논문 2020-2-16 http://dx.doi.org/10.29056/jsav.2020.12.16

신호상관함수를 이용한 3 상태 능동적 디지털 워터마크의 개발

최용수*†

Development of 3-State Blind Digital Watermark based on the Correlation Function

YongSoo Choi*†

요 약

인터넷을 기반으로 하는 디지털 컨텐츠 응용 분야에서 디지털 콘텐츠의 보안 및 인증은 매우 중요하다. 특히, 콘텐츠의 인증을 수행하는 방법에는 여러 가지가 있다. 디지털 워터 마킹은 이와같은 방법 중 하나이다. 특히, 논문은 디지털 이미지의 응용에서 동작하는 디지털 워터마크 기반 인증 방법을 제시한다. 제안 된 워터마크는 3 상태 정보를 가지고 있으며 참조 콘텐츠 없이 임베딩 및 검출을 수행하는 블라인드 워터마킹 방법이다. 디지털 콘텐츠의 소유자 정보를 인증 할 때 자기 상관 함수를 사용한다. 또한, 주파수 영역(DWT Domain)에서 원본 콘텐츠의 신호에 적응적이도록 대역 확산(Spread Spectrum) 방식을 사용한다. 따라서 은닉정보의 검출에서 일어나는 오류의 가능성을 줄였다. DWT에서의 워터마킹은 다른 변환 방법 (DFT, DCT 등)보다 빠른 은닉/검출의 장점을 가진다. 크기가 N=m X m인 이미지를 사용할 때의 계산량은 2상태 워터마크는 O (N · logN)인 반면 O (N)으로 줄어들 수 있다. 가장 우수한 장점은 비트 당 더 많은 정보를 숨길 수 있다는 것이다.

Abstract

The digital content's security and authentication are important in the field of digital content application. There are some methods to perform the authentication. The digital watermarking is one of authentication methods. Paper presents a digital watermark authentication method that works in the application of digital image. The proposed watermark has the triple status information and performs the embedding and the detection without original Content. When authenticating the owner information of digital content, an autocorrelation function is used. In addition, a spread spectrum method is used to be adaptive to the signal of the original content in the frequency domain(DWT Domain). Therefore, the possibility of errors occurring in the detection of hidden information was reduced. it also has a advantage what Watermarking in DWT has faster embedding and detection time than other transformation domains(DFT, DCT, etc.). if it has a an image of size N=mXm, the computational amount can be reduced from $O(N \cdot logN)$ to $O(N)$. The particular advantage is that it can hide more information(bits) per bit.

한글키워드 : 블라인드 워터마킹, 상관함수, 웨이블릿 변환, 3 상태

keywords : Blind Watermark, Correlation Function, Wavelet Transformation, 3 status

* 신한대학교 기계자동차융합공학과
† 교신저자: 최용수(email: ciechoi@shinhan.ac.kr)

접수일자: 2020.12.07. 심사완료: 2020.12.16.
게재확정: 2020.12.21.

1. 서 론

최근의 통신환경에서 Internet을 통해 정보의 무제한적인 공유가 일상화 되고 있다. 이러한 환경에서 디지털 이미지, 동영상, 오디오와 같은 디지털 컨텐츠에 대한 소유권 및 사용인증에 대한 요구는 점점 높아진다. 일반적으로 소유권 보호를 위한 목적으로 암호화 방법을 주로 사용하고 암호화에서는 원본 컨텐츠를 복원해내기 위해서는 보안키를 필요로 한다. 논문에서 사용하는 디지털 워터마킹 기술은 암호화와 틀리게 대상 컨텐츠의 내용물의 지각적 형태는 최대한 유지하고 특정 코덱에 의해 소유자의 소유권 정보를 컨텐츠에 삽입하고 공유 및 판매에 사용한다. 컨텐츠에 대한 인증이 필요할 때에는 콘텐츠에 포함된 소유자의 소유권 정보를 찾아냄으로써 인증을 가능하게 한다. 콘텐츠들의 공유 및 배포 과정에서 악의적 의도를 가진 외부로부터의 훼손이 가해지거나 상업적인 목적으로 사용이 되는데 이와같은 악의적 행동을 콘텐츠에 대한 공격이라고 할 수 있다. 공격 방법들로는 일반적인 신호처리(압축, 필터링, 복사, 크기변환 등) 방법이 사용된다. 신호처리 공격에 강인하게 워터마크 신호를 유지함으로써 인증의 오류를 줄일 수 있다.

이러한 목적으로 개발된 디지털 워터마킹 기술들은 다음과 같은 특성들을 만족해야 한다.

- Invisibility: 사용자 정보를 표시하는 워터마크 삽입 후에도 원 영상의 시각적 인지가 바뀌지 않아야 한다.
- Embedding Capacity: 시각적 훼손이 이루어지지 않는 한 영상에 많은 정보를 숨길 수 있어야 한다.
- Transparency: 컨텐츠에서 추출된 정보는 정확해야 한다.
- Robustness: 일반적인 신호처리 공격(Compression, Filtering, Noise Addition, Resize 등)들에 강인해야 한다.
- Multi Collusion attack: 워터마크가 삽입된 여러 개의 목적영상의 통계적인 평균에 의한 워터마크 검출이 불가능해야 한다.

워터마킹방법들은 다양한 형태들로 제안되어 왔으며 삽입되는 공간에 따라 공간영역·변환영역 워터마크로 나눌 수 있다. 변환영역을 얻을 때는 Discrete Fourier·Discrete Cosine·Discrete Wavelet Transform 과 같은 주파수영역 변환 방법들이 주로 사용된다. 특히 채널 통신 시스템에서 주로 사용되는 대역 확산[1]방법을 응용한 기술들이 많이 개발되고 있다. 워터마크 검출 시 원본 컨텐츠의 필요 여부에 따라 블라인드, 논-블라인드 방법으로 나누기도 한다. 원본 컨텐츠가 필요한 경우 Passive, 그렇지 않은 경우 Active 워터마킹이라고 나눌 수 있다. 실제 소유권 인증을 위한 환경에서 원본 컨텐츠의 보유 가능성이나 인증 처리 시간을 종합적으로 고려해 본다면 블라인드 방법이 더 유용한 은닉방법이다. 이 논문에서는 Blind 방법을 사용한 워터마크 시스템의 구현을 하게 된다.

블라이드 워터마킹 기법에서 컨텐츠에 가해주는 주요 신호처리 공격들은 필터링, 블러링, 저주파·고주파 통과 필터, 회전, 크기변환, 압축 등이 대표적이다. 워터마크의 검출을 위한 기법들도 유사도 측정, 신호 상관도 측정, 통계적인 값의 검출 등을 사용한다. 논문에서는 워터마크 신호의 상관도[2]를 통한 검출을 사용하였다.

2. 제안된 워터마킹 시스템

논문에서는 멀티미디어 중 디지털 이미지에 대한 워터마킹 기법을 제안하였다. 변환 영역에

서의 워터마크 검출을 위해 블라인드 워터마킹을 할 것이며 신호 상관 함수를 이용하여 워터마크의 비트를 검출한다. 신호의 변환을 위해 DWT를 사용한다. 웨이블릿 변환은 변환 후에도 원본 컨텐츠의 공간적인 특성을 유지한 채 주파수 특성을 보여준다는 장점을 가지고 있어 인식과 같은 영상처리 분야에서 활용도가 높다.

변환된 영역에서 워터마킹 정보를 삽입하면 목적 영상이 만들어 지고 IDWT(Inverse DWT)를 수행 함으로써 원래의 컨텐츠 상태로 복원이 된다. 인증정보를 포함한 복원 영상은 일반적인 콘텐츠 배포·유통 과정과 동일하게 사용이되어진다. 논문에서는 배포된 컨텐츠에서 최소의 정보만을 가지고 원본영상 보유와 상관없이 상관도를 이용하여 배포물에 대한 인증을 시도하여 진위여부를 판정할 것이다. 논문에서 사용한 신호처리 공격은 Filtering, Blurring, AWGN(Additive White Gaussian Noise), JPEG Compression,), Cropping and Scaling을 사용하였으며 공격들에 대한 워터마크 검출결과를 제시하였다.

3장과 4장에서는 워터마크 삽입·검출 과정을, 5장에서는 워터마크 검출의 결과를 보인다. 마지막으로 6장에서는 제안한 워터마킹 방법과 실험결과에 대한 결론을 제시하였다.

3. 워터마크 삽입 과정

변환 영역에 워터마크를 삽입하는 기법들은 다음과 같은 워터마크 신호 삽입 공식[1]을 이용한다.

$$V'_i = V_i + \alpha\, w_i \quad (1)$$

$$V'_i = V_i(1 + \alpha\, w_i) \quad (2)$$

$$V'_i = V_i(\, e^{\alpha w_i}) \quad (3)$$

여기서 V는 원본 영상, V′은 워터마크된 영상, w는 워터마크 신호 그리고 α는 워터마크의 강도를 결정하는 요소이다.

전체적인 삽입과정은 아래의 그림(1)에 도시하였다. 원본 영상에 HVS(Human Visual System)의 채용을 위해 RGB 컬러영역에서 HSV(Hue, Saturation & Value) 포멧으로 컬러 공간 변환을 수행한다. 변환된 영역 중에서 밝기(V) 영역을 추출하여 워터마크 삽입을 위한 공간으로 사용한다. 영상의 밝기 성분은 인간의 지각적인 모델에서 가장 둔감한 신호이므로 신호 변형과 같은 응용에서 많이 사용된다. 밝기 영역(V)에 대해 2-level DWT를 수행한다. 논문에서는 2차원의 형태인 디지털 영상을 사용하므로 2-Dimensional DWT를 두 번 사용한다. 대상 영상이 N X N 크기인 경우 2-level DWT를 수행하면 4개의 N/4 X N/4크기 영역이 생성되어진다.

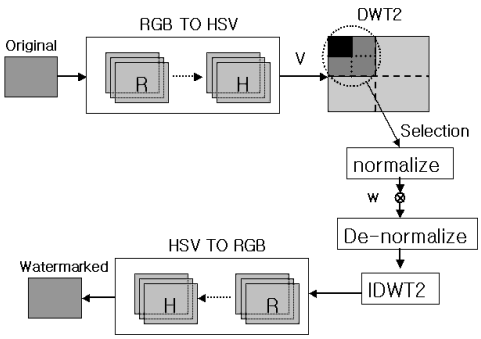

그림 1. 워터마크 삽입 과정
Fig. 1. Watermark Embedding Process

웨이블릿 영역 중에서 원본의 계수에 변형을 가했을 때 지각적 강인성, 신호의 강인성에서 가장 우수한 영역이 중-저주파수 계수(Coefficient)으므로 해당 영역들만을 추출하여 $I=\{I_1, I_2, I_3,..., I_i\}$집합을 생성한다.

선택된 영역의 계수들에 식(4)[2]를 적용하면 각 집합에 대한 정규화 집합 C가 생성되어진다.

$$C_i = \frac{I_i - \mu_I}{\sigma_I} \qquad (4)$$

$$where\ \mu_I : Mean\ of\ I,\ \sigma_I : \sqrt{Variance\ of\ I}$$

집합 C는 식(5)의 연산을 통해 워터마크가 삽입되어진 정규화 집합 C′로 생성된다. 계수들의 집합 C는 통계적으로 가우시안 분포를 따른다고 할 때, 나머지 W+와 W-도 마찬가지의 통계적 특성을 가질 것이다. 정규화된 집합 C의 통계적 특성은 워터마크 검출 과정에서 검출과 오류의 감소를 유도하는 논문의 주요 기여도가 된다.

$$C'_i = C_i + [(\alpha_1 * |C_i|) \cdot W^+_i] + [(\alpha_2 * |C_i|) \cdot W^-_i] \qquad (5)$$

식(5)에서 C는 정규화 된 원본 계수, C′는 워터마크된 계수 그리고 W+, W-는 정보비트의 부호를 결정짓는 워터마크 신호, 그리고 α는 W의 삽입강도를 결정하는 인자이다. 제안한 식(5)는 다음과 같은 특징을 가진다. 1) 원본에 적응적인 크기로 워터마크가 생성되어진다 2) 3개의 상태를 가지는 워터마크의 개발을 위해 직교하는 두 개의 워터마크 신호 집합(W+, W-)를 사용하는 것이다. 하나의 비트를 삽입한다고 할 때, +(positive)정보를 위하여 W+를 소유자 정보 삽입용 워터마크로 사용하고 상대적으로 작은 신호인 W-를 지연된 검증 집합(비트 검증용 워터마크)으로 쓰게 된다. 반대로 -(negative)정보에서는 집합의 순서를 바꾸어 W-과 W+를 사용한다. 검증용 워터마크의 사용을 하는 경우 단일하게 W 집합을 사용하는 것보다 워터마크 신호의 강인성이 높아진다. 워터마크 강도 계수 α는 SNR(Signal to Noise Ratio)=20 · log10(Signal/Noise)를 구하는 식에 의해 구해진다. 시각적인 지각성을 위해 영상의 화질저하를 일으키지 않는 임계치 SNR을 미리 정함으로써 강도 계수를 정할 수 있다. W가 삽입되어진 C′ 집합은 역정규화 과정을 거쳐 I′집합으로 변환이 되어진다. I′집합은 웨이블릿 영역에서 원래의 위치로 되돌려지고 역 웨이블릿 후에 Hue, Saturation 영역들과 함께 RGB 컬러영역으로 변환을 수행함으로써 목적영상(X′)이 만들어진다.

4. 실험 및 결과

그림2에서 논문에서 제안한 검출알고리즘을 보여준다. 배포 되어진 영상을 X^*로 정하였다. 워터마크 검출은 삽입과정과 동일한 순서를 가지며 RGB 컬러영역을 HSV 신호 영역으로 변환하고 2-level DWT를 수행한 후에 중-저주파수 영역 계수들을 선택한다. 계수들의 집합인 I^*에서 평균과 표준편차를 사용하여 신호들의 정규화가 이루어지고 워터마크 검출식 6에 대입이 된다. 식 6은 원본 워터마크신호인 W+(-)와의 자기 상관도를 구하게 된다.

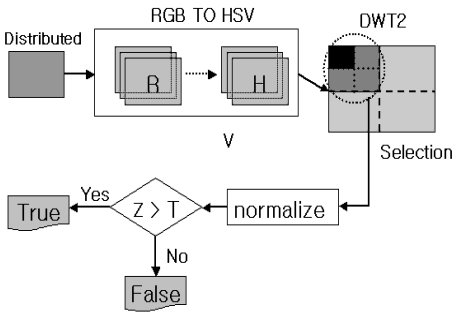

그림 2. 워터마크 검출 프로세스
Fig. 2. Watermark Detection Process

검출되는 상관값들은 집합의 평균이 이동하는 통계적인 특성을 가지게 된다. 이와 같은 통계적 특성은 여러 논문 (Modified Patchwork Algorithm[3], PatchWork Algorithm[6])에서 이미 검증이 되었다. 제안한 논문에서는 Correlation 값을 이용 집합의 통계적인 특성을

따르는 워터마크 신호 집합의 상관도를 통해 정보 비트의 상태를 결정 할 수 있다.

$$Z = \frac{1}{M} W \cdot C^* = \frac{1}{M} \sum_{i=1}^{M} W_i C^*_i \quad where, M: length\ of\ I = \frac{1}{M} \sum_{i=1}^{M} (W_i C_i + \alpha_1 * (|C_i| \quad W^+_i) W_i + \alpha_2 * (|C_i| \quad W^-_i) W_i) \quad (6)$$

검출식 6을 적용함으로써 워터마크의 자기상관도를 표시하는 Z의 +(Positive), -(Negative) 상태에 따라 정보비트 Wi의 상태 값이 결정되어 질 것이다. 예를 들어 + 정보일 경우 괄호 안의 첫 번째 항과 세 번째 항의 경우 0과 아주 작은 값을 가질 것이고 두 번째 항의 경우 Max Correlation값을 가지게 될 것이다.

다음의 식과 같이 임계값 T를 정함으로써 상관값의 크기에 따른 워터마크 판별정보를 상대적으로 변경할 수 있다.

$$T = \alpha \ \mu_{|C|}/2 \quad where, \alpha = \alpha_1 + \alpha_2 \quad (7)$$

식 7에서 α_2가 0인 경우 상관함수를 적용했을 때 Z는 임계치 T보다 클 것이다. 본 논문에서 제안한 방법은 각 신호 State에 대한 지연된 워터마크를 삽입하므로 보다 세부적인 검출 과정으로 나누어야 한다.

+ 정보일 경우 $Z^+ > (\alpha_1 \mu_{|C|})/2 = T_1$ 일 것이고 지연된 검증 집합에 대한 상관값은 $Z^- > (\alpha_2 \mu_{|C|})/2 = T_2$ 이어야 한다.

Z의 값들에 대한 가우시안 분포를 그려보게 되면 그림 3과 같이 통계적 차별성을 가짐을 알 수 있다. 그림은 + 정보에 대한 분포를 그려놓았다. 그러므로 Z+의 평균이 Z-의 평균보다 오른쪽에 위치할 것이다.

검출 정보의 상태를 결정할 때 먼저 Z+가 만족하다면 지연된 검증 집합에 대한 Z-가 만족한지를 보고, 두 조건이 모두 만족하다면 검출 비트의 상태 값은 +가 되는 것이다. - 상태 정보의 경우도 Z-의 평균이 더 높을 것이므로 Z-가 만족하면 Z+를 확인하여 두 가지 모두 만족한다면 - 상태 정보가 된다.

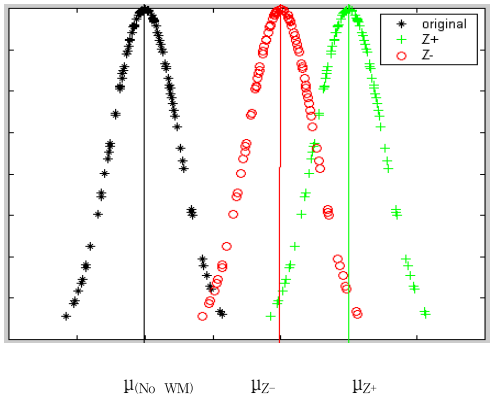

그림 3. 3상태 상관 워터마크 신호의 가우시안 분포
Fig. 3. Gaussian Distribution of 3-State Watermark Signal

실험에서는 각기 다른 주파수특성을 가지는 세 가지 영상을 원본영상으로 정했다. 정물화인 Im1, 인물화인 Im2 그리고 배경화인 Im3을 사용하였으며 신호처리 공격으로는 No Watermark, 잡음 첨가, JPEG 압축, Cropping, Scaling 공격을 사용하였다. 그리고 실험에서는 + 상태정보를 삽입하는 경우을 선택하였으므로 Z+와 Z-값의 통계적 특성을 예측하여 실험결과를 관찰 할 수 있다. - 정보에 대한 방법은 상대적인 값으로 나타나므로 생략하였다. 다음 장의 표(1)에서 그 결과들을 자세히 표기하였다.

실험에 사용된 그림들은 JPEG 포멧으로 정하였으며 신호처리 공격의 상세(Specification)는 다음과 같이 정하였다. 1) 2% Noise Addition, 2) 10% JPEG 압축, 3) 50% Scaling Up & Down, 4) 25% Cropping(워터마크된 영상의 1/4만큼을 원 영상으로 채워 넣은 경우). 공격을 받은 영상

들 중에서 지각적인 차이를 보여주는 영상들만을 표시하였다. 실험영상들에서 나타나는 지각적인 변화는 유사하므로 Im2와 Im3는 원본 영상만을 도시하였으며 결과는 표를 통하여 제시하였다.

그림 4. 원본 영상(Im1)
Fig. 4. Original Image(Im1)

그림 5. 복원된 영상
Fig. 5. Recovered Image

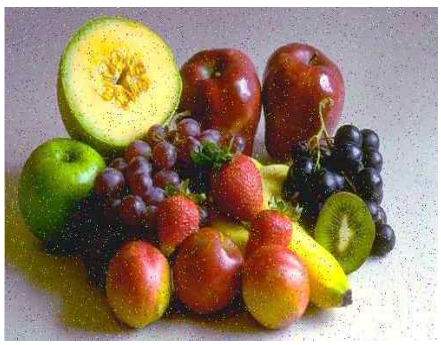

그림 6. 잡음 첨가 영상
Fig. 6. Noise Added Image

그림 7. JPEG 압축 영상
Fig. 7. JPEG Compression Image

그림 8. 자르기 영상
Fig. 8. Cropped Image

그림 9. 크기 변경 영상
Fig. 9. Scaled Image

그림 10. 원본 영상(Im2)
Fig. 10. Original Image(Im2)

그림 11. 원본 영상(Im3)
Fig. 11. Original Image(Im3)

아래의 그래프들은 각각의 공격에 대한 상관값들을 보여주고 있으며 소유자의 워터마크 ID를 25로 정하고 검증용 워터마크의 ID를 5로 정하였다. 신호처리 공격의 상황에서 워터마크의 강인성을 확인하기 위해 50개의 서로 다른 워터마크들에 대한 상관값들을 함께 도시하였다.

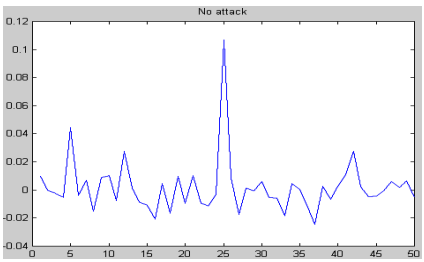

그림 12. 무공격 상관 값
Fig. 12. No Attack Correlation Values

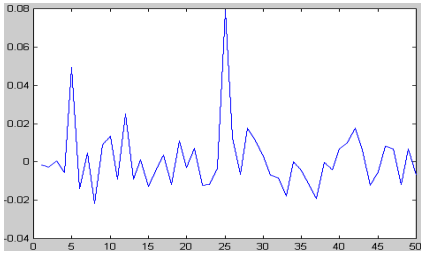

그림 13. 노이즈 공격 시 상관 값
Fig. 13. Correlation Values with Noise Addition

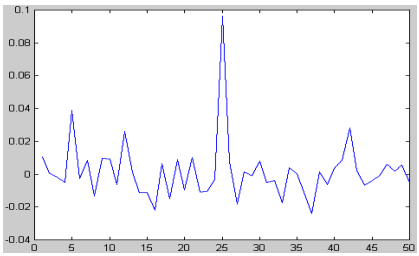

그림 14. Cropping 공격 시 상관 값
Fig. 14. Correlation Values with Cropping

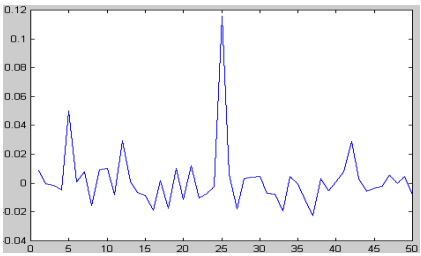

그림 15. JPEG 압축 시 상관 값
Fig. 15. Correlation Values with JEPG Compression

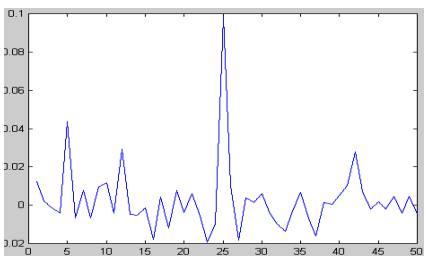

그림 16. 크기 변경 공격 시 상관 값
Fig. 16. Correlation Values with Image Scaling

실험결과의 그래프에서 미리 정한 소유자의 워터마크 ID는 강한 피크로 나타나며 검증용 워터마크 ID에서는 대부분의 상관값 보다는 크고 소유자 워터마크 상관값보다는 작은 값을 가진 것을 볼 수 있다. 표 1에서 3가지 실험영상에 대해 각종 공격 방법을 적용한 경우들의 워터마크 검출 결과를 표시하였다.

표 1. 공격 종류별 워터마크 검출 결과
Table 1. Watermark Detection Result against Attacks

공격	Im1		Im2		Im3	
	+Flag	-Flag	+Flag	-Flag	+Flag	-Flag
No	O	O	O	O	O	O
Noise	O	△	O	O	O	△
JPEG	O	O	O	O	O	O
Crop	O	O	O	O	O	O
Scaling	O	O	O	O	O	△

O: Satisfy △: Not Satisfy

5. 실험적 결론 및 향후 연구

실험에서 보였듯이 워터마크 삽입 시 + 상태 정보를 대표하는 값인 Z+(+Flag)는 모든 공격들에 대해 검출된 ID가 삽입된 소유자 워터마크 ID와 같음을 알 수 있다. Z-(-Flag)의 삽입하는 경우에도 대부분의 공격에 대해 대체로 삽입 ID와 일치하는 ID를 검출함을 볼 수 있다. 향후 강도계수인 α의 설정에 있어 보다 효과적인 도출이 이루어진다면 실험에서 발생한 검출 오류(일부 검출되지 않은 ID)도 개선이 될 것이다.

본문은 통계적인 방법과 함께 실제 검출해야 할 정보와 직교하는(Negative)값을 Flag로 사용하는 방법을 사용함으로써 정확성을 더 높일 수 있는 워터마크 삽입 알고리즘을 제안하였다. 기존의 이진 워터마크 삽입 기법들에 비해 동일 삽입영역 대비 정보은닉용량을 높일 수 있었다. 하지만 영상의 실험결과에서 확인된 것과 같이 대부분의 디지털 이미지 워터마킹 기법들처럼 회전 공격의 경우 검출률이 현저히 낮음을 알 수 있다. 이는 영상 회전의 경우 영상의 픽셀이 가지는 값의 통계적 형태가 무너져 검출신호의 상태에 오류를 보였으므로 회전공격에 대한 강인함을 위한 워터마크로 개선되어야 할 것이다.

"이 논문은 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. NRF-2017R1D1A1B03031465)"

참 고 문 헌

[1] Ingemar J. Cox, Joe Kilian, Tom Leighton and Tatal Shamoon. "Secure Spread Spectrum Watermarking for Multimedia", IEEE Trans on Image Proc, 1997. https://doi.org/10.1109/83.650120

[2] Alessandra Lumini, Dario Maio. "A blind Watermarking system for digital images in the wavelet domain", Proceeding of SPIE Vol. 3971, 2000. https://doi.org/10.1117/12.385008

[3] I. K. Yeo and H. J. Kim, "Modified Patchwork Algorithm: A Novel Audio Watermarking Scheme", ITCC in Las Vegas, Nevada, USA, 2001. https://doi.org/10.1109/ITCC.2001.918798

[4] Robert C. Dixon, WonHoo Kim(translate) "Spread Spectrum Communication System", SeHwa Press, 1995

[5] Xia-mu Niu and Sheng-he Sun. Multiresolution Digital Watermarking For Still Image, IEEE, 2000. https://doi.org/10.1109/NNSP.2000.890133

[6] W. Bender, D. Gruhl, N. Morimoto, A. Lu, “Technique for Data Hiding”, IBM System journal, Vol 35, NOS 3&4, 1996. https://doi.org/10.1147/sj.353.0313

저 자 소 개

최용수 (YongSoo Choi)

1998년 강원대 제어계측공학과 공학사
2000년 강원대 제어계측공학과 공학석사
2006년 강원대 제어계측공학과 공학박사
2006년～2007년 연세대학교 첨단융합건설 연구단 연구교수.
2007년～2013년 고려대학교 정보보호대학원 연구교수.
2013년～2020년 성결대학교 파이데이아대학(멀티미디어) 조교수
2020년~ 현재 신한대학교 기계자동차융합공학과 조교수
<주관심분야> Digital Forensics, Information Hiding, Multimedia Watermarking, Steganography