

분산 확대에 기반한 LSB Steganalysis에 관한 기술 연구

박태희*

요약

본 논문은 비트 분해에 의한 분산 확대에 기반한 변형된 LSB 스테고 영상의 steganalysis 기법을 제안한다. LSB 임베딩은 가장 간단하고 쉬운 steganography 방법이나 최근 변형된 임베딩 방법으로 인해 비밀 메시지의 삽입 유무를 판단하기 어렵다. 따라서 본 논문에서는 영상 픽셀을 상위 및 하위 4비트로 분해함으로써 영상의 잡음 분산을 확대한다. 또한 분해된 두 개의 부영상에 대해 3-레벨 웨이블릿 변환을 수행하고, 분해된 부영상 및 분할된 각 부대역의 특성 함수로부터 84차의 통계적 모멘트를 추출한다. 추출된 특징 벡터는 MLP 분류기에 입력되어 커버 영상과 스테고 영상을 학습하고 판별한다. 제안 방법의 성능 평가를 위해 변형된 LSB 임베딩에 의한 다양한 삽입률의 스테고 영상을 사용하였으며, 실험 결과 기존의 기법에 비해 삽입 정보 유무의 검출율을 향상시킬 수 있음을 확인하였다.

A Study on the LSB Steganalysis method based on variance extension

Tae-Hee Park*

ABSTRACT

This paper proposes LSB steganalysis method based on variance extension by the bit separation. LSB embedding is the simplest and easiest among many steganography schemes. But it is difficult to detect secret message due to modified embedding methods. To improve the detection rate to existence of hidden message, we expand the noise variance of image by separating pixels into upper and lower 4bits. Then we decompose each separated sub-images by using 3-level wavelet transform, extract 84D statistical moments from characteristic function of original and subbands. Multi-layer perceptron(MLP) classifier is used to discriminate between cover and stego image. We experiment our proposed scheme on modified LSB embedding method. The proposed method improves the detection rate to existence of hidden message.

Key Words : steganography, steganalysis, LSB embedding, noise variance, characteristic function, MLP

* 동명대학교 메카트로닉스공학과(✉thpark77@dreamwiz.com)

· 제1저자(First Author) : 박태희 · 교신저자(Correspondent Author) : 박태희

· 접수일(2010년 7월 6일), 수정일(1차 : 2010년 9월 7일), 게재확정일(2010년 9월 13일)

1. 서 론

최근 인터넷과 멀티미디어 기술이 발달함에 따라 이미지 파일이나 MP3 파일 등 다양한 디지털 매체에 비밀 메시지를 숨기기 위한 관심이 증가하고 있다.

비밀 메시지를 기록하기 위한 오래된 방법 가운데 steganography는 고대 그리스어인 “감춰진 기록(covered writing)”에서 유래한 것으로서 암호(encryption)와는 달리 메시지가 있다는 사실 자체를 숨김으로서 정보를 은닉하는 기술이다[1-4].

steganography는 과거에는 주로 군사적 또는 국가적 비밀을 보호할 목적으로 사용되었다. 고대 그리스인들이 페르시아의 공격을 알리기 위해 비밀 메시지를 목판에 보이지 않게 새겼던 것이나 제 1차 세계 대전시 독일군이 편지의 여백에 군사 비밀을 몰래 기록하여 보낸 것은 steganography를 이용한 대표적인 예이다[3-4]. 그러나 최근들어 기업의 핵심 기술 및 고객 정보가 외부로 유출되고, 디지털 영상이나 문서 또는 음악 파일에 불법 음란물을 숨겨 판매하는 사례가 증가되면서 첨단 기술에 대한 보안 사항 중 steganography를 통한 기술 유출은 중요한 사회적 문제로 대두되었다.

이러한 의미에서 steganalysis는 steganography를 사용하여 숨겨진 비밀 메시지의 존재를 감지하는 것을 목적으로 한다[5]. steganalysis의 기본적인 원리는 숨겨진 메시지가 인간의 시각에는 잘 보이지 않을지라도 신호의 통계적인 분포에는 영향을 줄 수 있다는 것이다. 따라서 steganalysis에서 커버 신호(비밀 메시지의 삽입 전)와 스테고 신호(비밀 메시지의 삽입 후)에 통계적인 특징(feature)의 차를 발견하는 것은 매우 중요하다.

steganalysis에 관한 대표적인 연구로서 Lyu 등은 칼라 영상의 각 RGB 채널에 대해 2-레벨의 QMF(Quadrature Mirror Filters)를 이용하여 영상을 분해하고, 웨이블릿 도메인으로부터 216차의 통계적

인 특징을 추출하였다. 이때 4차의 통계적 모멘트(mean, variance, skewness, kurtosis)는 각 부밴드 및 최적화된 선형 예측기로부터 추정된 계수값과 계수의 로그값들의 오차에 대해 계산되었다[5].

Xuan 등은 3-레벨 haar 웨이블릿 변환을 사용하여 영상을 4개의 부밴드로 분해하고 원 영상 및 각 부밴드 계수의 특성 함수에 대해 39차의 특징을 추출하였다[6].

2008년 Mehrabi 등은 영상으로부터 몇 개의 MSB(Most Significant Bit)를 제거하고 이로부터 39차의 통계적 모멘트를 추출하는 steganalysis 방법을 제안하였다[7].

그러나 기존의 steganalysis 방법들은 비밀 메시지에 따라 하위 1비트만 변화시키는 기본적인 LSB steganography 영상[8]에만 좋은 성능을 보일 뿐 LSB matching[9] 또는 하위 L 개의 비트를 변화시키는 Bit Plane Complexity Segmentation Steganography(BPCS)[10], 그리고 랜덤 LSB steganography[11] 등 변형된 LSB 스테고 영상에는 적용하기가 어렵다.

따라서 본 논문에서는 변형된 LSB 스테고 영상에서 비밀 메시지의 삽입 유무의 검출율을 높이기 위해 분산 확대에 기반한 개선된 steganalysis 기법을 제안한다. 먼저 영상의 잡음 분산을 확대하기 위해 각 픽셀을 상위 4비트와 하위 4비트로 분해하고, 비트 분해된 두 개의 독립된 영상에 대해 3-레벨 undecimated 웨이블릿 변환을 수행한다. 또한 분할된 각 부대역의 계수값으로부터 히스토그램의 특성 함수를 추출하고 이로부터 84차의 통계적 모멘트를 추출한다. 추출된 특징 벡터는 MLP(Multi-Layer Perceptron) 분류기에 입력되어 커버 영상과 스테고 영상을 학습하고 판별한다.

본 논문의 구성은 다음과 같다. 2장에서는 기본적인 LSB 및 변형된 LSB steganography 방법에 대해 살펴보고, 변형된 LSB 스테고 영상에 대한 비트 분해된 부영상의 공간 및 웨이블릿 부대역의 통계적인 특성에 대해 살펴본다. 3장에서는 비트 분해된 영상의 분산

확대에 대해 설명하고, 제안한 steganalysis 방법을 설명한다. 4장에서는 실험을 통해 제안 기법의 성능을 확인하고, 5장에서는 결론을 맺는다.

II. LSB steganography

2.1 변형된 LSB steganography

LSB 임베딩은 가장 일반적이고 자주 사용되는 steganography 방법이다. LSB 임베딩 기법은 커버 영상 픽셀의 LSB(Least Significant Bit)를 비밀 메시지 비트로 직접 대체함으로써 LSB 플레인을 조절한다.

그레이 영상의 픽셀값 p 를 이진 표현(binary representation)으로 표기할 때 식 (1)과 같다고 하자.

$$p = b_m b_{m-1} \cdots b_n \cdots b_1 \quad (1)$$

여기서 $p \in \{0, \dots, 255\}$ 는 영상의 픽셀값을 나타내며, b_m, b_n, b_1 은 픽셀값 p 를 이진수로 표현했을 때 각각 최상위 비트, n 번째 하위 비트, 최하위 비트를 의미한다.

대부분의 LSB 임베딩 방법은 b_1 을 변경하나 간단한 LSB와 BPCS(BitPlane Complexity Segmentation)는 비밀 메시지의 삽입을 위해 n 개의 하위 비트를 사용한다. 그러나 이러한 LSB 임베딩은 구현이 간단하다는 장점에도 불구하고 다양한 공격에 약하다는 단점이 있다.

변형된 LSB 방법은 다양한 공격에 강한 steganography 기법으로서 랜덤한 순서로 비밀 메시지를 n 개의 비트에 삽입하게 된다. 이를 통해 스테고 영상은 통계적으로 보다 더 잡음처럼 보이게 된다.

2.2 변형된 LSB 스테고 영상의 비트 분해

먼저 8비트 그레이 영상의 픽셀 $p = b_8 b_7 \cdots b_1$ 을

식 (2) 및 (3)과 같이 상위 4비트와 하위 4비트로 분해한 부영상을 각각 p_{MSB}, p_{LSB} 라 가정하자.

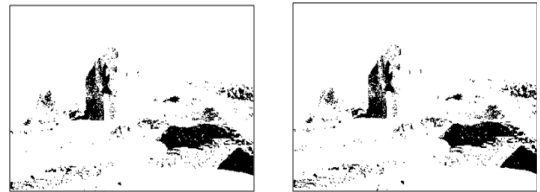
$$p_{MSB} = b_8 \cdots b_5 \quad (2)$$

$$p_{LSB} = b_4 \cdots b_1 \quad (3)$$

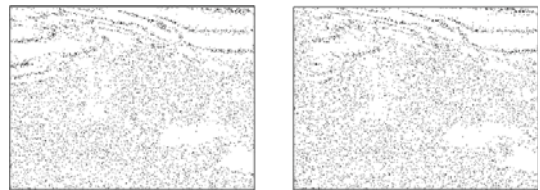
그림 1은 커버 영상 및 변형된 LSB 스테고 영상에 대한 부영상 p_{MSB}, p_{LSB} 을 보인 것이다.



(a) 원 영상



(b) p_{MSB} 영상

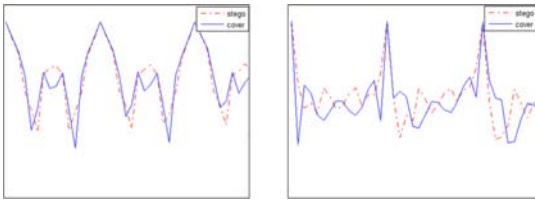


(c) p_{LSB} 영상

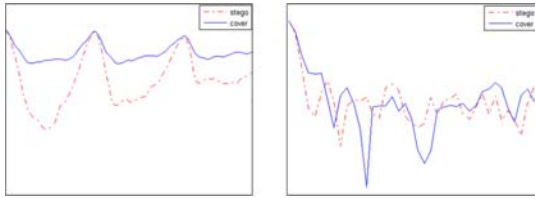
그림 1. 비트 분해된 커버 및 변형된 LSB 영상 (0.3bpp)(왼쪽: 커버영상, 오른쪽: 변형된 LSB 영상)
Fig. 1. bit decomposed cover and modified LSB image (0.3bpp)(left : cover, right : modified LSB stego image)

또한 그림 2는 비트 분해된 부영상 및 웨이블렛 LH 부대역의 특성 함수 그래프를 비교한 것이다. 여기서 실선은 커버 영상, 점선은 변형된 LSB 스테고 영상을

의미한다.



(a) 비트 분해 영상의 특성 함수



(b) 비트 분해 영상의 LH 부대역의 특성 함수

그림 2. 비트 분해된 커버 및 변형된 LSB 스테고 영상의 특성 함수 비교(왼쪽 : P_{MSB} , 오른쪽 : P_{LSB})

Fig. 2 Comparison of Characteristic function of bit decomposed cover and modified LSB stego image (left : P_{MSB} , right : P_{LSB})

그림 1 및 그림 2로부터 커버 및 변형된 LSB 스테고 영상은 비트 분해된 부영상 P_{MSB} 와 P_{LSB} 영상에서 공간적인 차이가 있으며, 특히 웨이블릿 변환 영역에서 특성 함수의 차이가 더 크음을 알 수 있다. 이는 상위 및 하위 부영상 모두에서 커버 및 스테고 영상을 분류하기 위한 뚜렷한 특징 벡터를 추출할 수 있음을 의미한다.

III. 제안 방법

앞 장에서는 커버 및 스테고 영상을 비트 분해했을 때 특성 함수의 차이가 크게 나타남을 알 수 있었다. 이 절에서는 비트 분해에 의한 커버 및 스테고 영상의 분산 확대를 수학적으로 증명하고, 비트 분해된 두 개의 부영상으로부터 독립적으로 통계적 모멘트를 추출한다.

먼저 변형된 LSB 스테고 영상의 8비트 픽셀이 그림 3과 같다고 하자.

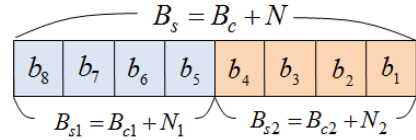


그림 3. 8비트 픽셀의 비트 분해

Fig. 3. bit decomposition of 8-bit pixel

여기서 B_s 는 스테고 영상의 픽셀이며, B_c 는 커버 영상의 픽셀, N 은 잡음 신호를 나타내며 식 (4)와 같이 표현할 수 있다.

$$B_s = B_c + N \quad (4)$$

이때 8비트 픽셀을 상위 4비트 B_{s1} 및 하위 4비트 B_{s2} 로 분해하면 식 (5) 및 (6)과 같다.

$$B_{s1} = B_{c1} + N_1 = (B_c - B_{c2}) + (N - N_2) \quad (5)$$

$$B_{s2} = B_{c2} + N_2 = (B_c - B_{c1}) + (N - N_1) \quad (6)$$

여기서 N_1 , N_2 는 분해된 픽셀에 포함된 잡음 신호이며, N_1 , N_2 , N 은 i.i.d(independent and identically distribution)이다. 또한 $B_c - B_{c1}$, $B_c - B_{c2}$ 는 8비트 픽셀에서 하위 4비트 및 상위 4비트를 제거한 픽셀을 의미한다.

그리고 식 (5) 및 (6)으로부터 분해된 픽셀의 분산을 구하면 각각 식 (7) 및 (8)과 같다.

$$Var(N - N_2) = 2 Var(N) + Var(N_2) \quad (7)$$

$$Var(N - N_1) = 2 Var(N) + Var(N_1) \quad (8)$$

위의 식으로부터 상위 및 하위 비트로 분해된 영상의 분산은 스테고 잠음 신호의 분산보다 두 배 이상 큼을 알 수 있으며, 이는 비트 분해시 커버 및 스테고 영상의 통계적 특성의 차이가 더 커지게 됨을 의미한다. 따라서 본 논문에서는 그림 4와 같이 영상이 주어질 때 상위 4비트 및 하위 4비트 부영상으로 분해한 후, 각각에 대해 3-레벨 웨이블릿 변환을 수행하고 비트 분해된 부영상 및 부대역의 특성 함수로부터 통계적 모멘트를 추출한다.

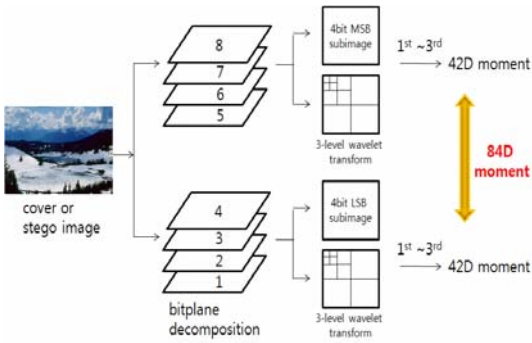


그림 4. 제안된 steganalysis 방법
Fig. 4. Proposed steganalysis method

이때 사용된 통계적 모멘트는 절대값 모멘트 (absolute moment)이며, 식 (9)와 같이 계산된다.

$$M_n^A = \frac{\sum_{k=0}^{M_p/2-1} k^n |H(k)|}{\sum_{k=0}^{M_p/2-1} |H(k)|}, \quad n = 1, 2, 3 \quad (9)$$

여기서 n 은 모멘트의 차수이며, M_p 는 히스토그램의 특성 함수 $H(k)$ 의 길이를 의미한다. 또한 $|H(k)|$ 은 특성 함수의 크기를 나타낸다.

본 논문에서는 분해된 부 영상 및 12개 부밴드의 특성 함수로부터 첫 3차 통계적 모멘트를 계산하여 총 84

차 특징 벡터를 추출한다. 추출된 특징 벡터는 분류기에 사용하기 전 식 (10)과 같이 스케일링 과정을 수행한다[11].

$$\tilde{m} = \frac{m - m_{\min}}{m_{\max} - m_{\min}} \quad (10)$$

여기서 모든 특징 벡터 m 에 대해 m_{\max} 는 최대값, m_{\min} 은 최소값을 나타낸다. 이러한 스케일링 과정은 모든 학습 및 테스트 영상의 특징 벡터를 0과 1 사이에 존재하게 함으로써 넓은 범위의 값을 가지는 특징 벡터들을 고르게 분산되도록 한다.

IV. 실험 결과

본 논문에서는 제안 방법의 성능을 평가하기 위하여 코렐드로우 영상 데이터베이스를 사용한다. 사용된 영상은 1,000개의 512×768 또는 768×512 의 압축되지 않은 bmp 파일이며, 본 논문에서는 변형된 LSB 기법을 사용하여 0.1bpp, 0.3bpp, 0.5bpp의 삽입률을 갖는 다양한 스테고 영상을 만들었다. 그중 800개의 영상은 MLP 분류기의 학습을 위해 사용되며, 200개는 테스트를 위해 사용한다.

표 1은 각 비트율의 스테고 영상에 대해 기존의 steganalysis 방법을 비교 실험한 결과를 보인다.

이때 메시지의 삽입 유무의 검출율은 10번 반복 실험한 후 그 평균값으로 하였다. 여기서 TP(true positive)는 스테고 영상에 삽입 정보가 있다고 옳게 판단하는 경우이며, TN(true negative)는 커버 영상을 커버 영상으로 옳게 판단하는 경우이다. 또한 ave.는 TP와 TN의 평균값을 나타낸다.

실험 결과 Xuan의 방법은 평균 55%, Mehrabi의 경우는 평균 61%의 검출율을 보이나 본 논문에서 제안한 방법은 평균적으로 약 76%의 보다 높은 검

출율을 보임을 알 수 있다. 특히 **Mehrabi**의 방법은 **TP**가 **TN**보다 훨씬 높기 때문에 전체적으로 볼 때 검출 성능이 좋다고 볼 수 없다.

또한 표 2는 200개의 테스트 영상에 대한 실험 성능을 보인 것이다. 여기서 **Precision**은 스테고 영상이라고 판단한 것 중에서 실제 스테고 영상일 확률을 나타내며, **Recall**은 실제 스테고 영상 중 얼마나 많은 스테고 영상을 찾아내느냐 하는 것이다. 또한 **F-measure**는 식 (11)과 같이 **Precision**과 **Recall**의 조화 평균으로서 값이 클수록 제대로 분류할 확률이 큼을 의미한다[12].

고 분류 성능이 우수함을 알 수 있다.

또한 그림 5는 각 비트율에서 **ROC(Receiver Operating Characteristic)** 곡선을 보여준다. **ROC** 곡선은 검출 임계치가 변화할 때 검출율(detection probability)과 오탐지율(false alarm probability)이 어떻게 변화하는가를 보여주는 그래프이다. 검출율은 스테고 영상이 정확히 스테고 영상으로써 분류될 확률을 말하며, 오탐지율은 커버 영상이 스테고 영상으로 분류될 확률을 말한다.

그림에서 각 **steganalysis** 방법에 기재된 수치는 **ROC**의 면적을 나타내는 것으로서 0.5에서 1의 값을 갖는 것이 좋다.

표 1. 기존의 방법과의 검출율 비교(단위 : %)

Table 1. Comparison of detection rate in the previous and proposed methods(units : %)

비트율(bpp)	Xuan's[6]			Mehrabi's[7]			Proposed method		
	TP	TN	ave.	TP	TN	ave.	TP	TN	ave.
0.1	49.5	51.5	50.5	91.5	33.0	62.3	88.5	63.0	75.8
0.3	57.5	53.0	55.3	94.5	26.0	60.3	89.0	62.0	75.5
0.5	52.5	62.5	57.5	85.5	38.0	61.8	86.0	68.0	77.0

표 2. 기존의 방법과의 F-measure 비교

Table 2. Comparison of F-measure in the previous and proposed methods

비트율(bpp)	Xuan's[6]			Mehrabi's[7]			Proposed method		
	Precision	Recall	F-measure	Precision	Recall	F-measure	Precision	Recall	F-measure
0.1	0.566	0.515	0.399	0.686	0.623	0.587	0.775	0.758	0.753
0.3	0.516	0.505	0.402	0.693	0.603	0.55	0.775	0.755	0.75
0.5	0.573	0.565	0.552	0.726	0.716	0.713	0.763	0.75	0.747

$$F-measure = \frac{2 * Precision * Recall}{Precision + Recall} \tag{11}$$

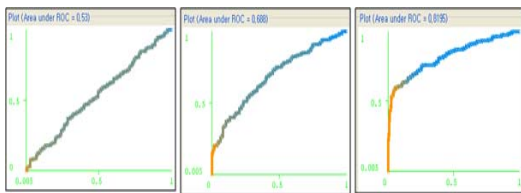
실험 결과 제안된 기법의 **F-measure**가 가장 크

본 논문에서는 식 (12)와 같이 사다리꼴 면적법을 이용하여 **ROC**의 면적을 계산하였으며, 표 3은 기존의 방법과의 **ROC** 면적을 비교한 것이다.

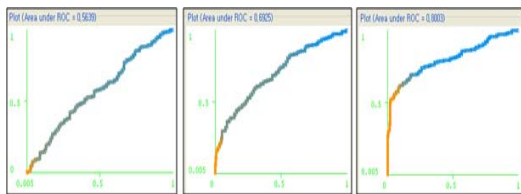
$$area = \sum_{i=2}^N (FAP_i - FAP_{i-1}) * (DP_i + DP_{i-1}) / 2 \quad (12)$$

여기서 FAP 는 오탐지율, DP 는 검출율을 의미한다.

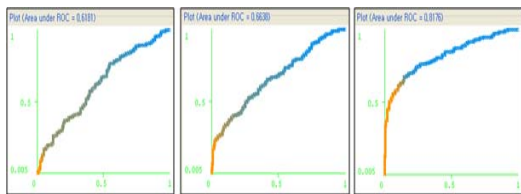
ROC 면적으로부터 본 논문에서 제안한 기법은 다른 방법에 비해 모든 비트율에서 평균 0.82에 가까운 면적을 가짐을 알 수 있다. 이는 제안한 기법에서 검출 결과의 신뢰도가 높음을 의미한다.



(a) 0.1bpp



(b) 0.3bpp



(c) 0.5bpp

그림 5. 각 비트율에서 ROC 곡선의 비교

(왼쪽:Xuan's, 중간:Mehrabi's, 오른쪽:Proposed method)

Fig. 5. Comparison of ROC curve in each bit rates

(left:Xuan's, middle:Mehrabi's, right:Proposed method)

표 3. 기존의 방법과의 ROC 면적 비교

Table 3. ROC area of previous and proposed methods

비트율 (bpp)	Xuan's [6]	Mehrabi's [7]	Proposed method
0.1	0.53	0.51	0.82
0.3	0.53	0.49	0.80
0.5	0.62	0.49	0.82

V. 결 론

컴퓨터 및 네트워크 환경이 발달함에 따라 불법 음란물이나 기업의 핵심 기술 유출, 그리고 테러 지령 전달 등과 같은 불법 정보의 유출 방법이 더욱더 다양화되며 지능화되고 있다. 따라서 비밀 메시지의 존재를 감지하고 유출을 방지할 수 있는 steganalysis에 관한 기술 개발이 보다 더 시급하다.

본 논문에서는 이러한 비밀 메시지의 검출율을 향상시키기 위한 개선된 steganalysis 기법을 제안하였다. 제안된 방법은 커버 영상과 스테고 영상의 분류 효과를 최대화하기 위해 영상의 각 픽셀을 상위 4비트와 하위 4비트로 분해하고, 비트 분해된 독립된 두 개의 영상에 대해 웨이블릿 변환을 수행한 후, 분해 영상 및 각 부대역의 계수값으로부터 84차의 통계적 모멘트를 추출하였다. 추출된 특징 벡터는 MLP 분류기에 입력되어 커버 영상과 스테고 영상을 학습하고 판단하는데 사용되었다. 다양한 삽입률의 변형된 LSB 임베딩 영상에 대해 제안된 기법을 실험한 결과 기존의 기법에 비해 비밀 메시지의 삽입 유무에 대한 검출율을 향상시킬 수 있음을 확인할 수 있었다. 향후 다양한 임베딩 방법에 의한 스테고 영상에 적용할 수 있는 방법과 검출율을 보다 더 향상시킬 수 있는 새로운 특징 벡터의 추출에 관한 연구가 제안된다.

참고문헌

[1] "<http://en.wikipedia.org/wiki/Steganography>"

[2] T. Sharp, "An implementation of key-based digital signal steganography", in: I. S. Moskowitz(eds.): 4th International Workshop on Information Hiding, LNCS 2137, Springer-Verlag, New York, pp. 13-26, 2001

[3] N. Provos, and P. Honeyman, "Hide and Seek : An Introduction to Steganography", IEEE Security & Privacy, 1(3): 32-44, maio/junho, 2003

[4] M. Kharrazi, H. T. Sencar, and N. Memon, "Performance study of common image steganography and steganalysis techniques", Journal of Electronic Imaging, vol. 15(4), 041104, 2006

[5] H. Farid, and S. Lyu, "Steganalysis using higher-order image statistic", IEEE Trans. on Info. Forensic and Security, vol. 1(1), pp. 111-119, 2006

[6] G. Xuan, Y.Q. Shi, and J. Gao, D. Zou, C. Yang, Z. Zhang, P. Chai, C. Chen and W. Chen, "Steganalysis Based on Multiple Features Formed by Statistical Moments of Wavelet Characteristic Functions", in Proceedings of Information Hiding Workshop, Baecelona, Spain, pp. 262-277, Jun, 2005

[7] M.A. Mehrabi, H. Aghaeinia, and M. Abolghasemi, "Image Steganalysis Based on Statistical Moments of Wavelet Subband Histogram of Images with Least Significant Bit Planes", Congress on Images and Signal Processing, IEEE, 2008

[8] G. Cancelli, G. Doërr, and M. Barni, I. Cox, "Detection of ± 1 LSB Steganography Based on the Amplitude of Histogram Local Extrema", in Proceedings of the IEEE International Conference of Image Processing, pp. 1288-1291, 2008

[9] M. Abolghasemi, H. Aghainia, and K. Faez, M. A. Mehrabi, "Steganalysis of LSB Matching Based on Co-Occurrence Matrix and Removing Most Significant Bit Planes", International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 1527-1530, 2008

[10] S. Torres, M. Nakano, and H. Perez, "An image steganography system based on BPCS and IWT", 16th

International Conference on Electronics, Communications and Computers, pp. 51-56, 2006

[11] Y. Sun, F. Liu, and B. Liu, P.Wang, "Steganalysis Based on Difference Image", IWDW, LNCS, 2008

[12] "http://en.wikipedia.org/wiki/Information_retrieval"



박태희(Tae-Hee Park)

1996년 부경대학교 일반대학원
전자공학과 공학석사
1999년 부산대학교 일반대학원
전자공학과 박사수료

2006년~현재 동명대학교 메카트로닉스공학과 전임강사
※ 관심분야: 비디오 영상처리, Steganalysis