Using Statistical Image Model for JPEG Steganography: Uniform Embedding Revisited

Linjie Guo, *Student Member, IEEE*, Jiangqun Ni, *Member, IEEE*, Wenkang Su, Chengpei Tang, and Yun-Qing Shi, *Fellow, IEEE*

***Abstract*— Uniform embedding was first introduced in 2012 for non-side-informed JPEG steganography, and then extended to the side-informed JPEG steganography in 2014. The idea behind uniform embedding is that, by uniformly spreading the embedding modifications to the quantized discrete cosine transform (DCT) coefficients of all possible magnitudes, the average changes of the first-order and the second-order statistics can be possibly minimized, which leads to less statistical detectability. The purpose of this paper is to refine the uniform embedding by considering the relative changes of statistical model for digital images, aiming to make the embedding modifications to be proportional to the coefficient of variation. Such a new strategy can be regarded as generalized uniform embedding in substantial sense. Compared with the original uniform embedding distortion (UED), the proposed method uses all the DCT coefficients (including the DC, zero, and non-zero AC coefficients) as the cover elements. We call the corresponding distortion function uniform embedding revisited distortion (UERD), which incorporates the complexities of both the DCT block and the DCT mode of each DCT coefficient (i.e., selection channel), and can be directly derived from the DCT domain. The effectiveness of the proposed scheme is verified with the evidence obtained from the exhaustive experiments using a popular steganalyzer with rich models on the BOSSbase database. The proposed UERD gains a significant performance improvement in terms of secure embedding capacity when compared with the original UED, and rivals the current state-of-the-art with much reduced computational complexity.**

***Index Terms*— JPEG steganography, minimal-distortion embedding, uniform embedding, distortion function design.**

Manuscript received March 31, 2015; revised July 1, 2015 and August 10, 2015; accepted August 17, 2015. Date of publication August 26, 2015; date of current version September 30, 2015. This work was supported in part by the Key Program through the Natural Science Foundation of Guangdong under Grant S2012020011114, in part by the National Research Foundation for the Doctoral Program of Higher Education of China under Grant 20120171110037, and in part by the National Natural Science Foundation of China under Grant 61379156 and Grant 60970145. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Patrick Bas. *(Corresponding author: Jiangqun Ni.)*

L. Guo is with the School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510006, China, and also with Huawei Technologies Company, Ltd., Shenzhen 518129, China (e-mail: g.linjie@gmail.com).

J. Ni is with the School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510006, China, and also with the State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China (e-mail: issjqni@mail.sysu.edu.cn).

W. Su is with the School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510006, China (e-mail: swk1004@163.com).

C. Tang is with the School of Engineering, Sun Yat-sen University, Guangzhou 510006, China (e-mail: tchengp@mail.sysu.edu.cn).

Y.-Q. Shi is with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102 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.2015.2473815

I. Introduction

STEGANOGRAPHY is the science and art of concealing a message within empirical cover media (image, audio or video etc.) [3]. Generally, a good steganographic method should have acceptable statistical imperceptibility and a sufficient payload, while these two objectives are generally conflicting with each other for a given algorithm. Therefore, the purpose of the steganographer is to lower the statistical detectability, i.e., to improve the security performance for a fixed payload. In modern steganography, numerous attempts have been made to achieve this purpose. Among them, preserving a chosen cover model has been proved to be a bad idea, while the most common and effective approach is minimizing a heuristically-defined embedding distortion for the empirical cover media. Such approach is also formulated as minimal distortion embedding framework.

The development of JPEG steganography benefits a lot from the application of channel coding, e.g., Hamming code for F5 [4], modified matrix encoding for MME [5] and BCH code for BCHopt [6], etc. In 2011, Filler *et al.* developed a breakthrough coding framework by utilizing the syndrome-trellis codes (STCs), which allows the steganographer to minimize an additive distortion function while embedding a near-maximal payload [8]. By letting the distortion of an element to be an infinite value, the framework can even prohibit modification of this element, thus allowing the implementation of the wet paper coding.

Over the past few years, the emerging JPEG steganographic schemes all focused on the design of the distortion function. Among them, distortion functions for JPEG steganography without any side-information are particularly desirable for practical applications, e.g., MOD [9], UNIWARD [10] and UED [2], even though UNIWARD and UED can also utilize the side-information. The distortion function of MOD was heuristically defined as a rich parametric model, which was then optimized to obtain the least detectability with respect to a selected feature set (cover model). Note the fact that the cover model involved in MOD is an incomplete statistical descriptor of the empirical source, MOD ended up being more detectable as the steganalyst can easily design a detector "outside the model" [11]. Despite the limitation with MOD [9], the effectiveness of its motivation, i.e. content-adaptivity, was demonstrated in a novel method called UNIWARD [10]. In [10], the distortion functions for spatial (S-UNI), JPEG (J-UNI) and side-informed JPEG (SI-UNI) domain are all derived from the wavelet domain. Unlike the conventional JPEG steganographic schemes which only embed the secret message

1556-6013 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

into non-zero AC coefficients, J-UNI and SI-UNI use all DCT coefficients (DCs, zero and non-zero ACs) as possible cover elements, and achieves so far the best security performance. However, the computational complexity of attaining distortion from the wavelet domain may be the major problem in implementation.

Under the circumstance that many content-adaptive image steganographic methods rigidly adhere to a given cover model, the idea of uniform embedding opens a new thought of JPEG steganography. In practice, the statistical detectability is measured as the detection results of the steganalysis, which often consists of a classifier and a feature set. It is exactly the feature set that significantly influences the detectability of the steganalysis. Most existing distortion functions are said to be heuristically defined, because one can hardly establish a direct relationship between the distortion function and the feature set. Guo et al. [2] provided a new approach for JPEG steganography, i.e., "spreading" the embedding modification uniformly to quantized DCT coefficients of all possible magnitudes. In this way, less statistical detectability can be achieved, owing to the reduction of the average changes of the first- and second-order statistics of DCT coefficients as a whole. By incorporating the uniform embedding (UED) strategy in the framework of minimal distortion embedding, UED substantially improved the previous state-of-the-art: nsF5 algorithm [7].

However, UED reported in [2] is only a preliminary attempt to implement the uniform embedding strategy. In this paper, we are trying to perfect the uniform embedding from the motivation to the actual distortion function. The newly proposed method takes into account the relative change of the statistical image model, which needs to be proportional to the "coefficient of variation", and can be regarded as the generalized uniform embedding. Compared with original UED, the proposed method utilizes all DCT coefficients (including the DC and zero AC coefficients) as the cover elements, and the corresponding distortion function is constructed based on the complexities of both the DCT block and the mode of each DCT coefficient (selection channel), which can be derived directly from the DCT domain. We call the new distortion function uniform embedding revisited distortion (UERD). It is showed that its computational complexity is almost negligible, which is more suitable for practical application. The effectiveness of the proposed scheme is verified with evidence obtained from exhaustive experiments using popular steganalyzer with rich models on the BOSSbase database [15]. The proposed UERD achieves considerable performance improvement in terms of secure embedding capacity when compared with the original UED, and has a comparable security performance with the current state-of-the-art UNIWARD.

The rest of this paper is organized as follows. In Section II, we propose the generalized uniform embedding strategy, and the corresponding distortion functions (UERD) for JPEG steganography with or without side-information of the original uncompressed image are presented in Section III. The new JPEG steganographic scheme is briefly described in Section IV, which is followed by the experimental results and analysis in Section V. Finally, the concluding remarks are summarized in Section VI.

II. Generalized Uniform Embedding Strategy

Nowadays, the most successful and practical approach to steganography is the minimal distortion embedding framework, which includes a tractably optimizable distortion function and a method for encoding the message to minimize the distortion. With the development of syndrome-trellis codes (STCs), which can reach the rate-distortion bound for additive distortion measure, the only thing left to the steganographers is the design of the distortion function. While most of the existing distortion functions are based on the complexity of image contents (both in spatial and JPEG domain) and built heuristically, Guo et al. [2] proposed to construct the distortion function based on the concept of uniform embedding, and the corresponding distortion function is called uniform embedding distortion (UED). Although UED substantially improved the security performance, we find there still exists possibility to further refine the uniform embedding strategy from the perspective of statistical image models. Below we will firstly take a brief review of the original uniform embedding strategy.

A. Review of the Original Uniform Embedding Strategy

By following the concept in the spirit of "spread spectrum communication", the motivation of the original uniform embedding strategy is to uniformly "spread" the embedding modification to the quantized DCT coefficients of all possible magnitudes. Specifically, let x denote the DCT coefficients, which are Laplacian distributed, the distortion function involved in UED should have the general form of $1/|x|$. Equiped with the corresponding distortion function, UED would result in possible minimal artifacts of the first- and second-order statistics for DCT coefficients as a whole. One can refer to [2] for detail.

The motivation of the original uniform embedding strategy can be easily understood from the perspective of the information theory, i.e., the maximum entropy principle for discrete source. However, for natural images, the tolerable relative changes of DCT coefficients with different values may differ from each other. Take the changes of the global histogram caused by data embedding as example, a similar change may be insignificant for the distribution of the a-valued coefficient, while it can be very serious for the one of the b-valued coefficient ($|a| \gg |b|$). This is the first problem with UED that must be carefully investigated again. Another one that is worth delving into is whether "uniform spread" of the modification really leads to the most difficult detection for the steganalysis, i.e., the best security performance.

B. Generalized Uniform Embedding Strategy

As mentioned in [2], the statistics of natural images indeed exhibit, to some extent, deviation away from their models of any kinds, which are what the potentials of natural images left for steganography. Fig. 1(a) and (b) show the mean value $\mu(x)$ and standard deviation $\sigma(x)$ of the histogram of DCT coefficients $p(x)$ over 10,000 JPEG images with $QF = 75$ from Bossbase [15]. All the experiments below are also implemented on the same database unless otherwise specified. As is shown in Fig. 1(a), $\mu(x)$ decreases with increasing $|x|$.

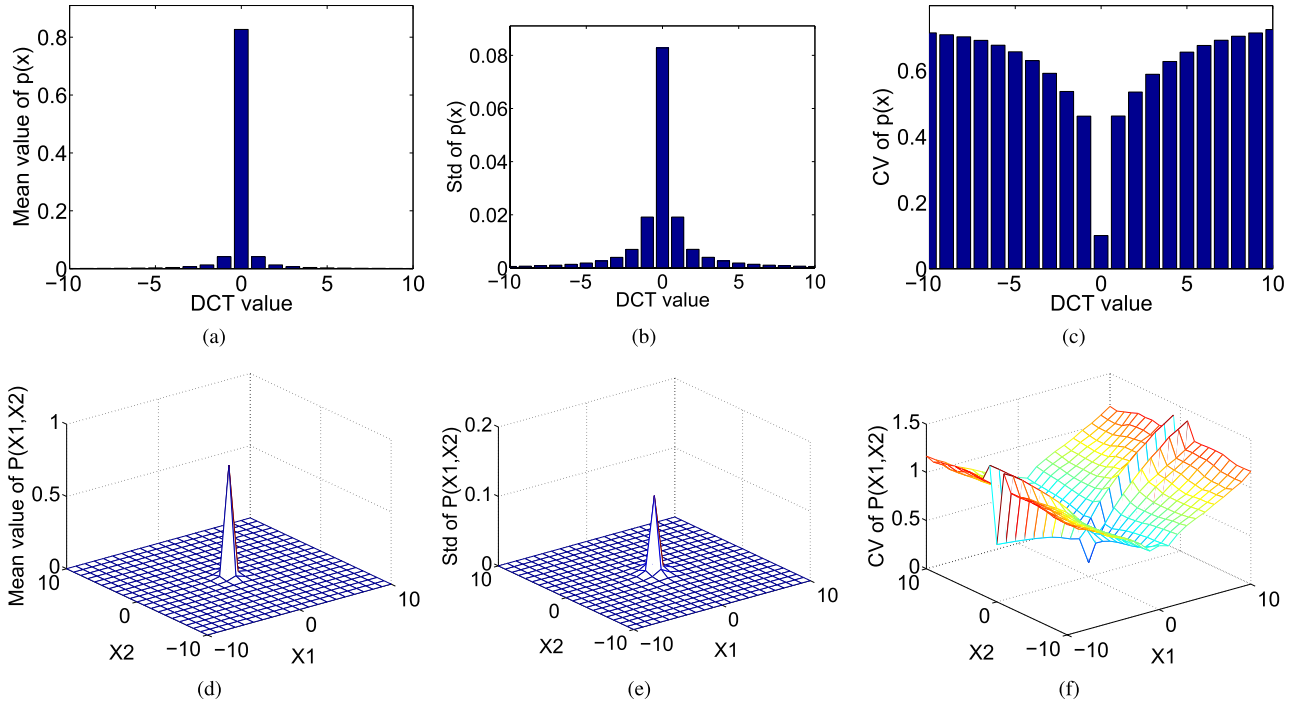


Fig. 1. (a)-(c) and (d)-(f) are the mean, standard deviation and CV of $p(x)$ and $P(X_1, X_2)$ over 10,000 JPEG image, respectively.

Therefore, when we study the changes of each bin after data embedding, the **relative change** will be a more appropriate measurement to figure out the real impact of each bin of the $p(x)$.

For steganalysis, a different model will provide different detection ability, furthermore, the detection ability of different bin in a model are also different. Although $\sigma(x)$ decreases with the increase of $|x|$ (Fig. 1(b)), we could not conclude that the detection ability of bin 0 is the worst (for the same relative change). Considering the fact that $\mu(x)$ also decreases with increasing $|x|$, a relative measurement will be more appropriate. In this paper, in order to investigate the detection ability of each bin of $p(x)$, we adopt the concept of **coefficient of variation** $CV(x)$ [12], which is defined as the ratio of the standard deviation to the mean:

$$CV(x) = \frac{\sigma(x)}{\mu(x)}. \tag{1}$$

For better comparison of the detection ability, an important assumption to make is that the relative change of each bin is the same. As is shown in Fig. 1(c), $CV(x)$ increases with the increase of $|x|$, which implies that, the detection ability of each bin of $p(x)$ decreases with the increase of $|x|$.

The similar effect is also observed with second-order statistics of natural images. Let $P(X_1, X_2)$ denote the co-occurrence matrix with offset $d = (0, 1)$. Fig. 1(d)-(f) show the mean $\mu(X_1, X_2)$, standard deviation $\sigma(X_1, X_2)$ and normalized *coefficient of variation* $CV(X_1, X_2)$ of $P(X_1, X_2)$, respectively. Obviously, $CV(X_1, X_2)$ also increases with the increase of $|X_1|+|X_2|$. In other words, the detection ability of each bin of $P(X_1, X_2)$ decreases with increasing $|X_1|+|X_2|$.

We believe that the similar effects can also be observed with other higher-order statistics. This indicates that uniformly spreading the modification to all possible bins does not necessarily imply the uniform detection ability of each element for a given model. Therefore, in order to achieve the best security performance, a reasonable choice is to uniformly spread the "detection ability" to each bin. To this end, we only need to control the relative change of each bin to be proportional to the $CV(x)$ or $CV(X_1, X_2)$, which is called generalized uniform embedding strategy.

C. Utilizing the Zero AC and DC Coefficients

For a long time, in order to make the embedding naturally content-adaptive, most JPEG steganographic schemes avoid to reduce the number of the zero AC coefficients and create new zero AC coefficients. With the generalized uniform embedding strategy, it is observed from Fig. 1(b) and (c) that the zero AC coefficients take up the largest proportions of the AC coefficients that can be modified. Recognizing that the zero AC coefficients can be used as the cover elements to increase the security performance, the key issue is to choose the appropriate zero AC coefficients for embedding. This is done by using the distortion function. Actually the zero AC coefficients in high frequency region have relatively small CVs as described later in Section III-A, and embedding message into these zero AC coefficients would also decrease the efficiency of JPEG compression. Therefore, those zero AC coefficients should be treated as the wet points in JPEG steganography. While the zero AC coefficients in low frequency region could indeed be explored for message embedding, in this paper, zero AC coefficients and other non-zero AC coefficients are treated equally. Whether a zero AC coefficient is a wet point or not will depend on a unified distortion function.

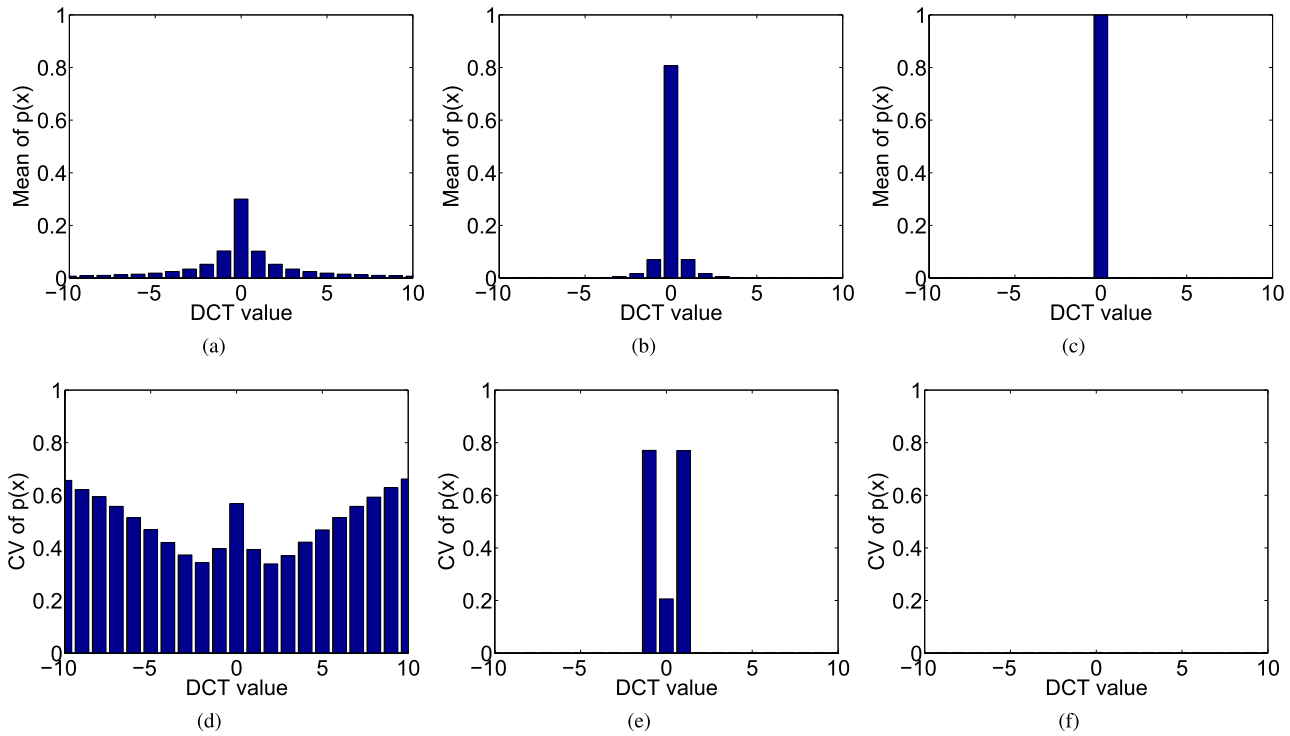


Fig. 2. (a) – (c) are the histograms $p(x)$ for the coefficients of modes 10, 33 and 66 (mode 00 is the DC mode), respectively. (d) – (f) are the corresponding CVs of (a) – (c). All experimental data is obtained on the same test database aforementioned.

The DC coefficients have been prohibited from embedding message in JPEG steganography for a long time, which is similar to what happened with zero AC coefficients. In fact, there is no explicit reason to specify why DC coefficients could no be used as cover elements. All the DC coefficients, when arranged together, can be considered as a down sampled image in spatial domain. Consequently, there is only a modest correlation between a DC coefficient and its adjacent ones, especially in textured region. In other words, the CVs of the model to characterize DC coefficients is quite large. Therefore, there also exists a possibility that messages can be securely embedded in the DC coefficients.

III. Distortion Function for Generalized Uniform Embedding Strategy

A. Uniform Embedding Revisited Distortion

Similar to the design of the original UED, the general approach to implement the generalized uniform embedding strategy is to increase the probability to be modified for those coefficients with larger CVs of the statistical model, while decrease the ones with smaller CVs. The corresponding distortion function, i.e., UERD (uniform embedding revisited distortion), should be developed to facilitate the minimal distortion embedding. To construct the UERD, we start with analyzing the CVs of first- and second-order statistics for different coefficients. Since a single coefficient itself does not exhibit any statistical property, we study the "position" of the coefficients instead, to figure out the coefficient in what position (which DCT block or which AC mode) has smaller or larger CVs.

TABLE I
Level of Texture Complexity for Different DCT Blocks

number of non-zero AC coefficients	level of texture complexity
$[0, 8)$	lowest
$[8, 18)$	lower
$[18, 30)$	higher
$[30, 63]$	highest

Firstly, we investigate the statistical properties of the different AC mode. For the sake of brevity, we only present the results of three AC mode, i.e., modes 10, 33 and 66 (mode 00 is the DC mode) for first-order statistics (Fig. 2) and modes 10, 21 and 32 for second-order statistics (Fig. 3). It is observed that, the higher frequency the AC mode is, the more zero coefficients the corresponding AC coefficients have and the smaller the corresponding CV becomes, which indicates the lower probability the coefficient to be modified, and vice versa.

Secondly, we investigate the statistical properties of the DCT blocks with different texture complexity. We artificially divide the DCT blocks of an image into four categories according to the number of their non-zero AC coefficients included, as in Table I. Put another way, the complexity of DCT block is measured by its number of non-zero AC coefficients. The first- and second-order statistics and their corresponding CVs for the four DCT block categories over 10,000 images from BOSSbase are shown in Fig. 4 and Fig. 5, respectively. It is easy to follow that, the more zero AC coefficients in a DCT block, the smaller the corresponding CVs of the first- and second-order statistics become, which means the lower the modification probability, and vice versa.

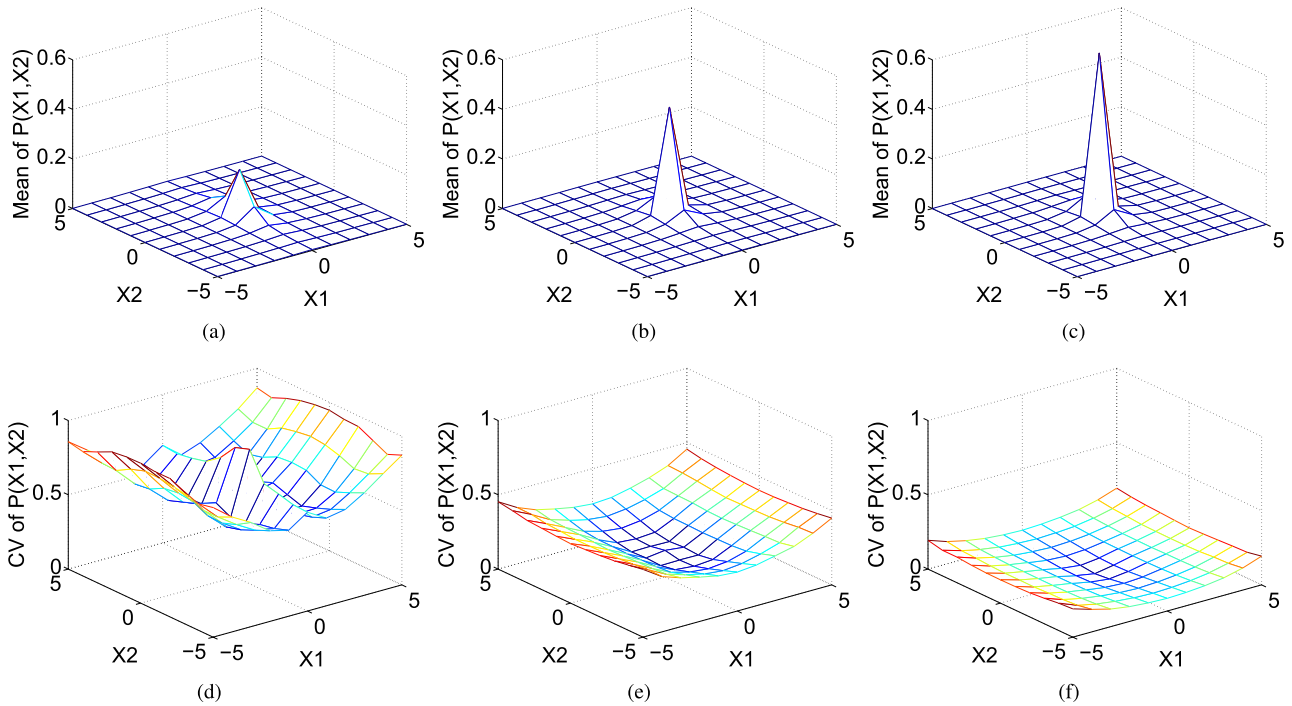


Fig. 3. (a) – (c) are the co-occurrence matrix $P(X_1, X_2)$ with offset $d = (0, 1)$ for the coefficients of modes 10, 21 and 32 (mode 00 is the DC mode), respectively. (d) – (f) are the corresponding CVs of (a) – (c). All experimental data is obtained on the same test database aforementioned.

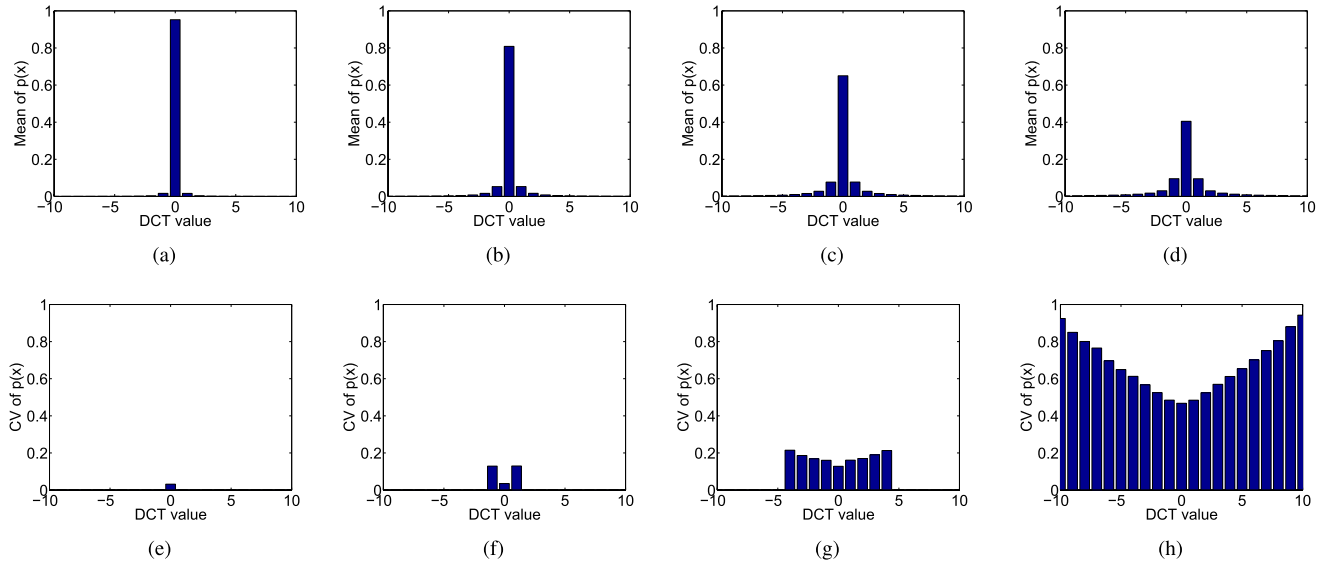


Fig. 4. (a) – (d) are the histograms for the coefficients of DCT blocks with different texture complexities, the corresponding texture complexity of (a) is the lowest while the one of (d) is the highest. (e) – (h) are the corresponding CVs of (a) – (d). All experimental data is obtained on the same test database aforementioned.

According to Fig. 2 and Fig. 4, the CV of bin 0 in lower frequency AC mode is higher than the one in higher frequency, and even higher than the one of some other bins in the same mode; while the CV of bin 0 in blocks with higher texture complexity is much higher than the one in lower texture complexity, and even higher than other non-zero bins in blocks with lower texture complexity. This also indicates that there exist zero coefficients that are far more suitable for data embedding.

Based on the above two observations, we now proceed to construct the corresponding distortion function UERD. Let x_{ij} be a coefficient in position (i, j) of a 8×8 DCT block in position (m, n), the distortion function ρ_{ij} for x_{ij} is defined as:

$$\rho_{ij} = \rho_{ij,\text{mode}} \cdot \rho_{ij,\text{block}}, \tag{2}$$

where $\rho_{ij,\text{mode}}$ and $\rho_{ij,\text{block}}$ are the distortion measures for the corresponding AC mode and DCT block, respectively. There are plenty of choices for the above two measures. In this paper, we define $\rho_{ij,\text{mode}} = q_{ij}$, where q_{ij} is the corresponding quantization step of x_{ij}. While $\rho_{ij,\text{block}}$ is defined as the function of the energies of the block where x_{ij} belongs to and

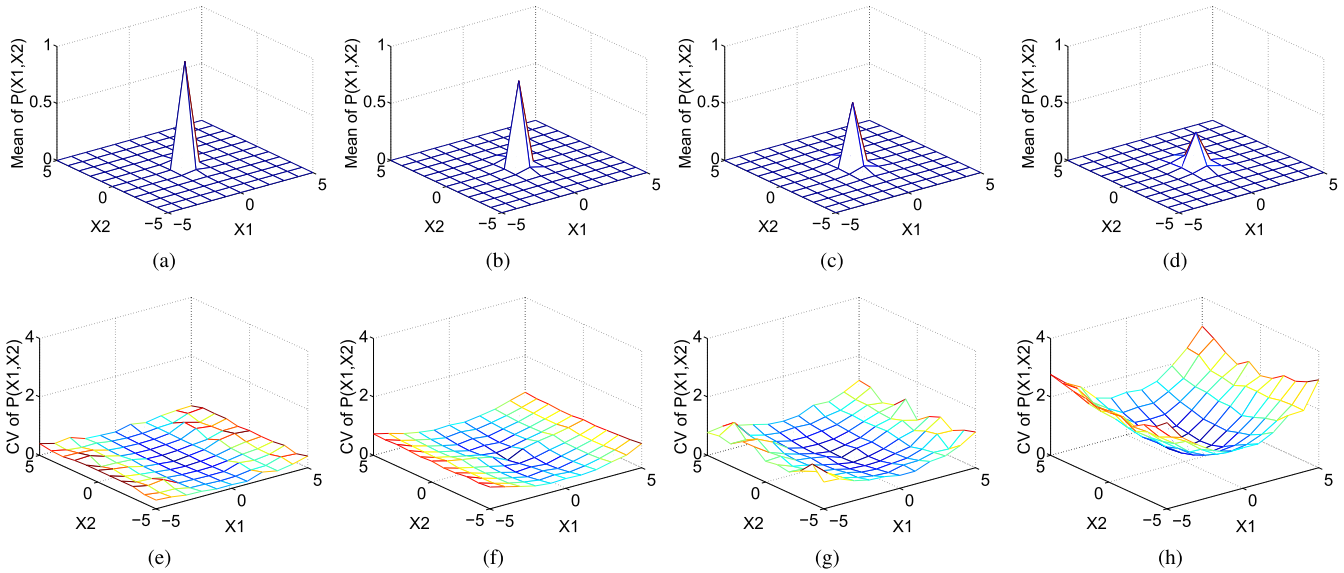


Fig. 5. (a) – (d) are the co-occurrence matrix $P(X_1, X_2)$ with offset $d = (0, 1)$ for the coefficients of DCT blocks with different texture complexities, the corresponding texture complexity of (a) is the lowest while the one of (d) is the highest. (e) – (h) are the corresponding CVs of (a) – (d). All experimental data is obtained on the same test database aforementioned.

its adjacent blocks. Let x_{ij} be in the mn^{th} block, its block energy D_{mn} is defined as

$$D_{mn} = \sum_{k=0}^{7} \sum_{l=0}^{7} |x_{kl}| \cdot q_{kl}, \tag{3}$$

where $x_{kl}, k, l \in \{0, \cdots, 7\}$ is the coefficient in the block, $x_{00} = 0$ to avoid the influence of the DC coefficient, and q_{kl} is its corresponding quantization step. We then have

$$\rho_{ij} = \begin{cases} \dfrac{0.5 * (q_{(i+1)j} + q_{i(j+1)})}{D_{mn} + 0.25 * \sum\limits_{d \in \hat{D}} d}, & \text{if } (i, j) \bmod 8 = (0, 0) \\ \dfrac{q_{ij}}{D_{mn} + 0.25 * \sum\limits_{d \in \hat{D}} d}, & \text{otherwise,} \end{cases} \tag{4}$$

where $\hat{D} = \{D_{(m-1)(n-1)}, D_{(m-1)n}, D_{(m-1)(n+1)}, D_{m(n-1)}, D_{m(n+1)}, D_{(m+1)(n-1)}, D_{(m+1)n}, D_{(m+1)(n+1)}\}$ are block energies of the eight-neighborhood of mn^{th} block. When the considered block is located in the image boundary, the nonexistent block are obtained by block padding. Considering the statistics of the DC coefficients is quite different from the AC coefficients, here we heuristically defined the distortions for the DC coefficients as the mean of their neighborhood AC coefficients in the same DCT block.

As is shown in Fig. 6(c), the distribution of the coefficients modified by our proposed UERD is quite different with the original UED (see Fig. 6(a) and (b)), i.e., the UERD modified more coefficients with small magnitude, especially the zero coefficients, while modified less coefficients with large magnitude. And Fig. 6(f) indicates that the distribution of corresponding relative changes of the global histogram is almost proportional to its CV (see Fig. 1(c)), which achieves the objective of the generalized uniform embedding strategy. The relative changes of both the first- and second-order statistics are also smaller than the ones for the original SC-UED and JC-UED, as is shown in Fig. 6(d)-(f) and Fig. 6(g) - (i), respectively.

B. Side-Informed Uniform Embedding Revisited Distortion

For the side-informed JPEG steganography, the most advantage is the use of the side-information, i.e., the rounding error during the process of JPEG compression. Such side-information can significantly improve the security performance of the JPEG steganography, which has been proved by many practical methods [2], [5], [6], [10], [13] and theoretical analysis [14]. According to our understanding, the distortion function for side-informed JPEG steganography can be factorized into two parts, as described in the form

$$\rho = \rho_{\text{fs}} \cdot \rho_{\text{si}}, \tag{5}$$

where ρ_{fs} is the distortion function designed against the feature space, while ρ_{si} denotes the distortion function that constructed with the rounding error.

In practice, ρ_{fs} can be defined just as the distortion function for the non side-informed JPEG steganography. While the ρ_{si} is always determined according to the additional rounding error, $E = |R' - R|$, where R represents the rounding error due to JPEG compression and R' is the embedding error caused by data embedding. Thus the proposed side-informed uniform embedding revisited distortion (SI-UERD) is defined as:

$$\rho_{ij} = e_{ij} \cdot \begin{cases} \dfrac{0.5 * (q_{(i+1)j} + q_{i(j+1)})}{D_{mn} + 0.25 * \sum\limits_{d \in \hat{D}} d}, & \text{if } (i, j) \bmod 8 = (0, 0) \\ \dfrac{q_{ij}}{D_{mn} + 0.25 * \sum\limits_{d \in \hat{D}} d}, & \text{otherwise,} \end{cases} \tag{6}$$

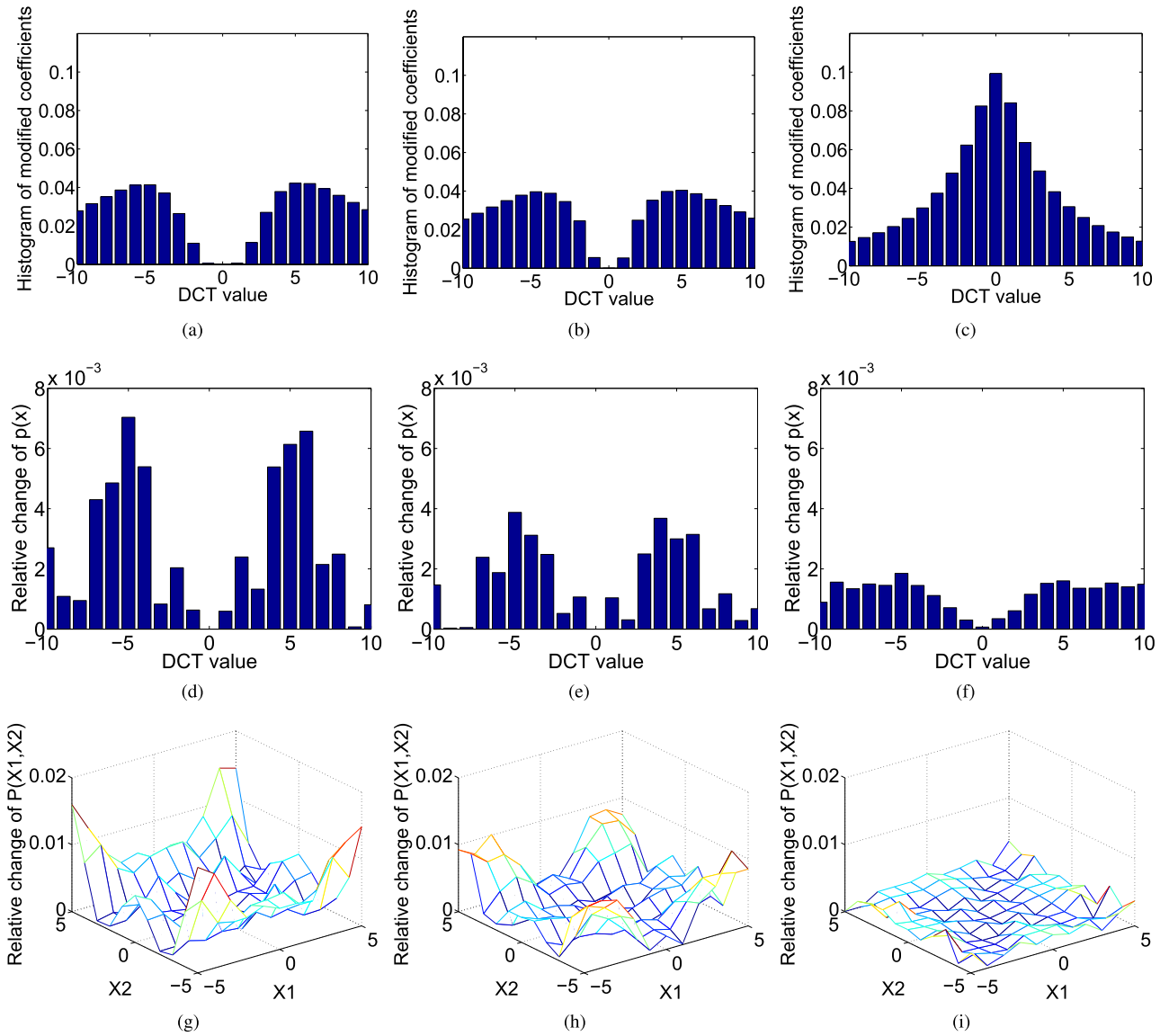


Fig. 6. (a) – (c) are the histograms for the modified coefficients with SC-UED, JC-UED and the proposed UERD at 0.2 bpnzac, respectively. (d) – (f) are the relative changes of the global histogram $p(x)$ of the same three methods. (g) – (i) are the relative changes of the co-occurrence matrix $P(X_1, X_2)$ with offset $d = (0, 1)$ of them. All experimental data are obtained on the same test database aforementioned.

where e_{ij} is the additional rounding error for x_{ij}, other parameters and the distortions of the DC coefficients are similarly defined as the ones in (3) and (4).

MME [5] is the first side-informed JPEG steganographic scheme that tried to improve its security performance by minimizing the additional rounding error. According to [5], modifying the coefficients with rounding error closer to 0.5 leads to less additional rounding errors. Except the consideration of rounding error, the distortion function of MME, however, acts as some kind of random embedding, say nsF5. Both SI-UED and SI-UERD try to combine the embedding strategies with the side-information to further increase the security performance. Fig. 7(a) – (c) show the average distribution of the rounding errors of the modified coefficients over 10,000 images for MME, SI-UED and SI-UERD, respectively. For fair comparison, all three schemes are STC coded. It is observed that SI-UERD modified more coefficients with large rounding error but less coefficients with small rounding error, leading to less additional rounding error on the whole. Obviously, for side-informed JPEG steganography, minimizing the rounding error is of particular importance. As mentioned before, the UERD modifies much more coefficients with small value (including the zero AC coefficients) than JC-UED. Considering the fact that the small AC coefficients account for the majority part of all coefficients and the rounding errors for all coefficients are uniformly distributed, SI-UED modifies less coefficients with large rounding error inevitably.

IV. Proposed JPEG Steganographic Scheme

The implementation of the proposed UERD is similar to the one of the original UED. By incorporating the UERD, the STC (with constraint height $h = 10$) based minimal distortion embedding framework is applied to embed the secret message to all DCT coefficients (DC, zero and non-zero

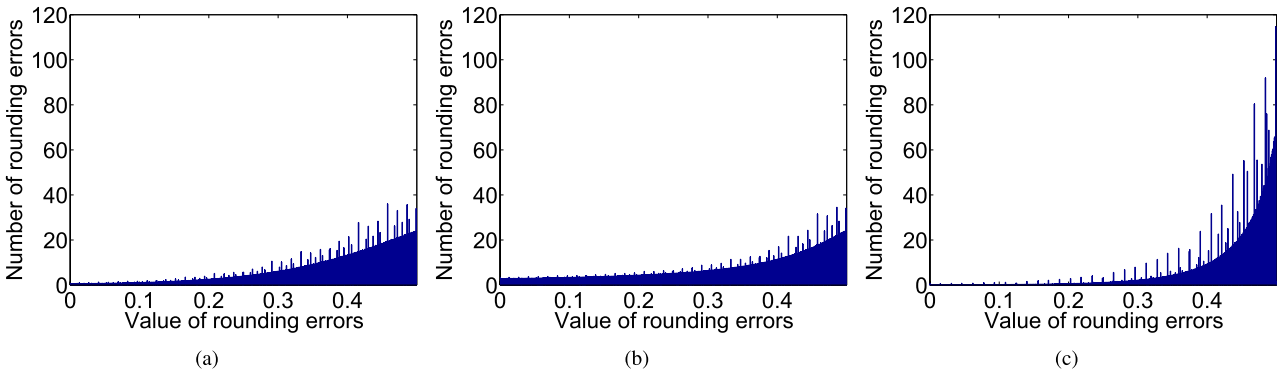


Fig. 7. (a) – (c) are average distribution of the absolute value of rounding errors of the modified coefficients over 10,000 images for MME, SI-UED and SI-UERD, respectively. All of the three schemes are STC coded with a payload of 0.3 bpnzac.

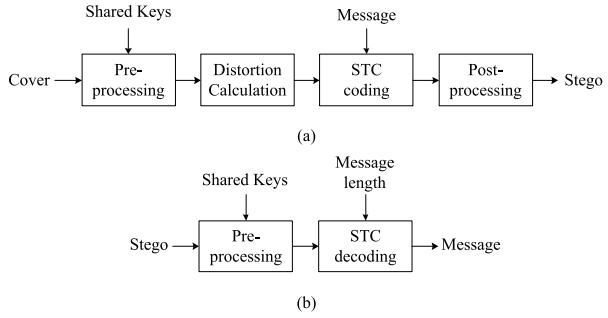


Fig. 8. Proposed scheme. (a) Data embedding. (b) Data extraction.

AC coefficients). Fig. 8 illustrates the unified framework for JPEG steganography with original BMP image or JPEG compressed image as input, which includes the process of data embedding and extraction.

In the stage of data embedding, since all DCT coefficients are utilized as the cover, there is no need to avoid modifying the zero AC coefficients and creating the new zero AC coefficients. Therefore, the embedding operations of all the coefficients are determined by ternary STC (UERD) and the binary STC (SI-UERD). As a result, the additional rounding error for the side-informed case should also be adjusted correspondingly.

For other processes such as the distortion calculation, the STC coding and decoding are just the same as in [2].

V. Experimental Results and Analysis

In this section, experimental results and analysis are presented to demonstrate the feasibility and effectiveness of the proposed UERD schemes. The comparisons with the original JC-UED and SI-UED [2], nsF5 simulator [7], Wang *et al.*'s block entropy weighted scheme (EBS) [13], Sachnev *et al.*'s BCH [6], and the state-of-the-art method — Holub *et al.*'s J-UNIWARD and SI-UNIWARD [10] for non side-informed embedding and side-informed embedding, are included. In our experiments, we concentrate on the comparison of the distortion functions for the involved schemes. Therefore, the Hamming code in MME and BCH code in BCH are replaced with STC in the interest of fair comparison.

A. Experiment Setup

1) Cover Source: All experiments are carried out on the image database BOSSbase ver. 1.01 [15]. The original database contains 10,000 images acquired by eight digital cameras in their RAW format (CR2 or DNG) and subsequently processed by converting to grayscale, and then resizing and cropping to the size of 512×512 pixels using the script available from [15]. The images in BOSSbase are then JPEG compressed using quality factors 75 and 95. Thus, we have three image databases each with 10,000 grayscale cover images of different texture characteristics in format of BMP and JPEG, which serve as the precover (BMP) and cover (JPEG) for side-informed and non side-informed JPEG embedding, respectively. In our experiments, the payloads for both non side-informed and side-informed embedding range from 0.1 to 0.5 bits per non-zero cover AC DCT coefficient (bpnzac) with a step of 0.1 bpnzac.

2) Steganalysis Features: There are several mainstream universal feature sets for JPEG steganalysis, such as the CC-PEV-548D [16], MP-486D [17] and the state-of-the-art CC-JRM-22,510D [18]. Since the CC-JRM-22,510D has been proved to be the best one by many prior works, the CC-PEV-548D and MP-486D are no longer included in our experiments. Instead, a down-scaled version of the SRM with a single quantization step $q = 1$ (SRMQ1-12,753D) are also employed to evaluate the security performances of the involved JPEG steganographic schemes. The final features will be the union of the CC-JRM-22,510D and SRMQ1-12,753D, which is also called J+SRM, giving the total feature dimensions of 35,263.

3) Machine Learning: The ensemble classifier with Fisher linear discriminant as the base learner in [19] is incorporated in our experiments as is done in [18], since it enables fast training in high-dimensional feature spaces and has a comparable performance to that of SVMs [20] working on low-dimensional feature sets. For the covers and their corresponding stego images with different steganographic schemes, embedding rates and QFs, the features for the steganalysis tool are firstly extracted. Typically, half of the cover and the stego images will be used as the training set for the ensemble classifiers, and the remaining half will be used as test set to evaluate the trained classifier. And, the security performance will be quantified as

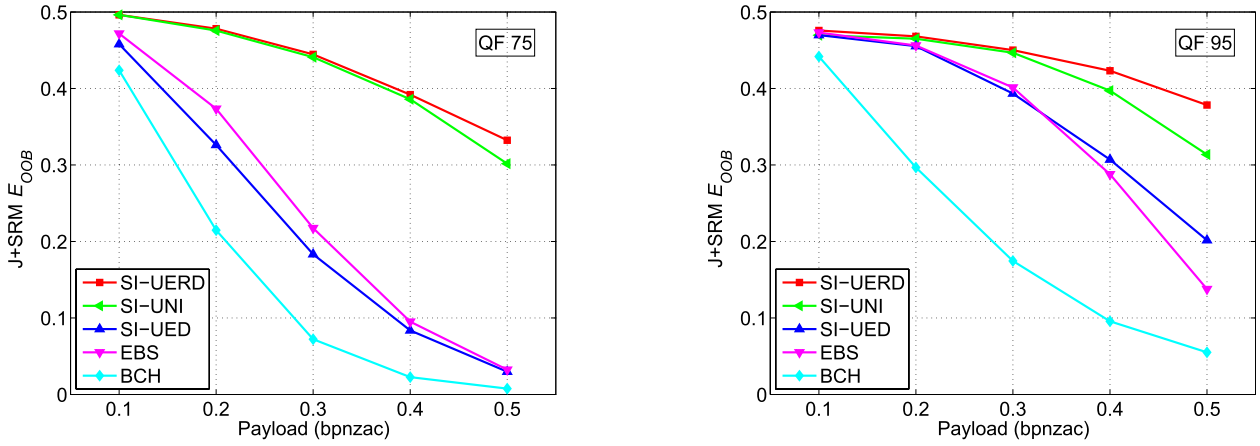


Fig. 9. (a)-(b) are detection error E_{OOB} for BCH, EBS, SI-UED, SI-UNIWARD (SI-UNI) and SI-UERD with the J+SRM-35,263D and the ensemble classifier for JPEG quality factors 75 and 95, respectively. BCH, EBS, SI-UED, SI-UNIWARD (SI-UNI) and SI-UERD are binary STC coded.

the minimal total error P_{E} under equal priors achieved on the test set, which is defined as

$$P_{\mathrm{E}} = \min_{P_{\mathrm{FA}}} \frac{P_{\mathrm{FA}} + P_{\mathrm{MD}}(P_{\mathrm{FA}})}{2}, \tag{7}$$

where P_{FA} is the false alarm rate and P_{MD} is the missed detection rate. The performance is evaluated using the median value of P_{E} over ten random tests and is denoted as $\bar{P}_E$. In this paper, we use the ensemble's 'out-of-bag' (OOB) error E_{OOB} to represent the P_{E}, since E_{OOB} is an unbiased estimate of the P_{E}. And, the security performance is displayed graphically by plotting E_{OOB} as a function of the relative payload.

B. Performance of Side-Informed UERD

For JPEG steganography with side-information, we compare the proposed SI-UERD with the original SI-UED [2], EBS [13], BCH [6], as well as the recently proposed SI-UNIWARD [10]. Fig. 9 shows the security performance of the involved schemes against the J+SRM-35,263D, for JPEG quality factors 75 and 95. For all cases, both SI-UERD and SI-UNIWARD clearly outperform the other schemes by a sizeable margin, especially for JPEG quality factors 75. SI-UERD, apparently, has the best security performance across the two JPEG quality factors. SI-UERD, however, has almost the same security performance with SI-UNIWARD before it starts to outperform SI-UNIWARD when the payload is greater than 0.3 bpnzac for all the tested JPEG images. Compared with SI-UED, the proposed SI-UERD has significant improvement in terms of security performance, indicating that the generalized uniform embedding is far more efficient than the original one. Fig. 9 also shows an interesting result that the EBS performs better than SI-UED, which is quite different with the result in [2], where the opposite results for SI-UED were observed. This could be due to the fact that the J+SRM-35,263D is better than CC-JRM-22,510D at detecting SI-UED.

C. Performance of Non Side-Informed UERD

For JPEG steganography without side-information, we compare the proposed UERD with original JC-UED [2], as well as nsF5 [7] and the state-of-the-art method — J-UNIWARD [10]. The security performances of the involved schemes against the J+SRM-35,263D, for JPEG quality factors 75 and 95, are illustrated in Fig. 10. Unlike the case of side-informed JPEG steganography, J-UNIWARD outperforms UERD this time, although the test results do not show a significant difference in performance. Note that for practical application, the performance loss of UERD in comparison with J-UNIWARD can be compensated to a considerable extent by its extremely low computational complexity, which will be discussed later in Section V-D. When compared with the original JC-UED, however, the proposed UERD shows significant improvement in terms of security as in the case of side-informed embedding.

D. Comparison of Computational Complexity

We then proceed to evaluate the computational complexity for UERD and UNIWARD. Note that both schemes utilize the same framework of minimal distortion embedding with STC as the coding method, the main difference of computational complexity is in the calculation of the distortions. To simplify the evaluation process, we assume that one operation of addition, subtraction, multiplication, division and signed magnitude arithmetic has the same computational cost.

For an $M \times N$ image, the distortion function for J-UNIWARD is defined as the sum of relative changes of the wavelet coefficients with respect to the cover image, i.e.,

$$D(\mathrm{X}, \mathrm{Y}) \triangleq \sum_{k=1}^{3} \sum_{u=1}^{n_1} \sum_{v=1}^{n_2} \frac{|W_{uv}^{(k)}(\mathrm{X}) - W_{uv}^{(k)}(\mathrm{Y})|}{\sigma + |W_{uv}^{(k)}(\mathrm{X})|}, \tag{8}$$

where X and Y are cover and stego images, one can refer to [10] for detail description and parameter setup. Changing a DCT coefficient X_{ij} will affect a block of 8×8 pixels and therefore a block of $(8 + s1) \times (8 + s1)$ (s = 16, leading to 23×23) wavelet coefficients, where $s \times s$ is the size of the 2-D wavelet support. Thus the computational complexity of (8) mainly includes the calculation of 2-D inverse DCT and 2-D DWT (Discrete Wavelet Transform). For 2-D inverse DCT transform, one pixel involves 8×8 multiplications and

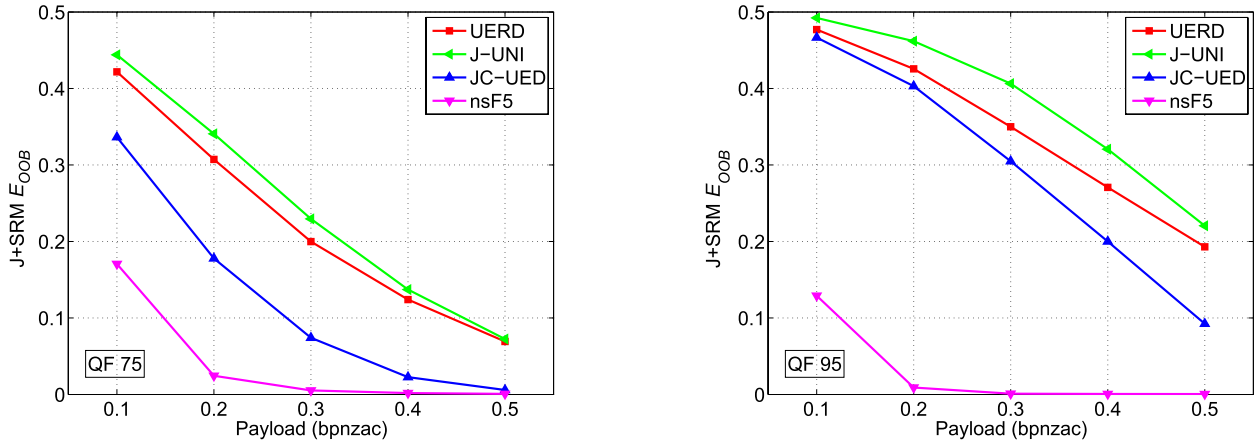


Fig. 10. (a)-(b) are detection error E_{OOB} for nsF5, JC-UED, J-UNIWARD (J-UNI) and UERD with the J+SRM-35,263D and the ensemble classifier for JPEG quality factors 75 and 95, respectively. JC-UED, J-UNIWARD (J-UNI) and UERD are ternary STC coded.

TABLE II

COMPARISON IN COMPUTATIONAL TIME OF THE UERD AND UNIWARD FOR BOTH NON SIDE-INFORMED AND SIDE-INFORMED EMBEDDING ON A SERVER WITH 2.4 GHz DUAL PROCESSORS AND 96 GB MEMORY. RESULTS OF BOTH C++ AND MATLAB IMPLEMENTATION ARE PROVIDED. THE UNIT OF TIME IS SECOND

	JPEG quality factor 75			JPEG quality factor 95		
	J-UNI	UERD	ratio	J-UNI	UERD	ratio
C++	1.7346	0.0184	94	1.7329	0.0189	92
matlab	16.6478	0.0359	464	16.7592	0.0366	458
	SI-UNI	SI-UERD	ratio	SI-UNI	SI-UERD	ratio
C++	1.7378	0.0271	64	1.7326	0.0275	63
matlab	32.5986	0.0353	923	32.5007	0.0360	903

8×8 additions, thus its computational complexity is $O(128MN)$. For the 2-D DWT, the image is first transformed to wavelet domain, and then the corresponding 23×23 wavelet coefficients can be easily obtained from the wavelet image directly. Thus the computational complexity for 2-D DWT (with three directional filters) is $O(3 \times 2 \times 16 \times 16 \times M \times N) = O(1536MN)$. Therefore, the total computational complexity for the distortion function of J-UNIWARD is roughly $O(1536MN + 128MN) = O(1664MN)$.

As for the UERD as in (4), the computation of the energy of the DCT block for the coefficients in the same DCT block is the same, which requires 8×8 signed magnitude arithmetics, 8×8 multiplications and 8×8 additions. The computation of the denominator also requires 1 multiplication and 8 additions. Thus the computational cost of the denominator for an image of size $M \times N$ is $O(3 \times 8 \times 8 \times (M/8) \times (N/8) + M \times N + 8 \times M \times N) = O(12MN)$. Taking into account the additional division operation for each pixel, the total computational complexity of the distortion function for UERD is $O(13MN)$, which is $1/128$ of the one of J-UNIWARD.

Table II reports the practical computational time of UERD and UNIWARD for both non side-informed and side-informed embedding under JPEG quality factor 75 and 95 on a server with 2.4 GHz dual processors and 96 GB memory. For all cases, the computation time for UERD is negligible. While for J-UNIWARD, when implemented with MATLAB, the computation times are more than 16 seconds for both JPEG quality factor 75 and 95, which are 464 and 458 times of the ones for the corresponding UERD. As for C++ implementation, both the computation times of UNIWARD and UERD drop significantly due to the optimized executable codes generated with C++ compiler. The computation time of UERD, however, is still roughly two orders of magnitude less than UNIWARD ($1/94$ and $1/92$ for both JPEG quality factor 75 and 95), which coincide with the theoretical evaluation given above.

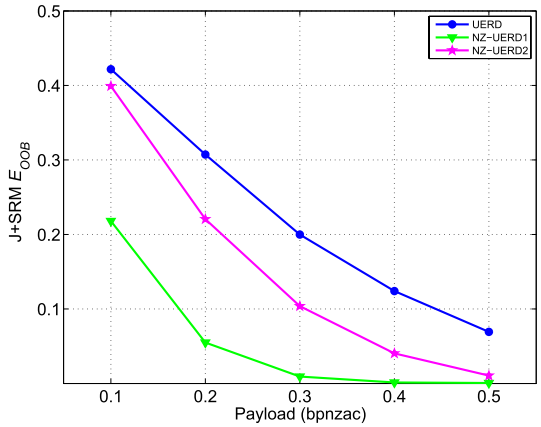


Fig. 11. Detection error E_{OOB} for the ordinary UERD and the NZ-UERD1, as well as NZ-UERD2 with the J+SRM-35,263D and the ensemble classifier for JPEG quality factors 75. The ordinary UERD utilizes all DCT coefficients, while NZ-UERD1 and NZ-UERD2 use only non-zero AC coefficients.

E. Importance of the Use of Zero AC Coefficients

For UERD, all the DCT coefficients including DC, zero and non-zero AC coefficients are treated equally using a unified distortion function. One may conjecture that the performance improvement of UERD is mainly attributable to the increase of cover element, so that more messages can be embedded.

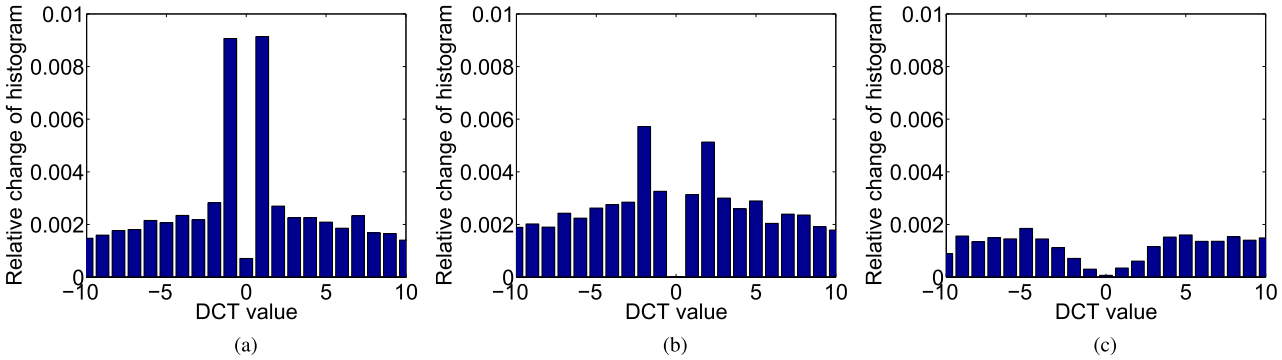


Fig. 12. (a) – (c) are the relative changes of the global histogram over 10,000 images with a payload of 0.3 bpnzac for NZ-UERD1, NZ-UERD2 and the ordinary UERD, respectively.

We then resort to additional experiment to explain the exact role that zero AC coefficients play in the proposed UERD. In the experiment, only the non-zero AC coefficients are utilized in the UERD as they do in UED. In this case, we have two choices for the embedding operation of the ± 1 AC coefficients, depending on whether new zero AC coefficients are created after embedding. For case 1 (NZ-UERD1), the ± 1 coefficients are allowed to be modified to 0, while for case 2 (NZ-UERD2), the implementation is the same as the original UED, i.e., avoiding to create new zero AC coefficients. In Fig. 11, we compare the security performance of the UERD and its two variants (NZ-UERD1 and NZ-UERD2). As seen, the NZ-UERD1 is completely broken by J+SRM-35263D, while the NZ-UERD2 is considerably outperformed by the UERD.

In Fig. 12, we show the relative changes of the global histogram for the three involved schemes over 10,000 images with a payload of 0.3 bpnzac, which clearly illustrates the reason why the performances of the two UERD variants drop so heavily. In order to implement the generalized uniform embedding, the UERD has to modify more small coefficients (Fig. 6). And the change of a bin in global histogram depends on how many coefficients are modified in such a bin and the distribution of its neighboring bins. Note that there is no coefficient modified in bin 0, the embedding operations result in sharp decrease of the ± 1 coefficients and fairly big increase of zero coefficients for NZ-UERD1 (Fig. 12(a)), or considerable increase of ± 2 coefficients for NZ-UERD2 (Fig. 12(b)). Therefore, the use of zero AC coefficients is of particular importance for the UERD. It is more than a matter that we make use of zero AC coefficients to increase the cover elements, but the one that we need to use more zero AC coefficients to better implement the proposed generalized uniform embedding strategy.

VI. Conclusion

Nowadays, perfect secure steganography has been proved to be unrealistic for empirical JPEG image. If the steganographer has to resort to imperfect steganography, currently, the most successful approach is the minimal distortion embedding framework with STC code. By properly defining the distortion function, the problem of message embedding can be formulated as the one of source coding with a fidelity constraint. Steganalysis, on the other hand, aims at detecting the presence of hidden data. By far the best image steganalyzer are constructed through feature based steganalysis and machine learning. Generally, the feature set for JPEG steganalysis is built from the first-, second- and other higher-order statistics of quantized DCT coefficients (histogram and co-occurrences).

While the previously proposed UED tries to uniformly embed the message to non-zero AC coefficients of different amplitudes, with the expectation to minimize the possible changes of statistics after embedding, this paper proposes a generalized uniform embedding strategy (UERD) by exploring the tolerable variation of image statistical model. Compared with UED, the most notable difference is that, in embedding, UERD attempts to uniformly "spread" the relative changes of statistics in a way that they are proportional to their CVs (coefficient of variation).

As for the implementation, the UERD uses all DCT coefficients (DC, zero and non-zero AC) as cover elements, and both zero and non-zero AC coefficients are treated equally using a unified distortion function. When the side-information of the original uncompressed image is available, the corresponding side-informed UERD (SI-UERD) is constructed by simply combining the additional rounding error with the non side-informed UERD. Extensive experiments have been carried out to demonstrate the superior performance of the proposed UERD in terms of secure embedding payload against steganalysis and show that the UERD has a comparable security performance with the state-of-the-art UNIWARD with markedly reduced computational complexity.

Acknowledgment

The authors would like to thank the anonymous reviewers and associate editor for their comments that greatly improved the manuscript.

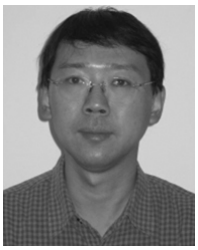
References

[1] L. Guo, J. Ni, and Y. Q. Shi, "An efficient JPEG steganographic scheme using uniform embedding," in *Proc. 4th IEEE Int. Workshop Inf. Forensics Secur.*, Tenerife, Spain, Dec. 2012, pp. 169–174.

[2] L. Guo, J. Ni, and Y. Q. Shi, "Uniform embedding for efficient JPEG steganography," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 5, pp. 814–825, May 2014.

[3] R. Böhme, "Improved statistical steganalysis using models of heterogeneous cover signals," Ph.D. dissertation, Faculty Comput. Sci., Tech. Univ. Dresden, Dresden, Germany, 2008.
[4] A. Westfeld, "F5—A steganographic algorithm," in *Proc. 4th Int. Workshop Inf. Hiding*, LNCS, vol. 2137, Springer-Verlag, 2001, pp. 289–302. [Online]. Available: http://wwwrn.inf.tu-dresden.de/~westfeld/f5.html
[5] Y. Kim, Z. Duric, and D. Richards, "Modified matrix encoding technique for minimal distortion steganography," in *Proc. 8th Int. Workshop Inf. Hiding*, LNCS, vol. 4437, Springer-Verlag, Jul. 2006, pp. 314–327.
[6] V. Sachnev, H. J. Kim, and R. Zhang, "Less detectable JPEG steganography method based on heuristic optimization and BCH syndrome coding," in *Proc. 11th ACM Workshop Multimedia Secur.*, Sep. 2009, pp. 131–140.
[7] J. Fridrich, T. Pevný, and J. Kodovský, "Statistically undetectable jpeg steganography: dead ends challenges, and opportunities," in *Proc. 9th ACM Workshop Multimedia Secur.*, Dallas, TX, USA, Sep. 2007, pp. 3–14. [Online]. Available: http://dde.binghamton.edu/download/nsf5simulator/
[8] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 920–935, Sep. 2011.
[9] T. Filler and J. Fridrich, "Design of adaptive steganographic schemes for digital images," *Proc. SPIE, Media Watermarking, Secur., Forensics III*, vol. 7880, p. 78800F, Jan. 2011.
[10] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP J. Inf. Secur.*, 2014(1), pp. 1–13, 2014.
[11] J. Kodovsky, J. Fridrich, and V. Holub, "On dangers of overtraining steganography to incomplete cover model," in *Proc. 13th ACM Multimedia Workshop Secur.*, New York, NY, USA, Sep. 2011, pp. 69–76.
[12] D. A. Freedman, *Statistical Models: Theory and Practice*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[13] C. Wang and J. Ni, "An efficient JPEG steganographic scheme based on the block entropy of DCT coefficients," in *Proc. IEEE ICASSP*, Kyoto, Japan, Mar. 2012, pp. 1785–1788.
[14] J. Fridrich, "On the role of side information in steganography in empirical covers," *Proc. SPIE, Media Watermarking, Secur., Forensics XV*, vol. 8665, pp. 86650I-1–86650I-11, Mar. 2013.
[15] P. Bas, T. Filler, and T. Pevný, "Break our steganographic system: The ins and outs of organizing BOSS," in *Proc. 13th Inf. Conf.*, 2011, pp. 59–70. [Online]. Available: http://www.agents.cz/boss/
[16] J. Kodovský and J. Fridrich, "Calibration revisited," in *Proc. 11th ACM workshop Multimedia Secur.*, New York, NY, USA, Sep. 2009, pp. 63–74.
[17] C. Chen and Y. Q. Shi, "JPEG image steganalysis utilizing both intrablock and interblock correlations," in *Proc. IEEE Int. Symp. Circuits Syst.*, Mar. 2008, pp. 3029–3032.
[18] J. Kodovský and J. Fridrich, "Steganalysis of JPEG images using rich models," *Proc. SPIE*, vol. 8303, p. 83030A, Feb. 2012.
[19] J. Kodovský, J. Fridrich, and V. Holub, "Ensemble classifiers for steganalysis of digital media," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 432–444, Apr. 2012.
[20] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, Apr. 2011, Art. ID 27. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm

Linjie Guo (S'12) received the Ph.D. degree from the School of Information Science and Technology, Sun Yat-sen University, Guangzhou, China, in 2015. Since 2015, he has been with Huawei Technologies Company, Ltd., where he is currently a Research Engineer. His current research interests include computer vision, machine learning, and steganography.



Jiangqun Ni (M'12) received the Ph.D. degree in electronics engineering from the University of Hong Kong, Hong Kong, in 1998.

He was a Postdoctoral Fellow for a joint program between the Sun Yat-sen University, Guangzhou, China, and the Guangdong Institute of Telecommunication Research from 1998 to 2000. Since 2001, he has been with the School of Information Science and Technology, Sun Yat-sen University, where he is currently a Professor. He has authored over 50 papers in his research areas. His research interests include data hiding and digital image forensics, image-based modeling and rendering, and image/video processing.

Wenkang Su received the B.S. degree in electronics engineering from Jinan University, Guangzhou, China, in 2014. He is currently pursuing the M.S. degree with the School of Information Science and Technology, Sun Yat-sen University, Guangzhou, China. His current research interests include steganography, steganalysis, and multimedia security.

Chengpei Tang received the Ph.D. degree in radio physics from Sun Yat-sen University, Guangzhou, China, in 2007. Since 2007, he has been with the School of Engineering, Sun Yat-sen University, where he is currently an Assistant Professor. His current research interests include multimedia security, indoor positioning algorithm, and wireless sensor networks.

Yun-Qing Shi (M'88–SM'92–F'05) received the M.S. degree from Shanghai Jiao Tong University, China, and the Ph.D. degree from the University of Pittsburgh, USA. He has been with the New Jersey Institute of Technology, Newark, NJ, USA, since 1987. His research interests include digital data hiding, forensics and information assurance, and visual signal processing and communications.

He has authored or coauthored 300 papers, one book, five book chapters, and edited ten books. He holds 28 U.S. patents. He received the Innovators Award 2010 from the New Jersey Inventors Hall of Fame for Innovations in Digital Forensics and Security. His U.S. patent 7 457 341 entitled "System and Method for Robust Reversible Data Hiding and Data Recovery in the Spatial Domain" received the 2010 Thomas Alva Edison Patent Award from the Research and Development Council of New Jersey.

He served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS (II), and a few other journals. He also served as the Technical Program Chair of the IEEE ICME07, the Co-Technical Chair of IWDW06, 07, 09, 10, 11, 12, and 13, and the IEEE MMSP05, the Co-General Chair of the IEEE MMSP02, and a Distinguished Lecturer of the IEEE CASS. He is a member of a few IEEE technical committees.