Blind Steganalysis : To Analyse the Detection Rate of Stego Images using Different Steganalytic Techniques with Support Vector Machine Classifier

B.Yamini M.Tech.,
Research Scholar
Sathyabama University
Old mamallapuram road
Chennai -103
TamilNadu,India.

R.Sabitha, Ph.D
Professor and Head
Department of IT
Jepiaar Engineering College
Chennai-119
TamilNadu,India.

ABSTRACT
Steganography is the art of hiding the secret message in a cover medium and the media could be a audio or video or image. Steganography aims at hiding data as stealthy as possible in a cover medium, Steganalysis aims to detect the presence of any hidden information in the stego media here it refers to the JPEG images. Current Steganalysis aims to focus more on detecting statistical anomalies in the stego images which are based on the features extracted from typical cover images without any modifications.Most steganalysis algorithms are based on exploiting the strong interpixel dependencies which are typical of natural images. Steganalysis can be classified into two broad categories: Specific/Targeted and Blind Steganalysis. Blind steganalysis also known as universal steganalysis.Steganalysis is the modern and powerful approach to attack a stego media since this method does not depend on knowing any particular embedding technique. A pattern recognition classifier is then used to differentiate between a cover images and a stego image.A number of algorithms were used to analyse the cover media.In proposed work blind steganalysis is done using J2 technique, inorder to analyse the performance of J2 with respect to capacity and stealthiness and also to compare the detection rate of J2 with other popular algorithms using first and second order steganalysis using Support Vector Machine.

Keywords
Steganalysis, Cover images, Stego Images, Support Vector Machine, Blind Steganalysis, Targeted Steganalysis & Adaptive Steganography.

1. INTRODUCTION
Steganalytic attack is the technique used to discover hidden messages in cover objects. Steganalysis [1] could be a blind steganalysis or a targeted steganalysis. Targeted steganalysis aims to break a particular known steganographic embedding scheme, and blind steganalysis, which attempts to identify the existence of steganography without a prior knowledge of the specific algorithm that was used. Images can be scanned for suspicious properties in a very basic way; detecting hidden messages usually require a more technical approach. Changes in size, file format, last modified timestamp and in the color palette might point out the existence of a hidden message, but this will not always be the case. A widely used technique for image scanning involves statistical analysis.

2. SUPPORT VECTOR MACHINE
Support Vector machine (SVM) [2] is a newer technique for data classification and regression. The Classification problem can be restricted to consideration of the two class problem without loss of generality. SVMs introduced in COLT-92 by Boser, Guyon & Vapnik.SVM are theoretically well motivated algorithms developed from Statistical Learning Theory (Vapnik & Chervonenkis).It has empirically good performance in many fields such as bioinformatics, text, image recognition, etc.SVM technique can be used in many types of pattern recognition in face detection [3], face authentication[4,5,6],face recognition [4], object detection [7],text classification [8],image classification [9] and voice identification [9].

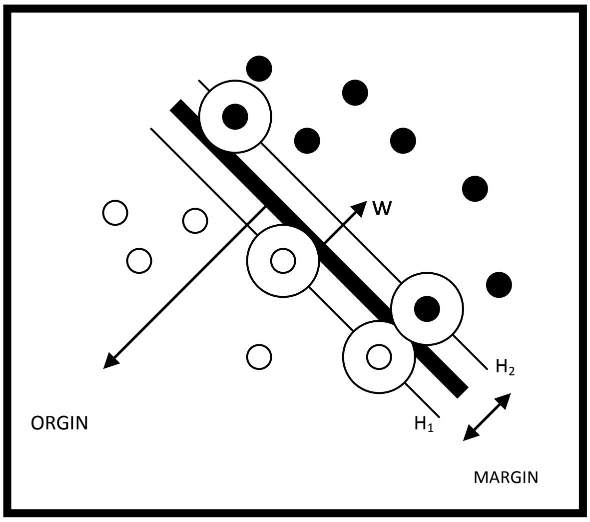

Figure 1. Linear Separating hyper planes for the separable case.

3. PROPOSED WORK
Fig.2 describes the method in detail. The method shows only one coefficient change per block for simplicity. The actual J2 can change more than one coefficient if the current block is not able to produce the desired datum on the spatial domain. To compare the detection rate of J2 with other popular algorithms using first and second order steganalysis using Support Vector Machine.

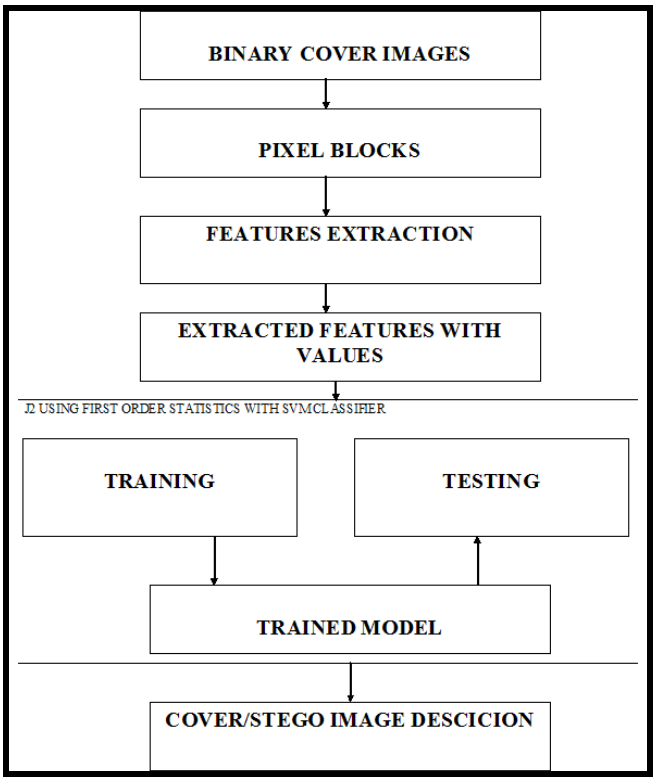

Fig. 2, Block Diagram for J2 with other popular algorithms using first order steganalysis with Support Vector Machine classifier.

4. FIRST ORDER STATISTICS

First order statistics look at single pixels in the image and do not take into account relations between neighbours. They correspond to the distribution of values within the analysed data set and its properties, making them easy to both compute and interpret [van der Schaaf 1998; Huang and Mumford 1999]. Several simple regularities have been established regarding the statistical distribution of pixel values; the gray-world assumption is often quoted in this respect. Averaging all pixel values in a photograph tends to yield a color that is closely related to the blackbody temperature of the dominant illuminant.

The integrated color tends to cluster around a blackbody temperature of 4800 K. For outdoor scenes, the average image color remains relatively constant from 30 minutes after dawn until within 30 minutes of dusk.

The correlated color temperature then increases, giving an overall bluer color.1 The average image color tends to be distributed around the color of the dominant scene illuminant according to a Gaussian distribution, if measured in log space . The eigenvectors of the covariance matrix of this distribution follow the spectral responsivities of the capture device, which for a typical color negative film are :

$L = 1/\sqrt{3}(\log(R) + \log(G) + \log(B))$,
$s = 1/\sqrt{2}(\log(R) - \log(B))$,
$t = 1/\sqrt{6}(\log(R) - 2\log(G) + \log(B))$.

For a large image database, the standard deviations in these three channels are found to be 0.273, 0.065, and 0.030. Note that this space is similar to the L color space. Although that space is derived from LMS rather than RGB, both operate in log space, and both have the same channel weights. The coefficients applied to each of the channels are identical as well, with the only difference being that the role of the green and blue channels has been swapped.

Place Tables/Figures/Images in text as close to the reference as possible (see Figure 1). It may extend across both columns to a maximum width of 17.78 cm (7").

Captions should be Times New Roman 9-point bold. They should be numbered (e.g., "Table 1" or "Figure 2"), please note that the word for Table and Figure are spelled out. Figure's captions should be centered beneath the image or picture, and Table captions should be centered above the table body.

5. SECOND ORDER STATISTICS

Second order statistics capture relations and regularities over pairs of pixels in the image. The remarkable and salient natural image statistic that has been discovered so far is that the slope of the power spectrum tends to be close to 2. The power spectrum of an M × M image is computed as

$$S(u,v) = \frac{|F(u,v)|^2}{M^2},$$

where F is the Fourier transform of the image. By representing the two-dimensional frequencies u and v in polar coordinates (u = f cos and v = f sin _) and averaging over all directions and all images in the image ensemble, it is found that on log-log scale amplitude as a function of frequency, f lies approximately on a straight line . This means that spectral power as a function of spatial frequency behaves according to a power law function. Moreover, fitting a line through the data points yields a slope of approximately 2 for natural images. Although this spectral slope varies subtly between different studies and with the exception of low frequencies , it appears to be extremely robust against distortions and transformations.

It is therefore concluded that this spectral behavior is a consequence of the images themselves, rather than of particular methods of camera calibration or exact computation of the spectral slope. We denote this behavior by

$$S(f) \propto \frac{A}{f^{\alpha}} = \frac{A}{f^{2-\eta}},$$

where A is a constant determining overall image contrast, is the spectral slope, and is its deviation from 2. However, the exact value of the spectral slope depends somewhat on the type of scenes that make up the ensembles. Most interest of the natural image statistics community is in scenes containing mostly trees and shrubs. Some studies show that the spectral slope for scenes containing man-made objects is slightly different.

Even in natural scenes, the statistics vary, dependent on what is predominantly in the images. The second-order statistics for sky are, for example, very different from those of trees. One of the ways in which this becomes apparent is when the power spectra are not circularly averaged, but when the log average power is plotted against angle.

The power spectrum is related to the auto-correlation function through the Wiener-Khintchine theorem, which states that the auto-correlation function and the power spectrum form a Fourier transform pair . Hence, power spectral behavior can be equivalently understood in terms of correlations between pairs of pixel intensities. A related image statistic is contrast, normally defined as the standard deviation of all pixel intensities divided by the mean intensity. This measure can either be computed directly from the image data, or it can be derived from the power spectrum through Parceval's theorem

$$\frac{\sigma^2}{\mu^2} = \sum_{(u,v)} S(u,v).$$

This particular contrast computation can be modified to compute contrast in different frequency bands. Frequency-conscious variance can then be thresholded, yielding a measure which can detect blur . This is useful as lack of contrast can also be caused by the absence of sufficient detail in a sharp image.

6. RESULTS

This table1 shows the result of the detection rate of J2 with other popular algorithms using first order steganalysis with Support Vector Machine classifier.

YASS	F5	Outguess	Jsteg	MB1	J2
0.04	0.2	0.6	0.8	0.2	0.82
0.09	0.5	0.9	0.9	0.39	0.93
0.18	0.9	0.99	0.9	0.7	0.97
0.39	1	0.99	0.5	0.82	0.98
0.55	1	1	0.32	0.89	1
0.79	1	1	0.5	0.9	1

Table 1

Fig 1 is the Graphical Representation of J2 with other popular algorithms using first order steganalysis with Support Vector Machine classifier.

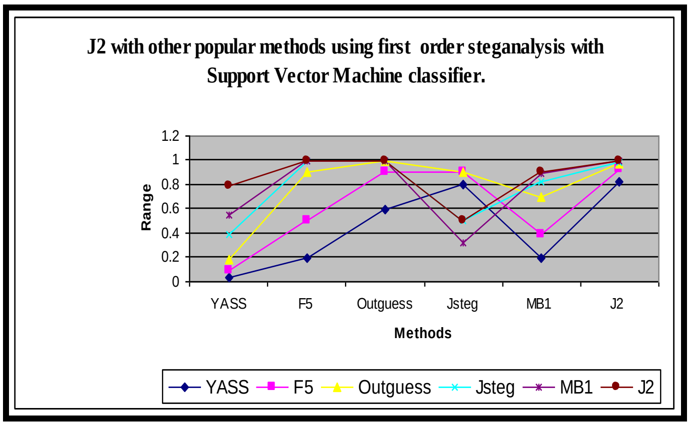

Fig. 2

This table2 shows the result of the detection rate of J2 with other popular algorithms using Second order steganalysis with Support Vector Machine.

YASS	F5	Outguess	Jsteg	MB1	J2
0.09	0.3	0.63	0.83	0.27	0.89
0.12	0.56	0.92	0.94	0.45	0.98
0.19	0.92	1	0.93	0.79	0.9
0.45	1.2	1.2	0.54	0.88	0.86
0.59	1.2	1.3	0.52	0.9	1
0.8	1.5	1.4	0.3	0.92	1

Table 2

Fig 2 is the Graphical Representation of J2 with other popular algorithms using first order steganalysis with Support Vector Machine classifier.

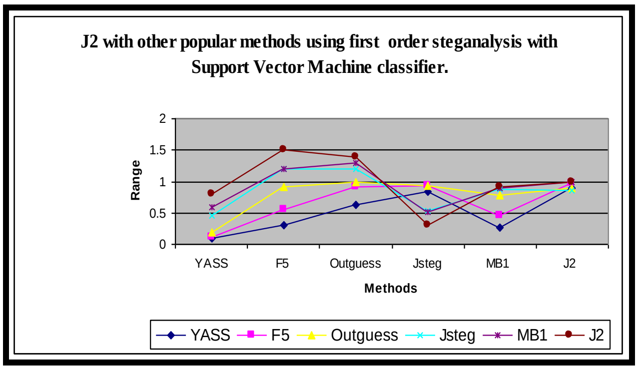

Fig.3

7. CONCLUSION

The experimental results were carefully evaluated and interpreted. Conclusions concerning current and future steganographic schemes for JPEGs were also drawn. The Detection rate of second order statistics using Support Vector Machine performs well than that of first order statistics. In future, experiments can be made on higher order statistics.

8. REFERENCES

[1]Mahdi Ramezani and Shahrokh Ghaemmaghami,“Adaptive Image Steganography with Mod-4 Embedding using Image Contrast”, Intl. Conf. on IEEE CCNC 2010 proceedings.

[2] M. Raja Sekar, Prof. P Prem Chand & Prof. I .V.Muralikrishna,“Image Authentication using Support Vector Machines”, Journals of Advanced Research in Computer Engineering”, vol.3, No.2, July-December 2009, pp. 367-374.

[3]Chen, X., Faltemier, T., Flynn, P., and Bowyer,K.,“Human face modeling and face recognition through multi-view high resolution stereopsis”,In Conference on computer vision
and pattern Recognition (2006) pp.50-56.

[4] Bicego, M., Grosso, E., and Tistarelli, M.,“ Face Authentication using One-Class Support Vector Machines”, In International Workshop on Biometric Recognition Systems(2005) 1522.

[5] Chen, Y., Zhou, X.S., and Huang, T.S., " One- Class SVM for Learning in Image Retrival", In IEEE International Conference on Image Processing, (2001) 34-37.

[6] Carson, C., and Greenspan, H., "Blobworld Image Segmentation using Expectation Maximization and its Application to Image Querying. IEEE Transaction on Pattern Analysis and Machine Intelligence, 24(8) (2002) 1026-1037.

[7] Burges, C.J., " A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery, 2(1998) 121-167.

[8] Ben-hur, A., Horn, D., Siegelmann, H., and Vapnik,V.,"Support Vector Clustering",Journal of Machine Learning Research, 2 (2001)125-137.

[9] Chapelle, O., and Haffner, p., and Vapnik, V.N,"Support Vector Machines For Histogram based image classification", IEEE transaction on Neural Networks, 10(5) (1999) 1055-1064.

[10] Burges, C.J., " A Tutorial on Support Vector Machines for Pattern Recognition", Bell Laboratories, 2(1998) 1-43.

[11] Kang Leng Chiew , Josef Pieprzyk, " Estimating Hidden Message Length in Binary Image Embedded By Using Boundary Pixels Steganography.Intl. Conf. On Availability Reliability and Security,(2010) 683-688.