

MP3 와 AAC 압축에 강인한 오디오 워터마킹 기술에 관한 연구¹

이한호*, 김종원*, 최종욱*
*마크애니 연구소
e-mail : sslh2@markany.com

A Study of Robust Watermarking Technique against MP3 and AAC Audio Compression

Han Ho Lee*, Jong Weon Kim**, Jong Uk Choi*
* MarkAny Research Institute

요 약

본 논문은 심리음향모델과 주파수변환을 이용하여 MP3 와 AAC 의 압축에서 강인하게 살아남을 수 있는 디지털 오디오 워터마킹 알고리즘에 관한 것이다. 워터마크를 의사난수열이나 이미지 등 외부 정보를 이용하지 않고 원본음악으로부터 생성시킨다는 것이 본 논문의 가장 큰 특징으로 원본 오디오로부터 생성된 워터마크는 음악과 융합 되어 워터마크의 삽입여부를 일반인의 청각으로는 인식할 수 없다

1. 서론

오디오 워터마크에 관한 연구는 크게 4 가지 방법으로 이루어져왔다. 하위비트 조작기법[1][2][3], 위상 부호화 기법(Phase coding)[1], 확산 스펙트럼 기법(Spread Spectrum)[1][4][5][6][7][8] 그리고 반향 은폐 기법[1]이 오디오 워터마크의 대표적인 기술이다. 위에서 언급한 워터마크 기술 외에도 저작권 정보를 더욱 강력하게 보존하기 위한 방법들이 연구되었다. Tsung-Li[9]는 워터마크 기술에 암호화를 첨부하여 저작권 보호를 강력하게 하고자 하였다. 그리고 Chun-Shien[10]은 한가지 방법의 워터마크를 삽입한 기존의 방법과 달리 2 가지 이상의 워터마크 삽입 방법을 이용하여 다중으로 워터마크를 삽입하였다. 워터마크 기술의 연구 개발과 달리 Bender[1]와 Fabien[11]은 기존의 워터마크 기술들을 종합 평가하는 연구를 진행하기도 했다.

본 논문에서 제시하고자 하는 것은 디지털 오디오에 저작권을 보호할 수 있는 워터마크 기술이다. 최근에 오디오 압축기술의 발달, 인터넷의 발달 그리고 휴대용 디지털 음악 재생기(PD: Portable Device)의 등장으로 디지털 오디오에 대한 저작권 보호기술의 필

요성이 높아지고 있다. 특히 오디오 압축기술의 발달은 워터마크 기술 개발의 필요성을 더욱 부추기고 있다. MP3 와 AAC 를 비롯한 대부분의 압축 기술들이 손실 압축방법이기 때문에 삽입된 워터마크도 압축 시에 손상을 입게 되므로 워터마크의 검지가 힘들어진다. 워터마크를 삽입하는 기술개발에서 몇 가지 주의 해야 할 사항이 있다.

- ① 사람의 귀에 들리지 않아야 한다.
- ② 원본 없이도 추출이 가능해야 한다.
- ③ MP3 와 AAC 와 같은 손실 압축에도 살아 남아야 한다.

사람의 귀에 들리지 않게 해야 하는 것은 오디오 데이터의 상품성에서 매우 중요한 부분이고, 원본을 필요로 한다면 그 관리가 어렵고 변조의 가능성이 높다. 그리고 마지막으로 앞에서 언급했던 바와 같이 손실 압축에 강해야만 한다.

기존의 오디오 워터마크 연구에서는 주로 워터마크로 의사난수[4][5][6][7][8]또는 육안으로 식별이 가능한 이미지 형태의 워터마크[12][13]를 사용해 왔다. 본 논문에서는 위에서 언급한 방법들과 다르게 매우 독특한 워터마크를 생성하여 사용한다. 각 오디오마다

¹ 본 논문은 NRL 과제에 의해서 지원되었으며, 특허 청구되어져 있습니다.

가지는 특징을 워터마크로 사용한다. 주파수 분석을 통해서 오디오 워터마크를 생성하고, 삽입하고 추출하게 된다.

2. 실험 평가 방법

워터마킹 기술력은 워터마킹의 강인성과 비인지성으로 평가하는데 이 두 항목은 상보성(tradeoff)관계가 있다. 따라서 심리음향모델(Pschoacoustic Model)[14][15]을 이용하여 강인성과 비가청성을 충족시키려는 논문들이 많이 발표되었다. 본 논문에서는 먼저, 사람의 귀에 들리는가의 측정을 위해서 워터마크를 삽입한 오디오데이터와 원본 데이터의 SNR(Signal to Noise Ratio)을 측정하여 데이터의 손상도를 측정하였다. SNR은 순음에 잡음이 얼마나 섞여 있는가를 측정할 수 있는 척도이다. 잡음이 많고 강하면 강할수록 사람의 귀에는 음질이 좋지 않게 들린다. 따라서 본 논문에서는 사람의 귀에 들리는가를 측정하기 위해서 PCM 파일 포맷에 워터마크를 삽입한 후에 SNR 값을 측정하여 본 논문의 오디오 워터마킹 방법의 비가청성을 평가하였다.

다음으로, 워터마크의 강인함을 알아보기 위해서 대표적인 손실 압축방법에 대해 얼마나 살아 남는가를 측정하였다. 손실 압축에 가장 많이 쓰이고 있는 MP3와 AAC를 각각 전송률 128kbps와 64kbps로 압축한 후에 워터마크를 추출하였다. 보편적으로 사용되어지고 있는 압축률은 128kbps이지만, 고압축을 원할 경우 64kbps도 사용되어진다. 두 가지 압축형식과 각각의 압축률에 대해서 워터마크의 생존성을 측정하여 본 연구에서 개발한 기술을 객관적으로 평가하였다.

3. 워터마크 삽입방법

기존의 워터마크 삽입 방법들은 의사난수를 발생시키거나, 이미지 데이터를 워터마크로 사용했다. 본 연구에서는 전혀 새로운 방식을 채택하기로 한다. 본 연구의 가장 핵심적인 아이디어이다. 의사난수 방법은 키에 의해서 발생한 의사난수를 사용한다. 본 연구에서는 서로 다른 오디오 데이터들이 각기 가지고 있는 특징을 이용하여 워터마크를 생성한다.

그림 1-(a)는 오디오 데이터를 FFT를 이용하여, 시간영역의 데이터를 주파수 영역으로 변환하여 표시한 그림이다. FFT를 이용하여 주파수영역 데이터로 변환을 하면, 실수부와 허수부가 나오게 되는데, 그림 1-(a)는 실수부를 절대값으로 표시한 것이다. 그림 1-(a)의 사각형 영역을 확대 표시한 그림이 그림 1-(b)이다. 본 연구에서는 그림 1-(b)에서 보이는 바와 같이 원본데이터(O)와 변형시킬 데이터(W)를 선택한다. 이때, 두 데이터의 지정은 바로 인접한 데이터로 정한다. 이 방법은 심리음향 모형에서 사람이 주파수 강도가 비슷한 인접된 두개의 주파수를 구분할 수 없다는 것을 이용한 것으로 워터마크를 들리지 않게 삽입한다. 사람의 귀는 이를 구분할 수 있을 정도로 예민하지 못하다. 그림 1-(c)에서 보이는 바와 같이, 인접한 실수부 데이터를 원본 데이터와 같은 값으로 바꾼다. 이

로써 워터마크를 구성하는 하나의 비트가 생성된 것이다.

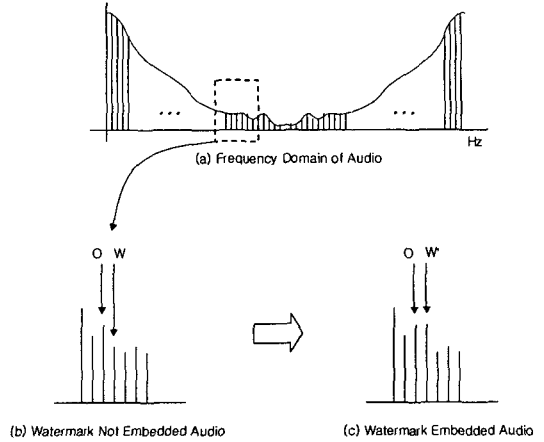


그림 1 워터마크 신호의 생성 예

그림 2-(a)는 최종적으로 생성된 워터마크를 나타낸다. 워터마크는 변형된 신호(W')의 집합으로 구성된다. 즉, 오디오 자체가 가지고 있는 특징을 가지고 워터마크를 생성시키는 것이다. 이 워터마크는 프레임단위로 그림 1의 방법으로 n번 반복 삽입되어 최종적으로 그림 2-(a)와 같은 난수(random number)의 성격을 가지게 되며, 노래마다 그 특성이 구별되는 차별성을 가진다. 이 차별성은 그동안 다른 워터마크 연구와는 다르게 음악과 워터마크가 오디오의 거부반응 없이 잘 결합될 수 있다는 점이 본 연구의 큰 장점이다. 이런 특징들로 인해 생성된 신호는 워터마크로 사용되기에 충분한 조건을 가지고 있다고 볼 수 있다.

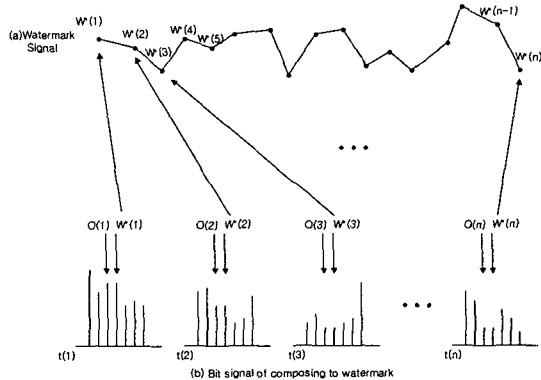


그림 2 워터마크의 생성 예

생성된 워터마크는 1초 단위로 삽입된다. 그림 2-(b)에 나타나 있는 단위시간(t)은 임의의 시간으로, 단위시간이 작을수록 워터마크를 구성하는 난수의 개수는 많아진다. 단위시간을 작게 하면 워터마크를 구성하는 난수열의 개수가 길어져서 워터마크가 오점지율이 적은 반면에 강인함(robustness)에 문제가 있다.

단위시간을 크게 해서 워터마크를 구성하는 난수의 개수를 줄이면, 강인함은 높아지지만 오검지율이 높아진다. 본 연구의 실험에서는 단위시간을 0.03 초로 실험하였고, 워터마크를 구성하는 열의 길이는 약 30 개로 하였다. 따라서 1 초에 하나의 워터마크열이 삽입되어진다.

그림 1-(b)에서와 같이 원본 데이터(O)와 변경할 데이터(W)의 선정을 하였다. Sampling rate 가 44100kHz 인 오디오데이터를 그림 3 과 같이 30Hz 의 주파수 해상도를 가지는 FFT 한 후에 원본 데이터(O) 주파수는 5670Hz 로 규정하였고, 변경될 데이터(W)는 바로 인접한 5700Hz 로 정하였다. 30Hz 의 간격의 주파수 변화는 심리음향모델에 의해서 인지하기가 불가능하기 때문이다.

그림 3 은 워터마크 삽입과정을 나타내고 있다. 그림에서 입력된 신호(O)는 시간영역에서 주파수영역으로 데이터가 변환되고, 워터마크 생성부로 데이터가 입력된다. 워터마크 생성부에서는 워터마크를 생성하는데, 삽입 대역선정에서 워터마크를 삽입 할 주파수 대역을 선정하게 된다. MP3 나 AAC 의 손실압축에서 사용되는 마스킹 기법을 고려하여 워터마크가 사라지지 않게 생성한다. 마지막으로 청각심리모형을 적용하여 워터마크를 사람의 귀에 들리지 않는 영역에 삽입한다.

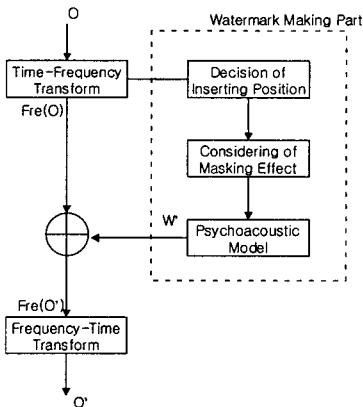


그림 3 워터마크 삽입 흐름도

4. 워터마크 추출 방법

워터마크 추출 알고리즘은 워터마크 삽입 알고리즘의 역과정으로 진행된다. 그림 4 에서와 이 단위시간(t)으로 FFT 변환 후에 원본데이터(O)의 신호와 변형된 워터마크 신호(W)열을 각각 뽑아낸다. 그림 4-(a)와 (b)에서 같이 추출된 두 신호의 상관도(Correlation)를 자기상관 함수 [식 1]에 적용하여 측정한다.

$$r = \frac{\sum(O - \bar{O})(W - \bar{W})}{\sqrt{\sum(O - \bar{O})^2} \sqrt{\sum(W - \bar{W})^2}} \quad [식 1]$$

여기서, \bar{O}, \bar{W} 는 평균값, r 은 상관도 값을 의미한

다. 계산된 상관도 값이 일정 임계값(threshold)을 넘으면 워터마크가 있는 것으로 판단하고, 임계값보다 낮으면 워터마크가 없는 것으로 판단한다.

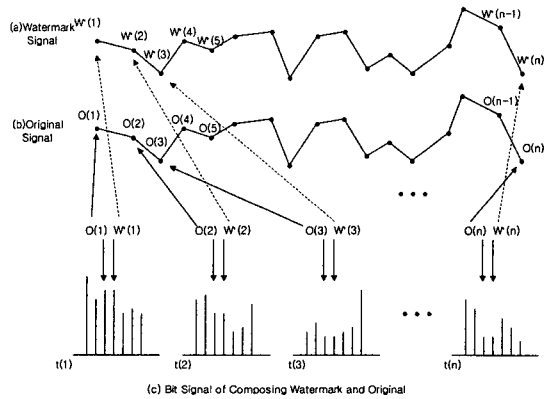


그림 4 원본신호와 워터마크의 신호의 추출 예

5. 실험결과

5.1 비가청성 측정

사람의 청각능력이 천차만별이기 때문에 Inaudibility 를 객관적으로 측정하는 것은 매우 어렵다. 보통사람에는 들리지 않는 소리가 일명, 골든이어(golden ear)들에게는 들릴 수도 때문이다. 이에 대한 객관적인 평가를 위해서 본 연구에서는 SNR(Signal to Noise Ratio)를 사용하였다. SNR 은 신호에 잡음이 얼마나 많이 포함되어 있는가를 측정하는 척도로서, Inaudibility 의 측정에 적합한 방법으로 생각된다.

$$SNR = 10 \log_{10} \left(\frac{Signal^2}{(Noise - Signal)^2} \right) \quad [식 2]$$

표 1 워터마크 된 오디오 데이터의 SNR 값

Channel	왼쪽	오른쪽
SNR 값	34.0319dB	34.1106dB

표 1 에서 알 수 있듯이 워터마크 된 오디오 신호의 SNR 이 평균 34dB 로서 오디오에 노이즈의 비율이 매우 적은것으로 평가되었으나, 실질적으로 비가청성 측정에 대한 가장 적절한 방법은 아니다.

5.2 강인성 측정

Robustness 의 객관적인 측정을 위해서 테스트 오디오를 3 가지의 장르로 구분하여 실험하였다. 본 알고리즘이 주파수 영역으로 변화하여 워터마크를 삽입하고 추출하므로 오디오의 특성에 따라 검지율이 달라질 것으로 판단했기 때문이다.

실험 결과 예상대로 음악에 따라서 검지율이 많은 차이를 보였다. 실험에서 검지율은 50%이상을 목표로 하였다. 최근 SDMI 나 STEP2000 과 같은 오디오 워터마크 표준협회의 robustness 측정 기준이 50%의 검지율을 기준으로 잡고 있다. 표 2 과 3 에 나타난 결과를 보면, Pop 과 Classic 은 모든 MP3 와 AAC 128kbps 의 압축에서 50%이상의 검지율을 보였다. 그

러나 AAC의 64kbps 압축에서는 50%에 미치지 못하는 결과를 산출했다.

표 2 MP3로 압축된 오디오 데이터의 각 장르별 워터마크 검지율

MP3 128kbps compression						
장르	Pop		Classic		Rock	
Channel	왼쪽	오른쪽	왼쪽	오른쪽	왼쪽	오른쪽
검지율	89.1	89.1	74.4	88.5	81.3	81.8

MP3 64kbps compression						
장르	Pop		Classic		Rock	
Channel	왼쪽	오른쪽	왼쪽	오른쪽	왼쪽	오른쪽
검지율	74.5	80.0	73.1	70.5	50.9	52.7

표 3 AAC로 압축된 오디오 데이터의 각 장르별 워터마크 검지율

AAC 128kbps compression						
장르	Pop		Classic		Rock	
Channel	왼쪽	오른쪽	왼쪽	오른쪽	왼쪽	오른쪽
검지율	90.9	89.1	92.3	97.4	74.5	87.3

AAC 64kbps compression						
장르	Pop		Classic		Rock	
Channel	왼쪽	오른쪽	왼쪽	오른쪽	왼쪽	오른쪽
검지율	41.8	38.2	35.9	25.6	23.6	23.6

6. 결론

연구 결과 AAC 64kbps를 제외한 MP3와 AAC 압축에 대해서 워터마크가 50%이상 살아 남는 좋은 결과를 보였다. 연구의 객관성을 보여주기 위해서 다양한 장르의 곡을 테스트한 결과 pop과 classic에서는 좋은 결과를 보였으나, rock에서는 결과가 다른 곡들보다 다소 떨어졌다. 주파수 영역으로 표시하여 rock계통의 음악을 분석한 결과 rock 음악에는 주파수가 전반적으로 높게 나타났다. 그 이유는 전자기타, 전자건반 등과 같은 악기들이 고주파의 음을 많이 내기 때문인 것으로 판단된다. 이 부분은 차츰 보완해 연구할 분야인 것으로 보인다.

오디오 워터마크가 해결해야 하는 부분은 매우 많다. 본 연구에서는 대표적인 압축 방식인 MP3와 AAC에 대해서 테스트를 했지만, 이 밖에도 ATRAC-3, ATRAC, Q-Design, VQF, RA 등 다양한 압축방법이 있으며, 이들 압축 방법들이 가지는 특성 역시 각기 다르다. 각각 다른 압축 방식에 모두 대응할 수 있는 워터마크를 만드는 일은 매우 힘든 작업이다.

압축에 대한 강인함 외에도 워터마크는 더욱 많은 공격에 살아 남아야 한다. 최근에는 DVD와 같은 새로운 미디어가 등장하여 상용화 되어지고 있는데, DVD 미디어 같은 경우 주파수 영역이 96kHz로 CD의 두 배 이상이며, sampling quantization이 24bit로 CD의 16bit보다 훨씬 높다. DVD 데이터를 48kHz 또는 44.1kHz와 같이 downsampling했을 때, 그리고 quantization level을 24bit에서 16bit로 requantization했을 때, 워터마크가 사라지지 않게 하는 연구도 이루어져야 한다.

참고문헌

[1] Bender W., D. Gruhl, N. Morimoto and A. Lu, "Techniques for Data Hiding", *IBM Systems Journal*, VOL

35, NOS 3&4, p.313-336, 1996

- [2] Fabien A. P. Petitcolas, Ross J. Anderson and Markus G. Kuhn. "Information hiding", *Proceedings of the IEEE, special issue on protection of multimedia content*, 87(7):1062--1078, July 1999
- [3] Basia, V. and I. Pitas, "Robust Audio Watermarking in the time-domain", *Proceeding of EUSIPCO'98*, September 8-11, Rhodes, Greece, 1998.
- [4] F. Hartung and B. Girod, "Watermarking of Uncompressed and Compressed Video", *Signal Processing*, vol. 66, no. 3 (Special issue on Watermarking), pp. 283-301, May 1998
- [5] M. Ramkumar, A.N. Akansu, X. Cai, "Floating Signal Constellations for Multimedia Steganography", *presented at IEEE ICC*, New Orleans, LA, June 2000.
- [6] M. Ramkumar, A.N. Akansu, "FFT-Based Signaling for Multimedial Steganography", *presented at IEEE ICCASP*, Istanbul, Turkey, June 2000.
- [7] I.J.Cox, J. Kilian, T. Leighton and T. Shamoan, "Secure Spread Spectrum Watermarking for Images, Audio and Video", *Proceeding Of 1996 International conference On Image Processing(ICIP '96)*, vol. III, pp. 243-246(1996)
- [8] I.J. Cox, J. Kilian, T. Leighton and T. Shamoan, "A Secure, Robust Watermark for Multimedia", *Workshop on Information Hiding*, Newton Institute, Univ. of Cambridge, May 1996.
- [9] Wu T.L. and S.F. Wu, "Selective Encryption and Watermarking of MPEG Video", in *Proceeding of International Conference on Image Science, Systems, and Technology, CISS'97*, June 1997.
- [10] Chun-Shien Lu, Hong-Yuan Mark Liao, and Liang-Hua Chen, "Multipurpose Audio Watermarking", *15th International conference On Pattern Recognition*, Spain, Sept. 3-8, 2000.
- [11] Fabien A. P. Petitcolas and Ross J. Anderson, "Evaluation of copyright marking systems", *In proceedings of IEEE Multimedia Systems'99*, vol. 1, pp. 574-579, 7-11 June 1999, Florence, Italy.
- [12] Zeng W. "Digital watermarking and data hiding: technologies and applications," *Invited Talk, Proceeding of International Conference on Information Systems, Analysis and Synthesis*, 1998
- [13] Wu M., H. Yu and A. Gelman: "Multi-level Data Hiding for Digital Image and Video", *SPIE Photonics East '99*, Boston, 1999
- [14] Laurence Boney, Ahmed H. Tewfik, and Khaled N. Hamdy, "Digital Watermarks for Audio Signals", *1996 IEEE International Conference on Multimedia Computing and Systems*, June 17-23, Hiroshima, Japan, p. 473-480.
- [15] Laurence Boney, Ahmed H. Tewfik, and Khaled N. Hamdy, "Digital Watermarks for Audio Signals", *EUSIPCO-96, VIII European Signal Proc. Conf.*, Trieste, Italy, Se