

RESEARCH

Open Access



Non-integer expansion embedding techniques for reversible image watermarking

Shijun Xiang* and Yi Wang

Abstract

This work aims at reducing the embedding distortion of prediction-error expansion (PE)-based reversible watermarking. In the classical PE embedding method proposed by Thodi and Rodriguez, the predicted value is rounded to integer number for integer prediction-error expansion (IPE) embedding. The rounding operation makes a constraint on a predictor's performance. In this paper, we propose a non-integer PE (NIPE) embedding approach, which can proceed non-integer prediction errors for embedding data into an audio or image file by only expanding integer element of a prediction error while keeping its fractional element unchanged. The advantage of the NIPE embedding technique is that the NIPE technique can really bring a predictor into full play by estimating a sample/pixel in a noncausal way in a single pass since there is no rounding operation. A new noncausal image prediction method to estimate a pixel with four immediate pixels in a single pass is included in the proposed scheme. The proposed noncausal image predictor can provide better performance than Sachnev et al.'s noncausal double-set prediction method (where data prediction in two passes brings a distortion problem due to the fact that half of the pixels were predicted with the watermarked pixels). In comparison with existing several state-of-the-art works, experimental results have shown that the NIPE technique with the new noncausal prediction strategy can reduce the embedding distortion for the same embedding payload.

Keywords: Reversible watermarking; Non-integer prediction error; Expansion embedding; Noncausal prediction

1 Introduction

Reversible watermarking techniques embed data in a host signal (for example, an audio/image) and allow for the original digital image to be exactly recovered. This is very useful in some applications, especially in medical, military, and legal domains. There are two important requirements for reversible watermarking techniques: 1) a large embedding capacity and 2) a low watermark distortion. The two requirements conflict with each other since embedding more data into a work will cause bigger distortion. A desirable technique should embed the same capacity with lower distortion or vice versa. During the recent 10 years, the reversible watermarking has been an active research domain. In the literature, several types of reversible watermarking algorithms has been proposed:

- Type I algorithms use modulo-arithmetic-based additive spread frequency techniques [1, 2], some of

which provide robustness but often cause salt-and-pepper artifacts due to the many wrapped around pixel intensity. In this direction, a different approach proposed by Vleeschouwer et al. [3] can reduce the artifacts by using the circular interpolation of the bijective transform of image histogram. Finally, good results are reported by using the integer-to-integer wavelet-based reversible watermarking in [4].

- The direct method to reversible watermarking is to compress a set of selected features from an image in a lossless way and substitute the selected features with their compressed versions and the watermark data [5–8], referred to as type II algorithms in the literature. In order to ensure the watermark imperceptivity, the selected features are usually in the least significant bit area, such as the generalized LSB (g-LSB) embedding algorithm [8], which is an extension of the work [5]. Theoretical analysis of reversible watermarking has been presented by Kalker and Willems in [7]. Compared with type I algorithms, type II ones can provide higher payload.

*Correspondence: xiangshijun@gmail.com
Department of Electronic Engineering, School of Information Science and Technology, Jinan University, Guangzhou 510632, China

- The third type of reversible watermarking algorithms can be classified as difference expansion (DE) embedding methods, in which a common feature is to use difference operators to create features with a small magnitude, and further expand these features in order to create vacancies for embedding the watermark and the auxiliary information. The DE embedding technique was originally developed by Tian [9] and improved in [10–13]. The capacity in Tian's method is close to 0.5 bpp in a single pass by using two adjacent pixels as a group. By extending the DE to a generalized integer transform, the auxiliary information can be reduced with groups of three or four pixels [10]. In [11], the sorting step in the wavelet domain was introduced to expand those smaller pixel difference. The DE scheme was introduced for 2-D vector map in [12], and the location map was reduced in [13].
- A fruitful research direction, proposed by Thodi and Rodriguez [14], is prediction-error expansion (PE) embedding technique. Comparing with the DE-based methods, one of the advantages of the PE technique is that it significantly adds the number of the feature elements that expanded for data embedding. The other advantage is that a predictor generates feature elements that are often smaller in magnitude than the feature elements generated by a difference operator. With embedding into each pixel, the PE embedding techniques provided the maximal capacity up to 1 bpp in a single pass. So far, several versions of PE reversible watermarking algorithms have been proposed in [14–19] and others in order to improve the performance of the PE schemes by focusing on the reduction of the size of the auxiliary information (with the use of the sorting and histogram shifting techniques [15]), the reduction of the prediction error (by using multiple predictors [16], the prediction by flooring the average value of the four immediate pixels [15] and adaptive prediction [17]), and the reduction of the embedding distortion (by the pixel selection [17] and the low distortion transform by splitting the difference between the current pixel and its prediction context [18, 19]). These existing PE-based schemes share a common property, that is, the predicted value or its variety was rounded to integer value for expansion embedding.

Reversible watermarking algorithms have also been proposed for digital audio files by expansion embedding [20–22]. There are other novel reversible watermarking approaches which are worth mentioning, too. Data embedded into the histogram bins has been proposed by Leest et al. in [23]. Additive embedding strategy by combining histogram shifting technique [24] and bilinear

interpolation prediction has been proposed by Chen et al. in [25, 26].

This paper proposes a non-integer PE (NIPE) embedding strategy, which can proceed non-integer prediction errors for data embedding by expanding only the integer element of a prediction error while keeping the fractional element unchanged. Furthermore, we propose a new non-causal image prediction method for the NIPE technique. In comparison with the classical PE technique [14], the NIPE strategy has lower embedding distortion for non-integer prediction values by using better predictor. The new predictor can provide better performance than existing causal predictors (such as difference predictor, MED predictor, and GAP) due to the fact that there is no rounding operation in the prediction phase. Comparing with Sachnev et al.'s noncausal double-set prediction method [15], the proposed prediction method can further reduce the prediction errors by predicting data in a single pass (which avoids the distortion problem due to the fact that half of the pixels in the second set predicted with the watermarked pixels in the first set). Experimental results for some standard test images have shown that the proposed NIPE method with the noncausal image predictor can embed more payload for the same watermark distortion.

The outline of this paper is as follows. In the next section, the proposed NIPE embedding technique is introduced. This is followed by a description of a new prediction strategy and an image predictor. We then address the proposed reversible watermarking scheme and compare the proposed prediction method with existing typical predictors. Furthermore, the watermarking scheme's performance is tested against other typical reversible image watermarking works. Finally, we draw the conclusions.

2 Prediction-error expansion embedding

PE embedding is a technique to expand a prediction error to create a vacant position and insert a bit into the vacant position, generally at the least significant bit (LSB). The PE-based scheme was originally developed by Thodi and Rodriguez [14] and later developed by other researchers (such as [16–19]). In these methods, causal predictive methods with only past pixels are often applied so that the predicted value or its variety can be rounded to integer value for integer prediction-error expansion (IPE) embedding.

In this section, we proposed a NIPE embedding technique, which really brings a predictor into full play to reduce the prediction errors in comparison with the IPE embedding technique. In the NIPE-based method, the predicted value is no longer rounded to integer number. This is beneficial to using more efficient predictor in the NIPE¹. The basic principles of the IPE method [14] and the proposed NIPE method are described as follows.

2.1 Classical IPE embedding

In the IPE embedding technique [14], the prediction error is the difference between a pixel intensity y and its estimate \hat{y} (which should be rounded to integer value if it is not), denoted by $e = y - \hat{y}$. After embedding a bit w , the watermarked prediction error is

$$e_w = 2 \times e + w, \tag{1}$$

and the marked pixel intensity is

$$y_w = \hat{y} + e_w = \hat{y} + 2 \times e + w. \tag{2}$$

Since y_w is an integer number, \hat{y} and e are required to be integer values for expansion embedding. This is why we denote Thodi and Rodriguez's method as the IPE embedding scheme in this paper. It is worth noting that the condition that \hat{y} should be integer makes an undesirable requirement, that is, the prediction context often contains only *causal pixels* so as to obtain the same predicted value \hat{y} in the decoder.

The hidden bit, w , is extracted from the LSB of e_w and the original pixel intensity y is recovered by

$$w = \text{mod}(e_w, 2), y = \hat{y} + \lfloor \frac{e_w}{2} \rfloor, \tag{3}$$

where $\text{mod}(e_w, 2)$ is the remainder on division of e_w by 2 and $\lfloor \frac{e_w}{2} \rfloor$ represents the greatest integer less than or equal to $\frac{e_w}{2}$.

2.2 NIPE embedding technique

Sections 2.1 shows that the IPE embedding approach [14] is suffering from an undesired requirement, that is, the prediction value should be rounded to integer number for expansion embedding. This will restrain a predictor's performance since the prediction context is restricted to only causal pixels (past pixels) in order to generate the same predicted value in the decoder. In this section, we propose a new expansion embedding approach, one of the advantages of which is able to deal with non-integer prediction values for data embedding. The other, more important, advantage of the approach is that the current pixel can be estimated in a single pass by using noncausal predictive way to improve prediction performance².

From the expression $y_w = \hat{y} + 2 \times e + w$ in Eq. (2), we find that in order to recover the marked pixel y_w , not exact of an integer \hat{y} is needed, but of the sum of $\hat{y} + 2 \times e$. Towards this direction, the basic idea of the proposed approach in this paper is to allow \hat{y} to take non-integer value but make sure that the combination of \hat{y} and e takes integer value for hiding a given bit w .

For a pixel in intensity y , the prediction error e is a non-integer value when its estimate \hat{y} takes non-integer number. In this case, split the non-integer error e into two parts: integer part ℓ ($\ell = \text{fix}(e)$) and fractional part δ ($\delta = e - \ell$). Here, $\text{fix}(\cdot)$ is a function to strip off the fractional part of its argument and returns the integer

part. The function does not perform any form of rounding or scaling, e.g., $\text{fix}(-4.4) = -4$ and $\text{fix}(5.4) = 5$. The basic idea of the NIPE embedding technique expands only the integer part of a prediction error for data embedding while keeping the fractional part unchanged. The detail is described as follows.

In the encoder, the watermarked prediction error is computed by

$$e_w = \begin{cases} 2 \times \ell + \delta + w = e + \ell + w, & \text{if } e \geq 0 \\ 2 \times \ell + \delta - w = e + \ell - w, & \text{otherwise,} \end{cases} \tag{4}$$

where w is a binary bit, taking either 0 or 1. Equation (4) can make sure that the fractional element of e_w is equal to that of e . This is beneficial to recover the watermark bit and the original pixel intensity in the decoder.

The resulting watermarked pixel is

$$y_w = \hat{y} + e_w = \begin{cases} \hat{y} + e + \ell + w = y + \ell + w, & \text{if } e \geq 0 \\ \hat{y} + e + \ell - w = y + \ell - w, & \text{Otherwise.} \end{cases} \tag{5}$$

Equation (5) shows that even though \hat{y} and e take non-integer values, the watermarked pixel y_w is an integer number.

In the decoder, the hidden bit w is extracted from e_w and the original pixel y is restored by

$$w = \text{mod}(\ell_w, 2), \text{ and } y = \hat{y} + \text{fix}\left(\frac{\ell_w}{2}\right) + \delta_w, \tag{6}$$

where ℓ_w is the integer element of e_w and $\delta_w = e_w - \ell_w$.

Take two simple examples to show the proposed NIPE technique:

- Case 1 Let $y = 100$, $w = 1$, and $e = 100 - 101.4 = -1.4$ when $\hat{y} = 101.4$. In the encoder,
 - $e_w = e + \ell - w = -1.4 - 1 - 1 = -3.4$,
 - $\ell_w = \text{fix}(e_w) = -3$, $\delta_w = e_w - \ell_w = -0.4$,
 - $w = \text{mod}(\ell_w, 2) = 1$,
 - $y = \hat{y} + \text{fix}(\frac{\ell_w}{2}) + \delta_w = 101.4 + \text{fix}(\frac{-3}{2}) - 0.4 = 100$.
- Case 2 : Let $y = 100$, $w = 1$ and $e = 100 - 97.4 = 2.6$ when $\hat{y} = 97.4$. In the encoder,
 - $e_w = e + \ell + w = 2.6 + 2 + 1 = 5.6$,
 - $\ell_w = \text{fix}(e_w) = 5$, $\delta_w = e_w - \ell_w = 0.6$,
 - $w = \text{mod}(\ell_w, 2) = 1$,
 - $y = \hat{y} + \text{fix}(\frac{\ell_w}{2}) + \delta_w = 97.4 + \text{fix}(\frac{5}{2}) + 0.6 = 100$.

We can see from these two cases that the NIPE scheme can effectively deal with the non-integer prediction values for reversible watermarking.

Equations (4), (5), and (6) form the proposed NIPE embedding strategy, which can avoid the rounding operation in the IPE.

2.3 Distortion analysis of NIPE and IPE embedding

Let y be a pixel and let \hat{y} be an estimate of y computed on a neighborhood of y . Let further w is the binary bit to be embedded. Let $y_w = y + p_w$, where p_w is the watermark distortion on y . For an image with the same predictor, the embedding distortion from the proposed NIPE scheme is lower than from the classical IPE scheme [14]. This can be explained as follows:

- Case 1 \hat{y} is integer: From (2) in the IPE, $p'_w = y - \hat{y} + w$. From (5) in the NIPE, $p''_w = \text{fix}(y - \hat{y}) + w = y - \hat{y} + w$. In this case, $p'_w = p''_w$, indicating that the NIPE has the same embedding distortion as the IPE. For example, if $y = 100$, $w = 1$, $\hat{y} = 102$. With NIPE, $p''_w = \text{fix}(100 - 102) + 1 = -1$; with IPE, $p'_w = 100 - 102 + 1 = -1$;
- Case 2 \hat{y} is non-integer and $y - \hat{y} > 0$: In this case, $p'_w = y - \lfloor \hat{y} \rfloor + w$, $p''_w = \text{fix}(y - \hat{y}) + w = y - \text{fix}(\hat{y}) - 1 + w$. Since the estimate of a pixel \hat{y} is always positive, we have $p''_w = y - \text{fix}(\hat{y}) - 1 + w = y - \lfloor \hat{y} \rfloor - 1 + w$, that is $p'_w = p''_w + 1$, meaning that the IPE scheme will introduce more distortion in this case. For example, if $y = 100$, $w=1$, $\hat{y} = 98.4$. With NIPE, $p''_w = \text{fix}(100 - 98.4) + 1 = 2$; With IPE, $p'_w = 100 - 98 + 1 = 3$.
- Case 3 \hat{y} is non-integer and $y - \hat{y} < 0$: In this case, $p'_w = y - \lfloor \hat{y} \rfloor + w$. Referring to (5), $p''_w = \text{fix}(y - \hat{y}) - w = y - \text{fix}(\hat{y}) - w = y - \lfloor \hat{y} \rfloor - w$. From $y - \lfloor \hat{y} \rfloor = p'_w - w$, we have $p''_w = p'_w - 2w$, indicating the distortion from the NIPE is higher when $w = 1$. When $w = 0$, the NIPE and IPE has the same embedding distortion. For example, if $y = 100$, $\hat{y} = 101.4$. With NIPE, $p''_w = \text{fix}(100 - 101.4) - w = -1 - w$; with IPE, $p'_w = 100 - 101 + w = -1 + w$. If w is 1, then $p''_w = -2$ and $p'_w = 0$; If w is 0, then $p''_w = p'_w = -1$.

In statistics, the predictor error ($e = y - \hat{y}$) is equal to the probability of positive or negative and the watermark bit (w) is also equal to the probability of 1 or 0. As a result, the NIPE has the same embedding distortion as the classical IPE scheme (e.g., the same PSNR value).

3 Pixel prediction

Image prediction is an important step in lossless compression coding applications [27–30]. For an image, each pixel is estimated from its neighborhood to generate the prediction error. Usually, the mean value and variance of the predicted errors are smaller than that of the original pixels. This is beneficial to improve coding efficiency. There are two main prediction ways: 1) causal prediction [27, 30]

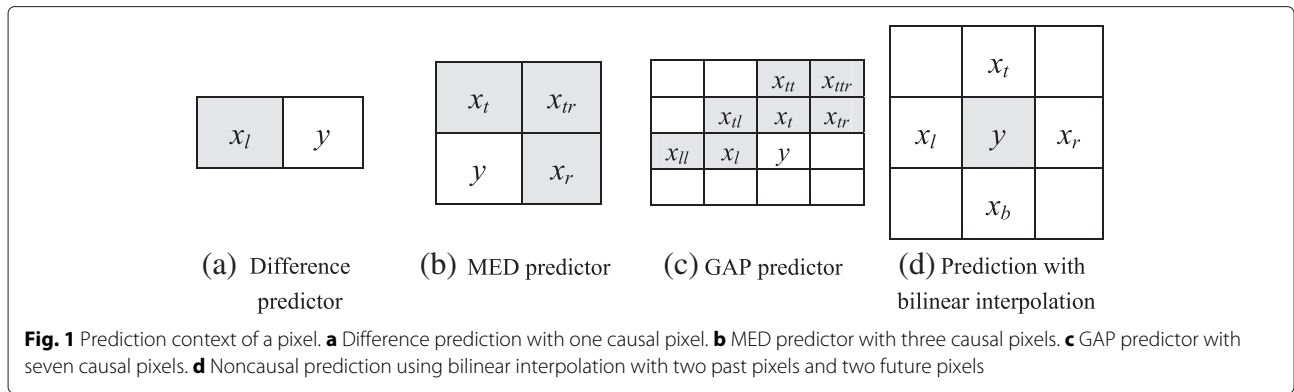
and 2) noncausal prediction [28, 29]. The main difference between causal and noncausal predictive ways is that the prediction context of the former is restricted to only past pixels while the latter uses past and future pixels. By using causal prediction, the predicted value can be rounded to integer number for enhancing coding efficiency since the prediction context of a pixel has been known, such as the median edge direction (MED) predictor in JPEG-LS standard [30]. Noncausal predictive ways are beneficial to reduce the prediction errors for image vector quantization coding [28, 29] and speech compression [31].

3.1 Image predictors for reversible watermarking

Image prediction is also an important step in the expansion embedding-based reversible watermarking [9, 10, 14, 15, 18, 19]. Usually, a better predictor can reduce the prediction error for improving the payload capacity with the same embedding distortion. Figure 1 plots four typical prediction operations by estimating the current pixel y with different neighboring pixels: (a) difference predictor used in [9, 10], (b) MED predictor used in [14, 17–19], (c) the gradient adaptive predictor (GAP) used in [18, 19], and d) prediction with four immediate pixels used in [15, 26]. As shown in Fig. 1, data prediction in the difference predictor, MED predictor, and GAP is restricted only causal pixels while scanning an image in a raster order. The predictor plotted in Fig. 1d is a noncausal predictive method using four immediate pixels of the current pixel as the context, including two past pixels, and two future pixels.

In the DE-based reversible watermarking [9], the difference of two adjacent pixels is expanded to create space to embed the data and the auxiliary information. The auxiliary information can be reduced by extending the DE on groups of three or four neighboring pixels [10]. In the PE-based reversible watermarking [14, 17–19], the difference is a prediction error between a pixel and its estimate. Since the predictor can apply several causal pixels (such as MED, GAP, and others designed for lossless data compression [30, 32]) as the context, the prediction error in magnitude is usually smaller than the difference of two adjacent pixels. In [14, 17, 18], the MED predictor was used to measure the payload capacity and the embedding distortion. Also in [18], the GAP and simplified GAP (SGAP) are applied to show how to reduce the embedding distortion by marking the current pixel and its context. No matter the MED predictor, GAP or SGAP, the data prediction is restricted only causal pixels for the IPE embedding. For example, the MED predictor combines three past pixels (x_t, x_{tr}, x_r) as the context of the current pixel (y) as illustrated in Fig. 1b. The output of the MED predictor is

$$\hat{y} = \begin{cases} \max(x_t, x_r), & \text{if } x_{tr} \leq \min(x_t, x_r) \\ \min(x_t, x_r), & \text{if } x_{tr} \geq \max(x_t, x_r) \\ x_t + x_r - x_{tr}, & \text{otherwise.} \end{cases} \quad (7)$$



The same predictor was also considered as the median of three simple linear predictors: x_l , x_r , and $x_l + x_r - x_{lr}$ [33], that is, $\hat{y} = \frac{2x_l + 2x_r - x_{lr}}{3}$.

Comparing with causal prediction ways, noncausal prediction approaches can improve the prediction precision by using more neighboring samples. The prediction method described in [15] was very different one, which can combine the IPE scheme and noncausal prediction with bilinear interpolation in a way that the image is divided into two sets (like a chess board, the *black* set, and the *white* set). In the first pass, the pixel in the black set is predicted with four immediate pixels in the white set to generate an integer difference for IPE-based expansion embedding. Then, the white set is predicted by using the marked version of the black set for the second pass of embedding. Similar idea has been used for additive PE embedding [26] which divided an image into four sets and further embedded the data into the sets one by one by using multi-passes embedding. The problem in the noncausal predictors [15, 26] is that part of the pixels are predicted with the watermarked pixels instead of the original ones. This will introduce some unnecessary distortion since the distribution of the prediction errors estimated from the marked pixels has a bigger variance than that estimated from the original pixels.

In the following section, we will present a noncausal predictive model for the NIPE embedding scheme. Comparing with the noncausal prediction methods presented in [15, 26], the predictor proposed in this paper can predict all the pixels before watermarking. This is beneficial to reducing the distortion due to part of the pixels predicted by the modified pixels.

3.2 Proposed noncausal prediction model

In this section, a new method incorporating data prediction not restricted to only causal pixels is designed for the proposed NIPE technique. This noncausal predictive method can predict data in a single pass.

Assume there is a time-discrete signal Y of length N , $Y = \{y_1, y_2, \dots, y_N\}$ with $y_i \in \{0, 1, \dots, 2^m - 1\}^N$,

and where m indicates the number of bits used to represent a point (could be a sample or a pixel). The signal after the prediction is denoted by \hat{Y} . The residual signal is $E = Y - \hat{Y}$. Here, the predicted value is a linear combination of past and future pixels/samples:

$$\hat{y}_i = \sum_{t=1}^p a_{i-t}y_{i-t} + \sum_{t=1}^p a_{i+t}y_{i+t}, p < i < N - p + 1, \quad (8)$$

where $\sum_{t=1}^p a_{i-t}y_{i-t}$ is the linear combination of p past pixels/samples, $\sum_{t=1}^p a_{i+t}y_{i+t}$ that of p future pixels/samples. The prediction error is computed as

$$d_i = y_i - \hat{y}_i = y_i - \sum_{t=1}^p a_{i-t}y_{i-t} - \sum_{t=1}^p a_{i+t}y_{i+t}. \quad (9)$$

Since there are no any rounded operations on the predicted value \hat{y}_i , Eq. (9) can be rewritten as

$$y_{i+p} = \frac{y_i - d_i - \sum_{t=1}^p a_{i-t}y_{i-t} - \sum_{t=1}^{p-1} a_{i+t}y_{i+t}}{a_{i+p}}. \quad (10)$$

Equation (10) shows that the information of the first $2p$ pixels/samples $\{y_1, y_2, \dots, y_{2p}\}$ is needed to be saved as part of the residual signal for the recovery of the original signal. From the data series $\{y_1, y_2, \dots, y_{2p}\}$ and d_{p+1} , the pixel/sample y_{2p+1} can be recovered. Then, $y_{2p+2}, y_{2p+3}, \dots, y_N$ can be restored in sequential order.

The resulting signal above can be denoted by $D = \{y_1, y_2, \dots, y_{2p}, d_{p+1}, d_{p+2}, \dots, d_{N-p}\}$. The data $\{y_1, y_2, \dots, y_{2p}\}$ can be further predicted to generate the difference $\{e_1, e_2, \dots, e_{2p}\}$ with the predictor described in Section 3.3. Denote $\hat{d}_i = e_{i+p}, p < i < N - p + 1$. As a result, the original signal Y can be further predicted as $E = \{e_1, e_2, \dots, e_{2p}, e_{2p+1}, e_{2p+2}, \dots, e_N\}$.

The above model shows that the noncausal prediction model can be performed in a single pass since there is no rounding operation on the predicted value \hat{y}_i in Eq. (8). For a two-dimensional image, it can be mapped into the one-dimensional form by using a scanning operation (e.g., in a raster scan or zigzag scan order).

3.3 Predictor with $p = 1$

The section above has addressed the basic principle of the proposed prediction model. When $p = 1$, the prediction model can be simplified as follows. Beginning from the second pixel, the pixel y_i is predicted by averaging its two closest neighbors (y_{i-1}, y_{i+1}):

$$\hat{y}_i = \frac{y_{i-1} + y_{i+1}}{2}, 1 < i < N. \quad (11)$$

The difference is computed as

$$d_i = y_i - \hat{y}_i = y_i - \frac{y_{i-1} + y_{i+1}}{2}, 1 < i < N. \quad (12)$$

From Eq. (12), the original pixel y_{i+1} is recovered by

$$y_{i+1} = 2y_i - y_{i-1} - 2d_i, 1 < i < N, \quad (13)$$

Equation (13) indicates that when the first two pixels y_1 and y_2 are saved, the third pixel y_3 can be recovered by referring to the prediction error d_2 in Eq. (13), then recovering y_4, y_5 and the other pixels in sequential order. Let $e_1 = y_1, e_2 = y_2 - y_1$ and $e_{i+1} = d_i$. Overall, the output of the predictor is denoted as $E = \{e_1, e_2, e_3, \dots, e_{N-1}\}$. The predictor in the case of $p = 1$ has been fully proven effective for digital audio in our earlier work [34].

3.4 Proposed noncausal prediction for image

In this section, we design a new image prediction method by referring to the proposed prediction model. Though this prediction method looks like the one described in [15], it is a different one with higher prediction accuracy. In [15], half of the pixels were predicted with the modified pixels instead of the original pixels. In the proposed method, all the pixels are predicted with the original pixels. This explains why the proposed NIPE watermarking scheme has better performance. The detail on the proposed image predictor is addressed as follows:

1) For a given 2-D image I in size $R \times C$, use the following projection into the 1-D form:

$$\begin{cases} y_{(i-1) \cdot R + j} = I(i, j) \\ N = R \times C, \end{cases} \quad (14)$$

where $I(i, j)$ is the pixel intensity in the i th row and the j th column, satisfying $1 \leq i \leq R$ and $1 \leq j \leq C$. Denote the resulting 1-D signal by $Y = \{y_1, y_2, \dots, y_N\}$.

2) After the projection in *step 1*, predict Y by referring to the proposed noncausal prediction model. Let $p = C$, that is to say, the estimate of the current pixel y_k is a linear combination of C past and C future neighboring pixels, formulated as

$$\hat{y}_k = \sum_{t=1}^C a_{k-t} y_{k-t} + \sum_{t=1}^C a_{k+t} y_{k+t}, C+1 \leq k \leq N-C. \quad (15)$$

Consider the strong correlation property between the current pixel (y_k) and its four immediate pixels in the top (y_{k-C}), left (y_{k-1}), right (y_{k+1}), and bottom (y_{k+C}). The current pixel (y_k) is predicted by

$$\hat{y}_k = \begin{cases} \frac{y_{k-C} + y_{k+C} + y_{k+1}}{3}, & \text{if } \text{mod}(k, C) = 1 \\ \frac{y_{k-C} + y_{k-1} + y_{k+C}}{3}, & \text{if } \text{mod}(k, C) = 0 \\ \frac{y_{k-C} + y_{k-1} + y_{k+C} + y_{k+1}}{4}, & \text{others,} \end{cases} \quad (16)$$

where the condition $\text{mod}(k, C) = 1$ means that the pixel in the first column is predicted with the average value of three neighbors in the top, right, and bottom, as shown in Fig. 2a. For the pixel in the last column (satisfying $\text{mod}(k, C) = 0$, see Fig. 2c), the prediction context includes three immediate neighbors in the top, left, and bottom. The other pixels are estimated by averaging four immediate pixels of the present pixel, as illustrated in Fig. 2b. *Step 2* will predict the pixels from the second row to the second last row. As analyzed in Section 3.2, the information of the first $2 \times C$ pixels (in the first and second rows) is saved for the inverse prediction. Let $e_k = y_{k-C} - \hat{y}_{k-C}, 2C+1 \leq k \leq N$. The resulting signal from step 2 is denoted by $D = \{y_1, y_2, \dots, y_{2C}, e_{2C+1}, e_{2C+2}, \dots, e_N\}$.

3) In order to further reduce the prediction errors, the first $2C$ elements in D (the pixels in the first two rows) are predicted before data embedding. This can be done by raster scanning the pixels in the first two rows and predicting the $2 \times C$ pixels by using the noncausal predictor in the case of $p = 1$, described in Section 3.3. The first $2C$ pixels are predicted and kept as $E_1 = \{e_1, e_2, \dots, e_{2C}\}$.

4) From steps 2 and 3, we have $E = \{e_1, e_2, \dots, e_{2C}, e_{2C+1}, e_{2C+2}, \dots, e_N\}$, which is the output of the proposed image predictors in this paper.

5) The original image can be restored from E by performing the inverse prediction operations. Referring to step 3 and Section 3.3, the first $2C$ pixels are recovered by

$$\begin{cases} y_1 = e_1 \\ y_2 = y_1 + e_2 \\ y_i = 2y_{i-1} - y_{i-2} - 2e_i, 3 \leq i \leq 2C. \end{cases} \quad (17)$$

Once the first $2C$ pixels are recovered, the other pixels can be recovered by referring to Eq. 16 in step 2 with the following expression:

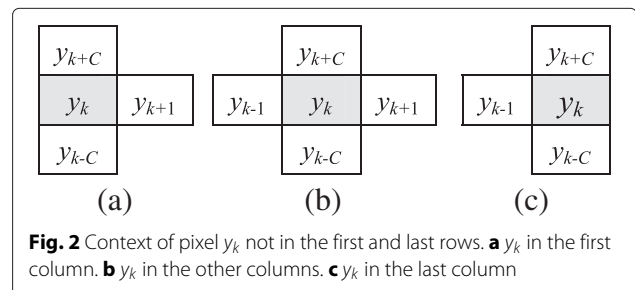


Fig. 2 Context of pixel y_k not in the first and last rows. **a** y_k in the first column. **b** y_k in the other columns. **c** y_k in the last column

$$y_{k+C} = \begin{cases} 3y_k - y_{k-C} - y_{k+1} - 3e_{k+C}, & \text{if } \text{mod}(k, C) = 1 \\ 3y_k - y_{k-C} - y_{k-1} - 3e_{k+C}, & \text{if } \text{mod}(k, C) = 0 \\ 4y_k - y_{k-C} - y_{k+1} - y_{k-1} - 4e_{k+C}, & \text{others,} \end{cases} \quad (18)$$

$$e_{wi} = \begin{cases} 2 \times \ell_i + \delta_i + w_i & \text{if } \ell_i \in [0, T] \\ 2 \times \ell_i + \delta_i - w_i & \text{if } \ell_i \in [-T, 0) \\ e_i + T + 1 & \text{if } \ell_i > T \\ e_i - T - 1, 7 & \text{if } \ell_i < -T, \end{cases} \quad (19)$$

where $C < k \leq N - C$. Finally, the original image is restored from the residual signal E .

3.5 Comments

From the analysis above, we can see that the proposed image predictor can predict pixels with its four immediate pixels which are not modified. Comparing with existing typical predictors, it can further improve the performance due to the following facts:

1. The data is predicted in a noncausal way that most of the pixels can be predicted with four immediate pixels. This improves prediction performance since some predictors only use causal pixels for data prediction, such as MED predictor.
2. All the pixels can be predicted in a single pass with original pixels as context. This is beneficial to avoid the the distortion problem due to part of the pixels predicted by the modified pixels, such as the predictor in [15].
3. In the literature, the IPE watermarking scheme [15] has a satisfactory performance. The NIPE technique with the predictor above can further reduce the embedding distortion for the same payload by avoiding part of the pixels predicted by the modified pixels.

4 Proposed watermarking scheme

The proposed watermarking scheme is a combination of existing techniques (histogram shifting in [14, 24]) and new techniques (NIPE embedding and noncausal image prediction).

4.1 Prediction expansion with histogram shifting

The histogram shifting method, introduced in [14, 24], is an efficient reversible watermarking technique to enhance fidelity of the marked signal and avoid overlapping problems caused by expansion embedding. The combination of histogram shifting and IPE has been previously addressed in [14, 15]. Here, we present how to combine the NIPE method with histogram shifting technique. We adopt a positive threshold value T to control the embedding distortion. Specifically, only those prediction values in $[-T, T]$ are selected for NIPE embedding (denoted as the expanding set S_1), the prediction errors not in the range $[-T, T]$ are going to be shifted (denoted as the shiftable set S_2) to avoid overlapping problems. The reversible watermarking rules are formulated as follows.

where ℓ_i is integer part of the i th prediction error, e_i , satisfying $e_i = \ell_i + \delta_i$. The marked prediction error is denoted by e_{wi} after the bit w_i is inserted.

The decoder recovers the original prediction error e_i and the bit w_i from e_{wi} by:

$$e_i = \begin{cases} \text{fix}(\frac{\ell_{wi}}{2}) + \delta_{wi} & \text{if } \ell_{wi} \in [-2T - 1, 2T + 1] \\ e_{wi} - T - 1 & \text{if } \ell_{wi} > 2T + 1 \\ e_{wi} + T + 1, & \text{if } \ell_{wi} < -2T - 1 \end{cases} \quad (20)$$

and

$$w_i = \text{mod}(\ell_{wi}, 2), \text{ if } e_{wi} \in [-2T - 1, 2T + 1], \quad (21)$$

where $\ell_{wi} = \text{fix}(e_{wi})$ is the integer element of e_{wi} and $\delta_{wi} = e_{wi} - \ell_{wi}$. It is worth noting that $\delta_{wi} = \delta_i$ since the encoder only expands the integer part while keeping the fraction element unchanged. Finally, the original pixels are recovered from the original prediction errors by performing inverse prediction operation in Section 3.4.

The ratio between the sets S_1 and S_2 can be controlled by changing the embedding threshold T . The bigger the threshold value T , the higher the embedding payload, the more the embedding distortion is.

4.2 Capacity analysis

The marked images may suffer from overflow and underflow problems due to expansion embedding and histogram shifting operations. Towards this direction, an embedding testing step is first performed to pick up those pixels with overflow or underflow problems. The testing process has been addressed in [14, 15] in detail. For an image with m -bit representation, when a watermarked pixel in intensity is not in the range $[0, 2^m - 1]$, labeled as a *bad pixel*. The bad pixels in position will be saved and embedded with the payload to indicate the expandable locations.

Usually, the size of images is smaller than 5000×5000 . After the mapping in (14), the length is $5000 \times 5000 = 25,000,000 < 2^{25}$. That is to say, 25 bits of information is required for indicating a bad pixel. In addition, 7 bits of information is required for sending the embedding threshold parameter T to the decoder. Without the consideration of recursive embedding, the maximal embedding rate of the proposed reversible watermarking method can be estimated by:

$$C = \frac{N_1 - 25 \times N_p - 25 - 7}{N}, \quad (22)$$

where C is the maximal embedding rate (bounded to 1), N_1 the length of the expandable set S_1 , N_p the number

of the bad pixels, and N the number of the cover image pixels.

5 Encoder and decoder

The proposed reversible watermarking scheme, as illustrated in Fig. 3, can be used for image or audio files. In this paper, we take some standard images for experimental testing. In the embedding, the maximal capacity (P_{max}) of an image is first computed by using the proposed reversible watermarking strategy, $P_{max} \leq N$. When an actual payload size $P (P \leq P_{max})$ is given, the embedding threshold T can be computed. For recovering the cover image, the information of P and T is needed to be sent to the decoder in a way that the LSB values of the first 32 prediction errors are kept (as part of the payload) and then replaced by the parameters P (25 bits) and T (7 bits).

Referring to Sections 2 and 3, the embedding process of the proposed scheme is described as follows:

1. Referring to Section 3.4, predict the cover image Y to get the prediction errors E ;
2. Find the bad pixels in position by using an embedding testing operation. The embedding testing step is similar to that in [14, 15]. Each bad pixel consumes 25 bits of payload to indicate the embedding position;
3. Referring to Section 2.2, embed the data (including P , T , and the bad pixels in position) into E to generate E_w ;
4. Reconstruct the marked image Y_w from E_w by using the inverse prediction operation as described in Section 3.4, step 5).

In the decoder, the same prediction operations are performed on Y_w to get E_w . Then, the information of P and T is extracted from the LSB values of the first 32 prediction errors. Furthermore, the hidden data and the

original prediction errors E are extracted from E_w . Finally, the original image Y is completely recovered from E by using the inverse prediction operations.

6 Experimental testing and analysis

In this section, we adopt 24 gray-level versions of Kodak test images (<http://r0k.us/graphics/kodak/index.html>) and four standard benchmark images (*baboon*, *barbara*, *f16*, and *lena* in Fig. 4) as data set. Firstly, the performance of the noncausal predictor proposed in Section 3.4 is tested by comparing with other several typical predictors. This is followed by a performance comparison of the proposed watermarking scheme against three existing state-of-the-art works [14, 15, 18]. All the algorithms were implemented in Matlab, and the experiments were performed by embedding and decoding randomly generated binary bitstreams on image data set for reversible watermarking.

6.1 Comparison of typical predictors

The shape of the histogram of the predicted errors is often used to measure the performance of the embedding scheme. In general, distribution of the prediction errors obeys a Laplacian distribution. The shape of the distribution is determined by the absolute mean and variance. If the mean is close to zero, the variance essentially determines the shape of the histogram. The smaller variance value, the better performance can be achieved for the reversible watermarking scheme.

In the literature, the predictor proposed in [15] provides a satisfactory performance by dividing an image into two sets (like a chess board divided into black and white sets) to achieve noncausal prediction in a way that a pixel can be predicted with its four immediate pixels. This prediction method is suffering from a distortion problem. That

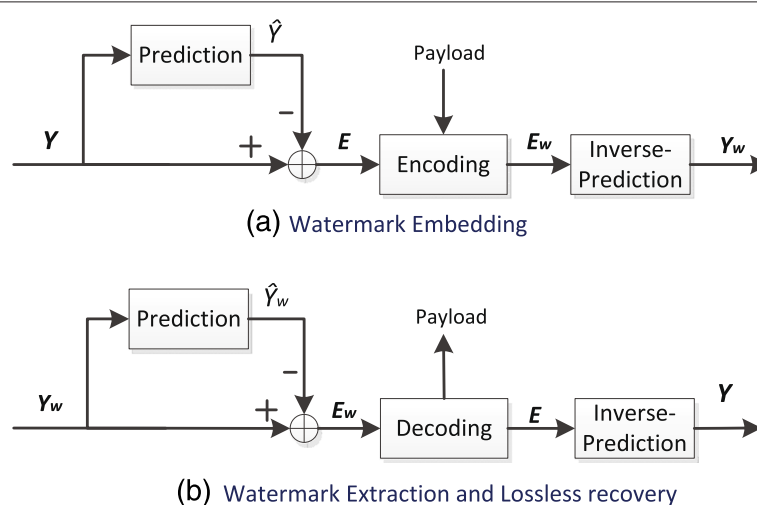
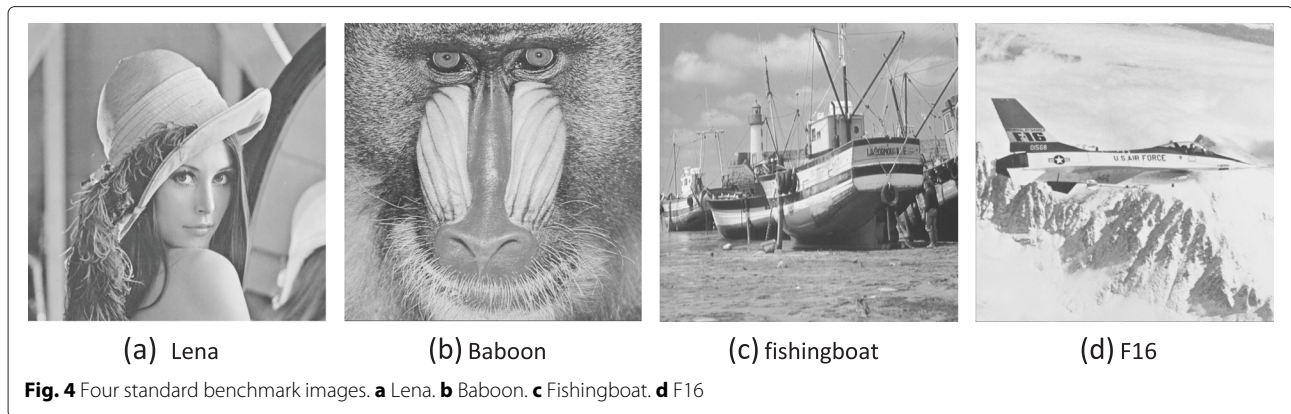


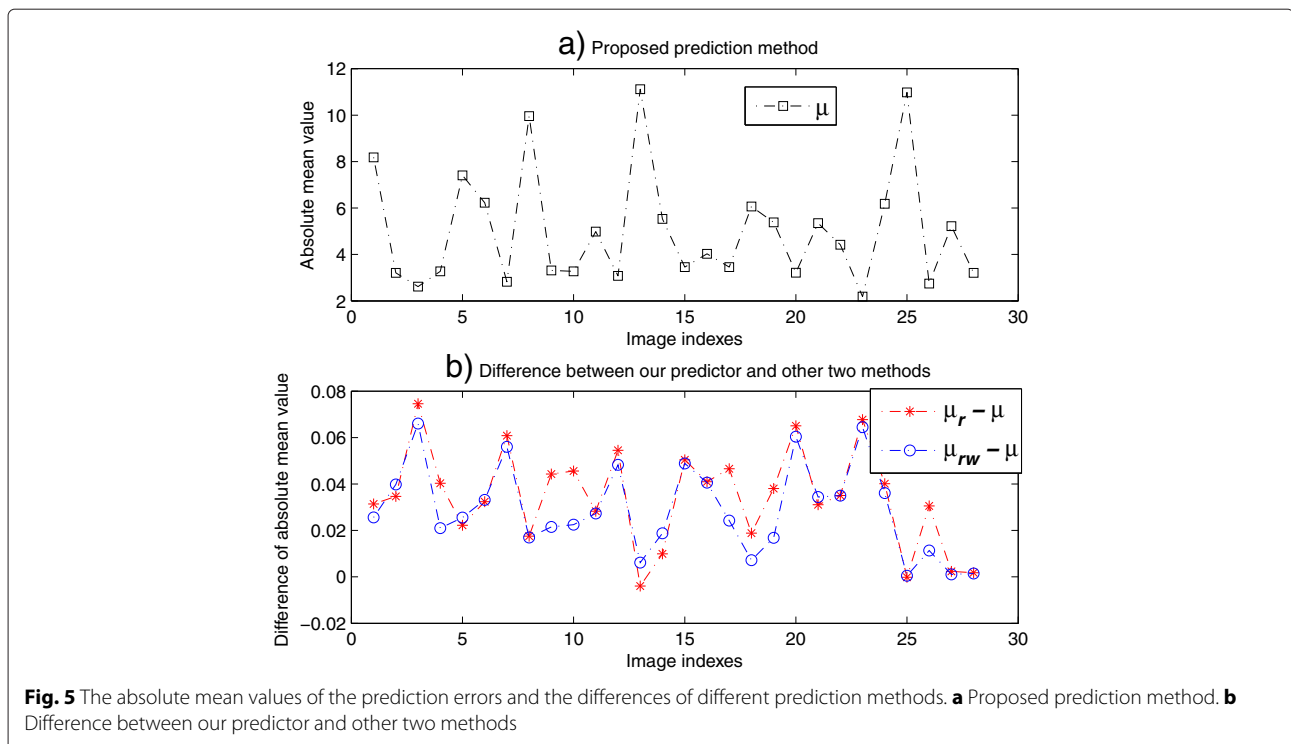
Fig. 3 Proposed reversible watermarking scheme. **a** Watermark embedding. **b** Watermark extraction and lossless recovery

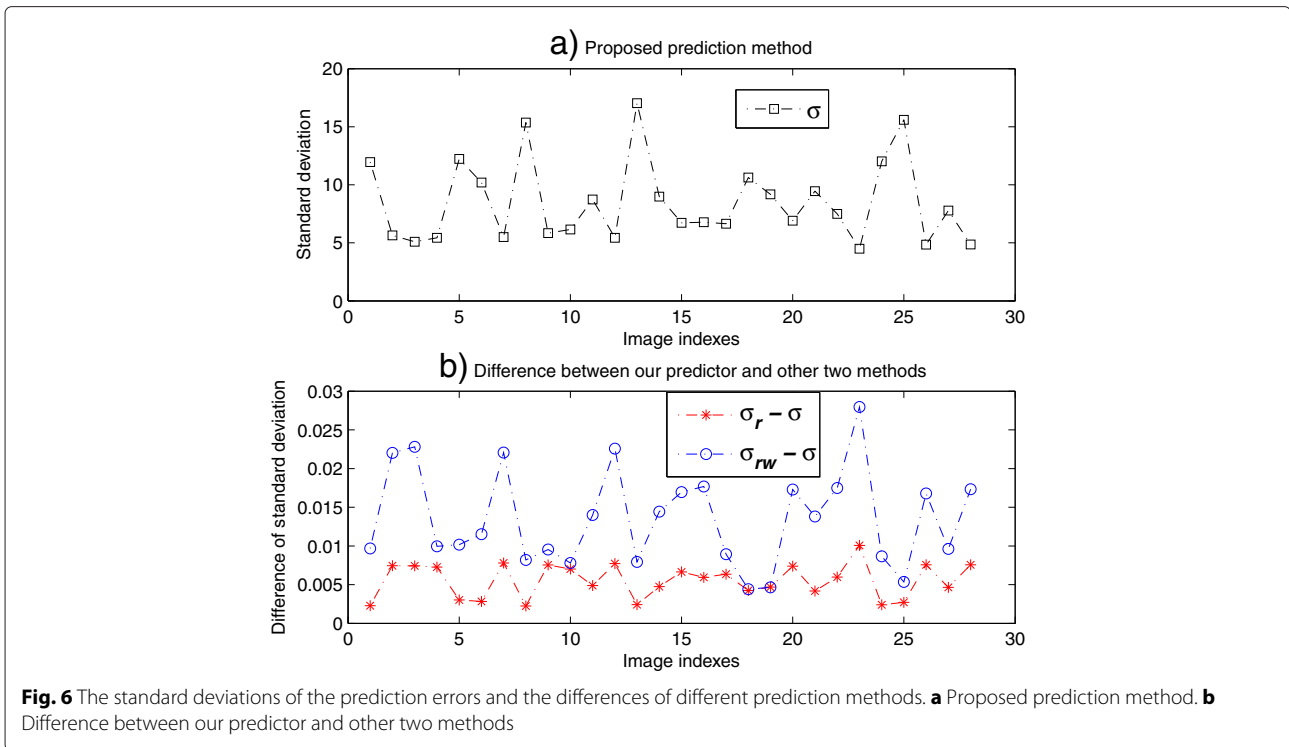


is, half of the pixels were predicted with modified pixels instead of the original ones. As shown in Section 3.4, the proposed prediction method can predict a pixel with its four immediate pixels, and all the pixels can be predicted completely with original pixels as context. As a result, the proposed prediction method has better prediction accuracy. In order to better evaluate the effect of the rounding and watermarking operations, we have computed the absolute mean values and the standard deviations of all the 28 test images which are predicted with the proposed predictor in this paper (denoted by μ and σ), the proposed predictor with an integer rounding operation (μ_r and σ_r) and the prediction method in [15] (μ_{rw} and σ_{rw}).

Figure 5 shows the absolute mean values of all the 28 test images which are predicted with the proposed

method in this paper and the differences among these three predictors. We can see from Fig. 5a that the mean values are close to zero. In Fig. 5b, the differences ($\mu_r - \mu$) plotted with the “asterisk” line are often positive, indicating the effect of the rounding operation on the predicted values. The “circle” line plots the differences ($\mu_{rw} - \mu$) that are also positive, indicating the effect of prediction with the modified pixels on the mean values. Figure 6 shows the corresponding standard deviations (σ) and their differences, denoted by $\sigma_r - \sigma$ and $\sigma_{rw} - \sigma$, respectively. We can see from the “asterisk” line in Fig. 6b that the difference values are always positive, indicating effect of the rounding operation and prediction with modified pixels on the standard deviations. Figures 5 and 6 show that the prediction errors with the proposed predictor





has smaller mean value and the standard variances for an image.

Furthermore, Table 1 lists the absolute mean values and the standard deviations of four benchmark images predicted with five different predictors (difference predictor used in [9, 10], integer MED predictor used in [14, 18], non-integer MED predictor [33], the prediction method in [15], and the proposed one in this paper). We can see from this table that the prediction method proposed in this paper can provide the smallest mean values and standard deviations.

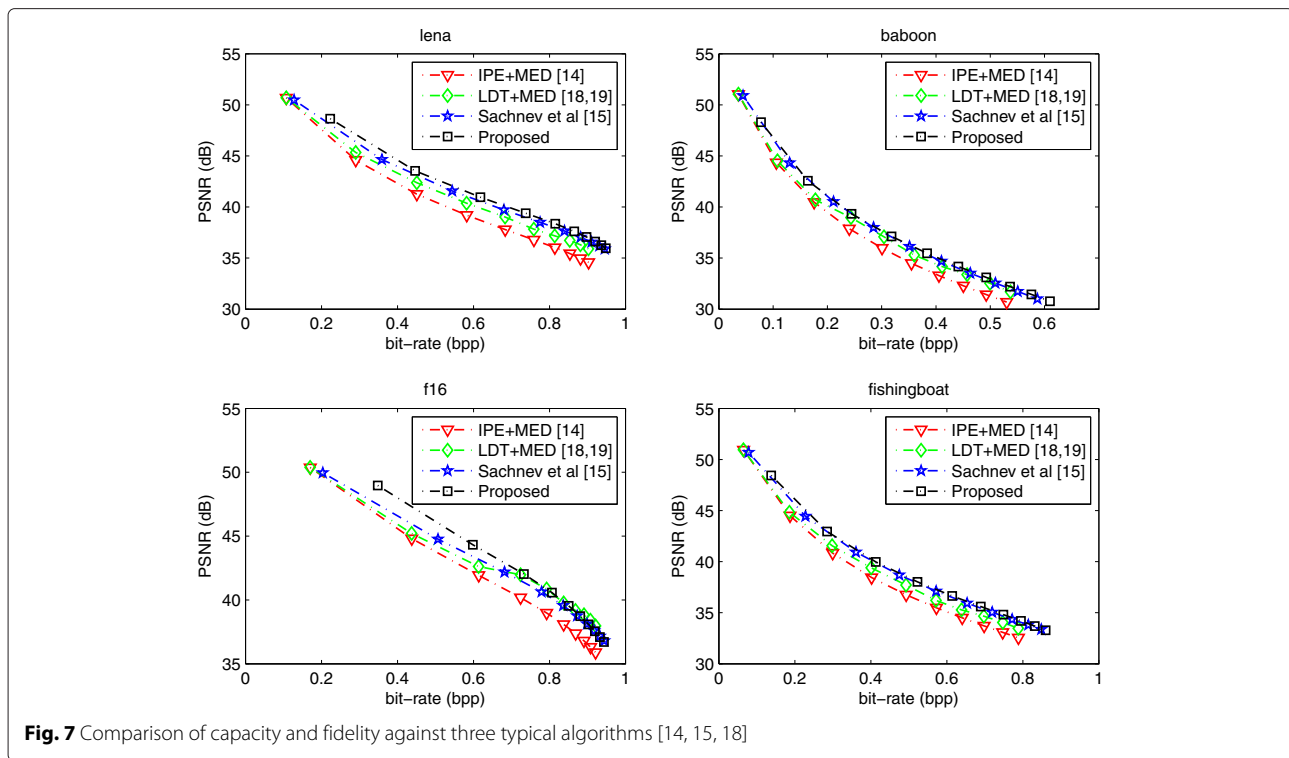
6.2 Comparison with other recent algorithms

We implemented three typical algorithms: Thodi and Rodriguez’s IPE algorithm with histogram shifting and flag bits (P3) [14], Coltuc’s low distortion transform method on MED predictor [18], and the IPE-based double-embedding scheme [15] and the proposed scheme

with the NIPE technique and the new prediction method. The scheme, proposed by Thodi and Rodriguez, is the classical IPE embedding strategy which includes a MED predictor to output integer prediction errors for expansion embedding. Reversible watermarking using low distortion transforms the same MED predictor proposed by Coltuc for the reduction of the embedding distortion by marking not only the current pixel but also its context [18]. In the literature, the double-embedding scheme proposed by Sachnev et al. has provided a satisfactory performance by dividing an image into two sets in a way that the pixels can be predicted in a noncausal way. Four standard benchmark images are adopted to report experimental results, as plotted in Fig. 7. Simulation results are similar for the other test images. We can see from Fig. 7 that the NIPE technique with the proposed noncausal prediction method performs better at all embedding rates. The detail is described as follows.

Table 1 Performance comparison of five predictors

Predictors	Lena		Baboon		Fishingboat		F16	
	μ	σ	μ	σ	μ	σ	μ	σ
Difference predictor [9, 10]	4.6724	7.8454	18.4130	27.6633	7.2441	11.1876	5.2971	11.9235
Integer MED predictor [14]	4.3407	6.9049	13.5220	19.6804	6.3836	9.3957	3.4273	6.2054
Non-integer MED predictor [33]	4.3060	6.7608	13.3652	19.2567	6.5863	9.7152	3.6089	6.6339
Sachnev’s predictor [15]	3.2063	4.8738	10.9786	15.6010	5.2220	7.7943	2.7578	4.8683
Proposed predictor	3.2049	4.8565	10.9782	15.5957	5.2209	7.7847	2.7465	4.8515



The classical IPE embedding scheme, proposed by Thodi and Rodriguez [14], is a high-capacity reversible watermarking algorithm by developing PE embedding technique with MED predictor. We can see from Fig. 7 that the NIPE technique with the proposed noncausal predictor can provide higher embedding payload in comparison with the IPE scheme. The basic reason is that the proposed noncausal predictor has higher prediction precision than the MED predictor used in [14].

Coltuc [18, 19] has also developed Thodi and Rodriguez's work and has presented the results for the MED predictor. The basic idea of the approach is to embed the entire expanded difference not only into the current pixel but also its context. Then, the minimization of the square errors is considered to reduce the embedding distortion. When the parameter α is 0.25 (referred to [18], Eq. (4)), we can see from Fig. 7 that Coltuc's embedding method has lower embedding distortion than the IPE for the same embedding rate. In [19], the embedding approach has been further generalized as a low distortion transform (LDT) for reversible watermarking. Comparing with the LDT technique with the MED predictor in [18], the proposed NIPE scheme can provide higher embedding payload or capacity for the same embedding distortion. The basic reason is that the proposed prediction strategy can better estimate the current pixel by incorporating data prediction not restricted to only causal pixels, as listed in Table 1.

Another important improvement on Thodi and Rodriguez's work has been proposed by Sachnev et al. [15] by introducing a high-precision prediction strategy and sorting technique. Since a pixel can be estimated with its four immediate pixels, the IPE-based scheme can provide satisfactory performance in the literature. Figure 7 shows that the reversible watermarking approach proposed in this paper has lower embedding distortion at all embedding rates than Sachnev et al.'s double embedding scheme. The reason is that the predictor used for NIPE can predict pixels with four original immediate pixels when one used in [15] predicted half of pixels with modified pixels as context.

7 Conclusions

This paper presents a NIPE embedding technique for reversible watermarking. The NIPE technique can remedy a major drawback of Thodi and Rodriguez's work (called IPE in this paper) that the predicted values should be rounded to integer number for data embedding. With the NIPE technique, the rounding operation in the prediction process (that often appears in IPE-based reversible watermarking algorithms to generate integer prediction errors) can be discarded. This is beneficial to use better predictor. In order to prove the advantage, we proposed a prediction model and designed an image predictor for the NIPE. The new predictor can predict pixels with four immediate pixels, and all pixels can be predicted with the original

pixels. With the proposed NIPE and predictor, the embedding distortion is smaller than that in [15] at all embedding rates. Experimental results have shown that the predictor designed in this paper can provide the best performance than several existing typical prediction methods. In comparison with other typical reversible watermarking algorithms, the proposed scheme (combining the NIPE technique and new prediction method) performs better.

Endnotes

¹In the IPE, the prediction errors should be rounded to integer value. This rounding operation brings a constraint on a predictor's performance.

²In the literature [15, 26], noncausal predictive methods have been used for the IPE by using multi-passes prediction for multi-layers embedding. Since part of the pixels were predicted by using the watermarked pixels instead of the original ones, some distortion has been introduced.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was supported by NSFC (No. 61272414). The authors would like to thank the reviewers' valuable comments.

Received: 11 October 2014 Accepted: 16 May 2015

Published online: 08 July 2015

References

- W Bender, D Gruhl, N Morimoto, A Lu, Techniques for data hiding. *IBM Syst. J.* **35**(3), 313–336 (1996)
- B Macq, in *Proc. the European Signal Processing Conf., Lossless multiresolution transform for image authenticating watermarking* (Tampere, Finland, 2000)
- C De Vleeschouwer, JE Delaigle, B Macq, Circular interpretation of bijective transformations in lossless watermarking for media asset management. *IEEE Trans. Multimedia.* **5**(1), 97–105 (2003)
- S Lee, CD Yoo, T Kalker, Reversible image watermarking based on integer-to-integer wavelet transform. *IEEE Trans. Inf. Forensics Secur.* **2**(3), 321–330 (2010)
- J Fridrich, M Goljan, R Du, in *Proc. SPIE Photonics West, Security and Watermarking of Multimedia Contents III*. Invertible authentication, vol. 3971 (San Jose, 2001), pp. 197–208
- J Fridrich, M Goljan, R Du, Lossless data embedding paradigm in digital watermarking. *EURASIP J. Appl. Signal Process.* **2**, 185–196 (2002)
- AACM Kalker, FMJ Willems, in *Proc. 14th Int. Conf. Digital Signal Processing*. Capacity bounds and constructions for reversible data-hiding, vol. 1 (Santorini, 2002), pp. 71–76
- MU Celik, G Sharma, AM Tekalp, E Saber, Lossless generalized-LSB data embedding. *IEEE Trans. Image Process.* **14**(2), 253–266 (2005)
- J Tian, Reversible data embedding using a difference expansion. *IEEE Trans. Circuits Syst. Video Technol.* **13**(8), 890–896 (2003)
- AM Alattar, Reversible watermark using the difference expansion of a generalized integer transform. *IEEE Trans. Image Process.* **13**(8), 1147–1156 (2004)
- L Kamstra, H Heijmans, Reversible data embedding into images using wavelet techniques and sorting. *IEEE Trans. Image Process.* **14**(12), 2082–2090 (2005)
- X Wang, C Shao, X Xu, X Niu, Reversible data-hiding scheme for 2-D vector maps based on difference expansion. *IEEE Trans. Inf. Forensics Secur.* **2**, 311–319 (2007)
- Y Hu, H-K Lee, J Li, DE-based reversible data hiding with improved overflow location map. *IEEE Trans. Circuits Syst. Video Technol.* **19**(2), 250–260 (2009)
- DM Thodi, JJ Rodriguez, Expansion embedding techniques for reversible watermarking. *IEEE Trans. Image Process.* **15**, 721–729 (2007)
- V Sachnev, HJ Kim, J Nam, S Suresh, Y Shi, Reversible watermarking algorithm using sorting and prediction. *IEEE Trans. Circuits Syst. Video Technol.* **19**(7), 989–999 (2009)
- H-W Tseng, C-P Hsieh, Prediction-based reversible data hiding. *Inf. Sci.* **179**, 246–2469 (2009)
- X Li, B Yang, T Zeng, Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection. *IEEE Trans. Image Process.* **20**(12), 3524–3533 (2011)
- D Coltuc, Improved embedding for prediction-based reversible watermarking. *IEEE Trans. Inf. Forensics Secur.* **6**(3), 873–882 (2011)
- D Coltuc, Low distortion transform for reversible watermarking. *IEEE Trans. Image Process.* **21**(1), 412–417 (2012)
- Veen van der M, F Bruekers, A van Leest, S Cavin, in *Proc. SPIE Photonics West, Electronic Imaging 2003, Security and Watermarking of Multimedia Contents V*. High-capacity reversible watermarking for audio, vol. 5020 (San Jose, California, 2003), pp. 1–11
- B Bradley, AM Alattar, in *Proc. SPIE Photonics West, Electronic Imaging 2005, Security and Watermarking of Multimedia Contents VII*. High-capacity, invertible, data-hiding algorithm for digital audio, vol. 5681 (San Jose, California, 2005), pp. 789–800
- D Yan, R Wang, in *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. Reversible data hiding for audio based on prediction error expansion (Harbin, 2008), pp. 249–252
- A Van Leest, M Van der Veen, F Bruekers, in *Proc. IEEE Conf. Image Processing*. Reversible image watermarking, vol. 3 (Barcelona, 2003), pp. 731–734
- Z Ni, Y Shi, N Ansari, S Wei, in *Proc. IEEE Int. Symp. Circuits and Systems*. Reversible data hiding, vol. 2 (Bangkok, 2003), pp. 912–915
- M Chen, Z Chen, X Zeng, Z Xiong, in *Proc. 11th ACM Workshop Multimedia and Security*. Reversible data hiding using additive prediction-error expansion (Princeton, 2009), pp. 19–24
- L Luo, Z Chen, M Chen, X Zeng, Z Xiong, Reversible image watermarking using interpolation technique. *IEEE Trans. Inf. Forensics Security.* **5**(1), 187–193 (2010)
- Anil K Jain, *Fundamentals of digital image processing*. (Prentice, Hall, Englewood Cliffs, NJ, 1989)
- A Asif, JMF Moura, Image codec by noncausal prediction, residual mean removal, and cascaded vector quantization. *IEEE Trans. Circuits Syst. Video Technol.* **6**(1), 42–55 (1996)
- N Balram, JMF Moura, Noncausal predictive image codec. *IEEE Trans. Image Process.* **5**(8), 1229–1242 (1996)
- M Weinberger, G Seroussi, Sapiro G, The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS. *IEEE Trans. Image Process.* **9**(8), 1309–1324 (2000)
- WR Gardner, BD Rao, Non-causal linear prediction of voiced speech. *IEEE Asilomar Conference on Signals, Systems and Computers*, (Pacific Grove, CA, Oct. 1992), pp. 1100–1104
- Wu X, Memon N, Context-based, adaptive, lossless image coding. *IEEE Trans. Commun.* **45**(4), 437–444 (1997)
- S Martucci, in *Proc. IEEE Int. Symp. Circuits and Systems*. Reversible compression of HDTV images using median adaptive prediction and arithmetic coding (New Orleans, 1990), pp. 1310–1313
- S Xiang, in *IH 2012, LNCS 7692*. Non-integer expansion embedding for prediction-based reversible watermarking (Springer Berkeley, 2013), pp. 224–239