

ІНТЕЛЕКТУАЛЬНА СИСТЕМА ВИЯВЛЕННЯ БЛИЗЬКОРОЗТАШОВАНИХ МАЛОРОЗМІРНИХ ОБ'ЄКТІВ: АРХІТЕКТУРА ТА РЕАЛІЗАЦІЯ

Крохмаль А.В., Стрілкова Т.О.

Кафедра мікроелектроніки, електронних приладів та пристроїв,
Харківський національний університет радіоелектроніки,
Україна

E-mail: andrii.krokhmal@nure.ua
tetiana.strilkova@nure.ua

Abstract

The article analyzes the development of an image formation simulation model for optoelectronic systems and a combined method for detecting small, low-contrast objects in digital images. The proposed model enables the generation of synthetic datasets that account for physical and technical factors affecting image quality, providing a valuable tool for testing image processing algorithms. Additionally, the detection method, based on an adapted MobileNet V1 architecture, demonstrates high accuracy and robustness to noise under limited computational resources, making it suitable for embedded applications.

Вступ

Сучасні військові загрози, з якими стикається Україна, характеризуються масованим застосуванням зокрема безпілотних літальних апаратів, високоточних боєприпасів та роїв дронів-камікадзе. Їх використання проти критичної інфраструктури створює серйозні виклики для систем протиповітряної оборони. У цих умовах особливе значення набувають системи технічного зору, здатні до точного й швидкого аналізу середовища. Вони забезпечують сприйняття, аналіз та інтерпретацію візуальних даних у режимі реального часу, що є необхідною умовою функціонування автономних комплексів, зокрема систем підтримки прийняття рішень. Одним із ключових викликів є виявлення близькорозташованих малорозмірних об'єктів на зображеннях, які можуть бути як елементами атаки, так і індикаторами її здійснення. Традиційні методи моніторингу часто не дають належної точності при роботі з малими об'єктами, що створює ризик пропуску критично важливих сигналів. Використання технологій штучного інтелекту, машинного навчання та обробки сигналів дозволяє створювати адаптивні системи, здатні до самонавчання та підвищення ефективності виявлення. Отже, дослідження, спрямовані на розвиток систем технічного зору, є актуальним науково-технічним завданням, яке має вагомий практичний значення.

Метою дослідження є аналіз проблеми виявлення близькорозташованих об'єктів у системах технічного зору для захисту критичної інфраструктури та визначення архітектурних рішень, здатних забезпечити стійкість і ефективність роботи в складних умовах.

Для розробки та удосконалення інтелектуальної системи обробки зображень при виявленні малорозмірних та малокоонтрастних об'єктів виникає необхідність у синтезі штучних зображень, що формуються системою технічного зору відповідно до реальних умов спостереження. Для цього було розроблено імітаційну модель відгуку багатоелементного фотоприймача системи технічного зору. Модель побудована на основі стохастико-детермінованого підходу та ґрунтується на хвильовій та корпускулярній теорії світла. Принципи формування сигналів в системах технічного зору розглядаються як стохастичний процес. Стохастико-детермінований підхід до обробки сигналів є одним з найпродуктивніших шляхів оцінки ефективності систем технічного зору формування та обробки сигналів та зображень [1-4].

Імітаційне моделювання відгуку фотоприймача

Запропонована модель базується на послідовному врахуванні фізичних процесів в електронних системах, що визначають формування відгуку багатоеlementних фотоприймачів. На першому етапі розглядається оптична передавальна функція (OTF) апертури, яка визначається автокореляцією функції пропускання. Для прямокутної апертури її можна представити як добуток одномірних трикутних функцій, що відображає обмеження просторових частот. Також враховується атмосферна турбулентність при великому інтервалі експозиції за моделлю Колмогорова [5]. Повна OTF системи визначається як добуток апертурної та атмосферної функцій.

$$\text{OTF}_{\text{ап}}(f_x, f_y) = \left(1 - \frac{|f_x|}{f_{c,x}}\right)_+ \cdot \left(1 - \frac{|f_y|}{f_{c,y}}\right)_+, \quad f_{c,x} = \frac{a}{\lambda}, \quad f_{c,y} = \frac{b}{\lambda},$$

$$\text{OTF}_{\text{атм}}(f) = \exp\left[-3.44 \left(\frac{\lambda f}{r_0}\right)^{5/3}\right], \quad f = \sqrt{f_x^2 + f_y^2},$$

$$\text{OTF}_{\text{сист}}(f_x, f_y) = \text{OTF}_{\text{ап}}(f_x, f_y) \cdot \text{OTF}_{\text{атм}}(f)$$

Функція розсіювання точки (PSF) отримується оберненим перетворенням Фур'є від OTF (рис. 1). Вона визначає просторове розмиття точкового джерела на зображенні. Формування зображення описується згортою початкового розподілу інтенсивності $I_{\text{вихідн}}$ з PSF [6].

$$\text{PSF}(x, y) = \mathbf{F}^{-1}\{\text{OTF}_{\text{сист}}(f_x, f_y)\},$$

$$I_{\text{вихідн}}(x, y) = I_{\text{вхідн}}(x, y) * \text{PSF}(x, y),$$

де \mathbf{F}^{-1} - обернене перетворення Фур'є, $I_{\text{вхідн}}$ - початковий розподіл інтенсивності,

* - двовимірна згортка.

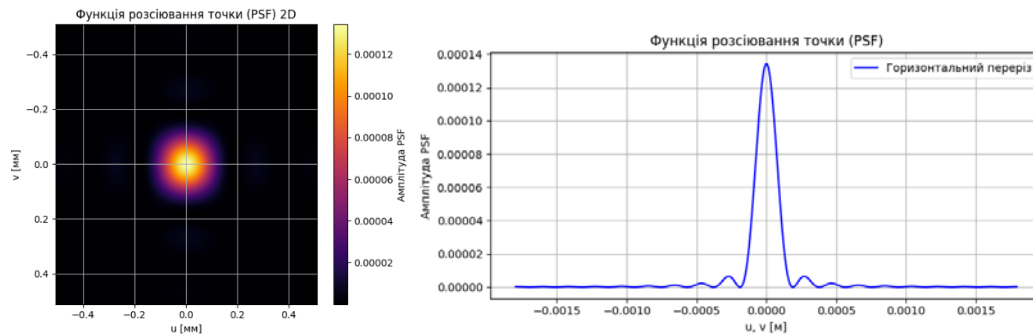


Рис. 1 – Моделювання функції розсіювання точки оптичної системи

Подальші етапи моделі враховують специфіку роботи багатоеlementного фотоприймача. По-перше, враховано оптичні втрати, зумовлені відбиттям, абсорбцією та іншими чинниками. Фотонний шум моделюється як пуассонівський процес, що відображає стохастичний характер надходження фотонів. Далі моделюється світлова характеристика багатоеlementного фотоприймача. При генерації фотоелектронів враховується квантова ефективність фотоприймача. Останній має скінчену ємність пікселя, яка визначає межу насичення: у робочій зоні відгук є лінійним, але поблизу насичення плавно переходить у плато [7].

Також модель враховує темновий струм, який виникає за відсутності світла внаслідок теплової генерації носіїв заряду. Його середнє значення пропорційне часу експозиції та залежить від температури, шум темного струму моделюється пуассонівським розподілом. Після зчитування сигнал підсилюється аналоговим підсилювачем із заданим коефіцієнтом. Даний процес супроводжується шумом зчитування, який було змодельовано як адитивний гаусівський шум [8].

Вихідний аналоговий сигнал обмежується діапазоном напруг багатоеlementного фотоприймача та переводиться у цифрові одиниці (ADU). Для цього застосовується аналогово-цифровий

перетворювач, що здійснює округлення до найближчого дискретного рівня, число яких визначається його розрядністю. Якщо сигнал виходить за межі діапазону, відбувається насичення АЦП. В результаті формується цифрове зображення, яке може бути збережене у стандартних растрових форматах (8-bit PNG, 12-bit RAW, 16-bit TIFF тощо).

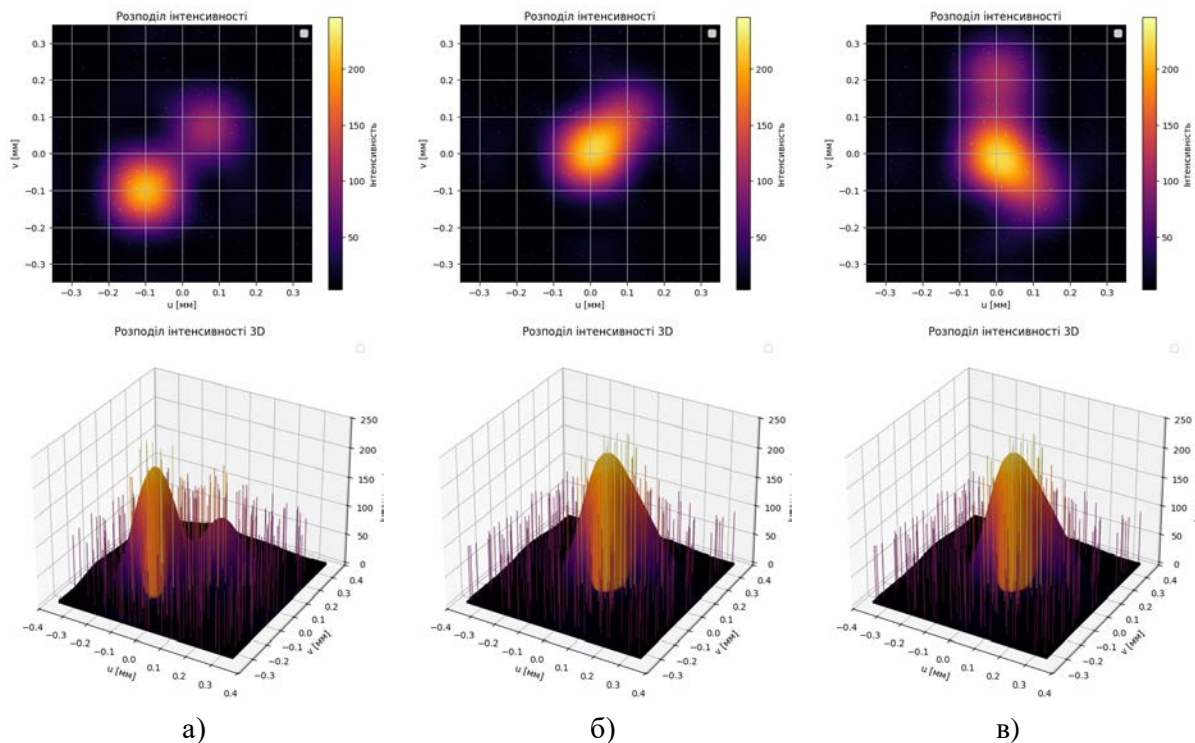


Рис. 2 – Результати моделювання Дифракційні плями близькорозташованих об'єктів: двох об'єктів на більшій (а) та меншій (б) відстанях, трьох об'єктів на різних відстанях (в)

Таким чином, розроблена імітаційна модель формування зображень в системах технічного зору дозволяє моделювати вихідне зображення на основі інтенсивності та положення спостережуваних близькорозташованих об'єктів (рис. 2). Модель враховує такі фактори, як дифракція Фраунгофера на апертурі, довгоекспозиційна атмосферна турбулентність, зовнішні та внутрішні шуми (темновий струм, фотонний шум та шум зчитування), нелінійність світлової характеристики фотоприймача, температурна залежність темного струму та вплив аналогово-цифрового перетворювача. Основними недоліками є можливість моделювання лише монохроматичних світлових сигналів, неврахування відстані до спостережуваного об'єкта, можливість моделювання впливу атмосферної турбулентності лише за умови, що інтервал експозиції набагато більше періоду «замороженості» атмосфери, а також врахування лише хвильової природи світла.

Розробка методу виявлення об'єктів

Було розроблено комбінований метод реєстрації близькорозташованих малорозмірних та мало контрастних об'єктів, що дозволяє визначити кількість об'єктів на зображенні, а також математичне сподівання координат центру для кожного з об'єктів (рис. 3). Він є комбінацією глибоких згорткових нейронних мереж, кожна з яких виконує специфічну задачу. За основу взято архітектуру MobileNet V1 [9], що має суттєві переваги для застосування у вбудованих системах завдяки своїй обчислювальній ефективності. Запропонований метод реєстрації об'єктів на цифрових зображеннях складається з послідовних етапів, кожен з яких реалізує окрему підзадачу обробки вхідного зображення. На першому етапі здійснюється виділення ознак за допомогою адаптованої архітектури MobileNet V1, де перша згортка модифікована для обробки зображення одного об'єкта.

Ознаки вилучаються з прихованого шару попередньо навченої нейронної мережі на наборі даних ImageNet, що дозволяє сформувати вектор ознак сталої розмірності, який узагальнює візуальні характеристики зображення. Такий підхід дає можливість сформувати стійкі до шумів масштабно-незалежні інтенсивнісні та просторові характеристики, які є ефективними для подальшої обробки.

Другий етап передбачає визначення кількості об'єктів на зображенні як вирішення задачі класифікації. Для цього використовується згортова нейронна мережа, побудована на основі наступних шарів MobileNet V1, з модифікованою класифікаційною «головою». Після навчання на синтетичному наборі даних модель показала високу точність класифікації: 0.917, а аналіз метрик (табл. 1) свідчить про її здатність до узагальнення та відсутність значної переваги одного класу над іншими.

На третьому етапі виконується регресія координат центрів виявлених об'єктів. Зважаючи на змінну кількість об'єктів, застосовуються окремі моделі-регресори для кожного можливого числа об'єктів (від 0 до 5) з відповідною розмірністю вихідного шару. Архітектура моделей, аналогічно, базується на незадіяних шарах MobileNet V1, з глибшою регресійною «головою», що включає два повнозв'язні шари. Для забезпечення пермутаційної інваріантності використано модифіковану середньоквадратичну функцію втрат у поєднанні з алгоритмом Hungarian matching [10]. Результати тестування (табл. 1) демонструють високу точність для поодиноких об'єктів, однак зі збільшенням їх кількості спостерігається монотонне погіршення метрик. Найбільший деградаційний стрибок спостерігається при переході від одного до двох об'єктів, що зумовлено появою ситуації їх близького розташування. Загалом, модель є адекватною для локації невеликої кількості об'єктів, однак її масштабування потребує подальшого удосконалення архітектури та процесу навчання.

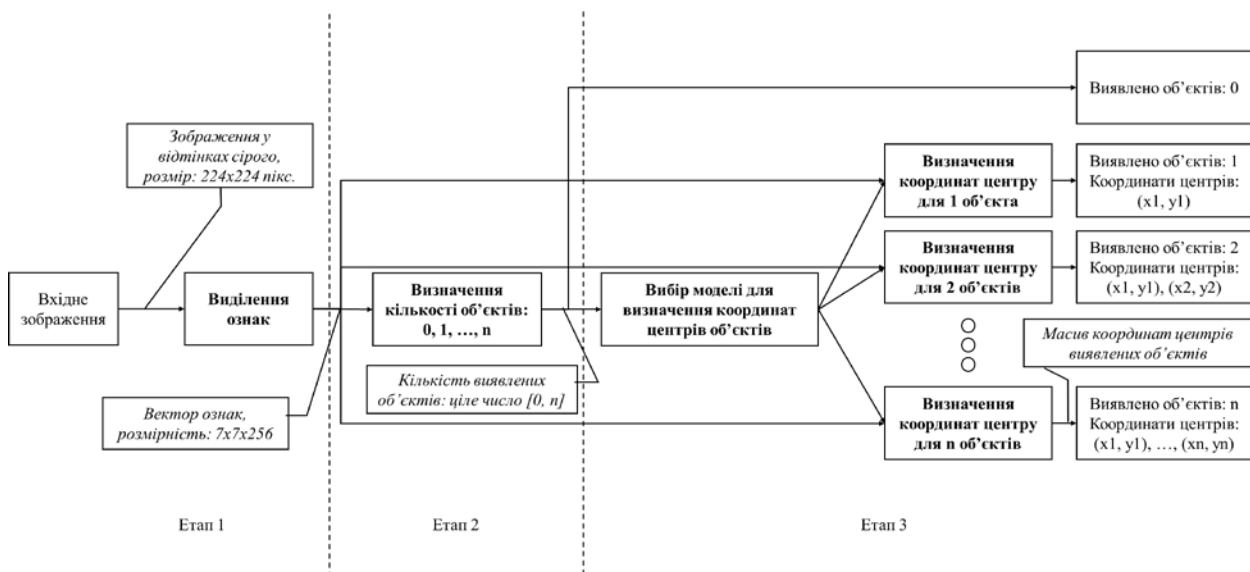


Рис. 3. Архітектура та реалізація інтелектуальної системи виявлення близькорозташованих малорозмірних об'єктів

Таблиця 1. Метрики якості роботи системи виявлення та локації

Кількість сигналів	Метрики визначення кількості виявлених об'єктів (класифікації)			Метрики визначення координат центрів об'єктів (регресія)		
	Precision	Recall	F1-score	MSE	MAE	R ²
0	0.98	0.99	0.99	-	-	-
1	0.96	0.96	0.96	0.000012	0.001565	0.9742
2	0.93	0.93	0.93	0.000039	0.004532	0.9147
3	0.88	0.90	0.89	0.000053	0.005465	0.8863
4	0.85	0.83	0.84	0.000065	0.006117	0.8602
5	0.91	0.88	0.90	0.000068	0.006281	0.8526

Висновки

В результаті проведеного дослідження отримано наступні основні висновки: запропоновано імітаційну модель формування зображень в системах технічного зору, що враховує комплекс фізичних явищ (дифракцію, атмосферну турбулентність, шумові процеси) та інші особливості роботи фотоприймача. Враховано вплив цих факторів на процес формування зображення, що надало можливість генерації великих наборів синтетичних даних для навчання та тестування систем, що є критично важливим за умов обмежених експериментальних даних. Розроблено комбінований метод виявлення та локації об'єктів, зокрема близько розташованих, малорозмірних та малоконтрастних, на основі глибоких згорткових нейронних мереж. Він поєднує високу стійкість до шумів та достатню точність із низькими вимогами до обчислювальних ресурсів, що робить його перспективним для застосування у вбудованих та автономних системах технічного зору. Основною перевагою розробленого методу є можливість розрізнення близько розташованих об'єктів завдяки врахуванню форми дифракційної плями, що стало можливим завдяки використанню методів глибокого навчання. Це також робить метод стійким до шумів та флуктуацій, отже, доцільним для використання в складних умовах спостереження. Метод є адаптивним та може бути переналаштований під параметри конкретної системи технічного зору для забезпечення максимальної точності виявлення. Завдяки поетапному виконанню запропонований метод дозволяє здійснювати виявлення та локацію з досить високою точністю за умов обмежених обсягів оперативної пам'яті та низької швидкодії процесора. Серед недоліків методу слід зазначити можливість виявлення та локації фіксованої кількості об'єктів, низьку роздільну здатність зображення, що аналізується (224x224 пікселі) та обмежену ефективність в реальних умовах через використання лише синтетичних даних для навчання нейронних мереж.

Література

1. Strelkova T.A., Lytyuga A.P., Kalmykov A.S. Statistical Characteristics of Optical Signals and Images in Machine Vision Systems Examining Optoelectronics in Machine Vision and Applications in Industry 4.0. 2021, Pages: 134-162. DOI: 10.4018/978-1-7998-6522-3.ch005.
2. Strelkova T.A., Strelkov A.I., Lytyuga A.P., Kalmykov A.S. Methods of Reception and Signal Processing in Machine Vision Systems in Examining Optoelectronics in Machine Vision and Applications in Industry 4.0., 2021, Pages: 71-102. DOI: 10.4018/978-1-7998-6522-3.ch003.
3. Стрількова Т. О., Крохмаль А. В. Формування зображень в оптико-електронних системах з урахуванням умов спостереження / Т. О. Стрількова, А. В. Крохмаль // Радіоелектроніка та молодь у XXI столітті : Матеріали 29-го Міжнародного молодіжного форуму, 16–19 квітня 2025 р. Харків: ХНУРЕ, 2025. – Т. 1. – С. 24–26.
4. Krokmal V., Krokmal A., Tarasov V., Rudniev Ye. Application of the combined method of raster image recognition in the computer vision system of unmanned vehicles in the mining industry // Geo-Technical Mechanics. 2024. № 171. DOI: 10.15407/geotm2024.171.098.
5. Ismail, M.Mohamed. (2015). A Study on atmospheric turbulence with Shearing Interferometer wavefront sensor. International Journal of Innovative Research in Science, Engineering and Technology. 04. 991-999. DOI: 10.15680/IJRSET.2015.0403039.
6. Fiete, R.D., Paul, B.D., 2014. Modeling the optical transfer function in the imaging chain. Opt. Eng. 53(8), 083103. 29 p. DOI: 10.1117/1.OE.53.8.083103.
7. Abarca, A.; Theuwissen, A. A CMOS Image Sensor Dark Current Compensation Using In-Pixel Temperature Sensors. Sensors 2023, 23, 9109. DOI: 10.3390/s23229109.
8. Bigas, M. & Cabruja, Enric & Forest, Josep & Salvi, Joaquim. (2006). Review of CMOS image sensors. Microelectronics Journal. 37. 433-451. DOI: 10.1016/j.mejo.2005.07.002.
9. Howard, Andrew & Zhu, Menglong & Chen, Bo & Kalenichenko, Dmitry & Wang, Weijun & Weyand, Tobias & Andreetto, Marco & Adam, Hartwig. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. DOI: 10.48550/arXiv.1704.04861.
10. Chen, Lingji & Sharma, Alok & Shirore, Chinmay & Zhang, Chengjie & Buddharaju, Balarama. (2022). How to Backpropagate through Hungarian in Your DETR?. DOI: 10.48550/arXiv.2211.14448.