Object recognition based on fractal coding using domain blocks dictionary

E Y Minaev

1Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

e-mail: eminaev@gmail.com

Abstract. This article investigated the method of fractal compression for obtaining image descriptors adapted for use on mobile platforms, optimizing the performance and the amount of stored fractal images. The main idea is to apply the fractal compression method based on iterated function systems to reduce the dimension of the original images. The proposed method is more preferable than the existing methods of fractal compression by the number of computational operations and by the volume of stored data for compressed images. Also, a new concept of fractal coding using domain blocks dictionary is proposed in this article.

1. Introduction
The problem of using existing fractal compression algorithms on mobile hardware and software platforms is noted in [1]. Traditionally, fractal compression methods have high computational complexity, and methods and algorithms for optimizing performance developed for desktop hardware platforms are not always applicable for mobile platforms [2] [3]. Modern performance solutions are based on the use of user-programmable gate arrays (FPGAs) and the use of GPUs, which makes it difficult to use these approaches for most mobile platforms. At the same time, the urgency of using fractal compression methods for mobile devices is emphasized in the article [4]. Among the existing compression methods, the most interesting for mobile platforms are algorithms "without search" [5] and methods using predefined sets of domain blocks [6]. For individual tasks, for example, regarding the recognition of objects on radar images in mobile systems of unmanned aerial vehicles, the computational complexity and the amount of stored data play a decisive role.

2. Adaptation of fractal compression for mobile platforms
This article explores the method of fractal compression for obtaining image descriptors, adapted for use on mobile platforms, optimizing the performance and the amount of stored fractal images. The main idea of the method is to apply fractal compression on the basis of iterated function systems to reduce the dimension of the original images [7]. Due to the narrower specialization for radar image recognition problems, the proposed method is preferable to the existing methods of fractal compression [8], the number of computational operations, and the volume of stored data for compressed images. To realize the recognition technology on mobile platforms, the volumes of stored data and the computational complexity of the recognition algorithms are crucial, as a basis. It is proposed to use the previously developed information recognition technology, including the stage of...
fractal compression of the original images and the stage of classification of the obtained fractal images, using the criterion of class separability according to the conjugacy index (Figure 1) [9]. The main modification of the technology affects the stage of fractal image compression as the most resource-intensive process of the whole information technology.

The original images in this technology are divided into square non-overlapping parts, named range blocks, and into larger square parts, named domain blocks. The scheme of image domain and range block partitioning is shown in Figure 2. Compression algorithm searches the best affine transformation from domain to range block for every range block. The computational complexity of full search, the volume of stored information, and the quality of fractal images obtained directly depends on the number of rank and domain blocks. The main goal of the research in this paper is to find a partition with a minimum number of blocks, which preserves the acceptable quality of recognition.

In the development of the proposed method, this article also offers a concept of new information technology for object recognition based on the use of the domain block dictionary. New technology assumes the formation of support subspaces at the stage of fractal image compression of the training sample, and a significant acceleration of the classification stage of the unknown image. Among the existing publications in which the use of the dictionary is a key stage of the technology, it should be noted [6] in which the authors propose to form a dictionary from the fractals of Mandelbrot and Julia sets. Such a scheme significantly speeds up the process of image compression, but for the technology of object recognition on images, the formation of a dictionary directly from the training samples is more preferable.

3. Classifier training
When we trained the classifier in Figure 1, the main problem was that the images forming the training sample of one class were compressed independently of each other, and were combined together only at the stage of construction of the support subspaces. At the same time, the recognition stage raises problems associated with the possible intersection of the support subspaces. Accordingly, it is necessary to apply methods that provide spatial separability, which further increases the computational complexity. In [10], a fractal compression scheme using several different images is proposed. In this article, it is proposed to apply this scheme on cyclic sequence of images from the training set, with the formation of a dictionary of rank and domain blocks and the corresponding transformations. Classic compression IFS algorithm searches the best affine
transformation from domain to range block for every range block (Figure 2). As a result, an input image is coded by several affine transformations:

\[ I' = F(1) = C_{1,4}1 + C_{5,6}, \]

where \( I' = (i', j')^T \), \( I = (i, j)^T \) is the coordinates of pixel from domain and range block accordingly,
\[ C_{1,4} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}, \quad C_{5,6} = \begin{bmatrix} c_5 \\ c_6 \end{bmatrix} \] – transformation coefficients, \( u_{i,j}^* = u_{i,j} \) is the pixel brightness from range and domain area, \( a, c_7, c_8 \) – contrast and brightness shift parameter.

We use eight different sets of parameters for fractal image transformation:
\[ C_{1,4} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}, \quad \begin{bmatrix} 0 & -0.5 \\ 0.5 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0.5 \\ -0.5 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0.5 & 0 \\ -0.5 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}, \quad \begin{bmatrix} 0 & -0.5 \\ -0.5 & 0 \end{bmatrix}. \]
\[ c_5, c_6 \] – shift coefficients of affine transformations.

These parameters correspond to different kinds of transformations, such as rotation, domain area mapping and compression with a rate of 0.5.

The transformation is conducted in a class of contraction mapping to obtain a unique and stable fractal image (the maximum of the transformation matrix eigenvalue is less than 1). Parameters of transformations \( c_1 - c_8 \) are computed by IFS fractal compression algorithm: \( c_1 - c_4 \) are selected from the possible sets, \( c_5, c_6 \) are calculated in the process of searching the best affine transformation from domain to range block, \( c_7, c_8 \) – are calculated on the average brightness of domain and range blocks.

Set of transformations for every range block can be written as:

\[ I_1 = \bigcup_{i} F_i(I_0). \]  

Using the Hutchinson operator, it can be written shortly as:

\[ I_1 = F I_0, \]

where \( I_0 – \) initial image, \( F – \) Hutchinson operator, representing set of affine transformations, \( I_1 – \) result image. The scheme of cyclic sequence of transformations for several images of training set is presented in Figure 3.

**Figure 3.** Scheme of cyclic sequence of transformations.
After searching the best affine transformation from domain to range block for every range block, we can compose the dictionary, including information concerning class number, range blocks division, images of range, and domain blocks with transformation coefficients (Figure 4), and with every class of image with different range blocks division training independently.

4. Classification process

Using the dictionary, we can realize the fractal coding of input images by this procedure. At first, the input image is divided into square non-overlapping range blocks. Then, for every range block we search similar appropriate range blocks with domain block and transformation for every class. As a result, we obtain a set of transformations and its initial data for every class of images. Using Hutchinson operator, it can be represented as $F_1^*, F_2^*, ..., F_m^*$, for $m$ classes. The distance between input image and each class can be written as:

$$D_i = \frac{d(F_i^* I', I^*)}{I_{w_i} I_{h_i}^*} = \left\| F_i^* I' - I^* \right\|,$$

where $I'$ – initial image, $I_{w_i}, I_{h_i}$ – width and height of initial image, $F_i^*$ – Hutchinson operator for transformations of $i$ class, $d$ - Euclidean norm. Class with minimal distance to input image is the result of classification.

Class 1 training set of images:

Class 2 training set of images:

Class m training set of images:

The details of whole information technology are as follows:

(1) Classifier training. For images set representing one of the classes, we obtain an acyclic sequence of transformations. The results are written to dictionary.

(2) Repeat step (1) for all classes, and all variants of range blocks division ($4\times4$, $8\times8$, etc.) for multi-scale recognition.
(3) Input test image is divided into square non-overlapping range blocks. For every range block, we search similar range blocks with domain block and transformation from the part of dictionary of class and certain range blocks division.

(4) Compute the distance between input image and class.

(5) Repeat steps (3), (4) for all classes, and all variants of range blocks division.

(6) Find class with minimal distance to input test image.

(7) Repeat steps (3), (4), (5), (6) for other input image.

5. Technology adaptation for mobile platforms

To implement the proposed technology in mobile platforms, the following scheme is proposed. The process of training the classifier is the most computationally complex, and the most suitable scheme of work involves the use of a separate computing server to create a dictionary from the sets of training images. The created dictionary of fractal transformations can be used directly on the mobile platform because the recognition stage is performed quite quickly. The process of additional training, for example, adding 1-2 images to the training set, can be realized and implemented both on the server and on the mobile platform. An extended scheme for using the technology on mobile platforms is shown in Figure 5.

5. Experiments results

In our recognition experiments, we used the MSTAR (moving and stationary target acquisition and recognition) public dataset. Objects BMP2, BTR70, T72 were used, and training and test samples were employed for each object from the dataset. These SAR images are collected using an X-band SAR sensor at two different depression angles (15° and 17°). The total number of SAR images in training set is 689, whereas it is 1365 in test set. At the stage of fractal compression, a different number of range blocks were used 16 (4×4), 64 (8×8), 256 (16×16), 1024 (32×32), 4096 (64×64) and accordingly, domain blocks 9 (3×3), 49 (7×7), 225 (15×15), 961 (31×31), 3969 (63×63). Examples of the obtained fractal images are shown in Figure 6.

Table 1 shows the recognition results for different range blocks division: the computational complexity essentially increases with the size of the block. In whole information technology, it means that it is useful to evaluate the recognition process initially at 4×4 range blocks division, and then refine the recognition results on 8×8, 16×16 range blocks division. Also for proof-of-concept, the object recognition technology with fractal coding using domain blocks dictionary was tested on

![Figure 5. Recognition technology on mobile platforms.](image-url)

![Figure 6. Examples of fractal images of an object, a – original image, b, c, d – fractal image for range blocks 32×32, 16×16, 8×8, accordingly.](image-url)
three-class classification task, with objects BMP2, BTR70, T72 (Table 2).

| Number of rangeblocks | Share of correctly recognized objects | Number of operations |
|-----------------------|---------------------------------------|----------------------|
| 4×4                   | 0.651                                 | 1584                 |
| 8×8                   | 0.832                                 | 4312                 |
| 16×16                 | 0.884                                 | 39600                |
| 32×32                 | 0.905                                 | 338272               |
| 64×64                 | 0.906                                 | 2794176              |

Table 2. Multiclass recognition results for 3 objects.

| Class name | Share of correctly recognized objects | Proposed method | Saliency Attention and SIFT[11] |
|------------|---------------------------------------|-----------------|---------------------------------|
| BMP2       | 0.891                                 | 0.64            |                                 |
| BTR70      | 0.882                                 | 0.75            |                                 |
| T72        | 0.904                                 | 0.74            |                                 |

The proposed method was compared with another experimental method [11]. The experimental conditions in this work are quite similar. The recognition quality is satisfactory. Hence, at this stage of the technology development, it is better to use it in limited computing resources applications insensitive to quality; for example, implementation on mobile platforms.

7. Conclusion

The results of the experiments can be interpreted as follows: the optimal ratio of the proportion of correctly recognized objects and the number of computational operations is 256 (16×16) rank areas. However, for some application problems it is possible to use a partition into 64 (8×8) regions: for real-time preliminary recognition with significantly limited computing resources, for example, on ultra-small unmanned aerial vehicles [12], in navigation tasks [13], etc. The proposed technology for using the dictionary allows to additionally reduce the requirements for computing resources on mobile platforms. Average recognition quality on the MSTAR dataset is 0.892, which is comparable to the results in actual papers.

8. References

[1] Srivastava S and Lall B 2015 Superresolution based medical image compression for mobile platforms Workshop on Machine Learning for HealthCare hal-01436138
[2] Chen Dand Singh D 2013 Fractal video compression in OpenCL: An evaluation of CPUs, GPUs, and FPGAs as acceleration platforms Design Automation Conference (ASP-DAC) 18th Asia and South Pacific IEEE 297-304
[3] Son T N, Hoang T M, Dzung N T and Giang N H 2014 Fast FPGA implementation of YUV-based fractal image compression IEEE Fifth International Conference on Communications and Electronics (ICCE) 440-445
[4] Lima V, Schwartz W and Pedrini H 2011 Fast low bit-rate 3D searchless fractal video encoding 24th SIBGRAPI Conference on Graphics, Patterns and Images IEEE 189-196
[5] Furao S and Hasegawa O 2004 A fast no search fractal image coding method Signal Processing: Image Communication 19(5) 393-404
[6] Sun Y, Xu R, Chen L, Kong R and Hu X 2014 A Novel Fractal Coding Method Based on MJ PloS one 9(7) e101697
[7] Minaev E Y and Nikonorov A V 2012 Object detection and recognition in the driver assistance system based on the fractal analysis Computer Optics 36(1) 124-130
[8] Das S and Ghoshal D 2016 Hybrid Fractal Image Compression Based on Graph Theory and Equilateral Triangle Segmentation *International Journal of Applied Engineering Research* **11**(6) 4467-4477

[9] Minaev E and Fursov V 2016 Support subspaces method for fractal images recognition *CEUR Workshop Proceedings* **1638** 379-385

[10] Ozawa K 2008 Dual fractals *Image and Vision Computing* **26**(5) 622-631

[11] Karine A, Toumi A, Khenchaf A and El Hassouni M 2017 Saliency attention and sift keypoints combination for automatic target recognition on MSTAR dataset *Advanced Technologies for Signal and Image Processing (ATSIP) International Conference on IEEE* 1-5

[12] Muraviev V S, Smirnov S A and Strotov V V 2017 Aerial vehicles detection and recognition for UAV vision systems *Computer Optics* **41**(4) 545-551 DOI: 10.18287/2412-6179-2017-41-4-545-551

[13] Konovalenko I, Miller A, Miller Band Nikolaev D 2015 UAV navigation on the basis of the feature points detection on underlying surface *Proceedings 29th European Conference on Modelling and Simulation (ECMS)* 499-505

**Acknowledgments**

The reported study was funded by RFBR according to the research projects No. 17-29-03112-OFI-m, No. 16-07-00729-A.