



## Повышение эффективности процесса интерпретации данных дистанционного зондирования Земли за счет анализа дескрипторов окрестности

С. А. Ямашкин\*, А. А. Ямашкин  
ФГБОУ ВО «МГУ им. Н. П. Огарёва» (г. Саранск, Россия)

\*yamashkinsa@mail.ru

*Введение.* В оценке пространственно-временной структуры земной поверхности большую значимость приобретают данные дистанционного зондирования Земли. Повышение эффективности инструментов анализа космической съемки возможно через изучение проблемы получения комплексной пространственно-временной характеристики состояния земель. Целью исследования является повышение точности автоматизированного анализа данных дистанционного зондирования Земли за счет учета инвариантных и динамических дескрипторов окрестности.

*Материалы и методы.* С целью повышения точности классификации данных дистанционного зондирования Земли проводился расчет комплексных пространственно-временных характеристик состояния земель на основе системного анализа данных, характеризующих динамические и инвариантные состояния территории, окружающей геофизический участок. Формализация данного процесса, приведенная в статье, включает методики расчета набора численных дескрипторов окрестности: локальной энтропии, локального диапазона, среднеквадратического отклонения, цветового момента, гистограммы оттенков, кортежа цветов. Описана методика расчета комплексного дескриптора на основе вектора Фишера. Для апробации решения составлен план эксперимента и проведена выборка исходных данных.

*Результаты исследования.* Апробация методики и созданного на ее основе алгоритма, реализованного в виде комплекса программ, на системе тестовых полигонов показала варьирование точности классификации в диапазоне 81–89 % (без учета дескрипторов окрестности) и 91–97 % (с учетом дескрипторов). Значительное увеличение радиуса анализируемой окрестности приводит к снижению точности классификации.

*Обсуждение и заключения.* Разработанный комплекс программ позволяет оперативно проводить моделирование пространственных систем с целью тематического картографирования землепользования и анализа развития чрезвычайных ситуаций, а созданная методика анализа земель с учетом дескрипторов окрестности дает возможность повысить точность классификации.

**Ключевые слова:** дешифрирование космических снимков, дистанционное зондирование Земли, анализ земель, дескриптор окрестности, инвариантные свойства, динамические свойства

**Для цитирования:** Ямашкин С. А., Ямашкин А. А. Повышение эффективности процесса интерпретации данных дистанционного зондирования Земли за счет анализа дескрипторов окрестности // Вестник Мордовского университета. 2018. Т. 28, № 3. С. 352–365. DOI: <https://doi.org/10.15507/0236-2910.028.201803.352-365>

© Ямашкин С. А., Ямашкин А. А., 2018



This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted reuse, distribution, and reproduction in any medium provided the original work is properly cited.



## Improving the Efficiency of Remote Sensing Data Interpretation by Neighborhood Descriptors Analyzing

A. A. Yamashkin, S. A. Yamashkin\*

*National Research Mordovia State University, Saransk, Russia*

\*yamashkinsa@mail.ru

*Introduction.* In the evaluation of the space-time structure of the Earth's surface, the data of remote sensing of the Earth (RSE data) become more important. Increasing the effectiveness of space survey analysis tools is possible through studying the problem of obtaining an integrated space-time characterization of the status of lands. The purpose of this study is to improve the accuracy of the automated analysis of remote sensing data by taking into account the invariant and dynamic descriptors of the vicinity.

*Materials and methods.* In order to improve the accuracy of the RSE data classification, a computation of complex space-time characteristics on the state of the earths was conducted, based on a system analysis of data characterizing the dynamic and invariant states of the territory surrounding the geophysical site. The formalization of this process includes methods for calculating a set of numerical descriptors of the neighborhood: local entropy, local range, standard deviation, color moment, histogram of hues, and color cortege. A technique for calculating a complex descriptor based on the Fisher vector is described. To approbate the solution, a plan for the experiment was drawn up and a sample of the initial data was sampled.

*Results.* Approbation of the methodology and the algorithm developed on its basis, implemented as a set of programs, on the test polygon system showed a variation in the classification accuracy in the range of 81–89% (without regard to the neighborhood), and taking into account the neighborhood, it increases to 91–97%. It is revealed that a significant increase in the radius of the analyzed neighborhood leads to a decrease in the classification accuracy.

*Conclusions.* The application of the developed set of programs allows the rapid implementation of modeling of spatial systems for the purpose of thematic mapping of land use and analysis of the development of emergency situations. The developed methodology for analysis of land with regard to the descriptors of the neighborhood makes it possible to improve the accuracy of classification.

**Keywords:** interpretation of space images, remote sensing, land analysis, neighborhood descriptors, invariant property, dynamic property.

**For citation:** Yamashkin A. A., Yamashkin S. A. Improving the Efficiency of Remote Sensing Data Interpretation by Neighborhood Descriptors Analyzing. *Vestnik Mordovskogo universiteta* = Mordovia University Bulletin. 2018; 28(3):352–365. DOI: <https://doi.org/10.15507/0236-2910.028.201803.352-365>

### Введение

Эффективное производственное и экономическое планирование, основанное на релевантной информации о качественных и количественных свойствах природных и антропогенных систем, способствует минимизации издержек хозяйственной деятельности от различных факторов. Выделение на картах и космических снимках относительно однородных по географическим, физическим и химическим свойствам участков земной по-

верхности позволяет не только оценить состояние территории, но и предсказать ее устойчивость к различным нагрузкам, а также спрогнозировать развитие стихийных природных и техногенных ситуаций. В этом вопросе ключевую значимость имеют получаемые со спутников данные дистанционного зондирования Земли (ДЗЗ) и их автоматизированный анализ с применением методов и алгоритмов классификации. В настоящее время существует большое количество методов анализа данных ДЗЗ, однако

вопрос повышения их точности, быстродействия и эффективности по-прежнему актуален ввиду специфики решаемой задачи. Главным направлением в решении данного вопроса является критическое исследование проблемы получения комплексных пространственно-временных характеристик состояния земель и развития стихийных процессов. При этом следует опираться на системный анализ массивов данных, которые в полной мере характеризуют инвариантные и динамические состояния территории, окружающей геофизический участок. Такая необходимость объясняется тем, что анализ свойств окрестности так же важен для определения класса территории, как и спектральные характеристики.

Целью исследования является повышение эффективности анализа материалов космической съемки для решения задачи мониторинга состояния земель и прогнозирования развития стихийных процессов с помощью расчета динамических и инвариантных дескрипторов окрестности. Для достижения поставленной цели в рамках исследования, представленного в статье, решены следующие задачи:

- проведен анализ предметной области и обобщен отечественный и мировой опыт исследования структуры и состояния земель на базе материалов космической съемки;

- создана новая методика анализа земель с применением синтетических дескрипторов окрестности;

- предложенный математический аппарат апробирован на тестовых полигонах с помощью созданного для проведения эксперимента прототипа.

По результатам исследований создан программный комплекс анализа космической съемки, который является инструментом мониторинга состояния земель и прогнозирования стихийных процессов.

### Обзор литературы

Инструментальное дешифрирование мульти- и гиперспектральных данных ДЗЗ в настоящее время основывается на математических и интеллектуальных алгоритмах<sup>1</sup> и методиках анализа сигналов данных (многомерных изображений)<sup>2</sup>. Для классификации материалов космической съемки применяются методы машинного обучения<sup>3</sup> [1] и цифровой обработки сигналов данных<sup>4-5</sup>. Детектирование характеристик, типов и свойств объектов осуществляется на базе выявления системных связей их структурных и спектральных свойств<sup>6-7</sup>, нейронных сетей<sup>8</sup> (самообучающихся [2] и обучающихся с учителем<sup>9</sup> [3]), статистической классификации [4], машин опорных векторов (Support Vector Machines)<sup>10</sup> [5]. Актуальность решения задачи повышения качественных и количественных харак-

<sup>1</sup> Хайкин С. [Нейронные сети](#): полный курс : пер. с англ. М. : Вильямс, 2006. 1104 с.

<sup>2</sup> Шовенгердт Р. А. [Дистанционное зондирование: модели и методы обработки изображений](#). М. : Техносфера, 2010. 560 с.

<sup>3</sup> Duda R. O., Hart P. E., Stork D. G. [Pattern classification](#). – 2<sup>nd</sup> ed. New York : Wiley, 2001. 738 p.

<sup>4</sup> Landgrebe D. A. [Signal theory methods in multispectral remote sensing](#). New York : Wiley, 2003. 508 p.

<sup>5</sup> Woodcock C. E., Strahler A. H., Smith J. A. [On the nature of models in remote sensing](#) // Remote sensing of environment. 1986. Vol. 20, no. 2. P. 121–139.

<sup>6</sup> Апанасевич П. А. [Основы теории взаимодействия света с веществом](#). Минск : Наука и техника, 1977. 496 с.

<sup>7</sup> Степанов Б. И. [Введение в современную оптику : фотометрия : о возможном и невозможном в оптике](#). Минск : Наука и техника, 1989. 254 с.

<sup>8</sup> Галушкин А. И. [Нейронные сети: основы теории](#). М. : Горячая линия-Телеком, 2012. 496 с.

<sup>9</sup> Rumelhart D. E., Hinton G. E., McClelland J. L. [A general framework for parallel distributed processing : parallel distributed processing](#). 1986. 235 p.

<sup>10</sup> Vapnik V. N. [The nature of statistical learning theory](#). Springer, 1995. 508 p.



теристик распознавания и дешифрирования материалов космической съемки для решения научных и практических задач в сфере оптимизации землепользования определяет актуальность поиска новых методов получения информации, а также повышения эффективности существующих. Результаты анализа отечественного и зарубежного опыта в области распознавания образов свидетельствуют о перспективности исследований и экспериментов в области концепции обучения ансамблей (Ensemble Learning, гибридного обучения) [6–8] в сочетании с расчетом контекстной синтетической информации (различных текстовых дескрипторов территорий).

Анализ земель на основе космической съемки Земли ориентирован на детектирование системных связей, которые существуют между физическими, химическими и биологическими свойствами территории и полем ее излучения в различных диапазонах (поверхности по-разному отражают и излучают электромагнитные волны). При этом зависимость спектральной яркости от длины волны представляет собой важнейшую характеристику свойств исследуемой территории. Системный анализ данной зависимости может предоставить информацию не только о типе территории (растительность, водные объекты, почвы, материнская порода), но и о ее свойствах (биологическом разнообразии, плодородии, чистоте водоемов).

Доступные на свободной и коммерческой основе мульти- и гиперспектральные космические снимки дают возможность проводить результативные исследования оптических свойств земель в разных спектральных диапазонах (как в видимых зонах спектра, так и в ультрафиолетовых и инфракрасных). При выборе данных ДЗЗ для анализа и проведения экспериментов необходимо основываться на специфике поставленной цели и задачах,

решаемых для ее достижения. Анализ рынка космических снимков позволил выбрать для экспериментов материалы, получаемые со спутников Landsat 7-го и 8-го поколений. Их пространственное разрешение позволяет исследовать отдельные лесные кварталы, сельскохозяйственные земли, районы населенных пунктов и учитывать геопространственные композиции земель различных категорий. Спектральное разрешение данных, получаемых с этих спутников, дает информацию о свойствах геофизических оболочек в информативных видимом, ближнем и коротковолновом ИК-диапазонах. При достаточной периодичности съемки эти данные свободно распространяются через интернет-портал Геологической исследовательской службы США (USGS).

Для анализа ДЗЗ целесообразно использовать сложные программные комплексы вычислительных расчетов и моделирования. Кроме предоставления возможности эффективной работы с матрицами большого объема (Big Data Processing), программный инструментарий для исследования должен включать средства статистического и интеллектуального анализа данных, обработки изображений (сигналов данных), оценки точности, ошибочности и эффективности экспериментальных исследований. Актуальный характер имеют качественные нефункциональные требования, такие как быстроедействие, надежность, отказоустойчивость, хорошая задокументированность. По этим характеристикам был выбран MATLAB – система прикладных программ для различных расчетов, сгруппированных по тематическим разделам: Neural Network, Image Processing, Statistics Toolboxes.

### **Материалы и методы**

Для получения наиболее информативных пространственных и временных характеристик состояния пространственных объектов и земель следует опи-

ратся не только на спектральные параметры территории, как это принято в мировой практике, но и на системный анализ данных, характеризующих инвариантные и динамические свойства территории, которая окружает геофизический объект (окрестность). Под окрестностью пространственного объекта в данном исследовании мы понимаем территорию, которая расположена в пределах установленного расстояния от него<sup>11-12</sup>.

Выявление инвариантных свойств земель возможно благодаря исследованию (с помощью морфометрических карт рельефа) таких характеристик земель, изменение которых происходит в течение длительного времени и является при этом необратимым<sup>13-14</sup> [9–11]. Сведения о динамических особенностях земель можно извлечь из данных космической съемки благодаря автоматизированному анализу спектральных свойств территорий [12–14]. При решении задачи автоматизированного детектирования типов и свойств земель следует учитывать параметры (дескрипторы) окрестности в совокупности со спектральными свойствами. Чтобы получить ответ на вопрос, какое влияние анализ свойств окружения оказывает на результат дешифрирования, необходимо провести эксперимент. Для этого нужно формализовать созданную методику анализа земель и реализовать ее в виде комплекса программ. Приведем описание ключевых территориальных дескрипторов в математическом виде<sup>15</sup>.

*Локальная энтропия окрестности.* Данный дескриптор характеризует геофизическую неоднородность территории и является идентифицирующей метрикой, определяющей тип геофизического комплекса. Расчет локальной энтропии окрестности в рамках реализованного алгоритма будет производиться по формуле

$$E = \sum_{i=1}^N \frac{n_i}{S} \log\left(\frac{n_i}{S}\right),$$

где  $n_i$  – количество пикселей, входящих в окрестность объекта и обладающих определенной спектральной яркостью  $i$ ;  $N$  – радиометрическое разрешение снимка;  $S$  – площадь окрестности в пикселях.

*Локальный диапазон окрестности.* Величина данного дескриптора одновременно проста в вычислении и информативна, т. к. позволяет определить сосредоточенность краевых и граничных элементов в границах исследуемой территории:

$$R = \max(X) - \min(X),$$

где  $X$  – совокупность значений спектральных яркостей пикселей окрестности.

*Среднеквадратическое отклонение.* Данный дескриптор характеризует величину разброса значений спектральных яркостей. Географические объекты разного типа имеют уникальную пространственную структуру, которая однозначно определяется

<sup>11</sup> Ямашкин С. А. [Методическое и алгоритмическое обеспечение процесса анализа структуры земель на базе данных дистанционного зондирования](#) : дис. ... канд. техн. наук. Пенза, 2016.

<sup>12</sup> Ямашкин С. А. [Методическое и алгоритмическое обеспечение процесса анализа структуры земель на базе данных дистанционного зондирования](#) : автореф. дис. ... канд. техн. наук. Пенза, 2016.

<sup>13</sup> Получение, хранение и распространение геоданных как единый информационный процесс / С. М. Вдовин [и др.] // Природные опасности: связь науки и практики : мат-лы II Междунар. науч.-практ. конф. Саранск, 2015. С. 82–90.

<sup>14</sup> Ямашкин С. А., Занозин В. В. Геоинформационные ресурсы как системы поддержки принятия управленческих решений // Современные проблемы географии : межвуз. сб. науч. ст. 2017. С. 31–37.

<sup>15</sup> GIS-diagnosis of relief forms based on remote sensing data / S. A. Yamashkin [et al.] // International Scientific Conference 150<sup>th</sup> Anniversary of Jovan Cvijić's Birth. Belgrade, 2015. P. 45.



отличными друг от друга величинами среднеквадратического отклонения. Данный параметр вычисляется по следующей формуле:

$$S = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$$

где  $\bar{x}$  – среднее арифметическое спектральной яркости пространственного окружения объекта.

Определение цветовых свойств отдельных пикселей космического снимка также является важной составляющей анализа земель. Целесообразно использовать признаки, инвариантные к нежелательным изменениям цвета, таким как засвечивание или наложение тени.

*Цветовой момент* – дескриптор, характеризующий вероятностное распределение цветов окрестности; определяется кортежем среднего арифметического, дисперсии и коэффициента асимметрии. В совокупности они дают устойчивую и аутентичную цветовую характеристику, которая сопоставляет распределение цвета с цветовым пространством:

$$M = [\bar{I}, D, A] = \left[ \frac{1}{N} \sum_{j=1}^N I(j), \sqrt{\frac{1}{N} \sum_{j=1}^N (I(j) - \bar{I})^2}, \sqrt[3]{\frac{1}{N} \sum_{j=1}^N (I(j) - \bar{I})^3} \right]$$

где  $M$  – цветовой момент;  $N$  – число пикселей окрестности;  $\bar{I}$  – среднее арифметическое яркостей пикселей снимка;  $D$  – дисперсия пиксельных яркостей снимка;  $A$  – коэффициент асимметрии пиксельной яркости снимка.

*Гистограмма оттенков*, рассчитываемых для всех окрестных пикселей, является информативным дескриптором геофизического объекта. Значение оттенка вычисляется исходя из величин яркости объекта в трех анализируемых спектральных диапазонах:

$$H = \begin{cases} 1) \text{ if } (\max(c_1, c_2, c_3) = \min(c_1, c_2, c_3)) \Rightarrow \\ \Rightarrow 0; \\ 2) \text{ if } (\max(c_1, c_2, c_3) = c_1 \text{ и } c_2 \geq c_3) \Rightarrow \\ \Rightarrow \frac{c_2 - c_3}{6(\max(c_1, c_2, c_3) - \min(c_1, c_2, c_3))}; \\ 3) \text{ if } (\max(c_1, c_2, c_3) = c_1 \text{ и } c_2 < c_3) \Rightarrow \\ \Rightarrow \frac{c_2 - c_3}{6(\max(c_1, c_2, c_3) - \min(c_1, c_2, c_3))} + 1; \\ 4) \text{ if } (\max(c_1, c_2, c_3) = c_2) \Rightarrow \\ \Rightarrow \frac{c_3 - c_1}{6(\max(c_1, c_2, c_3) - \min(c_1, c_2, c_3))} + \frac{1}{3}; \\ 5) \text{ if } (\max(c_1, c_2, c_3) = c_3) \Rightarrow \\ \Rightarrow \frac{c_1 - c_2}{6(\max(c_1, c_2, c_3) - \min(c_1, c_2, c_3))} + \frac{2}{3}. \end{cases}$$

где  $H$  – оттенок;  $c_i$  – спектральная яркость в  $i$ -м диапазоне.

*Кортеж цветов* – набор значений, которые описывают вероятность появления в окрестности объекта пикселей с определенной цветовой меткой; на основе отображения цветового пространства соотносятся в численный цветовой атрибут. Поскольку объекты одного типа могут характеризоваться различными оттенками, но одной и той же цветовой меткой, данный параметр характеризуется информативной фотометрической инвариантностью. Кортеж цветов заполняется следующим образом:

$$CC = [p(C_1), p(C_2), \dots, p(C_k)] = \left[ \frac{N_1}{N}, \frac{N_2}{N}, \dots, \frac{N_k}{N} \right],$$

где  $CC$  – кортеж (вектор) цветов;  $p(C_i)$  – вероятность нахождения пикселя  $i$ -го цвета в пределах окрестности;  $N_i$  – количество пикселей  $i$ -го цвета в окрестности;  $N$  – общее число пикселей в окрестности.

Комплексный учет свойств окрестности при классификации земель целесообразно осуществлять путем расчета вектора Фишера, который представляет собой частный случай ядра Фишера<sup>16-17</sup>, предназначенного для кодирования локальных особенностей сигнала в формате, подходящем для проведения эффективного и точного обучения. Вектор Фишера – это такое представление изображения, которое получено путем объединения локальных дескрипторов изображения. Именно поэтому его целесообразно использовать в качестве комплексного дескриптора изображения при классификации земель. Изображение можно представить в виде одномерного массива  $D$ -мерных векторов дескрипторов, автоматически рассчитываемых из изображения или сигнала:

$$I = [x_1, \dots, x_N],$$

где  $I$  – изображение (сигнал);  $x_i$  –  $i$ -й дескриптор.

При  $G = [a_j, \mu_j, \sigma_j; j = \overline{1, J}]$  – параметрах смеси гауссовых распределений ( $a$  априорная вероятность,  $\mu$  математическое ожидание,  $\sigma$  среднеквадратическое отклонение), сформированной на базе распределения дескрипторов, каждый вектор  $x_i$  сопоставляется с модой  $j$  с силой, заданной апостериорной вероятностью:

$$\rho_{ij} = \frac{\exp\left(-\frac{1}{2} \frac{1}{\sigma_j} (x_i - \mu_j)^2\right)}{\sum_{l=1}^J \exp\left(-\frac{1}{2} \frac{1}{\sigma_j} (x_i - \mu_l)^2\right)}.$$

Далее для каждой моды  $j$  следует рассчитать векторы математического

ожидания ( $M$ ) и среднеквадратического отклонения ( $\Sigma$ ):

$$\begin{aligned} [M_{kj}; \Sigma_{kj}] = \\ = \left[ \frac{1}{N\sqrt{a}} \sum_{i=1}^N \rho_{ij} \frac{x_{ki} - \mu_{kj}}{\sigma_{kj}}; \frac{1}{N\sqrt{2a}} \left( \sum_{i=1}^N \rho_{ij} \left( \frac{x_{ki} - \mu_{kj}}{\sigma_{kj}} \right)^2 - 1 \right) \right], \end{aligned}$$

где  $k = \overline{1, D}$  охватывает размерность вектора.

Основываясь на этих расчетах, вектор Фишера изображения  $I$  составляется следующим образом:

$$F(I) = [M_j; \Sigma_j], l = \overline{1, J}.$$

В процессе автоматизированной классификации земель с расчетом параметров окрестности с каждым элементарным участком (пикселем) данных ДЗЗ следует сопоставить как его спектральные параметры, так и дескрипторы его окрестности, расчет которых проводится по представленным выше формулам. На базе вектора этих дескрипторов предлагается осуществлять классификацию пространственных объектов для создания карт земель.

Эксперимент по тематической интерпретации данных ДЗЗ с учетом дескрипторов окрестности проводился по следующему плану:

- 1) определение проблемной ситуации и выявление целей создания карт земель;
- 2) подбор данных для автоматизированного анализа и определение формы номенклатуры классов конечного результата;
- 3) предварительная обработка данных ДЗЗ, включающая в себя расчет дескрипторов окрестности (для этого необходимо экспертно подобрать размер окрестности, ее форму, а так-

<sup>16</sup> Ямашкин С. А. [Методическое и алгоритмическое обеспечение процесса анализа структуры земель на базе данных дистанционного зондирования](#) : дис. ... канд. техн. наук. Пенза, 2016.

<sup>17</sup> Ямашкин С. А. [Методическое и алгоритмическое обеспечение процесса анализа структуры земель на базе данных дистанционного зондирования](#) : автореф. дис. ... канд. техн. наук. Пенза, 2016.



же определиться с самим перечнем рассчитываемых информативных дескрипторов);

4) сопоставление всех атомарных участков космической съемки вектора с его параметрами:  $X = [A M]$ , где  $A = [\lambda_1, \dots, \lambda_n]$  – вектор значений спектральных яркостей участка, а  $M = [\mu_1, \dots, \mu_k]$  – вектор дескрипторов его окрестности;

5) выбор алгоритма для классификации данных (гибридные системы, нейросетевые алгоритмы, машины опорных векторов (SVM), деревья принятия решений), определение его архитектуры, функциональных параметров и средств программной реализации;

6) формирование обучающих и тестовых выборок на базе априорной информации об анализируемой территории, полученной в ходе полевых исследований;

7) обучение классификатора его типу и архитектуре;

8) классификация материалов космической съемки с целью выделения интересующих типов земель;

9) оценка качества проведенной классификации на базе расчета численных метрик и экспертной оценки; формирование вывода о качестве решения поставленной задачи.

Для чистоты и полноты эксперимента выполнение пунктов 3–9 данного плана необходимо проводить с различными дескрипторами окрестности и архитектурой классифицирующих алгоритмов. На основе полученных результатов необходимо сделать вывод о влиянии значения дескрипторов окрестности на процесс проведения классификации и ее результат.

#### **Результаты исследования**

Апробация алгоритма анализа данных ДЗЗ с учетом дескрипторов окрестности была проведена на четырех тестовых полигонах: «Инерка», «Смольный», «Мамангино», «Чеберчинка» (рисунок). Полигон «Инерка» (центр:  $54^{\circ}3'52.21''$  с. ш.,  $45^{\circ}53'11.20''$  в. д.) отражает взаи-

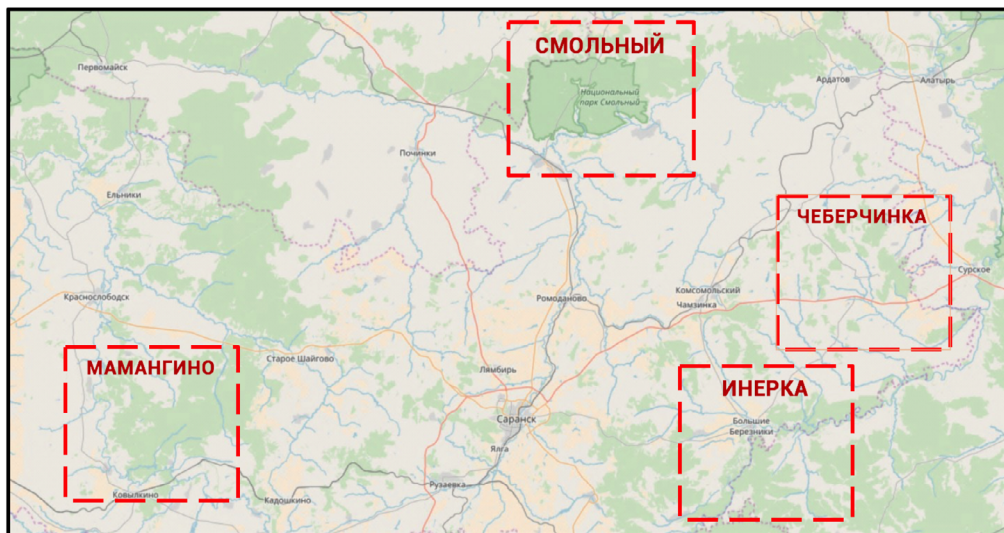
модействие парагенетических систем останцово-водораздельных массивов эрозионно-денудационной равнины и долины р. Сура [15]; «Смольный» (центр:  $54^{\circ}42'18.72''$  с. ш.,  $45^{\circ}18'33.79''$  в. д.) выделен в зоне взаимодействия лесостепных геосистем вторичных моренных, лесных ландшафтов водно-ледниковых и древнеаллювиальных равнин [16]; «Чеберчинка» (центр:  $54^{\circ}27'36.46''$  с. ш.,  $46^{\circ}18'50.60''$  в. д.) отражает своеобразие развития лесостепных ландшафтов эрозионно-денудационных равнин полигона, что связано с активными тектоническими инверсиями в неоген-четвертичное время [17]; полигон «Мамангино» (центр:  $54^{\circ}13'13.42''$  с. ш.,  $43^{\circ}53'30.05''$  в. д.) организован в парагенетических системах лесных ландшафтов водно-ледниковых и древнеаллювиальных равнин.

В качестве исходных данных выбраны космические снимки Landsat-7. Так, для анализа земель национального парка «Смольный» изначально предполагалось, что номенклатура результата классификации будет строиться на разделении классов растительности, в т. ч. хвойных, широколиственных, мелколиственных лесов, а также земель, не покрытых растительностью. Для эффективного анализа данных ДЗЗ разумно выбрать комбинацию коротковолнового ИК, ближнего ИК и красного каналов Landsat-7. Сочетание этих спектральных зон информативно при решении задачи классификации земель, покрытых растительностью.

Был проведен расчет дескрипторов окрестности. Для оценки влияния размера анализируемой окрестности на результат классификации были проведены эксперименты с учетом характеристик дескрипторов окрестности радиусов 10 px (300 м), 50 px (1 500 м), 100 px (3 000 м), а также без учета этих параметров.

В качестве классификатора данных была использована нейронная сеть прямого распространения; совместно





Р и с у н о к. Система тестовых полигонов на территории Республики Мордовия  
 F i g u r e. The system of test polygons in the Republic of Mordovia

с экспертом для нее были подготовлены обучающие, валидационные и тестовые выборки. Тематическая интерпретация материалов космической съемки нейронными сетями, обученными на этих выборках, позволила сгенерировать карты растительности и выделить участки, соответствующие землям, покрытым хвойными, широколиственными, мелколиственными лесами, а также песками. Классификация с учетом дескрипторов окрестности радиусом 300 м привела к получению генерализованного результата. Классификация данных ДЗЗ с учетом дескрипторов окрестности радиусом 1,5 км позволила выделить песчаные участки, не покрытые растительностью, которые были недостаточно заметны в результате остальных экспериментов.

Это стало возможным благодаря учету информативных дескрипторов окрестности, особенности которых стали проявляться в данном масштабе. Классификация земель с учетом дескрипторов окрестности радиусом 3 км напротив, позволила, добиться существенной генерализации, при которой

с результирующей карты были убраны незначительные территориальные объекты. Таким образом, анализ специфики окрестных территорий с опорой на вычисление пространственных закономерностей и особенностей позволяет с большей точностью осуществлять классификацию земель. Объективный учет дескрипторов окрестности возможен при комплексном исследовании различных свойств окружения (табл. 1), так как каждый из них предоставляет различную аутентичную информацию о пространственных комплексах.

Проведенные эксперименты позволили сделать вывод о том, что классификация геоданных с учетом дескрипторов окрестности позволяет осуществлять более гибкую настройку получаемого результата, который часто характеризуется большей точностью (табл. 2).

Использование в качестве особенностей классифицируемых территорий рассчитываемых дескрипторов окрестности позволяет в многомерных пространствах признаков анализируемых земель получить компактные кластеры признаков. Диагностика аутентич-



**Показатели точности при анализе различных типов геофизической поверхности**  
**Accuracy indicators for the analysis of different types of geophysical surface**

Дескриптор окрестности / Descriptor of neighborhood	Точность классификации земель, % / Accuracy of land classification, %		
	Растительность / Vegetation	Вода / Water	Почва / Soil
Граничный (локальный диапазон) / Boundary (local range)	91,82	90,04	93,11
Неоднородность (локальная энтропия, среднеквадратическое отклонение) / Inhomogeneity (local entropy, standard deviation)	94,27	92,95	91,01
Цветовой (цветовой момент, гистограмма оттенков, кортеж цветов) / Color (color moment, histogram of hues, tuple of colors)	91,96	92,30	93,24
Комплексный дескриптор / Complex descriptor	95,82	93,33	94,23

**Показатели точности классификации при учете дескрипторов разного радиуса  
(полигоны «Инерка», «Смольный», «Чеберчинка», «Момангино»)**  
**Indicators of classification accuracy, taking into account descriptors of a different radius  
(polygons “Inerka”, “Smolnyy”, “Cheberchinka”, “Momangino”)**

Радиус окрестности / Radius neighborhood	Точность классификации на полигонах, % / Accuracy of classification on polygons, %			
	Инерка / Inerka	Смольный / Smolnyy	Чеберчинка / Cheberchinka	Момангино / Momangino
Без учета окрестности (0 px) / Excluding the neighborhood (0 px)	84,04	88,01	81,12	82,35
10 px	89,27	92,32	87,23	86,91
50 px	93,72	93,45	91,02	92,99
100 px	90,65	96,07	93,71	92,87

ных свойств внутренней структуры пространственных объектов на основе анализа разнообразия динамических и инвариантных характеристик способствует подавлению шумового воздействия, генерализации результата и общего увеличения точности классификации.

Апробация разработанного алгоритма на четырех тестовых полигонах показала значение точности классификации в диапазоне 81–88 % (без учета дескрипторов окрестности). С учетом свойств окрестности точность возрастает до

96 %. В ходе экспериментов было выявлено, что существенное увеличение радиуса анализируемой окрестности ведет к понижению точности интерпретации космических снимков. Применение разработанных алгоритмов, основанных на анализе инвариантных и динамических дескрипторов окрестности, позволяет оперативно проводить автоматизированный анализ территориальных комплексов с целью картографирования систем землепользования и определения векторов развития стихийных процессов.

Данный комплекс алгоритмов реализован в виде набора скриптов для классификации материалов космической съемки на языке MATLAB. Алгоритмы условно подразделяются на два ключевых блока: 1) для анализа данных, описывающих инвариантные дескрипторы земель с выделением их границ; 2) для анализа динамических свойств территории с целью изучения аутентичных свойств внутренней структуры земель.

### Обсуждение и заключения

Проведенное исследование позволило сделать следующие выводы.

1. Изучена предметная область, отечественный и зарубежный опыт в области анализа материалов космической съемки для оперативного мониторинга состояния земель и прогнозирования развития стихийных природных процессов. Анализ данных источников показал, что при существовании множества решений эффективность существующего инструментария можно существенно повысить путем внедрения оптимизированных алгоритмов анализа пространственных данных.

2. Разработан математический аппарат, который впоследствии послужил основой программного комплекса анализа материалов космической съемки для исследования состояния земель и прогнозирования стихийных природных процессов, а также сформировал методику анализа геофизической оболочки с использованием синтетических дескрипторов окрестности.

3. Созданные алгоритмы реализованы в виде программного комплекса на языке MATLAB. Эффективность предлагаемых решений апробирована на тестовых полигонах. Анализ результатов эксперимента проводился с участием эксперта. Применение методики анализа земель с использованием дескрипторов окрестности на четырех тестовых полигонах показало возможность достижения точности классификации в 97 % с учетом параметров окрестности. Также был сделан вывод о том, что чрезмерное увеличение радиуса анализируемой окрестности неизбежно приводит к снижению точности дешифрирования.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. **Kruse F. A., Lefkoff A. B.** The spectral image processing system (SIPS) – interactive visualization and analysis of imaging spectrometer data // *Remote Sens. Environ.* 1993. Vol. 44. P. 145–163. DOI: [https://doi.org/10.1016/0034-4257\(93\)90013-N](https://doi.org/10.1016/0034-4257(93)90013-N)
2. **Kohonen T.** Self-organizing maps (third extended edition). New York, 2001. 501 p. DOI: <https://doi.org/10.1007/978-3-642-56927-2>
3. **Rosenblatt F.** Analytic techniques for the study of neural nets // *IEEE Trans. on Appl. and Industry.* 1964. Vol. 83, no. 74. P. 285–292. URL: <https://ieeexplore.ieee.org/document/5407758>
4. **Breiman L.** Bagging predictors // *Machine Learning.* 1996. Vol. 24, no. 2. P. 123–140. URL: <https://www.stat.berkeley.edu/~breiman/bagging.pdf>
5. **Foody G. M., Mathur A.** Toward intelligent training of supervised image classifications: direct training data acquisition for SVM classification // *Remote Sensing of Environment.* 2004. Vol. 93. P. 107–117. DOI: <https://doi.org/10.1016/j.rse.2004.06.017>
6. **Polikar R.** Bootstrap inspired techniques in computational intelligence: ensemble of classifiers, incremental learning, data fusion and missing features // *IEEE Signal Processing Magazine.* 2007. Vol. 24, no. 4. P. 59–72. URL: <http://users.rowan.edu/~polikar/RESEARCH/PUBLICATIONS/spm2007.pdf>
7. **Polikar R.** Ensemble based systems in decision making // *IEEE Circuits and Systems Magazine.* 2006. Vol. 6, no. 3. P. 21–45. DOI: <https://doi.org/10.1109/MCAS.2006.1688199>



8. Woods K., Kegelmeyer W. P. J., Bowyer K. Combination of multiple classifiers using local accuracy estimates // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1997. Vol. 19, no. 4. P. 405–410. DOI: <https://doi.org/10.1109/34.588027>

9. Ямашкин С. А., Ямашкин А. А., Фролов А. Н. Базовые геоинформационные ресурсы для оптимизации регионального землепользования // Международный научно-исследовательский журнал. 2016. Т. 53, № 11-2. С. 85–88. DOI: <https://doi.org/10.18454/IRJ.2016.53.009>

10. Ямашкин С. А. Использование ансамбль-систем в анализе данных ДЗЗ для мониторинга состояния земель и прогнозирования стихийных природных процессов // Фундаментальные и прикладные исследования: проблемы и результаты. 2016. № 28. С. 71–77.

11. Использование ансамбль-систем для картографирования ландшафтов / С. А. Ямашкин [и др.] // Геодезия и картография. 2016. № 7. С. 42–49. DOI: <https://doi.org/10.22389/0016-7126-2016-913-7-42-49>

12. Ямашкин С. А., Ямашкин А. А. Моделирование синтетической ландшафтной карты // Перспективы развития информационных технологий. 2016. № 32. С. 68–74. URL: <https://elibrary.ru/item.asp?id=27114514>

13. Ямашкин С. А. Мониторинг состояния земель и прогнозирования стихийных процессов на основе анализа инвариантных и динамических параметров территории // Международный научно-исследовательский журнал. 2016. Т. 51, № 9-2. С. 110–112. DOI: <https://doi.org/10.18454/IRJ.2016.51.129>

14. Синтетическое ландшафтное картографирование в региональной ГИС / С. А. Ямашкин [и др.] // Международный научно-исследовательский журнал. 2017. Т. 58, вып. 04-1. С. 54–56. DOI: <https://doi.org/10.23670/IRJ.2017.58.003>

15. Ямашкин А. А., Ямашкин С. А., Акашкина А. Г. ГИС-моделирование ландшафтного разнообразия // Геодезия и картография. 2013. № 11. С. 40–45. URL: [http://geocartography.ru/scientific\\_article/2013\\_11\\_40-45](http://geocartography.ru/scientific_article/2013_11_40-45)

16. Ямашкин А. А., Ямашкин С. А. Применение алгоритма выделения краев к решению задачи моделирования границ ландшафтов // Вестник Воронежского государственного университета (Сер. «География. Геоэкология»). 2013. № 2. С. 28–34. URL: [www.vestnik.vsu.ru/pdf/geo-graph/2013/02/2013-02-04.pdf](http://www.vestnik.vsu.ru/pdf/geo-graph/2013/02/2013-02-04.pdf)

17. Применение ГИС в анализе морфологической структуры ландшафтов / А. А. Ямашкин [и др.] // Вестник Удмуртского университета (Сер. «Биология. Науки о Земле»). 2013. № 6-3. С. 115–122. URL: <http://docplayer.ru/29381408-A-a-yamashkin-s-a-yamashkin-a-a-klkunov-a-g-akashkina-yu-s-shukshin-primenenie-gis-v-analize-morfologicheskoy-struktury-landshaftov.html>

*Поступила 01.12.2017; принята к публикации 20.02.2018; опубликована онлайн 20.09.2018*

*Об авторах:*

**Ямашкин Станислав Анатольевич**, старший преподаватель кафедры автоматизированных систем обработки информации и управления, ФГБОУ ВО «МГУ им. Н. П. Огарёва» (430005, Россия, г. Саранск, ул. Большевикская, д. 68/1), кандидат технических наук, Researcher ID: N-2939-2018, ORCID: <https://orcid.org/0000-0002-7574-0981>, [yamashkinsa@mail.ru](mailto:yamashkinsa@mail.ru)

**Ямашкин Анатолий Александрович**, декан географического факультета, ФГБОУ ВО «МГУ им. Н. П. Огарёва» (430005, Россия, г. Саранск, ул. Большевикская, д. 68/1), доктор географических наук, профессор, Researcher ID: N-2941-2018, ORCID: <https://orcid.org/0000-0001-9995-8371>, [yamashkin56@mail.ru](mailto:yamashkin56@mail.ru)

*Вклад соавторов:*

С. А. Ямашкин – разработка методики и алгоритма анализа земель, реализация математического аппарата в виде комплекса программ; А. А. Ямашкин – анализ экспериментальных результатов.

## REFERENCES

1. Kruse F. A., Lefkoff A. B. The spectral image processing system (SIPS) – interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ.* 1993; 44:145–163. DOI: [https://doi.org/10.1016/0034-4257\(93\)90013-N](https://doi.org/10.1016/0034-4257(93)90013-N)
2. Kohonen T. Self-organizing maps (third extended edition). New York, 2001. DOI: <https://doi.org/10.1007/978-3-642-56927-2>
3. Rosenblatt F. Analytic techniques for the study of neural nets. *IEEE Trans. on Appl. and Industry.* 1964; 83(74):285–292. Available at: <https://ieeexplore.ieee.org/document/5407758>
4. Breiman L. Bagging predictors. *Machine Learning.* 1996; 24(2):123–140. Available at: <https://www.stat.berkeley.edu/~breiman/bagging.pdf>
5. Foody G. M., Mathur A. Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. *Remote Sensing of Environment.* 2004; 93:107–117. DOI: <https://doi.org/10.1016/j.rse.2004.06.017>
6. Polikar R. Bootstrap inspired techniques in computational intelligence: ensemble of classifiers, incremental learning, data fusion and missing features. *IEEE Signal Processing Magazine.* 2007; 24(4):59–72. Available at: <http://users.rowan.edu/~polikar/RESEARCH/PUBLICATIONS/spm2007.pdf>
7. Polikar R. Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine.* 2006; 6(3):21–45. DOI: <https://doi.org/10.1109/MCAS.2006.1688199>
8. Woods K., Kegelmeyer W. P. J., Bowyer K. Combination of multiple classifiers using local accuracy estimates. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 1997; 19(4):405–410. DOI: <https://doi.org/10.1109/34.588027>
9. Yamashkin S. A., Yamashkin A. A., Frolov A. N. Basic information resources for optimisation of the regional land use. *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal = International Research Journal.* 2016; 53(11-2):85–88. DOI: <https://doi.org/10.18454/IRJ.2016.53.009> (In Russ.)
10. Yamashkin S. A. [Using ensemble-based systems in the analysis of remote sensing data for monitoring the state of lands and forecasting natural processes]. *Fundamentalnyye i prikladnyye issledovaniya: problemy i rezul'taty = Fundamental and Applied Research: Problems and Results.* 2016; 28:71–77. (In Russ.)
11. Yamashkin S. A., Radovanovich M. M., Yamashkin A. A., Vukovich D. V., Frolov A. N. Using ensemble-based systems for the landscapes mapping. *Geodeziya i kartografiya = Geodesy and Cartography.* 2016; 7:42–49. DOI: <https://doi.org/10.22389/0016-7126-2016-913-7-42-49> (In Russ.)
12. Yamashkin S. A., Yamashkin A. A. [Modeling a synthetic landscape map]. *Perspektivy razvitiya informatsionnykh tekhnologiy = Prospects for Development of Information Technology.* 2016; 32:68–74. Available at: <https://elibrary.ru/item.asp?id=27114514> (In Russ.)
13. Yamashkin S. A. Monitoring of land and natural processes prediction based on analysis of invariant and dynamic parameters of the territory. *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal = International Research Journal.* 2016; 51(9-2):110–112. DOI: <https://doi.org/10.18454/IRJ.2016.51.129> (In Russ.)
14. Yamashkin A. A., Yamashkin S. A., Frolov A. N., Zarubin O. A. Synthetic landscape mapping in the regional GIS. *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal = International Research Journal.* 2017; 58(04-1):54–56. DOI: <https://doi.org/10.23670/IRJ.2017.58.003> (In Russ.)
15. Yamashkin A. A., Yamashkin S. A., Akashkina A. G. [GIS-modeling of landscape diversity]. *Geodeziya i kartografiya = Geodesy and cartography.* 2013; 11:40–45. Available at: [http://geocartography.ru/scientific\\_article/2013\\_11\\_40-45](http://geocartography.ru/scientific_article/2013_11_40-45) (In Russ.)
16. Yamashkin A. A., Yamashkin S. A. [Application of the algorithm for edge selection to the solution of the problem of modeling the borders of landscapes]. *Vestnik Voronezhskogo gosudarstvennogo universiteta. Ser. Geografiya. Geoekologiya = Voronezh State University Bulletin. Series: Geography, Geoecology.* 2013; 2:28–34. Available at: [www.vestnik.vsu.ru/pdf/geograph/2013/02/2013-02-04.pdf](http://www.vestnik.vsu.ru/pdf/geograph/2013/02/2013-02-04.pdf) (In Russ.)



17. Yamashkin A. A., Yamashkin S. A., Klikunov A. A., Akashkina A. G., Shukshin Yu. S. Application of GIS in the analysis of morphological structure of the landscape. *Vestnik Udmurtskogo universiteta. Ser.: Biologiya. Nauki o Zemle* = Udmurt University Bulletin. Series: Biology, Earth Sciences. 2013; 6-3:115–122. Available at: <http://docplayer.ru/29381408-A-a-yamashkin-s-a-yamashkin-a-a-klikunov-a-g-akashkina-yu-s-shukshin-primenenie-gis-v-analize-morfologicheskoy-struktury-landshaftov.html> (In Russ.)

*Received 01.12.2017; revised 20.02.2018; published online 20.09.2018*

*About authors:*

**Stanislav A. Yamashkin**, Senior Lecturer of Chair of Automated Systems of Information Processing and Control, National Research Ogarev Mordovia State University (68/1, Bolshevistskaya St., Saransk 430005, Russia), Ph.D. (Engineering), Researcher ID: N-2939-2018, ORCID: <https://orcid.org/0000-0002-7574-0981>, [yamashkinsa@mail.ru](mailto:yamashkinsa@mail.ru)

**Anatoliy A. Yamashkin**, Dean of Geography Faculty, National Research Ogarev Mordovia State University (68/1, Bolshevistskaya St., Saransk 430005, Russia), D.Sc. (Geography), Professor, Researcher ID: N-2941-2018, ORCID: <https://orcid.org/0000-0001-9995-8371>, [yamashkin56@mail.ru](mailto:yamashkin56@mail.ru)

*Author's contribution:*

S. A. Yamashkin developed the methods and algorithms for the analysis of lands and implementation of mathematical apparatus in form of a program system. A. A. Yamashkin analyzed the experimental results.

*All authors have read and approved the final version of the paper.*