

Support Vector Machine을 이용한 오디오 워터마크 디코딩 모델 개발

Development of Audio Watermark Decoding Model Using Support Vector Machine

서예진, 조상진[†]

(Yejin Seo and Sangjin Cho[†])

울산대학교 전기공학부

(접수일자: 2014년 6월 18일; 수정일자: 2014년 9월 11일; 채택일자: 2014년 9월 26일)

초 록: 본 논문은 SVM(Support Vector Machine)을 이용하여 공격에 강인한 워터마크 디코딩 모델을 제안한다. 이 모델은 워터마크 된 신호에 대해 워터마크 삽입 과정을 역으로 수행한 후 SVM을 이용하여 워터마크를 검출한다. SVM을 생성하기 위해 먼저 4가지 워터마킹 알고리즘을 이용하여 삽입한 워터마크를 추출하여 데이터를 만들고, 이들의 BER(Bit Error Rate)을 이용하여 문턱값을 구한다. 이후, 이 문턱값을 기준으로 훈련 집합을 만든다. 강인성 검증을 위해 워터마크 된 신호에 StirMark, SMDI, STEP2000 벤치마킹 중에서 14개의 공격을 가하였는데, 그 결과 기존의 방법보다 PSNR(Peak Signal to Noise Ratio)과 BER이 모두 개선되었다. 특히, PSNR이 10 dB 이상인 경우에는 대부분의 공격에서 1% 이내의 BER을 갖는 우수한 성능을 보였다.

핵심용어: 오디오 워터마킹, Support Vector Machine, 디코딩 모델, 벤치마킹

ABSTRACT: This paper describes a robust watermark decoding model using a SVM(Support Vector Machine). First, the embedding process is performed inversely for a watermarked signal. And then the watermark is extracted using the proposed model. For SVM training of the proposed model, data are generated that are watermarks extracted from sounds containing watermarks by four different embedding schemes. BER(Bit Error Rate) values of the data are utilized to determine a threshold value employed to create training set. To evaluate the robustness, 14 attacks selected in StirMark, SMDI and STEP2000 benchmarking are applied. Consequently, the proposed model outperformed previous method in PSNR(Peak Signal to Noise Ratio) and BER. It is noticeable that the proposed method achieves BER 1% below in the case of PSNR greater than 10 dB.

Keywords: Audio watermarking, Support Vector Machine, Decoding model, Benchmarking

PACS numbers: 43.60.Dh

1. 서 론

워터마킹의 가장 중요한 과제는 저작권 보호이므로 강인성은 워터마킹 시스템의 중요한 요소이다. 따라서 강인한 워터마킹의 기술에 대한 관심과 연구가 증가하였다.^[1] 워터마킹 기술은 워터마크를 삽입하는 영역에 따라 일반적으로 공간 영역(spatial domain)

워터마킹 기술과 주파수 영역(frequency domain) 워터마킹 기술로 구분된다. 주파수 영역에서의 워터마킹 기술은 공간영역에서 적용되던 방식에 비해 공격에 강한 특징을 가지지만 주파수 특성상 워터마크 삽입 위치를 정확히 선정할 수 없고 계산 복잡도가 높다는 단점이 있다.^[2,3] 이에 따라 강인성을 높이는 공간영역에서의 워터마킹 기술이 발표되고 있다. 1998년에는 진폭 변조에 기반을 둔 방법이 소개되었고 2004년에는 진폭 변조의 기술을 SVM과 결합시켜서 더욱 발전시킨 방법이 발표되었다.^[4,5] 이 방법은 워

[†]Corresponding author: Sangjin Cho (sjcho75@ulsan.ac.kr)
School of Electrical Engineering University of Ulsan, 93
Daehak-Ro, Nam-Gu, Ulsan 680-749, Republic of Korea
(Tel: 82-52-259-2724, Fax: 82-52-259-1686)

터마크와 훈련시킨 데이터를 더하여 원신호에 삽입하고 훈련 집합을 SVM에 훈련시키면서 워터마크를 검출한다. SVM은 자료를 분류하는 기술인데 강인성과 정확도가 높기 때문에 큰 주목을 받고 있다.^[6] 본 논문은 이러한 연구의 일환으로 삽입 방법에는 구애받지 않으면서 SVM을 이용한 디코딩 모델을 제안한다. II장에서는 배경이론으로 기존 연구의 워터마크 삽입과 검출 과정을 묘사한다. 그리고 III장에서는 제안하는 방식을 설명하고 IV장에서는 제안하는 방식과 기존 방식의 실험결과를 비교하였으며 끝으로 V장에서는 결론을 맺는다.

II. 배경이론

본 논문은 워터마크 삽입 방법에 구애받지 않는, 기존의 방법에 SVM을 이용하는 디코딩 모델을 적용하여 워터마크를 검출하는 방법을 제안한다. 본 논문에서는 4가지 방법으로 워터마크를 삽입한 후 성능을 비교하였으며 그 방법은 다음과 같다.

2.1 주파수 영역에서 주요피크에 QIM을 적용한 오디오 워터마킹

주파수 영역에서 주요피크를 찾은 후 일반적인 QIM을 이용한 워터마킹에서 l 개의 워터마크를 삽입하기 위해서는 메시지 비트 m 이 디더벡터(dither vector) $\mathbf{d}(m) = [d_1^{(m)} d_2^{(m)} \dots d_k^{(m)} \dots d_l^{(m)}]$ 로 변환되어야 하며, 디더벡터의 원소는 Eq.(1)과 같다.

$$d_k^{(1)} = \begin{cases} d_k^{(0)} + \Delta_k/2, & d_k^{(0)} < 0 \\ d_k^{(0)} - \Delta_k/2, & d_k^{(0)} \geq 0 \end{cases} \quad (1)$$

여기서 k 는 삽입될 워터마크 인덱스이고, $d_k^{(0)}$ 은 $[-\Delta_k/2, \Delta_k/2]$ 에 균일하게 분포하는 유사 랜덤 시퀀스이며 Δ_k 는 균일 양자화기의 스텝 사이즈이다. 그러나 주요피크에 QIM을 적용한 방법에서는 적응적인 스텝사이즈를 사용한다.^[7] 적응적인 스텝 사이즈는 프레임 별 주파수 크기 응답의 최대값을 기준으로 결정하며 워터마크 m 은 Eq.(2)를 이용하여 원신호의 주파수 크기 응답 벡터 \mathbf{x}_m 에 삽입한다.

$$\mathbf{s}_m(\mathbf{x}_m; m) = \mathbf{q}[\mathbf{x}_m + \mathbf{d}(m)] - \mathbf{d}(m). \quad (2)$$

$\mathbf{s}_m(\mathbf{x}_m; m)$ 은 워터마크 된 신호이다. 워터마크 검출 시에는 워터마크를 삽입할 때와 동일한 방법으로 워터마크가 삽입된 위치와 대상 샘플을 찾아서 QIM 방식을 이용하여 수신된 신호 \mathbf{y} 에 0과 1의 워터마크를 각각 삽입한다. 그리고 Eq.(3)의 유클리디안 거리(Euclidean distance)를 이용한 블라인드 방식으로 워터마크를 검출한다.

$$\hat{m}(\mathbf{y}) = \underset{m}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{s}(\mathbf{y}; m)\| \quad (3)$$

2.2 주파수 위상 응답에 QIM을 적용한 오디오 워터마킹

주파수 위상 응답의 전체 샘플을 대상으로 Eq.(4)의 QIM 방법을 이용하여 워터마크를 삽입하는 방식이다.^[8] 프레임 별로 각각 메시지 비트 0과 1에 따라 1개의 메시지 비트를 워터마크로 삽입한다.

$$s_p(\mathbf{x}_p; m) = \begin{cases} \Delta_p(m) \times \left[\left(\frac{\mathbf{x}_p}{\Delta_p(m)} \right) \right], & m = 0 \\ \Delta_p(m) \times \left[\left(\frac{\mathbf{x}_p}{\Delta_p(m)} - 0.5 \right) + \frac{\Delta_p(m)}{2} \right], & m = 1. \end{cases} \quad (4)$$

$s_p(\mathbf{x}_p; m)$ 은 원신호의 주파수 위상 응답 벡터 \mathbf{x}_p 에 워터마크 m 이 삽입된 신호이고, Δ_p 는 스텝 사이즈이며 $\pi/32$ 를 사용한다. 워터마크를 삽입할 때와 같은 방법으로 수신된 신호를 FFT하여 각 프레임의 주파수 위상 응답의 전체 샘플에 0과 1의 워터마크를 각각 삽입한 후 Eq.(3)의 유클리디안 거리를 이용하여 워터마크를 검출한다.

2.3 웨이블릿 패킷을 이용한 오디오 워터마킹

공격에 강인하면서 사람이 인식하지 못하도록 신호의 피치가 포함된 대역을 워터마크 위치로 선정하여 워터마크를 삽입한다.^[9] 웨이블릿 패킷 변환을 이용하여 대역을 나누고 대역별 에너지를 각각 계산한 후 최대 에너지를 가지는 대역을 피치로 선정한다.

선정된 대역의 신호는 FFT를 이용하여 주파수 영역으로 변경한 후 주파수 크기 응답의 최대값을 1로 정규화 하여 프레임마다 각기 다른 적응적인 스텝 사이즈를 가지는 QIM으로 워터마크를 삽입한다. 원 신호의 주파수 크기 응답과 QIM을 이용하여 양자화한 신호를 비교하면, 원 신호의 주파수 크기 응답이 $\Delta_k/2$ 보다 작은 경우에는 음수로 양자화된다. 이러한 현상을 방지하기 위해 문턱값 이상의 값을 가지는 샘플에 대해 Eq.(2)의 방법으로 워터마크를 삽입한다. 검출은 유클리디안 거리를 이용한다.

2.4 반향 신호를 이용한 오디오 워터마킹

$[t_1, t_0]$ 의 오프셋을 가지는 워터마크 $[1, 0]$ 의 반향 커널을 각각 생성하고 양극 반향(bipolar echo)을 사용한다.^[10] 각 프레임에 할당된 워터마크에 따라 반향 커널, $h[n]$ 을 결정하여 오디오의 프레임 신호와 컨볼루션하여 반향을 삽입한다.

$$h[n] = \delta[n] + \alpha_1 \delta[n - d_1] - \alpha_1 \delta[n - (d_1 + 1)], \quad (5)$$

여기서 α_1 는 반향의 크기를 표시하는 스케일 팩터이고, d_1 는 지연시간인 오프셋이다. 반향 신호의 오프셋 값을 찾아서 워터마크를 검출한다.

III. 제안하는 방법

본 논문에서는 SVM을 이용한 새로운 워터마크 디코딩 모델을 제안한다. 워터마크 삽입 과정을 역으로 수행하여 $N \times N$ 의 데이터를 만들어서 각 행과 열의 BER을 구하고 이를 SVM에 적용하여 워터마크를 검출한다. SVM의 생성과정은 Fig. 1과 같다. 먼저 워터마크와 동일한 크기인 $N \times N$ 의 데이터를 n 개 준비한다. Fig. 1의 R-(a)와 C-(a)에 나타난 각 행과 열의 BER을 식(6)과 같이 계산한다.^[11]

$$BER = \frac{\text{오류가 생긴 비트수}}{\text{삽입된 총 비트수}}. \quad (6)$$

n 개의 데이터에서 동일한 방법을 수행하여 $n \times N$ 개의 BER을 각각 구한다. 그리고 극단적인 값에 대한 영향을 줄이기 위해 Fig. 1의 R-(b), C-(b)와 같이 훈련 집합의 각 행과 열의 BER의 중간값을 구하고 이를 SVM의 True/false의 기준으로 사용한다. Fig. 1의 각 (c)에서 보는 것처럼 Eq.(7)의 수식으로 가중치 α 를 곱하여 문턱값을 계산한다.

$$Thr_i = \alpha \cdot \text{median}(BER_j), \quad (7)$$

$$1 \leq i \leq n, 1 \leq j \leq N,$$

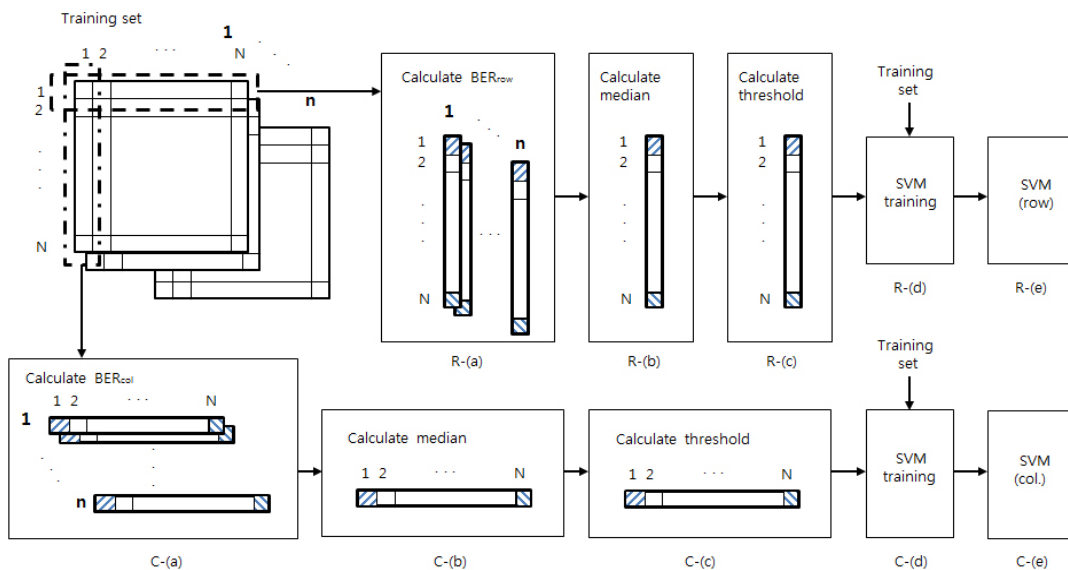


Fig. 1. Block diagram of proposed model.

계산한 문턱값을 판별기준으로 정하여 훈련 집합을 학습시켜서 행과 열에 대한 SVM을 생성한다. 워터마크된 수신 신호는 삽입과정을 역으로 수행하여 $N \times N$ 의 데이터 집합을 만든 후 각 행과 열에 대한 BER을 계산한다. 이를 앞에서 구한 행과 열의 SVM 모델에 적용하여 워터마크를 검출한다.

IV. 실험결과 및 토의

본 논문에서는 제안하는 워터마킹 시스템의 성능을 평가하기 위해 클래식(8:30), 팝(7:58), 락(9:50), 메탈(9:11)의 4가지 음악샘플을 사용하였다. 그리고 강인성 검증을 위해 4가지 워터마킹 알고리즘을 이용하여 워터마크를 삽입하였으며, StirMark, SMDI, STEP2000 벤치마크 중에서 Table 1에 설명한 14개의 공격을 가하였다.^[12-14] 훈련 집합은 여러 장르의 7곡을 사용하였으며 이는 워터마킹 시스템의 성능 평가에 사용한 곡과 중복되지 않는다. 그리고 다양한 혼

련 집합을 만들기 위해 4가지 워터마킹과 14개의 워터마킹 공격을 가하여 만든 392개 샘플을 사용하였다. 모든 신호는 44.1 kHz 샘플링에 16 비트로 양자화된 음원이다. SVM은 비선형 분류의 문제를 해결하기 위해 많이 사용하는 RBF 커널(Radial Basis Function kernel)을 사용하고 가중치 α 는 실험적으로 구한 1.2를 사용하였다. 워터마크는 Fig. 2와 같은 30×30 의 그레이 스케일의 영상을 사용하였으며 본 논문에서는 이를 8비트의 이진 데이터로 변환하여 워터마크로 삽입하였다.

워터마킹의 성능은 PSNR과 BER로 비교하였다. PSNR은 다음의 Eq.(8)을 이용하여 구할 수 있다.^[11]

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right), \quad (8)$$

여기에서 MSE 는 원래의 워터마크와 1차 검출한 워터마크 사이의 평균 제곱 오차이고 MAX 는 영상의 최대 픽셀 값을 의미하며, 본 논문에서는 255를 사용하였다.

새로운 디코딩 모델을 사용한 경우에는 대부분의 경우 기존의 방법을 사용했을 때 보다 PSNR은 약 1 dB 이상 개선되었다. 그리고 기존의 방법에서 BER이 5% 이내인 경우에는 0~2% 이내로 감소하였으며 5% 이상을 보이는 경우에는 3~20%가 개선되는 것을 확인할 수 있다. 2.1절 ~ 2.4절의 방법과 제안하는 방법의 결과는 Tables. 2와 3에서 각각 PSNR과 BER을 비교하여 나타내었다. 2.1절의 방법과 비교하면 BPF, HPF, MP3 공격에서는 BER이 1% 내외로 낮아지는 우수한 성능을 확인할 수 있었다. 또한 N/A 공격에서는 PSNR이 10 dB 이상으로 개선되는 것을 알 수 있었다. Echo, FFT RR 공격에서는 10% 미만의 PSNR과 20% 이상의 BER로 인해 워터마크를 인지하기에는 어려움이 있지만 PSNR과 BER이 많이 개선되는 현상을 볼 수 있었다. 2.2절의 방법과 제안하

Table 1. Attacks for robustness test.

Attack	Description
None	Bypass the audio signal.
LPF	Low pass filtering with cut-off frequency of 15kHz.
BPF	Band pass filtering with cut-off frequency of 100Hz ~ 6kHz.
HPF	High pass filtering with cut-off frequency of 150Hz.
N/A	Add white noise to the audio signal with -10dB signal-to-noise ratio.
Echo	Add an quite echo to the audio signal (Delay: 100ms, Feedback coefficient: 0.5).
Crop.	Cut and replace samples of arbitrary part of the audio signal whose duration is 2seconds.
Resam.	Change the sampling rate of the audio signal from 48kHz to 44.1kHz.
MP3	Compress the audio signal with MPEG-1 Audio Layer 3.
A/C	Compress the amplitude by changing quantization bit from 16bit to 8bit.
Norm.	Normalize the amplitude using the maximum value.
Inv.	Invert all samples in the audio signal, like a 180° phase shift.
FFT I	Invert the real and imaginary part of the sample values in the frequency domain (FFT size: 1024).
FFT R	Reverse only the real part of the audio signal in the frequency domain (FFT size: 1024).



Fig. 2. Watermark.

는 방법의 결과를 보면 Echo 공격에서는 PSNR이 10 dB 이상으로 개선된 워터마크를 확인할 수 있으며 Crop.와 FFT RR 공격에서는 거의 0%로 개선된 BER을 볼 수 있었다. 2.3절의 방법을 사용했을 때보다 BPF와 MP3 공격에서 BER이 개선된 결과를 보였다. 그리고 다른 공격에서도 PSNR과 BER이 크게 개선되었다. 특히 Echo 공격에서는 BER이 약 13% 감소하였으며 A/C 공격에서는 PSNR이 무한대로 개선되

는 우수함을 보였다. 2.4절의 방법과의 비교에서는 새로운 디코딩 방법을 사용한 경우의 대부분이 무한대의 PSNR과 0%의 BER을 가져서 우수한 워터마크의 성능을 보여준다.

제안하는 알고리즘은 다른 기존의 방법들과 비교하여 모든 공격에서 워터마크의 PSNR과 BER이 상당히 개선되는 것을 볼 수 있었다. 10 dB 이상의 워터마크가 기준에 검출된 경우에는 PSNR이 최소 2 dB

Table 2. Comparison of PSNR between method in Sec. 2.1-4 and proposed method.

Attacks		None	LPF	BPF	HPF	N/A	Echo	Crop.	Resam	MP3	A/C	Norm.	Inv.	FFT I	FFT R	
Sec.2.1	Classic	Ori.	24.04	24.04	19.64	18.18	8.98	7.41	22.90	24.04	17.44	21.76	24.04	24.04	24.04	7.31
		Prop.	24.89	24.89	21.35	19.72	10.52	8.21	24.87	24.89	19.82	22.61	24.89	24.89	24.89	8.19
	Pop	Ori.	22.17	22.11	14.28	13.62	10.07	7.84	21.24	22.17	14.39	21.05	22.17	22.17	22.17	7.48
		Prop.	27.17	27.17	15.23	14.41	10.36	8.86	25.72	27.17	15.16	23.74	27.17	27.17	27.17	8.70
	Rock	Ori.	32.51	32.51	5.40	4.31	17.41	7.91	25.06	32.51	15.74	32.51	32.51	32.51	32.51	8.58
		Prop.	Inf	Inf	6.25	5.29	19.11	8.89	26.52	Inf	17.06	Inf	Inf	Inf	Inf	9.39
	Metal	Ori.	29.52	29.52	13.59	7.37	8.05	7.46	26.19	29.59	12.52	25.36	29.52	29.52	29.52	7.44
		Prop.	Inf	Inf	14.94	8.79	8.55	8.62	27.44	Inf	13.79	Inf	Inf	Inf	Inf	8.39
Sec.2.2	Classic	Ori.	Inf	2.28	4.40	9.26	5.26	8.53	27.31	Inf	4.69	11.64	Inf	Inf	Inf	23.26
		Prop.	Inf	3.22	5.29	10.42	6.32	9.54	30.40	Inf	5.59	13.41	Inf	Inf	Inf	30.67
	Pop	Ori.	Inf	3.59	4.38	8.13	5.04	8.72	26.51	Inf	4.63	17.07	Inf	Inf	Inf	22.84
		Prop.	Inf	4.61	5.30	8.85	5.96	9.86	29.89	Inf	5.53	18.05	Inf	Inf	Inf	30.29
	Rock	Ori.	Inf	4.53	4.57	5.98	5.15	6.97	28.89	Inf	4.72	Inf	Inf	Inf	Inf	23.68
		Prop.	Inf	5.58	5.51	6.95	6.07	7.97	30.38	Inf	5.70	Inf	Inf	Inf	Inf	29.51
	Metal	Ori.	Inf	4.02	4.35	5.15	5.12	5.46	28.36	Inf	4.78	14.39	Inf	Inf	Inf	10.55
		Prop.	Inf	4.90	5.28	6.20	5.95	6.56	29.12	Inf	5.72	Inf	Inf	Inf	Inf	12.08
Sec.2.3	Classic	Ori.	Inf	Inf	20.49	4.34	Inf	5.98	Inf	Inf	16.77	30.38	Inf	Inf	Inf	5.65
		Prop.	Inf	Inf	22.82	5.34	Inf	7.00	Inf	Inf	18.85	Inf	Inf	Inf	Inf	6.57
	Pop	Ori.	Inf	35.51	14.27	4.44	Inf	6.14	Inf	Inf	12.70	32.05	Inf	Inf	Inf	5.65
		Prop.	Inf	Inf	15.44	5.38	Inf	7.13	Inf	Inf	13.70	35.53	Inf	Inf	Inf	6.63
	Rock	Ori.	Inf	Inf	5.74	4.03	Inf	5.83	Inf	Inf	12.63	Inf	Inf	Inf	Inf	6.13
		Prop.	Inf	Inf	6.73	5.08	Inf	6.84	Inf	Inf	13.77	Inf	Inf	Inf	Inf	7.12
	Metal	Ori.	Inf	Inf	15.18	4.65	41.28	6.20	Inf	Inf	11.65	33.85	Inf	Inf	Inf	5.74
		Prop.	Inf	Inf	16.69	5.61	Inf	7.32	Inf	Inf	12.77	Inf	Inf	Inf	Inf	6.74
Sec.2.4	Classic	Ori.	Inf	71.65	71.65	Inf	6.97	57.30	22.71	Inf	20.37	8.65	Inf	Inf	Inf	71.65
		Prop.	Inf	Inf	Inf	Inf	7.64	Inf	26.02	Inf	22.03	9.31	Inf	Inf	Inf	Inf
	Pop	Ori.	Inf	Inf	Inf	Inf	10.00	45.61	21.86	Inf	35.53	25.17	Inf	Inf	Inf	35.53
		Prop.	Inf	Inf	Inf	Inf	10.58	Inf	24.05	Inf	Inf	33.09	Inf	Inf	Inf	Inf
	Rock	Ori.	Inf	Inf	Inf	35.53	16.11	46.30	23.13	Inf	24.91	29.84	Inf	Inf	35.53	35.53
		Prop.	Inf	Inf	Inf	Inf	18.31	Inf	Inf	Inf	27.15	Inf	Inf	Inf	Inf	Inf
	Metal	Ori.	35.53	35.53	35.53	35.53	4.90	35.07	21.24	35.53	19.22	7.93	35.53	35.53	35.53	30.76
		Prop.	Inf	Inf	Inf	Inf	5.76	Inf	Inf	Inf	24.02	Inf	Inf	Inf	Inf	Inf

※ Ori.: Original method without SVM decoding model, Prop.: Proposed method, Inf: infinity

Table 3. Comparison of BER between method in Sec. 2.1-4 and proposed method

Attacks		None	LPF	BPF	HPF	N/A	Echo	Crop.	Resam.	MP3	A/C	Norm.	Inv.	FFT I	FFT R	
Sec.2.1	Classic	Ori.	0.56	0.56	2.25	2.83	20.63	32.65	0.82	0.56	4.67	1.01	0.56	0.56	0.56	33.53
		Prop.	0.22	0.22	0.78	1.38	14.14	26.72	0.25	0.22	1.74	0.49	0.22	0.22	0.22	27.36
	Pop	Ori.	1.60	1.63	8.33	9.50	17.89	30.25	1.85	1.60	8.69	1.83	1.60	1.60	1.60	31.94
		Prop.	0.21	0.21	5.44	6.97	15.99	24.25	0.39	0.21	6.39	0.36	0.21	0.21	0.21	25.43
	Rock	Ori.	0.06	0.06	43.96	53.24	4.49	30.53	0.38	0.06	6.15	0.06	0.06	0.06	0.06	27.75
		Prop.	0.00	0.00	35.71	43.15	1.88	24.32	0.22	0.00	3.51	0.00	0.00	0.00	0.00	22.50
	Metal	Ori.	0.49	0.49	10.01	27.49	24.43	33.31	0.78	0.47	12.49	0.88	0.49	0.49	0.49	32.61
		Prop.	0.07	0.07	6.40	19.36	21.57	25.82	0.18	0.07	9.03	0.14	0.07	0.07	0.07	25.83
Sec.2.2	Classic	Ori.	0.00	71.07	52.89	23.47	47.74	27.61	0.35	0.00	50.17	12.96	0.00	0.00	0.00	1.93
		Prop.	0.00	57.47	42.63	18.56	38.15	21.94	0.08	0.00	40.42	8.40	0.00	0.00	0.00	0.07
	Pop	Ori.	0.00	59.90	51.67	30.61	47.42	26.99	0.31	0.00	50.24	4.07	0.00	0.00	0.00	1.36
		Prop.	0.00	47.93	41.69	25.46	38.42	21.10	0.08	0.00	40.72	2.67	0.00	0.00	0.00	0.06
	Rock	Ori.	0.00	51.21	52.19	40.40	46.71	34.85	0.28	0.00	51.03	0.00	0.00	0.00	0.00	1.46
		Prop.	0.00	47.93	41.69	25.46	38.42	21.10	0.08	0.00	40.72	2.67	0.00	0.00	0.00	0.06
	Metal	Ori.	0.00	55.69	52.97	44.08	48.06	44.82	0.25	0.00	50.28	6.93	0.00	0.00	0.00	17.39
		Prop.	0.00	10.50	10.26	9.74	8.64	9.50	0.11	0.00	9.58	1.44	0.00	0.00	0.00	4.75
Sec.2.3	Classic	Ori.	0.00	0.00	1.89	52.83	0.00	42.44	0.00	0.00	5.65	0.35	0.00	0.00	0.00	44.01
		Prop.	0.00	0.00	0.50	42.04	0.00	33.71	0.00	0.00	2.47	0.00	0.00	0.00	0.00	35.33
	Pop	Ori.	0.00	0.03	7.74	52.50	0.00	39.28	0.00	0.00	12.49	0.14	0.00	0.00	0.00	43.49
		Prop.	0.00	0.00	5.13	42.32	0.00	31.74	0.00	0.00	9.94	0.01	0.00	0.00	0.00	35.06
	Rock	Ori.	0.00	0.00	40.56	55.67	0.00	43.01	0.00	0.00	14.38	0.00	0.00	0.00	0.00	40.89
		Prop.	0.00	0.00	32.79	44.65	0.00	34.54	0.00	0.00	10.89	0.00	0.00	0.00	0.00	33.13
	Metal	Ori.	0.00	0.00	6.92	50.13	0.07	39.42	0.00	0.00	15.10	0.14	0.00	0.00	0.00	42.01
		Prop.	0.00	0.00	3.85	40.32	0.00	31.11	0.00	0.00	11.38	0.01	0.00	0.00	0.00	33.79
Sec.2.4	Classic	Ori.	0.00	0.01	0.01	0.00	30.81	0.07	0.82	0.00	1.50	20.63	0.00	0.00	0.00	0.01
		Prop.	0.00	0.00	0.00	0.00	25.65	0.00	0.32	0.00	0.58	16.89	0.00	0.00	0.00	0.00
	Pop	Ori.	0.00	0.00	0.00	0.00	15.36	0.57	0.67	0.00	0.03	1.04	0.00	0.00	0.00	0.01
		Prop.	0.00	0.00	0.00	0.00	12.83	0.00	0.24	0.00	0.00	0.10	0.00	0.00	0.00	0.00
	Rock	Ori.	0.00	0.00	0.00	0.01	5.90	0.35	0.61	0.00	1.00	0.42	0.00	0.00	0.01	0.01
		Prop.	0.00	0.00	0.00	0.00	2.32	0.00	0.00	0.00	0.14	0.00	0.00	0.00	0.00	0.00
	Metal	Ori.	0.01	0.01	0.01	0.01	46.51	0.69	0.94	0.01	1.69	26.22	0.01	0.01	0.01	0.04
		Prop.	0.00	0.00	0.00	0.00	38.26	0.00	0.11	0.00	0.31	21.32	0.00	0.00	0.00	0.03

※ Ori.: Original method without SVM decoding model, Prop.: Proposed method, Inf: infinity

이상 개선되었으며 최대 3.4%의 BER이 개선효과를 가져왔다. 특히 대부분의 BER이 1% 이내의 BER을 보이는 우수한 결과를 보였다. 기존의 방법으로는 PSNR이 10 dB가 안되는 경우가 있었지만 제안하는 방법에서는 PSNR이 10 dB이상으로 개선되는 것을 확인할 수 있는 경우도 발생했다. 제안하는 방법은 워터마크의 행과 열에 대한 SVM 모델을 각각 만들어서 중복으로 사용하기 때문에 하나의 SVM을 사용

하였을 때 보다 워터마크의 PSNR은 최대 10 dB이상, BER은 1% 이내로 개선되는 결과를 확인할 수 있었다.

V. 결 론

본 논문에서는 워터마크 디코딩 모델에 SVM에 기반을 둔 분류 기술을 적용하여 공격에 강인한 새로운 오디오 워터마크 기술을 제안하였다. StriMark,

SDMI, STEP2000 벤치마킹에 사용되는 공격 중 주로 사용되는 14가지를 선택하여 강인성 검증을 한 결과, 기존 방법보다 PSNR과 BER이 모두 개선되는 결과를 확인할 수 있었다. 그리고 기존의 방법에서 10 dB 이상의 PSNR인 경우에는 대부분의 공격에서 1% 이내의 BER을 가지는 우수한 성능을 보여줬다. 또한 기존의 방법에서 10 dB 미만인 경우이지만 제안하는 방법에서는 PSNR이 10 dB 이상이 되는 것을 확인할 수 있었다. 그러나 PSNR이 10 dB 이하인 경우에는 성능이 많이 개선된다 할지라도 워터마크가 존재한다고 판단하기에는 어려운 경우가 많았다. 그리고 가중치에 따라 비례적으로 성능이 더 개선될 것으로 기대하였으나 이와 무관하게 삽입 방법과 공격에 따라 성능이 다르게 나오는 것을 확인할 수 있었다. 향후에는 PSNR이 10 dB 이하인 경우에도 성능을 개선할 수 있는 방법과 다양한 가중치에 대한 실험을 수행하여 이 알고리즘을 개선해 나가야 할 것이다.

References

1. Y. G. Fu, R. Shen, and H. Lu, "Watermarking scheme based on support vector machine for color images," *Electron. Lett.* **40**, 986-987 (2004).
2. C. V. Serdean, M. Tomlinson, J. Wade, and A. M. Ambroze, "Protecting intellectual rights: digital watermarking in the wavelet domain," in *Proc. IEEE Int. Workshop Trends and Recent Achiev. in IT*, 16-18 (2002).
3. H. J. Choi, Y. H. Seo, J. S. Yoo, and D. W. Kim, "Holomarking: digital watermarking method using fresnel hologram" (in Korean), *J. Commun. Inform. Sci. Kr.* **34**, 604-610 (2009).
4. Y. Fu, R. Shen and H. Lu, "Optimal watermark detection based on support vector machines," in *Proc. Int. Symp. Neural Netw.* 552-557 (2004).
5. M. Kutter, F. Jordan, and F. Bossen, "Digital signature of color images using amplitude modulation," *J. Electron. Imaging* **7**, 326-332 (1998).
6. C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. and Knowl. Discov.* **2**, 121-167 (1998).
7. 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).
8. 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. Multimed. Soc. Kr.* **14**, 1431-1437 (2011).
9. Y. J. Seo, S. J. Cho, and U. P. Chong, "Multiple audio watermarking using quantization index modulation on frequency phase and magnitude response" (in Korean), *J. Acoust. Soc. Kr.* **32**, 71-78 (2013).
10. K. P. Shine, and S. Krishna Kumar, "Extended bipolar echo kernel for audio watermarking," in *Proc. ARTCom*, 487-489 (2009).
11. I. Y. An, "Adaptive watermarking method using watermark detection rate" (in Korean), *J. Inst. Electron. Commun. Sci. Kr.* **5**, 465-470 (2010).
12. A. Lang, *StirMark benchmark for audio*, <http://www.witi.cs.uni-magdeburg.de/~alang/smba.php>, 2007.
13. Y. Lin and W. H. Abdulla, "Audio watermarking for copyright protection," University of Auckland, Auckland, New Zealand, Tech. Rep., 2007.
14. 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.

저자 약력

▶ 서 예 진 (Yejin Seo)



2002년 2월: 울산대 전기전자및자동화공학부(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