

오디오 워터마크를 이용한 방송 동기화 시스템 구현*

신동환, 신승원, 안종근, 신영일, 김종원, 최중욱
 ㈜마크애니 연구소
 서울 중구 쌍림동 151-11 쌍림빌딩 10층

Implementation of Synchronizing Broadcasting System Using Audio Watermark

Donghwan Shin, Seungwon Shin, Chongkeun Ahn,
 Keaton Shine, Jongweon Kim, and Jonguk Choi
 MarkAny Research Institute
 E-mail: dhshin@markany.com

요약

본 논문에서는 오디오 워터마킹 기술을 방송 동기화 시스템에 적용하여 방송시스템을 자동화 하기 위한 시스템을 구현했다. 방송시스템은 대개의 경우 1개의 중앙방송과 다수의 지역방송국으로 구성된다. 방송동기화 시스템은 방송소재가 중앙방송에서 지방방송으로 전환하는 시점을 알리는 신호에 워터마크를 삽입하여 방송하고 다수의 지방 방송에서는 워터마크를 검출해서 방송소재의 전환을 자동으로 하고자 한다.

본 논문에서 구현한 방송동기화 시스템은 방송시스템의 특성을 중시하여 설계 제작되었다. 방송시스템의 오류는 바로 방송사고로 이어지는 만큼 동기신호를 잘못 검지하는 오류는 없어야 한다. 즉 워터마크가 없는 오디오에서 워터마크가 있다고 검출되는 경우가 0% 이어야 한다. 본 논문에서 제안된 알고리즘으로 실제 방송되는 라디오 신호를 갖고 실험해본 결과 위의 조건을 만족하고 다양한 환경 즉 압축 및 A/D, D/A, 노이즈 삽입에 대한 신뢰성 테스트를 수행하여 우수한 결과를 얻었다.

1. 서론

방송시스템의 구성은 대개의 경우 1개의 중앙방송국과 다수의 지역방송국으로 구성된다. 지역방송국의 경우 주로 중앙방송국의 방송소재를 방송하고 필요한 경우에 지역방송에서 자체 제작한 방송소재를 중간중간 삽입하여 방송하고 있다. 여기서 방송소재의 방송 전환(중앙->지역, 지역->중앙)이 필요하게 됐다. 이런 동작은 여러 지역방송이 동시에 이루어져야 한다는 의미에서 방송 동기화라고 부르고자 한다. 방송 동기

화를 자동으로 하기위한 여러 방법들이 제시되어 왔다. 그림 1-1 은 방송동기화 시스템이 사용된 방송시스템을 나타낸다.

방송이 전환되는 시점에 특정 오디오(예: “문화방송입니다.”, “MBC FM” 등)를 방송하고 이 특정오디오를 음성인식기술을 이용하여 방송전환시점을 결정하게 된다. 이 방법의 단점은 아직 음성인식기술이 100%의 인식률을 갖지 못하여 오검지의 확률이 높다는 문제점을 안고있다. 즉 방송도중 방송전환 오디오와 같은 내용이 있을 경우 이를 전환신호로 잘못 인식 할 경우가 있다.

또 다른 시도는 전용망을 이용한 동기화 신호 전송이다. 이 방법도 전용망의 전송상태에 따라서 지연이 되는 경우가 있고 이 경우 방송사고로 이어질 수 있다

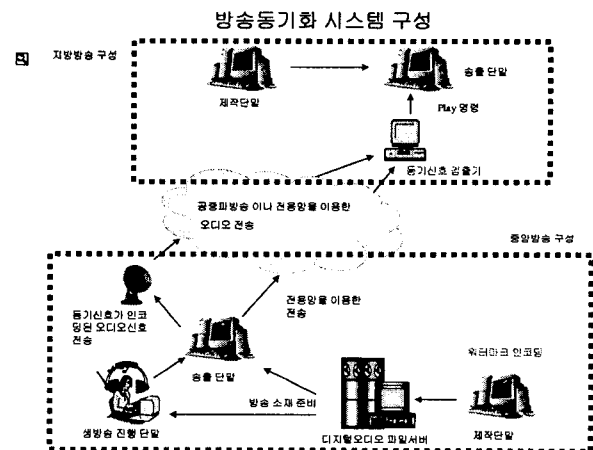


그림 1-1. 방송동기화 시스템 구성

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

본 논문은 오디오 워터마킹 기술을 방송장비의 동기화 시스템에 적용한 것이다. 오디오 워터마킹 기술을 방송에 이용하려는 시도는 여러 차례 있어왔다. 방송 내용을 모니터링하기 위한 방법으로 1997년 음악에 워터마크를 삽입하여 배포하는 기술이 MusiCode라는 이름으로 제안되었다. 또 VIVA(Visual Identity Verification Auditor) 프로젝트가 있다[1][2]. 이 프로젝트에선 주로 비디오 영상에 워터마킹 기술을 적용하여 방송을 모니터링하고자 했다.

본 논문에서는 오디오 워터마킹 기술을 방송 동기화 시스템에 적용하여 방송시스템을 자동화하기 위한 것이다. 방송시스템의 특성상 동기신호를 잘못 검지하는 오류는 없어야 한다. 즉 워터마크가 없는 오디오에서 워터마크가 있다고 검출되는 경우의 에러(false positive error)가 0% 이어야 한다.

본 논문에서 제안된 알고리즘은 오디오 워터마킹 알고리즘으로서 디지털 필터링을 이용해서 원본 오디오와 모든 면에서 동일하지만, 원본 오디오는 갖지 않는 스펙트럼 특징을 이용하여 정보를 삽입, 추출하는 기술을 이용하였다[3]. 구현된 시스템 특성상 5bit의 정보를 0.5초 이내에 검출할 수 있는 알고리즘을 이용하여 실시간으로 워터마크를 찾아내고 이를 이용하여 시스템 동기신호를 생성한다.

논문의 구성은 사용된 오디오 워터마킹 알고리즘에 대한 소개가 2장에 소개된다. 3장에선 알고리즘을 시스템으로 구현한 방법을 소개한다. 4장에서는 사용된 워터마크 알고리즘의 강인성을 시험한다. 실제 방송시스템에서 문제가 될 수 있는 잡음문제와 압축에 따른 워터마크 검지율을 비교한다.

2. 워터마킹 알고리즘

본 논문에서 사용된 오디오 워터마킹 알고리즘은 디지털 필터링을 이용하여 원본 오디오는 갖지 않는 스펙트럼 특성으로 워터마크를 만들어 삽입하고 추출하는 기술이다.

2.1 워터마크 삽입 절차

오디오 신호에 정보를 삽입하기 위한 모듈은 그림 2-1과 같다. 여기서, Cover Signal은 정보를 삽입하지 않은 원본 오디오 신호를, Stego Signal은 Cover Signal에 워터마크 정보가 삽입된 신호를 의미한다.

워터마킹을 삽입하기 위한 과정을 살펴 보면, 첫 번째 단계로 입력된 Cover Signal을 웨이블릿 변환을 통해서 필터링할 대상 대역 주파수와 필터링되지 않는 대역 주파수로 분해한다. 필터링 단계에서 cover signal을 웨이블릿으로 분해하여 대상 주파수만을 필터링하는 것은 디지털 필터링 과정에서 주변 주파수 성분에 불필요한 변화를 주지 않기 위함이다. 이상적인 디지털 필터를 설계하고 구현하는 것은 사실상 불가능하므로 정교한 필터를 설계하더라도, 필터링 과정에서 주

변 신호의 주파수에 영향을 미치는 것을 막을 수는 없다. 특히, 고주파 성분은 상대적으로 적은 에너지를 갖고 있는데, 이 부분에 발생하는 변화는 고주파를 왜곡시켜 오디오 전문가들이 느끼는 음색의 변화를 발생시킬 가능성이 높다. 이와 같은 왜곡 현상을 사전에 막기 위해서 시간-주파수 영역에서 신호의 분해능이 뛰어난 웨이블릿[5]을 이용한다.

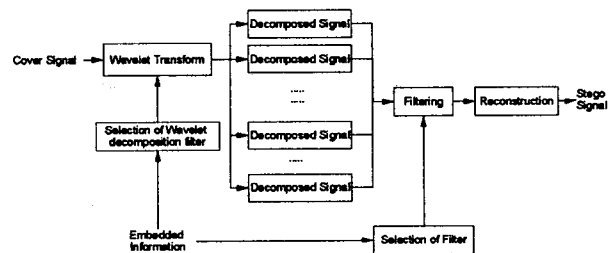


그림 2-1. 워터마크 삽입 구조도

두 번째 단계로 대상 주파수 신호를 미리 설계된 필터를 이용해서 필터링한다. 이때 여기서 설계된 디지털 필터가 IIR(infinite impulse response) 필터이기 때문에 비선형 위상 변화가 발생하고, 비선형 위상 변화는 음질 변화를 야기시킬 우려가 있다. 또, 이런 필터링 과정을 반복하게 되면 위상의 불규칙한 변화는 왜곡된 음을 만들기도 한다. 이런 위상 변화에 대한 문제를 해결하기 위하여 우리는 필터링을 거치는 동안 위상 변화가 발생하지 않도록 Zero-phase filtering[4,6]을 이용하였다. 이는 Fredrik[6]이 IIR 필터링을 전방향, 후방향으로 실시함으로써 위상이 변하지 않는 방식과 필터링의 초기값을 계산하는 방법을 이용하였다.

마지막으로, 웨이블릿으로 분해된 각각의 신호들과 필터링된 신호를 다시 합성하여 원래의 오디오 신호로 변환하면 정보가 삽입된 Stego signal이 만들어진 것이다.

본 논문에서 이용되는 디지털 필터는 사전에 정보량과 저작물의 특성에 따라 디지털 필터를 설계하여 이용한다. 구현하고자 하는 시스템은 방송시스템에 사용된다는 점과 워터마크가 삽입되는 신호가 주로 음성이라는 점을 고려하여 표 2-1과 같은 특징을 갖는 대역저지 디지털필터를 이용했다.

표 2-1. 디지털 필터(kHz)

No.	통과대역 (Passband)	차단대역 (Stopband)	통과대역 변화 (dB)	차단대역 완화율 (dB)
1	2.85-3.15	2.95-3.05	0.1	74
2	3.30-3.60	3.40-3.50	0.1	75
3	4.00-4.30	4.10-4.20	0.1	77
4	4.60-4.90	4.70-4.80	0.1	78

5	5.30-5.60	5.40-5.50	0.1	80
---	-----------	-----------	-----	----

본 논문에서는 시스템특성을 고려하여 강인성(robustness)을 중시하고 삽입대상이 오디오임을 고려해 워터마크 삽입영역을 3~6kHz 대역으로 했다. 필터링은 식(1)에 의해서 원하는 신호만을 추출했다.

$$y_k = \sum_{m=0}^M a_m x_{k-m} + \sum_{n=1}^N b_n y_{k-n} \quad \text{식(1)}$$

이 것의 주파수 특성은 식(2)를 이용하여 분석할 수 있다[4].

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{m=0}^M a_m z^{-m}}{1 - \sum_{n=1}^N b_n z^{-n}} \quad \text{식 (2)}$$

$$|H(e^{j\omega T})| = \frac{1}{\sqrt{\{H_R(e^{j\omega T})\}^2 + \{H_I(e^{j\omega T})\}^2}}$$

2.2 워터마크 추출 절차

워터마크 추출과정은 오디오의 진폭스펙트럼을 분석하여 찾아낼 수 있다. 그림2-2는 워터마크 삽입 과정을 보여준다. 워터마크 정보가 삽입된 Stego signal 의 샘플링 주파수, 삽입정보의 대역폭 등을 고려하여 1024를 주파수 분석길이로 하였다.

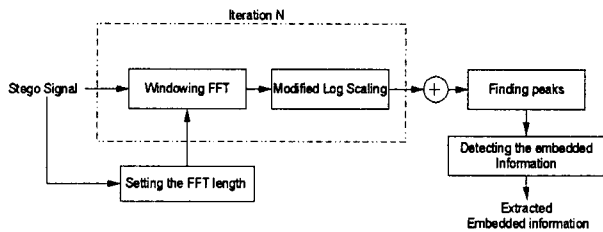


그림 2-2. 워터마크 추출 구조도

워터마크 신호를 제외한 나머지 신호의 주파수별 신호발생 빈도가 랜덤하다고 가정하면 반복하여 더하면 워터마크 신호는 강해지고 그 외의 신호는 약해지는 성질을 이용한다. 본 논문에서는 반복횟수 N 을 85회로 했다. 최대값이 제한된 수정된 로그스케일링을 이용하여 워터마크 정보를 강조한다. 이를 바탕으로 피크위치를 식별하고 식별된 위치정보를 이용하여 워터마크 삽입 유무를 판단한다.

3. 시스템 구현

본 논문에서는 시스템의 안정성 및 빠르고 정확한 시스템구현을 위해서 전용 하드웨어를 설계 구현

하였다. 방송동기화 시스템은 특성상 빠른 시간 안에 동기신호를 감지해서 감지된 동기신호를 제어신호 발생장치로 전송해야 한다. 적어도 0.5초내에 동기신호를 검출 가능해야 한다. 이와 같은 조건을 만족하기 위해서 PC 기반의 시스템을 구현할 수도 있지만 PC 특성상 안정화되지 못하고 실시간 처리를 보장하기 어렵다. 따라서 본 연구에서는 연산속도가 빠르고 안정된 DSP(digital signal processor)칩을 사용하여 시스템을 구현하였다.

본 연구에서 사용된 DSP 칩은 TI(Texas Instrument)의 TMS320VC5410-100MIPS(million instruction per second)을 사용했다. 그림3-1에 구현된 시스템 하드웨어 블럭도를 나타냈다.

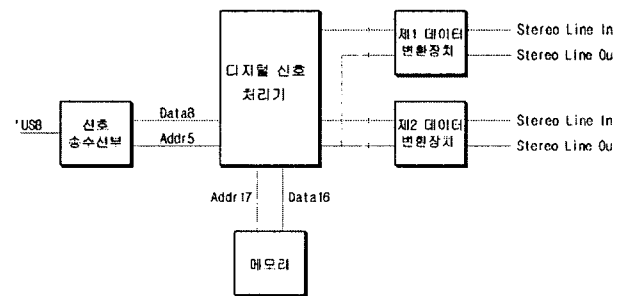


그림 3-1. 하드웨어 블럭도

시스템에서 하드웨어 변환장치는 A/D, D/A 변환기를 말하는 것으로 입력되는 아날로그 스테레오 오디오 신호를 44.1kHz 스테레오 16bit ADPCM 형태의 디지털 신호로 변환하고 동시에 디지털 신호 처리기에서 보내온 디지털 스테레오 오디오 신호를 아날로그 스테레오 오디오 신호로 변환하는 장치이다. 본 논문에서는 입력오디오의 loop back 신호를 생성하는데 사용했다. 즉 입력되는 오디오 신호를 디지털 신호로 바꾸고 다시 아날로그 출력으로 내보냈다. 이를 통하여 입력신호의 상태를 모니터링 할 수 있다.

메모리부는 128Kwords 플래시메모리로 구성되어 있다. 위 시스템은 2개의 방송신호를 동시에 처리할 수 있다. 한 방송 채널당 최대 40MIPS 까지 연산량을 소모한다.

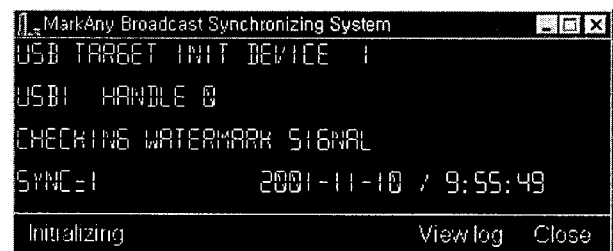


그림3-2. 모니터링 프로그램 실행화면

디지털 신호 처리기로 입력된 디지털 방송 오디오신호를 워터마크 검지 알고리즘을 이용하여 실시간으로 검색한다. 검색 중에 동기신호로써 삽입된 워터

마크 정보를 추출하게 되면 이 결과를 USB 포트를 이용해 모니터링 PC 에게 정보를 알려준다. 이때 모니터링 PC 에 전달되는 정보는 동기신호가 검출된 시각과 동기신호가 검출된 누적 횟수를 log 파일로 저장하게 된다. 그림 3-2는 PC 상에서 동작하는 모니터링 프로그램 실행화면을 나타냈다. 동기신호가 발생한 시각을 알려주고 있다.

4. 실험 결과 및 분석

본 논문에서 구현된 알고리즘의 강인성을 시험하기 위해서 2가지 방법으로 실험을 했다. 실제 방송상태와 유사한 제약조건을 만들기 위해서 잡음과 압축레벨을 변화시키면서 실험을 했다.

본 논문에서 사용된 워터마킹된 오디오는 0.7 초 길이 오디오 신호로 신호 뒤에는 0.5 초 이상의 묵음 구간이 이어진다. 이 묵음은 동기화 신호임을 알려주기 위한 신호이고 방송전환에 따른 잡음을 제거하기 위한 부분이다.

그림4-1에 본 실험에서 사용된 음성을 나타냈다. 남성 아나운서 음성으로 가로축은 시간(단위:초)을 나타내고 세로축은 파형(단위 dB)의 크기를 나타낸다.

-inf 는 완전한 묵음상태를 나타내고 0dB 를 16비트로 나타낼 수 있는 최대 진폭을 나타내고, 최소로 표현 가능한 크기는 -90.3dB 를 기준으로 한다.

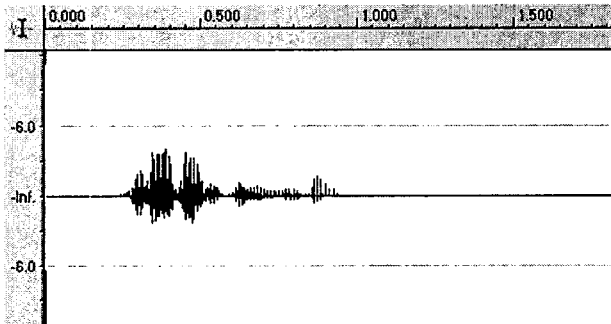


그림4-1. 사용된 음성

4.1 음질 검사(Listening Test)

수치적인 방법으로 음질을 평가할 수 있는 방법 가운데서 가장 널리 이용되고 있는 계산법은 신호 대 잡음 비율 (SNR, Signal to Noise Ratio)을 구하는 방법이 있다. SNR이란, 신호는 단독으로 존재하지 않고 대개 잡음과 섞여 있는데, 그 비율을 나타내는 척도로서 SNR이 쓰인다. 신호전력을 Signal, 잡음전력을 Noise이라 할 때 수식 (3)과 같이 나타내며, 단위는 데시벨 (dB)이다.

$$SNR = 20 \left(\log_{10} \frac{Signal}{Noise} \right) \text{ dB} \quad \text{식 (3)}$$

오디오의 경우에는 SNR에 많은 의미를 부여하지는 않으나, 일반적으로 60dB이상이 되면 거의 잡음을

감지할 수 없는 것으로 알려져 있다. (<표 4-1> 참조).

표 4-1. SNR 에 따른 품질[7,8]

S/N ratio dB	S/N ratio:1	Audio quality
60 dB	1,000	Excellent, no noise apparent
50 dB	316	Good, a small amount of noise but audio quality good.
40dB	100	Reasonable, fine grain or snow in the audio, some fine detail lost.
30 dB	32	Poor audio with a great deal of noise.
20 dB	10	Unusable audio.

본 논문에서 제작된 워터마크가 삽입된 오디오 신호의 경우 SNR을 측정한 결과 66.090dB로 워터마크 삽입으로 인한 오디오 신호 열화를 거의 느낄 수 없었다.

4.2 강인성 검사(Robustness Test)

검지율 검사를 위한 잡음 삽입 실험에서는 백색 잡음(white noise)을 -70dB, -60dB, -50dB 를 첨가했다.

압축률에 따른 검사를 위해서 MP3 압축으로 압축률은 128kbps, 96kbps 으로 압축 후 실험했다. 시스템 구성상 방송된 아날로그 신호를 디지털 신호로 변환하여 워터마크를 검출하기 때문에 기본적으로 D/A, A/D 변환에 의한 공격이 포함되어 있다.

실험결과를 표4-1에 표시했다. 실험은 워터마크가 삽입된 오디오를 100번 재생하여 워터마크를 검출한 결과를 표시했다.

표 4-1. 압축과 잡음첨가 공격후 워터마크 검출결과

잡음 강도	No Compression	128kbps	96kbps
-45dB	71	72	67
-50dB	97	99	96
-60dB	100	100	100
-70dB	100	100	100

실험결과를 살펴 보면 -60dB 잡음첨가까지는 미검지하는 경우가 발생하지 않았다. 그러나 잡음의 강도를 높여서 -50dB 이상의 강도로 첨가했을 때는 검지 못하는 경우가 증가하고 있다. 또 128kbps 의 압축률을 적용한 경우는 압축을 하지 않은 경우와 차이가 거의 없으나 96kbps 로 압축률을 높인 경우는 검지율이 떨어지고 있음을 알 수 있다.

본 논문에서 구현된 시스템의 신뢰성 검사를 위해서 실제 방송되는 라디오 신호를 갖고 실험해본 결과 240 시간 동안 false positive 에러가 발생하지 않았다.

5. 결론

본 논문에서는 워터마크를 이용하여 방송시스템의 방송전환시점을 자동으로 알려주는 방송동기화 시스템을 구현하였다. 특히 방송장비의 특성상 실시간으로 방송신호를 처리하여 0.5초 이내에 워터마크를 검출하여 방송전환시점을 알려주도록 설계 되었다. 또한 신뢰성 검사를 통하여 240시간 동안 false positive 오류검지가 없음을 확인했다.

본 논문에서 구현된 시스템은 잡음과 압축에 -60dB 이상의 백색잡음 공격과 압축에 대한 공격에도 강인함을 알 수 있었다. 또한 워터마크된 오디오의 음질 테스트 결과 원음과 정보가 삽입된 음을 거의 구별할 수 없는 것으로 분석되었다.

본 논문에서 사용한 워터마킹 기술을 발전시킬 경우 방송광고 모니터링 시스템이나 저작권료 지불을 위한 방송 모니터링 시스템에 응용 가능하다. 또한 라디오 수신기를 이용하여 다수의 시스템을 동시에 제어하는 장비에도 적용할 수 있다.

참고 문헌

- [1] T. Kalker, G. Depovere, J. Haitsma, M.J. Maes, "Video watermarking system for broadcast monitoring", Proceedings of the Society and Watermarking of Multimedia Contents, 103-112, 19990125.
- [2] De Strycker, Pascale Termont, et al, "The VIVA project: Digital watermarking for broadcast monitoring", Proceedings of the 1999 International Conference on Image Processing, volume 2, 202-205, 19990124
- [3] 신승원, 김종원, 최중욱, "디지털 오디오의 인터넷 방송 서비스를 위한 저작물 식별 코드 은닉 기술", 한국방송공학회 학술대회, 2001. 11.
- [4] Oppenheim, A.V., and R.W. Schaffer, *Discrete-Time Signal Processing*, Prentice-Hall, 1989, pp. 311-312
- [5] Yang Yan, Zhang Dong, "Enhancement of Angiograms via M-Band Wavelet Transform", Proceedings of SPIE Biomedical Photonics and Optoelectronic Imaging, 165-169, 2000.
- [6] Fredrik Gustafsson, "Determining the Initial States in Forward-Backward Filtering", Transactions on Signal Processing, V.44 N. 4, 1996. 4. 1.
- [7] M. Adinolfi et al., Nucl. Instrum. Methods A 329 (1993) 117.
- [8] The RD20 Collaboration, preprint CERN/DRDC 94-39 (Geneva,1994). Peter Weilhammer, private communication