

Информационные технологии, моделирование и управление

УДК: 637.1:65.011.56

Профессор В.К. Битюков, профессор А.А. Хвостов,
доцент Д.И. Ребриков, аспирант В.Е. Мерзликин
(Воронеж. гос. ун-т. инж. технол.) кафедра информационных и управляющих систем.
тел. (473) 255-38-75
E-mail: kafasu@vgta.vrn.ru

Professor V.K. Bitiukov, professor A.A. Khvostov,
associate professor D.I. Rebrikov, graduate V.E. Merzlikin
(Voronezh state university of engineering technologies) Department of information and control systems.
phone (473) 255-38-75
E-mail: kafasu@vgta.vrn.ru

Автоматизация обработки микрофотографий молочных продуктов с использованием ImageJ и Statistica

Automated processing of dairy product microphotos using ImageJ and Statistica

Реферат. В статье рассмотрены вопросы построения алгоритмов автоматизированной обработки микрофотографий молочных продуктов. Автоматическая обработка микрофотографий молочных продуктов актуальна при исследовании степени их гомогенизации. Микрофотографии молочных продуктов несут информацию о распределении жировых шариков по массовым фракциям. На сегодняшний день существует большое количество программных продуктов, осуществляющих обработку изображений и избавляющих исследователей от рутинных операций ручной обработки данных. Однако требуется их адаптация для осуществления обработки микрофотографий молочных продуктов. В данной работе предлагается использовать для обработки пакет прикладных программ ImageJ для обработки графических файлов, полученных с помощью цифрового микроскопа, а для расчета статистических характеристик предлагается использовать пакет прикладных программ Statistica. Алгоритм обработки состоит из последовательных стадий преобразования в формат оттенков серого цвета, шкалирования, фильтрации, бинаризации, распознавания объектов и статистической обработки результатов распознавания. Результатом работы реализованных алгоритмов обработки данных является функция распределения жировых шариков по объемным или массовым фракциям, а также статистические параметры распределения (математическое ожидание, дисперсия, коэффициенты асимметрии и эксцесса). Для осуществления проверки работы алгоритма и его отладки проведены экспериментальные исследования. Проведена гомогенизация фермерского молока при разных давлениях гомогенизации. Для каждого образца были сделаны микрофотографии и проведена обработка снимков в соответствии с предлагаемым алгоритмом. Проведенные исследования показали эффективность и возможность реализации предложенного алгоритма в виде скрипта на языке java для пакета программ ImageJ с последующей передачей данных в виде файла в пакет прикладных программ Statistica.

Summary. The article discusses the construction of algorithms for automated processing of microphotos of dairy products. Automated processing of micro photos of dairy products relevant in the study of the degree of homogenization. Microphotos of dairy products contain information about the distribution of fat globules in the mass fractions. Today, there are some of software products, offering image processing and relieving researchers from routine operations manual data processing. But it need to be adapted for performing the processing of microphotos of dairy products. In this paper we propose to use for processing the application package ImageJ for processing image files taken with digital microscope, and to calculate the statistical characteristics of the proposed use of the software package Statistica. Processing algorithm consists of successive stages of conversion to gray scale, scaling, filtering, binarization, object recognition and statistical processing of the results of recognition. The result of the implemented data processing algorithms is the distribution function of the fat globules in terms of volume or mass fraction, as well as the statistical parameters of the distribution (the mathematical expectation, variance, skewness and kurtosis coefficients). For the inspection of the algorithm and its debugging experimental studies were carried out. Carries out the homogenization of farm milk at different pressures of homogenization. For each sample were made microphoto sand image processing carried out in accordance with the proposed algorithm. Studies have shown the effectiveness and feasibility of the proposed algorithm in the form of java script for ImageJ and then send the data to a file for the software package Statistica.

Ключевые слова: автоматизация обработки изображений, качество молочных продуктов, гомогенизация молочных продуктов.

Keywords: automation of images processing, quality of dairy products, homogenization of dairy products.

© Битюков В.К., Хвостов А.А.,
Ребриков Д.И., Мерзликин В.Е., 2014

В последнее время информационные и цифровые технологии находят все большее применение в разнообразных видах деятельности человека. Одним из эффективных применений является анализ свойств оптически проницаемых сред. В задачах контроля качества и управления технологическим процессом гомогенизации молочных продуктов может использоваться микроскопия проб молочных продуктов для анализа степени гомогенизации, распределения жировых шариков по массе [1]. Для автоматизации обработки информации используются микроскопы, соответствующие цифровые камеры, обеспечивающие регистрацию изображений в заданные промежутки времени, устройства захвата изображений или видеопотока и управления цифровой камерой, а также программные средства для интерпретации полученной информации в виде, удобном для лица, принимающего решение (ЛПР).

Несмотря на наличие средств получения цифровых микроснимков проб молочных продуктов, остается актуальной задача разработки специального программного обеспечения (СПО), реализующего весь цикл обработки информации применительно к молочным продуктам. Для разработки СПО можно использовать имеющиеся библиотеки и среды обработки цифровых изображений, которые необходимо скомпоновать в соответствии с заданным алгоритмом [2]. Одной из распространенных и удобных сред, позволяющих создавать гибкие алгоритмы обработки графической информации является среда ImageJ, в которой реализовано большое количество типовых алгоритмов обработки и анализа цифровых изображений [3]. Кроме того, использование механизма макросов на языке Java позволяет создавать специфические алгоритмы обработки цифровых изображений в соответствии с заданной задачей, импортировать данные разных форматов, а также экспортировать данные в продукты семейства Microsoft Office и Statistica, делает этот продукт удобным для отработки алгоритмов анализа цифровых изображений в задачах анализа качества молочных продуктов.

Рассмотрим использование среды ImageJ для задачи исследования изменения распределения жировых шариков молока в процессах гомогенизации молока.

Процесс гомогенизации молока и жидких молочных продуктов служит для повышения дисперсности в них жировой фазы, что позволяет исключить отстаивание жира во время хранения молока, развитие окислительных процессов, дестабилизацию и подсыживание при интенсив-

ном перемешивании и транспортировании. Под степенью гомогенизации понимают средний размер жировых шариков, степень устойчивой во времени однородности (гомогенности) в смысле отсутствия микроконцентрационных неоднородностей, образующихся при смешивании взаимно-нерастворимых веществ [1].

Изменение распределения жировых шариков осуществлялось гомогенизацией исходного образца молочного продукта (фермерского негомогенизированного молока) на лабораторной установке гомогенизации научно-производственной лаборатории группы компаний «ЭФКО», позволяющей осуществлять процесс гомогенизации молочных продуктов с разным давлением гомогенизации. Для исследований использовались давления 30, 60, 90, 120, 150, 180 МПа. Полученные после гомогенизации пробы исследовались с помощью цифрового микроскопа с увеличениями 100, 400, 600 крат. Параллельно осуществлялось по 2 снимка, ввиду чего общее количество снимков составило 42 штук (таблица 1). Пример микроснимка пробы молока с увеличением в 100 раз представлен на рисунке 1.

Т а б л и ц а 1
Основные характеристики микроснимков

№ п/п	Давление гомогенизации, МПа	Увеличение, крат
1	Без гомогенизации	100
2	Без гомогенизации	400
3	Без гомогенизации	600
4	30	100
5	30	400
6	30	600
7	60	100
8	60	400
9	60	600
10	90	100
11	90	400
12	90	600
13	120	100
14	120	400
15	120	600
16	150	100
17	150	400
18	150	600
19	180	100
20	180	400
21	180	600

Для калибровки изображений и перехода к физическим единицам измерения осуществлялась микросъемка эталонного объекта (рисунок 1).

Поскольку цвет пробы не представляет интереса для дальнейших исследований, для упрощения дальнейших преобразований необ-

ходимо осуществить преобразование в оттенки серого с желаемой шкалой, например $Y_{ij} = 0.3R_{ij} + 0.59G_{ij} + 0.11B_{ij}$. Осуществим преобразование в режим 32 bit следующей командой: `run("32-bit")`.

Дальнейшим этапом является приведение цифрового изображения к физическим единицам измерения с использованием калибрования на основе цифрового изображения измерительной линейки (рисунок 2) с одинаковым коэффициентом увеличения. Для этого используется инструмент MakeLine с параметрами, которые автоматически задаются при создании линии с длиной, равной одному делению эталонной шкалы. Соответствующий скрипт макроса на языке Java имеет вид:

```
makeLine(x1, y1, x2, y2);
run("Set Scale...", "distance=26.019
known=0.000001 pixel=1 unit=m");
```

где $x1, y1, x2, y2$ – координаты точек начала и конца отрезка, соответствующего одному делению шкалы (рисунок 2), параметр *distance* – длина отрезка в пикселях, *known* – длина отрезка в единицах *unit*.

Например, одному делению шкалы на рисунке 1 при увеличении 100 раз соответствует 26 пикселей. В дальнейшем все снимки проходят процедуру установки масштаба `Analyze->SetScale...` с параметрами `Distanceinpixels = 26`, `Pixelaspectratio = 0.000001` и `Unitoflength = m`.

Далее необходимо избавиться от помех на фоне, так как некоторые мелкие объекты в дальнейшем могут ошибочно распознаться как жировые шарики. Для этого можно использовать функцию `Process->SubtractBackground...` с параметром `Rollingballradius = 170`.

```
run("SubtractBackground...",
"rolling=150");
```

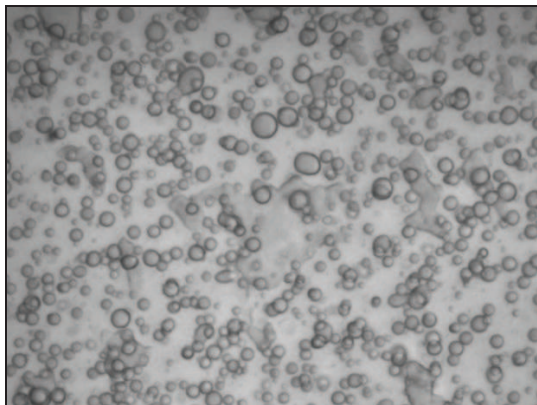


Рисунок 1. Снимок пробы молока со 100-кратным увеличением

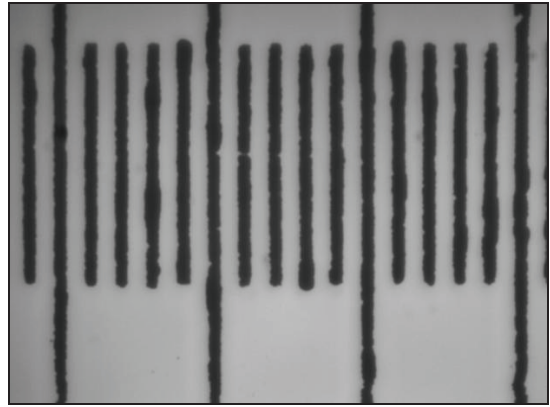


Рисунок 2. Снимок калибровочной линейки со 100-кратным увеличением

Альтернативным вариантом является использование полосового фильтра (bandpass-filter) на основе быстрого преобразования Фурье, чтобы отфильтровать объекты в заданном диапазоне их размеров (например, 5-600px). Алгоритм осуществляет переход в частотную область с использованием быстрого преобразования Фурье, осуществляет вырезку необходимого частотного диапазона, а затем выполняется обратное преобразование [4].

Изображение $f(x, y)$ измеряется с шагом $\Delta x, \Delta y$, измеренные значения $\tilde{f}_{j_x, j_y}$ содержат полезный сигнал и помеху [2]

$$\tilde{f}_{j_x, j_y} = f(x_j, y_j) + \xi_{j_x, j_y}, \quad (1)$$

где $1 \leq j_x \leq N_x, 1 \leq j_y \leq N_y, N_x, N_y$ – размер изображения в единицах элементов размером $\Delta x \times \Delta y$ (может быть один пиксель). Далее осуществляется вычисление коэффициентов дискретного преобразования Фурье (ДПФ):

$$\tilde{F}(l_x, l_y) = \frac{1}{N_x N_y} \sum_{j_x=0}^{N_x-1} \sum_{j_y=0}^{N_y-1} \tilde{f}_{j_x, j_y} \cdot e^{-\frac{2\pi i}{N_x N_y} (j_x l_x + j_y l_y)} \quad (2)$$

Каждому индексу l_x, l_y соответствует частота $\omega_x = \Delta_x l_x, \omega_y = \Delta_y l_y$.

Далее из полученного спектра изображения удаляются определённые участки за пределами полосы пропускания $\omega_{nn} \leq \omega_{nn} \leq \omega_{nn}$,

где ω_{nn}, ω_{nn} – нижний и верхний пределы полосы пропускания, ω_{nn} – частоты, попадающие в диапазон полосы пропускания. Для этого осуществляется операция умножения спектра изображения на функцию, определяющую характер «вырезания» определённых частот

$$\hat{F}(l_x, l_y) = \tilde{F}(l_x, l_y) \cdot W(l_x, l_y), \quad (3)$$

где $W(l_x, l_y)$ функция, реализующая удаление части спектра. В простейшем случае закон «вырезания» выглядит следующим образом:

$$L(\omega) = \begin{cases} 0, \forall \omega > \omega_{nn} \\ 1, \forall \omega_{en} \geq \omega \geq \omega_{nn} \\ 0, \forall \omega < \omega_{en} \end{cases} \quad (4)$$

Этому преобразованию соответствует оператор:

```
run("BandpassFilter...", "filter_large=600
filter_small=5 suppress=Nonetolerance=5 au-
toscalesaturate").
```

Результирующее изображение представлено на рисунке 3.

Следующим этапом идет процедура получения двухбитного изображения на основе пороговой классификации пикселей в соответствии с правилом:

$$T = \{p \in D | f(p) < h\}, \quad (5)$$

где p – характеристика «высоты» пикселя; h – величина порога.

Для этого используются процедуры:

```
setThreshold(1, 159);
run("ConverttoMask").
```

Следует отметить, что выбор пороговых значений в сильной степени определяет качество анализа частиц в дальнейших процедурах, т.к. в результате этой операции происходит как «склеивание» рядом лежащих частиц в одну, так и «незамыкание» колец для светлых частиц, что приводит к погрешностям в дальнейшем анализе. Такие проблемы вызваны как неравномерностью засветки по всей площади снимка, так и ввиду того, что проба молока имеет свою толщину и жировые шарики находятся на разном расстоянии от точки фокуса и поэтому имеют разную резкость и яркость. Результирующее изображение представлено на рисунке 4. Для автоматического определения порогов можно использовать процедуру `setAutoThreshold("Default")`.

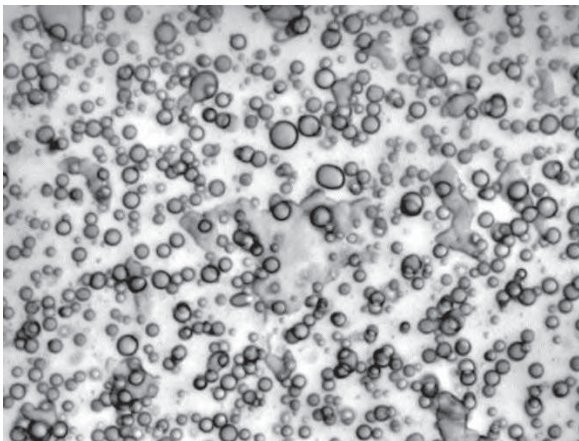


Рисунок 3. Вид микроснимка с настроенным диапазоном полосы пропускания

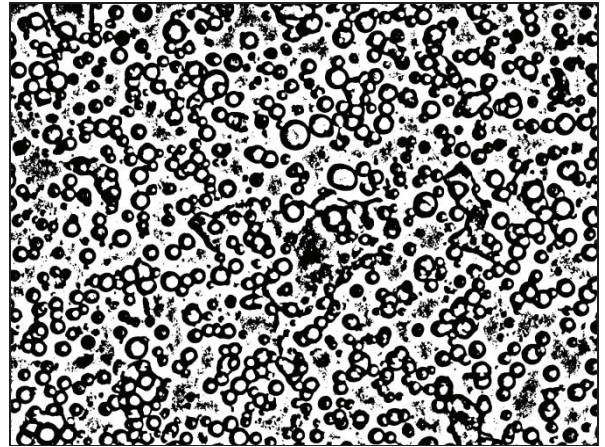


Рисунок 4. Микроснимок в двухбитном виде

Поскольку жировые шарики частично пропускают свет, после обработки получают кольцевые структуры. Для избавления от кольцевых структур можно использовать операцию заливки колец (рисунок 5) `run("FillHoles")`.

Наличие «слипшихся» элементов является причиной близкого расположения жировых шариков друг к другу и их частичного перекрытия. В дальнейшем изображение после операции `FillHoles` малоприспособно для последующей обработки, т.к. «слипшиеся элементы» в ряде случаев будут интерпретироваться как одно целое и необходима процедура, снижающая погрешности расчетов, вызванные «слипанием». Для этих целей можно использовать алгоритм преобразования графических изображений `Watershed`, осуществляющий декомпозицию отдельных элементов изображения по принципам геометрических характеристик их пикселей, аналогичных высоте в рельефе земли и моделирующий «водораздел» [5]. Для бинарных изображений алгоритм декомпозиции основан на вычислении расстояний каждой точки изображения из множества черных пикселей $b \in B$ до ближайшей точки из множества белых пикселей $w \in W$. Расстояние в простейшем случае может быть Евклидовым:

$$D_{Eu}([i_1, j_1], [i_2, j_2]) = \sqrt{(i_1 - i_2)^2 + (j_1 - j_2)^2}. \quad (6)$$

Преобразование осуществляется процедурой `run("Watershed")`. Результат преобразования представлен на рисунке 6.

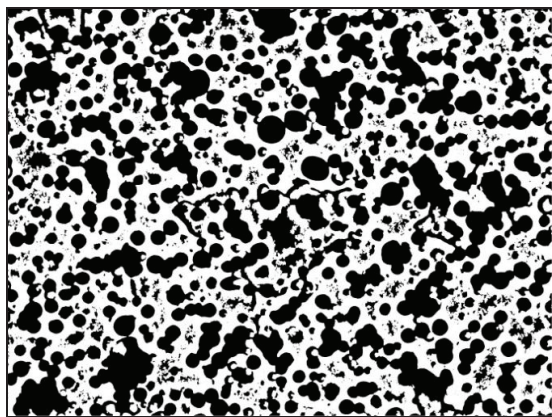


Рисунок 5. Изображение с залитыми кольцами

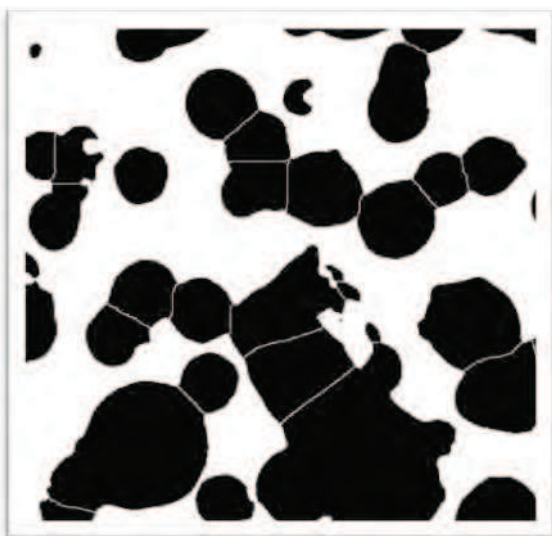


Рисунок 6. Изображение, преобразованное по принципу геометрических характеристик

Проведенные операции подготавливают изображение к проведению основной операции, которая осуществляет поиск и анализ частиц на изображении в соответствии с заданными параметрами AnalyzeParticles. При использовании этой функции можно задать диапазон размеров площадей частиц, которые будут анализироваться в пикселях в квадрате (Size, pixel^2), причем максимальный диапазон может быть от 0 до значения Infinity. Параметр Circularity отвечает за анализ близости найденных частиц к форме круга, значение 1 соответствует кругу, а ноль - полосе единичной толщины. Промежуточные значения позволяют настроить алгоритм поиска на обнаружение фигур, близких по форме к кругу (например, $\text{Circularity}=(0.6 \div 1)$).

Пример выходного графического изображения с опцией Outlines (контуры) представлен на рисунке 7. Также в результате работы процедуры AnalyzeParticles формируется таблица, содержащая перечень свойств всех обнаруженных объектов.

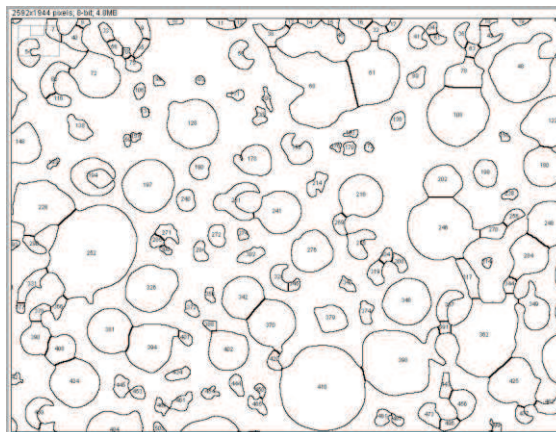


Рисунок 7. Изображение в виде контуров

Полученная таблица содержит информацию о площади обнаруженного объекта в выбранной на этапе шкалирования размерности (пиксель^2 , или мкм^2), среднем, минимальном и максимальном значении выборки, а также порядковый номер объекта. Формат Excel сохранения таблицы совместим с пакетом Statistica, что позволяет использовать всю вычислительную мощь этого пакета для дальнейшего анализа полученных данных.

Для дальнейшего анализа также целесообразно осуществить переход от единиц площади к радиусу объекта $r_{об} = \sqrt{\frac{S_{об}}{\pi}}$ и, зная плотность жира, к приближительному значению массы $m_{об} = \frac{4}{3} \pi \rho_{ж} r_{об}^3$.

Эти расчеты можно выполнить как в пакете Statistica, так и в системе Excel.

Обработка преобразованных данных в пакете Statistica имеет множество возможностей. Основными представляющими интерес с точки зрения свойств молочных продуктов и влияния технологического режима их обработки (в данном случае процесса гомогенизации) являются простейшие описательные статистики (descriptivestatistics). Выделением соответствующего столбца данных и выбора меню Statistics рассчитываются такие статистические характеристики исследуемой выборки как среднее значение, размах, дисперсия, коэффициенты асимметрии и эксцесса.

Кроме этого пакет Statistica строит гистограммы с аппроксимирующей распределение функцией и её параметрами (рисунок 8), которые наглядно демонстрируют распределение жировых шариков по выбранному параметру (площадь, объем или масса). Также информативными для изучения свойств молочных продуктов могут оказаться диаграммы,

показывающие отклонения от нормального распределения и «яшнковые» диаграммы.

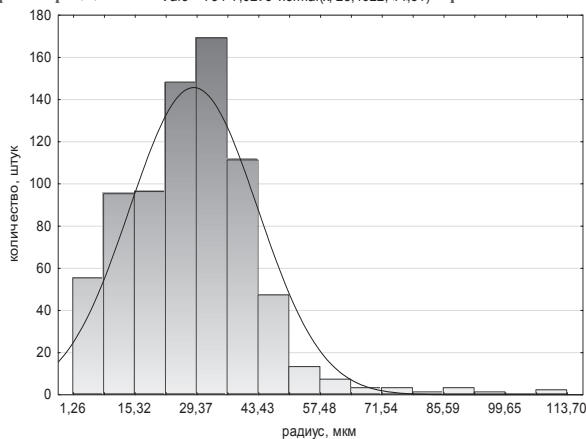


Рисунок 8. Гистограмма распределения жировых шариков в анализируемом образце по радиусу

ЛИТЕРАТУРА

- 1 Крусъ Г.Н. Технология молока и молочных продуктов. М.: Колосс, 2003. 315 с.
- 2 Gonzalez R.C., Woods R.E. Digital Image Processing. Prentice-Hall, 2002. P. 617-620.
- 3 Официальный сайт ImageJ [Электронный ресурс]. Режим доступа: <http://www.imagej.ru> (22 сентября 2014 г.).
- 4 Дженкинс Г., Ваттс Д. Спектральный анализ и его приложения. М.: Мир, 1971. 316 с.
- 5 Jos B., Roerdink T.M., Meijster A. The water shed transform: Definitions, algorithms and parallelization strategies // Fundament Informatics. 2001. № 41. P. 187-228.

В заключение можно отметить, что в сравнении с результатами визуального анализа, при обработке «потеряно» порядка 7% шариков, что можно считать приемлемым результатом, основной проблемой подхода является наличие незамкнутых кольцевых областей, которые не обрабатываются процедурой fillholes. Улучшить результаты можно многократным применением и комбинированием фильтров.

REFERENCES

- 1 Krus' G.N. Tekhnologiya moloka i molochnykh produktov [Technology of milk and dairy products]. Moscow, Koloss, 2003. 315 p. (In Russ.).
- 2 Gonzalez R.C., Woods R.E. Digital Image Processing. Prentice-Hall, 2002. pp. 617-620.
- 3 Ofitsial'nyi sait ImageJ [The official website for ImageJ]. Available at: <http://www.imagej.ru> (Accessed 22 September 2014). (In Russ.).
- 4 Dzenkins G., Vatts D. Spektral'nyi analiz i ego prilozheniia [Spectral analysis and its applications], Moscow, Mir, 1971. 316 p. (In Russ.).
- 5 Jos B., Roerdink T.M., Meijster A. The water shed transform: Definitions, algorithms and parallelization strategies. Fundament Informatics, 2001, no. 41, pp. 187-228.