A space increased reversible information hiding technique by reducing redundant recording

Jeanne Chen\textsuperscript{a}, Tung-Shou Chen\textsuperscript{b}, Wien Hong\textsuperscript{c}, Chi-Nan Lin\textsuperscript{d}, Mei-Chen Wu\textsuperscript{e}, Han-Yan Wu\textsuperscript{f}

\textsuperscript{a}\textsuperscript{b}Department of Computer Science and Information Engineering, National Taichung University of Science and Technology, Taichung 404, Taiwan
\textsuperscript{c}\textsuperscript{e}\textsuperscript{f}Department of Information Management, Yu Da University, Miaoli 361, Taiwan
\textsuperscript{d}Department of Management Information Systems, Central Taiwan University of Science and Technology, Taichung 406, Taiwan

Abstract

This paper proposes improvements to the reversible hiding technique proposed in by Hong et al. in 2012. The proposed technique is based on the characteristics of the human visual on neighboring pixel values to obtain the average for calculating the minimum value of just noticeable difference (JND). The JND is then used to determine the suitable embedding level to decrease image distortion. Hong et al.’s method will first shift pixels that caused overflow and underflow after finding the suitable embedding level. The values of the shifted pixel will be recorded in a location map the size of an image. The binary location map was JBIG2 compressed and embedded with the secret information. In the location map, pixels not shifted were recorded as 0. Therefore, many pixels were recorded. These resulted in the compressed length to be longer which affects the payload size. Experimental results showed that our proposed method decreases the size of the location map, payload is increased but the image quality is maintained similar to Hong et al.’s method.

© 2013 The Authors. Published by Elsevier B.V.
Selection and peer-review under responsibility of the organizers of the 2013 International Conference on Information Technology and Quantitative Management

Keywords: Reversible data hiding, human visual system, just noticeable difference (JND), histogram shifting

1. Introduction

Although the advancement in the World Wide Web has brought along a lot of conveniences, the capacity of information transmission has also escalated. Therefore, information security has been an important research by experts. Data hiding is used to carry information on digital media like text, sound and image. Data hiding in images can be differentiated between reversible [1-3] and non-reversible [4-5] techniques.

* Corresponding author. Tel.: +886 37 651188/6129; fax: +886 37 652143.
E-mail address: wienhong@ydu.edu.tw.
The reversible data hiding technique was originally proposed by Baton [6] in 1997. Following many more were being proposed. In 2003, Tian [7] proposed a new embedding technique by expanding the pixel difference for embedding data. However, the values of the modified were too large and caused high distortion. Alattar [8], in 2004, proposed an improvement to Tian’s method by first using \( n \) pixels to expand the pixel difference. Then embed the difference in \( n - 1 \) bits secret information. The other methods are similar to Tian’s method which resulted in high distortions. Ni et al. [9] in 2006 proposed a novel hiding technique by combining histogram shifting and embedding by the histogram peak. Although Ni et al.’s method resulted in high image quality, the embedded payload is affected by the peak’s height. In 2007, Thodi and Rodriguez [10] combined Tian and Ni et al.’s methods to expand the pixel difference and to embed by the histogram technique. However, similar to Tian’s method the method resulted in high distortions. Lin et al. [11] in 2008 made use of shifting to embed data. The method resulted in larger payload than Ni et al.’s method. However, some space was taken up by the need to record the pixels on the edges.

Tsai et al. [12] in 2009 proposed improving Ni et al.’s method by partitioning an image into \( 3 \times 3 \) blocks. The pixel located in the middle is the reference pixel for its surrounding 8 pixels. The 8 error values are obtained after deductions. The error values are calculated for all the blocks which were then used in the histogram statics. The histogram is differentiated for error values greater or equal 0 and less than 0 to locate two highest peaks for embedding. Although Tsai et al.’s method resulted in larger payload than Ni et al.’s method. However, too many pixels are shifted before embedded which downgrades the image. Within the same year, Tai et al. [13] proposed modifying pixels by embedding level parameters and to use predicted error values. Although the method has significant payload, the accuracy for predicted values affected the payload.

These techniques did not consider the human visual system when embedding. Normally, the human visual is sensitive to the complex region. Therefore, some research was aimed at this characteristic to develop related embedding technique. In 2011, Jung et al. [14] proposed a reversible data hiding technique based on the human visual system. The maximum allowed changes is calculated for the original image. Then Tai et al.’s method is used to calculate the maximum change for the embedding level for embedding information. Compared to previous research, Jung et al.’s method under the same payload resulted in higher image quality. However, their proposed method made used of the mean of previously accessed pixels to predict the current pixel value. This resulted in larger prediction error difference. Also, the method required a larger location map to record the information for extraction use. All these factors caused the payload size to decrease. In 2012, Hong et al. [15] proposed improvements to Jung et al.’s method by using the neighboring three pixels for predicting and to use the human visual characteristic to calculate the just noticeable difference (JND). The JND values were used on pixels that are suitable for shifting in order to decrease the size of the location map. They also used the embedding level adjustment mechanism to avoid the edges which would have resulted in more changes whereby incurring unnecessary location map. Hong et al.’s method resulted in image quality higher than Jung et al.’s method be it low or high payloads. At the same time, the smaller location map also resulted in increase payload size.

Although Hong et al.’s method resulted in considerable image quality and the payload is higher than Jung et al.’s method, the location map will take up some space which affects the payload size. Therefore, we proposed an improvement to the location map of Hong et al.’s method which is the size of the original image for recording pixels that were shifted and not shifted. In our proposed method, only the pixels that were shifted will be recorded. This significantly decreases the size of our location map.

2. Related Work

In 2012, Hong et al. [15] proposed a reversible data hiding histogram technique based on the human visual system. The median edge predictor (MED) was used to predict pixel values. The predicted error difference is
Jeanne Chen et al. / Procedia Computer Science 17 (2013) 229 – 236

shifted bidirectional for making modifications in order to increase the payload at low embedding level. They made use of the adaptable histogram to increase payload at low embedding level. To deal with the location map, the adaptable histogram is used. The embedding level $L_{i,j}$ is shifted $2^{k_{i,j}}$ to decrease the unnecessary pixel change. The selection mechanism for the embedding Level is based on decreasing the neighboring smooth region. The design is to persuade a smaller distortion when shifting. At the same time and can avoid the recording information in the location map whereby resulting in smaller location map.

Hong et al. proposed method made use of pixel, $I_{i,j}$, left, upper and left upper to be the three neighboring pixels with values, $\Omega_{i,j} = \{I_{i,j-1}, I_{i-1,j}, I_{i-1,j-1}\}$. The average $\bar{\Omega}_{i,j}$ is calculated for $\Omega_{i,j}$ and the variance is $\sigma^2(\Omega_{i,j})$. Eq. (1) is used to calculate smoothness and complexity for pixel.

$$E_{i,j} = \begin{cases} 0, & \text{var}(\Omega_{i,j}) \leq T_E; \\ 1, & \text{var}(\Omega_{i,j}) > T_E, \end{cases}$$

where 0 and 1 represents the pixel in smooth or edge region. $T_E$ is the threshold for prediction, $\text{var(*)}$ is the variance. Next, Eq. (2) is used to find the JND of pixel $I_{i,j}$.

$$J_{i,j} = TL_{i,j}(x) + \frac{T_{A_{i,j}}(x)}{TL_{i,j}(x)},$$

where $TL_{i,j}(x)$ and $T_{A_{i,j}}(x)$, respectively, represent the background brightness $x$ threshold and pixel $I_{i,j}$ visible activity threshold. $TL_{i,j}(x)$ is defined as in Figure 1, where when $E_{i,j} = 0$, then $a = 10$, $b = 20$, and $c = 24$. If $E_{i,j} = 1$, then $a = 8$, $b = 18$, and $c = 22$. $T_{A_{i,j}}(x)$ is defined as in all of $\Omega_{i,j}$ pixels which larger than the smallest pixel difference.

![Fig.1 Background brightness and visible thresholds](image)

After using Eq. (2) to obtain the JND value $J_{i,j}$ of pixel $I_{i,j}$, the embedding level $L_{i,j}$ is calculated for pixel $I_{i,j}$. If the image is in the smooth region and $\bar{\Omega}_{i,j}$ is not near the edge, then restrict $I_{i,j}$ during embedding to have changes smaller than JND in order to preserve image quality. Therefore, Eq. (3) is used to derive the embedding level.

$$L_{i,j} = \arg \max_k 2^k,$$

subject to $2^k < J_{i,j}, 2^{L_{\text{Max}}} < \bar{\Omega}_{i,j} < 255 - 2^{k_{\text{Max}}}, k \leq L_{\text{Max}}$. 
If the image is in the complex region and \( \Omega_{i,j} \) is not near the edge then restrict pixel \( I_{i,j} \) during embedding to have changes greater than JND so that the payload is increased. Therefore, \( L_{i,j} \) can use Eq. (4) to make the following decision.

\[
L_{i,j} = \arg \min_k 2^k,
\]

subject to \( 2^k > J_{i,j}, 2^{L_{\text{Max}}+1} < \Omega_{i,j} < 255 - 2^{L_{\text{Max}}+1}, k \leq L_{\text{Max}} \).

If the above two conditions do not match, then embedding will result in overflow or underflow. The location information has to be deleted. At the same time, shifting will cause large damage. At this time, no recording is required. The location must be recorded, to give \( L_{i,j} \) a smaller but permanent embedding level \( L_e \), in order to raise image quality and to decrease the size of location map.

After getting the embedding level, then based on the size of \( L_{i,j} \) the \( \hat{I}_{i,j} \) is shifted \( 2^{L_{i,j}} \) to avoid overflow or underflow. Hong et al.’s method is when \( \hat{I}_{i,j} < 2^{L_{i,j}} \), then \( I_{i,j} = \hat{I}_{i,j} + 2^{L_{i,j}} \) and \( LM_{i,j} = 1 \) is recorded. If \( \hat{I}_{i,j} > 255 - 2^{L_{i,j}} \), then \( I_{i,j} = \hat{I}_{i,j} - 2^{L_{i,j}} \) and record \( LM_{i,j} = 1 \). After completing for all pixels, the location map \( LM \) is the size of the original image. Finally, \( LM \) is JBIG2 compressed which is the compressed location map \( C_{LM} \). Then \( C_{LM} \) is concatenated with secret information \( S \) resulting in \( M = C_{LM} | S \).

3. Proposed Method

Hong et al.’s [15] method used a location map the size of the original image to record the shift locations as 1s and 0s for non-shift. The length of the location map is long even after compression and occupies the space meant for payload. We proposed an improvement to Hong et al.’s method by recording only the pixels that were shifted and pixels that might have been modified. Both were respectively recorded as 1s and 0s. This guarantees the recovery for pixels that were shifted or modified.

The embedding level is used to determine whether the pixel is in the shift range. Suppose the embedding level \( L \) is 2 and the pixels in the range \( \hat{I}_{i,j} < 2^2 \) or \( \hat{I}_{i,j} > 255 - 2^2 \) are to be shifted. The pixel locations are recorded as 1 in the location map. If pixels are in the ranges \( 2^2 \leq \hat{I}_{i,j} \leq 2 \times 2^2 - 1 \) or, \( 255 - 2^2 \geq \hat{I}_{i,j} \geq 255 - 2 \times 2^2 + 1 \) then the locations are recorded as 0s. As seen in Fig.2, if the pixel with value of 248 is not recorded as 0, then after 252 is shifted such that \( 252 - 2^2 = 248 \); it will be difficult to determine which pixel is the original pixel. Therefore, the pixels within the range will be recorded as 0s to guarantee recovery to original image. The range is described in Fig.2. After shifting, the information is embedded and as follows

\[
\begin{align*}
2^L &= 2^2 \\
255 - 2^L &= 255 - 2^2
\end{align*}
\]

<table>
<thead>
<tr>
<th>Cover pixel</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>Location map</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

![Fig.2 Recording the location map](image)

3.1 Embedding algorithm

The embedding algorithm used the three neighbor pixels \( \Omega_{i,j} \) to calculate the pixel \( I_{i,j} \) in the embedding
level $L_{i,j}$. Next, shift bidirectional by prediction error difference for embedding based on the size of $L_{i,j}$. MED is used to obtain the predict value of $I_{i,j}$. Fig. 3 shows the MED prediction.

$$p = \begin{cases} \min(a,b), & c \geq \max(a,b); \\ \max(a,b), & c \leq \min(a,b); \\ a + b - c, & \text{otherwise}. \end{cases}$$

Fig.3 MED prediction

The embedding steps are as follows.

Input: Original image $\hat{I}$, embedding level $L$, secret information $S$.
Output: Stego image $I'$, location map length $|LM|$, secret information length $|S|$.

Step 1: As mentioned earlier, from the original image $\hat{I}$ obtain the shifted image $I$ and the location map $LM$. Concatenate $LM$ with the secret information $S$ to get $M = LM | S$.

Step 2: Set $I' = I$. Scan from row two, column two and from upper to lower, left to right of image $I'$. After scanning all the pixels $I_{i,j}$, get the neighbor pixels $\Omega_{i,j}$ of $I_{i,j}$. Select the suitable embedding level to get the embedding level $L_{i,j}$ used for embedding information.

Step 3: Use $\Omega_{i,j}$ and the MED method to predict $I_{i,j}$ value $p$. Next, calculate the prediction error difference $d_{i,j} = I_{i,j} - p$. If $-2^{k_{i,j}} \leq d_{i,j} < 2^{k_{i,j}}$ then from $M$ extract $m$ bits as in Eq. (5).

$$I'_{i,j} = I_{i,j} + d_{i,j} + m.$$  \hspace{2cm} (5)

Otherwise, do as in Eq. (6) to shift $I_{i,j}$.

$$I'_{i,j} = \begin{cases} I_{i,j} + 2^{k_{i,j}}, & d_{i,j} \geq 2^{k_{i,j}}; \\ I_{i,j} - 2^{k_{i,j}}, & d_{i,j} < -2^{k_{i,j}}. \end{cases}$$  \hspace{2cm} (6)

Step 4: Repeat steps 2-3 until all of $M$ are embedded.

In order to extract the embedded information, the receiver must have $|LM|$ and $|S|$. Therefore, $|LM|$ and $|S|$ treated as a key and transmitted to the receiver by a secret channel.

### 3.2 Extraction algorithm

The secret information and image recover can be performed after receiving $I'$ and the key $|LM|$ and $|S|$. Detail is presented as follows.

Input: Stego image $I'$, $|LM|$, and $|S|$.
Output: Original image $\hat{I}$ and secret information $S$.

Step 1: Set $I = I'$. Scan from row two, column two and from upper to lower, left to right of image $I$. Next get the neighbor values $\Omega_{i,j}$ of $I_{i,j}$. Then based on the $\Omega_{i,j}$ and embedding level to get the embedding level $L_{i,j}$ used by $I_{i,j}$.

Step 2: Use the neighbor $\Omega_{i,j}$ and MED prediction method to predict $I'_{i,j}$ value $p$. Calculate the difference between $I'_{i,j}$ and $p$ to get $d_{i,j} = I'_{i,j} - p$. If $-2^{k_{i,j}+1} \leq d_{i,j} < 2^{k_{i,j}+1}$ then extract bits, $m = \text{mod}(d_{i,j}, 2)$ and set $I_{i,j} = I'_{i,j} - \lfloor d_{i,j} / 2 \rfloor$, otherwise set $I_{i,j}$ as in Eq.(7).
\[
I_{i,j} = \begin{cases} 
I'_{i,j} - 2^{L_{i,j}}, & d_{i,j} \geq 2^{L_{i,j}} + 1; \\
I'_{i,j} + 2^{L_{i,j}}, & d_{i,j} < -2^{L_{i,j}}. 
\end{cases}
\]

(7)

Step 3: Repeat steps 2-7 until all the embedded information are extracted. Extract from the information \(|LM|\) bits as the location map and \(|S|\) bits of embedded secret information.

Step 4: Scan the entire image \(I\). If pixel \(I_{i,j}\) is within the range \(2^2 \leq I_{i,j} \leq 2 \times 2^2 - 1\) or \(255 - 2^2 \geq I_{i,j} \geq 255 - 2 \times 2^2 + 1\) and \(LM_{i,j} = 1\), then this means that shifting had been done before the information was embedded. Eq.(8) is used to recover the pixel. However, if \(LM_{i,j} = 0\) then this means that the original pixel is preserved and no modification is required.

\[
\hat{I}_{i,j} = \begin{cases} 
I_{i,j} - 2^{L_{i,j}}, & 2^{L_{i,j}} \leq I_{i,j} \leq 2 \times 2^{L_{i,j}} - 1; \\
I_{i,j} + 2^{L_{i,j}}, & 255 - (2 \times 2^{L_{i,j}} - 1) \leq I_{i,j} \leq 255 - 2^{L_{i,j}}. 
\end{cases}
\]

(8)

After shifting is completed, the original image is recovered.

4. Experimental Results

In the experimental tests, eight 8-bit gray images of size \(512 \times 512\) were used. The original images are as shown in Fig.4 and accessed from [16] and [17]. The pure payload and location map length are used to calculate amount of payload after improvement. PSNRs are calculated to compare the quality of the two methods. The PSNR measures the difference between the original image and stego image. Eq.(9) calculates the PSNR.

\[
\text{PSNR} = 10 \log_{10} \frac{255^2}{\text{MSE}}
\]

(9)

Fig.4 Eight 8-bit gray images of size \(512 \times 512\)
Table 1 shows the results when the Embedding Level is 0. As seen in the table, we have made significant improvements to Hong et al.’s method. The results showed image quality is maintained and compressed location map length is shorter. The reason for the shorter length is that locations not shifted were not recorded which means that this part of location map length is 0. Since the compressed location map length is shorter, there is more room for payload.

Table 1 Comparison results when Embedding Level is 0

| Image | Pure payload | $|LM|$ | PSNR | Hong et al. | Proposed method |
|-------|--------------|------|------|------------|-----------------|
|       | Pure payload |     |      | PSNR       | Pure payload    | $|LM|$ | PSNR |
| A     | 53112        | 312 | 48.62| 53424      | 0               | 48.62 |
| B     | 32622        | 504 | 48.43| 33103      | 23              | 48.43 |
| C     | 38307        | 312 | 48.48| 38619      | 0               | 48.48 |
| D     | 80023        | 312 | 48.87| 80335      | 0               | 48.87 |
| E     | 29112        | 312 | 48.40| 29424      | 0               | 48.40 |
| F     | 29313        | 312 | 48.40| 29625      | 0               | 48.40 |
| G     | 49699        | 672 | 48.59| 50266      | 105             | 48.59 |
| H     | 127780       | 480 | 49.38| 128240     | 20              | 49.38 |
| Average | 54996      | 402 | 48.65| 55380      | 19              | 48.65 |

Table 2 shows the comparison results when the embedding level is 5. When the embedding is larger, more pixels will be shifted. Hong et al.’s method made use of the location map which is the size of the original image to record. Meanwhile, our proposed method records only the locations of pixels that were shifted. Obviously, our proposed method will be smaller than Hong et al.’s method. Therefore, the compressed location map length in our proposed method is also shorter. This created more space for payload.

Table 2 Comparison results when Embedding Level is 5

| Image | Pure payload | $|LM|$ | PSNR | Hong et al. | Proposed method |
|-------|--------------|------|------|------------|-----------------|
|       | Pure payload |     |      | PSNR       | Pure payload    | $|LM|$ | PSNR |
| A     | 226959       | 448 | 34.63| 227314     | 93              | 34.63 |
| B     | 190218       | 2624| 32.58| 192187     | 655             | 32.58 |
| C     | 200889       | 352 | 33.57| 201218     | 23              | 33.57 |
| D     | 247936       | 2080| 37.35| 249479     | 537             | 37.35 |
| E     | 162819       | 1056| 32.43| 163733     | 142             | 32.43 |
| F     | 174789       | 648 | 31.83| 175166     | 271             | 31.83 |
| G     | 182226       | 1512| 32.99| 182876     | 862             | 32.99 |
| H     | 241042       | 1120| 37.22| 241811     | 351             | 37.22 |
| Average | 203360      | 1230| 34.00| 240223     | 367             | 34.00 |

5. Conclusion

Hong et al.’s method required the location map which the size of the original image to record the locations of all pixels. The location map length is longer which competes with payload embedding space. In our proposed method, we record only locations that were shifted and those pixels that might cause overflow and underflow. Therefore, the length of the location map is shorter which creates more space for payload. Furthermore, our proposed method in comparison to Hong et al.’s method resulted in similar high image quality but higher payload.
Reference