Digital Video Steganalysis Based on Statistical Features

Thu Thu Htet, Khin Than Mya
University of Computer Studies, Yangon
thuthuhtet@gmail.com,khinthanmya@gmail.com

Abstract

Steganalysis is the art and science of detecting a secret communication. Hiding a message will most likely leave detectable traces in the cover medium. The information hiding process changes the statistical properties of the cover, which is a steganalyst attempts to detect. The process of attempting to detect statistical traces is called statistical steganalysis. This paper presents an improved blind steganalysis technique to detect the presence of hidden messages. In order to identify and classify the two types of statistic texture feature are used. The first type features derive from the average co-occurrence matrices. The second type features is the grey level histogram. Support Vector Machine is considered a state-of-the-art classification algorithm. SVM classifier is utilized as the classifier. Experimental results show that this approach is very successful in detecting the information-hiding in MSU Stego Video steganograms.

Keywords- *steganalysis, histogram characteristic function, co-occurrence matrices, SVM classifier.*

1. Introduction

Steganography is the art of hiding information in ways that prevent the detection of hiding message whereas cryptographic techniques try to conceal the contents of a message. . The object of communication is the hidden message and the cover data are only the means of sending it. Secret information as well as cover data can be any multimedia data like text, image, audio, video etc. Text, image, audio, and video can be represented as digital data. The explosion of Internet applications leads people into the digital world, and communication via digital data becomes recurrent. However, new issues also arise and have been explored, such as data security in digital communications, copyright protection of digitized properties, and invisible communication via digital media.

There are two types of steganalysis: targeted and blind. Targeted steganalysis is designed to attack one particular embedding algorithm. Targeted steganalysis can produce more accurate results, but it normally fails if the embedding algorithm used is not the target. Blind steganalysis can be considered a universal technique for detecting different types of steganography. Because blind steganalysis can detect a wider class of steganographic techniques, it is generally less accurate; however, blind steganalysis can detect new steganographic techniques where there is no targeted steganalysis available. In other words, blind steganalysis is an irreplaceable detection tool if the embedding algorithm is unknown or secret.

Nowadays, many video information hiding methods have been proposed. For video stream is first offered in compressed format, steganography algorithms that are not applicable in compressed bit-stream would require complete or at least partial decompression, which is an unnecessary burden best avoided. For this reason, a large number of video information hiding algorithms was designed in the compressed domain [1, 2] recently; which can be used as video steganography method with slightly modification. While on the other hand, the number of video steganalysis algorithm is very small [6, 9] for the complexity of the hidden message detection in video sequence, especially the compressed video bit streams.

The art of discovering hidden data in cover

objects, steganalysis has gained attention in the fields of computer forensics to decrease the serious consequences of covert communications. In this paper propose a scheme for detecting the information-hiding in videos. Two types of features are extracted from stego and original sequence such as co-occurrence matrices and the grey level histogram. These features are classified by Support Vector Machine.

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. The experiment results are reported in Section 4.The final conclusions are drawn in Section 5.

2. Related Work

Steganalysis is a process to detect the presence of hidden message in a cover media file. The hidden information can be detected with help of appropriate steganalysis algorithm. Many steganalysis researchers such as Neil F. Johnson and S. Jajodia [7] attempt to categories steganalysis attacks to recover modify or remove the message, based on information available. With cryptography, comparison is made between any possible parts of the plaintext and parts of the ciphertext. In steganography, comparisons may be made between the cover, the stego image, and possible parts of the message. The message in a stego object may or may not be encrypted. If it is encrypted and the message is extracted, the cryptanalysis techniques may be applied. In order to define attack techniques categories used for steganalysis, to consider the corresponding techniques in cryptanalysis. Attacks categories for cryptanalysis are cipher text-only, known plaintext, chosen plaintext, and chosen cipher text.

The difference in the challenge of steganalysis makes this categorization harder to be done. The idea to use a trained classifier to detect data hiding was first introduced in a paper by Avcibas et al. [4]. Farid [3] constructed the features from higher-order moments of distribution of wavelet coefficients and their linear prediction errors from several high-frequency sub-bands. The same authors also showed that SVMs generally provide better performance as classifiers compared to linear classifiers. Other authors have investigated the problem of blind steganalysis using trained classifiers [5].

3. Proposed System

The general structure of the proposed steganalysis method consists of two main stages: feature extraction and classification. The proposed system architecture is shown in Figure 1. Firstly visual feature is extracted from video file for feature calculation and each clip is divided into frames.

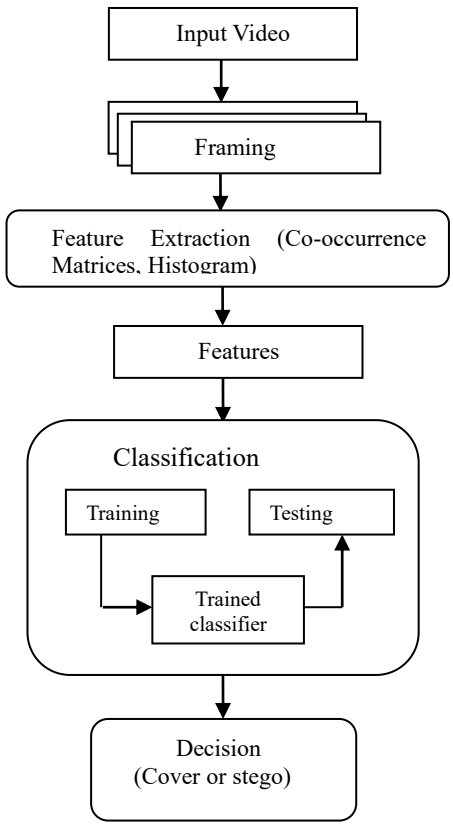

Figure 1. Proposed architecture

Features are analyzed and extracted by using visual features histogram moments, Co-occurrence matrices: contrast, energy, entropy,

homogeneity, and variance. Finally, these features are classified for stego video or original video by using SVM classifier. Classifier accuracy can be increased as a result of feature extraction. SVM is used to detect hidden data.

3.1. Feature Extraction methods

One of the application areas of texture analysis is defect detection in textured images. Texture defect detection can be defined as the process of determining the location and/or extend of a collection of pixels in a textured image with remarkable deviation in their intensity values or spatial arrangement with respect to the background texture.

3.1.1. Co-occurrence Matrices

Among all statistical methods, the most popular one which is based on the estimation of the second order statistics of the spatial arrangement of the gray values is the gray level co-occurrence matrices. A co-occurrence matrix is a square matrix whose elements correspond to the relative frequency of occurrence of pairs of gray level of pixels separated by a certain distance in a given direction. Formally, the elements of a G*G gray level co-occurrence matrix P_d for a displacement vector
$d = (d_x, d_y)$ is defined as :

$$P_d(i,j) = \left| \{((r,s),(t,v)) : I(r,s)=i \, , \, I(t,v)=j\} \right| \quad (1)$$

where $I(\cdot,\cdot)$ denote an image of size NxN with G gray values, (r, s), (t, v)∈ NxN, (t,v)=(r + dx, s + dy) and $|\,.\,|$ is the cardinality of a set. Haralick, Shanmugan and Dinstein [10] proposed 14 measures of textural features which are derived from the co-occurrence matrices, and each represents certain image properties as coarseness, contrast, homogeneity and texture complexity.

Those that are used, in this work, for extracting features in the defect detection of textured images are:

1) Entropy:

$$ENT = -\sum_i \sum_j p(i,j) \log\ p(i,j) \quad (2)$$

Entropy gives a measure of complexity of the image. Complex textures tend to have higher entropy.

2) Contrast :

$$CON = \sum_i \sum_j (i-j)^2\, p(i,j) \quad (3)$$

Contrast feature is a measure of the image contrast or the amount of local variations present in an image.

3) Angular Second Moment:

$$ASM = \sum_i \sum_j \{p(i,j)\}^2 \quad (4)$$

Angular second moment is a measure of the homogeneity of an image. Hence it is a suitable measure for detection of disorders in textures. For homogeneous textures value of angular second moment turns out to be small compared to non-homogeneous ones.

4) Inverse Difference Moment:

$$IDM = \sum_i \sum_j \frac{1}{1+(i-j)^2}\, p(i,j) \quad (5)$$

In Equation (2) - (5), *p(i,j)* refers to the normalized entry of the co-occurrence matrices. That is *p(i,j)* = P_d *(i,j)/R* where R is the total number of pixel pairs (i,j). For a displacement vector d = *(dx,dy)* and image of size NxM R is given by (N-dx)(M-dy).

3.1.2. Gray level histogram

The statistical approach is more useful than structural approach to texture analysis. The simplest statistical features like the mean (6) and standard deviation (8) can be computed indirectly in terms of the image histogram h.

Thus,

$$\mu = \frac{1}{N}\sum_{i=1}^{K} x_i h(x_i) \qquad (6)$$

$$N = \sum_{i=1}^{K} h(x_i) \qquad (7)$$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{K} (x_i - \mu)^2 h(x_i) \qquad (8)$$

Where N = n_1 n_2 is the image dimension, and K is the number of grey levels.

3.2. Support Vector Machine

SVM models the boundary between the classes instead of modeling the probability density of each class (Gaussian Mixture, Hidden Markov Models). SVM algorithm is a classification algorithm that provides state of the art performance in a wide variety of application domains. SVMs have been recently proposed as a new learning algorithm for pattern recognition. SVM learns an optimal separating hyper-plane from a given set of positive and negative examples.

Support Vector Machines (SVM) has recently gained prominence in the field of machine learning and pattern classification [8]. Classification is achieved by realizing a linear or non-linear separation surface in the input space. In Support Vector classification, the separating function can be expressed as a linear combination of kernels associated with the Support Vectors as

$$f(x) = \sum_{x_j \in S} \alpha_j y_j K(x_j, x) + b \qquad (9)$$

Where x_i denotes the training patterns, $y_i \in \{+1, -1\}$ denotes the corresponding class labels and S denotes the set of Support Vectors.The dual formulation yields

$min_{0 \le \alpha_i \le C} W = \frac{1}{2}\sum_{i,j} \alpha_i\, Q_{ij} \alpha_j - \sum_i \alpha_i + b \sum_i y_i \alpha_i$ (10)

where α_i are the corresponding coefficients, b is the offset, $Q_{ij} = y_i y_j K(x_i, x_j)$ is a symmetric positive definite kernel matrix and *C* is the parameter used to penalize error points in the inseparable case.

The Karush-Kuhn-Tucker (KKT) conditions for the dual can be expressed as

$$g_i = \frac{\partial W}{\partial \alpha_i} = \sum_i Q_{ij}\alpha_j + y_i b - 1 = y_i f(x_i) - 1 \qquad (11)$$

and

$$\frac{\partial W}{\partial b} = \sum_j y_j \alpha_j = 0 \qquad (12)$$

This partitions the training set into S the Support Vector set $(0 < \alpha_i < C, g_i = 0), E$ the error set $(\alpha_i < C, g_i < 0)$and R the well classified set $(\alpha_i = 0, g_i > 0)$.

If the points in error are penalized with a penalty factor C', then, it has been shown that the problem reduces to that of a separable case with $C = \infty$. The kernel function is modified as

$$K'(x_i, x_j) = K(x_i, x_j) + \frac{1}{C'}\delta_{ij} \qquad (14)$$

where δ_{ij} = 1 if $i = j$ and δ_{ij} = 0 otherwise. The advantage of this formulation is that the SVM problem reduces to that of a linearly separable case. It can be seen that training the SVM involves solving a quadratic optimization problem which requires the use of optimization routines from numerical libraries. This step is computationally intensive, can be subject to stability problems and is non-trivial to implement. Attractive iterative algorithms like the Sequential Minimal Optimization (SMO), Nearest Point Algorithm (NPA) etc. have been proposed to overcome this problem [8].

3.3. Material and Methods

MSU Stego Video allows hiding any file in a

video sequence. When some hidden data is embedded in a cover video sequence, the encoding of the cover video sequence is analyzed and an algorithm is chosen which provides small data loss after video compression. MSU Stego Video supports multiple video compression formats. Data embed in video sequence using MSU Stego Video. The features are extracted from the original and stego videos. After that, classifier train this features and classify the testing video sequences.

4. Experimental Result

In this section, all experiment based on different video sequences from different sources including the movies and on-line videos from CNN and YouTube. In the experiment, the total frame number of these video sequences is 10880. Figure 2 lists six frame samples of different video clips. In the first scheme texture and histogram features are extracted from each frame of the test sequence. The classifier is trained with feature vectors extracted 40% of the frames in each sequence and the remaining frames are used for testing the classifiers. In this system, utilize a SVM classifier with RBF kernel to construct the classification model on the training and testing dataset.

Figure 2.Six frame samples of different video chips

The result is measured by classification accuracy defined as the number of correctly classified clips over total number of clips. Table 1 lists testing results, consisting of True Negative (TN), False Positive (FP), True Positive (TP), False Negative (FN), and testing accuracy. Training samples and testing samples are from different video chips. That is, if some frames from a certain chips are trained for setting up the classification model, other frames in the same chip are not permitted as testing samples.

Table.1. Detection Accuracy of different video chips

Testing videos	TN	FP	TP	FN	Accuracy
Uncompressed	10428	0	10440	12	99.88%
Compressed	25935	1268	25194	527	96.61%

Figure 3.show the ROC curves under the condition of training samples and testing samples from different video chips which are uncompressed.

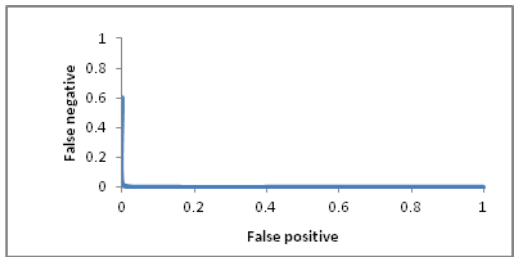

Figure 3. ROC curves for different video chips

Figure 4 lists six frame samples of same video chips and Table 2 lists the testing results. In each video chip, some frames are used for setting up the classification model, other frames are tested. In the same video chips, the accuracy result is slightly larger than the accuracy of different video chips.

Figure 4. Six frame samples of same video chips

Table 2. Detection Accuracy of same video chips

Testing videos	TN	FP	TP	FN	Accuracy
Uncompressed	10436	0	10440	4	99.96%
Compressed	26452	46	26416	10	99.89%

Figure 5 lists the ROC curves under the conditions of training samples and testing samples from same chips which are uncompressed.

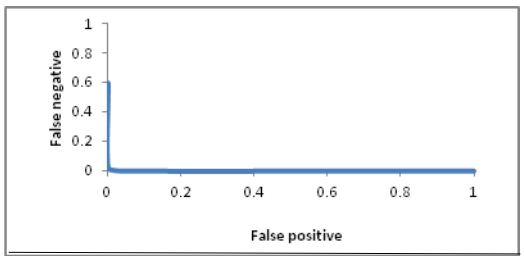

Figure 5.ROC curves for same video chips

5. Conclusion

This paper proposed a general blind steganalysis system for video sequences. The Features are analyzed and extracted from video sequences by using histogram moments, Co-occurrence matrices: contrast, energy, entropy, homogeneity, and variance. These features are classified by SVM classifier. SVM gave accuracy above 99.5%. The future directions in this work can concentrate more on the other statistics from curve let domain like higher order moments and apply this system to videos. The performance of the system can also be improved by using appropriate fusion techniques in the machine learning component.

References

[1] B. G. Mobasseri, M. P. Marcinak. Watermarking of MPEG-2 Video in Compressed Domain Using VLC Mapping. ACM Multimedia and Security Workshop 2005, New York, NY, August 2005. 91-94.

[2] C. S. Lu, J. R. Chen and K. C. Fan. Real-time frame-dependent video watermarking in VLC domain. Signal Processing: Image Communication 20, 2005, 624-642.

[3] H. Farid and L. Siwei. Detecting hidden messages using higher-order statistic and support vector machines. In F.A.P. Petitcolas, editor, Information Hiding 5th International Workshop, volume 2578 of Lecture Notes in Computer Science pages 340–354. Springer-Verlag, 2003.

[4] I. Avcibas, N. Memon, and B. Sankur. Steganalysis using image quality metrics. In Proceeedings of SPIE Electronic Imaging, Security and Watermarking of Multimedia Contents, volume 4314, pages 523–531, San Jose, CA, 2001.

[5]. J.J. Harmsen and W.A. Pearlman. Steganalysis of additive noise modelable information hiding. In Proceedings of SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents V, pages 131–142, Santa Clara, CA, 2003.

[6]. J. S. Jainsky, D. Kundur and D. R. Halverson. Towards Digital Video Steganalysis using Asymptotic Memoryless Detection. MM&Sec'07, September 2021, 2007, Dallas, Texas, USA.

[7].Natarajan Meghanathan and Lopamudra Nayak, Steganalysis Algorithms for Detecting the Hidden Information in Image,Audio and Video Cover Media, International Journal of Network Security & Its Application (IJNSA), Vol.2, No.1,
January 2010.

[8]. N.F. and Jajodia, Johnson, S. (1998) 'Steganalysis of Images Created Using Current Steganography Software', Workshop On Information Hiding Proceedings, Portland, Oregon, USA.

[9]. S.V.N. Vishwanathan, M. Narasimha Murty. SSVM: A Simple SVM Algorithm. Automation, Indian Institute of Science, Bangalore 560 012, INDIA

[10]. U. Budhia, D. Kundur, T. Zourntos. Digital Video Steganalysis Exploiting Statistical Visibility in the Temporal Domain. IEEE Trans. Information Forensics and security, 2006, (1)4:502-516.