УДК 621.321

М.В. Харинов

ОБОБЩЕНИЕ ТРЕХ ПОДХОДОВ К ОПТИМАЛЬНОЙ СЕГМЕНТАЦИИ ЦИФРОВОГО ИЗОБРАЖЕНИЯ

Харинов М.В. **Обобщение трех подходов к оптимальной сегментации цифрового изображения.**

Аннотация. В статье предлагается аналитически обоснованный метод кластеризации мультимножеств, названный *K*–методом, который в кластерном анализе позволяет превзойти традиционный метод *K*–средних. В области сегментации изображений предлагаемый метод решает проблему вычисления оптимальных приближений изображения в последовательном числе яркостных градаций, рассматриваемую в мультипороговом методе Оцу, и кардинально улучшает по суммарной квадратичной ошибке приближения изображения связными сегментами, рассматриваемые в модели Мамфорда–Шаха.

Если традиционный метод *K*–средних анализирует близость пикселей к центрам кластеров, то *K*–метод учитывает более сильный признак устойчивости оптимального разбиения относительно реклассификации пикселей из одного кластера в другой. При этом *K*–метод оказывается практичнее метода Оцу, т.к. при вычислении каждого последующего разбиения с очередным числом кластеров не ограничен экспоненциальным возрастанием продолжительности обработки. В сравнении с моделью Мамфорда–Шаха, основное преимущество *K*–метода состоит в снижении суммарной квадратичной ошибки за счет генерации последовательности перекрывающихся разбиений в комбинированном алгоритме слияния/дробления–коррекции сегментов изображения.

Ключевые слова: суммарная квадратичная ошибка, метод Оцу, метод *K*–средних, модель Мамфорда–Шаха.

Kharinov M.V. **A generalization of three approaches to an optimal segmentation of digital image.**

Abstract. The paper proposes an analytically justified method for clustering of multisets, called *K*–method, which in the cluster analysis provides to surpass the conventional *K*–means method. In the image segmentation domain it solves the problem of optimal image approximating with the sequential numbers of intensity gradations, which is posed in multi–threshold Otsu method, and essentially improves in the total square error the sequence of approximations of the image with connected segments that are treated in the Mumford–Shah model.

While the conventional *K*–means method analyzes the proximity of pixels to the cluster centers, our *K*–method treats much stronger feature of the optimal partition, namely stability relative to reclassification of pixels from one cluster to another. All other things being equal, *K*–method turns out more efficient than Otsu method, since in the calculation of the series of the partitions into sequentially increasing cluster number it doesn't face the exponential increase of processing time. In comparison with Mumford–Shah model, the main advantage of *K*–method consists in the reduction of total square error due to the generation of the sequence of overlapping partitions by means of merge–and–correct, split–and–correct or composite technique.

Keywords: total squared error, *K*–means method, Otsu method, Mumford–Shah model.

1. Введение. Есть основания предполагать, что теория и практика автоматического выделения объектов на цифровых изображениях в оптимизированном представлении [1–19] в ближайшем будущем выйдет на новые рубежи, опираясь на технологический скачок, который произошел в начале текущего столетия в развитии аппаратного и программного обеспечения. Если еще в последнем десятилетии прошлого века свободному программированию обработки изображений мешали ограниченные вычислительные ресурсы, то, по крайней мере, оперативная память современного компьютера оказывается достаточной и даже избыточной для решения аналогичных задач, хотя размеры изображений возросли. С другой стороны, за последнее десятилетия компьютеры научились «видеть», что выражается в наглядности результатов современного автоматического выделения объектов [5–7], сменивших более ранние результаты [19], которые на стадии выделения объектов нередко пропускали наиболее заметные из них. Поэтому в настоящее время для развития теории обработки изображений на первоначальной стадии выделения объектов становится особенно актуальным аналитическое обобщение и развитие существующих теоретических подходов [1, 2, 4], [8], [11, 12], [17, 18] и программно–аппаратных решений [5–7], [19].

В современной области автоматического анализа и распознавания цифровых изображений одной из наиболее критичных проблем является так называемая «проблема сегментации». С практической точки зрения она состоит в разработке робастных алгоритмов разделения пикселей изображения на *сегменты* из связных пикселей или *кластеры* из не обязательно связных пикселей, отвечающие изображенным объектам. Под робастностью обычно понимается свойство алгоритма эффективно выделять конкретные объекты независимо от модификации изображения при различных условиях съемки, например, при изменении освещения сцены или масштаба изображения*. В общем случае, при заранее не заданных объектах, требуется в результате сегментации выделить все объекты, имеющиеся на изображении. Но тогда нуждается в формальном уточнении постановка задачи создания робастного алгоритма, который, с одной стороны, не пропускает объектов, а, с другой стороны, для изменчивого изображения дает приблизительно одинаковые разбиения.

Для формализации постановки задачи сегментации изображений достаточно предположить, что свойством робастности, в первую очередь, обладают не алгоритмы разделения множества пикселей изобра-

---

* В литературе в обсуждаемом смысле, помимо термина «робастность», употребляется также термин «устойчивость», который в настоящей статье используется в самостоятельном значении.

жения на кластеры, а искомые разбиения, если они по качеству близки к некоторым оптимальным. Стандартной оценкой качества разбиения в кластерном анализе является суммарная квадратичная ошибка [17, 18], которая описывает отличие изображения от своего кусочно–постоянного (ступенчатого) приближения, в англоязычных источниках обозначаемого термином «piecewise constant approximation» [1–3, 5, 6]. Хотя современные компьютеры позволяют получать разбиения цифровых изображений, реально близкие к оптимальным по критерию суммарной квадратичной ошибки, однако эта возможность используется пока недостаточно, отчасти в силу недостатка аналитического обоснования сегментации (кластеризации) пикселей.

Целью настоящей статьи является аналитическое обобщение метода *K*–средних, метода Оцу и модели Мамфорда–Шаха, применяемых для вычисления оптимальных приближений цифрового изображения.

2. Задача оптимизации. Классифицируемыми элементами служат пиксели цветового или мультиспектрального изображения, которые составляют мультимножества [20], т. е. могут повторяться и в обсуждаемом контексте не образуют линейного пространства. Тем не менее, поскольку мы избегаем в формулах явного учета повторений пикселей, все аналитические выражения читаются так же, как и в случае обычных неповторяющихся векторов.

Пусть x_i — многомерное значение пикселя изображения, имеющего координату i, которая принимает значение от 0 до $N-1$, где N — число пикселей в изображении. Полагаем что, в общем случае, значение пикселя составляют вещественные числа, которые удобно представлять упорядоченными в столбец.

Пусть пиксели изображения разбиты на Q кластеров посредством функции $f(i)$, сопоставляющей каждой координате i номер кластера q в диапазоне от 1 до Q.

Тогда среднее значение I_q, пикселей, отнесенных к кластеру $q = 1, 2, ..., Q$, выражается в виде:

$$I_q = \frac{\sum_{f(i)=q} x_i}{n_q}, \tag{1}$$

где n_q — число пикселей в q-ом кластере, а суммирование выполняется по всем координатам i, помеченных значением q.

Пикселям y_i *кусочно–постоянного приближения* изображения приписываются значения $y_i = I_{f(i)}$. Под *суммарной квадратичной ошибкой* E понимается сумма квадратов отклонений значений x_i пикселей изображения от значений y_i пикселей кусочно–постоянного приближения:

$$E = \sum_{i=0}^{N-1} \left\| x_i - y_i \right\|^2 , \tag{2}$$

где знак $\| \ \|^2$ обозначает суммирование квадратов отклонений $x_i - y_i$ (евклидову норму «векторов» по цветовым или спектральным компонентам).

В аналитических выкладках и практических вычислениях удобно использовать аддитивное выражение суммарной квадратичной ошибки:

$$E = \sum_{q=1}^{Q} E_q , \tag{3}$$

в котором квадратичная ошибка E_q для q-го кластера подсчитывается по формуле:

$$E_q = \sum_{f(i)=q} \left\| x_i \right\|^2 - n_q \cdot \left\| I_q \right\|^2 . \tag{4}$$

При анализе и сравнении кусочно–постоянных приближений изображения вместо суммарной квадратичной ошибки E обычно пользуются среднеквадратичным отклонением приближения от изображения σ, которое выражается через E в виде:

$$\sigma = \sqrt{\frac{E}{N}} . \tag{5}$$

В базовой постановке задача оптимизации решается для каждого числа кластеров $Q = 1, 2, ..., N$ и состоит в получении оптимального разбиения изображения на кластеры $f(i)$ или соответствующего оптимального приближения y_i, для которого величина суммарной или среднеквадратичной ошибки достигает минимального значения:

$$E = \min \quad \Leftrightarrow \quad \sigma = \min . \tag{6}$$

Эквивалентным условием, активно используемым на практике [9], является максимальная величина взвешенной суммы квадратов средних по кластерам значений пикселей:

$$\sum_{q=1}^{Q} n_q \cdot \left\| I_q \right\|^2 = \max . \qquad (7)$$

В простейшем случае, к искомым кластерам не предъявляется дополнительных ограничений, как, например, в [8–10]. Тем не менее, даже в этом случае, точное решение сформулированной задачи оказывается для произвольного изображения далеко не всегда доступным. Поэтому речь идет о практическом вычислении оптимальных приближений с некоторой погрешностью, которая может быть сведена к предельному минимуму, например, для стандартных изображений из общеупотребительной базы изображений «The USC-SIPI Image Database» (http://sipi.usc.edu/database/).

В более сложном случае *условной* оптимизации, как в [1–7], на кластеры искомого оптимального разбиения накладываются дополнительные условия, например условие связности пикселей в кластерах. При этом различные пиксели изображения считаются *соседствующими* или *связными*, если находятся на минимальном расстоянии друг от друга. Тогда *условие связности* пикселей в кластере состоит в том, что каждый кластер либо содержит единственный пиксель, либо состоит из пикселей, каждый из которых соседствует с другим пикселем из этого кластера. Такие кластеры именуется *сегментами*. В обобщенном условии *ослабленной связности* требуется, чтобы каждый пиксель данного кластера либо был единственным, либо имел другой пиксель из данного кластера на расстоянии, не превышающем установленного порога. Тогда при минимальном значении порога искомые кластеры являются сегментами из связных пикселей, а при увеличении порога могут состоять из несвязных пикселей.

Таким образом, в общем случае результатом оптимизации является последовательность разбиений, вычисленных для каждого числа кластеров Q от 1 до N. Допускается, что для того или иного числа кластеров оптимальное разбиение может вычисляться неоднозначно. Очевидно, разбиения на $Q=1$ и $Q=N$ кластеров являются тривиальными оптимальными разбиениями. Нетрудно сообразить, что для оптимальных разбиений, суммарная квадратичная ошибка E и среднеквадратичное отклонение σ с ростом числа кластеров Q не возрастает и удовлетворяет условию монотонности:

$$f_Q(i),\ f_{Q+1}(i):\quad E\big|_Q \geq E\big|_{Q+1} \quad \Leftrightarrow \quad \sigma\big|_Q \geq \sigma\big|_{Q+1}. \tag{8}$$

Немонотонная зависимость $E\big|_Q < E\big|_{Q+1}$, вычисленных для разбиений $f_Q(i)$ и $f_{Q+1}(i)$, указывает на погрешность вычислений, которую можно компенсировать, если разделить надвое один из кластеров в разбиении $f_Q(i)$ и получившимся разбиением из $Q+1$ кластеров заменить разбиение $f_{Q+1}(i)$.

Пиксели в цифровом изображении обычно повторяются. Очевидно, что при безусловной оптимизации разбиений на кластеры множество, в котором встречаются M различных пикселей, оптимально разбивается на M кластеров с нулевой суммарной квадратичной ошибкой $E=0$, если одинаковые пиксели отнести к одному кластеру. Оптимальные приближения с числом кластеров от $M+1$ до N, точно аппроксимирующие изображение с $E=0$, получаются измельчением предыдущего, и нетривиальная задача сводится к получению последовательности оптимальных разбиений с числом кластеров от 1 до M.

Для практических целей наиболее важным является получение оптимальных или близких к оптимальным разбиений для начальных значений числа кластеров Q, ориентировочно, в пределах 30 усредненных градаций в приближениях изображения пикселями без учета связности, или в пределах 1000 сегментов в приближениях изображения разбиениями из связных пикселей.

3. Общее описание слияния и коррекции пикселей. Рассмотрим разбиение изображения на Q кластеров, которое описывается суммарной квадратичной ошибкой E. Преобразуем его в тривиальное посредством слияния $L \leq Q$ кластеров в один. При этом приращение ΔE_{merge} суммарной квадратичной ошибки E определяется разностью квадратичной ошибки $E(1 \cup 2 \cup \ldots \cup L)$ для результата слияния кластеров $1, 2, \ldots, L$ и суммы квадратичных ошибок $E(1),\ E(2), \ldots E(L)$ по каждому кластеру:

$$\Delta E_{merge} = E(1 \cup 2 \cup \ldots \cup L) - E(1) - E(2) \ldots - E(L). \tag{9}$$

Аналитически ΔE_{merge} вычисляется как взвешенная сумма квадратов расстояний между центрами кластеров $\left\| I_p - I_q \right\|^2$ с весовыми коэффициентами $n_p n_q$, деленная на удвоенное общее число пикселей в рассматриваемых кластерах:

$$\Delta E_{merge} = \frac{1}{2} \cdot \frac{\sum_{p,\, q=1}^{L} n_p n_q \left\| I_p - I_q \right\|^2}{\sum_{q=1}^{L} n_q}, \tag{10}$$

где номера кластеров p и q пробегают значения от 1 до L.

Рассмотрим объединение L кластеров, и преобразуем его разбиение на кластеры $1, 2, ..., L$ в некоторое разбиение $1', 2', ...$ того же самого подмножества пикселей: $1 \cup 2 \cup ... \cup L = 1' \cup 2' \cup ...$. Пусть ΔE — сопутствующее приращение суммарной квадратичной ошибки E. Тогда ΔE_{merge} есть сумма ΔE и приращения квадратичной ошибки $\Delta E'_{merge}$, отвечающего слиянию преобразованных кластеров:

$$\Delta E_{merge} = \Delta E + \Delta E'_{merge}. \tag{11}$$

Очевидно, величина ΔE может быть как положительной, так и отрицательной, в отличие от ΔE_{merge} и $\Delta E'_{merge}$.

В частном случае преобразования разбиения рассматриваемого подмножества пикселей из L кластеров в разбиение с тем же самым числом L штрихованных кластеров: $1 \to 1'$, $2 \to 2'$,..., $L \to L'$, величина ΔE в (11) обозначается как $\Delta E_{correct}$ и трактуется как приращение суммарной квадратичной ошибки в результате коррекции разбиения множества пикселей $1 \cup 2 \cup ... \cup L \equiv 1' \cup 2' \cup ... \cup L'$, где L' обозначает преобразованный кластер под номером L. При этом приращение $\Delta E_{correct}$ суммарной квадратичной ошибки при коррекции разбиения совпадает с разностью приращений суммарной квадратичной ошибки, отвечающих слиянию L кластеров до и после коррекции:

$$\Delta E_{correct} = \Delta E_{merge} - \Delta E'_{merge}, \tag{12}$$

где ΔE_{merge} и $\Delta E'_{merge}$ подсчитываются по формуле (10).

Положим, что преобразование коррекции выполняется только при отрицательном значении $\Delta E_{correct}$ и при неизменном числе кластеров сопровождается уменьшением суммарной квадратичной ошибки. Таким образом, критерием коррекции служит условие:

$$\Delta E_{correct} < 0\,. \tag{13}$$

Из (12), (13) следует, что оптимальные разбиения изображения при каждом числе кластеров $Q = 1, 2, ..., N$ характеризуются не только минимумом квадратичной ошибки, но также максимумом приращения суммарной квадратичной ошибки ΔE_{merge}, и задача минимизации суммарной квадратичной ошибки эквивалентна задаче максимизации ее приращения при слиянии всех кластеров изображения в единственный кластер тривиального разбиения. При этом, в силу аддитивности (3) суммарной квадратичной ошибки E, получаем, что в случае оптимального разбиения пикселей изображения на Q кластеров, обсуждаемые условия выполняются не только для всего множества кластеров, но и для любого его подмножества в виде объединения $L \le Q$ кластеров, которое можно рассматривать как множество пикселей самостоятельного изображения:

$$E\big|_L = \min \quad \Leftrightarrow \quad \Delta E_{merge}(1 \cup 2 \cup ... \cup L) = \max . \tag{14}$$

Нетрудно убедиться, что, если пиксели изображения разбиваются всего на два кластера ($L = Q = 2$, $N \equiv n_1 + n_2$), то условия минимизации $E\big|_2$ и максимизации $\Delta E_{merge}(1 \cup 2)$ в (14) оказываются равносильны условиям, установленным, в [8]:

$$\sigma_W^2 \equiv \frac{E\big|_2}{n_1 + n_2} = \min \quad \Leftrightarrow \quad \sigma_B^2 \equiv \frac{\Delta E_{merge}(1 \cup 2)}{n_1 + n_2} = \max , \tag{15}$$

где $\sigma_W^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2$ — так называемая, взвешенная суммарная дисперсия внутри классов 1 и 2, $\sigma_B^2 = w_1 w_2 \|I_1 - I_2\|^2$ — дисперсия между классами 1 и 2, $\sigma_1^2 = \frac{E(1)}{n_1}$ и $\sigma_2^2 = \frac{E(2)}{n_2}$ — дисперсии внутри классов 1 и 2, а $w_1 = \frac{n_1}{n_1 + n_2}$ и $w_2 = \frac{n_2}{n_1 + n_2}$ — вероятности встретить на изображении пиксели из классов 1 и 2, соответственно[†].

В продолжение сопоставления с известными решениями, здесь же уместно обратить внимание, что в версии [4] модели Мамфорда–Шаха без учета границ между сегментами величина $\Delta E_{merge}(1 \cup 2)$ служит минимизируемым критерием слияния смежных сегментов, что, на пер-

[†] В отличие от [8, 9], формулы (15) интерпретируют метод Оцу в обобщенном для многомерных пикселей виде.

вый взгляд, противоречит (15). На самом деле, слияние сегментов с минимальной величиной критерия $\Delta E_{merge}(1 \cup 2)$ необходимо для того, чтобы в результирующем разбиении оставались сегменты с максимальными значениями $\Delta E_{merge}(1 \cup 2)$.

4. Формула реклассификации. Рассмотрим реклассификацию k пикселей с центральным (средним) значением I, исключаемых из кластера 1 с центром I_1 и включаемых в число пикселей кластера 2 с центром I_2. Сопутствующее приращение ΔE суммарной квадратичной ошибки E выражается, так называемой, формулой реклассификации:

$$\Delta E = \begin{cases} \Delta E_{merge}(1 \cup 2), & k = n_1, \\ \Delta E_{correct}(1, 2), & k < n_1, \end{cases} \tag{16}$$

где n_1, n_2 — число пикселей кластерах 1 и 2, соответственно, а $\Delta E_{merge}(1 \cup 2)$ и $\Delta E_{correct}(1, 2)$ имеют вид:

$$\Delta E_{merge}(1 \cup 2) = \frac{\|I_1 - I_2\|^2}{\dfrac{1}{n_1} + \dfrac{1}{n_2}}, \tag{17}$$

$$\Delta E_{correct}(1, 2) = \frac{\|I - I_2\|^2}{\dfrac{1}{k} + \dfrac{1}{n_2}} - \frac{\|I - I_1\|^2}{\dfrac{1}{k} - \dfrac{1}{n_1}}, \tag{18}$$

где выражение (17) непосредственно следует из (10), а выражение (18) выводится из (10) и (12) посредством элементарных преобразований.

$\Delta E_{merge}(1 \cup 2)$ описывает приращение суммарной квадратичной ошибки при слиянии кластеров или, иначе, реклассификации всех пикселей из кластера 1 в кластер 2, и уменьшении числа множеств на единицу. $\Delta E_{correct}(1, 2)$ описывает приращение суммарной квадратичной ошибки при реклассификации части пикселей из кластера 1 в кластер 2 и сохранении числа кластеров.

Особенностью операции реклассификации пикселей, описываемой формулами (16)–(18), является то, что она распадается на две операции преобразования разбиения изображения, выполняемые посредством слияния и коррекции кластеров. Для получения последовательности оптимальных разбиений изображения, очевидно, необходимы обе операции, если искомая последовательность не является иерархической. В модели Мамфорда–Шаха [1–7] для сегментации изображений используется только слияние сегментов согласно критерию $\Delta E_{merge}(1 \cup 2) = \min$ в [4], что в приложении к произвольным цифровым изображениям заведомо ограничивает возможности модели и не компенсируется ни модификациями критерия ([3], [5, 6]), ни вариантами минимизируемого функционала ([1–4] и др.).

Пользуясь (18), раскроем критерий коррекции (13) в явном виде. При этом для реклассификации k пикселей из кластера 1 в кластер 2 получаем:

$$\|I - I_1\| > \alpha \cdot \|I - I_2\|, \tag{19}$$

где коэффициент $\alpha<1$ описывает соотношение количеств пикселей в рассматриваемых кластерах 1, 2 и определяется в виде:

$$\alpha \equiv \sqrt{\frac{n_2(n_1 - k)}{n_1(n_2 + k)}}\,. \tag{20}$$

5. *K*–метод vs. *K*–средних метод. Обсуждаемые формулы подсказывают очевидную альтернативу методу *K*–средних, которую мы назвали *K*–методом.

В *K*–методе реклассификация выполняется без выделения центров кластеров в качестве представителей кластера в целом, как это делается в методе *K*–средних. При этом мы предполагаем, что каждый кластер разбивается на вложенные подмножества, например, подмножества одинаковых пикселей или отдельные пиксели. Далее, в простейшем случае, мы рассматриваем пары кластеров или некоторые подмножества пар кластеров, например, пары смежных сегментов изображения, которые, в общем случае, перекрываются друг с другом.

Последовательность разбиений с различным числом кластеров, в частности, сегментов, получается посредством слияния или разбиения кластеров текущего разбиения, но с тем отличием, что выполнение (19) для той или иной рассматриваемой пары кластеров является условием начала или продолжения предварительной коррекции. Формула (18) используется для выбора варианта реклассификации из нескольких возможных, что по сравнению с методом *K*–средних позволяет получить более эффективную оптимизацию.

В отличие от версий [1–7] модели Мамфорда–Шаха, метод *K*–средних поддерживает оптимизацию перекрывающихся разбиений с различным числом кластеров, которые получаются благодаря добавлению кластерных центров [14] и итеративной реклассификации пикселей, относимых к кластеру с ближайшим центром. Тем не менее, в методе *K*–средних [11–16] при реклассификации пикселей не учитывается коэффициент α, что, вообще говоря, влечет преждевременный останов реклассификации, а при вариантах реклассификации препятствует наибольшему снижению суммарной квадратичной ошибки, которое оценивается по формуле (18) (см. таблицу).

Сравнение с методом *K*–средних

	метод *K*–средних	*K*–метод
Условие выполнения	$\|I-I_1\| > \|I-I_2\|$	$\|I-I_1\| > \alpha \cdot \|I-I_2\|$
Условие выбора	$\|I-I_2\| = \min$	$\dfrac{\|I-I_2\|^2}{\frac{1}{k}+\frac{1}{n_2}} - \dfrac{\|I-I_1\|^2}{\frac{1}{k}-\frac{1}{n_1}} = \min$

Таблица поясняет преимущества нашего метода в сравнении с методом *K*–средних [11–16] В первой строке указаны условия начала или продолжения процесса реклассификации пикселей, которая выполняется, если находится кластер 2. Вторая строка описывает выбор кластера 2 из нескольких возможных.

По поводу формулы (20) для коэффициента $\alpha < 1$ следует добавить, что коэффициент α уменьшается с ростом k. Поэтому реклассификация наборов пикселей способствует уменьшению суммарной квадратичной ошибки E по сравнению с ее уменьшением при реклассификации отдельных пикселей. В частности, если реклассификация данного пикселя из кластера 1 в кластер 2 уменьшает E, то реклассификация остальных пикселей, значения которых совпадают со значением данного пикселя, также уменьшает E. Поэтому в нашем методе предусматривается реклассификация множеств пикселей, что обеспечивает дополнительное преимущество по сравнению с методом *K*–средних, в котором для выполнения коррекции кластеры «рассыпаются» на отдельные пиксели.

6. Устойчивая сегментация. Введенная нами операция коррекции выполняется до тех пор, пока посредством предусмотренных преобразований оказывается возможным улучшить приближение изображения по суммарной квадратичной ошибке согласно критерию (13) или формуле (19), применяемой при анализе пар кластеров. Результатом коррекции является устойчивое разбиение изображения.

Условимся называть разбиение пикселей на кластеры, равно как и соответствующее приближение изображения, *устойчивым*, если предусмотренные при фиксированном числе кластеров преобразования коррекции разбиения увеличивают суммарную квадратичную ошибку E или оставляют ее неизменной, что выражается *условием устойчивости*:

$$\Delta E_{correct} \geq 0\,, \tag{21}$$

которое конкретизируется для реклассификации k пикселей из одного кластера в другой посредством логического отрицания (19):

$$\|I - I_1\| \leq \alpha \cdot \|I - I_2\|\,. \tag{22}$$

Обсуждаемую устойчивость следует отличать от неформального употребления этого термина в смысле робастности результатов алгоритма сегментации по отношению к изменчивым изображениям, настроечным параметрам и начальным условиям обработки. В отличие от алгоритмической устойчивости [21], мы рассматриваем устойчивость состояния, которая определяется, как устойчивость равновесия, по аналогии с устойчивостью механической системы в физике, если суммарную квадратичную ошибку E считать аналогом потенциальной энергии.

Очевидно, оптимальные разбиения на кластеры устойчивы, но устойчивые разбиения не обязательно оптимальны. При этом, согласно формулам в приведенной выше таблице, метод *K*–средних, хотя и предусматривает снижение суммарной квадратичной ошибки E при фиксированном числе сегментов, но, в общем случае, не гарантирует получения устойчивых и, тем более, оптимальных разбиений. В сравнении с методом *K*–средних, модель Мамфорда–Шаха [1–7] может уступать по минимизации суммарной квадратичной ошибки, так как вовсе не предусматривает механизма снижения E при фиксированном числе сегментов. Однако модель Мамфорда–Шаха разработана специально для сегментации изображений и усиливается в составе нашего метода за счет дополнительной операции коррекции сегментов и получения устойчивых разбиений.

7. Прототипы решения. В качестве основы для развития вычисления оптимальных приближений цифрового изображения мы рассматривали метод Оцу [8, 9], в котором выполняется кластеризация пикселей независимо от их геометрического распределения в изображении, и модель Мамфорда–Шаха [4–6] сегментации цифрового изображения, в которой вычисляется иерархия кусочно–постоянных приближений изображения последовательным числом связных сегментов.

Метод Оцу [8–10] направлен на вычисление оптимальных разбиений изображения, характеризуемых наименьшими значениями E, которые в принципе могут быть достигнуты. Во избежание полного перебора всевозможных разбиений, в методе Оцу учитывается то, что в оптимальных разбиениях одинаковые пиксели принадлежат одному кластеру, а значения пикселей заполняют «компактные» диапазоны шкалы яркости, число которых совпадает с числом кластеров или градаций средних по кластерам значений яркости пикселей изображения. При этом расчеты, хотя и ограничиваются перебором разделяющих диапазоны порогов яркости, но, тем не менее, при увеличении числа кластеров сталкиваются с экспоненциальным возрастанием продолжительности обработки и позволяют получить оптимальные приближения реальных изображений не более чем с десятком градаций [9]. С точки зрения продолжения вычислений при большем числе градаций, явным недостатком версии [9] мультипорогового метода Оцу является независимое вычисление оптимальных разбиений для различного числа кластеров, который, впрочем, нетрудно исправить, например, за счет вычисления очередного разбиения посредством разделения одного из кластеров надвое или слияния подходящей пары кластеров оптимизированного разбиения, полученного на предыдущем шаге обработки, как это делается в модели Мамфорда–Шаха [1–7] и др.

В модели Мамфорда–Шаха решается задача условной оптимизации разбиений пикселей изображения по свойствам геометрического распределения пикселей. Последовательность разбиений получают посредством слияния пар смежных сегментов, причем свойство «компактности» целевых разбиений непосредственно не принимается во внимание [1–7]. Помимо аналитически оправданного в [1, 2] критерия слияния сегментов (в виде (17) в [4] или вариантах [3], [5–7], и др.), вычисления в модели Мамфорда–Шаха отличают иерархичность получаемых разбиений изображения и связность сегментов. По всей видимости, указанные особенности необходимо принимать во внимание при творческом обобщении и развитии модели, но учитывать их как возможности, а не как ограничения.

По нашему опыту в программной реализации модели Мамфорда–Шаха удобно применять структуру данных динамических деревьев Слэйтора–Тарьяна [19, 22–24], которые поддерживают слияние произвольных множеств пикселей. Тогда при единственной операции слияния смежных сегментов [1–7] условие связности сегментов поддерживается автоматически, если начальным разбиением изображения служит разбиение на связные сегменты, например, на отдельные пиксели или сегменты из одинаковых пикселей. В то же время сохраняется работоспособность программ при использовании начальных разбиений множества пикселей изображения на кластеры из несвязных сегментов, а также при выполнении операции коррекции, которая, вообще говоря, нарушает связность сегментов. В случае необходимости, для сохранения связности сегментов при слиянии сегментов в сочетании с коррекцией разбиения изображения, достаточно блокировать реклассификацию пикселей, приводящую к нарушению связности сегментов. С другой стороны, появляются возможности учитывать связность сегментов в ослабленном варианте.

По поводу операции слияния сегментов, которой обычно ограничиваются в практике сегментации по Мамфорду–Шаху [5–7] следует заметить что, помимо операции коррекции, ее полезно дополнить операцией дробления сегментов, которая для любого сегмента изображения определяется обращением процедуры итеративного слияния пикселей. Действительно, для оптимального разбиения изображения на сегменты можем записать:

$$\Delta E_{split}(1) + \Delta E_{merge}(2 \cup 3) \geq 0\,, \tag{23}$$

где 1 — составной сегмент разбиения, разделение которого надвое сопровождается минимальной величиной неположительного приращения $\Delta E_{split}(1) \leq 0$ суммарной квадратичной ошибки, а $2, 3$ — пара смежных сегментов, слияние которых минимально повышает суммарную квадратичную ошибку на величину $\Delta E_{merge}(2 \cup 3) \geq 0$ [‡].

Формула (23), очевидно, справедлива, так как, в противном случае, разбиение может быть оптимизировано по суммарной квадратичной ошибке разделением надвое сегмента 1 и последующим слиянием сегментов $2, 3$. Следовательно, для снижения суммарной квадратичной ошибки перед слиянием сегментов имеет смысл обеспечить выполнение условия (23) за счет комбинированной операции дробления и слияния сегментов, применяемой при нарушении (23), например, при совпадении сегментов $2, 3$ по средней яркости и наличии составных сегментов, разделяющихся на два вложенных с различной средней яркостью. При этом комбинированная операция парного слияния и дробления сегментов надвое обеспечивает монотонное возрастание сум-

[‡] Чтобы не загромождать изложение деталями, мы опускаем обсуждение случая, когда оптимальным оказывается разделение сегмента надвое и слияние одной из полученных частей с другим сегментом.

марной квадратичной ошибки (8) и способствует также возрастанию приращения суммарной квадратичной ошибки при убывании числа сегментов.

Таким образом, в *K*–методе, наряду с операцией коррекции, предусматривается применение операций слияния и дробления кластеров, и сохраняется преемственность вычислений со связными сегментами и иерархическими разбиениями, рассматриваемыми в модели Мамфорда–Шаха, а также перекрывающимися разбиениями множества пикселей изображения на кластеры, рассматриваемыми в методе Оцу.

8. Назначение и отличительные особенности *K*–метода. *K*–метод предназначен для практического вычисления оптимальных и близких к оптимальным приближений многомерных данных (мультимножеств), в частности, кластеризации пикселей цветовых и многоспектральных изображений на основе аналитического обоснования минимизации суммарной квадратичной ошибки E.

В области кластерного анализа [17, 18] посредством анализа пар кластеров предлагаемый метод, при прочих равных условиях, является более сильной альтернативой общеупотребительному методу *K*–средних [11–16], в котором для оптимизации разбиений по суммарной квадратичной ошибке E при фиксированном числе кластеров анализируется расстояние до центров кластеров вместо более точной величины $\Delta E_{correct}(1,2)$ из (18), а генерация разбиений с изменением числа кластеров выполняется посредством эвристического приема добавления кластерных центров [14]. В модели Мамфорда–Шаха [4] для вычислений с изменением числа кластеров анализируется величина $\Delta E_{merge}(1 \cup 2)$ из (17), которая используется в качестве минимизируемого критерия слияния. В методе Оцу [8] эта величина рассматривается в качестве максимизируемой в процессе коррекции, что используется для разработки частного приема оптимизации расчетов. Однако, как в модели Мамфорда–Шаха, так и методе Оцу упускается возможность применения для коррекции выражения (18) для $\Delta E_{correct}(1,2)$, что в методе Оцу [9] компенсируется перебором яркостных порогов, который, в свою очередь, вытесняет аналитическое развитие метода.

В случае безусловной оптимизации по суммарной квадратичной ошибке *K*–метод направлен на вычисление реально близких к оптимальным приближений изображения последовательным числом градаций яркости, что отличает его от мультипорогового метода Оцу [9], в котором решение той же самой проблемы остается умозрительным, т.к. ограничивается несколькими оптимальными приближениями из-за продолжительного по времени перебора вариантов. При этом существенное использование стратегии перебора в методе Оцу ограничивает возможности его развития по сравнению с моделью Мамфорда–Шаха [1–7], применяемой при условной оптимизации приближений изображения связными сегментами.

В сравнении с моделью Мамфорда–Шаха основным преимуществом *K*–метода является то, что он поддерживает получение последовательности разбиений изображения с перекрытиями сегментов из различных разбиений. При этом эффект снижения суммарной квадратичной ошибки E зависит от того, насколько целевая последовательность оптимальных разбиений отличается от иерархической, что подлежит экспериментальному исследованию. Для оценки перспектив улучшения модели Мамфорда–Шаха немаловажным является также вопрос о наличии в иерархической последовательности разбиений по Мамфорду–Шаху, близких к оптимальным [25]. Следующая задача, важная для унификации решения и его интерпретации, заключается в сравнении и согласовании результатов оптимизации разбиений, получаемых в модели Мамфорда–Шаха и методом Оцу. Решение перечисленных задач необходимо для развития версии [4] модели Мамфорда–Шаха в составе *K*–метода и дальнейшей его унификации с целью эффективного выполнения как условной, так и безусловной оптимизации последовательности разбиений цифрового изображения.

Если в методе *K*–средних [11–16] в качестве рабочего признака оптимального разбиения рассматривается максимальная близость пикселей к центрам кластеров, к которым они отнесены, то в *K*–методе используется более сильный признак устойчивости разбиения относительно реклассификации отдельных пикселей и их предусмотренных подмножеств из одного кластера в другой в процессе итеративной коррекции. При этом в *K*–методе результатом коррекции являются устойчивые разбиения изображения, характеризуемые в (21) неотрицательной величиной приращения суммарного квадратичного отклонения $\Delta E_{correct}$, тогда как метод *K*–средних и модель Мамфорда–Шаха [4], вообще говоря, приводят к неустойчивым результирующим разбиениям с завышенной суммарной квадратичной ошибкой.

Для того, чтобы при расчете последовательности разбиений изображения на $Q = 1, 2, ..., N$ кластеров избежать с ростом Q экспоненциального возрастания времени вычисления очередного разбиения на Q кластеров, как это имеет место в мультипороговом методе Оцу [9], в K–методе применяется оптимизация разбиений изображения по его перекрывающимся частям.

Оптимизация по *перекрывающимся частям* состоит в том, что сначала из кластеров, в частности, сегментов текущего разбиения, формируются перекрывающиеся *кортежи* (или, так называемые, n-ки). Затем для каждого кортежа, как для самостоятельного изображения, выполняется коррекция его разбиения на кластеры, в процессе которой объединенные в кортеж кластеры обмениваются подмножествами пикселей из условия уменьшения суммарной квадратичной ошибки до получения устойчивого кортежа (см. (13) и (21) для $\Delta E_{correct}$).

Преобразование разбиения данного кортежа в устойчивое, вообще говоря, нарушает устойчивость других кортежей, и процесс снижения суммарной квадратичной ошибки E возобновляется до преобразования всех рассматриваемых кортежей в устойчивые и получения соответствующего *устойчивого* разбиения изображения.

В качестве конкретных примеров кортежей, активно используемых нами в вычислениях, можно привести всевозможные пары смежных сегментов, объединения каждого сегмента со смежными в модели Мамфорда–Шаха, множества пикселей со значениями, попадающими в два, три и т.д. последовательных диапазона яркостной шкалы в методе Оцу и пр. В качестве подмножеств, на которые разделяются содержащиеся в кластерах пиксели, мы, помимо отдельных пикселей, используем подмножества пикселей одинаковой яркости. Разумеется, приведенные примеры не исчерпывают возможных вариантов способов разделения кластеров на подмножества и формирования кортежей, которые определяют творческий потенциал интерпретации устойчивости, свойственной оптимальным разбиениям изображения.

В целом предложенный K–метод получения оптимальных и близких к оптимальным разбиениям изображения можно охарактеризовать как метод кластеризации (сегментации) многомерных пикселей цифрового изображения посредством итеративного слияния/дробления кластеров, отличающийся тем, что для практического получения последовательности оптимальных разбиений, разделенные на подмножества пикселей кластеры объединяют в перекрывающиеся кортежи, и кластеры каждого кортежа обмениваются подмножествами пикселей из условия уменьшения суммарной квадратичной ошибки до получения устойчивого разбиения.

9. Экспериментальные результаты. Поскольку результаты экспериментальных исследований по обсуждаемой теме, включающие сравнительный анализ расчетов по модели Мамфорда–Шаха (версиям [4] и [5, 6]), методу Оцу [8–10] и *K*–методу недавно опубликованы в общедоступном электронном источнике arXiv.org [26, 27], мы ограничимся здесь их обзором с упором на приоритетные результаты.

В [26, 27] на примере стандартного изображения показано, что оптимальные приближения цифровых изображений эффективно вычисляются на современном компьютере, например, *K*–методом.

В качестве экспериментального подтверждения высказанного тезиса в отношении оптимизации по версии [4] модели Мамфорда–Шаха демонстрируется приближение стандартного изображения двумя связными сегментами, которое по среднеквадратичному отклонению σ отличается от оптимального не более чем на несколько процентов и именуется приближением, «близким к оптимальному». Характерно, что на данном конкретном примере удается выявить структуру оптимального приближения, образуемого «площадками» и «соединениями» шириной в один пиксель. Для более простой задачи безусловной оптимизации по Оцу [8–10] демонстрируется полная последовательность приближений стандартного изображения в последовательном числе градаций и приводятся соответствующие оценки значений σ, рассчитанные до пятого десятичного знака после запятой [26].

В результате сравнительного анализа в [26, 27] показано, что в модели Мамфорда–Шаха результирующие разбиения на связные сегменты при каждом числе сегментов Q в диапазоне от 1 до 1000, благодаря *K*–методу, могут быть существенно улучшены по среднеквадратичному отклонению посредством преобразования их в устойчивые. Однако и улучшенная последовательность устойчивых разбиений оказалась не оптимальной, так как в частном случае двух сегментов не содержит «близкого к оптимальному» приближения, которое принято в качестве контрольного результата. В результате продолжения экспериментального исследования выяснилось, что требуемого улучшения приближений по σ удается добиться посредством ослабления требования связности пикселей в сегментах, что определяет основное направление развития программно–алгоритмической реализации модели Мамфорда–Шаха в составе *K*–метода.

10. Заключение. Таким образом, из трех подходов, объединяемых в приложении к сегментации изображений, наиболее интересной для дальнейшего развития в составе *K*–метода является модель Мамфорда–Шаха, в которой, по сравнению с оригинальной версией [1, 2], упрощается функционал по версии [4], добавляется операция коррекции, используется операция дробления сегментов и предусматривается ослабление связности пикселей в сегментах.

Обращаясь к модели Мамфорда–Шаха [1–6], не следует недооценивать вычислительной проблемы, связанной с необходимостью скоростного выполнения порядка $N-1$ итераций слияния сегментов, где N — число пикселей в изображении, которая решается за счет трудоемкой оптимизации вычислений [5–7], но усугубляется при встраивании в многочисленные итерации слияния циклической же коррекции множеств пикселей. Расчеты существенно упрощаются, если выполняются в терминах динамических деревьев Слэйтора–Тарьяна [22–24], которые позволяют экономично вычислять, запоминать и преобразовывать в оперативной памяти компьютера многочисленные иерархические разбиения изображения [19], необходимые для построения перекрывающихся оптимальных разбиений. Развитие структуры данных [22–24] в приложении к задачам оптимизации приближений изображения представляет самостоятельный интерес.

Посредством имеющейся структуры данных актуально, в первую очередь, решить следующие задачи:

1. Оценить перспективы применения *K*–метода вместо метода *K*–средних в задачах обработки изображений;
2. Программно реализовать сегментацию многомерных цветовых и мультиспектральных изображений;
3. Исследовать признаки пикселей в системе яркостей оптимальных приближений изображения;
4. Изучить возможности разложения перекрывающихся разбиений в систему иерархических разбиений (в терминах динамических деревьев Слэйтора–Тарьяна).

Особый интерес представляет дальнейшее развитие аналитического аппарата *K*–метода в процессе решения перечисленных задач.

Литература

1. *Mumford D., Shah J.* Boundary detection by minimizing functionals, I // Proc. IEEE Comput. Vision Patt. Recogn. Conf., San Francisco, 1985. — pp. 22–26.
2. *Mumford D., Shah J.* Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems // Communications on Pure and Applied Mathematics, Vol. XLII, No 4, 1989. — pp. 577–685.
3. *Koepfler G., Lopez C., Morel J.* A Multiscale Algorithm for Image Segmentation by Variational Method // SIAM Journal on Numerical Analysis, Vol. 31, No 1, 1994. — pp. 282–299.
4. *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. Том 1, — 140 с.
5. *Crisp D.J., Tao T.C.* Fast Region Merging Algorithms for Image Segmentation // The 5th Asian Conf. on Computer Vision (ACCV2002), Melbourne, Australia, 23–25 January 2002. — pp. 1–6.
6. *Robinson B.J., Redding N.J., Crisp D.J.* Implementation of a fast algorithm for segmenting SAR imagery // Scientific and Technical Report, Australia: Defense Science and Technology Organization, Australia, 01 January 2002. — 42 p.
7. Environment for Visualizing Images, ENVI Feature Extracting Module User's Guide, URL: http:// www.ittvis.com/portals/0/pdfs/envi/Feature_Extraction_Module.pdf
8. *Otsu N.* A Threshold Selection Method from Gray–Level Histograms // IEEE Transactions on systems, MAN, and CYBERNETICS, Vol. SMC–9, №. 1, 1979. — pp. 62–66.
9. *Ping-Sung Liao, Tse-Sheng Chen, Pau-Choo Chung* A Fast Algorithm for Multilevel Thresholding // J. Inf. Sci. Eng. Vol. 17 (5), 2001. — pp. 713–727.
10. *Sezgin M., Sankur B.* Survey over image thresholding techniques and quantitative performance evaluation// Journal of Electronic Imaging Vol. 13 (1), 2004. — pp. 146–165.
11. *Steinhaus H.* Sur la division des corps materiels en parties // Bull. Acad. Polon. Sci., C1. III Vol. IV, 1956. — pp. 801–804.
12. *Lloyd S.P.* Least squares quantization in PCM // IEEE Transactions on Information Theory No 28 (2), 1982. — pp. 129–137.
13. *MacQueen J.B.* Some Methods for classification and Analysis of Multivariate Observations // Proc. Fifth Berkeley Symp. Math. Stat. and Probab., Berkeley: University of California Press, Vol. 1, 1967. — pp. 281–297.
14. *Likas A., Vlassis N. and Verbeek J.* The Global *K*–Means Clustering Algorithm // Pattern Recognition. Vol. 36, 2003. — pp. 451–461.
15. *Jain A.K., Murthy M.N., Flynn P.J.* Data Clustering A Review // ACM Computing Surveys, Vol. 31, No. 3, September 1999. — pp. 264–323.
16. *Jain A.K.* Data Clustering: 50 Years Beyond K–Means // Pattern Recognition Letters, Vol. 31, No. 8, 2010. — pp. 651–666.
17. *Мандель И.Д.* Кластерный анализ. М.: Финансы и статистика. 1988. —176 с.
18. *Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д.* Прикладная статистика: Классификация и снижение размерности. М.: Финансы и статистика, 1989. — 607 с.
19. *Харинов М.В.* Разработка динамических структур данных системы автоматизированного распознавания изображений / руков. В.В. Александров / Автореф. Дис. канд. технич. наук. — С.П. 1993. –20с.
20. *Кнут Д.Э.* Искусство программирования для ЭВМ. Том 2. Получисленные алгоритмы, пер. с англ., М.:Мир, 1977. — 727 с.
21. *Kuncheva L.I., Vetrov D.P.* Evaluation of stability of k–means cluster ensembles with respect to random initialization // IEEE Transactions on Pattern Analysis and Machine Intelligence, 28 (11), 2006, — pp. 1798–1808.
22. *Tarjan R.E.* Efficiency of a Good But Not Linear Set Union Algorithm // Journal of the ACM Vol. 22 (2), 1975. — pp. 215–225.
23. *Sleator D.D., Tarjan R.E.* Self–Adjusting Binary Search Trees // Journal of the ACM Vol. 32 (3), 1985. — pp. 652–686.
24. *Nock R., Nielsen F.* Statistical Region Merging // IEEE Trans. Pattern Anal. Mach. Intell. Vol. 26(11), 2004. — pp. 1452–1458.
25. *Martínez-Usó A., Pla F., García-Sevilla P.* Unsupervised Image Segmentation Using a Hierarchical Clustering Selection Process // SSPR 2006 and SPR 2006, Proc. of Joint

IAPR International Workshops, Hong Kong, China, August 17–19, 2006. — pp. 799–807.
26. *Kharinov M.V.* Stable Segmentation of Digital Images // arXiv preprint, arXiv: 208.2655, 14 August 2012. — 6 p. (in Russian). URL: http://arxiv.org/ftp/arxiv/papers/1208/1208.2655.pdf
27. *Kharinov M.V.* Reclassification formula that provides to surpass *K*–means method // arXiv preprint , arXiv:1209.6204, 28 Sep 2012. — 10 p. URL: http://arxiv.org/ftp/arxiv/papers/1209/1209.6204.pdf

Харинов Михаил Вячеславович — канд. техн. наук, доцент; старший научный сотрудник лаборатории прикладной информатики СПИИРАН. Область научных интересов: анализ цифровой информации, количественная оценка, система числового представления, идемпотентные преобразования, иерархические структуры данных, единое представление аудио- и видеосигналов для хранения, обработки и передачи, цветовое преобразование изображений. Число научных публикаций — 120. khar@iias.spb.su; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р. т. +7(812)328-1910, факс +7(812)328-4450.

Kharinov Mikhail Vyacheslavovich — PhD., associate professor; senior researcher, Laboratory of Applied Informatics, SPIIRAS. Research interests: digital information analysis, information quantity estimation, numerical representation system, idempotent transformations, hierarchical data structures, unified algorithms for audio and videosignal processing, color transformations of images. The number of publications — 120. khar@iias.spb.su; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-1919, fax +7(812)328-4450.

Рекомендовано СПИИРАН, заведующий лабораторией биомедицинских исследований СПИИРАН дтн С.Б. Рудницкий.
Статья поступила в редакцию 11.06.2012.

РЕФЕРАТ

Харинов М.В. **Обобщение трех подходов к оптимальной сегментации цифрового изображения.**

В статье рассматривается проблема оптимальной сегментации цветовых и многоспектральных изображений, состоящих из пикселей (элементов мультимножеств), для которых характерны повторения. В аналитической форме проводится сравнительный анализ метода *K*–средних, метода Оцу и модели Мамфорда–Шаха, интенсивно применяемых в практике машинного зрения на первичной стадии выделения и распознавания объектов для сегментации цифровых изображений посредством вычисления кусочно–постоянных приближений, качество которых оценивается по величине суммарной квадратичной ошибки или среднеквадратичному отклонению приближения от изображения.

Целью работы является разработка и аналитическое обоснование метода сегментации цифрового изображения, названного *K*–методом, который объединяет достоинства и компенсирует недостатки трех перечисленных.

Посредством элементарных выкладок в статье показано, что метод *K*–средних, опирающийся на правило отнесения пикселя к кластеру с ближайшим по яркости центром, недостаточен для минимизации суммарной квадратичной ошибки. При этом более сильным признаком оказывается *устойчивость* разбиения относительно реклассификации пикселей из одного кластера в другой, которая определяется в *K*–методе по аналогии с устойчивостью равновесия в физике. В качестве дополнительного признака оптимальных разбиений в *K*–методе учитывается монотонность суммарной квадратичной ошибки в зависимости от числа кластеров.

С целью преодоления характерного для мультипорогового метода Оцу экспоненциального падения скорости вычислений с ростом числа кластеров, в *K*–методе предлагается способ получения полной последовательности оптимальных разбиений изображения, согласно которому разделенные на подмножества пикселей кластеры объединяются в перекрывающиеся кортежи, и кластеры каждого кортежа обмениваются подмножествами пикселей из условия уменьшения суммарной квадратичной ошибки до получения устойчивого разбиения изображения.

Как показано в статье, в модели сегментации Мамфорда–Шаха эффективной минимизации целевого функционала, препятствует недостаточность операции слияния сегментов. Указанное ограничение модели преодолевается в *K*–методе благодаря введению дополнительной операции коррекции сегментов, а также ослаблению условия связности пикселей внутри сегмента, что, в предельном случае несвязных пикселей выражается в совпадении результатов сегментации по Мамфорду–Шаху с результатами расчетов в разработанной версии мультипорогового метода Оцу.

В статье имеется экспериментальное обоснование предложенного метода, которое обсуждается со ссылкой на иллюстративные материалы, обнародованные в общедоступном электронном источнике arXiv.org.

SUMMARY

Kharinov M.V. **A generalization of three approaches to an optimal segmentation of digital image.**

In this paper the problem of optimal segmentation of color or multispectral images consisting of repeatable pixels that are treated as elements of the multiset, is under consideration. Comparative analysis of *K*–means method, Otsu method and Mumford–Shah model is carried out using elementary formulas. A generalization of the listed methods is justified by the fact that these techniques are extensively applied in the practice of machine vision at the early stage of object detection and recognition. All are designed for segmentation of digital images by calculating the piecewise constant approximations. Finally, they are based on an assessment of the quality of the partition by the total squared error or by the standard deviation of piecewise constant approximation from an image.

The aim is to develop and analytically justified a method of image segmentation, called *K*–method, which combines the advantages and overcomes the shortcomings of above three specified methods.

As shown by elementary calculations, the conventional *K*–means method, which consists in iteratively assigning a pixel to the cluster that has the closest center constituted of averaged intensities, is insufficient to best minimize the total squared error.

As it turned out, it is better to use a stronger feature of optimal approximation, namely, a stability of the partition relative to reclassification of pixels from one cluster to another, which is defined in *K*–method by analogy with the equilibrium stability in physics. As an additional feature of optimal partition *K*–method takes into account the monotony of the total squared error depending on the number of clusters.

To overcome an exponential fall of computing speed along with cluster number increasing that is typical for multi-threshold Otsu method *K*–method provides a practical way to obtain a complete sequence of optimal image approximations. In the proposed solution, the clusters form the overlapping tuples, and clusters within every tuple give or take away the subsets of pixels from each other to reduce the total squared error until the obtaining of stable image partition.

As shown in the paper, the insufficiency of segment merging, which only one is used in Mumford–Shah model prevents effective minimization of the objective function. This model restriction is overcome in *K*–method owing to introducing of an additional operation of segment correction, and the weakening of the connectivity conditions for pixels within the segment so that in the extreme case of disconnected pixels the results of the segmentation according to Mumford–Shah model reproduce the results of the segmentation according to generalized version of the multi-threshold Otsu method.

In the paper the experimental substantiation of the proposed method is discussed referring to illustrative applications published in an accessible electronic resource arXiv.org.