Cryptanalysis and improvement of a chaotic system based fragile watermarking scheme

Article *in* AEU - International Journal of Electronics and Communications · June 2013

DOI: 10.1016/j.aeue.2012.12.001

3 authors, including:

Xingyuan Wang
Dalian University of Technology
300 PUBLICATIONS **4,629** CITATIONS

SEE PROFILE

Int. J. Electron. Commun. (AEÜ) 67 (2013) 540–547

Contents lists available at SciVerse ScienceDirect

International Journal of Electronics and Communications (AEÜ)

journal homepage: www.elsevier.com/locate/aeue

Cryptanalysis and improvement of a chaotic system based fragile watermarking scheme

Lin Teng, Xingyuan Wang*, Xiukun Wang

Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China

ARTICLE INFO

Article history:
Received 20 January 2012
Accepted 1 December 2012

Keywords:
Fragile watermarking
Chaotic system
Attack
Improvement

ABSTRACT

In this paper, we analyze the security of a chaotic system based fragile watermarking scheme for image tamper detection proposed by Rawat et al. recently. Some errors and modification attack against Rawat et al.'s scheme are demonstrated. Both theoretical analysis and experimental results show that the fragile watermarking scheme is not security. Besides, improvement measure is presented to enhance the security of the fragile watermarking scheme.

1. Introduction

With the rapid growth of internet, a large amount of digital data is easily accessible and the copyright of digital content against piracy and malicious manipulation is more and more important. Digital watermarking has been proposed as an efficient way to protect the copyright. Digital watermarking is the means of embedding information into digital data which is imperceptible to human observer but easily detected by computer algorithm. It has been branched into two classifications: robust watermarking technique and fragile watermarking technique. Robust digital watermarking is used to protect ownership of the digital media. Fragile watermarking is to detect any possible modification of the pixel values. Many fragile watermarking schemes are proposed [1–4]. In recent years, chaotic maps have attracted more attention to develop digital watermark due to the properties such as sensitivity to initial conditions and random-like behavior. Chaotic maps have been used for digital watermark to increase the security [5–7]. Digital watermarks can be classified into two categories: spatial domain [7–9] and frequency domain [4,10,11] based. The former can embed a large number of bits without incurring noticeable visual artifacts; whereas, the latter has been shown to be quite robust.

Rawat et al. [8] recently proposed a fragile watermarking scheme aimed to detect image tamper areas. They use chaotic systems to disturb watermark and corresponding position between pixels in the watermarked image. The main defect of this scheme is to embed the watermark in the least significant bit plane of the host image, which is easily obtained and replaced. In this paper, we propose a novel method to modify the watermarked image which the scheme will not be able to detect the tampered areas. And improvement measure is presented to enhance the security of the fragile watermarking scheme.

The rest of this paper is organized as follows. Section 2 is a brief introduction to Rawat et al.'s fragile watermarking scheme based on chaotic system. Section 3 analyzes the weakness of the scheme and attacks it. Section 4 gives improvement measure to improve the scheme. In the last section, we conclude this paper.

2. Rawat et al.'s fragile watermarking scheme

The fragile watermarking scheme is composed of two processes: watermark embedding and watermark extraction. Firstly, the watermark is embedded as follows:

(1) Scramble the original image I of size $M \times N$, using Arnold cat map k times, the result denoted by I_{scr}.
(2) Divide the scrambled image I_{scr} into 8-bit planes.
(3) Generate a chaotic sequence S of length $m \times n$ using logistic map. Round off the chaotic sequence and rearrange to get the chaotic image pattern S_{cp}.
(4) Obtain a binary chaotic watermark W_p using exclusive-or operation between the watermark W of size $m \times n$ and S_{cp} as $W_p = S_{cp} \oplus W$.
(5) Replace the least significant bit plane of I_{scr} by W_p.

* Corresponding author. Tel.: +86 13889431686.
E-mail addresses: tenglin@mail.dlut.edu.cn (L. Teng), wangxy@dlut.edu.cn (X. Wang).

1434-8411/$ – see front matter © 2012 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.aeue.2012.12.001

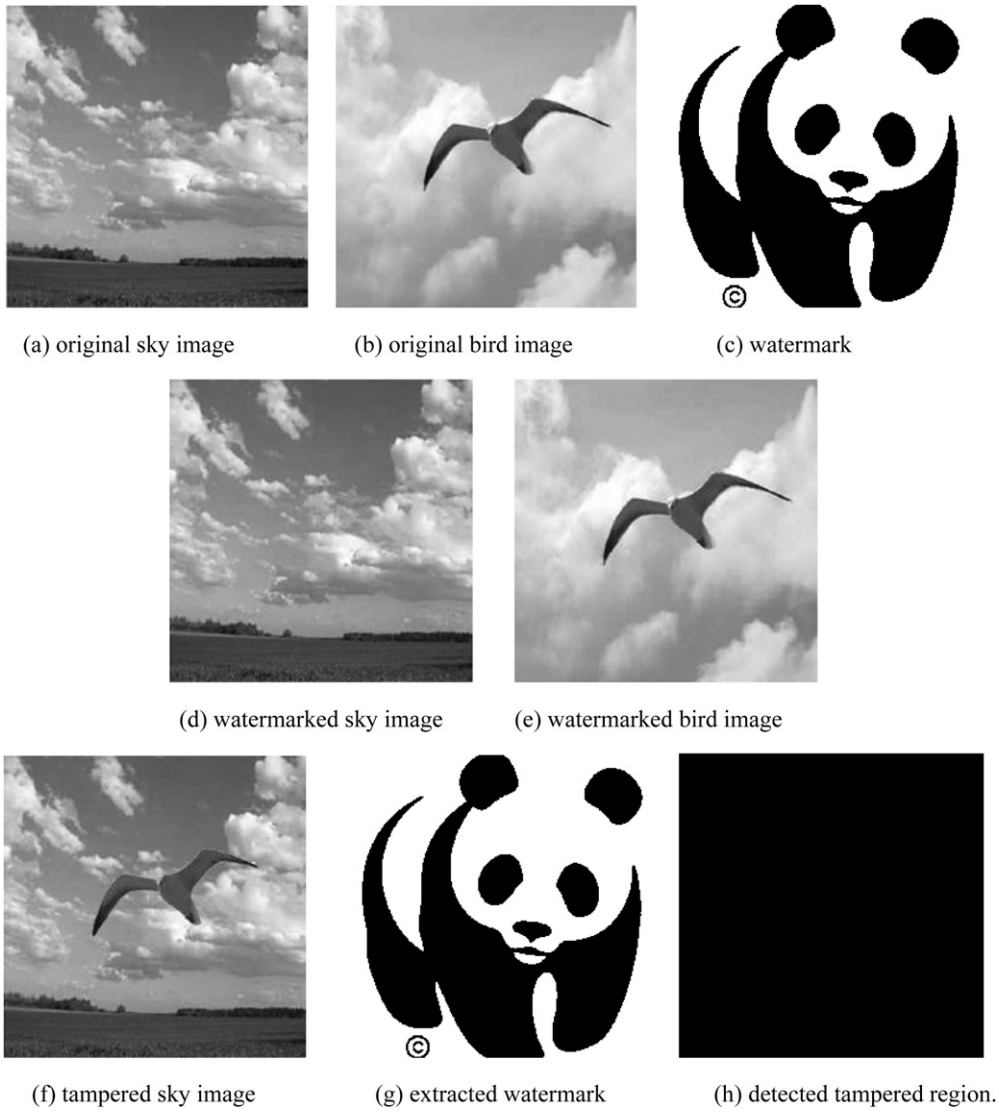


Fig. 1. Collage attack. (a) Original sky image, (b) original bird image, (c) watermark, (d) watermarked sky image, (e) watermarked bird image, (f) tampered sky image, (g) extracted watermark and (h) detected tampered region.

(6) Apply Arnold cat map $T-k$ times on modified I_{scr} to get the watermarked image, where T is the period of cat map.

The watermark is extracted as follows:

(1)–(4) Process watermarked image I_W similar as the step (1) to step (4) in the watermark embedding process to get the extracted watermark W^{ext}.
(5) Take the absolute difference of extracted watermark W^{ext} and original watermark W. Apply Arnold cat map $T-k$ times to locate the tampered areas of the watermarked image.

For a detail explanation of this scheme, please refer to Ref. [8].

3. The cryptanalysis and attack

According to Kerchoff's principle [12], when analyzing an encryption algorithm, an assumption is that the cryptanalyst knows exactly the design and working of the cryptosystem. Namely, cryptanalyst knows everything about the cryptosystem except for the secret keys.

In this paper, we attack the fragile watermarking scheme without knowing the secret keys. That is, we propose a novel method to modify the watermarked image which the scheme will not be able to detect it.

3.1. Mistakes in the scheme

There are some errors in Rawat et al.'s scheme:

First, the cat map is used to scramble the host image, but the cat map is only suitable for square image, the size of host image and the watermark image should be $N \times N$. Moreover, in order to detect all regions of the watermarked image, the watermark should be the same size of host image, so the size of watermark image is $N \times N$.

Second, the author claimed that the scheme can resist collage attack. But, same secret keys will generate same chaotic watermark. When a counterfeit image is formed by combining the portions of multiple water marked images with the same watermark while preserving their relative spatial locations within the target image, the watermark in counterfeit image is exactly the same as the watermark in original watermarked image. So, the scheme will not detect the tampered areas and cannot resist collage attack if the

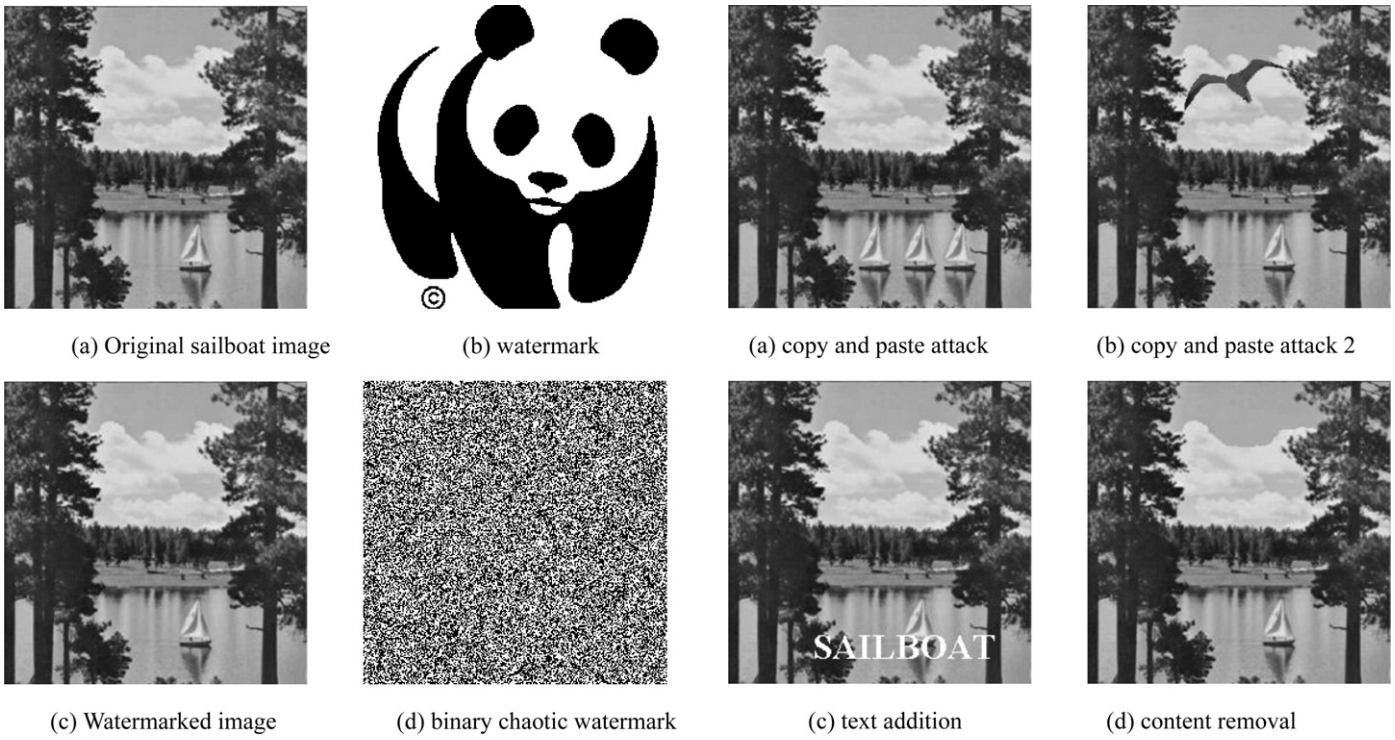


Fig. 2. Watermark embedded use Rawat et al.'s scheme. (a) Original sailboat image, (b) watermark, (c) watermarked image and (d) binary chaotic watermark.

Fig. 3. Tampered images use our scheme. (a) Copy and paste attack, (b) copy and paste attack 2, (c) text addition and (d) content removal.

secret keys remain unchanged. The experiment result is shown in Fig. 1.

3.2. Method to modify the watermarked image

The method to modify the watermarked image is described in detail below.

Step 1. Save the least significant bit plane of the original watermarked image I_W as W_o.
Step 2. Modify the watermarked image.
Step 3. Replace the least significant bit plane of the tampered watermarked image with W_o and get the fake watermarked image I_f.

Then use the Rawat et al.'s scheme to extract the watermark from image I_f, the extracted watermark is the same as the original watermark. That is, in step 5, the absolute difference of extracted watermark and original watermark is zero. So, the Rawat et al.'s scheme is not able to detect the modification and locate the tampered areas.

We use the same keys as the Rawat et al.'s scheme to embed the watermark. The experiments results are shown below. Fig. 2 shows the host image, watermark, the corresponding watermarked image obtained by Rawat et al.'s scheme and the extracted binary chaotic watermark in watermarked image. The tampered image after replace the least significant bit plane with binary chaotic watermark in Fig. 1(d) are shown in Fig. 3. Fig. 4 shows the results of detected tampered region of Fig. 3 use Rawat et al.'s scheme. From Fig. 4 we can see that the Rawat et al.'s scheme is not able to detect the tampered areas of different types of modified images which processed by our method.

4. Improvement measure

The main drawback in Rawat et al.'s scheme is to embed the scrambled watermark in the least significant bit (LSB) plane of the image which is easily obtained and replaced. Moreover, the scheme is not dependent on the plain image, which would increase the possibility of attacks. In order to overcome the above drawbacks

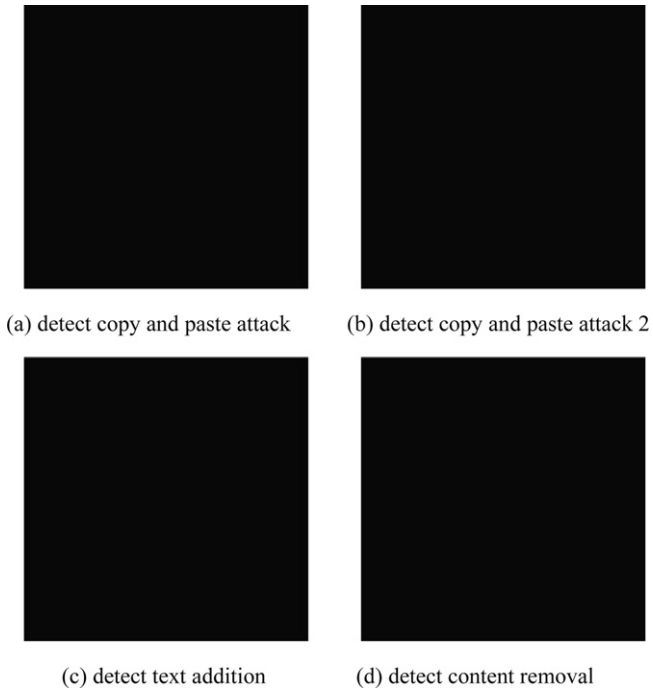
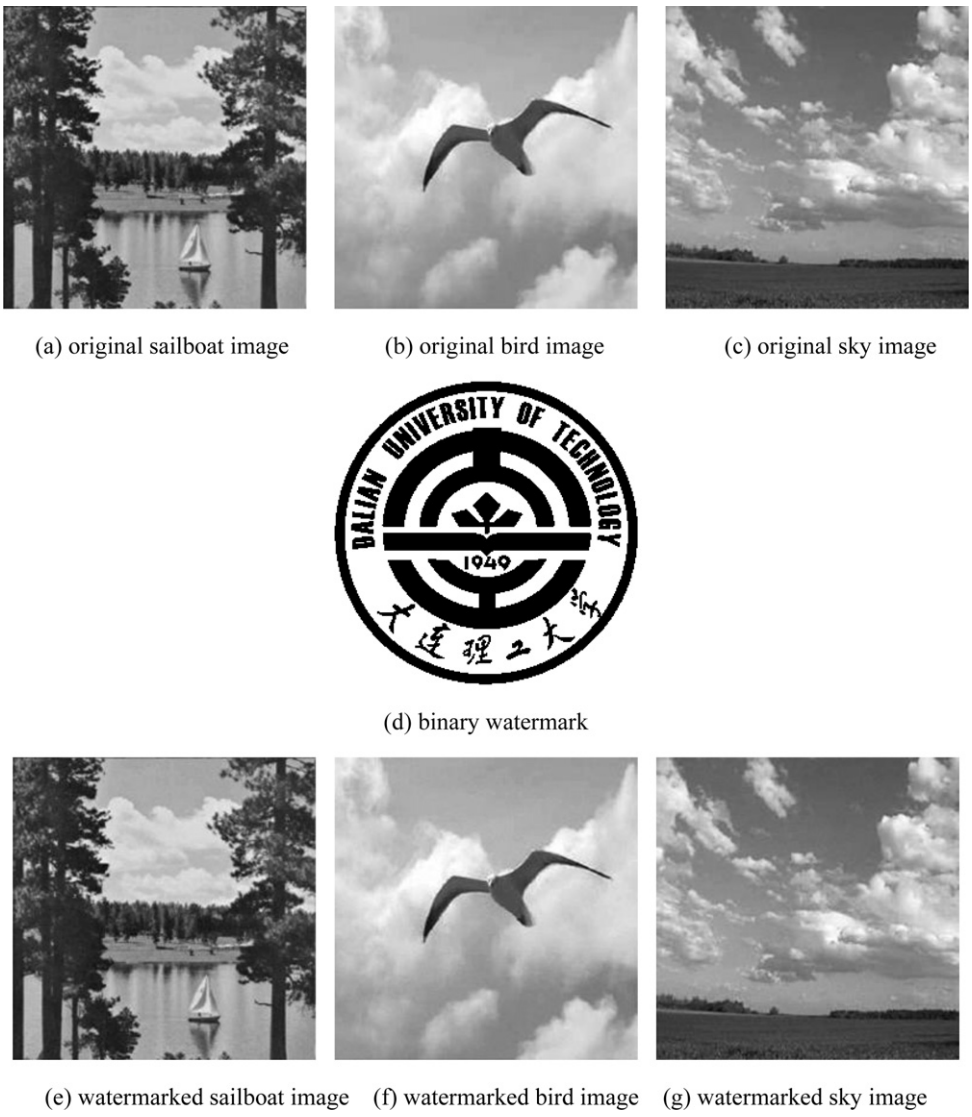


Fig. 4. Detected tampered region of Fig. 2. (a) Detect copy and paste attack, (b) detect copy and paste attack 2, (c) detect text addition and (d) detect content removal.



(a) original sailboat image (b) original bird image (c) original sky image

(d) binary watermark

(e) watermarked sailboat image (f) watermarked bird image (g) watermarked sky image

Fig. 5. Watermark embedded. (a) Original sailboat image, (b) original bird image, (c) original sky image, (d) binary watermark, (e) watermarked sailboat image, (f) watermarked bird image and (g) watermarked sky image.

and to improve security, we propose the improvement measure here.

4.1. Watermark embedding

The watermark is embedded as follows:

(1) Scramble the original image I of size $N \times N$, using Arnold cat map k times, the result denoted by I_{scr}.

(2) Divide the scrambled image I_{scr} into 8-bit planes. From the most significant bit plane I^8_{scr} to the least significant bit plane I^1_{scr}.

A bit can contain different amounts of information depending on its position in the pixel. The higher 4 bits (8th, 7th, 6th and 5th) carry 94.125% of the total information of the image. On the other hand, the lower 4 bits (4th, 3rd, 2nd and 1st) carry less than 6% of the image information.

(3) Generate a chaotic sequence S of length $N \times N$ using logistic map. Round off the chaotic sequence and rearrange to get the chaotic image pattern S_{cp}. Generate an integer sequence Z of length $N \times N$ through S use the following formula:

$$Z = floor(S \times 10^{14}) \bmod 3 \quad (3)$$

where $floor(a)$ returns the largest integer smaller than a, values of Z are integers range from 0 to 2. Rearrange integer sequence Z to get the pattern Z_p corresponding with the image.

(4) Obtain a binary chaotic watermark W_p using exclusive-or operation between the watermark W of size $N \times N$, S_{cp} and I^8_{scr}, I^7_{scr} as $W_p = S_{cp} \oplus W \oplus I^8_{scr} \oplus I^7_{scr}$.

The watermark W_p is dependent on the plain host image, so different host image generate different watermark and any modification on the watermarked image will affect the watermark even if the watermark is obtained before.

(5) Replace the lower three bit planes of I^1_{scr}, I^2_{scr} or I^3_{scr} by W_p according to Z_p. If $Z_p = 0$, replace the correspondence position bit in I^1_{scr}; if $Z_p = 1$, replace the correspondence position bit in I^2_{scr}; if $Z_p = 1$, replace the correspondence position bit in I^3_{scr}.

In Rawat et al.'s scheme, they only embed the scrambled watermark in the least significant bit plane of the image which is easily obtained and replaced. To overcome this drawback, we embed the W_p into the lower three bit planes according to chaotic pattern. So, the plane to embed each watermark pixel is not predictable.

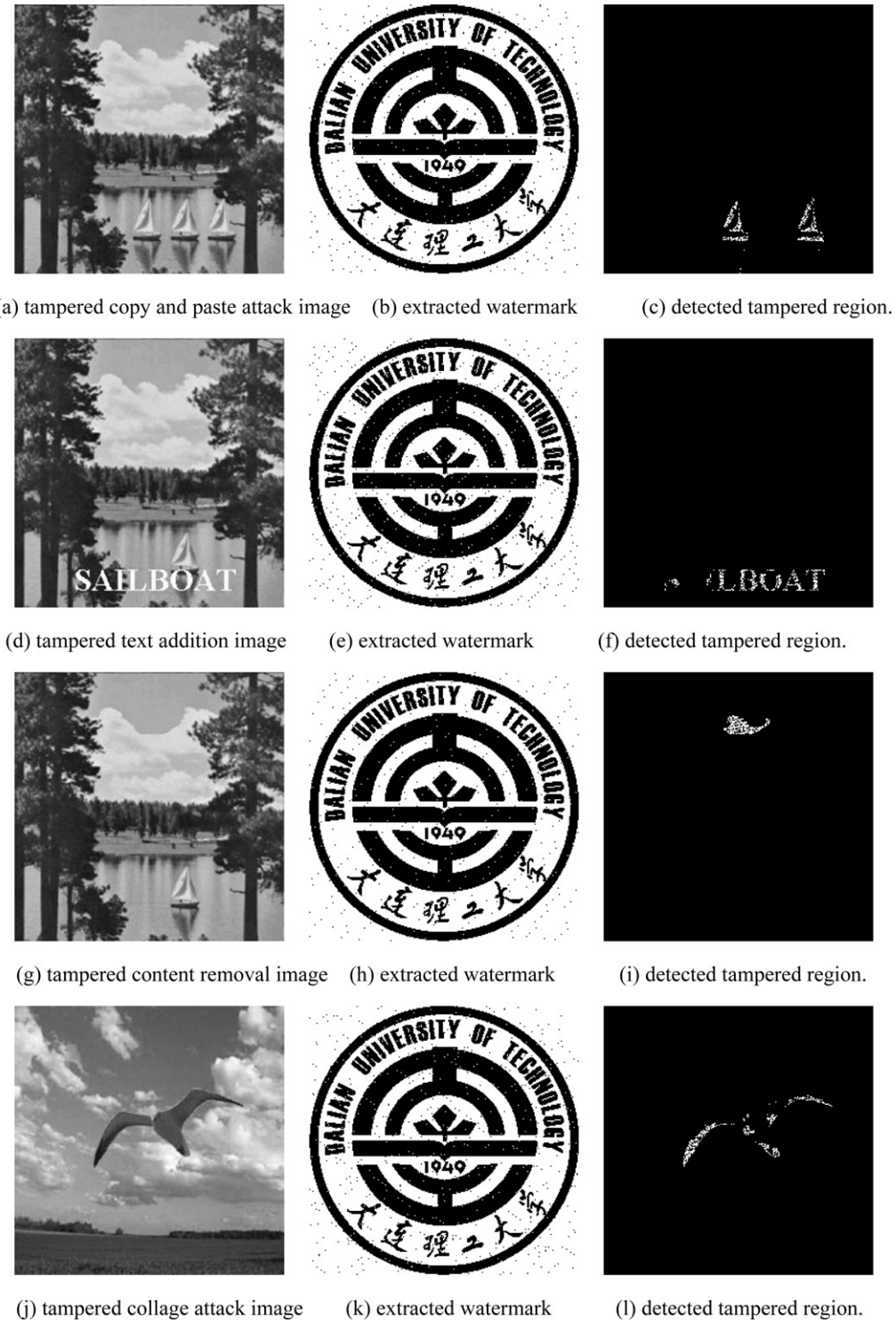
(6) Apply Arnold cat map $T - k$ times on modified I_{scr} to get the watermarked image, where T is the period of cat map.

Fig. 6. Copy and paste attack. (a) Tampered image, (b) extracted watermark and (c) detected tampered region.

Fig. 7. Text addition. (a) Tampered image, (b) extracted watermark and (c) detected tampered region.

Fig. 8. Content removal. (a) Tampered image, (b) extracted watermark and (c) detected tampered region.

Fig. 9. Collage attack. (a) Tampered image, (b) extracted watermark and (c) detected tampered region.



(a) tampered copy and paste attack image (b) extracted watermark (c) detected tampered region.

(d) tampered text addition image (e) extracted watermark (f) detected tampered region.

(g) tampered content removal image (h) extracted watermark (i) detected tampered region.

(j) tampered collage attack image (k) extracted watermark (l) detected tampered region.

Fig. 10. Detect tampered areas after modified by our attack. (a) Tampered copy and paste attack image, (b) extracted watermark, (c) detected tampered region, (d) tampered text addition image, (e) extracted watermark, (f) detected tampered region, (g) tampered content removal image, (h) extracted watermark, (i) detected tampered region, (j) tampered collage attack image, (k) extracted watermark and (l) detected tampered region.

4.2. Watermark extraction

The watermark is extracted as follows:

(1) Scramble the watermarked image I_w, using Arnold cat map k times. Let us denote the result by I_{Wscr}.
(2) Divide the scrambled watermarked image I_{Wscr} into 8-bit planes.
(3) Obtain the same chaotic image pattern S_{cp} and Z_p as in step 3 of embedding algorithm.
(4) Extract the binary chaotic watermark W_p according to Z_p similar as in step 4 of embedding algorithm.
(5) Apply exclusive-or operation between binary chaotic watermark W_p, chaotic image pattern S_{cp} and I^8_{scr}, I^7_{scr} as $W_p = S_{cp} \oplus W \oplus I^8_{scr} \oplus I^7_{scr}$ to get the extracted watermark W^{ext}.

(6) Take the absolute difference of extracted watermark W^{ext} and original watermark W. Apply Arnold cat map $T-k$ times to locate the tampered areas of the watermarked image.

4.3. Experimental results

Various experiments are carried out to assess the performance of the improvement algorithm. In our scheme, different host images should use different secret keys. A binary logo of size 256×256 is used as watermark in all the experiments. The host images are "sailboat", "sky", "bird" of size 256×256. The parameters of Arnold cat map are $a = 1$, $b = 1$ and $k = 75$. The parameters of logistic map are chosen as $\mu = 3.845$ and $x(0) = 0.654$ for "sailboat"; $\mu = 3.846$ and $x(0) = 0.654$ for "sky"; $\mu = 3.845$ and $x(0) = 0.655$ for "bird". PSNR (peak signal-to-noise ratio) is used in this paper to analyze the visual quality of the watermarked image in comparison with the original image I. PSNR is defined as:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{255^2}{MSE}\right) dB \tag{4}$$

where MSE is the mean squared error between the original image I and the attacked image $\hat{I}$, given by

$$\mathrm{MSE} = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[I(i,j) - \hat{I}(i,j)]^2 \tag{5}$$

Fig. 5 shows the host image, binary watermark and the corresponding watermarked image. The PSNR value of watermarked image Fig. 5(e), (f) and (g) are 42.3922, 42.7101 and 42.7049, which is better than the PSNR value in [5,7]. From Fig. 5 we can see that the watermark is imperceptible to human observer.

4.3.1. Performance under copy and paste attack

In the copy and paste attack, the watermarked sailboat image is modified by inserting two more boats in the image, where the boats are copied from the same watermarked image. The tampered image is shown in Fig. 6(a). Fig. 6(b) shows the extracted watermark from Fig. 6(a). The tamper detection result is shown in Fig. 6(c).

4.3.2. Performance under text addition

In this experiment, the watermarked image, shown in Fig. 7(a) is modified by adding the text "SAIL BOAT" at the bottom of the image. Extracted watermark from Fig. 7(a) is shown in Fig. 7(b). Detected tampered region is shown in Fig. 7(c).

4.3.3. Performance under content removal

In this experiment, some content of the watermarked image is removed without degrading the image quality. We have removed some portion of the cloud from the watermarked image. The tampered image is shown in Fig. 8(a). Fig. 8(b) shows the extracted watermark from Fig. 8(a). The tamper detection result is shown in Fig. 8(c).

4.3.4. Performance under collage attack

The counterfeit image, as shown in Fig. 9(a) was constructed by copying the bird from Fig. 5(b) and inserting it in relative spatial location in sky from Fig. 5(c). Fig. 9(b) shows the extracted watermark from Fig. 9(a) and Fig. 9(c) shows the detected tampered region.

4.4. Attack on our improved scheme

We use similar attack method as which used in Section 3.2 to analyze the performance of our improved scheme. The method is described in detail below.

Step 1. Save the lower three bit planes of the original watermarked image I_W as W_0^1, W_0^2 and W_0^3.
Step 2. Modify the watermarked image.
Step 3. Replace the lower three bit planes of the tampered watermarked image with W_0^1, W_0^2 and W_0^3 and get the fake watermarked image I_f.

Then use our improved scheme to extract the watermark from image I_f. The experiments results are shown in Fig. 10. From Fig. 10 we cam see that our improved scheme can resist the attack proposed above.

5. Conclusion

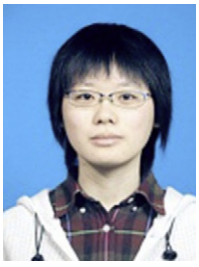
In this paper, we analyze the fragile watermarking scheme for image tamper detection proposed by Rawat et al. recently. Cryptanalysis shows that in Rawat et al.' scheme, the chaotic watermark can be easily obtained and be replaced, and modification attack is performed successfully, through which the scheme could not be able to detect the tamper. Besides, improvement measure is proposed to enhance security.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (Nos.: 61173183, 60973152, and 60573172), the Superior University Doctor Subject Special Scientific Research Foundation of China (No.: 20070141014), Program for Liaoning Excellent Talents in University (No.: LR2012003), the National Natural Science Foundation of Liaoning Province (No.: 20082165) and the Fundamental Research Funds for the Central Universities (No.: DUT12JB06).

References

[1] Liu SH, Yao HX, Gao W, Liu YL. An image fragile watermark scheme based on chaotic image pattern and pixel-pairs. Appl Math Comput 2007;185:869–82.
[2] Guo HP, Li YJ, Liu AY, Jajodia S. A fragile watermarking scheme for detecting malicious modifications of database relations. Inform Sci 2006;176:1350–78.
[3] Zhang XP, Wang SZ. Fragile watermarking scheme using a hierarchical mechanism. Signal Process 2009;89:675–9.
[4] Aslantas V, Ozer S, Ozturk S. Improving the performance of DCT-based fragile watermarking using intelligent optimization algorithms. Opt Commun 2009;282:2806–17.
[5] Chang CC, Chen KN, Lee CF, Liu LJ. A secure fragile watermarking scheme based on chaos-and-hamming code. J Syst Softw 2011;84:1462–70.
[6] Wu YT, Shih FY. Digital watermarking based on chaotic map and reference register. Pattern Recogn 2007;40:3753–63.
[7] Wu XY, Guan ZH. A novel digital watermark algorithm based on chaotic maps. Phys Lett A 2007;365:403–6.
[8] Rawat S, Raman B. A chaotic system based fragile watermarking scheme for image tamper detection. Int J Electron Commun 2011;65:840–7.
[9] Chang CC, Hu YS, Lu TC. A watermarking-based image ownership and tampering authentication scheme. Pattern Recogn Lett 2006;27:439–46.
[10] Bhatnagar G, Raman B. A new robust reference logo watermarking scheme. Multimedia Tools Appl 2011;52:621–40.
[11] Niu PP, Wang XY, Yang YP, Lu MY. A novel color image watermarking scheme in nonsampled contourlet-domain. Expert Syst Appl 2011;38:2081–98.
[12] Kerckhoffs A. La cryptographie militaire. J Sci Militaries 1883;4:161–91.



Lin Teng was born in Liaoning, China, in 1986. She received her B.S. degree in software engineering and her M.S. degree in computer software and theory from Dalian University of Technology, Liaoning, China, in 2009 and 2011, respectively. Currently, she is working towards her Ph.D. degree at the Faculty of Electronic Information & Electrical Engineering, Dalian University of Technology, Dalian, China. Her research interests include chaos, image encryption, secure communication and memristor.

Xingyuan Wang was born in Liaoning, China, in 1964. He received his B.S. degree in application physics and his M.S. degree in optics from Tianjin University, Tianjin, China, in 1987 and 1992, respectively, and his Ph.D. degree in computer software and theory from Northeastern University, Shenyang, China, in 1999. From 1999 to 2001, he was a Postdoctoral Fellow with the Department of Automation, Northeastern University. He is currently a Professor with the Faculty of Electronic Information & Electrical Engineering, Dalian University of Technology, Dalian, China. His research interests include biomedical information, computer graphics, image processing, complex network and chaos control and synchronization.

Xiukun Wang was born in 1945. She received her B.S. degree in applied mechanics and her M.S. degree in computer science and engineering from Dalian University of Technology, Liaoning, China, in 1970 and 1979, respectively. She is currently a Professor with the Faculty of Electronic Information & Electrical Engineering, Dalian University of Technology, Dalian, China. Her research interests include database systems and decision support systems.