

# 주파수 위상 응답과 크기 응답에 QIM을 이용한 다중 오디오 워터마킹

## Multiple Audio Watermarking using Quantization Index Modulation on Frequency Phase and Magnitude Response

서예진, 조상진, 정의필\*†

(Yejin Seo, Sangjin Cho, and Uipil Chong\*†)

울산대학교 전기공학부, \*울산대학교 고래연구소

(접수일자: 2012년 8월 16일; 수정일자: 2012년 10월 16일; 채택일자: 2012년 12월 14일)

**초 록:** 본 논문에서는 주파수 위상 응답과 크기 응답에 QIM을 이용하여 워터마크를 삽입하는 다중 오디오 워터마킹을 제안한다. 워터마크 삽입은 두 단계로 나뉘며 주파수 영역에서 수행된다. 1단계에서는 주파수 위상 응답의 전체 샘플을 대상으로 QIM을 이용하여 워터마크를 삽입하고 2단계에서는 먼저 무음 구간 여부에 따라 워터마크를 삽입할 프레임을 정하고 해당 프레임에서 주파수 크기 응답의 최대값을 기준으로 정규화한 후 고정된 스텝사이즈를 적용하는 방법으로 적응적인 QIM을 구현하여 워터마크를 삽입한다. 워터마크 삽입 단계를 거친 주파수 위상 응답과 크기 응답을 역푸리에 변환을 통해 시간 영역의 신호로 복원한다. 워터마크 검출은 유클리디안 거리(Euclidean distance)를 이용하여 블라인드 검출 방식으로 수행된다. 오디오 워터마크 벤치마킹을 통한 강인성 검증에서 대부분 강인한 특성을 보였으며 특히 푸리에 공격에서 복원율은 95% 이상으로 매우 강인한 결과를 보였다.

**핵심용어:** 다중 오디오 워터마킹, 주파수 크기 응답, 주파수 위상 응답, QIM

**ABSTRACT:** This paper describes a multiple audio watermarking using Quantization Index Modulation (QIM) on frequency phase and magnitude response. Proposed embedding procedure is composed of two stage. At the first stage, the watermark is embedded on the frequency phase response using QIM. In the second stage, the watermark is embedded using adaptive QIM with the step-size that is adaptively determined using the maximum value of the frequency magnitude response of every frame. The watermark is extracted by calculating the Euclidean distance as the blind detection. The proposed method is robust against most of attacks of audio watermark benchmarking. For the Fourier attacks, the proposed method shows over 95% recovery rate.

**Key words:** Multiple audio watermarking, Frequency magnitude response, Frequency phase response, Quantization index modulation

**PACS numbers:** 43.60.Dh

## 1. 서 론

디지털 워터마킹은 멀티미디어 콘텐츠의 지적재산권을 보호하는 역할을 한다. 오디오 워터마킹은 원 오디오 신호에 워터마크를 삽입하고 필요할 때 이를 검출하여 저작권을 증명함으로써 지적재산권

을 보호하는 기술이다. 워터마크는 삽입된 후에도 음질저하가 발생하지 않도록 인간이 인지할 수 없는 데이터를 사용해야 한다. 그리고 워터마크를 검출할 때는 저작자만 알고 있는 키를 사용하여 인증된 사람에게 의해서만 검출할 수 있어야 한다. 그런데 불법 사용자가 저작권을 주장하지 못하도록 악의적인 공격으로 삽입한 워터마크를 제거하거나 왜곡시키기 위해 워터마크를 공격하는 경우가 발생한다. 삽입한 워터마크가 검출되어야 소유권을 주장할 수 있으

†Corresponding author: Uipil Chong (upchong@ulsan.ac.kr)  
Whale Research Institute, University of Ulsan, 93 Daehak-ro,  
Nam-gu, Ulsan 680-749, Republic of Korea  
(Tel: 82-52-259-2220, Fax: 82-52-259-1687)

로 워터마킹 알고리즘은 워터마크 공격에 강인해야 한다. 이에 따라 워터마크의 강인성을 향상시키는 알고리즘이 많이 연구되고 있다.<sup>[1-4]</sup>

오디오 워터마킹 기술은 삽입 영역에 따라 시간 영역과 주파수 영역으로 분류할 수 있다. 시간 영역에서 워터마크를 삽입하는 방식은 삽입 과정이 비교적 간단하고 계산복잡도가 낮지만 압축, 필터링과 같은 신호처리 공격에 약하다. 이에 비해 주파수 영역에서 워터마크를 선형적으로 삽입하면 시간 영역에서 워터마크를 삽입하는 방식보다 워터마크 공격에 강인하지만 이 방식 또한 워터마크 공격을 받으면 여러 가지 신호 변형에 의해 워터마크가 민감하게 영향을 받는다.<sup>[5]</sup> 주파수 영역의 워터마킹 방식은 심리음향학의 주파수 마스킹 특성을 이용하여 인간이 인지할 수 없는 데이터를 워터마크로 삽입하는 방식이다. 특히 인간의 청각은 위상 변조에 둔감하여 절대적인 위상 값을 감지하는 능력이 부족하기 때문에 위상을 이용한 워터마킹 방식은 음질 저하 문제에 대해 효과적이다. 위상 양자화를 이용한 워터마킹 방식은 Bender 외 3명이 의해 처음으로 제안되었다.<sup>[6]</sup> 이 방식은 가장 좋은 신호 대 인지잡음 비율(signal-to-perceived noise ratio)을 가졌지만 사용자가 감지할 수 있는 왜곡을 발생시키는 단점이 있었다. Dong 외 2명은 주파수 영역에서 각 배음의 절대 위상에 QIM(Quantization Index Modulation)을 이용하여 사람이 인지하기 힘든 워터마크를 삽입하였다.<sup>[7]</sup> 음성 신호에서 이 방식을 적용하였을 때 MP3 공격에는 강인하였으나, 10 dB 이하의 가우시안 잡음 공격에서는 약한 특성을 보여 이를 보완하기 위해 Narimannejad 외 1명이 음성의 정현파 모델을 이용한 워터마킹 알고리즘을 제안하였다.<sup>[8]</sup> 이 방식은 음성 신호에서 정현파의 위상 성분을 양자화시키는 방법으로 왜곡을 최소화하였고, 가우시안 잡음 공격에 특히 강하였다. 그러나 음성 신호와 일부 워터마크 공격에 대해서만 제한하여 실험했기 때문에 강인성 여부를 검증하기에는 어려움이 있다.

본 논문에서는 주파수 위상 응답과 크기 응답에 QIM을 이용하여 워터마크를 삽입하는 강인한 다중 워터마킹 방식을 제안한다. 워터마크가 삽입된 주파수 위상 응답과 크기 응답은 워터마크 공격을 받았

을 때 강인성을 가지는 공격의 종류가 다르다. 그러므로 이들을 다중 워터마킹 방식으로 사용할 경우 서로 다른 종류의 공격에 대한 강인성을 보완할 수 있다. 그리고 단일 워터마킹 방식인 경우에는 워터마크 공격에 대한 강인성을 높이기 위해 가중치를 증가시켜서 사용하지만 제안한 방식에서는 보완할 수 있는 공격에 대한 강인성을 높이기 위하여 가중치를 무리하게 증가시킬 필요가 없기 때문에 상대적으로 작은 가중치를 사용할 수 있으므로 워터마크에 의한 음질저하를 감소시킬 수 있다. 본 논문에서는 FFT(Fast Fourier Transform)으로 주파수 위상 응답과 크기 응답을 구한다. 그리고 QIM을 이용하여 각각의 주파수 응답에 워터마크를 삽입한다. 본 논문에서는 제안하는 알고리즘의 강인성 검증을 위해 개수를 제한시킨 주파수 위상 응답을 이용한 방식, 주파수 크기 응답에 QIM을 적용한 방식, 웨이브렛(wavelet) 영역에서 적용한 QIM 워터마킹 방식, 여러 영역에서 Cox 방식<sup>[9-10]</sup>을 적용한 워터마킹 방식들을 비교하였다. 2장에서는 본 논문에서 제안한 방법을 설명하고 3장에서는 제안한 방식과 기존 방식의 실험 결과를 비교하였으며 끝으로 4장에서는 결론을 맺는다.

## II. 제안한 알고리즘

### 2.1 삽입 알고리즘

본 논문에서는 Fig. 1의 블록도와 같이 주파수 위상 응답과 크기 응답에 워터마크를 각각 삽입한다. 워터마크 삽입 과정은 다음과 같다. 먼저 오디오 신호에 사각 윈도우(rectangular window)로 오버랩(overlap)없이 2048 샘플을 취하고 푸리에 변환을 통해 주파수 영역에서의 위상 응답과 크기 응답을 구한다. 1단계는 주파수 위상 응답을 이용한 워터마킹 방식이다. 식(1)의 방법으로 QIM을 이용하여 각각 메시지 비트 0과 1에 따라 프레임별로 1개의 메시지 비트를 워터마크로 삽입한다.

$$s_{\text{phase}}(\mathbf{x}_{\text{phase}}; m) = \begin{cases} \Delta_{\text{phase}}(m) \times \lfloor \mathbf{x}_{\text{phase}} / \Delta_{\text{phase}}(m) \rfloor & , m = 0 \\ \Delta_{\text{phase}}(m) \times \lfloor (\mathbf{x}_{\text{phase}} / \Delta_{\text{phase}}(m) - 0.5) + \Delta_{\text{phase}}(m) / 2 \rfloor & , m = 1 \end{cases} \quad (1)$$

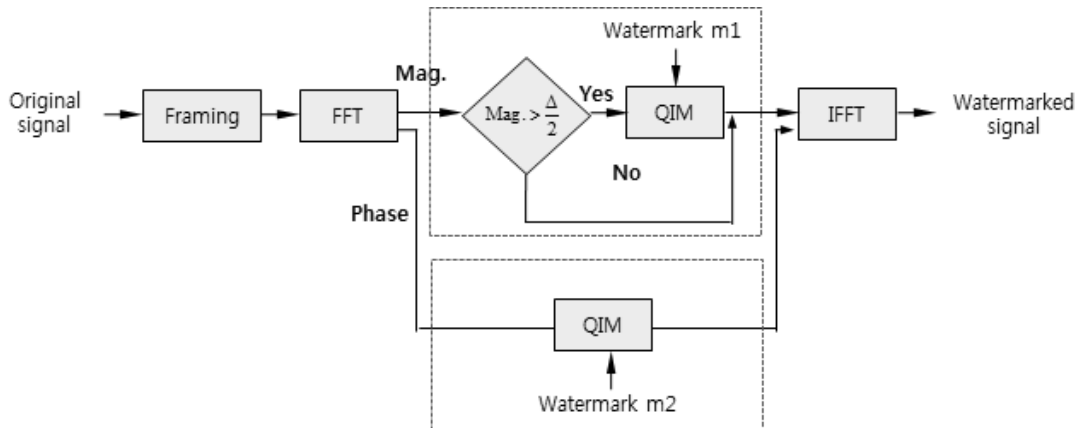


Fig. 1. Block diagram of the watermark embedding process.

여기서  $s_{phase}(\mathbf{x}_{phase}; m)$  은 원신호의 주파수 위상 응답 벡터  $\mathbf{x}_{phase}$  에 워터마크  $m$  이 삽입된 신호이고,  $\Delta_{phase}$  는 스텝 사이즈이며 본 논문에서는  $\pi/32$  를 사용한다. 그 다음에는 주파수 크기 응답에 워터마크를 삽입하는 방식으로 우선 워터마크를 삽입했을 때 음질 열화를 발생시킬 수 있는 무음구간인지를 판단하여 워터마크의 삽입 가능 여부를 결정한다. 무음구간은 식(2)의 영교차율을 통해 구할 수 있다.

$$z = \frac{1}{L} \sum_{i=1}^{L-1} |sign(\mathbf{x}_{mag}^i) - sign(\mathbf{x}_{mag}^{i-1})|. \quad (2)$$

여기서  $L$  은 프레임의 길이,  $\mathbf{x}_{mag}^i$  는  $i$  번째의 원신호 주파수 크기 응답 벡터이고  $sign(\cdot)$  은 샘플의 부호를 반환하는 함수이다. 이 단계에서는 디더 변조(dither modulation)와 양자화를 이용하여 프레임마다 1개의 메시지 비트를 워터마크로 삽입하며, 디더 벡터(dither vector)  $\mathbf{d}$  는 식(3)으로 정의한다.<sup>[6][11]</sup>

$$d[k,1] = \begin{cases} d[k,0] + \Delta_k/2, & d[k,0] < 0 \\ d[k,0] - \Delta_k/2, & d[k,0] \geq 0. \end{cases} \quad (3)$$

$k$  는 삽입될 워터마크 인덱스이고,  $d[k,0]$  은  $[-\Delta_k/2, \Delta_k/2]$  에 균일하게 분포하는 유사 랜덤 시퀀스이며  $\Delta_k$  는 균일 양자화의 스텝 사이즈이다. 그리고  $d[k,0]$  와  $d[k,1]$  은 각각 메시지 비트 0과 1을 삽입하기 위해 사용하는 디더 벡터이다. 본 논문에서는 음질 열화를 최소화하고 공격에 강인하도록 식(4)에 표시

한 식을 이용하여 프레임마다 주파수 크기 응답의 최대값을 기준으로 정규화하는 방법을 사용하여 적응적으로 변하는 스텝 사이즈를 사용한다.

$$\Delta_{mag,k} = \lfloor \frac{\max(|\mathbf{x}_{mag}^k|)}{\Delta_T} \rfloor. \quad (4)$$

$\Delta_T$  는 신호 전체에 대한 고정된 스텝 사이즈이다. 워터마크는 식(5)와 같이 신호를 양자화시키는 방법으로 삽입한다.

$$\mathbf{s}_{mag}(\mathbf{x}_{mag}; m) = \mathbf{q}(\mathbf{x}_{mag} + \mathbf{d}(m)) - \mathbf{d}(m). \quad (5)$$

여기서  $\mathbf{s}_{mag}(\mathbf{x}_{mag}; m)$  은 원신호의 주파수 크기 응답 벡터  $\mathbf{x}_{mag}$  에 워터마크  $m$  이 삽입된 신호이다.

워터마크를 각각 삽입한 후에는 주파수 위상 응답과 크기 응답에 대해 역푸리에 변환을 이용하여 시간영역에서의 워터마크된 신호를 구할 수 있다.

## 2.2 검출 알고리즘

워터마크 검출 시에는 Fig 2의 워터마크 검출 불럭도에 나타난 것처럼 워터마크를 삽입할 때와 동일한 방법으로 워터마크가 삽입된 위치와 대상 샘플을 찾아서 주파수 위상 응답과 크기 응답에 각각 사용한 QIM 방식을 이용하여 워터마크된 신호에 따라 0과 1의 워터마크를 삽입한다. 다양한 통신 채널을 통과하거나 공격을 받은 후 수신된 신호를  $\mathbf{y}$  라고 하면,

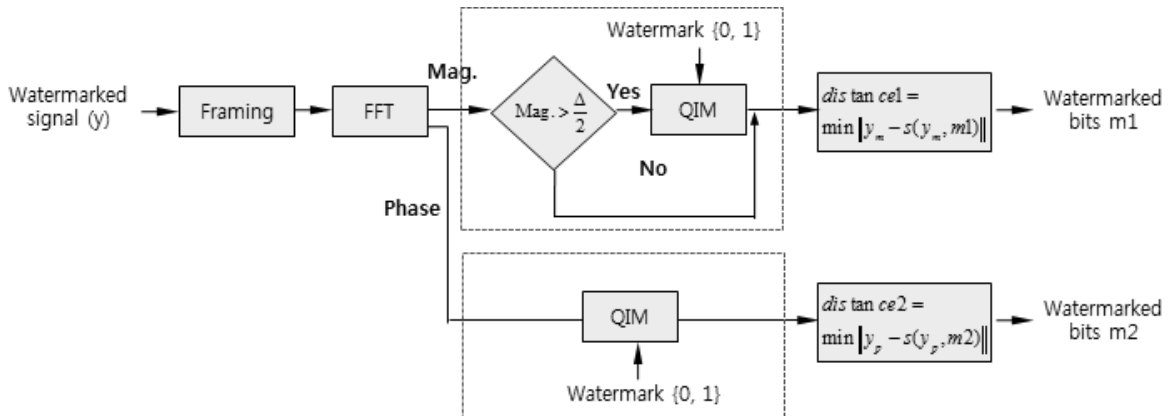


Fig. 2. Block diagram of the watermark detection process.

식(6)의 유클리디안 거리(Euclidean distance)를 이용한 블라인드 검출 방식으로 워터마크를 검출한다.

$$\hat{m}(y) = \underset{m}{\operatorname{argmin}} \|y - s(y; m)\| \quad (6)$$

주파수 위상 응답과 크기 응답에서 각각 워터마크를 검출한 후 복원율(recovery rate)을 계산하여 높은 복원율을 선택한다.

### III. 실험 결과 및 토의

본 논문에서는 제안한 워터마킹 시스템의 성능을 평가하기 위해 대상 음악으로 단일 악기를 사용한 피아노 연주곡(6초), 여러 악기를 연주한 팝(11초), 사람의 음성 신호(4초) 그리고 악기음과 음성이 모두 포함되어 전대역의 스펙트럼을 가지고 있는 클래식 음악(3분45초)을 사용하였다. 이들 신호는 44.1 kHz 샘플링에 16 비트로 양자화된 음원이며 사운드 포지 8.0에서 녹음한 디지털 오디오를 사용하였다.

#### 3.1 신뢰성 검사

제안하는 워터마킹 알고리즘은 주파수 위상 응답과 크기 응답 두 곳에서 각각 워터마크 검출 시도를 통해 워터마크의 복원율이 높은 쪽을 선택한다. 이러한 방식은 필연적으로 오경보 알람(false-alarm) 증가로 이어질 수 있으므로 이에 대한 성능을 분석하기 위해 식(7)의 오경보 알람 확률을 계산하여 신뢰

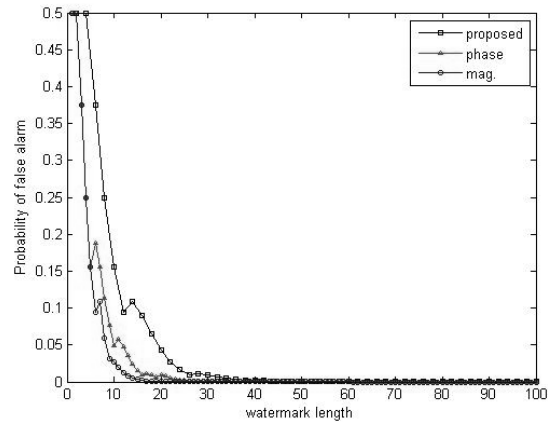


Fig. 3. Results of reliability test by probability of false-alarm : case of proposed method (square mark), case of using phase response that is the first stage of the proposed method (triangle mark) and case of using magnitude response that is the second stage of the proposed method (circle mark).

성 검사를 하였다.<sup>[12]</sup>

$$P_{fa} = 2^{-N} \sum_{r=Th}^N C_N^r \quad (7)$$

여기서  $N$ 과  $r$ 은 각각 삽입된 워터마크의 길이와 검출된 워터마크의 길이이다. 또한 문턱값은  $Th = \lceil (1 - BER) \times N \rceil$ 이며  $BER$ 은 다양한 워터마크 공격을 가한 후 계산한 전체 BER(Bit-Error-Rate)의 평균을 사용한다. 본 논문에서는 15개의 공격을 행하여 구한 BER을 사용하였다. Fig 3은 삽입된 워터마크의 길이가 (0,100]일 때 클래식 음악의 오경보 알람 확률을 나타낸 것으로 주파수 위상 응답과 크기 응

답에 삽입한 워터마크의 길이가 20비트 이상일 때 각각 오경보 알람 확률은 약 0의 값을 보인다. 그리고 제안하는 방식은 워터마크의 길이가 약 40비트 이상일 때 0의 오경보 알람 확률값을 보인다. 본 논문에서는 대상 음악인 피아노 연주곡, 팝, 음성 신호, 클래식 음악의 주파수 위상 응답에는 각각 142, 238, 106, 2325비트의 워터마크를 삽입하였고, 주파수 크기 응답에는 52, 97, 58, 946비트의 워터마크를 각각 삽입하였다. 결과적으로 제안하는 방식에는 40비트 이상인 174, 335, 142, 3271비트 길이의 워터마크가 각각 삽입되었기 때문에 본 제안하는 시스템의 오경보 확률은 0이 됨을 알 수 있다.

### 3.2 강인성 검사

본 논문에서는 강인성 검증을 위하여 Table 1과 같은 StirMark, SMDI, STEP2000 벤치마크의 시간 및 주파수 영역에서의 공격을 가하였다<sup>[13-15]</sup>. QIM을 이용한 방식은 식(8)의 복원율을 이용하여 정량화하였다.<sup>[16]</sup>

$$R = \frac{\sum_{k=1}^N B(m_k = \tilde{m}_k)}{N} \quad (8)$$

여기서  $m_i$ 는  $i$ 번째 프레임에 삽입된 워터마크이고,

$\tilde{m}_i$ 는  $i$ 번째 프레임에서 검출한 워터마크이다. 그리고  $N$ 는 전체 워터마크의 길이이며  $B()$ 는  $B(True) = 1$ 과  $B(False) = 0$ 의 불린(Boolean) 함수이다. cox 방식<sup>[9]</sup>은 식(9)의 유사도(similarity)를 계산하였으며

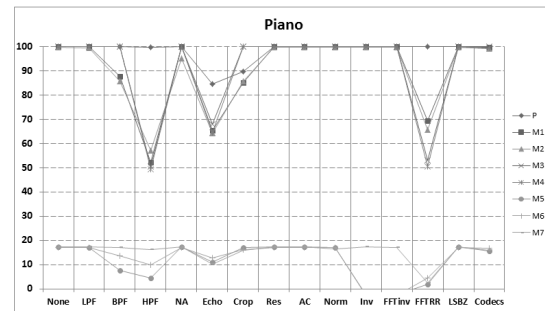


Fig. 4. Results of robustness test of piano sound.

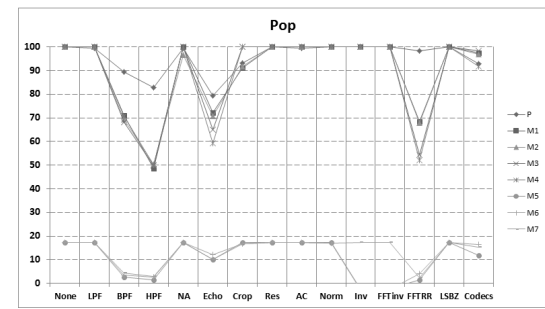


Fig. 5. Results of robustness test of pop sound.

Table 1. Attacks for robustness test.

Attack	Description	Benchmarking
None	This process does not affect the audio signal, it is a simple bypass.	StirMark
LPF	Low pass filtering with cut-off frequency of 15kHz.	StirMark
BPF	Band pass filtering with cut-off frequency of 100Hz~6kHz.	SDMI
HPF	High pass filtering with cut-off frequency of 150Hz.	StirMark
NA	Add white noise to the audio signal with -10dB signal-to-noise ratio.	STEP2000
Echo	Add an quite echo to the audio signal with parameters: 100ms delay and 0.5 feedback coefficient.	SDMI
Crop	Cut and replace samples of arbitrary part of the audio signal whose duration is 2 seconds.	-
Res	Change the sampling rate of the audio signal from 48kHz to 44.1kHz.	SDMI
AC	Compress the amplitude by changing quantization bit from 16bit to 8bit.	STEP2000
Norm	Normalize the amplitude using the maximum value.	StirMark
Inv	Invert all samples in the audio signal, like a 180 degree phase shift.	StirMark
FFTinv	Invert the real and imaginary part of the sample values in the frequency domain.	StirMark
FFTRR	Reverse only the real part of the audio signal in the frequency domain.	StirMark
LSBZ	Set all least significant bit (LSB) to zero.	StirMark
Codes	Compress the audio signal with MPEG-1 Audio Layer 3	SDMI

Fig. 4~7에 표시하였다.

$$sim(\mathbf{x}, \mathbf{x}^*) = \frac{\mathbf{x}^* \cdot \mathbf{x}}{\sqrt{\mathbf{x}^* \cdot \mathbf{x}^*}} \quad (9)$$

여기서  $\mathbf{x}^*$ 와  $\mathbf{x}$ 는 각각 워터마크된 신호와 원신호이다. 그리고 제안하는 방식에서는 주파수 위상 응답과 크기 응답에서 각각의 복원율을 구한 후 높은 복

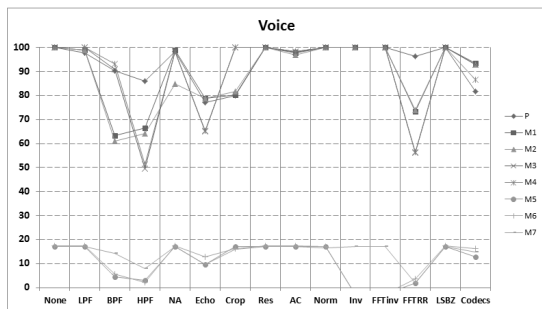


Fig. 6. Results of robustness test of voice.

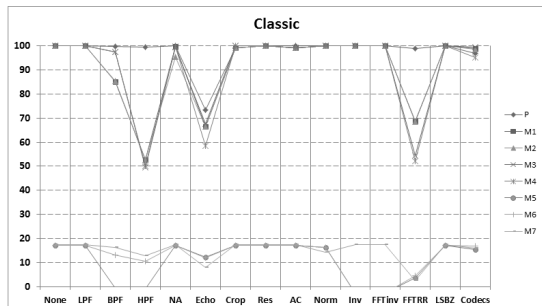


Fig. 7. Results of robustness test of classic sound.

원율을 선택하였다. 워터마크를 삽입하는 것은 소유권을 주장하기 위함이므로 다중 워터마크를 삽입한 경우에는 하나의 워터마크라도 검출이 된다면 그 소유권을 주장할 수 있기 때문이다. 제안한 알고리즘의 강인성을 검증하기 위해 주파수 영역의 대표적인 워터마킹 방식들과 비교하였으며 신뢰도를 높이기 위해 각 실험은 100번씩 수행하였다. 비교 대상은 cox가 제안한 확산대역 스펙트럼 방식을 이용한 3가지 방식(M5<sup>[9]</sup>, M6<sup>[9]</sup>, M7<sup>[10]</sup>), 확산대역 스펙트럼을 이용한 방법보다 더 강인하다고 알려진 QIM을 이용한 블라인드 워터마킹 방식 2가지(M3<sup>[17]</sup>, M4<sup>[18]</sup>) 그리고 마지막으로 음질 저하를 감소시킬 수 있는 주파수 위상 응답을 이용한 워터마킹 방식 2가지(M1<sup>[8]</sup>, M2<sup>[7]</sup>)이며 각각의 방식은 Table 2에 설명하였다.

제안하는 알고리즘은 다른 기존의 방법들과 비교하여 대부분의 공격에서 높은 복원율을 얻을 수 있었다. 특히 FFTRR 공격에는 비교 대상인 7가지 방법 모두에서 유사도는 6이하, 복원율은 약 70% 이하로 매우 취약했지만 제안하는 방식은 95%~100%의 복원율을 나타내어 매우 강인한 특성을 보였다. 이 공격은 푸리에 변환 후 실수부의 순서를 역으로 바꾸는 것으로 주파수 크기 응답에서 주요 피크의 위치가 바뀔 정도로 응답 특성이 달라지지만<sup>[19]</sup>, 주파수 위상 응답은 원 신호의 것과 대체적으로 비슷한 패턴을 보였다. 따라서 위상 응답에 워터마크를 삽입하면 주파수 크기 응답에 삽입할 때 보다 공격에 강인할 수 있다. 하지만, M1, M2와 같이 특정한 위상에

Table 2. Seven watermarking algorithms for comparison.

Comparison method	Description
P	Proposed method
M1	Embedding watermarks into the frequency phase response by phase quantization using sinusoidal modeling of the speech.
M2	Embedding watermarks into the absolute frequency phase response corresponding to each harmonic by QIM in the STFT domain.
M3	Embedding watermarks into the frequency magnitude response using adaptive QIM.
M4	Embedding watermarks into the frequency magnitude response of the selected signal using adaptive QIM. The target signal is decomposed into the subband by wavelet packet transform and the signal is selected by the subband power-based criterion.
M5	Cox's method
M6	Embedding watermarks into the largest 300 samples of DCT coefficients except DC value by using Cox's method.
M7	Embedding watermarks into the DCT coefficients whose value is distributed in the range from 0.3 to 0.4 by using Cox's method.

만 삽입한 경우에는 강인성을 보장할 수 없다. 워터마크를 삽입한 위상이 공격에 의해 달라진 경우, 워터마크를 검출할 수 없기 때문이다. 그러나 제안하는 방법은 주파수 위상 응답 전체를 대상으로 워터마크를 삽입하기 때문에 워터마크 삽입량이 많아서 일부가 공격에 의해 손실되더라도 잔재하는 워터마크의 양이 충분하여 FFTRR 공격에 매우 강인함을 보이는 것이다. 비교 대상들이 주파수 영역에서의 대표적인 워터마킹 방식임을 감안할 때 제안하는 방식이 FFTRR 공격에 강인하다는 특성은 주목할 만하다. 그러나 일부 음원과 공격의 종류에 따라 워터마크 복원율이 다소 떨어지는 현상을 볼 수 있었다. 예를 들어 피아노 연주곡인 경우 Echo 공격에서 84.65%, Crop 공격에서는 89.67%의 복원율을 보였다. 이는 M1~M4의 방식에 따른 복원율과 비교하면 공격에 강인하였으나 10 이상의 큰 유사도를 가지는 cox의 확산대역 스펙트럼 방식과 비교하면 강인성이 떨어지는 결과이다. 이때 제안한 방식의 표준편차를 계산하면 각각 4.48, 3.22로써 이것은 다른 공격에 대한 복원율의 표준 편차가 0~1인 것에 비하면 매우 높은 값이다. 이는 워터마크 복원율이 넓게 분포되어 있음을 의미하는 것이다. 100가지의 워터마크를 생성하여 실험한 결과, 복원율의 최대값은 각각 98.08%, 100%로써 매우 높게 나타났으며 이러한 현상은 팝, 음성 신호, 클래식 음악 모두에서 관찰할 수 있었다. 이러한 결과는 워터마크에 따라 제안한 방식의 복원율을 크게 높일 수 있다는 뜻도 된다. 그러나 표준편차가 크다고 해서 모든 복원율이 높아질 가능성이 있는 것은 아니다. 피아노 연주곡에서 M4의 방식을 사용하여 워터마크를 삽입하고 HPF 공격을 시도했을 때는 6.69의 가장 높은 표준편차 값이었으나 평균 복원율은 50.69%였고, 최대 복원율은 63.38%였다. 즉, 표준편차가 크다 하더라도 평균 복원율이 높아야 이러한 효과를 기대할 수 있는 것이다. 주파수 위상 응답을 이용한 최근의 워터마킹 방식인 M1과 M2의 결과와 비교했을 때 제안한 방식은 필터 공격과 푸리에 공격에서 월등히 높은 복원율을 나타냈으며 Echo와 Crop 공격에서도 더 강인한 결과를 보였다.

본 논문에서 제안한 워터마킹 방식은 기존의 대표적인 워터마킹 방식들과 비교하였을 때 대부분의 워

터마크 공격에 상대적으로 더 강인한 특성을 보였다. 특히 푸리에 공격에서 독보적인 강인성을 나타내었는데, FFTinv 공격에서는 100%의 복원율, FFTRR 공격에서는 약 98%이상의 복원율을 각각 얻을 수 있었다.

#### IV. 결론 및 향후과제

본 논문에서는 주파수 위상 응답과 크기 응답에 QIM 방식을 이용하여 워터마크를 삽입하는 강인한 다중 워터마킹 방식을 제안하였다. StriMark, SDMI, STEP2000 벤치마킹에 사용되는 공격 중 15가지를 선택하여 비교 대상들과 강인성 검증을 한 결과, 제안한 방식만 푸리에 공격에서 유일하게 강인한 특성을 보였으며 대부분의 공격에서에서도 대체로 강인한 결과를 보였다. 복원율이 다소 떨어지는 경우가 있었지만 이것은 실험적으로 워터마크를 생성하여 사용한다면 이를 보완할 수 있다. 세 가지 벤치마킹 프로그램에서 중복되는 공격을 제외하면 대략 50여 종이 되는데, 향후 다른 공격에 대해서도 실험이 수행되어야 할 것이며, 이를 바탕으로 제안한 알고리즘을 개선해 나가야 할 것이다.

#### 감사의 글

이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2011-0014554).

#### 참고문헌

1. M. D. Swanson, B. Z hu, and A. H. Tewfik, "Robust audio watermarking using perceptual masking," *Signal Process.* **66**, 337-355 (1998).
2. D. Kirovski and H. Malvar, "Robust spread-spectrum audio watermarking," in *Proc. IEEE ICASSP 2001*, **3**, 1345-1348 (2001).
3. N. Kalantari, S. Ahadi, and A. Kashi, "A Robust Audio Watermarking Scheme Using Mean Quantization in the Wavelet Transform Domain," in *Proc. IEEE ISSPIT*, 1673-1687 (1997).
4. W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Statistical characteristic-based robust audio watermarking

- for resolving playback speed modification,” DSP 1 **21**, 110-117 (2011).
5. J. S. Lee, W. Y. Kim, and E. T. Lee, “Digital watermark for protection of multimedia content” (in Korean) J. ETRI **13**, 42 (1998).
  6. W. Bender, D. Gruhl, and N. Morimoto, “Techniques for data hiding,” J. IBM Systems **35**, 313-336 (1996).
  7. X. Dong, M. F. Bocko, and Z. Ignjatovic, “Data hiding via phase manipulation of audio signal,” in Proc. IEEE ICASSP **5**, 377-380 (2004).
  8. L. Girin and S. Marchand, “Watermarking of speech signals using the sinusoidal model and frequency modulation of the partials,” in Proc. IEEE ICASSP **1**, 633-636 (2004).
  9. I. Cox, J. Killian, F. Leighton, and T. Shamoan, “Secure Spread Spectrum Watermarking for Multimedia Data,” IEEE-T-IP **6**, 1673-1687 (1997).
  10. D. H. Woo, “Audio Watermarking Using Specific Frequency Coefficients” (in Korean) J. Acoust. Soc. Kr. **22**, 269-275 (2003).
  11. B. Chen and G. W. Wornell, “Quantization index modulation: a class of provably good methods for digital watermarking and information embedding,” IEEE Trans. Inf. Theory, **47**, 1423-1443 (2001).
  12. M. Fan and H. Wang, “Chaos-based discrete fractional Sine transform domain audio watermarking scheme,” Comput. Electr. Eng., **35**, 506-516 (2009).
  13. A. Lang, Stirmark benchmark for audio, <http://www.witi.cs.uni-magdeburg.de/~alang/smba.php>, 2007.
  14. Y. Lin and W. H. Abdulla, *Audio watermarking for copyright protection* (University of Auckland, Auckland, New Zealand, Tech.1 Rep., No. 650, 2007).
  15. JASRAC and NRI Ltd., “Announcement of Evaluation Test Results for “STEP 2000,” *International Evaluation Project for Digital Watermark Technology for Music*,” <http://www.jasrac.or.jp/watermark/ehoukoku.htm>, 2000.
  16. Y. Lin and W. H. Abdulla, “An Audio Watermarking Method by Using Automatic Music Transcription Information,” in Proc. ISITA 2010, 303-308 (2010).
  17. Y. H. Lee, K. W. Lee, S. J. Kim, D. H. Won, H. K. Yang, “A Study on Optimizing Quantization Steps for QIM Watermarking Schemes,” J. Kr. Soc. Comput. Inf. **16**, 45-53 (2006).
  18. Y. J. Seo, S. J. Cho, and U. P. Chong, “Watermarking Algorithm using Power of Subbands Decomposed by Wavelet Packet and QIM,” (in Korean) J. Kr. Multimedia Soc. **11**, 1431-1437 (2011).
  19. J. S. Kang and S. J. Cho, “Audio Watermarking Using Quantization Index Modulation on Significant Peaks in Frequency Domain,” (in Korean) J. Acoust. Soc. Kr. **30**, 303-307 (2011).

## 저자 약력

### ▶ 서 예 진(Yejin Seo)



2002년 2월 울산대 전기전자및자동화 공학부(공학사)  
2004년 2월 울산대 전기전자정보시스템공학과(공학석사)  
2011년 2월 울산대 컴퓨터정보통신공학과(박사수료)  
<관심 분야> 디지털신호처리, 워터마크, DSP 응용

### ▶ 조 상 진(Sangjin Cho)



1997년 2월 울산대 전자공학과(공학사)  
2002년 2월 울산대 전자공학과(공학석사)  
2007년 2월 울산대 컴퓨터정보통신공학과(공학박사)  
2009년 3월 ~ 현재 울산대 전기공학부 연구교수  
<관심 분야> 디지털신호처리, 악기음향, DSP 응용

### ▶ 정 의 필(Uipil Chong)



1978년 2월 울산대 전기공학과(공학사)  
1980년 2월 고려대 전기공학과(공학석사)  
1986년 1월 미국 Oregon State University 전기 및 컴퓨터공학과(공학석사)  
1996년 12월 미국 뉴욕대(Polytechnic) 전자공학과(공학박사)  
1986년 9월 ~ 1990년 12월 미국 Bruckner Mobile Electronics, 전자 및 통신 엔지니어  
1994년 1월 ~ 1996년 5월 미국 뉴욕대(Polytechnic) CATT 연구소 Research Fellow  
1997년 3월 ~ 현재 울산대 전기공학부 교수  
2011년 3월 ~ 현재 울산대 고래연구소 소장  
<관심 분야> 신호처리, 고장진단, 3D사운드, 워터마킹