

MPEG-2 AAC 심리음향 모델을 이용한 오디오 워터마킹 기법

성종수, 강상구, 신재호
동국대학교 전자공학과

The Audio Watermarking method Using the MPEG-2 AAC Psychoacoustic Model

Jong-soo Sung, Sang-goo Kang, Jaeho Shin
Dept. of Electronic Eng., Dongguk University
E-mail : jhshin@cakra.dongguk.ac.kr

Abstract

In this paper, we present a method for embedding digital watermarks into digital audio signals. The watermarking must be imperceptible and should be robust to attacks, such as filtering and compression etc. In our method, we adaptively embedded the watermarks changing the scale factor using the spread spectrum and MPEG-2 AAC psychoacoustic model.

I. 서론

최근 인터넷 망을 통하여 멀티미디어 데이터가 급속하게 분배되고 있고, 현재 VOD등의 디지털 음성 및 영상 서비스를 위해 표준화 작업에 있는 DAVIC이 마무리 될 경우, 불법 복제 및 전파와 같은 소유권 문제는 더욱더 중요시 될 것이다. 혼존하는 대표적인 보호 시스템은 데이터를 암호화(Encryption)된 형태로 전송 및 저장하는 방법으로 반드시 복호화가 되어야 데이터를 이용할 수 있으나, 일단 복호화가 된 데이터는 불법적으로 복사가 되고 전송될 수 있는 문제점이 있다. 따라서, 멀티미디어 데이터 자체에 소유권 정보를 삽입하여 복사 방지 시스템에 적용될 수 있는 워터마킹 기법이 필요하다.

워터마킹의 기본적인 개념을 살펴보면, 오디오나 영상 등의 멀티미디어 데이터에 들리지 않거나 보이지 않도록 수정함으로써 데이터의 소유권 정보를 삽입하

는 과정으로 일단 워터마킹된 데이터는 권한이 없는 사용자에게 통계적으로 검출이 불가능하여야 하며 필터링이나 압축 등과 같은 공격에도 소유자에 의해서는 검출이 가능하여야 한다[2][3].

본 논문에서는 MPEG 압축 공격에 대해 개인성을 제공하기 위하여 주파수 영역에서의 대표적인 워터마킹 기법인 Ingemar J. Cox의 대역확산 (spread spectrum) 기법에 MPEG-2 AAC의 심리음향 모델을 적용하여 웨이브렛 영역에서 워터마크의 스케일 팩터를 적응적으로 변환시키면서 워터마크를 삽입하였다 [1][5].

본 논문의 2장 MPEG-2 AAC 심리음향 모델에서는 심리음향 모델의 전체 흐름에 관하여 설명하고, 3장 워터마크 알고리즘에서는 워터마크 삽입 및 검출 방법에 대해, 4장 실험 결과 및 고찰에서는 실험결과와 실험 중 문제점에 대해 살펴보고 5장에서 결론을 맺는다.

II. MPEG-2 AAC 심리음향 모델

ISO/IEC 13818-7 MPEG-2 AAC(Advanced Audio Coding)는 ITU-R의 평가 척도에 의한 투명 품질, 모노와 스테레오를 포함하여 7+1 채널까지의 다채널, 다국어를 지원하며 MPEG 오디오 표준 중 가장 고 압축율을 보장한다. 기존의 MPEG 오디오에 대한 후방 호환성을 가지지 않는다 하여 MPEG-2 NBC(Non-Backward Compatible standard)라 불리기도 한다[5].

MP3(MPEG-1 Layer III)는 128kbps에서 CD의 음질

을 보장하고 있지만 64kbps이하의 bitrate에서 음질이 저하되는 문제를 가지고 있다. AAC는 스테레오 모드에서 44.1Khz로 sampling하여 기존의 MP3 등을 대치 할 수 있고, 64Kbps에서 MP3와 동일한 음질을 보장하며 very low bitrate coding method인 MPEG-4(ISO/IEC 14496-3)의 T/F coder로 사용되고 있다[5].

심리음향 모델은 MPEG-1의 Psychoacoustic Model II에서 몇몇 파라미터가 추가되었다. 그림 1에서 심리음향모델의 전체 구조를 나타내었다.

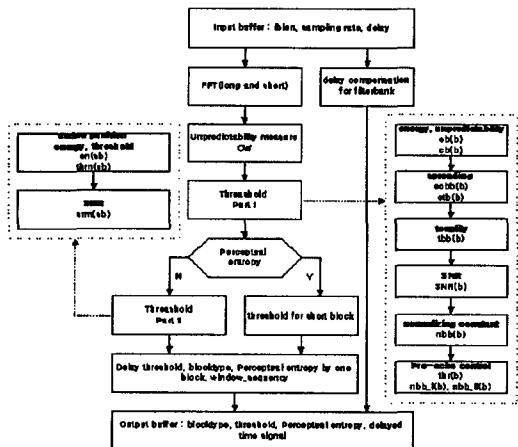


그림 1. Psychoacoustic model

심리음향 모델에서는 signal energy에 의해 masking 되는 masking curve를 구하게 되는데, 이 energy를 threshold라고 하며, *iblen*, delay, sampling rate등의 입력을 받아, SMR(Signal to Mask Ratio)과 threshold, delayed time domain data, window type, expectation bitrate의 출력 결과를 얻는다[5].

*Iblen*은 threshold 계산과정에 필요한 shift length로써 본 논문에서 사용한 long FFT의 경우는 1024이다.

III. 워터마크 알고리즘

Ingemar J. Cox에 의해 제안된 대역확산 기법은 주파수 영역에서의 대표적인 워터마킹 기법으로 협대역 신호를 넓은 대역폭에 걸쳐 확산시킴으로써 특정 주파수에서 볼 때 신호의 에너지가 아주 작게 나타난다. 즉, 워터마크가 여러 주파수 대역으로 확산되어 매우

작은 에너지를 갖게되며, 결국은 워터마크가 검출될 수 없다. 워터마크를 없애기 위해서는 높은 진폭을 갖는 잡음을 모든 주파수 대역에 추가해야 할 것이다. 그럼에도 불구하고 소유권자는 워터마크와 워터마크가 삽입된 위치를 알고 있기 때문에 많은 약한 신호들을 모으면 높은 신호 대 잡음비로 워터마크를 검출할 수 있다.

워터마크를 주파수 성분에 삽입하기 위해서 다음 식에 의해 수행된다[1].

$$V_i = V_i + \alpha x_i \quad (1)$$

$$V_i' = V_i(1 + \alpha x_i) \quad (2)$$

$$V_i' = V_i(e^{\alpha x_i}) \quad (3)$$

식에서 V_i 는 원신호, V_i' 는 워터마크된 신호, α 는 스케일 팩터, 그리고 x_i 는 워터마크를 각각 나타낸다. 식 (2)(3)에 의한 워터마크 삽입이 여러 스케일에 더욱 더 강인하기 때문에 본 실험에서는 식(2)를 이용했다.

본 논문에서는 워터마크로 Pseudo-Noise sequence를 사용하였는데, 간략하게 살펴보면 다음과 같다. 정 보신호를 변조하는데 사용되는 PN sequence는 잡음과 같은 특성을 갖고 간섭에 강하며 좋은 auto-correlation 특성을 갖는다. 그림 2와 같이 고정된 길이 m 의 linear feedback shift register에 의해 $N(0,1)$ 인 정규 분포에 따라 생성되며, 최대주기는 $N=2^m-1$ 가 된다[4].

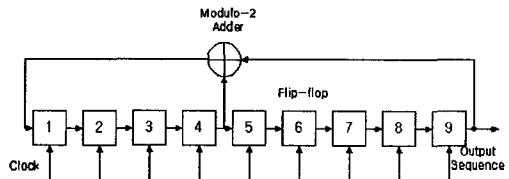


그림 2. Shift register($m=9$, $N=511$)

워터마크 삽입과 검출 알고리즘은 다음과 같다.

① 워터마크 삽입

AAC coding에 사용되는 masking curve에 의해서 원 신호는 인간이 느끼지 못하는 정도의 왜곡이 일어나게 된다. 이러한 왜곡은 저작권자가 삽입한 워터마크에 영향을 미치게되며 검출시 워터마크의 왜곡 및 손실을 가져오게 된다. 이러한 손실을 줄이기 위해 왜곡의 강도에 따라 스케일 팩터(α)를 적응적으로 변화 시켜 워터마크의 손실을 줄일 수 있다. 이러한 손실을 줄이기 위해 본 논문에서는 원 신호와 masking curve의 비율인 SMR을 사용하였다.

I. Cox가 제안한 DCT변환 또는 FFT변환 대신 웨이브렛 변환을 한 후 워터마크를 삽입하였다. 웨이브렛을 이용할 경우 주파수 영역에 대한 특성 및 시간 영역에 대한 특성 또한 고려하여 워터마크를 삽입할 수 있다.

심리음향모델의 적용은 마스킹값이 최고인 주파수 영역과 대칭되는 웨이브렛 변환계수영역을 선택하여 approximation영역에 삽입하였다. 그림 3에서는 제안된 워터마크 알고리즘을 보인다.

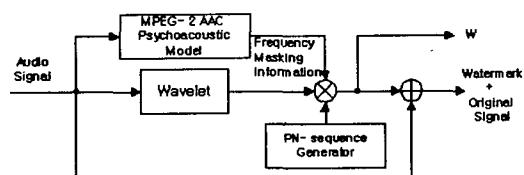


그림 3. 워터마크 삽입 알고리즘

② 워터마크 검출

오디오 신호는 MPEG 압축과 같은 잡음에 의해 변형될 수 있는데, 이러한 변형된 오디오 신호에서 워터마크를 검출하기 위해서 본 논문에서는 PN-sequence의 자기 상관 특성을 이용한다. 먼저 식 (4)에 의해서 식 (5)와 같이 워터마크가 삽입된 오디오 신호와 원 오디오 신호의 차 D 를 구하고, D 와 소유권자가 소유하고 있는 워터마크 W 와의 cross-correlation을 구함으로써 상관도를 비교할 수 있다.

$$\begin{aligned} V'_i &= V_i(1 + \alpha x_i) \\ &= V_i + V_i \alpha x_i \end{aligned} \quad (4)$$

$$\begin{aligned} D &= V'_i - V_i \\ &= V_i \alpha x_i \\ &= W \end{aligned} \quad (5)$$

IV. 실험결과 및 고찰

이 장에서는 3장에서 언급한 워터마크 삽입 및 검출 알고리즘을 이용하여 실험한 결과에 대해 논의한다.

샘플 음원으로는 20초 길이의 ‘비발디의 사계’ 중 ‘봄’을 이용하였으며, 웨이브렛의 기저로는 db1~db7 및 db30을 사용하였다. 먼저 샘플 음원을 1024개씩 나누고 PN-sequence 1024개를 나눈 샘플 음원에 각각 삽입하였다.

실험결과 5, 6 레벨에 삽입하였을 경우와 db가 높을 수록 음질이나 장인성면에서 좋은 결과를 얻었다. ‘db’

가 높을수록 프레임간의 overlap이 작기 때문에 검출율이 높게 나타나는 것으로 분석된다. Threshold를 0.45로 생각할 경우 AAC 압축에서 95%이상의 검출율을 나타내었다.

Level	Attack	
	MPEG Layer-III	AAC
3	0.3101	0.3634
4	0.3407	0.3821
5	0.3841	0.4280
6	0.3794	0.4159
7	0.2129	0.3051

표 1. 워터마크 검출(db30)

Level	Attack	
	MPEG Layer-III	AAC
3	0.2273	0.3153
4	0.2362	0.3276
5	0.2586	0.3409
6	0.2428	0.3321
7	0.2109	0.2499

표 2. 워터마크 검출(db3)

Level	Attack	
	MPEG Layer-III	AAC
3	0.2344	0.3196
4	0.2392	0.3314
5	0.2791	0.3872
6	0.2620	0.3716
7	0.2262	0.2807

표 3. 워터마크 검출(db5)

Level	Attack	
	MPEG Layer-III	AAC
3	0.2601	0.3234
4	0.3017	0.3421
5	0.3341	0.3880
6	0.3494	0.3976
7	0.2829	0.3074

표 4. 워터마크 검출(db7)

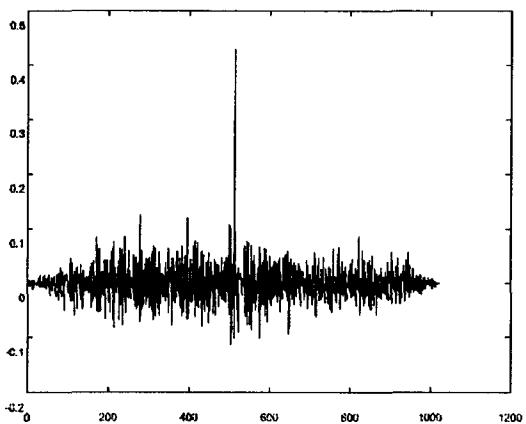


그림 4. 워터마크의 검출
(1024 samples, db30, level-5)

V. 결론 및 추후연구

Network 기반이 크게 발전하고 인터넷이 보편화됨에 따라 멀티미디어 데이터의 자유로운 정보교환이 이루어지게 되었지만 저작권을 침해하는 자유로운 복제나 도용이 쉬워졌다. 또한 128kbps에서 CD음질을 보장하는 MP3를 대신해서 64kbps에서 상용하는 성능을 발휘하는 AAC 압축기법이 차세대 오디오 압축기술로 주목받고 있고, 이로인해 급증하게 될 AAC의 사용에 있어서 불법적인 사용을 막기 위한 워터마킹 기법이 필요하다.

웨이브렛과 MPEG-2 AAC 심리음향 모델을 이용한 워터마크 모의 실험결과, 어떠한 웨이브렛 기저를 사용하였나에 따라 조금은 차이가 있지만 평균적으로 높은 레벨보다는 낮은 레벨에 워터마크를 삽입할 경우 검출률은 좋았으나 음질에 문제가 생겼으며, db30의 기저를 사용하여 레벨 5 또는 6에 워터마크를 삽입할 경우 검출과 음질면에서 가장 좋은 결과를 얻을 수 있었다.

본 논문에서 MPEG-2 AAC 심리음향 모델을 사용하여 AAC 오디오에 대한 워터마크 알고리즘을 제안하고, 이의 성능 및 문제점에 대하여 분석해 보았다. 성능을 좀더 개선하면 차세대 오디오 압축기술인 AAC에 적용할 수 있으리라 사료된다.

[* 본 연구는 '98 한국과학재단의 핵심전문연구비(981-0928-155-2) 지원으로 수행된 결과임]

참고문헌

- [1] Ingemar J. Cox, Joe Kilian, Tom Leighton and Talal Shamoon, "Secure Spread Spectrum Watermarking for Multimedia", NEC Research Institute, Technical Report 95-10, 1995
- [2] F. M. Boland, J. J. K. O Ruanaidh and C. Dautzenberg, "Watermarking Digital Images for Copyright Protection", Image Processing And its Applications, pp. 326-330, July 1995
- [3] Gerrit C. Langelaar, Jan C.A. van der Lubbe, Reginald L. Lagendijk, "Robust Labeling Method for Copy Protection of Images", Proceedings of SPIE ELECTRONIC IMAGING '97, storage and Retrieval for Image and Video Database V, pp. 298-309, February 1995.
- [4] Laurence Boney, Ahmed H. Tewfik, Khaled N. Hamdy, "Digital Watermarks for Audio Signals", http://lci.die.unifi.it/~piva/Watermarking/watermark_audio.html
- [5] ISO/IEC 11818-7 : 1993 Information technology-Generic coding of moving pictures and associated audio information part 7 : Advanced Audio Coding(AAC)