A. S.Vasyura, Cand. Sc.(Eng), Prof.; Y.A. Zolotavkin

DETECTION OF HIDDEN CONTENT IMAGES, COMPRESSED BY FRACTAL ALGORITHM

The criterion for detecting of hidden data in fractal code of images is suggested in the article. The approach consisting in estimation of steganographic threat of every record in the fractal code is used. This estimation is based on peculiarities of fractal compression. The number of positively estimated blocks is used for further general conclusion about the possibility of built-in data obtaining. The efficiency of the developed detector is found out experimentally and compared to existing ones.

Keywords: *steganography, information hiding, fractal image compression, detection, stegoanalytic criterion.*

Introduction

Steganography of images is unique branch of the information protection sphere and its efficiency is confirmed by rapid development. This is partially due to the popularity of cryptographic protection methods and to the necessity of secret key storing. Invisible transmission of secret messages also makes great part of practical facilities. Image steganography is widely applied in the copyright protection sphere, this fact result to separation of research direction and introduction of digital watermarks (DW).

Each practical task requires some specific approach to its solution but, in many cases steganographic efficiency is provided by the combination of confidentiality and robustness. Level of above-mentioned characteristics should be correlated with the amount of built-in data [1].

Methods of providing necessary level of resistance are described in literature [2]. Unlike the robustness the problem of confidentiality is ambiguous as a result of large number of disclosing features. This explains the complexity of the development of absolutely adequate detector. But relative index of correct detecting allows to estimate the steganalythical criterion efficiency. Resistance to disclosure supposes possessing the criterion better than that of the interceptor.

Though steno detectors diversity their common structural feature is a binary classifier employing specific characteristics sensitive to the presence of hidden content. Characteristics sequence of each image can be showed by vector, as in [3, 4], where classification was carried out by means of the support vector method (SVM). But to improve detection statistics the classifier training should take place within sample representing two image subsets: the first being formed by original images, the second – by stego images. Efficiency is determined by conformity degree of building in parameters in the training set to real parameters of test stegosystem. Steganalysis quality is reduced if the parameters of real stegosystem are unknown or if they vary in building in process.

The second special feature of stegocriterion is characteristics' structure, which have to cause the separation of stego images from original ones. To achieve this we should take into consideration characteristics inherent to a certain image format, which are determined by processing algorithm.

The compression algorithms are widely used due to significant achievements in the sphere of image processing. This explains the fact that the majority of image representation modern formats of image representation provide compression with losses. The formats files of compressed images are seldom subjected to additional processing and compression, therefore, they provide greater robustness with building-in data. On the other hand for the majority of compression algorithms with losses it is difficult to estimate changed peculiarities of original image, this fact favours information hiding.

Fractal algorithms provide the convenient relation between compression coefficient and quality and possess unique detailing feature while random scaling [5, 6]. The development of fractal compression makes formats on its basis rather popular (e.g. STING), this fact proves the expedience steganographic usage of their.

Principle of Fractal Image Compression

The account of particularities of image processing algorithm is important for determination of characteristics and further detection of hidden content. To solve this problem we have to consider the principle of fractal image compression.

Fractal archiving is based on image representation in compact form by means of iterated function system (IFS) coefficients. IFS – is a set of 3-D affine transformation converting one image into another. Points in 3-D space (2-D space of plane image and brightness) are subjected to transformaton.

Let the metric space of digital images be set by the pair (M, d), where d – given metric measure. To compress the image $I \in M$ it is necessary to find the reflection $\tau : M \to M$ meeting the following conditions:

$$\exists\, 0 < z < 1, \quad \forall \mu, \nu \in M, d(\mu, \tau(\nu)) \le z.d(\mu, \nu), \tag{1}$$

$$d(I, \tau(I)) \cong 0, \tag{2}$$

where μ and ν – different fragments of the image I.

Then, at condition of even division

$$\forall \mu_i \in I, i = \overline{1, n}, \quad I = \bigcup_i \mu_i, \quad \mu_i \cap \mu_j = 0, i \ne j, \tag{3}$$

and existance of image totality $T = \{\tau_i\}$, the ones that $d(\mu_i, \tau_{m,i}(\nu_m)) \le \varepsilon$, the following expression is valid

$$d(I, F^m(T)) \le \frac{\varepsilon}{1 - z}, m \to \infty, \tag{4}$$

where $F^l(T) = \bigcup_i (\mu_i^l \leftarrow \tau_{m,i}(\nu_m^{l-1}))$.

The reflection $\tau_{m,i}$ is an affine transformation and

$$\tau_{m,i} = N_{m,i} \circ S_{m,i} \circ G_{m,i}, \tag{5}$$

where G –operator of geometrical part providing the compression with the coefficient z, turns at specific angles and symmetric reflection of image fragments; S – transfer operator implementing the shift of each elementary part of image fragment in 2-D space; N – intensity operator of image fragment ν_m varying the intensity of each elementary part e_m (pixel) in the following way: $N_{m,i}(e_m) = s_{m,i} \cdot e_m + o_{m,i}$, where s – scaling, and o – brightness [5, 6].

In practice the number of iterations m is limited by small number sufficient for ensuring visual similarity if ε is satisfactory. Metric space (M, d) is determined by the method of division into rank and domain blocks μ and ν correspondently. Usually the image I is of rectangular form, ranke and domain blocks are squares with dimensions $k \times k$ and $2k \times 2k$ pixels, $z = 0{,}5$ and

$$d(\mu_l, \mu_m) = \sqrt{\sum_{i=1}^{k} \sum_{j=1}^{k} (e_{i,j}^{\mu_l} - e_{i,j}^{\mu_m})^2}. \tag{6}$$

The aim of image fractal algorithm of image compression is the search of the set of transformation T for some image I in the case ε is sufficiently small. The essential restriction of this process is the condition $inf(T) << inf(I)$, where the function inf determines the volume of information, necessary for argument descreption. But in practice interpretation (I, ε) into T is not unambigous. The possibility to manipulate T allows employing steganographic equipment utilizing the diversity of interchangable fragments of real images.

The peculiarities of image division into domain and rank blocks can significantly affect the given diversity of comparison variants. The most spread is the quad-tree diagram shown in, Fig. 1. Such division simplifies the search of correspondences between fragments and ensures great diversity. If, the domain block v_m, which due to the reflection $\tau_{m,i}$, corresponds to the rank one μ_i within the limits of ε is not found, the rank block is divided into four smaller blocks. From the point of view of compression such splitting of rank blocks is not efficient as compared with, for example, HV-division [7]. But for steganographic usage this becomes an advantage: greater number of rank blocks allows to hide more data.

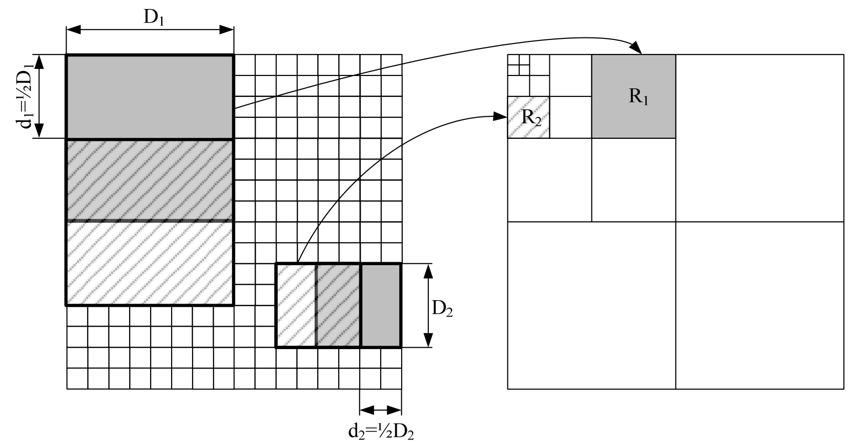

Fig. 1. Illustration of fractal compression by quad-tree diagram

Universality of the square blocks ensures the diversity of their interpretations by means of setting various reflections parameters. This allows to reduce the domain blocks number (increase d_1, d_2) at the expense of a wide set of parameters that is proved for compressing. Steganographic application is limited in certain sense by the peculiarity of above-mentioned division scheme.

Further important practical realization feature of fractal compression model is the method of organization of the search of correspondences between rank and domain blocks. The disadvantage of fractal compression is significant time complexity while processing the variants of blocks comparison [8]. For the solution of this problem the following procedure is to be performed: 1) instead of searching of a domain block reflected into the specific rank one is satisfied with the fist similar one, found within ε; 2) one applies various characteristics of blocks in order to classify them which allows to make search more simple. The first point determines major confidentiality requirements since less importance of domain block chosen for comparison cannot conform to any modification of fractal image compression method. Neglecting this property will result in arising of disclosing feature.

Some modifications of fractal image compression methods differ from described basic approach by the value of scaling coefficient z, by the set of affine transformations and by the operator N [7]. But even an indirect influence of such changes on features of building-in data is insignificant as it concerns only blocks reflection and is not connected with their choice sequence and quantity.

Building-in Data Method Based on Fractal Algorithm

A great variety of steganographic methods on fractal compression principles [9-11] are described in the literature, but greatest robustness is ensured by means of the methods [10, 11] since they directly manipulate the code of compressed image. Building in secrecy increase the given approaches (by means of efficient stegodetector development) will provide high level of protection.

Diversity of interchangeable image fragments allows coding of hidden data. For this purpose in every case of comparison with the rank block all domain-candidates should be described: one should determine the set of candidates and compare with each element of the set corresponding numerical index.

In this case it is sufficient to note, that the losses of secret information are due to the errors of non-correspondence of indices while building-in and restoration. On the other hand the increase of staganographic efficiency can also be achieved by increase of the number of indices.

Maximal amount of different indices equals to the number of elements in the set [12, 13]. In such indexation method the set of candidates should be determined not only in the process of building-in but also in the process of data restoration. The complexity of indices restoration connected with the lack of original image can result in above-mentioned non-conformity and loss of information.

In case of versatile method of indices definition their amount is equal for every set, therefore, they can iterate within one set. Schematic illustration of the two described above methods of indexation is given in Fig. 2. Probably, besides repetition of indices, the lack of elements needed for their interpretation while incorporation (for instance, the fragment with crossfield texture) is one more disadvantage of universal indexation.

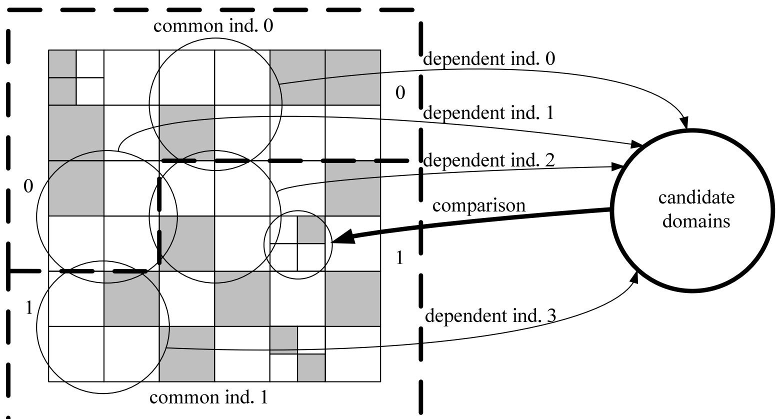

Fig. 2. Methods of Domain Blocks Indexation While Data Incorporation

Realization of methods both of universal and local indexation can differ for various stegomethods but universal method will provide essentially smaller calculation complexity and high reliability of incorporation.

In the process of incorporation to ensure the maximum data throughput on condition of confidentiality and using of universal indexation we should determine the following peculiarities: 1) number of indices; 2) in case of existence of several elements possessing the required index within one set of domain-candidates it is necessary to optimize the choice of one of them; 3) it is necessary to determine the number of steganographic comparisons (part of used rank blocks). Hence, while solving the given problem, the condition of confidentiality is the main restriction to be formalized by means of the criterion.

Stegoanalytical Criterion

The conclusion about the hidden content is suggested to make on the basis of intermediate characteristic of images which presents the number of “suspicious” rank blocks; which are considered to be “suspicious” being obtained despite the fractal compression algorithm. Decision about “suspiciousness” is made on classification basis of conditions (characteristics) of comparison of domain block with the given rank one.

Disadvantage of applying the current criteria on the SVM [3, 4] basis for detection of described stegosystem is the record of only local inter-pixel characteristics. To determine the “suspicious” blocks we suggest a set of characteristics relating to specific features of fractal compression only.

There has been made decision to perform the classification into “suspicious” and “unsuspicious” blocks by means of SVM as well, the efficiency of which is high enough.

Regarding supported conclusion about comparison conformity of definite domain and range blocks with fractal algorithm can be made only on the basis of the whole set of blocks. This can be explained by top-priority requirement of selected domain block and also the necessity to compare the degree of local conformity of rank block environment with the other comparison variants.

In [14] the criterion of distorted areas detection was developed on the basis of characteristics of image fragments self-similarity. To describe the typical image areas the methods of clustering and reducing the dimensional representation were applied. All fragments beyond the clusters were considered distorted.

On the other hand, the verification of domain block top-priority requires determination of the possibilities for comparison of preceding domains. The most complete idea regarding the conditions of "competition" of domain-candidates with the block chosen for comparison can be obtained only on basis of mutual arrangement of their points-characteristics.

The decision about local correspondence and top-priority of each selected domain block can be made on basis of vector-characteristics location of image fragments in features space. But even for small block of 8×8 pixels in size the total number of features equals 64, that makes the description by mutual arrangement of blocks (in their number is great) too bulky to be classified by means of SVM.

To decrease the dimensionality of the characteristic vector $\overline{\mathbf{q}}_{m,i}$ describing the specific features of the domain block v_m comparison into the position of the rank μ_i we suggest to form it of scalar values $q_{m,i}^j$ which are the distances between selected domain block and other fragments. At such simplification the pairwise distances between the rest of fragments are not considered except the fragment being considered. But if the number of fragments grows the characteristic efficiency increases. This is also facilitated by additional set of parameters representing the extent of the relative interpixel non-conformity along the perimeter of blocks.

The conformity degree of the selected domain block v_m with the original range one μ_i, being substituted by domain block, is described by g scalar values $q_{m,i}^j, 1\le j\le g$ – distances between them and g of the most similar blocks of whole image. To confirm the selection priority of the stated block, h of the scalar values $q_{m,i}^j, g+1\le j\le g+h$ – distances between it and h of the most similar domain blocks preceding the selected one are used.

For adequate realization the degree of conformity with the original range block μ_i we suggest to consider the selected domain (transformed) block $\tau_{m,i}(v_m)$ together with environment in the form of a frame f_i around it (Fig. 3) [15]. Then $q_{m,i}^j=\alpha_1 d(f_i,\dot{f}_i^{\,j})+\alpha_2 d(\tau_{m,i}(v_m),\dot{\mu}_i^j), i=\overline{1,n}, 1\le j\le g$, where $\alpha_1\ge1, \alpha_2\le1$ – are constants, $\dot{f}_i^{\,j}$ and $\dot{\mu}_i^j$ – are fragments replicating f_i and μ_i in size and being the constituent parts of the image block $\dot{b}_i^j=\dot{f}_i^{\,j}\circ\dot{\mu}_i^j$ that satisfies the condition

$$\begin{cases}\min\limits_{\dot{b}_i^j} q_{m,i}^j,\ j=1\\ \min\limits_{\dot{b}_i^j\notin\bigcup_{k=1}^{j-1}\dot{b}_i^k} q_{m,i}^j,\ j>1.\end{cases} \tag{7}$$

Using the frame around range block is explained by the fact that incorporation by means of varying dependences between range and domain blocks cannot cover all image blocks. And, since the frame pixels belong to eight different blocks, it is more stable to the distortions, due to incorporation. This specific feature explains choice of the coefficients $\alpha_1\ge\alpha_2$.

Fig. 3 shows that the location $\dot{b}_i^j$ is not limited only by the standard grid. On the other hand, while determining $q_{m,i}^j$ the application of the operators N and G is not provided for. To estimate the degree of similarity between selected domain and original range blocks only natural image properties are used [14].

The characteristic determining top-priority is the sequence of elements $q_{m,i}^{j} = d(\tau_{m,i}(v_m), \tau_{l,i}(v_l)), l < m, g+1 \le j \le g+h$, where index l must not be reiterated under different j and $q_{m,i}^{j} \to \min$. The substantiation of this characteristic is based on the following specific feature: the more domain blocks being located close enough to selected one, the higher probability of its minority.

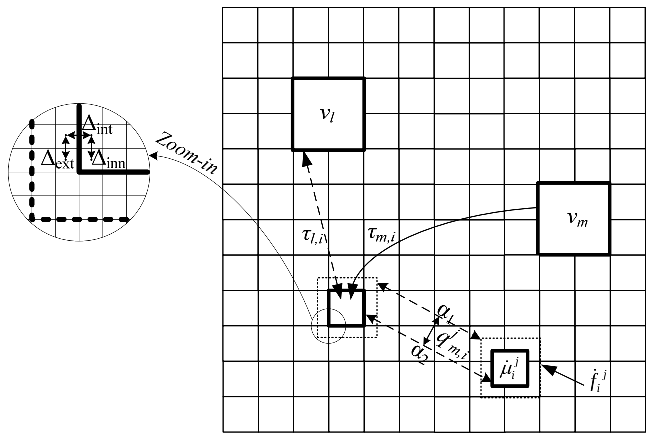

Fig. 3. Determination of Comparison Characteristics

In order to supplement two considered above characteristic types we suggest to consider the specific features of pixel intensity variation on the blocks perimeter. While comparing the difference by modulus between the pixels (Fig. 3) we find out that one of three pixel pairs ensures higher value of this parameter than the other two, therefore, transition in this direction is denoted as sharp (upper index shp). In case of correspondences failure between domain and range blocks this fact is confirmed by higher quantity of sharp transitions between the adjacent blocks. Application of sharp transitions quantity $\Delta_{\text{int}}^{shp}$ is reasonable for representing of coordination of block with its environment. Applying this characteristic to the blocks claimers $v_l, l < m$ for the place of ranking one μ_i allows to specify the degree of local coordination with the given image area and increase the quality of conclusion about top-priority. It is supposed to take into account this characteristic for determination of local coordination of domain block v_m.

The characteristic vector $\overline{\mathbf{q}}_{m,i}$ describing the specific features of comparison $\tau_{m,i}(v_m) \to \mu_i$ for further to perform further classification by means of SVM can structurally be represented in the following way:

$$\left\{ q_{m,i}^{j} = \alpha_1 d(f_i, \dot{f}_i^{j}) + \alpha_2 d(\tau_{m,i}(v_m), \dot{\mu}_i^{j}), 1 \le j \le g \right\} \cup$$
$$\cup \left\{ \begin{array}{l} q_{m,i}^{j} = d(\tau_{m,i}(v_m), \tau_{l,i}(v_l)) \\ q_{m,i}^{j+h} = \Delta_{\text{int}}^{shp}[\tau_{l,i}(v_l)] \end{array} \middle| \, l < m, g+1 \le j \le g+h \right\} \cup \quad (8)$$
$$\cup \left\{ q_{m,i}^{g+2h+1} = \Delta_{\text{int}}^{shp}(\tau_{m,i}(v_m)) \right\}.$$

Advantages of this characteristic are high sensitivity to violation of actions order provided for fractal image compression. Calculation complexity of training and classification by means of SVM can be decreased as compared to [3, 4], since the vector length $\overline{\mathbf{q}}_{m,i}^{j}$ makes $g + 2h + 1$ and, on condition $g = h = 10$, the necessary sensitivity for detection of hidden content is ensured (if 50% of rank blocks are used steganographically).

Experiment

The main aim of the experiment is determination of critical amount Q_S of „suspicious"blocks in the image, exceeding of which allows to make a conclusion that incorporation has been performed. To do this we should train the SVM using the described features set combined into characteristic vector. While training various images with hidden data should be used, the data being incorporated by means of described steganographic pattern as well as the image with no hidden content. The result of testing is not affected by the amount data, incorporated into stegoimage since each block is analyzed separately.

Optimization of threshold value Q_S is aimed at minimization of detection entropy. Task of detection is limited by the pre-defined incorporation pattern that provides invariant and known amount of hidden data in stego image. Comparative analysis of detection results of images mixed set on basis of the criteria [3, 4] and suggested criterion allows to evaluate its efficiency.

In order to classify blocks as „suspicious" or "unsuspicious" for training SVM 200 greyscale images 256×256 in size have been used. Amount of incorporated data into the images varied within the range from 300 to 1000 bits but the total number of range blocks agreed with fractal algorithm was equal to that of range blocks steganographically modified. In forming the feature vector $g=h=10$, its total length was 31, $\alpha_1=1.2, \alpha_2=0.8$, the frame thickness – 3 pixels. The classification efficiency determined in the result of testing was 70.3%.

When calculating Q_s for detection of stego images with the amount of hidden data equivalent to 400 bits two dependences have been used, the graphs of which are shown in Fig. 4. The first one shows the number of original images with corresponding amount of "suspicious" blocks. The second dependence shows the number of stego images with definite amount of "suspicious" blocks. Optimum value Q_S should ensure the minimum of crosshatched space corresponding to intersection point of the gaphs. For these incorporation conditions $Q_s=354$. Therefore, if the number of "suspicious" blocks exceeds Q_S the image is considered to be steganographically significant. The resultant probability of reliable detection on the basis of suggested steganalythical criterion reached 97.1%.

To verify detection efficiency by means of the criteria [3, 4] the training SVM for each of them has been occurred on training set, the half of which being stego images with hidden data amount 400 bits. The probability of reliable detection reached 56.7% and 52.5% for [3] and [4] correspondingly. This fact confirms significant advantage of suggested method of hidden content detection in fractal image code over the current criteria.

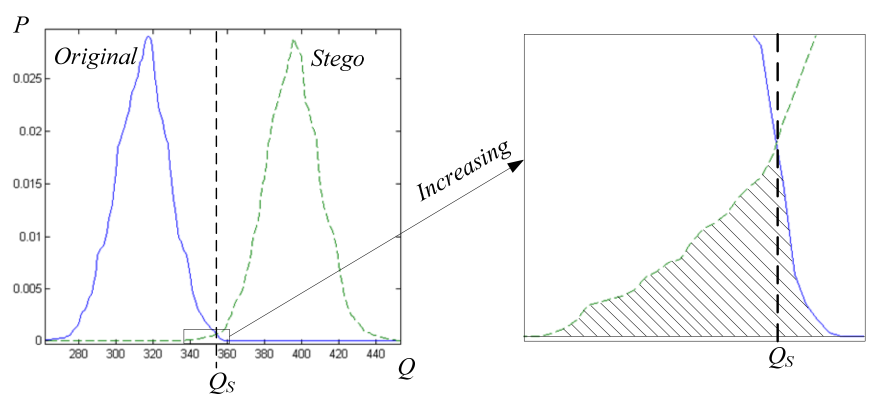

Fig. 4. The Hidden Content Detection Result

Conclusions

In the article the steganographic criterion intended for detection of images with hidden content built-in at the stage of compression by means of fractal algorithm is suggested. In order to synthesize the efficient criterion specific features of fractal compression that vary as the result of data incorporation were determined. On the basis of stated features the characteristics of individual image blocks were formed in order to perform further binary classification by means of SVM. The amount of “suspicious” blocks obtained as a result of classification is applied for establishing hidden content presence by means of comparison with the critical threshold Q_S.

Such method provides the considerable increase of stego images detection efficiency, subjected to incorporation as a result of sequence variation of operations specified by fractal compression algorithm. Due to application of suggested characteristics the accuracy of classification into “suspicious” and “unsuspicious” blocks reached 70.3% which allows to detect stegoimages with the amount of hidden data as small as 400 bits with accuracy of 97.1%.

On the other hand, the main disadvantage is the necessity to characterize and classify each image block. But small dimensionality of characteristic vector and parallel organization possibility of uniform calculations suggests efficient solution of the problem.

Study of adaptive methods of data incorporation into fractal code of images and increase of their efficiency is to be considered in further investigations.

REFERENCES

1. Johnson, N., Duric, Z., Jajodia, S. Information Hiding: Steganography and Watermarking - Attacks and Countermeasures. – New York: Kluwer Academic Pub., 2000. – 200p.

2. Грибунин В. Г., Оков И. Н., Туринцев И. В. Цифровая стеганография. – СПб.: Солон-Пресс, 2002. – 272с.

3. Zou D., Shi Y., Su W., Xuan G. Steganalysis based on Markov Model of Thresholded Prediction-Error Image // IEEE ICME Conference Record, 2006. – P. 1365-1368.

4. Chen, X., Wang, Y., Tan, T., Guo, L. Blind Image Steganalysis Based on Statistical Analysis of Empirical Matrix // IEEE ICPR’06, 2006. – № 3. – P. 1107-1110.

5. Barnsley, M., Hard, L. Fractal Image Compression. – Wellesley: A.K. Peters, Ltd., 1993. – 256p.

6. Уэлстид С. Фракталы и вейвлеты для сжатия изображений в действии. – М.: Триумф, 2003. – 320с.

7. Zhao, E., Liu, D. Fractal Image Compression Methods: A Review // ICITA’05, 2005. – P. 756-759.

8. Polvere, M., Nappi, M. Speed-up in Fractal Image Coding: Comparison of Methods // IEEE TIP, 2000. – № 6. – P. 1002-1009.

9. Bas, P., Chassery, J. M., Davoine, F. Using the Fractal Code to Watermark Images // Proc ICIP’98, 1998. – № 1. P. 469-473.

10. Li, C., Wang, S. Digital Watermarking Using Fractal Image Coding // IEICE Trans. Fund., 2000. – № 6. – P.1286-1288.

11. Liao, P., Chen, C., Chen, C., Pan, J. Interlacing Domain Partition for Fractal Watermarking // IIH-MSP’06, 2006. – P. 441-444.

12. Васюра А.С., Золотавкін Є.А., Лукічов В.В. Адаптивний метод вбудовування даних у фрактальний код зображень // Інформаційні технології та комп'ютерна інженерія. – Вінниця, 2006. – № 2. – С.105-110.

13. Васюра А.С., Золотавкін Є.А. Визначення та забезпечення стійкості методу таємної передачі даних на основі фрактального стиснення зображення // Вісник Хмельницького національного університету. – Хмельницький, 2007. – №2 .– С. 133-138.

14. Amano, T. Correlation Based Image Defect Detection // IEEE ICPR’06, 2006. – P. 163-166.

15. Liang, X., Chenrong, H., Haijun, L., Huizhong, W. Concealment of Damaged Block Coded Images Using Intelligent Two-Step Best Neighborhood Matching Algorithm // Proc CGIV’05, 2005. – P. 38-42.

Anatoliy Vasyura – Director of the Institute, Professor of the Department of Automation and Information Measuring Devices;

Yevhen Zolotavkin – Post-Graduate Student of the Department of Automation and Information Measuring Devices;

Vinnytsia National Technical University