A Video Watermarking Algorithm Based on Chaotic and Wavelet Neural Network

Jiadong Liang

Abstract**—**This paper presented a video watermarking algorithm based on wavelet chaotic neural network. First, to enhance binary image's security, the algorithm encrypted it with double chaotic based on Arnold and Logistic map, Then, the host video was divided into some equal frames and distilled the key frame through chaotic sequence which generated by Logistic. Meanwhile, we distilled the low frequency coefficients of luminance component and self-adaptively embedded the processed image watermark into the low frequency coefficients of the wavelet transformed luminance component with the wavelet neural network. The experimental result suggested that the presented algorithm has better invisibility and robustness against noise, Gaussian filter, rotation, frame loss and other attacks.

Keywords**—**Video watermark, double chaotic encryption, wavelet neural network.

I. Introduction

In recent years, with the development of network technology, the propagation of multimedia information has become more convenient but at the same time it has brought many copyright disputes. By embedding watermark information into the video, video watermarking technique could protect the copyright. In terms of video watermarking research, due to the more accurate human visual model including spatial masking effects has not been fully established, video watermarking technology is underdeveloped compared with image watermarking techniques and the existing standard video coding format has also limited the development of watermarking technology. Initially, video watermarking is used to protect the copyright of digital video products, but recently, its application has been extended due to its imperceptibility, robustness and security features [1].

Wavelet neural network is a multilayer feed-forward network which was proposed on the basis of wavelet analysis. It radically helps network to avoid local optimum and accelerates the convergent rate. Wavelet neural network has high learning and generalization ability. The difference between wavelet neural network and multilayer feed-forward network is that the former activation function is not Sigmoid nonlinear function but wavelet-based function [2].

This paper presented a video watermarking algorithm. The algorithm combines the chaotic map and wavelet chaotic neural network to embed the watermark into the key frames. Related simulation experiments show that the algorithm is simple, robust, and easy to operate.

Jiadong Liang is with the Department of mathematics and Physics, Qingdao University of Science & Technology, Qingdao City, Shandong Province, PR China (e-mail: dongdongdiandi@163.com).

II. Image Watermarking Pretreatment

This paper selected binary image as a watermark and scrambled it before embedding video. Specific procedures are described as [3]:

Step 1: We transformed image's data matrix into double-precision type. In order to apply two-dimensional cat map, we have to cut the double image into $N \times N$. For the sake of convenience, this paper chose 128×128 logo. bmp image.

Step 2: Cat map shuffled pixels: First we have to discretize cat map and generalize its phase space to $\{0,1,2,3,4,......,N-1\}$ to get the broad cat map as:

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & a \\ b & ab+1 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} (\text{mod N}) = C \begin{bmatrix} x_n \\ y_n \end{bmatrix} (\text{mod N}) \tag{1}$$

Among them, *a* and *b* are positive integers and obviously the determinant of matrix |C| is one. Geometrically, it could be stretched and folded so that it is able to part the adjacent pixels through transformation, that is to say the broad cat map still maintains its sensitivity of initial values.

Encryption's times (L), cat-map parameters (a and b) can be the keys. Here we set L=15, a=N-15, b=N-19 and keys are associated with image sizes;

Step 3: Logistic map diffuses gray-scale value [4].

The definition formula of logistic map is as:

$$x_{n+1} = \mu x_n (1 - x_n) \tag{2}$$

For this formula, $0 < \mu \le 4$ is branch parameter, $x_n \in (0,1)$ m =, n $- \{0,1,2,3,4......\}$. Studies [4] have shown that, if $3.5699456 < \mu \le 4$, logistic map will in a chaotic state. We set $\mu = 3.72$, $x_0 = 0.18$ and generated chaotic sequences through logistic map and gave XOR operation to the rearranged image gray value to get two-dimensional watermark sequences $w(x, y)$. Then, we reduced the two-dimensional watermark sequence to one-dimension sequence S_t.

Step 4: Repeat step 2 and step 3 for L times and complete the process of encryption, the result of the encryption is shown in Figs. 1 and 2;

Step 5: Decryption: First we xorred the encrypted image's gray-values with the Logistic-chaotic sequence, then inverse the location of the pixels by the cat-mapping with the same keys, the formula of the inverse matrix was shown as (3). Repeat L times we could decrypt the original watermark image.

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} ab+1 & -a \\ -b & 1 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} (\text{mod N}) = \mathrm{C}^{-1} \begin{bmatrix} x_n \\ y_n \end{bmatrix} (\text{mod N}) \tag{3}$$

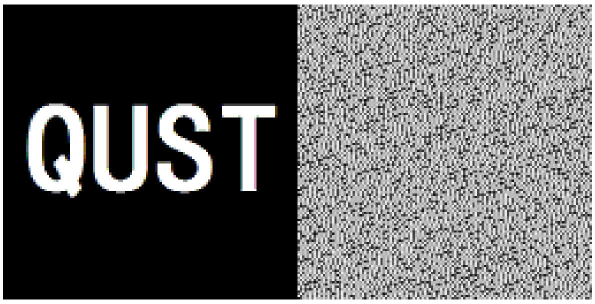

Fig. 1 Original watermark image and encrypted watermark image

III. Wavelet Neural Network [5]

Wavelet neural network is a multilayer feed-forward network which is proposed on the basis of wavelet analysis. It radically helps network to avoid local optimum and accelerates the convergent rate. Wavelet neural network has high learning and generalization ability. The difference between wavelet neural network and multilayer feed-forward network is that the former's activation function is not sigmoid nonlinear function but wavelet-based function.

We supposed wavelet neural network has M input nodes, N output nodes and n hidden layer nodes as well as used vector X and Y to represent network's input and output data respectively, that is:

$$X = (x_1, x_2, ..., x_m)^T \tag{4}$$

$$Y = (y_1, y_2, ..., y_N)^T \tag{5}$$

We set x_k as the Kth input sample of input layer, y_i as the i th input value of input layer, w_{ij} as the weighting to connect out layer i and hidden layer node j, w_{jk} as the weighting to connect hidden layer j and the output layer node k. If we set w_{i0} as threshold of the j th output layer node and w_{j0} as threshold of the j th hidden layer node (input $x_0 = -1$), a_j as scaling factor of the j th hidden layer node, b_j as shift factor of the j th hidden layer node, then wavelet neural network model can be expressed as:

$$y_i(\mathrm{t}) = \sigma\left[\sum_{i=0}^{n} w_{ij}\psi_{a,b}\left(\sum_{k=0}^{m} w_{jk}x_k(\mathrm{t})\right)\right] \quad , \quad \mathrm{i} = 1,2,...,\mathrm{N} \tag{6}$$

For this model, $\sigma(\mathrm{t}) = \dfrac{1}{1+e^{-t}}$. If we set $net_j = \sum_{k=0}^{m} w_{jk}x_k$, then:

$$\psi_{a,b}(net_j) = \psi\left(\frac{net_j - b_j}{a_j}\right) \tag{7}$$

$$y_i(\mathrm{t}) = \sigma\left[\sum_{i=0}^{n} w_{ij}\psi_{a,b}\left(net_j\right)\right] \tag{8}$$

IV. Design of Watermark Embedding

When processing the host video, first we divided the video into some equal frames to get frame numbers and numbered them from 1 to n, then used Logistic map to produce a *n* length chaotic sequence and corresponded it with the number of video frames. We valued the sequence which is greater than or equal to 0.9 as value 9 and valued the sequence which greater than 0.8 and less than 0.9 as value 8. By this analogy, we divided the sequence into 10 parts and got the 0-9 sequences with *n* length of each frame. Then exacted all the 0-9 frames and embedded watermarks into them. Furthermore, we set the logistic mapped coefficient μ, x_0 and video coefficient *r* as keys to exact video frames. Meanwhile, we could get better robustness by watermarking the most important part of vision and after the wavelet decomposition; low frequency coefficient which is the most important part of vision contained most of the energy. So, this algorithm selected the low frequency sub-band as embedded region and the embedding strength will be self-adaptive determined by low frequency coefficients which trained by wavelet neural network. In this algorithm, we used coefficient template of brightness component to train neural network and got the relationship between template feature data and digital video data which were embedded watermark through machine learning. Thus, we used this value and the neural network outputs as embedding strength and embedded them into digital video. Detailed procedures are as [6]:

Step1. Chose Haar as wavelet basis to decompose the embedded region into three layers wavelet and got the kth frame video image illumination's decomposition coefficient Y_k. Exacted neighbouring low frequency coefficient Y_k and $Y_k(\mathrm{i,j+1})$ from the approximation component $Y_k(\mathrm{i,j})$ of the third luminescent component layer, with $j = 4$, $i \geq 1$.

Step2. Here we introduced embedding strength T. Then got $T_{i,j}$ and $T_{i,j+1}$ through iterating low frequency coefficient $Y_k(\mathrm{i,j})$ and $Y_k(\mathrm{i,j+1})$ by (8). We set n as input key and got embedded intensity T's range of value is $0 < T < 1$ after several experiments. So, the embedded strength was automatically adjusted by the contents of video frame to realize adaptive embedding.

Step3. During the process, coefficient changed. When $S_t = 1$, $Y_k(i, j+1)' = Y_k(i, j+1) - T_{i,j+1}$; when $S_t = 0$, $Y_k(i, j)' = Y_k(i, j) - T_{i,j}$, $Y_k(i, j+1)' = Y_k(i, j+1) + T_{i,j+1}$.

Step4. Decompose watermarked wavelet coefficient Y_k' into three layers to get watermarked video frame's luminance component. Then reconstructed it with the unchanged luminance component to get watermarked video frames.

V. Design of Watermark Detection Scheme

Step1. Divide the watermarked video into equal frames. Use key μ and x_0 to generate chaos sequence and exact key frame according to video coefficient and then extract the luminance component.

Step2. Make triple wavelet transform to the luminance component of key frame.

Step3. Extract watermarked $Y_k(i, j)'$ and $Y_k(i, j+1)'$ from low frequency sub-band. If $Y_k(i, j)' > Y_k(i, j+1)'$, then set watermark information as 1, otherwise 0;

Step4. Repeat steps 2 and 3 to get all watermark information $w(x, y)$.

Step5. XOR the watermark information $w(x, y)$ and Logistic chaos sequence and then use inverse matrix which mapped by the same key two-dimensional cat map to inverse the pixel position. We could decode the image after repeating L times.

VI. Analysis of Simulation Results

During the simulation, we adopted foreman.yuv video sequences which are CIF format as host video and divided the video into equal frames to get 151 frames' video data. The 0-9 chaotic sequence is 7676……8592 which were produced by Logistic and key $\mu = 3.81$, $x_0 = 0.19$. When r=4, the key frame number is 6, 18, 22, 32, 36, 44, 92, 102, 120, 139, 143, 152, 173. Then extracted the corresponding key frames, each video image's size is 256×256 and the image watermark is a 128×128 binary image [7].

A. Experimental Test of Imperceptibility

The imperceptibility requires visual imperceptible of watermark. This paper embedded image watermark into video. Judged subjectively, the quality of carrier image and watermark almost didn't deteriorate, as shown in Figs. 2-5.

Fig. 2 Original video frame

Fig. 3 Watermarked video frame

Fig. 4 Original watermarking

Fig. 5 Extractive watermarking

B. Experimental Test of Robustness

We commonly attacked all frames of the watermarked video such as pepper and salt noise, Gaussian filter, rotation, frame deletion and so on. We adopted different intensity and coefficient during the attack to verify the robustness of this algorithm. Pepper and salt noise is also known as impulse noise and could randomly change some pixel values. It is a kind of black and white noise generated by image sensor, transmission channel and decoding process and so on. The noise attack coefficient is 0.001, 0.005 and 0.01, shown in

International Science Index, Computer and Information Engineering Vol:10, No:7, 2016 waset.org/Publication/10006055

Figs. 7-9. Gaussian filter is a process of taking weighted average of the whole image; every pixel value is determined by taking weighted average of its own and other neighboring pixel values. Gaussian coefficient is 5, 0.2; 5, 0.5; 5, 0.8, as shown in Figs. 10-12. Frame deletion attack is a situation of losing one or more frames during the process of compression, shear or transmission. It will cause great damage to original video sequences and loss data. In this paper, we randomly deleted the 5, 10, 15 frame of video sequence to finish the experiment, as shown in Figs. 12-14 [8].

Fig. 6 Noise 0.001

Fig. 7 Noise 0.00

Fig. 8 Noise 0.01

Fig. 9 5,0.2

Fig. 10 5,0.5

Fig. 11 5,0.8

Fig. 12 5 Frames

Fig. 13 Frame

Fig. 14 Frames

TABLE I
COMPARISON OF PSNR AND NC VALUE UNDER THE ATTACK OF NOISE

Noise Intensity	0.001	0.005	0.01
PSNR	26.2431	21.3141	16.9523
NC	0.9643	0.9343	0.8595

TABLE II
COMPARISON OF PSNR AND NC VALUE UNDER THE ATTACK OF GAUSSIAN

Gaussian Filter Coefficient	5,0.2	5,0.5	5,0.8
PSNR	28.343	15.5493	8.7844
NC	0.9932	0.7623	0.4572

TABLE III
COMPARISON OF PSNR AND NC VALUE UNDER THE ATTACK OF FRAME LOSS

Frame loss rate	3.3%	6.6%	9.9%
PSNR	28.8203	26.6546	21.2365
NC	0.9951	0.9876	0.9325

The above experimental results and data have shown that although the quality of watermarked video declined during the attack, the extractive watermark image still has a high similarity to original watermark. So, we can conclude that watermarking algorithm has a strong robustness to the above attacks.

VII. CONCLUSION

This paper presented a video watermarking algorithm based on wavelet chaotic neural network and implemented relevant attack experiments. The experiments concluded PSNR and NC value based on various attacks of the presented algorithm. Through the effect diagram we can see that watermark almost didn't influence the video frame quality and has a good invisibility; The PSNR and NC value shown that even with low PSNR value, the attacked watermarked image could still be tested and it has a better robustness.

REFERENCES

[1] J. Zhang, L. Zhao, S.Q. Yang. *Summary of Video Watermarking Technology*. Computer application, vol.25, 2015, pp.850-852.
[2] Zuo D G,Zhou S,Zhang X Y. *Wavelet Neural Network*.Journal of Sichuan Ordnance, vol.5,2012,pp.84-86.
[3] X. Li. *The Research of Chaos Image Encryption Scheme Based on Cat Map and Logistic Map*. Guangdong University of technology, vol.6, 2013. pp.35-38.
[4] X.F. Liao, D. Xiao, Y. Chen. *Principle and Application of Chaotic Cryptology (m)*. Beijing: Science Press, vol.4, 2009, pp.44-50.
[5] H.X. Fu, H. Zhao et al. *Practical Design of Application of MATLAB Neural Network*. Beijing: China machine press vol.7, 2010, pp.99-106.
[6] K.Q. Ren, K. Zhang, B. Xie. *Blind Adaptive Video Watermarking Algorithm Based on Low Frequency Coefficients of Wavelet Domain*. TV technology, vol.36, 2012, pp.47-51.
[7] Li Y, Gao X B, Ji H B. *A 3D Wavelet Based Spatial Temporal Approach for Video Watermarking(C)//Proc. of the 5th International Conf. on Computational Interlligence and Multimedia Applications*. China: Computer Society Press, Xi'an, vol.4, 2003, pp.260-265.
[8] K.H. Sun, W. Cheng, Y.S. Chen. *Video Watermarking Algorithm Based on Chaos and Motion Vector*. Application Research of Computers, vol.28, 2011, pp.3046-3048.