Journal of Engineering and Applied Sciences 12 (10): 2623-2627, 2017
ISSN: 1816-949X

An Improved Reversible Data Hiding Algorithm using Adjacent Pixel Difference Scheme

Soo-Mok Jng
Division of Computer Science and Engineering, Sahmyook University, 01795 Seoul, Korea

Abstract: Adjacent pixels of an image have very similar values. The APD (Adjacent Pixel Difference) algorithm embeds data using the difference sequence of adjacent pixels. In this study, we proposed an improved reversible data hiding algorithm that applies pixel value prediction algorithm based on similarity of adjacent pixel values and data compression algorithm to APD technique. By using the proposed algorithm in this study, more data can be embedded in cover image at various levels. The superiority of the proposed algorithm is verified by experiments.

Key words: Cover image, data hiding, reversible data hiding, adjacent pixel difference, locality

INTRODUCTION

Data hiding techniques are widely used to embed confidential data into cover images. If the quality of the generated stego-image is high by embedding the confidential data in the cover image, it is impossible for the person to recognize whether the confidential data is embedded. A technique for embedding confidential data as well as extracting confidential data and original cover image completely from a stego-image is called a reversible data hiding technique. Various reversible data hiding techniques have been developed in the past (Ni *et al*., 2006; Li *et al*., 2010, 2013; Tsai *et al*., 2004; Tseng and Hsieh, 2008; Lo *et al*., 2014; Jung and On, 2016).

NSAS was proposed to increase the number of confidential data to be embedded into cover image (Ni *et al*., 2006). In NSAS algorithm, the number of data bits embedded in cover image is limited to the sum of the frequency at Peak Point 1(PP_1) and the frequency at Peak Point 2(PP_2) in the histogram of the cover image.

In APD algorithm, a technique of using a differential sequence between adjacent pixels is proposed to effectively embed data (Li *et al*., 2010). Adjacent pixels have very similar values because there is locality in the image. Therefore, in the histogram for the difference sequence, the frequency at the peak point is greatly increased, so that a large amount of data can be embedded.

To increase the number of embedded bits of APD algorithm, an improved reversible data hiding algorithm was proposed in this study. The proposed algorithm is an improvement of the existing algorithm (Jung *et al*., 2016). Using the proposed algorithm, more data can be embedded in a cover image.

APD algorithm: The APD algorithm is a technique of concealing confidential data by shifting the histogram of the pixel value difference sequence. Adjacent pixel values are very similar because there is locality in the image. Therefore, the frequency at the peak point of the histogram of the adjacent pixel value difference sequence increases greatly. Since, the frequency at the peak point of the histogram is large, the number of data bits that can be concealed in the cover image increases. The embedding procedure of APD algorithm is as follows.

Step 1: The cover image sequence C is generated by scanning the cover image from top to bottom, left to right in inverse s-order.

Step 2: Generate a difference sequence D that satisfies (Eq. 1):

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

Step 3: After the histogram of D is generated, determine PP_1 and PP_2 with the first highest frequency and the second highest frequency. Determine CZP_1 and CZP_2 which are pixel values with a frequency value of 0 closest to PP_1 and PP_2.

Step 4: Generate a shifted sequence DS satisfying the Eq. 2 and 3:

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

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

Step 5: Generate sequence DE with embedded confidential data as shown in Eq. 4:

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

Step 6: Generate a sequence S of stego-image as shown in Eq. 5:

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

where, n = (image height) x (image width).

Step 7: Generate a stereo-image by placing the elements of the sequence S in from top to bottom, left to right inverse s-order. In APD algorithm, it is possible to extract confidential data and original cover image perfectly from stego-image in reverse of embedding procedure. The extraction procedure in the APD algorithm is as follows.

Step 1: Scan the stego-image from top to bottom, left to right in inverse s-order in order to generate a setgo-image sequence S of stego-image.

Step 2: Restore DE and cover image sequence C by repeatedly applying Eq. 6 and 7 with increasing i value:

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

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

Step 2: Confidential data hidden in the cover image is extracted from DE using Eq. 8:

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

Step 3: Generate the original cover image by placing the elements of the sequence C from top to bottom, left to right in inverse s-order. In the APD algorithm, the confidential data and original cover image can be extracted without loss and the maximum number of data bits to be embedded is limited to $h(PP_1)+h(PP_2)$.

MATERIALS AND METHODS

Proposed algorithm: In natural images, adjacent pixel values have similar characteristics. That is, there is high locality in natural images. Therefore, pixel value at arbitrary position can be predicted using adjacent pixel values at the point with high locality. In the pixel value prediction, first of all it is efficient to determine whether the locality is high and then pixel value can be predicted only when the region has high locality. Pixel value prediction is executed as follows. In Fig. 1a, the predicted pixel value PV at odd-numbered line is obtained by using 12 adjacent pixels. The average pixel value of the region is calculated by applying the weight at each position as shown in Eq. 9. The sum of the deviations from the average pixel value is obtained by taking the pixel values and weights at each position as shown in Eq. 10. If D which is the sum of the deviations is larger than the threshold value (*.), it is judged that the locality at the region is low. Otherwise, it is judged that the locality at the region is high. In case of high locality, the predicted pixel value PV_A is calculated as shown in Eq. 11. If the locality is low, the pixel value at that position becomes the predicted pixel value. A_1 through A_{12} are pixel values scanned in inverse s-order from the upper left corner of the cover image to the lower right corner. Using this pixel value prediction method, a prediction image is generated from top to bottom, left to right in inverse s-order. The predicted pixel value in the even-numbered row are obtained by replacing A with B in Eq. 9-11:

$$V = (A_6 + A_7 + A_8 + A_{12} + (A_3 + A_{11})*. + (A_2 + A_4 + A_5 + A_9)*. + (A_1 + A_{10})*.)/(4 + 2*. + 4*. + 2*.) \quad (9)$$

$$D = \text{abs}(V - A_6) + \text{abs}(V - A_7) + \text{abs}(V - A_8) + \text{abs}(V - A_{12}) + \{\text{abs}(V - A_3) + \text{abs}(V - A_{11})\}*. + \{\text{abs}(V - A_2) + \text{abs}(V - A_4) + \text{abs}(V - A_5) + \text{abs}(V - A_9)\}*. + \{\text{abs}(V - A_1) + \text{abs}(V - A_{10})\}*. \quad (10)$$

$$PV_A = (A_6 + A_7 + A_8 + A_{12})/4 \quad (11)$$

The cover image sequence C is generated by scanning the cover image in inverse s-order from top to bottom, left to right. The prediction image sequence P is generated by scanning the prediction image in the same manner. In the proposed algorithm, the difference sequence D is obtained as shown in Eq. 12 using the cover image sequence C and the prediction image sequence P:

$$D_i = C_i \text{ when } i = 0,\ C_i - P_i \text{ when prediction value was used},\ C_{i-1} - C_i \text{ otherwise} \quad (12)$$

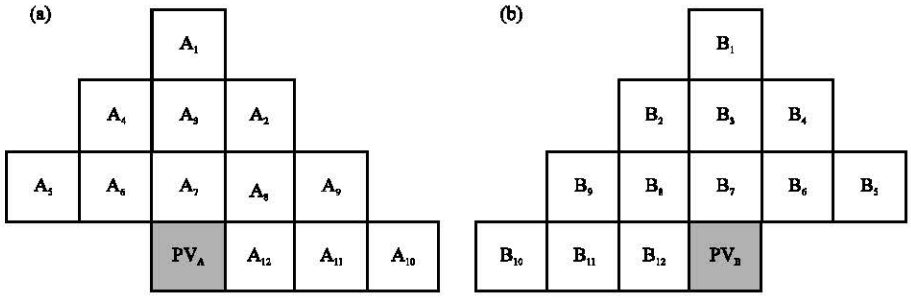

Fig. 1: Pixel value prediction using neighbor 12 pixel values: a) Odd-numbered line and b) Even-numbered line

As in the APD algorithm, a histogram for the difference sequence D is generated to determine PP_1, PP_2, CZP_1 and CZP_2. PP_1 is the largest Peak Point, PP_2 is the second largest peak point, CZP_1 is the Closest Zero Point to PP_1, CZP_2 is the Closest Zero Point to PP_2. The shifted difference sequence DS is generated as shown in Eq. 2-3 in APD algorithm. As in the APD algorithm, the sequence DE in which confidential data is concealed is generated as shown in Eq. 4 and the stego-image sequence S is generate as shown in Eq. 5. The difference from the APD algorithm in DE generation is that confidential data was compressed with Huffman coding and the compressed confidential data was used for DE generation. In order to increase the amount of data to be concealed, the data obtained by compressing confidential data by applying the Huffman coding technique is used to in Eq. 4. The data used in Eq. 4 consists of a Huffman code table and Huffman coded confidential data. The stego-image sequence S is generated as shown in Eq. 5. The stego-image in which confidential data is concealed is generated by arranging the data of the stego-image sequence in inverse s-order from top to bottom, left to right. The generated stego-image includes a Huffman code table, Huffman coded confidential data and an original cover image. If the image quality of the stego-image is similar to that of the original cover image, it is impossible to recognize whether the confidential data is embedded in the stego-image.

The procedure for extracting confidential data and original cover image from a stego-image without loss is as follows. Scan the stego-image in inverse s-order from top to bottom, left to right, to generate the stego-image sequence S. There are three modes for extraction. Extraction mode is determined at each position in sequence S from I = 0 to (image width)×(image height)-1 during the reverse procedure of the embedding.

Mode 1: In the case of prediction pixel value was not used at the current position and the just before position. This mode is the same as the APD algorithm. In the APD algorithm, only Mode 1 exists.

Mode 2: In the case of prediction pixel value was used at the current position, regardless of whether the prediction pixel value was used at the just before position.

Mode 3: In the case of prediction pixel value was not used at the current position but prediction pixel value was used at the just before position. The cover image sequence C and the sequence DE can be generated as follows. Let S_0 be C_0 and C_0 be DE_0. It is determined whether the position is a pixel value predictable position while increasing i. In the case of the pixel value predictable position, the extraction mode is determined by determining whether the pixel value of the cover image is used as the pixel value of prediction image at the corresponding position or whether the prediction pixel value is used.

In mode 1, DE and C are extracted in the same way as APD algorithm using Eq. 6 and 7. In mode 2, DE is restored using Eq. 6 and the cover image sequence C is restored using Eq. 13-14. In mode 3, DE is restored using Eq. 6 and the cover image sequence C is restored using Eq. 15. RF_i is a value calculated by Eq. 14. Equation 8 is applied to DE_i to extract confidential data embedded in the stereo-image. The extracted data is Huffman code table and Huffman coded confidential data.

The original confidential data can be extracted from the Huffman coded confidential data using the Huffman code table. The original cover image can be generated by placing the elements of the cover image sequence C from top to bottom, left to right in inverse s-order. The restored cover image and the confidential data are the same as the original cover image and the original confidential data:

$$C_i = RP_i + RF_i \tag{13}$$

$$RF_i = C_{i-1} - S_i - sd_j \text{ when } C_{i-1} - S_i \notin [PP_i + sd_j, CZP_j], \quad C_{i-1} - S_i \text{ in other cases} \tag{14}$$

$$Ci = RP_i - RF_i \tag{15}$$

RESULTS AND DISCUSSION

The data embedding performance of the proposed algorithm was evaluated using gray scale images with 512×512 pixels such as Barbara, Hara and sail boat images. The ., ., . values used in the experiment were 0.20, 0.15, 0.05, respectively. Experiments were performed while changing the threshold value (.) to 0, 5, 10 and 15. If the threshold value is 0, the proposed algorithm is the same as the APD algorithm. The abstract of this study was used as confidential data. The abstract was Huffman-coded to generate compressed confidential data. The Huffman code table used for compression was first embedded in the cover image and then the compressed confidential data was repeatedly embedded into the cover image. The cover images are shown in Fig. 2a and stego-images in which compressed confidential data is embedded using the proposed algorithm are shown in Fig. 2b-e. As shown in Fig. 2f-o because the visual quality of the stego-image is excellent, it is impossible to recognize whether the confidential data is embedded in the stego-image. Table 1 shows the experimental data. As shown in Table 1, since the average value of the prediction error is very small, the frequency at the peak point is greatly increased in the histogram of the difference sequence. Therefore, if the data is embedded using the proposed algorithm, data can be embedded at various levels and the number of embedded data bits can be greatly increased.

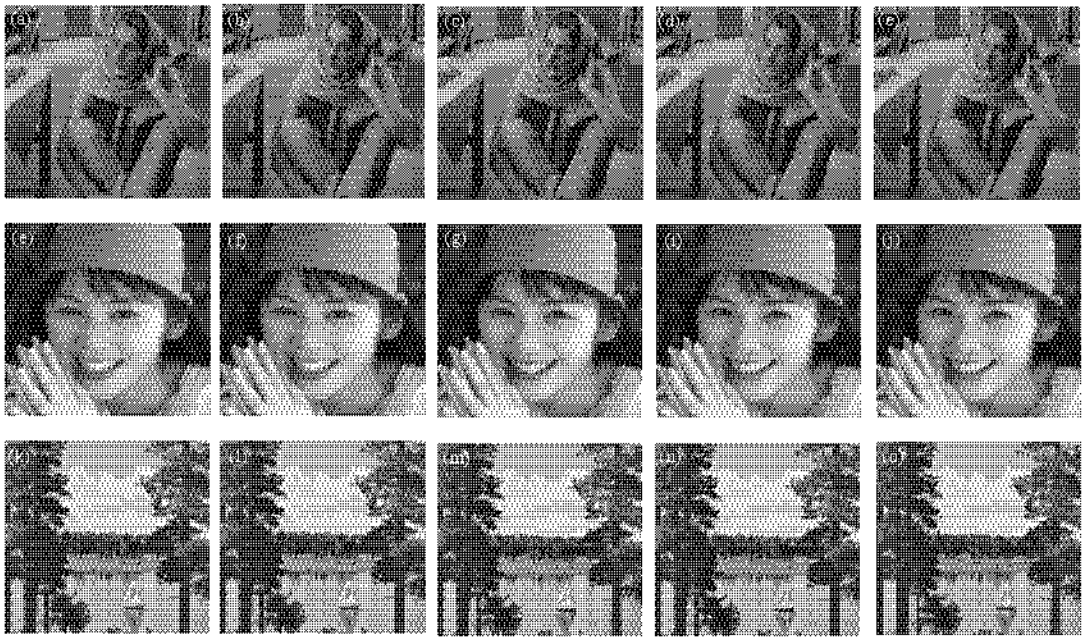

Fig. 2: Cover images, stego-images at various levels: a) (A-1) Barbaracover image; b) (A-2)APD = proposed (0); c) (A-3) proposed (5); d) (A-4) proposed (10); e) (A-5) proposed (15); f) (B-1) Haracover image; g) (B-2) APD = proposed (0); h) (B-3) proposed (5); i) (B-4) proposed (10); j) (B-5) proposed (15); k) ((C-1) Sail boat cover image; l) (C-2) APD = proposed (0); m) (C-3) proposed (5); n) (C-3) proposed (5); o) (C-5) proposed (15)

Table 1: Experimental results

Images	Algorithm	Embedded characters	PSNR (dB)	Prediction accuracy	Average prediction error	Increase rate of embedded characters (%)
Barbara	APD	4.1390	48.45	x	x	x
	Pro(5)	7.7980	45.16	18.98	1.67	88.4
	Pro(10)	8.7350	38.16	17.11	1.87	111.0
	Pro(15)	9.3090	35.58	15.80	2.08	124.9
Hara	APD	8.6070	48.82	x	x	x
	Pro(5)	15.8920	43.94	34.36	0.95	84.6
	Pro(10)	17.0920	38.28	25.80	1.37	98.6
	Pro(15)	18.1440	35.64	22.88	1.63	110.8
Sail boat	APD	3.8600	48.43	x	x	x
	Pro(5)	7.2440	46.00	27.64	1.22	87.7
	Pro(10)	7.8190	39.25	19.43	1.95	102.6
	Pro(15)	8.2420	34.43	15.32	2.62	113.5

CONCLUSION

In this study, an improved reversible data hiding algorithm was proposed. In the proposed algorithm, the performance of the APD algorithm is greatly improved by applying the pixel prediction method and the Huffman coding technique to the APD algorithm. Using the proposed algorithm, a large amount of data can be embedded in the cover image at various levels and the original cover image and confidential data can be completely restored from the stego-image. When Barbara was used as a cover image, the maximum increase rate of the data concealment rate of the proposed technique was 124.9%.

REFERENCES

Jung, S.M. and B.W. On, 2016. Reversible data hiding algorithm using spatial locality and the surface characteristics of image. J. Korea Soc. Comput. Inf., 21: 1-12.

Li, X., B. Li, B. Yang and T. Zeng, 2013. General framework to histogram-shifting-based reversible data hiding. IEEE Trans. Image Process., 22: 2181-2191.

Li, Y.C., C.M. Yeh and C.C. Chang, 2010. Data hiding based on the similarity between neighboring pixels with reversibility. Digital Signal Process., 20: 1116-1128.

Lo, C.C., Y.C. Hu, W.L. Chen and C.M. Wu, 2014. Reversible data hiding scheme for BTC-compressed images based on histogram shifting. Intl. J. Secur. Appl., 8: 301-314.

Ni, Z., Y.Q. Shi, N. Ansari and W. Su, 2006. Reversible data hiding. IEEE Trans. Circuits Syst. Video Technol., 16: 354-362.

Tsai, C.L., K.C. Fan, C.D. Chung and T.C. Chuang, 2004. Reversible and lossless data hiding with application in digital library. Proceedings of the 38th Annual 2004 International Carnahan Conference on Security Technology, October 11-14, 2004, IEEE, Albuquerque, New Mexico, ISBN:0-7803-8506-3, pp: 226-232.

Tseng, H.W. and C.P. Hsieh, 2008. Reversible data hiding based on image histogram modification. Imaging Sci. J., 56: 271-278.