Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105
Volume **11**, Number **1**, February **2026**

MHM-based Reversible Data Hiding with Adaptive Feature Selection

Jiehuang Chen[1]

[1]School of Computer Science and Mathematics
Fujian University of Technology, Fuzhou 350118 China
1402573734@qq.com

Tiancong Zhang[2,*], Wenbing Zheng[2], Wei Chen[2]

[2]School of Electronic, Electrical Engineering and Physics
Fujian University of Technology, Fuzhou 350118, China
kushentian@163.com, zhwb190715@163.com, zcy@xzmu.edu.cn

Tien-Szu Pan[3]

[3]Department of Electronic Engineering
Kaohsiung University of Science and Technology, Kaohsiung, Taiwan
tien.pan@msa.hinet.net

[*]Corresponding author: Tiancong Zhang

Received March 7, 2024, revised September 2, 2024, accepted December 9, 2024.

Abstract. *Considering that traditional multiple histogram modification (MHM)-based reversible data hiding (RDH) methods utilize fixed features to generate multiple prediction-error histograms (PEHs), we propose a novel MHM-based RDH method by adaptively selecting different features with considering the statistical properties of images, so that multiple sharper PEHs are constructed. Moreover, two predictors, namely the rhombus and gradient-based edge direction predictors, are exploited to further improve the prediction performance. The experimental results also demonstrate that the proposed method provides better embedding performance in terms of the visual quality and payload than several related works.*
Keywords: Reversible data hiding, Multiple histogram modification

1. **Introduction.** Different from encryption [1–3] that prevents data from being grabbed, data hiding refers to the process of hiding data into a carrier in an imperceptible manner to convey the presence of embedded data, or to protect the transmitted contents [4]. Among various carriers such as images, videos, and audios, digital images are one of the most popular carriers because they are widely transmitted over the Internet. The image used for carrying data is called the cover image while the image with embedded data is called the stego image. When data are embedded into a cover image, the pixels will be changed, and thus the distortion is inevitably introduced. If the pixels are distorted permanently after data embedding, they cannot be recovered to their original states after completely extracting the embedded data. However, for some sensitive applications such as medical or military imaging, any permanent distortion caused by data embedding is unacceptable due to high requirements for legal consideration or high-precision nature. To this end, reversible data hiding (RDH) is introduced to fulfill the requirements by restoring the original image from the stego image along with the extraction of the embedded data. The

embedding performance of an RDH method is evaluated using the tradeoff between the embedding capacity versus the image quality.

Early RDH methods mainly achieve data embedding by lossless compression [5, 6]. A common feature of these methods is to losslessly compress features extracted from the cover image for leaving embedding space; however, the obtained embedding capacity is limited. In 2003, Tian's first proposed an RDH method based on difference expansion (DE), in which the difference value between two adjacent pixels is expanded to carry 1 bit of data. By expanding difference values to embed data, Tian can achieve the maximum capacity of 0.5 bit per pixel (bpp) but introduces relatively large distortion. Ni *et al.*'s method [7] proposed to embed data using histogram shifting (HS). The primary idea of HS is to embed data into peak bins of an image histogram while shifting the other bins to leave embedding spaces for data embedding. Since pixels are modified by 1 at most, Ni *et al.*'s method can offer high visual quality. However, the embedding capacity is limited by the peak bin heights. Afterwards, Tian's method are extended from the following three aspects, namely DE [8–10], prediction error expansion (PEE) [11] and integer-to-integer transform [12–17]. PEE is the most effective RDH method owing to its efficient payload-distortion tradeoff. PEE was firstly proposed by Thodi *et al.* [11], which performs DE and HS on the prediction errors of pixels generated by a median edge detection (MED) predictor [18], rather than difference value of two adjacent pixels in [19]. Subsequently, PEE-based RDH has been widely investigated in research such as two-dimensional (2D) prediction error histogram(PEH) modification [20], two-layer embedding [21–24]. In addition, some PEE-based methods focus on generating sharply distributed prediction error histogram (PEH) by improving predictors such as MED [18], gradient-adjusted predictor [25], rhombus predictor [22, 26], interpolation technique [21], asymmetric predictor [24, 27] and multiple predictors [28]. Among the aforementioned techniques, Chen *et al.* [27] proposed to generate a asymmetric PEH using a predictor based on MED, thus reducing the embedding distortion by decreasing the number of shifted prediction errors. Kouhi *et al.* [24] proposed to generate two asymmetric PEHs using weibull generalized exponential distribution (WGED) [29], thereby increasing the visual quality by reducing the number of invalid modifications.

Note that, most existing PEE-based RDH methods are based on one-dimensional (1D) or two-dimensional (2D) PEH. Given that the 1D or 2D PEH modifies all prediction errors uniformly without considering their local properties, Li *et al.* [30] proposed an embedding mechanism based on multiple histograms modification (MHM), in which the embedding performance is optimized by adaptively determining two expansion bins (i.e., a pair of expansion bins) for each PEH according to the local properties of each PEH. The optimal expansion bin determination is formulated as a rate-distortion optimization problem. However, directly exploiting exhaustive search to solve the rate-distortion optimization would result in huge computational complexity. To reduce the computational complexity, Li *et al.* utilize exhaustive search to determine two optimal bins for each histogram after decreasing the solution space. Afterwards, MHM has received increasing attention [31–35]. Among them, Ou *et al.* [31] proposed to expand multiple pairs in each PEH for achieving high embedding capacity, rather than expanding a single pair of each PEH like in Li *et al.*'s method [30]. Moreover, an advisable expansion bin selection strategy is used to optimize the embedding performance. Differently from the expansion bin selection strategy used in Ou *et al.*'s method [31], Qi *et al.* [32] proposed a computationally efficient algorithm to adaptively determine optimal multiple pairs for each histogram. Wang *et al.* proposed to utilize the fussy C-means (FCM) clustering with multiple features for constructing multiple sharper PEHs. Weng *et al.* [34] proposed to employ K-means clustering with multiple features for constructing multiple sharper PEHs and utilize an

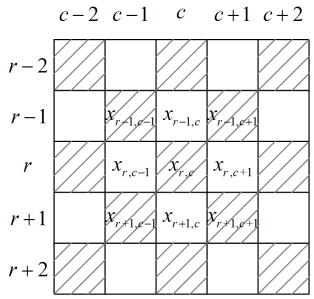


FIGURE 1. Two non-overlapping sets Ψ_+ and Ψ_-.

improved crisscross optimizing algorithm to determine optimal expansion bins for each histogram. Wang *et al.* [35] proposed to construct an RDH general framework based on MHM, which involves several key issues, namely generating multiple PEHs based on optimized multi-features, allocating the embedding capacity among multiple histograms and solving the optimization problem using the evolutionary algorithms.

Different from traditional MHM-based RDH methods that select fixed features of pixels to construct multiple PEHs, this paper focuses on adaptively selecting different features based on the statistical properties of images to construct sharper PEHs. Moreover, two predictors, including the rhombus and gradient-based edge direction predictors, are conductive to generate higher peak points for each PEH.

The rest of this paper is organized as follows. The proposed method is presented in Section 2. Section 3 provides the experimental results and the conclusions are given in Section 4.

2. **The proposed method.**

This paper focuses on generating sharper PEHs by adapting various strategies like multiple predictors or adaptive feature selection, thereby reducing unnecessary pixel modifications while improving the embedding performance. Different from traditional MHM-based RDH methods that construct multiple PEHs using fixed features of pixels, the proposed method generates sharper PHEs by adaptively choosing features based on the statistical properties of images, thereby enhancing the embedding performance. The proposed method is mainly composed of the following two parts: two predictors and adaptive feature selection. Subsequently, we will introduce each part separately in the following subsections.

2.1. **Two predictors.**

A cover image $\mathcal{O}$ is split into two non-overlapping sets Ψ_+ and Ψ_- according to the layout of a chessboard in Figure 1. The pixels in Ψ_+ are used for predicting the pixels in Ψ_- and the prediction errors of the pixels in Ψ_- are embedded with data, and vice versa. Since two sets are processed in the same manner, we only discuss the case of Ψ_+.

Different from traditional RDH methods that exploit an advanced predictor to generate a symmetric PEH, the proposed method utilizes two predictors, including the rhombus predictor [22] and gradient-based edge direction predictor (GEDP) [26], to generate two asymmetric PEHs with long and short tails. Specifically, for a pixel $x_{r,c}$ in Ψ_+, suppose its four neighboring pixels constitute a set denoted by $\mathcal{S}_{\mathcal{N}}$, namely, $\mathcal{S}_{\mathcal{N}} = \{x_{r-1,c}, x_{r,c-1}, x_{r,c+1}, x_{r+1,c}\}$. Let two prediction values be $\hat{x}^1_{r,c}$ and $\hat{x}^2_{r,c}$, namely $\hat{x}^1_{r,c} = \lfloor (x_{r-1,c} + x_{r+1,c} + x_{r,c-1} + x_{r,c+1})/4 \rfloor$, and $\hat{x}^2_{r,c}$ is calculated below:

$$\hat{x}^2_{r,c} = \left\lfloor \sum_{s,t\in\Phi} w_{s,t} \times x_{s,t} \Big/ \sum_{s,t\in\Phi} w_{s,t} \right\rfloor, \tag{1}$$

where $\Phi = \{(r \pm 1, c), (r, c \pm 1)\}$, $w_{s,t}$ is the weight assigned for $x_{s,t}$, i.e., $w_{r-1,c} = 1/(1 + |\Delta v_{r,c}| + 2|\Delta v_{r-1,c}| + |\Delta v_{r-2,c}|)$, $w_{r+1,c} = 1/(1 + |\Delta v_{r,c}| + 2|\Delta v_{r+1,c}| + |\Delta v_{r+2,c}|)$, $w_{r,c-1} = 1/(1 + |\Delta h_{r,c}| + 2|\Delta h_{r,c-1}| + |\Delta h_{r,c-2}|)$, and $w_{r,c-1} = 1/(1 + |\Delta h_{r,c}| + 2|\Delta h_{r,c-1}| + |\Delta h_{r,c-2}|)$. $w_{r,c-1} = 1/(1 + |\Delta h_{r,c}| + 2|\Delta h_{r,c-1}| + |\Delta h_{r,c-2}|)$.

$\Delta h_{r,c}$ is defined below:

$$\Delta h_{r,c} = \begin{cases} \frac{3}{2} \sum\limits_{u \in \{\pm 1\}} (x_{r+u,c+1} - x_{r+u,c-1}) \\ + \frac{1}{2} (x_{r,c+2} - x_{r,c-2}), \quad \text{if } x_{r,c} \in \Psi_+; \\ \frac{1}{4} \left(\sum\limits_{u \in \{\pm 2\}} (x_{r+u,c+1} - x_{r+u,c-1}) \right. \\ \left. + \sum\limits_{u \in \{\pm 1\}} (x_{r+u,c+2} - x_{r+u,c-2}) \right) \\ + \frac{5}{2} (x_{r,c+1} - x_{r,c-1}), \quad \text{otherwise.} \end{cases} \tag{2}$$

And, $\Delta v_{r,c}$ is calculated as follows:

$$\Delta v_{r,c} = \begin{cases} \frac{3}{2} \sum\limits_{u \in \{\pm 1\}} (x_{r+1,c+u} - x_{r-1,c+u}) \\ + \frac{1}{2} (x_{r+2,c} - x_{r-2,c}), \text{ if } x_{r,c} \in \Psi_+; \\ \frac{1}{4} \left(\sum\limits_{u \in \{\pm 2\}} (x_{r+1,c+u} - x_{r-1,c+u}) \right. \\ \left. + \sum\limits_{u \in \{\pm 1\}} (x_{r+2,c+u} - x_{r-2,c+u}) \right) \\ + \frac{5}{2} (x_{r+1,c} - x_{r-1,c}), \text{ otherwise.} \end{cases} \tag{3}$$

After comparing $\hat{x}^1_{r,c}$ with $\hat{x}^2_{r,c}$, the larger and smaller numbers are denoted by $\hat{x}_{\max}$ and $\hat{x}_{\min}$, respectively. The final prediction value of $x_{r,c}$, denoted by $x^*_{r,c}$, is obtained using the following rule:

$$x^*_{r,c} = \begin{cases} \hat{x}_{\min}, \text{ if } \hat{x}_{\min} == \hat{x}_{max}; \\ \hat{x}_{\min}, \text{ if } x_{r,c} \le \hat{x}_{min}; \\ \hat{x}_{\max}, \text{ if } x_{r,c} \ge \hat{x}_{max}; \\ null, \text{ otherwise.} \end{cases} \tag{4}$$

The prediction error $e_{r,c}$ is calculated as $e_{r,c} = x_{r,c} - x^*_{r,c}$ when $\hat{x}_{min} \ge x_{r,c}$ or $x_{r,c} \ge \hat{x}_{max}$.

2.2. **Adaptive feature selection.** To efficiently estimate the local smoothness, 16 features $\{f_1, f_2, \ldots, f_{16}\}$ derived from Weng *et al.*'s method [36] are generated for each pixel $x_{r,c}$. 16 features are extracted for $x_{r,c}$, and the k^{th} ($\omega_k \in \{1, 2, \ldots, 16\}$) feature is allocated to the k^{th} weight termed ω_k.

$$\ell s(x_{r,c}) = \sum_{k=1}^{16} \omega_k \times f^k_{r,c}, \tag{5}$$

where $\ell s_{r,c}$ denotes the weighted summation of 16 features.

For the cover image $\mathcal{O}$, the correlation between the absolute value of each prediction error and the corresponding is calculated using the following formula:

$$\begin{aligned} R_{\mathbf{p_e},\mathbf{l_s}}(a) &= \frac{\mathrm{cov}(\mathbf{p_e},\mathbf{l_s})}{\sigma_{\mathbf{p_e}} \times \sigma_{\mathbf{l_s}}} \\ &= \frac{\sum\limits_{e_{r,c} \in \Psi_+} (e_{r,c} - \mu_e) \times (\ell s(x_{r,c}) - \mu_{\ell s})}{\sqrt{\sum\limits_{e_{r,c} \in \Psi_+} (e_{r,c} - \mu_e)^2} \sqrt{\sum\limits_{x \in \Psi_+} (\ell s(x_{r,c}) - \mu_{\ell s})^2}}, \end{aligned} \tag{6}$$

where μ_e denotes the mean value of the total prediction errors in Ψ_+, $\mu_{\ell s}$ is the mean of the total weighted features in Ψ_+. $\mathbf{p_e}$ is the vector containing all the prediction errors in Ψ_+, and $\mathbf{l_s}$ denotes the vector containing all the weighted features in Ψ_+.

The sequential quadratic programming (SQP) algorithm [37] is used to calculate the maximum value of Equation 6, and therefore, the corresponding 16 weights are obtained accordingly. Each of 16 weights is normalized to a real number in the range of -255 through 255. Subsequently, the absolute values of the 16 normalized weights are sorted in descending order, so that the features corresponding to the first 8 sorted weights are chosen. For some cover image, there exist $C_8^1 + C_8^2 + C_8^3 + C_8^4 + C_8^5 + C_8^6 + C_8^7 + C_8^8 = 255$ combinations of 8 features. 255 combinations are sequentially visited to search for the optimal feature combination for each image, which generates the best power signal-to-noise ratio (PSNR).

After obtaining the optimal feature combination, all the pixels in Ψ_+ are sorted in ascending order of weighted summation of the selected feature combination and then split into 16 classes $\{h_1, h_2, \ldots, h_{16}\}$. The genetic algorithm (GA) is exploited to search for the optimal embedding points for each histogram. The detailed implementation details of GA are mentioned in [38].

2.3. **Data hiding.**

2.3.1. *The additional information.* For the κ^{th} histogram h_κ, the improved discrete particle swarm optimization (IDPSO) [36], rather than the exhaustive search, is used to search for two optimal bins. Suppose that the two optimal bins of h_κ are ℓ_κ^* and γ_κ^*. For blind data extraction and image restoration, the additional information has to be embedded into the cover image. The additional information for each of two non-overlapping sets, i.e., Ψ_+ and Ψ_-, is mainly composed of the following several parts.

- The selected feature combination of 8 features;
- The parameters associated with IDPSO;
- $\{\ell_\kappa^*\}_{\kappa=1}^{16}$ and $\{\gamma_\kappa^*\}_{\kappa=1}^{16}$ for all 16 histograms;
- The compressed location map (LM) to avoid overflow and underflow. If a pixel in Ψ_+ or Ψ_- is valued at 0 or 255, the corresponding bit in the location map is recorded by 1; otherwise, the corresponding bit is 0. Simultaneously, the pixel valued 0 or 255 is changed to be 1 or 254, respectively.

The additional information is embedded into the cover image along with the payload. The border pixels having no sufficient neighbors to predict, are simply skipped, and their least significant bits (LSBs) are used to carry the additional information. The replaced LSBs are appended to the payload. The detailed data embedding of $e_{r,c}$ in h_κ is given by

$$e'_{r,c} = \begin{cases} e_{r,c} - b, & \text{if } e_{r,c} = \ell_\kappa^*; \\ e_{r,c} - 1, & \text{if } e_{r,c} < \ell_\kappa^*; \\ e_{r,c} + 1, & \text{if } e_{r,c} > \gamma_\kappa^*; \\ e_{r,c} + b, & \text{if } x_{r,c} = \gamma_\kappa^*, \end{cases} \tag{7}$$

where b is one bit of data and $b \in \{0, 1\}$.

2.4. **Data extraction.** The additional information must be extracted prior to the payload from the border pixels. With the aid of the additional information, $e_{r,c}$ is retrieved by

$$e_{r,c} = \begin{cases} e'_{r,c} - 1, & \text{if } e'_{r,c} \le \ell_\kappa^*; \\ e'_{r,c} + 1, & \text{if } e'_{r,c} \ge \gamma_\kappa^*. \end{cases} \tag{8}$$

Simultaneously, the embedded bit b is extracted as

$$b = \begin{cases} 0, & \text{if } e'_{r,c} = \ell_\kappa^* \text{ and } e'_{r,c} = \gamma_\kappa^*; \\ 1, & \text{if } e'_{r,c} = \ell_\kappa^* - 1 \text{ and } e'_{r,c} = \gamma_\kappa^* + 1. \end{cases} \tag{9}$$

3. **Experimental results.** In this section, several experiments are conducted to demonstrate the superiority of the proposed method. The proposed method is compared with four related works including Kouhi *et al.*'s [24], Li *et al.*'s [30], Wang *et al.*'s methods [35]. Eight images taken from USC-SIPI [39] are utilized as test images.

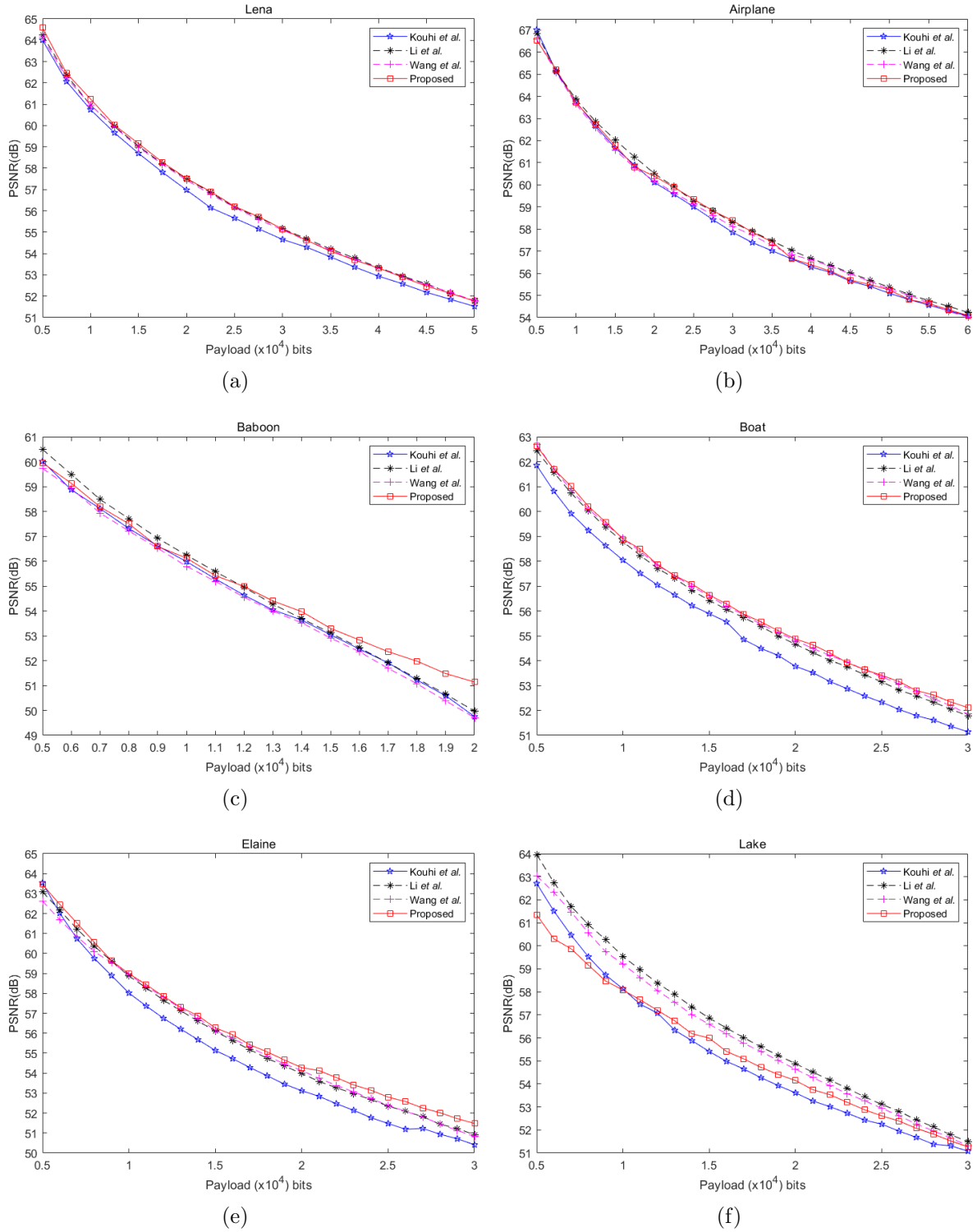
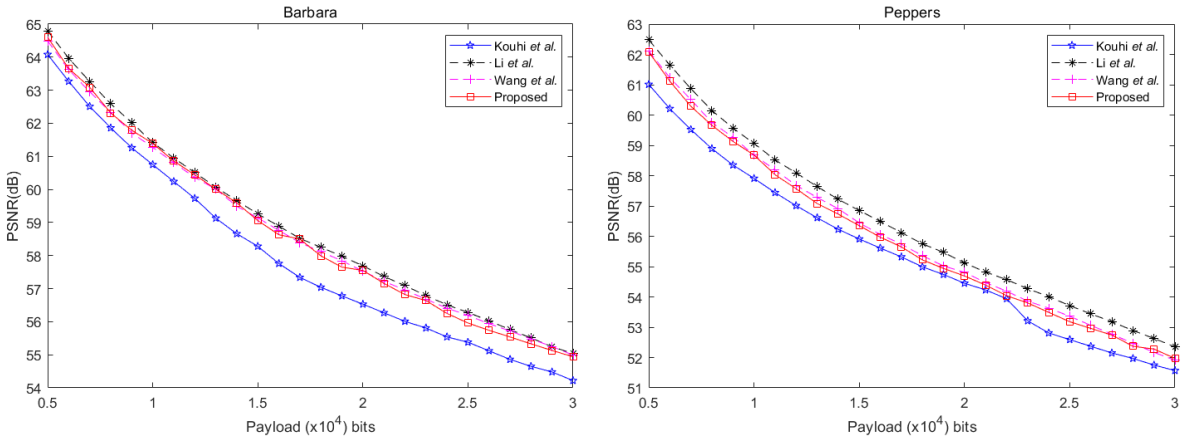


FIGURE 2. Comparison of PSNR versus payload among four methods.



(g) (h)

FIGURE 2. Comparison of PSNR versus payload among four methods (continued).

From Figure 2, it can be clearly observed that the proposed method achieves higher or comparable embedding capacity than the three compared methods.

4. **Conclusion.** This paper improves the embedding performing by constructing sharp PEHs from two aspects: two predictors and adaptive feature selection. Unlike fixed feature selection utilized in traditional MHM-based RDH methods, the adaptive feature selection strategy adaptively selects different features for images by considering the statistical characteristics of images. Two predictors are jointly utilized to include the smaller predictor errors into data embedding while excluding larger prediction errors from data embedding. The experimental results also demonstrate the proposed method offer higher or comparable embedding performance than several related works.

Acknowledgments. This work was supported in part by the National NSF of China under Grant 61872095, Grant 61571139, Grant 61872128, in part International Scientific and Technological Cooperation of Guangdong Province under Grant 2019A050513012, in part by the Open Project Program of Shenzhen Key Laboratory of Media Security under Grant ML-2018-03, in part by Fujian Science Fund for Distinguished Young Scholars under Grant 2020J06043.

REFERENCES

[1] T.-Y. Wu, X. N. Fan, K.-H. Wang, C.-F. Lai, N. X. Xiong, and M.-T. Wu, "A DNA computation based image encryption scheme for cloud CCTV systems," *IEEE Access*, vol. 7, pp. 181 434–181 443, 2019.

[2] T.-Y. Wu, X. N. Fan, K.-H. Wang, J.-S. Pan, and C.-M. Chen, "Security analysis and improvement on an image encryption algorithm using chebyshev generator," *Journal of Internet Technology*, vol. 20, no. 1, pp. 13–23, 2019.

[3] T.-Y. Wu, X. N. Fan, K.-H. Wang, J.-S. Pan, C.-M. Chen, and M.-T. Wu, "Security analysis and improvement of an image encryption scheme based on chaotic tent map," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 9, no. 4, pp. 1050–1057, 2018.

[4] W. Hong, T.-S. Chen, and J. Chen, "Reversible data hiding using delaunay triangulation and selective embedment," *Information Sciences*, vol. 308, pp. 140–154, 2015.

[5] J. Fridrich, M. Goljan, and R. Du, "Lossless data embedding-new paradigm in digital watermarking," in *EURASIP J. Appl. Signal Processing*, vol. 2002, no. 2, 2002, pp. 185–196.

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

[7] Z. Ni, Y.-Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 3, pp. 354–362, 2006.
[8] Y. Hu, H. K. Lee, and J. Li, "DE-based reversible data hiding with improved overflow location map," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 2, pp. 250–260, 2009.
[9] W. L. Tai, C. M. Yeh, and C. C. Chang, "Reversible data hiding based on histogram modification of pixel differences," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 6, pp. 906–910, 2009.
[10] X. L. Li, B. Yang, and T. Zeng, "Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3524–3533, 2011.
[11] D. M. Thodi and J. J. Rodriguez, "Expansion embedding techniques for reversible watermarking," *IEEE Transactions on Image Processing*, vol. 16, no. 3, pp. 721–730, 2007.
[12] A. M. Alattar, "Reversible watermark using the difference expansion of a generalized integer transform," *IEEE Transactions on Image Processing*, vol. 13, no. 8, pp. 1147–1156, 2004.
[13] D. Coltuc and J. M. Chassery, "Very fast watermarking by reversible contrast mapping," *IEEE Signal Processing Letters*, vol. 14, no. 4, pp. 255–258, 2007.
[14] S. W. Weng, Y. Zhao, J. S. Pan, and R. Ni, "Reversible watermarking based on invariability and adjustment on pixel pairs," *IEEE Signal Processing Letters*, vol. 15, pp. 721–724, 2008.
[15] X. Wang, X. Li, B. Yang, and Z. Guo, "Efficient generalized integer transform for reversible watermarking," *IEEE Signal Processing Letters*, vol. 17, no. 6, pp. 567–570, 2010.
[16] F. Peng, X. L. Li, and B. Yang, "Adaptive reversible data hiding scheme based on integer transform," *Signal Processing*, vol. 92, no. 1, pp. 54–62, 2012.
[17] D. Coltuc, "Low distortion transform for reversible watermarking," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 412–417, 2012.
[18] M. Weinberger, G. Seroussi, and G. Sapiro, "The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS," *IEEE Transactions on Image Processing*, vol. 9, no. 8, pp. 1309–1324, 2000.
[19] J. Tian, "Reversible data embedding using a difference expansion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 8, pp. 890–896, 2003.
[20] B. Ou, X. Li, Y. Zhao, R. Ni, and Y.-Q. Shi, "Pairwise prediction-error expansion for efficient reversible data hiding," *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 5010–5021, 2013.
[21] L. Luo, Z. Chen, M. Chen, X. Zeng, and Z. Xiong, "Reversible image watermarking using interpolation technique," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 1, pp. 187–193, 2010.
[22] V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y.-Q. Shi, "Reversible watermarking algorithm using sorting and prediction," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 7, pp. 989–999, 2009.
[23] S. W. Weng, J.-S. Pan, and L. D. Li, "Reversible data hiding based on an adaptive pixel-embedding strategy and two-layer embedding," *Information Sciences*, vol. 369, pp. 144–159, 2016.
[24] A. Kouhi and M. H. Sedaaghi, "Reversible data hiding based on high fidelity prediction scheme for reducing the number of invalid modifications," *Information Sciences*, vol. 589, pp. 46–61, 2022.
[25] M. Fallahpour, "Reversible image data hiding based on gradient adjusted prediction," *IEICE Electronics Express*, vol. 20, no. 5, pp. 870–876, 2008.
[26] W.-J. Yang, K.-L. Chung, H.-Y. M. Liao, and W.-K. Yua, "Efficient reversible data hiding algorithm based on gradient-based edge direction prediction," *Journal of Systems and Software*, vol. 86, pp. 567–580, 2013.
[27] X. Y. Chen, X. M. Sun, H. Sun, Z. Zhou, and J. Zhang, "Reversible watermarking method based on asymmetric-histogram shifting of prediction errors," *Journal of Systems and Software*, vol. 86, no. 10, pp. 2620–2626, 2013.
[28] X. X. Ma, Z. B. Pan, S. Hua, and L. F. Wang, "High-fidelity reversible data hiding scheme based on multi-predictor sorting and selecting mechanism," *Journal of Visual Communication and Image Representation*, vol. 28, pp. 71–82, 2015.
[29] A. Mustafa, B. S. El-Desouky, and S. AL-Garash, "Weibull generalized exponential distribution," 2016, arXiv:1606.07378.
[30] X. Li, W. Zhang, X. Gui, and B. Yang, "Efficient reversible data hiding based on multiple histograms modification," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 9, pp. 2016–2027, 2015.

[31] B. Ou, X. Li, W. Zhang, and Y. Zhao, "Improving pairwise PEE via hybrid-dimensional histogram generation and adaptive mapping selection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 7, pp. 2176–2190, 2019.

[32] W. F. Qi, X. L. Li, T. Zhang, and Z. M. Guo, "Optimal reversible data hiding scheme based on multiple histograms modification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 8, pp. 2300–2300, 2020.

[33] J. Wang, N. Mao, X. Chen, J. Ni, and Y.-Q. Shi, "Multiple histograms based reversible data hiding by using FCM clustering," *Signal Processing*, vol. 159, pp. 193–203, 2019.

[34] S. W. Weng, W. L. Tan, B. Ou, and J. S. Pan, "Reversible data hiding method for multi-histogram point selection based on improved crisscross optimization algorithm," *Information Sciences*, vol. 549, pp. 13–33, 2021.

[35] J. X. Wang, X. Chen, J. Q. Ni, N. X. Mao, and Y. Q. Shi, "Multiple histograms based reversible data hiding: Framework and realization," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 8, pp. 2313–2328, 2020.

[36] S. W. Weng, T. S. Hou, T. C. Zhang, and J.-S. Pan, "Adaptive smoothness evaluation and multiple asymmetric histogram modification for reversible data hiding," *Journal of Visual Communication and Image Representation*, vol. 90, p. 103732, 2023.

[37] P. T. Boggs and J. W. Tolle, "Sequential quadratic programming," *Actanumerica*, vol. 4, pp. 1–15, 1995.

[38] T. C. Zhang, C. J. Yang, S. W. Weng, and T. S. Hou, "Adaptive multi-histogram reversible data hiding with contrast enhancement," *Journal of Visual Communication and Image Representation*, vol. 89, p. 103637, 2022.

[39] "USC-SIPI image database," http://sipi.usc.edu/database/, 1977.