

신경망을 이용한 강인한 디지털 이미지 워터마킹 알고리즘

박성일, 한승수
 명지대학교 정보공학과

Robust Digital Image Watermarking Algorithm Using Neural Network

Cheng-Ri Piao and Seung-Soo Han
 Myongji University, Department of Information Engineering

Abstract - 본 논문에서는 디지털 영상의 소유권 보호를 위하여 양자화기법과 신경망을 적용하여 기존의 방법보다 강인한 워터마킹 기법을 제안하였다. 제안한 워터마킹 알고리즘은 시간영역에서 양자화 기법을 사용하여 워터마크를 삽입하고 추출하였고, back-propagation neural network(BPN)을 사용하여 워터마크를 검출하였다. 실험결과 압축공격에 강인하며, PSNR이 41dB이상이므로 비가시성을 만족함을 알 수 있었다.

1. 서 론

최근 데이터의 디지털화와 멀티미디어 기술의 발전 및 인터넷의 보급과 더불어 디지털 데이터가 갖는 복제의 용이성으로 인해 저작자의 소유권 보호와 인증에 대한 문제가 중요시되고 있다. 이에 따라 디지털 데이터에 워터마크를 삽입하여 소유권을 보호하고 데이터의 무결성을 보증하도록 하는 연구가 활발히 진행되고 있다. 워터마킹이란 디지털 데이터에 인간이 감지할 수 없는 정보를 삽입하는 기술이다[1]. 이러한 워터마크는 저작권 보호, 자료인증, 소유권 증명 등을 효율적으로 하기 위하여 권한이 없는 불법 사용자에게 시각적, 그리고 통계적으로 검출되지 않아야 한다. 그리고 다양한 신호처리(손실압축, 필터링, 양자화)와 기하학적 변형(회전, 확대 및 축소, 잘라내기, 변환)에도 워터마크의 검출이 가능해야 한다[2,3].

현재 존재하는 대부분의 워터마킹 알고리즘들은 영상을 DWT, DCT, FFT 등 주파수 영역으로 변환한 후에 워터마크를 삽입하고 역변환을 하여 워터마크된 영상을 얻고, 검출시에는 다시 주파수 영역으로 변환한 후에 검출한다[4,5]. 주파수 영역에서 워터마킹은 여러 가지 공격에 강인하나 연산 시간이 많이 걸린다. 그리고 시간 영역에서 워터마킹 기법은 연산 시간이 적으나 공격에 강인하지 않다.

본 논문에서는 여러 가지 공격에 강인한 양자화 기법을 이용하여 강인성을 보장하는 알고리즘을 제안하였다. 그리고 BPN신경망을 이용하여 삽입, 검출하였기 때문에 검출율과 비가시성이 우수함을 실험결과 알 수 있었다.

2. 신경망과 워터마킹 알고리즘

본 논문에서는 신경망을 이용하여 워터마크를 삽입, 추출하였다. 그리고 워터마크를 삽입할 때 여러 가지 공격에 강인한 양자화 기법을 이용하였다.

2.1 BPN신경망

현재 연구되고 있는 디지털 워터마킹은 워터마크를 삽입, 추출하기 위한 방법이나 응용의 목적에 따라 여러 가지로 분류할 수 있는데 그 중에서 워터마크 추출 시

원영상의 사용 여부에 따라서 원 영상 없이 추출 가능한 blind 워터마킹과 원영상과 워터마크영상 둘 다 있어야 워터마크가 추출 가능한 non-blind 워터마킹 기법이 있다. non-Blind 워터마킹 방법은 워터마크 검출이 용이하나 원 영상을 항상 필요로 하기 때문에 활용범위가 제한적인 문제점이 있다. 때문에 강인성이 상대적으로 떨어진 원본 영상 없이 워터마크된 영상만으로 워터마크를 검출하는 blind 워터마킹 방법이 더 유용하다. 본 논문에서는 신경망을 이용하여 원영상의 일정한 특징을 학습시켜서 워터마크 검출 시 사용하는 방법을 개발하였다.

본 논문에서 BPN신경망을 이용하였다. 신경망은 3개의 층으로 되어있는데, 입력층, 은닉층, 출력층이 있다. 아래의 그림 1은 신경망 구조이다.

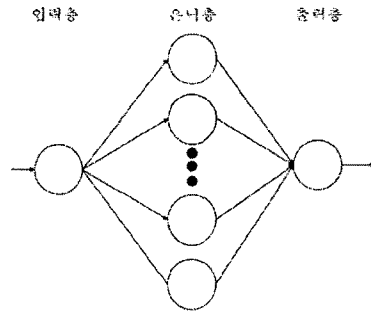


그림 1. BPN 신경회로망 구조

본 논문에서 신경망을 학습시키는 과정은 아래 그림 2와 같이 한다. 그림에서 C는 DCT계수이며, Q는 양자화 값이고, P는 신경망의 입력 값이며, T는 신경망의 출력 값이다. 이렇게 학습시킨 BPN은 아래에 워터마크 삽입과 추출과정에 이용된다.

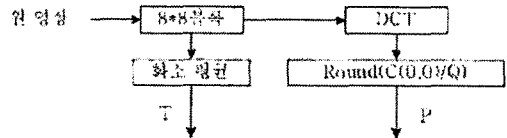


그림 2. BPN training 과정

본 논문에서 사용한 신경망 구조는 1-65-1이며, 신경망의 입력층에서는 비선형전달함수(sigmoid)를 사용하였고, 은닉층에서는 선형전달함수를 사용하였고, 훈련은 Levenberg-Marquardt규칙을 이용하였다. 그리고 학습오차를 0.001하였고, 최대 학습 횟수는 1000으로 하였다. 그림(3)은 256*256 크기의 lena 영상을 학습할 때 각 단계별 학습오차 그래프이다.

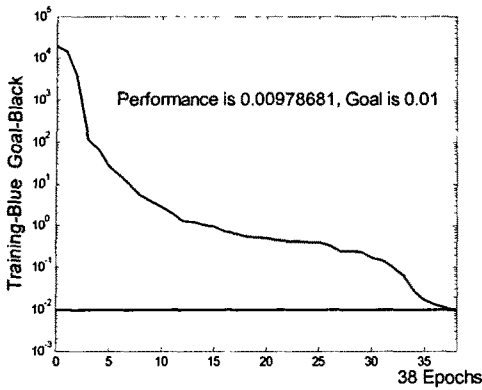


그림 3. 각 단계별 학습오차 그래프

2.2 제안한 워터마킹 알고리즘

워터마킹 알고리즘은 일반적으로 워터마크 생성, 워터마크 삽입, 워터마크 추출 등 3단계로 구성 된다. 본 논문에서는 육안으로 쉽게 식별 가능한 이진 비트맵을 워터마크로 삽입하였다.

2.2.1 워터마크 삽입

아래의 그림(4)은 워터마크 삽입 블록도이다.

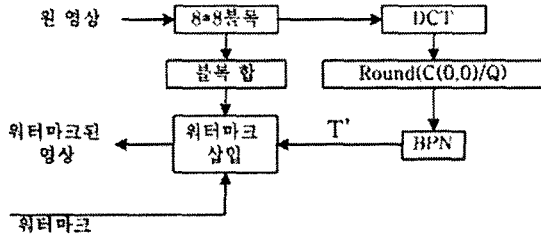


그림 4. 워터마크 삽입 블록도

워터마크 삽입과정은 아래와 같다.

Step1: 원 영상을 8*8로 나누어 블록 합과 DCT를 계산한다.

Step2: DCT계수에서 DC성분을 Q로 양자화 하여 BPN에 입력하여 출력 T'를 얻는다.

Step3: 출력 T'와 블록합을 가지고 아래의 수식을 이용하여 워터마크된 영상을 얻는다.

$$\Delta P_1 = 0.25 * 8 * Q + 64 * T' - \sum_{i,j=0}^7 p_{i,j} \quad (1)$$

$$\Delta P_0 = -0.25 * 8 * Q + 64 * T' - \sum_{i,j=0}^7 p_{i,j} \quad (2)$$

$$p'_{i,j} = \begin{cases} p_{i,j} + \text{round}(\Delta P_1 * p_{i,j} / \sum_{i,j} p_{i,j}) & \text{if } w_{m,n} = 1 \\ p_{i,j} + \text{round}(\Delta P_0 * p_{i,j} / \sum_{i,j} p_{i,j}) & \text{if } w_{m,n} = 0 \end{cases} \quad (3)$$

위 수식에서 $p_{i,j}$ 는 원 영상의 8*8블록의 픽셀을 표시하며, $p'_{i,j}$ 는 워터마크된 픽셀을 표시한다. 그리고 $w_{m,n}$ 는 워터마크를 표시한다.

2.2.2 워터마크 추출

워터마크 추출과정은 삽입과정의 역 과정으로서 아래의 그림(5)와 같이 추출한다.

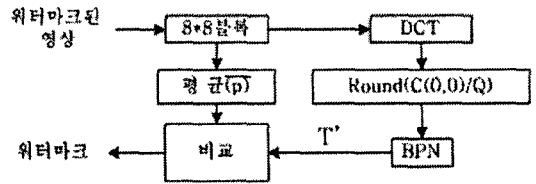


그림 5. 워터마크 추출 블록도

워터마크를 추출 할 때 아래의 수식을 이용한다.

$$w'_{m,n} = \begin{cases} 1 & \text{if } \bar{p} \geq T' \\ 0 & \text{else} \end{cases} \quad (4)$$

3. 실험

본 논문에서는 원 영상으로 256*256 크기의 lena 영상을 사용하였고, 워터마크는 시각적으로 쉽게 인지 할 수 있는 이진영상을 사용하였다. 그림(6)은 실험에 사용한 원 영상과 워터마크 영상을 나타낸다.



(a) 원영상

(b) 워터마크

그림 6. 원영상과 워터마크

영상의 훼손 정도를 보기 위하여 본 논문에서는 PSNR값을 이용하였다. 여기서 PSNR을 다음과 같이 계산한다.

$$PSNR(dB) = 10 \log_{10} \left(\frac{255^2}{\frac{1}{MN} \sum_{i,j=1}^M \sum_{i,j=1}^N (p'(i,j) - p(i,j))^2} \right) [dB] \quad (5)$$

아래의 그림(7)은 양자화 값과 PSNR의 관계를 보여준다. 실험의 결과로부터 PSNR이 41dB이상으로서 화질의 열화를 육안으로 쉽게 구별할 수 없는 수치이다.

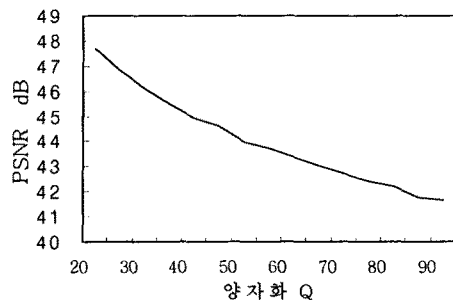


그림 7. 양자화Q와 PSNR 관계

알고리즘의 강인성을 실험하기 위하여 워터마크가 삽입된 영상을 JPEG으로 압축한 후 워터마크를 검출하였다. 아래의 표(1)는 양자화 Q의 변화에 따라 압축률을 변화하면서 육안으로 워터마크를 보고 저작권을 증명할 수 있을 때 압축률과 검출률을 나타낸다. 여기에서 압축률은 아래의 수식(5)을 사용하였다.

$$\text{압축률} = \frac{\text{원영상파일크기}}{\text{압축영상파일크기}} \quad (5)$$

표 1. 양자화 Q, 압축률 및 검출률 관계

Q	압축률	검출률
25	9.0092	90.4766
30	9.56	93.8477
35	10.0854	90.9648
40	11.6662	89.0117
45	12.2452	91.1055
50	12.7515	90.3789
55	13.8117	89.8438
60	14.5034	89.4258
65	15.2749	93.2617
70	15.6188	90.332
75	16.0787	89.0625
80	16.4479	91.8945
85	17.5346	93.8477
90	18.1115	89.3555

아래의 그림(8)에서 (a)는 양자화 값 Q=70일 때 워터마크된 영상을 보여주며, (b)는 JPEG으로 15.6의 압축률로 압축하였을 때 검출된 워터마크를 보여준다.



(a) 워터마크된 영상 (PSNR=42.5dB) (b) 검출된 워터마크 (검출률 90%)

그림 8. 워터마크된 영상과 검출된 워터마크

실험결과로부터 압축 후 검출률이 89%이상으로서 육안으로 쉽게 워터마크를 증명할 수 있음을 볼 수 있다.

4. 결 론

본 논문에서는 영상의 소유권 정보보호를 위하여 양자화 기법과 신경망에 기반한 새로운 디지털 워터마킹 알고리즘을 제안하였다. 제안한 알고리즘은 신경망을 사용하여 원 영상의 일정한 정보를 학습시키고 이 정보를 사용하여 워터마크를 삽입하고 추출한다. 실험결과 PSNR은 41dB 이상으로서 화질이 열화를 육안으로 감지할 수 없으며, 또한 높은 압축률에도 검출률이 89%이상으로서 육안으로 쉽게 워터마크를 증명할 수 있음을 보여주었다.

감사의 글

본 연구는 과학기술부 및 한국과학재단의 ERC 프로그램을 통한 지원으로 이루어졌으며 이에 감사를 드립니다.

[참 고 문 헌]

- [1] S. Craver, N. Memon, N. Yeo, and M. Yeung, "Can Invisible Watermarks Resolve Rightful Ownership?", IBM Research Report, RC 20509, July 25, 1996
- [2] I. J. Cox and M. L. Miller, "A review of watermarking and the importance of perceptual modeling", Proc. SPIE Conf. on Human Vision and Electronic Imaging II, Vol. 3016, Feb. 1997, pp. 92-99.
- [3] I. J. Cox, J. Kilian, T. Leighton and T. Shamoan, "Secure Spread Spectrum Watermarking for Multimedia", IEEE Trans. On Image Processing, Vol. 6, No. 12, 1997, pp. 1673-1687.
- [4] H. A. Peterson, A. J. Ahumada, Jr., and A. B. Watson, "Improved detection model for DCT coefficient quantization," in Proc. SPIE Conf. Human Vision, Visual Processing, and Digital Display IV, Feb. 1993, vol. 1913, pp. 191-201.
- [5] Kundur, D.; Hatzinakos, D. "A robust digital image watermarking method using wavelet-based fusion" Image Processing, 1997. Proceedings., International Conference on Volume 1, 26-29 Oct. 1997 Page(s):544 - 547 vol.1