A Modified Reversible Data Hiding in Encrypted Images Using Random Diffusion and Accurate Prediction

Ming Li, Di Xiao, Zhongxian Peng, and Hai Nan

A modified version of Zhang's reversible data hiding method in encrypted images is proposed in this letter. To make full use of spatial correlation in natural images, the former idea of block division is thoroughly abandoned, whereas the random diffusion strategy is used. Additionally, the fluctuation measurement of pixels containing embedded data is improved by accurate prediction. The experiment results reveal that our proposed method is superior to both Zhang's method and the later improved version proposed by Hong and others.

Keywords: Reversible data hiding, encrypted image, random diffusion, accurate prediction.

I. Introduction

Reversible data hiding is a technology to embed confidential information into a host image imperceptibly and enable the host image to be recovered without any error upon extraction of the embedded information. To meet the requirement of privacy protection simultaneously, the ability of embedding additional data reversibly into an encrypted image is demanded. However, this is difficult to achieve by using such common data hiding algorithms as [1]-[3]. Unlike some of the existing schemes concentrating on embedding information into the partial unencrypted data of the image [4], [5], in 2011, Zhang proposed a scheme for reversible data hiding in encrypted images [6].

Manuscript received Sept. 17, 2013; revised Oct. 4, 2013; accepted Oct. 16, 2013.

The work was supported by the Natural Science Foundation Project of CQ CSTC (Grant No. 2011jjjq40001).

Ming Li (phone: +86 15922618991, liming629@gmail.com), Di Xiao (corresponding author, xiaodi_cqu@hotmail.com), Zhongxian Peng (pengzhongxian@126.com) and Hai Nan (cqu.nn@163.com) are with the College of Computer Science, Chongqing University, Chongqing, China.

Zhang's scheme includes four phases: image encryption, data embedding, image decryption, and data extraction and image recovery. The gray value of each pixel, $p_{i,j}$, of the original image is first represented by 8 bits:

$$b_{i,j,c} = \left\lfloor \frac{p_{i,j}}{2^c} \right\rfloor \bmod 2, \quad c = 0,1,...,7. \tag{1}$$

Every pixel in the image is then encrypted in turn by stream cipher r to get the encrypted bit, as follows:

$$b'_{i,j,c} = b_{i,j,c} \oplus r_{i,j,c}, \tag{2}$$

where $\oplus$ is bitwise exclusive-or (XOR) and $r_{i,j,c}$ is the corresponding bit in r that needs to be XOR $b_{i,j,c}$. After image encryption, the additional data is embedded into the encrypted image by flipping the three least significant bits (LSBs) of the specified pixels in the divided non-overlapping blocks according to the data hiding key. The image decryption phase is similar to image encryption. By the XOR operation with the same r, the flipped LSBs will be equal to the flipped values of the corresponding LSBs of pixels in the original image, as shown in (3):

$$\overline{b'_{i,j,c}} \oplus r_{i,j,c} = \overline{b_{i,j,c} \oplus r_{i,j,c}} \oplus r_{i,j,c} = \overline{b_{i,j,c}}. \tag{3}$$

In the data extraction and image recovery phase, a fluctuation function is used to measure the smoothness of each block. Based on the spatial correlation in the natural image, the original pixels are smoother than the flipped pixels; so, the flipped pixels in the block can be found with the aid of the data hiding key. Thus, the embedded data can be extracted, and the image can be recovered. For further details, please refer to [6].

In 2012, by considering the pixel correlations at the border of

neighboring blocks, Hong and others [7] improved Zhang's work and lowered the extracted-bit error rate (ER) to some extent. However, both works are based on blocks, and the spatial correlation in the block cannot fully reflect the smoothness of a natural image, especially when the block is small. The limitation and centralization of local pixels will lead to the unreliability of fluctuation measurement, which is based on the smoothness of natural images. Therefore, in this letter, we modify Zhang's work by using random diffusion to eliminate the adverse effect of block division and employ accurate prediction [8] to improve the performance of fluctuation measurement. The experiment results reveal that our method is superior to those in [6] and [7].

II. Our Proposed Method

The image encryption phase and image decryption phase are identical with those of Zhang's work, but the data embedding phase, the data extraction and image recovery phase, are improved as follows.

1. Data Embedding

The pixels of the encrypted image are first divided into two disjoint sets, that is, black and white sets, as shown in Fig. 1. The white set is used to embed the additional data, and the number of pixels in the white set is denoted by w. The black set, which is not changed in the data embedding phase, is used to predict the pixel values in the white set.

A chaotic map, as shown in (4), is used to compute the positions for data embedding.

$$q_{n+1} = \alpha q_n (1 - q_n), \quad q_n \in (0,1). \tag{4}$$

With parameter $\alpha \in (3.5699456, 4]$, the system is in a chaotic state [9]. Initial parameter α and initial value q_0 are considered the data hiding key (herein, set $\alpha = 4$, $q_0 = 0.3$). First, array H including w random permuted integers ranging from 1 to w without repetition is obtained by sorting w random numbers generated by (4). Then, the pixel position (x_k, y_k) for embedding the k-th bit of the additional data can be obtained by (5).

$$\begin{cases} x_k = \left\lfloor \dfrac{2 \times H[k]}{N} \right\rfloor, \\ y_k = (2 \times H[k] - 1) \bmod N + ((N+1) \bmod 2) \times \left(\left\lfloor \dfrac{2 \times H[k]}{N} \right\rfloor \bmod 2 \right), \end{cases} \tag{5}$$

where N denotes the width of the image, and $H[k]$ denotes the k-th element in H. Clearly, the position (x_k, y_k) is distinctive and is randomly distributed in the white set. If the bit to be embedded is 0, flip three LSBs of the pixel value in position (x_k, y_k). If the bit is 1, do nothing.

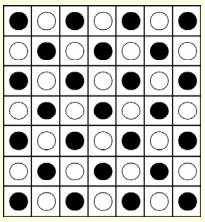

Fig. 1. Black and white sets separated from encrypted image.

To lower the extracted-bit ER, the additional data should be duplicated as many times as possible before embedding. However, the total length of the duplicated data cannot be more than w. Assume the bits of additional data constitute linear array A with length z, and the duplicated data is denoted by A'. Algorithm 1 is the duplication algorithm.

Algorithm 1: Additional data duplication.

For $i \leftarrow$ **0 to z–1 Do**

For $j \leftarrow$ **1 to** $\left\lfloor \frac{w}{z} \right\rfloor$ **Do**

$A'[i \times \left\lfloor \frac{w}{z} \right\rfloor + j] \leftarrow A[i+1]$;

End

End

2. Data Extraction and Image Recovery

After image decryption, the values of the unflipped pixels are not changed compared to the original pixels. Also, according to (3), the values of three LSB-flipped pixels are equal to the original pixel values with three flipped LSBs. There are two main steps in the data extraction and image recovery phase: accurate prediction and joint judgment.

A. Accurate Prediction

As shown in Fig. 2, each pixel in the white set has four neighboring pixels in the black set, and the four neighboring pixels, whose values are the same as the original pixels, can be used to predict the center pixel value in the white set. Pair 1 to

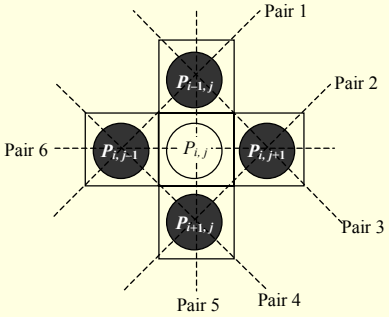

Fig. 2. Pixel pairs in one cross.

ETRI Journal, Volume 36, Number 2, April 2014
http://dx.doi.org/10.4218/etrij.14.0213.0449

pair 6 are six combinations of pairs. Each pair contains two pixels on the dotted line (for example, pair 1 contains $p_{i-1,j}$ and $p_{i,j-1}$). Then, divide the six pairs into three groups: (pair 1, pair 2), (pair 3, pair 4), and (pair 5, pair 6). Assuming $\hat{p}_{i,j}$ is the predicted value of the center pixel in the white set, then the accurate prediction algorithm of $\hat{p}_{i,j}$ is as follows.

Algorithm 2: Accurate prediction.
If (In group (pair 5, pair 6), one pair's difference value > 14, and the other pair ≤ 14) **Then**
$\hat{p}_{i,j} \leftarrow$ the mean value of the other pair (≤ 14)
Else
If (Either in group (pair 3, pair 4) or (pair 1, pair 2), one pair's difference value > 20, and the other pair ≤ 20) **Then**
$\hat{p}_{i,j} \leftarrow$ the mean value of the other pair (≤ 20)
Else
$\hat{p}_{i,j} \leftarrow (p_{i-1,j} + p_{i+1,j} + p_{i,j-1} + p_{i,j+1}) / 4$
End
End

Algorithm 2 is a slightly optimized version based on [8], and the numbers 14 and 20 in the algorithm are the given threshold values obtained by thorough experiments in [8]. For the border and corner pixels in the white set, the mean value of the existing neighboring pixels in the black set is used as $\hat{p}_{i,j}$.

B. Joint Judgment

Based on statistics, the original pixel value is much closer to $\hat{p}_{i,j}$ than the interfered one with three flipped LSBs. Assume $\hat{p}_{i,j}$ denotes the pixel value of the three flipped LSBs of $p_{i,j}$. Either $p_{i,j}$ or $\hat{p}_{i,j}$ must be the original pixel, and the other is the one with three flipped LSBs. Because the additional data was duplicated before embedding, the duplicated bits should represent a joint estimate of the original bits of the additional data.

To retrieve $A[n](n \in [1, z])$, assume t denotes $\left\lfloor \frac{w}{z} \right\rfloor$ if $\sum_{k=(n-1)\times t+1}^{(n-1)\times t+t} \left| p_{x_k,y_k} - \hat{p}_{x_k,y_k} \right| > \sum_{k=(n-1)\times t+1}^{(n-1)\times t+t} \left| \overline{p}_{x_k,y_k} - \hat{p}_{x_k,y_k} \right|$ and $\left\{ \overline{p}_{x_k,y_k} \right\}_{k=(n-1)\times t+1}^{(n-1)\times t+t}$ are the original pixels, and $\left\{ p_{x_k,y_k} \right\}_{k=(n-1)\times t+1}^{(n-1)\times t+t}$ must be flipped in the data embedding phase, so embedded bits $\left\{ A'[k] \right\}_{k=(n-1)\times t+1}^{(n-1)\times t+t}$ equal 0 (that is, $A[n]$ is 0). Otherwise, $\left\{ p_{x_k,y_k} \right\}_{k=(n-1)\times t+1}^{(n-1)\times t+t}$ are the original pixels, and $A[n]$ is 1. At last, all embedded data can be extracted successfully bit by bit in this way, and the obtained original pixels are collected to form the final recovered image.

III. Experiment Results

1. Test on Lena Image

Take image Lena (512 × 512) as an example. The length of the additional data is 4,096, which is the same as the number of embedded bits in the 8 × 8 block in [6]. As shown in Fig. 3, our scheme works well. The extracted-bit ER of our scheme is 0% (that is, all of the embedded data is exactly extracted), and the recovered image is perfect. It is much better than Zhang's 1.29% [6] and Hong and others' 0.34% [7].

2. Extracted-Bit Error Rate Comparison

The test images are all 18 grayscale images of different sizes obtained from the standard test images collection [10], including Baboon, Barbara, Boats, Bridge, Camera, Columbia, Couple, Crowd, Goldhill, Lake, Lax, Lena, Man, Milkdrop, Peppers, Plane, Woman1, and Woman2. For easy comparison, the data embedding capacity of our proposed method is set equal to [6] and [7], which is determined by the block size. When the block side length ranges from 40 to 8 in [6] and [7], step by 2, the 17 values of corresponding data embedding capacity between 144 and 4,096 are shown as the horizontal axis label of Fig. 4. A comparison of the average extracted-bit ER of 18 test images of Zhang's, Hong and others', and our proposed method is shown in Fig. 4(a), and that of the minimum and maximum ER is shown in Figs. 4(b) and 4(c), respectively.

Obviously, the average extracted-bit ERs of our proposed method are much lower than those in [6] and [7]. Notably, when the data embedding capacity does not exceed 529, the

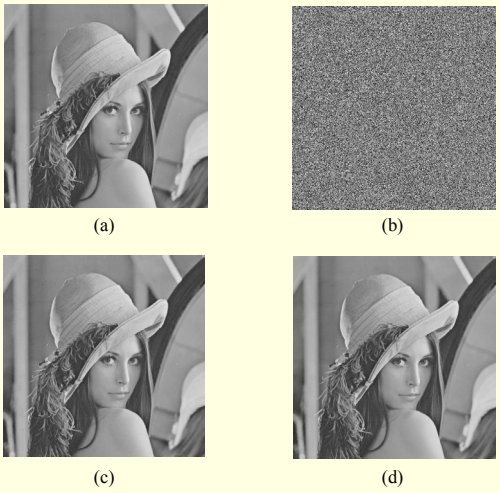

Fig. 3. (a) Original image, (b) encrypted image, (c) decrypted image containing embedded data, and (d) perfectly recovered image.

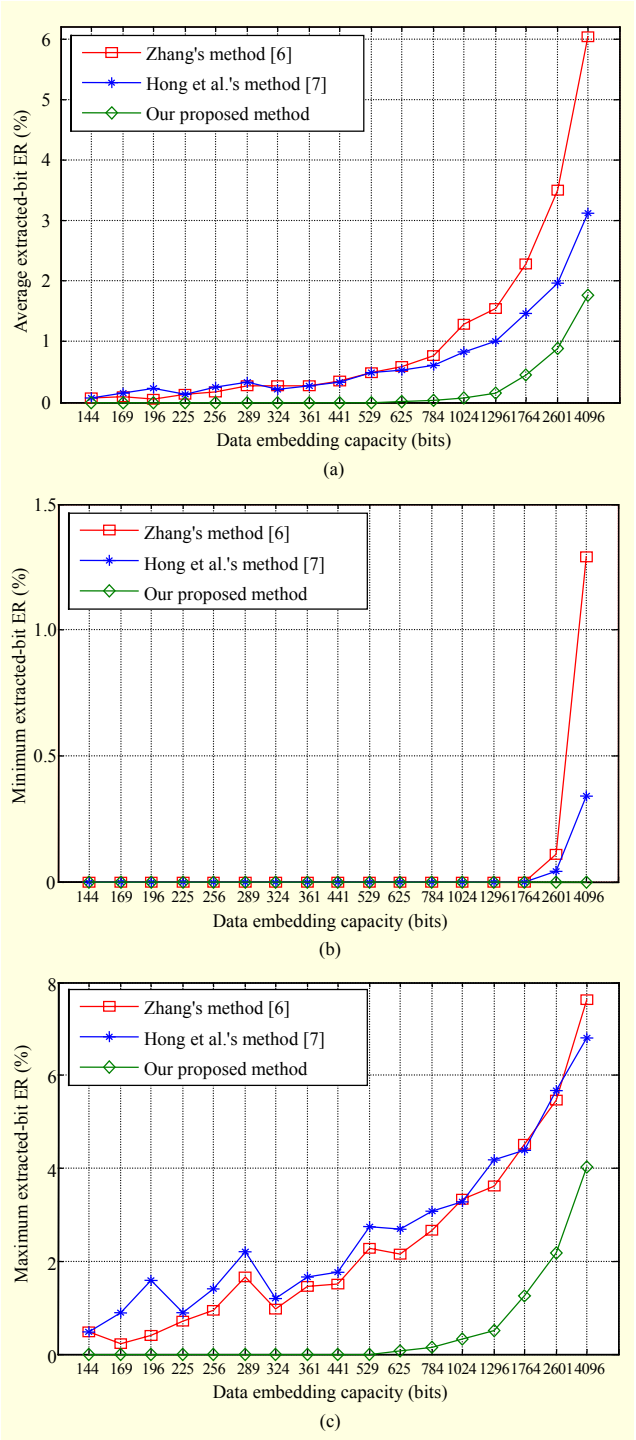

Fig. 4. Extracted-bit ER comparison: (a) average ER, (b) minimum ER, and (c) maximum ER.

average ERs using our proposed method equal 0. This means that all the test images are recovered without any error, and all of the hiding data is retrieved exactly. However, neither [6] nor [7] has an average ER of 0, even when the data embedding capacity takes the minimum value of 144 (that is, block size 40 × 40 in [6] or [7]). The minimum and maximum ER comparisons also ensure the superiority of our proposed method. Actually, our method outperforms the others for all images except Lax, which is a unique satellite image amongst the test images.

IV. Conclusion

Zhang's work was greatly improved in this letter. In the data embedding phase, the random diffusion strategy was used to eliminate the adverse effect of block division; in the data extraction and image recovery phase, we employed the idea of accurate prediction to improve the performance of fluctuation measurement. The experiment results show that our proposed method outperforms both Zhang's and Hong and others' methods. To meet the requirement of reversible data hiding, the recommended data embedding capacity is 529 bits.

References

[1] J. Tian, "Reversible Data Embedding Using a Difference Expansion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, Aug. 2003, pp. 890-896.

[2] Z. Ni et al., "Reversible Data Hiding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 8, Mar. 2006, pp. 354-362.

[3] X.-T. Zeng, Z. Li, and L.-D. Ping, "Reversible Data Hiding Scheme Using Reference Pixel and Multi-layer Embedding." *Int. J. Electr. Commun.*, vol. 66, no. 7, July 2012, pp. 532-539.

[4] S. Lian et al., "Commutative Encryption and Watermarking in Video Compression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 6, June 2007, pp. 774-778.

[5] M. Cancellaro et al., "A Commutative Digital Image Watermarking and Encryption Method in the Tree Structured Haar Transform Domain," *Signal Process., Image Commun.*, vol. 26, no. 1, Jan. 2011, pp. 1-12.

[6] X. Zhang, "Reversible Data Hiding in Encrypted Image," *IEEE Signal Process. Lett.*, vol. 18, no. 4, Apr. 2011, pp. 255-258.

[7] W. Hong, T.-S. Chen, and H.-Y. Wu, "An Improved Reversible Data Hiding in Encrypted Images Using Side Match," *IEEE Signal Process. Lett.*, vol. 19, no. 4, Apr. 2012, pp. 199-202.

[8] S. Kang, H.J. Hwang, and H.J. Kim, "Reversible Watermark Using an Accurate Predictor and Sorter Based on Payload Balancing," *ETRI J.*, vol. 34, no. 3, June 2012, pp. 410-420.

[9] K.T. Alligood, T.D. Sauer, and J.A. Yorke, *Chaos: An Introduction to Dynamical Systems*, New York: Springer-Verlag, 1996.

[10] Standard picture collection: grayscale images and color images. Accessed Apr. 19, 2013. http://media.cs.tsinghua.edu.cn/~ahz/digitalimageprocess/benchmark.htm