Reversible Data Hiding Algorithm Using Dual Domain Embedding

Po-Whei Huang
Department of Computer Science and Engineering
National Chung Hsing University
Taichung, Taiwan, R. O. C.
powhei.huang@msa.hinet.net

Yung-Kuan Chan
Department of Management Information Systems
National Chung Hsing University,
Taichung, Taiwan, R. O. C. ykchan@nchu.edu.tw

Chia-Yi Chuang
Department of Computer Science and Engineering
National Chung Hsing University
Taichung, Taiwan, R. O. C.
s9356056@cs.nchu.edu.tw

Hao-Cheng Wang
Innovative DigiTech-Enabled Applications & Services
(IDEAS)
Institute of Information Industry
Taipei, Taiwan, R. O. C.
kevinhcwang@iii.org.tw

Abstract—**The image data hiding technique can embed the secret message into the digital image. For military or medical applications the reversible data hiding technique is preferred because it offers the advantage to recover the original image after the message extraction, yet it still maintains the reversibility. In order to achieve reversibility, we propose a reversible data hiding algorithm that explores the spatial and frequency domains. Predictive coding is employed to embed the secret message in the spatial domain. The proposed algorithm applies an integer-to-integer transformation to convert the image originating in the spatial domain into the frequency domain. We then modify the high frequency coefficients only when embedding the secret message. Consequently, the low frequency coefficients are unchanged enabling the embedded image to preserve the image quality due to the hidden message. In addition, the algorithm further increases the embedding capacity because it takes advantage of both the spatial and frequency domains.**

Keywords-Data hiding; reversible data embedding; predictive coding; wavelet transform

I. Introduction

The reversible data hiding scheme is known as lossless data hiding. Despite the secret message embedding, the reversible data hiding scheme can produce a full recovery of the original image even when the secret message is extracted. This reversible data hiding scheme is preferred in some applications, such as images used for military and medical purposes, because the original image cannot tolerate any tiny faults after a secret message is extracted. A large amount of reversible techniques directly embed secret messages into the pixel values of the original image. We refer to them as spatial domain reversible data hiding schemes [1-8]. Tian proposed a spatial domain reversible method with a high embedding capacity by using the concept of different expansions for a pair of pixels [5]. The scheme embeds one bit into the expandable difference of two pixels. Unfortunately, since some pixel pairs cannot embed a secret

bit due to the problem of overflow or underflow, the scheme needs extra information for distinguishing whether a pixel pair has any hidden message. This extra information is an overhead, and is referred to as the location map, leading to the decrease of the embedding capacity. To reduce the size of the location map, Thodi and Rodriguez simplify the contents of the location map by using the concept of difference histogram shifting [3].

Dissimilar to the spatial domain schemes which directly embed messages into pixels of the original image, some reversible data hiding methods apply a trans-formation operation to the original image before the secret messages are concealed in the transformed results in order to accomplish reversibility [4, 8]. These methods are referred to as frequency domain schemes. Lee, Yoo and Kalker used the integer-to-integer wavelet transform to produce wavelet coefficients [4]. Then, they adopted the bit-shifting technique to embed secret bits into wavelet coefficients. Weng, Zhao, Pan and Ni proposed a reversible integer transformation scheme based on the invariability of the sum of pixel pairs enabling their scheme to embed secret bits into each pixel pair [8]. That scheme increases the compression ratio of the location map by using the pairwise difference adjustment (PDA).

II. Integer-to-Integer Wavelet Transform

Wavelet transform is a technique used to transform an image from a spatial domain into a frequency domain. In general, the conventional wavelet transform is based on the floating-point operation. When we apply this technique directly to embed messages into an image for reversible data hiding, a truncation error may be encountered, leading to the failure of the message extraction and image recovery. To overcome this drawback, the invertible integer-to-integer wavelet trans-form was proposed [9] which can convert an image to the frequency domain without any information loss, thereby causing it to be appropriate for reversible data hiding. In the following we briefly describe the concept of this

transformation. Given an original image, the integer-to-integer wavelet transform generates wavelet coefficients, which can further be categorized into four sub-bands denoted by LL, LH, HL and HH. Each sub-band represents different degrees of importance with respect to the original image. In particular, the low frequency sub-band (LL) represents the miniature of the image indicating the most important part of the image. The high frequency sub-bands (LH, HL, and HH) indicate pixel differences in different directions which represent the variation of the image.

III. THE PROPOSED ALGORITHM

In this paper, we propose a reversible data hiding algorithm using both the spatial domain and the frequency domain. Our algorithm explores dual domain embedding enabling us to provide a high embedding capacity while producing a good visual quality of the embedded image. The original image can be reproduced from the embedded image after extracting the secret message. In the spatial domain, we apply a median edge detector (MED) predictor to embed a large capacity of payload. An integer to integer wavelet transform is then employed in the frequency domain. This allows our algorithm to embed more amounts of secret messages into high frequency areas without degrading the quality of the embedded image. This section describes our dual domain reversible data hiding algorithm. The embedding procedure consists of two sub-procedures including spatial domain embedding and frequency domain embedding. We embed a large amount of the message in the spatial domain using the MED predictor. Then, we employ the frequency domain embedding procedure to embed more messages which increase the embedding capacity. During the message embedding in the frequency domain, we preserve the image quality by using an integer-to-integer transform. We predefine that the original image O, where each pixel $O(i, j)$ is referred to by two indexes i and j.

A. Spatial Domain Embedding

We can then embed the secret message SM_s into the O using our proposed embedding technique in the spatial domain. Here, the subscript "s" indicates that the secret message is embedded in the spatial domain. Similarly, we will use the subscript "f" to represent the secret message SM_f to be embedded in the frequency domain. We use the MED predictor to generate error values and then modify the error values in order to embed the secret messages. We describe the embedding procedure as follows:

Given the pixel $O(i, j)$, we employ the MED predictor, shown in (1), to generate the predictive values $PI(i, j)$.

$$PI(i,j) = \begin{cases} \min[O(i, j-1), O(i-1, j)] & \text{if } O(i, j) \geq \max[O(i, j-1), O(i-1, j)], \\ \max[O(i, j-1), O(i-1, j)] & \text{if } O(i, j) \leq \min[O(i, j-1), O(i-1, j)], \\ O(i, j-1) + O(i-1, j) - O(i, j) & \text{otherwise.} \end{cases} \quad (1)$$

We then produce the error sets ES by calculating the difference between O and PI, as shown in (2), where $ES(i, j)$ represents the error value in the i-th row and j-th column.

$$ES(i, j) = O(i, j) - PI(i, j). \quad (2)$$

We derive the absolute value of $ES(i, j)$ using (3), and denote it as $AES(i, j)$.

$$AES(i, j) = |ES(i, j)|. \quad (3)$$

We can determine the maximal appeared frequency value p_s from AES. Given this p_s, we can modify the error sets ES to embed the secret message SMs. (4) describes the rule of the modification, where $sms \in \{0, 1\}$ denotes the secret bits "0" or "1," and ES represents the error sets that have conveyed the secret message.

$$ES'(i,j) = \begin{cases} ES(i, j) + sms & \text{if } AES(i, j) = p_s \text{ and } ES(i, j) \geq 0, \\ ES(i, j) - sms & \text{if } AES(i, j) = p_s \text{ and } ES(i, j) < 0, \\ ES(i, j) + 1 & \text{if } AES(i, j) > p_s \text{ and } ES(i, j) \geq 0, \\ ES(i, j) - 1 & \text{if } AES(i, j) > p_s \text{ and } ES(i, j) < 0, \\ ES(i, j) & \text{otherwise,} \end{cases} \quad (4)$$

where $ES'(i, j)$ is the adjusted error value of ES.

Given the modified error value $ES'(i, j)$ and the $PI(i, j)$ pixel in the predictive image PI, we can use (5) to generate the corresponding pixel $TI(i, j)$ in the temporary image TI. This temporary image will serve as the input for message embedding in the frequency domain.

$$TI(i, j) = PI(i, j) + ES'(i, j). \quad (5)$$

B. Frequency Domain Embedding

The integer-to-integer wavelet transform is employed to transform the temporary image TI into the frequency domain. By modifying the wavelet coefficients of the transformed TI, we embed the secret message SM_f. There are two benefits to using the integer-to-integer wavelet transform for reversible data hiding. First, the integer-to-integer wavelet transform is operated in the integer domain, and causes no truncation problems for the inverse transform. This implies that the transform provides reversibility achieving the objective of reversible data hiding. Second, the wavelet transform can map the original image into four sub-bands representing different degrees of features in the original image. By embedding the secret message into the middle frequency and high frequency sub-bands, our algorithm can faithfully preserve the features in the original image and reduce the risks of attention by eavesdroppers. The frequency domain embedding procedure is described as follows:

The temporary image TI is translated into the wavelet coefficient set CS by forwarding the two-dimensional integer-to-integer transform. This produces four sub-bands, LL, LH, HL, and HH. It is obvious that changing the low frequency sub-band affects the visual quality of the image. Consequently, our scheme embeds information into the middle frequency and high frequency sub-bands, which are the less crucial sub-bands (LH, HL and HH), and we keep

the low frequency sub-band (LL) unchanged. For a wavelet coefficient $CS(i, j)$ that belongs to one of the three less crucial sub-bands, we can compute the absolute wavelet coefficient $ACS(i, j)$ by (6). We can further find the maximum appeared frequency value p_f for these absolute coefficients.

$$ACS(i, j) = |CS(i, j)|. \qquad (6)$$

In order to embed the secret message SM_f we modify the rule in the spatial domain to become (7), where $CS'(i, j)$ represents the embedded wavelet coefficients, and $smf \in \{0,1\}$ is the combined secret bit in the secret message SM_f The $ACS(i, j)$ and $CS(i, j)$ belong to the three less crucial sub-bands, LH, HL and HH.

$$CS'(i, j) = \begin{cases} CS(i, j) + smf & \text{if } AES(i, j) = p_f \text{ and } CS(i, j) \geq 0, \\ CS(i, j) - smf & \text{if } AES(i, j) = p_f \text{ and } CS(i, j) < 0, \\ CS(i, j) + 1 & \text{if } AES(i, j) > p_f \text{ and } CS(i, j) \geq 0, \\ CS(i, j) - 1 & \text{if } AES(i, j) > p_f \text{ and } CS(i, j) < 0, \\ CS(i, j) & \text{otherwise.} \end{cases} \qquad (7)$$

Then, we can apply the inverse integer-to-integer wavelet transform to translate the modified wavelet coefficients CS' into the embedded image E.

C. The Extraction Procedure

In the extraction procedure, our inverse processing to extract the secret message and recover the original image. The extraction process begins in the frequency domain followed by the spatial domain. In particular, the embedded image E is translated into embedded wavelet coefficients CS' by the forward integer-to-integer transform. We extract the secret message SM_f. We further recover the original wavelet coefficients CS using (9).

$$csm = \begin{cases} 0 & \text{if } ACS'(i, j) = p_f, \\ 1 & \text{if } ACS'(i, j) = p_f + 1. \end{cases} \qquad (8)$$

$$CS(i, j) = \begin{cases} CS'(i, j) - 1 & \text{if } ACS'(i, j) > p_f \text{ and } CS'(i, j) > 0, \\ CS'(i, j) + 1 & \text{if } ACS'(i, j) > p_f \text{ and } CS'(i, j) < 0, \\ CS'(i, j) & \text{otherwise.} \end{cases} \qquad (9)$$

Then, we employ the inverse integer-to-integer wavelet transform for the original wavelet coefficients CS to produce the temporary image TI. This ends the extraction process in the frequency domain where the temporary image TI

becomes the input in the spatial domain extracting procedure. Given the temporary image TI in the spatial domain extracting procedure, the secret message SM_s can be extracted by referring to the value p_s and the MED predictor. The predictive pixel $PI(i, j)$ is calculated from TI by the MED predictor by using (1), and then we compute the $ES'(i, j)$ using (10).

$$ES'(i, j) = TI(i, j) - PI(i, j). \qquad (10)$$

For each $ES'(i, j)$, the embedded secret bit sms and the original image O can be obtained by (11) and (12), respectively.

$$sms = \begin{cases} 0 & \text{if } AES'(i, j) = p_s, \\ 1 & \text{if } AES'(i, j) = p_s + 1. \end{cases} \qquad (11)$$

$$O(i, j) = \begin{cases} PI(i, j) + ES'(i, j) - 1 & \text{if } AES'(i, j) > p_s \text{ and } ES'(i, j) > 0, \\ PI(i, j) + ES'(i, j) + 1 & \text{if } AES'(i, j) > p_s \text{ and } ES'(i, j) < 0, \\ PI(i, j) + ES'(i, j) & \text{otherwise.} \end{cases} \qquad (12)$$

IV. EXPERIMENTAL RESULTS

Experimental results to show the performance of our algorithm in terms of the payload capacity and the image quality of the embedded image. We employ three images as our test images, each of which has a resolution of 512×512 pixels. We adopt the binary random sequences as the secret messages. In addition, we use the peak signal to noise ratio (PSNR) to evaluate the image quality of the embedded image. When deriving our results, we report the magnitude of the pure capacity in bits per pixel (bpp). Fig. 1 shows the payload capacity when the PSNR value is approximate to 30 dB. Observing these images, we find that the average payload capacity for Fig. 1(a)-(c) was 1.173 bpp, but the capacity for the "Baboon" image shown in Fig. 1(c) is smaller than 0.724 bpp. Since the "Baboon" image contains wayward textures, the distribution of the absolute error value generated by the MED predictor will not focus on specific values. Because our scheme offers the reversibility, the payload capacity can be accumulated by applying the embedding more than one time, which is referred to as multiple layer embedding. Table 1 shows the pure payload capacity in the dual domain at each layer of the embedding. For each test image, the pure payload capacity decreases gradually as the layers increase. These statistics demonstrate that the pure payload capacity offered in the spatial domain is larger than that provided in the frequency domain.

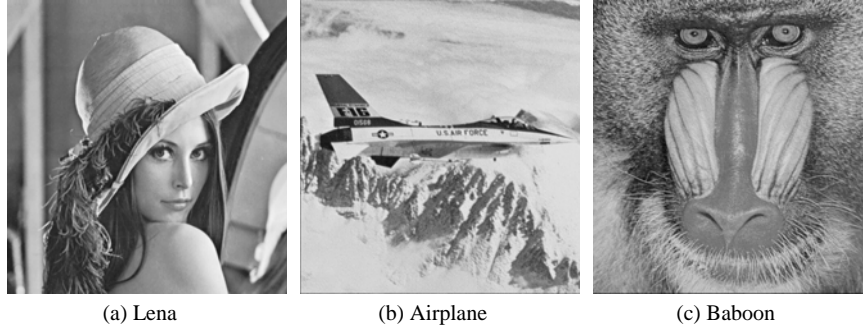|	(a) Lena	(b) Airplane	(c) Baboon

Figure 1.   Embedded images and their accumulated pure payload capacities when the PSNR value of the embedded image is approximate to 30 dB.

TABLE I.   THE PURE PAYLOAD CAPACITIES (BPP) IN THE SPATIAL AND FREQUENCY DOMAIN

Layer	Lena		Airplane		Baboon	
	Spatial	Freq.	Spatial	Freq.	Spatial	Freq.
1	0.327	0.176	0.336	0.197	0.097	0.067
2	0.194	0.096	0.218	0.113	0.083	0.052
3	0.120	0.080	0.126	0.079	0.065	0.044
4	0.102	0.063	0.098	0.062	0.060	0.035
5	0.079	0.049	0.075	0.050	0.048	0.034
6	0.067	0.045	0.064	0.044	0.045	0.030
Total	0.889	0.509	0.917	0.545	0.398	0.262
Percent	64%	36%	63%	37%	60%	40%

V.   CONCLUSIONS

In this paper, we propose a high embedding capacity reversible data hiding algorithm. Our algorithm takes advantage of both a spatial domain and a frequency domain. The algorithm not only embeds a large quantity of secret messages in the spatial domain, but also explores the frequency domain in order to increase capacity, while still preserving features of the images in the frequency domain. Experimental results show that when PSNR values of the embedded image are close to 30 dB, our algorithm provides the high payload capacity. It provides a large embedding capacity with good visual quality for the embedded image.

REFERENCES

[1] M. D. Adams, and R. K. Ward, "Symmetric-extension-compatible reversible integer-to-integer wavelet transforms," IEEE Transactions on Signal Processing, vol. 51, pp. 2624-2636, 2003.

[2] D. Coltuc, and J. M. Chassery, "Very fast watermarking by reversible contrast mapping," IEEE Signal Processing Letters, vol. 14, pp. 255-258, 2007.

[3] D. M. Thodi, and J. J. Rodriguez, "Expansion embedding techniques for reversible watermarking," IEEE Transactions on Image Processing, vol. 16, pp. 721-730, 2007.

[4] S. Lee, C. D. Yoo, and T. Kalker, "Reversible image watermarking based on integer-to-integer wavelet transform," IEEE Transactions on Information Forensics and Security, vol. 2, pp. 321-330, 2007.

[5] J. Tian, "Reversible data embedding using a difference expansion," IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, pp. 890-896, 2003.

[6] Y. Y. Tsai, D. S. Tsai, and C. L. Liu, "Reversible data hiding scheme based on neighboring pixel differences," Digital Signal Processing, vol. 23, pp. 919–927, 2013.

[7] H. W. Tseng, and W. B Ding, "Reversible data hiding scheme for colour images based on pixel clustering and histogram shifting," Imaging Science Journal, vol. 60, pp. 47-53, 2012.

[8] S. 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.

[9] A. R. Calderbank, I. Daubechies, W. Sweldens, and B. L. Yeo, "Wavelet transforms that map integers to integers," Applied Computational Harmonics Analysis, vol. 5, pp.332-369, 1998.