International Journal of Automation and Computing 7(4), November 2010, 531-542
DOI: 10.1007/s11633-010-0537-1

Passive Steganalysis Based on Higher Order Image Statistics of Curvelet Transform

S. Geetha[1] Siva S. Sivatha Sindhu[1] N. Kamaraj[2]

[1]Department of Information Technology, Thiagarajar College of Engineering, Madurai 625 015, Tamil Nadu, India

[2]Department of Electrical and Electronics Engineering, Thiagarajar College of Engineering, Madurai 625 015, Tamil Nadu, India

Abstract: Steganographic techniques accomplish covert communication by embedding secret messages into innocuous digital images in ways that are imperceptible to the human eye. This paper presents a novel passive steganalysis strategy in which the task is approached as a pattern classification problem. A critical part of the steganalyser design depends on the selection of informative features. This paper is aimed at proposing a novel attack with improved performance indices with the following implications: 1) employing higher order statistics from a curvelet sub-band image representation that offers better discrimination ability for detecting stego anomalies in images, as compared to other conventional wavelet transforms; 2) increasing the sensitivity and specificity of the system by the feature reduction phase; 3) realizing the system using an efficient classification engine, a neuro-C4.5 classifier, which provides better classification rate. An extensive experimental evaluation on a database containing 5600 clean and stego images shows that the proposed scheme is a state-of-the-art steganalyser that outperforms other previous steganalytic methods.

Keywords: Image steganalysis, curvelet higher order statistics, neuro-C4.5 classifier, information forensics, information security.

1 Introduction

Steganography, coined from the Greek words "stegos", meaning roof or covered and "graphia", which means writing, is the art and science of hiding the fact that communication is taking place. Through steganography, a secret message is embedded inside a piece of unsuspicious information and sent without anyone knowing of the existence of the secret message. Secrets can be hidden inside all sorts of cover information: text, image, audio, video, and more. Most steganographic utilities nowadays, hide information inside images, as this is relatively easy to implement. However, there are tools available to store secrets inside almost any type of cover source.

Since steganography is used to hide the occurrence of communication, it creates a potential problem when this technology is misused for planning criminal activities. Hence, it becomes very essential to distinguish between the benign image from the anomalous stego image, which is the motivation of this research. The art of discovering and rendering useless such covert messages is called as the steganalysis, the counter measure to steganography. With careful selection of an appropriate cover image and a stego-tool, it is possible to create a stego-image that does not appear to be different within the limits of human perception. However, electronically, each of these tools leaves a fingerprint or signature in the image that can be used to alert an observer to the presence of a hidden message. Discovering a hidden message is the first step in steganalysis and is considered an "attack" on the hidden information.

There are two common types of attacks against steganography. The first is the "known message" attack. In this case, the steganalyst has a known hidden message and the corresponding stego-image. Here, the objective is to determine patterns that result from hiding the message. These patterns can then be used to analyze other stego-objects in the future. The second type is the "chosen message" attack. In this case, the steganalyst will create a message and use a known stego-tool to create a stego-image. This, known stego image is then analyzed to determine patterns for later use against other stego-images. Once a stego-image has been discovered, there are several steps that can be taken to disable or destroy the hidden message.

Steganalysis can be also broadly classified into two categories: active and passive warden style. Active steganalysis deals with the estimation of the facts, such as the embedded message length, locations of the hidden message, secret key used in embedding, and finally, the extraction of the entire message that is hidden. Passive steganalysis on the other hand detect only the presence or absence of a hidden message.

On the outset, deciding whether the cover media contains any secret message embedded in it or not, is essential to steganalysis. Although it is not complicated to inspect suspicious objects and extract hidden messages by comparing them to the original versions, the restricted portability and accessibility of original cover-signals generally make blind steganalysis more attractive and reasonable in many practical applications. Blindness is meant to analyze stego-data without the knowledge of the original signal and without exploiting the embedding algorithm. Hence, detecting the existence of hidden information becomes quite difficult and complex without exactly knowing which embedding algorithm, hiding domain, and steganographic keys were used. This motivates our current research: extracting low-dimensional, informative features that are significantly sensitive to data hiding process and devising a feature-based algorithm to classify multimedia objects as bearing hidden data or not. Our objective is not to extract the hidden messages or to identify the existence of particular information (as it is in watermarking applications), but only to

Manuscript received February 11, 2009; revised September 24, 2009

determine whether a multimedia object was modified by information hiding techniques or not — passive steganalysis. Once classified, the suspicious objects can then be inspected in detail by any particular data embedding/retrieving algorithms. This pre-process would particularly contribute to save time in active steganalysis.

The paper is organized as follows. Related works from the state-of-the-art literature are presented in Section 2. The rationale for the choice of the curvelet features and the neuro-C4.5 classifier is discussed in Section 3. In Section 4, the setup for performance evaluation of the proposed system and the empirical results obtained are given. Section 5 concludes the paper, representing a number of issues for future research.

2 Related work

2.1 Steganographic domains

In recent years, there has been significant research effort in steganalysis with primary focus on digital images. All the popular data hiding methods can be divided into two major classes: spatial domain based and transform domain based. Spatial domain based techniques are easy to implement, providing high payload capacity but their robustness is weaker than their counter part. Some tools, such as StegoDos, S-tools, and EzStego, provide spatial-domain-based steganographic techniques[1, 2]. Least significant bit (LSB) addition[3, 4] or substitution[5, 6] method is the most popular hiding technique. Generally, these techniques operate on the principle of tuning the parameters of the cover signal (e.g., the payload or disturbance) so that the difference between the cover signal and the stego signal is little and imperceptible to the human eyes. Nevertheless, computer statistical analysis is still promising to detect such a distinction that is difficult for humans to perceive.

Transform domain based techniques include data embedding process mainly using discrete Fourier transform (DFT), discrete cosine transform (DCT), and discrete wavelet transform (DWT). Hiding can also be performed in the transform domain, e.g., DCT[7−12], or DWT domain[13, 14]. Regardless of the domain, "significant" transform coefficients are often selected to mix with secret/perturbing signal in a way such that information hiding or watermarking is transparent to human eyes. For instance, Cheng and Huang[11] proposed an additive approach to hiding secret information in the DCT and DWT domains. Lie et al.[7] proposed a two-level data embedding scheme, based on additive spread spectrum and spectrum partition, for applications in copy control, access control, robust annotation, and content-based authentication. Xiang and Wu[15] described a new content-based image watermarking scheme. The Harris-Laplace detector is adopted to extract feature points. The local characteristic regions (LCRs) are adaptively constructed based on scale-space theory. Then, the LCRs are mapped to geometrically invariant space by using image normalization technique. Finally, several copies of the digital watermark are embedded into the non-overlapped LCRs by quantizing the magnitude vectors of discrete Fourier transform (DFT) coefficients.

Passive and active warden styles. Xu et al.[16] viewed active steganalysis as blind sources separation (BSS) problem and solved it with independent component analysis (ICA) algorithm under the assumption that the embedded secret message is an independent identically distributed (i.i.d.) random sequence and independent of the cover image. Fridrich[17] described an improved version of passive steganalysis in which the features for the blind classifier are calculated in the wavelet domain from the higher-order absolute moments of the noise residual. The features are calculated from the noise residual because it increases the features′ sensitivity to embedding, which leads to improved detection results. Geetha et al.[18] presented a passive approach for image steganalysis using genetic-X-means classifier and content independent image quality metrics. They have employed different signal features and applied various machine learning techniques ranging from genetic algorithms to decision tree learners, for the task of blind image- and audio-steganalysis[19−22].

2.2 Steganalysis through signal processing

Some steganalytic methods[23−25] are proposed in the DCT domain. Manikopoulos et al.[24] estimated the probability density function (PDF) of DCT coefficients for the test image, and calculated its difference with respect to a reference PDF, which is then used as a feature input to a trained two-layer neural network for classification. In their work, the reference PDF derived by averaging PDFs from all plain images in the database is required for this similarity measure. Generally, the representation of a set of plain images in terms of a reference PDF is questionable. Fridrich and Goljan[23] described that a modified image block will most likely become saturated (i.e., at least one pixel with the gray value 0 or 255) in a JPEG-format. If no saturated blocks can be found, there will be no secret messages therein. Otherwise, a spatial-domain steganalytic method[23] mentioned earlier can be used to analyze these saturated blocks. In [25], the author modeled the common steganographic schemes as a linear transform between the cover and stego images, which can be estimated after at least two copies of a stego image, were obtained. In [26], a steganalytic scheme was devised to deal with information hiding schemes mixing a secret and a cover signal by an addition rule. The phenomenon that the centre of mass of the histogram characteristic function in a stego image moves left or remains the same to that of the cover image was observed and exploited to distinguish the stego images from the plain ones.

2.3 Steganalysis using distortion measures

Our research is based on the extension of the fact that hiding information in digital media requires alterations of the signal properties that introduce some form of degradation, no matter how small. These degradations can act as signatures that could be used to reveal the existence of a hidden message. The idea that the addition of a watermark or message leaves unique artifacts, which can be detected using the various distortion metrics i.e., image quality measures (IQM) is introduced in [27]. This may be considered to be a break-through contribution in the field of passive steganalysis.

It is noticed that most of the steganalytic schemes were designed either in specific operating domain, or even for particular steganographic algorithm. Building a universal steganalytic system is, up to now, a challenging exercise. Lie and Lin[28] modeled a universal passive steganalyser that operates to distinguish stego images from clean images using two features only, namely, gradient energy and statistical variance of the Laplacian parameter. The system lacks the ability to strongly attack a wavelet based stego system but that can be solved by using a feature that is more sensitive to such embedding strategy.

2.4 Steganalysis using high-order statistics

There are many works reporting that high-order statistics are very effective in differentiating stego-images from cover-images. Farid[29] proposed a general steganalysis algorithm based on image high-order statistics. In this method, a statistical model based on the first (mean) and higher-order (variance, skewness, and kurtosis) magnitude statistics, extracted from wavelet decomposition, is used for image steganography detection. In [30], a steganalysis method based on the moments of the histogram characteristic function was proposed. It has been proved that, after a message is embedded into an image, the mass center (the first moment) of histogram characteristic function will decrease. Holotyak et al.[31] used higher-order moments of the PDF of the estimated stego-object in the finest wavelet level to construct the feature vectors. Due to the limited number of features used in the steganalysis technique proposed in [30], Shi et al.[32] proposed the use of statistical moments of the characteristic functions of the wavelet sub-bands. Because the n-th statistical moment of a wavelet characteristic function is related to the n-th derivative of the corresponding wavelet histogram; the constructed 39-dimensional (39D) feature vector has proved to be sensitive to embedded data. Usually, the steganalysis algorithms based on the higher-order statistics achieved satisfactory performance on image files, regardless of the underlying embedding algorithm.

However, since the data-embedding method is typically unknown prior to detection, we focused on the design of a unified blind steganalysis algorithm to detect the presence of steganography independent of the steganography algorithm used. Moreover, we focused on passive detection as opposed to active warden steganalysis[27], which aims to detect and modify the hidden content. In this work, we employed the higher order statistical features that were collected from a new transform domain, i.e., curvelet transform domain. To our knowledge, this work is typically a novel and first attempt of employing these features for steganalysis. Experimental results also prove that our claim is justified.

3 Design of the steganalyser

For a group of given data samples (e.g., coefficients in any sub-band of the image multi-resolution representation), the first important step of machine-learning-based image steganalysis is to choose representative features. Then, a decision function is built based on the feature vectors extracted from the two classes of training images: photographic cover images and stego-images with hidden information. The performance of the classifier depends on the discrimination capabilities of the features. Also, if the feature vector has low dimension, the computational complexity of learning and implementing the decision function will decrease. In summary, we need to find informative, low-dimensional features.

3.1 Choice of curvelet statistics

A good steganalytic feature should be accurate, consistent, and monotonic in predicting the status of an image as stego or clean image. In the context of steganalysis, prediction accuracy can be interpreted as the ability of the measure to detect the presence of hidden message with minimum error on the average. Similarly, prediction monotonicity signifies that the features should ideally be monotonic in their relationship to the embedded message size or watermark strength. Finally, prediction consistency relates to the feature′s ability to provide consistently accurate predictions for a large set of watermarking or steganography techniques and image types. This implies that the spread of quality features due to factors of image variety, active warden, or passive warden steganography methods should not eclipse the score differences arising from message embedding artifacts.

The schematic diagram describing the embedding process in the steganography system is shown in Fig. 1. Let $C(x, y)$ denote a cover image object and $S(x, y)$ be its stego-version. Let $GN(x, y)$ be an independent and identically distributed (i.i.d.) Gaussian noise; then, the stego-image can be expressed as $S(x, y) = C(x, y) + GN(x, y)$ with the additive noise model. A feature may be regarded as a good stego clue when it magnifies the distance between $C(x, y)$ and $S(x, y)$, i.e., even a small perturbation introduced by the stego embedding process is proportionately enlarged. The steganographic algorithms hide more data along the edges and on the texture areas in an image since such locations offer appreciable imperceptibility to the stego images. Many works have been reported stating this fact[8, 12, 13].

Recently, researchers have changed their focus to the curvelet transform[33] due to its outstanding properties — sparse representation and random direction selectivity. The curvelet transform is a multi-scale pyramid with many directions and positions at each fine scale and has needle-shaped elements at the fine scales. These properties allow details in nearly all directions to be described and sparser representation to be obtained. The curvelet statistics efficiently represent the curves, i.e., edges and texture regions that exhibit changes. Any manipulations done on these locations could be easily captured through the higher order statistics. A detailed recapitulation on curvelet transformation is presented in the following section.

3.2 Image sub-band decomposition: choice of curvelet transforms

The decomposition of images using basis functions that are localized in spatial position, orientation, and scale have proven to be 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. A special member of this emerging family of multi-scale geometric transforms is the curvelet transform[34−36], which overcomes the inherent limitations of traditional multi-scale representations, such as wavelets[37].

To process 2D image signals, the 2D wavelet transform, composed of the tensor product of two 1D wavelet basis functions, takes advantage of the separable transform kernels to realize the wavelet transform horizontally first and then vertically. Such kernels of the 2D wavelet transform are isotropic, leading to that the local transform modulus maxima only reflect the positions of those maxima that are across the edge, instead of along the edge. However, singularities in most of the images are characterized by lines and curves, which seriously reduce the approximation efficiency of wavelet. In this circumstance, the traditional wavelet transform is limited in the field of image processing. However, curvelet transform can sparsely characterize the high-dimensional signals that have line, curve, or hyperplane singularities, and the approximation efficiency is one magnitude order higher than wavelet transform through its anisotropic property[36].

Since more data are embedded along the edges and curvelet statistics capture these small manipulations efficiently, the usage of these measures for steganalysis seems to be a clever choice. Also, the application of curvelet statistics for image steganalysis is relatively unexplored.

The curvelet transform[33] is introduced to address the problem of optimally finding sparse representations of objects with discontinuities along C^2 edges. This curvelet transform inherits the ridgelet conception[36] and is constructed without using ridgelets[33].

Let μ be the triple (j, l, k) in the frequency plane; $j = 0, 1, 2, \cdots$ is a scale parameter; $l = 0, 1, \cdots, 2^j$ is an orientation parameter; and $k = (k_1, k_2)$, $k_1, k_2 \in \mathbf{Z}$ is a translation parameter pair. A curvelet coefficient is simply the inner product between an element $f \in L^2(R^2)$ and a curvelet ϕ_μ given by

$$C_\mu \equiv \langle f, \phi_\mu \rangle = \int_{R^2} f(x)\overline{\phi_\mu(x)}\,\mathrm{d}x = \frac{1}{(2\pi)^2}\int \widehat{f}(\varpi)U_j(R_{\theta_l}\varpi)e^{i\langle x_k^J, \varpi\rangle}\mathrm{d}\varpi \qquad (1)$$

where R_θ is the distance covered due to rotation by θ radians, $J = (j, l)$ is the index of a wedge for all k within it, and U_j is a polar "wedge" window of radial dilation and angular translations[11]. Define coarse scale curvelets as

$$\phi_{j0,k}(x) = \phi_{j0}(x - 2^{-j0}k),\ \widehat{\phi}_{j0}(\varpi) = 2^{-j0}A_0(2^{-j0}|\varpi|) \qquad (2)$$

where A_0 is a radial window and the curvelet is isotropic. For $j \geqslant j_0$, we have a reconstruction formula

$$f = \sum_\mu \langle f, \phi_\mu \rangle \phi_\mu. \qquad (3)$$

Similar to other multiscale pyramids, curvelet transformation decomposes the images into several frequency scales. Suppose we have an object supported in $[0, 1]^2$, which has a discontinuity across a nice curve and which is otherwise smooth. The procedure used to define the m term approximation $\hat{f}_m^W$ in wavelet transform is also adapted here. i.e., a standard Fourier representation is employed and it is approximated with $\hat{f}_m^F$ built from the best m nonzero Fourier terms. Finally, we get the m term approximation $\hat{f}_m^C$ in curvelet transform, m term approximation as

$$\lim_{m\to\infty}\left\|f - \hat{f}_m^F\right\| = m^{-\frac{1}{2}}, \quad \lim_{m\to\infty}\left\|f - \hat{f}_m^W\right\| = m^{-1}, \quad \lim_{m\to\infty}\left\|f - \hat{f}_m^C\right\| = Cm^{-2}(\log_2 m)^3. \qquad (4)$$

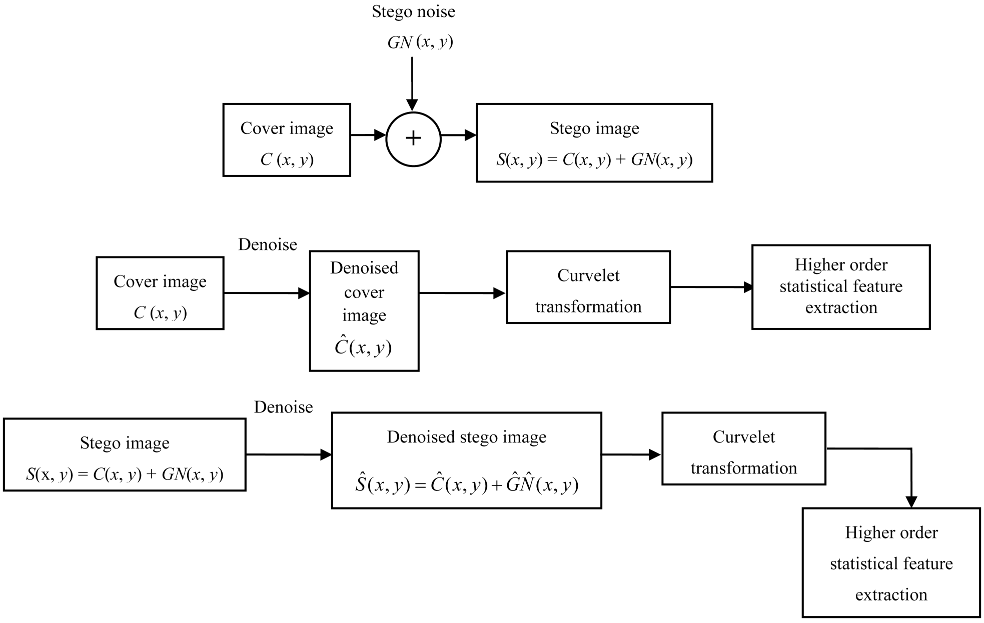


Fig. 1 Schematic descriptions of (a) additive noise image-steganography model, (b) denoising a cover-image object, and (c) denoising a stego-image object

Clearly, the last result is the smallest and nearly as good as m^{-2} term occurring in an adaptive asymptotic representation[33]. Even the hot dual-tree wavelet is not sparser than the ordinary wavelet[37]. Conceptually, the curvelet transform decomposes an image in many directions with a multiscale pyramid structure and produces needle-shaped elements at fine scales. These help to obtain a sparser representation than other multiscale representations, such as wavelets. For a detailed insight into the curvelet transforms, the reader is referred to[33].

We propose to decompose the given image into six subbands $b_k, k = 1, 2, 3, 4, 5, 6$ using curvelet transformation as in [33]. Let us denote by R_1 the set of these six curvelet sub-bands plus the image itself. The noise residual component for a cover image and its stego-image possess different statistics, which are useful in steganalysis. Since curvelet coefficients possess strong intra and inter subband dependencies, we propose to construct a set R_2 of six noise residual sub-band to exploit these dependencies as follows. Take a sub-band coefficient $b_k(i, j)$ as an example, where (i, j) denotes the spatial co-ordinates at band k. The magnitude of the denoised component of this band can be computed by applying Wiener filter over these coefficients.

3.3 Feature extraction

Given this image decomposition, the statistical model is composed of the higher order statistics like "Feature 1 – mean, Feature 2 – variance, Feature 3 – skewness, and Feature 4 – kurtosis" of the sub-band coefficients at each orientation. These statistics characterize the basic coefficient distributions. The second set of statistics is based on the noise component of the stego-image in the curvelet domain. The noise component was obtained using the denoising filter, as in [34]. We reiterate that the denoising step increases the signal to noise ratio (SNR) between the stego signal and the cover image, thus making the features calculated from the noise residual more sensitive to embedding and less sensitive to image content. The denoising filter is designed to remove Gaussian noise from images under the assumption that the stego image is an additive mixture of a non-stationary Gaussian signal (the cover image) and a stationary Gaussian signal with a known variance (the noise). As the filtering is performed in the curvelet domain, all our features (statistical moments) are calculated as higher order moments of the noise residual in the curvelet domain (see Algorithm 1).

Algorithm 1. Feature calculation

Step 1. Calculate the first level curvelet decomposition of the stego image. Denote the six sub-bands as $b_k(i, j)$ where k runs through the index 1 to 6 and (i, j) runs through some index set J.

Step 2. In each sub-band, estimate the local variance of the cover image for each curvelet coefficient using the maximun a posteriori (MAP) estimator for four sizes of a square $N \times N$ (i.e., N cross N) neighbourhood, for

$$N \in \{3, 5, 7, 9\}$$

$$\hat{\sigma}_N^2(i, j) = \max(0, \frac{1}{N^2} \sum_{(i,j)\in N} w^2(i, j) - \sigma_0^2), \quad (i, j) \in J. \tag{5}$$

Take the minimum of the four variances as the final estimate,

$$\hat{\sigma}^2(i, j) = \min(\hat{\sigma}_3^2(i, j), \hat{\sigma}_5^2(i, j), \hat{\sigma}_7^2(i, j), \hat{\sigma}_9^2(i, j)), \ (i, j) \in J. \tag{6}$$

Step 3. The denoised curvelet coefficients for the first sub-band are obtained using the Wiener filter

$$b_{1\_den}(i, j) = b_1(i, j) \frac{\hat{\sigma}^2(i, j)}{\hat{\sigma}^2(i, j) + \hat{\sigma}_0^2}, \quad (i, j) \in J \tag{7}$$

and, similarly, for the other five sub-bands also, the denoised curvelet coefficients are obtained.

Step 4. Calculate the noise residual in each sub-band like

$$r_{1\_den}(i, j) = b_1(i, j) - b_{1\_den}(i, j), \quad (i, j) \in J \tag{8}$$

and similarly $r_{k\_den}(i, j), (i, j) \in J$, and $k = 2, 3, \cdots, 6$.

Step 5. Calculate the mean, variance, skewness, and kurtosis of the six sub-band coefficients of the image and also of each noise residual.

We set the parameter $\sigma_0^2 = 0.5$, which is the same value as in [26] and corresponds to the variance of the stego signal for an image fully embedded with ±1 embedding.

The algorithm yields 24 (6 bands and 4 statistics from each band) statistics from the noise residue and 24 statistics from the image. Combining these statistics, we get a total of 48 statistics that form a feature vector that is used to train the steganalyser.

3.4 Choice of neuro-C4.5 classifier

This paper investigates the use of neuro-C4.5 algorithm to generate rules automatically instead of manual intervention. For any learning algorithm, generalization and comprehensibility could be considered as inevitable properties. Generalization accounts for accurate prediction of unobserved data. Neural network (NN) reports to have strong generalization ability[38−40].

Comprehensibility, i.e., the ability to explain the learned knowledge, is vital in terms of usage in reliable applications like steganalysis. Decision trees[41] are with good comprehensibility because the learned knowledge is explicitly represented in trees, while neural networks are with poor comprehensibility because the learned knowledge is implicitly encoded in a lot of connections. Therefore, these two virtues (comprehensibility and generalization) are incorporated into a single algorithm neuro-C4.5. This algorithm trains a neural network (NN) (back propagation) at first. Then, the trained NN is employed to generate a new training set through replacing the desired class labels of the original training examples, with those output from the trained NN. Some extra training examples are also generated from the trained NN and added to the new training set. Finally, a C4.5 decision tree is grown from the new training set. Since its learning results are decision trees, the comprehensibility of neuro-C4.5 is better than that of neural network. Moreover, experiments show that the generalization ability of neuro-C4.5 decision trees is better than that of simple C4.5 decision trees[42]. The pseudo-code of the model is provided in Algorithm 2.

Algorithm 2. Neuro-C4.5 algorithm applied to image steganalysis

Input. Training set $T = \{(x_{1,1}, x_{1,2}, \cdots x_{1,48}, y_1), (x_{2,1}, x_{2,2}, \cdots x_{2,48}, y_2), \cdots, (x_{n,1}, x_{n,2}, \cdots x_{n,48}, y_n)\}$, where $x_{i.j}$ represents a feature vector formed from the curvelet statistics of the image data set, in which i denotes the sample number, j represents the feature number and j ranges from 1 to 48, and y_i represents the class label (clean, stego type) of the i-th sample, extra data ratio α, neural learner (NL), trials of bootstrap[43] sampling β.

Output. The decision tree C4.5 DT model.

Step 1. Train the NN from T via bagging[44] $NL^* = Bagging(T, NL, \beta)$.

Step 2. $S' = \phi$.

Step 3. Process the original training set with the trained NL^*

```
For ($i = 1, \cdots, n$)
{
$y'_i = NL^*(x_i : (x_i, y_i) \in T)$
$S' = S' \cup \{(x_i, y'_i)\}$
}
```

Step 4. Generate extra training data from the trained NL^*

```
For ($j = 1, \cdots, \alpha \times n$)
{
$x'_j$ = Random()
$y'_j = NL^*(x'_j)$
$S' = S' \cup \{(x'_j, y'_j)\}$
}
```

Step 5. Read the records from S'.

Step 6. Tokenize each record and store it in an array.

Step 7. Determine whether the attribute is discrete or continuous.

Step 8. If (discrete attribute)

1) Find the probability of occurrence for each value for each class

2) Find the entropy:

$$I(P) = -(p_1 \times \log(p_1) + p_2 \times \log(p_2) + \cdots + pn \times \log_2(p_n)).$$

3) Calculate the information gain $\text{Gain}(X, T) = \text{Info}(T) - \text{Info}(X, T)$

Else

1) For continuous attributes, the values are sorted and the gain for each partition is found.

2) The partition with the highest gain is taken.

Step 9. Construct the tree C4.5 DT with the highest information gain attribute as the root node and values of the attribute as the arc labels.

Step 10. Repeat until categorical attributes or the leaf nodes are reached.

Step 11. Derive rules following each individual path from root to leaf in the tree.

Step 12. The condition part of the rules is built from the label of the nodes and the labels of the arcs: the action part will be the classification (e.g., pure, stego-scheme name, etc.).

3.5 Architecture of neuro-C4.5 image steganalyser model

The general architectural framework for neuro-C4.5 based image steganalysis system is illustrated in Fig. 2. The multimedia traffic (image, video, image, text, HTML pages, etc.) is keenly monitored by the system. Whenever the entry of image documents is sensed, the steganalyser system is triggered. The system consists of two main phases. They are training phase and testing phase. In the training phase, the knowledge base that contains rules to distinguish the

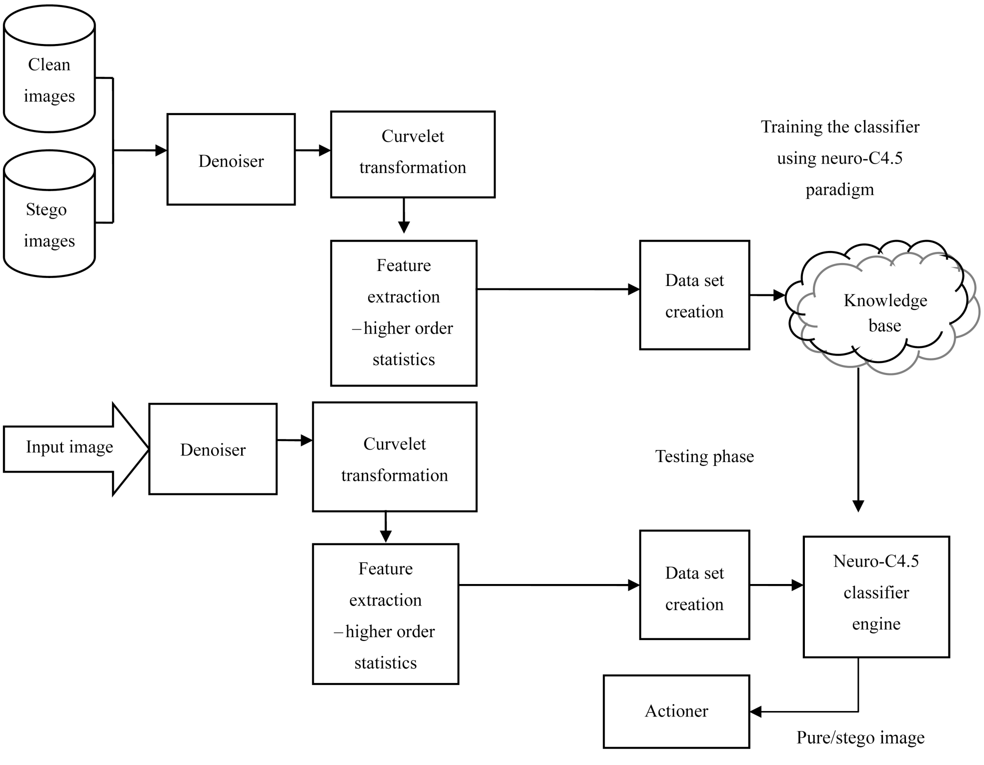


Fig. 2 Functional model of the proposed image steganalyser

stego objects from the clean files is constructed by the neuro-C4.5 model, from a set of known cover and stego image files. The classification performance of the evolved model is tested with fresh data during the detection stage. The real time detection is performed in the testing phase.

4 Experimental topology

In our experiments, the discrimination performance of higher order statistics gathered from curvelet coefficients as features is analyzed first. Then, the classification performance of our steganalyser under the prepared test image set is reported. The impacts of embedding rate and the effectiveness of the selected features are explored.

4.1 Preparation of test images and schemes

The design of experiments is important in evaluating the steganalytic algorithm. The key considerations include the following steps.

First, from the point of "generalization", the proposed image features and associated classifier should be capable of identifying the existence of hidden data that are possibly generated by using various kinds of embedding methods, regardless of steganography or watermarking, and regardless of spatial or transform-domain operations.

Second, from the outlook of "performance", the classifier should, on one hand, detect hidden data as likely as possible (regardless of how transparent the embedded secret information is), and on the other hand, keep false alarms to as few as possible for plain images.

Third, in view of "robustness", the classifier should be capable of differentiating the effect of ordinary image processing operations (such as filtering, enhancement, etc.) from that of data embedding.

On the grounds of the above considerations, six published methods based on two types of principles, LSB embedding, and spread spectrum, were chosen for evaluation:

Scheme #1: Digimarc[45]
Scheme #2: Pretty good signature (PGS)[46]
Scheme #3: Cox et al.′s[47]
Scheme #4: S-tools[48]
Scheme #5: Steganos[49]
Scheme #6: JSteg[50]

They can be further categorized into:

1) Steganography (#4, #5, #6) or watermarking (#1, #2, #3) purpose;

2) Spatial (#2, #4, #5) or transform (#1, #3, #6) domain operation.

For further testing and to verify the effectiveness of the features selected, we select an extra scheme based on the wavelet domain. The ability of the system to attack any new stego system is studied. The system is not trained with the stego patterns of scheme #7.

3) Scheme #7: Kim et al.′s method[14].

It is expected that the difference between a cover image and its stego version can be easily detected when more secret messages are embedded. Hence, the capacity of the payload of a steganography scheme should be taken into account in evaluating the detection capacity of a steganalytic classifier. To depict this, the embedding rate (ER) characterizing a scheme that is defined as the ratio between the number of embedded bits and the number of pixels in an image, is used.

To test the performance of the proposed method, our cover image dataset consists of 200 images with a dimension of 256×256 8-bit gray-level photographic images, including standard test images such as Lena, Baboon, and also images from [51]. Our cover images contain a wide range of outdoor/indoor and daylight/night scenes, including nature (e.g., landscapes, trees, flowers, and animals), portraits, manmade objects (e.g., ornaments, kitchen tools, architectures, cars, signs, and neon lights), etc. This database is augmented with the stego versions of these images using the above mentioned seven schemes, at various embedding rates. Also, a separate image set was generated by applying the image processing techniques, like JPEG compression (at several quality factors), low-pass filtering, image sharpening, etc. Our generation procedure is aimed at making even contributions to database images from different embedding schemes, from original or stego, and from processed or non-processed versions, so that the evaluation results can be more reliable and fair. To investigate the impact of varying ER, three different ERs are attempted for each scheme, like (#1) 5%, (#2) 10%, (#3) 20% of the size of the maximum payload capacity prescribed by the respective techniques. Finally, the entire database contains $200 \times 4 \times 7 = 5600$ (No. of images) × (three different ERs + 1 for the clean set) × (No. of schemes evaluated) images on the whole.

4.2 Feature extraction

The features are calculated based on the algorithm in Algorithm 1. We get an overall of 48 statistical features that represent a file.

4.3 Feature normalization

Before proceeding to evaluate the performance of the classifier, the discrimination capability of the proposed features is to be analyzed. The experiment involves breaking of different steganographic or watermarking strategies, which may adapt extremely different techniques for embedding ranging from LSB substitution to embedding inside the wavelet co-efficient.

Hence, the feature set formed has to be normalized before feeding into the classifier for training to achieve a uniform semantics to the feature values. A set of normalized feature vectors as per the data smoothing function[52], $\check{f}_i = (f_i - f_i^{\min})/(f_i^{\max} - f_i^{\min})$, are calculated for each seed image to explore relative feature variations after and before it is modified. $\check{f}_i$, $f_i^{\min}$, and $f_i^{\max}$ represent the i-th feature vector value, the corresponding feature′s minimum and maximum value, respectively.

4.4 Neuro-C4.5 classifier

In the sequel, the model is incorporated in Java[53], and the algorithm described in Algorithm 2 is implemented as per the method proposed. The ensemble classifier was trained and evaluated by using 4800 images out of the whole database, excluding those generated by using scheme #7 (employed as the test images to see how the proposed features behave when there is a mismatch between the operation domains). Here, two-thirds (3200) of the images were randomly chosen as the training set, and the others (1600 images) act as the validation set.

To demonstrate the increase in the detection performance, we compare the detection rate of the proposed steganalyser, (neuro-C4.5 classifier operated by the curvelet statistics) with a similar system proposed in [27], a passive steganalyser designed based on the image quality metrics. The classification and error rates obtained are listed in Table 1, and graphically shown in Fig. 3.

Results show that the average classification rate is 85.62 %. We are interested in analyzing the detectability of proposed features and classifier against embedding schemes of different applications or principles. Table 2 depicts the specificity and sensitivity of the model. The system offers an appreciable range like 84.12 % to 95.79 % sensitivity and 77.19 % to 94.27 % for specificity. Table 3 lists true positive (TP) rate and true negative (TN) rates to see differentiation in performances among: 1) six targeted embedding schemes, 2) steganographic or watermarking applications, 3) spatial or DCT operation domain, and 4) types of processed non-stego images. We also analyzed the TN rates for the original, smoothed, sharpened, and JPEG-compressed non-stego images. It has been found that our system has a better performance in recognizing the plainness of JPEG-compressed images. The higher TN rate for JPEG-compressed images is beneficial to real applications, since most images will be compressed in JPEG form.

Table 2 Sensitivity and specificity of the steganalytic classifier

Scheme	Sensitivity (TP)/(TP+FN)		Specificity (TN)/(TN+FP)	
	IQM	Proposed scheme	IQM	Proposed scheme
DigiMarc	80.00 %	87.47 %	80.00 %	88.36 %
PGS	80.00 %	94.27 %	90.00 %	86.06 %
Cox	66.67 %	77.19 %	75.00 %	85.79 %
S-tools	69.23 %	82.58 %	85.71 %	95.79 %
Steganos	66.67 %	81.79 %	75.00 %	89.57 %
JSteg	70.00 %	79.87 %	70.00 %	84.12 %

4.5 Influence of embedding rate

In this experiment, the images at various payload capacities were selected to see the influence on detectability. The ERs for the six embedding schemes were tried at 5 %, 10 %, and 20 % of the maximum hiding capacities in their proposed versions. The experimental results are listed in Table 4, which depicts that the average classification rate still remains above 81.85 % for 20 %, 74.33 % for 10 %, and 69.71 % for 5 % of maximum payload capacity. The results for steganographic schemes are more promising than for the watermarking schemes, as the steganographic schemes carry more hidden data than those of watermarking schemes, which makes the measured features more distinguishable for detection. The results, depicted in Fig. 4, clearly reveal that our proposed system, curvelet statistics coupled with neuro-C4.5 classifier, still yield reasonable results for stego images of less ER.

Table 3 Average PD/ND rates for performance differentiation between different target schemes, different applications, different operation domains, and different types of nonstego images

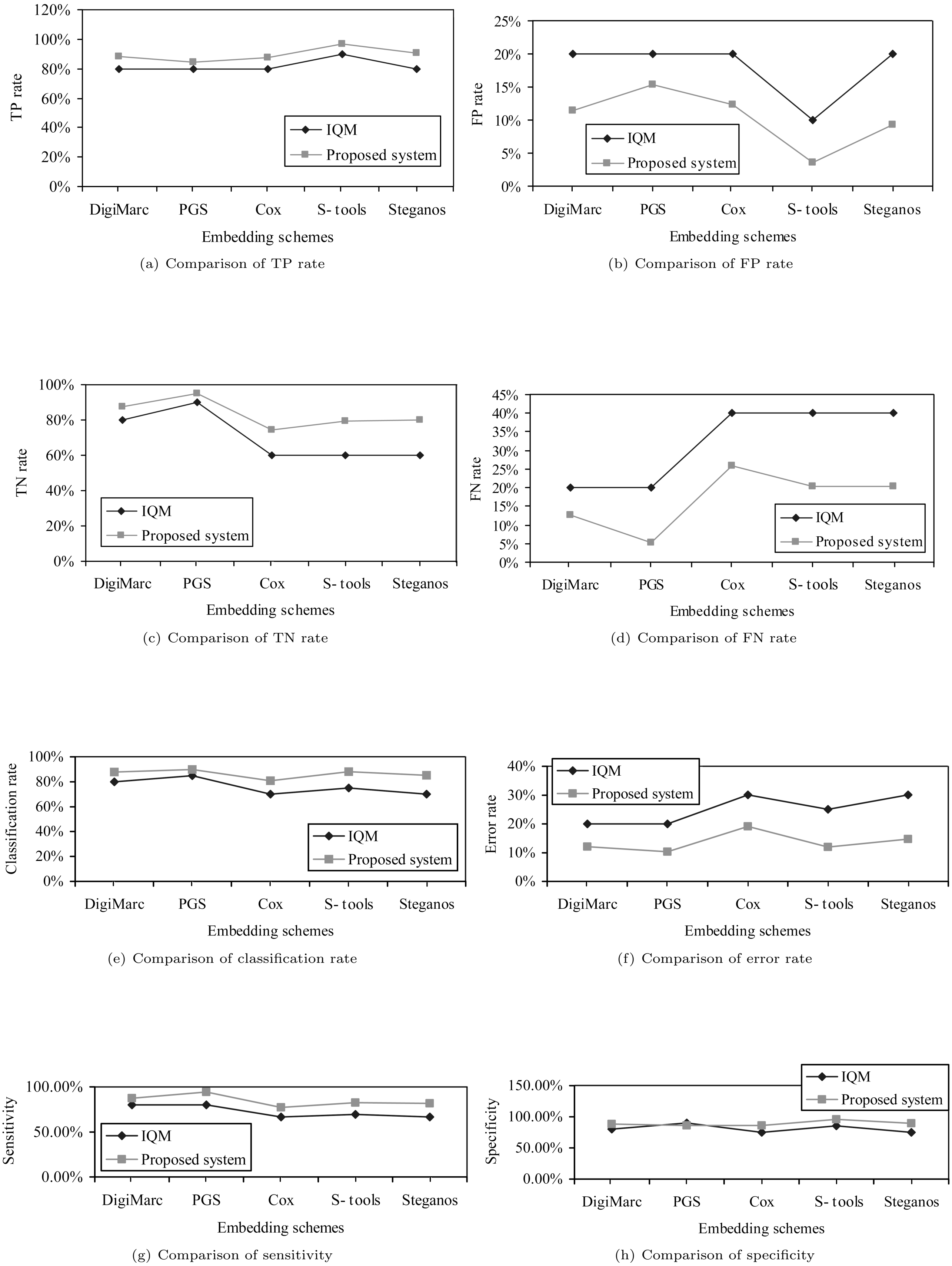
Differentiation categories		TP rate
Schemes	#1	88.50 %
	#2	84.63 %
	#3	87.73 %
	#4	96.50 %
	#5	90.71 %
	#6	85.17 %
Applications	Watermarking	86.95 %
	Steganography	90.79 %
Operation domain	Spatial	90.61 %
	DCT	87.13 %
	DWT	89.33 %
Type of processed non-stego images	Original	79.20 %
	JPEG-compressed	85.30 %
	Smoothed	79.50 %
	Sharpened	53.00 %

4.6 Application on a completely new steganography scheme

In order to show that the system is dynamic, i.e., adaptable to detect any new steganographic technique, the system was tested on scheme #7, which is based on the wavelet-domain techniques. It was found that the TP rate against scheme #7 is 88.23 %, as given in Table 3. This proves that the identified features are sensitive even to detect images generated by any new stego systems. To accommodate the identification of more hiding schemes, other kinds of image features should be explored further. The system is able to achieve a reasonably good TP rate of 89.23 % because the data hiding process done in the DWT domain leaves statistical artifacts in the higher order statistics of curvelet domain similar to other transform domain techniques. The system looks for these changes and thus is the competent of capturing these differences and classifying the images as stego-bearing or not.

Table 1 Performance comparison of the classifiers

Scheme	TP		TN		Classification rate (TP+TN)/2		Scheme	FP		FN		Error rate (FP+FN)/2	
	IQM	Proposed scheme	IQM	Proposed scheme	IQM	Proposed scheme		IQM	Proposed scheme	IQM	Proposed scheme	IQM	Proposed scheme
DigiMarc	80 %	88.50 %	80 %	87.32 %	80 %	87.91 %	DigiMarc	20 %	11.50 %	20 %	12.68 %	20 %	12.09 %
PGS	80 %	84.63 %	90 %	94.86 %	85 %	89.75 %	PGS	20 %	15.37 %	20 %	5.14 %	15 %	10.26 %
Cox	80 %	87.73 %	60 %	74.08 %	70 %	80.91 %	Cox	20 %	12.27 %	40 %	25.92 %	30 %	19.10 %
S-tools	90 %	96.50 %	60 %	79.65 %	75 %	88.08 %	S-tools	10 %	3.50 %	40 %	20.35 %	25 %	11.93 %
Steganos	80 %	90.71 %	60 %	79.80 %	70 %	85.26 %	Steganos	20 %	9.29 %	40 %	20.20 %	30 %	14.75 %
JSteg	70 %	85.17 %	70 %	78.54 %	70 %	81.86 %	JSteg	30 %	14.83 %	30 %	21.46 %	30 %	18.15 %



(a) Comparison of TP rate

(b) Comparison of FP rate

(c) Comparison of TN rate

(d) Comparison of FN rate

(e) Comparison of classification rate

(f) Comparison of error rate

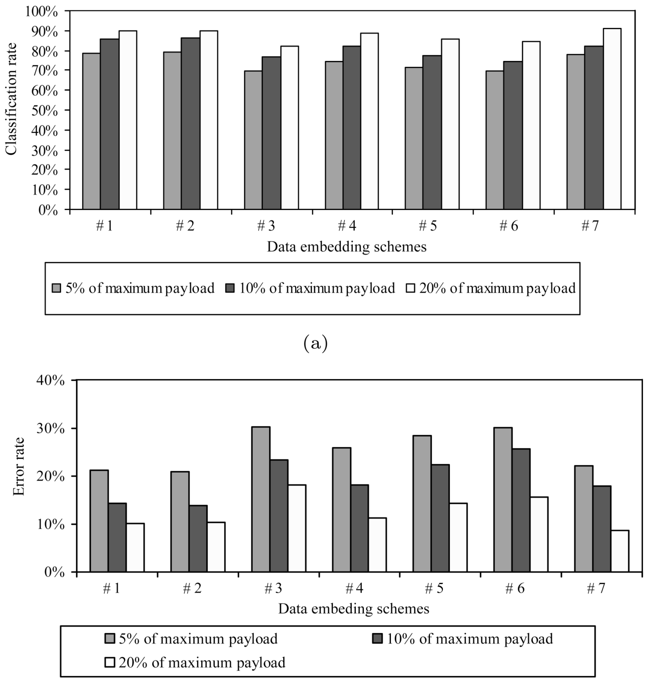
(g) Comparison of sensitivity

(h) Comparison of specificity

Fig. 3 Performance comparison curves

Table 4 Classification and error rates for test sets at various embedding rates

Schemes	Classification rate			Error rate		
	5 % of maximum payload	10 % of maximum payload	20 % of maximum payload	5 % of maximum payload	10 % of maximum payload	20 % of maximum payload
#1	78.80 %	85.66 %	89.87 %	21.20 %	14.34 %	10.13 %
#2	79.05 %	86.10 %	89.66 %	20.95 %	13.90 %	10.34 %
#3	69.71 %	76.66 %	81.85 %	30.29 %	23.34 %	18.15 %
#4	74.13 %	81.85 %	88.75 %	25.87 %	18.15 %	11.25 %
#5	71.60 %	77.66 %	85.66 %	28.40 %	22.34 %	14.34 %
#6	69.90 %	74.33 %	84.33 %	30.10 %	25.67 %	15.67 %
#7	77.85 %	82.05 %	91.33 %	22.15 %	17.95 %	8.67 %



(a)

(b)

Fig. 4 Influence of embedding rate on the performance of the steganalyser. (a) Classification rate at various embedding capacities; (b) Error rate at various embedding capacities

5 Discussion and conclusion

This paper presents a rationale for a blind image steganalytic model based on higher order statistics computed from curvelet coefficients. The feasibility of the proposed system is proved by systematic experiments. In our experiments, a database composed of processed plain images and stego images generated by using seven embedding schemes was utilized to evaluate the performance of our proposed features and classifier. A nonlinear classifier neuro-C4.5 that is easy to adapt to non-separable classes is adopted in our system. The average classification rate (85.62 %) for our proposed system is a promising discovery in blind steganalysis research. The future directions in this work can concentrate more on the other statistics from curvelet domain like higher order moments and apply this system to videos and compressed images. The performance of the system can also be improved by using appropriate fusion techniques in the machine learning component.

Acknowledgement

We sincerely thank the Management, Principal and Head of the Department of Information Technology of Thiagarajar College of Engineering, Madurai, India, for their support and encouragement. We sincerely thank the anonymous reviewers for the constructive comments which helped to improve the clarity and presentation of the paper.

References

[1] F. A. P. Petitcolas, R. J. Anderson, M. G. Kuhn. Information hiding — A survey. In *Proceedings of IEEE*, vol. 87, no. 7, pp. 1062–1078, 1999.

[2] S. Katzenbeisser, F. A. P. Petitcolas. *Information Hiding Techniques for Steganography and Digital Watermarking*, Norwood, MA, USA: Artech House, 2000.

[3] W. Bender, D. Gruhl, N. Morimot, A. Lu. Techniques for data hiding. *IBM Systems Journal*, vol. 35, no. 4, pp. 313–336, 1996.

[4] N. Nikolaidis, I. Pitas. Robust image watermarking in the spatial domain. *Signal Processing*, vol. 66, no. 3, pp. 385–403, 1998.

[5] W. N. Lie, L. C. Chang. Data hiding in images with adaptive numbers of least significant bits based on human visual system. In *Proceedings of IEEE International Conference on Image Processing*, IEEE, Kobe, Japan, pp. 286–290, 1999.

[6] Y. K. Lee, L. H. Chen. High capacity image steganographic model. *IEE Proceedings: Vision, Image and Signal Processing*, vol. 147, no. 3, pp. 288–294, 2000.

[7] W. N. Lie, G. S. Lin, C. L. Wu. Robust image watermarking on the DCT domain. In *Proceedings of IEEE International Symposium on Circuits and Systems*, Geneva, Switzerland, vol. 1, pp. 228–231, 2000.

[8] J. Huang, Y. Q. Shi. Adaptive image watermarking scheme based on visual masking. *Electronics Letters*, vol. 34, no. 8, pp. 748–750, 1998.

[9] T. Ogihara, D. Nakamura, N. Yokoya. Data embedding into pictorial with less distortion using discrete cosine transform. In *Proceedings of International Conference on Pattern Recognition*, Vienna, Austria, pp. 675–679, 1996.

[10] I. J. Cox, J. Kilian, F. T. Leighton, T. Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1673–1687, 1997.

[11] Q. Cheng, T. S. Huang. An additive approach to transform-domain information hiding and optimum detection structure. *IEEE Transactions on Multimedia*, vol. 3, no. 3, pp. 273–284, 2001.

[12] F Pérez-González, F. Balado, J. R. H. Martin. Performance analysis of existing and new methods for data hiding with known-host information in additive channels. *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 960–980, 2003.

[13] C. I. Podilchuk, W. Zeng. Image-adaptive watermarking using visual models. *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 4, pp. 525–539, 1998.

[14] Y. S. Kim, O. H. Kwon, R. H. Park. Wavelet based watermarking method for digital images using the human visual system. *Electronics Letters*, vol. 35, no. 6, pp. 466–468, 1999.

[15] X. Y. Wang, J. Wu. A feature-based robust digital image watermarking against desynchronization attacks. *International Journal of Automation and Computing*, vol. 4, no. 4, pp. 428–432, 2007.

[16] B. Xu, Z. B. Zhang, J. Z. Wang, X. Q. Liu. Improved BSS based schemes for Active steganalysis. In *Proceedings of ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel Distributed Computing*, IEEE, Qingdao, PRC, vol. 3, pp. 815–818, 2007.

[17] J. Fridrich. Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. In *Proceedings of International Workshop on Information Hiding, Lecture Notes in Computer Science*, Springer, vol. 3200, pp. 67–81, 2005.

[18] S. Geetha, S. S. Sivatha Sindhu, N. Kamaraj. Blind image steganalysis based on content independent statistical measures maximizing the specificity and sensitivity of the system. *Computers & Security*, vol. 28, no. 7, pp. 683–697, 2008.

[19] S. Geetha, S. S. Sivatha Sindhu, N. Kamaraj. Close color pair signature ensemble adaptive threshold based steganalysis for LSB embedding in digital images. *Transactions on Data Privacy*, vol. 1, no. 3, pp. 140–161, 2008.

[20] S. Geetha, S. S. S. Sindhu, N. Kamaraj. Steganalysis of LSB embedded images based on adaptive threshold close color pair signature. In *Proceedings of the 6th IEEE Indian Conference on Computer Vision, Graphics and Image Processing*, IEEE, pp. 281–288, 2008.

[21] S. Geetha, S. S. S. Sindhu, N. Kamaraj. StegoBreaker: Defeating the steganographic systems through genetic-X-means approach using image quality metrics. In *Proceedings of the 16th IEEE International Conference on Advanced Computing and Communication*, IEEE, pp. 382–391, 2008.

[22] S. Geetha, S. S. S. Sindhu, N. Kamaraj. StegoCracker: A genetic algorithm tuned neural network paradigm for breaking the audio steganographic utilities. In *Proceedings of IEEE Indicon*, pp. 180–186, 2007.

[23] J. Fridrich, M. Goljan. Practical steganalysis of digital images — State of the art. In *Proceedings of the SPIE International Conference on Security and Watermarking of Multimedia Contents*, San Jose, CA, USA, vol. 4675, pp. 1-13, 2002.

[24] C. Manikopoulos, Y. Q. Shi, S. Song, Z. Zhang, Z. Ni, D. Zou. Detection of block DCT-based steganography in gray-scale images. In *Proceedings of the 5th IEEE Workshop on Multimedia Signal Processing*, IEEE, pp. 355–358, 2002.

[25] R. Chandramouli. A mathematical approach to steganalysis. In *Proceedings of the SPIE International Conference on Security and Watermarking of Multimedia Contents*, San Jose, CA, USA, vol. 4675, pp. 14–25, 2002.

[26] J. J. Harmsen, W. A. Pearlman. Steganalysis of additive noise modelable information hiding. In *Proceedings of the SPIE*, vol. 5020, pp. 131–142, 2003.

[27] I. Avcibas, N. Memon, B. Sankur. Steganalysis using image quality metrics. *IEEE Transactions on Image Processing*, vol. 12, no. 2, pp. 221–229, 2003.

[28] W. N. Lie, G. S. Lin. A feature-based classification technique for blind image steganalysis. *IEEE Transactions on Multimedia*, vol. 7, no. 6, pp. 1007–1020, 2005.

[29] H. Farid. Detecting hidden messages using higher-order statistical models. In *Proceedings of International Conference on Image Processing*, Rochester, NY, USA, pp. 905–908, 2002.

[30] J. J. Harmsen. Steganalysis of Additive Noise Modelable Information Hiding, Master dissertation, Rensselaer Polytechnic Institute, Troy, New York, USA, 2003.

[31] T. Holotyak, J. Fridrich, S. Voloshynovskiy. Blind statistical steganalysis of additive steganography using wavelet higher order statistics. *Lecture Notes in Computer Science*, Springer, pp. 273–274, 2005.

[32] Y. Q. Shi, G. Xuan, C. Yang, J. Gao, Z. Zhang, P. Chai, D. Zou, C. Chen, W. Chen. Effective steganalysis based on statistical moments of wavelet characteristic function. In *Proceedings of IEEE International Conference on Information Technology: Coding and Computing*, IEEE, vol. 1, pp. 768–773, 2005.

[33] E. J. Candes, D. L. Donoho. New tight frames of curvelets and optimal representations of objects with C2 singularities. *Communications on Pure and Applied Mathematics*, vol. 57, no. 2, pp. 219–266, 2004.

[34] R. R. Coifman, D. L. Donoho. Translation-invariant denoising. *Lecture Notes in Statistics*, Springer, vol. 103, pp. 125–150, 1995.

[35] J. L. Starck, E. J. Candes, D. L. Donoho. The curvelet transform for image denoising. *IEEE Transactions on Image Processing*, vol. 11, no. 6, pp. 670–684, 2001.

[36] E. J. Candes, D. L. Donoho, C. R. A. Cohen, L. L. Schumaker. Curvelets — A surprisingly effective nonadaptive representation for objects with edges. *Curves Surfaces*, Nashville, TN, USA, pp. 105–120, 2000.

[37] N. Kingsbury, T. Reves. Redundant representation with complex wavelets: How to achieve sparsity. In *Proceedings of International Conference on Image Processing*, IEEE, Barcelona, Spain, vol. 1, pp. 45–48, 2003.

[38] Z. H. Zhou, J. Wu, W. Tang. Ensembling neural networks: Many could be better than all. *Artificial Intelligence*, vol. 137, no. 1–2, pp. 239–263, 2002.

[39] J. R. Quinlan. *C4.5: Programs for Machine Learning*, San Mateo, CA, USA: Morgan Kaufmann, 1993.

[40] S. S. S. Sindhu, S. Geetha, M. Marikannan, A. Kannan. A neuro-genetic based short-term forecasting framework for network intrusion prediction system. *International Journal of Automation and Computing*, vol. 6, no. 4, pp. 406–414, 2009.

[41] Z. H. Zhou, Z. Q. Chen. Hybrid decision tree. *Knowledge-based Systems*, vol. 15, no. 8, pp. 515–528, 2002.

[42] Z. H. Zhou, Y. Jiang. NeC4.5: Neural ensemble based C4.5. *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 6, pp. 770–773, 2004.

[43] B. Efron, R. Tibshirani, R. J. Tibshirani. *An Introduction to the Bootstrap*, New York, USA: Chapman & Hall, 1993.

[44] L. Breiman. Bagging predictors. *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[45] PictureMarc, Embed Watermark, v 1.00.45, Digimarc Corp.

[46] M. Kutterand, F. Jordan. JK-PGS (Pretty Good Signature), Signal Processing Laboratory at Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, [Online], Available: http://ltswww.epfl.ch/˜kutter/watermarking/JK_PGS.html, 1998.

[47] I. J. Cox, J. Kilian, F. T. Leighton, T. Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1673–1687, 1997.

[48] A. Brown. S-tools Version 4.0, [Online], Available: http://members.tripod.com/steganography/stego/s-tools4. html.

[49] *Steganos Security Suite*, [Online], Available: http://www.steganos.com/english/steganos/download.htm.

[50] J Korejwa. Shell 2.0, [Online], Available: http://www.tiac.net/users/korejwa/steg.htm.

[51] Images. [Online], Available: http://www.cl.cam.ac.uk/˜fapp2/watermarking/benchmark/image_database.html.

[52] J. H. Holland. *Adaptation in Natural and Artificial Systems*, University of Michigan Press, 1975.

[53] NeC45.zip, [Online], Available: http://lamda.nju.edu.cn/datacode/NeC4.5/nec45.zip.

S. Geetha received the B. Eng. and M. Eng. degrees in computer science and engineering in 2000 and 2004, respectively, from the Madurai Kamaraj University and Anna University of Chennai, India. In July 2004, she joined the Department of Information Technology at Thiagarajar College of Engineering, Madurai, India. She is a recipient of University Rank and Academic Topper Award in B. Eng. and M. Eng. in 2000 and 2004, respectively. She was an editor for *Proceedings the 1st Indian Conference on Computational Intelligence and Information Security*.

Her research interests include multimedia security, intrusion detection systems, machine learning paradigms, and information forensics.

E-mail: sgeetha@tce.edu (Corresponding author)

Siva S. Sivatha Sindhu received the B. Eng. and M. Eng. degrees in computer science and engineering from Maharaja Sayajirao University of Baroda and Anna University, India in 2002 and 2004, respectively. She is currently with the Department of Information Technology, Thiagarajar College of Engineering, Madurai, Tamilnadu, India.

Her research interests include information security, intrusion detection systems, and soft computing approaches.

E-mail: sivatha1277@yahpoo.com

N. Kamaraj received the B. Eng. degree in electrical and electronics engineering and the M. Eng. degree in power system engineering from Madurai Kamaraj University, India in 1988 and 1994, respectively. He received the Ph. D. degree in the power system security assessment in 2003 from Madurai Kamaraj University. He is an associate professor in Electrical Engineering Department, Thiagarajar College of Engineering, Madurai, Tamilnadu, India. Currently, he is heading the Department of Electrical and Electronics Engineering in Thiagarajar College of Engineering. He is the recipient of Merit Award from IEEE Computer Society for Computer Society International Design Competition (CSIDC) 2003 as best advisor for the team contested in CSIDC. Also, he has received Gold Medal and Corps Subject Award from Institution of Engineers (India) for the year 2003.

His research interests include security assessment using neural network, fuzzy logic, and genetic algorithm.

E-mail: nkeee@tce.edu