

УДК 528.46:711.14

## ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ МЕТОДІВ РОЗДІЛЕННЯ КЛАСІВ ПРИ КЛАСИФІКАЦІЇ ЛІСІВ

Ю. Денис

Національний університет “Львівська політехніка”

**Ключові слова:** класифікація, евклідова відстань, дивергенція, трансформована дивергенція, відстань Джефріса–Матусіти, достовірність.

### Постановка проблеми та аналіз літератури

Розвиток космічної галузі та геоінформаційних технологій дає змогу ефективніше використовувати матеріали космічного знімання для вирішення основних лісогосподарських завдань завдяки високій інформативності даних дистанційного зондування, їх об'єктивності, регулярності отримання та широким можливостям опрацювання цих матеріалів для великих територій.

Вирішення таких важливих завдань лісового господарства, як контроль за знищенням лісів, інвентаризація лісів, картографування лісів та моніторинг лісових ресурсів, потребує об'єктивної інформації щодо лісів [1, 3, 5]. Саме тому космічне знімання є найдоступнішим видом інформації. Космічні знімки об'єктивно відображають стан місцевості на момент знімання, вони актуальні, оскільки матеріали космічного знімання можна одержувати на різні дати з різним розрізненням. Сучасне знімання в оглядових масштабах дає змогу одночасно знімати величезні території з доволі високою деталізацією.

Моніторинг лісів являє собою систему регулярного спостереження, оцінювання і прогнозу динаміки кількісного та якісного стану лісів. Моніторинг лісів є складовою частиною державної системи моніторингу навколишнього природного середовища [7]. Дані дистанційного зондування необхідні для картографування та проведення моніторингу стану лісів. Вихідні продукти подаються у вигляді карт вирубок, карт лісового покриву, карт типів лісів, карт віку лісу тощо [2, 6, 9].

Важливою складовою картографування лісів є процедура класифікації. Зокрема, в [10] автори вказують на необхідність додаткових досліджень для ідентифікації та зменшення невизначеності під час оброблення зображень для підвищення точності класифікації. Саме тому вибір оптимальної кількості класів є важливим етапом виконання класифікації лісів для тематичного картографування.

### Постановка завдання

Завдання цього дослідження – порівняти ефективність таких статистичних методів розділення класів,

як евклідова відстань, дивергенція, трансформована дивергенція та відстань Джефріса–Матусіти, для класифікації лісів та зробити висновок про доцільність їх застосування.

### Виклад основного матеріалу та результати досліджень

Отримання тематичної інформації на основі ДЗЗ та із застосуванням автоматизованих методів їх оброблення є тематичною класифікацією [8].

Класифікаційні алгоритми ґрунтуються на тому, що кожному земному об'єкту властива своя специфіка відбивання в різних ділянках ЕМ-спектра. Основне завдання класифікації полягає у встановленні зв'язку спектральних класів з тематичними. В деяких випадках спектральні образи можуть бути доволі компактними і завдяки цьому їх легко порівнювати з тим чи іншим тематичним класом. Але здебільшого, через значну неоднорідність земної поверхні, одному тематичному класу може відповідати декілька спектральних, що виявляється у їх перекритті в просторі спектральних ознак. В такому разі однозначність зарахування пікселя до того чи іншого тематичного класу на основі неоднозначних спектральних ознак залежить від вибраного класифікаційного підходу та алгоритму його реалізації.

Роздільність сигнатур (еталонів) – метод розрахунку статистичної різниці між двома сигнатурами, величина різниці, яка показує, наскільки одна сигнатура віддалена від іншої. Роздільність можна розраховувати за такими методами:

- спектральна евклідова відстань між двома середніми значеннями;
- дивергенція;
- трансформована дивергенція;
- відстань Джефріса–Матусіти.

Однією з причин використання статичних методів для розрахунку роздільності класів є можливість перекриття класів у просторі ознак. Наслідком є те, що класифікатор, використовуючи доступну дистанційну інформацію, допускає помилки, оскільки частина еталонів класу 1 не відрізняється від еталонів класу 2, частина еталонів класу 2 не відрізняється від класу 3 і т. д. Основна мета статистичного підходу – мінімізувати частоту появи таких помилок, тобто мінімізувати можливість помилки.

**Евклідова відстань**

Евклідову відстань між середніми значеннями за класами зручно використовувати в тих випадках, якщо класи доволі однорідні за яскравостями, тобто утворюють невеликі компактні зони в просторі ознак. Чим більша відстань між середніми значеннями за класами, тим менша ймовірність помилки. Така оцінка, як і метод класифікації за мінімальною відстанню, в цьому випадку виявляється ефективною [4].

Формула для розрахунку спектральної евклідової відстані:

**Дивергенція**

Формула для розрахунку дивергенції:

$$D_{ij} = \frac{1}{2}tr((C_i - C_j)(C_i^{-1} - C_j^{-1})) + \frac{1}{2}tr((C_i^{-1} - C_j^{-1})(\mu_i - \mu_j)(\mu_i - \mu_j)^T) \quad (1)$$

де  $i, j$  – дві сигнатури (класи), що порівнюються;  $C_i$  – коваріаційна матриця сигнатури  $i$  (класу  $i$ );  $\mu_i$  – середнє значення класу  $i$ ;  $tr$  – функція сліду, слід матриці;  $T$  – функція транспонування.

Вираз складається з двох доданків. Перший зумовлений лише відмінностями між відповідними коваріаційними матрицями, а другий – нормалізована відстань між математичними сподіваннями.  $D_{ij}$  дорівнюватиме нулю лише у випадку, якщо математичні сподівання і коваріаційні матриці будуть рівними [4].

Під час використання цього методу виникають деякі труднощі, тому що поведінка дивергенції як функції нормалізованої відстані між класами значно відрізняється від поведінки ймовірності правильного розпізнавання.

**Трансформована дивергенція**

Формула для обчислення трансформованої дивергенції:

$$D_{ij} = \frac{1}{2}tr((C_i - C_j)(C_i^{-1} - C_j^{-1})) + \frac{1}{2}tr((C_i^{-1} - C_j^{-1})(\mu_i - \mu_j)(\mu_i - \mu_j)^T) \quad (2)$$

$$TD_{ij} = 2000 \left( 1 - \exp\left(\frac{-D_{ij}}{8}\right) \right) \quad (3)$$

де  $i, j$  – дві сигнатури (класи), що порівнюються;  $C_i$  – коваріаційна матриця сигнатури  $i$  (класу  $i$ );  $\mu_i$  – середнє значення класу  $i$ ;  $tr$  – функція сліду, слід матриці;  $T$  – функція транспонування.

Трансформована дивергенція має верхню і нижню межі. Трансформована дивергенція змінюється в межах від 0 до 2000. Якщо обчислені значення дорівнюють відповідній верхній межі, то можна вважати, що класи розділені добре. Якщо ж ці значення дорівнюють нулю, то між класами розподілу немає. Якщо значення трансформованої дивергенції становить 1550 і вище, то між класами буде чітке розділення, якщо ж нижче від 1550 – подільність буде гіршою.

**Відстань Джефріса–Матусіті**

Формула для розрахунку відстані Джефріса–Матусіті:

$$a = \frac{1}{8}(m_i - m_j)^t \left( \frac{c_i + c_j}{2} \right)^{-1} (m_i - m_j) + \frac{1}{2} \ln \left( \frac{|(c_i + c_j)/2|}{|c_i| \times |c_j|} \right) \quad (4)$$

$$JM_{ij} = \sqrt{2(1 - e^{-a})} \quad (5)$$

де  $i, j$  – дві сигнатури (класи), що порівнюються;  $C_i$  – коваріаційна матриця сигнатури  $i$  (класу  $i$ );  $\mu_i$  – середнє значення класу  $i$ ;  $|C_i|$  – визначник матриці  $C_i$ .

Відстань Джефріса–Матусіті – це міра статистичного розділення між класами. Загалом це міра середньої різниці між функціями щільності для двох класів. JM міститься між 0 і 1414.

Як і у випадку з дивергенцією, припущення, що класи відповідають чіткому розподілу, дає змогу отримати вираз, що не містить інтегралів і залежить від математичних сподівань і коваріаційних матриць. Відмінність полягає у компоненті з від’ємною експонентою. Вплив цієї компоненти дає експоненціально меншу вагу різницям між функціями щільності класів, що зростають [4].

**Експериментальні дослідження**

Для виконання досліджень ефективності методів розділення класів використовуємо знімок лісового масиву, розташованого в Яворівському районі Львівської області, біля Івано-Франкове. Космічний знімок отриманий знімальною системою QuickBird у червні 2010 р. (рис. 1).



Рис. 1. Вхідний знімок

Виконали класифікацію з поділом на 3, 5, 7, 9, 11, 13 і 15 класів. Для кожного відкласифікованого зображення розрахували матриці з евклідовою відстанню, дивергенцією, трансформованою дивергенцією та відстанню Джефріса–Матусіті.

Критерієм для аналізу методів розділення слугували дані, отримані у попередніх дослідженнях, наведених у [2]. За даними контрольованої класифікації найвищу точність визначення площ класів отримано за поділом зображення на вісім класів.

Проаналізовано значення матриць, отриманих у разі поділу зображення на 3, 5, 7, 9, 11, 13 і 15 класів. У табл. 1–6 подано матриці зі значеннями дивергенції, трансформованої дивергенції та відстані Джефріса–Матусіти у разі поділу на сім та дев'ять класів. Згідно з літературними джерелами, саме ці методи вважають точнішою метрикою. Аналіз табл. 1 і 2, у яких подано значення трансформованої дивергенції, показує, що у разі поділу на дев'ять класів є два значення, які менші від 1550, а це означає, що ці класи не будуть розділятися (класи 4 і 5, 5 і 6). Відповідно, для подальшого опрацювання

знімка доцільне використання поділу на сім–вісім класів. Порівнявши значення трансформованої дивергенції та дивергенції (табл. 3, 4), можна зробити висновок про їх подібність. Відмінність полягає у тому, що дивергенція не має встановлених верхньої та нижньої меж, що утруднює її безпосередній аналіз. Аналіз матриць за методом відстані Джефріса–Матусіти складніший через відсутність конкретних значень для критерію роздільності класів. У результаті аналізу матриць (табл. 5, 6) зауважено, що в разі поділу на дев'ять класів одне значення біля діагоналі наближається до 1414,21, що свідчить про нечітку роздільність класів (класи 8 і 9).

Загалом встановлено, що методи трансформованої дивергенції та дивергенції близькі за отриманими результатами подільності класів.

Таблиця 1

Матриця значень трансформованої дивергенції у разі поділу на сім класів

Signature Name	1	2	3	4	5	6	7
Class 1	0	1553,93	1995,44	2000	1999,99	2000	1999,99
Class 2	1553,93	0	1594,22	1999,94	1991,25	2000	2000
Class 3	1995,44	1594,22	0	1974,87	1651,81	1999,88	2000
Class 4	2000	1999,94	1974,87	0	1746,59	1739,48	1995,8
Class 5	1999,99	1991,25	1651,81	1746,59	0	1983,26	1999,98
Class 6	2000	2000	1999,88	1739,48	1983,26	0	1693,25
Class 7	1999,99	2000	2000	1995,8	1999,98	1693,25	0

Таблиця 2

Матриця значень трансформованої дивергенції у разі поділу на дев'ять класів

Signature Name	1	2	3	4	5	6	7	8	9
Class 1	0	1639,61	1997,44	2000	2000	2000	2000	2000	1999,99
Class 2	1639,61	0	1550,03	1995,5	1999,99	2000	2000	2000	1999,99
Class 3	1997,44	1550,03	0	1580,33	1994,33	1999,98	2000	2000	1999,99
Class 4	2000	1995,5	1580,33	0	1545,55	1992,13	1999,98	2000	2000
Class 5	2000	1999,99	1994,33	<b>1545,55</b>	0	1540,02	1993,28	2000	2000
Class 6	2000	2000	1999,98	1992,13	<b>1540,02</b>	0	1557,43	1997,74	2000
Class 7	2000	2000	2000	1999,98	1993,28	1557,43	0	1586,56	2000
Class 8	2000	2000	2000	2000	2000	1997,74	1586,56	0	2000
Class 9	1999,92	1999,99	1999,99	2000	2000	2000	2000	2000	0

Таблиця 3

Матриця значень дивергенції у разі поділу на сім класів

Signature Name	1	2	3	4	5	6	7
Class 1	0	12,0034	48,676	139,178	95,2013	145,039	95,5943
Class 2	48,676	0	12,7608	83,6221	43,452	120,845	144,506
Class 3	139,178	12,7608	0	35,0144	13,9851	77,9079	122,852
Class 4	139,178	83,621	35,0144	0	16,5273	16,3058	49,3277
Class 5	95,2013	43,452	13,9851	16,5273	0	38,2642	93,1424
Class 6	145,039	120,845	77,9079	16,3058	38,2642	0	14,9989
Class 7	95,5943	144,506	122,85	49,3277	93,1421	14,9989	0

Таблиця 4

**Матриця значень дивергенції у разі поділу на дев'ять класів**

Signature Name	1	2	3	4	5	6	7	8	9
Class 1	0	13,7098	53,2943	118,886	191,397	230,415	201,452	123,791	80,7561
Class 2	13,7098	0	11,9337	48,7765	102,857	154,375	181,59	197,828	100,068
Class 3	53,2943	11,9337	0	12,4915	46,9265	93,5165	139,014	195,983	97,4479
Class 4	118,886	48,7765	12,4915	0	11,8546	44,2978	90,703	163,566	108,992
Class 5	191,397	102,857	46,9265	<b>11,8546</b>	0	11,7577	45,5644	112,27	104,669
Class 6	230,415	154,375	93,5165	44,2978	<b>11,7577</b>	0	12,0664	54,2886	104,21
Class 7	201,452	181,59	139,014	90,703	45,5644	12,0664	0	12,6112	122,222
Class 8	123,791	197,828	195,983	163,566	112,27	54,2886	12,6112	0	220,879
Class 9	80,7561	100,068	97,4479	108,992	104,669	104,21	122,222	220,879	0

Таблиця 5

**Матриця відстаней Джефріса–Матусіти у разі поділу на сім класів**

Signature Name	1	2	3	4	5	6	7
Class 1	0	1101,08	1369,92	1413,66	1409,39	1414,19	1414,21
Class 2	1101,08	0	1250,3	1414,17	1409,58	1414,21	1414,21
Class 3	1369,92	1250,3	0	1404,66	1240,02	1414,13	1414,11
Class 4	1413,66	1414,17	1404,66	0	1253,69	1274,51	1388,81
Class 5	1409,39	1409,58	1240,02	1253,69	0	1405,9	1412,69
Class 6	1414,19	1414,21	1414,13	1274,51	1405,9	0	1212,22
Class 7	1414,21	1414,21	1414,11	1388,81	1999,98	1212,22	0

Таблиця 6

**Матриця відстаней Джефріса–Матусіти у разі поділу на дев'ять класів**

Signature Name	1	2	3	4	5	6	7	8	9
Class 1	0	1126,33	1368,58	1409,21	1413,78	1414,19	1414,21	1414,21	1408,34
Class 2	1126,33	0	1229,45	1410,88	1414,2	1414,21	1414,21	1414,21	1398,67
Class 3	1368,58	1229,45	0	1252,36	1411,85	1414,21	1414,21	1414,21	1372,01
Class 4	1409,21	1410,88	1252,36	0	1242,73	1411,38	1414,19	1414,16	1358,32
Class 5	1413,78	1412	1411,85	1242,73	0	1238,89	1409,48	1412,7	1359,14
Class 6	1414,19	1414,21	1414,21	1411,38	1238,89	0	1227,52	1390,56	1376,79
Class 7	1414,21	1414,21	1414,21	1414,19	1409,48	1227,52	0	1161,64	1400,04
Class 8	1414,21	1414,21	1414,21	1414,16	1412,7	1390,56	1161,64	0	<b>1414,46</b>
Class 9	1408,34	1398,67	1372,01	1358,32	1359,14	1376,79	1400,04	1414,46	0

**Висновки**

1. Дослідження ефективності розділення класів виконано за зображенням лісового масиву, отриманого з космічного супутника QuickBird (червень, 2010 р).

2. Важливим для проведення тематичної класифікації є вибір оптимальної кількості класів. Здійснено поділ на 3, 5, 7, 9, 11, 13 і 15 класів.

3. Дослідження виконано для таких методів розділення класів: евклідової відстані, дивергенції, трансформованої дивергенції та відстані Джефріса–Матусіти. Аналіз статистичних методів розділення

класів показав, що найефективнішим для класифікації лісів є метод трансформованої дивергенції.

4. Встановлено, що для цієї території оптимальна кількість класів – сім–вісім.

**Література**

1. Багатоспектральні методи дистанційного зондування Землі в задачах природокористування / за ред. В. І. Лялько та М. О. Попова. – К.: Наукова думка, 2006. – 360 с.
2. Бурштинська Х. Методика двоетапної класифікації лісів за космічними зображеннями

високого розрізнення / Х. Бурштинська, Ю. Денис, Ю. Мадяр, Б. Поліщук // Сучасні досягнення геодезичної науки та виробництва. – 2016. – № I (31). – С. 148–155.

3. Данилин И. М. Лазерная локация Земли и леса: учеб. пособ. / И. М. Данилин, Е. М. Медведев, С. Р. Мельников. – Красноярск: Ин-т леса им. В. Н. Сукачева СО РАН, 2005. – 182 с.
4. Дейвис Ш. М. Дистанционное зондирование: количественный подход: пер. с англ. / Ш. М. Дейвис, Д. А. Ландгребе, Т. Л. Филлипс. – М.: Недра, 1983. – 415 с.
5. Аэрокосмический мониторинг лесов / А. С. Исаев, В. И. Сухих, Е. Н. Калашников и др. – М.: Наука, 1993. – 241 с.
6. Миклуш С. І. Дистанційне зондування Землі в лісовому господарстві: навч. посіб. / С. І. Миклуш, С. А. Гаврилюк, О. Г. Часковський. – Львів: ЗУКЦ, 2012. – 324 с.
7. Лісовий кодекс України [Електронний ресурс]. – Режим доступу: <http://www.zakon1.rada.gov.ua>
8. Свідзінська Д. В. Методи геоекологічних досліджень: геоінформаційний практикум на основі відкритої ГІС SAGA: навч. посіб. / Д. В. Свідзінська. – К.: Логос, 2014. – 402 с.
9. Сесін В. А. Геоінформаційний підхід до картографування лісового господарства // Вісник геодезії і картографії. – 2003. – № 3. – С. 27–32.
10. Lu D. and Wenga Q. Survey of image classification methods and techniques for improving classification performance // International Journal of Remote Sensing – 10 March 2007. – Vol. 28, No. 5. – P. 823–870.

### Дослідження ефективності методів розділення класів при класифікації лісів Ю. Денис

Досліджено ефективність таких методів розділення класів, як евклідова відстань, дивергенція, трансформована дивергенція та відстань Джефріса–Матусіти, для класифікації лісів. Дослідження проводились на знімку QuickBird (червень, 2010 р). Визначено оптимальний метод розділення класів для цього знімка.

### Исследование эффективности методов разделения классов при классификации лесов Ю. Денис

Исследовано эффективность таких методов разделения классов, как евклидово расстояние, дивергенция, трансформированная дивергенция и расстояние Джефриса–Матуситы при классификации лесов. Исследования проводились на снимке QuickBird (июнь, 2010 г.). Определен оптимальный метод разделения классов для этого снимка.

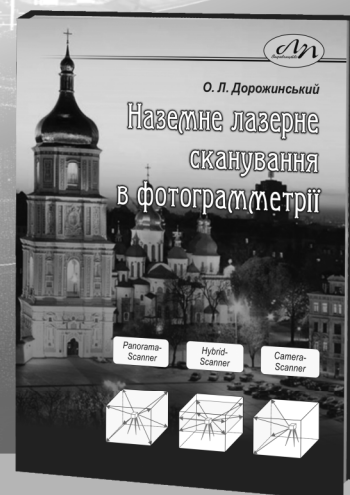
### Research of efficiency of separation of classes for the classification of forests J. Denus

Research of the effectiveness of such methods of separation of classes as Euclidean distance, divergence, transformed divergence and distance of Dzhefriss–Matusity for classification of forests was conducted. This research was conducted on the image QuickBird (June, 2010). The optimal method of separation of classes for this picture has been chosen.

- теоретичні засади наземного лазерного сканування
- аналіз найбільш вживаних приладів
- мобільне наземне лазерне сканування
- прикладне застосування сканування

Дорожинський О. Л.  
**НАЗЕМНЕ ЛАЗЕРНЕ СКАНУВАННЯ В ФОТОГРАММЕТРІЇ**  
Навчальний посібник.

Видавництво Львівської політехніки, 2014. 96 с.  
ISBN 978-617-607-617-9



**Нові технології –  
ключ до успіху геодезиста**