

## Формирование информативного индекса для различения заданных объектов гиперспектральных данных

Р.А. Парингер<sup>1,2</sup>, А.В. Мухин<sup>1</sup>, А.В. Куприянов<sup>1,2</sup>

<sup>1</sup> Самарский национальный исследовательский университет имени академика С.П. Королёва, 443086, Россия, г. Самара, Московское шоссе, д. 34;

<sup>2</sup> ИСОИ РАН – филиал ФНИЦ «Кристаллография и фотоника» РАН, 443001, Россия, г. Самара, ул. Молодогвардейская, д. 151

### Аннотация

Работа посвящена разработке подхода, позволяющему по малому числу наблюдений создавать правила различения заданных объектов гиперспектральных данных. Разработка подобного подхода способствовала бы развитию методов и алгоритмов для оперативного анализа гиперспектральных данных, применимых как для предварительной обработки, так и для выполнения разметки гиперспектральных данных. Для реализации подхода предлагается применять технологию, заключающуюся в совместном использовании общих правил вычисления индексов и критериев информативности. В рамках данной работы при реализации предлагаемой технологии индекс задается нормализованной разностной формулой, а информативность оценивается на основе значения критерия разделимости дискриминантного анализа. В результате проведенных исследований, было показано, что с использованием алгоритма, реализующего технологию, была решена задача различения областей гиперспектральных данных с разной растительностью. Сформированный алгоритмом индекс оказался близким по значениям к NDVI. Применяемая технология является генерализацией подхода к формированию правил анализа гиперспектральных данных по малому числу признаков и может быть использована для формирования индексов, информативных в различных задачах.

**Ключевые слова:** классификация, гиперспектральные данные, NDVI, дискриминантный анализ.

**Цитирование:** Парингер, Р.А. Формирование информативного индекса для различения заданных объектов гиперспектральных данных / Р.А. Парингер, А.В. Мухин, А.В. Куприянов // Компьютерная оптика. – 2021. – Т. 45, № 6. – С. 873-878. – DOI: 10.18287/2412-6179-CO-930.

**Citation:** Paringer RA, Mukhin AV, Kupriyanov AV. Formation of an informative index for recognizing specified objects in hyperspectral data. Computer Optics 2021; 45(6): 873-878. DOI: 10.18287/2412-6179-CO-930.

### Введение

Современные гиперспектральные сенсоры позволяют получать гиперспектральные снимки, иначе гиперспектры. Такие снимки состоят из нескольких сотен спектральных слоев и содержат информацию об объектах, недоступную для распространенных повсеместно RGB-камер. Учет этой информации открывает новые возможности в различных областях. Использование гиперспектральных камер в земледелии позволяет оценивать такие параметры, как содержание хлорофилла [1], количество биомассы [2], содержание нитрогенов [3], степень полива растений [4]. Также благодаря гиперспектральным снимкам можно произвести классификацию растений [5–7] и обнаружить их заболевания [8, 9]. Гиперспектральные изображения также успешно применяются при химическом анализе [10] для количественной оценки таких свойств, как содержание белка, фосфора, целлюлозы и т.д. Распространено применение гиперспектрометров также и в медицине [11], например, для измерения параметров рефузии.

Чаще всего для решения задач классификации гиперспектральные данные используются в совокупно-

сти с нейросетевыми алгоритмами [12–14]. Такие алгоритмы чувствительны к обучающему набору данных. На данный момент не существует больших наборов размеченных гиперспектральных данных, подобных наборам ImageNet [15] или COCO [16], содержащих RGB-изображения. Поэтому чаще всего исследователи обучают свои нейронные сети на довольно скромных по объему наборах данных. Наиболее популярным набором гиперспектральных данных является “Hyper spectral Remote Sensing Scenes” [17], который содержит всего 6 размеченных гиперспектров.

Основным препятствием при создании большого набора данных является сложность разметки гиперспектров. Сложность заключается в затрудненном восприятии как отдельных слоёв гиперспектральных данных, так и совокупности множеств слоёв. Кроме того, без проведения программного анализа не всегда возможна разметка некоторых классов объектов в силу их специфики. Так, некоторые классы становятся видны для человека только после применения метода k-средних или метода главных компонент.

Для гиперспектральных данных также свойственна проблема изменяющихся условий съемки. Так, спектры объектов на гиперспектрах, полученных с

одного ракурса, но при различных погодных условиях, могут иметь существенные различия. Эта существенность приводит к тому, что алгоритмы, обученные на наборе данных, снятых в одних условиях, неприменимы к другим наборам данных.

Для борьбы с двумя этими проблемами можно использовать «нормализованные разностные индексы». Такие индексы позволяют разделять заданные группы объектов и корректировать атмосферные искажения и погодные условия. Например, если вычислить нормализованный разностный индекс, используя слои, относящиеся к инфракрасной и красной областям спектра, можно различить в данных области фотосинтетически активной биомассы и определить количество данной массы. Такой индекс называется нормализованным разностным индексом растительности NDVI – Normalized Difference Vegetation Index [18]. Вычисляется данный индекс по формуле (1).

$$NDVI = \frac{NIR - RED}{NIR + RED}, \tag{1}$$

где NIR – слой гиперспектра с длиной волны 0,7–1,0 мкм, RED – слой гиперспектра с длиной волны 0,6–0,7 мкм.

Благодаря нормированности данного индекса он является устойчивым к некоторым видам искажений, влияющих на равномерность яркости.

Поиск индекса для конкретной задачи является нетривиальной проблемой и не имеет на данный момент однозначного решения. Трудности возникают из-за объема данных, неявных физических свойств объектов, выраженной зависимости от условий съёмки и необходимости предобработки данных. Существуют общеизвестные частные решения (NDVI), однако общего правила формирования индексов предложено не было. В связи с этим определение генерализирующей технологии формирования индекса является актуальной задачей, так как она позволит упростить поиск эффективных индексов в различных прикладных областях по аналогии с NDVI.

### 1. Технология формирования информативного индекса

Предлагаемая к применению технология формирования индекса основана на предположении о том, что, задав специальным образом формулу индекса и критерий информативности, можно реализовать алгоритм, способный сформировать индекс, который наилучшим образом разделяет заданные объекты. Индекс, найденный с применением технологии, будем называть информативным. Слои, используемые для вычисления информативного индекса, будем также называть информативными. Предлагаемая к применению технология является универсальной благодаря возможности произвольного задания индексов и возможности применения различных алгоритмов ранжирования признаков. Данная технология позво-

ляет разработать семейство алгоритмов формирования информативного индекса.

Благодаря возможности произвольного задания формулы индекса с помощью предлагаемой к применению технологии можно получить алгоритм, позволяющий найти информативный индекс. Благодаря возможности выбора формулы индекса можно добиться разных результатов для одной задачи. Это позволяет решать любые специфичные задачи за счет выбора наиболее подходящей формулы, дающей желаемый результат. Данная технология не ограничивает число слоёв, используемых для вычисления индекса.

Критерий информативности индекса также можно задать в произвольном виде.

Задав формулу индекса и критерий информативности, на основе применяемой технологии можно разработать алгоритм формирования информативного индекса.

Далее будет предложена одна из возможных реализаций алгоритма. Также будет экспериментально показана работоспособность и эффективность реализованного алгоритма.

### 2. Алгоритм формирования информативного индекса

Для реализации алгоритма на основе применяемой технологии на данном этапе исследования мы остановились на использовании только нормализованного разностного индекса NDI в силу его устойчивости к неравномерности яркости и распространённого применения. Индекс NDI вычисляется по формуле (2).

$$NDI(l_1, l_2) = \frac{l_1 - l_2}{l_1 + l_2}, \tag{2}$$

где  $l_1$  и  $l_2$  – слои гиперспектра.

Для формирования информативного индекса по формуле (2) необходимо найти слои  $l_1$  и  $l_2$ , которые ассоциируются с определенными длинами волн. Разные комбинации таких слоёв при расчете NDI могут позволить различать разные объекты в гиперспектре. Слои, составляющие информативный NDI, также являются информативными.

На данном этапе исследования для реализации алгоритма на основе применяемой технологии мы ограничимся использованием критерия разделимости дискриминантного анализа [19]. Использование данного критерия можно обосновать тем, что он рассматривает отношения центроидов классов, что придаёт дополнительной устойчивости к неравномерной яркости.

Далее алгоритм, основанный на применяемой технологии, перебирает все возможные комбинации пар слоёв гиперспектра. Для каждой пары вычисляет индекс по формуле (2). Оценивает информативность полученного индекса, различающего заданные объекты, на основе значений критерия разделимости дискриминантного анализа.

Далее под признаками будем подразумевать индексы, вычисленные по формуле (2). Исходными данными предлагаемого алгоритма являются: гиперспектр  $HS(l, h, w)$ , где  $l$  – число слоёв,  $h$  – высота,  $w$  – ширина гиперспектра; маска  $GT(h, w)$ , которая содержит разметку объектов. Данные объекты должны принадлежать одному из двух классов, т.к. в данной задаче рассматривается бинарное разделение объектов.

Общее число вычисляемых индексов равно  $(l^2 - l) / 2$ , так как, вычисляя индекс для пар вида  $(l_1, l_1)$ , получим нулевой индекс, а для пар  $(l_1, l_2)$  и  $(l_2, l_1)$  значения индекса будут отличаться лишь знаком.

Также будем рассматривать индексы независимо друг от друга, что позволит существенно сократить количество вычислений при оценке информативности критерием разделимости дискриминантного анализа.

Для каждого индекса вычислим значение  $\lambda_i$ , которое находится по формуле (3).

$$\Lambda = (\lambda_i) = b_i / w_i, \quad (3)$$

где  $B = (b_i)$  – межгрупповое рассеяние наблюдаемых переменных от средних,  $W = (w_i)$  – внутригрупповое рассеяние наблюдаемых переменных от средних.

Значение  $\lambda_i$  описывает отношение межгрупповой вариации к внутригрупповой. Максимальное её значение будет соответствовать признаку, который лучше всего разделяет заданные объекты. Иными словами, максимальное значение  $\lambda_i$  означает, что переменные, соответствующие группам, имеют наименьшее рассеяние относительно их центроидов при наибольшем расстоянии между центроидами.

Значения рассеяний находятся из следующих формул.

$$B = (b_i) = \sum_{k=1}^g n_k (\bar{x}_{ik} - \bar{x}_i)^2, \quad (4)$$

$$W = (w_i) = \sum_{k=1}^g \sum_{m=1}^{n_k} (x_{ikm} - \bar{x}_{ik})^2, \quad (5)$$

где  $g$  – число классов,  $p$  – число признаков,  $n_k$  – число наблюдений в  $k$ -й группе,  $n$  – общее число наблюдений,  $x_{ikm}$  – величина переменной  $i$  для  $m$ -го наблюдения в  $k$ -й группе,  $\bar{x}_{ik}$  – средняя величина переменной  $i$  в  $k$ -й группе,  $\bar{x}_i$  – среднее значение переменной  $i$  по всем группам.

Итого, шаги алгоритма можно сформировать следующим образом:

Исходные данные алгоритма:  $HS$  – гиперспектральный снимок размера  $(h, w, l)$ ;  $GT$  – маска размера  $(h, w)$ .

1. Для каждой пары слоёв  $(l_1, l_2)$  из  $HS$ , где  $0 < l_1 \leq l$  и  $l_2 < l_1$ :

1.1. Вычисляем признак  $NDI$  по формуле (2).

1.2. Вычисляем значение  $b$  для признака  $NDI$  по формуле (4), используя данные о классах из маски  $GT$ .

1.3. Вычисляем значение  $w$  для признака  $NDI$  по формуле (5), используя данные о классах из маски  $GT$ .

1.4. Вычисляем значение  $\lambda$  по формуле (3) и сохраняем его в список  $X$ .

2. Максимальное значение  $\lambda$  из списка  $X$  будет соответствовать наиболее информативному признаку, разделяющему заданные объекты.

Полученный с использованием алгоритма информативный индекс из числа нормализованных разностных будем называть INDI (Informative Normalized Difference Index).

### 3. Выделение объектов на INDI

Для упрощения процесса разметки гиперспектральных данных объекты на INDI можно выделить с помощью бинаризации по двойному порогу. Применение такого подхода для выделения объектов можно обусловить использованием критерия дискриминантного анализа для формирования INDI.

Так, значения порогов мы задавали перцентилями. В рамках проведенного исследования экспериментально были подобраны следующие значения: в качестве левого порога выбран 10-й перцентиль, в качестве правого порога – 90-й перцентиль. Они позволяют адекватно визуализировать результат работы алгоритма.

### 4. Результат работы алгоритма

Для анализа результатов работы предложенного алгоритма был использован гиперспектр “SalinasScene” из набора данных “Hyperspectral Remote Sensing Scenes” [17].

На данном гиперспектральном изображении, представленном на рис. 1а, вручную были выбраны области, соответствующие объектам разных классов. Выбранные объекты задают маску классов, изображенную на рис. 1б. Первый класс соответствует полю стерни (область красного цвета). Второй класс – полю сельдерея (область зеленого цвета). Можно также сказать, что первый класс – это мертвые растения, а второй – живые растения.

Далее, применив к данному гиперспектру и маске предложенный алгоритм, сформируем информативный индекс, изображенный на рис. 1в. Применив бинаризацию двойным порогом, была получена разметка, изображенная на рис. 1г.

Здесь мы можем также убедиться в применимости бинаризации изображения в силу максимизации расстояния между центроидами классов. Распределение классов в найденном информативном признаке изображено на рис. 2. На данном рисунке можно явно наблюдать разделимость заданных классов объектов.

В результате применения алгоритма для вычисления найденного информативного признака по формуле (1) были использованы: слой с длиной волны 750 нм и слой с длиной волны 700 нм. Заметим, что найденный алгоритмом индекс INDI соответствует индексу NDVI. Длина волны первого слоя соответствует инфракрасному излучению, а длина волны второго

слоя соответствует красному спектру видимого света. Также для найденного индекса значения в интервале от 0,1 до 0,2 соответствуют открытой почве, а значения от

0,7 до 0,8 соответствуют густой растительности. Это также доказывает утверждение о соответствии найденного информативного индекса индексу NDVI.

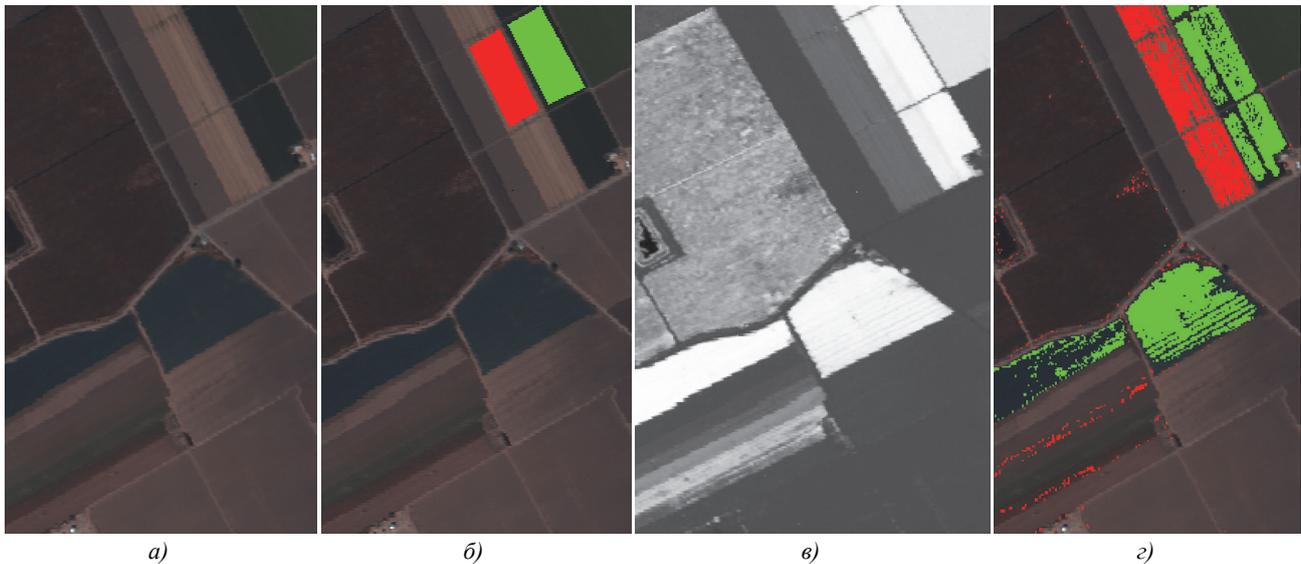


Рис. 1. Исходные данные и визуализация результатов работы алгоритма: визуализация исходного гиперспектра (а); маска классов (б); визуализация INDI (в); результат применения бинаризации двойным порогом (г)

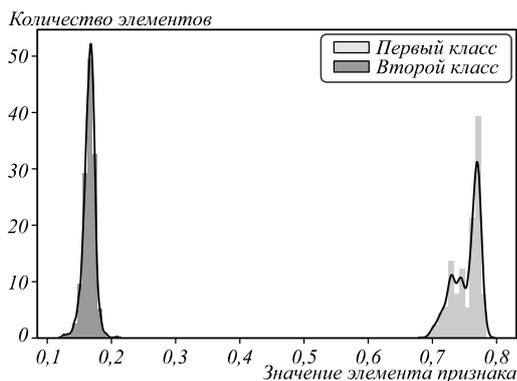


Рис. 2. Распределение классов в информативном признаке

На рис. 3 изображена матрица значений  $\lambda$  для каждой пары слов. Данная матрица симметрична относительно диагонали, поэтому далее будем рассматривать только верхний её треугольник. На рис. 3 мы можем видеть, что наибольшие значения  $\lambda$  находятся рядом или в области, соответствующей индексу NDVI (данная область ограничена зелёными линиями). Действительно, индекс растительности способен лучше всего определять зоны фотосинтетической активности. Следовательно, найденный предложенным алгоритмом индекс является одним из лучших вариантов для разделения таких объектов, как сельдерей и стержень.

### Заключение

В данной статье используется технология формирования информативного индекса для различения заданных объектов гиперспектральных данных. Было показано, что информативные индексы, сформиро-

ванные применяемой технологией, могут быть использованы для создания вычислительно эффективных средств анализа гиперспектральных данных. Благодаря возможности задания произвольных правил для вычисления индекса и оценки информативности, можно адаптировать технологию под любые специфичные задачи и данные.

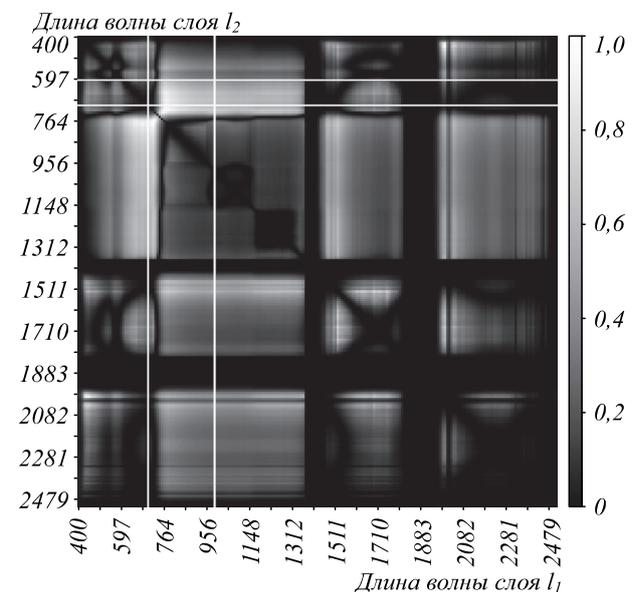


Рис. 3. Значения  $\lambda$  для всех пар слоёв исходного гиперспектра

Выл разработан алгоритм, в котором индекс задается нормализованной разностной формулой, а информативность оценивается на основе значения критерия разделимости дискриминантного анализа. В ра-

боте показано, что сформированный алгоритмом информативный индекс для задачи различения полей соответствует индексу растительности NDVI, что подтверждает практическую значимость применяемой технологии.

Применяемая технология является генерализацией подхода к формированию правил анализа гиперспектральных изображений по малому числу признаков, который, в частности, может быть сведён как к NDVI (при одном варианте задания опорных объектов), так и к выводу принципиально новых индексов.

Интерпретация правила проведения разделяющей границы, сформированного технологией в случае использования разностного нормализованного индекса, позволяет оценить свойства объекта по аналогии с тем, как интерпретируется NDVI. Совместное же использование большого числа индексов, информативных для разных задач, может быть применено для описания характеристик неизвестных объектов гиперспектральных данных.

### Благодарности

Результаты исследования были получены при поддержке государственного задания Минобрнауки России Самарскому университету в рамках работ НИЛ-602 "Фотоника для умного дома и умного города" тема 19в-Р001-602 43/21Б (экспериментальная часть), в рамках проекта № 0777-2020-0017 (программная реализация и разработка технологии), при частичной финансовой поддержке РФФИ в рамках научного проекта № 20-51-05008 (теоретическая часть).

### Литература

1. **Wu, C.** Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation / C. Wu, Z. Niu, Q. Tang, W. Huang // *Agricultural and Forest Meteorology*. – 2008. – Vol. 148, Issues 8-9. – P. 1230-1241.
2. **Cho, M.A.** Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression / M.A. Cho, A. Skidmore, F. Corsi, S.E. Van Wieren, I. Sobhan // *International Journal of Applied Earth Observation and Geoinformation*. – 2007. – Vol. 9, Issue 4. – P. 414-424.
3. **Knyazikhin, Y.** Hyperspectral remote sensing of foliar nitrogen content / Y. Knyazikhin, M.A. Schull, P. Stenberg, M. Möttöus, M. Rautiainen, Y. Yang, R.B. Myneni // *Proceedings of the National Academy of Sciences*. – 2013. – Vol. 110, Issue 3. – P. E185-E192.
4. **Zhang, F.** Estimation of vegetation water content using hyperspectral vegetation indices: A comparison of crop water indicators in response to water stress treatments for summer maize / F. Zhang, G. Zhou // *BMC Ecology*. – 2019. – Vol. 19, Issue 1. – P. 18.
5. **Zhang, X.** Crop classification based on feature band set construction and object-oriented approach using hyperspectral images / X. Zhang, Y. Sun, K. Shang, L. Zhang, S. Wang // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. – 2016. – Vol. 9, Issue 9. – P. 4117-4128.
6. **Boori, M.S.** Supervised and unsupervised classification for obtaining land use/cover classes from hyperspectral and multi-spectral imagery / M.S. Boori, P. Paringer, K. Choudhary, A. Kupriyanov // *Proceedings of SPIE*. – 2018. – Vol. 10773. – P. 107730L. – DOI: 10.1117/12.2322624.
7. **Boori, M.S.** Monitoring crop phenology using NDVI time series from Sentinel 2 satellite data / M.S. Boori, K. Choudhary, R. Paringer, A.K. Sharma, A. Kupriyanov, S. Corgne // 2019 5<sup>th</sup> International Conference on Frontiers of Signal Processing (ICFSP). – 2019. – P. 62-66. – DOI: DOI: 10.1109/ICFSP48124.2019.8938078.
8. **Golhani, K.** A review of neural networks in plant disease detection using hyperspectral data / K. Golhani, S.K. Balasundram, G. Vadamalai, B. Pradhan // *Information Processing in Agriculture*. – 2018. – Vol. 5, Issue 3. – P. 354-371.
9. **Boori, M.S.** Spatiotemporal ecological vulnerability analysis with statistical correlation based on satellite remote sensing in Samara, Russia / M.S. Boori, K. Choudhary, R. Paringer, A. Kupriyanov // *Journal of Environmental Management*. – 2021. – Vol. 285. – P. 112138. – DOI: 10.1016/j.jenvman.2021.112138.
10. **Pandey, P.** High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging / P. Pandey, Y. Ge, V. Stoerger, J.C. Schnable // *Frontiers in Plant Science*. – 2017. – Vol. 8. – P. 1348.
11. **Holmer, A.** Oxygenation and perfusion monitoring with a hyperspectral camera system for chemical based tissue analysis of skin and organs / A. Holmer, F. Tetschke, J. Marotz, H. Malberg, W. Markgraf, C. Thiele, A. Kulcke // *Physiological Measurement*. – 2016. – Vol. 37, Issue 11. – P. 2064-2078.
12. **Zhang, H.** Hyperspectral classification based on lightweight 3-D-CNN with transfer learning / H. Zhang, Y. Li, Y. Jiang, P. Wang, Q. Shen, C. Shen // *IEEE Transactions on Geoscience and Remote Sensing*. – 2019. – Vol. 57, Issue 8. – P. 5813-5828.
13. **Zhu, L.** Generative adversarial networks for hyperspectral image classification / L. Zhu, Y. Chen, P. Ghamisi, J.A. Benediktsson // *IEEE Transactions on Geoscience and Remote Sensing*. – 2018. – Vol. 56, Issue 9. – P. 5046-5063.
14. **Li, J.** Classification of hyperspectral imagery using a new fully convolutional neural network / J. Li, X. Zhao, Y. Li, Q. Du, B. Xi, J. Hu // *IEEE Geoscience and Remote Sensing Letters*. – 2018. – Vol. 15, Issue 2. – P. 292-296.
15. **Deng, J.** Imagenet: A large-scale hierarchical image database / J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, L. Fei-Fei // 2009 IEEE Conference on Computer Vision and Pattern Recognition. – 2009. – P. 248-255.
16. **Lin, T.Y.** Microsoft COCO: Common objects in context / T.Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C.L. Zitnick // *European Conference on Computer Vision*. – 2014. – P. 740-755.
17. **Graña, M.** Hyperspectral remote sensing scenes [Electronical Resource] / M. Graña, M.A. Veganzons, B. Ayerdi. – 2020. – URL: [http://www.ehu.es/ccwintco/index.php?title=Hyperspectral\\_Remote\\_Sensing\\_Scenes](http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes) (request date 18.10.2020).
18. **Asrar, G.Q.** Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat 1 / G.Q. Asrar, M. Fuchs, E.T. Kanemasu, J.L. Hatfield // *Agronomy Journal*. – 1984. – Vol. 76, Issue 2. – P. 300-306.
19. **Fukunaga, K.** Introduction to statistical pattern recognition / K. Fukunaga. – New York, London: Academic Press, 1972. – 369 p.

---

**Сведения об авторах**

**Парингер Рустам Александрович**, 1990 года рождения, доцент кафедры технической кибернетики Самарского национального исследовательского университета имени академика С.П. Королева (Самарский университет), научный сотрудник Института систем обработки изображений РАН – филиала ФНИЦ «Кристаллография и фотоника» РАН. В 2013 году окончил факультет информатики СГАУ. Кандидат технических наук с 2017 года. Круг научных интересов включает интеллектуальный анализ данных, распознавание образов и искусственные нейронные сети. E-mail: [rusparinger@ssau.ru](mailto:rusparinger@ssau.ru).

**Мухин Артем Владимирович**, 1999 года рождения, студент факультета информатики Самарского национального исследовательского университета имени академика С.П. Королева (Самарский университет), старший лаборант научно-исследовательской лаборатории автоматизированных систем научных исследований. Круг научных интересов включает интеллектуальный анализ данных, распознавание образов и искусственные нейронные сети. E-mail: [mukhin.av@ssau.ru](mailto:mukhin.av@ssau.ru).

**Куприянов Александр Викторович**, 1978 года рождения, заведующий кафедрой технической кибернетики Самарского национального исследовательского университета имени академика С.П. Королёва; старший научный сотрудник Института систем обработки изображений РАН – филиала ФНИЦ «Кристаллография и фотоника» РАН. Сфера научных интересов: цифровая обработка сигналов и изображений; распознавание образов и искусственный интеллект; анализ и интерпретация биомедицинских сигналов и изображений. E-mail: [akupr@ssau.ru](mailto:akupr@ssau.ru).

---

ГРНТИ: 28.23.15

Поступила в редакцию 27 мая 2021 г. Окончательный вариант – 8 сентября 2021 г.

---

---

# Formation of an informative index for recognizing specified objects in hyperspectral data

R.A. Paringer<sup>1,2</sup>, A.V. Mukhin<sup>1</sup>, A.V. Kupriyanov<sup>1,2</sup>

<sup>1</sup> Samara National Research University, 443086, Samara, Russia, Moskovskoye Shosse 34,

<sup>2</sup> IPSI RAS – Branch of the FSRC “Crystallography and Photonics” RAS, 443001, Samara, Russia, Molodogvardeyskaya 151

## Abstract

The paper is about the development of an approach which able to create rules for distinguishing between specified objects of hyperspectral data using a small number of observations. Such an approach would contribute to the development of methods and algorithms for the operational analysis of hyperspectral data. These methods can be used for hyperspectral data preprocessing and labeling. Implementation of the proposed approach are using a technology that harnesses both discriminative criteria and the general formulas of spectral indexes. In implementing the proposed technology, the index was defined with normalized difference formula. The Informativeness was estimated using separability criteria of discriminative analysis. The results show that the implemented algorithm can recognize areas of hyperspectral data with different vegetation. The index formed by the algorithm is similar to Normalized Difference Vegetation Index (NDVI). The proposed technology is the generalization of the approach of forming recognition rules using a small number of features. It has been shown that technology can form informative indexes in specified tasks of hyperspectral data analysis.

**Keywords:** classification, hyperspectral data, NDVI, discriminant analysis.

**Citation:** Paringer RA, Mukhin AV, Kupriyanov AV. Formation of an informative index for recognizing specified objects in hyperspectral data. *Computer Optics* 2021; 45(6): 873-878. DOI: 10.18287/2412-6179-CO-930.

**Acknowledgements:** The results of the study were obtained as part of the state program of the Ministry of Education and Science of the Russian Federation to Samara University research laboratory #602 «Photonics for Smart House and Smart City» (experiments), partially funded by project no. 0777-2020-0017 (software development and technology), and partially funded by the Russian Foundation for Basic Research under project # 20-51-05008 (theoretical research).

## References

- [1] Wu C, Niu Z, Tang Q, Huang W. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agric For Meteorol* 2008; 148(8-9): 1230-1241.
  - [2] Cho MA, Skidmore A, Corsi F, Van Wieren SE, Sobhan I. Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *Int J Appl Earth Obs Geoinf* 2007; 9(4): 414-424.
  - [3] Knyazikhin Y, Schull MA, Stenberg P, Möttus M, Rautiainen M, Yang Y, Myneni RB. Hyperspectral remote sensing of foliar nitrogen content. *Proc Natl Acad Sci U S A* 2013; 110(3): E185-E192.
  - [4] Zhang F, Zhou G. Estimation of vegetation water content using hyperspectral vegetation indices: A comparison of crop water indicators in response to water stress treatments for summer maize. *BMC Ecology* 2019; 19(1): 18.
  - [5] Zhang X, Sun Y, Shang K, Zhang L, Wang S. Crop classification based on feature band set construction and object-oriented approach using hyperspectral images. *IEEE J Sel Top Appl Earth Obs Remote Sens* 2016; 9(9): 4117-4128.
  - [6] Boori MS, Paringer R, Choudhary K, Kupriyanov A. Supervised and unsupervised classification for obtaining land use/cover classes from hyperspectral and multi-spectral imagery. *Proc SPIE* 2018; 10773: 107730L. DOI: 10.1117/12.2322624.
  - [7] Boori MS, Choudhary K, Paringer R, Sharma AK, Kupriyanov A, Corgne S. Monitoring crop phenology using NDVI time series from Sentinel 2 satellite data. 2019 5<sup>th</sup> Int Conf on Frontiers of Signal Processing (ICFSP) 2019: 62-66. DOI: 10.1109/ICFSP48124.2019.8938078.
  - [8] Golhani K, Balasundram SK, Vadamalai G, Pradhan B. A review of neural networks in plant disease detection using hyperspectral data. *Inf Process Agric* 2018; 5(3): 354-371.
  - [9] Boori MS, Choudhary K, Paringer R, Kupriyanov A. Spatiotemporal ecological vulnerability analysis with statistical correlation based on satellite remote sensing in Samara, Russia. *J Environ Manage* 2021; 285: 112138. DOI: 10.1016/j.jenvman.2021.112138.
  - [10] Pandey P, Ge Y, Stoerger V, Schnable JC. High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging. *Front Plant Sci* 2017; 8: 1348.
  - [11] Holmer A, Tetschke F, Marotz J, Malberg H, Markgraf W, Thiele C, Kulcke A. Oxygenation and perfusion monitoring with a hyperspectral camera system for chemical based tissue analysis of skin and organs. *Physiol Meas* 2016; 37(11): 2064-2078.
  - [12] Zhang H, Li Y, Jiang Y, Wang P, Shen Q, Shen C. Hyperspectral classification based on lightweight 3-D-CNN with transfer learning. *IEEE Trans Geosci Remote Sens* 2019; 57(8): 5813-5828.
  - [13] Zhu L, Chen Y, Ghamisi P, Benediktsson JA. Generative adversarial networks for hyperspectral image classification. *IEEE Trans Geosci Remote Sens* 2018; 56(9): 5046-5063.
-

- 
- [14] Li J, Zhao X, Li Y, Du Q, Xi B, Hu J. Classification of hyperspectral imagery using a new fully convolutional neural network. *IEEE Geosci Remote Sens Lett* 2018; 15(2): 292-296.
- [15] Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. *IEEE Conf on Computer Vision and Pattern Recognition* 2009: 248-255.
- [16] Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL. Microsoft COCO: Common objects in context. *ECCV* 2014: 740-755.
- [17] Graña M, Veganzons MA, Ayerdi B. Hyperspectral remote sensing scenes. Source: [http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral\\_Remote\\_Sensing\\_Scenes](http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes).
- [18] Asrar GQ, Fuchs M, Kanemasu ET, Hatfield JL. Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat 1. *Agron J* 1984; 76(2): 300-306.
- [19] Fukunaga K. *Introduction to statistical pattern recognition*. New York, London: Academic Press; 1972.
- 

### *Authors' information*

**Rustam Alexandrovich Paringer**, (born 1990), received Master's degree in Applied Mathematics and Informatics from Samara State Aerospace University (2013). He received his PhD in 2017. Associate professor of the Technical Cybernetics department of Samara National Research University and researcher of IPSI RAS – Branch of the FSRC "Crystallography and Photonics". Research interests: data mining, machine learning and artificial neural networks. E-mail: [rusparinger@ssau.ru](mailto:rusparinger@ssau.ru).

**Artem Vladimirovich Mukhin**, (born 1999), student of Samara National Research University. Senior laboratory assistant at the research laboratory of automated research systems of Samara National Research University. Research interests: data mining, machine learning and artificial neural networks. E-mail: [mukhin.av@ssau.ru](mailto:mukhin.av@ssau.ru).

**Alexander Victorovich Kupriyanov**, (born 1978) graduated with honors from Samara State Aerospace University (SSAU) (2001). Candidate's degree in Technical Sciences (2004) and Doctor of Engineering Science (2013). Currently, Head of the Technical Cybernetics department of Samara National Research University, and part-time position as Senior Researcher at the Image Processing Systems Institute of RAS – Branch of the FSRC "Crystallography and Photonics" RAS. Areas of interest: digital signals and image processing, pattern recognition and artificial intelligence, nanoscale image analysis and understanding, biomedical imaging and analysis. More than 90 scientific papers, including 42 published articles and 2 monographs. E-mail: [akupr@ssau.ru](mailto:akupr@ssau.ru).

---

*Received May 27, 2021. The final version – September 8, 2021.*

---