

СПИСОК ЛИТЕРАТУРЫ

1. Дмитриев Ю.Г. Непараметрическое условное оценивание функционалов плотности распределения // Математическое моделирование и теория вероятностей. — Томск: Изд-во «Пеленг», 1998. — С. 169–177.
2. Скрипин С.В. Комбинированная непараметрическая оценка регрессии // Проблемы компьютерной безопасности и криптография: Докл. IV Сибирской научной школы-семинара с международным участием. — Sibecrypt'05 // Вестник Томского государственного университета. — 2005. — Прил. № 14. — С. 311–313.
3. Скрипин С.В. О свойствах комбинированной оценки регрессии // Проблемы компьютерной безопасности и криптография: Докл. VI Сибирской научной школы-семинара с международным участием. — Sibecrypt'07 // Вестник Томского государственного университета. — 2007. — Прил. № 23. — С. 326–329.
4. Эфрон Б. Нетрадиционные методы многомерного статистического анализа. — М.: Финансы и статистика, 1988. — 261 с.

Поступила 09.06.2008 г.

УДК 504.064(4)

ПРИМЕНЕНИЕ ПРОСТРАНСТВЕННЫХ ХАРАКТЕРИСТИК ПРИ МОДЕЛИРОВАНИИ ИЗМЕНЕНИЙ ЛАНДШАФТНОГО ПОКРОВА С ИСПОЛЬЗОВАНИЕМ КЛЕТОЧНЫХ АВТОМАТОВ

А.В. Замятин

Томский политехнический университет
E-mail: zamyatin@tpu.ru

Предложен подход к моделированию изменений ландшафтного покрова с использованием клеточных автоматов и 7-ми пространственных характеристик, наиболее часто встречающихся при решении различных задач геоинформатики, позволяющий повысить адекватность моделирования. Проведен поиск информативности каждой из характеристик в различных размерах скользящего окна. Предложена методика оценки значимости характеристик в различных сочетаниях при определении правил функционирования клеточных автоматов, а также выполнена соответствующая оценка значимости характеристик. Результаты экспериментов получены с использованием типовых разновременных данных растровой геоинформационной системы Idrisi Kilimanjaro.

Ключевые слова:

Моделирование, клеточные автоматы, пространственные характеристики, численные эксперименты, экосистемный мониторинг.

Введение

Моделирование изменений ландшафтного покрова является одним из наиболее перспективных подходов к прогнозированию таких явлений как распространение лесных пожаров, обезлесение, эрозия почвы, изменение границ снежного покрова [1, 2]. Решают задачи моделирования с использованием как детерминированных, так и стохастических моделей [3]. С учетом случайной природы моделируемых процессов ландшафтного покрова и возрастающими возможностями компьютерной техники наиболее перспективными следует считать стохастические модели, учитывающие пространственное взаимодействие элементов с помощью клеточных автоматов (КА).

Одним из ключевых факторов, влияющих на адекватность моделирования с использованием КА, является определение в каждом конкретном случае правил взаимодействия элементов (правил функционирования) КА. В настоящее время все большее распространение получают подходы к использованию пространственных характеристик при решении различных задач геоинформатики [4–6], а предварительные исследования в [7] показывают перспективность использования таких характеристик для определения правил функционирования КА.

При моделировании практическое применение пространственных характеристик, позволяющих более полно учитывать особенности ландшафта, обычно затруднено в связи с их различной информативностью в различных размерах анализируемой окрестности, а также неопределенностью вариантов возможного сочетания характеристик и их необходимого количества в каждом конкретном случае. Так, в [2, 4] использована лишь одна, а в [7] — четыре пространственные характеристики при решении задач моделирования изменений ландшафтного покрова.

Данная работа направлена на повышение эффективности моделирования изменений ландшафтного покрова с использованием увеличенного набора из 7-ми пространственных характеристик, наиболее часто встречающихся при решении различных задач геоинформатики, анализ этих характеристик в варьируемом размере анализируемой окрестности и определение их результативной сочетаемости при совместном применении.

Правила функционирования КА при моделировании

В [7] автором правила функционирования КА в алгоритме моделирования изменений ландшафтного покрова предлагается определять в каждой

точке исследуемой территории на основе результирующей вероятности замещения i -го на j -ый тип ландшафта:

$$p_{ij}^{\text{рез}} \sim f(p_{ij}^{\text{вер}}, p_{ij}^{\text{пр}}, p_{ij}^{\text{доп}}), \quad (1)$$

где $p_{ij}^{\text{вер}}$ — вероятностная характеристика, получаемая на основе стохастической матрицы, $p_{ij}^{\text{доп}}$ — вероятность, рассчитываемая на основе априорной информации из карт соответствия (карты буферных зон, расстояний до магистралей, источников питьевой воды и пр.). Вероятность $p_{ij}^{\text{пр}}$ в выражении (1) предлагается рассчитывать с использованием 7-ми различных характеристик, расчет которых производится для каждого i -го элемента изображения в скользящем окне порядка d с количеством элементов в нем $(2d+1) \times (2d+1)$:

$$C_{\text{ЧВ}}^k = \frac{\sum_{j=1}^m n_j^k}{n}, \quad (2)$$

$$C_{\Phi}^k = \frac{m^k}{m}, \quad (3)$$

$$C_{\text{СФИ}}^k = \frac{2 \sum_{j=1}^m \frac{\ln(0,25 \cdot p_j^k)}{\ln(n_j^k)}}{m} - 1, \quad (4)$$

$$C_{\text{СР}}^k = \frac{\sum_{j=1}^m h_j^k}{m^k \cdot 2d}, \quad (5)$$

$$C_{\text{СД}}^k = \left[1 - \sum_{j=1}^m \left(\frac{n_j^k}{2d+1} \right) \right], \quad (6)$$

$$C_{\text{СВУ}}^k = \frac{\max(n_j^k)}{A^k}, \quad (7)$$

$$C_{\text{СФВПП}}^k = \frac{\left[m_j^k \sum_{j=1}^m (\ln p_j^k \cdot \ln n_j^k) \right] - \left[\sum_{j=1}^m \ln p_j^k \cdot \sum_{j=1}^m \ln n_j^k \right]}{\left(m_j^k \sum_{j=1}^m \ln p_j^k \right)^2 - \left(\sum_{j=1}^m \ln p_j^k \right)^2}, \quad (8)$$

где n — общее число элементов, m — общее число участков всех типов, индекс k определяют тип ландшафта, индекс j — номер участка (под участком понимается набор элементов (пикселей) одного типа ландшафта, окруженных элементами (пикселями) типов ландшафта, отличными от рассматриваемого), n_j^k — число элементов в окрестности, p_j^k — периметр участка, h_j^k — число элементов соответствующего участка до ближайшего соседнего элемента участка k -го типа, m^k — число участков соответствующего типа, A^k — общая площадь участков соответствующего типа, а условные обозначения характеристик: $C_{\text{ЧВ}}$ — частота встречаемости, C_{Φ} — степень фрагментации, $C_{\text{СФИ}}$ — среднее фрактальное измерение периметров, $C_{\text{СР}}$ — среднее расстояние между участками, $C_{\text{СД}}$ — степень деления, $C_{\text{СВУ}}$ — степень величины участка, $C_{\text{СФВПП}}$ — средняя фрак-

тальная величина «периметров-площадей».

Возможны различные варианты использования пространственных характеристик в задачах моделирования. В данном случае вероятность $p^{\text{пр}}$ предлагается рассчитывать для каждой i -ой точки изображения на основе выражения

$$p^{\text{пр}} = |1 - d(\mathbf{F}_{\text{Слок}}, \mathbf{F}_{\text{Сгл}})|, \quad (9)$$

где $d(\mathbf{F}_{\text{Слок}}, \mathbf{F}_{\text{Сгл}})$ — Евклидово расстояние между вектором «локальных» характеристик и вектором «глобальных» характеристик. Сравнивая значения этих векторов попарно для каждого из типов ландшафта можно определить, насколько значения пространственных характеристик в текущем скользящем окне схожи со значением пространственных характеристик k -го типа ландшафта на исследуемой территории в целом. Это позволит принять решение о вероятности принадлежности i -го центрального элемента анализируемой окрестности к каждому из типов ландшафта, и выбрать максимальную вероятность, соответствующую k -му типу ландшафта.

Локальные характеристики для каждого k -го типа ландшафта рассчитываются с использованием выражений (2)–(8) в текущей анализируемой окрестности и могут быть представлены в виде вектора

$$\mathbf{F}_{\text{Слок}} = [C_{\text{ЧВ}}^k \ C_{\Phi}^k \ C_{\text{СФИ}}^k \ C_{\text{СР}}^k \ C_{\text{СД}}^k \ C_{\text{СВУ}}^k \ C_{\text{СФВПП}}^k]^T.$$

Глобальные характеристики могут быть представлены в виде вектора

$$\mathbf{F}_{\text{Сгл}} = [\bar{C}_{\text{ЧВ}}^k \ \bar{C}_{\Phi}^k \ \bar{C}_{\text{СФИ}}^k \ \bar{C}_{\text{СР}}^k \ \bar{C}_{\text{СД}}^k \ \bar{C}_{\text{СВУ}}^k \ \bar{C}_{\text{СФВПП}}^k]^T. \quad (10)$$

Компоненты вектора (10) рассчитываются для каждого из k типов ландшафта путем перемещения скользящего окна по всему изображению и подсчета среднего значения для каждой из характеристик (2)–(8).

Предложенный способ нахождения правил функционирования КА на основе (1) и (9) был положен в основу алгоритма моделирования изменений ландшафтного покрова, реализованного в программной системе анализа динамики ландшафтного покрова.

Результаты исследований

Для определения эффективности и пределов применимости алгоритма моделирования изменений ландшафтного покрова был проведен ряд экспериментов, направленных на поиск информативности характеристик в варьируемых размерах анализируемой окрестности и их значимости в различных сочетаниях. В качестве исходных данных были использованы разновременные тематические карты, поставляемые в качестве примерных данных с растровой геоинформационной системой *Idrisi Kilimanjaro*, — файлы *Landuse71.rst*, *Landuse85.rst* и *Landuse91.rst* [8]. Наличие такого набора данных позволило выполнять расчет параметров моделирования на базе карт 1971 г. и 1985 г., а карта 1991 г.

выступила эталоном. В качестве критерия оценки точности моделирования использованы широко распространенные *Receiver/Relative Operating Characteristic* (ROC) и *Kappa Index of Agreement* (KIA) [9, 10].

Для определения информативности характеристик при различных размерах анализируемой окрестности (скользящего окна) был проведен ряд экспериментов, в которых размер окна был принят 3×3 , 5×5 , 7×7 , 9×9 , 15×15 и 29×29 элементов, а моделирование выполнялось как без использования характеристик, так и с использованием каждой из характеристик (2)–(8). Такие эксперименты позволили не только подтвердить целесообразность использования пространственных характеристик (оценка точности результатов моделирования без характеристик ниже, чем оценка с любой из характеристик), но и выявить «чувствительность» каждой отдельной характеристики к величине анализируемой окрестности (рис. 1). Анализ полученных результатов исследований показывает, что практически во всех случаях для критериев точности KIA и ROC размер окна 5×5 элементов позволяет добиться наилучшего результата моделирования, поэтому именно такой размер окна следует признать оптимальным.

Исключение составили результаты моделирования при использовании характеристики C_{CBV} (наиболее низкие результаты в сравнении с другими характеристиками для окна любого размера), а также характеристики C_{BVU} , которые не показали устойчивой зависимости от размера окна (результаты дополнительных исследований C_{BVU} – на рис. 2).

Дополнительные исследования эффективности применения C_{BVU} при моделировании, выполненные с меньшим шагом дискретизации размера окна, подтвердили существенную «чувствительность» этой характеристики к масштабу исследуемой ландшафтно-классовой структуры. В данном случае более высокие и стабильные результаты для ROC и KIA получены при размере окна небольших размеров – 5×5 , 9×9 , а также больших размеров – 21×21 , 23×23 и 29×29 элементов.

Кроме численного способа оценки информативности каждой из пространственных характеристик был использован и визуальный способ оценки. Для этого результаты расчета каждой характеристики, полученные с использованием изображения *Idrisi Kilimanjaro* (файл *Landuse85.rst*), были визуализированы в виде вероятностных изображений, каждый пиксель которых представлял собой результат расчета соответствующей характеристики в окне 5×5 элементов (рис. 3). Визуальный анализ изображений показывает, что каждая из характеристик в различной степени позволяет выявить некоторые уникальные пространственные особенности анализируемого изображения и подчеркнуть границы между типами ландшафта.

Очевидно, что характеристики C_{CF} и C_{CD} в сравнении с другими характеристиками выявляют су-

щественно меньшее количество пространственных особенностей, в связи с тем, что соответствующие изображения (ε и δ на рис. 3) содержат минимальное количество информации. Следует отметить, что даже такая минимальная информация позволяет получать удовлетворительные результаты моделирования при использовании характеристик (5) и (6) не только в сравнении с результатами, полученными без использования характеристик, но и с результатами, полученными при использовании прочих характеристик (рис. 1), что подтверждает перспективность использования самых различных пространственных характеристик в задачах моделирования изменений ландшафтного покрова.

Каждая из характеристик (2)–(8) позволяет выявить «собственные» уникальные особенности анализируемого ландшафта, поэтому их практическое использование следует осуществлять совместно, учитывая одновременно группу различных пространственных особенностей в выбранной анализируемой окрестности скользящего окна. Поиск наиболее значимого набора пространственных характеристик автором предлагается выполнить в соответствии со следующей методикой.

Учитывая необходимость оценки 7-ми характеристик, общее количество экспериментов, каждый из которых выполнен с использованием их уникального сочетания, составило $N_{exp} = 2^7 = 128$. Результат каждого эксперимента – прогнозная карта, сравнивая которую с имеющимся эталоном получаем оценку значения критерия точности KIA и ROC. После проведения экспериментов, их результаты следует упорядочить по убыванию значений одного из критериев точности (в данном случае произвольно выбран критерий точности KIA, рис. 4, а).

Значимость некоторой характеристики C в отдельном k -ом эксперименте предлагается оценить как $S_C^k = v_C^k \cdot \frac{acc^k}{M}$, где $k=1, 2, \dots, N_{exp}$, acc – значение критерия точности (в данном случае – значение KIA), M – общее количество характеристик, используемое при моделировании (в данном случае $M=7$), $v=\{0, 1\}$: $v=0$, если характеристика C отсутствует в данном сочетании характеристик и $v=1$ – если присутствует. Оценку значимости характеристики C по результатам N экспериментов (интегральную оценку) выполним как

$$S_C^N = \sum_{k=1}^N v_C^k \cdot \frac{acc^k}{M} / N. \quad (11)$$

На рис. 4, б, приведены результаты оценки значимости каждой из характеристик, полученные с использованием (11) при $N=1, 2, \dots, N_{exp}$.

Анализ приведенных на рис. 4 зависимостей показывает, что наиболее точные результаты моделирования при использовании сочетаний характеристик получены с использованием C_{BV} , C_{CF} , C_{BVU} и C_{CFVPP} , а это демонстрирует более высокую значимость этих характеристик в проведенных экспериментах. Меньшая значимость характеристики C_{CFI} ,

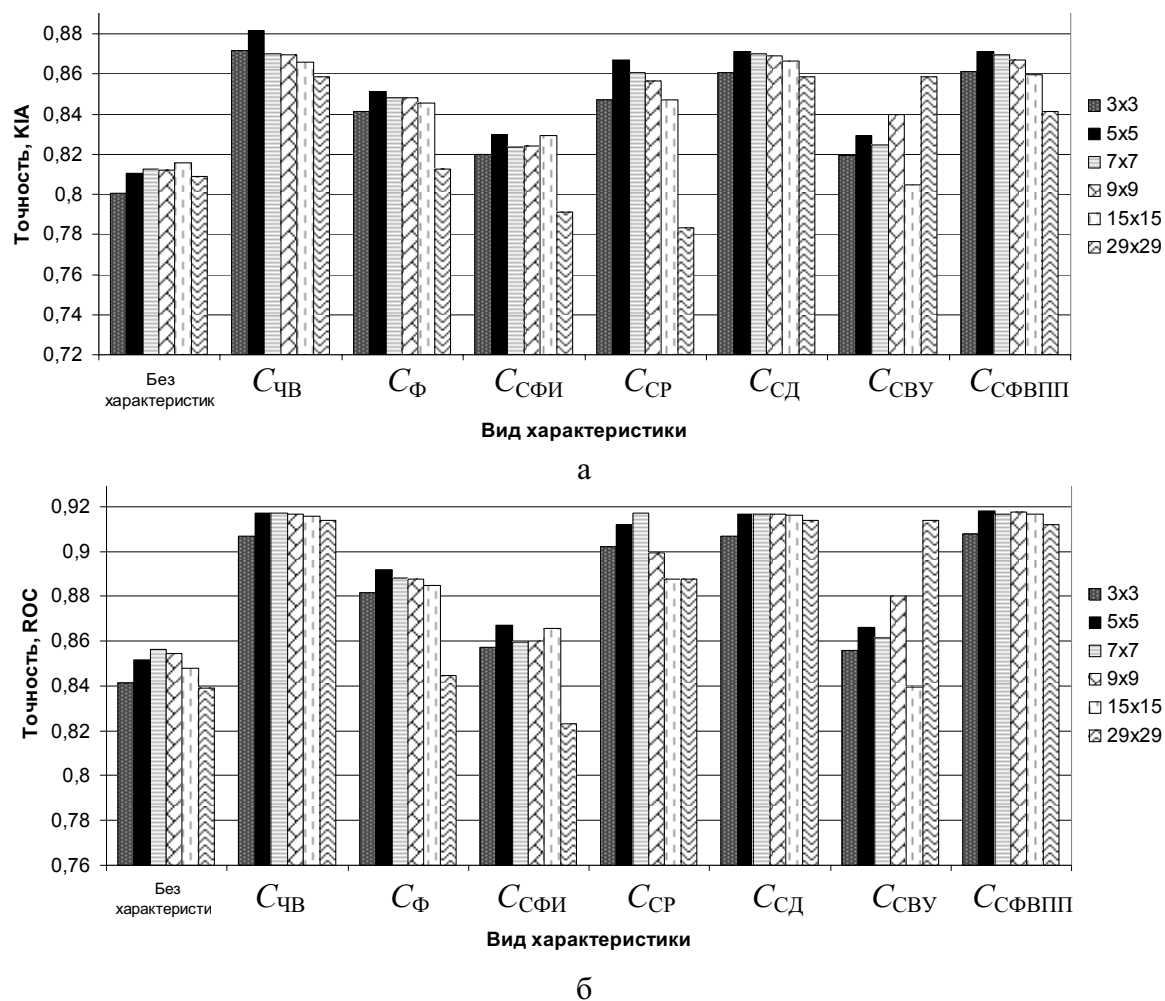


Рис. 1. Зависимость точности моделирования от вида используемой пространственной характеристики в скользящем окне различного размера. Критерий оценки: а) KIA; б) ROC

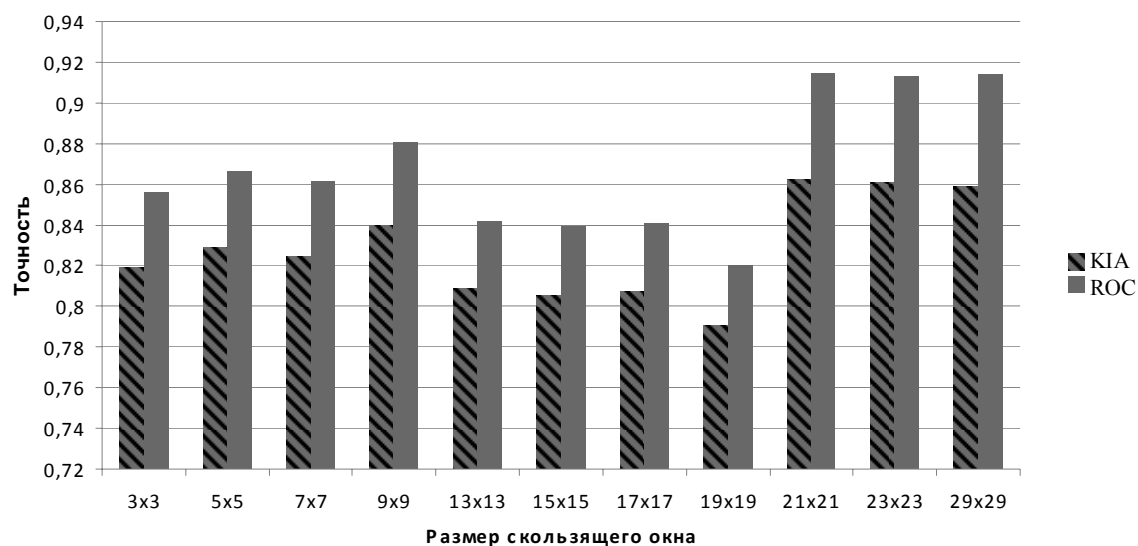


Рис. 2. Зависимость точности моделирования от размера скользящего окна при использовании характеристики $C_{сву}$

а также характеристик $C_{ср}$ и $C_{сд}$ подтверждается экспериментами, фрагменты результатов которых представлены выше, – на рис. 1 и на рис. 3, з, и д, соответственно.

Выводы

1. Результаты проведенных исследований подтвердили предварительные выводы, получен-

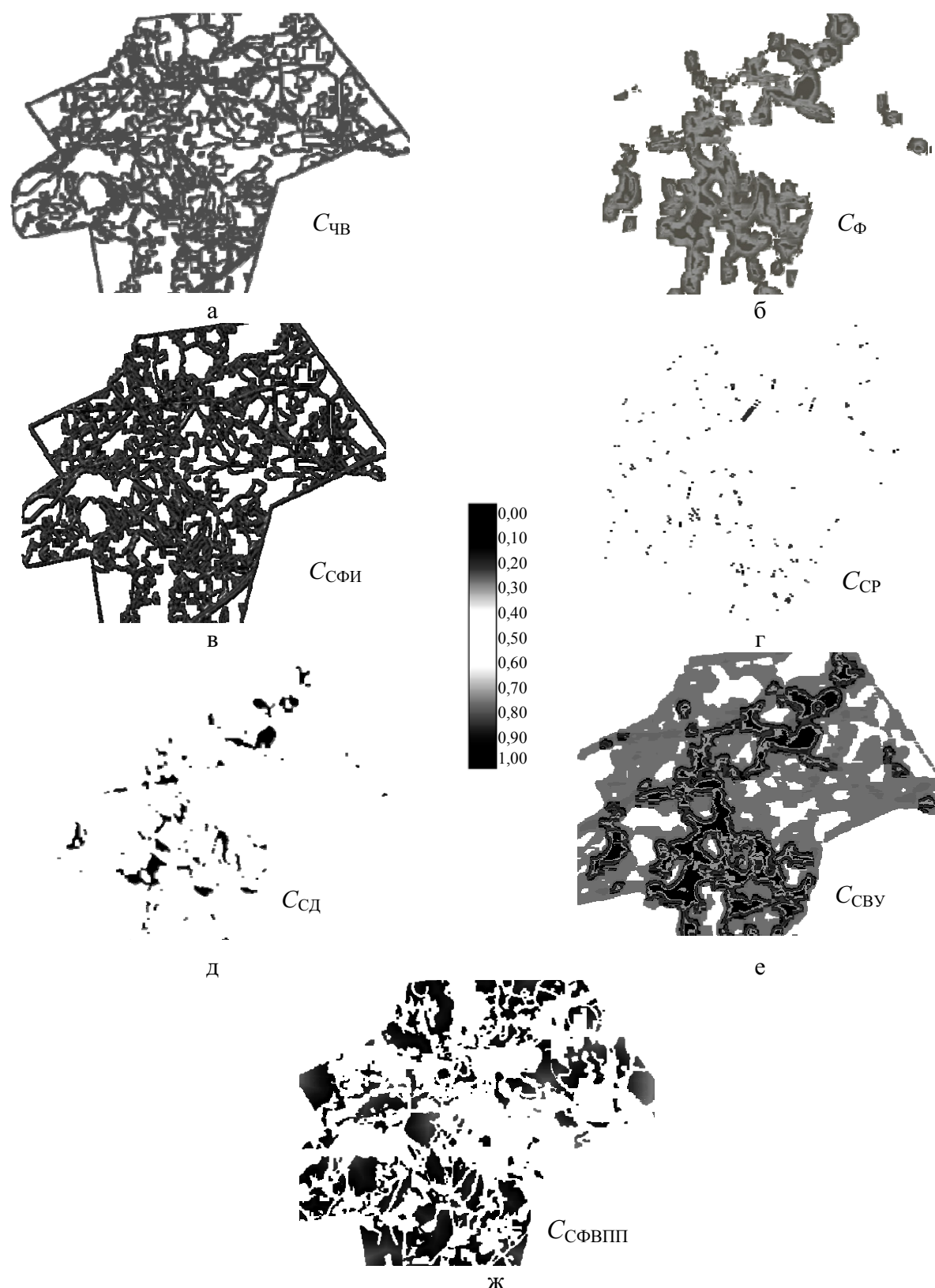


Рис. 3. Пространственные характеристики при $d=2$

ные автором ранее, и показали целесообразность использования пространственных характеристик для определения правил функционирования

клеточных автоматов с целью повышения точности результатов моделирования изменений ландшафтного покрова.

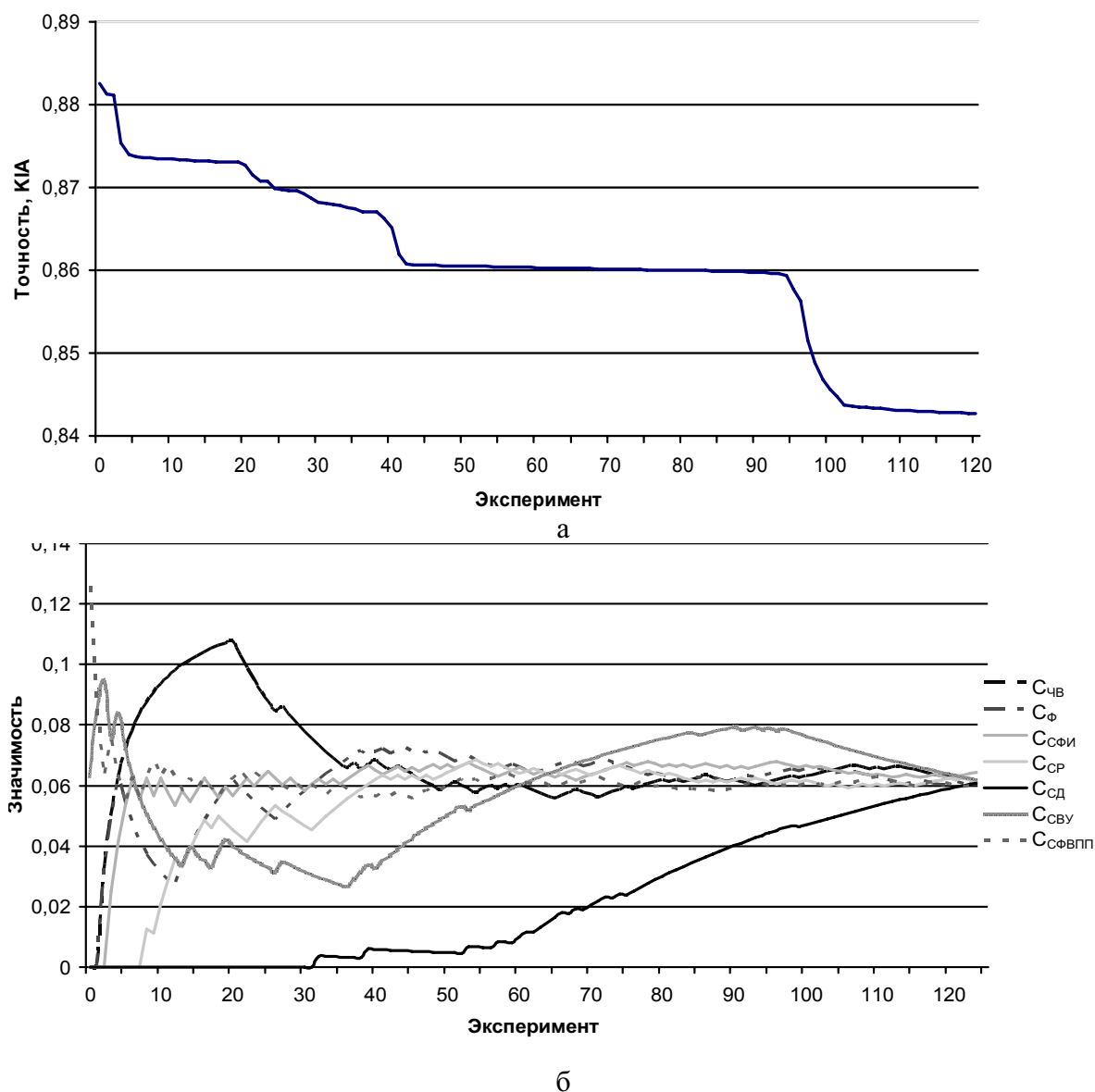


Рис. 4. Результаты экспериментов при различном сочетании пространственных характеристик, ранжированные по убыванию критерия точности KIA: а) значения критерия точности KIA в экспериментах; б) значимость каждой характеристики

2. С использованием типовых разновременных данных геоинформационной системы *Idrisi Kili-tanjaro* проведена оценка информативности 7-ми пространственных характеристик в анализируемой окрестности различного размера, по результатам которой показано, что расчет правил функционирования клеточных автоматов с использованием характеристик при решении практических задач моделирования изменений ландшафтного покрова наиболее целесообразно осуществлять в окне размером 5×5 элементов.
3. Предложена методика оценки значимости пространственных характеристик в различ-

ных сочетаниях при решении задачи моделирования с использованием клеточных автоматов. С использованием методики пространственные характеристики частота встречаемости, степень фрагментации, степень величины участка и средняя фрактальная величина «периметров-площадей» определены как наиболее значимые для обеспечения высоких показателей моделирования по критерию точности.

Автор выражает благодарность д.т.н., профессору Н.Г. Маркову за ценные замечания и советы при подготовке публикации.

СПИСОК ЛИТЕРАТУРЫ

1. Замятин А.В., Марков Н.Г. Анализ динамики земной поверхности с использованием данных дистанционного зондирования Земли. — М.: Физматлит, 2007. — 176 с.
2. Замятин А.В. Анализ динамики ландшафтного покрова на основе данных дистанционного зондирования Земли // Исследование Земли из космоса. — 2006. — № 6. — С. 50–64.
3. Тикунов В.С. Моделирование в картографии. — М.: Изд-во МГУ, 1997. — 405 с.
4. McGarigal K. Landscape pattern metric. [Электронный ресурс]. — Режим доступа: — www.umass.edu/landeco/pubs/Fragmetrics_short.pdf, свободный.
5. Verburg P.H. et al. A method to analyse neighborhood characteristics of land use patterns // Computers, Environment and Urban Systems, 2003. — № 24. — P. 354–369.
6. Wu J.G. Can landscape indices predict ecological processes consistently? // Landscape Ecology. — 2000. — V. 16. — № 3. — P. 235–254.
7. Замятин А.В., Михайлов П.В., Cabral P. Современные средства для решения задач анализа динамики и прогнозирования изменений ландшафтного покрова // Известия Томского политехнического университета. — 2006. — Т. 309. — № 7. — С. 80–86.
8. Clark Labs — IDRISI GIS and Image Processing Software. [Электронный ресурс]. — Режим доступа: — <http://www.clarklabs.org/>, свободный.
9. Westin L.K. Department of computer Science Umeå University. Receiver operating characteristic analysis. [Электронный ресурс]. — Режим доступа: — <http://www.cs.umu.se/research/reports/2001/018/part1.pdf>, свободный.
10. Richards J.A., Xiuping Jia. Remote Sensing Digital Image Analysis: An Introduction. — Berlin: Springer, 1999. — 400 p.

Поступила 24.10.2008 г.

УДК 004.627

СЖАТИЕ МНОГОЗОНАЛЬНЫХ АЭРОКОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ И УЧЕТОМ МЕЖДИАПАЗОННОЙ ЗАВИСИМОСТИ

А.В. Замятин, То Динь Чыонг

Томский политехнический университет

E-mail: zamyatin@tpu.ru

Предложен трехэтапный алгоритм сжатия многозональных аэрокосмических изображений, основанный на использовании вейвлет-преобразования и учете междиапазонной зависимости, позволяющий в большей мере учесть специфику и повысить степень сжатия данных дистанционного зондирования Земли. Проведены сравнительные исследования эффективности предложенного алгоритма и универсальных алгоритмов сжатия, подтвердившие его работоспособность и позволившие определить оптимальную глубину вейвлет-преобразования. Результаты исследований показали превосходство предложенного алгоритма в различной степени над аналогами в сжатии при более существенных вычислительных затратах.

Ключевые слова:

Аэрокосмические изображения, вейвлет-преобразование, сжатие изображений, междиапазонная корреляция.

Введение

В связи с постоянно растущими техническими возможностями систем дистанционного зондирования Земли (ДЗЗ), их более широким использованием потребителями при решении различных задач, объем данных, которыми оперируют эти системы, исчисляется терабайтами и продолжает неуклонно увеличиваться. Поэтому решение задачи сжатия таких данных с использованием различных подходов и программно-аппаратных средств с целью существенного повышения эффективности обработки, хранения, и передачи аэрокосмической информации по каналам связи становится все более актуальным как в России, так и за рубежом [1–5]. Учитывая необходимость предварительной обработки и автоматизированной классификации таких данных, наибольшую ценность представляют алгоритмы сжатия без потерь, не допускающие искажения статистических яркостных характери-

стик восстановленного аэрокосмического изображения (АИ).

Выделяют два принципиально различных подхода к сжатию данных ДЗЗ. Один подход предполагает использование универсальных и широко известных алгоритмов сжатия, не учитывающих их специфику, и представленных в универсальных архиваторах *WinRar* или *WinZip* [3]. Другой подход предполагает разработку новых алгоритмов сжатия, учитывающих при обработке не только данные каналов (диапазонов) как совокупность обычных черно-белых изображений или как неструктурированный массив информации, но и существующую зависимость (корреляцию) между диапазонами АИ. Несмотря на то, что такой подход является более сложным с вычислительной точки зрения, именно при его использовании можно добиться существенно более высоких показателей степени сжатия за счет учета специфики данных ДЗЗ.