

다운 믹싱에 강한 디지털 오디오 워터마킹 기법에 대한 연구

오승수 · 김기만

한국해양대학교 전파공학과

A Study on the Robust Digital Audio Watermarking against Down Mixing

Seung-soo Oh · Ki-man Kim

Dept. of Radio Science & Engineering, Korea Maritime University

E-mail : sunnyside@hananet.net

요 약

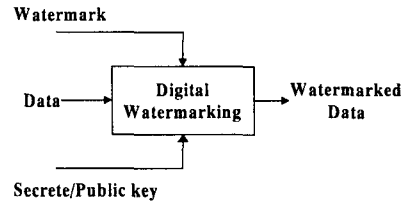
최근 디지털 기술의 발전으로 인해 여러 디지털 콘텐츠들의 불법복제가 확산되고 있다. 따라서 저작권 보호와 불법복제를 막기 위한 다양한 기술들이 연구되고 있다. 그 중 디지털 워터마킹 기법은 디지털 콘텐츠 내부에 사용자가 원하는 정보를 보이거나 들리지 않게 숨기는 것을 말하며, 이때 삽입되는 정보를 워터마크(Watermark)라 한다. 디지털 워터마크를 삽입 하는 방법으로는 크게 PN시퀀스를 삽입하는 방법, 위상을 변화시키는 방법, 반향신호를 삽입하는 방법이 있다. 이러한 워터마크는 다양한 공격에서도 검출할 수 있어야 한다. 본 논문은 미국음반산업협회(RIAA)에서 제안한 SDMI(Secure Digital Music Initiative) Phase II Screening중 다운 믹싱에 강한 워터마킹 방법에 대해서 연구하였으며, 시뮬레이션과 청취 실험을 통하여 성능을 검증하였다.

1. 서 론

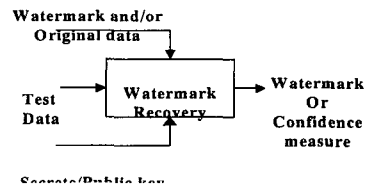
최근 몇 년 사이에 디지털 기술의 발전과 인터넷의 보급으로 인하여 디지털 멀티미디어 콘텐츠의 사용이 폭발적으로 증가하였고 그 시장도 급성장하고 있다. 하지만 디지털의 장점인 복제/배포의 용이성은 불법 복제를 확산시켜 넵스터(Napster)나 소리바다에 대한 소송문제와 같은 저작권에 대한 문제를 일으켰다. 따라서 이런 문제들을 해결하기 위한 디지털 저작권 보호에 관한 많은 연구가 활발히 진행되고 있다.

그 중 디지털 워터마킹은 디지털 콘텐츠 내부에 사용자가 원하는 정보를 보이거나 들리지 않도록 숨기는 것을 말하며, 이 때 삽입되는 정보를 워터마크(Watermark)라 한다[1][2][3]. 그림 1과 같이 워터마크를 삽입하고 추출해냄으로써 콘텐츠의 유통 제어, 복제 제어, 소유자 증명, 인증 등을 할 수 있다.

이러한 워터마킹은 기본적으로 두 가지 특성을 가진다. 첫째, 보이거나 들리지 않는 비지각성(Imperceptibility), 둘째, 신호처리나 의도적인 공격에 대한 강인성(Robustness)이다. 그러나 이 두 가지 조건은 서로 상충관계에 있다. 따라서 두 가지 조건을 만족하는 범위내에서 워터마크를 삽입해야 한다.



(a) Generic watermark embedding scheme



(b) Generic watermark Extracting scheme

그림 1. 일반적인 워터마킹의 삽입과 복구

최근 DVD의 보급으로 고음질의 멀티채널 오디오 재생이 보편화 되어가고 있다. 그러나 MPEG4 압축 알고리즘을 기반으로 한 DivX라는 codec이 등장하면서 DVD의 영상정보와 음향을

추출하여 인터넷을 통해 불법적으로 유통되고 있는 실정이다. 특히 멀티채널의 음향신호는 대개 2채널로 다운믹스되어 MP3형식으로 압축하여 영상신호와 함께 유통되고 있다.

본 논문은 미국음반산업협회(RIAA)에서 제안한 SDMI(Secure Digital Music Initiative) Phase II Screening에 만족하도록 멀티채널 오디오에 워터마크를 삽입하여 다운믹스를 한 경우에도 삽입된 정보를 추출할 수 있는 알고리즘을 제안하고자 한다.

2. 디지털 오디오 워터마킹[2]

디지털 오디오 워터마킹은 오디오 신호안에 저작권등의 정보를 들리지 않도록 은닉하는 것을 말한다. 오디오 콘텐츠의 저작권 보호를 위한 워터마킹은 다음과 같은 특성을 가진다.

- (1) 비지각성을 가져야 한다.
- (2) 원음이 손상되지 않고서는 제거되지 않아야 한다.
- (3) A/D - D/A변환, 복사, 편집, 압축 등의 신호처리 과정에도 강해야 한다.
- (4) 실시간 처리가 가능해야 한다.

인간의 지각능력은 청각이 시각보다 훨씬 더 민감하므로 삽입된 신호의 비지각성과 공격에 대한 강인성 두 조건을 만족시키기 어렵고, 또한 영상신호에 비해 데이터의 양이 적기 때문에 워터마크를 삽입할 수 있는 영역이 부족하다는 단점이 있다. 크게 디지털 오디오 워터마킹 방법은 다음과 같이 3가지로 나눌 수 있다.

2.1 Phase Coding

HAS(Human Auditory System)이 작은 위상 변화에 둔감하다는 특성을 이용하여 각 프레임별로 위상을 변화시킴으로서 워터마크를 삽입하는 방법이다. 그러나 이 방법은 오디오 신호가 안정적이어야 한다는 조건이 있으며 실제 오디오 신호는 안정적이지 않은 경우가 많기 때문에 음질에 많은 변화를 준다.

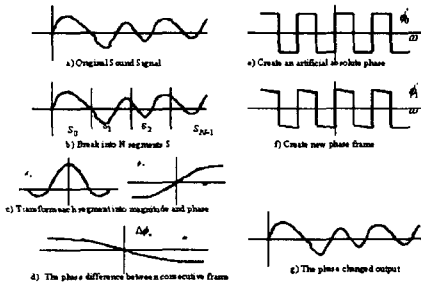
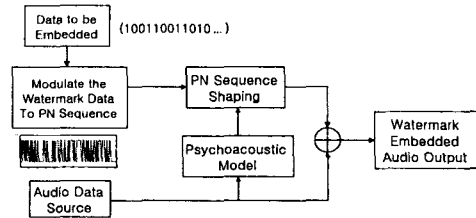


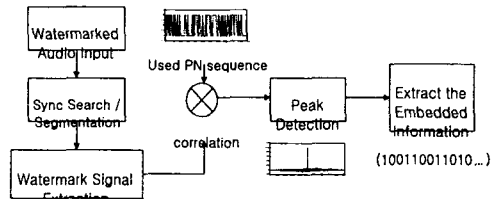
그림 2. Phase coding scheme

2.2 Spread Spectrum Coding

오디오 신호를 프레임 별로 나누어 각 프레임에 워터마크 신호로 PN코드를 삽입하는 방법이다. 이는 PN코드가 가지는 백색잡음과 유사한 통계적 특성과 좋은 자기상관 특성을 이용하는 것이다. 하지만 PN코드의 삽입은 오디오신호에 잡음을 추가하는 것이므로 검출이 가능한 한도내에 그 에너지가 들리지 않도록 삽입해야 한다.



(a) Watermark Embedding Scheme



(b) Watermark Extracting Scheme

그림 3. Spread Spectrum Coding 방식의 워터마크의 삽입과 추출

2.3 Echo Data Hiding

프레임별로 나누어진 오디오신호에 워터마크 신호로 인위적인 Echo신호를 삽입하는 방법으로 Echo신호의 크기와 지연시간을 조절하여 원음과 청각적으로 구분되지 않는 Echo신호를 삽입할 수 있다. 또한 이렇게 삽입된 워터마크 신호는 원신호 없이 복호화 할 수 있는 장점이 있다.

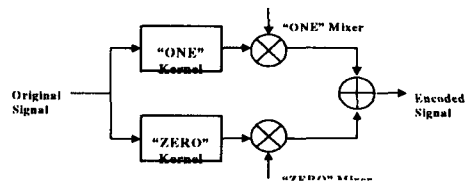


그림 4. Echo Embedding Scheme

3. 다운믹싱에 강한 오디오 워터마킹

앞에서 오디오정보가 영상정보에 비해 상대적으로 데이터가 적으므로 상대적으로 영상에 비해 삽입할 수 있는 워터마크 정보가 작다고 했지만,

일반적으로 스테레오라고 말하는 2채널, 그리고 DVD-Audio나 SACD(Super Audio Compact Disk)같은 멀티채널을 지원하는 음원에 대해서는 채널별로 다양한 정보를 삽입하여 워터마크 정보를 확대할 수 있다. 본 논문은 Spread Spectrum coding을 기반으로 하는 워터마킹 기법을 이용하여 멀티채널 오디오에 워터마크를 삽입하고 다운믹싱 하였을 때도 워터마크를 검출할 수 있는 알고리즘을 제안하였다.

3.1 Watermark Embedding Scheme

워터마크의 삽입과정은 그림 4와 같이 삽입하려는 정보 데이터가 '0', '1' 일 때 각각 다른 PN 코드를 선택하여 각 채널에 삽입한다. 채널별로 삽입할 때 랜덤하지만 '1'과 '1'을 같은 수로 가진 수열 값에 의해 PN코드를 각 채널에 인터리브(Interleave)하여 삽입한다. 이렇게 삽입함으로써 채널이 많을수록 더 많은 정보를 삽입할 수 있으며 따라서 다양한 부가정보를 제공할 수 있다. 또한 일반적인 방법으로는 복호를 할 수 없는 장점을 가지게 된다.

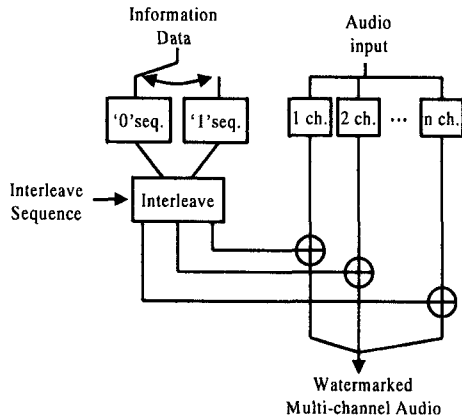


그림 4. 인터리브를 이용한 멀티채널 Watermark Embedding Scheme

3.2 Watermark Extracting Scheme

각 채널에 인터리브하여 삽입한 워터마크를 추출하기 위해서 삽입시 인터리브한 순서가 필요하다. 이 순서에 의해 PN코드를 재구성하여 워터마크된 정보를 추출할 수 있다. 그러나 채널 다운믹스를 한 경우에는 워터마크를 추출할 수 없다. 따라서 다운믹스 되었을 경우 그 여부를 파악하여 다운믹스된 오디오신호에서 삽입된 PN코드를 복구하여 워터마크를 추출할 수 있는 알고리즘을 수행한다. 그림 5는 인터리빙을 이용한 워터마크 삽입 알고리즘에 따른 추출 방법을 나타낸 것이고 그림 6은 워터마크가 삽입된 멀티채널 오디오가 다운믹싱이 되었을 때 삽입된 정보를 추출하

는 방법을 나타낸 것이다.

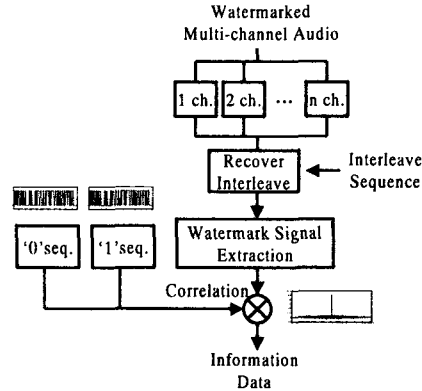


그림 5. 인터리브를 이용한 멀티채널 Watermark Extracting Scheme

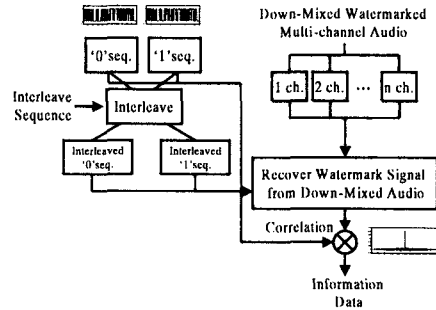


그림 6. 인터리브를 이용하여 삽입한 워터마크 오디오신호를 다운믹싱시킨 오디오 신호로부터 워터마크를 검출하는 알고리즘

4. 실험 및 결과

SDMI Phase II에서 규격 하는 신호처리에 관한 강인성 중 다운믹스에 대한 조건은 표 1과 같다[6].

표 1. 다운믹스에 관한 SDMI phase II의 규정

1) Persistent Assertion

Signal Processing	Description
Down Mixing and Surround Sound Processing	<ul style="list-style-type: none"> • 6 channel to stereo • SRS • Spatializer • Dolby Surround • Dolby Headphone

2) Typical Consumer Attack

Signal Processing	Description
Down Mixing and Surround Sound Processing	<ul style="list-style-type: none"> • Multi-channel to stereo • Stereo-to-mono

제안된 알고리즘을 이용한 시뮬레이션을 위해 위 조건 중 Typical Consumer Attack의 Stereo-to-Mono 다운믹스에 대하여 간단한 성능 테스트를 하였다. 10초 정도의 길이를 가지는 음원은 CD에서 추출한 CD음질(44100Hz, 16bit)의 테크노, 클래식, 록 장르의 음악을 사용하였다. PN코드의 1 프레임의 길이는 1024샘플이지만 2개의 채널로 인터리브를 시켜 삽입하므로 채널당 1프레임은 512개의 샘플을 가지게 되어있다. 삽입된 정보데이터의 길이는 860비트이고 워터마크는 시간영역에 삽입하였으며 워터마크 에너지는 충분히 사람의 귀에 들리지 않도록 하였으며 각 음원에 대한 삽입된 워터마크의 SNR (Signal - to - Noise Ratio)은 각각 30, 50, 70dB로 하였다.

성능을 검증하기 위해 인터리브를 이용한 워터마크가 삽입된 Stereo음원을 Mono로 다운믹스하고, 그 음원을 제안한 워터마크 추출 알고리즘을 이용하여 검출하고 정보데이터를 복구하였고 검출 알고리즘은 눈블라인드(Non-Blind) 검출방법을 사용하였으며 BER(Bit Error Rate)을 계산하였다. 실험의 결과는 표 2에 나타내었다.

표 2. 다운믹스된 워터마크 오디오를 이용하여 일 반방법과 제안된 방법으로 워터마크 신호 검 출한 결과

음원	SNR	BER(일반)(%)	BER(제안)(%)
Techno	30	46.9767	0
	50	54.6512	0
	70	50.9302	0
Classic	30	46.9767	0
	50	49.5349	0
	70	52.0930	0
Rock	30	46.9767	0
	50	46.9767	0
	70	48.3721	0

실험결과에서 알 수 있듯 다운믹스가 되었음에도 완벽하게 정보를 복원하는 것을 알 수 있다. 하지만 주파수영역에서의 워터마크의 삽입과 다른 신호처리 공격(압축, 필터 등)과 여러 입체음향환경에 대한 성능평가도 역시 필요하다.

5. 결 론

본 논문은 SDMI Phase II Screening에서 제안하는 규격 중, 채널 다운믹스에 강한 알고리즘을 제안하였다. 본 알고리즘은 멀티채널 오디오에 워터마크를 삽입할 때 랜덤시퀀스에 의한 인터리브를 시켜 정보삽입효율을 높이고 다운믹스한 경우에도 워터마크를 검출하여 삽입된 정보를 복구할 수 있도록 하였고 이를 실험으로 검증하였다.

하지만 실제 멀티채널 오디오의 경우 Dolby Digital이나 DTS같은 서라운드 시스템이 대부분이다. 이러한 멀티채널 서라운드 시스템은 입체음향 재현뿐만 아니라 압축알고리즘도 가지고 있다. 또한 Dolby Headphone과 같은 Multi - Channel을 스테레오로 다운믹스하며 입체음향을 재현하는 서라운드 시스템도 있다. 따라서 압축과 다운믹스 서라운드 등과 같은 다양한 조건에 부합하도록 하는 알고리즘의 보완이 필요하다.

참고문헌

- [1] L. Boney, A Tewfik and K. Hamdy, "Digital Watermarks for Audio Signal", *IEEE Proceedings of Multimedia '96*, 1996.
- [2] W. Bender, D. Gruhl, N. Morimoto, A. Lu, "Techniques for data hiding", *IBM System Journal*, Vol 35, Nos 3&4, 1996.
- [3] 김현욱, 오현오, 석종원, 홍진우, 윤대회, "강인한 오디오 워터마크의 블라인드 검출을 위한 PN시퀀스 삽입 및 추출에 관한 연구", 2001년도 한국음향학회 학술발표대회 논문집 제20 권 제1호 pp365~368, 2001.
- [4] P. Bassia and I. Pitas, "Robust audio Watermarking in the time domain", *Proc. EUSIPCO 98, vol. 1, pp. 22-28*, Rhodes, Greece, 1998.
- [5] J. D. Gordy and L. T. Bruton, "Performance Evaluation of Digital Audio Watermarking Algorithm", *MWSCAS2000, pp. 456-459*, Detroit, MI, USA, 2000.
- [6] SDMI(Secure Digital Music Initiative), "Call for Proposals for Phase II Screening Technology Version 1.0", 2000.