

무질서한 SMVQ 기반의 제로-워터마킹 Zero-Watermarking Based on Chaotic Side Match Vector Quantization

김형도(Kim, Hyoung Do)^a·박찬권(Park, Chan Kwon)^b

^a 한양사이버대학교 경영학부
133-791, 서울시 성동구 행당동 17
Tel: +82- 2-2290-2812, Fax: +82- 2-2290-2828, E-mail: hdkim@hycu.ac.kr

^b 한양사이버대학교 경영학부
133-791, 서울시 성동구 행당동 17
Tel: +82- 2-2290-2813, Fax: +82- 2-2290-2828, E-mail: chankwon@hycu.ac.kr

요약(Abstract)

디지털 워터마킹은 디지털 콘텐츠에 워터마크를 삽입함으로써 불법적인 복제를 방지하고, 지적재산권 및 저작권을 보호하며, 소유권을 주장할 수 있는 근거를 제시하는 기술이다. 기존의 워터마킹 기법에서는 워터마크를 삽입함으로써 데이터 왜곡과 품질 저하가 불가피하다는 단점이 있다. 이를 극복하기 위하여 원래 데이터를 변경하지 않는 제로-워터마킹 기법들이 제시되고 있다. 이 논문에서는 VQ(Vector Quantization) 방식의 블록 효과를 줄이고, 압축 비율과 품질을 향상시킨 SMVQ(Side Match Vector Quantization) 방식에 대한 제로-워터마킹 체계인 CSMVQ(Chaotic SMVQ)를 제안한다. SMVQ 이미지 압축에서는 동일하게 두 이웃 블록의 접면 정보를 이용하기 위하여 좌측 상단에서 우측 하단으로 진행되므로, 임의의 순서로 블록을 선택하여 워터마크를 삽입하기 어렵다. CSMVQ에서는 이전에 부호화된 이웃 블록들의 접면 정보의 일치성을 동적으로 고려하여 부호화를 진행한다. 그러므로 무질서한 방식으로 워터마크가 삽입되도록 지원할 수 있다.

키워드:

워터마크; 제로-워터마킹; 벡터양자화; 보안

Keywords:

Watermark; Zero-Watermarking, VQ; Security

서론 (Introduction)

디지털 워터마킹은 디지털 콘텐츠에 워터마크를 삽입하여 불법적인 복제를 방지하고, 지적재산권 및 저작권을 보호하며, 소유권을 주장할 수 있는 근거를 제시하는 기술이다. 이러한 디지털 워터마킹 기술은 오디오, 비디오, 그래픽 등 디지털 콘텐츠의 유형이나, MP3, JPEG, MPEG 등 압축 방식에 따라서 다양한 기법이 연구/개발되고 있다. 벡터

양자화 (VQ: Vector Quantization) 기법[4]은 이미지 압축에서 많이 사용되는 근사 해법이다. 초기에 VQ 기반의 워터마킹 기법[6, 9]은 벡터 코드북의 인덱스에 워터마크를 삽입하는 기법에 대한 연구가 주류를 이루었다. VQ 이미지 압축의 문제점인 블록화 현상을 최소화하기 위한 SMVQ[7, 13]가 등장하면서 이를 기반으로 하는 워터마킹 기법들[2, 3, 11]에 대한 연구도 활발하다.

전통적인 워터마킹 기법들은 공통의 약점이 있는데, 이것은 워터마크를 디지털 콘텐츠에 삽입함으로써 어느 정도의 품질 악화를 회피하기 어렵다는 것이다. 이를 극복하기 위하여 등장한 것이 제로-워터마킹[1, 5, 10]으로서, 디지털 콘텐츠에 대한 워터마크 삽입을 하지 않는 워터마킹 기법이다. 이 논문은 SMVQ에 제로-워터마킹을 적용하기 위한 CSMVQ(Chaotic SMVQ)를 제안한다. SMVQ 이미지 압축에서는 동일하게 두 이웃 블록의 접면 정보를 이용하기 위하여 좌측 상단에서 우측 하단으로 진행되므로, 임의의 순서로 블록을 선택하여 워터마크를 삽입하기 어렵다. CSMVQ에서는 이전에 부호화된 이웃 블록들의 접면 정보의 일치성을 동적으로 고려하여 부호화를 진행한다.

이 논문의 구성은 다음과 같다. 2장에서는 벡터 양자화 기반의 워터마킹 기법의 동향과 문제점을 정리하고, 3장에서는 이 논문이 제안하는 CSMVQ를 소개한다. 4장에서는 이 기법에 대한 실험방법 및 결과를 제시하고 분석/평가한다. 마지막으로, 5장에서는 이 논문의 의미를 정리하고, 향후 연구방향을 제시한다.

VQ 기반 워터마킹 (VQ-Based Watermarking)

VQ 기법은 이미지 압축에서 많이 사용되는 근사 해법으로서 3단계의 처리과정으로 구성된다. 1단계 (학습단계)에서는 학습용 입력 벡터 집합으로부터 N 개의 벡터로 이루어진 코드북 $C = \{c_0, c_1, \dots, c_{N-1}\}$

를 생성한다. 2단계 (압축단계)에서는 압축하고자 하는 소스 이미지 X 를 T 개의 벡터들 $\{x_0, x_1, \dots, x_{T-1}\}$ 로 표현한다. 각각의 입력 벡터 x_i 에 대해서는 이것과 가장 유사한 코드워드 c_i 를 코드북에서 찾아내고 x_i 대신에 i 를 배정한다. 3단계 (복원단계)에서는 각각의 i 에 대응하는 c_i 를 코드북에서 추출하여 x_i 대신에 사용하게 된다.

[그림1]은 8개의 코드워드 (c_0, \dots, c_7)로, 구성된 코드북을 사용하여 이미지를 압축/복원하는 과정을 설명하고 있다. 먼저 압축과정에서는 압축하고자 하는 이미지를 코드워드와 동일한 차원의 벡터들로 분해한 뒤, 각각의 벡터들 ($x_0, x_1, x_2 \dots$)을 코드워드와 비교하여 가장 가까운 거리에 있는 것을 선택하고, 이것의 위치를 압축된 이미지에서 사용하게 된다. x_0 의 경우 c_3 가 가장 가까운 코드워드인 것으로 판명되어, 이것의 위치를 나타내는 3이 압축된 이미지에서 사용된다. 복원과정에서는 압축된 이미지에 저장된 코드워드의 위치 값을 가지고 코드북에서 코드워드의 값을 추출하여 원래의 이미지와 유사한 복원된 이미지를 생성하게 된다. 이와 같이 벡터 양자화는 효율적인 압축이 가능하고, 복원속도가 빠르다는 장점을 가지고 있다.

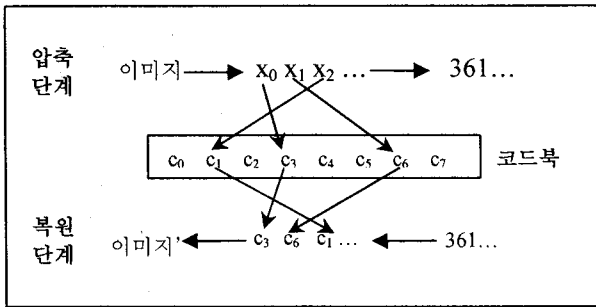


그림 1 - 벡터 양자화에서의 압축/복원 과정

VQ 기반 이미지 압축에서는 코드북 최적화를 통해서 블록의 중복을 최소화하게 되는데, 이것으로 인해서 VQ 압축 이미지에는 블록화 현상이 초래되게 된다. FSVQ(Finite State VQ)에서는 이웃하는 벡터간의 상관관계를 최적화하기 위하여 각 상태에서 사용할 수 있는 소형 코드북을 생성한다. 이 소형 코드북을 상태 코드북이라고 하는데, 이 코드북은 수퍼 코드북이라고 하는 공통의 코드북으로부터 선택하여 구성된다. SMVQ(Side Match VQ)는 FSVQ의 일종으로서, 주어진 블록의 좌측과 상단에 위치한 두 연결 블록의 정보를 사용하여 블록화 현상이 가장 적은 코드워드를 선택하게 된다. 이 기법은 블록화를 성공적으로 감소시켜 이미지 품질을 향상시키고, 압축 비율을 증가시킨다 [7]. 이와 같이 2면을 고려하는 것뿐만 아니라 3면이나 4면을 고려하는 SMVQ도 등장하였다[13].

VQ 기반의 워터마킹 기법[6, 9, 12, 14]은 대체적으로 벡터 코드북의 인덱스에 워터마크를 삽입하는 방법을 사용한다. 이를 위해서 학습단계에서는 코드북의 코드워드들을 유사한 것들끼리 그룹화하고 순서를 부여하게 된다. 그리고 압축단계에서 각 이미지의 벡터에 해당하는 코드워드를 확인하고, 삽입할 워터마크 정보에 따라서 유사한 다른 코드워드로 대체하여 표현하게 된다. 복원단계에서는 주어진 코드워드의 그룹을 구하고, 이 코드워드의 위치를 이용해서 워터마크 정보를 추출하게 된다.

SMVQ가 등장하면서 이를 기반으로 하는 워터마킹 기법들[2, 3, 11]도 최근에 제시되고 있다. 예를 들면, Chang, Wu, Tsou[3]의 연구에서는 먼저 압축된 블록들과 2면을 접하는 블록에 대하여 상태 코드북을 구한다. 그리고 이 코드북에 속한 코드워드들 중에서 그 위치가 워터마크 값과 같은 것을 결정하고, 이 위치로부터 $t, 2t$ 거리에 있는 코드워드를 함께 구한다. 이렇게 구한 세 개의 코드워드 중 블록과의 거리가 최소인 것을 결정하고, 그 거리가 사전에 설정된 한계치보다 작으면, 각각 01, 10, 11의 코드와 함께 코드워드의 값을 출력하게 된다. 블록과의 거리가 최소인 코드워드가 사전에 결정된 한계치보다 그 거리가 크다면, VQ를 이용하여 압축하게 된다.

SMVQ 기반의 제로-워터마킹 (SMVQ-Based Zero Watermarking)

전통적인 워터마킹 기법들은 공통의 약점이 있는데, 이것은 워터마크를 디지털 콘텐츠에 삽입함으로써 어느 정도의 품질 악화를 회피하기 어렵다는 것이다. 이를 극복하기 위하여 등장한 것이 제로-워터마킹[1, 5, 10]으로서, 디지털 콘텐츠에 대한 워터마크 삽입을 하지 않는 워터마킹 기법이다.

제로-워터마킹에서는 이미지를 압축하는 과정에서 워터마크와 관련된 비밀 키를 이미지의 블록별로 생성하게 된다. 이렇게 생성된 키는 신뢰할 수 있는 제3자에게 등록하여 이미지의 소유권을 증명하게 된다. 이 과정에서 압축이 적용되는 블록들은 랜덤하게 선택되게 되는데, SMVQ에서는 왼쪽 위부터 오른쪽 아래로 순차적인 압축이 진행되므로, 이러한 선택이 불가능하다. 이러한 이유로 아직까지 SMVQ에 대한 제로-워터마킹 스킴은 제안되고 있지 못하다. 이를 극복하기 위해서 이 논문에서는 이미 압축된 블록과의 접면에 대한 일치성을 동적으로 고려하는 CSMVQ를 제안한다. 이미지에서 블록을 랜덤하게 선택한다면, 접면의 수는 0개에서 4개까지 가능하다.

이 기법도 일반적인 VQ와 마찬가지로 3단계의 처리과정으로 구성된다. 첫번째인 학습단계에서는 주어진 이미지 X 를 일정한 크기의 블록으로

나누어서 코드북 $C = \{c_0, c_1, \dots, c_{N-1}\}$, $N=2^n$, $n>0$ 을 구성한다. 예를 들면 4×4 의 블록으로 나누어서 각각을 16차원의 벡터로 표현하고, 일반화된 로이드 알고리즘[8]을 이용하여 $N=512$, $n=9$ 인 코드북을 구성한다.

압축단계에서 이미지로부터 블록들을 무질서한 방식으로 추출한다. 블록 선택은 혼돈이론을 바탕으로 한 로지스틱 함수를 이용하게 되는데, 이 함수에서 생성된 값 s_i 에다 이미지의 총 블록 수 M 을 곱하여 0에서 $M-1$ 까지의 값을 생성하게 된다. 로지스틱 함수 $s_{i+1} = rs_i(1-s_i)$ 에서는 i 번째 값 s_i 를 이용하여 $(i+1)$ 번째 값 s_{i+1} 을 계산하는데, 여기서 r 은 성장률을, s_0 는 주어진 초기값을 나타낸다. 예를 들면, 16개의 블록으로 구성된 이미지는 $r=4$, $s_0=0.1$ 인 로지스틱 함수를 이용하여 아래 그림과 같이 순차적으로 선택된다.

10	7	11	14
3	1	12	8
16	5	15	13
4	9	2	6

그림2. 무질서한 블록 선택 사례

선택된 블록이 이미 압축된 블록을 하나도 이웃하고 있지 않다면, 코드북으로부터 가장 가까운 코드워드를 선택하여 설정하면 된다. [그림2]에서는 블록 1과 2가 이와 같은 순수한 VQ를 이용하여 압축된다. 만약 하나 이상의 압축된 블록이 이웃하고 있다면, 이들 블록의 접면을 이용하여 코드워드를 결정한다. [그림2]에서는 블록 1, 5, 9, 10이 두 면을 접하게 되고, 블록 14가 세 면을 접하게 되고, 나머지가 한 면을 접하게 되므로, 이들을 SMVQ를 이용하여 동적으로 압축하게 된다. SMVQ에 대한 상태 코드북은 VQ용 코드북을 수퍼 코드북으로 해서 생성된다. 접면의 수에 따라서 상태 코드북의 사이즈는 조정된다. 예를 들면, 단면 SMVQ에는 $16=2^4$ 개의 코드워드를, 양면 SMVQ에는 $8=2^3$ 개의 코드워드를, 삼면 SMVQ에는 $4=2^2$ 개의 코드워드를 사용할 수 있다.

이렇게 생성된 코드워드의 인덱스에다 XOR연산으로 워터마크를 적용하여 키를 생성하게 된다. 키를 생성하는데 사용되는 워터마크의 비트 수는 사용하는 코드북의 사이즈에 의해서 결정된다. 예를 들면, 코드북의 사이즈가 256이라고 하면, 3비트의 워터마크를 삽입할 수 있다. 어떤 블록의 인덱스가 01000001이라고 한다면, 3비트의 워터마크인 011은 10진수 3에 해당되므로, 세번째 비트를 1로 하는 이진수 00000100으로 표현할 수 있다. 이 두 이진수의 XOR를 취하여 01000101이라는 키를 생성하게 된다.

최종적으로 이렇게 생성된 비밀키들의 집합 K 는 신뢰할 수 있는 제3자에게 등록되어 그 소유권을

증명하는데 이용된다. 이미지에 대해서는 직접적으로 워터마크가 삽입되지 않았으므로, 이미지는 왜곡되지 않게 된다. CSMVQ의 전체적인 압축과정은 [그림3]과 같다.

한가지 유의할 점은, 만약 상태 코드북의 첫번째 코드워드와 블록간의 거리가 사전에 주어진 한계치를 넘어설 경우 비밀키 생성은 취소된다는 것이다. 이것은 이미지 품질의 악화와 전파를 방지하기 위한 것이다.

이상과 같이 상태 코드북을 전체적으로 이용할 경우 복원과정에는 디지털 이미지 X' 와 함께 s_0 , 비밀키 집합 K 가 필요하다. 먼저, s_0 를 이용하여 X' 로부터 블록들을 순차적으로 추출하고, 각 블록에 대하여 코드북으로부터 가장 가까운 코드워드를 결정한다. 결정된 코드워드에 XOR연산으로 워터마크를 적용한 값과 키를 비교하여 그 일치성을 확인한다.

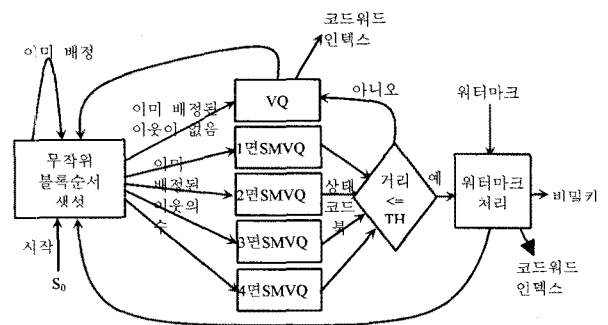


그림3. CSMVQ에 의한 압축과정

실험결과 (Experimental Results)

이 논문에서는 [그림4]와 같은 512×512 크기의 Lena와 Peppers 흑백 이미지를 사용하여 CSMVQ 기반 제로-워터마킹 기법을 테스트하였다. Lena 이미지의 경우에는 4×4 크기로 16384 개의 블록으로 분해되었다. 워터마크로는 [그림5]와 같은 64×64 픽셀로 이루어진 흑백 이진 이미지를 사용하였다. 추출된 워터마크의 효용성은 BCR(Bit Correction Rate)를 이용하여 평가하였다. BCR은 압축 시 사용된 워터마크 W 와 추출된 워터마크 W' 와의 유사도를 이용하여 다음과 같이 평가된다. 여기서 $W1$ 과 $W2$ 는 워터마크 이미지의 가로, 세로 픽셀 수를 나타낸다.

$$BCR = [\sum_i \sum_j \{W(i,j) \oplus W'(i,j)\} / W1 \times W2] \times 100\% \text{ -----(1)}$$

CSMVQ 기반 제로-워터마킹 기법의 강인함을 확인하기 위하여 JPEG 손실 압축을 이용한 공격에 대한 결과를 분석하였다.



그림4. 실험에 사용된 흑백 Lena와 Peppers 이미지



그림5. CSMVQ 압축 결과

표1. 실험 결과

Attacks	Lena	Peppers
SMVQ+JPEG(90%품질)	98.9%	98.3%
SMVQ+Sharpening	98.7%	97.2%

결론 (Conclusion)

SMVQ 이미지 압축에서는 동일하게 두 이웃 블록의 접면 정보를 이용하기 위하여 좌측 상단에서 우측 하단으로 진행되므로, 임의의 순서로 블록을 선택하여 워터마크를 삽입하기 어렵다. 이런 이유로 SMVQ에 대한 제로-워터마킹 연구가 진전되지 못하였다. 이 논문에서는 SMVQ에 제로-워터마킹을 적용하기 위한 CSMVQ(Chaotic SMVQ)를 제안하였다. 초기 실험 결과는 이 기법이 JPEG 손실 압축 등을 이용한 공격에도 강인함을 보여준다. 향후 연구로는 보다 다양한 실험을 통해서 이 체계의 특성을 심도 있게 분석하고, 활용을 위한 구체적인 개선 방안을 모색할 계획이다.

참고문헌 (References)

[1] Chang, C.-C. and Tseng, H.-W. (2004). "VQ-Based Image Watermarking Using Anti-/Gray Coding," *INFORMATICA*, Vol. 15, No. 2, pp. 147-160.

[2] Chang, C.-C., Tai, W.-L., and Lin, C.-C. (2006). "A Reversible Data Hiding Scheme Based on Side Match Vector Quantization," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 16, No. 10, pp. 1301-1308.

[3] Chang, C.-C., Wu, W.-C., and Tsou, C.-C. (2007). "Hiding Secret Information Using Adaptive Side-Match VQ," *IEICE Transactions on Information*

and Systems, Vol. E90-D, No. 10, pp. 1678-1686.

[4] Gray, R.M. (1984). "Vector Quantization," *IEEE ASSP Magazine*, Vol. 1, No. 2, pp. 4-29.

[5] Hanqiang, C., Hua, X., Xutao, L., Miao, L., Sheng, Y., and Fang, W. (2006). "A Zero-Watermarking Algorithm Based on DWT and Chaotic Modulation," *Proceedings of SPIE*, Vol. 6247, pp. 624716-1-624716-9.

[6] Jo, M.H. and Kim, H.D. (2002). "A Digital Image Watermarking Scheme Based on Vector Quantization," *IEICE Transactions on Information and Systems*, Vol. E85-D, No. 6, pp. 1054-1056.

[7] T. Kim (1992). "Side Match and Overlap Match Vector Quantizers for Images," *IEEE Transactions on Image Processing*, Vol. 1, No. 2, pp. 170-185.

[8] Lloyd, S.P., "Least Squares Quantisation in PCM," *IEEE Transactions on Information Theory*, Vol. 28, March 1982, pp. 127-135.

[9] Lu Z.M. and Sun, S.H., (2000). "Digital Image Watermarking Technique Based on Vector Quantization," *Electronic Letters*, Vol. 36, No. 4, pp. 303-305.

[10] Sang, J., Liao, X., and Alam, M.S. (2006). "Neural-Network-Based Zero-Watermarking Scheme for Digital Images," *Optical Engineering*, Vol. 45, No. 9, pp. 097006-1-097006-9.

[11] Shie, S.C., Lin, S.D., and Fang, C.M. (2006). "Adaptive Data Hiding Based on SMVQ Prediction," *IEICE Transactions on Information and Systems*, Vol. E89-D, No. 1, pp. 358-362.

[12] Wang, F.-H., Jain, L.C., and Pan, J.-S. (2007). "VQ-Based Watermarking Scheme with Genetic Codebook Partition," *Journal of Network and Computer Applications*, Vol. 30, No. 1, pp. 4-23.

[13] Wei, H.-C., Tsai, P.-C., and Wang J.-S. (2000). "Three-Sided Side Match Finite-State Vector Quantization," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 10, No. 1, pp. 51-58.

[14] Wu, H.-C. and Chang, C.-C. (2005). "A Novel Digital Image Watermarking Scheme Based on Vector Quantization Technique," *Computers and Security*, Vol. 24, pp. 460-471.