AN ARTIFICIAL NEURAL NETWORK-BASED SCHEME FOR FRAGILE WATERMARKING

Yu-Cheng Fan, Wei-Lung Mao and Hen-Wai Tsao
Integrated System Lab
Department of Electrical Engineering and Graduate Institute of Electronics Engineering
National Taiwan University, Taipei, Taiwan, 10617, R.O.C.

ABSTRACT

This paper proposes an artificial neural network based fragile watermarking scheme. Our method can detect tampering, locate where the tampering has occurred and recognize what kind of alteration has occurred. The experimental results have proven that our method is indeed effective.

INTRODUCTION

In this paper, we propose an artificial neural network-based scheme for fragile watermarking. A fragile watermark is useful in image authentication applications. It can detect slight changes in the image and prevent mark-transfer attacks [1]. In the past, most fragile watermark systems work by inserting watermark data or modifying some coefficients in the host image [2][3][4]. It is suitable to embed some data into the image as a robust watermark that carries proof of authorship. However, these fragile watermark systems cannot stand for the features of the image and recognize what kind of modification has occurred. These methods cannot detect all kinds of distortion. Sometimes, these fragile watermarking methods destroy the host images. In order to overcome these problems, we here propose an "artificial neural network-based fragile watermarking scheme" as shown in Fig. 1. Our method is designed to detect tampering of the host image, locate where the tampering has occurred and recognize what kind of alteration has occurred.

FRAGILE WATERMARKING PROCEDURES

At first, the host image is transformed by the discrete wavelet transform (DWT) to transfer image information from the spatial domain to the wavelet domain. We analyze the coefficients in the high-high band after DWT and embed the fragile watermark according to lookup table. The lookup table is base on the image's characteristics. The coefficients in the high frequency range stand for high resolution and represent a subtle difference in the image.

When the host image is modified, the slight changes can be detected easily according to analyze of the variations of the fragile watermark. We transform the modified image using DWT and extract the fragile watermark in high-high band. Comparing the extracted fragile watermark with the original fragile watermark, we can find the difference between the host image and modified image. Then we use artificial neural network to recognize what kind of modification has occurred [5]. We adopt back propagation model as shown in Fig. 2. We get the difference coefficients between the host image and modified image. (Fig. 3) Next, we analyze the horizontal energy distribution, vertical energy distribution, size, histogram variation and similarity. We use these characteristics as the neural network input signals. We use several kinds of modifying forms as output signals. After the initial values are set, the neural network is begun training until find the optimize weights. After the neural network is established, we can use this model to analyze the degree of changes and any modified image what kind of modification has occurred.

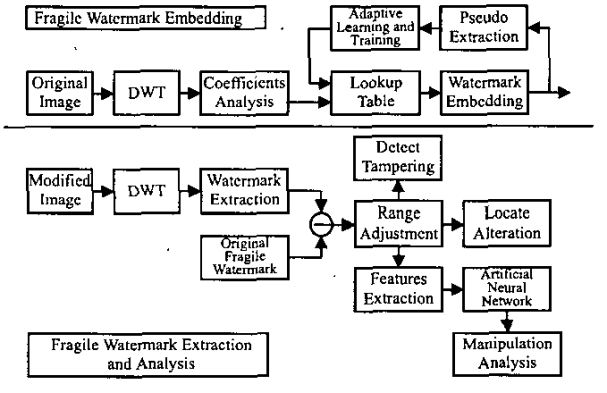

Fig.1 Artificial neural network-based fragile watermarking scheme

0-7803-7721-4/03 $17.00 © 2003 IEEE

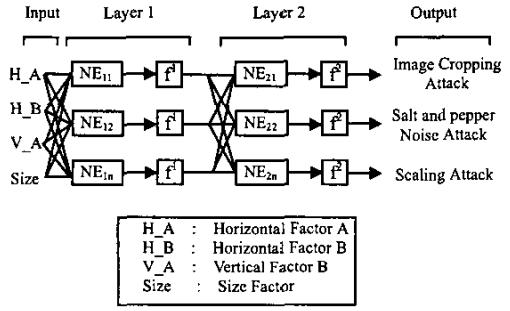

Fig.2 Back propagation model of artificial neural network

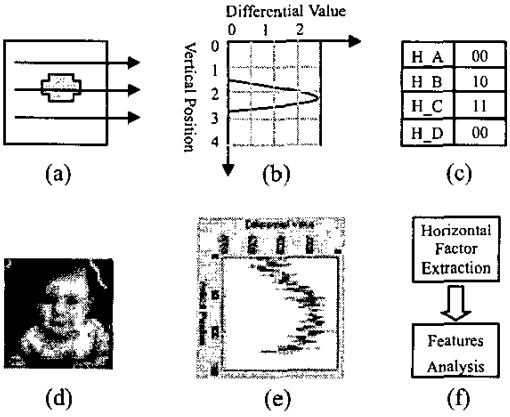

Fig.3 Features extraction and analysis (Horizontal Factor) (a) .The difference between the host image and modified image (b) Calculate the horizontal difference value (c) Calculate the horizontal factor (d) Host image after Salt and pepper noise attack (e) Calculate the horizontal difference value (f) Horizontal factor extraction and features analysis.

Table 1 Attack recognition summary

Picture Number	Attack Type				
	Blurred	Scaling	Median Filtered	Gaussian Noise	JPEG Compressed
50	94 %	100 %	93 %	95 %	89 %
100	96 %	100 %	95 %	97 %	92 %
150	99 %	100 %	99 %	99 %	94 %
200	100 %	100 %	99 %	100 %	95 %
250	100 %	100 %	100 %	100 %	95 %

(Unit: Recognition Rate %)

EXPERIMENTAL RESULTS

In order to prove the ability of the artificial neural network-based fragile watermark, a series of experiments were conducted. (Table 1) This scheme can easy recognize blurred attack, scaling attack, median filtered attack, Gaussian Noise attack and JPEG Compressed. This scheme also can recognize cropping attack, salt and pepper noise attack, slight changes, and so on. The experimental results have proven that our method is indeed effective.

CONCLUSIONS

An artificial neural network-based scheme for fragile watermarking has been developed in this work. This scheme analyzes the coefficients in the high-high band. This fragile watermarking represents the characteristic of the host image. It is easy to detect the slight changes and include the ability to locate and characterize alterations. We use this artificial neural network model to analyze the degree of changes and any tampered image what kind of alteration has occurred. This is a very convenient and feasible scheme.

ACKNOWLEDGMENT

The authors gratefully acknowledge NSC Chip Implementation Center (CIC), for supplying the SPW/HDS software used in the functional simulations.

REFERENCES

[1] Ingemar, J. Cox, Matthew L. Miller, and Jeffrey A. Bloom, "*Digital Watermarking,*" San Diego, CA: Academic Press, 2002

[2] Min Wu, Bede Liu, "Watermarking for image authentication," *Image Processing, 1998. Proceedings. 1998 International Conference on, Vol. 2*, Oct 1998, pp: 437 -441

[3] Alturki, F.; Mersereau, R., "Secure fragile digital watermarking technique for image authentication," *Image Processing, 2001. Proceedings. 2001 International Conference on, Vol.3*, 2001, pp: 1031 –1034

[4] Fridrich, J.; Goljan, M.; Baldoza, A.C., "New fragile authentication watermark for images," *Image Processing, 2000. Proceedings. 2000 International Conference on, Vol.1*, 2000, pp: 446 -449

[5] Simon Haykin, "*Neural Network: A Comprehensive Foundation,*" Prentice Hall Press, 2nd, 1999.