

A New Approach to Reducing the Distortion of the Digital Image Natural Model in the DCT Domain When Embedding Information According to the QIM Method

O.O. Evsutin^{1,2,3}, A.S. Melman³, R.V. Meshcheryakov², A.O. Ishakova^{2,3}

evsutin.oo@gmail.com | annakokurina94@yandex.ru | mrv@ipu.ru | shumskaya.ao@gmail.com

¹National Research University Higher School of Economics, Moscow, Russia;

²V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia;

³Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russia

One of the areas of digital image processing is the steganographic embedding of additional information into them. Digital steganography methods are used to ensure the information confidentiality, as well as to track the distribution of digital content on the Internet. Main indicators of the steganographic embedding effectiveness are invisibility to the human eye, characterized by the PSNR metric, and embedding capacity. However, even with full visual stealth of embedding, its presence may produce a distortion of the digital image natural model in the frequency domain. The article presents a new approach to reducing the distortion of the digital image natural model in the field of discrete cosine transform (DCT) when embedding information using the classical QIM method. The results of the experiments show that the proposed approach allows reducing the distortion of the histograms of the distribution of DCT coefficients, and thereby eliminating the unmasking signs of embedding.

Keywords: information security, steganography, digital images, discrete cosine transform.

1. Introduction

One of the promising areas for solving the problem of ensuring information security of multimedia data is the use of digital steganography and digital watermarking (DWM) technique, which allow you to hide additional information of various purposes in digital objects, in particular, in digital images.

The methods for embedding information in digital images are divided into methods for embedding in the spatial domain and in the frequency domain of digital images. In practice, the use of embedding methods in the frequency domain is more effective, since such embedding in the general case provides greater resistance to various destructive influences. However, most of the known frequency embedding algorithms lead to significant distortions of a digital image natural model in the frequency domain. Such distortions are an unmasking feature that reduces the stability of the steganographic algorithm before steganalysis, aimed at identifying the presence of embedded additional information in digital objects. Using steganalysis methods, an intruder can detect the presence of an embedded message in a given stego-image and subsequently compromise or destroy it.

In general case, steganalysis of digital images is considered as a two-class classification problem. Many modern methods of steganalysis are the development of a classical study [4], which proposes a set of features for conducting steganalysis. The studies presented in articles [5, 8] are aimed at minimizing distortions of the natural model of digital images by using various feature spaces. The articles [2, 6] are devoted to steganalysis, the improvement of steganographic algorithms to counteract it and the expansion of feature spaces.

This article proposes a new approach to reducing the distortions of the digital image natural model in the discrete cosine transform (DCT) domain by the steganographic embedding of information. Embedding is performed according to an algorithm based on the popular steganographic method of quantization index modulation (QIM). The idea of the proposed approach consists in adaptive selection of the quantization step (the main parameter of the classical QIM method) depending on the characteristics of a particular cover image. The aim of the work is to study the effectiveness of this approach and its specific algorithmic implementations.

2. The embedding operation

Embedding operation, i.e. operation of direct changes in the values of frequency coefficients, is based on the QIM method. The idea of the QIM method [7] is to modulate the brightness of the pixels or the values of the frequency coefficients depending on the values of the embedded bits. In this study, the QIM method is used to embed information in DCT coefficients. Images are processed in blocks of 8×8 pixels. The embedding area consists of 36 high- and mid-frequency AC-coefficients. The embedding of information is carried out by the formula

$$c' = q \cdot \left\lfloor \frac{c}{q} \right\rfloor + \frac{q}{2} \cdot b_i,$$

where c is the DCT coefficient before embedding, c' – is the DCT coefficient after embedding, b_i is the secret message bit,

q is the quantization step.

The algorithm used in this study is distinguished by the possibility of error-free extraction of embedded information due to an iterative embedding procedure. Since this feature of the algorithm does not affect the distortion of the digital image natural model in the frequency domain, typical for the QIM method, we shall omit its description. The authors of this study give more information on the principle of iterative correction of extraction errors by example of the discrete Fourier transform in [3].

3. Proposed Approach

The application of the QIM method to the DCT domain is associated with the problem of distortion of the digital image natural model. In the present work, as a digital image natural model in the DCT domain we mean a histogram of distribution of the DCT coefficient values. An example of a typical histogram of DCT-coefficients of the image is shown in Fig. 1 (a).

If we embed a message in the corresponding image using the QIM method with a predefined quantization step q , the histogram of the DCT coefficients will take the form shown in Fig. 1 (b). The obtained histogram is markedly different from the original. This is due to the fact that the classical QIM method narrows the number of possible variants of the DCT coefficient values [7].

One of the solutions to this problem was previously proposed by the authors of this study in [1], where to redistribute the

arising distortions of the histogram, the quantization step was variable and depended on the ratio of the average values of the AC-coefficients moduli of the one-dimensional DCT in and out of the embed area. This approach demonstrated a positive effect on the compensation of distortions of the histogram, but the problem was not completely solved. The form of the histogram for a number of images still made it possible to unambiguously determine the presence of a steganographic attachment. Fig. 1 (c) demonstrates the influence of the approach described in [1] on the reduction of histogram distortions.

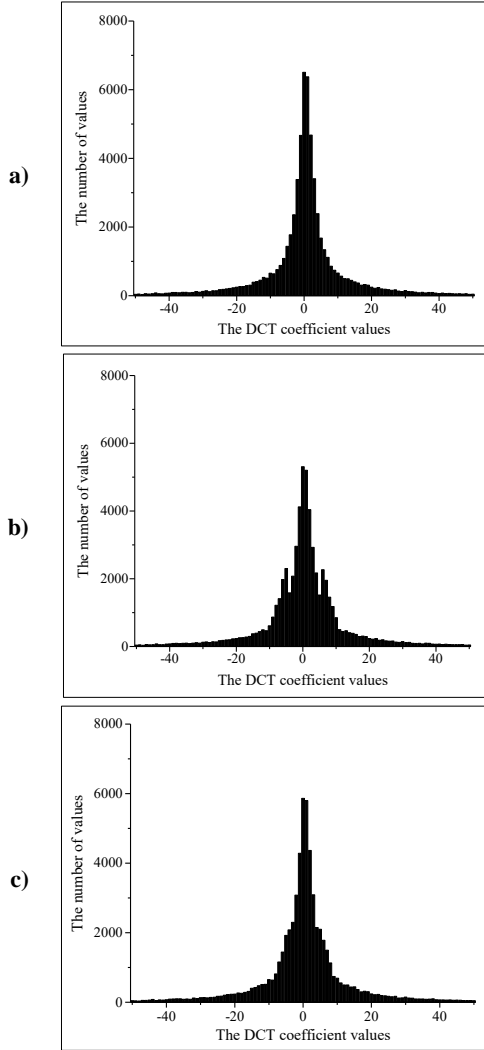


Fig. 1. The histogram of the image: a) before the embedding; b) after the embedding with a constant quantization step; c) after the embedding with a variable quantization step from [1].

This study proposes to develop the idea of using the ratio of domains within a block to select a quantization step.

Obviously, the embedding area undergoes the most significant distortions. At the same time, a change in the non-embedding area, i.e. in the other AC-coefficients of the block is negligible. Therefore, to select the quantization step, it is proposed to use the non-embedding area of the corresponding block. The invariance of the non-embedding area will allow to extract embedded data without errors, since the quantization step selected over the non-embedding area will be the same for both the cover and the stego-images. It does not require knowledge of any additional key information. This means that the transmission of the stego-image does not require a preliminary exchange of keys, the presence of which would contradict the very idea of

steganography due to the compromise of the fact of the hidden transmission of information.

In order to decide how the non-embedding area can be interconnected with the value q , it is necessary to pay attention on the nature of the distortions that arise when embedding information with a constant quantization step. The histogram in Fig. 1 (b) corresponds to the stego-image obtained with $q = 8$. The peaks that appear in the histogram correspond to the values ± 4 , i.e. $\pm \frac{q}{2}$. So, a certain predetermined value leads to the

fact that the frequency of occurrence of values equal to $\pm \frac{q}{2}$ increases. To reduce the probability that some new value of the quantization step for the next block will coincide with the most frequently occurring value of DCT coefficients, i.e. in order not to enhance the growth of possible peaks in the histogram, it is proposed to choose the least frequently encountered value of DCT coefficients from the non-embedding area as the value of the quantization step. It should be defined more exactly that the quantization step is a positive integer, while the DCT coefficients are real values; therefore, to select the next quantization step, the DCT coefficients from the non-embedding area must be taken modulo and rounded (an example is shown in Fig. 2). To determine the least common values of DCT coefficients, it is necessary to construct a histogram of the distribution of their converted values over the non-embedding area.

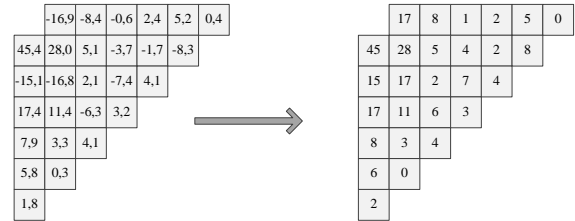


Fig. 2. Transformation of the non-embedding area.

The authors of the study considered two options for choosing the quantization step based on the obtained histogram of the non-embedding area:

- from a group of values with a frequency not exceeding the set one;
- from all values of the non-embedding area.

To implement the first option, it is necessary to pre-set the threshold for the frequency of occurrence of rounded absolute values of the DCT coefficients, among which the quantization step will be selected. Then the values that occur no more often than the quantity of the threshold value will form a group. The quantization step will be equal to the smallest value in the group. The second option does not require explicitly setting a threshold value. In this case, the smallest of the most rarely found values in the block is selected as the quantization step.

These approaches are somewhat similar, since in both cases the decision to select the quantization step is made using the threshold value of the frequency of occurrence of DCT coefficients. However, the fundamental difference between them is that the first option operates on a single threshold value for all image blocks, while the second option uses different threshold values for different blocks.

It is empirically found that the choice of a quantization step of less than three provides a very small capacity, insufficient for effective operation, and the selection of a quantization step of more than twenty leads to a significant deterioration in the visual quality of images, therefore we introduce the condition: $3 \leq q \leq 20$.

Fig. 3 shows an example of selecting a quantization step q over the non-embedding area for both variants. In the first case, the choice of q is made according to a group of values, the

occurrence frequency of which should be no more than two. The smallest among these values is three, therefore $q = 3$. In the second case, the lowest frequency of occurrence of individual values should be firstly determined, in this case it is equal to one. Then from the values encountered only once, the smallest is selected, and in the end $q = 7$.

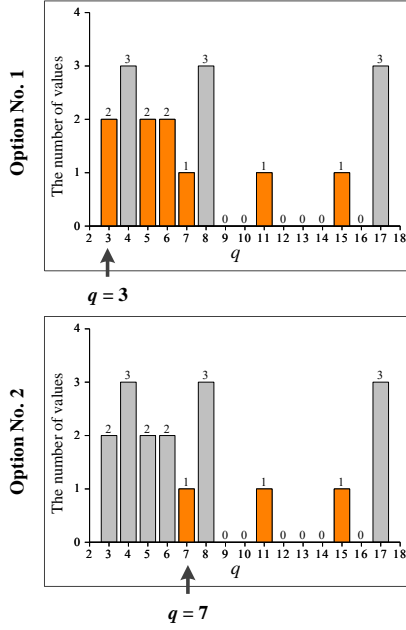


Fig. 3. The choice of the quantization step over the non-embedding area.

4. The results of the experiments

To evaluate the effectiveness of the proposed approach and to compare the two options described above, computing experiments were conducted. For the experiments, 20 classic test images of 512×512 pixels in grayscale were taken from the “USC-SIPI Image Database”.

When evaluating the effectiveness, such standard characteristics as capacity is used, i.e. the ratio of the number of embedded bits to the size of the container, and the PSNR metric, which is calculated by the formula

$$\text{PSNR} = 20 \log_{10} \left(\frac{255}{\text{RMSE}} \right),$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - Q_i)^2},$$

where n is the total number of pixels, P_i is the pixel value of the cover image, Q_i is the pixel value of the stego-image.

Since when choosing a quantization step over a group, its value is on average less than when choosing it over the entire non-embedding area, the algorithm capacity when using the first option is much lower. So, the maximum capacity for the first option is an average of 201 bit/pixel over 20 test images, with the corresponding average PSNR of 36.19 dB, while for the second option, the average capacity is 0.34 bits/pixel with a PSNR of 35.63 dB. The comparison of the histogram of the container, the histogram of the stego-image obtained with a constant quantization step $q=8$ at the maximum capacity, and the histogram of the stego-image obtained using the second option at the maximum capacity is shown in Fig. 4. The form of the histogram of the stego-image obtained using the selection of q over the non-embedding area is close to the form of the histogram of the container. However, for a correct assessment of the effectiveness of the proposed approach and both its variants, the

capacity in all cases was further set equal to the capacity obtained in the case of a selection over a group.

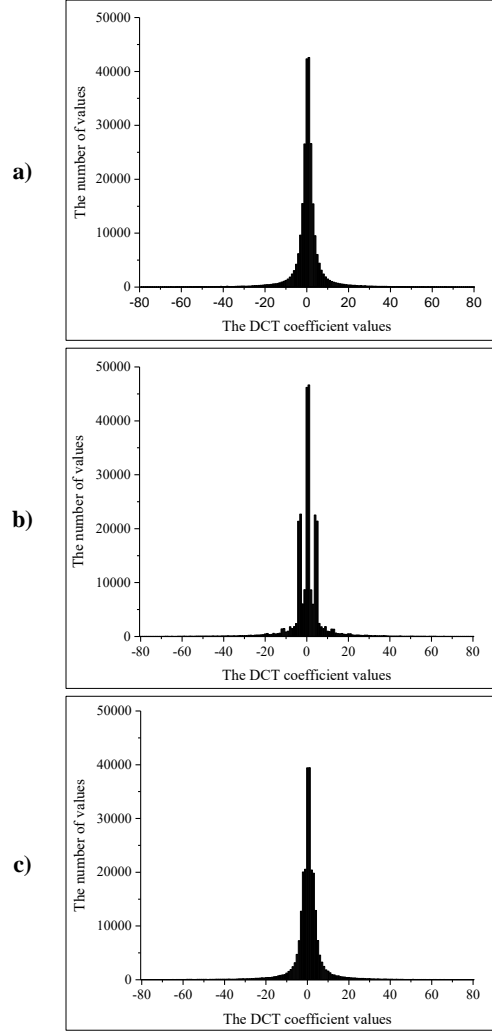


Fig. 4. Histograms for the “Airplane” image: a) container; b) stego-image (constant q , maximum capacity); c) stego-image (new approach, option No. 2, maximum capacity).

Fig. 5 and Fig. 6 present the results of comparing the histograms for the “Airplane” and “Baboon” images obtained for the cover and stego-images with different options for choosing the quantization step. For the “Airplane” test image, the embedding capacity is 0.29 bit / pixel, for the “Baboon” test image, it is 0.11 bit / pixel. These images are selected as examples, as they belong to different types: the “Airplane” image contains many one-colour areas, while the “Baboon” image has a high degree of detail.

According to Fig. 5 and Fig. 6, the application of the proposed approach significantly reduces the distortions that occur when embedding information. The histograms shown in Fig. 5 (b) and Fig. 6 (b) unambiguously show the presence of a steganographic embedding in the corresponding images ($q=8$) while histograms corresponding to stego-images obtained with the proposed approach conform with the natural form of the histograms of cover images. So we can conclude that the proposed approach can reduce the distortions of the digital image natural model in the frequency domain, and therefore, increase the imperceptibility of embedding.

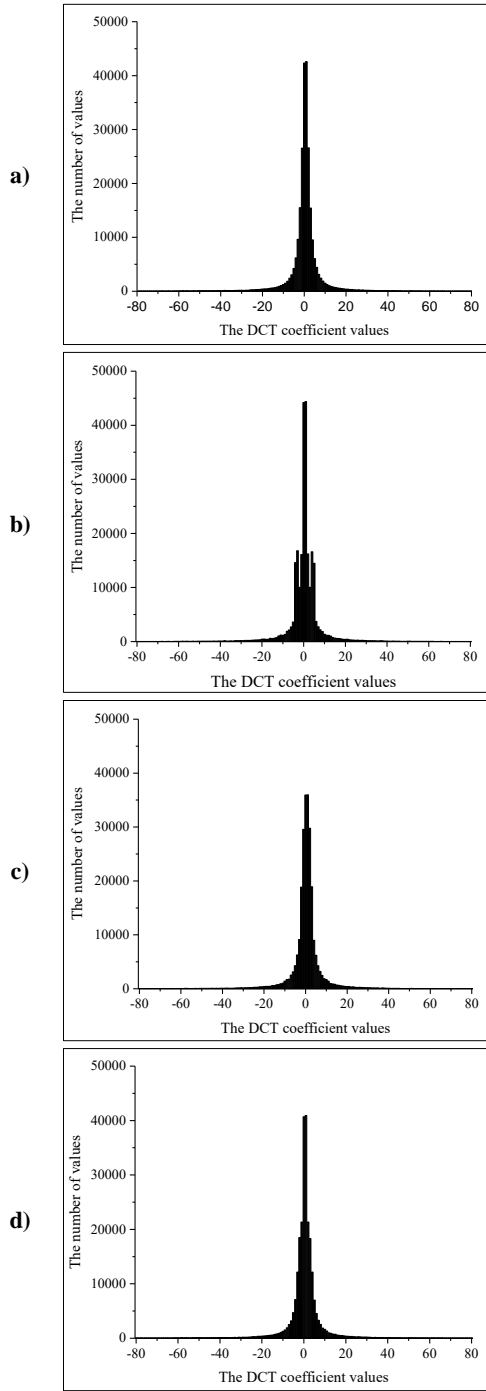


Fig. 5. Histograms for the “Airplane” image: a) container; b) stego-image (constant q); c) stego-image (new approach, option No. 1); d) stego-image (new approach, option No. 2).

To evaluate the differences of the histograms numerically, the value of the RMSE metric between the histograms of the cover images and the corresponding stego-images was calculated. On average, for 20 test images, the RMSE value between the histograms of cover and stego-images obtained with a constant $q = 8$ was 1017,80, for the “Airplane” and the “Baboon” images it was 513,13 and 268,55 respectively. The results for the first option of the variable q : 565.81 on average, 799.21 for “Airplane”, 184.76 for “Baboon”. Results for the second variant of the variable q : 598.06 on average, 680.40 for the “Airplane”, 288.72 for the “Baboon”. It can be concluded that the differences between the histograms of cover and stego-images are much smaller when using a variable q . The choice of

quantization step over a group (the first option) demonstrates an average RMSE value less than for all values (the second option), but the difference between them is not significant. It is also worth noting that the second option allows you to provide a larger embedding capacity, so the first option is preferable with a small embedding volume, but you should use the second option if you need to embed a larger message.

The analysis of the presented histograms showed that, on the one hand, they do not contain characteristic peaks, and, on the other hand, the proposed approach allows us to restore exactly the natural form of the initial histograms and does not lead to excessive “uniformity”, which could also become an unmasking sign.

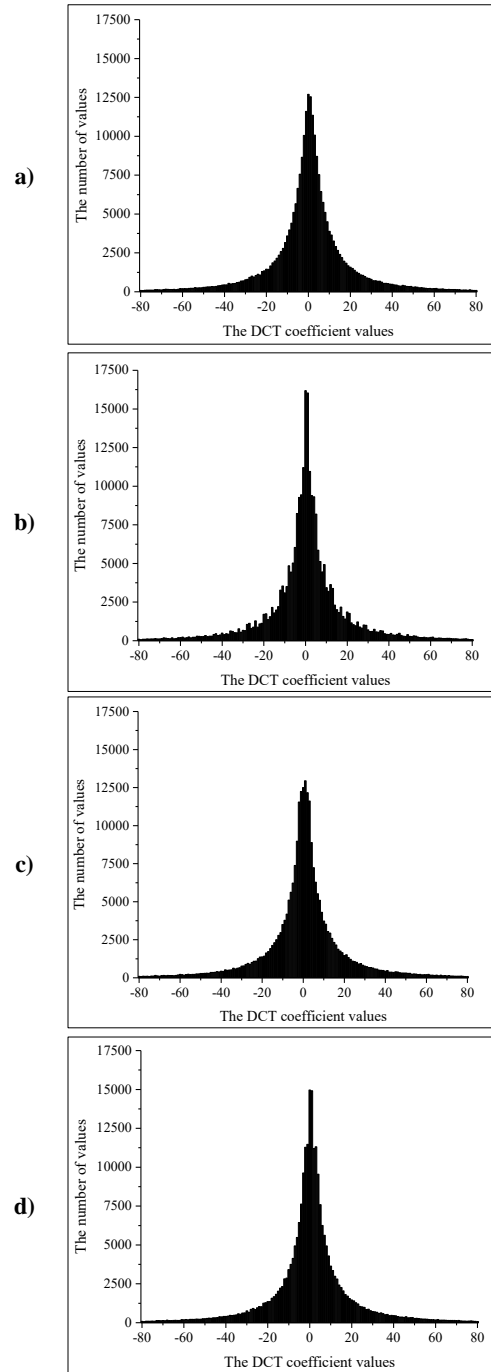


Fig. 6. Histograms for the “Baboon” image: a) container; b) stego-image (constant q); c) stego-image (new approach, option No. 1); d) stego-image (new approach, option No. 2).

5. Conclusion

The article presented and investigated a new approach to reducing the distortions of a digital image natural model in the DCT domain when embedding information using the QIM method. As it can be seen from the results of the experiments, the application of this approach has a positive effect on reducing the unmasking signs of the embedding in the frequency domain. In the future, it is planned to continue work to reduce the distortions of the natural model of images in the frequency domain by adapting the embedding parameters to a specific container.

6. Acknowledgments

This work was financially supported by a grant from the Russian Foundation for Basic Research and the Tomsk Region in the framework of project No. 19-47-703003 and financially supported by a grant from the Russian Foundation for Basic Research in the framework of project No. 18-29-22104.

7. References

- [1] Evsutin O.O. Steganographic embedding of information into the frequency domain of the DCT of digital images using the QIM method with variable quantization step / O.O. Evsutin, A.S. Kokurina, R.V. Meshcheryakov // Proceedings of the 28th International Conference on Computer Graphics and Vision «GraphiCon 2018». – 2018. – Russia, Tomsk. – P. 297-300.
- [2] Denemark T. Steganalysis Features for Content-Adaptive JPEG Steganography / T. Denemark, M. Boroumand, J. Fridrich // IEEE Transactions on Information Forensics and Security. – 2016. – Vol. 11(8). – P. 1736-1746.
- [3] Evsutin O. The adaptive algorithm of information unmistakable embedding into digital images based on the discrete Fourier transformation / O. Evsutin, A. Kokurina, R. Meshcheryakov, O. Shumskaya // Multimedia Tools and Applications. – 2018. – Vol. 77(21). – P. 28567-28599.
- [4] Fridrich J. Feature-Based Steganalysis for JPEG Images and its Implications for Future Design of Steganographic Schemes // Lecture Notes in Computer Science. – 2004. – Vol. 3200. – P. 67-81.
- [5] Holub V. Universal Distortion Function for Steganography in an Arbitrary Domain / V. Holub, J. Fridrich, T. Denemark // EURASIP Journal on Information Security. – 2014. – № 1. – P. 1-13.
- [6] Kodovsky J. Ensemble Classifiers for Steganalysis of Digital Media / J. Kodovsky, J. Fridrich, V. Holub // IEEE Transactions on Information Forensics and Security. – 2012. – Vol. 7(2). – P. 432-444.
- [7] Mitekin V. A new QIM-based watermarking algorithm robust against multi-image histogram attack / V. Mitekin, V. Fedoseev // Procedia Engineering. – 2017. – Vol. 201. – P. 453-462.
- [8] Sedighi V. Content-Adaptive Steganography by Minimizing Statistical Detectability // V. Sedighi, R. Cogranne, J. Fridrich // IEEE Transactions on Information Forensics and Security. – 2016. – Vol. 11(2). – P. 221-234.