Feature Based Steganalysis Using Wavelet Decomposition and Magnitude Statistics

Gireesh Kumar T
TIFAC CORE in Cyber Security
Amrita Vishwa Vidyapeetham
Coimbatore, India
gireeshkumart@gmail.com

Jithin R , Deepa D Shankar
TIFAC CORE in Cyber Security
Amrita Vishwa Vidyapeetham
Coimbatore, India
{jithinr550, sudee99}@gmail.com

***Abstract*—Steganography is broadly used to embed information in high resolution images, since it can contain adequate information within the small portion of cover image. Steganalysis is the procedure of finding the occurrence of hidden message in an image. This paper compares the efficiency of two embedding algorithms using the image features that are consistent over a wide range of cover images, but are distributed by the presence of embedded data. Image features were extracted after wavelet decomposition of the given image. These features were then given to a SVM classifier to identify the stego content.**

Keywords-Steganography; Steganalysis; SVM; Wavelet decomposition

I. INTRODUCTION

Steganography is the procedure of hiding information in a cover medium. Cover medium can be an image, audio or a video. The main aim of Steganography is to hide the data in a cover medium so as to transfer the data secretly through a public channel. Different steganographic scenarios can be distinguished based on what information is available to the steganalyst.

The goal of Steganalysis is to recognize whether the medium contains any hidden information. The Steganalysis is successful if it can find whether an image contains a hidden message or not with a probability higher than random guessing. The different types [4] of Steganalysis are:

1) Supervised learning based Steganalysis [2] [3]: It has two phases (a) In the training phase, the stego image features along with image type (stego or not) will be given to a statistical classifier. The classifier "learns" the best classification rule using these examples. (b) In the testing phase, unknown images are given as input to the trained classifier to decide whether a secret message is present or not.

2) Blind identification based Steganalysis [7]: Some statistical properties such as the independence of host and secret message are exploited. The embedding algorithm is represented as a channel and the goal is to invert this channel to identify the hidden message.

3) Parametric statistical Steganalysis [5]: This approach is formulated as a hypothesis testing problem, namely, null hypothesis (no message) and alternate hypothesis (message present). A statistical detection algorithm is then designed to test between the two hypotheses.

4) Hybrid techniques: Hybrid techniques overlap more than one of the above approaches.

Steganalysis also attempts to discover more information of the image and hidden message such as the type of embedding algorithm, the length of the message, the content of the message or the secret key used. A less theoretical and more practical categorization of Steganalysis is of the following

1) Targeted Steganalysis: In the case of a known algorithm, an attack that works for that specific algorithm is called Targeted Steganalysis.

2) Blind Steganalysis: Steganalysis attacks that can be appropriate on all steganographic algorithms are called blind Steganalysis.

3) Semi Blind Steganalysis: Steganalysis attacks that can apply on a selected set of steganographic algorithms are called semi-blind attacks.

The rest of this paper is organized as follows: In Section 2, some existing Steganalysis methods are explained. In Section 3, the proposed Steganalysis method is explained in detail. Section 4 includes the implementation and results. The final conclusions are drawn in Section 5.

II. EXISTING STEGANALYSIS METHODS

Visual attack is a type of Targeted Steganalysis method. The idea of visual attack is to eliminate any parts of the image that is distinguishable to the human eye. Fig. 2.1 shows the LSB plane of the cover image and its corresponding stego image. It clearly distinguishes between the two.

Histogram analysis attack works on the stego systems which are embedded sequentially or pseudo-random type in frequency domain. This is a Semi-blind Steganalysis method. It can efficiently estimate the length of the message embedded and it is based on the loss of histogram symmetry after embedding. This attack works on Outguess 0.1. The following example shows the histogram of a cover image (Fig. 2.2) and the histogram of corresponding stego image (Fig. 2.3).

Blockiness defines the sum of spatial discontinuities along the boundary of all 8x8 JPEG blocks. Blockiness calculates the difference between the pixel values at the

978-0-7695-4058-0/10 $26.00 © 2010 IEEE
DOI 10.1109/ACE.2010.33

boundaries of each JPEG block. The differences of the pixel values are calculated for both column and row boundaries and the sum of those gives our Blockiness value. The Blockiness value increases after the message is embedded into it. Blockiness value is directly proportional to the length of embedded message.

In a cover image, the Blockiness value increases quickly with respect to the length of the hidden data. But in a stego image, the Blockiness value increases gradually with respect to the message length. The change in the Blockiness value is inversely proportional to the length of the data embedded in the image.

In feature based Steganalysis, the first set of features are extracted from the image. Calibrated image will be formed using calibration technique. The calibrated image is perceptually similar to the original cover and hence the features would be similar to those of the original. A second set of features is extracted for this calibrated image. The two set of features will be extracted from a large database of images. The extracted features will be used to train a classifier (e.g.: Fisher Linear Discriminant).

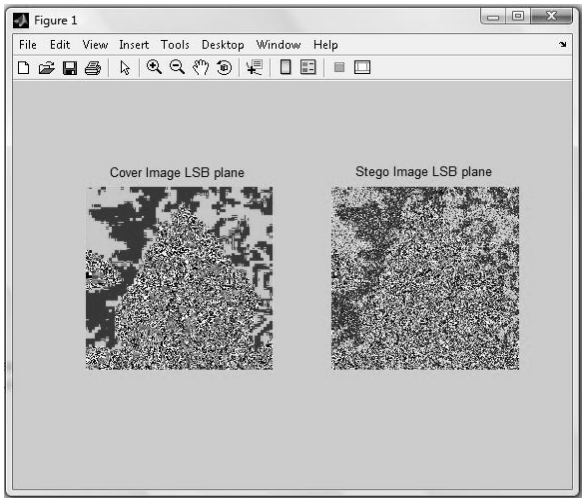

Figure 2.1 Visual Attack

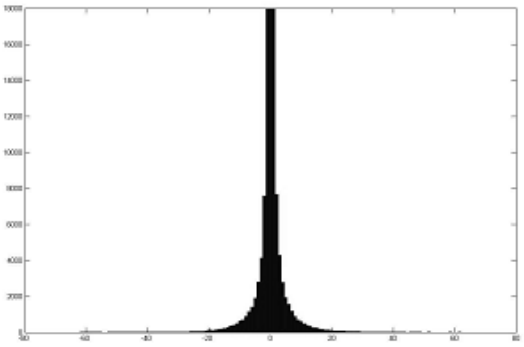

Figure 2.2 Histogram of the Cover Image

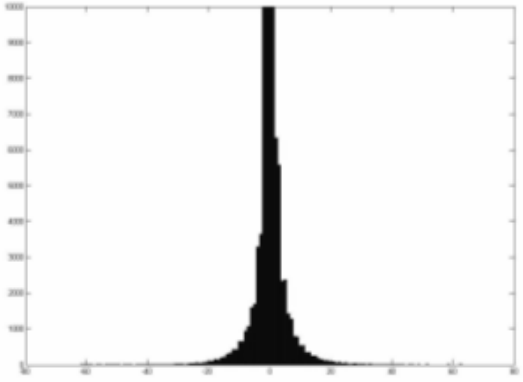

Figure 2.3 Histogram of the Stego Image

III. PROPOSED SYSTEM

The proposed system is a Universal Steganalysis method which uses a blind classifier. The classifier is based on a large set of feature vectors derived from an image. The proposed system architecture is shown in Fig. 3.1.

A. Transformation

Describe the image transformation that localize image structure in both space and frequency domain. This can be done using Wavelet decomposition [6]. The image decomposition employed here is based on separable Quadrature Mirror Filters. The separable QMFs consist of a pair of one-dimensional low-pass, l(.), and high-pass, h(.), filters.

The decomposition consisting of a vertical (1), horizontal (2), diagonal (3) and low-pass sub band (4) is generated by convolving each color channel with the low-pass and high-pass filters.

$$V^c_i(x,y) = I^c(x,y) * h(x) * l(y) \quad (1)$$

$$H^c_i(x,y) = I^c(x,y) * l(x) * h(y) \quad (2)$$

$$D^c_i(x,y) = I^c(x,y) * h(x) * h(y) \quad (3)$$

$$L^c_i(x,y) = I^c(x,y) * l(x) * l(y) \quad (4)$$

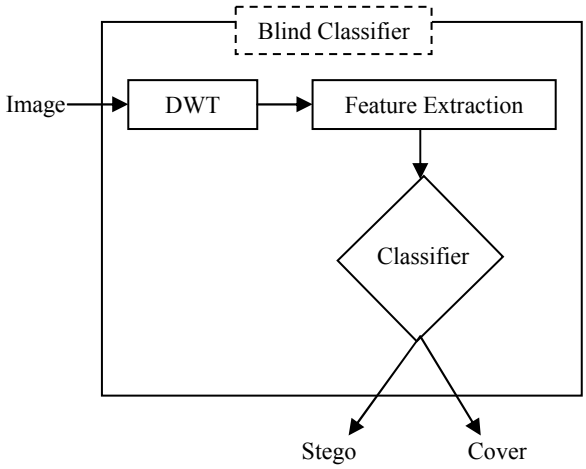

Figure 3.1 Proposed System Architecture

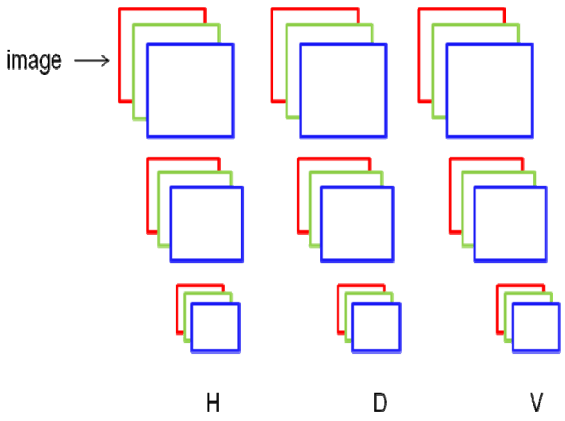

Figure 3.2: Three level wavelet decomposition of a RGB image

The entire process is repeated on the low-pass sub band to create as many scales as needed. Three levels of Wavelet decomposition have been used here.

B. Feature Extraction

Each image will be having red, green and blue channel. For each channel, three levels of wavelet decomposition have been done. A total of 27 sub bands were obtained as shown in Fig. 3.2. For each sub band the magnitude statistics is extracted, which consist of mean, variance, skewness and kurtosis. Thus a total of 108 features were obtained.

C. Linear Support Vector Machine (SVM)

There are many classifiers like neural network, Fisher Linear Discriminant, SVM etc. Out of these, SVM is considered to be more powerful, hence used here. Since the Steganalysis system was blind, the SVM has to be trained before testing.

IV. Implementation And Results

Firstly, 15 cover images were taken and corresponding stego images are created using F5 steganographic algorithm. The magnitude statistics of these 30 images were extracted .These features are used to train the linear SVM. The performance of the classifier was tested using 10 test images which contain 5 cover and 5 F5 stego images. The performance of the classifier was 80%

Secondly, 15 cover images were taken and created their corresponding stego images using Jpeg Hide and Seek (JPHS) steganographic algorithm. The magnitude statistics of these 30 images are extracted and used to train the SVM classifier. The performance of the classifier was tested using 10 test images containing 5 cover and 5 JPHS stego images. The performance of the classifier was only 50%. It is very less compared to F5 stego images.

Next, the features of 15 cover images, 15 F5 stego images and 15 JPHS stego images were combined. These 45 image features were given to a classifier for training.

TABLE I. Experimental Results

Images used for training	Payload Size (%)	Performance of classifier (%)
F5 embedding images + cover images	20%	80%
JPHS embedding images + cover images	20%	50%
Combined F5 and JPHS images + cover images	20%	66%

Then the classifier was tested using the test images, which include 5 cover images and their corresponding stego images created using F5 and JPHS steganographic scheme. The performance after combining the features was only 66%. The test results are shown in Table 1.

V. Conclusion And Future Work

From the experimental results, we concluded that JPHS embedding is more efficient than F5 embedding scheme. Also from the observation of feature values extracted for cover images, its corresponding F5 stego images and JPHS stego images, the difference between cover image features and F5 stego image features are very large compared to JPHS. The feature values were almost equal for cover and JPHS stego images. So it can be easily concluded that the JPHS is more efficient than F5 embedding scheme.

The higher order features like rotational invariant features and noise features values [1] are consistent for cover images and are distributed by the presence of stego content. These features can improve the performance of the classifier. Feature selection can be applied using projection pursuit algorithms to improve the detection efficiency. More embedding schemes can be used to analyse the features efficiency and hence a comparative study of each.

References

[1] Hongmei Gou, Ashwin Swaminathan and Min Wu, "Noise Features for Image Tampering Detection and Steganalysis," University of Maryland, 2006.

[2] I. Avcibas, N. Memon, and B. Sankur, "Steganalysis Using Image Quality Metrics," *IEEE Trans*. on Image Processing, vol. 12, no. 2, pp. 221–229, Feb. 2003.

[3] Jessica Fridrich, "Feature-Based Steganalysis for Jpeg Images and its Implications for Future Design of Steganographic Schemes," SUNY Binghamton, USA, 2005.

[4] R. Chandramouli and K.P. Subbalakshmi, "Current Trends in Steganalysis: A Critical Survey," Stevens Institute of Technology, 2005.

[5] R. Chandramouli, "A Mathematical Framework for Active Steganalysis," ACM Multimedia Systems, vol. 9, no. 3, pp. 303–311, September 2003.

[6] Siwei Lyu and Hany Farid, "Steganalysis Using Higher-Order Image Statistics," IEEE Trans. on Info. Forensics and Security, vol. 1, no. 1, pp. 111–119, Mar 2006.

[7] Taras Holotyak, Jessica Fridrich, Sviatoslav Voloshynovskiy, "Blind Statistical Steganalysis of Additive Steganography Using Wavelet Higher Order Statistics," USA, 2005.