Defining Cost Functions for Adaptive JPEG Steganography at the Microscale

Kejiang Chen, Hang Zhou, Wenbo Zhou, Weiming Zhang, and Nenghai Yu

***Abstract*— Minimal distortion steganography is the most successful model for adaptive steganography, in which the cost function determines the security. Texture complexity is the major factor in defining cost function in images. In this paper, we proposed a method to improve the cost function of JPEG steganography by exploiting the texture in microscale. The proposed scheme is designed by using a "microscope" to highlight details in an image, so that distortion definition can be more refined. Linear unsharp masking acts as the microscope, because it can accentuate the texture region as well as maintain the original characteristics of images. Inter-block spreading rule is proposed to further strengthen the security. We improve the state-of-the-art schemes, J-UNIWARD and UERD, as J-UNIWARD has outstanding performance on resisting detection while UERD has significant lower computational complexity. In order to keep high efficiency of UERD, filtering in the DCT domain is introduced. Extending experiments show that in most cases the proposed methods (J-MSUNIWARD and MSUERD) can achieve a higher level of security than the original methods.**

***Index Terms*— Steganography, distortion, JPEG, microscale.**

I. Introduction

STEGANOGRAPHY is the art of hiding messages in objects without drawing suspicion from steganalysis [1], [2]. Currently, the vast majority of work on steganography has focused on digital images. With the purpose of minimizing statistical detectability, modern steganography can be formulated as a source coding problem that minimizes embedding distortion [3]. The distortion is obtained by assigning a cost to each cover element, and the messages are embedded while minimizing the total function which is the sum of all elements' costs. Syndrome-trellis codes (STCs) provide a general methodology for embedding while minimizing an arbitrary additive distortion function with a performance near the theoretical bound [4].

As for content-adaptive steganography, how to define the cost function becomes one of the most important research issues. In the spatial domain, taking into account of adversary's attack method, the cost function of HUGO [5] is defined as the weighted sum of the difference between feature vectors extracted from a cover image and its stego version in SPAM [6] feature space. Holub and Fridrich [7] proposed the algorithm WOW (Wavelet Obtained Weights) which assigns high costs to pixels in regions that are easily predictable by a bank of directional filters. UNIWARD (UNIversal WAvelet Relative Distortion) [8] has a slightly modified filter bank from WOW to improve the versatility of the algorithm. Thus, it can be realized in arbitrary domain, including spatial domain, JPEG domain, and so on. HILL [9] improves the cost function by cooperating with spreading rule, which makes more embedding changes concentrated in texture areas. The above methods design cost function in an ad hoc or empirical manner, Sedighi *et al.* [10] proposed MiPOD under a model-driven framework with a comparable security to HILL.

Same as the spatial domain steganography, in the early period, a lot of non-adaptive schemes are developed for JPEG steganography, such as Jsteg [11], F5 [12], nsF5 [13], MME [14]. With the development of STC in steganographic code, the emerging JPEG steganographic schemes all focused on the design of the distortion function over the past few years. Filler and Fridrich *et al.* proposed MOD (Model Optimized Distortion) [15], whose distortion was heuristically defined as a rich parametric model, and then was optimized to obtain the least detectability with respect to a selected feature set (cover model). On the basis of UNIWARD, Holub *et al.* [8] developed it to the JPEG domain (J-UNIWARD) and the side-informed domain (SI-UNIWARD). Unlike the conventional JPEG steganographic schemes which only embed the secret message into non-zero AC coefficients, J-UNIWARD and SI-UNIWARD use all DCT coefficients including DCs, zero and non-zero ACs as possible cover elements, and achieve so far the best security performance. However, the computational complexity of obtaining distortion from the wavelet domain may be a major problem in implementation, especially when applied in the mobile terminal.

With regard to efficient JPEG steganography, a lightweight distortion metric known as uniform embedding distortion metric (UED) [16], which takes into account the magnitude of DCT coefficient as well as both its intra- and inter-block neighbourhood coefficients, is constructed to incorporate the uniform embedding. Guo *et al.* [17] improved UED, which is called UERD, by exploring the tolerable variation of image statistical model. It shows that the UERD has a close security performance to the state-of-the-art

Manuscript received November 9, 2017; revised April 4, 2018 and September 2, 2018; accepted September 4, 2018. Date of publication September 10, 2018; date of current version November 8, 2018. This work was supported by the National Natural Science Foundation of China under Grant U1636201, Grant 61572452, and Grant 61802357. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Pedro Comesana. *(Corresponding author: Weiming Zhang.)*

The authors are with the CAS Key Laboratory of Electro-Magnetic Space Information, University of Science and Technology of China, Hefei 230026, China (e-mail: zhangwm@ustc.edu.cn).

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.2018.2869353

J-UNIWARD with markedly reduced computational complexity. An improved version of UERD named IUERD [18] was proposed, by exploring the correlation among neighbouring DCT blocks more efficiently.

All the above methods will try to cluster the modifications into texture regions. To more precisely explore texture and describe costs, we proposed the microscale steganography scheme in the spatial domain in our previous work [19], utilizing image enhancement technique to highlight the texture regions before defining cost function, and then spreading rule is cooperated to further strengthen the security. Referring to JPEG steganography, the embedding distortion is highly related to the texture region as well. As for J-UNIWARD, the distortion relies on the wavelet filter residuals and the large residual leads to small cost. For UERD, the DCT energy represents the complexity of image to some extent. These indicate that JPEG steganography satisfies the essential requirement of microscale steganography. Therefore, we extend the framework from the spatial domain to the JPEG domain.

Spreading rule has been successfully utilized in spatial domain, which indicates that the costs of modifying neighbouring elements should be similar [20]. The underlying premise of spreading rule is that neighbouring elements own strong correlation. However, since the modification impact among neighbouring coefficients varies a lot, it cannot be directly utilized on the DCT plane. Inspired by the formation of JPEG images, once we collect the coefficients in the same frequency (DCT mode), the neighbour coefficients in the collected plane will own high correlation, and thus we can spread the distortion. Motivated by mentioned factor, inter-block spreading rule is proposed to enhance the security of JPEG steganography including the proposed microscale steganography.

Since the costs of J-UNIWARD are calculated in the spatial domain, the microscale steganography scheme can be directly applied. We highlight the decompressed image with the unsharp masking filter in the spatial domain which is similar to the prior work [19]. The ultimate wavelet filter residual would be calculated on both the original image and the enhanced image, which ensures the texture area would be assigned large residual and small distortion. The improved schemes are named J-MSUNIWARD and SI-MSUNIWARD corresponding to J-UNIWARD and SI-UNIWARD, respectively.

With the purpose of maintaining the low computational complexity of UERD, filtering in the DCT domain is introduced for microscale steganography of UERD. Additionally, spatial domain filtering is utilized for comparison. Analogously, The improved schemes are named MSUERD_DCT, MSUERD_SPA for JPEG steganography, and SI-MSUERD_DCT, SI-MSUERD_SPA for side-informed steganography, where the suffixes 'DCT' and 'SPA' mean which type of filter is used.

The security performance of proposed schemes are verified with exhaustive experiments using the state-of-the-art steganalyzers with DCTR [21], GFR [22] and J+SRM [23] on the BOSSbase database [24] and BOWS2 [25]. Experimental results show that in most cases the proposed methods can achieve higher level of security than the original methods.

The contributions of this work are summarized as follows.

1) We propose the scheme of microscale steganography in JPEG domain, which achieves higher security performance than seed methods.
2) Filtering in the DCT domain is introduced for improving the efficiency of MSUERD, which is valuable in practical.
3) Based on the property of JPEG image, inter-block spreading rule is proposed for JPEG steganography, which do enhance the security of microscale steganography for UERD.

The rest of this paper is organized as follows. After introducing notations, we review microscale steganography in the spatial domain. In Section III, JPEG steganography is subsequently reviewed. In Section IV and Section V, we propose microscale steganography for JPEG steganography and side-informed steganography, respectively. Results of experiments are elaborated in Section VI to demonstrate the effectiveness of the proposed schemes. Conclusion and future work are given in Section VII.

II. Preliminaries and Prior Work

A. Notations

Throughout the paper, matrices, vectors and sets are written in bold face. The cover image (of size $n_1 \times n_2$) is denoted by $\mathbf{X} = (x_{ij})^{n_1 \times n_2}$, where the signal x_{ij} is an integer, such as 8-bit pixels values, $x \in \{0, \ldots, 255\}$ or quantized JPEG DCT coefficients, $x \in \{-1024, \ldots, 1023\}$. $\mathbf{Y} = (y_{ij})^{n_1 \times n_2}$ denotes the stego image. For simplicity and without loss of generality, we will assume that n_1 and n_2 are multiples of 8.

For the sake of legibility, we try to keep the notations consistent to the former works. A precover image will be denoted as $\mathbf{P} = (P_{ij})$. When compressing $\mathbf{P}$, the DCT transform is executed for each 8×8 block from a fixed grid. Then the DCT coefficients are divided by quantization steps and rounded to integers. We use the symbols $\mathbf{D}$ and $\mathbf{X}$ to denote the matrices of all raw and quantized DCT coefficients. The symbol $J^{-1}(\mathbf{X})$ for the JPEG image represents the spatial image decompressed from DCT coefficient $\mathbf{X}$ [8].

The embedding operation on x_{ij} is formulated by the range I. An embedding operation is called binary if $|I| = 2$ and ternary if $|I| = 3$. For instance, the ± 1 embedding operation is ternary embedding with $I_{i,j} = \{x_{ij} - 1, x_{ij}, x_{ij} + 1\}$, where "0" denotes no modification.

B. Microscale Steganography in Spatial Domain

Generally speaking, content-adaptive steganography assigns low costs in texture regions, while high costs in smooth areas. From this point of view, grasping the distribution of the texture areas in an image counts for a lot. By comparing the cover image and the corresponding distortion, we are able to find some pixels with high cost values inside texture areas. However, these areas are probably suitable for concealing data, and should be assigned with low costs. These phenomena imply that the current steganographic distortions may not seize the detail of image precisely. Fortunately, image enhancement plays a role in exposing the detail of the image. In order

to highlight fine details as well as maintain the original characteristics of the image, unsharp masking (UM) is to the choice. We used 'microscope' as a metaphor for the unsharp masking.

With the help of a 'microscope', we can get the enhanced image and then utilize existing steganographic methods (WOW, UNIWARD, HILL, etc.) to define distortion on the enhanced image. The ultimate distortion will be obtained by cooperating with spreading rule and then assigned to the cover image. Finally, the information hiding would be well implemented by STCs. The experimental results showed that the scheme of the microscale steganography did improve the security of the current steganography methods in the spatial domain [19].

Since JPEG serves as one of the most popular adopted formats for image storage and transmission, JPEG steganography has become an important branch of information hiding. Naturally, we would like to extend microscale steganography to the JPEG domain, which will be expounded in subsequent sections.

III. Review of JPEG Steganography

Currently, J-UNIWARD and UERD become the mainstream embedding methods in JPEG images, for which the former achieves the state-of-the-art security performance and the latter owns the considerable security and much lower computational complexity. Since the process of JPEG compression is block-wised, the distortion definitions of the mentioned two schemes are block-wised as well, and can be formulated as one framework:

$$\rho = \frac{\textit{Inner block distinguishing factor (IF)}}{\textit{Block texture descriptor (TD)}}, \tag{1}$$

where the inner block distinguishing factor (*IF*) is to give different weights according to positions of DCT coefficients in an 8×8 block. Different positions in DCT block represent different frequencies, which impact the detector's performance in varying degrees. Actually, the inner block distinguish factor (*IF*) is independent with the image content. The block texture descriptor stands for the texture of the areas around the coefficient, closely linked to the content.

A. J-UNIWARD

J-UNIWARD's distortion is formed in the wavelet domain. As mentioned in [8], they utilized a set of linear shift-invariant filters represented with their kernels $\mathcal{K} = \mathbf{k}^{(1)}, \mathbf{k}^{(2)}, \mathbf{k}^{(3)}$. The kernels are used to compute directional residuals $\mathbf{W}^{(i)}(\mathbf{I}) = \mathbf{K}^{(i)} \star \mathbf{I}$, where $'\star'$ is a mirror-padded convolution, representing the smoothness of a given spatial image $\mathbf{I}$. We will denote with $\mathbf{W}^{(i)}_{pq}$, $p = 1, 2, \ldots, l_1, q = 1, 2, \ldots l_2$, their corresponding pqth wavelet coefficient in the ith subband of the first decomposition level, where l_1, l_2 are the width and height of the wavelet coefficient matrix. J-UNIWARD utilizes the Daubechies 8-tap wavelet directional filter bank built from one dimensional low-pass and high-pass filters, $\mathbf{h}$ and $\mathbf{g}$:

$$\mathbf{K}^1 = \mathbf{h} \cdot \mathbf{g}^T, \quad \mathbf{K}^2 = \mathbf{g} \cdot \mathbf{h}^T, \ \mathbf{K}^3 = \mathbf{g} \cdot \mathbf{g}^T. \tag{2}$$

We denote with $\mathbf{B}^{(k,l)}$, a derived matrix from 8×8 zero matrix by modifying the klth element to 1. Given the directional filters, the *IF* in J-UNIWARD is represented by 64 matrices, and can be formulated as

$$IF_{\text{UNI}} = \left| \mathbf{W}^{(i)}\left(J^{-1}(\mathbf{B}^{(k,l)})\right) \right|, \quad i = 1, 2, 3. \tag{3}$$

Every matrix represents the modification impact of the corresponding position (k, l) in an 8×8 DCT block, which does not have any relation to image content and can be precaculated. Since Daubechies 8-tap wavelet directional filter bank (16×16) is adopted, the size of $\mathbf{W}^{(i)}\left(J^{-1}(\mathbf{B}^{(k,l)})\right)$ will be 23×23, namely, $l_1 = 23, l_2 = 23$. Given a cover JPEG image $\mathbf{X}$, we can obtain its corresponding wavelet residual coefficient matrix $\mathbf{W}^{(i)}\left(J^{-1}(\mathbf{X})\right)$. The larger the absolute value of residual is, the more texture the image is. Therefore, let x_{kl} be a coefficient in position (k, l) of the mnth block of the image $\mathbf{X}$, and the *TD* of mnth block is defined as the absolute value of residual block $\mathbf{W}^{(i)}\left(J^{-1}(\mathbf{X}_{mn})\right)$ by collecting 23×23 wavelet coefficients around the corresponding position of mnth block in $\mathbf{W}^{(i)}\left(J^{-1}(\mathbf{X})\right)$. According to Eq. (1), the distortion of x_{kl} in the mnth block can be defined as:

$$\rho^{(k,l)}_{mn} = \sum_{i=1}^{3} \sum_{p=1}^{23} \sum_{q=1}^{23} \frac{\left| \mathbf{W}^{(i)}_{pq}\left(J^{-1}(\mathbf{B}^{(k,l)})\right) \right|}{\left| \mathbf{W}^{(i)}_{pq}\left(J^{-1}(\mathbf{X}_{mn})\right) \right| + \sigma}, \tag{4}$$

where $\sigma = 2^{-6}$ is a constant stabilizing the numerical calculations.

B. UERD

Though J-UNIWARD achieves state-of-the-art performance, it has high computational complexity. A lightweight distortion (UERD) for JPEG steganography was proposed [17]. When it comes to the *IF* in UERD, quantization table with an adjustment on the DC quantization step is adopted to distinguish the impact of different positions in a block. The weighted DCT energy is defined as *TD*, reflecting the texture of the block and its neighbours. And the DCT energy is formulated as follows:

$$D_{mn} = \sum_{k=0}^{7} \sum_{l=0}^{7} |x_{kl}| \cdot q_{kl}, \quad k, l \in \{0, \cdots, 7\}, \tag{5}$$

where x_{kl} is DCT coefficients in the mnth block, $x_{00} = 0$ to avoid the influence of DC coefficient, and q_{kl} is the quantization step. Then the distortion of UERD is defined as:

$$\rho^{(k,l)}_{mn} = \begin{cases} \dfrac{0.5 * (q_{(k+1)l} + q_{k(l+1)})}{D_{mn} + 0.25 * \sum_{d \in \hat{D}} d} & \text{if } (k, l) \bmod 8 = (0, 0) \\ \dfrac{q_{ij}}{D_{mn} + 0.25 * \sum_{d \in \hat{D}} d} & \text{otherwise,} \end{cases} \tag{6}$$

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 the block energies of the neighbourhood of the mnth block. When it comes to boundary blocks, the nonexistent blocks are obtained by block symmetric padding [17]. The distortions for the DC coefficients are defined as the mean of their neighbourhood AC coefficients in the same DCT block.

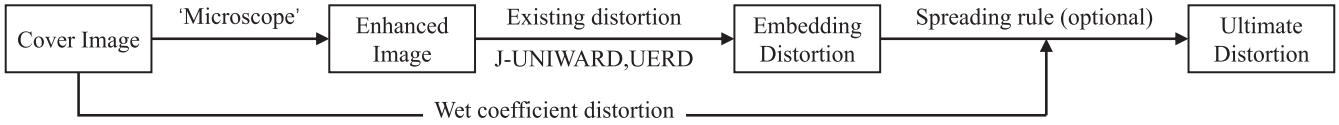


Fig. 1. The diagram of the microscale steganography in the JPEG domain.

C. Side-Informed JPEG Steganography

Given the raw DCT coefficient C_{ij} obtained from the precover $\mathbf{P}$, the steganographer has the choice of rounding C_{ij} up or down to modulate its parity. The rounding error is denoted with e_{ij}:

$$e_{ij} = \left|C_{ij} - X_{ij}\right|, \quad e_{ij} \in [0, 0.5]. \tag{7}$$

In SI-UNIWARD, a binary embedding scheme modulates the cost of changing $C_{ij} = [X_{ij}]$ to $[X_{ij}] + \text{sign}(e_{ij})$ by $1 - 2\left|e_{ij}\right|$, while prohibiting the change to $[X_{ij}] - \text{sign}(e_{ij})$ [26]:

$$\rho_{ij}^{(\text{SI})}\big(\text{sign}(e_{ij})\big) = (1 - 2\left|e_{ij}\right|)\rho_{ij} \tag{8}$$

$$\rho_{ij}^{(\text{SI})}\big(-\text{sign}(e_{ij})\big) = \Omega, \tag{9}$$

where $\rho_{ij}^{(\text{SI})}(u)$ is the cost of modifying the cover value by $u \in \{-1, 1\}$, ρ_{ij} are the costs of J-UNIWARD, and Ω is a large constant.

As for SI-UERD, the rounding error acts as an multiplicative factor in the distortion:

$$\rho_{ij}^{(\text{SI})} = e_{ij} \cdot \rho_{ij}. \tag{10}$$

IV. Microscale Steganography for JPEG Steganography

A. Motivation

After reviewing JPEG steganography, it is easy to see the *IF* in J-UNIWARD or UERD is fixed pattern, but the *TD* is closely linked to the image content reflecting the texture of the block. According to the review of J-UNIWARD and UERD, the *TD* of J-UNIWARD is defined more precisely than that of UERD. In J-UNIWARD, the *TD* is related to a 23×23 neighbour residuals of three filter banks, while UERD considers the *TD* merely counting on non-zero coefficients in one single block. As a result, the security of J-UNIWARD performs better than UERD. In Fig. 2, we contrast the modifications in the spatial domain caused by DCT embedding for UERD and for J-UNIWARD. The changes introduced by J-UNIWARD are distributed in texture areas, while there are many changes introduced by UERD in smooth pillars. The phenomenon indicates that if we describe the texture of image meticulously, the security performance will be strengthened. The requirement meets the aim of the microscale steganography, so we extend our previous work to the JPEG domain.

The diagram of Microscale Steganography (MS) in the JPEG domain is presented in Fig. 1. First, the cover image would be enhanced by a microscope into the enhanced image (filtering in spatial domain or DCT domain). As shown in Fig. 4, the enhanced image appears to contain more details.

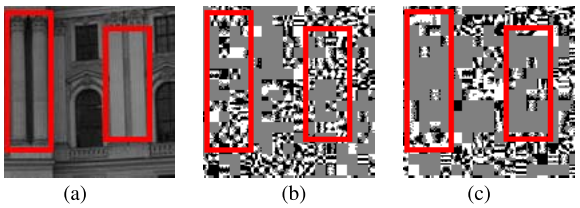


Fig. 2. The changes (b),(c) in the spatial domain caused by DCT embedding with respect to the cover image (a) with payload 0.5 bpnzac, QF=75, using UERD and J-UNIWARD, respectively. White pixels represent positive changes; dark pixels represent negative changes; gray pixels mean no changes. Regularly, fewer changes in smooth area mean better security. It can be seen that the changes in the red rectangle (smooth area) caused by J-UNIWARD are fewer than that caused by UERD. (a) Cover image (b) UERD (c) J-UNIWARD.

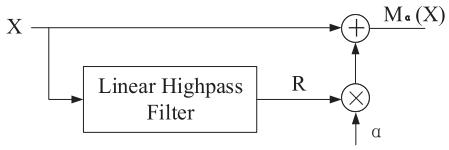


Fig. 3. Linear unsharp masking for image enhancement.

Fig. 4. The enhanced image (b) is sharpened by UM algorithms in the DCT domain. The detail in the enhanced image is clearer and sharper than that in the original image (a). Specifically, The edges are highlighted and the cloud owns more sense of hierarchy. (a) Original image. (b) Enhanced image.

Then utilize existing distortion methods to define the embedding distortion. The embedding distortion will be smoothed according to the inter-block spreading rule optionally. Finally, the distortion should be adjusted when it comes to the saturated coefficients, i.e. $X_{ij} = -1024$ or 1023.

Linear unsharp masking is adopted as the microscope, following the previous work in the spatial domain [19]. In the linear UM algorithm [27], as shown in Fig. 3, the enhanced image $M\alpha(X)$ is obtained from the input image $\mathbf{X}$ as

$$M_\alpha(\mathbf{X}) = \mathbf{X} + \alpha * \mathbf{R}, \tag{11}$$

where $\mathbf{R}$ is the correction signal as the output of a high-pass filter and α is the positive scaling factor that controls the level of contrast enhancement achieved at the output.

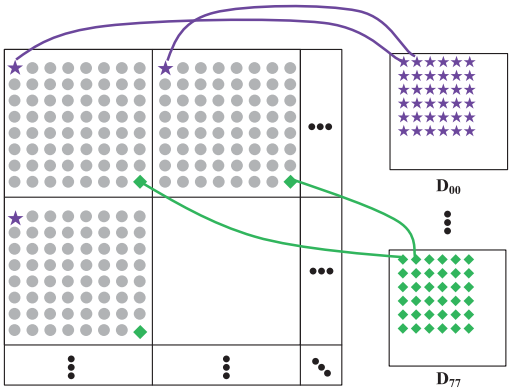


Fig. 5. The procedure of formulation of distortion subplane according to DCT mode. The distortion of coefficients with the same DCT mode will be divided into a subplane.

As mentioned before, the *IF* is fixed, therefore, the focus in the implementation of the microscale steganography in the JPEG domain is improving the description of *TD*. The more detail of the microscale steganography schemes will be presented in the following subsections.

B. Inter-Block Spreading Rule in JPEG Domain

Spreading rule is one part of the scheme of microscale steganography in the spatial domain [19]. However, spreading rule cannot be directly utilized on the DCT plane. As concluded in [17], the higher frequency of the AC mode is, the higher modification distortion the coefficient is, and vice versa. If we utilize the spreading rule directly on the DCT plane, the cost will be spread among different frequency components, which weakens the security performance of steganography.

Inspired by the formation of JPEG images, once we collect the coefficients in the same frequency (DCT mode), the collected plane can be seen as a spatial plane, where the neighbour coefficients own high correlation, so that we can spread the distortion. In this way, we propose a inter-block spreading rule to enhance the security of JPEG steganography.

Given the cover image $\mathbf{X}$ and the seed cost $\mathbf{D}$, the inter-block spreading rule can be summarized as follows:

Step 1: According to the DCT mode, $\mathbf{D}$ is grouped into 64 subcosts $\mathbf{D}_{ab}$ following the equation:

$$\mathbf{D}_{ab}(i, j) = \mathbf{D}\,(i + 8a, j + 8b)\,, \tag{12}$$

where $a = 0, 1, \ldots, \frac{n_1}{8} - 1, b = 0, 1, \ldots, \frac{n_2}{8} - 1, i, j = 0, 1, \ldots, 7$, and the process is shown in Fig. 5.

Step 2: Compute the filtered cost value by using a low-pass filter $\mathbf{L}$ to spread the embedding distortion

$$\widehat{\mathbf{D}}_{ab} = \mathbf{D}_{ab} \otimes \mathbf{L}, \tag{13}$$

where the symbol '$\otimes$' denotes mirror-padded convolution.

Step 3: Merge the filtered subcosts into the final cost $\widehat{\mathbf{D}}$ in the inverse process of Step 1, following the equation:

$$\widehat{\mathbf{D}}\,(i + 8a, j + 8b) = \widehat{\mathbf{D}}_{ab}(i, j). \tag{14}$$

Algorithm 1 Microscale Steganography for J-UNIWARD

Input: A cover image $\mathbf{X}$ with N DCT coefficients; L bits of message $\mathbf{m}$ which determines the relative payload of target $\gamma = L/N$.

Output: The stego image $\mathbf{Y}$.

1: Compute the IF:$\mathbf{W}^{(i)}\left(J^{-1}(\mathbf{B}^{(k,l)})\right)$ with respect to Eq. (3).
2: Decompress cover image $\mathbf{X}$ into spatial domain $J^{-1}(\mathbf{X})$.
3: Enhance the decompressed image $J^{-1}(\mathbf{X})$ into enhanced image $M_\alpha\left(J^{-1}(\mathbf{X})\right)$ by linear UM.
4: Acquire the joint wavelet residual $\mathbf{W}^{(i)^{(\mathrm{MS})}}\left(J^{(-1)}(\mathbf{X})\right)$ according to Eq. (15).
5: Compute embedding distortion ρ_{mn} of each DCT block with respect to Eq. (16).
6: Embed L bits of message $\mathbf{m}$ into cover image $\mathbf{X}$ with STCs according to the embedding distortion ρ, and finally output the stego image $\mathbf{Y}$.

C. Microscale Steganography for J-UNIWARD

The wavelet residual used as *TD* in J-UNIWARD can be improved utilizing microscale schemes. Here we denote with $M_\alpha(\mathbf{X})$ the enhanced image, by using linear UM algorithm to highlight the detail of the decompressed image $\mathbf{X}$. The joint wavelet residual can be defined as

$$\mathbf{W}^{(i)^{(\mathrm{MS})}}(\mathbf{X}) = max\left(\mathbf{W}^{(i)}(\mathbf{X}), \mathbf{W}^{(i)}\left(M_\alpha(\mathbf{X})\right)\right), \tag{15}$$

where the function $max(\mathbf{A}, \mathbf{B})$ creates a matrix and returns the largest value for every element between $\mathbf{A}$ and $\mathbf{B}$. The Eq. (15) guarantees that the elements that owns large residual will be assigned large residual. Therefore, the improved distortion (J-MSUNIWARD) can be denoted by:

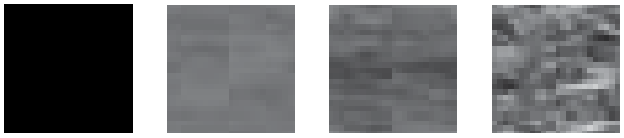
$$\rho_{mn}^{(k,l)} = \sum_{i=1}^{3}\sum_{p=1}^{l_1}\sum_{q=1}^{l_2} \frac{\left|\mathbf{W}_{pq}^{(i)}\left(J^{-1}(\mathbf{B}^{(k,l)})\right)\right|}{\left|\mathbf{W}_{pq}^{(i)^{(\mathrm{MS})}}\left(J^{-1}(\mathbf{X}_{mn})\right)\right| + \sigma}. \tag{16}$$

The pseudo-code of J-MSUNIWARD is presented in Algorithm 1. The inter-block spreading rule is not applied in J-MSUNIWARD, and the experimental result shows it does not reinforce the security performance. In practical, the spreading rule is always adopted in the distortion definition by low-pass filter, like in HILL and MiPOD. In J-UNIWARD, the wavelet filter bank ($\mathbf{K}^1, \mathbf{K}^2$ in Eq. (2)) includes the low-pass components, consequently, we infer the low-pass components serve the similar effect as the inter-block spreading rule.

D. Microscale Steganography for UERD

Actually, in UERD, the DCT energy (*TD*) represents the smoothness of the 8×8 block to some extent. The larger DCT energy is, the more complex the block is, as shown in Fig. 6. From this point of view, if we define the energy more precisely, the security of the steganography will be improved. Similarly, we propose microscale steganography for UERD by precisely defining the DCT energy.

Since the distortion of UERD is totally defined in the DCT domain, filtering in the spatial domain would require



(a) (b) (c) (d)

Fig. 6. For better visual perception, we crop 4 small images (a)-(d), composed of 9 neighbour 8×8 blocks, to explore the relationship between texture and DCT energy. The means of DCT energies of images (a)-(d) are 0, 75, 243, 885 in order. The larger DCT energy is, the more complex the image is.

modules for inverse discrete cosine transform (IDCT), spatial domain filtering, and discrete cosine transform, which will sharply slow down the embedding speed of UERD. To keep low computation complexity, filtering in the DCT domain is introduced.

E. Filtering in the DCT Domain

Like all unitary orthogonal transforms, the DCTs are distributive to matrix multiplications [28]. With this property, one can perform the filtering in the DCT domain. Let $\{f(k,l)\}$ be the 2-D filter, and further assume that 2-D filter $\{f(k,l)\}$ is separable, namely, $\{f(k,l)\}$ can be factorized as $\{f(k,l)\} = v_k h_l$, where v_k and h_l are 1-D filters. In addition, we assume that each component is symmetric, that is, $v_k = v_{-k}$ and $h_l = h_{-l}$. The supports of $\{v_k\}$ and $\{h_l\}$ are $|k| \leq M$ and $|l| \leq N$, and M, N should not exceed 8 as in the previous works in [28].

Let the original image $\mathbf{X}$ and the filtered image $\mathbf{F}$ be composed of non-overlapping 8×8 matrices $\mathbf{X}_{mn}$ and $\mathbf{F}_{mn}$, respectively. Let $\mathbf{V} \equiv [\mathbf{V}_{-1}\mathbf{V}_0\mathbf{V}_1]$ and $\mathbf{H} \equiv [\mathbf{H}_{-1}\mathbf{H}_0\mathbf{H}_1]$, where

$$\mathbf{V}_{-1} = \begin{bmatrix} v_8 & v_7 & \cdots & v_2 & v_1 \\ 0 & v_8 & \ddots & v_3 & v_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & v_8 & v_7 \\ 0 & 0 & \cdots & 0 & v_8 \end{bmatrix} \tag{17}$$

$$\mathbf{V}_0 = \begin{bmatrix} v_0 & v_1 & \cdots & v_6 & v_7 \\ v_1 & v_0 & \ddots & v_5 & v_6 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ v_6 & v_5 & \ddots & v_0 & v_1 \\ v_7 & v_6 & \cdots & v_1 & v_0 \end{bmatrix} \tag{18}$$

and $\mathbf{V}_1 = \mathbf{V}_{-1}^T$. The matrices $\mathbf{H}_{-1}$, $\mathbf{H}_0$ and $\mathbf{H}_1$ are defined similarly to $\mathbf{V}_{-1}$, $\mathbf{V}_0$ and $\mathbf{V}_1$, respectively, by replacing v_k with h_k for all k [29].

The 2-D separable symmetric linear filtering can be represented as a block-based matrix in the form

$$\mathbf{F}_{mn} = \sum_{i=-1}^{1} \sum_{j=-1}^{1} \mathbf{V}_i \mathbf{X}_{(m+i)(n+j)} \mathbf{H}_j{}^{\mathrm{T}}. \tag{19}$$

Note that $\mathbf{F}_{mn}, \mathbf{V}_i, \mathbf{H}_j$ are matrices of size 8×8.

Let $\mathbf{C}$ be the DCT transform matrix. Since the DCT is unitary, $\mathbf{C}^{-1} = \mathbf{C}^{\mathrm{T}}$. Let $\mathbf{X}_{mn}^{\mathrm{d}}$ and $\mathbf{F}_{mn}^{\mathrm{d}}$ be the DCT of $\mathbf{X}_{mn}$ and $\mathbf{F}_{mn}$, respectively, for example, $\mathbf{X}_{mn}^{\mathrm{d}} = \mathbf{C}\mathbf{X}_{mn}\mathbf{C}^{\mathrm{T}}$. Then the filtering in the DCT domain can be represented in

$$\mathbf{F}_{mn}^{\mathrm{d}} = \sum_{i=-1}^{1} \sum_{j=-1}^{1} \mathbf{V}_i^{\mathrm{d}} \mathbf{X}_{(m+i)(n+j)}^{\mathrm{d}} \mathbf{H}_j^{\mathrm{d}\,\mathrm{T}}, \tag{20}$$

where the filtering matrices $\mathbf{V}_i^{\mathrm{d}}$ and $\mathbf{H}_j^{\mathrm{d}}$ are DCT of $\mathbf{V}_i$ and $\mathbf{H}_j$, respectively, which can be precalculated given the filter coefficients $\{v_k\}$ and $\{h_l\}$. Fast computing methods for DCT domain filtering are provided in [28].

F. Definition of MSUERD

Unsharp masking filtering in the DCT domain and spatial domain are applied to UERD, named MSUERD_DCT, MSUERD_SPA, respectively. We define $\mathbf{X}'$ with the enhanced image which is acquired by unsharp masking filtering, and let x'_{kl} be a coefficient in position (k,l) of an 8×8 DCT block in position (m,n) of the filtered image $\mathbf{X}'$, and its block energy D'_{mn} is defined as

$$D'_{mn} = \sum_{k=0}^{7} \sum_{l=0}^{7} |x'_{kl}| \cdot q_{kl}, \quad k,l \in \{0, \cdots, 7\} \tag{21}$$

where x'_{kl}, is the coefficient in the block, $x'_{00} = 0$ to avoid the influence of DC coefficient, and q_{kl} is its corresponding quantization step.

According to the definition of UERD, the improved distortion function is given by

$$\rho_{mn}^{(k,l)} = \begin{cases} \dfrac{0.5 * (q_{(k+1)l} + q_{k(l+1)})}{D'_{mn} + 0.25 * \sum_{d \in \hat{D}'} d} & \text{if } (k,l) \bmod 8 = (0,0) \\ \dfrac{q_{kl}}{D'_{mn} + 0.25 * \sum_{d \in \hat{D}'} d} & \text{otherwise,} \end{cases} \tag{22}$$

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 the block energies of the neighborhood of the mnth block in the enhanced image. Then, the inter-block spreading rule will be adopted to spread the distortion to neighbour inter-block coefficients with a low-pass filter, to obtain the final distortion $\hat{\rho}_{mn}^{(k,l)}$.

V. Microscale Steganography for Side-Informed JPEG Steganography

The distortion function for side-informed JPEG steganography can be factorized into two parts [17]:

$$\rho = \rho_{\mathrm{ori}} \odot \rho_{\mathrm{si}} \tag{23}$$

where ρ_{ori} is the distortion function for the non-side-informed JPEG steganography, and ρ_{si} depends on the quantization rounding error in the process of JPEG compression. The symbol $\odot$ means modulation in SI-UNIWARD and multiplications operation in SI-UERD. In order to expand the microscale steganography to side-informed domain, the straight idea is to replace ρ_{ori} with the microscale version.

TABLE I
DIMENSIONALITY AND KERNELS OF DIFFERENT FEATURE SETS

Feature	Dimensionality	Kernel or composition
DCTR	8,000	DCT bases
GFR	17,000	Gabor filter
J+SRM	35,263	union of SRMQ1 and CC-JRM

VI. EXPERIMENT

A. Setups

In this paper, two disjoint image sets BOSSbase 1.01 [24] and BOWS-2-OrigEP3 [25] (simplified as BOWS2), both of which contain 10,000 grayscale 512×512 images, are adopted as the image database. The original images are then JPEG compressed using quality factors 75 and 95, so we have six image databases in the format of PGM and JPEG, which act as the precover (PGM) and cover (JPEG) for side-informed and non-side-informed JPEG embedding, respectively. All tested embedding algorithms are simulated at their corresponding payload-distortion bound for payloads $R \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ bit per non-zero cover AC coefficient (bpnzac).

Three state-of-the-art feature sets (DCTR [21], GFR [22], J+SRM [23]) are selected for steganalysis of JPEG image. The former two steganalyzers are formed from noise residuals computed by convolving the decompressed (non-rounded) JPEG image with different kernels. Then the residuals are quantized and the histograms of the quantized residuals are calculated as the final statistical feature [30]. J+SRM is the union of CC-JRM [23] and the spatial domain Rich Model (SRMQ1) [31]. The dimensionality and kernels of different feature sets are shown in Table I.

The detectors are trained as binary classifiers implemented using the FLD ensemble with default settings. A separate classifier is trained for each embedding algorithm and payloads. The ensemble by default minimizes the total classification error probability under equal priors $P_{\mathrm{E}} = min_{P_{\mathrm{FA}}} \frac{1}{2}(P_{\mathrm{FA}} + P_{\mathrm{MD}})$, where P_{FA} and P_{MD} are the false-alarm probability and the missed-detection probability respectively. The ultimate security is qualified by average error rate $\overline{P}_{\mathrm{E}}$ averaged over ten 5000/5000 database splits, and larger $\overline{P}_{\mathrm{E}}$ means stronger security.

B. Determining the Parameters of Microscale Steganography

BOSSbase is set as the final test set where we compare security performance of microscale steganography with other steganographic methods. As for parameter setting, a disjoint set BOWS2 is chosen and split into two sets: training (5000 images), validation (5000 images). The optimal parameters are determined on the validation set by traversal search with a step of 0.1, when payload is 0.4 bpnzac, against DCTR feature.

Unsharp masking consists simply of generating a sharp image by subtracting from an image a blurred version of itself, which can be seen as one filter operation. For efficiency and simplifying searching optimal parameters, we further consider linear unsharp mask filter as a symmetric 3×3 filter operator which can be product of two 1-D UM filters. Specifically,

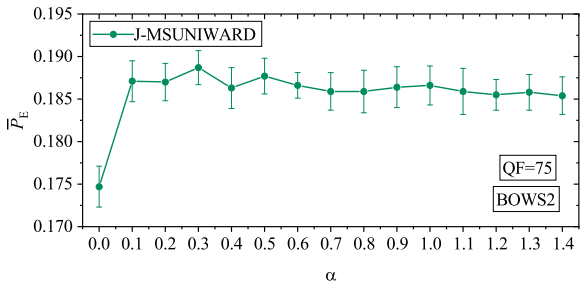


Fig. 7. Average detection error $\overline{P}_{\mathrm{E}}$ of J-MSUNIWARD as a function of the scaling factor α of unsharp masking at 0.4 bpnzac when steganalyzing with DCTR on BOWS2.

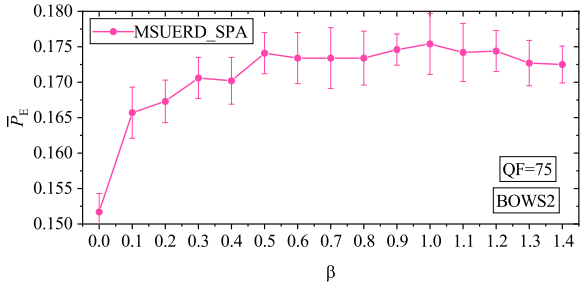


Fig. 8. Average detection error $\overline{P}_{\mathrm{E}}$ of MSUERD_SPA as a function of the scaling factor β of unsharp masking at 0.4 bpnzac when steganalyzing with DCTR on BOWS2.

$$\frac{1}{2.56}\begin{bmatrix} 0.09 & -0.48 & 0.09 \\ -0.48 & 2.56 & -0.48 \\ 0.09 & -0.48 & 0.09 \end{bmatrix} \quad \text{(a)} \qquad \frac{1}{9}\begin{bmatrix} 1 & -3 & 1 \\ -3 & 9 & -3 \\ 1 & -3 & 1 \end{bmatrix} \quad \text{(b)}$$

Fig. 9. The unsharp mask filter operator (a) is utilized in J-MSUNIWARD, while (b) is used in MSUERD.

1-D UM filter can be represented as $\frac{1}{1+2\epsilon}[-\epsilon \ 1+2\epsilon \ -\epsilon]$, where ϵ is the scaling factor. Therefore, the 2-D filter operator of linear UM filter can be represented as

$$\frac{1}{(1+2\epsilon)^2}\begin{bmatrix} \epsilon^2 & -\epsilon-2\epsilon^2 & \epsilon^2 \\ -\epsilon-2\epsilon^2 & (1+2\epsilon)^2 & -\epsilon-2\epsilon^2 \\ \epsilon^2 & -\epsilon-2\epsilon^2 & \epsilon^2 \end{bmatrix}. \quad (24)$$

With regard to filtering in the DCT domain, filter size of 3×3 means $M = N = 3$, namely, $\{v_k\} = 0$ and $\{h_l\} = 0$ when $|k| > 1$ or $|l| > 1$ mentioned in Section IV-E. The scaling factor ϵ of linear UM will be represented by α in J-MSUNIWARD, and β in MSUERD, respectively.

The results of searching for scaling factor are shown in Fig. 7 and Fig. 8. $\alpha = 0$ and $\beta = 0$ mean the seed algorithm J-UNIWARD and UERD. As for J-UNIWARD, the testing error has a large promotion from $\alpha = 0$ to $\alpha = 0.1$, and then it turns to be gradual. $\alpha = 0.3$ outperforms other values on the validation set. When it comes to MSUERD, $\beta = 1$ performs best, so the filter operators of UM for J-MSUNIWARD and MSUERD are presented in Fig. 9.

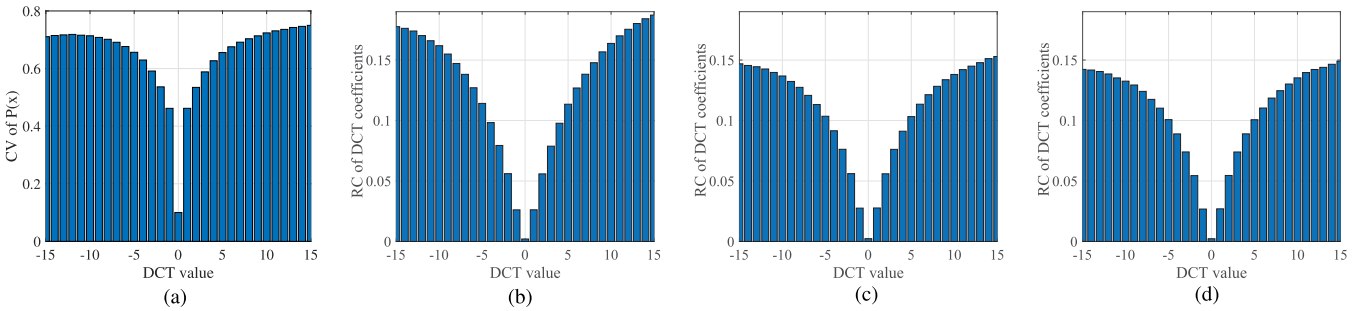


Fig. 10. (a) is CV of $p(x)$, and (b)-(d) are the the RC of the DCT coefficients with UERD, MSUERD_SPA, MSUERD_DCT at 0.3 bpnzac, respectively.

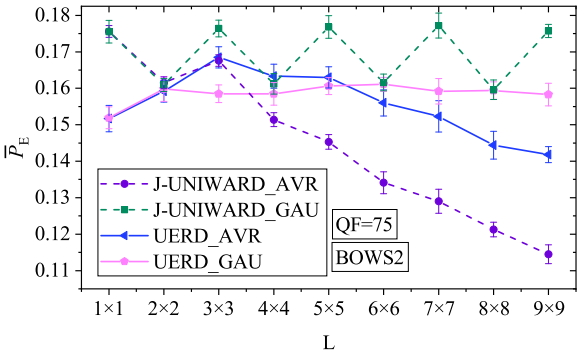


Fig. 11. Average detection error $\overline{P}_{\mathrm{E}}$ of UERD/J-UNIWARD applying inter-block spreading rule as a function of the filter size L and UM filter type at 0.4 bpnzac when steganalyzing with DCTR on BOWS2.

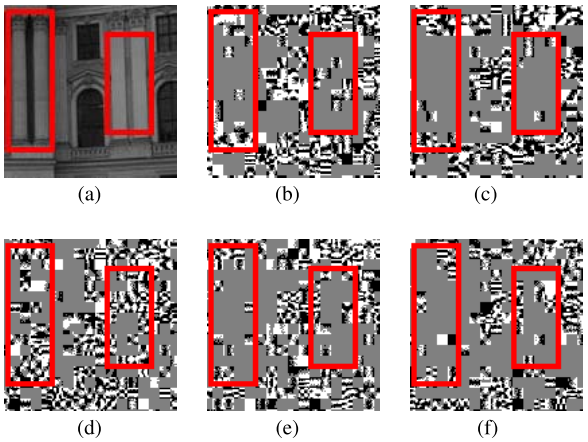


Fig. 12. The changes (b)-(f) in the spatial domain caused by DCT embedding with respect to the cover image (a) with payload 0.5 bpnzac, QF=75, using J-UNIWARD, J-MSUNIWARD, UERD, MSUERD_DCT, MSUERD_SPA, respectively. White pixels represent positive changes; dark pixels represent negative changes; gray pixels mean no changes. Regularly, fewer changes in smooth area mean better security. The area of gray zone of the improved method in the rectangle (smooth region) is larger than that of the seed method, like (c) to (b) and (e)(f) to (d), which indicates the proposed schemes are effective. (a) Cover image. (b) J-UNIWARD. (c) J-MSUNIWARD. (d) UERD. (e) MSUERD_DCT. (f) MSUERD_SPA.

In the case of inter-block spreading rule, two kinds of filters and several sizes are traversed under the same experimental condition. To distinguish different options, we use the syntax of names following the convention:

$$name = \{seed\ distortion\}\_\{filter\} \tag{25}$$

The field $seed\ distortion \in \{$UERD, J-UNIWARD$\}$ indicates the seed distortion used and $filter \in \{$AVR, GAU$\}$ indicates the type of filter used, AVG for using an average, GAU for using a Gaussian low-pass filter. Filter size of 1×1 means no filter used. As depicted in Fig. 11, the average filter with the size of 3×3 is more profitable than others when applied to UERD, and is adopted in MSUERD subsequently. However, the testing errors of J-UNIWARD cooperated with inter-block spreading rule are below that J-UNIWARD (filter of size 1×1 in Fig. 11), meaning that the inter-block spreading rule is not valid for J-UNIWARD, so we abandon it in J-MSUNIWARD.

To sum up, $\alpha = 0.3$ and no low-pass filter in J-MSUNIWARD, $\beta = 1$ and average filter with the size of 3×3 in MSUERD are set.

C. Visualizing Embedding Changes

To verify whether the proposed algorithm can effectively improve the distribution of embedding changes, we give an example to visualize the embedding changes. A sample cover image of size 128×128 pixels, containing smooth regions, edges, and textured regions, as shown in Fig. 12(a), is cropped from "1013.jpg" in BOSSbase. We show the changes in the spatial domain caused by DCT embedding with 0.5 bpnzac in Fig. 12(b-f). White pixels represent positive changes; dark pixels represents negative changes; gray pixels means no changes. Regularly, fewer changes in smooth area is better. The red rectangle part is the pillow, which is seen as the smooth region. The area of gray zone in the rectangle of the improved method is larger than that of the seed method, like (c) to (b) and (e)(f) to (d) in Fig. 12, which indicates the proposed schemes are effective.

D. Analysis Based on CV

In UERD, the distortion was derived from observation of the coefficient of variation (CV) [17] denoted by:

$$CV(x) = \frac{\sigma(x)}{\mu(x)} \tag{26}$$

where $\mu(x)$ and $\sigma(x)$ are the mean and standard deviation of the histogram of DCT coefficients $p(x)$ over 10000 JPEG images with QF=75 from BOSSbase. The motivation of UERD is the generative uniform embedding strategy, which

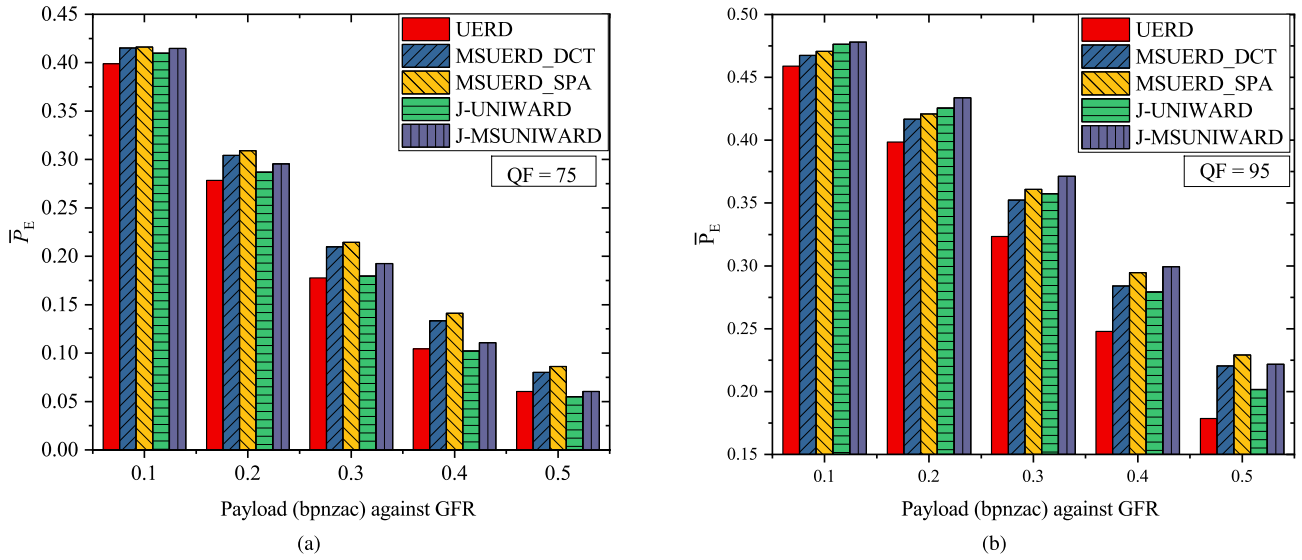


Fig. 13. Detection error $\overline{P}_E$ for five steganographic algorithms against GFR feature versus payload, JPEG quality factor 75 and 95.

stresses the relative change of each bin ought to be proportional to the *CV(x)*. We utilize the similarity between the *CV* of the cover images and the relative change (*RC*) of DCT coefficients to conjecture the security of the algorithm. Every algorithms has its *RC* figure, as shown in Fig. 10. The more similar the *CV* of cover images and the *RC* of DCT coefficients are, the more secure the algorithm is. It is not difficult to find that the bins of improved schemes are more gradual, which are more similar to the *CV*. In order to quantitatively describe the relationship between the similarity of the mentioned two factors, we introduce a histogram distance metric named BRD [32], which is robust to partial matching and histogram normalization. Given two normalized histograms $\mathbf{p}$ and $\mathbf{q}$ with n bins, the BRD $d_{\mathrm{BRD}}(\mathbf{p}, \mathbf{q})$ is defined as

$$d_{\mathrm{BRD}}(\mathbf{p}, \mathbf{q}) = n - \|\mathbf{p} + \mathbf{q}\|_2 \sum_{i=1}^{n} \frac{p_i q_i}{(p_i + q_i)^2} \tag{27}$$

Table VI shows the BRD distance between the histograms of *CV* and *RC*, and detection error rates of different embedding methods at 0.3 and 0.4 bpnzac. It can be seen that the detection error rates are inversely proportional to the BRD distance, which verifies the improvements of MSUERD with respect to UERD. However, the relationship between BRD distance and detection error rates is limited to the general uniform embedding strategy, excluding J-UNIWARD series.

E. Statistical Significance of the Improved Accuracy

In order to confirm the statistical significance of the improved accuracy, a 5×2 fold cross-validated paired t-test defined by Dietterich [33] is realized between the error rates of the original and the improved algorithms, which defines a statistic value t that has an approximately t distribution with 5 degrees of freedom in the null hypothesis. The hypotheses are denoted as follow:

$$H_0 : \mu_1 = \mu_2; \quad H_1 : \mu_1 > \mu_2.$$

in which μ_1 and μ_2 are the mean values of testing errors of the original and the improved schemes, H_0 represents that there is no significant differences between them, while H_1 means that the improved accuracy do exists rather than random chance [34].

The significance level for the test is set to 0.05 ($t_{0.025}(5) = 2.5706$), which is usually recommended as a convenient cutoff level to reject the null hypothesis, given that it were true. We underline the testing error in Table III-VII, where the improvement of the MS-version compared to the seed algorithm is statistically significant.

F. Performance of Non-Side-Informed JPEG Steganography

The performance between the improved algorithms and the original algorithms would be compared, respectively. As shown in Fig. 13, Table II and Table III. The improved schemes utilizing microscale steganography perform better than the original schemes in most cases, when steganalyzing DCTR, GFR and J+SRM features.

J-MSUNIWARD has slight improvements than J-UNIWARD with QF=75, and the increments become larger with QF=95 in terms of the average testing error. MSUERD_DCT, MSUERD_SPA and UERD are compared subsequently, where MSUERD_SPA and MSUERD_DCT perform better than UERD. MSUERD_SPA surpasses UERD with a maximum boost of 3.71% at 0.3 bpnzac when steganalyzed with GFR and QF=75. When QF changes to 95, the increment peak reaches 4.23% at 0.5 bpnzac against GFR. MSUERD_DCT should generate the same results as MSUERD_SPA mathematically, the varying experimental results are possibly due to that MSUERD_DCT suffers from the padding of boundary block and the blocking effect when filtering in the DCT domain. The improvements of MSUERD are achieved with microscale steganography and inter-block spreading rule.

Selection-channel-aware attack is also executed to verify the improvements of the proposed schemes. SCA-GFR [30]

TABLE II

NON-SIDE-INFORMED, QF=75: DETECTABILITY IN TERMS OF $\overline{P}_E$ VERSUS EMBEDDED PAYLOAD SIZE IN BITS PER NON-ZERO COVER AC COEFFICIENT (BPNZAC) FOR PRIOR ARTAND APPLIED TO OUR SCHEME ON BOSSBASE 1.01 USING THE FLD ENSEMBLE CLASSIFIER WITH THREE FEATURE SETS. BOLD FONT MEANS THE PROMOTION IS STATISTICALLY SIGNIFICANT WITH RESPECT TO SEED ALGORITHM

Feature	Embedding Method	0.1	0.2	0.3	0.4	0.5
GFR	UERD	.3983 ± .0015	.2796 ± .0035	.1774 ± .0030	.1045 ± .0037	.0592 ± .0021
	MSUERD_DCT	**.4151 ± .0035**	**.3042 ± .0027**	**.2099 ± .0034**	**.1333 ± .0022**	**.0801 ± .0021**
	MSUERD_SPA	**.4162 ± .0035**	**.3092 ± .0039**	**.2145 ± .0034**	**.1412 ± .0050**	**.0862 ± .0023**
	J-UNIWARD	.4086 ± .0033	.2871 ± .0020	.1783 ± .0016	.1016 ± .0032	.0546 ± .0014
	J-MSUNIWARD	.4146 ± .0027	**.2955 ± .0020**	**.1925 ± .0030**	**.1103 ± .0026**	**.0598 ± .0016**
DCTR	UERD	.4295 ± .0018	.3324 ± .0029	.2302 ± .0030	.1468 ± .0014	.0871 ± .0017
	MSUERD_DCT	**.4366 ± .0028**	**.3461 ± .0033**	**.2576 ± .0033**	**.1762 ± .0026**	**.1130 ± .0020**
	MSUERD_SPA	**.4401 ± .0030**	**.3555 ± .0028**	**.2616 ± .0045**	**.1790 ± .0026**	**.1108 ± .0019**
	J-UNIWARD	.4379 ± .0022	.3416 ± .0023	.2389 ± .0019	.1551 ± .0013	.0920 ± .0024
	J-MSUNIWARD	.4419 ± .0018	**.3507 ± .0028**	**.2511 ± .0030**	**.1671 ± .0025**	**.1002 ± .0013**
J+SRM	UERD	.4342 ± .0048	.3345 ± .0038	.2374 ± .0047	.1536 ± .0038	.0959 ± .0029
	MSUERD_DCT	.4416 ± .0036	**.3520 ± .0048**	**.2621 ± .0037**	**.1804 ± .0046**	**.1164 ± .0028**
	MSUERD_SPA	**.4485 ± .0035**	**.3664 ± .0039**	**.2732 ± .0045**	**.1901 ± .0054**	**.1212 ± .0025**
	J-UNIWARD	.4579 ± .0024	.3769 ± .0035	.2792 ± .0038	.1928 ± .0036	.1238 ± .0034
	J-MSUNIWARD	.4634 ± .0032	**.3861 ± .0033**	**.2936 ± .0035**	**.2050 ± .0032**	**.1384 ± .0030**

TABLE III

NON-SIDE-INFORMED, QF=95: DETECTABILITY IN TERMS OF $\overline{P}_E$ VERSUS EMBEDDED PAYLOAD SIZE IN BITS PER NON-ZERO COVER AC COEFFICIENT (BPNZAC) FOR PRIOR ARTAND APPLIED TO OUR SCHEME ON BOSSBASE 1.01 USING THE FLD ENSEMBLE CLASSIFIER WITH THREE FEATURE SETS. BOLD FONT MEANS THE PROMOTION IS STATISTICALLY SIGNIFICANT WITH RESPECT TO SEED ALGORITHM

Feature	Embedding Method	0.1	0.2	0.3	0.4	0.5
GFR	UERD	.4598 ± .0017	.3985 ± .0021	.3241 ± .0031	.2494 ± .0022	.1782 ± .0031
	MSUERD_DCT	.4674 ± .0026	**.4167 ± .0020**	**.3524 ± .0032**	**.2841 ± .0029**	**.2205 ± .0047**
	MSUERD_SPA	.4707 ± .0025	**.4209 ± .0024**	**.3608 ± .0024**	**.2946 ± .0049**	**.2292 ± .0026**
	J-UNIWARD	.4764 ± .0028	.4256 ± .0037	.3574 ± .0026	.2792 ± .0024	.2017 ± .0035
	J-MSUNIWARD	.4787 ± .0017	.4315 ± .0039	**.3712 ± .0044**	**.2964 ± .0043**	**.2212 ± .0024**
DCTR	UERD	.4770 ± .0025	.4335 ± .0026	.3755 ± .0045	.3011 ± .0040	.2256 ± .0031
	MSUERD_DCT	.4817 ± .0028	**.4421 ± .0021**	**.3905 ± .0026**	**.3315 ± .0027**	**.2642 ± .0026**
	MSUERD_SPA	.4836 ± .0016	**.4461 ± .0023**	**.3973 ± .0025**	**.3388 ± .0028**	**.2715 ± .0026**
	J-UNIWARD	.4884 ± .0018	.4536 ± .0026	.4027 ± .0026	.3352 ± .0032	.2607 ± .0031
	J-MSUNIWARD	.4897 ± .0018	.4605 ± .0022	**.4145 ± .0022**	**.3568 ± .0044**	**.2871 ± .0023**
J+SRM	UERD	.4839 ± .0030	.4423 ± .0033	.3775 ± .0050	.3021 ± .0049	.2223 ± .0021
	MSUERD_DCT	.4883 ± .0025	**.4531 ± .0041**	**.4065 ± .0047**	**.3411 ± .0046**	**.2720 ± .0040**
	MSUERD_SPA	.4897 ± .0030	**.4594 ± .0040**	**.4129 ± .0045**	**.3501 ± .0037**	**.2775 ± .0042**
	J-UNIWARD	.4935 ± .0026	.4725 ± .0025	.4336 ± .0022	.3782 ± .0031	.3066 ± .0036
	J-MSUNIWARD	.4962 ± .0027	.4755 ± .0037	**.4429 ± .0021**	**.3910 ± .0036**	**.3224 ± .0031**

is chosen as the detector due to its strong detectability, and analogous results are displayed in Fig. 15. MSUERD_SPA and MSUERD_DCT perform better than UERD. The promotion of J-MSUNIWARD is small when QF=75, and turns to be large when QF=95.

It is obvious that the improvement of MSUERD is more obvious than J-MSUNIWARD on both non-side-informed and side-informed steganography, which is possibly owing to that J-UNIWARD considers the smoothness of the cover image more seriously than MSUERD. In J-UNIWARD, the smoothness of the coefficients is related to a 23×23 neighbour residuals of three filter banks, while UERD considers the smoothness of a coefficient merely counting on non-zero coefficients in a single block.

Overall, MSUERD_SPA outperforms others under the payloads of 0.2-0.5 bpnzac when resisting GFR and DCTR with QF=75, and J-MSUNIWARD owns the most secure performance in most cases when QF=95. The statistic test results show that the improvements are statistically significant in most cases.

G. Performance of Side-Informed JPEG Steganography

For JPEG steganography with side-information, Table IV and Table V show the security performance of the involved schemes against three features with quality factor 75 and 95. For better visual experience, Fig. 14 displays the security performance of the involved schemes against GFR. SI-MSUNIWARD has mild promotion than SI-UNIWARD and the increment becomes considerable with the increase of payload against three steganalyzer features. Numerically, SI-MSUERD_DCT and SI-MSUERD_SPA perform better than SI-UERD, and SI-MSUERD_SPA has an obvious improvement at 0.5 bpnzac by around 4%∼5% with both QF=75 and QF=95 against GFR. SI-MSUNIWARD

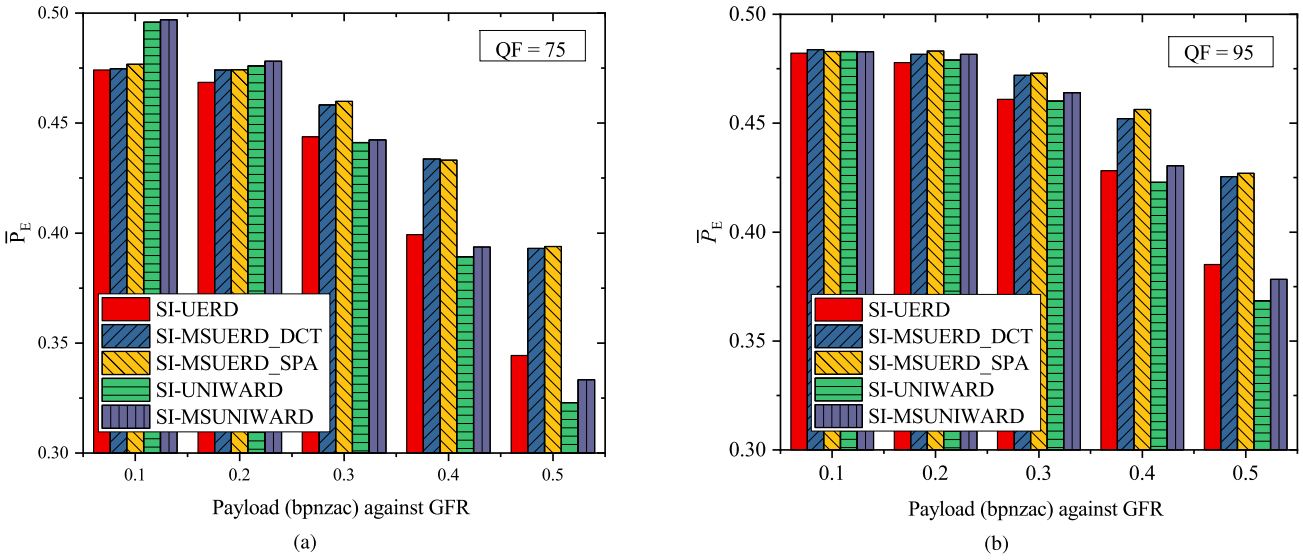


Fig. 14. Detection error $\overline{P}_E$ of the side-informed JPEG steganography against GFR feature versus payload, with quality factor 75 and 95.

TABLE IV

SIDE-INFORMED, QF=75: DETECTABILITY IN TERMS OF $\overline{P}_E$ VERSUS EMBEDDED PAYLOAD SIZE IN BITS PER NON-ZERO COVER AC COEFFICIENT (BPNZAC) FOR PRIOR ART AND APPLIED TO OUR SCHEME ON BOSSBASE 1.01 USING THE FLD ENSEMBLE CLASSIFIER WITH THREE FEATURE SETS. BOLD FONT MEANS THE PROMOTION IS STATISTICALLY SIGNIFICANT WITH RESPECT TO SEED ALGORITHM

Feature	Embedding Method	0.1	0.2	0.3	0.4	0.5
GFR	SI-UERD	.4742 ± .0023	.4685 ± .0023	.4438 ± .0030	.3993 ± .0039	.3443 ± .0027
	SI-MSUERD_DCT	.4747 ± .0022	.4742 ± .0022	**.4583 ± .0026**	**.4337 ± .0029**	**.3930 ± .0035**
	SI-MSUERD_SPA	.4767 ± .0027	.4743 ± .0026	**.4599 ± .0025**	**.4331 ± .0028**	**.3939 ± .0036**
	SI-UNIWARD	.4959 ± .0029	.4760 ± .0019	.4412 ± .0020	.3892 ± .0032	.3229 ± .0027
	SI-MSUNIWARD	.4973 ± .0027	.4792 ± .0029	.4426 ± .0041	.3957 ± .0016	**.3334 ± .0033**
DCTR	SI-UERD	.4716 ± .0027	.4700 ± .0018	.4573 ± .0032	.4256 ± .0027	.3831 ± .0030
	SI-MSUERD_DCT	.4748 ± .0031	.4743 ± .0016	**.4658 ± .0025**	**.4450 ± .0015**	**.4111 ± .0022**
	SI-MSUERD_SPA	.4817 ± .0028	.4725 ± .0020	**.4661 ± .0036**	**.4441 ± .0032**	**.4130 ± .0023**
	SI-UNIWARD	.4947 ± .0032	.4833 ± .0024	.4534 ± .0017	.4062 ± .0032	.3441 ± .0022
	SI-MSUNIWARD	.4969 ± .0028	.4842 ± .0025	.4553 ± .0028	.4107 ± .0039	**.3524 ± .0021**
J+SRM	SI-UERD	.4777 ± .0022	.4751 ± .0047	.4503 ± .0017	.4093 ± .0037	.3579 ± .0052
	SI-MSUERD_DCT	.4794 ± .0016	.4792 ± .0039	**.4627 ± .0038**	**.4334 ± .0040**	**.3887 ± .0040**
	SI-MSUERD_SPA	.4793 ± .0027	.4809 ± .0036	**.4665 ± .0037**	**.4354 ± .0034**	**.3943 ± .0032**
	SI-UNIWARD	.4980 ± .0037	.4875 ± .0023	.4607 ± .0023	.4235 ± .0035	.3661 ± .0037
	SI-MSUNIWARD	.4996 ± .0032	.4904 ± .0036	.4648 ± .0034	.4274 ± .0045	.3708 ± .0032

outperforms the original methods at 0.5 bpnzac by 1.4% against GFR for QF=95. As for small payloads, the promotions are not significant, because the seed algorithms are too safe to improve, where the testing error rates are near to 50%.

On the whole, SI-MSUNIWARD is more surreptitious in small payloads, while SI-MSUERD_SPA owns best security performance in large payloads.

H. Comparison With Other Image Enhancement Methods

The unsharp mask is selected as our image enhance methods, for it not only highlighting the detail but also maintaining the characteristics of image. We have also tried other image enhancement methods to prove the generalizability of microscale steganography. Histogram Equalization, Gamma Correction, Localcontrast Enhancement with default setting in Matlab Image Processing Toolbox are selected as the enhancement techniques in microscale steganography for comparison. Since the experiments are designed to test the effectiveness of image enhancement techniques, inter-block spreading rule is not utilized here. The payload is 0.3 bpnzac, the steganalyzer is GFR, and the results are shown in Table VII. We can see that most of the enhancement methods can improve the security performance, and linear unsharp mask behaves best among them.

I. Comparison of Computational Complexity

Now that all schemes utilize the same framework of minimal distortion embedding with simulate coding, we separately evaluate computational complexity of the distortions. Following the complexity calculation method in [17], the computational

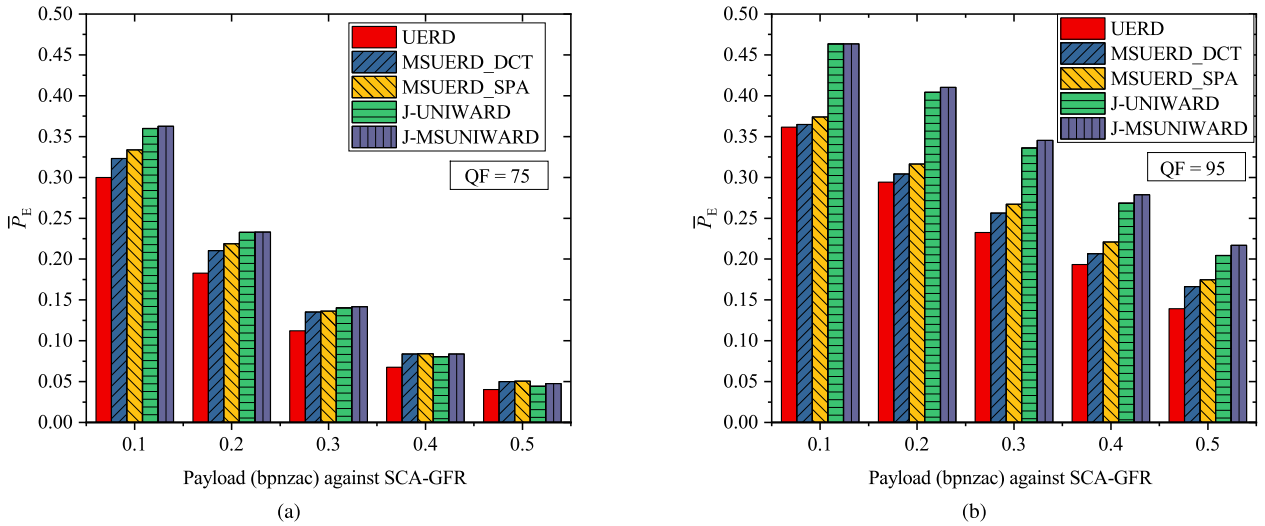


Fig. 15. Detection error $\overline{P}_E$ for five steganographic algorithms against SCA-GFR features versus payload, JPEG quality factor 75 and 95.

TABLE V

SIDE-INFORMED, QF=95: DETECTABILITY IN TERMS OF $\overline{P}_E$ VERSUS EMBEDDED PAYLOAD SIZE IN BITS PER NON-ZERO COVER AC COEFFICIENT (BPNZAC) FOR PRIOR ART AND APPLIED TO OUR SCHEME ON BOSSBASE 1.01 USING THE FLD ENSEMBLE CLASSIFIER WITH THREE FEATURE SETS. BOLD FONT MEANS THE PROMOTION IS STATISTICALLY SIGNIFICANT WITH RESPECT TO SEED ALGORITHM

Feature	Embedding Method	0.1	0.2	0.3	0.4	0.5
GFR	SI-UERD	.4822 ± .0028	.4778 ± .0035	.4610 ± .0043	.4282 ± .0040	.3851 ± .0040
	SI-MSUERD_DCT	.4838 ± .0025	.4816 ± .0034	**.4720 ± .0032**	**.4521 ± .0031**	**.4255 ± .0052**
	SI-MSUERD_SPA	.4829 ± .0034	.4831 ± .0031	**.4730 ± .0034**	**.4563 ± .0019**	**.4270 ± .0037**
	SI-UNIWARD	.4829 ± .0029	.4790 ± .0035	.4578 ± .0033	.4229 ± .0041	.3661 ± .0034
	SI-MSUNIWARD	.4828 ± .0027	.4816 ± .0027	.4640 ± .0021	.4304 ± .0029	**.3801 ± .0030**
DCTR	SI-UERD	.4754 ± .0017	.4739 ± .0023	.4710 ± .0026	.4491 ± .0028	.4149 ± .0028
	SI-MSUERD_DCT	.4748 ± .0031	.4743 ± .0016	.4709 ± .0023	**.4592 ± .0027**	**.4378 ± .0038**
	SI-MSUERD_SPA	.4792 ± .0028	.4739 ± .0030	.4686 ± .0031	**.4578 ± .0026**	**.4369 ± .0020**
	SI-UNIWARD	.4744 ± .0028	.4705 ± .0035	.4567 ± .0043	.4236 ± .0030	.3683 ± .0040
	SI-MSUNIWARD	.4747 ± .0018	.4728 ± .0030	.4616 ± .0017	**.4329 ± .0018**	**.3831 ± .0021**
J+SRM	SI-UERD	.4695 ± .0033	.4691 ± .0028	.4565 ± .0038	.4330 ± .0024	.3968 ± .0046
	SI-MSUERD_DCT	.4706 ± .0021	.4673 ± .0041	.4586 ± .0045	**.4438 ± .0035**	**.4211 ± .0050**
	SI-MSUERD_SPA	.4735 ± .0028	.4694 ± .0018	.4619 ± .0030	**.4455 ± .0040**	**.4192 ± .0035**
	SI-UNIWARD	.4711 ± .0030	.4678 ± .0025	.4570 ± .0036	.4347 ± .0032	.3905 ± .0028
	SI-MSUNIWARD	.4730 ± .0027	.4709 ± .0026	.4621 ± .0024	.4415 ± .0043	**.4004 ± .0030**

TABLE VI

THE BRD DISTANCE BETWEEN THE HISTOGRAMS OF CV AND RC, AND THE DETECTION ERROR RATES OF DIFFERENT EMBEDDING METHODS AT 0.3 BPNZAC AND 0.4 BPNZAC AGAINST GFR

Embedding method	Payload	BRD	Detection error rates
UERD	0.3	2.5677	0.1774
MSUERD_DCT	0.3	1.9157	0.2099
MSUERD_SPA	0.3	1.9093	0.2145
UERD	0.4	1.9565	0.1045
MSUERD_DCT	0.4	1.6086	0.1333
MSUERD_SPA	0.4	1.6065	0.1412

TABLE VII

THE MEAN TESTING ERROR OF MICROSCALE STEGANOGRAPHY ON BOSSBASE 1.01 USING DIFFERENT IMAGE ENHANCEMENT TECHNIQUES WHEN THE PAYLOAD = 0.3 BPNZAC AGAINST GFR

Image Enhance methods	Seed methods	
	UERD	J-UNIWARD
None	0.1774±0.0030	0.1783±0.0016
Histogram Equalization	0.1861±0.0033	0.1836±0.0020
Gamma Correction	0.1871±0.0028	0.1854±0.0031
Localcontrast Enhancement	0.1883±0.0029	0.1872±0.0026
Linear Unsharp Mask	0.2009±0.0032	0.1925±0.0030

complexity is represented by the number of mathematical operations. Since addition and subtraction are linear computational complexity that is far less than multiplication and division, we focus on the number of latters. The division has the same asymptotic complexity as multiplication [35], so computational complexity of the algorithms can be in the form of the number of multiplications. Since the *IF* of UERD, J-UNIWARD and their MS-version can be precalculated, we ignore the computational complexity of *IF* and just calculate that of *TD* and the division in Eq. (1).

TABLE VIII
COMPUTATIONAL COMPLEXITY OF ALGORITHMS FOR AN 8×8 BLOCK

Algorithm	Number of multiplications
UERD	129
MSUERD_DCT	865(non-sparse), 577(sparse)
MSUERD_SPA	1665
JUNIWARD	104209
J-MSUNIWARD	107414

UERD requires 8×8 multiplications in Eq. (5) and 1 multiplications and 8×8 divisions in Eq. (6), so the final computational complexity is roughly $64 + 1 + 64 = 129$. As concluded in [28], the spatial domain filtering for DCT coefficients would require 1536 multiplications, and DCT domain filtering needs 736 in the nonsparse case and 448 in the sparse case[1] for one 8×8 DCT block. Intuitively, MSUERD_DCT is $129 + 736 = 865$ in nonsparse case and $129 + 448 = 577$ in sparse case, and MSUERD_SPA is $129 + 1536 = 1665$.

As for J-UNIWARD, the computation mainly includes dequantization, 2-D IDCT, wavelet filtering and the division of Eq. (4). One block involves 64 multiplications for dequantization and 192 multiplications for fast 2-D IDCT [36]. The computational complexity of the wavelet filter using fast convolution product for an $n \times n$ image is $n^2 \ln(n^2)$ [37], which can be evenly divided into $64\ln(n^2)$ per block. Specifically, three-time wavelet filtering needs $3 \times 64\ln(512^2) \approx 2396$ per block for 512×512 image. The most expensive computational cost in J-UNIWARD is the division of Eq. (4). As for one block, the number of divisions of Eq. (4) is $3 \times 23 \times 23 \times 8 \times 8 = 101568$. Consequently, the computational complexity of J-UNIWARD is roughly $64 + 192 + 2396 + 101568 = 104220$. J-MSUNIWARD additionally needs one-time unsharp mask filtering and three-time wavelet filtering with respect to J-UNIWARD according to Eq. (15), so the computational complexity of J-MSUNIWARD is $104220 + 4 \times 64\ln(512^2) \approx 107414$.

In summary, we have the approximate computational complexity of presented algorithms for one 8×8 block in a 512×512 image in Table VIII. UERD owns the cheapest computation cost. In nonsparse case, the computation complexity of MSUERD_DCT is about 7 times of that of UERD and half of that of MSUERD_SPA. J-MSUNIWARD has similar computational complexity as J-UNIWARD.

Furthermore, we randomly selected 1000 512×512 images to measure the average distortion definition time of some of the mentioned steganographic method at 0.3 bpnzac with different quality factors. The time measurement is performed with Matlab 2017a on a 3.20 GHz Intel Core i5 desktop computer with 8GB of memory running a 64-bit Ubuntu. As shown in Fig. 16, the computational time of J-MSUNIWARD and J-UNIWARD is far larger than MSUERD_DCT, MSUERD_SPA and UERD. Numerically, shown in Table IX, the time of J-MSUNIWARD is nearly the same as that of J-UNIWARD. The time of

[1]The sparse case is very common in DCT-based compression. Actually, DCT blocks are typically sparser, i.e., less than 25% of coefficients are nonzero, especially for low-bit-rate compression.

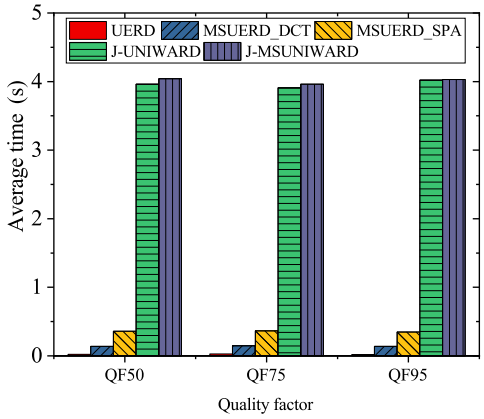


Fig. 16. Average computational time of 1,000 images using UERD, MSUERD_DCT, MSUERD_SPA, J-UNIWARD and J-MSUNIWARD for JPEG steganography (0.4 bpnzac), respectively.

TABLE IX
AVERAGE TIME OF COMPUTING DISTORTION ON RANDOMLY SELECTED 1000 IMAGES

Embedding method	QF50 (s)	QF75 (s)	QF95 (s)
UERD	0.0215	0.0269	0.0201
MSUERD_DCT	0.1398	0.1458	0.1376
MSUERD_SPA	0.3580	0.3535	0.3466
J-UNIWARD	3.9628	3.9081	4.0202
J-MSUNIWARD	4.0425	3.9621	4.0312

MSUERD_DCT is nearly 7 times of UERD and about half of MSUERD_SPA, which coincides with the theoretical evaluation given above approximately. It is worth noting that MSUERD_DCT achieves better security performance in most cases with considerable computational complexity.

VII. CONCLUSIONS

In this paper, we extend the former work, microscale steganography in the spatial domain, into adaptive JPEG steganography. Before distortion definition, the cover image would be preprocessed with a microscope, which can seize the texture areas more precisely and thus improve the security of adaptive steganography. Here, linear unsharp masking plays the role of the microscope. As for J-UNIWARD, the image is filtered in the spatial domain, for its distortion is already calculated on the spatial domain. When it comes to UERD, the DCT domain filtering was introduced in order to maintain the low computational complexity, and inter-block spreading rule is cooperated to further reinforce security. The experimental results verify that the proposed scheme does work. The improved schemes outperform the original steganography algorithm. It is worth mentioning that MSUERD_DCT achieves better security performance in large payloads (0.2-0.5 bpnzac) with lower computational complexity against steganalyzer GFR when QF=75 with respect to J-UNIWARD.

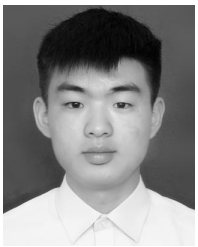
Since the improvements are based on block texture descriptor (*TD*) of existing distortion functions, the definition of *TD* will be reconsidered in the future. In addition, to design a better inner block distinguishing factor (*IF*) is also a part of our future work.

ACKNOWLEDGMENT

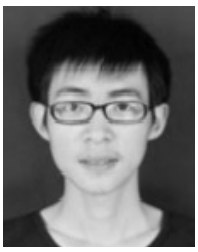
The authors would like to thank DDE Laboratory of SUNY Binghamton for sharing the source code of steganography, steganalysis and ensemble classifier on the webpage (http://dde.binghamton.edu/download/).

REFERENCES

[1] T. Pevný and J. Fridrich, "Benchmarking for steganography," in *Proc. 10th Int. Workshop Inf. Hiding*. Berlin, Germany: Springer, 2008, pp. 251–267.
[2] J. Fridrich, *Steganography in Digital Media: Principles, Algorithms, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[3] J. Fridrich and T. Filler, "Practical methods for minimizing embedding impact in steganography," *Proc. SPIE*, vol. 6505, p. 650502, Feb. 2007.
[4] 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.
[5] T. Filler and J. Fridrich, "Gibbs construction in steganography," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 4, pp. 705–720, Dec. 2010.
[6] T. Pevný, P. Bas, and J. Fridrich, "Steganalysis by subtractive pixel adjacency matrix," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 215–224, Jun. 2010.
[7] V. Holub and J. Fridrich, "Designing steganographic distortion using directional filters," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Dec. 2012, pp. 234–239.
[8] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP J. Inf. Secur.*, vol. 2014, no. 1, pp. 1–13, 2014.
[9] B. Li, M. Wang, J. Huang, and X. Li, "A new cost function for spatial image steganography," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 4206–4210.
[10] V. Sedighi, R. Cogranne, and J. Fridrich, "Content-adaptive steganography by minimizing statistical detectability," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 2, pp. 221–234, Feb. 2016.
[11] D. Upham. *Steganographic Algorithm JSteg*. [Online]. Available: https://zooid.org/~paul/crypto/jsteg/
[12] A. Westfeld, "F5—A steganographic algorithm," in *Proc. 4th Int. Workshop Inf. Hiding*, in Lecture Notes in Computer Science, vol. 2137. Springer-Verlag, 2001, pp. 289–302.
[13] 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/
[14] Y. Kim, Z. Duric, and D. Richards, "Modified matrix encoding technique for minimal distortion steganography," in *Proc. Int. Workshop Inf. Hiding*. Springer, 2006, pp. 314–327.
[15] T. Filler and J. Fridrich, "Design of adaptive steganographic schemes for digital images," *Proc. SPIE* , vol. 7880, p. 78800F, Feb. 2011.
[16] 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.
[17] L. Guo, J. Ni, W. Su, C. Tang, and Y.-Q. Shi, "Using statistical image model for JPEG steganography: Uniform embedding revisited," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 12, pp. 2669–2680, Dec. 2015.
[18] Y. Pan, J. Ni, and W. Su, "Improved uniform embedding for efficient JPEG steganography," in *Proc. Int. Conf. Cloud Comput. Secur.* Cham, Switzerland: Springer, 2016, pp. 125–133.
[19] K. Chen, W. Zhang, H. Zhou, N. Yu, and G. Feng, "Defining cost functions for adaptive steganography at the microscale," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Dec. 2016, pp. 1–6.
[20] B. Li, S. Tan, M. Wang, and J. Huang, "Investigation on cost assignment in spatial image steganography," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 8, pp. 1264–1277, Aug. 2014.
[21] V. Holub and J. Fridrich, "Low-complexity features for JPEG steganalysis using undecimated DCT," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 2, pp. 219–228, Feb. 2015.
[22] X. Song, F. Liu, C. Yang, X. Luo, and Y. Zhang, "Steganalysis of adaptive JPEG steganography using 2D Gabor filters," in *Proc. 3rd ACM Workshop Inf. Hiding Multimedia Secur.*, 2015, pp. 15–23.
[23] J. Kodovský and J. Fridrich, "Steganalysis of JPEG images using rich models," in *Proc. SPIE, Media Watermarking, Secur., Forensics, Int. Soc. Opt. Photon.*, vol. 8303, 2012, p. 83030A.
[24] P. Bas, T. Filler, and T. Pevný, "'Break our steganographic system': The ins and outs of organizing BOSS," in *Proc. Int. Workshop Inf. Hiding*. Berlin, Germany: Springer, 2011, pp. 59–70.
[25] P. Bas and T. Furon. (Jul. 2007). *BOWS-2(Break Our Watermarking System)*. [Online]. Available: http://bows2.ec-lille.fr/
[26] T. Denemark and J. Fridrich, "Model based steganography with precover," *Electron. Imag.*, vol. 2017, no. 7, pp. 56–66, 2017.
[27] A. Polesel, G. Ramponi, and V. J. Mathews, "Image enhancement via adaptive unsharp masking," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 505–510, Mar. 2000.
[28] R. Kresch and N. Merhav, "Fast DCT domain filtering using the DCT and the DST," *IEEE Trans. Image Process.*, vol. 8, no. 6, pp. 821–833, Jun. 1999.
[29] V. Kapinaiah, J. Mukherjee, and P. K. Biswas, "Block DCT to wavelet transcoding in transform domain," *Signal, Image Video Process.*, vol. 6, no. 2, pp. 179–195, 2012.
[30] T. D. Denemark, M. Boroumand, and J. Fridrich, "Steganalysis features for content-adaptive JPEG steganography," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 8, pp. 1736–1746, Aug. 2016.
[31] J. Fridrich and J. Kodovský, "Rich models for steganalysis of digital images," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 868–882, Jun. 2012.
[32] N. Xie, H. Ling, W. Hu, and X. Zhang, "Use bin-ratio information for category and scene classification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2313–2319.
[33] T. G. Dietterich, "Approximate statistical tests for comparing supervised classification learning algorithms," *Neural Comput.*, vol. 10, no. 7, pp. 1895–1923, 1998.
[34] W. Zhou, W. Zhang, and N. Yu, "A new rule for cost reassignment in adaptive steganography," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 11, pp. 2654–2667, Nov. 2017.
[35] S. A. Cook and S. O. Aanderaa, "On the minimum computation time of functions," *Trans. Amer. Math. Soc.*, vol. 142, pp. 291–314, Aug. 1969.
[36] N. IkCho and S. UkLee, "Fast algorithm and implementation of 2-D discrete cosine transform," *IEEE Trans. Circuits Syst.*, vol. 38, no. 3, pp. 297–305, Mar. 1991.
[37] J.-F. Couchot, R. Couturier, and C. Guyeux, "STABYLO: Steganography with adaptive, Bbs, and binary embedding at low cost," *Ann. Telecommun.-Ann. Télécommun.*, vol. 70, nos. 9–10, pp. 441–449, 2015.



Kejiang Chen received the B.S. degree from the School of Communication and Information Engineering, Shanghai University, in 2015. He is currently pursuing the Ph.D. degree in information security with the University of Science and Technology of China. His research interests include information hiding, image processing, and deep learning.



Hang Zhou received the B.S. degree from the School of Communication and Information Engineering, Shanghai University, in 2015. He is currently pursuing the Ph.D. degree in information security with the University of Science and Technology of China. His research interests include information hiding, image processing, and computer graphics.

Wenbo Zhou received the B.S. degree from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2014. He is currently pursuing the Ph.D. degree with the University of Science and Technology of China. His research interests include steganography, steganalysis, and multimedia security.

Weiming Zhang received the M.S. and Ph.D. degrees from the Zhengzhou Information Science and Technology Institute, Zhengzhou, China, in 2002 and 2005, respectively. He is currently a Professor with the School of Information Science and Technology, University of Science and Technology of China. His research interests include multimedia security, information hiding, and privacy protection.

Nenghai Yu received the B.S. degree from the Nanjing University of Posts and Telecommunications in 1987, the M.E. degree from Tsinghua University in 1992, and the Ph.D. degree from the University of Science and Technology of China in 2004. He is currently a Professor with the University of Science and Technology of China. His research interests include multimedia security, multimedia information retrieval, video processing, information hiding and security, and privacy and reliability in cloud computing.