

Reversible data hiding algorithm using spatial locality and the surface characteristics of image

Soo-Mok Jung*, Byung-Won On**

Abstract

In this paper, we propose a very efficient reversible data hiding algorithm using spatial locality and the surface characteristics of image. Spatial locality and a variety of surface characteristics are present in natural images. So, it is possible to precisely predict the pixel value using the locality and surface characteristics of image. Therefore, the frequency is increased significantly at the peak point of the difference histogram using the precisely predicted pixel values. Thus, it is possible to increase the amount of data to be embedded in image using the spatial locality and surface characteristics of image. By using the proposed reversible data hiding algorithm, visually high quality stego-image can be generated, the embedded data and the original cover image can be extracted without distortion from the stego-image, and the embedding data are much greater than that of the previous algorithm. The experimental results show the superiority of the proposed algorithm.

▶ Keyword : cover image, stego-image, reversible data hiding, locality, surface characteristics

I. Introduction

데이터 은닉(data hiding) 기법은 이미지나 동영상과 같은 cover media에 기밀 데이터를 삽입(embedding)하는 중요한 기법이다. 기밀 데이터를 커버 이미지(cover image)에 삽입하여 생성된 스테고 이미지(stego-image)로부터 기밀 데이터를 추출(extraction)하게 된다. 스테고 이미지에 은닉된 기밀 데이터를 사람이 인지할 수 없도록 하기 위하여 데이터 은닉기법에서는 비인지성이 요구된다[1][2].

비인지성 요구 조건(imperceptibility requirement)을 만족시키기 위해서는 스테고 이미지의 화질(quality)이 우수해야 한다. 스테고 이미지의 화질을 높이기 위하여 제안된 대부분의 데이터 은닉 기법들은 스테고 이미지로부터 기밀 데이터를 추출한 후에 얻게 되는 복원된 커버 이미지에 왜곡이 발생하여 원본 커버 이미지와 일치 하지 않는다[3].

스테고 이미지로부터 기밀 데이터를 추출한 후에 얻게 되는

복원된 커버 이미지가 원본 커버 이미지와 완전히 일치하는 가역적 인(reversible) 데이터 은닉 기법은 의학, 군사, 디지털 라이브러리(digital library) 등의 응용에서 매우 중요하다[4].

최근에 다양한 가역적 데이터 은닉 기법들이 제안되어 왔다 [2-10]. Ni 등은 이미지의 히스토그램(histogram)을 사용하는 가역적인 데이터 은닉기법(NSAS)을 제안하였다[2]. NSAS기법은 커버 이미지에 대한 히스토그램에서 peak point와 zero point 쌍(pair)을 조사하여 (peak point, zero point) 사이에 있는 픽셀들을 쉬프트(shift) 시킨 후, peak point에 해당하는 픽셀들에 기밀 데이터를 삽입한다. 따라서 삽입되는 최대 데이터 비트수는 커버 이미지에 대한 히스토그램의 peak point에서의 픽셀 수로 제한되는 한계가 있다.

Li 등은 NSAS기법을 개선한 APD(Adjacent Pixel Difference) 기법을 제안하였다[3]. APD기법에서는 커버 이미지에 삽입되는

• First Author: Soo-Mok Jung, Corresponding Author: Soo-Mok Jung
*Soo-Mok Jung(jungsm@syu.ac.kr), Division of Computer Science & Engineering, Sahmyook University
**Byung-Won On(bwon@kunsan.ac.kr), Department of Statistics & Computer Science, Kunsan National University
• Received: 2016. 04. 11, Revised: 2016. 04. 21, Accepted: 2016. 04. 28.
• This research was supported by Sahmyook University Research Fund, 2015

기밀 데이터 비트수를 증가시키기 위하여 커버 이미지로부터 인접 픽셀 간의 픽셀 값의 차(difference)로 구성되는 차분 시퀀스(difference sequence)를 생성한 후, 차분 시퀀스에 대하여 히스토그램을 구한다. 인접한 픽셀 값들은 유사성(similarity)을 갖기 때문에 차분 시퀀스를 구성하는 값(difference value)들은 매우 작은 값들을 갖게 되며, 생성된 히스토그램의 peak point에서의 빈도수가 매우 크게 되어 커버 이미지에 삽입할 수 있는 기밀 데이터 비트수가 크게 증가하게 된다. APD기법에서는 커버 이미지에 삽입되는 기밀 데이터 비트수가 차분 시퀀스에 대한 히스토그램의 peak point에서의 빈도수로 제한되는 단점이 있다.

APD기법의 삽입 데이터 비트수를 증가시키기 위하여, 본 논문에서는 기존의 기법[11]을 개선하여 자연 영상에 존재하는 지역성(locality)과 영상의 표면 특성(surface characteristics)을 효과적으로 활용하는 효율적인 가역 데이터 은닉 기법을 제안하였다. 제안된 기법에서는 두 단계를 거쳐 픽셀 값을 예측한다. 첫 번째 단계는 인접 픽셀 값들을 사용하여 지역성이 높을 것으로 예상되는 위치를 선정한다. 두 번째 단계는 지역성이 높을 것으로 예상되는 위치에서 영상의 표면 특성을 조사한 후, 표면 특성을 반영하여 픽셀 값을 정밀하게 예측한다.

제안된 기법을 사용하여 픽셀 값을 예측하고, 예측된 픽셀 값을 사용하여 차분 시퀀스를 구성하면 히스토그램의 peak point에서의 빈도수가 크게 증가한다. 픽셀 값 예측이 정확할수록 픽셀 값의 차로 구성되는 차분 시퀀스에 대한 히스토그램의 peak point에서의 빈도수가 증가하게 되어 커버 이미지에 삽입되는 기밀 데이터 비트수가 더욱 증가하게 된다. 제안된 기법을 사용하면 커버 이미지에 삽입되는 기밀 데이터 비트수가 APD기법에 비하여 크게 증가된다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 APD기법에 대하여 기술하였고, 3장에서는 제안된 가역 데이터 은닉기법에 대하여 기술하였다. 4장에서는 실험결과를 기술한 후, 이를 분석하였고, 5장에서 결론을 맺는다.

II. Adjacent Pixel Difference(APD)

512x512 크기를 갖는 gray scale 영상인 Lenna영상에 대한 히스토그램이 Fig. 1에 나타나있다. 히스토그램에서 가장 큰 빈도수, 두 번째로 큰 빈도수를 갖는 픽셀 값을 각각 PP₁(peak point 1), PP₂, 라고 둔다. PP₁, PP₂에 가장 가까이 위치한 빈도수가 0인 값을 갖는 픽셀 값을 각각 CZP₁(closest zero point 1), CZP₂로 표시하면 PP₁=153, CZP₁=33, PP₂=154, CZP₂=235가 되는 것을 Fig. 1에서 확인할 수 있다. Ni, Shi, Ansari, Su 등은 히스토그램 쉬프트를 이용한 데이터 은닉 기법(NSAS)을 제안하였다[2]. 이 기법에서는 임베딩 되는 기밀 데이터 비트 수가 최대 빈도수의 합(h(PP₁)+h(PP₂))인 5,701 bit로 제한되는 단점이 있다.

Li 등은 NSAS 기법을 개선한 APD 기법을 제안하였다[3]. APD 기법은 인접한 픽셀 값들이 비슷한 값을 갖는 특성을 효과적으로 이용하여 커버 이미지에 임베딩 되는 기밀 데이터 비트수를 크게

증가시킨다. APD기법에서는 inverse s-order로 커버 이미지를 스캔한다. 즉, 커버 이미지를 좌측에서 우측으로, 위에서 아래쪽으로 스캔하여 픽셀 값 시퀀스(sequence)인 커버 이미지 시퀀스 C를 생성한다. 인접한 픽셀 간의 픽셀 값 차이 값들로 구성되는 차분 시퀀스 D를 구성하기 위하여 식 (1)을 커버 이미지 시퀀스 C에 적용한다. 구해진 차분 시퀀스 D에 대한 히스토그램은 0과 0에 가까운 값들에서의 빈도수가 크게 증가하게 된다. 이러한 특성은 자연 영상(natural image)에 존재하는 지역적 유사성(local similarity)때문이다. 식 (1)에서 n은 (image height)*(image width)의 값을 갖는다.

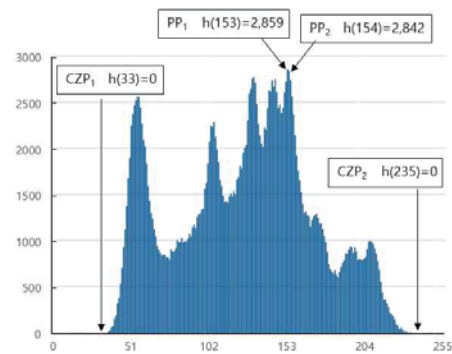


Fig. 1. Histogram of 512x512 Lenna image

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

Lenna영상을 커버 이미지로 사용하여 생성된 커버 이미지 시퀀스 C에 식 (1)을 적용하여 생성된 차분 시퀀스 D에 대한 히스토그램은 Fig. 2와 같다. Fig. 2에서 보는 바와 같이 h(PP₁)=25,010, h(PP₂)=23,414가 되어 그림 1에서의 h(PP₁)=2,859, h(PP₂)=2,842 값보다 매우 큰 값이 된다. APD기법에서는 차분 시퀀스 D에 대한 히스토그램을 쉬프트 시켜서 기밀 데이터를 임베딩하기 때문에 임베딩 되는 최대 데이터 비트수는 peak point에서의 빈도수 합인 h(PP₁)+h(PP₂)=48,424비트가 된다. 따라서 APD 기법은 NSAS 기법에 비하여 42,723비트를 더 많이 삽입할 수 있어, NSAS기법에 비하여 삽입되는 기밀 데이터 비트수가 8.5배 정도 증가한다.

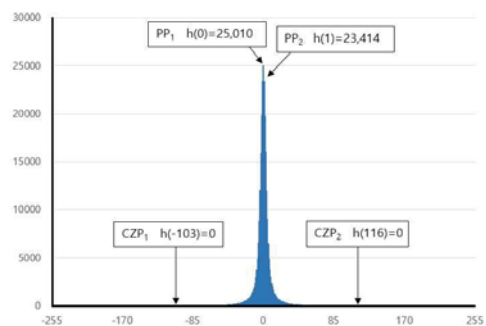


Fig. 2. Histogram of sequence F

APD기법의 data embedding process는 Fig. 3과 같고, APD기

법에서 커버 이미지에 삽입되는 최대 데이터 비트수는 Fig. 3의 step5에서 보는 바와 같이 $h(PP_1)+h(PP_2)$ 가 된다.

Step 1. 커버 이미지의 픽셀 값을 좌측에서 우측으로, 위에서 아래쪽으로 역 s-순(inverse s-order)으로 scan하여 픽셀 값들로 구성되는 커버 이미지 시퀀스 C 생성.

Step 2. 식 (1)을 만족하는, 인접 픽셀 간의 픽셀 값 차분을 갖는 차분 시퀀스 D 생성.

Step 3. 차분 시퀀스 D에 대한 히스토그램을 생성한 후, PP_1 , CZP_1 , PP_2 , CZP_2 값 결정.

Step 4. 차분 시퀀스 D를 식 (2), (3)에 따라 쉬프트(shift) 시켜, 쉬프트된 차분 시퀀스 DS 생성.

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

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

Step 5. 식 (4)와 같이, 쉬프트된 차분 시퀀스 DS에 기밀 데이터를 삽입(embedding)하여 임베딩 시퀀스 DE 생성.

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

Step 6. 식 (5)에 따라, 스테고 이미지 시퀀스 S 생성.

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

where $n=(\text{image height}) * (\text{image width})$

Step 7. S를 순차적으로 scan하여 역 s-순(inverse s-order)으로 스테고 이미지 생성

Fig. 3. The embedding process of APD

기밀 데이터를 커버 이미지에 삽입하여 생성된 스테고 이미지로부터 기밀 데이터를 추출하고 원본 커버 이미지를 복원하는 절차는 Fig. 4와 같다. Fig. 3의 절차를 따라 커버 이미지에 기밀 데이터를 삽입하여 스테고 이미지를 생성하고, Fig. 4의 절차를 따라 스테고 이미지로부터 기밀 데이터와 커버 이미지를 추출하면 원본 기밀 데이터와 원본 커버 이미지를 손실 없이 추출할 수 있다. 따라서 APD기법은 대량의 기밀 데이터를 은닉할 수 있는 우수한 가역적 데이터 은닉 기법이다.

Step 1. 스테고 이미지의 픽셀 값을 좌측에서 우측으로, 위에서 아래쪽으로 역 s-순(inverse s-order)으로 scan하여 픽셀 값들로 구성되는 은닉 이미지 시퀀스 S 생성.

Step 2. i 값을 증가 시키면서 step 2.1과 step 2.2를 반복적으로 수행하여 커버 이미지 시퀀스 C, 임베딩 시퀀스 DE 복원.

Step 2.1. 식 (6)을 만족하는 임베딩시퀀스 DE 복원.

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

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

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

Step 3. 임베딩 시퀀스의 각 요소 DE_i 에 식 (8)을 적용하여 스테고 이미지에 삽입되어 있는 기밀 데이터 추출.

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

Step 4. 복원된 커버 이미지 시퀀스 C를 scan하여 역s-순(inverse s-order)으로 원본 커버 이미지 생성.

Fig. 4. The extraction process of APD

APD기법을 적용한 data embedding process와 data extraction process에 대한 예제가 Fig. 5와 Fig. 6에 나타나 있다. Fig. 5의 커버 이미지에 대한 차분 시퀀스 D에서, $PP_1=0$, $PP_2=1$ 이며 $h(PP_1)=7$, $h(PP_2)=4$ 이다. 따라서 커버 이미지에 삽입될 수 있는 최대 데이터 비트수는 $h(PP_1)+h(PP_2)=11$ 비트이다. Fig 5에서 보는 바와 같이 기밀 데이터를 커버 이미지에 삽입하여 생성한 스테고 이미지는 원본 커버 이미지와 부분적으로 차이가 있어 스테고 이미지의 화질은 저하되지만, Fig. 6에서 보는 바와 같이 스테고 이미지로부터 기밀 데이터와 원본 커버 이미지가 손실 없이 완벽하게 추출되기 때문에 의료, 국방 등의 분야에 매우 적합하게 응용될 수 있다.

III. Proposed Algorithm(APDLS)

자연 영상에는 일반적으로 공간적 지역성(spatial locality)이 존재하고, 영상의 각 부분에는 표면 특성(surface characteristics)이 존재한다. 즉, 인접한 픽셀 값들은 매우 비슷한 값들을 갖는 특성이 있으며 곡면 혹은 평면 등의 다양한 표면 특성을 갖는다.

공간적 지역성과 표면 특성을 활용하는 제안기법(APDLS: Adjacent Pixel Difference using spatial Locality and Surface characteristics)에서는 인접 픽셀 값들을 이용하여 지역성이 높을 것으로 예상되는 위치를 선정한 후, 선정된 위치에서의 표면 특성을 조사하여 픽셀 값을 정밀하게 예측한다. 정밀하게 예측된 픽셀 값을 이용하여 픽셀 값 차로 구성되는 차분 시퀀스를 구성하면, 차분 시퀀스에 대한 히스토그램의 peak point에서의 빈도수가 크게 증가하여, 커버 이미지에 삽입되는 기밀 데이터 비트수를 크게 증가시킬 수 있다.

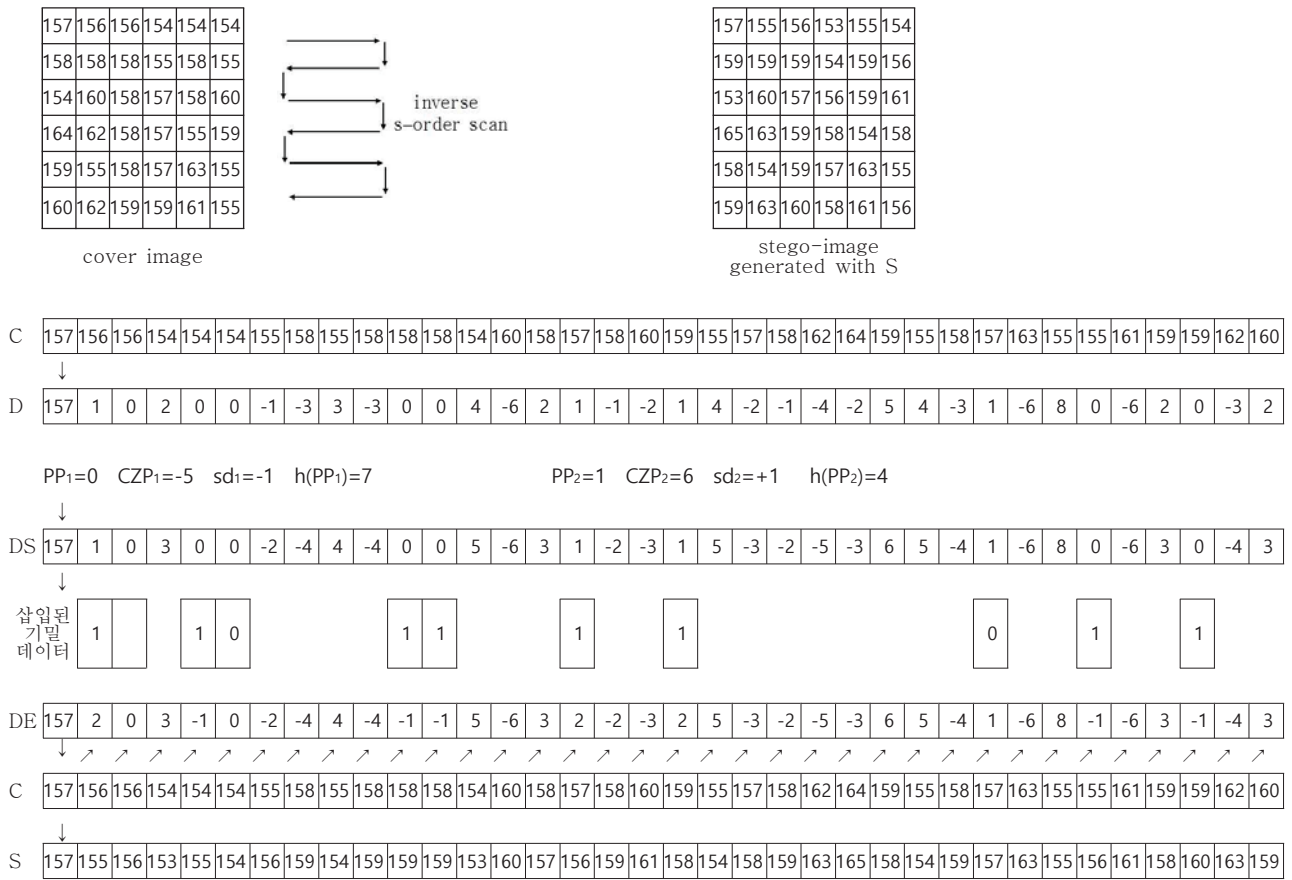


Fig. 5 Data embedding process of APD

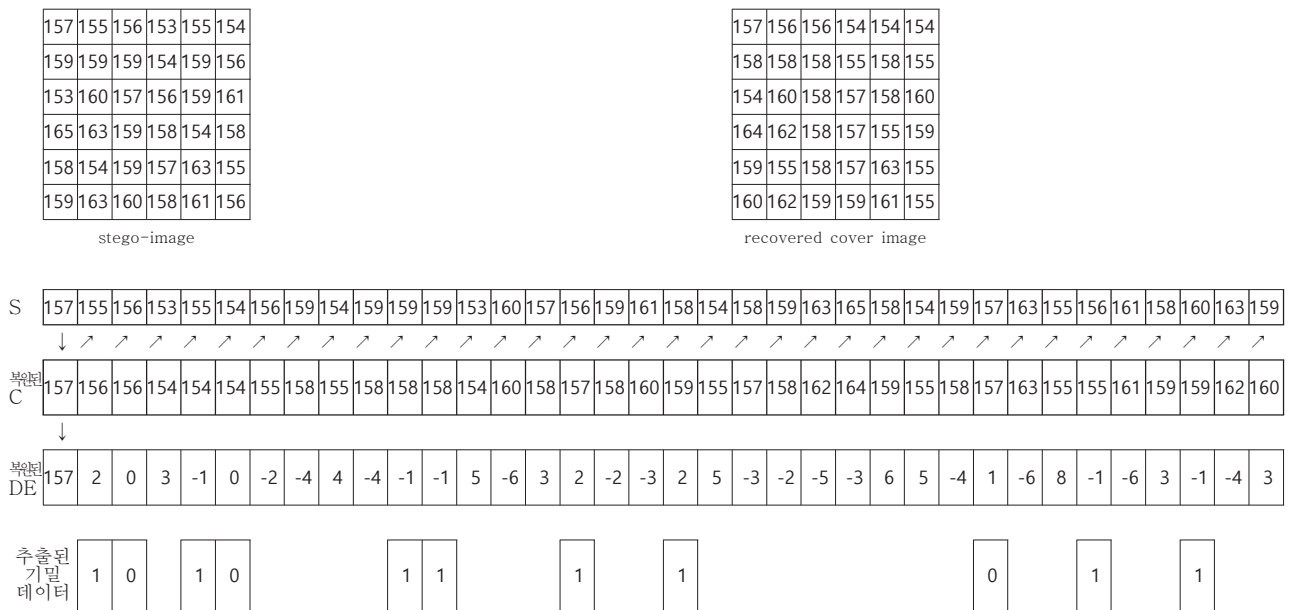


Fig. 6. Data extraction and cover image recovery process of APD

픽셀 값 예측이 정확할수록, 픽셀 값 차로 구성되는 차분 시퀀스에 대한 히스토그램의 peak point에서의 빈도수가 더욱 증가하게 되어 커버 이미지에 삽입될 수 있는 기밀 데이터 비트수가 더욱 증가하게 된다.

커버 이미지의 픽셀들을 사용하여 픽셀 값을 예측한 후, 예측 이미지를 생성하는 과정은 Fig. 7과 같다. Fig. 7의 (a)는 커버 이미지를 나타내고, (b)는 예측 이미지를 구성하기 위한 픽셀 값 예측 위치와 예측 순서를 나타낸다. 역 s-순(inverse-s order)으로 이미 scan된 12개의 커버 이미지의 픽셀 값들을 사용하여 굵은 선 안쪽에 있는 각 위치(①~⑧)에서의 픽셀 값들을 순차적으로 예측한다. Fig. 7의 (c)는 위치 ①과 ④에서의 픽셀 값을 예측하는데 사용되는 12개의 주변 픽셀 값들을 나타내며, 12개의 주변 픽셀 값들은 역 s-순으로 이미 scan된 픽셀 값들이다. (d)는 제안된 기법을 적용하여 생성된 예측 이미지이다. 예측 이미지에서 굵은 선 바깥쪽은 픽셀 값 예측 가능 영역이 아니기 때문에 커버 이미지의 픽셀 값을 예측 이미지에 그대로 사용한다.

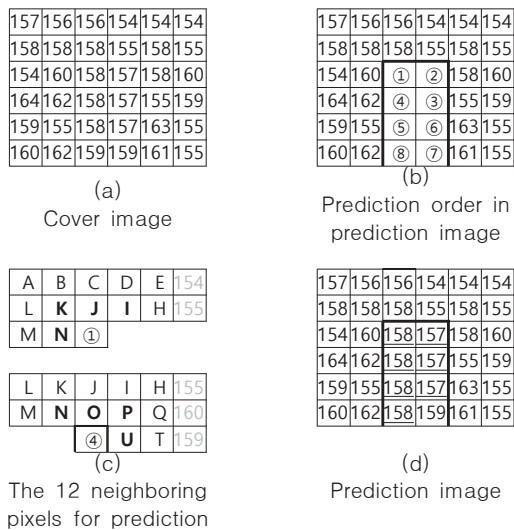
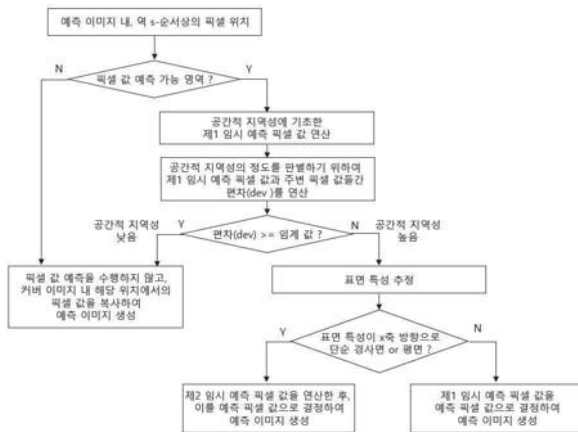


Fig. 7. Generation of prediction image

제안된 기법에서 예측 이미지를 생성하는 절차를 나타내는 흐름도는 Fig. 8과 같다.



8. The flowchart of prediction image generation

제안된 기법에서는 예측 이미지를 생성하기 위하여 역 s-순 (inverse s-order)으로 각 위치에서 픽셀 값 예측 가능 영역인지 여부를 판단한다. 픽셀 값 예측 가능 영역이 아닌 경우에는 커버 이미지 내의 해당 위치에서의 픽셀 값을 복사하여 예측 이미지를 생성한다. 픽셀 값 예측 가능 영역인 경우에는 아래와 같은 6단계의 픽셀 값 예측 절차를 수행하여 예측 이미지를 생성한다.

단계 1: 공간적 지역성에 기초하여, 주변 픽셀들의 위치별 영향을 고려한 제1 임시 예측 픽셀 값 V_1 을 구한다. 위치 ①에서의 제1 임시 예측 값 V_1 은 식 (9)와 같이 구한다. 위치 ①의 좌표는 (x_i, y_j) 이고, 픽셀 값이 I, J, K, N인 픽셀의 좌표는 각각 (x_{i+1}, y_{j-1}) , (x_i, y_{j-1}) , (x_{i-1}, y_{j-1}) , (x_{i-1}, y_j) 이다. $P(x, y)$ 는 좌표 (x, y) 에서의 픽셀 값을 나타낸다. I, J, K, N은 그림 (a)에서 보는 바와 같이 155, 158, 158, 160의 값을 갖는다. 식 (9)에서 보는 바와 같이 좌측과 우측과 위쪽으로 인접해 있는 픽셀들이 제1 임시 예측 픽셀 값 V_1 에 미치는 영향을 나타내는 계수는 1로 하였고, 한 픽셀 떨어져 있는 위치의 픽셀들이 제1 임시 예측 픽셀 값 V_1 에 미치는 영향을 나타내는 계수는 α 로 표시하였다.

단계 2: 공간적 지역성의 정도를 판별하기 위하여, 제1 임시 예측 픽셀 값 V_1 과 주변 픽셀 값들과의 편차(dev)를 식 (10)과 같이 구한다.

$$V_1(x_i, y_j) = [P(x_{i+1}, y_{j-1}) + P(x_i, y_{j-1}) + P(x_{i-1}, y_{j-1}) + P(x_{i-1}, y_j) + \{P(x_{i-2}, y_{j-2}) + P(x_{i-1}, y_{j-2}) + P(x_i, y_{j-2}) + P(x_{i+1}, y_{j-2}) + P(x_{i+2}, y_{j-2}) + P(x_{i+2}, y_{j-1}) + P(x_{i-2}, y_{j-1}) + P(x_{i-2}, y_j)\} * \alpha] / (4 + 8 * \alpha) \quad (9)$$

$$dev(x_i, y_j) = abs(V_1 - P(x_{i+1}, y_{j-1})) + abs(V_1 - P(x_i, y_{j-1})) + abs(V_1 - P(x_{i-1}, y_{j-1})) + abs(V_1 - P(x_{i-1}, y_j)) + \{abs(V_1 - P(x_{i-2}, y_{j-2})) + abs(V_1 - P(x_{i-1}, y_{j-2})) + abs(V_1 - P(x_i, y_{j-2})) + abs(V_1 - P(x_{i+1}, y_{j-2})) + abs(V_1 - P(x_{i+2}, y_{j-2})) + abs(V_1 - P(x_{i+2}, y_{j-1})) + abs(V_1 - P(x_{i-2}, y_{j-1})) + abs(V_1 - P(x_{i-2}, y_j))\} * \alpha \quad (10)$$

단계 3: 공간적 지역성의 정도를 판단하기 위하여 구해진 편차 (dev)를 임계 값(β)과 비교한다.

단계 4: 편차(dev)가 임계 값(β) 이상일 경우에는 해당 위치에서의 영상은 픽셀 값 변화가 큰 경계 영역 혹은 공간적인 지역성이 낮은 영역으로 판단하여 픽셀 값 예측 과정을 더 이상 수행하지 않고, 커버 이미지의 해당 위치에서의 픽셀 값을 복사하여 예측 이미지를 생성한다. 구해진 편차(dev)가 임계 값(β) 미만일 경우에는 해당 위치에서의 영상은 픽셀 값 변화가 적고, 공간적 지역성이 높은 것으로 판단하여 단계 5를 수행한다.

단계 5: 해당 위치에서의 표면 특성을 추정한다. 위치 ①에서의 표면 특성을 추정하는 과정은 다음과 같다. 임의의 두 지점 (x_1, y_1) 과 (x_2, y_2) 를 지나는 직선의 방정식은 식 (11)과 같다. 식 (11)을

이용하면 상호 인접한 픽셀사이의 직선 방정식을 간단히 구할 수 있다. 픽셀 값이 K인 픽셀과 X축으로 인접한 픽셀 값이 J인 픽셀 사이에서, 픽셀 값에 대하여 X축 방향만을 고려한 직선 방정식을 구한다. 즉, (x_{i-1}, K) 와 (x_i, J) 를 지나는 직선의 방정식 $P_L(x)$ 은 식 (12)와 같다. 동일한 방법으로 픽셀 값이 J인 픽셀과 픽셀 값이 I인 픽셀 사이에서, 픽셀 값에 대하여 X축 방향만을 고려한 직선 방정식 $P_R(x)$ 은 식 (13)과 같다. 픽셀 값이 N인 픽셀과 픽셀 값이 J인 픽셀 사이에서, 픽셀 값에 대하여 X축 방향만을 고려한 직선 방정식 $P_{LD}(x)$ 은 식 (14)와 같다. 식 (12)~ (14)의 기울기가 모두 0보다 큰 경우에는 왼쪽보다 오른쪽의 픽셀 값이 큰 값을 갖는 단순 경사면(simple inclined surface)이라고 추정한다. 또한 식 (12)~(14)의 기울기가 모두 0보다 적은 경우에는 오른쪽보다 왼쪽의 픽셀 값이 더 큰 값을 갖는 단순 경사면이라고 추정하고, 식 (12)~(14)의 기울기가 모두 0인 경우에는 왼쪽과 오른쪽이 평평한 평면(flat surface)이라고 추정한다.

단계 6: 해당 위치에서 X축 방향의 표면 특성에 따라 다음과 같이 예측 픽셀 값을 결정한다. 해당 위치에서 X축 방향의 표면 특성이 단순 경사면이거나 혹은 평면인 경우에는, 위치 ①에서의 제2 임시 예측 픽셀 값 V_2 는 식 (15)와 같이 구하고, 제2 임시 예측 픽셀 값 V_2 를 위치 ①에서의 예측 픽셀 값으로 하여 예측 이미지를 생성한다. 그 외의 표면특성을 갖는 경우에는 식 (9)에서 구해진 제1 임시 예측 값 V_1 을 위치 ①에서의 예측 픽셀 값으로 하여 예측 이미지를 생성한다. 따라서 위치 ①에서의 예측 픽셀 값 $V(x_i, y_j)$ 는 식 (16)과 같이 되고, 예측 픽셀 값 $V(x_i, y_j)$ 가 예측 이미지의 픽셀 값 $P(x_i, y_j)$ 가 된다. 식 (16)에서 $C(x_i, y_j)$ 는 좌표 (x_i, y_j) 에서의 커버 이미지의 픽셀 값을 나타내고, $V_1(x_i, y_j)$ 과 $V_2(x_i, y_j)$ 는 좌표 (x_i, y_j) 에서의 제1 임시 예측 픽셀 값과 제2 임시 예측 픽셀 값을 각각 나타내고, $P(x_i, y_j)$ 는 좌표 (x_i, y_j) 에서의 예측 이미지의 픽셀 값을 낸다.

$$y = \{(y_2 - y_1) / (x_2 - x_1)\}(x - x_1) + y_1 \quad (11)$$

$$P_L(x) = \{(J - K) / (x_i - x_{i-1})\}(x - x_{i-1}) + K \quad (12)$$

$$P_R(x) = \{(I - J) / (x_{i+1} - x_i)\}(x - x_i) + J \quad (13)$$

$$P_{LD}(x) = \{(J - N) / (x_i - x_{i-1})\}(x - x_{i-1}) + N \quad (14)$$

$$V_2(x_i, y_j) = (I + J + K + N) / 4 \quad (15)$$

$$V(x_i, y_j) = \begin{cases} C(x_i, y_j) & \text{if } dev \geq \beta \\ V_2(x_i, y_j) & \text{if } dev < \beta \text{ and} \\ & \text{(경사면 or 평면)} \\ V_1(x_i, y_j) & \text{if } dev < \beta \text{ and} \\ & \text{the other surface} \end{cases} \quad (16)$$

식 (16)에 따라 구해진 예측 픽셀 값을 사용하여 예측 이미지를 생성한 결과가 Fig. 7 (d)에 나타나 있다. Fig. 7 (d)에서 밀줄 친 픽셀 값들은 해당 위치에서의 예측 픽셀 값을 나타낸다. 밀줄이 없는, 위치 ⑦에서의 픽셀 값 159는 커버 이미지의 픽셀 값이다. 제안된 기법에서는 식 (9), (10), (16)을 적용하여 예측 이미지를

생성하기 때문에 예측 이미지의 상위 2개 행과 좌·우측 각 2개 열의 픽셀 값은 커버 이미지의 픽셀 값을 그대로 복사하여 사용하게 되고, 나머지 영역이 픽셀 값 예측 가능 영역이 되며, 예측 픽셀 값은 식 (16)에 따라 계산된다. 제안된 기법의 data embedding process는 Fig. 9와 같고, 제안된 기법에서의 data extraction process는 Fig. 10과 같다.

- Step 1. 예측 이미지의 상위 2개 행과 좌·우측 각 2개 열의 픽셀 값은 커버 이미지의 픽셀 값을 그대로 사용하고, 예측 이미지의 나머지 영역(픽셀 값 예측 가능 영역)에서의 예측 픽셀 값은 아래의 단계를 따라 구한 후, 예측 이미지를 생성.

Step 1.1 제1 임시 예측 픽셀 값 V_1 을 식 (9)에 따라 계산.

Step 1.2 제1 임시 예측 픽셀 값 V_1 과 주변 픽셀 값들과의 편차(dev)를 식 (10)에 따라 계산.

Step 1.3 if $(dev) \geq \beta$ then 커버 이미지 내 해당 위치에서의 픽셀 값을 복사하여 예측 이미지 생성.

Step 1.4 else if $((dev < \beta) \ \&\& \text{(경사면 or 평면)})$ then 식 (15)를 적용하여 구한 제2 임시 예측 픽셀 값 V_2 를 예측 픽셀 값으로 하여 예측 이미지 생성.

Step 1.5 else V_1 을 예측 픽셀 값으로 하여 예측 이미지 생성.

Step 2. 커버 이미지와 예측 이미지(prediction image)의 픽셀 값을 각각 좌측에서 우측으로, 위에서 아래쪽으로 역 s-순(inverse s-order)으로 scan하여 픽셀 값들로 구성되는 커버 이미지 시퀀스 C와 예측 이미지 시퀀스 P 생성.

Step 3. 커버 이미지 시퀀스 C와 예측 이미지 시퀀스 P를 사용하여 식 (17)에 따라 차분 시퀀스 D 생성.

$$D_i = \begin{cases} C_i & \text{if } i = 0 \\ C_i - P_i & \text{else if 예측 픽셀 값 사용} \\ C_{i-1} - C_i & \text{otherwise (예측 픽셀 값 미사용)} \end{cases} \quad (17)$$

Step 4. 차분 시퀀스 D에 대한 히스토그램을 생성하여 PP_1, CZP_1, PP_2, CZP_2 결정.

Step 5. 식 (2), (3)을 사용하여 차분 시퀀스 D를 쉬프트 시켜, 쉬프트된 차분 시퀀스 DS 생성.

Step 6. 차분 시퀀스 DS에 기밀 데이터를 식 (4)와 같이 삽입하여 임베딩 시퀀스 DE 생성.

Step 7. 커버 이미지 시퀀스 C와 임베딩 시퀀스 DE에 대하여 식 (5)를 적용하여 스테고 이미지 시퀀스 S 생성.

Step 8. 스테고 이미지 시퀀스 S를 순차적으로 scan하여 역 s-순(inverse s-order)으로 스테고 이미지 생성.

Fig. 9. Data embedding process of the proposed algorithm

Step 1. 스테고 이미지의 픽셀 값을 좌측에서 우측으로, 위에서 아래쪽으로 역 s-순(inverse s-order)으로 scan하여 픽셀 값들로 구성되는 스테고 이미지 시퀀스 S 생성.

Step 2. $C_0=S_0$ $DE_0=C_0$
 i 값을 증가 시키면서, 픽셀 값 예측 가능 영역인지 판단. 픽셀 값 예측 가능 영역인 경우에는 Fig. 9의 step 1.1~1.5의 과정을 수행하여, 해당 위치에서의 픽셀 값이 커버 이미지의 픽셀 값이 사용되었는지 혹은 예측 픽셀 값이 사용되었는지를 판단하여 extraction mode 결정.

Mode 1: 은닉 이미지의 상위 2개 행에 해당하는 경우 ($i < \text{imagewidth} * 2$) 혹은 은닉 이미지 시퀀스의 이전 위치와 현재 위치 모두에서 예측 픽셀 값이 사용되지 아니한 경우. 즉, Mode 2와 Mode 3를 제외한 전 영역.
 Fig. 4의 APD기법과 같이, 식 (6)과 식(7)을 수행하여 임베딩 시퀀스 DE와 커버 이미지 시퀀스 C를 복원.

Mode 2: 현재 위치에서 예측 값이 사용된 경우. 식 (6)을 사용하여 임베딩 시퀀스 DE를 복원하고, 식 (18)~(19)를 사용하여 커버 이미지 시퀀스 C를 복원한다. RP_i 는 복원과정 중, 추출 모드 판단결과가 Mode 2인 경우, i번째 위치에서 계산되는 복원 예측 픽셀 값이다. 식 (18)에 사용된 연관 값 AV(Associated Value)는 복원 과정 중, 추출 모드 판단 결과가 Mode 2 혹은 Mode 3인 경우에 식 (19)와 같이 계산되는 값이다.

$$C_i = RP_i + AV_i \tag{18}$$

$$AV_i = \begin{cases} C_{i-1} - S_i - sd_j & \text{if } C_{i-1} - S_i \in [PP_j + sd_j, CZP_j] \\ C_{i-1} - S_i & \text{otherwise} \end{cases} \tag{19}$$

Mode 3: 이전 위치에서 예측 픽셀 값이 사용되었으나, 현재 위치에서는 예측 픽셀 값이 사용되지 아니한 경우. 식 (6)을 사용하여 임베딩 시퀀스 DE를 복원하고, 식 (20)을 사용하여 커버 이미지 시퀀스 C를 복원한다. AV_i 는 식 (19)에 의해서 계산된 값이다.

$$C_i = RP_{i-1} - AV_i \tag{20}$$

Step 3. 임베딩 시퀀스의 각 요소 DE_i 에 대해 식 (8)을 적용하여 스테고 이미지에 삽입되어 있는 기밀 데이터를 추출.

Step 4. 복원된 커버 이미지 시퀀스 C를 scan하여 역 s-순 (inverse s-order)으로 원본 커버 이미지를 생성.

제안된 기법의 data embedding process와 data extraction process의 예제가 Fig. 11과 Fig. 12에 제시 되어 있다. Fig. 11과 Fig. 12는 식 (9), (10)에서 사용된 α 를 0.298, 식 (16)에 사용된 임계 값(β)을 15로 설정하여 생성된 예측 이미지의 예를 보이고 있다. Fig. 7의 ①~⑧까지의 각 위치에서 제1 임시 예측 픽셀 값 V_1 을 식 (9)에 따라 계산하면, 각각 157.0, 156.7, 157.3, 157.5, 158.0, 157.6, 158.7, 158.2 가 된다.

그리고 계산된 제1 임시 예측 값 V_1 을 식 (10)에 적용하여 편차 (dev)를 구하면, 각각 10.9, 10.3, 8.1, 7.3, 13.2, 8.4, 15.3, 12.2 가 된다.

식 (16)에서의 편차(dev)가 임계 값(β) 이상인 경우에는 예측 픽셀 값을 구하지 않고, 커버 이미지의 픽셀 값을 그대로 사용한다. 위치 ⑦에서의 편차(15.3)가 임계 값(β) 이상이기 때문에 Fig. 9의 step 1.3에서와 같이 예측 픽셀 값을 구하지 않고, 커버 이미지의 픽셀 값을 복사하고, 이를 위치 ⑦에서의 픽셀 값으로 하여 예측 이미지를 생성한다. 편차가 임계 값 미만인 나머지 위치 ①~⑥, ⑧에서는 영상의 표면 특성을 조사하여, 단순 경사면(simple inclined surface) 혹은 평면(flat surface)인 영역으로 추정되는 위치 ①, ④, ⑥을 찾는다. 위치 ①, ④, ⑥에서의 제2 임시 예측 픽셀 값 V_2 를 식 (15)에 따라 계산하면 각각 157.8, 158.0, 157.0이 되며, 이 값들을 반올림한 후, 반올림된 값들을 각 위치에서의 예측 픽셀 값으로 하여 예측 이미지를 생성한다. 그 외의 위치 ②, ③, ⑤, ⑧에서의 표면 특성은 단순 경사면 혹은 평면이 아닌 것으로 추정되었기 때문에 식(9)에 따라 구해진 제1 임시 예측 픽셀 값 V_1 을 반올림한 값이 해당 위치에서의 예측 픽셀 값이 되어 예측 이미지가 생성된다. 이러한 과정을 거쳐 구성된 예측 이미지가 Fig. 7 (d)와 Fig. 10에 나타나있다. Fig. 10에서 보는 바와 같이 제안기법을 적용하여 차분 시퀀스 D를 생성하면, 차분 시퀀스에 대한 히스토그램에서 $PP_1=0$, $PP_2=1$ 이며 $h(PP_1)=12$, $h(PP_2)=3$ 이 된다. 따라서 커버 이미지에 삽입되는 최대 데이터 비트수는 15비트로, APD기법에 비하여 4비트 증가하게 된다.

임계 값(β)을 20으로 두면, 위치 ⑦에서도 예측 픽셀 값이 사용되어 $h(PP_1)=13$, $h(PP_2)=3$ 이 되어 삽입되는 기밀 데이터 비트수가 1비트 더 증가 하게 된다. 제안된 기법에서, 식 (16)에서의 임계 값(β)을 0으로 두면 예측 이미지의 모든 픽셀은 커버 이미지의 픽셀 값을 갖게 된다. 따라서 임계 값(β)이 0인 경우에는 제안 기법은 APD기법과 동일하게 동작하게 된다.

Fig. 11과 Fig. 12에서 보는 바와 같이 제안기법을 적용하여 커버 이미지에 기밀 데이터를 삽입하게 되면 APD기법에 비하여 더 많은 기밀 데이터를 삽입할 수 있으며, 스테고 이미지로부터 원본 기밀 데이터와 원본 커버 이미지를 손실 없이 완벽하게 복원할 수 있어 군사, 의학 등의 응용분야에 효과적으로 활용될 수 있다.

Fig. 10. Data extraction process of the proposed algorithm

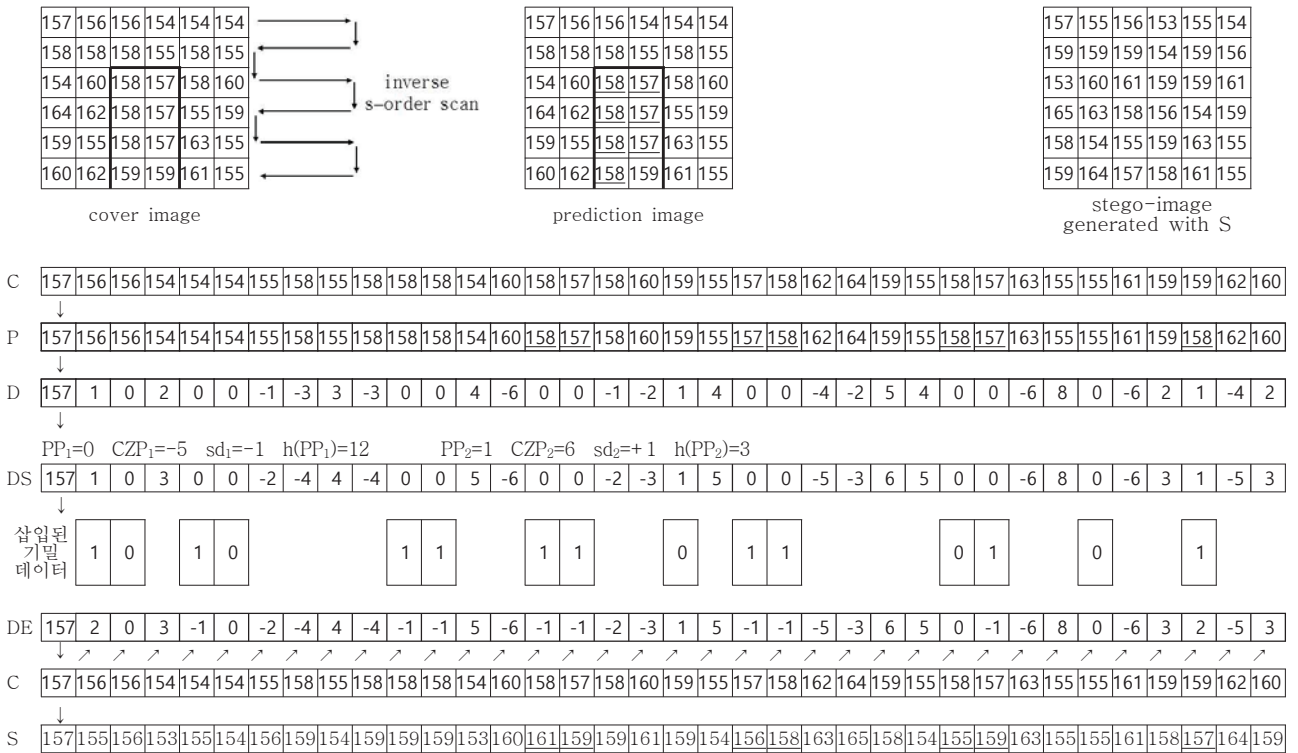


Fig. 11. Data embedding process of APDLS

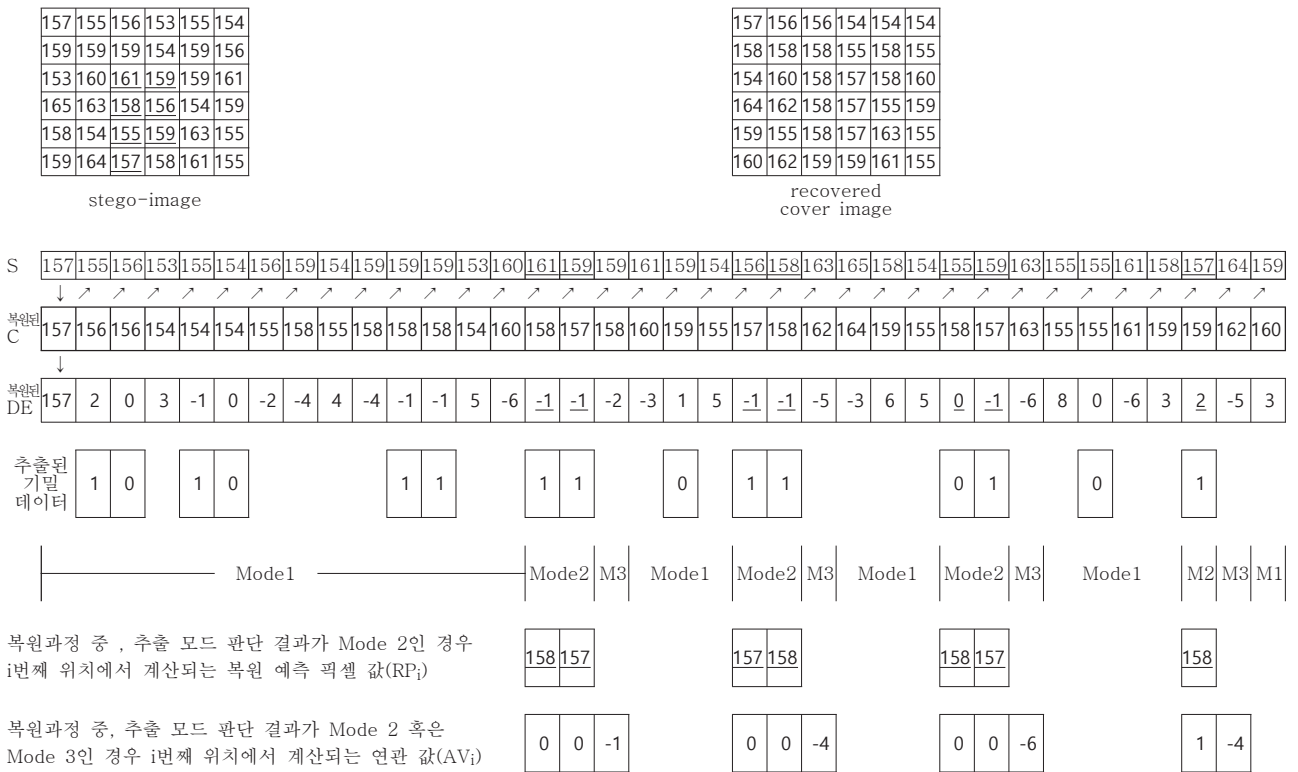


Fig. 12. Data extraction and cover image recovery process of APDLS

V. Experimental Results

본 논문에서 제안된 APDLS 기법의 성능을 평가하기 위하여 512x512 gray scale 영상인 Lenna, sail-boat, Barbara, Elaine을 커버 이미지로 사용하여 실험을 수행하였다. 본 논문의 abstract를 ASCII코드로 변환한 것을 기밀 데이터로 사용하여 커버 이미지에 삽입하였다. 식 (9), (10)에 사용된 α 는 0.298로 하였다. 수식 (16)에서의 임계 값(β)을 적용하여 실험을 수행한 경우를 APDLS(β)로 표기하였다.

커버 이미지로 사용된 Lenna, sail-boat, Barbara, Elaine에 대한 실험결과 영상이 Fig. 13~16에 나타나 있다. 각 커버 이미지에 대하여 APD기법과 제안기법 APDLS를 적용하여 실험을 수행한 결과가 Table 1에 나타나 있다.

임계치(β)를 0으로 두면, 예측 이미지는 커버 이미지와 동일하게 된다. 따라서 커버 이미지 시퀀스 C와 예측 이미지 시퀀스 P가 동일하게 된다. 그러므로 임계 값(β)이 0인 경우에 제안 기법은 APD기법과 동일하게 동작한다.

임계 값(β)이 커질수록 예측을 수행하는 횟수가 증가하게 되어 예측 이미지에 예측 픽셀 값이 많이 사용되어, 예측 이미지 시퀀스 P에 많은 예측 픽셀 값들이 포함되게 된다. 또한 픽셀 값 예측이 정확하면, 예측 픽셀 값이 적용된 경우의 D_i 값은 0이 되고, 픽셀 값 예측이 정확할수록 D_i 값은 0에 근접하게 된다. 이는 식 (17)에서 보는 바와 같이, 픽셀 값 예측이 정확한 경우에는 P_i 가 C_i 와 동일하게 되어 $D_i=C_i-P_i=0$ 이 되기 때문이다.

제안기법을 적용한 경우, Table 1의 실험 결과에서 보는 바와 같이 평균 예측 오차가 매우 적기 때문에 D_i 는 매우 작은 값이 된다. 따라서 차분 시퀀스 D에 대한 히스토그램을 생성하면 peak point에서의 빈도수가 크게 증가하게 된다. 그러므로 커버 이미지에 삽입되는 기밀 데이터 비트수가 크게 증가 하게 된다.

임계 값(β)을 0부터 35까지 증가시키면 스테고 이미지의 PSNR 값은 감소하지만, 스테고 이미지에 기밀 데이터가 은닉되어있는지 여부를 시각적으로 분간하기 어려울 정도로 스테고 이미지와 커버 이미지가 동일함을 Fig. 13~16에서 확인 할 수 있다. 제안된 APDLS 기법을 적용하여 커버 이미지에 기밀 데이터를 은닉하면, 스테고 이미지의 화질은 시각적으로 커버 이미지와 구별이 거의 어려울 정도로 매우 우수하게 유지되고, 삽입되는 비트수가 APD기법에 비하여 최대 25.9% 증가 하였다.

제안된 APDLS 기법을 적용하면, 커버 이미지에 다양한 레벨로 기밀 데이터를 은닉 할 수 있다. 또한 스테고 이미지로부터 원본 기밀 데이터와 원본 커버 이미지를 손실 없이 완벽하게 복원할 수 있어 의료, 국방, 디지털 라이브러리 등의 응용분야에 매우 유용하게 사용될 수 있다.

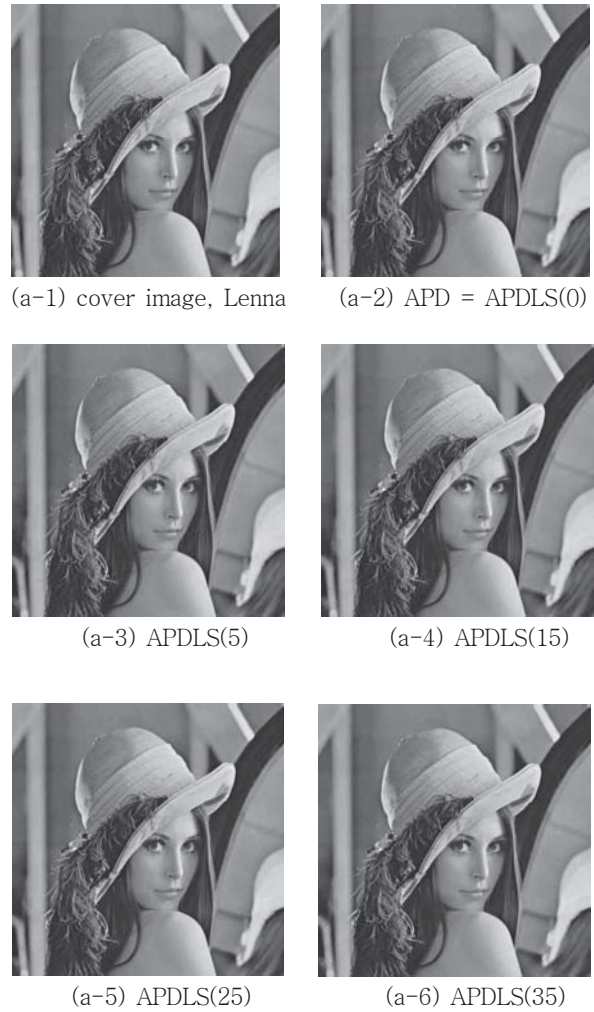
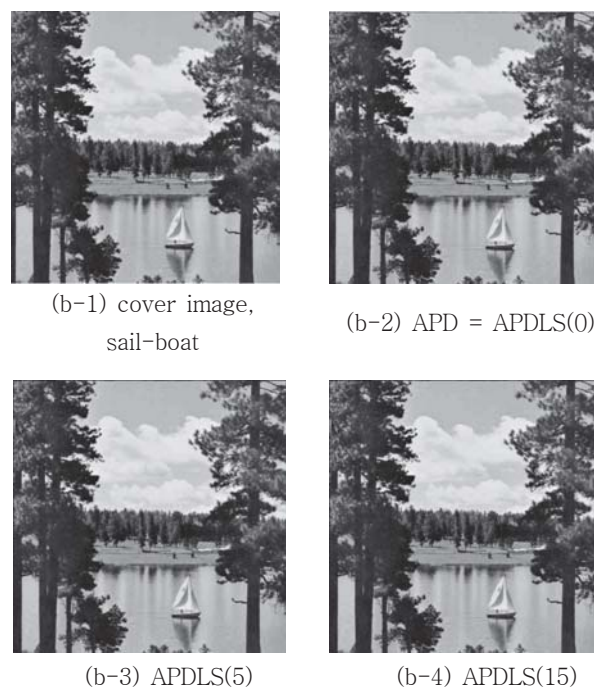


Fig. 13. Lenna image & stego-images(2~6)





(b-5) APDLS(25)



(b-6) APDLS(35)

Fig. 14. sail-boat image & stego-images(2~6)



(d-1) cover image, Elaine



(d-2) APD = APDLS(0)



(c-1) cover image,
Barbara



(c-2) APD = APDLS(0)



(d-3) APDLS(5)



(d-4) APDLS(15)



(c-3) APDLS(5)



(c-4) APDLS(15)



(d-5) APDLS(25)



(d-6) APDLS(35)

Fig. 16. Elaine image & stego-images(2~6)



(c-5) APDLS(25)



(c-6) APDLS(35)

Fig. 15. Barbara image & stego-images(2~6)

Table 1. The experimental results of the proposed algorithm

영상	기법	은닉 비트수	PSNR (dB)	예측 적중률 (%)	평균 예측 오차	예측 횟수	은닉 비트수 증가율 (%)
Lenna	APD	48424	48.60	x	x	x	x
	APDLS(0)	48424	48.56	x	x	x	x
	APDLS(5)	48705	47.58	27.47	1.28	3112	0.6
	APDLS(10)	52447	39.82	20.23	1.64	54367	8.3
	APDLS(15)	56369	36.31	18.29	1.82	112861	16.4
	APDLS(20)	57985	34.73	17.25	1.96	143593	19.7
	APDLS(25)	59246	33.67	16.49	2.10	162581	22.3
	APDLS(30)	60257	32.83	15.98	2.21	176288	24.4
	APDLS(35)	60985	32.11	15.58	2.33	187023	25.9
sail boat	APD	30880	48.42	x	x	x	x
	APDLS(0)	30880	48.42	x	x	x	x
	APDLS(5)	31140	47.60	31.35	1.02	3777	0.8
	APDLS(10)	32701	42.98	23.12	1.54	21900	5.9
	APDLS(15)	34272	38.31	18.15	2.05	46552	11.0
	APDLS(20)	35587	34.80	15.06	2.57	72686	15.2
	APDLS(25)	36563	32.28	13.20	3.01	98555	18.4
	APDLS(30)	37229	30.61	12.02	3.35	121069	20.6
	APDLS(35)	37497	29.42	11.24	3.64	139059	21.4
Barbara	APD	33113	48.44	x	x	x	x
	APDLS(0)	33113	48.44	x	x	x	x
	APDLS(5)	33246	47.88	19.60	1.57	1260	0.4
	APDLS(10)	36079	41.09	18.26	1.73	34199	9.0
	APDLS(15)	38603	37.39	16.70	1.93	75647	16.6
	APDLS(20)	39906	35.69	15.67	2.10	100018	20.5
	APDLS(25)	40563	34.64	14.99	2.25	114407	22.5
	APDLS(30)	40895	33.83	14.49	2.38	124187	23.5
	APDLS(35)	41112	33.08	14.05	2.52	131723	24.2
Ela-line	APD	30487	48.42	x	x	x	x
	APDLS(0)	30487	48.42	x	x	x	x
	APDLS(5)	30539	48.11	59.53	0.66	1853	0.2
	APDLS(10)	30846	45.18	29.48	1.44	10414	1.2
	APDLS(15)	31266	40.10	19.56	2.29	24221	2.6
	APDLS(20)	31891	35.37	14.19	3.09	48404	4.6
	APDLS(25)	32664	31.89	11.32	3.71	81479	7.1
	APDLS(30)	33498	29.50	9.86	4.15	117891	9.9
	APDLS(35)	34245	27.89	8.95	4.48	151365	12.3

V. Conclusions

본 논문에서는 커버 이미지에 기밀 데이터를 은닉하는 가역적 데이터 은닉기법인 APDLS기법을 제안하였다. 제안된 APDLS기법을 적용하면 다양한 레벨에서 기밀 데이터 은닉이 가능하다. 또한 생성된 스테고 이미지의 시각적 화질은 커버 이미지와 구별이 거의 어려울 정도로 높게 유지되면서, 삽입되는 기밀 데이터의 비트수가 APD기법에 비하여 최대 25.9% 증가하였다.

제안된 APDLS기법을 적용하여 커버 이미지에 기밀 데이터를 은닉하면, 기존의 APD 기법에 비하여 커버 이미지에 보다 많은 기밀 데이터를 효율적으로 은닉할 수 있다. 또한 스테고 이미지로부터 원본 기밀 데이터와 원본 커버 이미지를 손실 없이 완벽하게 복원할 수 있어 의료, 국방, 디지털 라이브러리 등의 응용분야에 매우 유용하게 사용될 수 있다.

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] C. C. Lo, Y. C. Hu, W. L. Chen, and C. M. Wu, "Reversible Data Hiding Scheme for BTC-compressed Images based on Histogram Shifting," *International Journal of Security and Its Applications*, Vol. 8, No. 2, pp. 301-314, March

2014.

- [11] S. M. Jung, "An advanced reversible data hiding algorithm based on the similarity between neighboring pixels," Journal of The Korea Society of Computer and Information, Vol. 21, No. 2, pp. 33-42, February 2016.

Authors



Soo-Mok Jung received the B.S., M.S. degrees in Electronics Engineering from Kyungpook National University, Korea, in 1984 and 1986 respectively. He received the Ph.D. degree in Computer Science and Engineering from Korea University, Korea, in 2002.

Dr. Jung joined the faculty of the Department of Computer Information at Sahmyook College, Seoul, Korea, in 1991. He is currently a professor in the Department of Computer Science & Engineering, Sahmyook University. He is interested in image processing and parallel processing.



Byung-Won Ong received M.S. degree in Computer Science and Engineering from Korea University, in 2000. He received the Ph.D. degree in Computer Science and Engineering from Pennsylvania State University, USA, in 2006.

Dr. On joined the faculty of the Department of Statistics & Computer Science, in 2014. He is interested in Data Mining, Information Retrieval, and Databases.