An Efficient Semi-fragile Watermarking Scheme for Tamper Localization and Recovery

Xiang Hou¹, Hui Yang¹, Lianquan Min¹,*

¹Information Engineering University, Zhengzhou, China

*Corresponding author e-mail: rainman_mlq@163.com

Abstract. To solve the problem that remote sensing images are vulnerable to be tampered, a semi-fragile watermarking scheme was proposed. Binary random matrix was used as the authentication watermark, which was embedded by quantizing the maximum absolute value of directional sub-bands coefficients. The average gray level of every non-overlapping 4×4 block was adopted as the recovery watermark, which was embedded in the least significant bit. Watermarking detection could be done directly without resorting to the original images. Experimental results showed our method was robust against rational distortions to a certain extent. At the same time, it was fragile to malicious manipulation, and realized accurate localization and approximate recovery of the tampered regions. Therefore, this scheme can protect the security of remote sensing image effectively.

1. Introduction

As an important carrier of geospatial information, remote sensing images play an increasingly crucial role in numerous fields. The security of remote sensing images has really come to the forefront. Semi-fragile watermarking is mainly designed for content authentication. Different from the characteristics of completely fragile watermarking [1, 2], semi-fragile watermarking is able to achieve selective authentication. Hence, semi-fragile watermarking is an important technical means to ensure the data security.

Compared with the vigorous development of robust watermarking [3, 4], semi-fragile watermarking for remote sensing image remains to be prosperous. Caldelli et al.[5] divided the host images into blocks and inserted watermark into least significant bit (LSB). Ho et al.[6] employed the texture information of host remote sensing image as authentication watermark. However, the localization accuracy for tampered regions needed to be improved. Jiang et al.[7] designed a method in DCT domain, which modified an odd number of AC coefficients randomly in the low frequency to embed watermark. Nevertheless, it did not have the capability of correcting tampered regions. Ruiz et al.[8] employed the iterative approach to modify the low frequency coefficients until meeting the specific requirements. However, it had to save specific key for each sub-block as auxiliary information.

A semi-fragile watermarking method with tamper localization and recovery capability is proposed in this paper. The remainder of our paper is organized as the following order. In Sec. 2, we present the designed semi-fragile watermarking technique. Sec. 3 shows experimental results and performance of our method. In Sec. 4, we make a conclusion about this paper.
2. Proposed Scheme

2.1. Watermark Generation
Authentication watermark is designed for locating tampered regions, while recovery watermark is designed for restoring the tampered content. The scheme uses secret keys to generate pseudo random binary matrix as the authentication watermark \( w_s \). Then, the generating stage of recovery watermark \( w_s \) takes place in the following steps:

Step 1: The original carrier is divided into four equal Blocks A, B, C, D and then each is further divided into several \( 4 \times 4 \) sub-blocks. Next, the average gray value of every non-overlapping \( 4 \times 4 \) block is calculated.

Step 2: The rounded integral average gray level of each \( 4 \times 4 \) sub-block is represented by an 8-bit binary sequence, so the recovery watermarks of Blocks A, B, C and D can be generated. Then the Baker Transform is applied to scramble the four recovery watermarks, and the processed watermarks are represented as \( w_a', w_b', w_c', w_d' \). Combine them together, and the final recovery watermark \( w_s' \) of the whole image can be generated.

2.2. Watermark Embedding
Watermarks are embedded according to the detailed steps as follows:

(1) Set the LSB of every pixel in original remote sensing image \( I \) to zero.

(2) Here, the processed image \( I \) is applied to Contourlet Transform (CT). Let \( C_k(i,j) \) be the coefficients in the location \((i,j)\) in directional sub-bands of the \((k-1)\)th level \((k\) is the direction of the sub-band). Select \( |C_k(i,j)| \) that has the maximum value of all sub-bands, and embeds a single bit of \( w_k(i,j) \) by modifying \( C_k(i,j) \) as the following:

\[
g(i,j) = \begin{cases} 0, & \text{if floor} \left[ \frac{C_k(i,j)}{Q} \right] \text{ is even} \\ 1, & \text{if floor} \left[ \frac{C_k(i,j)}{Q} \right] \text{ is odd} \end{cases}
\]

\[
C_k'(i,j) = \begin{cases} \text{sign}(C_k(i,j)) \cdot (\text{floor} \left[ \frac{C_k(i,j)}{Q} \right] \cdot Q + 2Q), & \text{if } g(i,j) = W(i,j) \\ \text{sign}(C_k(i,j)) \cdot (\text{floor} \left[ \frac{C_k(i,j)}{Q} \right] \cdot Q), & \text{if } g(i,j) \neq W(i,j) \end{cases}
\]

where \( Q \) is the quantization step, \( \text{sign}() \) is the sign function. Repeat this operation until all bits of the authentication watermark \( w_s \) have been embedded.

(4) Apply the inverse CT to obtain the image containing authentication watermark \( w_s \). Set the LSB of each pixel in this image to zero, and obtain the processed image \( I' \).

(5) \( I' \) is divided into blocks. The scrambled recovery watermarks \( w_{a'}, w_{b'}, w_{c'}, w_{d'} \) of Block A, B, C and D, are embedded in the LSB of Block B, C, D and A respectively. Then the image \( I' \) containing the authentication and recovery watermark can be obtained.

2.3. Watermark Detection
Detailed process is described in the following:

(1) Extract the LSB of each pixel in the to-be-checked image \( I' \), and apply inverse scrambling to obtain the recovery watermark \( w_s' \) of the image.

(2) Set the LSB of each pixel in image \( I' \) to zero, and obtain the processed image \( I'' \).

(3) Perform CT on \( I'' \) the same way employed in the embedding process. To extract the bit at the position \((i,j)\), \( C_k'(i,j) \) with maximum absolute value is quantized using the following equation:

\[
w_k'(i,j) = \begin{cases} 1, & \text{if } \text{mod} \left( \text{round} \left( C_k'(i,j) \right) / Q \right), 2 = 1 \\ 0, & \text{if } \text{mod} \left( \text{round} \left( C_k'(i,j) \right) / Q \right), 2 = 0 \end{cases}
\]

Repeat this operation until all bits of authentication watermark \( w_s' \) are extracted.
2.4. Tamper Localization and Recovery

The detected watermark $w_\mathcal{A}'$ and the original authentication watermark $w_\mathcal{A}$ are compared bit by bit according to the following equation:

$$D(i, j) = |W_\mathcal{A}'(i, j) - W_\mathcal{A}(i, j)|$$  \hspace{1cm} (4)

where $D$ is the tamper detection matrix. The 0 and 1 mean the match and mismatch between extracted watermark and original authentication watermark, respectively.

Then we use the following strategy to process Matrix $D$. If $D(i, j) = 1$ and the number of non-zero pixels in its eight neighbors is less than or equal to 1, we restore $D(i, j)$ to 0. If $D(i, j) = 0$ and the number of non-zero pixels in its eight neighbors is greater than or equal to 4, we restore $D(i, j)$ to 1. If 0 and 1 show a high density alternative distribution, then this overall region is judged to be tampered, the above strategy for processing is not adopted. Finally, authentication result is amended appropriately combining with the morphological operators.

After the tampered region of image is determined, the extracted recovery watermark $w_\mathcal{A}'$ is restored to gray level and magnified by four times using bicubic interpolation to obtain approximate images. The tampered region is replaced by the content of corresponding position in the approximate image to complete the recovery procedure.

3. Experimental Results and Analysis

Various simulation experiments are shown in this section to evaluate the performance of our algorithm, using panchromatic remote sensing image with a size of $512 \times 512$ pixels. The applied CT adopts 9-7 pyramid filter and pkva directional filter.

3.1. Robustness to Rational Distortions

Various kinds of noise with different intensity are added and JPEG compressions with different quality factors (QF) are performed respectively. Normalized correlation (NC) is employed to assess the similarity.

| Parameter | Gaussian noise | Salt and pepper noise | JPEG compression |
|------------|----------------|-----------------------|------------------|
| Parameter  | NC 0.97 6      | Parameter 0.005       | QF 90 0.992      |
| 8          |                | 0.968                 |                  |
| 16         | 0.93 9         | 0.01 0.931            | 80 0.973         |
| 32         | 0.86 1         | 0.02 0.873            | 70 0.955         |

Table 1 shows that watermark can still be extracted even though the watermarked remote sensing images undergo a certain intensity of noise attack and JPEG compression. This proves that the proposed scheme has certain robustness to rational distortions.

3.2. Tampered Regions Localization and Recovery

In order to demonstrate the performance of tampered regions localization and recovery, various malicious attacks are performed respectively on the watermarked remote sensing images. Fig.1(a), Fig.1(d) and Fig.1(g) are tampered remote sensing images. Fig.1(b), Fig.1(e) and Fig.1(h) are the tamper detection results, respectively. Similarly, the recovered images are presented in Fig.1(c), Fig.1(f) and Fig.1(i), respectively.
As is proved in Fig. 1, our method is fragile to malicious tampering attacks. It can discover the tampered region accurately and restore the tampered content approximately. The restored remote sensing image can meet the basic visual requirement and can be applied efficiently when the request for data accuracy is not too rigid.

3.3. Data Accuracy Analysis
In order to detect the effects of embedded watermarks on data accuracy, we compare and analyze the statistical information before and after embedding the watermarks.

|                              | Original remote sensing image | Watermarked remote sensing image |
|------------------------------|-------------------------------|----------------------------------|
| minimum gray value           | 49                            | 44                               |
| maximum gray value           | 255                           | 255                              |
| average gray value           | 115.902                       | 115.373                          |
| standard deviation           | 36.586                        | 36.971                           |

As is presented in Table 2, the changes of gray value extrema in the watermarked remote sensing image are five gray levels, moreover, the changes of average gray value and standard deviation are miniscule compared with the original image. Thus, the results demonstrate that watermarks embedding have little effect on data accuracy.

4. Conclusion
In this paper, a semi-fragile watermarking scheme of satellite image for tamper localization and recovery has been proposed. This scheme is not only able to distinguish the reasonable distortion from malicious manipulation exactly, but also achieves accurate localization and approximate recovery of
the tampered regions. Therefore, the method is capable of protecting remote sensing images security effectively. Nevertheless, how to reduce the effect on data accuracy as much as possible remains to be further studied.

Acknowledgement
This work was supported by the National Natural Science Foundation of China (Project No. 41471337), State Key Laboratory of Geo-Information Engineering Foundation of China (Project No. SKLGIE2014-M-4-6).

References
[1] L. Teng, X. Y. Wang, X. K. Wang, “Cryptanalysis and improvement of a chaotic system based fragile watermarking scheme,” Int. J. Electron. Commun. 67, 540-547 (2013)
[2] J. P. Zhang, Q. F. Zhang, H. L. Lv, “A novel image tamper localization and recovery algorithm based on watermarking technology,” Optik 124, 6367-6371 (2013)
[3] B. P. Kumari, V. P. S. Rallabandi, “Modified patchwork-based watermarking scheme for satellite imagery,” Signal Process. 88, 891-904 (2008)
[4] P. Zhu, F. Jia, J. L. Zhang, “A copyright protection watermarking algorithm for remote sensing image based on binary image watermark,” Optik 124, 4177-4181 (2013)
[5] R. Caldelli, G. Macaluso, M. Barni, “Joint near-lossless watermarking and compression for the authentication of remote sensing images,” in: Proceedings of IEEE International Geoscience and Remote Sensing Symposium, pp. 297-300 (2004)
[6] A. T. S. Ho, X. Zhu, W. M. Woon. “A semi-fragile pinned sine transform watermarking system for content authentication of satellite images,” in: Proceedings of IEEE International Geoscience and Remote Sensing Symposium, pp. 737-740 (2005)
[7] L. Jiang, Z. Q. Xu, “DCT semi-fragile watermarking algorithm for remote sensing image,” Journal of Huazhong University of Science and Technology 40, 47-51 (2012)
[8] J. S. Ruiz, D. Megias, “A novel semi-fragile forensic watermarking scheme for remote sensing images,” Int. J. Remote Sens. 32(19), 5583-5606 (2011)