

Reversible Data Hiding Based on Combined Predictor and Prediction Error Expansion

Xiaochao Qu¹, Suah Kim¹, Run Cui¹, Fangjun Huang²,
and Hyoung Joong Kim¹ (✉)

¹ Graduate School of Information Security and Management,
Korea University, Seoul 136-701, Korea

{quxiaochao,suahnkim}@gmail.com, {cuirun,khj-}@korea.ac.kr

² School of Information Science and Technology, Sun Yat-Sen University,
Guangdong 510-006, China
huangfj@mail.sysu.edu.cn

Abstract. This paper presents a novel reversible data hiding method which uses a combined predictor. The proposed combined predictor combines five base predictors according to their global and local predicting performance. The weights to combine the base predictors are calculated with a pixel by pixel manner that they adjust to the local image patch characteristics. The proposed predictor is shown to have high prediction precision which is beneficial for the following prediction error expansion (PEE). Observing that our predictor performs well even for images with complex textures, a novel pixel selection criterion that bases on the prediction errors is proposed, which can accurately select the pixels that have small prediction errors to use. Extensive experiments are conducted to verify the superior performance of the proposed method.

Keywords: Reversible data hiding · Combined predictor · Pixel selection · Prediction error expansion

1 Introduction

Reversible data hiding is a technique that can embed data into host image and the data can be extracted from the marked image. After the extraction of data, the host image can be recovered perfectly. The ability of recovering the original host image is required in applications such as military and medical image processing where any distortion can not be tolerated.

Many reversible data hiding methods have been proposed that can be categorized as lossless compression based [5], integer transform based [15, 22], histogram shifting based [12, 13, 19] and difference expansion based [21]. Among the above mentioned methods, difference expansion based method is the most popular one due to its high embedding capacity and low embedding distortion.

Difference expansion (DE) was first proposed by Tian [21] which expands the difference value between a pair of pixels to embed 1 bit data. Later on,

DE was improved in various different ways. The first way is to generalize DE where $n - 1$ bits can be embedded into a n pixel vector [1]. The second way is to reduce the size of the location map. Kamastra [8] proposed sorting to make the location map more compressible. Kim [9] proposed a novel difference expansion transform with a simplified location map. The location map in their method has a better compression ratio thus the true capacity of their method is increased. Hu [7] constructs a payload-dependent overflow location map which reduces unnecessary alteration to image and has a better compressibility. Thodi [20] uses histogram shifting to separate the embedded pixel and the un-embedded pixel such that the location map size is significantly reduced. The third way is to use prediction error expansion (PEE) proposed by Thodi [20] instead of DE. PEE has two advantages over DE: larger embedding capacity and smaller distortion. Predictors used in PEE can better utilize the correlation of pixels and the obtained prediction error is usually smaller compared with pixel difference. Many predictors are proposed such as rhombus predictor [18], orthogonal projection [6], gradient adjusted predictor (GAP) [11], pixel value ordering based predictor [10, 16], partial differential equation based predictor [14], edge based difference expansion [17] and local predictor [4]. Recently, a novel improvement of DE called context embedding was proposed [2, 3]. Context embedding distributes the embedding distortion to the context pixels which reduces the overall embedding distortion.

In this paper, we propose a high performance reversible data hiding method using a novel combined predictor and a new pixel selection criterion. Five base predictors are linearly combined into a final predictor. The weights for combining those five base predictors are determined by the base predictor's global performance and local performance. The base predictor is texture selective that it performs well when certain texture appears. By adjusting the weight pixel by pixel, the proposed combined predictor performs well with different textures, including very complex textures. To realize the full potential of the combined predictor, a novel pixel selection criterion is used. The new criterion is based on the characteristics of the combined predictor's prediction errors. It selects the image regions with small prediction errors to use such that the texture regions can be utilized. In the experiment, the proposed method is compared with the methods in [11, 18], and it is shown that our method has the best performance.

The rest of the paper is organized as follows. The reversible data hiding based on PEE is introduced in Sect. 2. Section 3 describes the proposed combined predictor, the new pixel selection method and the way to avoid overflow and underflow problems. The experiment result is shown in Sect. 4. Finally, Sect. 5 is the conclusion.

2 Reversible Data Hiding Based on Prediction Error Expansion

Without loss of generality, the predictor in [18] is used in this section (Fig. 1). Assume the current pixel is x and its four context pixels are x_1 , x_2 , x_3 and x_4 .

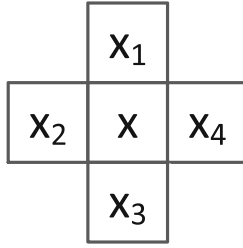


Fig. 1. The four context pixels of the pixel x .

The prediction of x is calculated as:

$$\hat{x} = \left\lfloor \frac{(x_1 + x_2 + x_3 + x_4)}{4} + 0.5 \right\rfloor \tag{1}$$

where the $\lfloor \cdot \rfloor$ is the floor operation which rounds down the value to the next lowest integer value. The prediction error is calculated as:

$$e = x - \hat{x} \tag{2}$$

The prediction error can be expanded to embed a bit $b \in \{0, 1\}$ as:

$$\hat{e} = 2 \times e + b \tag{3}$$

The marked pixel is obtained after the embedding as:

$$X = \hat{x} + \hat{e} \tag{4}$$

To control the embedding capacity and distortion, a pair of threshold $T_p \in [0, +\infty)$ and $T_n \in (-\infty, 0)$ determines which pixel to use. The pixels where $e > T_p$ or $e < T_n$ are shifted as:

$$X = \begin{cases} x + T_p + 1, & \text{if } e > T_p \\ x + T_n, & \text{if } e < T_n. \end{cases} \tag{5}$$

At data extraction, the same prediction \hat{x} is calculated and \hat{e} is obtained as $X - \hat{x}$. If $\hat{e} \in [2 \times T_n, 2 \times T_p + 1]$, the extracted data can be calculated as:

$$b = \hat{e} - 2 \times \left\lfloor \frac{\hat{e}}{2} \right\rfloor \tag{6}$$

The original pixel value x can be recovered as:

$$x = \hat{x} + \left\lfloor \frac{\hat{e}}{2} \right\rfloor \tag{7}$$

If $\hat{e} \in (-\infty, 2 \times T_n) \cup (2 \times T_p + 1, \infty)$, no data can be extracted and the original pixel value is recovered as:

$$x = \begin{cases} X - T_p - 1, & \text{if } \hat{e} > 2 \times T_p + 1 \\ X - T_n, & \text{if } \hat{e} < 2 \times T_n. \end{cases} \tag{8}$$

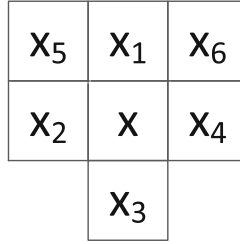


Fig. 2. The proposed six context pixels of the pixel x .

3 Proposed Method

3.1 Combined Predictor

The proposed predictor uses the nearest six context pixels to predict the center pixel x as shown in Fig. 2. x_5 and x_6 can not be used in the predictor of [18] due to the sorting operation. Pixel selection is used instead of sorting in our method which allows x_5 and x_6 to be included into the context pixels. When extracting data from the current pixel x , all its context pixels have been recovered which guarantees the predictions of the x in the extracting process and embedding process have the same value.

Using those context pixels, five base predictors can be defined as:

$$\begin{aligned}
 p_1 &= \frac{x_1+x_2+x_3+x_4}{4} \\
 p_2 &= \frac{x_2+x_4}{2} \\
 p_3 &= \frac{x_1+x_3}{2} \\
 p_4 &= x_5 \\
 p_5 &= x_6
 \end{aligned} \tag{9}$$

The performance of these five base predictors are different for different kinds of textures. For example, p_2 performs well when encountering an horizontal edge in the image. Therefore, it is not optimal to select one of these five base predictors to use for the whole image since there are different textures in different regions of the image. It is better to assign weights to these five base predictors and linearly combine them. Higher weights are assigned to the predictors that are proper for the current image texture.

For determining the weights, the global and the local performance of the base predictors are considered together. The global performance of the base predictors are measured by the prediction errors of the whole host image. Assume the accumulated absolute prediction error of the base predictor p_i for the host image I is E_i^{global} (the summation of all the absolute prediction errors for a host image), the global weight assigned to base predictor p_i is calculated as:

$$w_i^{global} = \frac{\frac{1}{E_i^{global}}}{\sum_{i=1}^5 \frac{1}{E_i^{global}}} \tag{10}$$

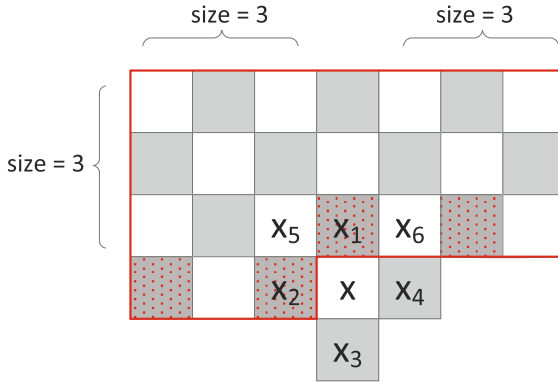


Fig. 3. The local region with red border is used to evaluate the predictor’s local performance. The pixels with red dots are not included due to incomplete context pixels.

Clearly, the base predictor with small global prediction error will be assigned with large global weight. However, the global weight itself is not enough to capture the local texture characteristics. Image texture varies from region to region, therefore, local performance of each base predictor should also be considered. The local performance is measured by the prediction error of this base predictor in a small local region. Figure 3 shows an example of the small region with size equals to three. Notice that the pixels with red dots can not be used due to incomplete context pixels, which guarantees that the embedding process and the extracting process has the same pixel information. The current pixel x is white pixel so that all gray pixels are available, and all white pixels that above x are also available. x_1 and x_2 hve incomplete context pixel because of x is not available. The other two red dot pixels have incomplete context pixel because of the lower vertical pixels are not available Assume the accumulated absolute prediction error in the local region is E_i^{local} for the base predictor p_i , the local weight assigned to base predictor p_i is calculated as:

$$w_i^{local} = \frac{\frac{1}{E_i^{local}}}{\sum_{i=1}^5 \frac{1}{E_i^{local}}} \tag{11}$$

Notice that if the pixel in the local region does not have full context available, it will be ignored when computing the accumulated absolute error. The base predictor with small local prediction errors will be assigned with large local weight. With the global weight and local weight at hand, the next step is to combine these two weights into the final weights as:

$$w_i = \frac{w_i^{local} \times w_i^{global}}{\sum_{i=1}^5 w_i^{local} \times w_i^{global}} \tag{12}$$

The rationale of the proposed predictor can also be explained with Bayes’ theorem. The weight reflects the probability of selecting each base predictor.

The global weight can be regarded as the prior of each base predictor and the local weight is the likelihood of each base predictor. If there is no local region pixels available, the prior (global weight) will be used directly. By observing the local region prediction errors, the likelihood is estimated and it transforms the prior into the posteriori.

The proposed combined predictor can be calculated as:

$$p = \left[\sum_{i=1}^5 w_i \times p_i + 0.5 \right] \quad (13)$$

3.2 Pixel Selection Using Prediction Errors

Pixel selection separates the host image into two parts: usable part and unusable part. The prediction errors in the usable part tends to be smaller than that in the unusable part. For each pixel, a pixel selection measurement can be calculated. Then, a proper pixel selection threshold T is determined given a certain payload. The pixels with the pixel selection measurement smaller than T are used to embed data, and all other pixels are skipped.

Conventionally, pixel selection uses neighboring pixel value variance or difference as the metric. The assumption behind is that predictor performs well in smooth regions and those regions should be embedded with priority. However, with the proposed combined predictor, this assumption is not true any more. The proposed predictor adaptively assigns weights to the five base predictors such that it performs well in complex texture regions (e.g., strong edges).

In order to use the texture region as well (which is skipped when using pixel value variance), a new pixel selection metric is proposed. The new metric considers the prediction error values instead of the pixel values. The idea is that the prediction error in the small local region near the current pixel can reflect the performance of the combined predictor to some extent. We can assume that the performance of the combined predictor is similar for the current pixel and its neighbor pixels. Hence, the prediction errors in the small local region can be used to decide whether to use the current pixel to embed data. Assume the absolute prediction errors in the small local region are grouped into one vector denoted as PE, the pixel selection metric is calculated as $\max(\text{PE}) - \min(\text{PE})$, where $\max()$ and $\min()$ compute the maximum and minimum value of PE, respectively. The advantage of the new metric is that it predicts the predictor's performance using that predictor's historic predicting performance, which can better utilize those pixels in texture but easy to predict region.

3.3 Overflow and Underflow Prevention and Side Information Construction

In the embedding process, some modifications to original pixels may occur overflow and underflow problems that the modified pixel value may be out of the intensity level, for example, $[0, 255]$ for eight-bit grayscale image. Usually,

a binary location map with the same size of the host image is needed to record such locations and transmitted to the decoder side. Due to the large size of the location map, compression is needed to reduce the location map's size. The usage of compression makes the algorithm complex and the determination of parameters for the embedding algorithm is hard. We use a strategy to construct a location map that compression is avoided as in [11].

Before embedding, all pixels should be checked if an overflow or underflow problem may occur, if it is possible to have such problems, then the location of this pixel needs to be recoded and this pixel is not used to embed data. For a $W \times H$ grayscale image, each location needs $\left\lceil \log_2(W \times H) \right\rceil + 1$ bits. Besides the location map, some other side information should also be recorded:

1. Compressed location map size ($\left\lceil \log_2(W \times H) \right\rceil + 1$).
2. Global weights for five base predictors (16 bits for each weight).
3. Local region size (8 bits).
4. Embedding threshold T (8 bits).
5. Pixel selection threshold T_{ps} (16 bits).

All the side information concatenated with the location map is embedded using least significant bit (LSB) substitution as in [11].

4 Experiment

The test images used in our experiment have the size of 512×512 , namely *Lena*, *F16*, *Baboon*, *Barbara*, *Sailboat* and *Elaine* which can be downloaded from SIPI image database (sipi.usc.edu/database/) except *Barbara*. Capacity-distortion performance is used to evaluate the performance of reversible data hiding method. The embedding capacity is measured with bits per pixel (bpp) and distortion is measured with peak signal-to-noise ratio (PSNR).

The performance of the PEE based reversible data hiding is highly dependent on the predictor's performance. An accurate predictor produces small prediction errors which can achieve large embedding capacity with low distortion. The performance of the predictor can be measured by the entropy of the prediction errors which is calculated as:

$$entropy = - \sum_{i=1}^N p_i \log_2(p_i) \quad (14)$$

where p_i is the probability of the i th prediction error and N is the total number of possible values of the prediction error. A small entropy means that the prediction errors are more concentrated which is beneficial for the following expansion operation. The entropy comparison of the proposed combined predictor with other predictors is shown in Table. 1. The window size is set to 3 since we found that the window size of 3 has the best trade-off of prediction performance and computation complexity. Henceforth, all the window size used is 3.

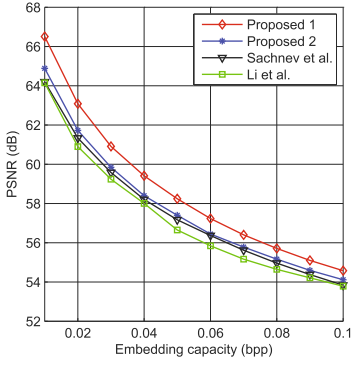
Table 1. Comparisons in terms of entropy for different predictors.

Image	MED [20]	GAP [11]	Rhombus [18]	Combined predictor
Lena	4.55	4.39	4.10	3.99
F16	4.18	4.12	3.86	3.80
Baboon	6.27	6.21	5.96	5.82
Barbara	5.48	5.38	5.14	4.75
Sailboat	5.38	5.25	4.97	4.92
Elaine	5.34	5.15	4.89	4.77
Average	5.20	5.08	4.82	4.67

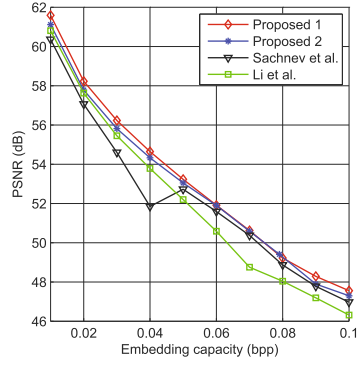
To show the capacity-distortion performance of the proposed reversible data hiding method, we conduct two separate experiments for small size payload and large size payload. The experiment for small size payload can also show the performance of the proposed new pixel selection method.

Pixel selection is more important with small size payload than with large size payload. With large size payload, pixel selection has to select most pixels to use. While with small size payload, the pixel selection has more freedom to decide which pixel to use. The proposed pixel selection based on prediction error distribution is compared with pixel selection based on pixel variance as in [18]. The capacity-distortion performance comparison with small size payload is shown in Fig. 4. The embedding capacity tested is from 0.01 bpp to 0.1 bpp. *Proposed 1* utilizes the combined predictor and the pixel selection based on prediction error distribution. *Proposed 2* utilizes the combined predictor and the pixel selection based on pixel variance. The only difference of *Proposed 1* and *Proposed 2* is the pixel selection method, hence, the performance difference of these two methods can reflect the effectiveness of the respective pixel selection method. As can be seen, *Proposed 1* performs better than *Proposed 2* for all the test images. Compared with other two methods in [11] and [18], *Proposed 1* and *Proposed 2* both performs better.

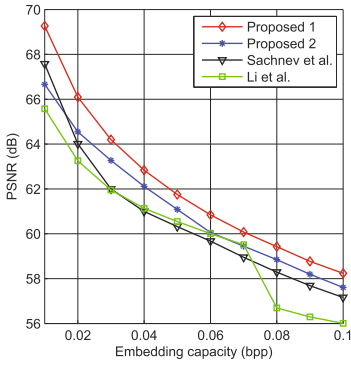
With large embedding capacity, we compare the proposed method (combined predictor + pixel selection based on prediction error) with [11] and [18] given the payload size from 0.1 bpp to the maximum embedding capacity. The proposed method outperforms [11] and [18] for all the test images. However, the performance gap between the proposed method with [11] and [18] is different from image to image. For the relatively smooth images, e.g., *F16*, the performance gain of the proposed method is not significant. The reason is that the advantage of the proposed combined predictor is less with smooth images. Given the images with complex texture patterns, e.g., *Barbara*, the proposed combined predictor performs well, therefore, the overall performance of the proposed reversible data hiding method is excellent (Fig. 5).



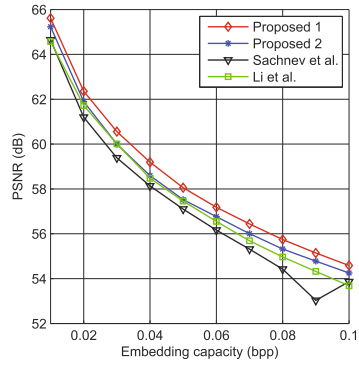
(a) Lena



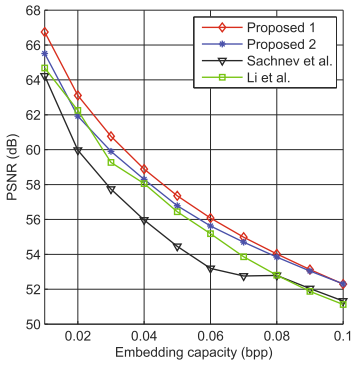
(b) Baboon



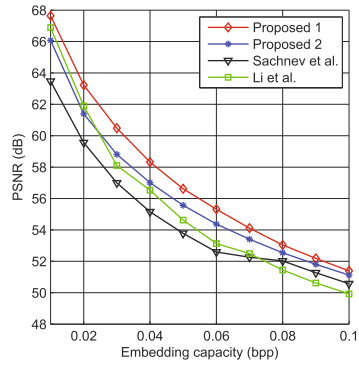
(c) F16



(d) Barbara

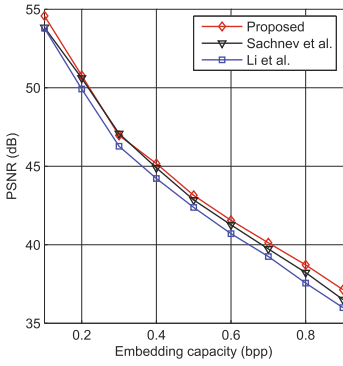


(e) Sailboat

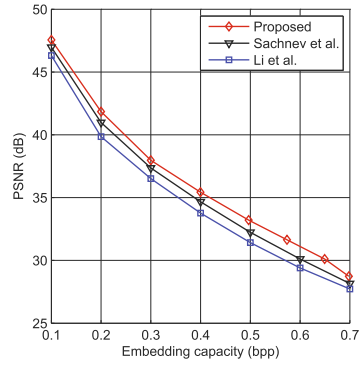


(f) Elaine

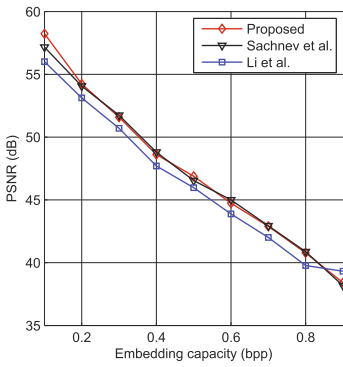
Fig. 4. The capacity-distortion performance of the proposed method compared with the methods of Li et al. [11] and Sachnev et al. [18] with low payload size.



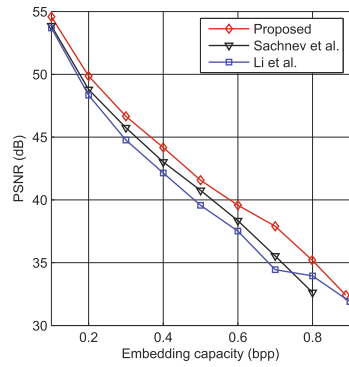
(a) Lena



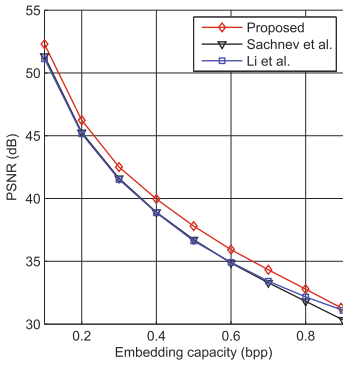
(b) Baboon



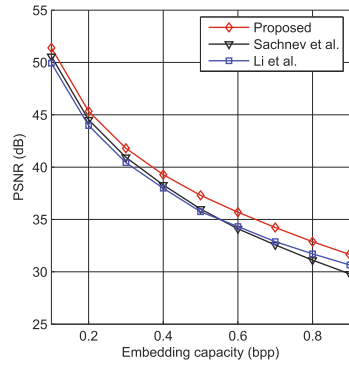
(c) F16



(d) Barbara



(e) Sailboat



(f) Elaine

Fig. 5. The capacity-distortion performance of the proposed method compared with the methods of Li et al. [11] and Sachnev et al. [18] with large payload size.

5 Conclusion

In this paper, a high performance combined predictor and a novel pixel selection method is proposed. It is shown that the proposed combined predictor performs well especially for complex textured images and the proposed pixel selection can better choose the proper pixel to use. Combining the combined predictor and the proposed pixel selection together, the proposed reversible data hiding method achieves excellent result.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012015587), and the National Natural Science Foundation of China (no. 61173147).

References

1. Alattar, A.M.: Reversible watermark using the difference expansion of a generalized integer transform. *IEEE Trans. Image Process* **13**(8), 1147–1156 (2004)
2. Coltuc, D.: Improved embedding for prediction-based reversible watermarking. *IEEE Trans. Inf. Forensics Secur.* **6**(3), 873–882 (2011)
3. Coltuc, D.: Low distortion transform for reversible watermarking. *IEEE Trans. Image Process* **21**(1), 412–417 (2012)
4. Dragoi, I., Coltuc, D.: Local prediction based difference expansion reversible watermarking. *IEEE Trans. Image Process* **23**(4), 1779–1790 (2014)
5. Fridrich, J., Goljan, M., Du, R.: Lossless data embedding new paradigm in digital watermarking. *EURASIP J. Adv. Sig. Process.* **2002**(2), 185–196 (1900)
6. Hong, W., Chen, T.S., Chang, Y.P., Shiu, C.W.: A high capacity reversible data hiding scheme using orthogonal projection and prediction error modification. *Sig. Process.* **90**(11), 2911–2922 (2010)
7. Hu, Y., Lee, H.K., Li, J.: De-based reversible data hiding with improved overflow location map. *IEEE Trans. Circ. Syst. Video Technol.* **19**(2), 250–260 (2009)
8. Kamstra, L., Heijmans, H.J.: Reversible data embedding into images using wavelet techniques and sorting. *IEEE Trans. Image Process* **14**(12), 2082–2090 (2005)
9. Kim, H.J., Sachnev, V., Shi, Y.Q., Nam, J., Choo, H.G.: A novel difference expansion transform for reversible data embedding. *IEEE Trans. Inf. Forensics Secur.* **3**(3), 456–465 (2008)
10. Li, X., Li, J., Li, B., Yang, B.: High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion. *Sig. process.* **93**(1), 198–205 (2013)
11. Li, X., Yang, B., Zeng, T.: Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection. *IEEE Trans. Image Process* **20**(12), 3524–3533 (2011)
12. Li, X., Zhang, W., Gui, X., Yang, B.: A novel reversible data hiding scheme based on two-dimensional difference-histogram modification. *IEEE Trans. Inf. Forensics Secur.* **8**(7), 1091–1100 (2013)
13. Ni, Z., Shi, Y.Q., Ansari, N., Su, W.: Reversible data hiding. *IEEE Trans. Circ. Syst. Video Technol.* **16**(3), 354–362 (2006)
14. Ou, B., Li, X., Zhao, Y., Ni, R.: Reversible data hiding based on pde predictor. *J. Syst. Softw.* **86**(10), 2700–2709 (2013)

15. Peng, F., Li, X., Yang, B.: Adaptive reversible data hiding scheme based on integer transform. *Sig. Process.* **92**(1), 54–62 (2012)
16. Peng, F., Li, X., Yang, B.: Improved pvo-based reversible data hiding. *Digit. Sig. Process.* **25**, 255–265 (2014)
17. Qu, X., Kim, S., Kim, H.: Reversible watermarking using edge based difference modification. In: Fifth International Conference on Graphic and Image Processing. pp. 90690Q–90690Q. International Society for Optics and Photonics (2014)
18. Sachnev, V., Kim, H.J., Nam, J., Suresh, S., Shi, Y.Q.: Reversible watermarking algorithm using sorting and prediction. *IEEE Trans. Circ. Syst. Video Technol.* **19**(7), 989–999 (2009)
19. Tai, W.L., Yeh, C.M., Chang, C.C.: Reversible data hiding based on histogram modification of pixel differences. *IEEE Trans. Circ. Syst. Video Technol.* **19**(6), 906–910 (2009)
20. Thodi, D.M., Rodríguez, J.J.: Expansion embedding techniques for reversible watermarking. *IEEE Trans. Image Process.* **16**(3), 721–730 (2007)
21. Tian, J.: Reversible data embedding using a difference expansion. *IEEE Trans. Circ. Syst. Video Technol.* **13**(8), 890–896 (2003)
22. Weng, S., Zhao, Y., Pan, J.S., Ni, R.: A novel reversible watermarking based on an integer transform. In: IEEE International Conference on Image Processing, 2007, ICIP 2007, vol. 3, pp. III-241–III-244. IEEE (2007)