데이터 매트릭스와 암호 키를 이용한 하이브리드 워터마킹 기법

論 文
55D-9-5

Hybrid Watermarking Scheme using a Data Matrix and Cryptograph Key

全 盛 求* · 金 明 東** · 金 一 煥†
(Seong-Goo Jeon · Myung-Dong Kim · Il-Hwan Kim)

Abstract - In this paper we propose a new watermarking scheme using a data matrix and a cryptograph key. The data matrix of two-dimensional bar codes is a new technology capable of holding relatively large amounts of data compared to the conventional one-dimensional bar code. And a cryptograph key is used to prevent a watermark from malicious attacks. We encoded the copyright information into a data matrix bar code, and it was spread as a pseudo random pattern using the owner key. The experimental results show that the proposed scheme has good quality and is robust to various attacks, such as JPEG compression, filtering and resizing. Also the performance of the proposed scheme is verified by comparing the copyright information with the information which is extracted from the watermark.

Key Words : Two-Dimensional Bar Code, Data Matrix, Cryptograph Key, Copyright Information, Watermarking

1. 서 론

최근 멀티미디어의 발달, 초고속 통신망의 보급 그리고 디지털 데이터를 쉽게 조작할 수 있는 강력한 도구로 인해서 디지털 컨텐츠에 대한 저작권 보호 문제가 크게 부각되고 있다. 이를 해결하기 위한 솔루션으로 디지털 워터마킹(digital watermarking) 기법이 다양하게 제안되고, 연구가 활발하게 진행되고 있다. 일반적으로 디지털 워터마킹은 디지털 컨텐츠의 저작권 보호를 위해 소유권 및 저작권 정보를 사람의 육안이나 청각으로는 구별할 수 없게 디지털 컨텐츠에 삽입하는 기법으로 디지털 컨텐츠에 대한 무단 배포 및 불법적인 사용을 막고 소유권을 주장할 수 있는 근거를 제시할 수 있도록 하는 기술이다[1][2]. 이와 같이, 소유권 정보인 디지털 워터마크가 효과적이기 위해서는 다음의 특징을 가져야 한다[2][3]. 첫 번째로 지각적, 통계적으로 비가시성을 가져야 한다. 두 번째로 손실 압축을 포함하는 공격들에 대해서 강인해야 한다. 즉, 어떠한 공격에도 워터마크가 검출되어야 한다. 세 번째로 저작권자가 아닌 불법적인 사용자로부터 안전해야 한다. 네 번째로 분명하게 소유자를 확인할 수 있어야 한다.

기존의 디지털 워터마킹 기법들은 의미 있는 로고 또는 랜덤 시퀀스를 워터마크로 사용하여 왔다[1-3][6-9]. 의미 있는 로고는 저작권자가 자신의 특정한 정보를 이미지로 만들어 사용하기 때문에 저작권을 공표하기 위한 아주 좋은 툴이다. 랜덤 시퀀스는 평균이 0이고 분산이 1인 정규분포를 가지는 신호를 만들어서 이미지 스펙트럼 전체에 넓게 펼치는 것으로 의도적 또는 비의도적 공격에 대해서 보안성을 높일 수 있는 특징을 가지고 있다. 그러나 이러한 워터마크가 여러 종류의 공격에 의해서 손상되었을 때 그것은 더 이상 저작권을 증명할 수 없을 것이다. 그리고 의미 있는 로고의 경우 디지털 서명 또는 저작권을 표시하기 위해 많은 양의 데이터를 사용할 경우 워터마크의 크기가 커지는 문제점을 가지고 있다[4].

본 논문에서는 이러한 문제점을 해결하기 위해서 데이터 매트릭스와 암호 키를 이용한 하이브리드 워터마킹 기법을 제안하였다. 데이터 매트릭스는 고밀도의 데이터 저장능력과 바코드가 손상이 되더라도 오류 검출 및 복원 알고리즘으로 원본 데이터를 복원할 수 있는 특징을 가지고 있다. 한편, 저작권을 증명할 수 있는 정보를 데이터 매트릭스 생성 알고리즘을 이용하여 암호화하고 소유자 키를 사용하여 바코드를 랜덤화하여 확산된 워터마크 신호를 만든다. 그리고 워터마크 삽입위치와 워터마크 삽입 패턴을 암호 키를 이용하여 생성함으로써 공격에 보다 강인하게 하였다. 워터마크 삽입은 512×512 이미지를 DCT(Discrete Cosine Transform) 변환한 후 주파수 영역에서 64×64 크기의 워터마크를 삽입하였으며, 검출은 원 영상을 사용하지 않는 블라인드 워터마킹 방법을 사용하였다. 제안된 방법의 성능을 평가하기 위하여 워터마킹된 영상의 PSNR(Peak Signal to Noise Ratio)과 여러 가지 공격 후에 추출된 워터마크의 NC(Normalized Correlation)를 측정하였고, 2차원 바코드 스캐너로 인식하여 워터마킹 기법의 타당성을 확인하였다.

본 논문은 다음과 같이 구성되어 있다. 2장에서 데이터 매트릭스 2차원 바코드에 대해서 간략하게 알아볼 것이다.

† 교신저자, 正會員 : 江原大學校 電氣電子工學部 敎授 · 工博
E-mail : ihkim@kangwon.ac.kr
* 學生會員 : 江原大學校 制御計測工學科 博士課程
** 學生會員 : 江原大學校 電氣電子工學部 碩士課程
接受日字 : 2006年 4月 20日
最終完了 : 2006年 7月 24日

3장은 워터마크 삽입과 추출에 대해서 알아볼 것이다. 4장에서는 실험을 통해서 워터마크를 삽입하고 워터마크가 삽입된 영상에 여러 가지 공격을 가한 후에 워터마크를 추출한 결과를 보여줄 것이다. 마지막으로 5장에서 최종적으로 결론을 맺는다.

2. 데이터 매트릭스

2.1 데이터 매트릭스 구조

데이터 매트릭스 2차원 바코드는 1차원 바코드 심벌로지가 가지는 문제점인 데이터 표현의 제한성을 보완하기 위하여 1980년대 중반에 제안 되었다. 본 논문에서는 2차원 바코드에서 널리 사용되고 있는 고밀도의 데이터 저장능력과 오류수정 기능이 포함된 데이터 매트릭스 코드를 채택하였다. 데이터 매트릭스에는 오류검출 및 복원(Error Checking & Correction) 알고리즘으로 Convolutional 방법을 사용하는 ECC00-140과 Reed-Solomon방법을 사용하는 ECC200이 있다. 본 논문에서는 ECC200을 사용하였다[5].

그림 1은 데이터 매트릭스 2차원 바코드를 나타낸다. 데이터 매트릭스는 규칙적인 배열로 설계된 정사각형 모듈을 포함하는 데이터 영역으로 구성된다. 데이터 영역은 정렬 패턴에 의해서 분리되어 있다. 데이터 영역은 finder 패턴으로 둘러싸여 있으며, 이 패턴은 빈 여백으로 사방이 둘러싸여 있다. 그림 1의 (b)와 (c)는 finder 패턴을 나타낸다. 이것은 데이터 영역의 주변에 있고 폭이 1인 모듈이다. 데이터 매트릭스에서 데이터의 한 비트를 암호화하기 위해 사용된 하나의 셀을 모듈이라고 한다. (b)의 L자 모양으로 구성된 왼쪽과 아래의 인접한 테두리는 검은색 선이다. 이들은 주로 실제 크기, 방위, 심벌 뒤틀림을 결정하는데 사용된다. (c)는 검은색과 흰색 모듈이 교대로 나타나도록 구성되어 있다. 이들은 심벌의 셀 구조를 정의하는데 사용되고 바코드의 실제 사이즈와 왜곡을 결정하는데 도움을 줄 수 있다. (d)의 빈 여백은 인식 패턴을 둘러싸고 있는 스페이스 영역으로 최소 1 모듈 이상의 폭을 가져야 한다. (a)는 데이터 영역을 나타낸다. 데이터 영역은 입력된 데이터에 의해서 생성된 코드워드와 그 코드워드에 의해서 생성된 오류 정정 코드워드를 표시한다. 바코드는 짝수 개의 열과 줄의 수를 갖는다. 이것은 빈 여백을 포함하지 않는 10x10에서 144x144 크기를 가지는 정사각형이다. 데이터 매트릭스는 최대 2334개의 문자 숫자식의 문자 표현이 가능하다. 오류정정 기능은 바코드의 약 30%가 손상이 되어도 복원이 가능하다[5][10].

2.2 데이터 암호화

데이터 매트릭스에서 데이터는 6가지 암호화 schemes의 조합을 사용해서 암호화된다[5]. 표1은 6가지 암호화 scheme을 나타낸다.

데이터 매트릭스 바코드를 생성하는 암호화 절차는 다음과 같이 크게 3단계로 분류할 수 있다. 각 단계별로 설명하기로 한다.

○ 단계 1 : 데이터 암호화

입력된 데이터 열을 암호화하기 위하여 분석한다. 입력된 데이터 집합에 대한 최고의 암호화 scheme은 문자당 최소의 비트 수를 가지는 하나의 scheme이 아닐 수도 있다. 데이터 매트릭스 바코드는 기본적인 scheme보다 더 효율적으로 입력된 데이터 열을 코드워드로 변환할 수 있는 다양한 암호화 scheme을 허용한다.

○ 단계 2 : 오류검출 및 정정 코드워드 생성

오류정정 코드워드는 Reed-Solomon 알고리즘을 사용하여 생성되어지고 암호화된 데이터 열에 덧붙여진다.

○ 단계 3 : 매트릭스 안에 모듈 배치

심볼 문자배치 규칙에 의하여 코드워드 모듈을 배치한다.

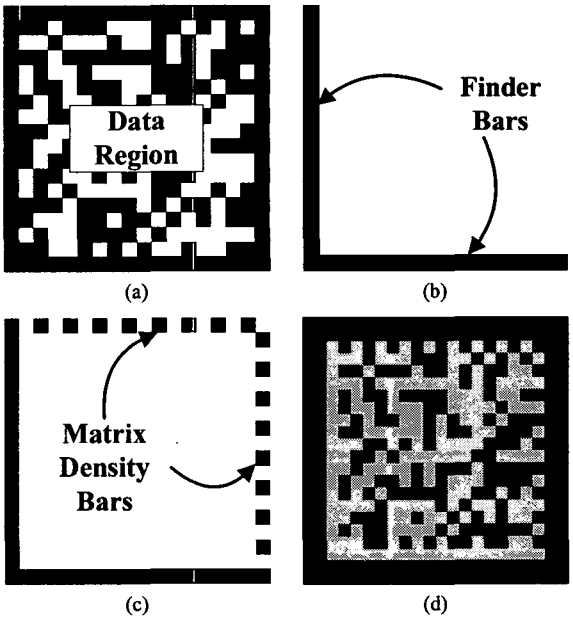


그림 1 데이터 매트릭스 바코드 구조

Fig. 1 Data Matrix 2D barcode structure

표 1 데이터 매트릭스 암호화 schemes

Table 1 Encodation schemes for Data Matrix 2D barcode

암호화 schemes	Characters
ASCII	Double digit numerics ASCII values 0~127 Extended ASCII values 128~255
C40	Primarily upper-case alphanumeric
Text	Primarily lower-case alphanumeric
X12	ANSI X12 EDI data set
EDIFACT	ASCII values 32~94
Base256	All byte values 0~255

3. 제안한 워터마킹 기법

본 논문에서 DCT계수의 크기를 변화시키는 삽입 가중치 인자 α 값을 각 DCT블록의 평균 절대 편차에 따라 적응하

여 워터마크를 삽입하고 추출하는 방법을 사용하였다. 또한 암호 키를 이용하여 워터마크 삽입 위치와 삽입 패턴을 결정하였다.

3.1 워터마크 삽입

본 논문에서 제안한 워터마크 삽입 알고리즘은 크게 워터마크 생성부와 워터마크 삽입부로 나눌 수 있다. 그림 2는 전체적인 워터마크 삽입 알고리즘을 나타낸다.

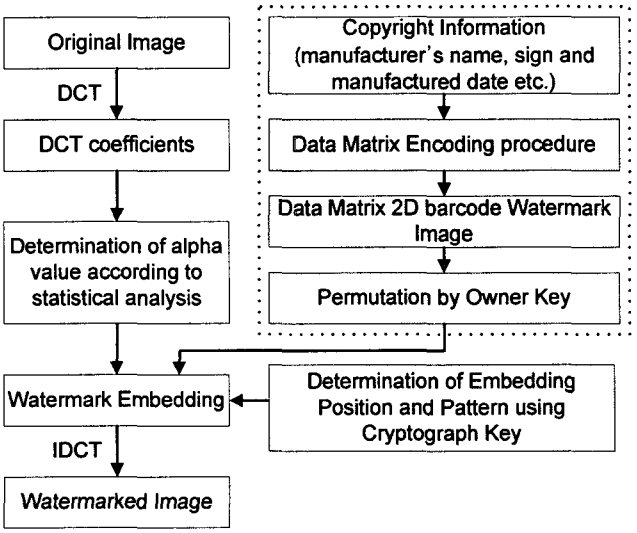


그림 2 워터마크 삽입 알고리즘
Fig. 2 Watermark embedding algorithm

먼저 워터마크 생성부에서는 입력 정보를 이용하여 랜덤화 된 워터마크 신호를 최종적으로 생성한다. 워터마크 생성부는 다음과 같이 4단계로 알고리즘이 구성되어 있다.

- ○ 단계 1 : 제작자의 이름, 서명, 제작된 날짜 등의 저작권 정보를 입력받는다.
- ○ 단계 2 : 입력된 정보는 데이터 매트릭스 암호화 과정을 통해서 데이터 매트릭스 2차원 바코드로 생성한다.
- ○ 단계 3 : 단계 2에서 생성된 데이터 매트릭스 2차원 바코드를 워터마크 이미지로 생성한다.
- ○ 단계 4 : 단계 3에서 생성된 워터마크 이미지를 소유자 키를 이용하여 랜덤화 된 워터마크 신호로 생성한다.

그리고 워터마크 삽입부에서는 앞서 생성된 워터마크를 원본 영상에 삽입하는 부분으로 4단계로 알고리즘이 구성되어 있다.

- ○ 단계 1 : 원본 영상을 8×8 블록 크기로 나누어서 DCT를 수행한다.
- ○ 단계 2 : 그림 3과 같이 각각의 8x8 DCT 블록에서 비가시성과 견고성을 고려하여 22개의 삽입위치를 선정한다. 그리고 암호 키를 이용하여 선정된 22개의 삽입 위치 중에서 15개의 워터마크 삽입 위치와 삽입 패턴을 결정한다.
- ○ 단계 3 : 각각의 8×8 DCT 블록에서 앞서 결정된 삽입 위치의 계수 값들을 가지고 평균 절대 편차에 비례하는 워터마크 삽입 강도 α 값을 계산한다.
- ○ 단계 4 : 식 (1)에 의해서 각 워터마크 비트는 각각의 8×8 DCT 블록에서 선정된 15개의 계수에 α값과 앞서 결정된 삽입 패턴에 따라 삽입한다.

본 논문에서 제안한 워터마크 삽입 알고리즘에서 저작권 정보를 데이터 매트릭스 2차원 바코드로 만드는 워터마크 생성 알고리즘과 암호 키에 의해서 삽입 패턴과 삽입 위치를 결정하는 알고리즘이 가장 중요한 요소이다.

$$W_i = I_i + sgn(I_i)\alpha E_i^j,\ i=1,2\cdot\cdot\cdot15,\ j=0,1$$
$$\alpha = C*\sigma,\ E_i^j \in [E_i^0, E_i^1] \quad (1)$$

식(1)은 워터마크 삽입 수식이다. I_i는 원본 영상의 DCT 계수, W_i는 워터마크가 삽입된 후에 DCT 계수, C는 비례 상수, σ는 8×8 DCT 블록에서 선정된 15개 계수의 평균 절대 편차, E_i^0와 E_i^1은 워터마크 비트의 "0"과 "1"의 삽입 패턴이다.

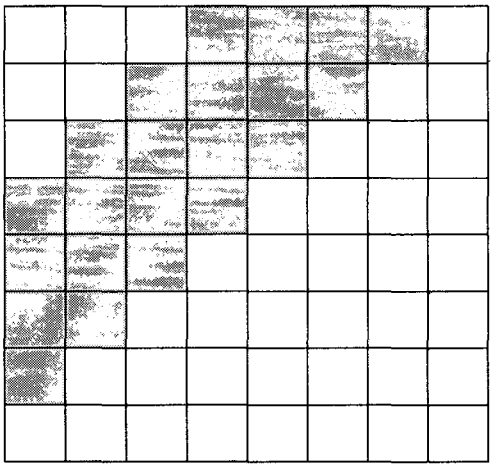


그림 3 워터마크 삽입 위치
Fig. 3 Watermark embedding positions

3.2 워터마크 추출

그림 4는 워터마크 추출 알고리즘을 나타낸다. 본 논문의 워터마크 추출은 원 영상이 필요하지 않은 블라인드 워터마킹 기법을 사용하고 있다. 알고리즘의 대략적인 과정은 아래와 같다.

- ○ 단계 1 : 워터마크가 삽입된 영상을 8×8 블록 크기로 나누어서 DCT를 수행한다.
- ○ 단계 2 : 암호 키를 이용하여 삽입 위치와 삽입 패턴을 결정한다.
- ○ 단계 3 : 식 (2)에 의해서 각각의 8×8 DCT 블록에서 앞서 결정된 삽입 위치의 계수 값과 삽입 패턴과의 상관도를 구하여 워터마크를 추출한다.

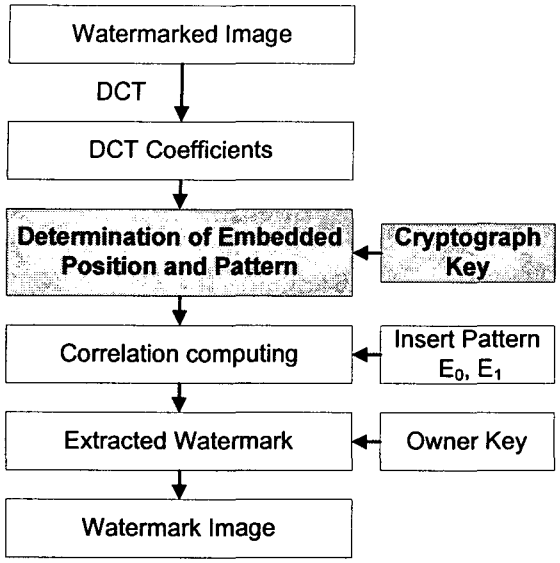


그림 4 워터마크 추출 알고리즘
Fig. 4 Watermark extraction algorithm

식(2)는 워터마크 검출 수식이다.

$$D_0 = \sum | W_i | E_i^0 \ , \ D_1 = \sum | W_i | E_i^1$$
$$\text{if} \begin{pmatrix} D_0 \geq D_1 \ , \ bit = 0 \\ D_0 < D_1 \ , \ bit = 1 \end{pmatrix} \quad (2)$$

4. 실험 및 결과

본 논문에서는 제안된 방법의 성능을 평가하기 위해서 워터마크 삽입 후, 영상의 손실정도를 측정하기 위해 PSNR을 사용하였고, 원본 워터마크와 추출된 워터마크의 객관적인 유사성 측정을 위하여 NC을 사용하였다. 식(3)은 PSNR 계산식을 나타낸다.

$$PSNR(dB) = 10 \log_{10} \frac{MN \max I(x,y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\gamma(x,y) - I(x,y)]^2} \quad (3)$$

여기서 $I(x,y)$ 는 원 영상이며, $\gamma(x,y)$ 는 워터마크가 삽입된 영상이다.

$$NC = \frac{\sum_i \sum_j w(i,j) \hat{w}(i,j)}{\sum_i \sum_j w(i,j)^2} \quad (4)$$

그리고 식(4)는 원본 워터마크와 추출된 워터마크 사이의 유사성 측정을 위한 수식이다. $w(i,j)$ 는 원본 워터마크이고, $\hat{w}(i,j)$는 추출된 워터마크이다.

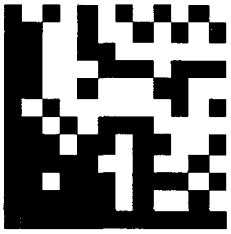


그림 5 워터마크 이미지
Fig. 5 Watermark image

표 2 실험 결과
Table 2 The details of simulation result

테스트 영상	PSNR (dB)	공격(Attacks)	NC	인식한 결과
Barbara	39.34	No attack	1.0	1234567890
		JPEG (Quality 80)	0.98	
		Sharpening	0.98	
		Blurring	0.98	
		Resizing(BO)	0.98	
		Resizing(SO)	0.97	
Boat	40.31	No attack	1.0	
		JPEG (Quality 80)	0.96	
		Sharpening	0.98	
		Blurring	0.98	
		Resizing(BO)	0.97	
		Resizing(SO)	0.95	
Pepper	42.44	No attack	1.0	
		JPEG (Quality 80)	0.96	
		Sharpening	0.97	
		Blurring	0.98	
		Resizing(BO)	0.98	
		Resizing(SO)	0.97	

본 논문에서 우리는 실험 영상으로 512×512 크기의 Barbara 영상을 포함한 3개의 표준 실험 영상을 사용하였다. 그리고 워터마크는 "1234567890"을 앞서 언급한 방법으로 생성한 데이터 매트릭스 바코드 이진영상을 사용하였다. 강인성 실험을 위하여 JPEG 압축, 필터링 그리고 크기 변화 공격에 대한 실험을 하였다. 실험을 통해 추출된 워터마크를 2차원 바코드 스캐너로 인식하여 결과를 확인하였다. 표 2는 공격을 가한 후 영상손실, 추출된 워터마크의 유사성을 측정한 결과와 바코드 스캐너로 확인한 결과를 나타낸다. 스캐너로 확인한 결과 입력 데이터가 모두 확인되었다. 그림5는 "1234567890" 정보를 암호화하여 워터마크로 만든 64×64 크기의 데이터 매트릭스 2차원 바코드 이미지이다.

표 2에서 Resizing(BO)은 워터마크가 삽입된 영상을 2배로 크기를 변화시킨 후에 다시 원래의 이미지 크기로 축소하는 공격이고, Resizing(SO)은 0.5배로 크기를 변화시킨 후에 다시 원래의 이미지 크기로 확대하는 공격이다.

그림 6은 3개의 표준 영상의 실험 중에서 Barbara 이미지에 대한 강인성 실험 결과이다.

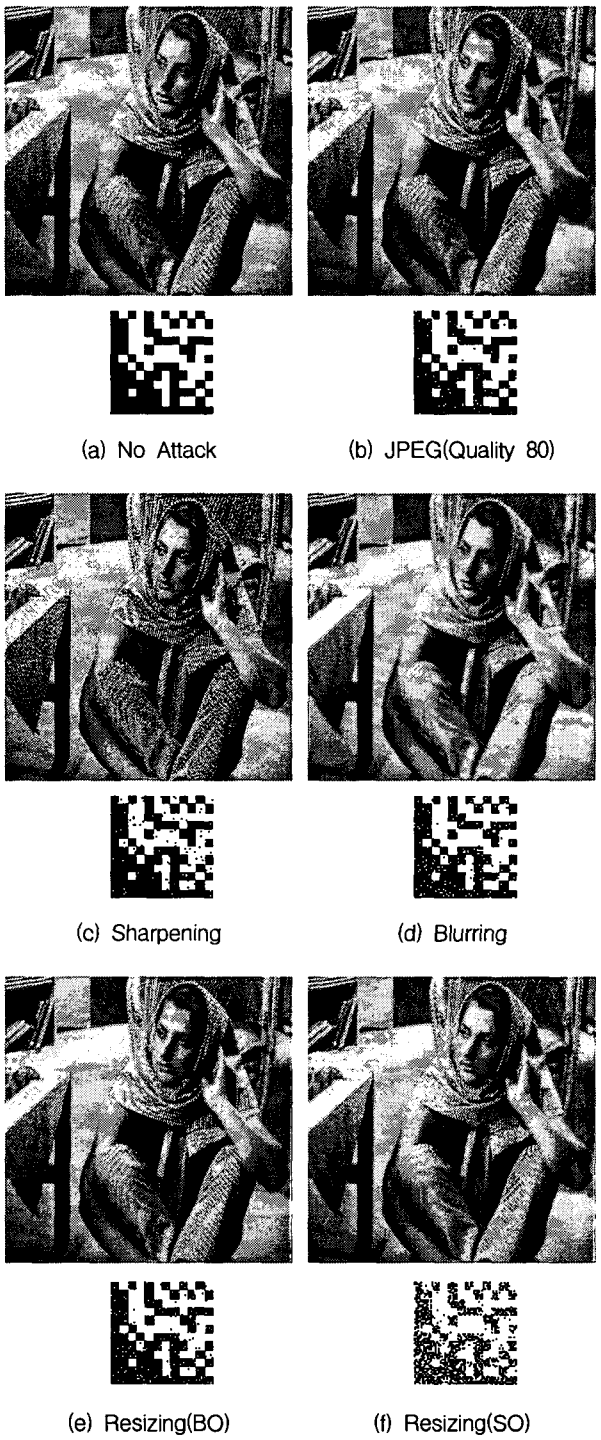


그림 6 강인성 실험 결과의 한 예
Fig. 6 An example of various attacks

5. 결 론

본 논문에서는 기존의 연구에서 워터마크가 여러 가지 공격에 의한 손상이나 저작권을 나타내기 위해 워터마크의 용량이 증가하는 문제를 해결하기 위해 데이터 매트릭스와 암호 키를 이용한 하이브리드 워터마킹 기법을 제안하였다. 저작권 정보를 데이터 매트릭스 2차원 바코드 생성 알고리즘에 적용하여 보안성이 강화된 바코드를 생성하였다. 생성된 바코드를 암호 키를 이용하여 삽입, 추출하였다. 실험 결과에서 알 수 있듯이, 저작권 정보를 데이터 매트릭스 생성 알고리즘을 사용하여 암호화하여 바코드로 생성함으로써 워터마크가 손상이 되어도 저작권 정보를 정확히 찾아내는 것을 볼 수 있었다. 또한 암호 키를 사용하여 삽입 패턴을 결정함으로써 원본 영상 없이도 워터마크를 추출할 수 있었다. 그리고 시각적으로도 영상 손실 정도가 대략 40dB이상을 유지하였고, 여러 가지 공격 후에 추출된 워터마크와 원본 워터마크의 유사성은 평균적으로 약 0.97의 결과를 보였다.

감사의 글

이 논문은 강원대학교 BK21사업 및 정보통신연구소 지원에 의하여 이루어진 연구로서, 관계부처에 감사를 드립니다.

참 고 문 헌

[1] I. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for images, audio and video", in Proc. Int. Conf. Image Processing, Vol. 3, pp. 243-246, Sept. 1996.

[2] W. Zeng and B. Liu, "A statistical watermark detection technique without using original images for resolving rightful ownerships of digital images", IEEE Transactions on Image Processing, Vol.8, pp. 1534-1548, Nov. 1999.

[3] N. Nikolaidis and I. Pitas, "Copyright protection of images using robust digital signatures", in Proc. Int. Conf. Acoustics Speech and Signal Processing, Vol. 4, pp. 2168-2171, May 1996.

[4] Ji-Hong Chang and Long-Wen Chang, "A new image copyright protection using digital signature of trading message and bar codewatermark", Proceeding of Image and Vision computing, pp. 205-209, Nov. 2003.

[5] ISO/IEC 2000, International symbology specification - Data Matrix, 2000.

[6] J. J. Chae and B. S. Manjunath, "A Robust Embedded Data from Wavelet Coefficients", Proceedings of the SPIE : Storage and Retrieval for Image and Video Databases VI, Vol. 3312, pp. 308-317, Jan. 1998.

[7] M. Barni, F. Bartolini, A. De Rasa and A. Piva, "Capacity of full frame DCT image watermarks", IEEE Transactions on Image Processing, Vol 9, pp. 1450-1455, Aug. 2000.
[8] J. O. Ruanaidh, H. Petersen, A. Herrigel, S. Pereira, and T. Pun, "Cryptographic copyright protection for digital images based on watermarking techniques", Theoretical Computer Science, Vol. 226, pp. 117-142, Sept. 1999.
[9] Tomokazu Onuki, Takeharu Adachi, Madoka Hasegawa, and Shigeo Kato, "A study on a digital watermarking method for still images", International Technical Conference on Circuits/Systems, Computers and Communications, pp. 19-22, July 2000.
[10] 김병찬, 정성훈, 임재홍, "유비쿼터스 환경에서 2차원 바코드와 RFID 응용에 관한 연구", 한국항해항만학회 04 춘계학술대회 논문집, pp. 49-54, 2004.

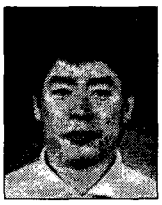
저 자 소 개

전 성 구 (全 盛 求)

강원대학교에서 제어계측 학사, 석사학위를 각각 2000년과 2002년에 받았으며, 현재 동 대학원에서 박사과정 중에 있다. 관심 연구 분야는 제어 시스템, 시스템 프로그래밍, 영상처리이다.

김 명 동 (金 明 東)

강원대학교에서 제어계측 학사학위를 2005년에 받았으며, 현재 동대학원에서 석사과정 중에 있다. 관심 연구 분야는 제어 시스템, 지능 제어, 마이크로프로세서 응용이다.



김 일 환 (金 一 煥)

서울대학교에서 제어계측 학사, 석사 학위를 각각 1982년과 1988년에 받았으며, 1993년에 일본 토호쿠 대학에서 공학 박사 학위를 받았다. 1995년 강원대학교 전기전자공학부 교수로 임용되어 현재 동 학부 교수로 재직 중이다. 관심 연구 분야는 제어, 메카트로닉스 및 휴먼 인터페이스이다.