STEGANOGRAPHY USING SENSOR NOISE AND LINEAR PREDICTION SYNTHESIS FILTER

Xiaoyi Yu[1]*, Xinshan Zhu*[2] *and Noboru Babaguchi*[1]

[1]Graduate School of Engineering, Osaka University, Japan
[2]Peking University, Beijing, China

ABSTRACT

This paper presents a new approach utilizing the sensor's pattern noise and linear prediction synthesis filter for steganography. The pattern noise is extracted from the images using denoising filter (for example wavelet-based filter). Then the approach introduces the linear prediction synthesis filter, whose parameters are derived from the extracted noise. After being filtered by such a filter, the secret message can be embedded by adapting the characteristics of the sensor's pattern noise. As a result, the embedding process violate little of the natural image statistics, and hence the detectability of steganalytic method is noticeably decreased. The experimental results prove the effectiveness of the new approach.

Index Terms— Image processing, Image analysis

1. INTRODUCTION

Steganography aims to hide the very presence of communication. That is to say, the essential goal of steganography is to conceal the facts of a hidden message. Similar to cryptanalysis, steganalysis attempts to defeat the goal of steganography. The most popular and frequently used steganographic method is the Least Significant Bit embedding (LSB). LSB steganography is based on manipulating the LSB planes by directly replacing the LSBs of the cover-image with the message bits. However this technique is fragile to steganalysis attack. A series of steganalysis techniques [1] have been developed to attack this steganography.

Many steganographic methods have been proposed for anitsteganalysis. For example, Provos' Outguess algorithm [2] attempted to use histogram compensation for LSB hiding, and Eggers et al [3] suggest the same idea which called histogram preserving data mapping (HPDM). However, many steganalysis tools [4] [5] have been proposed to counter such histogram preserving hiding methods. For example, Fridrich et al [4] proposed a re-quantization method to defeat F5 and Outguss. Sallee's model-based steganography [6]has been seen a landmark of steganography. It provides an interesting and different perspective in the design of steganographic systems, with the hider ensuring that the stego signal conforms to a given model. However, it is hard to find a good model for the cover media. This is indeed practically shown in [7], where Sallee's Cauchy-model based JPEG steganography is broken by using only the first order statistics.

In this paper, we propose a novel embedding schemes that can evade statistical steganalysis. We are motivated by the notion of pattern noise proposed by Lukas et. al. [8]. The pattern noise is an equivalent of gun identification from bullet scratches for identification of digital cameras from images. Using the denoising filter described in [8], authors can extract the high frequency part of the pattern noise. A detailed description will be given in section 2. Pattern noise can be utilized for steganography. The secret message is embedded into the image by trying to adapting the pattern noise. A linear prediction (LP) synthesis filter is introduced to adapt the pattern. Section 3 describes the construction and property of the filter. Section 4, we introduce the concept of our proposed steganogrphic method, and analyze some practical implementation issues and anti-steganalysis performance. A specific experiment is designed according to the proposed scheme and a serial of tests are done to evaluate it in Section 5. Finally, we conclude the paper.

2. SENSOR PATTERN NOISE

As explained in the introduction, the pattern noise is a kind of systematic noise that does not change from image to image and is relatively stable over the camera life span and a reasonable range of conditions (temperature)[8].There are many sources of noise in images obtained using charge-coupled device (CCD) arrays, such as dark current, shot noise, circuit noise, fixed pattern noise, etc. The only noise components that are not reduced by frame averaging [8]are fixed pattern noise and photoresponse non-uniformity noise.

As described in [8], the main problem is how to obtain the sensor's pattern noise. The authors in [8], utilize wavelet-based image denoising method to characterize the pattern noise. The problem investigated in [8] is to identify the sensor that was used to obtain a given digital image. The authors claim that the high-medium frequency component of the sensor pattern noise is an equivalent of "bullet scratches" for digital images and can be used for reliable forensic identification. In

this paper, we use another method to obtain the pattern noise which we show in the next section.

3. LINEAR PREDICTION SYNTHESIS FILTER

In this section, the 2-D LP synthesis filter [11]is described concisely. Suppose that $x(n_1, n_2)$ represents a 2-D discrete spatial signal. Applying optimum linear prediction theory (i.e. forming and solving 2-D Yule-Walker normal equations) to $x(n_1, n_2)$, the linear prediction coefficients (LPC), $A = \{a(m_1, m_2) \in R|0 \leq m_1 \leq m_2, 0 \leq m_2 \leq p_2\}$ can be obtained, where p_1 and p_2 represent the row order and column order respectively [11]. And the linear prediction error $e(n_1, n_2)$ with variance $G^2 = \varepsilon\{|e(n_1, n_2)|^2\}$ has the form

$$e(n_1, n_2) = x(n_1, n_2) - \sum_{\substack{m_1=0 \\ m_2 \neq 0}}^{p_1} \sum_{\substack{m_2=0 \\ m_1 \neq 0}}^{p_2} a(m_1, m_2) x(n_1 - m_1, n_2 - m_2) \quad (1)$$

The LP synthesis filter is defined as an IIR filter using the LPC array A as the filter taps. Therefore, it is expressed in the form of I/O difference equation as

$$y(n_1, n_2) = \sum_{\substack{m_1=0 \\ m_2 \neq 0}}^{p_1} \sum_{\substack{m_2=0 \\ m_1 \neq 0}}^{p_2} \phi$$
$$\phi = a(m_1, m_2) y(n_1 - m_1, n_2 - m_2) + G\nu(n_1, n_2) \quad (2)$$

It has been shown in [11] that a nice property of this filter is that it can adapt the spectral structure of the processing noise, $\nu(n_1, n_2)$ to that of the host signal $x(n_1, n_2)$, which is illustrated in 1-D in the bottom of Fig. 1. According to the conclusion, this filter is also able to adapt the spectral structure of the processing noise to the signal $x(n_1, n_2)$ itself, as long as A is derived by applying optimum linear prediction theory to the 2-D inverse Fourier transform of $x(n_1, n_2)$, which is denoted with $\widetilde{X}(f_1, f_2)$. In fact, $x(n_1, n_2)$ is viewed as the spectrum to be simulated. The design of LP synthesis filters for the latter purpose is depicted in Fig. 2.More detail can be found in [11].

4. STEGANOGRAPHY PROCEDURE

LP synthesis filter is applied broadly to speech coding, image coding and compression.We use LP synthesis filter to maintain the spectral structure of the extracted pattern noise for steganography. We first design a basic operation for secret message embedding. It is illustrated in Fig. 3. We consider the basic operation as a blak box which takes 2 inputs and 1 output. The random signal (not secret message), R_o is a set of random numbers of independently Normal distribution with zero mean. Therefore

$$R_o(i, j) \sim N(0, \sigma^2_{Ro}) \quad (3)$$

It is used to adjust the output I_s . The input I_o is an image block randomly choosen from the original image. Inside the

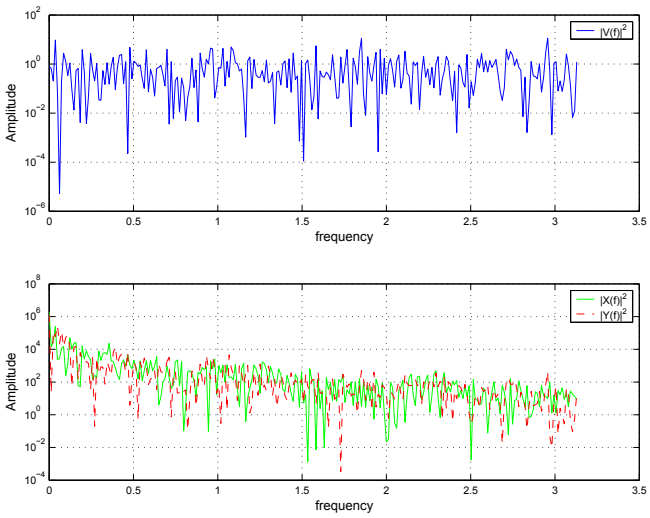

Fig. 1. The spatial domain LP synthesis filter adapts the structure of the noise spectrum to that of the signal spectrum.

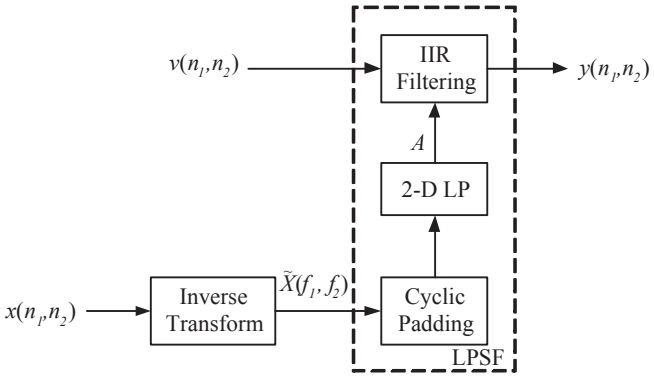

Fig. 2. Basic design of LP synthesis filter to adapt the spectral structure of the processing noise to the signal $x(n_1, n_2)$ itself

black box,$I_o - F_\sigma$ is the extracted pattern noise. According to the property of LP synthesis filter, which is described in Section 3, R_o is reshaped as R_s, the spectral coefficients of which, R_s is near to $I_o - F_\sigma$, the pattern noise. The LP synthesis filter is implemented as shown in Fig. 2. Therefore, the reshaped signal R_s has the form

$$R_s(n_1, n_2) = \sum_{\substack{m_1=0 \\ m_2 \neq 0}}^{p_1} \sum_{\substack{m_2=0 \\ m_1 \neq 0}}^{p_2} \phi$$
$$\phi = a(m_1, m_2) R_s(n_1 - m_1, n_2 - m_2) + GR_o(n_1, n_2) \quad (4)$$

where $a(m_1, m_2)$ is the LPC of $I_o - F_\sigma$.

Then, we will describe our new steganographic protocol which attempt to defeat these statistical steganalysis. The embedding and extraction procedure is described as follows:

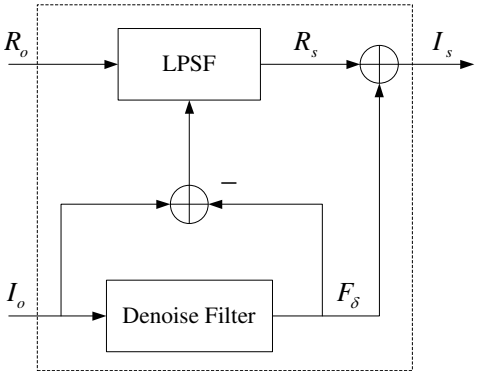

Fig. 3. Basic embedder with LP synthesis filter.

(a) (b)

Fig. 4. (a) Original image "Lena"; (b) The steg-image "Lena"

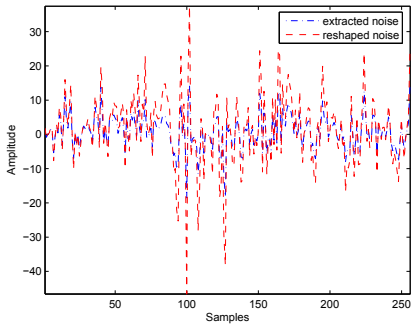

Fig. 5. Extracted pattern noise & reshaped noise

4.1. Embedding Procedure

First, we describe the embedding procedure.

1. The original image are broken into blocks of arbitrary size according to a predefined pattern;

2. Using a steganography key, pseudo-randomly map the secret message bits to image blocks in the host image.

3. For any image block I_o, which is mapped to a secret message bit, randomly choose a signal R_o, run the basic operation, which described in Section 4, with inputs R_o and I_o, obtain the output I_s; Calculate the cipher text or hash value of I_s, using encryption algorithm such as DES, or hash function such as MD5.

4. If the parity of cipher text or hash value is consistent with the corresponding secret message, the pixel value remains unchanged. Otherwise, choose another random signal R_o, go to step 2.

4.2. Extraction Procedure

The extraction procedure is very simple. First, with a steganography key, we find out the image blocks which carry secret message bits. Then for any image block which is mapped to a secret message bit, calculate the cipher text or hash value of that block, using encryption algorithm such as DES, or hash function such as MD5. Last, combine the parity of cipher texts or hash values, we obtain the secret message.

5. EXPERIMENTAL RESULTS

In order to verify the proposed steganographic method, two sets of experiments are performed here. One experiment is carried on to test the performance of data hiding in images using our proposed method, while the other is to test the anti-steganalysis performance of the proposed method. In our experiments, the image block size was taken as 8x8 JPEG blocks. The random signal R_o is pseudo-randomly generated and is reshaped as an 8x8 block. Fig. 4 shows the original image and stego-image with full capacity message embedded.

There is no visual artifact in the stego-image, and we can tell no difference of these 2 images from visual analysis. Since the embedding procedure only modify the extracted noise, we can compare the extracted noise and reshaped noise. Black line in Fig.5 shows the extracted pattern noise $I_o - F_\sigma$ and doted line shows the reshaped noise R_s. In Fig.5, we only shows a line of whole blocks. Fig. 6 gives the histogram of original image and stego-image obtained from embedding full capacity messages using the proposed method. The 2 cures are almost identical as the difference introduced by the embed-ding procedure is very small. Although we tested our method on grayscale images, nothing prevents its application to color images.

The second experiment is set up to test the anti-steganalysis performance. we focus our attention on so called specific techniques and blind techniques.

For specific techniques, since we only change the pixel value through filtering image block, this might change the blockiness of a image, we turn our attention to the blockiness due to block embedding[5]. We compare the detection results of steg-images, which have embedded with secret message using our proposed method, with the detection results of original cover images. The emebedding capacity is full embedding, i.e. every block carries one secret message bit. We have generated a stego image for each image in Corel

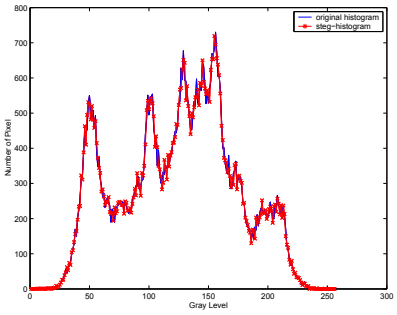

Fig. 6. Histogram of Original Image and Stego-Image

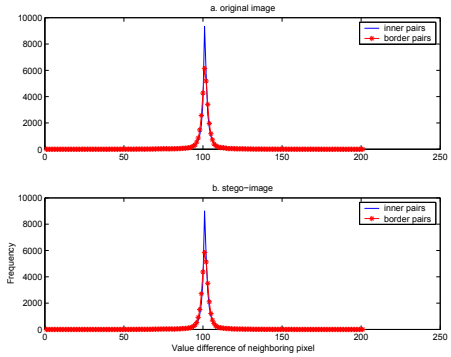

Fig. 7. Blockiness steganalysis test results

Image Database (10000 JPEG images totally) using our proposed steganographic method, then run the detection test described in [5] both on stego image and original image. Fig.7 shows the histograms for value difference from inner and border pairs based on original image and stego-image. We find that there is no much difference of the detection results.

For blind techniques, we used the universal steganalysis technique proposed by Farid et. [10] . Fisher classifier was trained with the statistics from the 300 image training subset . Then the trained classifier was tested against new 300 images(including 150 original images and 150 steg-images). experiments results show that the classifier was only able to correctly classify 45% of the original image and 54% of the steg-images with a false negative rate of 55% and a false positive rate of 46% respectively.

6. CONCLUSION

In this paper, a newsteganography approach using sensor's pattern noise and linear prediction synthesis filter for anti-steganalysis performance is proposed. The secret message is embedded through the modifying of the extracted pattern noise, whosespectrum structure of pattern noise is preserved. For anti-steganalysis performance of the proposed method, one specific steganalytic method and one blind steganalysis method are test. The experiment results show the effectiveness of the proposed method. However, our scheme still has room for improvements. What kinds of filter operations are optimal for data hiding in images, while cause least image visual and statistical distortion. The data hiding capacity is not so large, future research also should focus on this matter.

7. REFERENCES

[1] X. Yu, T. Tan and Y. Wang, "Isotropy-Based Detection and Estimation: A General Framework of LSB Steganalysis," submitted to IEEE Trans. on Information Forensics and Security.

[2] N. Provos, "Defending against statistical steganalysis, " 10th USENIX Security Symp., Washington DC, USA, 2001.

[3] J. J. Eggers, R. Bauml, and B. Girod, " A communication approach to image steganography, " in Proceedings of SPIE, San Jose, CA, 2002.

[4] J. Fridrich, M. Goljan, and D. Hogea, " Steganalysis of JPE images: Breaking the F5 algorithm, " in LNCS: 5th Int. Workshop on Info. Hiding, 2002, vol. 2578, pp. 3103

[5] Y. Wang and P. Moulin, " Steganalysis of block-DCT imag steganography," in IEEE workshop on Statistical Signal Processing, St Louis, MO, USA, Sept. 2003.

[6] P. Sallee, " Model-based steganography, " in IWDW 200 LNCS 2939, Oct. 2003, pp. 154167.

[7] R. Bohme and A. Westfeld, " Breaking cauchy model-based jpeg steganography with first order statistics, " P. Samarati et al (Eds.): ESORICS 2004, LNCS 3193, pp. 125140, 2004.

[8] J. Lukas, J. Fridrich, and M. Goljan, " Determining Digital Image Origin Using Sensor Imperfections, "Proc. SPIE Electronic Imaging San Jose, CA, January 16-20, pp. 249-260, 2005

[9] J. Lukas, J. Fridrich, and M. Goljan, " Digital Bullet Scratches for Images, " Proc. ICIP 2005, Sep. 11-14, 2005, Genova, Italy.

[10] S. Lyu and H. Farid, " Detecting hidden messages using higher-order statistics and support vector machines, " 5th International Workshop on Information Hiding., 2002.

[11] Kuo, Shyh-Shiaw, Johnston, J.D., " Spatial Noise Shaping Based on Human Visual Sensitivity and Its Application to Image Coding, " IEEE Transactions on Image Processing. **Vol. 11, no. 5** (May 2002) 509–517