REVERSIBLE DATA HIDING SCHEME BASED ON PREDICTION ERROR SORTING AND DOUBLE PREDICTION

Ling-ling WAN^{1} , Fan $CHEN^{1}$, Hong-jie HE^{1} , Lei $ZHANG^{2}$

¹ the School of Information Science and Technology, Southwest Jiao-tong University, Chengdu, Sichuan ² Computer and Information Engineering College, Henan Normal University, Xinxiang, Henan

mrchenfan@126.com

ABSTRACT

To consider both hiding payload and stego-image quality, a reversible data hiding scheme is proposed in this work. The prediction error of each block of size 2×2 pixels is firstly computed and sorted to obtain the ordered sequence of image blocks according to the prediction errors absolute value. And then the first $\lfloor k/2 \rfloor$ blocks with the least absolute value of errors are selected and used to hide the secret data with the length of k bits. Two bits secret data are hidden in one block by twice prediction error expansion (PEE) with different reference pixel. These two strategies not improve the maximum hiding payload, but also improve the quality of stego-image carrying the same bits of secret data. Experimental results demonstrate the proposed algorithm possesses superior properties in terms of hiding capacity and image quality.

Index Terms—Reversible data hiding, error sorting, prediction error expansion, high capacity

1. INTRODUCTION

Reversible data hiding refers to extract the hidden data correctly, at the same time, the condition to recover the host image. It as an important branch of data hiding, attracted much attention in recent years, is widely used in medical, military and law, etc, and it has important theoretical significance and practical value. Reversible data hiding algorithm is mainly divided into three categories: lossless compression ^[1-2], histogram shifting ^[3-8] and difference expansion ^[9-17].

Lossless compression does not take into account the image itself, which leads to the low coding rate and is difficult to meet the requirements of image processing. So, in later years, many approaches have been proposed, Tian's difference expansion $(DE)^{[9]}$ as a kind of classical reversible data hiding scheme is known, the large auxiliary data need to be hide to increase image distortion. Thodi ^[4] proposed histogram shifting method to solve the problems of Tian's DE algorithm and improved the algorithm performance. PEE^[13-17] is used to improve DE scheme. Many

mathematical models are used in the PEE system, such as partial differential ^[13], interpolation ^[14], least square method ^[17] etc, to predict the center pixel in order to decrease the error, but the reversibility of algorithm is guaranteed with the help of the auxiliary data. The auxiliary data can greatly reduce the effective capacity of the algorithm. Therefore, how to reduce the auxiliary information and the best is to have no auxiliary information become more important. Coltuc's [16] algorithm employed the reversibility of mathematical formula to achieve reversibility of algorithm without the auxiliary information, and also after hiding a certain amount of secret data, it can still maintain a good quality of image. However, due to the limitations of the algorithm is only a quarter of the image size.

In this paper, we propose a reversible scheme by using prediction error sorting and double expansion, achieving the reversibility of image as same as hiding large number of secret data and also maintain the image quality. We will briefly review prediction error expansion scheme in Section 2. Section 3 will present the proposed scheme. The experimental results will be discussed in Section 4. Finally, conclusions will be given in Section 5.

2. PREDICTION ERROR EXPANSION AND ANALYSIS

The prediction error expansion is efficient for hiding data at the same time not too much influencing the image quality. Next, we can introduce the basic principle of the prediction error ^[16] expansion below. We take an image block such as Fig.1 as an example.

<i>a</i> ₁	<i>a</i> ₂
<i>a</i> ₃	a_4
1 an image block with four nive	

Fig.1 an image block with four pixels

Pixel a_4 is estimated as

$$a_4' = a_2 + a_3 - a_1 \tag{1}$$

Let b be the data bit to be hidden, and let p be the prediction error. Let a_4 be reference pixel, i.e. $p = a_4 - a_4'$. Let $p_b = p + b$. The classical prediction error expansion reversible watermarking additively embeds p_b into the current pixel. Let d be a fraction of p_b , and let us transform a_4 and its entire context as follows:

$$A_{4} = a_{4} + p_{b} - 3d \qquad A_{3} = a_{3} - d$$

$$A_{2} = a_{2} - d \qquad A_{1} = a_{1} + d \qquad (2)$$

The square error is $e^2 = (A_4 - a_4)^2 + (A_3 - a_3)^2 + (A_2 - a_2)^2 + (A_1 - a_1)^2$. When $\partial E / \partial d = 0, e^2$ gets minimum error value $p_b^2/4$. According to the minimum error value, we can easily knowledge that p_b^2 is the smaller, e^2 is the smaller. Therefore, when we hide secret data, it prefers to choose smaller absolute value of p_b to expand in order to achieve better image quality.

The classical prediction error expansion reversible watermarking additively embeds p_b into the current pixel, however, we transform a_4 and its entire context. In order to split p_b as evenly as possible in four parts, we take

$$\begin{aligned} d_4 &= \lfloor p_b / 4 \rfloor & d_3 &= \lfloor (p_b + 1) / 4 \rfloor \\ d_1 &= \lfloor (p_b + 2) / 4 \rfloor & d_2 &= \lfloor (p_b + 3) / 4 \rfloor \end{aligned}$$

the stego-image can be computed as shown below

$$\begin{array}{ll} A_4 = a_4 + d_4 & A_3 = a_3 - d_3 \\ A_2 = a_2 - d_2 & A_1 = a_1 + d_1 \end{array} \tag{3}$$

As a result, $\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$ is a block of the stego-image. By changing the value of each pixel, we can obtain the stego-

image.

The algorithm transforms is reversible. However, from the hiding model, although once expansion can ensure better image quality, the maximum hiding capacity is inadequate. We can see that each block can hide one bit data. If we want to hide more secret data, the better way for us is to employ expansion technique again. We can observe that four pixels of a block, there are two pixels to plus the difference and the other to subtract. We employ expansion technique again in same block, the image quality is better than in different block. Based on the capacity and quality into consideration, we propose reversible data hiding scheme of prediction error sorting and double expansion.

3. PROPOSED REVERSIBLE DATA HIDING SCHEME

In this section, we proposed reversible data hiding scheme with prediction error sorting and double expansion. Firstly, the prediction errors of all blocks - are computed and sorted according to the prediction errors absolute value. Secondly, in order to increase the hiding capacity, the expansion technique is employed with two times. The goal of the proposed scheme is to provide a higher hiding capacity while keeping the good quality of the stego-image.

3.1. Data hiding procedure

The detailed depiction of the hiding procedure is as follows:

Step 1: Divide $I(m \times n)$ into non-overlapping 2×2 blocks $\{I_1, I_2, ..., I_N\}$, and N is the total number of pixel blocks, $N = (m \times n)/4$.

Step 2: Let the secret data *S* include *k* bits $\{s_1, s_2, ..., s_k\}$, where $1 \le k \le 2N$, and s_i (i=1,2,...,k) denotes the *i*th bit of secret data. Since one block carries two bit secret key in the proposed scheme, the number of image blocks which are used to embedded the secret data equals to $\lfloor k/2 \rfloor$. Note that the last block is embedded one bit if *k* is odd.

Step 3: For all blocks $\{I_1, I_2, ..., I_N\}$ in the original image, the prediction errors are firstly computed according to (1), in which the fourth pixel is considered as the reference pixel. And then all blocks are sorted according to the corresponding absolute value of prediction errors. Lastly, the $\lfloor k/2 \rfloor$ blocks with the least errors absolute values are selected to be hidden the secret data.

Step 4: For each selected blocks, two bits secret data is embedded. Firstly, the first bit is hidden by adopting the PEE method and obtain the temporary stego-block. And then the stego-block is generated by reusing the PEE method to embed the second bit in the temporary stegoblock. The difference using the PEE method twice is that the reference pixel is not same. Specifically, the reference pixel of the first PEE is the fourth pixel in block, and that of the second PEE is the second pixel in block.

3.2. Data extraction and recovery procedure

The data extraction and recovery procedure is relatively simple.

Step 1: The same as the Step 1 in data hiding procedure, the stego-image is divided into N non-overlapping sized blocks noted as $\{I_1^2, I_2^2, \dots, I_N^2\}$. $N = (m \times n)/4$.

Step 2: By employing the inverse transformation of prediction error expansion technique and recover the image I_1 , and record extracted data $w^1 \cdot w^1 = \{w_i^1 | i = 1, 2, ..., N\}$, $w_i^1 = \{1, 0\}$, w_i^1 is equal to 1 or 0. Employ inverse transformation again to recover image *I* based on I_1 . Extracted data $w^2 = \{w_i^2 | i = 1, 2, ..., N\}$, $w_i^2 = \{1, 0\}$,

Step 3: According the recover image I, calculate error values.

Step 4: Error values and k are known, the original secret data $s_j = \{(w_i^1, w_i^2) | i = 1, 2, ..., N\}, j = 1, 2, ..., k/2$. but *i* depends on errors values and k. $S = \{s_1, s_2, ..., s_i\}$.

Next, we give an example to explain our algorithm process shown in Fig.2. Get an image I with size of -4×4 pixels as shown in the upper part of Fig.2. Here, we do not discuss the case of hiding maximum capacity, because the PEE is directly adopted twice for all blocks in this case. We

focus on the situation the hiding capacity is less than maximum capacity. According to first part of Fig.2, the four 2×2 blocks in the original image is computed to generate four errors, respectively -1, -2, 0 and 0. And we need to hide 6 bits data, only three blocks with blue areas can be needed.

According to our rules, the blocks with error of 0, 0 and -1 can be selected to hide secret data. If the error of many blocks is equal, and the hidden data is less than the number of blocks, we will follow the scanned order of the image to hide sequentially. After two expansions, we find the values



Fig. 2 The diagram of data hiding

of some pixels are changed, such as 163 to 164. We can look at the pixel which is how to change from the lower half of Fig.2. The lower right corner block is selected to hide 2 bits data 1 and 0. In first expansion, the prediction error is 0, reference pixel is 159, and p_b is 1. The conclusion is the upper right corner pixel minus 1. The same calculation process is in second expansion, but there is little difference, we select different reference pixel such as the right corner pixel 152 in second expansion. Using the method, the image quality is a little better than using the same reference pixel, such as 0.28dB (Lena).

Next, the extracted secret data and recovering process are very simple. According the first part, after the first reverse transformation, the extracted data is 1, 0, 1 and 0 in scanning order. And after second reverse transformation, the data is 0, 0, 0 and 1. The original image is recovered, and the error can be calculated -1, -2, 0 and 0. Because we hide 6 bits data, the blocks which have smaller errors preferentially are selected. Therefore, -1, 0 and 0 can be selected, the blocks of hidden secret data are known. According to the selection rules of secret data, the extracted data can be reconstructed into complete secret data. Up to this point, the entire process of hiding, extracting and recovering is over.

4. EXPERIMENTAL RESULTS

The proposed algorithm possesses many advantages what other algorithms don't have. Next, we will prove these unique merits of proposed algorithm by giving concrete experimental data. Here, we first present the evaluation index of algorithm performance. The PSNR vs. payload is used for comparison. PSNR is calculated by

$$PNSR = 10 \cdot log_{10} \left(\frac{255^2}{MSE}\right) \tag{4}$$

where MSE (mean square error) is calculated by

$$MSE = \frac{\sum_{i=1}^{512} \sum_{j=1}^{512} (x'(i,j) - x(i,j))^2}{512 \cdot 512}$$
(5)

To validate the effectiveness of proposed algorithm, we choose six natural images such as "Goldhill", "barnara", "Baboon", "Peppers", "Lena" and "F-16" as shown in Fig. 3.



Fig.3 Six natural images for testing: Barnara, Baboon, Peppers, Goldhill, Lena, F-16



Fig.4 PSNR vs. Hiding Capacity for six images

Next, we will analyze the specific performance of the algorithm by experimental data. From the Fig.4, we give six natural images with 512*512 sizes to show the performance of proposed algorithm. These images include smooth images such as 'Lena' and texture image such as 'Baboon'. These curves are roughly the same direction. But from top to bottom, the rightmost of fourth curve has a relatively steep drop. The reason is because the 'Barbara' image has more and larger absolute value of error value. The overall trend is quite satisfactory.





We compare our proposed method against literature [16] because our proposed algorithm employs prediction error expansion technique which is from literature [16] to hide the secret data. In Coltuc's algorithm, it don't consider the overhead, we give the two results of [16], one is casually to choose a place to hide by scanning order, and the other is smaller errors to hide by selecting image blocks not to consider overhead. The original prediction error expansion technique only expands once to hide one bit secret data in each block. In accordance with the rules of blocking, the size of hiding secret data is equal to a quarter of the image size. The hiding capacity is seriously limited. Therefore, the double expansion technique is proposed.

Also, when we hide secret data, the smaller prediction error can be preference. We keep the same conditions with the literature [16]. The combination of smaller error preference and double expansion are greatly enhancing the performance of the algorithm. At the Fig.5, we can give an example to analyze specifically such as 'Baboon' image in Fig.5 (2) and to choose a specific capacity value 0.25bpp to analyze. we need keep the same condition. if want to hide 0.25 bpp secret data, the whole image can be used. Each block hides one bit data. However, proposed algorithm does not need to use the entire image. We can only employ an half of entire image to expand twice to hide 0.25 bpp. From the experimental result, the proposed algorithm is better than Coltuc's algorithm about 3dB in PSNR. Obviously, PSNR of 'Lena' image is more higher. In fact, with easy to understand, the 'Lena' image has more smooth block and with smaller prediction error, image quality can maintain better very naturally. Overall, the performance of this algorithm is better than the performance of comparative literature.

5. CONCLUSION

In this paper, a reversible data hiding based on error sorting and double prediction errors expansion scheme has been proposed for natural images. The prediction error expansion is used to hide abundant secret data, and in order to maintain the image quality better, the prediction error need to be sorted to hide data from small to big. In case of not affecting other image blocks, the blocks which had been hidden data will be expanded again to hide more data. As a result, our proposed data hiding scheme has shown better payload-PSNR performance compared with the existing prediction error expansion scheme.

6. ACKNOWLEDGEMEN

This research was supported by the National Natural Sciences Foundation of China (No. 61373180), the Science and Technology Innovation Talent Project of Sichuan province (No. 2014-058), and the Science and Technology Key Project of Henan (No. 142102210451).

7. REFERENCES

[1] M. Goljan, J. Fridrich, and R. Du, "Distortion free data embedding for images," Inf. Hiding, pp. 27-41, Apr. 2001,

[2] M.U. Celik, G. Sharma, A.M. Tekalp, and E. Saber, "Lossless generalized LSB data embedding," IEEE Transactions on image processing, pp. 253–266, Vol. 14, No. 2, Feb. 2005

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

[4] D.M. Thodi and J.J. Rodríguez, "Expansion embedding techniques for reversible watermarking," IEEE Transactions on image processing, pp. 721-730, Vol. 16, No. 3, March 2007

[5] X.L. Li, B. Li, B. Yang, and T. Zeng, "General framework to histogram shifting based reversible data hiding," IEEE Transactions on image processing, pp. 2181-2191, Vol. 22, No. 6, June 2013

[6] Z.H Wang, C.F. Lee, and C.Y. Chang, "Histogram shifting imitated reversible data hiding," The Journal of Systems and Software, pp. 315–323, 2013

[7] P. Tsai, Y.C. Hu, and H.L. Yeh, "Reversible image hiding scheme using predictive coding and histogram shifting," Signal Processing, pp. 1129–1143, 2009

[8] H.T. Wu, J.W. Huang, "Reversible image watermarking on prediction errors by efficient histogram modification," Signal Processing, pp. 3000–3009, 2012

[9] T. Jun, "Reversible data embedding using a difference expansion [J]," IEEE Transaction on circuits and systems for video technology, pp. 890-896, 2003

[10] G. Coatrieux, W. Pan, N.C. Boulahia, and F. Cuppens, "Reversible watermarking based on invariant image classification and dynamic histogram shifting," IEEE Transaction on information forensics and security, pp. 111-119, 2013

[11] X. Zhang, "Reversible data hiding with optimal value transfer," IEEE Transactions on multimedia, pp. 316-325, Vol. 15, No. 2, Feb 2013

[12] B. Ou, X.L. Li, Y. Zhao, R. Ni and Y.Q. Shi, "Pairwise prediction error expansion for efficient reversible data hiding," IEEE Transactions on image processing, pp. 5010-5021, Vol. 22, No. 12, Dec 2013

[13] B. Ou, X.L. Li, Y. Zhao, and R. Ni, "Reversible data hiding based on PDE predictor," The Journal of Systems and Software, pp. 2700–2709, 2013

[14] L. Luo, Z. Chen, M. Chen, X. Zeng, and Z. Xiong, "Reversible image watermarking using interpolation technique," IEEE Transactions on information forensics and security, pp. 187-193, Vol. 5, No. 1, March 2010

[15] D. Coltuc, "Improved embedding for prediction based reversible watermarking," IEEE Transactions on information forensics and security, pp. 873-882, Vol. 6, No. 3, Sept 2011

[16] D. Coltuc, "Low distortion transform for reversible watermarking," IEEE Transaction on image processing, pp. 412-417, Vol. 21, No. 1, Jan 2012

[17] I.C. Dragoi, D. Coltuc, "Local prediction based difference expansion reversible watermarking," IEEE Transactions on image processing, pp. 1779-1790, Vol. 23, No. 4, April 2014