

УДК 528.77

## Порівняльний аналіз методів класифікації лісів гірської місцевості за матеріалами знімання супутника RapidEye

А.В. Шпак

Національний авіаційний університет

*В статті розглядаються сучасні методи контрольованої класифікації лісів з використанням багатоспектральних зображень супутникової системи RapidEye. Проведено порівняльний аналіз класифікації методом мінімальних відстаней та відстаней Махаланобіса. Встановлено, що для класифікації лісів ефективнішим є метод відстаней Махаланобіса.*

*СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ КЛАССИФИКАЦИИ ЛЕСОВ ГОРНОЙ МЕСТНОСТИ ПО МАТЕРИАЛАМ СЪЕМОК СПУТНИКА RAPIDEYE, Шпак А.В. — В статье рассматриваются современные методы контролируемой классификации лесов с использованием многоспектральных снимков спутниковой системы RapidEye. Проведен сравнительный анализ классификации методом минимальных расстояний и расстояний Махаланобиса. Установлено, что для классификации лесов более эффективным есть метод расстояний Махаланобиса.*

*COMPARATIVE ANALYSIS OF SUPERVISED METHODS FOR THE CLASSIFICATION OF MOUNTAIN FORESTS BASED ON SPACE IMAGERY FROM THE SATELLITE RAPIDEYE, by Shpak A.V. — This article is consider the current methods of supervised forest classification using multispectral RapidEye satellite imagery. The comparative analysis of Mahalanobis distance classification and minimum distance classification was made. Found that method of Mahalanobis distances is more efficient for forest classification.*

**Ключевые слова:** дистанционное зондирование; классификация лесов; контролируемые методы классификации; метод минимальных расстояний; метод расстояний Махаланобиса; снимки RapidEye; лесное хозяйство.

**Key words:** remote sensing; forest classification; supervised classification methods; minimum distance; Mahalanobis distance; RapidEye images; forestry.

### 1. ПОСТАНОВКА ПРОБЛЕМИ

За останні десятиліття спостерігається стрімкий розвиток комп'ютерних систем, засобів вимірювальної техніки і систем дистанційного зондування Землі. Сукупність досягнень в даних галузях науки і техніки дає можливість отримувати інформацію про навколишнє середовище, в тому числі про характеристики і стан лісових масивів.

На сьогоднішній день існує велика кількість методів дешифрування і класифікації рослинності, а також програмних засобів, за допомогою яких можна проводити класифікацію рослинності та вивчення характеристик лісів. Одними з найвідоміших програмних засобів, що використовуються для обробки космічних зображень, в тому числі й для класифікації рослинності, є автоматизовані програмні комплекси, такі як ERDAS Imagine та ENVI.

Використання аерокосмічних знімків в поєднанні з програмними продуктами по обробці зображень, а також програмними продуктами ГІС дає змогу вирішувати наступні задачі моніторингу лісів [5]:

- уточнення видового складу лісової рослинності;
- виявлення несанкціонованих вирубок;
- контроль фітосанітарного стану лісів;
- моніторинг лісових пожеж та оцінка пожежонебезпечності території;
- виявлення вітровалів;
- контроль та планування відновлення лісів з метою запобігання паводків, селі та ерозійних процесів;

Методи обробки і класифікації супутникових знімків залежать як від сенсора, яким було отримано супутниковий знімок, так і від характеристик досліджуваної місцевості.

### 2. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Класифікація є одним з способів тематичної обробки космічних знімків, в процесі якого відбувається дешифрування і розпізнавання об'єктів і явищ на космічному знімку. Сьогодні класифікація космічних знімків включає в себе комп'ютерне дешифрування отриманих космічних зображень або процес автоматизованого поділу всіх пікселів знімка на класи, які відповідають різним об'єктам.

Процес розподілу пікселів на класи здійснюється на основі значень спектральних характеристик. Здаються критерії, згідно яких проводиться поділ пікселів на класи; якщо піксел задовольняє певний набір умов, то його відносять до класу, що відповідає заданому критерію. Яскравість кожного пікселу залежить не тільки від об'єкту, який відповідає даному пікселу, а також від умов знімання. Основні труднощі класифікації полягають в правильному врахуванні всіх умов знімання. Для проведення класифікації лісів гірської місцевості слід враховувати наступні параметри: час доби, пору року, рельєф місцевості, погодні умови та ін. За значенням спектральної яскравості пікселі можна одночасно віднести до декількох класів об'єктів, в такому випадку класифікація є невизначеною та неточною. Зазвичай дана ситуація викликана тим, що піксель відповідає відразу декільком різним об'єктам, відбивна здатність яких відрізняється. Даній ситуації можна уникнути декількома шляхами: 1) збільшити кількість класів об'єктів — при цьому значно збільшиться час і вартість класифікації; 2) після проведення комп'ютерної класифікації знімків проводити порівняння з наземними даними, вибираючи тестові проблемні ділянки території.

Розробкою нових методів класифікації дистанційних зображень і апробацією вже існуючих методів займається багато вітчизняних та зарубіжних вчених. Серед вітчизняних вчених, які займаються вивченням лісів за допомогою дистанційних методів, є Лялько В.І., Попов М.О., Бурштинська Х.В., Часковський О.Г. та ін.

По характеру реалізації алгоритми класифікації поділяють на два типи: алгоритми контрольованої і неконтрольованої класифікації.

Контрольована класифікація або так звана класифікація з навчанням — це процес, при якому відбувається порівняння значень яскравості кожного пікселу з еталонами (тестовими ділянками), в результаті чого, кожен піксел відноситься до класу об'єктів, який найбільше відповідає заданим критеріям. Набір еталонних пікселів — це правила, згідно яких проводиться класифікація всього зображення; дані правила вводить фахівець-дешифрувальник, а потім їх автоматично застосовують для всього зображення.

Для контрольованої класифікації важливо знати наперед, які об'єкти є на знімку, кількість класів, що виокремлюються, і щоб ці класи чітко відрізнялись на знімку. Загалом потрібно, щоб кількість класів об'єктів не перевищувала 30. Також потрібно мати критерії і способи розпізнавання класів за пікселями.

Контрольована класифікація здійснюється за наступними етапами:

- 1) визначення задач обробки зображення і вибір способу класифікації;
- 2) вибір еталонних ділянок;
- 3) проведення класифікації і оцінка якості результатів.

Неконтрольована класифікація, або так звана класифікація без навчання — це процес, при якому розподіл пікселів зображення відбувається автоматично на основі аналізу статистичного розподілу яскравості пікселів [3].

Варто відмітити, що перед початком класифікації невідомо, скільки і які об'єктів є на знімку, а після проведення класифікації необхідно провести дешифрування отриманих класів, щоб визначити, яким об'єктам вони відповідають. Також неконтрольовану класифікацію використовують як попередній етап перед контрольованою класифікацією.

### 3. ПОСТАНОВКА ЗАВДАННЯ

Завдання цієї роботи — порівняти точність і ефективність використання методів контрольованої класифікації лісів гірської місцевості. Вихідним матеріалом для класифікації лісів слугував космічний знімок, отриманий із супутника RapidEye. Територія досліджень розташована в Надвірнянському районі Івано-Франківської області поблизу села Бистриця.

### 4. ТЕОРЕТИЧНИЙ МАТЕРІАЛ

Розпізнавання типів підстиляючої поверхні для видового складу лісової рослинності на багатоспектральних знімках середнього і високого просторового розрізнення здійснюють автоматично за допомогою сучасних комерційних програмних засобів (в даній роботі використовувався програмний продукт ENVI), використовуючи алгоритми контрольованої або неконтрольованої класифікації. В даній роботі випробувано два алгоритми контрольованої класифікації.

Мірою близькості між спектральними характеристиками пікселів та середніми значеннями яскравості еталонів може бути евклідова відстань між векторами яскравості  $ij$ -го пікселу зображення  $(\bar{B}_{ij} = B_{ij}^1, B_{ij}^2, \dots, B_{ij}^L)$  та вектором середніх еталонних значень яскравості спектральних зон, яка для об'єктів класу  $m$  розраховується за формулою:

$$\bar{R}_m \bar{M}_m \bar{B}_{ij} = \sqrt{\sum_{l=1}^L (M_m^l - B_{ij}^l)^2},$$

де  $\bar{M}_m = (M_1, M_2, \dots, M_L)$  — вектор середніх значень яскравості  $m$ -го класу об'єктів, що має розмірність  $L \times 1$  ( $L$  — кількість спектральних зон космічного знімку).

Евклідова відстань розраховується для всіх можливих класів  $M_{kl}$ , і піксел приписується до класу з найменшим значенням  $\bar{R}_m, \bar{M}_m, \bar{B}_{ij}$ . Так як для кожного пікселя існує найближчий еталон, то всі пікселі класифікуються. Такий результат не зовсім коректний, тому іноді застосовують порогові значення спектральних відстаней.

При розрахунку евклідової відстані не враховується мінливість (дисперсія) яскравості пікселів у класі, тому частина пікселів територій з великою дисперсією яскравості можуть бути пропущеними, а до території з низькою мінливістю може бути віднесено значно більше пікселів, ніж їй належить.

Використання відстані Махаланобіса є значно ефективнішим для відображення міри близькості між спектральними характеристиками пікселів та середніми значеннями яскравості еталонів. Формула розрахунку відстані Махаланобіса (квадратична форма) крім середнього значення містить коефіцієнти коваріаційної матриці:

$$D_M = (\bar{B} - \bar{M}_m)^T \cdot \bar{K}_m^{-1} \cdot (\bar{B} - \bar{M}_m),$$

де  $\bar{K}_m^{-1}$  — обернена коваріаційна матриця (розміром  $L \times L$ ) яскравості пікселів зображення для  $m$ -го класу об'єктів;

$T$  — знак операції транспонування матриці;

$\bar{B} = (B^1, B^2, \dots, B^L)$  — вектор розміром  $L \times 1$  значень яскравості пікселя  $L$  спектральних зон;

$\bar{M}_m$  — вектор середніх значень яскравості розміром  $L \times 1$  для  $m$ -го класу об'єктів, що зображені на космічному знімку.

Коваріаційна матриця  $\bar{K}_m$  характеризує ступінь статистичного зв'язку між яскравостями пікселів зображення різних спектральних зон  $m$ -го класу об'єктів. Її  $kl$ -ий елемент (коваріація яскравості  $k$ -го й  $l$ -го спектральних каналів)  $m$ -го класу об'єктів, обчислюються за еталонними вибірками згідно формули:

$$K_{mkl} = \frac{1}{n_m - 1} \sum_{j=1}^{n_m} (B_{jk} - M_{mk}) \cdot (B_{jl} - M_{ml})^T,$$

де  $n_m$  — кількість пікселів зображення еталону  $m$ -го класу об'єктів;

$B_{jk}, B_{jl}$  — яскравість  $j$ -го пікселя зображення  $k$ -го (та  $l$ -го) спектрального каналу  $m$ -го класу об'єктів;

$M_{mk}, M_{ml}$  — середнє значення яскравості пікселів зображення еталону  $m$ -го класу об'єктів  $k$ -го (та  $l$ -го) спектрального каналу, яке розраховується за формулою:

$$M_{mk(l)} = \frac{1}{n_m} \sum_{j=1}^{n_m} B_{jk(l)}.$$

Кількість пікселів навчальної вибірки для статистичних методів класифікації теоретично має бути не менше  $(L + 1)$ , в іншому випадку коваріаційна матриця буде виродженою. Практично для отримання надійних оцінок параметрів класів об'єктів розпізнавання необхідно не менше  $10L$  пікселів еталонних зображень кожного класу об'єктів [2, 7].

Врахування дисперсії і коваріації спектральних ознак дозволяє класифікувати пікселі, що утворюють на зображенні області з великою частотою змін значень яскравості, відносячи їх до класу з подібними характеристиками. Недоліком алгоритму Махаланобіса є більші витрати машинного часу, ніж методу евклідової відстані, до того ж він реалізує статистичний підхід за припущенням нормального розподілу даних у кожній спектральній зоні.

## 5. ХАРАКТЕРИСТИКА ВИХІДНИХ МАТЕРІАЛІВ ТА ТЕРИТОРІЇ ДОСЛІДЖЕННЯ

Вихідними матеріалами для класифікації лісів Надвірнянського району Івано-Франківської області слугував космічний знімок, отриманий оптико-електронною знімальною системою із супутника RapidEye. Система дає зображення в п'яти спектральних діапазонах (блакитний — 440–510 нм; зелений — 520–590 нм; червоний — 630–690 нм; крайній червоний — 690–730 нм; ближній ІЧ — 760–890 нм) [6]. Роздільна здатність системи становить 5 метрів, знімання виконане в вересні 2009 року.

Ділянка дослідження розташована в Надвірнянському районі Івано-Франківської області поблизу с. Бистриця. Ліси на ділянці хвойні та змішані. Територія ділянки гірська, максимальні перепади висот 600 метрів. Для врахування особливостей рельєфу використовувалась цифрова модель рельєфу (ЦМР) з точністю по висоті 30 м.

## 6. ВЕРИФІКАЦІЙНІ ДАНІ

Для проведення контрольованої класифікації необхідно використовувати інформацію, яка дає змогу точніше визначити еталони навчальних вибірок. Звірка інформації виконувалась двома шляхами:

1) При виділенні еталонних ділянок використовувались різні комбінації спектральних каналів зображення, в залежності від характеристик об'єктів еталонного класу. Кожна еталонна ділянка порівнювалась з супутниковими знімками з системи Google Earth.

Таблиця 1

	Загальна характеристика ділянки (рельєф, тип лісу)	Відсоток породи лісу, що переважає	Експозиція схилу	Середня висота над рівнем моря, м
Ділянка-1	переважає хвоя	80% — хвоя, 20% — листяні ліси	схід	945
Ділянка-2	змішані	60% — хвоя, 40% — листяні	схід	997
Ділянка-3	переважає хвоя, в пн.сх. частині змішані	80% — хвоя	пд.зх.	1004
Ділянка-4	переважає хвоя	90% — хвоя	пн.сх.	892
Ділянка-5	рідколісся	40% — хвоя	пд.зх.	1054
Ділянка-6	рідколісся, ґрунтова дорога, змішаний ліс	—	пд.зх.	899
Ділянка-7	50% — рідколісся та чагарники	50% — вкрито хвойними лісами	більша частина на південь	876
Ділянка-8	хвойний, з западиною по серед ділянки	—	більша частина на захід	906
Ділянка-9	листяний	—	пд.сх та схід	1042
Ділянка-10	по середині балка, по контурах з листяним лісом	50–60% — хвойний ліс	більша частина на пн.сх.	989
Ділянка-11	рідколісся, переважно хвоя, з невеликою галювиною	—	північ, схід та пд.сх.	875
Ділянка-12	в пд.зх. Частині ділянки хвойний ліс — 35%, все інше рідколісся і дорога	—	пн.сх., пд.сх. та пд.зх.	829

Таблиця 2. Результат класифікації по виділених еталонних класах

Клас	Частка згідно методу евклідових відстаней, %	Частка згідно методу відстаней Махаланобіса, %
Хвойна рослинність	54,22	57,00
Змішані ліси	11,23	17,9
Трава	6,12	4,12
Вода	1,12	1,02
Рідколісся, чагарники	23,72	15,99
Ґрунти, ґрунтові дороги	1,97	2,36
Забудова, антропогенні об'єкти	1,62	1,60

2) Для отримання задвіркової інформації на досліджувану територію, проведено польові дослідження, з поділом знімка на ділянки і описом кожної з ділянок (табл. 1). Загалом виділено 12 ділянок з різними просторовими і рослинними характеристиками (рис. 1).

Кожна з ділянок описувалась за наступними критеріями (табл. 1):

- загальна характеристика ділянки (рельєф, тип лісу);
- відсоток породи лісу, що переважає;
- експозиція схилу;
- середня висота над рівнем моря.

## 7. РЕЗУЛЬТАТИ РОБОТИ

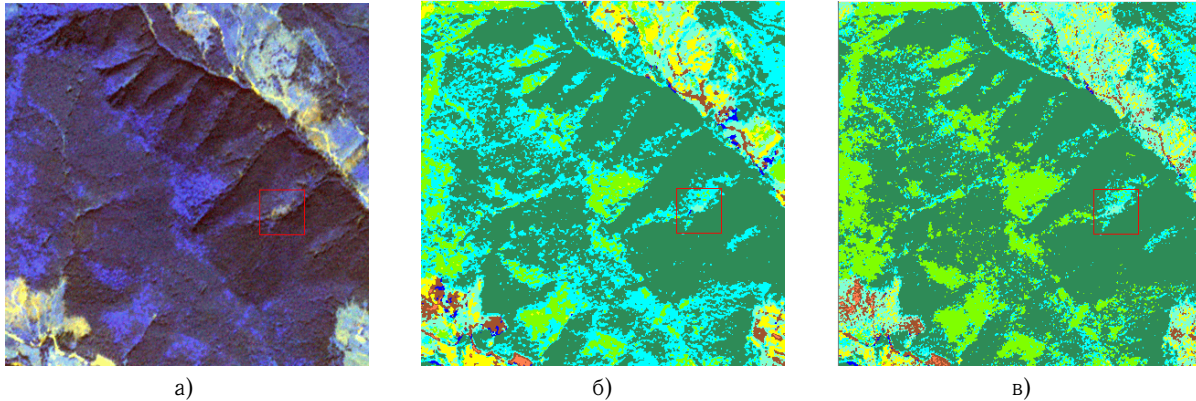
В результаті виконання роботи було виділено 7 класів об'єктів, для яких виділено еталонні ділянки (табл. 2). На основі створених еталонних ділянок, було проведено класифікацію з навчанням за двома методами — мінімальною відстані та відстані Махаланобіса (рис. 2б, 2в).

Отримані результати класифікації (табл. 2) показують, що територія населених пунктів, доріг та водойм мають значні похибки. Це зумовлено в першу чергу характеристиками місцевості (по-перше, в вересні місяці русла гірських річок мілкі і в багатьох місцях проглядається каміння; по берегах річок ростуть дерева і чагарники, які в деяких місцях перекривають все русло річки; по-друге, мережа доріг — ґрунтова з гравієм і за своїми відбивними характеристиками подекуди схожа на русла річок або ділянки з вигорілою травою). Для досягнення кращих результатів класифікації населених пунктів варто використати більшу кількість еталонних ділянок, а також провести звірку з наземними даними.

Щодо отриманих результатів класифікації рослинності, то вони є задовільними. При обох методах класифікації добре виділяється клас з хвойними насадженнями. Натомість при класифікації методом евклідових відстаней значна частина ділянок з змішаним лісом розпізналась як чагарникова рослинність, рідколісся. Це зумовлено тим, що при розрахунку евклідової відстані не враховується мінливість (дисперсія) яскравості пікселів у класі, тому частина пікселів з низькою мінливістю може бути віднесена



**Рис. 1.** Зображення ЦМР з тестовими ділянками



**Рис. 2.** а) — зображення досліджуваної ділянки місцевості знімок RapidEye (комбінація 1, 2, 5 каналів); б) — класифікація методом евклідових відстаней; в) — класифікація методом Махаланобіса.

до одного класу.

## 8. ВИСНОВКИ

1. За допомогою супутникової системи RapidEye можна визначати багато характеристик лісу, зокрема проводити класифікацію ліс по видах рослинності.

2. Вибір методу класифікації дистанційного зображення залежить від типу об'єктів, які класифікуються, вихідних даних, особливостей місцевості та звіркової інформації. Для методів контрольованої класифікації необхідною умовою є попереднє створення еталонних ділянок. Еталонні ділянки слід створювати для об'єктів з чіткою спектральною характеристикою.

3. Неконтрольовану класифікацію використовують для визначення класів об'єктів або як попередній етап перед контрольованою класифікацією. Після проведення неконтрольованої класифікації необхідно провести дешифрування отриманих класів, щоб визначити, яким об'єктам вони відповідають.

4. Порівняльний аналіз класифікації методом евклідової відстані і відстані Махаланобіса з використанням верифікаційних даних показав, що метод відстані Махаланобіса є точнішим для класифікації об'єктів з незначними спектральними відмінностями. Порівнявши інформацію з наземними даними по 12 тестових ділянках, можна сказати, що отриманий результат класифікації на 70–80% відповідає дійсності.

5. Результати класифікації можна покращити шляхом: а) збору більш детальних наземних даних, які слід враховувати при процесі класифікації; б) шляхом детальнішого створення еталонних ділянок, для чого потрібно затратити більше часу.

1. Багатоспектральні методи дистанційного зондування Землі в задачах природокористування / За ред. В.І.Лялько та М.О. Попова. — К.: Наукова думка, 2006. — 369 с.
2. Шовенгерд Р.А. Дистанционное зондирование. Модели и методы обработки изображений. — М.: Техносфера, 2010. — 560 с.
3. Лабутина И.А. Дешифрирование аэрокосмических снимков. Учебное пособие. — М.: Аспект Пресс, 2004. — 184 с.
4. Кашкин В.Б. Дистанционное зондирование Земли из космоса. Цифровая обработка изображений. Учебное пособие. — М.: Логос, 2001. — 264 с.
5. Чухланцев А.А., Саворский В.П. Задачи ДЗ лесов при изучении климатических изменений // Современные проблемы дистанционного зондирования Земли из космоса. — 2011. — 8, № 2. — С. 246–256.
6. Marx A. Мониторинг лесов с помощью группировки спутников RapidEye // Геоматика. — 2011. — № 3. — С. 58–66.
7. Richards J.A. Remote Sensing Digital Image Analysis. — Springer-Verlag, Berlin, Heidelberg, 2006. — 454 с.

Надійшла до редакції 31.10.2012