情報保護學會論文誌
第14券 第4號, 2004. 8

블록 왜곡도를 이용한 JPEG 기반의 심층암호분석

장 정 아[†], 유 정 재, 이 상 진[‡]
고려대학교 정보보호기술연구센터

A Steganalysis using Blockiness in JPEG images

JungAh Chang[†], JeongJae Yu, Sangjin Lee[‡]
Center for Information Security Technologies, Korea University

요 약

JPEG 영상을 이용하여 심층암호 통신을 하는 알고리즘의 대부분은 양자화된 DCT 계수의 최하위 비트를 치환하여 메시지를 삽입하는데, 대표적인 심층암호 알고리즘으로는 Jsteg[1], JP Hide & Seek[2], F5[3], OutGuess[4] 등이 있다. Jsteg, JP Hide & Seek 는 χ^2-테스트[4]로도 비밀데이터 삽입 여부를 탐지할 수 있지만, 탐지율이 낮은 편이다. 본 논문에서는 Fridrich의 블록 왜곡도 분석 기법[5]을 보완하여 탐지과정을 단순화하였으며 탐지율도 기존의 방식보다 향상시켰다. 또한 Jsteg, JP Hide & Seek를 이용한 실험 결과, 데이터 삽입 여부를 100%로 탐지하였다.

ABSTRACT

In general, the steganographic algorithm for embedding message in JPEG images, such as Jsteg[1], JP Hide & Seek[2], F5[3], OutGuess[4] replaces the LSB of DCT coefficients by the message bits. Both Jsteg and JP Hide & Seek are detected by χ^2-test, steganalytic technique[4], the rate of detection is very low, though.

In this paper, we propose a new steganalysis method that determine not only the existence of hidden messages in JPEG images exactly, but also the steganographic algorithm used. This method is advanced from the technique Blockiness[5]. It has many advantages that include a computational efficiency, correctness and that can detect without knowing steganographic algorithm. Experiment results show the superiority of our approach over Blockiness[5].

Keywords : *Stegangraphy, Blockiness, DCT(Discrete Cosine Transform)*

Ⅰ. 서 론

암호 통신은 데이터의 내용을 보호하고자 하는 것이 목적이므로, 데이터가 전송되고 있다는 사실 자체는 노출될 수 있다. 따라서 감시자가 암호 통신이 이루어지고 있다는 사실을 알게 되면 우선적인 감시 대상이 될 수밖에 없다. 심층 암호 기술은 암호 통신이 은닉채널을 통해 이루어지도록 함으로써 더욱 안전하게 비밀통신을 할 수 있도록 한다. 즉, 심층 암호 기술에서는 비밀통신의 존재성에 대한 의심이 들지 않도록 하는 것이 가장 중요하므로, 심층 암호 기술을 적용한 후, 은닉 데이터만 남기고 원본은 없애는 것이 일반적이다. 그러므로 심층 암호 분석기술은 탐지과정에서 은닉 데이터만 이용하여 은닉 데이터임을 구분해야 한다.

공간영역 심층 암호 기술은 BMP나 GIF와 같은 영상포맷을 이용하는데, 이러한 영상포맷은 삽입 가

접수일 : 2004년 3월 10일 ; 채택일 : 2004년 7월 15일
* 본 연구는 정보통신부 대학 IT연구센터 육성·지원사업의 연구결과로 수행되었습니다.
† 주저자 : eosjung@cist.korea.ac.kr
‡ 교신저자 : sangjin@korea.ac.kr

능한 메시지 용량이 크고 화질도 좋지만, 영상의 크기가 커서 네트워크 상에서 통신하기에 무리가 따른다. 따라서 요즘에는 압축률이 큰 JPEG과 같은 영상포맷을 네트워크 상에서 많이 이용하며, BMP나 GIF와 같은 영상포맷을 네트워크 상에서 통신하는 것은 일반적이지 않다. 그러므로 BMP나 GIF와 같은 영상포맷은 통신하는 것 자체로 의심을 일으킬 수 있으며, 이것은 심층 암호의 본래 목적에 위반된다.

따라서 심층 암호용의 영상 포맷은 널리 사용되고 있는 JPEG이 적당하다. JPEG은 공간 영역을 주파수 영역으로 변환하여 저장하기 때문에 데이터를 은닉할 때에도 주파수 영역의 값을 변경하는 것이 바람직하다. JPEG 영상을 이용하는 대부분의 심층 암호 알고리즘 역시 양자화된 DCT 계수의 최하위 비트를 치환하여, 메시지를 삽입하는데, 대표적인 것으로 Jsteg[1], JP Hide & Seek[2], F5[3], Out-Guess[4] 등이 있다

현재 JPEG 은닉영상을 탐지하는 방법으로는 Fridrich[5]의 블록 왜곡도를 이용한 탐지방법과 Westfeld 의 χ^2-테스트[4]가 있다. 본 논문에서는 JPEG 심층 암호 기술의 대표적인 삽입 알고리즘인 Jsteg, JP Hide & Seek으로 메시지를 삽입한 은닉영상을 기존의 방법과 다른 방법으로 효과적으로 탐지할 뿐 아니라, 삽입 알고리즘까지 판별할 수 있는 방법을 제안한다.

본 논문은 탐지의 대상이 되는 삽입알고리즘을 설명한 후, Fridrich의 블록 왜곡도를 응용한 새로운 은닉영상 탐지방법을 제안하고, 실험결과를 논하는 순서로 이루어진다.

II. 주파수 영역 심층 암호

JPEG 영상을 이용하는 심층 암호 알고리즘의 대부분은 양자화된 DCT 계수의 최하위 비트에 메시지를 삽입하며, 대표적인 심층 암호 알고리즘으로는 Jsteg[1], JP Hide & Seek[2], F5[3], Out-Guess[4] 등이 있다. 본 논문에서는 Jsteg, JP Hide & Seek 삽입 알고리즘을 대상으로 하는 분석방법을 제안하고자 한다.

2.1. Jsteg

Upham이 제안한 이 알고리즘[1]은 엄폐영상 전체 크기의 약 12.8% 에 해당하는 큰 삽입량을 가진다. Jsteg는 양자화 변환 후 DCT 계수의 최하위 비트를 비밀메시지로 치환하는 방식이다. 이때 DCT 계수의 값이 0, 1인 DCT 계수는 비밀메시지 삽입대상에서 제외한다.

심층 암호 통신은 암호 통신을 감추기 위한 수단으로 사용되며, 심층 암호가 해독되어 삽입된 메시지가 노출되더라도 보내고자 하는 메시지는 노출되지 않아야 하기 때문에 암호화를 거친 비밀메시지가 일반적으로 삽입된다. 따라서 비밀메시지는 암호문의 특성인 난수성이 있다고 가정한다. 따라서 JPEG 영상에 비밀메시지를 삽입하면 DCT 계수는 난수성을 갖는 비밀메시지로 치환되므로 (그림 1)과 (그림 2)에서와 같이 삽입이 이루어진 후의 DCT 계수는 이웃하는 DCT 계수의 빈도수와 비슷하게 바뀐다. 이 특성을 이용해서 탐지하는 방법이 χ^2-테스트 이다.

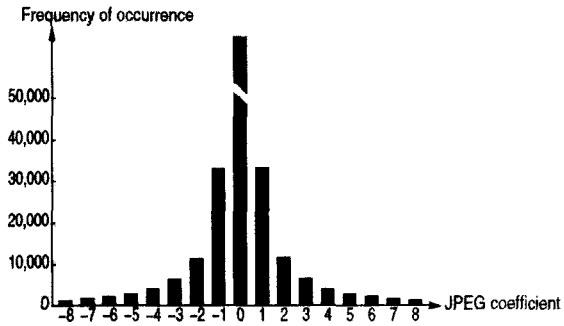


그림 1. 일반 JPEG 영상의 DCT 계수 분포

(그림 1)에서와 같이, 일반 JPEG 영상에서 0의 DCT 계수 분포가 다른 DCT 계수의 분포보다 매우 높음을 알 수 있다. 따라서 0과 1에도 비밀메시지를 삽입하면, χ^2-테스트를 적용하지 않아도, 비밀메시지의 난수성을 이용하여 0과 1의 히스토그램만 비교함으로써 은닉영상의 탐지가 가능하기 때문에 메시지 삽입대상에서 제외한다. Jsteg는 비밀메시지의 삽입위치를 정하는데 있어서 엄폐영상에 순차적으로 삽입하는 방식과 랜덤하게 삽입하는 방식이 있다. 순차적으로 삽입하는 방식은 χ^2-테스트[6]로 탐지할 수 있고 랜덤하게 삽입하는 방식은 확장된 χ^2-테스트[4]에 의해 탐지할 수 있으나 〔표 1〕에서 보는 것처럼 확률적 탐지이며, 신뢰도 p가 높아질수록 탐지율이 떨어짐을 알 수 있다.

표 1. χ^2-테스트에 의한 은닉영상 탐지율

구 분		퀄러티 팩터 = 60 %		퀄러티 팩터 = 75 %	
메시지 크기	임계값	Jsteg	JPhide	Jsteg	JPhide
6 KB	p = 75	100% (80/80)	8% (6/80)	93% (74/80)	90% (72/80)
9 KB		100% (80/80)	54% (43/80)	95% (76/80)	95% (76/80)
12 KB		100% (80/80)	33% (33/80)	100% (80/80)	100% (80/80)
6 KB	p = 91	95% (76/80)	1% (1/80)	93% (74/80)	69% (55/80)
9 KB		99% (79/80)	12% (9/80)	94% (75/80)	54% (43/80)
12 KB		100% (80/80)	28% (22/80)	95% (76/80)	90% (40/80)

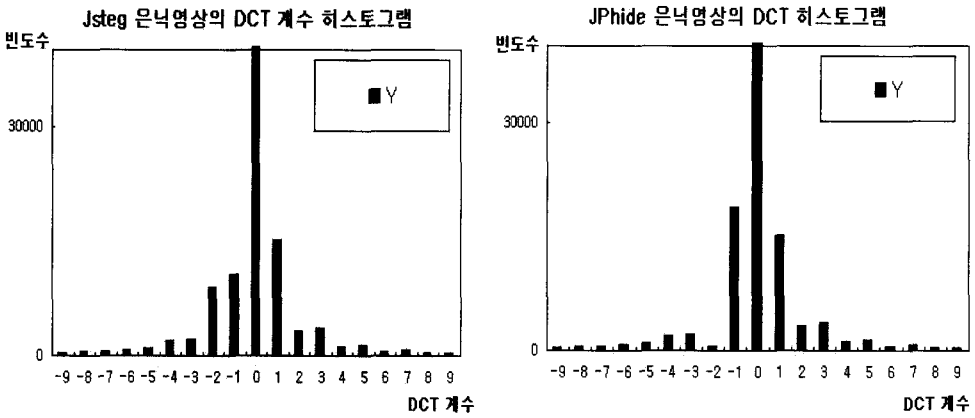


그림 2. Jsteg(좌) 및 JP Hide & Seek(우) 은닉영상의 DCT 계수 분포

2.2. JP Hide & Seek

Jsteg 삽입 알고리즘을 탐지하는 방법에는 χ^2-테스트 외에도 JSteg-like 방식 탐지 기법 [7]이 있다. (그림 1)에서와 같이 원본 JPEG 영상의 히스토그램은 -1과 1의 DCT 계수의 빈도수가 비슷하지만, Jsteg 삽입 알고리즘을 적용한 은닉영상의 히스토그램은 (그림 2)의 (좌)에서와 같이 -1과 1의 DCT 계수의 빈도수에서 차이가 발생한다. Jsteg-like 방식 심층 암호 분석 기법은 이 사실을 이용하여 은닉영상을 탐지하는 것으로, 다른 탐지기법에 비하여 탐지 속도가 빠른 장점을 갖는다. JP Hide & Seek 삽입 알고리즘[2]은 Jsteg-like 방식 탐지 기법을 보완한 것으로, Jsteg 삽입 알고리즘과 거의 유사하나 다음과 같은 차이점이 있다. Jsteg는 0과 1인 DCT 계수를 삽입대상에서 제외하고 있지만, JP Hide & Seek는 0, +1, -1인 DCT 계수를 삽입대상에서 제외하고 있다. -1과 1에서 DCT 계수의 빈도수를 삽입 전과 후에 동일하게 유지함으로써 Jsteg-like 방식 심층 암호 분석을 피하고 있으나, χ^2-테스트[4]에 의해 확률적으로 탐지된다. JP Hide & Seek는 비밀메시지를 순차적으로 삽입하는 방식을 취하고 있다.

Ⅲ. Fridrich의 블록 왜곡도 심층 암호 분석

JPEG 영상 압축 기법은 공간 영역상의 영상을 8×8 블록으로 분할하고, DCT(Discrete Cosine Transform) 변환을 거쳐 시각적으로 민감하지 않은 고주파 성분만을 제거하는 손실 압축의 일종이다. 따라서 주파수 영역의 DCT 계수 형태로 저장된 JPEG 파일을 공간영역의 영상으로 복원하게 되면, 시각적으로는 구별할 수 없지만 JPEG 영상 압축 이전의 원 영상에 비해 다소 왜곡된 형태의 영상이 된다. 이러한 영상 왜곡은 8×8 블록간의 경계에서 가장 쉽게 관찰할 수 있으며, 왜곡의 정도는 압축률이 클수록 커진다. 그리고 JPEG 영상을 이용하여

DCT 계수에 비밀메시지를 삽입하는 심층 암호의 효과도 블록들 간의 왜곡을 심화시키는 요인으로 작용한다. Fridrich는 이러한 영상 왜곡의 정도를 블록 왜곡도 라는 통계량으로 정량화하였고[5], 이 사실을 주파수 영역 심층 암호기술 분석에 활용하였다. Fridrich의 블록 왜곡도 탐지 기법은 은닉영상의 탐지뿐 아니라 삽입된 메시지의 크기를 추정할 수 있다는데 그 의의를 둔다.

블록 왜곡도 B는 M×N 크기의 영상에서 8×8 블록 경계를 이루는 이웃 화소 값들의 차를 모두 합한 것으로 식 (1)과 같다. 여기서 $g_{(i,j)}$는 (i,j)번째의 화소값 이다.

$$B=\sum_{i=1}^{\lfloor (M-1)/8 \rfloor}\sum_{j=1}^{N}|g_{(8i,j)}-g_{(8i+1,j)}| + \sum_{j=1}^{\lfloor (N-1)/8 \rfloor}\sum_{i=1}^{N}|g_{(i,8j)}-g_{(i,8j+1)}| \quad (1)$$

심층 암호 분석기술은 원본영상 없이, 실험영상만을 가지고 은닉영상임을 탐지해야 하므로, Fridrich는 원본영상에 대한 블록 왜곡도 통계량을 추정하기 위하여 비교적 DCT 변환에 덜 민감한 부분인 실험영상 블록의 중간에 위치하는 화소들, 8i+4 와 8i+5 간의 편차합을 탐지대상 영상에서의 원본영상에 대한 블록 왜곡도로 대체하였다. Fridrich의 블록 왜곡도 통계분석은 그 실행과정에 있어서, 삽입량을 추정하기 위하여 해당 삽입알고리즘을 통해 삽입과정을 반복하여야 하는 단점이 있다. 탐지과정은 다음과 같다.

1 단계. 실험영상의 블록 왜곡도, $B_s(0)$를 계산한다.

2 단계. 해당 삽입알고리즘을 이용하여 삽입 가능한 최대 메시지 크기로 삽입한 후, 영상의 블록 왜곡도, $B_s(1)$를 계산한다. 실험 영상과 이에 대응하는 최대 메시지 삽입 영상과의 블록 왜곡도 편차 $S=B_s(0)-B_s(1)$를 계산한다.

3 단계. 실험영상을 공간영역에서 4열을 잘라낸 후, 실험 영상과 같은 양자화 테이블을 이용하여 주파수 영역으로 변환한다. 4열을 자른 후의 영상에 대한 블록 왜곡도, $B(0)$를 계산한다.

4 단계. 4열을 자른 영상을 해당 삽입 알고리즘을 이용하여, 삽입 가능한 최대 메시지 크기로 삽입한 후, 영상의 블록 왜곡도, $B(1)$을 계산한다. 원본 추정 영상과 이에 대응하는 최대 메시지 삽입 영상과의 블록 왜곡도 편차, $S_0=B(1)-B(0)$을 계산한다. S_0는 원본 영상과 최대 메시지 삽입 은닉 영상과의 블록 왜곡도의 편차이므로 가장 큰 값의 편차를 가질 것이라고 예상할 수 있다.

5 단계. 4단계의 영상에 또다시 해당 삽입 알고리즘을 이용하여, 삽입 가능한 최대 메시지 크기를 삽입한 후, 영상의 블록 왜곡도, $B1(0)$를 계산한다. 이 때 Out-Guess와 같이 랜덤 구간 삽입 방식의 알고리즘의 경우에는 같은 메시지를 삽입하여도 무방하지만, Jsteg나 JP Hide & Seek과 같은 삽입 알고리즘의 경우에는 서로 다른 내용의 메시지를 사용하여 중복 삽입하자. 최대 메시지를 한번 삽입한 영상과 두 번 삽입한 영상간의 블록 왜곡도 편차, $S_1=B1(1)-B(1)$를 계산한다. 여기에서 S_1는 가장 작은 값을 가지는 블록 왜곡도의 편차로서 활용될 것이다.

6 단계. 원본으로 추정한 영상과 최대 메시지 크기로 한번 삽입한 영상 간 블록 왜곡도 편차, S_0는 블록 왜곡도가 가장 큰 두 영상들의 편차이므로, 5단계에서 계산한 S_1보다 크다. 임의의 영상과 최대 메시지 삽입 은닉 영상과의 블록 왜곡도 편차, S는 영상에 삽입된 메시지 길이에 비례한다고 볼 수 있으므로 $S\in[S_1, S_0]$라고 가정하여도 무방하다. 즉, 임의의 원본 영상에 메시지의 삽입량을 증가시킬 수록 실험 영상과 최대 메시지 삽입 은닉 영상과의 블록 왜곡도 편차는 감소하게 되므로 메시지 삽입률 p를 변수로 정점 S_0로부터 음의 기울기 $-(S_0-S_1)$를 가지는 1차 보간법으로 추정하면, $S=S_0-p(S_0-S_1)$와 같이 정리할 수 있다.

7 단계. 실험 영상에 대한 삽입 메시지 추정량 p

를 식 (2)에 의해 계산하며, 삽입 메시지 추정량 p가 0보다 크면 은닉영상이다.

$$p = \frac{S_0 - S}{S_0 - S_1} \quad (2)$$

이 탐지방법은 실험 영상의 압축률을 고려해야만 한다. 왜냐하면, 영상의 압축률이 클 경우에는-퀄러티 팩터값이 작을수록-0과 1을 제외한 DCT 계수가 감소하므로 삽입할 수 있는 최대 메시지의 길이가 짧아지게 된다. 따라서 분석 과정에 필요한 블록 왜곡도의 편차도 그만큼 감소하여 오탐지의 가능성이 증가하게 되기 때문이다. [표 2]는 퀄러티 팩터값을 75%로 고정한 JPEG 영상에 Jsteg, JP Hide & Seek 알고리즘을 이용하여 9KB의 메시지를 삽입한 각각의 은닉영상 20개를 Fridrich 의 블록 왜곡도 탐지 기법으로 실험한 결과이다. [표 2]에서 보는 것과 같이 탐지율도 매우 높고, 메시지 추정율도 매우 정확한 편이다.

표 2. Fridrich의 블록 왜곡도를 이용한 은닉영상 탐지율 (퀄러티 팩터 : 75%, 메시지 삽입량 : 9KB, 메시지 추정의 오차범위 : 10%)

구 분	탐 지 율	메시지 추정의 정확도
원본영상	100%	100%
Jsteg 은닉영상	100%	96%
JP Hide & Seek 은닉영상	90%	91%

그러나 2, 4, 6 단계에서 사용하는 해당 삽입 알고리즘을 정확히 알지 못했을 경우에는 오탐지의 가능성이 증가하게 된다. [표 3]은 해당 삽입 알고리즘을 잘못 판단하였을 경우에 Fridrich의 블록 왜곡도를 이용한 탐지 결과이다. 이 표에서 보는 것과 같이 Fridrich의 분석 기법은 적용된 삽입 알고리즘을 모르는 경우에는 은닉영상을 탐지하는 것조차 매우 어렵다.

표 3. 삽입 알고리즘 예측이 잘못된 경우 Fridrich의 블록 왜곡도를 이용한 은닉영상 탐지율(퀄러티 팩터 : 75%, 메시지 삽입량 : 9KB)

구분	탐지율
원본영상	100%
Jsteg	0%
JPhide	0%

Ⅳ. 블록 왜곡도의 특성 분석 및 새로운 탐지 방법

4.1. Jsteg 와 JP Hide & Seek의 블록 왜곡도 특성 분석

Fridrich의 블록 왜곡도[5]를 이용한 탐지기법과 제안하는 탐지방법의 큰 차이는 Fridrich가 은닉영상으로부터 원본영상을 추정하기 위하여 공간영역으로 변환한 후에 4열을 자른 영상의 블록 왜곡도를 이용한 것에 비교하여, 제안하는 방법은 이와 같은 변환 없이 JP Hide & Seek 방식으로 은닉영상에 특정 메시지를 삽입하는 것만으로, 원본영상과 은닉영상에 대한 통계적 구분을 할 수 있다.

앞 절에서 잠시 언급했듯이, 원본영상에 메시지를 삽입하는 것은 8×8 블록 경계에서 왜곡도를 심화시키는 결과를 낳는다. 따라서 메시지가 삽입되어 있지 않은 JPEG 영상에 암호문을 삽입하면, 블록 왜곡도가 증가한다. 그러나 이미 난수성이 있는 암호문이 삽입된 은닉영상이라면, DCT 계수의 최하위 비트에 또 다시 난수를 삽입하더라도 원본영상에서 증가한 것 보다 적게 블록 왜곡도가 증가하는 결과가 나타날 것이며, 원본영상에 난수를 삽입했던 영상보다 적은 영상 블록간의 왜곡 편차를 가지게 된다. 이러한 영상의 블록간 왜곡 편차를 이용하면 원본영상과 은닉영상을 효과적으로 구별할 수 있다.

Jsteg[1]와 JP Hide & Seek[2] 삽입 알고리즘은 DCT 계수의 최하위 비트를 비밀메시지로 치환하는 방식이다. 따라서 삽입 가능한 최대 메시지 크기로 암호문을 Jsteg나 JP Hide & Seek 방식으로 삽입한다는 것은 영상 블록간의 왜곡을 심화 시키는 요인이라고 볼 수 있다. 따라서 원본영상의 블록 왜곡도는 은닉영상의 블록왜곡도보다 항상 작다.

원본영상에 Jsteg나 JP Hide & Seek 방식으로 0(또는 1로)만 이루어진 메시지를 삽입하는 것은 암호문을 삽입하는 것과 같이 블록 왜곡도를 증가시키는 방향으로 작용하기 때문에, 원본영상의 블록 왜곡도는 Jsteg 방식이나 JP Hide & Seek 방식으로 0(또는 1로)만 이루어진 메시지를 삽입한 영상의 블록 왜곡도 보다 항상 작다.

그러나 은닉 영상은 이미 암호문을 포함하고 있기 때문에 상황이 다르다. 은닉영상에 Jsteg방식이나 JP Hide & Seek방식으로 0(또는 1)로만 이루어진 메시지를 삽입하는 것은 암호문을 삽입하는 것과

는 다르게, 블록 왜곡도를 증가시킬 수도, 감소시킬 수도 있다. 이러한 블록 왜곡도의 증감은 삽입된 메시지 크기에 영향을 받게 되는데, 실험한 바에 의하면 삽입 가능한 최대 메시지 크기의 33%(3KB) 이상에 해당하는 메시지를 Jsteg로 삽입한 은닉 영상의 경우 그 영상에 JP Hide & Seek 방식을 이용해 1로만 이루어진 메시지를 삽입했을 때 블록 왜곡도가 감소하는 결과를 나타내었다.

제안하는 방식처럼 0(또는 1)로만 구성된 메시지를 이용하여 DCT계수의 최하위 비트에 삽입하는 이유는 다음에서 설명할 특성탐지를 하기 위해서이다.

〔그림 1〕은 일반영상의 DCT 계수 히스토그램이고 〔그림 2〕는 Jsteg 은닉영상과 JP Hide & Seek 은닉영상의 DCT 계수 히스토그램이다. 이 3개의 그림에서 주목해야 할 부분은 -2와 -1에서 DCT 계수의 빈도수 차이다. 이러한 차이는 DCT 계수 -1을 메시지 삽입에 이용하느냐, 마느냐에 따라 발생하게 되므로 이 차이를 이용하면 메시지 삽입 알고리즘을 구별할 수 있게 된다.

JPEG DCT 계수 중 영상의 화질에 가장 큰 영향을 미치는 계수는 〔그림 1〕에서 보는 바와 같이 가장 많은 빈도 분포를 가지는 -1, 0과 1 이다. 그런데 Jsteg방식으로 메시지를 삽입한 은닉 영상의 경우에는 DCT 계수 -1(0XFF, LSB=1)를 메시지 삽입에 이용하게 되므로 〔그림 2〕의 좌측과 같이 -1에 대한 LSB의 보수인 -2(0XFE, LSB=0)도 메시지 삽입량에 따라 증가함을 볼 수 있다. 이때 이미 Jsteg방식으로 메시지를 삽입한 은닉 영상에 JP Hide & Seek 방식으로 전부 0으로만 이루어진 메시지 0", 혹은 전부 1로만 이루어진 메시지 1"를 중복해서 삽입한 후 발생되는 변화를 살펴보도록 하자. 주지하는 바와 같이 JP Hide & Seek는 메시지 삽입에 DCT 계수 -1, 0, 1을 이용하지 않는다. 먼저 0"을 삽입하는 경우에는 DCT 계수 -1, 0, 1을 제외한 홀수의 DCT 계수들이 그와 LSB의 보수 관계에 있는 짝수 DCT 계수로 변환되므로 적은 양만의 계수들만 변화하고 DCT 계수 -2는 변환되지 않는다. 따라서 그림 3에서 보는 바와 같이 원래 은닉 영상의 블록 왜곡도와 그다지 큰 차이를 가지지 않는다. 그러나 같은 은닉 영상에 JP Hide & Seek 방식으로 1"을 삽입하는 경우에는 상황이 달라진다. -2를 포함한 짝수의 DCT 계수들이 그와 LSB의 보수 관계에 있는 홀수 DCT 계수로 변환되므로 매우 큰 빈도 분포를 가졌던 DCT 계수 -2 전부가 -1로 변환되면서 메시지를 삽입하지 않은 원본 영상의 DCT 계수 분포와 유사한 분포를 가지게 되고 〔그림 3〕에서 보는 바와 같이 블록 왜곡도는 크게 감소하는 결과로 나타난다.

반면, JP Hide & Seek방식으로 메시지를 삽입한 은닉 영상의 경우에는, 메시지 삽입 시 DCT 계수 -1, 0과 1을 이용하지 않으므로 0"이나 1"인 메시지를 삽입한 후 -1, 0과 1 다음으로 많은 빈도 분포를 가지는 DCT 계수 -2와 2에 주목하여야 한다. 이 때 고려해야 될 사항은 DCT 계수 -2(0XFE, LSB=0)와 LSB의 보수 관계에 있는 DCT 계수는 -1(0XFF, LSB=0)이고, DCT 계수 2(0X02, LSB=0)와 LSB의 보수 관계에 있는 DCT 계수는 3(0X03, LSB=1)이라는 점이다. JP Hide & Seek방식으로 1"인 메시지를 삽입한다면 DCT 계수 -2의 경우에는 전부가 DCT 계수 -1로 변환되지만 DCT 계수 2는 3으로 변환되어 DCT 계수 1에는 변화가 없으므로 영상의 화질에 민감한 DCT 계수인 -1과 1의 편차는 은닉 영상에 존재하였던 편차보다 더욱 증가하게 된다. 따라서 〔그림 4〕에서와

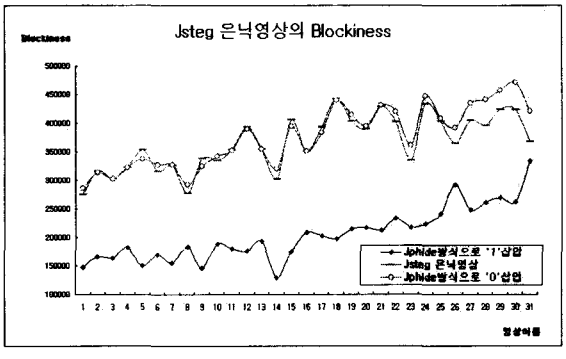


그림 3. Jsteg 은닉 영상의 블록 왜곡도

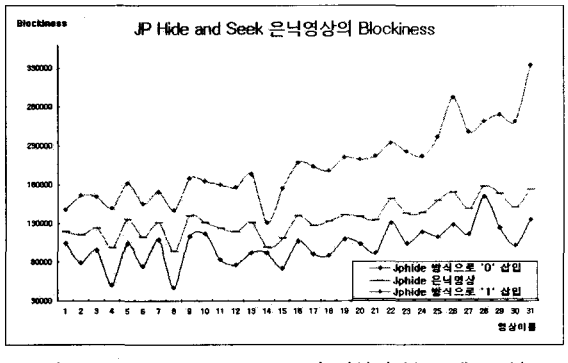


그림 4. JP Hide & Seek 은닉 영상의 블록 왜곡도분포

같이 JP Hide & Seek방식으로 1"인 메시지를 삽입한 은닉 영상의 블록 왜곡도가 원래의 은닉 영상보다 큰 것을 볼 수 있다. 그러나 JP Hide & Seek 방식으로 0"인 메시지를 삽입하면 절대값이 홀수인 DCT 계수들은 그와 LSB의 보수 관계에 있는 짝수 DCT 계수로 변환되고 원래 영상에서의 DCT 계수 -1과 1의 편차도 유지시킬 수 있다. 또한 DCT 계수 -2를 제외하면 짝수의 DCT 계수만이 남은 상황에서는 절대값이 증가함에 따라 빈도수가 감소하는 일반적인 JPEG 영상 특성을 가지게 되어 〔그림 4〕에서 보는 바와 같이 원래의 은닉 영상의 블록 왜곡도보다 적은 값을 나타낸다 - 〔그림 2〕의 (우)에서 보면 JP Hide & Seek 방식은 메시지 추출을 위하여 특히 DCT 계수 -2의 LSB의 보수는 -1이므로 삽입할 메시지 1인 경우에는 1/2의 확률로 다른 DCT 계수에 중복하여 메시지 비트 1을 삽입한다. 따라서 자세히 살펴보면 절대값이 홀수인 DCT 계수의 발생 빈도가 그와 LSB의 보수 관계에 있는 짝수 DCT 계수보다 큼을 볼 수 있다.

순차-Jsteg, 랜덤-Jsteg 및 JP Hide & Seek을 적용한 은닉영상을 대상으로, 제안하는 방법으로 은닉영상을 탐지하는 실험을 해보았다. 원본영상의 블록 왜곡도는 은닉영상의 블록 왜곡도보다 작을 것이라는 예상 및 Jsteg, JP Hide & Seek 은닉영상에 JP Hide & Seek 삽입알고리즘 방식으로 0, 1을 삽입하기 전과 후의 블록 왜곡도의 관계에 대한 예상을 〔그림 3〕과 〔그림 4〕에서 확인할 수 있다.

따라서 제안하는 방법으로 은닉영상과 원본영상을 구분할 수 있으며, Jsteg 심층 암호 기술을 적용한 은닉영상과 JP Hide & Seek 심층 암호 기술을 이용한 은닉영상을 구분할 수 있다. 즉, 원본영상의 블록 왜곡도는 은닉영상의 블록 왜곡도보다 작으며, Jsteg 심층 암호 기술을 적용한 은닉영상의 블록 왜곡도는 실험영상에 JP Hide & Seek 방식으로 1을 삽입한 후의 블록 왜곡도 보다 크며, JP Hide & Seek 심층 암호 기술을 이용한 은닉영상의 블록 왜곡도는 실험영상에 JP Hide & Seek 방식으로 0을 삽입한 후의 블록 왜곡도 보다 크다.

4.2 새로운 탐지 방법 및 실험결과

Jsteg 및 JP Hide & Seek 심층 암호 기술을 이용하여 JPEG 영상의 주파수 영역에 비밀메시지를 삽입한 은닉영상를 효과적으로 탐지하는 방법은 다음과 같으며, Fridrich 의 블록 왜곡도 계산식 (1)을 이용한다.

1 단계. 실험영상의 블록 왜곡도, $B(test)$를 계산한다.

2 단계. 실험영상에 JP Hide & Seek 방식으로 최대 삽입량 크기의 메시지를 삽입한 후, 영상의 블록 왜곡도, $B(JPhide_0)$를 계산한다. 이때 메시지는 0으로만 이루어진 문자열이다.

3 단계. 실험영상에 JP Hide & Seek 방식으로 최대 삽입량 크기의 메시지를 삽입한 후, 영상의 블록 왜곡도, $B(JPhide_1)$를 계산한다. 이때 메시지는 1로만 이루어진 문자열이다.

4 단계. 영상의 블록 왜곡도 $B(test)$가 $B(JPhide_0)$,

그림 5. 원본 lena.jpg

그림 6. Jsteg 삽입 알고리즘을 적용한 lena.jpg

$B(JPhide_1)$의 2가지 값보다 작으면 원본 영상으로 출력한다. 만약 $B(test)$가 $B(JPhide_0)$보다 크면 JP hide & seek 심층 암호 기술을 이용한 은닉영상으로, $B(test)$가 $B(JPhide_1)$보다 크면 Jsteg 심층 암호 기술을 이용한 은닉영상으로 출력한다.

다음은 인터넷 웹사이트 http://www.cs.washington.edu/research/imagedatabase/groundtruth/[9]에서 얻은 768×512 크기의 JPEG 영상과 디지털 카메라 MINOLTA DIMAGE F200로 획득한 영상을 대상으로, 퀄러티 팩터 60%, 75% 별로 각각 80 개씩 Jsteg와 JP Hide & Seek 삽입 알고리즘을 이용하여 6KB, 9KB, 12KB의 메시지를 삽입한 후 제안하는 탐지기법으로 실험한 결과이다.

〔표 4〕는 제안하는 방법으로 탐지한 은닉영상 탐지율을 나타낸 것이고, 〔표 5〕는 제안하는 방법으로 탐지한 특성 탐지율을 나타냈다. 〔표 6〕은 원본영상의 퀄러티 팩터를 60%로 조정한 후, Jsteg, JP Hide &Seek 삽입 알고리즘을 이용하여 비밀메시지 12KB를 삽입한 은닉영상 각각 80개를 동일한 퀄러티 팩터를 적용한 원본영상 80개와 섞은 후, 총 240개의 영상을 제안하는 방법으로 탐지한 결과를 나타낸 것이다. 〔표 6〕에서 보는 것과 같이 원본영상과 은닉영상은 100%의 정확도로 구분하며, 특성 탐지는 98%의 정확도로 탐지한다.

표 4. 제안하는 방법에 의한 은닉영상 탐지율

메시지 크기	퀄러티 팩터 = 60%		퀄러티 팩터 = 75%	
	Jsteg	JPhide	Jsteg	JPhide
6KB	100% (80/80)	100% (80/80)	100% (80/80)	100% (80/80)
9KB	100% (80/80)	100% (80/80)	100% (80/80)	100% (80/80)
12KB	100% (80/80)	100% (80/80)	100% (80/80)	100% (80/80)

표 5. 제안하는 방법에 의한 특성 탐지율

메시지 크기	퀄러티 팩터 = 60%		퀄러티 팩터 = 75%	
	Jsteg	JPhide	Jsteg	JPhide
6KB	97% (78/80)	90% (72/80)	79% (63/80)	65% (52/80)
9KB	97% (78/80)	90% (72/80)	97% (78/80)	95% (76/80)
12KB	99% (79/80)	100% (80/80)	97% (78/80)	100% (80/80)

표 6. 원본영상 대비 은닉영상 탐지율(퀄러티 팩터 : 60%, 메시지 크기 : 12KB)

원본 영상	Jsteg	JP Hide & Seek
100%	100% (98%)	100% (98%)

또한 False positive와 False negative는 0%이다.

V. 결 론

JPEG 영상을 이용한 대표적인 심층 암호 알고리즘으로는 Jsteg[1], JP Hide & Seek[2], F5[3], OutGuess[4] 등이 있다. 본 논문에서는 Fridrich의 블록 왜곡도 분석 기법[5]을 보완하여 새로운 탐지기법을 제안하였다. 본 논문에서 제안한 블록 왜곡도를 이용한 특성 탐지기법으로 Jsteg, JP Hide & Seek 삽입알고리즘을 대상으로 하여 은닉영상을 탐지한 결과 메시지 삽입 여부는 100%로 탐지하였으며, 특성 탐지율은 98% 정도의 매우 높은 확률로 탐지한다.

Fridrich의 블록 왜곡도를 이용한 심층 암호 분석방법에 비교하여, 본 논문에서 제안하는 방법은 다음과 같은 장점이 있다.

원본영상 없이 은닉영상임을 탐지하는 것이 일반적인 현실에 비추어 보았을 때, 적용된 삽입 알고리즘을 알아야 하는 Fridrich의 분석기법은 실제로 은닉 영상을 탐지하는 방법으로는 적합하지 않다. 이에 비교하여, 본 논문에서 제안한 탐지방법은 삽입 알고리즘을 미리 예측하는 것과 같은 사전 작업 없이도, 매우 높은 확률로 은닉영상 탐지와 동시에 사용한 심층 암호의 종류를 탐지할 수 있다는 장점이 있다.

또한, 제안하는 방법은 공간영역과 주파수 영역간의 변환과정이 없기 때문에 탐지과정이 간단하고, 공간영역에서의 연산량도 적으므로 효율적이다.

앞으로 다양한 실험을 통하여 OutGuess[8]나 F5[3] 같은 χ^2-테스트[4]에 강한 심층 암호 기술에도 적용하여, 은닉영상 탐지 및 특성탐지와 메시지 삽입량 추정을 위한 연구가 계속될 예정이다.

참 고 문 헌

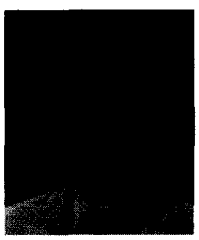
〔1〕 D. Upham, "Jsteg", http://www. security-focus.com/tools/1434, 1997.

[2] A. Latham, "JP Hide & Seek", http://linux01.gwdg.de/~alatham /stego.html. 1997.
[3] A. Westfeld and A. Pfitzmann, "High Capacity Despite Better Steganalysis F5-A Steganographic Algorithm", Information Hiding-4th International Workshop Lecture Notes in Computer Science 2137, pp. 289-302, 2001.
[4] N. Provos, "OutGuess - Universal Steganography",http://www. outguess. org / ,August 1998.
[5] J. Fridrich and M. Goljan and D. Hogea, "Attacking the OutGuess", Proc. of the ACM Workshop on Multimedia and Security, 2002.
[6] A. Westfeld and A. Pfitzmann, "Attacks on Steganographic Systems", Information Hiding-Third International Workshop, IH'99, Lecture Notes in Computer Science 1768, pp. 61-76, 1999.
[7] T. Zhang and X. Ping, "A Fast and Effective Steganalytic Technique against JSteg-like Algorithms ", In Proceedings of ACM Symposium on Applied Computing, ACM Press, pp. 307-331, 2003.
[8] N. Provos, "Defending Aganist Statistical Steganalysis", Proceedings of the 10th USENIX Security symposium, pp. 323-335, 2001.
[9] http://www.cs.washington.edu/research /imagedatabase/groundtruth/

〈著者紹介〉

장 정 아 (Jung-Ah Chang) 학생회원
2002년 2월 : 인하대학교 컴퓨터 공학부 졸업
2003년 3월 : 고려대학교 정보보호 대학원 석사 과정
〈관심분야〉 정보은닉이론, 디지털 워터마킹, 패턴 인식

유 정 재 (Jeong-Jae Yu) 학생회원
1998년 8월 : 고려대학교 수학과 졸업
2003년 8월 : 고려대학교 정보보호 대학원 석사
2003년 9월~현재 : 고려대학교 정보보호 대학원 박사 과정
〈관심분야〉 정보은닉이론, 디지털 워터마킹, 정보보호 프로토콜



이 상 진 (Sangjin Lee) 정회원
1987년 2월 : 고려대학교 수학과 학사
1989년 2월 : 고려대학교 수학과 석사
1994년 2월 : 고려대학교 수학과 박사
1989년 2월~1999년 2월 : 한국전자통신연구원 선임 연구원
1999년 2월~2001년 8월 : 고려대학교 자연과학대학 조교수
2001년 9월~ 현재 : 고려대학교 정보보호대학원 부교수
〈관심분야〉 대칭키 암호의 분석 및 설계, 정보은닉이론, 컴퓨터 포렌식스