Stirmark공격에 강한 DCT 기반 디지털 워터마킹에 관한 연구

손윤경*, 강현호**, 박지환**
*부경대학교 교육대학원 전산교육전공
**부경대학교 대학원 전산계산학과
**부경대학교 전자컴퓨터정보통신공학부

A Study on DCT-based Digital Watermarking Robust Against Stirmark Attack

Yun-Kyung Son*, Hyun-Ho Kang**, Ji-Hwan Park***
*Dept. of Computer Education, PuKyong Nat'l University
**Dept. of Computer Science, PuKyung Nat'l University
***Dept. of Electronic, Computer and Telecommu. Eng., PuKyong Nat'l University

요 약

디지털 데이터의 저작권 보호를 위하여 최근 디지털 워터마킹에 관하여 많은 연구가 이루어지고 있다. 저작권 보호를 위해 멀티미디어 컨텐츠에 인지되지 않도록 기밀정보를 삽입하는 기술을 워터마킹이라고 하고 삽입되는 마크를 워터마크라고 한다. 본 논문에서는 DCT-Block을 이용한 디지털 워터마킹 기법을 제안한다. 이 기법을 이용하여 의사난수형의 잡음신호를 눈에 보이지 않도록 삽입하고, Stirmark공격에 강인하도록 구성하였다. 성능을 평가하기 위하여 영상에 워터마크를 삽입하고 각종 공격을 가하여 그 유효성을 보인다.

1. 서론

컴퓨터와 인터넷에 대한 기술의 비약적인 발전으로 인하여 각종 멀티미디어의 디지털화가 이루어졌고 디지털 데이터는 아날로그 데이터에 비해 저장 및 편집이 용이할 뿐만 아니라 누구나 디지털 데이터의 내용을 쉽게 변형 및 복제가 가능하기 때문에 각종 멀티미디어 서비스와 환경이 개인에게까지 제공되고 있다. 또한, 디지털 데이터는 원본과 복사본의 구분이 불가능하고 이로 인해 저작권의 보호문제가 심각하게 대두되었다.

멀티미디어 컨텐츠에 지적 소유권자의 마크를 삽입하여 저작권을 확인할 때 삽입한 마크를 검출하여 멀티미디어 컨텐츠의 저작권을 주장하기 위한 디지털 워터마킹(digital watermarking) 기법에 대한 연구가 계속해서 진행되고 있다.

본 논문에서는 블록분할 DCT 영상에 대한 워터마킹 기법을 제안한다. 이 방법은 공격에 강인한 특성을 가지며, 저작권 정보를 추출하는데 원 영상을 필요로 하지 않는 장점을 가지고 있다.

논문의 구성은 2장에서 워터마킹 기법들에 대해 간략히 소개하고, 3장에서는 DCT를 이용한 기존의 워터마킹 기법 및 문제점에 대해서 살펴보고, 기존 방법의 문제점 해결을 위하여 DCT 분할블록을 응용한 워터마킹 기법을 제안한다. 4장에서는 기존의 워터마킹 기법과 제안방법의 시뮬레이션을 수행한 결과를 나타내고, 여러 가지 평가를 통해 제안방법의 효율성을 확인한다. 마지막으로 5장에서는 결론 및 향후 연구과제를 제시한다.

2. 디지털 워터마킹 기법

영상에 워터마크를 삽입하는 방법에서는 공간 영역에서의 워터마킹기법과 주파수 영역에서의 워터마킹 기법으로 구분할 수 있다.

2.1 공간 영역에서의 워터마킹 기법

공간 영역에서의 워터마킹 기법은 원 영상에 직접

워터마크를 삽입하여 밝기의 세기를 직접 변화시키는 방법으로 다양한 영상신호처리 과정시 워터마크의 손실이 크고 삽입되는 워터마크의 양이 적다는 점에서 제3자의 고의적 공격에 취약하다는 단점이 있다. 공간 영역에서의 워터마킹 기법 중 하나로 Schyndel과 Osborne[1]는 삽입되는 워터마크를 키에 의해 발생된 이진 난수로 생각하여 원 영상의 픽셀들을 임의적으로 선택하여 밝기의 LSB (Least Significant Bit)를 변형시키는 방법을 제시하였다. 영상의 모든 픽셀 중 평균 50%의 화소만 워터마크가 삽입이 되고 워터마크의 삽입위치가 LSB임을 파악하면 공격에 취약한 상태에 놓이게 된다. 이처럼 공간 영역에서의 워터마킹 기법은 삽입될 수 있는 워터마크의 크기가 작을 뿐만 아니라, 잡음에 민감하고 기하학적인 변환, 데이터 압축과 같은 영상 처리에 강인하지 못하기 때문에 주파수 영역의 워터마크 연구가 활발히 진행되고 있다[2].

2.2 주파수 영역에서의 워터마킹 기법

주파수 영역에서의 워터마킹은 영상을 Fourier, DCT, DWT 등에 의해 변환된 주파수 성분의 계수에 워터마크를 삽입하는 것으로 시각적으로 덜 민감한 고주파 성분에 워터마크를 삽입하는 방법이다[3][4]. 그러나 시각적으로 덜 민감한 고주파 성분을 많이 압축하기 때문에 손실압축에 의해 워터마크가 쉽게 제거될 수 있다. 그래서 시각 특성상 중요한 부분에 워터마크를 삽입하는 방법으로 중간대역에 삽입하여 워터마크의 제거시 현저한 화질 저하를 가져오도록 하고 있다.

대표적인 주파수 영역 워터마킹 기법으로 Hartung [5]에 의해 제안된 Spread Spectrum원리에 기반한 방법이 있다. 삽입될 워터마크의 비트 시퀀스를 chip rate라 불리는 큰 계수에 의해 확산시킨다. 이 신호는 다시 진폭계수에 의해 확대되고, 이진의사잡음계열에 의해 변조된다. 이렇게 대역 확산을 시켜 변조된 신호는 spread spectrum watermark가 된다.

삽입된 워터마크의 검출은 원 영상에 대한 정보 없이 워터마크된 영상의 상관관계 합을 이용하여 쉽게 이루어진다. 상관관계 합의 의미는 원 영상과 삽입된 워터마크 사이의 상관관계가 양수이면 삽입된 정보비트가 양의 값이고, 상관관계가 음수이면 삽입된 정보비트는 음의 값이 되는 것이다.

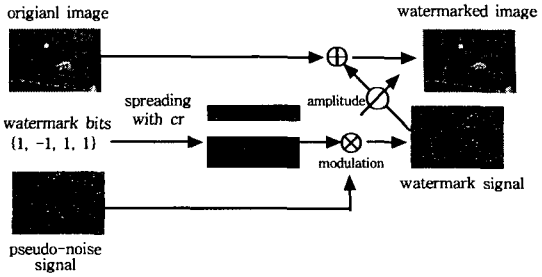


그림1. 대역확산을 이용한 워터마크 삽입

3. DCT-based 워터마킹 기법

3.1 기존의 DCT-based 워터마킹 기법

Barni[6] 등에 의해 제안된 DCT-based 워터마킹 기법은 원 영상에 삽입하고자 하는 워터마크가 M개의 pseudo random sequence로 구성되어 있고, 그 값을 $X=\{x_1, x_2, \cdots, x_M\}$라 한다. X를 이루고 있는 각 x_i의 값은 표준정규분포 $N(0, 1)$에 의해 발생된 랜덤한 실수 값이다. 워터마킹의 처리방법은 일반적으로 통신로 상에서 일어나는 것과 유사한데, 원 영상에 삽입되는 워터마크는 채널을 통해 전송되어 신호가 되고, 채널 잡음은 워터마크에 대해 행해지는 고의적인 공격이나 왜곡에 비유될 수 있다. 워터마크의 검출은 통신로 상에서 잡음이 부가된 영상을 수신하여 검출하는 것이다.

가. 워터마크의 삽입

워터마크를 영상에 삽입하는 방법은 크기가 $N \times N$인 원 영상 I를 $N \times N$ DCT를 취하여 DCT 계수를 구한다. 계산된 DCT 계수를 zig-zag scan하여 정렬된 DCT 계수 중 워터마크 시퀀스 $X=\{x_1, x_2, \cdots, x_M\}$가 삽입될 대역 L과 M을 결정하게 된다. 만약, L의 크기가 작아지면 저주파 영역에 워터마크가 삽입되어 시각적으로 쉽게 인식되고, M이 커지면 고주파 영역에 워터마크가 삽입되어 시각적으로 인식이 어려워진다. 처음부터 $L+M$번째까지의 계수를 선택하여 벡터 $T=\{t_1, t_2, \cdots, t_L, t_{L+1}, \cdots, t_{L+M}\}$를 구성한다.

그리고 나서, 시각적인 인지도와 강인성간의 trade-off를 고려하여 워터마크 시퀀스 X를 가장 저역인 L은 제외하고 M까지 삽입하여 새로운 벡터

$T'=\{t_1, t_2, \cdots, t_L, t'_{L+1}, \cdots, t'_{L+M}\}$을

아래의 식(1)과 같이 생성한다.

$$t'_{L+i} = t_{L+i} + \alpha \mid t_{L+i} \mid x_i \qquad (1)$$

이러한 방법은 원래의 신호에 워터마크된 신호를 더하여 진폭계수 α를 증가시켜도 워터마크에 의한 영상의 시각적인 감지를 방지하고, 공격자로부터 워터마크가 지워지는 것을 방지할 수 있다. α는 평균 0.2의 값을 가진다. 그런 다음에 계수 T'를 inverse zig-zag scan 하고, 다시 IDCT를 하여 워터마크된 이미지 I'를 생성한다.

나. 워터마크의 검출

워터마크의 검출은 변조된 영상 I^*을 $N \times N$ DCT를 하여 DCT 계수를 구하고, 구해진 DCT 계수를 zig-zag scan한 후, 그 계수 중 워터마크가 삽입된 대역인 $L+1$번째부터 $L+M$번째까지를 택하여 벡터 $T^* = \{ t^*_{L+1}, t^*_{L+2}, \cdots, t^*_{L+M} \}$를 구성한다. 워터마크된 DCT 계수에 삽입할 때 사용된 워터마크 X를 곱하여 상관계수 z를 구한다.

$$z = \frac{1}{M} \sum_{i=1}^{M} y_i \ t^*_{L+i} \qquad (2)$$

문턱치 S_z를 미리 정의하여 상관계수 z와 비교함으로서 워터마크가 존재하는지 그렇지 않은지를 판단한다. S_z는 워터마크된 영상을 이용하여 다음과 같이 계산된다.

$$S_z = \frac{\overline{\alpha}}{3M} \sum_{i=1}^{M} \mid t^*_i{}' \mid \qquad (3)$$

3.2 DCT 블록을 이용한 워터마킹 기법의 응용

기존의 DCT-based 워터마킹 기법은 기본적으로 주파수 영역에 워터마크를 삽입하여 워터마크가 잡음 형태로 삽입되므로 화질저하를 감소시키고 공격에 강한 장점이 있다. 그러나, DCT 계수를 이용하기 때문에 워터마크가 삽입될 영상은 크기는 $N \times N$ DCT를 해야하는 제약이 있다. 또한, 삽입 알고리즘이 DCT 계수를 이용하고 중간대역에 삽입된 워터마크라는 것이 공개된다면 영상에 대한 공격이 쉽게 이루어질 수 있다. 원 영상에 대한 화질열화를 최소화하면서 다양한 영상에 적용하기 위하여 워터마크를 삽입할 영상을 $n \times n$의 크기를 가지는 블록단위로 나누어 각 블록에 대해 DCT를 구하고 워터마크를 삽입하여 검출할 수 있는 방법을 제안한다. 즉, 제안방식은 그림2와 같이 m개의 워터마크를 삽입하기 위하여 먼저, 영상을 동일한 크기의 블록으로 분할하고, 분할된 블록내의 각 픽셀에 대한 DCT 계수를 구하여 원 영상에 워터마크를 삽입하게 된다.

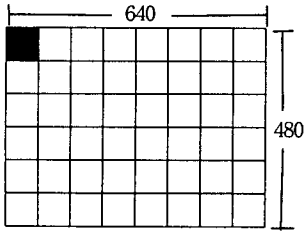


그림2. 원 영상의 $n*n$블록 분할

원 영상의 크기가 640×480일 때 한 블록당 크기를 80×80으로 하고(블록당 픽셀수 6,400개) 총 블록수를 48개로 했을 때, 각 블록에 대하여 DCT값을 구한다. 삽입될 워터마크의 수는 1,600개로 하고 삽입될 위치는 중간대역인 2,400번째부터 4,000번째까지 선택하여 그림 3과 같이 나타낼 수 있다.

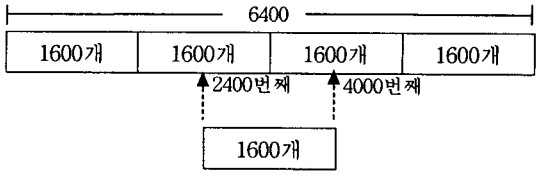


그림3. 워터마크 삽입 Scheme

삽입될 워터마크는 영상의 크기와 특성에 따라 적절한 block 크기로 분할하여 각 block에 대해 삽입될 주파수대역과 워터마크 수를 선택하는 것이 중요하다.

4. 실험결과

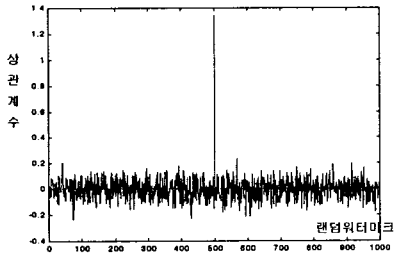
제안 방식의 유용성을 확인하기 위하여 256×256의 gray scale 영상 Lena, Bridge, Cronkite, Couple영상, House, F16영상을 대상으로 하여 시뮬레이션하고, 여러 가지 기하학적 왜곡을 가하여 공격으로부터의 강인성을 시험하고자 한다. 그림4(a)는 기존의 삽입 알고리즘에 따라 $\overline{\alpha} = 0.2$, $L = 2{,}5000$, $M = 1{,}6000$을 적용하여 삽입하였다. 그림4(b)는 제안방법에 따라 64×64크기의

블록으로 나누고 워터마크는 $\overline{\alpha} = 0.2$, L은 중간주파수 대역인 1,536번째부터 시작되고, M은 실험을 통해 블록 크기의 1/4에 해당하는 1,024개를 적용하여 삽입되어진 영상을 나타내었다. 두 방식 모두 시각적인 변화는 보이지 않는다.

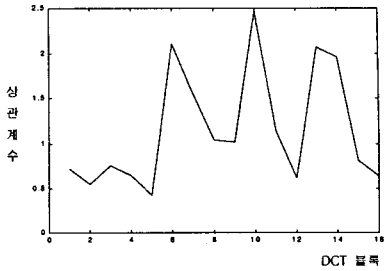
(a) 기존 방식 (b) 제안 방식

그림4. 워터마크 삽입된 영상

기존 방식의 워터마크 검출 알고리즘에 의한 결과로 워터마크 삽입시 더해진 워터마크에 대한 결과와 워터마크를 1000회 정도 랜덤하게 발생시켜 검출한 결과를 그림5(a)에 나타낸다. 그 결과로 나타난 상관계수 값은 삽입시에 사용된 정확한 워터마크를 가지고 검출하였을 때 가장 큰 값을 나타내고, 다른 신호 값에 대해서는 아주 낮은 값을 나타내었다. 이는 삽입시에 랜덤계열에 의해 발생된 워터마크와 똑같은 워터마크를 임의로 만들어 내는 것이 어렵다는 것을 보여준다. (b)는 제안방식의 검출 알고리즘에 의한 블록당 상관계수값을 그래프로 나타내었다. 최대 값은 2.4690이고, 최소 값은 0.4274로 나타났다. 부당한 사용자로부터 공격이 행해져서 몇 개의 블록에 포함되어 있는 워터마크가 사라지더라도 나머지 블록에 나타난 워터마크 검출하여 보임으로써 소유권을 주장할 수가 있다.



(a) 기존 방식



(b) 제안 방식

그림 5. 워터마크 검출시의 응답값

제안된 워터마킹 기법으로 워터마크된 영상에 대하여 현재 벤치마크 툴로 잘 알려진 Stirmark Ver.3.1 [7]의 대표적인 공격 파라메터를 이용하여 평가하였다. 기존 방식에 대해 각각의 공격을 가한 후, 검출 알고리즘에 의한 상관계수 값을 표1에 정리하였다. JPEG은 30%, Copping은 50%, 축소와 확대는 각각 75%, 110%한 후, 원 영상의 크기로 resizing하였다. 결과 값은 대체로 높은 상관계수 값을 보여 외부의 공격에 대해 강인함을 보여주고 있으나, cropping과 general linear geometric에 대해서는 낮은 상관계수 값을 나타내고 있다.

(a) JPEG 압축(30%) (b) Cropping(50%)

(c) 축소(0.75) (b) 확대(1.1)

(e)General linear geometric (f) Gaussian Filter
(1.010, 0.013, 0.009, 1.011)

(g) Sharpening (h) Median Filter(3×3)

그림6. 워터마크된 영상에 대한 공격 예

표1. 각 공격영상에 대한 상관계수(기존 방식)

공격파라메터	상관계수값	공격파라메터	상관계수값
JPEG (30%)	0.8855	General linear geometric	0.0745
Cropping (50%)	0.0727	Gaussian	0.2181
축소.(75%)	0.6015	Sharpening	6.0237
확대(110%)	0.6039	Median Filter	0.6738

표2는 제안방식에 의해 워터마크된 영상에 동일한 공격을 가하고 각 공격 파라메터에 의한 상관계수 중 최대 값을 정리하였다. 각 파라메터에 대한 결과는 기존방식보다 높은 값이 나타났고, 특히 cropping과 general linear geometric 공격에 대해 높은 값을 나타내었다. 각 분할블록의 특성에 따라 낮은 값을 보이는 것도 있으나 이는 각 블록에 따라 삽입되는 워터마크의 위치와 개수를 변화시키면 개선이 가능할 것으로 보인다.

표2. 각 공격영상에 대한 상관계수(제안 방식)

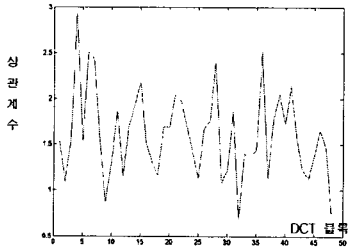
공격파라메터	상관계수 최대값	공격파라메터	상관계수 최대값
JPEG	2.4133	General linear geometric	0.3089
Cropping	2.1212	Gaussian	0.4304
축소	0.9736	Sharpening	10.4655
확대	1.1186	Median Filter	1.1922

또한, 실험영상을 달리하여 워터마크된 영상 각각에 대해 PSNR 값을 비교해 보았다. 기존 방식과 제안 방식 모두 PSNR값이 42dB이상을 보여 워터마크 삽입에 의한 화질의 열화가 거의 없으며, 기존 방식과 제안 방식의 PSNR 값에 큰 차이가 없음을 보여주고 있다.

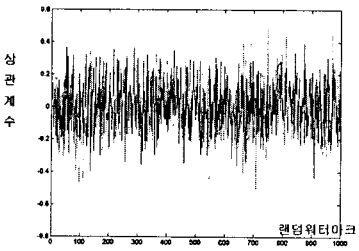
표3. 기존방식과 제안방식의 PSNR 비교

실험영상	기존방식	제안방식
Lena	47.07[dB]	46.69[dB]
Bridge	42.43[dB]	42.40[dB]
Cronkite	50.79[dB]	50.80[dB]
Couple	50.22[dB]	49.95[dB]
House	44.02[dB]	43.99[dB]
F16	45.49[dB]	45.24[dB]

제안 방식을 정방형이 아닌 640×480의 크기를 가지는 gray scale의 Soccer영상에 시뮬레이션 하여 검출된 response 값을 그래프로 나타내었다. 블록의 크기는 80×80로 하고 블록당 삽입되는 워터마크는 블록크기의 1/4크기 수준인 1,600개를 삽입하였다.



(a) 각 블록에 대한 응답



(b) 랜덤 signal에 대한 응답

그림7. 워터마크 검출 응답

표4는 영상을 분할하는 블록의 크기를 달리하여 상관계수와 PSNR 값을 살펴보았다.

표4. 블록의 크기에 따른 상관계수와 PSNR

블록의 크기	상관계수 최대값	PSNR
32×32	3.8303	46.50[dB]
40×40	4.6072	46.48[dB]
80×80	2.9323	46.41[dB]
160×160	2.3032	46.09[dB]

블록의 크기를 달리한 결과, 상관계수 값과 PSNR의 값이 차이를 보였다. 워터마크를 삽입할 영상의 종류와 특성에 따라 분할 블록의 크기와 워터마크의 시퀀스의 크기를 변화시켜 화질의 열화와 공격에 강하도록 구성하는 것이 요구된다.

5. 결론

본 논문에서는 Stirmark에 강인한 블록분할 DCT 영상을 이용한 워터마킹 기법을 제안하고, 그 유용성에 대해 살펴보았다. 워터마크는 잡음과 같은 형태로 워터마크된 영상의 화질에 영향을 미치지 않고, 삽입과 검출 알고리즘이 간단하여 멀티미디어의 유통시에 비용을 감소시킬 수 있다. 또한, 통신로 상에 발생 가능한 각종 공격에 대해서도 워터마크를 검출할 수 있음을 보였다. 제안 방식은 다양한 크기의 영상에 블록 단위로 워터마크의 삽입이 이루어져 기존 방식보다 계산량은 다소 증가하지만, 화질열화를 최소화하면서 공격에 대한 안전성을 보장할 수 있다. 향후 연구과제로는 DCT 기반 워터마킹 알고리즘을 응용하여 비디오, 오디오 등 다양한 멀티미디어에 대해서도 적용 가능한 워터마킹에 대한 연구가 필요하다.

[참고문헌]

[1] R. G. van Schyndel, A. Z. Trikel and C. F. Osborne, "A Digital Watermark" Proc. IEEE Int. Conf. Image Processing, Vol. 2, pp. 86-90, Austin, TX, 1994.

[2] I. Pitas, "A Method for Watermark Casting on Digital Images", IEEE Trans. on CSVT vol.8, no.6, pp775-780, Oct 1998

[3] I. J. Cox, J. Kilian, T, Leighton and T. Shamoon, "Secure Spread Spectrum Watermarking for Images, Audio and Video", Proc. of Int'l Conf. on Image Processing, Vol.3, pp.243-246, 1996

[4] J. Cox, "Secure Spread Spectrum Watermarking for Multimedia", IEEE Trans on Image Processing Vol.6, No.12, pp.1673-1687, Dec. 1997

[5] Frank Hartung, Bernd Girod, "Watermarking of Uncompressed and Compressed Video", Signal Processing 66(1998), pp.283-301, Feb 1997

[6] A. Piva, M. Barni, F. Bartolini, V. Cappellini, "DCT-based Watermarking Recovering without Resorting to the Uncorrupted Original Image", IEEE Signal Processing Society 1997 International Conference on Image Processing (ICIP'97)

[7] http://www.cl.cam.ac.uk/~fapp2/watermarking/stirmark