

디지털 영상을 위한 DWT 기반의 강인성 블라인드 워터마킹

*이용석, **서영호, **김동욱
광운대학교

stucom12@kw.ac.kr, yhseo@kw.ac.kr, dwkim@kw.ac.kr

A Robustness Blind Watermarking for Digital Image based on DWT

*Yong-Seok Lee, **Young-Ho Seo, **Dong-Wook Kim
Kwangwoon University

요 약

본 논문에서는 디지털 영상 콘텐츠의 저작권 보호를 위해 DWT 를 이용하여 강인성(robustness), 비가시성(invisibility), 보안성(security)을 만족하는 blind 워터마킹 방법을 제안한다. 이 방법은 영상정보에 대한 변화를 원본 영상의 특정 부분에 국한하지 않게 하기 위해서, 특정 주파수 성분에 워터마크를 삽입하는 것이 아닌 전체 영상에 워터마크를 삽입하는 방법이다. 이 방법에서 저작권에 대한 판단은 추출한 워터마크 정보를 통합 판단하는 방법을 사용한다. 이 방법은 저작권 보호를 위해 원본 영상을 보관할 필요가 없으며, Gaussian filtering, average filtering, sharpening 의 영상 필터링, JPEG 손실 영상 압축, resizing, 영상 품질 향상을 위한 Contrast 나 histogram equalization 등의 공격에 충분한 강인성을 보여준다.

1. 서론

디지털 영상 미디어의 발달로 콘텐츠들이 쉽게 전달될 수 있고, 불법복제되어 상업적으로 이용될 가능성이 커지고 있다. 이런 환경에서 콘텐츠 제작자들은 자신의 콘텐츠에 대해 저작권 주장이 필요하다. 이런 디지털 콘텐츠를 보호하는 방법에는 콘텐츠의 유통 자체를 제한하는 방법과 워터마크를 삽입해 후에 저작권을 주장하는 방법이 있다. 하지만 요즘 디지털 녹화기술의 발달로 recording 이나 capture 등으로 불법복제가 손쉽게 가능하다. 이런 이유로 디지털 워터마킹 방법이 저작권 주장을 위해 필요하다. 디지털 워터마킹을 수행하는 방법에는 spatial domain 에서 워터마크를 삽입하는 방법, frequency domain 에서 워터마크를 삽입하는 방법이 있다. 최근에는 frequency domain 중에서도 다양한 resolution 에 적응적으로 워터마크를 삽입할 수 있는 DWT(Discrete Wavelet Transform)를 이용한 워터마크 삽입 방법이 많이 연구되고 있다[1-5].

이에 본 논문에서는 DWT 로 변환된 frequency domain 에서 워터마크를 삽입하는 알고리즘을 제안한다. 강인성을 가지기 위해 3-level DWT 영역에서 워터마크를 삽입하며, LL, HL, LH, HH 영역에 반복적으로 암호화된 워터마크를 0 과 1 에 해당하는 특정한 상태로 변환하는 알고리즘이다. 이 알고리즘은 워터마크 추출에 원본영상이 필요없는 blind 방법이다. 이 알고리즘의 강인성과 비가시성을 보이기 위해 여러 공격에 대한 강인성 실험을 수행한다.

2. 워터마킹 방법

본 논문에서 제안하는 워터마킹은 512x512 color image 를 이용해 진행한다. 컬러영상은 워터마킹할 때 RGB channel 중 한 channel 에서 워터마크를 삽입하면 JPEG 등의 압축공격에 약하다는 실험결과가 발표된 적이 있으며[6] 따라서 luminance 성분인 Y 성분을 사용할 수 있는 YCbCr 이나 YIQ channel 로 변환 후 워터마크를 삽입하는 방법이 이용되고 있다[7]. 또한 DWT 를 이용해 워터마크를 삽입하는 경우 가시성의 이유로 LL 영역은 피해서 high frequency 대역 중 일부에 삽입하는 방법이 주로 사용되고 있으나, scaling factor 를 HL, LH, HH 영역과 다르게 설정해 LL 영역에도 약하게 워터마크를 삽입하는 방법이 연구되어 왔다[7]. 이런 기존의 연구결과들을 바탕으로 본 논문에서 제안하는 워터마크 삽입 알고리즘은 컬러영상을 YCbCr channel 로 변환해 Y 성분을 추출하고, 그림 1 과 같이 3-level DWT 영역에서 LL, HL, LH, HH 모두에 scaling factor 를 다르게 하여 워터마크를 삽입하는 방법을 사용한다.

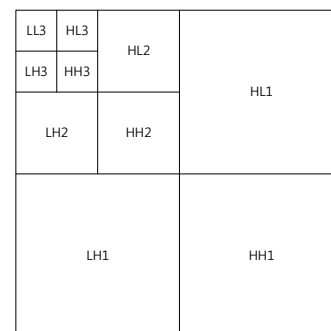


그림 1. 3 레벨 DWT 트리 구조

2.1 워터마크 삽입

워터마크를 삽입하는 방법은 그림 2 에 흐름도로 나타내었는데, 32×32 binary 워터마크를 key 를 사용하는 암호화 방법으로 암호화하여 random 신호로 변환한 후 삽입한다. 암호화로 random sequence 를 생성하는 방법은 key-dependent 한 방법으로 워터마크의 security 를 높여줄 수 있다.

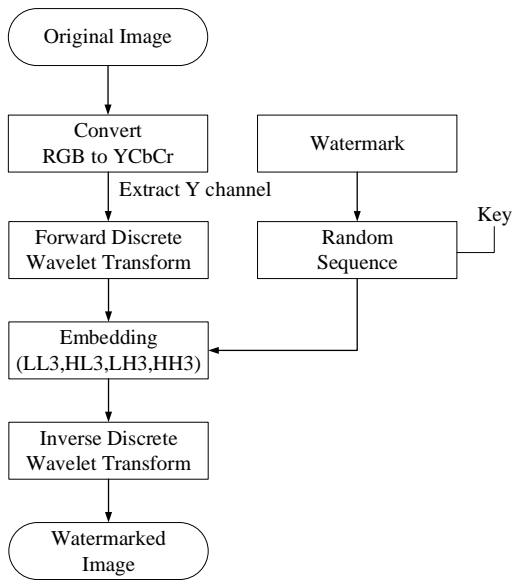


그림 2. 워터마크 삽입 알고리즘

이렇게 생성된 워터마크는 3-level DWT 각 영역에 삽입되는데, 512x512 영상을 3-level DWT 하면 각 영역은 64x64 의 크기를 갖는다. 따라서 각 영역에 워터마크를 반복적으로 4 번 삽입할 수 있고, 4 개의 영역을 모두 합하면 16 번의 32x32 워터마크를 삽입하게 된다. 워터마크를 삽입하는 방법은 식(1)과 같이 나타낼 수 있다. 값의 변화에 민감한 $-\alpha < W < \alpha$ 영역은 특별히 취급하고, 나머지 범위에는 같은 수식을 적용시킨다. 식(1)에서 $\lfloor \frac{W}{\alpha} \rfloor$ 는 $\frac{W}{\alpha}$ 보다 작은 가장 큰 정수를 나타내고, mod2 는 2 로 나눈 나머지를 나타낸다.

$$\begin{cases} \text{if } -\alpha < W < \alpha & \begin{cases} \text{if } wm = 0, & W' = +0.5\alpha \\ \text{if } wm = 1, & W' = -0.5\alpha \end{cases} \\ \text{else,} & W' = \alpha \cdot \left(\left\lfloor \frac{W}{\alpha} \right\rfloor + coef \right) \end{cases} \quad (1)$$

$$\begin{cases} \text{if } \left\lfloor \frac{W}{\alpha} \right\rfloor \bmod 2 = wm, & coef = 0.5 \cdot \frac{W}{|W|} \\ \text{if } \left\lfloor \frac{W}{\alpha} \right\rfloor \bmod 2 \neq wm, & coef = 1.5 \cdot \frac{W}{|W|} \end{cases} \quad (2)$$

여기서 W 는 DWT 영역에서 주파수 값이며, W' 은 워터마크 삽입 후 주파수 값, α 는 scaling factor, wm은 삽입될 워터마크 0,1 을 의미한다. 식(2)는 coef 를 판단하는 과정을 나타내었다.

2.2 워터마크 추출

워터마크를 추출하는 알고리즘은 그림 3 에 나타내었다. 워터마크된 영상에서 Y 성분을 DWT 하기 위해 color domain 을 변환하고, 3-level DWT 영역을 픽셀 단위로 추출한다.

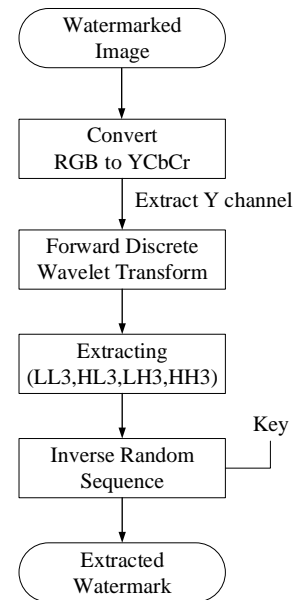


그림 3. 워터마크 추출 알고리즘

자세한 워터마크 추출방법은 식(3)과 같이 표현할 수 있다. 이 식을 이용해 32x32 영역별로 key 를 이용한 암호화된 random sequence 를 복호화하여 워터마크 데이터를 복원할 수 있다. 값은 식(3)을 통해 0 과 1 로 구분된다. wm'은 추출된 워터마크이다.

$$wm' = \left\lfloor \frac{W'}{\alpha} \right\rfloor \bmod 2 \quad (3)$$

워터마크는 그림 4(a)에서 볼 수 있듯이 LL3, HL3, LH3, HH3 영역에서 총 16 개가 추출되며 이는 일련의 판단과정을 거쳐 최종 1 개의 워터마크로 도출한다. 최종 도출 방법은 영역 통합 판단으로 영역을 구분하지 않고 16 개의 워터마크를 모두 같은 가중치로 두고 9 개 이상이 같은 값을 가지면 그 값을 대표값으로 추출한다. 이 같은 방식으로 워터마크를 도출하면, 특정 부분이 회손되더라도 보완적인 방법으로 워터마크를 찾을 수 있어 공격에 더 강인성을 가지게 된다. 또한, 특정 주파수가 아니라 전체적인 주파수를 취합해 판단하므로 특정 주파수를 제거하는 filtering 등 다양한 공격에 더 강인성을 가질 수 있다. 최종으로 추출된 결과는 그림 4(b)와 같다.

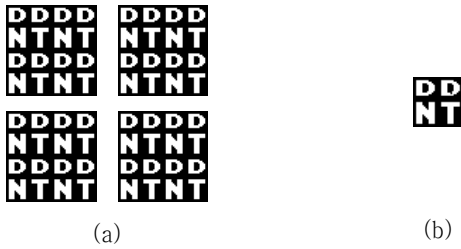


그림 4. 추출된 워터마크: (a) 3-level DWT LL, HL, LH, HH 에서 추출된 16 개의 워터마크, (b) 영역 통합 판단 후 최종으로 추출된 32x32 워터마크

3. 실험 및 결과

본 실험에서 사용한 영상은 512x512 크기의 컬러 영상들이며, 워터마크 데이터는 32x32 크기의 binary 영상이다. 워터마킹에 사용하는 scaling factor 인 α 값은 $\{\alpha_{LL3}, \alpha_{HL3}, \alpha_{LH3}, \alpha_{HH3}\} = \{2, 4, 4, 8\}$ 로 정하였다. 원본영상과 워터마크된 영상은 그림 5(a),(b)에 Lena 영상을 예로 비교하였고, 이 예에서 추출한 워터마크는 그림 5(c)에 보였다. 워터마크된 영상의 Y channel PSNR 은 40.05 이었다. 실험에서 워터마크 추출정도를 판단하기 위해 NCC(Normalized Correlation Coefficient)를 사용했으며, 제안한 알고리즘으로 워터마크된 영상을 공격하지 않고 추출했을 경우 NCC 가 1.0000 으로 워터마크가 완벽하게 추출되었다.

실험결과는 그림 6 에 다양한 공격에 대한 실험 결과를 기존 논문[1-5]와 비교해 그래프로 나타내었다. 가로축은 실험한 공격의 종류를 나타내었고, 세로축은 추출된 워터마크를 NCC 로 수치화한 결과이다. 그래프에서 꺾은선형 그래프는 각 실험에 대해 제안한 알고리즘의 결과이며, 막대그래프는 각각 기존 논문[1-5]의 결과를 나타낸다. 실험한 공격의 종류로는

Gaussian Blur, Average Blur, Sharpening, Median Filtering 같은 Filtering, JPEG 같은 손실 압축, Resizing(Shrink, Magnify), Crop(Cutting), Rotation 같은 기하학적인 변환, Contrast, Histogram Equalization 같은 화질변환, Gaussian Noise, Pepper & Salt 등 일반적인 영상 공격에 대해 다양하게 실험했다.



그림 5. 비가시성 예: (a) 원본영상, (b) 워터마크된 영상, (c) 영상에서 추출한 워터마크

실험 결과를 비교한 논문[1-5]는 DWT 를 사용하여 제안한 알고리즘이다. [1]은 3-level DWT 에서 LL, HL 영역에 SVD 로 워터마크를 삽입하고, [2]역시 3-level DWT 후 특정 고주파수 영역에 2x2 블록단위로 DCT 한 워터마크를 삽입한다. [3]은 1-level DWT 에서 LL 영역을 8x8 블록단위로 SVD 를 통해 워터마크를 삽입하며, [4]는 1-level DWT 의 LL 에서 8x8 블록을 DCT 한 후 SVD 로 워터마크 삽입한다. [5]는 기하학적 공격에 강한 워터마크로 DWT 후 4x4 단위로 SVD 를 통해 워터마크를 삽입한다. 이처럼 대부분이 일부 영역인 LL 만 사용하고 블록단위의 워터마크 삽입한다. 이렇게 특정 주파수만을 사용하면 영상에서 특정영역이 변조될 가능성이 높아진다. 또한, 블록단위로 워터마크를 삽입하면 영상의 일부 데이터가 급격히 변화하는 crop 이나 시점이동에 취약해진다. 이에 제안한 알고리즘은 LL, HL, LH, HH 전체영역에 픽셀단위로 삽입한다. 이는 본 논문에 차별성을 부과하는 특징이며 여러 공격에 대처할 수 있는 장점을 가진다.

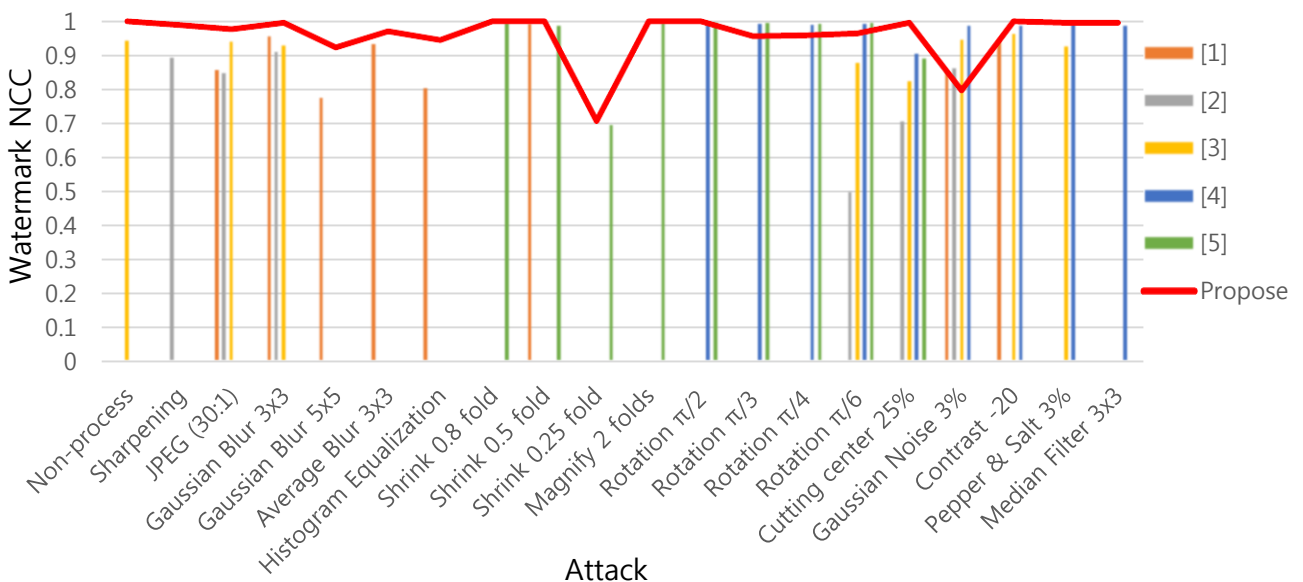


그림 6. 공격 후 추출된 워터마크 데이터

표 1. 공격 후 추출된 워터마크와 데이터

Attack	Avg. PSNR	WM NCC					
		Prop.	[1]	[2]	[3]	[4]	[5]
Non-process	40.05	1.0000	-	-	0.9473	-	-
Sharpening	33.62	0.9895	-	0.8962	-	-	-
JPEG (30:1)	31.72	0.9770	0.8617	0.8518	0.9439	-	-
Gaussian Blur 3x3	35.10	0.9958	0.9612	0.9132	0.9335	-	-
Gaussian Blur 5x5	32.64	0.9237	0.7805	-	-	-	-
Average Blur 3x3	33.07	0.9709	0.9384	-	-	-	-
Histogram Equalization	22.65	0.9457	0.8085	-	-	-	-
Shrink 0.8 fold	39.54	1.0000	-	-	-	-	1.0000
Shrink 0.5 fold	34.34	1.0000	0.9984	-	-	-	0.9911
Shrink 0.25 fold	29.56	0.7095	-	-	-	-	0.6996
Magnify 2 folds	46.70	1.0000	-	-	-	-	1.0000
Rotation $\pi/2$	36.93	1.0000	-	-	-	0.9986	1.0000
Rotation $\pi/3$	16.36	0.9570	-	-	-	0.9960	1.0000
Rotation $\pi/4$	16.02	0.9592	-	-	-	0.9926	0.9978
Rotation $\pi/6$	16.31	0.9649	-	0.5034	0.8819	0.9950	1.0000
Cutting center 25%	10.06	0.9958	-	0.7089	0.8286	0.9081	0.8965
Gaussian Noise 3%	30.30	0.7987	0.8574	0.8643	0.9502	0.9916	-
Contrast -20	36.63	1.0000	0.9484	-	0.9666	0.9897	-
Pepper & Salt 3%	34.22	0.9958	-	-	0.9312	0.9969	-
Median Filter 3x3	34.21	0.9958	-	-	-	0.9905	-
Average		0.9589	0.8943	0.7896	0.9229	0.9843	0.9538

결과적으로 [4]는 실험한 대부분의 공격에 강인성을 보였지만, Crop(Cutting)에는 약한 결과를 나타냈으며 실제로 특정부분에서 워터마크가 집중적으로 손실되며 로고 인식에 어려움을 보였다. [5]역시 기하학적 공격에 대비해 제안되었지만 Crop에는 여전히 약한 결과를 보였다. 이에 반해 본 논문에서 제안한 알고리즘은 표 1에서 볼 수 있듯이 [4,5]보다 Crop에서 우수한 결과를 보이며 Gaussian noise를 제외한 대부분의 공격은 [4,5]와 비슷한 결과를 보였다. 특히 기하학적 공격을 위해 제안된 [5]에 비해 resizing(shrink, magnify)에서도 더 좋은 추출율을 보였다. 전체적으로 본 논문에서 제안한 워터마크 알고리즘은 Gaussian filtering, average filtering, sharpening 등의 필터링, JPEG 손실 압축, resizing, 영상 품질 향상을 위한 contrast, histogram equalization 등의 공격에 대부분 높은 추출율을 보이며 평균 NCC 0.95로 높은 강인성을 나타냈다.

4. 결론

본 논문에서는 디지털 콘텐츠 저작권 보호를 위한 방법 중 하나인 워터마크 삽입 방법에대한 알고리즘을 제안했다. DWT를 이용해 LL, HL, LH, HH 영역이라는 비교적 많은 영역에 반복적으로 워터마크를 삽입하지만, scaling factor를 이용해 영역별로 값의 변화를 최소화하는 방법을 사용하였다. 이런 방법으로 blind한 워터마크임에도 다양한 공격에 강인할 수 있었고, 통합 판단 추출 방법으로 적은 값의 변화로도 강인하게 추출할 수 있다는 실험 결과를 이끌어냈다.

감사의 글

이 연구는 2014년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-

2013R1A1A2057798)

참고문헌

[1] Yavuz, Erkan, and Ziya Telatar. "Improved SVD-DWT based digital image watermarking against watermark ambiguity." Proceedings of the 2007 ACM symposium on Applied computing. ACM, 2007.

[2] Jiansheng, Mei, Li Sukang, and Tan Xiaomei. "A digital watermarking algorithm based on DCT and DWT." Proceedings of the 2009 International Symposium on Web Information Systems and Applications. 2009.

[3] Wang, Ben, et al. "An image watermarking algorithm based on DWT DCT and SVD." Network Infrastructure and Digital Content, 2009. IC-NIDC 2009. IEEE International Conference on. IEEE, 2009.

[4] Ye, Xueyi, et al. "A Multiple-Level DCT Based Robust DWT-SVD Watermark Method." Computational Intelligence and Security (CIS), 2014 Tenth International Conference on. IEEE, 2014.

[5] Ye, Xueyi, et al. "A SIFT-based DWT-SVD blind watermark method against geometrical attacks." Image and Signal Processing (CISP), 2014 7th International Congress on. IEEE, 2014.

[6] Kutter, Martin, Frederic D. Jordan, and Frank Bossen. "Digital signature of color images using amplitude modulation." Electronic Imaging'97. International Society for Optics and Photonics, 1997.

[7] Sun, Guangmin, and Yao Yu. "DWT based watermarking algorithm of color images." Industrial Electronics and Applications, 2007. ICIEA 2007. 2nd IEEE Conference on. IEEE, 2007.