

Сравнение методов структурного анализа двухмерных спектральных данных дистанционного зондирования лесных массивов, повреждённых сибирским шелкопрядом

А.В. Лапко^{1,2}, В.А. Лапко^{1,2}

¹ Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева, 660037, г. Красноярск, пр. «Красноярский рабочий», д. 31;

² Институт вычислительного моделирования СО РАН, 660036, Россия, г. Красноярск, ул. Академгородок, д. 50, стр. 44

Аннотация

Представляется подход структурного анализа данных дистанционного зондирования, основанный на разбиении области значений двухмерных спектральных признаков с использованием компонент коэффициента корреляции. Структурный анализ данных основывается на анализе знаков составляющих компонент коэффициента корреляции, которые соответствуют нормированным значениям спектральных признаков данных дистанционного зондирования. На этой основе определяются четыре класса, которым соответствуют значения составляющих компонент коэффициента корреляции: положительные, знакопеременные и отрицательные. Для анализа данных дистанционного зондирования формируется решающее правило, которое позволяет относить контрольные ситуации к одному из четырёх классов. В зависимости от исследуемых пар спектральных признаков каждый класс характеризует уникальную сигнатуру, соответствующую определённому типу земной поверхности. Для проведения дополнительных исследований полученных классов предлагается методика декомпозиции спектральных данных. Методика основана на преобразовании составляющих компонент коэффициента корреляции в виде произведения или их евклидова расстояния. Для визуального представления полученных результатов используется ядерная оценка плотности вероятности. Приводится методика её оптимизации. Эффективность предложенных методик декомпозиции сравнивается с анализом данных дистанционного зондирования, основанного на использовании значений нормализованного разностного инфракрасного индекса. Этот индекс преобразует информацию спектральных каналов ближнего и коротковолнового инфракрасного. Проведено сравнение результатов применения рассматриваемых методик структурного анализа на тестовом участке лесного массива, повреждённого сибирским шелкопрядом, расположенного на территории Красноярского края. Приводятся картосхемы и числовые характеристики полученных результатов.

Ключевые слова: структурный анализ данных, автоматическая классификация, компоненты коэффициента корреляции, ядерная оценка плотности вероятности, данные дистанционного зондирования, спектральные признаки, лесной массив, индекс NDII, сибирский шелкопряд.

Цитирование: Лапко, А.В. Сравнение методов структурного анализа двухмерных спектральных данных дистанционного зондирования лесных массивов, повреждённых сибирским шелкопрядом / А.В. Лапко, В.А. Лапко // Компьютерная оптика. – 2026. – Т. 50, № 4. – 1800. – doi:10.18287/COJ1800.

Citation: Lapko AV, Lapko VA. Comparison of methods for structural analysis of two-dimensional spectral remote sensing data of forests damaged by the Siberian silkmoth. Computer Optics 2026; 50(4): 1800. doi:10.18287/COJ1800.

Введение

Совершенствование систем дистанционного мониторинга приводит к появлению инновационных подходов к анализу данных, использующих автоматизированную обработку спектральных признаков. Исследование закономерностей, характерных для объектов различной природы, начинается с выявления компактных групп статистических данных. Эта задача решается с использованием алгоритмов автоматической классификации [1–3]. Обзор и систематизация методов автоматической классификации представлены в работах [4–5]. Важнейшим этапом изучения закономерностей функционирования природных объектов является структурный анализ их спектральных характеристик, включающий предварительную кластеризацию данных дистанционного зондирования и использование методов декомпозиции. Полученные результаты являются основой создания систем мониторинга состояний природных объектов.

В настоящее время в качестве альтернативного подхода структурного анализа данных дистанционного зондирования используются значения вегетационных индексов, например, NDVI [6], GNDVI [7], NDII [8, 9]. Они

основаны на нелинейном преобразовании спектральных каналов сенсоров космических аппаратов. Например, нормализованный относительный индекс растительности NDVI использует контраст характеристик в двух спектральных каналах: поглощение пигментом хлорофилла в красной спектральной зоне (RED) и высокую отражательную способность в ближнем инфракрасном спектре (NIR). Он является безразмерным и определяется в диапазоне $[-1; 1]$. Методика его использования при анализе данных дистанционного зондирования состоит в определении пороговых значений на интервале изменения при участии эксперта.

Методика структурного анализа данных дистанционного зондирования, основанная на использовании составляющих коэффициента корреляции, предложена в работах [10, 11]. Её применение позволяет осуществлять декомпозицию спектральных данных дистанционного зондирования на четыре класса наблюдений. В соответствии со спецификой прикладных задач возникает необходимость проведения дополнительной декомпозиции полученных классов в пространстве показателей спектральных признаков.

Цель работы состоит в развитии метода структурного анализа двумерных спектральных данных дистанционного зондирования и его сравнении с нормализованным разностным инфракрасным индексом NDII при обработке информации лесного массива, повреждённого сибирским шелкопрядом.

1. Методика структурного анализа данных дистанционного зондирования

Пусть имеется выборка $V = (x_1^i, x_2^i, z_1^i, z_2^i, i = \overline{1, n})$ объёма n данных дистанционного зондирования, составленная из значений спектральных признаков x_1, x_2 и соответствующих им координат земной поверхности (z_1, z_2) . Для анализа зависимости между спектральными признаками x_1 и x_2 используем компоненты статистической оценки коэффициента корреляции.

Статистическая оценка коэффициента корреляции между спектральными признаками x_1 и x_2 представима в виде

$$R = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_1^i - \bar{x}_1}{\bar{\sigma}_1} \right) \left(\frac{x_2^i - \bar{x}_2}{\bar{\sigma}_2} \right) = \frac{1}{n-1} \sum_{i=1}^n r_1^i r_2^i, \quad (1)$$

где $\bar{\sigma}_v = \left(\frac{1}{n-1} \sum_{i=1}^n (x_v^i - \bar{x}_v)^2 \right)^{1/2}$ – оценка среднего квадратического отклонения x_v , а $\bar{x}_v = \frac{1}{n} \sum_{i=1}^n x_v^i$ – среднее значение признака x_v , $v = 1, 2$; $r_1^i = (x_1^i - \bar{x}_1) / \bar{\sigma}_1$, $r_2^i = (x_2^i - \bar{x}_2) / \bar{\sigma}_2$.

Статистическая оценка коэффициента корреляции (1) рассчитывается на основе составляющих $r_1^i, r_2^i, i = \overline{1, n}$ (1). Значения произведения r_1^i, r_2^i характеризуют относительную зависимость между случайными величинами x_1, x_2 в конкретных условиях x_1^i, x_2^i .

Сформируем решающее правило на основе составляющих r_1^i, r_2^i компонент $r^i = r_1^i r_2^i$ оценки коэффициента корреляции [12]

$$m(x_1^i, x_2^i): \begin{cases} (x_1^i, x_2^i) \in \Omega_1, & \text{если } r_1^i > 0 \text{ и } r_2^i > 0; \\ (x_1^i, x_2^i) \in \Omega_2, & \text{если } r_1^i < 0 \text{ и } r_2^i < 0; \\ (x_1^i, x_2^i) \in \Omega_3, & \text{если } r_1^i > 0 \text{ и } r_2^i < 0; \\ (x_1^i, x_2^i) \in \Omega_4, & \text{если } r_1^i < 0 \text{ и } r_2^i > 0, i = \overline{1, n}. \end{cases} \quad (2)$$

Введём четыре типа $\Omega_j, j = \overline{1, 4}$, качественной зависимости между спектральными признаками x_1, x_2 , которые характеризуются положительными и отрицательными значениями составляющих r_1, r_2 компонент r оценки коэффициента корреляции R . Назовём области $\Omega_j, j = \overline{1, 4}$, классами, которые определяются условиями (2).

На основе решающего правила (2) будем формировать массивы данных $V_j = (r^i, i \in I_j), j = \overline{1, 4}$, двумя способами:

$$V_j^1 = (r^i = r_1^i r_2^i, i \in I_j), \quad (3)$$

$$V_j^2 = \left(r^i = \sqrt{(r_1^i)^2 + (r_2^i)^2}, i \in I_j \right), j = \overline{1, 4}. \quad (4)$$

Здесь I_j – множество номеров элементов выборки V_j , принадлежащих классу Ω_j , а n_j – их количество. Каждый j -й массив данных V_j соответствует обнаруженным классам $\Omega_j, j = \overline{1, 4}$.

При исследовании природных объектов возникает задача дополнительного анализа классов $\Omega_j, j = \overline{1, 4}$.

Для этого воспользуемся ядерными оценками плотностей вероятностей [13, 14]

$$\bar{p}_j(r) = \frac{1}{n_j c_j} \sum_{i \in I_j} \Phi \left(\frac{r - r^i}{c_j} \right), j = \overline{1, 4}. \quad (5)$$

Ядерные функции $\Phi(u)$ в статистике (5) удовлетворяют условиям [13]:

$$\Phi(u) = \Phi(-u), \quad 0 \leq \Phi(u) < \infty, \int_{-\infty}^{\infty} \Phi(u) du = 1, \int_{-\infty}^{\infty} u^m \Phi(u) du < \infty, \quad 0 \leq m < \infty.$$

Оценки плотности вероятности (5) формируются на основе массивов данных V_j^1 или V_j^2 , $j = \overline{1, 4}$. Выбор коэффициентов размытости ядерных функций в статистике (5) осуществляется из условия минимума оценок среднеквадратических отклонений непараметрических оценок плотностей вероятностей $\bar{p}_j(r)$ от $p_j(x)$ с использованием метода случайного поиска наилучшей пробы [15 – 17].

При восстановлении плотностей вероятностей будем использовать оптимальную ядерную функцию В.А. Епанечникова [14]

$$\Phi(u) = \begin{cases} \frac{3}{4\sqrt{5}} - \frac{3[u]^2}{20\sqrt{5}}, & \text{если } |u| < \sqrt{5} \\ 0, & \text{если } |u| \geq \sqrt{5}, \end{cases}$$

где $u = (r - r^i)/c_j$.

Развитие метода структурного анализа данных основано на учёте вида плотностей вероятностей (5) при определении пороговых границ λ на значения r (3), (4) для классов Ω_j , $j = \overline{1, 4}$, позволяющих разделить их элементы на однородные группы наблюдений. Тогда правило $m_j(x_1^i, x_2^i)$ дополнительной декомпозиции элементов Ω_j , например, для класса Ω_1 ($j = 1$) (2), при наличии одной пороговой границы будет иметь вид

$$m_1(x_1^i, x_2^i): \begin{cases} (x_1^i, x_2^i) \in \Omega_1^1, & \text{если } r_1^i > 0 \text{ и } r_2^i > 0 \text{ и } r > \lambda; \\ (x_1^i, x_2^i) \in \Omega_1^2, & \text{если } r_1^i > 0 \text{ и } r_2^i > 0 \text{ и } r \leq \lambda; \end{cases}$$

где Ω_1^1, Ω_1^2 – однородные области Ω_1 ; значение r может вычисляться одним из двух способов – (3) или (4).

В общем случае однородные группы элементов из класса Ω_j соответствуют интервалам значений компоненты r с одномодальными фрагментами плотности вероятности $\bar{p}_j(r)$, что допускает наличие нескольких пороговых значений r (3), (4).

Для проведения вычислительных экспериментов разработано программное обеспечение в среде программирования Delphi-7. Программное обеспечение позволяет считывать значения спектральных признаков и их координаты, вычислять традиционные числовые характеристики оценки математического ожидания, дисперсии, среднеквадратического отклонения, коэффициентов асимметрии и эксцесса, моды распределений спектральных признаков. Программа позволяет проводить структурный анализ данных дистанционного зондирования и рассчитывать значения исследуемых вегетационных индексов в каждом i -м элементе исходной выборки V

$$index^i = (x_q^i - x_g^i)/(x_q^i + x_g^i),$$

где x_q^i, x_g^i – значения спектральных признаков под номерами q и g .

С учётом вида плотностей вероятностей программное обеспечение позволяет экспертно определять пороговые значения λ и формировать на этой основе картосхемы в пространстве географических координат, а также рассчитывать количество пикселей, принадлежащих разным состояниям объекта исследования.

2. Анализ результатов вычислительного эксперимента

Для проведения исследований рассматривался тестовый участок лесного массива в Ирбейском районе Красноярского края, на территории которого зарегистрирована в 2017–2019 гг. вспышка массового размножения сибирского шелкопряда. В этих условиях повреждёнными являются в основном темнохвойные древостой, такие как пихта, ель и кедр [18 – 21]. Площадь тестового участка является частью космического снимка исследуемого района от 08 августа 2019 года с аппарата Landsat 8 / OLI. Площадь участка составила 6,24 тыс. га (рис. 1). Тестовый участок состоял из 69342 пикселей с пространственным разрешением 30 м/пиксель. При проведении исследований для каждого пикселя использовались семь спектральных признаков $x = (x_1, \dots, x_7)$ (табл. 1). Снимок атмосферно скорректирован и представлен в единицах отражательной способности. При обработке снимка применен алгоритм С-коррекции для топографической нормализации данных [22].

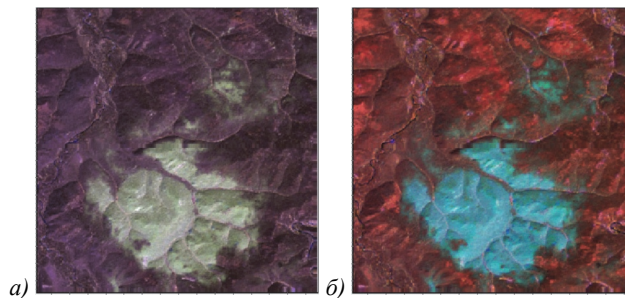


Рис. 1. Фрагмент съемки Landsat-8/OLI от 8 августа 2019 года. Изображение, соответствующее комбинации каналов (SWIR 1, SWIR 2, Green) в композите RGB – (а) и (NIR, SWIR 2, Blue) – (б)

Табл. 1. Характеристика используемых спектральных каналов аппарата Landsat-8 / OLI

Спектральный канал	Диапазон спектра (мкм)	Название
x_1	0,43–0,45	синяя область спектра (Coastal/Aerosol)
x_2	0,45–0,51	синий (Blue)
x_3	0,53–0,59	зеленый (Green)
x_4	0,64–0,67	красный (Red)
x_5	0,85–0,88	ближний инфракрасный (NIR)
x_6	1,57–1,65	коротковолновый инфракрасный (SWIR 1)
x_7	2,11–2,29	коротковолновый инфракрасный (SWIR 2)

В результате сформирована выборка значений спектральных признаков $V = (x_v^i, v = \overline{1, 7}, z_1^i, z_2^i, i = \overline{1, n})$ объёма $n=69342$, каждый элемент которой определяется географическими координатами z_1^i, z_2^i . На основе разработанного программного обеспечения рассчитывались оценки коэффициентов корреляции между парами спектральных признаков (табл. 2).

Табл. 2. Оценки коэффициентов корреляции между спектральными признаками

	x_2	x_3	x_4	x_5	x_6	x_7
x_1	0,992	0,888	0,924	-0,286	0,885	0,924
x_2		0,913	0,954	-0,302	0,9	0,96
x_3			0,904	-0,013	0,947	0,919
x_4				-0,393	0,894	0,973
x_5					-0,018	-0,284
x_6						0,956

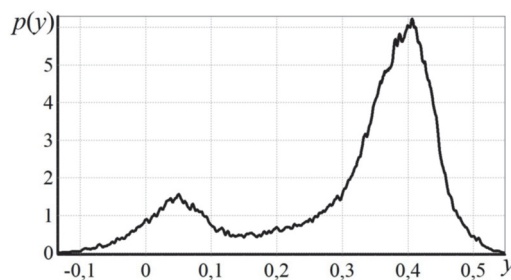
Из табл. 2 следует, что парам признаков (x_1, x_5) , (x_2, x_5) , (x_3, x_5) , (x_4, x_5) , (x_5, x_6) , (x_5, x_7) соответствуют относительно низкие значения оценок коэффициентов корреляции.

Рассмотрим нормализованный разностный инфракрасный индекс NDII (Normalized Difference Infrared Index), который применяется для анализа состояния и динамики лесных территорий в зонах ограниченного влагообеспечения [23]. В отличие от широко известного метода NDVI для расчёта индекса NDII вместо спектрального канала Red (x_4) используются значения в области спектра SWIR1 (x_6). Он является безразмерным в диапазоне изменения $[-1; +1]$. Обозначим значения индекса NDII через переменную y .

Значения индекса NDII (y) можно вычислить в соответствии с выражением:

$$y^i = (x_5^i - x_6^i) / (x_5^i + x_6^i), i = \overline{1, n}, n = 69342.$$

Следуя предложенной методике, определим пороговые значения индекса y . Для этого построим ядерную оценку плотности вероятности типа (5) значений y (рис. 2).

Рис. 2. Ядерная оценка плотности вероятности индекса y

Область определения значений индекса NDII соответствует диапазону $y \in [-0,17; 0,56]$. С учётом двухмодального вида плотности вероятности и введённого выше определения однородного интервала случайной величины установлены пороговые значения индекса NDII, которые соответствуют усохшим $y \in [-0,17; 0,1)$ и усыхающим $y \in [0,1; 0,3]$ темновойным древостоям [20]. Интервал значений $y > 0,3$ соответствует неповреждённым древостоям лесного массива. Усыхающие древостои, характеризующиеся одномодальными фрагментами плотности вероятности $p(y)$, занимают промежуточное положение между двумя интервалами значений y . Картограмма повреждённых древостоев представлена на рис. 3.

На основе картограммы (рис. 3) и оценки плотности вероятности (рис. 2) установлено, что усохшим темновойным древостоям соответствует 10578 пикселей, а усыхающим – 10255. Общее количество

повреждённых темнохвойных древостоев соответствует 20833 пикселям (1875 га). Из картосхемы видно, что часть пикселей с кустарниковой растительностью вдоль ручьёв и малых рек ошибочно отнесены к повреждённым темнохвойным древостоям. Эта особенность свойственна большинству вегетационных индексов, включая NDVI.

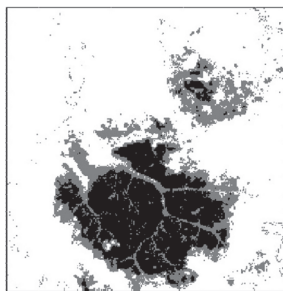


Рис. 3. Картосхема дешифрирования тестового участка с использованием значений индекса NDII (γ). На картосхеме усохшие темнохвойные древостои обозначены черным цветом, а находящиеся в процессе усыхания – серым

Результаты использования методики структурного анализа данных дистанционного зондирования (2) представлены на рис. 4.

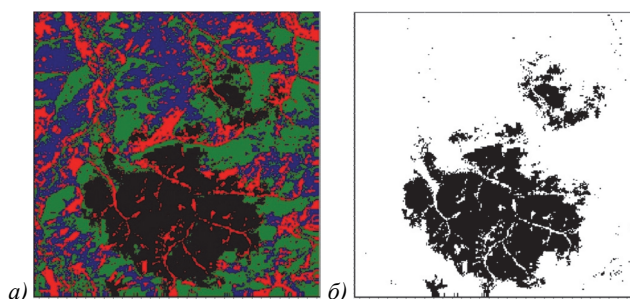


Рис. 4. Картосхемы дешифрирования тестового участка на основе методики структурного анализа в пространстве признаков (x_5 , x_6) (а). Чёрным цветом выделен участок лесного массива, повреждённый сибирским шелкопрядом (б)

На рис. 4 чёрным цветом отмечены участки лесного массива, повреждённые сибирским шелкопрядом. В соответствии с решающим правилом (2) данная область принадлежит классу Ω_4 , который включает усохшие и усыхающие темнохвойные древостои (15384 пикселя, площадью 1384 га).

Проведём дополнительное исследование класса Ω_4 при использовании данных (3) и (4). Соответствующие им ядерные оценки плотностей вероятностей $p(r)$ представлены на рис. 5.

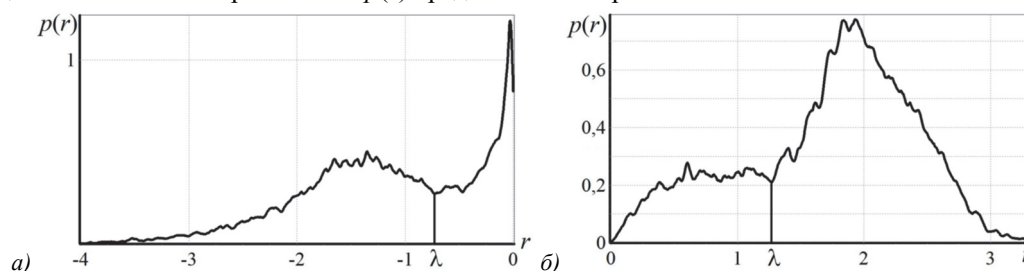


Рис. 5. Ядерные оценки плотности вероятности $p(r)$, восстанавливаемые по данным (3) (а) и (4) (б)

С учётом результатов формирования значений r определены картосхемы усохших и усыхающих темнохвойных древостоев (рис. 6).

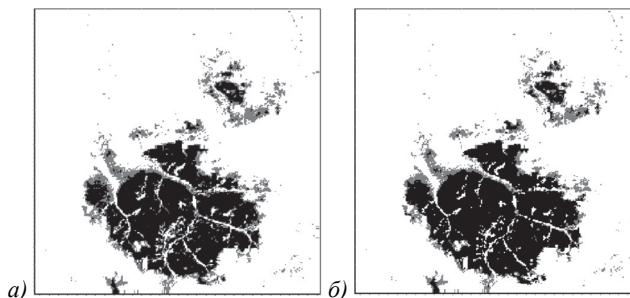


Рис. 6. Картосхемы дешифрирования участка лесного массива, повреждённого сибирским шелкопрядом, в пространстве признаков (x_3 , x_6) при формировании значений r с использованием данных (3) (а) и данных (4) (б). Чёрным цветом выделен участок лесного массива с усохшими древостоями, серым – с находящимися в процессе усыхания

В процессе анализа данных (3), соответствующих рис. 6а, установлен диапазон изменения значений $r \in [-4,53; 0]$. При пороговом значении $\lambda = -0,726$ количество усохших древостоев равно 10013 пикселям (901 га) и усыхающих – 5371 пикселю (483 га). Из анализа данных (4), соответствующих рис. 6б, следует, что диапазон изменения значений $r \in (0; 3,87)$. В этих условиях при пороговом значении $\lambda = 1,266$ количество усохших древостоев соответствует 11609 пикселям (1044 га) и усыхающих – 3775 пикселям (340 га). Приведённые пороговые значения λ на рис. 5а и рис. 5б определяются минимальными значениями плотности вероятности $p(r)$ между двумя модами законов распределения.

Из анализа картосхем следует, что методика (4) формирования значений r при анализе спектральных данных более предпочтительна по сравнению методикой (3). Методика (4) основана на евклидовом расстоянии между значением ноль и контрольной ситуации в пространстве (r_q^i, r_g^i) спектральных признаков с номерами q и g . Поэтому получаемые на её основе результаты обладают устойчивостью к малым значениям составляющих выражения (4). При использовании методики (3) могут возникать ситуации, для которых значения $r_q^i \rightarrow 0$ и $r_g^i \rightarrow \infty$, а их произведение $r_q^i r_g^i \approx 0$ близко к нулю.

Заключение

Упрощённым направлением структуризации данных дистанционного зондирования природных объектов являются методы, основанные на использовании значений вегетационных индексов NDVI, GNDVI, NDII. Их использование связано с экспертным назначением пороговых значений на основе опыта полевых исследований. При использовании значений вегетационных индексов возникает проблема выделения кустарниковой растительности вдоль ручьёв и малых рек как повреждённых сибирским шелкопрядом лесных массивов.

Преимущество предложенной методики структурного анализа состоит в отсутствии необходимости назначения пороговых значений показателей спектральных признаков, используемых для декомпозиции данных дистанционного зондирования. Для проведения дополнительных исследований выделенного класса спектральных данных на участке с однородной растительностью можно использовать два метода расчёта значений показателя r , основанных на произведении составляющих компоненты оценки коэффициента корреляции и на их евклидовом расстоянии от центра координат. С учётом вида плотности вероятности $p(r)$ определяются пороговые значения компоненты r . Основным критерием назначения порогового значения является наличие двухмодального или многомодального закона распределения компонент r . Определены условия компетентности рассматриваемых методов расчёта показателей r , и установлена сопоставимость результатов их применения при обнаружении повреждённых сибирским шелкопрядом древостоев.

Использование значений индекса NDII сопряжено с завышенной оценкой повреждённой сибирским шелкопрядом площади лесного массива по сравнению с предлагаемым методом структурного анализа спектральных данных. Пороговые значения индекса NDII определялись по результатам анализа непараметрической оценки плотности вероятности и экспертных знаний об объекте исследования.

Полученные результаты направлены на автоматизацию процесса структурного анализа двумерных спектральных данных дистанционного зондирования природных объектов.

Дальнейшее развитие предлагаемой методики состоит в декомпозиции многомерных спектральных признаков данных дистанционного зондирования на разработку двухуровневой системы структурного анализа данных дистанционного зондирования. На первом уровне системы осуществляется анализ данных дистанционного зондирования по предлагаемой методике в пространстве одной пары спектральных признаков, например (x_5, x_6) , с выделением класса Ω_j . На втором уровне выделенный класс Ω_j дополнительно исследуется с использованием предложенной методики в пространстве другой пары спектральных признаков, например (x_3, x_4) .

References

- [1] Borzov SM, Potaturkin OI. Selection of the informative feature system for crops classification using hyperspectral data [in Russian]. *Avtometriya* 2020; 56(4): 134-144. doi:10.15372/AUT20200413.
- [2] Shipko VV, Borzov SM. Analysis of the efficiency of hyperspectral data classification under constraints on the quantization bit depth, the number of spectral channels, and spatial resolution [in Russian]. *Avtometriya* 2022; 58(3): 79-87. doi:10.15372/AUT20220309.
- [3] Tuboltsev VP, Lapko AV, Lapko VA. Modified nonparametric algorithm for automatic classification of large-volume statistical data and its application. *Sci Tech Inf Process* 2024; 51(6): 592-597. doi:10.3103/S0147688224700552.
- [4] Dorofeyuk AA. Algorithms of automatic classification (review) [in Russian]. *Avtomatika i Telemekhanika* 1971; 12: 78-113.
- [5] Dorofeyuk AA. Methodology of expert classification analysis in the management and processing of complex data (history and prospects of development) [in Russian]. *Control Sci* 2009; 3(1): 19-28.
- [6] Shevymogov AP, Botvich IYu, Pisman TI, Volkova AI, Kononova NA, Ivanov SA. Informative value of spectral vegetation indices for the meadow and steppe vegetation monitoring of Khakassia by ground and satellite data [in Russian]. *Izv Atmos Ocean Phys* 2024; 60(1): 16-28. doi:10.31857/S0205961424010028.
- [7] Sahu G, Garanayak M, Paikaray BK. Green vegetation detection in satellite imagery using the normalised difference vegetation index method. *Int J Bioinform Res Appl* 2023; 19(4): 327-342. doi:10.1504/IJBRA.2023.135367.

- [8] Sriwongsitanon N, Jandang W, Williams J, Suwawong T, Maekan E, Savenije HHG. Using normalised difference infrared index patterns to constrain semi-distributed rainfall–runoff models in tropical nested catchments. *Hydrol Earth Syst Sci* 2023; 27(11): 2149-2171. doi:10.5194/hess-27-2149-2023.
- [9] Wilson NR, Norman LM. Analysis of vegetation recovery surrounding a restored wetland using the normalized difference infrared index (NDII) and normalized difference vegetation index (NDVI). *Int J Remote Sens* 2018; 39(10): 3243-3274. doi:10.1080/01431161.2018.1437297.
- [10] Lapko AV, Lapko VA. Procedure for decomposing the values of two-dimensional spectral features of remote sensing based on the analysis of correlation coefficient components. *Meas Tech* 2024; 67(6): 427-432. doi:10.1007/s11018-024-02362-6.
- [11] Lapko AV, Lapko VA, Im ST. Decomposition of spectral features of remote sensing based on the components of the correlation coefficient [in Russian]. *Avtometriya* 2025; 61(3): 28-36. doi:10.15372/AUT20250303.
- [12] Lapko AV, Lapko VA, Yuronen YuP. Structural analysis of spectral remote sensing data based on the use of correlation coefficient components. *Computer Optics* 2026; 50(3): 1746. DOI: 10.18287/COJ1746.
- [13] Parzen E. On estimation of a probability density function and mode. *Ann Math Stat* 1962; 33(3): 1065-1076. doi:10.1214/aoms/1177704472.
- [14] Epanechnikov VA. Non-parametric estimation of a multivariate probability density. *Theory Probab Appl* 1969; 14(1): 153-158. doi:10.1137/1114019.
- [15] Rudemo M. Empirical choice of histogram and kernel density estimators. *Scand J Stat* 1982; 9(2): 65-78. Available at: <https://www.jstor.org/stable/4615859>.
- [16] Bowman AW. Comparative study of some kernel-based non-parametric density estimators. *J Stat Comput Simul* 1985; 21(4): 313-327. doi:10.1080/00949648508810822.
- [17] Hall P. Large-sample optimality of least squares cross-validation in density estimation. *Ann Stat* 1983; 11(4): 1156-1174.
- [18] Ponomarev EI, Yakimov ND, Tretyakov PD, Sultson SM. Estimation of defoliation features in dark coniferous tree stands after the impact of Siberian silk moth according to remote data. *Cosmic Res* 2024; 62(S1): S73-S80. doi:10.1134/S0010952524601208.
- [19] Sultson SM, Ponomarev EI, Shvetsov EG, Tretyakov PD, Goroshko AA, Kulakova NN, Mikhaylov PV. Using remote sensing for forecasting damage to dark coniferous forests after Siberian silkmouth outbreak [in Russian]. *Biosfera* 2023; 15(1): 21-32. doi:10.24855/biosfera.v15i1.790.
- [20] Kharuk VI, Im ST, Soldatov VV. Siberian silkmouth outbreaks surpassed geoclimatic barrier in Siberian mountains. *J Mt Sci* 2020; 17(8): 1891-1900. doi:10.1007/s11629-020-5989-3.
- [21] Kharuk VI, Im ST, Ranson KJ, Yagunov MN. Climate-induced northerly expansion of Siberian silkmouth range. *Forests* 2017; 8(8): 301. doi:10.3390/f8080301.
- [22] Riano D, Chuvieco E, Salas J, Aguado I. Assessment of different topographic corrections in Landsat-TM data for mapping vegetation types. *IEEE Trans Geosci Remote Sens* 2003; 41(5): 1056-1061. doi:10.1109/TGRS.2003.811693.
- [23] Hardisky M, Klemas V, Smart R. Influences of soil salinity, growth form, and leaf moisture on the spectral reflectance of *Spartina alterniflora* canopies. *Photogramm Eng Remote Sens* 1983; 49(1): 77-83.

Сведения об авторах

Лапко Александр Васильевич, 1949 года рождения, в 1971 году окончил Фрунзенский политехнический институт по специальности «Автоматика и телемеханика», доктор технических наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник Института вычислительного моделирования Сибирского отделения Российской академии наук, профессор кафедры космических средств и технологий Сибирского государственного университета науки и технологий имени академика М.Ф. Решетнева. Область научных интересов: непараметрическая статистика; распознавание образов и анализ изображений; моделирование и оптимизация неопределённых систем, дистанционное зондирование. E-mail: lapko@icm.krasn.ru, <https://orcid.org/0000-0002-0664-3870>, <https://www.researcherid.com/rid/KFA-9046-2024>

Лапко Василий Александрович, 1974 года рождения, в 1996 году окончил Красноярский государственный технический университет по специальности «Управление и информатика в технических системах», доктор технических наук, профессор, ведущий научный сотрудник Института вычислительного моделирования Сибирского отделения Российской академии наук, заведующий кафедрой космических средств и технологий Сибирского государственного университета науки и технологий имени академика М.Ф. Решетнева. Область научных интересов: непараметрическая статистика; распознавание образов и анализ изображений; моделирование неопределённых систем, дистанционное зондирование. E-mail: valapko@yandex.ru, <http://orcid.org/0000-0001-6938-9323>, <https://www.researcherid.com/rid/C-9863-2018>

Поступила в редакцию 08 августа 2025 г.



Окончательный вариант – 26 сентября 2025 г.

Comparison of methods for structural analysis of two-dimensional spectral remote sensing data of forests damaged by the Siberian silkmoth

A.V. Lapko^{1,2}, V.A. Lapko^{1,2}

¹ Reshetnev Siberian State University of Science and Technology,
Prospekt Krasnoyarsky Rabochoy 31, Krasnoyarsk, 660037, Russia;

² Institute of Computational Modelling SB RAS, Akademgorodok Str. 50, Bldg. 44, Krasnoyarsk, 660036, Russia

Abstract

An approach to structural analysis of remote sensing data is presented, based on the partitioning of the domain of values of two-dimensional spectral features using components of the correlation coefficient. Structural data analysis is based on the analysis of the signs of the correlation coefficient components, which correspond to the normalized values of the spectral features of remote sensing data. On this basis, four classes are defined, to which the values of the components of the correlation coefficient correspond: positive, alternating and negative. To analyze remote sensing data, a decision rule is formed that allows classifying control situations into one of four classes. Depending on the studied pairs of spectral features, each class characterizes a unique signature corresponding to a certain type of earth's surface. To conduct additional studies of the obtained classes, a spectral data decomposition technique is proposed. The method is based on the transformation of the components of the correlation coefficient in the form of a product or their Euclidean distance. For the visual presentation of the obtained results, a kernel probability density estimate is used. A method for its optimization is presented. The effectiveness of the proposed decomposition methods is compared with the analysis of remote sensing data based on the use of normalized difference infrared index values. This index transforms information from spectral channels of near and short-wave infrared. A comparison of the results of applying the considered structural analysis methods was carried out on a test plot of a forest area damaged by the Siberian silkmoth, located in the Krasnoyarsk Territory. Maps and numerical characteristics of the obtained results are provided.

Keywords: structural data analysis, automatic classification, components of the correlation coefficient, kernel probability density estimation, remote sensing data, spectral features, forest massif, NDII index, Siberian silkmoth.

Citation: Lapko AV, Lapko VA. Comparison of methods for structural analysis of two-dimensional spectral remote sensing data of forests damaged by the Siberian silkmoth. *Computer Optics* 2026; 50(4): 1800. doi:10.18287/COJ1800.

About authors

Aleksander Vasilievich Lapko (b. 1949), graduated from Frunze Polytechnic Institute on speciality "Automation and telemechanics" in 1971. Doctor of Technical Sciences, professor, honored worker of science of the Russian Federation, chief researcher of the Institute of Computational Modelling of the Siberian Branch of the Russian Academy of Sciences; Professor of Space Facilities and Technologies department of the Reshetnev Siberian State University of Science and Technology. Research interests: nonparametric statistics; pattern recognition and image analysis; modeling and optimization of uncertain systems; remote sensing. E-mail: lapko@icm.krasn.ru, <https://orcid.org/0000-0002-0664-3870>, <https://www.researcherid.com/rid/KFA-9046-2024>

Vasily Aleksandrovich Lapko (b. 1974), graduated from Krasnoyarsk State Technical University on speciality "Management and Informatics in Technical Systems" in 1996. Doctor of Technical Sciences, professor, leading researcher at the Institute of Computational Modelling of the Siberian Branch of the Russian Academy of Sciences; Head of Space Facilities and Technologies department of the Reshetnev Siberian State University of Science and Technology. Research interests: nonparametric statistics; pattern recognition and image analysis; modeling of uncertain systems; remote sensing. E-mail: valapko@yandex.ru, <http://orcid.org/0000-0001-6938-9323>, <https://www.researcherid.com/rid/C-9863-2018>

Received August 08, 2025. The final version – September 26, 2025.
