High Embedding Reversible Data Hiding Scheme for JPEG

Wen-Chung Kuo
Depart. of Computer Science and Information Engineering,
National Formosa University,
Yunlin 632, Taiwan
simonkuo@nfu.edu.tw

Shao-Hung Kuo, Lih-Chyau Wuu
Graduate School of Engineering Science and Technology Doctoral Program,
National YunLin University of Science and Technology,
Yunlin 640, Taiwan
{g9810814, wuulc}@yuntech.edu.tw

Abstract—Both Data hiding and data compression are very important technologies in the field of image processing. It seems that there is no relationship between data hiding and data compression because most of the data hiding mechanisms focus on information security and the data compression mechanisms emphasize the compression ratio. In fact, they are closely related. Until now, there are many literatures to discuss the secret data how to be embedded into the media image. Unfortunately, the stego-image’s size will be increased when the media image is hidden a lot of secret data. In order to overcome this disadvantage, a reversible data hiding scheme based on EMD data hiding and JPEG compression technology will be proposed in this paper. According to the experimental results, we can prove that our proposed scheme still keeps high embedding capacity, security and good compression ratio.

Keywords: EMD, JPEG compression, Cover image, Stego-image.

I. INTRODUCTION

Digital information is copied or tampered to spread easily by malicious people. It brings about many risks during information transmission. In order to deliver the cover image which secret information is embedded in to the receiver safely and extract secret information completely, there are several major requirements for data hiding: security, capacity, robustness, imperceptibility, unambiguousness and nonremovable. Many methods achieve above requirements and the most common data hiding method is least significant bit replacement [3], “LSB replacement” for short. Another method is least significant bit matching [4], “LSB matching” for short. Both methods are simple and effective. However, we adopt a general high embedding capacity EMD data hiding method [5] which is improved by a high embedding capacity EMD data hiding method [6] based on security, capacity and imperceptibility. The data hiding method based on EMD was proposed by Zhang et al.[8] in 2006, and the general high embedding capacity data hiding method based on improving EMD was proposed by Kuo et al.[5] in 2009.

Most accessible digital information are digital images like digital photos or digital videos due to popularity of digital cameras. The file size of digital images it’s increase as the requests of high quality increase. We must use some efficient image processing technology to store a large number of images to avoid wasting time and space. The most common approach is that store images with compression format. Two famous international standard compression format are JPEG (Joint Picture Experts Group)[1] and JPEG2000[2] compression technology. In 2007, Seki et al. proposed a random location data hiding method based on JPEG[7]. It hides data in random designate locations based on Discrete Cosine Transform (DCT) and addition and subtraction of remainder quantification. Because of random designate hiding locations it is irreversible. We improve this drawback in this paper. We propose a high capacity and good security reversible data hiding method by combining the general high embedding capacity EMD data hiding method proposed by Kuo et al.[5].

The rest of this paper is organized as follows. Section II reviews the related method including the JPEG compression technology, the random location data hiding method based on JPEG and the data hiding method based on EMD. The proposed scheme and experimental results are given in Section III and Section IV, respectively. Finally, conclusions are given in Section V.

II. RELATE WORK

As we known, JPEG compression is a lossy compression technique, the image will be made perpetual distortion. In other words, after compression and decompression procedure, the original image will produce the distortion and reduce the image quality. However, it can bring high compression ratio to improve data storage and accelerate the transmission speed. The JPEG encoding method can be explained as following steps and the flow chart shown as Fig.1.

Step 1) Image is broken down into 8 × 8 blocks.
Step 2) Each block is put through the DCT (Discrete Cosine Transform) transform, which DCT transform as Eq.(1).

\[
F(u, v) = \frac{C(u)C(v)}{4} \sum_{j=0}^{7} \sum_{k=0}^{7} f(j, k) \cos \left( \frac{(2j+1)u\pi}{16} \right) \cos \left( \frac{(2k+1)v\pi}{16} \right)
\]

(1)

Where \( F(0, 0) \) is Direct Current, the other is Alternating Current. The transformed coordinate was \( F(u, v) \), and \( f(j, k) \) is the original coordinate.

\[
\begin{align*}
C(u) & \text{ or } C(v) = \sqrt{2}^{(-1)} \quad u \text{ or } v = 0, \\
C(u) & \text{ or } C(v) = 1 \quad u \text{ or } v \neq 0.
\end{align*}
\]

(2)
Step 3) The DCT coefficients are turned into integers by using the quantization table (shown as Fig.2).

Step 4) Use the zigzag scan method to scan AC coefficient and the symbol EOB represent the AC coefficient values are zero behinds it in this block (shown as Fig.3).

Step 5) Use Huffman code to complete encoding, whose in order all of block to do JPEG compression.

Decoding steps as long as the decoding stage according to Fig.1, it can get the compressed images.

A. A Data Hiding Method without Specifying Embedded Positions for JPEG Image)

2007, Seki et al.[7] (SKFK-scheme for short) proposed a data hiding method to embed the secret data into a JPEG image. The major contribution of their scheme does not need to point out the suitable location to embed the secret data. In other words, the secret data embedded position is randomly selected between the DC value and EOB. The implementation embedding steps as follows and flow chart are shown as Fig.4:

Step 1) Calculate the sum S for the DCT coefficients after quantization of the block.

Step 2) Compute the remainder , where N is the maximum embedding value.

Step 3) Calculate the difference g between r and the secret data, i.e.,

Step 4) Compute and choose .

Step 5) The embedded sum can be obtained by using Eq(3).

\[
S' = \begin{cases} 
S + d & \text{if } g < g' \\
S - d & \text{if } g \geq g'
\end{cases}
\]

B. The Improve High Capacity EMD Hiding Technique

Recently, Kuo et al.[5] proposed the modified data hiding scheme to enhance the security of LWC-scheme[6]. The main ideal of improved version of the high capacity EMD scheme is through the look-up table can quickly to get changed pixel from hidden image, and also can quickly to get secret information from the extraction function. However, the drawback of this proposed scheme is the need for additional storage of eight tables. In order to extract the secret message successfully, these additional tables must be transmitted to the receiving by the secret channel. Hence, it will consume many spaces for storage and many bandwidths, respectively. Kuo et al. modify these additional eight tables condensed into the two characteristics tables shown in Fig.5, respectively. The detail the relationship between (α, β) and seed is also shown in the Table I, in other words, each seed
also has their corresponding weighting value \((a, b)\), and \((\alpha, \beta)\). Finally, the relationships between all the corresponding values are shown in Table II. And Kuo et al. introduce a strategy that improves the safety of high capacity EMD hiding. The extraction function \(f_s(X,Y)\) is similar to that of the LWC-scheme shown as Eq.(4).

\[
f_s(X,Y) = (X \times a + Y \times b) \mod 8
\]  

Figure 5. (a)The seed from 0 to 3. (b)The seed from 4 to 7.

TABLE I. RELATIONSHIP OF GENERALIZED EQUATION

<table>
<thead>
<tr>
<th>Seed : 0~3</th>
<th>Seed : 4~7</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f_s(X,Y))</td>
<td>(X', Y')</td>
</tr>
<tr>
<td>(f_s(X,Y)=0)</td>
<td>(X' = X)</td>
</tr>
<tr>
<td>(f_s(X,Y)=1)</td>
<td>(X' = X+\alpha)</td>
</tr>
<tr>
<td>(f_s(X,Y)=2)</td>
<td>(X' = X-\alpha)</td>
</tr>
<tr>
<td>(f_s(X,Y)=3)</td>
<td>(X' = X+\alpha)</td>
</tr>
<tr>
<td>(f_s(X,Y)=4)</td>
<td>(X' = X-\alpha)</td>
</tr>
<tr>
<td>(f_s(X,Y)=5)</td>
<td>(X' = X)</td>
</tr>
<tr>
<td>(f_s(X,Y)=6)</td>
<td>(X' = X+\alpha)</td>
</tr>
<tr>
<td>(f_s(X,Y)=7)</td>
<td>(X' = X-\alpha)</td>
</tr>
</tbody>
</table>

TABLE II. RELATIONSHIP BETWEEN \(a\) AND \(b\)

<table>
<thead>
<tr>
<th>Seed</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>((a,b))</td>
<td>((-1,3))</td>
<td>((1,3))</td>
<td>((-1,7))</td>
<td>((1,7))</td>
<td>((-1,1))</td>
<td>((1,1))</td>
<td>((-1,5))</td>
<td>((1,5))</td>
</tr>
<tr>
<td>(\beta)</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>

Consequently, this scheme uses the seeds to decide the weighting value. A random weighting value is used to check the modulo table to determine which way KWSK-scheme should use to change the pixel pair \((X, Y)\) in the cover image. About KWSK-scheme embedding flow chart and 8 kinds of adjustment table, are shown in Fig.6.

From Fig.5, EMD picks a pair of pixels, gets the remainder by passing the pair through modulo, and finds the value in the surrounding area. About modulo table, all of them are use 0 as the center, then the different weighting value generalized function including \(\alpha\) and \(\beta\) is designed to match the Fig.5 (a), (b) condition.

III. THE PROPOSED METHOD

Seki et al.[7] proposed by using random location to hidden the secret, so their technique had high degree of freedom. That is also a disadvantage for their proposed which is a non-reversible data hiding technique. Here, we will explain a method for hiding and extracting information in a JPEG image using the high capacity EMD hiding technique, and is a novel scheme with reversible data hiding technique. The major difference point between our proposed method and Seki et al. is that our proposed scheme is not care about the biggest hiding value of \(N\). Finally, this study will be focused on security, hiding capacity and PSNR analysis of three aspects to make a detailed comparison in Section IV.

The proposed method is also a selection of two blocks, and take two blocks of the DC value is used as a general formula of high-hiding capacity of EMD extraction function (Eq.4) of the input values \((X, Y)\). After calculate to got value \(fs\), there will be decision EOB how many space needs to fall back. Thus this method can hide a block of time from 1 to 8 bits. The hiding data procedure as following steps:

Step 1) Select two sequence quantized blocks, and then choose the DC value as input values \((X, Y)\).

Step 2) Using Eq.(4) and \((X, Y)\) to calculate value \(fs\) and weights \((a, b)\) is determined for the Seed.

Step 3) The EOB will be fall back \(fs+1\) space.

Step 4) The secret bits will be hiding in the EOB of fall back.

The details of proposed extracting method are described as following steps:

Step 1) Compute the \(fs\) by using Eq.(3) the two blocks of the DC values.

Step 2) Forward counting the \(fs+1\) space from EOB to extract the secret data.

Step 3) Find out the original of the EOB position.

IV. EXPERIMENTAL RESULTS

The proposed scheme was tested on the eight 512×512 gray images (Airplane, Baboon, Boat, Elaine, Gold Hill, Lena, Pepper, and Tiffany) shown in Fig.7. The
corresponding stego images of [7] and our proposed scheme are shown in Fig.8 and Fig.9, respectively. Research analysis and comparison of experimental results are included in Table III. According to the Table III results, we can find out three major contributions in our proposed scheme. Firstly, from Table III, the hiding secret information's capacity of proposed scheme is 6,000 bits more than Seki et al.[7]. In other words, there is a high-hiding capacity property in our proposed scheme. Secondly, from the security view, the embedding function can’t be recovered by using the different weights in the extraction function. Therefore, our proposed scheme is more secure than [7].

<table>
<thead>
<tr>
<th>File size used JPEG compressed (KB)</th>
<th>Seki[7]</th>
<th>Our proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airplane</td>
<td>31.6</td>
<td>35.0</td>
</tr>
<tr>
<td>Baboon</td>
<td>60.2</td>
<td>66.1</td>
</tr>
<tr>
<td>Boat</td>
<td>36.8</td>
<td>40.4</td>
</tr>
</tbody>
</table>

V. CONCLUSION

In this paper, a new reversible data compression hiding scheme based on JPEG and formal improvement high-hiding capacity of EMD [5] has been proposed. According to the experimental results, our proposed scheme is not only to achieve reversible data hiding techniques but also to maintain the good compressed image quality.

ACKNOWLEDGMENT

This work was supported by the National Science Council of Taiwan under grant NSC 98-2219-E-150-001.

REFERENCES