

피치 스케일링을 이용한 강인한 오디오 워터마킹 알고리즘

신승원*, 김완진*, 김종원*, 최종욱*

*㈜ 마크애니 부설연구소

e-mail : {swshin, aistero,jwkim,juchoi}@markany.com

A Robust Audio Watermarking Algorithm Using Pitch Scaling

Seungwon Shin*, Oanjin Kim*, Jongweon Kim*, and Jonguk Choi*

*MarkAny Inc., MarkAny Research Institute

요 약

본 논문은 손실 압축(lossy compression)에 강인한 오디오 워터마킹(audio watermarking) 기술을 제안한다. 제안된 워터마킹 기술은 온라인으로 유통되고 있는 MP3, AAC, WMA, RA, ATRAC 등과 같은 대부분의 오디오 압축 포맷으로 변형되더라도 삽입된 정보를 검출할 수 있다. 실험 결과 BER 이 9% 이내로 오류 정정 코드(error correction code)를 이용해서 삽입 정보를 100% 검출 가능하였다.

1. 서론

인터넷을 통한 디지털 저작물 시장을 위해서는 저작물의 불법 복제, 전송, 배포 등을 방지하고 모니터링할 수 있는 기술이 절대적으로 요구된다. 최근 이러한 행위를 금지하는 법안을 각국에서 제정하고 있으나, 실제 행위를 적발하고 동기 유발을 억제하기 위해서는 기술적인 뒷받침이 필요하다. 이에 적합한 기술로는 상호 보완적인 관계를 갖고 있는 암호화 기술과 워터마킹 기술이 있다. 암호화 기술은 원천적으로 불법 복제된 저작물은 재생이 불가능하도록 하기 때문에 불법 복제 행위를 막거나 억제하는 역할을 하며, 워터마킹 기술은 디지털에서 아날로그로 변환하는 단계에서 이미지 캡처링, 오디오 녹음 등을 통해서 불법 복사된 저작물의 유통 경로를 추적하고 불법 저작물을 식별하는 역할을 한다[1].

본 논문은 디지털 저작물 중에서 오디오 저작물의 저작권 정보를 삽입하는 워터마킹 기술을 제안한다. 이 기술은 A/D 변환과 인터넷에서 활용되고 있는 MP3, AAC, WMA, RealAudio, ATRAC 등 다양한 손실압축 기술에 매우 강인하다.

본 논문의 구성은 2 장에서는 현재 디지털 오디오 워터마킹 기술의 개발 동향과 선행 연구에 대해서 살펴보고, 3 장에서는 본 논문에서 제안하고자 하는 워터마킹 방법을 설명하고, 4 장에서는 제안한 기술에 대

한 실험 결과를 제시한다. 마지막으로 5 장에서는 간략한 연구의 성과를 살펴보고 향후 개선 방향을 제시한다.

2. 선행 연구 분석

지금까지 국내의 논문을 통해서 알려진 오디오 워터마킹 기술을 주요 응용 기술로 구분하면 Spread Spectrum 을 이용한 방법[2-4], 반향 삽입법(echo hiding)[5-7], 디지털 필터를 이용한 방법[8], 복제 신호(replica Signal)를 이용한 워터마킹 방법[9] 등 매우 다양한 연구가 진행되어 왔다.

Spread Spectrum 을 이용한 연구[2-4]는 워터마킹 분야에서 가장 먼저 활용된 응용 기술로 PN-sequence 를 심리음향 모델 또는 선행 예측 모델로 마스킹하여 삽입함으로써 견고성과 비가청성을 만족시키고, 추출은 상호 상관함수를 이용한다. Spread Spectrum 은 손실 압축에는 강한 것으로 알려져 있으나 압축율을 높일 경우 에러가 빈번이 발생하며 비가청성을 위한 연산 시간이 길다는 문제점이 있다. 또한, 디지털 필터링 등에 의해서 워터마크 신호가 제거될 위험이 매우 큰 단점이 있다.

반향 삽입법(echo hiding)[5-7]은 매우 짧은 시간 간격을 두고 발생하는 반향(echo)이 사람의 귀에 둔감하다는 점에 착안한 워터마킹 기술이다. 오디오 신호에 대한 이해와 분석이 매우 철실하게 요구되는 기술이

다. 압축에 대한 검출율이 높고, 구현이 쉬우나, 캡스 트럼 분석을 통해서 나타나는 반향의 위치를 이용해서 삽입 정보를 검출하기 때문에 해킹 위험이 높다.

디지털 필터링[8]을 이용한 방법은 오디오 신호의 특정 주파수 신호를 주파수 차단 필터로 제거하고, 특정 주파수의 성분이 감쇄된 위치를 확인하여 은닉된 정보를 찾는다. 특히, 이 방법은 linear speed change, time scaling 에 영향을 받지 않으나 많은 정보를 삽입하기가 어려운 단점이 있다. 또한, 반향 삽입법과 마찬가지로 주파수 분석을 통해서 워터마크 정보를 판별해 낼 수 있기 때문에 해킹에 대한 우려가 있다.

마지막으로 복제 신호 (Replica signal) 변조를 이용한 워터마킹 기술[9]은 오디오 신호와 모든 면에서 동일한 특성을 갖고 있으나, 주파수의 위치 또는 위상 지연 등의 일부 특징만을 변경한 복제 신호를 워터마크로 이용한다. 이 기술은 아직 구체적인 실험 결과와 분석 결과가 제시되고 있지 않아 그 유용성을 단정짓기 어려우나, 원 오디오 신호로부터 워터마크를 생성한다는 점에서 손실 압축에 견고하며 동시에 비가청성이 확보될 것으로 판단된다. 또한, 복제 신호를 생성하는 매개 변수가 매우 다양하기 때문에 해킹에 대한 안전성 또한 높다고 판단된다.

요컨대, Spread Spectrum 은 비가청성을 만족시키기 위한 연산과정이 복잡하고 많은 정보를 삽입하기가 어렵기 때문에 실제 시스템에 적용하기 어렵고, 반향 은닉법과 디지털 필터링을 이용한 워터마킹 기술은 은닉 정보의 추출과정에 있어서 안전성이 결여되어 있기 때문에 디지털 저작물의 저작권 보호 시스템에 적용하기 위해서는 문제점이 있다.

본 논문에서는 손실 압축과 A/D 변환에 대한 강인성 (robustness), 원음과 정보를 은닉한 신호와 구별되지 않는 비가청성 (transparency), 해킹에 대한 안전성 (security), 그리고 실시간 처리 등을 만족하는 워터마킹 기술을 제안하고자 한다. 우리는 앞서 언급된 기술들에 대한 논문 조사와 실증적 구현을 통해서 원하는 성능을 만족시키는 워터마킹 기술을 제안한다.

3. 제안된 기술

3.1 워터마크 생성 절차

오디오 저작물에 정보를 은닉시키기 위해서는 정보를 다른 형태로 표현을 해야 한다. 예를 들어 Spread spectrum 방식에서는 은닉 정보를 PN 열로 표현한다. 은닉하고자 하는 정보가 "100110" 이라면, "0"을 나타낼 때는 PN-0 열을 이용하고, "1"을 나타낼 때는 PN-1 열을 이용한다. 이런 방식으로 오디오 신호를 PN 열의 길이마다 1 비트씩 정보를 차례로 삽입을 한다. 이때, PN-0 과 PN-1 열을 정보를 운반하는 신호라 해서 운반신호(carrier)라고 부른다. 즉, 워터마크는 은닉 정보를 운반하기 위해서 어떤 신호를 이용하느냐에 따라 그 성능과 구현의 복잡성이 결정된다.

Petrovic(2001)은 운반신호를 오디오 신호로부터 생성할 수 있는 몇가지 방법을 [11]에서 제안하였다. 이 논문에서 그는 오디오 신호 성분중에서 시간/주파수

영역의 일부 신호 성분을 운반신호로 이용하고 있으며, 운반신호를 주파수 이동, 위상 지연, 직교 신호의 곱을 통해서 변조시켜 워터마크 신호를 생성하였다. Petrovic 의 방법으로 생성된 워터마크 신호는 오리지널 오디오 신호와 모든 신호적 특성이 유사하기 때문에 워터마크 신호를 원본 오디오 신호에 더하다라도 음질에 큰 영향을 미치지 않는다. 또한, 삽입된 워터마크 신호와 오디오 신호의 특성이 동일하기 때문에 필터링, 오디오 압축 등에 의한 열화가 상대적으로 적은 장점이 있다.

본 논문에서는 이러한 선행 연구를 토대로 운반신호는 오디오 신호의 일부 신호를 이용하는 Petrovic 이 제안한 워터마크 생성방법을 응용하였다.

우리는 실험을 통해서 주파수 이동(frequency shift)을 통해서 생성된 워터마크를 이용한 워터마킹 삽입/추출 방법이 높은 검출율을 유지할 수 있음을 알 수 있었다. 그러나, 그의 논문에서 제안한 주파수 이동으로 오디오 신호를 변조시켜 워터마크로 이용할 경우에 몇몇 음악에서 음색의 변화와 음의 변화가 관찰되었다. 이는 주파수 이동이 오디오의 주기적인 특성파음의 하모닉을 훼손시키기 때문인 것으로 예상된다.

따라서, 우리는 단순히 모든 주파수를 일정량만큼 이동시키기 보다는 Pitch scaling 을 이용하여 워터마크를 생성하는 것이 견고성과 비가청성을 만족시키는 데에 더욱 효과적이다.

첫번째 단계로, 입력 신호 $s(t)$ 로부터 운반신호를 생성해 내어야 한다. 가장 쉽게는 $s(t)$ 자체를 이용할 수도 있으나, 입력 신호가 외부 변형이 가해질 여지가 많으므로 필터링을 통해서 일부 주파수 성분만을 운반신호로 이용하는 것이 좋다. 일반적인 변형으로는 고주파 성분은 쉽게 제거될 수 있으므로 저주파나 중간 주파수 대역을 이용할 수 있는 Low-pass 또는 band-pass 필터를 이용한다.

오디오 신호 $s(t)$ 를 대역통과필터 (bandpass filter)를 통과시켜 운반신호 $c(t)$ 를 생성한다.

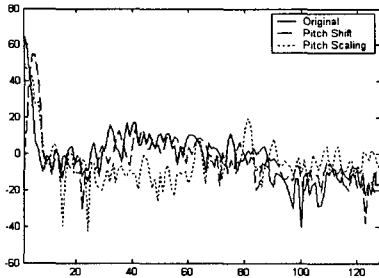
$$c(t) = \sum_{\tau} s(t)h(t-\tau) \quad (1)$$

여기서, $h(t)$ 는 입력 신호로부터 얻고자 하는 성분을 얻기 위한 필터 계수이다. 이때, 필터링을 통해서 얻어진 신호 $p(t)$ 는 입력 신호 $s(t)$ 로부터 상속받은 신호 특성을 지니게 된다.

두번째 단계로, 운반신호 $c(t)$ 가 $s(t)$ 와 다른 신호적 특성을 갖도록 변조를 한다. Petrovic 는 주파수 이동을 이용한 반면 우리는 여기서 주파수 이동 대신에 pitch scaling 를 통해서 운반신호 $c(t)$ 가 원 오디오 신호 $s(t)$ 와 다른 특성을 갖도록 한다.

Pitch scaling 은 오디오 샘플의 길이는 그대로 유지하면서 오디오의 하모닉 관계를 확장하거나 축소하고, pitch shift 는 오디오의 스펙트럼을 이동시킨다.

<그림 1>은 원본 오디오 신호와 pitch shift, pitch scaling 한 신호의 주파수 파형의 변화를 도식화 한 것이다.



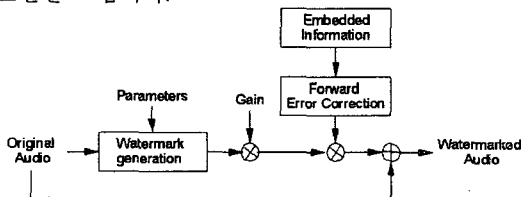
<그림 1> Pitch scaling 과 Pitch shit 의 주파수 변화

Pitch scaling 은 STFT(short time Fourier transform)을 이용해서 주파수 영역에서 쉽게 구현이 가능하다. 이에 대한 자세한 기술은 본 논문의 범위를 벗어나므로 참고문헌[10-11]을 통해서 자세한 내용을 찾을 수 있다.

3.2 워터마크 삽입 절차

우리는 3.1 절에 설명한 워터마크 생성절차에 따라 반복적인 실험을 거쳐서 일반 오디오에 적합한 워터마크를 생성하였다. $p_i(t)$ 는 1.5kHz~5.0kHz 의 대역통과필터를 이용하여 필터링 하였으며, Pitch scaling 은 "0.7"을 이용하였다. 또한, 오디오 신호를 일정한 크기의 프레임 사이즈로 나누어서 정보를 은닉하는데 프레임 사이즈는 오디오 샘플링 주파수가 44.1kHz 인 경우에 1764 개를 이용하며, 각 프레임에 1 비트 정보를 삽입하면 정보 삽입률은 약 25bps - 35bps 가 된다.

<그림 2>는 정보를 실제로 은닉하는 과정을 간략히 표현한 그림이다.



<그림 2> 정보 은닉 절차도

워터마크된 신호 $\tilde{s}(t)$ 는 $w(t)$ 에 삽입하고자 하는 워터마크 정보 m 을 곱하여 더해준다. 워터마크 정보 m 은 '1' 또는 '-1'의 값을 갖는 상수가 될 수도 있으며, 시간축에 따라 변하는 신호가 될 수도 있다. 만약, '1' 또는 '-1'을 삽입하는 경우에는 식(3)과 같이 된다. '1'은 비트 정보 '1'을, '-1'은 비트정보 '0'을 의미한다.

$$\tilde{s}(t) = s(t) + \sum_k gm_k w(t) \quad (3)$$

여기서, g 삽입강도를 의미한다. 운반신호 $w(t)$ 는 $s(t)$ 의 크기에 비례적이므로, 삽입강도는 모든 프레임에 동일하게 적용되는 상수로 설정한다. 우리는 실험적으로 g 를 0.2 로 설정하였다.

은닉 정보의 구성은 외부에 의한 공격에 대한 사전적인 대비와 오류가 발생한 경우에 오류를 수정하기

위한 정보로 구성된다. 외부의 공격에 대해서는 대표적으로 A/D, D/A 변환, Time Scaling 에 대해서 고려할 필요가 있다. A/D, D/A 변환 과정에서 일부 샘플 데이터의 누수가 발생하기 때문에 정보를 "1"을 삽입한 블록에서 "-1"이 검출되는 경우가 있다. 이러한 문제를 해결하기 위해서 우리는 지금까지 고정된 크기의 프레임 사이즈 단위로 반복해서 삽입하는 방법을 버리고, 워터마크를 삽입하는 위치를 오디오로부터 추출하여 삽입하는 방법을 채택하였다.

A/D, D/A 변환 과정에서 오디오 데이터가 일부 유실되는 문제와 time scaling 등에 의해서 워터마크를 삽입한 위치를 대략적으로 유추하기 위해서는 오디오의 어떤 특징 위치를 기준으로 워터마크를 삽입하는 것이 좋다.

워터마크의 삽입 위치를 결정하는 방법은 다음과 같다. 먼저, 오디오 신호를 저주파 대역 통과 필터(lowpass filter)를 이용해서 100Hz 이하의 신호를 구한 다음에, 국소 극대점과 국소점을 찾는다. 첫번째 극대점 또는 국소점을 워터마크 삽입의 시작위치로 삼는다. 그 다음 워터마크 삽입 시작 위치는 선행 프레임의 마지막 지점에서 가장 근접한 국소 극대점(local maxima) 또는 국소점(local minima)을 워터마크 삽입 시작 위치로 결정한다. 이때, 워터마크의 프레임은 고정되어 있으나, 워터마크가 삽입되는 위치를 추출 과정에서도 동일하게 찾아낼 수 있기 때문에 A/D 변환 등과 같은 소량의 데이터 손실에 의한 오검지는 방지할 수 있다. 그러나, 이 방법으로는 10%이내로 변화를 가하는 time scaling 에 대처하기는 어렵다.

워터마킹 기술로 정보를 은닉한 이후에 오디오 컨테츠에는 많은 공격이 가해지게 된다. 이런 공격을 받게 되면 워터마크 신호가 일부 손실이 발생하는 것이 일반적이기 때문에 반드시 오류에 대한 대응이 필요하다. 최근 워터마킹 기술을 적용할 때 오류 정정 코드(error correction code)를 이용하는 경우가 많은데, 1-4 비트 정도의 적은 데이터를 삽입하던 이전과 달리 최근에는 64 비트 이상의 많은 정보의 삽입을 시도하고 있기 때문이다.

오류에 대한 대응책으로 본 논문에서는 터보 코드(turbo code)를 이용하였으며, 10%-14%이내의 오류가 발생할 경우에 오류를 정정할 수 있도록 삽입 정보를 변환하였다. 터보 코드를 이용해서 정보를 변환하게 되면 64 비트 정보가 198 비트 정보로 인코딩된다.

3.3 워터마크 추출 절차

은닉된 정보의 추출은 삽입과정과 유사하다. 삽입 과정에서 이용한 운반 신호를 워터마크가 삽입된 결과 신호 $\tilde{s}(t)$ 로부터 삽입 절차에서 이용한 방법과 동일한 방법으로 구한다. 이 신호 $w'(t)$ 의 성분은 식(5)와 같이 $w(t)$ 와 잡음 신호 $e(t)$ 로 구성된다. 이때, 워터마크를 검출하기 위해서 운반신호 $w'(t)$ 와 정보가 은닉된 신호 $\tilde{s}(t)$ 곱하면 식(6)과 같이 표현되나 $w'(t)$ 의 신호와 $w(t)$ 가 거의 유사하므로 식(7)과 같

이 축약이 가능하다.

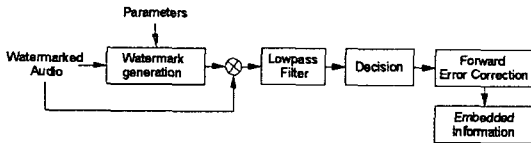
삽입 정보의 추출은 식 (7)에서 워터마크 $w(t)$ 와 $\tilde{s}(t)$ 의 곱에 의해서 강한 DC 성분과 저주파 신호를 얻을 수 있는데, 이 신호의 부호는 삽입 정보의 부호와 일치하게 된다. 따라서, 식 (7)로부터 얻어진 신호를 저역통과필터를 통과시켜 얻어낸 신호의 합이 양이면 "1", 음이면 "-1"의 정보가 삽입된 것으로 판단하면 은닉된 정보를 복원할 수 있다.

$$w'(t) = w(t) + ge(t) \quad (5)$$

$$\tilde{s}(t) \cdot w'(t) = \left(s(t) + gm \sum_k w(t) \right) \cdot (w'(t)) \quad (6)$$

$$\tilde{s}(t) \cdot w(t) = s(t)w(t) + gm \sum_k w^2(t) \quad (7)$$

아래 <그림 3>은 추출 절차를 도식적으로 설명한 것이다.



<그림 3> 은닉 정보 추출 과정도

4. 실험 결과

실험 항목은 MP3, AAC, WMA, RA 등 오디오 압축과 A/D 변환과 일반적인 신호처리에 대한 견고성을 평가하였다. Real Audio 와 WMA 는 사운드 카드를 통해서 나오는 출력 신호를 22.05kHz, 16bits 로 녹음해서 검출한 결과이다. 실험한 곡은 총 4 곡이다.

<표 1> 실험 결과 (BER : Bit Error Rates (단위 %))

Signal Music Processing	TA1	TA2	TA3	TA4	합계	BER (%)
Payload (bits)	1024	3584	3840	3072	11520	-
MP3-64kbps	7	9	74	44	134	1.42
AAC-64kbps	8	20	119	14	161	1.40
WMA-64kbps	43	127	231	106	507	4.40
RA-44kbps	73	299	316	283	971	8.43
Re-sampling	11	17	44	6	78	0.68
A/D Conversion	35	102	188	87	412	3.57

<표 1>에서 볼 수 있듯이 BER(bit error rate)값이 가장 큰 경우인 Real Audio 로 압축한 경우에도 6.75%로 오류 정정 코드로 수정할 수 있는 범위 10%를 넘지 않는 것을 알 수 있다. 따라서 일반적으로 인터넷 상에서 통용되는 대부분의 압축된 오디오 저작물의 저작권 보호에 적용이 가능한 기술임을 알 수 있다.

5. 결론 및 향후 연구 과제

본 논문에서는 오디오 신호로부터 생성한 운반신호를 이용한 워터마크 생성에 관하여 설명하였으며, 제안된 알고리즘을 이용한 실험결과에서는 A/D 변환과 다양한 형태의 손실 압축 후에도 은닉 정보를 완벽하게 추출할 수 있음을 보여주고 있다.

향후, 제안된 기술이 pitch shift, time scaling 등의 공

격을 받은 저작물에서 워터마킹을 검출하기 위해서는 오디오 segmentation 및 워터마크의 삽입 위치를 결정하는 알고리즘을 보다 세밀하게 적용할 수 있어야 할 것이다.

Acknowledgement

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

참고문헌

- [1] C. I., Podilchuk, E.j. Delp, "Digital watermarking: algorithms and applications", IEEE Signal Processing Magazine, Volume: 18 Issue: 4, July 2001 pp.33-46
- [2] P. Bassia, I. Pitas, and N. Nikolaidis, "Robust audio watermarking in the time domain", IEEE Transactions on Multimedia, Vol. 3, June 2001, pp. 232-241.
- [3] D. Kirovski and H. Malvar, "Robust spread-spectrum audio watermarking", 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 3, 2001, pp. 1345-1348.
- [4] L. Boney, A.H. Tewfik, and K.N. Hamdy, "Digital watermarks for audio signals", In International Conference on Multimedia Computing and Systems, pp. 473-480, IEEE, Hiroshima, Japan, 17-23 Jun. 1996.
- [5] D. Gruhl, A. Lu, W. Bender, "Echo hiding", in Proc. Information Hiding Workshop, University of Cambridge, U.K., 1996, pp. 295-315.
- [6] S. W. Foo, T. H. Yeo, and D. Y. Huang, "An Adaptive Audio Watermarking System", Electrical and Electronic Technology, 2001. TENCON. Proceedings of IEEE Region 10 International Conference on, Volume: 2, 2001 pp.509-513
- [7] H. O. Oh, J. W. Seok, J. W. Hong and D. H. Youn, "New echo embedding technique for robust and imperceptible audio watermarking", Acoustics, Speech, and Signal Processing, 2001. Proceedings. 2001 IEEE International Conference on, Volume: 3, 2001, pp.1341-1344
- [8] 신승원, 김종원, 최종욱, "디지털 필터를 이용한 오디오 워터마킹", 한국정보보호학회 종합학술발표회 논문집, Vol. 11. No.1 pp.464-468.
- [9] R. Petrovic, "Audio signal watermarking based on replica modulation", Telecommunications in Modern Satellite, Cable and Broadcasting Service, 2001. TELSIKS 2001. 5th International Conference on, Volume: 1, 2001 pp. 227-234.
- [10] J. Laroche, "Autocorrelation method for high-quality time/pitch-scaling", Applications of Signal Processing to Audio and Acoustics, 1993. Final Program and Paper Summaries., 1993 IEEE Workshop on, 1993, pp. 131-134
- [11] P.J.B. Jackson, C.H. Shadle, "Performance of the pitch-scaled harmonic filter and applications in speech analysis", Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on, Vol. 3, 2000 pp. 1311-1314 vol.3