공간영역에 기반한 블라인드 워터마킹

강현호, 박영란, 박지환
부경대학교 전자계산학과

Blind Watermarking based on the Spatial Domain

Hyun-Ho Kang, Young-Ran Park, Ji-Hwan Park
Dept. of Computer Science, PuKyong Nat'l University
hhkang@shannon.pknu.ac.kr, young@shannon.pknu.ac.kr, jpark@pknu.ac.kr

요 약

본 논문은 디지털 컨텐츠의 저작권 보호를 위한 공간영역기반 워터마킹을 다루고 있다. 특히, 기존 연구 중 Lee[1]의 방법에서 nonblind 워터마킹의 한계를 확장시켜 원 영상과 워터마크 없이도 삽입된 워터마크를 추출할 수 있는 blind 워터마킹을 제안한다. 제안된 기법은 기존기법에서 다루어진 다양한 공격에서 동등이상의 성능을 보이고 있다.

1. 서론

인터넷과 네트웍의 발달, 컴퓨터 산업의 성장으로 인해 수많은 정보를 얻는 것이 매우 쉽게 되었고, 이러한 정보중심의 사회에서는 여러 가지 다양한 멀티미디어 정보가 주축이 될 것이다. 그러나, 디지털화된 멀티미디어 정보는 제한 없이 복제 및 배포가 가능하다는 장점이 있으나 동시에 디지털 컨텐츠의 소유권 문제의 불명확성은 새로운 디지털 컨텐츠 개발 의욕을 떨어뜨릴 것이다. 따라서, 디지털 컨텐츠(문서, 오디오, 영상, 비디오 등)에 대한 저작권 보호를 위해 암호화 기술과 워터마킹(watermarking) 기술이 연구되고 있다. 암호화 기술은 디지털 데이터를 전송하는 동안 데이터를 보호할 수 있으나 수신자가 복원한 데이터는 원 데이터와 같으므로 더 이상 보호할 수 없게 된다. 그러나 워터마킹 기술은 워터마크(watermark)라고 하는 인지할 수 없는 소유자 정보를 원 데이터에 숨겨놓기 때문에 암호화 기술만의 문제점을 보완할 수 있게 된다. 즉, 디지털 데이터에 대한 저작권 문제가 발생하였을 경우 삽입된 워터마크를 추출함으로써 원 소유자의 저작권을 보호할 수 있게 된다.

이러한 디지털 워터마킹 기술은 저작권 보호의 용도뿐만 아니라 다양한 곳에서 이용되고 있다[2]. 예를 들면, 컨텐츠 소유자가 일련번호등의 구매자 정보를 워터마크로 삽입하여 계약을 위반하고 불법배포한 사용자를 추적하는 데 사용되는 핑거프린팅(fingerprinting) 기술, 복제방지 비트 정보로 워터마크를 구성하고 기록장치의 워터마크 검출기가 컨텐츠의 복제 가능여부를 결정하게 하는 복제방지(copy protection)기술, 상업성 광고나 TV 프로그램 속에 워터마크를 삽입하여 광고나 프로그램이 계약대로 방송되고 있는지를 확인할 수 있는 방송 모니터링(broadcast monitoring)기술, 데이터 위조 여부와 위조된 위치에 대한 정보를 제공하는 연성(fragile) 워터마크를 이용한 데이터 인증(data authentication)기술 등에 이용할 수 있다.

워터마킹 기술에 요구되는 사항으로 워터마크가 원본 데이터의 품질에 영향을 미치지 않도록 삽입해야 하는 지각적 투명성(perceptual transparency), 워터마크에 들어갈 수 있는 정보량을 기술의 응용분야에 따라 종속적으로 조절해야 하는 워터마크 삽입량(payload of the watermark), 원본에 숨겨진 워터마크를 제거하려는 공격에 대해서 워터마크의 손실이 생겨서는 안 되는 강인성(robustness), 워터마크의 삽입과 검출 알고리즘이 알려져도 워터마크의 존재를 검출하거나 제거하는데 어려운 보안성(security), 많은 응용 에서는 워터마크의 검출을 위하여 원본 데이터를 요구하지만, 불특정 다수를 대상으로 하는 블라인

드(blind) 워터마킹 기술이 다양한 분야에 응용하기 위해서는 원본 없이 검출하는 방식이 검토되어야 할 것이다.

워터마크를 삽입하는 영역의 종류에 따라 공간영역 기반과 주파수영역 기반으로 분류 할 수 있는데, 여기서는 공간영역에 기반한 기존연구들 중 하위 비트 부호화 방법[3]과 대표적인 통계적 방법으로 알려진 패치웍(patchwork)방법[4]을 살펴본다. 이 중 하위 비트 부호화 방법은 이미지 픽셀 값의 최하위 비트를 숨기는 방식으로 구현이 용이한 반면 일반적인 공격에 손상되기 쉽기 때문에 연성워터마크에 속하여 인증이나 무결성 검증에 활용할 수 있다. 패치웍 방법은 키값에 의해 n개의 픽셀을 갖는 두 블록 (A, B)를 랜덤하게 선택하고, 선택된 블록의 휘도값(a_i, b_i)을 아래와 같이 조작한다.

$$\tilde{a}_i = a_i + 1$$
$$\tilde{b}_i = b_i + 1$$

통계적으로 원 영상의 평균값이 i'라고 하면, A 블록의 영상 평균값은 $i'+1$이 되고, B 블록의 평균값은 $i'-1$이 된다. 하지만 전체 영상 평균값은 변함이 없이 i'가 된다. 워터마크의 삽입여부를 확인할 때 키값에 의해 임의로 선택된 영역 (A, B)에 대해 아래 식의 연산을 수행하면 S는 $2n$의 값을 갖고, 워터마크가 삽입되지 않은 경우는 0에 가까운 값이 나올 것이다.

$$S = \sum_{i=1}^{n} (\tilde{a}_i - \tilde{b}_i)$$

본 논문은 워터마크의 삽입을 공간영역(spatial domain)에 기반한 연구 중 특히 Lee[1]의 제안을 기반으로 하였고, 저작권 보호를 위한 목적으로 위의 요구사항을 만족하면서 blind 워터마킹의 구현에 초점을 두고 있다. 논문의 구성은 2장에서 Lee[1]의 기법에 대해서 기술하고 3장에서는 blind 워터마킹 기법을 제안을 하고 4장에서 제안한 기법에 대해서 기존의 방법과의 비교 실험을 행한다. 마지막으로 5장의 결론에서 향후의 과제를 제시한다.

2. 기존기법(non-blind watermarking)

이 장에서는 공간영역(spatial domain)에 기반한 다양한 이전의 연구들 중 Lee[1]의 논문을 고찰한다.

워터마크 삽입 방법은 원 영상(256 Gray level)에 2진 로고영상을 삽입하는 것이다. 워터마크 영상을 먼저 스크램블 한 후에 워터마크를 삽입한다. 스크램블 단계는 워터마크 이미지의 각 픽셀의 인덱스를 의사난수 순열(permutation)을 이용하여 재배열하는 것으로, 예를 들면 워터마크 영상을 주사선 방향으로 스캔한 순서로 0에서 $(M \times N-1)$이라면, 새로운 스캔 시퀀스의 인덱스로 난수 순열을 이용하여 0에서 $(M \times N-1)$사이의 $(M \times N)$개의 난수 시퀀스를 생성한다. 그런 다음 원 워터마크 영상 시퀀스를 랜덤 시퀀스에 의해 주어진 인덱스에 해당하는 새로운 시퀀스에서의 위치의 픽셀 값을 할당하는 것이다.

스크램블을 수행한 워터마크 영상을 원 영상에 삽입할 때 이미지의 크기는 원 영상의 4×4블록에 워터마크 1비트를 삽입하므로 워터마크 영상은 원 영상의 1/4 크기로 한다.

또한, 워터마크 정보가 삽입될 블록의 선택도 워터마크 영상을 스크램블 한 것처럼 재배열한다. 이것은 삽입 위치를 순서대로 하지 않고 랜덤하게 삽입하기 위한 것이다.

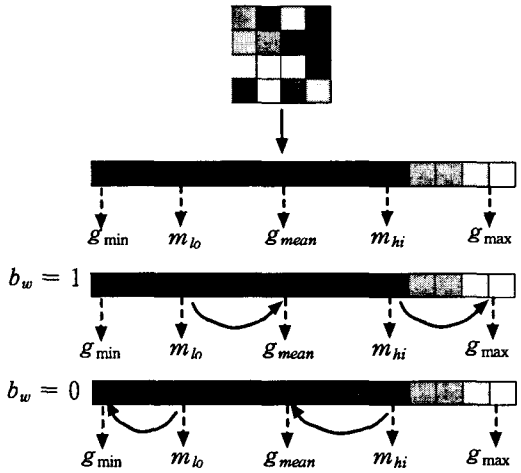

그림1. 워터마크 삽입 방법

스크램블된 워터마크 1비트를 b_w, 선택되어진 삽입될 위치의 4×4블록을 B라고 가정한다면 삽입 절차는 다음과 같다.

단계1 : B의 각 픽셀 값 16개를 오름차순으로 정렬을 하여 최소값(g_{min}), 최대값(g_{max}), 평균값(g_{mean})을 계산한다.

단계2 : 평균값을 기준으로 블록 B의 i,j번째 위치의 픽셀 $b_{i,j}$의 값 g가 평균보다 크면 G_H 그룹으로, 작으면 G_L 그룹으로 구분한 다음 그룹 G_H의 평균 m_{hi}와 G_L 그룹의 평균 m_{lo}를 계산한다.

$$\text{if} \quad g \geq g_{mean}, \quad b_{i,j} \in G_H$$
$$\text{if} \quad g < g_{mean}, \quad b_{i,j} \in G_L$$

단계3 : 주어진 워터마크 b_w의 값을 블록 B에 삽입을 하는 과정으로 다음의 조건에 따라 수행된다.

$$\begin{aligned} &if\ b_w = 1: \\ &g' = g_{max} \quad if\ g > m_H \\ &g' = g_{mean} \quad if\ m_L \leq g < g_{mean} \\ &g' = g + \delta \quad otherwise \\ &if\ b_w = 0: \\ &g' = g_{min} \quad if\ g < m_L \\ &g' = g_{mean} \quad if\ g_{mean} \leq g < m_H \\ &g' = g - \delta \quad otherwise \end{aligned}$$

여기에서 g'는 새로 부여될 픽셀 값, α는 상수, δ는 0에서 C_B사이의 난수로 C_B의 계산은 아래와 같다.

$$C_B = \max\ (C_{\min}, \alpha(g_{\max} - g_{\min}))$$

이것을 블록 B내의 각 픽셀마다 수행하면 원 영상의 한 블록 B에 대하여 B_{new}가 생성되고, 이와같은 절차를 원 영상 전체의 블록에 대하여 수행하면 워터마크가 삽입된 영상를 얻을 수 있다.

추출 방법은 원 영상 블록의 합과 워터마크 삽입 영상의 블록의 합을 비교하여 워터마크 삽입 영상의 블록의 합계가 크면 워터마크를 1로, 작으면 0으로 추정한다.

3. 제안기법(blind watermarking)

기존의 워터마크 삽입 방법은 한 블록의 픽셀 값들의 차가 큰 경우나 최대값과 최소값의 범위가 큰 경우에는 변경되는 값이 원래의 픽셀 값과의 차이가 크기 때문에 워터마크 삽입 영상의 열화가 다소 생긴다는 것을 알 수 있다.

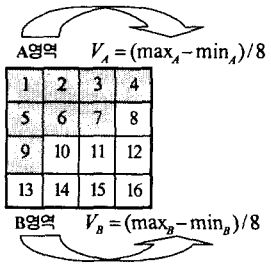

그림.2 블록분할

$$V_A = (\max_A - \min_A)/cnt_A$$
$$V_B = (\max_B - \min_B)/cnt_B$$

여기에서 $\max_A$ 와 $\max_B$ 는 A영역과 B영역의 최대값이고, $\min_A$ 와 $\min_B$ 는 최소값이며 cnt_A 와 cnt_B 는 분할 블록의 픽셀 수이다.

본 논문에서는 4×4 의 각 블록에서 그림2와 같이 대각선으로 A영역과 B영역으로 두 부분으로 분할한다. 그리고 두 영역의 최대값과 최소값을 각각 계산하여, 최대값과 최소값의 차를 구한 후, 그 차의 평균을 각 영역에 동일하게 분산을 시키는 방법이다. 특히 blind 워터마킹을 수행하기 위해 두 영역에 적용되는 값을 합과 차로 나누어서 블록간의 전체 휘도 성분의 차이를 조절하여 워터마크 정보를 삽입하는 방법이다.

삽입과정은 삽입 비트가 1이면 각 영역의 차인 V_A와 V_B를 A영역과 B영역의 각 픽셀 값에 아래와 같이 합과 차를 계산하고, 삽입 비트가 0이면 1반대로 한다.

$$\begin{aligned} &if\ b_w = 1: \\ &\quad A'_{area} = A_{area} + V_A \\ &\quad B'_{area} = B_{area} - V_B \\ &if\ b_w = 0: \\ &\quad A'_{area} = A_{area} - V_A \\ &\quad B'_{area} = B_{area} + V_B \end{aligned}$$

이 방법은 각 영역의 픽셀 값의 범위를 고려하여 워터마크를 삽입하였기 때문에 워터마크가 삽입된 영상의 열화에 큰 영향을 주지 않으면서 blind 워터마킹을 수행할 수 있다.

워터마크의 추출은 기존연구와 달리 원 영상의 정보 없이 워터마크된 영상 각각의 4×4 블록에 대하여 그림2에서와 같은 A영역과 B영역에 대해서 각각의 합을 계산한다. 각각의 합계를 서로 비교하여 A영역의 합계가 B영역의 합계보다 크면 1을 워터마크로 간주하고, 작으면 0으로 간주한다.

$$\begin{aligned} &if\ sum(A'_{area}) \geq sum(B'_{area}) \\ &\quad b_w = 1 \\ &if\ sum(A'_{area}) < sum(B'_{area}) \\ &\quad b_w = 0 \end{aligned}$$

여기에서 b_w가 추출된 워터마크 비트이다.

4. 실험 및 결과

실험에 사용된 영상은 512×512 크기의 Lenna영상(256 gray)에 128×128크기의 logo영상(binary)을 삽입하였다. 삽입된 워터마크의 공격을 위한 신호처리는

Matlab의 신호처리 툴을 사용하였다.

그림3. 원본 Lenna영상

그림4. Logo영상

그림5. 워터마크된 Lenna영상 (PSNR=37.3493)

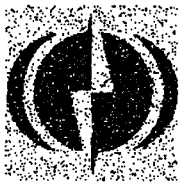

그림6. 추출된 logo영상 (NC=0.95101)

워터마크된 영상에 대해 9가지의 공격을 수행하였다. 각 결과에 대한 PSNR과 NC값은 표1과 표2에 각각 나타내었다. 약간의 PSNR감소가 생기지만 제안방법을 한번 더 수행하므로써 좀 더 높은 NC값을 얻을 수 있다. 이를 제안방식2로 나타낸다. 아래 실험 결과는 제안방식2를 바탕으로 한 결과이다.

$$PSNR = 10\log_{10}\frac{x_p^2}{\frac{1}{N}\sum_{k=0}^{N-1}(x_k - y_k)^2}$$

위 식과 같이 원 영상을 x, 신호 처리된 영상을 y로 두어서 원 신호값이 잡음값을 초과하는 정도를 나타낼 수 있다. 여기서 x_p는 영상 픽셀값의 최고값을 나타낸다.

$$NC = \frac{\sum_i \sum_j W_{ij} W'_{ij}}{\sum_i \sum_j [W_{ij}]^2}$$

위 식과 같이 원 로고를 W, 추출된 로고를 W'로 두어서 두 신호간의 상관도(NC) 값을 구할 수 있다.

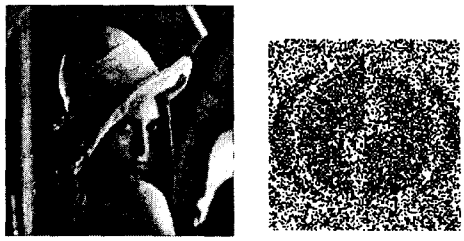

그림7. lowpass (PSNR=31.5780) (NC=0.82043)

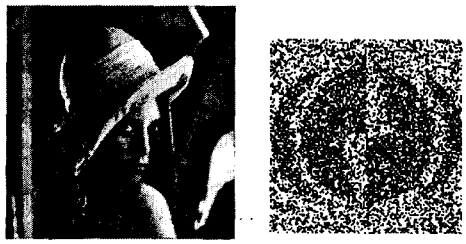

그림8. median (PSNR=33.9530) (NC=0.82431)

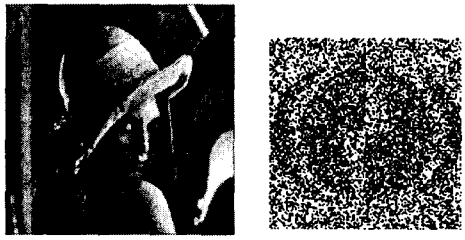

그림9. scaled(R) (PSNR=31.3385) (NC=0.81837)

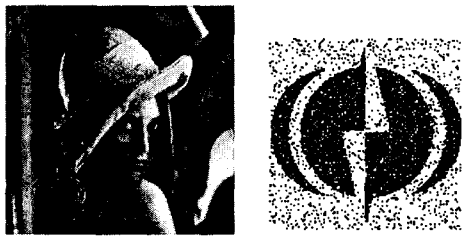

그림10. JPEG100%(PSNR=35.0014) (NC=0.95152)

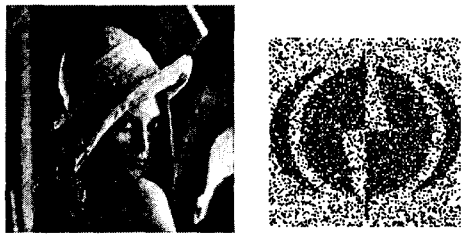

그림11. JPEG75% (PSNR=35.0014) (NC=0.89389)

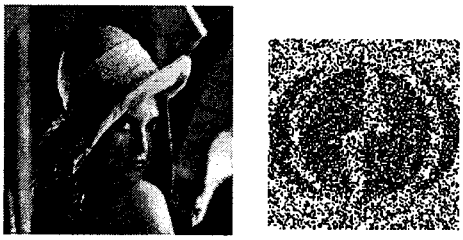

그림12. JPEG50% (PSNR=33.1842) (NC=0.80755)

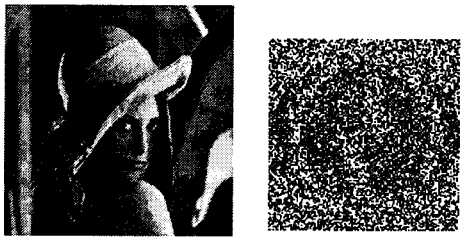

그림13. JPEG25% (PSNR=32.1533) (NC=0.67874)

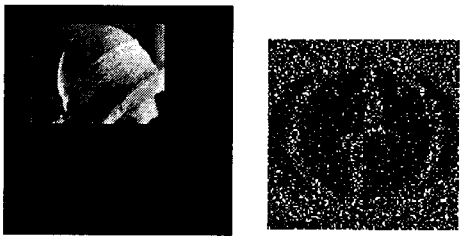

그림14. cropped (PSNR=7.0548) (NC=0.2552)

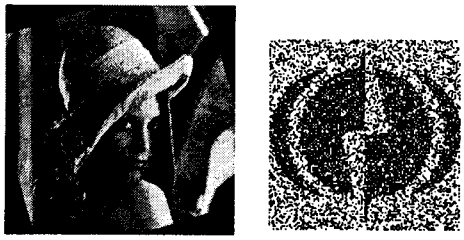

그림15. rotate17도(PSNR=15.7221) (NC=0.78488)

표1. 각종공격에 대한 PSNR

공격처리	기존방법	제안1	제안2
1. No pro	36.4252	37.3493	35.0953
2. Lowpass	31.5752	31.8690	31.5780
3. Median	33.7105	34.6806	33.9530
4. Scale(r)	31.2613	31.5685	31.3385
5. JPEG100%	36.3904	37.1979	35.0014
6. JPEG75%	34.8789	35.2038	33.8100
7. JPEG50%	34.0792	34.2519	33.1842
8. JPEG25%	32.7915	32.8980	32.1533
9. Cropped	7.0552	7.0555	7.0548
10. Rotate(17)	15.7261	15.7327	15.7221

표2. 각종공격에 대한 워터마크의 NC

공격처리	기존방법	제안1	제안2
1. No pro	1	0.90056	0.95101
2. Lowpass	0.52855	0.81784	0.82043
3. Median	0.73465	0.82318	0.82431
4. Scale(r)	0.76740	0.81758	0.81837
5. JPEG100%	0.97960	0.90260	0.95152
6. JPEG75%	0.84553	0.86856	0.89389
7. JPEG50%	0.81353	0.79256	0.80755
8. JPEG25%	0.81542	0.67240	0.67874
9. Cropped	0.25720	0.23457	0.25520
10. Rotate(17)	0.79012	0.76797	0.78488

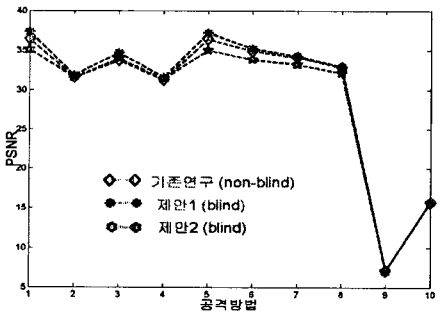

그림16. 각종공격에 대한 PSNR

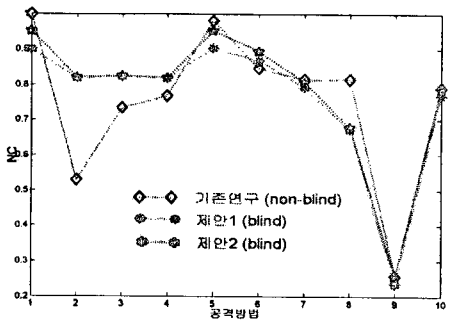

그림17. 각종공격에 대한 워터마크의 NC

5. 결론

본 논문에서 공간영역에 기반한 영상 워터마킹 중에서 blind 워터마킹 기법에 중점을 두었다. 워터마크로는 로고영상을 삽입함으로써 인간의 눈으로 바로 확인이 가능하게 하였다. 또한, 다양한 공격처리에 의한 기존연구와의 비교를 통해서 blind 방법으로 인한 성능저하는 없음을 확인할 수 있었다. 이러한 blind 워터마킹 기법은 기존연구의 제한된 응용 영역을 한층 넓게 확장시킬 수 있을 것이다. 최근은 주파수 영역에의 워터마킹 기법이 주류를 이루고 있으나 공간영역에서의 워터마킹을 함께 고려한다면 보다 효율적이고 넓은 응용에 활용될 수 있을 것으로 기대한다.

[참고문헌]

[1] C. H. Lee and Y. K. Lee, "An Adaptive Digital Image Watermarking Technique for Copyright Protection", IEEE Trans. on Consumer Electronics, Vol.45, No.4, November 1999.

[2] G. C. Langelaar, I. Setyawan, and R. L. Lagendijk, "Watermarking Digital Image and Video Data", IEEE Signal Processing Magazine, Vol.17, No.5, September 2000.

[3] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne, "A Digital Watermark", in Proceedings of ICIP'94, Vol.1, pp.86-90, 1994.

[4] W. Bender, D Gruhl, N. Morimoto, and A. Lu, "Techniques for Data Hiding", IBM Systems Journal, Vol.25, pp.313-335, 1996.

[5] N. Memon and P. W. Wong, "Protecting Digital Media Content", Communications of the ACM, Vol.41, No.7, July 1998.

[6] I. Pitas, "A method for signature casting on digital images", in Proceedings of ICIP'96, Vol.3, pp.215-218, 1996.