International Journal of
Science and Engineering Investigations vol. 1, issue 2, March 2012

A New Steganalysis Method for Steganographic Images on DWT Domain

Saeid Fazli[1], Maryam Zolfaghari-Nejad [2]
[1,2] Electrical Dept., Eng. Faculty, Zanjan University
([1] fazli@znu.ac.ir, [2] Maryam_zolfaghari@znu.ac.ir)

Abstract- In this paper, we introduce a new method for steganalysis of grey-scale images. First, we analyzed the effect of various steganographic processes on the statistical properties of the image. So we extracted the optimal features from the images, which have high ability in make differentiated between two groups of normal and stego images. In this method, high order statistics in discrete wavelet transform (DWT) coefficients are used. Then the pre-processing of principal component analysis (PCA) is done on extracted features. The support vector machines (SVM) is used to classify image segments into stego or non-stego cases. The proposed method with comprehensive look into current steganographic techniques in DWT domain is able to detect the presence of hidden messages with more than 90% accuracy in different embedded rates.

Keywords- *Steganalysis, statistics features, principal component analysis (PCA), support vector machines (SVM).*

I. Introduction

The aim of steganography is to pass hidden information without any suspicion to the existence of the message. Many steganographic techniques are proposed in papers. These techniques embed the data in spatial domain, DCT (Discrete Cosine Transform), DFT (Discrete Fourier Transform) and DWT (Discrete Wavelet Transform). While the aim of steganalysis is to detect and estimate hidden information from observed data with little or no knowledge about the steganography algorithm and parameters [5]. None of the existing steganographic algorithms achieves perfect security and in principle, all are detectable.

In reference [11], can be considered an overview of the types of public steganalysis methods. Some of the most important of them are: the use of Image Quality Matrices (IQM), Higher Order Moments of Wavelet Sub-band Coefficients, Moments of Wavelet Characteristic Function and Histogram Characteristic Function Centre of Mass (HCF-COM). In [8], a steganalysis algorithm based on statistical moments of wavelet characteristic function proposed. The authors extract the first and second statistical moments of the characteristic functions from all the sub-bands to form an 18 dimensional feature vector for steganalysis and use Bayes classifier for classification. Ref. [9] proposed two active steganalysis schemes for spread spectrum image steganography.

In spatial domain, LSB method has long been used by steganographers, because the eye cannot detect the very small perturbations. In ref. [8], Fridrich and long proposed an algorithm for steganalysis of the LSB embedding in 24-bit colour images. This method is based on statistical analysis of the image colours in RGB. Ref. [9] have proposed a new method for detection of LSB data hiding based on Gray Level Co.ocurrence Matrix (GLCM).

In this paper, we proposed the new steganalysis method as the feature-based classification to devise a blind detector specific to grey-scale images. The most important advantage is calculating the GLCM features in the DWT domain that it caused more accuracy and lower error.

The rest of this paper is organized as follows. In section 2, we introduce GLCM features and analyze the effect of various steganographic processes on the statistical properties of the image. Section 3 describes the details for proposed method. Section 4 gives experimental results. This paper concludes in section 5.

II. GLCM Features

This section presents the features we selected to use for classification. Choosing discriminating and independent features is a key to detecting algorithm being successful in classification. The effect of various steganographic methods can be observed on first and second order statistics. The effect of steganography on first order statistics is softening the histogram. Also, the effect of steganography on second order statistics can be observed as reducing the spatial domain correlation or transform domain (DWT, DCT, and DFT).

The co-occurrence matrix is essentially a two-dimensional histogram of the number of times that pairs of intensity values occur in a given spatial relationship [1]. An element P of a GLCM is defined as the joint probability that the gray levels i and j occur separated by distance d and along direction θ in an image. These are properties of pairs of pixel values. The various co-occurrence features such as contrast, energy, entropy, local homogeneity, maximum probability, dissimilarity, cluster shade, cluster prominence, variance, inverse difference moment, Information Measures of Correlation, Autocorrelation, Angular Second Moment and etc are calculated depending on a series of second order statistics computed using the GLCM. These features will be computed in the DWT domain. The co-occurrence matrix describes the probability distribution of pairs of neighboring DWT coefficients.

The discrete wavelet transform (DWT) highlights structural, geometrical and directional features of objects in an image. The Algorithm is based on the fact that the behavior of the carrier image under a smoothing filter is different from the behavior of the stego-image under the same filter [7].

These features obtained from references [1-3]. Where assume $P(i,j)$ to be the (i,j) th entry in a normalized GLCM. The mean and standard deviations for the rows and columns of the matrix are:

$$\mu_x = \sum_i \sum_j i.P(i,j), \quad \mu_y = \sum_i \sum_j j.P(i,j)$$

$$\sigma_x = \sum_i \sum_j (i-\mu_x)^2.P(i,j),$$

$$\sigma_y = \sum_i \sum_j (i-\mu_y)^2.P(i,j),$$

Then the features are as follows:

1) Energy [1]:

$$f_1 = \sum_i \sum_j \{P(i,j)\}^2,$$

2) Contrast [1]:

$$f_2 = \sum_{n=0}^{N-1} n^2 \{\sum_{i=1}^{N} \sum_{j=1}^{N} P(i,j) \| i-j |= n\},$$

3) Correlation [1]:

$$f_3 = \frac{\sum_i \sum_j (ij)P(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y},$$

4) Entropy [2]:

$$f_4 = -\sum_i \sum_j P(i,j) \log(p(i,j)),$$

5) Homogeneity [2]:

$$f_1 = \sum_i \sum_j \frac{1}{1+(i-j)^2} P(i,j),$$

6) Autocorrelation [2]:

$$f_6 = \sum_i \sum_j (ij)P(i,j),$$

7) Dissimilarity [2]:

$$f_7 = \sum_i \sum_j | i-j | .P(i,j),$$

8) Cluster Shade [2]:

$$f_8 = \sum_i \sum_j (i+j-\mu_x-\mu_y)^3.P(i,j),$$

9) Cluster Prominence [2]:

$$f_9 = \sum_i \sum_j (i+j-\mu_x-\mu_y)^4.P(i,j),$$

Cluster shade and cluster prominence can be modified to a sum histogram problem.

10) Maximum Probability:

$$f_{10} = \underset{i,j}{MAX} \; p(i,j)$$

11) Sum of Squares: Variance [1]:

$$f_{11} = \sum_i \sum_j (i-\mu)^2 P(i,j),$$

μ is the mean value of *P*.

12) Inverse Difference Moment [3]:

$$f_{12} = \sum_i \sum_j \frac{1}{1+(i-j)^2} P(i,j),$$

13) Sum Average [1]:

$$f_{13} = \sum_{i=2}^{2N} iP_{x+y}(i),$$

N is the number of gray levels used.

14) Sum Variance [1]:

$$f_{14} = \sum_{i=2}^{2N} (i-f_s)^2 P_{x+y}(i),$$

15) Sum Entropy [3]:

$$f_{15} = -\sum_{i=2}^{2N} P_{x+y}(i) \log\{P_{x+y}(i)\},$$

16) Difference Variance [3]:

$$f_{16} = Variance\ of\ \ p_{x-y},$$

17) Difference Entropy [3]:

$$f_{17} = -\sum_{i=0}^{N-1} P_{x-y}(i) \log\{p_{x-y}(i)\},$$

18) Angular Second Moment [1]:

$$f_{18} = \sum_{i}\sum_{j}\{P(i, j)\}^2,$$

19) Gray Level Co-occurrence Mean [2]:

$$f_{19} = \sum_{i,j=0}^{N-1} iP(i, j),$$

20) Information Measures of Correlation [3]:

$$f_{20} = \frac{HXY - HXY1}{\max\{HX, HY\}},$$
$$HXY = -\sum_{i}\sum_{j} P(i, j)\log(P(i, j)),$$
$$HXY1 = -\sum_{i}\sum_{j} P(i, j)\log\{P_x(i)P_y(j)\}$$

Where *HX* and *HY* are entropies of P_x and P_y .

III. Proposed method

In this paper, we proposed an algorithm for targeted steganalysis that capable to detect specific steganographic algorithms. The algorithms which embed the data in spatial domain such as lsb encoding, lsb matching and methods in ref. [10-12]. We found that even embedding in spatial domain, changes the correlation between wavelet coefficients. With this fact, we have analyzed the sub-bands of DWT and extract the features of classifier.

The proposed method is an algorithm for steganalysis as block wise where each block has been classified independently. So each image is divided into 8×8 blocks and for each sub-image the DWT were applied. The Discrete Wavelet Transform (DWT) of images produces a non-redundant image representation which provides better spatial and spectral localization of image formation. The sub-image is decomposed using a 1-level two-dimensional DWT to obtain four sub-bands as cA, cH, cV and cD. In our method, we select the diagonal detail coefficient (cD) for obtaining the necessary features. The difference between the DWT coefficients of each block and the same stego-block reveals the existence of hidden information. Through cD, cH and cV, the diagonal detail coefficient is suitable for steganalysis of steganography methods in spatial domain such as LSB or proposed steganography method in ref. [10-12].

Every test sub-image's cD was filtered by Gaussian kernel. Then from the resulting filtered cD, some quality features were calculated using the GLCM.

These effective features are proposed in section 2. Feature selection is an important issue in classification. We have applied PCA to improve the detection accuracy of SVM classifier and reduced the false positive rate.

Principal Component Analysis (PCA) is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components. We used the PCA, to 20 features be reduced to 10 features.

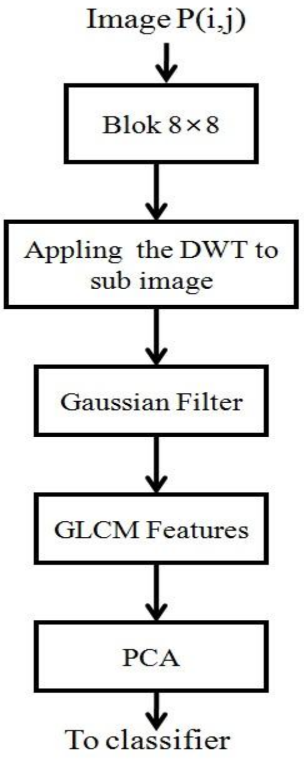

Fig. 1. Block diagram of proposed method

For classification, Support Vector Machine (SVM) was selected. One of the reasons for the popularity of SVMs is that they are considered resistant to the curse of dimensionality and to uninformative features [13]. SVM is a powerful tool for pattern classification. With introduction of kernel ticks in SVM, it has become a very popular in machine learning community [4]. Fig. 1 shows the block diagram of steganalysis method. This steganalysis technique can detect the existence of hidden massage in images that are embedded data with various algorithms in spatial. Each algorithm that changes the pixel values in spatial domain can be detected with proposed method. Some of these algorithms are ref. [10-12].

IV. EXPRIMENTAL RESULTS

To evaluate the algorithm, we first created a set of stego images using embedding algorithms in ref. [10-12] in the spatial domain using different relative massage length. In this section, we demonstrate the detestability of proposed method using a classifier based on the feature set. Statistical delectability is evaluated by support vector machine (SVM) with Gaussian kernel. We used a data base of 1000 grey-scale images from stego and non-stego images with 512×512 size. The database was divided into 500 training images and 500 testing images. We created stego images from normal images and obtained training and testing images. Table I shows the result of classification for steganography algorithms. The result demonstrates that our method can achieve good distinction between stego and non-stego images. For all the steganography algorithms, the detection is about 97%.

TABLE I
THE RESULT OF STEGANALYSIS

Steganography Methods	False Reject Rate	False Accept Rate	Correct Detection Rate
LSB	**0.6**	**0.8**	**98.6**
Method in [10]	**0.3**	**0.5**	**99.2**
Method in [11]	**0.9**	**1.1**	**98**
Method in [12]	**1.1**	**1.4**	**97.5**

V. CONCLUSIONS

By studying the steganography algorithms, we also discover important consequences that influence the development and future design of steganalysis. By realizing the limits and by improving the practical stegosystems, we can design the more secure steganography algorithms. In this paper, we developed a new targeted steganalysis method based on GLCM features for grey-scale images. Each feature is calculated in DWT coefficients. We have applied the detection to several current steganographic schemes. The experimental results were carefully evaluated and we concluded that secure steganographic schemes must preserve as many statistics of DWT coefficients as possible.

REFERENCES

[1] R. M. Haralick, K. Shanmugam, and I. Dinestein, “Textural Features of Image Classification,” *IEEE Transactions on Systems*, vol. SMC-3, No. 6, Nov. 1973.

[2] L. Soh and C.Tsatsoulis, “Texture Analysis of SAR Sea Tce Imagery Using Gray Level Co-occurrence Matrices,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, No. 2, March 1999.

[3] D A. Clausi, “An Analysis of Co.occurrence Texture Statistics as a Function of Gray Level Quantization,” *Can. J. Remote Sensing*, Vol. 28, No. 1, pp. 45-62, 2002.

[4] J. Fridrich, M. Goljan and D. Soukal, “Higher-order statistical steganalysis of palette images,” *Proc. EISPIE Santa Clara*, CA, Jan 2003, pp. 178-190.

[5] R. Popa, “An analysis of steganographic techniques,” *Master Thesis, Polytechnic University of Timisoara. Department of computer science and software engineering*, 1998, pp. 34-50.

[6] T.Pevny, J.Fridrich, “Multi-Class Detector of Current Steganographic Methods for JPEG Format,” *Department of Electrical and Computer Engineering,* Bingamton, 2008.

[7] T.pevny, J.Fridrich, “Merging Markov and DCT Features for Multi-Class JPEG Steganalysis,” *In E.J. Delp and P.W. Wong. Editors, Proc, SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Content IX*, vol. 6505, pp.03-04, January 2007.

[8] J. Fridrich, M. Goljan and R. Du, “Detecting LSB steganography in color and gray-scale mages,” *IEEE multimedia*, vol. 8, no. 4, pp. 22-28, 2001.

[9] M. Abolghasemi, H. aghainia, K.faez, M.A.Mehrabi, “steganalysis of LSB matching based on co-occurrence matrix and removing most sighnificant planes,” *International Conference on Intelligent Information Hiding and Multimedia Signal Processing,* 2008, pp. 1527-1530

[10] Jain A.K and Uludag U., “Hiding Biometric Data,” *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 25, NO. 11, pp. 1494-1498, November 2003.

[11] A. Sur, P. Goel and J. Mukhopadhyay, “A Spatial Domain Steganographic Scheme For Reducing Embedding Noise*,” 3rd international symposium on communications, control and signal processing,* ISCCSP March 12-14,2008, ST. Julians, Malta, on page(s): 1024-1028.

[12] U. Uludag, B. Gunsel, M. Ballan, “A Spatial Method For Watermarking of fingerprint images,” *Proceedings of the first international workshop on pattern recognition in information systems*, PRIS 2001, Sebutal, Portugal, ICEIS Press, July 2001, pp. 26-33.

[13] A. Abu-EErrub, A. Al-Haj, “Optimized DWT Based Image Watermarking,” *IEEE*, 2008

[14] R.safarbakhsh, S.zzaboli, A. Tabibiazar, “Digital Watermarking on Still Images Using Wavelet Transform,” *in Proc. International Conference on Information Technology: Coding and Computing ITCC’04,IEEE,* 2004.

[15] J. Fridrich and M.Goljan. “Digital image steganography using stochastic modulation”. In E. J. Delp and P.W. Wang, editors, *Proceeding SPIE, Electronic Imaging, Security, steganography and Watermarking of multimedia contents V,* volume 5020, pages 191-202, Santa Clara, CA, January 21-24, 2003.