INTRA BLOCK AND INTER BLOCK NEIGHBORING JOINT DENSITY BASED APPROACH FOR JPEG STEGANALYSIS

Arun R[1], Nithin Ravi S[2] and Thiruppathi K[3]

[1,2,3]TIFAC CORE in Cyber Security, Amrita Vishwa Vidhyapeetham, Coimbatore
[1]csarunr@gmail.com , [2]nithinravi535@gmail.com, [3]ktirupathi1988@gmail.com

ABSTRACT

Steganalysis is the method used to detect the presence of any hidden message in a cover medium. A novel approach based on feature mining on the discrete cosine transform (DCT) domain based approach, machine learning for steganalysis of JPEG images is proposed. The neighboring joint density on both intra-block and inter-block are extracted from the DCT coefficient array. After the feature space has been constructed, it uses SVM like binary classifier for training and classification. The performance of the proposed method on different Steganographic systems named F5, Pixel Value Differencing, Model Based Steganography with and without deblocking, JPHS, Steghide etc are analyzed. Individually each feature and combined features classification accuracy is checked and concludes which provides better classification.

KEYWORDS

Steganography, Steganalysis, DCT, PVD, MB1, MB2,F5, JPHS, Steghide.

1. INTRODUCTION

Steganography is the art of passing information through apparently innocent files in a manner that the very existence of the message is unknown. Steganalysis is the art of discovering hidden data in cover objects. As in cryptanalysis, we assume that the steganographic method is publicly known with the exception of a secret key. The method is secure if the stego-images do not contain any detectable artifacts due to message embedding. In other words, the set of stego-images should have the same statistical properties as the set of cover-images [1]. If there exists an algorithm that can guess whether or not a given image contains a secret message with a success rate better than random guessing, the steganographic system is considered broken.

Steganalysis is the art and science to detect whether a given medium has hidden message in it. Also steganalysis can serve as an effective way to judge the security performance of steganographic techniques. Steganalysis can be mainly classified into two-Blind Steganalysis and Targeted Steganalysis [2]. Targeted Steganalysis are designed for a particular steganographic algorithm. Blind Steganalysis are schemes which are independent of any specific embedding technique are used to alleviate the deficiency of targeted analyzers by removing their dependency on the behavior of individual embedding techniques. To remove their dependency a set of distinguishing statistics that are sensitive to a wide variety of embedding operations are determined and collected. These statistics are taken from both cover and stego images and are

DOI: 10.5121/ijsc.2012.3206

used to train a classifier, which is latter used to distinguish between cover and stego images. Hence, the dependency on a specific embedder is removed using these statistics.

Universal steganalysis is composed of feature extraction and feature classification. In feature extraction, a set of distinguishing statistics are obtained from a data set of images. In feature classification the obtained distinguishing statistics from both stego and cover images are used to train a classifier and finally the trained classifier is used to classify an input image as either being a stego image which carrying a hidden data or a clear image. The above statistics are obtained by observing general image features that exhibit strong variation under embedding.

In [3] Fridrich proposed a universal steganalysis scheme specially designed for JPEG steganography. A set of 23 distinguishing features from the block discrete cosine transform(BDCT) domain and spatial domain is proposed. In [4] Shie et al. presented a new universal steganalysis secheme in which all the 324 features are calculated directly from the quantized DCT coefficients. The Markov process is applied to modeling the difference of JPEG 2D arrays along horizontal, vertical and diagonal directions so as to utilize the second order statistics for steganalysis. In [4] Fu et al. presented another universal JPEG steganalysis scheme totally based on quantized DCT coefficients which extracted 200 features. The Markov empirical transition matrices are used to exploit the magnitude correlations between BDCT coefficients in both intra and inter block. By extending the feature set in [3] and applying calibration to the Markov features a new JPEG steganalysis scheme is developed by Penvy et al. [5] with 274 features. In [6] Qingzhong et al. proposed a new approach based on feature mining on the discrete cosine transform (DCT) and machine learning for steganalysis of JPEG images. The neighboring joint density features on both intra and inter block are extracted from the DCT coefficient array and the absolute coefficient array.

In this paper we propose a new steganalysis scheme to attack some latest developed JPEG steganographic schemes. In this features from neighboring joint probabilities of the DCT coefficients on intra and inter block are extracted. The paper is organized as follows. In the next section the statistical models and information hiding is explained. In the section 3 the neighboring joint density features on both intra and inter block features are explained. In section 4 proposed method is explained followed by the experimental results of the proposed steganalysis method in section 5. And paper is concluded in section 6.

2. STATISTICAL MODELS AND INFORMATION HIDING

A model Probability density function (PDF) can characterize the statistical behavior of a signal. For multimedia signals, the Generalized Gaussian distribution (GGD) is often used. GGD can be applied to model the distribution of Discrete Cosine Transform (DCT) coefficients, the wavelet transform coefficients, pixels difference, etc. Thus, it might be used in video and geometry compression, watermarking, etc. GGD is also known in economy as Generalized Error Distribution (GED). Probability density function of the continuous random variable of GGD takes the form [5]

$$p(x;\alpha,\beta) = \frac{\beta}{2\alpha\cdot\Gamma(1/\beta)}\, e^{\{-(|x|/\alpha)^{\beta}\}}$$

where $\Gamma(z) = \int_0^{\infty} e^{-t}t^{z-1}dt, z > 0$

$\Gamma(z)$ is the Gamma function, scale parameter α models the width of the PDF peak and shape parameter β models the shape of the distribution. Their exists the dependency between the

compressed DCT coefficients and their neighbors. The information hiding will modify the neighboring joint density of the DCT coefficients. Let the left or upper adjacent DCT coefficient be denoted by random vector X_1 and the right or lower adjacent DCT coefficients be denoted by random vector X_2; let $X = (X_1,X_2)$. When hidden data are embedded in the compressed DCT domain in JPEG images by using any steganographic algorithms the DCT neighboring joint probability density coefficients is affected and these changes will be helpful for steganalysis.

The change in joint density due to message embedding is shown by the following example. Figure 1 shows the cover image, F5 embedded image and the steghide embedded image. Figure 2 shows the compressed DCT neighboring joint density probability, the neighboring joint density distribution of a F5 steganogram carrying some hidden data and the neighboring joint density distribution of a steghide steganogram carrying some hidden data. From figure 2 it is clear that the neighboring joint density is approximately symmetric about the origin. Figure 3 shows the difference of neighboring joint density of F5 steganogram and steghide steganogram with cover image. So by embedding message the neighboring joint density get modified.

Figure 1. Cover Image, F5 Steganogram and Steghide Steganogram in (a), (b) and (c).

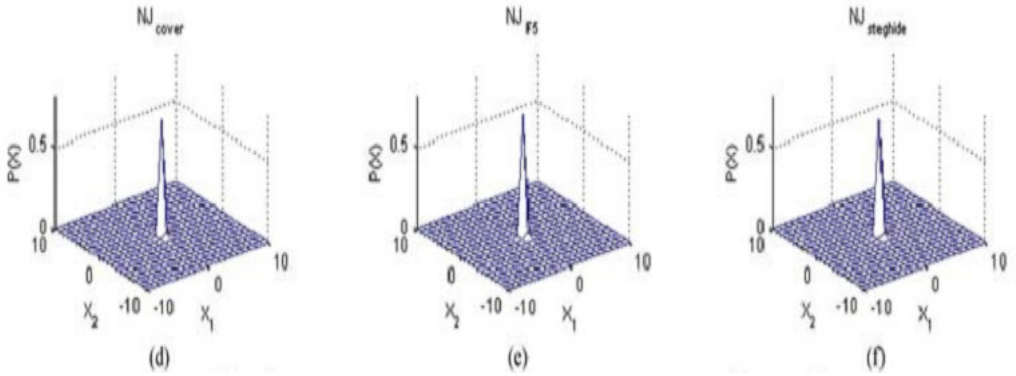

Figure 2. Compressed DCT neighboring joint density probability, the neighboring joint density distribution of a F5 steganogram and the neighboring joint density distribution of a steghide steganogram in (d), (e) and (f).

Figure 3. Difference of neighboring joint density of F5 steganogram and steghide steganogram with cover image in (g) and (h).

3. NEIGHBORING JOINT DENSITY FEATURES

The dependency between compressed DCT coefficients and their neighbours is explained in [5]. The information hiding will modify the neighboring joint density. When messages are embedded in the compressed DCT domain in JPEG images by any of the steganographic algorithms the DCT neighboring joint density probability density is affected which will gives a way for steganalysis. The modification of joint densities as a result of data embedding is shown in [6].

3.1. Feature Extraction

The neighboring joint features are extracted on intra-block and inter-block from the DCT coefficient array respectively.

From the DCT coefficient array the neighboring joint density of intra block and inter block features are extracted as shown below. Let F denote the compressed DCT coefficient array of a JPEG image, consisting of M × N blocks F_{ij} (i = 1, 2.. M; j = 1, 2.. N). Each block has a size of 8 × 8. The intra-block neighboring joint density matrix on horizontal direction $NJ_{1h}$and the matrix on vertical direction NJ_{1v} are constructed as

$$NJ_{1h}(x,y) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{m=1}^{8}\sum_{n=1}^{7}\delta(C_{ijmn}=x, C_{ijm(n+1)}=y)}{56MN}$$

$$NJ_{1v}(x,y) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{m=1}^{7}\sum_{n=1}^{8}\delta(C_{ijmn}=x, C_{ij(m+1)n}=y)}{56MN}$$

where C_{ijmn} stands for the compressed DCT coefficient located at the m^{th} row and the n^{th} column in the block F_{ij} ; δ = 1 if its arguments are satisfied, otherwise, δ = 0; x and y are integers. For computational efficiency, we define NJ_1 as the neighboring joint density features on intra-block, calculated as follows:

$$NJ_1(x,y) = \frac{NJ_{1h}(x,y) + NJ_{1v}(x,y)}{2}$$

Here the values of *x* and *y* are in the range of [−6, +6], so NJ_1 has 169 features. Similarly the inter-block neighboring joint density matrix on horizontal direction NJ_{2h} and the matrix on vertical direction NJ_{2v} are constructed as follows:

$$NJ_{2h}(x,y) = \frac{\sum_{m=1}^{8}\sum_{n=1}^{8}\sum_{i=1}^{M}\sum_{j=1}^{N-1}\delta(C_{ijmn} = x, C_{i(j+1)mn} = y)}{64M(N-1)}$$

$$NJ_{2v}(x,y) = \frac{\sum_{m=1}^{8}\sum_{n=1}^{8}\sum_{i=1}^{M-1}\sum_{j=1}^{N}\delta(C_{ijmn} = x, C_{(i+1)jmn} = y)}{64(M-1)N}$$

We define NJ_2 as the neighboring joint density features on inter-block, calculated as follows:

$$NJ_2(x,y) = \frac{NJ_{2h}(x,y) + NJ_{2v}(x,y)}{2}$$

Similarly, the values of *x* and *y* are in the range of [−6, +6] and NJ_2 has 169 features

Hence we extract 169 features from both neighboring joint density of intra and inter block. So totally 338 features are extracted from neighboring joint density DCT array.

4. FEATURE BASED JPEG STEGANALYSIS USING NEIGHBORING JOINT DENSITY BASED FEATURES

By combining the features obtained from the neighboring joint densities, a new feature based JPEG steganalysis scheme is proposed. From the neighboring joint density of intra block 169 features and from neighboring joint density of inter block another 169 features are extracted and totally 338 distinguishable statistics are extracted for better steganalysis.

After the features are extracted from both stego and clear images it will be given to SVM like binary classifier for training. After the training is completed the features from test images are given for classification.

5. EXPERIMENTAL RESULTS

Five hundred and eighty five natural images were collected and these color images span a range of indoor and outdoor scenes and typically are 256 x 256 pixels in size. Another five hundred and eighty five stego images were generated by embedding messages of various sizes into the cover images. The payload corresponding to 100%, 75%, 50%, 25%, 20% and 10% of total cover capacity. The total cover capacity is defined to be the maximum size of a message that can be embedded by the embedding algorithm. Messages were embedded using F5, Model Based

Steganography (MB1 and MB2), Pixel Value Differencing (PVD), JPHS and Steghide algorithms.

Individually each feature set is used for steganalysis and the combined one is also used. Neighboring joint density of intra block, Neighboring joint density of inter block and Combined Neighboring joint density of intra and inter block are extracted and used for steganalysis of all the above mentioned stego algorithms. Features will be extracted from each images yielding to a 169, 169 and 338 feature vector respectively. These features are used to train the linear SVM classifier separately. The performance of the classifier was tested using 250 test images which contain 25 cover and 25 stego images for F5, Model Based Steganography (MB1 and MB2 each), Pixel Value Differencing (PVD) respectively, 15 cover and 15 stego images for JPHS and 10 cover and 10 stego images for Steghide algorithm.

Table 1 shows the classification accuracy of the neighboring joint density of intra block feature based steganalysis, Table 2 shows the classification accuracy of the neighboring joint density of inter block feature based steganalysis and Table 3 shows the classification accuracy of the neighboring joint density of intra block and inter block feature based steganalysis. All the individual features and the combined features are used for steganalysis. The neighboring joint density of intra and inter block features will combined and used for feature based steaganalysis which will give better result when compared to other feature based steganalysis. Different payload can be embedded and used for steganalysis. While for lower payload also this feature based steganalysis gives better results than other features. Strong steganographic algorithms like steghide and JPHS will also gives better result in these features than other.

Table 1. Classification accuracy of Neighboring joint density of intra block features

Algorithms	**Payload**	**Classification Accuracy (%)**
PVD	25-50-75-100	100
F5	25-50-75-100	100
MB1	25-50-75-100	100
MB2	25-50-75-100	88
JPHS	10-20	70
Steghide	10-20	95

Table 2. Classification accuracy of Neighboring joint density of inter block features

Algorithms	**Payload**	**Classification Accuracy (%)**
PVD	25-50-75-100	96
F5	25-50-75-100	94
MB1	25-50-75-100	100
MB2	25-50-75-100	96
JPHS	10-20	63.33
Steghide	10-20	90

Table 3. Classification accuracy of Combined neighboring joint density of intra and inter block features

Algorithms	Payload	Classification Accuracy (%)
PVD	25-50-75-100	98
F5	25-50-75-100	100
MB1	25-50-75-100	100
MB2	25-50-75-100	98
JPHS	10-20	73.33
Steghide	10-20	100

6. CONCLUSIONS AND FUTURE WORKS

From the above experiments we concluded that the combination of all neighboring joint density features used steganalysis will gives better result when compared with other features. For strong steganographic algorithms like steghide which uses graph theoretical approach for embedding this feature based steganalysis performs better detection. The results of this paper demonstrate that, with judicious and sophisticated feature mining, it is possible to simultaneously achieve faster detection time, and higher detection performance for JPEG image steganography.

The future work is to do the feature selection by ranking the feature vector using some ranking algorithms and the optimum features has to be discovered out. These optimum features can reduce the miss classification. Feature selection can also be applied using projection pursuit algorithms to improve the detection efficiency. More embedding schemes can be used to analyse the features efficiency.

REFERENCES

[1] Gireesh Kumar T, Jithin R, Deepa D Shankar, (2010) " Feature Based Steganalysis using Wavelet Decomposition and Magnitude Statistics", *International Conference on Advances in Computer Engineering*.

[2] Deepa.D.Shankar, Gireeshkumar T, (2010) "Feature Based Classification System for Image Steganalysis", *International Conference on Computer Communications and Networks(CCN-10)*.

[3] J. Fridrich, (2004) " Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes", *Information Hiding, 6th International Workshop,LNCS 3200* ,pages 67-81.

[4] D.Fu, Y.Q.Shi and D.Zou, (2006) "JPEG steganalysis using empirical transition matrix in block DCT domain", *International Workshop on Multimedia Signal Processing* , Victoria, BC, Canada.

[5] T.Pevny and J.Fridrich, (2007) " Merging Markov and DCT features for multi-class JPEG steganalysis", *Proceedings of SPIE Electronic Imaging, Security, Steganography and Watermarking of Multimedia Contents IX*, volume 6505, pages 650503-1 to 650503-13

[6] Qingzhong Liu, Andrew H Sung, Mengyu Qiao, (2011) " Neighboring Joint Density-Based JPEG Steganalysis", *ACM Transactions on Intelligent Systems and Technology.* volume 2, No 2, Article 16.

Authors

Arun R. received B.Tech degree in Computer Science and Engineering from St Josephs College of Engineering and Technology Palai, Mahatma Gandhi University, Kerala in 2009. He is pursuing his M.Tech in Cyber Security at Amrita Vishwa Vidhyapeetham University, Coimbatore. His research interests are Steganography, Cryptography and Image Processing

Nithin Ravi S. received B.Tech degree in Computer Science and Engineering from Government Engineering College Wayanad, Kannur University, Kerala in 2009. He is pursuing his M.Tech in Cyber Security at Amrita Vishwa Vidhyapeetham University, Coimbatore. His research interests are Steganography and Image Processing

Thiruppathi K. received B.Tech degree in Computer Science and Engineering from Velammal Engineering College Chennai, Anna University, Tmil Nadu in 2010. He is pursuing his M.Tech in Cyber Security at Amrita Vishwa Vidhyapeetham University, Coimbatore. His research interests are Steganography and Image Processing