

심리음향모델 기반의 디지털 오디오 워터마킹

송유수* · 김종환* · 신경욱*

*금오공과대학교

Digital Audio Watermarking Based on Psychoacoustic Model

You-Su Song* · Jong-Hwan Kim* · Kyung-Wook Shin*

*School of electronic Eng., Kumoh National Institute of Technology

E-mail : cain96@kumoh.ac.kr

요 약

본 논문에서는 대역확산 기반 오디오 워터마킹 삽입에 대한 연구를 하였다. 최근 디지털 저작권을 확인하는데 디지털 워터마크가 사용되고, 영상, 문서뿐만 아니라 디지털 오디오 분야에서도 연구가 이루어지고 있다. MP3(MPEG Audio Layer-III)에 사용하는 심리음향모델을 이용함으로써 음질에 영향을 주지 않고 지각 할 수 없는 디지털 저작권을 삽입 할 수 있다. 본 논문에서는 대역확산 기반 오디오 워터마킹을 위한 심리음향모델을 분석하고 Matlab으로 모델링하여 시뮬레이션 하였으며, 이를 대역확산 기반 오디오 워터마크 삽입에 적용하였다.

ABSTRACT

This paper describes a study on the digital watermarking algorithm which is used to confirm the copyright protection of digital audio data. The digital audio watermarking algorithm based on psychoacoustic model is used for the inaudibility of the watermark data. The psychoacoustic model which is a key algorithm in MP3 audio compression is analyzed by MATLAB simulation, and is applied to digital audio watermark insertion.

키워드

대역확산 워터마킹, 심리음향모델, MP3, 디지털 오디오

I. 서 론

최근 인터넷 보급과 디지털 멀티미디어 기술의 발달에 의해 영상, 음향, 비디오, 출판물 등의 지적재산권 정보가 네트워크상에서 다양한 형태로 액세스 되고, 무단 복조 및 변조가 가능하게 됨에 따라 디지털 콘텐츠의 지적재산권 보호가 사회적으로 중요한 이슈로 대두되고 있다. 디지털 저작물의 저작권을 보호하기 위한 방법은 크게 나누어 콘텐츠를 제3자가 알아볼 수 없도록 암호화 기술과 저작권을 확증해 주는 기술로 구분된다. 저작권 보호 기술은 멀티미디어 콘텐츠에 저작권자만이 아는 정보를 사람의 육안이나 귀로 구별할 수 없게 변형하여 삽입하고, 불법 복제나 유통 시에 이를 추출/검지하여 저작권을 입증하는 방법을 일컫는다. 최근, MP3 플레이어의 보편화에 따라 음악 파일에 대한 저작권 문제가 대두됨에 따라 디지털 오디오 워터마킹(watermarking) 저작권 보호 기법에 대한 연구가 활발히 이루어지고 있다.

본 논문에서는 디지털 오디오 워터마킹 시스템을 개발하기 위한 기초연구로서 사람이 지각할 수 없는 워터마킹 삽입방법을 고안하기 위해 MP3 압축에 사용되고 있는 심리음향모델에 대해 분석하였다.

디지털 오디오 데이터의 저작권 보호를 위한 방안으로 오디오 워터마킹 기법에 대한 연구가 활발하게 이루어지고 있다. 대표적인 오디오 워터마킹 방법으로는 위상 부호화(phase coding) 기법, 반향은폐(echo hiding) 기법, 대역확산(spread spectrum) 기법 등이 발표되고 있다.[1]

오디오 데이터는 크기와 위상의 두 가지 성분으로 구성되며, 인간의 귀는 신호의 크기에는 민감하지만 위상의 변화에는 둔감하다는 성질을 이용하는 방법이 위상 부호화 기법[2]이며, 이 방법은 정보량이 제한적이고 신호처리 과정에서 쉽게 손상되는 단점이 있다. 반향은폐 기법[3]을 이용한 워터마킹은 귀에 들리지 않을 만큼 작은 반향을 오디오 신호에 첨가함으로써 워터마크 정보를

삽입하는 방법이다. 워터마크 강인성을 향상시키기 위하여 충분히 큰 반향 신호를 삽입하면 왜곡을 초래할 수 있고, 복호화기의 복잡도가 높다는 단점을 갖는다. 대역확산 기반의 오디오 워터마킹 기법[4,5]은 PN (Pseudo-random Noise) 시퀀스에 의해 대역확산된 워터마크 정보를 심리음향모델에 의해 변형시켜 오디오 신호에 삽입하는 방법이다. 이 방법의 특징은 심리음향모델을 이용하여 삽입된 워터마크 데이터가 인간의 귀에 들리지 않도록 하는 기술이다.

II. 대역확산 기반 오디오 워터마킹

그림 1은 대역확산 기반 오디오 워터마킹의 부호화 과정을 보인 것이다.[4,5] PN 시퀀스의 자기상관 특성을 이용하여 워터마크 신호와 PN 시퀀스를 곱하여 워터마크 데이터를 대역확산시킨 후, FFT 연산을 통해 주파수 영역으로 변환한다. 한편, 입력 버퍼에 저장된 오디오 데이터는 일정한 블록 크기로 나뉘어 FFT 연산을 통해 주파수 영역으로 변환되며, 심리음향모델을 거쳐 masking threshold 값이 추출된다. 대역확산된 워터마크 데이터는 masking threshold와 곱해지는 변형과정을 거친 후 IFFT를 거쳐 시간영역으로 역변환되고 원래 오디오 데이터와 합해짐으로써 워터마크의 삽입이 이루어진다.

본 논문에서 선택한 워터마크 삽입 방법은 PN 시퀀스에 의해 워터마크 데이터를 대역확산 시킴으로써 PN 시퀀스의 자기상관 특성에 의해 삽입된 워터마크 정보의 추출이 가능하며, 또한 심리음향모델에 의한 변형을 통해 삽입된 워터마크 정보가 인간의 청각에 들리지 않게 하는 강인한 워터마킹 방법이다.

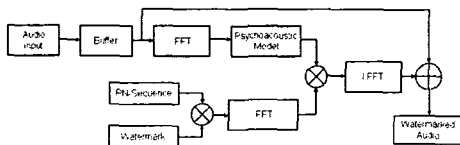


그림 1. 오디오 워터마킹 부호화기

그림 2는 오디오 워터마크 복호화 과정을 보인 것이며, 복호화에서는 워터마크 삽입 시에 사용된 PN 시퀀스가 검출 키(key)로 사용된다. 검출 시에 원 신호를 사용하게 되면 적은 연산량을 갖는 복호화기만으로도 신뢰도가 매우 높은 복호화가 가능하지만, 실시간 인증형 워터마킹과 같이 원 신호를 사용하는 것이 불가능한 대부분의 시스템에는 적합하지 않다.

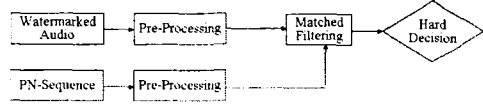


그림 2. 오디오 워터마킹 복호화기

III. 오디오 워터마킹을 위한 심리음향모델

1. 심리음향모델

MPEG 오디오 코딩에서 오디오 신호의 압축 시에 음질을 저하시키지 않기 위해 사용되는 심리음향모델은 최소가청한계(absolute hearing threshold)와 사운드 마스킹(sound masking) 특성을 이용한다. 그림 1의 디지털 오디오 워터마킹 방식에서도 인간의 청각에 들리지 않도록 워터마크 정보를 삽입하기 위해 심리음향모델을 이용한다.

최소가청한계는 잡음이 없는 환경에서 사람의 청각이 감지할 수 있는 음의 최소 한계를 나타내며, 고요할 때의 최소가청한계는 음의 주파수에 따라 다르며, 최소가청한계 보다 작은 음은 들을 수 없다.

사운드 마스킹은 강력한 신호와 약한 신호가 동시에 발생하였을 때 약한 신호를 들을 수 없게 되는 현상을 말한다. 사운드 마스킹에는 주파수 마스킹(frequency masking)과 순시적 마스킹이 있는데 일반적으로 심리음향모델에서는 주파수 영역에서의 마스킹을 모델링한다.[6]

주파수 마스킹은 오디오 신호의 주파수 영역에서의 사운드 마스킹 현상을 말하며, 이는 주파수 영역에서 인접한 두 오디오 신호 중 큰 음압을 갖는 주파수 신호는 작은 음압의 주파수 신호가 들리지 않도록 만드는 현상이다. 이때 큰 음압을 갖는 주파수 성분을 마스커(masker)라고 하며, 작은 음압을 갖는 주파수 성분을 마스kee(maskee)라 한다. 마스커의 masking threshold는 마스커와 마스kee의 주파수에 의해 결정된다.

2. 심리음향모델 분석

본 절에서는 MPEG ISO 문서[6]에서 제안된 심리음향모델을 분석하고, 이를 디지털 오디오 워터마크의 삽입에 적용하였다. 본 논문에서는 보다 정교한 masking threshold의 추출이 가능하여 MPEG Audio Layer-III에서 사용되는 심리음향모델 II에 대해 분석하였다.

그림 3은 심리음향모델 II의 처리 과정을 보인 것이다[6]. 버퍼에 저장된 오디오 데이터는 1024 샘플 단위로 Hann window 연산을 거친 후, 윈도우 크기에 따라 1024-point 또는 256-point FFT 연산에 의해 주파수 영역으로 변환된다.

FFT에 의해 주파수 영역으로 변환된 오디오 신호

호의 크기성분(r_w)과 위상성분(f_w)을 이용하여 식 1에 의해 정의되는 unpredictability measure(C_w)가 계산된다. 계산된 unpredictability measure는 신호의 tonality를 예측하기 위해 사용된다.

$$C_w = \frac{((r_w \cos f_w - \hat{r}_w \cos \hat{f}_w)^2 + (r_w \sin f_w - \hat{r}_w \sin \hat{f}_w)^2)^{0.5}}{r_w + \text{abs}(\hat{r}_w)} \quad (1)$$

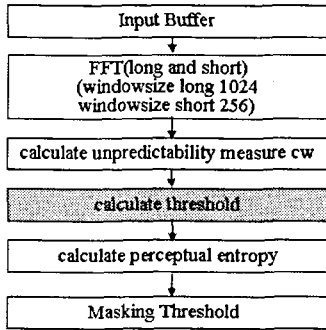


그림 3. 심리음향모델 II의 흐름도

식 1에 의해 계산된 unpredictability measure는 threshold 계산을 도우며, 각 threshold에 대해 TMN(Tone Masking Noise) 효과와 NMT(Noise Masking Tone) 효과를 weighting한다. 한편, 이 과정에서 구해지는 tonality는 신호(tonal)와 잡음(non-tonal)을 구분하는 기준이 된다. 구해진 tonality는 tonal한 성분만큼 TMN을 weighting하고, non-tonal한 만큼 NMT를 weighting하여 식 2와 같이 SNR이 계산되고, threshold 값은 구해진 SNR에 의해 식 3과 같이 계산된다.

$$SNR = \text{MAX}[\text{minval}(b), TMN(b), \text{tbb}(b) + NMT(b), (1 - \text{tbb}(b))] \quad (2)$$

$$\text{nbb}(b) = \text{ecbb}(b) \times \text{norm}(b) \times 10^{-\frac{SNR(b)}{10}} \quad (3)$$

최종 masking threshold 값은 현재의 masking threshold, 이전 블록의 masking threshold, 두 블록 이전의 masking threshold 그리고 최소가칭한 제의 4가지 값들 중 최대값으로 결정된다.

앞에서 설명된 심리음향모델을 Matlab으로 모델링하여 시뮬레이션을 하였으며, 그림 4는 시뮬레이션에 사용된 PCM 오디오 데이터 1024 샘플의 파형이다. 그림 5은 오디오 샘플을 FFT 변환한 후 식 1에 의해 unpredictability measure를 계산한 결과 파형이다. 그림 5의 unpredictability measure를 이용하여 tonality를 구하면 그림 6와 같이 되며, 파형에서 1에 가까운 샘플이 tonal하다는 것을 나타낸다. 그림 7은 최종으로 구해진 masking threshold 파형을 보인 것이다.

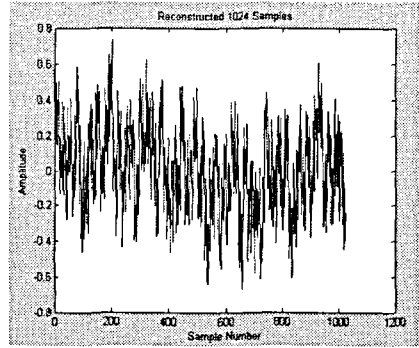


그림 4. 입력 1024 샘플의 파형

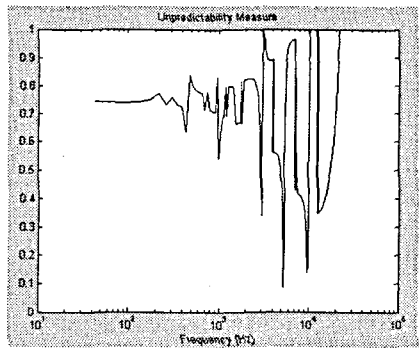


그림 5. Unpredictability measure 파형

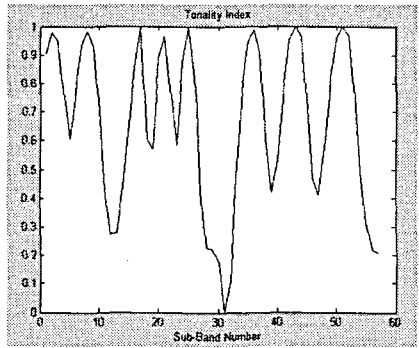


그림 6. Tonality 파형

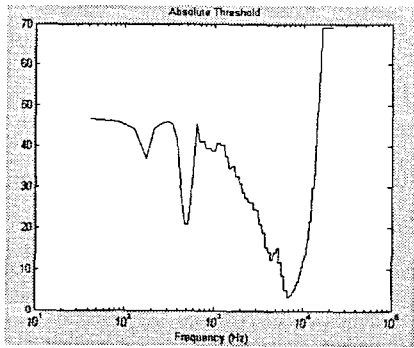


그림 7. Masking threshold 파형

3. 워터마크 삽입

앞에서 분석된 심리음향모델을 그림 1의 대역 확산 기반 오디오 워터마킹 부호화에 적용하기 위하여 워터마크 삽입과정을 Matlab으로 모델링하고 시뮬레이션 하였으며, 그림 8는 그림 7의 masking threshold를 이용하여 PN 시퀀스로 대역확산된 워터마크 데이터를 변형한 결과 파형이다. 그림 9에서 위의 파형은 원래 오디오 신호이며, 아래 파형은 워터마크가 삽입된 오디오 신호의 파형이다.

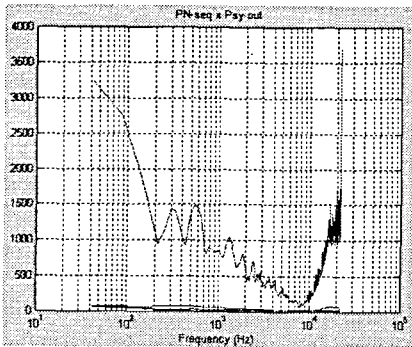


그림 8. Masking threshold로 변형된 워터마크

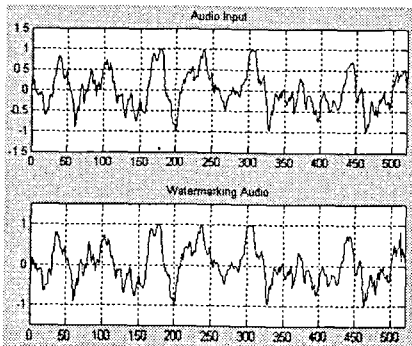


그림 9. 워터마크가 삽입된 오디오 파형

IV. 결 론

본 논문에서는 디지털 오디오 워터마킹에 사용되는 심리음향모델을 분석하고 Matlab으로 모델링하여 시뮬레이션 하였으며, 이를 대역확산 기반 오디오 워터마크 삽입에 적용하였다. 향후 이를 적용한 디지털 오디오 워터마킹 시스템을 구현할 예정이다.

참고문헌

- [1] 오현오, 윤대희, 석종원, 홍진우 "방송 콘텐츠 보호를 위한 디지털 오디오 워터마킹" 한국방송공학회논문지, 1226-7953, 제6권1호, pp.3-12, 2001
- [2] C. Xu, J. Wu, Q. Sun and K Xin, "Applications of digital watermarking technology in audio signals", J. Audio Eng. Soc., vol. 47, no. 10, 1999.
- [3] W. Bender and D. Gruhl, "Techniques for data hiding", IBM Sys. Journal, vol.35, no.3&4, 1996.
- [4] I. J. Cox, J. Kilian, T. Leighton and T. Shamoon, "Secure spread spectrum watermarking for images, audio and video", ICIP'96, vol.III, 1996.
- [5] L. Boney, A. Tewfik and K. Hamdy(1996.), "Digital Watermarks for Audio Signals." Processdings of the Third IEEE International Conference on Multimedia Computing and Systems.
- [6] ISO/IEC JTC1/SC29/WG11, "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s-CD 11172-3(Part3. MPEG-Audio)", no. 71.

※ 2004년도 IT SoC 핵심설계인력양성 사업의 SoC 전공실습프로젝트 지원에 의한 연구결과의 일부임.
 ※ 반도체설계교육센터(IDECE)의 CAD Tool 지원에 감사드립니다.