터의 교환을 위한 ADIF 헤더 구조와 실시간 방송이나 스트리밍을 위한 ADTS 헤더 구조를 제공한다.
- ▶ ATDS 헤더 구조를 이용하면 AAC로 압축된 비트스트림은 블록별로 복호화가 가능하므로 전송에러가 존재하는 환경에 적합하다.
- ▶ 다양한 비트율와 샘플링 주파수를 지원한다.
- ▶ 우수한 압축 능력을 제공한다. 스테레오 오디오에 대하여 96 kbps AAC의 음질은 128 kbps MP3 음질과 유사하거나 우수하다.
- ▶ 적용되는 알고리즘의 복잡도와 기능에 따라서 메인 프로파일, Low-Complexity (LC)O 프로파일, Scalable Sampling Rate (SSR) 프로파일을 제공한다.

한국전자통신연구원(ETRI)에서는 Frauhoffer-IIS 와의 국제 공동연구를 통하여 TMS320C6701 부동소수점 DSP를 이용한 MPEG-2 AAC H/W 코덱을 개발한 바 있다[9]. 또한 MPEG-2 AAC 오디오 부호화기와 복호화기를 이용하여 인터넷을 이용한 MPEG-2 AAC 스트리밍 서비스 시스템을 개발하였고[10], MPEG-2 AAC 오디오 디코더를 범용 고정소수점 DSP를 이용하여 ASIC 칩을 개발하였다[11].

2.2. Spectral Band Replication (SBR)

부호화된 오디오 신호의 청각적 특성을 향상시키기 위한 연구는 예전부터 많이 있어 왔다. 대표적인 것으로 Kenwood의 D.R.I.V.E., pioneer의 Legato-link 그리고

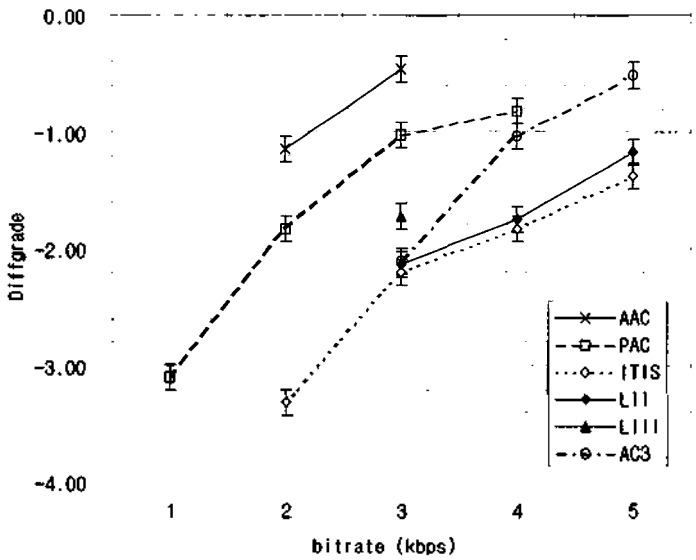

그림 1. MPEG-2 AAC 음질 테스트 결과
Fig. 1. MPEG-2 AAC verification test result.

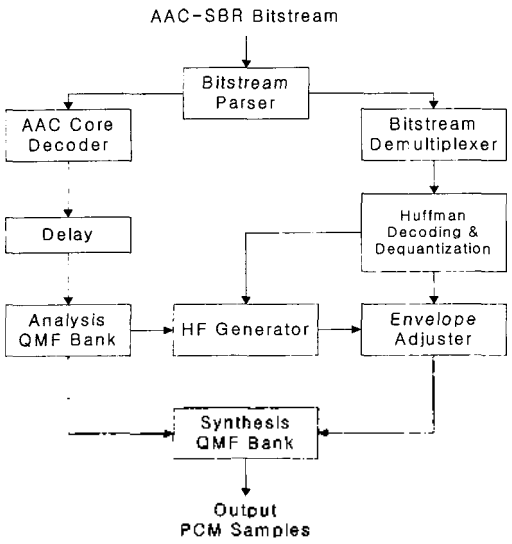

그림 2. AAC-SBR 복호화기의 블록도
Fig. 2. AAC-SBR decoder block diagram.

최근에 발표된 SBR이 있다. 이들 방법의 공통점은 저주파 대역의 특성을 부호화 과정에서 손상된 고주파 대역을 복원하는데 이용하는 것이다. SBR의 특징은 인코더에서 부가정보를 보내고 이를 디코더에서 오디오의 청각적 특성을 향상시키는데 사용하였다는 점이다. SBR 방식의 전형적인 예는 22.05 kHz로 샘플링되고 MP3 부호화기를 이용하여 64 kbps로 압축된 오디오 신호를 44.1 kHz로 샘플링되고 128 kbps로 압축된 오디오 신호와 비슷한 음질을 갖는 것이다.

SBR 알고리즘은 현재 MPEG-4 오디오 Extension에서 BWE (Bandwidth Extension)이란 이름으로 표준화가 진행 중이다[12]. BWE은 MPEG-4 일반 오디오 부호화 방식과 호환성을 유지하면서 대역폭을 늘려주는 방식으로 그림 2는 BWE의 복호화 과정을 나타내는 블록도이다.

그러나 SBR 알고리즘은 복호화 과정에서 추가적인 QMF 해석 및 합성필터가 사용되므로 AAC 복호화 과정과 비슷하거나 더 많은 연산량을 요구하게 된다. 따라서 이동통신 단말과 같이 제한된 연산능력과 저전력의 하드웨어 구조를 요구하는 환경에서 사용하기에는 적절치 않다.

III. MPEG-2 AAC를 이용하여 낮은 샘플링 주파수와 비트율로 압축한 오디오 신호의 음질 분석

MPEG-2 AAC를 이용해서 채널당 16 kbps 이하로 코딩하면서 적당한 음질을 유지하기 위해서는 16 kHz 이하의 샘플링 주파수를 이용하여야만 적절한 음질을 유지할 수 있다. 이와 같이 낮은 샘플링 주파수와 낮은 비트율로 압축된 MPEG-2 AAC 오디오는 음질 면에서 두 가지 큰 문제점을 나타낸다[4].

3.1. 고주파 성분의 손실

샘플링 이론에 의하여 임의의 신호를 fs의 주파수로 샘플링하면 fs/2 이상의 주파수 성분은 복구할 수 없게 된다. 따라서 16 kHz 로 입력 오디오 신호를 샘플링하면 8 kHz 이상의 신호는 복구가 불가능하며, 또한 MPEG-2 AAC 부호화 과정 중에서 양자화 비트수의 제약으로 인하여 상대적으로 음질에 영향을 적게 주는 높은 주파수 대역의 신호에는 비트를 적게 할당하여 추가적인 고주파

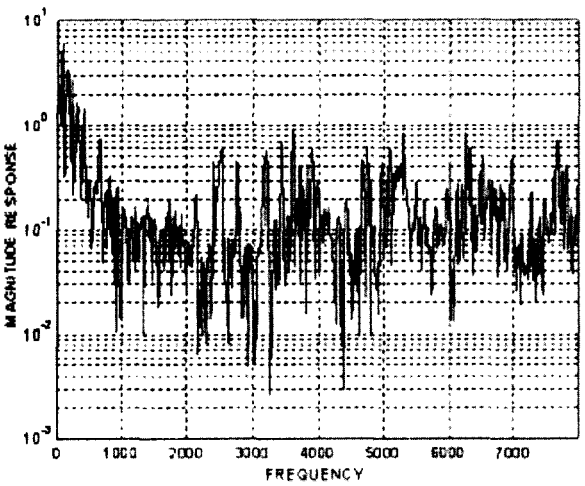

그림 3. 16kHz로 샘플링된 오디오
Fig. 3. The spectrum of 16 kHz sampled audio.

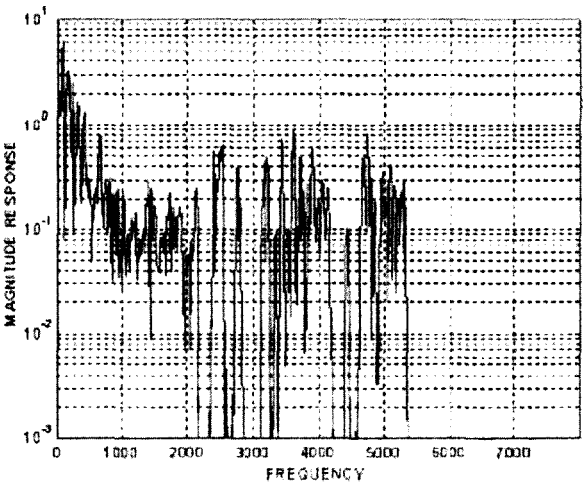

그림 4. 16 kbps로 MPEG-2 AAC를 이용하여 압축된 오디오
Fig. 4. The spectrum of 16 kbps MPEG-2 AAC coded audio.

대역의 손실이 발생한다. 이를 그림으로 설명하면 그림 3은 16 KHz로 샘플링된 오디오 신호이며, 그림 4는 MPEG-2 AAC LC 프로파일 부호화기와 복호화기로 채널당 16 kbps의 비트율로 압축된 오디오의 스펙트럼이다. MPEG- 2 AAC 부호화 과정에서 양자화 비트 수를 줄이기 위해 5 kHz 이상의 오디오 성분들은 거의 제거되었다. 따라서 5 kHz 이상의 오디오 성분을 들을 수 없게 된다. 비록 오디오 신호의 대부분의 정보가 수 kHz 부근에 집중되어 있지만 합성 오디오의 음질의 명료도에 심각한 영향을 끼친다.

3.2. 울림 (reverberation) 효과

오디오 신호를 압축하면 할수록, 양자화 잡음이 커지게 된다. 양자화 잡음이 피치 사이에 들리게 되면 압축된 오디오는 울림 (reverberation)을 가지게 된다. 이러한 울림 효과는 오디오가 음성과 같은 1 kHz 정도의 중간 주파수 대역에서 큰 피크들을 가질 경우 압축된 오디오의 음질을 크게 저하시킨다. 이러한 현상을 줄이려는 목적으로 MPEG-2 AAC 에서는 Temporal Noise Shaping (TNS) 기술을 사용한다[6]. 실제 울림효과는 압축 비트율에 비해 입력 오디오의 샘플링 주파수가 클 경우 심하게 나타난다.

IV. 낮은 비트율로 압축된 오디오에서 고주파 스펙트럼 복원

오디오, 비쥬얼 멀티미디어 데이터를 인터넷이나 이동통신망을 이용하여 전송하거나 스트리밍하려고 할 경우 제한된 대역폭과 전송환경의 가변성 때문에 전송 비트율을 최소화하여야 한다. 그러나 대부분의 대역폭이 상대적으로 데이터량이 많은 비쥬얼 데이터에 할당되고, 오디오에는 상대적으로 적은 대역폭이 할당되게 된다. 따라서 오디오 데이터에 대해서는 더욱 적은 비트율로 압축하여야 한다. 그러나 앞 장에서 살펴본 바와 같이 낮은 비트율로 압축된 오디오는 만족할만한 음질을 가질 수 없다. 본 논문에서는 MPEG-2 AAC로 압축된 오디오의 저주파 대역성분을 미리 구해진 부가정보인 에너지비에 따라 스케일링하여 고주파 대역에 반복하여 고주파 대역 성분의 손실을 줄이는 방법을 제안하였다.

4.1. 바크밴드 스펙트럼

바크밴드 (Bark-band)는 인간의 청각 특성에 따라 24개의 임계 대역 (critical band)으로 나누어지며, 24개의 임계 대역의 경계는 허츠 (Hertz) 단위로 [0, 100, 200, 300, 400, 510, 630, 770, 920, 1080, 1270, 1480, 1720, 2000, 2320, 2700, 3150, 3700, 4400, 5300, 6400, 7700, 9500, 12000, 15500] 라는 것이 실험적으로 알려져 있다 [5]. 바크 스케일은 위 임계 밴드에 따라 주파수를 1~24 bark로 나타낸 것이다. 위에서 알 수 있듯이 고주파 대역에서는 바크밴드의 대역폭이 지수적으로 증가하므로, 고주파 대역을 모델링하거나 처리하는데 저주파대역에 비하여 상대적으로 작은량의 정보를 사용할 수 있다. 부호화기에서 MPEG-2 AAC로 압축된 오디오와 원래의 오디오의 바크밴드 에너지의 비를 구하고 이것을 부가정보로 AAC 비트스트림에 추가하여 전송하게 되며, 복호화기에서는 MPEG-2 AAC로 복호화하고 업샘플링 (up-sampling) 한 후 저주파 대역성분을 고주파 대역에 부가 정보로 주어지는 에너지비로 스케일링하여 반복하여 복원하게 된다. 복원된 스펙트럼을 inverse-FFT를 취하면, 고주파수 성분이 복원된 오디오 신호를 얻게 된다. 그러므로 바크밴드를 이용하면 고주파 신호를 복원하는데 필요한 부가정보의 양을 크게 줄일 수 있었다.

4.2. 제안된 부호화기

바크밴드에서의 MPEG-2 AAC로 부호화된 오디오 신호와 원본 오디오 신호 사이의 에너지 비를 이용하여 고주파수 대역의 신호를 복원하였다. 고주파수 대역을 복원하는데 필요한 부가 정보의 양을 줄이기 위해 바크밴드에서의 에너지비를 이용하였다. 그림 5와 6에 나타난 바

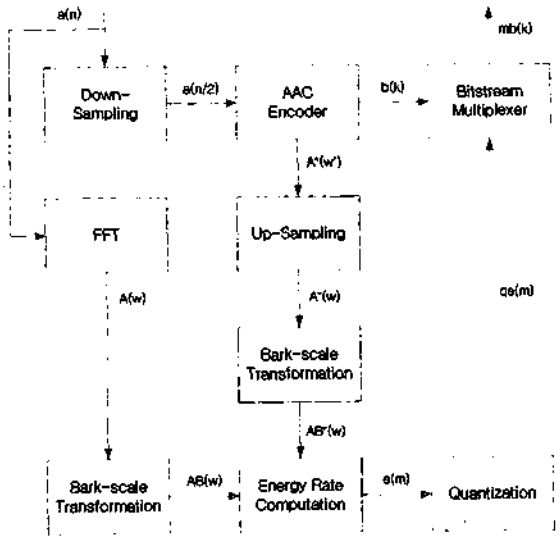

그림 5. 제안된 부호화기의 블럭도
Fig. 5. The block diagram of encoder.

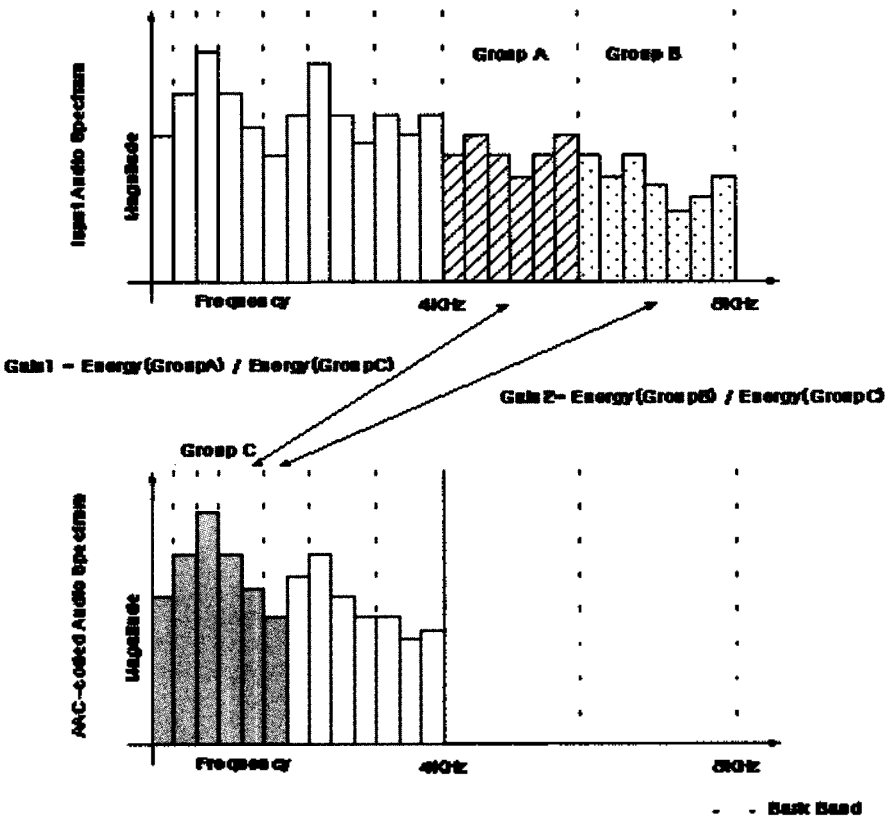

그림 6. 바크밴드에서 에너지비를 구하는 방법
Fig. 6. The diagram of calculating the rate of energy in bark band.

와 같이 제안된 오디오 부호화의 세부 과정은 다음과 같다. 이때 입력 오디오 신호는 16-bit, 16 kHz로 샘플링하였다.

(1) 16-bit, 16 kHz로 샘플링된 입력 오디오 신호를 8 kHz로 다운샘플링한다.
(2) 입력 오디오 신호를 FFT를 이용하여 주파수축으로 변환한 후 바크밴드 스케일로 대역을 분할한다.
(3) 8 kHz로 다운샘플링된 신호를 MPEG-2 AAC 알고리즘을 이용하여 채널당 8 kbps로 압축한다. 이때, 프레임 단위로 MPEG-2 AAC 비트열과 부가비트열을 다중화하기 위하여 ADTS (Audio Data Transport Stream) 형식으로 프레임 헤더를 사용한다.
(4) 압축된 오디오 신호의 스펙트럼을 16 kHz로 업샘플링한 후 바크밴드 스케일로 대역을 분할한다.
(5) 압축된 오디오 신호의 저주파수 대역의 에너지와 입력 오디오 신호의 고주파수 대역의 에너지비를 바크밴드 대역에서 계산한다. MPEG-2 AAC 부호화 과정에서 손실되는 고주파수 대역을 바크밴드에서의 에너지비를 계산하여 이를 복호화 과정에서 복구한다.
(6) (5)의 과정에서 계산된 에너지비를 양자화한 후 MPEG-2 AAC 비트열에 부가비트열로 다중화한 후 최종적인 부호화 비트열을 구성한다.

비교적 짧은 시간인 한 프레임 내에서 오디오 신호의 하모닉 특성이 정상적 (stationary)이라고 생각할 수 있다. 따라서 저주파수 대역의 하모닉 특성을 고주파수 대역으로 적절한 이득값으로 스케일링하면서 복사하면 다운샘플링 과정과 AAC 압축 과정에서 손상된 고주파 신호를 복원할 수 있다. 이와 같이 저주파 대역 신호의 스펙트럼에 데이터량이 적은 에너지비를 부가정보로 이용하여 고주파 신호를 복원하는 것은 부호화 비트율를 줄이면서도 오디오 음질을 크게 개선하는 효과가 있다.

4.3. 제안된 복호화기

ADTS 헤더 포맷을 지원하는 MPEG-2 AAC 복호화기는 먼저 ADTS 헤더 내의 동기신호를 참고하여 헤더의 시작점을 찾고, 헤더 내의 프레임 길이를 읽어서 헤더의 전체 길이를 찾게 된다. 제안된 복호화기는 MPEG-2 AAC 부호화 프레임의 마지막에 위치하는 다중화 비트스트림을 읽

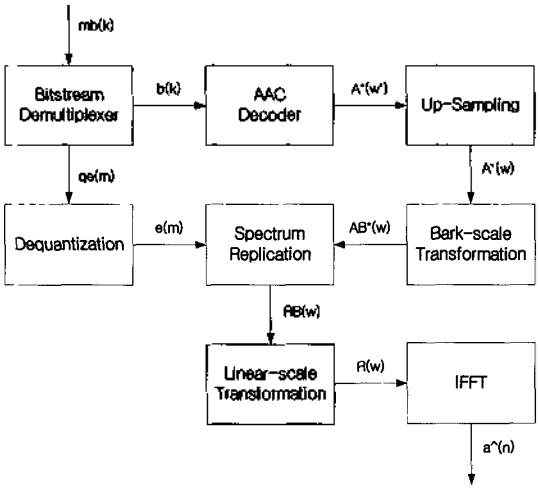

그림 7. 제안된 복호화기의 블럭도
Fig. 7. The block diagram of decoder.

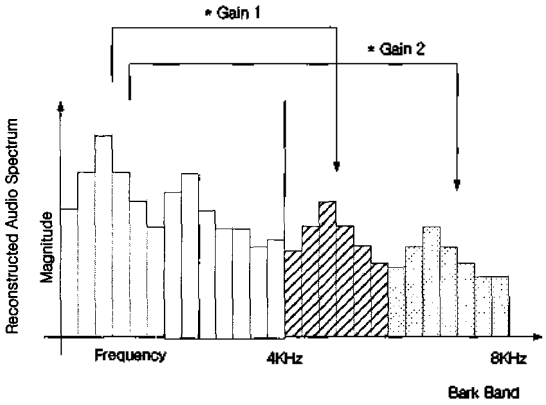

그림 8. 에너지비를 이용한 고주파수 복원방법
Fig. 8. The replication procedure in the decoder.

어서 고주파수 복원시에 이용하게 되고, 일반적인 MPEG-2 AAC 복호화기는 ADTS 헤더 정보로부터 전체 프레임을 복호화한 후와 동기신호를 찾는 과정에서 부가비트열을 무시하게 된다. 그림 7과 8에 나타난 바와 같이 제안된 오디오 복호화기의 세부 과정은 다음과 같다.

(1) 입력 부호화 오디오 비트열 *mb (k)*를 MPEG-2 AAC 부호화 비트열 *b (k)*와 에너지비 비트열 *qe (m)*으로 역다중화한다.
(2) MPEG-2 AAC 부호화 비트을 복호화하여 복호화 스펙트럼을 구한 후 16 kHz로 업샘플링한다.
(3) (2)의 과정에서 계산된 복호화 스펙트럼을 바크 스케일로 변환한다.
(4) 에너지비 비트열을 이용해 그림 8에 나타난 바와 같이 복호화 스펙트럼의 저주파 대역과 에너지비를 이용하여 고주파 신호를 복원한다.
(5) 크기 (magnitude)값은 상기에서 설명한 방법으로 모델링 하지만 위상 (phase)값은 별도의 모델링 과정을 거치지 않는다. 오디오 신호에서의 위상은 음질에 큰 영향을 끼치지 않으므로 저주파수 대역의 위상을 고주파수 대역까지 확장하여 전체주파수 대역의 위상을 얻었다.
(6) (5)의 과정에서 계산된 주파수 신호를 역퓨리어 변환을 이용하여 시간축 신호로 변환하여 고주파수 신호가 복원된 오디오 신호를 얻는다.

제안된 고주파수 복원 알고리즘은 MPEG-2 AAC 복호화 과정에서 얻어지는 주파수 스펙트럼 성분을 이용하여 스펙트럼을 구하고 이를 간단한 스케일링 과정을 통하여 고주파수 성분을 복원한 후 역퓨리어 변환을 이용하여 최종적인 출력 오디오 신호를 생성하는 간단한 방법이다. 따라서 제안된 주파수 복원 알고리즘은 MPEG-2 AAC 복호화 알고리즘과 비슷하거나 다소 많은 연산량을 요구하게 되므로 저전력을 필요로 하는 이동통신 단말과 같은 환경에서 적용하기에 적절하다.

V. 실험결과

제안된 알고리즘의 성능을 검증하기 위해, 다양한 오디오 신호에 대해서 정량적/정성적인 실험을 수행하였다.

5.1. 스펙트럼 분석

먼저 제안된 방법의 주파수 성분 복원 능력을 확인하기 위하여 입력 오디오 신호와 제안된 방법으로 복원한 오디오 신호의 스펙트럼들을 그림 9에서 12까지 나타내었다. 그림 9는 16 kHz로 샘플링된 입력 오디오 신호의 스펙트럼이다. 그림 10은 입력 오디오 신호를 8 kHz로 다운샘플링하고, 8 kbps의 비트율로 MPEG-2 AAC로 압축한 오디오 신호의 스펙트럼이다. 그림 10에서 나타난 바와 같이 4 kHz 이상의 성분은 다운샘플링 과정에서 상실되며, MPEG-2 AAC 부호화 과정에서 3.5 KHz 이상의 스펙트럼 성분이 상쇄됨을 알 수 있다. 그림 11은 제안된 방법으로 저주파 신호를 이용하여 고주파 신호를 복원한 오디오 신호의 스펙트럼이다. 그림 9와 그림 11을 비교하면 저주

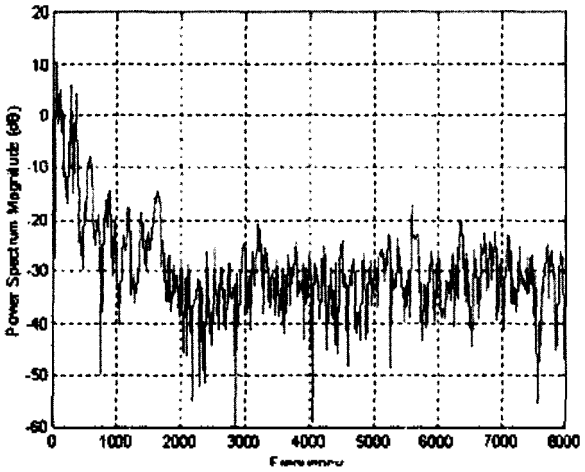

그림 9. 16 kHz로 샘플링된 원래의 오디오의 스펙트럼
Fig. 9. The spectrum of 16 kHz sampled audio.

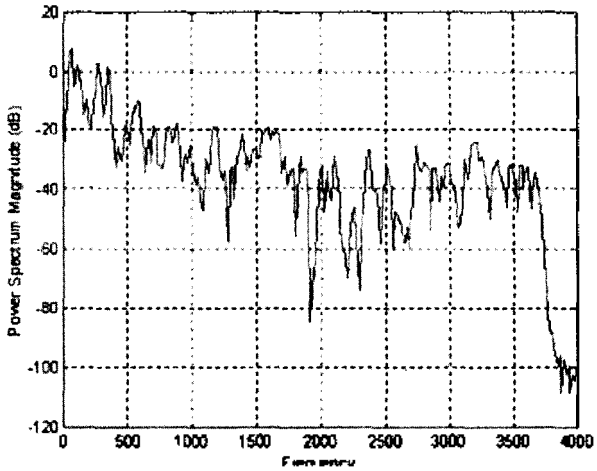

그림 10. 8kbps MPEG-2 AAC 압축된 오디오의 스펙트럼
Fig. 10. The spectrum of 8 kbps MPEG-2 AAC coded audio.

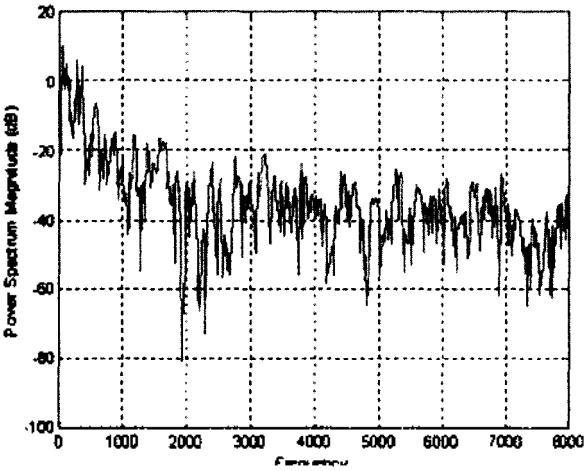

그림 11. 제안된 방법으로 음질을 개선한 오디오의 스펙트럼
Fig. 11. The spectrum of enhanced AAC-coded audio.

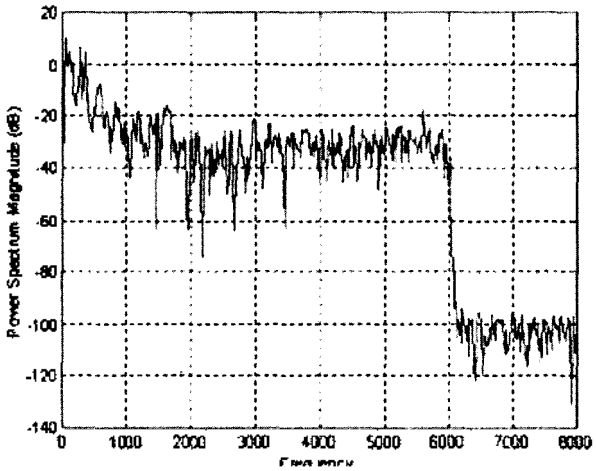

그림 12. 16 kbps AAC로 압축된 오디오
Fig. 12. The spectrum of 16 kbps AAC-coded audio.

파수 성분은 MPEG-2 AAC 부호화 과정에서 복원되고, 고주파수 성분의 전체적인 하모닉 특성은 저주파수 성분을 스케일링하면서 복사함으로써 8 kHz까지 고주파수 성분이 복원됨을 확인할 수 있다. 이는 출력 오디오 신호의 명료도를 크게 향상시키는 결과를 나타내었다. 그림 12는 원래의 오디오를 다운샘플링하지 않고 16 kbps의 비트율로 AAC압축한 오디오 신호의 스펙트럼이다. 그림에서 알 수 있듯이 MPEG-2 AAC 부호화 과정 중에서 6 kHz 이상의 고주파 성분은 상실됨을 알 수 있다. 따라서 제안된 방법으로 고주파 성분을 복원하는 방법의 주파수 복원 능력이 우수함을 확인할 수 있다.

5.2. 세그멘탈 신호대 잡음비와 MOS 테스트

세그멘탈 (segmental) 신호대 잡음비 [13, pp. 586]를 이용하여 제안된 방법이 얼마나 성능을 개선시키는지 정량적으로 알아보았다. 이 경우 대부분의 오디오에서 8 kbps의 비트율로 AAC로 압축한 오디오에 비해 0.5~1 dB 정도 향상되었다. 이에 비해 16 kbps의 비트율로 AAC로 압축한 오디오 신호의 경우 3 dB 정도 세그멘탈 신호대 잡음비 면에서 8 kbps의 비트율로 AAC로 압축한 오디오에 비해 우수하므로 제안된 방법이 세그멘탈 신호대 잡음비 면에서는 그다지 좋은 결과를 보여주지 못한다. 그러나 세크멘탈 신호대 잡음비는 객관적 음질 향상 척도로서 사람이 직접 듣는 결과와는 다를 수 있다. 따라서 주관적 듣기 평가를 위해서 MOS (Mean Opinion Score) 테스트를 수행하였다. 10명의 시험자를 이용하여 1점에서 5점까지의 점수로 음질을 평가하였다.

표 1에 나타난 바와 같이 듣기 평가 결과에서는 8 kbps AAC 압축하고 제안된 방법으로 부가정보를 이용해 음질을 개선한 오디오와 16 kbps AAC로 압축한 오디오가 비슷

표 1. MOS 테스트 결과
Table 1. The result of subjective audio quality test.

Audio Source	MOS
16 KHz로 샘플링된 입력 오디오	4.5
8 KHz 샘플링, 8 kbps/ch MPEG-2 AAC	3.72
8 KHz 샘플링, 8 kbps/ch MPEG-2 AAC + 고주파수 성분 복원	3.97
16 KHz 샘플링, 16 kbps/ch MPEG-2 AAC	4.03

한 음질을 보였다. 따라서 부가정보의 비트 레이트를 고려하더라도 수 kbps 정도의 이득을 얻을 수 있음을 알 수 있다.

VI. 결 론

MPEG-2 AAC 압축 과정에서 손실되는 고주파 성분을 바크밴드에서의 에너지비를 이용하여 복원하는 방법을 제안하였다. MPEG-2 AAC 비트열과 부가 비트열을 다중화할 때 MPEG-2 AAC 비트열의 각 프레임의 마지막에 부가 비트열을 삽입하여 ADTS 헤더 포맷을 지원하는 MPEG-2 AAC 복호화기에도 복호화가 가능하도록 하였다. 제안된 방법의 성능을 검증하기 위해 세그멘탈 신호대 잡음비를 계산하여 약 0.5~1 dB 정도 향상됨을 확인하였고, 10명의 시험자를 이용한 MOS 테스트를 수행한 결과 주관적인 음질이 향상됨을 확인하였다. 또한 고주파수 대역의 복원에 필요한 부가적인 연산량을 최소화함으로써 이동통신 단말과 같이 저전력을 요구하는 환경에서도 적용이 가능하도록 하였다.

낮은 비트율로 압축된 오디오에 크게 나타나는 두 가지 큰 문제점인 고주파 대역 손실, 울림 현상 중 제안된 방법은 고주파 대역 손실에는 큰 효과가 있었으나 울림 현상은 제거할 수 없었다. 앞으로 오디오 소리의 울림 현상을 제거하기 위한 추가적인 연구가 필요하다.

감사의 글

본 논문은 과학기술부 지원 "디지털 오디오 방송기술 연구" 과제 연구비에 의해 수행되었습니다.

참 고 문 헌

1. ISO/IEC JTC1/SC29/WG11 (MPEG), International Standard ISO/IEC 14496-1, Generic Coding of Audio-Visual Object: System.
2. ISO/IEC JTC1/SC29/WG11 (MPEG), International Standard ISO/IEC 14496-3, Generic Coding of Audio-Visual Object: Audio.
3. O. Kunz, "SBR Explained: White Paper," www.coding-technologies.com, 2001.
4. S. Levine, *Audio Representations for Data Compression and Compressed Domain Processing*, Ph. D. thesis, Stanford University, 1999.
5. J. O. Smith and J. S. Abel, "Bark and ERB Bilinear Transforms," *IEEE Trans. on Speech and Audio Processing*, 7 (6), 1999.
6. ISO/IEC JTC1/SC29/WG11 (MPEG), International Standard ISO/IEC 13818-7, Generic Coding of Moving Pictures and Associated Audio Information: Advanced Audio Coding.
7. D. Meares, K. Watanabe, and E. Scheirer, "Report on the MPEG-2 AAC Stereo Verification Tests," ISO/IEC JTC1/SC29/WG11 N2006, Feb. 1998.
8. ISO/IEC JTC1/SC29/WG11 N1420, Overview of the Report on the Formal Subjective Listening Tests of MPEG-2 AAC Multichannel Audio Coding, Marceio, Brazil, Nov., 1996.
9. D.-Y. Jang, K.-O. Kang, J.-S. Kwak, S.-H. Kim, J.-W. Hong, S. Geyersberger, W. Fiesel, H. Gerenhard, and D. Huhn, "A Multichannel Audio Codec System for Multichannel Audio Authoring," 106th AES Convention, May, 1999.
10. J.-I. Seo, T.-J. Lee, J.-W. Seok and J.-W. Hong, "Internet Audio Streaming Servie Technology Integrated with Copyright protection," 111th AES Convention, Nov., 2001.
11. 서정일, 장대영, 홍진우, "MP3/AAC 복합복호화기를 이용한 오디오 플레이어의 구현," 2001년도 추계 한국음향학회 학술발표대회, 2001.
12. ISO/IEC JTC1/SC29/WG11 M7947, "Technical description of Coding Technologies' Proposal for MPEG-4 v3 General Audio Bandwidth Extension: pectral Bandwidth Replication (SBR)," Jeju-Island, Korea, March, 2002.
13. J. R. Deller, J. G. Proakis and J. H. Hansen, *Discrete-Time Processing of Speech Signals*, Macmillan Publishing, 1993.

저자 약력

● 서 정 일 (Jeong-Il Seo)

1994년 2월: 경북대학교 전자공학과 졸업 (공학사)
1996년 2월: 경북대학교 대학원 전자공학과 졸업 (공학석사)
1996년 3월~현재: 경북대학교 대학원 전자공학과 박사과정
1998년 3월~2000년 11월: 현대전자 (연구원)
2000년 11월~현재: 한국전자통신연구원 방송미디어연구부 (연구원)

● 서 진 수 (Jin-Soo Seo)

1998년 2월: 한국과학기술원 전자전산학과 졸업 (공학사)
2000년 2월: 한국과학기술원 대학원 전자전산학과 졸업 (공학석사)
2000년 3월~현재: 한국과학기술원 대학원 전자전산학과 박사과정

● 홍 진 우 (Jin-Woo Hong)

1982년 2월: 광운대학교 응용전자공학과 (공학사)
1984년 2월: 광운대학교 대학원 전자공학과 (공학석사)
1993년 8월: 광운대학교 대학원 전자공학과 (공학박사)
1998년~1999년: 독일 프라운호퍼 연구소 (교환연구원)
1984년 3월~현재: 한국전자통신연구원 방송컨텐츠보호연구팀장 (책임연구원)
2000년 1월~현재: 한국음향학회 홍보이사, 뉴미디어음향학술분과 위원장, 한국방송공학회 편집위원

● 강 경 옥 (Kyeoung-Ok Kang)

1985년 2월: 부산대학교 물리학과 졸업 (이학사)
1988년 2월: 부산대학교 대학원 물리학과 (이학석사)
1993년 3월~현재: 한국항공대학교 대학원 전자공학과 박사과정
1991년~현재: 한국전자통신연구원 방송미디어연구부 방송컨텐츠응용연구팀장 (선임연구원)