

УДК 681.5(042.3)

Oleksiy Semko¹, PhD, Researcher of Information Security Department
ORCID ID: <https://orcid.org/0000-0001-6473-1329> **e-mail:** semalek@meta.ua

Denys Vinnichuk², PhD student
ORCID ID: <https://orcid.org/0009-0001-9441-7696> **e-mail:** vinniboy2709@gmail.com

¹Institute of Telecommunications and Global Information Space of NASU, Kyiv, Ukraine

²National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine

HYBRID METHOD FOR TRACKING MOVING OBJECTS IN VIDEO STREAMS UNDER DYNAMIC OBSERVATION CONDITIONS

Abstract. *A hybrid method for tracking objects in dynamic observation is proposed. Such conditions arise when the camera rotates and changes the zoom factor. The method uses the metadata of the camera position, including pitch, yaw, roll angles and zoom factors. The YOLO v8 artificial convolutional neural network was used as a detector. The main method of tracking the two is the stage of comparing detections and previous trajectories. At the first stage, the matching performed using the area overlap coefficient in pixel coordinates. At the second stage, the calculation of global coordinates is used based on the position of the object in the frame and the camera metadata. These global coordinates are using to predict the next position and compare with previous trajectories. The proposed method also allows determining the position of objects using data from altimeter sensors. The advantage of the proposed approach is the possibility of observation during sharp camera movements and changes in perspective. The method was experimentally tested on complex dynamic traffic scenes. The proposed method demonstrates higher metrics $IDF1 = 0.84$ and $MOTA = 0.81$ than standard algorithms on complex dynamic scenes. The method can be used in dynamic surveillance systems.*

Keywords: *computer vision, convolutional neural networks, classification, tracking of moving objects, dynamic observation.*

О.В. Семко¹, Д.О. Віннічук²

¹Інститут телекомунікацій і глобального інформаційного простору НАН України, м. Київ, Україна

²Національний університет біоресурсів і природокористування України, м. Київ, Україна

ГІБРИДНИЙ МЕТОД ВІДСТЕЖЕННЯ РУХОМИХ ОБ'ЄКТІВ У ВІДЕОПОТОЦІ В УМОВАХ ДИНАМІЧНОГО СПОСТЕРЕЖЕННЯ

Анотація. *Запропоновано гібридний метод відстеження об'єктів при динамічному спостереженні. Це умови, коли камера обертається та змінює коефіцієнт зумування. Метод використовує метадані положення камери, а саме кути тангажу, курсу та крену, а також коефіцієнт зумування. Як детектор було використано штучну згорткову нейронну мережу YOLO v8. Основною методу відстеження є двохетапне зіставлення детекцій та попередніх траєкторій. На першому етапі зіставлення відбувається класично за допомогою коефіцієнта площі перекриття в піксельних координатах. На другому етапі використовується розрахунок глобальних координат на основі*

положення об'єкта на кадрі та метаданих камери. Ці глобальні координати використовуються для прогнозування наступного положення та зіставлення з попередніми траєкторіями. Запропонований метод також дозволяє визначати положення об'єктів відносно камери за допомогою даних від датчиків висотоміра. Перевагою запропонованого підходу є можливість здійснювати спостереження при різких рухах камери та зміні ракурсу. Було проведено експериментальне випробування методу на складних динамічних сценах дорожнього руху. Запропонований метод демонструє вищі метрики $IDF1 = 0.84$ та $MOTA = 0.81$, ніж стандартні алгоритми на складних динамічних сценах. Метод може бути використаний у системах динамічного спостереження.

Ключові слова: комп'ютерний зір, згорткові нейронні мережі, класифікація, відстеження рухомих об'єктів, динамічне спостереження.

<https://doi.org/10.32347/2411-4049.2026.2.230-237>

Вступ

Технології комп'ютерного зору активно використовуються в системах безпеки, робототехніці, автономному транспорті, військовій справі тощо [1]. Зокрема, актуальним є завдання автоматичного відеоспостереження та інтелектуального аналізу просторових сцен, де наявні рухомі цілі. Виконання цих задач стає можливим як завдяки широкій доступності і здешевленню виробництва камер, так і значному збільшенню обчислювальних потужностей для обробки даних. Іншим фактором є значний прогрес у використанні методів штучного інтелекту. Основними етапами інтелектуального аналізу просторових сцен є: детекція, класифікація, відстеження, визначення характеристик об'єктів.

Особливу роль відіграє вирішення задачі відстеження об'єктів між кадрами відеопотоку, що дозволяє отримати відомості про поведінку об'єкта спостереження в часі та синтезувати оптимальну стратегію управління відповідно до функції цілі.

Аналіз літературних даних. Згорткові нейронні мережі (ЗНМ) зарекомендували себе як потужний інструмент, який здатний ефективно вирішувати задачі як детекції, так і класифікації [2]. Перевагою ЗНМ є здатність виділяти ознаки елементів зображення за умов зміни масштабу зображення і завад. При цьому слід враховувати те, що різноманітні архітектури моделювання інтелектуальної діяльності дозволяють визначити співвідношення між швидкістю, продуктивністю, точністю і ефективністю залежно від особливостей цієї задачі. Одноетапні згорткові нейронні мережі, такі як YOLO [3], дозволяють працювати в режимі реального часу за рахунок одноразової обробки зображення. Також слід враховувати те, що двоетапні нейронні мережі демонструють кращі значення основних метрик, але більший час швидкодії.

Такі характеристики переміщення об'єктів в просторі спостереження, як швидкість або траєкторія руху практично неможливо визначити без вирішення задачі відстеження, що дозволяє призначити об'єктам спостереження унікальні символічні ідентифікатори, які зберігаються в послідовності кадрів відеопотоку [4]. Таким чином, встановлюється відповідність між спостереженнями одного і того ж об'єкта і накопичується інформація про його переміщення. Цю задачу вирішують шляхом використання трекерів. Основним методом при вирішенні

такої задачі є “відстеження за детекцією”, тобто зіставлення результатів роботи детектора на окремих кадрах відеопотоку між собою. Процес відстеження умовно можна розділити на декілька етапів:

- прогнозування розташування об’єкта на кадрі з урахуванням інформації з попередніх кадрів;
- зіставлення з новими даними від детектора за допомогою деякої метрики та алгоритмів оптимізації;
- видалення незіставлених траєкторій та створення нових траєкторій.

Прикладами сучасних трекерів є ByteTrack [5], SORT, DeepSORT [6]. Сучасні трекери можуть використовувати як класичні, так і нейромереві алгоритмічні підходи для асоціації об’єктів та траєкторій.

Мета та задачі дослідження. Сучасні трекери переважно використовують тільки піксельне відображення сцени. Однак, одним з найбільших їх недоліків є неврахування динаміки спостереження, наприклад при обертанні камери. У такому разі обмежувальна рамка буде активно змінювати свої координати та розміри попри те, що об’єкт спостереження не змінював своє положення, а змінювався лише ракурс камери. Це може призвести до хибної зміни ідентифікатора, хибного визначення траєкторії та швидкості об’єкта спостереження. Для унеможливлення цього доцільним є використання трекеру, який враховує динаміку спостереження, а саме використовує простір ознак, що сформований з урахуванням обертання та зумування камери, що дозволяє вирішувати задачу відстеження об’єкта при різких рухах камери, або при тимчасовій втраті об’єкта спостереження.

Методика врахування особливостей динамічного спостереження

Як основний детектор для отримання координат обмежувальних рамок було обрано архітектуру *YOLOv8* [7]. Зазвичай під час спостереження камера активно обертається та виконує зумування. Обертання може виконуватись за осями курсу (*Yaw*), тангажу (*Pitch*), крену (*Roll*) (Рис. 1). Зазвичай послідовність осей один від одного визначається в порядку *Yaw* -> *Pitch* -> *Roll*, що зумовлене конструкцією шарнірів сервоприводів, які обертають камеру. Однак воно може бути і інше в залежності від конструкції камери. Важливим параметром камери є фокусна відстань, яка відображає відстань від оптичного центру об’єктива до сенсора. Коротка фокусна відстань (одиниці мм) дозволяє отримати широкий кут зору, охоплюючи багато об’єктів з меншою деталізацією. Довша фокусна відстань (десятки мм) дозволяє отримати вузький кут зору, що може бути використано для спостереження за віддаленими об’єктами.

Камера може активно змінювати коефіцієнт зумування *zoom*, що фізично відображається як зміна фокусної відстані камери:

$$\begin{cases} f_{new} = f * zoom \\ \operatorname{tg}(\xi / 2) = \frac{w}{2f} \end{cases}, \quad (1)$$

де f – фокусна відстань без зумування; f_{new} – фокусна відстань з урахуванням зумування; ξ – кут поля зору камери; w – ширина сенсора.

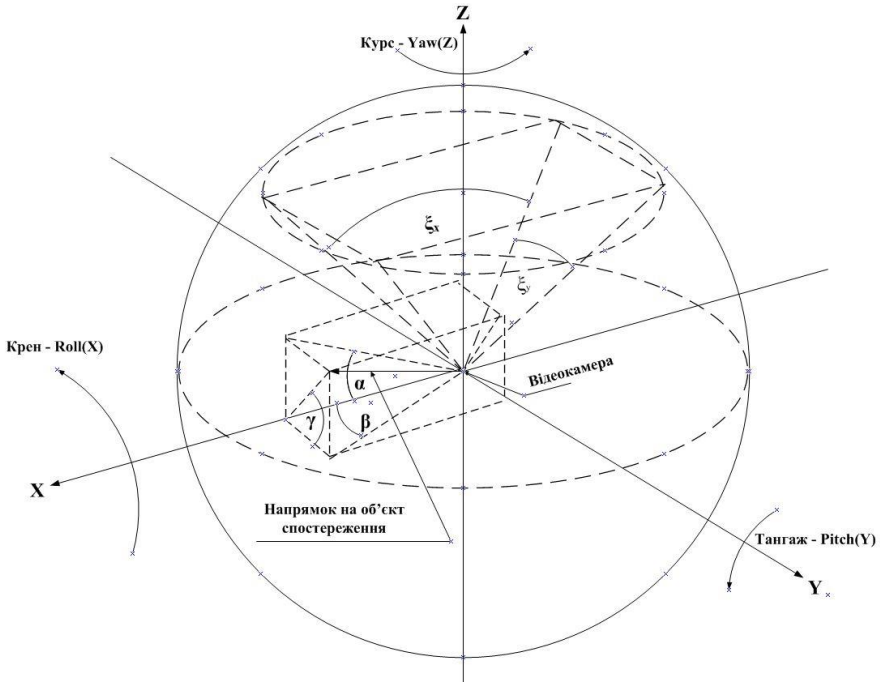


Рис. 1. Схема спостереження, де камера розглядається як центр сфери

Таким чином, виходячи зі співвідношення (1), можна визначити зміни поля зору камери:

$$\xi_{new} = 2\arctg\left(\frac{w}{2 * f_{new}}\right) = 2\arctg\left(\frac{w}{2 * f * zoom}\right) = 2\arctg\left(\frac{\text{tg}(\xi / 2)}{zoom}\right), \quad (2)$$

де ξ_{new} – кут огляду камери з урахування зумування. Співвідношення (2) показує залежність поля зору камери від величини коефіцієнта зумування.

Для того, щоб врахувати динаміку камери, було використано модель одиничної сфери. Нехай камера знаходиться в центрі та отримує зображення прямокутної проекції частини цієї сфери. Розмір зображення залежить від кутів огляду, а його положення залежить від значень кутів курсу, тангажу та крену камери. Якщо положення об'єкта на кадрі є центром обмежувальної рамки, яку повернула нейромережа, тоді, знаючи значення кутів огляду камери за горизонталлю й вертикаллю, можна знайти одиничний напрямний вектор на об'єкт спостереження:

$$\begin{cases} x_{image} = \frac{(u - w / 2)}{(w / 2)}; y_{image} = \frac{(v - h / 2)}{(h / 2)} \\ \theta = \arctg(x_{image} * \text{tg}(\xi_{x_new})) \\ \varphi = \arctg(y_{image} * \text{tg}(\xi_{y_new})) \end{cases}, \quad (3)$$

де x_{image}, y_{image} – відносні координати об'єкта на зображенні; u, v – піксельні координати об'єкта; w, h – ширина і висота кадру в пікселях; ξ_{x_new}, ξ_{y_new} – кути огляду камери за горизонталлю та вертикаллю з урахуванням коефіцієнта зумування; θ – полярний кут об'єкта спостереження; φ – азимутальний кут об'єкта спостереження. Таким чином, співвідношення (3) визначає сферичні координати кожного об'єкта спостереження в моделі одиничної сфери.

Для врахування обертання камери, потрібно послідовно врахувати матриці обертання по осях курсу, тангажу та крену камери.

$$R = R_y(\beta) R_z(\gamma) R_x(\alpha), \quad (4)$$

де R_x – матриця обертання навколо осі x ; α – кут тангажу; R_y – матриця обертання навколо осі y ; β – кут курсу; R_z – матриця обертання навколо осі z ; γ – кут крену. Таким чином, співвідношення (4) дозволяє визначити загальну матрицю обертання. У такому разі положення об'єкта з урахуванням повороту камери можна визначити як

$$V_{rot} = R * V, \quad (5)$$

де V_{rot} – вектор координат на одиничній сфері з урахуванням обертання; V – початковий вектор координат на одиничній сфері. Припускаючи, що спостереження ведеться за плоскою поверхнею, та виходячи зі значення висоти h розташування камери, визначимо координати об'єкта спостереження на плоскій поверхні:

$$\begin{cases} t = -y_{cam} / y_{obj} \\ x_{ground} = t * x_{obj} \\ y_{ground} = 0 \\ z_{ground} = t * z_{obj} \end{cases}, \quad (6)$$

де $x_{cam}; y_{cam}; z_{cam}$ – положення камери в просторі, $x_{ground}; y_{ground}; z_{ground}$ – реальне положення об'єкта на площині, $x_{obj}; y_{obj}; z_{obj}$ – положення об'єкта на одиничній сфері з урахуванням обертання та зумування камери. Таким чином, співвідношення (6) дає змогу визначити координати об'єкта відносно камери.

Відстеження рухомих об'єктів з використанням глобальних координат

У класичних підходах при вирішенні задачі відстеження використовуються піксельні координати об'єкта. У випадку рухомої камери доцільно також враховувати глобальні координати, оскільки вони не змінюються при обертанні або зумуванні камери. Розглядаючи найпростішу модель спостереження за плоскою поверхнею та припускаючи, що всі об'єкти інтересу знаходяться на ній, для відстеження можна використовувати тільки x_{ground} та z_{ground} . Основою

роботи методу відстеження є порівняння координат, отриманих на поточному кадрі з існуючими траєкторіями, які були розраховані на основі попередніх кадрів (Рис. 2).

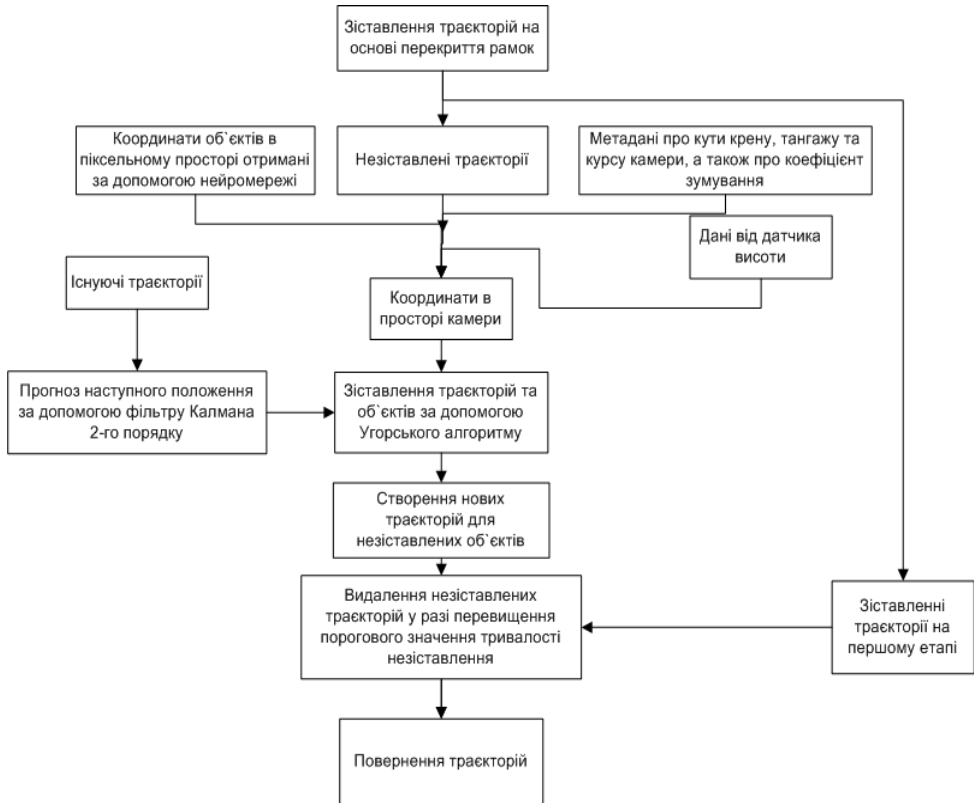


Рис. 2. Схема роботи алгоритму відстеження об'єктів в умовах динамічного спостереження

При ініціалізації алгоритму визначаються початкові траєкторії і ідентифікатори об'єктів спостереження. У подальшому на кожному етапі відстеження визначається передбачення координат об'єкта з використанням фільтра Калмана другого порядку [8]. Наступним кроком є зіставлення передбачених координат та об'єктів, що детектовані на поточному кадрі. Зіставлення відбувається у два етапи. На першому етапі як метрику використано IoU (коефіцієнт перекриття рамок), а як алгоритм оптимізації було обрано Угорський алгоритм. Використання такого підходу є класичним та дозволяє провести ефективно зіставлення у випадку, якщо камера є статичною. Рамки, які не були зіставлені з траєкторіями на першому етапі, проходять зіставлення на другому етапі, який враховує динаміку камери. Як метрику, яка використовується для формування матриці ціни, було обрано евклідову відстань між центрами об'єктів, розраховану на основі глобальних координат у просторі камери. Це дозволяє провести зіставлення при різких рухах камери або зумуванні. У випадку, якщо об'єкт не був знайдений на поточному кадрі, то траєкторія певну кількість кадрів зберігається в пам'яті. Значення кількості кадрів, яке визначає інтервал збереження траєкторії об'єкта в пам'яті, визначається динамічно та залежить від тривалості "життя" траєкторії:

$$N = \min(N_{min} + ka; N_{max}), \quad (7)$$

де N_{min} , N_{max} – мінімальна та максимальна кількість кадрів пам’яті; k – тривалість “життя” траєкторії (кількість попередніх кадрів упродовж яких об’єкт спостерігався); $a = 1$ – коефіцієнт зростання.

Застосування підходу з використанням співвідношення (7) дозволяє збільшити час пошуку об’єктів, які зазвичай швидко зникають із зображення і мають швидше видалятися з пам’яті. У випадку, якщо об’єкт не було зіставлено, то для нього створюється нова траєкторія та новий ідентифікатор.

Експериментальні дослідження

У результаті динамічного спостереження за складною сценою з декількома рухомими об’єктами, як за умови статичної камери, так і при змінах ракурсу та коефіцієнта зумування, було виміряно точність (*Identity Precision*) та повноту ідентифікації (*Identity Recall*), а також метрики точності багатооб’єктної відстежуваності (*Multiple Object Tracking Accuracy – MOTA*) та *F1*-метрику ідентичності (*IDF1*) (Таблиця 1). У результаті алгоритм показав приріст на всіх метриках, крім точності, порівняно з стандартним методом *ByteTrack* при вимірюванні на тій самій сцені.

Таблиця 1. Порівняння метрик

Назва алгоритму	IDF1	MOTA	Точність ідентифікації	Повнота ідентифікації
OCSORT	0.71	0.64	0.87	0.61
ByteTrack	0.81	0.76	0.87	0.74
Запропонований алгоритм	0.84	0.81	0.84	0.83

Головним недоліком запропонованого алгоритму є необхідність отримання метаданих від камери, що не завжди є можливим чисто технічно. Також метадані мають бути синхронізовані з зображенням, що не завжди вдається досягти на практиці.

Висновки

У результаті роботи запропоновано гібридний метод відстеження рухомих об’єктів з урахуванням динамічної моделі камери. Метод поєднує класичні підходи піксельного аналізу та алгоритм обчислення реальних глобальних координат відносно камери. Він використовує метадані від камери, а саме значення кутів крену, курсу, тангажу, а також коефіцієнт зумування, який являє собою відношення початкової та поточної фокусної відстані. Перевагою методу є його можливість відстеження об’єктів при активному зумуванні та зміні ракурсів камери. Алгоритм також дає змогу обчислити реальні, а не відносні швидкості та траєкторії об’єктів спостереження на основі моделі плоскої поверхні та даних від датчика висотоміра. Метод демонструє високі показники метрик порівняно з іншими класичними алгоритмами, а саме $IDF1 = 0.84$, $MOTA = 0.81$. Серед недоліків варто виділити необхідність постійного отримання метаданих від камери.

СПИСОК ЛІТЕРАТУРИ \ REFERENCES

1. Shen, Y. (2025). Computer vision: Technologies and applications. *Applied and Computational Engineering*, 163(1), 35–41. <https://doi.org/10.54254/2755-2721/2025.23817>
2. Krichen, M. (2023). Convolutional neural networks: A survey. *Computers*, 12(8), 151. <https://doi.org/10.3390/computers12080151>
3. Dasi, M., & Deepa. (2025). Real time object detection using You Only Look Once (Yolo) algorithm. *International Journal of Engineering Technology and Management Sciences*, 9(4), 103–109. <https://doi.org/10.46647/ijetms.2025.v09i04.012>
4. Kadam, P., Fang, G., & Zou, J. J. (2024). Object tracking using computer vision: A review. *Computers*, 13(6), 136. <https://doi.org/10.3390/computers13060136>
5. You, L., et al. (2024). Multi-object vehicle detection and tracking algorithm based on improved YOLOv8 and ByteTrack. *Electronics*, 13(15), 3033. <https://doi.org/10.3390/electronics13153033>
6. Kusumah, A. P., et al. (2023). Counting various vehicles using YOLOv4 and DeepSORT. *Journal of Integrated and Advanced Engineering (JIAE)*, 3(1), 1–6. <https://doi.org/10.51662/jiae.v3i1.68>
7. Terven, J., Córdova-Esparza, D.-M., & Romero-González, J.-A. (2023). A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Machine Learning and Knowledge Extraction*, 5(4), 1680–1716. <https://doi.org/10.3390/make5040083>
8. Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1), 35–45. <https://doi.org/10.1115/1.3662552>

Стаття надійшла до редакції 17.02.2026, надійшла після рецензування 17.03.2026, прийнята 07.04.2026

The article was received 17.02.2026, received after revision 17.03.2026, accepted 07.04.2026

Семко Олексій Вікторович

кандидат технічних наук, науковий співробітник відділу інформаційної безпеки Інституту телекомунікацій і глобального інформаційного простору НАН України
Адреса робоча: 02000, Київ, Україна, Чоколівський бульвар, 13
ORCID ID: <https://orcid.org/0000-0001-6473-1329> e-mail: semalek@meta.ua

Віннічук Денис Олександрович

аспірант кафедри комп'ютерних наук Національного університету біоресурсів і природокористування України
Адреса робоча: вул. Героїв Оборони, 15, Київ, 03041
ORCID ID: <https://orcid.org/0009-0001-9441-7696> e-mail: vinniboy2709@gmail.com