
**МАТЕМАТИЧЕСКИЕ МЕТОДЫ
И МОДЕЛИРОВАНИЕ В ПРИБОРОСТРОЕНИИ**

УДК 519.688; 543.087.9

© Л. В. Новиков

**ДЕТЕКТИРОВАНИЕ ПИКОВ НАНОМАСШТАБНЫХ
ИЗОБРАЖЕНИЙ В ШУМАХ**

Предлагается новый помехоустойчивый алгоритм оценки положения и интенсивности пиков нанообъектов. Метод основан на обработке изображения по столбцам и строкам с использованием алгоритма поиска экстремумов по трем точкам в скользящем окне данных. Метод позволяет оценить положение и интенсивность пиков при отношении сигнал/шум до трех с вероятностью обнаружения до единицы, нулевой вероятностью пропуска пика и ложной тревоги.

Кл. сл.: наномасштабные изображения, обработка изображений, детектирование пиков

ВВЕДЕНИЕ

В последнее время значительно расширился круг задач, решаемых с помощью анализа изображений самых различных объектов в области криминалистики, биологии, медицины, микроэлектроники, фармацевтики, аналитической химии и в других приложениях. В работе [1] предложены новые подходы анализа химического состава поверхности гетерогенных образцов, названные "химическая визуализация" (chemical imaging). В основе лежит анализ корреляции между, как минимум, двумя изображениями, полученными различными методами, например время-пролетной вторичной ионной масс-спектрометрией (ToF-SIMS) и сканирующей электронной микроскопией (SEM), рентгеновской фотоэлектронной спектроскопией и атомной силовой микроскопией (AFM) и т. п. с последующей обработкой многомерных данных [2–4].

Концепция многосенсорного гиперспектрального образа анализируемого объекта объединяет неразрушающие методы анализа на основе рамановской (RMS), энергодисперсионной рентгеновской (EDX) и электронно-колебательной спектроскопии, а также SEM и ToF-SIMS. Полученные изображения обрабатываются методами многомерной статистики для получения обобщенных спектральных признаков с использованием методов главных компонент (PCA) и кластерного анализа. Эта концепция находит применение в области науки о жизни, материаловедения и геофизики, классификации клеток и аэрозолей в пробах окружающей среды [1].

Одной из задач, решаемых в теории обработки изображений, является получение данных о морфологии поверхности, например морфологии на-

ноструктур, используемых в различных приложениях от распознавания отпечатков пальцев и диагностики печатных плат [5] до оценки шероховатости поверхности [6–8], определения размеров наночастиц [9] и качества биосимуляров в фармацевтической промышленности [10].

Достоверность принимаемых решений по оценке морфологических параметров, в частности наноструктур, зависит от точности оценок этих параметров в условиях сильных шумов и других помех. Настоящая работа посвящена оценке интенсивности и положения пиков изображений, характерных для комплексных исследований с использованием атомной силовой, туннельной или расстровой электронной микроскопии.

Традиционный подход в анализе структуры поверхности при наличии шумов основан на предварительной фильтрации шума с использованием тех или иных хорошо известных алгоритмов. Однако эта операция неизбежно ведет к искажению характера поверхности — нарушению формы пиков, уменьшению их интенсивности и иногда положения максимумов. В предлагаемом алгоритме не требуется предварительная фильтрация. Он основан на методе поиска начала, максимума и конца пиков по производной в трех точках скользящего окна данных, предложенного в работе [11]. Поиск максимумов изображения нанообъектов выполняется в два этапа. На первом — путем обнаружения экстремумов по строкам и столбцам матрицы данных формируются кластеры точек в зоне вершины пика. Затем, используя иерархический кластерный анализ, локализуются эти кластеры и определяются их центры, величина и расположение которых принимаются за параметры пиков в изображении.

ОПИСАНИЕ АЛГОРИТМА

Изображение, содержащее множество пиков, может быть представлено в виде матрицы Z , элементы которой $z(i, j)$ определяются формулой

$$z(i, j) = \sum_{l=1}^L A_l f_l((i - i_l), (j - j_l)) + \rho(i, j),$$

где A_l — интенсивность; f_l — функция формы; i_l , j_l — положение l -го пика; L — число пиков; $\rho(i, j)$ — аддитивный шум; $i = 1 \dots I$, $j = 1 \dots J$.

В простейшем случае в качестве модели функции f_l может быть выбран двумерный гаусс

$$f_l(\cdot) = A_l \exp \left\{ -\frac{r_l^2}{\mu_l^2} \right\},$$

где $r_l^2 = (i - i_l)^2 + (j - j_l)^2$, μ_l — среднеквадратическая ширина пика.

В настоящей работе для проверки алгоритма используется более сложная модель пика, представляющая собой гауссов пик с эллипсоидальной формой сечения горизонтальной плоскостью. Тогда

$$r_l^2 = \left((i - i_l)^2 + (j - j_l)^2 \right) / A_l^2 - \frac{\mu_l^2}{\beta_l^2} \left((i - i_l) \cos \alpha_l - (j - j_l) \sin \alpha_l \right)^2 / A_l^2.$$

Здесь $2\mu_l$ — фокальное расстояние и $2\beta_l$ — большая полуось эллипса, $\beta_l > \mu_l$, α_l — угол наклона оси эллипса в плоскости (i, j) ,

$$A_l = \sqrt{\beta_l^2 - \mu_l^2}.$$

Каждый j -й столбец матрицы Z образует I -мерный вектор-столбец $\mathbf{Z}_j = [z_j(i)]$ с элементами $z_j(i)$.

Каждая i -я строка матрицы Z образует J -мерный вектор-строку $\mathbf{Z}_i = [z_i(j)]$ с элементами $z_i(j)$.

Алгоритм обнаружения положения (i_l, j_l) и амплитуды A_l пиков изображения производится последовательной обработкой по столбцам и строкам матрицы Z с последующим анализом полученных данных с целью уточнения искомых параметров.

При обработке по столбцам $\mathbf{Z}_j$ производится поиск перегибов (хребтов) в каждом столбце j матрицы Z . Для этой цели используется подход, подробно изложенный в работе [11] для одномерного случая. Суть предлагаемого алгоритма состоит в следующем.

Сформируем текущую выборку $\mathbf{y}$ из $K + 1$ отсчетов с центром в точке i вектора $\mathbf{Z}_j$

$$\mathbf{y} = [y(k)] = (z_j(i - K/2), \dots, z_j(i), \dots, z_j(i + K/2)),$$

где $k = 1 \dots K$ и K выбирается всегда четным числом из условия 0.5–0.7 от ширины пика на половине высоты в отсчетах. При этом во избежание скачков на границах $\mathbf{Z}_j$ полагаем, что если $(i - K/2) \leq 0$, то $y(i - K/2) = \mathbf{Z}_j(1)$ и, если $(i + K/2) > I$, то $y(i + K/2) = \mathbf{Z}_j(I)$.

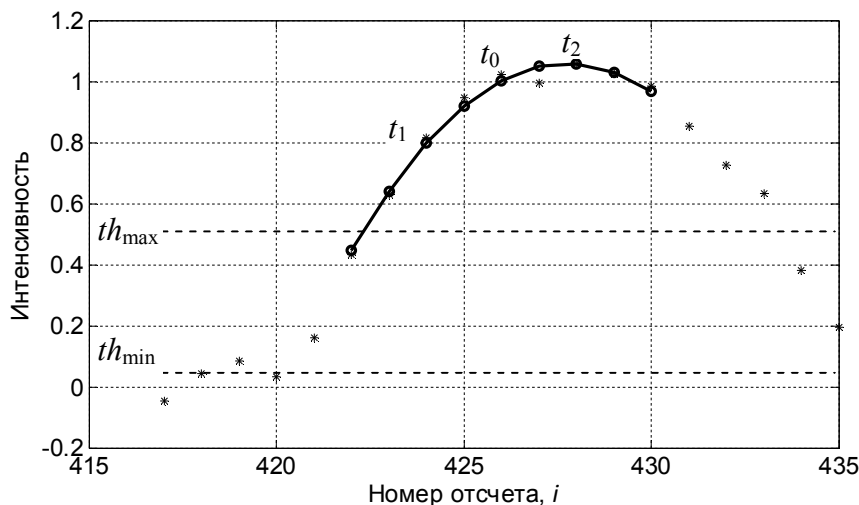


Рис. 1. Фрагмент столбца матрицы изображений с окном данных из девяти отсчетов ($K = 8$) и аппроксимирующая парабола (сплошная линия) с точками t_1 , t_0 , t_2 ; $th_{\min}$, $th_{\max}$ — минимальный и максимальный пороги обнаружения

Аппроксимируем отсчеты в окне y полиномом второй степени (параболой) $\tilde{y}$ на сетке $k = (-K/2, \dots, 0, \dots, K/2)$:

$$\tilde{y}(k) = p_1 k^2 + p_2 k + p_3,$$

где p_1, p_2, p_3 — коэффициенты полинома.

Выберем в интервале $\left[-\frac{K}{2}, \frac{K}{2}\right]$ три точки t_1, t_0, t_2 такие, что $t_1 = -\left[\frac{K}{4}\right]$, $t_0 = 0$, $t_2 = \left[\frac{K}{4}\right]$, где $\left[\frac{K}{4}\right]$ — целое число, большее или равное $\frac{K}{4}$.

На рис. 1 показан фрагмент столбца с окном данных из девяти отсчетов ($K=8$) и аппроксимирующая парабола с точками t_1, t_0, t_2 .

Величина первой производной параболы в точке t_1 будет равна

$$d_1 = 2p_1 t_1 + p_2 \quad (1)$$

и в точке t_2 :

$$d_2 = 2p_1 t_2 + p_2. \quad (2)$$

Величина второй производной параболы равна $2p_1$.

Величина параболы в точке $k=0$ равна средневзвешенному значению в скользящей выборке исходных данных в точке i :

$$\bar{Z}_j(i) = \tilde{y}(0) = p_3. \quad (3)$$

Используя формулы (1)–(3) рассмотрим алгоритм обнаружения хребтов в каждом столбце матрицы Z . В начале обработки исходных данных устанавливаются признаки обнаружения начала и вершины хребта:

$$PS = \begin{cases} 0, & \text{начало хребта не обнаружено;} \\ 1, & \text{начало обнаружено.} \end{cases} \quad (4)$$

$$PSS = \begin{cases} 0, & \text{вершина хребта не обнаружена;} \\ 1, & \text{вершина обнаружена.} \end{cases} \quad (5)$$

Условие начала хребта имеет вид:

$$PS = 0; d_1 > 0; d_2 > d_1; p_3 > th_{\min}, \quad (6)$$

где $th_{\min}$ — порог нижнего уровня, выбираемого по стандартному отклонению шума σ : $th_{\min} = (1 \div 2)\sigma$.

После обнаружения начала хребта устанавливаются признаки $PS=1$ и $PSS=0$.

Обозначим номера строк, в которых обнаружены начала хребтов, как $I_{\min \max}$.

Условие вершины хребта имеет вид:

$$PS=1; d_1 > 0; d_2 < 0; p_1 < 0; \max(\tilde{y}) > th_{\max}, \quad (7)$$

где $th_{\max}$ — порог максимального уровня, выбираемый из условия $th_{\max} = (4 \div 6)\sigma$. После обнаружения вершины хребта устанавливаются признаки $PS=0$ и $PSS=1$. Обозначим номера строк, в которых обнаружены максимумы хребтов, как $I_{\max}$.

Сформируем новую матрицу Z_m данных, состоящую из столбцов матрицы Z и строк с номерами $I_{\max}$, т. е. $Z_m = Z(I_{\max}, 1:J)$.

Выполним анализ матрицы Z_m по строкам $I_{\max}$ с использованием алгоритмов (4)–(7) с целью обнаружения номеров столбцов $J_{\max}$, в которых обнаружены максимумы.

В результате приведенного анализа матрицы данных Z получаем координаты пиков в виде таблицы $[I_{\max} J_{\max}]$.

Амплитуда $\hat{A}(I_{\max}, J_{\max})$ обнаруженных пиков оценивается по точке максимума параболы при условии $PSS=1$. Если форма пиков несимметричная, то в окрестности их вершины даже при отсутствии шума обнаруживается несколько максимумов, образующих кластер. За оценку положения такого пика принимается среднее взвешенное значение по каждой координате, а за интенсивность — максимальное значение интенсивности пиков, входящих в кластер. Признаком принадлежности кластеру принимается расстояние между точками, которое должно быть меньше минимального расстояния между пиками анализируемой поверхности, задаваемого по априорным данным. За оценку положения кластера принимается среднее взвешенное значение по каждой координате, а за интенсивность — максимальное значение интенсивности пиков, входящих в кластер.

МОДЕЛИРОВАНИЕ

Проверка эффективности алгоритма производилась по модельному спектру, состоящему из 81 пика гауссовой формы с эллипсоидальной формой сечения и добавлением белого шума (рис. 2, а). На рис. 2, б, показаны кластеры пиков изображения и их вершины, полученные на основе предлагаемого

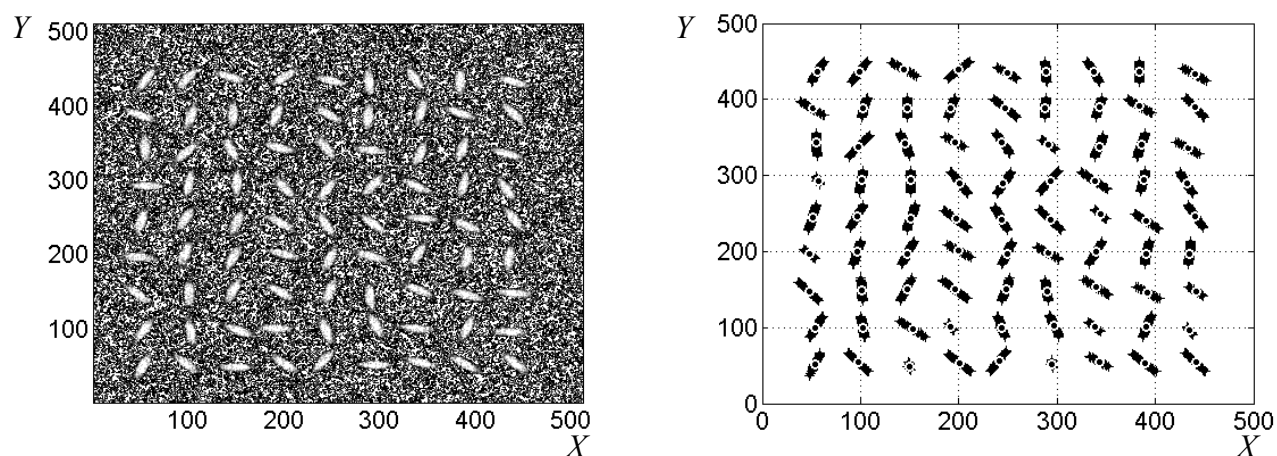


Рис. 2. Изображение из 81 модельных пиков.

а — в шумах при отношении сигнал/шум, равном трем; б — кластеры пиков изображения с отметкой обнаруженных вершин пиков (О)

Абсолютные (для положения пиков по осям X и Y) и относительные среднеквадратические ошибки (СКО) оценки параметров пиков при различных отношениях сигнал / шум (c/π) для метода поиска экстремумов, изложенного в [12, 13], — 1 и предлагаемого метода — 2.

Параметр	Отношение с/ш							
	3.0		10		20		100	
	Метод поиска экстремумов							
	1	2	1	2	1	2	1	2
Положение пика (в отсчетах):								
	X	0	0	0	0	0	0	0
	Y	1	1	1	0	0.8	0	0
Положение пика (СКО):								
	X	1.65	0.85	1.44	0.97	1.40	0.9	1.40
	Y	1.51	1.13	1.38	0.87	1.33	0.68	1.30
Амплитуда		0.15	0.075	0.12	0.029	0.120	0.021	0.120
Вероятность правильного обнаружения		0.907	0.931	0.976	0.999	0.970	1	0.986
Вероятность ложного обнаружения		0.727	0.034	0.028	0.002	0.020	0	0.013

алгоритма. Результаты сравнивались с программой определения экстремумов в изображении — extrema2.m, используемой в разложении по эмпирическим модам (BEMD) [12, 13], с предварительной фильтрацией шума на основе мультивейвлетов [14]. Относительная погрешность оценки координат и амплитуды пиков производилась по десяти реализациям шума. Погрешность положения пика определялась по его отклонению от истинного значения. Максимальные значения погрешностей для обоих методов обработки сведены в таблицу.

Кроме того, вычислялась вероятность правильного обнаружения как отношение числа обнаруженных пиков по всем десяти реализациям, положение которых отличается от заданных в модели не более, чем на четыре отсчета, к общему числу модельных пиков. Вероятность ложного обнаружения вычислялась как отношение числа обнаруженных пиков по всем десяти реализациям, не заданных в модели, к общему числу модельных пиков. Сравнение данных в столбцах методов 1 и 2 таблицы показывает явное преимущество предлагаемого алгоритма как по уровню погрешности оценки параметров пиков, так и по величине вероятности ложного обнаружения.

ЗАКЛЮЧЕНИЕ

Алгоритм детектирования предложенным методом позволяет повысить достоверность обнаружения пиков благодаря низкому уровню ошибки оценок и вероятности ложного обнаружения при малых отношениях сигнал/шум. Это позволяет значительно понизить порог чувствительности приборов и расширить диапазон оцениваемых параметров.

Работа поддержана государственным заданием № 007-00229-18-00.

СПИСОК ЛИТЕРАТУРЫ

1. Ofner J., Brenner F., Wieland K. et al. Image-based chemical structure determination // Scientific reports. 2017. Doi: 10.1038/s41598-017-07041-x.
2. Ofner J., Kamilli K.A., Eitenberger E. et al. Chemometric analysis of multisensor hyperspectral images of precipitated atmospheric particulate matter // Anal. Chem. 2015. Vol. 87, no. 18. P. 9413–9420. Doi: 10.1021/acs.analchem.5b02272.
3. Gowen A.A., Dorrepaal R.M.. Multivariate chemical image fusion of vibrational spectroscopic imaging modalities // Molecules. 2016. Vol. 21, no. 7. Doi: 10.3390/molecules21070870.
4. Sobol O., Holzlechner G., Holzweber M. et al. First use of data fusion and multivariate analysis of ToF-SIMS and SEM image data for studying deuterium-assisted degradation processes in duplex steels // Surf. Interface Anal. 2016. Wiley Online Library. Doi: 10.1002/sia.6015.
5. Kaur B., Kaur S.P. Applications of mathematical morphology in image processing: a review // International Journal of Electronics & Communication Technology. 2013. Vol.4, no. 3. P. 15–17.
6. Толстихина А.Л. Атомно-силовая микроскопия кристаллов и пленок со сложной морфологией поверхности. Дис. ... д-ра физ.-мат. наук. М.: ФГБУН Институт кристаллографии им. А.В. Шубникова РАН, 2013. 332 с.
7. Орлов А.А., Антонов Л.В. Алгоритмы обработки снимков промышленных изделий // Современные проблемы науки и образования. 2012. № 6. URL: <http://science-education.ru/ru/article/view?id=7844>.
8. Соифер В.А., Куприянов А.В. Анализ и распознавание наномасштабных изображений: традиционные подходы и новые постановки задач // Компьютерная оптика. 2011. Т. 35, № 2. С. 136–144.
9. Чукланов А.П., Бородин П.А., Зиганишина С.А., Бухарев А.А. Алгоритм для анализа АСМ-изображений поверхностей со сложной морфологией // Ученые записки Казанского государственного университета. 2008. Т. 150, № 2. С. 220–227.
10. Reichelt W.N., Kaineder A., Brillmann M. et al. High throughput inclusion body sizing: Nano particle tracking analysis // Biotechnol. J. 2017. Vol. 12, № 6. Doi: 10.1002/biot.201600471.
11. Новиков Л.В., Куркина В.В. Метод оценки параметров спектральных пиков // Научное приборостроение. 2017. Т. 27, № 3. С. 99–106. URL: <http://213.170.69.26/mag/2017/abst3.php#abst12>.
12. Dong W., Li X., Lin X. and Li Z. A bidimensional empirical mode decomposition method for fusion of multispectral and panchromatic remote sensing images // Remote Sens. 2014. Vol. 6, no. 9. P. 8446–8467. Doi: 10.3390/rs6098446.
13. MathWorks. URL: <http://www.mathworks.com/matlabcentral/fileexchange>.
14. Rahman M.A., Khan A.S. A spatial spectral filtration (SSF) based correlated coefficients thresholding approach for image denoising // International Journal of Computer Science and Information Technologies. 2016. Vol. 7, № 3. P. 1097–1101.

**Институт аналитического приборостроения РАН,
г. Санкт-Петербург**

Контакты: Новиков Лев Васильевич,
novik38@mail.ru

Материал поступил в редакцию 7.05.2018

DETECTION OF PEAKS NANOSCALE IMAGES IN NOISES

L. V. Novikov

Institute for Analytical Instrumentation of RAS, Saint-Petersburg, Russia

A new noise-immune algorithm for estimating the position and intensity of the peaks of nanoobjects is proposed. The method is based on image processing by columns and rows using the algorithm of three-point extremes search in a sliding data window. The method allows to estimate the position and intensity of peaks with a signal-to-noise ratio up to three with probability of detection up to unity, zero probability of peak miss and false alarm.

Keywords: nanoscale images, image processing, detection of peaks

REFERENCES

- Ofner J., Brenner F., Wieland K. et al. Image-based chemical structure determination. *Scientific reports*, 2017. Doi: 10.1038/s41598-017-07041-x.
- Ofner J., Kamilli K.A., Eitenberger E. et al. Chemometric analysis of multisensor hyperspectral images of precipitated atmospheric particulate matter. *Anal. Chem.*, 2015, vol. 87, no. 18, pp. 9413–9420. Doi: 10.1021/acs.analchem.5b02272.
- Gowen A.A., Dorrepaal R.M. Multivariate chemical image fusion of vibrational spectroscopic imaging modalities. *Molecules*, 2016, vol. 21, no. 7. Doi: 10.3390/molecules21070870.
- Sobol O., Holzlechner G., Holzweber M. et al. First use of data fusion and multivariate analysis of ToF-SIMS and SEM image data for studying deuterium-assisted degradation processes in duplex steels. *Surf. Interface Anal.*, 2016, Wiley Online Library. Doi: 10.1002/sia.6015.
- Kaur B., Kaur S.P. Applications of mathematical morphology in image processing: a review. *International Journal of Electronics & Communication Technology*, 2013, vol.4, no. 3, pp. 15–17.
- Tolstihina A.L. *Atomno-silovaya mikroskopiya kristallov i plenok so slozhnoj morfologiej poverhnosti*. Diss. ... d-ra fiz.-mat. nauk [Atomic and power microscopy of crystals and films with difficult morphology of a surface. Drs. phys.-mat. sci. diss.]. Moscow, FGBUN Institute of a crystallography A.V. Shubnikova of RAS, 2013. 332 p. (In Russ.).
- Orlov A.A., Antonov L.V. [Image processing algorithms of the industrial products]. *Sovremennye problemy nauki i obrazovaniya* [Modern problems of science and education], 2012, no. 6. (In Russ.). URL: <http://science-education.ru/ru/article/view?id=7844>.
- Soifer V.A., Kupriyanov A.V. [Analysis and recognition of the nanoscale images: conventional approach and novel problem statement]. *Komp'yuternaya optika* [Computer Optics], 2011, vol. 35, no. 2, pp. 136–144. URL: <http://www.computeroptics.smr.ru/KO/PDF/KO35-2/01.pdf>. (In Russ.).
- Chuklanov A.P., Borodin P.A., Ziganshina S.A., Buhaev A.A. [Algorithm for the analysis of ASM-images of surfaces with difficult morphology]. *Uchenye Zapiski Kazanskogo Universiteta* [Scientific notes of the Kazan state university], 2008, vol. 150, no. 2, pp. 220–227. (In Russ.).
- Reichelt W.N., Kaineder A., Brillmann M. et al. High throughput inclusion body sizing: Nano particle tracking analysis. *Biotechnol. J.*, 2017, vol. 12, № 6. Doi: 10.1002/biot.201600471.
- Novikov L.V., Kurkina V.V. [The method for estimation of spectral peak parameters]. *Nauchnoe Priborostroenie* [Scientific Instrumentation], 2017, vol. 27, no. 3, pp. 99–106. Doi: 10.18358/np-27-3-i99106. (In Russ.).
- Dong W., Li X., Lin X. and Li Z. A bidimensional empirical mode decomposition method for fusion of multispectral and panchromatic remote sensing images. *Remote Sens.*, 2014, vol. 6, no. 9, pp. 8446–8467. Doi: 10.3390/rs6098446.
- MathWorks. URL: <http://www.mathworks.com/matlabcentral/fileexchange>.
- Rahman M.A., Khan A.S. A spatial spectral filtration (SSF) based correlated coefficients thresholding approach for image denoising. *International Journal of Computer Science and Information Technologies*, 2016, vol. 7, no. 3, pp. 1097–1101.

Contacts: Novikov Lev Vasil'evich,
novik38@mail.ru

Article received in edition 7.05.2018