

웨이브릿 변환 영역에서 스토케스틱 영상 모델을 이용한 내용기반 적응 워터마킹

김현천, 강균호*, 권기룡*, 김종진
부경대학교 전자정보통신공학부, 부산외국어대학교 컴퓨터전자공학부*

Content Adaptive Watermarking Using a Stochastic Image Modeling Based on Wavelet Transform Domain

Kim Hyun-Chun, Kang Kyun-Ho*, Kwon Ki-Ryong*, Kim Jong-Jin
Division of Electronic and Telecommunication Eng., Pukyung Nat'l Univ.
*Division of Computer and Electronic Eng., Pusan Univ. of Foreign Studies

요약

본 논문에서는 보다 효과적이고 강인한 워터마크 은닉을 위한 방법으로 웨이브릿 변환 영역에서 영상의 통계적 특성에 기초한 비정상상태(non-stationary)에서와 정상상태(stationary) 일반화 가우스(generalized Gaussian: GG)모델을 이용한 적응 워터마크 은닉 기술을 제안한다. 워터마크는 고주파 영역에서 연속 부대역 양자화(successive subband quantization: SSQ)를 이용하여 다해상도 영상의 웨이브릿 계수 중에서 시각적 중요 계수(perceptual significant coefficients: PSC)를 선택하여 삽입한다. 워터마크 은닉을 위한 지각 모델은 NVF(noise visibility function)함수에 의해 계산된다. 이것은 비정상상태와 정상상태의 통계적 특성을 이용하고, 국부영상 특성을 가진다. 은닉모델은 다해상도내의 각 부대역별 분산과 형상계수(shape parameter)를 사용한다. Stirmark benchmark test에 근거하여 여러 가능한 왜곡에 대한 실험에서 강인성과 비가시성에서의 우수함을 확인하였고, 비정상상태의 경우와 정상상태의 경우를 비교하였다.

1. 서론

최근 정보통신의 발달에 힘입어 향상된 네트워크 망을 통한 디지털 데이터의 보급이 급격히 증가하고 있다. 디지털 데이터를 복사할 경우 복사본은 원본 데이터와의 구별이 거의 불가능하다. 디지털 데이터의 이러한 특성 때문에 불법적인 복제 및 조작에 의한 저작권자의 지적 재산권에 위협이 가해지고, 저작권에 해당하는 보호와 보안의 필요성이 대두되었다. 따라서, 디지털 워터마킹(digital watermarking) 기술이 디지털 미디어 데이터의 저작권 보호를 위하여 개발되었다.

워터마킹은 인간이 시각적으로 인지할 수 없는 의미 있는 표시인 워터마크를 디지털 미디어에 은닉한 후, 그 표시를 찾아냄으로써 저작자임을 입증할 수 있는 기술이다 [1]. 따라서, 워터마크가 삽입된 미디어는 삽입된 워터마크로 인해 품질의 저하가 없어야하고, 추출할 때 완벽하게 워터마크를 추출할 수 있어야 소유권을 주장할 수 있다. 다시 말해 호스트 데이터의 품질의 열화를 피하기 위해 투명하게 은닉되어야 하고, 손실압축, 필터링, 기하학적 처

리 등에 의해 변형된 미디어에 대해서도 은닉된 워터마크를 검출할 수 있어야 한다.

Swanson 등[2]은 인간 지각 시스템의 특성을 이용한 블록 DCT 영역을 사용한 방법을 제안하였다. 그리고, Kutter [3]는 HVS(human visual system)의 luminance sensitivity function을 기반으로 한 내용기반 적응 시스템을 개발하였다. 이 함수는 효과적인 워터마크의 삽입을 위하여 영상의 휘도 추정을 기본으로 하지만 웨이브릿 압축 및 디노이징과 같은 공격에는 효과적이지 못하였다. Podilchuk 등[4]은 DCT 영역 및 웨이브릿 영역에 대해서 워터마크의 크기와 강도를 조절할 수 있는 내용 적응 구조를 개발하였다. Voloshynovskiy 등[5]은 내용 적응 디지털 워터마크를 위해 통계적 모델링 기법을 제안하였다. 이는 원영상 및 워터마크의 통계적 모델을 이용하여 고전적 Bayesian 및 다해상도 paradigm에 따른 워터마크를 추정 및 검출하였다.

기존의 워터마킹 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 워터마크 신호

본 연구는 한국과학재단 목적기초연구(R01-2002-000-00589-0)지원으로 수행되었음.

를 삽입한다. 그러므로 이것의 삽입 방법은 작은 변화에 대하여 민감한 평탄 영역에 artifact를 초래하게 한다. 이러한 artifact를 줄이기 위하여 주어진 워터마크의 강도를 줄여야 한다.

본 논문에서는 웨이브릿 변환영역에서 비정상상태 가우시안 모델과 정상상태 GG 모델을 이용한 적응 워터마크 은닉 기술을 제안한다. 워터마크는 고주파 영역에서 SSQ를 이용하여 다해상도 영상의 웨이브릿 계수 중에서 PSC에만 은닉한다. 워터마크를 은닉하기 위한 지각모델은 비정상상태와 정상상태의 통계적 특성을 이용한다. 이것은 국부영상 특성을 갖는 NVF 함수에 의하여 계산되어진다. 은닉모델은 다해상도 내의 각 서브밴드별 평균과 분산에 대한 추정기법과, 분산과 형상계수를 사용한다. 실험에서 비정상상태 가우시안 모델과 정상상태 GG 모델을 비교하였고, 제안한 두 모델의 비가시성 및 강인성이 우수함을 확인하였다.

II. 적응 워터마크 은닉 기술

1. 연속 부대역 양자화

적응 워터마크 삽입 모델은 쌍직교 웨이브릿 변환을 한 다음 손실 압축이나 기하학적 변환과 같은 일반 영상처리 기법에 의해 변형되기 쉬운 최상위 레벨의 고주파 부대역들에 위치한 웨이브릿 계수는 워터마크 삽입 단계에서 제외한다. 고주파 대역의 PSC는 서브밴드 적응 문턱값에 의한 SSQ에 의하여 선택되어진다. 이것은 각 서브밴드의 최대 계수값의 반으로 결정되어진다. PSC를 선택하기 위한 서브밴드 적응 문턱값[6]은

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

이다. 여기서 T_i 는 각 서브밴드의 최대값을 나타낸다. 워터마크는 서브밴드 적응 문턱값보다 더 큰 PSC에만 삽입된다. Lena 영상에 대한 선택된 PSC는 그림 1과 같고, 워터마크는 고주파 서브밴드의 PSC에 대한 비가시성 및 강인성을 제공하는 NVF를 기본으로 하여 삽입된다.



(a) PCS (b) Lena 영상
그림 1. Lena 영상에 대한 PSC

2. 통계적 모델에 의한 NVF 추정

1) non-stationary Gaussian model

비정상상태의 가우시안 모델의 경우 NVF 함수는

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

이다. 여기서 $\sigma_x^2(i, j)$ 는 화소좌표 (i, j) $1 \leq i, j \leq M$ 를 중심으로 한 국부 분산을 나타낸다. 워터마크는 단위 분산, $N(0, 1)$ 을 갖는 i.i.d. Gaussian 분포함수이다. 국부 분산을 추정하기 위하여 ML(maximum likelihood) 추정 기법을 사용한다. 국부영상에 대한 평균과 분산에 대한 ML 추정 기법은

$$\sigma_x^2(i, j) = \frac{1}{(2L+1)^2} \sum_{k=-L}^L \sum_{l=-L}^L (x(i+k, j+l) - x(i, j))^2 \quad (3)$$

$$x(i, j) = \frac{1}{(2L+1)^2} \sum_{k=-L}^L \sum_{l=-L}^L x(i+k, j+l) \quad (4)$$

여기서 $(2L+1) \times (2L+1)$ 은 추정을 위한 윈도우 크기를 나타낸다.

2) stationary GG model

최적의 NVF를 결정하기 위해 정상 상태의 NVF는 영상의 국부특성에 따라서 복원영상의 작은 변화에도 visible artifact를 줄이기 위하여 웨이브릿 영역의 각 서브밴드의 분산과 형상계수를 이용한다.

정상상태의 일반화 가우스 모델의 경우 NVF 함수는

$$NVF(i, j) = \frac{w(i, j)}{u(i, j) + \sigma_x^2(i, j)} \quad (5)$$

$$w(i, j) = \gamma [\eta(\gamma)]^\gamma \frac{1}{\|a(i, j)\|^{2-\gamma}} \quad (6)$$

$$a(i, j) = \frac{x(i, j) - \overline{x(i, j)}}{\sigma_x} \quad (7)$$

이다. 여기서 $\sigma_x^2(i, j)$ 는 웨이브릿 변환된 계수들의 분산이고, $\overline{x(i, j)}$ 는 영상의 평균값이다. γ 는 형상계수이다. 이 형상계수를 추정하기 위하여 모멘트 정합방법을 사용한다. 본 논문에서 추정한 형상계수 γ 는 0.67이다. 워터마크는 비정상 상태의 경우와 같이 단위 분산, $N(0, 1)$ 을 갖는 i.i.d. Gaussian 분포함수이다.

3. 내용 기반 적응 워터마크 삽입

비정상상태 모델과 정상상태 GG 모델을 이용하여 추정된 NVF 함수를 이용한 적응 워터마크 삽입 식은 다음과 같다.

$$x' = x + (1 - NVF)nA + NVFnB \quad (8)$$

여기서 x' , x 및 n 은 워터마크가 삽입된 영상, 원 영상 및 워터마크를 나타낸다. 그리고 A 는 텍스처영역과 예지영역의 워터마크 삽입강도이고 B 는 평탄영역의 삽입강도이다.

본 논문에서의 워터마크 삽입방법은 워터마크의 강도를 증가함에 따라서 강인성을 증가시키는 목적과 visual artifact를 동시에 줄일 수 있는 방법을 제안하였다. 제안한 적응 워터마크 은닉 모델을 간략화한 블록도는 그림 2와 같다.

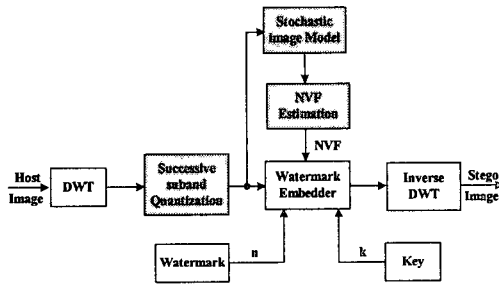


그림 2. 제안한 적응 워터마킹 모델

III. 실험결과 및 고찰

제안한 웨이브릿 영역에서 내용 적응 워터마크 삽입 방법의 주된 특징을 설명하기 위하여 512×512 크기의 여러 영상에 대하여 실험을 하였다. 사용한 웨이브릿 기법은 9/7 쌍직교 웨이브릿이며, 4레벨로 분해하였다. 단위 분산을 갖는 Gaussian random sequence 1000개를 워터마크로 사용하였다. 최상위 고주파 서브밴드는 워터마크 삽입과정에서 제외한다. 웨이브릿으로 다해상도 분해된 원 영상에 대해 첫 단계로 SSQ에 의해 PSC를 선택한다. 두 번째 단계로 비정상상태와 정상상태의 두 모델을 사용하여 각각 NVF를 추정한다. 다음으로 추정된 NVF를 삽입강도를 달리하여 두 모델에 대하여 각각 수행한 후 PSNR을 비교한다. 마지막으로 워터마크를 추출하여 유사도를 구하고, 두 모델을 비교한다. 강인성에 대한 척도로 원래의 워터마크와 추출한 워터마크와의 유사도(correlation response)를 사용하고[1], 비가시성의 척도로 PSNR을 사용하였다.

그림 3은 식(8)에서 삽입강도에 해당하는 S와 A, B를 변화시키면서 워터마크가 삽입된 영상의 PSNR을 나타냈다. 표 1은 Lena영상에 대한 Stirmark benchmark에 근거한 공격에 대한 PSNR과 유사도를 보였다. 그리고, 200 번째 seed를 사용하여 워터마크가 삽입된 영상에 공격이 없이 추출된 워터마크의 유사도를 그림 4와 그림 5에 보였다. JPEG 공격 하에 강인성을 평가하기 위하여 JPEG Q-factor를 10%에서 90%로 변화시켰을 때의 PSNR을 그림 6에, 유사도를 그림 7에 나타내었다. 그 결과 JPEG 압축에 대한 워터마크 구조의 탄성(resilience)을 알 수 있었다. 또한, 그림 8과 그림 9에 JPEG Q-factor 80%에서 워터마크에 해당하는 데이터를 10%에서 90%까지 훼손된 경우에 대한 실험 결

과를 보였다. 지금까지의 실험에서 비정상 상태의 Gaussian 모델의 경우는 강인성에서, 정상상태의 GG 모델의 경우는 비가시성에서 더 좋은 결과를 확인하였다. 그림 10에서는 워터마크가 삽입된 Lena영상과 Barbara영상을 Stirmark attack에 대해서 화질의 열화가 없음을 보였다.

IV. 결론

본 논문에서는 다해상도를 갖는 웨이브릿 변환영역에서 지각 모델 및 통계적 특성의 비정상상태 모델과 정상상태 GG 모델을 이용한 적응 워터마크 은닉 기술을 제안하였다. 4레벨로 분해된 웨이브릿 계수들 중에서 시각적으로 중요한 계수에만 워터마크를 삽입하였고, 워터마크를 은닉하기 위한 지각모델은 SSQ와 통계적 특성을 이용하였으며, 은닉 모델은 국부영상의 분산과 형상계수로 구해진 NVF 함수에 의해 설계되었다. 실험결과 은닉 모델은 비정상상태와 정상상태의 두 모델에서 비가시성과 강인성에 대해 우수함을 확인하였다.

[참고 문헌]

- [1] I.J. Cox, J. Kilian, T. Leighton, and T. Shamoon "Secure Spread Spectrum Watermarking for Multimedia," in *IEEE Trans. on Image Processing*, vol. 6, no. 12, pp. 1673~1687, Dec. 1997.
- [2] M. Swanson, B. Zhu, and A. Tzefik, "Transparent Robust Image Watermarking," in *IEEE International Conference on Image Processing ICIP96*, vol. 3, pp.211-214, 1996.
- [3] M. Kutter, "Watermarking Resisting to Translation, Rotation and Scaling," in *Proc. of SPIE*, Boston, USA, Nov. 1998.
- [4] C. Podilchuk and W. Zeng, "Image Adaptive Watermarking Using Visual Models," in *IEEE Journal on Selected Areas in Communication*, vol. 16, no. 4, pp. 525-539, May 1998.
- [5] Sviatoslav Voloshynovskiy, A. Herrigel, N. Baumgaertner, and T. Pun, "A Stochastic Approach to Content Adaptive Digital Image Watermarking," in *Proc. of 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.

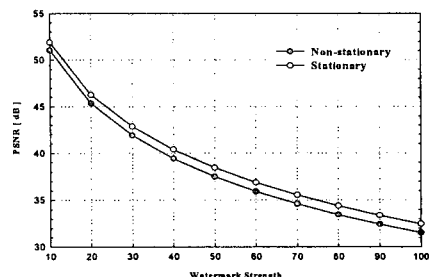


그림 3. 삽입강도에 따른 PSNR 비교

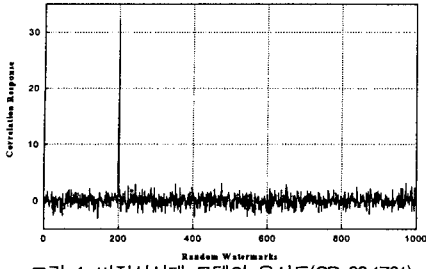


그림 4. 비정상상태 모델의 유사도(CR=32.1761)

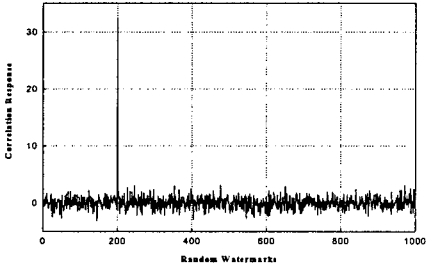


그림 5. 정상상태 모델의 유사도(CR=32.1785)

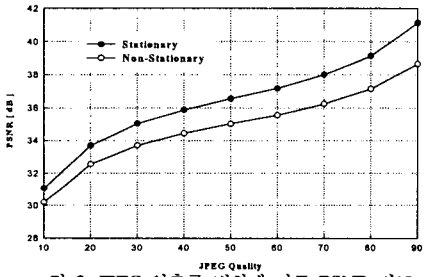


그림 6. JPEG 압축률 변화에 따른 PSNR 비교

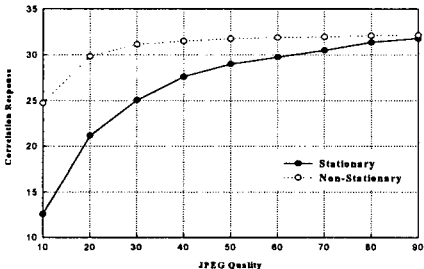


그림 7. JPEG 압축률 변화에 따른 유사도 비교

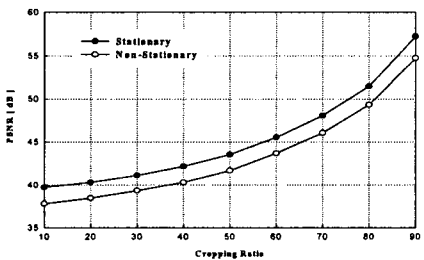


그림 8. JPEG 압축률 80%에서 Cropping을 변화에 따른 PSNR 비교

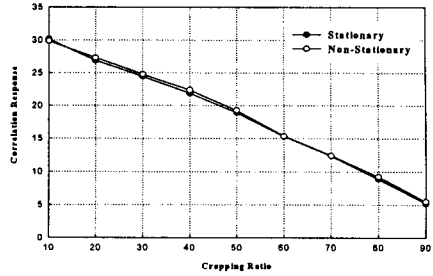


그림 9. JPEG 압축률 80%에서 Cropping을 변화에 따른 유사도 비교



(a) 비정상상태 모델 (b) 정상상태 모델

그림 10. Sirmark benchmark test영상

표 1. Lena 영상에 대한 여러 가지 공격 실험

Lena 512×512				
구분	비정상상태 모델		정상상태 모델	
	PSNR	유사도	PSNR	유사도
Sharpening	21.74	16.75	22.50	6.61
Median	30.47	20.90	31.50	8.31
Gaussian	33.23	27.77	34.53	17.40
JPEG Q=50%	35.02	31.75	37.18	29.74
Cropping 50%	49.47	19.31	47.51	19.34