Journal of Information Hiding and Multimedia Signal Processing ©2012 ISSN 2073-4212
Ubiquitous International Volume **3**, Number **2**, April **2012**

Reversible Data Hiding for High-Quality Images Based on Integer Wavelet Transform

Ching-Yu Yang[a] Chih-Hung Lin[b] and Wu-Chih Hu[a]

[a]Department of Computer Science and Information Engineering
National Penghu University of Science and Technology
Penghu, Taiwan
chingyu@npu.edu.tw; wchu@npu.edu.tw

[b]Graduate Institute of Mathematics and Science Education
National Chiayi University
Chiayi, Taiwan
chuck@mail.ncyu.edu.tw

Received March 2011; revised April 2012

Abstract. *This article presents a simple reversible data hiding scheme based on the integer wavelet transform (IWT). By adjusting the coefficient values, data bits are effectively embedded into the low-high (LH) and high-low (HL) subbands of the IWT domain. Simulations show that both the host media and secret message can be completely recovered, without distortion, if the stego-images remain intact. Moreover, the resulting perceived quality of the image is highly satisfactory, as is the hiding capacity. In addition, the resulting Peak Signal-to-Noise Ratio (PSNR) of the proposed method is better than that of existing techniques. Furthermore, the stego-images generated by the proposed method have a certain degree of robustness which is resistant to image processing operations.*

Keywords: reversible data hiding, integer wavelet transform, watermarking

1. Introduction.

Recently, data hiding has played an important role in protecting (or securing) sensitive data [1, 2, 3, 4, 5]. Most data hiding techniques perform data embedment by altering the contents of a host media. As a result, the host image cannot be completely recovered after the bits extraction. These types of data hiding techniques are thus irreversible. To preserve the originality of valuable or priceless media, particularly those associated with law enforcement, medical and military image systems, and media with geographic information, many authors have developed reversible data hiding techniques. Reversible data hiding, also known as lossless data hiding, enables the recovery of both the original host media and the secret message from the stego-images. These methods can be divided into four classifications, as outlined in the following subsections.

1.1. Difference Expansion Techniques.

Tian [6] introduced a difference expansion (DE) technique to derive high capacity, low-distortion reversible data hiding. This technique divided the image into pairs of pixels, and a secret message was embedded into the difference between the pixels of each pair that was not expected to cause overflow/underflow issues. Simulation showed that the payload size and the perceived quality of the marked images generated by this technique were better than those achieved by the existing research at that time. Alattar [7] used DE with vectors instead of pixel pairs to extend and improve the performance of Tian's algorithm. In a single pass, Alattar's algorithm could embed several bits into every vector. Hsiao *et al.* [8] suggested an algorithm to hide data bits in two areas: data embedding and auxiliary information areas. The former was further divided into three categories of the blocks. Namely, a message was hidden in both the smooth and normal blocks, whereas the complex blocks contained no data bit. The overhead bits were embedded in the auxiliary information area. To increase hiding capacity, Hu *et al.* [9] proposed a variant DE-based technique that improved the compressibility of the location map. Compared to conventional DE-based schemes, this technique provided increased embedding storage and performed well with a variety of image types. Yang *et al.* [10] utilized the coefficient-bias scheme and proposed a reversible data hiding method. The secret message, which can be divided into two parts, was embedded into the spatial domain and/or frequency domain of a host image. The stego-images generated by the method were, to a certain degree, resistant to manipulation.

1.2. **Histogram-based Schemes.** Lin *et al.* [11] proposed a multilayer scheme for reversible data hiding based on a modification of the difference histogram. By combining the peak point of a difference image with a multilevel hiding strategy, the scheme could maintain a high capacity while keeping distortion low. Tai *et al.* [12] used a histogram modification technique with a binary tree structure to present a lossless data hiding approach. By using the difference between adjacent pixels rather than a single pixel, their algorithm could operate at a high capacity while maintaining low distortion. Based on a modification of the difference histogram between sub-sampled images, Kim *et al.* [13] developed an efficient lossless data hiding algorithm. The algorithm shifted the difference histogram and then embedded data bits into the modified pixel values. The algorithm prevented any overflow/underflow issues and did not require overhead information during data extraction. Hong *et al.* [14] utilized the modification of prediction errors (MPE) and proposed a lossless data hiding method. First, pixel values would be determined from an input image, and then error values would be obtained. A message could then be embedded by modifying the prediction errors. MPE can keep the distortion low if few messages are embedded, and leave sufficient vacant space to hide the desired payload.

1.3. **Prediction-based Methods.** Based on prediction-error expansion and the histogram shifting technique, Thodi and Rodriguez [15] presented an effective and reversible method of data hiding. Simulation showed that prediction-error expansion doubled the maximum embedding capacity in comparison to that provided by difference expansion. The perceived quality of the marked images was good at a moderate embedding capacity. By exploiting the expansion of the difference between a pixel and its predictive value, Tseng and Hsieh [16] proposed a high capacity and low distortion reversible data hiding scheme. Lee *et al.* [17] developed an adaptive reversible data hiding approach based on the prediction of difference expansion. Simulation demonstrated that this approach can achieve a high perceived quality of the marked image. Moreover, the location map would not be required during data extraction.

1.4. **Interpolation-based Algorithms.** By using the scaling-up neighbor mean interpolation method, Jung and Yoo [18] proposed a new lossless data hiding method. The interpolation technique provides the advantages of both low-time complexity and high-computing speed. Experiments indicated that this method can embed a large amount of bits into the host image while still keeping distortion low. Luo *et al.* [19] presented a novel reversible data hiding scheme using an interpolation technique, which can embed a secret message into images with imperceptible modification. Owing to the only slight modification of pixels, a high perceived quality of the resulting image would be preserved.

2. **Proposed Method.** To provide a good perceived quality and robustness, the proposed method only embeds a secret message into the low-high (LH) and high-low (HL) subbands of integer wavelet transform (IWT) domain. First, an input image is decomposed to the IWT domain by using the following two formulas:

$$d_{1,k} = s_{0,2k+1} - s_{0,2k} \tag{1}$$

and

$$s_{1,k} = s_{0,2k} + \left\lfloor \frac{d_{1,k}}{2} \right\rfloor, \tag{2}$$

where $s_{j,k}$ and $d_{j,k}$ are the kth low-frequency and high-frequency wavelet coefficients at the jth level, respectively [20]. The $\lfloor x \rfloor$ is a floor function. The details of the scheme are described in the following sections.

2.1. **Data Embedding.** Let $C_j = \{c_{jk}\}_{k=0}^{n^2-1}$ be the jth block of size $n \times n$ taken from the LH (or HL) subband of IWT domain. Also let $C_{jp} = \{c_p | \beta \leq c_p < 2\beta\}$ and $C_{jm} = \{c_m | -2\beta \leq c_m < -\beta\}$ be two subsets of C_j. The β used here is a control parameter. The main steps of bits embedding are specified as follows:

step1: Input a block C_j that has not been processed yet.

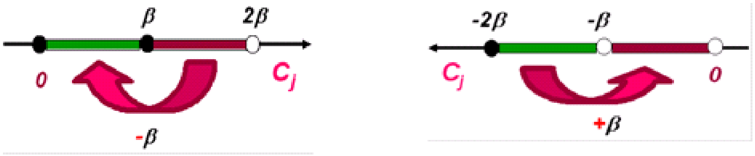
step2: If $|C_{jp}| \neq \varnothing$ then subtract β from each coefficient of C_{jp} and mark a flag to the modified coefficient.

step3: If $|C_{jm}| \neq \varnothing$ then add β to each coefficient of C_{jm} and mark a flag to the modified coefficient.

step4: After adjustment, for a coefficient $c_i \in C_j$ with $0 \leq c_i < \beta\,(or - \beta \leq c_i < 0)$, multiply c_i by 2 to obtain $\hat{c}_i$, and add an input bit to $\hat{c}_i$.

step5: Repeat from Step 1 until all blocks have been processed.

The purpose of Steps 3 and 4 are to further dig out hiding space from the selected coefficients. The schema for the adjustment of the coefficient values for the above two steps is illustrated in Fig. 1.



(a) (b)

Figure 1. The schema of the coefficients adjustment. (a) The positive part and (b) negative part.

To increase payload size, multiple bits can be hidden in each IWT coefficient. In this case, the above Steps 2-4 are rewritten as follows:

step2a: For a coefficient $c_t \in C_j$ with $-\beta < c_t < \beta$, multiply c_t by 2^k to obtain $\hat{c}_t$, and mark a flag to the modified coefficient.

step3a: For each $\hat{c}_t$, add data bits $\varnothing$ to $\hat{c}_t$ if $\hat{c}_t \geq 0$, otherwise, subtract $\varnothing$ from $\hat{c}_t$

The parameter k is an integer. To maintain a good perceived quality in the resulting image, the value of k should be no more than 2 The idea behind Steps 2-3 (and Steps 2a-3a) is to dig out an extra hiding space while keeping distortion low. From the above procedure we can see that the number of bits used for recording the indices of the modified coefficients is $\lfloor M/2n \rfloor \times \lfloor N/2n \rfloor \times n^2 \times 2 \leq MN/2$, where M and N are the size of a host image. The overhead information can be losslessly compressed by using either the run-length coding algorithm or JBIG2, and the resulting bit stream can then be sent by an out-of-band transmission to the receiver.

2.2. **Data Extraction.** Without loss of generality, let D_j be the jth hidden block of size $n \times n$ taken from the LH (or HL) subband of IWT domain which is derived from a stego-image, and let $\hat{D}_j = \left\{ \hat{d}_j | -2\beta \le \hat{d}_j < 2\beta \right\}$ with $\hat{D}_j \subseteq D_j$ The procedure of bits extraction can be summarized by the following steps.

step1: Input a block D_j that has not been processed yet.

step2: A data bit can be extracted by performing modulus -2 to $\hat{d}_j$

step3: The IWT coefficients $\tilde{d}_j$ which hide a data bit can be restored by performing either $\tilde{d}_j = \left\lfloor \hat{d}_j/2 \right\rfloor$ if $\hat{d}_j \ge 0$ or $\tilde{d}_j = \left| \left(\hat{d}_j/2 \right) - 0.5 \right|$ if $\hat{d}_j < 0$.

step4: The original IWT coefficients can be recovered by adding (or subtracting) β to (or from) $\tilde{d}_j$ if $\tilde{d}_j \ge 0$ (or $\tilde{d}_j < 0$) while the flag of $\tilde{d}_j$ was marked.

step5: Repeat from Step 1 until all data bits have been extracted.

Note that $\lfloor x \rfloor$ and $\lceil x \rceil$ in Step 3 denote the floor and ceiling functions respectively. To perform multiple bits extraction for each coefficient, the above Steps 2-4 are rewritten as follows:

step2b: A data bit can be extracted by performing modulus -2^k to d'_j with $-2^k\beta \le d'_j < 2^k\beta$.

step3b: The IWT coefficients $\tilde{d}_j$ which hide data bits can be restored by performing $\tilde{d}_j = \left\lfloor \hat{d}_j/2^k \right\rfloor$ if the flag of $\tilde{d}_j$ was marked.

step4b: The original IWT coefficients can be recovered by adding (or subtracting) $\left(2^k - 1\right)\beta$ to (or from) $\tilde{d}_j$ if $\tilde{d}_j \ge 0$ (or $\tilde{d}_j < 0$).

3. **Experimental Results.** Several 512 × 512 gray-scale images were used as the host images. A quarter of the host image *Lena* was used as test data. The size of the block was 4×4. The control parameter β was not set at a fixed value. The stego-images generated by the proposed method using β of value 10 were illustrated in Fig. 2. It can be seen from the figure that the perceived quality is good. Their average payload size in bits per pixel (bpp) and Peak Signal-to-Noise Ratio (PSNR) is 0.455 bpp and 38.727 dB, respectively. The PSNR is defined by

$$PSNR = 10 \times log_{10} \frac{255^2}{MSE}, \tag{3}$$

where $MSE = \frac{1}{MN} \sum_{i=1}^{N} \sum_{j=1}^{M} (\hat{x}(i,j) - x(i,j))^2$. Here $x(i,j)$ and $\hat{x}(i,j)$ denote the pixel values of the original image and the stego-image. In addition, the trade-off between PSNR and payload for the proposed method was depicted in Figure 3. The figure indicates that the average PSNR achieved by the proposed method was approximately 56 dB at a bit rate of 0.178 bpp. Conversely, the optimal PSNR value of 36.92 dB can be achieved in image *Jet* with a bit rate of 0.701 bpp. In addition, the relationship between bit rate and control parameter β is drawn in Fig. 4. From the figure we can see that the larger the value of β, the higher the bit rate that is achieved.

For the purpose of clarity, our method was compared with four successful schemes, namely, the Kim *et al.*'s approach [13], the Hong *et al.*'s technique [14], the Tseng and Hsieh's algorithm [16] and the Lee et al.'s scheme [17]. The performance comparisons between these methods are given in Table 1. From this table it is obvious that our

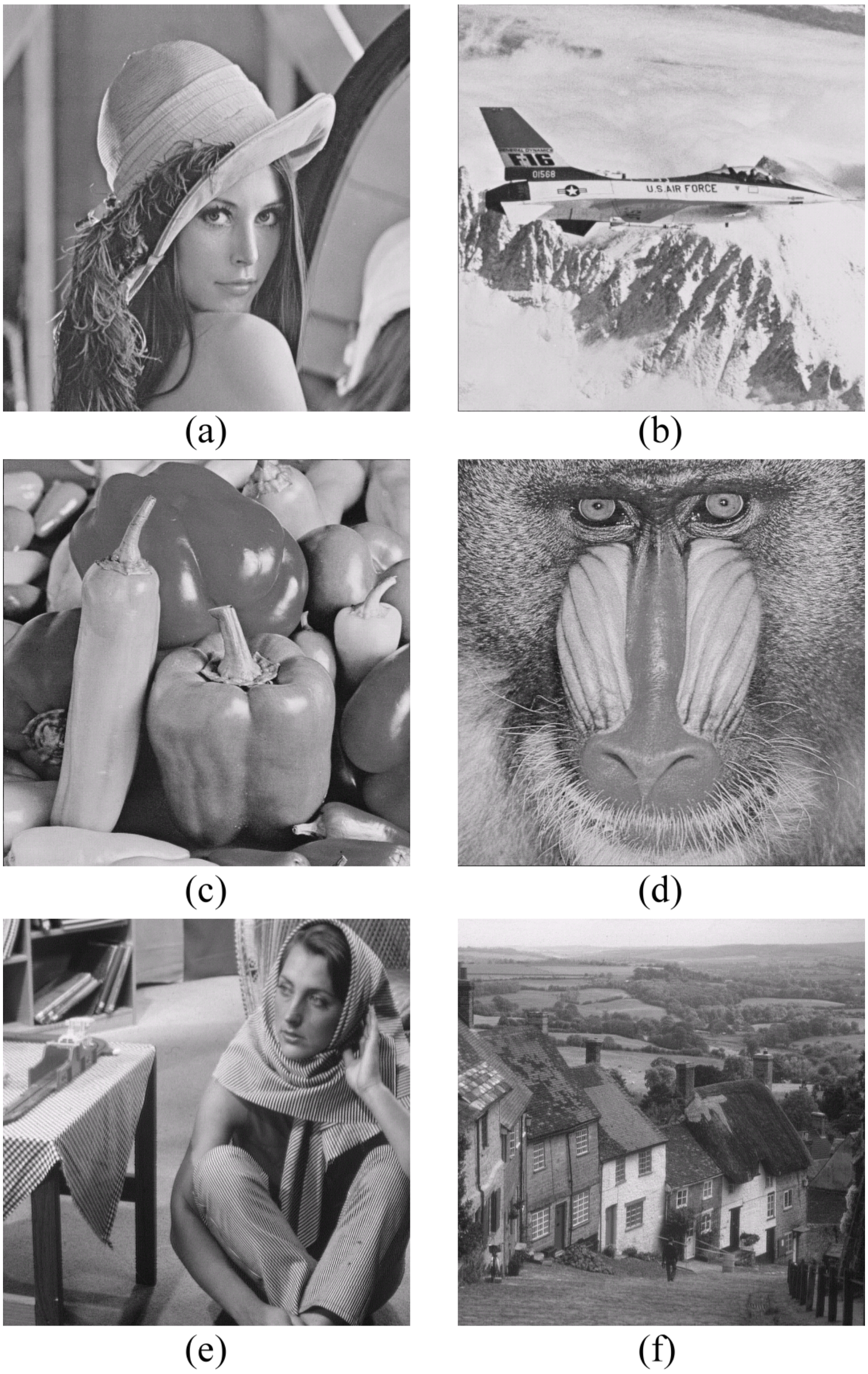


Figure 2. The stego-images generated by the proposed method. (a) Lena, (b) Jet, (c) Pepper, (d) Baboon, (e) Barbara, and (f) Goldhill.

proposed method provides the largest payload among these methods, while the PSNR for our proposed method is also superior to that of the other four algorithms. Moreover, Table 1 implies that the hiding capacity provided by our proposed method is nearly four times that achieved by the Kim *et al.*'s scheme [13], and is approximately double that achieved by the approach of Lee *et al.*'s scheme [17].

To demonstrate the robustness of the proposed method, examples of extracted watermarks (with a size of 124 124 at 8 bits/pixel, 2 colours), after various manipulations of

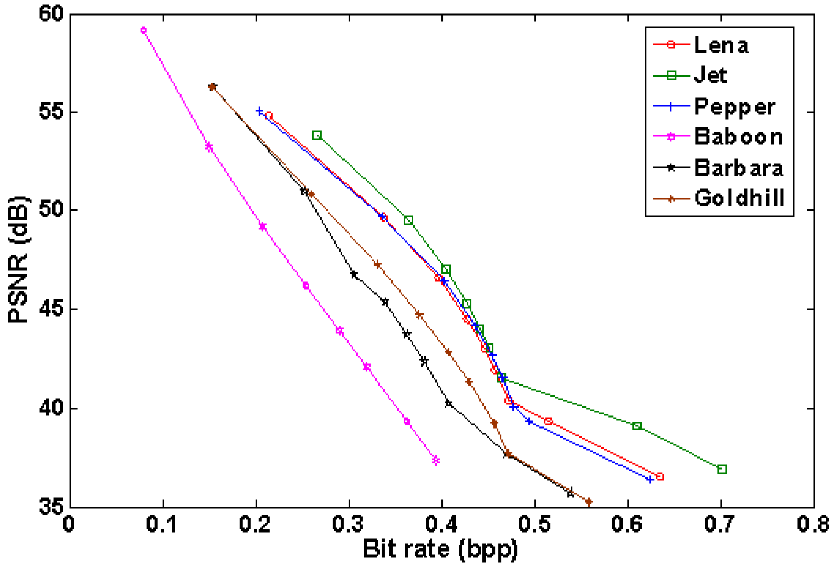


Figure 3. The hiding performance for the proposed method.

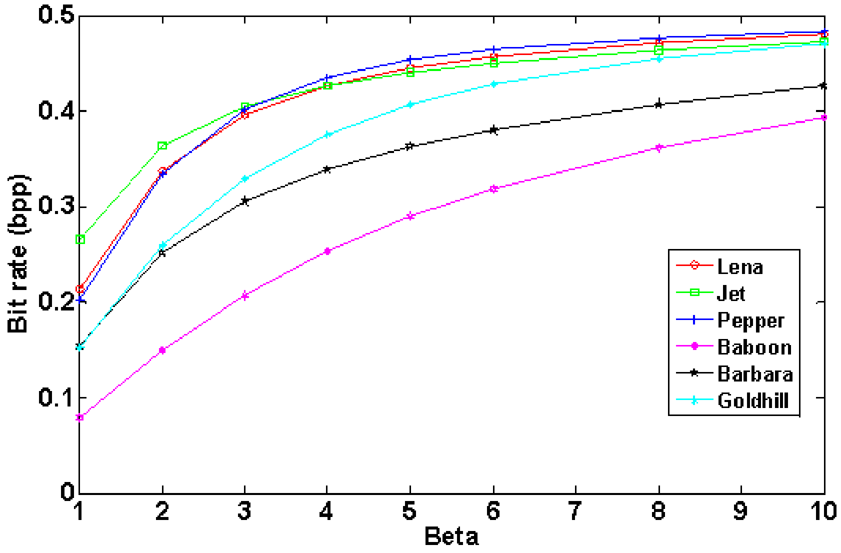
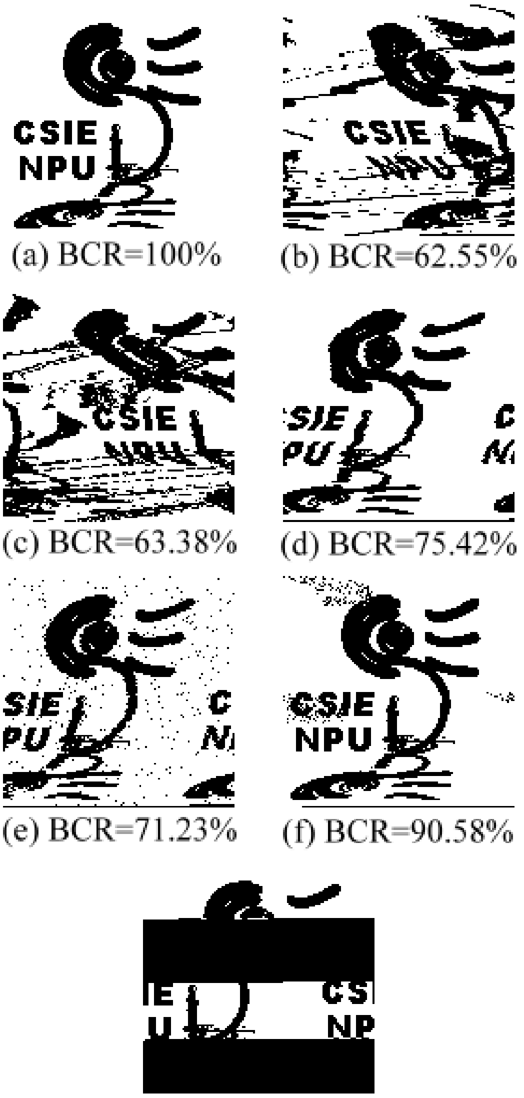


Figure 4. The relationship between and bit rate for the proposed method.

the image, are given in Fig. 4. The bit correct ratio (BCR) is also included. The BCR is defined by

$$BCR = \left(\sum\nolimits_{i=0}^{ab-1} \overline{w_i \oplus \tilde{w}_i} / \left(a \times b\right)\right) \times 100\%, \tag{4}$$

where w_i and $\tilde{w}_i$ represent the values of the original watermark and the extracted watermark respectively, as well as the size of a watermark is $a \times b$. Notice that a majority-vote decision was used during bits extraction. The BCR for an extracted watermark is 100% if a stego-image has not been attacked or altered. From Fig. 5 we can see that most of the extracted watermarks are recognizable. Although the BCR of the watermarks in Figs. 5(c) and 5(g) are not high, they are identifiable. From the above demonstration,



(a) BCR=100%

(b) BCR=62.55%

(c) BCR=63.38%

(d) BCR=75.42%

(e) BCR=71.23%

(f) BCR=90.58%

(g) BCR=52.11%

Figure 5. Examples of extracted watermarks after various manipulations. (a) Attack free, (b) Brightness (+70%), (c) Brightness (-70%), (d) Inversion, (e) JPEG (CR=1.47), (f) JPEG2000 (CR=1.72), and (g) Cropping (50%).

Table 1. Performance comparison between various methods.

Images	Bit rate/ PSNR				
	Kim et al.'s alg. [13]	Hong et al.'s alg. [14]	Tseng & Hsieh's alg. [16]	Lee et al.'s alg. [17]	Our method
Lena	0.07/ 48.9	0.33/ 48.93	0.22/ 47.31	0.14/ 48.54	0.34/ 49.70
Jet	0.12/ 49	0.27/ 48.79	0.29/ 47.24	0.19/ 48.54	0.36/ 49.53
Baboon	0.02/ 48.7	0.06/ 48.29	-	0.14/ 48.54	0.21/ 49.23
Average	0.08/ 48.87	0.22/ 48.67	0.26/ 47.28	0.16/ 48.54	0.30/ 49.49

we can conclude that the stego-images generated by the proposed method are resistant to attacks or alterations, including JPEG, JPEG2000, cropping, brightness, and inversion. Simulations also validate that the term β can be viewed as robustness parameter. Namely, the larger the value of β, the better the robustness for the proposed method.

4. **Conclusion.** The main reason that our proposed method outperforms existing schemes can be attributed to two main items. First, a smart adjustment of IWT coefficients employed in the proposed method can effectively embed data bits into the IWT block while keeping distortion low. Second, based on IWT domain, the stego-images generated by the proposed method are equipped with a certain degree of robustness to protect against image processing operations. Notably, of all of the above schemes performed in spatial domain, data extraction can fail at the receiver if even a slight alternation is intentionally (or unintentionally) imposed upon the stego-images. In addition, if an adversary stole (or eavesdropped) the stego-images generated by existing schemes, the hidden messages (or watermarks) have the potential to be extracted. This places the security and confidentiality of the embedded message at a severe risk. However, as the overhead of our method can be independently transmitted to the receiver, it is nearly impossible for any third parties (or malicious users) to extract the hidden message, and recover the original host image, when they only have access to the stego-images.

REFERENCES

[1] I. J. Cox, M. L. Miller, J. A. Bloom, J. Fridrich, and T. Kalker, *Digital Watermarking and Steganography*, 2^{nd} Ed., Morgan Kaufmann, 2007.

[2] S. Wang, B. Yang and X. Niu, A secure steganography method based on genetic algorithm, *Journal of Information Hiding and Multimedia Signal Processing*, vo1. 1, no. 1, pp. 28-35, 2010.

[3] K. Yamamoto and M. Iwakiri, Real-time audio watermarking based on characteristics of PCM in digital instrument, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 1, no. 2, pp. 59-71, 2010.

[4] C. C. Lin and P. F. Shiu, High capacity data hiding scheme for DCT-based images, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 1, no. 3, pp. 220-240, 2010.

[5] R. M. Noriega, M. Nakano, B. Kurkoski, and K. Yamaguchi, High payload audio watermarking: toward channel characterization of MP3 compression, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, no. 2, pp. 91-107, 2011.

[6] J. Tian, Reversible data embedding using a difference expansion, *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 8, pp. 890-896, 2003.

[7] A. M. Alattar, Reversible watermark using the difference expansion of a generalized integer transform, *IEEE Trans. Image Processing*, vol. 13, no. 8, pp. 1147-1156, 2004.

[8] J. Y. Hsiao, K. F. Chan, and J. M. Chang, Block-based reversible data embedding, *Journal of Signal Processing*, vol. 89, pp. 556-569, 2009.

[9] Y. Hu, H. K. Lee, and J. Li, DE-based reversible data hiding with improved overflow location map, *IEEE Trans. Circuits and Systems for Video Technology*, vol. 19, no. 2, pp. 250-260, 2009.

[10] C. Y. Yang, W. C. Hu, and C. H. Lin, Reversible data hiding by coefficient-bias algorithm, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 1, no. 2, pp. 91-100, 2010.

[11] C. C. Lin, W. L. Tai, and C. C. Chang, Multilevel reversible data hiding based on histogram modification of difference images, *Journal of Pattern Recognition*, vol. 41, pp. 3582-3591, 2008.

[12] W. L. Tai, C. M. Yeh, and C. C. Chang, Reversible data hiding based on histogram modification of pixel differences, *IEEE Trans. Circuits and Systems for Video Technology*, vol. 19, no. 6, pp. 906-910, 2009.

[13] K. S. Kim, M. J. Lee, H. Y. Lee, and H. K. Lee, Reversible data hiding exploiting spatial correlation between sub-sampled images, *Journal of Pattern Recognition*, vol. 42, pp. 3083-3096, 2009.

[14] W. Hong, T. S. Chen, and C. W. Shiu, Reversible data hiding for high quality images using modification of prediction error, *Journal of Systems and Software*, vol. 82, pp. 1833-1842, 2009.

[15] D. M. Thodi and J. Rodriguez, Expansion embedding techniques for reversible watermarking, *IEEE Trans. Image Processing*, vol. 16, no. 3, pp. 721-730, 2007.

[16] H. W. Tseng and C. P. Hsieh, Prediction-based reversible data hiding, *Journal of Information Science*, vol. 179, pp. 2460-2469, 2009.

[17] C. F. Lee, H. L. Chen, and H. K. Tso, Embedding capacity raising in reversible data hiding based on prediction of different expansion, *Journal of Systems and Software*, vol. 83, pp. 1864-1872, 2010.

[18] K. H. Jung and K. Y. Yoo, Data hiding method using image interpolation, *Journal of Computer Standard and Interfaces*, vol. 31, pp. 465-470, 2009.

[19] L. Luo, Z. Chen, M. Chen, Q. Zeng, and Z. Xiong, Reversible image watermarking using interpolation technique, *IEEE Trans. Information Forensics and Security*, vol. 5, no. 1, pp. 187-193, 2010.

[20] A. R. Calderbank, I. Daubechies, W. Sweldens, and B. L. Yeo, Wavelet transforms that map integers to integers, *Journal of Applied and Computational Harmonics Analysis*, vol. 5, no. 3, pp. 332-369, 1998.