**Image Steganography Based on Modified LSB Substitution Method and Data Mapping**

Pranab Kumar Dhar†*, Abdul Kaium†, and Tetsuya Shimamura††

†Department of Computer Science and Engineering, Chittagong University of Engineering and Technology, Chittagong-4349, Bangladesh
††Graduate School of Science and Engineering, Saitama University, Saitama 338-8570, Japan

*Corresponding author

Summary
This paper introduces an image steganography method based on modified least significant bit (LSB) substitution technique and data mapping. Initially, the secret image is pre-processed using a new data mapping scheme to reduce the number of 1’s in secret data. The secret data is then embedded into the host image using a modified LSB substitution technique. In this technique, each pixel value of host image is changed if value of secret bit is 1 otherwise the LSB of each pixel value will remain unchanged which is different from conventional LSB technique. This improves the visual quality of the stego image by changing the less number of bits in image pixel. Experimental results indicate that the proposed method provides high quality stego image. Moreover, it outperforms state-of-the-art methods in terms of peak signal-to-noise ratio (PSNR), mean square error (MSE), and structural similarity (SSIM).

Key words: Image steganography, Least Significant Bit, Data mapping, Peak signal-to-noise ratio, structural similarity.

1. Introduction

Internet has become an essential part of our life due to the advancement in the field of information and communication technology. Hence, it is important to provide sufficient security measure of the transmitted data over the internet considering the fact that the data transmitted via internet consists of not only casual data but also may include confidential information like medical, financial and military data. It is obvious that unwanted access of this data poses a constant threat to the confidentiality of such sensitive information. Thus it is important to transmit sensitive data by a means whose contents can only be revealed by the intended recipient. Steganography is one of the data hiding techniques which solves this problem. It can be used for a wide range of applications such as, in defense organizations for safe circulation of secret data, in military and intelligence agencies, in smart identity cards where personal details are embedded in the photograph for copyright control of materials, in medical imaging where patient’s details are embedded within image providing protection of information and reducing transmission time and cost, in an online voting system for making the online election secure and robust against a variety of fraudulent behaviors, in improving mobile banking security, in tamper proofing to prevent or detect unauthorized modifications and other numerous applications. Steganography comes from the Greek words Steganós (Covered) and Grapto (Writing)[1]. Steganography refers to information or a file that has been concealed inside a digital picture, video or audio file. The goal of modern steganography is to keep its information undetectable. Many steganographic methods have been introduced in literature [2]-[5]. The least significant bits (LSB) substitution technique is widely used for image steganography [6]-[7]. The main limitation of this method is that it has poor visual quality and it is less secured. Authors of [8] proposed a variable rate image steganography method in which some pixels are highly modified but other will remain unchanged so that attacker can easily detect if there is any hidden information. In [9], a DCT-based image steganography method was introduced which does not provide good quality stego image and it gives information regarding the presence of message in cover image. Another LSB based image steganography method is presented in [10]. In this method, a fixed number of bits are embedded in smooth areas and a variable number of bits are embedded into the edged areas. LSB replacement (LSBR) is proposed in [11], where only the LSB plane of the cover image is replaced with the secret image according to a pseudorandom number generator. As a result, some structural asymmetry is introduced and thus it is very easy to detect the existence of hidden message. Another method is proposed in [12] based on PDV, in which cover image area are classified into two types: one is smoother region and other one is hard regions. The smoother pixels are not suitable for data embedding. Therefore, the pixels with large differences and their neighbor pixels are used for data embedding. This method provides good embedding capacity but stego images are statically detectible one. Wang et al. [13] proposed a method to embed data by using genetic algorithm to improve the quality of the stego image. However, genetic
algorithm takes more computational time. To overcome these limitations, in this paper we propose a modified LSB technique that embeds less number of secret bits in cover image with high capacity. In this technique, modification in LSB of pixel value is done if the value of secret bit is 1 otherwise, the pixel value will remain unchanged which is different from LSB substituting technique in which every pixel value irrespective of 0 or 1 will be changed [1]. Therefore, the number of change in cover image depends on number of 1’s in secret bit. We use secret key for selection of pixels so that any one can not detect the message sequence without key. In addition, we also introduce data mapping for better performance. Mapping of secret data is done to minimize the number of 1’s and maximize the number of 0’s in secret bit stream so that changes in stego image will be minimum.

The rest of this paper is organized as follows. Section 2 presents the image steganography method including embedding procedure and detection procedure of hidden information. Section 3 shows the performance of the proposed method with some recent methods in terms of PSNR (peak signal-to-noise ratio), MSE (mean square error), and SSIM (structural similarity). Lastly, the conclusion of this paper is discussed in Section 4.

2. Proposed Steganography Method

Let \( A = \{a(i,j), 1 \leq i \leq M, 1 \leq j \leq M\} \) be a host image and \( W = \{w(i,j), 1 \leq i \leq N, 1 \leq j \leq N\} \) be a secret image to be embedded into the host image.

2.1 Embedding Procedure of Hidden Information

Step 1: A 24-bit color scheme uses 24 bits per pixel and each byte represents the intensity of the three primary colors red, green, and blue (RGB), respectively. Therefore, each pixel of the secret image is divided into three matrices each having 8 bit binary values as shown in Fig. 1. This 2D array is reshaped into a 1D array Data Stream (DS). This 1D array matrix is also called bit stream of hidden information. The length of bit stream DS is \( N \times N \times 24 \). The cover image is also divided into three colors red, green and blue (RGB), respectively. Total number of pixel is \( M \times M \times 3 \). Each pixel is numbered in the range 1 to \( (M \times M \times 3) \) uniquely using secret key random generator.

Step 2: The 1D bit stream DS of the secret message is divided into 4-bit groups for mapping purpose. Mapping is done in several iterations with maximum possibility of 16 iterations. This mapping sequence generated by considering number of one in 4-bit group. If the 4-bit group that occurred maximum replaced by minimum number of one 4-bit group, then the total number of one in mapped message will be reduced. This is the purpose of using mapping in our proposed method.

In each iteration, the maximum occurrence of 4-bit groups are searched and replaced by the sequence according to Table 1.

![Fig.1 Binary representation of RGB matrix of secret image](image)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000</td>
<td>0001</td>
<td>0010</td>
<td>0100</td>
<td>1000</td>
<td>0011</td>
<td>0101</td>
<td>0110</td>
</tr>
<tr>
<td>1001</td>
<td>1010</td>
<td>1100</td>
<td>1101</td>
<td>1110</td>
<td>1111</td>
<td>1112</td>
<td>1113</td>
</tr>
</tbody>
</table>

Consider 1D bit stream of the secret message is

\[
1110,1101,0110,1111,0101,1101,1110,1111
\]

In first iteration, maximum occurred 4-bit group is 1101. Thus, it will change by 0000 as mentioned earlier. Now 1D bit stream

\[
1110,0000,0110,1111,0101,0000,1110,0000
\]

After several iterations, 1D bit stream will be

\[
0001,0000,0010,0100,1000,0000,0001,0000
\]

This mapped data is embedded in the cover image in the subsequent steps. The mapped table is generates as:

- 1101 replace by 0000
- 1110 replace by 0001
- 0110 replace by 0010
- 1111 replace by 0100
- 0101 replace by 1000 etc.
We observed that before mapping the number of one in the bit stream is 23 and after mapping the number of one in the bit stream is just 5. This process reduces the number of 1’s resulting in high quality stego image.

Let CP be the cover image pixel and SP be the stego image pixel.

Step 2: Calculate the difference of pixel value of each SP and CP. The secret message bit is retrieved using the following rule:
If SP=CP, then the secret bit is 0, otherwise the secret bit is 1.

Step 3: The 1D secret bit stream is obtained by utilizing the Table 1.

Step 4: The secret image is obtained by converting the 1D secret bit stream into 2D matrix.

Fig. 2  Embedding procedure of hidden information

Step 3: Select pixels sequentially from the cover image where secret data is to be embedded.

Step 4: Each secret bit is checked sequentially. LSB of randomly located pixel will be modified using the following rule:
(i) If secret bit is 1 then;
   pixel value of cover image is checked,
   if the pixel value < 128;
   increase the pixel value by x
   if pixel overflows (value>=128),
   decrease pixel value by x.

Step 5: After embedding the secret data, the stego image A’ is found.

2.2 Recovery Technique of Hidden Information

Step 1: Select the stego image A’. Using the random key the sequence of the pixels of the stego image is identified.

Fig. 3  Recovery procedure of hidden information

3. Experimental Results and Analysis

In this section, several experiments were carried out to demonstrate the performance of our proposed method. We used some standard RGB images (Lena, House, Baboon, Peppers, Airplane and Sailboat on lake) of size 512×512 as the host image shown in Fig. 5. IEEE logo image of size 64×64 is used as secret image shown in Fig. 6. Therefore, the size of secret bit stream is (64×64×3×8) = 98304 bit.

The hidden information used in our proposed method is shown below:
After embedding secret data, the stego images are shown in Fig. 6. We observed that original images and stego images are quite similar. In other word, the proposed method provides high quality stego image.

To measure the imperceptibility of the proposed algorithm, (i) PSNR (ii) MSE and (iii) structural similarity (SSIM) between the original host image and stego images are used. These terms are defined as follows:

$$PSNR = 10 \log_{10} \left( \frac{255^2}{MSE} \right)$$  \hspace{1cm} (1)

$$MSE = \frac{1}{MM} \sum_{i=1}^{M} \sum_{j=1}^{M} (A - A')^2$$  \hspace{1cm} (2)

$$SSIM(A,A') = \frac{(2\mu_A\mu_{A'} + c_1)(2\sigma_{AA'} + c_2)}{\mu_A^2 + \mu_{A'}^2 + \sigma_{AA'}^2 + c_1}$$  \hspace{1cm} (3)

Here, $\mu_A, \sigma_A, \mu_{A'}, \sigma_{A'}$, and $\sigma_{AA'}$ indicate the mean of $A$, variance of $A$, mean of $A'$, variance of $A'$, covariance of $A$ and $A'$, respectively. $c_1$ and $c_2$ are the two variables used in Eq. (2) and (3).

In this study, we have compared our proposed method with some recent methods in terms of PSNR, MSE, and SSIM as shown in Table 2. From this Table, we observed that the PSNR, MSE, and SSIM of the proposed method range from 42.37 to 48.41, 0.938 to 1.510, and 0.835 to 0.899, respectively. On the other hand, the PSNR, MSE, and SSIM of the existing methods range from 35.47 to 44.67, 2.219 to 18.45, and 0.721 to 0.956, respectively. The payload of the proposed method is 3.0, whereas the payload of the existing methods ranges 1.45 to 3.12. The existing methods [2], [10], [12] embed more pixels of a cover image in hiding information with same capacity whereas our proposed method embeds a smaller number of bits by utilizing modified LSB substitution technique. As we have to change less number of bits of the cover image, for this reason, the resulting PSNR value of our proposed method is better than other methods. This is done by using the proposed mapping scheme. In Distortion technique, the bit stream is changed when a 1 is occurred. By reducing the number of 1’s also reduces the number of changes required. On the other hand, Chan-Cheng’s and Four Neighbor method modifies the bit stream according to LSB where it is necessary to change the bit stream irrespective whether the bit is 0 or 1. Overall, we can conclude that the proposed method provides better result than the existing methods in terms of PSNR, MSE, SSIM, and payload. In other words, the proposed method shows an effective way to hide secret information into the host image while providing high quality stego image.

4. Conclusion

In this paper, we have presented an image steganography method based on modified LSB substitution technique and data mapping. Experimental results indicate that the proposed method effectively embed hidden information into host image while maintaining sufficient visual quality.
of stego image. Moreover, it shows superior performance than the state-of-the-art methods in terms of PSNR, MSE, and SSIM because of the modified LSB substitution method and data mapping. In future, we will extend our research to develop a better way of data mapping to improve the trade-off among storage, complexity and performance.

Table 1: Comparison between the proposed and several recent methods

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>PSNR</td>
<td>38.45</td>
<td>41.23</td>
<td>43.24</td>
<td>44.5</td>
<td>48.4</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>9.29</td>
<td>4.898</td>
<td>3.08</td>
<td>2.27</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.73</td>
<td>0.844</td>
<td>0.844</td>
<td>0.86</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>payload</td>
<td>3.0</td>
<td>1.59</td>
<td>3.016</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>House</td>
<td>PSNR</td>
<td>37.65</td>
<td>40.51</td>
<td>40.65</td>
<td>42.3</td>
<td>46.3</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>11.17</td>
<td>5.782</td>
<td>5.60</td>
<td>3.76</td>
<td>1.51</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.72</td>
<td>0.797</td>
<td>0.810</td>
<td>0.87</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>payload</td>
<td>3.0</td>
<td>1.48</td>
<td>3.12</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Baboon</td>
<td>PSNR</td>
<td>35.47</td>
<td>38.43</td>
<td>41.89</td>
<td>43.3</td>
<td>48.3</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>18.45</td>
<td>9.334</td>
<td>4.21</td>
<td>2.37</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.82</td>
<td>0.751</td>
<td>0.816</td>
<td>0.84</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>payload</td>
<td>3.0</td>
<td>1.50</td>
<td>3.10</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Peppers</td>
<td>PSNR</td>
<td>39.38</td>
<td>36.67</td>
<td>41.56</td>
<td>44.6</td>
<td>48.3</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>7.50</td>
<td>13.99</td>
<td>4.45</td>
<td>2.21</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.79</td>
<td>0.754</td>
<td>0.799</td>
<td>0.86</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>payload</td>
<td>3.0</td>
<td>1.44</td>
<td>3.01</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Airplane</td>
<td>PSNR</td>
<td>36.80</td>
<td>37.49</td>
<td>40.67</td>
<td>44.2</td>
<td>47.5</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>13.58</td>
<td>11.59</td>
<td>5.572</td>
<td>2.45</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.77</td>
<td>0.751</td>
<td>0.812</td>
<td>0.85</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>payload</td>
<td>3.0</td>
<td>1.36</td>
<td>3.01</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Sailboat</td>
<td>PSNR</td>
<td>35.72</td>
<td>37.23</td>
<td>41.45</td>
<td>44.2</td>
<td>47.1</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>17.42</td>
<td>12.3</td>
<td>4.66</td>
<td>2.43</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.79</td>
<td>0.801</td>
<td>0.719</td>
<td>0.81</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>payload</td>
<td>3.0</td>
<td>1.45</td>
<td>3.02</td>
<td>3.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>

References


Abdul Kaium received the Bachelor of Science (B.Sc.) Degree in Computer Science and Engineering from Chittagong University of Engineering and Technology (CUET), Chittagong, Bangladesh in 2016. His research interest includes Digital Watermarking, Multimedia Systems, and Digital System Design.

Tetsuya Shimamura received the B.E, M.E., and Ph.D Degrees in Electrical Engineering from Keio University, Yokohama, Japan, in 1986, 1988, and 1991, respectively. In 1991, he joined Saitama University, Saitama, Japan, where he is currently a Professor. He was a visiting researcher at Loughborough University, U.K. in 1995 and at Queen’s University of Belfast, U.K. in 1996, respectively. Prof. Shimamura is an author and coauthor of 6 books. He serves as an editorial member of several international journals and is a member of the organizing and program committees of various international conferences. His research interests are in digital signal processing and its application to speech, image, and communication systems.