

# МОДЕЛЬ КВАЗИОПТИМАЛЬНОЙ ИЕРАРХИЧЕСКОЙ СЕГМЕНТАЦИИ ЦВЕТОВОГО ИЗОБРАЖЕНИЯ

© 2015 г. М. В. Харинов, канд. техн. наук

Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации РАН, Санкт-Петербург

E-mail: khar@iias.spb.su

Обсуждается решение проблемы сегментации цифрового изображения посредством иерархической последовательности кусочно-постоянных приближений, которые минимально отличаются от изображения по среднеквадратическому отклонению. Дается аналитическое обоснование вычислений на стадии предварительной автоматической обработки изображения без использования управляющих параметров. Предлагается алгоритм комбинированной сегментации/улучшения качества/кластеризации, который поясняется наглядными примерами.

**Ключевые слова:** *кластеризация пикселов, сегментация изображений, метод Уорда, модель Мамфорда–Шаха, кусочно-постоянные приближения, суммарная квадратическая ошибка, минимизация, иерархическая последовательность.*

Коды OCIS: 150.0150, 100.2000.

Поступила в редакцию 06.02.2015.

## Введение

В ближайшие годы за рубежом, а также в России, ожидается создание унифицированных программно-алгоритмических решений так называемой “проблемы сегментации” изображения. Проблема возникает при распознавании цифровых изображений на первоначальной стадии “приведения данных к удобному для распознавания виду” [1], “локализации объектов” [2], извлечения и упорядочения “глобально-локальной информации” [3], “квазиоптимальной разметки” изображения [4]. Суть проблемы состоит в недостаточной формализации понятия “объектов”, которые “видят” или выделяет на изображении компьютер в виде иерархического множества связных сегментов или *кластеров* из несмежных сегментов для последующего анализа признаков и идентификации в системах машинного зрения.

В наиболее формализованных методах сегментации качество разбиения изображения на сегменты или кластеры оценивается независимо от способа получения сегментированного изображения. При заполнении кластеров средними значениями пикселов разбиения преобразуются в кусочно-постоянныеприближения

изображения. В качестве целевых обычно рассматривают оптимальные разбиения и приближения, минимально отличающиеся от изображения по среднеквадратическому отклонению  $\sigma$ . Поскольку большая алгоритмическая сложность вычисления оптимальных приближений в случае произвольного изображения не позволяет получить точного решения, возникает задача аппроксимации оптимальных приближений близкими квазиоптимальными приближениями, которые оказывается возможным эффективно вычислять для современных изображений из десятков миллионов пикселов. Сегментация посредством квазиоптимальных приближений, трактуемая как стадия первичной автоматической обработки изображения в системах машинного зрения, является темой данной статьи.

## Квазиоптимальная иерархическая сегментация

Модель квазиоптимальной иерархической сегментации разрабатывается исходя из трех основных положений:

- изображение состоит из вложенных изображений, именуемых “объектами”;

- объекты образуют иерархию, что позволяет учесть неоднозначность изображения;
- объектам приписываются целочисленные обозначения, и для пикселов изображения автоматически генерируются метки, которые кодируют иерархию объектов [4, 5].

В нашей модели понятия “изображение”, “объект” и “кластер сегментов” формально не разграничиваются между собой. Сегментация трактуется как разделение изображения на вложенные изображения объектов. В частном случае полагается, что пиксели объекта составляют единственный сегмент. В общем случае допускается, что объект состоит из нескольких сегментов. В результате сегментации рассчитывается квазиоптимальная бинарная иерархия кластеров связных сегментов, которая отражает неоднозначную классификацию пикселов изображения. Тривиальные кластеры, состоящие из одинаковых пикселов, считаются элементарными изображениями, которые полагаются неделимыми. Типичным примером неделимых кластеров являются сегменты из одинаковых пикселов.

Таким образом, в отличие от традиционной сегментации, результатом обсуждаемой иерархической квазиоптимальной сегментации считается не единственное разбиение, а последовательность разбиений изображения на  $g = 1, 2, \dots, N'$  кластеров, где  $N'$  – количество неделимых кластеров в изображении из  $N > N'$  пикселов. При этом задача оптимизации разбиений по среднеквадратическому отклонению  $\sigma$  или суммарной квадратической ошибке  $E = 3N\sigma^2$  трактуется как задача минимизации  $E$  или  $\sigma$  для каждого числа кластеров  $g$ , что исключает характерную для традиционной сегментации необходимость задавать число кластеров или сегментов в качестве управляющего параметра.

Условие иерархической организации объектов обеспечивает принципиальную возможность снижения алгоритмической сложности вычислений за счет исключения анализа повторений, так как при бинарной иерархии последовательность из  $N'$  разбиений порождается всего  $2N' - 1$  различными кластерами, которые повторяются в различных разбиениях, содержащих от 1 до  $N'$  кластеров.

Метками объектов, выделяемых в модели квазиоптимальной иерархической сегментации, служат инвариантные целочисленные обозначения, которые получаются в алгоритме изотонного

преобразования средних по кластерам яркостей пикселов изображения, составляют инвариантные представления квазиоптимальных приближений и не меняются при предусмотренных преобразованиях или искажениях изображения. При этом инвариантные представления визуально воспроизводят квазиоптимальные приближения, задают иерархию последних и конвертируются одно в другое арифметическими преобразованиями значений своих пикселов [4, 5].

## Формулы для алгоритмов аппроксимации изображения

Для разработки алгоритмов аппроксимации изображения квазиоптимальными приближениями мы получили аналитическое выражение, которое в общем виде описывает изменение суммарной квадратической ошибки при слиянии нескольких кластеров.

Если исходное или вложенное изображение является результатом слияния  $j$  вложенных кластеров, то интегральное квадратическое отклонение  $E_0$  значений пикселов этого изображения от среднего значения получается суммированием интегральных квадратических отклонений  $E_p$  для вложенных кластеров и некоторой неотрицательной добавки, характеризующей средний взвешенный разброс квадратов евклидовых попарных расстояний  $\|I_p - I_q\|^2$  между центрами  $I_p, I_q$  кластеров  $p$  и  $q$  из  $n_p$  и  $n_q$  пикселов в цветовом RGB пространстве [6]:

$$E_0 = \sum_{p=1}^j E_p + \frac{1}{2} \frac{\sum_{p=1}^j \sum_{q=1}^j n_p n_q \|I_p - I_q\|^2}{n}. \quad (1)$$

Здесь  $\sum_{p=1}^j E_p = E$  – суммарная квадратическая ошибка для разбиения изображения на  $j$  кластеров,  $E_0$  – суммарная квадратическая ошибка для разбиения изображения с единственным кластером,  $n = \sum_{p=1}^j n_p$  – число пикселов в рассматриваемом изображении. Если в качестве  $j$  вложенных кластеров иметь в виду тривиальные кластеры, на которые распадается изображение, то сумма  $E_p$ , очевидно, обращается в 0. Если в качестве тривиальных кластеров рассматриваются отдельные пиксели, то в формуле (1) следует также положить  $n_p = n_q = 1$  и  $j = n$ . Тогда она, с точностью до множителя,

совпадает с выражением для дисперсии значений пикселов, которое активно применяется в [7–9] для теоретической и практической разработки методов кластеризации. Таким образом, формула (1) обобщает аналогичную формулу [7–9] на случай тривиальных кластеров из нескольких или многих пикселов.

Для случая  $j = 2$  пары вложенных кластеров, из (1) получаются выражения для приращений суммарной квадратической ошибки при трех операциях с кластерами пикселов:

- операции слияния пары кластеров;
- операции коррекции кластеров посредством реклассификации подмножества пикселов из одного кластера в другой;
- операции дробления кластера за счет выделения подмножества пикселов в отдельный кластер [6].

Приращение  $\Delta E_{\text{merge}} = E(1 \cup 2) - E(1) - E(2)$  суммарной квадратической ошибки  $E$  при слиянии кластеров 1 и 2 с числом пикселов  $n_1, n_2$  выражается через квадрат евклидова расстояния  $\|I_1 - I_2\|^2$  между трехмерными средними яркостями  $I_1, I_2$  в виде

$$\Delta E_{\text{merge}} = \frac{\|I_1 - I_2\|^2}{\frac{1}{n_1} + \frac{1}{n_2}} \geq 0. \quad (2)$$

Именно выражение (2) служит критерием слияния сегментов изображения в версиях [6, 10] модели сегментации Мамфорда–Шаха [11] и либо с аддитивной добавкой, либо с дополнительным множителем используется в версиях [12, 13] модели. В отличие от модели Мамфорда–Шаха, в нашей модели сегментации минимизация суммарной квадратической ошибки  $E$  достигается не только за счет слияния смежных сегментов изображения, а может выполняться также и на множество всевозможных пар кластеров, как в методе Уорда [14], что существенно улучшает качество приближений, особенно при малом числе кластеров или сегментов изображения. Характерным свойством иерархической последовательности приближений цветового изображения, получаемых методом Уорда, является выпуклость соответствующей последовательности значений  $E_g$  суммарной квадратической ошибки:

$$E_g \leq \frac{E_{g-1} + E_{g+1}}{2}, \quad g = 2, 3, \dots, N' - 1, \quad (3)$$

где  $g$  – число кластеров в очередном приближении изображения. Однако следует иметь в виду, что резкое падение скорости вычислений, про-

порциональное четвертой степени линейного размера, препятствует непосредственному применению метода Уорда для кластеризации пикселов актуальных цифровых изображений.

Аналитическое выражение для приращения  $\Delta E_{\text{correct}}$  суммарной квадратической ошибки при коррекции кластеров посредством реклассификации  $k$  пикселов из кластера-донора 1 в кластер-акцептор 2 выражается через расстояния  $\|I - I_1\|$  и  $\|I - I_2\|$  между трехмерной средней яркостью  $I$  реклассифицируемых пикселов и средними яркостями  $I_1, I_2$  донорского и акцепторного кластеров формулой

$$\Delta E_{\text{correct}} = \frac{\|I - I_2\|^2}{\frac{1}{k} + \frac{1}{n_2}} - \frac{\|I - I_1\|^2}{\frac{1}{k} + \frac{1}{n_1}}, \quad (4)$$

где  $n_1, n_2$  – число пикселов в кластерах 1 и 2.

При коррекции вложенных кластеров множество пикселов в них меняются. Поэтому обе иерархии множеств пикселов вложенных кластеров приходится каждый раз обновлять, обрабатывая кластеры как самостоятельные изображения. Выражение (4) продуктивно тем, что позволяет вывести классический метод  $K$ -средних ( $K$ -means, [15]) и обобщенный метод “ $K$ -meanless” [7], который в нашей версии [6], наряду с операциями с отдельными пикселами, предусматривает операции с наборами пикселов.

Приращение  $\Delta E_{\text{split}}$  суммарной ошибки при дроблении кластера 1 со средним значением пикселов  $I_1$  за счет выделения в новый кластер его  $k < n_1$  пикселов со средним значением  $I$  описывается формулой

$$\Delta E_{\text{split}} = -\frac{\|I - I_1\|^2}{\frac{1}{k} + \frac{1}{n_1}} \leq 0. \quad (5)$$

При применении формулы для оптимизации дробления кластеров по критерию  $\Delta E_{\text{split}} = \min$  важно, что множество кластеров в (5) полагается сформированным, например, по алгоритмам слияния кластеров, использующим (2), и при дроблении учитывается расширенное множество кластеров, которое включает кластеры, полученные на различных итерациях. Поэтому дробление кластера, полученного итеративным слиянием вложенных кластеров, при обратном увеличении числа кластеров на единицу снижает или, по крайней мере, не увеличивает значение  $E$ . Особенностью алгоритмов дробления на основе формулы (5) является обновление иерархии кластеров, которое при очередной

модификации выполняется для кластера, как для отдельного изображения.

Если в частном случае разделения кластера надвое, допускается только разделение кластера 1 на пару тех кластеров 1' и 1'', слиянием которых он ранее получен, то соответствующее отрицательное приращение  $\Delta E_{\text{divide}}(1)$  суммарной квадратической ошибки, очевидно, выражается формулой (2), взятой с обратным знаком,

$$\Delta E_{\text{divide}}(1) \equiv -\Delta E_{\text{merge}}(1', 1''). \quad (6)$$

Тогда, при наличии в приближении изображения кластеров 2 и 3, которые отличаются от кластера 1 и при слиянии не компенсируют падение  $E$  при его разделении надвое:

$$\Delta E_{\text{divide}}(1) + \Delta E_{\text{merge}}(2, 3) < 0, \quad (7)$$

обеспечивается тривиальная возможность улучшения качества приближения изображения за счет итеративного выполнения операции разделения надвое в комбинации с операцией слияния. По сравнению с методом  $K$ -средних, метод улучшения качества по формуле (7) интересен тем, что позволяет улучшать качество приближений изображения связными сегментами в алгоритмах, которые не выводят из множества связных сегментов.

Таким образом, по сравнению с моделью сегментации Мамфорда–Шаха, в нашей модели комбинированной сегментации/кластеризации число базовых операций с множествами пикселов изображения утроено, что обеспечивает достоверную оптимизацию качества приближений изображения по суммарной квадратической ошибке или среднеквадратическому отклонению приближения от изображения. Судя по нашему опыту [6, 16], система элементарных формул (2), (4), (5)

обеспечивает эффективную разработку и развитие алгоритмов достоверной оптимизации качества приближений изображения и, что не менее важно, позволяет уточнить интерпретацию классических методов [11, 14, 15] минимизации суммарной квадратической ошибки, потенциал которых в приложении к цифровым изображениям, по всей видимости, еще далеко не исчерпан.

## Экспериментальные результаты

Руководствуясь тремя основными формулами (2), (4), (5), можно разработать множество алгоритмов квазиоптимальной кластеризации пикселов, в частности сегментации цифровых изображений. Разнообразие существенно различных алгоритмов зависит от количества признаков, рассчитываемых непосредственно в процессе построения целевой иерархической последовательности разбиений изображения, и организации вычислений для сочетания агломеративных и дивизионных приемов обработки. Комбинирование различных алгоритмов минимизации суммарной квадратической ошибки способствует эффективному приближению к ее минимально достижимым при каждом числе кластеров (сегментов) значениям. Данный раздел на примере стандартного изображения иллюстрирует алгоритмы, testируемые в настоящее время. В нем приводятся характерные результаты оптимизации качества приближений изображения, которые, по сравнению с алгоритмами вычислений, могут быть менее вариабельны, поскольку оптимальные приближения определяются независимо от алгоритмов их получения.

Рисунок 1 демонстрирует результат сегментации изображения в модели Мамфорда–Шаха

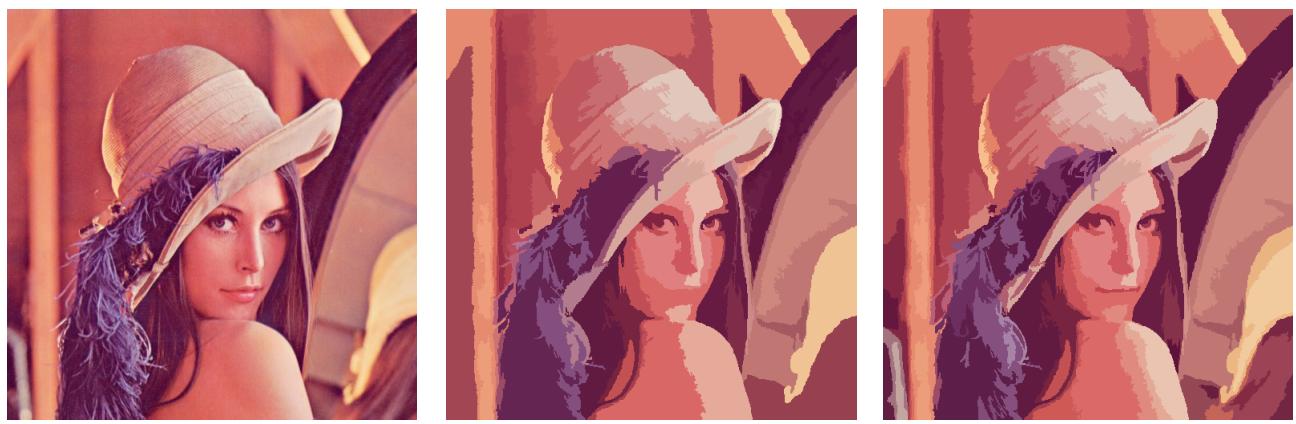


Рис. 1. Улучшение качества приближения цветового изображения.



Рис. 2. Квазиоптимальные приближения в 2, 3 и 20 цветах.

и эффект улучшения качества приближения изображения за счет итеративного слияния двух смежных сегментов по очереди с разделением некоторого третьего сегмента изображения по критерию (7).

Слева на рис. 1 показано цветовое изображение “Лена” из  $512 \times 512$  пикселов. В центре показано его приближение  $g = 100$  сегментами, полученное в скоростной версии [17] модели Мамфорда–Шаха.

Улучшенное приближение, показанное справа, совпадает с центральным по количеству сегментов, но отличается меньшим значением среднеквадратического отклонения  $\sigma$  пикселов приближения от пикселов изображения. Характерно, что формальная оценка улучшения качества приближения согласуется с визуальной оценкой – на обработанном приближении проявляется большее количество объектов.

Рисунки 2 и 3 описывают генерацию последовательности квазиоптимальных приближений для каждого числа кластеров  $g = 1, 2, 3, \dots$  по комбинированному алгоритму сегментации/улучшения качества/кластеризации, в котором

- на первом шаге выполняется иерархическая сегментация изображения по Мамфорду–Шаху;
- на втором шаге производится улучшение качества приближения изображения, как на рис. 1, но для  $g = 1000$  сегментов;
- на третьем, заключительном шаге выполняется кластеризация по Уорду, при которой в качестве начальных кластеров берутся сегменты улучшенного по качеству приближения изображения.

Таким образом, в результате получается иерархия кластеров *суперпикселов*, для которой *суперпиксели* вычисляются в комбинированном алгоритме сегментации/улучшения качества.

На рис. 2 показаны примеры квазиоптимальных приближений стандартного цветового изображения, которые образованы 2, 3 и 20 кластерами пикселов различных цветов.

Квазиоптимальные приближения обеспечивают аппроксимацию изображения с хорошим для современной обработки изображений качеством. Примечательно, что различные парные объекты (глаза) на рис. 2, в крайнем правом квазиоптимальном приближении в 20 цветах, помечаются сгруппированными суперпикселями одинаковых цветов, что способствует решению проблемы “локализации объектов” [2] в автоматическом режиме, без предварительного обучения.

Рисунок 3 графически описывает первую тысячу квазиоптимальных приближений.

Графики на рис. 3 в логарифмическом масштабе по горизонтальной оси иллюстрируют поведение среднеквадратического отклонения  $\sigma$  приближения от изображения в зависимости от числа  $g$  кластеров. Кривая 1 представляет

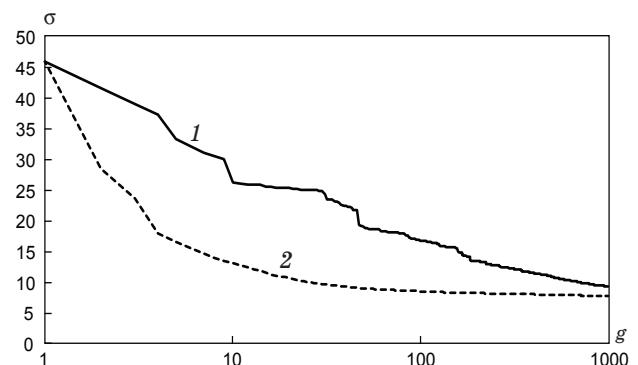


Рис. 3. Среднеквадратическое отклонение  $\sigma$  в зависимости от числа кластеров  $g$ . 1 – сегментация в модели Мамфорда–Шаха, 2 – кластеризация пикселов по Уорду.

результаты скоростной сегментации [17] в модели Мамфорда–Шаха, кривая 2 – результирующую кластеризацию изображения в алгоритме сегментации/улучшения качества/кластеризации. Кривая 2 получается в алгоритме слияния кластеров по Уорду.

При сегментации по Мамфорду–Шаху и кластеризации по Уорду пары объединяемых кластеров выбираются из условия  $\Delta E_{\text{merge}} = \min$ , где  $\Delta E_{\text{merge}}$  вычисляется по формуле (2). При кластеризации по Уорду в качестве начального берется приближение изображения тысячей сегментов, полученное в модели Мамфорда–Шаха, качество которого затем улучшается по алгоритму рис. 1, использующему критерий (7).

В отличие от кривой 1, кривая 2 при пересчете в координаты  $E \times g$  преобразуется в выпуклую (3). Выпуклость последовательности значений суммарной квадратической ошибки при изменении числа кластеров от 1 до установленного числа сегментов является непременным свойством решения согласно модели квазиоптимальной иерархической сегментации. При этом тривиальное свойство монотонной зависимости значений  $E$  от числа кластеров  $g$  дополняется нетривиальной монотонностью значений производной  $E$  по числу кластеров  $g$ .

## Заключение

Обобщенная постановка и решение проблемы сегментации посредством аппроксимации цифрового изображения иерархической последовательностью квазиоптимальных приближений в задачах распознавания необходимы для создания унифицированных алгоритмов автоматического выделения объектов. Помимо автоматизации распознавания цифровых изобра-

жений, обсуждаемые решения полезны также для

- развития стеганографических приложений;
- автоматической обработки оцифрованных звуковых и пр. сигналов;
- фильтрации и подавления шумов;
- минимизации потерь в задачах сжатия сигналов;
- оптимизации порядка передачи информации и др.

Еще 35–40 лет назад типичной задачей при расчетах на ЭВМ являлось сжатие данных без потерь и этим занималось множество программистов для решения конкретных проблем хранения и передачи информации. В настоящее время этой задачей продолжает заниматься узкий круг профессионалов, а остальные специалисты пользуются готовыми программами. Вполне вероятно, что автоматическая иерархическая сегментация цифровых изображений окажется следующей проблемой, для которой будут найдены общеупотребительные решения, превосходящие известные по установленным критериям качества. В нашем варианте аппроксимации изображения иерархической последовательностью квазиоптимальных приближений требуемым критерием является максимальная близость приближений к изображению по суммарной квадратической ошибке  $E$  или среднеквадратическому отклонению  $\sigma$ . Без тщательного исследования возможностей применения указанного критерия, вероятно, нельзя обойтись, поскольку он является простейшим.

Особенности целочисленной оцифровки [4, 5] вещественных средних значений яркости при преобразовании квазиоптимальных приближений в инвариантные представления изображения планируется рассмотреть в самостоятельной статье.

\* \* \* \* \*

## ЛИТЕРАТУРА

1. Gurevich I., Trusova Yu., Yashina V. The challenges, the problems and the tasks of the descriptive approach to image // Proc. 11-th. Int. Conf. on Pattern Recognition and Image Analysis: New Information Technologies (PRIA-11-2013), Sept. 23-28. 2013. V. 1. P. 30–35.
2. Визильтер Ю.В., Желтов С.Ю. Проблемы технического зрения в современных авиационных системах // Механика, управление и информатика. 2011. № 3. С. 11–41.
3. Абламайко С.В., Недзьведь А.М., Белоцерковский А.М., Руцкая Е.А. Сегментация трехмерных изображений компьютерной томографии на основе глобально-локальной информации // Вестник Белорусского государственного университета. 2009. № 1. С. 58–64.

4. Мельников П.А., Копылов А.В. Алгоритм поиска квазиоптимальной разметки для обработки изображений с построчным комбинированием переменных // Известия Тульского государственного университета. Естественные науки. 2012. № 1. С. 119–129.
5. Харинов М.В. Запоминание и адаптивная обработка информации цифровых изображений / Под ред. Юсупова Р.М. СПб.: СПбГУ, 2006. 138 с.
6. Харинов М.В. Обобщение трех подходов к оптимальной сегментации цифрового изображения // Труды СПИИРАН. 2013. Вып. 2. № 25. С. 294–316.
7. Dvoenko S.D. Meanless k-means as k-meanless clustering with the bi-partial approach // Proc. of the 12th Int. Conf. “Pattern Recognition and Information Processing (PRIP’2014)”. Minsk, 2014. P. 50–54.
8. Dvoenko S.D. Clustering of a set of objects // Proc. of the 9th Int. Conf. “Pattern Recognition and Information Processing (PRIP’2007)”. Minsk, 2007. P. 93–97.
9. Двоенко С.Д. Неиерархический дивизимный алгоритм кластеризации // Автоматика и телемеханика. 1999. № 4. С. 117–124.
10. Бугаев А.С., Хельвас А.В. Поисковые исследования и разработка методов и средств анализа и автоматического распознавания потоковой информации в глобальных информационных системах. Шифр “Лацкан” // Отчет по НИР. М.: МФТИ, 2001. Т. 1. 140 с.
11. Mumford D., Shah J. Boundary detection by minimizing functionals, I // Proc. IEEE Comput. Vision Patt. Recogn. Conf., San Francisco. 1985. P. 22–26.
12. Redding N.J., Crisp D.J., Tang D.H., Newsam G.N. An efficient algorithm for Mumford–Shah segmentation and its application to SAR imagery // Proc. Conf. “Digital Image Computing Techniques and Applications (DICTA’99)”. 1999. P. 35–41.
13. Koepfler G., Lopez C., Morel J. A Multiscale Algorithm for Image Segmentation by Variational Method // SIAM J. on Numerical Analysis. 1994. V. 31. № 1. P. 282–299.
14. Ward J.H., Jr. Hierarchical grouping to optimize an objective function. // J. Am. Stat. Assoc. 1963. V. 58. Issue 301. P. 236–244.
15. Jain A.K. Data Clustering: 50 Years Beyond K-Means // Pattern Recognition Letters. 2010. V. 31. № 8. P. 651–666.
16. Kharinov M.V. Image Segmentation Using Optimal and Hierarchical Piecewise-Constant Approximations // Pattern Recogn. Image Anal.: Adv. Math. Theory Appl. 2014. V. 24. № 2. P. 409–517.
17. Kharinov M.V. Adaptive Dichotomous Image Segmentation Toolkit // Pattern Recogn. Image Anal.: Adv. Math. Theory Appl. 2012. V. 22. № 1. P. 228–235.