효율적인 PN 시퀀스 삽입을 통한 오디오 워터마킹*

변영배, 박창목, 김종원, 최종욱
(주) 마크애니 부속 연구소

Audio Watermarking Using An Effective PN Sequence Embedding Method

Youngbae Byun, Changmok Park, Jongweon Kim, and Jonguk Choi
MarkAny Inc., MarkAny Research Institute
Email : {byun, cmpark, jwkim, juchoi}@markany.com

요약 : 의사잡음 시퀀스를 이용한 대역확산 기반의 오디오 워터마킹은 들리지 않으면서도 강인한 워터마크를 만들기 위해 심리음향모델이나 고정필터를 사용하여 의사잡음 시퀀스를 변형시킨다. 그런데, 이러한 방법을 이용하여 스펙트럼 변형된 의사잡음은 고주파 영역에 대부분의 에너지를 갖게 되므로 인위적으로 오디오 신호의 고주파 영역을 잘라내는 공격에 취약하다는 단점이 있다. 본 논문에서는 이러한 단점을 보안하고 강인성 및 잡음의 최소화를 위하여 중간값의 성질을 이용하여 의사잡음을 변형 후 삽입하는 워터마킹 시스템을 제안한다. 중간값 성질을 이용하여 변형한 의사잡음은 원 오디오 신호와의 상관성이 높으며 전 주파수 대역에 고르게 분포하는 성질이 있으므로 고주파 영역의 공격에 강인하다. 제안 방법은 의사잡음의 고유성질을 최대로 살린 방법으로 각종 오디오 부호화, 부가잡음, 다운/업 샘플링, 채널변경, 진폭 공격과 같은 다양한 공격에도 워터마크 신호의 검출이 가능하다.

1. 서론

대역 확산(Spread Spectrum) 기반의 워터마킹은 영상 워터마킹에서부터 널리 사용되며, 오디오 워터마킹에서도 현재 가장 많이 연구되고 있는 대표적인 워터마킹 방법으로, 이진수로 표현된 워터마크 신호를 PN 시퀀스(Pseudo-Noise Sequence)로 대역확산하여 오디오 신호에 삽입하고, PN 시퀀스의 높은 자기 상관특성을 이용하여 삽입된 워터마크 정보를 추출하는 방법으로 간섭에 강한 특징을 갖는다[1,2]. 이때 복호화기에서는 삽입에 사용된 PN 시퀀스를 알고 있어야 검출이 가능하며, 이것이 검출을 위한 변수로 사용된다. 강인성의 향상을 위해서는 큰 에너지를 갖도록 삽입해야 하는데, 삽입되는 에너지가 증가할수록 원 신호의 SNR이 나빠지므로 음질의 왜곡을 초래한다.

이때 심리음향모델 등을 통해 얻을 수 있는 지각적 마스킹 효과를 이용하여 PN 시퀀스를 변형하여 삽입하게 되면, 지각되지 않으면서도 비교적 큰 에너지의 워터마크 삽입이 가능하여 강인성을 보다 향상시킬 수 있다[3,4]. 그런데, 워터마크를 삽입하는 모든 프레임에 대해 심리음향모델을 계산해야하기 때문에 부호화 과정이 매우 복잡하다는 문제가 있다. 이러한 문제점을 보안하고자 부호화기에서 심리음향모델의 역할을 대신하도록 만든 고정필터로 PN 시퀀스를 변형하면 간단하면서도 효율적인 워터마킹 시스템이 가능해진다[5].

하지만, 심리음향모델이나 고정필터를 이용하는 워터마킹 방법은 고대역통과 특성을 가지므로 스펙트럼 변형된 PN 시퀀스는 고주파 영역에 대부분의 에너지를 갖게되어 인위적으로 오디오 신호의 고주파 영역을 잘라내는 공격에 취약하다는 단점이 있다.

본 논문에서는 이러한 단점을 보안하고 강인하고 음질에 왜곡을 초래하지 않도록 하기 위해서 중간값의 성질을 이용하여 의사잡음을 변형 후 삽입하는 워터마킹 시스템을 제안한다. 제안한 PN 시퀀스 변형 방법을 사용하면 변형된 PN 시퀀스는 전 주파수 대역에 고르게 분포하는 성질을 최대로 살릴 수 있으므로 각종 오디오 부호화, 부가 잡음, 다운/업 샘플링, 채널변경, 진폭 공격 등과 같은 다양한 공격에도 강인하면서, 간단한 연산과정으로 PN 시퀀스를 삽입할 수 있다.

2장에서는 대역확산 기반의 오디오 워터마킹의 기본 원리에 대해 알아보고, 3장에서는 중간값을 이용한 새로운 오디오 워터마킹 방법을 제안한다. 4장에서는 모

* 본 논문은 국가지정 연구실 사업 (과제번호 : 2000N-NL-01-C-286)의 지원으로 수행되었음.

의실험을 통하여 제안한 오디오 워터마킹의 성능을 확인하고, 마지막 5장에서 결론을 맺는다.

2. 대역확산 기반의 오디오 워터마킹

대역확산 기반의 워터마킹은 이진수로 표현된 워터마크 신호를 PN 시퀀스(Pseudo-Noise Sequence)로 대역확산하여 오디오 신호에 삽입하고, 높은 자기 상관특성을 이용하여 삽입된 정보를 추출하는 방법이다.

PN 시퀀스는 평균이 0 이고 분산이 1 인 정규분포 $N(0,1)$ 과 같이 주어진 확률 분포로 만들어진 N 샘플의 시퀀스 $w(n)$ 으로 구성된다. 워터마크는 신호 $x(n)$ 에 삽입되어 워터마크가 삽입된 신호 $y(n)$ 을 만든다.

$$y(n) = x(n) + \alpha w(n) \tag{1}$$

여기서 α 는 워터마크의 크기를 결정하며, α 값을 크게 하면 워터마크 신호의 에너지가 커지므로 강인하고 복호화 성능이 좋아지지만 어느 정도 이상의 에너지로 삽입되면 보이거나 들릴 수 있기 때문에 비지각성을 만족시키지 못하므로, 워터마크의 두 가지 기본 특성인 "비지각성"과 "강인성"을 만족시키기 위해서는 적당한 α 값을 선정해야한다.

대역확산 기반의 워터마킹 방법은 복호화기에서 워터마크 정보의 삽입 시에 사용한 PN 시퀀스를 알고 있어야 검출이 가능하므로 높은 암호화 특성을 가지며 일반적인 신호처리 공격에 대해 대체로 강인한 장점이 있다.

3. 중간값의 성질은 이용한 오디오 워터마킹

3-1 부호화기

본 논문에서 제안한 워터마킹 삽입 알고리즘은 중간값의 성질을 이용한다. 중간값의 성질을 이용하여 변형한 의사잡음은 원 오디오 신호와의 상관성이 높으며 전 주파수 대역에 고르게 분포하는 성질이 있다. 따라서 비의도적 또는 의도적 공격에 대해 높은 보안성을 제공한다. 또한, 워터마크가 삽입된 신호는 어떠한 지각적 왜곡도 발생하지 않는다.

N 샘플의 오디오 신호 $x(n)$ 과 의사잡음 시퀀스 $w(n)$ 을 고려해보자. 워터마크가 삽입된 샘플 $y(n)$ 은 다음과 같이 표현된다.

$$\begin{aligned} y(n) &= x(n) + f(x(n), w(n)) \\ &= x(n) + \tilde{w}(n) \end{aligned} \tag{2}$$

여기서 $f(\cdot)$ 는 중간값의 성질을 이용하여 삽입할 워터마크를 생성하는 함수이다.

중간값의 성질을 이용하여 워터마크를 생성하기 위한 함수 $f(\cdot)$ 를 구체화하면 다음과 같이 표현된다.

$$\tilde{w}(n) = \begin{cases} (1)\ If\ I\times\Delta\times w(n)\geq 0\ then & \\ \quad 2\beta I|\Delta| w(n), & \text{if } \alpha > 2\times|\Delta| \\ \quad \alpha\beta I w(n), & \text{if } \alpha > |\Delta| \\ \quad \beta I|\Delta| w(n), & otherwise \\ (2)\ If\ I\times\Delta\times w(n)<0\ then & \\ \quad \beta I|\Delta| w(n), & \text{if } \alpha > |\Delta| \\ \quad \alpha\beta I w(n), & otherwise \end{cases} \tag{3}$$

여기서 Δ 는 전후 비트의 중간값과 처리하고자 하는 비트의 차분값으로

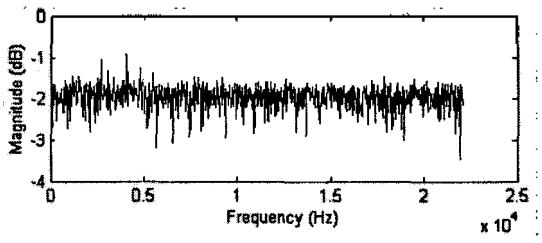
$$\Delta = (x(n-1) + x(n+1))/2 - x(n) \tag{4}$$

이고, α 는 전체 오디오 신호와 에너지에 비례하는 변수이며, β 는

$$\beta = 0.01\times(\alpha - \gamma) \tag{5}$$

이다. γ 는 한 프레임의 오디오 신호 에너지에 비례하는 변수이다.

함수 $f(\cdot)$ 에 의해서 변형된 PN 시퀀스는 그림 1 과 같이 전 주파수 대역에 고르게 분포하므로 인위적으로 오디오 신호의 고주파 영역을 잘라내는 공격, 즉 각종 오디오 부호화(압축)에도 워터마킹 신호를 제거할 수 없다.



[그림 2] Power Density Spectrum of filtered PN sequence

3-2 복호화기

PN 시퀀스의 상관특성을 이용하여 삽입된 워터마크 정보를 검출한다. 검출시에 원신호를 사용하게 되면 적은 연산량을 갖는 복호화기만으로도 신뢰도가 매우 높은 복호화가 가능하지만, 실시간 인증형 워터마킹과 같이 원 신호를 갖는 것이 불가능한 대부분의 응용 시스템에서는 사용할 수 없으므로 블라인드(blind) 검출 방식을 이용하여 워터마크 정보를 검출한다. 워터마크가 삽입된 오디오 신호로부터 원래의 오디오 신호만을

제거하기 위하여 선형예측(Linear Prediction)으로 AR 모델을 만들고 이것을 이용하여 워터마크가 삽입된 오디오 신호로부터 워터마크 신호를 추출하도록 하였다 [5, 6].

$$\begin{aligned} \tilde{w}'(n) &= y(n) - \tilde{x}(n) \\ &= x(n) + \tilde{w}(n) - \tilde{x}(n) \\ &= \tilde{w}(n) + e(n) \end{aligned} \tag{6}$$

워터마크 정보는 식(6)에서와 같이 추출된 워터마크 신호 $\tilde{w}'(n)$ 과 부호화기에서 삽입한 PN시퀀스 $\tilde{w}(n)$ 과의 상호상관관계(cross correlation)를 측정하여 결정한다. 상호상관관계는 다음과 같이 정의된다.

$$r(m) = \sum_{n=1}^{N} \tilde{w}'(n)\,\tilde{w}(n+m) \tag{7}$$

4. 성능평가

제안한 중간값을 이용한 워터마킹 시스템의 성능을 알아보기 위해서 표 1의 6가지 오디오 샘플에 대해 부호화 성능 측정을 수행하였다. 실험에 사용한 오디오 샘플은 CD에서 추출한 것으로 각각의 곡 전체에 대해서 실험하였다.

각각의 샘플 오디오의 샘플링 주파수는 44.1kHz 이고, 양자화 비트는 16bit인 스테레오 음악이다. 워터마킹 정보는 1초에 25비트씩 삽입하였다. 즉, 워터마킹 정보의 삽입율은 25bps 이다.

실험에서 설정한 변수들은 다음과 같다.

$$\alpha = \frac{\overline{P}_a}{10^4} \tag{8}$$

$$\gamma = \frac{\overline{P}_f}{10^4} \times 0.1 \tag{9}$$

여기서, $\overline{P}_a$ 은 오디오 신호의 전체 평균 전력이고, $\overline{P}_f$ 는 한 프레임 내에서의 오디오 신호의 평균 전력이다.

제안한 워터마킹 삽입 방법의 성능을 평가하기 위해서 위한 대표적인 공격들을 대상으로 BER을 측정하였다. 즉, 잔폭변조, 채널 변환, 업/다운 샘플링, 부가잡음, 부호화 등과 같은 다양한 공격에 대하여 성능 측정을 행하였다.

복호화 성능 측정은 식 (10)에 의해 BER을 계산하였다.

$$BER = \left(\frac{\text{발생한 에러 비트 수}}{\text{삽입한 전체 비트 수}}\right) \times 100\ (\%) \tag{10}$$

각각의 BER은 표 2에 나타내었다.

5. 결론

본 논문에서는 대역확산 기반의 워터마킹 시스템의 성능을 향상시킬 수 있는 효율적인 부호화 방법, 즉 중간값을 이용한 디지털 오디오 워터마킹 방법을 제안하였다. 제안 방법은 의사잡음의 고유성질을 최대로 살린 방법으로, 부호화가 간단하고, 변형한 의사잡음은 원 오디오 신호와의 상관성이 높으며, 전 주파수 대역에 고르게 분포하는 성질이 있으므로 고주파 영역의 공격에 강인하다. 또한, 각종 오디오 부호화, 부가잡음, 다운/업 샘플링, 채널변경, 진폭 공격과 같은 다양한 공격에도 거의 모든 워터마킹 정보의 검출이 가능하다. 또한 고음질의 높은 데이터율을 얻을 수 있으므로 소유권 증명을 위한 응용 등에 활용될 수 있다.

실험 결과는, 다양한 공격에도 소유자에 의해서는 검출이 가능하다는 것을 보여주고 있다. 차후에는 제안한 기법에 적합한 복호화기에 대한 연구 및 실제적인 워터마킹 시스템의 구현에 대한 연구가 진행될 것이다.

참고문헌

[1] W. Bender, D. Gruhl, N. Morinoto, and A. Lu, "Techniques for data hiding", IBM System Journal. Vol 35, No 3&4, 1996.

[2] I. J. Cox, J. Kilian, T. Leighton and T. Shamoon, "A secure, robust watermark for multimedia", Proc. Workshop on Information Hiding, Univ. of Cambridge, U.K., May 30-Jun. 1, 1996.

[3] L. Boney, A. Tewfik and K. Hamdy, "Digital Watermarks for Audio Signals", IEEE Int. Conference on Multimedia Computing and Systems, pp. 473-480, 1996.

[4] M. Swanson, B. Zhu, and A. Tewfik, "Current state of the art, challenges and future directions for audio watermark-ing", IEEE Int. Conference on Multimedia Computing and Systems, pp 19-24, 1999.

[5] 김현욱, 오현오, 김연정, 윤대희, "고음질 디지털 오디오 워터마킹을 위한 효율적인 PN 시퀀스 삽입 및 검출 방법", 방송공학회지, 6권, 1호, 2001년

[6] Jon Won Seok and Jin Woo Hong, "Audio watermarking for copyright protection of digital audio data", Electronics Letters, Vol. 37, No. 1, pp. 60-61, Jan. 2001.

표 1. 실험에 사용한 6곡의 오디오 샘플.

샘플	곡 명	작곡가 & 가수	설 명	길 이
S1	떠나가는 배	박인수 노래	가곡	3분05초
S2	Cartoon Heroes	Aqua	락	3분38초
S3	Living in the country (Summer)	George Winston	피아노 연주곡	3분45초
S4	Song bird	Kenny G	조용한 연주곡 (재즈)	4분00초
S5	Thank You	Boyz Men	힙팝	4분34초
S6	Tombe La Neige	Adamo	샹송	2분53초

표 2. 부호화 및 복호화 후의 BER 측정 결과.

Attack	S1	S2	S3	S4	S5	S6
No Attack (44.1kHz/16bit/2ch)	0.00 %	0.00 %	0.00 %	0.03 %	0.00 %	0.00 %
Amplitude compression (44.1kHz/8bit/2ch)	1,88 %	0.22 %	18.16 %	6.27 %	0.61 %	2.01 %
Channel conversion (44.1kHz/16bit/1ch)						
Left 90%, Right 10%	0.00 %	0.00 %	0.00 %	0.02 %	0.00 %	0.00 %
Left 50%, Right 50%	0.00 %	0.00 %	0.00 %	0.02 %	0.00 %	0.00 %
Left 10%, Right 90%	0.00 %	0.00 %	0.00 %	0.02 %	0.00 %	0.00 %
Up/Down sampling						
48kHz/16bit/2ch	0.00 %	0.00 %	0.00 %	0.03 %	0.00 %	0.00 %
32kHz/16bit/2ch	0.00 %	0.00 %	0.00 %	3.38 %	0.53 %	0.28 %
22kHz/16bit/2ch	0.63 %	0.50 %	0.32 %	13.17 %	7.61 %	7.50 %
Linear data compression						
MPEG1 Audio Layer3 (MP3) 128kbps	0.02 %	0.00 %	0.00 %	1.69 %	0.07 %	0.01 %
MPEG1 Audio Layer3 (MP3) 112kbps	0.03 %	0.00 %	0.00 %	3.32 %	0.20 %	0.10 %
MPEG1 Audio Layer3 (MP3) 96kbps	0.08 %	0.04 %	0.06 %	10.57 %	2.99 %	0.74 %
MPEG1 Audio Layer3 (MP3) 80kbps	2.54 %	0.85 %	3.31 %	20.13 %	10.47 %	12.86 %
MPEG1 Audio Layer3 (MP3) 64kbps	5.12 %	1.21 %	5.52 %	23.21 %	12.23 %	14.41 %
ATRAC 3 66kbps	0.00 %	0.00 %	0.00 %	7.28 %	0.33 %	0.06 %
ATRAC 3 105kbps	0.04 %	0.00 %	0.03 %	7.78 %	0.52 %	0.15 %
ATRAC 3 132kbps	0.01 %	0.00 %	0.00 %	4.37 %	0.07 %	0.00 %
Noise						
White noise -30dB	0.12 %	0.00 %	6.59 %	0.61 %	0.00 %	0.06 %
White noise -40dB	0.00 %	0.00 %	0.02 %	0.02 %	0.00 %	0.00 %