An Improved Data Hiding Technique Using Bit Differencing and LSB Matching

Aditya Kumar Sahu and Gandharba Swain

Abstract—This paper proposes an improved image steganographic technique based on the principle of modified least significant bit (LSB) substitution and LSB matching, to improve the capacity and peak signal to noise ratio (PSNR). The proposed technique has been divided into 3 variants such as variant-1, variant-2, and variant-3. A block consisting of 2 pixels has been considered for data embedding in all the three variants. The variant-1 initially uses the 6th and 7th bit to hide 2 bits of secret data in the first pixel of a block. Further modification to the pixel is done by ±1 or 0 in order to hide 2 more bits in a block. Similarly, the variant-2 hides 3 bits and variant-3 hides 2 bits respectively in a block. The experimental results prove that the variant-1 offers better capacity whereas variant-3 offers better peak signal to noise ratio (PSNR). The results with respect to the steganographic parameters such as embedding capacity, PSNR, and the universal image quality index (Q) has been presented and it is found that the performance of the proposed technique is superior compared to existing techniques.

Index Terms—Steganography, Least significant bit (LSB), Capacity, PSNR.

I. INTRODUCTION

Information security is the dominant field in digital data communication [1]. Security to the digital data during transmission is the supreme importance in this field. Data hiding is the only answer in this aspect. Data hiding can be achieved through prominent approaches like watermarking, cryptography, and steganography [2]. Watermarking is used to determine the ownership and copyright protection. It is divided into 2 types such as visible and invisible watermarking. In case of cryptography, it scrambles the data before transmission to the intended receiver, so that any unauthorized party can’t recognize the data [3].

As opposed to cryptography, steganography uses carrier files to deliver the data to an intended receiver so that the attacker will get no idea about the transmission [4]. The carrier files could be an image, video, audio, or text. Digital images are the preferable choice for delivery of secret data to the intended receiver. When data is hidden inside an image, this is called image steganography [5]. The image before carrying the data is called original image and when the data is loaded inside it, then it is called as stego image. The classification of various data hiding approaches is shown in Fig. 1.

Image steganography can be performed in two domains, namely, spatial domain and frequency domain. Frequency domain techniques use various transformation and inverse transformation techniques. Techniques such as Fourier transform, Z transform, and Laplace transform can be used to transform the image into the frequency domain [6]. The discrete cosine transform (DCT) and discrete wavelet transform (DWT) are the two most widely used transforms for steganography. The homogenous frequency corresponds to smooth pixel region and the high frequencies give an indication of edge or texture regions in an image. The spatial domain directly performs the embedding of secret data on the pixels of the cover image or media. Techniques like least significant bit (LSB) substitution, pixel value differencing (PVD), exploiting modification directions (EMD), histogram shift, spread spectrum, and pixel indicator are popular steganography techniques [7].

LSB substitution technique is the most common and simple technique for digital data transmission. In this technique, the least significant bits of the pixels can be exploited to hide the data [8]. A good number of articles [9-12] have been proposed based on LSB substitution. The LSB substitution greatly increases the capacity of the technique by embedding more number of data bits in the LSB bits. Although this increases the capacity, the distortion to the stego image is higher. When more number of bits in a pixel are changed, then the pixel value changes and that can be easily be identified by the human visual system (HVS) [9].

The steganography technique by Wu and Tsai [13] is called pixel value differencing (PVD). The difference value d between the two consecutive pixels decides how many bits to be embedded inside the pixels of a block. If the difference value between the two consecutive pixels is large, then the pixels can be utilized to hide more number of secret bits. There are varieties of PVD techniques proposed in [14-18].

The LSB techniques provide good capacity whereas PVD techniques have better resistance to steganalytic attacks. Jung [19] has combined the advantage of LSB and PVD techniques to increase the capacity as well as the security of the steganographic system. Khodaei et al. [20] proposed an adaptive PVD and LSB technique to hide secret data. The
difference between two consecutive pixels identifies the number of bits to be hidden inside a block. Again, the pixels are divided into 2 levels i.e. one is from 0 to 191 and other is from 192 to 255. If both the pixels values of the block are less than or equals to 191, then 3-bit LSB substitution is applied. Otherwise, if one of the pixels value is more than 191 then using the predetermined range table the number of bits of secret data to be embedded is selected. Techniques combining LSB and PVD have been discussed in [21-23].

In 2016, Islam et al. [24] proposed a bit differencing technique using most significant bits (MSB) of the pixel. The difference between 5th and 6th MSB bits of the consecutive pixels is computed and if the difference value is same as the secret data bit to be embedded then there is no change to the pixels, otherwise, the 5th MSB bit of the first pixel will be flipped from 0 to 1 or 1 to 0. At the extraction side, the hidden data are extracted by computing the subtraction between the 5th and 6th MSB bits of the consecutive pixels.

II. PROPOSED WORK

The proposed technique has three variants such as variant-1, variant-2, and variant-3. These variants hide 4, 3 and 2 bits respectively in a block of two consecutive pixels. An image is divided into blocks, each block consisting of two consecutive pixels. The embedding and extraction algorithm for the proposed variants are discussed below.

A. Proposed Technique – Variant-1

The embedding algorithm of variant-1 is as described below.

Step-1: Suppose the two pixels of a block are \( P_1 \) and \( P_{i+1} \) and the four bits to be hidden are \( s_1, s_2, s_3, s_4 \).

Step-2: Let the eight bits for the pixel \( P_1 \) and \( P_{i+1} \) are \( b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8 \) and \( c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8 \).

Step-3: Obtain the difference value \( d_1 \) and \( d_2 \) using (1).

\[
(d_1, d_2) = ((b_6 - c_6), (b_7 - c_7))
\]  

Step-4: Obtain the new values for \( b_6 \) and \( b_7 \) using (2), where \( \sim \) is the complement operator.

\[
(b_6, b_7) = \begin{cases}
(b_6, \sim b_7), & \text{if } d_1 = s_1 \text{ and } d_2 = \sim s_2 \\
(b_6, b_7), & \text{if } d_1 = s_1 \text{ and } d_2 = s_2 \\
(\sim b_6, b_7), & \text{if } d_1 = \sim s_1 \text{ and } d_2 = s_2 \\
(\sim b_6, \sim b_7), & \text{if } d_1 = \sim s_1 \text{ and } d_2 = \sim s_2
\end{cases}
\]  

Step-5: After \( b_6 \) and \( b_7 \) values are updated in \( P_1 \) and \( P_{i+1} \) using (2), suppose the new values of two pixels are \( P'_1 \) and \( P'_{i+1} \).

Step-6: Now, use (3) and (4) to update the pixels of \( P'_1 \) and \( P'_{i+1} \) respectively. Rename the pixels as \( P^*_1 \) and \( P^*_{i+1} \) respectively.

\[
P_i^* = \begin{cases}
(P'_1 + 1), & \text{if } s_3 = 1 \text{ and } P'_1 \text{ is even} \\
(P'_1 - 1), & \text{if } s_3 = 0 \text{ and } P'_1 \text{ is odd}
\end{cases}
\]  

\[
P_{i+1}^* = \begin{cases}
(P'_{i+1} + 1), & \text{if } s_4 = 1 \text{ and } P'_{i+1} \text{ is even} \\
(P'_{i+1} - 1), & \text{if } s_4 = 0 \text{ and } P'_{i+1} \text{ is odd}
\end{cases}
\]  

Thus the stego block consists of two pixels \( P^*_1 \) and \( P^*_{i+1} \).

The extraction algorithm is described below.

Step-1: Suppose the pixels of the stego-block are \( P^*_1 \) and \( P^*_{i+1} \).

Step-2: Let the eight bits of \( P^*_1 \) and \( P^*_{i+1} \) are \( b'_1, b'_2, b'_3, b'_4, b'_5, b'_6, b'_7, b'_8 \) and \( c'_1, c'_2, c'_3, c'_4, c'_5, c'_6, c'_7, c'_8 \) respectively. Obtain the difference \( d'_1 \) and \( d'_2 \) using equation 5.

\[
(d'_1, d'_2) = ((b'_6 - c'_6), (b'_7 - c'_7))
\]  

Step-3: Find \( d'_3 = b'_8 \) and \( d'_4 = c'_8 \).

Step-4: Now concatenate \( d'_1, d'_2, d'_3 \) and \( d'_4 \). These four bits are the extracted bits.

The embedding example for the proposed variant-1 is shown in Fig. 2.
The embedding algorithm of variant-2 is as described below.

Step-1: Suppose the two pixels of a block are $P_i$ and $P_{i+1}$ and the three bit of data to be hidden are $s_1, s_2, s_3$.

Step-2: Let the eight bits for the pixel $P_i$ and $P_{i+1}$ are $b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8$ and $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$ respectively.

Step-3: Obtain the difference value $d_1$ and $d_2$ using (6).

\[ (d_1, d_2, d_3) = ((b_6 - c_6), (b_7 - c_7), (b_8 - c_8)) \]  

Step-4: Obtain new values for $b_6, b_7, b_8$ using (7), where $\sim$ is the complement operator.

\[ (b_6, b_7, b_8) = \begin{cases} (b_6, b_7, b_8), & \text{if } d_1 = s_1, d_2 = s_2 \text{ and } d_3 = s_3 \\ (b_6, \sim b_7, \sim b_8), & \text{if } d_1 = s_1, d_2 = s_2 \text{ and } d_3 = s_3 \\ (b_6, b_7, \sim b_8), & \text{if } d_1 = s_1, d_2 = s_2 \text{ and } d_3 = s_3 \\ (\sim b_6, b_7, b_8), & \text{if } d_1 = s_1, d_2 = s_2 \text{ and } d_3 = s_3 \\ (\sim b_6, \sim b_7, b_8), & \text{if } d_1 = s_1, d_2 = s_2 \text{ and } d_3 = s_3 \\ (\sim b_6, b_7, \sim b_8), & \text{if } d_1 = s_1, d_2 = s_2 \text{ and } d_3 = s_3 \\ (b_6, \sim b_7, \sim b_8), & \text{if } d_1 = s_1, d_2 = s_2 \text{ and } d_3 = s_3 \\ (\sim b_6, \sim b_7, \sim b_8), & \text{if } d_1 = s_1, d_2 = s_2 \text{ and } d_3 = s_3 \end{cases} \]  

Step-5: Generate the stego pixels $P_i'$ and $P_{i+1}'$ for the block by considering the updated values of $b_6, b_7,$ and $b_8$.

The extraction algorithm is described below.

Step-1: The stego pixels for the block are $P_i'$ and $P_{i+1}'$.

Step-2: Let the eight bits for $P_i'$ and $P_{i+1}'$ are $b'_1, b'_2, b'_3, b'_4, b'_5, b'_6, b'_7, b'_8$ and $c_1', c_2', c_3', c_4', c_5', c_6', c_7', c_8'$ respectively. Observe the difference $d'_1$ and $d'_2$ using (8).

\[ (d'_1, d'_2, d'_3) = ((b'_6 - c'_6), (b'_7 - c'_7), (b'_8 - c'_8)) \]  

Step-3: Now concatenate $d'_1, d'_2$ and $d'_3$. These are the bits extracted.

C. Proposed Technique – Variant -3

The embedding algorithm of variant-3 is as described below.

Step-1: Suppose the two pixels of a block are $P_i$ and $P_{i+1}$ and the two data bits to be hidden are $s_1, s_2$.

Step-2: Let the eight bits for the pixel $P_i$ and $P_{i+1}$ are $b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8$ and $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$ respectively.

Step-3: Obtain the difference value $d_1$ and $d_2$ using (9).

\[ (d_1, d_2) = ((b_7 - c_7), (b_8 - c_8)) \]  

Step-4: Obtain new values for $b_7, b_8$ using (10), where $\sim$ is the complement operator.

\[ (b_7, b_8) = \begin{cases} (b_7, b_8), & \text{if } d_1 = s_1 \text{ and } d_2 = s_2 \\ (b_7, \sim b_8), & \text{if } d_1 = s_1 \text{ and } d_2 = s_2 \\ (\sim b_7, b_8), & \text{if } d_1 = s_1 \text{ and } d_2 = s_2 \\ (\sim b_7, \sim b_8), & \text{if } d_1 = s_1 \text{ and } d_2 = s_2 \end{cases} \]  

Step-5: Generate the stego pixels $P_i'$ and $P_{i+1}'$ for the block by considering the updated values of $b_7, b_8$.
The universal quality index (Q) can be computed using (14). The universal quality index (Q) suggests the similarity between the original and stego image. If two images are completely identical, then the ‘Q’ value is 1.

\[
Q = \frac{4 \sigma_{xy} \tilde{p}^2}{\sigma_x^2 + \sigma_y^2} \left[ \sigma (\tilde{p})^2 + (\tilde{q})^2 \right]
\]

Where,

\[
\tilde{p} = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} (p_{ij} - \bar{p})
\]

\[
\tilde{q} = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} (q_{ij} - \bar{q})
\]

\[
\sigma_x^2 = \frac{1}{m \times n - 1} \sum_{i=1}^{m} \sum_{j=1}^{n} (p_{ij} - \bar{p})^2
\]

\[
\sigma_y^2 = \frac{1}{m \times n - 1} \sum_{i=1}^{m} \sum_{j=1}^{n} (q_{ij} - \bar{q})^2
\]

\[
\sigma_{xy} = \frac{1}{m \times n - 1} \sum_{i=1}^{m} \sum_{j=1}^{n} (p_{ij} - \bar{p})(q_{ij} - \bar{q})
\]

The PSNR speaks about the quality of stego image, the higher the PSNR the lesser is the distortion. Capacity speaks the total amount of secret data an image can hide [26]. BPP speaks about the capacity of individual pixels [27]. The proposed variants have been compared with Wu & Hwang [9], 1-bit and 2-bit LSB substitution. The proposed variant-1 offers a better capacity of 2 BPP compared to other techniques whereas the variant-3 offers PSNR of 54.38 dB. Although the PSNR for Wu & Hwang [9] is 49.50 dB it has a limited capacity of 1 BPP only. Similarly, 1-bit and 2-bit LSB substitution method offer PSNR of 47.44 dB and 43.94 dB with BPP of 1 and 2 respectively. Hence from the above analysis, it has been proved that the proposed variant-1 offers better capacity and variant-3 offers better PSNR.

### Table I

<table>
<thead>
<tr>
<th>Images</th>
<th>Variant-1</th>
<th>Variant-2</th>
<th>Variant-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>42.95</td>
<td>0.96</td>
<td>524288</td>
</tr>
<tr>
<td>Baboon</td>
<td>43.90</td>
<td>0.98</td>
<td>524288</td>
</tr>
<tr>
<td>Peppers</td>
<td>44.96</td>
<td>0.97</td>
<td>524288</td>
</tr>
<tr>
<td>Boat</td>
<td>43.55</td>
<td>0.98</td>
<td>524288</td>
</tr>
<tr>
<td>House</td>
<td>43.67</td>
<td>0.98</td>
<td>524288</td>
</tr>
<tr>
<td>Baby</td>
<td>44.89</td>
<td>0.98</td>
<td>524288</td>
</tr>
<tr>
<td>Barbara</td>
<td>44.90</td>
<td>0.97</td>
<td>524288</td>
</tr>
<tr>
<td>Bird</td>
<td>43.78</td>
<td>0.93</td>
<td>524288</td>
</tr>
<tr>
<td>Average</td>
<td>44.07</td>
<td>0.96</td>
<td>524288</td>
</tr>
</tbody>
</table>

### Table II

<table>
<thead>
<tr>
<th>Images</th>
<th>Wu &amp; Hwang [9], 1-bit and 2-bit LSB substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>50.04 0.99 262144 1 45.35 0.99 524288 2</td>
</tr>
<tr>
<td>Baboon</td>
<td>50.89 0.98 262144 1 42.62 0.96 524288 2</td>
</tr>
<tr>
<td>Peppers</td>
<td>50.01 0.98 262144 1 44.39 0.96 524288 2</td>
</tr>
<tr>
<td>Boat</td>
<td>49.10 0.97 262144 1 43.20 0.94 524288 2</td>
</tr>
<tr>
<td>House</td>
<td>49.09 0.98 262144 1 43.47 0.96 524288 2</td>
</tr>
<tr>
<td>Baby</td>
<td>48.92 0.98 262144 1 44.36 0.96 524288 2</td>
</tr>
<tr>
<td>Barbara</td>
<td>48.98 0.99 262144 1 44.80 0.95 524288 2</td>
</tr>
<tr>
<td>Bird</td>
<td>49.01 0.98 262144 1 43.94 0.95 524288 2</td>
</tr>
<tr>
<td>Average</td>
<td>49.50 0.98 262144 1 43.44 0.97 524288 2</td>
</tr>
</tbody>
</table>
This paper proposes an improved image steganographic technique using the principle of modified LSB substitution and LSB matching. The proposed technique has 3 variants. The experimental results prove that the variant-1 provides better capacity with 4 bits per block whereas variant-3 provides better PSNR with 54.38 dB. Further, the proposed variants have been compared with existing techniques with respect to the image steganographic parameters such as PSNR, capacity and quality index and found to be satisfactory.

REFERENCES


[23] G. Swain, “Digital image steganography using eight directional PVD against RS analysis and PDH analysis” Accepted for publication, Advances in Multimedia, to be published.


Mr. Aditya Kumar Sahu was born and brought up in Odisha, India. He received B.Tech (IT) degree from G.I.E.T Gunupur, Odisha, India, under Biju Pattanaik University in 2007, M.Tech (CS) degree from M.I.T.S Rayagada, Odisha, India, under Berhampur University in 2011. He is pursuing Ph.D. (CSE) in Koneru Lakshmaiah Education Foundation, Vaddeswaran, Guntur, Andhra Pradesh, India. He is working as an assistant professor in the Department of Computer Science & Engineering, GMRIT, Guntur, Andhra Pradesh, India. He has more than 10 years of teaching experience and published more than 10 research articles in international journals and conferences. His research interests include network security and image steganography. Mr. Sahu is a life member of Computer Society of India (CSI).

Dr. Gandharba Swain was born and brought up in Odisha. He has received B.Sc. (Hons) degree from Berhampur University in 1995, MCA degree from VSS University of Technology (VSSUT), Burla, in 1999, M.Tech (CSE) degree from NIT, Rourkela in 2004. He was awarded with Ph.D. degree from SOA Deemed University, Bhubaneswar in 2014. He is working as a Professor in the Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaran, Guntur, Andhra Pradesh, India. Previously he had worked at IETE, Rajam and Andhra Pradesh, India. He has more than 10 years of teaching experience and more than 10 research articles in international journals and conferences. His research interests include network security and image steganography. Mr. Sahu is a life member of Computer Society of India (CSI).