Reversible Data Hiding in the Spatial and Frequency Domains

Ching-Yu Yang chingyu@npu.edu.tw
Dept. of Computer Science and Information Engineering
National Penghu University
Penghu, 880, Taiwan

Wu-Chih Hu wchu@npu.edu.tw
Dept. of Computer Science and Information Engineering
National Penghu University
Penghu, 880, Taiwan

Abstract

Combinational lossless data hiding in the spatial and frequency domains is proposed. In the spatial domain, a secret message is embedded in a host medium using the min-max algorithm to generate a stego-image. Subsequently, the stego-image is decomposed into the frequency domain via the integer wavelet transform (IWT). Then, a watermark is hidden in the low-high (LH) and high-low (HL) subbands of the IWT domain using the coefficient-bias approach. Simulations confirm that the hidden data is successfully extracted and the host image is completely recovered. In addition, the perceptual quality of the mixed image generated by the proposed method is good. Moreover, the mixed images are robust against attacks such as JPEG2000, JPEG, brightness adjustment, and inversion.

Keywords: Reversible data hiding, IWT, Min-max algorithm, Coefficient-bias approach.

1. INTRODUCTION

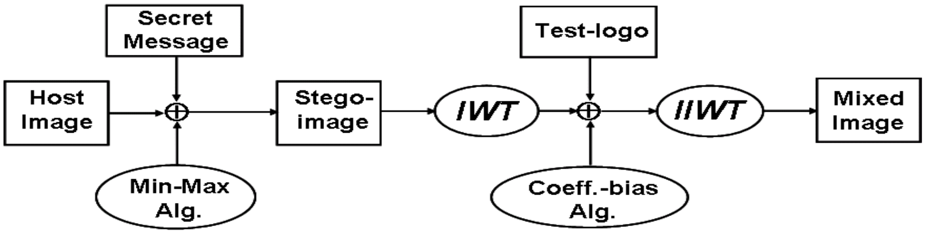
A stable and efficient data switching network makes it easy for individuals and organizations to exchange (or share) their resources on the Internet. Business-to-business (B2B), business-to-consumer (B2C), and customer-to-customer (C2C) commerce are three popular services provided over the Internet. However, data can be eavesdropped on, illicitly tampered, or falsified during transmission. Most commercial parties (or organizations) utilize encryption to protect important (or private) data during transactions. However, confidential data can become insecure if a private key is exposed or stolen by a third party. Data hiding techniques are an alternative solution to data protection. Generally speaking, data hiding can be classified into fragile watermarking and robust watermarking [1-2]. Fragile watermarking approaches [3-5] have the capability of hiding a large amount of data in a host medium while obtaining good resultant perceived quality. However, the marked images generated by these approaches are vulnerable to manipulations. Robust watermarking schemes [6-8] that can resist image processing attacks have been presented. However, most of the schemes allow a limited payload size.

Host media are important objects, such as law enforcement, military maps, and medical images, so they must not be damaged after digital watermarking. Several researchers presented lossless watermarking techniques [9-16]. Tian [9] implemented the difference expansion (DE) technique for lossless data hiding. To obtain extra storage space, Tian employed the DE technique to explore redundancy in the image content. Simulations showed that both the hiding capacity limit and the perceptual quality of the marked images were among the best at that time. Alattar [10] extended Tian's algorithm with DE of vectors, instead of pairs, to improve hiding efficiency. Using a generalized integer transform, Alattar presented a reversible watermarking algorithm, which has a very high-bit hiding capacity, along with high peak signal-to-noise ratio (PSNR) performance. Ni et al. [11] utilized the ideal of the zero (or the minimum) points of the histogram to embed data bits into a host medium. Although the average PSNR was 48.20 dB, the payload size was insufficient. Based on the idea of three-pixel block differences, Lin and Hsueh [12] suggested a high performance reversible hiding algorithm. The average (pure) payload was 1.79 bit per pixel (bpp), but the resultant PSNR was 22.06 dB. Lin et al. [13] presented a multilevel reversible data hiding scheme based on difference image histogram modification. By employing the peak point of a difference image with a multilevel hiding policy, the scheme allows a large number of embedded bits while maintaining good resultant perceptual quality. Using a location map, auxiliary information, and a novel LSB substitution, Hsiao et al. [14] employed a block-based reversible data hiding method. The average PSNR generated by the method was about 30 dB with an embedding rate of 1.02 bpp. Tseng and Chang [15] proposed a reversible watermarking algorithm using the idea of shiftable pixel pairs. The extended difference expansion algorithm has a great hiding capacity without producing noticeable distortion. Tsai et al. [16] utilized predictive coding and histogram shifting to further improve the performance of Ni et al.'s method. The technique has good hiding capability and resulting perceived quality for stego-images produced from medical images.

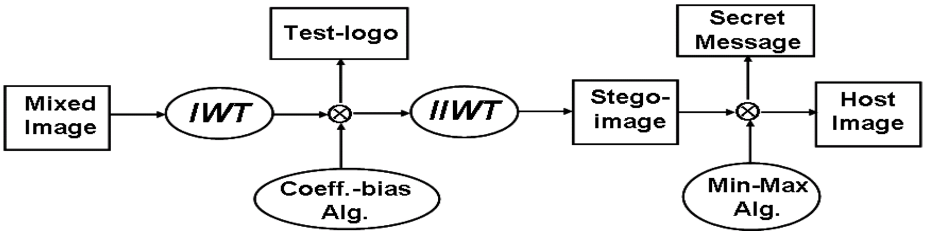
The above lossless data hiding schemes [9-16], which are conducted in the spatial domain, provide a large number of hiding bits. In the present study, we develop a reversible data hiding method based on the spatial and frequency domains that has the capability of resisting manipulations. The rest of the paper is organized as follows. The proposed min-max algorithm and coefficient-bias algorithm are described in Section 2. Section 3 presents the simulations. The conclusion is given in Section 4.

2. PROPOSED METHOD

In the proposed method, a secret message is first embedded in the spatial domain using the min-max algorithm, and then a watermark is hidden in the integer wavelet transform (IWT) domain [3] using the coefficient-bias approach. More specifically, the watermark is embedded in the low-high (LH) and high-low (HL) subbands of the L1 IWT domain. A schematic overview of the proposed method is shown in Fig. 1. Note that 'Secret Message' and 'Test-logo' as shown in Fig. 1 denote two various attributes of input data. However, it can be replaced by a single input message or a piece of icon. Notice as well the IIWT appeared in Fig. 1 stands for inverse integer wavelet transform.



(a) Transmitter



(b) Receiver

FIGURE 1: Block diagram of the proposed method. (a) Transmitter and (b) receiver.

2.1 Min-max algorithm

To provide extra storage space for hiding data bits, the proposed min-max algorithm was employed in the spatial domain. Without loss of generality, let $P = \{p_j\}_{j=0}^{(n\times n)-1}$ be an *n×n* nonoverlaping block divided from a host image. Let $p_{\min} = \arg Min\{p_j\}_{j=0}^{(n\times n)-1}$ and $p_{\max} = \arg Max\{p_j\}_{j=0}^{(n\times n)-1}$ be the minimum value and maximum value of the pixel in block *P*, respectively. Also let σ be a control parameter and *k* is a postive multiplier. The main steps of the min-max algorithm are as follows:

Step 1. Input a block *P* from a host image.

Step 2. Compute $p_{\min}$ and $p_{\max}$ of *P*.

Step 3. If $p_{\min} \geq 128$, then subtract $p_{\min}$ from p_j to obtain q_j; otherwise, subtract $p_{\max}$ from p_j to obtain q_j.

Step 4. If there exists a pixel $q_j \geq \sigma$, then add σ to q_j to obtain $\tilde{q}_j$. (The pixels are not qualified to carry bits.)

Step 5. If there exists $q_j < \sigma$, then multiply *k* by q_j to obtain $\hat{q}_j$. If an input bit is 1, add 1 to $\hat{q}_j$; otherwise, do nothing.

Step 6. If $p_{\min} \geq 128$, then add $p_{\min}$ to $\tilde{q}_j$ and $\hat{q}_j$, respectively; otherwise, subtract $p_{\max}$ from $\tilde{q}_j$ and $\hat{q}_j$, respectively. (The marked block contains the hidden bits.)

Step 7. Repeat Step 1 untill all data bits have been embedded in the block.

At the receiver, all of the data bits are sequentially extracted from the hidden block in a stego-image using a reverse procedure of the above algorithm. The host image can thus be completely recovered. Fig. 2 shows an example of bits being embedded using the min-max algorithm. In Fig. 2(a), we assume that the divided block has a size of 4×4 and that the input bit stream is "11101001011." *k* and σ are set at 2 and 5, respectively. Note that the minimum value, $p_{\min}$, and the maximum value, $p_{\max}$, of the block are 163 and 168, respectively. Step 3 of the algorithm produces a difference block, as shown in Fig. 2(b). To further alleivate distortion, the coefficients q_j which satisfy $q_j \geq \sigma$ are isolated from others in the block, as shown in Fig. 2(c), by adding σ *to* q_j via Step 4. The hidden block shown in Fig. 2(d) was obtained in Step 5. Finally, the marked block shown in Fig. 2(e) was generated in Step 6. Note that the mean square error (MSE) computed from the original block and the marked one is 8.69. An example of bit extraction is shown in Fig. 3. The figure shows a reverse procedure conducted on the marked block. The hidden bits are successfully extracted and the original block is completely recovered.

168	164	165	163
168	164	165	163
168	164	165	163
168	164	165	163

(a)

5	1	2	163
5	1	2	0
5	1	2	0
5	1	2	0

(b)

10	1	2	163
10	1	2	0
10	1	2	0
10	1	2	0

(c)

10	3	5	163
10	3	4	1
10	2	4	1
10	2	5	1

(d)

173	166	168	163
173	166	167	164
173	165	167	164
173	165	168	164

(e)

FIGURE 2: Example of bit embedding. (a) 4×4 block of the original block, (b) a difference block, (c) isolated-coefficients, (d) the hidden block, and (e) the marked block.

173	166	168	163
173	166	167	164
173	165	167	164
173	165	168	164

(a)

10	3	5	163
10	3	4	1
10	2	4	1
10	2	5	1

(b)

10	1	2	163
10	1	2	0
10	1	2	0
10	1	2	0

(c)

5	1	2	163
5	1	2	0
5	1	2	0
5	1	2	0

(d)

168	164	165	163
168	164	165	163
168	164	165	163
168	164	165	163

(e)

FIGURE 3: Example of bit extraction. (a) An input marked block, (b) coefficient subtraction, (c) bit extraction, (d) restored differecnce block, and (e) recovered original block.

2.2 Coefficient-bias approach

As described previously, the purpose of the coefficient-bias approach (with pixel adjustment) is to embed a watermark in the frequency domain. The details are given in the following three subsections.

2.2.1 Data embedding

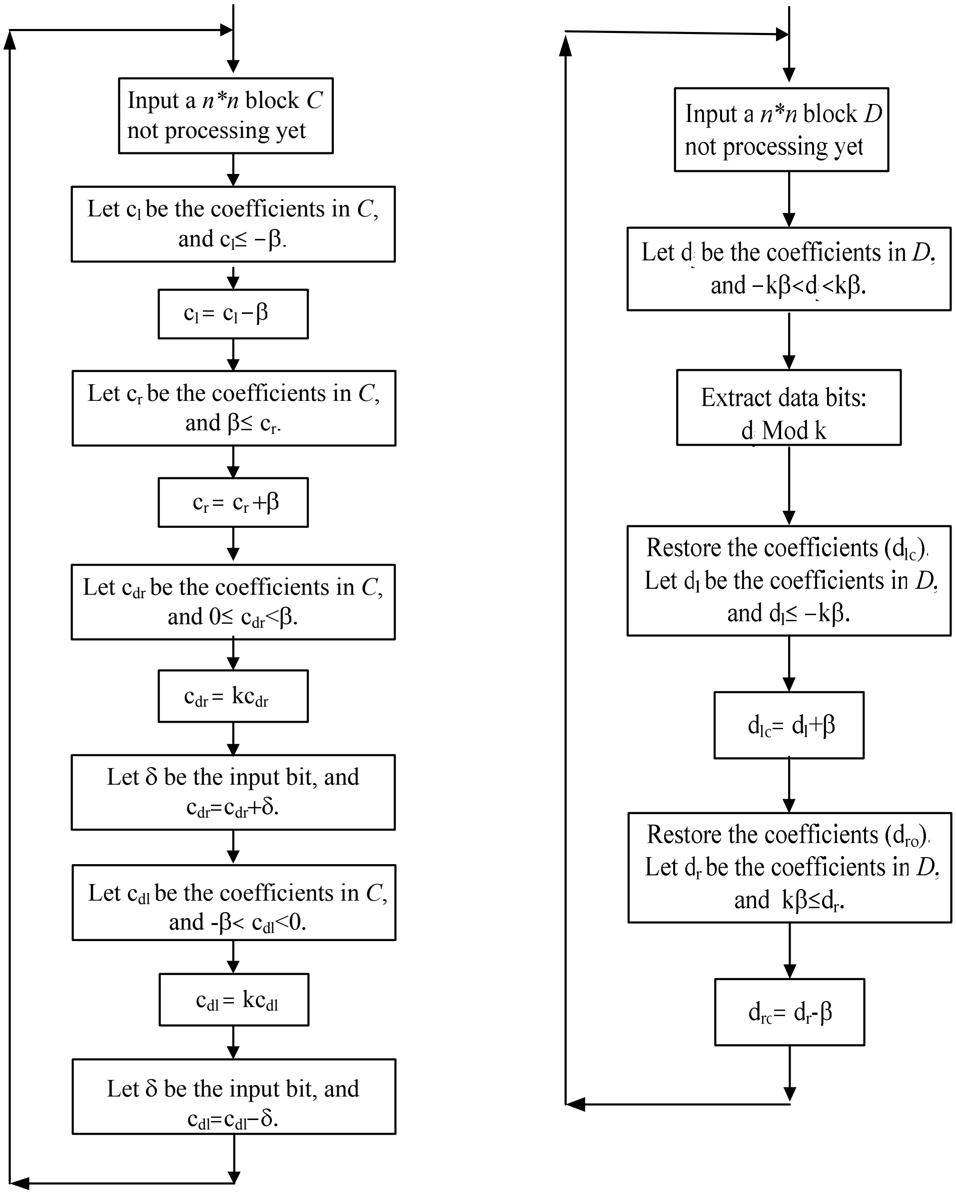
Decompose a stego-image into IWT domain. Input an *n*×*n* block $C = \{c_j\}_{j=0}^{(n\times n)-1}$ from the LH (or HL) of the IWT coefficient and δ be the input data. If there exists a coefficient $c_l \in C$ and $c_l \le -\beta,$ then subtract β from c_l. If there also exists a coefficient $c_r \in C$ and $\beta \le c_r,$ then add β to c_r. The payload provided by a host image is determined by the parameter β. Let $\hat{c}$ be the resultant coefficient of a IWT block. The above rules can be summarized as follows:

$$\hat{c} = \begin{cases} c_l - \beta, & \text{if } c_l \le -\beta; \\ c_r + \beta, & \text{if } c_l \ge \beta. \end{cases} \tag{1}$$

After coefficient adjustments, data bits are ready to be embedded in blocks. Multiply coefficients $c_{dr} \in C$ which satisfy $0 \le c_{dr} < \beta$ by *k* to obtain $\hat{c}_{dr}$. *k* is an integer. Add δ to $\hat{c}_{dr}$. Then, multiply coefficients $c_{dl} \in C$ which satisfy $-\beta < c_{dl} < 0$ by *k* to obtain $\hat{c}_{dl}$. Subtract δ from $\hat{c}_{dl}$. Normally, to embed a data bit into each of the candidate coefficients, the value of *k* is set at 2. The procedure is repeated until all data bits have been processed.

2.2.2 Data extraction

At the receiver, a marked image is first decomposed into the IWT domain. Then, read in a block *D* of size *n*×*n* from the LH and HL subbands of IWT, respectively. If there exists a coefficient $d_j \in D,$ which satisfies $-k\beta < d_j < k\beta,$ divide d_j by *k*. The hidden bits can be obtained from the residual. Subsequently, restore the coefficients which were originally located between -β and β by dividing d_j by *k*. Then, restore the coefficients which were originally less than or equal to -β by adding $d_l,$ which satisfies $d_l \le -2\beta$, to β and restore the coefficients which were originally greater than or equal to β by subtracting $d_r,$ which satisfies $d_r \ge 2\beta$, from β. The procedure is repeated until all data bits are extracted. The coefficient-bias approach is summarized in Fig. 4.



(a) Encoding part (b) Decoding part

FIGURE 4: Flowchart of the proposed approach. (a) Encoder and (b) decoder.

2.2.3 Pixel adjustment

The aim of the pixel adjustment used here is to ensure lossless data hiding. To determine whether the goal of a successful recovery of a mixed image is achieved or not, a prior data extraction is performed before the mixed image transmitted to the receiver. More specifically, If a stego-image cannot be losslessly recovered from a mixed image, pixel adjustment is utilized. That is, if pixel *p* in a host medium satisfies either $p < \phi_1$ or $\phi_2 < p$, then the pixel is adjusted to a new value by adding or subtracting pixel-offset γ(the value of γcan be set to be the same as that of the parameter β). The new pixel $\hat{p}$ is obtained using:

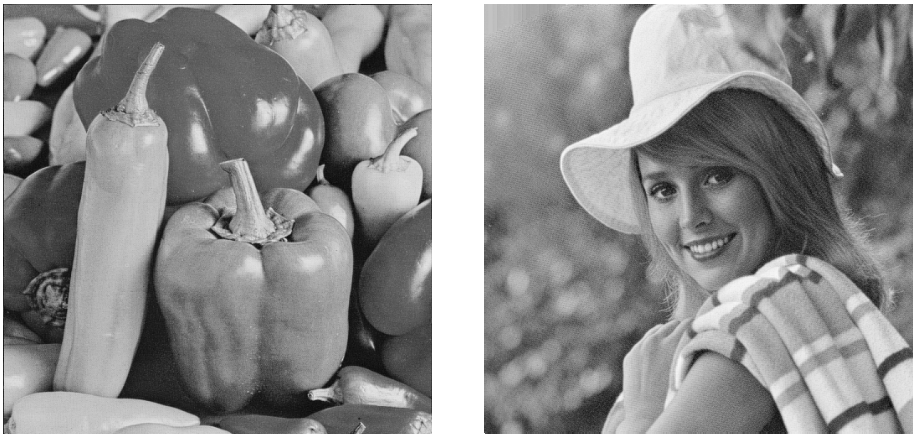
$$\hat{p} = \begin{cases} p+\gamma & \text{if } p < \phi_1; \\ p-\gamma, & \text{if } \phi_2 < p. \end{cases} \tag{2}$$

The overhead information, which is used to record the position of each adjusted-pixel, can be losslessly compressed [17] and out-of-band transmission to the receiver. The stego-images can be recovered completely at the receiver by a reverse pixel adjustment.

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 mixed images generated by embedding parts of the test data in the host images using the proposed method are shown in Fig. 5. The block size is 4×4. The control parameters σ and β were set to 3 and 8, respectively. The multiplier factor k used here is 2. Fig. 5 shows that the perceptual quality of the mixed images is good. Their hiding performance is listed in Table 1. Most of the images required no pixel adjustment during data embedding. 1-pixel and 10-pixel adjustments were required for the images Elaine and Sialboat, respectively. Note that the two sets of parameters (ϕ_1, ϕ_2) used in these two images were (1, 255) and (7, 255), respectively. The average PSNR is 34.59 dB with an embedding rate of 0.457 bpp. In addition, payload size generated by the proposed method in the spatial and frequency domain, respectively, is given in Table 2. It is obvious that payload size provided in IWT domain is about seven times larger than that provided in spatial domain. The trade-off between PSNR and hiding rate for the proposed method is shown in Fig. 6. To obtain higher PSNR performance with an embedding rate of less than 0.2 bpp, the value of σ should be set at 1, and that of β be set below 3. On the other hand, better bits-hiding capability is obtained when larger values of σ and β are used in the proposed method.

(a) Lena (b) Jet



(c) Peppers (d) Elaine

FIGURE 5: Mixed images generated by the proposed method. (a) Lena, (b) Jet, (c) Peppers, (d) Elaine, (e) Goldhill, and (f) Sailboat.

(e) Goldhill (f) Sailboat

FIGURE 5: Continued.

Images	Embedding rate (bpp)	PSNR (dB)	No. of pixel-adj.
Lena	0.491	34.99	0
Jet	0.527	35.29	0
Peppers	0.487	35.00	0
Elaine	0.423	34.09	1
Goldhill	0.410	34.12	0
Sailboat	0.403	34.06	10

TABLE 1: Hiding performance for Fig. 5.

Images	IWT domain	Spatial domain	Total payload	PSNR
Lena	110,557	18,246	128,803	34.99
Jet	110,794	27,476	138,270	35.29
Peppers	112,724	14,866	127,590	35.00
Elaine	99,650	11,210	110.860	34.09
Goldhill	95,877	11,599	107,476	34.12
Sailboat	94,154	11,019	199,329	34.06

TABLE 2: Payload size generated by the proposed method in the spatial and IWT domain.

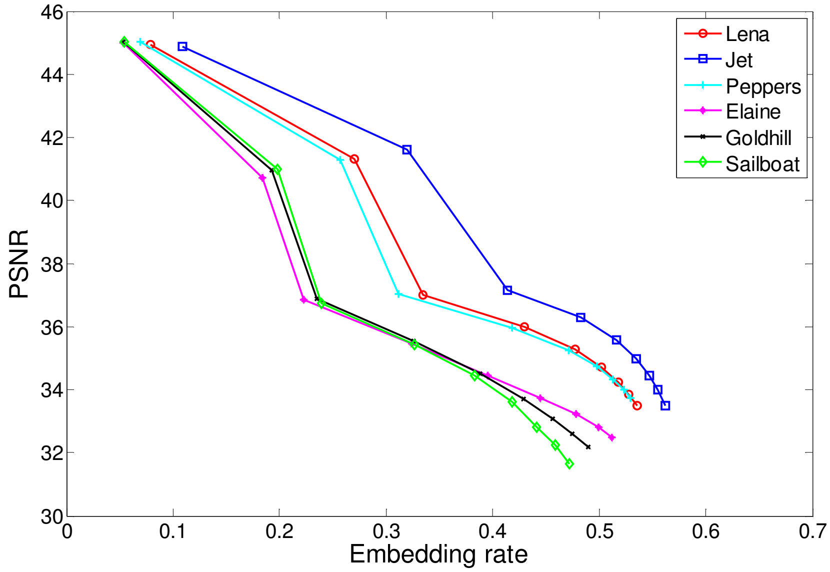


FIGURE 6: Trade-off between PSNR and hiding rate for the proposed method.

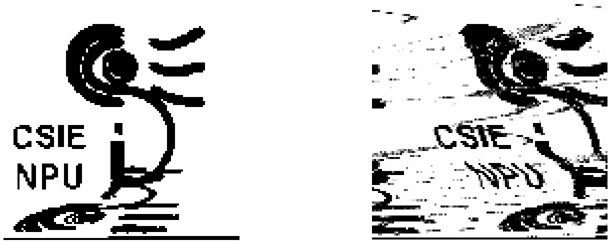
Performance comparison between our method and several lossless data hiding schemes [12-14] is listed in Table 3. It can be seen from Table 3 that the schemes (performed in the spatial domain) provide a large hiding capacity, but their average PSNR is about 30 dB. Since the perceived quality is not so good that it might be attarcted by the third parties. In other words, the resulting images generated by these scheme are vulnerable to attack. However, the resultant images generated by our method are more robust against attack than those generated by spatial-domain methods. Fig. 7 shows that the mixed images produced by the proposed method (using β=8, σ=3, and k=2 for image Lena) can resist attacks such as brightness adjustment (±45%), JPEG2000 coding with a compression ratio (CR) of 1.58, JPEG coding (with CR=1.36), and inversion. Although the bit correct ratio (BCR) for the watermarks in Fig. 7(b) and 7(c) are a bit low, the extracted watermarks are recognizable. Although the BCR of Fig. 7(e) is only 18.65%, the extracted watermark is still recognizable. It is interesting that the BCR of Fig. 7(f) is 100%, which means that the mixed images generated by our method are immune to an inversion attack. BCR is defined as:

$$BCR = \left(\sum_{i=0}^{MN-1} \overline{w_i \oplus w_i'} \Big/ M \times N \right) \times 100\% \qquad (3)$$

where w_i and w_i' represent the values of the original watermark and the extracted watermark, respectively. The watermark has a size of $M{\times}N$.

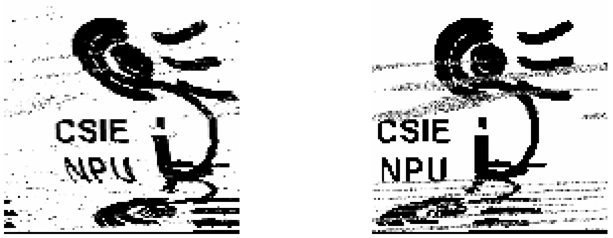
Images	Lin et al.'s tech. [12]	Lin et al.'s appr. [13]	Hsiao et al.'s alg. [14]	Proposed method
Lena	30.0/ 1.18	30.19/ 1.322	30.00/ 1.159	34.99/ 0.491
Jet	30.3/ 1.40	30.19/ 1.384	30.00/ 1.093	35.29/ 0.527
Peppers	30.2/ 1.36	30.19/ 1.305	30.00/ 1.159	35.00/ 0.487
Goldhill	30.1/ 1.16	-	30.00/ 0.936	34.12/ 0.410

TABLE 3: PSNR and embedding rate for the proposed method and other schemes.



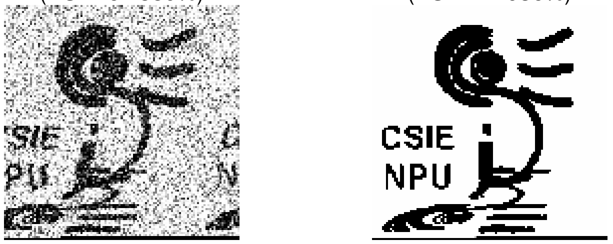
(a) Attack-free
(BCR=100%)

(b) Brightness (+45%)
(BCR=60.801%)



(c) Brightness (-45%)
(BCR=67.880%)

(d) JPEG2000 (CR=1.58)
(BCR=74.936%)



(e) JPEG (CR=1.36)
(BCR=18.65%)

(f) Inversion
(BCR=100%)

FIGURE 7: Examples of extracted watermarks (size of 117×117 with 8 bits/pixel, 2 colors) after various attacks. (a) Attack-free, (b) Brightness (+45%), (c) Brightness (-45%), (d) JPEG2000, (e) JPEG, and (f) Inversion.

4. CONCLUSION

An effective lossless data hiding scheme that embeds data bits in the spatial and frequency domains was proposed. The proposed method consists of two approaches, namely, the min-max algorithm and coefficient-bias approach. The min-max algorithm is used to hide a secret message in a host media in the spatial domain. In the frequency domain, a watermark is embedded in the LH- and HL-subbands of IWT using the coefficient-bias approach. Experiments indicate that not only a hidden data is successfully extracted but also a host image is losslessly restored. Moreover, the resultant perceptual quality generated by the proposed method is good. The mixed images can survive various manipulations, such as JPEG2000 and JPEG brightness adjustment, and inversion.

5. REFERENCES

1. F. Y. Shih. "*Digital watermarking and steganography: fundamentals and techniques*". CRC Press., FL (2008).

2. I. J. Cox, M. L. Miller, J. A. Bloom, J. Fridrich and T. Kalker. "*Digital watermarking and steganography, 2^{nd} Ed.*". Morgan Kaufmann., MA (2008).

3. G. Xuan, J. Zhu, J. Chen, Y. Q. Shi, Z. Ni and W. Su, "*Distortionless data hiding based on integer wavelet transform*". Electronics Letters, 38(25): 1646-1648, 2002.

4. H. C. Wu, N. I. Wu, C. S. Tsai and M. S. Hwang. "*Image steganographic scheme based on pixel-value differencing and LSB replacement methods*". IEE Proc. Vision Image Signal Processing, 152:611-615, 2005.

5. R. Z. Wang and Y. S. Chen. "*High-payload image steganography using two-way block matching*". IEEE T. Signal Processing Letter, 13(3):161-164, 2006.

6. H. M. Al-Otum and N. A. Samara. "*Adaptive blind wavelet-based watermarking technique using tree mutual difference*". Journal of Electronic Imaging, 15(4):043011-1~12, 2006.

7. X. Zhu, A. T. S. Ho and P. Marziliano. "*A new semi-fragile image watermarking with robust tampering restoration using irregular sampling*". Signal Processing: Image Communications, 22: 515-528, 2007.

8. Y. Govindarajan and S. Dakshinamurthi, "*Quality-security uncompromised and plausible watermarking for patent infringement*". International Journal of Image Processing, 1(2):11-20, 2007.

9. J. Tian. "*Reversible data embedding using a difference expansion*". IEEE T. Circuits and Systems for Video Technology, 13(8):890-896, 2003.

10. A. M. Alattar. "*Reversible watermark using the difference expansion of a generalized integer transform*". IEEE T. Image Processing, 13(8):1147-1156, 2004.

11. Z. Ni, Y. Q. Shi, N. Ansary and W. Su, "*Reversible data hiding*," IEEE T. Circuit and System for Video Technology, 16:354-362, 2006.

12. C. C. Lin and N. L. Hsueh. "*A lossless data hiding scheme based on three-pixel block differences*". Pattern Recognition, 41:1415-1425, 2008.

13. C. C. Lin, W. L. Tai and C. C. Chang. "*Multilevel reversible data hiding based on histogram modification of difference images*". Pattern Recognition, 41:3582-3591, 2008.

14. J. Y. Hsiao, K. F. Chan and J. M. Chang. "*Block-based reversible data embedding*". Signal Processing, 89:556-569, 2009.

15. H. W. Tseng and C. C. Chang. "An extended difference expansion algorithm for reversible watermarking". Image and Vision Computing, 26:1148-1153, 2009.

16. P. Tsai, Y. C. Hu and H. L. Yeh. "*Reversible image hiding scheme using predictive coding and histogram shifting*". Signal Processing, 89:1129-1143, 2009.

17. C. Saravanan and R. Ponalagusamy, "*Lossless grey-scale image compression using source symbols*". International Journal of Image Processing, 3(5):246-251, 2009.