원본정보 없이 씰영상의 추출이 가능한 이미지 워터마킹 기법

김 원 겸[†] · 이 종 찬[††] · 이 원 돈[†††]

요 약

디지털 이미지 기법과 디지털 네트웍의 출현으로 예술적 작품의 복사가 더욱 쉬워지고 있다. 이러한 창작품을 보호하기 위해 데이터 안에 저작권을 표시할 수 있는 표식이나 인식 가능한 데이터를 삽입하는 기술이 필요해지고 있고 지난 몇 년간 디지털 이미지나 오디오, 비디오 등의 멀티미디어 데이터에 저작권을 표시하기 위한 데이터나 기타 다른 정보를 삽입할 수 있는 많은 기법들이 제안되어 왔다. 본 논문에서는 이미지의 주파수 영역에 인식 가능한 패턴을 삽입하고 추출하는 워터마킹 기법을 제안한다. 또한 삽입된 워터마크를 원본이미지의 정보 없이 추출할 수 있도록 하여 임의의 사람이 워터마크 된 이미지로부터 가짜원본을 생성하는 것이 어렵도록 한다. 원본정보 없이 워터마크를 추출하기 위해서 화소의 원래 값을 예측하는 방법을 사용한다. 예측기법은 구하고자 하는 화소의 주위값을 평균한다는 것을 의미한다. 부가적으로 워터마크를 이미지의 주파수 영역에 삽입함으로써 JPEG같은 손실압축방법에도 견딜 수 있도록 한다.

A Watermarking Scheme to Extract the Seal Image without the Original Image

Won-Gyum Kim[†] · Jong-Chan Lee[††] · Won-Don Lee[†††]

ABSTRACT

The emergence of digital imaging and digital networks has made duplication of original artwork easier. In order to protect these creations, new methods for signing and copyrighting visual data are needed. In the last few years, a large number of schemes have been proposed for hiding copyright marks and other information in digital image, video, audio and other multimedia objects. In this paper, we propose a technique for embedding the watermark of visually recognizable patterns into the frequency domain of images. The embedded watermark can be retrieved from the decoded sequence without knowledge of the original. Because the source image is not required to extract the watermark, one cannot make the fake original that is invertible to watermarking scheme from the watermarked image. In order to recover the embedded signature data without knowledge of the original, a prediction of the original value of the pixel containing the information is needed. The prediction is based on a averaging of amplitude values in a neighborhood around the pixel itself. Additionally the proposed technique could survive several kinds of image processings including JPEG lossy compression.

1. 서 론

관용적 암호 방식은 암호화(encryption)된 데이터에 대해 키를 소유한 사람에게만 정상적으로 접근이 허용된다. 따라서 키를 소유한 사람이 암호화 된 데이터를 복호화(decryption) 해서 얻은 데이터를 임의로 배포하거나 불법적으로 사용한다면 복호화된 이후부터는 전혀 소유권 보장을 할 수 없게 된다. 따라서 관용적 암호 방식은 데이터에 대해 약간의 보호밖에 할 수 없다. 그런 암호 방식을 보완하기 위한 것이 워터마킹

※ 본 논문은 SOREC과 한국과학재단과제(99-11-01-02-A-2)의 연구비 지원에 의한 것임.
† 준 회 원 : 충남대학교 대학원 컴퓨터과학과
†† 정 회 원 : 충남 청운대학교 인터넷컴퓨터학과 교수
††† 정 회 원 : 충남대학교 컴퓨터과학과 교수
논문접수 : 1999년 5월 4일, 심사완료 : 2000년 11월 24일

(watermarking) 기술이라 할 수 있다. 이런 워터마킹 기술은 또한 data hiding, electronic watermarking, image labeling등으로 불리고 있다.

워터마킹 기술은 판매하거나 배포할 자료들 중, 특히 이미지, 오디오, 텍스트, 비디오 및 멀티미디어 데이터 등에 원래의 소유주를 표시할 수 있는 특정 데이터, 즉 워터마크를 넣어 소유주를 확인할 수 있게 하는 최근의 기법이다. 일단 삽입된 워터마크는 항구적으로 데이터에 계속 남아 있어야 한다. 다시 말해 워터마크가 삽입된 데이터는 워터마크를 추출한 후에도 워터마크를 포함하고 있어야 하며 또 여러 가지 신호처리 후에도 역시 워터마크를 계속 포함하고 있어야 한다. 이는 워터마크가 계속 남아 있어야만 언제든 소유권 확인이 가능해지기 때문이다.

워터마크의 형태는 삽입하는 방식에 따라 다르지만 주로 소유주를 표시하는 특정한 코드나 패턴, 혹은 초기값(seed)을 갖는 무작위비트열(random sequence) 등으로 이루어진다. 즉 일정 길이의 비트 열이나 임의의 초기값으로 만들어진 실수(random number)들 혹은 특정한 그림 등으로 워터마크를 구성한다. 일정 길이의 비트열로 구성된 워터마크는 비트 단위로 이미지상에 삽입되는 형식으로 비트열을 구성한 모든 비트에 대해서 중복 삽입하는 방식으로 이루어진다. 임의의 초기값으로 만들어진 실수들의 순서로 이루어진 워터마크의 경우는 위의 비트열 방식처럼 중복 삽입되지 않고 실수들의 순서대로 혹은 임의의 순서대로 차례로 삽입된다. 이미지로 구성 된 워터마크의 경우는 이미지 안에 이미지를 삽입하는 방식으로 이루어진다. 물론 추출한 워터마크의 형태는 삽입한 워터마크의 형태와 동일하거나 거의 흡사해야 한다.

워터마크는 우선 눈으로 볼 수 있는 (visible) 것과 볼 수 없는 (invisible) 것으로 나뉜다. 일반적으로 눈으로 볼 수 없는 워터마크를 많이 사용하는데 이는 눈에 보이는 것은 우선 불법 공격자에 의해 공격받기 쉽고 또한 없어지기 쉽기 때문이다. 따라서 워터마크는 여러 가지 의도적, 비의도적 공격에 강인한 성질을 가진 눈에 보이지 않는 워터마크를 사용해야 한다. 이미지 워터마크는 여러 가지 이미지 처리 기술에 강인하여야 하며 정확한 저작권 보호를 위해 다음과 같은 특성을 만족하여야 한다.

- **Invisibility**

삽입된 워터마크는 눈에 보이지 않아야 한다. 혹은, 보이더라도 워터마크의 존재가 저작권 보호를 하는데 방해가 되어서 안 된다. 즉, 눈에 전혀 보이지 않거나 보이더라도 제거하지 못하도록 워터마크를 삽입하여야 한다.

- **Lossless**

원본 이미지에 워터마크를 넣은 후라도 원본의 데이터와 비교해 정보의 손실이 없어야 한다. 즉, 이미지에 워터마크를 삽입하고 난 후, 이미지의 일부분이라도 손상이 가거나 혹은 변형됨이 없어야 한다는 것을 의미한다.

- **Security**

각 소유주는 삽입된 워터마크에 대한 비밀키를 가져야 되고 이 비밀키 외에 다른 어떤 것으로도 원래 소유주의 워터마크가 데이터로부터 추출이 불가능해야 한다. 만약 추출이 가능하다면, 원래의 소유주 외에 다른 불법사용자도 신호를 가질 수 있으므로 저작권보호(copyright protection)의 기능을 수행 할 수 없다.

- **Robustness**

의도적, 비의도적인 처리(processing)에도 삽입된 워터마크는 제거되지 않아야 된다. 삽입 방식의 일부를 알고 있고 이미지상의 어떤 위치들에 워터마크가 삽입되어 있다는 것을 모를 때, 워터마크를 제거하거나 파괴하려는 시도에 대해 워터마크의 훼손 혹은 파괴 이전에 자체 이미지의 훼손 정도가 심해서 정상적인 이용이 불가능하게 되어야 한다. 또한 JPEG 혹은 MPEG 등의 손실(lossy) 압축 방법으로 압축해도 워터마크가 없어지지 않아야 한다.

본 논문에서는 위의 조건을 만족하는 워터마크의 삽입/추출을 위해 FFT 변환을 이용한 워터마크 삽입기법을 제안한다. 또한 삽입되는 워터마크의 형태도 인식할 수 있는 패턴이나 로고등을 이용한다. 우선 삽입할 패턴을 비밀키와 합쳐 비트순열로 변환한 다음 이미지를 FFT를 이용해 주파수영역으로 바꾸어 진폭(amplitude)계수를 구한 후 그 값에 워터마크를 삽입한다. 추출시에는 원본이 없이도 추출이 가능하도록 하기 위해 워터마크가 삽입된 이미지로부터 근사원본(approximate original)을 만든다. 제안한 워터마킹 기법은 256x256 크기의 흑백 Lenna영상과 6개의 일반적인 컬러영상에 구현되었으며 견고성 검증을 위해 JPEG 압축

/해제에 대해 실험하였다.

2. 워터마킹 기법

기존의 워터마킹 기법들을 크게 나누면 공간영역(spatial domain)에서의 워터마킹 기법[2, 10, 18]과 주파수영역(frequency domain)에서의 워터마킹 기법[3, 4, 9, 15-17], 그리고 두 영역에 모두 워터마크를 삽입하는 혼합형(Hybrid) 워터마킹 기법[6-8]으로 나눌 수 있다. 공간영역에서의 워터마킹 기법은 화소값(pixel value)에 대한 미세한 변화를 워터마크로 삽입하는 방법으로 삽입과 추출은 간단하지만 손실압축, Filtering같은 이미지 처리에 약하다는 면이 있다. 주파수영역에서의 워터마킹 기법은 이미지 데이터를 주파수 성분의 신호로 변환한 다음 워터마크를 삽입하는 방식으로 일반적으로 DCT, FFT, Wavelet 변환 등을 사용한다. 이 방법은 삽입되는 워터마크 계수들이 데이터의 전 영역으로 분포되기 때문에 제3자로 하여금 영상의 왜곡이나 변형에 보다 강력하다는 장점을 가지나 계수들의 값에 따라 얼룩이나 찌그러짐 같은 이미지 손상이 생길 수 있다. 최근에 많은 연구가 이루어지고 있는 혼합형(Hybrid) 워터마킹 기법은 워터마크를 공간영역이나 주파수영역 모두에 삽입함으로써 각각의 장점을 모두 가지도록 하는 방식이다.

공간영역(spatial domain)에서의 워터마킹 기법으로 Bender[2]가 제시한 방법이 있다. 이는 2가지의 삽입방법으로 이루어지는데, 첫번째는 "패치웍(Patchwork)"라는 통계적 방법이다. 패치웍은 임의로 이미지의 두 점 (a_i, b_i)을 n개를 선택하여 a_i과 b_i의 밝기를 각기 단위 크기씩 증가, 감소시킨다. 이 방법은 이미지의 특정 통계적 성질이 사실이라고 할 때, 선택된 n개의 점의 쌍들의 합의 기대치는 2n이라고 하여 저작권을 확인하는 방식이다. 특히, 이 방식은 모든 밝기 정도가 균등하다. 다시 말해, 휘도(intensity)가 균등하게 분포되어 있다는 가정을 하고 있다. 그러나 실제적으로 이미지에서 휘도가 균등하게 분포되어 있는 것은 드문 경우이다. 또한 이 기법은 단위 크기의 증가 혹은 감소에 의해 휘도(intensity)의 파형이 순간적으로 변하는 현상(jittering)에 약할 수 있고 또한 기하학적인 변환에 아주 민감한 단점이 있다. 두 번째 방식은 "texture block coding" 이라 불리는 방식으로 임의의 texture pattern이 포함된 지역을 비슷한 texture를 가진 지역으로 복사하는 방식이다. 각 texture 지역은 자기 상관(autocorrelation)에 의해 복구된다. 그러나 이 방식의 가장 큰 문제는 큰 영역의 임의의 texture를 포함한 이미지에만 적용할 수 있다는 점이다. 또한 텍스트로 구성된 이미지에서는 사용할 수 없고 오디오에도 적용할 수 없다.

Koch, Rindfrey 와 Zhao[3]는 주파수영역(frequency domain)에서의 이미지 워터마킹 기법으로 두 가지를 제안하였다. 첫 번째 방식은 이미지를 8×8의 블록 단위로 구분하고 각 블록에 대해 DCT(Discrete Cosine Transform)을 계산한다. 한 블록을 임의의 작은 부분으로 나누고 그 작은 부분에서 주파수 3개를 선택한다. 각각이 3개의 주파수로 구성된 18개의 3중 주파수들 중에 하나의 3중 주파수를 선택하여 각각의 상대적 강도(strength)를 1 또는 0으로 만든다. 18개의 3중 주파수들은 8x8의 DCT 블록의 미리 결정된 8개의 주파수들 중에 3에 선택함으로써 구성된다. 이 방식은 Cox, Killan, Leighton과 Shamoon[1]이 제안한 대역 확산 방식과 비슷하다. 그렇지만 시각적 중요성과 상대적 에너지를 고려하지 않은 것이 차이가 난다. 더구나 8개의 주파수들의 계수(coefficient)들의 편차가 작기 때문에 잡음 및 distortion에 취약할 수가 있다. 또한 Cox 등의 방식이 JPEG 압축에 대해 quality factor가 5%까지 워터마크가 강인성을 지닌 반면에 이 방식은 50%까지밖에 지원하지 못하는 단점이 있다. 두 번째 방식은 흑백 이미지에 적용되는 것이며 주파수영역으로의 전환을 하지 않는 특징이 있다. 또한 선택된 각 블록들을 흰 픽셀과 검정 픽셀의 상대적 빈도수로 코딩하는 방식이다. 이 두 가지 방식은 모두 다중 삽입에 의한 공격에 강인(robust)하기 위하여 이미지로부터 임의의 64개의 픽셀을 샘플링 하여 분산된 8×8 블록을 만드는 방법을 제안하였다. 그러나 이 방법으로 생기는 DCT 결과가 원본 이미지에 의해 생기는 DCT 결과와 관련이 없어 이미지에 인공적인 흔적이 남을 수도 있고 또한 잡음에 민감할 수도 있는 단점이 있다.

주파수영역(frequency domain)에 워터마크를 삽입하는 기법중에서 FT(Fourier Transform)변환을 사용한 예로는 Solachidis와 Pitas[4]가 제안한 대칭적(symmetric) 워터마킹 삽입기법이 있다. 이는 이미지를 DFT (Discrete Fourier Transform)을 사용하여 변환한 다음 주파수영역의 magnitude 계수에 워터마크를 삽입하는 방식이다. 다른 방식과의 차이점은 워터마크의 삽입영역이 주파수 영역의 전체가 아닌 중간주파수 부분을

대칭적 원형으로 선택한 후에 이를 다시 등간격의 섹터(sector)로 나누어 워터마크를 삽입한다는 점이다. 등간격의 섹터로 나누어 삽입하는 이유는 회전(rotation)공격에 강인하도록 하기 위해서이다. 워터마크의 추출은 통계적인 방법에 의해서 이루어진다. 삽입할 때 +와 -를 무작위(random)로 삽입함으로써 그 총합이 0에 근사하게 되는 성질을 이용하여 워터마크를 원본의 정보없이 추출한다.

최근에는 혼합형(Hybrid) 워터마킹 기법이 많이 연구되고 있는데 D. Swanson et. al.[6]이 제안한 방법은 마스킹(masking) 특성을 이용하여 워터마크를 공간영역과 주파수 영역에 삽입하는 기법이다. 주파수영역에 삽입하기 위해서 DCT변환을 사용했으며, 이렇게 두 개의 영역에 워터마크를 삽입함으로써 noise첨가나 JPEG압축에 견고하도록 하였다. 또한 H. Kii et. al.[7] 등은 기존에 이미 제안되었던 공간영역에서의 워터마킹 기법인 patchwork 기법과 DCT 변환후의 워터마킹 삽입기법을 동시에 적용하는 방법을 제안하였다. J. Fridrich[8]는 공간영역과 주파수영역에 각각 다른 종류의 워터마크를 삽입하는 방법을 제안하였다. 공간영역에는 fragile 워터마크를 삽입하고 DCT변환 후의 계수에는 robust 워터마크를 삽입하여 원본데이터의 변형에 민감하도록 한 방식이다. 이런 혼합형 방식은 많은 변형과 공격에 대한 견고성이 높다라는 장점은 있지만 워터마크를 중복해서 삽입하는 효과로 인해 삽입된 워터마크의 정보가 서로 상쇄된다는 것과 워터마크의 삽입/추출 과정이 복잡하다는 단점을 가지고 있다.

3. Rightful ownership

워터마킹 기법은 주로 워터마크가 삽입되는 영역에 따라 분류되긴 하나 또 다른 분류 개념으로 추출시 원본이 필요한 기법인지 아닌지에 따라 분류되기도 한다. 삽입한 워터마크의 추출시 원본이 필요한 기법들은 워터마크 추출에 있어서 전형적이고도 간단한 방식이지만 소유자가 비밀키와 동시에 원본까지 관리해야 한다는 점과 의도적인 공격자(attacker)가 이점을 악용하여 쉽게 가짜원본을 만들 수 있다는 점이 큰 단점으로 지적되고 있다. 즉 공격자가 추출시 원본과 비교하여 그 차로 워터마크를 추출하는 방식을 역이용하여 기존의 워터마크된 이미지에 자기의 워터마크를 이중으로 삽입하여 원소유주의 원본에서도 자기의 워터마크가 삽입된 것 같은 효과를 냄으로서 소유권의 구별을 어렵게 만드는 경우라 할 수 있다. 이를 rightful ownership problem이라고 하며 IBM의 Craver[5]와 L. Qiao[13]등에 의해 제기되었다. 최근에는 그러한 공격자체가 워터마크의 저작권 보호라는 본래 의도를 방해하는 가장 큰 공격의 하나로 인식되고 있으며 따라서 원본정보 없이도 삽입된 워터마크를 추출할 수 있는 워터마킹 기법[4, 9, 10]에 대한 연구가 활발히 진행중이다.

Craver[5]에 의한 문제제기는 다음과 같다.

이미 소유주의 워터마크가 삽입된 이미지가 있을 때, 그 이미지에 불법의 attacker가 자기만의 방법으로 워터마크를 다시 삽입했다고 하면 데이터의 소유주는 어떻게 구별할 수 있을까 하는 문제가 제기 된다. 하지만 이 문제는 이외로 쉽게 풀릴 수 있다. 원 소유자나 attacker가 가지고 있는 자기의 원본을 대상으로 상대방의 원본에서 자기의 워터마크를 추출할 수 있는지를 살펴보면 된다. 만약 원소유자의 원본이 유출되지 않았다면 attacker는 원소유자의 원본에서 자기의 워터마크를 검출하지 못할 것이다. 그러나, 원본을 잘 관리한다고 해서 문제가 해결되는 것은 아니다. 또 다음과 같은 시나리오를 가정해 보자. 원 소유자가 한 이미지 I에 대해 워터마크 S를 삽입하여 I′라는 워터마크된 이미지를 만들고, 그 이미지를 인터넷 같은 공용 매체에 공개했다. 그런데 attacker가 워터마크 된 이미지 I′를 인터넷 같은 공용 매체에서 가지고 가서 자신의 워터마크 S′를 삽입하여 이미지 I″를 만든다. 삽입하는 방법은 원소유자의 추출알고리즘에 사용되는 추출 함수의 반대되는 연산을 가진 역함수를 사용한다. 이렇게 만들어진 이미지 I″를 attacker 자신의 원본 이미지라고 주장하게 되고 그런 이미지를 가짜 원본 이미지(fake original source 혹은 counterfeit original)라고 부른다. 이렇게 되면 원래의 소유주는 attacker가 원본 이미지라 우기는 가짜 원본에서 자신의 워터마크를 추출할 수 있겠지만 반대로 attacker도 원소유자의 원본 이미지에서 자신의 워터마크를 추출할 수 있게 되어 소유권을 확인할 수 없게 된다.

이와 같이 추출함수의 반대연산이 존재하여 attacker가 이 반대연산을 이용, 원소유자의 원본과 유사한 가짜원본을 생성할 수 있을 때 이 워터마크 알고리즘은 invertible 하다고 한다. 만약 워터마크 알고리즘이 invertible하고 삽입알고리즘에 사용된 연산과 전체적인 알고리즘이 공개되어 있다면 누구라도 워터마크된

이미지를 가져가서 자신의 가짜원본을 생성할 수 있어 저작권보호를 할 수 없게 된다.

일반적으로 attacker가 가짜원본을 생성하여 자기의 워터마크를 원 소유주의 원본에서도 추출 가능하게 하는 이유는 워터마크 추출알고리즘이 원본과 비교해 그 상대적인 값으로 자기의 워터마크를 추출하는데 있다. 예를 들어 attacker가 워터마크 이미지에 ⊖연산을 이용, 자기만의 워터마크를 삽입하여 가짜 원본을 만든다. 물론 attacker는 삽입한 위치를 알고 있다. 이런 경우 원소유자의 원본과 attacker의 가짜원본을 비교하여 워터마크를 추출하게 되면 원소유자의 원본이 비록 attacker의 손에 들어가지 않았다 하더라도 원소유자의 원본에서도 attacker의 워터마크가 ⊕연산의 효과로 인하여 추출이 가능하게 된다.

본 논문에서는 attacker가 자기의 워터마크를 원소유주의 원본에서도 쉽게 추출할 수 없도록 하기 위해 추출 시 원본과의 비교가 필요 없는 워터마킹 알고리즘을 제안한다. 물론 추출시 원본이 필요없는 워터마킹 기법[4, 9, 10]들이 이미 제안되었지만 대부분 삽입되는 정보를 +와 -의 무작위(random) 비트열과 합하여 삽입함으로써 그 합이 0에 근사한다는 통계적 방법에 의한 방법들이다. 하지만 본 논문에서 제안하는 추출방법은 주변화소(pixel)값을 평균하여 근사원본을 생성한 뒤 이 근사원본과 워터마크된 이미지와의 차를 구함으로써 추출이 이루어진다는 점에서 기존의 방법과 다르다고 할 수 있다. 보다 원본에 근사한 근사원본을 생성하기 위해서는 주변화소값들이 서로 밀접한 상관관계하에 있어야 되는데 이를 위해서 이미지를 FFT를 이용해 주파수영역으로 변환한 후 워터마크를 삽입한다.

4. 제안한 워터마킹 기법

4.1 삽입(Embedding) 알고리즘

본 논문에서 제안하는 워터마킹 기법은 이미지의 주파수영역에 인식 가능한 패턴을 삽입하는 방식이다. 인식 가능한 패턴이라 함은 소유주를 상징하는 부가이미지로 소유주의 이름이나 로고(logo)등을 의미하며 씰영상(seal image)이라고도 한다. 씰영상을 워터마크로 삽입하면 추출시에 소유주의 독특한 마크가 추출됨으로써 소유권에 대한 시각효과를 증대시킬 수 있다.

주어진 이미지 데이터에 워터마크를 삽입하기 위해서는 먼저 이미지를 주파수 영역으로 변환해야 한다. 본 논문에서는 FFT(Fast Fourier Transform)변환을 사용하였다. 삽입알고리즘의 단계는 다음과 같다

- **단계 1** : 주어진 이미지 데이터를 FFT를 사용하여 주파수영역으로 변환한 후 진폭(amplitude)과 위상(phase)을 구한다.

$$A(x) = \sqrt{R(x)^2 + I(x)^2}$$
$$P(x) = \tan^{-1}\left(\frac{I(x)}{R(x)}\right)$$

$R(x)$와 $I(x)$는 FFT후의 실수와 허수 각각을 나타내며, $A(x)$와 $P(x)$는 x에 대한 진폭과 위상의 값이다.

- **단계 2** : 씰영상과 비밀키로부터 삽입할 워터마크 신호를 구한다. 먼저 씰영상을 1과 -1의 조합으로 바꾸어 그 비트열(bit stream)을 S라 하면,

$$S = \{s_i, 0 \le i < N\},\ \ s_i \in \{-1, 1\}$$

이 된다. 다음으로 비밀키에 의해 생성된 무작위(random)비트열, p_i를 다음과 같이 정의한다.

$$p_i,\ p_i \in \{-1, 1\},\ i \in N$$

마지막으로 삽입하고자 하는 워터마크 W_i는 s_i와 p_i를 조합하여 다음과 같이 정의한다.

$$W_i = \alpha * s_i * p_i$$

α는 워터마크의 세기를 결정하는 strength factor로 그 값이 클 경우 견고성은 높아지나 이미지 데이터의 훼손정도가 심해진다.

- **단계 3** : 진폭(amplitude)계수에 워터마크를 삽입한다. 진폭계수를 $A(x_i)$라 할 때, 워터마크 된 $A'(x_i)$는 다음과 같이 구해진다.

$$A'(x_i) = A(x_i) + W_i,\ \ 0 \le i < N$$

- **단계 4** : 워터마크된 진폭계수 $A'(x_i)$와 단계 1에서 구한 위상계수 $P(x)$를 이용하여 Inverse FFT를 위한 $R'(x)$와 $I'(x)$를 구한다.

$$R'(x) = A'(x)\cos(P(x))$$
$$I'(x) = A'(x)\sin(P(x))$$

- **단계 5** : 단계 4에서 구한 $R'(x)$와 $I'(x)$를 가지고

Inverse FFT를 실행하여 이미지를 복원한다.

4.2 추출(Extracting) 알고리즘

삽입된 셀영상을 추출하기 위해서는 소유주만이 알고 있는 비밀키가 필요하다. 본 알고리즘은 추출 시 소유주의 원본을 필요로 하지 않도록 하기 위해서 추출하는 과정에서 워터마크 된 이미지로부터 진폭계수 값의 특성을 이용하여 근사원본(approximate original)을 만든다. 추출알고리즘의 단계는 다음과 같다.

- **단계 1** : 소유주의 비밀키를 사용하여 무작위 비트열, p_i'를 구한다.
- **단계 2** : 워터마크 된 이미지를 주파수영역으로 변환한 뒤 진폭(amplitude)계수, $A'(x)$를 구한다.

$$A'(x) = \sqrt{R(x)^2 + I(x)^2}$$

- **단계 3** : 진폭계수에서 워터마크를 추출하기 위해 근사원본, $A''(x)$를 구한다. 진폭계수의 근사원본은 자신을 포함한 주변 계수값의 평균으로 구할 수 있다.

$A'(x_{i-4})$	$A'(x_{i-3})$	$A'(x_{i-2})$
$A'(x_{i-1})$	$A'(x_i)$	$A'(x_{i+1})$
$A'(x_{i+2})$	$A'(x_{i+3})$	$A'(x_{i+4})$

(그림 1) 근사원본을 구하기 위한 윈도우

(그림 1)에서 윈도우의 크기가 3인 경우 근사원본을 구하기 위한 주변계수값을 나타내고 있다. 수식으로 표현하면,

$$A''(x_i) = \frac{1}{c^2} \sum_{j=i-c^2/2}^{i+c^2/2} A'(x_j)$$

이 된다. c는 평균을 구하고자 하는 윈도우의 크기를 나타낸다. 여기서 $A'(x)$가 워터마크된 이미지의 진폭계수라면 삽입알고리즘의 단계 3에 의하여 다음과 같이 나타내질 수 있다.

$$\begin{aligned} A''(x_i) &= \frac{1}{c^2} \sum_{j=i-c^2/2}^{i+c^2/2} A'(x_j) \\ &= \frac{1}{c^2} \sum_{j=i-c^2/2}^{i+c^2/2} (A(x_j) + \alpha * s_j * p_j) \\ &= \frac{1}{c^2} \sum_{j=i-c^2/2}^{i+c^2/2} (A(x_j)) \\ &\quad + \frac{1}{c^2} \sum_{j=i-c^2/2}^{i+c^2/2} (\alpha * s_j * p_j) \end{aligned} \quad (1)$$

위 식에서 식 (1)의 두 번째 항은 p_i가 {+1, -1}의 무작위 비트열이므로 0값에 근사하게 된다. 따라서 수식을 정리하면 다음과 같다.

$$\begin{aligned} A''(x_i) &= \frac{1}{c^2} \sum_{j=i-c^2/2}^{i+c^2/2} (A(x_j)) + \Delta \ (\approx 0) \\ &\approx \frac{1}{c^2} \sum_{j=i-c^2/2}^{i+c^2/2} (A(x_j)) \end{aligned} \quad (2)$$

이는 워터마크된 이미지의 주변계수를 평균하여 근사원본을 구하는 것은 워터마크가 삽입되지 않은 원본의 주변계수들을 평균내는 것과 근사하다는 것을 보여주고 있다.(2)

또한 진폭계수 $A(x)$는 데이터의 주파수 성분을 분석하여 표현하는 Power Spectrum으로 그 값이 서로 연관성 있는 비슷한 값들로 분포하는 성질을 가진다. 즉 주변값들의 평균은 자기자신의 값과 근사하게 된다. 수식으로 정리하면 아래와 같다.

$$\begin{aligned} A''(x_i) &\approx \frac{1}{c^2} \sum_{j=i-c^2/2}^{i+c^2/2} A(x_j) \approx A(x_i) \\ \therefore A''(x_i) &\approx A(x_i) \end{aligned} \quad (3)$$

식 (3)에서 워터마크된 이미지의 진폭계수의 평균은 (만약, 소유주의 워터마크가 제대로 되어 있다면) 원본 이미지의 진폭계수와 거의 근사하다는 것을 알 수 있다. 이는 진폭계수의 상관성과 워터마크 신호인 p_i의 무작위성(randomness)을 이용하여 원본 없이도 거의 원본에 근사한 또 다른 원본을 구할 수 있다는 것을 보여주는 것이다.

- **단계 4** : 근사원본의 진폭계수인 $A''(x)$와 워터마크 이미지의 진폭계수인 $A'(x)$를 비교하여 워터마크 신호 W_i'를 추출한다.

$$W_i' = \begin{pmatrix} 1 & \text{if}(A'(x_i) - A''(x_i) \geq 0) \\ -1 & otherwise \end{pmatrix}$$

● **단계 5 : 추출된 워터마크 신호 W_i'에 단계 1에서 구한 p_i'를 곱하여 씰영상 s_i'를 구한다.**

$$s_i' = W_i' * p_i'$$

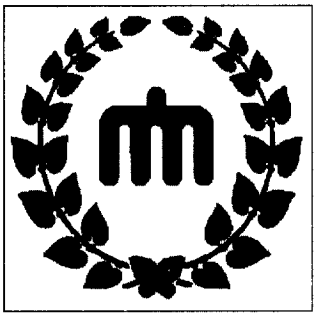
지금까지 **설명한 5단계를 거쳐 씰영상** s_i'를 추출한 후 소유권 **주장을 위하여 원래의 삽입 영상인** s_i와 유사성 정도를 **측정해야 한다. 다시 말해 추출된 워터마크(씰영상)가 본래의 삽입한 워터마크와** 유사성 정도가 크면 클수록 소유권이 분명한 것을 나타낸다. 본 논문에서는 두 씰영상의 상관성을 계산해 봄으로써 유사성을 구했다. 상관계수(Correlation Value : CV)를 수식으로 정의하면 아래와 같다.

$$CV = \frac{\sum_{i=0}^{N-1} X_i' * \sum_{i=0}^{N-1} X_i}{\sum_{i=0}^{N-1} X_i * \sum_{i=0}^{N-1} X_i}$$

수식에서 X_i'는 추출한 씰영상과 삽입한 씰영상 사이에서 일치하는 비트, X_i는 삽입한 씰영상의 비트를 나타낸다. 수식에서와 같이 유사성 비교를 위한 상관계수는 삽입한 씰영상과 추출한 씰영상의 비트를 비교함으로써 계산되어 지는데, 삽입한 씰영상이 손상 없이 전부 복구(retrieval)될 경우 그 값은 1이 됨을 알 수 있다.

5. 실험 및 결과

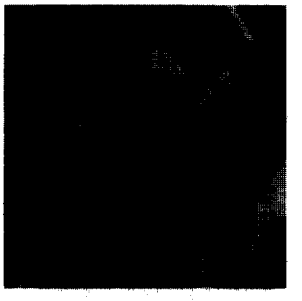
본 논문에서 사용한 씰영상은 256×256크기의 흑백 비트맵(bitmap)이미지로 대학교의 로고를 사용하였다.



(그림 2) 워터마크로 사용된 씰영상

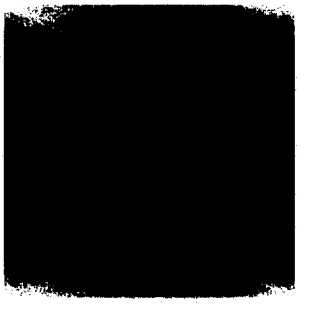
(그림 2) 흑백이미지를 사용한 이유는 씰영상이 워터마크로 삽입되기 위해서는 {1, -1}의 비트순열로 변환되어야 하기 때문이다.

첫번째 실험으로 씰영상을 (그림 3)의 256×256크기의 흑백 Lenna 이미지에 삽입해 보았다.



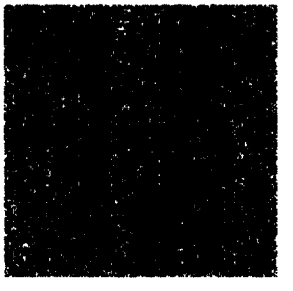
(그림 3) 원본이미지

먼저, 원본이미지를 FFT를 실시하여 주파수 영역으로 바꾼다. 주파수영역의 진폭계수를 그림으로 나타내면 (그림 4)와 같다.



(그림 4) 원본이미지의 주파수영역

다음으로 주파수영역에 삽입할 워터마크 신호를 만든다. 워터마크 신호는 씰영상과 비밀키에 의해 생성된 무작위 비트열과 strength factor, α를 곱하여 생성한다. 본 실험에서 사용된 α는 2000이다.



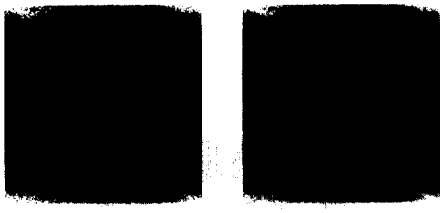
(그림 5) 씰영상이 포함된 워터마크 신호

(그림 5)는 씰영상이 함축된 워터마크 신호를 보여주고 있다. 비밀키에 의해 생성된 무작위 비트열 때문에 워터마크 신호는 원본 데이터에 삽입되는 잡음처럼 취급되어진다. 본 실험에서 사용한 비밀키는 500이다.

(그림 6)은 (그림 5)의 워터마크 신호가 (그림 4)의 주파수 영역에 삽입된 후 역 FFT를 통한 후의 이미지이다. 워터마크 된 이미지의 quality는 α 값에 의해 조정이 가능하다.

(그림 6) 워터마크가 삽입된 이미지

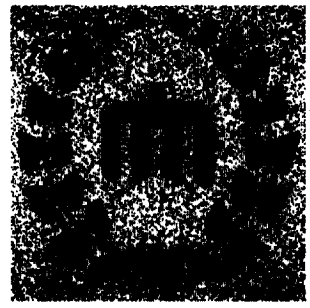
(그림 6)으로부터의 씰영상 추출은 원본이미지 필요없이 비밀키만으로도 추출이 가능하다. 우선 비밀키를 이용하여 무작위 비트열 p_i를 생성한다. 다음으로 워터마크를 추출할 이미지를 주파수 영역으로 바꾼 후 근사원본을 구한다. 위에서 언급한 것 과 같이 근사원본은 주변계수들의 평균으로 구할 수 있는데, 이는 이미지를 blurring한 것과 같은 효과라 할 수 있다. (그림 7)은 주파수영역의 주변계수들로부터 구한 근사원본을 영상으로 표현한 것이다. (b)의 영상이 (a)로부터 만들어진 근사원본이다.



(a) (b)

(그림 7) 주파수영역의 영상과 근사원본

마지막으로 주파수영역의 영상과 근사원본의 차를 구하고 무작위 비트열 p_i와 곱하여 구하고자 하는 씰영상을 얻는다. (그림 8)은 최종적으로 추출된 씰영상이다.

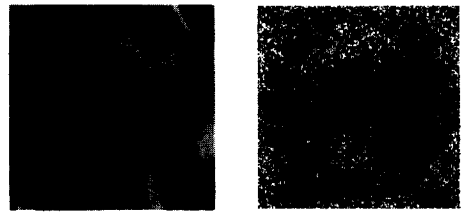


(그림 8) 추출된 씰영상

추출된 씰영상은 시각적으로 원래의 영상임을 지각할 수 있는 정도의 quality를 가짐을 볼 수 있다. (그림 7)에서 보여주고 있는 씰영상의 상관계수(correlation value : CV)는 0.721이다.

5.1 JPEG압축에 대한 견고성 실험

견고성(robustness)에 대한 실험으로 워터마크 된 이미지를 대표적인 손실(lossy)압축방법인 JPEG압축을 이용하여 압축/복원한 후 삽입한 워터마크를 추출하는 실험을 실시하였다. (그림 9)는 JPEG압축/해제 후에 추출된 씰영상을 보여주고 있다



(그림 9) JPEG압축/해제후 추출된 씰영상

본 실험에서 사용한 JPEG의 압축비는 1 : 10 (Quality factor : 85)정도이고, 상관계수는 0.645이다. (그림 9)의 씰영상은 (그림 8)에서 보여주는 씰영상보다는 훼손이 많이 되어 있는 것을 볼 수 있다. 하지만 씰영상의 시각적 패턴은 지각할 수 있을 정도이다.

두 번째 실험으로 본 논문에서 제안한 알고리즘을 320×256 크기의 6개의 비트맵(bitmap) 컬러이미지에 적용해 보았다. 실험에 사용한 컬러이미지는 흑백이미지와는 다르게 RGB 컬러스페이스를 사용하고 있으므로 먼저 YCbCr 컬러스페이스로 전환하는 작업이 필요하다. 다음으로 Y값(혹은 Cb값)을 FFT를 사용해 주파

수영역으로 변환한 다음 워터마크를 삽입하였다.

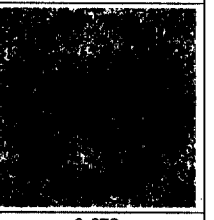
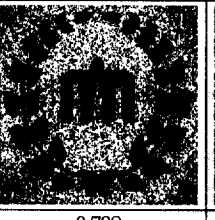
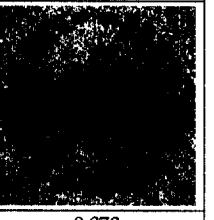
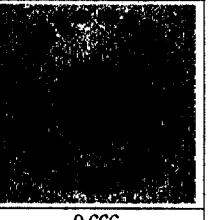
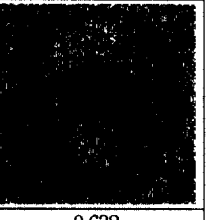
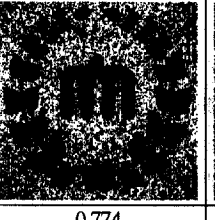
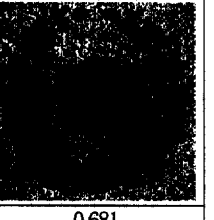
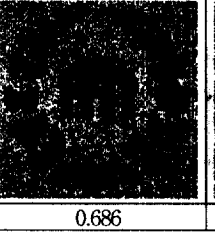
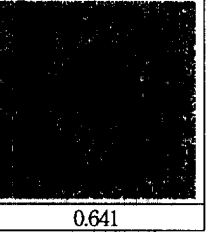
(그림 10)은 6개의 컬러이미지에서 추출한 워터마크를 보여 주고 있다. 삽입 후 아무런 처리도 하지 않은 이미지에 대해 워터마크를 추출한 경우 평균적으로 0.75정도의 상관계수값이 측정되었다. 삽입한 워터마크가 100퍼센트 복구가 되지 않는 이유는 본 논문에서 제안한 방법이 원본과 비교해서 워터마크를 추출하는 방법이 아니기 때문이다. 즉 근사원본을 추출 대상 이미지로부터 만들기 때문에 실험에서 측정된 상관계수값이 가지는 의미는 원래의 원본과 주변계수값을 평균해서 생성하는 근사원본의 유사성을 나타내는 정도라 분석된다.

JPEG압축/복원한 후 추출한 씰영상의 경우 CV 값이 현저하게 떨어졌지만 추출된 씰영상은 소유주의 씰영상임을 인식할 수 있을 정도로 복구가 가능했다.

6. 결론 및 향후연구방향

본 논문에서는 현재 저작권보호(copyright protection) 기술로 주목받고 있는 이미지 워터마킹(Image Watermarking)방법을 제안하였다. 워터마킹 기법은 특정 데이터에 소유주만이 인식할 수 있는 정보(hidden information)를 따로 숨겨두는 기술로 후에 이 정보를 소유주만의 방법으로 추출하여 소유권을 주장하게 된다. 하지만 정보를 삽입하는 과정에서 그 삽입연산에 대한 반대 연산이 존재할 경우 이를 invertible이라 하며, 불법 attacker가 이 점을 이용하여 소유권 보호를 어렵게 만드는 문제점이 있다. 본 논문에서는 이 문제를 해결하기 위하여 워터마크 추출시 원본이 필요 없는 방법을 제안하였고, 삽입하는 워터마크도 인식가능한 일정 패턴을 이용함으로써 무작위비트열(random sequence)을 워터마크 패턴으로 삽입하는 기법에 비해 추출 후 소유권에 대한 시각적인 강조효과를 증대시켰다. 특히 원본정보 없이 삽입된 워터마크를 추출하는 다른 기법들과의 차별성을 위해 주파수영역을 하나의 이미지로 취급하여 주변화소(pixel)값을 평균 내어 근사원본을 생성하는 알고리즘을 제안하고 실험하였다. 또한 씰영상을 흑백이미지는 물론 컬러이미지의 주파수영역에까지 삽입하고, JPEG 같은 손실압축/해제 후에도 견고함을 보였다.

향후연구방향으로는 기타 원본정보손실이 많은 변형에 더욱 강인할 수 있도록 하는 연구가 필요하다.

	삽입후 추출	JPEG압축/해제후 추출
PIC 1		
CV	0.772	0.652
PIC 2		
CV	0.798	0.676
PIC 3		
CV	0.779	0.666
PIC 4		
CV	0.724	0.632
PIC 5		
CV	0.774	0.681
PIC 6		
CV	0.686	0.641

(그림 10) 추출된 씰영상과 상관계수값

이를 위해 주파수영역의 저주파 혹은 중간주파수영역으로만 삽입영역을 제한하거나 워터마크를 이미지의 공간(spatial)영역에 삽입하는 방법, 혹은 공간, 주파수 두 개의 영역에 동시에 삽입/추출할 수 있는 방법 등으로 연구가 이루어져야 할 것이다.

참 고 문 헌

[1] I. J. Cox, J. Killian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," Tech. Rep. 95-10, NEC Research Institute, 1995.

[2] W. Bender, D. Gruhl, and N. Morimoto. "Techniques for data hiding," *Proc. of SPIE*, Vol.2420, pp.40, FEB. 1995.

[3] E. Koch, J. Rindfrey, and J. Zhao. "Copyright protection for multimedia data," *Proc. of the Int. Conf. on Digital Media and Electronic Publishing*, 1994.

[4] V. Solachidis, I. Pitas, Circularly Symmetric Watermark Embedding in 2-D DFT domain, *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'99)*, Vol.VI, pp.3469-3472, Phoenix, USA, March, 1999.

[5] Scott Craver, Nasir Memon, Boon-Lock Yeo and Minerva Yeung. "Can invisible watermarks resolve rightful ownerships?," *Proc. of the IS&T/SPIE Conference on Storage and Retrieval for Image and Video Databases V*, Vol.3022, pp.310-321, San Jose, CA, USA, FEB. pp.13-14, 1997.

[6] M. Swanson, B. Zhu and A. Tewfik, Robust Data Hiding for Images, *IEEE Digital Signal Processing Workshop (DSP'96)*, pp.37-40, Sep., 1996.

[7] H. Kii, J. Onishi, and S. Ozawa, The Digital Watermarking Method by Using both Patchwork and DCT, *Proceedings of the International Conference on Multimedia Computing and Systems*, Vol.1, pp.895-899, Florence, Italy, June, 1999.

[8] J. Fridrich, A Hybrid Watermark for Tamper Detection in Digital Images, *Proceedings of the Fifth International Symposium on Signal Processing and its Applications(ISSPA'99)*, Vol.1, pp.301-304, Australia, Aug., 1999.

[9] F. Hartung and B. Girod. "Digital watermarking of raw and compressed video," *In N. Ohta, editor, Digital Compression Technologies and Systems. For Video communications, volume 2952 of SPIE Processings Series*, pp.205-213, Oct., 1996.

[10] V. Basia and I. Pitas, Robust Audio Watermarking in the time-domain, *Proc. of EUSIPCO'98*, Sep., Rhodes, Greece, 1998.

[11] N. Nikolaidis, I. Pitas, Digital Image Watermarking : an Overview, *Proceedings of the International Conference on Multimedia Computing and Systems*, Vol.1, pp.1-6, Florence, Italy, June, 1999.

[12] S. Kang and Y. Aoki, Image Data Embedding System for Watermarking Using Frensnel Transform, *Proceedings of the International Conference on Multimedia Computing and Systems*, Vol.1, pp.885-889, Florence, Italy, June, 1999.

[13] L. Qiao and K. Nahrstedt, Watermarking Methods For MPEG Encoded Video : Towards Resolving Rightful Ownership, Technical Report UIUCDCS-R-97-2032, Dec., 1997.

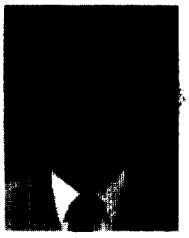
[14] E. Kock and J. Zhao. "Towards robust and hidden image copyright labeling," In *Proc. of 1995 IEEE Workshop on Nonlinear Signal and Image Processing*, JUNE 1995.

[15] A.G. Bors and I. Pitas, "Image watermarking using DCT domain constraints," *Proc. IEEE Int. Conference on Multimedia Computing and Systems*, pp.86-90, 1994.

[16] Ju Han Kim, Won Don Lee and Jin Hyeong Park, "A Watermarking Scheme Robust to IBM Attack," *Proc. of The International Conference on Multimedia and Telecommunications Management (ICMTM)*, Hong Kong, Dec., 1998.

[17] Ju Han Kim, Won Don Lee and Jin Hyeong Park, "A Watermarking with Two Signatures," *Proc. of IEEE Second Workshop on Multimedia Signal Processing*, pp.394-399 LA, Dec., 1998.

[18] M. Kutter, F. Jordan, F. Bossen, "Digital signature of color images using amplitude modulation," *Journal of Electronic Imaging*, Vol.7, No.2, pp. 326-332, April, 1998.



김 원 겸

e-mail : wgkim@brain.cs.chungnam.ac.kr

1992년 충남대학교 전산학과 졸업(학사)

1994년 충남대학교 대학원 컴퓨터과학과 졸업(이학석사)

1995년~1997년 현대전자(주) 생산기술연구소 주임연구원

1997년~현재 충남대학교 컴퓨터과학과 박사과정

관심분야 : 워터마킹, 음성처리, 데이터압축

이 종 찬

e-mail : jclee@cwunet.ac.kr

1988년 충남대학교 계산통계학과 졸업(학사)

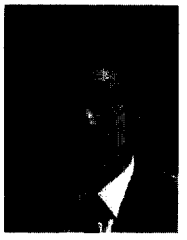
1990년 충남대학교 대학원 계산통계학과 졸업(이학석사)

1996년 충남대학교 대학원 전산학과 졸업(이학박사)

1996년~1999년 충남 청운대학교 인터넷컴퓨터학과 전임강사

1999년~현재 충남 청운대학교 인터넷컴퓨터학과 조교수

관심분야 : 신경회로망, 인공지능, 워터마킹



이 원 돈

e-mail : wdlee@cnu.ac.kr

1979년 서울대학교 화학과 졸업(학사)

1982년 일리노이대(Urbana) 화학과 졸업(석사)

1986년 일리노이대(Urbana) 전산학과 졸업(박사)

1987년 텍사스대(Arlington) 전산공학과 조교수

1987년~현재 충남대학교 정보통신공학부 교수

관심분야 : 신경회로망, 음성인식, 데이터압축, 입체음향