

# Исследование метода машинного обучения на основе фрактального сжатия

Е.Ю. Минаев<sup>1,2</sup>

<sup>1</sup>Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

<sup>2</sup>Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

**Аннотация.** В данной статье исследован метод машинного обучения с циклическим фрактальным кодированием и использованием словаря доменных блоков, адаптированный для применения на мобильных платформах, с оптимизацией производительности и объема хранимых фрактальных образов изображений. Основная идея метода заключается в применении метода фрактального сжатия на основе систем итерированных функций для понижения размерности исходных изображений, и использовании циклического фрактального кодирования для представления класса изображений в целом. В результате исследований метода получено, что доля верно распознанных объектов по базе MSTAR в среднем составляет 0.892, время распознавания в среднем 254 мс. Достигнутые показатели являются приемлемыми для использования в мобильных платформах, в т.ч. для БПЛА и наземных автономных роботов.

## 1. Введение

Проблема использования существующих алгоритмов фрактального сжатия на мобильных программно-аппаратных платформах отмечена в [1]. Традиционно методы фрактального сжатия обладают высокой вычислительной сложностью, и для мобильных платформ не всегда применимы методы и алгоритмы оптимизации производительности разработанные для настольных программно-аппаратных платформ [2][3]. Современные решения проблемы производительности основаны на применении программируемых пользователем вентильных матриц(FPGA) и использовании графических процессоров, что затрудняет использование данных подходов для большинства мобильных платформ. При этом актуальность использования методов фрактального сжатия для мобильных устройств подчеркнута в статье [4].

## 2. Реализация машинного обучения на циклическом фрактальном сжатии

Один из перспективных подходов к реализации классификатора на основе фрактального сжатия предложен в [5] При обучении классификаторов описанных в [6], основная проблема заключалась в том, что изображения, формирующие обучающую выборку одного класса, сжимались независимо друг от друга, и объединялись только на этапе построения опорного подпространства. При этом на этапе распознавания возникают проблемы связанные с возможным пересечением опорных подпространств. Соответственно необходимо применять

методы обеспечивающие пространственную разделимость, что дополнительно повышает вычислительную сложность. В [7] предложена схема фрактального сжатия использующая несколько различных изображений. В данной статье исследуется классификатор построенный по схеме на замкнутой цепочке изображений из обучающей выборки, с формированием словаря ранговых и доменных областей [8] и соответствующих им преобразований.

Классический алгоритм сжатия на основе систем итерированных функций осуществляет поиск наилучшего аффинного преобразования из доменной области в ранговую. В результате входное изображение кодируется несколькими аффинными преобразованиями:

$$\begin{aligned} \mathbf{I}^* &= F(\mathbf{I}) = \mathbf{C}_{1,4}\mathbf{I} + \mathbf{c}_{5,6}, \\ u_{i,j}^* &= c_7 \cdot u_{i,j} + c_8, \end{aligned} \tag{1}$$

где  $\mathbf{I}^* = (i^*, j^*)^T$ ,  $\mathbf{I} = (i, j)^T$  – координаты пикселей из доменных и ранговых областей соответственно,  $\mathbf{C}_{1,4} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}$ ,  $\mathbf{c}_{5,6} = \begin{bmatrix} c_5 \\ c_6 \end{bmatrix}$  – коэффициенты аффинного преобразования,  $u_{i,j}^*$ ,  $u_{i,j}$  – значения яркостей пикселей в из доменных и ранговых областях,  $c_7, c_8$  – параметры контраста и сдвига яркости.

В работе используется 8 различных типов преобразований:

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

$c_5, c_6$  – коэффициенты сдвига и устанавливаются в зависимости от положения блока на изображении. Эти параметры соответствуют различным видам преобразований, таким как вращение, отражение доменных областей с коэффициентом сжатия 0,5, что в соответствии с теоремой Банаха обеспечивает вычислительную сходимость процесса.

Параметры преобразований  $c_1 - c_8$  вычисляются в результате фрактального сжатия:  $c_1 - c_4$  выбираются из наборов,  $c_5, c_6$  вычисляются в процессе поиска самоподобных блоков,  $c_7, c_8$  – вычисляются по средним значениям яркостей блоков.

Набор преобразований для каждого блока диапазона может быть записан как:

$$I_1 = \bigcup_i F_i(I_0). \tag{2}$$

Используя оператор Хатчинсона, это выражение можно записать в виде:

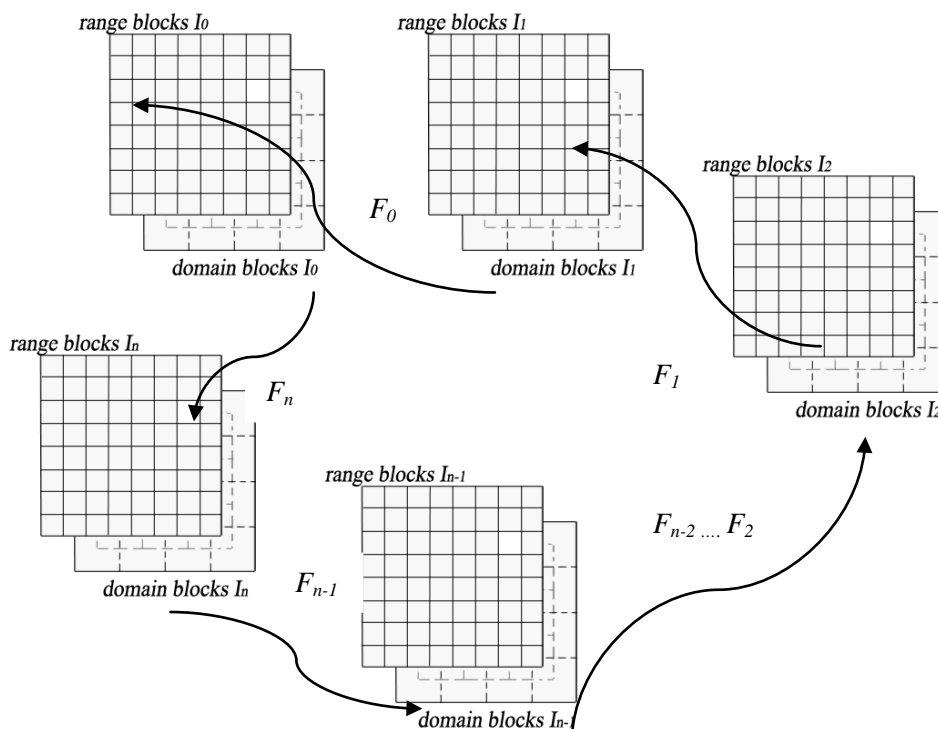
$$I_1 = FI_0, \tag{3}$$

где  $I_0$  – исходное изображение,  $F$  – оператор Хатчинсона, представляющий набор аффинных преобразований,  $I_1$  – результирующее изображение. Схема циклической последовательности преобразований для нескольких изображений обучающего набора представлена на рисунке 1.

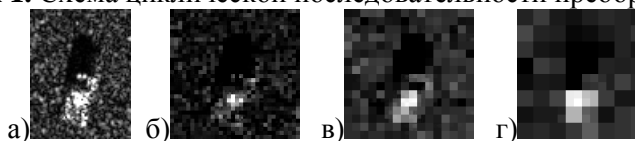
После поиска наилучшего аффинного преобразования из доменных областей в ранговые, формируется словарь, включающий информацию о номерах классов изображений, коэффициентах преобразований и схемой деления изображения на ранговые области.

### 3. Экспериментальные исследования

В качестве исходных экспериментальных данных была выбрана общеизвестная база радиолокационных изображений MSTAR (moving and stationary target acquisition and recognition). Использовались объекты BMP2, BTR70, T72, для каждого объекта из базы были задействованы обучающие и контрольные выборки. Данные SAR изображения получены с помощью радиолокатора в X-диапазоне под двумя разными углами места (15° и 17°). На этапе построения классификатора для обеспечения многомасштабности было использовано различное количество ранговых блоков 16 (4×4), 64 (8×8), 256 (16×16), 1024 (32×32) и соответственно доменных блоков 9 (3×3), 49 (7×7), 225 (15×15), 961 (31×31). Примеры полученных фрактальных изображений представлены на рисунке 2.



**Рисунок 1.** Схема циклической последовательности преобразований.



**Рисунок 2.** Примеры фрактальных изображений объекта, а – исходное изображение, б,в,г – фрактальное изображение для разбиений ранговых блоков  $32 \times 32$ ,  $16 \times 16$ ,  $8 \times 8$  соответственно.

Проведено исследование метода распознавания объектов с циклическим фрактальным кодированием с использованием словаря доменных блоков на трехклассовой классификации, с объектами BMP2, BTR70, T72 (Таблица 1).

**Таблица 1.** Мультиклассовое распознавание для трех типов объектов.

Класс	Доля верно распознанных объектов	
	Предлагаемые метод	Saliency Attention and SIFT[11]
BMP2	0.891	0.64
BTR70	0.882	0.75
T72	0.904	0.74

Предложенный метод сравнивался с другим экспериментальным методом [9], условия экспериментов идентичны. Также было исследовано быстродействие метода распознавания на мобильных платформах на базе процессора Qualcomm Snapdragon 625, 2 ГГц. Получено, что среднее время распознавания объектов по базе MSTAR составляет 254 мс, что в целом соответствует скорости обработки в реальном времени.

#### 4. Заключение

В результате исследований метода машинного обучения с циклическим фрактальным кодированием и использованием словаря доменных блоков получено, что доля верно распознанных объектов по базе MSTAR в среднем составляет 0.892, время распознавания в

среднем 254 мс. Достигнутые показатели являются приемлемыми для использования в мобильных платформах, в т.ч. для БПЛА и наземных автономных роботов.

## 5. Литература

- [1] Srivastava, S. Superresolution based Medical Image Compression for Mobile Platforms / S. Srivastava, B. Lall // Workshop on Machine Learning for HealthCare. – 2015. – P. 01436138.
- [2] Chen, D. Fractal video compression in OpenCL: An evaluation of CPUs, GPUs, and FPGAs as acceleration platforms / D. Chen, D. Singh // Design Automation Conference (ASP-DAC), 2013. – P. 297-304.
- [3] Son, T.N. Fast FPGA implementation of YUV-based fractal image compression / T.N. Son, T.M. Hoang, N.T. Dzong, N.H. Giang // Communications and Electronics (ICCE), 2014. – P. 440-445.
- [4] Lima, V. Fast low bit-rate 3D searchless fractal video encoding / V. Lima, W. Schwartz, H. Pedrini // 24th SIBGRAPI Conference on Graphics, Patterns and Images (Sibgrapi), 2011. – P. 189-196.
- [5] Minaev, E. Object recognition based on fractal coding using domain blocks dictionary // Journal of Physics: Conference Series. – 2018. – Vol. 1096. – P. 012099.
- [6] Minaev, E. Support subspaces method for fractal images recognition / E.Y. Minaev, V.A. Fursov // CEUR Workshop Proceedings. – 2016. – Vol. 1638. – P. 379-385.
- [7] Ozawa, K. Dual fractals // Image and Vision Computing. – 2008. – Vol. 26. – P. 622-631.
- [8] Sun, Y. A Novel Fractal Coding Method Based on MJ Sets / Y. Sun, R. Xu, L. Chen, R. Kong, X. Hu // PloS one. – 2014. – Vol. 9(7). – P. e101697.
- [9] Karine, A. Saliency attention and sift keypoints combination for automatic target recognition on MSTAR dataset / A. Karine, A. Toumi, A. Khenchaf, M. El Hassouni // Advanced Technologies for Signal and Image Processing (ATSIP), International Conference on IEEE, 2017. – P. 1-5.

## Благодарности

Разработка методов и алгоритмов выполнена при поддержке РФФИ (проект № 17-29-03112-офи-м), экспериментальные исследования - в рамках госзадания ИСОИ РАН - филиал ФНИЦ "Кристаллография и Фотоника" РАН (соглашение № 007-ГЗ/ЧЗ363/26).

# An investigation of machine learning method based on fractal compression

E.Yu. Minaev<sup>1,2</sup>

<sup>1</sup>Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

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

**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.