



Efficient Pairwise Reversible Data Hiding Technique Using in Image Authentication

¹Nidhi Mishra, ²C. M. Tiwari

¹MCGGV Chitrakoot, Satna (M.P) India

²Dept. of Physics, APS University, Rewa (M. P.) India

Abstract: Some area such as the law enforcement, medical and military use image system. It is important to restore the original image without any distortions. The watermarking techniques complacency those requirements are referred to as 'reversible watermarking'. Reversible watermarking is mainly data hiding technique. It restores the real image without any distortion. It has become a vital research area in the field of data security, confidentiality, and integrity. In this paper we propose a novel reversible data hiding framework algorithm which is using prediction-error expansion (PEE) in image authentication. This algorithm makes use of prediction errors to embed data into an image. We propose that every two adjacent prediction-errors compositely make pair and consider to generate a sequence which consisting of prediction-error. A sorting technique is used to make an entry of prediction error pairs. This is based on connection of prediction error. Then, based on the sequence and the resulting 2D prediction-error histogram, a more efficient embedding strategy, can be designed.

Keywords: Pairwise prediction-error expansion (pairwise PEE), reversible data hiding (RDH), 2D prediction-error histogram (2D PEH), Authentication, pixel selection.

I. INTRODUCTION

Image Authentication is used to evidence that image is really what the user deems it is. For example, during image watermarking of patient's diagnostic image may occur errors in diagnosis and treatment, which may direction to possible life-blusterous outcome. Thus, to take over the problem of happening of artefacts and to making zero distorted or noise free watermarked medical images. Several reversible watermarking schemes have been initiated. Data hiding offering a way to situate data into cover medium for the motive of ownership fingerprinting, protection, secret communication authentication, and annotation [1]. In most data hiding algorithms, the cover data is ruinous abiding and cannot be exactly or same restored after the embedded message is extracted. At the latest, a new data hiding technique, that is, reversible data hiding (RDH) [2], is proposed. With the help of RDH the cover data and the embedded message both we can be extracted from the marked content.

Reversible data hiding techniques wrap data in a host signal. For example, extraction allows of the original host signal and extracts an image also the embedded message. There are two main things are required for reversible data hiding techniques: the embedding capacity should be huge; and distortion should be minimum. These two requirements strife with each other. Generally, a higher embedding capacity results a higher degree of distortion. An improved Technique interpolates the same capacity with lower distortion or vice versa.

Until now, various RDH algorithms have been proposed, e.g., compression based algorithms [3] [4], difference expansion (DE) based algorithms [5] [6] [7], histogram shifting (HS) based algorithms [8] [9], prediction-error expansion (PEE) based algorithms [10]–[15] and integer-to-integer transform based algorithms [16], some other RDH analysis have also been given etc. all of these techniques, PEE is widely applied nowadays because of its efficient capacity distortion trade-off.

PEE was first proposed by Thodi and Rodriguez [10]. They were proposed that the difference between the pixel and its prediction is expanded for data embedding. When they compared between DE and HS based methods, PEE has a better performance. Hence the derived prediction-error histogram (PEH) is more rapidly distributed. Afterwards, Hu *et al.* [11] proposed PEE with an effective location map of smaller size, and thus growth of the capacity. Sachnev *et al.* [12] were proposed to sort prediction errors ethically to the local variance. Then prediction errors are process in the sorted order. Compared with former methods, the sorted prediction-errors embedding introduce minimum distortion especially for low capacities. Li *et al.* [14] were proposed to adaptive embed method which selects two bits in a smooth pixel and one bit in a rough one. Hereby they obtained a better execution for high capacities. Coltuc [15] proposed a novel PEE which was based on transformation of modifying. This modification was not only the current pixel but it also applies its next three context pixels. Since the total

changes on four pixels are sure to be smaller than that of only modification one pixel. This method introduces less distortion than the other ones which based on high-performance predictors

The prior PEE methods mainly focus on make use of inter pixel correlations, but have not contemplate the correlations within next prediction-errors. That is, a prediction-error may be presumption from its neighbors. And such type of correlations can also be make use to afore reduce the distortion. Yet, the one-dimensional prediction-error histogram (1D PEH) engaged in the conventional PEE, as a low-dimensional projection of image data. It is here upon vital to make a new pattern of PEE in a higher dimensional to better exploit for correlations.

In this paper, we proposed a novel RDH framework that called pairwise PEE. In resist to the prior PEE methods, we take every two adjacent prediction-errors firstly. Then generate a sequence of unit which consists of prediction-error pairs. Afterward we obtain a two-dimensional prediction-error histogram (2D PEH). Finally, embed data by using pairwise PEE, by shifting or expanding the bins of 2D PEH. Comparing between the conventional PEE, and pairwise PEE. And pairwise PEE is better exploit image redundancy than conventional PEE and achieves an improved performance. In addition, a refined pixel-selection technique is used to firstly process of selection. The pixels location select in smooth image regions, and this may further develop the embedding performance.

The remaining part of this paper is organized as follows. In Section II, the main idea of pairwise PEE is introduced. Then, the proposed RDH method is described in details in Section III. Finally, Section IV concludes this paper.

II. RDH FRAMEWORK BASED ON PAIRWISE PEE

In this segment, the basic principle of the conventional PEE is used for 1D PEH is scrutiny at first in Section II-A. Then, the advantage of 2D PEH over 1D PEH in designing efficient RDH is explored in Section II-B. Finally, the proposed pairwise PEE technique used for 2D PEH is detailed in Section II-C. The effectiveness of pairwise PEE in terms of capacity-distortion behavior is also discussed in Section II-C, and show that the proposed embedding technique normally represents a better performance rather than the conventional PEE.

A. Conventional PEE

The PEE embedding procedures are having following steps.

1) First of all predict image pixels which are gain a prediction-error sequence. Pixels are choosing under specific scan sequencing. The cover image pixels are together into a one-dimensional sequence as $(x_1, x_2, x_3, \dots, x_N)$. Then, we are using a predictor to specified the prediction of x_i notified as $\bar{x}_i$. Next, the prediction-error calculated by $e_i = x_i - \bar{x}_i$ (suppose here for simplicity that $\bar{x}_i$ is an integer). Finally, the prediction-error series $(e_1, e_2, \dots, e_N)$ are obtained.

2) Produce the prediction error histogram (PEH) by counting the repetition of prediction-errors. That is, the PEH is determined as

$$h(k) = \#\{1 \leq i \leq N: e_i = k\} \quad (1)$$

Where # denotes the cardinal number of a set. Usually, the PEH execute a Laplacian-like distribution centered at 0 or close to 0. When the PEH is more sharp dispenses, the distortion is lower for embedding the same amount of bits.

3) Modifying through shifting and expansion in PEH and embed data. Specifically, for each prediction-error e_i . It is expanded or shifted as e'_i as

$$e'_i = \begin{cases} 2e_i + b, & \text{if } e_i \in (-T, T) \\ e_i + T, & \text{if } e_i \in (T, +\infty) \\ e_i - T, & \text{if } e_i \in (-\infty, -T) \end{cases} \quad (2)$$

Where T is a parameter value of capacity-dependent integer and $b \in \{0, 1\}$ is to be embedded data bit. Here, the bins in $[-T, T]$ are enlarged to embed data and those in $(-\infty, -T) \cup [T, +\infty)$ are shifted from outside to create vacancies. Finally, each pixel value x_i is improved to $x'_i = \bar{x}_i + e'_i$ to obtain the marked image

For above PEE embedding, the embedding capacity EC is

$$EC = \sum_{k=-T}^{T-1} h(k) \quad (3)$$

And the embedding distortion ED in terms of l2-error can be formulated as.

$$ED = \sum_{k=-T}^{T-1} \left(k^2 + k + \frac{1}{2} \right) h(k) + T^2 \left(\sum_{k=-\infty}^{-T-1} h(k) + \sum_{k=T}^{+\infty} h(k) \right) \quad (4)$$

In the PEE extraction procedure, the original prediction-error e_i are recoverable from the present marked prediction-error e'_i as

$$e_i = \begin{cases} \lfloor e'_i/2 \rfloor, & \text{if } e'_i \in [-2T, 2T] \\ e'_i - T, & \text{if } e'_i \in [2T, +\infty) \\ e'_i + T, & \text{if } e'_i \in [-\infty, -2T] \end{cases} \quad (5)$$

Where $\lfloor \cdot \rfloor$ is the floor function. The embedded bits are extracted as the least significant bit (LSB) of those $e'_i \in [-2T, 2T]$. Finally, the cover image is restored using the recovered prediction-errors. To guaranteed the reversibility that, the prediction values used in extraction should be the same as it is in embedding.

B. 2D PEH

We first show that the prediction-error series $(e_1, e_2, \dots, e_N)$ used in PEE. It is not independently distributed. In fact, adjacent prediction-errors are normally highly correlated. For example, the correlation of two adjacent prediction-errors e_i and e_{i+1} , is given. This is computed by

$$\frac{|\sum_{i=1}^{N-1} (e_i - \bar{e})(e_{i+1} - \bar{e})|}{\sqrt{\sum_{i=1}^{N-1} (e_i - \bar{e})^2} \sqrt{\sum_{i=1}^{N-1} (e_{i+1} - \bar{e})^2}} \quad (6)$$

is 0.213 for the standard 512×512 sized gray-scale image Lena, 0.294 for Baboon, 0.293 for Airplane and 0.575 for Barbara, where $\bar{e}$ is the mean of $(e_1, e_2, \dots, e_N)$. This correlation coefficient in equation (6) ranges from 0 to 1, and the larger the value, the more correlated the two prediction-errors. It is verified that e_i and e_{i+1} are correlated to some range, or even highly correlated for some specific images (e.g., Barbara). Here in this method, rhombus prediction is used to generate the prediction error sequence that means the prediction of a pixel is the average value of its four adjacent neighboring pixels. To improve the embedding performance, previous PEE based works mainly focus on how to better make use of inter-pixel correlations to make a sharply distributed PEH. When PEH is generated, the prediction-errors are improved individually for data embedding. However, the correlations within prediction-errors are not making use of in this individually modification strategy. Hence, to obtain a better performance in RDH, we propose to consider 2D PEH where the correlations within prediction-errors can be utilized.

We now explain the 2D PEH and show its advantage in RDH over the 1D PEH. By considering every two adjacent prediction-errors together, the sequence $(e_1, e_2, \dots, e_N)$ can be transformed into a new one $(e_1, e_2, \dots, e_{N/2})$ with $e_i = (e_{2i-1}, e_{2i})$ (suppose for simplicity that N is even), and the associated 2D PEH is

$$g(k_1, k_2) = \#\{1 \leq i \leq \frac{N}{2} : e_{2i-1} = k_1, e_{2i} = k_2\} \quad (7)$$

For convenience, the sequences $(e_1, e_2, \dots, e_N)$, $(e_1, e_3, \dots, e_{N-1})$ and $(e_2, e_4, \dots, e_{N/2})$ are denoted as E , E_{odd} , E_{even} and E^* , respectively. Clearly, it can be assumed that

$$H(E) = H(E_{odd}) = H(E_{even}) \quad (8)$$

Where the function H calculates the entropy of a distribution. Then, the entropy of E^* satisfies

$$H(E^*) = H(E_{odd}, E_{even}) \leq H(E_{odd}) + H(E_{even}) = 2H(E) \quad (9)$$

And the equality holds if and only if E_{odd} and E_{even} are independent. However, as mentioned above, E_{odd} and E_{even} are usually correlated.

Conventional PEE can be implemented in a similar way by modifying the 2D PEH. This structure of PEE gives us a new way to design RDH. By changing the expansion and shifting techniques used in 1D PEH to 2D PEH, the conventional PEE can be modified and improved in a more reasonable histogram modification strategy on 2D PEH.

C. Pairwise PEE

The result of 2D PEH can better perform the complex dependencies in image data than 1D PEH. It is precious to make use of 2D PEH to design RDH. In this work, the histogram change strategy based on 2D PEH is called Pairwise PEE. Where the data embedding is applied by expanding and shifting the bins of 2D PEH. Pairwise PEE is a natural expansion of the conventional PEE in upper dimensional space. Based on 2D PEH, various histogram correction strategies are designed with different performance.

The capacity-distortion performance in proposed Pairwise PEE is examined as follows. The conventional PEE capacity and the proposed Pairwise PEE capacity, denoted as EC_{con} and EC_{pro} , can be written as

$$EC_{con} = 2 \sum_{e \in A} g(e) + \sum_{e \in C} g(e) \quad (10)$$

And

$$EC_{pro} = \log_2 3 \sum_{e \in A} g(e) + \sum_{e \in B} g(e) + \sum_{e \in C} g(e) \quad (11)$$

Where g is used in 2D PEH explained in equation (7). And the other term, the distortion in views of l^2 -error, the conventional PEE distortion and the proposed pairwise PEE distortion, formulated as ED_{con} and ED_{pro} , can be formulated as

$$ED_{con} = \sum_{e \in A} g(e) + 2 \sum_{e \in B} g(e) + \frac{3}{2} \sum_{e \in C} g(e) + 2 \sum_{e \in D} g(e) \quad (12)$$

And

$$ED_{pro} = \frac{2}{3} \sum_{e \in A} g(e) + \sum_{e \in B} g(e) + \frac{3}{2} \sum_{e \in C} g(e) + 2 \sum_{e \in D} g(e) \quad (13)$$

In equation (12) and (13), comparing both the conventional PEE and proposed Pairwise PEE, we saying that the pairwise PEE distortion is minimize.

III. PROPOSED PAIRWISE PEE SCHEME

The equilateral parallelogram pattern [12] is used in our work for prediction

A. Pixel Prediction Using a Rhombus Pattern

We propose a rhombus pattern prediction scheme. In order to predict the pixel value of position $x_{i,j}$ four surrounding neighboring pixels ($y_{i,j-1}$, $y_{i+1,j}$, $y_{i,j+1}$, and $y_{i-1,j}$) are used. The five pixels including $x_{i,j}$ making a cell. A cell is used to hide one bit of data. In a cover image all pixels are divided into two sets: the “blank” set and “shadow” pixel pair set. The shadow set is using for embedding data and blank set is using for computing predictors. The encoder of the shadow pixel sets embedding scheme for a single cell is as given follows. The center pixel $x_{i,j}$ of the cell can be predicted from its four surrounded neighboring pixels. The pixel x_{ij} and its predicted value $\bar{x}_{i,j}$ is computed as follows:

$$\bar{x}_{i,j} = \left\lfloor \frac{y_{i,j-1} + y_{i+1,j} + y_{i,j+1} + y_{i-1,j}}{4} \right\rfloor \quad (14)$$

Note that the sets that are shadow set and blank sets are independent of each other. Independence means if we change in one set then cause do not affect the other set, and vice versa.

Doing of something the embedding scheme to rely on the shape of the histogram. In normally, doing distribution of the prediction errors has a Laplacian distribution. The shape of the distribution is too decided by the mean and variance. Finally, after the prediction error sequence ($e_1, e_2, \dots, e_N$) is produced, it is divided into pairs of sequence by taking $e_i = (e_{2i}, e_{2i-1})$.

B. Prediction Error Expansion Technique

We have to offer for consideration the linear relationship of the pixels in four directions. Predictors, is bring into the optimization procedure for finding out exactly the smooth and rough region.

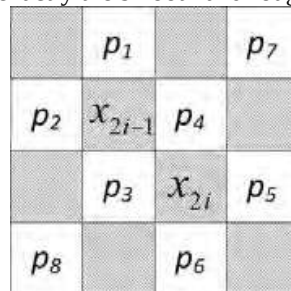


Fig. 1: Context of shadow pixel pair

Finally, after the prediction error sequence ($e_1, e_2, \dots, e_N$) is produced, it is partitioned into pairs by taking $e_i = (e_{2i}, e_{2i-1})$. A usual way to find that a prediction is accurate or not. Accuracy is calculated through prediction's local complexity. See the figure1. Specifically, for a pair e_i its local complexity $LC(e_i)$ is computed as follows

$$LC(e_i) = |p_1 - p_2| + |p_2 - p_3| + |p_3 - p_4| + |p_4 - p_1| + |p_3 - p_6| + |p_6 - p_5| + |p_5 - p_4| + |p_4 - p_7| + |p_3 - p_8| \quad (16)$$

Local complexity is the sum of absolute differences between diagonal blank pixels in the 4×4 sized surrounded neighborhood. A small LC shows that the pair is situated in a smooth image area and should be used firstly choose for data embedding. The main advantages of this pixel-selection strategy are its simplicity and efficiency. Yet a more typical measure may have a more successful selection on pixels. By setting a threshold ρ , the pairs satisfying $LC(e_i) \leq \rho$ are utilized in data embedding. For a described payload, ρ is found out as the smallest integer such that it can be make sure the enough capacity. For other extraction and restoration, some side information requires to be embedded into cover image as well. In our method, the side information considers

- Location map: throughout the time of embedding process, some pixels may found the overflow/underflow problem, for example, the gray-scale value of a pixel after data embedding may be out of the range $[0, 255]$. To stop this, preprocess should be done to change these pixel values into a trusted range. Since change the quality to each pixel value is at most in our method,
- Three parameters are used: first the pixel-selection threshold ρ , second is compressed location map size and third is message size. For example we take a 512×512 sized gray-scale image, and $12 + 18 + 18 = 48$ bits are sufficient to encode these parameters.

In our method, the LSBs embedded three parameters. It is the first line of cover image using LSB replacement, and should be pull out first at data extraction phase. The replaced LSBs of the first-line pixels and the compressed location map will be embedded into the cover image. Compressed location map is as a part of payload.

Now, we describe in details the embedding and extraction procedures for the blank layer. The shadow and blank layers are embedded equally. Each layer is complete embedding process with half of secret message bits.

Embedding Procedure

- Step 1. Leaving out first-line pixels better suited the values of boundary-valued pixels into the reliable range, and build up the location map accordingly. Then minimum lossy compress the location map.
- Step 2. Predict shadow pixels in the scan order and then find out the prediction error pair series ($e_1, e_2, \dots, e_N$) as mentioned above. For each e_i , compute its local complexity $LC(e_i)$ using equation (17).

- Step 3. Empty LSBs of few first-line pixels. To create space for embedding of the three parameters. Put the replaced LSBs and compressed location map as a part of payload.
- Step 4. Find the smallest integer ρ . In which there are enough pairs to embed the payloads. It includes the replaced LSBs, the compressed location map and the secret message bits.
- Step 5. With the use of LSB replacement, the values of ρ , the compressed location map size and the message size are embedded into LSBs of few first-line pixels.
- Step 6. Process the prediction-error pairs, if satisfying $LC(e_i) \leq \rho$ to embed the payload. After this step, the shadow layer embedding is completed.

Extraction Procedure

- Step 1. Through reading LSBs of few first-line pixels, find out the values of the three parameters.
- Step 2. Use same prediction and scan order to get marked prediction-error pair sequence $(e_1', e_2', \dots, e_N')$. For each e_i' calculate its local complexity $LC(e_i')$ which is the same as the one used in the embedding phase.
- Step 3. Process the pairs which is satisfying $LC(e_i') \leq \rho$. To become well again of these pairs is applied by the inverse mapping of the proposed pairwise PEE, and the embedded payload is extracted from the pairs.
- Step 4. After the embedded payload is pull out, the location map and replaced LSBs can be getting.
- Step 5. To returns the first-line pixels by the extracted LSBs. Finally, the original shadow pixels are recovered.

IV. Conclusion

The proposed reversible watermarking algorithm is a mixture of both efficient well-known existing techniques and new techniques. Which is enables performance significantly. Using a rhombus prediction scheme based PEE enables the efficient advantageous use of hiding. A set of sorted prediction errors can be efficiently used for low distortion data hiding. The 2D prediction error expansion method advantageous over the sorted prediction errors. PEE method proposed that it produce excellent ratio between capacity and distortion. In addition, the 2D PEE method in the proposed scheme significantly decreases location map size. Thus, capacity will be considerably increased. The Pairwise PEE is a reversible mapping method that is uses the correlations among prediction errors. The double embedding scheme prevents using each pixel for data hiding in the ideal case. Such designs are expected to further reduce the distortion or increase the capacity to achieve better RDH schemes. It is valuable to investigate this issue in the further work.

References

- [1] M. Wu and B. Liu, "Data hiding in image and video .I. fundamental issues and solutions," *IEEE Trans. Image Process.*, vol. 12, no. 6, pp. 685–695, Jun. 2003.
- [2] Y. Q. Shi, Z. Ni, D. Zou, C. Liang, and G. Xuan, "Lossless data hiding: Fundamentals, algorithms and applications," in *Proc. IEEE ISCAS*, vol. 2, pp. 33–36, May 2004.
- [3] J. Fridrich, M. Goljan, and R. Du, "Invertible authentication," *Proc. SPIE*, vol. 4314, pp. 197–208, Aug. 2001.
- [4] M. U. Celik, G. Sharma, and A. M. Tekalp, "Lossless watermarking for image authentication: A new framework and an implementation," *IEEE Trans. Image Process.*, vol. 15, no. 4, pp. 1042–1049, Apr. 2006.
- [5] J. Tian, "Reversible data embedding using a difference expansion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 890–896, Aug. 2003.
- [6] A. M. Alattar, "Reversible watermark using the difference expansion of a generalized integer transform," *IEEE Trans. Image Process.*, vol. 13, no. 8, pp. 1147–1156, Aug. 2004.
- [7] L. Kamstra and H. J. A. M. Heijmans, "Reversible data embedding into images using wavelet techniques and sorting," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2082–2090, Dec. 2005.
- [8] S. K. Lee, Y. H. Suh, and Y. S. Ho, "Reversible image authentication based on watermarking," in *Proc. IEEE ICME*, pp. 1321–1324, Jul. 2006.
- [9] J. Hwang, J. Kim, and J. Choi, "A reversible watermarking based on histogram shifting," in *Proc. 5th IWDW*, pp. 348–361, 2006.
- [10] D. M. Thodi and J. J. Rodriguez, "Expansion embedding techniques for reversible watermarking," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 721–730, Mar. 2007.
- [11] Y. Hu, H. K. Lee, and J. Li, "DE-based reversible data hiding with improved overflow location map," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 2, pp. 250–260, Feb. 2009.
- [12] V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y. Q. Shi, "Reversible watermarking algorithm using sorting and prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 7, pp. 989–999, Jul. 2009.
- [13] L. Luo, Z. Chen, M. Chen, X. Zeng, and Z. Xiong, "Reversible image watermarking using interpolation technique," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 1, pp. 187–193, Mar. 2010.
- [14] X. Li, B. Yang, and T. Zeng, "Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3524–3533, Dec. 2011.
- [15] D. Coltuc, "Improved embedding for prediction-based reversible watermarking," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 873–882, Sep. 2011.
- [16] S. Lee, C. D. Yoo, and T. Kalker, "Reversible image watermarking based on integer-to-integer wavelet transform," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 3, pp. 321–330, Sep. 2007.