제33회 한국정보처리학회 춘계학술발표대회 논문집 제17권 제1호 (2010. 4)

디지털 멀티미디어 컨텐츠 불법 사용 차단을 위한 스테가노그라피 연구

김대경*, 이상순**, 이병수*
*인천대학교 컴퓨터공학과
**가천의과학대학교 컴퓨터공학과
e-mail:{techntip, bsl}@incheon.ac.kr*, sslee@gachon.ac.kr**

A Study on Steganography for Block Illegal Use of Digital Multimedia Content

Dae-Kyeong Kim*, Sang-Soon Lee**, Byung-Soo Lee*
*Dept of Computer Engineering, University of Incheon
**Dept of Computer Engineering, Gachon University of Medicine and Science

요 약

본 연구는 디지털 멀티미디어 컨텐츠에 대한 불법 사용 차단 및 저작권 보호 기술에 대한 연구이다. 본 연구 주제와 관련해서 스테가노그라피 기술 자체에 대한 연구와 저작권을 위한 워터마킹 기술, 정당한 수신자의 권리를 위한 스크램블링 연구로 나뉜다. 본 논문은 저작권 증명, 원본 증명, 정보 추출, 정보 은닉을 위해 활용되어질 수 있는 스테가노그라피 기술을 다룬다.

1. 서론

1.1 연구배경 및 필요성

디지털 멀티미디어 컨텐츠와 관련 저작물을 안전하게 전달하고 사용을 통제하는 DRM 암호기술, 불법 복제물의 탐색 및 색출을 위한 검색엔진, 불법 복제 및 사용을 막기 위한 복제 관리 기술 등이 연구되고 있다[1]. 특히 암호화와 스테가노그라피, 그리고 워터마킹 기술이 주를 이루나 각 장단점이 있다. 암호화는 암호화와 복호화를 통해 효과적인 보안을 꾀하지만, 암호화정보라는 것을 숨길 수 없기 때문에 쉽게 공격대상이 된다. 워터마킹은 그것이 해적작품임을 확인해 저작권 보호에 일조하나 불법복제를 방지해주지는 못한다. 스테가노그라피에 있어 스크램블링 연구는 그것이 정당한 수신자의 권리 보호에는 일조하나 또한 디지털 멀티미디어 컨텐츠와 관련 복합적 문제, 곧 저작권 증명, 원본 증명, 정보 추출, 정보 은닉 기능을 모두 제공을 해결해주지 못한다.

1.2 연구목적

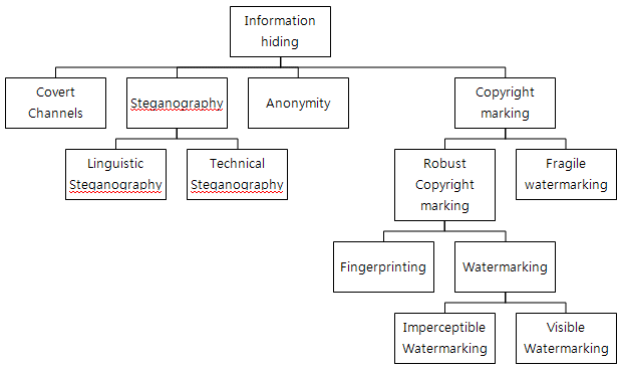
따라서 본 연구는 디지털 멀티미디어 컨텐츠와 관련한 복합적 문제, 저작권 보호와 정당한 수신자 보호에 대한 스테가노그라피 기술을 연구한다.

1.3 연구범위

본 논문의 나머지 부분은 다음과 같이 구성된다. 2장에서는 스테가노그라피 관련 연구들에 대해 설명하고 암호화와 워터마킹, 스테가노그라피에 대해 비교·설명한다. 3장에서는 디지털 멀티미디어 컨텐츠와 관련한 복합적 문제, 저작권 증명, 원본 증명, 정보 추출, 정보 은닉을 위한 방법을 소개한다. 4장에서는 실험결과를 보여주고, 5장에서 결론을 맺는다.

2. 관련 연구

스테가노그라피는 정보은닉 방법 중 하나이다. 스테가노그라피의 구성은 다음 (그림 1)과 같다.



(그림 1) 정보은닉 기술의 분류[2]

2.1 스테가노그라피 관련연구

스테가노그라피의 연구는 다음과 같이 이뤄졌다.

○ 커버에 비밀 메시지를 삽입하는 방법에 대한 연구

-텍스트 스테가노그라피와 관련된 연구로는 Line-Shift Coding, Word-Shift Coding 등이 있다[3].

-오디오 스테가노그라피와 관련된 연구로는 LSB Encoding, Phase Encoding, Spread Spectrum, Echo Data Hiding 등이 있다[4,5].

-이미지 스테가노그라피와 관련되 연구로는 첫째 LSB(Least Significant Bit)[6,7,8]와 같은 공간적 스테가노그라피, 그리고 비트 플래인(bit plane) 삽입방법이 있다. 이외 정보를 주파수 영역과 같은 변환영역에 삽입하는 방

법, Spread Spectrum 통신의 개념을 이용한 방법, 표본의 평균등을 이용한 통계적인 방법, 원래의 Cover 데이터에 왜곡을 주고 수신단에서 Cover 데이터와의 편차를 측정하여 정보를 판독하는 왜곡 방법[9]등이 있다.

○ 압축에 대한 연구

압축과 관련해서는 벡터 양자화(Vector Quantization)[10]가 있으며, PSNR을 스테가노그라피에 적용한 연구[9] 등이 있다.

○ 스크램블링하는 방법에 대한 연구

스크램블링은 크게 공간 영역에서의 스크램블링, 주파수 영역에서의 스크램블링 방법, 움직임 벡터를 이용한 방법, 암호화 알고리즘을 이용한 방법 등으로 나눈다[11].

○ steganalysis 연구

steganalysis은 스테가노그라피에 의해 만들어진 Stego 데이터에 은닉정보가 있는지를 알아내는 기술이다. 이 기술은 감지, 파괴 또는 검출 및 해독의 과정으로 분류한다[12].

2.2 암호화, 워터마킹, 스테가노그라피

<표 1>은 디지털 멀티미디어 컨텐츠 보호 기술인 스테가노그라피, 워터마킹, 암호화, 스크램블링의 특징을 정리한 것이다.

<표 1> 디지털 멀티미디어 컨텐츠 보호 기술의 비교

내용 / 종류	3자에 의한 의심	통신의 주체	은닉 정보	저작권 보호	수신자의 권리 보호
Cryptography	○	원본	×	○	×
Watermarking	△	원본	△(소유권, 라이센스)	○	×
Scrambling	△	원본	×	×	○
Steganography	×	메시지->원본	○(소유권, 라이센스)	○	○

<표 1>에서 보면 스테가노그라피는 정보은닉이라는 곧, 통신의 주체가 메시지라는 사고에서 보면 영상내용이 실제정보가 아닌 탓에 수신자의 권리보호에 일조하지 못하는 것처럼 보인다. 하지만 통신의 주체가 메시지가 아니라 원본, Cover 데이터라면 정보은닉 내용은 실제 정보인증도구로 사용될 수 있다. 또한 저작권 증명, 원본 증명, 정보 추출, 정보 은닉 기능을 모두 제공할 수 있다.

3. 제안 스테가노그라피 방법

3.1 스테가노그라피 개요

스테가노그라피는 일명 정보은닉 기술로 불리는 것으로서, 원본 이미지 파일의 크기 변화 및 손상이 전혀 없는 상태에서 다른 형태의 파일을 40% 이내의 범위에서 숨길 수 있는 기술이다[13].

스테가노그라피의 일반적 식은 다음과 같다.

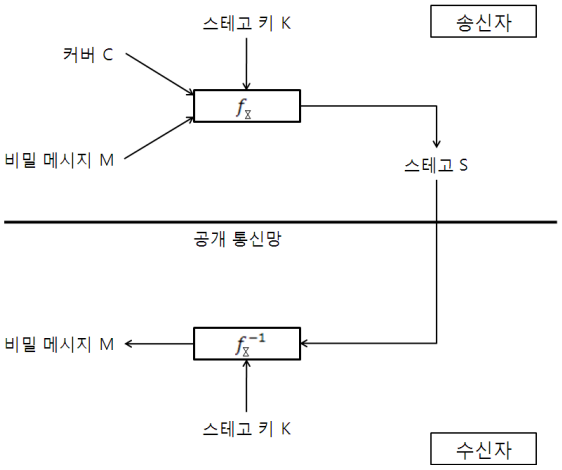
$$C=Q+Q' \qquad (식1)$$

커버 영상(C), 인식할 수 있는 정보의 양(Q), 정보를 변형하여도 인식할 수 없는 정보의 양(Q′)으로 나눌 수 있다. 비밀 정보 통신을 위한 스테가노그라피는 Q′부분을 변형하여 정보를 삽입한다.

다음은 스테가노그라피의 일반적 식을 위한 구성 요소이다[14].

비밀 메시지 : M
커버 : C
스테고 : S
스테고 키 : K
삽입 알고리즘 : $f_E(C,M,K) \rightarrow S$
추출 알고리즘 : $f_E^{-1}(S,K) \rightarrow M$

스테가노그라피 시스템의 일반적인 모델은 (그림2)와 같이 나타낸다.



(그림 2) 스테가노그라피 시스템의 일반적 모델

스테가노그라피는 비가시성, 삽입용량, 보안성, 강인성, 비검출성을 특징으로 한다. 스테가노그라피 기술은 비가시성과 삽입용량을 고려하기 위해 인간 시각 시스템(HVS : Human Visual System)에 기반하여 시각적으로 덜 민감한 부분 등에 비밀 정보를 대치하며 많은 양을 처리한다.

3.2 스테가노그라피 분류

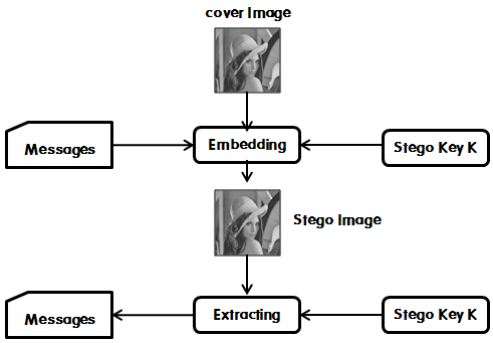
따라서 스테가노그라피는 정보 삽입 및 추출 시 비밀키인 스테고 키 사용 유무와 키 운용 방법에 따라 3가지 형태로 분류한다. 순수 스테가노그라피(pure steganography), 비밀키 스테가노그라피(private key steganography), 공개키 스테가노그라피(public key steganography)가 그것이다[15].

<표 2> 스테가노그라피의 분류

키 분류	식 구성요소
순수 스테가노그라피	비밀 메시지 : M 커버 : C 스테고 : S 삽입 알고리즘 : $f_E(C,M) \rightarrow S$ 추출 알고리즘 : $f^{-1}_E(S) \rightarrow M$
비밀키 스테가노그라피	비밀 메시지 : M 커버 : C 스테고 : S 스테고 키 : K 삽입 알고리즘 : $f_E(C,M,K) \rightarrow S$ 추출 알고리즘 : $f^{-1}_E(S,K) \rightarrow M$
공개키 스테가노그라피	비밀 메시지 : M 커버 : C 스테고 : S 공개키 : P 스테고 키 : K 삽입 알고리즘 : $f_E(C,M,K) \rightarrow S$ 추출 알고리즘 : $f^{-1}_E(S,K) \rightarrow M$

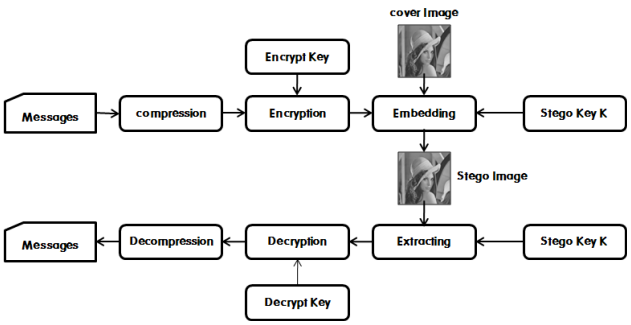
3.3 제안 스테가노그라피 시스템 모델

처음 순수 스테가노그라피를 처리하는 일반적 모델 처리 프로세스를 그려보면 다음 (그림 3)과 같다.



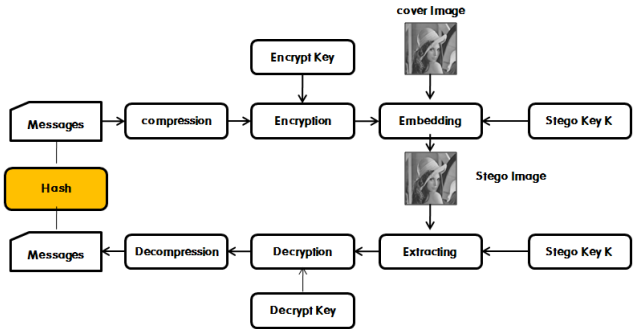
(그림 3) 순수 스테가노그라피 처리 프로세스 블록 다이어그램

(그림 3)은 메시지와 Cover 이미지만 있는 일반적 형태의 스테가노그라피 처리를 위한 블록 다이어그램이다. 반면 (그림 4)는 압축 및 암호화 연구를 통해 보다 더 진전된 스테가노그라피 처리 블록다이그램이다.



(그림 4) 보다 진전된 스테가노그라피 처리 프로세스 블록 다이어그램

본 연구는 다음 (그림 5)와 같이 저작권 증명, 원본 증명, 정보 추출, 정보 은닉 기능 처리를 위한 알고리즘을 위해 해시를 적용한다.



(그림 5) 본 연구에서 제시한 해시 적용 스테가노그라피 처리 프로세스 블록 다이어그램

해시를 통해 판독된 메시지에 대한 처리 알고리즘은 (그림 6)과 같다.

```
1: Hash 처리를 통해 암호화 파일을 복호화한다.
2: if 암호화 파일 ≠복호화 파일 then
3:    RETURN FALSE;
4: else
5:    저작권 증명, 정보추출, 정보은닉 가능;
6:end if
```

(그림 6) 판독된 메시지에 대한 처리 알고리즘

4. 실험 및 결과

4.1 실험모델

본 연구에서 실험은 은닉하기 전 파일과 은닉된 후 추출되었을 때의 파일 동일성을 다룬다. 이를 위해 MD5와 SHA 알고리즘의 비교 분석을 한다.

4.2 실험조건

(1) 원 영상-cover Image

512X512 Lena 영상 이 영상의 경우 수많은 이미지 프로세싱 연구에서 기준으로 삼는 영상이므로 다른 알고리

즘과 비교하기 위한 경우 좋은 대상이 된다.

(2) 암호화 파일을 삽입한 영상-stego Image

실험의 구현 원리를 설명하기 위해서 Cover 이미지에 암호를 집어넣고 위·변조 감별을 위해 복호화를 해본다.

(3) Hash 알고리즘을 통한 암·복호화 파일의 동일성 검증

감별을 위해서 MD5, SHA 해시 알고리즘을 이용한다. SHA(Secure Hash Algorithm) 함수들은 서로 관련된 암호학적 해쉬 함수들의 모음이다. SHA 함수군에 속하는 최초의 함수는 공식적으로 SHA라고 불리지만, 나중에 설계된 함수들과 구별하기 위하여 SHA-0이라고도 불린다. 2년 후 SHA-0의 변형인 SHA-1이 발표되었으며, 그 후에 4종류의 변형, 즉 SHA-224, SHA-256, SHA-384, SHA-512가 더 발표되었다. 이들을 통칭해서 SHA-2라고 하기도 한다.

4.3 실험결과

<표 3> 실험 결과

구분	Embedding 전 암호화 파일	Extracting 후 복호화 파일
MD5	Calculating hash of 1083 bytes file 'C:\works\stego\encrypt.txt'... MD5 : 1DB41368FF7CE6E8CCD7D7F7DAE94F57 Calculation took 0.016 seconds	Calculating hash of 1083 bytes file 'C:\works\stego\dec\decrypt.txt' ... MD5 : 1DB41368FF7CE6E8CCD7D7F7DAE94F57 Calculation took 0.015 seconds
SHA 512	Calculating hash of 1083 bytes file 'C:\works\stego\encrypt.txt'... SHA-512 : A4FC4A04150D25890A0A5A59B198AD4BFAA109673CAC6450E397EFAF8D056C5A308D3179460E69F4038AEC547061BFABAC0273E9AF621F0AEAB085027919078A Calculation took 0.016 seconds	Calculating hash of 1083 bytes file 'C:\works\stego\dec\decrypt.txt' ... SHA-512 : A4FC4A04150D25890A0A5A59B198AD4BFAA109673CAC6450E397EFAF8D056C5A308D3179460E69F4038AEC547061BFABAC0273E9AF621F0AEAB085027919078A Calculation took 0.016 seconds

본 실험을 통해 암호화 파일과 복호화 파일 사이의 동일성이 유지됨을 알 수 있었다. 또한 본 연구의 실험을 통해 암호화 파일이 훌륭하게 복호화되며 원본과 같음을 확인할 수 있었다. 이렇게 검출된 복호화 파일은 저작권 증명, 원본 증명, 정보 추출, 정보 은닉을 위해 활용되어 질 수 있다.

5. 결론

스테가노그라피는 데이터 은닉을 위한 암호화 방법의 한계와 저작권 인증을 위한 워터마킹 방법의 삽입 용량에 대한 한계를 극복할 수 있는 방법이다 암호학의 경우는 전송 사실을 쉽게 알 수 있다. 또 최근의 처리 능력으로 인해 해독이 쉬워지고 있다. 하지만 스테가노그라피는 정보의 전달 자체를 감추기 때문에 보다 안전하다.

이후 스테그어날리시스와 스테고 파일 감지에 대비 역어셈블 방지와 디버깅 방지를 위한 조작 방지 기술과 코드 혼란 기술에 대한 연구와 좀 더 향상된 기능의 스테고 기술, 정책적인 암호화 그리고 보다 더 진전된 해시 처리 기술 등에 대한 연구가 필요하다.

참고문헌

[1] 마크 스탬프, 안태남, 손용락, 이광석 역, “정보보안 이론과 실제:암호, 접근제어, 프로토콜, 소프트웨어”, 한빛미디어(주), 2008.

[2] B. Pfitzmann, "Information Hiding Terminology, "Information Hiding: first international workshop, 1996.

[3] Stefan Katzenbeisser, and Fabien A. P. Petitcolas "Information hiding techniques for steganography and digital watermarking," Artech House Publishers, 2000.

[4] S.K. Pal, P.K. Saxena, and S.K. Mutoo, "The Future of Audio Steganography," STEG'02, 2002.

[5] 오종은, “웨이브 오디오 스테가노그래피의 차이점 분석”, 숭실대학교 정보과학대학원, 2004.

[6] Y.K. Lee and L. H. Chen, "An Adaptive Image Steganographic Model based on Minimum-Error LSB Replacement," Proc of ICMCS, vol. 2, pp. 506-511, 1999.

[7] D.C. Wu and W.H. Tsai, "A Steganographic Method for Images by Pixel-Value Differencing," Pattern Recognit, Lett, 2003, 24, (9-10), pp. 1613-1626, 2003.

[8] R.Z. Wang, C. F. Kin, and J. C. Lin, "Image hiding by optimal LSB substitution and genetic algorithm", Pattern Recognition, vol. 34. no. 3, pp. 671-683, 2001.

[9] 최용수, "확장된 컬러 매핑 방법을 이용한 색인화된 영상에서의 압축과 스테가노그라피", 강원대학교대학원, 2006.

[10] Yoseph Linde, Andres Buzo, and Robert M.Gray, "An Algorithm for Vector Quantizatizer Design," IEEE Trans. on Comm, VOL-COM, no. 1, 1980.

[11] Hobbs and G. Lamont, "Video Scrambling," US patent, No.5815572, 1998.

[12] 김형중, “스테가노그라피의 이론적 배경과 검출기법”, 정보보호학회지 제12권, 제1호, pp. 35-48, 2002.

[13] G. simmons, "The Prisoners Problem and the Subliminal Channel," Proc of CRYPTO, pp. 51-67, 1983.

[14] W.Bender, D. Gruhl, N. Mormoto, and A. Lu, "Techniques for Data Hiding," IBM System Journal, vol. 35, no. 3-4, pp. 313-336, 1996.

[15] R.J.Anderson, "Stretching the Limits of Steganography," Information Hiding, Springer Lecture Notes in Computer Science, vol. 1174, pp. 39-48, 1996.