

УДК 528.88.042:528.852.5

**ИСПОЛЬЗОВАНИЕ ПОЛНОГО НАБОРА НОРМАЛИЗОВАННЫХ
МЕЖКАНАЛЬНЫХ ИНДЕКСОВ МНОГОСПЕКТРАЛЬНЫХ
КОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ ПРИ КЛАССИФИКАЦИИ ПОКРЫТИЙ
ЛАНДШАФТА**

Попов М.А., Станкевич С.А., Сахацкий А.И., Козлова А.А.

Научный центр аэрокосмических исследований Земли ИГи НАН Украины

E-mail: pop@casre.kiev.ua, st@casre.kiev.ua, sakhatsky@casre.kiev.ua, ak@casre.kiev.ua

Для повышения качества классификации покрытий ландшафта на многоспектральных космических изображениях за счет инвариантности к аддитивным и мультипликативным искажениям регистрируемых оптических спектров предложено использовать полный набор нормализованных межканальных индексов – отношений, в которых числитель и знаменатель являются соответственно разностью и суммой радиометрических величин всех возможных пар спектральных диапазонов.

Показано, что при известных спектральных сигнатурах классов возможна оптимизация состава нормализованных межканальных индексов, используемых при классификации, по критерию минимума суммарной вероятности ошибки. Использование оптимального набора нормализованных межканальных индексов обеспечивает повышение точности классификации покрытий ландшафта на девятиканальных космических снимках EOS/ASTER на 15–20 %.

Ключевые слова: космические изображения, классификация покрытий ландшафта, нормализованные межканальные индексы.

ВВЕДЕНИЕ

Надежная и универсальная классификация покрытий ландшафта является основой решения большинства тематических задач природопользования и пространственного планирования дистанционными методами. Существующие технологии классификации покрытий ландшафта базируются на автоматизированной интерпретации многоспектральных аэрокосмических изображений по характерным опознавательным признакам классов [1]. При классификации полезная информация определяется физическими полями объектов ландшафта. Различные классы имеют различающиеся спектральные характеристики оптического отражения или собственного излучения. Использование многоспектральных аэрокосмических изображений позволяет более точно восстанавливать спектры классов, что повышает качество классификации [2].

По составу и типам классов различают два рода задач классификации покрытий ландшафтов. В задачах первого рода цель классификации сводится к выявлению и распознаванию компактных объектов сложной формы. К ним относится большинство искусственных сооружений. Основными опознавательными признаками таких объектов являются форма контура и детали. Поэтому автоматизированная классификация для задач первого рода должна проводиться с использованием структурных признаков [3].

К задачам второго рода относится классификация протяженных площадных

объектов неправильной формы, в основном природного происхождения, с приблизительно равномерными статистическими свойствами. Основные опознавательные признаки таких объектов, как правило, содержатся в спектральных оптических характеристиках [4]. Ниже под классификацией покрытий ландшафта подразумевается как раз статистическая классификация многомерных радиометрических полей, которые соответствуют объектам второго рода на многоспектральных аэрокосмических изображениях.

1. ПОДХОД

Основу подхода к классификации многоспектральных аэрокосмических изображений составляет описание оптического сигнала в каждом элементе (пикселе) отдельной дискретной функцией $E(\lambda)$, $\lambda = 1..m$ в многомерном спектральном пространстве $\{\lambda\}$. Очевидно, что средняя точность классификации m -мерных радиометрических полей будет асимптотически возрастать с увеличением количества спектральных отсчетов m [5].

Важной проблемой является экстраполяция ограниченных априорных знаний о сцене многоспектрального изображения на процесс детальной попиксельной классификации. Обычно это производится определением обучающих последовательностей статистических выборок областей классов. Теория классификации требует полноты набора и принципиальной разделимости классов в пространстве признаков, то есть с математической точки зрения распределение плотности вероятности многоспектральных данных должно быть сепарабельной смесью распределений плотностей вероятности классов [6].

Известные алгоритмы автоматизированной классификации многоспектральных аэрокосмических изображений формируют некоторые количественные оценки принадлежности каждого пикселя каждому из рассматриваемых классов. Для классических статистических алгоритмов это вероятности классов или отношения правдоподобия, для других – более или менее обоснованные собственные меры сходства [7]. При дальнейшем применении некоторого решающего правила об окончательном отнесении пикселя к определенному классу эта информация, как правило, теряется. Между тем, ее использование при последующей тематической обработке и интерпретации результатов классификации, могло бы дать дополнительные данные и повысить общую точность [8]. Поэтому, на наш взгляд, конечным результатом классификации покрытий ландшафта должен быть набор пиксельных вероятностных карт всех рассматриваемых классов.

При классификации покрытий ландшафта широко используются всевозможные нормализованные индексы различий спектральных оптических сигналов – такие, как NDVI, SAVI и другие – отношения, в которых числитель и знаменатель являются соответственно разностью и суммой радиометрических величин в пикселях спектральных изображений двух разных спектральных диапазонов [1, 8]. Использование нормализованных межканальных индексов позволяет обеспечить полную или частичную инвариантность спектральных признаков классов к нежелательным аддитивным и мультипликативным искажениям регистрируемых оптических спектров, например, вследствие вариаций спектральных характеристик объектов, состояния атмосферы, изменения условий освещенности и т. д. Практика решения многих тематических задач дистанционного зондирования свидетельствует

о большей информативности нормализованных межканальных индексов по сравнению с “сырыми” спектральными сигналами [9].

Однако, в настоящее время использование нормализованных межканальных индексов основывается на эмпирии, а выбор конкретных индексов для решения той или иной тематической задачи является своего рода искусством. Недостатки такого подхода особенно заметны при решении комплексных задач, в том числе и при классификации покрытий ландшафта. Между тем, для всякого многоспектрального аэрокосмического изображения существует конечное множество – полный набор нормализованных межканальных индексов размерностью $\frac{m \cdot (m-1)}{2}$. Возможности современных систем обработки многоспектральных аэрокосмических изображений позволяют перевести процесс классификации из пространства спектральных оптических сигналов в пространство нормализованных межканальных индексов и тем самым повысить ее качество.

2. МОДЕЛЬ

При классификации покрытий ландшафта на многоспектральных аэрокосмических изображениях соответствующие m -мерные оптические сигналы рассматриваются как нестационарные эргодические случайные процессы с дискретными отсчетами [10].

Классификация покрытий ландшафта на многоспектральном аэрокосмическом изображении может проводиться на основе двух основных моделей – статистической и информационной. В первой используются обобщенные статистические метрики, такие как метрика Чернова или статистическое расстояние Бхаттачария B_j [11]:

$$B_j = -\ln \int \sqrt{p(\lambda) \cdot p_j(\lambda)} d\lambda, \quad (1)$$

где $p(\lambda)$, $p_j(\lambda)$ – плотности вероятности распределений спектральных интенсивностей m -мерных оптических сигналов $E(\lambda)$ и $E_j(\lambda)$ соответственно.

Вторая модель основывается на взаимных информационных метриках статистических распределений $p(\lambda)$ и $p_j(\lambda)$. Чаще всего используется дивергенция Кульбака-Лейблера D_j [12]:

$$D_j = \int p(\lambda) \cdot \log_2 \frac{p(\lambda)}{p_j(\lambda)} d\lambda. \quad (2)$$

Обе модели позволяют оценить сверху вероятности ошибочного отнесения ε_j спектральной сигнатуры текущего пикселя многоспектрального аэрокосмического изображения к любому j -му классу:

$$\varepsilon_j = \sqrt{A_j \cdot (1 - A_j)} \cdot e^{-B_j} - \quad (3)$$

для расстояния Бхаттачария, где A_j – априорная вероятность j -го класса на изображении [11], или

$$\varepsilon_j = \sqrt{A_j \cdot (1 - A_j)} \cdot 2^{-m \cdot D_j} - \quad (4)$$

для дивергенции Кульбака-Лейблера, где m – размерность оптических сигналов $E(\lambda)$ и $E_j(\lambda)$ [13]. Обобщенные метрики (1) и (2) легко конкретизируются для случая

нормальных распределений статистики классов [10].

Если размерность многоспектральных оптических сигналов достаточно велика, то статистические выборки, как правило, будут нерепрезентативными, а оцениваемые по ним распределения плотности вероятности могут быть неточными. Кроме того, с ростом числа спектральных отсчетов классические статистические алгоритмы классификации становятся практически неработоспособными вследствие быстрой сингуляризации ковариационных матриц. В этом случае предпочтительной является информационная модель [14].

После того, как для текущего пикселя найдены вероятности ошибочного отнесения к каждому из s классов, становится возможным по формуле Байеса определить апостериорные вероятности P_k , $k = 1 \dots s$ его принадлежности к этим классам:

$$P_k = \frac{A_k \cdot (1 - \varepsilon_k)}{\sum_{j=1}^s A_j \cdot (1 - \varepsilon_j)} \quad (5)$$

Соотношение (5) является основой для построения пиксельных вероятностных карт классов. Правда, статистические метрики в своих обобщенных формах (1) и (2) не учитывают шумы многоспектральных изображений, что может привести к определенным трудностям при реализации изложенной модели.

3. АЛГОРИТМ

Пусть классификация покрытий ландшафта включает s классов, каждый из которых описывается m -мерной спектральной сигнатурой $E_j(\lambda)$, $\lambda = 1 \dots m$, $j = 1 \dots s$. В рамках изложенного подхода прежде всего строится полный набор нормализованных межканальных индексов $z(\lambda)$, $\lambda = 1 \dots \frac{m \cdot (m-1)}{2}$:

$$z(\lambda+l) = \frac{E(\lambda) - E(l)}{E(\lambda) + E(l)}, \quad \lambda = 1 \dots m-1, \quad l = \lambda+1 \dots m. \quad (6)$$

Далее в соответствии с (2), (4) и (5) уже могут быть построены вероятностные карты классов. Вместе с тем, вполне возможна ситуация, когда полный набор нормализованных межканальных индексов будет иметь существенную информационную избыточность, которая может стать не только бесполезной, но и вредной – за счет “растворения” полезной информации о классах в общем ее шумовом фоне. Поэтому обоснованный отбор информативных нормализованных межканальных индексов многоспектральных аэрокосмических изображений позволит существенно повысить информационную ценность и сократить вычислительные затраты при классификации покрытий ландшафта.

Целью оптимизации состава нормализованных межканальных индексов является обеспечение максимальной в смысле (2) разделимости классов. Рассмотрим вектор выбора нормализованных межканальных индексов $W = \{ W(\lambda) \}$, $\lambda = 1 \dots \frac{m \cdot (m-1)}{2}$, элементы которого равны единице, если соответствующий индекс выбран, или нулю – в противном случае. Полная разделимость классов D будет

зависеть от вектора выбора и эквивалентного отношения “сигнал-шум” для пар классов ψ_{jk} [15] и составит:

$$D(W) = \sum_{j=1}^s \sum_{k=1}^s D_{jk}(W) \cdot tX_k. \quad (7)$$

В этом случае задачей оптимизации состава нормализованных межканальных индексов гиперспектрального аэрокосмического изображения станет определение такого вектора оптимального выбора W^* , который обеспечит максимум целевой функции (7):

$$W^* = \operatorname{argmax} D(W). \quad (8)$$

При известных полных наборах нормализованных межканальных индексов классов это – задача аналитической безусловной дискретной оптимизации. Ее вычислительная сложность экспоненциально зависит от размерности и при числе исходных спектральных диапазонов $m \leq 7..8$ решается прямым перебором. При большем числе спектральных диапазонов оптимизация может быть выполнена методом псевдоградиентного поиска на регулярной решетке в пространстве переменных вектора W :

$$\begin{aligned} W_0 &= \{ 1 \}, \\ W_n &= W_{n-1} + \Delta W_{n-1} \cdot \operatorname{grad} D(W_{n-1}), \\ \operatorname{grad} D(W^*) &\leq 0, \end{aligned} \quad (9)$$

где $\operatorname{grad} D(W)$ – оценка градиента целевой функции $D(W)$, ΔW – скорость изменения координат в выбранном направлении [16].

Порядок проведения классификации покрытий ландшафта с использованием нормализованных межканальных индексов многоспектральных аэрокосмических изображений поясняется схемой рис. 1.

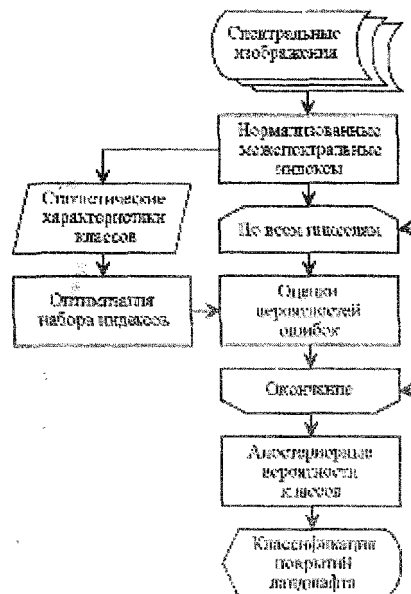


Рис. 1. Алгоритм классификации покрытий ландшафта с использованием нормализованных межканальных индексов многоспектральных аэрокосмических изображений

4. РЕЗУЛЬТАТЫ

Проведено тестовая классификация восьми классов покрытий ландшафта на космическом снимке EOS/ASTER (девять спектральных диапазонов, пространственное разрешение 30м, рис. 2) с использованием традиционной технологии на основе спектральных сигнатур классов, а также с использованием описанного алгоритма на основе как полного набора нормализованных межканальных индексов, так и с предварительной его оптимизацией. Результаты классификации иллюстрируются рис. 3.

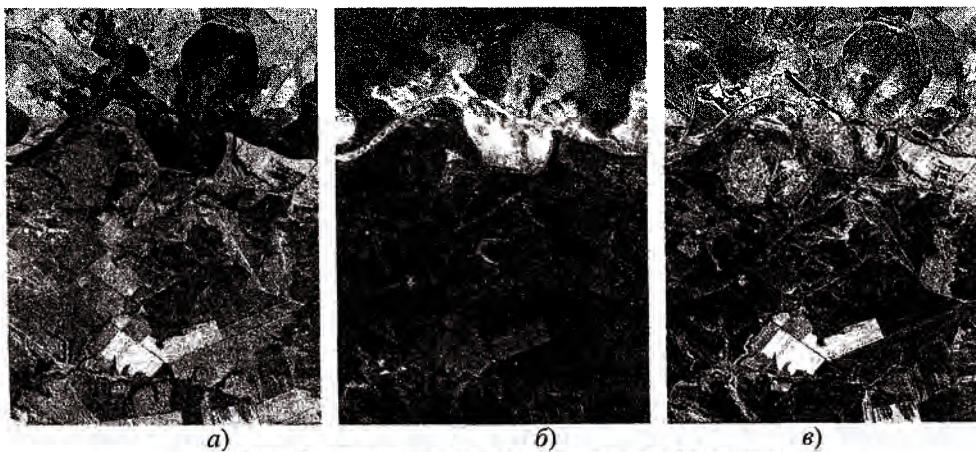


Рис. 2. Исходные космические изображения

а – цветосинтезированный (спектральные каналы 2,209, 0,807 и 0,556 мкм) снимок EOS/ASTER (Украина, 17.04.2004); *б* – нормализованный межканальный индекс 0,556–1,656 мкм (индикатор влагосодержания); *в* – нормализованный межканальный индекс 0,661–2,167 мкм

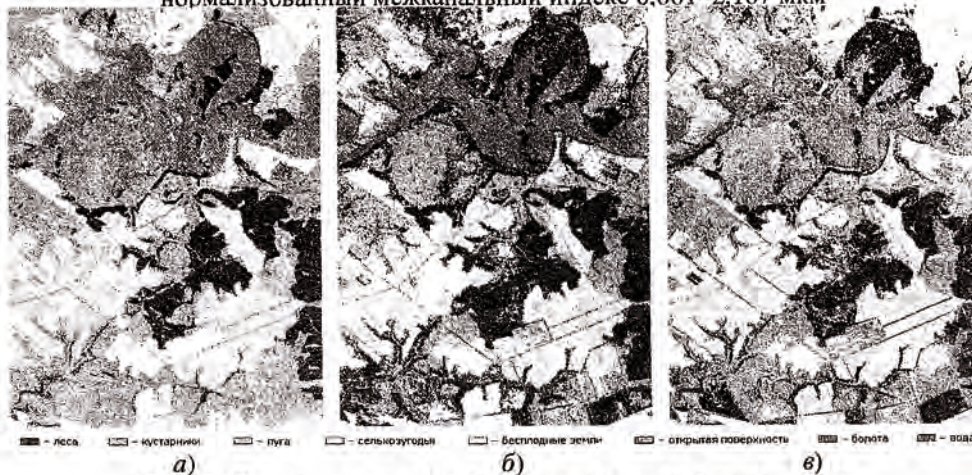


Рис. 3. Результаты классификации покрытий ландшафта

а – классификация покрытий ландшафта на основе спектральных сигнатур;
б – классификация покрытий ландшафта на основе полного множества нормализованных межканальных индексов;
в – классификация покрытий ландшафта на основе оптимального набора нормализованных межканальных индексов

Полный набор включает 36 нормализованных межканальных индексов, оптимизация методом псевдоградиентного поиска (9) позволяет сократить его до 9. В оптимальную комбинацию входят нормализованные межканальные индексы 0,556–0,661 мкм, 0,556–0,807 мкм, 0,556–1,656 мкм, 0,661–0,807 мкм, 0,661–2,167 мкм, 0,661–2,262 мкм, 0,661–2,4 мкм, 0,807–2,336 мкм, 0,807–2,4 мкм, в том числе и приведенные на рис. 2б, в.

Точность классификации оценивалась теоретически по матрицам ошибок классов. Результаты приведены в табл. 1. Приведенные оценки являются ориентировочными, а для получения точных величин обязательно потребуется привлечение данных наземной заверки по тестовым участкам.

Таблица 1.

Сравнительная точность алгоритмов классификации

Алгоритм классификации	Точности классов, %	Средняя точность, %
На основе спектральных сигнатур	48,54 – 88,27	66,54
На основе полного набора нормализованных межканальных индексов	51,44 – 96,67	74,85
На основе оптимального набора нормализованных межканальных индексов	63,61 – 98,85	84,60

Использование полного набора нормализованных межканальных индексов многоспектрального аэрокосмического изображения позволяет повысить среднюю точность классификации покрытий ландшафта на 5–10 %, а оптимизированного набора нормализованных межканальных индексов – на 15–20 %.

ВЫВОДЫ

Таким образом, использование полного набора нормализованных межканальных индексов многоспектральных космических изображений с последующей оптимизацией под состав классов, присутствующих в сцене, позволяет повысить точность и устойчивость классификации покрытий ландшафта и может быть рекомендовано для включения в состав соответствующих перспективных дистанционных технологий.

БЛАГОДАРНОСТИ

Данное исследование выполнено в рамках международного проекта "The Model and Automated Technology for LandCover Classification (TLLC)" при финансовой поддержке The International Association for the Promotion of Co-operation with Scientists from the New Independent States (NIS) of the Former Soviet Union (INTAS), Centre National d'Etudes Spatiales (CNES) и Национального космического агентства Украины (НКАУ) – грант INTAS-CNES-NSAU Ref. No. 06-1000024-9100.

Список литературы

1. *Багатоспектральні методи дистанційного зондування Землі в задачах природокористування* / За ред. В.І. Лялька та М.О. Попова. - Київ: Наукова думка, 2006. - 360 с.
2. *Попов М.О.* Сучасні погляди на інтерпретацію даних аэрокосмічного дистанційного зондування Землі // *Космічна наука і технологія*, 2002. - Т. 8. - № 2/3. - С. 110-115.
3. *Gomes A., Marcal A.R.* Land Cover Revision Through Object Based Supervised Classification of ASTER Data // *Proceedings of the ASPRS 2003 Annual Conference*. - Anchorage, 2003. - P. 45-49.
4. *Кононов В.И., Станкевич С.А.* Сравнительная оценка информативности цифровых аэрокосмических изображений высокого и низкого разрешения // *Ученые записки Таврического национального университета им. В.И. Вернадского*, 2004. - Т. 17 (56). - С. 88-95.

5. *Liang S.* Quantitative Remote Sensing of Land Surfaces. - Hoboken: John Wiley @ Sons, 2003. - 560 p.
6. *Патрик Э.* Основы теории распознавания образов: Пер. с англ. - М.: Сов. радио, 1980. - 408 с.
7. *Попов М.А., Лялько В.И., Подорван В.Н., Сахацкий А.И.* Методика классификации площадных объектов на многоспектральных космических изображениях на основе последовательного слияния информации // *Материалы Второй открытой Всероссийской научной конференции "Современные проблемы дистанционного зондирования Земли из космоса"*. - М.: ИКИ РАН, 2005. - С. 88-94.
8. *Станкевич С.А., Козлова А.О.* Особливості розрахунку індексу видового різноманіття за результатами статистичної класифікації аерокосмічних знімків // *Ученые записки Таврического национального университета им. В.И. Вернадского*, 2006. - Т. 19 (58). - С. 144-150.
9. *Norman J.M., Divakarla M., Goel N.S.* Algorithms for Extracting Information from Remote Thermal-IR Observations of the Earths Surface // *Remote Sensing of Environment*, 1995. - Vol. 51. - No. 1. - P. 157-168.
10. *Landgrebe D.A.* Signal Theory Methods in Multispectral Remote Sensing. - Hoboken: John Wiley @ Sons, 2003. - 512 p.
11. *Фукунага К.* Ведение в статистическую теорию распознавания образов: Пер. с англ. - М.: Наука, 1979. - 368 с.
12. *Chang C.-I.* An Information Theoretic-Based Approach to Spectral Variability, Similarity, and Discriminability for Hyperspectral Image Analysis // *IEEE Transactions on Information Theory*, 2000. - Vol. 46. - P. 1927-1932.
13. *Georgiou T.T., Lindquist A.* Kullback-Leibler Approximation of Spectral Density Functions // *IEEE Transactions on Information Theory*, 2003. - Vol. 49. - P. 2910-2917.
14. *Станкевич С.А.* Кількісне оцінювання інформативності гіперспектральних аерокосмічних знімків при вирішенні тематичних задач дистанційного зондування Землі // *Доповіді НАН України*, 2006. - № 10. - С. 136-139.
15. *Станкевич С.А.* Оптимизация состава спектральных каналов гиперспектральных аерокосмических изображений при решении тематических задач дистанционного зондирования Земли // *Космическая наука и технология*, 2007. - Т. 13. - № 2. - С. 25-28.
16. *Попов М.А., Станкевич С.А.* Методы оптимизации числа спектральных каналов в задачах обработки и анализа данных дистанционного зондирования Земли // *Материалы Четвертой открытой Всероссийской научной конференции "Современные проблемы дистанционного зондирования Земли из космоса"*. - М.: ИКИ РАН, 2006. - Т. 1. - С. 106-112.

Попов М.О., Станкевич С.А., Сахацкий О.И., Козлова А.О. Використання повного набору нормалізованих міжканальних індексів багатоспектральних космічних зображень при класифікації покриття ландшафту // *Вчені записки Таврійського національного університету ім. В. І. Вернадського*. – 2007. – Серія «Географія». - Т. 20 (59).- № 1. - С. 175-182.

Для підвищення якості класифікації покриття ландшафту на багатоспектральних космічних зображеннях за рахунок інваріантності до адитивних і мультиплікативних викривлень оптичних спектрів, що реєструються, запропоновано використовувати повний набір нормалізованих міжканальних індексів – відношень, в яких чисельник і знаменник є відповідно різницею і сумою радіометричних величин усіх можливих пар спектральних діапазонів.

Показано, що за відомих спектральних сигнатурах класів можлива оптимізація складу нормалізованих міжканальних індексів, які використовуються при класифікації за критерієм мінімуму сумарної ймовірності помилки. Використання оптимального набору нормалізованих міжканальних індексів забезпечує підвищення точності класифікації покриття ландшафту на дев'ятиканальних космічних знімках EOS/ASTER на 15–20 %.

Ключові слова: космічні зображення, класифікація покриття ландшафту, нормалізовані міжканальні індекси.

Popov M.A., Stankevich S.A., Sakhatsky A.I., Kozlova A.A. Use of a full set of the normalized band-difference indexes for land cover classification on multispectral space imagery // *Uchenye zapiski Tavricheskogo Natsionalnogo Universiteta im. V.I. Vernadskogo*. – 2007. – Series «Geography». – V. 20 (59). - № 1. – С. 175-182.

For quality of land cover classification on multispectral space imagery improvement due to invariance against additive and multiplicative distortions of registered optical spectra is offered to use a full set of the normalized band-difference indexes – the ratios, in which numerator is a difference and denominator is a sum of appropriate radiometric magnitudes for all possible pairs of spectral bands. If classes spectral signatures are known then optimization of the normalized band-difference indexes set used for classification is possible by minimum of error summary probability criteria.

The use of optimal set of the normalized band-difference indexes increases land cover classification accuracy on the nine-bands EOS/ASTER space imagery to 15–20 %.

Keywords: space imagery, land cover classification, normalized band-difference indexes.

Поступила в редакцію 16.04.2007г.