# Shortening the Cover for Fast JPEG Steganography

Weixiang Li, Wenbo Zhou, Weiming Zhang, Chuan Qin, Huanhuan Hu, Nenghai Yu

***Abstract*—Recently, the most effective steganographic schemes for JPEG images are based on minimal distortion model with Syndrome-Trellis Codes (STCs) as the coding method. However, the execution time of STCs will be severe for message embedding to the cover object of large size, which cannot meet the demand of real-time communication in real-world application. According to the time complexity $O(2^h n)$, it is suggested in STCs to accelerate the embedding process by decreasing the constraint height $h$. However, smaller $h$ corresponds to lower steganographic security. In this paper, we investigate the possibility of shortening the cover (reducing the length $n$) for speeding up the execution of STCs without weakening the steganographic security. After introducing some properties of cover selection with proofs, we propose several algorithms designed for JPEG images to construct a preferable shortened cover containing DCT coefficients of smaller costs as much as possible. Experimental results display the superiority of the proposed algorithm on the speed profit and the security when compared with the method of decreasing $h$. With confidence, a JPEG image of arbitrary quality factor can be safely shortened as 1/4 of the original, and correspondingly the execution of STCs can be 4 times faster.**

***Index Terms*—Steganography, JPEG images, minimal distortion, syndrome-trellis codes, cover selection, fast embedding.**

## I. Introduction

Modern steganography aims to embed a covert message in a digital cover object by slightly changing its elements without drawing suspicions from steganalysis [1], [2]. During the development of image steganography, some steganographic methods pursued large embedding capacity by using multi-bit technique (e.g., [3]–[5]), and some valued strong robustness (e.g., [5]–[8]). With the advance of modern steganalysis based on high-dimensional feature sets [9]–[14], the undetectability (security) of these methods are relatively weak. Currently, the most effective steganographic schemes are categorized as adaptive steganography that is based on minimal distortion model [15], [16]. The adaptive steganography strives for high-level undetectability in resisting modern steganalysis and lacks robustness, which is only applicable to noiseless channel [1].

The minimal distortion based adaptive steganography usually comprises of a heuristically designed distortion function and a method for encoding the message to minimize the distortion. The distortion function element-wisely evaluates the effect of individual embedding modification. And the minimal distortion can be realized in practice using a general methodology of syndrome coding [16], [17] via the parity-check matrix of error-correcting codes.

As a widely adopted format for image storage and transmission, JPEG steganography has become a research hotspot over the past few years. There emerge various content-adaptive distortion functions [18]–[25] designed for JPEG steganography. Meanwhile, microscale steganography [26], cost spreading rule [26], controversial-pixel-prior rule [27] and non-additive principle of block-boundary-continuity [28] are extended from spatial image steganography to help improve the performance of above JPEG distortion functions.

Syndrome-Trellis Codes (STCs) [16] are vital for adaptive steganography, since they provide a general and efficient syndrome coding method that can asymptotically approach the theoretical bound of average embedding distortion for arbitrary additive distortion function. The time and space complexities of STCs are $O(2^h n)$, where $n$ is the length of cover object and $h$ is the constraint height of the parity-check submatrix. The height $h$ is a design parameter that affects the coding performance and speed, and larger $h$ corresponds to better coding performance but lower speed.

The execution speed of STCs is an important issue for real-world steganography. Since multimedia files of large size can be transmitted rapidly in the upcoming 5G Networks, steganography using cover object of large size will be more and more common, and thus STCs need to be fast enough to meet the demand of real-time communication. Although the time complexity of STCs is linear with the length of cover, the time consumption of message embedding is still severe for the large cover. In a laboratory environment, ternary STCs with common $h = 10$ requires 0.7 seconds [1] to embed message into a JPEG image of size $512 \times 512$. But in the real world, the size of high-resolution images taken by smart phones would reach $4160 \times 3120$. And for that large image, STCs will consume a terrible 38.3 seconds or even more time if on a smart phone's processor, which is clearly unacceptable for real-time communication. Therefore, to move steganography from the laboratory into the real world, it is of practical significance to accelerate the execution of STCs.

Since the constraint height $h$ is a design parameter for STCs, it has been recommended to speed up the embedding process by decreasing $h$ [16]. However, $h$ is related to the capacity of minimizing distortion, i.e., decreasing $h$ will do harm to the steganographic security. From the complexity $O(2^h n)$, here comes a question: is it possible to achieve the fast embedding by reducing the cover length $n$ while maintaining the steganographic security? Obviously, a simple approach is to



This work was supported in part by the Natural Science Foundation of China under Grant U1636201 and 61572452, and by Anhui Initiative in Quantum Information Technologies under Grant AHY150400.

The authors are with the 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).

[1] The source codes of STCs are downloaded from the webpage of DDE Laboratory (http://dde.binghamton.edu/download/), and the result is obtained by Intel(R) Core(TM) i5-4590 CPU @ 3.30GHz.

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.2019.2908689, IEEE Transactions on Circuits and Systems for Video Technology



select part of DCT coefficients as the shortened cover (e.g., the left half of the image). However, such sequential embedding may not withstand the low order statistical attacks [29]. And in view of adaptive steganography, the shortened cover cannot ensure to be a texture-complex region that is more suitable for modifications. Moreover, the embedding payload of the shortened cover increases, so that the steganographic security would be inevitably lowered.

Intuitively, based on the minimal distortion model, we can minimize the negative influence of shortening the cover as long as the top coefficients of smallest costs are selected as the shortened cover. However, the costs calculated from the stego image are different from that of the cover image, so the receiver cannot obtain the accurate shortened stego from the costs for extracting the message correctly. Therefore, it is critical to find some cover selection strategies for obtaining a preferable shortened cover, which can preserve the steganographic security after message embedding and from which the receiver can correctly extract the message.

This paper focus on investigating the possibility of shortening the cover (reducing the length $n$) to achieve secure and faster embedding when using STCs for adaptive steganography. In this paper, we firstly study the impact of cover selection on the average distortion under the minimal distortion model, and derive the optimality of cover selection. To approach that optimality, several cover selection algorithms are proposed for JPEG images and then compared with each other via their minimal average distortions. Experimental results show that the best algorithm displays the superiority on speeding up the embedding process with the minimum negative impact in steganography when compared with the method of decreasing $h$. For the faster execution of STCs, a JPEG image of arbitrary quality factor can be shortened as 1/4 of the original size without weakening the steganographic security.

The rest of this paper is organized as follows. In Section II, the framework of minimal additive distortion and syndrome-trellis codes are briefly reviewed. We introduce the abstract problem and some properties of cover selection with proofs in Section III. To obtain a preferable shortened cover, four cover selection algorithms are proposed for JPEG steganography in Section IV. Experimental results and comparisons are presented in Section V, and the paper is concluded in Section VI.

## II. Preliminaries

In this paper, sets, vectors and matrices are always written in boldface. The cover and stego sequences are denoted by $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ and $\mathbf{y} = (y_1, y_2, \cdots, y_n)$ respectively, where $x_i$ is the value of a quantized DCT coefficient in JPEG images. The embedding operation on $x_i$ is formulated by the dynamic range $\mathcal{I}_i$. For binary embedding, $\mathcal{I}_i = \{x_i, \bar{x}_i\}$ where $\bar{x}_i$ is $x_i$ after flipping its Least Significant Bit (LSB), and $\mathcal{I}_i = \{x_i - 1, x_i, x_i + 1\}$ is for ternary embedding. $k$-ary entropy function is denoted by $H(\pi_1, \cdots, \pi_k)$ for $\sum_{i=1}^{k} \pi_i = 1$, where binary entropy function is $H(\pi) = -\pi \log \pi - (1-\pi)\log(1-\pi)$.

### A. Minimal distortion model

Under an additive distortion scenario, the impacts of embedding changes are assumed to be mutually independent, so the total distortion $D(\mathbf{x}, \mathbf{y})$ for embedding is the sum of the costs $\rho(y_i)$ at $x_i$ changed to $y_i$ [16], that is

$$D(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \rho(y_i). \quad (1)$$

Denote $\pi(y_i)$ as the probability of modifying $x_i$ to $y_i$, the sender wants to embed message of $m$ bits while minimizing the average distortion $E_{\boldsymbol{\pi}}(D)$ by the following optimization problem:

$$\underset{\boldsymbol{\pi}}{\text{minimize}}\ E_{\boldsymbol{\pi}}(D) = \sum_{i=1}^{n} \sum_{t_i \in \mathcal{I}_i} \pi(t_i)\rho(t_i) \quad (2)$$

$$\text{subject to } H(\boldsymbol{\pi}) = -\sum_{i=1}^{n} \sum_{t_i \in \mathcal{I}_i} \pi(t_i) \log \pi(t_i) = m. \quad (3)$$

Following the maximum entropy principle, the optimal modification probability $\boldsymbol{\pi}_\lambda$ has been proven to have a Gibbs distribution [15], [30], which is the best mapping from the cost to the modification probability for the additive minimal distortion model. Concretely, the optimal $\boldsymbol{\pi}_\lambda$ is calculated by

$$\pi_\lambda(y_i) = \frac{\exp(-\lambda\rho(y_i))}{\sum_{t_i \in \mathcal{I}_i} \exp(-\lambda\rho(t_i))}, 1 \le i \le n, \quad (4)$$

where the scalar parameter $\lambda > 0$ is determined by (3).

The optimality of $\boldsymbol{\pi}_\lambda$ implies that $E_{\boldsymbol{\pi}}(D)$ of any probability distribution $\boldsymbol{\pi}$ satisfying (3) cannot be smaller than $E_{\boldsymbol{\pi}_\lambda}(D)$, i.e.,

$$E_{\boldsymbol{\pi}}(D) \ge E_{\boldsymbol{\pi}_\lambda}(D) \quad (5)$$

with equality iff $\boldsymbol{\pi} = \boldsymbol{\pi}_\lambda$.

### B. Syndrome coding and syndrome-trellis codes

For a binary embedding operation, the optimization problem (2)-(3) can be realized in practice using syndrome coding with the message $\mathbf{m}$ embedding and extraction mappings:

$$\begin{aligned} \text{Emb}(\mathbf{x}, \mathbf{m}) &= \arg \min_{\mathcal{P}(\mathbf{y}) \in \mathcal{C}(\mathbf{m})} D(\mathbf{x}, \mathbf{y}) \\ \text{Ext}(\mathbf{y}) &= \mathcal{P}(\mathbf{y})\mathbb{H}^{\mathrm{T}} = \mathbf{m}, \end{aligned} \quad (6)$$

where $\mathcal{P} : \mathcal{X} \to \{0, 1\}$ is a parity function shared between the sender and the receiver (e.g., $\mathcal{P}(x) = x \bmod 2$), $\mathbb{H}^{\mathrm{T}} \in \{0, 1\}^{n \times m}$ is a parity-check matrix of the binary code $\mathcal{C}$, and $\mathcal{C}(\mathbf{m}) = \{\mathbf{z} \in \{0, 1\}^n | \mathbf{z}\mathbb{H}^{\mathrm{T}} = \mathbf{m}\}$ is the coset corresponding to syndrome $\mathbf{m}$.

Syndrome-Trellis Codes (STCs) [16] provide a kind of syndrome coding methods that can approach the minimal average distortion for arbitrary additive distortion. In other words, the total distortion $D(\mathbf{x}, \mathbf{y})$ after performing STCs can approach the minimal average distortion $E_{\boldsymbol{\pi}_\lambda}(D)$. For STCs, each solution can be represented as a path through the syndrome trellis of $\mathbb{H}^{\mathrm{T}}$, which is constructed by a submatrix $\hat{\mathbb{H}}$ of size $h \times w$ ($w = n/m$). The height $h$ determines the number of paths, and there are $k^h$ choices in each grid of the trellis for $k$-ary embedding. Therefore, larger $h$ means more powerful capacity to minimize distortion but also higher computational complexity. Thanks to the multi-layered construction [16], STCs can also be fast implemented for ternary embedding in time complexity $O(2^h n)$. This implies that the time complexity of STCs is proportional to $2^h$ and $n$.

## III. Abstract and Properties of Cover Selection

Denote $\theta$ $(0<\theta<1)$ as the shorten-rate, the cover selection is selecting $\theta n$ elements from the cover of length $n$ to construct a shortened cover for faster message embedding. For illustration purposes, we assume $\theta n$ and $\theta^{-1}$ are integers in this paper. Suppose that the execution time of STCs on the original cover is $T$, the execution time of STCs on the shortened cover can be reduced to $\theta T$ according to the complexity $O(2^h n)$. In this section, we will abstract the problem of cover selection and introduce some properties that can be used to instruct the cover selection. Without loss of generality, the following cover selection properties are proved using the case of binary embedding in which the cost of element not being changed is 0. The properties can be similarly proved for $k$-ary embedding. Note that the proofs are based on the optimality of Gibbs distribution (4) with the minimal average distortion (5).

### A. Abstract problem and lossy property

For the message of $m$ bits and the cover sequence $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ with cost $\rho(\bar{\mathbf{x}}) = \big(\rho(\bar{x}_1), \rho(\bar{x}_2), \cdots, \rho(\bar{x}_n)\big)$, the average distortion introduced by message embedding on $\mathbf{x}$

$$E_{\boldsymbol{\pi}}(D_{\mathbf{x}}) = \sum_{i=1}^{n} \pi(\bar{x}_i)\rho(\bar{x}_i) \tag{7}$$

can be minimized when the probability distribution $\boldsymbol{\pi}$ follows Gibbs distribution (4). In this manner, the minimal average distortion is computed by

$$E_{\boldsymbol{\pi}_\lambda}(D_{\mathbf{x}}) = \sum_{i=1}^{n} \pi_\lambda(\bar{x}_i)\rho(\bar{x}_i) \tag{8}$$

where

$$\pi_\lambda(\bar{x}_i) = \frac{\exp(-\lambda\rho(\bar{x}_i))}{1+\exp(-\lambda\rho(\bar{x}_i))} \quad \text{and} \quad m = \sum_{i=1}^{n} H(\pi_\lambda(\bar{x}_i)). \tag{9}$$

For a given shorten-rate $\theta$, $\theta n$ elements are selected from $\mathbf{x}$ to form a shortened cover $\mathbf{s} = (s_1, s_2, \cdots, s_{\theta n})$, and its modification cost is $\rho(\bar{\mathbf{s}}) = \big(\rho(\bar{s}_1), \rho(\bar{s}_2), \cdots, \rho(\bar{s}_{\theta n})\big)$ where $s_i \in \mathbf{x}$ and $\rho(\bar{s}_i) \in \rho(\bar{\mathbf{x}})$. Similarly, to embed the same message into $\mathbf{s}$, the minimal average distortion of the optimal distribution $\boldsymbol{\mu}_{\lambda'}$ about $\rho(\bar{\mathbf{s}})$ becomes

$$E_{\boldsymbol{\mu}_{\lambda'}}(D_{\mathbf{s}}) = \sum_{i=1}^{\theta n} \mu_{\lambda'}(\bar{s}_i)\rho(\bar{s}_i) \tag{10}$$

where

$$\mu_{\lambda'}(\bar{s}_i) = \frac{\exp(-\lambda'\rho(\bar{s}_i))}{1+\exp(-\lambda'\rho(\bar{s}_i))} \quad \text{and} \quad m = \sum_{i=1}^{\theta n} H(\mu_{\lambda'}(\bar{s}_i)). \tag{11}$$

Since the probability distribution about the complementary set $\rho(\bar{\mathbf{x}}) \setminus \rho(\bar{\mathbf{s}})$ can be regarded as $\mathbf{0}$, we can consider that the distribution $\boldsymbol{\mu}_{\lambda'}$ complemented with $\mathbf{0}$ is a distribution about $\rho(\bar{\mathbf{x}})$, i.e., $\boldsymbol{\pi} = (\boldsymbol{\mu}_{\lambda'}, \mathbf{0})$. Obviously, $E_{\boldsymbol{\pi}}(D_{\mathbf{x}}) = E_{\boldsymbol{\mu}_{\lambda'}}(D_{\mathbf{s}})$ and $H(\boldsymbol{\pi}) = m$. From (5), we thus have

$$E_{\boldsymbol{\mu}_{\lambda'}}(D_{\mathbf{s}}) > E_{\boldsymbol{\pi}_\lambda}(D_{\mathbf{x}}). \tag{12}$$

This implies that the minimal average distortion will be increased by the cover selection, i.e, shortening the cover is a lossy operation on the average distortion of steganography.

As pointed out in [31], embedding efficiency directly influences the steganographic security, and larger embedding efficiency leads to stronger steganographic security. An established way of evaluating coding algorithms in steganography is to compare the embedding efficiency $e(m) = m/E_{\boldsymbol{\pi}}(D)$ (in bits per unit distortion) [16] for a fixed message of $m$ bits with the upper bound derived from (2). The definition formula reveals the equivalence relation between the embedding efficiency and the steganographic distortion. With the increase of $E_{\boldsymbol{\pi}}(D)$ that STCs can approach, the corresponding embedding efficiency is decreased, and thus the steganographic security of performing STCs on the shortened cover would be weakened to some extents.

The proof also reveals the abstraction of the cover selection problem: for a given message and any shorten-rate, the probability distribution about the cost of the shortened cover corresponds to a nonoptimal probability distribution about the cost of the original cover. The nonoptimal distribution about the complete cost is the combination of the optimal (Gibbs) distribution about the selected costs and the zero-distribution about the unselected costs.

### B. The optimality of cover selection

For a given shorten-rate $\theta$, selecting appropriate $\theta n$ elements from the cover is vital for reducing the negative impact of cover selection. Here we prove the optimality of cover selection: selecting the top $\theta n$ elements of smallest costs to construct a shortened cover.

Assume that the top $\theta n$ elements of smallest costs are selected to form the shortened cover $\mathbf{s}_t = (s_1^t, s_2^t, \cdots, s_{\theta n}^t)$, and the cost of $\mathbf{s}_t$ is $\rho(\bar{\mathbf{s}}_t) = \big(\rho(\bar{s}_1^t), \rho(\bar{s}_2^t), \cdots, \rho(\bar{s}_{\theta n}^t)\big)$. For a arbitrary shortened cover $\mathbf{s}$ with (10) and (11), when $\rho(\bar{\mathbf{s}}_t)$ uses the optimal distribution $\boldsymbol{\mu}_{\lambda'}$ about $\rho(\bar{\mathbf{s}})$, we have

$$\sum_{i=1}^{\theta n} \mu_{\lambda'}(\bar{s}_i)\rho(\bar{s}_i^t) \le \sum_{i=1}^{\theta n} \mu_{\lambda'}(\bar{s}_i)\rho(\bar{s}_i) = E_{\boldsymbol{\mu}_{\lambda'}}(D_{\mathbf{s}}) \tag{13}$$

because $\rho(\bar{s}_i^t) \le \rho(\bar{s}_i)$ for $1 \le i \le \theta n$. Obviously, $\boldsymbol{\mu}_{\lambda'}$ is the optimal distribution about $\rho(\bar{\mathbf{s}})$ but may not be the optimal $\boldsymbol{\mu}^t_{\lambda''}$ about $\rho(\bar{\mathbf{s}}_t)$, so that from (5)

$$E_{\boldsymbol{\mu}^t_{\lambda''}}(D_{\mathbf{s}_t}) \le \sum_{i=1}^{\theta n} \mu_{\lambda'}(\bar{s}_i)\rho(\bar{s}_i^t). \tag{14}$$

From (13) and (14), we can derive

$$E_{\boldsymbol{\mu}^t_{\lambda''}}(D_{\mathbf{s}_t}) \le E_{\boldsymbol{\mu}_{\lambda'}}(D_{\mathbf{s}}) \tag{15}$$

with equality iff $\rho(\bar{\mathbf{s}}_t) = \rho(\bar{\mathbf{s}})$. It reveals that the smallest minimal average distortion of cover selection can be achieved if the top elements of smallest costs are selected to form the shortened cover. In other words, the negative impact of shortening the cover can be minimized by the optimality of cover selection.

As mentioned before, the receiver cannot obtain the accurate stego elements to extract the message because of the non-synchronous costs calculated from the cover and stego image. Thus the optimality of cover selection is inaccessible in practice. But it indicates the goal of shortening the cover: selecting the elements of smaller costs as much as possible.

### C. The smaller the shorten-rate, the worse the optimality

This proof focuses on the optimality of cover selection w.r.t. different shorten-rates. Suppose a smaller shorten-rate $\theta'$ ($\theta' < \theta$), the optimal shortened cover $\mathbf{s}'_t$ of length $\theta' n$ is obviously a subset of the optimal shortened cover $\mathbf{s}_t$ of length $\theta n$. So the relation between $\mathbf{s}'_t$ and $\mathbf{s}_t$ can be regarded as the abstract problem of cover selection where $\mathbf{s}_t$ is shortened to $\mathbf{s}'_t$. From (12), we derive

$$E_{\boldsymbol{\mu}^{t'}_{\lambda'''}}(D_{\mathbf{s}'_t}) > E_{\boldsymbol{\mu}^{t}_{\lambda''}}(D_{\mathbf{s}_t}), \quad (16)$$

which indicates that with the decrease of shorten-rate, the minimal average distortion of optimal cover selection is inevitably increasing. Although the proof is for the optimality that cannot be achieved, we believe that the coding performance of STCs on different shorten-rates will follow a relatively consistent trend. Therefore, the shorten-rate $\theta$ is a design parameter that affects the security and speed of performing STCs, like the height $h$.

Through the above analysis, we can conclude that

- Shortening the cover with any shorten-rate will increase the average distortion of steganography.
- For a given shorten-rate, selecting the top elements of smallest costs is optimal for minimizing the negative impact of cover selection; since the optimality is inaccessible, selecting the elements of smaller costs as much as possible is recommended.
- The shorten-rate affects the performance of shortening the cover that, smaller shorten-rate comes with greater embedding speed but larger steganographic distortion.

## IV. Algorithms of Cover Selection for JPEG Steganography

For a JPEG image of quality factor QF and size $M \times N$, denote $q_{uv}$ as the quantization step at frequency $(u, v)$, and $x^{(kl)}_{uv}$ as the quantized coefficient at frequency $(u, v)$ of the $(k, l)$th DCT block, where $k \in \{0, 1, \cdots, M/8 - 1\}$, $l \in \{0, 1, \cdots, N/8 - 1\}$ and $u, v \in \{0, 1, \cdots, 7\}$. Defining the DC quantization step $q_{00} = (q_{01} + q_{10})/2$ as done in [20], we sort the 64 $q_{uv}$s in an ascending order (zigzag-scanning-order prior for the equal quantization steps). Denote $f_i$ as the frequency $(u, v)$ of $i$th smallest $q_{uv}$, the cover sequence $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ of $n = M \times N$ can be rewritten as $\mathbf{x} = (\mathbf{x}_{f_1}, \cdots, \mathbf{x}_{f_{64}})$, where $\mathbf{x}_{f_i} = (x^{(00)}_{f_i}, \cdots, x^{(kl)}_{f_i}, \cdots, x^{((M/8-1)(N/8-1))}_{f_i})$ is the vector of all coefficients at frequency $f_i$. Fig. 1 illustrates the arrangement of a JPEG image into $\mathbf{x} = (x_1, x_2, \cdots, x_n) = (\mathbf{x}_{f_1}, \cdots, \mathbf{x}_{f_{64}})$.

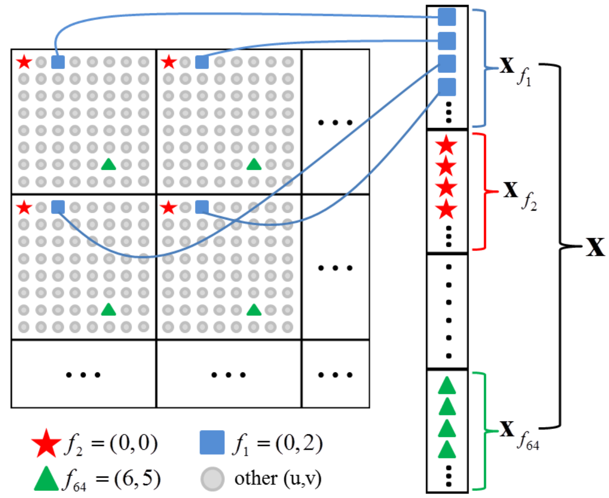



Fig. 1: Arrangement of a JPEG image into cover sequence $\mathbf{x} = (x_1, x_2, \cdots, x_n) = (\mathbf{x}_{f_1}, \cdots, \mathbf{x}_{f_{64}})$. Assume that frequency $(0, 2)$ has the smallest quantization step, frequency $(0, 0)$ has the second smallest quantization step, and frequency $(6, 5)$ has the largest quantization step.

Assume the modification $\xi_i = y_i - x_i \in \{-1, 0, +1\}$ for ternary embedding, the relation $\rho(x^+_i) = \rho(x^-_i) = \rho(\bar{x}_i)$ is common in additive distortion functions [18]–[27], where $\rho(x^+_i)$ and $\rho(x^-_i)$ are the cost of modifying $x_i$ by $+1$ and $-1$ respectively. To simplify the description, we use $\rho(\bar{x}_i)$ to represent $\rho(x^+_i)$ and $\rho(x^-_i)$ in the following algorithms unless otherwise stated.

To minimize the negative impact of cover selection, it is highly recommended to select the elements of smaller costs as much as possible. Under the premise of ensuring that the receiver can correctly extract the message, we design several algorithms for selecting more appropriate DCT coefficients as follows.

### A. Quantization-step partition algorithm

In JPEG distortion functions [19]–[25], the modification costs of coefficients at the same DCT block are distinguished by the quantization steps $q_{uv}$. That is to say within a DCT block, a coefficient of smaller $q_{uv}$ has a smaller cost. Naturally, it is reasonable to select the coefficients of small $q_{uv}$ as the shortened cover for message embedding. It ensures that the relatively appropriate coefficients of small costs can be selected regardless of the texture complexity of the DCT block.

In this manner, we can select the front $\theta n$ coefficients of $\mathbf{x}$ in Fig. 1 as the shortened cover, i.e., $\mathbf{s} = (x_1, \cdots, x_{\theta n})$, because $\mathbf{x}$ is arranged by the quantization steps in an ascending order. After obtaining $\mathbf{s}$ and its cost $\rho(\bar{\mathbf{s}})$, we can perform STCs in a reduced execution time $\theta T$ ($T$ is the original execution time). Note that the receiver can construct the shortened stego using the same operation of quantization-step partition for extracting the message correctly. We denote this algorithm by **Method-1**.

### B. Segment-sum based algorithms

In segment-sum based algorithms, the cover is segmented into $\theta^{-1}$ non-overlapping parts of same length $\theta n$. The sum 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.2019.2908689, IEEE Transactions on Circuits and Systems for Video Technology



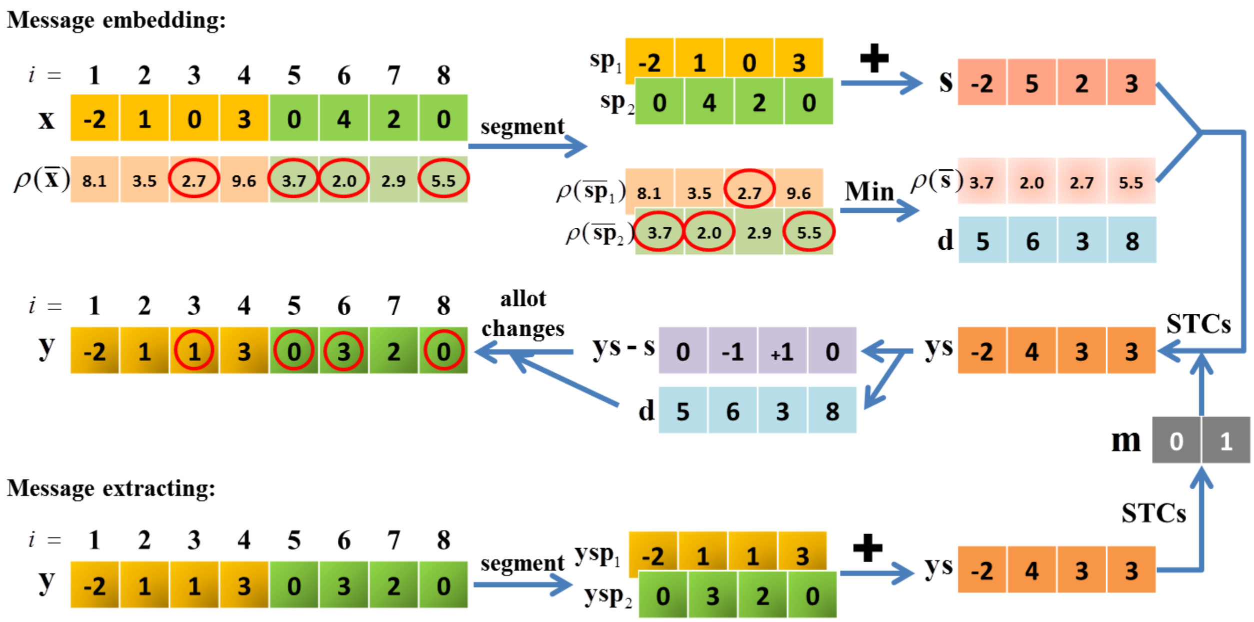



Fig. 2: An example of segment-sum algorithm with shorten-rate $\theta = 1/2$ on the cover of length $n = 8$.

coefficients at the same location of these parts is then served as the element of the shortened cover. Inspired by the optimality of cover selection, the cost of the shortened cover is defined as the smallest cost among coefficients at the same location. And after message embedding, the modifications of the shortened cover will be allotted to the coefficients with the smallest costs at the corresponding locations to generate the stego image. In this manner, the receiver can construct the shortened stego for extracting message via the same operations of segmenting and summing. Fig. 2 provides an example of the segment-sum algorithm with $\theta = 1/2$ on the cover of $n = 8$.

*1) General segment-sum algorithm:* The cover $\mathbf{x}$ should be scrambled (with a key shared with the receiver) before the segmenting operation, which can help select the smaller costs as much as possible due to the effect of randomness. Denote the scrambled cover by $\mathbf{x}' = (x'_1, \cdots, x'_n)$, and its cost by $\rho(\bar{\mathbf{x}}')$. For a given shorten-rate $\theta$, $\mathbf{x}'$ is segmented into $\theta^{-1}$ parts of length $\theta n$,

$$\mathbf{sp}_i = (x'_{(i-1)\theta n+1}, \cdots, x'_{i\theta n}), 1 \le i \le \theta^{-1}. \tag{17}$$

The shortened cover $\mathbf{s}$ is the sum of these $\mathbf{sp}_i$,

$$\mathbf{s} = \sum_{i=1}^{\theta^{-1}} \mathbf{sp}_i = (\sum_{i=1}^{\theta^{-1}} x'_{(i-1)\theta n+1}, \cdots, \sum_{i=1}^{\theta^{-1}} x'_{i\theta n}). \tag{18}$$

And the cost $\rho(\bar{\mathbf{s}}) = \min_{1\le i\le\theta^{-1}} \{\rho(\overline{\mathbf{sp}}_i)\}$ is the smallest one of these $\rho(\overline{\mathbf{sp}}_i)$,

$$\rho(\bar{\mathbf{s}}) = (\min_{1\le i\le\theta^{-1}} \{\rho(\bar{x}'_{(i-1)\theta n+1})\}, \cdots, \min_{1\le i\le\theta^{-1}} \{\rho(\bar{x}'_{i\theta n})\}), \tag{19}$$

where the indices of the selected costs are denoted by $\mathbf{d} = (d_1, \cdots, d_{\theta n})$.

Denote $\mathbf{ys} = (ys_1, \cdots, ys_{\theta n})$ as the shortened stego after performing STCs on $\mathbf{s}$, we can construct the complete scrambled stego $\mathbf{y}' = (y'_1, \cdots, y'_n)$ by

$$y'_i = \begin{cases} x'_i & \text{if } i \notin \mathbf{d} \\ x'_i + (ys_j - s_j) & \text{if } i = d_j \in \mathbf{d} \end{cases}, 1 \le j \le \theta n, 1 \le i \le n, \tag{20}$$

where the modifications are assigned to the coefficients of smaller costs. We finally obtain the stego $\mathbf{y}$ via an inverse operation of scrambling. Note that the receiver can construct the shortened stego using (17) and (18) on the same scrambled stego for extracting the message correctly. We denote this algorithm by **Method-2**. Note that **Method-2** can be applied to general cover object.

*2) Segment-sum algorithm combined with quantization-step partition:* For JPEG images, the costs of coefficients within a DCT block are distinguished by the quantization steps, thus the shortened cover in **Method-1** is constructed by the coefficients of smaller quantization steps. Since the cost of coefficient of larger quantization step in a texture-complex DCT block may be smaller than that of coefficient of smaller quantization step in a texture-smooth DCT block, we can combine **Method-1** with **Method-2** to select more coefficients of smaller costs by the effect of randomness.

For the cover $\mathbf{x} = (x_1, \cdots, x_n)$ in Fig. 1, we firstly segment it into $\theta^{-1}$ parts

$$\mathbf{sp}_i = (x_{(i-1)\theta n+1}, \cdots, x_{i\theta n}), 1 \le i \le \theta^{-1}, \tag{21}$$

where $\mathbf{sp}_i$ of smaller $i$ corresponds to smaller quantization step, and $\mathbf{sp}_1$ is the selected shortened cover in **Method-1**. However, we cannot directly apply (18)-(20) to (21) because the cost $\rho(\bar{\mathbf{s}})$ in (19) would remain the same as $\rho(\overline{\mathbf{sp}}_1)$. When $64\theta$ is an integer, $\rho(\bar{x}_{(i-1)\theta n+1})$ for $1 \le i \le \theta^{-1}$ are in the same DCT block and thus $\min_{1\le i\le\theta^{-1}} \{\rho(\bar{x}_{(i-1)\theta n+1})\} = \rho(\bar{x}_1)$. So $\rho(\bar{\mathbf{s}}) = \rho(\overline{\mathbf{sp}}_1)$ in (19) implies that **Method-2** does not help obtain a better shortened cover. Similarly, each $\mathbf{sp}_i$ should be scrambled so that (18)-(20) can work for the scrambled $\mathbf{sp}'_i$.

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.2019.2908689, IEEE Transactions on Circuits and Systems for Video Technology



Due to the effect of randomness, coefficients from diverse DCT blocks (include texture-complex and texture-smooth blocks) have chances to be summed as an element of the shortened cover. And thus the cost of the element may adopt the smaller cost of coefficient of larger quantization step in the texture-complex DCT block. In this manner, more coefficients of smaller costs can be selected to construct a better shortened cover. Note that the receiver can extract the message correspondingly. We denote this algorithm by **Method-3**.

Obviously, **Method-3**, that combines the extra knowledge of cost characteristic in JPEG images, is an enhanced version of **Method-2**. This indicates that **Method-2** can be improved for general cover object if combining some useful cost characteristics of the cover object.

### *C. Cover-smoothed and cost-wetted algorithm*

As proved in Section III, the optimality of cover selection cannot be achieved because the costs calculated from the images before and after message embedding are non-synchronous. So the above algorithms are well designed under the premise that the receiver can construct the same shortened stego with no need for the costs. Here we propose to synchronize an approximative cover between the sender and receiver so as to compute the synchronous approximative costs. With the synchronous approximative costs, the sender can select the top coefficients of smallest approximative costs as the shortened cover. And the receiver can correspondingly construct the accurate shortened stego for extracting the message. Because the cost of the original cover is more precise, it is better for the sender to use the original costs of the selected top coefficients for message embedding. In this way, the optimality of cover selection is approximatively reachable.

For binary embedding, we can define the approximative cover $\mathbf{x}_2 = \mathbf{x} - (\mathbf{x} \mod 2)$ ignoring the Least Significant Bit (LSB) of coefficient. Because only the LSB of coefficient may be modified during the embedding process, the receiver can easily obtain the approximative cover $\mathbf{x}_2 = \mathbf{y} - (\mathbf{y} \mod 2)$. Having the same cover ensures that the sender and the receiver can compute the approximative enough and synchronous cost $\rho(\bar{\mathbf{x}}_2)$. Thus the sender can construct a shortened cover according to the top smallest approximative costs and use the original costs of the selected coefficients for message embedding.

It is known that the common case in adaptive steganography is ternary embedding which can achieve the smaller embedding impact. For ternary embedding, all bits of coefficient may be changed by the modification of $+1$ or $-1$, indicating that the above approximation operation for binary embedding is not suitable for the ternary case. Fortunately, we can construct a sharable cover by smoothing the cover and introducing some wet coefficients with infinite costs.

Firstly, the sender smoothes the cover and obtains the approximative cover $\mathbf{x}_\beta = |\mathbf{x}| - (|\mathbf{x}| \bmod \beta)$. The receiver can obtain the same smoothed cover $\mathbf{y}_\beta = |\mathbf{y}| - (|\mathbf{y}| \bmod \beta) = \mathbf{x}_\beta$, i.e.,

$$|x_i + \xi_i| - (|x_i + \xi_i| \mod \beta) = |x_i| - (|x_i| \mod \beta), \quad (22)$$

under the condition that the modifications of $\xi_i = +1$ and $\xi_i = -1$ violating (22) are forbidden by assigning them infinite costs. Therefore, we update the costs $\rho'(\mathbf{x}^+)$ and $\rho'(\mathbf{x}^-)$ by

$$\rho'(x_i^+) = \begin{cases} \infty & \text{if } x_i > 0 \text{ and } x_i \bmod \beta = \beta - 1 \\ \infty & \text{if } x_i < 0 \text{ and } x_i \bmod \beta = 0 \\ \rho(x_i^+) & \text{otherwise} \end{cases}, 1 \le i \le n, \quad (23)$$

$$\rho'(x_i^-) = \begin{cases} \infty & \text{if } x_i > 0 \text{ and } x_i \bmod \beta = 0 \\ \infty & \text{if } x_i < 0 \text{ and } x_i \bmod \beta = \beta - 1 \\ \rho(x_i^-) & \text{otherwise} \end{cases}, 1 \le i \le n, \quad (24)$$

which can be proved to hold (22). After calculating the smoothed cost $\rho(\bar{\mathbf{x}}_\beta)$ of $\mathbf{x}_\beta$, we sort it in an ascending order and select its top $\theta n$ smallest costs, of which the indices are denoted by $\mathbf{d} = (d_1, \cdots, d_{\theta n})$. We then construct the shortened cover $\mathbf{s} = (x_{d_1}, \cdots, x_{d_{\theta n}})$ and its original cost $\rho(\mathbf{s}^+) = \left(\rho'(x_{d_1}^+), \cdots, \rho'(x_{d_{\theta n}}^+)\right)$, $\rho(\mathbf{s}^-) = \left(\rho'(x_{d_1}^-), \cdots, \rho'(x_{d_{\theta n}}^-)\right)$, so that STCs can be executed in a reduced time. Obviously, the receiver can compute the same cost $\rho(\bar{\mathbf{x}}_\beta)$ because of $\mathbf{y}_\beta = \mathbf{x}_\beta$, to obtain $\mathbf{d}$ for extracting the message. We denote this algorithm by **Method-4**.

Note that the modulus $\beta \ge 2$ used in (22)-(24) determines not only the approximation degree of the smoothed cost to the original cost but also the number of wet coefficients. Obviously, smaller $\beta$ will make the smoothed cost closer to the original but the larger wetness (the ratio of the number of wet coefficients to $n$). The impact of $\beta$ will be discussed in Section V.

As elaborated above, four cover selection algorithms are proposed utilizing the frequency knowledge of JPEG images, the effect of randomness and approximation, so as to select coefficients of smaller costs as much as possible. Different algorithms lead to different shortened covers and their corresponding minimal average distortions (10). We will compare the advantages of algorithms via their minimal average distortions. As mentioned in Section III, the embedding distortion is equivalent to the embedding efficiency for a given message, and larger embedding efficiency results in stronger steganographic security. Obviously, the smaller the distortion, the larger the embedding efficiency, the securer the algorithm.

## V. Experimental Results and Analysis

There exist two approaches for accelerating the execution of STCs: decreasing the height $h$ or reducing the length $n$. According to the complexity $O(2^h n)$, $h-1$ and $n/2$ have the equivalent capacity in theory to accelerate the embedding process. Thus we choose shorten-rates $\theta = 1/2, 1/4, 1/8$ in comparisons with $h = 9, 8, 7$ on the basic of the common $h = 10$ for illustration. Under the minimal distortion model, the steganographic security can be measured in two ways. One is the embedding efficiency (or the distortion) introduced by message embedding, and the other is the undetectability against the actual steganalysis. The former can determine the latter. As mentioned before, if the distortion increment is negligible, the embedding efficiency loss can be neglected and thus the steganographic security can be preserved. We will compare the above acceleration methods using the embedding distortion/efficiency. Finally, the steganographic performance

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.2019.2908689, IEEE Transactions on Circuits and Systems for Video Technology



will be verified by the empirical security via the actual steganalysis including the traditional paradigm and the state-of-the-art network architecture.

In the following experiments, the method using the original cover (length $n$) for message embedding is denoted as **Originality** for short, which is the benchmarking method for performance comparisons. The method using the theoretically optimal shortened cover (top $\theta n$ coefficients of smallest costs) for message embedding is denoted as **Optimality** for short. Since we have proved in Section III that the negative impact of shortening the cover can be minimized by the optimality of cover selection (inaccessible), the performance of **Optimality** is the theoretic upper bound of that of cover selection algorithms. All **Originality**, **Optimality** and cover selection algorithms (**Method-1**, **Method-2**, **Method-3**, **Method-4**) execute ternary STCs with $h = 10$ by default.

### A. *Performance comparison among algorithms*

We will compare the minimal average distortions brought by the proposed cover selection algorithms. And the best algorithm with smallest distortion will be compared with the method of decreasing $h$. UERD [20] is used as the sample distortion function with relative embedding payloads $\alpha = 0.1, 0.3, 0.5$ bpnzac (bit per nonzero AC coefficient). The experimental results are qualified by the average value over the fixed 100 JPEG images, which are randomly selected from BOSSBase 1.01 [32] compressed with quality factor 75.

*1) Comparison of cover selection algorithms on minimal average distortion:* We firstly investigate the impact of the modulus $\beta$ in **Method-4**. As mentioned before, smaller $\beta$ will make the approximative cost closer to the original cost, but meanwhile, more coefficients of small costs may be wetted to hold (22). Just as shown at 0.3bpnzac in Fig. 3, $\beta$ cannot be too large or too small. The optimum range of $\beta$ is 8-10, which can be verified experimentally to work for other payloads. To further demonstrate the concrete effect of $\beta$ on selecting smaller costs, we plot in Fig. 4 the cost-distribution curves of the typical image "1013.jpg" [28] using **Method-4** w.r.t. different $\theta$ and $\beta$ at 0.3bpnzac. It is clear that the cost-distribution curve of $\theta = 1/2$ and $\beta = 10$ is closest to the cost-distribution curve of the original cover, and helps obtain the smallest $E(D)$ computed from (10). For the same $\beta$, the performance of **Method-4** goes worse with the decrease of $\theta$. For the same $\theta$, the performance of $\beta = 10$ is better than that of too large or small $\beta$. Consequently, $\beta = 10$ is used in **Method-4** for the following comparisons.

To study the negative impact of shortening the cover, the performance of **Optimality** and four cover selection algorithms (**Method-1**, **Method-2**, **Method-3**, **Method-4**) is compared with the benchmarking method **Originality** in Fig. 6. Clearly in Fig. 6, when the cover is shortened largely even by $\theta = 1/8$, the $E(D)$ of the theoretically optimal shortened cover (**Optimality**) is still approximately equal to that of the original cover (**Originality**) at all payloads. This indicates that the probability distribution of the selected top smallest costs carries almost all of the entropy $m$ in (3). Just as proved in Section III, the $E(D)$ of **Optimality** is smaller

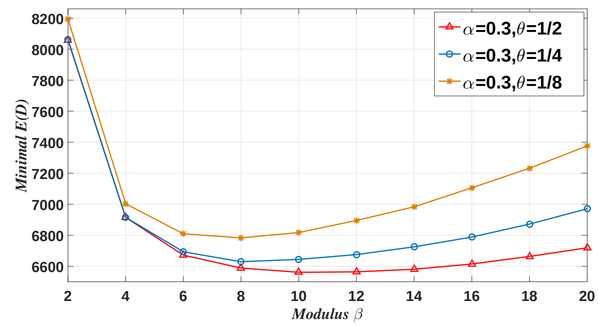



Fig. 3: The impact of modulus $\beta$ on the minimal average distortion for **Method-4** at 0.3bpnzac under 100 images from BOSSBase of QF=75.

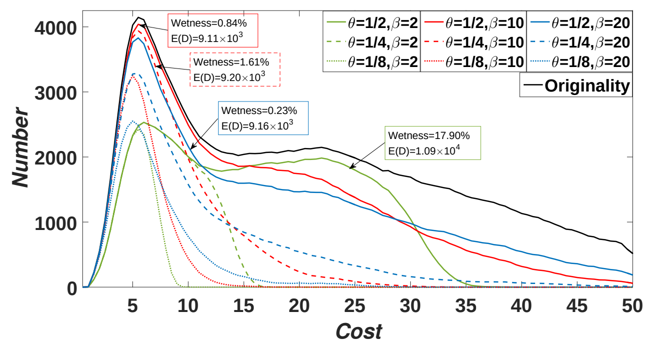



Fig. 4: Cost distributions of **Method-4** w.r.t. shorten-rate $\theta = 1/2, 1/4, 1/8$ and modulus $\beta = 2, 10, 20$ at 0.3bpnzac under "1013.jpg" from BOSSBase of QF=75.

than that of each cover selection algorithm. With the decrease of shorten-rate, the $E(D)$ of each cover selection algorithm is increased, which also verified the negative effect of smaller shorten-rate on the performance of shortening the cover. For the same shorten-rate, the distortion increments to **Originality** are relatively outstanding at larger payloads.

Since STCs aim to approach the theoretical $E(D)$, an algorithm leading to smaller $E(D)$ is better for executing STCs on the corresponding shortened cover. As depicted in

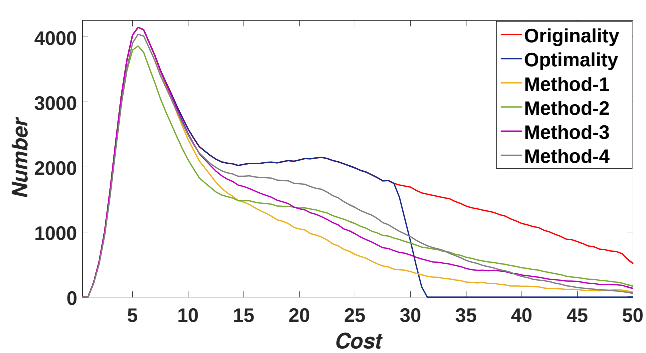
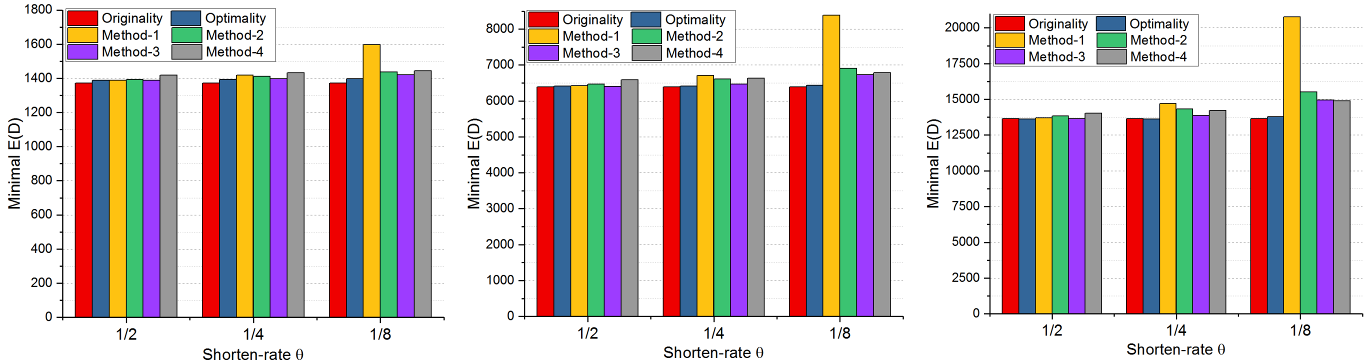
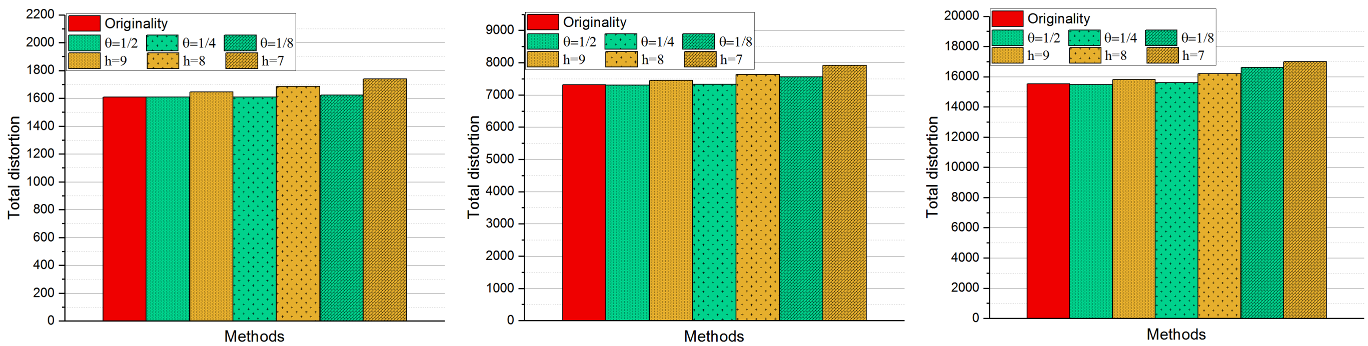



Fig. 5: Cost distributions of cover selection algorithms for shorten-rate $\theta = 1/2$ under "1013.jpg" from BOSSBase of QF=75.




(a) Minimal average distortion at 0.1bpnzac (b) Minimal average distortion at 0.3bpnzac (c) Minimal average distortion at 0.5bpnzac

Fig. 6: Minimal average distortion of cover selection algorithms w.r.t. shorten-rate $\theta = 1/2, 1/4, 1/8$ at payloads $\alpha = 0.1, 0.3, 0.5$ bpnzac under 100 images from BOSSBase of QF=75.




(a) Total distortion at 0.1bpnzac (b) Total distortion at 0.3bpnzac (c) Total distortion at 0.5bpnzac

Fig. 7: Comparisons between **Method-3** and the method of decreasing $h$ on the total distortion when performing ternary STCs under 100 images from BOSSBase of QF=75. Shorten-rates $\theta = 1/2, 1/4, 1/8$ are compared with $h = 9, 8, 7$ respectively at payloads $\alpha = 0.1, 0.3, 0.5$ bpnzac.

TABLE II: Embedding efficiency and its loss of **Method-3** and the method of decreasing $h$ corresponding to Fig. 7. Note that because $e_\theta$ may be slightly larger than $e_{\mathbf{x}}$ due to the effect of randomness and computation precision, the efficiency loss $\mathcal{L}_\theta$ may be a very small negative value, which can be considered as 0 representing a negligible loss.

| Payload | Metric | **Originality** | $\theta = 1/2$ | $h = 9$ | $\theta = 1/4$ | $h = 8$ | $\theta = 1/8$ | $h = 7$ |
|---|---|---|---|---|---|---|---|---|
| 0.1bpnzac | $e$ | 2.1012 | 2.1073 | 2.0663 | 2.1065 | 2.0079 | 2.0892 | 1.9500 |
| | $\mathcal{L}$ | / | -0.29% | 1.66% | -0.25% | 4.44% | 0.57% | 7.20% |
| 0.3bpnzac | $e$ | 1.4115 | 1.4202 | 1.3837 | 1.4108 | 1.3533 | 1.3828 | 1.3059 |
| | $\mathcal{L}$ | / | -0.62% | 1.97% | 0.05% | 4.12% | 2.03% | 7.48% |
| 0.5bpnzac | $e$ | 1.1193 | 1.1350 | 1.0975 | 1.1191 | 1.0729 | 1.0706 | 1.0386 |
| | $\mathcal{L}$ | / | -1.40% | 1.95% | 0.02% | 4.15% | 4.35% | 7.21% |

TABLE I: Average PSNR and SSIM between cover-stego image pairs embedded by **Originality**, **Method-3** ($\theta = 1/8$) and the method of decreasing $h$ ($h = 7$) when performing ternary STCs at 0.5 bpnzac under 100 images from BOSSBase of QF=75.

| Method | **Originality** | $\theta = 1/8$ | $h = 7$ |
|---|---|---|---|
| PSNR(dB) | 46.9350 | 46.8965 | 46.6060 |
| SSIM | 0.9915 | 0.9913 | 0.9910 |

Fig. 6 that among four cover selection algorithms, **Method-3** has the smallest $E(D)$ in all cases of different shorten-rates and payloads. This can be verified in Fig. 5 where **Method-3** does select the maximum number of smallest costs. As shown in TABLE I, the high PSNR (Peak Signal to Noise Ratio) and SSIM (Structural Similarity Index) signify a negligible impact of adaptive steganography on image quality. We also test the robustness of **Method-3** via StirMark [33]. Like other adaptive steganographic schemes [18]–[28] with STCs [16] as the coding method, even weak noise will cause the failure of message extraction. So how to extend the idea of shortening

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.2019.2908689, IEEE Transactions on Circuits and Systems for Video Technology



the cover to robust steganography deserves our research in the future. We now select **Method-3** as the best cover selection algorithm for the following experiments.

*2) Comparison with method of decreasing $h$ on total distortion and time consumption:* Here we compare the method of reducing $n$ with the method of decreasing $h$ in terms of the total embedding distortion/efficiency and execution time of performing ternary STCs [16]. $\theta = 1/2$ vs. $h = 9$, $\theta = 1/4$ vs. $h = 8$, and $\theta = 1/8$ vs. $h = 7$, are three pairs for comparisons because each pair has the theoretical capacity to accelerate the execution of STCs according to the complexity $O(2^h n)$ under the benchmark height $h = 10$.

The total distortions of performing STCs by **Originality**, **Method-3** and the method of decreasing $h$ are denoted by $D_{\mathbf{x}}$, $D_\theta$ and $D_h$ respectively. Their corresponding embedding efficiencies are computed by $e_{\mathbf{x}} = m/D_{\mathbf{x}}$, $e_\theta = m/D_\theta$ and $e_h = m/D_h$ (as in [16]). More intuitively, we define the efficiency loss, $\mathcal{L}_\theta = (e_{\mathbf{x}} - e_\theta)/e_{\mathbf{x}}$ and $\mathcal{L}_h = (e_{\mathbf{x}} - e_h)/e_{\mathbf{x}}$, so as to further measure the lossy degree of the cover selection algorithm and the method of decreasing $h$. As displayed in Fig. 7, for the method of decreasing $h$, the total distortion is rapidly boosted with the decrease of $h$ at all payloads. But for **Method-3** of $\theta = 1/2, 1/4$, the distortions stay the same as that of the original cover at all payloads. These differences are quantized in TABLE II where shortening the cover does not lead to any noticeable loss in embedding efficiency. Although the distortion under $\theta = 1/8$ is gradually enlarged with the increase of payload, the efficiency losses are much smaller than that under $h = 7$. Therefore, the proposed method of shortening the cover is able to better restrain the negative impact caused by speeding up the embedding process.

In TABLE III and Fig. 8, the time consumption of STCs of **Method-3** is proportional to the cover length, i.e., the execution time is cut down strictly by the shorten-rate $\theta$. However, when $h$ decreases by 1, the execution time reduces less than a half, which does not confirm to the complexity $O(2^h n)$. We attribute this phenomenon to the additional time of calculating $\boldsymbol{\pi}_\lambda$ (4) before performing double-layered embedding [16] in ternary STCs. In (4), the computational complexity is related to the cover length, so **Method-3** can also accelerate the process of searching $\lambda$. From the above analysis, the proposed algorithm of shortening the cover displays the superiority on speeding up the embedding process with the minimum negative impact in steganography.

TABLE III: Comparisons between **Method-3** and the method of decreasing $h$ on the average time consumption when performing ternary STCs at 0.3 bpnzac under 100 images from BOSSBase of QF=75.

| Method | Time/s | Ratio on **Originality** | Method | Time/s | Ratio on **Originality** |
|---|---|---|---|---|---|
| **Originality** | 0.757 | / | **Originality** | 0.757 | / |
| $\theta = 1/2$ | 0.375 | 49.5% | $h = 9$ | 0.512 | 67.6% |
| $\theta = 1/4$ | 0.188 | 24.8% | $h = 8$ | 0.392 | 51.7% |
| $\theta = 1/8$ | 0.089 | 11.8% | $h = 7$ | 0.339 | 44.8% |

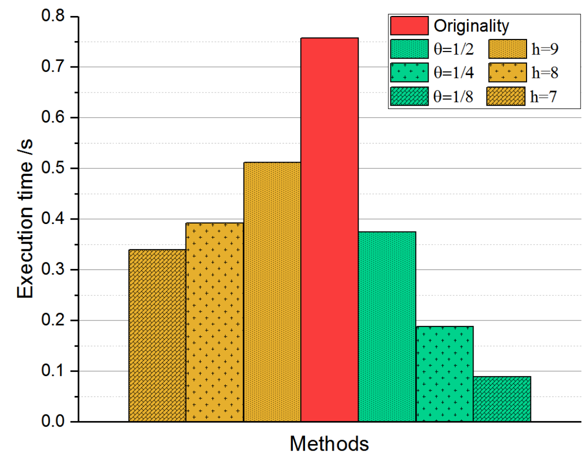



Fig. 8: Illustration of average time consumption corresponding to TABLE III.

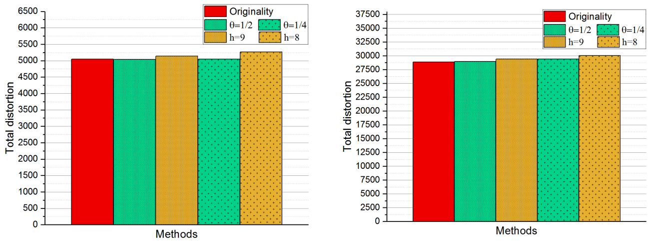
## B. Steganographic security of steganalysis experiments

We now examine the empirical securities of **Method-3** and the method of decreasing $h$ in resisting the detection of actual steganalysis. In addition to the traditional detection paradigm of classifiers [34]–[36] combined with rich models [9]–[14], we also make use of the recent advance of CNN-based network architectures [37]–[43] for verification.

*1) Steganalysis of ensemble classifier with rich features:* Experiments are conducted on BOSSBase 1.01 [32], which contains 10,000 gray-scale images of size $512 \times 512$ pixels. All of the images are compressed into JPEG domain with quality factor QF=50,75,95 respectively, which are then adopted as datasets for experimental comparisons. We use the mainstream distortion functions UERD [20] and J-UNIWARD [18] for message embedding, and the relative embedding payloads range from 0.1 to 0.5 bpnzac (bit per nonzero AC coefficient) with a step of 0.1bpnzac. The steganalyzer is trained by using state-of-the-art DCTR-8,000D [12] and GFR-17,000D [13] with the FLD ensemble [34] by default. The FLD ensemble can minimize 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 $P_{\mathrm{FA}}$ and $P_{\mathrm{MD}}$ are the false-alarm (FA) probability and the missed-detection (MD) probability respectively. The ultimate security is qualified by average error rate $\overline{P_{\mathrm{E}}}$ averaged over 10 random 5000/5000 splits of the dataset, and larger $\overline{P_{\mathrm{E}}}$ means stronger security.

Note that, **Method-3** with too small shorten-rate (e.g., $\theta = 1/8$) would cause failure of performing STCs at too large payload. For ternary embedding, each coefficient can carry $\log_2 3$ bits of message at most, and thus the cover of length $n$ can carry up to $n \log_2 3$ bits of message in theory. When the cover is shortened by $\theta$, the maximum bits of message carried by the shortened cover become $\theta n \log_2 3$, which can be seen as the maximum entropy of arbitrary $\boldsymbol{\mu}$ in (11). Obviously, the optimal $\boldsymbol{\mu}_{\lambda'}$ satisfying (11) exists under the premise of $\theta n \log_2 3 \geq m = \alpha N_{nzac}$ ($N_{nzac}$ is the number of nonzero AC coefficients of the original cover), which can be rewritten as

$$n \log_2 3 \geq \frac{m}{\theta} = \frac{\alpha N_{nzac}}{\theta}. \tag{25}$$




(a) QF=50 (b) QF=95

Fig. 11: Comparisons between **Method-3** and the method of decreasing $h$ on the total distortion at 0.3bpnzac under 100 images from BOSSBase of QF=50,95.

Since small $\theta$, large $\alpha$ or texture-complex image with large $N_{nzac}$ may go against this premise (25), STCs would fail to embed message to some images. To make a fair comparison, we only exhibit the experimental results of successfully embedding on all images of a dataset. Therefore, the securities of **Method-3** with small $\theta = 1/8$ are missing at large $\alpha = 0.4, 0.5$bpnzac in Fig. 9 and at 0.3bpnzac in Fig. 10(b) with large $N_{nzac}$ of QF=95.

As depicted in Fig. 9, the change of steganographic security has the same trend as the change of total distortion in Fig. 7. This verifies the strong correlation between security and distortion that, the stronger steganographic security comes with the smaller total distortion. It is clear that under quality factor 75, the securities have not been weakened by **Method-3** of shorten-rate $\theta = 1/2, 1/4$. And **Method-3** still works for $\theta = 1/8$ at small payloads ($\leq$0.3bpnzac) with slight or negligible security losses. However, the impact of decreasing $h$ on the security should not be ignored. With the decrease of $h$, the security losses are gradually enlarged, and are outstanding by larger than $1\%$ for $h = 8, 7$ at most payloads. Therefore, **Method-3**, which is designed for shortening the cover by selecting smaller costs as much as possible to achieve smaller steganographic distortion, is much securer for accelerating the embedding process.

To verify the universal validity of **Method-3** for JPEG images, we test the steganographic security for other quality factors at payload 0.3bpnzac as shown in Fig. 10. For small quality factor 50, the security advantage of **Method-3** stays the same as that for QF=75. But for large quality factor 95, **Method-3** has the similar securities as the method of decreasing $h$. The reason we conjecture is that larger quality factor results in larger $N_{nzac}$, which enlarges the distortion difference between **Originality** and **Method-3**. Thus the distortion advantage of **Method-3** to the method of decreasing $h$ is becoming weak as shown in Fig. 11. Note that in Fig. 7, the total distortion of **Method-3** also grows faster with the larger payload. So we can infer that too large $m = \alpha N_{nzac}$ in (25) may have a bad effect on the performance of **Method-3**.

*2) Steganalysis of deep residual network:* We select the state-of-the-art steganalysis network architecture, called SRNet [43], to further verify the security advantage of **Method-3** in comparison with the method of decreasing $h$. Experiments are conducted on the union of BOSSBase 1.01 [19] and BOWS2 [44], each containing 10,000 grayscale images resized from original size $512 \times 512$ to $256 \times 256$ and JPEG compressed with quality factor 75. Randomly chosen 14,000 images from the total 20,000 JPEG images are used for training with 1,000 images aside for validation. And the remaining 5,000 images are used for testing. We also use UERD [20] and J-UNIWARD [18] for message embedding at payload 0.3bpnzac. The SRNet is run on a single GPU of NVIDIA GTX 1080Ti in our experiments, and is trained in default settings as described in [43], except that the Adamax [45] is used with minibatches of 8 cover-stego pairs (instead of 16 pairs [43]) and therefore the training is run for 800k iterations (instead of 400k [43]). Similarly, the ultimate security is measured with the total classification error probability on the testing set under equal priors $P_{\mathbf{E}} = \min_{P_{\text{FA}}} \frac{1}{2}(P_{\text{FA}} + P_{\text{MD}})$ where $P_{\text{FA}}$ and $P_{\text{MD}}$ are the false-alarm (FA) probability and the missed-detection (MD) probability respectively.

TABLE IV: Detection errors of **Method-3** ($\theta = 1/4$) and the method of decreasing $h$ ($h = 8$) on UERD and J-UNIWARD against SRNet at 0.3bpnzac under BOSSBase+BOWS2 of QF=75.

| Method | **Originality** | $\theta=1/4$ | $h=8$ |
|---|---|---|---|
| UERD | 0.0808 | 0.0820 | 0.0788 |
| J-UNIWARD | 0.1337 | 0.1329 | 0.1257 |

As verified in TABLE IV, the empirical security in resisting the detection of SRNet is not weakened by shortening the cover with **Method-3**. And **Method-3** is more secure than the method of decreasing $h$, which stays consistent with the aforementioned conclusion of resisting the traditional steganalysis paradigm.

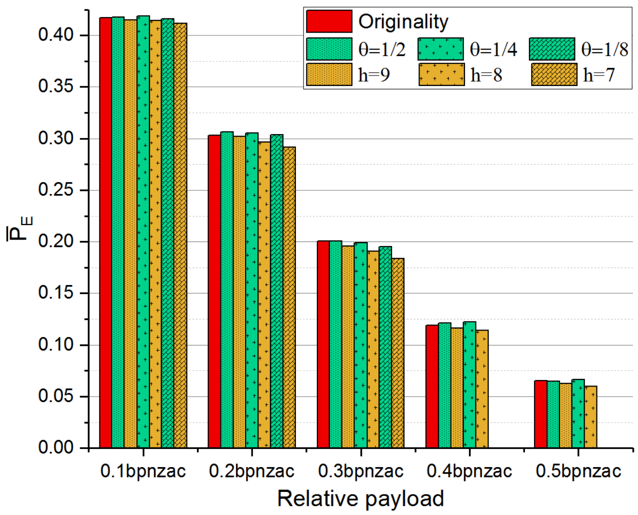
From the above experimental results and analysis, we are confident that a JPEG image of arbitrary quality factor can be safely shortened as 1/4 of the original by **Method-3** to speed up the execution of STCs. With preserving the steganographic security, **Method-3** runs much faster for message embedding than the method of decreasing $h$. We believe from (25) that, when the message is short enough, the shorten-rate could be safely even smaller, and thus the execution of STCs can be even faster.

## VI. Conclusion

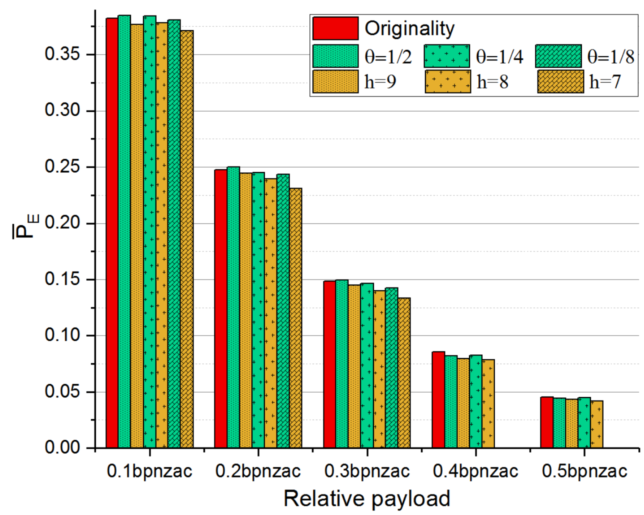
The execution speed of STCs is an important issue for the real-world steganography. In this paper, we investigate the possibility of shortening the cover for accelerating the embedding process without weakening the steganographic security. Since the cover selection for constructing a shortened cover will increase the steganographic distortion, we attempt to minimize this negative impact by selecting DCT coefficients of smaller costs as much as possible. When compared with the method of decreasing $h$, the proposed algorithm of shortening the cover shows superiority not only on the speed profit but also the steganographic security. With confidence, a JPEG image of arbitrary quality factor can be safely shortened as

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.2019.2908689, IEEE Transactions on Circuits and Systems for Video Technology

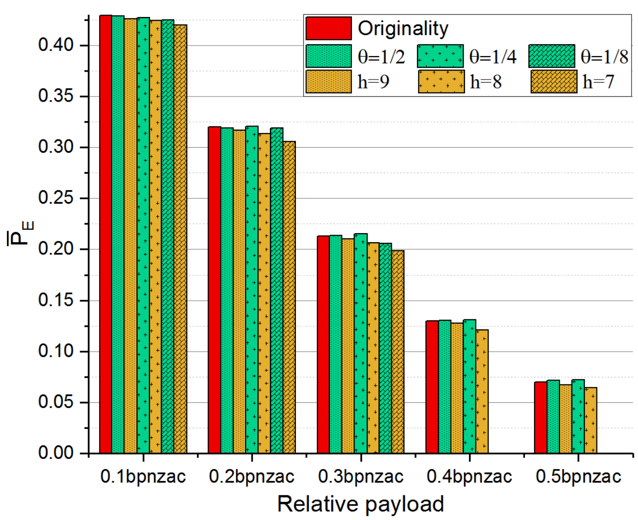




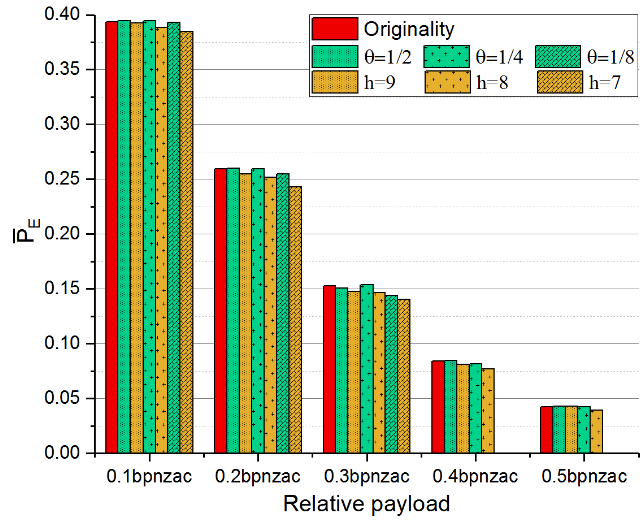

(a) UERD against DCTR




(b) UERD against GFR

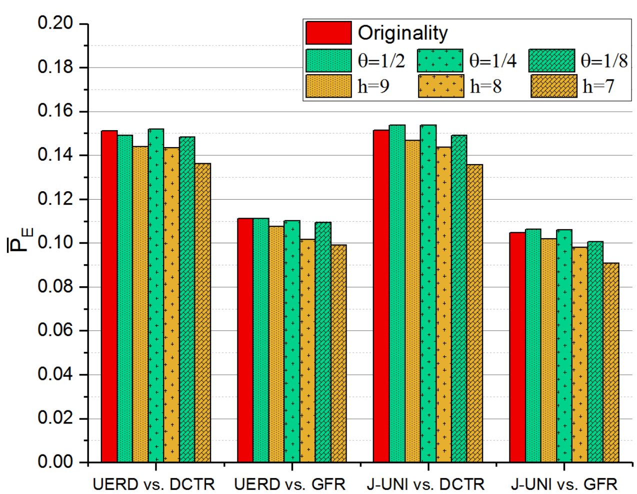



(c) J-UNIWARD against DCTR

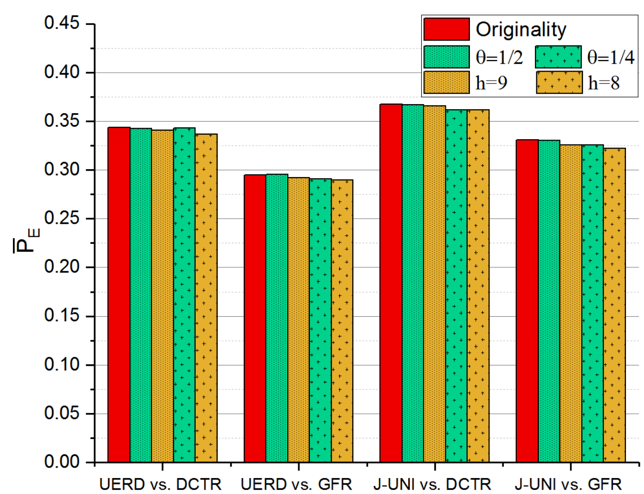



(d) J-UNIWARD against GFR

Fig. 9: Detection errors of **Method-3** and the method of decreasing $h$ on UERD and J-UNIWARD against DCTR and GFR under BOSSBase of QF=75.




(a) QF=50




(b) QF=95

Fig. 10: Detection errors of **Method-3** and the method of decreasing $h$ on UERD and J-UNIWARD against DCTR and GFR at 0.3bpnzac under BOSSBase of QF=50,95.

1/4 of the original, and correspondingly the execution of STCs can be 4 times faster.

Although the proposed cover selection algorithm is designed for reducing the execution time of STCs, it is also suitable for the memory-limited scenario in real-world application. Moreover, like the additive distortion steganography in this paper, the proposed algorithm can also be applied to non-additive distortion steganography [28] or batch steganography [46] for the popular JPEG images where STCs are used to embed message. Note that STCs may not be the only choice for adaptive steganography in the future, we believe that the proposed algorithm can also work for other steganographic coding methods.

Since the relationship between steganographic distortion and security is far from clear [16], it is difficult to precisely measure the impact of cover selection on the steganographic security. In future work, we will investigate the secure payload of shortening the cover. Also, generalizing the idea of shortening the cover to robust steganography and other media objects, e.g., spatial images, audio and video, is an important and interesting issue.

## 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/). Specifically, they are grateful to the AE Prof. Gian Luca Foresti and the anonymous reviewers for their valuable comments and helpful suggestions.

## References

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

[2] B. Li, J. He, J. Huang, and Y. Q. Shi, "A survey on image steganography and steganalysis," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, no. 2, pp. 142–172, 2011.

[3] G. Paul, I. Davidson, I. Mukherjee, and S. Ravi, "Keyless dynamic optimal multi-bit image steganography using energetic pixels," *Multimedia Tools and Applications*, vol. 76, no. 5, pp. 7445–7471, 2017.

[4] N. Mukherjee, G. Paul, and S. K. Saha, "An efficient multi-bit steganography algorithm in spatial domain with two-layer security," *Multimedia Tools and Applications*, vol. 77, no. 14, pp. 18 451–18 481, 2018.

[5] B. Datta, S. Roy, S. Roy, and S. K. Bandyopadhyay, "Multi-bit robust image steganography based on modular arithmetic," *Multimedia Tools and Applications*, pp. 1–36, 2018.

[6] Y. Zhang, C. Qin, W. Zhang, F. Liu, and X. Luo, "On the fault-tolerant performance for a class of robust image steganography," *Signal Processing*, vol. 146, pp. 99–111, 2018.

[7] J. Tao, S. Li, X. Zhang, and Z. Wang, "Towards robust image steganography," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 2, pp. 594–600, 2019.

[8] Z. Zhao, Q. Guan, H. Zhang, and X. Zhao, "Improving the robustness of adaptive steganographic algorithms based on transport channel matching," *IEEE Transactions on Information Forensics and Security*, 2018.

[9] J. Fridrich and J. Kodovsky, "Rich models for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 868–882, 2012.

[10] J. Kodovskỳ and J. Fridrich, "Steganalysis of jpeg images using rich models," in *Media Watermarking, Security, and Forensics 2012*, vol. 8303. International Society for Optics and Photonics, 2012, p. 83030A.

[11] L. Chen, Y. Q. Shi, P. Sutthiwan, and X. Niu, "A novel mapping scheme for steganalysis," in *The International Workshop on Digital Forensics and Watermarking 2012*. Springer, 2013, pp. 19–33.

[12] 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.

[13] 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.

[14] 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.

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

[16] 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.

[17] R. Crandall, "Some notes on steganography," *Posted on steganography mailing list*, pp. 1–6, 1998.

[18] 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.

[19] L. Guo, J. Ni, and Y. Q. Shi, "Uniform embedding for efficient jpeg steganography," *IEEE transactions on Information Forensics and Security*, vol. 9, no. 5, pp. 814–825, 2014.

[20] 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.

[21] Y. Pan, J. Ni, and W. Su, "Improved uniform embedding for efficient jpeg steganography," in *International Conference on Cloud Computing and Security*. Springer, 2016, pp. 125–133.

[22] Z. Wang, X. Zhang, and Z. Yin, "Hybrid distortion function for jpeg steganography," *Journal of Electronic Imaging*, vol. 25, no. 5, pp. 050 501–050 501, 2016.

[23] Q. Wei, Z. Yin, Z. Wang, and X. Zhang, "Distortion function based on residual blocks for jpeg steganography," *Multimedia Tools and Applications*, pp. 1–14, 2017.

[24] 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.

[25] 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.

[26] 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.

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

[28] 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.

[29] N. Provos, "Defending against statistical steganalysis." in *Usenix security symposium*, vol. 10, 2001, pp. 323–336.

[30] G. Winkler, *Image analysis, random fields and Markov chain Monte Carlo methods: a mathematical introduction*. Springer Science & Business Media, 2012, vol. 27.

[31] J. Fridrich, P. Lisoněk, and D. Soukal, "On steganographic embedding efficiency," in *International Workshop on Information Hiding*. Springer, 2006, pp. 282–296.

[32] 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.

[33] F. A. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Attacks on copyright marking systems," in *International workshop on information hiding*. Springer, 1998, pp. 218–238, available: https://www.petitcolas.net/watermarking/stirmark/.

[34] 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.

[35] R. Cogranne, V. Sedighi, J. Fridrich, and T. Pevnỳ, "Is ensemble classifier needed for steganalysis in high-dimensional feature spaces?" in *Information Forensics and Security (WIFS), 2015 IEEE International Workshop on*. IEEE, 2015, pp. 1–6.

[36] T. Pevny and A. Ker, "Towards dependable steganalysis," in *IS&T/SPIE Electronic Imaging*, 2015.

[37] Y. Qian, J. Dong, W. Wang, and T. Tan, "Deep learning for steganalysis via convolutional neural networks," in *Media Watermarking, Security, and Forensics 2015*, vol. 9409. International Society for Optics and Photonics, 2015, p. 94090J.

[38] G. Xu, H.-Z. Wu, and Y. Q. Shi, "Ensemble of cnns for steganalysis: An empirical study," in *proceedings of the 4th ACM workshop on information Hiding and Multimedia security*. ACM, 2016, pp. 103–107.

[39] G. Xu, H.-Z. Wu, and Y.-Q. Shi, "Structural design of convolutional neural networks for steganalysis," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 708–712, 2016.

[40] J. Ye, J. Ni, and Y. Yi, "Deep learning hierarchical representations for image steganalysis," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 11, pp. 2545–2557, 2017.

[41] G. Xu, "Deep convolutional neural network to detect j-uniward," in *Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security*. ACM, 2017, pp. 67–73.

[42] J. Zeng, S. Tan, B. Li, and J. Huang, "Large-scale jpeg image steganalysis using hybrid deep-learning framework," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 5, pp. 1200–1214, 2018.

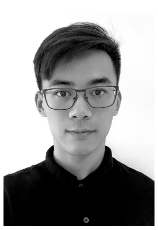
[43] M. Boroumand, M. Chen, and J. Fridrich, "Deep residual network for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, 2018.

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

[45] D. Kinga and J. B. Adam, "A method for stochastic optimization," in *International Conference on Learning Representations (ICLR)*, vol. 5, 2015.

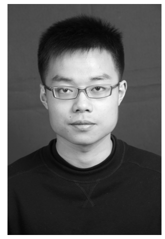
[46] A. D. Ker and T. Pevnỳ, "The steganographer is the outlier: realistic large-scale steganalysis," *IEEE Transactions on information forensics and security*, vol. 9, no. 9, pp. 1424–1435, 2014.


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



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

**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.



**Chuan Qin** received his B.S. degree from Northwest University, Xi'an, China, in 2016. He is currently pursuing the Ph.D. degree with University of Science and Technology of China. His research interests include steganography, steganalysis and adversarial examples.

**Huanhuan Hu** received his B.S. degree in 2015 and M.S. degree in 2018 from University of Science and Technology of China. His research interests includes information hiding, multimedia security and privacy protection.

**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.