

디지털 오디오의 인터넷 방송 서비스를 위한 저작물 식별 코드 은닉 기술*

신승원, 김종원, 최종욱
(주) 마크애니 연구소
서울시 중구 쌍림동 151-11

A Unique Identification hiding technique for Internet Broadcast service of Digital Audio

Seungwon Shin, Jongweon Kim, and Jonguk Choi
MarkAny Research Institute
E-mail: swshin@markany.com

요약

본 논문은 인터넷을 이용한 방송 서비스에서 저작권 보호 및 저작물 식별을 위한 고유 식별 정보(unique identification)를 삽입, 추출하는 기술을 제안한다. 인터넷을 통한 방송 서비스의 활성화를 위해서는 저작권의 보호와 저작물의 불법 사용 방지, 저작물에 대한 저작권료 지불, 그리고 차후에 불법 사용자의 적발 등을 할 수 있는 기술이 필요하다.

본 논문에서 제안하는 기술은 디지털 음악을 소비자에게 서비스할 때, 사용자의 아이디(ID)와 신상 정보를 이용해서 생성된 정보와 음악의 식별 정보를 같이 저작물에 삽입하여 다운로드 받은 음악 파일이나 또는 청취 중에 녹음한 음악 파일을 불법으로 유통시킬 경우에 유통된 음악 파일에서 삽입된 식별코드를 추출하여 음악 파일의 불법 유통경로를 추적하여 불법 사용자를 적발할 수 있는 기능을 갖고 있다. 이러한 기능을 만족시키기 위해서 인터넷 서비스에서 널리 이용되는 MP3, AAC, WMA 등과 같은 손실 압축을 거친 이후에도 삽입된 워터마크를 검출할 수 있으며, 일반 사용자들이 손쉽게 접할 수 있는 여러 신호처리에도 강인한 특성을 갖는 기술을 제안한다.

1. 서론

디지털 워터마킹 기술은 디지털로 제작된 저작물에 저작권의 고유 정보를 삽입하여 자신의 권리를 추후에 주장할 수 있도록 하는 기술로 온라인 멀티미디어 시장에 핵심적인 기술로 자리잡고 있다. 워터마킹 기술과 함께 멀티미디어 저작물 보호를 위해서 암호화(encryption) 기술이나 스크램블링(scrambling) 기술이 널리 이용되고 있으나, 멀티미디어 저작물의 특성상 저작물을 보거나 듣기 위해서는 저작물을 보호하고 있는 보호장치를 해제해야 하기 때문에 디지털 저작물을

재생하는 도중에는 이들 기술로 보호 받을 수가 없다. 따라서, D/A, A/D변환을 통해서 쉽게 원래의 저작물을 얻어 불법 복사 또는 배포가 가능하게 된다. 이런 상황이 발생했을 때, 사후적으로 저작권을 보호하고 저작물에 대한 권리를 주장하기 위해서는 디지털 저작물 속에 사용자들은 인지가 불가능한 워터마킹 정보를 삽입해야 한다[1].

사후에 저작물에 대한 권리를 주장하기 위해서는 은닉된 식별코드가 어떤 의미를 가져야 하고, 의미를 적절히 표현하기 위해서는 정보량이 증가할 필요가 있다. 단순히 워터마킹으로 보호되고 있다는 것만을 의미할 수 있는 1~2비트 정보로는 부족하다. 지금까지의 워터마킹 기술이 어떤 저작물이 워터마킹이 삽입되어 있는가, 아닌가 또는 복사가 가능한 저작물인가, 아니면 복사나 유통이 제한되어 있는 저작물인가 등 1비트 혹은 4비트 이내의 제한된 정보를 삽입하는 기술이 대부분이었다. 이는 이미지, 비디오, 오디오 등 거의 대부분의 기술들이 이러한 종류에 포함된다. 위 기술들은 매우 견고하여 저작물의 변화에 강인한 특성을 이용해서 하드웨어 장치로 제작해서 원천적으로 저작물의 불법 유통 및 사용을 억제하기 위한 기술들이다. 대표적인 기술이 SDMI(secure digital music initiative)에 참여하고 있는 기술들이다.

그러나, 워터마킹 기술의 궁극적인 목적은 어떤 이용자에 의해서 불법적으로 배포가 되었다면 어떤 이용자에 의해서 불법 배포가 이루어졌는지 식별하기 위한 것이 첫번째이다. 즉, 저작물 속에 배포된 사용자 ID가 삽입 가능해야 하며, 온라인으로 서비스를 할 경우에는 어떤 저작물이 사용자들에게 판매가 되었으며, 그 저작물의 저작권자가 누구인지를 식별하기 위한 저작물 식별 코드가 삽입되어야 한다. 식별 코드는 최소한 64비트 이상으로 구성되기 때문에 근래에 온라인 멀티미디어 서비스를 위한 전체 조건으로 워터마크의 견고성, 비가시성 또는 비가청성, 그리고 삽입 가능한

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

정보량이 우선시 되고 있다.

본 논문에서는 웨이브렛과 디지털 필터를 이용한 워터마킹 삽입 및 추출 알고리즘을 이용해서 30초 이내의 오디오 저작물에 64비트 정보를 삽입하고, 추출할 수 있는 기술을 제안하고자 한다.

2. 기술동향 및 접근 방법

오디오 워터마킹 기술은 대부분 불완전한 인간의 청각 시스템 (HAS: Human Auditory System)을 이용한다. 우리가 매일 접하고 있는 MP3, WMA, AAC 등과 같은 압축 파일 형태도 인간의 청각 시스템을 이용한 기술들이다.

오디오 워터마킹 기술의 초창기 연구는 스프레드 스펙트럼 기법(Spread Spectrum Technique)을 이용한 워터마킹 기술이 널리 활용되었는데[2-4], 이 기술은 비동기화 공격 (desynchronization attack)에 치명적으로 약한 약점을 갖고 있는 것으로 밝혀져, 비동기화 공격에 대비한 워터마크의 설계와 동기화를 실시간으로 재조정할 수 있는 기술들이 제안되기도 하였다[5-6]. 그러나, 이러한 노력에도 불구하고 비동기화 공격에 취약한 성질은 실시간 워터마킹 추출 시스템을 구현하거나 많은 정보량을 삽입하는 가장 큰 걸림돌이 되고 있다. 예를 들어 플레이 도중에 일정 부분을 건너뛰거나 빨리 진행시키면 일부 데이터를 손실하게 되고, 워터마킹 정보를 정상적으로 추출할 수 없게 된다.

이러한 단점을 극복하기 위해서 디지털 통신 시스템에서 이용되고 있는 기술들이 도입되었는데, 대표적인 방법이 디지털 필터를 이용한 방법[7-8]으로, [7]에서는 all pass filter를 이용하여 위상 변조를 이용한 워터마킹 방법을 제시하였고, [8]에서는 이미지 워터마킹 알고리즘으로 임의의 두 지점에 위상변화를 가하여 그 위상 변화를 찾아서 워터마킹을 식별하는 방법을 제안하였다. 이외에도 Auto-correlation 을 이용한 방법과 잔향 또는 반향을 이용한 기술들이 제안되기도 하였다.

본 논문에서 제안한 방법은 디지털 필터를 이용하여 time scaling, pitch shift와 같은 비동기화 공격뿐만 아니라 손실 압축, 샘플링 주파수 변환과 같은 공격에 강한 특성을 갖는 워터마킹 기술이다. 먼저 우리는 디지털 필터에 의한 오디오 음원 변화를 최소화할 수 있도록 HAS의 임계 대역(critical band)을 고려하여 필터를 설계하는 방법을 제안한다. 워터마크를 삽입하기 전과 삽입한 후에 음질 변화를 느끼지 못하도록 하기 위해서는 사람의 청각 시스템의 신호 해상도를 고려한 디지털 필터를 설계해야 한다. 우리가 본 연구에서 이용한 디지털 필터는 대역 저지 필터로 특정 대역의 신호 성분을 제거 또는 왜곡하는 방법으로 대역 저지폭은 인간의 분해능을 넘지 않도록 설계하였다.

또한, 삽입 과정에서 필터링을 하는 대상 주파수 영역 이외의 저주파/고주파 대역에 미치는 역효과를 최소화하기 위하여 대상 주파수 영역만을 웨이브렛으로 분해하여 필터링하도록 설계하였으며, 필터링 과정에서는 위상변화가 발생하지 않도록 zero phase filtering을 이용하였다.

워터마크 정보 검출 과정에서는 윈도우 고속 푸리에 변환을 이용한 스펙트럼 분석과 정보 복원 절차에

대해서 설명을 한다.

마지막으로, 기술의 성능 평가 결과를 통해서 본 논문에서 제안된 기술의 유용성을 보여준다.

3. 정보 은닉을 위한 워터마킹 기술

3.1 워터마킹 기술의 조건

본 논문에서 제안하는 기술은 다음과 같은 성능을 만족하도록 설계하였다.

첫째, 비가청성. 오디오에 다량의 정보를 삽입함으로써 오디오의 음질 저하가 소비자에게 인식되어서는 안된다. 이는 기존 오디오 워터마킹 기술에서도 추구하는 것과 동일하다.

둘째, 강인성. 온라인으로 서비스되는 대부분의 오디오 포맷은 MP3, AAC, WMA, VQF, AC3 등 거의 모두가 손실 압축으로 압축된 포맷이기 때문에 다른 공격보다도 특히 손실 압축에 강한 특성을 가져야만 한다.

셋째, 신뢰성. 이 항목은 매우 중요하다. 추출된 정보는 어떤 소비자의 ID이거나 또는 저작물의 고유 ID이기 때문에 검출된 정보는 반드시 정확해야만 한다 즉, 검출된 정보가 삽입된 정보와 다른 정보를 검출하게 되면 안된다. 검출된 정보는 불법 사용자의 적발 또는 저작권자에게 비용을 지불하기 위한 정보로 활용되기 때문에 부정확한 정보의 검출이 발생해서는 안되며 반드시 정확하게 정보를 추출한 경우에만 결과를 출력해야 한다.

넷째, 대량 삽입 가능. 사용자 ID 또는 저작물의 식별 코드를 삽입하기 위해서는 최소한 64비트 정보가 필요하며, 부가적으로 오류 정정을 위한 정보와 추출 정보의 정확성을 인증하기 위한 CRC 정보 등을 포함하면 80비트 이상의 정보를 30초 이내에 삽입할 수 있는 기술이어야 한다.

3.2 제안된 기술

본 논문은 디지털 필터링을 이용해서 원본 오디오와 모든 면에서 동일하지만, 원본 오디오는 갖지 않는 스펙트럼 특성으로 워터마킹 신호를 만들기 위한 기술이다. 이를 위해서 음질의 변화는 최소화하면서 워터마킹 검출장치를 통해서 식별할 수 있는 성질의 필터를 설계한 다음에 각각의 필터 조합에 따라 <표 1>에 제시한 것과 같은 정보를 부여하였다. 디지털 필터는 저역통과필터, 고역통과필터, 대역통과필터, 대역저지필터 등으로 설계 특성에 따라 구분되는데, 본 논문에서 제안하는 기술에 가장 적합한 특징을 갖는 대역저지필터를 이용하여 구현하였다.

워터마킹 알고리즘의 삽입 과정을 간략히 설명하면, 만약 임의의 필터 H로 오디오 신호 S를 필터링하여 필터링된 오디오 신호 \tilde{S} 를 만들었다면, 이 필터링된 오디오 신호 \tilde{S} 의 스펙트럼은 원본 오디오 신호 S 구별이 가능한 특징을 갖게 된다. 즉, 어떤 필터를 통과한 신호인지 워터마킹 검출장치에서 식별할 수 있으며, <표 1>에서 정보를 찾아낼 수 있게 된다.

표 1. 필터 조합에 따른 삽입 정보

Em. Info	No. of filter	Em. Info	No. of filter
0000	NONE	1000	2,6
0001	1,3	1001	2,7
0010	1,4	1010	3,5
0011	1,5	1011	3,6
0100	1,6	1100	3,7
0101	1,7	1101	4,6
0110	2,4	1110	4,7
0111	2,5	1111	5,7

본 논문에서 이용되는 디지털 필터는 사전에 정보량과 저작물의 특성에 따라 디지털 필터를 설계하여 이용한다. <표 1>과 같은 정보를 표현하는 데에는 7개의 디지털 필터를 설계해야 한다. 디지털 필터는 특정 대역의 주파수 성분을 제거 또는 완화시키는 기능을 한다. <표 2>는 본 논문에서 적용한 IIR 디지털 필터 설계 요소(factor)들로, 4bit 입력정보를 삽입하기 위한 예이다. 여기서, 필터의 차단 대역의 중심 주파수는 4.4kHz, 4.8kHz, 5.2kHz, 5.6kHz, 6.0kHz, 6.4kHz, 6.8kHz이고, 차단대역(stopband)의 진폭 dB는 주변에 비해서 30dB 감소한다. 우리는 Elliptic 모형을 이용하여 Matlab Signal toolbox를 이용해서 필터를 설계하였다.

표 2. 디지털 필터

No.	통과대역 (Passband)	차단대역 (Stopband)	통과대역 변화 (dB)	차단대역 완화율 (dB)
1	4.25- 4.55	4.35-4.45	0.05	30
2	4.65- 4.95	4.75-4.85	0.05	30
3	5.05-5.35	5.15-5.25	0.05	30
4	5.45-5.75	5.55-5.65	0.05	30
5	5.85-6.15	5.95-6.05	0.05	30
6	6.25-6.55	6.35-6.45	0.05	30
7	6.65-6.95	6.75-6.85	0.05	30

대역 저지 필터의 대상 주파수 영역을 4~7kHz 대역으로 선택한 것은 음질과 강인성을 동시에 고려하여 선택한 것이다. 4kHz안쪽 주파수 대역은 오디오 신호중에서 가장 많은 정보를 표현하고 있기 때문에 사람의 귀에 매우 민감할 뿐만 아니라 미세한 변화도 음질에 많은 영향을 준다. 물론, 4kHz대역 안쪽을 이용하게 되면 강인성은 증가될 수 있으나, 압축(compression), 주파수 변환(re-sampling)을 고려하여 최소 64kbps, 16kHz 정도의 음질을 보존하는 오디오 신호에서는 검출이 가능한 범위를 필터링 대상으로 설정하였다.

$$y_k = \sum_{m=0}^M a_m x_{k-m} + \sum_{n=1}^N b_n y_{k-n} \quad \text{식(1)}$$

위 (1)식으로부터 우리는 원하는 주파수 성분만을 추출하여 사용하게 된다. 필터의 주파수 특성을 알아보기 위해서는 우선 (1)식의 양변을 z 변환을 하고 $z = e^{j\omega T}$ 을 대입하여 진폭특성과 위상특성을 구하여 필터의 주파수 특성을 분석하게 된다[9].

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{m=0}^M a_m z^{-m}}{1 - \sum_{n=1}^N b_n z^{-n}} \quad \text{식 (2)}$$

$$|H(e^{j\omega T})| = \sqrt{\{H_R(e^{j\omega T})\}^2 + \{H_I(e^{j\omega T})\}^2}$$

디지털 필터를 이용하여 오디오 신호에 정보를 삽입하는 데에 모든 디지털 필터가 가능한 것은 아니다. 워터마킹 기술에서 이용되는 디지털 필터에 의해서 원 오디오와 차별화된 특징을 갖는 디지털 필터를 이용해야 하며, 필터링에 의한 음질 차이가 사람의 청각 시스템에 의해서 인지되어서는 안된다. 따라서, 본 논문에서도 가급적 오디오 신호에 미치는 영향을 최소화하고, 필터에 의한 음질 변화를 거의 느끼지 못할 수 있는 필터로 대역 저지 필터를 이용하였다.

대역 저지 필터로 오디오 신호를 필터링 하게 되면, 특정 대역 주파수 성분이 제거되는데 이때 대역폭이 사람의 귀로 구분할 수 있는 임계 대역(critical band)보다 훨씬 작으면 필터링된 신호와 원본 오디오와 음질의 차이를 구별할 수 없게 된다. <표 2>는 주파수에 따른 임계 대역으로, 1kHz대역에서는 임계 대역은 190Hz로 폭이 좁으나 4kHz대역에서는 900Hz로 폭이 점점 더 넓어 지고 있는 것을 알 수 있다. 즉, 임계대역은 저주파에서 고주파로 갈수록 더 넓고 그 만큼 오디오 신호의 주파수 성분을 인간의 귀로 세밀하게 분석해낼 수 없음을 의미한다[10].

표 3. Zwicker 가 분석한 임계 대역[10]

Bark	f	Δf	Bark	f	Δf
0	0	100	12	1720	280
1	100	100	13	2000	320
2	200	100	14	2320	380
3	300	100	15	2700	450
4	400	110	16	3150	550
5	510	120	17	3700	700
6	630	140	18	4400	900
7	770	150	19	5300	1100
8	920	160	20	6400	1300
9	1080	190	21	7700	1800
10	1270	210	22	9500	2500
11	1480	240	23	12000	3500

가. 삽입 정보의 구성 (payload)

디지털 오디오에 사용자 ID 또는 저작물 식별 코드를 삽입하는 과정에서 부가적으로 필요한 정보가 많다. 아래 <그림 1>은 실제 삽입되는 정보의 구조로 Sync_Info, Sync_Test, CMI_Info(contents ID 또는 User ID),

CCI_Info, 그리고 CRC info 등이 포함되어 있다.

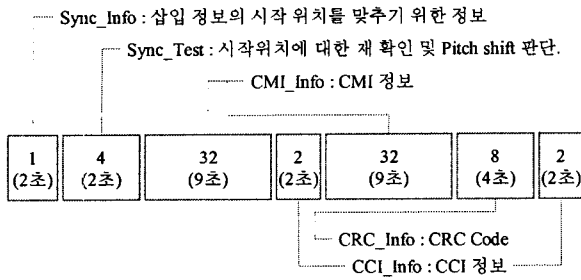


그림 1. Payload 구조

Sync_Info는 A/D, D/A 변환시에 발생하는 시간 지연에 따른 검출위치가 틀러지는 것을 막기 위해서 삽입되는 정보로 실제 정보를 표현하는 필터와 차별화된 필터를 이용해서 고유한 성질을 갖는 1bit 정보를 삽입한다. Sync_Test는 앞서 검출된 Sync_Info가 정확한지 아닌지를 판단하는 과정으로 '1001'의 값을 추출해서 판단한다. 이후에는 CMI_Info, CCI_Info와 같은 우리가 삽입하고자 하는 실제 정보를 삽입한다. 또한, CRC 정보는 추출된 정보가 오류가 전혀 없는 완벽한 정보인지 아닌지를 판단해서 정확한 정보인 경우에는 출력을 하고, 그렇지 않은 경우에는 오점지임을 알려주는 역할을 한다. CRC정보는 본 기술이 현장에 적용될 경우에 시스템의 신뢰성을 향상시키기 위해서 반드시 필요한 정보이다.

나. 정보 삽입 절차

오디오 신호에 정보를 삽입하기 위한 모듈은 <그림 2>와 같다. 여기서, Cover Signal은 정보를 삽입하지 않은 원본 오디오 신호를, Stego Signal은 Cover Signal에 삽입 정보가 포함된 신호를 의미한다.

워터마킹을 삽입하기 위한 과정을 살펴 보면, 첫 번째 단계로 입력된 Cover Signal을 웨이블릿 변환을 통해서 필터링할 대상 대역 주파수와 필터링되지 않는 대역 주파수로 분해한다. 필터링 단계에서 cover signal을 웨이블릿으로 분해하여 대상 주파수만을 필터링하는 것은 디지털 필터링 과정에서 주변 주파수 성분에 불필요한 변화를 주지 않기 위해서이다. 이상적인 디지털 필터를 설계하고 구현하는 것은 사실상 불가능하므로 정교한 필터를 설계하더라도, 필터링 과정에서 주변 신호의 주파수에 영향을 미치는 것을 막을 수는 없다. 특히, 고주파 성분은 상대적으로 적은 에너지량을 갖고 있는데, 이 부분에 발생하는 변화는 고주파를 왜곡시켜 오디오 전문가들이 느끼는 음색의 변화를 발생시킬 가능성이 높다. 이와 같은 왜곡 현상을 사전에 막기 위해서 시간-주파수 영역에서 신호의 분해능이 뛰어난 웨이블릿[11]을 이용한다.

두 번째 단계로 대상 주파수 신호를 미리 설계된 필터를 이용해서 필터링한다. 이때 여기서 설계된 디지털 필터가 IIR(infinite impulse response) 필터이기 때문에 비선형 위상 변화가 발생하고, 비선형 위상 변화는 음질 변화를 야기시킬 우려가 있다. 또, 이런 필터링 과정을 반복하게 되면 위상의 불규칙한 변화는 왜곡된

음을 만들기도 한다. 이런 위상 변화에 대한 문제를 해결하기 위하여 우리는 필터링을 거치는 동안 위상 변화가 발생하지 않도록 Zero-phase filtering[9,12]을 이용하였다. 이는 Fredrik[12]이 IIR 필터링을 전방향, 후방향으로 실시함으로써 위상이 변하지 않는 방식과 필터링의 초기값을 계산하는 방법을 이용하였다.

마지막으로, 웨이블릿으로 분해된 각각의 신호들과 필터링된 신호를 다시 합성하여 원래의 오디오 신호로 변환하면 정보가 삽입된 Stego signal이 만들어진다.

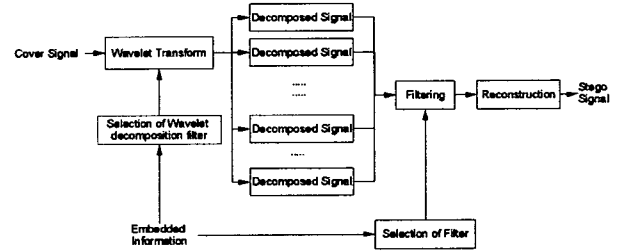


그림 2. 정보 삽입 구조도

위 <그림 2>는 워터마킹을 삽입하기 위한 각 모듈의 진행 과정을 설명한 것이다. 그리고, 아래 <그림 3>은 실제 구현을 위한 절차도이다.

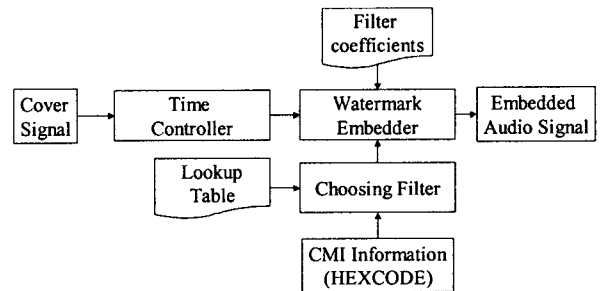


그림 3. 워터마킹 삽입 시스템

다. 정보 추출 절차

워터마크의 추출과정은 오디오의 진폭 스펙트럼을 분석해서 찾아낼 수 있다. 정확한 추출을 위해서는 워터마킹 삽입장치에서 이용한 필터 테이블을 갖고 있어야 하며 일반적인 스펙트럼을 분석하는 방법으로 구현이 가능하다. 본 연구에서는 가장 일반적인 FFT를 이용한 방법을 설명한다.

삽입과정에서 설명하였듯이 필터링은 오디오의 특정 대역 주파수 성분을 제거 또는 왜곡시키기 때문에 주파수 성분을 계속해서 합산하게 되면, 제거된 신호가 상대적으로 낮은 에너지량을 보이기 때문에 점점 더 뚜렷하게 차이를 나타내게 된다.

<그림 4>는 추출 과정을 보여주고 있는데, 우선 정보가 삽입된 오디오 신호(Stego signal)의 샘플링 주파수, 채널 수 등을 고려하여 충분히 필터링된 주파수 신호가 주파수 변환에 의해서 보일 수 있도록 분석 길이를 설정해야 한다. 우리가 필터링할 때 대역 차단된 주파수 범위가 80Hz이므로, 44.1kHz인 경우에는 1024, 또는 2048 데이터 샘플 단위로 분석해야 만이 필터링 효

과를 뚜렷이 관찰할 수 있다. 또, 필터링 효과는 로그 스케일로 변환된 상태에서만 정확하게 관찰할 수 있기 때문에 오디오 신호에 윈도우를 곱하여 Gibbs현상을 제거하고 진폭 스펙트럼을 dB단위로 표현한다. 이 과정을 미리 설정한 반복횟수 N회 실시하여 합산된 신호로부터 우리는 워터마크가 삽입된 위치를 식별할 수 있으며, 식별된 위치로부터 어떤 필터에 의해서 생겨난 피크인지 알 수 있기 때문에 <표 1>과 같이 정의된 lookup table을 이용하여 정보를 복원할 수 있다.

<그림 4>에서 수정된 로그 스케일링 과정(modified log scaling)은 최대값이 제한된 로그 스케일링을 의미한다.

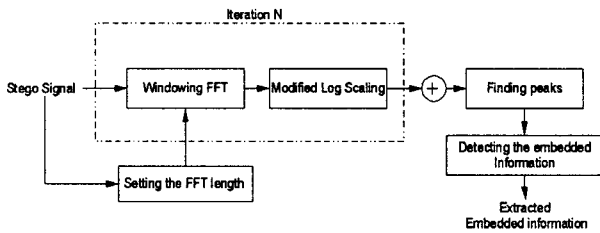


그림 4. 삽입 정보 추출 구조

<그림 5>는 워터마킹 추출을 위한 전체 시스템을 보여준다.

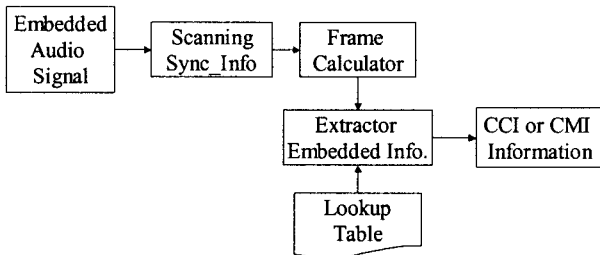


그림 5. 워터마킹 추출 시스템

4. 평가 결과

본 논문에서 제안한 기술의 성능을 평가하기 위하여 5곡의 음악에 워터마크를 삽입하고 추출하는 테스트를 실시하였다. 평가는 강인성 평가와 음질평가에 대해서 실시하였다.

4.1 강인성 평가 (Robustness Test)

아래 <표 7>은 각 음악에 30초 이내의 프레임에 반복적으로 64비트 정보를 삽입한 곡을 신호처리와 손실 압축으로 변환한 다음에 실험한 결과이다. 각 곡마다 11, 9, 6, 8, 8번의 정보가 삽입되었다. 실험결과에서 볼 수 있듯이, 대부분의 공격에서 성공적으로 64비트 정보를 추출하는 데에는 성공하였으나, 부분적으로 오류가 발생하게 됨을 알 수 있다. 본 연구에서는 CRC를 이용해서 검출된 정보에 오류가 발생하면 워터마킹이 검출되지 않았다는 메시지만 출력하도록 되어 있기 때문에 아래 <표 5>의 결과는 64bit 정보를 모두 정확하게 검출한 횟수에 대한 결과이다.

표 5. 신호처리 공격후 워터마크 검출 결과 (64bit, %)

Signal Processing	M1	M2	M3	M4	M5	Detection rates	
No Attack	10	9	6	8	7	95%	
Channel Conversion	9	9	6	8	6	90%	
Down Sampling	10	9	6	3	5	79%	
Pitch Shift	+4	8	6	6	7	74%	
	-4	10	9	5	5	86%	
Time Scaling	+10	10	9	6	8	90%	
	-10	10	9	6	8	95%	
Add Noise	-36dB	10	9	6	8	90%	
MPEG1 Layer3	128kbps	10	9	5	4	6	81%
	96kbps	10	9	5	2	4	71%
MPEG2 AAC	128kbps	10	9	5	4	6	81%
	96kbps	10	9	2	3	4	60%
WMA	128kbps	10	9	4	3	4	71%
ATRAC3	108kbps	10	9	5	5	5	81%

4.2 음질 평가 (Listening Test)

수치적인 방법으로 음질을 평가할 수 있는 방법 가운데서 가장 널리 이용되고 있는 계산법은 신호 대 잡음 비율 (SNR, Signal to Noise Ratio)을 구하는 방법이다. SNR이란, 신호는 단독으로 존재하지 않고 대개 잡음과 섞여 있는데, 그 비율을 나타내는 척도로서 SNR이 쓰인다. 신호전력을 Signal, 잡음전력을 Noise이라 할 때 수식 (3)과 같이 나타내며, 단위는 데시벨 (dB)이다.

$$SNR = 20 \left(\log_{10} \frac{Signal}{Noise} \right) \text{dB} \quad \text{식 (3)}$$

오디오의 경우에는 SNR에 많은 의미를 부여하지는 않으나, 일반적으로 60dB이상이 되면 거의 잡음을 감지할 수 없는 것으로 알려져 있다. 이와 유사하게 이미지의 경우에 SNR에 대한 이미지의 화질을 살펴 보면, 역시 60dB이상에서는 거의 어떤 노이즈도 발견할 수 없고, 최소한 40dB를 유지해야만 저작물의 가치를 유지할 수 있다(<표 6> 참조).

표 6. SNR에 따른 품질[13-14]

S/N ratio dB	S/N ratio:1	Audio quality
60 dB	1,000	Excellent, no noise apparent
50 dB	316	Good, a small amount of noise but audio quality good.
40dB	100	Reasonable, fine grain or snow in the audio, some fine detail lost.
30 dB	32	Poor audio with a great deal of noise.
20 dB	10	Unusable audio.

본 연구에서 테스트 대상으로 되는 5곡에 각각 정

보를 삽입하고 SNR를 측정한 결과 아래 <표 7>과 같은 수치를 얻었다. 이 값들을 보면, 80dB로 정보 삽입으로 인한 오디오 신호의 열화가 거의 없는 것으로 나타났다.

표 7. Signal to Noise Ratio (단위: dB)

	Pop-1	Pop-2	Classic-1	Classic-2	Peculiar
SNR	85.080	84.261	80.056	79.648	80.384

5. 결론 및 향후 연구 과제

본 논문에서는 오디오 신호에 저작물 식별 코드 또는 사용자 ID를 삽입 및 추출할 수 있는 방법을 소개하였다. 본 논문에서 제안한 이 기술은 매우 강인할 뿐만 아니라 음질 테스트 결과 원음과 정보가 삽입된 음을 거의 구별할 수 없는 것으로 분석되었다. 또한, 이 기술은 1~2비트에서부터 64비트 이상까지 많은 정보를 삽입할 수 있기 때문에 저작물의 고유 코드번호를 부여하여 삽입하여 추후에 어떤 음악이 누구에 의해서 사용되고 있는지, 현재 듣고 있는 음악의 고유 ID를 판별해 낼 수 있기 때문에 사후 저작물의 관리와 저작권자에게 정당한 이용료 지불 등에 활용될 수 있을 것으로 판단된다. 또한, 이 기술은 광고방송에 고유 식별 코드를 삽입하여 광고방송 모니터링 시스템이나 방송 동기화 장비에도 활용될 수 있을 것으로 판단된다.

참고문헌

[1] Katzenbeisser S., Petitcolas, F.A.P., Information Hiding Techniques for Steganography and Digital Watermarking, Artech House, Boston (2000).

[2] Bassia, P., Pitas, I., Robust audio watermarking in the time domain, In Proc. EUSIPCO 98, vol. 1. Rodos, Greece, IEEE (1998) 25-28.

[3] Cox. I. J., Kilian, J., Leighton, T., Shamoan, T., A secure robust watermark for multimedia, In Anderson, R. Information Hiding, Cambridge, UK. Lecture Notes in Computer Science, vol. 1174, Springer-Verlag, Berlin/Heidelberg (1996) 185-206.

[4] Swanson, M.D., Zhu, B., Tewfik, A.H., Boney, L., Robust audio watermarking using perceptual masking, Signal Processing 66, 337-355, 1998.

[5] Darko Kirovski and Henrique Malvar, Robust Covert Communication over a Public Audio Channel Using Spread Spectrum, Pre-proceedings of the 4th Information Hiding Workshop, Holiday Inn Select Pittsburgh University Center, 2001.

[6] Chung-Ping Wu, Po-Chyi Su, C. -C. Jay Kuo, Robust and Efficient Digital Audio Watermarking Using Audio Content Analysis, Security and Watermarking of Multimedia Contents II, 382-392, 2000. 12. 4.

[7] Tolga Ciloglu, S. Utku Karaaslan, An Improved All-Pass Watermarking Scheme for Speech and Audio,

Proceedings of the 2000 IEEE International Conference on Multimedia and Expo – vol. 2, 1017-1020.

[8] Faisal Alturki, Russell Mersereau, An Oblivious Robust Digital Watermark Technique for Still Images Using DCT Phase Modulation, Proceedings of the 2000 IEEE International Conference on Acoustics, Speech and Signal Processing - Volume 4, 1975-1978.

[9] Oppenheim, A.V., and R.W. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall, 1989, pp. 311-312.

[10] E. Zwicker, *Psychoacoustics*, Springer-Verlag, Berlin, 1990.

[11] Yang Yan, Zhang Dong, Enhancement of Angiograms via M-Band Wavelet Transform, Proceedings of SPIE Biomedical Photonics and Optoelectronic Imaging, 165-169, 2000.

[12] Fredrik Gustafsson, Determining the Initial States in Forward-Backward Filtering, Transactions on Signal Processing, V.44 N. 4, 1996. 4. 1.

[13] M. Adinolfi et al., Nucl. Instrum. Methods A 329 (1993) 117.

[14] The RD20 Collaboration, preprint CERN/DRDC 94-39 (Geneva,1994). Peter Weilhammer, private communication.