

고음질 디지털 오디오 워터마킹을 위한 효율적인 PN 시퀀스 삽입 및 검출 방법

김 현 옥*, 오 현 오*, 김 연 정*, 윤 대 회*

An Efficient PN Sequence Embedding and Detection Method for High Quality Digital Audio Watermarking

Hyun-Wook Kim*, Hyen-O Oh*, Yeon-Jung Kim* and Dae-Hee Youn*

요 약

PN 시퀀스를 삽입하는 오디오 워터마킹은 들리지 않으면서도 강인한 워터마크를 만들기 위해 심리음향모델을 사용하여 PN 시퀀스를 변형시킨다. 하지만 워터마크를 삽입하는 모든 프레임에 대해 심리음향모델을 계산하기 위해서는 부호화 과정이 매우 복잡해지는 문제가 있다. 부호화기에서 심리음향모델의 역할을 대신 하도록 만든 고정필터로 PN 시퀀스를 변형하면 훨씬 간단하면서도 효율적인 워터마킹 시스템이 가능해진다. 본 논문에서는 고정된 지각필터를 도입한 효율적인 워터마킹 시스템을 제안한다. 심리음향모델을 대체하는 고정필터를 사용함으로써 PN 시퀀스를 들리지 않게 만들어주고 따라서 강인한 워터마크를 삽입할 수가 있다. 이와 함께 복호화기에서는 PN 시퀀스를 매칭시켜 상관도를 높여주는 보상필터 구조를 제안하여 원신호를 사용하지 않는 복호화기의 복호화 성능을 향상시켰다.

Abstract

In the PN-sequence based audio watermarking system, the PN sequence is shaped by a filter derived from the psychoacoustic model to increase robustness and inaudibility. The psychoacoustic model calculated in each audio segment, however, requires heavy computational loads. In this paper, we propose an efficient watermarking system adopting a fixed-shape perceptual filter that substitutes psychoacoustic model derived filter. The proposed filter can shape the PN-sequence to be inaudible and enable to embed the robust watermark in a simple manner. Moreover, we propose an architecture for the PN-sequence compensation filter in the watermark detector to increase correlation between the watermark and the PN-sequence. With the proposed architecture, the blind watermark detection performance has been enhanced.

I. 서 론

디지털 워터마킹이란 디지털 매체 내부에 사용자가 원하는 정보를 보이거나 들리지 않도록 숨기는 것을 말하며 이때 삽입되는 정보를 워터마크라고 한다. 즉, 디지털 워터마크를 이용하여 매체 내부에 저작권 등의 부가정보를 지각되지 않도록 삽입, 숨김으로써 매체의 유통을 제어하

고, 불법 복제를 막기 위한 법적 제재 근거를 만들 수가 있다. 저작권 보호를 목적으로 하는 디지털 워터마킹은 방송 모니터링, 소유자 확인, 소유권 증명, 인증, 유통 추적, 복제 제어, 접근 제어 등의 용도로 활용될 수 있다^{[1][2]}. 이러한 워터마킹은 활용 목적에 따라 요구되는 특성이 달라지지만 다음의 두 가지 기본적인 특성을 가지고 있어야 한다.

- (1) 보이거나 들리지 않는 비지각성(Imperceptibility).
- (2) 공격이나 신호처리 과정에서도 없어지지 않는 강인성(Robustness).

* 연세대학교 전기전자공학과 음향·영상·신호처리연구실
ASSP Lab., Dept. of Electronic and Electrical Eng., Yonsei Univ.

삽입되는 워터마크 신호의 에너지가 클수록 공격에 대해 강인하고 복호화 성능이 좋아지지만 어느 정도 이상의 에너지로 삽입되면 보이거나 들릴 수 있기 때문에 비저각성을 만족시키지 못한다. 이렇게 워터마크의 두 가지 기본 특성인 “비저각성”과 “강인성”은 서로 상충관계를 갖는다. 워터마킹 시스템은 크게 부호화기와 복호화기로 나누어진다.

부호화기에서는 삽입 알고리즘을 통해 워터마크 정보를 매체 내에 지각되지 않도록 삽입하고, 복호화기에서 검출 알고리즘으로 삽입했던 워터마크 정보를 추출해 낸다. 부호화기 및 복호화기의 구조는 시스템에서 사용하는 알고리즘에 의해 결정되는데 먼저 워터마크에 요구되는 강인성과 비저각성에 따라 삽입 알고리즘을 설계하게 되며, 삽입된 워터마크 신호를 효율적으로 추출할 수 있도록 검출 알고리즘을 설계하게 된다. 대표적인 워터마킹 알고리즘은 대역확산 방법, 위상변형 방법, 반향삽입 방법 등이 있다^{[1][3]}.

대역확산 방법은 워터마크 신호로써 PN 시퀀스를 삽입하고 그 상관특성을 이용하여 워터마크를 검출한다^[4]. 부호화기에서는 PN 시퀀스를 들리지 않도록 변형한 후 이를 오디오 신호에 더하는데 이때 주로 심리음향모델을 사용한다^[5]. 심리음향모델을 사용하는 부호화기는 그 구조가 복잡하지만 복호화기는 PN 시퀀스의 상관도를 이용하여 간단하게 구현될 수 있는 장점이 있다.

본 논문에서는 대역확산 방식의 워터마킹 시스템에서 새로운 PN 시퀀스 변형 삽입 방법과 효율적인 워터마크의 복호화 기법을 제안한다. 기존의 대역확산 방식 워터마킹 시스템에서 PN 시퀀스를 변형하기 위해 사용하는 심리음향모델은 많은 연산을 필요로 하기 때문에 부호화 과정이 상당히 복잡하고 따라서 워터마크를 삽입할 수 있는 휴대용 기기 등의 구현을 어렵게 하는 요인이 된다.

본 논문에서 제안한 PN 시퀀스 변형 방법을 사용하면 심리음향모델의 복잡한 연산과정 없이도 심리음향모델을 사용한 것과 유사한 워터마킹 성능을 갖도록 만들 수가 있다. 즉, 심리음향모델 대신 고정된 필터를 사용하여 삽입되는 워터마크를 들리지 않게 삽입하는 동시에 효율적인 워터마크 추출이 가능하게 된다.

먼저 2장에서는 기존의 심리음향모델을 적용한 PN 시퀀스 삽입 워터마킹 시스템에 대해 알아보고, 3장에서는 부호화 과정에서 심리음향모델의 PN 시퀀스 변형 효과를 분석하여 심리음향모델을 대신할 수 있는 새로운 고정필터를 제안한다. 4장에서 원신호가 필요 없는 Blind detection이 되도록 하기 위해 적용한 전처리 과정을 설명하며, 이때 복호화 성능을 향상시킬 수 있는 효율적인 상관도 검출

구조를 제안한다. 5장에서는 기존의 시스템과 제안된 시스템의 워터마킹 성능을 비교하기 위해 음질 및 복호화 성능을 측정하여 제안한 고정필터를 사용한 새로운 워터마킹 시스템이 비저각성을 만족하면서 효율적인 워터마크의 검출이 이루어짐을 확인하고, 마지막 6장에서 결론을 맺는다.

II. 대역확산 기반 오디오 워터마킹

대역확산(Spread-Spectrum) 기반 워터마킹은 PN 시퀀스(Pseudo-Noise Sequence)를 워터마크 신호로 삽입하고 그 상관특성을 이용하여 삽입된 워터마크 정보를 추출하는 방법이다^[5]. PN 시퀀스는 백색잡음과 유사한 통계적 특성을 갖도록 만들어진 주기신호로 우수한 자기상관특성을 가지며 간섭에 대한 영향이 적다^[7]. 또한 PN 시퀀스 자체만으로도 암호화가 이루어지기 때문에 워터마크로써 유리하게 사용될 수가 있다. 이러한 PN 시퀀스는 들리지 않을 정도의 작은 에너지로 오디오에 삽입되어도 자기상관특성에 의한 추출이 가능하지만 PN 시퀀스를 심리음향모델에 의해 변형하여 삽입하면 훨씬 큰 에너지를 가지면서도 들리지 않게 삽입할 수 있기 때문에 더욱 강인한 워터마킹이 될 수 있고 효율적인 검출이 이루어지게 된다. PN 시퀀스 변형에 사용하는 심리음향모델은 인간의 청각특성인 마스킹 효과를 모델링한 것으로 MPEG Audio와 같은 고음질 오디오 부호화 방식에서 제안하는 방법을 사용할 수 있다^{[8][9]}. 심리음향모델로 PN 시퀀스를 변형 삽입하는 대역확산 기반 워터마킹 시스템의 부호화기와 복호화기는 다음과 같다.

1. 부호화기

부호화기는 PN 시퀀스를 해당 오디오 신호에 더함으로써 워터마크 정보를 삽입한다. 이때 PN 시퀀스를 심리음향모델로 변형해주어 워터마크의 강인성을 높여주게 된다. 워터마크 신호를 오디오에 더해지는 부가잡음이라고 하면 식 (1)로 정의되는 SNR에 의해서 삽입되는 워터마크 신호의 에너지를 나타낼 수 있다.

$$SNR = 10 \log_{10} \left(\frac{\text{Audio Signal Energy}}{\text{Watermark Signal Energy}} \right) \text{ (dB)} \quad (1)$$

SNR이 크면 오디오 신호에 비해 워터마크 신호가 상대적으로 작게 삽입된 것이므로 비저각성 측면에서는 좋지만 강인성 및 복호화 성능은 떨어지게 된다. 따라서 들

리지 않는 한도 내에서 워터마크의 에너지가 최대가 되도록 삽입해 주어야 효율적인 복호화가 이루어질 수 있다. PN 시퀀스 삽입 워터마킹의 부호화 과정은 그림 1의 (a)와 같다. 워터마크를 삽입하려는 원신호를 $x(n)$ 이라고 하고 삽입되는 PN 시퀀스 워터마크 신호를 $w(n)$ 이라고 할 때 부호화기의 출력인 워터마크가 삽입된 오디오 신호 $y(n)$ 은 식 (2)와 같이 얻을 수 있다.

$$y(n) = x(n) + S(x(n), w(n)) \quad (2)$$

$$= x(n) + \tilde{w}(n)$$

이때 $\tilde{w}(n)$ 은 비지각성을 향상시키기 위해 변형된 PN 시퀀스이고 식 (3)과 같이 처리된다.

$$\tilde{w}(n) = S(x(n), w(n)) \quad (3)$$

$$= scf \times \text{IFFT} [M(i) \times W(i)]$$

식 (3)에서 $M(i)$ 는 원신호 $x(n)$ 에 대해 심리음향모델로 구한 마스킹 곡선이고 $W(i)$ 는 주파수 변환된 PN 시퀀스이며 scf 는 최종적으로 워터마크 신호의 에너지를 결정하기 위해 공급하는 상수이다. 심리음향모델에 의해 구해지는 마스킹 곡선은 스펙트럼 정보이므로 이를 주파수축 상에서 PN 시퀀스에 곱함으로써 정확한 스펙트럼의 변형이 가능하다. 하지만 이 경우 워터마크를 삽입하는 모든 프레임마다 주파수축에서 스펙트럼 변형된 PN 시퀀스를 시간축으로 역변환 해주어야 하기 때문에 추가적인 연산이 필요하게 된다.

2. 복호화기

삽입된 워터마크 정보는 PN 시퀀스의 상관특성을 이용하여 검출하게 된다. 즉, 워터마크된 오디오에서 추출해낸 워터마크 신호와 삽입시에 사용한 PN 시퀀스의 상관도를 구하여 워터마크 정보를 복호화한다^[5]. 이때 PN 시퀀스를 정확하게 알아야 상관도를 구할 수 있으므로 삽입시에 사용한 PN 시퀀스는 복호화를 위한 키 값이 된다. 복호화기에서 원신호를 사용할 수 있는 경우에는 단순히 워터마크가 삽입된 오디오 신호에서 원신호를 빼는 것으로 워터마크를 정확하게 추출할 수가 있고 가장 효율적인 복호화가 이루어질 수 있다. 하지만 원신호를 사용하지 않는 Blind detection 복호화기는 오디오 신호에 비해서 아주 작은 에너지로 삽입되어 있는 워터마크 신호를 별도로 추출해내는 과정이 필요하다.

$$y'(n) = (x(n) + \tilde{w}(n))' \quad (4)$$

$$= x'(n) + \tilde{w}'(n)$$

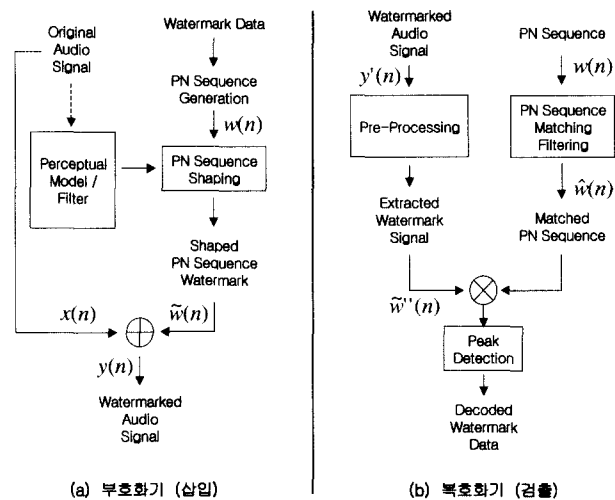


그림 1. PN 시퀀스 삽입 워터마킹 시스템
Fig. 1. PN-sequence embedding watermarking system (a) Encoder (b) Decoder

Blind detection 방식의 워터마크 복호화 과정은 그림 1의 (b)와 같다. 먼저 워터마크가 삽입된 오디오 신호 $y(n)$ 이 채널을 통과하면서 식 (4)와 같이 $y'(n)$ 으로 변형되어 복호화기의 입력으로 들어오게 된다. 복호화기의 입력 $y'(n)$ 이 전처리 과정을 통과하면서 삽입되었던 워터마크 신호 $\tilde{w}'(n)$ 이 $\tilde{w}''(n)$ 으로 추출된다. 마지막으로 이렇게 추출된 $\tilde{w}''(n)$ 과 삽입에 사용했던 PN 시퀀스 $w(n)$ 의 상관도를 구함으로써 워터마크 데이터를 검출하고 복호화한다.

III. PN 시퀀스 변형을 위한 새로운 고정필터

부호화기에서 PN 시퀀스를 들리지 않도록 변형하기 위해 사용하는 심리음향모델은 매우 복잡한 연산 과정이다. 또한 워터마크를 삽입하는 모든 프레임마다 심리음향모델을 적용하기 위해서는 부호화 과정에서 많은 연산이 요구되므로 구현이 어려워진다. 이러한 문제는 부호화기에서의 심리음향모델 작용을 대신할 수 있는 효율적인 필터를 사용함으로써 해결할 수 있다. 본 장에서는 그 한 가지 방법으로 적용할 수 있는 고정필터를 제안한다. 고정필터는 심리음향모델로 구해진마스

표 1. 실험에 사용한 여섯 가지 오디오 샘플
Table 1. Test audio samples for experiments

샘플	곡명	설명
CE1	Maiden Voyage - Leny Andrade	베이스, 심벌, 여성보컬 연주
CE2	Played Twice - The Fred Hersch	조용한 피아노 연주곡
CE3	I Love Paris - Johnny Frigo	경쾌한 바이올린 연주곡
CE4	Sweet Georgia Brown - Monty Alexander	대편성(Big-Band) 연주
CE5	Grandma's Hand - Livingston Taylor	남성 4명 아카펠라 곡
CE6	Flute Concerto in D - Vivaldi	관현악과 플룻 연주곡

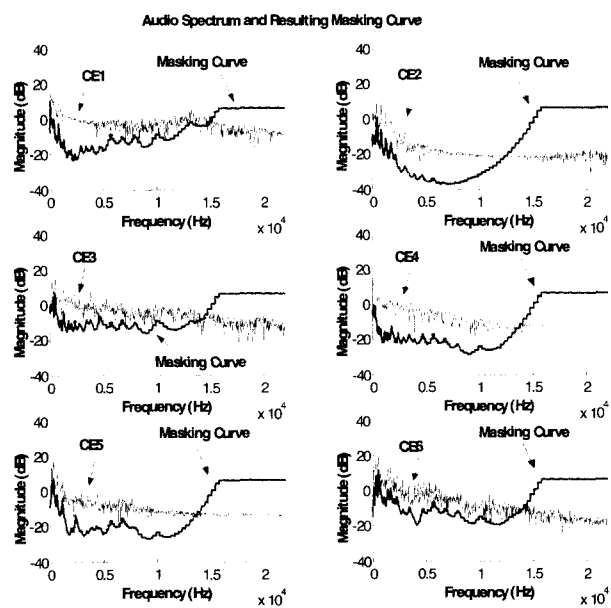


그림 2. 실험한 오디오 샘플로 구한 마스킹 곡선
Fig. 2. Masking curves for several test audio samples

킹곡선의 특성을 갖도록 만들어진다. 실제로 마스킹 곡선은 오디오 신호에 의해 형태가 결정되기 때문에 다양한 종류의 오디오 신호로부터 얻은 마스킹 곡선에서 공통적인 특성을 찾고 이를 바탕으로 고정필터를 만들게 된다. 이때 삽입되는 PN 시퀀스가 모든 오디오 신호에 의해 마스킹되어 잡음으로 들리지 않도록 하기 위해서는 특정 오디오 신호에 의한 마스킹 형태는 배제하여야 한다.

그림 2는 표 1의 여섯 가지 오디오 샘플에 대해 심리음향모델로 구한 마스킹 곡선의 예이다. 그림 2에서와 같이 마스킹 곡선이 중간 주파수 영역에서는 오디오 신호에 따라 달라지지만, 16kHz 이상의 주파수 영역에서는 모두가 일정한 형태를 가지며 이 영역에 대부분의 에너지를 가지고 있음을 볼 수 있다. PN 시퀀스는 이러한 특성의 마스킹 곡선에 의해 큰 에너지를 가지면서도 들리지 않게 변형되어 오디오 신호에 삽입될 수 있게 된다. 마찬가지로 심리음향모델을 대신하는 고정필터는 마스킹 곡선과 동일하게 PN 시퀀스를 큰 에너지를 가지면서도 들리지 않도록 변형해줄 수 있어야 한다.

그림 3(a)는 고정필터를 설계하기 위해 구한 마스킹 곡

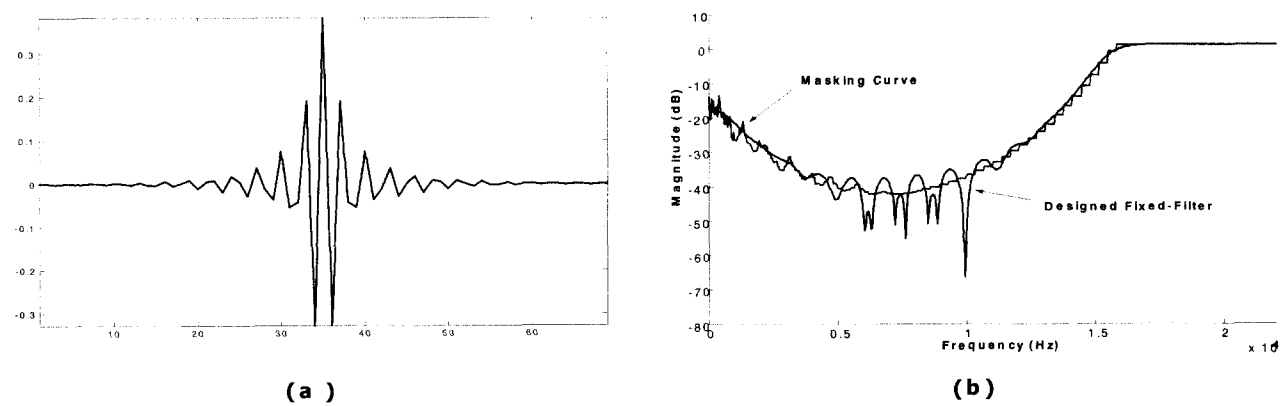


그림 3. 설계된 고정필터의 계수(a)와 주파수 응답 곡선(b)
Fig. 3. Fixed filter coefficients(a) and frequency response(b)

표 2. 900 프레임에 대한 부호화 과정 수행시간 비교
Table 2. Function processing time of encoding of 900 frames

기 존 방 법 (심리음향모델)	수행 횟수	심리음향모델			고정필터	
		900	900	900	N/A	N/A
	시간(msec)	1116.610	2558.633	1160.665	790.244	5626.125
	백분율 (%)	19.8	45.5	20.6	14.1	100
제안한 방법 (고정필터)	수행 횟수	0	0	0	N/A	N/A
	시간(msec)	0	0	0	631.814	631.814
	백분율	0	0	0	100	100

선에 대해서 주파수 샘플링 방법으로 설계한 68차 FIR 고정필터의 예이다. 그리고 그림 3(b)는 설계된 고정필터의 주파수 응답과 마스킹 곡선을 그린 것이다. 설계된 고정필터의 차수는 부호화기의 경우에는 미리 변형된 PN 시퀀스를 삽입하기 때문에 필터 차수가 부호화기의 구현에 문제가 되지 않지만, 다음 IV장에서 다루는 내용과 같이 복호화 성능 향상을 위해 고정필터를 복호화기에서 사용하는 경우에는 차수를 고려하여 설계해야 한다. PN 시퀀스는 이렇게 설계된 고정필터에 의해서 스펙트럼이 변형되고 최종적으로 삽입하는 과정에서 오디오 신호의 크기에 비례하여 에너지를 조절함으로써 오디오 신호에 들리지 않는 워터마크로 삽입된다.

고정필터를 사용한 부호화기는 심리음향모델을 사용하지 않고 미리 변형된 PN 시퀀스를 삽입할 수 있기 때문에 부호화 과정이 훨씬 간단해진다. 표 2는 부호화 과정의 실제 연산량 측정을 위해 기존의 심리음향모델을 사용한 부호화기와 본 논문에서 제안한 고정필터를 사용한 부호화기를 각각 C로 구현하여 부호화 과정 수행시간을 측정 한 것이다. 공정한 수행시간 비교를 위해 각 시스템에 대해 동일한 환경에서 900 프레임의 오디오에 워터마크를 삽입하도록 하여 과정별 수행시간을 측정하였다. 여기에서 프레임은 워터마크 정보 한 비트가 삽입되는 오디오 신호의 단위 블록이다.

측정 결과를 보면 기존 방식의 부호화기에서는 심리음향모델과 FFT, IFFT가 매 프레임마다 한 번씩 수행되는데, 이 과정이 부호화 과정의 대부분의 연산량을 차지하고 있다. 이때 FFT는 심리음향모델 과정 내부에서 수행되는 것이므로 심리음향모델 자체가 전체 부호화 과정의 65% 이상을 차지함으로 알 수 있다. 또한 기존의 부호화에서는 식 (3)에서와 같이 스펙트럼 변형된 PN 시퀀스를 시간 영역으로 역변환하기 위해 IFFT를 사용하지만 고정필터를 사용할 경우에는 미리 변형된 PN 시퀀스를 시간 영역

에서 바로 삽입하므로 IFFT와 같은 추가적인 연산이 필요 없어진다. 즉, 고정필터를 사용하면 심리음향모델 과정과 IFFT 등의 추가 연산 과정이 없어지면서 900 프레임 전체에 대한 부호화 수행시간이 5.626초에서 0.632초로 줄어들게 되었고 결과적으로 심리음향모델을 대체하는 고정필터를 사용하면 워터마크 부호화의 총 연산량이 기존 것의 11.2%로 크게 줄어들음을 알 수 있다.

IV. 복호화 성능 향상 기법

1. 전처리 과정 (Pre-Processing)

전처리 과정은 워터마크가 삽입된 오디오 신호 $y'(n)$ 에서 워터마크 신호 $\tilde{w}''(n)$ 을 추출해 내는 과정으로 식 (5)와 같이 $y'(n)$ 으로 원신호를 추정하여 $\hat{x}(n)$ 을 만들고 $y'(n)$ 에서 이 추정된 신호를 제거하는 것과 같다. 신호의 추정은 오디오 신호 모델링을 통해 할 수 있다. 본 논문에서는 선형예측(Linear Prediction)으로 AR 모델을 만들고 이것을 이용하여 $y'(n)$ 으로부터 워터마크 신호를 추출하도록 하였다.

$$\begin{aligned} \tilde{w}''(n) &= y'(n) - \hat{x}(n) \\ &= x'(n) + \tilde{w}'(n) - \hat{x}(n) \\ &= \tilde{w}'(n) + e(n) \end{aligned} \quad (5)$$

워터마크 정보는 전처리 과정에서 추출된 워터마크 신호 $\tilde{w}''(n)$ 과 부호화기에서 삽입한 PN 시퀀스 $w(n)$ 의 구하여 검출하는데, 이때 $\tilde{w}''(n)$ 이 $w(n)$ 과 어느 정도의 상관도를 가지느냐에 따라 워터마크 복호화 성능이 결정된다. 따라서 전처리, 과정의 워터마크 추출 성능이 복

호화 성능을 좌우한다고 할 수 있다. 한편, 부호화기에서도 전처리 과정에 의한 워터마크 추출이 잘 이루어지도록 워터마크를 삽입하여 전체적인 복호화 성능을 향상시킬 수 있다.

2. 복호화 성능 향상을 위한 PN 시퀀스의 변형

복호화의 마지막 과정은 전처리 과정에서 추출해낸 워터마크 신호 $\tilde{w}''(n)$ 과 삽입시에 사용한 PN 시퀀스 $w(n)$ 의 상관도를 구하여 워터마크 데이터를 검출하는 것이다. 전처리 과정에서 추출된 워터마크 신호 $\tilde{w}''(n)$ 은 이미 원래의 PN 시퀀스 $w(n)$ 과는 많이 달라져있기 때문에 $\tilde{w}''(n)$ 과 $w(n)$ 의 상관도를 구하는 것보다는 $w(n)$ 이 $\tilde{w}''(n)$ 으로 변형된 경로를 보상해준 다음 상관도를 구하는 것이 더욱 효과적이다. 앞의 식 (3), 식 (4), 식 (5)로부터 워터마킹 과정에서 $w(n)$ 이 $\tilde{w}''(n)$ 으로 변형되는 요인을 살펴보면 다음과 같다.

- ① 부호화기에서 마스크 곡선에 의한 변형 : $w(n) \rightarrow \tilde{w}(n)$
- ② 채널을 통과할 때의 공격 등에 의한 변형 : $\tilde{w}(n) \rightarrow \tilde{w}'(n)$
- ③ 복호화기에서 전처리 과정에 의한 변형 : $\tilde{w}'(n) \rightarrow \tilde{w}''(n)$

여기에서 ②의 채널에서의 변형은 공격 등에 의한 것으로서 복호화기에서는 알지 못하기 때문에 보상이 어렵지만 ①과 ③은 그 변형 과정을 알고 있으므로 보상이 가능하다. 그림 4는 두 가지 변형 과정을 필터링을 통해 복호화단의 PN 시퀀스 $w(n)$ 에 적용하는 과정을 나타낸다. 그림 4에서 첫 번째 필터는 부호화기에서 마스크 곡선에 의한 PN 시퀀스의 변형을 보상한다. 부호화기에서 심리음향

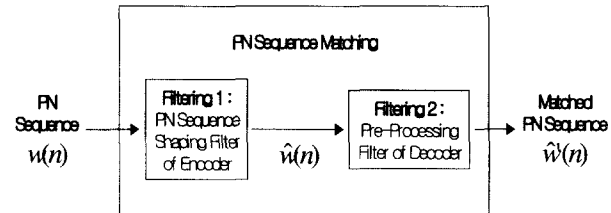


그림 4. 상관도 향상을 위한 PN 시퀀스 보상 과정
Fig. 4. PN-sequence compensation process in the decoder

모델을 사용하여 변형시킨 경우에는 심리음향모델로 구한 마스크 곡선이 변형필터가 되므로 복호화기에서도 매 프레임에 대해 심리음향모델 연산을 수행해야 한다. 또한 이때 복호화기에서는 원신호를 가지고 있지 않으므로 워터마크가 삽입된 오디오 신호로부터 마스크 곡선을 계산해야 하므로 부호화기에서 적용한 필터와 정확하게 일치하지 않는 문제가 발생한다.

그러나 본 논문에서 제안한 고정필터를 이용한 시스템의 경우에는 부호화기에서 사용한 변형필터가 고정되어있기 때문에 이를 그대로 사용할 수가 있다. ③의 변형을 보상해주는 두 번째 필터는 복호화기의 전처리 과정에서 만들어진 것이므로 이 필터를 그대로 사용하여 보상해줄 수 있다. 보상 과정에서 PN 시퀀스 $w(n)$ 은 첫 번째 필터에 의해 $\hat{w}(n)$ 으로 변형되고, $\hat{w}(n)$ 은 두 번째 필터에 의해 다시 $\hat{w}'(n)$ 으로 변형되어 최종적으로 추출된 워터마크 신호와의 매칭이 이루어지게 된다.

그림 5(a)는 추출된 워터마크 신호 $\tilde{w}''(n)$ 과 원래의 PN 시퀀스 $w(n)$ 의 상관도 결과이고 그림 5(b)는 워터마크 신호 $\tilde{w}''(n)$ 과 제안된 보상필터 구조를 통과하여 보상된 PN 시퀀스 $\hat{w}'(n)$ 과의 상관도이다. 두 결과 그림

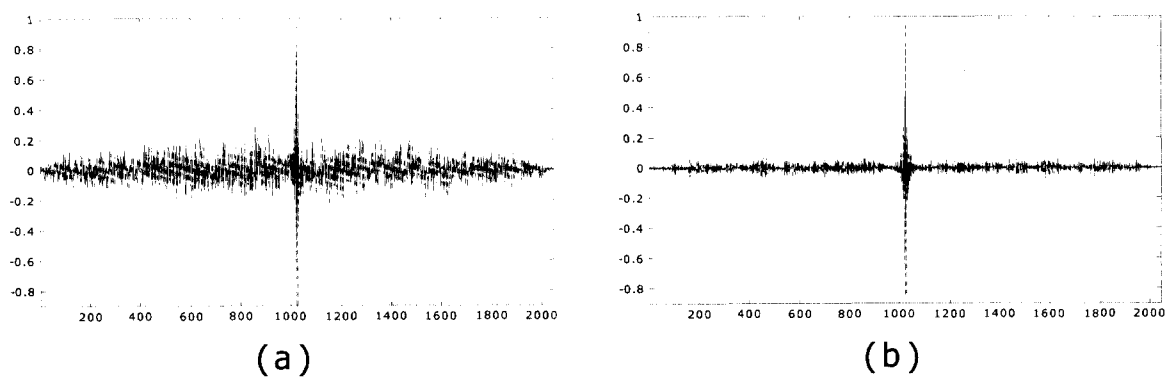


그림 5. 추출된 워터마크 신호와 (a) 원래의 PN 시퀀스, (b) 매칭된 PN 시퀀스와의 상관도 결과
Fig. 5. Correlation of extracted watermark signal (a) with plain PN-sequence, (b) with matched PN-sequence

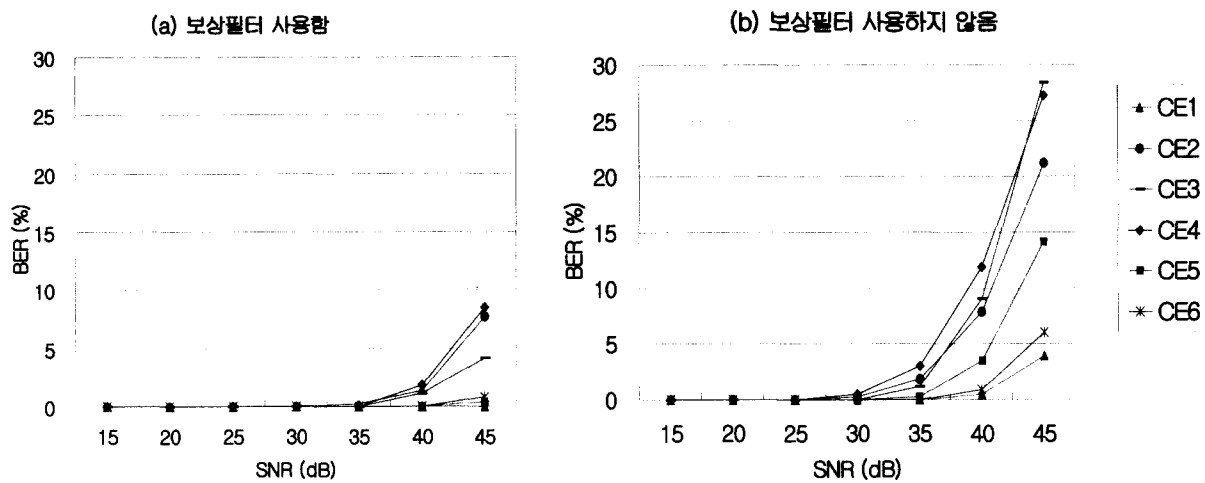


그림 6. 보상필터를 한 경우(a)와 하지 않은 경우(b)의 복호화 BER 비교
 Fig. 6. BERs with(a) and without(b) compensation filtering

을 비교하면 보상필터에 의해서 상관도가 향상되는 것을 볼 수 있다. 그림 6의 복호화 성능을 비교한 BER 결과에서도 PN 시퀀스 워터마크가 35dB SNR 이상의 작은 에너지로 삽입되었을 경우 보상필터로 PN 시퀀스를 매칭시켜 줌으로써 복호화 성능이 많이 향상되는 것을 볼 수 있다.

V. 실험 및 평가

1. 실험 내용

제안한 고정필터를 사용한 워터마킹 시스템의 성능을 알아보기 위해서 표 1의 여섯 가지 오디오 샘플에 대해 표 3에 정리한 세 가지 방법으로 워터마킹 시스템을 구성하여 각각에 대한 복호화 성능 측정 및 음질 평가를 수행하고 그 결과를 비교하였다. 실험에 사용한 오디오 샘플은 CD에서 추출한 것으로 각각 10초 길이의 곡에 대해서 실험하였다.

표 3. 성능 평가 실험을 위한 세 가지 워터마킹 시스템
 Table 3. Three systems for watermarking performance evaluation

시스템 구성	부 호 화 기	복 호 화 기	
	PN 시퀀스 변형	보상필터 1	보상필터 2
방법_1	0 (심리음향모델)	0 (심리음향모델)	0
방법_2	X	X	0
방법_3	0 (고정필터사용)	0 (고정필터사용)	0

세 가지 워터마킹 시스템 중 먼저 “방법_1”은 기존의 워터마킹 시스템으로서 부호화기에서 심리음향모델을 이용하여 마스킹 곡선을 구하고 이것으로 PN 시퀀스를 변형한다. 그리고 복호화기에서도 PN 시퀀스를 매칭시키기 위해 심리음향모델로 복호화기 입력 신호에 대한 마스킹 곡선을 구한다. 구해진 마스킹 곡선과 전처리 과정 필터로 PN 시퀀스를 변형하여 상관도를 높여준 다음 워터마크 데이터를 복호화한다.

“방법_2”는 PN 시퀀스를 변형하지 않고 시간축에서 그 크기만 조절하여 삽입한다. 따라서 복호화기에서는 전처리 과정 필터만으로 PN 시퀀스를 매칭한 후 워터마크 데이터를 추출하게 된다.

“방법_3”은 제안한 고정필터로 PN 시퀀스를 변형하여 삽입한다. 즉, “방법_1”에서처럼 워터마크를 삽입하는 모든 프레임에 대해 심리음향모델로 마스킹 곡선을 구하는 것이 아니라 고정필터로 PN 시퀀스를 보상해준다. 따라서 복호화기에서 PN 시퀀스를 매칭시킬 때 심리음향모델 과정 필요 없이 부호화기의 고정필터를 사용한다.

이와 함께 각각의 오디오 샘플 및 시스템에 대해 SNR을 변화시키면서 실험하였다. SNR은 15dB부터 45dB까지 5dB 간격으로 하여 총 7 가지의 경우에 대해 실험하였으며 이때, SNR에 따른 워터마킹 성능 비교가 공정하도록 하기 위해 식 (1)의 SNR을 각 오디오 샘플의 일정한 구간에 대해 측정하여 맞추어 주었다. 복호화 성능은 모든 SNR 경우에 대해 측정하였으며 단, 음질 평가는 이 중에서 15dB, 30dB, 40dB 세 가지에 대해서만 실시하였다. 복

호화 성능 측정은 식 (6)에 의해 BER을 계산하였다. 실험에서 각 오디오 샘플에 삽입한 정보량은 총 430 비트이고 따라서 약 43bps의 워터마크 데이터율을 갖는다.

$$BER = \left(\frac{\text{발생한 에러 비트 수}}{\text{삽입한 전체 비트 수}} \right) \times 100 (\%) \quad (6)$$

음질 평가는 ITU-R의 권고안 BS.1116을 참고하여 이에 준한 주관적 음질 평가를 실시하였고 이를 통해 얻은 점수를 평균 내어 측정하였다^[6]. 실험은 10여 명의 방음 및 잔향 설계를 한 오디오 룸에서 하였으며 3명의 오디오, 음악 평론가를 포함한 오디오 및 음성 신호처리 전공자 9명에 대해 실시하였다. PN 시퀀스를 이용한 워터마킹에서는 삽입되는 PN 시퀀스가 오디오 신호에 대해 추가적인 잡음이므로 최대한 들리지 않는 것이 좋다. BS.1116의 음질 평가 점수 기준에 의하면 4점 미만인 경우 잡음이 약간의 거슬림이 있는 것이므로 워터마크로써 사용되기 어렵다.

2. 음질 평가 결과

먼저 세 가지 워터마킹 시스템에서 SNR 15dB, 30dB, 40dB 각각으로 워터마크를 삽입한 후 측정된 음질 평가 결과는 그림 7과 같다. 평가 점수는 오디오 신호에 따라 조금 다르게 나왔지만 전체적으로 SNR에 따라 비슷한 점수를 나타내었다. 우선 심리음향모델을 사용한 워터마킹 시스템(방법_1)은 테스트한 모든 SNR에 대해서 원신호와 거의 구별이 되지 않는 5점에 가까운 점수를 받았다. 특히 실험 중에서 워터마크 신호가 최대의 크기로 들어간 15dB SNR에서도 잡음이 거의 들리지 않음을 알 수 있다.

고정필터를 사용한 워터마킹 시스템(방법_3)은 15dB의 경우 약간의 잡음이 들리는 것이 있었지만 심리음향모델

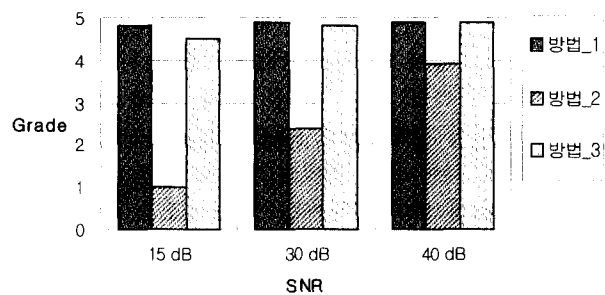


그림 7. 주관적 음질 평가 결과
Fig. 7. Results of subjective quality assessment

을 사용한 것과 비슷한 결과가 나왔으며 충분히 워터마크로써 사용될 수 있음을 알 수 있다. 하지만 “방법_2”의 PN 시퀀스를 변형하지 않고 삽입한 경우에는 15dB와 30dB에 대해서는 전혀 워터마크로 사용할 수 없을 정도로 잡음이 크게 거슬렸고 40dB SNR에서도 약간의 잡음이 들렸다.

결국 PN 시퀀스를 변형하지 않고 삽입하는 경우에는 SNR이 40dB보다 커야 워터마크로써 사용될 수가 있다. 이와 같은 음질 평가 결과를 바탕으로 PN 시퀀스를 변형한 후 삽입함으로써 SNR 측면에서 25dB 이상의 이득을 가져올 수 있으며, 또한 제안한 고정필터를 사용한 방법의 경우 심리음향모델을 사용한 것과 거의 유사한 음질 특성을 나타내는 것을 확인할 수 있다.

3. MPEG Layer-III 부호화에 대한 강인성 측정

제안된 고정필터를 사용했을 때의 MPEG Layer-III (MP3) 부호화에 대한 강인성을 테스트하기 위해 앞에서 사용한 표 1의 오디오 샘플에 워터마크를 삽입한 후 MP3 부호화/복호화를 하여 워터마크 복호화 BER을 측정하였다. 이때 워터마크는 들리지 않은 범위에서 큰 에너지를 갖도록 삽입하여 워터마크 검출 효율을 높여주었다. MP3 부호화/복호화는 Fraunhofer의 코덱을 사용하여 64kbps, 80kbps, 96kbps의 비트율에 대해 실험하였고 MP3 부호화/복호화 과정에서 생기는 시간 지연을 보상하여 모든 동기를 맞추어준 후 워터마크를 검출하였다.

표 4의 결과에서 80kbps 비트율로 압축한 경우에는 비교적 낮은 에러율을 보였고 96kbps에 대해서는 거의 완벽한 복호화가 가능하였다. 하지만 64kbps의 경우 에러율이 40% 정도로 증가하여 낮은 비트율의 압축에 의해 워터마크가 많이 없어짐을 알 수 있다. 그리고 이때는 오디오 샘플에 따라서 그 결과 차이가 많이 나는 것을 볼 수 있다.

표 4. MP3 부호화/복호화 후의 워터마크 복호화 BER 측정 결과
Table 4. BERs of watermark detection after MP3 encoding/decoding process

구분	삽입률	64kbps	80kbps	96kbps
CE1	0.0 %	42.2 %	0.75 %	0.5 %
CE2	0.0 %	2.5 %	0.25 %	0.0 %
CE3	0.0 %	29.0 %	8.75 %	0.5 %
CE4	0.0 %	39.5 %	2.0 %	0.0 %
CE5	0.0 %	44.0 %	8.25 %	0.0 %
CE6	0.0 %	42.25 %	3.5 %	0.0 %

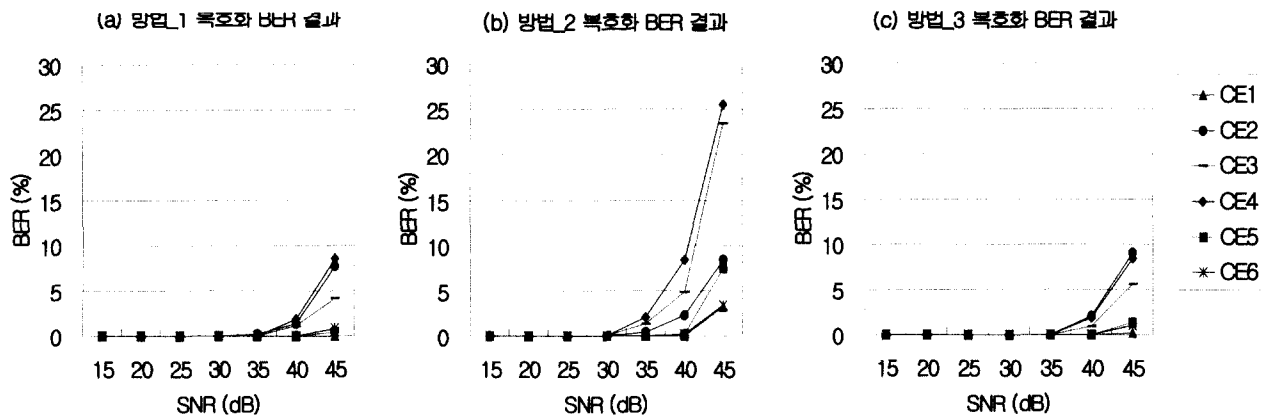


그림 8. 세 가지 방법의 복호화 BER 결과 비교
Fig. 8. Comparison of BER results for three methods

낮은 비트율의 압축에서는 고주파 성분을 제거하고 부호화 하므로 고정필터로 인해 높은 주파수 영역으로 삽입된 워터마크 신호가 대부분 제거되기 때문인데, 낮은 비트율의 압축 후에도 워터마크가 살아남기 위해서는 저주파 영역에 워터마크를 삽입하는 추가적인 장치가 필요할 것이다.

4. 성능 측정 결과

세 가지 워터마크 시스템에 대해 워터마크를 각각의 SNR로 삽입한 후 추출함으로써 워터마크 데이터의 복호화 BER을 측정하였다.

그림 8은 각 시스템의 SNR에 따른 BER 결과이다. 그림 8에서 30dB SNR까지는 공격이 없는 경우 세 가지 방법 모두 에러 없는 워터마크 복호화가 가능하였다. 그러나 "방법_2"의 '직접 삽입' 시스템에서는 음질 평가 결과에서 30dB SNR에서도 잡음이 많이 들렸기 때문에 워터마크로써 사용되기가 어렵다는 것을 고려해야 한다. 35dB SNR 이상부터는 조금씩 복호화 오류가 발생하였는데, "방법_1"과 "방법_3"의 경우가 "방법_2"보다 더 좋은 BER 결과를 나타내었다. 즉, 같은 SNR로 삽입한 경우에 PN 시퀀스를 변형하여 삽입함으로써 비직각성과 함께 강인성 측면에서도 성능 향상을 가져올 수 있다. 실험에서 "방법_1"(그림 8(a))과 "방법_3"(그림 8(c))은 서로 비슷한 BER 결과를 보여 주었는데 이로써 제안한 고정필터를 사용한 워터마크 시스템이 심리음향 모델을 사용한 것과 유사한 음질 및 복호화 성능을 얻을 수 있다.

VI. 결론

본 논문에서는 대역확산 기반 워터마크 시스템의 성능을 향상시킬 수 있는 효율적인 부호화기 및 복호화기 구조를 제안하였다. 먼저 부호화기에서는 고정필터로 PN 시퀀스를 변형하여 심리음향모델을 사용하는 것에 비해 전체 연산량이 11% 정도로 줄어든 간단한 부호화기 이루어지도록 하였다. 그리고 제안한 고정필터를 사용한 워터마크 부호화기가 심리음향모델을 사용한 것과 유사한 워터마크 성능을 가짐을 음질 평가 및 복호화 성능 측정 실험을 통해 확인하였다.

복호화기에서는 원신호 없이도 워터마크의 복호화가 가능하도록 전처리 과정을 도입하여 삽입된 워터마크 신호를 추출하도록 하였다. 또한 추출된 워터마크 신호와의 상관도를 향상시키기 위한 PN 시퀀스 보상필터 구조를 제안하여 복호화 효율을 높여 주었다. 보상필터에 따른 복호화기의 성능 측정 비교에서 보상필터를 사용한 복호화기의 BER이 사용하지 않은 복호화기의 BER보다 10% 이상 줄어드는 것을 확인하였다.

PN 시퀀스를 삽입하는 대역확산 방법 워터마크 시스템은 PN 시퀀스의 자기상관특성을 이용하여 워터마크를 효율적으로 검출할 수가 있고 잡음에 강인하다는 장점을 갖는다. 그러나 PN 시퀀스를 들리지 않도록 변형하기 위한 심리음향모델은 그 복잡성으로 인해 구현에 많은 제약이 된다.

따라서 본 논문에서 제안한 고정필터를 사용하면 휴대 기기 등에서의 워터마크 삽입 구현이 수월해진다. 그렇지만, 이렇게 심리음향모델의 공통된 마스킹 특성

에 의해 설계된 고정필터는 그림 3처럼 고역통과의 특성을 갖게 되므로 스펙트럼 변형된 PN 시퀀스는 15KHz 이상의 고주파 영역에 대부분의 에너지를 갖게 된다. 이러한 경우 인위적으로 오디오 신호의 고주파 영역을 잘라내 워터마크를 없앨 수 있으므로 소유권 증명을 위한 응용 등에는 사용되기가 어렵지만, 부호화 및 복호화 과정이 매우 간단하고 고음질과 높은 데이터율을 얻을 수 있기 때문에 휴대 기기 등에서 워터마크 정보를 계속적으로 변경해야 하는 경우나 부가정보를 오디오 신호에 삽입하여 전달하는 목적의 어플리케이션으로써 활용될 수 있는 효율적인 워터마킹 방법이 될 수 있다. 차후에는 제안한 기법을 적용한 실제적인 워터마킹 시스템의 구현에 대한 연구가 이루어질 것이다.

참 고 문 헌

- [1] F.A.P. Petitcolas, R.J. Anderson, M.G. Kuhn, "Information Hiding-A Survey," *Proc. of the IEEE* Vol. 87, 7, pp. 1062-1078, Jul. 1999.
- [2] I.J. Cox, M.L. Miller, J.A. Bloom, "Watermarking applications and their properties," *Proc. of Int'l Conf. on Information Technology Coding and Computing*, pp. 6-1, 02000.
- [3] W. Bender, D. Gruhl, N. Morimoto, A. Lu, "Techniques for Data Hiding," *IBM System Journal*, Vol. 35, No 3&4, 1996.
- [4] I.J. Cox, J. Kilian, T. Leighton, and T. Shamon, "A Secure, Robust Watermark for Multimedia," *Proc. of Workshop on Information Hiding, Univ. of Cambridge, U.K.*, pp. 175-190, May 30 Jun. 1, 1996.
- [5] L. Boney, A.H. Tewfik, K.N. Hamdy, "Digital Watermarks for Audio Signals," *Proc. of the Third IEEE Int'l Conf. on Multimedia Computing and Systems*, pp. 473-480, 1996.
- [6] 오현오, 김현욱, 윤대희, 석종원, 홍진우, "강인한 오디오 워터마킹을 위한 새로운 반향 커널 설계," *한국음향학회지*, 20권, 2호, 2001년, 2월.
- [7] Bernard Sklar, *Digital Communications Fundamentals and Applications*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1988.
- [8] ISO/IEC JTC1/SC29/WG11 No.71 "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s-CD 11172-3 (Part3, MPEG-Audio)", 1992.
- [9] E. Zwicker, *Psychoacoustics*, Springer-Verlag, New York, 1982.
- [10] F.A.P. Petitcolas, "Towards 'Robust' Watermarks," *IEE Seminar on Secure Images and Image Authentication (Ref. No. 2000/039)*, pp. 1/1-1/7, 2000.

저 자 소 개



김 현 욱

2000년 2월 : 연세대학교 전자공학과 졸업 (공학사)
 2000년 3월~현재 : 연세대학교 대학원 전기전자공학과 석사과정
 주관심분야 : 오디오 워터마킹, 오디오 신호처리



오 현 오

1996년 2월 : 연세대학교 전자공학과 (공학사)
1998년 2월 : 연세대학교 대학원 전자공학과 (공학석사)
1998년 3월~현재 : 연세대학교 대학원 전기전자공학과 박사과정
주관심분야 : 디지털 신호처리, 오디오 워터마킹, 오디오 부호화



김 연 정

2001년 2월 : 연세대학교 전자공학과 졸업 (공학사)
2001년 3월~현재 : 연세대학교 대학원 전기전자공학과 석사과정
주관심분야 : 오디오 워터마킹, 오디오 신호처리



윤 대 회

1977년 2월 : 연세대학교 전자공학과 (공학사)
1979년 8월 : Kansas State Univ. (공학석사)
1982년 8월 : Kansas State Univ. (공학박사)
현재 : 연세대학교 전기전자공학과 교수
주관심분야 : 적응 신호처리, 음성 신호처리, 오디오 신호처리