

An investigation of machine learning method based on fractal compression

E Y Minaev^{1,2}

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

²Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

e-mail: eminaev@gmail.com

Abstract. In this article the method of machine learning with cyclic fractal coding and the use of domain block dictionary, adapted for use on mobile platforms, with optimization of performance and volume of stored fractal images is investigated. The main idea of the method is to use the fractal compression method based on iterated function systems to reduce the dimension of the original images, and to use cyclic fractal coding to represent the class of images. As a result of research of the method it was found that the share of correctly recognized objects on MSTAR averages 0.892, the recognition time averages 254 ms. The achieved results are acceptable for use in mobile platforms, including UAVs and ground autonomous robots.

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].

2. Implementation of machine learning method based on cyclic fractal compression

One of the promising approaches to the implementation of the classifier based on fractal compression is proposed in [5]. When we trained the classifier described in [6], 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 [7], 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 [8] 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:

$$\mathbf{I}^* = F(\mathbf{I}) = \mathbf{C}_{1,4}\mathbf{I} + \mathbf{c}_{5,6},$$

$$u_{i,j}^* = c_7 \cdot u_{i,j} + c_8,$$
(1)

where $\mathbf{I}^* = (i^*, j^*)^T$, $\mathbf{I} = (i, j)^T$ – is the coordinates of pixel from domain and range block accordingly, $\mathbf{C}_{1,4} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}$, $\mathbf{c}_{5,6} = \begin{bmatrix} c_5 \\ c_6 \end{bmatrix}$ – 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:

$$\mathbf{C}_{1,4} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}, \begin{bmatrix} 0 & -0.5 \\ 0.5 & 0 \end{bmatrix}, \begin{bmatrix} -0.5 & 0 \\ 0 & -0.5 \end{bmatrix}, \begin{bmatrix} 0 & 0.5 \\ -0.5 & 0 \end{bmatrix}, \begin{bmatrix} 0.5 & 0 \\ 0 & -0.5 \end{bmatrix}, \begin{bmatrix} -0.5 & 0 \\ 0 & 0.5 \end{bmatrix}, \begin{bmatrix} 0 & 0.5 \\ 0.5 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -0.5 \\ -0.5 & 0 \end{bmatrix}.$$

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:

$$I_1 = \bigcup_i F_i(I_0).$$
(2)

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

$$I_1 = FI_0,$$
(3)

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 1.

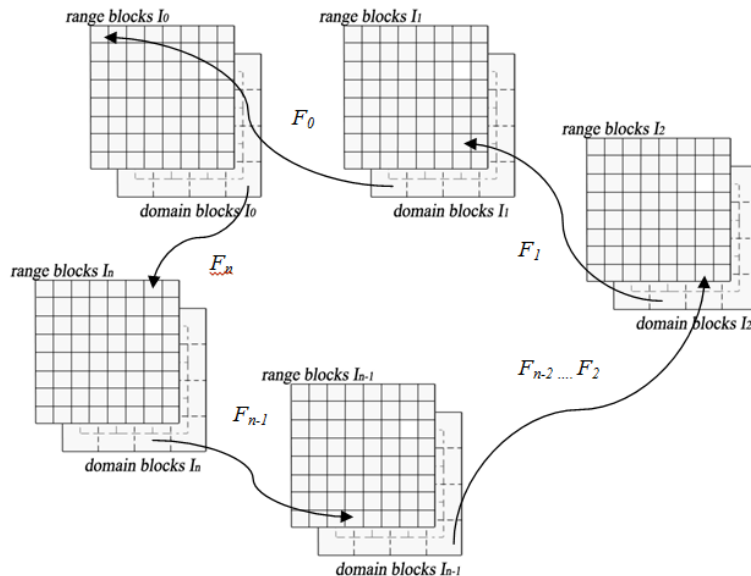


Figure 1. Scheme of cyclic sequence of transformations.

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, and with every class of image with different range blocks division training independently (Figure 2).

3. 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 $F_1^*, F_2^* \dots F_m^*$, for m classes. The distance between input image and each class can be written as:

$$D_i = \frac{d(F_i^* I^*, I^*)}{I_w^* I_h^*} = \frac{\|F_i^* I^* - I^*\|_2}{I_w^* I_h^*}, \quad (4)$$

where I^* – initial image, I_w^*, I_h^* – 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.

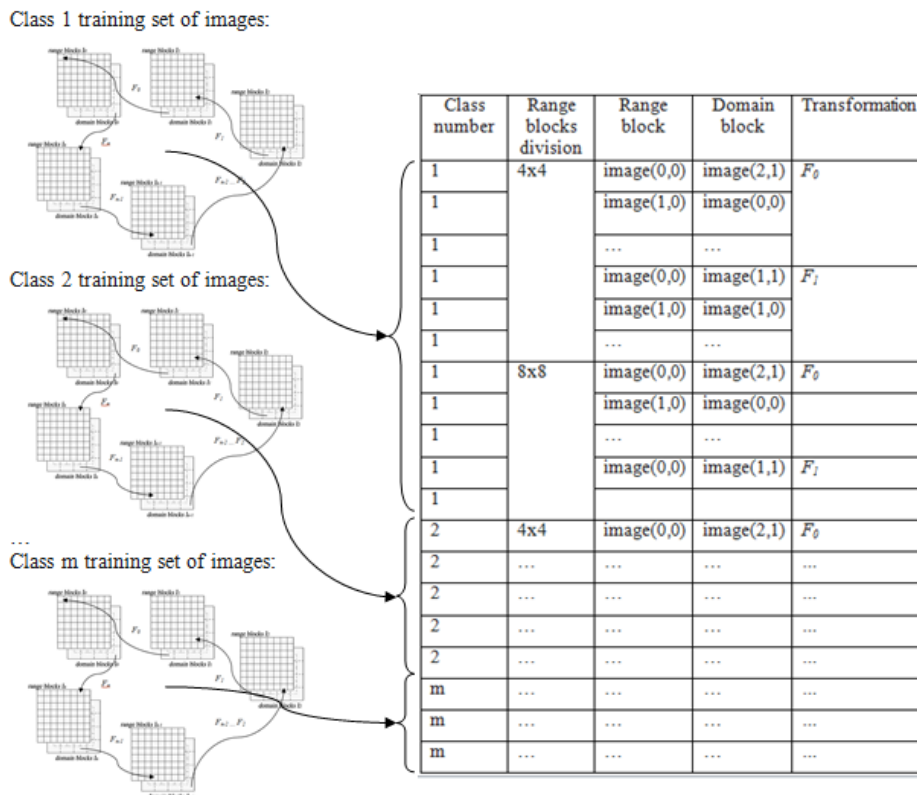


Figure 2. Scheme of training process.

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 (4x4, 8x8, 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.

4. 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), and accordingly, domain blocks 9 (3×3), 49 (7×7), 225 (15×15), 961 (31×31). Examples of the obtained fractal images are shown in Figure 3.

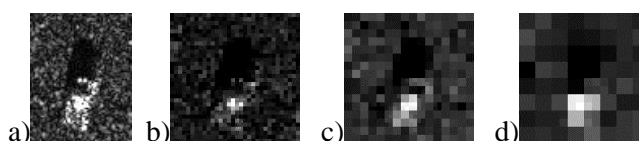


Figure 3. Examples of fractal images of an object, a – original image, b, c, d – fractal image for range blocks 32×32, 16×16, 8×8, accordingly.

An investigation of the object recognition method with cyclic fractal coding using domain blocks dictionary was tested on three-class classification task, with objects BMP2, BTR70, T72 (Table 1).

Table 1. Multiclass recognition results for 3 objects.

Class name	Share of correctly recognized objects	
	Proposed method	Aliency 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 [9]. The experimental conditions in this work are quite similar. The performance of the recognition method on mobile platforms based on the Qualcomm Snapdragon 625, 2 GHz processor was also investigated. It was found that the average recognition time of objects on the MSTAR dataset is 254 ms, which generally corresponds to the speed of processing in real time.

In another experiment we used the MNIST image database. The MNIST database (Modified National Institute of Standards and Technology database) is a large database of handwritten digits that is commonly used for training various image processing systems. The MNIST database contains 60,000 training images and 10,000 testing images (Figure 4). The purpose of the experiment is to show the stability of the object recognition method with cyclic fractal coding using domain blocks dictionary on the data with a large number of instances of the class.



Figure 4. Sample images from MNIST test dataset.

The results of comparison with other classical recognition algorithms (without boosting, preprocessing and a combinations of several methods[10]) are given in the table 2.

Table 2. Multiclass recognition results for 10 objects.

Method name	Test error rate(%)
pairwise linear classifier	7.6
K-nearest-neighbors	2.83
SVM, Gaussian Kernel	1.4
Proposed	2.45

The results show that the proposed method shows comparable results for large databases. A promising direction of future research is to improve the quality of recognition due to additional combinations of methods of boosting, preprocessing, augmentation and reduction of dimensionality.

5. Conclusion

As a result of investigation of the machine learning method with cyclic fractal coding and using the domain block dictionary, it was found that the share of correctly recognized objects on the MSTAR dataset averages 0.892, the recognition time averages 254 ms. The achieved results are acceptable for use in mobile platforms, including UAVs and ground autonomous robots.

6. References

- [1] Srivastava S, Lall B 2015 Superresolution based Medical Image Compression for Mobile Platforms *Workshop on Machine Learning for HealthCare* 01436138
- [2] Chen D, Singh D 2013 Fractal video compression in OpenCL: An evaluation of CPUs, GPUs, and FPGAs as acceleration platforms *Design Automation Conference (ASP-DAC)* 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 *Communications and Electronics (ICCE)* 440-445
- [4] Lima V, Schwartz W and Pedrini H 2011 Fast low bit-rate 3D searchless fractal video encoding *Graphics, Patterns and Images (Sibgrapi)* 189-196
- [5] Minaev E 2018 Object recognition based on fractal coding using domain blocks dictionary *Journal of Physics: Conference Series* **1096(1)** 012099
- [6] Minaev E, Fursov V A 2016 Support subspaces method for fractal images recognition *CEUR Workshop Proceedings* **1638** 379-385
- [7] Ozawa K 2008 Dual fractals *Image and Vision Computing* **26** 622-631
- [8] Sun Y, Xu R, Chen L, Kong R and Hu X 2014 A Novel Fractal Coding Method Based on MJ Sets *PloS one* **9(7)** e101697
- [9] Karine A, Toumi A, Khenchaf A and El Hassouni M 2017 Saliency attention and sift keypoints combination for automatic target recognition on MSTAR dataset *International Conference Advanced Technologies for Signal and Image Processing (ATSIP)* 1-5
- [10] The MNIST Database of handwritten digits URL: <http://yann.lecun.com/exdb/mnist/> (2019-05-25)
- [11] Dmitriev E A, Myasnikov V V 2018 Comparative study of description algorithms for complex-valued gradient fields of digital images using linear dimensionality reduction methods *Computer Optics* **42(5)** 822-828 DOI: 10.18287/2412-6179-2018-42-5-822-828

Acknowledgments

The reported study was funded by RFBR according to the research projects No. 17-29-03112-OFI-m, 19-29-01235-mk, experimental studies - in the framework of the state assignment of the IPSI RAS - a branch of the Federal Scientific-Research Center "Crystallography and Photonics" of the RAS (agreement № 007-Г3/Ч3363/26).