Pairwise Prediction-Error Expansion for Efficient Reversible Data Hiding

Bo Ou, Xiaolong Li, Yao Zhao, *Senior Member, IEEE*, Rongrong Ni, *Member, IEEE*, and Yun-Qing Shi, *Fellow, IEEE*

***Abstract*— In prediction-error expansion (PEE) based reversible data hiding, better exploiting image redundancy usually leads to a superior performance. However, the correlations among prediction-errors are not considered and utilized in current PEE based methods. Specifically, in PEE, the prediction-errors are modified individually in data embedding. In this paper, to better exploit these correlations, instead of utilizing prediction-errors individually, we propose to consider every two adjacent prediction-errors jointly to generate a sequence consisting of prediction-error pairs. Then, based on the sequence and the resulting 2D prediction-error histogram, a more efficient embedding strategy, namely, pairwise PEE, can be designed to achieve an improved performance. The superiority of our method is verified through extensive experiments.**

***Index Terms*— Pairwise prediction-error expansion (pairwise PEE), reversible data hiding (RDH), 2D prediction-error histogram (2D PEH).**

I. Introduction

DATA hiding offers a way to embed data into cover medium for the purposes of ownership protection, authentication, fingerprinting, secret communication and annotation [1]–[3], etc. In most data hiding algorithms, the cover data is destroyed permanently and cannot be exactly restored after the embedded message is extracted. Recently, a new data hiding technique, namely, reversible data hiding (RDH) [4]–[6], is proposed, in which both the cover data and the embedded message can be extracted from the marked content. This specific data hiding technique has been found to be useful in the military, medical and legal fields, where the recovery of the original content is required after data extraction.

Manuscript received March 16, 2013; revised July 14, 2013; accepted September 1, 2013. Date of publication September 11, 2013; date of current version October 2, 2013. This work was supported in part by 973 Program under Grant 2011CB302204, in part by the National Natural Science Funds for Distinguished Young Scholar under Grant 61025013, in part by the National NSF of China under Grants 61210006 and 61272355, and in part by PCSIRT under Grant IRT 201206. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Chun-Shien Lu.

B. Ou and R. Ni are with the Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China, and also with the Beijing Key Laboratory of Advanced Information Science and Network Technology, Beijing 100044, China (e-mail: 09112055@bjtu.edu.cn; rrni@bjtu.edu.cn).

X. Li is with the Institute of Computer Science and Technology, Peking University, Beijing 100871, China (e-mail: lixiaolong@pku.edu.cn).

Y. Zhao is with the Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China, and also with the State Key Laboratory of Rail Traffic Control and Safety, Beijing 100044, China (e-mail: yzhao@bjtu.edu.cn).

Y.-Q. Shi is with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102-1982 USA (e-mail: shi@njit.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2013.2281422

Up to now, various RDH algorithms have been proposed, e.g., compression based algorithms [7]–[11], difference expansion (DE) based algorithms [12]–[16], histogram shifting (HS) based algorithms [17]–[19], prediction-error expansion (PEE) based algorithms [20]–[34] and integer-to-integer transform based algorithms [35]–[41], etc. Besides, some RDH analyses about theoretical capacity limit subjected to admissible distortion have also been given [42]–[44]. Among these techniques, PEE is widely applied nowadays due to its efficient capacity-distortion trade-off.

PEE was first proposed by Thodi and Rodriguez [20], in which the difference between the pixel and its prediction is expanded for data embedding. Compared with DE and HS based methods, PEE has a superior performance since the derived prediction-error histogram (PEH) is more sharply distributed. Afterwards, Hu *et al.* [22] proposed an effective location map for PEE with smaller size, and thereby increased the capacity. Sachnev *et al.* [23] proposed to sort prediction-errors according to the local variance and then process them in the sorted order. Compared with previous methods, the embedding on the sorted prediction-errors introduces less distortion especially for low capacities. Li *et al.* [29] proposed to adaptively embed two bits in a smooth pixel and one bit in a rough one, and thus obtained a better performance for high capacities. Coltuc [39] proposed a novel PEE based transform by modifying not only the current pixel but also its three context pixels. Since the total changes on four pixels are guaranteed to be smaller than that of only modifying one pixel, this method introduces less distortion than the ones based on high-performance predictors such as median-edge-detector (MED) and gradient-adjusted-predictor (GAP).

The prior PEE methods mainly focus on exploiting inter-pixel correlations, but have not considered the correlations within adjacent prediction-errors. That is, a prediction-error may be guessed from its neighbors, and this type of correlations can also be utilized to further reduce the distortion. However, the one-dimensional prediction-error histogram (1D PEH) employed in the conventional PEE, as a low-dimensional projection of image data, is not capable of reflecting such complex dependencies. From this point of view, it is therefore important to formulate a new paradigm of PEE in a higher dimensional space to better exploit these correlations.

In this paper, a novel RDH framework based on the so-called pairwise PEE is proposed. In contrast to the prior PEE methods, we first take every two adjacent prediction-errors as a unit to generate a sequence consisting of prediction-error pairs, then obtain a two-dimensional prediction-error histogram (2D PEH), and finally, embed data by using pairwise PEE, i.e., by expanding or shifting the 2D PEH bins. Compared with the conventional PEE, pairwise PEE can better exploit image redundancy and achieve an improved performance. In addition, inspired by our previous work [29], a refined pixel-selection technique is utilized to preferentially process the pixels located in smooth image regions, and this may further enhance the embedding performance. Experimental results verify that the proposed method outperforms some state-of-the-art works.

The remainder 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. The experimental results are given in Section IV. Finally, Section V concludes this paper.

II. RDH Framework Based on Pairwise PEE

In this section, the basic principle of the conventional PEE used for 1D PEH is reviewed at first in Section II-A. Then, the superiority of 2D PEH over 1D PEH in designing efficient RDH is explained in Section II-B. Finally, the proposed pairwise PEE technique used for 2D PEH is presented in Section II-C. The effectiveness of pairwise PEE in terms of capacity-distortion behavior is also discussed in Section II-C, and one can see that the proposed embedding technique usually provides a better performance compared with the conventional PEE.

A. Conventional PEE

The PEE embedding procedure contains three steps.

1) Predict image pixels to obtain a prediction-error sequence. First, under a specific scan sequencing, the cover image pixels are collected into a one-dimensional sequence as $(x_1, x_2, \ldots, x_N)$. Then, a predictor is used to determine the prediction of x_i denoted as $\widehat{x}_i$. Next, the prediction-error is computed by $e_i = x_i - \widehat{x}_i$ (suppose here for simplicity that $\widehat{x}_i$ is an integer). Finally, the prediction-error sequence $(e_1, e_2, \ldots, e_N)$ is derived.
2) Generate the PEH by counting the frequencies of prediction-errors. That is, the PEH is defined as

$$h(k) = \#\{1 \le i \le N : e_i = k\} \tag{1}$$

where # denotes the cardinal number of a set. Usually, the PEH obeys a Laplacian-like distribution centered at 0 or close to 0. The more sharply the PEH distributes, the less distortion is for embedding the same amount of bits.
3) Embed data by modifying the PEH through expansion and shifting. Specifically, for each prediction-error e_i, it is expanded or shifted as

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

where T is a capacity-dependent integer-valued parameter, and $b \in \{0, 1\}$ is a to-be-embedded data bit. Here, the bins in $[-T, T)$ are expanded to embed data, and those in $(-\infty, -T) \cup [T, +\infty)$ are shifted outwards to create vacancies. Finally, each pixel value x_i is modified to $x_i' = \widehat{x}_i + e_i'$ to obtain the marked image.

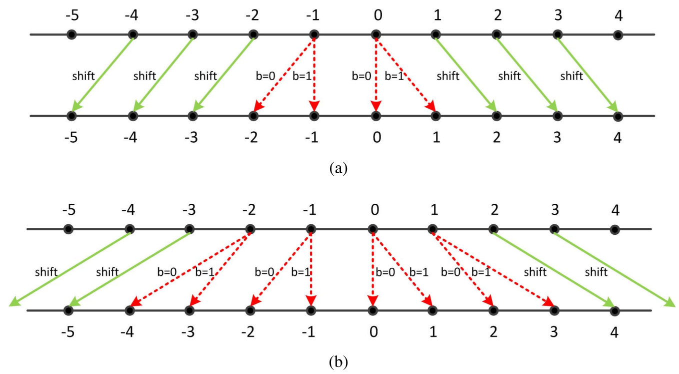


Fig. 1. Transformation of prediction-errors in conventional PEE. (a) $T = 1$. (b) $T = 2$.

For the above PEE embedding, the capacity EC is

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

And the embedding distortion ED in terms of l^2-error can be formulated as

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

since the average distortion for a pixel is

$$\|x_i' - x_i\|^2 = \|e_i' - e_i\|^2 = \begin{cases} \frac{1}{2}\sum_{b\in\{0,1\}}(e_i + b)^2 = e_i^2 + e_i + \frac{1}{2}, & \text{if } e_i \in [-T, T) \\ T^2, & \text{otherwise.} \end{cases} \tag{5}$$

Intuitively, setting a larger expansion range $[-T, T)$ results in the increase of both capacity and distortion.

In PEE extraction procedure, the original prediction-error e_i is recovered from the marked prediction-error e_i' as

$$e_i = \begin{cases} \lfloor e_i'/2 \rfloor, & \text{if } e_i' \in [-2T, 2T) \\ e_i' - T, & \text{if } e_i' \in [2T, +\infty) \\ e_i' + T, & \text{if } e_i' \in (-\infty, -2T) \end{cases} \tag{6}$$

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. Notice that, to guarantee the reversibility, the prediction values used in extraction should be the same as in embedding.

To better illustrate PEE, the transformation of prediction-errors in terms of mapping is shown in Fig. 1, for the cases of $T = 1$ and $T = 2$. For example, when $T = 1$, the mapping between e_i and e_i' is

- $e_i = 0 \longrightarrow e_i' \in \{0, 1\}$,
- $e_i = -1 \longrightarrow e_i' \in \{-1, -2\}$,
- $e_i > 0 \longrightarrow e_i' = e_i + 1$,
- $e_i < -1 \longrightarrow e_i' = e_i - 1$.

B. 2D PEH

We first show that the prediction-error sequence $(e_1, e_2, \ldots, e_N)$ used in PEE is not independently distributed.

In fact, adjacent prediction-errors are usually highly correlated. For example, the correlation within two adjacent prediction-errors e_i and e_{i+1}, which is computed by

$$\frac{|\sum_{i=1}^{N-1}(e_i-\overline{e})(e_{i+1}-\overline{e})|}{\sqrt{\sum_{i=1}^{N-1}(e_i-\overline{e})^2}\sqrt{\sum_{i=1}^{N-1}(e_{i+1}-\overline{e})^2}} \tag{7}$$

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 $\overline{e}$ is the mean of $(e_1, e_2, \ldots, e_N)$. This correlation coefficient (7) 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 extent, or even highly correlated for some specific images (e.g., Barbara). Here, rhombus prediction is used to generate the prediction-error sequence, i.e., the prediction of a pixel is the average value of its four nearest neighboring pixels.

To our knowledge, to improve the embedding performance, previous PEE based works mainly focus on how to better exploit inter-pixel correlations to derive a sharply distributed PEH. After PEH is generated, the prediction-errors are modified individually for data embedding. However, the correlations within prediction-errors are not utilized in this individually modification strategy. Hence, to achieve a better performance in RDH, we propose to consider 2D PEH where the correlations within prediction-errors can be exploited.

We now define the 2D PEH and prove its advantage in RDH over the 1D PEH. By considering every two adjacent prediction-errors together, the sequence $(e_1, e_2, \ldots, e_N)$ can be transformed into a new one $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{N/2})$ with $\mathbf{e}_i = (e_{2i-1}, e_{2i})$ (suppose for simplicity that N is even), and the associated 2D PEH is

$$g(k_1, k_2) = \#\{1 \le i \le N/2 : e_{2i-1} = k_1, e_{2i} = k_2\}. \tag{8}$$

For convenience, the sequences $(e_1, e_2, \ldots, e_N)$, $(e_1, e_3, \ldots, e_{N-1})$, $(e_2, e_4, \ldots, e_N)$ and $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{N/2})$ are denoted as E, E_{odd}, E_{even} and E^*, respectively. Clearly, it can be assumed that

$$H(E) = H(E_{odd}) = H(E_{even}) \tag{9}$$

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

$$H(E^*) = H(E_{odd}, E_{even}) \le H(E_{odd}) + H(E_{even}) = 2H(E) \tag{10}$$

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. So, we have

$$\frac{H(E^*)}{2} < H(E). \tag{11}$$

It demonstrates that the average number of bits required to represent a prediction-error in E^* is less than in E. In other words, compared with E, E^* can better decorrelate the image data and has a potential to yield an improved RDH.

Moreover, to clarify (11), the entropies $H(E)$ and $H(E^*)/2$ for six standard test images are given in Table I. One can see that (11) holds for these commonly used test images. The theoretical result in (11) is then verified experimentally.

TABLE I

ENTROPIES $H(E)$ AND $H(E^*)/2$, FOR SIX IMAGES INCLUDING LENA, BABOON, AIRPLANE, BARBARA, LAKE AND BOAT

Image	Lena	Baboon	Airplane	Barbara	Lake	Boat
$H(E)$	4.106	5.970	3.868	5.111	4.972	4.810
$H(E^*)/2$	4.021	5.832	3.731	4.835	4.874	4.713

Finally, we illustrate that the conventional PEE in 1D PEH can also be implemented in 2D PEH. Specifically, the transformation of prediction-errors in the conventional PEE (see Fig. 1) can be redefined for prediction-error pairs in an equivalent way. In particular, in the case of $T = 1$, the mapping in Fig. 1(a) is in fact equivalent to the one shown in Fig. 2, in which the new mapping represents the transformation of prediction-error pairs. To better describe Fig. 2, some examples illustrating the transformation of prediction-error pairs are given below.

- For the pair $(e_{2i-1}, e_{2i}) = (0, 0)$, in the 1D mapping shown in Fig. 1(a), e_{2i-1} is expanded to 0 or 1 to embed a data bit $b_1 \in \{0, 1\}$, e_{2i} is also expanded to 0 or 1 to embed another data bit $b_2 \in \{0, 1\}$. Accordingly, in the 2D case shown in Fig. 2, the pair $(0, 0)$ is expanded to $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$ when (b_1, b_2) is $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$, respectively.
- For $(e_{2i-1}, e_{2i}) = (2, 0)$, in 1D case, e_{2i-1} is shifted to 3, and e_{2i} is expanded to 0 or 1 to embed a data bit $b \in \{0, 1\}$. So, in 2D case, the pair $(2, 0)$ is expanded to $(3, 0)$ if $b = 0$, and $(3, 1)$ if $b = 1$.
- For $(e_{2i-1}, e_{2i}) = (1, 2)$, in 1D case, e_{2i-1} and e_{2i} are shifted to 2 and 3, respectively. So, in 2D case, the pair $(1, 2)$ is shifted to $(2, 3)$.

With the above illustration, we see that the conventional PEE can be implemented in an equivalent way by modifying the 2D PEH. This paradigm of PEE provides us a new way to design RDH. We will see later that, by extending the expansion and shifting techniques used in 1D PEH to 2D PEH, the conventional PEE can be improved according to a more reasonable histogram modification strategy on 2D PEH.

C. Pairwise PEE

Since 2D PEH can better reflect the complex dependencies in image data than 1D PEH, it is valuable to utilize 2D PEH to design RDH. In this work, the histogram modification strategy based on 2D PEH is called pairwise PEE, in which the data embedding is implemented by expanding and shifting the bins of 2D PEH. Pairwise PEE is a natural extension of the conventional PEE in higher dimensional space. Based on 2D PEH, various histogram modification strategies can be designed with different performance. Actually, the conventional PEE shown in Fig. 2 is a special case of pairwise PEE. However, we will see later that this specific pairwise PEE is far from efficient. In fact, a reasonable histogram modification strategy directly contributes to a superior RDH performance.

We now present our pairwise PEE strategy to enhance the conventional PEE shown in Fig. 2. Notice that, we do not

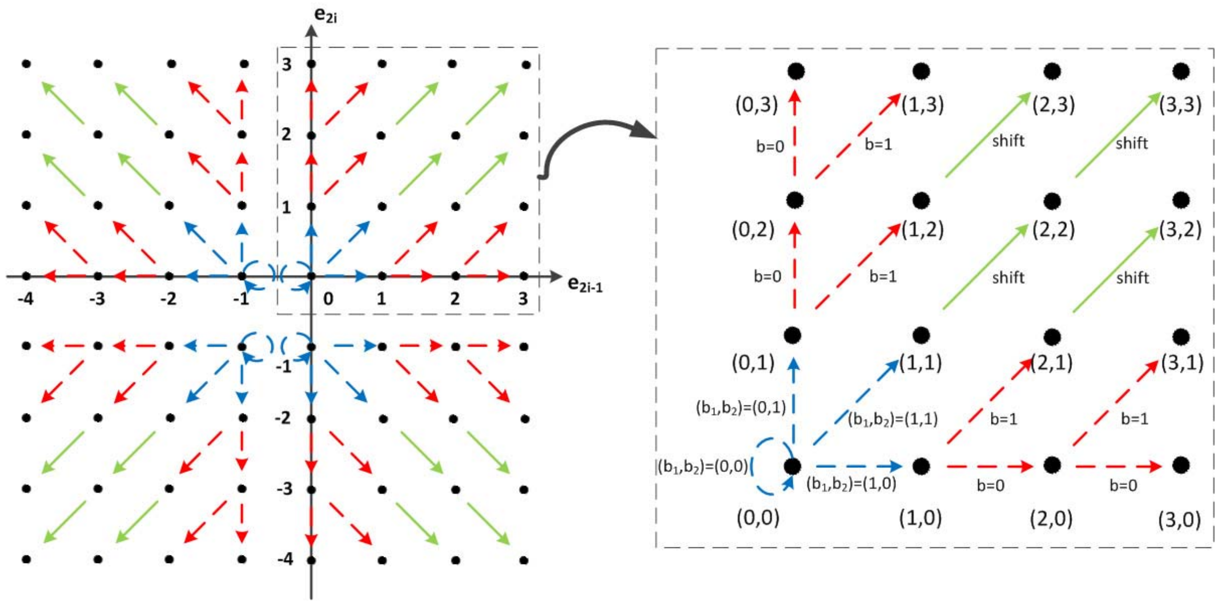


Fig. 2. Transformation of prediction-error pairs in conventional PEE, for $T = 1$.

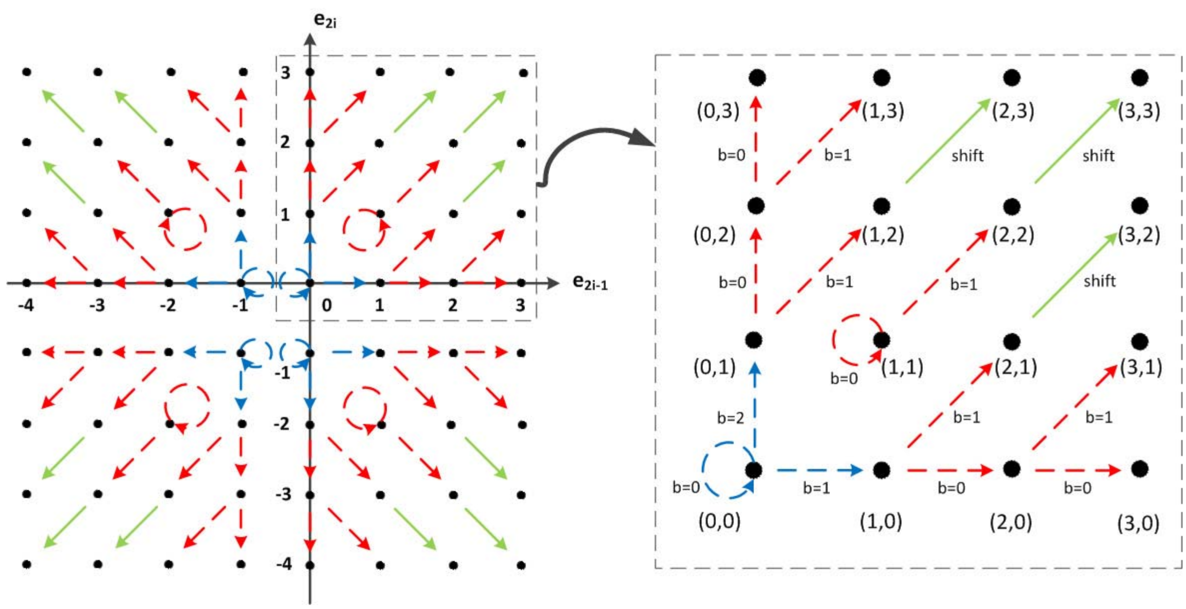


Fig. 3. Proposed pairwise PEE.

intent to analyze every possible case of pairwise PEE, but investigate the generic case that leads to insightful results. In doing so, here we only consider the case that the maximum modification to each pixel value is restricted to 1.

Our idea is straightforward. We want to expand or shift bins in a less distorted direction as much as possible. For example, for the pair $(0, 0)$ in Fig. 2, it is embedded with 2 bits (b_1, b_2), and the distortion is 0, 1, 1 and 2 when (b_1, b_2) is $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$, respectively. In this case, the cost for mapping $(0, 0)$ to $(1, 1)$ is 2, while mapping $(0, 0)$ to $(0, 0)$, $(0, 1)$ or $(1, 0)$ costs much less. So, to reduce the distortion, the directions with high distortion, e.g., mapping $(0, 0)$ to $(1, 1)$, should be discarded. Based on this consideration, we propose a new 2D mapping as shown in Fig. 3. In this new mapping, for each pair in $\{(0, 0), (0, -1), (-1, 0), (-1, -1)\}$, it is embedded with $\log_2 3$ bits instead of 2 bits in the conventional PEE. Moreover, since the mappings between $(0, 0)$ and $(1, 1)$, $(0, -1)$ and $(1, -2)$, $(-1, 0)$ and $(-2, 1)$, $(-1, -1)$ and $(-2, -2)$, are discarded, each pair in $\{(1, 1), (1, -2), (-2, 1), (-2, -2)\}$ can be embedded with 1 bit in the new mapping while they are just shifted in the conventional PEE.

The capacity-distortion performance of the proposed pairwise PEE is analyzed as follows. One can see later that the new method outperforms the conventional PEE. For convenience, the 2D bins are classified into four types A, B, C and D. Fig. 4 shows the classification for the first quarter of 2D PEH. The classifications for the other three quarters are similar

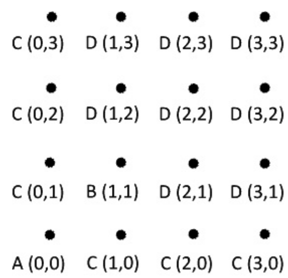


Fig. 4. Classification of 2D bins for the first quarter of 2D PEH.

TABLE II

CAPACITIES (IN BITS) OF THE CONVENTIONAL PEE ($T = 1$) AND THE PROPOSED PAIRWISE PEE, FOR SIX IMAGES INCLUDING LENA, BABOON, AIRPLANE, BARBARA, LAKE AND BOAT

Image	Lena	Baboon	Airplane	Barbara	Lake	Boat
Conventional PEE	68,866	21,988	100,961	54,417	39,338	41,081
Proposed pairwise PEE	71,299	22,725	97,596	56,500	40,304	42,558

to the first one and the illustrations are omitted. With this classification, the capacities of the conventional PEE and the proposed pairwise PEE, denoted as EC_{con} and EC_{pro}, can be written as

$$EC_{con} = 2\sum_{\mathbf{e}\in A} g(\mathbf{e}) + \sum_{\mathbf{e}\in C} g(\mathbf{e}) \tag{12}$$

and

$$EC_{pro} = \log_2 3 \sum_{\mathbf{e}\in A} g(\mathbf{e}) + \sum_{\mathbf{e}\in B} g(\mathbf{e}) + \sum_{\mathbf{e}\in C} g(\mathbf{e}) \tag{13}$$

where g is the 2D PEH defined in (8). On the other hand, the distortion in terms of l^2-error of the conventional PEE and the proposed pairwise PEE, denoted as ED_{con} and ED_{pro}, can be formulated as

$$ED_{con} = \sum_{\mathbf{e}\in A} g(\mathbf{e}) + 2\sum_{\mathbf{e}\in B} g(\mathbf{e}) + \frac{3}{2}\sum_{\mathbf{e}\in C} g(\mathbf{e}) + 2\sum_{\mathbf{e}\in D} g(\mathbf{e}) \tag{14}$$

and

$$ED_{pro} = \frac{2}{3}\sum_{\mathbf{e}\in A} g(\mathbf{e}) + \sum_{\mathbf{e}\in B} g(\mathbf{e}) + \frac{3}{2}\sum_{\mathbf{e}\in C} g(\mathbf{e}) + 2\sum_{\mathbf{e}\in D} g(\mathbf{e}). \tag{15}$$

According to (14) and (15), compared with the conventional PEE, we claim that the distortion of the proposed pairwise PEE is less, since

$$ED_{con} - ED_{pro} = \frac{1}{3}\sum_{\mathbf{e}\in A} g(\mathbf{e}) + \sum_{\mathbf{e}\in B} g(\mathbf{e}) > 0. \tag{16}$$

For the capacity, according to (12) and (13), we have

$$EC_{con} - EC_{pro} \approx 0.415 \sum_{\mathbf{e}\in A} g(\mathbf{e}) - \sum_{\mathbf{e}\in B} g(\mathbf{e}). \tag{17}$$

By this equation, although we cannot demonstrate theoretically that the capacity could be increased by our method, one can expect that the capacity will not change too much. Referring to Table II, our capacity is in fact larger than that of conventional PEE in most cases. For this table, we apply (12) and (13) directly to a prediction-error sequence obtained by using the rhombus prediction.

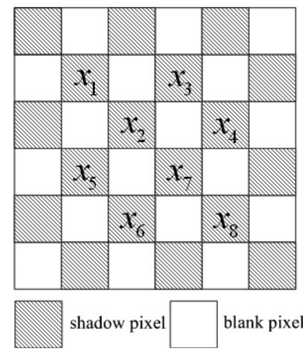


Fig. 5. Scan order for shadow pixels (from left to right and top to bottom, two lines by two lines).

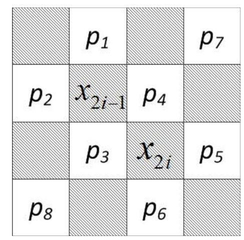


Fig. 6. Context of a shadow pixel pair.

Based on the above analysis, we conclude that the proposed pairwise PEE is better than the conventional PEE in terms of capacity-distortion trade-off. The detailed data embedding and extraction procedures are given in next section, and the experimental results confirming our superiority will be reported in Section IV.

III. IMPLEMENTATION OF PROPOSED PAIRWISE PEE SCHEME

The equilateral parallelogram pattern in Sachnev *et al.*'s method [23] is used in our work for prediction (see Fig. 5), in which the cover image is divided into two sets denoted as "shadow" and "blank". Notice that the first line of cover image is excluded in the partition of shadow and blank since it will be used to embed some parameters for the purpose of blind extraction.

The embedding based on equilateral parallelogram pattern is also called double-layered embedding as the twice embedding should be processed to cover the whole image, and the prediction of blank pixels are processed only when the embedding of shadow pixels is completed. Since the two layers' embedding are processed similarly, we only take the shadow layer for illustration.

First, the shadow pixels are collected into a sequence $(x_1, x_2, \ldots, x_N)$ in a defined order as shown in Fig. 5. Then, for a shadow pixel, the rhombus prediction is computed using its four nearest blank pixels. Here, the prediction value is rounded up if it is not an integer. Finally, after the prediction-error sequence $(e_1, e_2, \ldots, e_N)$ is generated, it is divided into pairs by taking $\mathbf{e}_i = (e_{2i}, e_{2i-1})$.

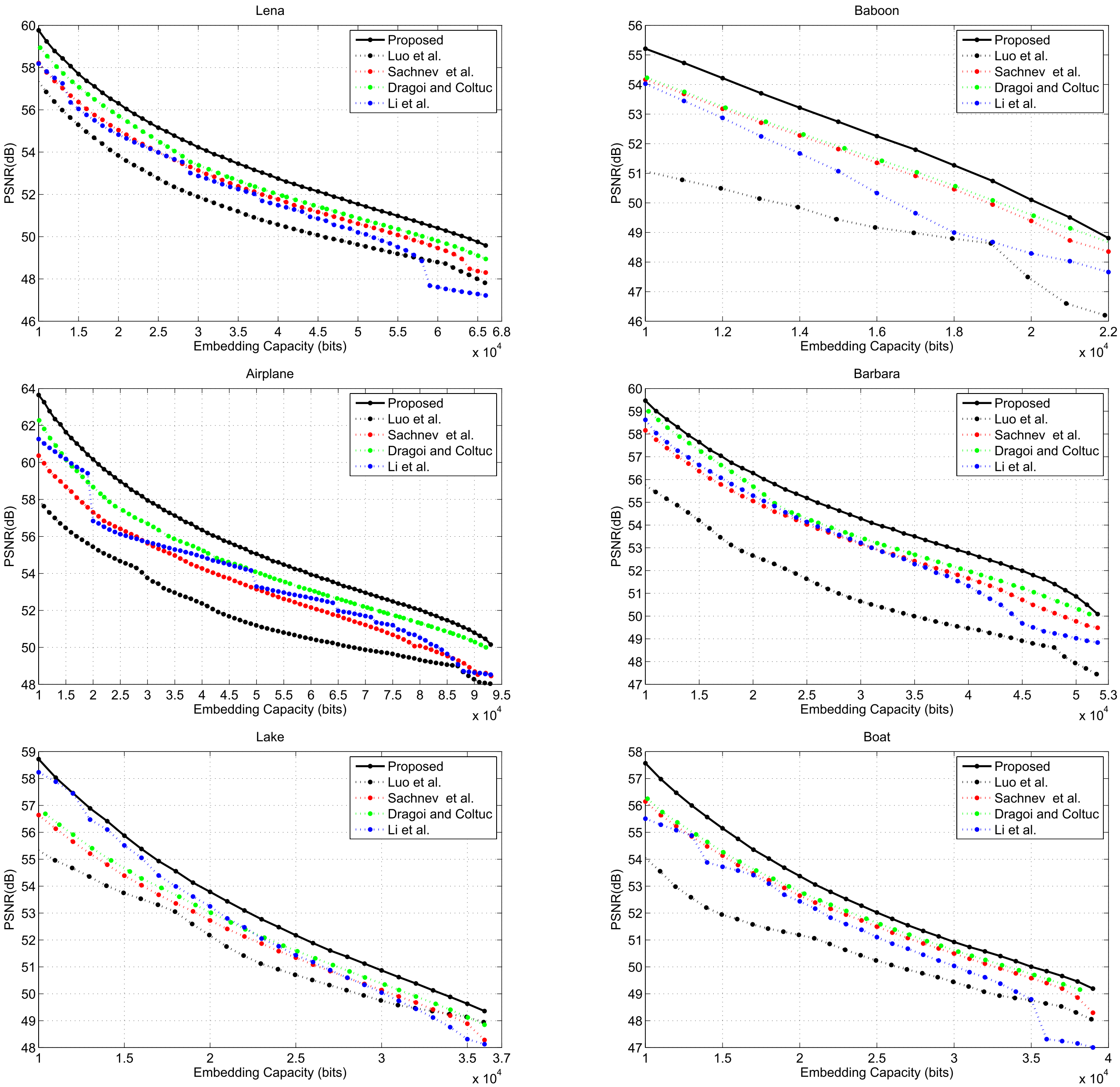


Fig. 7. Comparisons in terms of capacity-distortion performance.

In addition to pairwise PEE, a refined pixel-selection technique motivated by our previous work [29] is also utilized in the proposed method. The objective of pixel-selection is to put prediction-error pairs in a decreasing order of prediction accuracy, so that the pairs with small magnitudes are processed first. A common way to determine whether a prediction is accurate or not is to calculate its local complexity. Specifically, for a pair $\mathbf{e}_i$, its local complexity $LC(\mathbf{e}_i)$ is computed as follows (see Fig. 6)

$$LC(\mathbf{e}_i) = |p_1 - p_2| + |p_2 - p_3| + |p_3 - p_4| + |p_4 - p_1| + |p_3 - p_6| + |p_6 - p_5| + |p_5 - p_4| + |p_4 - p_7| + |p_3 - p_8|. \quad (18)$$

It is the sum of absolute differences between diagonal blank pixels in the 4×4 sized neighborhood. A small LC indicates that the pair is located in a smooth image region and should be used preferentially for data embedding. The simplicity and efficiency are the main advantages of this pixel-selection strategy, yet a more complex measure may have a more successful selection on pixels. By setting a threshold ρ, the pairs satisfying $LC(\mathbf{e}_i) \leq \rho$ are utilized in data embedding while the others are skipped. For a specific payload, ρ is determined as the smallest integer such that it can ensure the enough capacity.

For blind extraction and restoration, some side information need to be embedded into cover image as well. In our method, the side information includes

- Location map: During the embedding process, some pixels may encounter the overflow/underflow problem, i.e.,

TABLE III
COMPARISONS IN TERMS OF PSNR (dB) FOR A CAPACITY OF 10,000 BITS

Image	Luo *et al.*	Sachnev *et al.*	Dragoi and Coltuc	Li *et al.*	Proposed
Lena	57.35	58.22	58.94	58.17	**59.75**
Baboon	51.08	54.19	54.22	54.02	**55.21**
Airplane	57.97	60.56	62.28	61.27	**63.76**
Barbara	55.76	58.16	59.00	58.62	**59.48**
Lake	55.34	56.69	56.69	58.23	**58.72**
Boat	54.07	56.17	56.25	55.51	**57.55**
Average	55.26	57.30	57.90	57.64	**59.08**

TABLE IV
COMPARISONS IN TERMS OF PSNR (dB) FOR A CAPACITY OF 20,000 BITS

Image	Luo *et al.*	Sachnev *et al.*	Dragoi and Coltuc	Li *et al.*	Proposed
Lena	53.85	55.06	55.70	54.83	**56.29**
Baboon	47.50	49.42	49.56	48.29	**50.12**
Airplane	55.44	57.33	58.67	56.84	**60.20**
Barbara	52.66	55.06	55.69	55.29	**56.27**
Lake	52.18	52.73	53.01	53.25	**53.76**
Boat	51.19	52.68	52.72	52.44	**53.34**
Average	52.14	53.71	54.23	53.49	**55.00**

the gray-scale value of a pixel after data embedding may be out of the range $[0, 255]$. To prevent this, a preprocess should be done to adjust these pixel values into a reliable range. Since the modification to each pixel value is at most 1 in our method, only border-valued pixels need to be adjusted as follows. We modify x_i to 254 if it is 255 and to 1 if it is 0. Meanwhile, each modified pixel is marked with 1 in a location map while the others are marked with 0. Then the location map is losslessly compressed to further reduce its size.

- Three parameters: the pixel-selection threshold ρ, the compressed location map size and the message size. For a 512×512 sized gray-scale image, $12 + 18 + 18 = 48$ bits are sufficient to encode these parameters.

In our method, the three parameters are embedded into LSBs of the first line of cover image using LSB replacement, and should be extracted 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 as a part of payload. That is, the embedded payload includes three parts: the replaced LSBs, the compressed location map and the secret message.

We now describe in details the embedding and extraction procedures for the shadow layer. Notice that here, the same as the method of Sachnev *et al.* [23], the shadow and blank layers are embedded equally, i.e., each layer is embedded with half of secret message bits.

Embedding Procedure

- Step 1. Except first-line pixels, adjust the values of boundary-valued pixels into the reliable range, and construct the location map accordingly. Then losslessly compress the location map.

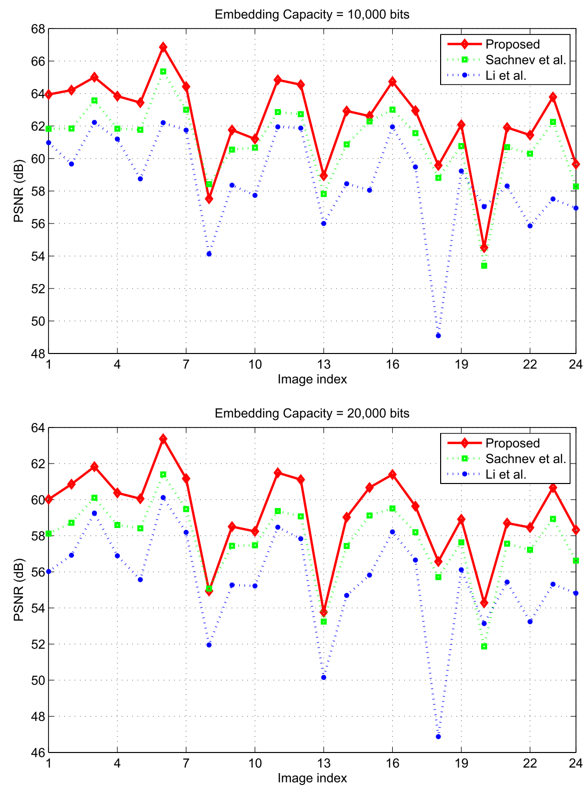


Fig. 8. Comparisons on the Kodak image database, for capacities of 10,000 and 20,000 bits.

- Step 2. Predict shadow pixels in the scan order as shown in Fig. 5, and then determine the prediction-error pair sequence $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N)$ as mentioned above. For each $\mathbf{e}_i$, calculate its local complexity $LC(\mathbf{e}_i)$ using (18).
- Step 3. Empty LSBs of some first-line pixels to make room for the embedding of the three parameters. Take the replaced LSBs and compressed location map as a part of payload.
- Step 4. Find the smallest integer ρ such that there are enough pairs to embed the payload including the replaced LSBs, the compressed location map and the secret message bits.
- Step 5. Using LSB replacement, embed the values of ρ, the compressed location map size and the message size into LSBs of some first-line pixels.
- Step 6. Process the prediction-error pairs satisfying $LC(\mathbf{e}_i) \leq \rho$ to embed the payload. The modifications on these pairs are based on the proposed pairwise PEE shown in Fig. 3. After this step, the shadow layer embedding is completed.

Extraction Procedure

- Step 1. By reading LSBs of some first-line pixels, determine the values of the three parameters.
- Step 2. Use the same prediction and scan order to obtain the marked prediction-error pair sequence $(\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_N)$. For each $\mathbf{e}'_i$, compute its local complex-

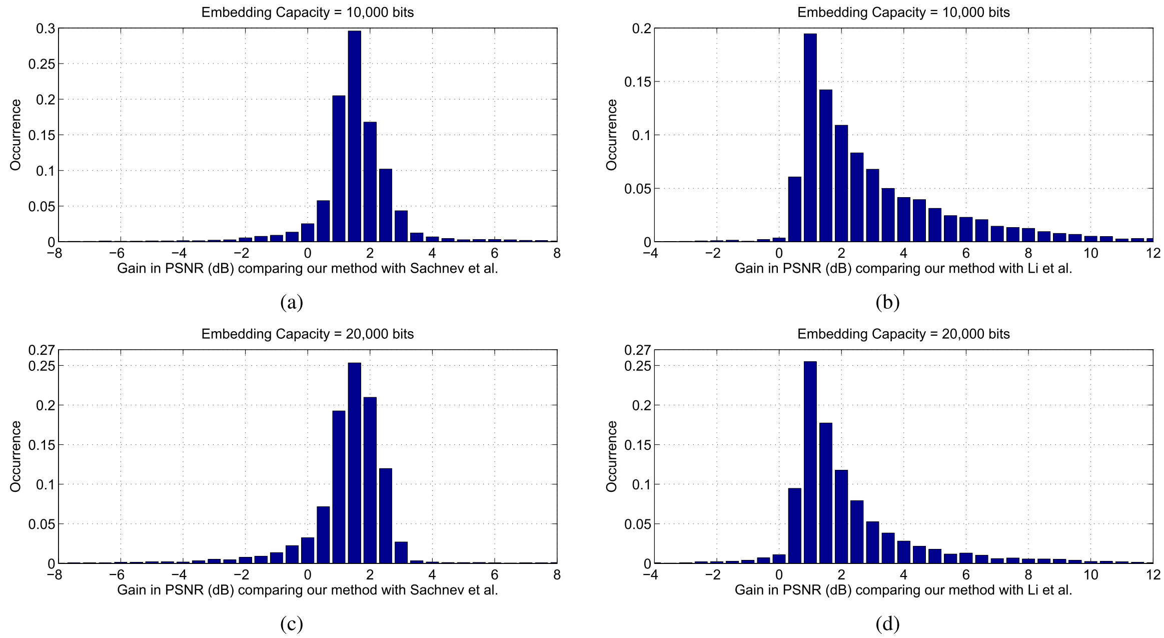


Fig. 9. Histograms of gain in PSNR by comparing our method with Sachnev *et al.*'s and Li *et al.*'s on BOSSbase v1.01 image database, for capacities of 10,000 and 20,000 bits. (a) Comparison with Sachnev et al.'s method. (b) Comparison with Li et al.'s method. (c) Comparison with Sachnev et al.'s method. (d) Comparison with Li et al.'s method.

ity $LC(\mathbf{e}'_i)$ which is the same as the one used in the embedding phase.

- Step 3. Process the pairs satisfying $LC(\mathbf{e}'_i) \leq \rho$. The recovery of these pairs is implemented by the inverse mapping of the proposed pairwise PEE, and the embedded payload is extracted from the pairs with type A, B and C (see Fig. 4).
- Step 4. After the embedded payload is extracted, the location map and replaced LSBs can be obtained. According to the location map, the pixels marked with 1 will be further revised and their original values can be restored.
- Step 5. Recover the first-line pixels by the extracted LSBs. Finally, the original shadow pixels are recovered.

IV. Experimental Results and Discussion

In this section, the capacity-distortion performance of the proposed method is evaluated by comparing it with some state-of-the-art works including Luo *et al.* [25], Sachnev *et al.* [23], Dragoi and Coltuc [31] and Li *et al.* [29].

The first experiment, as shown in Fig. 7, is enforced on six standard 512×512 sized gray-scale images: Lena, Baboon, Airplane, Barbara, Lake and Boat. Here, except Barbara, the other five images are downloaded from the USC-SIPI database[1]. Since the maximum modification of each pixel value is 1 in the proposed method, our capacity is limited (e.g., about 0.263 bits per pixel for Lena image). For comparison, only the low capacity cases are considered. That is, the capacity range is limited from 10,000 bits to the maximum of our method with a step size of 1,000 or 2,000 bits. According to Fig. 7, one can

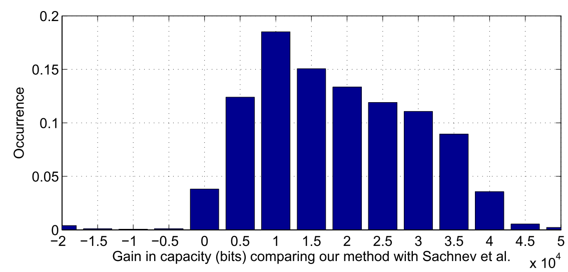


Fig. 10. Histogram of gain in capacity for the equal distortion rate of 55 dB.

see that the proposed method outperforms these state-of-the-art works. Our method can provide a larger PSNR whatever the test image or capacity is. Moreover, experimental results for two given capacities 10,000 and 20,000 bits, are listed in Tables III and IV, and our PSNR gains in average are at least 1.18 and 0.77 dB, respectively.

Among the four comparison methods, Luo *et al.*'s is a typical PEE method which pursues the prediction accuracy by designing a more accurate predictor. Sachnev *et al.*'s is a hybrid PEE method using a sorting technique to preferentially process the smooth pixels. Although the predictor is simple, the sorting technique makes the data embedding more efficient and hence Sachnev *et al.*'s method significantly outperforms the typical PEE method of Luo *et al.*. Dragoi and Coltuc's method also utilizes a sorting technique. This method improves the rhombus prediction by considering local gradients, and it is slightly better than Sachnev *et al.*'s. Li *et al.*'s method utilizes GAP for prediction, and mainly focuses on the performance of very high capacity by embedding more bits into smooth

[1]http://sipi.usc.edu/database/database.php?volume=misc

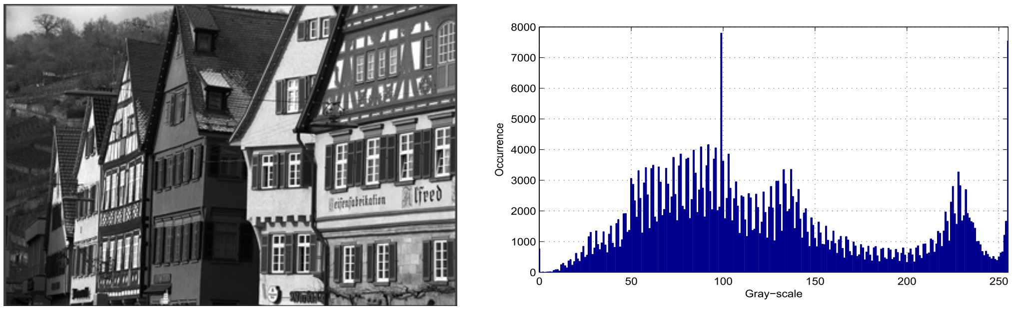


Fig. 11. The 8-th Kodak image and its gray-scale histogram.

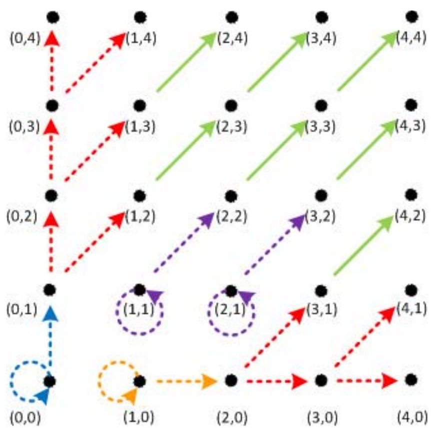


Fig. 12. Another 2D mapping for pairwise PEE, where $T = 1$.

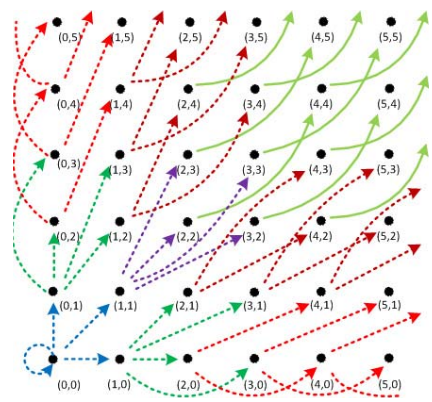


Fig. 13. 2D mapping for the conventional PEE with $T = 2$.

pixels. Although employing a pixel selection technique, for low capacity, their method usually yields a smaller PSNR than Sachnev *et al.*'s. Typically, compared with Sachnev *et al.*'s method, although the same predictor and a similar sorting technique are utilized, our method is better by increasing PSNR with 1.78 and 1.29 dB in average for a capacity of 10,000 and 20,000 bits (see Tables III and IV), respectively. Our main advantage is the utilization of 2D PEH and pairwise PEE.

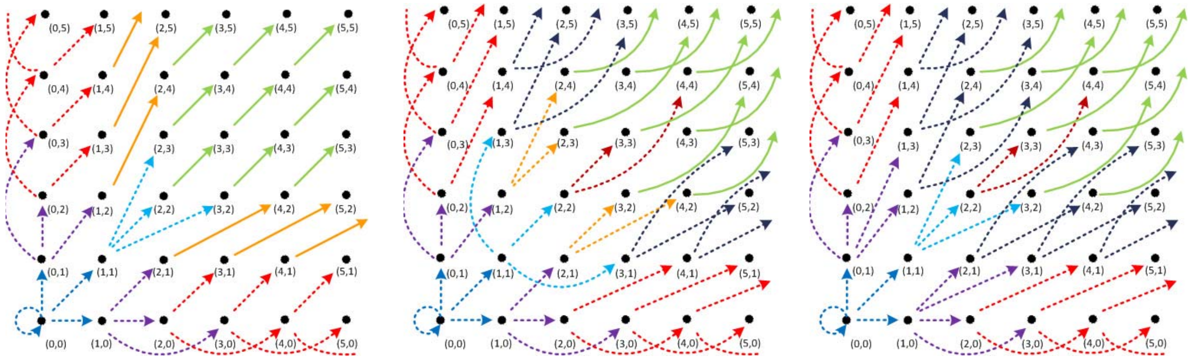
To further verify the superiority of our method, another experiment as shown in Fig. 8 is enforced on the Kodak image database[2] which contains 24 color images sized 512×768 or 768×512. For testing, the color images are transformed into gray-scale. In Fig. 8, compared our method with Sachnev *et al.*'s, the average PSNR gains are 1.34 and 1.50 dB for a capacity of 10,000 and 20,000 bits, respectively. Compared with Li *et al.*'s, the corresponding gains are 3.67 and 3.76 dB, respectively.

In addition to the standard and Kodak images, we also compare our method with Sachnev *et al.*'s and Li *et al.*'s on a large-scale database BOSSbase v1.01[3] [45]. BOSSbase contains 10,000 512×512 sized gray-scale images and has been widely utilized in information hiding society nowadays. Figs. 9(a) and (c) give the histograms of PSNR gains, in which the gain is calculated by subtracting Sachnev *et al.*'s PSNR from ours, while Figs. 9(b) and (d) give the comparison with Li *et al.*'s method. For the capacity of 10,000 bits, the average PSNRs of Sachnev *et al.*'s, Li *et al.*'s and ours are 61.5, 59.98 and 63.08 dB, respectively. Thus, compared with these two methods, the average PSNR gains are 1.58 and 3.1 dB, respectively. For 20,000 bits, the corresponding average PSNRs for the three methods are 58.34, 57.43 and 59.67 dB, respectively. And, our gains are 1.33 and 2.24 dB in this case.

For the equal distortion rate, Fig. 10 plots the histogram of capacity gain by comparing our method with Sachnev *et al.*'s on BOSSbase images. Here, for each image, we calculate the capacities by using the two methods when the PSNR of marked image is 55 dB. In this situation, our method can embed 58,999 bits per image in average, while Sachnev *et al.*'s can embed 40,545 bits. The capacity is increased by 45.5% in average.

It is worth mentioning that pairwise PEE has a lot of freedom in design, and various mappings based on 2D PEH can be obtained. The proposed pairwise PEE is a generic extension for the conventional PEE with the case of $T = 1$, but may not be optimal. For example, for the 8-th Kadak image, our PSNR is lower than that of Sachnev *et al.* for a capacity of 10,000 bits (see Fig. 8 (a)). The image is textured and contains a large number of boundary-valued pixels (see

[2]http://www.r0k.us/graphics/kodak/
[3]http://www.agents.cz/boss/BOSSFinal/



(a) (b) (c)

Fig. 14. Three 2D mappings for pairwise PEE with $T = 2$. (a) 2D mapping: M1. (b) 2D mapping: M2. (c) 2D mapping: M3.

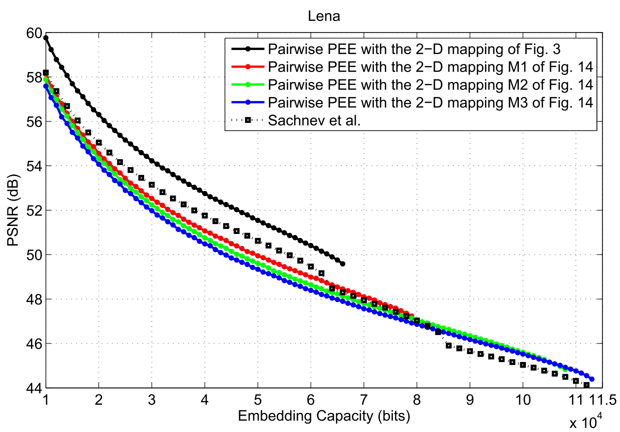


Fig. 15. Performance comparison between pairwise PEE (with the mapping in Fig. 3 and the three new mappings in Fig. 14) and Sachnev *et al.*'s method.

Fig. 11). To get a better performance on this image, we may design another mapping shown in Fig. 12 for pairwise PEE. Here, for simplicity, only the mapping for the first quarter is plotted. By this new mapping, compared with the one in Fig. 3, the PSNR can be increased from 57.63 to 57.96 dB for a capacity of 10,000 bits. For the same capacity, the PSNR of Sachnev *et al.*'s method is 57.98 dB. One can see that the PSNR of this new mapping is comparable with that of Sachnev *et al.*'s. In this light, we may conclude that a more better RDH scheme can be obtained by designing an image-dependent pairwise PEE.

So far, only the low capacity cases with $T = 1$ are considered. In fact, by modifying the pixel values larger than 1, higher capacities can be achieved in pairwise PEE as well. Considering the case of $T = 2$, the conventional PEE in terms of 2D mappings is plotted in Fig. 13. The same idea described in Section II-C can be then applied to the mapping shown in Fig. 13 to get better RDH schemes. We discard the mapping direction with large distortion and propose three new mappings for $T = 2$ as shown in Fig. 14. Compared with the mapping in Fig. 3, these new ones introduce larger distortion but higher capacities. The performance comparison between pairwise PEE with different mappings and Sachnev *et al.*'s method, for Lena image, is plotted in Fig. 15. One can observe that, by using the new mappings, the capacity of pairwise PEE is enlarged, and a larger PSNR over Sachnev *et al.*'s method is also derived for the case of higher capacities. Moreover, for pairwise PEE, different mappings may result in different performance. So, for a given capacity, to get the best performance, one should adaptively employ a specific mapping to maximize the PSNR. In this way, referring to Fig. 16, pairwise PEE with adaptive 2D mapping can improve Sachnev *et al.*'s method in an enlarged capacity range. Here, only the four mappings shown in Figs. 3 and 14 are employed as candidates to optimize the pairwise PEE.

As aforementioned, there exists many 2D mappings, and the proposed ones are serval effective attempts for low capacities. We remark that, finding the optimal 2D mapping for a given capacity and providing a way to adaptively alternate 2D mappings for different capacities would be a valuable work. This is beyond the scope of this paper and would be done in the future work.

V. Conclusion

In this paper, an efficient RDH scheme based on pairwise PEE is proposed. The pairwise PEE is a novel reversible mapping that utilizes the correlations among prediction-errors. With the help of this type of correlations, the distortion can be controlled at a low level, and thereby the proposed scheme outperforms some state-of-the-art RDH algorithms.

So far, the proposed pairwise PEE is only designed for simple cases that the modification to a pixel value is at most 2. In fact, some other histogram modification strategies for different designs of pairwise PEE could also be explored, e.g., an image-dependent design or modify the pixel values larger than 2. Such designs are expected to further reduce the distortion or increase the capacity to achieve more better RDH schemes. It is valuable to investigate this issue in the further work.

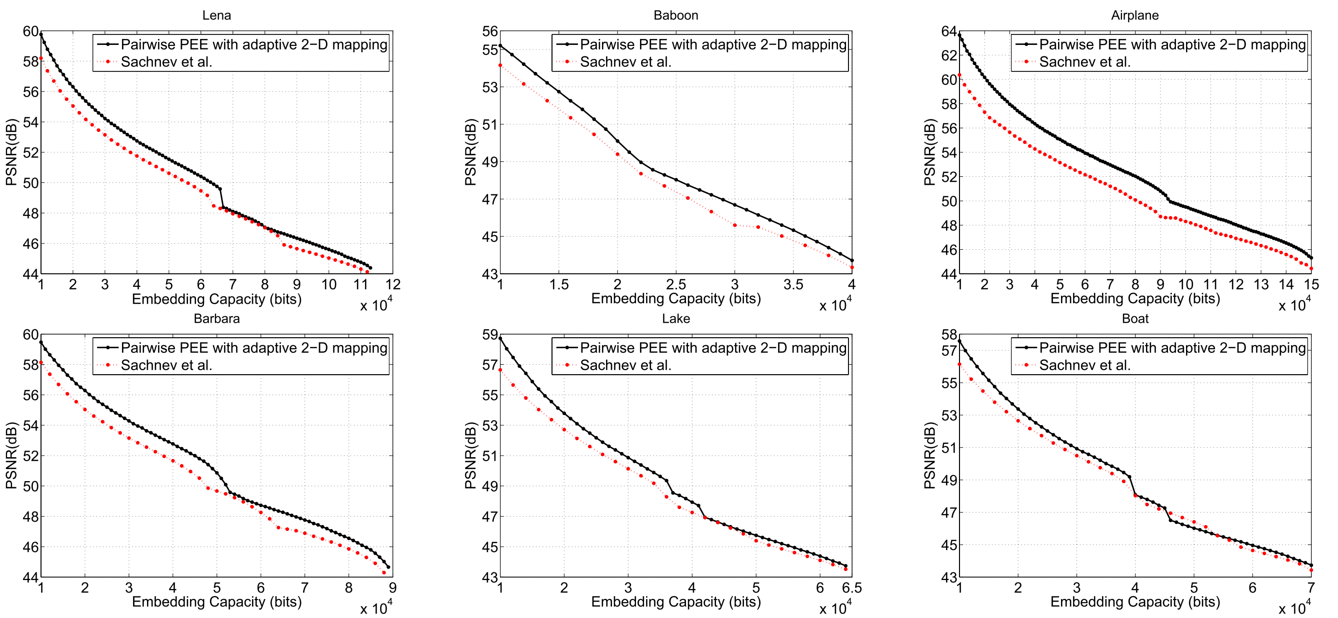


Fig. 16. Performance comparison between pairwise PEE with adaptive 2D mapping and Sachnev *et al.*'s method.

ACKNOWLEDGMENT

The authors would like to thank Prof. D. Coltuc of Valahia University for providing us their experimental results.

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] M. Wu, H. Yu, and B. Liu, "Data hiding in image and video .II. designs and applications," *IEEE Trans. Image Process.*, vol. 12, no. 6, pp. 696–705, Jun. 2003.

[3] I. Cox, M. Miller, J. Bloom, J. Fridrich, and T. Kalker, *Digital Watermarking Steganography*, 2nd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2007.

[4] 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. May 2004, pp. 33–36.

[5] Y. Q. Shi, "Reversible data hiding," in *Proc. IWDW*, 2004, pp. 1–12.

[6] R. Caldelli, F. Filippini, and R. Becarelli, "Reversible watermarking techniques: An overview and a classification," *EURASIP J. Inf. Security*, vol. 2010, Jan. 2010.

[7] J. Fridrich, M. Goljan, and R. Du, "Invertible authentication," *Proc. SPIE*, vol. 4314, pp. 197–208, Aug. 2001.

[8] G. Xuan, J. Zhu, J. Chen, Y. Q. Shi, Z. Ni, and W. Su, "Distortionless data hiding based on integer wavelet transform," *IEE Electron. Lett.*, vol. 38, no. 25, pp. 1646–1648, 2002.

[9] J. Fridrich, M. Goljan, and R. Du, "Lossless data embedding—New paradigm in digital watermarking," *EURASIP J. Appl. Signal Process.*, vol. 2002, no. 2, pp. 185–196, Feb. 2002.

[10] M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber, "Lossless generalized-LSB data embedding," *IEEE Trans. Image Process.*, vol. 14, no. 2, pp. 253–266, Feb. 2005.

[11] 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.

[12] J. Tian, "Reversible data embedding using a difference expansion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 890–896, Aug. 2003.

[13] 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.

[14] 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.

[15] H. J. Kim, V. Sachnev, Y. Q. Shi, J. Nam, and H. G. Choo, "A novel difference expansion transform for reversible data embedding," *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 3, pp. 456–465, Sep. 2008.

[16] W. L. Tai, C. M. Yeh, and C. C. Chang, "Reversible data hiding based on histogram modification of pixel differences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 6, pp. 906–910, Jun. 2009.

[17] Z. Ni, Y. Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 3, pp. 354–362, Mar. 2006.

[18] S. K. Lee, Y. H. Suh, and Y. S. Ho, "Reversible image authentication based on watermarking," in *Proc. IEEE ICME*, Jul. 2006, pp. 1321–1324.

[19] J. Hwang, J. Kim, and J. Choi, "A reversible watermarking based on histogram shifting," in *Proc. 5th IWDW*, 2006, pp. 348–361.

[20] 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.

[21] M. Fallahpour, "Reversible image data hiding based on gradient adjusted prediction," *IEICE Electron. Exp.*, vol. 5, no. 20, pp. 870–876, Oct. 2008.

[22] 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.

[23] 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.

[24] W. Hong, T. S. Chen, and C. W. Shiu, "Reversible data hiding for high quality images using modification of prediction errors," *J. Syst. Softw.*, vol. 82, no. 11, pp. 1833–1842, Nov. 2009.

[25] 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.

[26] W. Hong, "An efficient prediction-and-shifting embedding technique for high quality reversible data hiding," *EURASIP J. Adv. Signal Process.*, vol. 2010, Feb. 2010.

[27] G. Xuan, Y. Shi, J. Teng, X. Tong, and P. Chai, "Double-threshold reversible data hiding," in *Proc. IEEE ISCAS*, Jun. 2010, pp. 1129–1132.

[28] X. Gao, L. An, Y. Yuan, D. Tao, and X. Li, "Lossless data embedding using generalized statistical quantity histogram," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 8, pp. 1061–1070, Aug. 2011.

[29] 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.

[30] D. Coltuc, "Improved embedding for prediction-based reversible watermarking," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 873–882, Sep. 2011.

[31] C. Dragoi and D. Coltuc, "Improved rhombus interpolation for reversible watermarking by difference expansion," in *Proc. EUSIPCO*, 2012, pp. 1688–1692.

[32] H.-T. Wu and J. Huang, "Reversible image watermarking on prediction errors by efficient histogram modification," *Signal Process.*, vol. 92, no. 12, pp. 3000–3009, Dec. 2012.

[33] C. Qin, C.-C. Chang, Y.-H. Huang, and L.-T. Liao, "An inpainting-assisted reversible steganographic scheme using histogram shifting mechanism," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 7, pp. 1109–1118, Jul. 2013.

[34] D. Coltuc and I.-C. Dragoi, "Context embedding for raster-scan rhombus based reversible watermarking," in *Proc. ACM IH & MMSEC*, 2013, pp. 215–220.

[35] 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.

[36] D. Coltuc and J. M. Chassery, "Very fast watermarking by reversible contrast mapping," *IEEE Signal Process. Lett.*, vol. 14, no. 4, pp. 255–258, Apr. 2007.

[37] S. Weng, Y. Zhao, J. S. Pan, and R. Ni, "Reversible watermarking based on invariability and adjustment on pixel pairs," *IEEE Signal Process. Lett.*, vol. 15, no. 1, pp. 721–724, Jan. 2008.

[38] X. Wang, X. Li, B. Yang, and Z. Guo, "Efficient generalized integer transform for reversible watermarking," *IEEE Signal Process. Lett.*, vol. 17, no. 6, pp. 567–570, Jun. 2010.

[39] D. Coltuc, "Low distortion transform for reversible watermarking," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 412–417, Jan. 2012.

[40] F. Peng, X. Li, and B. Yang, "Adaptive reversible data hiding scheme based on integer transform," *Signal Process.*, vol. 92, no. 1, pp. 54–62, 2012.

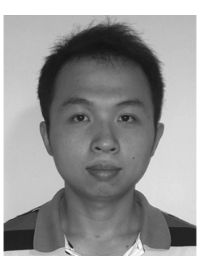
[41] X. Gui, X. Li, and B. Yang, "A novel integer transform for efficient reversible watermarking," in *Proc. ICPR*, 2012, pp. 947–950.

[42] T. Kalker and F. Willems, "Capacity bounds and constructions for reversible data-hiding," in *Proc. DSP*, vol. 1. 2002, pp. 71–76.

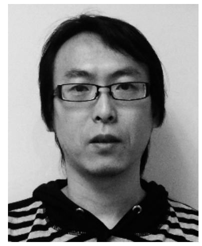
[43] W. Zhang, B. Chen, and N. Yu, "Improving various reversible data hiding schemes via optimal codes for binary covers," *IEEE Trans. Image Process.*, vol. 21, no. 6, pp. 2991–3003, Jun. 2012.

[44] S.-J. Lin and W.-H. Chung, "The scalar scheme for reversible information-embedding in gray-scale signals: Capacity evaluation and code constructions," *IEEE Trans. Inf. Forens. Security*, vol. 7, no. 4, pp. 1155–1167, Aug. 2012.

[45] P. Bas, T. Filler, and T. Pevny, "Break our steganographic system—The ins and outs of organizing BOSS," in *Proc. 13th Int. Workshop Inf. Hiding*, vol. 6958, 2011, pp. 59–70.

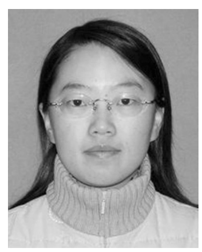


Bo Ou was born in Hunan, China, in 1985. He received the B.S. degree from Beijing Jiaotong University, Beijing, China, in 2008, where he is currently pursuing the Ph.D. degree with the Institute of Information Science. His current research interests include image processing and digital watermarking.



Xiaolong Li received the B.S. degree from Peking University, Beijing, China, the M.S. degree from Ecole Polytechnique, Palaiseau, France, and the Ph.D. degree in mathematics from ENS de Cachan, Cachan, France, in 1999, 2002, and 2006, respectively. Before joining Peking University, he was a Post-Doctoral Fellow with Peking University, Beijing, from 2007 to 2009. His current research interests include image processing and information hiding.

Yao Zhao (M'06–SM'12) received the B.S. degree from Fuzhou University, Fuzhou, China, in 1989, and the M.E. degree from Southeast University, Nanjing, China, in 1992, both from the Radio Engineering Department, and the Ph.D. degree from the Institute of Information Science, Beijing Jiaotong University (BJTU), Beijing, China, in 1996. He became an Associate Professor at BJTU in 1998 and became a Professor in 2001. From 2001 to 2002, he was a Senior Research Fellow with the Information and Communication Theory Group, Faculty of Information Technology and Systems, Delft University of Technology, Delft, The Netherlands. He is currently the Director of the Institute of Information Science, BJTU. His current research interests include image/video coding, digital watermarking and forensics, and video analysis and understanding. He is now also leading several national research projects from the 973 Program, 863 Program, and the National Science Foundation of China. He serves on the editorial boards of several international journals, including as an Area Editor of *Signal Processing: Image Communication*, and as an Associate Editor of *Circuits, System, and Signal Processing* (Springer). He was named a Distinguished Young Scholar by the National Science Foundation of China in 2010.



Rongrong Ni received the Ph.D. degree in signal and information processing from the Institute of Information Science, Beijing Jiaotong University, Beijing, China, in 2005. Since 2005, she has been on the faculty of the School of Computer and Information Technology and the Institute of Information Science, where she has been an Associate Professor since 2008. Her current research interests include image processing, data hiding and digital forensics, pattern recognition, and computer vision. She was selected in Beijing Science and Technology Stars Projects in 2008, and was awarded the Jeme Tien Yow Special Prize in Science and Technology in 2009. She is currently the Principal Investigator of three projects funded by the Natural Science Foundation of China. She participates in the 973 Program, the 863 Program, and international projects as the backbone. She has published more than 80 papers on academic journals and conferences, and holds six national patents. She is a member of IET and EURASIP.

Yun-Qing Shi (F'05) has been with the New Jersey Institute of Technology, Newark, NJ, USA, since 1987. He received the M.S. degree from Shanghai Jiao Tong University, Shanghai, China, and the Ph.D. degree from University of Pittsburgh, Pittsburgh, PA, USA. His current research interests include digital data hiding, forensics and information assurance, and visual signal processing and communications. He is an author or co-author of 300 papers, one book, and five book chapters, and an editor of ten books. He holds 27 U.S. patents. He received the Innovators Award 2010 by the New Jersey Inventors Hall of Fame for Innovations in Digital Forensics and Security. His U.S. patent 7 457 341 entitled *System and Method for Robust Reversible Data Hiding and Data Recovery in the Spatial Domain* won the 2010 Thomas Alva Edison Patent Award by the Research and Development Council of New Jersey. He served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS (II), and a few other journals; he also served as the Technical Program Chair of IEEE ICME in 2007, a Co-Technical Chair of IWDW in 2006, 2007, 2009, 2010, 2011, 2012, and 2013, and IEEE MMSP in 2005, a Co-General Chair of IEEE MMSP in 2002, and a Distinguished Lecturer of IEEE CASS. He is a member of a few IEEE technical committees.