

고음질 오디오를 위한 디지털 오디오 워터마킹에 관한 연구

°김 정 태, 구 대 성, 이 강 현

조선대학교 전자정보통신공학부, 멀티미디어 ASIC 설계 실험실

Tel : (062) 230-7066/Fax : (062) 233-1120

E-mail : {space,orion,khrhee}@vlsi.chosun.ac.kr

A Study on the Audio watermarking for High Quality Digital Audio

°Jung-Tae KIM, Dae-Sung KU, Kang-Hyeon RHEE

School of Electronics and Info-Comm. Engineering, Chosun University

Abstract

In this paper, we proposed the high quality digital audio watermarking algorithm in the frequency domain.

The spread spectrum technique is used to encrypted a stream of information by spreading the data as much of the frequency spectrum as possible. It's technique adapt well to data hiding in audio signal.

We have used the perceptual model and MDCT/IMDCT for the high quality digital audio watermarking. The proposed watermark algorithm preserved high quality audio data from watermark signal.

I. 서 론

오늘날 디지털 세계에서, 정보에 대한 가치가 매우 크다. 그 정보라는 것은 인터넷의 발전으로 다양한 형태로 액세스 될 수 있어서 사용자가 데이터를 복사, 변조하여 저작권이나 소유권을 주장하는 것이 가능하다. 현재 디지털 저작물에 대한 저작권을 확인하는 방법으로 워터마크(watermark)가 쓰이고 있으며, 영상뿐만 아니라 디지털 오디오 분야에서도 연구가 활발하다. 그림 1은 일반적인 워터마크 삽입과정을 나타낸 것이다.

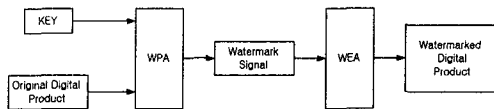


그림 1. 워터마크 삽입 과정

보통 음성 데이터에서는 샘플링 이론에 근거하여 음성 데이터를 표본화하고, 양자화해서 디지털화 되며, 특히 오디오 데이터는 44.1Khz 샘플링 율, 16bit 양자화, PCM으로 디지털화 된다. 이 방법에 의해서 인간의 가청 주파수 대역을 모두 커버할 수 있기 때문에, 매우 높은 음질을 보존한 채 디지털화 할 수 있다. 한편, 이 디지털 데이터는 쉽게 완전한 형식으로 복제할 수 있는 특징이 있다. 이 충실한 재현성은 디지털화의 큰 장점이 되는 반면, 디지털 저작물의 저작권 보호가 필수 조건으로 되어 왔다. Boney는 사람이 지각할 수 없는 형식으로 저작권 정보를 오디오 데이터에 삽입하는 방법을 사용하였다.[1] 청각의 마스킹 현상을 이용한 Boney의 디지털 워터마크의 삽입 방법은 음의 진폭성분을 직접 이용해서 삽입하기 때문에 고품질의 오디오 데이터에는 부적절하고 동시에 워터마크의 은닉성에 문제가 된다. 그래서, 이 논문에서는 고품질을 요구하는 오디오 데이터에 워터마크를 삽입하는 적절한 방법으로서 주파수 성분에 저작권 정보를 삽입하는 방법을 사용한다.[2][3]

II. 오디오 워터마크

오디오 워터마크 방법에는 가장 기본적인 Low-bit

coding과 Phase coding, Spread spectrum, Echo hiding 등의 방법이 있다. 본 논문에서는 삽입된 워터마크의 은닉성을 확보하기 위하여 Spread spectrum 방법을 사용하고, 오디오 데이터의 고음질을 보장하기 위하여 MPEG 오디오에 채택된 Perceptual model과 MDCT/IMDCT(Modified Discrete Cosine Transform/Inverse Modified Discrete Cosine Transform)를 사용한다.[4-6] 그림 2는 오디오 워터마킹을 위한 전체적인 블록도이다.

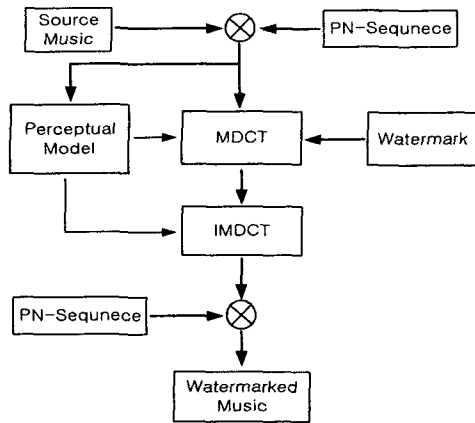


그림 2. 오디오 워터마크 삽입 블록도

2.1 워터마크 삽입

그림 2에 의한 워터마크 삽입 처리 과정은 다음 순서를 따른다.

- 가. PN-sequence
- 나. Psychoacoustic model
- 다. 오디오 신호의 주파수 변환(MDCT)
- 라. 워터마크 삽입
- 마. 워터마크된 신호 시간영역 변환(IMDCT)
- 바. 워터마크된 오디오 데이터

디지털 워터마크의 은닉성을 확보하기 위하여 Spread spectrum 방식의 한 방법인 직접 스펙트럼 확산(DSSS : Direct Sequence Spread Spectrum)을 이용한다. DSSS에는 PN-sequence가 필요하기 때문에 이 PN-Sequence의 생성기를 디지털 워터마크의 고유키로 정한다. PN-sequence와 오디오 데이터를 곱하여 확산된 오디오 신호는 Perceptual model과 MDCT의 입력 값이 되어 변환된다. 그림 3은 본 논문에서 사용된 10개의 레지스터를 갖고 3, 10번 출력 레벨에서 변환이 이루어지는 PN-sequence이다.

Perceptual model에서는 입력 신호의 마스킹 쓰레스

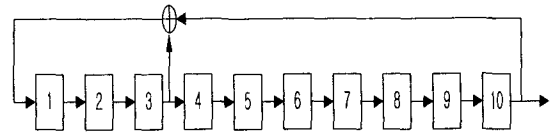


그림 3. PN-sequence

홀드를 계산하고, 마스킹 쓰레스홀드는 Hanning 윈도우에 비중을 두어 연속적인 블록에 50% 중첩을 위하여 2048 또는 512 샘플 오디오 세그먼트를 위해 입력된 오디오 데이터를 분석하여 MDCT/IMDCT를 위해 윈도우 정보를 제공한다. 이 윈도우 정보를 가지고 MDCT/IMDCT는 주파수 변환을 한다. 그림 4는 MDCT/IMDCT를 위한 윈도우 형태이다.

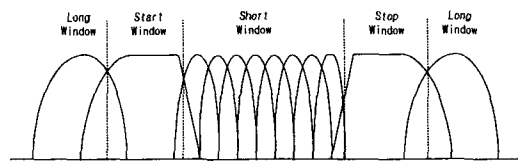


그림 4. MDCT 분석 윈도우

MDCT/IMDCT는 각각 식 (1), 식 (2)과 같다

$$X(i, k) = 2 \cdot \sum_{n=0}^{N/2-1} x(i, n) \cos\left(\frac{2\pi}{N}(n+n_0)\left(k+\frac{1}{2}\right)\right)$$

for $0 \leq n < N/2$

$$\begin{cases} x_{in} = \text{spectral coefficient} \\ n = \text{샘플 인덱스} \\ k = \text{cosine계수인덱스} \\ i = \text{블록 인덱스} \\ N = \text{윈도 길이} \\ n_0 = \left(\frac{N}{2} + 1\right) \frac{1}{2} \end{cases} \dots (1)$$

$$X(i, n) = \frac{2}{N} \sum_{k=0}^{N/2-1} x(i, k) \cos\left(\frac{2\pi}{N}(n+n_0)\left(k+\frac{1}{2}\right)\right)$$

for $0 \leq n < N$

$$\begin{cases} x_{in} = \text{윈도된 입력} \\ n = \text{샘플 인덱스} \\ i = \text{윈도 인덱스} \\ k = \text{spectral계수인덱스} \\ N = \text{윈도 길이} \\ n_0 = \left(\frac{N}{2} + 1\right) \frac{1}{2} \end{cases} \dots (2)$$

MDCT/IMDCT는 블록연산이며 따라서 블록간의 연결을 부드럽게 하기 위해 윈도를 사용한다. 변환을 위한 합성윈도의 길이 N 은 2비트로 표현되는 window_sequence에 의해 결정된다. 표 1은 window_sequence에 따른 윈도 길이 N 이다.

표 1. Window sequence

| | |
|-------|--------------------------------|
| $N =$ | 2048, ONLY_LONG_SEQUENCE(0x0) |
| | 2048, LONG_START_SEQUENCE(0x1) |
| | 256, EIGHT_SHORT_SEQUENCE(0x2) |
| | 2048, LONG_STOP_SEQUENCE(0x3) |

표 1에서 EIGHT_SHORT_SEQUENCE만 제외하고 모두 2048개의 샘플을 사용하는 LONG_BLOCK이다, 따라서 window_sequence=0x3일 때만 SHOT_BLOCK으로 256개의 샘플을 블록으로 처리하고 나머지는 모두 LONG_BLOCK으로 처리한다.

본 논문에서의 MDCT/IMDCT는 KBD(Kaiser-Bessel derived)윈도와 sine 윈도 2가지를 사용하기 때문에 윈도 길이 뿐만 아니라 윈도형태도 결정해야 한다. 이것은 현재 블록의 WINDOW_SHAPE와 과거 window shape인 WINDOW_SHAPE_PREVIOUS_BLOCK으로 결정한다. 그리고 윈도의 전체 길이를 1/2하여 전반부(left)는 WINDOW_SHAPE으로 결정된 윈도형태를 사용하고, 후반부(right)는 WINDOW_SHAPE_PREVIOUS_BLOCK에 의해 결정된 윈도형태를 사용한다. WINDOW_SHAPE / WINDOW_SHAPE_PREVIOUS_BLOCK=0x1이면 KBD window을 사용한다.

이렇게 주파수 변환 영역에 집적 워터마크를 삽입하게 되고, 다시 복원후 PN-sequence를 곱하면 음질의 열화 없는 워터마크된 오디오 데이터를 얻을 수 있다.

III. 오디오 워터마크 추출

워터마크 삽입량 M 은 오디오 표본화 주파수 f_s 와, 샘플수 N 을 이용해서 식 (3)과 같이 구해진다.

$$M = \frac{f_s}{N} \dots \dots (3)$$

따라서, 음질 열화에 비해서, 각 프레임에 대해 복수 비트를 삽입한 것이 쉽게 가능하고 삽입량은 각 프레임마다 삽입 비트수에 비례한다.

오디오 워터마크의 추출과정은 워터마크된 오디오 데이터에 워터마크 삽입 과정에서 사용된 것과 동일한 모듈인 PN-sequence를 곱하여 동기시킨후 다시 한번 MDCT를 하여 주파수 영역에서 워터마크를 추출한다. 이 과정은 그림 5와 같다.

만약 워터마크 삽입단계와 다른 PN-sequence를 사용하면 올바른 워터마크를 검출할 수 없을 뿐만 아니라 오디오 음질 자체가 열화된다.

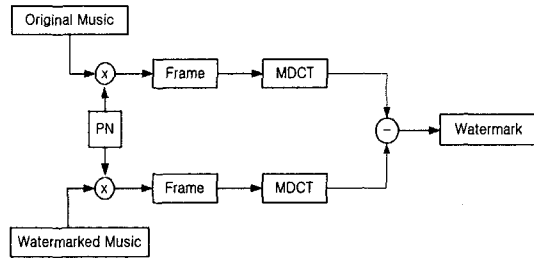


그림 5. 워터마크 추출 과정

IV. 실험 결과

고음질에 가치가 있는 오디오 데이터에 워터마크를 삽입할 경우, 워터마크에 의해 음질이 떨어지는 않는 것이 매우 중요하고, 실제로 고품질의 오디오 데이터에 삽입을 해서 음질이 미치는 영향을 알아본다.

실험 데이터는 44.1Khz 16비트이고, 양자화 한 것이며, 오디오 데이터는 항상 스테레오 음이기 때문에 그 한쪽 성분에만 워터마크를 삽입하였다.

그림 6은 실험s에 사용된 캐스터네츠 오디오 데이터

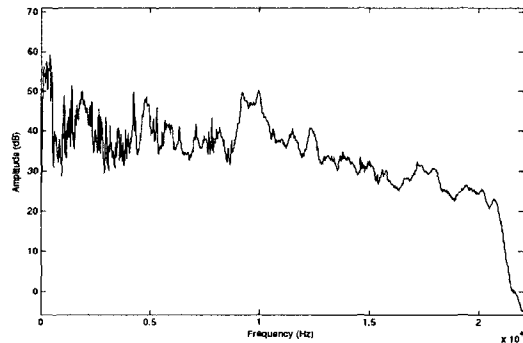


그림 6. 캐스터네츠 오디오 데이터 파형

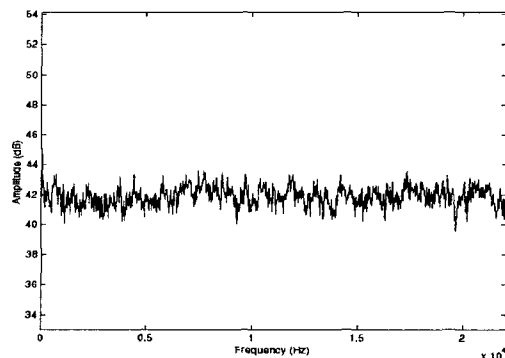


그림 7. PN-sequence를 곱한 데이터의 주파수 영역

를 주파수 영역에서 파형을 나타낸 것이고, 그림 7은 PN-sequence를 곱한후 오디오 데이터의 주파수 영역에서 파형을 나타낸 그림이다. 그림 7에서 알 수 있듯이 PN-sequence를 곱하여 주파수 변환된 원 데이터는 전 주파수 영역에 확산됨을 알 수 있다.

그림 8은 이렇게 확산된 주파수 영역에 워터마크를 삽입하고 IMDCT로 복원한후 워터마크를 추출한 결과이다.

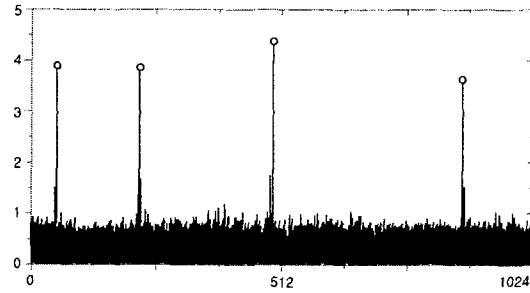


그림 8. 워터마크 검출

그림 9는 워터마크된 오디오 데이터를 MPEG-1 Layer 3 압축후 다시 복원한 다음 워터마크를 검출한 것이다.

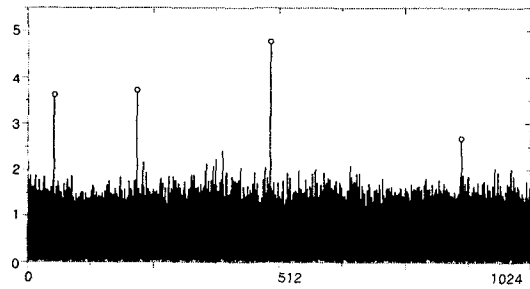


그림 9. MPEG-1 Layer 3 압축에 대한 워터마크 의 추출

V. 결 론

본 논문에서는 고음질을 위한 오디오 워터마킹 방법을 제안하였는데, 입력 데이터 값인 워터마크를위한 오디오 파일은 샘플링율 8Khz~96Khz, 비트 율 64Kbps~256Kbps, 스테레오 및 모노를 지원하고, 주파수 영역에 직접 워터마크를 삽입하는 방법이다. 실험 결과 PN-sequence 동기가 이루어지지 않는 오디오 데이터에서 워터마크를 검출하기위해 공격하면 워터마크를 검출할 수 없을 뿐만아니라 오디오 데이터의 음질에도 심각한 영향이 있음을 알 수 있다.

참고문헌

- [1] L. Boney, A.H. Tewfik, & K.N. Hamdy, " Digital watermark for audio signals", Proc IEEE Int. Conf. on Multimedia Computing and Systems, June 1996.
- [2] Ingemar J. Cox, Joe Kilian, Tom Leighton and Talal Shamoan "Secure Spread Spectrum Watermarking for Multimedia", IEEE Trans. on Image Processing , 6, 12, 1997
- [3] V. Basia and I. Pitas, "Robust audio watermarking in the time-domain", Proc. Europ. Signal Processing Conf., Sept 1998
- [4] Jack Lacy, Schuyler R. Quackenbush, Amy R. Reibman and James H. Snyder, " Intellectual property protection systems and digital watermarking" OPTICS EXPRESS. 7 Dec. 1998/Vol. 3, No. 12
- [5] "ISO/IEC MPEG-2 Advanced Audio Coding 4382(N-1)"-Presented at the 101st Convention 1996 November 8-11 Los Angeles, California, AN AUDIO ENGINEERING SOCIETY PREPRINT, 1996..
- [6] Brandenburg K, Stoll G, "ISO-MPEG-1 Audio: A Generic Standard for Coding of High Quality Digital Audio, Journal of the Audio Engineering Society", No. 10, Oct 1994, pp. 780-792.