

# 랜덤신호를 이용한 오디오 워터마킹 Audio Watermarking by use of Random Sequences

박창목, 신승원, 김종원, 신동환, 최종욱  
(주) 마크애니 부설 연구소  
E-mail: cmpark@markany.co.kr

본 연구는 디지털 오디오 데이터의 전송 및 배포과정에서 필요한 워터마크기술에 관한 것이다. 삽입 방법은 일종의 확산 스펙트럼 방식으로 특정 Key 에 의해 생성된 의사난수신호를 인간의 가청주파수모델과 오디오 신호의 특성에 따라 조작 한 후 삽입하게 된다. 이러한 주파수 조작은 데이터 은닉의 inaudibility 와 audio compression 에의 견고성을 위하여 필요한 사항이다. 워터마크 추출과정에서는 일정한 길이의 오디오 신호 앙상블 평균을 구하고, 이 신호와 워터마크와의 상호 상관함수를 구하여 워터마크의 삽입 유무를 판단하게 된다. 알고리즘의 테스트 결과 본 기술은 오디오 압축 및 오디오 조작에 강인한 것으로 나타났다.

차례

1. Introduction	5. Watermark Extracting
2. Literature Review	6. Robustness Testing
3. Generating Watermark Signals	6. Conclusion
4. Watermark Embeding	Reference

## 1. Introduction

디지털 워터마크는 네트워크상에서 사용 가능한 상태로 널리 분포, 유통될 수 있는 멀티미디어 데이터 및 출판물과 같이 지적 재산권 보호 대상 성격을 지니는 자료(intellectual copyright property)에 대해 원 데이터에 권리자 및 인증과 같은 추가적인 정보를 삽입하여 데이터에 대한 지적재산을 보호하기 위한 기법이다. 즉, Watermarking 저작권 보호(Copyright Protection) 기술의 일종으로서 영상, 비디오, 오디오 등의 멀티미디어 데이터에 소유주만이 아는 신호를 사람의 육안이나 귀로는 구별할 수 없게 삽입하고 불법 복제 유통시 이를 추출/검지하여 저작권을 입증할 수 있다. 따라서 영상, 음향, 비디오 등의 자료에 저작자 고유의 신호를 넣고, 후에 사용자들에 의해 데이터가 불법적으로 복제 유통되었을 때 원 소유주의 신호를 추출/검지하는 기술을 의미한다. 이는 CP(Contents Provider)서버로부터 사용자에게 디지털 콘텐츠 데이터를 제3자가 알아볼 수 없도록 암호화하는 Front-End 기술과는 달리 저작권을 확증해주는 Back-End 기술이다.

## 2. Literature Review

워터마크 삽입 알고리즘은 하위 비트 조작기법(Low-bit coding), 위상 부호화 기법(Phase coding), 확산 스펙트럼 기법(Spread Spectrum), 반향 은폐 기법(Echo hiding)으로 분류할 수 있으며, 데이터 삽입을 위한 Transform 방법은 푸리에 변환, Hartley 변환, 웨이블릿 변환 등을 사용한다. 적합한 변환의 선택은 삽입된 데이터의 견고성(robustness), 비가청성(inaudibility)등에 큰 영향을 주는 것으로 알려지고 있다.

### 2.1. 하위 비트조작기법(Low-bit coding)[2]

하위 비트 조작기법은 가장 단순한 오디오 워터마크 삽입 방법이다. 각 sampling 지점에 이진정보를 추가하고 삭제함으로써 워터마크를 삽입하는 방법이다. 전송률에 따라서 그 데이터 양이 조절될 수 있다. 예를 들어, 44.1khz의 sampling 지점은 모두 44100개이다. 따라서 44100개의 비트 정보가 들어갈 수 있는 것이다. 이 방법의 장점은 많이 데이터를 삽입할 수 있는 것이다. 가장 큰 단점은 워터마크 데이터가 너무 약하다는 것이다. 데이터에 noise의 발생이나, resampling과 같은 간단

한 변화나 조작에 의해서 쉽게 그 워터마크 데이터가 손실된다는 것이다. 하위 비트 조작기법은 주로 시간영역(Time domain)에서 워터마크의 삽입과 추출이 이루어진다. 시간영역에서 워터마크 삽입추출에 관한 알고리즘은 의사난수를 워터마크를 사용하는 방법과 이미지데이터를 워터마크로 사용하는 방법이 있다. 의사난수 방법과 이미지를 워터마크로 사용하는 방법은 하위 비트 조작기법 임에도 불구하고, resampling과 noise의 발생에 비교적 강인함(robustness)을 나타낸다.

## 2.2. 위상 부호화 기법(Phase coding)[2]

오디오 데이터는 위상(phase)과 크기(magnitude)의 두 가지로 구성되어 있다. 사람의 귀는 크기에는 민감하지만, 위상의 변화에는 둔감하다. 위상 부호화 기법은 이런 이론을 이용하여 워터마크 정보를 삽입하는 방법이다. 이 방법은 SNR(Signal to Noise Ratio)을 줄일 수 있는 가장 효과적인 방법이다.

위상 부호화 기법은 SNR에 거의 영향이 없기 때문에 음질이 좋긴 하지만, 위상의 변화는 위상의 예리함을 가져오게 된다. 이를 해결하기 위해서 평활화(smooth) 기법을 활용하게 되고, 이런 과정은 결국 대역폭(bandwidth)의 감소를 가져오게 되는 단점을 가지고 있다.

## 2.3. 확산스펙트럼 기법(SpreadSpectrum)[5]

확산 스펙트럼 방식은 주파수 스펙트럼에 조작성을 가함으로 해서 대역폭은 유지하고, 워터마크의 힘은 감소시킬 수 있는 방법이다. 이 방법의 기본 개념은 가능한 한 주파수 스펙트럼의 넓은 영역으로 데이터를 확산시키는 것이다. 이 방법에서 주로 사용되는 데이터는 의사난수(pseudo random number)이다. 키(key)값에 의해서 생성된 의사난수를 워터마크로 사용한다. 키는 워터마크를 삽입할 때, 그리고 추출할 때 모두 같은 값이어야만 한다.

확산 스펙트럼 기법은 위상 부호화 기법과 다르게, 난수를 오디오 신호에 노이즈로 삽입한다는 것이다. 워터마크를 삽입할 때, 귀에 들리지 않는 낮은 노이즈 레벨을 유지하기 위해서 오디오 신호의 유동 범위(dynamic range) 0.5퍼센트 이내로 워터마크 신호를 조정해야 한다.

확산 스펙트럼 기법은 가장 많이 연구되어지고 있는 분야로서 워터마크로 사용되는 데이터와 사람의 귀에 들리지 않게 하기 위해서 그 삽입 방법이 매우 다양하다. 많은 연구 논문들이 워터마크로 의사난수(Pseudo Random Noise)를 사용하였다. 의사난수는 발생시키는

키 값에 따라서 각기 다른 특징을 나타내고 화이트 노이즈의 성격을 띄고 있기 때문에 사람이 인식하기가 어렵다는 장점을 가지고 있어서 워터마크로 많이 사용되어 지고 있다. 그러나 낮은 노이즈 레벨을 유지해야 한다는 문제가 있다고 언급했던 것처럼 이에 대한 연구도 많이 진행되었다. 의사난수를 워터마크로 사용되 인간의 심리음향모델(Psychoacoustic Model)을 적용한 방법이 있다.

## 2.4. 반향 은폐 기법(Echo hiding)[4]

반향 은폐 기법을 설명하기 위해서는 먼저 반향(Echo)이 무엇인지를 알아야 한다. 워터마크는 반향의 세가지 성분 초기 진폭(initial amplitude), 쇠퇴율(decay rate), 갈라짐(offset)에 의해서 숨겨진다. 사람은 어떤 시점에서 원본 오디오 신호와 반향 신호를 구분하지 못한다. 그러나 이 방법은 원 신호에 의해서 나타나는 반향 효과(Cepstrum)로 워터마크 추출이 까다로운 단점이 있다.

## 2.5. 기타 오디오 워터마크 기법[6]

위에서 언급한 워터마크 기술 외에도 저작권 정보를 더욱 강력하게 보존하기 위한 방법들이 연구되어지고 있으며, 워터마크 기술에 암호화를 첨부하는 방법과 한가지 방법의 워터마크를 삽입한 기존의 방법과 달리 2가지 이상의 워터마크 삽입 방법을 이용하여 다중으로 워터마크를 삽입하는 방법도 연구되어지고 있다.

# 3. Generating Watermark Signals

## 3.1. Psychoacoustic modeling

워터마크의 설계는 첫째, 워터마크에 의해서 원 오디오의 음질에 영향을 최소한도로 하며, 동시에 워터마크의 검지를 위해서 가능한 범위 내에서 가장 강하게 삽입할 수 있도록 한다. 최소가청한계[7]는 고요할 때 청각이 감지할 수 있는 음의 최소 레벨로서 워터마크의 삽입 강도의 상한선을 의미하게 되며, Linear-Filter는 워터마크를 오디오 신호에 삽입할 때 가능한 동질적인 특성을 갖도록 하기 위해서 사용된다.

기저 신호를 최소 가청 한계 곡선으로 필터링하여 변경한 워터마크를 이용하는 것이 워터마크의 견고성이나 비가청성에 모두 효과적이다. 위의 <Figure 1>은 가장 고요한 상황에서 사람의 귀로 인지할 수 있는 한계곡선을 주파수 축으로 표시한 것으로 이 곡선보다 작은 에너지를 갖고 있는 워터마크는 들을 수 없다.

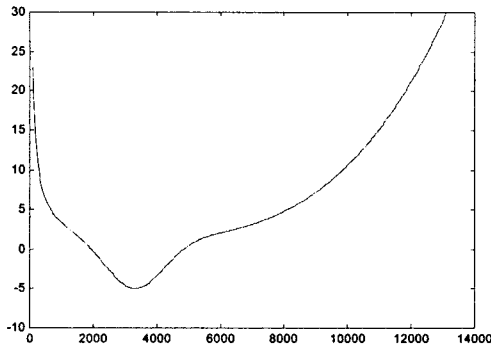


Figure 1 최소 가청 한계 곡선  
(absolute threshold of hearing)

최소 가청 한계 곡선의 생성식은 다음과 같다.

$$M = 3.64(f/1000)^{-0.8} - 6.5e^{-0.6(f/1000-3.3)^2} + 0.001(f/1000)^4$$

$f$  : frequency , Hz  
 $M$  : absolute threshold of hearing

그러나, 최소 가청 한계 곡선은 주변이 매우 조용한 상태를 기준으로 설정된 것이기 때문에 음악이 흘러 나오거나 다른 소리가 들리는 상황에서의 가청 한계 곡선은 이와는 다르다. <Figure 1>의 최소 가청 한계 곡선을 그대로 이용하면 일반적인 음악에 워터마크를 삽입할 경우에 워터마크의 성분이 지나치게 작게 되므로 일반적인 오디오에는 적합하지 않다. 그렇다고 모든 오디오 프레임을 분석해서 가청 한계 곡선을 구하여 랜덤 비트열을 변조하면 엄청난 계산량이 요구되기 때문에 현실적으로 적용이 불가능하다.

워터마크를 최적으로 필터링하기 위해서는 오디오에 적합한 가청 한계 곡선을 실험적으로 측정하여 이용하는 것이 가장 현실성이 있다. <Figure 2>는 실증적 분석을 토대로 구한 오디오 최소가청한계곡선으로서 클래식과 세미 클래식의 오디오를 MPEG-1 layer3의 심리청각모델로 분석한 결과이다.

$$P = [1.056801742606838e-026, -7.214332602361358e-022, 0.809126572761631e-017, -1.941502598267307e-013, 5.982813623951169e-010, 4.211560372433627e-006, -3.420594737587419e-002, 8.533065083348841e+001]$$

$$M = P(1)f^7 + P(2)f^6 + \dots + P(7)f + P(8)$$

$f$  : frequency, Hz

$M$  : Modified absolute threshold of hearing

변형된 최소 가청 한계 곡선의 생성식은 위 식과 같다.

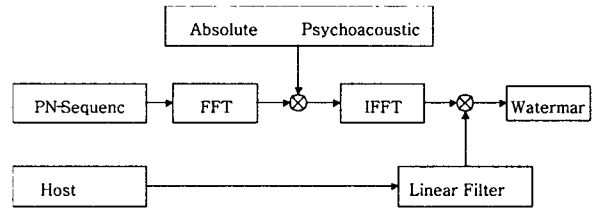


Figure 2. 워터마크 생성 과정

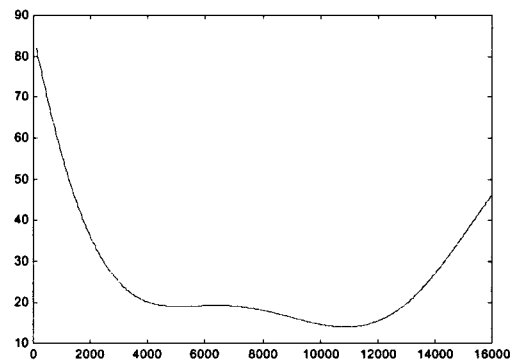


Figure 3. 실험에 의해 구해진 오디오에 대한 최소 가청 한계 곡선 Watermark Embedding

### 3.2. Variable Gain of the Watermark

가청한계를 고려하여 생성된 워터마크는 오디오 신호에 적절한 삽입 강도를 오디오의 특성에 맞게 결정하여 오디오에 삽입하게 된다. [Figure 4]에서 Gain Calculator는 원 오디오 신호의 최대/최소 크기를 이용하여 삽입강도를 계산한 후 이 삽입강도를 워터마크에 곱하여 오디오 신호와 더해 준다.

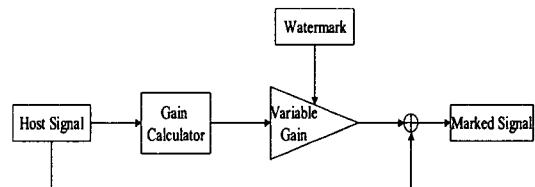


Figure 4. watermark 삽입강도 계산

$$s'(n+iN) = s(n+iN) + \alpha u(n)$$

$$n: 1, 2, \dots, N$$

$s(n+iN)$ :  $i$  번째 frame의  $n$  번째 신호

$s'(n+iN)$ : Embedded Signal

$u(n)$ : Watermark Signal

$\alpha = \text{max value in current frame} / 4000$

$N$ : frame size

### 3.3. The Overall Process

[Figure 5]는 워터마크의 삽입과정을 종합하여 나타낸 그림이다. 이 과정은 오디오 신호 전체에 워터마크가 삽입될 때까지 반복해서 진행된다. 단, 워터마크의 생성과정은 단 한번만 시행한다.

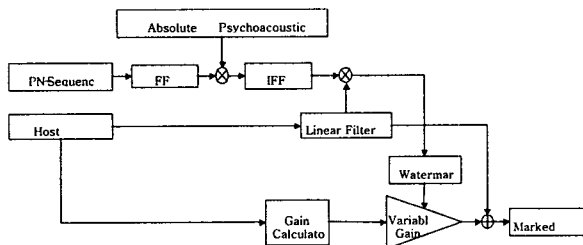


Figure 5. watermark 삽입 전체 구조

## 4. Watermark Extracting

### 4.1. Calculation of Ensemble Mean

일정한 길이의 오디오 신호의 앙상블 평균을 구하고, 이 신호와 워터마크와의 상호 상관 함수를 구하여 워터마크의 유무를 판별하게 된다. 앙상블 평균은 일반적으로 잡음을 감소하기 위한 방법으로 사용되나, 여기서는 오디오 신호를 잡음으로 보고, 우리가 삽입된 신호를 부각시키기 위한 방법으로 사용되어 진다. 상호 상관 함수는 FFT를 이용하여 순환 상호 상관 함수(circular cross correlation)를 구하게 된다.

만일  $P$ 개의 프레임을 이용하여 워터마크의 유무를 판별하게 된다면  $N$  개의 앙상블 평균 sequence 는 다음과 같이 구하여 진다.

$$f(n) = \frac{1}{P} \sum_{i=0}^{P-1} s(n+iN)$$

$f(n)$ :  $n$  번째 앙상블 평균

$s(n+iN)$ :  $i$  번째 frame의  $n$  번째 신호

$N$ : frame size

$P$ : 워터마크 유무를 파악하는데 필요한 frame 수

$n: 1, 2, \dots, N$

### 4.2. circular cross-correlation

앙상블 평균 sequence와  $k$ 만큼 위치 이동된 워터마크 sequence와의 circular cross correlation을 구하게 되면 correlation 함수의  $k$ 번째 위치에 상대적으로 큰 값이 존재하게 된다.

$$r = f(x) \circ u(x)$$

$$= \text{inverse fft of } [F^*(X) \cdot U(X)]$$

$r$ : correlation functions

$\circ$ : correlation

$u(x)$ : watermark sequences

$F^*(X)$ : the complex conjugate of  $F(X)$

## 5. Robustness Test

아래 결과는 16bit, 44.1kHz, Stereo 음악 5 곡에 대한 테스트 결과로 15초 간격의 Time-window에서 워터마크를 추출한 결과임

Signal Process	Description	Channel	
		Left-ch	Righth-ch
no process	Not attacked	100%	100%
Equalization	Freq[Hz] + 6/-6dB	100%	100%
Band-pass filtering	100Hz-6kHz	70%	70%
Codec	ISO/IEC 11172-3:1993 (MP3) - 128kbps	100%	100%
	Q-Design - 64kbps	0%	100%
	Windows Media Audio	100%	100%
Addition echo	Maximum delay: 100ms	100%	100%

## 6. Conclusion

본 알고리즘은 Sampling rate 변경을 고려할 경우, 워터마크 추출과정에서 프레임사이즈가  $2^n$ 이 아니므로 DFT를 이용하게 되고, 이에 따른 빠른 연산이 어렵다. 특히, 하드웨어 구현 시에 그 실현가능성이 문제가 되고 있다. 또한 Linear speed change, time scaling modification 과 같은 신호처리를 한 경우에 프레임사이즈가 변하므로 검지가 거의 불가능하다. 또한 프레임 사이즈와 관계가 깊은 공격 (re-sampling, time-scaling modification, linear speed change)에 강한 기술 개발이 필요하다

7. D. Pan, "A tutorial on mpeg audio compression", IEEE Multimedia Journal, Summer 1995

## Reference

1. 지적재산권 보호를 위한 정보은닉 기술 및 표준화 연구, 한국 전산원, 2000
2. Changsheng Xu, Jiankang Wu, Qibin Sun, and Kai xin, "Applications of Digital Watermarking Technology in Audio Signals", J. Audio eng. Soc. Vol 47, No.10, 1999
3. L. R. Rabiner/R. W. Schafer, "Digital Processing of Speech Signals", Prentic-Hall, 1978
4. W. Bender/D. Gruhl, "Techniques for data hiding", IBM Systems Journal vol 35, nos 3&4, 1996
5. Ingemar J. Cox, Joe Kilian, Tom Leighton, and Talal Shamoon, "Secure spread spectrum watermarking for images, audio and video", ICIP'96, vol III, 1996
6. Stefan Katzenbeisser, Fabien A. P. Petitcolas, "Information Hiding Techniques for Steganography and Digital Watermarking", Artech House, 2000