©2010 International Journal of Computer Applications (0975 - 8887)
Volume 1 – No. 19

Universal Steganalysis using Feature Selection Strategy for Higher Order Image Statistics

Sonali S.Ekhande
Walchand College of Engg
Sangli

Prof. S.P.Sonavane
Walchand College of Engg.
Sangli.

Dr. P .J .Kulkarni
Walchand College of Engg.
Sangli

ABSTRACT
The purpose of image steganalysis is to detect the presence of hidden message in cover photographic images. Supervised learning is an effective and commonly used method to cope with difficulties of unknown image statistics and unknown steganography. Present paper proposes; a universal approach for steganalysis for detecting presence of hidden messages embedded within digital images. This paper describes wavelet like decomposition to build higher order statistical model of natural images. Feature selection technique like ANOVA is used to select relevant features. SVM are then used to discriminate between clean and stego images. Study of the effect of relevant features on classification accuracy may help to improve the complexity.

Categories and Subject Descriptors
Information Security: Features – secret communication, data hiding, steganography, universal steganalysis, supervised learning, Feature Selection..

General Terms
Security, Verification.

Keywords:
Information Hiding, Steganography, Steganalysis, Image statistics, Support Vector Machine (SVM), Feature selection, ANOVA

1. INTRODUCTION
Information hiding has been a hot research area in recent years. Its application lies in Military and Intelligence agencies, Law enforcement and counter intelligence. Early research has been focused on steganography to establish secret channel between two parties. The goal of steganography is to embed within an innocuous looking cover medium (image, audio, video, etc.) a message so that casual inspection of the resulting medium will not reveal the presence of message [1-2]. Today steganography is an active research area due to abundance of digital media serving as cover signals, and availability of public communication network as internet. By secretly embedding information into innocent cover signal, transmitter hopes that message will reach the receiver without suspicion.

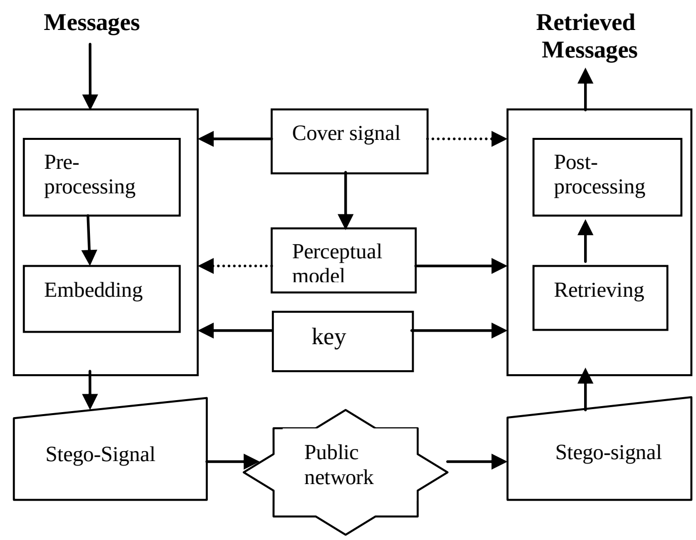

Figure 1: Illustration of information hiding process

Fig. 1 illustrate the flow of most information hiding [3] processes. If sender wants to transmit a message to a receiver through a public channel, a chosen material called cover-signal, is modified to create the stego signal by embedding desired messages. In practical application, messages are usually hidden in random position after being ciphered by using an encryption algorithm (with secret key) to achieve the goals of statistical undetectability and authorized access. After the stego signal is received, the hidden message can be extracted by using the same key with/without knowledge of the original cover signal.

Steganalysis aims at detecting presence of hidden information from stegosignals. Steganography introduces different artifact into images and leaves specific fingerprints. Hence steganographer faces difficult challenge of preserving statistics of all image features after data embedding and the steganalyzer faces the opposite problem of finding same features whose statistics are changed by data embedding.

Current steganalysis method fall broadly into two categories: embedding specific [4] or universal [5].Universal steganalysis attempts to detect the presence of an embedded message independent of embedding algorithm and ideally image format. With ever growing number of steganography tools, universal approaches are necessary to perform generic steganalysis. Aim of this paper [6] is to describe scheme for detailed analysis of steganographic procedure i.e. a universal steganalysis method for detecting presence of hidden messages. The approach taken here relies on building higher-order statistical models for natural images and looking for deviations from these models. Across a large number of natural images, there exist strong higher-order statistical regularities within a wavelet-like decomposition. The embedding of a message significantly alters these statistics and thus becomes detectable. Support vector machines are employed to detect these statistical deviations. Feature selection strategy is then used to select relevant features.

2. IMAGE STATISTICS

The decomposition of images using basis functions that are localized in spatial position, orientation and scale (e.g., wavelet) have proven extremely useful in image compression, image coding, noise removal and texture synthesis. One reason is that such decompositions exhibit statistical regularities that can be exploited.

2.1 Image Decomposition

The multi-scale image decomposition employed in this work is the wavelet decomposition, based on Daubechies wavelet family(db4). An advantage of wavelet transforms is that the windows vary. In order to isolate signal discontinuities one would like to have some very short basis functions. At the same time, in order to obtain detailed frequency analysis, one would like to have some very long basis functions. A way to achieve this is to have short high-frequency basis functions and long low-frequency ones. This is exactly what we get with wavelet transforms. Some of the wavelet bases have fractal structure. The Daubechies wavelet family is one example.

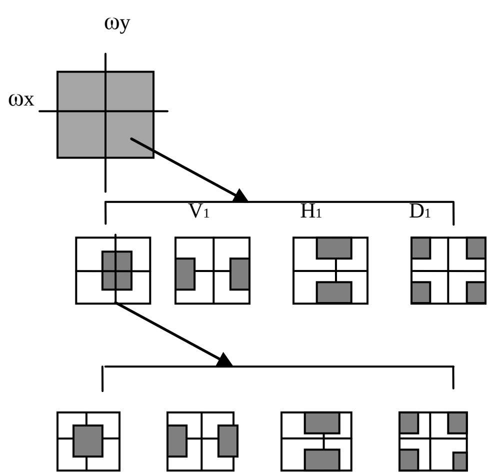

Figure 2. Decomposition of frequency space

2.2 Magnitude Statistics

Starting with the simpler case of grayscale images, it is known that sub band coefficients of a multi-scale image decomposition of a natural image have specific distributions which are characterized by a sharp peak at zero and long symmetric tails. An intuitive explanation is that natural images contain large smooth regions and abrupt transitions (e.g., edges). The smooth regions, though dominant, produce small coefficients near zero, while the transitions generate large coefficients. This property holds for the multiscale image decompositions (e.g., wavelets).

Given this image decomposition, the statistical model is composed of the mean, variance, skewness and kurtosis of the sub band coefficients at each orientation, scale and color channel.
The first four cumulants (i.e., the mean, variance, skewness and kurtosis) of the coefficients in each sub band of all orientations, scales and color channels are obtained to characterize the marginal distributions of the coefficients. The cumulants determine the distribution indirectly; distributions sharing similar cumulants will have similar shapes.

$$\mu_x \stackrel{def}{=} E_{p_x(x)}\{x\}$$
$$\sigma_x^2 \stackrel{\text{def}}{=} E_{p_x(x)}\left\{\left(x-\mu_x\right)^2\right\}$$
$$s_x \stackrel{\text{def}}{=} E_{p_x(x)}\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^3\right\}$$
$$k_x \stackrel{\text{def}}{=} E_{p_x(x)}\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^4\right\}$$

While these statistics characterize the basic coefficient distributions, they are unlikely to capture the strong correlations that exist across space, orientation, scale and color. For example, edges tend to extend spatially and across multiple scales. As such, if a large coefficient is found in a horizontal sub band, then it is likely that its left and right spatial neighbors in the same sub band will also have a large value. Similarly, a large coefficient at scale i might indicate a large value for its parent at scales $i + 1$. Since wavelet coefficient possess strong intra and inter sub band dependencies, prediction error sub bands exploit these dependencies as follows [6].

Take a sub band coefficient $H_i(j,k)$ as an example , where, (j, k) denotes the spatial coordinates at scale i . The magnitude of $H_i(j, k)$ can be linearly predicted by those of its parent $H_{i+1}(j/2, k/2)$; neighbors $H_i(j+1, k)$, $H_i(j, k+1)$, $H_i(j-1, k)$ and $H_i(j, k-1)$; cousins $D_i(j,k)$ and $V_i(j, k)$; and aunts $D_{i+1}(j/2, k/2)$ and $V_{i+1}(j/2, k/2)$. If we denote the predicted magnitude as $p|Hi(j, k)|$.

Then, the logarithmic error $eH_i(j,k)$ is given by

$$e\,H_i(j,k) = \log(|Hi(j,k)|) - \log(p|H(j,k)|)$$

This defines an error sub band eHi that corresponds to H_i . One can similarly define the error sub bands eV_i and eD_i. at scales $i=1,2,3$. The prediction errors for a cover image and its stego image have different statistics, which are useful in steganalysis .It is from this error the additional statistics namely mean, variance, skewness and kurtosis are collected. For each orientation, scale and color sub band, a similar error metric and error statistics are computed.

For a decomposition with scales $i = 1,\ldots,n$, the total number of basic coefficient statistics is $36(n - 1)$ ($12(n - 1)$ per color channel) and the total number of error statistics is also $36(n - 1)$, yielding a total of $72(n - 1)$ statistics. These statistics form the first set of feature vector to be used to discriminate between clean and stego image.

3. CLASSIFICATION

From the measured statistics of a training set of clean and stego image, the goal is to determine whether a test image contains a hidden message. Support vector machine (SVM)

Classifier is employed [6, 8] for classification. Here are briefly described, in increasing complexity, three classes of SVMs.

In the binary classification, a linear SVM classifier seeks a hyper plane that separates training data of two classes with the largest classification margin, which provably has the best generalization ability.

The nonlinear classification, contrary to the arbitrarily complicated non-linear classification techniques such as the neural network, is achieved by first embedding training data into a higher (possibly infinite) dimensional space. A linear separation is then found in that space by the linear SVM algorithm and is mapped back to the original data space as a non-linear classification surface. Such a non-linear classification, though more flexible, inherits the stability and generalization ability of linear SVM, thus effectively reduces the chance of over-fitting the training data

An OC-SVM first projects these data into a higher, potentially infinite, dimensional space with the mapping: $\Phi : R_d \rightarrow F$. In this space, a bounding hyper sphere is computed that encompasses as much of the training data as possible, while minimizing its volume.

4. FEATURE SELECTION USING ANOVA

Support Vector Machine (SVM) is an effective classification method, but it does not directly obtain the feature importance. There are various feature selection strategies, some of them are “filters”: general feature selection methods independent of SVM. That is, these methods select important features first and then SVM is applied for classification. On the other hand, some are wrapper-type methods: modifications of SVM which choose important features as well as conduct training/testing [10].

ANOVA is used as a statistical tool to show whether data from several groups’ stego images could be accounted for by the hypothesized factor. The objective of single-factor ANOVA problem is to decide whether the means for more than two treatments are identical. In this experiment, single-factor ANOVA should be used to find out whether data from different steganographic groups have a common mean. The assumption of ANOVA is that test data are normally distributed

The ANOVA produces an F statistic, the ratio of the variance among the means to the variance within the samples, this ratio of variance is a comparison of the variance amongst the different groups to the variance amongst all the individuals within those groups. A higher ratio implies significant differences between the groups.

The degrees of freedom for the numerator is I-1, where I is the number of groups (means) The degrees of freedom for the denominator is N - I, where N is the total of all the sample sizes

5. RESULTS AND ANALYSIS

The cover images in the experiments are the 9000 JPEG natural images as described. From these 9000 natural images, 9000 stego images (1500 per embedding message type and per steganography tool) were generated by embedding random messages of various sizes into the full-resolution cover images. The messages consisted of central regions of randomly chosen images from the same image database corresponding to 100% and 50% 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 (steghide, jphide & seek, outguess).

The same statistical feature vector (144 features) as described above was computed from the central 256X256 region of each stego image. In all of the results presented below, 15,000 of the clean and stego images were used to train a SVM, and the remaining 3000 images were used in testing throughout, results from the testing stage are presented

The confusion matrix for above classification in percentage is as

Embedding Algorithm	**%of Embedd**	**Classified as (in percent)**			
		Clean	**steghide**	**jphide& seek**	**outguess**
Steghide	100	23.	**66.6**	4.8	5.6
JpHide&Seek	100	23.2	6.8	**65.6**	4.4
Outguess	100	34	7.6	9.2	**49.2**
Steghide	50	35.6	**46.8**	8.4	9.2
JpHide&Seek	50	32.8	10.4	**49.6**	7.2
Outguess	50	39.2	10	10.4	**40.4**

A good classification-based technique must have a high detection rate, and at the same time, a small false positive rate.

However outguess has high false alarm .Some of the JPEG-based steganographic embedding techniques (outguess) recompress the JPEG image before embedding the message in them, which may be the cause of false alarm (classifier misclassifying images because of the recompression artifact).

Similarly data set includes images with variety of qualities as well as sizes as opposed to constant quality and size. JPEG image quality factor affects the steganalyzers performance. Cover and stego images with high quality factors are less distinguishable than cover and stego image with lower quality [5]. As our dataset includes images with various quality factors, it may be reason for less detection accuracy

As discussed above computational complexity along with less detection rate is important issue to be solved. Hence it is necessary to find minimum features which would reduce complexity and may lead to improvement in detection accuracy

In order to identify the wavelet statistics effective in steganalysis, ANOVA is used to test which statistics are consistent and accurate against the effects of various steganography tools.

The ANOVA produces an F statistic, the ratio of the variance among the means to the variance within the samples, this ratio of variance is a comparison of the variance amongst the different groups to the variance amongst all the individuals within those groups. A higher ratio implies significant differences between the groups.

100 color images were downloaded and stego images were created separately using Steghide, Hide and seek, outguess algorithms (100% embedding). ANOVA is used to test the 144 feature values and 4 groups of each class (clean, steghide, jphide&seek, outguess)

The Result shows increase in accuracies of steghide (66.8% /45.6%) ,Hide seek 50%(51.2%) but little degradation in outguess

.However reducing the number of Statistics (90 out of original 144) significantly lowers the training complexity .

Table 1. Average detection accuracies for 100% and 50% embedding at 74 D, 90 D , 128 D and 144 D feature vector

All Feature Embd	74D	90D	128D	144D
100	55.1	56	58.5	60.06
50	41.7	44.4	45.2	45.8

6. DISCUSSION

Presented is an universal approach to steganalysis which detect the embedding algorithm independent of cover image format, compression ratio, embedding capacity and embedding algorithm used.

It relies on building a statistical model of first- and higher-order magnitude statistics extracted from multiscale, multiorientation image decompositions. It is shown that these statistics are relatively consistent across a broad range of images, but are disturbed by the presence of hidden messages.

Feature selection strategy implemented here gives relevant features to be used for training and thus reduces the training complexity.

A single feature may provide only small indication of presence of steganography while several feature may face conflict in diagnosis .Hence it is necessary to find minimum feature which would reduce complexity and may lead to improve detection accuracy . One way ANOVA is used to test which statistics are consistent and accurate against the effect of various steganographic tools. N dimensional features selections are based on Fscore selected at different significant levels. Selection of appropriate significant level may give optimum features which reduce complexity and increase accuracy.

However the project also has its own limitations.

As the message size becomes smaller, the chance of detection falls. The results are adversely affected by various JPEG image compression ratio used in training and testing dataset.

The future work could be done on following points

Minimizing the effect of compression ratio(recompression artifact) on detection accuracy by proper training of the classifier or by taking quality factor as separate feature[9].

Irrelevant image manipulations which may be the cause of false alarm, must be detected and then removed from the training set of the classifier, and in classification, rejected by the classifier.

Strategy to select the significant levels of Fscore which will help in reduction in computational complexity with increase in detection accuracies.

Higher order statistics are computationally costlier then first order hence representing higher order in terms of lower order must be studied.

7. REFERENCES

[1] D. Kahn, "The history of steganography", in Proc. Information HidingFirst International Workshop, Cambridge, U.K., 1996.

[2] R. Anderson and F. Petitcolas, "On the limits of steganography", *IEEE J. Sel. Areas Commun.*, vol. 16, no. 4, pp. 474–481, 1998.

[3] Wen-Nung Lie and Guo-Shiang Lin "A feature based classification technique for blind image steganalysis", IEEE Transaction on multimedia, vol.7 ,no.6 ,Dec 2005 .

[4] N. Johnson and S. Jajodia, "Steganalysis of images created using current steganography software", in *Lecture Notes in Computer Science*, vol 1525, 1998, pp. 273–289.

[5] S. Lyu and H. Farid, "Detecting hidden messages using higher-order statistics and support vector machines", presented at the 5th Int.Workshop on Information Hiding, Noordwijkerhout, The Netherlands, 2002.

[6] S. Lyu and H. Farid "Steganalysis using higher order image statistics", IEEE Transaction on information forensic and security ,vol 1,no.,1 March 2006.

[7] Ying Wang and Pierre Moulin "Optimized feature extraction for learning based image steganalysis" IEEE Transaction on information forensic and security ,vol 2,no 1, March 2007.

[8] C. Burges, "A tutorial on support vector machines for pattern recognition".

[9] J. Fridrich, "Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes", presented at the 6th International Workshop on Information Hiding, Toronto, ON, Canada,

[10] "Combining SVMs with Various Feature Selection Strategies" Yi-Wei Chen and Chih-Jen Lin Department of Computer Science, National Taiwan University, Taipei 106, Taiwan.