

상용 CDDA 와 하위 호환성을 가지는 고해상도 부호화방식의 제안

A Proposal for High-Resolution Encoding System with Backward Compatibility in CDDA

문동욱*, 김낙교**, 남문현**
(Dong-Wook Moon, Lark-Kyo Kim, Moon-Hyun Nam)

Abstract - Conventional CDDA (Compact Disc Digital Audio) system has limitations come from sampling frequency and quantization bit, 44.1kHz and 16 bit respectively. So, new medium is developed for high-resolution audio recording, like as DVD-audio etc. But CDDA is a widely used medium for high fidelity audio yet, because new medium has complexity and difficulty in manufacturing system. In this paper, we design a new encoding system for high-resolution audio signal. The system is backward compatible with conventional CDDA. By evaluation for encoding and decoding process, we describe practicability of our proposal system.

Key Words : Audio Coding, Spectrum Reconstruction, High Resolution Audio, High Fidelity Audio

1. 서론

선형 (LPCM; Linear Pulse Code Modulation) 방식에 의한 기존의 디지털 오디오 시스템은 표본화 주파수와 양자화 비트 수에 의해 신호 대역과 해상도가 결정되는 한계를 가지고 있다. 82 년에 등장한 CD 기반의 디지털 오디오 시스템은 44.1kHz 의 샘플링 주파수와 16 비트의 양자화로 인해, 기록 재생 가능한 신호가 20kHz 및 98.1dB 의 다이내믹 레인지 (Dynamic Range)로 제한되고 있다. 고전적인 이론에서 가청 주파수대역인 20kHz 까지를 기록 재생하기 위해서는 44.1kHz 의 표본화주파수로 충분한 것이 사실이지만, CD 가 상용화된 이후 이 이론은 논란의 대상이 되어왔다[1].

90 년대 이후, 저장매체의 폭발적인 집적도 향상으로 DVD (Digital Versatile Disc) 등 고용량 매체를 등장하면서 논란의 대상이 되었던 오디오 신호의 표본화 문제는 96/192kHz 24 비트 등 고해상도 포맷을 수용하는 것으로 해결되었다. CDDA 를 대체할 오디오 저장 매체로는, DVD 포럼 진영의 DVD-audio 와 Sony/Philips 진영의 SACD (Super Audio CD)가 유력하지만, 두 매체가 서로 경쟁하면서 시장에서는 기대만큼의 성공을 거두지 못하고 있는 것이 현실이다. 두 매체 모두 오소링 (Authoring) 절차로 인한 제조 공정상의 복잡성 및 관련 장비의 설계상의 어려움 등으로 인해, 등장한지 오랜 시간이 지났음에도 아직 음반업계에서 조차 어느 매체의 성공도 낙담하지 못하고 있다. 때문에, CDDA 포맷은 여전히 44.1kHz, 16 비트 표본화에서 벗어나지 못하고 있음에도 불구하고 아직도 가장 대중적인 오디오 저장 매체로 사용되고 있는 것이다.

본 논문에서는 기존의 CD 기반의 디지털 오디오 시스템과 호환성을 유지하며 88.2kHz 의 표본화 주파수 및 24 비트의 양자화를 갖는 새로운 부호화 방식을 소개함으로써 기존 시스템에서 고해상도의 오디오 신호를 재생할 수 있는 방법을 제안하고자 한다. 또한 이에 상응하는 복호화 과정을 통해 본 논문에서 제안하는 방식의 유용성을 검증하였다.

2. 새로운 고해상도 부호화 방식의 제안

2.1. CD 의 한계에 대한 개선 시도

CD의 한계를 벗어나고자 하는 움직임이 없었던 것은 아니다. CD 의 한계에 대한 논의는 90 년대가 들어오면서부터 각 오디오 메이커마다 보다 향상된 해상도의 기술을 소개하게 된다. 일본의 파이오니어(Pioneer) 의 레가토링크 (legato link) 는 CD 에 수록된 20kHz 까지의 신호만으로 20 kHz 이상의 신호를 합성해내고자 하였으며, 와디아 (Wadia) 의 디지마스터 (Digi-Master) 등은 시간축 상의 곡선 보간 기술로 누락된 신호 정보를 추정해내려고 하였다.

1996 년 퍼시픽 마이크로소닉스 (Pacific Microsonics) 에서 발표한 HDCD (High Definition Compatible Digital) 역시 이러한 CDDA 포맷의 한계를 넘어서고자 시도된 방식이다. HDCD 는 88.2kHz 및 24 비트로 레코딩된 음원을 44.1kHz 16 비트로 저표본화하는 데시메이션 (Decimation) 필터의 과정에서 부수적인 다이내믹 프로세싱을 행한 후, 이를 부호화하여 부가적으로 기록하고, 복호화 과정에서는 다시 이에 상응하는 과표본화 처리를 행하는 방법을 통해 고해상도 음원을 기록 재생하였다. 또한, 부가적인 제어 신호는 LPCM 의 LSB 영역에 기록함으로써 기존 CDDA 와 호환성을 유지할 수 있었다.

그러나, HDCD는 가청 주파수 영역에서의 신호 크기에 대

저자 소개

* 正 會 員 : 建國大學 電氣工學科 博士課程

** 正 會 員 : 建國大學 電氣工學科 正教授 · 工博

한 다이나믹 프로세싱과, 저표본화 및 과표본화 과정에서의 과도 영역의 필터링 특성을 개선하고자 하는 것에 초점이 맞추어진 방식으로 고해상도 음원이 그대로 부호화/복호화 과정을 거친다고는 보기에 무리가 있는 방식이다. 그럼에도 불구하고, HDCD 는 소비자들에게 CDDA 의 한계를 넘어서는 최초의 상용 매체로 주목을 받았으며, 적절한 마케팅과 기술 지원으로 고음질 CD 로서는 가장 크게 성공한 케이스가 되고 있다[2].

2.2 제안하는 부호화 방식

본 논문에서 제안하고자하는 고해상도 부호화 방식은 88.2kHz 및 24 비트로 레코딩된 음원을 44.1kHz 로 저표본화하는 과정에서 손실되는 음원에 대해 별도의 오디오 부호화기를 거쳐 저비트율로 압축하여 이를 기록한 후, 재생 시에는 복호화하여, 88.2kHz 및 24 비트의 음원으로 재생하고자하는 것이 기본 개념이다. 기존 CDDA 와 호환성을 유지하기 위해 HDCD 와 유사하게 LPCM 영역에서 LSB 및 LSB+1 영역의 2 비트를 압축 정보 영역으로 할당하고, MSB부터 MSB-13 까지의 14 비트는 기존 44.1kHz 표본화 주파수를 갖는 LPCM 영역으로 사용하고자 한다. 이러한 구조는 기존의 CD 플레이어에서 14 비트 LPCM 영역을 재생할 수 있을 뿐만 아니라, SPDIF (Sony-Philips Digital-audio Interface Format) 출력으로 신호의 전송이 가능하므로 외부기기와의 연결에 유연할 수 있게 된다. 물론, 본 논문에서 제안하는 복호화기를 장착한 기기에서는 LSB 와 LSB+1 영역에 기록된 압축 오디오 신호를 함께 재생함으로써, 88.2kHz 및 24 비트의 고해상도 음질을 재생할 수 있게 된다. 부호화 및 복호화 알고리즘은 그림 1, 2 (후면 참조) 와 같다.

2.2.1 부호화 과정

입력된 88.2kHz 및 24 비트 오디오 신호는 저역통과필터 및 데시메이션 필터를 통해 44.1kHz, 14 비트의 LPCM 신호로 변환된다. 이 과정에서 손실된 오디오 신호는 LPCM 신호를 다시 과표본화와 저역통과필터를 통해 88.2kHz 및 24 비트의 신호로 변환하여 원 신호와 (-) 연산을 함으로써 구할 수 있게 된다. 이에 따라 구해진 신호는 오디오 부호화기를 통해 채널당 88.2kbps 이하의 전송률을 가지는 압축신호로 변환된 후 LPCM 신호와 함께 기록된다.

2.2.2 복호화 과정

CDDA에 기록된 16 비트 신호로부터 MSB에서 MSB-13 까지의 LPCM 신호와 LSB, LSB+1 에 기록된 압축 신호를 분리해낸다. LPCM 신호는 과표본화와 저역통과필터를 통해 88.2kHz 및 24 비트 신호로 변환되고, 압축 신호 역시 복호화 과정을 통해 88.2kHz 및 24 비트 신호로 변환되어진다. 두 신호는 연산기를 통해 합쳐져 최종 출력되어진다.

2.2.3 오디오 부호화 과정

본 논문에서 사용되어지는 오디오 부호화 과정은 다음과 같은 조건들을 필요로 한다. 우선, 88.2kHz 및 24 비트 표본화를 지원해야하며, 특히 20kHz부터 40kHz 까지의 대역에 대해 88.2kbps 이하의 저전송률에서도 신뢰성 높은 부호화 및 복호화가 행해질 수 있어야 한다. 그리고 기존 CD 플레이

어의 FF/FR (Fast Forward/Fast Reverse) 및 랜덤 액세스 (Random Access) 기능과 호환될 수 있는 프레임 구조로 이루어져야하며, 부호화 및 복호화 과정에 과도한 프로세서 부하를 가지지 않는 효율적인 방식이 필요하다. 본 논문에서 이에 적합한 오디오 부호화 방식으로 MPEG2 에서 사용되어지고 있는 AAC (Advanced Audio Coding) 부호화기를 이용하였으며, 위의 조건에 부합하도록 AAC 사양의 일부를 변형하여 적용하였다.

MPEG2 에서 처음 제안된 AAC 는 MP3 와 같은 MPEG1 에서의 오디오 압축과는 달리 8 에서 96 kHz 의 표본화 주파수를 가질 수 있으며, 독립적인 부호화 방식을 적용하기 때문에 상대적으로 안정된 음질을 보여주는 것으로 평가되는 등 오디오 신호의 손실압축에 있어 압축율과 재생 품질에 있어 가장 효과적인 방법으로 알려지고 있다[3].

본 논문에서 사용한 부호화 방식은 AAC 를 기본 골격으로 하여, 앞서 언급한 필요조건을 만족하도록 일부 항목을 변형시킨 방식을 사용하였다. 특히 CD 의 프레임 처리와 호환될 수 있도록 프레임 당 샘플 수는 588 샘플에 비례하는 147, 294, 588, 1176 샘플을 취했으며, 88.2kbps 이하의 저전송률에서도 40kHz 까지의 신호가 부호화 될 수 있도록 필터 배열을 수정하였다.

3. 결과 고찰

본 논문에서 제안하는 방법을 검증하고자, 상기의 부호화 과정 및 복호화 과정을 PC 기반의 레코딩 시스템에서 구현하였다. 검증을 위해 88.2kHz 및 24 비트로 표본화된 음원을 사용하여, 부호화 및 복호화 과정을 처리한 후 스펙트럼 분석 결과를 비교하여 성능 평가를 하였다.

사용된 음원은 체스키 레코드 (Chesky Records) 에서 발매한 Chuck Mangione 의 「The Felling's Back」 에 수록된 'Quase'를 표본으로 사용하였다. 이 음원은 원래 DVD 매체로 발매된 것이라, 96kHz, 24 비트 음원이지만, 88.2kHz, 24 비트 음원으로 재표본화를 거친 후 검토에 사용하였다.

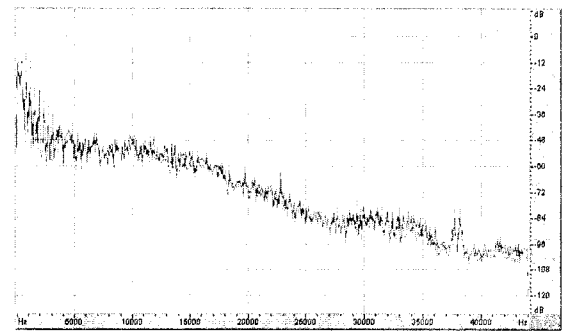


그림 3. 원 음원의 스펙트럼 (88.2 kHz 24 비트)

그림 3 은 88.2kHz, 24 비트의 원음원의 스펙트럼이다. 44kHz 근방까지 음원이 넓게 분포되어있는 것을 알 수 있다. 그림 4 는 본 논문에서 제안하는 방식으로 부호화를 거친후, 일반 CD 플레이어에서 재생할 시의 FFT 스펙트럼이다. 20kHz 이상의 신호가 손실되는 것을 알 수 있다. 그림 5

는 복호화한 과정을 거친 최종 출력 신호의 FFT 스펙트럼이다. 그림 3 과 비교하면, 비록 40kHz 이상의 신호가 손실된 것을 확인할 수 있지만 40kHz 까지의 신호는 매우 충실하게 재생되는 것을 확인할 수 있다.

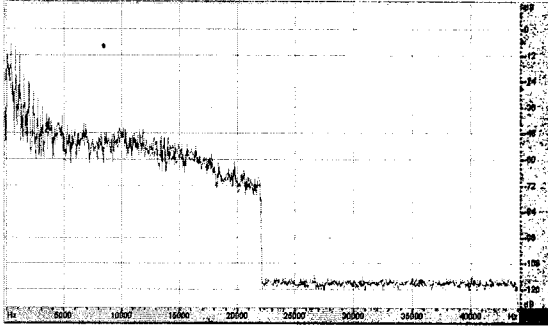


그림 4. 부호화하여 일반 CDP 에서 재생한 스펙트럼

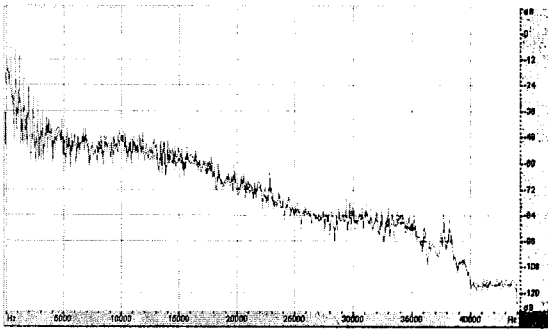


그림 5. 복호화 과정을 거친 후의 스펙트럼

4.

본 논문에서는 기존의 CDDA 포맷과 하위 호환성을 유지 하면서 고해상도 오디오 신호를 기록하는 시스템을 제안하였다. 또한 제안하는 방식에 대한 구현을 통해 그 유용성을 검증하였다. 그러나, 본 제안이 실용화가 되기 위해서는 보다 포괄적인 검증이 필요하다. 특히 LPCM 영역의 2 비트를 압축 데이터로 이용하는 만큼 손실되는 SNR (Signal to Noise Ratio) 을 보상할 수 있는 방법과 압축 부호화 방법에 대한 효율성을 향상시킬 수 있는 방법이 추가적으로 연구되어야 한다. 또한 실제 청각상의 효율성을 검증할 수 있는 피시험자를 대상으로 한 MOS (Mean Opinion Score) 등의 검토도 필요로 한다.

참 고 문 헌

[1] T. Oohashi, E. Nishina, N. Kawai, Y. Fuwamoto and H. Imai, "High-frequency sound above the audible range affects brain electric activity and sound perception", AES 91st Convention, Preprint 3207, October 1991
 [2] 문동욱, "다시 생각하는 오디오상식-주파수 영역에서 시간 영역까지 HDCD(3)", 월간 오디오, 제 314 호, pp. 150-153, 1998. 4.
 [3] 서정일, 서진수, 홍진우, 강경옥 "저비트율로 압축된 오디오의 음질 개선 방법", 한국음향학회지, 6 호 제 21 권, pp.566-575, 2002.

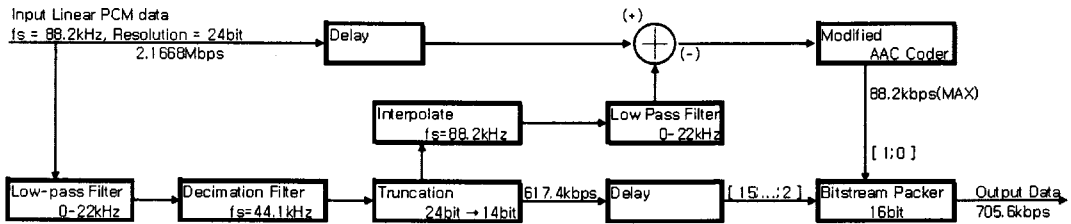


그림 1. 제안하는 고해상도 부호화 알고리즘

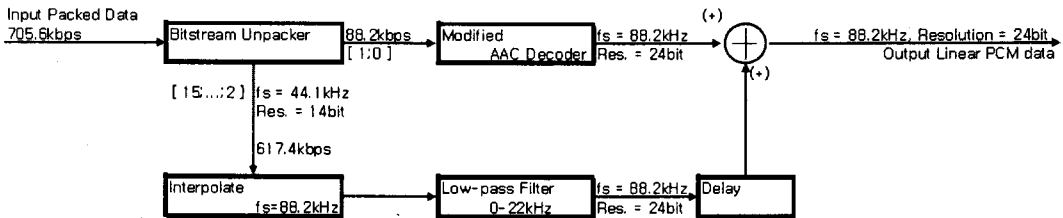


그림 2. 제안하는 고해상도 복호화 알고리즘