기하학적 공격 및 비트율 변화 공격에 강한 MPEG-2 동영상 워터마크 시스템 개발

논 문
59P-3-3

A Robust Watermark System for MPEG-2 Video Against Geometric and Bitrate Varying Attacks

황 선 철†

(Seon-Cheol Hwang)

Abstract - This paper describes a watermarking system that has the robust characteristics against the geometric attacks of resizing and the bitrate varying attacks. While the existing watermark methods of marking on AC in DCT coefficients are weak against the geometric attacks, the methods of tiny increasing the DC in DCT coefficients and the system implementations are suggested in this paper. We measure the differences between the original images and the watermarked images and the detecting rate from the geometric and bitrate attacked images. Results of experiments show that our method is robust against that attacks.

Key Words : Watermarking, Geometric Attacks, Bitrate-Varying Attacks, AC and DC in DCT Coefficients

1. 서 론

오늘날 우리는 콘텐츠가 공업제품보다 더 부가가치가 높은 세상에 살고 있다. 한편의 잘 만들어진 콘텐츠가 전 세계적으로 한류를 만들어내고 있으며 한편의 영화 수출이 자동차 수 천대 수출을 능가하는 이익을 내고 있다.

현재 디지털 방송 전반과 인터넷 등과 같은 통신 매체에서 가장 많이 유통되고 있는 데이터는 영상 콘텐츠이다. 유통되는 영상 콘텐츠로서는 JPEG을 이용한 정지영상(사진등)과 MPEG으로 대변되는 동영상 콘텐츠가 주종을 이루고 있다.

이러한 디지털 콘텐츠들의 불법 유통을 근본적으로 해결하기 위해 도입된 기술이 디지털 워터마킹 기술이다[1]. 디지털 워터마킹 방식에는 영상에 직접 삽입하는 공간영역 방식과 주파수 영역 방식이 있는데 공간영역 방식은 공격에 매우 취약하여 주파수 영역 방식을 주로 사용하고 있다. 주파수 영역 방식에는 웨이블릿 변환을 이용한 방식[3]과 이산 코사인 변환을 이용한 방식이 있는데 현재 JPEG과 MPEG의 영상압축 방식은 이산 코사인 변환 방식이 사용되며, 본 논문에서는 이 방식을 사용하여 워터마킹하게 된다[4][5].

Cox에 의해 제안된 워터마킹 방식은 DCT 계수의 AC 영역에 대한 워터마킹방식인데 이 방식은 기하학적 공격에 취약한 양상을 보였다[10]. 그러므로 본 논문에서는 기하학적 공격에 강한 특성을 보이는 방식으로 DCT의 DC 영역에 워터마킹하는 방식을 채택하여 비가시성의 가능성과 견고성 및 명확성, 범용성 여부를 실험하고자 한다. 이를 위해 본 논문에서는 워터마킹 삽입 용이성을 실험하고 원영상과의 영상비교를 시행한 후 크기 변형(resizing) 및 비트율 변환에 의해 변형된 영상에서 워터마크 검출율을 실험하여 그 결과를 제시하고자 한다.

2. MPEG-2 워터마크 삽입기

2.1 DCT 계수 마킹 기술

DCT 계수에 대한 워터마킹은 일반적으로 Y 신호에 대해서만 이루어진다. 색차신호는 정밀도가 매우 떨어지며 양자화도 데이터를 대폭 줄이는 방향으로 진행되므로 데이터의 많은 부분을 잃게 된다. 그러므로 대부분의 알고리즘은 보다 정밀한 Y 신호의 DCT 계수에 대한 워터마킹을 수행하고 있다.

DCT는 8×8 영상신호에 대해 행하여져서 8×8 주파수 데이터를 만들어낸다. 이때 주파수 데이터는 DC 성분의 저주파수에서부터 고주파수 성분으로 분해되어 분포한다. DCT는 분해된 주파수 계수가 저주파수 성분은 매우 큰 값으로 나타나고 고주파수 성분은 매우 작은 값으로 나타난다는 특성을 이용하여 DCT 계수를 일정한 숫자로 나눗셈을 하게 되어 작은 값을 0으로 만들게 된다. 이때 생성된 0의 값을 갖는 데이터는 없는 데이터로 간주하게 되는데 이로써 데이터의 개수가 감소하는 효과를 얻게 되어 결과적으로 데이터의 압축 효과를 얻게 되는 것이다.

변환된 DCT 계수는 좌상에서 우하 방향으로 지그재그(zig-zag) 형태로 저수파~고주파로 분포하게 된다. 여기서 저주파수는 영상에 큰 영향을 미치게 되고 고주파 성분은 영상에 별 영향을 못 미치게 된다.

이 사실에 착안하여 디지털 워터마크가 주입되게 되는데 Cox는 논문에서 디지털 워터마크가 들어갈 수 있는 위치를 지정하였다[12]. 일반적인 영상에서 통계적으로 분석한 결과 워터마크를 행할 수 있는 위치와 행해서는 안되는 위치를 지정한 것이다. 반면 Koch나 Zhao에 의해서 다른 위치에

---

† 교신저자, 정회원 : 인덕대학 방송영상미디어학과 부교수·공박
E-mail: sthwang@induk.ac.kr

접수일자 : 2010년 6월 15일
최종완료 : 2010년 7월 30일

대한 논문이 발표되었는데 Cox의 위치보다 줄어든 위치를 지정하고 있다[13].

그러나 본 논문에서는 저주파 영역 위치에 상관없이 DC 값의 크기를 일정한 레벨 +/- 변형시키는 방식을 도입하였다. 이는 다음과 같은 이유에서 도입되었다.

1) DC 레벨의 변형은 전체 영상 휘도의 ±1 이내의 증감 변화를 가져온다.
2) 8×8 블록 중 ±1 변화 픽셀은 대개 그 수가 40~80% 이내의 변화를 가져온다.

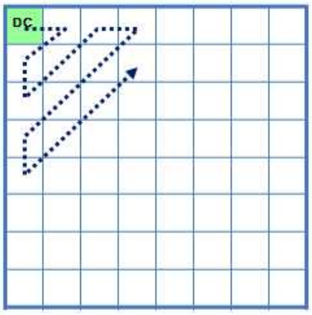

그림 1 본 연구에서 제안하는 워터마킹 위치
Fig. 1 Watermark position suggested in this paper

본 연구에서는 Cox 또는 Koch 등이 제안하는 방식과 다른 방식으로 워터마킹을 삽입한다. Cox 등이 제안하는 방식은 DCT 계수를 교체한 후 다시 인코딩을 수행해야하지만 본 연구에서 제안하는 방식은 전혀 재인코딩 없이 워터마킹을 수행하므로 공격자들이 위치를 발견할 수 없고, 또한 변환 흔적이 없으므로 워터마킹 여부를 알아내기 어려운 방식이다. 또한 많은 수의 워터마킹이 존재하므로 워터마크 감지 성공율이 현저히 향상되는 효과도 얻을 수 있었다.

2.2 워터마크 삽입기 설계

본 과제에서 개발된 HDTV 셋톱용 워터마크 삽입기는 다음 그림2와 같다. 영상콘텐츠가 DRM (Digital Right Management)에 의해 사용자에게 안전하게 전달된 다음 사용자는 이를 셋톱박스(STB)를 통해 재현하게 된다. 이때 STB에서는 사용자의 정보를 MPEG-2 데이터에 삽입하게 된다.

MPEG-2에서 워터마킹이 삽입되는 부분은 I-타입의 프레임에서만 가능하다. 이유는 P-타입이나 B-타입은 두 장면 혹은 세 장면 사이의 차이 나는 부분만을 압축하는 반면 I-타입의 영상은 화면 전체를 압축하는 알고리즘을 사용하기 때문이다.

본 과제에서는 MPEG-2 디코더를 개선하여 워터마크 삽입기를 탑재하였다. 원하는 정보에 대해 암호화를 수행하고 8×8 영상에 대한 디코딩에서 원하는 위치의 슬라이스(slice)에 대해 DC 계수를 원하는 레벨만큼 증감시킨다. 이때 증감폭은 +/- 1로 하는데 이유는 DC 부분이 영상에 민감한 영향을 미치기 때문이다. 본 논문에서는 DC 계수의 범위를 면밀히 조사하여 영상에 미치는 영향을 최소화하는 증가폭을 선택하여 적용하였다.

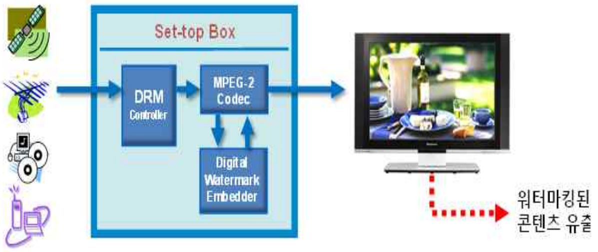

그림 2 HDTV STB용 워터마크 삽입기의 설계
Fig. 2 Design of Watermark embedder for HDTV STB

3. 실험 및 고찰

본 연구에서는 워터마킹 삽입 및 검출 실험을 위해 공중파 방송을 TS(Transport Stream) 형태로 저장한 다음 비디오 스트림만 분리하여 *.m2v'의 형태로 저장한 다음 실험을 수행하였다. 실험의 종류는 (1) 워터마크 삽입 실험, (2) 공격받지 않은 HD 동영상에서 워터마크 검출 실험, (3) 비트율 변환 공격 후 워터마크 검출 실험, (4) 해상도 저하 공격 후 워터마크 검출 실험을 실시하고 결과를 고찰하였다.

3.1 워터마크 삽입 실험

본 연구에서는 DCT 계수 중 DC 계수에 대해 1 만큼 증가시키는 실험을 실시하였다. 워터마크 삽입 실험은 차후 검출 실험으로 삽입 여부를 알 수 있겠으나 본 실험에서는 삽입 시간에 대한 실험 결과를 도출하고자 한다.

실험을 위해서 파일을 읽어서 저장하는 실험을 실시하였다. 실제 시스템에서는 실시간으로 삽입되지만 연구실 실험에서는 CPU의 연산 속도에 의한 삽입 실험만 가능하다. 본 실험에서는 워터마크를 수행하지 않았을 경우 시작 타이밍과 종료 타이밍을 mSec 단위로 기록한 후 워터마크 삽입시의 타이밍과 비교하는 실험을 실시하였다. 실험 결과는 4492 프레임으로 구성된 동영상을 10회에 걸쳐 실험하였으며 149.733 sec에 대해 삽입하는 동안 평균 0.091sec 가량 증가한 것으로 나타나 연산 시간이 0.06% 만 증가한 결과를 얻을 수 있었으며 이는 워터마킹 삽입이 전체 연산에 거의 영향을 미치지 않는 것으로 나타났다.

3.2 워터마크 검출 실험

본 연구에서는 세 가지 방법으로 검출 실험을 실시하였다. 첫 번째는 공격을 받지 않은 상태에서의 검출이고 두 번째는 비트율 저하 공격, 세 번째는 캡쳐 공격으로 1920×1080의 HD 영상을 720×480으로 아날로그 화면 캡처 공격의 경우이다.

3.2.1 공격 없는 HDTV 영상에 대한 검출 실험

공격을 받지 않은 원 상태에서의 검출 실험이다. 이 실험에서 워터마킹된 영상과 원영상 간의 차분을 하면 워터마킹된 영상이 +1 혹은 -1의 증감된 것이 검출된다. 이것을 눈으로 볼 수 있도록 약 100배를 곱하면 차이가 없는 부분은 0을, 차이가 나는 부분은 약 100의 값을 나타내며 결과 영상은 그림 5와 같다.

그림 3 원본 영상 (HDTV)
Fig. 3 Original Image (HDTV)

그림 4 워터마크 삽입 영상 (HDTV)
Fig. 4 Watermarked Image (HDTV)

그림 5 차분 영상 (원영상 - 워터마크영상)
Fig. 5 Differential Image (Original - Watermarked)

원본 영상에 대한 워터마크 삽입 및 검출 실험은 정확하게 결과를 얻을 수 있다. 이 경우의 검출율은 모든 경우에서 100%을 얻을 수 있다. 다음 그림 6과 그림 7은 DC 값을 +1 증가시켰을 때와 +3 증가시켰을 때의 비교 결과이다. 결과에서 볼 수 있듯이 픽셀 단위로 확대해서 보았을 때 증가값은 1차이로 대부분의 경우 동일하였으나 증가 픽셀의 밀도가 차이를 보이는 것으로 나타났다. 본 연구에서는 +1과 +2, +3의 경우만 실시하였으며 DC 값의 범위에 따라 +2를 사용하는 경우가 있었으나 대개의 경우 +1만을 사용하도록 제한하였다.

그림 6 DC+1의 경우
Fig. 6 Result image when DC+1

그림 7 DC+3의 경우
Fig. 7 Result image when DC+3

3.2.2 비트율 저하 공격 영상에 대한 검출 결과

비트율 저하 공격은 기존의 HD영상의 비트율을 낮추어 재구성하는 공격을 말한다. 이 경우는 HD영상의 파일이 커서 저장이 어려울 때와 전송로 상황이 좋지 않아 비트율을 낮추어 전송해야 할 경우 주로 사용되는데 DCT 계수의 양자화수치를 높이면 DCT 계수의 발생량이 대폭 줄어들게 되지만 화질은 저하된다. 하지만 디지털 데이터의 범주에서의 작업이므로 눈으로 보는 화질이 일정 수준 떨어지는 것 이외에는 시청에 거부감은 적은 편이다. 비트율 저하 공격에 의해 훼손된 영상에 대한 워터마킹 검출은 다음 그림과 같다.

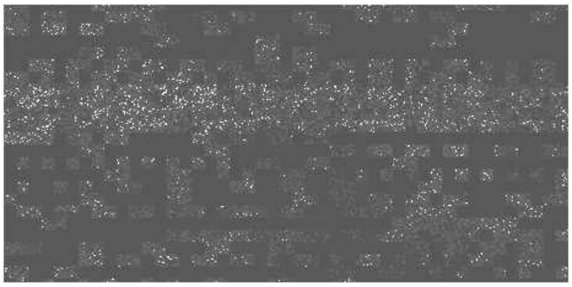

그림 8 비트율 저하 공격의 경우 검출 결과
Fig. 8 Result of bitrate-down attack

결과에서 볼 수 있듯이 비트율 저하 공격을 하면 복잡한 부분의 매크로블록의 데이터가 변하여 원하는 데이터를 얻을 수 없게 된다. 하지만 변화되는 정도가 인근 영역에서 동일하게 적용되므로 본 연구에서 개발한 검출 방식에 의해 검출할 경우 약 99% 이상의 검출율을 얻을 수 있었다.

3.2.3 캡쳐 공격 영상에 대한 검출 결과

캡쳐 공격은 가장 손실이 큰 공격이다. 이 공격은 동영상을 아날로그신호 형태로 캡쳐 보드를 통해 저장하기 때문에 원본의 성질은 모두 버리게 되고 오로지 캡쳐 보드의 특성을 따르는 화질로 바뀌게 된다. 이렇게 되면 대개의 경우 블러링(흐림효과)이 발생하고 영상이 흐려지며 색감도 바뀌게 된다. 또한 캡쳐보드가 전혀 새로운 시점에서 MPEG-2 압축을 실시하게 되어 DCT 계수의 구성도 완전히 새롭게 수행된다. 그림 9는 원영상과 캡쳐공격된 영상의 차분에 대한 영상으로 차이 값이 매우 커서 100배를 곱하지 않아도 차분 에러영상이 가시적이다.

하지만 본 연구에서 개발한 검출 방식인 확률적 매칭 방식(statistic matching method)을 사용하여 검출한 결과 그림 10 및 그림 11과 같은 결과를 얻을 수 있었다. 결과는 프로

그램 내부에서 처리되므로 볼 수 없지만 데이터를 추출하여 '마이크로소프트사의 Excel' 프로그램의 도움으로 결과를 출력하였다. 결과에서 보는 바와 같이 워터마킹이 삽입되지 않은 부분과 삽입된 부분이 현저하게 차이가 나는 것을 볼 수 있다. 본 실험에서 워터마킹이 삽입된 매크로블록 부분은 4, 5, 6번 부분인데, 삽입되지 않은 경우에는 연관성이 없음을 보여주지만 그림 11의 결과에서는 삽입된 부분이 뚜렷하게 나타남을 알 수 있다. 위에서 나타낸 그림 9에서는 전혀 나타나지 않던 워터마킹이 본 연구의 방식으로는 뚜렷하게 검출될 수 있음을 알 수 있게 된다. 본 실험의 경우 검출율은 영상 특성에 따라 약간씩 변하는데 토크쇼와 같이 화면 변화가 적은 경우는 약 95% 이상의 결과를 얻을 수 있었고, 드라마의 경우 약 92%, 월드컵 시민잔치의 경우 약 85% 이상의 결과를 얻을 수 있었다. 하지만 동영상의 경우는 콘텐츠의 길이가 최소한 30분 이상이며, 워터마크 삽입시에 일정한 주기를 가지고 지속적으로 삽입되므로 전체적인 검출률은 99% 이상의 높은 검출률을 나타낸다.

그림 9 캡쳐 공격 후 원본영상과의 차분 영상
Fig. 9 Differential image between original and watermarked image after capturing attack

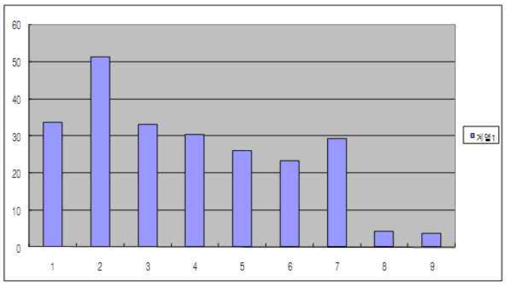

그림 10 워터마크가 삽입되지 않은 경우
Fig. 10 Result of non-WM

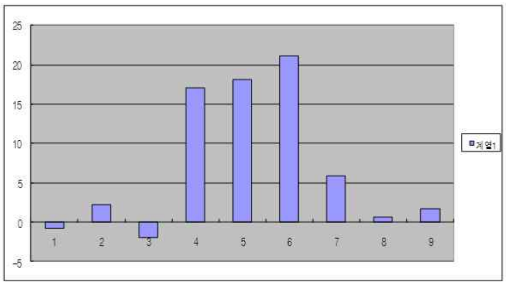

그림 11 4, 5, 6번째 매크로블록 그룹에 워터마크가 삽입된 경우
Fig. 11 Result of WM in 4, 5, 6th macroblock group

4. 결 론

본 논문에서는 화면의 확대와 축소를 통한 기하학적 공격, 압축율을 변화시켜 화질을 저하시키는 비트율 변화 공격 및 아날로그 신호형태로 캡쳐하는 캡쳐공격 등에 강한 특성을 갖는 워터마킹 알고리즘을 개발하고 MPEG-2 코덱에 적용시켜 상기의 결과를 얻었다. 기존에 존재하는 워터마킹들이 화면 비트맵에 워터마킹을 삽입하는 방식이나 DCT 계수에 삽입하는 방식을 택하여 화질의 변화 또는 기하학적 공격에 취약한 경우가 많이 발생하였으나 본 논문에서는 워터마킹의 특성인 강인성을 이룰 수 있었다. 검출률의 경우 한 프레임에 대해서는 동영상 특성에 따라 95%, 92%, 85% 이상의 검출율을 달성했으며 장시간 주기적으로 삽입할 경우 99% 이상의 높은 검출율을 얻을 수 있었다.

향후 본 논문에서 개발한 방식의 워터마킹을 기반으로 사용자의 정보를 원하는 크기만큼 삽입할 수 있는 핑거프린팅 방안의 기초가 될 수 있을 것으로 판단된다. 그러므로 본 연구는 앞으로 정확한 사용자의 정보를 검출할 수 있는 방안으로 알고리즘을 개발할 계획이다.

감사의 글

본 연구는 2010학년도 인덕대학 학술연구비 일부 지원에 의하여 이루어진 연구로서, 관계부처에 감사드립니다.

참 고 문 헌

[1] Chun-Shien Lu, Multimedia Security: Steganography and Digital Watermarking Techniques for Protection of Intellectual Property, Idea Group Publishing, 2005
[2] Clark, R. (2000). An introduction to JPEG 2000 and watermarking. *IEE Seminar on Secure Images & Image Authentication,* 3/1-3/6.
[3] Kim, Y.S., Kwon, O.H., & Park, R.H. (1999, March). Wavelet based watermarking method for digital images using the human visual system. *Electronics Letters, 35*(6), 466-468.
[4] Piva, A., Barni, M., Bartolini, F., & Cappellini, V. (1997, September). DCTbased watermark recovering without resorting to the uncorrupted original image. International Conference on Image Processing Proceedings, ICIP 97, (pp. 520-523).
[5] Tao, B., & Dickinson, B. (1997). Adaptive watermarking in the DCT domain. IEEE International Conference on Acoustics, Speech, and Signal Processing, CASSP 97, 4, 2985-2988.
[6] Wolfgang, P., & Delp, E. (1996, September). A watermark for digital images. *International Conference on Image Processing Proceedings,* ICIP 96, 219-222.
[7] 정제창 역, 최신 MPEG, 교보문고, 2004
[8] 유시룡 외, MPEG 시스템, 브레인코리아, 2004
[9] 이호석 외, 알기쉬운 MPEG-2, 홍릉과학출판사, 2003

[10] 김현승 역, Ingermar J. Cox, Matthew L. Miller, Jeffrey A. Bloom, 디지털 워터마킹, 도서출판 그린, 2005
[11] William B. Pennebaker, Joan L. Mitchell, JPEG: Still Image Data Compression Standard, 1993.
[12] I.J. Cox, Joe Kilian, F.T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia", IEEE Trans. Image Processing, Vol.6, No.12, pp.1673-1687, Dec., 1997.
[13] E. Koch and J. Zhao, "Towards robust and hidden image copyright labeling," Nonlinear Signal Processing Workshop, pp.460-463, 1995.

저 자 소 개

황 선 철 (黃 善 哲)
1965년 2월 15일생. 1999년 연세대학교 대학원 전기전자공학과 졸업(공학박사). 1991년 ~ 1998년 LG전자. 1999년 ~ 현재 인덕대학 방송영상미디어과 부교수
Tel : (02) 950-7444
Fax: (02) 950-7459
E-mail : sthwang@induk.ac.kr