A chaotic system based fragile watermarking scheme for image tamper detection

Sanjay Rawat*, Balasubramanian Raman1

Department of Mathematics, Indian Institute of Technology Roorkee, Roorkee 247 667, India

A R T I C L E   I N F O

Article history:
Received 23 September 2010
Accepted 24 January 2011

Keywords:
Tamper detection
Chaotic map
Authentication
Digital watermarking

A B S T R A C T

In the past few years, various fragile watermarking techniques have been proposed for image authentication and tamper detection. In this paper, a novel chaos based watermarking scheme for image authentication and tamper detection is proposed. Tamper localization and detection accuracy are two important aspects of the authentication watermarking schemes. Our scheme can detect any modification made to the image and can also indicate the specific locations that have been modified. To improve the security of the proposed scheme two chaotic maps are employed. Since chaotic maps are sensitive to initial values, the corresponding position relation between pixels in the watermarked image and the watermark get disturbed, which helps the watermarking scheme to withstand counterfeiting attacks. Initial values of the chaotic maps are used as secret keys in our scheme. The effectiveness of the proposed scheme is checked through a series of attacks. Experimental results demonstrate that the proposed scheme is not only secure but also achieves superior tamper detection and localization accuracy under different attacks. For instance in copy-and-paste attack and collage attack.

© 2011 Elsevier GmbH. All rights reserved.

1. Introduction

With the rapid growth of internet technologies, a large amount of digital data is easily accessible to everyone these days. This digital data can be easily manipulated, tampered and distributed with the help of powerful image processing tools. The ease and extent of such manipulations emphasize the need for image authentication techniques in applications where verification of integrity and authenticity of the image content is essential. Therefore, various authentication schemes have recently been proposed for verifying the integrity and authenticity of the image content. The authentication schemes can be divided into two categories: digital signature based schemes and digital watermark based schemes. A digital signature can be either an encrypted or a signed hash value of image contents and/or image characteristics. The major drawback of signature based schemes is that they can detect if an image has been modified, but they cannot locate the regions where the image has been modified [1–3]. To solve this problem, many researchers have proposed watermarking based schemes for image authentication [4–7]. One of the first watermarking-based authentication schemes was proposed by Walton [9]. He divided the image into $8 \times 8$ blocks and embedded the checksum in the LSB of each block. The main drawback of Walton’s scheme is that there is a possibility of exchanging the blocks with the same position in two different authenticated images without affecting the checksum of the image. Yeung and Mintzer [10] proposed a watermarking scheme for image authentication that uses a pseudo random sequence and a modified error diffusion method to embed a binary watermark into an image, so that any change in pixel values of the image can be detected. Fridrich et al. [11] analyzed the security issue in the scheme proposed by Yeung and Mintzer [10] and proposed an improved scheme with localization capability, where a block cipher defined on a local neighborhood rather than on a single pixel is used to replace the binary look-up tables. Thus, attacker could not deduce the binary look-up table. At the same time, authors embedded an image index into all non-overlapping sub-blocks of each image to prevent the collage attack proposed in [12,13]. Wong [14] proposed a public key fragile watermarking scheme for image authentication. He divided the image into non-overlapping blocks and inserted a digital signature for authentication. In his scheme, a key is used to generate a signature using the seven most significant bits of the pixels in each image block together with a logo to form a watermark, and embed the watermark into the least significant bits of the corresponding blocks. The blockwise independence of the authentication schemes, proposed in the literature was exploited by Holliman and Memon [15]. They proved that these scheme are vulnerable to vector quantization attack. According to them, a counterfeit image can be constructed using a vector quantization codebook generated from a set of watermarked images. Since each block is authenticated by itself, the counterfeit image appears authentic to the watermarking scheme. To withstand the vector quantization attack, a number of schemes has been proposed. Wong and Memon [16] proposed an improved blockwise authentication scheme by

* Corresponding author. Tel.: +91 01332 285852; fax: +91 01332 285852.
E-mail addresses: sanjuda@gmail.com (S. Rawat), balairitr@ieee.org (B. Raman).

1 Tel.: +91 1332 285852.

1434-8411/S – see front matter © 2011 Elsevier GmbH. All rights reserved.
doi:10.1016/j.aeue.2011.01.016
adding an image index and a block index to the inputs of the hash function. However, this idea works at the expense of requiring the verifier to have a priori knowledge about the image index, which limits its applicability to some extent. Celik et al. [17] proposed a hierarchically structured watermarking scheme based on Wong’s scheme [14] which provides a blockwise authentication with highly overlapping blocks. In Celik et al.’s scheme, the original image is partitioned into blocks in a multi-level hierarchy and then block signatures in this hierarchy are calculated. Based on this hierarchy structure, the scheme can effectively thwart vector quantization attack. Suthaharan [18] proposed a fragile watermarking scheme in which the security against vector quantization attack is achieved using a gradient image and its bits distribution properties to generate a large key space. Chang et al. [19] proposed a block-based image authentication scheme which can withstand counterfeiting attacks by combining the local and global features to obtain the authentication data. Chen et al. [20] proposed a fuzzy c-means clustering based watermarking scheme to resist counterfeiting attacks. To break the block wise independency, they applied the fuzzy c-means clustering technique to cluster all the image blocks, so that the relationship between blocks can be created. The authentication data is embedded into two least significant bits of each image block.

In this paper, a novel watermarking scheme based on chaotic maps is proposed. The pixels of the cover image are disturbed with the help of Arnold’s cat map. The image is further divided into 8-bit planes and the least significant bit (LSB) plane is used for watermark embedding. A binary logo is used as watermark in our scheme. A chaotic image pattern is generated by using logistic map. A scrambled watermark is obtained by using exclusive-or (XOR) operation between chaotic image pattern obtained by using logistic map and the binary watermark. The scrambled watermark is then embedded in the least significant bit (LSB) plane of the image. Watermarked image is obtained by performing an inverse cat map.

The rest of the paper is organized as follows. In Section 2, Arnold’s cat map and logistic map are briefly described. In Section 3, the proposed watermarking scheme is explained. Experimental results are given in Section 4. Conclusions are drawn in Section 5.
2. Chaotic maps

In recent years, chaotic maps have been used for digital watermarking to increase the security. The most attractive features of chaos in information hiding are its extreme sensitivity to initial conditions and the outspreading of orbits over the entire space. These special characteristics make chaotic maps excellent candidates for watermarking and encryption [8].

2.1. Arnold’s cat map

In order to shuffle the pixel positions of the host image, two dimensional Arnold cat map is employed in our scheme. The classical Arnold cat map is a two-dimensional invertible chaotic map described by

\[
\begin{bmatrix}
  x_{n+1} \\
  y_{n+1}
\end{bmatrix} = \begin{bmatrix}
  1 & 1 \\
  1 & 2
\end{bmatrix} \begin{bmatrix}
  x_n \\
  y_n
\end{bmatrix} \mod 1
\]

(1)

The map is area-preserving since the determinant of its linear transformation matrix is equal to 1. The map is chaotic in real space and in addition it is one to one map. The above 2D cat map can be extended as follows. First, the phase space is generalized to \([0, 1, 2, \ldots, N−1] \times [0, 1, 2, \ldots, N−1]\), i.e. only positive integers from 0 to \(N−1\) are taken; then above map is generalized as:

\[
\begin{bmatrix}
  x_{n+1} \\
  y_{n+1}
\end{bmatrix} = \begin{bmatrix}
  1 & a \\
  b & ab + 1
\end{bmatrix} \begin{bmatrix}
  x_n \\
  y_n
\end{bmatrix} \mod N = A \begin{bmatrix}
  x_n \\
  y_n
\end{bmatrix} \mod N
\]

(2)

where \(a\) and \(b\) are positive integers and \(\det(A) = 1\).

The generalized cat map given by Eq. (2) is also chaotic and area preserving. Since we have restricted the phase space to positive integers, the generalized cat map becomes periodic in nature. We say that the equation has period \(T\) if the pixel at location \((x, y)\) returns to its original position after being transformed \(T\) times. The period \(T\) depends on the parameters \(a, b\) and size of \(N\) of the original image. Thus parameters \(a, b\) and the number of iterations \(k\), all can be used as secret keys. Fig. 1 shows the periodicity property of the cat map where the parameters of Eq. (2) are \(a = 1, b = 1\) and size of the image is \(N = 128\). It shows that for these parameters the period is equal to 96.

2.2. Logistic map

Logistic map is one of the simplest chaotic maps, described by

\[x_{k+1} = \mu x_k (1 - x_k)\]

(3)

where \(0 < \mu \leq 4\). When \(3.5699456 < \mu \leq 4\), the map is in chaotic state. All the sequences generated by the logistic map are very sensitive to initial conditions, in the sense that two logistic sequences generated from different initial conditions are statistically uncor-

Fig. 3. Block diagram of extraction process.

related. Moreover, all the orbits of the logistic map are dense in the range of the map \([0, 1]\).

3. The proposed scheme

In this section, we explain the proposed watermarking scheme. Let us consider, \(I\) is the host image of size \(M \times N\) and \(W\) is the binary watermark of size \(m \times n\).

3.1. Watermark embedding

The watermark is embedded as follows:

1. Scramble the original image \(I\), using Arnold cat map \(k\) times, where \(k\) is the encryption key of the chaotic mixing process for \(I\). Let us denote the result 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 (XOR) operation between the watermark \(W\) and \(X_{cp}\) as follows:

\[W_p = S_{cp} \oplus W\]

5. Replace the least significant bit plane of \(I_{scr}\) by \(W_p\).
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 block diagram of the embedding process is shown in Fig. 2.

Fig. 4. (a) Original sailboat image, (b) binary watermark and (c) watermarked image.
3.2. 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 $I_{W}^{\text{sc}}$.
2. Divide the scrambled watermarked image $I_{W}^{\text{sc}}$ into 8-bit planes.
3. Obtain the same chaotic image pattern $S_{cp}$ as in step 3 of embedding algorithm.
4. Apply exclusive-or (XOR) operation between the least significant bit plane of $I_{W}^{\text{sc}}$ and chaotic image pattern $S_{cp}$ to get the extracted watermark $W^{\text{ext}}$.
5. Take the absolute difference of extracted watermark $W^{\text{ext}}$ and original watermark $W$. Apply Arnold cat map $(T - k)$ times to locate the tampered areas of the watermarked image.
4. Experimental results

Various experiments are carried out in this section, to assess the performance of the proposed algorithm. A binary logo of size $256 \times 256$ is used as watermark in all the experiments. The parameters of Arnold cat map used in our scheme are, $a = 1$, $b = 1$, and $k = 75$. The parameters of logistic map are chosen as $\mu = 3.854$ and $x(0) = 0.654$. PSNR (peak signal-to-noise ratio), is used in this paper to analyze the visual quality of the watermarked image $\hat{I}$ in comparison with the original image $I$. PSNR is defined as:

$$PSNR = 10\log_{10} \frac{255^2}{MSE} \text{ dB}$$

(4)
where MSE is the mean squared error between the original image \( I \) and the attacked image \( \hat{I} \), given by

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

\[(5)\]

4.1. Performance evaluation

4.1.1. Performance under copy and paste attack

In this experiment, ‘Sailboat’ image of size 256 \( \times \) 256 is used. A binary logo image of size 256 \( \times \) 256 is used as watermark. Fig. 4 shows the host image, binary watermark and the corresponding watermarked image. The PSNR value of watermarked image (Fig. 4(c)) is 50.7261 dB. Two kinds of copy and paste attacks are performed in our scheme. In first kind of copy and paste attack the watermarked sailboat image is modified by inserting two more boats in the image, where the boat is copied from the same watermarked image. The tampered image is shown in Fig. 5(c). Fig. 5(d) shows the extracted watermark from Fig. 5(c). The tamper detection result is shown in Fig. 5(e).

In second kind of copy and paste attack the watermarked sailboat image is modified by inserting an aeroplane in the image, where the aeroplane is copied from some other watermarked image. The tampered image is shown in Fig. 6(c). Fig. 6(d) shows the extracted watermark from Fig. 6(c). Detected tampered region is shown in Fig. 6(e).

4.1.2. Performance under text addition

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

4.1.3. 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(c). Fig. 8(d) shows the extracted watermark from Fig. 8(c). The tamper detection result is shown in Fig. 8(e).

4.1.4. Performance under collage attack

To evaluate the performance under collage attack, a counterfeit image is formed by combining the portions of multiple water-
marked images while preserving their relative spatial locations within the target image. We have performed this attack for two sets of images. The simulation results are as follows:

- Sky and bird: The original sky and bird images are shown in Fig. 9(a) and (b). Both the images are of size 256 × 256. The corresponding watermarked images are shown in Fig. 9(c) and (d), where the PSNR values are 51.1552 dB and 51.0992 dB, respectively. The counterfeit image, as shown in Fig. 9(e) was constructed by copying the bird from Fig. 9(d) and pasting it in Fig. 9(c). Fig. 9(f) shows the extracted watermark from Fig. 9(e) and Fig. 9(g) shows the detected tampered region.

- Sofa and doll: The original sofa and doll images are shown in Fig. 10(a) and (b). Both the images are of size 256 × 256. The corresponding watermarked images are shown in Fig. 10(c) and (d), where the PSNR values are 50.6722 dB and 51.0752 dB, respectively. The counterfeit image, as shown in Fig. 10(e) was constructed by copying the teddy bear from Fig. 10(d) and pasting it in Fig. 10(c). Fig. 10(f) shows the extracted watermark from Fig. 10(e) and Fig. 10(g) shows the detected tampered region.

5. Conclusion

A novel fragile watermarking scheme for image authentication and locating tampered regions, is presented in this paper. Chaotic maps are used in our scheme to make the scheme highly secure. Since chaotic maps are sensitive to initial values, they are used as keys in our scheme. Extracting the right watermark is only possible if someone has correct keys. A person with wrong keys will not be able to forge the watermark. As, in order to thwart counterfeiting attacks it is essential to break pixel wise independency, the proposed scheme employs chaotic maps to break the corresponding position relation between pixels in the watermarked image and the watermark. Experimental results show that our scheme has high fidelity and is capable of localizing modified regions in watermarked image.

References


