

# 다중 반향 커널을 이용한 부대역 신호 기반 오디오 워터마킹

\* 오인정, 조재원, 정현열, 정호열  
영남대학교 정보통신공학과

## Audio Watermarking in Sub-band Signals Using Multiple Echo Kernel

\* In-Jung Oh, Jae-Won Cho, Hyun-Yeol Chung, Ho-Youl Jung  
Department of Information and Communication Engineering, Yeungnam University  
{elena99, ram56}@yumail.ac.kr, {hychung, hoyoul}@yu.ac.kr

### 요약

최근 웨이블릿 변환을 이용하여 저주파수 부대역 신호를 반복적으로 분해하고 삽입된 워터마크 비트에 따라 특정 부대역 신호에 두개의 다른 반향 중 하나를 삽입하는 반향 삽입 기법을 제안하였다. 이 기법은 오디오 신호의 주파수 특성을 고려했기 때문에 시간 영역에서의 기존의 반향 삽입 기법에 비해 비가청성 (inaudibility), 워터마크의 강인성(robustness) 및 용적성(capacity)을 모두 증가시킬 수 있다.

기존의 부대역 신호 기반 반향 삽입 기법이 단일 반향 커널만을 이용하였으나, 본 논문에서는 부대역 신호에 다중 반향 커널을 이용한 반향 삽입 기법을 제안한다. 이 기법은 주관적 왜곡(perceptibility)과 SNR(signal to Noise Ratio)과 같은 객관적인 왜곡 면에서, 원본 오디오 신호의 왜곡을 줄일 수 있다.

실험 결과를 통하여 기존의 방법에 비해 제안된 방법이 SNR과 DR(Detection Rate) 모두 좋은 성능을 보여 제안된 방법의 유효성을 입증하였다.

### 1. 서론

최근, 디지털 멀티미디어 데이터의 현저한 발달로 데이터의 복제 및 재전송이 용이해졌다[1]. 이러한 추세로 저작권 보호의 필요성이 증대되고 있다. 암호화와 같은 기존의 데이터 보호 기술들은 데이터를 해독해 내면 데이터의 보호를 보장할 수 없기 때문에 저작권 보호를 위해서는 적절하지 않다[2]. 암호화와는 다르게, 디지털 워터마킹은 원본 데이터에 제한을 가하는 것이 아니라, 숨겨진 데이터를 손상시키지 않고 복원이 가능하도록 보장해 주는 기술이다[3]. 워터마킹은 원본 데이터에 워터마크라고 불리는 정보를 숨기는 저작권 보호 기술이다. 오디오 워터마킹 기법에는 위상 부호화(phase coding) 기법[4], 확산 스펙트럼(spread spectrum) 기법[2],[4], 하위 비트 부호화(low bit coding) 기법[4], 반향 삽입(echo hiding) 기법[4],[5] 등이 있다.

반향 삽입 워터마킹은 원본 오디오 신호에 인간의 청각에 인지되지 않을 정도의 작은 반향을 이용하여 워터마크 데이터를 은닉한다. 기존의 반향 삽입 기법은

원본 오디오 신호를 시간영역에서 프레임 단위로 나누어, 삽입된 워터마크 비트에 따라 각각의 프레임에 반향을 삽입한다[5]. 즉, 워터마크가 삽입된 신호는 삽입된 워터마크 비트에 따라 프레임 간에 두 개의 반향 필터 중 하나를 선택하여 이를 통해 원본 신호를 컨볼루션하여 얻을 수 있다. 워터마크 부호화 과정은 시간영역에서 적용되므로, 모든 주파수 대역에 동일한 반향이 삽입된다. 이로 인해 반향 삽입 과정에서 원본 신호의 주파수 특성이 고려되어야 할 필요가 있다.

이전 연구에서[7], 웨이블릿 변환(Wavelet Transform)을 이용하여 원본 신호를 여러 개의 부대역 신호로 분해하여, 각각의 부대역에 초기 진폭을 달리하여 반향 커널 쌍을 적용한 방법을 제안했다. 이 때, 반향 초기 진폭은 원본 신호의 에너지 분포 정도와 주파수 특성에 따라 각각의 부대역에 맞게 조절되므로, 원본 오디오 신호의 음질 왜곡을 최소화하면서도 높은 에너지의 반향을 삽입할 수 있다. 또한 이 기법은 여러 개의 워터마크 비트들을 동시에 삽입할 수 있으므로, 워터마크의 용적성을 증가시킨다[7].

본 논문에서는 부대역 신호 기반 반향 삽입 기술의 성능을 증가시키는 것을 목적으로 부대역 신호에 다중 반향 커널을 이용한 반향 삽입 기법을 제안한다. 이 기법은 주관적 왜곡(perceptibility)과 SNR(signal to Noise Ratio)과 같은 객관적인 왜곡 측면에서, 원본 오디오 신호의 왜곡을 줄일 수 있다.

본 논문의 구성은 다음과 같다. 먼저 2장에서 시간 영역에 반향을 삽입하는 일반적인 기법을 설명하고, 3장에서 단일 반향 커널 및 다중 반향 커널을 이용한 부대역 신호 기반 반향 삽입 기법을 소개한다. 4장에서는 단일 반향 커널을 이용한 이전의 기법과 제안된 기법의 실험 결과를 보인다. 이 실험은 블라인드 기법을 사용하였다. 마지막으로 제 5장에서 결론을 내린다.

### 2. 시간 영역 기반 반향 워터마킹

일반적으로, 반향 삽입은 오디오 신호와 반향 커널의 컨볼루션으로 표현이 가능하며, 이진 데이터의 삽입을 위해 반향 커널 쌍이 사용된다. 그림 1은 반향 삽입 과정을 나타낸다. 원본 오디오 신호  $x[n]$ 은 적절한 길이를 갖는 프레임 단위의 신호  $x_m[m]$ 으로 분할된다.

여기에서  $m=1,2,\dots,M$  이고,  $M$  은 프레임의 총 개수이다. 각각의 프레임 신호는 이진 워터마크 비트 열을 참조하여  $e_i[n]$  과 컨볼루션 함으로써 반향이 삽입된 오디오 신호를 얻을 수 있다. 여기에서  $i \in \{0,1\}$  이다. 식(1)은 임의의 프레임 신호에 반향이 삽입되는 과정을 나타낸다.

$$\hat{x}_m[n] = x_m[n] * e_i[n] \quad (1)$$

여기에서  $\hat{x}_m[n]$  은 워터마크가 삽입된  $m$  번째 프레임 신호를 의미하고 \* 은 컨볼루션을 나타낸다. 양의 단일 반향 커널(positive single echo kernel), 양의 다중 반향 커널(positive multiple echo kernel), 음의 단일 반향 커널(negative single echo kernel), 음의 다중 반향 커널(negative multiple echo kernel), 그리고 의사 랜덤 잡음을 이용한 반향 커널 등 지금까지 다양한 종류의 반향 커널의 설계에 대한 연구가 이루어졌다[3][4][5][6]. 일반적으로 반향 커널의 오프셋(offset)은 0.9msec에서 3.4msec 사이에 존재할 때 인간의 청각에 가장 인지되지 않는 것으로 알려져 있다[7]. 본 논문에서는 제안한 방법의 유효성을 판단하기 위해 양의 단일 반향 커널을 사용하였으며, 본 논문에서 사용된 반향 커널의 전달함수(transfer function)는 식(2)와 같이 표현할 수 있다.

$$E_i(z) = 1 + \alpha \cdot z^{d_i}, \quad \text{for } i \in \{0,1\} \quad (2)$$

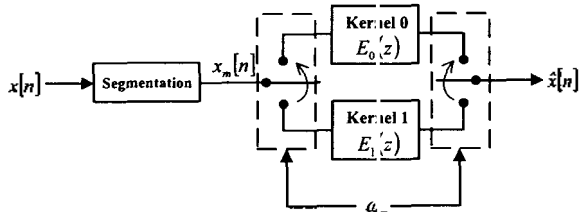


그림 1. 시간영역에서의 일반적인 반향 삽입 과정

여기에서  $\alpha$  와  $d_i$  는 반향의 초기진폭과 오프셋을 각각 나타낸다. 워터마크가 삽입된 오디오 신호  $\hat{x}[n]$  은

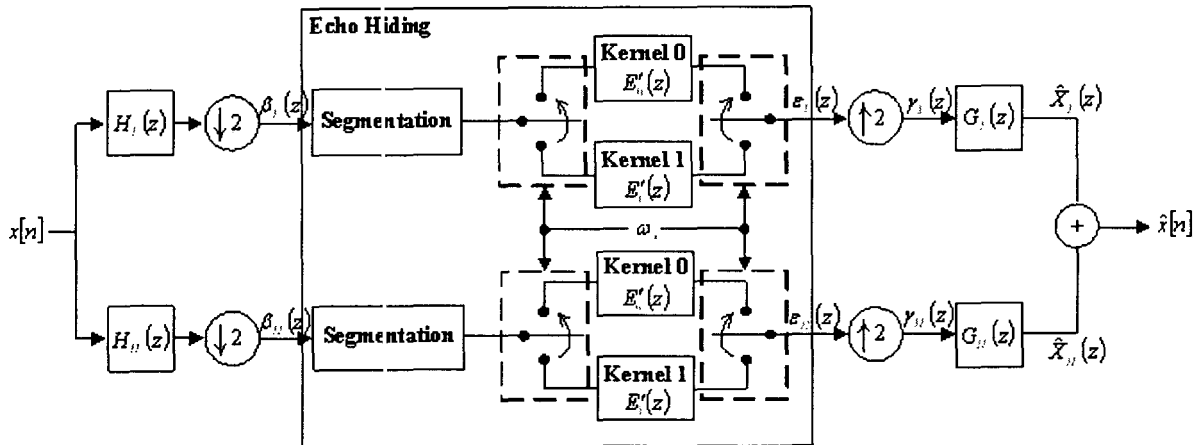


그림 2. 부대역 신호 기반 워터마크 삽입 과정(2채널)

반향이 삽입된 프레임 신호를 조합해서 얻을 수 있다. 이 때, 프레임과 프레임 사이의 경계 전후에 중첩구간을 두고 헤닝 또는 헤밍 창 함수(Hanning or Hamming Window Function)와 같이 양 끝이 완만한 필터를 거침으로써 반향이 삽입된 프레임들 간 경계에서의 급격한 변화로 인한 음질 왜곡을 방지할 수 있다. 그러나 일반적인 반향 삽입 방법은 시간 영역에서 이루어지기 때문에 근본적으로 주파수 특성을 고려하여 삽입하는 것이 어렵다.

삽입된 워터마크의 검출은 워터마크가 삽입된 오디오 신호의 각각에 프레임에서 반향 커널의 오프셋을 찾는 것을 의미한다. 이 경우 반향 커널의 오프셋 위치에서 피크를 나타내는 오토캡스트럼(auto-cepstrum)을 사용하는 것이 합리적이다[4]. 오토캡스트럼은 식(3)과 같이 표현된다.

$$F^{-1}(\ln(F(\hat{x}_m[n])))^2 \quad (3)$$

여기에서  $F$  와  $F^{-1}$  는 푸리에 변환(Fourier Transform, FT)과 그 역변환을 나타낸다. 삽입된 워터마크는 프레임마다 오토캡스트럼의 피크의 위치를 찾음으로써 구할 수 있다.

### 3. 부대역 신호 기반 반향 워터마킹

이 장에서는 부대역 신호에 반향을 삽입하는 방법을 설명한다. 원본 신호는 웨이블릿 변환(Wavelet Transform)을 이용하여 여러 개의 부대역 신호로 분해되고, 반향은 각각의 부대역 신호에 삽입된다. 간단하게 그림 2와 같은 2채널 반향 삽입 시스템을 고려해보자. 원본 신호  $x[n]$  은 분해 필터 군  $H_L(z)$  와  $H_H(z)$  에 의해 저주파수와 고주파수로 분해된다. 각 부대역 신호들은 더 작은 프레임으로 분해되고, 반향 커널 쌍  $E'_0(z)$ ,  $E'_1(z)$  은 부대역 신호의 각 프레임에 적용된다. 반향이 삽입된 부대역 신호는 복원 필터 군  $G_L(z)$  과  $G_H(z)$  을 통해 반향이 삽입된 오디오 신호로 변환된다. 만약 동일한 워터마크 비트 열이 모든 대역 신호에 적용된다면, 이 시스템은 그림 1과 같은 일반적인 반향 삽입과 동일한 결과를 얻을 수 있다.

이전 연구에서[7], 시간 영역에서의 반향 삽입 방법이 부대역 신호와 적절한 오프셋을 갖는 반향 커널과 컨볼루션으로 표현될 수 있음을 보였다. 즉, 제안된 반향 삽입 기법이 일반적인 반향 삽입 기법의 하나인 것이다. 또한 각 대역에 삽입된 반향의 진폭은 주파수 특성을 고려하여 적절하게 조정될 수 있다. 그러므로 반향 삽입 기법은 원본 오디오 신호의 음질 왜곡을 최소화하면서 높은 에너지의 반향 삽입이 가능하다.

이전 연구에서는[7] 단일 반향 커널만을 이용한 반향 삽입 기법을 보였으나, 본 논문에서는 다중 반향 커널을 이용한 반향 삽입 기법을 제안한다. 그림 3에서 양의 단일 반향 커널과 양의 다중 반향 커널을 보였다.

그림 4는 워터마크 검출 과정을 나타낸다. 반향이 삽입된 오디오 신호는 삽입할 때와 동일한 분해 필터 군을 통해 부대역 신호로 분해된다. 워터마크는 오토캡스트럼을 이용함으로써 각각의 부대역 신호로부터 검출된다. 워터마크 비트 열이 각각의 부대역 신호로부터 독립적으로 검출되기 때문에, 반향 삽입 과정에서 또 다른 비트 열들이 각각의 부대역 신호에 삽입될 수 있다. 이는 제안된 기법이 워터마크의 용적성을 증가시킬 수 있음을 의미한다. 만약 다른 워터마크 비트 열들이 같은 반향 커널 쌍을 이용하여 각각의 부대역에 삽입된다면, 삽입된 반향 신호들 간의 간섭이 발생할 수 있다. 이는 워터마크 검출율의 감소를 야기할 수 있으므로, 다른 반향 커널 쌍을 각각의 부대역 신호에 적용하여 이를 피할 수 있다.

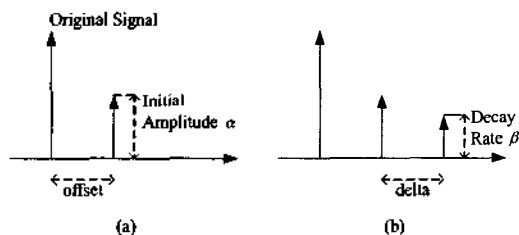


그림 3. (a) 단일 반향 커널 (b) 다중 반향 커널

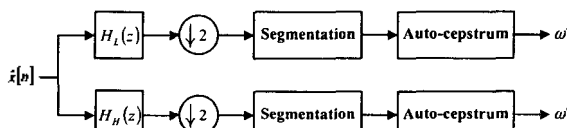


그림 4. 부대역 신호기반 워터마크 검출 과정(2채널)

#### 4. 실험결과

본 논문에서 제안된 반향 삽입 방법의 유효성을 판별하기 위해서 원본 오디오 데이터로 44.1Khz의 샘플링율(sampling rate)을 가지는 16비트 모노(mono)의 클래식(classic) 데이터를 사용하였다. 또한 반향이 삽입된 오디오 신호의 음질 측정을 위해 SNR(Signal to Noise Ratio)을 사용하였다.

$$SNR = 10 \log_{10} \left( \frac{\sum_{n=0}^L x[n]^2}{\sum_{n=0}^L (x[n] - \hat{x}[n])^2} \right) \quad (4)$$

여기에서  $L$ 은 오디오 신호의 전체 샘플 길이를 나타낸다. 본 논문에서 워터마크의 검출은 원본 오디오 신호 없이 삽입된 워터마크를 검출할 수 있는 블라인드(blind) 기법을 사용하였고, 강인성 평가를 위해 DR(Detection Rate)을 사용하였다.

$$DR = \frac{\# \text{ of watermark bits correctly extracted}}{\# \text{ of watermark bits placed}} \quad (5)$$

또한 부대역별 신호로 분해하기 위해서 5/3-탭 쌍 직교 완전 복원 필터군(5/3-tap bi-orthogonal perfect reconstruction filter bank)이 저주파수 대역 신호에 반복적으로 적용되었다. 원본 오디오 신호는 많은 실험을 통해 5개의 부대역 신호로 분해될 때, 즉 웨이블릿 분해 레벨(wavelet decomposition level)이 4일 때 워터마크의 강인성과 비가청성을 동시에 충족시킬 수 있었다. 웨이블릿 필터군은 고속화를 위해 리프팅(lifting) 기법을 이용하였다. 표 1은 각각의 대역에 따른 주파수 범위를 나타낸다[12]. 또한 SDMI(the Secure Digital Music Initiative)에서 권고한 워터마크 비트율이 약 20bits/sec임[8]을 감안하여 초당 워터마크를 22 비트로 삽입하기 위해 한 프레임의 길이를 46.44msec(2048samples)로 설정하였다. 반향의 오프셋 및 감쇄율(decay rate)은 실험을 통해 0.9-3.4msec 범위 내에 두었다. 웨이블릿 영역에서 시간 영역에서와 동일한 반향 효과를 얻기 위해서는 프레임의 길이와 반향의 오프셋, 감쇄율은 웨이블릿 분해 레벨에 의존하여 2의 지수 승으로 감소시켜야 한다. 삽입된 워터마크의 강인성을 평가하기 위해서 잡음 첨가(add noise), 진폭변경(amplify, 0.5), 반향삽입(초기진폭 0.5/오프셋 100msec) 및 MP3 압축(64Kbps/128Kbps) 공격 실험을 수행하였다.

실험 시 부대역 신호에 반향을 삽입할 때 사용된 단일 반향 커널과 다중 반향 커널 사이의 성능을 비교하였다. 표 2와 표 3은 단일 반향 커널과 다중 반향 커널에 유사한 SNR을 얻기 위하여 선택된 초기 진폭 값을 나타낸다. 표 4는 부대역 신호에 단일 반향 커널과 다중 반향 커널을 사용하였을 때의 평가 결과이다.

표 1. 부대역에 따른 주파수 범위(KHz)

1st-band	2nd-band	3rd-band	4th-band	5th-band
0.0~1.3	~2.7	~5.5	~11.0	~22.0

표 2. 각 부대역에 따른 단일 반향의 진폭

1st-band	2nd-band	3rd band	4th band	5th band
0.33	0.45	0.45	0.70	0.70

표 3. 각 부대역에 따른 다중 반향의 진폭

	1st-band	2nd-band	3rd band	4th band	5th band
$\alpha$	0.25	0.42	0.52	0.60	0.60
$\beta$	0.13	0.21	0.26	0.30	0.30

## 5. 결 론

본 논문에서는 부대역 신호에 양의 다중 반향을 삽입한 반향 워터마킹 기법을 제안하였다. 실험을 통해, 제안된 기법이 원본 오디오 신호의 음질 왜곡을 최소화하면서도 높은 에너지를 갖는 반향을 삽입할 수 있음을 보였다. 또한 제안된 기법이 여러 개의 워터마크 비트 열을 동시에 삽입할 수 있으므로 워터마크의 용적성을 증가시킬 수 있다. 실험 결과 제안된 기법이 단일 반향 삽입 기법 보다 우수한 성능을 보였다.

## 6. 참고문헌

[1]Paraskevi Bassia, Ioannis Pitas, Nikos Nikolaidis, "Robust Audio Watermarking in the Time Domain", IEEE Transactions on Multimedia, Vol.3, No.2, 232-241, 2001  
 [2]Darko Kirovski, Henrique Malvar, "Robust Spread-Spectrum Audio Watermarking", Proceedings of IEEE ICASSP 01, Vol.3, 1345-1348, 2001  
 [3]Say Wei FOO, Theng Hee YEO, Dong Yan HUANG, "An Adaptive Audio Watermarking System", Proceedings of IEEE Region 10, Vol.2, 509-513, 2001  
 [4]W.Bender, D.Gruhl, N.Morimoto, A.Lu, "Techniques for data hiding", IBM Systems Journal, Vol.35, Nos 3&4, 313-336, 1996  
 [5]D.Gruhl, W.Bender, "Echo Hiding", Proceedings of Information Hiding Workshop, 295-315, 1996  
 [6]Hyen-O Oh, Jong-Won Seok, Jin-Woo Hong, Dae-Hee Youn, "New Echo Embedding Technique for Robust and Imperceptible Audio Watermarking", Proceedings of IEEE ICASSP01, Vol.3, 1341-1344, 2001

[7]Jae-Won Cho, Ha-Joong Park, Young Huh, Hyun-Yeol Chung, and Ho-Youl Jung, "Echo Watermarking in Sub-band Domain", Proceeding of IWDW 2003, 447-455, 2003  
 [8]Hyen-O Oh, Hyun-Wook Kim, Dae-Hee Youn, Jong-Won Seok, and Jin-Woo Hong, "New Echo Embedding Technique for Robust Audio Watermarking", The Journal of the Acoustical Society of Korea, Vol.20, No.2, 66-76, 2003  
 [9]Byeong-Seob KO, Ryouichi NISHIMURA, Yoiti SUZUKI, "Time-spread Echo Method for Digital Audio Watermarking using PN sequences", Proceedings of IEEE ICASSP 02, 1317, 2002  
 [10]C.Xu, "Applications of Digital Watermarking Technology in Audio Signals", J.Audio Eng. Soc., Vol.47, No.10, 1999  
 [11]Gilbert Strang, Truong Nguyen, "Wavelets and Filter Banks", Wellesley Cambridge Press, 1997  
 [12]Sweldens W, "The lifting Scheme : A Custom-Design Construction of Biorthogonal Wavelets", Applied and Computational Harmonic Analysis, Vol.3, No.2, 186-200, 1996  
 [13]Xin Li, Hong Heather Yu, "Transparent and Robust Audio Data Hiding in Sub-band Domain", Proceedings of IEEE Coding and Computing, 7479, 2000

표 4. 단일 반향 삽입 방법과 제안한 방법의 SNR 및 DR 비교

Attack	Sub-band Domain										
	1st-band		2nd-band		3rd-band		4th-band		5th-band		
	single	multiple	single	multiple	single	multiple	single	multiple	single	multiple	
	echo	echo	echo	echo	echo	echo	echo	echo	echo	echo	
	SNR	SNR	SNR	SNR	SNR	SNR	SNR	SNR	SNR	SNR	
No Attack	9.90	9.12	23.12	23.51	29.69	28.24	34.41	36.01	45.65	47.33	
	0.82	0.82	0.88	0.89	0.93	0.93	0.98	0.99	1.00	1.00	
Add Noise	9.87	8.67	22.72	17.89	28.14	18.76	30.85	19.19	33.13	19.27	
	0.82	0.96	0.87	0.97	0.90	0.99	0.93	1.00	0.80	1.00	
Amplify	5.03	5.11	5.99	6.00	6.02	6.01	6.02	6.02	6.02	6.02	
	0.82	0.95	0.88	0.97	0.93	0.98	0.98	1.00	1.00	1.00	
Echo	3.34	2.27	4.63	2.83	4.71	2.85	4.72	2.86	4.73	2.86	
	0.74	0.93	0.87	0.97	0.86	0.98	0.93	1.00	0.95	1.00	
MP3	64Kbps	9.89	9.06	22.86	23.00	28.79	27.40	31.94	36.67	35.67	47.89
		0.82	0.93	0.87	0.87	0.90	0.94	0.96	0.96	0.87	0.96
	128Kbps	9.90	9.06	23.13	22.99	29.70	27.28	34.37	35.17	44.81	47.16
		0.82	0.87	0.88	0.87	0.92	0.95	0.98	0.99	0.99	0.99
Average	7.96	7.22	16.02	16.04	19.57	18.42	21.62	22.65	25.71	28.42	
	0.81	0.90	0.86	0.92	0.90	0.96	0.94	0.99	0.89	0.99	