# Local Linear Transform and New Features of Histogram Characteristic Functions for Steganalysis of Least Significant Bit Matching Steganography

**Ergong Zheng, Xijian Ping and Tao Zhang**
Zhengzhou Information Science and Technology Institute
Zhengzhou, Henan 450002 - P. R. China
[e-mail: zeg_1980@tom.com, pingxijian@yahoo.com.cn, brunda@163.com]
*Corresponding author: Ergong Zheng

*Received September 28, 2010; revised November 15, 2010; revised February 15, 2011; accepted April 19, 2011; published April 29, 2011*

## *Abstract*

In the context of additive noise steganography model, we propose a method to detect least significant bit (LSB) matching steganography in grayscale images. Images are decomposed into detail sub-bands with local linear transform (LLT) masks which are sensitive to embedding. Novel normalized characteristic function features weighted by a bank of band-pass filters are extracted from the detail sub-bands. A suboptimal feature set is searched by using a threshold selection algorithm. Extensive experiments are performed on four diverse uncompressed image databases. In comparison with other well-known feature sets, the proposed feature set performs the best under most circumstances.

***Keywords:*** Steganalysis, LSB matching, local linear transform, characteristic function

This work was supported by the National Natural Science Foundation of China under grant No. 60970142

**DOI: 10.3837/tiis.2011.04.012**

## 1. Introduction

In recent years, information hiding has become an important subject in the field of multimedia information security. Steganography, as a major branch of information hiding, is the art of embedding secret message imperceptibly within innocuous-looking cover data (e.g, digital images in this paper) so as to hide the presence of communication. On the contrary, steganalysis aims to discover the existence of secret message and can also serve as an effective way to test the security of steganographic techniques. The contest between the two sides promotes the progress of information hiding techniques.

Least significant bit (LSB) replacement is the most simplistic form of steganography, which replaces the LSBs of selected cover pixels by the corresponding secret message bits. However, a major shortcoming of this steganographic scheme is the phenomenon of pairs of values. By exploring the statistical anomaly, some reported methods can not only detect the presence of secret message with high reliability, but also estimate the length of message accurately [1][2] [3]. LSB matching (also known as ±1 embedding) is a trivial modification of LSB replacement, which randomly adds or subtracts pixel values by one to match the LSBs with the secret message bits. Despite the slight modification, LSB matching is much harder to be detected, especially in the case of low embedding rate. Furthermore, the performances of most LSB matching steganalyzers vary considerably across different cover sources [4][5].

The state-of-the-art specific stganalytic schemes for LSB matching can be divided into two categories as follows.

(1) Steganalysis based on smoothing effect of histogram.

In the context of additive noise steganography model, Harmsen and Pearlman proved that the LSB steganographic methods were equivalent to a low-pass filtering of histograms that was quantified by a decrease in the center of mass of the histogram characteristic function (HCF-COM) [6]. Afterwards, Ker extended HCF-COM to the center of mass of the adjacency histogram characteristic function (AHCF-COM), and proposed the first reliable LSB matching detector by incorporating the calibration (downsample) technique with the HCF-COM or AHCF-COM [7]. Following Ker's work, Li et al. suggested calculating the calibrated HCF-COM on the difference image and the experimental results showed the new detector outperformed Ker's detector [8]. On the other hand, based on the fact that the local maxima of histogram would decrease and the local minima would increase after LSB matching embedding, Zhang et al. proposed a detector based on the statistics of the amplitude of local extrema in the graylevel histogram [9]. Thereafter, Cancelli et al. extended Zhang et al.'s work to the two-dimensional histogram (further called the ALE detector) and the results demonstrated significantly improved performance compared to the original one [10]. In [11], by observing the uncertainty of local extrema of histogram, Gao et al. suggested calculating the sum of the amplitude of each point in the histogram instead of the ALE and they demonstrated experimentally that the new steganalyzer outperformed the previous ones proposed in [9] [10]. In addition, based on statistical modeling of pixel difference distributions, a novel detector was proposed in our previous work [12] and achieved better performance than the steganalyzer proposed in [9]. Recently, Cai et al. proposed another novel steganalyzer based on renormalized histograms [13].

(2) Steganalysis based on spatial dependences of natural images.

Generally, natural images are regarded as regional stationary sources and have strong spatial dependences among adjacent pixels, bit-plances and image blocks. However, these

dependences are violated by steganographic embedding where stego noise is an independent and identically distributed (i.i.d) sequence independent of the cover image. Huang et al. proposed an algorithm to detect the statistical changes of those overlapping flat blocks with 3×3 pixels in the least two significant bit planes after re-embedding operations [14]. Liu et al. described a detector based on correlation intra- and inter- bit-planes, and experimental results indicated that the significance of features and the detection performance depended not only on the embedding rate, but also on the image complexity [15]. Wang et al. introduced a method on the basis of spacing statistics of short and repeated sub-sequences [16]. Recently, Penvy et al. presented an algorithm based on Markov transition probability matrix [17]. In addition, there were two other methods for estimating the embedding rate [18][19]. Unfortunately, they are only effective for decompressed JPEGs.

Besides those specific steganalyzers, there exists many universal/blind steganalytic algorithms, and most of them can be used to detect LSB matching steganography as well. Most universal methods extract features from the histograms of image pixels or wavelet coefficients and used tow kinds of statistical moments as features. The first is empirical probability density function (PDF) moments, for example, Farid [20], Holotyak et al. [21], and Goljan et al. [22]. The second is empirical characteristic function (CF) moments, for example, Xuan et al. [23] and Shi et al. [24]. Based on Farid's work and Xuan et al.'s work, Wang et al. proposed optimized steganalytic features and demonstrated both qualitatively and quantitatively that the CF moments were superior to the PDF moments under a reasonable additive embedding model [25]. Furthermore, Li et al. regarded the task of image steganalysis as a texture classification problem and extracted features from the normalized histograms of the local linear transform (LLT) coefficients of an image [26].

Although various steganalyzers have been presented, the detection of LSB matching steganography remains unresolved, especially for uncompressed grayscale images and low embedding rates. In this paper, motivated by [25] and [26], we propose an effective steganalytic method for detecting LSB matching steganography. Specifically, we decompose images into a group of detail coefficients using LLT instead of discrete wavelet transform (DWT). Then, based on the local stationary characteristics of natural images and our observation that the largest difference between the CFs of most cover images and stego images lies at mid-frequencies, we construct a bank of band-pass filters for calculating new weighted features of the CFs. Lastly, a threshold selection algorithm is applied to feature selection. The experiments, performed on uncompressed grayscale images from four diverse image databases, analyse the impact of databases on performance. Comparisons to prior art reveal that the proposed approach performs the best under most circumstances.

## 2. Proposed Method

### 2.1 Image Representation with Local Linear Transform

LLT is a general computational framework in which an image is convolved with a bank of relatively small size masks for all local neighborhoods in a sliding window fashion. Local discrete cosine transform (DCT) is one kind of such local linear transform by using the orthogonal masks derived from DCT. The local DCT masks act as spatial band-pass filters in a very much the same way as Gabor filters [27]. Thus, they can provide powerful insight into an image's spatial and frequency characteristics. In aforementioned algorithms, DWT is usually chosen to decompose the image. However, we decide to choose LLT over DWT for steganalysis of LSB matching. One reason for our choice is that the LLT could capture the

changes of the local stochastic textures introduced by data embedding, which had been validated by the experimental results in [26]. Another is that the downsample operation involved in DWT will decrease the statistical difference between cover images and stego images. For example, Ker proposed a kind of downsample technique to estimate the cover image [7]. The downsampled stego image can be regarded as the stego image of the downsampled cover image by LSB matching with a reduced embedding rate, thus the procedure of downsample can reduce the embedding noise [11].

In addition to six of the local 2D DCT masks and four second derivative masks suggested by Li et al. in [26], our LLT masks consist of another two Laplace masks ( $w_{11}$ and $w_{12}$ ) which are actually derived from second derivative mask. We prefer second derivative masks to first derivative masks because of their stronger response to image details such edges, lines, and isolated points. These masks are given as:

Six local 2D DCT masks:

$$w_1=\begin{pmatrix}1&-2&1\\1&-2&1\\1&-2&1\end{pmatrix}\quad w_2=\begin{pmatrix}1&1&1\\-2&-2&-2\\1&1&1\end{pmatrix}\quad w_3=\begin{pmatrix}1&0&-1\\0&0&0\\-1&0&1\end{pmatrix}$$

$$w_4=\begin{pmatrix}1&-2&1\\0&0&0\\-1&2&-1\end{pmatrix}\quad w_5=\begin{pmatrix}1&0&-1\\-2&0&2\\1&0&-1\end{pmatrix}\quad w_6=\begin{pmatrix}1&-2&1\\-2&4&-2\\1&-2&1\end{pmatrix}$$

Six second derivative masks:

$$w_7=\begin{pmatrix}0&0&0\\1&-2&1\\0&0&0\end{pmatrix}\quad w_8=\begin{pmatrix}0&1&0\\0&-2&0\\0&1&0\end{pmatrix}\quad w_9=\begin{pmatrix}0&0&1\\0&-2&0\\1&0&0\end{pmatrix}$$

$$w_{10}=\begin{pmatrix}1&0&0\\0&-2&0\\0&0&1\end{pmatrix}\quad w_{11}=\begin{pmatrix}0&1&0\\1&-4&1\\0&1&0\end{pmatrix}\quad w_{12}=\begin{pmatrix}1&1&1\\1&-8&1\\1&1&1\end{pmatrix}$$

Denote *I* as an image under analysis. The *i*-th detail subband $D_i$ is obtained by convolving *I* with $w_i$, i.e.,

$$D_i = I * w_i, \quad 1 \le i \le 12 \tag{1}$$

where, the convolution process consists simply of moving the center of the mask $w_i$ from point to point in the image *I*, and at each point, the response of the mask at that point is the sum of products of the mask coefficients and the correspoinding neighborhood pixels in the area spanned by the mask.

Totally, these 12 subbands constitute an image representation with local linear transform. Features, such as CF weighted features to be defined in Section 2.2, are extracted from each detail subband $D_i, 1 \le i \le 12$.

## 2.2 New Features of Histogram Characteristic Functions

First, we give the definition of the CF of random variable. Let $p(x)$ denote PDF of a random variable *X*, the characteristic function of PDF $p(x)$ is expressed as

$$\Phi_X(\omega) = \int_{-\infty}^{\infty} p(x) e^{j\omega x} dx, \quad \omega \in \mathbb{R} \tag{2}$$

where $j=\sqrt{-1}$, $\omega$ is the frequency of random variable $x$ and an arbitrary real number.

For an additive noise steganography model, it is well known that the CF amplitude will

decrease after embedding, i.e.

$$\left|\Phi_S(\omega)\right| \le \left|\Phi_C(\omega)\right|, \quad \forall \omega \in \mathbb{R} \tag{3}$$

where $|\Phi_C(\omega)|$ and $|\Phi_S(\omega)|$ are the amplitudes of CFs of cover image and stego image respectively.

In practice, in order to obtain the empirical CF from a detail subband $D_i$, we first estimate the PDF $p(x)$ from an M-bin histogram $\{h(m)\}_{m=0}^{M-1}$. Let $K = 2^{\lceil \log_2 M \rceil}$. Then, the $K$-point discrete CF $\{\Phi(k)\}_{k=0}^{K-1}$ is defined as

$$\Phi(k) = \sum_{m=0}^{M-1} h(m) \exp\left\{\frac{j2\pi mk}{K}\right\}, \quad 0 \le k \le K-1 \tag{4}$$

which is similar to $\Phi_X(\omega)$ defined in (3) and can be calculated by fast Fourier transform (FFT) algorithms.

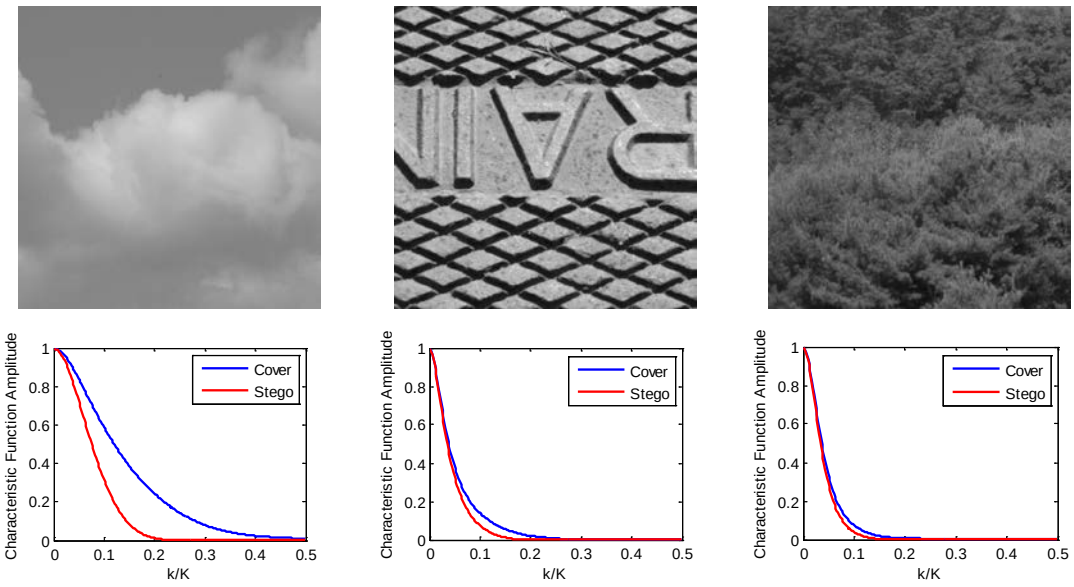


**Fig. 1**. Sample images and their corresponding CFs

**Fig. 1** shows three images and their corresponding CF amplitude plots of detail subband $D_7$, in which the blue and red lines represent the cover and stego counterparts, respectively. We can observe that the largest differences between the CFs of cover images and stego images lie at mid-frequencies, and the locations of the largest differences vary across images with diverse underlying image content. Furthermore, we also give the distributions of the locations of the largest differences between the CFs of $D_7$ for cover images and stego images on three self-built databases, as shown in **Fig. 2**, which validates our previous observation. Specifically, locations of the largest differences move to low-frequencies as the complexities of images increase. Intuitively, the extracted steganalytic features should capture the largest differences between the CFs of cover images and stego images.

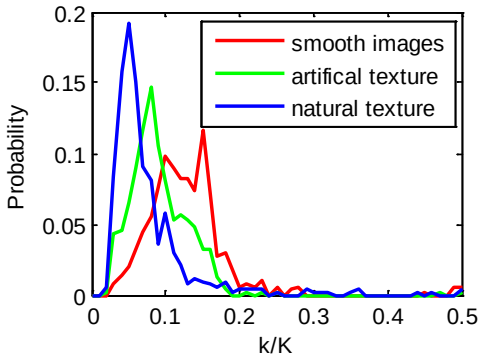


**Fig. 2**. The distributions of the locations of the largest differences between the CFs of cover images and stego images with 50% LSB matching embedding. The three databases consist of 500 smooth, artifical texture and natural texture images, respectively.

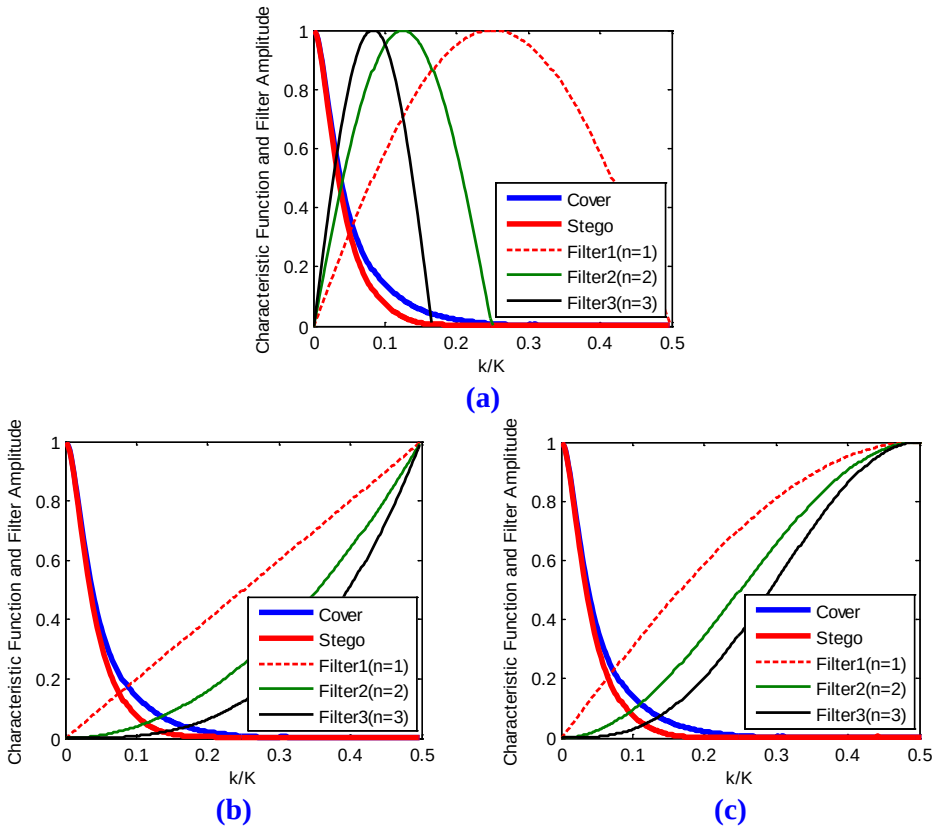


**Fig. 3**. Amplitude plots of characteristic functions and filters for calculating features: **(a)** CFs and our proposed new filters used in (5), **(b)** CFs and high-pass filters used in (6) and **(c)** CFs and high-pass filters used in (7)

On the basis of above observation, we propose a new kind of CF weighted features, the *n*-th CF weigthed feature is defined as

$$\hat{M}_n^* = \sum_{k=0}^{\lfloor K/2n \rfloor} \left|\Phi(k)\right| \sin\left(\frac{2n\pi k}{K}\right), \quad n \geq 1 \tag{5}$$

In the above equation, $|\Phi(k)|$ is weighted by $\sin\left(\frac{2n\pi k}{K}\right)$. Essentially, the features $\hat{M}_n^*$ are obtained by band-pass filtering the CFs, as shown in **Fig. 3-(a)**. These CF weighted features convey more information about the low-to mid-frequency components as the number $n$ increases. Hence, they can better capture the largest differences of CFs for images with diverse image content.

There are two another CF features presented in [23] and [25], known as CF moments, i.e.

$$\hat{M}_n^{'} = \sum_{k=0}^{K/2-1} |\Phi(k)| k^n, \quad n \geq 0 \tag{6}$$

$$\hat{M}_n^{A} = \sum_{k=0}^{K-1} |\Phi(k)| \sin^n\left(\frac{\pi k}{K}\right), \quad n \geq 0 \tag{7}$$

These features are obtained by essentially high-pass filtering the CFs, as shown in **Fig. 3-(b)** and **(c)** (Since $|\Phi(k)|$ has central symmetry structure, for the convenience of comparison, we only plotted the left half of $|\Phi(k)|$ in **Fig. 3-(c)**).

Moreover, we define the normalized CF weighted features as

$$\tilde{M}_n^* = \frac{\hat{M}_n^*}{\hat{M}_0^*}, \quad n \geq 1 \tag{8}$$

where $\hat{M}_0^* = \sum_{k=0}^{\lfloor K/2n \rfloor} |\Phi(k)|$. The normalization is equivalent to a self-calibration which reduces the dynamic range of CF weighted features and the overlap between the range of $\tilde{M}_n^*$ of cover images and that of stego images. A similar normalization was used in [25] and [23]

$$\tilde{M}_n^{A} = \frac{\hat{M}_n^{A}}{\hat{M}_0^{A}}, \quad n \geq 1 \tag{9}$$

$$\tilde{M}_n^{'} = \frac{\hat{M}_n^{'}}{\hat{M}_0^{'}}, \quad n \geq 1 \tag{10}$$

The advantage of $\tilde{M}_n^*$ over $\tilde{M}_n^A$ and $\tilde{M}_n^{'}$ will be evident in Section 2.3.

### 2.3 Feature Evaluation

To evaluate the usefulness of a feature in discriminating between classes, several criteria such as the Bhattacharyya distance may be used. The Bhattacharyya distance is defined as [25]

$$B(p_0, p_1) = -\log \int \sqrt{p_0(x) p_1(x)} \tag{11}$$

where $x$ is a feature, and $p_0(x)$ and $p_1(x)$ are the feature PDFs under Class 0 and Class 1, respectively. The larger the $B(p_0, p_1)$ is for a feature, the better that feature is for classification. When $p_0 = p_1, B(p_0, p_1) = 0$ and the feature is useless. In practice, $p_0(x)$ and $p_1(x)$ are often unavailable, and we estimate them using their histogram of training features and compute the empirical Bhattacharyya distance.

As shown in **Fig. 4**, we compare the empirical Bhattacharyya distance of features $\tilde{M}_n^*$ from (8), $\tilde{M}_n^A$ from (9), and $\tilde{M}_n^{'}$ from (10). The features are calculated from the subband $D_7$ of the LLT of 3164 cover images from CAMERA database [28] and their corresponding stego images with 50% LSB matching embedding. It is observed that the proposed CF weighted features $\tilde{M}_n^*$ are better features than the CF moment features $\tilde{M}_n^A$ and $\tilde{M}_n^{'}$ since the empirical

Bhattacharyya distance of $\tilde{M}_n^*$ is larger than that of the others. The above phenomena have been fairly consistently observed across all of the subbands $D_1 \sim D_{12}$.

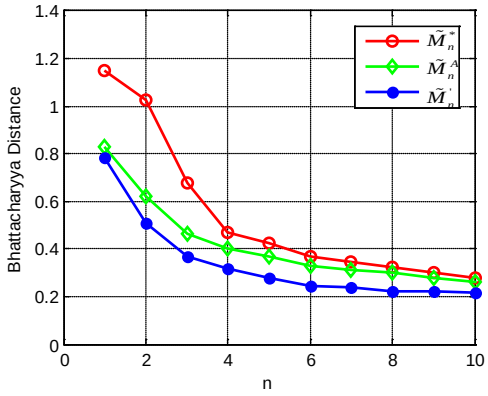


**Fig. 4**. Empirical Bhattacharyya distance for features $\tilde{M}_n^*$, $\tilde{M}_n^A$, and $\tilde{M}_n^{'}$, $1 \le n \le 10$. Data are gathered from the subband $D_7$ of the LLT of 3164 cover images, and their corresponding stego images with 50% LSB matching embedding.

## 2.4 Feature Selection

Given an image representation with LLT, we can calculate innumerous CF weighted features from each subband $D_i (1 \le i \le 12)$ by increasing the number *n*. Soppose we have a finite set of training samples and a total of *J* available features. The optimal feature set can be found by an exhaustive search over $2^J$ possibilities, which could achieve better performance at the cost of higher computational complextiy. To achieve a tradeoff between performance and computational complexity, we propose an improved threshold selection algorithm proposed by Wang et al. [25] to search a suboptimal feature set. The procedure of threshold selection algorithm is summarized as follows:

(1) Define a cost function *C*, and set $N = 1, C_0 = 0$;
(2) Divide the training set into 5 subsets of equal size;
(3) For each training image, extract $\tilde{M}_n^* (1 \le n \le N)$ from each subband $D_i$ as features. The feature set size is *lN* with *l* = 12 being the number of detail subbands;
(4) Each subset is tested using the classifier trianed on the remaining 4 subsets for 5-fold cross-validation and get the average value of *C*;
(5) If $C > C_0$, $C_0 = C$, $N = N + 1$, repeat steps (3) through (4); else $N_p = N - 1$, take the $lN_p$ features to form the feature set.

where, the AUC value is used as the cost function in [25]. In this paper, we define a more robust cost function as

$$C = \frac{AUC + TP\_E}{FP\_80 + FP\_50 + Min\_E} \tag{12}$$

where the abbreviative notations in the right half of the equation are five performance evaluation criteria defined as follows:

(1) AUC: area under the receiver operating characteristic (ROC) curve;
(2) TP_E: the true positive rate when the false positive rate equals to the false negative rate;

(3) FP_80: the false positive rate at the true positive rate of 80%;
(4) FP_50: the false positive rate at the true positive rate of 50%;
(5) Min_E: the minimal average probability of two errors (false positive rate and false negative rate).
where true positive rate represents the detection rate of stego images correctly classified as stego images; false positive rate represents the detection rate of cover images incorrectly classified as stego images; false negative rate represents the detection rate of stego images incorrectly classified as cover images, and the sum of true positive rate and false negative rate is 1.

The closer to 1 the first two measures are, the better the performance of the steganalyzer, and the closer to 0 the latter three measures are, the better the performance of the steganalyzer. Hence, the steganalyzer has better performance as $C$ increase. The computational complexity is $O(N)$ for the threshold selection algorithm, while it is $O(2^N)$ for an exhaustive search. The experimental results of feature selection will be reported in Section 3.2.

### 2.5 Procedure of Proposed Method

The whole procedure of proposed steganalysis algorithm is summarized as follows:
(1) For any given image $I$, compute the detail subbands $D_i(1 \le i \le 12)$ according to equation (1);
(2) For all images in the training set, extract CF weighted features $\tilde{M}_n^*, 1 \le n \le N$ from each subband according to equation (8), and get a collection of $12N$ features;
(3) Perform the threshold selection algorithm using the features of training set to search an suboptimal feature set with a size of $12N_p$;
(4) Train a classifier using the selected features of all images in trainng set;
(5) For all images in the test set, extract CF weighted features $\tilde{M}_n^*, 1 \le n \le N_p$ from each subband according to equation (8), and get a collection of $12N_p$ features;
(6) Classify images in test set using the trained classifier.

## 3. Experimental Results and Analysis

### 3.1 Experimental Setup

#### 3.1.1 Image Databases

It is well known that steganalysis of LSB matching is very sensitive to the type of cover images. To evaluate our proposed feature set and compare it to prior art under different conditions, the experiments were run on the following four uncompressed image databases.
(1) CAMERA contains 3164 TIFF images with the size of 512×512 [28].
(2) BOWS contains 10000 grayscale images with the size of 512×512 [29].
(3) UCID consists of 1338 TIFF images with the size of 384×512 or 512×384 [30].
(4) NRCS contains 3185 TIFF images with the size of 1500×2100 or 2100×1500 [31].
(5) HYBRID HYBRID contains images from all four databases, approximately 17700 images.

Where necessary, all images have been converted to grayscale, and in each database, two sets of stego images were created with the embedding rates of 50% and 25%.

It should be noted that these image datasets are very different: CAMERA images are central cropped images captured by 24 different digital cameras in the raw format, while BOWS images are rescaled and cropped natural images of various sizes. UCID images are small and

undergone downsampling operation, while NRCS images, coming from raw scans of film, are large and very noisy.

#### 3.1.2 Steganalytic Algorithms for Comparison

We compare our proposed method with the following four steganalytic algorithms.

(1) SPAM. Penvy et al. extracted a 686-dimensional feature set from the second-order subtractive pixel adjacency model with T=3 [17], and demonstrated that it outperformed the previous detectors in [10] and [22].

(2) RDIH. Cai et al. extracted a 70-demensional feature set from the peak-values and the renormalized histograms of difference images [13].

(3) OPTCF. In [25], Wang et al. extracted both PDF moment and CF moment features from wavelet subbands and prediction error subbands, and demonstrated that this algorithm was superior to the previous detectors in [20] and [23]. In addition, the threshold selection algorithm is also applied to feature selection in our experiments.

(4) LLTPDF. Li et al. extracted a 110-dimensional feature set from the PDFs of the LLT coefficients [26].

#### 3.1.3 Classifier and Evaluation Criteria

Fisher Linear Discriminant (FLD) is applied for all feature sets except SPAM which is linear non-separable from our experimental results. Instead, support vector machine (SVM) [32] is applied for SPAM feature set. All the features are linearly scaled before applying classifier. For a feature $f$, we find its maximum value $f_{\max}$ and its minimum value $f_{\min}$ from all the training images. For any training or test images, $f$ will be scaled as follows:

$$\tilde{f} = \frac{f - f_{\min}}{f_{\max} - f_{\min}} \tag{13}$$

For all the training images, $\tilde{f} \in [0,1]$; for most test images, $\tilde{f} \in [0,1]$. The main advantage of this scaling step is to avoid features with large numerical ranges dominate those in smaller numerical ranges. Another advantage is to avoid numerical difficulties during the calculation [32]. Five measures (AUC, TP_E, FP_80, FP_50 and Min_E) defined in Section 2.4 are used to evaluate the steganalyzer performance on the test set.

### 3.2 Results of Feature Selection

**Fig. 5** illustrates the threshold selection procedure described in Section 2.4 for a training subset from CAMERA database: 633 cover images and their stego counterparts with an embedding rate of 50%. The features are $\tilde{M}_n^*$ from $D_i$, $1 \le i \le 12$. We set $N = 20$. The feature set size is $lN = 12 \times 20 = 240$. We observe that the steganalysis performance measured by $C$ improves when $N$ increases, peaks at $N_p$ = 6. In this way, the threshold selection algorithm identifies $N_p$ and forms a steganalytic feature set that consists of those $12N_p$ = 72 features. The above similar peaking effect can be observed for other databases. Hence, we can use the threshold selection algorithm to search the suboptimal feature sets for other databases. Our experimental results indicate that the same or approximate $N_p$ is searched for a specific database under different embedding rates. The values of $N_p$ are 6, 7, 3 and 6 for CAMERA, BOWS, UCID and NRCS respectively. For HYBRID database, the value of $N_p$ is 6 under the embedding rate of 25%, while the value of $N_p$ is 7 under the embedding rate of 50%. It is generally observed that the best performance is achieved when $N_p$ = 6 as long as the number of training images is large enough.

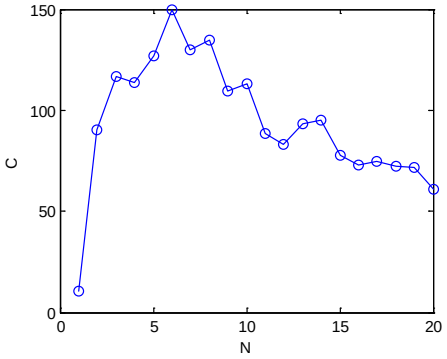


**Fig. 5**. The cost function $C$ for a training subset from CAMERA database with an embedding rate of 50%, using the threshold selection procedure. The performance peaks at $N_p = 6$.

### 3.3 Impact of Image Databases

In this Section, we report the experimental results that test the variability in performance across databases. The fixed embedding rate of 25% is used both during training and testing. Similar behavior is observed for the embedding rate of 50% but data is omitted for brevity and clarity. For comparison purposes, the experiments are performed under the following three conditions.

Specific. Both training and testing are performed on the same database.

Disjoint. Training is performed on three databases, and testing is performed on the fourth database which is not used during training.

Joint. Training is performed on portions of all four databases, and testing is performed on the portion of one of the four databases, which is not used during training.

The experimental results are summarized in **Table 1**. From the table, we can observe that: (1) For each setup, the performance varies considerable across the four image databases, and the performances on CAMERA and BOWS databases are consistently better than that on the other two databases as steganalysis is more accurate on relatively smooth images (e.g. images captured by cameras) than on images with more stochastic texture (e.g. images recaptured by scanner). (2) For different setups, as expected, the best performances are achieved when the training and testing databases match each other ("Specific"). However, when trainging database mismatches testing database ("Disjoint"), we observe considerable performance degradation for all four databases. This indicates that each individual database is not representative of the other databases. Moreover, the most drastic degradation occurs on CAMERA database. This may be due to the fact that this database actually consists of images captured by different cameras (i.e. different image sources). (3) If we do not know anything about the image source, our best strategy is to train the steganalyzer on as diverse image database as possible. The results under this condition are shown in **Table 1** in rows captioned by "Joint". Comparing to "Disjoint", we observe a considerable performance increase on the former two databases and a slight increase on the latter two databases. Whereas, comparing to "Specific", we are still able to observe performance degradation since training is no longer specific database. An alternative way of constructing a steganalyzer that is less sensitive to image source is to train a bank of classifiers for several image sources and recognize the image source of a test image using source identification technique in digital image passive forensics [33][34] and send the image to the appropriate classifier.

**Table 1**. The impact of databases on performance

| Test Database | Setup | AUC | TP_E | FP_80 | FP_50 | Min_E |
|---|---|---|---|---|---|---|
| CAMERA | Specific | 0.9889 | 0.9577 | 0.0073 | 0.0029 | 0.0409 |
| | Disjoint | 0.7782 | 0.6815 | 0.4211 | 0.1790 | 0.2992 |
| | Joint | 0.8797 | 0.8134 | 0.1783 | 0.0738 | 0.1806 |
| BOWS | Specific | 0.9872 | 0.9452 | 0.0126 | 0.0026 | 0.0538 |
| | Disjoint | 0.9589 | 0.9017 | 0.0634 | 0.0238 | 0.0932 |
| | Joint | 0.9780 | 0.9204 | 0.0301 | 0.0052 | 0.0763 |
| UCID | Specific | 0.8809 | 0.8041 | 0.1910 | 0.0511 | 0.1923 |
| | Disjoint | 0.8259 | 0.7418 | 0.3323 | 0.0833 | 0.2519 |
| | Joint | 0.8280 | 0.7439 | 0.3233 | 0.0890 | 0.2494 |
| NRCS | Specific | 0.8436 | 0.7665 | 0.2605 | 0.1044 | 0.2276 |
| | Disjoint | 0.7649 | 0.6859 | 0.4193 | 0.1849 | 0.3076 |
| | Joint | 0.7733 | 0.6925 | 0.4031 | 0.1775 | 0.2990 |

### 3.4 Comparison to Prior Art

For each image database, we randomly select 1 out of 5 cover images and their stego counterparts for training, and the remaining cover and stego images are used for testing. The procedures are repeated 20 times for cross-validatin and the results are averaged to obtain more reliable performance.

We report the experimental results with the embedding rates of 25% and 50% in **Table 2** and **Table 3**, respectively (Note: The values in bold indicate the best performance for each database). From the tables, we can see that: (1) All of algorithms perform more accurately as the embedding rate increases. (2) As described in Section 3.3, all of algorithms exhibit the variability in performance across databases. First, a given algorithm has different performances on different databases. For example, for the LLTPDF method, the AUC difference between CAMARA and UCID is up to 20% at the embedding rate of 25%. Furthermore, one steganalytic algorithm having better performance on one database than on another does not imply that another algorithm will have the same behavior on the two image databases. For instance, OPTCF has better performance on BOWS than on CAMERA, but our proposed approach has better performance on CAMERA than on BOWS. Hence, we should perform the performance comparison of different algorithms on a large number of images. (3) Compared with RDIH, OPTCF and LLTCF, our proposed method performs the best under most circumstances, especially in the low embedding rate cases. (4) Compared with SPAM, our proposed feature set performs better on CAMERA and BOWS, comparative on UCID and less accurate on NRCS and HYBRID. However, it should be noted that different classifiers are applied to these two feature sets. If SVM is also applied to our feature set, our proposed approach will perform better. For instance, we get a 72-dimentional feature set by setting $N_p$ to be 6, and then apply SVM classifier to this feature set on HYBRID, and get the results as shown in **Table 4**. It is evident that our proposed feature set with SVM has a considerable increase in performance compared with the same feature set with FLD and has comparative performance of SPAM. (5) Generally, the proposed algorithm has superiority over others on relatively smooth images (e.g. CAMERA and BOWS), while SPAM performs best on very noisy images (e.g. NRCS).

**Table 2**. Comparison of the performance at an embedding rate of 25%

| Database | Algorithm | AUC | TP_E | FP_80 | FP_50 | Min_E |
|---|---|---|---|---|---|---|
| CAMERA | SPAM | 0.9772 | 0.9277 | 0.0186 | **0.0024** | 0.0719 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | RDIH | 0.9713 | 0.9208 | 0.0278 | 0.0072 | 0.0777 |
| | OPTCF | 0.9071 | 0.8360 | 0.1380 | 0.0389 | 0.1638 |
| | LLTPDF | 0.9777 | 0.9241 | 0.0263 | 0.0065 | 0.0742 |
| | Proposed | **0.9889** | **0.9577** | **0.0073** | 0.0029 | **0.0409** |
| BOWS | SPAM | 0.9826 | 0.9311 | 0.0156 | 0.0030 | 0.0689 |
| | RDIH | 0.9646 | 0.9107 | 0.0417 | 0.0151 | 0.0884 |
| | OPTCF | 0.9504 | 0.8879 | 0.0643 | 0.0189 | 0.1106 |
| | LLTPDF | 0.9744 | 0.9220 | 0.0316 | 0.0100 | 0.0769 |
| | Proposed | **0.9872** | **0.9452** | **0.0126** | **0.0026** | **0.0538** |
| UCID | SPAM | 0.8671 | **0.8065** | **0.1897** | 0.0692 | **0.1921** |
| | RDIH | 0.8257 | 0.7620 | 0.2804 | 0.0972 | 0.2346 |
| | OPTCF | 0.8112 | 0.7389 | 0.3257 | 0.1177 | 0.2574 |
| | LLTPDF | 0.8084 | 0.7258 | 0.3474 | 0.1187 | 0.2687 |
| | Proposed | **0.8809** | 0.8041 | 0.1910 | **0.0511** | 0.1923 |
| NRCS | SPAM | **0.8968** | **0.8238** | **0.1597** | **0.0581** | **0.1748** |
| | RDIH | 0.8274 | 0.7495 | 0.2953 | 0.1097 | 0.2458 |
| | OPTCF | 0.7598 | 0.6912 | 0.4240 | 0.1745 | 0.3050 |
| | LLTPDF | 0.7599 | 0.6878 | 0.4272 | 0.1780 | 0.3086 |
| | Proposed | 0.8436 | 0.7665 | 0.2605 | 0.1044 | 0.2276 |
| HYBRID | SPAM | **0.9492** | **0.8848** | **0.0659** | **0.0185** | **0.1147** |
| | RDIH | 0.8860 | 0.8009 | 0.1986 | 0.0599 | 0.1957 |
| | OPTCF | 0.8735 | 0.7946 | 0.2088 | 0.0783 | 0.2009 |
| | LLTPDF | 0.8920 | 0.8055 | 0.1900 | 0.0629 | 0.1896 |
| | Proposed | 0.9211 | 0.8403 | 0.1340 | 0.0391 | 0.1549 |

**Table 3**. Comparison of the performance at an embedding rate of 50%

| **Database** | **Algorithm** | **AUC** | **TP_E** | **FP_80** | **FP_50** | **Min_E** |
|---|---|---|---|---|---|---|
| CAMERA | SPAM | 0.9979 | 0.9818 | 0.0008 | 0.0000 | 0.0182 |
| | RDIH | 0.9942 | 0.9674 | 0.0049 | 0.0018 | 0.0306 |
| | OPTCF | 0.9686 | 0.9151 | 0.0362 | 0.0105 | 0.0828 |
| | LLTPDF | 0.9901 | 0.9570 | 0.0091 | 0.0028 | 0.0417 |
| | Proposed | **0.9984** | **0.9859** | **0.0005** | **0.0001** | **0.0132** |
| BOWS | SPAM | 0.9947 | 0.9693 | 0.0048 | 0.0018 | 0.0296 |
| | RDIH | 0.9850 | 0.9521 | 0.0167 | 0.0081 | 0.0463 |
| | OPTCF | 0.9816 | 0.9427 | 0.0196 | 0.0100 | 0.0562 |
| | LLTPDF | 0.9871 | 0.9547 | 0.0180 | 0.0067 | 0.0434 |
| | Proposed | **0.9955** | **0.9739** | **0.0049** | **0.0016** | **0.0252** |
| UCID | SPAM | **0.9500** | **0.8860** | **0.0626** | **0.0112** | **0.1112** |
| | RDIH | 0.9014 | 0.8363 | 0.1366 | 0.0396 | 0.1601 |
| | OPTCF | 0.9078 | 0.8317 | 0.1443 | 0.0362 | 0.1644 |
| | LLTPDF | 0.8909 | 0.8062 | 0.1872 | 0.0503 | 0.1888 |
| | Proposed | 0.9463 | 0.8764 | 0.0675 | 0.0113 | 0.1202 |
| NRCS | SPAM | **0.9838** | **0.9486** | **0.0204** | **0.0086** | **0.0500** |
| | RDIH | 0.9405 | 0.8760 | 0.0840 | 0.0302 | 0.1199 |
| | OPTCF | 0.8965 | 0.8197 | 0.1648 | 0.0553 | 0.1764 |
| | LLTPDF | 0.8808 | 0.7974 | 0.2045 | 0.0632 | 0.1983 |
| | Proposed | 0.9269 | 0.8642 | 0.1072 | 0.0515 | 0.1254 |
| HYBRID | SPAM | **0.9838** | **0.9421** | **0.0203** | **0.0069** | **0.0551** |
| | RDIH | 0.9392 | 0.8675 | 0.0997 | 0.0354 | 0.1236 |
| | OPTCF | 0.9474 | 0.8797 | 0.0771 | 0.0256 | 0.1166 |
| | LLTPDF | 0.9359 | 0.8628 | 0.1048 | 0.0353 | 0.1315 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Proposed | 0.9661 | 0.9069 | 0.0531 | 0.0145 | 0.0884 |

**Table 4**. Performance of our feature set with SVM on HYBRID

| Database | Embedding rate | AUC | TP_E | FP_80 | FP_50 | Min_E |
|---|---|---|---|---|---|---|
| HYBRID | 25% | 0.9606 | 0.8986 | 0.0541 | 0.0155 | 0.0989 |
| | 50% | 0.9873 | 0.9474 | 0.0175 | 0.0045 | 0.0479 |

## 4. Conclusion

From the view of image textures, we regard the embedded secret message as a kind of stochastic texture in a fine scale and decompose images into a group of detail subbands with a bank of LLT masks which are very sensitive to textures. In order to capture the largest statistical difference before and aflter embedding, a bank of band-pass filters is constructed to extract novel features from the characteristic functions of the detail subbands. A feature selection algorithm is adopted to search an informative and low dimensional feature set. The impact of image databases is investigated on four diverse uncompressed image databases. Compared with the state-of-the-art steganalytic algorithms, the proposed feature set performs the best under most circumstances. Furthermore, although LSB matching steganography is only discussed in the paper, since the principle of this method is based on additive noise embedding, it is expected applicable to other steganographic methods as well. Future work will take other steganographic algorithms in the spatial or transform domain into consideration.

## References

[1] S. Dumitrescu, X. Wu and Z. Wang, "Detection of LSB steganography via sample pair analysis," in *Proc. of 5th Int. Workshop on Information Hiding*, pp. 355-372, Oct. 2002. Article (CrossRef Link).

[2] J. Fridrich, M. Goljan and R. Du, "Detecting LSB steganography in color and gray-scale images," *IEEE Multimedia*, vol. 8, no. 4, pp. 22-28, Oct. 2001. Article (CrossRef Link).

[3] T. Zhang and X. Ping, "A new approach to reliable detection of LSB steganography in natural images," *Int. Journal Signal Processing*, vol. 83, no. 10, pp. 2085-2093, Oct. 2003. Article (CrossRef Link).

[4] G. Cancelli, G. Doërr, I. J. Cox and M.Barni, "A comparative study of ±1 steganalyzers," in *Proc. of IEEE Workshop on Multimedia Signal Processing*, pp. 791-796, Oct. 2008. Article (CrossRef Link).

[5] A. D. Ker and I. Lubenko, "Feature reduction and payload location with WAM steganalysis," in *Proc. of SPIE Electronic Imaging, Media Forensics and Security*, pp. 72540A-01-72540A-13, Jan. 2009. Article (CrossRef Link).

[6] J. J. Harmsen and W. A. Pearlman, "Steganalysis of additive noise modelable information hiding," in *Proc. SPIE Security, Steganography, and Wartermarking of Multimedia Contents*, pp. 131-142, Jan. 2003. Article (CrossRef Link).

[7] A. D. Ker, "Steganalysis of LSB matching in grayscale images," *IEEE Signal Processing Letters*, vol. 12, no. 6, pp. 441-444, Jun. 2005. Article (CrossRef Link).

[8] X. Li, T. Zeng and B. Yang, "Detecting LSB matching by applying calibration technique for difference image," in *Proc. of 10th ACM Multimedia & Security Workshop*, pp. 133-138, Sep. 2008. Article (CrossRef Link).

[9] J. Zhang, I. J. Cox and G. Doërr, "Steganalysis for LSB matching in images with high-frequency noise," in *Proc. of IEEE Workshop on Multimedia Signal Processing*, pp. 385-388, Oct.

2007. Article (CrossRef Link).
[10] G. Cancelli, G. Doërr, I. J. Cox and M. Barni, "Detection of ±1 LSB steganography based on the amplitude of histogram local extrema," in *Proc. of IEEE Conf. on Image Processing*, pp. 1288-1291, Oct. 2008. Article (CrossRef Link).
[11] Y. Gao, X. Li, B. Yang and Y. Lu, "Detecting LSB matching by characterizing the amplitude of histogram," in *Proc. of IEEE Conf. on Acoustics Speech, Signal Processing*, pp. 1505-1508, Apr. 2009. Article (CrossRef Link).
[12] T. Zhang, W. Li, Y. Zhang, E. Zheng and X. Ping, "Steganalysis of LSB matching based on statistical modeling of pixel difference distributions," *Information Sciences*, vol. 180, no. 23, pp. 4685-4694, Dec. 2010. Article (CrossRef Link).
[13] K. Cai, X. Li and T. Zeng, "Reliable histogram features for detecting LSB matching," in *Proc. of IEEE Conf. on Image Processing*, pp. 1761-1764, Sep. 2010. Article (CrossRef Link).
[14] F. Huang, B. Li and J. Huang, "Attack LSB matching steganography by counting alteration rate of the number of neighbourhood gray levels," in *Proc. of IEEE Conf. on Image Processing*, pp. 401-404, Sep. 2007. Article (CrossRef Link).
[15] Q. Liu, A. H. Sung, B. Ribeiro, M. Wei, Z. Chen and J. Xu, "Image complexity and feature mining for steganalysis of least significant bit matching steganography," *Information Sciences*, vol. 178, no. 1, pp. 21-36, Jan. 2008. Article (CrossRef Link).
[16] G. Wang, X. Ping, M. Xu, T. Zhang and X. Bao, "Steganalytic method based on short and repeated sequence distance statistics," *Science in China Series F: Information Sciences*, vol.51, no.10, pp. 1466-1474, Oct. 2008. Article (CrossRef Link).
[17] T. Pevný, P. Bas and J. Fridrich, "Steganalysis by subtractive pixel adjacency matrix," *IEEE Trans. on Information Forensics and Security*, vol. 5, no.2, pp. 215-224, Mar. 2010. Article (CrossRef Link).
[18] J. Fridrich, D. Soukal and M. Goljan, "Maximum likelihood estimation of length of secret message embedded using ±K steganography in spatial domain," in *Proc. of SPIE Security, Steganography, and Wartermarking of Multimedia Contents*, pp. 595-606, Mar. 2005. Article (CrossRef Link).
[19] P. W. Wong, H. Chen and Z. Tang, "On steganalysis of plus-minus one embedding of continuous tone image," in *Proc. of SPIE Security, Steganography, and Wartermarking of Multimedia Contents*, pp. 643-652, Mar. 2005. Article (CrossRef Link).
[20] H. Farid, "Detecting hidden messages using higher-order statistical models," in *Proc. of IEEE. Conf. on Image Processing*, pp.905-908, Sep. 2002. Article (CrossRef Link).
[21] T. Holotyak, J. Fridrich and S. Voloshynovskiy, "Blind statistical steganalysis of additive steganography using wavelet higher order statistics," in *Proc. of 9th IFIP Conf. on Communications and Multimedia Security*, pp. 273-274, Sep. 2005. Article (CrossRef Link).
[22] M. Goljan, J. Fridrich and T. Holotyak, "New blind steganalysis and its implications," in *Proc. of SPIE Security, Steganography, and Watermarking of Multimedia Contents*, pp. 1-13, Jan. 2006. Article (CrossRef Link).
[23] G. Xuan, Y. Q. Shi, J. Gao, D. Zou, C. Yang, Z. Zhang, P. Chai, C. Chen and W. Chen, "Steganalysis based on multiple features formed by statistical moments of wavelet characteristic functions," in *Proc. of Int. Workshop on Information Hiding*, pp. 262-277, Jun. 2005. Article (CrossRef Link).
[24] Y. Q. Shi, G. Xuan, D. Zou, J. Gao, C. Yang, Z. Zhang, P. Chai, W. Chen and C. Chen , "Image steganalysis based on moments of characteristic functions using wavelet decomposition, prediction-error image,and neural network," in *Proc. of IEEE Conf. on Multimedia and Expo*, pp. 269-272, Jul. 2005. Article (CrossRef Link).
[25] Y. Wang and P. Moulin, "Optimized feature extraction for learning-based image steganalysis," *IEEE Trans. on Information Forensics and Security*, vol. 2, no. 1, pp. 31-45, Mar. 2007. Article (CrossRef Link).
[26] B. Li, J. Huang and Y. Q. Shi, "Textural features based universal steganalysis," in *Proc. of SPIE Security, Forensics, Steganography and Watermarking of Multimedia*, pp. 681912-681912-12, Jan. 2008. Article (CrossRef Link).
[27] I. Ng, T. Tan and J. Kittler, "On local linear transform and gabor filter representation of texture," in

*Proc. of 11th Int. Conf. on Pattern Recognition*, pp. 627-631, 1992. Article (CrossRef Link).
[28] G. Doërr, "Image Database for Steganalysis Studies," [Online]. Available: http://www.cs.ucl.ac.uk/ staff/I.Cox/Content/Downloads.html
[29] P. Bas and T. Furon, BOWS-2, Jul. 2007 [Online]. Available: http:// bows2.gipsa-lab.inpg.fr
[30] G. Schaefer and M. Stich, "UCID – An Uncompressed Colour Image Database," in *Proc. of SPIE Storage and Retrieval Methods and Applications for Multimedia*, pp.472-480, Jan. 2004. Article (CrossRef Link).
[31] [Online]. Available: http://photogallery.nrcs.usda.gov/
[32] C.-C. Chang and C.-J. Lin, LIBSVM: A Library for Support Vector Machines [Online]. Available: http://www. csie.ntu.edu.tw/~cjlin/libsvm
[33] N. Khanna, A. K Mikkilineni, A. F. Martone, G. N. Ali, G. T.-C. Chiu, J. P. Allebach and E. J. Delp, "A survey of forensic characterization methods for physical devices," *Digital Investigation*, vol. 3, no. s1, pp. 17-28, Sep. 2006. Article (CrossRef Link).
[34] T. V. Lanh, K. S. Chong, S. Emmanuel and M. S Kankanhalli, "A survey on digital camera image forensic methods," in *Proc. of IEEE Conf. on Multimedia and Expo*, pp. 16-19, Jul. 2007. Article (CrossRef Link).

**Ergong Zheng** received his B.S. and M.S. degrees in Electrical Engineering from Wuhan University of Technology, Wuhan, China, in 2003 and 2006, respectively, and is currently pursuing the Ph.D. degree in Zhengzhou Information Science and Technology Institute. His research interests include image processing and information hiding.

**Xijian Ping** received his M.S degree in Signal and Information Processing from Beijing University ofAeronautics and Astronautics, Beijing, China, in 1982. He is currently a Professor with Department of Information Science, Zhengzhou Information Science and Technology Institute, where he is also the supervisor of Ph.D. candidates from 1998. His research interests include image processing, pattern recognition and information hiding.

**Tao Zhang** received his M.S. and Ph.D. degrees in Signal and Information Processing from Zhengzhou Information Science and Technology Institute, Zhengzhou, China, in 2000 and 2003, respectively. He is currently an Associate Professor with Department of Information Science, Zhengzhou Information Science and Technology Institute. His research interests include information hiding, image processing and pattern recognition.