

# 청각시스템을 기반으로 한 새로운 오디오 워터마킹 시스템 설계

論 文

51D-7-5

## Design of a New Audio Watermarking System Based on Human Auditory System

申東煥\* · 申承源\*\* · 金鍾元\*\* · 崔鐘昱\*\* · 金德暎\*\*\* · 金成煥§  
(Dong-Hwan Shin · Seung-Won Shin · Jong-Weon Kim · Jong-Uk Choi  
Duck-Young Kim · Sung-Hwan Kim)

**Abstract** -In this paper, we propose a robust digital copyright-protection technique based on the concept of human auditory system. First, we propose a watermarking technique that accepts the various attacks such as, time scaling, pitch shift, add noise and a lot of lossy compression such as MP3, AAC, WMA. Second, we implement audio PD(portable device) for copyright protection using proposed method. The proposed watermarking technique is developed using digital filtering technique. Being designed according to critical band of HAS(human auditory system), the digital filters embed watermark without nearly affecting audio quality. Before processing of digital filtering, wavelet transform decomposes the input audio signal into several signals that are composed of specific frequencies. Then, we embed watermark in the decomposed signal (0kHz-11kHz) by designed band-stop digital filter. Watermarking detection algorithm is implemented on audio PD(portable device). Proposed watermarking technology embeds 2bits information per 15 seconds. If PD detects watermark '11', which means illegal song, PD displays "Illegal Song" message on LCD, skips the song and plays the next song. The implemented detection algorithm in PD requires 19 MHz computational power, 7.9kBytes ROM and 10kBytes RAM. The suggested technique satisfies SDMI(secure digital music initiative) requirements of platform3 based on ARM9E core.

**Key Words** : Audio Watermarking, PD(Potable Device), HAS(Human Auditory System).

### 1. 서 론

1990년대 이후 개인용 컴퓨터의 기능이 증대되고 네트워크의 전송속도가 향상되면서 시공간을 초월하여 다양한 디지털 데이터들을 주고받는 것이 매우 손쉽게 이루어지게 되었다. 이러한 환경은 기존의 아날로그 환경에서 불가능하다고 여겼던 다양한 디지털 콘텐츠(오디오, 이미지, 비디오, 문서 등)의 창작과 조작을 가능하게 만들었다. 이러한 네트워크를 통한 디지털 데이터의 교류 활성화는 긍정적인 측면과 함께 부정적인 측면을 유발하게 되었는데 불법적으로 복제된 콘텐츠(contents)가 인터넷을 통해서 매우 빠른 속도로 많은 사람들에게 퍼져나가면서 콘텐츠 제공자의 수익구조를 해치고 창작의욕을 떨어뜨리는 결과를 가져오게 되었다. 인터넷을 통한 다양한 콘텐츠의 유료서비스를 계획하거나 시행하기 위해서는 콘텐츠의 보안에 대한 요구가 증대되고 있으며, 몇 가지 방법을 이용하여 보안을 시도하고 있다.

대표적인 콘텐츠 저작권 보호기술로서 암호화 기술과 디지털

워터마킹(watermarking) 기술이 있다. 암호화 기술은 적법한 사용자에게만 암호화된 콘텐츠를 재생할 수 있도록 하는 기술로서 역시 복호화된 콘텐츠에 대해서는 보호할 방법이 없다. 워터마킹 기술은 1990년대 중반이후 각광을 받는 기술로서 콘텐츠 자체에 다양한 정보를 은닉하기 때문에 콘텐츠에 항상 따라다니는 정보라는 장점을 갖고 있다. 즉 워터마킹된 멀티미디어 콘텐츠를 암호화한 후 배포함으로써 복호화된 콘텐츠의 보호도 가능하게 되었다. 워터마킹 기술이 사용되는 응용분야는 다양하다. 삽입되는 워터마크의 용도에 따라서 방송되는 콘텐츠의 저작권료 지불을 위한 방송모니터링(broadcast monitoring) 시스템, 저작권 정보를 워터마크로써 삽입하여 추후 법정 분쟁에서 자신의 소유권을 주장할 수 있는 저작권인증(proof of ownership) 시스템, 콘텐츠의 불법 배포자를 찾아내기 위한 핑거프린팅(fingerprinting) 시스템, 콘텐츠의 위변조 식별을 위한 콘텐츠인증(content authentication) 시스템, 불법적인 복제를 원천적으로 막는 복제방지(copy protection) 시스템, 워터마크를 기기제어를 위해 사용하는 기기제어(device control) 시스템 등으로 구분된다 [1].

본 논문에서는 불법적으로 복제된 오디오 저작물을 휴대용 오디오 재생기에서 재생하지 못하도록 제어하는 복제방지 시스템을 제안하였다. 워터마크를 이용하여 복제방지를 하려는 노력은 세계적인 표준화 기구에서 주도하는 기술로서, 비디오 DVD의 경우 CPTWG(copy protection technical working group)에서 표준화를 진행하고 있고 오디오의 경우 SDMI

\* 正 會 員 : (주)마크텍 研究所

\*\* 正 會 員 : (주)마크에니 研究所

\*\*\* 正 會 員 : 富川大 電子科

§ 正 會 員 : 서울市立大學校 電子電氣컴퓨터工學部

接受日字 : 2002年 2月 26日

最終完了 : 2002年 4月 10日

(secure digital music initiative)에서 표준화 노력을 기울이고 있다[2]. 이러한 기구에서 워터마킹 기술을 이용하여 복제기와 재생기에 워터마크 검출장치를 내장하도록 하여 불법적으로 콘텐츠를 복사하는 것과 여기서 복사된 콘텐츠의 재생을 막고자 노력하고 있다.

오디오 콘텐츠에 사용되는 워터마킹 기술은 많은 연구자들에 의해서 개발되어져 왔다. 대표적인 방법으로 Cox등[4][12]에 의해 제안된 스프레드 스펙트럼 기법(spread spectrum technique)이 있다. 스프레드 스펙트럼 기법은 보내려는 정보를 넓은 주파수 대역으로 확산시키는 방식으로 수신기에서는 어느 한 주파수 대역에 간섭이 일어나도 신호를 수신할 수 있는 방법이다. 워터마킹에서는 삽입하려는 워터마크 정보를 광범위한 주파수 대역에 확산시킴으로써 사람이 인지하기 어렵게 만들어 정보를 은닉하고 있다. 이는 비동기화 공격(desynchronization attack)에 치명적인 약점을 갖고 있다. 최근에는 이와 같은 비동기화 공격에 대비하여 워터마크를 설계하고 삽입할 때 비동기화가 발생하더라도 상관도가 나타나도록 워터마크를 설계하거나[5], 워터마크를 삽입할 때 미리 설정된 간격으로 단순히 반복 삽입하지 않고 오디오 신호의 특성을 분석하여 오디오를 여러 개의 작은 블록으로 구분해 삽입하는 방법[6]을 시도하기도 하였다. 이 두가지 방법은 스프레드 스펙트럼이 갖고 있는 가장 큰 문제점을 해결하는 중요한 연구라 할 수 있다. 그러나 이러한 노력에도 불구하고 비동기화 공격에 취약한 특징은 실시간 워터마킹 추출 시스템을 구현하는데 큰 걸림돌로 작용한다. 예를 들어 재생중에 일정부분을 건너뛰거나 빨리 진행시키면 일부 정보를 오검지하게 된다.

디지털 필터를 이용한 방법으로는 Ciloglu등[7]이 all pass 필터를 이용하여 위상변조를 이용한 워터마킹 방법을 제시했으나, 워터마크를 검출하기 위해서는 동기를 맞추어야 하는 단점을 나타냈다.

본 논문에서 제안한 오디오 워터마킹 방법은 디지털 필터를 이용하여 비동기화 공격뿐만 아니라 시간 스케일링(time scaling), 피치이동(pitch shift), 손실압축과 같은 다양한 공격에 강인한(robust) 특성을 갖도록 하였다. 먼저 오디오 음원 변화를 사람이 거의 인지할 수 없도록 HAS(human auditory system)의 임계대역(critical band)을 고려하여 필터를 설계하는 방법을 제시하였다. 워터마크를 삽입하기 전과 삽입한 후에 음질변화를 느끼지 못하도록 하기 위해서 사람의 청각 시스템의 신호 해상도를 고려한 디지털 필터를 설계하였다. 본 연구에서 설계한 디지털 필터는 대역저지 필터(band stop filter)로 특정대역의 신호성분을 제거 또는 왜곡하는 방법으로 대역 저지폭은 100 Hz이하로 했다. 다음으로 필터링 과정에서 워터마크 삽입에 따른 음질저하를 최소화하기 위하여 대상 주파수 영역을 웨이블릿(wavelet)으로 분해하도록 했으며, 위상변화가 발생하지 않도록 하였다.

효율적으로 워터마크 정보를 검출하기 위하여 푸리에 변환(Fourier transform)으로 구한 진폭 스펙트럼을 로그 스케일(log scale)로 변환하여 워터마크를 추출했다. 제안한 방법의 효율성을 검증하기 위하여 워터마크 검출부분을 휴대용 오디오 재생기(portable device)에 구현하였다.

## 2. 장 새로운 오디오 워터마킹 알고리즘 제안

### 2.1 워터마킹 설계조건

본 논문에서 제안된 복제방지 시스템의 재생제어(playback control) 기능을 수행하기 위한 워터마킹 알고리즘은 아래와 같은 조건을 만족하도록 설계했다.

첫번째 조건은 비가청성(inaudibility)이다. 일반적인 오디오 워터마킹 알고리즘에서 요구되는 사항과 같다. 즉 오디오에 정보를 삽입함으로써 오디오의 음질저하가 소비자에게 인식되어서는 안된다. 이 경우 대부분 주관적인 판단으로 결정된다. 비교적 객관적인 판단 수단으로는 신호대 잡음비율(SNR, signal to noise ratio)을 이용한다.

두번째 조건은 강인성이다. 원본에 숨겨진 워터마크를 제거하기 위한 다양한 공격이 있을 수 있고 고의적이 아니더라도 음악을 듣기 위해선 보통 유손실 압축(MP3, AAC, WMA 등)을 행한다. 또한 원본음악을 재생하고 이를 마이크로폰이나 라인 입력을 이용한 복제가 행해질 수 있다. 즉 A/D, D/A와 시간스케일링과 피치이동과 같은 공격에도 워터마크가 살아 남아야 한다.

세번째로 요구되는 것은 알고리즘의 실시간 구현 가능성이다. 복제방지 기능을 휴대용 오디오 재생장치에 적용하기 위해서는 알고리즘이 간단해야하고 적은 연산량으로 워터마크를 검출 가능해야 한다. 또한 하드웨어 제약조건에 따라서 검출 알고리즘이 차지하는 프로그램 메모리 양과 데이터 메모리 양이 되도록 적어야 하고 정수연산으로도 워터마크 검출이 가능하도록 설계되어야 한다.

### 2.2 청각시스템의 주파수 특성

본 논문에서 설계된 워터마킹 알고리즘은 사람이 인지할 수 없도록 정보를 삽입하기 위해서 청각시스템의 개념을 사용하였다. 청각시스템은 크게 외이(外耳), 중이(中耳), 내이(內耳)로 구성된다. 귀에 입력된 음파는 청도(聽道)를 거쳐 고막을 진동시키고 이 진동이 중이를 거쳐 내이의 달팽이관 안에 있는 기저막(basilar membrane)을 진동시키게 된다. 기저막의 진동이 청신경을 자극하여 뇌에서 소리를 감지하게 된다.

기저막은 입력된 소리에 따라서 진동하고 주파수 분석기와 같이 음파신호를 주파수 성분으로 분해하는 역할을 한다. 기저막의 공진 주파수는 기저막의 두께에 따라 변화하며 그림 2.1 (a)와 같이 두꺼운 부분에서는 저주파 신호에 의해서 진동하고 끝으로 갈수록 두께가 얇아져서 고주파 신호에 의해서 진동하게 된다. 기저막이 진동하는 점과 공진주파수의 관계를 살펴보면 저주파 영역은 간격이 크고 고주파 영역은 간격이 작은 것을 알 수 있다. 이러한 특징에 의해서 그림 2.1 (b)와 같이 고주파에서 주파수 분해능이 저주파의 주파수 분해능보다 성능이 떨어지게 된다. 이와 같이 기저막 상의 위치와 반응하는 공진 주파수 관계에 대한 척도(scale)를 Bark 척도, 혹은 임계대역 척도라고 한다[8].

표 2.1은 Zwicker[9]가 분석한 주파수에 따른 임계대역으로 1kHz 대역에서는 임계대역이 190Hz로 폭이 좁으나 4kHz 대역에서는 900Hz로 폭이 점점 더 넓어지고 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.

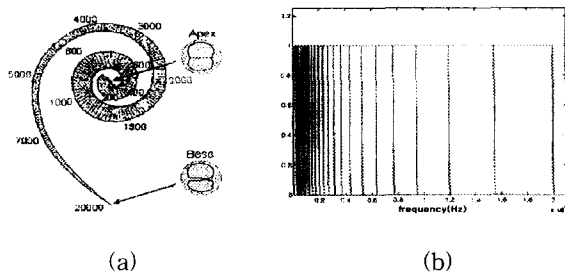


그림2.1 청각시스템의 임계대역 특성 : (a) 달팽이관 내의 기저막 위치에 따른 공진 주파수[Hz], (b) 임계대역

Fig. 2.1 Critical band characteristic of auditory system : (a) resonant frequencies[Hz] along the basilar membrane of cochlea, (b) critical band

표 2.1 Zwicker[9]가 분석한 임계대역

Table 2.1 Critical band by Zwicker[9]

Bark	f(Hz)	Δf(Hz)	Bark	f(Hz)	Δf(Hz)
0	0	100	12	1270	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

2.3 제안된 알고리즘의 구조

본 논문은 디지털 필터링을 이용해서 원본 오디오와 모든 면에서 동일하지만, 원본 오디오는 갖지 않는 스펙트럼 특성으로 워터마킹 신호를 만들기 위한 알고리즘을 제안한다. 이를 위해서 음질의 변화는 최소화하면서 워터마킹 검출장치를 통해서 식별할 수 있는 성질의 필터를 사람의 청각시스템을 기반으로 설계한 다음, 각각의 필터조합에 따라 표2.2에 제시한 것과 같은 정보를 부여했다. 본 논문에서는 대역저지 필터를 사용하였다. 대역저지 필터로 오디오 신호를 필터링하게 되면, 특정 대역 주파수 성분이 제거되는데 이때 대역폭이 사람의 귀로 구분할 수 없는 임계대역보다 훨씬 작으면 필터링된 신호와 원본 오디오 간의 음질 차이를 구별할 수 없게 된다.

표 2.2 필터 조합에 따른 워터마크 정보

Table 2.2 Watermark information according to filter combination

삽입정보	필터 조합	삽입정보	필터 조합
00	NONE	01	1,4
10	1,3	11	2,4

워터마크 삽입과정을 간략히 설명하면, 만약 임의의 필터 H(z)로 입력 오디오 신호 x를 필터링하여 출력신호 y를 생성

했다면 이 필터링된 오디오 신호 y의 스펙트럼은 x와 다른 특징이 있으며 이 특징을 표2.2와 비교하여 어떤 필터가 사용됐는지 판단하여 워터마크 정보를 추출할 수 있다.

그림2.2에 본 논문에서 사용한 필터링 방법에 대해 나타냈다. 본 논문에서는 필터응답의 신속성과 천이대역(transition band)을 좁게 하기 위해서 IIR(infinite impulse response)필터를 사용했다. 사용한 필터가 IIR 필터이기 때문에 위상특성이 선형이 아니다. 따라서 위상보상(phase compensation)을 하지 않고 그대로 사용하면 필터링 하지 않은 부분과 필터링 한 부분사이의 연결부위에서 잡음이 발생하게 되고 신호의 왜곡이 생기게 된다. 이러한 단점을 극복하기 위해서 zero-phase 필터링을 이용했다[10].

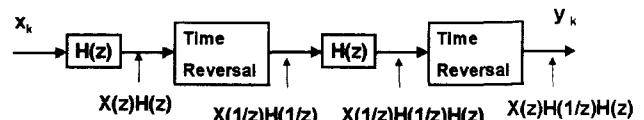


그림 2.2 워터마크 삽입을 위한 zero phase 필터링

Fig 2.2 Zero phase filtering for embedding watermark

본 논문에서 설계된 시스템은 2비트(bit)의 정보를 이용한 다. 표2.2에 표시된 것과 같이 2비트 정보를 삽입하기 위해서 4개의 필터를 구성했다. 각각의 2비트 정보를 표현하기 위해서 서로 다른 2개의 필터조합을 사용했다. 즉 오디오 신호에 2개의 필터정보를 각인시켜, 그 오디오 신호에 어떤 변형이 가해지더라도 삽입된 정보를 추출할 수 있게 된다.

본 논문에서 제안된 알고리즘은 사전에 디지털필터를 설계해 두어야 한다. 앞서 설명한 바와 같이 표2.2에서 4개의 필터가 필요하기 때문에 먼저 4개의 대역저지 필터를 설계했다. 설계된 IIR필터들의 차단대역(stop-band)의 중심 주파수가 5.35kHz, 6.15kHz, 6.95kHz, 7.65kHz인 필터이고 해당 차단대역의 감쇠율을 -30dB로 하고 통과대역리플(pass-band ripple)은 0.05dB로 했다. 타원필터(Elliptic filter)방법으로 Matlab Signal toolbox를 이용하여 설계했다. 필터차수는 모두 11차의 탭(식(1)의 M,N)을 갖는 필터구조를 갖고 있다. 표2.3은 본 논문에서 설계한 디지털필터 사양(specification)이다.

표 2.3 사용된 디지털 필터 사양

Table 2.3 Digital filter specification, used

No.	통과대역[kHz] (passband)	차단대역[kHz] (stopband)	통과대역 변화율(dB)	차단대역 감쇠율(dB)
1	5.25-5.45	5.30-5.40	0.05	30
2	6.05-6.25	6.10-6.20	0.05	30
3	6.85-7.05	6.90-7.00	0.05	30
4	7.55-7.75	7.60-7.70	0.05	30

대역저지 필터의 대상 주파수 영역을 5~8kHz로 선택한 것은 음질과 강인성을 동시에 고려하여 선택한 것이다. 5kHz 안쪽 주파수 대역은 오디오신호 중에서 가장 많은 정보를 표현하고 있기 때문에 사람의 귀에 매우 민감할 뿐만 아니라

미세한 변화도 음질에 영향을 준다. 물론 5kHz 대역 안쪽을 이용하게 되면 강인성은 증가될 수 있으나, 압축(compression), 샘플링 변화(variation of sampling rate)을 고려하여 최소 64 kbps 압축률과, 16 kHz 정도의 음질을 보존하는 오디오 신호에서는 검출이 가능한 범위를 필터링 대상으로 설정하였다.

디지털 필터링은 식(1)과 같이 구성되고 원하는 주파수만을 추출하여 사용하게 된다.

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

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 분석한다.

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

$$|H(e^{j\omega T})| = \sqrt{\{H_R(e^{j\omega T})\}^2 + \{H_I(e^{j\omega T})\}^2} \quad (3)$$

디지털필터를 이용하여 오디오 신호에 정보를 삽입하는 데에 모든 디지털 필터가 사용 가능한 것은 아니다. 워터마킹 기술에서 이용되는 디지털 필터에 의해서 원래의 오디오와 차별화된 특징을 갖는 디지털 필터를 이용해야 하며, 필터링에 의한 음질차이가 사람의 청각 시스템에 의해서 인지되어서는 안된다. 따라서, 본 논문에서도 가급적 오디오 신호에 미치는 영향을 최소화하고, 필터에 의한 음질 변화를 거의 느끼지 못하도록 하는 대역저지 필터를 사용했다.

2.3.1 워터마크의 삽입

오디오 신호에 워터마크를 삽입하기 위한 모듈은 그림 2.3와 같다. 여기서, cover signal은 정보를 삽입하지 않은 원본 오디오 신호를, stego signal은 cover signal에 삽입 정보가 포함된 워터마크된 신호를 의미한다.

그림 2.3에서 첫번째 단계로 워터마크 삽입 결정부가 있다. 이 부분에서 워터마크의 보안성을 높이기 위해서 cover signal의 워터마킹 삽입 유무를 결정한다.

두번째 단계로 입력된 cover signal을 웨이블릿 변환(wavelet transform)으로 필터링 대상 대역 주파수를 추출한다. cover signal을 웨이블릿으로 분해된 여러 신호중에 대상 주파수를 포함하는 신호만 필터링하는 것은 디지털 필터링에 의해서 주변 주파수에 미치는 side effect를 최소화하기 위해서이다. 이상적인 디지털 필터를 설계하는 것은 불가능하기 때문에 아무리 정교하게 필터가 설계되었어도 필터링에 의한 주변 신호와 위상 변화는 발생할 수밖에 없다. 특히, 고주파 성분은 상대적으로 적은 에너지량을 갖고 있는데, 이 부분에 발생하는 변화는 고주파를 왜곡시킬 가능성이 크다. 이런 변화를 사전에 막기 위한 작업으로, 시간-주파수 영역에서 신호의 분해능이 뛰어난 웨이블릿을 이용하였다.

세번째 단계로 대상 신호를 미리 설계된 필터를 이용해서 필터링한다. Fredrik[10]은 필터링을 전방향, 후방향으로 실시함으로써 위상이 변하지 않도록 하였으며 디지털 필터의 안

정적인 초기값을 결정하는 방법을 제안하였다. 이 방법은 위상변화를 제거할 뿐만 아니라, 필터링된 신호간에 자칫 발생할 수 있는 불연속점의 발생을 막는데 효과적이므로 본 연구에서는 이러한 방법을 사용하였다.

마지막으로, 웨이블릿으로 분해된 각 신호들과 필터링된 신호를 다시 합성하여 원래의 오디오 신호로 변환하면 워터마크가 삽입된 stego signal이 만들어진다.

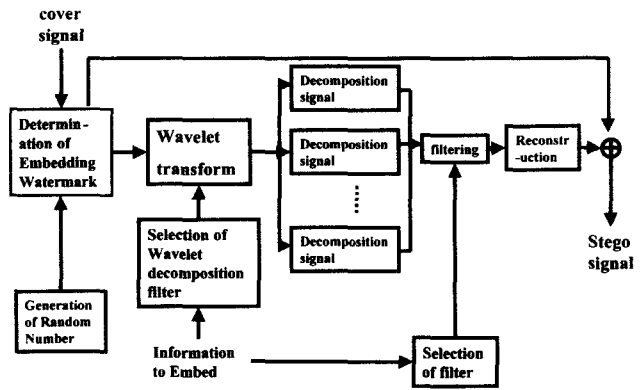


그림 2.3 워터마크 삽입 구조  
Fig. 2.3 Structure of embedding watermark

그림 2.4는 웨이블릿변환을 이용하여 워터마크를 삽입하는 예를 보여주고 있다. 여기서 사용된 웨이블릿은 haar 웨이블릿을 사용하였다. 여기서 WT는 웨이블릿 변환(웨이블릿 변환+역 웨이블릿 변환)을 이용하여 신호가 대역 분할된 것을 나타낸다. 웨이블릿 변환은 3 레벨까지 수행했다.

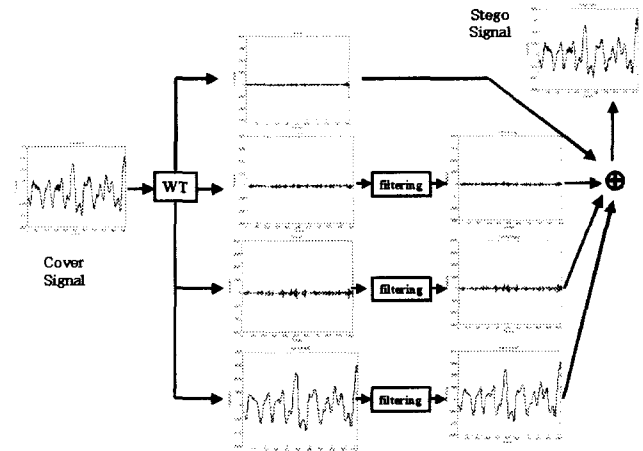


그림 2.4 웨이블릿을 이용한 워터마크 삽입 예  
Fig. 2.4 Example of embedding watermark using wavelets

위 과정을 일정 길이의 오디오 신호 단위로 반복적으로 진행하여 정보를 삽입한다. 만약, 64bit 정보를 삽입한다면, 위 과정을 32번 반복해야 모든 정보를 한번 삽입하게 된다. 많은 정보를 삽입할 때는 정보단위를 구분하기 위하여 추가 정보를 삽입한 다음에 반복적인 sync 비트 정보를 삽입하여 추출시에 전체 정보의 크기를 판단할 수도 있다.

또한, 적은 비트 정보를 삽입할 경우에는 오디오의 일부 신호에만 삽입해도 충분히 검출이 가능하므로 난수(random number) 발생기를 이용해서 난수값이 '0'일 경우에는 삽입하지 않고, '1'일 경우에는 삽입하는 과정을 반복하여 정보가 오디오 신호의 어느 부분에 삽입되어 있는지 감출 수 있다. 본 논문에서는 보안성을 위해서 오디오의 일부분에만 워터마크를 삽입하였다.

2.3.2 워터마크의 추출

워터마크의 추출과정은 오디오의 진폭 스펙트럼을 분석해서 찾아내도록 하였다. 삽입과정에서 설명하였듯이 필터링은 오디오의 특정 대역 주파수 성분을 제거 또는 왜곡시키는 효과를 주기 때문에 주파수 성분을 계속해서 합산하게 되면, 제거된 신호가 상대적으로 낮은 에너지량을 보이게 되고 뚜렷하게 차이를 나타내게 된다.

그림2.5은 추출 절차를 보여주고 있는데, 우선 워터마크가 삽입된 오디오 신호(stego signal)의 샘플링 주파수, 채널 수 등을 고려하여 필터링된 신호의 주파수 성분이 샘플링 변화에 의해서도 충분히 보일 수 있도록 분석 길이를 설정해야 한다. 대역 차단된 주파수 범위가 실질적으로 80Hz정도 넓으므로, 44.1kHz인 경우에는 1024, 또는 2048 데이터 샘플 단위로 분석해야 필터링 효과를 뚜렷이 관찰할 수 있다. 그러나 실제 하드웨어 구현시 실시간으로 처리하기 위해서 많은 샘플들을 블록(block) 프로세싱 할 수 없다. 이런 문제점 때문에 512개 샘플단위로 처리하였다. 샘플 수를 작게 하면 주파수 분해능(resolution)이 떨어져 정확한 피크 위치를 찾기 어려워진다. 필터링 효과는 로그 스케일로 변환된 상태에서만 정확하게 관찰할 수 있기 때문에 오디오 신호에 해밍 창함수(Hamming window function)를 곱하여 Gibbs현상을 제거하고 진폭 스펙트럼을 dB단위로 표현하였다. 이 과정을 미리 설정한 반복횟수 N회(워터마크 삽입비에 따라 달라질 수 있다) 실시하여 합산된 신호로부터 워터마크가 삽입된 위치를 식별할 수 있으며, 식별된 위치로부터 어떤 필터에 의해서 생겨난 피크(peak)인지 알 수 있기 때문에 표2.2과 같이 정의된 lookup table을 이용하여 정보를 복원할 수 있다. 본 논문의 휴대용 재생기 구현시에는 1204회 반복하여 신호를 합산하고 이 것으로부터 워터마크를 검출하였다.

그림2.5에서 로그 스케일링 과정은 일반적인 로그 변환에 문턱치(threshold value)를 적용하였다. 신호가 평균 음량 이상의 경우에는 문턱치로 변경하는 방식으로서 문턱치는 진폭 스펙트럼의 평균값을 이용하였다.

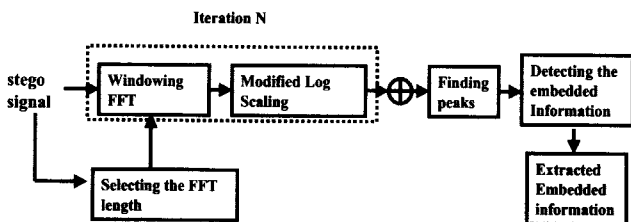


그림 2.5 삽입된 워터마크 추출 구조  
Fig. 2.5 Structure of extracting the embedded watermark

3. 제안된 알고리즘의 하드웨어 구현

복제방지 기능은 크게 2가지 기능으로 분류되며 하나는 복제장치에 워터마크 검출기를 내장하여 불법적인 멀티미디어 콘텐츠의 복제기능을 제어한다. 또 하나는 재생하는 단계에서 워터마크 검출기능을 내장하여 콘텐츠 안에 포함된 워터마크를 인식하고 불법 복제된 콘텐츠의 재생을 막는 기능이다. 그림3.1에 워터마크를 기반으로 한 복제제어 기능 흐름도를 나타냈다. 그림에서 복제방지를 위해 암호화를 이용했는데 인터넷이나 온라인 상에서 오디오 콘텐츠의 배포는 정당한 사용자만이 이용할 수 있게 보통 암호화 기법을 적용한다. 따라서 1차적으로 암호화에 의해 콘텐츠의 무단복제가 억제되고 정당한 사용자가 악의적으로 콘텐츠를 복제해 배포한 경우, 이를 막기 위해 워터마크를 이용한 복제방지 기능을 구현한다. 그림3.1에서 'X'표시는 사용이 억제되는 경우로 제어 수단(encryption, record control, playback control)을 함께 표시하고 있다. 본 논문에서는 워터마크를 기반으로 하는 재생제어(playback control)기능을 휴대용 재생기에 구현했다. 서론에서 밝힌 세계적인 표준화 기구인 CPTWG나 SDMI가 그림3.1의 방법으로 불법복제를 제어하고자 하였다. 그림3.1에 compliant이란 용어가 나오는데 이것은 표준화 기구의 표준을 따른다는 것이다. 즉, CPTWG나 SDMI가 인정하는 기술을 채택한 기기를 부를 때 사용한다. 본 논문에서 구현된 휴대용 재생기에는 복호화 알고리즘과 워터마크 검출 알고리즘이 구현되어 있다. 암호화 알고리즘은 미국 정부가 2001. 12월에 발표한 차세대 암호표준 알고리즘인 고등암호화표준(AES, advanced encryption standard)으로 채택된 린달(Rijndael)알고리즘을 사용하였다.

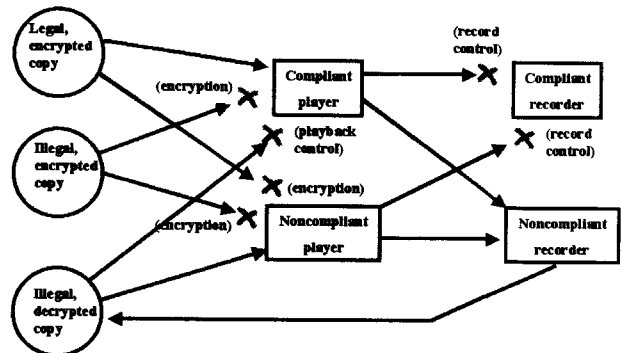


그림 3.1 암호화와 워터마크 기반의 복제제어 흐름도[1]  
Fig. 3.1 Flowchart of encryption and watermark-based copy control[1]

3.1 하드웨어 구성

본 논문에서 구현된 휴대용 재생기는 SDMI가 요구하는 사양[3]을 만족하도록 설계되어 있다. SDMI에서 요구하는 휴대용 재생기(PD:portable device)의 요구사항과 본 논문에서 구현한 사양을 표3.1에 나타냈다. SDMI의 요구사항 중 플랫폼3인 ARM-9E코어를 갖는 RISC칩을 사용한 휴대용 재생기의 요구사항을 나타냈다. 구현된 휴대용 재생기의 프로세서는 cirrus logic사의 ARM7을 코어로 하는 EP7209를 사용했

다. ARM7 코어가 ARM9E보다 성능이 떨어진다는 점을 감안하면 구현된 알고리즘의 성능은 SDMI의 요구사항을 충분히 만족함을 알 수 있다.

표 3.1 SDMI와 구현된 휴대용 재생기의 사양(ARM core)  
Table 3.1 Specification of SDMI and implemented PD(ARM core)

항 목	SDMI	implemented PD spec.
computational power	30 MHz	19 MHz
ROM	8 kB	7.9 kB
RAM	16kB	10 kB

그림3.2는 구현된 시스템의 구성도를 나타낸다. 4MByte의 프로그램 메모리와 16MByte 데이터 메모리로 구성된다. 현재 구현된 오디오 압축포맷은 MP3, AAC, WMA파일을 대상으로 구성되어 있다. 음원파일의 확장명을 이용해 해당하는 압축코덱을 실행하여 압축을 풀게된다. 오디오 음악 파일의 저장장치로는 32MByte MMC(multimedia card)를 사용했다. MMC에서 읽혀지는 오디오 신호는 마이크로 컨트롤러에서 오디오 압축코덱에 의해서 소프트웨어적으로 풀려진다. 풀려진 오디오 음악파일은 보통 44.1kHz 16bit 샘플링된 신호로 구성된다. 이 풀려진 오디오 신호는 DAC(digital analog converter)를 통해서 소비자가 스피커나 이어폰을 이용해서 들을 수 있는 신호가 된다. 또 ADC(analog digital converter)를 이용해서 외부 음악이나 음성을 ADPCM형태로 녹음할 수 있다. 마이크로 컨트롤러는 신호 송수신부에 의해서 RS232 시리얼 통신을 구현하고 이를 통하여 PC와 연결되어 컨트롤러의 운영 프로그램을 업데이트 하거나 프로그램 개발할 때 디버깅도 가능하게 한다. 사용자 인터페이스를 위해서 LCD와 여러 개의 키로 구성된 키패드를 마이크로 컨트롤러와 연결하여 사용할 수 있게 하였다.

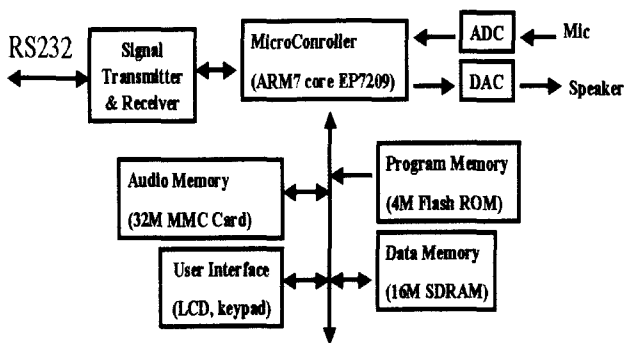


그림 3.2 구현된 휴대용 재생기 구조  
Fig. 3.2 The structure of implemented PD(portable device)

### 3.2 기능 구성

그림 3.3는 휴대용 재생기의 기능 블록선도이다. 워터마크 디코더에서 오디오에 삽입된 워터마크 정보를 추출하고 추출된 워터마크 정보를 바탕으로 재생기의 재생기능을 제어한다. 15초마다 한번씩 워터마크 2비트를 검출하도록 구현되었

으며 2번 이상 재생금지(구현된 재생기에서는 11 신호를 검지하면 재생금지로 판단한다) 신호가 검출되면 재생을 멈추고 LCD에 "Illegal Song"이라는 문구를 약 3초 정도 표시해준다. 3초 정도 표시한 후 현재의 오디오 파일은 건너뛰어 다음 오디오 음악을 재생하도록 하고 있다.

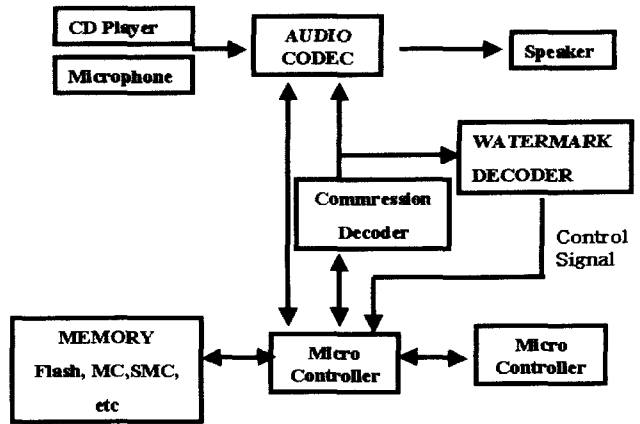


그림 3.3 휴대용 재생기 기능 블록선도  
Fig. 3.3 Functional block diagram of PD

본 논문에서 제안된 워터마크 추출 알고리즘의 하드웨어 구현은 빠른 연산시간을 위해서 실수형 계산을 정수형 계산으로 변환해서 사용하였다.

## 4. 실험결과 및 고찰

실험은 표준화 기구인 SDMI의 요구사항에 대해서 중점적으로 수행하였다. 실험은 대중가요 2곡(pop-1, pop-2), 클래식 1곡(classic), 음향효과음 1곡(sound effect)에 각각 워터마크 정보를 삽입하고, 15초 이내에 삽입된 워터마크를 반복적으로 추출하여 정확한 값을 찾은 경우의 수를 추출 횟수로 나누어서 검지율을 측정하였다.

### 4.1 음질평가

워터마크가 삽입된 오디오에 대한 음질평가는 주관적인 방법과 SNR을 이용한 객관적인 방법으로 수행되었다. 두가지 방법 중 주관적 음질평가 방법이 실제적으로 더 많이 사용되고 있는 방법이다.

#### 4.1.1 주관적 음질평가

오디오 음질의 평가는 ABX 음질 테스트 절차[11]에 따라 오디오 전문가(golden ear)들이 원음과 워터마크가 삽입된 음원을 얼마나 구별할 수 있는가를 확률적으로 평가하였다. 각 곡당 8번씩 원곡(A), 워터마크된 곡(B) 그리고 원곡과 워터마크된 곡중에서 임의의 곡(X)을 들려주고, 세번째 들려준 곡(X)이 앞서 들려준 A, B중에 어느 곡과 일치하는 지를 찾는 테스트이다.

표4.1은 3명의 전문가에 의해서 평가된 결과로 56% 원음을 정확히 찾아냈음을 알 수 있다. 그러나, 우연히 원음을 발견할 수 있는 확률이 50%라고 했을 때, 오디오 전문가들에

의해서 56% 찾아낼 수 있다는 것은 일반인들에 의해서는 거의 구별이 될 수 없다는 것을 의미한다.

표 4.1 워터마크 음질 테스트 결과 (ABX 시험)

Table 4.1 Result of watermark audibility test(ABX test)

전문가 \ 음악	1	2	3
pop-1	3/8	5/8	6/8
pop-2	5/8	5/8	5/8
classic	5/8	4/8	3/8
sound effect	4/8	5/8	5/8
검지율(%)	53%	59%	56%

4.1.2 객관적 음질평가

사람이 인식하는 특징과는 다소간의 차이는 있지만 정량적으로 표시할 수 있는 근사적인 방법이 신호대 잡음비(SNR, signal to noise ratio)로 측정하는 것이다.

$$SNR=20(\log_{10} \frac{Signal}{Noise}) \text{ dB} \quad (4)$$

오디오의 경우에는 SNR에 많은 의미를 부여하지 않으나, 일반적으로 60dB 이상이 되면 거의 잡음을 감지할 수 없는 것으로 알려져 있다. 표4.2에 SNR에 따른 오디오 성능을 나타냈다.

표 4.2 SNR에 따른 오디오 성능

Table 4.2 Audio quality to SNR

S/N ratio dB	S/N ratio : 1	audio quality
60dB	1,000	Excellent, no noise apparent
50dB	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
30dB	32	Poor audio with a great deal of noise
20dB	10	Unusable audio

본 연구에서 테스트 대상으로 되는 4곡에 대해서 각각 워터마크를 삽입하고 SNR를 측정된 결과 표4.3과 같은 수치를 얻을 수 있었다. 이 값들을 보면, 80dB로 정보삽입으로 인한 오디오 신호의 열화가 거의 없는 것으로 나타났다.

표 4.3 4곡에 대한 실험결과(SNR, dB)

Table 4.3 Experiment result for 4 songs(SNR, dB)

	pop-1	pop-2	classic	sound effect
SNR	85.080	84.261	80.056	80.384

4.2 강인성 평가(robustness test)

본 실험에서는 표4.4와 같이 신호처리 공격에 대한 강인성 테스트를 실시하였다. 이 실험은 총 4곡으로 15초마다 워터마크 정보를 검출하였고, 전체 검지횟수는 68번이다.

워터마크 검출을 어렵게 하는 공격에 대해서 간단히 설명하면, amplitude compression은 44.1kHz/16bit/2channel 음원을 하위 8bit를 잘라내고 상위 8bit를 이용하여 44.1kHz/8bit/2channel 음원으로 변환하는 공격이다. 워터마크를 하위 LSB(least significant bit)에 워터마크를 삽입한 경우 워터마크를 제거하기 위한 공격이다. channel conversion은 스테레오 신호를 더하여 모노로 만드는 공격방법이다. 스테레오 신호는 좌, 우 신호가 약간의 차가 있는데 이러한 성질을 이용하여 워터마크를 삽입한 경우 워터마크를 사라지게 하는 공격이다. down sampling은 오디오신호의 샘플링 주기를 변화시키는 공격으로 44.1kHz/16bit/2channel을 16kHz/16bit/2channel으로 변환했다. 워터마킹 알고리즘 중에 워터마크를 검출하기 위해서 동기를 맞춰서 계산해야 하는 경우가 많이 있다. 대표적인 것이 스프레드 스펙트럼 기술을 사용한 워터마크 알고리즘 방식이다. 이 방법으로 삽입된 워터마크를 상관도를 이용하여 검출하기 위해서는 동기를 맞추는 것이 필수적이다.

이러한 워터마킹 기술을 무력화시키기 위한 방법이 시간 스케일링(time scaling)과 피치이동(pitch shift)이다. 시간 스케일링은 오디오 데이터의 특정부분을 복사해서 늘리거나 제거하는 방법으로 오디오의 길이를 증감시키는 방법이다. 피치이동은 선형속도변화(linear speed change)라고도 하며 오디오의 길이를 선형적으로 늘이거나 줄여서 전체길이를 증감시키는 공격방법이다. 본 논문에서 제안된 방법의 장점은 이런 비동기화 공격(desynchronizing attack)에 강하다는 것이다.

표4.4에서 볼 수 있듯이, amplitude compression 공격과 WMA 압축률 64kbps에서 가장 검지율이 낮게 나타났으며, 그 이외의 공격에 대해서는 모두 90%이상의 높은 검지결과를 보였다. amplitude compression은 16비트 오디오 음원을 8비트 음원으로 변형하는 공격으로 오디오의 질적 저하로 인해서 일반 오디오에서는 거의 사용하지 않기 때문에 크게 문제되지 않을 것으로 판단된다.

본 논문에서 제안된 알고리즘의 우수성을 검증하기 위해서 스프레드 스펙트럼 기법에 의한 오디오 워터마킹 알고리즘 [12]과 비교했다. 스프레드 스펙트럼 기법을 워터마킹에 이용하는 방법은 여러 가지 방법이 있을 수 있으나 공통적인 중요한 특징은, 워터마크 추출단계에서 워터마크와 오디오 신호의 앙상블 평균(ensemble mean)과의 상관함수(correlation function)를 이용하기 때문에 추출 프레임간의 동기(synchronization)를 맞추어 검출해야 한다는 점이다. 따라서 표 4.5에 나타난 것과 같이 동기를 맞추지 못하게 하는 공격인 피치이동과 시간 스케일링에 매우 취약한 특징을 보인다. 본 논문에서 제안한 알고리즘이 이런 단점을 해결하고 있음을 알 수 있다.

**표 4.4** 다양한 공격후 워터마크 검출 정확도(회, %)  
**Table 4.4** Watermark detection ratio after various attacks (count, %)

attack item	pop-1	pop-2	classic	sound effect	detection rate		
no attack	21	18	11	16	66	97%	
amplitude compression	21	14	0	1	36	53%	
channel conversion	20	17	11	16	64	94%	
down sampling	21	18	12	16	67	99%	
pitch shift	+10%	20	16	11	14	61	90%
	-10%	21	17	11	15	64	94%
time scaling	+10%	21	17	12	15	65	96%
	-10%	21	17	12	15	65	96%
add noise	-36dB	21	17	11	15	64	94%
MPEG1 MP3	128kbps	21	17	10	16	64	94%
	96kbps	21	17	9	15	62	91%
MPEG2 AAC	128kbps	21	16	11	16	64	94%
	96kbps	21	16	9	16	62	91%
WMA	64kbps	21	16	10	12	59	87%

AAC: Advanced Audio Coding  
 WMA: Window Media Audio

**표 4.5** 스프레드 스펙트럼 방법과 제안된 방법의 성능비교  
**Table 4.5** Performance comparison with spread spectrum and proposed method

attack item	detection rate		
	spread spectrum method	proposed method	
amplitude compression	46%	53%	
channel conversion	91%	94%	
pitch shift	+10%	27%	90%
	-10%	13%	94%
time scaling	+10%	2%	96%
	-10%	2%	96%
add noise	-36dB	88%	94%
MPEG1 MP3	128kbps	90%	94%
	96kbps	84%	91%
MPEG2 AAC	128kbps	92%	94%
	96kbps	86%	91%

**5. 결 론**

본 논문에서는 오디오 콘텐츠의 불법복제 및 배포를 차단하기 위하여 새로운 오디오 워터마킹 알고리즘을 제안하고 휴대용 재생기를 설계하였다. 제안된 알고리즘은 미국음반협회의 주도로 이루어진 표준화기구 SDMI의 요구사항을 만족하는 방법으로서 특히 비동기화 공격에 강한 특성을 갖고 있다. 따라서 실시간으로 워터마크를 검출하여 하드웨어를 제

어하기에 매우 우수한 성능을 나타낸다. 워터마크 검출 알고리즘을 ARM7 core를 갖는 마이크로 컨트롤러에 구현했을 때 연산량이 19MHz, 사용된 저장공간은 ROM 7.9kByte, RAM 10kByte로 구현하였다. 이 정도의 리소스는 충분히 휴대용 재생기에서 수용할 수 있는 양이다. 본 논문에서는 MP3 플레이어에 적용했지만 PDA(personal digital assistants), mobile 폰 등과 같은 적은 용량의 리소스를 갖는 기기에 사용하기 적합한 알고리즘이다.

본 논문에서 제안된 워터마킹 알고리즘은 HAS(human auditory system)에 기본으로 한 임계대역을 기반으로 해서 설계된 대역저지필터를 이용하여 워터마크를 삽입하였다. 음질평가는 주관적인 평가로서 전문가에 의한 ABX 검사를 수행했으며, 결과는 56%정도 음원을 정확히 찾아냈기 때문에 우연히 찾을 수 있는 확률이 50%를 감안하면 일반인이 워터마킹된 음악을 구별하기가 쉽지 않으리라 판단된다. 객관적인 음질 평가로 SNR을 보면 80dB로써 오디오 신호의 질적 저하가 거의 없음을 나타내고 있다. 강인성 평가에서는 다양한 압축(MP3, AAC, WMA), 잡음첨가, 샘플링 변화 등의 공격에도 90%이상의 검출률로 양호한 결과를 보였으며 특히 본 논문의 알고리즘 특징으로 비동기화 공격(피치이동과 시간 스케일링)에 강인함을 보였다.

본 논문에서는 15초당 2bit 정보를 추출하도록 되어 있는데 약간의 알고리즘 변경에 의해서 15초 내에 64bit 이상까지 삽입이 가능하다. 이것은 디지털 콘텐츠의 저작물 식별코드로 사용함으로써 디지털 콘텐츠의 인터넷 방송에서 이용가능하고 또한 빠른 알고리즘 수행시간과 우수한 실시간 처리능력 때문에 방송광고 모니터링 시스템이나 방송동기화 시스템에 적용 될 수 있다고 사료된다.

**감사의 글**

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

**참 고 문 헌**

- [1] Ingemar J. Cox, Matthew L. Miller, Jeffrey A. Bloom, Digital Watermarking, Academic Press, pp.1-40, 2002.
- [2] J. A. Bloom, I. J. Cox, T. Kaller, et al, "Copy protection for DVD video," Proceedings of the IEEE, vol. 87, pp.1267-1276, July 1999.
- [3] [http://www.sdmi.org/port\\_device\\_spec\\_overview.htm](http://www.sdmi.org/port_device_spec_overview.htm)
- [4] I. Cox, J. Kilian, T. Leighton, and T. Shamon, "Secure Spread Spectrum Watermarking for Multimedia," IEEE Transactions on Image Processing, Vol. 6, pp. 1673-1687, Dec. 1997.
- [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, pp.363-378, 2001. 4.
- [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, pp. 382-392, 2000. 12.

[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, pp. 1017-1020. 2000.

[8] D. J. M. Robinson & M. J. Hawksford, "Time-Domain Auditory Model for the Assessment of High-Quality Coded Audio," The 107th Conference of the Audio Engineering Society, Sept. 1999.

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

[10] Fredrik Gustafsson, "Determining the Initial states in forward-backward filtering," IEEE Transactions on Signal Processing, vol.44 no.4, pp.988-992, 1996. 4.

[11] Burlington R., and May D. Jr., Handbook of Probability and Statistics with Tables, Second Edition, McGraw Hill NY, 1970.

[12] 박창복, 신승원, 김종원, 신동환, 최종욱, "랜덤신호를 이용한 오디오 워터마킹," 한국방송공학회 학술대회, pp. 37-41, 2000. 11.

저 자 소 개



**신 동 환 (申 東 煥)**  
 1995년 서울시립대학교 대학원 전자공학과 공학석사, 2000년 동대학원 박사과정 수료, 1992~1994년 LG전자 Video 사업부 연구원, 1996~2000 체육과학연구원 시스템개발팀 선임연구원. 2000~현재 (주)마크텍 부설연구소 선임연구원  
 Tel : 02-2267-6112  
 Fax : 02-2267-6113  
 E-mail : dhshin@marktek.co.kr



**신 승 원 (申 承 源)**  
 1996년 한양대학교 교통공학과(공학사), 1998년 한양대학교 대학원 교통공학과(공학석사), 1998~현재 (주)마크애니 부설연구소 선임연구원  
 Tel : 02-2262-5340  
 Fax : 02-2262-5333  
 E-mail : swshin@markany.com



**김 종 원 (金 鍾 元)**  
 1995년 서울시립대학교 대학원 전자공학과(공학박사), 1995~1996년 한국과학기술정보연구원 선임연구원, 1996~2000년 주성대학 정보통신학과/음향학과 조교수, 2000~현재 (주)마크애니 부설연구소 연구소장  
 Tel : 02-2262-5310  
 Fax : 02-2262-5333  
 E-mail : jwkim@markany.com



**최 종 욱 (崔 鍾 昱)**  
 1989년 University of South Carolina, Ph. D., 1989~1992년 KIST 인공지능연구실장, 1992~현재 상명대학교 정보통신학부 교수, 2000~현재 (주)마크애니 대표이사  
 Tel : 02-2262-5300  
 Fax : 02-2262-5333  
 E-mail : jwkim@markany.com



**김 덕 영 (金 德 暎)**  
 연세대학교 의용전자공학과(공학사), 서울시립대학교 전자공학과(공학석사), 서울시립대학교 전자공학과(공학박사), 현 부천대학 전자과 전임강사  
 E-mail : dykim@bc.ac.kr



**김 성 환 (金 成 煥)**  
 1952년 6월 12일 생. 1975년 2월 연세대 전기공학과 졸업. 1980년 2월 동 대학원 전기공학과(공박). 1977~1982년 대전 기계창. 1982년~현재 서울시립대 전자·전기공학부 교수.  
 Tel : 2210-2449, Fax : 2249-6802  
 E-mail : shkim1@uoscc.uos.ac.kr