2001년 한국정보처리학회 추계 학술발표논문집 제8권 제2호

인증과 무결성을 위한 연성 워터마킹

이혜란*, 박지환*
*부경대학교 전자계산학과
E-mail:hrlee@mail1.pknu.ac.kr

Fragile Watermarking for Integrity and Authentication

Hye-Ran Lee*, Ji-Hwan Park*
*Dept of Computer Science, PuKyong University

요약

본 논문에서는 디지털 영상의 변조를 확인함과 동시에 변조의 위치를 확인하는 연성 워터마킹(fragile watermarking)을 위하여 DCT를 통해 블록의 에너지를 계산한 후, 에너지의 단계별로 워터마크의 삽입량을 조절하는 방법을 제안한다. 디지털 영상의 소유권 확인을 위해 디지털 서명을 사용하며, 영상에 DCT를 수행함으로서 모든 픽셀에 워터마크를 삽입하지 않고서도 변조의 유무를 확인하는 것이 가능한 방식이다. DCT 계수로 각 블록의 에너지를 계산하여 블록의 단계를 분류하며 에너지가 작은 블록들과 큰 블록들은 인간의 시각에 민감한 부분이므로 워터마크의 삽입 양을 줄이고, 중간 단계의 블록일수록 워터마크의 삽입 양을 늘린다. 에너지의 단계 분류에 의해 가변적으로 워터마크를 삽입함으로 워터마크의 비가시성과 연성을 만족시키며 변조의 유무와 위치를 확인할 수 있게 된다.

1. 서론

최근에 디지털 미디어와 통신망의 급격한 발달로 정보교환이 신속하게 이루어지고 있고, 멀티미디어 데이터의 사용이 증가되고 있다. 영상 데이터의 경우에도 기존의 데이터가 디지털화 됨에 따라 많은 편리성을 제공해 주고 있지만, 디지털 영상은 복제가 용이하다는 것과 복제된 영상은 원 영상과 동일하다는 것, 디지털 영상에 대한 조작이 용이하다는 등의 부작용이 심각한 실정이다. 따라서 디지털 영상의 정보보호가 더욱 필요하게 된다.

디지털 영상의 보호에는 크게 암호화 방법, 사이트 보호방법, 디지털 워터마킹 등이 있다. 암호화 방법에는 공개키 방식의 암호 알고리즘 및 비밀키 방식의 암호 알고리즘이 메시지의 조작이나 변형을 방지하기 위하여 여러 분야에서 사용되고 있다. 사이트 보호방법은 패스워드를 통하여 사용자의 접근을 제어하고 있으며, 디지털 워터마킹은 사람의 눈으로 식별할 수 없는 정보를 영상 내에 삽입, 추출하는 과정으로 영상에 대하여 손실이 발생할 수 있지만 소유권자가 워터마크를 쉽게 추출하여 자신의 영상에 대한 소유권을 주장할 수 있는 방법을 제공한다. 디지털 워터마크는 다음과 같은 목적으로 사용되어질 수 있다[1].

· Copyright protection : 지적 재산권을 보호하기 위해서 소유자는 데이터에 저작권 정보를 나타내는 워터마크를 숨길 수 있다.

· Fingerprinting : 불법적인 복사를 추적하기 위해서 소유자는 구매자의 데이터에 각각 다른 워터마크를 숨길 수 있다.

· Copy control : 워터마크는 복사 보호 목적을 위한 digital recording device를 직접적으로 제어할 수 있다.

· Broadcast monitoring : 상업적인 광고에 워터마크를 삽입함으로서 자동화된 monitoring 시스템은 광고가 계약대로 방송되어졌는지를 증명할 수 있다.

· Data integrity : fragile watermark는 데이터의 무결성을 체크하기 위해 사용되어질 수 있다.

본 제안기법은 데이터의 무결성과 인증을 목적으로 하며 데이터의 변형 여부 및 데이터의 변형 위치 정보를 알 수 있는 장점을 갖는다.

디지털 워터마킹은 크게 공간 영역 워터마킹(spatial watermarking)과 주파수 영역 워터마킹

(frequency watermarking)으로 분류할 수 있으며, 공간영역 워터마킹 기술은 인간 시각이 영상의 밝기에 민감하지 않다는 것을 이용하여 영상의 픽셀 값에서 LSB를 조작하여 윤곽선의 밝기 값을 변화시키는 방법이다. 주파수 영역 워터마킹은 영상을 DCT, DWT, DFT 등으로 변화된 주파수 계수에 워터마킹하는 방법이다.

본 논문에서는 DCT를 이용한 주파수 영역 워터마킹 방법을 적용한다. 2장에서는 연성 워터마킹(fragile watermarking)에 대한 개념과 기존의 방법들에 대해 기술하고, 3장에서는 DCT 변환을 이용하여 블록의 에너지를 구하고, 블록의 단계를 정하여 워터마크의 양을 가변적으로 삽입하는 제안방법을 기술한다. 그리고 4장에서는 실험 결과를 통하여 제안방법을 평가하고, 5장에서는 결론과 향후의 과제에 대하여 기술한다.

2. 관련연구

디지털 영상을 위한 연성 워터마킹의 필요 조건은 다음과 같다.

- 영상의 변조 여부가 검출 가능해야 한다.
- 영상의 변조된 위치를 지정 가능해야 한다.
- 원 영상 없이 워터마크의 추출이 가능해야 한다.
- 워터마크는 비가시적이어야 한다.

첫 번째와 두 번째 조건은 강성 워터마킹(robust watermarking)과 구별되는 연성 워터마킹의 특징이라고 할 수 있다.

디지털 영상의 무결성은 수신된 영상이 전송 도중 변형되었는지를 확인할 수 있는 기능으로 디지털 서명(digital signature)에 의한 방법과 연성 워터마킹에 의한 방법이 있다[2]. 연성 워터마킹은 영상을 다소 훼손시키는 단점은 있으나 다음과 같은 장점이 있다.

- 워터마크는 영상에 직접 삽입되므로 부가적인 데이터를 보관할 필요가 없다.
- 디지털 서명은 영상을 단순한 데이터 열로 간주하므로 영상의 독특한 구조를 활용하지 못하지만 연성 워터마킹은 영상의 구조적인 특성을 활용할 수 있어 영상공간상 변조된 위치나 변조의 종류 등을 알 수 있다.

Wolfgang등에 의해 제안된 방법[2]은 의사 랜덤 이진 계열인 워터마크 W를 모든 블록에 식(1)과 같이 삽입한다.

$$Y(\mathbf{b})=X(\mathbf{b})+W(\mathbf{b}) \quad (1)$$

여기서, $X(\mathbf{b})$는 원 영상의 블록이며 $W(\mathbf{b})$는 워터마크 블록, $Y(\mathbf{b})$는 워터마크 된 영상의 블록이다. 변조의 여부를 확인하는 과정은 식(2)와 같이 계산된다.

$$\delta(\mathbf{b})=Y(\mathbf{b})\cdot W(\mathbf{b})-Z(\mathbf{b})\cdot W(\mathbf{b}) \quad (2)$$

$Z(\mathbf{b})$는 테스트 영상이며, 임계값 T에 대해 $\delta<T$ 이면 $Z(\mathbf{b})$는 변조되지 않은 것으로 간주한다.

Ng등에 의해 제안된 방법[3]은 영상의 DCT를 구하여 아래와 같이 워터마킹하고 정확한 추출을 위해 해밍코드를 이용한다.

step 1 : 공간 영역에서 8×8블록의 64픽셀의 합을 계산한다

step 2 : 합을 16으로 나눈 나머지를 구한다.

step 3 : 구해진 나머지를 워터마크 삽입을 하기 위한 출발점으로 사용한다.

step 4 : 출발점 이후의 DCT 계수를 정렬한다.

step 5 : DCT 계수에 식(3)과 같이 워터마크를 삽입한다.

$$f'=f\times(100\%\pm\alpha) \quad (3)$$

여기서, f는 원 영상의 DCT 계수, f'는 워터마크 된 영상의 DCT 계수, α는 워터마크의 강도를 나타낸다.

Ng등의 방법은 영상의 공간도메인상의 각 블록값으로 워터마크의 삽입위치가 정해지므로 블록마다 워터마크의 삽입위치가 가변적인 특징을 가지며, 워터마크된 영상은 픽셀값이 변경되므로 공격자는 정확한 출발점을 찾을 수 없다는 특징을 가지고 있다.

Wong에 의해 제안된 방법[4]은 영상의 변조 여부 및 블록 단위의 변조 위치 확인이 가능하다.

step 1 : 영상을 8×8블록으로 분할한다.

step 2 : 블록내 각 픽셀의 LSB를 떼어낸다.

step 3 : 나머지 비트와 영상의 크기정보를 입력으로 하여 해쉬함수를 수행한다.

step 4 : 해쉬함수의 출력값과 watermark를 XOR 연산한다.

step 5 : 비밀키로 서명한 후 영상의 LSB에 삽입한다.

Wong의 방법은 Scaling이나 Cropping등에 의한 영상의 크기변화의 검출이 가능하며, 대응되는 공개키를 사용하면 해당 워터마크를 추출할 수 있다. 특정 부위의 영상이 변조되면 워터마크 추출 과정에서 해당 부위가 랜덤잡음과 같은 신호를 출력한다.

3. 제안 알고리즘

Wong의 방법은 워터마크에 디지털 서명을 하여

영상 내에 삽입하므로 대응되는 공개키로 복호화하면 영상의 소유권을 확인할 수 있으며, 영상의 무결성을 증명할 수 있는 워터마킹 기법이지만, 공간 영역에서의 워터마킹 기법을 사용하고 있으며, LSB에 워터마크가 삽입되기 때문에 LSB 공격에 취약하다는 단점이 있다.

제안 알고리즘은 Wong이 사용한 것과 같이 디지털 서명 알고리즘을 사용하여 소유권을 확인하는 방법을 도입하면서 Wong 방식이 LSB 공격에 취약하다는 단점을 개선한다. 먼저 영상을 일정한 크기의 블록으로 나눈 후 DCT 변환을 수행하고 식(4)와 같이 각 블록의 에너지를 구한다.

$$E_b = \sum_{i=1}^{n} (a_i)^2 \qquad (4)$$

b는 영상의 각 블록이며 E_b는 블록b의 에너지를 나타낸다. a_i는 블록의 DCT 계수이며, n은 블록의 크기이다. 각 블록의 에너지를 정렬하여 전체 블록을 몇 개의 단계로 분류하여 낮은 단계의 블록과 높은 단계의 블록에는 적은 양의 워터마크를 삽입하며, 중간 단계의 블록으로 갈수록 더 많은 양의 워터마크를 삽입한다. 에너지가 낮은 블록과 높은 블록은 영상에서 아주 밝거나 아주 어두운 영역으로 인간 시각에 민감한 부분이므로 워터마크의 비가시성을 만족시키기 위해 블록에 최소한의 워터마크를 삽입한다.

기존의 Wong의 방법에서는 모든 픽셀의 LSB에 워터마크를 삽입하여 변조를 확인하게 되지만, 제안 방식은 각 블록의 에너지를 이용하여 가변적으로 일부 저주파 계수에 워터마크를 삽입하므로 Wong의 방법보다 훨씬 적은 량의 워터마크를 삽입하고 변조를 확인할 수 있는 방법이다.

그레이 레벨 영상을 기준으로 워터마크가 삽입 및 추출이 되며, 원 영상 X의 크기는 $M \times N$이다. 삽입될 워터마크 영상 B는 $I \times H$의 크기를 가지는 2진 영상이다. 먼저, 영상 X를 일정한 크기의 블록으로 분할하고, DCT 변환을 수행한다. 워터마크 영상 B는 비밀키로 암호화되는데, 이는 대응되는 공개키로만 삽입된 워터마크를 검출할 수 있기 때문에 소유권자를 알 수 있는 인증 기능을 갖게 된다. 워터마크의 삽입과정은 그림1과 같이 개괄적으로 나타내어진다.

워터마크의 추출은 먼저 테스트 영상을 DCT 변환한 후, 에너지를 계산하여 블록의 단계를 분류한다. 블록의 단계에 따라 워터마크의 추출량이 결정

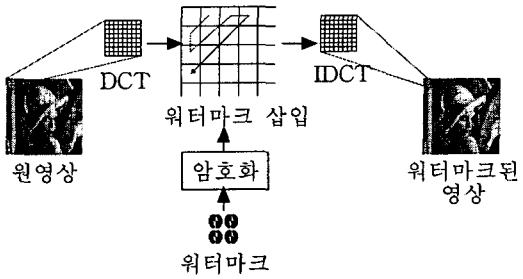

그림1. 워터마크 삽입 과정

되고, 추출된 워터마크를 복호화하면 삽입한 워터마크 영상을 얻을 수 있다. 변조가 일어났다면 변조된 위치에 워터마크 영상이 깨어진 것을 알 수 있으며, 깨어진 부분을 테스트 영상의 해당 위치에 대응시키면 영상의 어느 부분이 변조되었는지 위치 측정이 가능하다.

4. 실험 및 결과

본 논문에서 제안한 알고리즘의 효용성을 확인하기 위하여 그림2와 그림3에 표시된 256×256 크기의 그레이 레벨 Lena(8bits/pixel)영상과 같은 크기의 Camera영상을 각각 원 영상Ⅰ과 원 영상Ⅱ로 사용한다. 이진 워터마크 영상은 32×32 크기의 로고로 그림4에 표시되어 있다. 인증을 위해서 공개키 암호방식을 이용한 디지털 서명방식을 사용한다.

그림2. 원 영상Ⅰ

그림3. 원 영상Ⅱ

그림4. 워터마크

한 블록의 크기는 8×8로 하여 DCT를 수행한다. 각 DCT 블록은 식(1)을 사용하여 에너지를 구하고, 에너지를 기준으로 블록을 정렬하여 10단계의 블록으로 분류한다. 1, 2단계와 9, 10단계는 인간 시각에 민감한 부분이므로 각 블록에 1비트의 워터마크를 삽입한다. 워터마크의 삽입은 DCT 계수를 지그재그

주사하여 DC계수를 제외한 AC 계수에 이루어진다. 워터마크가 삽입될 AC계수의 1의 자리를 '0'으로 바꾼 후 워터마크를 삽입하게 된다. 3, 4단계와 7, 8단계의 블록에는 2비트의 워터마크, 5, 6단계 블록에는 3비트의 워터마크가 삽입된다. 그림5와 6은 워터마크된 영상이며, 그림7과 8은 영상을 간단히 변조하였다. 그림9와 10은 각각의 영상에서 추출한 워터마크이며, 특정 블록의 로고가 깨어진 것을 알 수 있다. 로고가 깨어진 부분을 영상의 블록으로 나타낸 것이 그림11과 12이다.

그림5. 워터마크된 영상 I

그림6. 워터마크된 영상 II

그림7. 변조된 영상 I

그림8. 변조된 영상 II

그림9. 추출된 워터마크 I

그림10. 추출된 워터마크 II

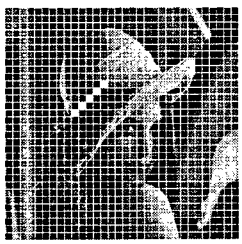

그림11. 영상의 변조 위치 I

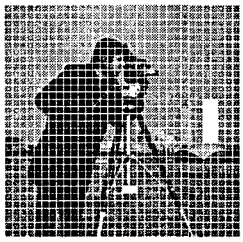

그림12. 영상의 변조 위치 II

식(5)를 사용하여 Wong의 방법과 제안방법의 PSNR(Peak Signal to Noise Ratio)을 계산하면 Wong의 방법은 50.549[dB]이며, 제안방법은 51.135[dB]가 나왔다.

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} [dB]$$
$$MSE = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (x(i,j) - \hat{x}(i,j))^2 \quad (5)$$

Wong의 방법은 영상에 65,536비트의 워터마크가 삽입되며, 제안방법은 1,844비트가 삽입된다. Wong의 방법은 공간 도메인에서 모든 픽셀에 워터마크가 삽입되는 반면 제안방법은 블록의 에너지 크기에 따라 워터마크를 가변적으로 삽입한다. 제안방법은 Wong의 방법보다 더 적은 양의 워터마크를 삽입하여 화질을 떨어뜨리지 않으면서 연성 워터마킹의 기능을 달성하는 장점이 있음을 알 수 있다.

5. 결론

본 논문에서는 블록의 에너지를 계산하여 블록 내에 적은 양의 워터마크를 가변적으로 삽입하여 변조의 유무와 위치를 측정할 수 있는 연성 워터마킹을 제안하였다. 단순히 하위 비트와 워터마크의 대체가 아니라 영상에 DCT 변환을 통하여 적은 양의 워터마크를 영상의 중요한 정보를 가지는 영역으로 확산하는 방법을 제안하였다. 제안방식은 영상의 화질을 기존의 방법보다 향상시키면서 변조를 검출할 수 있는 방법이다. 향후의 과제로는 블록 내에 워터마크가 삽입될 계수 선정에 관한 연구가 수행되어야 할 것이다.

[참고문헌]

[1] G. C. Langelaar, "Real-time Watermarking Techniques for Compressed Video Data," Ph.D. dissertation, Delft Univ. Technol., Delft, The Netherlands, Jan. 2000

[2] R. B. Wolfgang, D. J. Delp, "Fragile Watermarking Using the VW2D Watermark" Security and Watermarking of Multimedia Contents, Proc. of SPIE, vol. 3657, pp.204-213, 1999

[3] K. S. Ng, L. M. Cheng, L. L. Cheng and M. K. Wong, "Adaptive Watermarking by Using Pixel Position Shifting Technique," Proc. IEEE Transaction on Consumer Electronics, vol. 45, no. 4, pp. 1057-1064, Nov, 1999

[4] P. W. Wong, "A Public Key Watermark for Image Verification and Authentication," in Proceedings of ICIP 98, Oct. 1998