

영상의 특성을 효과적으로 이용하고 CZP 문제를 해결하여 영상에 가역적으로 데이터를 은닉하는 기법에 대한 연구

정 수 목*

A Study on reversible data hiding using the characteristics of image and solving CZP problem

Soo-Mok Jung*

요 약 본 논문에서는 영상에 존재하는 곡면 특성과 지역적 유사성을 효과적으로 이용하고, 극히 일부 영상에서 발생하는 CZP(Closest Zero point)가 존재하지 않는 문제를 해결하여 히스토그램 이동을 통해 가역적으로 기밀데이터를 커버 이미지(cover image)에 은닉할 수 있는 기법을 제안하였다. 제안된 기법을 적용하면 커버 이미지에 비가시적으로 데이터를 은닉할 수 있으며, 데이터가 은닉된 스테고 이미지(stego-image)로부터 기밀 데이터를 손실 없이 추출하고, 원본 커버 이미지를 온전히 복원할 수 있다. 제안된 기법을 적용하여 구성한 스테고 이미지는 커버 이미지와 차이를 구별할 수 없을 정도로 시각적 화질이 우수하기 때문에 기밀 데이터가 스테고 이미지에 은닉되었는지를 알 수 없다. 제안된 기법은 기존의 APD(Adjacent Pixel Difference)기법에 비하여 다양한 레벨로 기밀 데이터를 은닉할 수 있으며, APD 기법 대비 최대 25.1% 많은 기밀 데이터를 커버 이미지에 은닉 할 수 있었다.

Abstract In this paper, we have effectively used the surface characteristics and local similarity existing in image, solved the problem that there is no CZP(Closest Zero point) that occurs in a very few images to hide secrete data into cover image by using histogram shift. By applying the proposed technique, it is possible to hide secrete data invisibly into the cover image, extract secrete data from the stego-image with no data loss, and completely restore the original cover image. It is impossible to know whether the secrete data is hidden in the stego-image because the stego-image constructed by applying the proposed technique has a good visual quality that can not distinguish the difference from the cover image. The proposed method is able to hide secrete data at various levels compared to conventional APD(Adjacent Pixel Difference) technique, and hide secrete data up to 25.1% more than APD in cover image.

Key Words : Cover Image, Curved Surface Characteristics, Reversible Data Hiding, Stego-image, Regional Similarity

1. 서론

영상(image)에 기밀 데이터를 숨기기 위하여 데이터 은닉(data hiding) 기법이 사용되고 있다. 데이터 은닉 기법을 사용하여 커버 이미지(cover image)에 기밀 데이터를 삽입한 스테고 이미지(stego-image)

로부터 기밀 데이터를 추출하고 커버 이미지를 복원하게 된다. 데이터 은닉기법에서는 스테고 이미지에 숨겨져 있는 기밀 데이터를 사람이 인지 할 수 없어 야한다. 따라서 데이터 은닉 기법에서는 비인지성(imperceptibility)이 매우 중요하다.[1][2]

데이터를 커버 이미지에 은닉한 후 생성되는 스테

This paper was supported by the Sahmyook University Research Fund in 2018

*Division of Computer Science & Engineering, Sahmyook University

Received February 05, 2019

Revised February 12, 2019

Accepted February 13, 2019

고 이미지의 화질을 우수하게 하여 커버 이미지와 스테고 이미지간의 차이를 인지 할 수 없도록 하면 비인식성을 만족시킬 수 있다. 스테고 이미지의 화질을 높게 유지하기 위하여 개발된 대부분의 데이터 은닉 기법들은 스테고 이미지로부터 기밀 데이터를 추출하고 커버 이미지를 복원한 경우에 복원된 커버 이미지와 원본 커버 이미지가 불일치하는 비가역적인 특징을 갖는 단점이 있다.[3] 스테고 이미지로부터 기밀 데이터를 추출한 후에 얻게 되는 복원된 커버 이미지가 원본 커버 이미지와 완전히 일치하는 데이터 은닉 기법을 가역적(reversible) 데이터 은닉 기법이라고 한다. 이러한 가역적 데이터 은닉 기법은 보안, 군사, 소유권 보호 등의 응용분야에 매우 중요한 기술이다.[4] 그동안 가역적 데이터 은닉 기법들이 제안되어 왔다.[2-9] N_i 등은 이미지의 픽셀 값들에 대한 히스토그램(histogram)을 이동시켜 데이터를 은닉하는 기법(NSAS)을 제안하였다.[2] NSAS기법의 단점을 보완하기 위하여 APD (Adjacent Pixel Difference)기법이 제안되었다.[3]

APD기법에서 은닉되는 데이터 비트수가 차분 시퀀스의 히스토그램의 피크 포인트에서의 픽셀수로 제한되고, 차분 시퀀스의 히스토그램에서 zero point가 없는 경우에는 적용이 불가능한 문제가 있다. 이러한 단점들을 보완하기 위하여, 본 논문에서는 기존의 기법[10]을 개선한 가역 데이터 은닉 기법을 제안하였다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 연구인 APD기법에 대하여 기술하였고, 3장에서는 제안된 가역 데이터 은닉기법에 대하여 기술하였으며 4장에서는 실험결과와 분석결과를 기술하였다. 마지막으로 5장에서 결론을 맺는다.

2. APD(Adjacent Pixel Difference)

NSAS의 단점을 개선하기 위하여 Li 등이 APD 기법을 제안하였다.[3] APD 기법은 인접한 픽셀 값들이 대부분 비슷한 값을 갖는 특성을 효과적으로 이용하여 커버 이미지에 은닉되는 기밀 데이터 비트수를 크게 증가시키는 효과적인 기법이다. APD기법에

서는 커버 이미지를 좌측에서 우측으로, 위에서 아래 쪽으로 스캔하는 역 S-순(inverse s-order)으로 커버 이미지를 스캔하여 커버 이미지 시퀀스 C를 구성한다. 식 (1)을 C에 적용하여 차분 시퀀스 D를 구성한다. 자연 영상에서는 인접한 픽셀 값들이 서로 유사하기 때문에, 차분 시퀀스 D의 요소들은 0과 매우 작은 값을 주로 가지게 된다. 따라서 차분 시퀀스 D에 대한 히스토그램은 0과 0에 가까운 값들에서의 빈도수가 크게 증가하게 된다. (1)에서 n은 영상의 폭과 영상의 높이를 곱한 값이다.

$$D_i = \begin{cases} C_i & \text{if } i=0 \\ C_{i-1} - C_i & \text{if } 1 \leq i \leq n-1 \end{cases} \quad (1)$$

Lenna 영상을 커버 이미지로 사용하여 생성된 차분 시퀀스 D에 대한 히스토그램은 그림 1과 같다. 그림 1에서 보는 바와 같이 피크 포인트에서의 빈도수가 크게 증가하여 은닉 가능한 기밀 데이터의 비트수는 48,424비트($(h(PP_1)+h(PP_2))$)이다. 이는 피크 포인트에 해당하는 픽셀들에 기밀 데이터를 비트 단위로 은닉시키기 때문이다. 따라서 APD 기법은 NSAS 기법에 비하여 8.5배 더 많은 기밀 데이터를 은닉할 수 있다. APD기법에서 기밀 데이터를 은닉하여 스테고 이미지를 생성하는 절차는 그림 2와 같고, 스테고 이미지로부터 기밀 데이터와 원본 커버 이미지를 추출하는 절차는 그림 3과 같다. APD기법의 절차를 따라 커버 이미지에 기밀 데이터를 은닉하여 스테고 이미지를 생성하면, 스테고 이미지로부터 기밀 데이터와 원본 커버 이미지를 손실 없이 복원할 수 있다.

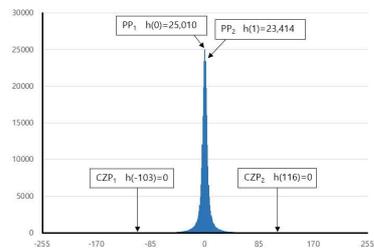


그림 1. D에 대한 히스토그램

Fig. 1. Histogram of sequence D

단계 1. 커버 이미지를 역 S-순으로 스캔하여 픽셀 값들로 구성되는 커버 이미지 시퀀스(C) 구성.

단계 2. 식 (1)에 따라 인접 픽셀 값과의 차분을 갖는 차분 시퀀스(D) 생성.

단계 3. D에 대한 히스토그램을 생성하여 PP₁, CZP₁, PP₂, CZP₂ 결정.

단계 4. 식 (2), (3)을 D에 적용하여 쉬프트된 차분 시퀀스(DS) 생성.

$$DS_i = \begin{cases} D_i & \text{if } i=0 \text{ or } D_i \notin [PP_j + sd_j, CZP_j] \\ D_i + sd_j & \text{if } D_i \in [PP_j + sd_j, CZP_j] \end{cases} \quad (2)$$

$$sd_j = \begin{cases} 1 & \text{if } PP_j < CZP_j \\ -1 & \text{if } CZP_j < PP_j \end{cases} \quad \text{where } j \in \{1,2\} \quad (3)$$

단계 5. 식 (4)를 DS에 적용하여 기밀 데이터가 삽입(embedding)된 은닉 시퀀스(DE) 생성.

$$DE_i = \begin{cases} DS_i & \text{if } i=0 \text{ or } DS_i \neq PP_j \text{ or } \text{data}=0 \\ DS_i + sd_j & \text{if } DS_i = PP_j \text{ and } \text{data}=1 \end{cases} \quad (4)$$

단계 6. 식 (5)를 적용하여 스테고 이미지 시퀀스(S) 생성. n의 값은 영상의 폭과 영상의 높이를 곱한 값이다.

$$S_i = \begin{cases} DE_i & \text{if } i=0 \\ C_{i-1} - DE_i & \text{if } 1 \leq i \leq n-1 \end{cases} \quad (5)$$

단계 7. S를 역 S-순으로 구성하여 스테고 이미지 생성.

그림 2. APD 기법에서의 은닉 절차
Fig. 2. The embedding process of APD

단계 1. 스테고 이미지를 역 S-순으로 scan하여 은닉 이미지 시퀀스 (S) 구성.

단계 2. i 값을 증가 시키면서 단계2.1~단계2.2를

반복 수행하여 커버 이미지 시퀀스(C)와 은닉 시퀀스(DE) 복원.

단계 2.1. 식 (6)을 만족하는 은닉시퀀스 (DE) 복원.

$$DE_i = \begin{cases} C_i & \text{if } i=0 \\ C_{i-1} - S_i & \text{otherwise} \end{cases} \quad (6)$$

단계 2.2. 식 (7)을 만족하는 커버 이미지 시퀀스 (C) 복원.

$$C_i = \begin{cases} S_i & \text{if } i=0 \\ S_i + sd_j & \text{else if } 1 \leq i \leq n-1 \text{ and } \\ & C_{i-1} - S_i \in [PP_j + sd_j, CZP_j] \\ S_i & \text{otherwise} \end{cases} \quad (7)$$

단계 3. 식 (8)을 DE의 각 요소에 적용하여 기밀 데이터 추출.

$$\text{Extraction bit} = \begin{cases} 0 & \text{if } DE_i = PP_j \\ 1 & \text{else if } DE_i = PP_j + sd_j \end{cases} \quad (8)$$

단계 4. C를 스캔하여 역 S-순으로 원본 커버 이미지 구성.

그림 3. APD기법에서의 추출 절차
Fig. 3. The extraction process of APD

3. 제안기법

APD 기법에서 생성된 차분시퀀스 D의 값들은 -255 ~ +255사이의 정수 값들을 갖는다. APD기법에서는 차분 시퀀스 D에 대한 히스토그램을 구성한 후 PP₁, PP₂에 가장 가까운 곳에 위치하는 차분의 빈도수가 0인 지점을 찾아 그 값을 CZP₁, CZP₂로 설정한다.

대부분의 영상의 차분 시퀀스에 대한 히스토그램에서 CZP가 존재한다. 그러나 인위적으로 제작된, 픽셀 값의 변화가 다양한 초대형의 영상의 극히 일부에서는 차분 시퀀스에 대한 히스토그램에서의 빈도수가 0이 되는 곳이 없는 경우가 있어서 CZP가 존재하지 않기 때문에 APD 기법과 같은 히스토그램 이동을 통한 데이터 은닉 기법을 적용할 수 없는 문제

가 발생한다. 이러한 문제를 해결하기 위하여, 본 논문에서는 CZP가 존재하지 않는 경우에도 히스토그램 이동을 통한 기밀 데이터 은닉을 가능하게 하는 기법을 제시한다. 제시된 기법은 아래의 4 단계의 절차를 수행하여 CZP 문제를 해결한다.

CZP 문제 해결 단계1: PP₁과 PP₂ 중에서 작은 값을 PP_L로 두고 큰 값을 PP_R로 둔다.

-255 ~ PP_L 사이의 값들에서 히스토그램의 빈도수(h[-255] ~ h[PP_L])가 가장 작은 값을 찾아 CZP_L로 둔 다음, 차분 시퀀스에서 CZP_L 값을 갖는 위치 i값들을 저장한다.

PP_R ~ 255사이의 값에서 히스토그램의 빈도수(h[PP_R] ~ h[255])가 가장 작은 값을 찾아 CZP_R로 둔 다음, 차분 시퀀스에서 CZP_R 값을 갖는 위치 j 값들을 저장한다.

CZP 문제 해결 단계2: 차분 시퀀스가 CZP_L, CZP_R 값을 갖는 위치 i, j에 대한 커버 이미지의 화면 좌표 계상의 y좌표와 x좌표를 식 (9)와 식(10)을 적용하여 찾는다. 식 (9)에서의 연산은 정수 나눗셈이다.

$$y=i/img_Heigt \tag{9}$$

$$\begin{aligned} \text{if}(y\%2==0) \quad x=i-y*img_Width \\ \text{else} \quad x=y*img_Width+(img_Width-1)-i \end{aligned} \tag{10}$$

512x512영상의 경우, i와 j는 0~262,143 사이의 값을 갖고, y와 x는 0~511 사이의 값을 갖는다.

CZP 문제 해결 단계3: 차분 시퀀스의 위치 i에 대응되는 화면 좌표계상의 y, x 좌표에 해당되는 커버 이미지의 픽셀 값을 1만큼 감소시키고, 차분 시퀀스의 위치 j에 대응되는 y, x 좌표상의 커버 이미지의 픽셀 값은 1만큼 증가시킨다. 이와 같이 하는 이유는 차분 시퀀스의 빈도수가 많은 쪽으로 커버 이미지의 값들을 1만큼 변경하여 전체 이미지에서의 밝기 값을 거의 동일하게 유지하기 위함이다.

CZP 문제 해결 단계4: 변경된 커버 이미지의 픽셀 값을 반영하여 차분 시퀀스를 생성한 후, PP₁과 PP₂ 를 구하고, CZP₁과 CZP₂를 조사한다.

하나의 CZP라도 존재하지 않으면 단계1로 넘어가서 위의 과정을 반복한다.

히스토그램에서 빈도수가 최소가 되는 픽셀 값을 갖는 픽셀들의 수는 초대형 영상인 경우에도 지극히 적기 때문에, CZP 문제를 해결하기 위하여 원본 이미지의 픽셀 값을 1만큼 변화 시켜도 커버 이미지의 화질은 거의 동일하게 유지된다.

일반적인 영상에는 지역적 유사성(regional similarity)이 존재하고, 영상의 각 부분에는 곡면 특성(curved surface characteristics)이 있다. 제안된 기법에서는 이러한 지역적 유사성과 곡면 특성을 고려하여 영상의 픽셀 값을 정밀하게 예측하여 예측 이미지를 생성하고, 예측 이미지를 데이터 은닉에 사용한다. 예측 이미지를 생성하기 위하여 먼저 커버 이미지의 픽셀들을 역 S-순으로 스캔하면서 각 위치에서 예측 가능 여부를 체크한다. 예측 불가능한 위치인 경우에는 커버 이미지의 픽셀 값을 예측 이미지의 픽셀 값으로 복사한다. 픽셀 값 예측 가능 위치인 경우에는 제안된 절차를 따라 예측 값을 구한 다음 이를 예측 이미지의 픽셀 값으로 한다.

그림 4에서 보는 바와 같이 커버 이미지를 역 S-순으로 스캔한 12개의 픽셀 값들을 사용하여 지역적 유사성이 높을 것으로 예상되는 위치를 선정한다. 선정된 위치에서의 곡면 특성을 조사하여 곡면 특성에 따라 픽셀 값을 정밀하게 예측한다. 그림 4에서 보는 바와 같이, 커버 이미지의 상위 2줄과 좌측과 우측의 각 2줄에 해당되는 영역은 픽셀 값 예측 불가능영역이고, 그 이외의 영역은 픽셀 값 예측 가능영역이 된다. 그림 4 표시된 화살표는 역 S-순으로 스캔하는 스캔 방향을 나타낸다. 따라서 커버 이미지의 화면 좌표계의 y값이 0, 2, 4, ... 등 인 경우에는 나머지 연산인 y%2의 값이 0이 되며 이러한 행들은 좌측에서 우측으로 스캔한다. y값이 1, 3, 5, 등 인 경우에는 y%2의 값이 1이 되는데 이러한 행들에 대하여는 우측에서 좌측으로 스캔을 하면 커버 이미지를 역 S-순으로 스캔할 수 있다.

→	A	B	C	D	E
←	J	I	H	G	F
→	K	L	예측 위치		

그림 4. y%2=1인 행에서의 예측 픽셀 값 예측
Fig 4. Pixel value prediction at the row with y%2=1

예측 가능위치에서의 픽셀 값 예측 과정은 아래의 6단계로 구성된다.

픽셀 값 예측 단계 1: 그림 4(a)의 예측 위치에서, 임시 예측 픽셀 값(V)을 식 (9)와 같이 계산한다. 식 (9)에서 보는 바와 같이 V는 이미 스캔된 주위의 픽셀들을 사용하여 계산되고, 위치별 가중치를 고려하여 계산된다. 복원 시 스테고 이미지를 역 S-순으로 스캔한 후 기밀 데이터와 커버 이미지를 복원하기 위하여 은닉 과정에서 이미 스캔된 픽셀들만을 사용하여 셀 값을 예측한다. 식 (9)에서, 픽셀 값이 L, I, H, G인 픽셀의 좌표는 각각 (x_{i-1}, y_j) , (x_{i-1}, y_{j-1}) , (x_i, y_{j-1}) , (x_{i+1}, y_{j-1}) 이다. 식 (9)에서 보는 바와 같이, 예측 위치와 인접해 있는 픽셀들(픽셀 값이 L, I, H, G인 픽셀)이 임시 예측 값에 미치는 영향을 나타내는 영향의 가중치를 1로 하였고, 그 이외의 픽셀들이 V에 미치는 영향의 가중치를 α 로 표시하였다.

$$V(x_i, y_j) = [P(x_{i+1}, y_{j-1}) + P(x_i, y_{j-1}) + P(x_{i-1}, y_{j-1}) + P(x_{i-1}, y_j) + P(x_{i-2}, y_{j-2}) + P(x_{i-1}, y_{j-2}) + P(x_i, y_{j-2}) + P(x_{i+1}, y_{j-2}) + P(x_{i+2}, y_{j-2}) + P(x_{i+2}, y_{j-1}) + P(x_{i-2}, y_{j-1}) + P(x_{i-2}, y_j)] * \alpha / (4 + 8 * \alpha) \quad (9)$$

픽셀 값 예측 단계 2: 지역적 유사성의 정도를 판단하는 단계이다. 임시 예측 픽셀 값(V)과 주변 픽셀 값들과의 편차(dev)를 식 (10)과 같이 구한 다음 임계값(β)과 비교한다.

$$dev(x_i, y_j) = \text{abs}(V - P(x_{i+1}, y_{j-1})) + \text{abs}(V - P(x_i, y_{j-1})) + \text{abs}(V - P(x_{i-1}, y_{j-1})) + \text{abs}(V - P(x_{i-1}, y_j)) + \text{abs}(V - P(x_{i-2}, y_{j-2})) + \text{abs}(V - P(x_{i-1}, y_{j-2})) + \text{abs}(V - P(x_{i+1}, y_{j-2})) + \text{abs}(V - P(x_{i+2}, y_{j-2})) + \text{abs}(V - P(x_{i+2}, y_{j-1})) + \text{abs}(V - P(x_{i-2}, y_{j-1})) + \text{abs}(V - P(x_{i-2}, y_j)) * \alpha \quad (10)$$

픽셀 값 예측 단계 3: 편차(dev)가 임계값(β)을 초과할 경우에는 해당 예측 위치에서의 영상은 지역적 유사성이 낮은 영역으로 판단하여 픽셀 값 예측 과정을 더 이상 수행하지 않고 커버 이미지의 픽셀

값을 예측 이미지의 픽셀 값으로 설정 한다. 계산된 편차(dev)가 임계값(β) 미만일 경우에는 지역적 유사성이 높은 것으로 판단하여 픽셀 값 예측 단계 4를 수행한다.

픽셀 값 예측 단계 4: 해당 위치에서의 영상의 곡면 특성을 예측하는 단계이다. 그림 4(a)의 예측 위치에서, 인접한 4개의 픽셀들을 사용하여 곡면 특성을 추정하는 과정은 다음과 같다. G가 H보다 크고, H가 I보다 크고, H가 L보다 큰 경우에는 해당 예측 위치에서의 곡면 특성은 왼쪽보다 오른쪽의 픽셀 값이 큰 값을 갖는 단순 경사면(simple inclined surface)의 특성을 갖는다고 판단한다. 또한 그 반대인 경우에는 해당 예측 위치에서의 곡면 특성은 오른쪽보다 왼쪽의 픽셀 값이 더 큰 값을 갖는 단순 경사면의 특성을 갖는다고 판단한다. 마지막으로 G, H, I, L의 값이 동일한 경우에는 해당 위치에서의 영상은 평면(flat surface)의 특성을 갖는다고 판단한다.

픽셀 값 예측 단계 5: 예측 위치에서, 영상의 곡면 특성에 따라 예측 픽셀 값을 계산하는 단계이다. 영상의 곡면 특성이 단순 경사면이거나 평면으로 예측된 경우에는 최종 예측 픽셀 값(V_i)을 식 (11)과 같이 계산하여 예측 이미지의 예측 픽셀 값으로 한다. 예측 위치에서의 곡면 특성이 평면도 아니고 단순 경사면도 아닌 경우에는 식 (9)에서 구해진 임시 예측 값(V)을 예측 위치에서의 예측 픽셀 값으로 하여 예측 이미지를 생성한다.

따라서 픽셀 값 예측 단계 1부터 픽셀 값 예측 단계 5의 과정을 따라 예측 값을 계산하면 정밀하게 예측 값을 계산할 수 있다.

$$V_i(x_i, y_j) = (G + H + I + L + (C + K) * \alpha) / (4 + 2 * \alpha) \quad (11)$$

제안된 기법의 데이터 은닉 절차는 그림 5와 같고, 제안된 기법에서의 기밀 데이터 추출과 원본 커버 이미지의 복원 절차는 그림 6과 같다.

단계 1. 예측 불가능영역(이미지의 상위 2개 행과 좌우측 각 2개 열)에서는 커버 이미지의 픽셀 값

을 예측 이미지에 그대로 사용. 예측 가능 영역에 서는단계1.1 ~ 단계1.5에 따라 예측 픽셀 값은 구 하여 예측 이미지를 생성.

단계 1.1. 식 (9)에 따라 임시 예측 픽셀 값(V_1)을 계산.

단계 1.2. 식 (10)에 따라 V_1 과 주변 픽셀 값들과의 편차(dev)를 계산.

단계 1.3. 편차가 임계값(θ)보다 크거나 같은 경우에는 커버 이미지 내 해당 위치에서의 픽셀 값을 복사하여 예측 이미지를 생성.

단계 1.4. 편차가 임계값(θ)보다 작으면서 경사면이거나 평면인 경우에는 식 (11)을 적용하여 구한 제 2 임시 예측 픽셀 값(V_2)을 예측 이미지의 픽셀값으로 함.

단계 1.5. 그 이외의 경우에는 제1 임시 예측 픽셀 값(V_1)을 예측 픽셀 값으로 하여 예측 이미지 생성.

단계 2. 커버 이미지와 예측 이미지의 픽셀 값을 역 S-순으로 스캔하여 픽셀 값들로 구성되는 커버 이미지 시퀀스(C)와 예측 이미지 시퀀스(P) 생성.

단계 3. 식 (12)를 C와 P에 적용하여 차분 시퀀스(D) 생성.

$$D_i = \begin{cases} C_i & \text{if } i = 0 \\ C_i - P_i & \text{else if 예측 픽셀 값 사용 (12)} \\ P_{i-1} - C_i & \text{otherwise (예측 픽셀 값 미사용)} \end{cases}$$

단계 4. PP_1, PP_2 를 구하고, ZP_1, CZP_2 를 찾는다. 만약 CZP 중에서 하나라도 없으면 단계4.1 ~ 단계4.4를 수행하고, CZP 가 모두 존재하면 단계5로 간다.

단계 4.1. CZP 문제 해결 단계1 수행

단계 4.2. CZP 문제 해결 단계2 수행

단계 4.3. CZP 문제 해결 단계3 수행

단계 4.4. CZP 문제 해결 단계4를 수행하여 차분

시퀀스 D에 대한 히스토그램을 생성하여 PP_1, CZP_1, PP_2, CZP_2 결정.

단계 5. 식 (2), (3)을 D에 적용하여, 쉬프트된 차분 시퀀스(DS) 생성.

단계 6. 식 (4)와 같이, DS에 기밀 데이터를 삽입하여 은닉 시퀀스(DE) 생성.

단계 7. 식 (5)를 C와 DE에 적용하여 스테고 이미지 시퀀스(S) 생성.

단계 8. S를 순차적으로 스캔하여 역 S-순으로 스테고 이미지 생성.

그림 5. 제안기법의 데이터 은닉 절차

Fig. 5. Data embedding process of the proposed algorithm

단계 1. 스테고 이미지의 픽셀 값을 역 S-순으로 스캔하여 스테고 이미지 시퀀스(S) 생성.

단계 2. $C_0=S_0, DE_0=C_0, i$ 값을 증가 시키면서, 픽셀 값 예측 가능 영역인지 판단한다. 픽셀 값 예측 가능 영역인 경우에는 그림 5의 단계1.1 ~ 단계1.5의 과정을 수행하여, 해당 위치에서의 픽셀 값이 커버 이미지의 픽셀 값인지 예측 픽셀 값을 판단하여 추출모드를 결정.

모드 1: Mode 2와 Mode 3를 제외한 전 영역에서의 추출 모드. APD 기법과 동일하게, 식 (6)과 식(7)을 수행하여 은닉 시퀀스(DE)와 커버 이미지 시퀀스(C) 복원.

모드 2: 현재 위치에서 예측 값이 사용된 경우에 해당. APD 기법과 동일하게 식 (6)을 사용하여 은닉 시퀀스(DE)를 복원. 식 (13)와 식 (14)를 사용하여 커버 이미지 시퀀스(C)를 복원. RP_i 는 추출 모드 판단결과가 Mode 2인 경우, i 번째 위치에서 계산되는 복원 예측 픽셀 값. 식 (13)에 사용된 연관 값 AV(Associated Value)는 식 (14)와 같이 계산되는 값으로, 복원 과정 중 추

출 모드가 Mode 2 혹은 Mode 3인 경우에 사용됨.

$$C_i = RP_i + AV_i \quad (13)$$

$$AV_i = \begin{cases} C_{i-1} - S_i - sd_j & \text{if } C_{i-1} - S_i \in [PP_j + sd_j, CZP_j] \\ C_{i-1} - S_i & \text{otherwise} \end{cases} \quad (14)$$

모드 3: 이전 위치에서 예측 픽셀 값이 사용되었으나, 현재 위치에서는 예측 픽셀 값이 사용되지 아니한 경우. 식APD 기법에서와 같이 식 (6)을 사용하여 은닉 시퀀스(DE)를 복원하고, 식 (15)를 사용하여 커버 이미지 시퀀스(C)를 복원.

$$C_i = RP_{i-1} - AV_i \quad (15)$$

단계 3. 식 (8)을 DE_i 에 적용하여 스테고 이미지에 삽입되어 있는 기밀 데이터를 추출.

단계 4. 복원된 커버 이미지 시퀀스(C)를 스캔하여 역 S-순(inverse s-order)으로 원본 커버 이미지 생성.

그림 6. 제안기법에서의 데이터 추출 절차
 Fig. 6. Data extraction process of the proposed algorithm

3. 실험결과

512x512 gray scale 영상인 Lenna, splash, girl을 커버 이미지로 사용하여 제안된 기법의 성능을 평가하였다. 영문 초록에 대한 ASCII코드를 기밀 데이터로 사용하였으며, 이를 반복적으로 커버 이미지에 삽입하였다. 가중치 값인 α 는 0.298로 하였고, 임계값(β)을 적용하여 실험을 수행한 경우를 L(β)로 표기하였다. 그림 7의 (a)는 실험에 사용된 커버 이미지들이고, (b)는 제안된 기법을 적용하여 커버 이미지에 기밀 데이터를 은닉하여 생성한 스테고 이미지들이다.

각 커버 이미지에 대하여 APD 기법과 제안 기법을 적용하여 기밀 데이터를 은닉한 결과가 표 1에 나타나 있다. 임계값(β)을 0으로 하면, 픽셀 예측을 수행하지 않기 때문에 APD기법과 동일하게 동작한

다.

표 1에서 보는 바와 같이 임계값(β)이 커지면 더 많은 위치에서 픽셀값 예측이 이루어지기 때문에 은닉되는 기밀 데이터수는 증가하나, 지역적 유사성이 낮은 경우에도 픽셀 값을 예측하기 때문에 예측의 정확도가 낮아지게 된다. 픽셀 값 예측이 정확할수록, 식 (13)에 따른 차분시퀀스의 D_i 값이 0에 가깝게 되어 차분 시퀀스에 대한 히스토그램의 피크 포인트에서의 빈도수가 크게 증가하게 된다. 따라서 커버 이미지에 은닉되는 기밀 데이터 비트수가 크게 증가한다. 표 1에서 보는 바와 같이 제안기법을 적용하여 픽셀 값을 예측하면, 평균 예측 오차가 매우 적기 때문에 D_i는 매우 작은 값을 갖게 되어, 차분 시퀀스 D에 대한 히스토그램을 생성하면 peak point에서의 빈도수가 크게 증가하게 된다. 따라서 커버 이미지에 은닉되는 기밀 데이터 비트수가 크게 증가한다.

임계값(β)을 0부터 36까지 증가시키면, 스테고 이미지의 PSNR 값은 감소하지만, 그림 7에서 보는 바와 같이 스테고 이미지의 시각적 화질이 매우 우수한 상태로 유지되고 있기 때문에 스테고 이미지에 기밀 데이터가 은닉되어 있는지를 인지할 수 없다. 제안 기법을 적용하는 경우, 은닉되는 기밀 데이터의 비트수가 APD 기법에 비하여 최대 25.1% 증가하였다.

제안된 기법을 적용하면, 모든 종류의 커버 이미지에 다양한 레벨로 기밀 데이터를 은닉 할 수 있으며, 스테고 이미지로부터 기밀 데이터를 손실 없이 추출할 수 있고, 원본 커버 이미지를 온전히 복원할 수 있다.



(a-1) cover image, Lenna



(a-2) APD=L(0)



(a-3) L(18)



(a-4) L(36)



(b-1) cover image, splash



(b-2) APD=L(0)



(b-3) L(18)



(b-4) L(36)



(c-1) cover image, girl



(c-2) APD=L(0)



(c-3) L(18)



(c-4) L(36)

그림 7. 커버 이미지와 스테고 이미지

Fig. 7. Cover images & stego-images

표 1. 제안된 기법에 대한 실험 결과

Table 1. The experimental results of the proposed algorithm

영상	기법	은닉 비트수	PSNR (dB)	예측 적중률 (%)	평균 예측 오차	은닉 비트수 증가율 (%)
Lenna	APD	48424	48.60	x	x	x
	L(0)	48424	48.60	x	x	x
	L(18)	57302	35.28	17.63	1.90	18.3
	L(36)	60595	31.93	15.46	2.37	25.1
splash	APD	68023	48.81	x	x	x
	L(0)	68023	48.81	x	x	x
	L(18)	73795	33.95	19.29	2.02	8.5
	L(36)	74929	31.67	16.80	2.43	10.2
girl	APD	60775	48.72	x	x	x
	L(0)	60775	48.72	x	x	x
	L(18)	64883	35.82	20.42	1.78	6.8
	L(36)	67479	32.13	17.59	2.29	11.0

3. 결론

본 논문에서는 커버 이미지에 대한 차분 시퀀스의 히스토그램에서 CZP를 포함하지 않는 극히 일부 영상에서의 CZP 문제를 해결하는 기법을 제시하였다. 최소한의 픽셀에 대하여 픽셀 값을 1만큼 변경하여 CZP 문제를 해결하기 때문에, CZP문제를 해결하도

록 구성 된 커버 이미지는 원본 커버 이미지와 거의 동일하게 유지된다.

제안된 기법을 적용하면, 커버 이미지에 기밀 데이터를 영상의 특성에 따라 효과적으로 은닉할 수 있고, 스테고 이미지로부터 기밀데이터를 손실 없이 추출할 수 있으며 원본 커버 이미지를 온전히 복원할 수 있다. 생성된 스테고 이미지의 시각적 화질이 우수하여 커버 이미지와의 차이를 인지 할 수 없기 때문에 스테고 이미지에 기밀데이터가 은닉되어 있는지를 식별 할 수 없다. 제안 기법을 적용하여 기밀 데이터를 은닉하면, 기존 APD기법 대비 최대 25.1% 증가한 기밀 데이터를 은닉 할 수 있었다.

REFERENCES

[1] H. C. Huang, C. M. Chu, and J. S. Pan, "The optimized copyright protection system with genetic watermarking," *Soft Computing*, Vol. 13, No. 4, pp. 333-343, Feb. 2009.

[2] Z. Ni, Y. Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 16, No. 3, pp. 354-362, March 2006.

[3] Y. C. Li, C. M. Yeh, and C. C. Chang, "Data hiding based on the similarity between neighboring pixels with reversibility," *Digital Signal Processing*, Vol. 20, No. 4, pp. 1116-1128, July 2010.

[4] C. L. Tsai, K. C. Fan, C. D. Chung, and T. C. Chuang, "Reversible and lossless data hiding with application in digital library," *Proc. 38th Annual 2004 Int. Canahan Conf. on Security Technology*, pp. 226-232, Albuquerque, New Mexico, USA, Oct. 2004.

[5] M. U. Celik, G. Sharma, A. M. Tekalp, E. Saber, "Reversible data hiding," *Proc. 2002 Intl. Conf. on Image Processing*, Vol 3, pp. 157-160, Rochester, New York, USA, Sep. 2002.

[6] C. C. Chang, W. L. Tai, and C. C. Lin, "A reversible data hiding scheme based on side match vector quantization," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 16, No. 10, pp. 1301-1308, Oct. 2006.

[7] L. Kamstra, H.J.A.M. Heijmans, "Reversible data embedding into images using wavelet techniques and sorting," *IEEE Trans. on Image Process*, Vol. 14, No. 12, pp. 2082-2090, Dec. 2005.

[8] H. W. Tseng and C. P. Hsieh, "Reversible data hiding based on image histogram modification," *Imaging Science Journal*, Vol. 56, No. 5, pp. 271-278, Oct. 2008.

[9] X. Li, B. Li, B. Yang, and T. Zeng, "General Framework to Histogram shifting based Reversible Data Hiding," *IEEE Trans. on Image Processing*, Vol. 22, No. 6, pp. 2181-2191, Feb. 2013.

[10] Soo-Mok Jung, "Reversible Data Embedding Algorithm using the Locality of Image and the Adjacent Pixel Difference Sequence," *The Journal of Korea Institute of Information, Electronics, and Communication Technology*, Vol. 9, No.6, pp.573-577, Dec. 2016.

저자약력

정 수 목(Soo-Mok Jung)

[종신회원]



1984 경북대학교 전자공학과 공학사
 1986 경북대학교 대학원 공학석사
 2002 고려대학교 대학원 컴퓨터학과
 이학박사
 현 재: 삼육대학교 컴퓨터·메카트로닉스
 공학부 교수
 관심분야: 멀티미디어, 영상처리
 E-mail: jungsm@syu.ac.kr