Application of Threshold Secret Sharing to Image Authentication and Recovery

Shu-Fen Tu*, Ching-Sheng Hsu** & Fu-Hsing Wang***

*Associate Professor, Department of Information Management, Chinese Culture University, Taipei, TAIWAN.
E-Mail: dsf3@ulive.pccu.edu.tw
**Associate Professor, Department of Information Management, Ming Chuan University, Taoyuan County, TAIWAN.
E-Mail: cshsu@mail.mcu.edu.tw
***Professor, Department of Information Management, Chinese Culture University, Taipei, TAIWAN. E-Mail: wang.fuhsing@gmail.com

Abstract—The proposed method borrowed the concept of threshold secret sharing to propose an image recovery scheme. The recovery message of each image block was split into three shares and spread out over the image. When a tampered block needs to be repaired, gathering at least two shares can reconstruct the recovery message. The main merit is that each tampered block has second chance for recovery; moreover, the length of recovery message carried by each pixel can be shortened. As regards to the image authentication, we employed Hsu and Tu’s scheme, which has low false detection rate. To increase the quality of recovery images, we add a third stage on Hsu and Tu’s scheme to decrease false negative rate. The experiment results show the proposed scheme can detect the tampering with low false negative rate and recover the tampered image with good quality.

Keywords—Fragile Watermarking; Image Authentication and Recovery; Polynomial Interpolation; Tampering Detection; Threshold Secret Sharing.

I. INTRODUCTION

NOWADAYS, multimedia data are more and more accessible to everyone, but are also easily exposes to malicious modification. Therefore, integrity authentication becomes an important issue for multimedia data. Digital signature is a feasible solution to authenticate the integrity of digital images. It can confirm if the image is tampered or not, but cannot locate the tampering area. Another feasible solution is fragile watermarking scheme, which cannot only confirm the tampering status, but also can point out which part is modified. Moreover, some fragile watermarking schemes can recover the tampering area to a certain extent [Chan & Cheng, 2004; Lin et al., 2005; Chang et al., 2006; Yeh & Lee, 2006; Liu et al., 2007; Chang et al., 2008; Lee & Lin, 2008; Zhang & Wang, 2009; Hsu & Tu, 2010; Yang & Shen, 2010].

According to the working domain, fragile watermarking scheme can be classified into spatial-domain based and frequency-domain based. At transmitter site, the image feature is extracted and transformed into authentication message. For a spatial-domain based scheme, the authentication message is embedded into the least significant bits of pixels. If the scheme intends to provide capability for recovering, rough pixel intensities of the image will be record and embedded before the authentication message is generated. When detecting tampering, the receiver can compare the authentication message inside the image with the regenerated authentication message to find out the modified region. After detecting the tampering area, the receiver can utilize the recovery message inside to fix them.

Generally speaking, the quality of the recovery image depends on the detect accuracy. The detect accuracy means that if a pixel is tampered, the tampering status has to be detected successfully; however, if a pixel is not tampered, it cannot be recognized as a tampered one. The former accuracy means true positive, and the latter one is true negative. Therefore, a good tampering detection scheme should decrease the rate of false positive and false negative as more as possible. Researchers usually adopt multi-stage tampering detection to decrease the rate of false negative, such as Lin et al.’s scheme [Lee & Lin, 2008]. Lin et al., proposed a 3-stage hierarchical scheme to detect tampered pixels. In Lin et al.’s scheme, more and more possible tampered pixels are found out stage by stage. However, the rate of false positive increases consequently. The increase of false positive is also caused by tampering unit. In Lin et al.’s scheme, the tampering detection is performed block by block, and the size of is 4 × 4 pixels. If only one pixel within the block is tampered actually, the whole block will be recognized as tampered. Another problem is that if a block is marked as invalid and its mapping block is also invalid, the block will never be recovered. In view of this, Lee & Lin (2008) reduced the size of tampering unit to 2×2 pixels. In addition, dual copies of watermark are embedded into the last three bits
of each pixel, so that every block has second chance to be recovered if it is tampered. The idea of dual copies of watermark is distinguished, but the length of bits carried by a pixel increases as well. Another problem of Lee and Lin’ scheme is that only the last three bits of a pixel is used to detect tampering; moreover, the block number does not involve in the calculation of watermark. If the last three bits does not act on the modification, such tampering can never be detected. For example, if a region of the image is copied and pasted to another region, Lee and Lin’ scheme is not sensitive to such tampering since the last three bits are not changed.

In 2012, Hsu and Tu proposed a probability-based tampering detection scheme for digital images [Hsu & Tu, 2010]. Unlike Lee and Lin’s scheme, Hsu and Tu’s scheme can discover the copy-and-paste attack because the authentication message includes the location information of the block. The detection accuracy of Hsu and Tu’s scheme is better than other researchers’ as shown in the experimental results, but they did not provide a further solution to recover the tampered image. In 2013, we proposed a hierarchical tampering detection and recovery scheme to reach better detection error rate and better image quality of recovered images [Tu & Hsu, 2013]. Our proposed scheme augmented Hsu and Tu’s image authentication scheme with the ability of restoring tampering area. For each image block, we evaluate the average pixel as its recovery message. The key feature of our scheme is that the recovery message was split into four 6-bit shares by means of Shamir’s three-out-of-four threshold secret sharing scheme at first. Then, the four shares were spread over the image to raise their survivability. Averagely, the bits carried by a pixel is 2, which is shorter than 2.5 bits of Lee and Lin’s scheme.

The purpose of this paper is to improve our previous work by further shortening the average bits carried by a pixel. The main feature is that the recovery message will be split to three shares according to Shamir’s two-out-of-three threshold secret sharing scheme. To protect the three shares from being destroyed simultaneously, the three shares are distributed apart from each other. When recovering a tampered image, the receiver can gather any two of the three shares to reconstruct the recovery message. By doing so, we can get double benefits: first, the length of bits carried by a pixel can be shortened; second, a tampered block may have more chance to be recovered since the recovery message still can be reconstructed successfully if one share is lost. In addition, we redesign the embedding rule and algorithm corresponding to the two-out-of-three secret sharing scheme. The rest of this paper is organized as follows. A review of Shamir’s threshold secret sharing scheme is provided for readers as preliminary knowledge in section 2. As regards Hsu and Tu’s scheme, we do not plan to review due to the limitation of paper length. Interesting readers can refer to [Hsu & Tu, 2010] for the detail. Next, the detail of the proposed scheme is explained in section 3. Then, the experimental results are shown in section 4. Finally, some conclusions are given in section 5.

II. RELATED WORKS

2.1. Shamir’s Threshold Secret Sharing Scheme

In 1979, Shamir proposed a \((k, n)\)-threshold secret sharing scheme based on polynomial interpolation [Shamir, 1979]. In a \((k, n)\)-threshold secret sharing scheme, the secret \(D\) is split into \(n\) shares denoted as \(D_1, D_2, \ldots, D_n\), and gathering at least \(k\) shares can reconstruct the secret successfully. To split the secret, we have to randomly choose \(k-1\) integers \(a_1, a_2, \ldots, a_{k-1}\) between 0 and \(p-1\) to build the following polynomial \(q(x)\):

\[
q(x) = D + a_1 x + a_2 x^2 + \ldots + a_{k-1} x^{k-1} \mod p,
\]

where \(p\) is a prime number. The then shares are evaluated:

\[
D_1 = q(1), \quad D_2 = q(1), \ldots, \quad D_n = q(n).
\]

The inverse evaluation, i.e., determining the coefficients of \(q(x)\) from a set of \(k\) point-value pairs \(\{(x_0, y_0), (x_1, y_1), \ldots, (x_k, y_k)\}\) such that all of the \(x_i\) are distinct and \(y_i = q(x_i)\) for \(i=0..(k-1)\) can be performed by the following Lagrange’s formula [Whittaker & Robinson, 1967]:

\[
q(x) = \sum_{i=0}^{k-1} y_i \prod_{j \neq i} \frac{(x - x_j)}{(x_i - x_j)}.
\]

2.2. Lin and Lee’s Dual Watermarking Scheme

In 2008, Lin and Lee proposed a tampering detection and recovery scheme for digital images [Lee & Lin, 2008]. At the transmitter site, the original image was split into nonoverlapping blocks of \(2 \times 2\) pixels. For each time, a pair of two blocks, say \(A\) and \(B\), were selected from upper half and lower half of the image, respectively. Then, the average pixel values of block \(A\) and \(B\) were calculated, truncated and combined into 10-bits recovery message. According to the 10-bits message, two parity bits were generated and appended to the end of the 10-bits message. The joint 12-bits watermark was duplicated and embedded into the least significant three bitplanes of two other blocks. At the receiver site, the tampering status of each block was judged according to the parity bits. If a block, say \(D\), is judged as tampered, the block \(D\) was recovered with the average value of its watermark. Since the watermark was duplicated and resided into different two blocks, any one copy can be used to recover \(D\). If one of the two resident blocks was judged as tampered, the block and the watermark inside was recognized as invalid. When such situation happened, the other copy of the watermark was used to recover \(D\). It is a novel idea to duplicate the watermark. Since the two copies of the watermark can be backup for each other, the survivability of the watermark can be enhanced. However, the duplication of the authentication message, i.e. the two parity bits, is not necessary. Besides, if the tampering only make modification to the first five bits of the pixels, a tampered block may be misjudged as untampered because the parity bits are generated according to the 10-bit recovery message. Another problem is that their scheme may not resist to copy-and-paste attacks since the watermark does not include the block location.
2.3. Hsu and Tu’s Probability-based Tampering Detection Scheme

In 2012, Hsu and Tu proposed a probability-based tampering detection scheme for digital images [Hsu & Tu, 2010]. At the transmitter site, both the coordinate and pixel values of the whole block involve in calculating the authentication message. At the receiver site, the tampering detection goes through two stages. Similar to other researchers’ method, Hsu and Tu regenerated the authentication message and compared them with the one inside the image to find out invalid blocks. The different part from other researchers’ is the second stage of detection. At the second stage, Hsu and Tu refined the detection result of the first stage by calculating the probability that a block recognized as valid at the first stage is actually a tampered one. If the probability is more than 0.5, the detection result of that block will be modified to be invalid. The detection accuracy of Hsu and Tu’s scheme is better than other researchers’ as shown in the experimental results, but a fly in the ointment was that they did not provide a further solution to recover the tampered image.

III. THE PROPOSED SCHEME

The proposed scheme is divided into two parts: one is the generation and embedding of recovery and authentication message at the transmitter site; the other is the detection and recovery of tampered blocks at the receiver site. Below we will explain these two parts in detail.

3.1. Generation and Embedding of Recovery and Authentication Message

Supposed that the original image I is an M×N gray-level image and is divided into four equal-sized regions as shown in figure 1. Every region is divided into nonoverlapping blocks of size 2 × 2 pixels. For each region, two blocks are randomly picked and grouped to calculate their average pixel values as recovery message. Then, the recovery message is split into three shares, each of which is embedded into one of the other three regions. For example, if the two blocks are located at region R1, then the three shares are embedded into R2, R3, and R4, respectively. The reason to embed shares to separated regions is to protect them from being destroyed simultaneously. Figure 2 illustrates the process of generating and splitting recovery message. Let block A and B denote the two randomly picked blocks from a region. Next, calculate the average pixel values of A and B, respectively. Let avg_A and avg_B denote the first five MSBs of the two average values. Then, construct the following polynomial:

\[ q(x) = \text{avg}_A + \text{avg}_B \mod 31. \]  

According to the above equation, the three shares Share 1, Share 2, and Share 3 are evaluated as q(1), q(2), and q(3), respectively. If block A and B are located at region Rm, share i will be embedded into two randomly picked blocks at region \( R_{(k+j) \mod 4} \). Supposed that block C and D denote as the two selected blocks at other region. Spreading out the eight bits of each pixel of two blocks, we can get the space as shown in Figure 3. Let \( s_j \) denote bit j of share i as shown in Figure 2. Then, the embedding position of each bit \( s_j \) is arranged as Figure 3. Note that legal range of coefficients in Eq.(3) is between 0 to 30, but the possible maximum of avg_A and avg_B is out of the range. When such situation happens, avg_A and avg_B will be modified to 30.

After the process of embedding recovery message is completed, Hsu and Tu’s scheme is used to generate 4-bit authentication message of each block of size 2 × 2 pixels. Each bit of the authentication message is embedded into the last bit of each pixel of the block self. Finally, we can get a watermarked image H". The whole process of generation and embedding is illustrated in Figure 4. The detail algorithm is as follows.

Input: A gray-level image I of size M×N pixels
Output: A gray-level image I’ with recovery and authentication message

Step 1: Split image I into four subimage \( R_1, R_2, R_3, \) and \( R_4 \). Let the matrix \( R \) denote the four subimages:

\[ R = \begin{bmatrix} R_1 & R_2 \\ R_3 & R_4 \end{bmatrix}. \]

Step 2: For each region \( R_k \), split \( R_k \) into nonoverlapping 2×2 blocks; then, scramble each two successive blocks randomly, where \( i = 1..4 \).

Step 3: Set \( i = 1 \).

Step 4: Set \( k = 1 \).

Step 5: Calculate the average pixel values \( b_k^i \) and \( b_{k+1}^i \) of block \( B_k^i \) and \( B_{k+1}^i \), respectively. Let

\[ \text{avg}_A = \begin{cases} b_k^i \gg 2 & \text{if } (b_k^i \gg 2) < 31 \\ 30 & \text{otherwise}, \end{cases} \]

and

\[ \text{avg}_B = \begin{cases} b_{k+1}^i \gg 2 & \text{if } (b_{k+1}^i \gg 2) < 31 \\ 30 & \text{otherwise}. \end{cases} \]

Step 6: Calculate three shares \( s_1, s_2, \) and \( s_3 \) by Eq.(3), where \( s_j = f(j) \) and \( j = 1..3 \).

Step 7: Set \( s_j = (s_{j0} s_{j1} s_{j2} s_{j3}), \) where \( j = 1..3 \). Let \((x_k^i, y_k^i)\) and \((x_{k+1}^i, y_{k+1}^i)\) denote the coordinates of blocks \( B_k^i \) and \( B_{k+1}^i \) in the subimage \( R_k \). According to the embedding rule of figure 3, embed \( s_j \) into the two blocks located at \((x_{k+j}^i \mod 3, y_k^i \mod 3, (i+j) \mod 3)\) and \((x_{k+1+j}^i \mod 3, y_{k+1}^i \mod 3, (i+j) \mod 3)\), respectively, where \( j = 1..3 \).

Step 8: Set \( k = k+2 \).

Step 9: Repeat Step 5 to Step 8 until \( k \geq (M/4) \times (N/4) \).

Step 10: Set \( i = i+1 \).

Step 11: Repeat Step 4 to Step 10 until \( i \geq 4 \).

Step 12: Rearrange all blocks according to the original order.

Step 13: For each 2×2 block of image I, generate 4-bit authentication message by means of Hsu and Tu’s scheme [Hsu & Tu, 2010]. Then, embed it to the last bit of each pixel. Output the image with recovery and authentication message I’.
scheme does not think this block needs to be fixed. But if it is a false negative, this block actually needs to be repaired. Therefore, decreasing false negative rate as much as possible can make most tampered block be recovered and hence increase the quality of the recovered image. Therefore, we proposed another stage of authentication to eliminate those isolated false in the matrix \( S \). Let \( W \) denote a window of size \( m \times m \), and \( t \) denote a predefined threshold. Moving and putting \( W \) on \( S \) from the up-left to the right-bottom corner, we count the number of false in the area falling within \( W \). If the number is less than \( t \), correct those false to true. After the window reaches the end of \( S \), the correction is completed, and we can get a modified matrix \( S' \).

According to \( S' \), we start stage-1 recovery process to repair each invalid block. For an invalid block, we examine its mapping blocks at other regions. If not all mapping blocks are invalid and any two shares can be retrieved successfully, the recovery message of this invalid block can be reconstructed. Supposed that the retrieved share \( y_i \) corresponds to argument \( x_i \), where \( i = 0, 1 \). Using Eq.(2), we can reconstruct the polynomial as shown in Eq.(3). Let \( c \) denote the coefficient corresponding to the invalid block. Then, shift \( c \) left by three bits to get the recovery value and replace every pixel of the invalid block with this new value. At the same time, correct the detected result of this invalid block from true to false.

After stage-1 recovery process is completed, the matrix \( S' \) is modified and denoted as \( S'' \). If there are some invalid blocks cannot be repaired at stage 1, we start stage-2 recovery process, which utilizes the neighboring blocks to repair the invalid block. Supposed that \( N_i \) denotes the eight blocks around an invalid block \( IB \), where \( i = 1, 2, \ldots, 8 \). Normally, there are total 12 pixels adjacent to \( IB \) within \( N_i \) to \( N_8 \). Calculate the average value of the 12 pixels, and replace every pixel of \( IB \) with this average value. Note that the average calculation excludes pixels of invalid neighboring blocks.

### IV. Experimental Results

In order to verify the performance of our method, we use the measurements PSNR (Peak Signal to Noise Ratio) to evaluate the similarity between the watermarked and recovered image to show the quality of the recovered image.

\[
PSNR = 10 \times \log \frac{255^2}{MSE} \text{ (dB)}
\]

where

\[
MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (p_{i,j} - p'_{i,j})^2
\]

In Eq. (3), \( p_{i,j} \) and \( p'_{i,j} \) are pixels located at \( (i, j) \) of the watermarked images and recovery image, respectively. Besides, the measurements FNR (False Negative Rate) and FPR (False Positive Rate) are used to evaluate the detect accuracy [Hsu & Tu, 2010]:

\[
FNR = FN/(FN + TP).
\]

\[
FPR = FP/(FP + TN).
\]
In Eq.(6) and Eq.(7), FN is the number of False Negative pixels, TP is the number of True Positive pixels, FP is the number of False Positive pixels, and TN is the number of True Negative Pixels. Consequently, tampering ratio can be defined as follows.

\[ \rho = \frac{(FN + TP)}{(M \times N)}. \]  

(8)

Figure 5 shows the experimental result of copy-and-paste attack of ours and Lee and Lin’s scheme. Observing the result, we can see that Lee and Lin’s scheme cannot detect the tampering area since their scheme uses only the last three bits of pixels for tampering detection. Besides, their authentication message does not include location information of blocks; hence, their scheme cannot sense the movement of blocks. Figure 6 shows the experiments, which simulate attacks with different tampering ratio. The results shows our scheme has low FNR and FPR, and most recovered images have good quality.

In our scheme, recovery message of each two blocks are generated and split into three shares, each of which is of 5-bit length. Simply speaking, the length of recovery message for each two blocks is of 15-bit length. The three shares are embedded into other two blocks, so two blocks covers 15-bit recovery message in our scheme. Therefore, the payload for each pixel is \(15/8 = 1.875\) bits per pixel. In our previous work [Qi & X.X., 2011], four 6-bit shares of the recovery messages are embedded into three \(2 \times 2\) blocks; hence, 2 bits are carried by a pixel averagely. In Lee and Lin’s scheme, 10-bit authentication message is embedded into two blocks, so the payload for each pixel is \(10/4 = 2.5\) bits per pixel. Compared to our previous work and Lee and Lin’s scheme, the proposed scheme has lower payload of each pixel.

<table>
<thead>
<tr>
<th>Watermarked Image</th>
<th>Tampered Image</th>
<th>Detect Result</th>
<th>Recovery Image</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ours</strong></td>
<td><img src="image1" alt="Image" /></td>
<td><img src="image2" alt="Image" /></td>
<td><img src="image3" alt="Image" /></td>
</tr>
<tr>
<td>(\rho = 7.51%)</td>
<td>FNR = 0.02012</td>
<td>FPR = 0.00336</td>
<td>PSNR = 36.5068</td>
</tr>
<tr>
<td><strong>Lee and Lin’s</strong></td>
<td><img src="image4" alt="Image" /></td>
<td><img src="image5" alt="Image" /></td>
<td><img src="image6" alt="Image" /></td>
</tr>
<tr>
<td>(\rho = 8.12%)</td>
<td>FNR = 0.98311</td>
<td>FPR = 0.00149</td>
<td>PSNR = 22.1138</td>
</tr>
</tbody>
</table>

Figure 5: Detect and Recovery Result of Copy-and-paste Attack
Figure 6: Experimental Results of Different Tampering Ratio

<table>
<thead>
<tr>
<th>Watermarked Image</th>
<th>Tampered Image</th>
<th>Detect Result</th>
<th>Recovery Image</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Watermarked Image" /></td>
<td><img src="image2" alt="Tampered Image" /></td>
<td><img src="image3" alt="Detect Result" /></td>
<td><img src="image4" alt="Recovery Image" /></td>
</tr>
<tr>
<td>ρ = 11.12%</td>
<td>FNR = 0.00288</td>
<td>FPR = 0.01643</td>
<td>PSNR = 29.2409</td>
</tr>
<tr>
<td>ρ = 13.96%</td>
<td>FNR = 0.00273</td>
<td>FPR = 0.00860</td>
<td>PSNR = 24.295</td>
</tr>
<tr>
<td>ρ = 6.19%</td>
<td>FNR = 0.01085</td>
<td>FPR = 0.01127</td>
<td>PSNR = 31.6334</td>
</tr>
<tr>
<td>ρ = 3.77%</td>
<td>FNR = 0.00850</td>
<td>FPR = 0.00531</td>
<td>PSNR = 37.9251</td>
</tr>
<tr>
<td>ρ = 4.56%</td>
<td>FNR = 0.00301</td>
<td>FPR = 0.00515</td>
<td>PSNR = 35.2196</td>
</tr>
</tbody>
</table>

V. CONCLUSIONS

Integrity authentication becomes an important issue for multimedia data. Many researchers proposed fragile watermarking scheme to resolve such issue. Some fragile watermarking scheme can only detect tampering, and some schemes can further recover the tampering area. In this paper, we proposed an image authentication and recovery scheme based on the concept of threshold secret sharing. Our scheme have two advantages: one is that the payload of each pixel of our scheme is lower than that of other researchers’ scheme;
the other is that the tampered block has second chance for recovery. In addition, the authentication message in our scheme includes the location information and pixel intensity of a block. Therefore, our scheme can be against copy-and-paste attack as shown in the experimental results.

Our scheme employs Hsu and Tu’s tampering detection method. Any image processing, such as cropping, compression, contrast adjust, …, etc. is recognized as a kind of attack. Recently, more and more researchers tend to not recognize image compression as a malicious attack and embed the watermark in the wavelet domain [Qi & X.X., 2011; Preda, 2013]. In the future, we expect to modify Hsu and Tu’s work to discriminate common image processing from malicious attacks.

ACKNOWLEDGEMENT

This work was supported in part by a grant from Chinese Culture University and from the National Science Council of the Republic of China under the projects NSC 102-2221-E-034-011-.

REFERENCES


Shu-Fen Tu was born in Taiwan in 1974. She received the BS degree in Management Information System from National Cheng-Chi University, Taiwan in 1996, the MS degree in Information Management from National Chia-Nan University, Taiwan in 1998, and the Ph.D. degree from the Institute of Information Management, National Central University, Taiwan in 2005. From 1998 to 1999, she is a software engineer of the Syscom Group Co., Taiwan. From February 2005 to July 2005, she is an assistant professor of Department of Information Management, Chaoyang University of Technology. Currently, she is an associate professor of Department of Information Management, Chinese Culture University, Taiwan. Her current research interests include steganography and secret sharing.

Ching-Sheng Hsu was born in Taiwan in 1971. He got his B.A. degree from the Department of Information Management, National Cheng-Chi University, Taiwan in 1994, M.A. degree from the Institute of Information Management, National Chia-Nan University, Taiwan in 1998, and Ph.D. degree from the Institute of Information Management, National Central University, Taiwan in 2005. From 1998 to 1999, he was a software engineer at the Syscom Group Co., Taiwan where his work focused on the Web-based stock trading systems. From 2000 to 2004, he was a part-time lecturer of the National Open University, Taiwan. Currently, he is an associate professor of Department of Information Management, Ming Chuan University. His current research interests include digital watermarking and information hiding, visual cryptography, optimization algorithms, and intelligent computer-assisted learning and testing systems.

Fu-Hsing Wang received the Ph.D. degree in Information Management from the National Taiwan University of Science and Technology. From 1992 to 2005, he was with the Chungyuan college, Taiwan. Since 2005, he has been with the faculty of the Department of Information Management, Chinese Culture University, Taipei, where he is currently a Professor. His research interests include distributed and stabilizing algorithm, graph theory, and interconnection networks.