# 영상의 인증과 무결성을 위한 블록 연성 워터마킹

주은경*, 박지환**
*부경대학교 전산정보학과
**부경대학교 전자컴퓨터멀티미디어공학부
E-mail:jooek@pknu.ac.kr

# Block Fragile Watermarking for Image Authentication and Integrity

Eun-Kyong Joo*, Ji-Hwan Park**
*Dept. of Computer and Information, Pukyong National University
**Div. of Electronic, Computer & Telecom. Eng., Pukyong National University

요 약

본 논문에서는 영상의 인증과 무결성을 위한 새로운 블록 연성 워터마킹을 제안한다. 원 영상의 인증과 무결성을 확인하기 위하여 이진로고 영상을 워터마크로 사용한다. 원 영상의 각 픽셀의 정보뿐만 아니라 해당 블록의 정보로 binary look-up table을 선택하여 워터마크 비트를 삽입한다. 기존 방식의 문제점으로 제시된 공격[6,7]을 효율적으로 막으면서도 시각적으로 변조유무 및 픽셀단위 또는 블록단위로 변조위치를 감지할 수 있음을 보인다.

## 1. 서론

최근 컴퓨터의 발전과 인터넷 보급에 따라 음악, 영상, 동영상 등과 같은 여러 형태의 멀티미디어 데이터가 디지털화되어 누구나 쉽게 저장 및 전송을 할 수 있게 되었다. 전송 받은 디지털 데이터는 각종 편집 프로그램의 편집도구를 이용하여 다양하게 변형을 할 수 있다. 또한, 이는 컴퓨터에서 다량의 복사가 가능하며 복사 후에도 원본과 동일하게 유지되므로 누구든지 저자의 동의없이 복사, 배포할 수 있다. 따라서, 이러한 디지털 데이터는 불법 복제 및 변조 등의 저작권 침해라는 문제를 야기시켰다.

일반적으로 디지털 데이터를 보호하기 위한 방법으로는 암호화(encryption)가 있다. 암호화는 암호화키를 이용하여 디지털 데이터를 암호화하여 전송하면 인가받은 사용자만이 알고있는 복호화키로 복호화(decryption) 해서 데이터를 복원할 수 있다. 그러나, 복호화된 이후에 데이터를 불법 복제하거나 배포한다면 디지털 데이터의 저작권 보호는 사실상 어려운 문제가 있다. 이에 대한 해결책으로 영상이나 오디오 등의 멀티미디어 데이터에 대해 저작권을 부여할 수 있는 방법으로 디지털 워터마킹(digital watermarking) 기술이 제시되었다.

디지털 워터마킹은 텍스트, 이미지, 비디오, 오디오 등의 디지털 데이터에 원 소유주만이 아는 마크(mark), 즉 워터마크를 사람의 육안이나 귀로는 구별할 수 없게 삽입하여 자신의 미디어에 대하여 저작권을 주장할 수 있는 방법을 제공한다. 이러한 디지털 워터마킹은 견고성에 따라 강성 워터마킹(robust watermarking)과 연성 워터마킹(fragile watermarking)으로 분류할 수 있다. 강성 워터마킹은 강인성을 목적으로 하기 때문에 통상적인 영상처리 기법에 의해 워터마크가 잘 지워지지 않으나, 연성 워터마킹은 인증 및 무결성을 목적으로 하므로 영상이 변조된 경우에는 워터마크가 깨어지도록 하여 변조여부를 검출하게 된다.

연성 워터마킹 중 디지털 영상에 대한 인증(authentication)과 무결성(integrity)을 위한 기법은 디지털 데이터의 내용이 조작되거나 변형되지 않았다는 것을 확인하면서 그 영상물의 송신자나 소유자를 확인할 수 있는 방법을 제공한다.

본 논문에서는 영상의 인증과 무결성을 목적으로 로고를 워터마크로 삽입하며, 워터마크를 추출하면서 시각적으로 영상의 변조 여부 및 변조 위치를 쉽게 확인할 수 있는 블록 연성 워터마킹을 제안한다.

본 논문의 2장에서는 기존의 연구에 대하여 살펴보고, 3장에서는 제안 알고리즘에 대하여 설명한다. 4장에서는 제안한 방법을 실험을 통하여 결과를 확인한다.

## 2. 기존의 연구

초기에 연성 워터마킹은 LSB(least significant bit)를 수정하면서 워터마크를 삽입하는 방법을 제안하였다[1][2]. 이 방법으로 워터마크가 삽입된 영상은 약간의 변경에도 워터마크가 깨어져서 변조된 부분이 정확하게 나타났다. 그러나, 이 방법은 공격자가 LSB를 손상하지 않으면서 영상을 변경하는 공격에 약점을 가지고 있었다.

Wong은 암호학적 해쉬함수로 디지털 서명을 생성한 후, 영상에 각 블록별로 LSB를 수정하여 워터마크를 삽입하는 방법을 제안하였다[4][5]. 이 방법은 영상의 변조 여부 및 블록 단위로 변조 위치를 확인할 수 있으나 워터마크가 삽입되는 위치가 노출되어 공격자가 영상내의 LSB를 삭제하고 자신의 비밀키를 사용하여 서명을 만든 뒤 LSB 부분에 삽입할 수 있는 단점이 있다.

Yeung and Mintzer는 각 픽셀에 대하여 이진로고를 워터마크로 삽입하는 방법(YM방식)을 제안하였다[3]. 이 방식은 먼저 비밀키에 의존적인 binary function f, f: {0, 1, ···, 255} → {0, 1} 즉, 0에서 255까지의 그레이 스케일을 0 또는 1로 사상하는 함수를 생성한다.

$$L_{(i,j)} = f(g_{(i,j)})$$

위의 식의 binary logo L과 이진로고에 따라 각 픽셀의 그레이 스케일을 수정하여 워터마크 삽입한다. 추출된 로고를 통해서 시각적으로 영상의 변조 여부 및 변조 위치를 확인하는데 좋은 성능을 가지고 있다. 그러나, 워터마크 삽입을 위조하기 쉽고 몇가지 공격은 성공적으로 수행된다. 첫 번째 공격은 동일한 비밀키와 이진로고 영상으로 여러 영상에 워터마크를 삽입할 경우에는 워터마크의 위치가 대응되는 이진로고의 위치이므로 공격자가 이진로고와 binary function f를 정확히 추정하여 워터마크된 영상을 수정하거나 위조할 수 있다. 두 번째는 "collage" 공격, 즉 워터마크 비트는 영상에 종속적이지 않으므로 공격자가 워터마크가 삽입된 여러 다른 영상의 일부분들 잘라내고 붙여서 결합해도 영상 내에서의 상대적인 위치를 유지하고만 있으면 정상적으로 인증이 된다[6,7].

Fridrich는 YM방식[3]을 기반으로 하여 이웃하는 그레이 스케일을 결합하여 원 영상에 종속적으로 워터마크 비트를 삽입하는 방법을 제안하였다[8]. 이 방법은 카메라 키로 블록 암호 알고리즘 $E_k$를 위한 비밀키를 생성하여 각 픽셀에 대해 이웃하는 픽셀들의 그레이 스케일을 블록 암호화하여 binary logo L을 얻는다.

$$L_{(i,j)} = f_k(g_1, \ldots g_{a\times a}) = \text{Parity}(E_k(g_1, \ldots g_{a\times a}))$$

위에서 언급한 공격[6,7]을 효율적으로 견디면서 영상의 변조 여부 및 블록 단위로 변조 위치를 확인할 수 있으나 알고리즘의 계산량이 많은 단점이 있다.

Hua Zhong은 각 영상의 특성을 이용하여 워터마크가 원 영상에 종속적인 방법을 제안하였다[9]. 이 방법은 원 영상에서 이미지의 특성을 가장 잘 나타내는 블록(예: Lena영상의 오른쪽 눈)과 카메라 키를 해시하여 비밀키를 생성한다. 이 비밀키에 의존적인 encryption matrix $\{B_{i,j}\}$를 원 영상의 크기와 동일하게 생성한다. 각 픽셀에 대해 binary function f, f: {0, 1, ···, 255} → {0, 1}와 Bi,j를 XOR하여 binary logo L을 얻는다.

$$L_{(i,j)} = f(g_{(i,j)}) \oplus B_{i,j}$$

위에서 언급한 공격[6,7]을 견디면서 영상의 변조 여부 및 픽셀 단위로 변조 위치를 확인 할 수 있으나 이미지의 특성인 블록을 변조하면 로고는 추출되지 않는 단점이 있다.

본 논문에서는 영상의 인증과 무결성을 위하여 기존의 방법들을 기반으로 위에서 언급한 공격을 견디면서도 알고리즘이 간단한 새로운 방식을 제안한다.

## 3. 제안 방식

### 3.1 삽입 알고리즘

본 논문에서는 아래 값에 따라 원 영상에 이진워터마크를 삽입한다.

- 각 픽셀의 MSB 값
- 해당 블록 모든 픽셀의 MSB의 암호값
- 각 픽셀의 값

1단계: 의사 무작위 이진 비트열의 생성

binary function을 작성하기 위하여 비밀키 K에 의존적인 8bit 그레이 스케일×4(256*4)의 크기인 의사 무작위 이진 비트열(bs: bit stream)을 생성시킨다.

$bs_1$ = bs{ 0, 1, ··· ,255}: 0 1 ··· 0
$bs_2$ = bs{256, 257, ··· ,511}: 1 0 ··· 1
$bs_3$ = bs{512, 513, ··· ,767}: 0 0 ··· 1
$bs_4$ = bs{768, 769, ··· ,1023}: 1 1 ··· 0

2단계: binary function과 LUT 생성

1단계에서 생성한 $bs_i$로 8bit 그레이 스케일 크기인 binary function $f_1$, $f_2$, $f_3$, $f_4$을 생성하여 각 function별로 0과 1의 LUT(binary look-up table)을 작성한다. binary function은 0에서 255까지의 그레

이 스케일을 0 또는 1의 이진값으로 사상하는 함수이고, 0과 1의 LUT은 각 이진 결과값의 위치에 따른 새로운 그레이 스케일표 이다.

$f_1$:{0, 1, …. 255} → {0, 1} ⇒ LUT00
$f_2$:{0, 1, …. 255} → {0, 1} ⇒ LUT01
$f_3$:{0, 1, …. 255} → {0, 1} ⇒ LUT10
$f_4$:{0, 1, …. 255} → {0, 1} ⇒ LUT11

3단계: 워터마크 삽입

원 영상의 각 픽셀 $g_{(i,j)}$에 대해 픽셀의 MSB값($f_1$)과 원 영상을 m×n으로 나눈 해당 블록별로 모든 픽셀의 MSB를 암호한 값($f_2$)으로 LUT를 선택한 후 이진로고에 따라 각 픽셀의 그레이 스케일을 수정하여 이진워터마크를 삽입한다.

(표1) $g_{(i,j)}$의 워터마크 삽입을 위한 LUT

| $g_{(i,j)}$의 LUT | $f_1$값 | $f_2$값 | $bs_i$ |
|---|---|---|---|
| LUT00 | 0 | 0 | $bs_1$ |
| LUT01 | 0 | 1 | $bs_2$ |
| LUT10 | 1 | 0 | $bs_3$ |
| LUT11 | 1 | 1 | $bs_4$ |

$f_1$값 = $f_1(g_{(i,j)'s\ MSB})$
$f_2$값 = $f_2(E_k(g'_{s\ MSB},\cdots,g_{m\times n's\ MSB}))$
여기서 $E_k$는 암호화하는 과정으로 Parity를 사용

### 3.2 추출 알고리즘

1단계: 의사 무작위 이진 비트열의 생성

비밀키 K를 seed값으로 하여 8bit 그레이 스케일 ×4(256*4)의 크기인 의사 무작위 이진 비트열(bs: bit stream)을 생성시킨다.

2단계: 워터마크 추출

워터마크가 삽입된 영상의 각 픽셀 $g_{(i,j)}$에 대해 픽셀의 MSB값($f_1$)과 해당 블록별로 모든 픽셀의 MSB를 암호한 값($f_2$)에 따라 bs를 선택한 후 픽셀의 값($f_3$)의 bs내 위치에 따라 이진값인 워터마크를 추출한다.

(표2) $g_{(i,j)}$의 워터마크 추출을 위한 $bs_i$

| $g_{(i,j)}$의 $bs_i$ | $f_1$값 | $f_2$값 | LUT |
|---|---|---|---|
| $bs_1$ | 0 | 0 | LUT00 |
| $bs_2$ | 0 | 1 | LUT01 |
| $bs_3$ | 1 | 0 | LUT10 |
| $bs_4$ | 1 | 1 | LUT11 |

$f_1$값 = $f_1(g_{(i,j)'s\ MSB})$
$f_2$값 = $f_2(E_k(g'_{s\ MSB},\cdots,g_{m\times n's\ MSB}))$
여기서 $E_k$는 암호화하는 과정으로 Parity를 사용

워터마크가 삽입된 영상의 인증시에는 워터마크된 영상에서 로고영상을 추출하여 소유권을 확인할 수 있으며 워터마크가 삽입된 영상이 변조시에는 추출된 로고영상에서 시각적으로 그 위치를 확인할 수 있다. 또한 삽입한 워터마크 영상과 추출된 워터마크 영상의 차이를 이용한 NC(Normalized Correlation)로도 그 무결성을 검증할 수 있다.

$$NC = \frac{\sum_{i=1}^{i<M}\sum_{j=1}^{j<N} w(i,j)w'(i,j)}{\sum_{i=1}^{i<M}\sum_{j=1}^{j<N} [w(i,j)]^2}$$

여기에서 w(i,j)는 삽입한 워터마크 영상이고, w'(i,j)는 추출한 워터마크 영상이다. NC는 0에서 1사이의 값을 가지는데 워터마크가 삽입된 영상이 변조되지 않았을 경우에는 그 값이 1이 되며, NC가 1이 아닌 경우에는 영상이 변조되었음을 알 수 있다.

## 4. 실험 및 결과

제안한 방법에서 실험에 사용한 영상은 256×256 크기의 Lena 및 House 8-bit 그레이 영상과 16×16 크기의 본 대학교 로고의 이진영상(그림1)이다. 워터마크 삽입과정 중 해당 블록별로 MSB를 암호화하기 위한 블록의 크기를 32×32로 실험하였다.

워터마크가 삽입된 영상(그림2)은 시각적으로 워터마크의 삽입의 여부를 구별하기 어려우며, PSNR도 43dB이상으로 원 영상과의 화질의 차이가 크지 않다는 것을 알 수 있다.

워터마크가 삽입된 영상으로부터 올바른 비밀키로 추출된 워터마크 영상(그림3(a))은 시각적으로 원래의 삽입한 로고임을 지각할 수 있고, NC의 결과값도 1로서 삽입한 영상과 추출된 영상이 동일함을 확인할 수 있다. 그러나, 비밀키가 다를 경우에는 그림3(b)와 같이 전혀 알아 볼 수 없는 영상이 추출되어 제3자가 소유권을 주장할 수 없게 된다.

공격자가 워터마크가 삽입된 영상에 그림4(a)의 우측 하단부와 같이 변조를 가할 경우에는 그 변조된 위치가 그림4(b)와 같이 시각적으로 표시되며, NC=0.98으로 변조가 가해졌음을 알 수 있다. 첫 번째 공격인 이진로고와 binary function f를 추정하는 공격은 이진로고가 동일하더라도 binary function f의 수가 4개이므로 이를 추정하기가 YM방식에 비하여 매우 어렵다. 또한 두 번째 공격인 collage 공격은 그림5(a)와 같이 Lena영상에 상대적 위치가 동일한 House영상의 일부분을 잘라내어 붙인 공격에도 변조위치가 블록으로 추출이 되어 실패하게 된다(그림5(b)).

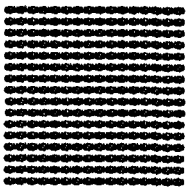
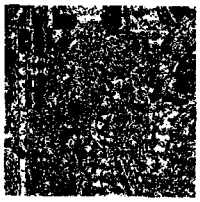


(a) Lena 영상 (b) House 영상 (c) 워터마크 영상

(그림1) 실험 영상

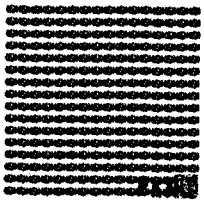

(a) Lena 영상 (PSNR=43.00dB) (b) House 영상 (PSNR=43.32dB)

(그림2) 워터마크가 삽입된 영상

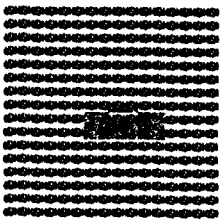
(a) 올바른 비밀키 사용 (NC=1) (b) 다른 비밀키 사용 (NC=0.53)

(그림3) Lena 영상에서 추출된 영상

(a) 변조 Lena 영상 (b) 추출 영상(NC=0.98)

(그림4) 변조된 영상과 추출 영상

(a) collage 공격 영상 (b) 추출 영상(NC=0.97)

(그림5) collage 공격 영상과 추출 영상

## 5. 결론

본 논문에서는 영상의 인증과 무결성을 위한 새로운 블록 연성 워터마킹 방법을 제안하였다. 워터마크 삽입시에 원 영상의 각 픽셀의 정보뿐만 아니라 해당 블록의 정보도 함께 가지는 블록 연성 워터마킹을 제안하여 이전에 지적된 두 공격[6,7]에 강인하면서 영상에 대한 소유권을 확인하고, 변조 여부와 픽셀단위 및 블록단위로 변조된 위치를 확인할 수 있었다.

## 참고문헌

[1] G. L. Friedman, "The trustworthy digital camera: restoring credibility to the photographic image," IEEE Trans. Consumer Electron., Vol. 39. pp. 905-910, Nov. 1993.

[2] R. G. van Schyndel, A. Z. Tirkel, C. F. Osborne, "A digital watermark," In Proc. IEEE, Vol. 2, pp. 86-90, 1994

[3] M. M. Yeung and F. Mintzer, "An invisible watermarking technique for image verification," In Proc of ICIP, pp. 680-683, Oct. 1997.

[4] P. W. Wong "A watermark for image integrity and ownership verification," In Proc. IS&T PIC, May 1998.

[5] P. W. Wong, "A public key watermark for image verification and authentication" In Proc. IS&T PIC, May 1998.

[6] J. Fridrich, M. Goljan, N. Memon, "Further attacks on Yeung-Mintzer watermarking scheme," In Proc. SPIE, pp. 428-437, Jan. 2000.

[7] M. Holliman and N. Memon, "Counterfeiting attacks for block-wise independent watermarking techniques," In Proc. IEEE trans. Vol. 9, No. 3, pp 432- 441, Mar. 2000.

[8] J. Fridrich, M. Goljan, A. C. Baldoza, "New fragile authentication watermark for images" In Proc of ICIP, Sep. 2000.

[9] Hua Zhong, Fang Liu, Li-cheng Jiao, "A new fragile watermarking technique for image authentication," In Proc. ICSP, pp. 792-795, 2002