

HVS 모델과 웨이블릿 변환을 이용한 블라인드 워터마킹

정회원 주 상 현*, 이 선 화*

Blind Watermarking Using HVS and Wavelet Transform

Sang-Hyun Joo*, Seon-Hwa Lee* *Regular Members*

요 약

본 논문에서는 인간의 심리시각적 특성을 고려한 HVS(Human Visual System) 모델과 웨이블릿 중간주파수 대역의 계수쌍을 이용한 블라인드 워터마킹 기법을 제안한다. 중간주파수 대역의 계수쌍은 각종 영상처리를 이용한 공격과 압축 공격 등에 대해서 비슷한 왜곡 현상을 보인다. 따라서 웨이블릿 계수를 일정한 크기로 양자화하는 기존 방법들보다 중간주파수대역의 계수쌍 사이의 관계를 이용한 양자화 방법이 보다 안정적인 성능을 나타낸다. 또한 Noise Visibility Function(NVF)을 이용한 HVS 모델을 적용함으로써 워터마크 삽입 후의 비가시성을 확보하였다. 실험을 통하여 이진 워터마크를 삽입한 영상이 평균 44dB 이상의 우수화질을 유지하면서도 각종 영상 처리 공격 및 JPEG 압축 공격에 대해서 견고성이 우수함을 확인하였다

Key Words : HVS(human visual system); wavelet transform; middle frequency pair; blind watermarking.

ABSTRACT

In this paper, we propose a blind watermarking that embeds watermarks into wavelet middle frequency subbands using human visual system. Wavelet middle frequency pairs(MFP) show similar distortion against general image processing attacks such as compression and filtering. So the quantization between MFPs is more robust than conventional methods that directly quantize DWT coefficients. We use a noise visibility function(NVF) to obtain a good visual quality. This is able to preserve embedding positions after many attacks. Our experimental results show that the proposed scheme is robust to various image processing such as JPEG, while preserving good visual quality above 44dB.

I. 서 론

최근, 디지털 콘텐츠에 대한 유료화가 가속화되면서, 이들에 대한 저작권 보호를 위한 연구가 활발히 진행되어 왔다. 디지털 워터마킹^{[1]-[10]}은 가장 널리 사용되는 저작권 보호 기술로서, 저작권자의 고유정

보를 사람의 인지능력으로는 감지하기 어렵도록 삽입한 후, 차후 저작권 분쟁이 발생할 경우 저작권 정보를 다시 추출함으로써 저작권자의 소유권을 입증할 수 있는 기술이다.

특히 강인한 워터마킹(Robust Watermarking) 기법은 영상 또는 동영상의 저작권 보호를 위하여 활발한 연구가 진행되고 있으며, 이러한 기법이 갖추어야

* 한국전자통신연구원 디지털저작권관리연구팀 (joos@etri.re.kr)
 논문번호 : 030405-0917, 접수일자 : 2003년 월 일

할 조건은 다음과 같다^[10]: 1) 비가시성, 2) 손실압축에 대한 강인성, 3) 각종 영상처리에 대한 강인성, 4) 허용 가능한 오검출률.

강인한 워터마킹은 워터마크가 삽입된 데이터에 대하여 공격이 가해진 후에도 워터마크 추출이 강인하도록 설계되어야 한다. 원본 데이터 없이 워터마크를 추출할 수 있는 기법을 블라인드 워터마킹이라고 하며^{[8][11]}, 원본이 필요한 기법을 널블라인드 워터마킹이라고 한다. 블라인드 워터마킹은 원본 데이터를 위한 별도의 저장공간이 필요없고, 공인 인증센터에서의 활용 때문에 최근에는 이 방식에 대한 연구가 활발히 진행되고 있다.

블라인드 워터마킹 기법은 추출 과정에서 원본 데이터를 필요로 하지 않지만, 몇몇 제안된 방법에서는 워터마크가 삽입된 위치 정보나, 임계치 등의 부가적인 정보를 필요로 하는 경우도 있다. 이러한 부가 정보들은 대개 검출 정확성을 높이기 위해서 사용되는데, 이러한 방법은 엄밀히 말하면 세미-블라인드 워터마킹이라고 할 수 있다^[9].

블라인드 워터마킹 기법은 추출시 참조할 데이터가 없기 때문에 활용이 용이한 반면, 공격에 더 취약할 수 있다는 단점이 있다. 데이터 압축 공격^{[12][14]}과 필터링과 같은 영상처리 공격^[15]을 고려한 연구들이 많이 진행되어 왔으며, 블라인드 워터마킹에 있어서 특히 매우 강력한 공격으로 손꼽히는 기하학적 공격, 즉 RST(Rotation, Scaling, Translation) 공격에 강인한 기법에 대한 많은 연구가 진행되고 있다^[16]. RST 공격을 고려한 대부분의 기법들은 DCT(Discrete Cosine Transform), DWT(Discrete Wavelet Transform), Log-Polar Fourier Transform 등을 이용한 주파수 영역에서의 워터마킹 기법을 제안하고 있다.

본 논문에서 제안하는 방식은 불연속 웨이블릿 변환(DWT)의 중간주파수 대역의 계수쌍(MFP : Middle Frequency Pair)에 인간의 시각 시스템(HVS : Human Visual System)을 고려한 NVF 모델을 결합하고, 웨이블릿 계수값의 중요도에 따라 적응적으로 변경되는 양자화 스텝을 사용하여 워터마크를 삽입하는 블라인드 방식의 워터마킹이다. HVS 모델은 널블라인드 워터마킹 기법에서 널리 사용되고 있는 방법으로서 본 논문에서는 이를 블라인드 워터마킹에 적용하고 그 강인성과 비가시성에 대하여 고찰하고자 한다.

본 논문의 구성은 다음과 같다. II장에서는 웨이블릿 중간주파수 대역의 계수쌍에 적응적으로 변경

된양자화 스텝을 이용하여 워터마크를 삽입하는 방법을 제안하고, III장에서는 본 알고리즘에서 사용한 HVS 모델인 NVF기법을 소개한다. IV장에서는 양자화 기법과 HVS 모델을 결합한 워터마크 삽입 및 추출 과정을 제안하며, V장에서는 제안한 알고리즘의 성능 평가를 위해 여러가지 공격을 가한 실험 결과 제시한다. 그리고 마지막으로 VI장에서 결론을 맺는다.

II. 웨이블릿 중간주파수 대역에서의 워터마킹

1. 불연속 웨이블릿 변환

불연속 웨이블릿 변환은 기본적으로 피라미드 알고리즘에 의해 효과적으로 수행되는데, 저주파통과 필터와 고주파통과 필터의 조합을 이용하여 영상을 LL(low-low), LH(low-high), HL(high-low), HH(high-high) 대역으로 분해하게 된다. LL 대역에 한번 더 웨이블릿 변환을 수행하여 한 단계 높은 근사 계수를 구할 수 있으며, 이러한 과정은 필요에 따라 반복적으로 수행될 수 있다. 또한, 웨이블릿 역변환(Inverse DWT)을 이용하여 웨이블릿 계수로부터 원영상을 재구성 할 수 있다.

LH 대역과 HL 대역은 수직방향의 상세정보와 수평방향의 근사정보 또는 그 반대의 정보를 포함하기 때문에 중간주파수 대역이라고 한다. 일반적으로 저주파수 대역인 LL 대역 근사계수에 워터마크를 삽입하는 방식은 공격에는 강인하지만 화질열화의 문제점이 있고, 고주파수 대역인 HH 대역의 상세계수에 삽입하는 방식은 비가시성은 좋지만 공격에 취약하다는 단점이 있다. 따라서, 중간주파수 대역을 선택함으로써 비가시성과 강인성 간의 균형을 맞추어 워터마크를 삽입하고자 한다.

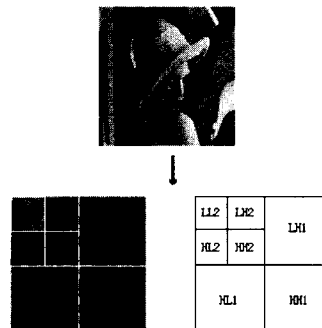


그림 1. 2단계 불연속 웨이블릿 변환
Fig. 1. 2-Level Wavelet Transform

2. 중간주파수 대역에서의 워터마킹

웨이블릿 중간주파수 대역의 계수쌍(MFP : Middle Frequency Pair)은 그림 2에서와 같이 LH와 HL 대역에서 같은 위치의 계수 쌍으로 정의된다.

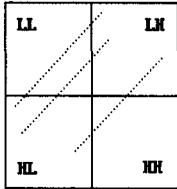


그림 2. 중간주파수 대역의 계수쌍
Fig. 2. Middle Frequency Pair

중간주파수 계수는 각종 영상처리와 압축 등의 공격을 가한 후에도 공격전과 비슷한 계수값을 나타내는데, 이러한 특성은 중간주파수 계수에 워터마크를 삽입할 경우 공격에 대하여 강인성을 유지할 수 있음을 의미한다. 이를 증명하기 위하여 식 (1)과 같은 절대값 비율을 사용하여 공격전후의 중간주파수 계수값의 변화를 살펴보았으며, 웨이블릿 기저로는 Haar 직교 기저를 사용하였다.

$$Ratio(LH(i, j)) = \frac{ABS(LH(i, j)) \text{ after attacks}}{ABS(LH(i, j)) \text{ before attacks}} \quad (1)$$

위 식에서 ABS는 절대값을 의미하며, 공격을 가한 영상에 대하여 웨이블릿 변환을 한 후, 모든 중간주파수 계수값들의 절대값 중 큰 1/2에 해당하는 계수들에 대하여 $Ratio(LH2(i, j))$ 와 $Ratio(HL2(i, j))$ 의 상관관계를 구하였다. 또한, 원영상과 공격이 가해진 영상간의 왜곡 정도를 나타내기 위해 PSNR(Peak Signal-to-Noise Ratio)를 이용하였다. PSNR은 식 (2)와 같이 구할 수 있다.

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \quad (2)$$

MSE(Mean Square Error)는 두 영상간의 평균 제곱 오차를 의미하며, 표 1은 512x512 크기의 Lena 영상에 대하여 실험한 결과이다.

중간주파수 대역의 계수값이 공격 후에도 비슷한 값을 나타낸다는 것은 그러한 계수쌍을 이용하여 양자화를 하는 것이 공격에 강인할 수 있다는 것을 의미한다. 본 논문에서 제안하는 알고리즘의 주요

표 1. 공격 전후의 중간주파수 계수값의 변화율

Table 1. Ratio of LH2 and HL2 bands before and after attacks

Image Operation	PSNR	Average Ratio(LH2(i, j))	Average Ratio(HL2(i, j))
Blurring	35.84	0.8895	0.9049
Sharpening	31.97	1.1198	1.0850
Histogram Equalization	18.70	1.5852	1.5770
JPEG Compression	39.82	0.9570	0.9649
	35.22	0.9010	0.9667

특징은 워터마크 삽입 위치에 해당하는 웨이블릿 계수값의 중요도에 따라 양자화 스텝을 적응적으로 변경하고, 그 양자화 스텝을 사용하여 중간주파수 대역의 각 계수값을 워터마크값에 따라 양자화함으로써 워터마크를 삽입하는 것이다. 또한, 중간주파수 대역의 같은 위치에 해당하는 계수쌍 중 하나의 계수값을 이용하여 다른 하나의 계수값을 양자화하고자 하는 것이다. 이를 위해서 먼저 중간주파수 대역의 모든 계수값들 중에서 그 절대값이 가장 큰 1/4에 해당하는 계수를 '중요계수(significant coefficient)'로 선택하고, 각 계수쌍에서 적어도 둘 중 하나의 계수가 '중요계수'보다 크다면 양자화 스텝은 둘 중 큰 계수를 이용하여 정한다. 그림 3에서와 같이 양자화 스텝은 큰 계수의 1/4로 정하고 그 값으로 작은 계수를 양자화한다. 또한, 계수쌍에서 두 계수의 값이 모두 '중요계수'보다 작다면 작은 값으로 미리 정해놓은 양자화 스텝을 이용하여 두 계수를 모두 양자화한다. 이렇게 양자화를 한 후 워터마크를 삽입하고 추출하는 방법에 대해서는 IV장에서 보다 자세히 기술하도록 한다.

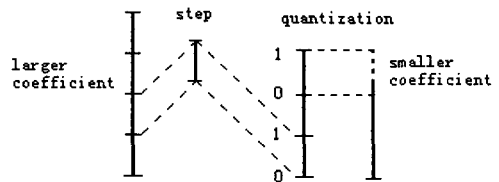


그림 3. 중간주파수 계수의 양자화
Fig. 3. The quantization for middle frequency pair

III. HVS 모델

워터마크 삽입 후에도 워터마크의 비가시성을 유

지하기 위하여 HVS(Human Visual System) 모델을 적용하였다. 워터마크가 삽입된 영상의 질은 대부분 PSNR 수치를 이용하여 측정하지만, 이는 인간의 심리시각적(psycho-visual) 특성을 고려하지 못하는 측정 방법이다. 또한, 단순히 LH와 HL 대역의 계수에 워터마크 정보를 삽입한다면 영상 왜곡이 눈에 띄게 나타날 수 있기 때문에 HVS 모델을 이용하여 비가시성을 향상시키고자 하였다. 본 논문에서는 참고문헌 [16]에 제안된 NVF(Noise Visibility Function)에 기반한 perceptual model을 적용하였다. NVF의 입력은 2단계 웨이블릿 변환에서의 LL2 영역이며, 출력은 λ 이다. 예를 들어 영상의 크기가 512×512일 경우 LL2, λ , NVFAcpt의 크기는 모두 128×128이다. NVF(noise visibility function)는 영상에 잡음(noise)을 넣었을 때, 잡음이 보이는 정도를 나타내는 함수로써, 영상의 지역적 영역에 따라 다른 값을 가지며 그 값은 0과 1사이의 값을 갖는다. 즉, 영상의 변화가 없는 평탄영역에서 NVF는 1이고, 윤곽 부근과 같은 변화가 많은 영역에서는 0에 근접한다. NVF를 이용한 마스크는 라플라시안 필터에 비해 노이즈 제거 공격에 강하고 워터마크가 삽입된 영상의 화질의 척도인 PSNR(Peak signal to noise ratio)을 구할 수 있는 특징을 가지고 있다. 먼저 nvf는 식 (3)에 의해 구할 수 있다. 여기서, $\sigma_x^2(i,j)$ 는 지역별 분산(local variance), $\sigma_{x_{max}}^2$ 는 영상에서의 최대 지역별 분산(maximum local variance)을 나타낸다.

$$nvf(i, j) = \frac{1}{1 + \theta \sigma_x^2(i, j)}, \quad \left(\theta = \frac{D}{\sigma_{x_{max}}^2} \right) \quad (3)$$

인지시각 마스크 λ 는 식 (4)에 의해 계산된다. 이 식에서 S_0 와 S_1 은 워터마크의 강도를 나타낸다. 평탄영역에서는 nvf의 세기(S_1)는 약하게, 변화가 심한 지역 성분($1-nvf$)의 세기(S_0)는 크게 설정하여 인지시각 마스크를 구한다.

$$\lambda = S_0(1 - nvf) + S_1 \cdot nvf \quad (4)$$

이렇게 구해진 λ 에 대하여 식 (5)를 이용하여 워터마크를 강하게 삽입할 위치를 구하고, 식 (6)에 따라 워터마크 삽입을 위한 적응적으로 변경되는 양자화 스텝을 구한다. 여기서, 임계치(threshold)는 전체 λ 의 평균값을 나타내며, MFP(i, j)는 중간주파수대역의 계수쌍을 의미한다. 또한, S는 상대적으로

워터마크를 약하게 넣기 위하여 작은 값으로 미리 정해놓은 양자화 스텝이다.

$$NVFAcpt(i, j) = \begin{cases} 1 & \lambda(i, j) > threshold \\ 0 & otherwise \end{cases} \quad (5)$$

$$Quantize_step(i, j) = \begin{cases} Maxcoef/4 & NVFAcpt(i, j) = 1 \text{ and} \\ S & \text{one of MFP}(i, j) > \text{significant coefficient} \\ & otherwise \end{cases} \quad (6)$$

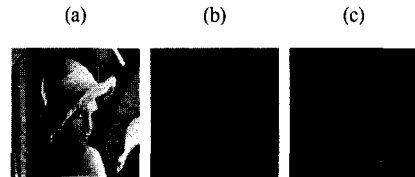


그림 4. NVF와 '중요계수'에 의해 산출된 양자화 크기 비교
(a) 원영상
(b) 워터마크 삽입시 산출된 양자화 스텝 영상
(c) 워터마크 추출시 산출된 양자화 스텝 영상

그림 4는 워터마크 삽입 및 추출시 계산되는 양자화 스텝을 정규화된 영상으로 나타낸 것으로서 워터마크 삽입 전후의 양자화 스텝이 거의 비슷하게 나타남을 알 수 있다. 또한, JPEG 압축이나 기타 영상처리 공격 등을 가한 후에도 비슷하게 유지됨을 알 수 있다.

IV. 워터마크 삽입 및 추출 알고리즘

1. 워터마크 삽입 방법

1) 1 단계 : 워터마크열의 생성 및 웨이블릿 분해
워터마크는 사용자키로부터 생성되는 1 또는 0의 값을 가지는 무작위 순열(random sequence)이며, 원영상의 웨이블릿 분해는 2단계까지 수행한다. 단계 2의 중간주파수대역인 LH2, HL2 대역에 워터마크를 삽입하는데, LH2의 크기가 $m \times n$ 이고, 워터마크열의 길이가 l 이라고 할 때, 워터마크를 적어도 $\left\lfloor \frac{m \times n}{l} \right\rfloor$ 번 만큼 반복적으로 삽입함으로써 강인성을 향상시킬 수 있다.

2) 2 단계 : 적응적 양자화 스텝의 결정 및 워터마크 열 삽입

LH2와 HL2 대역의 모든 웨이블릿 계수들 중에서 그 절대값이 가장 큰 1/4에 해당하는 값을 '중요 계수(significant coefficient, T)'로 선택한다. 식 (5)에서 인지시각 마스크에 의해 확인된 $NVFAcpt(i,j)$ 의 값이 1이고 LH2와 HL2의 계수쌍 중에서 어느 한 계수의 값이 '중요계수' T보다 크게 되면, 큰 계수의 절대값의 1/4을 양자화 스텝으로 정하고 이를 이용하여 작은 계수를 나눈다.

그 외의 경우 즉, $NVFAcpt(i,j)$ 의 값이 0이거나 계수쌍의 두 계수의 절대값이 모두 '중요계수' T보다 작을 경우는 작은 값으로 설정된 양자화 스텝(S)으로 두 계수를 모두 나눈다. 이러한 양자화 스텝(S)은 '중요계수'의 1/4 이하의 정수 값으로 선정하는 것이 바람직하며, 실험적으로 '중요계수'가 보통 15 내지 30정도이므로 양자화 스텝(S)을 2 내지 5의 값으로 선정하였다.

양자화 스텝으로 나눈 결과에 대해 삽입할 워터마크의 값이 1이면 가장 가까운 홀수로 반올림하고, 0인 경우에는 가장 가까운 짝수로 반올림한 후, 이 양자화된 값에 다시 각각의 양자화 스텝을 곱함으로써 워터마크 값에 따라 대상 계수값을 변경한다

```

For all the LH2 and HL2 coefficients
If (abs(LH2(i,j))<T && abs(HL2(i,j))<T) || NVFAcpt(i,j)==0
    Quantize LH2(i,j) and HL2(i,j) by S;
Else
    Maxcoef = max(abs(LH2(i,j)), abs(HL2(i,j)));
    Step = Maxcoef / 4;
    If Maxcoef == abs(LH2(i,j))
        Quantize abs(HL2(i,j)) by Step;
    Else
        Quantize abs(LH2(i,j)) by Step;
    End if
End if
End for
    
```

3) 3 단계 : 워터마킹된 영상의 생성

중간주파수 대역 LH2와 HL2에 워터마크열이 반복적으로 삽입되면, 전체 계수에 대하여 역웨이블릿 변환을 수행함으로써 워터마킹된 영상을 생성한다.

2. 워터마크 추출 방법

본 논문에서 제안하는 워터마크 추출방법은 블라인드 방식으로서, 추출시 워터마크가 삽입된 영상과 사용자키 만을 입력으로 하여 워터마크의 존재여부를 판별한다. 다음은 워터마크의 추출과정을 단계별로 서술하고 있다.

1) 1 단계 : 웨이블릿 분해

워터마크 삽입시와 마찬가지로 워터마크가 삽입된 영상에 대하여 2단계 웨이블릿 변환을 한 후, 단계 2의 중간주파수대역인 LH2, HL2 대역으로부터 워터마크를 추출한다. 한 번 이상 반복적으로 추출되는 각 워터마크값을 평균함으로써 보다 신뢰성있는 워터마크열을 산출할 수 있다.

2) 2 단계 : 양자화 스텝의 결정 및 워터마크열 추출

삽입시와 동일한 방법으로 LL2 대역의 계수들로부터 $NVFAcpt(i,j)$ 를 구하고 LH2와 HL2 대역의 웨이블릿 계수들 중에서 '중요계수' T를 구한다. 그리고 각 계수쌍의 절대값과 '중요계수'의 값을 비교하여 워터마크열 W^* 를 추출한다.

계수쌍의 절대값이 모두 '중요계수'보다 작거나 $NVFAcpt(i,j)$ 가 0일 경우 약속된 작은 값의 양자화 스텝(S)으로 각각의 계수값을 나누고, 그 결과에 대해 2로 나눈 나머지 값을 평균하여 워터마크를 추출한다. 이때, 반올림을 적용하여 워터마크를 0 또는 1의 정수값으로 추출한다. 반면, 둘 중 어느 한 계수값이 '중요계수'보다 크고 $NVFAcpt(i,j)$ 가 1일 경우는 큰 계수값의 1/4을 양자화 스텝으로 정하고 그 값으로 작은 계수값을 나눈 후 이에 대해 다시 2로 나누어 그 나머지 값을 워터마크로 추출한다.

이 때, 워터마크열이 반복 삽입된 경우, 그 반복 횟수만큼 워터마크열을 추출한 후 동일 위치의 각 비트를 평균함으로써 보다 신뢰성 있는 워터마크열을 추출할 수 있으며, 단순 평균이 아닌 각 비트의 양자화 스텝을 기중치로 두어 평균값을 계산한다.

```

For all the LH2 and HL2 coefficients
If (abs(LH2(i,j))<T && abs(HL2(i,j))<T) || NVFAcpt(i,j)==0
    W*(i,j) = (LH2(i,j)/S mod 2 + HL2(i,j)/S mod 2) / 2;
Else
    Maxcoef = max(abs(LH2(i,j)), abs(HL2(i,j)));
    Step = Maxcoef / 4;
    If Maxcoef == abs(LH2(i,j))
        W*(i,j) = HL2(i,j) / Step mod 2;
    Else
        W*(i,j) = LH2(i,j) / Step mod 2;
    End if
End if
End for
    
```

3) 3 단계 : 워터마크의 유효성 판별

추출된 워터마크 W^* 의 유효성을 판별하기 위하여 유사도 측정기법을 사용한다. 추출된 워터마크열 W^*

와 원래 삽입된 워터마크열 W 와의 유사도는 식 (7)에 의하여 구할 수 있다.

$$sim(W, W^*) = \frac{W^* \cdot W}{W^* \cdot W^*} \quad (7)$$

여기서, $W^* \cdot W$ 는 $\sum W_i^* W_i$ 이며, W_i^* , $W_i \in \{-1, +1\}$ 이다. 위 식을 이용하여 유사도를 계산하기 위하여 삽입 또는 추출된 워터마크열을 -1과 +1값을 가지는 열로 만들기 위하여 0을 -1로 바꾸어준다.

위 식에 의하여 계산된 유사도가 소정의 임계값보다 크면 워터마크, 즉 저작권이 존재하는 것으로 판단할 수 있는데, 유사도 판정의 임계값은 참고문헌 [17]에서 제안하는 서로 다른 키에 의해 발생된 워터마크열의 유사도들에 대한 표준편차(σ)의 6배에 해당하는 값으로 정하였다. 본 논문에서는 1000개의 서로 다른 키에 의해서 생성되는 워터마크열과 추출된 워터마크열에 대한 유사도를 각각 계산하여 그로부터 산출된 표준편차의 6배에 해당하는 값을 임계값으로 산정하였으며, 추출시 입력된 사용자키에 의해 생성된 워터마크열과 실제 추출된 워터마크열에 대한 유사도가 그 임계값보다 클 경우 추출 성공으로 간주하였다.

V. 실험 결과

실험에 사용된 영상은 USC-SIPI Image Database에서 제공하는 영상들이며, 크기별로 256×256 13개, 512×512 32개, 1,024×1,024 3개의 8bit 그레이 영상들을 사용하였고, 워터마크열의 길이는 1,024로 설정하였다. 고정된 양자화 스텝 S 는 3으로 설정하였으며, 원영상에 대한 워터마크 삽입 영상의 평균 PSNR은 44.59로서 워터마크가 비가시적으로 삽입됨을 알 수 있었다.

표 2는 워터마크 영상에 대한 공격 후 워터마크를 추출한 결과를 나타낸다. Blurring 공격은 가우시안 스무딩 필터($\sigma = 0.5$, 3×3)를 이용하였고, sharpening 공격은 식 (8)과 같이 원영상으로부터 저주파 통과 필터링한 결과를 감하는 unsharp 필터 ($\alpha = 0.2$, 3×3)를 사용하였으며, 공격을 가한 영상의 정량적인 수치를 나타내기 위하여 PSNR을 측정하였다.

$$\frac{1}{(\alpha+1)} \begin{bmatrix} -\alpha & \alpha-1 & -\alpha \\ \alpha-1 & \alpha+5 & \alpha-1 \\ -\alpha & \alpha-1 & -\alpha \end{bmatrix} \quad (8)$$

표 2에서의 임계치는 검출 성공여부를 판단하기 위한 임계치로서, 서로 다른 키에 의하여 계산된 유사도들의 표준편차(σ)의 6배에 해당하는 값이며, 추출된 워터마크와 삽입된 워터마크간의 유사도가 임계치보다 클 경우 검출성공으로 간주하였다. 또한, PSNR, 임계치 및 검출 유사도에 해당하는 값은 모든 테스트 영상에 대한 각각의 결과값을 평균한 값이다. Blurring의 경우 모든 영상이 워터마크 검출에 성공하였으며, JPEG 압축 공격의 경우 Quality 30 정도에서 높은 검출률을 나타내었다. 반면, sharpening 공격을 할 경우, 영상의 고주파성분이 강조되므로 중간주파수 대역의 계수의 절대값이 전체적으로 상승하게 되며, 이에 따라 '중요계수'가 삽입시보다 높게 산출된다. 따라서, 이 경우 '중요계수' 값이 커짐에 따라 양자화 스텝의 변화가 커져 검출성공률이 낮게 나타남을 알 수 있었으며, sharpening 후 다시 blurring을 가한 경우에는 고주파 성분이 blurring 과정을 통하여 제거되므로, 중간주파수 대역 계수들의 절대값과 '중요계수' 값이 다시 작아지게 되어 양자화 스텝의 변화가 작아져 검출성공률이 더 높게 나타난다. 표 3은 sharpening 공격 후에만 검출에 실패한 대표적인 5개의 영상에 대한 중요계수의 변화를 보여주고 있으며, 모두 sharpening 공격시 중요계수의 값이 가장 큰 것을 알 수 있다.

표 2. 공격에 대한 검출 성능

Table 2. Detection Performance for the attacks

공격	PSNR (dB)	평균 임계치	평균 검출 유사도	검출 성공률 (검출성공계수 / 실험영상개수 × 100)	
Blurring	37.15	0.190	0.942	100 %	
Sharpening	20.42	0.188	0.166	36.2%	
Sharpening + Blurring	25.46	0.188	0.322	78.7%	
JPEG compression	QF 15	29.34	0.187	0.156	31.9%
	QF 20	30.25	0.187	0.225	48.9%
	QF 25	30.95	0.188	0.327	74.5%
	QF 30	31.51	0.187	0.395	85.1%
	QF 35	32.0	0.187	0.459	91.5%

표 3. 공격 후 중요계수의 변화
Table 3. Significant coefficient changes after attacks

삽입시	추출시	공격후 추출시		
		Blurring	Sharpening	Sharpening + Blurring
13	14.5	14	23.5	20
9.5	10.5	10.25	26.75	21
32	33.25	31.25	58	49.5
9.75	10.5	11.75	20.5	16.25
38.25	39	35.5	70.25	59.25

그림 5는 Lena 영상으로부터 추출된 워터마크와 1000개의 서로 다른 키로부터 생성된 워터마크와의 유사도를 보여준다. 삽입된 워터마크는 키값 500으로부터 생성된 것이므로 500의 위치에서 가장 높은 피크가 검출되며, 이 값이 임계값(0.19, 가로선)보다 크므로 검출 성공으로 판단한다. 그림 6은 Quality factor 25로 JPEG 압축 공격을 가한 영상에 대한 유사도 측정 결과로서, 공격을 가하지 않은 영상에 비하여 유사도가 낮게 측정되지만 임계값보다 높으므로 검출에 성공한 경우이다.

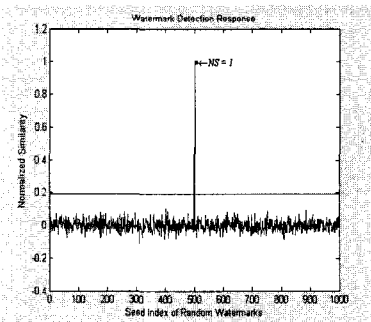


그림 5. 워터마크가 삽입된 Lena 영상에서의 워터마크 검출 결과
Fig. 5. Watermark Detection Result from Watermarked Lena Image

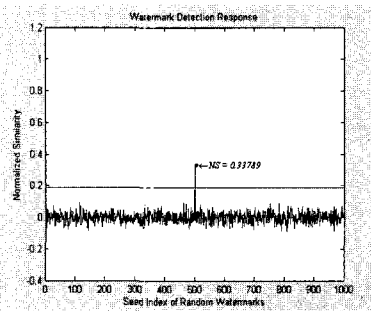


그림 6. JPEG 공격 후 워터마크 검출 결과 (QF 25)
Fig. 6. Watermark Detection Result after JPEG attack (QF 25)

VI. 결론 및 향후과제

본 논문에서는 웨이블릿 중간주파수 대역의 계수 쌍 과 HVS 모델을 결합한 블라인드 워터마킹 기법을 제안하였다. 중간주파수 대역은 영상처리 공격과 압축 공격 등에 대해서도 비슷한 계수값을 나타내기 때문에 제안한 양자화 방법은 이러한 공격에 강인함을 나타내며 비가시성과 강인성이라는 상반된 요구사항을 적절히 만족시킬 수 있는 특성을 가진다. 그리고 본 논문에서 사용된 NVF를 이용한 HVS 모델은 공격 후에도 중요한 계수를 대부분 그대로 보존함을 알 수 있었다. 또한 실험을 통하여 우수한 화질을 유지하면서도 공격에 대하여 강인함을 확인하였다. 그러나 웨이블릿 변환 자체가 기하학적 변환에 매우 민감하기 때문에 본 알고리즘은 RST 공격에는 약한 단점을 가지고 있다. 따라서 RST 공격에도 강인한 워터마킹 기법에 대한 연구가 향후 진행되어야 할 것이다.

참고 문헌

- [1] 서용석, 주상현, 정호열, "JPEG2000을 위한 디지털 워터마킹", 한국방송공학회 논문지 6권 1호, pp. 32-40, 2001년 6월
- [2] 김철기, 차의영, "웨이블릿과 특이치 분해에 기반한 디지털워터마킹", 한국통신학회 논문지, 27권 6호, pp. 602-609, 2002년 6월
- [3] A. Nikolaidis, I. Pitas, "Region-based image watermarking," IEEE Transactions on Image Processing, Volume: 10 Issue: 11 Page(s): 1726 -1740, Nov. 2001.
- [4] Jiwu Huang, Y.Q. Shi, Yi Shi, "Embedding image watermarks in dc components," IEEE Transactions on Circuits and Systems for Video Technology, Volume: 10 Issue: 6, Page(s): 974 979. Sept, 2000.
- [5] Ming-Shing Hsieh, Din-Chang Tseng, Yong-Huai Huang, "Hiding Digital Watermarks Using Multiresolution Wavelet Transform," IEEE Transactions on Industrial Electronics, Volume: 48, No. 5, Pages: 875-882, October 2001.
- [6] C-Y Lin, M. Wu, J.A. Bloom, I.J. Cox, M.L. Miller, Y.M. Lui, "Rotation, scale, and translation

resilient watermarking for images," IEEE Transactions on Image Processing, Volume: 10 Issue: 5, Page(s): 767 -782, May 2001.

[7] Chun-Hsien Chou, Yun-Chin Li, "A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," IEEE Transactions on Circuits and Systems for Video Technology, Volume: 5 Issue: 6 , Page(s): 467 476, Dec. 1995.

[8] N Kaewkamnerd, K.R. Rao, "Wavelet based watermarking detection using multiresolution image registration," Proceedings, TENCON 2000, Volume: 2,Page(s): 171 -175 vol.2, 2000.

[9] D Kundur, D Hatzinakos, "Semi-blind image restoration based on telltale watermarking,"Conference Record of the Thirty-Second Asilomar Conference on Signals, Systems & Computers, Volume: 2, Page(s): 933 937, 1998.

[10] Chion-Ting Hsu, Ja-Ling Wu, "Multiresolution watermarking for digital images," IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, Volume: 45, Issue: 8, Page(s): 1097 1101, Aug. 1998.

[11] Chun-Hsiang Huang, Ja-Ling Wu, Ding-Yun Chen, "A blind watermarking algorithm with semantic meaningful watermarks," Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers, Volume: 2,Page(s): 1827 -1830, 2000.

[12] Ikpyo Hong, Intaek Kim, Seung-Soo Han, "A blind watermarking technique using wavelet transform," Proceeding of IEEE International Symposium on Industrial Electronics, Volume: 3, Page(s): 1946 -1950, 2001.

[13] E. Koch, J. Zhao, "Towards Robust and Hidden Image Copyright Labeling," Proceeding of IEEE Workshop on Nonlinear Signal and Image Processing, pp: 456-459, Neos Marmaras, Greece, June 20-22, 1995.

[14] J.J. Eggers, J.K Su, B. Girod, "Robustness of a blind image watermarking scheme," Proceedings. 2000 International Conference on Image Processing, vol. 3, Page(s): 17 20, 2000.

[15] Gwo-Jong Yu, Chun-Shien Lu, Hong-Yuan

Mark Liao, Jang-Ping Sheu, "Mean quantization blind watermarking for image authentication," Proceedings of 2000 International Conference on Image Processing, vol. 3, Page(s): 706 709, 2000.

[16] S. Voloshynovskiy, A. Herrigel, N. Baumgaertner, and T. Pun, "A stochastic approach to content adaptive digital image watermarking", In Lecture Notes in Computer Science: Third International Workshop on Information Hiding, Vol. 1768, pp. 211-236, Sep. 1999.

[17] I.J.Cox, J.Kilian, F.T. Leighton, and T.Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. Image Processing, vol.6, no.12, pp.1672-1687, Dec.1997.

주 상 현(Sang-Hyun Joo)

정회원



1989년 동국대학교 전자공학과 공학사

1994년 동국대학교 전자공학과 공학석사

1999년 일본니이가타대학 자연과학연구과 공학박사

1994년~1996년:

생산기술연구원 연구원

1999년~2001년 니이가타대학 전기전자공학과 조교수

2001년~현재 한국전자통신연구원 가상현실연구부 선임연구원

<주관심분야> 저작권 관리/보호, 영상/비디오 처리

이 선 화(Seon-Hwa Lee)

정회원



2000년 2월 : 부산대학교 전자계산학과 졸업

2002년 2월 : 부산대학교 전자계산학과 석사

2002년 3월~현재 : 한국전자통신연구원 가상현실연구부 연구원

<주관심분야> 영상/신호처리, 워터마킹, 핑거프린팅