

논문 2010-47SP-6-7

웨이블릿 부대역의 히스토그램 특성과 통계적 모멘트를 이용한 스테그분석

(Steganalysis Using Histogram Characteristic and Statistical Moments of Wavelet Subbands)

현 승 화*, 박 태 회****, 김 영 인***, 김 유 신**, 엄 일 규**

(Seung Hwa Hyun, Tae Hee Park, Young In Kim, Yoo Shin Kim, and Il Kyu Eom)

요 약

본 논문은 스테가노그래피 알고리즘에 대한 블라인드 스테그분석 기법을 제안한다. 제안하는 스테그분석기법은 두 가지 형태의 특징 벡터를 추출한다. 첫 번째로, 영상에 정보를 은닉한 후 웨이블릿 부대역의 히스토그램 특성이 변한다는 것을 관찰하고 히스토그램의 위치 변화를 특징으로 이용한다. 두 번째로, 웨이블릿 특성 함수의 통계적 모멘트를 특징으로 이용한다. 첫 번째 형태의 특징은 영상을 3-레벨 웨이블릿 변환하여 9개의 고주파 부대역에서 각각 하나의 특징을 추출하여 총 9개의 특징 벡터를 얻는다. 두 번째 형태의 특징은 각 부대역별로 3차 모멘트까지 추출하여 39개의 특징 벡터를 얻는다. 총 48개의 특징 벡터를 교차학습을 이용하여 학습한 후 스테고 영상과 커버 영상을 분류한다. 다층 퍼셉트론 신경망 분류기를 이용하여 두 가지 형태의 특징을 입력으로 하여 삽입 데이터의 존재유무를 판별한다. 제안 방법의 성능을 평가하기 위하여 CorelDraw 데이터 베이스 영상이 사용되었고 LSB 방법과 SS방법, blind SS방법, F5방법으로 다양한 삽입률의 스테고 영상을 생성하여 실험한다. 민감도와 특이도, 에러율, ROC 커브 면적 등을 이용하여 제안 방법이 기존의 스테그분석 방법보다 삽입 정보 유무를 검출하는데 효과적임을 보여준다.

Abstract

In this paper, we present a universal steganalysis scheme. The proposed method extract features of two types. First feature set is extracted from histogram characteristic of the wavelet subbands. Second feature set is determined by statistical moments of wavelet characteristic functions. 3-level wavelet decomposition is performed for stego image and cover image using the Haar wavelet basis. We extract one features from 9 high frequency subbands of 12 subbands. The number of second features is 39. We use total 48 features for steganalysis. Multi layer perceptron(MLP) is applied as classifier to distinguish between cover images and stego images. To evaluate the proposed steganalysis method, we use the CorelDraw image database. We test the performance of our proposed steganalysis method over LSB method, spread spectrum data hiding method, blind spread spectrum data hiding method and F5 data hiding method. The proposed method outperforms the previous methods in sensitivity, specificity, error rate and area under ROC curve, etc.

Keywords : steganalysis, wavelet subband, probability density function, characteristic function, statistical moment

I. 서 론

스테가노그래피(steganography)는 통신의 존재를 숨기면서 통신하는 기술이라고 할 수 있는데, 어떤 송신자와 수신자 사이에 통신이 일어나고 있다는 사실을 숨기는 것이다. 스테가노그래피라는 용어는 고대 그리스로부터 유래된 덮어진 쓰기(covered writing)라는

* 학생회원, ** 정회원, 부산대학교 전자공학과
(School of Electronics Eng., Pusan National University)

*** 정회원, 부산대학교 바이오메디컬공학과
(Dept. Biomedical Eng., Pusan National University)

**** 정회원, 동명대학교 메카트로닉스공학과
(Dept. Mechatronics Eng., TongMyong University)

접수일자: 2010년4월27일, 수정완료일: 2010년7월14일

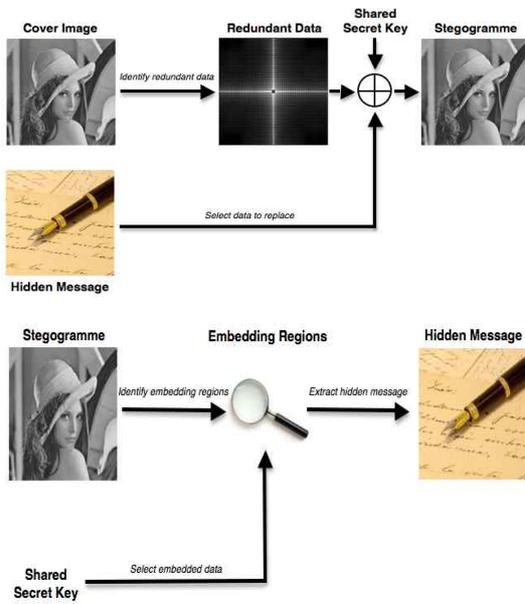


그림 1. 스테가노그래피의 전형적인 시나리오.
Fig. 1. Typical scenario of steganography.

의미로 다른 정보 안에 정보를 숨기는 것을 말한다. 특히, 스테가노그래피는 인간의 시각뿐만 아니라 컴퓨터 분석으로도 감지할 수 없도록 데이터를 숨겨야 하는 것을 목표로 하고 있다.

스테가노그래피는 군사적, 국가적 기밀 통신, 회사의 영업비밀의 보호 등의 긍정적인 목적으로 사용될 수 있으나, 이메일을 통한 데이터의 유출, 안티 바이러스 프로그램의 방해, 테러를 위한 비밀 통신 등의 부정적이고 위협적인 목적으로 사용될 수 있다. 특히, 미국의 911 테러에 스테가노그래피를 이용한 비밀 통신이 사용되어 그 위협성이 세계적으로 널리 알려진 바 있다. 그림 1은 전형적인 스테가노그래피의 시나리오를 나타내고 있다.

스테가노그래피의 목적이 어떤 메시지 정보의 은닉 자체를 숨기는 것이라면, 스테그분석(steganalysis)은 비밀 메시지의 은닉 여부를 판별하는 기술이다. 만약 비밀 메시지가 숨겨져 있다면, 그 메시지를 판독하거나 비밀 통신 자체를 파괴할 수 있다. 그러나 비밀 메시지를 판독하는 것은 상당히 어렵기 때문에 비밀 통신 자체를 무력화시키는 방법이 일반적으로 사용된다. 현재 영상, 오디오, 비디오 등이 파일의 형태로 개방형 네트워크를 통해 범람하고 있기 때문에, 이러한 디지털 콘텐츠에 대한 테러 정보, 국가 및 회사의 기밀 정보의 포함여부 및 검출, 복원 기술을 개발하여 유출 여부를 판단하고 결과적으로 불법적인 정보의 유출을 방지할

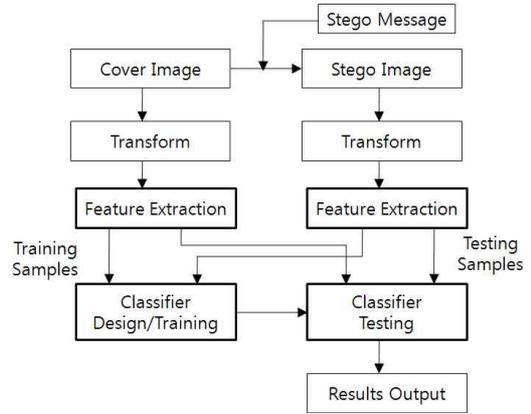


그림 2. 교사학습에 기반한 스테그분석 시스템의 전형적인 모델.

Fig. 2. A typical steganalysis model based on supervised learning.

수 있는 스테그분석에 대한 기술 개발이 시급하다고 하겠다.

스테그분석의 기본 개념은 커버 영상과 스테고 영상 간에 차이가 존재한다는 것이다. 일반적으로 자연 영상은 연속적이며, 인접한 화소 간에 상관관계가 강하게 나타난다. 그러나 정보가 은닉된 스테고 영상은 이러한 연속성을 변화시키거나 인접 화소간, 그리고 영상 블록간의 상관도가 줄어들게 된다. 따라서 스테그분석은 커버 영상과 스테고 영상 사이의 통계적 특성의 차이를 크게 하는 특징을 추출하는 것이 가장 중요하며, 이러한 특징을 추출하기 위한 다양한 연구가 현재까지 계속해서 이루어지고 있다.

본 논문에서는 스테그분석을 위해 두 가지 형태의 특징 벡터를 추출한다. 첫 번째로, 웨이블릿 부대역의 히스토그램 특성이 영상에 정보를 은닉한 후 변한다는 것을 관찰하고 히스토그램의 위치 변화를 특징으로 이용한다. 두 번째로, 웨이블릿 특성 함수의 통계적 모멘트를 특징^[5]으로 이용한다. 첫 번째 형태의 특징은 영상을 3-레벨 웨이블릿 변환하여 9개의 고주파 부대역에서 각각 하나의 특징을 추출하여 총 9개의 특징 벡터 얻는다. 두 번째 형태의 특징은 각 부대역별로 3차 모멘트까지 추출하여 39개의 특징 벡터를 얻는다. 총 48개의 특징 벡터를 교사학습을 이용하여 학습한 후 스테고 영상과 커버 영상을 분류한다. 본 논문의 방법은 히스토그램의 위치 변화를 특징 벡터로 사용하기 때문에 기존의 연구와 차별성을 가지며, 성능 또한 우수한 것으로 나타났다.

본 논문의 구성은 다음과 같다. II장에서는 교사학습

에 기반한 스테그분석 기법을 살펴보고, III장에서 본 논문에서 사용한 두 가지의 특징 벡터를 설명한다. 즉, 웨이블릿 부대역의 히스토그램 특성 변화를 분석하여 이를 나타낼 수 있는 특징을 추출하고 웨이블릿 특성 함수의 통계적 모멘트를 스테그분석을 위한 특징 벡터로 추출하는 방법에 대해 설명한다. IV장에서는 제안 방법과 이전 방법을 실험하여 비교하고 V장에서 결론을 맺는다.

II. 교사학습에 기반한 스테그분석 방법

현재의 스테그분석 기술은 2002년 H. Farid의 연구^[1]를 시작으로 알려지지 않은 스테고 시스템과 영상의 통계적 특성에 대한 어려움을 극복하기 위하여 교사학습(supervised learning)에 기반한 방법들이 주류를 이루고 있다. 아울러 이에 대한 많은 발전적인 방법들이 개발 중에 있다. 그림 2는 현재 가장 주류를 이루고 있는 스테그분석 시스템을 나타낸다. 그림 2에서 볼 수 있는 바와 같이 스테그분석 시스템은 영상의 변환, 특징 추출, 분류기로 이루어지는데, 현재 가장 중요한 이슈는 특징을 어떻게 추출할 것인가에 모아져 있다.

스테그분석에 사용되는 특징은 주로 영상의 화소 또는 변환 영역의 계수들의 히스토그램으로부터 추출된다.^[1~6] 특히, Harmsen과 Pearlman^[4]은 히스토그램 특성 함수의 무게 중심(Histogram characteristic function center of mass: HCF COM)을 스테그분석을 위한 특징으로 하여 스테그분석 방법을 제안하였다. 그리고 히스토그램 특성 함수의 무게 중심을 이용한 방법을 웨이블릿 영역으로 확장하여 웨이블릿 계수의 특성 함수의 통계적 모멘트를 특징으로 하는 방법^[5]이 Xuan 등에 의해 제안되었다. 또한 확률 밀도 함수(Probability density function)의 모멘트보다 특성 함수(Characteristic function)의 모멘트가 임베딩 데이터에 더욱 민감한 특징임을 [6]에서 Wang 등이 증명하였다.

스테그분석을 위한 특징 중 모멘트를 제외한 특징으로는 이웃 화소들 간의 마코프 모델 또는 계수의 동시 발생 행렬(co-occurrence matrix)을 이용한 방법이 존재한다^[7~9]. Ker^[10]는 동시 발생 행렬을 통계적으로 분석하여 히스토그램 특성 함수의 무게 중심 특징을 제안했다. [11~12]는 히스토그램 특성 함수의 무게 중심을 이용한 방법을 확장하여 특징 벡터를 구하여 스테그분석에 사용하였다.

본 논문에서는 웨이블릿 부대역의 특성함수 모멘트를 기본적으로 사용하고, 아울러 은닉된 데이터 때문에 변화된 히스토그램을 특징을 사용한 스테그분석 방법을 제안하고자 한다.

III. 제안 방법

커버 영상에 어떤 정보를 은닉하게 되면 커버 영상의 통계적 특성이 변하게 된다. 삽입 정보에 따른 변화가 민감한 특징 일수록 은닉 정보에 대한 검출률이 높아지게 된다. 본 논문에서는 커버 영상과 스테고 영상의 웨이블릿 부대역의 히스토그램의 변화를 관찰하고 이를 잘 나타낼 수 있는 특징을 선택하고자 한다. 또한 웨이블릿 부대역의 특성 함수의 모멘트를 특징 벡터로 추출하여 두 가지 형태의 특징을 분류기의 입력으로 하는 스테그분석 방법을 제안한다.

3.1 히스토그램 특성을 이용한 특징 추출

커버 영상과 스테고 영상을 Haar 웨이블릿 변환하면 4개의 부대역이 생기게 된다. 그림 3은 커버 영상과 스

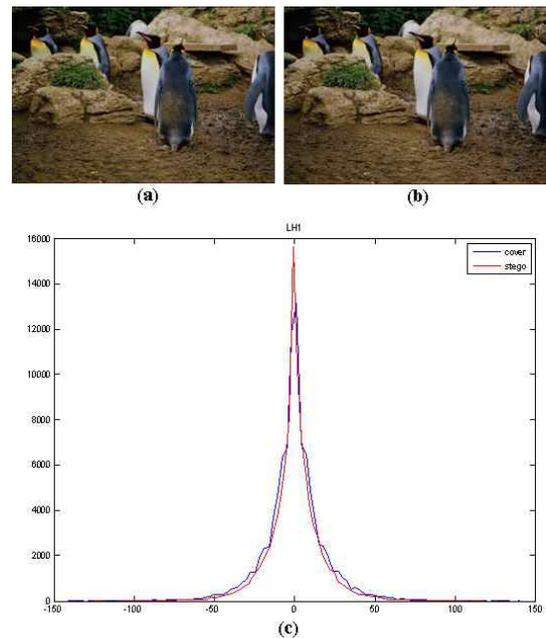


그림 3. 커버 영상과 스테고 영상의 1-레벨 LH 부대역의 히스토그램 비교 (a) 커버 영상 (b)스테고 영상 (c) 1-레벨 LH 부대역의 히스토그램

Fig. 3. The comparison of histogram of 1-level subbands with cover image and stego image (a) cover image (b) stego image (c) histogram of 1-level LH subband.

테고 영상에 대한 웨이블릿 부대역(1-레벨 LH 부대역)의 히스토그램을 각각 나타내고 있다. 그림 3(a)는 CorelDraw 영상 데이터베이스 중 하나를 나타내고, 3(b)는 F5 방법으로 정보를 은닉한 스테고 영상을 나타낸다. 그림 3(c)는 그림 3(a)와 3(b)를 Haar 웨이블릿 변환하여 1-레벨 LH 부대역 계수의 히스토그램을 나타내었다. 3(a)와 (b)를 시각적으로는 구분하기 어렵고 정보가 은닉되어있다는 사실 또한 알 수 없다. 하지만 그림 3(c)를 보면 스테고 영상의 계수들은 커버 영상의 계수보다 0에서 높은 빈도수를 가지고 큰 계수값에서는 커버 영상보다 적은 빈도수를 가진다는 것을 알 수 있다. 정보 은닉으로 인하여 웨이블릿 부대역의 히스토그램 분포가 변화하였음을 그림 3을 통해 알 수 있다.

본 논문에서는 정보가 은닉된 스테고 영상의 히스토그램의 변화에 주목한다. 즉, 히스토그램이 확률밀도함수를 표현한다고 가정하면, 면적 K 를 가지는 확률밀도함수의 값이 다를 것이다. 먼저 커버 영상과 스테고 영상이 동일한 확률밀도 함수 $f_X(x)$ 를 가지고, 다른 통계 파라미터를 가진다고 가정한다. 그리고 $f_X(x)$ 의 면적 K 는 다음과 같이 구해진다.

$$\int_0^x f_X(x)dx = K. \quad (1)$$

식 (1)에서 같은 면적 K 를 가지는 커버 영상과 스테고 영상의 x' 값이 달라질 것이다. 즉,

$$c_x' \neq s_x', \quad (2)$$

여기서 c_x' 은 면적 K 를 가지는 커버 영상의 x' 이고 s_x' 은 면적 K 를 가지는 스테고 영상의 x' 값을 나타낸다.

웨이블릿 부대역의 계수값은 다음과 같이 라플라시안 분포를 따른다고 모델링할 수 있다. 라플라시안 확률밀도함수^[6]는 다음과 같이 나타낼 수 있다.

$$f_X(x) = \frac{\alpha}{2} \exp(-\alpha|x|), \quad (3)$$

여기서 α 는 라플라시안 확률밀도 함수의 통계 파라미터이다. 확률밀도함수가 결정되면 다음과 같이 면적 K 를 가지는 확률 밀도 함수 상의 값 x' 을 구할 수 있다.

$$\int_0^x \frac{\alpha}{2} \exp(-\alpha x) dx = K. \quad (4)$$

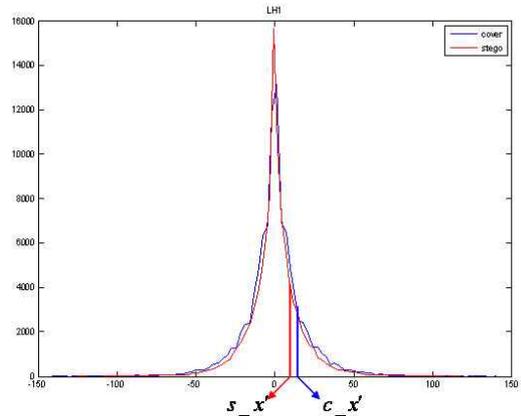


그림 4. 1-레벨 LH 부대역의 히스토그램에서의 특징 추출의 예

Fig. 4. An example of feature extraction by histogram of 1-level LH subband.

커버 영상과 스테고 영상의 히스토그램 분포가 다르므로 K 가 되는 x' 또한 다른 값을 가진다. 따라서 같은 면적 K 를 가지는 서로 다른 x' 을 정보 은닉 유무를 판단하는 특징으로 사용할 수 있고, 그림 4에서 그 예를 볼 수 있다.

x' 값을 구하기 위하여 식 (4)를 풀면 다음과 같다.

$$x' = -\frac{\log(1-2K)}{\alpha}. \quad (5)$$

즉, 면적 K 를 가지는 x' 값은 라플라시안 확률밀도함수의 통계 파라미터 α 의 함수로 나타난다. K 의 값은 실험적으로 구하였고 본 논문에서는 0.1의 값으로 실험하였다. 또한 α 는 다음과 같이 분산의 함수로 나타낸다.

$$\text{VAR}[X] = \frac{2}{\alpha^2} = \sigma^2. \quad (6)$$

따라서 식 (5)와 (6)을 정리하면, 다음과 같이 x' 을 구할 수 있다.

$$x' = -\frac{\sigma \log(1-2K)}{\sqrt{2}}. \quad (7)$$

즉, x' 는 표준편차의 함수가 된다.

위의 과정을 종합해 보면, 정보가 은닉된 스테고 영상은 커버 영상과 같은 확률밀도함수를 가지지만 다른 통계파라미터를 가지고 있다. 즉, 서로 다른 표준편차를 가지게 된다. 따라서 서로 다른 x' 을 가지기 때문에 커버 영상과 스테고 영상을 분리할 수 있는 특징이 될 수 있다.

본 논문에서는 영상을 3-레벨 웨이블릿 변화하여 10개의 부대역 중 LL밴드를 제외한 고주파 부대역 9개로부터 각 밴드별로 1개씩의 특징을 추출한다. 스테고 영상과 커버 영상의 히스토그램을 0에서부터 적분하여 같은 면적값 K 를 가지는 랜덤 변수의 위치 x 를 찾아 특징 벡터로 사용한다. 9개의 부대역에서 하나의 특징을 추출하므로 총 9개의 특징을 얻을 수 있다.

3.2 웨이블릿 특성 함수를 이용한 특징 추출^[5]

앞서 제안한 9개의 특징과 아울러 스테그분석을 위한 특징을 추출하기 위해서 영상을 3-레벨 웨이블릿 변환하여 13개의 부대역(LL1, HL1, LH1, HH1, LL2, HL2, HH2, LL3, HL3, LH3, HH3)을 이용한다. 즉, 13개 부대역의 특성 함수의 1차, 2차, 3차 모멘트를 특징으로 추출한다. 특성 함수의 n 차 통계적 모멘트 M_n 은 다음 식을 이용해서 구할 수 있다.

$$M_n = \left(\sum_{k=0}^{N/2} f_k^n |H(f_k)| \right) / \left(\sum_{k=0}^{N/2} |H(f_k)| \right), \quad (8)$$

여기서 $|H(f_k)|$ 은 특성 함수의 크기이고, 특성 함수는 히스토그램을 푸리에 변환한 것이다. 13개의 부대역에서 1차, 2차, 3차 모멘트를 추출하므로 39개의 특징을 얻는다. 앞에서 구한 히스토그램 특성을 이용한 특징 추출에서 9개의 특징을 추출하였으므로, 본 논문에서는 총 48개의 특징을 스테그분석에 이용한다.

IV. 모의실험 및 결과

스테그분석을 위한 제안 방법의 성능을 평가하기 위하여 CorelDraw 영상 데이터베이스^[13]를 사용한다. 데이터베이스는 1096개의 다양한 영상으로 구성되어있다. CorelDraw 영상 1096개 중 800개의 영상은 분류기의 학습을 위해 사용하고 200개는 테스트를 위해 사용한다. 검출률은 10번 실험하고 그 평균값으로 하였다. 실험을 위한 첫 번째 단계는 데이터를 삽입하여 스테고 영상을 만드는 것이다. 데이터 삽입은 LSB 방법과 SS 방법, blind SS 방법, F5방법^[14]을 사용하였다. 데이터 삽입률 각각 0.1bpp, 0.3bpp, 0.5bpp로 하여 LSB 방법, SS 방법, blind SS 방법으로 스테고 영상을 생성하였고, F5방법으로 80X80 영상을 커버 영상에 삽입하였다.

1096개의 CorelDraw 영상으로부터 각각의 방법으로 800개씩의 스테고 영상을 만든다. 800개의 스테고 영상

과 커버 영상으로부터 총 48개의 특징이 각각 추출된다. 이 특징들이 스테고 영상과 커버 영상의 분류를 위해서 신경망(Multi Layer Perceptrons : MLP)을 분류기로 사용하였다. 신경망은 웨카 툴을 이용하여 실험하였다.^[17]

본 논문의 방법과 비교를 위하여 차분 영상의 히스토그램 특성함수의 모멘트를 특징으로 하는 Sun^[15]의 방법과 웨이블릿 부대역의 히스토그램 특성함수의 모멘트를 특징으로 하는 Xuan^[5]의 방법에 대해서 각각 4개의 피쳐와 39개의 특징 벡터를 구하고 삽입 데이터에 대한 민감도를 객관적으로 비교하기 위하여 제안 방법과 같은 신경망 분류기를 사용하여 실험하였다. Xuan^[5]의 논문에서 사용한 베이지안 분류기를 웨카 툴을 이용하여 제안 방법과 Sun의 방법에 대해서 모두 실험하였다. 신경망 분류기를 사용한 것과 전체적인 경향은 같았지만 신경망보다 낮은 정확도를 나타내었다. 따라서 본 논문에서는 더 높은 정확도를 얻기 위하여 신경망 분류기를 사용하였다.

스테그분석의 성능 평가를 위하여 본 논문에서는 sensitivity 와 specificity, classification rate를 사용하였다. 즉,

$$Sensitivity = \frac{TP}{TP + FN}, \quad (9)$$

$$Specificity = \frac{TN}{TN + FP}, \quad (10)$$

$$classification\ rate = \frac{TP + TN}{TP + FP + TN + FN}, \quad (11)$$

여기서 TP(true positive)는 커버 영상에 삽입 정보가 없다고 판단하는 경우이고, TN(true negative)은 스테고 영상에 삽입 정보가 있다고 판단하는 경우이다. FP(false positive)는 스테고 영상을 커버 영상으로 판단하는 경우이고 FN(false negative)은 커버 영상을 스테고 영상으로 판단하는 경우이다.

표 1은 각 임베딩 방법별로 200개의 테스트 영상을 사용하여 구한 검출률을 나타내고 있다. 표 1에서 볼 수 있는 바와 같이 제안 방법은 Sun의 방법보다 여러 삽입 방법과 다양한 삽입률에 대해서 높은 검출률을 보여준다. Xuan의 방법과 비교할 때 SS 임베딩 방법에서만 낮은 검출률을 나타내고 나머지 삽입 방법에 대해서는 높은 검출률을 나타낸다. 특히 LSB 방법으로

표 1. 기존의 방법과의 검출률 비교(단위:%)

Table 1. Detection rate comparison of previous works and our proposed method for various embedding methods and embedding rate,

스테그 분석방법 삽입 방법 및 삽입률		Xuan's [5]			Sun's[15]			Proposed method		
		Sensitivity	Specificity	Classification rate	Sensitivity	Specificity	Classification rate	Sensitivity	Specificity	Classification rate
LSB	0.1	51.5	49.5	50.5	45.5	57.0	51.3	60.0	66.5	63.3
	0.3	53.0	57.5	55.3	53.0	49.5	51.3	44.5	60.5	52.5
	0.5	62.5	52.5	57.5	51.5	42.5	47.0	89.5	91.5	90.5
SS	0.1	90.5	81.5	86.0	52.5	56.5	54.5	76.5	84.5	80.5
	0.3	90.5	85.5	88.0	51.5	52.5	52.0	73.5	85.5	79.5
	0.5	89.0	86.5	87.8	61.0	53.0	57.0	77.5	89.5	83.5
blind SS	0.1	71.5	82.5	77.0	48.0	60.5	54.3	75.0	85.5	80.3
	0.3	71.8	81.3	76.4	52.0	65.5	58.8	83.5	88.5	86.0
	0.5	73.8	82.3	78.1	56.5	61.5	59.0	81.5	89.0	85.3
F5	0.13	95.8	96.0	95.9	53.5	47.5	50.5	98.5	99.5	99.0

표 2. 기존의 방법과의 에러율 비교(단위:%)

Table 2. Error rate comparison of previous works and our proposed method for various embedding methods and embedding rate.

스테그 분석방법 삽입 방법 및 삽입률		Xuan's [5]			Sun's[15]			Proposed method		
		FN rate	FP rate	Error rate	FN rate	FP rate	Error rate	FN rate	FP rate	Error rate
LSB	0.1	48.5	50.5	49.5	54.5	43.0	48.7	40.0	33.5	36.7
	0.3	47.0	42.5	44.7	47.0	50.5	48.7	55.5	39.5	47.5
	0.5	37.5	47.5	42.5	48.5	57.5	53.0	10.5	8.5	9.5
SS	0.1	9.5	18.5	14.0	47.5	43.5	45.5	23.5	15.5	19.5
	0.3	9.5	14.5	12.0	48.5	47.5	48.0	26.5	14.5	20.5
	0.5	11.0	13.5	12.2	39.0	47.0	43.0	22.5	10.5	16.5
blind SS	0.1	28.5	17.5	23.0	52.0	39.5	45.7	25.0	14.5	19.7
	0.3	28.2	18.7	23.6	48.0	34.5	41.2	16.5	11.5	14.0
	0.5	26.2	17.7	21.9	43.5	38.5	41.0	18.5	11.0	14.7
F5	0.13	4.2	4.0	4.1	46.5	52.5	49.5	1.5	0.5	1.0

0.5bpp의 삽입률로 데이터를 삽입 했을 때 Xuan 방법의 평균 검출률이 57.5%, Sun 방법의 평균 검출률이 47%인데 반해 제안 방법은 90.5%로 상당히 좋은 검출률을 가진다.

커버 영상을 스테고 영상으로 판단하는 FN도 중요하지만 스테고 영상을 아무것도 삽입되지 않은 커버 영상으로 판단하는 경우 심각한 문제가 발생할 수 있다. 우리의 목표는 커버 이미지에 삽입 정보가 있는지를 검출하여 스테고 이미지의 통신 자체를 무산시키는 것이다.

따라서 삽입정보가 국가 기밀 또는 군사 보안등에 관련된 민감한 정보라면 FP가 FN보다 더 중요한 요소라는 것을 알 수 있다. 본 논문에서는 다음과 같이 'FP rate', 'FN rate', 'Error rate'를 스테그분석의 성능 평가를 위한 추가적인 척도로 사용하였다.

$$FP \text{ rate} = \frac{FP}{FP + TN}, \tag{12}$$

$$FN \text{ rate} = \frac{FN}{FN + TP}, \tag{13}$$

$$Error\ rate = \frac{FP + FN}{FP + TN + FN + TP} \quad (14)$$

표 2는 기존 방법과 제안 방법의 'FP rate', 'FN rate', 'Error rate'를 비교하였다. 제안 방법과 Xuan의 방법을 비교해 볼 때 제안 방법이 SS 임베딩 방법에서 낮은 검출률을 나타내었지만 표 2에서 보여주는 'FP rate', 즉 스테고 영상을 커버 영상으로 잘못 판단하는 경우를 볼 때 제안 방법의 에러율이 더 낮은 것을 알 수 있다.

표 3은 제안 방법과 Xuan^[5] 방법의 TP rate와 FP rate의 차이 값을 나타낸다. ΔTP rate는 값이 클수록, ΔFP rate는 작은 값 일수록 제안 방법의 성능이 우수함을 나타낸다. 표 3을 보면 ΔTP rate는 LSB방법으로 0.3bpp의 삽입률로 삽입한 경우 Xuan 방법보다 8.5% 작게 나타났고 SS 방법으로 삽입한 경우는 제안 방법이 더 낮은 TP rate를 나타낸다. 하지만 ΔFP rate를 보면 Xuan^[5]이 제안한 방법보다 최대 39%까지 개선된 것을 볼 수 있다. 우리는 Xuan^[5]이 제안한 특징 벡터에 웨이블릿 부대역 히스토그램의 특성 변화를 나타내는 특징을 함께 분류기의 입력으로 사용하여 LSB(0.1bpp, 0.5bpp), BSS, F5 삽입 방법에 대해서 TP rate가 3.2%~26%의 성능이 향상됨을 볼 수 있었고, 4가지 삽입 방법 모두에 대해서 FP rate가 0%~39%의 성능이 향상됨을 알 수 있었다.

표 3. Xuan^[5]의 방법과 제안 방법의 TP rate와 FP rate의 차이 값 비교

Table 3. Performance comparison of ΔTP rate and ΔFP rate of Xuan's method and proposed method

삽입 방법 및 삽입률	Δ값	ΔTP rate	ΔFP rate
LSB	0.1	8.5	-17.0
	0.3	-8.5	-3.0
	0.5	26.0	-39.0
SS	0.1	-14.0	-3.0
	0.3	-17.0	0
	0.5	-11.5	-3.0
blind SS	0.1	3.5	-3.0
	0.3	11.7	-7.2
	0.5	7.7	-6.7
F5	0.13	3.2	-3.5

표 4. 기존의 방법과의 ROC 커브 면적 비교

Table 4. Performance comparison of area under ROC curve of previous works and our proposed method.

스테그 분석 방법 / 삽입 방법 및 삽입률		Xuan's [5]	Sun's [15]	Proposed method
LSB	0.1	0.53	0.50	0.65
	0.3	0.56	0.50	0.55
	0.5	0.62	0.50	0.99
SS	0.1	0.95	0.73	0.90
	0.3	0.95	0.78	0.91
	0.5	0.96	0.79	0.89
blind SS	0.1	0.86	0.48	0.85
	0.3	0.87	0.45	0.92
	0.5	0.87	0.61	0.92
F5	0.13	0.99	0.74	1.0

마지막으로 표 4는 ROC 커브^[16] 면적을 나타내고 있다. ROC (Receiver-Operating Characteristic) 커브는 스테그분석 방법의 민감도(sensitivity)와 특이도(specificity)가 컷 포인트에 따라서 어떤 관계를 갖고 있는지를 표현한 그래프이다. 민감도는 스테그 방법이 커버 영상을 커버 영상으로 얼마나 잘 분류하는가를 나타내고 특이도는 스테고 영상을 스테고 영상으로 얼마나 잘 분류하는가를 나타낸다. 그림 5에 나타나있는 ROC 커브 그래프는 x축은 1-specificity (false positive rate)를 나타내고 y축은 sensitivity (true positive rate)를 나타낸다. 스테그분석 방법의 성능을 평가하기 위한 것으로 커브 아래 면적을 이용한다. 그래프가 상단 좌측에 위치 할수록 유용한 분류방법이다. 완벽한 방법은 면적 1.0을 가지고 분류 능력이 없는 방법은 0.5의 면적을 보인다. 그림 5를 보면 기존의 방법과 제안 방법의 ROC 커브를 LSB방법으로 삽입률 0.5bpp로 삽입한 데이터에 대해서 나타내고 있다. 그림 5(a)는 제안 방법의 ROC 커브인데 앞에서 설명한 대로 좌측 상단에 위치한 것을 볼 수 있다. 이 경우 분류 방법의 성능이 좋다고 말할 수 있다. 그림 5(b)와 (c)는 Xuan과 Sun이 제안한 방법의 ROC 커브로 대각선에 가까운 모양임을 알 수 있다. 대각선 모양에 가까우면서 면적이 0.5에 가까우면 분류 방법이 좋은 성능을 가지지 못함을 말한다. 표 3을 보면 제안 방법은 LSB방법과 blind SS방법으로 데

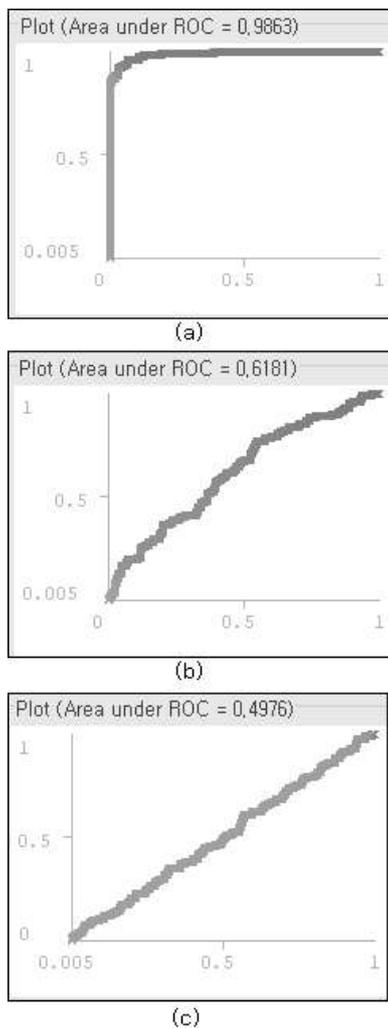


그림 5. ROC 커브 (a) 제안 방법 (b) Xuan의 방법 (c) Sun의 방법, (LBS 0.5bpp).

Fig. 5. ROC curve (a) Proposed method (b) Xuan's method (c) Sun's method, (LSB 0.5bpp).

이터를 삽입한 경우 다른 방법들보다 큰 ROC 커브 면적을 나타내고 SS방법에 대해서도 0.9정도의 면적을 나타내며 좋은 성능을 보여준다. F5 삽입 방법에 대해서는 1.0의 완벽한 ROC 커브 면적을 나타낸다. 이것은 표 1과 표 2에서 보여준 것과 같이 일관성 있는 결과를 나타내며 제안 방법의 성능에 대한 객관성을 높여준다.

IV. 결 론

본 논문에서는 스테그분석을 위하여 두 가지 형태의 특징을 사용하였다. 먼저, 스테고 영상의 웨이블릿 부대역의 히스토그램을 분석하여 특징 벡터를 구하고 두 번째로 웨이블릿 특성 함수의 통계적인 모멘트를 이용하여 특징 벡터를 구하였다. 또한 다층 퍼셉트론 신경망

분류기를 이용하여 스테고 영상과 커버 영상에서 추출된 두 가지 형태의 특징들을 함께 입력으로 하여 삽입 데이터의 존재유무를 판별하였다. 제안 방법의 성능을 평가하기 위하여 CorelDraw 영상 데이터베이스로부터 LSB 방법과 SS방법, blind SS방법, F5방법으로 다양한 삽입률의 스테고 영상을 생성하여 사용하였다. 실험 결과는 제안 방법은 삽입 데이터에 민감한 특징을 추출하여 삽입 정보 유무를 검출하는데 기존의 방법보다 효과적임을 알 수 있었다. 아울러 에러율 또한 비교적 우수하게 나타났다.

참 고 문 헌

- [1] H. Farid. "Detecting Hidden Messages Using Higher-order Statistical Models," in Proceedings of IEEE ICIP, pp.905-908, 2002.
- [2] T. Holotyak, J. Fridrich and S. Voloshynovskiy. "Blind Statistical Steganalysis of Additive Steganography Using Wavelet Higher Order Statistics," CMS 2005, LNCS, vol. 3677, pp.273-274. Springer, Heidelberg, 2005.
- [3] M. Goljan, J. Fridrich and T. Holotyak. "New Blind Steganalysis and its Implications," in Proceedings of Security, Steganography, and Watermarking of Multimedia Contents. SPIE, pp.1-13, 2006.
- [4] J. Harmsen and W. Pearlman. "Steganalysis of Additive Noise Modelable Information Hiding," in Proceedings of Security, Steganography, and Watermarking of Multimedia Contents VI. SPIE, pp.131-142, 2003.
- [5] G. Xuan, Y. Q. Shi, 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, Barcelona, Spain, pp. 262-277, Jun. 2005.
- [6] Y. Wang and P. Moulin. "Optimized Feature Extraction for Learning-based Image Steganalysis," IEEE Trans. Information Forensics and Security, vol. 2, no.1, pp.31-45, March, 2007.
- [7] K. Sullivan, U. Madhow, S. Chandrasekaran and B. Manjunath. "Steganalysis for Markov cover data with applications to images," IEEE Trans. Information Forensics and security, vol.1, no.2, pp.275-287, Jun. 2006.
- [8] G. Xuan, X. Cui, Y. Shi., et al. "JPEG

Steganalysis Based on Classwise Non-principal Components Analysis and Multi-directional Markov Model,” in Proceedings of IEEE ICME, pp.903-906, 2007.

[9] Y. Shi, C. Chen and W. Chen. “A Markov Process Based Approach to Effective Attacking JPEG Steganography,” IH 2006, LNCS, vol. 4437, pp.249-264. Springer, Heidelberg, 2007.

[10] A. Ker. “Steganalysis of LSB Matching in Grayscale Images,” IEEE Signal processing letters, vol.12, no.6, pp.441-444, 2005.

[11] C. Chen, Y. Shi and W. Chen, et al. “Statistical Moments Based Universal Steganalysis Using JPEG 2-D Array and 2-D Characteristic Function,” in Proceedings of IEEE ICIP, pp.105-108, 2006.

[12] Q. Liu, A.H. Sung, Z. Chen and J. Xu. “Feature Mining and Pattern Classification for Steganalysis of LSB Matching Steganography in Grayscale Images,” Pattern Recognition 41, pp.56-66, 2008.

[13] CorelDraw Software, www.corel.com.

[14] A. Westfeld. “High Capacity Despite Better Steganalysis (F5-A Steganographic Algorithm),” IH 2001, LNCS, vol.2137, Springer, Heidelberg, pp.289-302, 2001.

[15] Y. Sun, F. Liu, B. Liu and P. Wang. “Steganalysis Based on Difference Image,” IWDW 2008, LNCS, vol.5450, pp.184-198, Springer, Heidelberg, 2009.

[16] A. Westfeld. “Roc Curves for Steganalysts,” in Proceedings of the Third WAVILA Challenge(Wacha’07), Saint Malo (France), pp. 39-45 2007.

[17] WEKA tool
<http://www.cs.waikato.ac.nz/ml/weka/>

— 저 자 소 개 —



현 승 화(학생회원)
2003년 밀양대학교 정보통신 공학과 학사 졸업.
2005년 부산대학교 전자공학과 석사 졸업.
2007년 부산대학교 전자공학과 박사과정 수료.

현재 부산대학교 전자공학과 박사과정.
<주관심분야 : 영상처리, 오류은닉, 웨이블릿 변환 신호처리, 스테그 분석>



박 태 희(정회원)
1993년 부경대학교 정보통신 공학과 학사 졸업.
1996년 부경대학교 전자공학과 석사 졸업.
1999년 부산대학교 전자공학과 박사과정 수료

현재 동명대학교 메카트로닉스공학과 전임강사
<주관심분야 : 영상처리, 영상압축, 스테그 분석>



김 영 인(정회원)
1996년 명지대학교 컴퓨터공학과 박사 졸업.
1990년~1991년 (주)한국컴퓨터 기술연구소(연구원)
1996년~2006년 밀양대학교 컴퓨터공학부 부교수

2007년~2008년 Univ. of Missouri-Columbia (Visiting Scholar)

현재 부산대학교 바이오메디컬공학과 교수
<주관심분야 : 데이터베이스, 데이터마이닝>



김 유 신(정회원)
1974년 서울대학교 전자공학과 학사 졸업.
1980년 U.C. Berkeley 전자공학과 석사 졸업.
1983년 Stanford University 전자재료 박사과정 수학.

1995년 Cornell University Ph.D
1983년~현재 부산대학교 전자공학과 교수
<주관심분야 : 영상처리, 신경회로망, 공학윤리, 과학기술철학>



엄 일 규(정회원)-교신저자
1990년 부산대학교 전자공학과 학사 졸업.
1992년 부산대학교 전자공학과 석사 졸업.
1998년 부산대학교 전자공학과 박사 졸업.

현재 부산대학교 전자공학과 교수
<주관심분야 : 통계신호처리, 영상처리, 웨이블릿 변환 신호처리, 신경회로망>