

УДК 004.932

В.Г. Худов

Харківський національний університет радіоелектроніки, Харків

## ВИДІЛЕННЯ ГРАНИЦЬ СТРУКТУРОВАНИМ ДЕТЕКТОРОМ НА ЗОБРАЖЕННІ, ЩО ОТРИМАНЕ З БОРТОВОЇ СИСТЕМИ ОПТИКО-ЕЛЕКТРОННОГО СПОСТЕРЕЖЕННЯ

Встановлено, що результат обробки зображень, що отримані з бортових систем оптико-електронного спостереження, залежить від якості методу сегментування зображення, що, в свою чергу, поставляє перед розробниками систем обробки зображень проблему розробки методик, методів та вибору показників оцінки якості сегментування зображень. Розглянута можливість використання структурованого детектору з методом машинного навчання *Random Forest* для виділення границь на зображенні, що отримане з бортової системи оптико-електронного спостереження. Проведено експериментальне дослідження щодо виділення границь структурованим детектором на зображенні, що отримане з бортової системи оптико-електронного спостереження, з використанням машинного навчання *Random Forest*. Відмічаються основні переваги використання методу машинного навчання *Random Forest* при виділенні границь об'єктів на оптико-електронному зображенні.

**Ключові слова:** оптико-електронне зображення, виділення границь, метод, машинне навчання, бортова система, детектор, *Random Forest*, сегментування, дерево рішень, вибірка, навчання.

### Вступ

#### Постановка проблеми у загальному вигляді.

Відомо [1–4], що в теперішній час не існує загальної теорії оптимального представлення та обробки зображень. Вибір конкретної технології обробки зображень залежить від проблеми, що вирішується, та вимог, що пред'являються до результату обробки. Існує багато кількості практичних задач, що потребують виявлення візуально невидимих областей (об'єктів інтересу) слабоконтрастних зображень. Складність їх обробки, як правило, полягає в тому, що, з однієї аномалії частіше за все представляють невелику область, яку можна прийняти за шум або дефект зображення, а, з іншої сторони, параметри об'єкту інтересу можуть незначно відрізнятися один від одного та від фону, мати невідому форму та нечіткі границі.

Викладене вище стосується і зображень, що отримані з бортових систем оптико-електронного спостереження, що вирішуються завдання в інтересах безпеки і оборони [5–8]. Зростання потреби у даних, що отримуються з бортових систем оптико-електронного спостереження, супроводжується [7–9]:

- постійним збільшенням загального обсягу і доступності інформації бортових систем оптико-електронного спостереження;
- збільшенням інформації з високим розрізненням;
- впровадженням при вирішенні завдань інформаційного забезпечення в інтересах безпеки і оборони, комплексного використання даних відрізних джерел;

– збільшенням кількості суб'єктів у сфері отримання, розповсюдження, обробки та застосування даних, що отримані з бортових систем спостереження.

Результат обробки зображень, що отримані з бортових систем оптико-електронного спостереження, залежить від якості методу сегментування зображення, що, в свою чергу, поставляє перед розробниками систем обробки зображень проблему розробки методик, методів та вибору показників оцінки якості сегментування зображень.

**Аналіз останніх досягнень і публікацій.** Систематизація та аналіз ефективності сучасних методів сегментування зображення наведено в [10–14]. Виділяють наступні ознаки якісного сегментування [10]:

- однорідність області по характеристикам (в першу чергу, по кольору та текстурі);
- відмінність значень обраних характеристик для суміжних областей зображення;
- гладкість границь кожного сегменту зображення;
- незначна кількість «дірок» у сегменті.

Враховуючи перераховане вище, витікають три основні види можливих недоліків сегментування зображень, що отримані з бортових систем оптико-електронного спостереження [10]:

- неправильне сегментування, коли контури розподілу не співпадають з границями об'єктів на зображенні;
- пересегментування, коли має місце збільшений розподіл зображення на області;

– недосегментування, коли має місце недостатній розподіл зображення на області.

Зазвичай методи сегментування використовують декілька параметрів, підбираючи які можна уникнути останніх двох недоліків. Однак, перший недолік можна уникнути лише вибором методу сегментування. Відомо, що найбільш ефективні методи сегментування розроблені для конкретних завдань з урахуванням специфіки зображення [1–4; 10; 15–17].

**Мета статті** – розглянути можливість використання структурованого детектору з методом машинного навчання Random Forest для виділення границь на зображенні, що отримане з бортової системи оптико-електронного спостереження.

## Постановка задачі та викладення матеріалів дослідження

Метод машинного навчання Random Forest заснований на побудові великої кількості (ансамблю) дерев рішень (їх число – параметр методу), кожне з яких будується на виборці, що отримується з вихідної навчальної вибірки за допомогою вибірки з поверненням (bootstrap) [18]. При побудові кожного дерева на стадіях розщеплення вершин використовуються тільки фіксоване число ознак навчальної вибірки, що обираються випадково (ще один параметр методу), та будується повне дерево без відсічень. Класифікація відбувається за допомогою голосування класифікаторів, що визначаються окремими деревами. Точність (імовірність коректної класифікації) залежить від різноманіття класифікаторів, що складають ансамбль (від того, наскільки є кореляційними їх рішення). Чим більш різноманітні є класифікатори ансамблю (менше кореляція між їхніми рішеннями), тим вище імовірність коректної класифікації. В випадкових лісах рішення дерев, що їх складають, слабо корельовані завдяки подвійній «випадковості» в алгоритмах побудови випадкового лісу – на стадії вибірки з поверненням та на стадії випадкового відбору ознак, що використовуються при розщепленні вершин дерев.

Ансамбль класифікаторів представляє множину класифікаторів, рішення яких комбінуються деяким чином для отримання кінцевої класифікації спостережень. Існує декілька груп методів побудови ансамблів класифікаторів [18]:

- маніпулювання прикладами навчальної вибірки;
- маніпулювання ознаками;
- включення випадковості в індуктивний алгоритм;
- байєсове голосування.

Метод машинного навчання Random Forest використовує методи перших трьох груп. Перша група методів складається у використанні базового індуктивного алгоритму на різних підвибірках вихідної

навчальної вибірки, або в ітеративному переважанні спостережень. Найбільш простий спосіб формування підвибірок запропонований Брейманом [19]. Метод заснований на формуванні навчальної вибірки для кожного класифікатора за допомогою bootstrap (випадкової вибірки того ж об'єму, що й вихідна) з поверненням до вихідної навчальної вибірки та використанні методу голосування для агрегування рішень окремих класифікаторів. Цей метод отримав назву bagging або агрегований bootstrap [18; 19]. Указаний метод приводить до скорочення середньої квадратичної помилки класифікації. Однак, це вірно не для всіх базових моделей класифікаторів. До методів першої групи відносяться також метод, що заснований на формуванні диз'юнктивних підвибірок навчальної вибірки та використанні методики крос-валідації [20] для формування ансамблю класифікаторів.

Метод, що заснований на ітеративному переважанні спостережень навчальної вибірки, називається boosting [21]. В цьому методі класифікатори ансамблю будуються послідовно, і на кожній ітерації проводиться корекція (переважування) спостережень навчальної вибірки (на першій ітерації ваги всіх спостережень є рівними). Корекція здійснюється таким чином, щоб відповідний класифікатор робив менше помилок на тих спостереженнях, на яких частіше робили помилки класифікатори, що побудовані на попередніх ітераціях алгоритму.

До методів другої групи відноситься метод, в якому вихідна множина ознак розбивається на декілька диз'юнктивних множин та будується ансамбль нейронних мереж, кожна з яких включає ознаки тільки з однієї підмножини [22].

Методи третьої групи засновані на тому, що випадковість вводиться безпосередньо в базовий алгоритм.

Метод виділення границь структурованим детектором на зображенні з використанням машинного навчання Random Forest представимо у наступному вигляді:

1) для  $i = 1, 2, \dots, B$  ( $B$  – кількість дерев у ансамблі) необхідно:

- сформувати bootstrap вибірку  $S$  розміру  $l$  по вихідній навчальній вибірці  $D = \{x_i, y_i\}_{i=1}^l$ ;
- по bootstrap вибірці  $S$  індукувати неусічене дерево рішень  $T_i$  з мінімальною кількістю спостережень у вершинах, що дорівнює  $n_{\min}$ , за наступним рекурсивним алгоритмом:
  - з вихідного набору  $n$  ознак випадково обрати  $p$  ознак;
  - з  $p$  ознак обрати ознаку, яка забезпечує найкраще розщеплення;

– розщепити вибірку, що відповідає вершині, що оброблюються, на дві підвибірки;

2) в результаті виконання шагу 1 отримуємо ансамбль дерев рішень  $\{T_i\}_{i=1}^B$ ;

3) передбачення нових спостережень здійснюється наступним чином:

– для регресії (вираз (1)):

$$\hat{f}_{rf}^B(x) = \frac{1}{B} \sum_{i=1}^B T_i(x); \quad (1)$$

– для класифікації: нехай  $\hat{w}_i(x) \in \{w_1, w_2, \dots, w_c\}$  – клас, що обраний деревом рішень  $T_i$ , тобто  $T_i(x) = \hat{w}_i(x)$ ; тоді  $\hat{w}_{rf}^B(x)$  – клас, що найбільш частіше присутній в множині  $\{\hat{w}_b(x)\}_{i=1}^B$ .

Схематично використання машинного навчання Random Forest наведено на рис. 1.

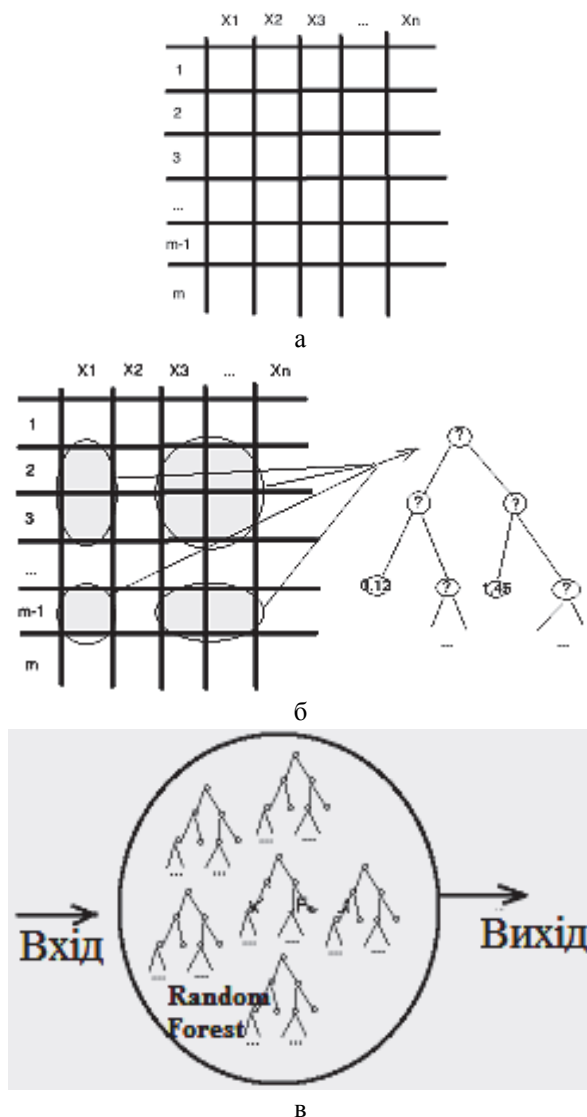


Рис. 1. Схематичне зображення використання методу машинного навчання Random Forest

На рис. 1, а зображено вхідні дані (кожен стовбець відповідає деякому параметру, кожна строчка відповідає деякому елементу). На рис. 1, б показано обирання випадковим чином з усього набору даних деякої кількості стовбців та строк та побудова по ним дерева прийняття рішень. Далі така процедура повторюється багато разів и отримується множина різних дерев (рис. 1, в). При цьому всі ці дерева є випадковими, тому що для створення кожного з них обиралося випадкова підмножина даних.

Проведемо виділення границь структурованим детектором на зображенні, що отримане з бортової системи оптико-електронного спостереження, з використанням машинного навчання Random Forest. У якості вихідного зображення оберемо зображення, що наведено на рис. 2 [23].



Рис. 2. Вихідне зображення [23]

На рис. 3. наведений результат виділення границь структурованим детектором на зображенні (рис. 2), що отримане з бортової системи оптико-електронного спостереження, з використанням машинного навчання Random Forest.



Рис. 3. Виділення границь структурованим детектором на зображенні з використанням машинного навчання Random Forest

## Висновки і напрями подальших досліджень

Таким чином, проведено виділення границь структурованим детектором на зображенні, що отримано з бортової системи оптико-електронного спостереження, з використанням машинного навчання Random Forest

Використання при виділенні границь структурованим детектором методу машинного навчання Random Forest має наступні переваги:

- висока точність класифікації;
- захист від «перепідгонки» (випадки, коли класифікатор добре класифікує спостереження навчальної вибірки, але непридатний для класифікації спостережень, що не входять до неї) навіть у випадку, коли кількість ознак значно перевищує кількість спостережень;

– для побудови Random Forest по навчальній вибірці необхідно завдання всього двох параметрів, які вимагають мінімального налаштування;

– метод out-of-bag [24] забезпечує отримання природної оцінки імовірності помилкової класифікації Random Forest на основі спостережень, що не входять у навчальні bootstrap вибірки, що використовуються для побудови дерев;

– навчальна вибірка для побудови Random Forest може вміщувати ознаки, що виміряні у різних шкалах: числовий, порядковий та номінальний, що є неприпустимим для багатьох інших класифікаторів.

Напрямок подальших досліджень є оцінка ефективності сегментування багатомасштабної послідовності зображень, що отримані з космічних систем оптико-електронного спостереження, удосконаленням методом сегментування.

## Список літератури

1. Wang Y.S. A New Image Threshold Segmentation based on Fuzzy Entropy and Improved Intelligent Optimization Algorithm / Y.S. Wang // *Journal of Multimedia*. – 2014. – Vol. 9, № 4. – P. 499-505.
2. Zhu S.J. Rival Penalized Image Segmentation / S.J. Zhu, J.Y. Zhao, L.J. Guo // *Journal of Multimedia*. – 2014. – Vol. 9, № 5. – P. 736-745.
3. Злобин В.К. Обработка аэрокосмических изображений / В.К. Злобин, В.В. Еремеев. – М.: ФИЗМАТЛИТ, 2006. – 288 с.
4. Ахметшина Л.Г. Сегментация слабоконтрастных изображений на основе фазификации и ортогональной декомпозиции / Л.Г. Ахметшина // *Теорія прийняття рішень: Праці VIII школи-семінару ДВНЗ «Ужгородський національний університет» (Ужгород 26 вересня – 1 жовтня 2016)*. – Ужгород: ДВНЗ «Ужгородський національний університет, 2016. – С. 36-37.
5. Алімпієв А.М. Особливості гібридної війни РФ проти України. Досвід, що отриманий Повітряними Силами Збройних Сил України / А.М. Алімпієв, Г.В. Певцов //

*Наука і техніка Повітряних Сил Збройних Сил України*. – 2017. – № 2 (27). – С. 19-25.

6. Світова гібридна війна: український фронт: моногр. / за заг. ред. В.П. Горбуліна. – К.: НІСД, 2017. – 496 с.

7. Гудима О.П. Проблеми використання інформації космічних систем та удосконалення геоінформаційного забезпечення при виконанні завдань АТО / О.П. Гудима, Г.В. Певцов, М.Ф. Пічугін // *Новітні технології – для захисту повітряного простору: Тези допов. 13 наук. конф. Харк. нац. ун-ту Повітряних Сил ім. Івана Кожедуба (Харків 12–13 квіт. 2017)*. – Х.: ХНУПС, 2017. – С. 452.

8. Кухарський І.А. Використання супутникових даних в інтересах ведення антитерористичних операцій / І.А. Кухарський, Г.В. Худов, В.О. Павлій // *Новітні технології – для захисту повітряного простору: Тези допов. 13 наук. конф. Харк. нац. ун-ту Повітряних Сил ім. Івана Кожедуба (Харків 12–13 квіт. 2017)*. – Х.: ХНУПС, 2017. – С. 453-454.

9. Болобан С.І. Підхід до оцінювання якості космічних знімків щодо можливості їх дешифрування / С.І. Болобан, Р.М. Осадчук, О.М. Перегуда, І.В. Рассохіна // *Збірник наукових праць ЖВІ НАУ*. – 2012. – Вип. 6. – С. 107-112.

10. Барталев С.А. Анализ возможностей применения методов сегментации спутниковых изображений для выявления изменений в лесах / С.А. Барталев, Т.С. Ховратович // *Современные проблемы дистанционного зондирования Земли из космоса*. – 2011. – Т. 8. № 1. – С. 44-62.

11. Смеляков К.С. Модели и методы сегментации границ изображений нерегулярного вида на основе адаптивных масок: дис. ... канд. техн. наук: 09.03.05 / Смеляков Кирилл Сергеевич. – Х., 2005. – 162 с.

12. Смеляков К.С. Методы сегментации изображений объектов нерегулярного вида, особенности их применения и перспективы развития / К.С. Смеляков, И.А. Романенко, И.В. Рубан, Н.И. Кириллова, О.В. Шитова // *Збірник наукових праць ХУПС*. – 2010. – Вип. 2 (24). – С. 92-97.

13. Стругайло В.В. Обзор методов фильтрации и сегментации цифровых изображений / В.В. Стругайло // *Наука и образование. Научное издание МГТУ им. Н.Э.Баумана. [Электронный ресурс]*. – URL: <http://technomag.edu.ru/doc/411847.html/> (дата обращения 01.07.2017).

14. Самоїленко Д.Е. Структурная сегментация изображений / Д.Е. Самоїленко // *Штучний інтелект*. – 2004. – № 4. – С. 521-528.

15. Барталев С.А. Исследование возможностей оценки состояния поврежденных пожарами лесов по данным многоспектральных спутниковых измерений / С.А. Барталев, В.А. Егоров, А.М. Крылов, Ф.В. Стыценко, Т.С. Ховратович // *Современные проблемы дистанционного зондирования Земли из космоса*. – 2010. – Т. 7, № 3. – С. 215-225.

16. Златопольский А.А. Выделение на изображении однородных участков с неполными границами / А.А. Златопольский // *Исследование Земли из космоса*. – 1985. – № 1. – С. 94-102.

17. Левашикина А.О. Исследование супервизорных критериев оценки качества сегментации / А.О. Левашикина, С.В. Поршнев // *Известия Томского политехнического университета*. – 2008. – Т. 313, № 5. – С. 28-33.

18. Чистяков С.П. Случайные леса: обзор / С.П. Чистяков // Труды Карельского научного центра РАН. – 2013. – № 1. – С. 117-136.
19. Breiman L. Bagging predictors / L. Breiman // Machine Learning. – 1996. – Vol. 24, № 2. – P. 123-140.
20. Parmanto B. Improving committee diagnosis with resampling technique / B. Parmanto, P.W. Munro, H.R. Doyle // Advanced in Neural Information Processing Systems. – 1996. – Vol. 8. – P. 146-148.
21. Freud Y. Experiments with a new boosting algorithm / Y. Freud, R.E. Shapire // Proceedings of the 13rd International Conference on Machine Learning. – P. Morgan Cauffman, 1996. – P. 15-21.
22. Cherkauer K.G. Human expert-level performance on a scientific image analysis task by a system using combined artificial neural networks. Eds. P. Chan Working Notes on the AAAI Workshop on Integrating Multiple Learned Models. 2006. – P. 15-21.
23. [Електронний ресурс]. Режим доступу <http://www.satimagingcorp.com/gallery/ikonos>.
24. Breiman L. Out-of-bag estimation / L. Breiman // Technical report, Statistics Department University of California. – Berkeley. 1996. – P. 1-13.

Надійшла до редколегії 6.07.2017

**Рецензент:** д-р техн. наук проф. І.В.Рубан, Харківський національний університет радіоелектроніки, Харків.

## ВЫДЕЛЕНИЕ ГРАНИЦ СТРУКТУРИРОВАННЫМ ДЕТЕКТОРОМ НА ИЗОБРАЖЕНИИ, ПОЛУЧЕННОМ С БОРТОВОЙ СИСТЕМЫ ОПТИКО-ЭЛЕКТРОННОГО НАБЛЮДЕНИЯ

В.Г. Худов

Установлено, что результат обработки изображений, которые получены с бортовых систем оптико-электронного наблюдения, зависит от качества метода сегментации изображения, что, в свою очередь, ставит перед разработчиками систем обработки изображений проблему разработки методик и выбора показателей оценки качества сегментации изображений. Рассмотрена возможность использования структурированного детектора с методом машинного обучения Random Forest для выделения границ на изображении, полученном с бортовой системы оптико-электронного наблюдения. Проведено экспериментальное исследование выделения границ структурированным детектором на изображении, полученном с бортовой системы оптико-электронного наблюдения, с использованием машинного обучения Random Forest. Отмечаются основные преимущества использования метода машинного обучения Random Forest при выделении границ на оптико-электронном изображении.

**Ключевые слова:** оптико-электронное изображение, выделение границ, метод, машинное обучение, бортовая система, детектор, Random Forest, сегментация, дерево решений, выборка, обучение.

## ALLOCATION OF EDGES BY A STRUCTURED DETECTOR ON AN IMAGE OBTAINED FROM A SHIPBORNE OPTICAL-ELECTRONIC SURVEILLANCE SYSTEM

V. Khudov

It has been established that the result of processing images obtained from on-board optical-electronic surveillance systems depends on the quality of the image segmentation method, which in turn poses the problem of the development of techniques and the selection of indicators for assessing the quality of image segmentation before the developers of imaging systems. The possibility of using a structured detector with the method of machine learning Random Forest for isolating boundaries on an image obtained from an on-board optical-electronic surveillance system is considered. An experimental study of the delineation of boundaries by a structured detector in an image obtained from an on-board optical-electronic surveillance system using Random Forest's computer training was carried out. The main advantages of using the method of machine learning Random Forest when allocating borders on the opto-electronic image are noted.

**Keywords:** optical-electronic image, border allocation, method, machine learning, on-board system, detector, Random Forest, segmentation, decision tree, sampling, training.