Multimed Tools Appl (2016) 75:5999–6019
DOI 10.1007/s11042-015-2708-0

A steganalysis method in the DCT domain

Mao Jia-Fa[1] · **Niu Xin-Xin**[2] · **Xiao Gang**[1] ·
Sheng Wei-Guo[1] · **Zhang Na-Na**[3]

Received: 9 September 2014 / Revised: 15 May 2015 / Accepted: 20 May 2015 /
Published online: 14 June 2015
© Springer Science+Business Media New York 2015

Abstract In this paper, we propose a steganalysis method based on the principle of additive operator, which chooses non-zero AC coefficients as carriers, with secret information independent of the carrier information flow. In the proposed method, AC coefficient statistical and energy features are initially extracted and used to construct a 3D feature vector. By employing the principle of Fisher linear discriminate analysis, a flexible discriminate classifier suitable for the extracted features is designed to improve detection performance. We infer and confirm theory of change in the statistical and energetic characteristics of the AC coefficient before and after additive steganography. The effectiveness of the proposed method is proven by the experiments. Moreover, the proposed method consistently outperforms related methods.

Keywords Steganalysis · Fisher linear discriminate analysis · AC coefficient energy · AC coefficient statistical characteristic

✉ Mao Jia-Fa
maojia@zjut.edu.cn

Niu Xin-Xin
xxniu@bupt.edu.cn

Xiao Gang
xg@zjut.edu.cn

Sheng Wei-Guo
wsheng@zjut.edu.cn

Zhang Na-Na
nanazhang2004@163.com

[1] College of Computer Science and Technology, ZheJiang University of Technology, Liuhe Road, No 180, Hang Zhou, ZheJiang Province 310023, Peoples Republic of China

[2] Information Security Center, Beijing University of Posts and Telecommunications, BeiJing 100876, China

[3] Department of Information Technology, Shanghai Jianqiao University, ShangHai 201315, China

1 Introduction

With the rapid development in Internet technology, communication has become convenient and efficient. Meanwhile, the quantity of data transmitted through networks is increasing substantially [18]. Under such conditions, network information security is becoming increasingly important. To address this critical problem, steganography and steganalysis have recently become key research areas. Compared with the temporal/spatial domain steganographic algorithm, steganography in the DCT domain distributes its energy into the local pixels of an image, thus realizing invisibility. More importantly, this method does not only combine certain characteristics of the human perception system with steganographic algorithm, but also complies with international data compression standards to realize steganographic encoding in the compression domain. Therefore, steganography in the DCT domain is one of the most popular approaches among steganographers.

Numerous steganalytic methods have been proposed in the literature. Li et al. [16] proposed a steganalysis method called yet another steganography scheme (YASS). The success of YASS suggests that a properly selected SO-domain is beneficial for steganalysis. Liu et al. [21] proposed an improved approach for the steganalysis of JPEG images. They extracted a 3950-D feature vector and applied support vector machine to detect covert images. The method has exhibited good detection performance on several JPEG-based steganographic systems. However, the performance of this approach depends on the number of features, and a large number of features decrease detection efficiency. Chen et al., [4] proposed an Improved Kernel Linear Discriminate Analysis algorithm to analyze the distribution differences between cover images and stego-images in the reduced dimensional space. They observed that the hidden information, the information hidden in the cover images, of stego-images are clustered in a plane while all other information of cover images are scattered more evenly in the whole space and have no other clusters. Awrangjeb and Lu [1] proposed micro and macro calibration methods that detect hidden information by calibrating the local and global distribution of the DCT coefficients of the image. All these methods employ high-dimensional feature vectors to describe the difference between cover and covert images, thus significantly affecting their performance in engineering applications. Based on the DCT coefficient generalized Gaussian distribution (GGD) statistical model, Natarajan and Anitha [22] put forward Universal Steganalysis Using Contourlet Transform. Pevny and Fridrich [24] proposed a JPEG image steganalysis that combines the features of Markov and DCT. Shi et al. [25] designed a Markov process-based approach to apply JPEG steganography effectively.

This study analyses the basic principle of steganography. For additive steganography that employs non-zero AC coefficients as carriers and a secret information stream independent of the carrier information stream, we extract the statistical and energetic features of the AC coefficients as the steganalytic features, according to the change in these characteristics before and after steganography. Then, by conducting Fisher linear discriminate (FLD) analysis, we design a flexible discriminate classifier suitable for the extracted features to improve detection performance. After experimenting on various cover and covert images using different embedding rates with two typical kinds of steganography, our proposed detection method has been exhibited to be effective.

2 Basic principles of steganography in the DCT domain

Steganographic technology can achieve covert communication [6, 14]. Research has been widely conducted on carrier positions relative to the suitability for embedding information.

The research results show that in certain areas or bands of an image, embedded information is effective against several conventional signal processing and geometric attacks. However, embedding information in such areas or frequency bands can result in poor visual quality of the carrier. This finding shows the contradiction between robustness and imperceptibility.

According to References [5, 13, 14, 19, 20, 29, 32, 33], steganography in the DCT domain is often described by the following formulas:

$$S = X + \alpha W, \tag{1}$$

$$S = X*(1 + \alpha W), \tag{2}$$

$$S = X*\exp(\alpha W), \tag{3}$$

where X refers to the DCT coefficients that carry secret information. This term is also called the actual carrier data set, such as AC coefficients, except for zero or zero and one; α refers to the intensity factor used to control embedding intensity and S refers to the covert image DCT coefficient. Formulas (1) and (2) are actually equal, and α in Formula (1) is equivalent to αX in Formula (2). Considering that $\ln S$=$\ln X$+αW is obtained after taking the logarithm of Formula (3), we can conclude that Formula (3) is equivalent to Formula (1) in the logarithm coordinate.

A common steganography process in the DCT domain is shown in Fig. 1. The DCT coefficients of the cover image often adopt quantized DCT coefficients because unchastised DCT coefficients result in the quantization and elimination of secret information during the quantizing encoding process. In Fig. 1, the key selecting model is mainly employed to choose the DCT coefficient as the information carrier and to take down the position of the coefficient to prepare to extract information. After obtaining the covert DCT coefficient, such coefficient is encoded according to JPEG standards, and then, the covert image is saved.

3 Histogram model of still image DCT coefficients

In the field of steganalysis, numerous researchers have investigated the statistical distribution model of image DCT coefficients and have proposed a number of statistical models. For example, Reference [10] developed DCT coefficients for JPEG images following the Gaussian distribution model. Lie and Lin [17] proposed DCT coefficients following the Laplacian distribution model. Yang and Kot [30] and Fridrich [7] devised DCT coefficients following the GGD statistical model.

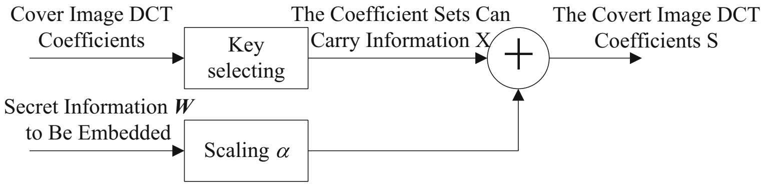
Cover Image DCT Coefficients
Key selecting
The Coefficient Sets Can Carry Information X
+
The Covert Image DCT Coefficients S
Secret Information W to Be Embedded
Scaling α

Fig. 1 Common steganography process in the DCT domain

The characteristics of the GGD model are determined by parameter c, sample mean value μ and standard deviation σ. The probability distribution function (PDF) is defined as

$$f_X(x) = A\exp(-\beta|x-\mu|^c), \tag{4}$$

where

$$\beta = \frac{1}{\sigma}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2}, \ A = \frac{\beta c}{2\Gamma(1/c)} \tag{5}$$

$\Gamma(\bullet)$ is a gamma function; and $\Gamma(1) = 1, \Gamma(1/2) = \sqrt{\pi}$. When c=1, the GGD model becomes a Laplacian distribution model, whereas when c=2, it becomes a Gaussian distribution model. Parameter c in GGD is flexible and changes with the actual data model.

A sample image with 256×256 pixels is shown in Fig. 2. An image for the DCT coefficient probability density (excluding DC coefficients in the DCT domain) is shown in Fig. 3. From this figure, we see that the distribution of DCT coefficients follows GGD.

4 Changes in the statistical characteristics of DCT coefficients after steganography and steganalysis feature selection

4.1 Related steganographic methods in the DCT domain

As shown in Section 1, steganography in the DCT domain has numerous advantages over steganography in the spatial domain. Thus, numerous existing steganographic algorithms are usually based on the DCT domain. A common characteristic of these steganographic algorithms is that the carrier coefficient set is composed of AC coefficients, except for zero [8, 9, 11, 12, 23]. Actual embedding methods are often according to Formulas (1), (2) and (3), among which Formula (1) is the simplest, most convenient and most useful. Therefore, Formula (1) is used in many steganographic algorithms, such as least significant bit (LSB)

Fig. 2 A sample image

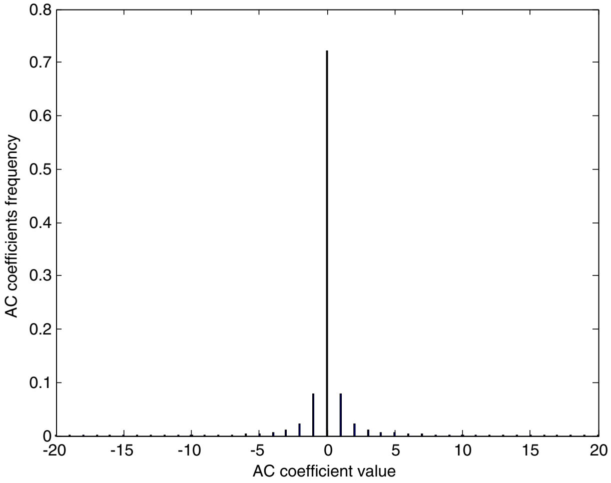

Fig. 3 The AC coefficient histogram of the sample image

and perturbed quantization (PQ) steganography. The embedding process can be described as follows.

1) Addition: Suppose that the secret information bit stream {0,1} is converted into a set $W=\{w|-1,1\}$. Embedding according to Formula (1) is called additive frequency band steganography in this paper. The PDF of the secret information bit stream is as follows:

$$p(w)=\begin{cases}1/2, w=-1\\1/2, w=1\end{cases}. \tag{6}$$

2) Substitution: Suppose that the secret information bit stream {0,1} is converted into a set $W=\{w|-1,0,1\}$. Embedding according to Formula (1) is called a substituting operator in this paper. Similar to LSB steganography, the PDF of the secret information bit stream is as follows:

$$p(w)=\begin{cases}1/4, w=-1\\1/2, w=0\\1/4, w=1\end{cases}. \tag{7}$$

These two steganographic methods are referred to as additive embedding operations in the present work.

4.2 Changes in histogram characteristics after image additive steganography

In this section, we introduce the definition of the embedding rate of the steganography operation and account for the changes in characteristics after image steganography.

Definition1. Suppose that the secret information flow is $W$and its length is *length (W)*. The actual carrier data set X is composed of the AC coefficients of JPEG images, except zero, and its length is *length (X)*. The bits per non-zero AC coefficient (bpc) of the embedding rate R is then defined as

$$R = \frac{length(W)}{length(X)}. \tag{8}$$

Suppose that a covert JPEG image I^s is obtained after the secret information is embedded into a cover image I^c, then the following changes happen to the statistical characteristics of the covert image.

THEOREM1. Suppose that the frequency of the AC coefficients of the cover image being zero is f_0^c, and that the frequency of the AC coefficients of the covert image being zero after additive embedding operation is f_0^s. When the mean value of f_0^s is larger than that of f_0^c, the following formula can be obtained:

$$E\left(f_0^s\right) > E\left(f_0^c\right). \tag{9}$$

PROOF. Suppose that the frequencies of the AC coefficients of a cover image being 1 and −1 are f_1^c and f_{-1}^c, respectively. Given that the secret information flow W is −1,1 or −1,0,1, according to additive steganographic principles, then AC coefficients with a value of 1 may changed its value to 0 or 2 and those with a value of −1 may changed its value to 0 or −2. From Formulas (6) and (7), a value of $p(-1)=p(1)=1/2$ *or* 1/4 can be obtained. Suppose that the embedding rate is R, then we have:

$$\begin{aligned} E\left(f_0^s\right) &= E\left(f_0^c + f_{-1}^c p(-1 \rightarrow 0)R + f_1^c p(1 \rightarrow 0)R\right) \\ &= E\left(f_0^c + f_{-1}^c p(1)R + f_1^c p(-1)R\right) \\ &= E\left(f_0^c\right) + E\left(f_{-1}^c + f_1^c\right)p(1)R \\ &\quad > E\left(f_0^c\right). \end{aligned}$$

THEOREM2. Suppose that the frequency of the AC coefficient absolute value of a cover image being 1 is $f_{|\eta|=1}^c$, and that of the AC coefficient absolute value of a covert image being 1 after additive embedding operation is $f_{|\eta|=1}^s$. Then, the following formula can be obtained:

$$E\left(f_{|\eta|=1}^s\right) < E\left(f_{|\eta|=1}^c\right). \tag{10}$$

PROOF. Given that $f_{|\eta|=1}^s = f_1^s + f_{-1}^s$ and $f_{|\eta|=1}^c = f_1^c + f_{-1}^c$, then $p(-1)=p(1)$ can be derived from Formulas (6) and (7). Suppose that the embedding rate is R, then:

$$\begin{aligned} f_1^s &= f_1^c - f_1^c p(1 \rightarrow 0)R - f_1^c p(1 \rightarrow 2)R + f_2^c p(2 \rightarrow 1)R \\ &= f_1^c - f_1^c p(-1)R - f_1^c p(1)R + f_2^c p(-1)R \quad , \\ &= f_1^c - \left(2f_1^c - f_2^c\right)p(1)R \end{aligned}$$

$$\begin{aligned} f_{-1}^s &= f_{-1}^c - f_{-1}^c p(-1 \rightarrow 0)R - f_{-1}^c p(-1 \rightarrow -2)R + f_{-2}^c p(-2 \rightarrow -1)R \\ &= f_{-1}^c - f_{-1}^c p(1)R - f_{-1}^c p(-1)R + f_{-2}^c p(1)R \quad . \\ &= f_{-1}^c - \left(2f_{-1}^c - f_{-2}^c\right)p(1)R \end{aligned}$$

Adding these values yields:

$$\begin{aligned} E\left(f^s_{|\eta|=1}\right) &= E\left(f^s_1+f^s_{-1}\right) = E\left(f^c_1+f^c_{-1}-\left(2\left(f^c_1+f^c_{-1}\right)-\left(f^c_2+f^c_{-2}\right)\right)p(1)R\right) \\ &= E\left(f^c_{|\eta|=1}\right)-\left(2E\left(f^c_{|\eta|=1}\right)-E\left(f^c_{|\eta|=2}\right)\right)p(1)R \\ &< \ E\left(f^c_{|\eta|=1}\right) \end{aligned}.$$

From Section 3, we can see that the AC coefficients follow zero-mean GGD. Then, $E(f^c_{|1|})>E(f^c_{|2|})$ and Formula (10) are derived.

THEOREM3. Let the maximum AC coefficient be max(*AC*) and the minimum AC coefficient be min(*AC*). To determine their minimum absolute value, γ=min(*abs*(max(*AC*)), *abs*(min(*AC*))) should be obtained. Suppose $\eta\in X$ and $\gamma>|\eta|>1$, then:

$$E\left(f^s_{\gamma>|\eta|>1}\right) > E\left(f^c_{\gamma>|\eta|>1}\right). \tag{11}$$

PROOF. Given that zero-mean GGD is symmetrical, $E(f^s_\eta)>E(f^c_\eta)$ is obtained when $\gamma>\eta>1$. Then, Formula (10) is also obtained. Suppose the embedding rate is R, then:

$$\begin{aligned} f^s_\eta &= f^c_\eta(1-p(\eta\rightarrow\eta-1)R-p(\eta\rightarrow\eta+1)R)+f^c_{\eta-1}p(\eta-1\rightarrow\eta)R+f^c_{\eta+1}p(\eta+1\rightarrow\eta)R \\ &= f^c_\eta(1-p(-1)R-p(1)R)+f^c_{\eta-1}p(1)R+f^c_{\eta+1}p(-1)R \\ &= f^c_\eta(1-2p(1)R)+\left(f^c_{\eta-1}+f^c_{\eta+1}\right)p(1)R \end{aligned}.$$

Taking the mean value, then:

$$\begin{aligned} E\left(f^s_\eta\right) &= E\left(f^c_\eta(1-2p(1)R)+\left(f^c_{\eta-1}+f^c_{\eta+1}\right)p(1)R\right) \\ &= E\left(f^c_\eta\right)+\left(\left(E\left(f^c_{\eta-1}-f^c_\eta\right)-E\left(f^c_\eta-f^c_{\eta+1}\right)\right)p(1)R\right. \end{aligned}.$$

Suppose that $\Delta_1=f^c_{\eta-1}-f^c_\eta$ and $\Delta_2=f^c_\eta-f^c_{\eta+1}$. Given that $\eta<\gamma$, $\eta+1\leq\gamma$, then we have:

$$E\left(f^s_\eta\right)=E\left(f^c_\eta\right)+(E(\Delta_1-\Delta_2))p(1)R.$$

The two sides of the GGD model show an exponential decline (Fig. 4). Thus, we obtain $\Delta_1>\Delta_2$. $(E(\Delta_1)-E(\Delta_2))p(1)R>0$ can also be obtained, and the preceding formula can be

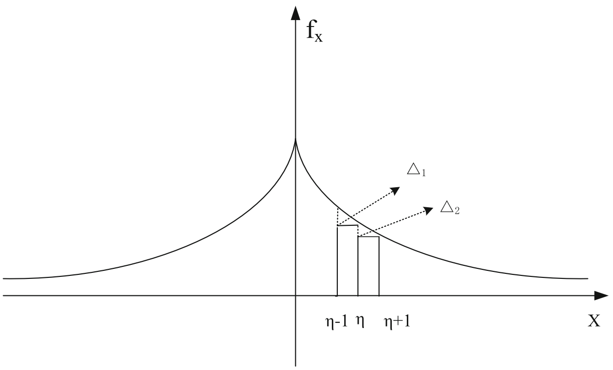

Fig. 4 The GGD model ($\Delta_1>\Delta_2$)

converted into $E(f_\eta^s)>E(f_\eta^c)$. Given that the zero-mean GGD model is symmetrical, then Formula (11) is derived.

Formula (11) shows that the frequency of AC coefficients increases if they have a value of more than 2 after the additive embedding operation, which is called the 'heavy tail' phenomenon [3]. The frequency of the AC coefficients after being embedded with the additive method at an embedding rate of 0.2 bpc in the sample image (Fig. 2) is shown in Fig. 4. Compared with the frequency of the AC coefficients of the cover sample (Fig. 3), the frequency of the AC coefficients is zero, as shown in Fig. 5, which is higher than that in Fig. 3. Although other changes are not evident, the tail is heavy after careful observation.

4.3 AC coefficient energy changes after image additive steganography

Definition2. Suppose that the AC coefficient flow of an image is $A=\{\eta_i|i=1,2,\cdots,length(A)\}$. The AC coefficient energy can be defined as:

$$En = E\left(A^2\right) = \frac{1}{length(A)}\sum_{i=1}^{length(A)}\eta_i^2, \tag{12}$$

Where *length*(*A*) refers to the length of the AC coefficients and *E*(•) refers to the mean operator. The carrier data set *X* is a subset of AC coefficient set A, that is, $X\subset A$. According to References [3, 12, 27, 30, 31], the carrier information is an original signal, whereas the secret information is a noise signal, and these two kinds of signals are independent of each other.

THEOREM 4. Suppose that En^c is the AC coefficient energy of a cover image, and En^s is that of a covert image after additive embedding operation, then:

$$En^s > En^c. \tag{13}$$

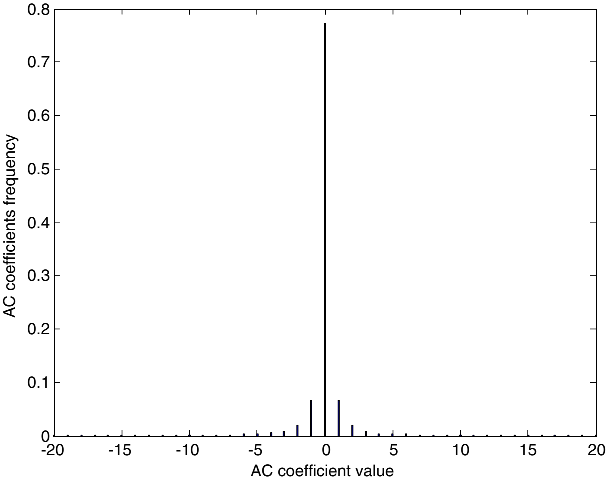

Fig. 5 Sample image of the distribution frequency of AC coefficients after additive embedding operation at an embedding rate of 0.2 bpc

PROOF. Suppose that the secret information flow is W. The AC coefficient energy of the covert image after additive embedding operation can be written as

$$\begin{aligned} En^s &= E\left((A+W)^2\right) = E\left(A^2+2AW+W^2\right) \\ &= E\left(A^2\right)+2E(AW)+E\left(W^2\right) \end{aligned}.$$

Given that the carrier information and the secret information are independent of each other, the following equation can be derived:

$$E(AW) = E(A)^*E(W).$$

The secret information flow consists of −1,1 or −1,0,1. By combining Formulas (6) and (7), $E(W)$=0 can be obtained, $E(W^2)$ is the mean energy value of the secret information flow. Evidently, $E(W^2)$>0, and the preceding formula can be transformed into:

$$En^s = E\left(A^2\right)+E\left(W^2\right) = En^c+E\left(W^2\right) > En^c.$$

THEOREM5. Suppose that the AC coefficient energies of a cover image and a corresponding covert image with absolute values of 0 and 1 are $En^c_{|\eta|\le 1}$ and $En^s_{|\eta|\le 1}$, respectively. Then, we have:

$$En^s_{|\eta|\le 1} < En^c_{|\eta|\le 1}. \tag{14}$$

PROOF. Suppose that the probability of the AC coefficient value of a cover image being a is f^c_η, and that of a covert image is f^s_η, then:

$$\begin{aligned} En^s_{|\eta|\le 1} &= E\left(\left(A^s_{|\eta|\le 1}\right)^2\right) = (-1)^2 f^s_{-1}+0^2 f^s_0+1^2 f^s_1 \\ &= f^s_{-1}+f^s_1 = f^s_{|\eta|=1} \end{aligned},$$

$$\begin{aligned} En^c_{|\eta|\le 1} &= E\left(\left(A^c_{|\eta|\le 1}\right)^2\right) = (-1)^2 f^c_{-1}+0^2 f^c_0+1^2 f^c_1 \\ &= f^c_{-1}+f^c_1 = f^c_{|\eta|=1} \end{aligned}.$$

Formula (14) can be obtained according to Theorem 2.

The AC coefficient energy of the covert image after additive embedding operation is larger than that of the cover image. However, the situation is reversed when the coefficient value is −1, 0, or 1. Therefore, the following conclusion can be inferred.

Inference: When the absolute value of the AC coefficient is greater than 1, the energy of the AC coefficient of a cover image is less than that of a covert image, which is:

$$En^s_{|\eta|>1} > En^c_{|\eta|>1}. \tag{15}$$

PROOF. Suppose that the energy of the AC coefficient absolute value of a cover image is greater than 1, that is, $En^c_{|\eta|>1}$, and the AC coefficient absolute value of a covert image is also greater than 1, that is, $En^s_{|\eta|>1}$, then:

$$\begin{aligned} En^c_{|\eta|>1} &= En^c - En^c_{|\eta|\le 1}, \\ En^s_{|\eta|>1} &= En^s - En^s_{|\eta|\le 1}. \end{aligned}$$

Formula (15) can be easily obtained according to Theorems 4 and 5.

Table 1 shows the AC coefficient energy of a sample image (Fig. 2), i.e., the AC coefficient energy with different embedding rates (bpc) after additive embedding operation. The AC coefficient energy increases when the general energy and absolute value are greater than 1 along with the increasing embedding rate. By contrast, when the absolute value is less than or equal to 1, the AC coefficient energy decreases along with the increasing embedding rate. This finding complies with our previous conclusion.

4.4 Feature selection of steganalysis

From the preceding discussion, we learn that AC coefficient characteristics change after additive embedding operation in the DCT domain. Some characteristic values increase, such as zero frequency, frequency with absolute value greater than 1 and AC coefficient energy. By contrast, other characteristic values are reduced, such as 1 and −1 frequencies and energy with an absolute value smaller than 1. Thus, the question is which characteristics should be selected and constructed as the steganalysis feature. This important issue will be addressed in the succeeding paragraphs.

The classifier of steganalysis is, in nature, a binary classifier that distinguishes between cover and covert images. According to the pattern recognition principle [2], if we aim to distinguish between two samples clearly, then the inner area of similar samples should be concentrated as much as possible. That is, the intra degree of scatter should be as small as possible, whereas the inter degree of scatter should be as large as possible. Based on this principle and on several other theorems obtained previously, we construct the following characteristics as our steganalysis features:

$$\text{Feature 1}: F(1) = E\left(f_0\right) \Big/ E\left(f_{|\eta|=1}\right), \tag{16}$$

$$\text{Feature 2}: F(2) = \left(\sum\nolimits_{|\eta|>1} E\left(f_\eta\right)\right) \Big/ E\left(f_{|\eta|=1}\right), \tag{17}$$

$$\text{Feature 3}: F(3) = En_{|\eta|>1} / En_{|\eta|\le 1}. \tag{18}$$

According to Theorems 1, 2, 3 and 5, and based on inference, the feature value of the covert image is larger than that of the cover image. That is:

$$F^s(i) > F^c(i), i = 1, 2, 3. \tag{19}$$

Thus, 500 covert images can be obtained from 500 cover images after additive spread spectrum operation at an embedding rate of 0.2 bpc. We select the 3D features of Formulas

Table 1 Ac coefficient energy of the sample image and different additive embedding rates

AC coefficient energy	Cover images	Covert (bpc)		
		0.1	0.2	0.3
En	2.4579	2.4581	2.4587	2.4595
$En_{\|\eta\|\le 1}$	0.1595	0.1576	0.1563	0.1550
$En_{\|\eta\|>1}$	2.2984	2.3005	2.3024	2.3045

(16), (17) and (18), as shown in Fig. 6, where '.' and '*' refer to the feature points of the cover and covert images, respectively. The feature value of the covert image is evidently larger than that of the cover image.

5 Design of the classifier

Selecting a feature is an essential step to detect information. A classifier suitable for features with good discrimination performance is required. After the features are selected, distinguishing whether such features include secret information is a problem of the binary classifier.

FLD analysis is a typical method employed in this type of work. We combine the single characteristic values of *F(1)*, *F(2)* and *F(3)* into a 3D feature vector x to obtain:

$$x = (F(1), F(2), F(3))^T. \tag{20}$$

According to FLD analysis [3], the best projection direction w should be:

$$w = S_w^{-1}(m^s - m^c), \tag{21}$$

where m^s and m^c in Formula (21) refer to the feature mean vector of the cover and covert images, respectively. S_w^{-1} refers to the inverse matrix of the intra matrix of scatter S_w:

$$S^i = \sum \left(x - m^i\right)\left(x - m^i\right)^T, i = s, c, \tag{22}$$

$$S_w = S^s + S^c, \tag{23}$$

$$y = w^T x \bigg| \begin{matrix} \geq \\ < \end{matrix} y_0 \rightarrow x \in \left\{ \begin{matrix} covert & image \\ cover & image \end{matrix} \right. . \tag{24}$$

Thus, according to Formula (24), any unknown sample x can be distinguished as a cover or covert image. y_0 in Formula (24) is the discrimination threshold.

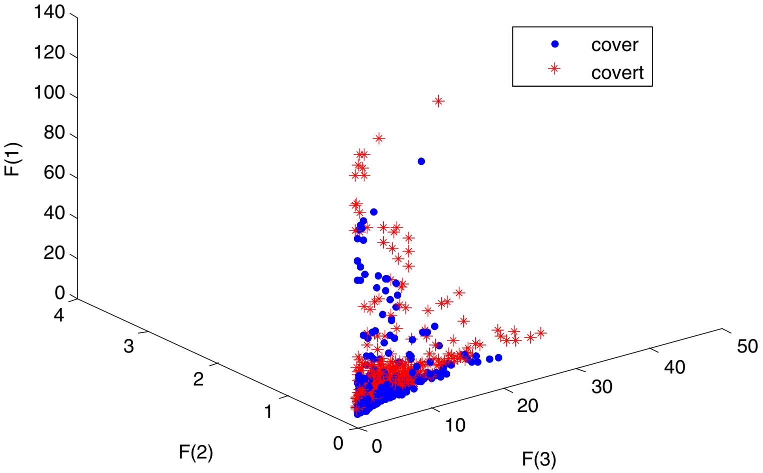

Fig. 6 Comparison between the feature values of 500 cover images and those of 500 covert images

6 Simulation experiment

6.1 Experimental setup

The experiment design is critical to evaluate a steganalysis method. A good experiment design should have the following key properties.

1) Generalization. The proposed image features and associated classifiers should be capable of identifying the presence of hidden data, which are possibly generated by various kinds of embedding methods, regardless of steganography.
2) Good performance. The classifier should, on one hand, have a detection rate of hidden data that is as high as possible and, on the other hand, keep false alarms as low as possible for cover images.
3) Robustness. The classifier should be capable of differentiating ordinary image-processing operations (such as smoothing, sharpening, recompression, rotation and cropping) from data embedding.

Starting from the aforementioned considerations, we choose four steganographic methods, PQ [9], Jsteg [15], and EBS [28] with representative steganography for the experimental evaluation of our additive embedding operation. PQ, Jsteg, and EBS are representative methods for adding and substituting embedding operators, respectively.

PQ: Fridrich [9] designed a method called PQ, which uses the wet paper code to develop a steganographic methodology for digital media. PQ embeds a secret message while downgrading the cover object by using information-reducing operations that involve quantization. PQ uses the knowledge of the unprocessed object and embeds data into elements with the most uncertain values after processing.

Jsteg: Jsteg [15] can use the LSB of the quantized DCT coefficients as redundant bits in which to embed the hidden message. The modification of a single DCT coefficient affects all 64 image pixels. The steganographic systems that modify LSBs of these image formats are often susceptible to visual attacks. The modifications are in the frequency domain instead of in the spatial domain, such that no visual attack occurs against JPEG image format.

EBS: Wang [28] presents an efficient JPEG steganography scheme based on the block entropy of OCT coefficients and syndrome trellis coding (STC). The proposed cost function explores both the block complexity and distortion effects due to flipping and rounding errors. The STC provides multiple solutions to embed messages to a block of coefficients.

When more information is embedded into an image, such image is easy to detect. By contrast, when less information is embedded, the image is difficult to distinguish from the cover image. Therefore, embedding rate is also an important factor that affects correct discrimination.

Our experimental data consist of two parts: JPEG cover images and covert images. We have 2000 cover images, including classic images widely used in the existing literature, such as Lena and Cameraman, digital camera images, and images downloaded from an online image library (http://sipi.usc.edu/database/database.cgi?volume=textures). All images are transformed into 256×256 JPEG grey images. These images are called cover images in this paper, and half of the images (i.e., 1000) are used for training and the other half are used for testing.

Moreover, PQ and Jsteg are employed. The covert image library is obtained after additive embedding operations at different embedding rates are conducted. We adopt five processing operations: 3×3 smoothing (averaging) spatial filter, 3×3 Laplacian sharpening spatial filter, recompression with 75, 65, and 50 quality factors, rotation of 25° anticlockwise and cropping to 1/4 of the original image to obtain 100 processed images (shown in Table 2).

6.2 Experimental results

In the present experiment, the DCT coefficients of the cover and covert images in the training library are chosen, and their feature values comprise the 3D cover and covert feature vector library according to Formulas (16), (17) and (18). These coefficients will be used as training data in FLD analysis to determine the projection direction $w = (\theta_1, \theta_2, \theta_3)^T$ and discrimination threshold y_0. FLD analysis is designed. The process of distinguishing an image is shown in Fig. 7.

When determining the projection direction, we input the feature vector set of the training images into the Fisher discriminate, that is, Formulas (21), (22) and (23), to obtain θ_1=31.3024, θ_2=0.1335 and θ_3=3.3993. We identify the threshold for determination. As we randomly select images, the feature values y follow GGD according to the central limit theorem in probability and statistics. For two class samples, suppose y of the cover and covert images follows GGD with a mean value of μ^i and a standard deviation of σ^i, $i = c$ or $i = s$. As shown in Fig. 8, the selection of the discrimination threshold y_0 is important for excellent detection effects. In general, false positive and false negative probability contradicts each other. When y_0 changes from large to small, the false negative probability decreases, whereas the false positive probability increases. When y_0 changes from small to large, the false positive probability decreases, whereas the false negative probability increases. Given that steganalysis is concerned with information security, we aim for a small false negative probability. Therefore, we design a flexible method for threshold selection, that is:

$$y_0 = \mu^s - \lambda\sigma^s, \tag{25}$$

where μ^s and σ^s in Formula (25) refer to the mean value and standard deviation of the projection feature value y_0 of the covert image, respectively. λ is the regulatory factor called the false negative probability controlling factor. The mean value and standard deviation of the training sample projection feature in this paper are: μ^c=347.341, μ^s=459.1301, σ^c=45.6131 and σ^s=64.2332. Table 3 shows the testing results with varying regulatory factors λ. When the projection feature probability of the cover image is the same as that of the covert image, that is, when $p(y^c)=p(y^s)$, the false positive and false negative probability attain balance, i.e., y_0=395.4431 and λ=0.9915.

Table 2 Composition of the image library

App.	Ori.	Smo.	Sha.	Rec.			PQ (bpc)			Jsteg (bpc)			UNIWADD	EBS
				75	65	50	0.1	0.2	0.3	0.1	0.2	0.3	0.2	0.2
Training	500	0	0	0	0	0	500	500	500	500	500	500	0	0
Testing	500	100	100	100	100	100	500	500	500	500	500	500	500	500

Fig. 7 The detection process

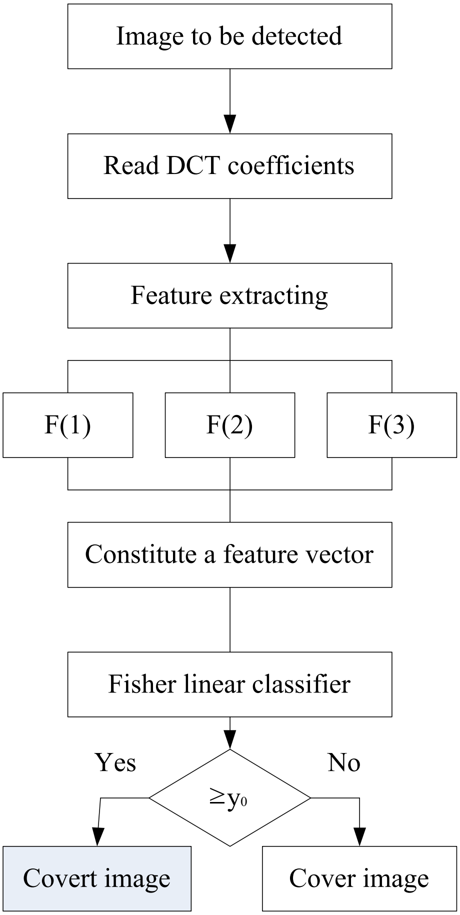

Table 3 indicates that a high embedding rate makes an image easy to detect. Moreover, detection and false positive probability increase simultaneously with increasing regulatory factor λ. We should control the regulatory factor according to practice requirements. To detect processed images, the images produced after sharpening the spatial filter exhibit the best results because non-zero AC coefficients are increased, which is contrary to that of the proposed method wherein zero AC coefficients increase. Although the detection results after smoothing the spatial filter or recompression are not as good as the original results, these two kinds of processed images increase the number of zero AC coefficients. However, the detection results of geometric attacks, such as rotation and cropping, is the same as that in the original images, thus indicating that the number of non-zero AC coefficients slightly change under a fixed image quality factor.

The experimental results show that higher compression quality factor leads to lower false positive probability while lower quality factor results in higher false positive probability. This is mainly due to when the quality factor is lowed, the compression rate will become higher, and at this time more zero coefficients will appear, which could easily lead to incorrect judgment.

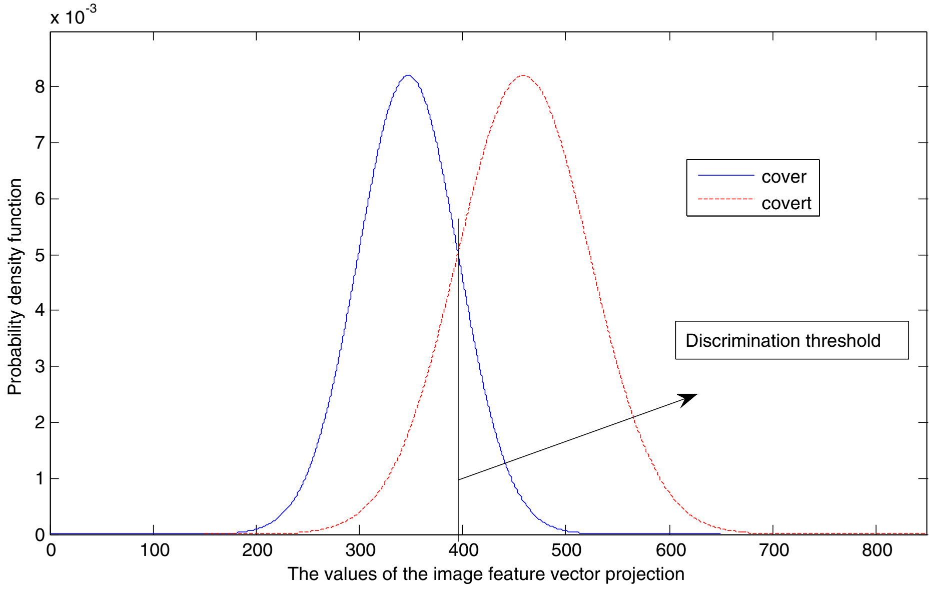

Fig. 8 FLD analysis

6.3 Detection performance analysis

There are two approaches [2, 26] to measure the performance of detecting steganography: the first one is the ROC curve while the other one is the minimal total detection error [28]. The

Table 3 Detection results

Covert images		The positive detection probability (%)				
		λ=0.8	λ=0.9	λ=0.9915	λ=1.0	λ=1.1
PQ	0.1	82.2	84.6	85.4	85.4	87.6
	0.2	82.8	84.8	85.8	85.8	88.0
	0.3	83.8	85.0	86.0	86.2	88.4
Jsteg	0.1	74.5	78.2	79.9	80.4	83.6
	0.2	79.3	82.6	84.2	84.7	86.9
	0.3	79.9	83.2	85.9	86.8	87.6
EBS	0.2	83.6	84.2	85.8	87.4	88.2
Average		80.9	83.1	84.6	85.3	87.1
Cover images		The false positive probability (%)				
		λ=0.8	λ=0.9	λ=0.9915	λ=1.0	λ=1.1
Orig.		10.4	12.2	13.8	14.0	16.6
Smoo.		10.0	11.0	12.0	14.0	16.0
Shar.		16.0	16.0	16.0	19.0	21.0
Reco.(75)		12.0	12.0	13.0	15.0	16.0
Reco.(65)		12.0	13.0	13.0	15.0	17.0
Reco.(50)		13.0	13.0	14.0	16.0	17.0
Average		10.46	11.46	12.45	13.55	15.46

ROC curve is a relation curve that describes false positive and positive detection probability. The minimal total detection error is calculated as follows:

$$P_E = \underset{P_{FP}}{\min}(P_{FP} + P_{MD}(P_{FA}))/2 \tag{26}$$

where P_{FP} and P_{MD} denote the probability of false positive and false negative misdetection, respectively. The false negative is also called misdetection. Several ROC curvesare needed to show the detection performance while one figure is enough by using minimal error. So, we use minimal total detection error to show the performance of detecting steganography.

The FLD method of the threshold is demonstrated and shown in Fig. 7. As shown in Fig. 8, this study analyses the detection system. Different P_{FP} and P_{MD} are obtained by changing the regulatory factor λ. In our work, we change λ from −4 to 4 with a step of 0.001. Therefore, a sequence of P_{FP} and P_{MD} are obtained. Thus, the minimal total detection error is calculated based on Eq. (26), as shown in Fig. 9.

We calculate minimal detection error of the PQ(0.1, 0.2,and 0.3), Jsteg(0.1, 0.2, and 0.3), NUIWARD(0.2), EBS(0.2), as well as the value of the regulatory factor λ when the detection error minimum. This figure clearly shows that our minimal detection error is lower than 18 %, thus our proposed method has good detection performance for additive steganography.

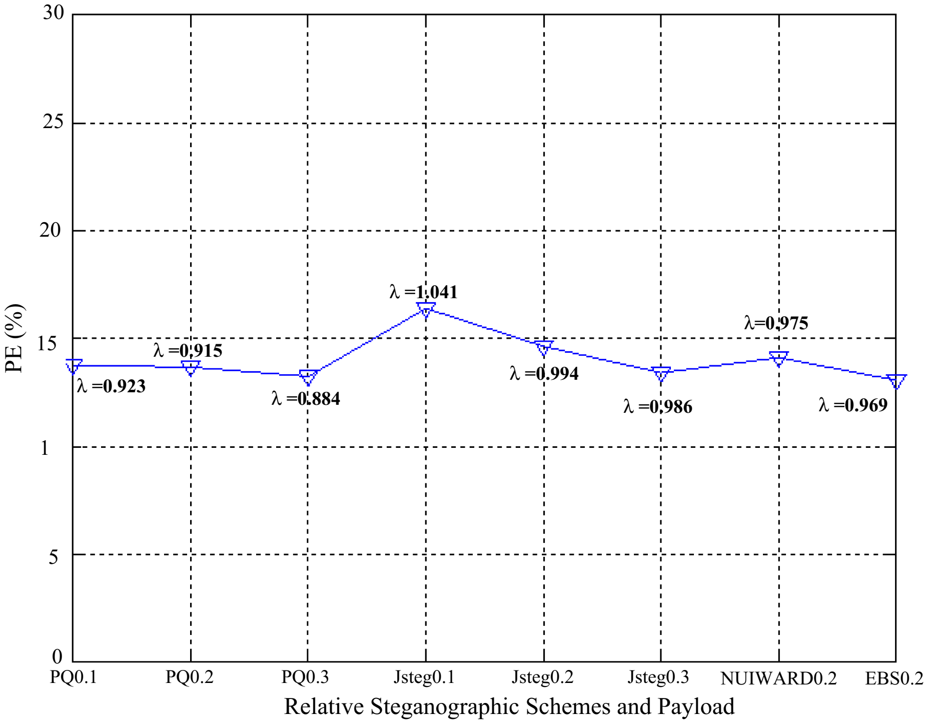

Fig. 9 The minimal total detection error of our system

Table 4 Summary of previous works and our proposed scheme

Works	NF	ER	APDR	AFPR	Classifier
[16]	140	0.03–0.18 (bpc)	NS	NS	URP
[21]	3950	0.29–0.35 (bpc)	78.37 %	18.54 %	SVM
[4]	150	0.1–0.3 (bpp)	NS	NS	KLDA
[24]	324	25–100 %	NS	NS	SVM
[25]	274	0.05 (bpc)	URP	URP	SVM
[17]	42	0.01–2.66 bpp	90.28 %	29.44 %	Neural
[12]	36	25–50 %	81.07 %	14.16 %	GMM
[27]	156	−4–11 dB	URP	URP	FLD
This Paper	3	0.1–0.3 (bpc)	84.4 %	13.75 %	FLD

NF The Number of Features, *ER* The Embedding Rate, *APDR* The Average Positive Detection Rate, *AFPR* The Average False Positive Rate, *GMM* The Gaussian Mixture Model, *SVM* The Support Vector Machine, *URP* Unreported, *NS* No Statistics, *KLDA* Kernel Linear Discriminate Analysis, *FLD* Fisher Linear Discriminator

Table 4 shows the results of the comparison of our detection method with a previous related work. The advantages of our detection method are summarized as follows.

(1) In terms of feature extraction, existing methods employed high dimensional features (the number of features range from 36 dimensional features [12] to 3950 dimensional features [21]) while our method use only 3 features. Therefore, our method of feature extraction is significantly simpler than existing methods.

(2) Although the PD in the reference [10] is slightly higher than that of our method, our approach achieves a significantly lower FP. Further, our minimal total detection error has much better performance, specifically, 10 % lower than the results reported in [10].

(3) Our system design is more reasonable compared with earlier work. This is due to we use the regulatory factor λ to control the false positive and false negative probability, while exiting methods do not use the regulatory factor at all.

(4) We infer and prove the feature changes of images before and after the embedding operation. Compared with previous work, in our paper we give a detailed theoretical derivation, which could help the readers to understand our work.

7 Conclusions

Additive embedding operation has been widely applied to information hiding during the last decade. For example, technologies such as PQ, Jsteg, and EBS are all based on additive embedding operation in the DCT domain. This study proposes a special steganalysis method by choosing several addition steganography or substitution steganography with non-zero AC coefficients as carrier and secret information, which are independent of carrier information flow. The statistical characteristics of AC coefficients change after steganography. We first select the statistical and energy characteristics of AC coefficients to construct a 3D feature vector. Then, by conducting FLD analysis, we design a flexible classifier with matching particular features, thus significantly improving detection performance. After evaluating cover and covert images with different embedding rates by using two typical kinds of steganography, the proposed method is found to be effective.

Acknowledgments The authors thank Jiangqun Ni research team in Sun Yat-sen University, China, they provide EBS steganography tool software. This work is supported by the National Natural Science Foundation of China (No. 61170271, 61272310, 61203288) and the ZheJiang province Natural Science Foundation of China (No. LY15F020032,LY12F02031).

References

1. Awrangjeb M, Lu G (2008) A robust content-based watermarking technique. MMSP 2008,MMSP 2008, Cairns, Queensland, Australia, pp. 713–718
2. Bian ZQ, Zhang XG (2005) Pattern recognition. Tsinghua University Press, BeiJing, pp 87–90
3. Briassouli A, Tsakslides P, Stouraitis A (2007) Hidden messages in heavy-tails: DCT-domain watermark detection using alpha-stable models. IEEE Trans Multimedia 7(3):700–715
4. Chen GM, Chen Q, Zhang D, Zhou DN (2014) Steganalysis based on distribution characters of stego-images in reduced dimension space. Multimedia Tools Appl 71(2):497–515
5. Cheng Q, Huang TS, Leighton T, Shamoon T (2001) An additive approach to transform-domain information hiding and optimum detection structure. IEEE Trans Multimedia 3(3):273–284
6. Dai ZH and Qi X (2012) Research on the large scale image steganalysis technology based on cloud computing and BP neutral network. 2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 415–419
7. Fridrich J (2004) Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. Proc.of the 6th Information Hiding Workshop, Springer-Verlag Berlin Heidelberg, pp. 67–81
8. Fridrich J, Goljan M, Hogea D (2002) Setganalysis of JPEG image: breaking the F5 algorithm. Information Hiding 5th International Workshop, Netherlands, pp. 310–323
9. Fridrich J, Goljan M, Soukal D (2005) Perturbed quantization steganography. Multimedia System. Proc. of the 6th Information Hiding Workshop 11(2):98–107.
10. Giannlual A, Boulgouris NV, Hatzinakos D, Platanitis KN (2006) Watermark detection for noisy interpolated images. IEEE Trans Circuits Syst 53(5):359–403
11. Hernadez JR, Amado M, Gonzalez FP (2000) DCT-domain watermarking techniques for still images: detector performance analysis and a new structure. IEEE Trans Image Process 9(1):55–68
12. Hou XD, Tao Z, Xiong G (2014) A novel steganalysis framework of heterogeneous images based on GMM clustering. Signal Process Image Commun 29(3):143–154
13. Jafari R, Ziou D, Rashidi MM (2013) Increasing image compression rate using steganography. Expert Syst Appl 40(17)
14. Kazem Q, Reza S (2014) A new steganography method which preserves histogram: generalization of LSB++. Inf Sci 277:90–101
15. Lee YK, Hwang SY, Ou ZH (2006) A novel quantity based on clipping statistics for Jsteg steganalysis. 8th IASTED Int. Con. On Signal & Image Processing (SIP 2006), Honolulu, Hawaii, USA, pp. 14–16
16. Li B, Shi Y, Huang JW (2009) Steganalysis of YASS. IEEE Trans Inf Forensics Secur 4(3):369–382
17. Lie WN, Lin GS (2005) A Feature-based classification technique for blind image steganalysis. IEEE Trans Multimedia 7(6):1077–1020
18. LiFang Y, Yao Z, RongRong N (2014) A channel selection rule for YASS. Sci Chin Inf Sci 87(8):1–10
19. Lingyun X, Xingming S, Gang L (2014) Linguistic steganalysis using the features derived from synonym frequency. Multimedia Tools Appl 71(3)
20. Liu QZ, Cooper PA, Chen L (2013) Detection of JPEG double compression and identification of smartphone image source and post-capture manipulation. Appl Intell 39(4):705–726
21. Liu Q, Sung A, Qiao M, Chen Z, Ribeiro B (2010) An improved approach to steganalysis of JPEG images. Inf Sci 180(9):1643–1655
22. Natarajan V, Anitha R (2012) Universal steganalysis using contourlet transform. Adv Comput Sci Eng Appl AISC 167:727–735
23. Ogihara T, Nakamura D, Yokoya N (1996) Data embedding into pictorial with less distortion using discrete cosine transform. In Proc.ICPR'96, Vienna, Austria 1996, pp. 675–679
24. Pevny T, Fridrich J (2007) Merging Markov and DCT features for mutli-class JPEG Steganalysis. Proceedings of SPIE Electronimc Imaging, Secruity, Steganography, and Watermarking of Multimedia Contents IX. San Jose, CA, USA, 6505, pp. 650503-1–650503-13
25. Shi YQ, Chen C, Chen W (2006) A Markov process based approach to effective attacking JPEG steganography. Information Hiding 8th international workshop, Berlin, Germany: Springer Berlin, 4437, pp. 249–264
26. Swaminathan A, Wu MK, Liu JR (2008) Digital image forensics via intrinsic fingerprints. IEEE Trans Inf Forensics Secur 3(1):101–117

27. Wang Y, Moulin P (2007) Optimized feature extraction for learning-based image steganalysis. IEEE Trans Inf Forensics Secur 2(1):31–45
28. Wang C, Ni J (2012) An efficient JPEG steganographic scheme based on the blook entropy of DCT coefficients. Proceeding of IEEE ICASSP 2012, Kyoto, Japan, pp. 1785–1788
29. Wu M, Yu H, Lui B (2003) Data hiding in image and video: pat-designs and applications. IEEE Trans Image Process 12(6):696–705
30. Yang HJ, Kot A (2007) Pattern-based data hiding for binary image authenticationby connectivity-preserving. IEEE Trans Multimedia 9(3):475–486
31. Yang CH, Weng CY, Wang SJ, Sun HM (2008) Adative data hiding in edge of images with spatial LSB domain systems. IEEE Trans Inf Forensics Secur 3(3):488–497
32. Yih-Kai L (2014) A data hiding scheme based upon DCT coefficient modification. Comput Stand Interfaces 36(5):855–862
33. Zhan-He O, Ling-Hwei C (2014) A steganographic method based on tetris games. Inf Sci 276:343–353

Mao Jia-Fa received the Ph.D. degree in pattern recognition from East China University of Science and Technology, China, in 2009. Since then, he has worked as a Post-doc Researcher at Beijing University of Posts and Telecommunications, China. In July 2011, he joined Zhejiang University of Technology, China, where he is currently an Associate Professor in the School of Computer Science & Technology. His research interests include pattern recognition, digital image processing and information hiding. He has published over 30 papers in the scientific literature.

Niu Xin-Xin received the Ph.D. degree in signal and information processing for the Chinese University of Hong Kong, Hong Kong, China, in 1997. She joined Beijing University of Posts and Telecommunications, China, where she is currently a Professor in the School of Computer Science & Technology. Her research interests include information hiding, digital image processing, signal processing. She has published over 100 papers in the scientific literature.

Xiao Gang received the Master degree from Tsinghua University, China, in 1992 and the Ph.D. degree from Zhejiang University of Technology, China, in 2011. In July 1985, he joined Zhejiang University of Technology, China, where he is currently a Professor in the School of Computer Science & Technology. His research interests include digital image processing, water quality testing and computer-aided design.

Sheng Wei-Guo received the M.Sc. degree in information technology from the University of Nottingham, U.K., in 2002 and the Ph.D. degree in computer science from Brunel University, U.K., in 2005. Since then, he has worked as a Researcher at the University of Kent, U.K. and Royal Holloway, University of London, U.K. In March 2011 he moved to Zhejiang University of Technology, China, where he is now a Professor in Computer Science. His research interests include evolutionary computations, data mining/clustering, pattern recognition and machine learning.

Zhang Na-Na received the M.Sc. degree in computer application technology from the Shanghai Normal University, China, in 2006. Since then she joined Shanghai Jianqiao University, China, where she is currently an Associate Professor in School of Information Technology. She research interests include digital image processing, information hiding, and pattern recognition.