

디지털 오디오의 저작권 보호를 위한 워터마크 삽입방법

석종원, 홍진우
한국전자통신연구원 방송미디어연구부

Watermark Embedding Method for Copyright Protection of Digital Audio

Jong Won Seok, Jin Woo Hong
Broadcasting Media Technology Department, ETRI

요약

최근들어 디지털 워터마킹(watermarking) 기법이 디지털 멀티미디어 콘텐츠 저작권 보호를 위한 새로운 해결책으로 제시되고 있으며, 국내외에서 이와 관련된 연구가 활발히 진행되고 있다. 본 논문에서는 이와 같이 최근 들어 많은 관심을 끌고있는 디지털 오디오의 저작권 보호를 위한 워터마킹 삽입방법을 소개한다. 개발된 오디오 워터마킹 기술은 원본 오디오에 비해 오디오의 품질이 떨어지지 않을 뿐 아니라 워터마크를 제거하고자 해지는 다양한 공격에도 워터마크가 제거되지 않는 특징을 가지고 있다.

1. 서론

인터넷과 같은 컴퓨터 망과 컴퓨터 이용의 급격한 발달로 인하여 문서, 음성, 오디오, 영상, 동영상 등의 멀티미디어 데이터의 이용 및 보급이 일반화 되고있다. 그러나 이러한 멀티미디어 데이터들은 디지털이라는 속성으로 인하여 복사를 하게 되면 또 하나의 원본이 만들어지게 되므로 누구나 손쉽게 불법적인 복제를 통해서 이들 디지털 데이터를 획득할 수 있게 된다.

지금까지 가장 대표적이고 널리 사용되는 데이터 보호기법은 데이터를 암호화 하는 방법으로 암호를 알지 못하면 데이터에 접근이 불가능하다는 장점이 있다. 하지만 일단 암호가 해독된 데이터는 아무런 제약 없이 불법적으로 복사되고 배포될 수 있다는 문제점을 가지고 있다. 이와 같은 이유로 인해서 최근에 디지털 워터마킹 기법이 디지털 멀티미디어 콘텐츠 저작권 보호를 위한 새로운 해결책으로 제시되고 있으며, 있으며 MPEG (Moving Picture Experts Group), DVD (Digital Versatile Disc) Forum, SDMI (Secure Digital Music Initiative) 등에서 표준화 기법을 규정하기 위한 연구가 진행되고

있으며, 국내외에서 이와 관련된 연구가 활발히 진행되고 있다[1-5].

현재까지 발표된 워터마크 알고리즘의 대다수는 영상 데이터를 대상으로 하였고, 오디오 데이터를 대상으로 한 경우는 극히 일부에 지나지 않는다. 하지만 최근 들어 문제가 되고 있는 MP3 오디오 파일의 불법복제 등을 생각해 볼 때 오디오 데이터의 워터마킹 기법 역시 시급히 해결되어야 할 과제이다. 현재까지 발표된 주요한 오디오 워터마킹 방법은 다음과 같다.

Bender[6] 등은 스프레드 스펙트럼, 에코우, 그리고 위상정보를 이용하는 방법의 3가지 방식의 오디오 워터마킹 방법을 제안 하였다. 스프레드 스펙트럼 방식은 저작권 정보를 PN 시퀀스를 이용하여 확산 시킨 후 이에 적절한 가중치를 두어 원본 오디오 신호에 더해 주는 방식이다. 에코우 방식은 청각적으로 감지할 수 없을 정도의 지연과 크기를 가지는 에코우를 오디오 신호에 삽입하는 방식이다. 삽입 시, 삽입되는 0과 1의 정보는 각각의 지연정도를 다르게 하여 삽입 하였다. 검출 시에는 자기상관 함수나 캡스트럼을 이용하여 지연된 위치에서의 피크를 찾아내어 삽입된 정보

가 1 인지 0 인지를 판별하게 된다. 위상을 이용하는 방법은 청각적 특성이 위상성분에 덜 민감하다는 점을 이용하여 위상성분에 저작권 정보를 삽입하여 주는 방법이다. Cooperman[7] 등은 오디오 신호에 FFT (Fast Fourier Transform)를 취하여 얻은 계수 값의 LSB (Least Significant Bit)에 워터마크를 삽입하고 다시 Inverse FFT를 취하여 워터마크가 삽입된 오디오 신호를 얻는 방법을 제안하였다. [8]에서는 오디오 신호를 주파수 영역으로 변환한 다음 6.5 kHz 이상의 대역을 워터마크 신호로 교체해 준다. 이 방법은 고주파 영역에만 워터마크 정보를 삽입하여 워터마크 삽입 후 오디오 신호의 음질이 떨어지는 것을 막고자 하였다. Solana Technology [9]에서는 스프레드 스펙트럼 방식과 선형예측분석 방법을 이용하는 방법을 제안하였으며 공격에 상당히 강인한 것으로 알려져 있다. 최근에는 Swanson[10] 등이 스프레드 스펙트럼 방식과 심리음향모델을 이용하는 방법을 제안하였다.

본 논문에서는 최근 들어 많은 관심을 끌고있는 디지털 오디오의 저작권 보호를 위한 워터마킹 삽입 방법에 대해 소개한다. 개발된 오디오 워터마킹 삽입 기술은 원본 오디오에 비해 오디오의 품질이 떨어지지 않을 뿐 아니라 워터마크를 제거하고자 가해지는 다양한 공격에도 워터마크가 제거되지 않는 특징을 가지고 있다.

II. 워터마킹 기법이 갖추어야 할 요건

(1) Imperceptibility

현재 워터마킹 기술에 대한 연구는 대부분 워터마크를 보이지 않게 또는 들리지 않게 영상이나 오디오 신호 속에 숨기는 방식들에 대한 것이다. 이것은 지적 소유권의 주장을 위해 워터마크를 삽입하면서도 서비스의 품질을 떨어뜨리지 않게 하기 위함이다. 예를 들어, 워터마크가 음악에 삽입되었을 때 원래의 음악과 워터마크된 음악 사이의 차이를 청취자가 구별할 수 없을 정도여야 하며, 영상 또는 비디오의 경우에도 마찬가지로 화질의 차이를 느낄 수 없어야 한다.

(2) Robustness

디지털 형태의 음악, 영상, 비디오 등은 손실 부호화, 필터링, 기하학적 변환 등의 일반적인 신호처리에 의해 쉽게 변형될 수 있다. 워터마킹 기술이 그 기능

을 발휘하기 위해서는 그 워터마크가 위와 같은 신호 처리 후에도 검출이 가능해야 한다. 특히, 오디오의 경우 위에서 언급한 필터링이나 압축공격 외에도 속도변화 (linear speed change), 스케일 변화 (pitch-invariant time-scaling change) 등 다양한 공격에도 워터마크가 살아 남아야 한다. 워터마킹 기술이 영상, 오디오, 비디오와 같은 멀티미디어 저작물에 대한 지적 소유권 보호를 위해 성공적으로 적용되기 위해서는 강인성이 무엇보다 중요하다.

(3) 워터마크의 검출 시 원 신호의 사용 여부

워터마킹 기법은 원 신호를 사용하는지 여부에 따라 크게 두 부류로 나누어볼 수 있다. 워터마크 검출 시 원 신호를 필요로 하는 경우는 실용적으로 적용하기 부적합하다. 즉, 원 신호가 유효하지 않은 경우가 존재할 수 있고, 신호처리 없이도 가짜 워터마크를 단순히 삽입함으로써 올바른 소유권자를 구별할 수 없도록 만들 수 있음이 지적되었기 때문이다[11]. 따라서 방송이나 PD (Portable Device) 등의 환경에 사용되기 위해서는 원 신호 없이 워터마크 검출이 가능해야 한다.

III. 오디오 워터마크 삽입방법

최근까지 연구된 오디오 데이터 워터마킹 방식 중 공격에 강하면서도 상대적으로 오디오의 품질을 손상시키지 않는 것으로 알려진 방식은 인간의 청각특성을 이용하는 방법이다. 즉, MPEG 오디오 부호화의 심리음향모델에서 사용되는 마스킹 임계치를 이용함으로써 워터마크가 삽입된 후에도 오디오 데이터의 품질을 떨어뜨리지 않게 되며 임의의 공격에도 강인한 특성을 지니게 된다.

이 장에서는 오디오 워터마킹 삽입방법으로 두 가지 방법을 제안한다. 첫 번째 방법은 DSSS (Direct Sequence Spread Spectrum) 기술에 기반한 알고리즘이며, 저작권 정보를 포함하게 되는 템플레이트로서 PN (Pseudo-noise) 시퀀스를 사용하는 방법이다. 두 번째 방법은 스프레드 스펙트럼 방식을 사용하지 않으며, 저작권 정보를 포함하게 되는 템플레이트로서 오디오 신호 자체를 이용하는 방법이다. 두 가지 방법 모두 심리음향모델을 이용하여 워터마크 신호를 삽입하게 되며 삽입한 후에도 오디오의 품질을 저하시키지 않는 장점을 가지고 있다. 그림 1은 오디오 세그먼트의 스펙트럼과 이에 해당하는 마스킹 임계치를 나타내고 있

다. 심리음향모델에서 마스킹 임계치를 구하는 과정은 아래와 같다.

- 1) 오디오 신호를 일정크기의 처리 단위인 프레임으로 나눈다.
- 2) 프레임 단위의 오디오 데이터에 FFT를 취하여 전력 스펙트럼을 구한다.
- 3) 구해진 전력 스펙트럼에서 순음 (Tonal)과 잡음 (Non-tonal) 성분을 구한다.
- 4) 미리 정의된 가청한계곡선을 이용하여 가청한계곡선 이하의 스펙트럼 성분을 제거한다.
- 5) 각각의 순음 성분과 잡음 성분에 대한 마스킹 곡선을 구한다.
- 6) 구해진 각각의 순음 및 잡음 성분의 마스킹 곡선을 이용하여 전체적인 마스킹 임계치를 구한다.

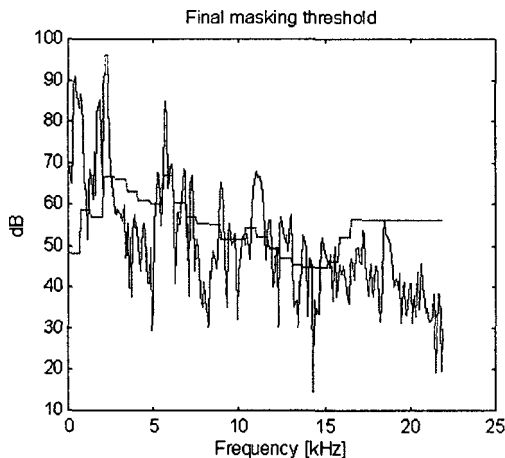


그림 1. 오디오 스펙트럼과 마스킹 임계치

첫 번째 오디오 워터마크 삽입 방법은 스프레드 스펙트럼 기술을 사용하며, 저작권 정보를 포함하게 되는 템플레이트로서 PN (Pseudo-noise) 시퀀스를 사용한다. 일반적으로 PN 시퀀스는 스프레드 스펙트럼 시스템에서 정보신호를 변조하여 확산시키는데 사용되며 잡음과 유사한 특성을 가지며 특히, 간섭에 매우 강한 성질을 가지며 자기상관이 우수한 것으로 알려져 있다. PN 시퀀스의 발생방법은 미리 정의된 길이의 linear feedback shift register를 이용하여 만들게 된다. 이때 PN 시퀀스의 길이는 shift register의 수에 의해 결정된다. 만일 shift register의 수가 m 이라면 최대 2^m-1 의 주기를 갖는 PN 시퀀스를 얻게 되며 2^m-1 개의 PN 시퀀스 중 하나를 워터마크 신호를 발생시키는 템플레이

트로 사용하게 된다. 그리고 이때 shift register의 상태 값이 템플레이트 PN 시퀀스를 발생시키는 키 (Key)가 되는 것이다. 따라서, 실제 삽입될 저작권 정보는 특정 키 값을 이용해 발생된 PN 시퀀스와 곱하여 확산시킨 다음 마스킹 임계치를 이용하여 가중치를 준 다음 원본 오디오 신호에 더해지게 된다. 자세한 삽입과정은 다음과 같다.

- 1) 심리음향모델을 이용하여 프레임 단위의 오디오 데이터에 대한 마스킹 임계치를 구한다.
- 2) 오디오 프레임 길이와 동일한 PN 시퀀스 신호를 발생시킨다.
- 3) 발생된 PN 시퀀스와 비트 스트림 형태의 저작권 정보를 곱하여 확산시켜 저작권 정보를 은닉시킨다.
- 4) PN 시퀀스를 이용하여 확산된 저작권 정보는 FFT를 취하여 주파수 영역의 데이터로 변환시킨다.
- 5) 1)번 과정에서 구해진 오디오 신호의 마스킹 임계치를 이용하여 주파수 영역으로 변환된 저작권 정보를 최대한 마스킹 임계치에 근접하도록 크기를 조절한다.
- 6) 마스킹 임계치에 근접하도록 크기가 조절된 저작권 정보에 Inverse FFT를 취하여 시간영역의 신호로 변환하여 최종적인 워터마크 신호를 얻어낸다.
- 7) 원본 오디오 신호와 6)번 과정을 통해 얻어진 워터마크 신호를 더하여 워터마크가 삽입된 오디오 신호를 얻는다.

일반적으로 스프레드 스펙트럼 방식의 워터마크 방식은 대부분의 공격에 상당히 강한 것으로 알려져 있고 실제 검출실험 시 검출율 역시 상당히 양호한 결과를 보여준다. 하지만 스프레드 스펙트럼 방식을 이용하는 워터마크 방법은 시간축 변형 공격에만은 상당히 취약한 단점을 가지고 있다. 즉, 속도 변화 (linear speed change), 스케일 변화 (pitch-invariant time-scaling change), Wow and Flutter 등의 시간축 상의 변화를 가져오는 공격이 가해지는 경우 검출이 불가능해진다. 이는 워터마크 검출 시 템플레이트로서 PN 시퀀스를 사용하기 때문인데 시간축 공격이 가해질 경우 오디오에 포함된 워터마크 신호와 템플레이트로서 사용되는 PN 시퀀스 사이의 동기가 맞지 않게 되기 때문이다.

두 번째 방법은 위에서 언급한 스프레드 스펙트럼 방식의 문제점을 해결하기 위해서 저작권 정보를 포함하게 되는 템플레이트로서 오디오 신호 자체를 이용하는 방법이다. 즉, PN 시퀀스를 사용하지 않고 변형된

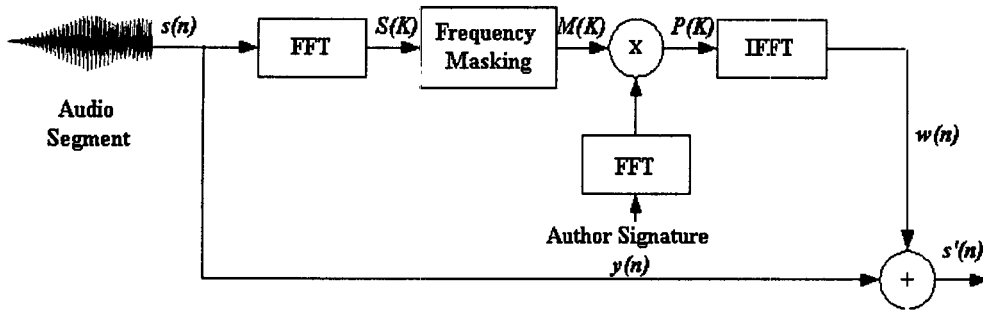


그림 2 오디오 워터마킹 과정

형태의 오디오 신호를 워터마크로서 사용하는 방법이다. 원본 오디오에 변형을 가하여 워터마크로 사용될 신호를 만든 다음 워터마크 신호에 약간의 지연(Delay)을 주어 원본 오디오 신호에 더해지게 된다. 이때 지연되는 정도가 바로 키 값이 되는 것이다. 검출 시에는 단순히 자기상관함수나 캡스트럼을 취하여 키 값인 지연된 위치에서 피크정보를 추출하여 저작권 정보를 얻어내게 된다. 삽입과정은 첫 번째 방법에서 PN 시퀀스 대신 변형된 형태의 원본 신호를 사용하는 것 외에는 동일하다. 그림 2는 심리음향모델을 이용한 워터마크 삽입과정을 나타내고 있다.

IV. 결론

본 논문에서는 최근 들어 많은 관심을 끌고있는 디지털 오디오의 저작권 보호를 위한 워터마크 삽입방법을 소개하였다. 개발된 오디오 워터마크 삽입 방법은 심리음향모델을 이용하여 원본 오디오에 비해 오디오의 품질이 떨어지지 않을 뿐 아니라 마스킹 임계치를 이용하여 워터마크의 강도를 조절하여 워터마크를 제거하고자 가해지는 다양한 공격에도 워터마크가 제거되지 않는 특징을 가지고 있다.

감사의 글

본 연구는 정보통신부의 "대화형 디지털 방송기술 연구" 사업의 일환으로 수행된 연구결과입니다

참고문헌

[1] M. Swanson, M. Kobayashi, and A. Tewfik, "Multimedia

data embedding and watermarking technologies," *Proceedings of IEEE*, Vol. 86, No. 6, pp. 1064-1087, June 1998.

[2] C. Cox, J. Killian, T. Leighton and T. Shamon, "Secure spread spectrum communication for multimedia", *Technical report, N.E.C. Research Institute*, 1995.

[3] R. J. Anderson and F. Peticolas, "On the limit of steganography," *IEEE J. Select. Areas Commun.*, Vol. 16, pp. 474-481, May 1998.

[4] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking," *Signal Processing*, Vol. 66, no. 3, pp. 357-372, May 1998.

[5] F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proceedings of IEEE*, Vol. 86, No. 6, pp. 1079-1107, June 1998.

[6] W. Bender, D. Gruhl, and N. Morimoto, "Techniques for data hiding," in *Proc. SPIE*, Vol. 2420, p. 40, San Jose, CA, Feb. 1995.

[7] M. Cooperman and S. Moskowitz, "Steganographic method and device," U.S. Patent 5,613,004, March 1997.

[8] J. Wolosewicz, "Apparatus and method for encoding and decoding information in audio signal," U.S. Patent 5,774,452, June 1998.

[9] C. Lee, K. Moallemi, and R. Warren, "Method and apparatus for transporting auxiliary data in audio signals," U.S. Patent 5,822,360, Oct., 1998.

[10] M. Swanson, B. Zhu, A. Tewfik, and L. Boney, "Robust audio watermarking using perceptual masking," *Signal Processing*, Vol 66, no. 3, pp. 337-355, May 1998.

[11] J. Brassil et al., "Electronic marking and identification techniques to discourage document copying," *IEEE J. Select. Areas Commun.*, Vol. 13, pp. 1495-1504, 1995.