

복소 켈스트럼 영역에서의 오디오 워터마킹 방법

A Digital Audio Watermarking Algorithm in the Complex Cepstrum Domain

이 상 광, 호 요 성

Sang-Kwang Lee, Yo-Sung Ho

Kwangju Institute of Science and Technology

1 Oryong-Dong Puk-Gu, Kwangju, 500-712, Korea

Phone: (62) 970-2263, Fax: (62) 970-2204, E-mail: sklee@kjist.ac.kr

In this paper, we propose a digital audio watermarking algorithm in the cepstrum domain, where we insert a watermark into the cepstral components of the audio signal using techniques analogous to spread spectrum communications, hiding a narrow band signal in a wideband channel. Because the complex cepstrum takes the complex logarithm to the Fourier transform, it is robust against the differences in scale. Also, the multiple scaling parameters are used for the appropriate perceptual system.

Keywords: Watermark, Spread Spectrum, Cepstrum, Mel-Cepstrum

1. 서론

최근 워터마킹에 관한 관심이 두드러지고 있다. 워터마킹은 주로 오디오, 영상, 비디오 신호와 같은 디지털 데이터의 저작권 보호를 제공하기 위한 필요성으로 개발되기 시작했다. 디지털 데이터는 아날로그 데이터에 비해 여러 가지 이점을 제공해 주지만, 원본과 동일한 사본이 불법적으로 제작될 수 있다는 사실이 소유권에 심각한 위협을 준다. 워터마킹은 소유자 확인(owner identification), 특허권 사용료 지불(royalty payments), 그리고 데이터가 원본을 변화시켰는지 아닌지를 결정하는 인증(authentication) 등에 사용된다.

워터마킹이란 오디오, 영상, 비디오 신호 등의 멀티미디어 데이터에 들리지 않거나 보이지 않도

록 데이터의 소유권 정보를 삽입하는 과정으로, 일단 워터마킹된 데이터는 권한이 없는 사용자에게 통계적으로 소유권 정보의 검출이 불가능하여야 하며, 필터링이나 압축 등과 같은 공격에도 소유자에 의해서는 소유권 정보의 검출이 가능하여야 한다.

영상과 비디오 신호에 비교해 볼 때, 오디오 신호는 단위 시간 간격당 매우 적은 수의 표본으로 표현된다. 이것은 강인하고 들리지 않게 삽입할 수 있는 정보량이 영상 신호에 비해 매우 적다는 사실을 말해준다. 오디오 워터마킹의 또다른 문제는 인간 청각 시스템(Human Audible System, HAS)이 인간 시각 시스템(Human Visual System, HVS)보다 매우 민감하고, 들리지 않는 부분이 보이지 않는 부분보다 매우 얻기 힘들다는 것이다.

Boney *et al.* [1, 2]은 오디오 워터마킹에 대한 대역확산(spread spectrum) 방법을 제안하였다. 그들은 여러 단계에 거쳐 필터링된 의사 잡음 시퀀스(pseudo-random sequence)를 이용하여 HAS의 장기간 및 단기간 마스킹 효과를 얻게 된다. Bassia와 Pitas [3]는 시간 영역에서 오디오 신호에 대역확산 방법을 이용하여 워터마크를 직접 삽입하였다. 하지만, 이 방법들은 인간의 청각 신경에 대한 수평 주파수 해상도에 대해서는 고려하지 않았다.

본 논문에서는 오디오 신호의 복소 켈스트럼(complex cepstrum) 성분에 워터마크를 삽입하고 이를 위해 대역확산 방법을 이용하였다. 이는 인간의 청각 신경이 소리 크기의 로그값에 비례해서 반응한다는 것을 고려한 것이다. 또한, 멜켈스트럼(mel-cepstrum)을 이용하여 인간 청각 신경의 수평 주파수 해상도를 고려하여 워터마크를 삽입할 수 있다.

2. 워터마크 삽입

주파수 영역에서의 워터마킹은 Cox *et al.* [4]에 의해 처음으로 도입되었는데, 변조를 기반으로 각각 적응적 방법이 독립적으로 개발되었다. 이 방법에서 워터마크 신호는 시각적으로 중요한 주파수 성분에 확산되기 때문에 대역확산 통신 방식과 유사한 특성을 갖는다.

워터마크 삽입 방법에서, 워터마크 신호는 몇 개의 주파수 성분에 확산되므로 어느 한 주파수 성분에서의 에너지가 매우 작으며 그 검출이 용이하지 않다. 따라서, 오디오 신호의 스펙트럼에 워터마크 신호를 확산시켜 비의도적인 또는 의도적인 공격에 높은 보안성을 보장받을 수 있다. 워터마크가 삽입된 신호는 인간이 감지할 수 있는 오디오 신호의 열화를 나타내지 않았다.

워터마크가 삽입된 신호 $z(n)$ 은 다음과 같이 표현된다.

$$z(n) = f(x(n), w(n)) \quad (1)$$

여기서 $x = x(1), \dots, x(N)$ 은 N 개의 표본을 갖는

오디오 신호를, $w = w(1), \dots, w(N)$ 은 의사 잡음 시퀀스를 나타낸다. $f(\cdot)$ 은 워터마크를 삽입하는 함수이다.

주파수 영역에서, 식 (1)은 다음과 같이 표현될 수 있다.

$$Z(w) = F(w, X(w), W(w)) \quad (2)$$

여기서 $Z(w)$ 과 $F(w)$ 은 각각 유한 길이 N 에 취해진 $z(n)$ 과 $f(x, w)$ 의 이산 푸리에 변환(Discrete Fourier Transform, DFT)이다.

$z(n)$ 의 복소 켈스트럼 $c_z(n)$ 은 다음과 같이 정의된다 [5].

$$\begin{aligned} c_z(n) &= F^{-1}\{\log(Z(w))\} \\ &= F^{-1}\{\log(Z(w))\} \end{aligned} \quad (3)$$

여기서 $F^{-1}(\cdot)$ 는 1차 역 이산 푸리에 변환(inverse discrete Fourier Transform)을 나타낸다. DFT에 복소 로그(complex logarithm)를 취하는 이유는 인간의 청각 신경이 소리 크기의 절대값에 비례해서 반응하지 않고, 그것의 로그값에 비례해서 반응하기 때문이다. 또한, 로그값을 취하기 때문에 오디오 신호와 워터마크 신호 사이에 일정한 비율을 유지시켜 준다. $w(n)$ 을 $x(n)$ 에 삽입할 때, $w(n)$ 이 $x(n)$ 을 변화시키는 정도를 결정하는 α 를 정의한다. 워터마크 신호를 삽입하기 위해서, 식 (3)은 다음과 같이 된다.

$$c_z(n) = c_x(n) + \alpha(c_x(n))w(n) \quad (4)$$

여기서 $c_x(n)$ 은 $x(n)$ 의 복소 켈스트럼을 나타낸다. $\alpha(\cdot)$ 은 다중 비례 계수(multiple scaling parameter)를 나타내며, $w(n)$ 이 $x(n)$ 의 지각적 특성을 얼마나 변화시키는 지의 상대적인 측정자이다. 본 논문에서는 복소 켈스트럼의 분포가 한 곳에 집중되어 분포한다는 것을 이용하여 켈스트럼 대역에 적응적으로 다중 비례 계수를 적용하였다.

마지막으로 식 (4)에 역 복소 켈스트럼(inverse complex cepstrum)을 취하여 워터마크가 삽입된 신호 $z(n)$ 을 얻게 된다. 그림 1은 워터마크 삽입 과정을 보여주고 있다.

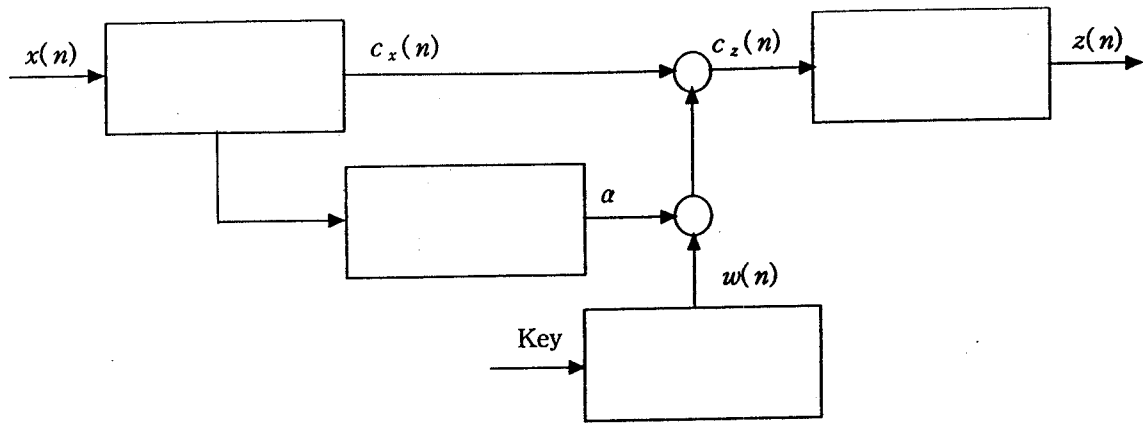


그림 1: 워터마크 삽입 알고리즘

3. 워터마크 검출

워터마크가 삽입된 오디오 신호는 여러 가지 공격으로 열화될 수 있는데, 이렇게 열화된 오디오 신호에서 워터마크 신호의 존재 여부를 조사하기 위해, 복원된 워터마크 신호 $w^*(n)$ 과 원래의 워터마크 신호 $w(n)$ 사이의 상호 상관 관계(cross correlation)를 계산하였다. 상호 상관 계수는 다음과 같이 주어진다.

$$\begin{aligned} r(m) &= \sum_{n=0}^N \alpha^{-1}(c_x(n)) w^*(n)w(n+m) \\ &= \sum_{n=0}^N w(n)w(n+m), 0 \leq m \leq N-1 \end{aligned} \quad (5)$$

여기서

$$\begin{aligned} w^*(n) &= c_z(n) - c_x(n) \\ &= \alpha(c_x(n))w(n) \end{aligned} \quad (6)$$

이상적으로, 의사 잡음 시퀀스는 백색 잡음(white noise)과 유사한 상관 관계 특성을 갖는 자기 상관 관계 함수(autocorrelation)를 가져야만 한다.

4. 실험 결과

본 논문에서 제안한 워터마크 삽입 및 검출 알

고리즘을 이용한 실험에는 몇 개의 16비트 스테레오 44.1kHz 오디오 신호가 이용되었으며, 1000개의 무작위로 발생된 워터마크 신호에 대한 워터마크 신호 검출 응답(watermark detector response)을 측정하였다.

그림 2는 삽입된 워터마크 신호의 유일성을 보여주고 있다. 올바른 워터마크 신호에 대한 응답이 그렇지 않은 신호에 대한 응답보다 매우 큰 값을 나타내고 있다.

그림 3은 연속으로 5번의 워터마킹을 수행한 신호에 대한 응답을 보여주고 있다. 이것은 워터마크가 삽입된 주파수 성분에 잡음을 추가하는 공격의 한 형태로 볼 수 있으며 반복적으로 워터마크가 삽입되는 경우 오디오 신호에 심각한 왜곡을 일으킬 수 있다. 실험 결과는 5개의 워터마크 신호 검출 응답이 대체로 낮은 값을 나타내었지만 어떠한 간섭없이 뚜렷이 구별되어 검출됨을 보여주고 있다.

그림 4는 MPEG-I Layer III의 압축이 워터마크 신호가 삽입된 오디오 신호에 어떠한 영향을 미치는가를 보여주고 있다. 그림 2와 비교해볼 때 거의 유사한 워터마크 신호 검출 응답을 보여주고 있으며 이는 MPEG-I Layer III의 압축에 강인함을 나타낸다.

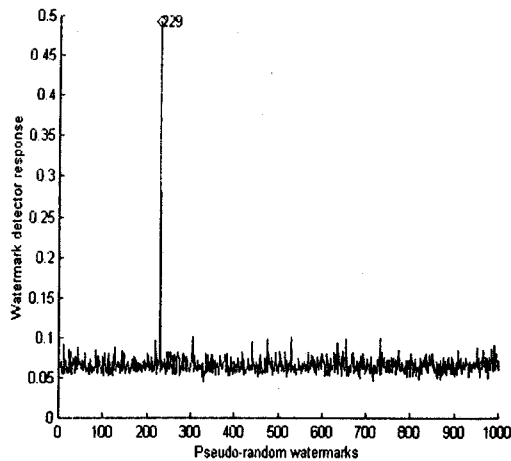


그림 2: 워터마크 신호 검출 응답

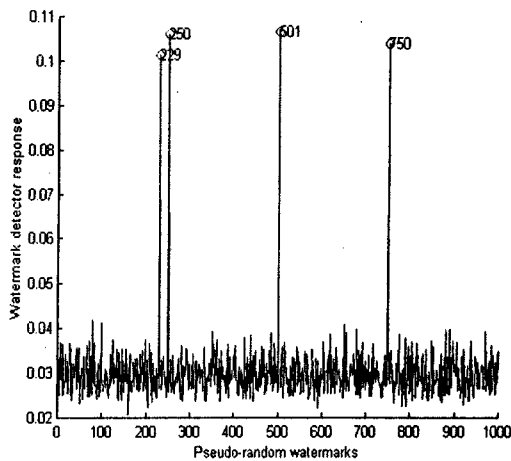


그림 3: 멀티 워터마크 공격

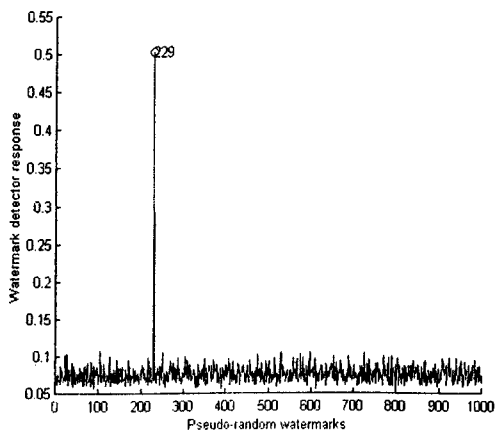


그림 4: MPEG-I Layer III 압축에 대한 응답

5. 결론

본 논문에서는 캡스트럼 영역에서의 디지털 오디오 워터마킹 방법을 제안하였다. 이 방법에서는 넓은 대역의 채널에 좁은 대역의 신호를 숨기는 대역 확산 방법을 이용하여 오디오 신호의 캡스트럼 성분들에 워터마크 신호를 삽입하였다. 실험 결과는 워터마크 신호는 권한이 없는 사용자에게 통계적으로 검출이 불가능하며 다양한 공격에도 소유자에 의해서는 검출이 가능하다는 것을 보여주고 있다.

감사의 글

본 연구는 광주과학기술원(K-JIST) 초고속광네트워크연구센터(UFON)를 통한 한국과학재단 우수연구센터(ERC)와 교육부 두뇌한국21(BK21) 정보기술사업단의 지원에 의한 것입니다.

참고 문헌

- [1] L. Boney, A. H. Tewfik, and K. N. Hamdy, "Digital watermarks for audio signals," *Proc. IEEE Int. Conf. on Multimedia Computing and Systems*, June 1996.
- [2] L. Boney, A. H. Tewfik, and Khaled N. Hamdy, "Digital watermarks for audio signals," *Proc. Europ. Signal Processing Conf.*, Sept. 1996.
- [3] V. Basia and I. Pitas, "Robust audio watermarking in the time-domain," *Proc. Europ. Signal Processing Conf.*, Sept. 1998.
- [4] I. Cox, J. Kilian, T. Leighton and T. Shamon, "Secure spread spectrum watermarking for images, audio and video," *Proc. IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, Sept. 1996.
- [5] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall, Englewood Cliffs, NJ, 1978.