

오디오 워터마킹 프로세서 구조 설계에 관한 연구

김기영*, 김영섭**, 이상범***

* 단국대학교 전자컴퓨터공학과

** 단국대학교 전자 컴퓨터학부

*** 단국대학교 전기전자컴퓨터공학부

초록

A number of watermark insertion method is proposed for the protection of audio contents such as MP3 music. In this paper, we propose a VLSI architecture that performs embedding watermark to the audio signal based on the scheme that proposed by XUEYAO LI[1]. This architecture is implemented and simulated in Verilog HDL. This watermark embedding method used a visually recognizable binary image. Despite a unit that determines the watermark embedded intensity is removed to archive low complexity of H/W, our experimental results show that watermarked signal is perceptually transparency and robust to several known attacks.

1. 서론

멀티미디어 산업이 발전하고 디지털 휴대용 멀티미디어 기기의 보급 확대, 초고속 네트워크 기술의 대중화로 대용량 디지털 멀티미디어 콘텐츠의 유통이 활발히 이루어지고 있다. 그러나 디지털 콘텐츠의 경우, 인터넷을 통한 불법복제 및 불법배포가 쉽게 이루어지고 있어 저작권 침해 논란이 끊임없이 제기되고 있는 실정이다. 디지털 콘텐츠 산업의 활성화를 위해서는 디지털 콘텐츠의 지적 소유권이 적극 보호 되어야 하는데 소유권자에게 정당한 지적 재산을 보장해 줄 수 있는 기술적 해결책으로서 워터마킹 기법이 제시되고 있다.

디지털 워터마크는 원본 오디오에 삽입되는 디지털 정보나 패턴으로써 원본 신호의 음질에 영향을 주지 않아야 한다. 오디오 신호를 감지하는 인간의 청각시스템(HAS)는 시각(HVS)에 비하여 매우 민감하여 오디오 신호에 정보를 삽입하는 것은 상대적으로 더 어렵다. 일반적으로 오디오 워터마킹은 비지각성(Inaudible), 강인성(robustness)의 특성을 만족해야 한다.

그러나 이 두 가지 특성은 서로 적절한 절충점에서 조정되어야 한다. 즉, 비지각성을 더 좋게 하기 위해서 워터마크의 삽입 강도를 낮추면 워터마크의 강인성이 떨어지게 되고 반대로 강인성을 향상시키기 위해 삽입 강도를 증가 시키면 워터마크된 오디오의 음질이 저하된다. 따라서 삽입할 워터마크의 삽입 강도를 결정하는 방법이 매우 중요한 사항이다. 이 같은 워터마크의 삽입강도 결정에 있어서 최근의 연구는 바로 MPEG 오디오 부호화에 이용되는 심리음향 모델을 적용하는 것이다. 이 기법은 PN 시퀀스를 이용하여 변조된 워터마크 신호에 심리음향 모델의 마스킹 임계값을 곱하여 원본신호에 더해진다.[2]

MPEG 심리음향 모델을 이용하면 워터마크된 오디오 신호의 음질 저하를 최소화 할 수 있으며 압축이나 대역 통과 필터링 등의 공격에도 상당한 강인성을 나타낸다. 그러나 MPEG 에서 정의한 심리음향 모델(I, II)을 구현하기 위해서는 서브밴드 필터 뱅크, 1024-512 점 FFT, 제곱근 연산, 로그 연산 등 많은 계산 복잡도를 필요로 한다.[3]

2. 워터마킹 프로세서의 H/W 구현에 관한 고찰

최근의 오디오 워터마킹에 관한 연구는 스프레드 스펙트럼과 심리음향 모델을 이용한 워터마킹 기법이 음질과 강인성 면에서 비교적 우수하여 많은 주목을 받고 있다. 그러나 앞서 언급하였듯이 심리음향 모델 처리와 워터마크 처리를 하나의 전용 H/W(Dedicated H/W) 모듈로 구현하기 위해서는 많은 비용이 소요되고 비효율적이다.

기존에 국내에서 수행된 관련 연구의 결과를 살펴보면 이강현[4]의 연구에서는 심리음향 처리기(PAM)를 S/W 로 구현하여 마스킹 임계값과 블록 스위칭을 위한 제어신호를 PC 로부터 전달 받는다.

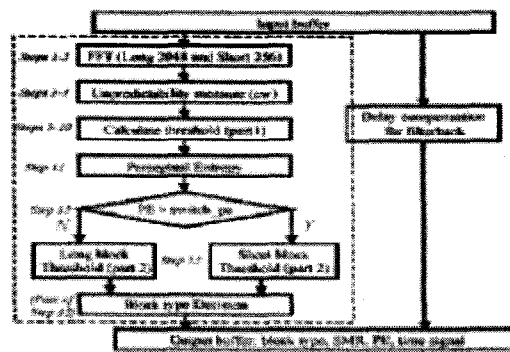


그림 1. 심리음향 모델(II)의 블록 다이어그램[5]

정남훈[6]의 MPEG-2 부호화기 설계의 연구에서는 프로세서 코어를 설계하여 마이크로-프로그램 방식으로 심리음향 모델을 처리하는 하이브리드 구조를 채택하여 구현하였다.

본 논문에서는 비교적 작은 사이즈의 전용 H/W IP(Intellectual Property)로 응용 될 수 있는 구조를 설계하기 위하여 복잡한 심리음향 모델을 이용하지 않았으며 시간-주파수 변환방법에 있어서도 서브밴드 필터링이 필요한 FFT 에 비하여 계산복잡도가 적은 DWT 를 이용한 워터마킹 기법을 적용하였으며 MATLAB 을 이용한 알고리즘 시뮬레이션 결과 음질과 강인성 면에서 비교적 좋은 결과를 얻었다.

3. 제안하는 워터마크 삽입 및 검출 방법

본 논문에서는 XUEYAO LI[1]가 제안한 방법을 기반으로 워터마크를 삽입하게 된다. XUEYAO LI 는 워터마크의 삽입 강도를 적응적으로 결정하기 위해 각 세그먼트 별로 스케일링 값을 계산하여 적용하지만 본 논문에서는 H/W 구현의 복잡성을 줄이고자 고정된 임계값을 적용하여 워터마크를 삽입한다. 삽입 기법을 각 단계별로 기술하면 다음과 같다.

Step1. 워터마크 이미지 처리

시각적으로 식별성이 뛰어난 이진 이미지를 워터마크로 사용한다. 이미지는 2 차원 신호이므로 식(3)에 의해 $M_1 * M_2$ 사이즈의 1 차원 신호로 변환한다.

$$V = \{v(k) = u(i, j), k = i * M_1 + j, 0 \leq i < M_1, 0 \leq j < M_2\} \quad (1)$$

또한 워터마크 시퀀스의 이웃 데이터와 상관성을 제거하기 위해 선형 피드백 시프트 레지스터(LFSR)를 이용하여 시퀀스를 재배열한다.

Step2. 오디오 신호 분할

원본 오디오 신호를 식(4)와 같이 2 개의 부분으로 나누어 워터마크를 삽입하게 되며 이때 A_e 를 워터마크로 사용되는 이진 이미지의 사이즈와 같은 개수의 세그먼트로 나눈다.

$$A = A_e + A_r \quad (2)$$

$$A_e = \{A_e(k), 0 \leq k < (M_1 * M_2)\} \quad (3)$$

Step3. DWT(Discrete Wavelet Transform) 변환

다른 시간 주파수 변환 방법과 비교하여 DWT 를 사용한 워터마킹은 매우 강인하다. 본 논문에서는 Daubechies D4 필터 계수를 이용하여 FIR(Finite Impulse Response) 필터 구조로 설계하였다.

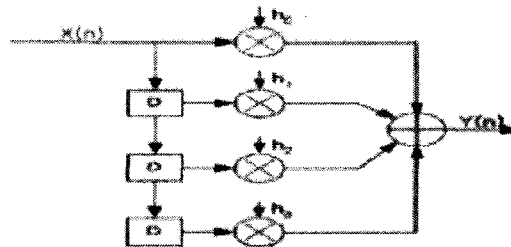


그림 2. DWT 변환을 위한 4Tap FIR 필터 구조[7]

Step4. 워터마크 삽입 기법

1) DWT 의 다해상도 특성에 따라 $A_e(k)$ 를 L 번째 레벨로 분해하기 위한 적절한 웨이브릿 패밀리를 선택한다. 본 논문에서는 Daubechies D4 필터계수를 이용하였다. 웨이브릿

계수들은 $A_{\frac{L}{k}}, D_{\frac{L}{k}}, D_{\frac{L}{k}-1}, D_{\frac{L}{k}-2}, \dots, D_{\frac{L}{k}-L+1}$ 이다. 여기서 $A_{\frac{L}{k}}$ 은 L 번째 레벨 근사

파트이고 $D_{\frac{L}{k}}, D_{\frac{L}{k}-1}, D_{\frac{L}{k}-2}, \dots, D_{\frac{L}{k}-L+1}$ 은 1 번째부터 L 번째까지의 디테일 파트이다.

2) 워터마크된 신호가 들리지 않게 하기 위해서는 HAS 의 마스킹 효과를 이용하여 오디오 신호의 가장 높은 에너지를 갖는 고주파수 부분에 삽입하여야 한다. 따라서 워터마크를

삽입하기 위해 각 세그먼트 $D^L(k)$ 의 가장 큰 계수의 절대값을 선택한다. 워터마크 삽입 공식은 아래 식(4)와 같이 표현된다.

$$D^L \begin{cases} D^L_i(k) + v(k), & D^L_i(k) = \max_j (D^L_j(k)) \\ D^L_i(k), & \text{otherwise} \end{cases} \quad (\text{식 4})$$

워터마크 $v(k)$ 는 k 번째 세그먼트 신호에 삽입된다. 따라서 워터마크된 신호의 L 레벨 디테일 파트 $D^L(k)$ 를 얻을 수 있다.

3) 워터마크가 삽입된 웨이브릿 계수들 $A^L(k), D^L(k), D^{L-1}(k), \dots, L^1(k)$ 을 이용하여 IDWT에 의해 오디오 신호 세그먼트들을 복원할 수 있다. 복원된 워터마크 신호는 Ae 이다.

4) Ae 를 Ae 로 교체하여 식(4)에 의해 워터마크된 오디오 신호를 얻을 수 있다.

워터마크 검출을 위해서는 원본 오디오 신호를 필요로 한다. 워터마크된 신호의 입력을 식(4)와 같이 $M1 * M2$ (워터마크 이미지의 사이즈)개의 세그먼트로 분할하여 웨이브릿 변환을 수행한 후 삽입영역의 계수와 원본 신호의 웨이브릿 계수와의 차를 이용하여 다음 식(10)과 같이 워터마크를 검출한다.

$$v(k) = D^L_i(k) - D^L_i(i) \quad (5)$$

여기서 $D^L_i(k)$ 은 워터마크된 신호의 L 레벨 디테일 파트이고 $D^L_i(i)$ 은 원본 신호의 웨이브릿 계수이다.

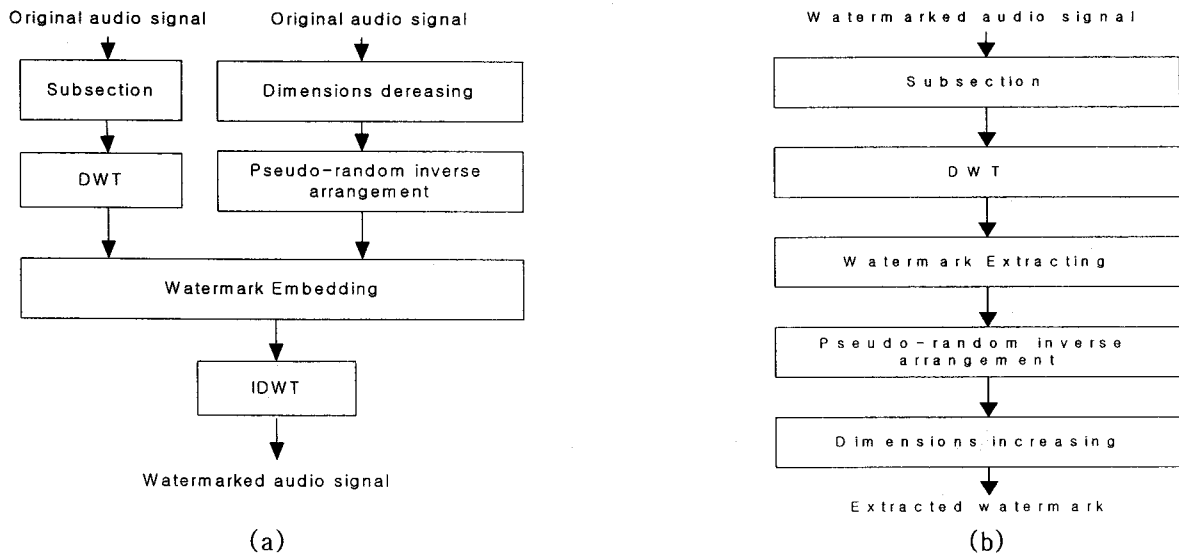


그림 3. 워터마크 삽입 알고리즘(a) 및 검출 알고리즘(b)

4. 제안하는 워터마킹 프로세서 구조의 설계

웨이브릿 변환을 이용한 워터마크 삽입 알고리즘을 수행하는 H/W 의 블록다이어그램은 다음과 같다.

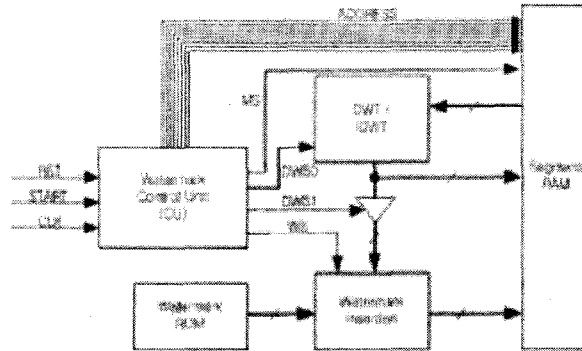


그림 4. 제안하는 워터마킹 프로세서 구조

Segment RAM 은 입력된 원본 신호를 2.1 의 Step2 에 따라 워터마크 사이즈만큼의 세그먼트로 분할하여 저장하는 세그먼트 메모리이다. DWT/IDWT 모듈은 세그먼트 메모리로부터 각 세그먼트 단위로 샘플들을 입력받아 3 레벨 이산 웨이브릿 변환과 워터마크가 삽입된 후 복원을 위해 역 웨이브릿 변환을 수행한다. 2.1 의 Step1 에 의해 전처리된 워터마크는 WM_ROM 에 저장되어 있으며 Watermark Insertion 모듈에 의해 DWT 모듈에 의해 수행된 원본신호의 웨이브릿 계수에 삽입된다. Watermark Control Unit(CU)는 전체 워터마크 삽입 과정을 제어하며 각 서브 모듈에 대한 제어 신호와 Segment RAM 접근을 필요로 하는 연산에 대해 적절한 어드레스를 발생 시킨다.

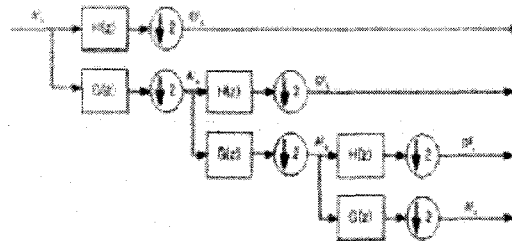


그림 5. 3단계 DWT 연산 구조

5. 실험 결과

제안한 워터마크 알고리즘의 음질 실험과 강인성 실험을 위해 44.1Khz 의 샘플링 주파수와 16 비트 양자화 비트수를 갖는 Mono 신호에 대해 16384 개의 샘플들을 취하여 MATLAB 을 이용하여 실험하였다. 또한 그림 7(a)와 같이 32*32 사이즈의 이진이미지를 워터마크로 사용하였다.

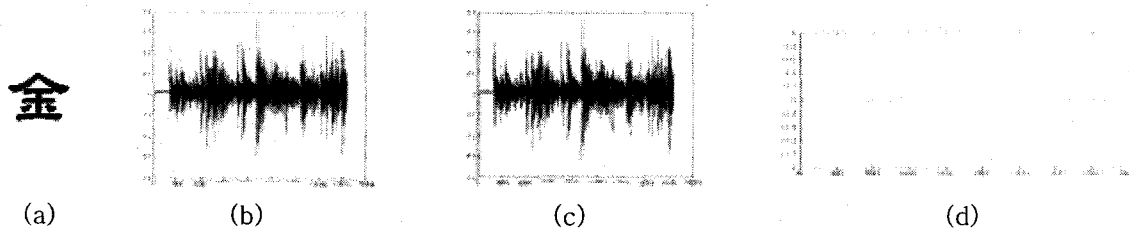


그림 6. 워터마크(a)와 원본신호(b), 워터마크된 결과 신호(c), (b)와 (c)의 차분파형(d)

본 논문에서 제안한 H/W 구조는 Verilog HDL 을 이용하여 설계하였으며 합성틀은 Synplify7.0 을 이용하였다.

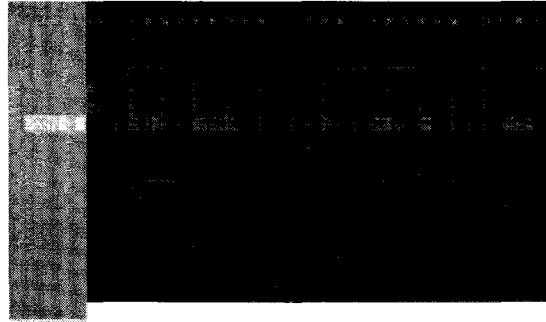


그림 7. DWT 모듈의 Modelsim 시뮬레이션 결과

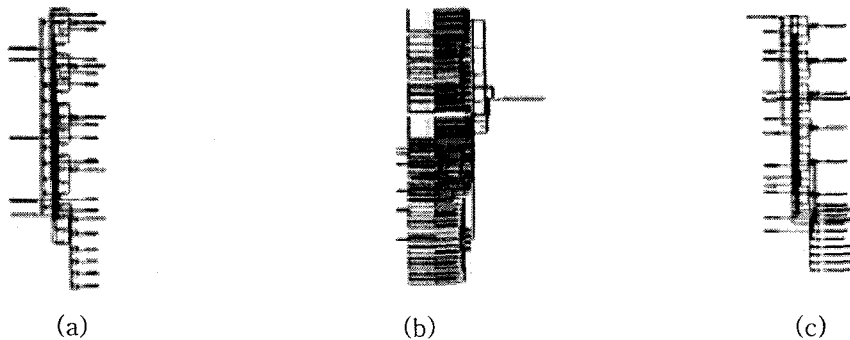


그림 8. 각 기능블럭의 합성 결과

6. 결론

그림 7(c)의 워터마크된 신호에 대한 잡음비는 약 42.0264dB 로서 원음과 비교하여 듣기를 하였을 경우 거의 구분이 되지 않는다. 또한 워터마크의 강인성 실험은 워터마크가 삽입된 신호에 대하여 샘플링율을 높이거나 낮추는 공격과 MP3 압축공격을 실험하였으며 앞의 2 가지 공격에 대하여는 매우 강인한 특성을 보였으나 압축공격에 대하여는 검출된 워터마크를 시각적으로 거의 판별하기 어려울 정도로 취약하였다.

이 같은 문제점을 해결하기 위해서는 압축공격에도 강인한 알고리즘뿐만 아니라 이를 구현할 H/W 구조에 관한 연구가 필요하다.

참고문헌

- [1] Xueyao Li "A New Adaptive Audio Watermarking Algorithm" Proceedings of the 5th. World Congress on Intelligent Control and Automation, June 15-19, 2004
- [2] M. Swanson, B. Zhu, A. Tewfik, and L. Boney, Robust audio watermarking using Perceptual masking, Signal Processing, Vol 66.
- [3] MPEG. Coding of moving pictures and associated audio for digital storage media at up to 1.5 Mbit/s, part 3: Audio, International Standard IS 11172-3, ISO/IEC JTC1/SC29 WG11, 1992.

한국반도체및디스플레이장비학회 2005년도 춘계학술대회 논문집

[4] 이강현, “고음질 오디오를 위한 워터마크 프로세서의 설계,” 2000 년도 대학기초사업 최종 연구개발 결과보고서, 2001

[5] MPEG. Information technology-Coding of audio-visual objects - Part3: Audio, International Standard IS 14496-3, ISO/IEC JTC1/SC29 WG11,1999.

[6] 정남훈, “VLSI를 이용한 MPEG-2오디오 부호화기 설계,” 연세대학교 대학원 석사학위 논문 집, 1996.

[7] Michael D. Ciletti, “Advanced Digital Design with the Verilog HDL,” Prentice Hall, 2003.