Steganalysis of YASS

Bin Li, *Student Member, IEEE*, Jiwu Huang, *Senior Member, IEEE*, and Yun Qing Shi, *Fellow, IEEE*

***Abstract*—A promising steganographic method—Yet Another Steganography Scheme (YASS)—was designed to resist blind steganalysis via embedding data in randomized locations. In addition to a concrete realization which is named the YASS algorithm in this paper, a few strategies were proposed to work with the YASS algorithm in order to enhance the data embedding rate and security. In this work, the YASS algorithm and these strategies, together referred to as YASS, have been analyzed from a warden's perspective. It is observed that the embedding locations chosen by YASS are not randomized enough and the YASS embedding scheme causes detectable artifacts. We present a steganalytic method to attack the YASS algorithm, which is facilitated by a specifically selected steganalytic observation domain (SO-domain), a term to define the domain from which steganalytic features are extracted. The proposed SO-domain is not exactly, but partially accesses, the domain where the YASS algorithm embeds data. Statistical features generated from the SO-domain have demonstrated high effectiveness in detecting the YASS algorithm and identifying some embedding parameters. In addition, we discuss how to defeat the above-mentioned strategies of YASS and demonstrate a countermeasure to a new case in which the randomness of the embedding locations is enhanced. The success of detecting YASS by the proposed method indicates a properly selected SO-domain is beneficial for steganalysis and confirms that the embedding locations are of great importance in designing a secure steganographic scheme.**

***Index Terms*—JPEG, steganalysis, steganography, YASS.**

I. Introduction

STEGANOGRAPHY is the technique of hiding secret data into an innocuous-looking cover medium to achieve the goal of covert communication [1]. A secure steganographic algorithm should guarantee that no one except data senders and data receivers is aware of the existence of the secret data in a stego medium [2].

Steganalysis, on a warden's behalf, is the art of revealing the presence of the secret data [3]. Once a rate better than random guessing can be achieved to decide whether secret data have been embedded into the media by a particular steganographic algorithm, such a steganographic algorithm is considered to be broken. The success of steganalysis relies on the fact that steganography alters some inherent statistics of cover media and the deviated statistics fall outside the normal scope where the statistics of cover media belong.

There are at least three reasons why the research on steganalysis has recently received intensive attention. First, detecting the presence of secret data can be used to deter covert communications launched by terrorists or illegal groups. Second, as the relation between cryptology and cryptanalysis, steganography and steganalysis are in a cat-and-mouse game [4]. The development of steganalysis helps improve the security of information hiding. Third, the research of steganalysis stimulates the attempts to build up better statistical models for multimedia contents, which can lead to successful applications in other related research fields, such as digital forensics [5]–[8].

Steganalytic methods fall into two categories, namely, blind steganalysis (also called universal steganalysis) and specific steganalysis, according to their application fields. Blind steganalysis can be used to detect various types of steganographic algorithms and it can even be adapted to attack new steganographic schemes. It often models the intrinsic nature of cover media via mapping the high dimensional medium space to a relatively low dimensional feature space. Specific steganalysis is targeted to break a particular steganographic algorithm by exploring how a given steganographic algorithm works and how it changes natural statistics of cover media. Some advanced specific steganalytic methods is able to get some more useful information from stego media, such as the rough size of embedded data, the embedding locations, and/or some embedding parameters of the steganography.

The data embedding rate is an important index for comparing the performance of different steganographic algorithms as well as different steganalytic methods. It is commonly measured by the ratio of the total number of embedded information data bits over the total number of possible embedding locations of a cover medium. It is intuitive that the higher the data embedding rate, the more artifacts are introduced into a cover medium, and the more probable such a steganographic scheme will be detectable.

Due to the prevalent usage of joint photographic experts group (JPEG) [9] images, the competition between JPEG steganography and JPEG steganalysis has escalated over the past few years. Several steganographic schemes [10], [11] have been proposed to embed data through replacing the least significant bits of the JPEG quantized alternating current (ac) discrete cosine transform (DCT) coefficients by secret data bits. Learning lessons from steganalysis, some data embedding methods have evolved with adding some advanced techniques [12]–[17]. In order to defeat steganography, some specific JPEG steganalytic schemes [18]–[20] as well as blind JPEG

Manuscript received September 16, 2008; revised May 17, 2009. First published June 30, 2009; current version published August 14, 2009. This work was supported in part by the NSFC (60633030, 90604008), and by the 973 program (2006CB303104). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jessica J. Fridrich.

B. Li was with the School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510275, China, and is now with the College of Information Engineering, Shenzhen University , Shenzhen 518060, China (e-mail: lib3366@gmail.com).

J. Huang is with the School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510275, China (e-mail: isshjw@mail.sysu.edu.cn).

Y. Q. Shi is with the Department of Electronic and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07029 USA (e-mail: shi@njit.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2009.2025841

1556-6013/$26.00 © 2009 IEEE

steganalytic methods [21]–[25] have been proposed. One of the most powerful strategies is to estimate the statistics of a cover image by cropping a few pixel rows and columns of a JPEG decompressed stego image and then recompressing the cropped image with the same compression parameters as the cover image. This strategy is often referred to as self-calibration and is used in [18], [19], [21], and [23]. Another powerful method is to model the inherent correlation between the coefficients by the Markov process [22]–[25]. These two kinds of methodology are not exclusive and can be jointly made up of a powerful blind steganalyzer [23]. Steganalysis has progressed to such an extent that many current steganographic methods have to reduce the data embedding rate in order to avoid causing a warden's suspicion. Fridrich *et al.* [26] concluded from their experimental results that the highest undetectable data embedding rate for a JPEG image with quality factor 70 is around 0.05 bpac (bits per nonzero quantized ac DCT coefficient of the cover image) under the blind JPEG steganalyzer in [23].

Yet Another Steganography Scheme (YASS) [27] intends to resist the existing blind steganalytic algorithms, especially the self-calibration-based methods. A concrete realization (which is named the YASS algorithm) has been described in [27] and a few strategies [27]–[29] also have been proposed to work with the YASS algorithm in order to enhance the data embedding rate and security. In this work, the YASS algorithm and these strategies are together referred to as YASS. The philosophy in YASS is quiet simple yet effective. Since the self-calibration-based method relies on estimating the macroscopic characteristics of a cover image from its stego version, YASS tries to distort the estimation by hiding data in randomized locations. More specifically, in the YASS algorithm, an image is first divided into blocks of size larger than 8×8. Then within each block, a randomly selected 8×8 sub-block, referred to as embedding host block (or H-block), is performed for DCT, data hiding, and inverse DCT. JPEG compression is applied to the whole image afterward.

According to the embedding manner of YASS and the format of its stego images, YASS can be regarded as either a spatial image steganography, or a JPEG steganography. Specifically, since YASS does not directly manipulate the JPEG quantized DCT coefficients, it can be classified as a steganography hiding data in the spatial domain in a robust manner [30]. However, as reported in [27] and [28], steganalytic methods [31], [32] which work well in detecting spatial domain steganography are not effective in detecting YASS. On the other hand, since YASS distributes stego images in JPEG format, it can also be considered as a JPEG steganography, as claimed by its authors. But the existing JPEG steganalytic schemes do not achieve good performance consistently, judging from the results obtained in [27], [28], and [30].

Motivated by the challenge of YASS and based on our previous work [33], we propose a specific steganalytic method in this paper. In doing so, we introduce a term—*steganalytic observation domain* (*SO-domain*)—as the domain for extracting steganalytic features. Following the success made in [21]–[25] by using the JPEG domain, where JPEG steganography takes place, as an SO-domain, we analyze the domain where YASS embeds data. Unlike the case that the data embedding domains of many steganographic methods are clearly known by a warden, the data embedding domain of YASS is carefully designed so that it cannot be fully accessed unless having a secret key. However, we have found out that the locations of the H-blocks in YASS may not be randomized enough. A specially designed SO-domain has been proposed to partially access the data embedding domain of the YASS algorithm. Some sets of the proposed SO-domain access the possible embedding locations with a certain probability, while other sets of the proposed SO-domain access the impossible embedding locations. The statistics of these two kinds of sets may show slight, or even no, difference in a cover image but great difference in a stego image. Consequently, we extract some statistical features in the SO-domain and detect the presence of the YASS algorithm by using supervised learning classifiers. Simulations are performed and the experimental results indicate that the proposed method is robust to some embedding parameters which make the data embedding rate of the YASS algorithm still appealing. The strategies in [27]–[29] that were proposed to enhance the capability of the YASS algorithm also suffer a similar security problem as the YASS algorithm, namely, the locations of H-blocks may not be randomized enough. Hence the idea of the proposed steganalytic method can also be adapted to defeat these further strategies. A possible way to increase the randomness of the H-blocks' location has been discussed and its countermeasure also has been demonstrated. This work confirms that the embedding locations are very important to a secure steganographic scheme.

This paper is organized as follows. To make this paper self-contained, Section II briefly covers the basic operations of the YASS algorithm as well as provides an analysis on its data embedding rate. Section III focuses on the essential idea of the proposed SO-domain against the YASS algorithm and describes the process of steganalytic feature extraction. We verify the effectiveness of the proposed method under different practical scenarios with experimental results in Section IV. Section V discusses possible ways to defeat some further strategies of YASS and a countermeasure to a new strategy in which the randomness of the embedding locations is enhanced. It also addresses the limitation of the proposed method. Contributions that have been made in this paper are summarized in Section VI.

II. Overview of the YASS Algorithm

A. Fundamental Operations of the YASS Algorithm

Given an image $\boldsymbol{I}$ of size $M \times N$, the data embedding procedure of the YASS algorithm [27] may be described as follows.

1) Encode the secret information data with error correction codes. YASS uses the repeat-accumulate (RA) codes. Here we refer to the encoded data as *RA-coded data*.
2) Divide $\boldsymbol{I}$ into consecutive disjoint blocks of size $B \times B$ ($B > 8$), which are referred to as *big blocks* (or *B-blocks*).
3) For each B-block, an 8×8 sub-block is randomly selected using a secret key only shared by data senders and receivers. The sub-block is named *embedding host block* (or *H-block*).
4) Perform 2-D DCT on each H-block. The resultant DCT coefficients, denoted by $D_{u,v} (u, v \in \{0, 1, \dots, 7\})$, are

respectively divided by the corresponding quantization steps, which are denoted by $q_{u,v}$ and specified by a *design quality factor* QF_h. The coefficients after division are called *un-rounded coefficients*, denoted by $Q_{u,v}$. That is,

$$Q_{u,v} = \frac{D_{u,v}}{q_{u,v}}. \tag{1}$$

5) The un-rounded coefficients whose rounding values are nonzero from some predetermined sub-bands (called *candidate embedding sub-bands*) are modified for hosting RA-coded data. All un-rounded coefficients that are not in candidate embedding sub-bands or their rounding values are zeros will not be touched to avoid introducing unnecessary artifacts. Quantization index modulation (QIM) [34], a commonly used data embedding technique, is employed to hide data in order to enhance the robustness. YASS uses a scalar QIM scheme [35] in which the RA-coded data are embedded by choosing uniform quantizers. Two quantizers with a step size of 2Δ are used, denoted by $\mathbf{Q}_0$ and $\mathbf{Q}_1$, respectively. If a to-be-embedded bit is "0", an even quantizer $\mathbf{Q}_0$ is employed to quantize the un-rounded coefficient to the nearest even reconstruction point, whose value is an even multiple of Δ. Similarly, an odd quantizer $\mathbf{Q}_1$ will be used to quantize the coefficient to the nearest odd reconstruction point if the to-be-embedded bit is "1". In other words, if we denote the set of the candidate embedding sub-bands as $\mathbb{Q}$, the coefficients after the data hiding as $Q'_{u,v}$, and a to-be-embedded bit as m, we will have (2) and (3), shown at the bottom of the page, where $\lfloor\cdot\rfloor$ is the floor operation. The quantization parameter Δ controls the trade-off between the robustness of embedded data and the distortions introduced by quantization. Typically, $\Delta = 1$.
6) After hiding data, multiply the coefficients in H-blocks with the quantization steps associated with QF_h and perform 2-D inverse DCT. The resultant pixel values are rounded to integers.
7) The entire image is compressed to JPEG with an *advertised quality factor* QF_a.

In data extraction, the image is first JPEG decompressed. Then retrieve all H-blocks with the secret key and perform 2-D DCT on them. Next, the achieved DCT coefficients are quantized by the quantization steps associated with QF_h. The resultant quantized coefficients from the candidate embedding sub-bands are further processed to recover the embedded data.

In the final step of data embedding, JPEG compression will inevitably introduce disturbance to the embedded data. In addition, the extracted zero quantized coefficients may be from two sources. Some of them may be generated from the un-rounded coefficients whose rounding values were zeros (e.g., $Q'_{u,v} = Q_{u,v} = 0.3$). Others are from the un-rounded coefficients that have been changed to zeros after QIM hiding (e.g., $\Delta = 1, m = 0, Q_{u,v} = 0.9, Q'_{u,v} = 0$). Therefore, YASS regards all extracted zero quantized coefficients on the data extraction's side as erasure symbols, no matter whether they have been processed by a QIM quantizer or not. The positions of the erasure symbols are known on the data extraction side. The communication channel used by YASS thus can be considered as a binary erasure channel. With the help of the technique of erasure and error correction codes, correct data extraction can be ensured. Therefore, in the first step of YASS embedding, the secret information data are encoded by RA codes with a redundancy factor q, which involves q-fold repetition, interleaving, and accumulation [27], [29]. A sum-product algorithm [36] is used to decode the encoded data.

Some further strategies of YASS, such as using more than one H-block per B-block (when $B > 16$) [27], using variable design quality factors [28], applying iterative embedding process [28], or correctly estimating the redundant factor on the data extraction's side [29], can further enhance the data embedding rate and the security of YASS. To avoid obscuring the main idea, we focus on detecting the YASS algorithm in this part. The discussions on how to adapt the proposed method to counter the further strategies are presented in Section V-A.

B. Analysis of Data Embedding Rate

As mentioned in Section I, the data embedding rate is crucial to evaluate the performance of a steganography. Here we give a simple analysis on the factors that may influence the data embedding rate of the YASS algorithm. We follow [30] to measure the data embedding rate in terms of *bpac*. The data embedded to the un-rounded coefficients are actually RA-coded data. The maximum capacity for the RA-coded data and that for the secret information data are $M' \times N' \times |\mathbb{Q}|$ and $\lfloor M' \times N' \times |\mathbb{Q}|/q \rfloor$, respectively, where $M' = \lfloor M/B \rfloor, N' = \lfloor N/B \rfloor$, and $|\cdot|$ is the cardinality of a set.

For a given image, it is obvious that when B or q increases and other conditions remain, the data embedding rate decreases in the YASS algorithm (in which only one H-block is used per B-block). The redundancy factor q is dependent on the properties of the binary erasure channel, which is mainly affected by QF_h and QF_a. Denote the embedded symbol as $m (m \in \{0,1\})$ and the extracted symbol as $\widehat{m}$ ($\widehat{m}\in \{0, e, 1\}$, where e denotes the erasure symbol). The 2×3 channel transition probability matrix $p(\widehat{m} \;|m)$ describes the characteristics of the channel [29]. Due to the fact that some un-rounded coefficients will be

$$Q'_{u,v} = \begin{cases} \mathbf{QIM}(Q_{u,v}), & \text{if } (u,v) \in \mathbb{Q}\text{and}\lfloor Q_{u,v} + 0.5 \rfloor \neq 0 \\ Q_{u,v}, & \text{otherwise} \end{cases} \tag{2}$$

$$\mathbf{QIM}(Q_{u,v}) = \begin{cases} \mathbf{Q}_0(Q_{u,v}) = 2\Delta \left\lfloor \frac{Q_{u,v}+\Delta}{2\Delta} \right\rfloor, & \text{if } m = 0 \\ \mathbf{Q}_1(Q_{u,v}) = 2\Delta \left\lfloor \frac{Q_{u,v}}{2\Delta} \right\rfloor + \Delta, & \text{if } m = 1 \end{cases} \tag{3}$$

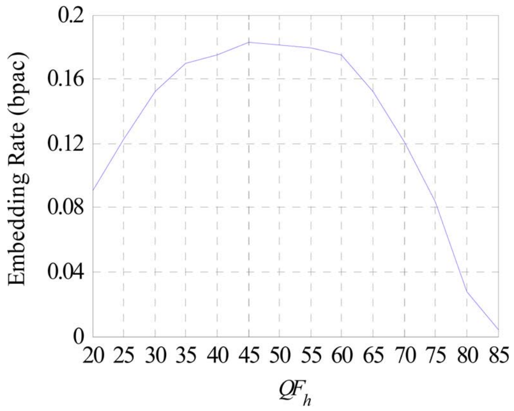

Fig. 1. Relation between the data embedding rate (averaged from 100 images with $B = 9$, $\mathrm{QF}_a = 75$) and QF_h (ranging from 20 to 85 with a step size of 5).

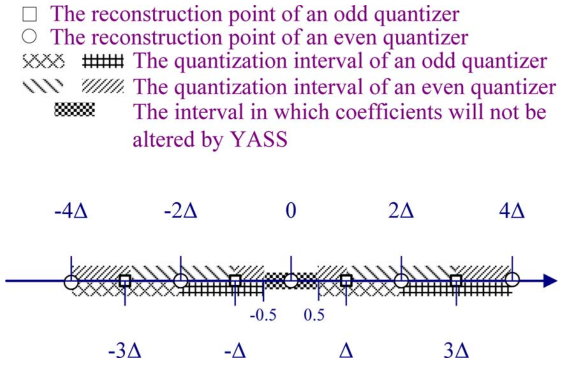
□ The reconstruction point of an odd quantizer
○ The reconstruction point of an even quantizer
The quantization interval of an odd quantizer
The quantization interval of an even quantizer
The interval in which coefficients will not be altered by YASS
-4Δ
-2Δ
0
2Δ
4Δ
-0.5
0.5
-3Δ
-Δ
Δ
3Δ

Fig. 2. Quantization interval of QIM used in YASS for the unrounded coefficients.

changed to zeros (symbol e's at the receiver's side) by embedding "0" while they will never be changed to zeros by embedding "1", such a channel is asymmetric. That is,

$$p(\widehat{m}= e|m = 0) > p(\widehat{m}= e|m = 1). \tag{4}$$

For a given QF_a, when QF_h gets larger, the embedded data are less robust to resist the JPEG compression. Hence the transition probabilities $p(\widehat{m}= 1|m = 0)$ and $p(\widehat{m}= 0|m = 1)$ will be larger, and more redundancy is needed for error correction. At the other extreme, when QF_h gets smaller, more erasure symbols are produced due to the larger quantization steps in (1). Therefore, $p(\widehat{m}= e|m = 0)$ and $p(\widehat{m}= e|m = 1)$ will be larger, and more redundancy is needed for erasure recovery. Usually, the erasures are easier than the errors to deal with by erasure and error correction codes, since the exact locations of the erasures are known, while that of the errors are not [37]. One or more than one QF_h value should exist to maximize the data embedding rate under a given QF_a. Take $\mathrm{QF}_a = 75$ and $B = 9$ for example, we vary QF_h from 20 to 85 with a step of 5. The possible largest data embedding rate for each QF_h averaged from 100 images (randomly selected from the image database described in Section IV-A) is depicted in Fig. 1. The figure indicates that the data embedding rate is maximized around $\mathrm{QF}_h = 45$, which is close to the results reported in [28]. When $\mathrm{QF}_h > \mathrm{QF}_a$, the data embedding rate is much lowered than its maximum. This is because the embedded data under a large QF_h are not robust enough to survive after a relatively severe JPEG compression. Empirically, $\mathrm{QF}_h \leq \mathrm{QF}_a$ holds to ensure the robustness of the embedded data as well as an appealing data embedding rate. Therefore, as [27]–[30], we are interested in the case of $\mathrm{QF}_h \leq \mathrm{QF}_a$ in this paper.

III. Steganalyzing the YASS Algorithm

A. Effect of QIM Embedding

The artifacts introduced to a stego image by YASS are attributed to the QIM embedding operation applied to the H-blocks. As described in Section II-A, Step 5, QIM embedding can be considered as an operation of employing two quantizers, namely, an odd quantizer for quantizing signal to carry "1", and an even quantizer for quantizing signal to carry "0". In this way, a coefficient whose value is in the interval of $[(2k-1)\Delta, (2k+1)\Delta)$ will be quantized to $2k\Delta$ if an even quantizer is used, while a coefficient whose value is in the interval of $[2k\Delta, (2k+2)\Delta)$ will be quantized to $(2k+1)\Delta$ if an odd quantizer is used, where k denotes an integer. An additional note of the data embedding by YASS is that the un-rounded coefficients whose values are in the interval of $[-0.5, 0.5)$ will not be altered. Fig. 2 demonstrates the quantization interval of QIM used in YASS for the un-rounded coefficients.

Suppose we do not apply the QIM embedding to the un-rounded coefficients. Instead, we perform rounding on the un-rounded coefficient. The rounding process thus can be considered as an operation with only one quantizer. A coefficient whose value falls in $[k-0.5, k+0.5)$ will be quantized to k in rounding. As a result, the coefficient values after QIM embedding and that after rounding may be different.

B. Inspecting the Locations of H-Blocks

Judging from the performance of previous steganalytic algorithms [21]–[25] in detecting JPEG steganography, extracting steganalytic features directly from the domain where the steganography takes place is straightforward and effective. As YASS does not embed data in JPEG quantized DCT coefficients, we are not going to find steganalytic features for YASS in the JPEG domain as some previous methods did [21]–[25]. We should look for some deviated statistics in H-blocks because the union of all H-blocks after 2-D DCT can be regarded as the data embedding domain of YASS. But without knowing the secret key, the locations of H-blocks cannot be accessed. Even though QIM embedding performs differently compared to rounding and alters some inherent statistics of images [38], [39], it is not straightforward for a warden to detect YASS.

From the data embedding procedure described in Section II-A, we know that each H-block should reside entirely inside a B-block. Although we cannot tell the exact location of each H-block because it may be controlled by a secret key, the locations where H-blocks cannot reside can be determined, as analyzed below.

We start our analysis with assuming the B-block size is known. Based on this assumption, the basic idea of steganalyzing the YASS algorithm is obtained. But we would like to stress that the proposed steganalytic method does not

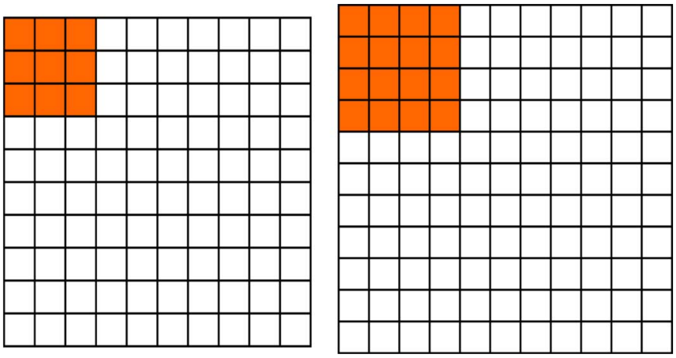

Fig. 3. Possible locations of the origin (filled with color) and impossible locations of the origin (blank) of an 8×8 H-block in a 10×10 B-block (left) and that in an 11×11 B-block (right).

need to know the exact B-block size beforehand. In fact, the proposed method can identify the B-block size when the B-block size is not too large and the data embedding rate is not too low. Denote the pixel elements in a B-block as $b_{s,t}(s,t \in \{0,1,\ldots,B-1\})$ and the pixel elements in an H-block as $d_{x,y}(x,y \in \{0,1,\ldots,7\})$. Define the *origin* of a block as the upper-left element of the block. Hence the origin of an H-block $d_{0,0}$ can only be coincided with $b_{i,j}$, where $i,j \in \{0,1,\ldots,B-8\}$. In other words, only a total number of $(B-7)^2$ elements in the region of the upper-left corner of a B-block can be the origin of an H-block. Other $B^2-(B-7)^2$ elements in a B-block are definitely impossible to be the origin of an H-block. Statistics extracted from 8×8 blocks with the origins in these two kinds of regions may show difference due to their different possibility in YASS data embedding. Fig. 3 shows the possible locations and impossible locations of the origin of an H-block in a B-block. This observation reveals that the embedding location of the YASS algorithm is not randomized enough. Such a shortcoming will shed light in detecting the YASS algorithm and its further strategies.

In our proposed method for detecting the YASS algorithm, instead of analyzing the statistics of all 8×8 blocks with the origins at $b_{s,t}(s,t \in \{0,1,\ldots,B-1\})$, we only focus on analyzing the statistics of 8×8 blocks with the origins at $b_{i,i}(i \in \{0,1,\ldots,B-1\})$. This procedure reduces the feature extraction space from B^2 to B and is yet still possible to meet the need of having two different kinds of regions for feature extraction. In fact, the 8×8 blocks with the origins at $b_{0,0}, b_{1,1}, \ldots$, or $b_{(B-8),(B-8)}$ are possible to fully coincide with some H-blocks, with a probability of $1/(B-7)^2$, whereas all 8×8 blocks with the origins at $b_{(B-7),(B-7)}, b_{(B-6),(B-6)}, \ldots$, or $b_{(B-1),(B-1)}$ are definitely impossible to fully coincide with any H-block. These two kinds of blocks bear different characteristics and thus steganalytic features extracted from them should be different. More details will be covered in the following subsections.

C. Steganalysis Through JPEG Requantization

From Section III-A, we know that QIM embedding and rounding do have differences. Obtaining some statistical features that can reflect the differences between stego images and cover images is a crucial step for steganalysis. Note that owning to the robust QIM embedding, the statistics of the coefficients in the data embedded H-blocks can be well preserved to some extent even though there is a JPEG compression in the final stage of YASS data embedding, especially when $\mathrm{QF}_h \leq \mathrm{QF}_a$. Suppose we have collected all H-blocks from a stego image and all corresponding H-blocks from a cover counterpart. (In fact, there is no actual H-block in a cover image. Here the H-blocks in a cover image are the 8×8 sub-blocks whose locations are the same as those in a corresponding stego image.) And suppose we use a JPEG quantizer at QF_h to requantize them, namely, perform 2-D DCT on H-blocks, then divide the DCT coefficients with quantization steps specified by a quantization table at QF_h, and finally round the coefficients. We refer to the resultant coefficients as *H-block requantized coefficients*. The statistics of H-block requantized coefficients from a stego image and that from a cover image will show differences. This is because the H-blocks in the stego image have been embedded data by QIM, but the corresponding H-blocks in the cover image have never been embedded data. In this way, the statistical abnormality of QIM embedding by YASS can be exposed.

But these two assumptions are in ideal conditions and often do not hold in practice. One difficulty lies in that a warden cannot access all H-blocks without a secret key. Another is that a warden sometimes may not know the exact value of QF_h. However, we can overcome these difficulties to some extent as shown below.

First of all, from Section III-B, we know that there are some locations that are possible for holding H-blocks and some locations that are definitely not possible for holding H-blocks. Even though we cannot access all H-blocks, the statistical features extracted from possible locations of H-blocks and that from impossible locations of H-blocks should have differences. In fact, the statistical features extracted from impossible locations of H-blocks can serve as a calibration signal, whose role is similar to the features extracted from the self-calibrated version of an image [18], [21], [23]. Second, since QF_a is known in a given JPEG image, we can use a JPEG quantizer at QF_a to replace a JPEG quantizer at QF_h for requantization. We can even identify the QF_h in some cases with the use of the modern techniques in double JPEG compression detection [40], [41].

D. Domain for Extracting Steganalytic Features

We call the domain from which the statistical features are extracted for steganalysis the *SO-domain*. It is efficient to use the data embedding domain as the full, or a part of, the SO-domain, as observed from previous methods [21]–[25]. As the evolving steganography becomes more complicated, for example, the data embedding domain is adaptive to the image content or hardly accessible, the need for an explicit definition of the SO-domain may rise. The SO-domain in this work to detect the YASS algorithm is constructed as follows.

1) Given an input JPEG image of size $M \times N$, decompress it to spatial representation and then separate it into disjoint $B \times B$ blocks. Denote the B-block as $K_{m,n}(m \in \{0,1,\ldots,\lfloor M/B \rfloor -1\}, n \in \{0,1,\ldots,\lfloor N/B \rfloor -1\})$ and the pixel elements in $K_{m,n}$ as $b_{m,n}^{s,t}(s,t \in \{0,1,\ldots,B-1\})$.

2) Denote $S^i_{m,n}$ as an 8×8 sub-block with the origin that is coincided with the $b^{i,i}_{m,n}(i \in \{0,1,\ldots,B-1\})$ of $K_{m,n}$.
3) Perform 2-D DCT on $S^i_{m,n}$. The resulting block is named the *feature generation block* (or *F-block*) and denoted by $C^i_{m,n}$.
4) Denote the set that contains all $C^i_{m,n}$ by $\mathbb{C}^i$. That is,

$$\mathbb{C}^i = \{C^i_{m,n} | m \in \{0,1,\ldots,\lfloor M/B \rfloor - 1\}, \\ n \in \{0,1,\ldots,\lfloor N/B \rfloor - 1\}\} \, (i \in \{0,1,\ldots,B-1\}).$$

5) The union of $\mathbb{C}^i$, namely, $\cup_{i=0}^{B-1}\mathbb{C}^i$, forms the proposed SO-domain for detecting YASS.

Under such definitions, blocks in the set $\cup_{i=0}^{B-8}\mathbb{C}^i$ are possible to overlap with the data embedded H-blocks of a stego image with a probability of $1/(B-7)^2$, while blocks in the set $\cup_{i=B-7}^{B-1}\mathbb{C}^i$ are definitely impossible to overlap with any H-block. Hence steganalytic features extracted from $\cup_{i=B-7}^{B-1}\mathbb{C}^i$ can be served as calibration signals, because no data is directly embedded into them. Apparently, some sets of the proposed SO-domain partially access the data embedding domain of the YASS algorithm with a larger probability than some other sets. As the B-block size B increases, the proposed SO-domain accesses the data embedding domain of the YASS algorithm with a smaller probability. And the performance of the proposed method will drop, as seen from the results in Section IV.

E. Steganalytic Features for Detecting YASS

After obtaining the SO-domain, we quantize the DCT coefficients in F-blocks with the quantization steps specified by QF_a. The resulting coefficients are referred to as *F-block requantized coefficients* (or *requantized coefficients*). In the following, we introduce three sets of statistical features that extracted from the requantized coefficients. They are applicable in detecting the YASS algorithm as well as the further strategies of YASS.

1) Frequency of Zero Requantized Coefficients: Consider using a JPEG quantizer at QF_h to obtain H-block requantized coefficients, as described in Section III-C. The zero-valued H-block requantized coefficients of the cover image are mainly generated from the un-rounded coefficients in H-blocks whose values are in the interval of $[-0.5, 0.5)$. Meanwhile, as illustrated in Fig. 2, we learn from the QIM embedding scheme used in YASS that the un-rounded coefficients in H-blocks whose values are in the interval of $[-0.5, 0.5)$ will not be altered. If they were requantized in the requantization, they would probably become 0's. The un-rounded coefficients in the interval of $[-\Delta, -0.5) \cup [0.5, \Delta)$ will have chances to be altered to 0's in the QIM embedding process. For instance, the chance is close to 50% if "0" and "1" are uniformly distributed in the RA coded data. The larger the Δ, the more robust the embedding is, and the more extra zero H-block requantized coefficients are introduced to a stego image by QIM when compared to a cover image. If we use a JPEG quantizer at QF_a to obtain H-block requantized coefficients, more zero H-block requantized coefficients can still be observed in a stego image than in a cover image. Now consider the F-block requantized coefficients. Since some F-blocks overlap with the data embedded H-blocks in a stego image, similarly, more zero F-block requantized coefficients appear in a stego image than in a cover image.

We define *the frequency of zero requantized coefficients* as the ratio of the amount of zero requantized coefficients in the candidate embedding sub-bands over the total number of requantized coefficients in the candidate embedding sub-bands. Denote the frequency of the zero requantized coefficients from candidate embedding sub-bands in $\mathbb{C}^i$ as z^i, where $i \in \{0,1,\ldots,B-1\}$. It is expected that the mean value of $z^0, z^1, \ldots, z^{B-8}$ is larger than the mean value of $z^{B-7}, z^{B-6}, \ldots, z^{B-1}$ in a stego image whose B-block size is B. But such an abnormal phenomenon is not expected to appear in a cover image. So we propose to use

$$\alpha^B = \left(\sum_{i=0}^{B-8} z^i\right) \bigg/ (B-7) \tag{5}$$

and

$$\beta^B = \left(\sum_{i=B-7}^{B-1} z^i\right) \bigg/ 7 \tag{6}$$

as two steganalytic features.

2) Probabilities of the First Significant Digits of Requantized Coefficients: The distribution of the first significant digits (also called first digits) of quantized DCT coefficients can be utilized for some forensic purposes, such as identifying the quality factor in the primary compression of a doubly JPEG compressed image [41], [42]. QIM embedding alters the distribution of the quantized DCT coefficients in H-blocks, and it may also alter the distribution of the first digits of quantized DCT coefficients in H-blocks. Although the QIM quantizers do not quantize the coefficients as a JPEG quantizer does, the resulting coefficient values are also multiples of the quantization steps specified by QF_h. If a JPEG quantizer with QF_a ($\mathrm{QF}_h \neq \mathrm{QF}_a$) is applied to the data embedded H-blocks, a double quantization phenomenon [40]–[42] may occur. When we use a JPEG quantizer at QF_a to quantize the F-blocks, a double quantization phenomenon may still exist because some F-blocks and some H-blocks are overlapped. Thus the distribution of first digits of requantized coefficients can be utilized to identify the QF_h in YASS data embedding.

Denote the probabilities of the first digits of requantized coefficients that are in the candidate embedding sub-bands of $\mathbb{C}^i$ as $p^i_d, (d \in \{1,2,\ldots,9\}, i \in \{0,1,\ldots,B-1\})$. Then $p^i_d(i \in \{0,1,\ldots,B-8\})$ should be different from $p^i_d(i \in \{B-7, B-6,\ldots,B-1\})$ because $\cup_{i=0}^{B-8}\mathbb{C}^i$ and $\cup_{i=B-7}^{B-1}\mathbb{C}^i$ bear different characteristics as mentioned in Section III-D. In order to reduce the feature dimension, we use the features

$$\gamma^B_d = \left(\sum_{i=0}^{B-8} p^i_d\right) \bigg/ (B-7) - \left(\sum_{i=B-7}^{B-1} p^i_d\right) \bigg/ 7 \tag{7}$$

where $d \in \{1,2,\ldots,9\}$, with nine dimensions for a given B.

3) Joint Probabilities of Requantized Coefficients in Neighboring F-Blocks: It is known that DCT coefficients of the same sub-band in two neighboring DCT blocks are highly correlated and their correlation can be exploited to design powerful blind steganalytic methods [25]. In our case, two neighboring F-blocks also have some correlation due to their spatial vicinity. Denote a coefficient of sub-band (u,v) in the F-block $C^i_{m,n}$ by

$c^i_{m,n}(u,v)$. We investigate the joint probability of $c^i_{m,n}(u,v)$ and $c^i_{m,n+1}(u,v)$, denoted by $p(c^i_{m,n}(u,v), c^i_{m,n+1}(u,v))$. Since the statistics in H-blocks are changed by QIM embedding whereas the statistics in blocks from the impossible locations of H-blocks almost remain unchanged, the joint probabilities in $\cup_{i=0}^{B-8}\mathbb{C}^i$ are expected to be different from that in $\cup_{i=B-7}^{B-1}\mathbb{C}^i$. In addition, the distribution of block DCT coefficients follows a Laplacian-like distribution, which means the coefficient values concentrate around 0. Hence, we are interested in the joint probability of coefficients whose values are -1, 0, and 1, namely, $p(c^i_{m,n}(u,v) = t_1, c^i_{m,n+1}(u,v) = t_2)$, where $t_1, t_2 \in \{-1,0,1\}$ and $(u,v) \in \mathbb{Q}$. We propose to use

$$\eta^B_{t_1,t_2} = \left(\sum_{i=0}^{B-8} p\Big(c^i_{m,n}(u,v) = t_1, c^i_{m,n+1}(u,v) = t_2\Big)\right)\Big/(B-7) - \left(\sum_{i=B-7}^{B-1} p\Big(c^i_{m,n}(u,v) = t_1, c^i_{m,n+1}(u,v) = t_2\Big)\right)\Big/7 \quad (8)$$

as steganalytic features with nine dimensions.

4) Feature Vector: We extract a set of features with a total number of 20 dimensions for a given B. We denote the steganalytic feature set for detecting YASS of a specific B-block size B by $F_B = \{\alpha^B, \beta^B, \gamma^B_d, \eta^B_{t_1,t_2}\}(d \in \{1,2,\ldots,9\}, t_1, t_2 \in \{-1,0,1\})$. In the proposed steganalytic scheme, we only consider detecting $B \leq 15$ due to the low data embedding rate when $B > 15$. In order to detect YASS for $9 \leq B \leq 15$, we use $\boldsymbol{F} = \{\boldsymbol{F}_9, \boldsymbol{F}_{10}, \ldots, \boldsymbol{F}_{15}\}$ to generate a final feature vector with a dimension of 140. Therefore, the proposed method does not need the knowledge of B-block size. In addition, the feature $\boldsymbol{F}_B$ will show abnormality for a stego image the B-block size of which is exactly equal to B. Thus, the proposed feature set $\boldsymbol{F}$ can help identify the B-block size.

F. Steganalytic Classifiers

Classifiers based on supervised learning theory are useful tools in classifying cover objects and stego objects. Steganalytic feature vectors extracted from training data are fed into a supervised learning based classifier to train a statistical model. Such a trained model can effectively help to decide to which class the testing objects belong.

A simple two-class Fisher's linear discriminant (FLD) [43] classifier is employed in the proposed scheme. Other classifiers such as support vector machine can also be used. We consider stego images as a positive class and cover images as a negative class. In the two-class FLD classification scheme, the feature vectors extracted from a training image set containing both classes are used to determine a projection matrix that maximizes the distance between the means of the two classes while minimizing the variance within each class. The projection matrix is then employed to project the feature vector of a testing image into a value that is compared with a threshold in order to determine whether the testing image belongs to a positive class or a negative one.

A multiclass classifier based on "one-against-one" criterion [43] is also used in the proposed scheme to identify some parameters utilized by YASS in data embedding. In this "one-against-one" approach, we construct $N\ (N-1)/2$ two-class FLD classifiers for a total number of N classes. For example, cover images can be regarded as a class and stego images with different B can be regarded as different classes. Each two-class classifier is trained by two classes and it can discriminate between such two classes after the training. A feature vector, extracted from the testing image, is assigned to a target class using each two-class classifier in turn, and a majority vote is taken. The class with the maximum votes is selected as the target class for the testing image.

The two-class classifier and the multiclass classifier are employed according to the practical scenarios, which are discussed in Section IV.

IV. Practical Considerations and Performance Evaluation

A. Experimental Setup

To evaluate the performance of the proposed methods on detecting the YASS algorithm, a total number of 2667 images are used as source images. They have never been JPEG compressed. Our group members take 1124 of them with a Panasonic DMZ-FZ30 camera, while the other 1543 images are downloaded from the NRCS website [44]. They are central-cropped to the size of 512×512 for experimental purpose. The first 19 ac sub-bands in zigzag order of luminance channel are selected as the candidate embedding sub-bands in the data embedding process. The advertised quality factor QF_a is set to 75 for cover images and stego images in all cases of our simulations. We focus on detecting the stego images with the B-block size from 9 to 15. Six types of QF_h, ranging from 50 to 75 with a step size of 5, are used for each image and each B-block size. The QIM embedding parameter is set to $\Delta = 1$.

Each image is embedded by its possible maximum data embedding rate.[1] The mean of the data embedding rate under different B and different QF_h are listed in Table I. It can be observed that the averaged data embedding rate decreases as B or QF_h increases.

B. Classifying Cover Images and Stego Images if B-Block Size and Design Quality Factor are Known

When the B-block size B and the design quality factor QF_h are assumed to be known, a two-class classification scheme can be employed to discriminate stego images from cover images. Previous reported results [27], [30] based on blind steganalytic methods belong to this case.

[1] We use the implementation code of the YASS algorithm downloaded from the website http://vision.ece.ucsb.edu/data_hiding/resist_blind_steganalysis.shtml and incorporate it with our implementation of the encoding and decoding of RA codes. The embedding rate is computed under the condition when a zero bit-error rate is achieved at the data extraction's side with a 30-time iterative RA decoding.

TABLE I
MEAN OF THE DATA EMBEDDING RATE (IN bpac)

B	**9**	**10**	**11**	**12**	**13**	**14**	**15**
***QF_h*=50, *QF_a*=75**	0.180	0.152	0.122	0.105	0.089	0.076	0.068
***QF_h*=55, *QF_a*=75**	0.175	0.149	0.119	0.103	0.087	0.074	0.066
***QF_h*=60, *QF_a*=75**	0.169	0.143	0.116	0.099	0.083	0.071	0.063
***QF_h*=65, *QF_a*=75**	0.148	0.129	0.102	0.091	0.074	0.064	0.056
***QF_h*=70, *QF_a*=75**	0.117	0.103	0.081	0.075	0.059	0.051	0.045
***QF_h*=75, *QF_a*=75**	0.078	0.071	0.054	0.052	0.039	0.035	0.030

TABLE II
TWO-CLASS CLASSIFICATION RESULTS (IN PERCENTAGE). TPR STANDS FOR TRUE POSITIVE RATE, TNR FOR TRUE NEGATIVE RATE, AND ACC FOR ACCURACY RATE

B		**9**	**10**	**11**	**12**	**13**	**14**	**15**
***QF_h*=50 *QF_a*=75**	**TPR**	99.98	99.95	99.66	99.39	98.31	97.50	95.14
	TNR	100.00	100.00	100.00	99.95	99.80	98.98	97.55
	ACC	99.99	99.97	99.83	99.67	99.06	98.24	96.34
***QF_h*=75 *QF_a*=75**	**TPR**	99.75	98.85	97.63	95.04	91.83	86.65	80.84
	TNR	100.00	99.96	99.59	97.65	94.12	87.87	82.67
	ACC	99.88	99.40	98.61	96.34	92.97	87.26	81.76

In the experiments, we test for the case of $(\mathrm{QF}_h, \mathrm{QF}_a) = (50, 75)$ and $(\mathrm{QF}_h, \mathrm{QF}_a) = (75, 75)$. For each $(\mathrm{QF}_h, \mathrm{QF}_a)$ pair and each B, a number of 667 images (about 1/4 of the total number of the images) from the stego image set are randomly selected and used for training with their corresponding cover counterparts from the cover image set. The rest of the stego images and their corresponding cover images are used for testing. As mentioned before, stego images are considered to be in a positive class, while cover images in a negative one. Classification results are averaged by 20 times of randomly selecting the training and testing images. The results are demonstrated in Table II in terms of true positive rate, true negative rate, and accuracy rate. In addition, the standard deviations of the 20-time results are quite small. All are less than 1%, which means the reported results are quite stable.

The performance of the proposed method is highly effective for $\mathrm{QF}_h = 50$, even in the case of the B-block size being as large as $B = 15$, if B is known. Note that when $\mathrm{QF}_h = \mathrm{QF}_a = 75$, our method greatly outperforms the prior schemes reported in [27] and [30]. The better performance of the proposed mehtod on $(\mathrm{QF}_h, \mathrm{QF}_a) = (50, 75)$ than on $(\mathrm{QF}_h, \mathrm{QF}_a) = (75, 75)$ is mainly attributed to the fact that the embedded data at $\mathrm{QF}_h = 50$ are more robust than $\mathrm{QF}_h = 75$ to the JPEG compression at $\mathrm{QF}_a = 75$ in YASS data embedding. Hence, the statistical abnormality caused by data embedding is preserved better for detection. We also use $\mathrm{QF}_h = 55, 60, 65$, and 70 in our experiments and verify that the smaller the QF_h, the easier the detection.

C. Identifying B-Block Size if the Design Quality Factor is Known

When the B-block size is unknown but the design quality factor is known, a multiclass classification scheme can be applied to identifying the B-block size and also differentiating cover images and stego images. We test for the case of $(\mathrm{QF}_h, \mathrm{QF}_a) = (50, 75)$ and $(\mathrm{QF}_h, \mathrm{QF}_a) = (75, 75)$, respectively. Cover images are considered as a class while stego images with a particular $B (9 \le B \le 15)$ are considered as a particular class. As a result, we have a total number of eight different classes for each kind of QF_h in our experiments. A multiclass classifier is trained by an image set containing 667 randomly selected cover images and their various kinds of stego counterparts of different B. The testing is performed on the remaining images. The 20-time-averaged detection results are shown in Tables III and IV.

It is not surprising that our method does a great job of identifying the B-block size, especially for the case of $(\mathrm{QF}_h, \mathrm{QF}_a) = (50, 75)$. When the steganalytic method is tested for the case of $(\mathrm{QF}_h, \mathrm{QF}_a) = (75, 75)$, its performance drops a little for large B-block sizes. The reason is also due to the less obvious data embedding artifacts by a larger QF_h and more random embedding locations by a larger B.

D. Identifying Design Quality Factor if B-Block Size is Known

The proposed method can be used to identify the design quality factor QF_h when B is known. We test on six kinds of QF_h, ranging from 50 to 75 with a step size of 5. Stego images are embedded with the B-block size of $B = 9$. A multiclass classifier is trained by an image set containing 667 randomly selected cover images and their various kinds of stego counterparts of different QF_h. The testing is conducted on the remaining images. Table V shows the averaged classification results.

The performance of the method for $B = 9$ is very encouraging. All cover images can be correctly identified whereas stego images with different QF_h can be discriminated with high confidence. The outstanding accomplishment is owed to the features $\gamma_d^B (d \in \{1, 2, \ldots, 9\})$, since they can be used to identify double JPEG compression [41], as explained in Section III-E.

TABLE III
MULTICLASS CLASSIFICATION RESULTS (IN PERCENTAGE) FOR IDENTIFYING B IN THE CASE OF $\mathrm{QF}_h = 50$ AND $\mathrm{QF}_a = 75$

Predicted / Actual	Cover	Stego (*B*=9)	Stego (*B*=10)	Stego (*B*=11)	Stego (*B*=12)	Stego (*B*=13)	Stego (*B*=14)	Stego (*B*=15)
Cover	**96.43**	0.00	0.00	0.00	0.05	0.17	0.92	2.43
Stego(*B*=9)	0.02	**99.98**	0.00	0.00	0.00	0.00	0.00	0.00
Stego(*B*=10)	0.02	0.00	**99.95**	0.00	0.00	0.00	0.00	0.03
Stego(*B*=11)	0.35	0.00	0.00	**99.65**	0.00	0.00	0.00	0.00
Stego(*B*=12)	0.62	0.00	0.00	0.00	**99.36**	0.00	0.00	0.02
Stego(*B*=13)	1.58	0.00	0.00	0.00	0.00	**98.27**	0.01	0.13
Stego(*B*=14)	2.44	0.00	0.00	0.00	0.00	0.02	**97.32**	0.23
Stego(*B*=15)	4.74	0.00	0.00	0.00	0.00	0.04	0.27	**94.95**

TABLE IV
MULTICLASS CLASSIFICATION RESULTS (IN PERCENTAGE) FOR IDENTIFYING B IN THE CASE OF $\mathrm{QF}_h = 75$ AND $\mathrm{QF}_a = 75$

Predicted / Actual	Cover	Stego (*B*=9)	Stego (*B*=10)	Stego (*B*=11)	Stego (*B*=12)	Stego (*B*=13)	Stego (*B*=14)	Stego (*B*=15)
Cover	**69.83**	0.00	0.03	0.28	1.62	4.40	9.50	14.34
Stego(*B*=9)	0.13	**99.70**	0.00	0.00	0.03	0.04	0.00	0.10
Stego(*B*=10)	0.56	0.00	**98.59**	0.00	0.03	0.12	0.28	0.44
Stego(*B*=11)	1.67	0.00	0.00	**96.86**	0.07	0.43	0.50	0.48
Stego(*B*=12)	3.17	0.00	0.01	0.04	**93.38**	0.77	0.93	1.71
Stego(*B*=13)	6.20	0.00	0.00	0.15	0.48	**88.95**	1.74	2.48
Stego(*B*=14)	10.21	0.00	0.02	0.17	0.80	1.84	**82.25**	4.72
Stego(*B*=15)	15.72	0.00	0.02	0.19	1.14	2.31	5.36	**75.27**

TABLE V
MULTICLASS CLASSIFICATION RESULTS (IN PERCENTAGE) FOR IDENTIFYING QF_h IN THE CASE OF $\mathrm{QF}_a = 75$ AND $B = 9$

Predicted / Actual		Cover	Stego					
			***QF**_h*=50	***QF**_h*=55	***QF**_h*=60	***QF**_h*=65	***QF**_h*=70	***QF**_h*=75
Cover		**100.0**	0.00	0.00	0.00	0.00	0.00	0.00
Stego	***QF**$_h$*=50	0.00	**98.60**	1.34	0.00	0.00	0.00	0.07
	***QF**$_h$*=55	0.01	0.30	**99.30**	0.00	0.00	0.16	0.24
	***QF**$_h$*=60	0.02	0.00	0.00	**94.97**	2.01	2.32	0.68
	***QF**$_h$*=65	0.03	0.00	0.00	0.86	**93.00**	5.38	0.75
	***QF**$_h$*=70	0.02	0.00	0.00	0.60	3.31	**84.89**	11.18
	***QF**$_h$*=75	0.25	0.00	0.00	0.05	0.06	7.66	**91.99**

The experiments are also conducted on stego images with $B = 12$ and $B = 15$, of which the results are reported in Tables VI and VII. The rate of correct identification drops as B increases. The QF_h of stego images with $B = 15$ are seldom correctly identified. However, stego images are unlikely to be identified as cover images and vice versa for cover images. This is because the feature pattern of a stego image is different from that of a cover image, regardless of the QF_h. Hence the multiclass classifier is still reliable for differentiating stego images from cover images only.

E. Classifying Cover Images and Stego Images if B-Block Size and Design Quality Factor are Unknown

When the B-block size B and the design quality factor QF_h are unknown at the same time, it is impractical to regard each combination of B and QF_h as a class and use a multiclass classifier for classification. This is because the amount of the constructed two-classifiers for the multiclass classifier will be too large. Since the primary goal of steganalysis is to discriminate stego images from cover images, to this end we simply ignore identifying the unknown QF_h, just focusing on finding out the stego images and identifying their B-block size. In this scenario, stego images with the same B, even with different QF_h, are considered as one class. We use cover images and stego images with $(\mathrm{QF}_h, \mathrm{QF}_a) = (75, 75)$ for training a multiclass classifier. Then the trained classifier is capable of identifying stego images with the same B, even if the QF_h may be different.

In the experiments, a multiclass classifier is trained by 667 cover images and their seven kinds of stego images, corresponding to different B $(9 \leq B \leq 15)$. Note that they all use $(\mathrm{QF}_h, \mathrm{QF}_a) = (75, 75)$. The testing images consist of 2000 cover images and their corresponding stego images with B ranging from 9 to 15 and with QF_h ranging from 50 to 75 (with a step size of 5). In other words, for each B, there are 6×2000 stego images for testing. The averaged results are demonstrated in Table VIII.

It can be seen that the classification results in Table VIII for identifying B are less accurate than those in Table III but more accurate than those in Table IV. The reason can be explained by Fig. 4, which demonstrates the averaged difference of the extracted steganalytic features between stego images and cover

TABLE VI
MULTICLASS CLASSIFICATION RESULTS (IN PERCENTAGE) FOR IDENTIFYING QF_h IN THE CASE OF $\mathrm{QF}_a = 75$ AND $B = 12$

Predicted Actual		Cover	Stego					
			QF_h=50	QF_h=55	QF_h=60	QF_h=65	QF_h=70	QF_h=75
Cover		**97.64**	0.00	0.00	0.04	0.03	0.06	2.24
Stego	QF_h=50	0.02	**78.48**	19.97	0.48	0.32	0.45	0.28
	QF_h=55	0.03	20.38	**61.99**	10.94	3.28	2.34	1.05
	QF_h=60	0.15	0.29	13.11	**60.66**	18.05	5.88	1.87
	QF_h=65	0.51	0.21	4.07	18.36	**54.41**	18.02	4.44
	QF_h=70	1.76	0.09	1.50	5.85	18.77	**47.86**	24.17
	QF_h=75	4.99	0.06	0.40	0.96	2.86	20.48	**70.26**

TABLE VII
MULTICLASS CLASSIFICATION RESULTS (IN PERCENTAGE) FOR IDENTIFYING QF_h IN THE CASE OF $\mathrm{QF}_a = 75$ AND $B = 15$

Predicted Actual		Cover	Stego					
			QF_h=50	QF_h=55	QF_h=60	QF_h=65	QF_h=70	QF_h=75
Cover		**81.52**	0.38	0.53	1.05	1.65	3.46	11.40
Stego	QF_h=50	0.44	**63.33**	20.70	7.49	3.52	2.90	1.62
	QF_h=55	1.13	34.70	**27.65**	17.60	9.13	5.83	3.97
	QF_h=60	2.07	12.51	20.22	**30.34**	18.11	9.74	7.01
	QF_h=65	4.33	6.40	12.13	21.35	**26.69**	16.74	12.36
	QF_h=70	9.71	3.76	6.60	11.31	19.03	**24.06**	25.54
	QF_h=75	17.96	1.95	3.08	5.67	10.44	20.05	**40.86**

TABLE VIII
MULTICLASS CLASSIFICATION RESULTS (IN PERCENTAGE) FOR IDENTIFYING B IN THE CASE OF UNKNOWN QF_h (FROM 50 TO 75) AND $\mathrm{QF}_a = 75$

Predicted Actual	Cover	Stego (*B*=9)	Stego (*B*=10)	Stego (*B*=11)	Stego (*B*=12)	Stego (*B*=13)	Stego (*B*=14)	Stego (*B*=15)
Cover	**69.68**	0.00	0.02	0.34	1.69	4.22	9.34	14.71
Stego(*B*=9)	0.04	**99.93**	0.00	0.00	0.00	0.01	0.00	0.02
Stego(*B*=10)	0.16	0.00	**99.67**	0.00	0.01	0.04	0.05	0.08
Stego(*B*=11)	0.37	0.00	0.00	**99.34**	0.02	0.09	0.08	0.09
Stego(*B*=12)	0.79	0.00	0.00	0.01	**98.38**	0.19	0.26	0.37
Stego(*B*=13)	1.65	0.00	0.00	0.07	0.26	**96.74**	0.60	0.69
Stego(*B*=14)	2.97	0.00	0.01	0.10	0.41	0.68	**94.16**	1.67
Stego(*B*=15)	5.26	0.00	0.03	0.10	0.76	1.27	2.62	**89.96**

images. For demonstration purpose, we have rearranged the features into three groups in Fig. 4. The first group contains a total number of 63 features, that is, γ_d^B $(d \in \{1, 2, \ldots, 9\}, B \in \{9, 10, \ldots, 15\})$. The second group contains the features η_{t_1,t_2}^B $(t_1, t_2 \in \{-1, 0, 1\}, B \in \{9, 10, \ldots, 15\})$, with the feature sequential number from 64 to 126 in Fig. 4. The last 14 features are α^B and β^B $(B \in \{9, 10, \ldots, 15\})$. For a given B, it can be observed from Fig. 4 that the shape of feature pattern of the first group is determined by QF_h. Specifically, the peak location of the first feature pattern group is different under the different QF_h. Such a phenomenon justifies that the proposed features are capable of identifying QF_h. It can also be observed that the shape of the feature pattern of the second and the third group is similar under different QF_h. But its magnitude increases as QF_h decreases. Therefore, when we use stego images with $\mathrm{QF}_h = 75$ for training a classifier, testing stego images with a smaller QF_h would be more recognizable using such a classifier due to the strengthened magnitudes of the second and third group of the features. As a result, the overall performance reported in Table VIII will be between that in Table III and that in Table IV.

Note that once the B of the stego image is correctly identified, we can use the multiclass classifier described in Section IV-D to further identify the QF_h. Due to the limited space and plenty of combinations of B and QF_h, we do not report the detailed results here. It can be expected that the QF_h of a stego image with a small B can be identified more reliably than that with a large B.

V. DISCUSSION

A. Defeating Further Strategies of YASS

Some further strategies have been proposed in the literature [27]–[29] to enhance the YASS algorithm's capability, such as the data embedding rate and the security level. In this part, we investigate how the proposed features can be adapted to defeat these further strategies.

In the YASS algorithm, as the B-block size increases, the data embedding rate decreases. One further strategy to improve the data embedding rate is to use more than one H-block per B-block [27]. For example, using n^2 H-blocks in one B-block with the size $B = (n \times 8 + 1)$ is very effective [27]. As the parameter n increases, the B-block size B increases, and the data embedding rate will increase. The proposed method can be adapted to work under this case. Denote the pixel elements in a B-block as $b_{s,t}$ $(s, t \in \{0, 1, \ldots, B-1\})$. The possible origin of

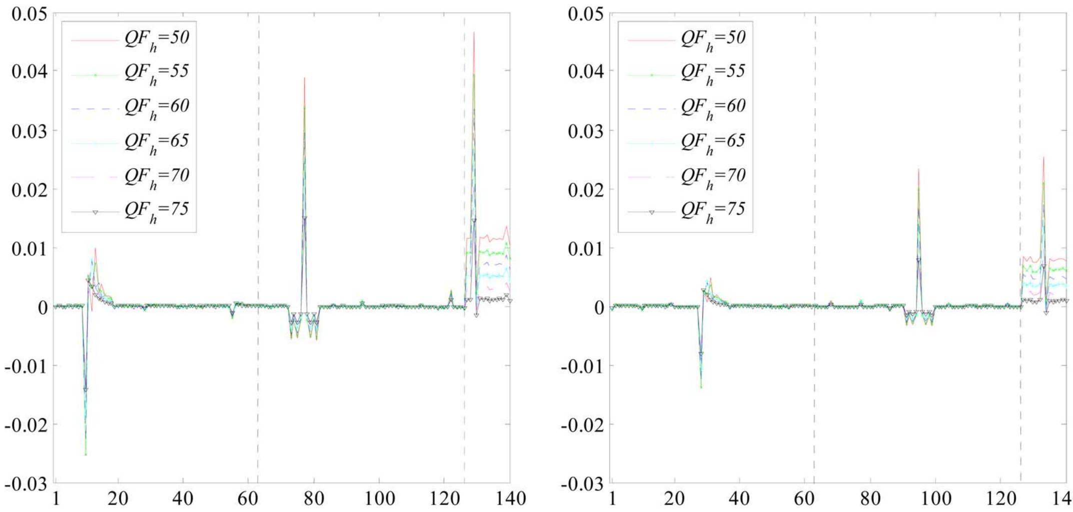

Fig. 4. Averaged feature difference between stego images (left: $B = 10$, $\mathrm{QF}_a = 75$; right: $B = 12$, $\mathrm{QF}_a = 75$) and cover images. The vertical axis is the magnitude of the feature. The horizontal axis is the sequential number of the feature. The features are arranged and divided into three groups. The first group with the sequential number from 1 to 63 contains the features of the probabilities of the first significant digits of requantized coefficients. The second group with the feature sequential number from 64 to 126 contains the features of the joint probabilities of requantized coefficients in neighboring F-blocks. The last 14 features correspond to the features of the frequencies of zero requantized coefficients.

an H-block is limited to coincide with $b_{m'\times 8+u',n'\times 8+v'}$, where $m', n' \in \{0, 1, \ldots, n-1\}$ and $u', v' \in \{0, 1\}$. As a result, the possible locations of H-blocks and the impossible locations of H-blocks can be obtained to form an updated SO-domain. Then the proposed steganalytic features can be extracted from the updated SO-domain for attacking.

Another further strategy is to use different designed quality factors for different H-blocks in data embedding [28]. A notable advantage of our proposed feature extraction process is that it does not rely on knowing the designed quality factor. It has been demonstrated that the proposed steganalytic method works even under the situation of lacking the prior knowledge of the B-block size and the designed quality factor. Therefore, even if the designed quality factors are not constant, our proposed method is still expected to work, especially when the averaged value of the designed quality factors is much smaller than the advertised quality factor. We can use stego images with a constant designed quality factor $\mathrm{QF}_h = \mathrm{QF}_a$ for training a multiclass classifier as described in Section IV-E. Then it can be generalized to detect stego images with varying designed quality factors as well as to identify the B-block size. From our preliminary experiments, we find out that the smaller the averaged value of the designed quality factors, the easier the stego images to be detected.

In [29], a method to estimate the redundant factor q of the RA codes at decoder's side has been proposed. Knowing the correct q can help maximize the data embedding rate. Since no fundamental change has been made in the data embedding operation, the output stego image is the same as the one output by the YASS algorithm. Therefore, the proposed method which is applicable to detect the YASS algorithm can be directly applied in this case.

If a strategy, as suggested in [28], uses iterative hiding to improve robustness of the embedded data and thus increasing the data embedding rate, the proposed method is expected to work better because embedding artifacts will be strengthened in H-blocks due to the iterative data embedding.

B. *Steganalyzing as the Origin of the B-Block Grid is Randomized*

Since the B-blocks in YASS are consecutive and nonoverlapping, they form the so-called *B-block grid*. Define the pixel element at the upper-left corner of the B-block grid as the origin of the B-block grid. While the locations of H-blocks within B-blocks are randomized in order to resist blind steganalytic methods, the origin of the B-block grid remains fixed at the upper-left corner of the image 2-D array in YASS [27]–[29]. That is, the origin of the B-block grid coincides with the origin of the image 2-D array. If the origin of the B-block grid is randomized, the task of steganalysis is more complicated. In this part, we examine how our proposed steganalytic method works when the location of the origin of the B-block grid is unknown.

Clearly, under such a circumstance, the strategy of our proposed method should have some change accordingly. Denote an image by $\boldsymbol{I}$ and its pixel element at position (x, y) by $\boldsymbol{I}(P_{x,y})$. Assume the image has been embedded by the YASS algorithm with B-block size B and the origin of the B-block grid is located at $\boldsymbol{I}(P_{a,b})$. We crop the image by s rows and t columns. Therefore, the origin of the cropped image is at $\boldsymbol{I}(P_{s,t})$ in the coordinate system of the original image $\boldsymbol{I}$. Denote the cropped image by $\boldsymbol{I}_{s,t}$ and divide it into $B \times B$ grid from its origin. Apparently, the B-block grid of $\boldsymbol{I}_{s,t}$ will overlap the B-block grid of $\boldsymbol{I}$ if both of the following two conditions are satisfied:

$$\mathrm{mod}(s, B) = \mathrm{mod}(a, B) \tag{9}$$

$$\mathrm{mod}(t, B) = \mathrm{mod}(b, B). \tag{10}$$

Obviously, after at most B^2 times of searching, one can find out the solution of s and t in (9) and (10) for a given pair of a and b. Accordingly, the strategy of our proposed method can be changed to use a two-class classifier to perform classification on the cropped image $I_{s,t} (s \in \{0, 1, \ldots, B-1\}, t \in \{0, 1, \ldots, B-1\})$. In this way, it can identify if each cropped image is data-embedded by the YASS algorithm. Therefore, an

TABLE IX
DETECTION RESULTS (IN PERCENTAGE) FOR CLASSIFYING COVER IMAGES AND STEGO IMAGES WITH TWO-CLASS CLASSIFIER WHEN THE ORIGIN OF THE B-BLOCK GRID, THE B-BLOCK SIZE, AND THE DESIGN QUALITY FACTOR ARE UNKNOWN. THE SECOND ROW SECOND COLUMN REPORTS THE OVERALL DETECTION RATE FOR ALL STEGO IMAGES. THE FOURTH ROW REPORTS THE DETECTION RATE FOR STEGO IMAGES WITH EACH SPECIFIC B-BLOCK SIZE

Cover	**Stego**						
77.6	63.9						
	***B*=9**	***B*=10**	***B*=11**	***B*=12**	***B*=13**	***B*=14**	***B*=15**
	86.4	77.6	76.8	63.2	60.8	48.0	37.2

image under scrutiny will be under the classification for a total number of B^2 times. Suppose we have a perfect two-class classifier (with a zero false positive rate and a zero false negative rate), the B^2-time classification results for a cover should be all negative and that for a stego image should be positive for at least one time. Thus a detection criterion can be set as "a stego image should get at least one positive result from the B^2-time classification." But a practical classifier usually has a nonzero false positive rate and a nonzero false negative rate. Therefore, for practical considerations, one feasible approach is that we train a classifier with a zero false positive rate by lowering the true positive rate. Hence a cover image under B^2-time classification will likely be classified in a negative class. Since we aim to detect a stego image with $9 \leq B \leq 15$, each image is under $15^2 = 225$ times of classification. If one of the classification results is positive, the image is labeled as a stego image, otherwise a cover image.

In the simulation, we randomly select a number of 667 cover images and their seven types of stego images, corresponding to seven different B-block sizes, to train a two-class classifier. Here the stego images with different B are all considered in a positive class. The origin of the B-block grid of each training stego image coincides with the origin of the image 2-D array. The training stego images use $(\mathrm{QF}_h, \mathrm{QF}_a) = (75, 75)$ and thus the trained classifier is capable of detecting stego images with different QF_h, due to the same reason as explained in Section IV-E. We adjust the bias of the two-class classifier to have a zero false positive rate for the training images. The testing image set contains 2000 images with different contents. Among them, a number of 250 are cover images and others are stego images. There are a number of 250 stego images for each kind of $B (9 \leq B \leq 15)$. In the experimental works, the QF_h of the testing stego images are uniformly distributed from 50 to 75, with a step of 5. The origin of the B-block grid of the testing image is selected to be randomly located in the area where the first 20 rows intersect the first 20 columns of the image. The process of steganalytic feature extraction and classification is performed for each cropped version of a testing image. If at least one of the $15^2 = 225$ testing results is positive, the image is considered as a stego image.

The overall true negative rate is 77.6% and the overall true positive rate is 63.9%, better than random guessing. We show the detection results for each particular B-block size in Table IX. It can be seen that the overall performance is dragged down mainly by the poor performance as B is large. It confirms that using a larger B is less likely to be detected. The reason is due in part to the more randomized locations of the H-blocks and in part to the rather low data embedding rate when B is large. Specifically, readers can refer to Table I which shows that the averaged data embedding rate is as low as around, and even below, 0.05 bpac as the B-block size B is equal to 15 or 14 when $\mathrm{QF}_h = 75$ or 70 and $\mathrm{QF}_a = 75$. At such a low data embedding rate other existing steganographic schemes, say, some proposed in [13]–[17], are also undetectable as reported in [22], [23], and [26]. It indicates YASS may lose its superiority at such a low data embedding rate.

C. Limitation of the Proposed Method

From the experimental results, we can conclude that when the B-block size of the YASS algorithm gets large, the performance of the proposed method will drop. This is due to the fact that the embedding locations of the H-blocks are more randomized as well as the data embedding rate of the YASS algorithm decreases. Thus, the proposed SO-domain will access the data embedding domain of the YASS algorithm with a smaller probability.

It is clear that the proposed steganalytic technique is specifically designed for detecting YASS [27]–[29]. It shows us a feasible method to construct an effective SO-domain by finding two kinds of sets, namely, one kind of set that can maximize the probability of accessing the data embedding domain and another kind of set that can minimize the probability of accessing the data embedding domain.

If a YASS-like scheme, which may be referred to as the next generation of YASS, is designed in the future to further randomize the data hiding locations while keeping an appealing data embedding rate, the proposed method may be restricted or even fail to construct an SO-domain with the aforementioned two kinds of sets. Therefore, a new SO-domain should be investigated, and some new features may be extracted from the new SO-domain for this new steganography. But the proposed method is still beneficial for steganalysis by showing a possible direction on the selection of the new SO-domain. That is, maximizing the probability of accessing the data embedding domain.

VI. CONCLUSION

In this paper, YASS [27]–[29] has been under investigation. A specific method for detecting the YASS algorithm has been proposed and possible ways to defeat the further strategies of YASS have been discussed. The contributions made in this paper are summarized as follows.

First, a new term, steganalytic observation domain (SO-domain), is defined in this paper as the domain from which the statistical features are extracted for steganalysis. We have constructed an SO-domain, which has been proved to be effective in detecting the YASS algorithm. The proposed SO-domain partially accesses the data embedding domain of the YASS algorithm and it catches the very defect that the locations of the H-blocks are not randomized enough. On the one hand, it is clear to see that the selection of an effective SO-domain is critical for steganalysis. On the other hand, the concept of the SO-domain provides insight on enhancing the security of a steganographic scheme. Steganography in the next generation should avoid leaking information about the embedding locations to a warden.

Second, we extract the steganalytic features with high efficiency for steganalysis of YASS while maintaining a low dimensionality. The features of the frequency of zero requantized coefficients (first-order statistics) are proposed with innovation. They are highly effective in discriminating stego images from cover images, since they are designed purposely to capture the QIM artifacts which are caused by YASS. Inspired by the ideas from [25] and [41], we proposed the features of the probabilities of the first significant digits of requantized coefficients (first-order statistics) and the features of the joint probabilities of requantized coefficients in neighboring F-blocks (second-order statistics). Not only can the features be used to enhance the performance of detecting stego images, but also do they help identify some parameters in embedding.

Third, combining the steganalytic features with a two-class classification scheme, we differentiate cover images and stego images embedded by the YASS algorithm with high accuracy, if the B-block size and the design quality factor of the stego image are known as prior knowledge. The performance outperforms the results reported in prior arts [27], [30] by a large margin. When the designed quality factor is known but the B-block size is unknown, a multiclass classification strategy can be employed to resolve the problem of identifying B-block size easily. When the B-block size in YASS is known but the design quality factor is unknown, the design quality factor can be identified by using the proposed features with a multiclass classifier trained by cover images and stego images with different design quality factors. If the B-block size and the designed quality factor are unavailable at the same time, a specially designed multiclass strategy is still able to identify the B-block size with a satisfactory accuracy. Identifying the embedding parameters will give a warden more options to deter the covert communication. Experimental results demonstrate the effectiveness of the proposed method.

Fourth, we have discussed how to generalize the proposed steganalytic method to counter the further strategies of YASS. For some strategies, the proposed method targeted to detect the YASS algorithm can be employed to defeat these strategies without fundamental change. For some other strategies, the SO-domain may need to be changed accordingly. In addition, shifting the origin of the B-block grid is not taken into consideration in YASS. However, randomizing the origin of the B-block grid is a possible way to increase the randomness of the locations of H-blocks, and therefore enhance the security of YASS. A multiple-time-applied two-class classification strategy with the proposed steganalytic features is still able to differentiate cover images from stego images under the circumstance that the origin of the B-block grid is arbitrarily located and the B-block size and the design quality factor are both unknown.

In summary, the proposed method sheds light on the insecure aspect of YASS. It should stimulate the design of a more secure steganography and a more capable steganalysis.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable comments.

References

[1] S. Wang and H. Wang, "Cyber warfare: Steganography vs. steganalysis," *Commun. ACM*, vol. 47, no. 10, pp. 76–82, 2004.

[2] C. Cachin, "An information-theoretic model for steganography," in *Proc. 2nd Int. Workshop Information Hiding*, Portland, OR, Apr. 1998, pp. 306–318.

[3] R. Chandramouli, M. Kharrazi, and N. Memon, "Image steganography and steganalysis: Concepts and practices," in *Proc. 2nd Int. Workshop Digital Watermarking*, Seoul, Korea, Oct. 2003, pp. 35–49.

[4] X. Luo, D. Wang, P. Wang, and F. Liu, "A review on blind detection for image steganography," *Signal Process.*, vol. 88, no. 9, pp. 2138–2157, 2008.

[5] H. Farid and S. Lyu, "Higher-order wavelet statistics and their application to digital forensics," in *Proc. IEEE Workshop Statistical Analysis in Computer Vision*, Madison, WI, 2003.

[6] S. Bayram, I. Avcibas, B. Sankur, and N. Memon, "Image manipulation detection," *J. Electron. Imaging*, vol. 15, no. 4, pp. 1–17, 2006.

[7] Y. Q. Shi, C. Chen, and W. Chen, "A natural image model approach to splicing detection," in *Proc. ACM Workshop on Multimedia and Security*, Dallas, TX, Sep. 2007, pp. 51–62.

[8] Y. Q. Shi, C. Chen, and G. Xuan, "Steganalysis versus splicing detection," in *Proc. 7th Int. Workshop on Digital Watermarking*, Guangzhou, China, Dec. 2007, pp. 158–172.

[9] G. K. Wallace, "The JPEG still picture compression standard," *Commun. ACM*, vol. 34, no. 4, pp. 30–44, 1991.

[10] *JSteg*. [Online]. Available: http://zooid.org/~paul/crypto/jsteg/

[11] S. Hetzl and P. Mutzel, "A graph theoretic approach to steganography," in *Proc. 9th Int. Conf. Communications and Multimedia Security*, Salzburg, Austria, 2005, pp. 119–128.

[12] N. Provos, "Defending against statistical steganalysis," in *Proc. 10th USENIX Security Symp.*, Washington, DC, Aug. 2001, pp. 323–325.

[13] P. Sallee, "Model-based steganography," in *Proc. 2nd Int. Workshop Digital Watermarking*, Seoul, Korea, Oct. 2003, pp. 154–167.

[14] A. Sarkar, K. Solanki, U. Madhow, S. Chandrasekaran, and B. S. Manjunath, "Secure steganography: Statistical restoration of the second order dependencies for improved security," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Honolulu, HI, Apr. 2007, pp. 277–280.

[15] A. Westfeld, "High capacity despite better steganalysis (F5—A steganographic algorithm)," in *Proc. 4th Int. Workshop on Information Hiding*, Pittsburgh, PA, Apr. 2001, pp. 289–302.

[16] J. Fridrich, M. Goljan, and D. Soukal, "Perturbed quantization steganography," *ACM Multimedia Security J.*, vol. 11, no. 2, pp. 98–107, 2005.

[17] Y. Kim, Z. Duric, and D. Richards, "Modified matrix encoding technique for minimal distortion steganography," in *Proc. 8th Int. Workshop Information Hiding*, Old Town Alexandria, VA, Jul. 2006, pp. 314–327.

[18] J. Fridrich, M. Goljan, and D. Hogea, "Steganalysis of JPEG images: Breaking the F5 algorithm," in *Proc. 5th Int. Workshop Information Hiding*, Noordwijkerhout, The Netherlands, Oct. 2002, pp. 310–323.

[19] J. Fridrich, M. Goljan, and D. Hogea, "Attacking the outguess," in *Proc. ACM Workshop Multimedia and Security*, Juan-les-Pins, France, Dec. 2002, pp. 3–6.

[20] R. Bohme and A. Westfeld, "Breaking Cauchy model-based JPEG steganography with first order statistics," in *Proc. 9th Eur. Symp. Research Computer Security (ESORICS)*, Sophia Antipolis, France, Sep. 2004, pp. 125–140.

[21] J. Fridrich, "Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes," in *Proc. 6th Int. Workshop Information Hiding*, Toronto, ON, Canada, May 2004, pp. 67–81.

[22] Y. Q. Shi, C. Chen, and W. Chen, "A Markov process based approach to effective attacking JPEG steganography," in *Proc. 8th Int. Workshop Information Hiding*, Old Town Alexandria, VA, Jul. 2006, pp. 249–264.

[23] T. Pevny and J. Fridrich, "Merging Markov and DCT features for multi-class JPEG steganalysis," in *Proc. Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX SPIE*, San Jose, CA, Jan. 2007, pp. 1–13.

[24] D. Fu, Y. Q. Shi, D. Zou, and G. Xuan, "JPEG steganalysis using empirical transition matrix in block DCT domain," in *Proc. IEEE 8th Workshop Multimedia Signal Processing*, BC, Canada, Oct. 2006, pp. 310–313.

[25] C. Chen and Y. Q. Shi, "JPEG image steganalysis utilizing both intrablock and inter-block correlations," in *Proc. IEEE Int. Symp. Circuits and Systems*, Seattle, WA, May 2008, pp. 3029–3032.

[26] J. Fridrich, T. Pevný, and J. Kodovský, "Statistically undetectable JPEG steganography: Dead ends challenges, and opportunities," in *Proc. 9th ACM Workshop Multimedia and Security*, Dallas, TX, Sep. 2007, pp. 3–14.

[27] K. Solanki, A. Sarkar, and B. S. Manjunath, "YASS: Yet another steganographic scheme that resists blind steganalysis," in *Proc. 9th Int. Workshop Information Hiding*, Saint Malo, France, Jun. 2007, pp. 16–31.

[28] A. Sarkar, K. Solanki, and B. S. Manjunath, "Further study on YASS: Steganography based on randomized embedding to resist blind steganalysis," in *Proc. Electronic Imaging, Security, Forensics, Steganography, and Watermarking of Multimedia Contents X, SPIE*, San Jose, CA, Jan. 2008.

[29] A. Sarkar, L. Nataraj, B. S. Manjunath, and U. Madhow, "Estimation of optimum coding redundancy and frequency domain analysis of attacks for YASS—A randomized block based hiding scheme," in *Proc. IEEE Int. Conf. Image Processing*, San Diego, CA, Oct. 2008, pp. 1292–1295.

[30] J. Kodovský and J. Fridrich, "Influence of embedding strategies on security of steganographic methods in the JPEG domain," in *Proc. Electronic Imaging, Security, Forensics, Steganography, and Watermarking of Multimedia Contents X, SPIE*, San Jose, CA, Jan. 2008.

[31] S. Lyu and H. Farid, "Detecting hidden messages using higher-order statistics and support vector machines," in *Proc. 5th Int. Workshop Information Hiding*, Noordwijkerhout, The Netherlands, Oct. 2002, pp. 340–354.

[32] G. Xuan, Y. Q. Shi, J. Gao, D. Zou, C. Yang, 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. 7th Int. Workshop Information Hiding*, Barcelona, Spain, Jun. 2005, pp. 262–277.

[33] B. Li, Y. Q. Shi, and J. Huang, "Steganalysis of YASS," in *Proc. 10th ACM Workshop Multimedia and Security*, Sep. 2008, pp. 139–148.

[34] B. Chen and G. W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.

[35] K. Solanki, N. Jacobsen, U. Madhow, B. S. Manjunath, and S. Chandrasekaran, "Robust image-adaptive data hiding based on erasure and error correction," *IEEE Trans. Image Process.*, vol. 13, no. 12, pp. 1627–1639, Dec. 2004.

[36] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[37] C.-E. W. Sundberg, "Erasure and error decoding for semiconductor memories," *IEEE Trans. Comput.*, vol. C-21, no. 8, pp. 696–705, Aug. 1978.

[38] K. Sullivan, Z. Bi, U. Madhow, S. Chandrasekaran, and B. Manjunath, "Steganalysis of quantization index modulation data hiding," in *Proc. IEEE Int. Conf. Image Processing*, Singapore, Oct. 2004, vol. 2, pp. 1165–1168.

[39] H. Malik, K. P. Subbalakshmi, and R. Chandramouli, "Nonparametric steganalysis of QIM-based data hiding using kernel density estimation," in *Proc. 9th ACM Workshop Multimedia and Security*, Dallas, TX, Sep. 2007, pp. 149–160.

[40] T. Pevny and J. Fridrich, "Detection of double-compression in JPEG images for applications in steganography," *IEEE Trans. Inf. Security Forensics*, vol. 3, no. 2, pp. 247–258, Jun. 2008.

[41] B. Li, Y. Q. Shi, and J. Huang, "Detecting doubly compressed JPEG images by using mode based first digit features," in *Proc. IEEE Int. Workshop Multimedia Signal Processing*, Cairns, Queensland, Australia, Oct. 2008, pp. 730–735.

[42] D. Fu, Y. Q. Shi, and W. Su, "A generalized Benford's law for JPEG coefficients and its applications in image forensics," in *Proc. SPIE, Security, Steganography and Watermarking of Multimedia Contents IX*, San Jose, CA, Jan. 2007.

[43] A. Webb, *Statistical Pattern Recognition*, 2nd ed. Hoboken, NJ: Wiley, 2002.

[44] NRCS Photo Gallery [Online]. Available: http://photogallery.nrcs.usda.gov

Bin Li (S'07) received the B.E. degree in communications engineering and the Ph.D. degree in communications and information systems from Sun Yet-sen University, China, in 2004 and 2009, respectively.

He is currently a lecturer with the College of Information Engineering, Shenzhen University, China. From 2007 to 2008, he received a scholarship from the China Scholarship Council and conducted research as a visiting scholar in New Jersey Institute of Technology. His current research interests include multimedia signal processing, pattern recognition, and information security.

Jiwu Huang (M'98–SM'00) received the B.S. degree from Xidian University, China, in 1982, the M.S. degree from Tsinghua University, China, in 1987, and the Ph.D. degree from the Institute of Automation, Chinese Academy of Science, in 1998.

He is currently a Professor with the School of Information Science and Technology, Sun Yat-Sen University, Guangzhou, China. His current research interests include multimedia security and data hiding.

Dr. Huang serves as a member of IEEE CAS Society Technical Committee of Multimedia Systems and Applications and the chair of IEEE CAS Society Guangzhou chapter.

Yun Qing Shi (M'90–SM'93–F'05) received the B.S. and M.S. degrees from Shanghai Jiao Tong University, Shanghai, China, and the M.S. and Ph.D. degrees from the University of Pittsburgh, PA.

He has been with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, since 1987, and is now a professor there. His research interests include digital multimedia data hiding, steganaysis, forensics and information assurance, visual signal processing and communications, motion analysis, theory of multidimensional systems, and signal processing. He is an author/coauthor of more than 200 papers, a book, and four book chapters. He holds five U.S. patents and has an additional 30 U.S. patents pending.

Dr. Shi is the founding Editor-in-Chief of *LNCS Transactions on Data Hiding and Multimedia Security* (Springer), and an Associate Editor of *Journal on Multidimensional Systems and Signal Processing* (Springer). He served as an Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART II, an Associate Editor of *International Journal of Image and Graphics*, and a guest editor of special issues for a few journals, the technical program chair of ICME07, co-technical chair of IWDW06, 07, 09, MMSP05, and co-general chair of MMSP02. He is a member of a few IEEE technical committees.