데이터 매트릭스와 비밀 키를 이용한 하이브리드 워터마킹 방법

Hybrid Watermarking Scheme using a Data Matrix and Secret Key

전 성 구*, 김 일 환**
Seong-Goo Jeon*, Il-Hwan Kim**

Abstract - 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 which is just a key that can access detailed information to the host computer database. A secret key is used to prevent a watermark from malicious attacks. We encoded copyright information into a Data Matrix bar code for encoding process and it was spread a pseudo random pattern using owner key. We embedded a randomized watermark into the image using watermark's embedding position, pattern generated with a secret key. The experimental results have shown that the proposed scheme has good quality and is very robust to various attacks, such as JPEG compression and noise. Also the performance of the proposed scheme is verified by comparing the copyright information with the information which is extracted from a bar code scanner.

Key Words : Data Matrix, Secret Key, Copyright, Robustness

1. 서론

최근 빠르게 증가하고 있는 멀티미디어 데이터에 부응하여 멀티미디어 데이터를 조작할 수 있는 강력한 도구로 인해서 저작권 보호 문제가 크게 부각되고 있다. 디지털 워터마크는 멀티미디어 데이터의 저작권 보호를 위해 멀티미디어 데이터에 삽입하는 식별 코드이다. 이러한 디지털 워터마크를 사람의 육안이나 청각으로는 구별할 수 없게 멀티미디어 데이터에 삽입하는 디지털 워터마킹 기술이 저작권 보호 솔루션으로 최근에 지속적으로 제안되고 있다[1][2].

기존의 디지털 워터마킹 연구들은 의미 있는 로고 또는 랜덤 시퀀스를 워터마크로 사용하여 왔다[1-3][6][7]. 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. 만약 이러한 워터마크가 여러 종류의 공격에 의해서 손상되었을 때 그것은 더 이상 저작권을 증명할 수 없을 것이다. 그리고 의미 있는 로고의 경우 디지털 키 또는 저작권을 표시하기 위해 많은 양의 데이터를 사용할 경우 워터마크의 크기가 커지는 문제점을 가지고 있다[4].

저자 소개
* 전성구 : 江原大學校 制御計測工學科 碩士課程
** 김일환 : 江原大學校 傳記電子工學部 敎授

본 논문에서는 이러한 문제점을 해결하기 위해서 2차원 바코드와 비밀 키를 이용한 하이브리드 워터마킹 방법을 제안하였다. 저작권을 증명할 수 있는 정보를 2차원 바코드 알고리즘을 이용하여 암호화하고 소유자 키를 사용하여 바코드를 랜덤화하여 확산된 워터마크 신호를 만든다. 그리고 워터마크 삽입위치와 워터마크 삽입 패턴을 비밀 키를 이용하여 생성함으로써 보다 공격에 강인하게 하였다. 워터마크 삽입은 이미지를 DCT 변환한 후 주파수 영역에서 워터마크를 삽입하였으며, 검출은 원 영상을 사용하지 않는 블라인드 워터마킹 방법을 사용하였다. 제안된 방법의 성능을 평가하기 위하여 워터마킹된 영상의 PSNR과 여러 가지 공격 후에 추출된 워터마크의 NC를 측정하였고, 2차원 바코드 스캐너로 인식하여 워터마킹 방법의 타당성을 확인하였다.

2. 데이터 매트릭스

2.1 데이터 매트릭스 구조

데이터 매트릭스 2차원 바코드는 1차원 바코드 심벌로지가 가지는 문제점인 데이터 표현의 제한성을 보완하기 위하여 1980년대에 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[5]. 본 논문에서는 ECC200을 사용하였다.

그림 1은 데이터 매트릭스 바코드를 나타낸다. 데이터 매트릭스 바코드는 규칙적인 배열로 설계된 정사각형 모듈을 포함하는 데이터 영역으로 구성된다. 데이터 영역은 정렬 패

턴에 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. (b)의 L자 모양으로 구성된 왼쪽과 아래의 인접한 테두리는 검은색 선이다. 이들은 주로 실제 크기, 방위, 심벌 뒤틀림을 결정하는데 사용된다. (c)는 검은색과 밝은색 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 짝수개의 열과 줄의 수를 갖는다. 이것은 빈 여백을 포함하지 않는 10x10에서 144x144 크기를 가지는 정사각형이다. 데이터 매트릭스는 최대 2334개의 문자 숫자식의 문자 표현이 가능하다. 오류 정정 기능은 바코드의 30%가 손상이 되어도 복원이 가능하다

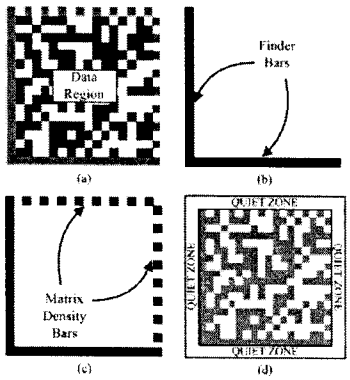

그림 1. 데이터 매트릭스 바코드 구조

2.2 데이터 암호화

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

표 1. 데이터 매트릭스 암호화 schemes

암호화 schemes	Characters
ASCII	Double digit numerics
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

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

○ 단계 1 : 데이터 암호화

입력된 데이터 열을 암호화 하기 위하여 분석한다. 입력된 데이터 집합에 대한 최고의 암호화 scheme은 문자당 최소의 비트 수를 가지는 하나의 scheme이 아닐 수도 있다. 데이터 매트릭스 바코드는 기본적인 scheme 보다 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..

3. 제안한 워터마킹 방법

본 논문에서 DCT계수의 크기를 변화시키는 삽입 가중치 인자 α값을 각 DCT블록의 평균 절대 편차에 따라 적응하여 워터마크를 삽입하고 추출하는 방법을 사용하였다. 또한 비밀 키를 이용하여 워터마크 삽입 위치와 삽입 패턴을 결정하였다.

3.1 워터마크 삽입

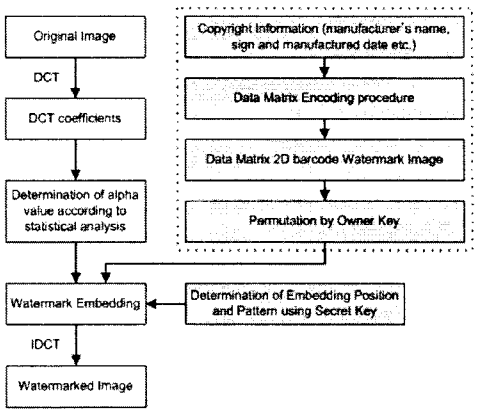

그림 2. 워터마크 삽입 과정

그림 2는 워터마크 삽입 과정을 나타낸다. 제작자의 이름, 서명, 제작된 날짜 등의 저작권 정보를 데이터 매트릭스 암호화 과정을 통해서 2차원 바코드 워터마크 이미지로 만들고 소유자 키를 이용하여 랜덤화된 워터마크 신호를 만든다. 그리고 원본 이미지를 8×8 블록 DCT를 수행한다. 각 블록의 중간 주파수 대역에서 선정된 계수들의 평균 절대 편차에 비례하는 α값을 계산한다. 생성된 워터마크와 α값을 이용하여 앞서 결정된 삽입 패턴에 따라 워터마크를 삽입한다. 삽입과정에서 저작권 정보를 데이터 매트릭스 2차원 바코드로 만드는 과정과 비밀 키에 의한 삽입 위치와 패턴 결정이 핵심적인 요소이다.

$$W_i = I_i + sgn(I_i)\alpha E_{i,a}\ ,\ i = 1,2\cdot\cdot\cdot 15\ ,\ a = 0,1$$
$$\alpha = C*\sigma\ ,\ E_{i,a} \in [E_0, E_1] \quad (1)$$

식(1)은 워터마크 삽입 수식이다. I_i는 원본 영상의 DCT 계수, W_i는 워터마크가 삽입된 후에 DCT 계수, C는 비례상수, σ는 8×8 DCT 블록에서 선정된 15개 계수의 평균 절대 편차, E_0, E_1은 워터마크 비트의 "0"과 "1"의 삽입 패턴이다.

3.2 워터마크 추출

그림 3은 워터마크 추출 과정을 나타낸다. 워터마크가 삽입된 이미지를 DCT한 후에, 삽입 패턴과 앞서 블록에서 선정한 삽입 위치의 값과의 상관도를 구하여 워터마크를 추출한다. 워터마크 추출은 원 영상이 필요하지 않은 블라인드 워터마킹 방법을 사용하고 있다.

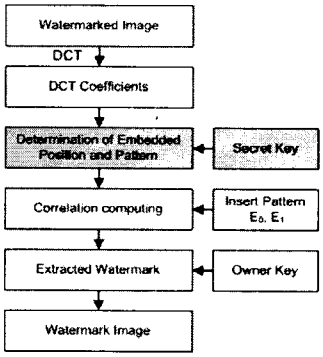

그림 3. 워터마크 추출 과정

4. 실험 및 결과

본 논문에서는 제안된 방법의 성능을 평가하기 위해서 워터마크 삽입 후, 영상의 손실정도를 측정하기 위해 PSNR(Peak Signal to Noise Ratio)를 사용하였고, 원본 워터마크와 추출된 워터마크의 객관적인 유사성 측정을 위하여 NC(Normalized Correlation)을 사용하였다. 식(2)는 PSNR 계산식을 나타낸다.

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

여기서 $I(x,y)$는 원 영상이며, $\gamma(x,y)$는 워터마킹된 영상이다. 그리고 식(3)은 원본 워터마크와 추출된 워터마크 사이의 유사성 측정을 위한 수식이다.

$$NC = \frac{\sum_i\sum_j w(i,j)\hat{w}(i,j)}{\sum_i\sum_j w(i,j)^2} \quad (3)$$

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, 노이즈(salt & pepper) 공격에 대한 실험을 하였다. 실험을 통해 추출된 워터마크를 2차원 바코드 스캐너로 인식하여 결과를 확인하였다. 표2는 공격을 가한 후 영상손실, 추출된 워터마크의 유사성을 측정하였고, 바코드 스캐너로 확인한 결과를 나타낸다. 스캐너로 확인한 결과 입력 데이터가 모두 확인되었다.

5. 결론

본 논문에서는 기존의 연구에서 워터마크가 여러 가지 공격에 의한 손상이나 저작권을 나타내기 위해 워터마크의 용량이 증가하는 문제를 해결하기 위해 데이터 매트릭스와 비밀 키를 이용한 하이브리드 워터마킹 방법을 제안하였다. 실험 결과에서 알 수 있듯이, 데이터 매트릭스 2차원 바코드와 비밀 키를 사용함으로써 일정부분 워터마크가 손상이 되어도 저작권 정보를 정확히 찾아내는 것을 볼 수 있었다. 또한 비밀 키를 사용하여 삽입 패턴을 결정함으로써 원본 영상 없이도 워터마크를 추출할 수 있었다. 그리고 시각적으로도 영상 손실 정도가 대략 40dB이상을 유지하였다.

표 2. 실험 결과

영상	PSNR(dB)	공격	NC	인식 결과
Barbara	39.71	No attack	1.0	1234567890
		JPEG (Quality 80)	0.97	
		Noise (salt&pepper)	0.87	
Boat	40.56	No attack	1.0	
		JPEG (Quality 80)	0.94	
		Noise (salt&pepper)	0.86	
Pepper	42.64	No attack	1.0	
		JPEG (Quality 80)	0.97	
		Noise (salt&pepper)	0.85	

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 code watermark", 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 International Conference on Storage and Retrieval for Image and Video Databases VI, Vol. 3312, pp. 308-317, Jan. 1998.

[7] 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.