

적응적 워터마크 삽입강도를 갖는 지각적 데이터 은닉 모델

조영웅, 장봉주, 김웅수, 문광석*, 권기룡
부산외국어대학교 전자컴퓨터공학과, 부경대학교 전자정보통신공학부*

Perceptual Data Hiding Model with Adaptive Watermark Strength

Young-Ung Jo, Bong-Ju Jang, Eung-Soo Kim, Kwang-Seok Moon*, Ki-Ryong Kwon
Division of Computer and Electronic Eng., Pusan Univ. of Foreign Studies
*Division of Electronic and Telecommunication Eng., Pukyung Nat'l Univ.

요 약

본 논문에서는 디지털 콘텐츠 저작권 보호를 위해 강인성과 비가시성의 유지를 위한 보다 효과적인 방법으로 웨이브릿 변환에서 적응적 워터마크 삽입강도를 갖는 지각적 데이터 은닉 모델을 제안한다. 먼저 영상을 9/7 쌍직교 웨이브릿 필터를 사용해 4레벨로 다해상도 분해한다. 다음으로 연속부대역 양자화(successive subband quantization)를 통한 시각적 중요계수(perceptually significant coefficient: PSC)들을 선정하여 선택된 계수들에 대해서만 워터마크 정보를 삽입한다. 지각 모델은 정상상태의 일반화 가우시안 모델(generalized gaussian model)로 추정된 NVF(noise visibility function)로 에지와 텍스처영역 그리고 평탄영역에 따라 각각 적응적으로 삽입되게 한다. 이는 각 서브밴드 내의 분산과 형상계수(shape parameter)에 의해 결정된다. 적응적 워터마크의 삽입강도를 갖기 위해 에지와 텍스처영역의 삽입강도는 각 서브밴드의 주파수 감도(frequency sensitivity)로 결정되고, 평탄영역의 삽입강도는 영상의 국부적 특성에 근거한 통계적 가중치를 사용한다. 삽입되는 워터마크는 랜덤시퀀스로 $N(0,1)$ 이다. 여러 가지 공격에 대한 실험으로 제안한 방법의 비가시성과 강인성을 확인한다.

1. 서론

오늘날 디지털 콘텐츠와 멀티미디어 데이터의 급격한 증가로 인하여 불법 복제와 무단 배포를 쉽게 한다. 디지털 워터마킹은 불법복제로부터 멀티미디어 데이터의 저작권을 보호하기 위해 사용될 수 있다. 워터마크는 미디어 데이터로부터 검출하기 위해 그 미디어 내용 안에 인지할 수 없는 신호로 직접 삽입된다. 워터마킹 시스템의 은닉과 추출은 미디어의 소유권과 원본의 확인, 사용 허가 등에 사용될 수 있다.

데이터 은닉 시스템의 중요한 요구조건은 은닉 알고리즘의 비가시성과 강인성이며 이 조건은 trade-off 관계에 있다[1]. 무엇보다 먼저 워터마크는 편집, 저장 혹은 워터마크 검출을 교묘하게 제거하려는 목적으로 미디어 콘텐츠에 접근하는 공격에 대해 강인성을 가지고 있어야 한다. 이런 공격들은 손실압축, 필터링,

잡음 첨가, 기하학적 변형 등이 해당한다. 두 번째로 워터마크는 원 영상의 질을 저하시키지 않도록 하기 위해 투명하게 은닉되어, 사용자가 워터마크가 은닉된 영상을 봤을 때 워터마크의 존재여부를 알 수 없어야 한다.

Swanson 등[2]은 인간 시각시스템과 DCT 영역에서의 블록들을 이용한 방법을 이용한 워터마킹 방법을 제안하였다. Watson 등[3]은 웨이브릿 각 부대역에 따른 시각시스템의 JND 특성을 이용한 주파수 감도를 제시하여 영상 압축에 이용했다. Kutter 등[4]은 HVS의 luminance sensitivity function을 기반으로 한 내용기반 적응 시스템을 개발하였다. 이 함수는 워터마크의 삽입을 위하여 영상의 휘도 추정을 기본으로 하지만 웨이브릿 압축 및 디노이징과 같은 공격에는 효과적이지 못하였다. Voloshynovskiy 등[5]은 내용 적응 디지털 워터마킹을 위하여 통계적 모델링 기법을 제안하였다. 원 영상 및 워터마크의 통계적 모델을 이용하여 고전적 Bayesian 및 다해상도 paradigm에 따라서 워터마크의 추정 및 검출을 하였다.

본 과제는 정보통신부의 정보통신기술기초연구지원사업(정보통신연구진흥원)으로 수행한 연구결과입니다.

기존의 워터마크 기법은 영상 특성에 대한 전체(global) 정보를 기반으로 하여 영상의 국부 특성에 관계없이 전체 원 영상에 대하여 동일한 워터마크 강도(strength)를 갖는 랜덤 노이즈로서 워터마킹 신호를 삽입한다. 그러므로 이것의 삽입 방법은 작은 변화에 대하여 민감한 평탄 영역에 artifact를 초래하게 한다. 이러한 artifact를 줄이기 위하여 주어진 워터마크의 세기를 줄여야 한다.

본 논문에서는 웨이브릿영역에서 연속 부대역 양자화[6]와 지각 모델을 이용한 적응적 워터마킹 기법을 제시한다. 이 모델은 워터마크를 더 강하게 삽입하기 위한 최적의 워터마크 삽입 위치를 결정하기 위해 영상의 국부적 특성을 이용한다. 원본 영상을 9/7 쌍직교 웨이브릿 필터를 사용해 4레벨로 다해상도 분해한다. 다음으로 연속부대역 양자화를 통한 시각적 중요 계수들을 선택하여 이들 계수들에 대해서만 워터마크 정보를 삽입한다. 지각 모델은 정상상태의 일반화 가우시안 모델로 추정된 NVF로 에지와 텍스처영역 그리고 평탄영역에 따라 각각 적응적으로 삽입되게 한다. 이는 각 서브밴드 내의 분산과 형상계수에 의해 계산된다. 적응적 워터마크의 삽입강도를 갖기 위해 에지와 텍스처영역의 삽입강도는 각 서브밴드의 주파수 감도로 결정되고, 평탄영역의 삽입강도는 영상의 국부적 특성에 근거한 통계적 가중치를 사용한다. 삽입되는 워터마크는 랜덤시퀀스로, i.i.d(independent identically distributed) 특성을 가지고 $N(0,1)$ 이다. 여러 가지 공격에 대한 실험으로 제안한 방법의 비가시성과 강인성을 확인한다.

2. 적응적 워터마크 삽입

2.1 연속 부대역 양자화

본 논문에서는 워터마크의 강인성의 보존을 위해 최고주파 부대역은 제외한 기적대역과 중간주파수 부대역에서 시각적으로 중요한 계수를 선택한다. 여기서 선택되는 계수들은 크기값이 큰 계수들을 선정하는데, 이는 워터마크 삽입에 따른 계수값의 변동율을 작게 하여 삽입된 영상의 열화를 최소화하기 위한 것이다. 시각적으로 중요한 계수를 선택하는 과정은 다음과 같다.

1 단계: 각 고주파 부대역들에 대하여 크기가 가장 큰 계수값의 절반을 각 부대역에서의 문턱값으로 정한다.

2 단계: 1 단계에서의 문턱값들 중 가장 큰 문턱값을 지닌 부대역에서 그 문턱값보다 큰 계수를 그 부

대역에서의 시각적으로 중요한 계수로 선택한다.

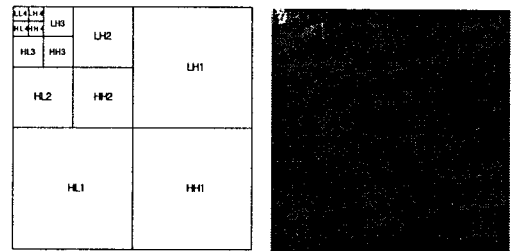
3 단계: 2 단계에서 선택된 계수들이 속한 부대역의 문턱값의 절반을 그 부대역의 새로운 문턱값으로 재설정한다.

4 단계: 원하는 계수들이 선택될 때까지 2 단계 및 3 단계를 반복한다.

따라서, 서브밴드 적용 문턱값은 각 부대역에 대해 식 (1)과 같이 적용한다[6].

$$TH_i = 2^{\lfloor \log_2 T_i \rfloor} - 1 \quad (1)$$

여기서, T_i 는 각 부대역의 가장 큰 계수를 나타내고, $\lfloor X \rfloor$ 는 X 보다 크지 않은 정수를 나타낸다. 워터마크는 부대역 적응적 문턱값보다 큰 PSC들에만 삽입한다. 그림 1은 각부대역과 레나 영상에 대한 선택된 PSC들을 나타낸다.



(a) 부대역 (b) PSCs

그림 1. Lena 영상에 대한 PSCs

2.2 지각 모델

이번 절에서는 웨이브릿영역에서 워터마크 은닉에 사용되는 통계적 다해상도를 갖는 지각 모델을 소개한다. 이 은닉방법은 [5]에서 제시한 국부 영상 특성을 갖는 NVF에 기초한다. 적응적 삽입 알고리즘을 갖는 모델은 SSQ에 의해 선택된 계수들에 대해서만 에지나 텍스처영역에서 더 강하게 삽입한다. 이 방법은 워터마크가 노이즈 특성을 갖기 때문에 변형된 정상상태 일반화 가우시안 모델 특성을 이용한다. 최적의 NVF를 결정하기 위해 우리는 워터마크를 노이즈로서 간주한다. 워터마크 삽입은 각 부대역의 국부 분산과 형상계수를 이용한다. 이는 에지나 텍스처영역과 평탄영역에 따라 적응적 criteria를 제공한다.

(1) 정상상태의 일반화 가우시안 모델에서의 NVF

정상상태의 일반화 가우시안 모델의 경우 NVF는 식 (2), (3)와 같다.

$$NVF(i, j) = \frac{\alpha \cdot w(i, j)}{\alpha \cdot w(i, j) + (1 - \alpha) \sigma_x^2} \quad (2)$$

$$w(i, j) = \gamma[\eta(\gamma)]^\gamma \frac{1}{\|\eta(i, j)\|^{2-\gamma}}, \quad \eta(\gamma) = \sqrt{\frac{\Gamma(\frac{3}{\gamma})}{\Gamma(\frac{1}{\gamma})}}$$

$$\eta(i, j) = \frac{x(i, j) - \bar{x}(i, j)}{\sigma_x} \quad (3)$$

여기서 $\Gamma(t)$ 는 감마함수이고, γ 는 형상계수로서 본 논문에서는 0.64를 사용하였다[5]. NVF를 예지나 텍스처영역에 더 강하게 삽입할 수 있게 $\alpha(=0.01)$ 를 첨가하여 변형된 NVF를 계산한다.

(2) 내용 적응적 워터마크 삽입

앞에서 계산한 NVF를 이용한 내용 적응적 워터마크의 삽입은 다음 식과 같다.

$$y = x + (1 - NVF) \cdot n \cdot S_1 + NVF \cdot n \cdot S_2 \quad (4)$$

여기서, y , x 및 n 는 워터마크가 삽입된 영상, 원 영상 및 워터마크를 나타내고, S_1 은 표 1과 같이 인간 시각시스템에 근거한 Watson 등[3]이 제시한 부대역 적응 삽입 강도로 예지와 텍스처영역의 삽입 강도이고, S_2 는 표 2와 같은 값으로 평탄영역에서의 삽입 강도이다. 본 논문에서는 워터마크 삽입방법은 워터마크의 강도를 증가함에 의해서 강인성을 증가시키는 목적과 화질의 열화를 동시에 줄일 수 있는 방법으로 제안한다. 그림 2는 제안한 방법의 블록도 이다.

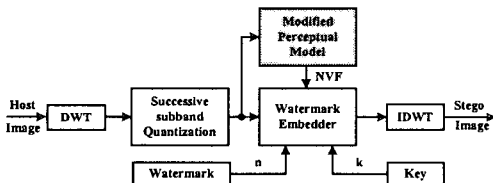


그림 2. 제안한 방법의 은닉 모델의 블록도

3. 컴퓨터 시뮬레이션 결과

제안한 웨이브릿 영역에서 내용 적응 워터마크 삽입 방법의 주된 특징을 설명하기 위하여 512×512 크기의 여러 영상에 대하여 컴퓨터 시뮬레이션 실험하였다. 실험은 StirMark 3.1 benchmark test를 사용하여 강인성과 비가시성의 정도를 측정하였다.

실험의 첫 번째 단계로 원 영상을 9/7 쌍직교 웨이브릿변환으로 4 레벨로 분해하였다. 사용되는 워터마크의 길이는 1000으로 하였고, 이는 단위 분산을 갖는 가우시안 랜덤 시퀀스를 사용하였다. 그리고, 강인성을 위해 가장 고주파에 해당하는 부대역은 삽입과정에서 제외하였다. 두 번째 단계로 각 부대역의 최대값을 갖는 계수의 반을 문턱값으로 하여 식(1)을 이용한 연속 부대역 양자로 시각적 중요계수 PSC를 선택한다. 다음 단계로 원 영상의 스토케스틱 모델에 따라

NVF를 계산한다. 계산된 NVF를 적용하여 식(4)에 따라 워터마크를 삽입한다.

그림 3은 Peppers 영상에 대한 제안한 방법으로 워터마크가 삽입된 영상이며, StirMark random bend 공격후의 영상을 그림 4에 보였다. 제안한 방법이 비가시성을 가지므로 삽입된 영상과 공격 후의 영상 모두 화질의 열화가 없음을 알 수 있었다. 그림 5는 영상의 필터링처리 후에도 삽입된 워터마크가 살아있음을 나타낸 그림이다. (a)는 워터마크를 삽입한 후 곧바로 추출한 값으로 원래의 워터마크와 추출한 워터마크와의 유사도(correlation response)이다. 추출한 워터마크가 원래의 워터마크와 상관성(correlation)이 있으므로 그 값이 32.14였고, (b)에서 median 필터링 후의 유사도는 12.69, (c)에서 gaussian 필터링 후에는 18.31, (d)에서 sharpening 필터링 후는 6.91이었다. 그림 6은 JPEG 손실 압축 하에서 제안한 방법의 비가시성과 워터마크의 강인성을 보기 위해 워터마크가 삽입된 영상을 JPEG 압축률을 변화시키면서 각각에 따른 PSNR과 유사도를 대역확산 방법과 비교하였다. 높은 압축률(10%)에서의 유사도에서 제안한 방법이 강인성이 떨어지지만 비가시성과의 trade-off면에서 볼 때 제안한 방법이 더 투명함을 알 수 있었다. 그림 7은 워터마크가 삽입된 영상을 JPEG 압축률 80%로 압축한 후 cropping율을 10%~90%으로 변화시키면서 각각의 PSNR과 유사도를 측정된 결과를 그림으로 나타내었다. cropping율이 증가할수록 삽입된 워터마크의 양이 감소하므로 강인성이 감소하였다. 결과 JPEG 압축의 손실압축과 기하학적 공격이 결합된 공격에서는 대역확산 방법과 비교해서 더 비가시성 있었지만서는 강인성 면에서 다소 떨어짐을 알 수 있었다.

4. 결론

본 논문은 웨이브릿 변환영역에서의 적응적 워터마크 은닉모델을 제시하였다. 이 연구는 최적의 워터마크 삽입 위치를 결정하기 위하여 연속부대역 양자화를 이용하였고, 시각적 모델을 위해 영상의 국부적 특성을 가지는 정상상태에서의 변형된 NVF를 계산함으로써 예지와 텍스처영역에 더 강하게 삽입하여 강인성을 증대시키고, 영상의 평탄영역의 화질열화를 최소화했다. StirMark benchmark test결과 제안한 방법의 강인성과 비가시성에서 우수함을 확인하였다.

[참고문헌]

[1] I.Cox, J.Kilian, T.Leighton and T.Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," in NEC Research Institute Tech Rep. 95-10, 1995.

- [2] M. Swanson, B. Zhu, and A. Tewfik, "Transparent Robust Image Watermarking," in *ICIP96*, vol. 3, pp.211-214, 1996.
- [3] A.B.Watson, G.Y.Yang, J.A.Solomon, and J.Villasenor, "Visual Thresholds for Wavelet Quantization error," *Proc. SPIE Human Vision and Electronic Imaging*, 1996, vol.2657, pp.381-392.
- [4] M.Kutter, "Watermarking Resisting to Translation, Rotation and Scaling," *SPIE*, Nov. 1998.
- [5] S.Voloshynovskiy, A.Herrigel, N.Baumgaertner, and T.Pun, "A Stochastic Approach to Content Adaptive Digital Image Watermarking," *Third Information Hiding Workshop*, 1999.
- [6] K.R. Kwon and A.H. Tewfik, "Adaptive Watermarking using Successive Subband Quantization and Perceptual Model based on Multiwavelet Transform," *SPIE*, vol. 4675-37, Jan. 2002.

표 1. 각 부대역에 따른 에지와 텍스처영역 삽입강도

orientation \ level	level			
	1	2	3	4
HL	23.028	14.485	12.707	14.156
LH	23.028	14.485	12.707	14.156
HH	58.756	28.408	19.54	17.864

표 2. 각 부대역에 따른 평탄영역 삽입강도

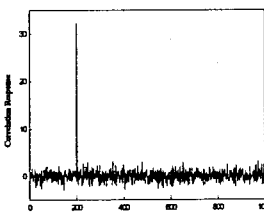
orientation \ level	level			
	1	2	3	4
HL	2.0	0.5	0.2	0.1
LH	2.0	0.5	0.2	0.1
HH	4.0	1.0	0.5	0.2



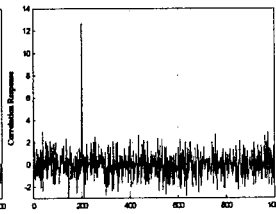
그림 3. 워터마크 삽입 영상



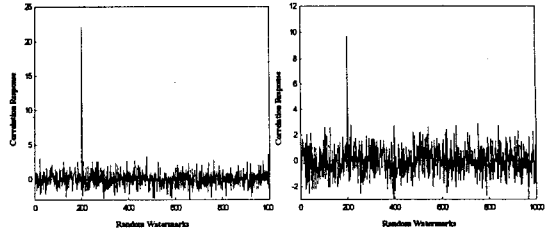
그림 4. 공격 후의 영상



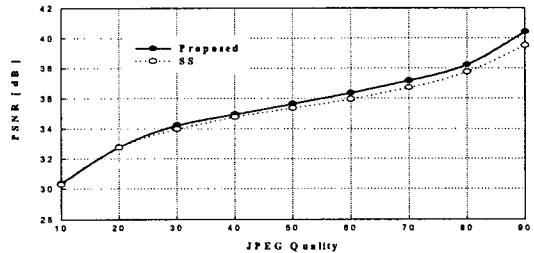
(a) No Attack



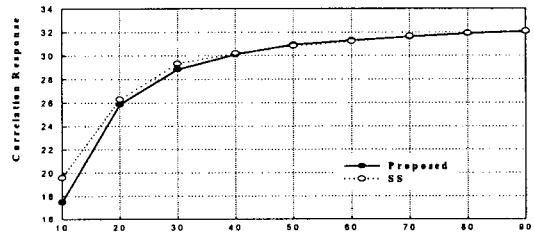
(b) median 필터링



(c) gaussian 필터링 (d) sharpening 필터링
그림 5. 일반적인 영상처리 후의 유사도

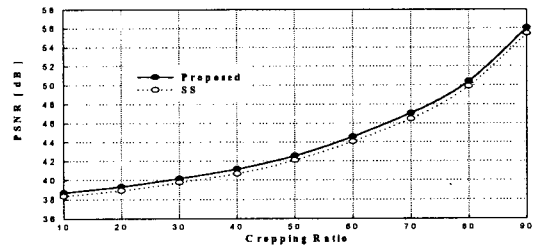


(a) PSNR

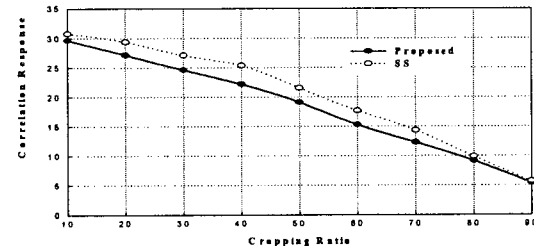


(b) 유사도

그림 6. JPEG 압축률 변화에 따른 PSNR과 유사도



(a) PSNR



(b) 유사도

그림 7. Cropping을 변화에 따른 PSNR과 유사도