

대용량 텍스트를 위한 손실 없는 영상 은닉기술

이근무*

*위덕대학교 정보통신공학부

e-mail: kmrhee@uu.ac.kr

Simple Image Steganography Technology for Large Scale Text

Keun-Moo Rhee*

*School of Information & Communication Ui-Duk University

요 약

이미지 혹은 문서 은닉기술은 문서, 영상, 오디오 등 모든 종류의 디지털 데이터에 대하여 연구가 진행되고 있다. 이들은 다양한 목적과 용도로 이용되고 있다. 본 연구에서는 낮은 수준의 보안을 요하는 텍스트를 대용량으로 은닉하여 전달할 수 있는 단순하고 단순한 기법을 구현하였다. 먼저 텍스트 이미지를 결합하고 이를 24 비트 심도의 칼라이미지에 인코딩하여 복구하였다. 결과는 상관기법을 이용하여 분석하였으며 텍스트 이미지의 손실이 미미한 것으로 판명되었다.

1. 서론

인간의 오관에 의한 지각능력은 지극히 제한적이다. 청각의 경우 20 db에서 최대 20000 db 이외의 정보는 인간의 인식능력을 넘어서는 것이며 시각 역시 초당 16 프레임을 넘어서면 잔상에 의해 정지화상을 동영상으로 인식하게 된다. 이러한 인간의 인식능력의 무능함에 호소하는 암호화 기법이 이미지, 오디오 동영상등에 문자등의 정보를 은닉해 전달하는 기법이며 이를 스테가노그래피(steganography)라 한다. 즉 스테가노그래피는 비밀이라는 그리스어원(steganos)에서 유래하며 잘 보이지 않게 데이터를 은닉하는 방법을 이른다. 역사적으로 스테가노그래피는 군대, 범인, 연구원들의 범위였다. 메시지의 존재 자체를 감추는 비밀통신 방법 17세기 영국 청교도 혁명시대에 크롬웰에 의해 체포된 왕당파 충신 존 트레베리언 경에게 전달된 숨겨진 메시지가 대표적 예이다. 스테노그래피는 정치군사적 이유와 기업의 비밀유출사건등에도 이용되면서 기업사회에서도 큰 관심을 끌고 있다. 스테가노그래피를 사용해 감춰진 기밀이 탐지되지 않거나 보고되지 않기

때문에 손실을 정량화하기가 힘들다. 스테가노그래피의 일부 형태가 이미 오래전부터 수천 년 동안 사용되어 오고 있지만, 컴퓨터 기술과 인터넷이 이런 유형의 은밀하게 전달되는 커뮤니케이션을 완전히 새로운 수준으로 바꿔 놓았다. 일반적으로 이미지 혹은 문서 은닉기술은 문서, 영상, 오디오 등 모든 종류의 디지털 데이터에 대하여 연구가 이루어지고 있으나, 본 논문에서는 영상 데이터에 대해서 논의한다. 본 논문의 구성은 다음과 같다. 먼저, 2장에서는 관련연구와 대표적 기법들을 소개하고 3장에서는 간단한 방법으로 이미지 데이터에 대용량텍스트를 은닉하고 복구하는 방법을 제안하고 4장에서는 제안한 은닉 기법에 대하여 구현 결과 와 그 유효성을 확인하며 5장의 결론에서는 연구의 의미와 향후의 연구 과제를 제시한다.

2. 관련연구

이미지인식에서 인간의 시지각의 능력은 제한적이다. 인간의 시지각 능력은 잔상이나 고주파영역에서 특히 변화에 둔감하다. 따라서 인간이 인식할 수 있는 범위 내에서의 조작될 수 있는 정보들을 이용

할 수 있는 것이다. 이를 이용하는 방법들은 마스크링과 필터링, 수학적 함수 등을 이용하는 변형기법, 이미지의 변형을 적게 주는 LSB에 치환 기법들이 연구되어 왔다.[1]-[4] 본 연구에서는 치환기법을 이용하여 구현하였다. [5]

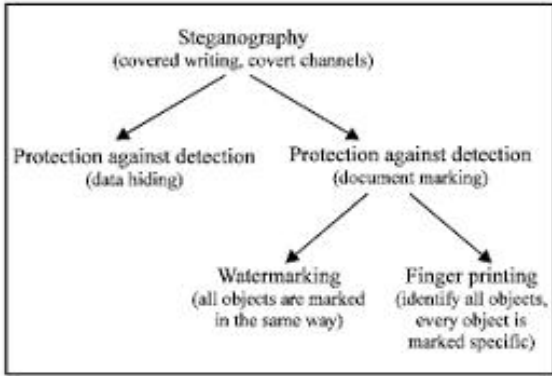


그림 2 Schematic representation of steganographic procedure[2]

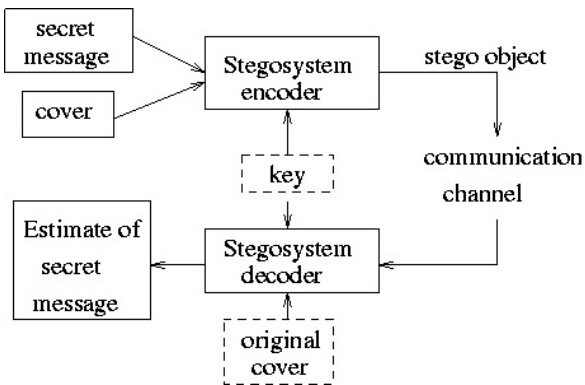


그림 2 정보은닉의 일반모형

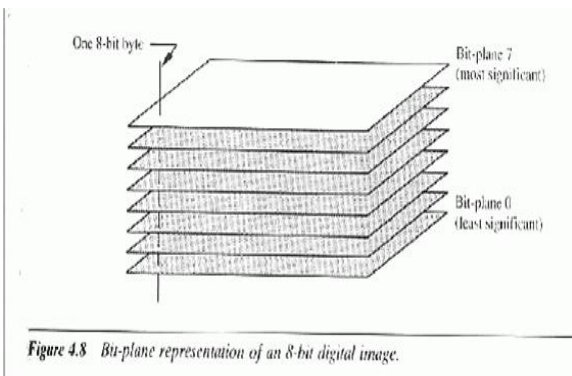


그림 4 Image bit plane

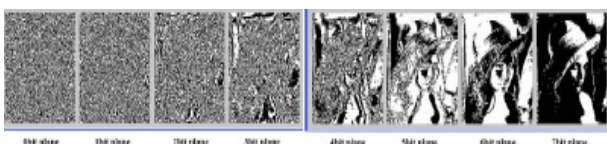


그림 5 bit plane 당 이미지의 중요도

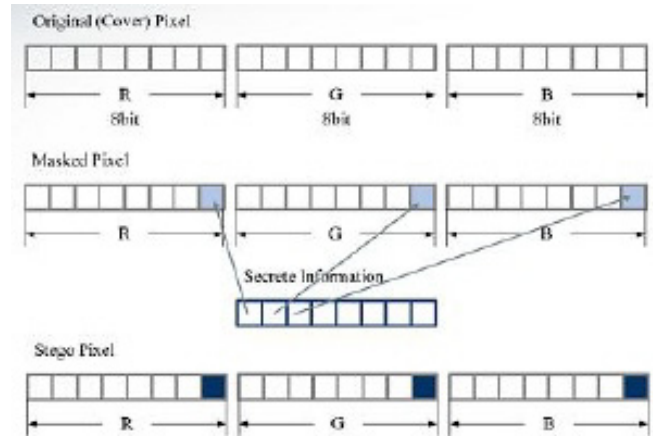


그림 1 LSB 정보은닉 방법

3. 모델과 구현 이미지 인코딩

전체 구현 절차는 Rajan[8]과 Emamm[9]의 알고리즘에 의하여 구현되었다.

step1 : text 결합

첫단계에서 두장의 동일사이즈의 텍스트 이미지는 하나의 이밋로 결합된다. 이 때 각 이미지의 픽셀은 쌍으로 번갈아 선택되게 된다.

```

for x=1:size(i1,1)
    for y=1:size(i1,2)
        if(mod((x+y),2)==0)
            alt(x,y,:)=i1(x,y,:);
        else
            alt(x,y,:)=i2(x,y,:);
        end
    end
end
step2
    
```

step2 color 이미지 텍스트 은닉
텍스트 이미지를 칼라이미지의 LSB 에 할당된다.

```

for i=1:size(org,1)
    for j=1:size(org,2)
        if text(i,j)<=128
            hide(i,j)=bitset(org(i,j),1,0);
        end
    end
end
    
```

```

else
    hide(i,j)=bitset(org(i,j),1);
end
end
end
    
```

step3 : decoding

역순으로 디코딩되며 과정은 지면상 생략한다.

image evaluation scale [ica][6]
전통적인 피어슨 상관계수에서부터

$$r_p = \frac{\sum((R_i - R_{avg})(G_i - G_{avg}))}{\sqrt{\sum(R_i - R_{avg})^2 \sum(G_i - G_{avg})^2}}$$

Ravg and Gavg the averages of the Red and Green channel, i 는 화소

(Pearson correlation coefficient)

$$r_o = \frac{\sum(R_i G_i)}{\sqrt{\sum R_i^2 \sum G_i^2}}$$

Mander's overlap coefficient

$$k_1 = \frac{\sum(R_i G_i)}{\sum R_i^2}$$

$$k_2 = \frac{\sum(R_i G_i)}{\sum G_i^2}$$

Manders' coefficients(channel 당 계수)

$$M_1 = \frac{\sum R_{i,color}}{\sum R_i}$$

$$M_2 = \frac{\sum G_{i,color}}{\sum G_i}$$

4. 결과

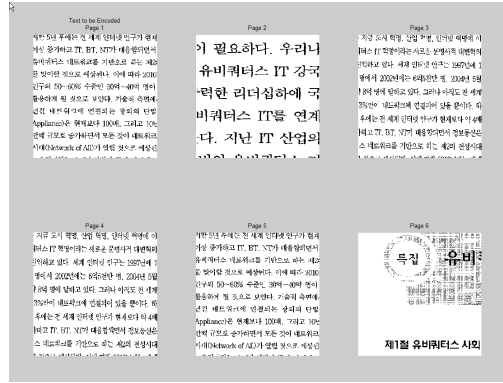


그림6 처리전 text

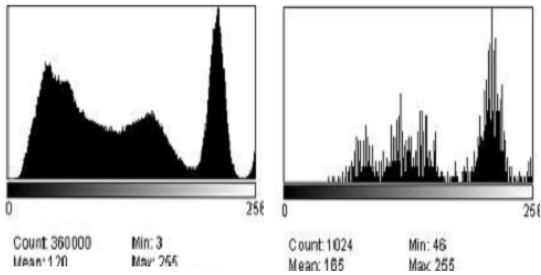


그림8 원이미지와 텍스트 포함 이미지의 히스토그램 변



그림 7 원 이미지와 6 장의 텍스트를 포함한 이미지

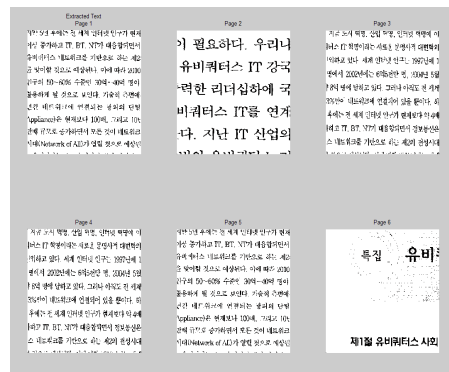


그림9 처리후 추출된 텍스트

Image	Rr	R	chl:ch2	M1	M2	N+ve	Ntotal	ICQ	Ch1 Thresh	Ch2 Thresh
1.bmp and tx1.bmp	1.000	1.000	1.006	0.980	0.984	-	322,791	- 0.5	1: 255	1: 255
2.bmp and tx2.bmp	1.000	1.000	1.000	0.991	0.991	-	324,030	- 0.5	1: 255	1: 255
3.bmp and tx3.bmp	1.000	1.000	1.008	0.981	0.986	-	317,917	0.5	1: 255	1: 255
4.bmp and tx4.bmp	1.000	1.000	1.007	0.982	0.986	-	322,340	- 0.5	1: 255	1: 255
5.bmp and tx5.bmp	1.000	1.000	1.006	0.980	0.984	-	322,791	- 0.5	1: 255	1: 255
6.bmp and tx6.bmp	1.000	1.000	1.007	0.970	0.975	-	337,897	- 0.5	1: 255	1: 255

표1 피어슨 상관계수등 코딩과 복구후 각텍스간의 일치도 통계계수들

5 결론

이미지 혹은 문서 은닉기술은 문서 영상, 오디오 등 모든 종류의 디지털 데이터에 대하여 연구가 진행되고 있다 이들은 다양한 목적과 용도로 이용되고 있다. 본 연구에서는 낮은 수준의 보안을 요하는 텍스트를 대용량으로 은닉하여 전달할 수 있는 단순하고 단순한 기법을 구현하였다. 먼저 텍스트 이미지를 결합하고 이를 24 비트 심도의 칼라이미지에 LSB를 이용하여 인코딩하고 이미지에 인코딩하여 복구하고 그 결과를 상관기법을 이용하여 분석하였으며 텍스트 이미지의 손실율이 미미한 것으로 판명되었다.

참고문헌

- [1] Shawn D. Dickman, An Overview of Steganography, James Madison University Infosec Techreport Department of Computer, Science ,JMU-INFOSEC-TR-2007-002, July 2007
- [2]Kefa Rabah, Steganography-The Art of Hiding Data, Information Technology Journal 3 (3): 245-269, 2004
- © 2004 Asian Network for Scientific Information
- [3]Sabu M Thampi, Information Hiding Techniques: A Tutorial Review, ISTE-STTP on

- Network Security & Cryptography, LBSCE 2004
- [4]Gleen, Watt, Steganography, 2006 CTA Conference.
- [5] Leo, Lee, LSB Steganography:Information Within Information, Computer Science 265, Section 2 Professor Stamp April 5, 2004
- [6] Analysis of protein co-localization using wide-field fluorescence microscopy and image-restoration for co-visualisation of CFP and YFP conjugated signalling proteins inside living cells. J. Weitzman, R. Lizundia, B. Blumen, M. Marchand, S. Shorte. http://www.pasteur.fr/recherche/unites/Pfid/html/un_coloc/?en
- [7]Zifu, Wang, How to quantify the colocalization in your images?, Optical Biology Core, Developmental Biology Center, University of California, Irvine 2006.
- [8]A. Cherma Rajan, Hide The Six Text Pages In A Color Image - <http://www.mathworks.com/matlabcentral/files/13698/content/html/main.html>
- [9]Nameer N. EL-Emam,Hiding a Large Amount of Data with High Security Using Steganography Algorithm , Journal of Computer Science 3 (4): pp. 223-232, 2007.

영문초록

These people where generally the image or the document nik technique silver document image, against the digital data of audio back all type the research is advanced being used with objective and the use which are various, is a d. Needs a low-end leveling instrument security text from the research which it sees and with substitution quantity the silver nik being simple it will be able to deliver the technique which is simple it embodied. It combined the text image first and the nose which is in the collar image of 24 bit depth which will reach ting it did and it rehabilitatedded and a higher officer technique and the result it used that the loss ratio of the text image to analyze is slight it was ascertained.