International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 3 Issue 4, April - 2014

Blind Image Steganalysis using Co-Occurrence Matrix of Difference Wavelet Coefficients

Syama S.
Department of computer Engineering
College of Engineering, Chengannur
Alappuzha, India

Deepa J.
Associate Professor, Department of Electronics Engineering
College of Engineering, Chengannur
Alappuzha, India

***Abstract*— Steganalysis is the art and science of discrimination between stego objects and cover objects. By means of steganalysis message does not have to be extracted, just its presence is detected. Steganalysis performed without any knowledge of the embedding algorithm and embedding secret key is referred to as blind steganalysis. The aim of this project is to develop a new blind steganalysis approach using wavelet transform domain. The work consists of two phases-feature extraction and classification. In the first phase of the work, features for classification are extracted based on the correlation of two level wavelet coefficients. In the second phase backpropagation neural network is trained using these extracted features to classify cover objects and stego objects. Although there are many approaches to blind steganalysis, those may require more features and thus calculations are complex. This proposed method uses minimum number of features for the classification, thus computational complexity can be reduced. In experiments, the proposed method is applied to different sizes of embedding message. Detection accuracy is also checked with different feature combinations. The images obtained from wavelet and DCT based embedding techniques are used for training and testing. Experimental results shows that the proposed method gives better accuracy with minimum number of features.**

Keywords—Steganalysis, Blind steganalysis, wavelet transform, Correlation, Backpropagation neural network

I. INTRODUCTION

Steganography literally means, "covered writing" and encompasses methods of transmitting secret messages through innocuous cover carriers in such a manner that the existence of the embedded messages is undetectable. The medium to which the message is embedded is called cover object. The result obtained after embedding the message into the cover object using any stegnographic algorithm is called as the stego object.

As the opposite technology of stegnography, steganalysis is the art and discrimination between stego objects and cover objects. Steganalysis needs to be done without any knowledge of secret key used for embedding and maybe even the embedding algorithm. The vast majority of work in steganalysis focuses on detection of secret messages rather than extraction.

Hiding information within electronic media requires alterations of the media properties that may introduce some form of degradation or unusual characteristics. These characteristics may act as signatures that broadcast the existence of the embedded message, thus defeating the purpose of steganography. These information are used for steganalysis.

Steganalysis can be broadly classified into two classes: signature steganalysis and statistical steganalysis [1]. The division is based on whether the signature of the steganography technique or the statistics of image is used to identify the presence of concealed messages in images embedded using steganography. The statistics of an image undergo alterations due to information hiding, statistical steganalysis, as the name implies, analyses this underlying statistics of an image to detect the secret embedded information. Information hiding within any electronic media using steganography requires alterations of the media properties that may introduce some form of degradation or unusual characteristics and patterns. These patterns and characteristics may act as signatures that broadcast the existence of embedded message. This information is used in signature steganalysis. Based on its application fields, statistical and signature steganalysis can be further divided into specific methods and universal (blind) methods. A specific steganalytic method utilizes the knowledge of a targeted steganographic technique and may only be appropriate to such a kind of steganography. The detection accuracy of specific steganalysis methods will be high, but it is useless for other stegnographic algorithms. A universal or blind steganalytic method is used to detect several kinds of steganography. Usually universal methods do not require the knowledge of the embedding operations. Hence, it is also called blind method. The blind steganalysis does not try to determine the specific hiding method, but makes the correct identifications of detected object that may contain hidden information. Therefore compared with specific steganalysis methods, even if the accuracy is less blind steganalysis is more practical.

The goal of steganalysis is to collect the ample evidence about the presence of embedded message. The importance of steganalytic techniques is increasing. Steganalysis finds its use in computer forensics for screening and tracking documents that are suspect of criminal activities, cyber warfare, tracking the criminal activities over the internet. Steganalysis system that is capable of classifying cover image and stego image which helps the steganalyst to reduce the set of data files to the subset most likely to have been altered. Steganalysis also provides an effective way to evaluate the security performance of a data hiding method. That is, it can be used to improve the security of a data hiding algorithm. Thus, a good data hiding method should be able to hide data imperceptibly not only to human eyes, but also to computer analysis.

II. EARLY METHODS OF BLIND STEGANALYSIS

Various blind steganalysis approaches can be broadly classified into spatial domain approach and transform domain approach. Spatial domain approaches uses various image characteristics as feature vector. Transform domain approaches uses various image statistics in transform domain as features. Most of the previous works uses support vector Machine as the classifier.

Spatial Domain Approaches in blind steganalysis includes runlength histogram analysis and steganalysis using rate distortion characteristics. Runlength Histogram analysis is based on the assumption that the statistical moments of the characteristic function (CF) of certain histogram are very effective features and utilizing multi-order moments of CF of the three types of image runlenght histograms (RLH) as features is more effective [2] .In blind steganalysis using Rate distortion characteristics, feature vector for the discrimination between cover image and steg image is taken based on the rate distortion characteristics. The features are taken from two observations 1) data hiding techniques increases image entropy to hide hidden message 2) Data hiding is limited to a small, imperceptible distortions .[3]

Various transform domain approaches based on fourier transform, discrete cosine transform and wavelet transform are proposed in various works. There are various fourier transform methods are proposed for blind image steganalysis. One of the famous work proposes the use of histogram and discrete fourier transform in blind image steganalysis. For calculating the feature vector for classifier training, histogram of each color plane is taken, then mean and variance of histogram coefficients of each color plane are calculated. These are the first set of features. Then DFT for histogram of each color component is applied, then mean, variance, skewness, kurtosis ,energy are calculated as features.1st order moment of difference of histogram and DFT is also calculated as feature . In this work they used SVM as the classifier. One of the main disadvantage of this work is that more number of features used to train SVM [4].

Discrete cosine transform based blind steganalysis approaches are common in blind steganalysis. One of the steganalysis approach is based on markov process. They proposed that the difference JPEG 2-D arrays along horizontal, vertical, diagonal and minor diagonal directions have enlarged changes caused by steganographic methods. JPEG 2-D arrays is made up of magnitude of DCT block coefficients. Then Markov process is applied to model the difference JPEG 2D arrays. Transition probability matrices is then used to represent markov process. All the elements of transition probability matrices are taken as features. Thresholding technique is used to reduce dimensionality of feature vectors to a manageable extent [5]. Another DCT domain approach is based on the correlations between block DCT coefficients in both intra-block and inter-block sense. Then Markov empirical transition matrices to capture these dependencies. The elements of these transition matrices are taken as the feature vector [6]. There is another approach that combines the features from DCT domain as well as spatial domain. Histogram characteristic function which is the image histogram in frequency domain and centre of mass , which is the energy distribution in HCF is taken as features. They proposed to take the histogram of DCT coefficients and the co-occurrence matrix of coefficients in adjacent DCT blocks [7].

The statistical model for the detection of stego objects can be created using wavelet transform. For extracting the features first the image is decomposed using wavelet decomposition. Then mean, variance, skewness, kurtosis of subband coefficients at each orientation and scale is calculated. These features forms the first set of features. These statistics characterize the basic coefficient distribution. Second set of statistics is calculated based on the errors in an optimal linear predictor of coefficient magnitude. Both these statistics are combined to get the feature set for classification [8]. This same method is modified by applying this concept to color images. There wavelet decomposition is independently applied to each of the color channel [9].Then mean, variance, skewness, kurtosis of subband coefficients at each orientation and scale is calculated for each color channel. The statistics based on the errors in the linear predictor is also applied to each of the color channels. Another method to calculate the features for steganalysis only from the noise component of the stego image in the wavelet domain is also proposed [10]. The noise component was obtained using the denoising filter. It is assumed that the denoising step increases the SNR between the stego signal and the cover image, thus making the features calculated from the noise residual more sensitive to embedding and less sensitive to image content. The denoising filter is designed to remove Gaussian noise from images under the assumption that the steg image is an additive mixture of a non-stationary Gaussian signal (the cover image) and a stationary Gaussian signal with a known variance (the noise). As the filtering is performed in the wavelet domain, all our features (statistical moments) are calculated as higher order moments of the noise residual in the wavelet domain. The higher order absolute moments of the noise residual is actually calculated as features. One of the improved approach to blind image steganalysis is done by using both the magnitude statistics and phase statistics [11]. The magnitude statistics include mean, variance, skewness, kurtosis of the subband coefficients at each orientation, scale and color channel. While these statistics characterize the basic coefficient distributions, they are unlikely to capture the strong correlations that exist across space, orientation, scale and color. In order to capture some of these higher-order statistical correlations, a second set of statistics that are based on the errors in a linear predictor of coefficient magnitude is also extracted. Phase statistics calculated using local angular harmonic decomposition is used as the second half of the feature vector.

Even if there exist many spatial domain and frequency domain approaches to blind steganalysis, they are not efficient since those methods require extraction of more features for the classification and thus the computational complexity is high. Most of the steganalysis methods uses only one dimensional and local features for the detection of stego objects, thus the accuracy is also less. This project proposes a blind steganalysis using minimum number of features that gives better accuracy.

III. PROPOSED BLIND STEGANALYSIS METHOD

We proposes a new multi feature blind steganalysis approach that is based on the assumption that, there exists some correlation or relationship between the wavelet coefficients within each subband and also between the coefficients in the adjacent subbands.

There are two phases included in the proposed method: feature extraction and classification. In the feature extraction s phase, features for classification are extracted using inter and intra subband difference co-occurrence matrices. A backpropagation neural network is used for classification. The extracted features are fed as input to a trained backpropagation neural network (BPNN) to obtain the class label (cover image or steg image) .The features of a set of cover images and steg images are used for training the backpropagation network. The proposed system architecture is shown in Fig 1.

The left side of the Fig 1 represents the training phase and right side represents the testing phase. During training phase, the training images are first decomposed in the wavelet domain. Then inter and intra wavelet subband difference matrices are calculated, co-occurrence matrices are calculated for these difference matrices. The features are then extracted from the co-occurrence matrices. Then features are fed as input to BPNN to train the network. In the testing phase, the test image is processed similarly to extract features. The extracted features are then fed as input to the trained BPNN to get the class label, ie cover image or steg image.

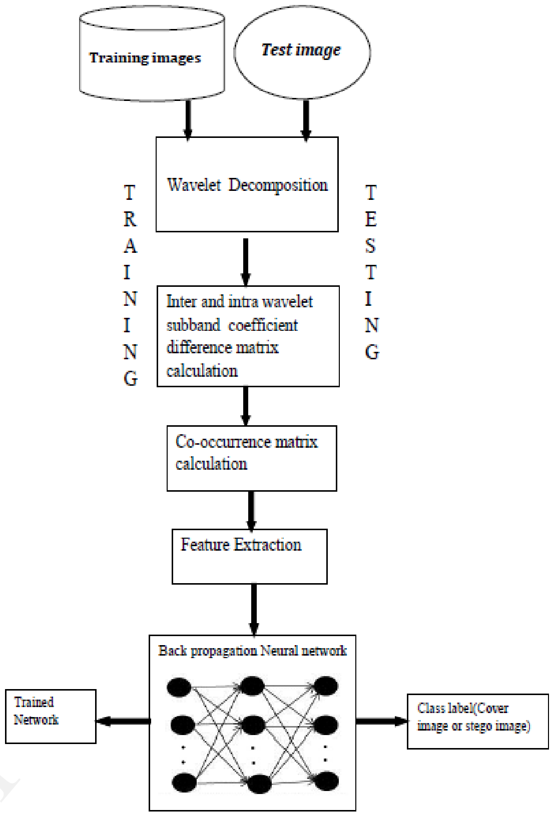


Fig 1.Proposed Blind steganalysis Method

A. *Feature extraction*

The first step in the feature extraction is wavelet decomposition. The image is first decomposed into two levels in the wavelet domain. Then the adjacent coefficient difference in horizontal, vertical, diagonal directions are calculated with in each subband. Then difference matrices are calculated between subbands. Next step is co-occurrence matrix calculation of these difference matrices. Then features are extracted from the co-occurrence matrices. Next set of features are calculated from the three level decomposed subbands of the image.

Step1: Wavelet Decomposition

After the secret message is embedded in a given image, there will be some changes in the pixel values of the image. According to wavelet theory, these changes may be reflected in the high frequency wavelet subbands. In wavelet domain subtracting adjacent coefficient reflect the relationship between those coefficients as well as the relationship between the pixel blocks constructing these coefficients. We use this concept for steganalysis. The high frequency subbands are processed to extract the features. Since the embedding process changes the higher frequency wavelet coefficients, the features extracted from those coefficients will be sufficient to perform steganalysis. The two level decomposition of standard Cameraman image is shown in Fig 2.

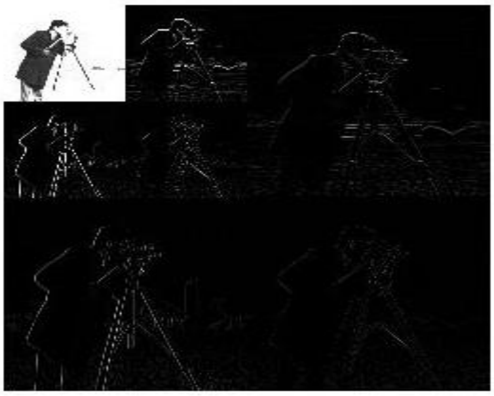


Fig 2. Two-level wavelet decomposition map of Cameraman image

Step 2: Difference matrices calculation

a) Difference matrices within subbands

Difference matrices with in each high frequency subbands in horizontal, vertical, diagonal directions are calculated after two level decomposition. Since there are six higher frequency subbands after two level decomposition, 18 difference matrices will be obtained. The following figure Fig 3 shows intra subband difference matrix calculation along three directions.

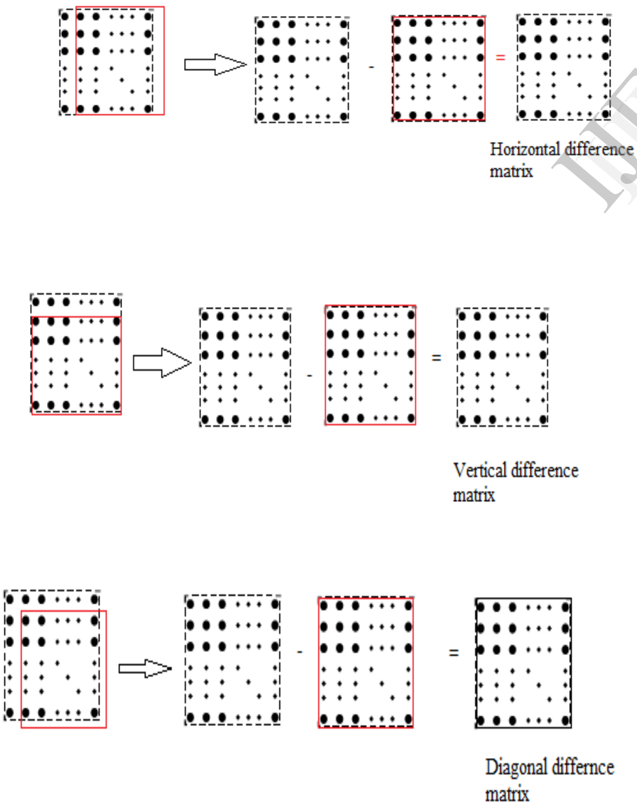
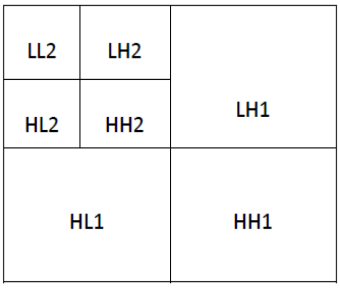


Fig 3. Difference matrix calculation

b) Difference matrices between subbands

Difference matrices are calculated between the higher frequency subbands in the first level and the difference matrices between all the subbands are calculated in the second level subbands.

Three difference matrices are calculated in the first level, and six difference matrices are calculated in the second level. Totally nine difference matrices will be obtained. This is shown in Fig 4.



1st level difference matrices

D1=LH1-HL1

D2=LH1-HH1

D3=HL1-HH1

2nd level difference matrices

D4=LL2-LH2

D5=LL2-HL2

D6=LL2-HH2

D7=LH2-HL2

D8=LH2-HH2

D9=HL2-HH2

Fig 4: Difference matrix calculation between subbands

Step 3: Co-occurrence matrix calculation and feature extraction

Co-occurrence matrices are used to represent the distribution of co-occurring values at a given offset. Here the co-occurrence matrices of the difference matrices within and between subbands are calculated. There are 18 difference matrices calculated within subbands, 18 co-occurrence matrices then calculated. From the 9 difference matrices between subbands, co-occurrence matrices are calculated along three direction obtaining 27 co-occurrence matrices.

Entropy and energy are extracted from all the co-occurrence matrices. From the co-occurrence matrices calculated for within subband difference matrices 36D feature is thus obtained. From the co-occurrence matrices calculated for between subband difference matrices 54D feature is obtained.

The PDF moments of all the higher frequency subbands of three level decomposition up to 4th order is also calculated as feature obtaining 36D feature. Finally 126D feature vector is computed.

B. Classification using Backpropagation network

A backpropagation neural network is used for of input classification. The backpropagation neural network consists of input layer, hidden layer and output layer. The number of input layer neuron corresponds to the dimension of the feature vector. There is one hidden layer that consists of 5 neurons. There is one output layer neuron and output range of the output layer neuron is [0,1]. The decision threshold is set as 0.5. If the output value is less than that threshold the image is classified as steg image, otherwise it is classified as cover image.

IV. RESULTS AND DISCUSSIONS

A. Experimental Setup

The original images used in experiments are taken from various image databases available on Web. The steg images are created using Wavelet based and DCT based embedding methods. Totally 600 images are used for training and 480 images are used for testing. 180 steg images are created by embedding message of size 256×256 by Wavelet based embedding method and 180 steg images are also created by DCT based embedding method. Similarly steg images are created by embedding message size 128×128 and 64×64.

B. Experimental Results

The experiments were conducted separately for DCT based embedding method and Wavelet based embedding method. The experimental results are shown in the following table.

Table 1. Accuracy for different watermark size

Watermark size	Wavelet transform based embedding method	DCT based embedding method
256×256	97.5	96.25
128×128	96.25	96.25
64×64	93.75	96

Table shows that the method gives better results for wavelet based and DCT based embedding method. As the embedding message size decreases the detection accuracy decreases.

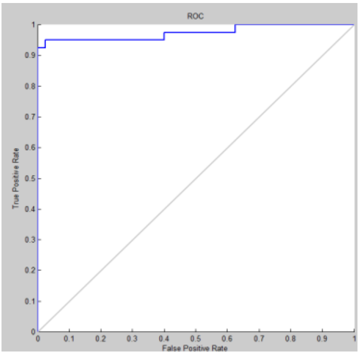


Fig 5. ROC curve for DCT based embedding steganalysis

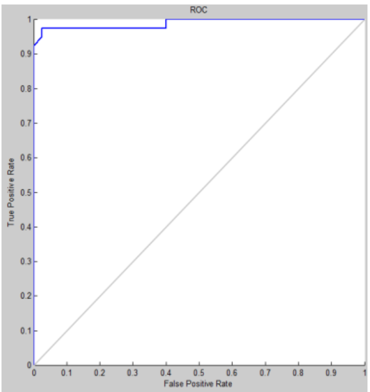


Fig 6. ROC curve for wavelet based embedding steganalysis

The results shows that if the transform domain used for embedding and steganalysis are same, then the detection accuracy will be same.

V. CONCLUSIONS

A new multifeature blind image steganalysis approach based on the correlation between wavelet coefficients with in subbands and correlation between the coefficients in adjacent subbands is proposed in this work. The backpropagation neural network is used for classification. The results shows that the method gives better accuracy for Wavelet based and DCT based embedding method. As the embedding message size increases the detection accuracy also increases.

ACKNOWLEDGMENT

We would like to take this opportunity to express our gratitude to all those who have guided in the successful completion of this endeavor. We are very thankful to the staffs of College of Engineering, Chengannur for their valuable suggestions.

REFERENCES

1. Han Zong, Fen-lin Liu, Xiang-yang Luo "Blind image steganalysis based on wavelet coefficient correlation " Digital Investigation 9 (2012)
2. Jing Dong and Tieniu Tan "Blind image steganalysis based on run-length histogram analysis."
3. M.U Celik,Gaurav Sharma,A.Murat Tekalp 'Universal Image Steganalysis using rate-distortion curves" In: Proceedings, SPIE security, steganography, and water-marking of multimedia contents VI, vol. 5306. 2004.
4. T. H. Manjula Devi "Detecting Original Image Using Histogram, DFT and SVM" International Journal of Recent Trends in Engineering Vol. 1, No. 1, May 2009
5. Shi YQ, Chen C, Chen W. "A Markov process-based approach to effectively attacking JPEG steganography."In Proceedings, 8th information hiding workshop. LNCS, vol. 4437; 2007. p. 24964.
6. Dongdong Fu, Yun Q. Shi, Dekun Zou,Guorong Xuan"JPEG Steganalysis Using Empirical Transition Matrix in Block DCT Domain"
7. Zhuo Li, Kuijun Lu, Xianting Zeng, Xuezeng Pan \A Blind Steganalytic Scheme Based on DCT and Spatial Domain for JPEG Images" Journal of multimedia vol 5,no 3,June 2010
8. Farid H, Lyu S. "Detecting hidden messages using higher-order statistics and support vector machines." In: Proceedings, 5th international information hiding workshop. LNCS, vol. 2578; 2002. p. 34054.
9. Lyu S, and Farid H. "Steganalysis using color wavelet statistics and oneclass support vector machines." In: Proceedings, SPIE symposium on electronic imaging, vol. 5306. 2004. p. 3545.
10. Lyu S, Farid H. "Steganalysis using higher-order image statistics." IEEE Transactions on Information Forensics and Security 2006;1(1):1119.
11. Goljan M, Fridrich J, and Holotya T. "New blind steganalysis and itsimplications." In: Proceedings, SPIE security, steganography, and watermarking of multimedia contents VIII, vol. 6072. 2006. p. 113.
12. Robi Polikar,"The Engineer's Ultimate Guide to Wavelet Analysis-The Wavelet Tutorial"