논문 2004-41SP-6-1

영상제거 공격에 강인한 LBX 인터리빙 워터마킹 방법

(An LBX Interleaving Watermarking Method with Robustness against Image Removing Attack)

고 성 식*, 김 정 화*

(Koh Sung Shik and Kim Chung Hwa)

요 약

디지털 미디어와 통신 네트워크의 급속한 발전으로 적절한 지적소유권(IPR) 보호 기술인 데이터 인증방법에 대한 필요성이 절실히 요구되고 있다. 본 논문에서는 정보량이 많은 워터마크를 삽입하여도 워터마크가 삽입된 영상의 화질을 열화시키지 않고, 특히 영상의 일부가 제거되는 공격에서 강인성을 갖기 위해, 제안한 선형비트확장(LBX) 인터리빙을 이용하여 마킹 공간인 이산 웨이브렛 변환(DWT)영역의 자주파수 계수에 그레이스케일 워터마크 로고를 삽입하는 새로운 영상 워터마킹 기술을 제안하였다. 실험결과 영상절단과 영상회전 등과 같은 영상의 일부가 제거되는 공격에 대해서 특히 높은 강인성을 가짐을 검증하였다.

Abstract

The rapid growth of digital media and communication networks has created an urgent need for self-contained data identification methods to create adequate intellectual property right(IPR) protection technology. In this paper we propose a new watermarking method that could embed the gray-scale watermark logo in low frequency coefficients of discrete wavelet transform(DWT) domain as the marking space by using our Linear Bit-eXpansion(LBX) interleaving of gray-scale watermark, to use lots of watermark information without distortion of watermarked image quality and particularly to be robust against attack which could remove a part of image. Experimental results demonstrated the high robustness in particular against attacks such as image cropping and rotation which could remove a part of image.

Keywords : gray-scale watermark, bit-expansion interleaving, wavelet transform

I. 서 론

인터넷을 통한 멀티미디어 데이터 배포가 보다 쉽고 빠르게 이용되고 있고, 전자상거래 분야와 온라인 서비스도 역시 빠르게 성장하고 있다. 그렇지만 이러한 발전 이면에는 디지털 컨텐츠 분쟁의 잠재성이 내포되어 있다. 디지털 워터마킹은 제한된 접근이 아니더라도 데이터에 직접 사용자나 다른 소유권 정보를 삽입함으로서 소유권 분쟁 시 이를 해결하기 위한 하나의 방편이다.

워터마킹의 종류는 영상의 소유권을 주장할 수 있도록 눈에 잘 띄는 메시지나 회사 로고를 영상에 포함하는 시각적 워터마킹 방법과, 눈에 띄지 않게 삽입하므로 워터마크 정보가 삽입된 영상과 원영상이 동일하게 보이는 비시각적 워터마킹 방법이 있다. 영상에 워터마크를 비시각적으로 삽입하는 방법 중 공간영역 워터마킹 방법은 일반적으로 영상 컨텐츠의 LSB를 수정하지만 저역통과 필터와 같은 연산에 강인성이 매우 약하다는 단점이 있다[1][2]. 그래서 대부분의 워터마킹 방법은 퓨리어[3], DCT[4]나 웨이브렛 변환[5][6]과 같은 변환영역의 데이터를 수정하여 워터마크 정보를 입력한다. 그렇지만 이 방법 역시 컨텐츠 일부가 제거되는 영상절단이나 회전공격에 대해서 강인성이 약하다는 문제점이 있다. 게다가 디지털 워터마크 데이터는 데이터 페이로드로 초래하는 시각적 화질열화 때문에 원영상의 데이터

* 정회원, 조선대학교 전자공학과
(Dept. of Electronics, Chosun University)
접수일자: 2004년3월13일, 수정완료일: 2004년7월16일

에 비해 아주 적은 데이터 양을 갖는 2진 영상 또는 2진 데이터 값을 주로 이용하고 있다. 본 방법은 워터마크된 영상의 화질을 최소화하면서 정보량이 많은 그레이 스케일 사진 워터마크를 이용할 수 있고, 영상을 75%까지 제거하는 공격에 대해서도 디인터리버에 의해 워터마크의 화질왜곡을 최소로하여 추출할 수 있는 기술을 제안한다.

본 논문의 구성은 다음과 같다. Ⅱ장에서는 제안한 LBX 인터리빙에 의해 워터마크를 삽입하는 방법을 살펴보고, Ⅲ장에서는 공격에 대한 LL_1강인성을 유지하면서 워터마크를 추출할 수 있는 디인터리빙 과정을 기술한다. Ⅳ장에서는 제안한 방법의 성능 검증을 위해 실험 영상으로 시뮬레이션하고 5장에서 결론과 활용방안에 대해서 제시한다.

Ⅱ. 제안한 선형비트확장(LBX) 인터리빙에 의한 워터마크 삽입 방법

인터리빙은 통신용어로서 송신하는 반복되는 비트열을 섞어서 재배분하면 공격으로 인해 데이터 비트에 비트에러가 발생하더라도 인터리빙의 역과정(디인터리빙)을 통하여 배열된 데이터를 원래 데이터 비트열로 배열하여 반복 패턴의 유무를 조사하여 손실된 비트열을 원래 비트열로 복구할 수 있는 방식이다. 워터마킹 시스템 역시 외부 잡음이나 불법적인 디지털 저작물의 변형은 삽입된 워터마크 데이터에 비트 오류가 발생되므로 본 논문의 워터마킹 시스템에서는 통신상의 인터리빙 기법을 응용해서 새로운 선형비트확장 인터리빙을 적용한 워터마킹 기법을 제안하였다. 이렇게 인터리빙을 워터마킹 시스템에 적용하면 집중된 비트 오류가 발생하더라도 디인터리빙으로 집중된 비트 오류를 모두 분산시켜 독립된 오류로 바꿈으로써 워터마크의 오류 정정 효율을 높일 수 있다.

제안한 인터리빙 방법은 Chae와 Manjunath[7]의 워터마킹 방법에서 이용한 선형비트확장 식(1)을 응용하였다. 비트열을 선형비트 확장을 하면 확장된 비트열은 일정한 패턴의 규칙성을 가지고 반복하게 된다는 점을 응용해서 워터마크를 인터리빙하였다.

$$W_N(x,y) = (2^N-1)\left(\frac{W(x,y)-W_{\min}}{W_{\max}-W_{\min}}\right) \qquad (1)$$

여기서 $W_N(x,y)$은 워터마크의 화소 $W(x,y)$를 선형 N비트확장한 결과이고, W_{max}과 W_{min}은 워터마크를 구성하는 화소값 중 최대값과 최소값이다.

제안한 인터리빙식은 화소당 n 비트 그레이 스케일 워터마크일 때 $W_{\max}$가 2^n-1이고 $W_{\min}$ 가 zero이므로 식(1)을 다음 식(2)과 같이 유도시킬 수 있다.

$$\begin{aligned} W_N(x,y) &= \frac{(2^N-1)}{(2^n-1)}W(x,y) = \frac{(2^{nK}-1)}{(2^n-1)}W(x,y) \\ &= [1+2^n+2^{2n}+2^{3n}+\cdots+2^{(K-1)n}]\cdot W(x,y) \\ &= \left[\sum_{k=1}^{K}2^{n(k-1)}\right]\cdot W(x,y) \\ &= R_n \cdot W(x,y), \quad K=\text{정수} \end{aligned} \qquad (2)$$

인터리빙을 위한 선형비트확장 식(2)을 이용해 워터마크 화소 당 n 비트를 (N=nK)로 선형비트 확장하면 n 비트가 동일한 패턴으로 K번 반복하게 된다. 이때 R_n을 n 비트 반복 인자라 정의한다. 즉 화소 당 8비트(n=8)인 256 그레이 스케일 워터마크 화소를 화소마다 32비트(N=nK=32)로 비트 확장(W_N=A_1,A_2,A_3,A_4)하면, 확장된 비트 열은 A_1=A_2=A_3=A_4가 성립된다. 이때 32비트의 반복인자 R_{32}은 10진수 "65793" 값이 되고 2진수로는 "00000001000000010000000100000001" 값을 갖는다. 따라서 MSB에서 LSB까지 동일하게 반복되는 4개 8비트열을 얻을 수 있다. 이러한 원리를 이용해 워터마크(128×128)를 선형 32비트 선형비트확장하고 그림 1의 배열방법으로 새로운 공간(256x256)으로 각각 재배치하여 인터리빙을 하였다.

인터리빙된 워터마크 크기는 인터리빙전의 워터마크 크기보다 **4**배 크기가 되어서 원영상의 1레벨 웨이브렛 변환 영역 중 저주파 LL밴드 마킹공간과 동일한 크기가 된다.

그림 2는 워터마크가 삽입된 영상을 얻기 위한 처리

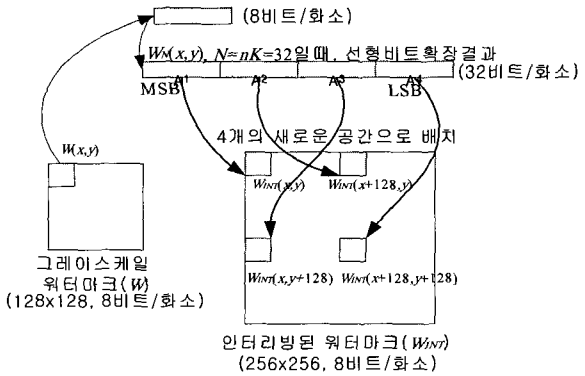


그림 1. 선형비트확장 인터리빙 절차
Fig. 1. Linear bit-expansion interleaving sequence.

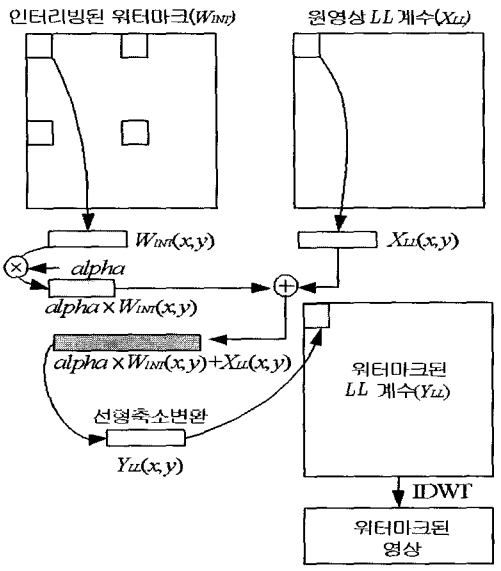


그림 2. 워터마크 삽입 절차
Fig. 2. Watermark Insertion Sequence.

절차를 나타낸다. 워터마크가 삽입된 저주파 계수(Y_{LL})는 위해 인터리빙된 워터마크를 원영상의 저주파 계수에 다음 식(3)을 이용하여 삽입한 후 선형축소변환 하여 얻을 수 있다.

$$Y_{LL}(x,y) = scale_{min(X_{LL})}^{max(X_{LL})}\{X_{LL}(x,y) + alpha \times W_{INT}(x,y)\} \quad (3)$$

여기서 $X_{LL}(x,y)$은 원영상의 웨이브렛 변환 영역 중 저주파 영역인 LL밴드 계수이고, $W_{INT}(x,y)$은 인터리빙된 워터마크이다. $alpha$는 워터마크의 삽입 강도를 결정하는 파라미터이고 시각적으로 화질열화를 인지할 수 없는 범위로 설정한다.

원영상의 웨이브렛 저주파 계수에 워터마크 정보를 삽입하게 되면 저주파 계수의 최대값과 최소값의 변화가 발생하게 된다. 이를 그대로 역웨이브렛 변환하여 워터마크가 삽입된 영상을 얻게 되면 화질열화가 많이 발생하게 된다. 따라서 워터마크된 영상의 화질을 최소화하기 위해 워터마크를 삽입한 저주파 계수를 삽입 전의 최대값과 최소값의 스케일로 선형변환 한다. 또한 워터마크를 추출할 때 역시 워터마크가 삽입된 저주파 계수의 원래 크기로 선형확장변환할 때 이용된다. 이때 선형변환을 하므로 데이터 자체 변형은 없다. 그래서 워터마크의 변형은 없게된다. 마지막으로 식(3)에 의해 얻어진 결과를 역웨이브렛 변환하면 워터마크가 삽입된 영상을 얻을 수 있다.

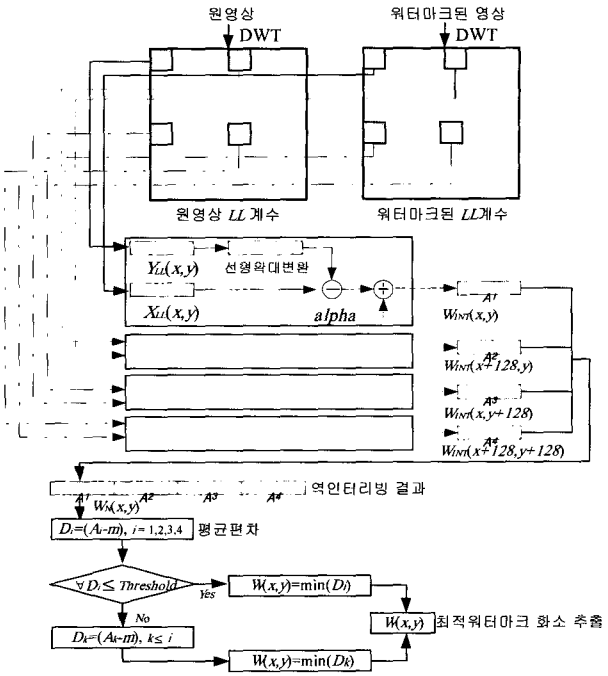


그림 3. 디인터리빙을 이용한 워터마크 추출
Fig. 3. Watermark extraction using de-interleaving.

Ⅲ. 워터마크 추출을 위한 디인터리빙 방법

삽입된 워터마크 추출은 원영상을 필요로 한다. 추출 순서는 워터마크를 삽입하는 절차의 역으로 그림 3과 같이 처리한다. 워터마크 추출을 위한 첫 번째 과정은 워터마크가 삽입된 영상과 원영상을 각각 웨이브렛 변환하여 저주파 영역 계수를 각각 얻는다. 다음은 삽입의 역 과정으로 인터리빙 시 재배치된 4곳으로부터 식(4)을 이용해 화소값(A_1,A_2,A_3,A_4)을 얻는다.

$$W_{INT}(x,y) = \frac{Y_{LL}(x,y) - X_{LL}(x,y)}{alpha} \quad (4)$$

추출된 A_1,A_2,A_3,A_4은 외부공격으로 인해 정보가 변형될 수 있기 때문에 최적의 워터마크 화소 선택은 제안한 디인터리빙 과정으로 다음과 같이 진행한다.

먼저 추출된 A_1,A_2,A_3,A_4의 평균편차(D_i)를 구한다. 평균 편차는 산술 평균값에서 각 측정치의 편차에 대한 절대치를 산술 평균하여 분산도를 나타내는 것으로 A_1,A_2,A_3,A_4의 평균 값(m)을 이용해 식(5)과 같이 나타낼 수 있다.

$$D_i = |A_i - m|, \quad i = 1,2,3,4 \quad (5)$$

그리고 식(5)을 이용해 표준편차(σ)를 계산한다. 표준편

차는 편차 제곱의 평균에 대해 그 제곱근을 계산함으로써 분산도를 나타내는 것으로 식(6)과 같이 나타낼 수 있다.

$$\sigma = \sqrt{\frac{1}{4}\sum_{i=1}^{4} D_i^2} \tag{6}$$

정규분포의 경우 표준편차는 사례들(cases)의 68.27%가 $m-\sigma$와 $m+\sigma$사이에 포함되고, 95.45%가 $m-2\sigma$와 $m+2\sigma$사이에 포함되고, 99.73%가 $m-3\sigma$와 사이에 포함되는 성질을 가진다.

제안한 방법에서는 추출된 A_1,A_2,A_3,A_4의 데이터 오류 발생 유무를 결정하는 역치값(Threshold)은 $m-\sigma$와 $m+\sigma$사이의 68.27%로 설정하였다. 만약 추출된 4개 데이터의 평균 편차 D_i가 모두 역치값보다 적다면 최소 평균편차값($\min\{D_i\}$)을 갖는 화소값이 최적 워터마크 데이터로 선택된다. 그러나 역치값보다 큰 편차가 존재하면 외부 공격에 의해 손실이 큰 것으로 간주하여 그때 화소값을 제외한 나머지 화소값($A_k|k\leq i$)으로 다시 평균 편차(D_k)를 구해서 최소 평균 편차값($\min\{D_k\}$)의 화소값이 최적 워터마크 화소값으로 선택된다.

또한 영상이 무작위로 제거되어도 LBX 인터리빙으로 4개의 워터마크 정보를 생성시켰기 때문에 나머지 25%의 영상정보만으로도 워터마크를 복원시킬 수 있다.

Ⅳ. 실험영상에 대한 시뮬레이션 결과

이 장에서는 다양한 실험영상(512x512, 8비트/화소)을 이용하여 제안한 워터마킹 방법의 강인성 평가를 한다. 그림 4(a)는 Lenna 실험영상이고 그림 4(b)는 가중치(*alpha*)가 0.02일 때 워터마크가 삽입된 영상이다. 워터마크 로고(128x128, 8비트/화소)는 그림 5(a)의 컨텐츠 소유자 사진을 이용하였고 그림 5(b)는 그림 4(b)의 Lenna 실험영상으로부터 공격을 하지 않은 상태에서 제안한 방법으로 추출한 워터마크이다. 추출된 워터마크의 화질열화는 웨이브렛 변환과정에서 유한어장 때문에 미소하게 발생되었지만 시각적 손실은 거의 없다.

유사성 평가는 크기가 $2^n\times2^n$이고 256 그레이레벨인 영상일 때 $PSNR$ 식 (7)을 이용한다.

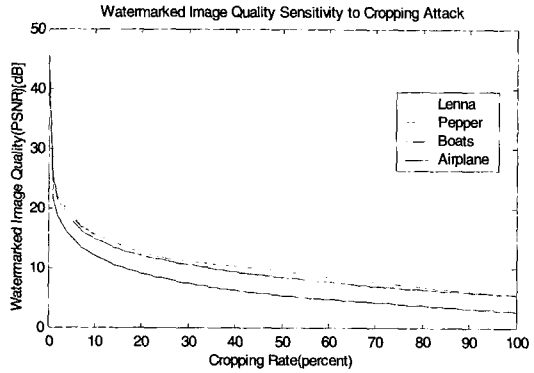
$$PSNR = 10\log\left\{\sum_{i=0}^{2^n-1}\sum_{j=0}^{2^n-1}\frac{255^2}{[x(i,j)-y(i,j)]^2}\right\} \tag{7}$$

(a) Lenna영상 (b) 워터마크된 영상(40.08dB)

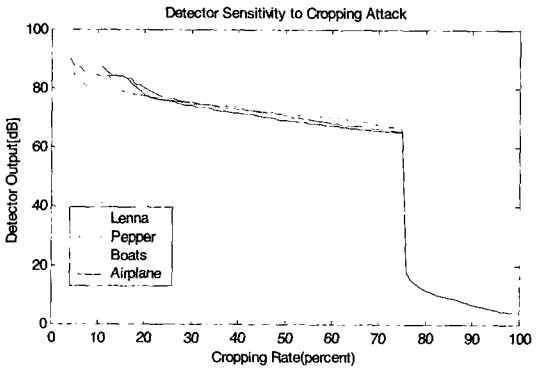
그림 4. 실험영상과 워터마크된 영상(512x512, 8bit/pixel)
Fig. 4. Test image and watermarked image.

(a) 워터마크 (b) 추출된 워터마크(87.26 dB)

그림 5. 워터마크와 추출된 워터마크(128x128, 8bit/pixel)
Fig. 5. Watermark and extracted watermark.



(a) 워터마크된 영상의 화질열화

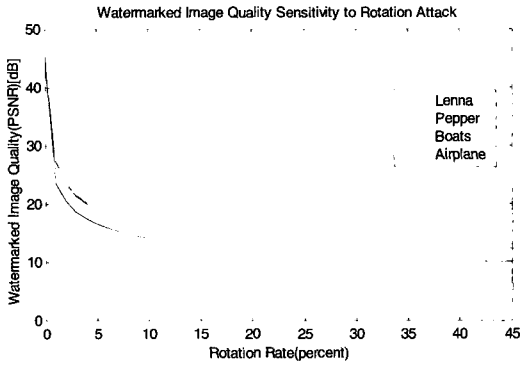


(b) 워터마크 검출기 민감도

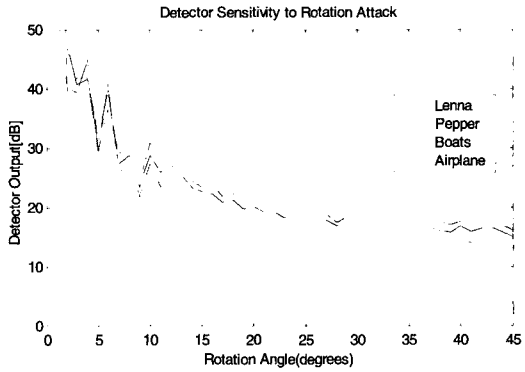
그림 6. 영상절단 공격에 대한 워터마크 추출
Fig. 6. Watermark extraction from cropping attacks.

여기서 $x(i,j)$는 원영상의 i, j 좌표의 화소값이고 $y(i,j)$는 평가 대상 영상의 화소값이다.

제안한 방법의 우수성을 평가하기 위해서 다양한 실험영상을 영상절단, 영상회전, 척도변환, 그리고 손실

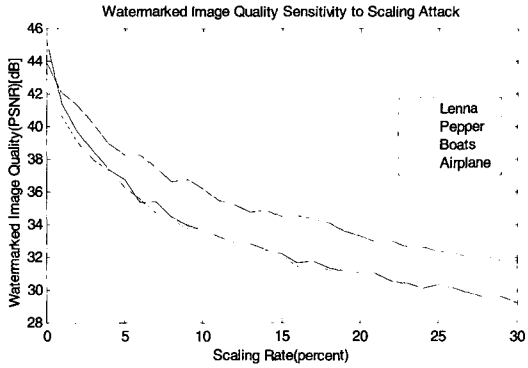


(a) 워터마크된 영상의 화질열화

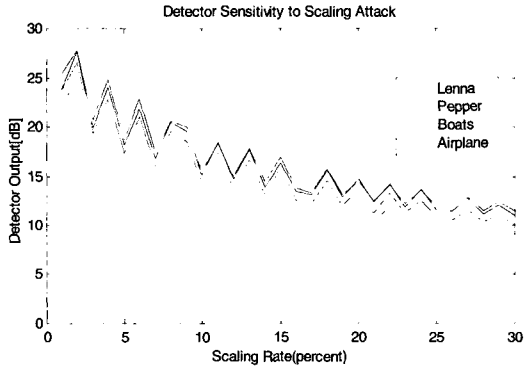


(b) 워터마크 검출기 민감도

그림 7. 영상회전 공격에 대한 워터마크 추출
Fig. 7. Watermark extraction from rotation attacks.



(a) 워터마크된 영상의 화질열화



(b) 워터마크 검출기 민감도

그림 8. 영상척도변환 공격에 대한 워터마크 추출
Fig. 8. Watermark extraction from scaling attacks.

JPEG압축 공격으로 실험를 하였다. 그림 6(a)은 워터마크가 삽입된 영상을 영상 절단 공격을 했을 때 화질열화를 나타내는 그래프이고 그림 6(b)은 절단 공격으로부터 추출된 워터마크의 화질을 나타낸 그래프이다. 영상절단이 75%일 때 8dB이하 실험 영상의 화질열화 조건에서도 워터마크는 60dB 이상으로 추출할 수 있었다.

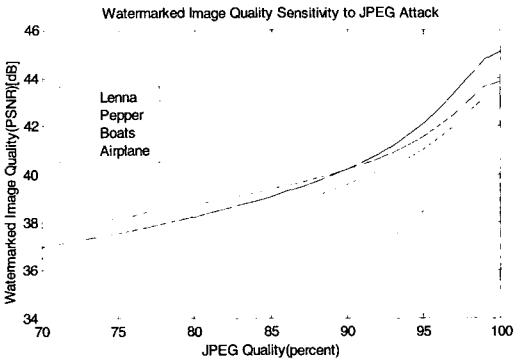
그림 7(a)은 워터마크가 삽입된 영상을 영상 회전 공격을 했을 때 화질열화를 나타내는 그래프이고 그림 7(b)은 회전 공격으로부터 추출된 워터마크의 화질을 나타낸 그래프이다. 영상회전은 영상의 일부를 제거시킬 뿐만 아니라 화소값을 변형시키는 공격이다. 회전에 따른 화소의 디지털 좌표 격자 구조의 변화는 회전각도에 비례하게된다. 영상회전이 20도일 때 18dB이하 실험 영상의 화질열화 조건에서도 워터마크는 20dB 이상으로 추출할 수 있었다.

그림 8(a)은 워터마크가 삽입된 영상을 척도변환 공격을 했을 때 화질열화를 나타내는 그래프이고 그림 8(b)은 척도변환 공격으로부터 추출된 워터마크의 화질을 나타낸 그래프이다. 척도변환 역시 영상의 일부를 제거시키면서 화소값을 변형시키는 강력한 공격이다. 그렇지만 척도변환은 주변화소를 이용하여 손실된 격자 구조 선상의 실험영상 정보를 복원시키기 때문에 워터마크 정보가 실험영상보다 더욱 제거시킨다. 척도변환이 10%일 때 35dB이하 실험 영상의 화질열화 조건에서 워터마크는 20dB 이상으로 추출할 수 있었다.

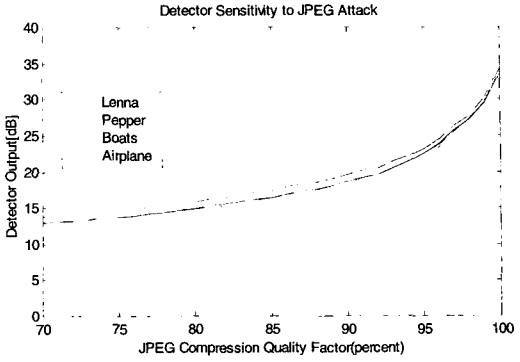
그림 9(a)는 워터마크가 삽입된 영상을 손실 JPEG압축 공격을 했을 때 화질열화를 나타내는 그래프이고 그림 9(b)는 압축 공격으로부터 추출된 워터마크의 화질을 나타낸 그래프이다. 모든 실험 영상에 대해서 JPEG quality 80%일 때 30dB 이하의 실험영상 화질열화 조건에서 15dB이상의 워터마크를 추출할 수 있었다.

그림 10은 "Lenna" 실험영상에 대한 시뮬레이션 결과의 예를 나타낸다. 특히 영상절단공격으로부터 추출된 워터마크가 강인성이 견고하다는 것을 알 수 있다.

따라서 본 논문에서 제안한 방법은 워터마크된 영상의 화질을 최소화하면서 워터마크 페이로드를 증가시킬 수 있고, 특히 영상의 일부를 제거하는 공격에 대해서 디인터리버에 의해 추출된 워터마크의 화질을 유지시킬 수 있었다. 그렇지만 화소의 변형을 가져오는 공격에 대해서는 추출된 워터마크의 화질이 다소 감소하였다.



(a) 워터마크된 영상의 화질열화



(b) 워터마크 검출기 민감도

그림 9. JPEG 압축 공격에 대한 워터마크 추출
Fig. 9. Watermark extraction from JPEG attacks.

7.10dB

65.83dB

(a) 영상절단 공격(70%)과 추출된 워터마크

17.28dB

28.76dB

(b) 영상회전 공격(10도)과 추출된 워터마크

34.02dB

17.46dB

(c) 영상척도변환 공격(10%)과 추출된 워터마크

34.79dB

15.64dB

(d) JPEG압축 공격(80%)과 추출된 워터마크

그림 10. 다양한 공격으로부터 추출된 워터마크
Fig. 10. Watermarks extracted from various attacks.

V. 결 론

본 논문에서는 LBX 인터리빙으로 그레이 스케일 워터마크 로고를 삽입할 수 있는 새로운 워터마킹 방법을 제안하였다. 정보량이 많은 워터마크 로고를 삽입을 위해 웨이브렛 변환영역을 마킹공간으로 이용하였고, 외부공격으로부터 워터마크의 강인성을 유지하기 위해 인터리빙과 디인터리빙 과정을 이용하였다. 시뮬레이션 결과, 특히 영상의 일부가 제거 공격에 대해서는 성능이 우수하였다. 영상이 75%이상 제거되는 공격에 대해서는 워터마크를 복원시킬수 없는 한계점이 있었지만 나머지 영상은 정보로서의 가치가 적기 때문에 논쟁의 의미가 없다. 따라서 본 논문에서 제안한 알고리즘은 공격에 대한 높은 강인성을 유지, 정보량이 많은 워터마크 이용 등에서 다양한 멀티미디어 저작권 보호에 활용될 수 있으리라 사료된다.

참 고 문 헌

[1] W. Bender, D. Gruhl, and N. Morimoto, "Techni -ques for Data Hiding," in Proc. of SPIE, Stora -ge and Retrieval for Image and Video Database III, vol. 2420, pp.164-173, San Jose, Feb. 1995.

[2] R. G. van Schyndel, A. Z. Tirkel and C. F. Osbor -ne, "A Digital Watermark," in Proc. of IEEE Int. Conf. Image Processing, vol. 2, pp.86-90, Austin, Nov. 1994.
[3] J. J. K. O Ruanaidh, W. J. Dowling, and F. M. Boland, "Phase Watermarking of Digital Images," in Proc. of IEEE Int. Conf. Image Processing, vol. 3, pp. 239-242, Switzerland, Sept. 1996.
[4] A. Piva, M. Barni, F. Bartolini, and B. Cap -pellini, "DCT-based Watermark Recovering Without Resorting to the Uncorrupted Original Image," in Proc. of IEEE Int. Conf. Image Processing, pp. 520-52, Santa Barbara, CA, Oct. 1997.
[5] M.Barni, F.Bartolini, V.Cappellini, A.Lippi, A.Piva, "A DWT-based Technique for Spatio-Frequency Masking of Digital Signatures," in Proc. of SPIE Security and Watermarking of Multimedia Contents, vol.3657, pp.31-39, California, January 1999.
[6] D. Kundur and D. Hatzinakos, "Digital Water -marking using Multiresolution Wavelet Decom -position," in Proc. of IEEE Int. Conf. Acoustics, Speech, Signal Processing, vol. 5, pp. 2969-2972, Seattle, WA, May 1998.
[7] Chae J. J., and Manjunath B.S., "A Robust Embedded Data from Wavelet Coefficients", in Proc. of SPIE EI'98, vol. 3312, pp. 308- 317, San Jose, Feb. 1998.

저 자 소 개

고 성 식(정회원)
1994년 조선대학교 전자공학과 (공학사).
1996년 조선대학교 전자공학과 (공학석사).
2002년 조선대학교 전자공학과 (공학박사).
2003년~현재 조선대학교 전자공학과 겸임교수.
<주관심분야: 지능형 교통시스템관련 운전지원시스템, 영상압축, 신호처리, 영상 워터마킹>

김 정 화(정회원)
1979년 조선대학교 전자공학과 (공학사).
1981년 조선대학교 전자공학과 (공학석사).
1991년 숭실대학교 전기공학과 (공학박사).
1979년~현재 조선대학교 전자공학과 교수.
<주관심분야: 영상처리, 신호처리, 신호처리 및 시스템, 영상 워터마킹, 유비쿼터스>