

工學碩士 學位論文

다운 믹싱에 강한 디지털 오디오 워터마킹
기법에 관한 연구

A Study on the Robust Digital Audio Watermarking against
Down Mixing

指導教授 金 基 萬

2001 年 11 月

韓國海洋大學校 大學院

電波工學科

吳 承 洙

本 論文을 吳承洙의 工學碩士 學位論文으로 認准함.

委員長 : 工學博士 趙 炯 來 (印)

委 員 : 工學博士 姜 仁 鎬 (印)

委 員 : 工學博士 金 基 萬 (印)

2002 年 2 月

韓國海洋大學校 大學院

電波工學科

吳 承 洙

차 례

그림차례.....	ii
표차례.....	iii
기호표.....	iv
Abstract.....	v
제 1 장 서 론.....	1
제 2 장 디지털 오디오 워터마킹.....	5
제 2-1 절 위상 정보 삽입 기법	7
제 2-2 절 대역 확산 기법	9
제 2-3 절 반향 신호 삽입 기법	13
제 3 장 다운믹싱에 강한 디지털 오디오 워터마킹.....	18
제 3-1 절 워터마크 삽입 과정.....	18
제 3-2 절 워터마크 검출 과정.....	20
제 3-3 절 다운믹싱 한 워터마크 삽입 오디오에서의 검출 과정.....	21
제 4 장 실험 결과.....	23
제 4-1 절 SDMI Phase II Screening.....	23
제 4-2 절 실험 방법 및 결과.....	27
제 5 장 결 론.....	30
참고문헌.....	31

그림 차례

그림 2-1	일반적인 워터마크의 삽입과 복구.....	6
그림 2-2	위상 정보 삽입 기법.....	9
그림 2-3	대역 확산 워터마크 삽입 과정.....	10
그림 2-4	변형 삽입 된 PN 코드 워터마크.....	11
그림 2-5	대역 확산 워터마크 검출 과정.....	12
그림 2-6	PN 코드 워터마크의 검출.....	13
그림 2-7	반향 신호의 충격 응답.....	14
그림 2-8	'0'과 '1'에 대한 반향 커널.....	15
그림 2-9	반향 신호 워터마크 삽입 인코더.....	15
그림 2-10	오토 캡스트럼을 이용한 워터마크 검출.....	16
그림 2-11	반향 신호 워터마크의 검출.....	17
그림 3-1	멀티 채널 오디오의 워터마크 삽입.....	20
그림 3-2	멀티 채널 오디오에 삽입된 워터마크 추출.....	21
그림 3-3	다운믹스한 오디오 신호에서의 워터마크 검출.....	22

표 차례

표 4-1	신호처리에 대한 워터마크의 강인성에 관한 규정.....	24
표 4-2	일반 사용자들의 신호처리.....	25
표 4-3	워터마크 삽입 및 검출 제한 조건.....	27
표 4-4	다운믹스 한 워터마킹 오디오를 이용하여 일반방법과 제안된 방법으로 워터마크 신호를 검출한 결과.....	28

ABSTRACT

In recent year, the use of digital multimedia content has increased explosively because of rapid progress of digital and network technologies. However, this increasing of the use of digital multimedia content were raised a problem that distribute illegal copied digital multimedia content. Therefore, it was necessary to research the copyright protection technique.

The digital watermarking is one of protection technique for the digital multimedia content, which is to embed the copyright information and additional information into the digital multimedia content. And the embedded information data is called to the watermark.

The Digital audio watermarking is to embed information into the digital audio content. There are three kinds of way for embedding the digital audio watermark signal; spread spectrum coding, echo coding, and phase coding. These embedded watermark signals must be extracted although it comes under various attacks; the A/D-D/A converting, cropping, down mixing, low pass filtering, and so on.

In this thesis, we suggest the new robust watermarking technique against the down mixing attack which is satisfied with SDMI(Secured Digital Music Initiative) the Phase II Screening. The proposed watermarking algorithms are the watermark embedding and extracting algorithm in multi-channel audio data and the extracting algorithm even though multi-channel audio data was down-mixed. The proposed embedding and extracting algorithms have high information embedding efficiency as embedding PN code in each channel using interleave sequence. Also, we propose the algorithm that is able to return to original watermark signals perfectly when the multi-channel audio data applied the proposed algorithm is attacked down-mix.

제 1 장 서론

최근 몇 년 사이에 디지털 기술의 발전과 인터넷의 보급으로 인하여 디지털 멀티미디어 콘텐츠의 사용이 폭발적으로 증가하였고 그 시장도 급성장하고 있다. 초기 디지털 콘텐츠들은 단순히 CD-ROM 과 같은 물리적인 매체를 이용하여 사용자들이 접근을 했지만 현재는 초고속 인터넷의 발전에 힘입어 네트워크에 연결된 컴퓨터만 있다면 언제 어디서나 다양한 형태의 디지털 멀티미디어 콘텐츠를 이용할 수 있게 되었다. 이렇게 미디어 자체의 유동성과 양방향성이 증가하면서 콘텐츠의 질적향상이 이루어졌으며 이용자의 증대로 인해 많은 이용자들이 양질의 콘텐츠를 요구하고 있다. 이러한 이용자의 증대는 이용 형태의 변화를 가져왔다. 초기에 단순 시청형 이었던 이용 형태는 양방향성을 가진 정보 선택형, 정보 요구형을 거쳐 직접 사용자가 콘텐츠를 제작하는 정보 창조형으로 변화했다. 사용자의 욕구는 콘텐츠 공급자들이 양질의 콘텐츠를 제작하도록 하였고 공급자들은 양질의 콘텐츠를 이용하여 일반 사용자들에게 유료로 제공하려는 시도를 하였다. 하지만 디지털 콘텐츠가 가지는 문제점으로 인하여 실제 상업적인 성과를 보지 못하고 있는 것이 현실이다. 디지털 콘텐츠가 가지는 문제점은 다음과 같이 정리 할 수 있다.

1. 디지털화 한 정보에 접근이 용이함
2. 복제가 용이하며 복제비용이 저렴
3. 복제된 디지털 콘텐츠는 손실 없이 원본과 동일
4. 재사용 및 조작성 쉬움
5. 네트워크(인터넷)을 통한 배포가 쉽고 빠름
6. 하나의 콘텐츠가 다양한 매체(방송, 인터넷)에서 사용

디지털 콘텐츠의 장점이라 할 수 있는 이러한 점들이 오히려 디지털 콘텐츠 보호를 어렵게 하는 요인으로 작용하고 있다. 불법 복제된 디지털 콘텐츠의 유통은 온라인, 오프라인 콘텐츠 산업을 위축시키고 콘텐츠의 질을 떨어뜨리고

있으며 냅스터(Napster)나 소리바다와 같은 저작권 분쟁을 일으켰다. 앞으로 서비스가 될 디지털 방송이나 IMT-2000 서비스를 통한 고속 무선 멀티미디어 서비스가 시작하게 되면 콘텐츠의 불법 복제 및 배포와 보안은 가장 큰 문제로 대두될 것임에 틀림없다. 따라서, 디지털 콘텐츠 보호를 위한 표준작업이 필요하게 되었다.

디지털 콘텐츠를 보호하기 위한 기술로는 크게 접속 제어 (Access Control) 사용 제어 (Usage Control), 내용 제어 (Content Control)로 나눌 수 있다.

첫째, 접속 제어는 소유권을 가지지 않은 사용자의 콘텐츠에 대한 접근을 막는 방법으로 인증 기술을 이용하여 접근을 방어한다. 하지만 저작권 보호는 불가능하고 일단 인증키가 풀린 디지털 콘텐츠의 경우에는 제어할 방법이 없으며 권한이 있는 사용자가 이를 불법적으로 배포하고자 한다면 이를 제어하기가 불가능한 단점이 있다.

둘째, 사용 제어는 정당하지 않은 사용자의 콘텐츠에 대한 사용을 막는 방법으로 주로 인터넷 상에서 많이 사용된다. 암호화 키(Key)나 디지털 서명 등을 이용하여 콘텐츠의 사용 행위(재생, 복사 등)의 조절할 수 있지만 실제 현실적으로 복사, 배포 등의 콘텐츠 사용과 관련된 모든 행위를 제어하기가 어렵다. 현재 DRM(Digital Right Management)이라는 암호키 인증 관리 솔루션이 개발 되어 있으며 내용제어 기술인 워터마킹 기술과 함께 강력한 불법복제 방지 솔루션으로 음악, 영상, 정지화상 등의 다양한 디지털 콘텐츠에 적용할 수 있다.

셋째, 내용 제어는 콘텐츠 내용 안에 정보를 은닉하는 방법을 이용하여 저작권 정보를 제공하고 불법적인 복제를 막는 방법으로 디지털 워터마킹(Digital Watermarking)와 같은 방법으로 지각적으로 감지되지 않는 저작권 정보를 콘텐츠 자체에 삽입함으로써 내용제어나 사용제어에서 암호가 풀리더라도 소유권을 주장할 수 있도록 한다. 하지만 신호처리 같은 모든 의도적인 공격에 살아 남아야 한다는 점이 어렵다.

콘텐츠 보호는 이 방법 중 하나만으로 되는 것이 아니라 세 가지의 콘텐츠 보호 방법이 적절하게 결합함으로써 보다 완벽한 콘텐츠 보호를 할 수 있도록

하게 한다. 이중 내용제어 기술에 속하는 디지털 워터마킹은 디지털 콘텐츠 내부에 사용자가 볼 수 없게 또는 들을 수 없도록 정보를 삽입하는 것이다.

원래 워터마크(Watermark)는 종이에 보이지 않는 잉크를 이용하여 글씨를 쓴 비밀편지에서 유래되었다. 아날로그 워터마크는 고대와 중세에 사용되던 일종의 암호 기법으로 종이에 특수한 잉크를 사용하여 보이지 않게 문서를 작성한 후 나중에 특수한 방법으로 그 내용을 확인하는 것을 말하며 상당히 오랜 역사를 가지고 있다. 워터마크의 다른 말인 그리스어 "Steganography"는 Stegano(Covered)와 Graphos(write)라는 말의 합성어로 "숨겨진 글"을 의미한다. 디지털 워터마킹의 기본 목적도 이와 같다고 볼 수 있으며, 다만 그 데이터를 숨기고 추출하는 대상과 방법이 디지털화 되었다고 볼 수 있다.

디지털 워터마킹은 다양한 형태의 디지털 콘텐츠에 대해 연구가 되고 있다. 처음에는 정지화상에 대한 워터마킹 기법들이 연구되기 시작해서 현재는 문서(전자책), 음악, 동영상과 같은 다양한 콘텐츠로 확대되었다. 따라서 다양한 형태의 디지털 콘텐츠를 보호하기 위해 워터마크를 삽입하고 숨김으로써 저작권, 소유자 확인, 소유권 증명, 사용자 인증, 유통경로 추적, 복제횟수 제어, 접근 제어, 방송 모니터링 등에 활용 할 수 있다.

디지털 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 있으므로 두 가지 조건을 만족하는 범위내에서 워터마크를 삽입해야 한다.

최근 DVD(Digital Versatile Disc)의 보급으로 고음질의 멀티채널 오디오 재생이 보편화 되어가고 있다. 그러나 MPEG4 압축 알고리즘을 기반으로 한 DivX 라는 codec 이 등장하면서 DVD 의 영상정보와 음향을 추출하여 인터넷을 통해 불법적으로 유통되고 있는 실정이다. 특히 멀티채널의 음향신호는 대개 2 채널로 다운믹스되어 MP3 형식으로 압축하여 영상신호와 함께 유통되고 있다.

본 논문에서는 미국음반산업협회(RIAA)에서 제안한 SDMI(Secure Digital Music Initiative) Phase II Screening 에 만족하도록 멀티채널 오디오에 워터마크를 삽입하여 다운믹스를 한 경우에도 삽입된 정보를 추출할 수 있는 알고리즘을

제안하고자 한다. 제안한 알고리즘은 워터마크 신호를 인터리브(Interleave)시켜 각 채널에 삽입함으로써 정보 삽입 효율을 높이고 제 3 자에 의한 공격에도 강인하게 하였으며 워터마크된 오디오 신호를 다운믹싱 하였을 때도 워터마크를 검출할 수 있도록 하였다.

본 논문의 제 2 장에서는 기존에 연구되어온 디지털 오디오 워터마킹 기법에 대해서 알아보고 제 3 장에서는 멀티 채널 오디오에 워터마크를 삽입하고 다운믹스 했을 때도 워터마크 신호를 검출 할 수 있는 알고리즘을 제안하였다. 제 4 장에서 제안한 방법의 성능을 고찰하기 위해 컴퓨터 시뮬레이션으로 성능을 평가하였으며 제 5 장에서 결론 및 향후 연구방향을 제시하였다.

제 2 장 디지털 오디오 워터마킹

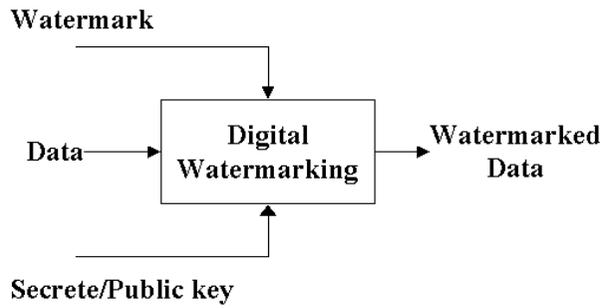
효율적인 오디오 데이터 압축 기술인 MP3 와 인터넷의 대중화 및 고속화에 의해 CD 음질에 필적하는 고음질의 오디오 데이터가 네트워크를 통하여 불법적으로 엄청난 양이 유통되고 있다. 미국의 "넵스터(Napster)"나 우리나라의 "소리바다" 같은 음악파일 공유프로그램에 대한 법정분쟁에서 볼 수 있는 것처럼 네트워크를 통한 디지털 오디오 콘텐츠의 불법유통은 음반산업에 심각한 타격을 주었으며 창작자의 의욕을 떨어뜨렸다. 따라서 디지털 오디오 콘텐츠에 대한 저작권보호 및 불법복제 방지를 위한 기술이 절실히 요구되었다. 따라서 저작권 보호를 할 수 있는 기술로서 디지털 오디오 워터마킹 기법이 연구되기 시작하였다.

디지털 오디오 워터마킹은 오디오 신호 안에 저작권 등의 정보를 들리지 않도록 은닉하는 것을 말한다.

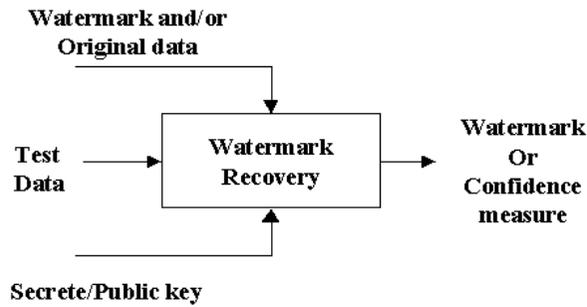
오디오 콘텐츠의 저작권 보호를 위한 워터마킹은 다음과 같은 특성을 가진다.

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

인간의 지각능력은 시각보다 청각이 훨씬 더 민감하고 영상신호에 비해 데이터의 양이 적기 때문에 워터마크를 삽입할 수 있는 영역이 부족하므로 영상신호를 위한 워터마크 기술을 그대로 적용하기 어렵다. 또한, 삽입된 워터마크 신호의 비지각성과 공격에 대한 강인성 두 조건을 만족시키는 것 역시 어렵다. 그 이유는 공격에 강하기 위해서 워터마크는 충분히 검출할 수 있는 에너지 크기를 가져야 하지만 그렇게 한 경우에는 워터마크 신호가 인간의 귀에 들릴 수 있기 때문이다.



(a) Generic watermark embedding scheme



(b) Generic watermark extracting scheme

그림 2-1. 일반적인 워터마킹의 삽입과 복구

Fig. 2-1. Generic watermark scheme.

그림 2-1 일반적인 워터마킹 방법을 나타낸 것이다. 먼저 워터마크 신호는 (2-1)식과 같이 삽입할 정보 I 와 키(Key) 값 k 로 생성한다.

$$W = f_0(I, k) \quad (2-1)$$

이렇게 만든 워터마크 신호는 (2-2)식과 같이 삽입될 원래의 신호 X 에 삽입하여 워터마킹한 데이터 Y 를 만든다.

$$Y = f_1(X, W) \quad (2-2)$$

삽입된 워터마크 신호는 키값 k 를 이용하여 (2-3)식과 (2-4)식처럼 추출 할 수 있다.

$$\hat{I} = g(X, Y, k) \quad (\text{원신호가 필요한 경우}) \quad (2-3)$$

$$\hat{I} = g(Y, k) \quad (\text{원신호가 필요없는 경우}) \quad (2-4)$$

현재 연구되고 있는 디지털 오디오 워터마킹 방법은 크게 3 가지로, 위상 정보 삽입 기법(Phase coding), 대역 확산 기법(Spread spectrum coding), 반향 삽입 기법(Echo coding) 이 있다.

제 2-1 절 위상 정보 삽입 기법 [2]

HAS(Human Auditory System)이 작은 위상 변화에 둔감하다는 특성을 이용하여 각 프레임별로 위상을 변화시킴으로써 워터마크를 삽입하는 방법이다.

워터마크를 삽입 할 원래의 오디오 신호를 $s(i)$ ($0 \leq i \leq I-1$)라 할 때, 위상 정보를 삽입하는 방법은 다음과 같다.

- 1) 오디오 신호 $s(i)$ ($0 \leq i \leq I-1$)를 N 개의 세그먼트로 잘라낸다. 이 때 잘려진 오디오 신호의 세그먼트를 $s_n(i)$ ($0 \leq n \leq N-1$) 이라 한다. (그림 2-1 (a)(b))
- 2) n 번째 세그먼트 $s_n(i)$ 를 K -point 이산 푸리에 변환을 한다.($K=I/N$) 그리고 위상 행렬(phase matrix) $\phi_n(\omega_k)$ 와 크기 행렬(Amplitude matrix) $A_n(\omega_k)$ 값을 구한다. ($0 \leq k \leq K-1$) (그림 2-1 (c))
- 3) $0 \leq n \leq N-1$ 에 대하여 각각 인접한 세그먼트 사이의 위상차를 구한다.

$$\Delta\phi_{n+1}(\omega_k) = \phi_{n+1}(\omega_k) - \phi_n(\omega_k) \quad (2-5)$$

- 4) 삽입할 부가정보 데이터 0 과 1 을 $\phi_{data} = \pi/2$ 또는 $-\pi/2$ 로 나타낸다. (그림 2-1 (e))

$$\phi'_0 = \phi'_{data} \quad (2-6)$$

- 5) 위상차를 이용하여 $n>0$ 에 대하여 위상행렬을 재구성한다.

$$\left[\begin{array}{l} \phi'_1(\omega_k) = \phi'_0(\omega_k) + \Delta\phi'_1(\omega_k) \\ \Lambda \\ \phi'_n(\omega_k) = \phi'_{n-1}(\omega_k) + \Delta\phi'_n(\omega_k) \\ \Lambda \\ \phi'_N(\omega_k) = \phi'_{N-1}(\omega_k) + \Delta\phi'_N(\omega_k) \end{array} \right] \quad (2-7)$$

- 6) 수정된 위상행렬 $\phi'_n(\omega_k)$ 와 원래의 크기행렬 $A_n(\omega_k)$ 를 이용하여 역이산 푸리에 변환을 수행하여 오디오 신호를 복원한다. (그림 2-1 (g))

그러나 이 방법은 오디오 신호가 안정적이어야 한다는 조건이 있으며 실제 오디오 신호는 안정적이지 않은 경우가 많기 때문에 음질에 많은 변화를 준다.

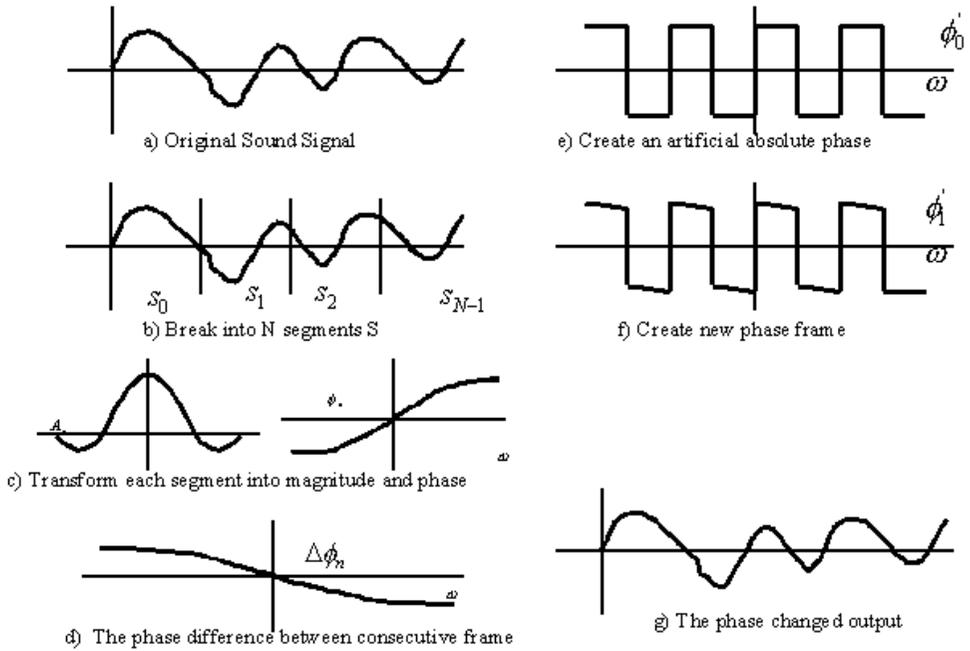


그림 2-2. 위상 정보 삽입 기법 과정

Fig. 2-2. Phase coding schematic.

제 2-2 절 대역 확산 기법

오디오 신호를 프레임 별로 나누어 각 프레임에 워터마크 신호로 PN(Pseudo-Noise)코드를 삽입하는 방법이다. 이는 PN 코드가 가지는 백색잡음과 유사한 통계적 특성과 좋은 자기상관(Auto-Correlation) 특성을 이용하는 것이다. 특히 자기상관 특성이 좋은 점은 PN 코드가 작은 에너지로 삽입되더라도 다른 신호에 비해 상관도가 상대적으로 큰 값을 가지기 때문에 검출 할 때 용이하다. 워터마크로 사용한 PN 코드값은 그 자체가 키(Key) 값이 삽입 할 때 사용한 PN 코드 값을 알지 못하면 워터마크를 검출 할 수 없게 된다.

하지만 PN 코드의 삽입은 오디오신호에 잡음을 추가하는 것이므로 검출이 가능한 한도 내에 그 에너지가 들리지 않도록 삽입해야 한다.

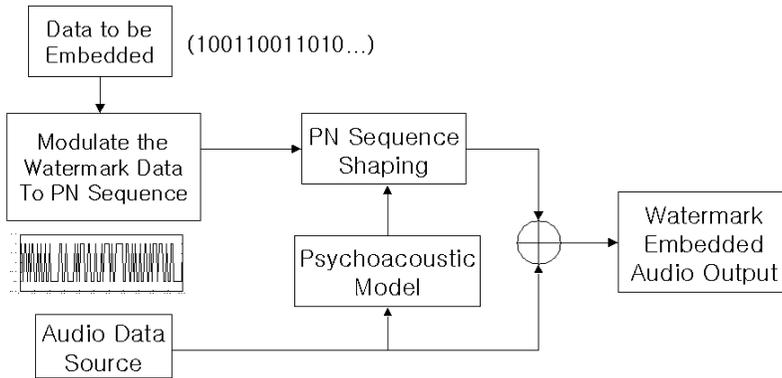
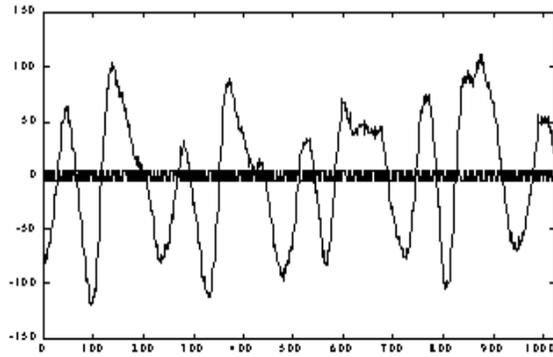


그림 2-3. 대역확산 워터마크 삽입 과정

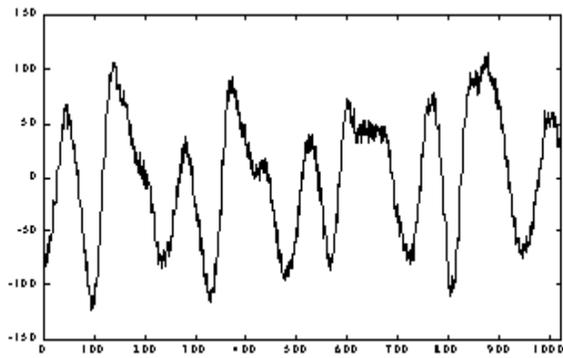
Fig. 2-3. The watermark embedding scheme using spread spectrum.

대역확산 기법을 이용한 워터마크의 삽입과정은 그림 2-3 과 같다.

워터마크 삽입 과정을 정리하면, 먼저 입력되는 오디오 신호를 프레임 단위로 나눈다. 그 후 각 프레임에 사람의 청각적 특성에 따라 작은 소리는 큰소리에 묻혀 들리지 않는다는 마스킹(Masking) 현상을 이용한 심리 음향 모델(Psychoacoustic model)을 적용하여 마스킹 레벨을 결정한다. 이 마스킹 레벨에 맞추어 삽입하려는 데이터에 의해 생성된 PN 코드를 변형시켜 오디오 신호에 삽입한다. 이렇게 삽입 된 심리음향의 마스킹 레벨에 의해 변형된 PN 코드는 마스킹 레벨 아래의 크기를 가지기 때문에 사람의 귀에는 들리지 않게 된다.



(a) 오디오 신호와 PN 코드 워터마크
 (a) Audio signal and PN code watermark.



(b) 삽입된 워터마크신호
 (b) Watermarked audio signal.

그림 2-4. 변형 삽입 된 PN 코드 워터마크 (시간영역)

Fig. 2-4. The watermark embedded PN code.

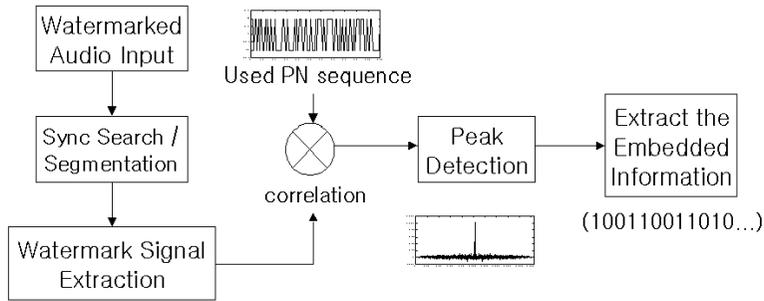
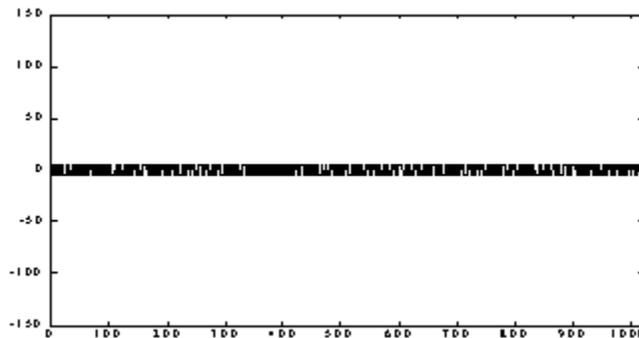


그림 2-5. 대역 확산 워터마크 검출과정

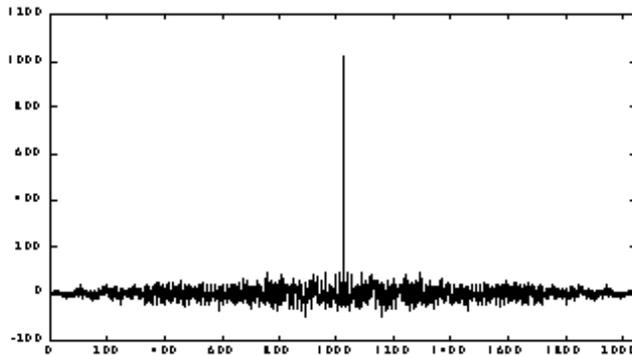
Fig. 2-5. The watermark extracting scheme using spread spectrum.

워터마크의 검출 과정은 그림 2-5 와 같이 워터마크가 삽입된 오디오 신호가 입력되면 삽입할 때 사용한 PN 코드와 삽입한 오디오의 프레임의 삽입 동기점을 찾아 내서 워터마크를 추출한다. 삽입한 워터마크 신호인 PN 코드를 복원하는 방법은 원래의 오디오 신호를 필요로 하는 널블라인드(Non-blind) 방법과 원래의 오디오 신호가 필요없는 블라인드(Blind) 방법이 있다. 이렇게 복원한 PN 코드와 키(Key)값으로 가지고 있는 PN 코드와 상관도를 이용하여 피크(Peak)값을 검출하여 삽입한 정보를 추출해낸다.



(a) 검출한 워터마크 신호

(a) Extracted watermark signal.



(b) 워터마크 신호와 사용한 PN 코드와의 상관도
 (b) Correlation of watermark signal and used PN code.

그림 2-6. PN 코드 워터마크의 검출
 Fig. 2-6. The Extraction of PN code watermark.

제 2-3 절 반향 신호 삽입 기법

프레임별로 나누어진 오디오 신호에 워터마크 신호로 인위적인 반향 신호를 삽입하는 방법이다. 적절하게 삽입한 반향 신호는 사람에게 왜곡으로 인지되지 않으며 오히려 음색을 풍부하게 해주는 효과를 가지고 있다. 따라서 삽입하는 반향 신호의 크기와 지연시간을 조절하면 원음과 청각적으로 구분되지 않는 반향 신호를 삽입할 수 있다. 또한 이렇게 삽입된 워터마크 신호는 원 신호 없이 복호화 할 수 있는 장점이 있으나 사용 권한이 없는 제 3 자에 의한 워터마크 추출이 쉽다는 단점이 있다.

반향 신호는 그림 2-7 과 같은 충격응답 모델로 모델링을 할 수 있다. 이 때 반향의 특성을 결정하는 요소는 초기진폭(Initial amplitude), 감쇄율(Decay rate), 오프셋(Offset), 델타(Delta) 등이 있다. 초기 진폭은 첫번째 반향이 갖는 크기이고 감쇄율이란 두 번째 반향 신호로부터 적용되는 초기 진폭으로 부터의 상대적인

크기의 비율을 나타낸다. 오프셋은 원래 신호로부터 첫번째 반향까지의 시간 지연을 의미하고 이후의 감쇄되는 반향 신호 사이의 간격을 델타로 정의한다.

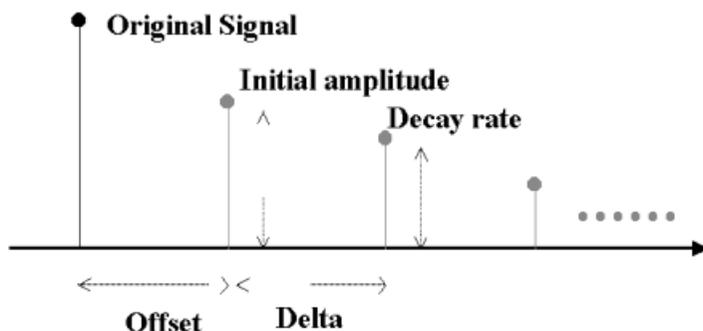


그림 2-7. 반향 신호의 충격 응답

Fig. 2-7. The impulse response of echo signal.

오디오 신호에 반향 신호를 삽입하는 것은 그림 2-7 과 같은 충격 응답 함수에 통과시키는 것과 같다. 충격 응답 함수에 오디오 신호를 통과시킴으로써 반향신호를 생성한다. 반향 충격 응답 함수의 특성을 보면 오프셋이 200msec 이상이 되면 반향 신호가 사람의 귀로 인지할 수 있게 되며, 10msec 이하가 되면 반향은 생기지 않고 주파수응답의 영향에 의해 착색현상(Coloration)이라는 음색(Timbre)의 변화가 인지된다. 따라서 10~200msec 내에서 오프셋을 결정해야 한다.

반향 신호를 워터마크 신호로 삽입하기 위해서는 그림 2-8 과 같이 우선 삽입할 내용의 데이터의 이진수 '0'과 '1'에 대하여 각각 다른 오프셋을 가지는 충격 응답 함수를 생성하고, 프레임별로 한 비트(Bit)씩 정보를 할당한 뒤, 그림 2-8 과 같이 그에 해당하는 충격 응답 함수를 통과시켜 워터마크 정보를 삽입한다.

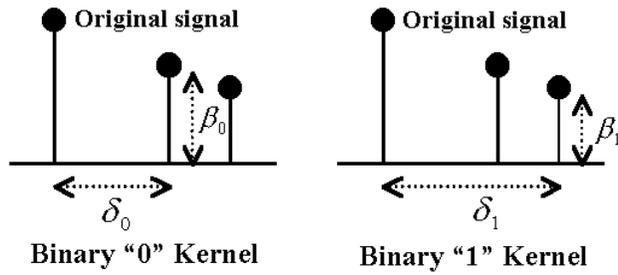


그림 2-8. '0'과 '1'에 대한 반향 커널

Fig. 2-8. The echo kernel for information data '0' and '1'.

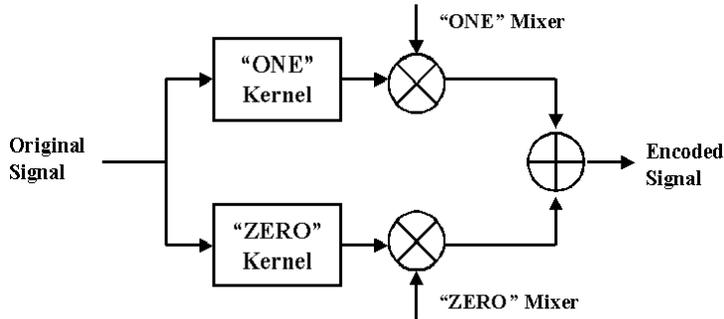


그림 2-9. 반향 신호 워터마크 삽입 인코더

Fig. 2-9. The echo embedding watermark encoder.

반향신호를 삽입한 워터마크를 검출하는 방법은 각 프레임에서 반향신호의 오프셋 값을 찾아내는 것이다. 오디오 신호에서 삽입된 반향신호의 시간 지연을 구할 수 있는 가장 좋은 방법은 켈프스트럼(Cepstrum)의 자기상관(Autocorrelation)을 구하는 것이다. 일반적으로 오디오 신호에 대한 켈프스트럼은 0의 값에 많은 에너지가 몰려있는 형태로 나타나고 반향 충격 응답은 오프셋 및 그의 배수에 해당하는 위치에 지수적으로 감쇄하는 주기적인 펄스가 존재하는 이산지수(Discrete time exponential)함수의 형태를 나타낸다. 반향신호가 첨가

되어 있는 오디오 신호의 cepstrum은 이 두 cepstrum의 합 형태로 나타나므로 오프셋의 배수 위치에 주기적으로 피크(Peak) 값을 가지는 파형을 가지게 된다. 따라서 cepstrum 신호에 자기상관함수(Autocorrelation)을 취하게 되면 반향이 존재하는 신호의 경우 오프셋 위치에서 상대적으로 큰 피크 값을 검출할 수 있으므로 이를 통해 워터마크 신호를 검출할 수 있다.

cepstrum에 대한 자기상관함수 연산을 오토 cepstrum(Autocepstrum)이라고 정의하며 자기상관과 파워 스펙트럼(Power spectrum)사이의 관계에 의해 (2-8)식과 같이 간단히 계산할 수 있다.

$$\tilde{x}(n) = F^{-1}((\log(F(x(n))))^2) \quad (2-8)$$

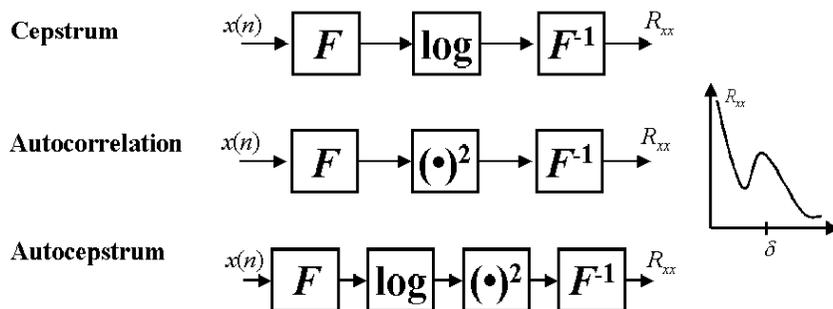
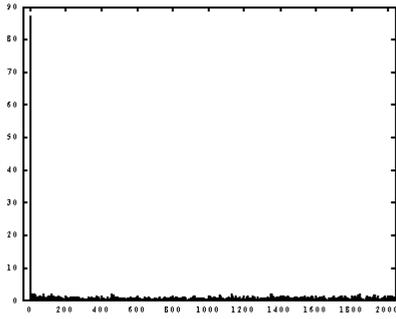
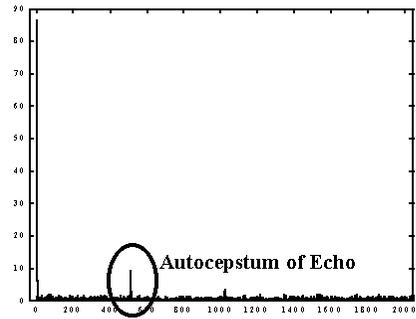


그림 2-10. 오토 cepstrum을 이용한 워터마크 검출

Fig. 2-10. Extraction of echo signal watermark using autocepstrum.



(a) 원신호의 오토 켈스트럼
 (a) Autocepstrum of original signal.



(b) 워터마크된 신호의 오토 켈스트럼
 (b) Autocepstrum of watermarked signal.

그림 2-11. 반향 신호 워터마크의 검출
 Fig. 2-11. Extraction of echo signal watermark.

제 3 장 다운믹싱에 강한 디지털 오디오 워터마킹

오디오 정보는 영상 정보에 비해 상대적으로 정보량이 적으므로 상대적으로 영상에 비해 삽입할 수 있는 워터마크 정보가 적다. 하지만, 일반적으로 상업적인 용도로 만들어지는 오디오 신호는 CD(Compact Disk)를 이용하는 소위 스테레오라고 말하는 2 채널 오디오이고, 차세대 오디오 매체로서 각광 받고 있는 DVD-Audio (Digital Versatile Disc Audio)나 SACD(Super Audio Compact Disk)는 멀티채널을 지원하고 있다. 이러한 2 채널 이상의 멀티 채널 오디오 음원에 대해서는 채널별로 워터마크 신호를 삽입하여 삽입할 수 있는 정보량을 확장할 수 있으며 또한 다양한 부가 정보를 삽입할 수 있다.

하지만 멀티채널에 워터마크를 삽입하였을 경우, 각 채널별로 나누어진 신호를 하나 또는 두개의 채널로 합치는 다운믹싱을 할 경우 각 채널 별로 삽입된 워터마크 신호역시 다운믹싱이 되므로 워터마크를 검출 할 수 없게 된다. 디지털 오디오 워터마킹에 관한 규격을 제안하는 단체인 SDMI(Secured Digital Music Initiative)에서도 신호 처리 공격에 대한 강인성에 대한 규정 중에 다운믹싱 공격에 대해서도 언급하고 있다.

본 연구는 DSSS(Direct Sequence Spread Spectrum) coding 을 기반으로 하는 워터마킹 기법을 이용하여 멀티채널 오디오에 워터마크를 삽입하고, 다운믹싱 하였을 때도 워터마크를 검출 할 수 있는 알고리즘을 제안하였다.

제 3-1 절 워터마크 삽입 과정

워터마크의 삽입과정은 그림 3-1 와 같이 삽입하려는 정보 데이터가 '0', '1' 일 때 각각 다른 PN 코드를 선택하여 각 채널에 삽입한다. 채널별로 삽입할 때 랜덤하지만 채널 수만큼의 숫자를 같은 수로 가진 수열 값에 의해 PN 코드를 각 채널에 인터리브(Interleave)하여 삽입한다.

워터마크를 삽입하기 위해 프레임 별로 나누어진 m 개의 채널을 가지는 오디오

오 신호 $x(n)$ 에 대하여 삽입되는 PN 코드 워터마크 신호를 $w(n)$ 이라 두면 워터마크가 삽입된 오디오 신호 $y(n)$ 은 다음과 같다.

$$y(n) = x(n) + \alpha w(n) \quad (3-1)$$

여기에서 α 는 워터마크의 크기이다. 각 채널에 대한 신호는 $x_1(n)$, $x_2(n)$, \dots , $x_m(n)$ 라 하고, 삽입 할 워터마크 중 '0'에 대한 값을 $w_0(n)$, '1'에 대한 값을 $w_1(n)$ 이라 한다면, 워터마크의 삽입은 Information data I 의 값에 따라 (3-2)식과 같이 표현할 수 있다.

$$\begin{cases} y(n) = x(n) + \alpha w_0(n) & (I = 0) \\ y(n) = x(n) + \alpha w_1(n) & (I = 1) \end{cases} \quad (3-2)$$

Information data I 에 의해 선택된 워터마크 신호는 m 개의 채널 수와 같은 1 에서 m 의 숫자를 동일하게 가지는 인터리브 시퀀스(Interleave sequence)에 의해 PN 코드 $w(n)$ 의 칩(Chip) 값 하나하나를 채널별로 삽입한다.

이렇게 삽입 함으로서 채널이 많을수록 더 많은 정보를 삽입할 수 있으며 따라서 다양한 부가정보를 제공할 수 있다. 또한 워터마크 신호를 인터리브 시퀀스(Interleave sequence)에 따라 삽입하므로 일반적인 방법으로는 복호를 할 수 없으므로 제 3 자의 공격에 대한 강인성을 가지게 된다.

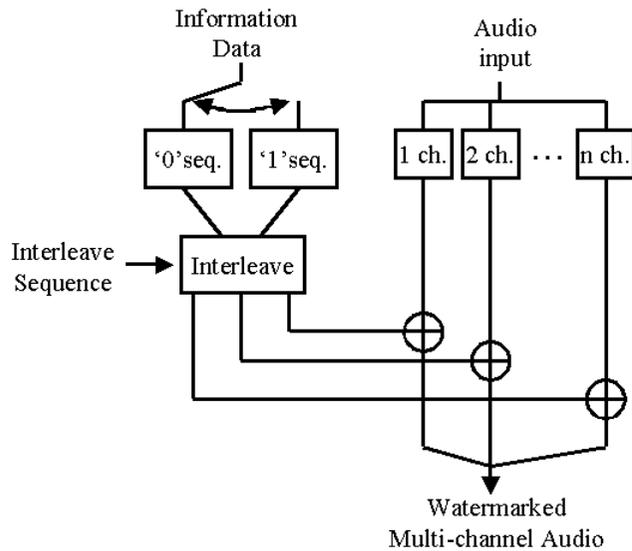


그림 3-1. 멀티 채널 오디오의 워터마크 삽입

Fig. 3-1. The watermark embedding scheme for Multi-channel audio.

제 3-2 절 워터마크 추출 과정

각 채널에 인터리브(Interleave)하여 삽입한 워터마크를 추출하기 위해서 삽입 시 인터리브한 순서가 필요하다. 워터마크한 멀티채널 오디오신호가 입력되면 인터리브 시퀀스(Interleave sequence)의 순서에 의해 PN 코드를 디인터리브(Deinterleave)하여 워터마크 정보를 추출하고 원래 삽입한 PN 코드와 상관도를 구해서 삽입한 정보를 복원할 수 있다.

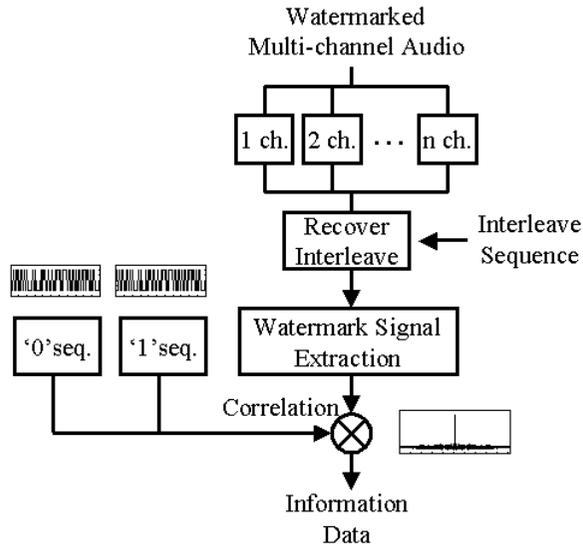


그림 3-2. 멀티 채널 오디오에 삽입된 워터마크 추출
 Fig. 3-2. The watermark extracting scheme from multi-channel audio.

제 3-3 절 다운 믹싱한 오디오 신호의 워터마크의 검출

멀티 채널 오디오 신호를 다운믹스를 한 경우에는 각 채널 별로 삽입한 워터마크 신호 역시 다운믹스 되므로 일반적인 워터마크 추출 알고리즘으로는 워터마크를 추출할 수 없다. 따라서 다운믹스 되었을 경우 그 여부를 파악하고, 다운믹스 한 오디오 신호라는 것을 확인하였을 때 다운 믹스된 오디오신호에서 PN 코드를 복구하여 워터마크를 추출할 수 있는 알고리즘을 수행한다.

그림 3-3 과 같이 다운믹스 한 워터마킹 오디오 신호가 입력 되면 인터리브 시퀀스(Interleave sequence)에 의해서 삽입한 워터마크와 상관 값을 구하기 위한 키 값으로 가지고 있는 PN 코드를 이용하여 채널 별로 삽입한 워터마크 신호를 재구성한다. 다운믹스한 워터마킹 오디오 신호는 가상의 m 개의 채널로 만들어 재구성한 워터마크 값을 이용하여 한 개의 채널의 오디오 신호에 다른채널의 재구성한 워터마크 값을 빼줌으로서 구할 수 있다.

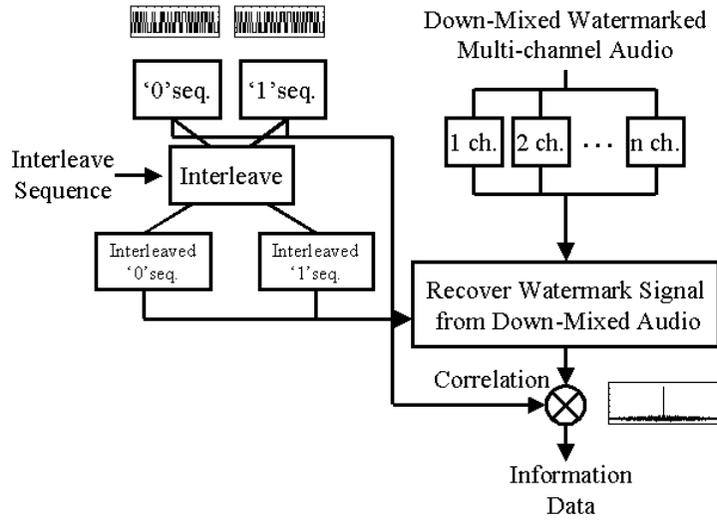


그림 3-3. 다운믹스한 오디오신호에서의 워터마크 검출

Fig. 3-3. The watermark extracting algorithm for down-mixed audio signal.

제 4 장 실험 결과

제 4-1 절 SDMI(Secured Digital Music Initiative) Phase II Screening [10]

SDMI(Secured Digital Music Initiative)는 미국 음반 산업 협회(RIAA)에 설립한 단체로 MP3 등과 같은 고음질, 저용량의 디지털 오디오 콘텐츠의 불법 유통을 막고 음반 시장을 지키기 위해 설립하였다. SDMI 에서 디지털 오디오 워터마킹 알고리즘에 대한 규격을 발표하고 전세계 기업 등을 대상으로 공모하였다. 1999 년에 phase I screening 에 대해 공모결과 5 개 업체가 참여하여 성능을 테스트한 결과 ARIS Technology 사의 알고리즘이 채택되었다. 현재는 phase II screening 을 진행 중에 있는데 Phase II screening 은 2000 년 2 월에 제안한 규격으로 멀티채널 오디오 포맷을 가진 DVD-audio 같은 차세대 오디오 미디어를 위한 오디오 워터마킹 규격을 정하고 있다.

Phase II screening 에서의 워터마킹 알고리즘의 성능 평가를 위한 기술 요구 사항을 정리하면 다음과 같다.

1. 비지각성 (Inaudibility)

청취자가 삽입한 워터마크 신호를 들 수 없어야 한다는 특성이며 소위 골든 이어(Golden ears)라고 불리는 청각이 뛰어난 사람들과 함께 청취 실험으로 확인.

2. 강인성 (Robustness)

신호처리에 대한 강인성에 대한 규정

a) 변하지 않아야 할 특성 (Persistent Assertion)

표 4-1 의 규정에 따른 다양한 신호처리에 도 워터마크를 검출 할 수 있어야 한다.

표 4-1. 신호처리에 대한 워터마크의 강인성에 관한 규정

Table. 4-1. The Screening about robustness of watermark for signal processing.

Signal Processing	Description
D/A, A/D	D/A, A/D, converting twice
Equalization	Typical case : 10-band graphic equalizer with the following characteristics Freq.[Hz] 31 62 125 250 500 1k 2k 4k 8k 16k Gain[dB] -6 +6 -6 +6 -6 +6 -6 +6 -6 +6
Band-pass filtering	100Hz ~ 6kHz, 12dB/oct
Linear speed change	+/- 10%
Codecs (at typically used rates)	ISO/IEC 13818-7:1997 ("AAC") ISO/IEC 14496-3:1999 (MPEG-4 AAC with Perceptual noise substitution) ISO/IEC 11172-3:1993 Layer III (MPEG-1 Audio Layer 3 "MP3") Q-Design Window Media Audio Twin-VQ ATRAC-3 Dolby Digital AC-3 ATSC A_52 ePAC
Noise addition	Adding white noise with constant level of 36dB lower than total averaged music power (S/N: 36dB)
Time scale modification	Pitch-invariant time scaling: +/- 4%
Wow and flutter	0.5% rms, from DC to 250Hz
Addition echo	Maximum delay: 100ms Feedback coefficient: up to 0.5
Down mix and	6 channel to stereo

Surround sound processing	SRS Spatializer Dolby Surround Dolby Headphone
Sampling Rate Conversion	48kHz to 44.1kHz 96kHz to 48/44.1kHz
Dynamic Range Reduction	Threshold: 50dB 16dB max compression Rate: 10ms attack, 3s recovery

b) 일반 사용자들의 신호처리 (Typical Consumer Signal Process)

표 4-2 와 같은 규격을 만족하도록 일반 사용자들의 신호처리를 했을 때에도 워터마크를 검출 할 수 있어야 한다.

표 4-2. 일반 사용자들의 신호처리

Table. 4-2. The signal processing of typical consumer.

Signal Processing to be applied to 44.1/48kHz 16bit content	Description
D/A, A/D	D/A followed by A/D converting once on typical consumer PC sound hardware w/ 60 dB SNR on full scale signals
Equalization	Typical case : 10-band graphic equalizer with the following characteristics Freq.[Hz] 31 62 125 250 500 1k 2k 4k 8k 16k Gain[dB] -3 +3 -3 +3 -3 +3 -3 +3 -3 +3
Down Mixing and Surround Sound Processing	Multi-channel to stereo Stereo to Mono

Sample Rate	48 kHz to 44.1 kHz
Conversion	96 kHz to 48/44.1 kHz

3. 고의적인 공격 (Malicious Attacks)

다음과 같은 고의적인 사용자들의 공격에 강해야 한다.

- a) 전체 지연(Group delay), 역재생(Inversion), LPC(Linear predictive coding)과 같은 신호처리 공격
- b) 워터마크 된 오디오 신호를 평균화(Averaging), 잘라낸 신호의 연결(Splicing), 믹싱(Mixing), 일부신호만 잘라냄(Subtraction)을 하는 사용자들이 공모하는 공격(Collusion attack)

4. 신뢰성 (Reliability)

워터마크 검출시 에러발생이 적어야 한다.

5. 갱신성 (Renewability)

워터마크가 공격에 의해 깨어지더라도 복구 할 수 있어야 한다. 알고리즘 제안자는 공격이 성공하더라도 복구할 수 있음을 보여야 한다.

6. 수행 효율 (Efficiency of Operation)

다음과 같은 플랫폼에서 수행하였을 때 표 4-3 에 제시한 시간 내에 검출 할 수 있어야 한다.

- a) 플랫폼 1 : 400MHz Pentium III, 64K cache, 64MB RAM, 100Mhz frontside bus, 5400 RPM disk, Windows 2nd Edition Operating System.
- b) 플랫폼 2: Texas Instrument C54x(16bit fixed point DSP)
- c) 플랫폼 3 : ARM 720 processor (32bit fixed point RISC processor)

표 4-3. 워터마크 삽입 및 검출 제한 조건

Table 4-3. The limited conditions of watermark embedding and extracting.

	Platform 2	Platform 3
Computational Power	12MIPS	9MHz
Data RAM	6KB	2KB
Data ROM	512B	8KB
Instruction Memory	8KB	8KB

제 4-2 절 실험 방법 및 결과

제 4-1 절에서 SDMI 가 제시한 조건 중 다운믹싱 한 신호에서의 워터마크 추출은 6 채널을 스테레오로 또는 멀티채널을 스테레오로, 또는 스테레오를 모노로 했을 때를 제시하고 있다. 제안된 알고리즘을 이용한 시뮬레이션을 위해 위 조건 중 Typical Consumer Signal Processing 의 Stereo-to-Mono 다운믹스에 대하여 소프트웨어 기반으로 간단한 성능 테스트를 하였다.

음원은 10 초 정도의 길이를 가지는 CD(Compact disk)에서 추출한 CD 음질 (44100Hz, 16bit)의 테크노, 클래식, 록 장르의 음악을 사용하였다. PN 코드의 1 프레임의 길이는 1024 샘플이지만 2 개의 채널로 인터리브를 시켜 삽입하므로 채널당 1 프레임은 512 개의 샘플을 가지게 되어있다. 삽입된 정보데이터의 길이는 860 비트이고 워터마크는 시간영역에 삽입하였다

워터마크 에너지는 충분히 사람의 귀에 들리지 않도록 하기 위해 (4-1)식을 이용하여 SNR 값을 구하였다.

$$SNR = 10 \log_{10} \left(\frac{\text{audio power}}{\text{watermark power}} \right) \quad (4-1)$$

(4-1)식을 이용하여 각 음원에 SNR (Signal - to - Noise Ratio)은 각각 30, 50, 70dB 의 에너지값을 가지도록 하여 PN 코드 워터마크를 삽입하였다.

성능을 검증하기 위해 인터리브를 이용한 워터마크가 삽입된 Stereo 음원을 Mono 로 다운믹스하고, 그 음원을 제안한 워터마크 추출 알고리즘을 이용하여 검출하고 정보데이터를 복구하였다. 검출 알고리즘은 워터마크 검출 시에 원래의 신호를 필요로 하는 널블라인드(Non-Blind) 검출방법을 사용하였으며 (4-2)식을 이용하여 삽입한 워터마크와 검출한 워터마크의 값을 비교하여 BER(Bit Error Rate) 값을 계산하였다.

$$BER = \frac{100}{B} \sum_{n=0}^{B-1} \begin{cases} 1, & \tilde{w}(n) \neq w(n) \\ 0, & \tilde{w}(n) = w(n) \end{cases} \quad (4-2)$$

여기서 B 는 삽입한 PN 코드 워터마크의 길이이고 $w(n)$ 은 삽입한 워터마크 신호, $\tilde{w}(n)$ 은 추출한 워터마크 신호이다.

표 4-4. 다운믹스 한 워터마킹 오디오를 이용하여 일반방법과 제안된 방법으로 워터마크 신호를 검출한 결과

Table. 4-4. The results of extracting watermark using down-mixed watermarked audio.

Sound source	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 볼 수 있는 것처럼 워터마킹한 오디오 신호가 다운믹스가 되었을 때도 일반적인 추출 알고리즘은 사용한 모든 음원에 대해 45 퍼센트를 넘는 많은 에러 값을 가지는데 비해 제안한 알고리즘은 완벽하게 워터마크를 검출하고 정보를 복원하는 것을 알 수 있다.

제 5 장 결 론

디지털 워터마킹은 콘텐츠 보호의 마지막 방어선으로 활발한 연구가 진행되고 있다. 콘텐츠에 워터마크를 삽입함으로써 내용기반 콘텐츠 검색 등과 같은 다양한 부가적인 서비스와 콘텐츠의 유통의 투명성을 보장할 수 있다.

본 논문은 SDMI Phase II Screening 에서 제안하는 규격 중, 채널 다운믹스에 강한 알고리즘을 제안하였다. 본 알고리즘은 대역 확산 워터마킹 기법을 기반으로 멀티채널 오디오에 워터마크를 삽입할 때 인터리브 시퀀스(Interleave sequence)에 의한 각 채널별로 인터리브(Interleave)해서 삽입함으로써 정보삽입효율을 높이고 제 3 자에 의한 공격을 어렵게 하였고, 워터마킹 한 오디오 신호를 다운믹스한 경우에도 워터마크를 검출하여 삽입된 정보를 복구할 수 있도록 하였고 이를 실험으로 검증하였다.

현재 일반적인 멀티채널 재생 방식의 오디오의 경우 Dolby Digital 이나 DTS 같은 재생 시스템이 대부분이며 이러한 방식의 멀티채널 재생 오디오 시스템은 멀티채널 재현뿐만 아니라 정보를 효율적으로 매체에 저장하기 위한 압축 알고리즘도 가지고 있다. 또한 Dolby Headphone 과 같은 멀티채널 오디오 신호를 스테레오로 다운믹스하며 멀티채널 음향효과를 재현하는 시스템도 있다. 따라서 압축과 멀티채널 다운믹스 입체음향 시스템, 각종 신호처리 공격 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*, pp. 473-480, 1996.
- [2] W. Bender, D. Gruhl, N. Morimoto, A. Lu, "Techniques for data hiding," *IBM System Journal*, Vol 35, Nos 3&4, pp. 313-336, 1996.
- [3] M. Swanson, M. Kobayashi, and A. Tewfik, "Multimedia data-embedding and watermarking technologies," *Proc. IEEE*, vol. 86, pp. 1064-1087, June 1998.
- [4] M. Swanson, B. Zhu, A. Tewfik, and L. Boney, "Robust watermarking using perceptual masking," *Signal Processing*, pp. 337-355, 1997.
- [5] F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proc. IEEE*, vol. 87, pp. 1079-1107, July 1999.
- [6] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoan, "A Secure, robust watermark for multimedia," *Proc. Workshop on Information Hiding*, Univ. of Cambridge, U.K., May 30-June 1, 1996.
- [7] C. Xu and *et al.*, "Applications of Digital Watermarking Technology in Audio Signals," *J. Audio Eng. Soc.*, Vol. 47, No.10, pp. 805-812, Oct. 1999.
- [8] P. Bassia and I. Pitas, "Robust audio Watermarking in the time domain," *Proc. EUSIPCO 98*, vol. 1, pp. 22-28, Rhodes, Greece, 1998.
- [9] J. D. Gordy and L. T. Bruton, "Performance Evaluation of Digital Audio Watermarking Algorithm," *MWSCAS2000*, pp. 456-459, Detroit, MI, USA, 2000.
- [10] SDMI(Secure Digital Music Initiative), "Call for Proposals for Phase II Screening Technology Version 1.0," 2000.

- [11] 김현욱, 오현오, 석종원, 홍진우, 윤대회, "강인한 오디오 워터마크의 블라인드 검출을 위한 PN 시퀀스 삽입 및 추출에 관한 연구", *2001년도 한국음향학회 학술발표대회 논문집*, 제 20 권 제 1 호, pp. 365-368, 2001.
- [12] 오현오, 김현욱, 윤대회, 차일환, "반향신호를 이용한 오디오 워터마킹에 관한 연구", *제 13 회 신호처리합동학술대회 논문집*, 제 13 권 1 호, pp. 369-372, 2000.
- [13] 김현욱, 오현오, 윤대회, 차일환, "공격에 강인한 고음질 디지털 오디오 워터마킹 시스템에 관한 연구", *제 13 회 신호처리합동학술대회 논문집*, 제 13 권 1 호, pp. 771-774, 2000.
- [14] 오승수, 김기만, "다운믹싱에 강한 디지털 오디오 워터마킹 기법에 대한 연구", *한국해양정보통신학회 추계학술대회*, pp781-784, 2001.

감사의 글

지난 2 년 동안 저에게 학문적으로 뿐만 아니라 인간으로도 성숙 할 수 있게, 그리고 소리에 대한 호기심을 호기심에서 끝나지 않도록 지도 해 주신 지도교수 김기만 교수님께 진심으로 감사드립니다. 또한 학부시절부터 지도해 주시고 격려해 주신 김동일 교수님, 조형래 교수님, 강인호 교수님, 정지원 교수님, 민경식 교수님께도 감사드립니다.

실험실에서 생활하며 같이 부대끼면서 고생하고 울고 웃었던 인식, 용주, 승용이, 아직 갈 길이 멀기만 한 것 같은 외형, 윤준이, 충고와 격려를 아끼지 않았던 지금은 군복무 중인 동현, 상용, 장가가서 행복하게 살고 있는 동원, 김준환 형에게 고마움을 전합니다.

태어나서 처음으로 소리에 대한 호기심을 주셨던, 항상 격려로 지금까지 공부할 수 있도록 버팀목이 되어주신 아버지 어머니, 음향에 대해 많은 도움을 준 동생에게 이 작은 결실을 바칩니다.