Available online at www.sciencedirect.com

ScienceDirect

Signal Processing 87 (2007) 1251–1263

SIGNAL PROCESSING

www.elsevier.com/locate/sigpro

A DCT-based Mod4 steganographic method

KokSheik Wong[a,*], Xiaojun Qi[b], Kiyoshi Tanaka[a]

[a]*Faculty of Engineering, Department of Electrical and Electronics Engineering, Shinshu University, 4-17-1 Wakasato, Nagano 380-8553, Japan*

[b]*Department of Computer Science, Utah State University, 84322 Logan, Utah, USA*

Received 9 February 2006; received in revised form 23 October 2006; accepted 26 October 2006
Available online 12 December 2006

Abstract

This paper presents a novel Mod4 steganographic method in discrete cosine transform (DCT) domain. Mod4 is a blind steganographic method. A group of 2×2 spatially adjacent quantized DCT coefficients (GQC) is selected as the valid message carrier. The modulus 4 arithmetic operation is then applied to the valid GQC to embed a pair of bits. When modification is required for data embedding, the shortest route modification scheme is applied to reduce distortion as compared to the ordinary direct modification scheme. Mod4 is capable in embedding information into both uncompressed and JPEG-compressed image. To compare Mod4 with other existing methods, carrier capacity, stego image quality, and results of blind steganalysis for 500 various images are shown. Visual comparison of three additional metrics is also presented to show the relative performance of Mod4 among other existing methods.

Keywords: Steganography; GQC; vGQC; Shortest route modification; JPEG; Modulus arithmetic; Steganalysis

1. Introduction

Steganography is a data hiding technique that has been mainly used in information security applications. It is similar to watermarking and cryptography techniques, but these three techniques are different in some aspects. Firstly, watermarking mainly tracks illegal copies or claims of the ownership of digital media. It is not geared for communication. Secondly, cryptography scrambles the data with the mixture of permutation(s) and substitution(s) so that unintended receivers cannot perceive the processed information. However, the fact that information has been embedded into a medium (i.e., watermarking) and communication has been carried out (i.e., cryptography) is known to everyone, or at least it is acceptable to reveal such a fact. Finally, steganography transmits information by embedding messages into innocuous-looking cover objects, such as digital images, to conceal the very existence of communication. As a result, steganography is the art and science of data smuggling since its goal is to hide the presence of communication [1].

Despite the fact that steganography has been notoriously used in spying and evil plots, there are many friendly applications of steganography. For instance, a photographer may record the aperture size, shuttle speed, and other settings for future references when capturing a picture. A person may

*Corresponding author. Tel.: +81 26 269 5236; fax: +81 26 269 5220.

E-mail addresses: koksheik@shinshu-u.ac.jp (K. Wong), xqi@cc.usu.edu (X. Qi), ktanaka@shinshu-u.ac.jp (K. Tanaka).

0165-1684/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2006.10.014

want to embed a file containing private information into a multimedia file to save storage space. Other applications include signature, authentication, history recording and so on.

For such reasons, many imagery steganographic methods have been invented. Here we briefly review research carried out particularly in the DCT domain. Upham invented JSteg that hides information sequentially in LSBs of the quantized DCT coefficient*s* (qDCTCs) while skipping 0's and 1's [2]. Provos developed OutGuess to scatter information into the LSB of qDCTCs [3]. Embedding is followed by a correction procedure to ensure that the distributions of any related pair of the qDCTCs are unchanged. Westfeld employs the technique of matrix encoding to hold secret information using LSB of qDCTCs in F5 [4]. Whenever a modification is needed, the magnitude of a coefficient is decremented.

Sallee proposed a model-based steganographic scheme that treats a cover medium as a random variable that obeys some parametric distribution such as Cauchy or Gaussian [5]. The medium is divided into two parts, i.e., the deterministic, and the indeterministic part where the secret message is embedded. Miyake et al. define diagonal bands for each 8×8 DCT block, and the number of zeros (in a zero sequence) in each band are modified to embed exactly one bit [6]. If the frequency of zeros within a zero sequence in a band is odd, that band stores the message bit 1, and vice versa. Seki et al. apply modulus arithmetic on the sum of all 64 qDCTCs from a block to embed information [7]. The coefficient(s) is (are) modified by exploiting the quantization errors so that the sum yields the desired remainder when divided by a fixed number d.

On the other hand, along with the evolution of steganographic methods, many steganalysis methods are invented. Westfeld and Pfitzmann invented the χ^2-statistical test to detect the message embedded sequentially into a cover medium by LSB flipping methods [8]. Fridrich et al. employ a macroscopic measure, i.e., increment of blockiness, to determine the length of the embedded message [9]. This measure detects the stego of OutGuess [3]. However, this macroscopic measure is limited to LSB flipping approaches. Extending their work to non-LSB flipping methods, Fridrich et al. successfully detect the message embedded using F5 [4] by exploiting the awkward concentration of zero coefficients in the distribution of qDCTCs [10].

While the above steganalysis methods target at some specific steganographic methods, Farid invented a blind steganalysis method that detects stegos regardless of the embedding algorithm in use [11]. This steganalyzer employs features extracted from the image after a series of filterings, and the errors collected from an optimal linear predictor. This method is recently extended to non-linear support vector machine to classify a given image with higher accuracies [12]. Avcibas et al. invented a blind steganalysis method utilizing various image quality metrics as the features [13]. This method is based on the observation that an embedded and filtered image differs statistically from a non-embedded but simply filtered image. Recently, Fridrich proposed an effective feature-based steganalysis method for JPEG images [14]. This classifier uses 23 features that are possibilly altered during data embedding. It detects the stego generated by JP Hide & Seek [15], OutGuess [3], F5 [4] and model-based steganography [5]. This method is further extended to include the ability to associate a detected stego with a known steganographic method [16].

In this paper, we propose a novel DCT-based Mod4 steganographic method for still images. It extends our previous work [17], which presents preliminary results, by adding more flexible constraints on a set of parameters and comparing with existing methods using six metrics and the state-of-the-art blind steganalysis method [14]. Our work is presented with the passive warden scenario [1] and assumes that there is no channel noise during data transmission. The remainder of the paper is organized as follows. Section 2 presents the Mod4 steganographic method. Section 3 discusses the features of the proposed method. Section 4 shows the experimental results to compare Mod4 with some selected existing steganographic methods in DCT domain. Visual comparison is presented in Section 5 to show the relative performance of Mod4 among other existing methods. Finally, conclusions are given in Section 6.

2. The proposed DCT-based Mod4 steganographic method

2.1. The embedding scheme

The block diagram of the proposed method is shown in Fig. 1, where we consider the JPEG compressed image for presentation purpose. In the

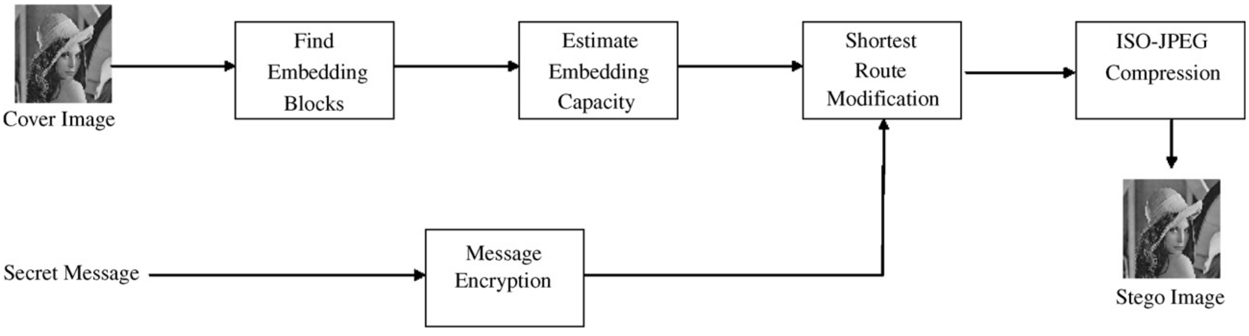

Fig. 1. Block diagram of Mod4 embedding.

following subsections, we discuss each block in detail.

2.1.1. Find embedding blocks

The 8 × 8 blocks of qDCTCs (each denoted by F_{QT}) are decompressed and extracted from the JPEG data stream. Each F_{QT} is then divided into 16 groups of 2 × 2 spatially adjacent quantized DCT coefficients (GQC). A GQC (which consists of four elements) is characterized as a valid GQC (vGQC) if it satisfies the following conditions:

$$|\{x : x \in GQC,\ x > \phi_1\}| \geqslant \tau_1,$$
$$\text{and}\quad |\{x : x \in GQC,\ x < -\phi_2\}| \geqslant \tau_2, \qquad (1)$$

where $|X|$ denotes the cardinality of the set X. ϕ_1 and ϕ_2 are magnitudes that govern the coefficients to be modified. τ_1 and τ_2 are the positive thresholds indicating if the candidate GQC is a vGQC.

2.1.2. Estimation of embedding capacity

The vGQCs are extracted from the F_{QT}s, and loaded into an empty dynamic array β in the order determined by a secret key $\mathscr{K}$. This process randomizes the order of extraction of vGQCs and thus makes Mod4 satisfy the Kerckhoff's principle [18]. When all vQGCs are exhausted, the length of β determines the maximum embedding capacity Ω of the cover image. In particular, Ω depends on the cover image, the JPEG quality factor, two magnitudes ϕ_1 and ϕ_2, and two positive thresholds τ_1 and τ_2. Ω is twice of the number of the vGQCs because each vGQC holds exactly two bits.

2.1.3. Shortest route modification

Mod4 embeds pairs of message bits into the ordered vGQCs in β. Denote the ith vGQC from β by B_i, and the ith pair of message bits from the secret message by $xy_i \in \{00, 01, 10, 11\}$. The shortest route modification (SRM) is employed to embed each xy_i into the corresponding B_i. Unlike the ordinary direct modification method, SRM leads to lower expected number of modification on the coefficients (clarified in 3.3).

The embedding scheme constrained by SRM is illustrated in Table 1, where we show a representative example to embed $xy_i = 00$. The first column $\mathrm{mod}(b_i, 4)$ indicates the remainder of the sum of the qDCTCs in B_i (i.e., $b_i := \sum_{z \in B_i} z$) divided by 4.

To ease the discussion, we define the following sets:

$$P := \{p : p \in B,\ p > \phi_1\} \quad \text{and}$$
$$N := \{n : n \in B,\ n < -\phi_2\} \qquad (2)$$

and relabel the elements in P and N so that[1]

$$p_1 \geqslant p_2 \geqslant p_3 \geqslant p_4 \qquad (3)$$

and

$$|n_1| \geqslant |n_2| \geqslant |n_3| \geqslant |n_4|. \qquad (4)$$

It is obvious that

$$\tau_1 + \tau_2 \leqslant |N| + |P| \leqslant 4. \qquad (5)$$

The second column $\oplus$ records the number to be added to the positive coefficients p_j to achieve $\mathrm{mod}(b'_i, 4) = 00$, where b'_i denotes the new sum after modification(s). Similarly, the third column $\ominus$ indicates the number to be subtracted from the negative coefficients n_j to achieve $\mathrm{mod}(b'_i, 4) = 00$.

SRM determines the route for modification(s) as follows:

(1) If $\oplus = \ominus = 0$, no changes are made. Done !
(2) If $\oplus > \ominus$, subtracting 1 from n_1 is the shortest subtraction route.
(3) If $\oplus < \ominus$, adding 1 to p_1 is the shortest addition route.

[1] Some of the p_i's and n_j's might not exist, depending on the values of ϕ_1 and ϕ_2.

(4) If $\oplus = \ominus = 2$, there are six cases to consider:

Case 1: $(|P|,|N|) = (1,1)$. Only p_1 and n_1 exist. If $p_1 \geqslant |n_1|$, $p_1' = p_1 + 2$. Otherwise, $n_1' = n_1 - 2$.

Case 2: $(|P|,|N|) = (2,1)$. p_1, p_2, and n_1 exist. Let $p_1' = p_1 + 1$ and $p_2' = p_2 + 1$. n_1 is left unmodified.

Case 3: $(|P|,|N|) = (2,2)$. p_1, p_2, n_1, and n_2 exist. If $p_1 \geqslant |n_1|$, $p_1' = p_1 + 1$ and $p_2' = p_2 + 1$. Otherwise, $n_1' = n_1 - 1$ and $n_2' = n_2 - 1$.

Case 4: $(|P|,|N|) = (3,1)$. p_1, p_2, p_3, and n_1 exist. Let $p_1' = p_1 + 1$ and $p_2' = p_2 + 1$. p_3 and n_1 are left unmodified.

Case 5: $(|P|,|N|) = (1,2)$. p_1, n_1, and n_2 exist. Let $n_1' = n_1 - 1$ and $n_2' = n_2 - 1$. p_1 is left unmodified.

Case 6: $(|P|,|N|) = (1,3)$. p_1, n_1, n_2, and n_3 exist. Let $n_1' = n_1 - 1$ and $n_2' = n_2 - 1$. p_1 and n_3 are left unmodified.

We made such choices based on two observations:

(1) Each coefficient should undergo the least number of modifications possible. For that, in case of $|P| > |N|$, we deal with p_1 and p_2 instead of n_1 even if $|n_1| > p_1$.

(2) Always modify the coefficients with larger magnitudes first. For that, in case of $|P| = |N|$, we deal with the positive coefficients if $p_1 > |n_1|$, and vice versa. The logic behind this particular choice of modification is presented in 3.2.

The extensions to $xy_i = 01, 10$, and 11 could be similarly derived. The modified vGQC B_i' is reinjected into the image where it is originally extracted. This process is repeated until all pairs of message bits are considered.

Table 1
Modification scheme — shortest route for $xy_i = 00$

mod(b_i, 4)	$\oplus$	$\ominus$	Possible routes	Shortest route
00	0	0	No change	N/A
01	3	1	−1 or +3	−1
10	2	2	+2 or −2	Conditional
11	1	3	+1 or −3	+1

2.1.4. ISO–JPEG compression

The modified B_i''s and GQCs (with location association) form blocks F_{QT}'. Each F_{QT}' then undergoes the same zigzag scan, and differential and run length coding as in the JPEG compression scheme [19]. The resulting steganogram is transmitted to the intended receiver as a JPEG image.

2.2. The extraction scheme

The extraction process can be carried out by reversing the embedding procedure and the block diagram is shown in Fig. 2. From this figure, it is obvious that Mod4 is blind. The receiver first identifies all vGQCs using ϕ_1, ϕ_2, τ_1, and τ_2. He/she then extracts B_i''s from F_{QT}'s, and sequentially orders them using the same secret key $\mathscr{K}$. Extraction of the embedded message is just a series of summations of qDCTCs in vGQCs and divisions by 4. This task could be efficiently accomplished with a 4-state machine, having exactly 4 edges (00, 01, 10, 11) originating from each state.

3. Features of Mod4

3.1. Coefficient distribution and migration

Since Mod4 is a non-LSB flipping-based steganographic method, it can withstand χ^2-statistical-based tests [8]. With the construction of SRM, it is obvious that the magnitude of a coefficient is incremented when modification is required. Thus the Laplacian distribution property of the AC-components is preserved. Also, coefficients in $[-\phi_2, \phi_1]$ are left unmodified, hence leading to the undetectability by histogram-based steganalysis that seeks for awkward concentration of coefficients at low value histogram bins [10].

Elaborating the aforementioned point, set $\phi_1 = \phi_2 = 1$. The discussion is still valid for $\phi_i > 1$. Let $h_k(i,j)$, $i,j \in \{0,1,\dots,7\}$, denote the count of

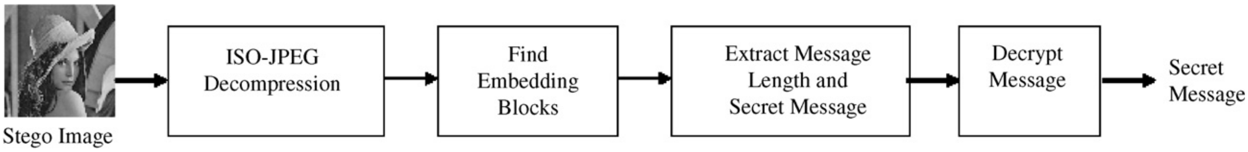

Fig. 2. Block diagram of Mod4 extraction.

the (i,j)-mode qDCTCs with value k, and let $h_k := \sum_i \sum_j h_k(i,j)$. We show the distribution of the (1,2)th AC component as the representative component for a natural image in Fig. 3. As observed in [20–22], $h_{-1}(1,2), h_0(1,2)$, and $h_1(1,2)$ outnumber the rest of the counts $h_k(1,2)$. If qDCTCs in $[-1,1]$ are employed for modifications, it will leave traces of modification, such as value h_1 will be far too low as compared to h_0 for an JPEG image. Such reasoning applies to each individual component $h_k(i,j)$. For this reason, the qDCTCs in $[-\phi_2, \phi_1]$ are left unmodified in the proposed Mod4.

3.2. Modify larger coefficients first

Larger coefficients are modified first based on the following three observations made from JPEG compression scheme: (a) The values in a quantization table (QT) increase from the upper left to the bottom right corner; (b) A qDCTC with a relatively large magnitude (compared with the average value of the whole block) corresponds to a small divisor in QT; (c) $h_{|k|}$ is inversely proportional to $|k|$, i.e., $h_{|k|} \propto |k|^{-1}$.

Such modification rule does not only ensure the minimum distortion in the resultant stego images but also guarantee that first-order histogram-based steganalysis tools fail to find any abnormal statistical property for detecting the embedded message. The rationales are further elaborated:

First, suppose that we have two qDCTCs c_1 and c_2 such that $c_1 < c_2$. For a natural image, their corresponding QT divisors d_1 and d_2 generally satisfy the condition of $d_1 > d_2$ based on the second observation (b). Consequently, the difference between the original and modified de-quantized coefficients is smaller for c_2, if c_1 and c_2 undergo the same number of modifications during the embedding phase. That is:

$$d_1|c_1 - c_1'| > d_2|c_2 - c_2'|, \tag{6}$$

where c_1' and c_2' are the newly modified quantized DCT values for c_1 and c_2, respectively. This inequality ensures the overall difference before and after embedding in each vGQC is the minimum if the coefficients with the largest magnitude is modified first. This difference can be expressed as

$$\|(F_{\mathrm{QT}} * \mathrm{QT}) - (F'_{\mathrm{QT}} * \mathrm{QT})\|_1, \tag{7}$$

where $\|\cdot\|_1$ denotes a matrix norm and $*$ denotes the element-wise multiplication operation. Thus, the minimal difference as introduced by our modification rule ensures less distortion in the resulting stego image.

Second, modifying the coefficients with the largest magnitudes first maintains certain statistical properties, such as: (a) Laplacian distribution of the AC-components; (b) No concentration of counts for bins labeled with $-\phi_2, \ldots, 0, \ldots, \phi_1$ since no changes occur at these bins; (c) It does not cause spike phenomena (see 4.4.1) since the bin counts always maintain their relative order, i.e., $h'_k > h'_{k+1}, k > 0$ and $h'_k > h'_{k-1}, k < 0$.

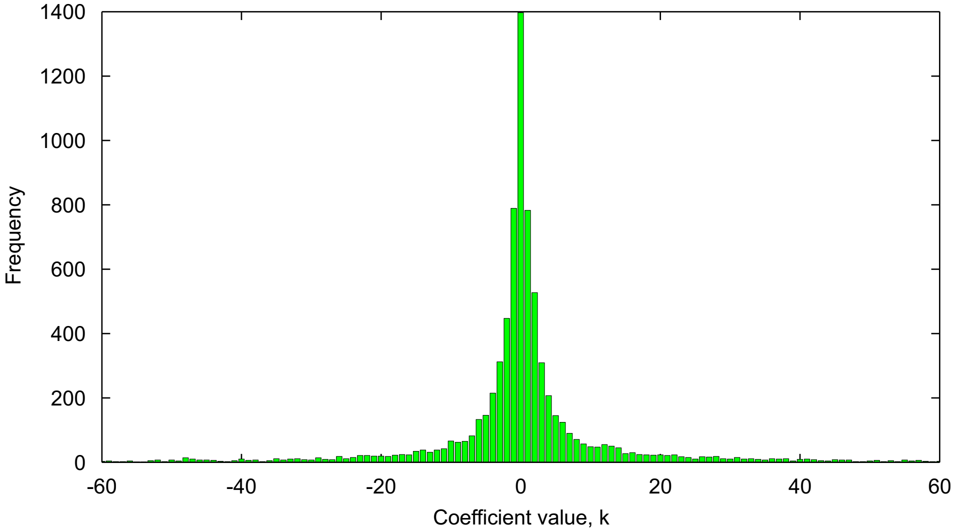

Fig. 3. Distribution of the (1,2)th AC component.

Table 2
Expected number of modifications

Condition	$\oplus$ Route	$\ominus$ Route	SRM
$\text{mod}(b,4)_2 = xy$	0	0	0
$\text{mod}(b,4)_2 = xy$	3	1	1
$\text{mod}(b,4)_2 = xy$	2	2	2
$\text{mod}(b,4)_2 = xy$	1	3	1

These statistical properties ensure that our proposed modification rule leads to the difficulties of steganalyzers in finding abnormalities.

3.3. Minimum number of modifications

Mod4 leads to minimum number of modifications by distributing a pair of bits among a group of qDCTCs within a vGQC. Here, we briefly prove that our expected number of modifications for one single embedded bit is less than 0.5 per coefficient.

Let π be the expected number of modifications that occurs within a vGQC while embedding a single message bit. Table 2 summarizes the three possible routes to modify the qDCTCs for embedding a pair of message bits xy. When the shortest route is not considered (e.g., only the $\oplus$ route is employed), the total changes is 6, and π is be calculated as

$$\pi = \pi_{\oplus} = \pi_{\ominus} = (0 + 1 + 2 + 3) \times 0.25 \div 2 = 0.75. \tag{8}$$

The multiplication of 0.25 is carried out because each case occurs with an equal probability, and the division by 2 is due to the fact that a pair of two bits is distributed into a vGQC.

Now consider SRM as presented in the same table. The numbers listed under the last column (i.e., SRM) equals to the number of changes for a valid embedding. For the SRM case, π is calculated to be:

$$\pi = (0 + 1 + 2 + 1) \times 0.25 \div 2 = 0.5. \tag{9}$$

SRM outperforms the ordinary direct embedding scheme by having a lower expected number of modifications, and the improvement is 0.25 for a single vGQC. Last but most importantly, when SRM is employed, the expected number of modifications on a single coefficient is obviously at most 0.5 since $1 \leqslant \tau_1 + \tau_2 \leqslant 4$.[2]

3.4. Adaptivity

The message carriers (i.e., vGQCs) are generally associated with the noisy regions in the cover image since high variation in pixel intensity values may lead to large high frequency components, i.e., large qDCTCs in GQCs. The chance of a GQC being a vGQC in a noisy block increases, leading to higher embedding capacity. Four adjustable parameters ϕ_1, ϕ_2, τ_1, and τ_2 also control the depth of the adaptivity. In general, the following conditions are imposed:

$$\phi_1, \phi_2 > 0 \quad \text{and} \quad 0 < \tau_1, \tau_2 < 3. \tag{10}$$

On the other hand, we can relax the adjacency condition to construct a GQC in any manner. However, it becomes difficult to associate large qDCTCs and noisy spatial blocks. By doing so, the adaptive property of Mod4 is lost.

4. Experimental results and discussions

4.1. Preparation

To compare Mod4 with the existing methods in terms of embedding capacity, image quality, and robustness against blind steganalysis, a database of 500 images is generated with Sony DSC-828 digital camera. Each image is converted to grayscale, and resized to 800 × 600 pixels using bicubic interpolation. These images are used as uncompressed original cover images. The considered methods are OutGuess [3], F5 [4], model-based steganography [5], and ZeroSequence [6], respectively referred as $OG, F5, MB$, and ZS for the rest of this paper. Let $M4(\phi_1, \phi_2, \tau_1, \tau_2)$ denote the proposed method operating with parameters as specified. Unless specified otherwise, the JPEG quality factor is set to 80.

4.2. Embedding capacity

To ease the discussion, let $\Omega(\varepsilon, A_k)$ denote maximum embedding capacity of image A_k with respect to method ε in terms of bits embedded per non-zero qDCTC (*bpc*) [14]. Fig. 4 shows the graph of $\Omega(\varepsilon, A_k)$ versus image number k. Here, we cropped the y-axis of the graph to the range of $[0, 1.6]\,bpc$ for better display since most images exhibit embedding capacity in this range. From this result, ZS has the highest embedding capacity by far, followed by MB, $F5$, OG, and our method.

[2] $1 \leqslant \tau_1 + \tau_2$ must hold otherwise there is no qDCTCs to modify.

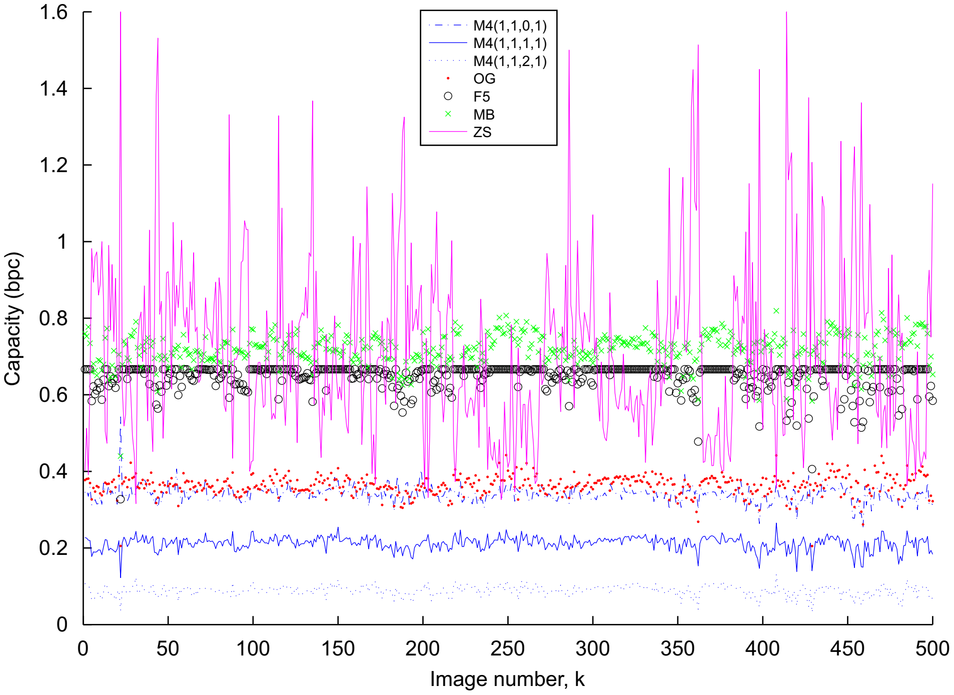

Fig. 4. Embedding capacity (*bpc*) for various images.

However, MB and F5 constantly achieve high embedding capacity while the performance of ZS fluctuates depending on the cover image. The embedding capacity of Mod4 is relatively low among the methods considered, but is comparable to OG, and adjusting the parameters allows us to control $\Omega(M4, A_k)$.

4.3. *Image quality*

To quantify the invisibility of the embedded secret message, we compute PSNR and Universal image Quality Index (UQI) [23] for stegos generated by method ε using image A_k. To have a fair evaluation, we embed a message of length $\Omega(M4(1,1,1,1), A_k)$, about 0.2*bpc*, into image A_k using $\varepsilon \in \{M4(1,1,1,1), M4(1,1,0,1), M4(1,1,2,1)$, OG, $F5$, MB, ZS$\}$.[3] Table 3 shows the average of PSNR and UQI for 500 stego images along with the JPEG compressed images. We observe that embedding the same amount of information into A_k using different methods produces almost similar PSNR and UQI values.

4.4. *Steganalysis*

Since Mod4 is not a LSB flipping method in DCT domain, it makes no sense to consider steganalyzers tools such as attacking OutGuess [9], and χ^2-statistical test [8]. Also, because there is no shrinkage [10] in Mod4, attacking F5 [10] is irrelevant as well. (Nevertheless, we verified the robustness of Mod4 against such classical steganalyzers.) In the following, we consider spike, and Fridrich's feature-based steganalysis [14] in this paper.

4.4.1. *Spike*

Most DCT-based steganographic methods try to maintain the global statistical distribution but neglect the local distributions for qDCTCs. Here, we introduce spike-based screening method. A spike occurs when the original counts satisfy

$$h_k \geqslant h_{k+1} \quad \text{for } k > 0 \quad \text{or} \quad h_{k-1} \leqslant h_k \quad \text{for } k < 0, \tag{11}$$

and the modified counts simultaneously satisfy

$$h'_k < h'_{k+1} \quad \text{for } k > 0 \quad \text{or} \quad h'_{k-1} > h'_k \quad \text{for } k < 0. \tag{12}$$

Even though no correction procedure is carried out to maintain the statistical property as in the case of

[3] In case of $M4(1,1,2,1)$, we embed the maximum capacity of $M4(1,1,2,1)$ in A_k since $\Omega(M4(1,1,2,1), A_k) < \Omega(M4(1,1,1,1), A_k)$, as shown in Fig. 4.

Table 3
Average PSNR and UQI values

Measure	JPEG	$M4(1,1,0,1)$	$M4(1,1,1,1)$	$M4(1,1,2,1)^3$	OG	F5	MB	ZS
PSNR	38.311	37.138	37.754	38.143	37.224	37.224	37.690	37.216
UQI	0.901	0.891	0.898	0.900	0.893	0.896	0.896	0.873

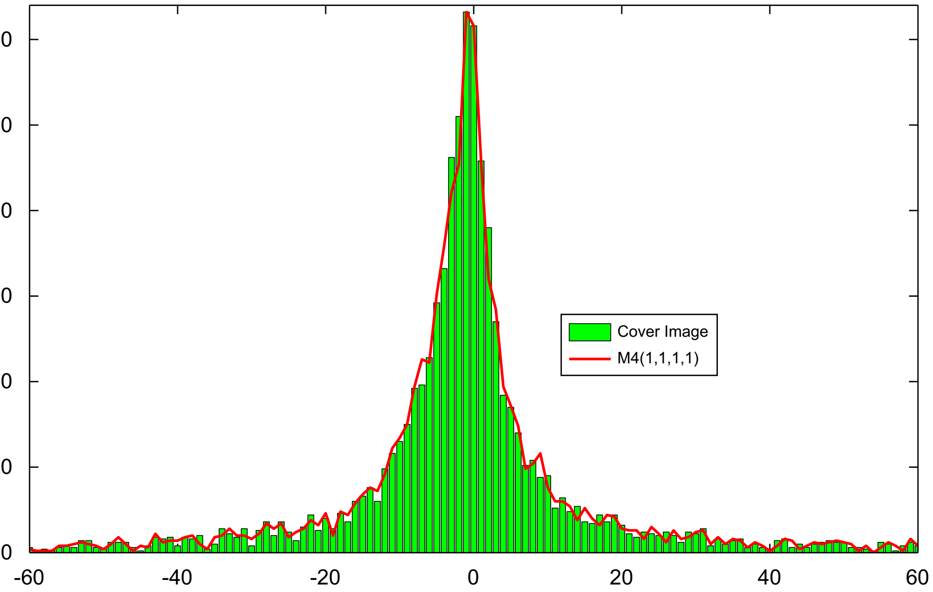

Fig. 5. Peppers-(1,2)-mode AC qDCTCs distribution.

OutGuess [3], Mod4 successfully maintains the Laplacian-like distribution both locally and hence globally. The "tail-based" modification in Mod4 complicates histogram-based steganalysis due to insufficient statistical patterns found at both ends of a distribution.

As a representative example, Fig. 5 plots the (1,2)-mode AC qDCTCs distribution for cover image "Peppers" and the corresponding $M4(1,1,1,1)$ stego that holds a message of length $\Omega(M4(1,1,1,1),$ " Peppers"). The line denotes the stego image, and bars denote the cover image. We can see that $M4$ maintains the component-wise distribution for the $(1,2)$-AC components between cover and stego images during data embedding. In fact, the other significant components (e.g., mode (2,1) and (2,2)) also exhibit similar behavior, and the same observations are obtained for other images as well. There is no awkward accumulation of coefficients in the bins labeled by $-1, 0$ or 1. Furthermore, a spike could hardly be seen and the general shape of the distribution (i.e., Laplacian) is maintained for each steganogram. Therefore, we conclude that Mod4 shows no abnormality with respect to the spike-based screening method.

Spike acts more as a precaution measure that each DCT-based steganographic method should take into consideration. The current implementation of Spike itself would not function as a steganalyzer, but it could be used as one of the preliminary screening tools to classify if a given JPEG image is suspicious. However, the potentials of expanding spike to function as a steganalyzer for the DCT-based method should be further investigated.

4.4.2. Blind steganalysis

We apply Fridrich's feature-based steganalysis method [14] to stegos generated by method $\varepsilon \in \{M4(1,1,1,1), M4(1,1,0,1), M4(1,1,2,1), OG, F5, MB, ZS\}$. Instead of computing the receiving operation characteristic curve, we consider a simple quotient of stego detection rate (SDR) defined as follows:

$$\mathrm{SDR} = \frac{\text{Total number of detected stego}}{\text{Actual number of stego } (= 200)}. \tag{13}$$

For each method ε and each of the 500 images A_k, we generate a stego $A'_k(\varepsilon, bpc)$ by embedding a message at rate bpc. The mixture of 300 cover images A_k and the corresponding 300 stegos $A'_k(\varepsilon, bpc)$ are fed into the classifier for training purposes. Specifically, we employ two discriminant functions f_0 and f_1 after the computation of the covariance matrix using 300 A'_k's and the corresponding A_k's, respectively. The rest of the 200 $A'_k(\varepsilon, bpc)$ are used for verifying the detection rate of the classifier. An image A'_k is classified using

$$A'_k = \begin{cases} \text{cover} & \text{if } f_0(A'_k) \geqslant f_1(A'_k), \\ \text{stego} & \text{otherwise.} \end{cases} \tag{14}$$

The experiment is iterated for $bpc \in \{0.2, 0.1, 0.05, 0.025\}$, and repeated for each ε.

Since it is usually the case that an adversary does not have the knowledge of the parameter values, and it is reasonable (and more realistic) to perform the aforementioned experiments for $M4$ assuming the following parameter values selected randomly:

$$\phi_1, \phi_2 \in \{0, 1, 2\} \quad \text{and} \quad \tau_1, \tau_2 \in \{1, 2\}.$$

That is, for a specific bpc, the set of stegos A'_k is generated with one of the $3 \times 3 \times 2 \times 2$ parameter combinations in Mod4. We denote this method by $M4(RP)$. The detection rate for each feature and the multi-dimensional detection rate are summarized in Tables 4–7 for 0.2, 0.1, 0.05, and 0.025bpc, respectively. Here, F5_ME and F5_111 denote the F5 method with and without matrix encoding, respectively. It should be noted that we ignore A_k in the computation for $M4(RP)$ if $\Omega(M4(RP), A_k) < 0.2bpc$ for a particular set of parameter $(\phi_1, \phi_2, \tau_1, \tau_2)$. In addition, experiments with 0.2bpc cannot be conducted for $M4(1,1,2,1)$ since embedding capacity of $M4(1,1,2,1)$ is less than 0.2bpc.

When we embed at the rate of 0.2bpc, all the methods listed are detectable by Fridrich's blind steganalyzer. However, if we decrease the embedding rate, the detection rate decreases gradually. From the result, for all the methods considered, it is difficult to survive Fridrich's blind steganalyzer. All methods are detectable (i.e., SDR > 0.5) even at embedding rate 0.05bpc. We observe that at each embedding rate, $M4(RP)$ and $M4(1,1,2,1)$ achieve

Table 4
Detection rate for Fridrich's blind steganalyzer (0.2bpc)

	$M4(1,1,0,1)$	$M4(1,1,1,1)$	$M4(1,1,2,1)$	$M4(RP)$	OG	F5_111	F5_ME	MB	ZS
Global hist.	0.835	0.720	N/A	0.681	0.880	0.995	0.785	0.585	0.590
Indi. hist. for (2,1)	0.475	0.600	N/A	0.538	0.510	0.930	0.730	0.445	0.500
Indi. hist. for (3,1)	0.865	0.735	N/A	0.692	0.780	0.965	0.700	0.585	0.490
Indi. hist. for (1,2)	0.350	0.560	N/A	0.527	0.515	0.955	0.675	0.475	0.470
Indi. hist. for (2,2)	0.535	0.730	N/A	0.681	0.820	0.980	0.870	0.635	0.560
Indi. hist. for (1,3)	0.850	0.705	N/A	0.659	0.740	0.910	0.685	0.570	0.740
Dual hist. for −5	0.960	0.470	N/A	0.484	0.535	0.490	0.580	0.560	0.615
Dual hist. for −4	0.715	0.470	N/A	0.352	0.575	0.565	0.440	0.455	0.570
Dual hist. for −3	0.545	0.535	N/A	0.396	0.495	0.425	0.535	0.500	0.495
Dual hist. for −2	0.985	0.535	N/A	0.418	0.970	0.435	0.490	0.580	0.465
Dual hist. for −1	0.430	0.415	N/A	0.374	0.735	0.560	0.540	0.545	0.845
Dual hist. for −0	0.845	0.710	N/A	0.703	0.795	0.990	0.855	0.675	0.730
Dual hist. for 1	0.535	0.405	N/A	0.363	0.610	0.635	0.445	0.535	0.855
Dual hist. for 2	0.465	0.475	N/A	0.440	0.535	0.430	0.460	0.560	0.435
Dual hist. for 3	0.430	0.510	N/A	0.429	0.480	0.495	0.475	0.510	0.590
Dual hist. for 4	0.410	0.480	N/A	0.418	0.520	0.395	0.415	0.500	0.450
Dual hist. for 5	0.625	0.415	N/A	0.593	0.540	0.400	0.505	0.595	0.590
Variation	0.415	0.555	N/A	0.505	0.540	0.735	0.550	0.500	0.520
L_1 blockiness	0.625	0.645	N/A	0.604	0.635	0.910	0.695	0.705	0.725
L_2 blockiness	0.430	0.410	N/A	0.505	0.430	0.875	0.685	0.615	0.670
Co-occurrence N_{00}	0.605	0.585	N/A	0.582	0.965	0.995	0.940	0.620	0.755
Co-occurrence N_{01}	0.745	0.665	N/A	0.571	0.980	0.925	0.665	0.750	0.500
Co-occurrence N_{10}	0.590	0.675	N/A	0.659	0.940	1.000	0.765	0.710	0.570
SDR	1.000	0.960	N/A	0.824	1.000	1.000	0.985	0.980	1.000

Table 5
Detection rate for Fridrich's blind steganalyzer (0.1*bpc*)

	$M4(1,1,0,1)$	$M4(1,1,1,1)$	$M4(1,1,2,1)$	$M4(RP)$	OG	F5_111	F5_ME	MB	ZS
Global hist.	0.640	0.560	0.490	0.570	0.720	0.930	0.645	0.525	0.525
Indi. hist. for (2,1)	0.440	0.480	0.470	0.411	0.445	0.815	0.620	0.445	0.475
Indi. hist. for (3,1)	0.660	0.605	0.535	0.525	0.610	0.885	0.545	0.505	0.520
Indi. hist. for (1,2)	0.325	0.495	0.400	0.456	0.475	0.735	0.595	0.465	0.460
Indi. hist. for (2,2)	0.465	0.560	0.570	0.551	0.700	0.920	0.690	0.530	0.570
Indi. hist. for (1,3)	0.555	0.585	0.455	0.532	0.635	0.800	0.595	0.545	0.615
Dual hist. for -5	0.710	0.475	0.665	0.551	0.485	0.470	0.570	0.430	0.550
Dual hist. for -4	0.535	0.450	0.355	0.418	0.625	0.550	0.410	0.550	0.580
Dual hist. for -3	0.425	0.470	0.310	0.462	0.470	0.450	0.560	0.535	0.500
Dual hist. for -2	0.680	0.475	0.370	0.456	0.820	0.450	0.525	0.505	0.475
Dual hist. for -1	0.290	0.435	0.305	0.405	0.530	0.460	0.465	0.505	0.590
Dual hist. for -0	0.620	0.530	0.545	0.519	0.585	0.945	0.665	0.535	0.605
Dual hist. for 1	0.370	0.390	0.645	0.348	0.435	0.495	0.415	0.500	0.580
Dual hist. for 2	0.365	0.465	0.385	0.405	0.435	0.420	0.545	0.480	0.435
Dual hist. for 3	0.360	0.465	0.375	0.392	0.485	0.475	0.505	0.520	0.570
Dual hist. for 4	0.340	0.555	0.650	0.367	0.460	0.415	0.360	0.505	0.425
Dual hist. for 5	0.635	0.550	0.335	0.418	0.565	0.410	0.540	0.565	0.580
Variation	0.380	0.555	0.665	0.544	0.500	0.595	0.500	0.465	0.520
L_1 blockiness	0.630	0.630	0.660	0.633	0.625	0.820	0.575	0.665	0.650
L_2 blockiness	0.395	0.400	0.660	0.437	0.410	0.810	0.585	0.590	0.610
Co-occurrence N_{00}	0.590	0.525	0.375	0.557	0.910	0.975	0.705	0.530	0.575
Co-occurrence N_{01}	0.575	0.590	0.435	0.544	0.880	0.750	0.605	0.600	0.525
Co-occurrence N_{10}	0.515	0.525	0.450	0.563	0.800	0.870	0.640	0.585	0.540
SDR	0.955	0.865	0.710	0.747	0.980	1.000	0.865	0.795	0.890

Table 6
Detection rate for Fridrich's blind steganalyzer (0.05*bpc*)

	$M4(1,1,0,1)$	$M4(1,1,1,1)$	$M4(1,1,2,1)$	$M4(RP)$	OG	F5_111	F5_ME	MB	ZS
Global hist.	0.525	0.485	0.460	0.470	0.580	0.770	0.535	0.430	0.510
Indi. hist. for (2,1)	0.445	0.440	0.480	0.415	0.380	0.655	0.595	0.435	0.470
Indi. hist. for (3,1)	0.585	0.520	0.510	0.490	0.510	0.730	0.500	0.500	0.565
Indi. hist. for (1,2)	0.345	0.455	0.355	0.435	0.435	0.590	0.595	0.445	0.430
Indi. hist. for (2,2)	0.450	0.515	0.505	0.480	0.555	0.720	0.585	0.435	0.545
Indi. hist. for (1,3)	0.460	0.525	0.430	0.510	0.540	0.705	0.555	0.500	0.580
Dual hist. for -5	0.450	0.525	0.395	0.515	0.520	0.445	0.570	0.575	0.430
Dual hist. for -4	0.450	0.400	0.615	0.630	0.590	0.550	0.455	0.450	0.585
Dual hist. for -3	0.315	0.525	0.375	0.460	0.490	0.555	0.520	0.490	0.505
Dual hist. for -2	0.445	0.450	0.310	0.460	0.505	0.595	0.485	0.460	0.470
Dual hist. for -1	0.300	0.415	0.315	0.430	0.450	0.400	0.470	0.445	0.405
Dual hist. for -0	0.545	0.470	0.490	0.435	0.480	0.790	0.540	0.485	0.505
Dual hist. for 1	0.355	0.385	0.665	0.370	0.410	0.435	0.430	0.470	0.410
Dual hist. for 2	0.325	0.435	0.345	0.430	0.430	0.525	0.555	0.525	0.555
Dual hist. for 3	0.335	0.525	0.370	0.455	0.425	0.465	0.495	0.485	0.550
Dual hist. for 4	0.345	0.440	0.335	0.455	0.495	0.470	0.555	0.535	0.465
Dual hist. for 5	0.640	0.525	0.605	0.455	0.560	0.405	0.530	0.555	0.390
Variation	0.355	0.445	0.655	0.545	0.470	0.545	0.485	0.450	0.515
L_1 blockiness	0.635	0.630	0.645	0.630	0.625	0.730	0.525	0.620	0.590
L_2 blockiness	0.375	0.385	0.665	0.395	0.395	0.735	0.555	0.575	0.590
Co-occurrence N_{00}	0.455	0.460	0.380	0.510	0.755	0.880	0.575	0.495	0.535
Co-occurrence N_{01}	0.505	0.510	0.380	0.490	0.730	0.640	0.540	0.535	0.505
Co-occurrence N_{10}	0.465	0.485	0.430	0.485	0.620	0.740	0.535	0.510	0.550
SDR	0.785	0.690	0.590	0.595	0.880	0.995	0.630	0.695	0.710

Table 7
Detection rate for Fridrich's blind steganalyzer (0.025*bpc*)

	$M4(1,1,0,1)$	$M4(1,1,1,1)$	$M4(1,1,2,1)$	$M4(RP)$	OG	F5_111	F5_ME	MB	ZS
Global hist.	0.470	0.405	0.420	0.385	0.525	0.630	0.470	0.390	0.455
Indi. hist. for (2,1)	0.440	0.410	0.490	0.415	0.500	0.565	0.485	0.530	0.575
Indi. hist. for (3,1)	0.560	0.485	0.530	0.475	0.505	0.605	0.535	0.520	0.435
Indi. hist. for (1,2)	0.355	0.485	0.380	0.465	0.520	0.580	0.535	0.515	0.500
Indi. hist. for (2,2)	0.420	0.435	0.475	0.445	0.505	0.650	0.550	0.455	0.530
Indi. hist. for (1,3)	0.450	0.470	0.435	0.475	0.500	0.630	0.555	0.455	0.545
Dual hist. for −5	0.400	0.565	0.630	0.385	0.565	0.545	0.570	0.445	0.605
Dual hist. for −4	0.390	0.445	0.390	0.450	0.525	0.470	0.490	0.530	0.550
Dual hist. for −3	0.315	0.480	0.325	0.460	0.470	0.485	0.520	0.605	0.580
Dual hist. for −2	0.390	0.490	0.310	0.440	0.510	0.500	0.500	0.445	0.530
Dual hist. for −1	0.310	0.535	0.305	0.530	0.530	0.380	0.400	0.625	0.455
Dual hist. for −0	0.510	0.395	0.480	0.410	0.505	0.620	0.455	0.405	0.535
Dual hist. for 1	0.350	0.530	0.660	0.450	0.440	0.325	0.325	0.375	0.455
Dual hist. for 2	0.340	0.440	0.355	0.445	0.420	0.545	0.520	0.540	0.450
Dual hist. for 3	0.350	0.510	0.645	0.475	0.520	0.470	0.500	0.545	0.550
Dual hist. for 4	0.625	0.530	0.630	0.515	0.510	0.460	0.460	0.455	0.625
Dual hist. for 5	0.660	0.490	0.370	0.505	0.535	0.560	0.555	0.425	0.555
Variation	0.350	0.415	0.645	0.590	0.390	0.500	0.475	0.420	0.490
L_1 blockiness	0.635	0.610	0.640	0.620	0.560	0.560	0.530	0.635	0.565
L_2 blockiness	0.360	0.375	0.660	0.625	0.425	0.565	0.515	0.600	0.635
Co-occurrence N_{00}	0.405	0.530	0.500	0.400	0.730	0.780	0.640	0.550	0.530
Co-occurrence N_{01}	0.450	0.545	0.400	0.495	0.660	0.645	0.620	0.530	0.610
Co-occurrence N_{10}	0.440	0.465	0.415	0.465	0.620	0.665	0.560	0.485	0.505
SDR	0.615	0.525	0.495	0.470	0.790	0.895	0.580	0.590	0.600

the minimum SDR. On top of that, if we further decrease the embedding ratio to 0.025*bpc*, only $M4(1,1,2,1)$ and $M4(RP)$ attain SDR$<$0.5. Also, $M4(1,1,1,1)$ yields SDR close to 0.5. In addition, there is an interesting trend in the results where $\Omega(M4, A_k)$ is consistently traded-off with the parameters $(\phi_1, \phi_2, \tau_1, \tau_2)$. That is, the stronger constraints on GQC to be vGQC, the lower SDR, and vice versa. As the conclusion, among the steganographic methods considered in this paper, the proposed Mod4 with appropriate parameter shows the lowest detection ratio against this particular blind steganalyzer. However, we should further improve our method to stay undetectable against this kind of effective steganalyzer without sacrificing embedding capacity as our future work.

5. Visual comparison

Three additional evaluation metrics are computed to further compare the proposed Mod4 method with other existing methods. These three metrics are filesize ratio (FSR), histogram product (HP), and embedding efficiency (EE). Specifically, FSR measures the difference between the filesizes of original image A (i.e., unmodified JPEG) and its stego image A' that contains a certain message. It is computed as

$$\mathrm{FSR}(A_k, A'_k) = 1 - \frac{8 \times |\mathrm{FS}(A'_k) - \mathrm{FS}(A_k)|}{bpc \times (\sum_{k \neq 0} h_k) \times 2}, \tag{15}$$

where h_k represents the count (from global histogram) of qDCTCs with value k, and FS is the filesize computed in bytes. Next, HP quantifies the similarity between the histograms of A and A'. It is computed as

$$\mathrm{HP} := \prod_{\substack{y=-1\\,0,1}} \left[\frac{\min\{h_y,\ h'_y\}}{\max\{h_y,\ h'_y\}} \prod_{\substack{i,j=1,2\\ i=j\neq 1}} \frac{\min\{h_y(i,j),\ h'_y(i,j)\}}{\max\{h_y(i,j),\ h'_y(i,j)\}} \right], \tag{16}$$

where $h_y(i,j), i,j \in \{0,1,\ldots,7\}$, denotes the count of the (i,j)-mode qDCTCs with value y for A, $h_y = \sum_i \sum_j h_y(i,j)$, and $h'_y(i,j)$ and h'_y are defined similarly for A'. Last but not least, EE measures the embedding efficiency in terms of the number of embedded bits per modification (say δ bits) and it is scaled into the interval [0, 1] by

$$\mathrm{EE} = 1 - \frac{1}{\delta}. \tag{17}$$

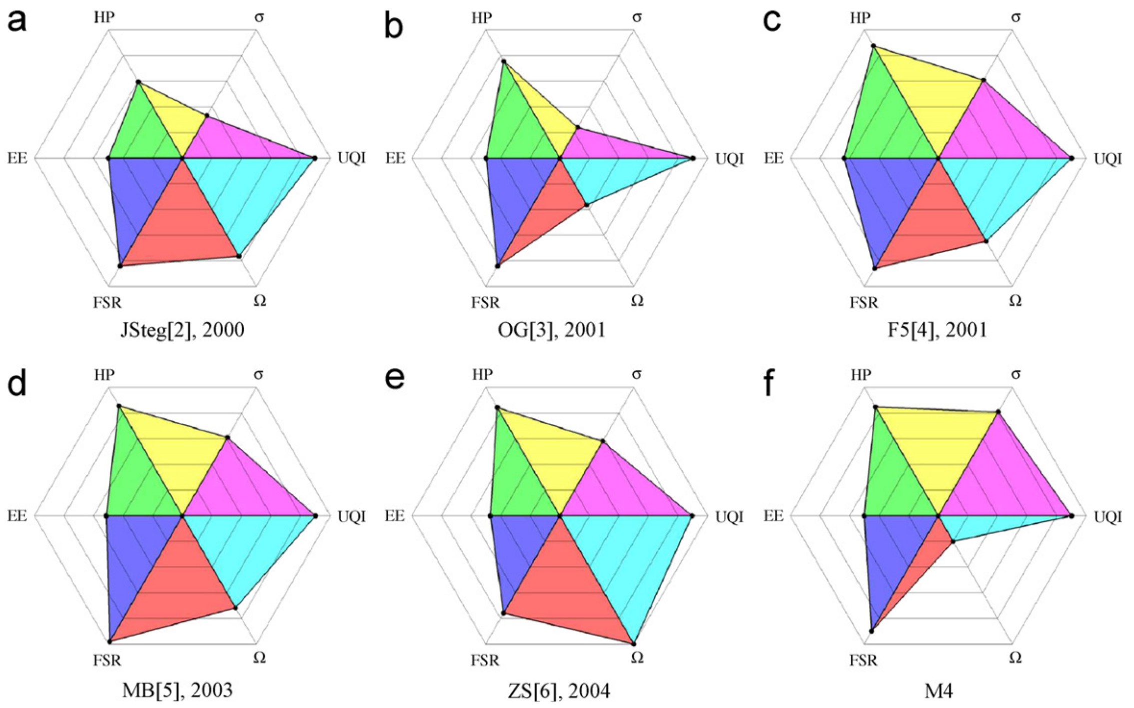

Fig. 6. Graphical comparison among DCT-based steganographic methods.

Fig. 6 compares six steganographic methods (i.e., five existing methods and Mod4) using six metrics, namely, filesize ratio FSR, histogram product HP, embedding efficiency EE, embedding capacity Ω, scaled stego detection rate σ (defined as $2 \times (1 - \mathrm{SDR})^4$), and image quality UQI. Specifically, it plots these six metrics as vectors in an ordinary polar coordinate system for each steganographic method. In each plot, six vectors (each separated at regular angle $\pi/3$) associate to six metrics and each vector is of length equal to the metric value computed. To ensure a fair comparison, all data is collected at the embedding rate of 0.05*bpc*. We also connect the tip of two adjacent vectors and fill each triangle with different color to aid the visual comparison process. Overall, F5 and MB have well-balanced performance. M4 is comparable to other existing methods in terms of HP, UQI, EE, and FSR. It achieves the highest σ (i.e., the lowest SDR) with the sacrifice of Ω.

6. Conclusions

A DCT-based steganographic method called Mod4 is proposed. A pair of message bits is hidden among qDCTCs in a vGQC. SRM is employed to ensure the expected number of modifications is minimal. Also, adjustable parameters (i.e., ϕ_1, ϕ_2, τ_1, and τ_2) are utilized to select message carriers adaptively. We have conducted experiments to compare Mod4 with other existing steganographic methods in terms of carrier capacity (*bpc*), image quality (PSNR and *Q*-metric), and detectability for blind steganalysis. In addition to these metrics, histogram product, filesize ratio, and embedding efficiency are considered to generate graphs for comprehensive visual comparisons among Mod4 and other methods. Visual comparisons indicate that the performance of Mod4 is comparable to other methods. While Mod4 provides a relatively low embedding capacity, it achieves the lowest detection ratio against blind steganalyzer. Mod4, with appropriate parameter settings, attains $\mathrm{SDR} < 0.5$ when the message length is reduced to 0.025*bpc*. At the same time, the embedding capacity in Mod4 is readily traded for the detection ratio against blind steganalysis.

[4] The multiplication of two is for the fact that all methods are detectable by the blind steganalyzer at embedding rate 0.05*bpc*, i.e., $\mathrm{SDR}(\varepsilon) > 0.5$.

As future works, we should further improve Mod4 so that it maintains the important features targeted by blind steganalyzers during data embedding without sacrificing embedding capacity. We should also consider the extension of Mod4 to embed multiple messages in an image by redefining vGQC. Furthermore, the extension of Mod4 to color images and moving pictures should also be investigated.

Acknowledgments

The authors are grateful to the anonymous reviewers for their invaluable comments and recommendations to improve our paper. Special thanks to Prof. Jessica Fridrich for generously providing the execution code for feature-based steganalysis.

References

[1] S. Katzenbeisser, F. Petitcolas, Information Hiding Techniques for Steganography and Digital Watermarking, Artech House Publishers, 2000.

[2] D. Upham, Jsteg, Software available at ⟨ftp.funet.f1⟩, 2000.

[3] N. Provos, Defending against statistical steganalysis, in: Proceeding of the 10th USENIX Security Symposium, 2001, pp. 323–335.

[4] A. Westfeld, F5 — a steganographic algorithm — high capacity despite better steganalysis, Information Hiding, Fourth International Workshop, Lecture Notes in Computer Science vol. 2137 (2001) 289–302.

[5] P. Sallee, Model based steganography, in: International Workshop on Digital Watermarking, Seoul, October 2003, pp. 154–167.

[6] K. Miyake, M. Iwata, A. Shiozaki, Digital steganography utilizing features of JPEG images, IEICE Trans. Fundam. E87-A (April 2004) 929–936.

[7] Y. Seki, H. Kobayashi, M. Fujiyoshi, H. Kiya, Quantization-based image steganography without data hiding position memorization, in: IEEE International Symposium on Circuits and Systems ISCAS, May 2005, pp. 4987–4990.

[8] A. Westfeld, A. Pfitzmann, Attacks on steganographic systems, in: Proceedings of the Third International Workshop on Information Hiding, August 1999, pp. 61–76.

[9] J. Fridrich, M. Goljan, D. Hogea, Attacking the outguess, in: Proceedings of the ACM Workshop on Multimedia and Security, 2002, pp. 967–982.

[10] J. Fridrich, M. Goljan, D. Hogea, Steganalysis of JPEG images: breaking the F5 algorithm, in: Fifth Information Hiding Workshop, Noordwijkerhout, Netherlands, October 2002, pp. 310–323.

[11] H. Farid, Detecting hidden messages using higher-order statistical models, in: International Conference on Image Processing, Rochester, New York, 2002, pp. 905–908.

[12] S. Lyu, H. Farid, Steganalysis using higher-order image statistics, IEEE Trans. Inform. Forensics Secur. 1 (March 2006) 111–119.

[13] I. Avcibas, N. Memon, B. Sankur, Steganalysis using image quality metrics, IEEE Trans. Image Process. 12 (2) (February 2003) 221–229.

[14] J. Fridrich, Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. in: Sixth Information Hiding Workshop, Lecture Notes in Computer Science, vol. 3200, New York, 2004, pp. 67–81.

[15] A. Latham, JP Hide & Seek, Software available at ⟨http://linux01.gwdg.de/alatham/stego.html⟩, 1999.

[16] T. Pevny', J. Fridrich, Multiclass blind steganalysis for jpeg images, in: Proceedings of the SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents, vol. VIII, San Jose, CA, January 2006, pp. 257–269.

[17] X. Qi, K. Wong, An adaptive DCT-based mod-4 steganographic method. in: IEEE Proceedings of ICIP, vol. II, September 2005, pp. 297–300.

[18] P. Garrett, Making, Breaking Codes: An Introduction to Cryptology, Prentice-Hall, Upper Saddle River, NJ, 2001.

[19] W. Pennebaker, J. Mitchell, JPEG Still Image Data Compression Standard, Van Nostrand Reinhold, New York, 1992.

[20] E. Lam, Analysis of the DCT coefficient distributions for document coding, IEEE Signal Process. Lett. 11 (February 2004) 97–100.

[21] E. Lam, J. Goodman, A mathematical analysis of the DCT coefficient distributions for images, IEEE Trans. Image Process. 9 (October 2000) 1661–1666.

[22] S. Smoot, L. Rowe, Study of DCT coefficient distributions. in: Proceedings of the SPIE Symposium on Electronic Imaging, vol. 2657, San Jose, CA, 1996, pp. 403–411.

[23] Z. Wang, A. Bovik, A universal image quality index, IEEE Signal Process. Lett. 9 (March 2002) 81–84.