НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
ДЕРЖАВНА УСТАНОВА «НАУКОВИЙ ЦЕНТР АЕРОКОСМІЧНИХ ДОСЛІДЖЕНЬ ЗЕМЛІ ІНСТИТУТУ ГЕОЛОГІЧНИХ НАУК НАЦІОНАЛЬНОЇ АКАДЕМІЇ НАУК УКРАЇНИ»

АЛЬПЕРТ СОФІЯ ІОГАНІВНА

УДК 528.85(477.41)

МЕТОДИ КОНТРОЛЬОВАНОГО КЛАСИФІКУВАННЯ ГІПЕРСПЕКТРАЛЬНИХ КОСМІЧНИХ ЗОБРАЖЕНЬ В УМОВАХ НАВЧАЛЬНИХ ВИБІРОК ОБМЕЖЕНОЇ ЯКОСТІ

05.07.12 – Дистанційні аерокосмічні дослідження

АВТОРЕФЕРАТ
дисертації на здобуття наукового ступеня
кандидата технічних наук

Київ-2016

Дисертацією є рукопис.

Робота виконана у Державній установі «Науковий центр аерокосмічних досліджень Землі ІГН НАН України»

Науковий керівник: доктор технічних наук, професор, заслужений діяч науки і техніки України **Попов Михайло Олексійович**, заступник директора з наукової роботи Наукового Центру аерокосмічних досліджень Землі Інституту геологічних наук НАН України

Офіційні опоненти: доктор технічних наук, старший науковий співробітник **Різник Олександр Михайлович**, завідувач відділу Нейротехнології Інституту проблем математичних машин та систем НАН України

кандидат технічних наук, доцент **Зайцев Олександр Вікторович**, заступник начальника 3 кафедри Воєнно-дипломатичної академії імені Євгенія Березняка

Захист дисертації відбудеться 6 грудня 2016 р. о 11 годині на засіданні Спеціалізованої вченої ради Д 26.162.03 Державної установи «Науковий Центр аерокосмічних досліджень Землі ІГН НАН України» за адресою: Україна, 01601, м. Київ, вул. Олеся Гончара, 55-Б.

З дисертацією можна ознайомитися у бібліотеці Державної установи «Науковий Центр аерокосмічних досліджень Землі ІГН НАН України» (м. Київ, вул. Олеся Гончара, 55-Б)

Автореферат розіслано 31 жовтня 2016 р.

Вчений секретар спеціалізованої
вченої ради Д 26.162.03
кандидат геологічних наук,
старший науковий співробітник О.В. Седлерова

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми.

Поступово, з кожним роком все більше актуальних і важливих задач вирішується із залученням методів і технологій дистанційного зондування Землі (ДЗЗ). Можливості ДЗЗ сьогодні використовуються при екологічному моніторингу територій, вивченні природних ресурсів, аналізі причин та наслідків надзвичайних ситуацій, при оперативному оцінюванні стану лісів і сільськогосподарських земель, тощо. Новітні досягнення науки та техніки, такі, як нанотехнології, поява принципово нових видових технічних засобів і методів цифрового оброблення зображень, удосконалення глобальних інформаційних комунікацій та ін., сприяють подальшому розвитку ДЗЗ. Тому за останнє десятиліття суттєво збільшилися якість, різноманіття та об’єми матеріалів ДЗЗ, що, в свою чергу, сприяє розробці та впровадженню нових ефективних методів вивчення та дослідження геосистем.

Серед видових даних, які одержуються шляхом ДЗЗ, гіперспектральні космічні зображення (ГКЗ) є найбільш змістовними і вміщують в собі надзвичайно великий обсяг інформації про об'єкти зйомки та спостереження. Ця інформація дозволяє виявляти та розпізнавати (класифікувати) об'єкти на знімках, оцінювати їх стан, фіксувати зміни, що відбуваються на місцевості, надавати прогнозні оцінки тощо.

Розв’язання подібних актуальних і важливих практичних завдань вимагає залучення великої кількості різних логіко-обчислювальних процедур, при цьому найбільш складною процедурою серед них є процедура класифікування об'єктів на ГКЗ, оскільки, по-перше, розмірність ознакового простору при цьому є дуже великою (загальна кількість спектральних каналів для сучасних бортових сенсорів обчислюється сотнями) і, по-друге, спектральні ознаки не відрізняються сталістю.

Тому останні роки активно розробляються методи і алгоритми, які орієнтуються саме на контрольоване класифікування (класифікування з навчанням) гіперспектральних зображень, проте їх точність поки залишається недостатньою, особливо в умовах, коли вхідна інформація, яка необхідна для проведення класифікування, має елементи невизначення та/або неповноти.

Слід зазначити, що контрольоване класифікування потребує наявності навчальних вибірок, до яких при цьому висуваються певні вимоги щодо їх повноти, достатності та чистоти (забрудненості). На практиці ці вимоги задовольнити можна далеко не завжди, що впливає на точність результатів класифікування й, далі, на достовірність вирішення поставленої тематичної задачі ДЗЗ. Тому дисертаційне дослідження спрямовано на розробку нових ефективних методів контрольованого класифікування ГКЗ в умовах навчальних вибірок обмеженої якості.

Математичною основою виконаних розробок є положення теорії свідчень Демпстера-Шейфера (ТСДШ). Такий вибір обумовлений тим, що апарат ТСДШ дозволяє будувати інтервальні оцінки достовірності гіпотез в умовах відсутності частини вхідних даних, пропонує просте правило для комбінування даних (об'єднання інформацій) від різних джерел, може у складних ситуаціях

класифікування надавати багатоальтернативний розв'язок задачі (у вигляді кортежу з кількох гіпотез), а також має деякі інші переваги перед традиційним ймовірнісним підходом.

Запропоновані в роботі методи контрольованого класифікування відрізняються від відомих розв'язків однойменних задач введенням процедури оцінювання класифікаційної цінності спектральних каналів за допомогою спеціальної емпіричної функції і особливим способом розбиття спектрального ознакового простору, що, у сукупності, дозволяє застосовувати при класифікуванні лише найбільш інформативні спектральні канали і ознаки і, таким чином, суттєво скоротити розмірність ознакового простору при одночасному збільшенні точності класифікування.

Практичне застосування результатів роботи забезпечує покращення повноти та якості вирішення тематичних задач ДЗЗ за допомогою гіперспектральної космічної інформації.

Зв'язок роботи з науковими програмами, планами, темами. Дослідження, що складають зміст дисертаційної роботи, виконувались у період 2012-2016 рр. у відповідності з планом наукових досліджень Наукового Центру аерокосмічних досліджень Землі ІГН НАН України і, зокрема з планами роботи відділу геоінформаційних технологій в ДЗЗ. Результати досліджень одержано автором в рамках виконання таких НДР: "Методологія комплексного геоінформаційного аналізу матеріалів аерокосмічних спостережень Землі і наземних даних в інтересах підвищення ефективності природокористування" (№ держреєстрації 0111U000030), "Методи інформаційного інтегрування даних багатоспектрального аерокосмічного знімання, польової спектрометрії і геолого-геофізичних даних при вирішенні завдань дистанційного пошуку покладів корисних копалин на суходолі і морському шельфі" (№ держреєстрації 0112U000702).

Мета та задачі дослідження. Метою дисертаційної роботи є підвищення точності контрольованого класифікування ГКЗ в умовах навчальних вибірок обмеженої якості, зокрема, в умовах забруднення та малої кількості складових елементів.

Мета реалізується через постановку загального завдання розробити нові ефективні методи контрольованого класифікування ГКЗ в означених умовах.

Для вирішення поставленого завдання було сформульовано і досліджено такі задачі:

- аналіз сучасних методів класифікування ГКЗ;

- огляд відомих методів і підходів до класифікування, які основані на використанні положень ТСДШ;

- розробка нового методу контрольованого класифікування ГКЗ в умовах забруднення навчальних вибірок;

- розробка нового методу контрольованого класифікування ГКЗ в умовах забруднення навчальних вибірок та обмежень на їх кількість.

- алгоритмічна та програмна реалізація запропонованих методів контрольованого класифікування;

- проведення експериментальних досліджень розроблених методів контрольованого класифікування на реальних ГКЗ, зокрема, оцінювання точності пропонованих методів та їх порівняння із відомими сучасними методами і алгоритмами, такими як метод опорних векторів (Support Vector Machine - SVM) та алгоритм класифікування об'єктів за їх спектрально-топологічними характеристиками.

Об'єкт дослідження: методи контрольованого класифікування гіперспектральних космічних зображень.

Предмет дослідження: точність методів контрольованого класифікування гіперспектральних космічних зображень в умовах навчальних вибірок обмеженої якості.

Методи дослідження. При проведенні досліджень було залучено методи розпізнавання образів, математичний апарат ТСДШ, методи статистичного оцінювання, комп'ютерне моделювання. Методи розпізнавання образів використано при формуванні загальних схем класифікування ГКЗ в умовах забруднених навчальних вибірок та обґрунтуванні організації оброблення даних. Математичний апарат та положення ТСДШ дозволили запропонувати нові алгоритми формування ознакового простору та комбінування гіпотез, а також нове правило прийняття рішень. Методи статистичного оцінювання були покладені в основу критеріального вибору та розрахунків показників ефективності запропонованих методів класифікування, а також дозволили впровадити нову процедуру оцінювання класифікаційної цінності спектральних каналів за допомогою спеціальної емпіричної функції і скоротити таким чином розмірність ознакового простору при одночасному підвищенні точності класифікування. Застосування комп'ютерного моделювання дозволило провести експериментальне дослідження розроблених методів на реальних ГКЗ.

Наукова новизна отриманих результатів. У результаті виконання дисертаційної роботи отримано нові наукові результати, які виносяться на захист.

1. Вперше, на основі математичного апарату та положень теорії свідчень Демпстера-Шейфера розроблені та обґрунтовані такі методи:

- метод класифікування гіперспектральних космічних зображень в умовах забруднених навчальних вибірок;

- метод класифікування гіперспектральних космічних зображень в умовах забруднених навчальних вибірок та обмежень на їх об'єм.

2. Впроваджено нову функцію класифікаційної цінності спектральних каналів, що дозволяє на кількісній основі здійснювати відбір найбільш інформативних зональних зображень.

3. Запропоновано новий спосіб визначення репрезентативної комбінації зональних зображень при класифікуванні гіперспектрального знімка, який мінімізує навантаження на експерта і дає можливість зменшити вимоги до необхідного обчислювального ресурсу.

Практичне значення отриманих результатів полягає у тому, що:

- показана можливість підвищення точності класифікування ГКЗ в умовах, коли існують обмеження на об'єм навчальних вибірок, а також на їх чистоту (забрудненість). Саме з такими умовами дослідник має справу при дослідженні змішаних лісів, вивченні природних та штучних об'єктів на урбанізованих територіях, у важко доступних місцях (наприклад, в гірській місцевості) тощо;
- сталість розроблених методів до забрудненості навчальних вибірок (робасність) дозволяє (зрозуміло, в певних границях) компенсувати помилки експерта-дослідника на етапі навчання системи і далі зберігати достатньо високу точність класифікування;
- розроблене спеціальне програмне забезпечення може бути використано при вирішенні багатьох тематичних задач ДЗЗ: оцінюванні стану змішаних лісів і сільськогосподарських земель, при вивченні природних ресурсів, екологічному моніторингу забруднених територій тощо.

Особистий внесок здобувача. Основою дисертаційної роботи стали результати особистих досліджень та персональних напрацювань автора. В опублікованих зі співавторами наукових роботах, у яких знайшли відображення деякі наукові результати дисертації, внесок здобувача полягав у наступному:

- участь у розробці методів відбору інформативних зональних зображень при класифікуванні ГКЗ;
- участь у розробці нових методів класифікування ГКЗ в умовах обмеженої якості навчальних вибірок, які базуються на використанні математичного апарату та прийомів, розвинених у ТСДШ;
- алгоритмічна та програмна реалізація розроблених методів класифікування.

Апробація результатів дисертації. Результати дисертаційних досліджень доповідались на наступних наукових конференціях і семінарах: XII Міжнародна конференція "Геоінформатика: теоретичні та прикладні аспекти" (Київ, 2013); XIII Міжнародна конференція "Геоінформатика: теоретичні та прикладні аспекти" (Київ, 2014); IV Всеукраїнська конференція "Аерокосмічні спостереження в інтересах сталого розвитку та безпеки GEO-UA" (Київ, 2014); VII Міжнародна науково-технічна конференція студентів та аспірантів "Перспективи розвитку інформаційно-телекомунікаційних технологій та систем" (Київ, 2015); XIV Міжнародна конференція "Геоінформатика: теоретичні та прикладні аспекти" (Київ, 2015); 16 Українська конференція з космічних досліджень (Одеса, 2016); науково-учбові семінари Наукового Центру аерокосмічних досліджень Землі ІГН НАН України (Київ, 2013-2015).

Публікації. Основні результати досліджень опубліковані у 15-ти друкованих роботах, серед яких 7 наукових статей у фахових виданнях та 8 публікацій в збірниках і працях конференцій [1–15].

Структура і обсяг дисертації. Дисертація складається зі вступу, чотирьох розділів, загальних висновків, списку використаної літератури та додатку. Робота містить 153 сторінки, з них 123 сторінки основного тексту, в тому числі 36 рисунків

та 12 таблиць. Список використаних джерел літератури нараховує 112 найменувань на 13 сторінках та 4 додатки на 17 сторінках.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступній частині роботи обґрунтовано актуальність теми, сформульовано мету і задачі досліджень, Описано новизну та практичне значення отриманих результатів, надано відомості про апробацію роботи і публікацію наукових результатів.

Наголошується на тому, що основним недоліком вищеописаних методів є те, що їх застосування стає недостатньо ефективним при наявності навчальних вибірок обмеженої якості, оскільки при цьому страждає точність класифікування. Тому пропонується розробка таких методів класифікування, які можуть працювати в умовах відсутності “чистих” навчальних вибірок та недостатньої апріорної інформації.

У *першому розділі* **“Аналіз методів класифікування гіперспектральних космічних зображень. Постановка завдання на дослідження”** обґрунтовується необхідність використання процедури класифікування даних, які отримуються за допомогою ДЗЗ. Методи класифікування використовуються для вирішення природоресурсних задач, моніторингу навколишнього середовища, для розв’язання численних задач з геології, лісогосподарства, сільського господарства, при здійсненні пошуку корисних копалин, проведенні оцінки стану ґрунтів, тощо. Також у розділі стисло розглянуто сучасні методи контрольованого класифікування: метод мінімальної відстані, метод дистанції Махаланобіса, метод максимальної правдоподібності та метод неконтрольованого класифікування ISODATA. Проаналізовано основні переваги та недоліки зазначених методів.

На основі цього аналізу робиться висновок, що точність відомих методів при класифікуванні ГКЗ часто залишається недостатньою, особливо в умовах, коли вхідна інформація, необхідна для проведення класифікування, має елементи невизначенності та/або неповноти. Одним з таких обмежень є недостатня якість навчальних вибірок.

Наголошується, що у подібних умовах заслуговує на увагу підхід, оснований на використанні ТСДШ, і перелічуються основні його переваги перед традиційним теоретико-ймовірнісним підходом.

Формулюються задачі дисертаційного дослідження і розглядаються статистичні показники, за якими може бути оцінена точність класифікування.

У *другому розділі* **“Розробка методу контрольованого класифікування гіперспектральних космічних зображень в умовах забруднених навчальних вибірок”** розглянуто основні положення ТСДШ, на яких базується розробка двох нових методів контрольованого класифікування ГКЗ, що запропоновані в дисертації. Один з них детально розглянуто у цьому розділі, інший метод описано у наступному розділі.

В роботі прийнята модель ГКЗ вигляді

$$HSI = \{ S_k \mid k = 1,2,...,K \} = \{ (\pi_n, \mathbf{u_n}) \mid n = 1,2,..., N_\pi \}, \quad (1)$$

де S_k – k-те зональне зображення; K – загальна кількість зональних зображень; π_n – n-ий піксел з повним (векторним) сигналом $\mathbf{u}_n = (u_{1n}, u_{2n}, \cdots u_{kn}, \ldots, u_{Kn})$; u_{kn} – величина сигналу піксела π_n у k-му спектральному каналі (зоні); N_π – загальна кількість пікселів в кадрі ГКЗ.

Сформульовано деякі припущення та обмеження, які притаманні процедурі класифікування.

Схема організації виконання логіко-обчислювальних процедур при класифікуванні ГКЗ в умовах забруднених навчальних вибірок представлена на рис. 1.

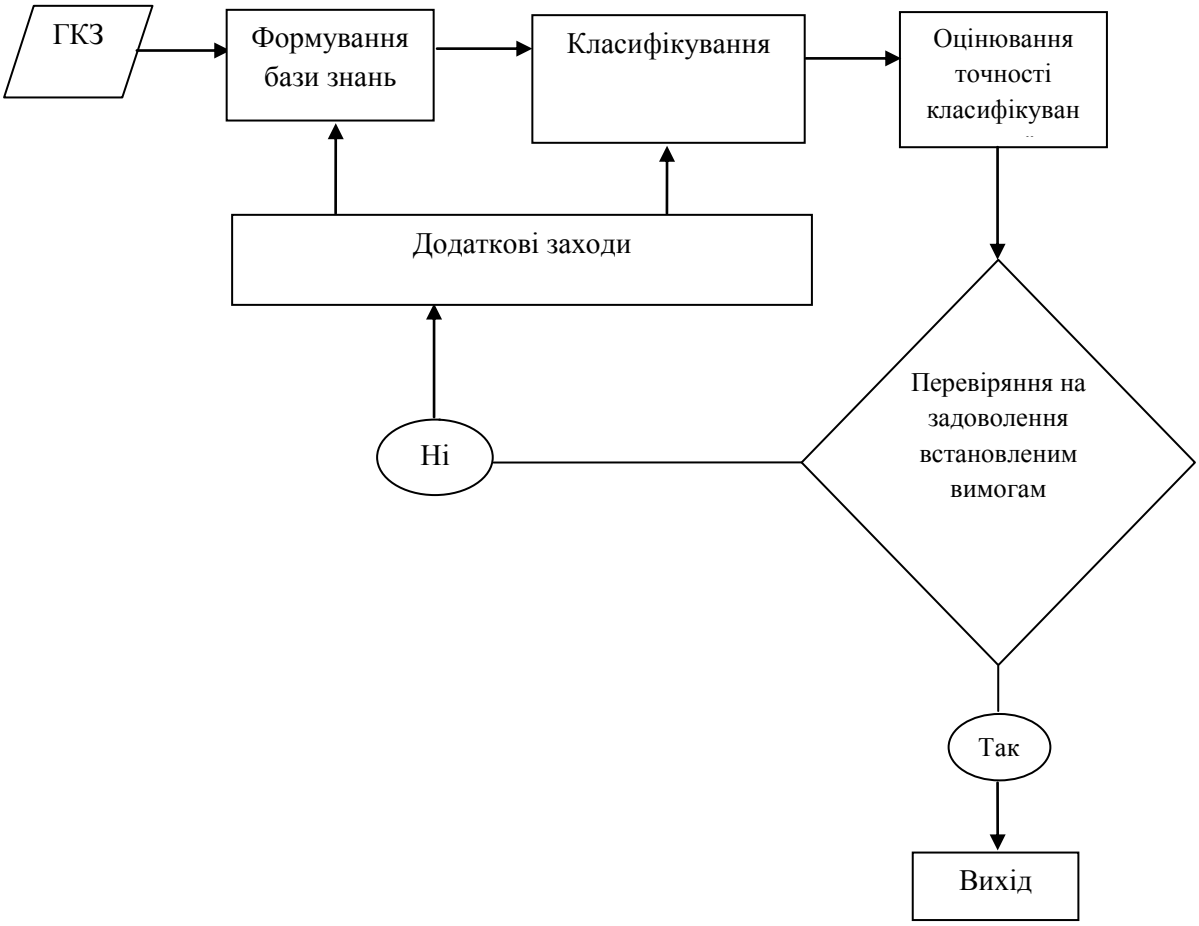


Рис. 1. Схема організації виконання логіко-обчислювальних процедур при класифікуванні ГКЗ в умовах забруднених навчальних вибірок

Метод включає такі основні процедури: формування бази знань (БЗ); здійснення класифікування ГКЗ; оцінювання точності класифікування; перевіряння на задоволення встановленим (заданим) вимогам.

Формування БЗ здійснюється на основі навчальної вибірки, яка одержується експертом-дослідником шляхом аналізу ГКЗ і виділення на ньому пікселів-репрезентів для кожного з наявних класів. Комплексний аналіз сигналів пікселів

одержаної навчальної вибірки дозволяє створити БЗ, інформація з якої надалі використовується при класифікуванні ГКЗ.

Процес класифікування ГКЗ здійснюється в автоматичному режимі. За результатами класифікування, при наявності референсних (завіркових) даних, будується матриця помилок (error matrix), на основі елементів якої обчислюються оцінки досягнутої точності. Якщо ті оцінки задовольняють встановленим (заданим) вимогам, процедура класифікування вважається завершеною. В іншому випадку здійснюються додаткові заходи, які можуть сприяти підвищенню точності класифікування (збільшення навчальної вибірки, уточнення описів класів і т. д.).

Процедура формування БЗ включає наступні роботи:

1. Відбір зональних зображень;
2. Виділення навчальної вибірки;
3. Фільтрація навчальної вибірки;
4. Розбиття ознакового простору;
5. Внутрішньоінтервальні оцінки для класів;
6. Визначення класифікаційної цінності спектральних каналів.

Також в другому розділі описано процедуру відбору найбільш інформативних зональних зображень $\{S_k\}$ з використанням коефіцієнту взаємної кореляції та процедуру виділення навчальної вибірки. Оскільки дуже рідко вдається виділити підвибірки, кожна з яких містить об'єкти тільки одного відповідного класу, та присутність сторонніх пікселів у навчальній вибірці негативно впливає на точність класифікування, тому для усунення їх зі складу вибірки було застосовано та описано спеціальну процедуру фільтрації.

Далі наведено процедуру розбиття ознакового простору. При цьому положення інтервалу на осі спектральних значень задається середнім значенням сигналів пікселів даного класу у спектральному каналі (наприклад, у k-ому спектральному каналі (зоні) для класів l та $l+1$ це будуть відповідно величини $\overline{u}_{k,l}$ та $u_{k,l+1}$, рис. 2), а положення граничної точки $A_{l,l+1}$, визначається зі співвідношення

$$\frac{a_l}{a_{l+1}} = \frac{\sigma_{k,l}}{\sigma_{k,l+1}}, \qquad (2)$$

де $\sigma_{k,l}$ і $\sigma_{k,l+1}$ – середньоквадратичне відхилення сигналів пікселів відповідних класів у k-ому спектральному каналі (зоні).

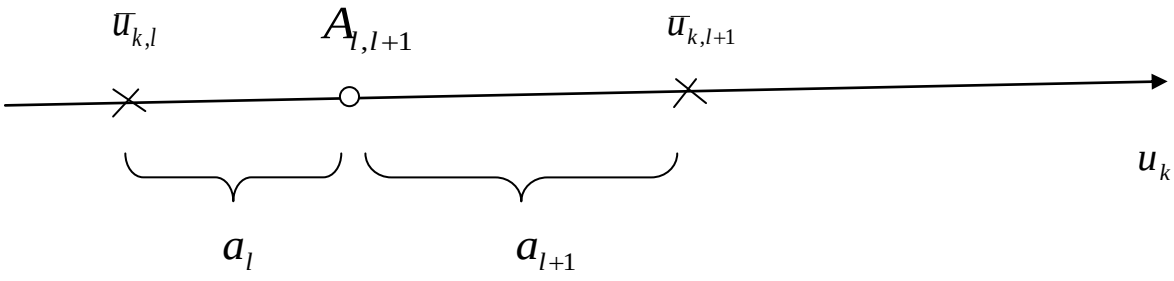


Рис. 2. До визначення положення границі, що розділяє класи на спектральній вісі.

Описано процедуру обчислення внутрішньоінтервальних оцінок для класів. Нехай, у деякому *j*-му інтервалі нараховується у цілому M_j пікселів, з них M_1 - це піксели класу, для якого даний інтервал – власний, і M_2 - загальне число пікселів інших класів. Тоді оцінка базової ймовірності піксельних об'єктів у власному інтервалі вказаного класу розраховується як

$$m^{(c)} = \frac{M_1}{M_j}, \tag{3}$$

а сумарна базова ймовірність піксельних об'єктів інших класів у цьому ж інтервалі розраховується як

$$m^{(нс)} = \frac{M_2}{M_j}. \tag{4}$$

Очевидно, що у кожному інтервалі $m^{(c)} + m^{(нс)} = 1$.

Для визначення класифікаційної цінності спектральних каналів введено функцію інформативності у вигляді

$$F = 1 - \frac{1}{L\ (L-1)} \cdot \sum_{m=1}^{L} \frac{\sum_{j=1}^{L} \left(\delta_{mj} \cdot \sum_{\substack{k=1 \\ k \neq m}}^{L} \delta_{kj} \right)}{\sum_{j=1}^{L} \delta_{mj}}, \tag{5}$$

де δ_{kj} – показник зв'язку класу, $F \in [0;1]$.

$$\delta_{kj} = \begin{cases} 1, & \textit{якщо об'єкт класу } k \textit{ пов'язаний з інтервалом } j; \\ 0, & \textit{в інших випадках}. \end{cases}$$

Класифікування здійснюється за послідовним алгоритмом. Припустимо, піксельний об'єкт π_n характеризується векторним сигналом $\mathbf{u}_n = (u_{1n}, u_{2n}, \ldots u_{kn}, \ldots, u_{Kn})$. Тоді алгоритм містить наступні кроки.

Крок 1. У кожному спектральному каналі визначається інтервал, з яким пов'язаний відповідний сигнал u_{kn} піксела π_n (*k=1,2,...,K).*

Крок 2. Встановлюються імена усіх класів об'єктів, що знаходяться у кожного такому інтервалі, після чого отримуються значення базової ймовірності власних об'єктів інтервалу та сумарної базової ймовірності об'єктів інших класів всередині інтервалу.

Крок 3. Розраховують за допомогою правила Демпстера комбіноване значення мас для результуючих фокальних елементів.

Крок 4. Використовуючи правило максимальної правдоподібності, приймається рішення про найбільш ймовірний клас (класи) даного поточного піксельного об'єкту.

У *третьому розділі* **“Розробка методу контрольованого класифікування гіперспектральних космічних зображень в умовах забруднення навчальних вибірок та кількісних обмежень”** запропонований метод класифікування ГКЗ в умовах, коли навчальна вибірка є не тільки забрудненою, а ще й малою за кількістю елементів, що її складають. Обмежена кількість елементів у навчальній виборці заважає застосовувати процедуру її фільтрації, яка входила до складу схеми метода, запропонованого у розділі 2 дисертації. Тому особливостями метода є удосконалена процедура навчання та впровадження правила прийняття рішення на основі пігністичної ймовірності.

Загальна схема організації виконання логіко-обчислювальних процедур при класифікуванні ГКЗ в умовах забруднення навчальних вибірок та кількісних обмежень представлена на рис. 3.

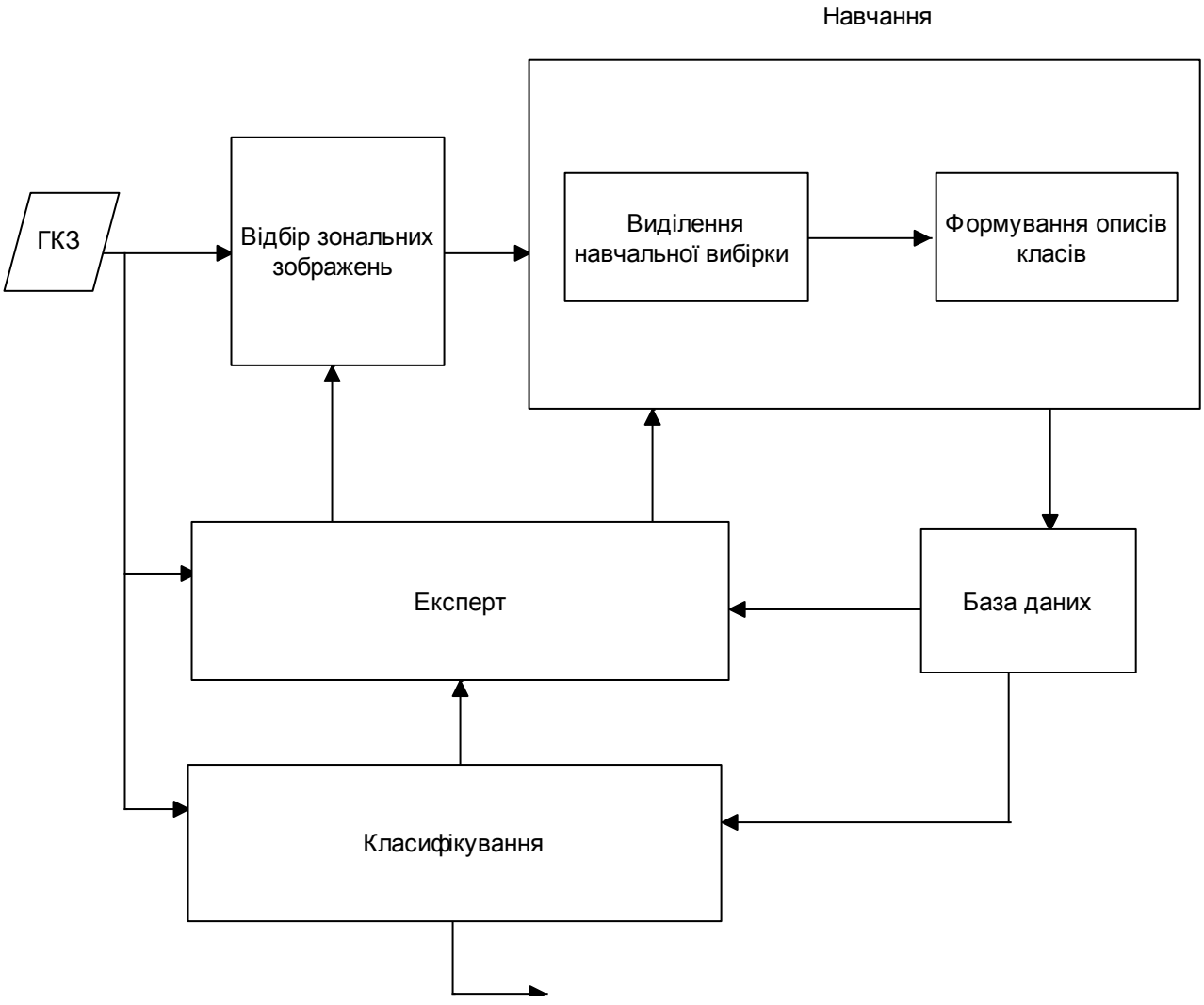


Рис. 3. Загальна схема організації виконання логіко-обчислювальних процедур при класифікуванні ГКЗ в умовах забруднення навчальних вибірок та кількісних обмежень

Метод включає такі основні процедури: відбір зональних зображень, навчання, формування бази даних (БД); здійснення класифікування ГКЗ.

Розроблено спосіб визначення репрезентативної комбінації зональних зображень при класифікуванні гіперспектрального знімка, який мінімізує навантаження на експерта і дає можливість об’єктивізувати і прискорити процес вирішення задачі за рахунок зменшення вимог до необхідного обчислювального ресурсу. Спосіб реалізується шляхом послідовного виконання чотирьох операцій.

- відбору зональних зображень за шумовими властивостями;

- побудови функції розподілу величин сигналів пікселів для кожного відібраного зонального зображення;

- кластеризації відібраних зображень;
- встановлення у кожному кластері відповідного зображення з максимальною репрезентативністю і формування таким чином репрезентативної комбінації зональних зображень.

Розрахунки показали, що застосування даного способу дозволяє у випадку знімків EO-1 / Hyperion зменшити необхідні обчислювальні витрати в 2,78-4,64 рази, а у випадку знімків AVIRIS – в 2,92-3,64 рази.Особливістю процедури навчання у методі є те, що після побудови класових інтервалів у всіх відібраних спектральних каналах виконується фокусизація. Під фокусизацією класового інтервалу мається на увазі створення списку фокальних підмножин на множині пікселів, сигнали яких проектуються на цей інтервал, та розрахунків для них базових ймовірностей.

Покажемо, як ці розрахунки виконуються на наступному прикладі. Нехай у деякому власному інтервалі класу l_1 знаходяться, крім Q_1 пікселів, які породжують цей інтервал (тобто віднесених експертом на етапі навчання до однойменного класу), ще Q_2 пікселів, яким експертом раніше була надана мітка належності до класу l_2, і Q_3 пікселів, яким експертом була надана мітка належності до класу l_3 $(l_1, l_2, l_3 \in L)$. При цих умовах список фокальних підмножин на даному класовому інтервалі буде містити три фокальних підмножини: $\{l_1\}$, $\{l_1, l_2\}$, $\{l_1, l_3\}$. Базові ймовірності для цих фокальних підмножин розраховуються за формулами:

$$\left.\begin{aligned} m(\{l_1\}) &= \frac{Q_1}{Q_\Sigma} \\ m(\{l_1, l_2\}) &= \frac{Q_2}{Q_\Sigma} \\ m(\{l_1, l_3\}) &= \frac{Q_3}{Q_\Sigma} \end{aligned}\right\}, \qquad (6)$$

де $Q_\Sigma = Q_1 + Q_2 + Q_3$.

Аналогічно виконуються розрахунки при наявності будь-якого іншого списку фокальних підмножин. Таким чином за результатами навчання отримуємо БД, що містить:

- опис (у вигляді номерів та границь) класових інтервалів у всіх K^* відібраних спектральних каналах;

- фокальні підмножини та їх маси для всіх класових інтервалів кожного з відібраних спектральних інтервалів.

На цьому процес навчання завершується і далі може виконуватися процедура класифікування. Класифікування пікселів ГКЗ здійснюється послідовно та незалежно, тому достатньо розглянути процедуру класифікування одного піксела. Нехай даний піксел π_n характеризується векторним сигналом $\mathbf{u}_n = \{ u_{nk} \}_{k=1}^{K^*}$. Знаючи цей сигнал та використовуючи сформовану при навчанні БД, можна визначити, до якого із *L* класів даний піксел найбільш ймовірно належить.

Алгоритм класифікування складається із наступних кроків.

Крок 1. У кожному *k*-ому спектральному каналі $(k = 1,2,... K^*)$ позначається номер (клас) інтервалу, на який проектується відповідна компонента векторного сигналу u_{nk} пиксела π_n.

Крок 2. З БД вилучається список фокальних підмножин та їх маси для зазначеного в кожному із спектральних каналів класового інтервалу.

Крок 3. Виконується об'єднання мас фокальних підмножин із вилученого списку із використанням комбінаційного правила Демпстера і отримуються оцінки мас цих підмножин із врахуванням всіх спектральних каналів.

Крок 4. Розраховуються пігністичні ймовірності розглянутих фокальних підмножин.

Крок 5. Розраховані значення пігністичних ймовірностей фокальних підмножин проранжовуються, і за критерієм максимуму пігністичної ймовірності визначається найбільш ймовірний клас піксела π_n.

Даний алгоритм застосовується послідовно до кожного піксела і результатом є відкласифіковане в цілому ГКЗ. По завершенні класифікування ГКЗ проводяться оцінювання точності класифікування і перевіряння на задоволення встановленим (заданим) вимогам.

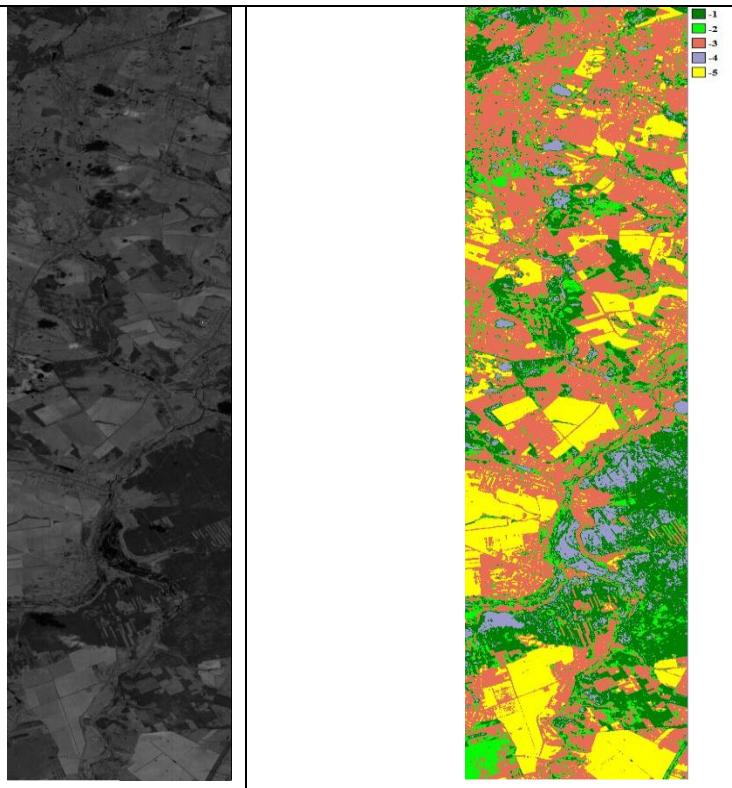
Ч*етвертий розділ* **“Експериментальне дослідження ефективності розроблених методів контрольованого класифікування гіперспектральних космічних зображень”** містить матеріали експериментальних досліджень ефективности розроблених методів контрольованого класифікування на основі використання реальних ГКЗ, отриманих 220-канальною оптико-електронною апаратурою Hyperion супутника ЕО-1.

Для реалізації та дослідження ефективності розроблених методів контрольованого класифікування космічних зображень було створено спеціалізоване програмне забезпечення (СПЗ) DST у середовищі пакету Visual Basic

6.0. У розділі наводиться його архітектура та описано функціональні особливості. Результати класифікування запропонованих методів порівнюються із результатами класифікування інших відомих методів, таких як метод SVM та алгоритм класифікування об’єктів за їх спектрально-топологічними характеристиками. Оцінювання точності класифікування здійснюється за допомогою таких показників: повна точність класифікування, точність користувача, точність виробника, помилки пропуску класу, помилки ототожнення з класом. Всі ці показники розраховуються на основі матриць помилок.

Експериментальне дослідження складалось з двох частин. У першій частині оцінювалась точність методу контрольованого класифікування, запропонованого і описаного у другому розділі, а також досліджено вплив процедур фільтрації навчальної вибірки і відбору певних спектральних каналів шляхом ранжування за їх класифікаційною цінністю на точність класифікування.

Для цього використовувалося гіперспектральне зображення EO1H1810252013112110KF, отримане 22.04.2013 р. (рис. 4).



а б

Рис. 4: а – знімок Hyperion за 22.02.2013 р.;

б –класифікаційна карта (розподіл класів)

У табл. 1 наведені значення точності класифікування ГКЗ трьома методами: запропонованим методом на основі ТСДШ, методом SVM та алгоритмом класифікування об’єктів за їх спектрально-топологічними характеристиками.

Таблиця 1

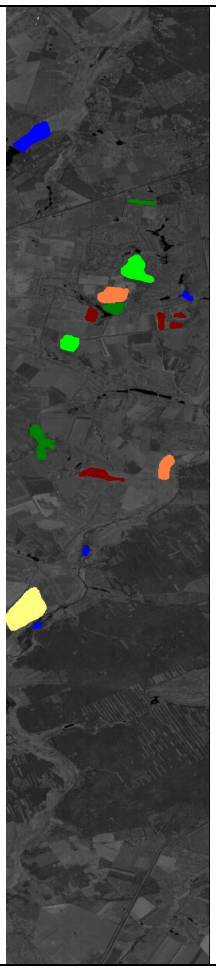
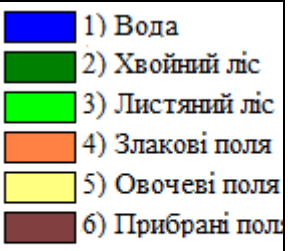
Точність класифікування ГКЗ різними методами

	Запропонований метод на основі ТСДШ	Метод SVM	Алгоритм класифікування об’єктів за їх спектрально-топологічними характеристиками
Точність (Overall Accuracy)	0,802	0,757	0,734

У другій частині експериментального дослідження оцінювалась точність методу контрольованого класифікування, запропонованого і описаного у третьому розділі дисертації. Для оцінки ефективності цього методу було використано два зображення: гіперспектральний знімок Hyperion EO1H1810252014211110KF, отриманий 30 липня 2014 р. (табл.2), та гіперспектральний знімок Hyperion EO1H1810252013157110K3-Refint від 6 червня 2013 р.

Таблиця 2

Класи об’єктів, навчальні підвибірки та контрольні ділянки

Класи об’єктів	**Навчальні підвибірки**		**Контрольні ділянки**		**Легенда**
Вода	335		1394		
Сосновий ліс	397		1471		
Листяний ліс	344		1541		
Злакові поля	525		1349		
Овочеві поля	440		2059		
Прибрані поля	315		1347		
Усього пікселів	2356		9161		

Результати класифікування показали, що метод DST дав середню за усіма класами точність (Overall Accuracy) 0,859 при навчанні на незабрудненій вибірці і 0,629 - на 50% забрудненій вибірці. У випадку методу SVM були отримані відповідно такі точності класифікування: 0,778 і 0,475. Тобто точність методу DST виявилася на 8-16% вище у порівнянні із методом SVM. Усереднені за класами результати розрахунків для випадку вихідної навчальної вибірки (без забруднення) наведено у табл. 3.

Таблиця 3

Усереднені помилки пропуску та ототожнення

Метод класифікування	Усереднена помилка пропуску класу (Average Omission Error)	Усереднена помилка ототожнення із класом (Average Commission Error)
DST	15%	13%
SVM	24%	22%

Порівняння усереднених за класами результатів розрахунків у табл. 3 показує, що й за цим показником метод DST виявляється на 9% більш ефективним, ніж метод SVM.

Подальше дослідження ефективності розробленого і описаного в розділі 3 методу класифікування ГКЗ для малих за кількістю складових пікселів та значно забруднених навчальних вибірок проводилося з використанням ще одного гіперспектрального знімку Hyperion EO1H1810252013157110K3-Refint, отриманого 6 червня 2013 р. На цьому знімку представлені ділянки Іванківського та Вишгородського районів Київської області та дельта річки Тетерів при впаданні у Київське водосховище.

При цьому, для повноти дослідження і одержання більш репрезентативних оцінок, процедура класифікування повторювалась при різних способах відбору інформативних спектральних каналів:

- з використанням функції кореляції (Correlation split/merge);
- на основі дивергенції Кульбака-Лейблера (KL selection);
- з використанням кількісної міри інформативності набору спектральних каналів на основі дивергенції Кульбака-Лейблера між спектральними розподілами оптичних сигналів об’єктів і фонів (C selection);
- з використанням функції похибок (Erf selection).

Загальна точність класифікування (Overall Accuracy) кожного методу при різних способах відбору спектральних каналів відображена в табл. 4.

Таблиця 4

Точність класифікування (Overall Accuracy)

Метод / алгоритм класифікування	Метод відбору спектральних каналів				
	Radiometric calibration and surface reflectance (172 bands total)	Correlation split/merge (12:34, 37:74, 87:96, 100:115, 134:163, 191:217)	KL selection (23, 45, 72, 66, 103, 108, 145, 157, 199)	C selection (12:13, 40, 43:45, 130, 132, 181, 182:184, 185:186, 189:190)	Erf selection (20, 33, 40, 130, 153, 181, 183, 211)
DST	0,900	0,905	0,888	0,883	0,888
Алгоритм класифікування об’єктів за їх спектрально-топологічними характеристиками	0,817	0,772	0,794	0,685	0,735

Порівняння одержаних в дослідженні оцінок показує, що запропонований в розділі 3 дисертації метод має хоча й невелику, але стабільну перевагу в точності класифікування перед топологічним алгоритмом.

ВИСНОВКИ

1. Розглянуто місце і роль процедури класифікування космічних зображень при розв’язанні тематичних задач дистанційного зондування Землі. Проведено стислий огляд сучасних методів і алгоритмів класифікування гіперспектральних космічних зображень. Показано, що сучасні алгоритми контрольованого класифікування гіперспектральних космічних зображень досить чутливі до характеристик навчальних вибірок. Більш конкретно, точність відомих алгоритмів суттєво залежить від об’ємів навчальних вибірок та рівня їх забрудненості.

2. Сформульовано завдання розробити та обґрунтувати нові ефективні методи контрольованого класифікування гіперспектральних космічних зображень в умовах, коли навчальна вибірка є забрудненою, в тому числі сильно, а її об’єм є дуже обмеженим. Наведено критерії, за якими передбачається оцінювати ефективність нових методів контрольованого класифікування. Математичною платформою для розроблення нових методів контрольованого класифікування обрано апарат та положення теорії свідчень Демпстера-Шейфера. Вибір пов'язаний з тим, що математичний апарат ТСДШ дає можливість в умовах забрудненості та неповноти даних працювати зі складними багатоальтернативними гіпотезами, пропонує просте правило для комбінування даних від різних джерел (у тому числі

зображень), а також має деякі інші переваги перед традиційним теоретико-ймовірнісним підходом.

3. Розроблено новий метод класифікування гіперспектральних космічних зображень в умовах забруднених навчальних вибірок. Введено модель гіперспектрального космічного зображення та припущення. Детально розглянуто організацію і зміст методу класифікування гіперспектральних космічних зображень в умовах забруднених навчальних вибірок. Запропонований метод відрізняється від відомих розв'язків однойменних задач введенням процедури оцінювання класифікаційної цінності спектральних каналів за допомогою спеціальної емпіричної функції і особливим способом розбиття спектрального ознакового простору, що, у сукупності, дозволяє застосовувати лише найбільш інформативні спектральні канали і, таким чином, суттєво скоротити розмірність ознакового простору при одночасному підвищенні точності класифікування. Детально описано процедуру відбору зональних зображень для виключення зашумлених зображень із використанням коефіцієнту кореляції між сусідніми зональними зображеннями, наведено процедури фільтрації навчальної вибірки, розбиття ознакового простору.

4. Запропоновано новий метод контрольованого класифікування гіперспектральних космічних зображень в умовах забруднення навчальних вибірок та обмежень на кількість елементів у виборці. Наводяться обмеження, припущення та визначення, на яких побудовано метод. Пропонований метод враховує невизначеності, які можуть виникати при навчанні та класифікуванні. Детально описано організацію метода і основні процедури, з яких він складається, у тому числі процедуру відбору зональних зображень і схему класифікування гіперспектральних космічних зображень в умовах забруднення навчальних вибірок та кількісних обмежень.

5. Запропоновано новий спосіб відбору зональних зображень і формування репрезентативної комбінації зональних зображень при класифікуванні гіперспектрального аерокосмічного знімка, який дозволяє об'єктивізувати і прискорити процес вирішення задачі за рахунок зменшення вимог до необхідного обчислювального ресурсу. Дана перевага досягається за рахунок автоматизації процедури оцінювання рівню шумів кожного з вихідних зональних зображень і відсіювання з подальшого аналізу зображень, які мають високий рівень зашумленості.

Показано, що введення операції відбору зашумлених зональних зображень у випадку знімків EO-1 / Hyperion дозволяє зменшити необхідні обчислювальні витрати в 2,78-4,64 рази, а у випадку знімків AVIRIS – в 2,92-3,64 рази.

6. Проведено експериментальну оцінку ефективності обох запропонованих в дисертаційній роботі методів контрольованого класифікування на реальних гіперспектральних космічних зображеннях, отриманих апаратурою Hyperion супутника EO-1, на базі спеціально розробленого програмного забезпечення DST.

7. Проведено оцінку ефективності розробленого методу класифікування гіперспектральних космічних зображень в умовах забруднених навчальних вибірок з використанням гіперспектрального знімка EO1H1810252014211110KF, отриманого 30 липня 2014 р. бортовою апаратурою Hyperion супутника EO-1. Результати

оцінювання даного методу показали, що його загальна точність (Overall Accuracy) складає 0,802 і перевищує загальну точність відомого методу SVM (склала 0,757) та алгоритму класифікування об’єктів за їх спектрально-топологічними характеристиками (склала 0,734).

8. Проведено оцінку ефективності розробленого методу DST класифікування гіперспектральних космічних зображень Hyperion в умовах забруднених малих навчальних вибірок. Дослідження показало, що метод DST дав середню за всіма класами загальну точність 0,859 при навчанні на незабрудненій вибірці і 0,629 на 50% забрудненій вибірці. У випадку методу SVM були отримані відповідно такі точності класифікування: 0,778 і 0,475. Тобто точність методу DST в умовах забруднених малих навчальних вибірок є на 8-16% вище у порівнянні із методом SVM. Встановлено, що усереднена помилка пропуску та усереднена помилка ототожнення із класом для методу DST становлять 15% та 13% відповідно, а для методу SVM - 24% та 22% відповідно, тобто запропонований метод DST в умовах забруднених малих навчальних вибірок показує перевагу перед методом SVM.

9. Здійснено порівняльну оцінку точності запропонованого метода DST та алгоритма класифікування об’єктів за їх спектрально-топологічними характеристиками з урахуванням способу відбору спектральних каналів. Для методу DST при відборі спектральних каналів із використанням функції кореляції загальна точність методу DST склала 0,905; при відборі на основі дивергенції Кульбака-Лейблера - 0,888; при відборі з використанням кількісної міри інформативності на основі дивергенції Кульбака-Лейблера між спектральними розподілами оптичних сигналів об’єктів і фонів - 0,883; при відборі з використанням функції похибок або функції Лапласа - 0,888. Для алгоритму класифікування об’єктів за їх спектрально-топологічними характеристиками значення загальної точності класифікування при різних способах відбору спектральних каналів складало відповідно: 0,772, 0,794, 0,685, 0,735. Тобто, точність запропонованого метода DST в умовах забруднених малих навчальних вибірок є на 11-17% вище у порівнянні із тою, що забезпечує алгоритм класифікування об’єктів за їх спектрально-топологічними характеристиками.

10. Таким чином, результати проведених наукових досліджень свідчать, що розроблено і обґрунтовано нові ефективні методи контрольованого класифікування гіперспектральних космічних зображень, які здатні забезпечити підвищення точності класифікування в умовах забруднення та малої кількості складових елементів, тобто мета дисертаційної роботи *досягнута.*

11. Подальші дослідження мають бути спрямовані на оптимізацію обчислювальних витрат на відбір найбільш інформативних зональних зображень, а також на пошук методів попереднього перетворення ознакового простору, які забезпечать можливість краще розрізняти класи об’єктів у просторі спектральних ознак.

СПИСОК ОПУБЛІКОВАНИХ РОБІТ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Альперт С.І. Сучасні критерії оцінки точності класифікації аерокосмічних зображень / С.І. Альперт // Математичні машини і системи. – 2013. – № 4. – С. 187–197.
2. Альперт С.І. Оцінка якості класифікації аерокосмічних зображень на основі матриці помилок та коефіцієнтів точності / С.І. Альперт // Математичні машини і системи. – 2014. – № 1. – С. 101–107.
3. Альперт С.І. Аналіз методів відбору найбільш інформативних спектральних каналів для вивчення об'єктів дистанційного зондування / С. І. Альперт // Математичні машини і системи. – 2014. – № 2. – С. 71–77.
4. Попов М.О. Метод класифікування гіперспектральних космічних зображень з використанням комбінаційного правила Демпстера / М.О. Попов., С.І. Альперт, В.М. Подорван //Космічна наука і технологія. – 2015. – Т. 21. – №1. – С. 25–37. *(Особистий внесок: постановка задачі, проведення математичних розрахунків, обговорення із співавторами висновків та підготовка статті до публікації.)*
5. Альперт С.І. Методи селекції інформативних зональних зображень при класифікації гіперспектральних супутникових зображень / С.І. Альперт // Математичні машини і системи. – 2015. – № 2. – С. 40–48.
6. Popov M. Method of Hyperspectral Satellite Image Classification under Contaminated Training Samples Based on Dempster-Shafer's Paradigm / M. Popov, S. Alpert, V. Podorvan, M. Topolnytskyi, S. Mieshkov // Central European Researchers Journal. – 2015. – Vol. 1. – № 1. – P. 86–97.
7. Попов М.А. Метод классификации космических изображений с использованием подхода Демпстера Шейфера / М.А. Попов, С.И. Альперт, В.Н. Подорван // Исследование Земли из космоса (РФ). – 2016. – № 5. – С. 26–37.

8. Альперт С.І. Оцінка точності класифікації космічних зображень на основі теорії Демпстера-Шафера / С.І. Альперт // Збірник праць XI-ої Міжнародної молодіжної науково-практичної конференції "Історія розвитку науки, техніки та освіти" за темою "Розбудова дослідницького університету", (Київ, 25 квітня 2013 р,). – Київ, 2013. – С. 242–245.
9. Альперт С.І. Сучасні методи оцінки точності класифікації/ С.І. Альперт // Геоінформатика – теоретичні і прикладні аспекти: матеріали 12-ї Міжнародної конференції, (Київ, 13–16 травня 2013). – К.: Наук. думка, 2013. – С. 1– 4.
10. Альперт С.І. Огляд різних методів відбору найбільш інформативних спектральних діапазонів / С.І. Альперт // Геоінформатика – теоретичні та прикладні аспекти: матеріали 13-ої Міжнародної конференції, (Київ, 12–15 травня 2014). – К.: Наук. думка, 2014. – С. 1–5.
11. Альперт С.І. Огляд оцінок точності та неточності класифікації земляного покриву та їх властивостей / С.І. Альперт // Тези доповідей IV Всеукраїнської конференції "Аерокосмічні спостереження в інтересах сталого розвитку та безпеки GEO-UA", (Київ, 26–30 травня 2014). – Київ, 2014. – С. 6–8.
12. Альперт С.І. Методи відбору найбільш інформативних спектральних каналів/ С.І. Альперт //Збірник праць VII-ої Міжнародної науково-технічної конференції

студентів та аспірантів “Перспективи розвитку інформаційно-телекомунікаційних технологій та систем”, (Київ, 21–24 квітня 2015). – Київ, 2015. – С. 335–337.
13. Альперт С.І. Нова модифікована критеріальна функція інформативності спектральних каналів/ С.І. Альперт // Геоінформатика – теоретичні та прикладні аспекти: матеріали 14–ої Міжнародної конференції (Київ, 11–14 травня 2015). – К.: Наук. думка, 2015. – С. 1–6.
14. Попов М.А. Метод автоматической классификации гиперспектральных космических изображений на основе теории свидетельств Демпстера–Шейфера/ М.А. Попов, В.Н. Подорван, С.И. Альперт // Тези доповідей XIV-ої Української конференції з космічних досліджень, (Київ, 8–12 вересня 2014). – Київ, 2014. – С. 152. *(Особистий внесок: участь у постановці задачі та опис використаних методів)*
15. Попов М.О. Оптимізація кластерної процедури визначення репрезентативної комбінації зональних зображень при класифікуванні гіперспектральних аерокосмічних знімків / М.О. Попов, С.А. Станкевич, С.П. Ковальчук, М.В. Топольницький, В.М. Подорван, В.В. Пилипчук, С.І. Альперт // Abstracts of 16th Ukrainian Conference on Space Research (Odesa, 2016, August, 22–27). – Kyiv: SSAU. 2016. – P. 219.

АНОТАЦІЯ

Альперт С.І. Методи контрольованого класифікування гіперспектральних космічних зображень в умовах навчальних вибірок обмеженої якості. – На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.07.12 – Дистанційні аерокосмічні дослідження. – Державна установа «Науковий центр аерокосмічних досліджень Землі ІГН НАН України», Київ, 2016.

Метою дисертаційної роботи є підвищення точності контрольованого класифікування гіперспектральних космічних зображень (ГКЗ) в умовах навчальних вибірок обмеженої якості, зокрема, в умовах їхньої забрудненості та малої кількості складових елементів.

Проведено огляд методів і алгоритмів класифікування космічних зображень і показано, що сучасні алгоритми контрольованої класифікації ГКЗ достатньо чутливі до характеристик навчальних вибірок. Тому була сформульована задача розробити і обґрунтувати нові методи контрольованого класифікування ГКЗ, ефективні в ситуаціях, коли навчальна вибірка є забрудненою, в тому числі сильно, а її обсяг є обмеженим. Математичною платформою для нових методів контрольованого класифікування гіперспектральних зображень є апарат і положення теорії свідчень Демпстера-Шейфера.

Запропоновано новий метод класифікування ГКЗ в умовах забруднення навчальних вибірок. Метод відрізняється від відомих розв’язків однойменних задач введенням процедури оцінювання класифікаційної цінності спектральних каналів за допомогою спеціальної емпіричної функції і особливим способом розбиття спектрального ознакового простору, що дозволяє застосовувати лише найбільш інформативні спектральні канали.

Також запропоновано новий метод контрольованого класифікування гіперспектральних зображень, що є ефективним в умовах забруднення навчальних вибірок та обмежень на їх об’єм.

Обидва розроблені методи контрольованого класифікування ГКЗ реалізовані алгоритмічно та програмно. Оцінювання ефективності розроблених методів проводилось із використанням реальних космічних зображень, отриманих гіперспектральною знімальною апаратурою Hyperion/EO-1.

Результати досліджень та розрахунків вказують на досить високу точність запропонованих методів контрольованого класифікування. Проведені порівняльні оцінки показали, що запропоновані методи є точнішими за відомий метод опорних векторів (Support Vector Machine) і алгоритм класифікування об'єктів за їх спектрально-топологічними характеристиками.

Запропоновані методи можуть бути використані при комп'ютерному аналізі гіперспектральних і багатоспектральних космічних знімків в задачах пошуку корисних копалин, при вирішенні природоресурсних та екологічних задач тощо.

Ключові слова: гіперспектральне космічне зображення, класифікування зображень, теорія свідчень Демпстера-Шейфера, навчальна вибірка, якість вибірки, точність класифікування.

АННОТАЦИЯ

Альперт С. И. Методы контролируемой классификации гиперспектральных космических изображений в условиях обучающих выборок ограниченного качества. – На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.07.12 – Дистанционные аэрокосмические исследования. – Государственное учреждение «Научный центр аэрокосмических исследований Земли ИГН НАН Украины», Киев, 2016.

Целью диссертационной работы является повышение точности контролируемой классификации гиперспектральных космических изображений (ГКИ) в условиях обучающих выборок ограниченного качества, в частности, в условиях их загрязненности и малого количества составных элементов.

Проведен обзор методов и алгоритмов классификации космических изображений и показано, что современные алгоритмы контролируемой классификации ГКИ весьма чувствительны к характеристикам обучающих выборок. Поэтому была сформулирована задача разработать и обосновать новые методы контролируемой классификации ГКИ, эффективные в ситуациях, когда обучающая выборка загрязнена, в том числе сильно, а ее объем является ограниченным. Математической платформой для новых методов контролируемой классификации гиперспектральных изображений является аппарат и положения теории свидетельств Демпстера-Шейфера.

Разработан новый метод классификации ГКИ в условиях загрязненных обучающих выборок. Метод отличается от известных решений одноименных задач введением процедуры оценивания классификационной ценности спектральных каналов с помощью специальной эмпирической функции и особым способом разбиения спектрального признакового пространства, что позволяет задействовать лишь наиболее информативные спектральные каналы и, таким образом, существенно сократить размерность признакового пространства при одновременном повышении точности классификации.

Также предложен новый метод контролируемой классификации ГКИ, который является эффективным в условиях загрязненных обучающих выборок и ограничений на их объем.

Оба разработанных метода контролируемой классификации ГКИ были реализованы алгоритмически и программно, что позволило провести экспериментальное исследование их точности. Оценка эффективности разработанных методов классификации проводилась с использованием реальных космических изображений, полученных гиперспектральной съемочной аппаратурой Hyperion / EO-1.

Результаты исследований и расчетов указывают на достаточно высокую точность разработанных в диссертации методов контролируемой классификации ГКИ. Проведенные сравнительные оценки показали, что предложенные методы более точные, чем известный метод опорных векторов (Support Vector Machine) и алгоритм классификации объектов по их спектрально-топологическим характеристикам, и являются достаточно простыми с вычислительной точки зрения.

Предложенные методы могут быть использованы при компьютерном анализе гиперспектральных и многоспектральных космических снимков в задачах поиска полезных ископаемых, при решении природоресурсных и экологических задач и др.

Ключевые слова: гиперспектральное космическое изображение, классификация изображений, теория свидетельств Демпстера-Шейфера, обучающая выборка, качество выборки, точность классификации.

ANNOTATION

Alpert S.I. Methods for hyperspectral satellite image supervised classification under training samples of limited quality. – Manuscript.

On obtain the Candidate of Sciences Degree Thesis, specialty 05.07.12 – Remote Aerospace Research. – State Institution "Scientific Centre for Aerospace Research of the Earth of IGS NAS of Ukraine", Kyiv, 2016.

The aim of the present dissertation is to improve an accuracy of supervised classification hyperspectral satellite images (HSI) under training samples of limited quality, particularly when the samples are contaminated and include only small number of constituent elements.

The-state of-the art of approaches to satellite image classification was considered and it was shown, that modern algorithms for HSI supervised classification are very sensitive to the characteristics of training samples. So formulated the aim to develop and justify new methods for HSI supervised classification under the aforementioned conditions.

The mathematical platform of new methods for HSI supervised classification is the evidence theory of Dempster-Schafer.

The new method for HSI supervised classification under contaminated training samples is proposed. This method differs from the known solutions of same problems by introducing a procedure of assessing classification value of the spectral bands by a special empirical function and by specific approach to partition of the spectral feature space, that allows to use only the most informative spectral bands.

A new effective method for HSI supervised classification under simultaneously contaminated training samples and restrictions on their volume was offered in the dissertation also.

Both proposed methods for HSI supervised classification are implemented algorithmically and on the software. Assessing the effectiveness of the developed methods was conducted using real satellite images obtained by the Hyperion/EO-1 on-board sensor.

Results show that the developed methods are more accurate than known method Support Vector Machine and algorithm for classifying objects based on spectral topological characteristics.

The proposed methods can be used for computer analysis of hyperspectral and multispectral satellite imagery in remote searching for minerals and hydrocarbons, solving environmental problems, etc.

Keywords: hyperspectral satellite image, image classification, evidence theory of Dempster-Shafer, training sample, the quality of a sample, classification accuracy.