

# 스프레드 스펙트럼 기반 디지털 오디오 워터마킹 기법 연구

진 장 윤, 최 창 렬, 정 재 창  
한양대학교 전자통신공학과  
전화 : 02-2292-6457 / 핸드폰 : 016-707-1401

## Digital Audio Watermarking Based on Spread Spectrum Techniques

Chang-yune Jin, Changryoul Choi, Jechang Jeong  
Department of Electronic Communications Engineering, Hanyang University  
E-mail : geoffrey@ece.hanyang.ac.kr

### Abstract

In this paper, we propose a robust audio watermarking method. The proposed watermarking algorithm is composed of a psychoacoustic model to achieve perceptual transparency and spread spectrum technique to embed watermark. The watermark is embedded in each audio frame by adding a perceptually-shaped pseudo-random sequence. We demonstrate the robustness of the watermarking algorithm.

### I. 서론

최근 몇 년간 디지털 기술의 발전으로 멀티미디어 정보가 급증하고 있고 네트워크의 발전에 의해 멀티미디어 정보들을 손쉽게 얻을 수 있게 되었다. 특히 디지털 오디오 데이터는 MP3, AAC 등의 기술로 압축되어 인터넷을 통하여 간단히 전송할 수 있게 되어 CD 수준의 고품질 압축 오디오 데이터는 네트워크 상에서의 AOD (Audio On Demand) 서비스를 가능하게 하였다. 그러나 압축된 오디오 데이터는 디지털 데이터의 특성상 복제 및 복사가 쉬워 네트워크 상에서 불법적으로 유통되는 사례가 빈번하게 발생한다. 이러한 불법 복제를 막기 위한 장치가 필요한데, 워터마킹(wa-

-termarking)은 오디오뿐만 아니라, 문서, 그림, 동영상 등의 지적 재산권 보호 대상 성격을 지니는 자료에 대해 원 데이터에 권리자의 정보나 배포자의 정보 및 인증과 같은 추가적인 정보를 삽입하여 데이터에 대한 지적재산을 보호하기 위한 기법이다. 즉, 워터마킹은 저작권 보호(Copyright protection) 기술의 일종으로 멀티미디어 데이터에 소유주만이 아는 신호를 사람의 육안이나 귀로는 구별할 수 없게 삽입하고 불법 복제 유통 시 이를 추출해서 워터마크의 유무를 판별하여 저작권을 입증할 수 있는 기술이다. 또한 콘텐츠 제공(Content provider) 서버로부터 사용자에게 디지털 콘텐츠 데이터를 제3자가 알아볼 수 없도록 암호화하는 Front-end 기술과는 달리 차후에 저작권을 증명해주는 Back-end 기술이다.

본 논문에서는 기존에 제안된 디지털 오디오 워터마킹 알고리즘을 분석하고, 확산 대역(Spread spectrum) 기법과 심리음향모델(Psychoacoustic model)을 적용한 워터마크 삽입 알고리즘을 제안한다. 제안한 워터마크 삽입 알고리즘을 구현하여, 디지털 오디오에 워터마크(watermark)를 삽입한 후 MP3 압축, Band-pass 필터링, 반향(echo) 삽입, 등화기(Equalization) 공격에 대한 워터마크의 검출 특성을 실험 및 검증하였다.

## II. 기존 오디오 워터마킹 기법들

### 2.1 하위 비트 조작 기법(Low-bit coding)[4]

하위 비트 조작 기법은 가장 단순한 워터마킹 삽입 방법으로 각 샘플링 지점에 이진정보를 추가하고 삭제하여 워터마크를 삽입하는 방법으로 데이터를 많이 삽입할 수는 있으나 워터마크의 삭제가 쉽게 일어나 강인하지 못한 약점을 지닌 방법이다.

### 2.2 위상 부호화 기법(Phase coding)[5]

디지털 오디오 데이터는 크기(magnitude)와 위상(phase)의 두 가지로 구성되어 있다. 인간의 귀는 크기에는 민감하지만, 위상의 변화에는 둔감한 성질을 이용한 워터마크 삽입 방법이다.

### 2.3 반향 은폐 기법(Echo hiding)[5]

반향은 일정한 시간 이후에 일정한 진폭으로 재현되는 신호로 원래 신호가 울리는 효과가 나타나게 된다. 반향 은폐 기법은 반향 신호는 어떤 시점에서 원본 오디오 신호와 반향 신호를 구별하지 못하게 되는 성질을 이용한 방법이다.

### 2.4 기타 오디오 워터마크 기법[5]

이 논문에서 제안한 확산 대역 기법과 심리 음향 모델을 이용한 기법을 비롯하여 워터마크에 암호화 기법이나 채널 코딩(channel coding)을 적용한 방법과 2가지 이상의 다중 워터마크를 삽입한 방법들이 연구되고 있다.

## III. 확산 대역 워터마킹 기법

워터마크가 각종 왜곡과 변형에 강인한 특성을 가지기 위해서는 청각적으로 중요한 위치에 삽입되어야 한다. 일례로 오디오 신호의 고주파 대역에 삽입된 워터마크는 오디오 신호의 약간의 음질 열화를 감수하며 저주파 통과 필터링(Low-pass filtering)을 통해서 쉽게 제거될 수 있다.

따라서, 문제는 원래 오디오 신호에 큰 변형을 주지 않으면서도 주파수 상에서 청각적으로 가장 중요한 위치에 어떻게 워터마크를 삽입할 것인가 하는 것이다. 주파수 영역에서 저주파에 해당하는 계수(coefficient) 값을 매우 작게 바꾸어 준다면 그러한 조건을 만족하겠지만 이런 작은 변화는 잡음(noise)에 매우 약한 문제를 일으켜서 워터마크가 쉽게 삭제되게 된다.

이러한 문제를 해결하기 위해서 워터마크를 채널

(channel)을 통해서 전달되는 신호로 생각하고 무선 통신에서 사용되는 확산 대역 통신(spread spectrum communication)기법을 적용하는 방법을 생각할 수 있다[1]. 확산 대역 통신에서는 협대역(narrowband) 신호를 하나의 주파수 값에서 신호의 에너지가 인식되지 않을 정도로 확산(spreading)시켜 훨씬 더 큰 대역폭(wideband)에 전송한다. 이처럼 저작권 정보 신호의 에너지를 넓은 주파수 영역으로 확산시켜 잡음과 비슷한 상태로 만들어 주고 이 신호를 오디오 신호에 더하여 저작권 정보를 삽입하는 방법이 확산 대역 워터마킹 기법이다.

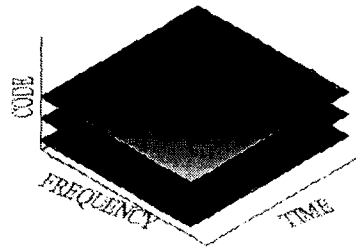


그림 1. 주파수 확산 방법

## IV. 심리음향모델

Perceptual Coding이나 MPEG Audio Coding에서는 오디오 신호의 음질을 저하시키지 않기 위해 심리음향 모델을 이용한다. 심리음향모델은 최소가청한계(absolute hearing threshold)와 마스킹(masking) 특성이 이용되고 있는데, 최소가청한계는 고요할 때 인간의 청각이 감지할 수 있는 음의 최소 레벨로서, 고요할 때 청각이 감지할 수 있는 잡음의 한계와 관계가 있다. 그림 2와 같이 고요할 때의 최소가청한계는 음의 주파수에 따라 다르며, 최소가청한계보다 작은 음은 들을 수 없다.

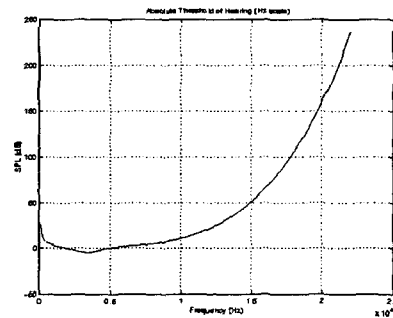


그림 2. 최소가청 한계

마스킹 특성은 그림 3과 같이 특정 밴드(band)의 강한 음이 마스크(masker)가 되어 근처 주파수의 음이 들리지 않게 되는 것을 말한다. 마스킹 효과가 일어나는 주파수 폭을 임계 대역(critical band)라고 부르며 임계 대역은 마스크와 마스킹 되어 들리지 않게 되는 음인 마스크(maskee)의 주파수에 따라 달라진다.

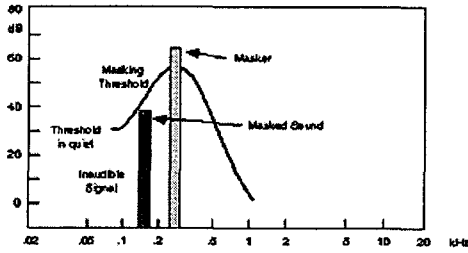


그림 3. 마스킹 효과

이 논문에서는 MPEG ISO 문서[3]에 제안된 심리음향모델을 이용하여 마스킹 곡선을 계산하였고, 워터마크를 삽입할 때 워터마크에 의한 오디오 음질에 영향을 최소한으로 하며, 워터마크의 검출을 위해서 가능한 범위 내에서 가장 강하게 삽입할 수 있도록 하기 위한 weighting factor로 심리음향모델이 사용되었다. 워터마크를 삽입하고자 하는 오디오 신호를 프레임별로 분석하여 심리음향모델의 가청한계곡선(masking threshold)을 구하여 적용하게 된다.

### V. 워터마크 삽입 및 검출

본 논문에서는 I. J. Cox가 제안한 확산 대역 기법 [1]과 M. Swanson이 제안한 심리음향모델[2], MPEG ISO 문서의 심리음향모델[3]을 적용하여 워터마크를 삽입하였다.

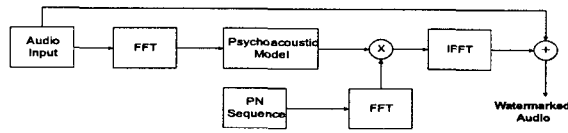


그림 4. 워터마크 삽입 과정

우선 워터마크를 넣고자 하는 오디오 신호를 1024개의 샘플 단위로 frame을 나눈다. 각 frame에 대하여 아래와 같은 과정으로 워터마크 신호를 삽입한다.

1. 오디오 frame  $s_i(k)$ 에 hamming window를 씌운 뒤, power spectrum  $S_i(k)$ 를 구한다.

$$S_i(k) = \text{Mag}(FFT(s_i(k) * h(w)))$$

2.  $S_i(k)$ 의 주파수 마스킹 함수  $M_i(k)$ 를 계산한다.
3. 삽입하고자 하는 워터마크 신호  $y_i(k)$ 의 power spectrum  $Y_i(k)$ 을 구한다.
4.  $Y_i(k)$ 에  $M_i(k)$ 를 곱하여  $P_i(k) = Y_i(k)M_i(k)$ 를 구한다.
5.  $P_i(k)$ 의 역 푸리에 변환을 구하여

$$p_i(k) = IFFT(P_i(k)) \text{를 구한다.}$$

6. 생성된 워터마크 신호  $p_i(k)$ 를 오디오 신호  $s_i(k)$ 에 더하여 워터마크가 삽입된 오디오 신호를 만든다.

$$s'_i(k) = s_i(k) + IFFT(Y_i(k)M_i(k)) = s_i(k) + p_i(k)$$

워터마크 검출에는 원본을 사용하지 않으며, 오디오 신호가 워터마크 신호보다 훨씬 크기 때문에, 오디오 신호의 성분을 작게 줄여줄 필요가 있다. 이 과정으로 pre-processing 과정을 거치는데 대표적인 방법은 Kutter의 방법[6]과 LPC[7]가 있다. 본 논문에서는 pre-processing을 거친 오디오 신호와 워터마크 신호를 Matched filter를 거쳐서 미리 정해진 false alarm rate( $10^{-3}$ )에 따른 threshold로 Hard-decision하여 워터마크의 여부를 판별하였다.

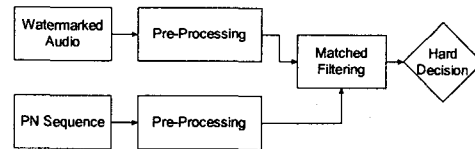


그림 5. 워터마크 검출 과정

### VI. 실험 결과

5가지의 여러 가지 오디오 신호에 강인성 실험을 했으며, 각 오디오 신호는 16bit, 44.1kHz, stereo 포맷이다. 다음 페이지의 표는 워터마크 검출 여부를 판별한 결과이다. 워터마크가 삽입된 오디오 신호의 음질은 주관적인 판단으로 원 신호와 차이가 없었다. 10초 동안의 오디오 신호를 Hard-decision을 하여 워터마크의 유무를 판별하였다.

각 샘플의 음악은 아래와 같다.

- Sample1 Bach Goldberg Variations (Piano Solo)
- Sample2 Glinka Ruslan & Ludmilla (Orchestral)
- Sample3 Mack the knife (Jazz Vocal)
- Sample4 Somebody to love (Rock Vocal)
- Sample5 취중진담 (Korean Vocal)

|                       | sample1 | sample2 | sample3 | sample4 | sample5 |
|-----------------------|---------|---------|---------|---------|---------|
| no attack             | 86 / 86 | 86 / 86 | 86 / 86 | 86 / 86 | 86 / 86 |
| Amplitude Compression | 86 / 86 | 86 / 86 | 86 / 86 | 86 / 86 | 86 / 86 |
| Equalization          | 86 / 86 | 86 / 86 | 86 / 86 | 86 / 86 | 86 / 86 |
| Band-pass Filtering   | 86 / 86 | 84 / 86 | 86 / 86 | 86 / 86 | 86 / 86 |
| Addition Echo         | 86 / 86 | 86 / 86 | 86 / 86 | 86 / 86 | 86 / 86 |
| MP3                   | 84 / 86 | 80 / 86 | 81 / 86 | 82 / 86 | 75 / 86 |

## VII. 결론

본 논문에서는 심리음향모델과 확산 대역 기법을 이용한 워터마킹 알고리즘을 연구하였다. 주파수 영역 상에서 삽입하고자 하는 워터마크 데이터를 효과적으로 변환하여 삽입하는 방법을 살펴보고, 여러 가지 신호처리 공격을 통해 강인성을 검증하였다. 하지만, 주파수 영역에서 워터마크 신호를 삽입하였기 때문에 시간 영역에서 프레임 사이즈가 변하는 신호처리를 한 경우엔 워터마크 검출이 매우 힘들 것으로 보인다. 앞으로 프레임 사이즈가 변화하는 공격에 강한 기술 개발이 필요하다고 생각된다.

## Reference

- [1] I. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia", *NEC Technical Report*, 95-10, NEC Research Institute, Princeton, NJ, 1995.
- [2] M. Swanson, B. Zhu, A. Tewfik, and L. Boney, "Robust Audio Watermarking using Perceptual Masking", *Signal Process*, 1998.
- [3] ISO/IEC IS 11172-3, "Information technology coding of moving pictures and associated audio for digital storage up to about 1.5Mbits/s".
- [4] Changsheng Xu, Jiankang Wu, Qibin Sun, and Kai xin, "Appciations of Digital Watermarking Technology in Audio Signals", *J. Audio eng. Soc.* Vol 47, No. 10, 1999.
- [5] W. Bender, D. Gruhl, "Techniques for data hiding", *IBM Systems Journal* vol.35, Nos 3&4, 1996.

- [6] M. Kutter, "Digital Image Watermarking : Hiding Information in Images", *Ph. D. Thesis*, Ecole Polytechniques Federale de Lausanne, 1999.
- [7] J. W. Seok, J. W. Hong, "Audio watermarking for copyright protection of digital audio data", *Electronics Letters*, 4th, January 2001, Vol. 31, No. 1.