

논문 2011-48SP-3-8

웨이블렛 부밴드의 조인트 모멘트를 이용한 스테그분석

(Steganalysis Using Joint Moment of Wavelet Subbands)

박 태 희*, 현 승 화**, 김 재 호***, 엄 일 규***

(Tae Hee Park, Seung Hwa Hyun, Jae Ho Kim, and Il Kyu Eom)

요 약

본 논문은 웨이블릿 도메인 상에서 부모와 자식 부밴드간의 비독립성에 기반한 영상 스테그분석 방법을 제안한다. 제안한 방법은 커버 영상과 비밀 메시지가 삽입된 스테고 영상에 대해 3-레벨 Haar UWT 웨이블릿 변환을 수행하여 12개의 부밴드로 분해한 후 부모와 자식 부밴드간의 통계적 의존성을 분석한다. 이러한 통계적 의존성은 비밀 메시지가 삽입된 스테고 영상의 경우 커버 영상과 상당한 차이를 보이므로 커버 및 스테고 영상을 구분하기 위한 특징으로 사용될 수 있다. 따라서 본 논문에서는 분해된 12개의 각 부모와 자식 부밴드간의 조인트 특성 함수에 대해 첫 9차의 통계적 모멘트를 추출함으로써 총 72차의 통계적 조인트 모멘트를 특징 벡터로 사용한다. 추출된 특징 벡터는 MLP(다층 퍼셉트론 신경망) 분류기에 입력되어 커버 영상과 스테고 영상을 학습하고 판별한다. 제안 방법의 성능 평가를 위해 LSB 및 SS, BSS 삽입 방법에 의한 다양한 삽입률의 스테고 영상을 사용하였으며, 실험 결과 제안한 기법은 기존의 기법에 비해 삽입 정보 유무의 검출율을 향상시킬 뿐만 아니라 판별의 정확도가 높음을 확인할 수 있었다.

Abstract

This paper propose image steganalysis scheme based on independence between parent and child subband on the multi-layer wavelet domain. The proposed method decompose cover and stego images into 12 subbands by applying 3-level Haar UWT(Undecimated Wavelet Transform), analyze statistical independency between parent and child subband. Because this independency is appeared more difference in stego image than in cover image, we can use it as feature to differentiate between cover and stego image. Therefore we extract 72D features by calculation first 3 order statistical moments from joint characteristic function between parent and child subband. Multi-layer perceptron(MLP) is applied as classifier to discriminate between cover and stego image. We test the performance of proposed scheme over various embedding rates by the LSB, SS, BSS embedding method. The proposed scheme outperforms the previous schemes in detection rate to existence of hidden message as well as exactness of discrimination.

Keywords : steganalysis, undecimated wavelet transform, statistical moment, characteristic function, MLP

I. 서 론

스테가노그래피(steganography)란 그리스어인 “감춰

진 글”에서 유래된 것으로서 텍스트, 오디오, 비디오 파일과 같은 디지털 매체에 비밀 메시지를 은닉하고 통신 자체를 감추기 위한 기술을 말한다^[1~4]. 스테가노그래피의 원리는 비밀 정보를 삽입함으로써 디지털 매체는 약간 변할지라도 사람의 눈이나 귀에는 비밀 메시지의 삽입 전이나 후가 전혀 차이가 없고 전송시에도 일반 파일과 동일하게 보이도록 하는 것이다. 그림 1은 전형적인 스테가노그래피의 과정을 보인 것이다.

이때 영상에 메시지를 숨기기 위한 일반적인 임베딩 방법으로는 공간 영역에서 샘플 픽셀 값의 LSB(Least Significant Bit)를 비밀 메시지 비트로 치환하는 LSB^[5]

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

** 학생회원, *** 정회원, 부산대학교 전자전기공학부
(School of Electrical Eng., Pusan National University)

※ 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. 2010-0011809)

접수일자: 2011년2월21일, 수정완료일: 2011년4월7일

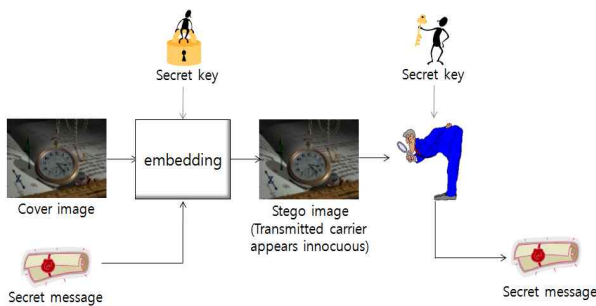


그림 1. 전형적인 스테가노그래피의 과정
Fig. 1. Typical steganography process.

방법과, 영상의 DCT(Discrete Cosine Transform) 영역에서 DC 계수를 제외한 첫 L 개의 중요 계수에 비밀 메시지를 삽입하는 SS(Spread Spectrum) 방법^[6], 그리고 시각적 아티팩트(artifact)를 줄이기 위해 첫 M 개의 계수는 버리고 다음의 L 개의 주파수 계수를 비밀 메시지로 치환하는 BSS(Blind SS)^[7] 방법 등이 있다.

이러한 스테가노그래피는 문서의 저작권을 보호하고 국가의 기밀을 보호하는 등 많은 응용에 있어서 유용하다. 그러나 인터넷을 통해 스테가노그래피 툴을 쉽게 구할 수 있고, 응용 분야도 풍부해 짐에 따라 반사회적 또는 범죄에 사용될 수 있다는 지적이 나왔다.

만일 이러한 기법이 테러 리스트에 의해 지령 정보를 전달하는데 사용되거나, 이메일을 통해 컴퓨터 바이러스 또는 트로이 목마와 같은 악성 프로그램을 보낸다면 컴퓨터를 크게 손상시킬 수 있다. 특히 회사에 불만을 품은 사원에 의해 회사의 기술 정보 파일을 외부로 유출하는데 사용된다면 기업의 경쟁력을 약화시킬 뿐만 아니라 많은 금전적 피해가 발생할 수 있다.

이러한 스테가노그래피의 모든 위협에 대응하는 가장 큰 방법은 스테그분석이다. 즉 스테그분석은 무결하게 보이는 신호로부터 숨겨진 정보의 존재를 감지하는 것이 목적이다^[8].

스테그분석의 기본 원리는 숨겨진 비밀 메시지가 사람의 눈에는 쉽게 감지되지 않으나 히스토그램의 분포 또는 인접 화소와 블록 간의 상관도 등 통계적인 분포에는 차이가 있다는 것이다. 따라서 메시지의 삽입 전후의 영상에서 통계적인 특징(feature)의 차를 발견하는 것은 매우 중요하며, 이러한 특징을 추출하기 위한 다양한 연구가 현재까지 계속되고 있다^[9].

Avcibas 등은 숨겨진 메시지의 존재 여부를 결정하기 위한 특징 벡터로서 연속된 비트 플레인 간의 이진 유사도 척도(binary similarity measure)를 이용하였다

^[10]. 이는 메시지의 삽입 전후 비트 플레인 간의 상관관계가 다를 것이라는 데 있으며, 특히 중요도가 낮은 7번째와 8번째 비트 플레인을 사용하여 10차의 특징을 추출하였다.

Lyu 등은 칼라 영상의 각 RGB 채널에 대해 4-레벨의 QMF(Quadrature Mirror Filters)를 이용하여 영상을 분해하고, 각 부밴드에 대해 총 72차의 특징 벡터를 추출하였다. 이때 4차의 통계적 모멘트(mean, variance, skewness, kurtosis)는 각 부밴드 및 최적화된 선형 예측기로부터 추정된 계수값과 계수의 로그값들의 오차에 대해 계산되었다^[11].

Xuan 등은 3-레벨 Haar 웨이블릿 변환을 사용하여 영상을 4개의 부밴드로 분해하고, 원 영상 및 각 부밴드 계수의 특성 함수에 대해 39차의 통계적 모멘트를 특징 벡터로 추출하였다^[12].

또한 Shu 등은 다양한 잡음의 특성을 반영하기 위해 Xuan 등이 제안한 특징에 예측 에러 영상에 대한 39차의 특징을 추가함으로써 총 78차의 특징 벡터를 사용하였다^[13]. 그러나 이 방법은 학습되지 않은 영상에 대해서는 성능이 저하되는 단점이 있다.

최근 Sun 등은 4 방향(수평, 수직, 왼쪽, 오른쪽 대각 방향)으로의 차분 영상(difference image)에 대해 총 4차의 특성 함수 모멘트를 특징 벡터로 추출하는 스테그 분석 방법을 제안하였다^[14].

본 논문에서는 웨이블릿 영역에서 부모와 자식 부밴드간의 계수들이 서로 독립이지 않다는 통계적 의존성을 이용하여 커버 영상과 스테고 영상의 조인트 특성 함수에 기반한 특징 벡터를 추출한다. 먼저 주어진 영상에 대해 3-레벨 undecimated Haar 웨이블릿 변환을 사용하여 12개의 부밴드로 분해한다. 그리고 부모와 자식 부밴드의 각 조인트 히스토그램의 특성 함수에 대해 첫 3차의 조인트 통계적 모멘트를 계산함으로써 총 72차 특징 벡터를 추출한다. 추출된 특징 벡터는 다층 퍼셉트론 신경망(MLP, multilayer perceptron) 분류기에 입력되어 커버 영상과 스테고 영상을 분류하기 위해 학습된다.

표 1은 기존의 스테그분석 방법 및 본 논문에서 제안한 방법의 특징 벡터 추출 방법 및 차원을 비교 요약한 것이다.

본 논문의 구성은 다음과 같다. II장에서는 웨이블릿 부밴드의 통계적 성질을 살펴보고, 본 논문에서 제안한 스테그분석 방법을 설명한다. III장에서는 실험을

표 1. 기존의 스테그분석 방법 및 제안한 방법에서 특징 벡터의 추출 방법 및 차원의 비교
Table 1. Comparison of feature extraction method and dimesion in the previous steganalysis and proposed methods.

	추출 영역	추출 방법	차원
Avcibas's ^[10]	공간 영역(7, 8번째 비트 플레인의 연속성 이용)	이진 유사도(binary similarity)	10
Lyu's ^[11]	4-레벨 QMF 변환 영역 (부밴드의 계수 및 에러 영상의 통계성 이용)	mean, variance, skewness, kurtosis	72
Xuan's ^[12]	3-레벨 웨이블릿 영역(부밴드의 통계성 이용)	특성 함수(characteristic function) 모멘트	39
Shu's ^[13]	3-레벨 웨이블릿 영역 (부밴드의 계수 및 에러 영상의 통계성 이용)	특성 함수 모멘트	78
Sun's ^[14]	공간영역(4 방향으로의 차분영상의 통계성 이용)	특성 함수 모멘트	4
proposed	3-레벨 웨이블릿 영역 (부모와 자식 부밴드의 통계적 의존성 이용)	조인트(joint) 특성 함수 모멘트	72

통해 제안 기법의 성능을 확인하고, IV장에서는 결론을 맺는다.

II. 제안 방법

1. 통계적인 웨이블릿 모델

본 논문에서는 웨이블릿 영역에서 부모와 자식 부밴드간의 의존성을 관찰하기 위해 먼저 주어진 영상에 3-레벨 UWT(Undecimation Wavelet Transform) 변환을 수행한다. UWT는 DWT(discrete wavelet transform)와는 달리 변환 과정에서 한 개의 트리 이상에서 같은 계수가 나타나므로 계수의 변화를 추정하는 과정에서 매우 작은 변화를 만들게 되고, 이로 인해 부모와 자식 부밴드는 같은 특성을 보유할 수 있다.

그림 2와 같이 UWT는 1차원 신호 c_j 에 대해 필터뱅크 (h, g)를 사용하여 $W = \{w_{j+1}, c_{j+1}\}$ 의 집합을 생성한다^[15].

여기서 w_{j+1} 는 스케일 $j+1$ 에서의 웨이블릿 계수이며, c_{j+1} 는 coarsest 분해도에서의 계수이다.

UWT 변환의 분해 과정에서 각 웨이블릿 계수의 값은 다음에 의해 결정된다.

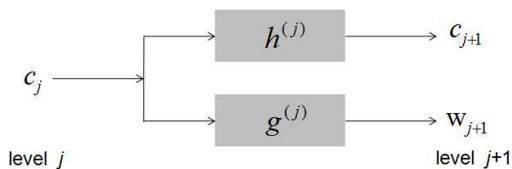


그림 2. 1차원 UWT
Fig. 2. One dimensional UWT.

$$c_{j+1}[l] = (h^{(j)} * c_j)[l] = \sum_k h[k]c_j[l + 2^j k] \quad (1)$$

$$w_{j+1}[l] = (g^{(j)} * c_j)[l] = \sum_k g[k]c_j[l + 2^j k] \quad (2)$$

여기서 만일 $l/2^j$ 가 정수이면 $h^{(j)}[l] = h[l]$ 이고, 정수가 아니면 0이다.

식 (1), (2)는 다음과 같이 2차원으로 확장될 수 있다.

$$c_{j+1}^{LL}[k, l] = (h^{(j)} h^{(j)} * c_j)[k, l] \quad (3)$$

$$w_{j+1}^{LH}[k, l] = (g^{(j)} h^{(j)} * c_j)[k, l] \quad (4)$$

$$w_{j+1}^{HL}[k, l] = (h^{(j)} g^{(j)} * c_j)[k, l] \quad (5)$$

$$w_{j+1}^{HH}[k, l] = (g^{(j)} g^{(j)} * c_j)[k, l] \quad (6)$$

여기서 $hg * c$ 는 분리 가능한 필터 hg 에 의해 c 를 컨볼루션한 것이다.

각 스케일에서 세 개의 부대역 신호 w^{LH}, w^{HL}, w^{HH} 을 얻게 되며, UWT는 DWT와는 달리 다운 샘플링(down-sampling) 과정을 거치지 않는다. 따라서 각 스케일의 부대역 신호는 원래의 신호와 동일한 크기를 가지며 시불변성(time-invariance)을 유지하게 된다. 만일 메시지의 삽입 비트율이 낮거나 특정 비트에 집중되어 있는 스테고 영상이라면 DWT에 의해 비밀 메시지의 손실이 발생할 수 있으며, 이로 인해 커버와 스테고 영상을 판별하기 어려울 수 있다.

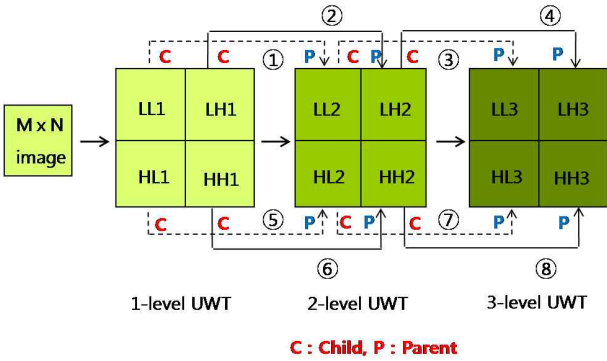


그림 3. 3-레벨 haar UWT에서 8개의 부모와 자식 부밴드
 Fig. 3. 8 parent and child subbands in 3-level haar UWT.

따라서 본 논문에서는 그림 3과 같이 주어진 영상에 대해 3-레벨 Haar UWT를 수행하여 크기가 모두 같은 12개의 부밴드(subband)로 분해한다.

이때 주어진 영상에 대해 UWT 분해한 부대역을 LL1, LH1, HL1, HH1이라 하고, 저주파 대역 LL1에 대해 한번 더 UWT 분해한 부밴드에 대해서는 LL2, LH2, HL2, HH2라 하였다. 부밴드 LL3에서 어떤 한 계수가 갖는 영상의 정보는 동일한 위치의 정보가 LL2에서도 존재하게 되며, 마찬가지로 LL2 밴드에서도 어떤 한 계수가 갖는 정보는 동일한 위치의 정보가 LL1에도 존재한다. 따라서 LL3을 부모 부밴드라 하면 LL2을 자식 부밴드라 하고, 마찬가지로 부밴드 LL2가 부모일 때 LL1은 LL2의 자식 부밴드가 된다.

본 논문에서는 그림 3과 같이 3-레벨 UWT를 통해 생성된 8개(①~⑧)의 부모(parent)와 자식(child) 부밴드에 대한 조건부 히스토그램(joint histogram)으로부터 통계적인 분포를 관찰하였다. 그림 4는 512×512 Lena 영상에 대해 자식 부밴드 C(LH1)와 부모 부밴드 P(LH2)의 조인트 히스토그램을 보인 것이다. 여기서 가로축은 $\log_2|P|$, 세로축은 $\log_2|C|$ 을 나타낸다.

전형적인 영상에서 웨이블릿 계수의 크기는 스케일 간에 강한 관계를 지니게 된다. 즉 부모 밴드가 작은 크기의 값을 가지게 되면 그 자식 밴드 또한 작은 값을 가지는 경향이 있다. 그림 4에서 보는 바와 같이 커버 영상의 경우 자식 부밴드(C)의 분포가 부모 부밴드(P)에 선형적으로 비례함으로써 부모와 자식 부밴드가 통계적으로 강한 의존성을 가짐을 알 수 있다^[16]. 그러나 영상에 랜덤 메시지를 삽입할 경우 이러한 통계적 의존성은 많은 영향을 받게 된다.

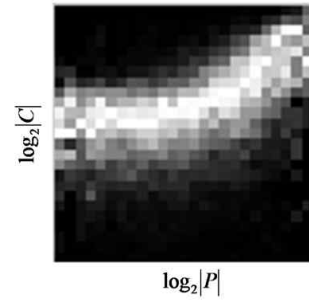


그림 4. 부모(P)와 자식(C) 부밴드의 조인트 히스토그램
 Fig. 4. joint histogram of parent and child subband.

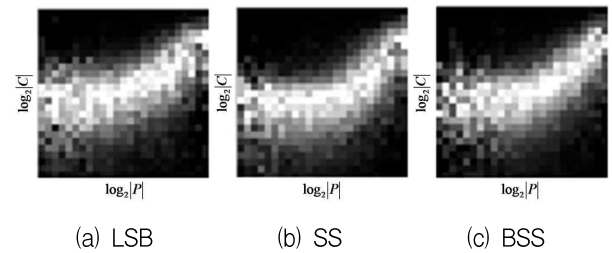


그림 5. 스테고 영상에서의 조인트 히스토그램
 Fig. 5. joint histogram in the stego images.

그림 5는 512×512 Lena 영상에 대해 0.1bpp의 삽입률을 가지는 LSB, SS, BSS 스테고 영상을 생성하고 2-레벨 LH 부밴드의 조인트 히스토그램을 비교한 것이다.

위의 그림에서 알 수 있는 바와 같이 메시지를 삽입하지 않은 그림 4의 커버 영상과 비교해 볼 때 스테고 영상은 자식 부밴드의 계수값이 보다 넓은 영역에 분산되어 부모와 자식 부밴드의 통계적 의존성이 감소함을 알 수 있다. 따라서 이러한 부모와 자식 부밴드의 통계적 의존성은 커버와 스테고 영상을 분류할 수 있는 중요한 특징으로 사용될 수 있다.

2. 특징 추출

본 논문에서는 부모와 자식 부밴드의 통계적 의존성에 기반하여 커버 영상과 스테고 영상의 조인트 히스토그램의 특성 함수로부터 조인트 모멘트를 추출하는 스테그분석 방법을 제안한다.

조인트 히스토그램의 특성 함수란 부모와 자식 부밴드의 조인트 히스토그램을 푸리에 변환한 것이다. 이때 영상으로부터 특징 벡터를 추출하기 위해 특성 함수를 사용한 이유는 조인트 히스토그램에 비해 특성 함수가 메시지 삽입에 보다 민감한 특성을 가지기 때문이다^[12].

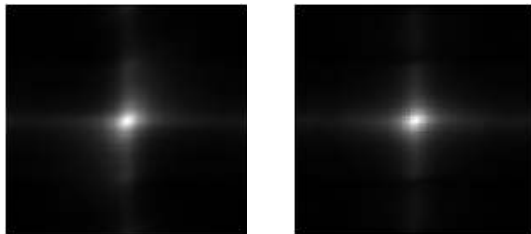
그림 6은 CorelDraw 영상 가운데 하나의 영상에서 2-레벨 LL 및 HL 부밴드의 특성 함수를 보인 것이다.



(a) 커버 및 LSB 스테고 영상



(c) 2-레벨 LL 부밴드의 특성 함수



(c) 2-레벨 HL 부밴드의 특성 함수

그림 6. 커버(왼쪽) 및 0.3bpp의 LSB 스테고 영상(오른쪽)의 특성 함수

Fig. 6. characteristic function of cover(left) and 0.3bpp LSB stego(right) image.

그림으로부터 알 수 있듯이 조인트 히스토그램의 특성 함수는 특히 고주파 영역에서 커버 영상과 뚜렷한 차이를 보인다. 이러한 통계적 특징들의 차이를 계산하기 위해 본 논문에서는 부모와 자식 부밴드의 조인트 특성 함수로부터 식 (7)과 같이 절대값 조인트 모멘트 (absolute joint moment)를 추출한다.

$$M_{mn}^A = \frac{\sum_{k,l=0}^{N/2-1} k^m l^n |H(k,l)|}{\sum_{k,l=0}^{N/2-1} |H(k,l)|} \quad (7)$$

여기서 m, n 은 모멘트의 차수이며, N 는 조인트 히스토그램의 특성 함수인 $H(k,l)$ 의 길이를 의미한다. 또한 $|H(k,l)|$ 은 조인트 히스토그램의 특성 함수의 크기를 나타낸다.

본 논문에서는 그림 3에서 설명한 8개의 부모와 자식 부밴드의 각 조인트 히스토그램의 특성 함수에 대해 첫 9차 조인트 모멘트, 즉 $M_{11}^A, M_{12}^A, M_{13}^A, M_{21}^A, M_{22}^A, M_{23}^A, M_{31}^A, M_{32}^A, M_{33}^A$ 를 계산하여 총 72차 특징 벡터를 추출한다.

III. 실험 및 결과

제안한 기법의 성능을 평가하기 위해 본 논문에서는 그림 7과 같이 1000개의 압축되지 않은 다양한 CorelDraw 영상을 사용하였다^[17]. 데이터베이스 영상의 크기는 512×768 또는 768×512이며, LSB^[5] 및 SS^[6] 그리고 BSS^[7] 임베딩 방법을 사용하여 0.1bpp, 0.3bpp, 그리고 0.5bpp의 삽입률을 갖는 다양한 스테고 영상을 만들었다. 이 가운데 800개의 영상은 MLP(다층 퍼셉트론 신경망) 분류기의 학습을 위해 사용하며, 200개는 테스트를 위해 사용하였다.

성능 비교를 위해 사용된 기존의 스테그분석 방법은 웨이블릿 영역 기반의 대표적인 스테그분석 방법인 Xuan의 방법^[12]과 최근 발표된 공간 영역의 스테그분석 방법인 Sun의 방법^[14]이다.

표 2는 각 비트율의 LSB와 SS 그리고 BSS의 스테고 영상에 대해 기존의 스테그분석 방법과 본 논문에서 제안한 기법을 비교 실험한 결과를 보인 것이다. 이때 비밀 메시지의 삽입 유무에 대한 검출율은 10번 실험한 후 평균값으로 나타내었다. 여기서 민감도(Sensitivity)는 커버 영상에 삽입 정보가 없다고 판단하는 경우이며, 특이도(Specificity)은 스테고 영상에 삽입 정보가 존재한다고 판단하는 경우이다^[18]. 또한 정확도(Accuracy)는 커버 영상과 스테고 영상을 정확하게 판단한 정도를 의미하는 것으로서 민감도와 특이도의 평균값이다.

실험 결과 Xuan의 방법은 평균적으로 76%, Sun의 경우 58.5%, 제안한 기법은 80.4%의 검출율을 보임을 알 수 있다. 특히 SS 스테고 영상의 경우 제안된 기법



그림 7. 비압축된 CorelDraw 데이터베이스 영상의 예
Fig. 7. Sample images from the uncompressed CorelDraw Database Image.

표 2. 기존의 스테그분석 방법과의 성능 비교(단위 : %)

Table 2. Performance comparison in the previous and proposed steganalysis methods(units : %).

삽입 방법	스테그분석 방법 비트율	Xuan's method ^[12]			Sun's method ^[14]			Proposed method		
		민감도	특이도	정확도	민감도	특이도	정확도	민감도	특이도	정확도
LSB	0.1	17.0	87.5	52.3	51.0	50.0	50.5	58.0	94.5	76.3
	0.3	34.5	84.0	59.3	52.0	49.0	50.9	73.5	95.0	84.3
	0.5	50.5	76.0	63.3	51.0	49.0	50.0	72.0	96.5	84.3
SS	0.1	85.0	84.0	84.5	48	85.5	66.8	74.0	86.5	80.3
	0.3	82.5	94.0	88.3	58.5	86	72.3	76.0	86.0	81.0
	0.5	88.5	82.0	85.3	61.0	53.0	57.0	77.5	83.0	80.3
Blind SS	0.1	70.0	80.5	75.3	48.0	60.5	54.3	60.5	87.0	73.8
	0.3	71.5	80.0	75.8	52.0	65.5	58.8	71.0	88.0	79.5
	0.5	69.5	88.5	79.0	56.5	61.5	59.0	67.5	93.5	80.5
Average		63.2	84.1	76.0	53.1	62.2	58.5	70.2	89.8	80.4

은 Xuan의 방법에 비해 평균적인 검출율은 낮으나 여전히 80% 이상의 검출율을 보이고 있으며, 거의 모든 비트율의 LSB 및 BSS 스테고 영상에서 Xuan과 Sun의 방법에 비해 검출율이 향상되었다.

스테그분석에서 커버 영상을 커버 영상으로 판단하는 민감도도 중요하지만 스테고 영상을 커버 영상으로 잘못 판별하는 경우 심각한 보안 문제가 발생할 수 있다. 따라서 스테고 영상을 정확하게 스테고 영상으로 판별하는 특이도의 평가가 더욱더 중요하다. 이러한 점에서 볼 때 본 논문에서 제안한 방법은 0.3bpp의 SS 방법을 제외하고는 기존의 방법에 비해 거의 모든 비트율의 스테고 영상에서 특이도가 더 높음을 확인할 수 있다.

또한 한가지 스테가노그래피 방법에 대한 스테그분석은 실제 응용에서 사용되기 어렵다. 따라서 본 논문에서는 스테고 영상의 종류를 모르는 경우에 있어서 검출 능력을 확인하기 위해 커버 영상과 0.3bpp의 비트율을 가지는 LSB, SS 그리고 BSS 스테고 영상을 200개씩 랜덤하게 추출하였다. 그리고 이를 한꺼번에 학습시킨 후 400개의 테스트 영상에 대한 삽입 정보 유무의 검출율을 살펴 보았다. 표 3은 이의 검출 성능을 보인 것이다.

표 3으로부터 다양한 스테고 영상들에 대해 기존의 방법은 약 50%의 정확도를 보이는 반면 제안한 기법은 평균적으로 91.5%의 높은 정확도를 보임을 알 수 있다.

표 3. 0.3bpp의 삽입률에 대한 검출율 비교(단위 : %)

Table 3. Performance Comparison to 0.3bpp(units: %).

스테그분석 평가 기준 방법	Xuan's ^[12]	Sun's ^[14]	Proposed
Sensitivity	32.0	46.0	94.0
Specificity	84.0	53.0	89.0
Precision	58.0	49.5	91.5

본 논문에서 제안한 기법의 또다른 성능 평가를 위해 ROC(Receiver Operating Characteristic) 곡선을 비교 분석하였다. ROC 곡선^[19]은 예측된 검출 결과와 실제 검출 결과를 비교하기 위한 그래프이며, 분류기의 검출 임계치(threshold)가 변화할 때 검출율(detection probability)과 오탐지율(false alarm probability)이 어떻게 변하는가를 보여준다.

검출율은 스테고 영상을 정확히 스테고 영상으로 분류할 확률을 말하며, 오탐지율은 커버 영상이 스테고 영상으로 분류되는 확률을 나타낸다. 따라서 좋은 분류기는 작은 오탐지율에서도 검출율이 높아야 하며, ROC 곡선은 왼쪽 상단에 위치하는 것이 좋다.

그림 8은 0.3bpp의 삽입률을 갖는 LSB, SS, BSS 스테고 영상에 대한 ROC 곡선을 보여준다. 그림에서 제안한 방법은 모든 스테고 영상에서 ROC 곡선이 왼쪽 상단에 위치함을 알 수 있으며, 특히 LSB와 BSS 스테고 영상의 경우 기존의 방법에 비해 더욱 향상됨을 알

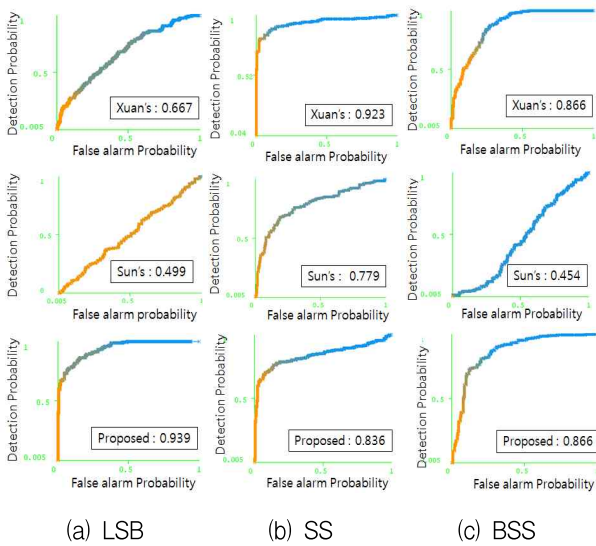


그림 8. 기존의 방법과 제안한 방법의 ROC 곡선의 비교

Fig. 8. Comparison of ROC curve in the previous and proposed method.

표 4. 기존의 방법과의 ROC 면적 비교

Table 4. ROC area of previous and proposed methods.

삽입 방법		Xuan's ^[12]	Sun's ^[14]	Proposed method
LSB	0.1	0.584	0.496	0.869
	0.3	0.667	0.499	0.939
	0.5	0.724	0.498	0.941
SS	0.1	0.923	0.725	0.869
	0.3	0.923	0.779	0.836
	0.5	0.933	0.789	0.875
Blind SS	0.1	0.857	0.475	0.800
	0.3	0.866	0.454	0.866
	0.5	0.895	0.610	0.896
Average		0.819	0.592	0.877

수 있다.

각 스테그분석 방법에 기재된 수치는 ROC 곡선 방향의 면적을 나타낸다. ROC 면적은 스테고 영상의 판별에 대한 정확도를 나타내는 척도로서 1에 가까울수록 분류 능력이 우수하며, 0.5 이하의 값을 가질 때 스테고 영상을 정확히 판별할 수 없게 된다.

본 논문에서는 Mann-Whitney 통계적 분석법을 이용하여 ROC 면적을 계산하였으며^[21], 표 4는 기존의 방법과의 ROC 면적을 비교한 것이다.

표의 결과로부터 알 수 있듯이 본 논문에서 제안한

기법은 모든 비트율의 스테고 영상에서 0.8 이상의 높은 ROC 면적을 가지며, Xuan의 방법은 평균 0.819, Sun의 방법은 0.592, 그리고 제안한 방법은 0.877의 면적을 보인다. 이는 기존의 방법에 비해 제안한 방법에서 검출 결과의 신뢰도가 보다 더 높음을 의미한다.

IV. 결 론

본 논문에서는 커버 영상과 스테고 영상 사이의 통계적인 특성을 구하기 위해 웨이블릿 부밴드의 조인트 히스토그램에 기반한 스테그분석 기법을 제안하였다. 제안한 기법은 웨이블릿 영역에서 부모와 자식 부밴드의 계수들이 서로 독립적이지 않다는 통계적인 성질을 이용하여 부모와 자식 부밴드의 조인트 히스토그램의 특성 함수로부터 72차의 특징 벡터를 추출하였다. 추출된 특징 벡터는 MLP 분류기에 입력되어 커버 영상과 스테고 영상을 구분하기 위해 학습되었다. 다양한 삽입률의 LSB, SS, BSS 스테고 영상에 대해 제안된 기법을 실험한 결과 제안된 기법은 기존의 기법에 비해 평균적으로 높은 검출율과 정확도를 가짐을 확인할 수 있었다. 특히 LSB의 스테고 영상의 경우 기존의 방법에 비해 평균 20% 이상의 높은 검출율을 보였다. 또한 모든 스테고 영상들을 랜덤하게 선택하여 학습한 결과 91.5% 이상의 높은 검출율을 보였다.

참 고 문 헌

- [1] <http://en.wikipedia.org/wiki/Steganography>
- [2] D. Kahn, "The history of steganography," in Proc Information Hiding, First International Workshop, Cambridge, U.K., 1996.
- [3] R. Anderson and F. Petitcolas, "On the limits of steganography," IEEE J. Sel. Areas Commun., vol. 16, no. 4, pp. 474-481, 1998.
- [4] N. Johnson and S. Jajodia, "Exploring steganography: Seeing the unseen," IEEE Computer, vol. 31, no. 2, pp. 26-34, 1998.
- [5] T. Zhang, X. Ping, "A new approach to reliable detection of LSB steganography in natural images," Signal Processing, vol. 83, no. 10, pp 2085-2093, 2003.
- [6] I.J. Cox, J. Kilian, T. Leighton and T. Shamon, "Secure Spread Spectrum Watermarking for Multimedia," IEEE Trans. On Image Processing, vol. 6, pp. 1673-1687, 1997.

- [7] A. Piva, M. Barni and E. Bartolini, "DCT-based watermark recovering without resorting to the uncorrupted original image," Proc. ICIP 97, vol. 1, pp. 520
- [8] <http://en.wikipedia.org/wiki/Steganalysis>
- [9] M.A. Mehrabi, H. Aghaeinia, 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.
- [10] I. Avcibas, N. Memon, and B. sankur, "Image steganalysis with binary similarity measures," in IEEE International Conference on Image Processing, September, 2002.
- [11] S. Lyu, H. Farid, "Detecting hidden messages using higher order statistics and support vector machines," presented at the 5th Int. Workshop on Information Hiding, Noordwijkerhout, The Netherlands, 2002.
- [12] 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.
- [13] Y. Shi, G. Xuan, C. Yang, J. Gao, Z. Zhang, P. Chai, D. Zou, C. Chen, W. Chen, "Effective Steganalysis Based on Statistical Moments of Wavelet Characteristic Functions," Intl. Conf. on Infor.Tech., Las Vegas, USA, April 4-6, 2005.
- [14] Y. Sun, F. Liu, B. Liu and P. Wang, "Steganalysis Based on Difference Image," IWDDW 2008, LNCS, vol. 5450, pp. 184-198, Springer, Heidelberg, 2009.
- [15] J.-L. Starck, J. Fadili, F. Murtagh, "The Undecimated Wavelet Decomposition and its Reconstruction," IP(16), No. 2, February 2007, pp. 297-309. IEEE DOI Link 0702
- [16] R.W. Buccigrossi, E.P. Simoncelli, "Image Compression via Joint Statistical Characterization in the Wavelet Domain," IEEE Trans Image Proc, vol. 8(12), pp. 1688-1701, Dec 1999.
- [17] CorelDraw Software, www.CorelDraw.com
- [18] I. H. Witten and E. Frank, *Data Mining*, Elsevier, 2005.
- [19] A. P. Bradley. *The use of the area under the ROC curve in the evaluation of machine learning algorithms*, *Pattern Recognition*, 30:1145-1159, 1997.
- [20] 원태연, 통계조사분석, 한나래아카데미, 2010.

 저 자 소 개



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, 스테그분석>



현 승 화(학생회원)
2003년 밀양대학교 정보통신 공학과 학사 졸업.
2005년 부산대학교 전자공학과 석사 졸업.
2011년 부산대학교 전자공학과 박사 졸업.

<주관심분야 : 영상처리, 오류은닉, 웨이블릿 변환 신호처리, 스테그분석>



김 재 호(정회원)
1980년 부산대학교 전기기계 공학과 학사 졸업.
1982년 한국과학기술원 산업 전자공학과 석사 졸업.
1990년 한국과학기술원 전기및 전자공학과 박사 졸업.

현재 부산대학교 전자공학과 교수
<주관심분야 : 영상처리, 영상압축, VLSI, SOC 설계>



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, 스테그분석>