This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2019.2925118, IEEE Transactions on Circuits and Systems for Video Technology

JPEG Steganography with Estimated Side-information

Weixiang Li, Kejiang Chen, Weiming Zhang, Hang Zhou, Yaofei Wang, Nenghai Yu

***Abstract*—Previous studies have exhibited that incorporating side-information, e.g., a high-quality "precover" image, can significantly improve steganographic security for JPEG images. This motivates us to estimate the side-information for traditional steganographic scenario in which only a JPEG image is available. It is expected to achieve high-level security by utilizing the estimated side-information similar to side-informed steganography, even though the estimated side-information is not perfectly precise. In this paper, a general framework of side-information estimated (SIE) JPEG steganography is proposed, under which the core problems are how to better estimate the precover and modulate the distortion function correspondingly. To address the two problems, we test several denoising filters and a deblocking filter to obtain the estimated precover, and we introduce two implementation models for modulating the costs. We finally recommend the combination of the deblocking filter and the modulation model using the polarity of the estimated rounding error. The experimental results show that the proposed method dramatically improves the existing additive distortions for images of an arbitrary quality factor and outperforms state-of-the-art methods based on estimating side-information when resisting modern steganalysis.**

***Index Terms*—Steganography, JPEG images, minimal distortion, side-information, denoising, deblocking.**

I. Introduction

Modern steganography is a science and art of covert communication that slightly modifies a digital cover object to transmit a covert message without drawing suspicions from steganalysis [1]. Since JPEG images are the widely adopted format for image storage and transmission, steganography on JPEG images has become a research hotspot over the past few years. Based on the minimal distortion model [2], various content-adaptive distortion functions [3]–[6] are designed for JPEG steganography by preferably exploiting image texture complexity to strengthen the steganographic security. Meanwhile, microscale steganography and the cost spreading rule [7], the JPEG controversial-pixel-prior rule [8] and the block-boundary-continuity principle [9] are extended from spatial image steganography to help improve the performance of the above additive distortion functions.

This work was supported in part by the Natural Science Foundation of China under Grant U1636201 and 61572452, by Anhui Initiative in Quantum Information Technologies under Grant AHY150400, and by the Fundamental Research Funds for the Central Universities under Grant WK6030000135 and WK6030000136.

The authors are with CAS Key Laboratory of Electro-magnetic Space Information, University of Science and Technology of China, Hefei 230026, China. Corresponding author: Weiming Zhang (Email: zhangwm@ustc.edu.cn).

Denote x as the cover element and $\mathcal{I}$ as the range of the embedding operation at x. In the context of digital image steganography, ternary embedding (±1 embedding with $\mathcal{I} = \{x-1, x, x+1\}$) is more commonly used than binary embedding (with $\mathcal{I} = \{x, \bar{x}\}$ where $\bar{x}$ is x flipping its Least Significant Bit (LSB)), since it can achieve a smaller embedding impact. In ±1 embedding, the costs of changing the quantized JPEG coefficient by $+1$ and -1 are equivalent [3]–[8]. It is widely recognized that incorporating side-information at the sender can significantly improve steganographic security in practice, where the costs of $+1$ and -1 are not the same but modulated by some additional information. For JPEG steganography, the side-information may be in the form of an uncompressed image (called the precover [10]) or concretely as the non-rounded DCT coefficient, which partially compensates for the lack of knowledge of the cover model when it is highly non-stationary. Numerous heuristic cost-modulated schemes were introduced in [3], [7], [11]–[14]. Side-informed (SI) steganography [13] allowed a ternary embedding operation rather than a binary approach and computed the costs from the uncompressed cover, both of which appeared to improve the empirical security of existing distortion functions by a rather large margin. Under the condition of the sender lacking access to a precover, [14] used a set of multiple JPEG images of the same scene to modulate the costs, achieving a high-level steganographic security. By utilizing the additional information that is unavailable to the Warden, the hope is that the embedding will disturb the statistical properties of the cover source less.

In the traditional and common steganographic scenario (non-SI steganography [3]–[9]), the sender has no access to a precover but experiences only one JPEG image for message embedding. The high-level security of SI steganography benefits from the premise that the sender has the precover [13] or multiple images of the same scene [14]. Motivated by SI steganography, relatively high security of non-SI steganography can be expectedly achieved if we can estimate the precover from the JPEG cover as precisely as possible. Recently, [15], [16] attempted to estimate the side-information with the average and the Wiener filter, respectively, for improving the distortion functions. However, the filters for estimating the precover seem to be primitive, and the method of minimizing the distance of steganalytic feature space leads to excessive time consumption for message embedding. Therefore, a universal and efficient framework of estimating side-information for JPEG steganography needs to be established.

In this paper, we focus on the traditional non-SI steganographic scenario and propose a general framework of side-information estimated (SIE) JPEG steganography, with its

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2019.2925118, IEEE Transactions on Circuits and Systems for Video Technology

several implementation methods for improving the existing distortion functions. Under the SIE framework, the core problems for strengthening the steganographic security include estimating the precover as precisely as possible and modulating the distortion according to the estimated non-rounded coefficient. Based on the simple SIE model that only considers the polarity of the rounding error, we examine the performance of various denoising and deblocking filters on estimating the precover. The experimental results show that the proposed SIE model equipped with the selected deblocking filter can significantly improve the empirical security of existing JPEG additive distortion functions. The proposed method also outperforms state-of-the-art methods using estimated side-information.

The rest of this paper is organized as follows. In Section II, we briefly review side-informed steganography and introduce its degradation model. The general framework of estimating side-information for JPEG steganography is proposed with several implementations in Section III. The experimental results and comparisons are presented in Section IV, and the paper is concluded in Section V.

II. JPEG Steganography with Precover

With the helpful additional information of precover [13] or the same scene-based multiple JPEG images [14] that are used to modulate the costs, side-informed (SI)-based JPEG steganography significantly improved the steganographic security of the existing additive distortion functions. In [13], a ± 1 embedding version of SI-based steganography was studied using the non-rounded DCT coefficients after compressing the precover.

Denote $\mathbf{x} = (x_1, x_2, \cdots, x_n)$, $\mathbf{u} = (u_1, u_2, \cdots, u_n)$ and $\mathbf{y} = (y_1, y_2, \cdots, y_n)$ as the cover (quantized DCT coefficients for embedding), the non-rounded coefficients of the precover, and the stego, respectively. The rounding error $e_i = u_i - x_i$ ($|e_i| \leq 0.5$, $1 \leq i \leq n$) is used to adjust the original cost $\rho_i^{(\mathcal{A})}$ of changing x_i by ± 1 with

$$\begin{cases} \rho_i^{(\mathrm{SI})+} = (1 - 2|e_i|)\rho_i^{(\mathcal{A})} & \text{if } y_i = x_i + \mathrm{sign}(e_i) \\ \rho_i^{(\mathrm{SI})-} = \rho_i^{(\mathcal{A})} & \text{if } y_i = x_i - \mathrm{sign}(e_i) \end{cases}, \tag{1}$$

where $\rho_i^{(\mathcal{A})}$ can be defined by any distortion function $\mathcal{A}$ [3]–[8]. It is recommended in [13] to compute the costs from the precover $\mathbf{u}$ instead of the cover $\mathbf{x}$. Intuitively, the modulated costs (1) not only reflect the local image complexity but also account for the distortion w.r.t. the precover. With the near-optimal STCs [2], the actual embedding will approach the minimal average distortion $E_{\boldsymbol{\pi}}(D) = \sum_{i=1}^{n}(\pi_i^+ \rho_i^{(\mathrm{SI})+} + \pi_i^- \rho_i^{(\mathrm{SI})-})$ by modifying x_i by $\pm\mathrm{sign}(e_i)$ with probability

$$\pi_i^{(\pm)} = \frac{\exp(-\lambda \rho_i^{(\mathrm{SI})\pm})}{1 + \exp(-\lambda \rho_i^{(\mathrm{SI})+}) + \exp(-\lambda \rho_i^{(\mathrm{SI})-})} \tag{2}$$

with λ ($\lambda > 0$) determined by the message length of m bits

$$\begin{aligned} m = -\sum_{i=1}^{n} \big(& \pi_i^+ \log_2 \pi_i^+ + \pi_i^- \log_2 \pi_i^- \\ & + (1 - \pi_i^+ - \pi_i^-) \log_2 (1 - \pi_i^+ - \pi_i^-) \big). \end{aligned} \tag{3}$$

We mark the specific SI-based method as SI-$\mathcal{A}(*)$ where $* \in \{\mathbf{u}, \mathbf{x}\}$ represents the cost computation from the precover $\mathbf{u}$ or the cover $\mathbf{x}$.

Here, we introduce a degradation model of the SI steganography that only focuses on the polarity of the rounding error. In the SI-polarity (SIp) model, the cost modulation neglects the amplitude of the rounding error by adjusting each cost with the same parameter α ($0 \leq \alpha \leq 1$), that is,

$$\begin{cases} \rho_i^{(\mathrm{SIp})+} = \alpha \cdot \rho_i^{(\mathcal{A})} & \text{if } y_i = x_i + \mathrm{sign}(e_i) \\ \rho_i^{(\mathrm{SIp})-} = \rho_i^{(\mathcal{A})} & \text{if } y_i = x_i - \mathrm{sign}(e_i) \end{cases}. \tag{4}$$

Correspondingly, we mark the specific SIp-based method as SIp-$\mathcal{A}(*)$. The SI and SIp models introduced here inspire the similar models of the proposed framework of estimating side-information for JPEG steganography as described below.

III. Framework of Estimating Side-information for JPEG Steganography and its Implementations

Under the traditional and common steganographic scenario, the sender possesses only a JPEG image without any side-information, i.e., the sender has no access to a precover. Inspired by the SI steganography, a clever sender may estimate some sufficiently accurate side-information from the JPEG image to aid the steganography, i.e., he or she may be able to modulate the costs with the help of the estimated side-information.

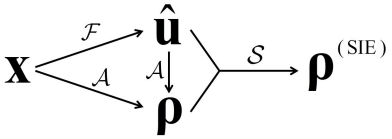


Fig. 1: General framework of side-information estimated (SIE) JPEG steganography.

In this section, we propose a general framework of side-information estimated (SIE) JPEG steganography as depicted in Fig. 1. The SIE framework is formulated as

$$\mathcal{S}(\mathcal{F})\text{-}\mathcal{A}(*).$$

In Fig. 1, the JPEG image $\mathbf{x}$ is used to estimate the precover $\hat{\mathbf{u}}$ (the estimated non-rounded coefficient) with a filter $\mathcal{F}$, and the original cost $\boldsymbol{\rho}$ can be defined by an arbitrary distortion function $\mathcal{A}(*)$ with $* \in \{\mathbf{x}, \hat{\mathbf{u}}\}$ representing the cost computation of $\mathbf{x}$ or $\hat{\mathbf{u}}$. Moreover, in terms of the update strategy $\mathcal{S}$, the SIE-based cost $\boldsymbol{\rho}^{(\mathrm{SIE})}$ is obtained by modulating $\boldsymbol{\rho}$ with $\hat{\mathbf{u}}$. Obviously, under the SIE framework, the two core issues for improving the steganographic security are the following: 1) how to precisely estimate $\hat{\mathbf{u}}$, i.e., choose better $\mathcal{F}$, and 2) how to adjust the costs according to $\hat{\mathbf{u}}$, i.e., design better $\mathcal{S}$. If the estimated $\hat{\mathbf{u}}$ is sufficiently close to the precover $\mathbf{u}$, we can simply refer to the SI method (1). However, it is of practical significance to design an appropriate update strategy $\mathcal{S}$ when $\hat{\mathbf{u}}$ cannot be perfect enough.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2019.2925118, IEEE Transactions on Circuits and Systems for Video Technology

A. Estimating the precover with various filters

The first critical problem under the SIE framework, that is, how to better estimate the precover, can be regarded as the image restoration problem. It is known that the information loss for JPEG compression takes place in the stage of quantization, leading to round-off errors in each block, which inevitably produces blocking artifacts. To suppress blocking artifacts and obtain a high-quality estimated precover, several denoising and deblocking filters are studied in this paper.

1) Denoising filters: The denoising filters are applied in the spatial domain, which is the decompression image of the JPEG image. We introduce the average filter, median filter, Gaussian filter and Wiener filter, which are abbreviated as Avg, Med, Gau and Wie, respectively, and each has window sizes of 3×3 and 5×5. The output of the average filter is simply the average value of pixels contained in the neighborhood of the filter mask. In the median filter, the current pixel is replaced with the median value among its neighboring pixels. The Gaussian filter is a non-uniform low-pass filter, which modifies the input signal by convolution with a Gaussian function. The Wiener filter is optimal in terms of the mean square error by using a pixelwise adaptive Wiener filtering method based on statistics estimated from a local neighborhood of each pixel. Specifically, the Gaussian and Wiener filtering outputs are obtained using `fspecial('gaussian')` and `wiener2` in MATLAB, respectively.

2) Deblocking filter: We also select a type of image-restoration-based deblocking filter, called $SSRQC$ [17], to obtain a better estimated precover. Image deblocking is usually formulated as an ill-posed image inverse problem by exploiting the information in the JPEG compressed bit-stream, such as the decompressed image and the quantization matrix. $SSRQC$ [17] was proposed for image deblocking by using a structural sparse representation (SSR) prior and a quantization constraint (QC) prior with a new split Bregman iteration-based method, which greatly improved the existing image-deblocking quality.

From the above introduction, the candidate filter $\mathcal{F} \in \{Avg(w), Med(w), Gau(w), Wie(w), SSRQC\}$ with window size $w \in \{3 \times 3, 5 \times 5\}$ will be investigated in this paper.

B. Modulating the costs with two implementation models

To further enrich the SIE framework, we provide here two heuristic update strategies that are inspired by (1) and (4). Denote $\hat{e}_i = \hat{u}_i - x_i$ as the estimated rounding error. The first strategy, which we name SIEg model, utilizes $\hat{e}_i$ to modulate the original cost $\rho_i^{(\mathcal{A})}$ by $g(\hat{e}_i)$, that is,

$$\begin{cases} \rho_i^{(\text{SIE}g)+} = g(\hat{e}_i) \cdot \rho_i^{(\mathcal{A})} & \text{if } y_i = x_i + \text{sign}(\hat{e}_i) \\ \rho_i^{(\text{SIE}g)-} = \rho_i^{(\mathcal{A})} & \text{if } y_i = x_i - \text{sign}(\hat{e}_i) \end{cases}. \tag{5}$$

Note that the value range of $\hat{e}_i$ is erratic (not in the value range $[0, 0.5]$ as SI) and determined by the precision of $\hat{\mathbf{u}}$. Therefore, we carefully design a heuristic form of $g(\cdot)$,

$$g(\hat{e}_i) = \begin{cases} 1 - 2|\hat{e}_i| & \text{if } |\hat{e}_i| \le 0.5 \\ \beta & \text{otherwise} \end{cases}, \tag{6}$$

where β ($0 \le \beta \le 1$) is to ensure that the cost corresponding to $|\hat{e}_i| > 0.5$ remains non-negative when using the modulation according to (1).

Since it is easier to estimate the accurate polarity of the rounding error than the amplitude of the rounding error, we introduce a compromising strategy inspired by the SIp model (4). The second strategy, called the SIE-polarity (SIEp) model, solely considers the polarity of $\hat{e}_i$ by assigning the same parameter γ ($0 \le \gamma \le 1$) for modulating the cost of $\pm\text{sign}(\hat{e}_i)$,

$$\begin{cases} \rho_i^{(\text{SIEp})+} = \gamma \cdot \rho_i^{(\mathcal{A})} & \text{if } y_i = x_i + \text{sign}(\hat{e}_i) \\ \rho_i^{(\text{SIEp})-} = \rho_i^{(\mathcal{A})} & \text{if } y_i = x_i - \text{sign}(\hat{e}_i) \end{cases}. \tag{7}$$

Therefore, the cost update strategy $\mathcal{S} \in \{\text{SIE}g, \text{SIEp}\}$ will be investigated in this paper.

C. Generalizing other related methods into the SIE framework

The methods in [15], [16] also attempted to utilize the estimated side-information for embedding, which can be considered as two implementation instances of the proposed SIE framework. Specifically, the method in [15] obtained the estimated precover by replacing the decompressed boundary pixels with the 3×3 average filtered boundary pixels (i.e., another $\mathcal{F}$) and adjusted the costs by the estimated quantized DCT coefficients instead of the estimated non-rounded coefficients (i.e., another $\mathcal{S}$). Similarly, the method in [16] used the 3×3 Wiener filter to estimate the precover and searched the best parameter for cost modulation by minimizing the steganalytic feature distance (i.e., another $\mathcal{S}$), which is time-consuming with 20-message embedding and 20-feature extraction. Comparatively, the proposed general SIE framework enables the combination of various $\mathcal{S}$, $\mathcal{F}$ and $\mathcal{A}$ for designing a better steganographic scheme.

IV. Experimental Results and Analysis

In this section, we will demonstrate the performance of different SIE-based methods and the comparison with the methods in [15], [16]. The experiments are mainly conducted on BOSSBase 1.01 [18], which contains 10,000 gray-scale images of size 512×512 pixels. All of the images are compressed into the JPEG domain with quality factors QF=50, 75 and 95, which are adopted as datasets for experimental comparisons. To verify the generalization of the proposed method, we perform experiments on another popular image set BOWS2 [19], which also contains 10,000 gray-scale images of size 512×512 and is JPEG compressed by QF=75. We use the mainstream distortion functions UERD [4] and J-UNIWARD [3] with the optimal simulator [20] for message embedding, and the steganalyzer is trained by using state-of-the-art DCTR-8,000D [21] and GFR-17,000D [22] with the FLD ensemble [23] by default. The FLD ensemble can minimize the total classification error probability under equal priors $P_{\text{E}} = \min_{P_{\text{FA}}} \frac{1}{2}(P_{\text{FA}} + P_{\text{MD}})$ where P_{FA} and P_{MD} are the false-alarm probability and the missed-detection probability, respectively. The ultimate security is qualified by the average error rate $\overline{P_{\text{E}}}$ averaged over 10 random 5000/5000 splits of the dataset, and larger $\overline{P_{\text{E}}}$ means stronger security.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2019.2925118, IEEE Transactions on Circuits and Systems for Video Technology

TABLE I: Performance of filter $\mathcal{F}$ on estimating the precover.

Filter $\mathcal{F}$	$Avg(3\times3)$	$Med(3\times3)$	$Gau(3\times3)$	$Wie(3\times3)$	$SSRQC$
PSNR (dB)	33.36	34.25	34.99	37.12	38.92
R_p (%)	55.79	55.03	56.02	56.09	57.39
Filter $\mathcal{F}$	$Avg(5\times5)$	$Med(5\times5)$	$Gau(5\times5)$	$Wie(5\times5)$	
PSNR (dB)	29.89	30.79	31.65	34.62	
R_p (%)	53.19	53.98	54.46	54.08	

TABLE II: Detection errors $\overline{P_{\mathrm{E}}}$ (%) of SIEp($\mathcal{F}$)-UERD($\mathbf{x}$) w.r.t. γ in (7) at payload 0.3bpnzac on BOSSBase of QF=75 against DCTR-8,000D feature.

Filter $\mathcal{F}$	$Avg(3\times3)$	$Med(3\times3)$	$Gau(3\times3)$	$Wie(3\times3)$	$SSRQC$
$\gamma=0.55$	27.77	30.75	27.85	32.74	34.85
$\gamma=0.60$	29.56	31.88	30.09	33.69	36.28
$\gamma=0.65$	30.50	**31.90**	**31.59**	**33.92**	**37.42**
$\gamma=0.70$	**31.46**	31.81	31.43	33.63	37.06
$\gamma=0.75$	31.32	31.15	31.37	32.30	34.91
Filter $\mathcal{F}$	$Avg(5\times5)$	$Med(5\times5)$	$Gau(5\times5)$	$Wie(5\times5)$	
$\gamma=0.55$	26.09	29.04	27.20	31.17	
$\gamma=0.60$	27.08	**29.62**	28.01	**31.55**	
$\gamma=0.65$	28.07	29.39	29.12	31.14	
$\gamma=0.70$	**28.33**	29.14	**29.46**	30.31	
$\gamma=0.75$	28.05	28.67	29.15	29.50	

A. Comparison of different SIE-based methods

1) Performance of different $\mathcal{F}$ on estimating the precover: The goal of employing $\mathcal{F}$ on the JPEG image is to obtain an estimated precover as close as possible to the real precover. Intuitively, we can evaluate the effect of different $\mathcal{F}$ via the peak signal-to-noise ratio (PSNR) between the real precover and the estimated precover, where $\mathcal{F}$ with larger PSNR may have a better ability to estimate the precover. In addition, we measure the ratio of correct polarities of the estimated rounding errors, i.e., $R_p = \left(\sum_{i=1}^{n}[\mathrm{sign}(\hat{e}_i) = \mathrm{sign}(e_i)]\right)/n$ where the Iverson bracket $[I]$ is defined to be 1 if the logical expression I is true and 0 otherwise. Because the polarity of the rounding error, which directs the DCT coefficient to be rounded to "the other side" [13], is significantly important for the steganographic security, $\mathcal{F}$ of higher R_p with more correct polarities may also correspond to stronger security.

We randomly select 1,000 JPEG images from BOSSBase of QF=75 and perform different $\mathcal{F}$ on these images to observe the average PSNR and R_p. As shown in TABLE I, each denoising filter of the window size 3×3 has a larger PSNR and R_p than that of the 5×5 window, which implies that the filter size should be sufficiently small because of the strong correlation among neighboring pixels. Among denoising filters, the 3×3 Wiener filter achieves the best performance. However, the deblocking filter $SSRQC$ is even better than the 3×3 Wiener filter. Therefore, we believe that the deblocking filter designed for technically removing blocking artifacts and restoring the image is more suitable for estimating the precover.

Since the SIEp model is easier to be implemented, we first investigate the impact of different $\mathcal{F}$ on improving the steganographic security when combined with the SIEp model. As shown in TABLE II, the security w.r.t. $\mathcal{F}$ has a consistent trend that is similar to the PSNR and R_p in TABLE I. The 3×3 Wiener filter is the best denoising filter, and it is still worse than the deblocking filter $SSRQC$. Therefore, we select the best filter $SSRQC$ while assigning the optimal $\gamma = 0.65$ to SIEp in (7) for the following experiments.

TABLE III: Detection errors $\overline{P_{\mathrm{E}}}$ (%) of SI-based and SIE-based methods using UERD and J-UNIWARD at 0.3bpnzac on BOSSBase of QF=75 against DCTR-8,000D feature.

$\mathcal{A}$	Scheme	$*=\mathbf{x}$	$*=\mathbf{u}/\hat{\mathbf{u}}$
UERD	$\mathcal{A}(*)$	23.06	/
	SI-$\mathcal{A}(*)$	45.07	46.25 ($\mathbf{u}$)
	SIp-$\mathcal{A}(*)$	32.27	34.12 ($\mathbf{u}$)
	SIE$g(SSRQC)$-$\mathcal{A}(*)$	17.50	16.54 ($\hat{\mathbf{u}}$)
	SIEp$(SSRQC)$-$\mathcal{A}(*)$	37.42	37.52 ($\hat{\mathbf{u}}$)
J-UNIWARD	$\mathcal{A}(*)$	24.35	/
	SI-$\mathcal{A}(*)$	44.82	45.41 ($\mathbf{u}$)
	SIp-$\mathcal{A}(*)$	31.60	30.61 ($\mathbf{u}$)
	SIE$g(SSRQC)$-$\mathcal{A}(*)$	17.54	17.67 ($\hat{\mathbf{u}}$)
	SIEp$(SSRQC)$-$\mathcal{A}(*)$	**40.45**	**37.97** ($\hat{\mathbf{u}}$)

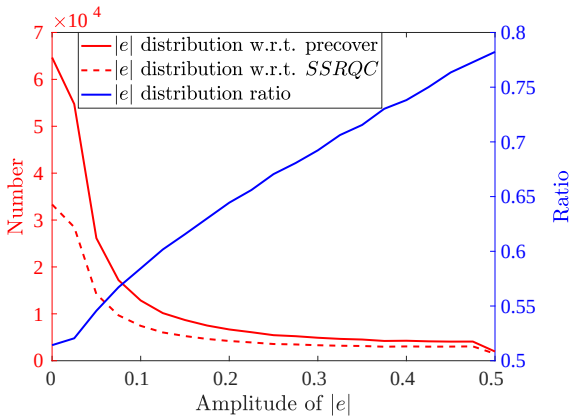


Fig. 2: Distributions of the rounding errors $|e|$ w.r.t the precover and the $SSRQC$-estimated precover, averaged over 1,000 images randomly selected from BOSSBase of QF=75. The curve of $|e|$ distribution w.r.t. precover means the distribution of all $|e|$s of the real precover, while the curve of $|e|$ distribution w.r.t. $SSRQC$ represents the distribution of the selected part of $|e|$s whose polarities are correctly estimated by $SSRQC$. More intuitively, the curve of $|e|$ distribution ratio depicts the ratio of the $|e|$ distribution w.r.t. $SSRQC$ to the $|e|$ distribution w.r.t. precover.

2) Investigation on SI-based and SIE-based models: We now study the gap between the SI-based and SIE-based methods as well as the impact of cost computation of $\mathbf{x}$ or $\mathbf{u}/\hat{\mathbf{u}}$. Similar to the choice of γ as shown in TABLE II, the optimal parameters in TABLE III, such that $\alpha = 0.4$ for SIp in (4) and $\beta = 0.65$ for SIEg in (6), are determined at 0.3bpnzac UERD on BOSSBase of QF=75 against DCTR by traversal search with a step of 0.05. According to the experimental results, the selected optimal parameters ($\alpha = 0.4$, $\beta = 0.65$, $\gamma = 0.65$) are applicable to other distortion functions, relative payloads, quality factors and image sources.

The security of SIp-$\mathcal{A}(*)$ is not as superior as that of SI-$\mathcal{A}(*)$, but it is still far better than the non-SI $\mathcal{A}(*)$ even though it only utilizes the direction of the rounding error. SIEp$(SSRQC)$-$\mathcal{A}(*)$ is naturally worse than SI-$\mathcal{A}(*)$ (the upper bound with the real precover), and the security of

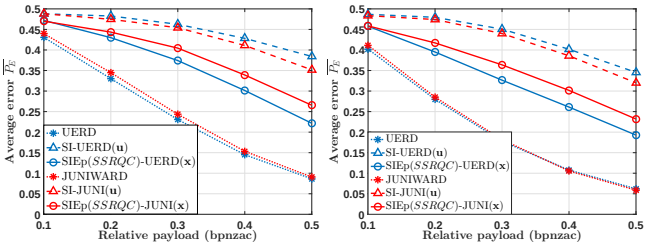
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2019.2925118, IEEE Transactions on Circuits and Systems for Video Technology

SIE$g(SSRQC)$-$\mathcal{A}(*)$ is terrible, given the fact that estimating the precise amplitude of the rounding error is more difficult. What is beyond our expectation is that SIEp$(SSRQC)$-$\mathcal{A}(*)$ outperforms SIp-$\mathcal{A}(*)$. Since SIp and SIEp are focusing on the polarity of the rounding error, SIp should be the ideal bound that SIEp can approach with more correct polarities. Interestingly, this phenomenon implies that simply pursuing larger R_p may not be the best choice for improving the SIEp method. We attempt to explore and explain this phenomenon via the distribution of the rounding errors whose polarities are correctly estimated. Suppose a DCT coefficient x_1 with the cost ρ_1 and the rounding error $|e_1| = 0.4$ and another DCT coefficient x_2 with the same cost $\rho_2 = \rho_1$ and the rounding error $|e_2| = 0.2$. According to SI (1), the modulated cost of x_1 is smaller than that of x_2 (i.e., x_1 is more suitable for modification), which distinguishes the modification priorities of these two coefficients with different rounding errors and thus leads to the high-level security of SI. However, in view of SIp (4), these two coefficients still experience the same modulated costs regardless of the amplitudes of their rounding errors. Since only a part (w.r.t. R_p) of coefficients whose polarities of the rounding errors are correctly estimated (i.e., $\text{sign}(\hat{e}_i) = \text{sign}(e_i)$) can be selected for cost modulation, the modification priorities of the coefficients with different $|e|$s will be reasonably reflected if more coefficients with larger $|e|$s are selected for cost modulation. As verified in Fig. 2, with increasing the amplitude of $|e|$, the ratio of the selected $|e|$ w.r.t. $SSRQC$ increases. In this way, it is very likely that x_1 with $|e_1| = 0.4$ is selected by SIEp and x_2 with $|e_2| = 0.2$ is not, such that the modification priorities of them are reflected similar to the optimal SI. Overall, unlike SIp that gives the same priorities to the coefficients with different $|e|$s, SIEp will focus more on the coefficients with larger $|e|$s that are more suitable for modification, making the modification priorities of the coefficients with different $|e|$s more reasonable. And we believe that this contributes to the secure advantage of SIEp to SIp.

As pointed out in [13], the costs of SI-$\mathcal{A}(*)$ computed from the real precover **u** achieve better security than that from the JPEG cover **x** because more information about the precover source is lost due to JPEG compression. However, since $\hat{\mathbf{u}}$ does not perfectly approach the real precover, the costs of $\hat{\mathbf{u}}$ do not show superior performance for the SIE-based methods when compared with the costs of **x**. Instead, the bold data in TABLE III demonstrate the advantage of calculating the costs of **x** for the SIEp method. Therefore, SIEp$(SSRQC)$-$\mathcal{A}(\mathbf{x})$ is the recommended implementation method of the proposed SIE framework for the following experiments.

B. Universality verification of the proposed SIE method

We test the universality of the proposed SIEp$(SSRQC)$-$\mathcal{A}(\mathbf{x})$ via using the mainstream distortion functions UERD [4] and J-UNIWARD [3] with relative embedding payloads $\{0.1, 0.2, 0.3, 0.4, 0.5\}$ bpnzac (bit per nonzero AC coefficient), on two image sets of different sources and quality factors in resisting the detections of state-of-the-art steganalytic features DCTR-8,000D [21] and GFR-17,000D [22]. Fig. 3 demonstrates that the proposed method can well approach the SI method and improve the traditional non-SI distortion functions by a large margin. TABLE IV verifies that the proposed method can be applied to image sets of relatively small or large quality factors and another image source.



(a) DCTR (b) GFR

Fig. 3: Performance of the proposed method SIEp$(SSRQC)$-$\mathcal{A}(\mathbf{x})$ using UERD and J-UNIWARD on BOSSBase of QF=75 against DCTR-8,000D and GFR-17,000D features.

TABLE IV: Detection errors $\overline{P_E}$ (%) of the proposed SIEp$(SSRQC)$-$\mathcal{A}(\mathbf{x})$ using UERD/J-UNIWARD at 0.3bpnzac on other image sets of different sources and quality factors.

Set-QF	Scheme	DCTR	GFR
BOSS-QF50	UERD	18.19	13.79
	SIEp$(SSRQC)$-UERD$(\mathbf{x})$	35.85	30.31
	J-UNIWARD	18.51	13.10
	SIEp$(SSRQC)$-JUNI$(\mathbf{x})$	37.65	32.78
BOSS-QF95	UERD	37.20	32.14
	SIEp$(SSRQC)$-UERD$(\mathbf{x})$	44.63	42.84
	J-UNIWARD	40.36	35.55
	SIEp$(SSRQC)$-JUNI$(\mathbf{x})$	45.54	44.77
BOWS2-QF75	UERD	23.93	17.33
	SIEp$(SSRQC)$-UERD$(\mathbf{x})$	37.23	33.08
	J-UNIWARD	26.62	18.77
	SIEp$(SSRQC)$-JUNI$(\mathbf{x})$	39.65	35.15

C. Comparison with other related methods

As mentioned before, the methods in [15], [16] can be generalized as two implementation instances of the proposed SIE framework. As shown in TABLE V, SIEp$\big(Avg(3\times3)\big)$-UERD$(\mathbf{x})$ using the same filter outperforms the method in

TABLE V: Detection errors $\overline{P_E}$ (%) and execution time (seconds) of different methods using estimated side-information at 0.3bpnzac on BOSSBase of QF=75. The overall execution time is obtained by MATLAB R2015b on an Intel(R) Core(TM) i5-4590 CPU @ 3.30GHz when using STCs with $h = 10$ [2].

Scheme	DCTR	GFR	Time
UERD	23.06	17.81	0.77
Method in [15] upon UERD	28.87	21.61	0.79
Method in [16] upon UERD	34.79	28.28	36.69
SIEp$\big(Avg(3\times3)\big)$-UERD$(\mathbf{x})$	30.50	25.35	0.78
SIEp$\big(Wie(3\times3)\big)$-UERD$(\mathbf{x})$	33.92	27.57	0.80
SIEp$(SSRQC)$-UERD$(\mathbf{x})$	37.42	33.05	14.15
SIE$(SSRQC)$-UERD$(\mathbf{x})$ + MinDistFea	37.88	34.14	50.11

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2019.2925118, IEEE Transactions on Circuits and Systems for Video Technology

[15], which indicates the reasonability of directly using the filtered image as the estimated precover. The method in [16] is slightly better than SIEp$\big(Wie(3\times3)\big)$-UERD($\mathbf{x}$) but at the cost of heavy time consumption, which is unacceptable in real-time online applications. Instead, the execution time of SIEp$\big(Wie(3\times3)\big)$-UERD($\mathbf{x}$) only performing STCs once is negligible. When employing the better filter $SSRQC$, SIEp($SSRQC$)-UERD($\mathbf{x}$) outperforms the method in [16] by 2.63% and 4.77%, respectively, in resisting DCTR and GFR. Although SIEp($SSRQC$)-UERD($\mathbf{x}$) takes 14.15s for message embedding because of the high computational complexity of $SSRQC$, it is still much faster than the method in [16]. Obviously, SIEp($SSRQC$)-UERD($\mathbf{x}$) can combine the idea of minimizing the feature distance as done in [16], but the corresponding profit is marginal. Therefore, we recommend the efficient and safe method SIEp$\big(Wie(3\times3)\big)$-$\mathcal{A}(\mathbf{x})$ and the safest method SIEp($SSRQC$)-$\mathcal{A}(\mathbf{x})$ for real-world steganography.

V. Conclusion

In this paper, we proposed a general framework of estimating side-information for JPEG steganography. To solve the two critical problems under the SIE framework, we employed several denoising and deblocking filters for better estimation of the precover, and we introduced two implementation models for modulating the costs. The experimental results validated that the proposed method SIEp($SSRQC$)-$\mathcal{A}(\mathbf{x})$ improved additive distortion functions by a large margin for different image sets of several quality factors, and it outperformed the state-of-the-art estimated side-information-based methods. Obviously, how to precisely estimate the side-information is critical for improving the security of the SIE-based method and thus needs further investigation. Furthermore, how to incorporate the estimated side-information as the knowledge of the selection channel [24] for steganalysis is another important and interesting issue.

References

[1] J. Fridrich, *Steganography in digital media: principles, algorithms, and applications*. Cambridge University Press, 2009.

[2] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 920–935, 2011.

[3] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP Journal on Information Security*, vol. 2014, no. 1, p. 1, 2014.

[4] L. Guo, J. Ni, W. Su, C. Tang, and Y.-Q. Shi, "Using statistical image model for jpeg steganography: uniform embedding revisited," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2669–2680, 2015.

[5] X. Hu, J. Ni, and Y.-Q. Shi, "Efficient jpeg steganography using domain transformation of embedding entropy," *IEEE Signal Processing Letters*, vol. 25, no. 6, pp. 773–777, 2018.

[6] W. Su, J. Ni, X. Li, and Y.-Q. Shi, "A new distortion function design for jpeg steganography using the generalized uniform embedding strategy," *IEEE Transactions on Circuits and Systems for Video Technology*, 2018.

[7] K. Chen, H. Zhou, W. Zhou, W. Zhang, and N. Yu, "Defining cost functions for adaptive jpeg steganography at the microscale," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 4, pp. 1052–1066, 2019.

[8] W. Zhou, W. Li, K. Chen, H. Zhou, W. Zhang, and N. Yu, "Controversial pixelprior rule for jpeg adaptive steganography," *IET Image Processing*, 2018.

[9] W. Li, W. Zhang, K. Chen, W. Zhou, and N. Yu, "Defining joint distortion for jpeg steganography," in *Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Security*. ACM, 2018, pp. 5–16.

[10] A. D. Ker, "A fusion of maximum likelihood and structural steganalysis," in *International Workshop on Information Hiding*. Springer, 2007, pp. 204–219.

[11] Y. Kim, Z. Duric, and D. Richards, "Modified matrix encoding technique for minimal distortion steganography," in *International Workshop on Information Hiding*. Springer, 2006, pp. 314–327.

[12] V. Sachnev, H. J. Kim, and R. Zhang, "Less detectable jpeg steganography method based on heuristic optimization and bch syndrome coding," in *Proceedings of the 11th ACM workshop on Multimedia and security*. ACM, 2009, pp. 131–140.

[13] T. Denemark and J. Fridrich, "Side-informed steganography with additive distortion," in *Information Forensics and Security (WIFS), 2015 IEEE International Workshop on*. IEEE, 2015, pp. 1–6.

[14] T. Denemark and J. Fridrich, "Steganography with multiple jpeg images of the same scene," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 10, pp. 2308–2319, 2017.

[15] Z. Wang, Z. Yin, and X. Zhang, "Asymmetric distortion function for jpeg steganography using block artifact compensation," *International Journal of Digital Crime and Forensics (IJDCF)*, vol. 11, no. 1, pp. 90–99, 2019.

[16] Z. Wang, Z. Qian, X. Zhang, M. Yang, and D. Ye, "On improving distortion functions for jpegsteganography," *IEEE Access*, 2018.

[17] C. Zhao, J. Zhang, S. Ma, X. Fan, Y. Zhang, and W. Gao, "Reducing image compression artifacts by structural sparse representation and quantization constraint prior," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 10, pp. 2057–2071, 2017.

[18] P. Bas, T. Filler, and T. Pevnỳ, " break our steganographic system: The ins and outs of organizing boss," in *Information Hiding*. Springer, 2011, pp. 59–70.

[19] P. Bas and T. Furon, "Break our watermarking system," 2008, available: http://bows2.ec-lille.fr/.

[20] T. Filler and J. Fridrich, "Gibbs construction in steganography," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 4, pp. 705–720, 2010.

[21] V. Holub and J. Fridrich, "Low-complexity features for jpeg steganalysis using undecimated dct," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 2, pp. 219–228, 2015.

[22] X. Song, F. Liu, C. Yang, X. Luo, and Y. Zhang, "Steganalysis of adaptive jpeg steganography using 2d gabor filters," in *Proceedings of the 3rd ACM Workshop on Information Hiding and Multimedia Security*. ACM, 2015, pp. 15–23.

[23] J. Kodovsky, J. Fridrich, and V. Holub, "Ensemble classifiers for steganalysis of digital media," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 432–444, 2012.

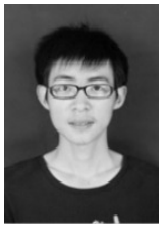
[24] T. D. Denemark, M. Boroumand, and J. Fridrich, "Steganalysis features for content-adaptive jpeg steganography," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 8, pp. 1736–1746, 2016.

Weixiang Li received the B.S. degree from Xidian University, Xi'an, China, in 2016. He is currently pursuing the Ph.D. degree with the University of Science and Technology of China. His research interests include steganography and steganalysis. He was a recipient of the Best Student Paper Award of 6th ACM IH&MMSec in 2018.

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

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2019.2925118, IEEE Transactions on Circuits and Systems for Video Technology

Weiming Zhang received his M.S. degree and PH.D. degree in 2002 and 2005 respectively from the Zhengzhou Information Science and Technology Institute, Zhengzhou, China. Currently, he is 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.



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

Yaofei Wang received the B.S. degree from the School of Physical Science and Technology, Southwest Jiaotong University, in 2017. 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.

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