

## МЕТОДИ ВІДБОРУ НАЙБІЛЬШ ІНФОРМАТИВНИХ СПЕКТРАЛЬНИХ КАНАЛІВ

Альперт С.І.

*Науковий Центр аерокосмічних досліджень Землі*

*ІГН НАН України, Київ, Україна*

*E-mail: sofia52010@yandex.ru*

### **Methods for selection of the most informative spectral bands**

In this work we analyzed different methods for selection of the most informative and most useful spectral bands. It was considered Fisher criterion, its modifications and criterion function of information capability of spectral bands. It was also proposed a new method to select spectral bands, which is based on the concept of new modified criterion function of information capability of spectral bands. It was shown, that the proposed new modified criterion function of information capability of spectral bands has advantages and gives more accurate results than other methods.

Як відомо, гіперспектральні супутникові зображення складаються з великої кількості зональних зображень, що, в свою чергу, є досить складним для обробки, тому актуальною є процедура відбору найбільш інформативних спектральних каналів. В даній роботі розглядаються різні критерії для оцінки класифікаційної цінності спектральних зображень, такі як: критерій Фішера, його модифікації та критеріальна функція інформативності. Також наводиться нова критеріальна функція інформативності, яка враховує об'єм об'єктів кожного класу, і дає більш точну оцінку інформативності спектральних каналів.

Лінійний дискримінантний аналіз (Linear Discriminant Analysis, LDA) - це одна з найбільш ефективних методик, що оптимізує використання критерію Фішера [1]. Дана методика використовується для розпізнавання та класифікації різних видів дерев хвойних та тропічних лісів, для ідентифікації одиниць земляного покриття, при розв'язанні екологічних задач, тощо.

Розглянемо множину з  $n$  ділянок,  $\{x_i, y_i\}_{i=1}^n$ , де  $x_i \in R^m$  являє собою вектор характеристик розмірності  $m$  для кожного  $i$ -го пікселя із позначенням  $y_i \in \Omega$ , де  $m$  - кількість спектральних каналів,  $\Omega$  - простір класів, які присутні на зображенні.

Метод класифікації LDA шукає таке лінійне перетворення (матрицю  $G$ ), яке понижує розмірність вектора характеристик  $x$ , оскільки цей метод не може працювати, коли розмірність простору характеристик набагато більша за кількість навчальних ділянок. Тобто ми перетворюємо вектор характеристик  $x$  розмірності  $m$  на вектор характеристик розмірності  $l$ ,  $a = G^T x \in R^l$ , де  $l < m$ . Ми понижуємо розмірність простору характеристик для того, щоб максимізувати відстань між класами.

Критерій Фішера полягає у максимізації відстані між середніми значеннями класів та мінімізації внутрішньокласових дисперсій за допомогою наступного функціоналу:

$$J(w) = (\mu_2 - \mu_1)^2 / (\delta_2^2 + \delta_1^2). \quad (1)$$

Розглянемо методику відбору спектральних каналів, що заснована на застосуванні критеріальної функції інформативності [2]:

$$F = 1 - \frac{1}{M(M-1)} \cdot \sum_{m=1}^M \frac{\sum_{j=1}^{nz} \left( l_{mj} \cdot \sum_{\substack{k=1 \\ k \neq m}}^M l_{kj} \right)}{\sum_{j=1}^{nz} l_{mj}}, \quad (2)$$

де  $M$  – кількість класів,  $nz$  – число інтервалів для ознаки спектрального каналу, що розглядається,  $l_{mj}$  – індикатор зіставлення,

$$l_{mj} = \begin{cases} 1, & \text{якщо об'єкт класу } K_m \text{ пов'язаний з інтервалом } j; \\ 0, & \text{в інших випадках.} \end{cases}$$

Область значень функції  $F$  лежить у діапазоні від 0 до 1. При цьому нуль відповідає випадку, коли об'єкти всіх класів зосереджені всередині одного інтервалу, а одиниця - випадку, коли об'єкти кожного із класів виявляються розподіленими чітко за відповідними власними інтервалами. В останньому випадку класифікаційна цінність спектрального каналу максимальна, тоді як у першому випадку такий спектральний канал ніякої цінності як класифікатор не має.

Розглянемо нову критеріальну функцію інформативності, яка враховує кількість об'єктів кожного класу.

Для кожного інтервалу  $j$  обчислюємо відсоток (долю) “неправильних” пікселів, тобто тих пікселів, що помилково потрапили в даний інтервал. Для цього, кількість “неправильних” пікселів ділимо на загальну кількість пікселів, що знаходяться в інтервалі. Отримуємо  $S_j$  – відношення кількості “неправильних” пікселів до загальної кількості пікселів в інтервалі  $j$ . Після того, як ми знайдемо значення всіх величин  $S_j$  для кожного інтервалу, нам треба їх просумувати та поділити на загальну кількість інтервалів.

Нова модифікована критеріальна функція інформативності буде мати вигляд:

$$F^* = 1 - \left( \frac{S_1 + S_2 + S_3 + \dots + S_j}{J} \right). \quad (3)$$

Дана функція інформативності приймає значення в діапазоні від 0 до 1. Чим ближче значення функції  $F^*$  буде до 1, тим більш інформативним буде даний спектральний канал.

Також формулу (3) ми можемо записати у іншому вигляді, через індикатори:

$$F^* = 1 - \frac{1}{J} (S_1 + S_2 + S_3 + \dots + S_j) = 1 - \frac{1}{J} \sum_{j=1}^J \left( \frac{\sum_{\substack{k=1 \\ k \neq j}}^M l_{kj} n_{kj}}{\sum_{k=1}^M l_{kj} n_{kj}} \right), \quad (4)$$

де  $M$  – кількість класів,  $J$  – число інтервалів для ознаки спектрального каналу, що розглядається,  $l_{kj}$  – індикатор зіставлення,

$$l_{kj} = \begin{cases} 1, & \text{якщо об'єкт класу } K_k \text{ пов'язаний з інтервалом } j, \\ 0, & \text{в інших випадках.} \end{cases}$$

$n_{kj}$  – кількість елементів з  $k$ -го класу, що потрапили в інтервал  $j$ .

Дана функція інформативності приймає значення в діапазоні від 0 до 1. Чим ближче значення функції  $F^*$  буде до 1, тим більш інформативним буде даний спектральний канал.

В даній роботі аналізується використання гіперспектральних зображень при розв'язанні тематичних задач ДЗЗ, зокрема, задач класифікації. Акцентується на тому, що більша частина зональних зображень, що складає гіперспектральне зображення, є зашумленими та несуть недостатньо корисної інформації. Тому була поставлена задача, яка полягає у виборі найбільш корисних спектральних каналів.

Було розглянуто різні методи відбору найбільш інформативних спектральних каналів, такі як: критерій Фішера, його модифікації та критеріальна функція інформативності. В статті була також наведена нова модифікована критеріальна функція інформативності, яка, на відміну від інших критеріїв, враховує кількість об'єктів кожного класу та дає більш точну оцінку інформативності спектральних каналів.

Експериментальна перевірка запропонованої модифікованої критеріальної функції інформативності доводить її ефективність та надійність при розв'язанні задач класифікації [3].

## Література

1. Bandos T.V. Classification of Hyperspectral Images with Regularized Linear Discriminant Analysis / T.V. Bandos, L. Bruzzone, G. Camps-Valls // IEEE Transactions on Geoscience and Remote Sensing, Vol. 47, №3, 2009, 862-873 pp.
2. Попов М. Методы оптимизации числа спектральных каналов в задачах обработки и анализа данных дистанционного зондирования Земли / М. Попов, С. Станкевич // Современные проблемы дистанционного зондирования земли из космоса. М.: ИКИ РАН.-2006.- Т. 2, №1.- С. 61-63.
3. Mc Coy R.M. Fields Methods in Remote Sensing / R.M. Mc Coy // New York: Guilford Press, pp. 150-160, 2005.