AN IMPROVED JPEG2000 STEGANOGRAPHY USING QIM AND ITS EVALUATION BY STEGANALYSIS

Takayuki Ishida, Kazumi Yamawaki, Hideki Noda, Michiharu Niimi

Kyushu Institute of Technology, Dept. of Systems Design and Informatics,
680-4 Kawazu, Iizuka, 820-8502 Japan
E-mail: {ishida, yamawaki, noda, niimi }@mip.ces.kyutech.ac.jp

ABSTRACT

This paper presents a modified QIM-JPEG2000 steganography which improves the previous JPEG2000 steganography using quantization index modulation (QIM). Post-embedding changes in file size and PSNR by the modified QIM-JPEG2000 are smaller than those by the previous QIM-JPEG2000. Steganalysis experiments to determine whether messages are embedded in given JPEG2000 images show that the modified QIM-JPEG2000 is more secure than the previous QIM-JPEG2000.

Keywords: steganography, steganalysis, JPEG2000, quantization index modulation

1. INTRODUCTION

Steganography is the process of hiding secret data in an innocent looking dummy container. This container may be a digital still image, audio file, or video file. Once the data has been embedded, it may be transferred across insecure lines or posted in public places. Therefore, the dummy container should seem innocent under most examinations. On the other hand, steganalysis is the task of attacking steganographic systems. Considering the aim of steganography, it might be sufficient if an attacker can only detect the presence of hidden data in a container. The main requirement of steganography is undetectability, which means that no steganalysis algorithm exists that can determine whether data is embedded in a given container.

In steganography using digital images, data embedding into compressed images should be primarily considered since images are usually compressed before being transmitted. The JPEG compression using the discrete cosine transform (DCT) is now the most common compression standard for still images, and therefore many steganographic methods have already been proposed for JPEG images including [1]-[6]. Several steganalysis methods for JPEG steganography have also been proposed to detect whether messages are embedded or not in a JPEG image [2],[7]. Steganalysis methods in [2],[7] exploit some changes in the histogram of quantized DCT coefficients caused by embedding. Steganalysis in [8] exploits higher order statistics as well as the first order statistics such as the histogram of DCT coefficients.

JPEG2000 using the discrete wavelet transform (DWT) is an incoming image coding standard which has improved features over the JPEG and is believed to be used widely. Since steganographic methods for JPEG2000 images might be commonly used in the near future, development of secure JPEG2000 steganography will be required soon. Among already proposed JPEG2000 steganographic methods [9]-[12], QIM-JPEG2000 steganography [12], which uses quantization index modulation (QIM) [13] in DWT domain, has a significant feature that it approximately preserves histograms of quantized DWT coefficients. The histogram preservation should be a necessary requirement for secure JPEG2000 steganography since steganalysis for JPEG2000 steganography will be likely to exploit firstly histogram changes by embedding. The QIM-JPEG2000 steganography, however, has a drawback that the file size of a post-embedding stego image increases significantly compared with that of its cover image; the increase in file size is much more than the size of embedded data. The increase of stego image size is not directly related with detectability of the presence of hidden data since in steganography it is assumed that an attacker cannot access its original cover image. However, the excessive increase in file size is not desirable since it should cause some sort of distortion to a stego image.

This paper presents a modified QIM-JPEG2000 steganography which does not increase the post-embedding file size while still keeping the post-embedding histogram almost unchanged. It is realized by embedding data without changes of quantized DWT coefficients between 0 and ± 1.

2. QIM-JPEG2000 STEGANOGRAPHY

In this section, we briefly review the QIM-JPEG2000 steganography [12]. In the QIM-JPEG2000 steganography, QIM [13] with two different quantizers is used to embed binary data at the quantization step of DWT coefficients. Each bit (zero or one) of binary data is embedded in such a way that one of two quantizers is used for quantization of a DWT coefficient, which corresponds to embed zero, and the other quantizer is used to embed one. In the following, it is assumed that the probabilities of zero and one are same in binary data to be embedded. This assumption is quite natural since any compressed data has such property.

Assuming that DWT coefficients belonging to a codeblock[1] are divided by its quantization step size in advance, two codebooks, C^0 and C^1, for two quantizers can be defined as $C^0 = \{0, \pm(2j+0.5); j \in \{1, 2, \ldots\}\}$ and $C^1 = \{\pm(2j+1.5); j \in \{0, 1, 2, \ldots\}\}$ for all frequency subbands. Let N_i and N_{-i}, $i \in \{1, 2, \ldots\}$ denote the number of DWT coefficients whose values w are in the interval $i \leq w < i+1$ and $-i-1 < w \leq -i$, respectively, and N_0 in the interval $-1 < w < 1$. Let N_i^L and N_i^H denote the number of DWT coefficients in the lower and higher half interval of N_i, respectively, and therefore $N_i^L + N_i^H = N_i$. After embedding by QIM, the histogram N_i is changed to $N_i^{'}$ as

$$N_i^{'} = \frac{1}{2}N_i + \frac{1}{2}(N_{i-1}^H + N_{i+1}^L). \quad (1)$$

Eq. (1) indicates that if $N_i = N_{i-1}^H + N_{i+1}^L$, then the number in the bin i does not change. In particular for $i = 0, \pm 1$, however, much difference between N_i and $N_{i-1}^H + N_{i+1}^L$ causes the significant change on $N_i^{'}$ after embedding. That is, since N_0 is usually larger than N_1 and N_{-1}, the most significant changes are decrease of N_0 and increase of N_1 and N_{-1}. In order to preserve N_0, N_1 and N_{-1} after embedding, a dead zone for DWT coefficients w, $t_d^- < w < t_d^+$ $(-1 < t_d^- < 0 < t_d^+ < 1)$ is introduced, where DWT coefficients are not used for embedding. Let N_d^+ and N_d^- denote the number of positive DWT coefficients and that of negative coefficients in the dead zone, i.e., the number of coefficients in the interval $0 < w < t_d^+$ and $t_d^- < w < 0$, respectively. t_d^+ and t_d^- are determined by optimum N_d^+ and N_d^- values which minimize the histogram changes for the bins 0 and ± 1. Note that in the QIM-JPEG2000 steganography, quantized coefficients 0s cannot be treated as zeroes embedded in them, because they cannot be discriminated from 0s in the dead zone. Also note that in data extraction stage, information on the dead zone (t_d^+ and t_d^-) is not necessary and data extraction is simply carried out based on whether non-zero coefficients are even or odd.

3. MODIFIED QIM-JPEG2000 STEGANOGRAPHY

We investigate the reason why the file size of post-embedding image by the QIM-JPEG2000 steganography increases significantly compared with that of its cover image. Fig. 1(a) shows the relation between the file size increase and the number of quantized DWT coefficients which changes from 0 to ± 1 after embedding, and Fig. 1(b) shows the relation between the file size increase and the number of changes from ± 1 to 0 after embedding. Fig. 1(c) shows the relation between the file size increase and the increase of ± 1 after embedding. Data in these figures are derived using eight standard images described in 4.1. These figures show that the file size increase is correlated with the number of changes between 0 and ± 1 and is not correlated with the increase of ± 1 after embedding. This evidence may indicate that the file size increase is caused by violating adaptive encoding of the arithmetic encoder in JPEG2000 which considers context of nearby pixels. That is, the change between 0 and ± 1 by embedding is made independently of the context and it may cause the increase.

In order to avoid the changes of quantized DWT coefficients between 0 and ± 1, we modify the previous QIM-JPEG2000 as follows.

(1) DWT coefficients in the interval $-1 < w < 1$ whose quantized values are 0s are not used for embedding.

(2) For DWT coefficients in the interval $1 < w < 2$ and $-2 < w < -1$, dead zones, $1 < w < t_d^+$ and $t_d^- < w < -1$ $(1 < t_d^+ < 2, -2 < t_d^- < -1)$ are introduced, where DWT coefficients are not used for embedding. The two dead zones are introduced to make histogram changes as small as possible for the bins 1 and 2 and for -1 and -2. The dead zones can be set by a similar way to one in the previous QIM-JPEG2000 [12]. For DWT coefficients outside the dead zones, half of the coefficients in $t_d^+ < w < 2$ and half of the coefficients in $-2 < w < t_d^-$ are quantized to 2 and -2, respectively, for embedding zeros.

Note that in the modified QIM-JPEG2000 steganography, quantized coefficients ± 1s cannot be treated as ones embedded in them, because they cannot be discriminated from ± 1s in the dead zones. Also note that in data extraction stage, information on the dead zones (t_d^+ and t_d^-) is not necessary and data extraction is simply carried out based on whether coefficients other than 0 and ± 1 are even or odd.

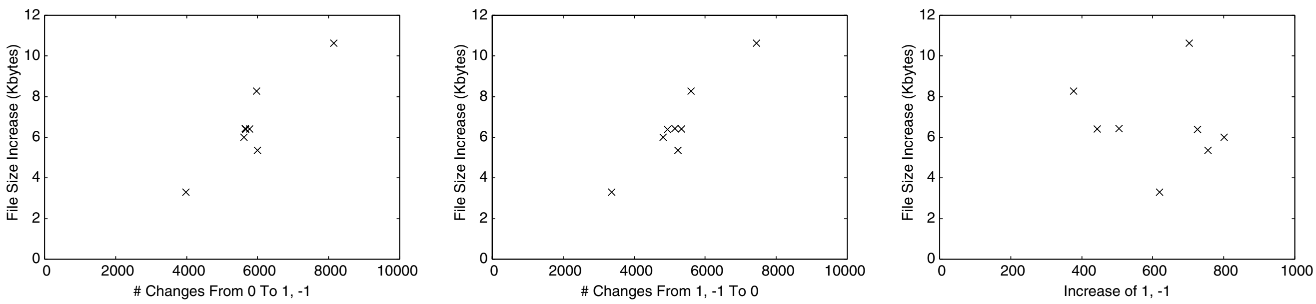
4. EXPERIMENTS

4.1 General Performance Evaluation

The modified QIM-JPEG2000 was evaluated by comparing it with the previous QIM-JPEG2000 steganography and the least significant bit (LSB) flipping steganography. These three methods were tested using eight standard images: Lena, Barbara, Mandrill, Airplane, Boat, Goldhill, Peppers, and Zelda from a database (http://sampl.eng.ohio-state.edu/~sampl/database.htm). These images are 512×512 pixels in size, 8 bit per pixel (bpp) gray images, and were compressed with 1 bpp as the pre-embedding target bit rate. The histogram change was measured by Kullback-Leibler divergence [14]. Smaller KL divergence values represent better histogram preservation. Experiments were carried out 100 times using different random data to be embedded into each image. Experimental results are shown in Table 1, where each result is the mean value for eight images. The KL divergence in the table are those averaged over three subbands (LH, HL, and HH subband) of third-level in five-level wavelet transform used. The third-level subbands are here selected considering the balance between the total number of DWT coefficients and the number of non-zero DWT coefficients in a subband.

Embedded data size by the modified QIM-JPEG2000 becomes smaller than that by the previous QIM-JPEG2000 because in the modified QIM-JPEG2000, neither quantized coefficients 0s nor ± 1s are used for embedding. The same

[1]The codeblock is a unit processing block in JPEG2000 coding. The quantization step size can be different from codeblock to codeblock.



(a) (b) (c)

Fig. 1: File size increase by QIM-JPEG2000 steganography: (a) file size increase vs. the number of changes from 0 to ± 1, (b) file size increase vs. the number of changes from ± 1 to 0, (c) file size increase vs. increase of ± 1.

Table 1: Results of embedding experiments.

method	embedded data size (bytes)	compressed image size (bytes)	file size increase (bytes)	PSNR (dB)	KL divergence
(no embedding)	-	32793	-	38.0	-
QIM-JPEG2000 (max. amount)	3865	39403	6610	35.3	0.0030
QIM-JPEG2000 (equal amount)	2446	37234	4441	36.0	0.0019
Modified QIM-JPEG2000	2446	33249	456	37.1	0.0022
LSB	2425	33101	308	36.6	0.0095

is true for LSB where flipping is carried out for quantized coefficients other than 0s and ± 1s. Therefore, experiments with equal amount of embedding were also performed for the previous QIM-JPEG2000 (see Table 1). The amount was adjusted by randomly selecting DWT coefficients used for embedding. It is seen that the file size increase by the modified QIM-JPEG2000 and LSB is much smaller than that by the previous QIM-JPEG2000. Additionally, the modified QIM-JPEG2000 produces the highest PSNR stego images among the three methods.

The KL divergence value for the modified QIM-JPEG2000 is comparable to or only a little bit larger than that for the previous QIM-JPEG2000 with equal amount of embedding. It is probably because in the previous QIM-JPEG2000, histogram preservation is considered only for the bins 0 and ± 1, but less amount of embedding than maximum reduces the KL divergence value to that for the modified QIM-JPEG2000 where histogram preservation is considered for ± 2 as well as 0 and ± 1. Regarding histogram preservation, LSB is much worse than the other two methods.

4.2 Steganalysis

Steganalysis experiments were carried out to determine whether messages are embedded in JPEG2000 images. For the experiments, 500 natural images were used which are 408×306 pixels in size, 8 bpp gray images and were compressed with 1 bpp as the pre-embedding target bit rate. A classifier using Fisher linear discriminant analysis [15] was used to classify whether a given image is a stego or cover image. Two kinds of features were used for the classifier: higher-order image statistics extracted from a wavelet decomposition [16] and histograms of wavelet coefficients. The former feature consists of 72 components which are the mean, variance, skewness and kurtosis of wavelet coefficients at three subbands of the first three (first to third) levels, and those of errors in an optimal linear predictor of coefficient magnitude [16]. The latter consists of 55 components which represent positive part histogram of wavelet coefficients, i.e., N_0 to N_{10} for all of five levels[2].

Given 500 cover images, corresponding stego images were generated by each of the above four methods in 4.1. Randomly selected 250 cover images and corresponding stego images were used for training the classifier, and the remaining 250 cover and 250 stego images were used for testing. At the training stage, the decision threshold for the classifier was set so that the false positive rate becomes 2%, i.e., correct detection rate for a cover image is 98%. Then the derived decision threshold was used in testing. Experiments were carried out 100 times using randomly selected images for training and testing, and the average of correct detection rates is shown in Table 2.

Results in Table 2 show that regarding the features, steganalysis using histograms works better than that using higher-order statistics. The results by the previous QIM-JPEG2000 with maximum and equal amount of embedding show that correct detection rate for a stego image naturally decreases with less amount of embedding. Comparing the two correct detection rates for a stego image by the previous and the modified QIM-JPEG2000, it is confirmed that the modified QIM-JPEG2000 is more secure than the previous QIM-JPEG2000. Embedding by LSB flipping is completely de-

[2]Considering nearly symmetrical shape of a histogram, negative part was omitted.

Table 2: Correct detection rates (%) in steganalysis experiments.

method	higher-order statistics cover	higher-order statistics stego	histograms cover	histograms stego
QIM-JPEG2000 (max. amount)	92.5	41.9	95.0	73.8
QIM-JPEG2000 (equal amount)	91.9	29.9	94.6	43.7
Modified QIM-JPEG2000	91.1	10.3	95.1	36.1
LSB	90.9	13.5	97.0	100.0

tected using the histogram feature since it causes significant histogram changes.

5. CONCLUSIONS

We have presented a modified QIM-JPEG2000 steganography by which the file size increase is much less than that by the previous QIM-JPEG2000, while keeping post-embedding histogram change comparable to that by the previous QIM-JPEG2000. Furthermore, post-embedding decrease of PSNR value by the modified QIM-JPEG2000 is smaller than that by the previous QIM-JPEG2000. Steganalysis experiments show that the modified QIM-JPEG2000 is more secure than the previous QIM-JPEG2000.

Acknowledgements This work was partly supported by Grants-in-Aid for Scientific Research (No.18360183).

6. REFERENCES

[1] D. Upham, http://ftp.funet.fi/pub/crypt/cypherpunks/steganography/jsteg/

[2] A. Westfeld, "F5 - A steganographic algorithm: high capacity despite better steganalysis", Lecture Notes in Computer Science, vol.2137, pp.289-302, 2001.

[3] N. Provos, "Defending against statistical steganalysis", 10th USENIX Security Symposium, 2001.

[4] J.J. Eggers, R. Bauml and B. Girod, "A communications approach to image steganography", Proceedings of SPIE, vol.4675, pp.26-37, 2002.

[5] P. Sallee, "Model-based steganography", Lecture Notes in Computer Science, vol.2939, pp.154-167, 2004.

[6] H. Noda, M. Niimi and E. Kawaguchi, "High performance JPEG steganography using quantization index modulation in DCT domain", Pattern Recognition Letters, vol.27, pp.455-461, 2006.

[7] J. Fridrich, M. Goljan and D. Hogea, "New methodology for breaking steganographic techniques for JPEGs", Proceedings of SPIE, vol.5020, pp.143-155, 2003.

[8] J. Fridrich, "Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes", Lecture Notes in Computer Science, vol.3200, pp.67-81, 2004.

[9] K. Ando, H. Kobayashi and H. Kiya, "A method for embedding binary data into JPEG2000 bit streams based on the layer structure", Proceedings of 2002 EURASIP EUSIPCO, vol.3, pp.89-92, 2002.

[10] H. Noda, J. Spaulding, M.N. Shirazi and E. Kawaguchi, "Application of bit-plane decomposition steganography to JPEG2000 encoded images", IEEE Signal Processing Letters, vol.9, no.12, pp.410-413, 2002.

[11] P.C. Su and C.C.J. Kuo, "Steganography in JPEG2000 compressed images", IEEE Trans. on Consumer Electronics, vol.49, no.4, pp.824-832, 2003.

[12] H. Noda, Y. Tsukamizu and M. Niimi, "JPEG2000 steganography which preserves histograms of DWT coefficients", IEICE Trans. Information and Systems, vol.E90-D, no.4, pp.783-786, 2007.

[13] B. Chen and G.W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding", IEEE Trans. on Information Theory, vol.47, no.4, pp.1423-1443, 2001.

[14] D.J.C. MacKay, Information theory, inference, and learning algorithm, Cambridge University Press, 2003.

[15] R.O. Duda, P.E. Hart and D.G. Stork, Pattern classification, Wiley-Interscience Publication, 2001.

[16] S. Lyu and H. Farid, "Detecting hidden messages using higher-order statistics and support vector machines", Lecture Notes in Computer Science, vol.2578, pp.340-354, 2003.