Improving the Perturbed Quantization Steganography by Modified Matrix Encoding

Xuexiu Zhu
Department of Information Research
Information Science and Technology Institute
Zhengzhou, China
xiuzi0305@163.com

Weiming Zhang
Department of Information Research
Information Science and Technology Institute
Zhengzhou, China
zwmshu@gmai.com

Jianqing Qi
Department of Network
Electronic Engineering Institute
Hefei, China
jianqing@sohu.com

Jiufen Liu
Department of Information Research
Information Science and Technology Institute
Zhengzhou, China
jiufenliu@163.com

***Abstract*—Perturbed Quantization (PQ) and Modified Matrix Encoding (MME) are two efficient embedding methods for JPEG steganography, in which the sender uses the side information of quantizing procedure of DCT coefficients to minimize the embedding distortion. In this paper, we propose a novel embedding method which can be viewed as a generalization of the MME to utilize the quantizing information by combining the PQ and the MME method. Experimental results show that the proposed method can achieve stronger security by keeping lower average distortion for the whole cover image than the PQ and introducing less large distortion in local than the MME.**

Keywords- wet paper codes; MME; PQ; LT process

I. INTRODUCTION

Steganography is the art about covert communication [1], which hides the secret message in an innocent cover object, such as digital images, audios, and videos. To resist detection, the primary goal of steganography is to communicate as many bits as possible by as small as embedding distortion on the cover object.

The most popular cover objects used in steganography is JPEG images, in which the secret messages are embedded into the least significant bits (LSBs) of the quantized DCT coefficients [1]. One of the most famous JPEG steganographic algorithms is F5 [2], which uses the matrix encoding (ME) [3] to decrease the number of changes.

Kim et al. [4] improved the F5 by modifying matrix encoding (MME) to allow more than one change in a $(2^k - 1)$-length block and choosing a change model with small embedding distortion for each block. This method [4] can significantly decrease the embedding distortion because it uses the side information from quantizing procedure of DCT coefficients.

Perturbed Quantization (PQ) [5] steganography is another efficient method to reduce the embedding distortion by using the quantizing information. PQ steganography is based on wet paper codes, which allow the sender to embed messages with an arbitrary selection channel. Only coefficients in the selection channel will be changed and this channel is not shared with the recipient. Therefore, the sender can freely avoid changes on coefficients with large embedding distortion.

MME [4] and PQ [5] have different merits for steganographic applications. The former can avoid large distortion in each cover block, and the latter can reduce the average distortion for the whole cover. In the present paper, we will fuse the both merits by combining the PQ steganography with the MME method.

The paper is organized as follows. In Section Ⅱ we briefly introduce wet paper codes, the PQ steganography and MME method. We describe the proposed method In Section Ⅲ, and compare its performance with the MME and PQ methods in Section Ⅳ. The paper is concluded in Section Ⅴ.

II. TECHNICAL BACKGROUND

A. Wet Paper Codes

Wet paper codes were originally proposed to construct steganographic schemes with an arbitrary embedding path that is not shared by the recipient [6]. The sender first defines "dry" or "wet" positions in the cover object, and then communicates message by changing only "dry" positions. Wet paper codes allowed the recipient to extract the messages without any knowledge about the dry positions.

Assume that the cover object X consists of n elements and X has k changeable (dry) pixels. Let vector $\boldsymbol{b}$ denote the LSBs of X. After embedding messages in X, we obtain the stego- object Y and let $\boldsymbol{b}'$ be the LSBs of Y.

In wet paper coding, the sender embedding the message $\boldsymbol{m}$ by modifying $\boldsymbol{b}$ to $\boldsymbol{b}'$, satisfying (1)

$$\boldsymbol{D}\boldsymbol{b}' = \boldsymbol{m} \quad . \tag{1}$$

978-1-4244-5849-3/10/$26.00 ©2010 IEEE

Where $\boldsymbol{D}$ is a $m \times n$ pseudo-random matrix shared by the sender and the recipient. The recipient can extract $\boldsymbol{m}$ from $\boldsymbol{b}'$ by only computing (1).

If a wet paper code can embed l bits of message into a cover with k dry positions, we say that the code has embedding rate $\alpha = l/k$. If the average number of changes is R_a in the embedding process, we define the embedding efficiency as $e = l/R_a$.

It is shown in [6] that for a random binary matrix $\boldsymbol{D}$, the sender can averagely communicate the maximal message bits is $l_{\max}$

$$l_{\max} = k + O(2^{-k/4}) \quad . \tag{2}$$

when k is sufficiently large, and $k < n$. Therefore, $l_{\max}$ can be close to k, i.e., the embedding rate is about 1.

The wet paper coding can be fast implemented by LT process [7], which can embed about k message bits with on average $k/2$ changes.

B. PQ and MME

As we know, when a raw image is compressed into the JPEG format, several steps as follows are needed. First, do the two-dimensional Discrete Cosine Transform (DCT) on the pixel values. Second, use the quantization table to divide the DCT coefficients and round them to integers. Third, use the standard of JPEG compressing to encode the quantized coefficients.

PQ (Perturbed Quantization) [5] and MME (Modified Matrix Encoding) [4] try to minimize the embedding distortion using side information (the quantizing procedure of DCT coefficients) only available to the sender. Next we first explain PQ.

In general, JPEG steganography embeds message by flipping the DCT coefficients after rounded. Let ρ_i be the embedding distortion of the quantized DCT coefficient x_i if it has to be changed during embedding. Denote by x_i' the coefficient x_i before rounded, let $r_i = x_i' - [x_i'] = x_i' - x_i$ denote the rounding distortion, where $[x_i']$ is the integer that is closest to x_i' and $r_i \in [-1/2, 1/2]$, Note that we can either add 1 to x_i or subtract 1 from it to encode a secret bit. The embedding distortion may be $\rho_i = 1 - |r_i|$ or $\rho_i = 1 + |r_i|$. Then we choose the way which has the smaller embedding distortion $\rho_i = 1 - |r_i|$. For example, if we want to use the quantized coefficient 3 to embed the message bit 0 and the coefficient 3 before rounded is 2.7, then we change 3 to 2 to denote the bit 0 and the embedding distortion is 0.7. This means that changing amplitude 1 for different quantized DCT coefficients may introduce different embedding distortions. Therefore the sender can make use of the side information (e.g., DCT coefficients before rounded) to minimize the embedding distortion in JPEG steganography.

In order to embed m bits using PQ, we select m DCT coefficients with smallest embedding distortion ρ_i as the changeable coefficients. But when the recipient wants to extract the secret information, (s)he does not know the changed coefficients, so the sender must use wet paper codes while embedding message, and label the m DCT coefficients with smallest embedding distortion ρ_i as dry. Therefore, PQ is a JPEG steganography which chooses the elements with minimal embedding distortion (maximal rounding distortion) as the selection channel to embed message by using wet paper codes.

MME exploits the side information by modifying the matrix encoding [3]. MME allows more than one embedding change in matrix encoding using Hamming codes. Matrix encoding embeds messages by syndrome coding based on Hamming codes, which can embed k bits of messages into the LSBs of $(2^k - 1)$ DCT coefficients by at most one change. Taking [7, 4] Hamming code as an example, we explain how to embed and extract 3 bits of messages into 7 pixels. Let $\boldsymbol{H}$ be the parity check matrix of the [7, 4] Hamming code

$$\boldsymbol{H} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}. \tag{3}$$

Given a length-7 block of cover $\boldsymbol{b}$ which is the LSBs of 7 non-zero quantized AC DCT coefficients and a 3 bits message block $\boldsymbol{m}$, for instance $\boldsymbol{b}^T = (1\,0\,0\,1\,0\,0\,0)$, and $\boldsymbol{m}^T = (1\,1\,0)$, compute

$$\boldsymbol{Hb} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \oplus \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}. \tag{4}$$

Note that the obtained result $(0\,1\,1)$ is the binary representation of 3, and the third column of $\boldsymbol{H}$. By flipping the third bit of $\boldsymbol{b}$ we get $\boldsymbol{b}' = (1\,0\,1\,1\,0\,0\,0)$, the embedding process is completed. To extract the messages, we only need to compute $\boldsymbol{Hb}'$.

Assume that the corresponding non-zero quantized AC DCT coefficients are $\boldsymbol{x} = (x_1, x_2, \cdots, x_7)$. Let r_i be the rounded distortion at the i-th DCT coefficient. Therefore, in the above example, the embedding distortion is $\rho_3 = 1 - |r_3|$.

In MME method, the sender can find column pairs (h_i, h_j) from $\boldsymbol{H}$ such that $h_i + h_j = h_3$, and change the i-th and the j-th bits of the cover block to finish the embedding. In the example above, there are three such pairs, $(h_1, h_2), (h_4, h_7), (h_5, h_6)$. The sender calculates the embedding distortion for each pair, and finally finds the column pairs with minimal distortion denoted by $\rho' = 1 - |r_i| + 1 - |r_j|$. Change the i-th and the j-th DCT coefficients instead of the 3-th DCT coefficient if $\rho' < \rho_3$. Because the recipient reads message bits only from non-zero AC DCT coefficients, if the value of the DCT coefficient x_i' has to be changed to zero, MME usually turns it to be 2 for $x_i' > 0$ or -2 for $x_i' < 0$.

III. THE PROPOSED METHOD

In this paper we assume that the cover object is a JPEG image which is transformed by an 8-bit grayscale bmp image.

It is assumed that the sender knows the original cover object which consists of N elements, let x_i' be the non-zero AC DCT coefficient of the raw JPEG image before rounded during compression. We denote by x_i the rounded AC DCT coefficients corresponding to x_i'. Then the rounding distortion is given by $r_i = x_i' - x_i$. Let ρ_i denote the embedding distortion when x_i needs to be modified during embedding.

Denote the LSBs of N rounded AC DCT coefficients $\{x_i\}_{i=1}^N$ by $\boldsymbol{A} = \{a_i\}_{i=1}^N$. The secret message $\boldsymbol{M}$ will be embedded into $\boldsymbol{A} = \{a_i\}_{i=1}^N$ by LSB flipping. The LSB flipping is implemented by changing $\{x_i\}_{i=1}^N$ with minimal embedding distortion.

Let $\boldsymbol{H}$ be the parity check matrix of binary Hamming codes $[2^k-1, 2^k-1-k]$, which can map 2^k-1 bits to k bits as shown in Subsection Ⅱ.

We first divides the rounded AC DCT coefficients $\{x_i\}_{i=1}^N$ and their LSBs $\boldsymbol{A}$ into $n_B = \lfloor N/(2^k-1) \rfloor$ disjoint pseudo-random blocks, and each block has size of 2^k-1. Each block is then mapped into a k-bit segment by $\boldsymbol{H}$. Then we combine all the k-bit segments to be a new cover sequence denoted by $\boldsymbol{C} = \{\boldsymbol{c}^{(1)}, \boldsymbol{c}^{(2)}, \cdots, \boldsymbol{c}^{(n_B)}\}$, where $\boldsymbol{c}^{(i)} = \{c_1^{(i)}, \cdots, c_k^{(i)}\}$, $1 \le i \le n_B$. Then we will embed message $\boldsymbol{M}$ into $\boldsymbol{C}$ by bit replacement, i.e., replace the bit of $\boldsymbol{C}$ by the message bit. Denote by $\boldsymbol{C}' = \{\boldsymbol{c}'^{(1)}, \boldsymbol{c}'^{(2)}, \cdots, \boldsymbol{c}'^{(n_B)}\}$ the modified $\boldsymbol{C}$, where $\boldsymbol{c}'^{(i)} = \{c_1'^{(i)}, \cdots, c_k'^{(i)}\}$. We use a scalar value $\hat{d}^{(i)}$ to measure the distortion of the i-th block.

Take the first block as an example, we explain how to calculate the distortion $\hat{d}^{(1)}$. Assume that the first cover block is $(x_1, x_2, \ldots, x_7)$ with LSBs such that $(a_1, a_2, \ldots, a_7)$, which can be mapped into 3-bit segment $(c_1^{(1)}, c_2^{(1)}, c_3^{(1)})$ by $\boldsymbol{H}$. Assumed that the first message block is $(m_1^{(1)}, m_2^{(1)}, m_3^{(1)})$, which has 2^3 possible states with equal probability 1/8. For example, if $(m_1^{(1)}, m_2^{(1)}, m_3^{(1)}) = (0,0,0)$, then we can make $(c_1^{(1)}, c_2^{(1)}, c_3^{(1)}) = (0,0,0)$ by changing $(x_1, x_2, \ldots, x_7)$ with the MME method. Denote the possible distortion by $d_0^{(1)}$. In this manner, we can calculate $d_1^{(1)}, d_2^{(1)}, \cdots, d_7^{(1)}$ that correspond to $(m_1^{(1)}, m_2^{(1)}, m_3^{(1)}) = (0,0,1), \cdots (1,1,1)$. We define the maximal distortion $\hat{d}^{(1)} = \max(d_0^{(1)}, d_1^{(1)}, \cdots, d_7^{(1)})$, which is used to measure the possible embedding impact in the first cover block. We calculate $\hat{d}^{(2)}, \cdots, \hat{d}^{(n_B)}$ for each block before embedding.

Assume that the sender wants to embed the message $\boldsymbol{M} = \{m_i\}_{i=1}^m$, that is, the embedding rate is $\alpha = m/N$. We define the relative embedding rate α' for the proposed method such that $\alpha' = m/n_B k$. When $\alpha' < 1$, we can choose the blocks with lower distortion $\hat{d}^{(i)}$ as the selection channel. Therefore, according to the embedding rate α', we introduce the distortion threshold T, which makes the proportion of blocks having $\hat{d}^{(i)} < T$ is α', i.e., $P(\hat{d}^{(i)} < T) = \alpha'$. If $\hat{d}^{(i)} < T$, we label $(x_1^{(i)}, x_2^{(i)}, \ldots, x_k^{(i)})$ as "dry", otherwise we label the block as "wet". With wet paper codes, we can embed the message $\boldsymbol{M}$ by only modifying the "dry" cover block.

Note that the distortion measurement $(\hat{d}^{(1)}, \hat{d}^{(2)}, \cdots, \hat{d}^{(n_B)})$ is used to avoid changing blocks with large embedding impact. The actual distortion in the i-th block caused by the embedding may be small than $\hat{d}^{(i)}$.

If there are defective pixels in the original cover $\boldsymbol{C}$ that are forbidden to be modified, the MME method can not be used any more. By defining infinite distortion on the defective pixels, the proposed method is still feasible for such kind of covers.

IV. EXPERIMENTS

Fig.1 shows that the embedding efficiency of various embedding rate for Perturbed Quantization, MME and the proposed method. The results were obtained by averaging 100 embeddings in a cover image whose size is 512×512. From Fig. 1, we can see that the performance of the proposed method is significantly better than that of PQ and close to that of MME. Note that in Section 3, if we do not set the distortion threshold T and mark all the block as "dry", then we just obtain the MME. Therefore the proposed method is a generalization of the MME.

Table Ⅰ shows that the number of blocks with embedding distortion larger than T when we use MME to embed the same message as used in the proposed method. Where R denotes the embedding rate, and T is the distortion threshold chosen in the proposed method.

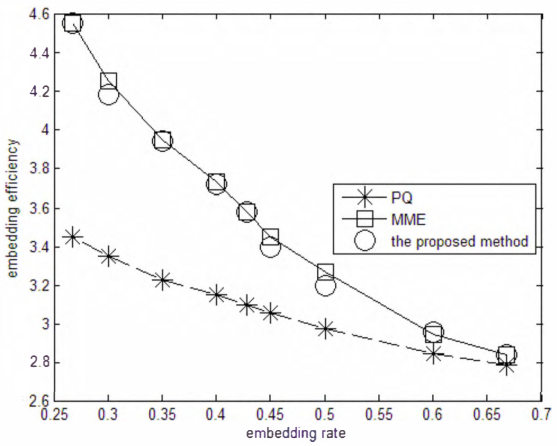

Figure 1. The comparison of embedding efficiency for the PQ, MME and the proposed method.

TABLE I. THE NUMBER OF BLOCKS WITH LARGE DISTORTION IN MME

R	0.6	0.5	0.45	0.4	0.35	0.3
$>T$	760	650	550	120	110	107

Table I also shows that the MME method can not avoid large distortion in local, which may be collected by the attacker to detect the existence of the secret message, while the proposed method can skip these blocks by wet paper coding.

V. CONCLUSIONS

In this paper we propose a novel method to embed message in the least significant bits of JPEG coefficients. By fusing modified matrix encoding (MME) and wet paper codes, the proposed method can reach low average distortion and avoid modification in some sensitive areas of the cover, and therefore improve the security of the steganography. On the other hand, this method can be used to wet paper cover, that is, a cover with defective element, and therefore it can be viewed as an extension of the MME method.

REFERENCES

[1] J. Fridrich, T. Pevny, and J. Kodovsky, "Statistically Undetectable JPEG Steganography: Dead Ends, Challenges, and Opportunities," Proc. ACM MM&Workshop, Dallas, TX, Sept. 2007, pp. 3-14, doi:10.1145/1288869.1288872.

[2] A. Westfeld, "High Capacity Despite Better Steganalysis(F5-A Steganography Algorithm)," Proc. Information Transactions Hiding, 4th International Workshop, Pittsburgh, PA, USA, LNCS 2137, Springer-Verlag, New York, 2001, pp. 289-320.

[3] R. Crandall. "Some notes on steganography," Posted on Stegano - graphy Mailing List, 1998, http://os.inf.tu-drenden de/~crandall.pdf.

[4] Y. Kim, Z. Duric, D. Richards, "Modified Matrix Encoding Technique for Minimal Distortion Steganography," In N. Johnson and J. Camenisch editors, Information Hiding 8th International Workshop, volume 4437 of Lecture Notes in Computer Science. Springer-Verlag, New York, 2006, pp. 314-327, doi:10.1007/978-3-540-74124-4_21.

[5] J. Fridrich, M. Goljan, D. Soukal, "Perturbed Quantization Steganography with Wet Paper Codes," Proc. ACM Multimedia Security Workshop, Magdeburg, Germany, Sept. 2004, pp.4-15, doi:10.1145/1022431.1022435.

[6] J. Fridrich, M. Goljan, P. Lisonek, D. Soukal, "Writing on Wet Paper," IEEE Trans. on Sig. Proc., Special Issue on Media Secutity, Eds. T. Kalker and P. Moulin, vol. 53, Oct. 2005, pp. 3926-3935, doi:10.1117/12.583160.

[7] J. Fridrich, M. Goljan, D. Soukal, "Efficient Wet Paper Codes," Information Hiding. 7th International Workshop, LNCS vol. 3727, Springer-Verlag, New York, 2005, pp. 204-218, doi:10.1007/ 11558859_16.