

МЕТОД ЛОКАЛІЗАЦІЇ ТА ПОБУДОВИ МАПИ ДИНАМІЧНОГО СЕРЕДОВИЩА ДЛЯ БЕЗПІЛОТНИХ СИСТЕМ

Moroz M.V. Simultaneous localization and mapping method of dynamic environment for unmanned systems. This paper is devoted to the study of simultaneous localization and mapping (SLAM) methods for unmanned systems operating in complex and dynamic environments. The paper considers the challenges associated with the need to adapt to environmental changes. The use of the Kalman filter and its extended versions is proposed, in particular with optimization using nonlinear programming methods. The obtained scientific results include the optimization of the Kalman filter taking into account nonlinearities using nonlinear programming methods, as well as the integration of classical computer vision algorithms (ORB) with convolutional neural networks. To achieve these results, the tasks of analyzing existing SLAM methods, developing new adaptive localization algorithms, and their experimental verification on existing datasets were solved. The results obtained can be recommended for use in projects on autonomous navigation of unmanned ground and air platforms, especially in conditions of limited environmental predictability, such as rescue operations, military missions, automated manufacturing and transportation systems.

Keywords: SLAM, Kalman filter, dynamic environment, localization, unmanned systems, convolutional neural networks

Мороз М.В. Метод локалізації та побудови мапи динамічного середовища для безпілотних систем. Стаття присвячена дослідженню методів одночасної локалізації та побудови мапи (SLAM) для безпілотних систем, що функціонують у складних та динамічних середовищах. У роботі розглядаються виклики, пов'язані з необхідністю адаптації до змін навколишнього середовища. Запропоновано використання фільтра Калмана та його розширених версій, зокрема з оптимізацією за допомогою методів нелінійного програмування. Отримані наукові результати включають оптимізацію фільтра Калмана з урахуванням нелінійностей за допомогою методів нелінійного програмування, а також інтеграцію класичних алгоритмів комп'ютерного зору (ORB) із згортковими нейронними мережами. Для досягнення зазначених результатів були вирішені задачі аналізу існуючих методів SLAM, розробки нових адаптивних алгоритмів локалізації, а також їхньої експериментальної перевірки на існуючих датасетах. Отримані результати можуть бути рекомендовані до використання у проектах з автономної навігації безпілотних наземних і повітряних платформ, особливо в умовах обмеженої передбачуваності середовища, таких як рятувальні операції, військові місії, автоматизоване виробництво та транспортні системи.

Ключові слова: SLAM, фільтр Калмана, динамічне середовище, локалізація, безпілотні системи, згорткові нейронні мережі

Вступ

Актуальність розробки та вдосконалення методів SLAM (simultaneous localization and mapping) в сучасній науці важко переоцінити. Завдання SLAM полягає в одночасній локалізації роботи й створенні мапи середовища [1], що є критичною функцією для багатьох автономних безпілотних систем, які оперують в нестабільних або невідомих умовах. Вдосконалення цих технологій відкриває можливість їх застосування в множині важливих галузей, таких як автоматизовані транспортні системи, військові та аварійно-рятувальні операції, сільське господарство, будівництво та промисловість.

В умовах постійної динаміки та непередбачуваності середовищ, здатність безпілотних систем ефективно локалізуватися та картографувати навколишнє середовище реалізує основоположну потребу в їх автономності. Наприклад, у автоматизації виробництва точне розміщення роботів та їх взаємодія зі складними промисловими установками може вирішуватись за допомогою продвинутих SLAM-систем [2]. Автономні транспортні засоби, такі як безпілотні системи, також залежать від високоточних SLAM-методів для навігації в неконтрольованих умовах.

Подальше вдосконалення методів SLAM може сприяти підвищенню надійності і безпеки

безпілотних систем, зокрема в критичних застосуваннях, як-от роботи в зонах стихійного лиха, зоні бойових дій, або в інших недоступних для людей місцях.

Аналіз останніх досліджень. В [3] представлено метод візуального SLAM, який інтегрує семантичну сегментацію об'єктів за допомогою Mask R-CNN, згорткової нейронної мережі для ефективного відокремлення динамічних об'єктів від статичного оточення. Метод використовує отримані маски об'єктів разом із оптичним потоком для одночасної оцінки траєкторії камери та повного 3D руху об'єктів. Запропонований підхід дозволяє точно відстежувати рухомі об'єкти без попередніх знань про їх форму, що особливо ефективно в динамічних сценах, таких як міське середовище. Реалізація методу демонструє значне покращення точності порівняно з існуючими статичними і лінійними методами.

В [4, 5, 6] представлено методи, які на відміну від [3], яка спираються на конвенційні методи сегментації зображень, таких як ORB (Oriented FAST and Rotated BRIEF) для виявлення ключових точок і побудови розрідженої мапи статичного оточення. Метод витягує ORB-особливості зі статичних частин зображень, виключаючи динамічні об'єкти, що дозволяє будувати точну розріджену мапу.

У статті [7] представлено метод для багатооб'єктного відстеження (MOT), яка інтегрує дані з різних сенсорів, зокрема LiDAR та камери, для покращення точності локалізації об'єктів у 3D-просторі. Метод використовує двоетапний процес асоціації: спочатку об'єднуються спостереження з обох сенсорів, а потім оновлюються траєкторії об'єктів навіть за відсутності даних з одного з них. Цей підхід дозволяє відстежувати об'єкти за межами діапазону LiDAR за допомогою візуальних даних, а також точно локалізувати їх у 3D-просторі, коли вони потрапляють у зону дії LiDAR. Метод використовує камери для ідентифікації віддалених об'єктів, а LiDAR – для точної оцінки траєкторії, що дозволяє динамічно зменшувати похибки локалізації та картографування. Реалізація також підтримує різні конфігурації сенсорів, що робить її універсальною для різних..

Постановка завдання. Конвенційні методи SLAM зазвичай забезпечують одометричну оцінку середовища та його наближення за допомогою імовірнісних та ітераційних методів оцінки розпізнаваних точок. Хоча отримані результати можна використовувати для таких завдань, як автономна навігація, їм не вистачає семантичної інформації про середовище. Натомість семантично-орієнтовані методи мають багато потенційних застосувань, як от розрізнення пішоходів, автомобілей, велосипедистів тощо в системах автономного керування транспортом, виявлення безпілотником навігаційних циклів на основі дескрипторів об'єктного рівня та розрізнення типів спостережуваної техніки для військових систем.

У сучасних реалізаціях SLAM методи локалізації та побудови мапи здебільшого базуються на використанні сенсорного злиття даних (sensor fusion) з лінійною асоціацією вимірювань, що формуються з декількох джерел, таких як камери, лідари, інерційні сенсори та GPS. Крім того, широко застосовується попереднє розділення (сегментація) середовища на статичні та динамічні об'єкти для підвищення точності побудови мапи. Проте, такі підходи мають низку обмежень, пов'язаних із недостатньою гнучкістю у змінних умовах та неврахуванням нелінійних залежностей між сенсорними даними, що особливо критично в динамічному середовищі з великою кількістю рухомих об'єктів.

У зв'язку з цим, завдання полягає у створенні нового методу, який би поєднував найкращі практики сегментації (зокрема, із використанням згорткових нейронних мереж) з нелінійною оцінкою фактору злиття даних. Такий підхід має ґрунтуватися на ітеративному аналізі похибок оцінювання, що дозволяє адаптивно налаштовувати ваги сенсорів залежно від якості та надійності вхідної інформації в кожен момент часу. Це дозволить значно покращити точність локалізації та побудови мапи навіть у складних та змінних умовах.

Метою роботи є розробка комбінованого методу локалізації та картографування в динамічних середовищах, який поєднує сучасні підходи до сегментації об'єктів та нелінійного сенсорного злиття даних. Такий метод має забезпечити підвищену точність локалізації та стабільність побудови мапи в умовах постійних змін оточення, високої насиченості рухомих об'єктів і варіативної якості сенсорної інформації. Для досягнення цього передбачається використання ітеративного аналізу похибки, що дозволяє адаптивно коригувати модель злиття

даних та мінімізувати вплив невизначеностей. Запропонований підхід спрямований на забезпечення надійної автономної навігації безпілотних систем у сценаріях, де класичні SLAM-методи втрачають ефективність.

Виклад основного матеріалу дослідження

Запропонований метод повинен використовувати комбінований підхід для сегментації зображень та побудови мапи. Для статичних об'єктів було застосовано алгоритм ORB (Oriented FAST and Rotated BRIEF), який виявляє ключові точки на нерухомих ділянках зображення. Ці точки формують розріджену мапу статичного оточення. Для динамічних об'єктів використовується нейромережа Mask R-CNN, яка точно сегментує рухомі об'єкти (наприклад, автомобілі чи техніку) з високою точністю, створюючи для них маски з унікальними ідентифікаторами.

Оптичний потік для запропонованого методу обчислюється за допомогою PWC-Net, що дозволяє відстежувати всі точки в межах масок динамічних об'єктів. Це особливо важливо, коли сегментація дає збій – оптичний потік допомагає відновити маски завдяки унікальним ідентифікаторам точок. Для визначення руху об'єктів використовується сценний потік (scene flow), який відрізняє рухомі об'єкти від статичних (рис. 1).

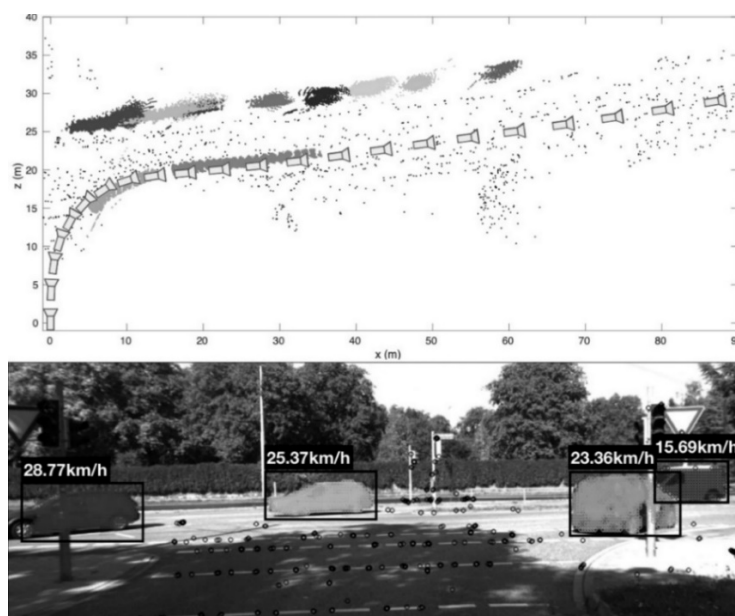


Рис. 1. Приклад сегментації зображення: а – двовимірна мапа з положенням потоку камери та розпізнаваних точок, б – зображення з камери з масками рухомих об'єктів

Запропонований метод будує дві окремі мапи: статичну та динамічну. Статична мапа формується за допомогою алгоритму Bundle Adjustment [8], який оптимізує позиції ORB-точок і позицію камери. Динамічна мапа відображає траєкторії рухомих об'єктів і постійно оновлюється з часом. Обидві мапи асоціюються між собою, що дозволяє отримати цілісне уявлення про навколишнє середовище (рис. 2).



Рис. 2. Схематичне зображення комбінованого методу сегментації зображення

Такий підхід дозволяє ефективно відокремлювати статичні та динамічні об'єкти, точно оцінювати їхній рух і будувати комплексну мапу навколишнього середовища. Це особливо важливо для застосувань у динамічних сценаріях.

Злиття даних від сенсорів. Однією з ключових складових SLAM алгоритмів є злиття даних від сенсорів для оцінки поточного положення платформи є використання алгоритму злиття (fusion). Запропонований метод SLAM має підвищену точність позиціонування та надійність шляхом злиття кількох джерел асоційованих даних від сенсорів за допомогою нелінійної оптимізації – така архітектура [9] демонструє найкращу точність та надійність. Фільтра Калмана є потужним інструментом оцінки та злиття даних у багатьох SLAM системах. Розширені версії фільтра апроксимують стан системи за допомогою нелінійних рівнянь, що робить його особливо ефективним для динамічних систем [10].

Стан фільтра визначається як:

$$x_{k+1} = x_k + u_k + \xi_k \quad (1)$$

$$z_k = x_k + \eta_k \quad (2)$$

$$x_{k+1}^{opt} = K \cdot z_{k+1} + (1 - K)(x_k^{opt} + u_k) \quad (3)$$

де x_k – оцінка вектору стану системи у певний момент часу k ;

z_k – стан системи отриманий з сенсора;

u_k – відома управляюча функція;

ξ_k – похибка моделі, а σ_ξ^2 її дисперсія;

η_k – похибка сенсора, а σ_η^2 її дисперсія;

x_k^{opt} – відфільтрований стан системи у певний момент часу k ;

K – змінний коефіцієнт Калмана;

Задача зводиться до знайдення оптимального відфільтрованого значення x_{k+1}^{opt} (3), де середньоквадратична похибка моделювання $E(e_{k+1}^2)$ (4) буде мінімізованою та обчислюється оптимальний коефіцієнт K для кожної ітерації:

$$E(e_{k+1}^2) \rightarrow \min \quad (4)$$

$$E(e_{k+1}^2) = (x_{k+1} - x_{k+1}^{opt})^2 = (1 - K)^2 (Ee_k^2 + \sigma_\xi^2) + K^2 \sigma_\eta^2 \quad (5)$$

Запропонований метод використовує оптимізацію фільтра за допомогою методів нелінійного програмування. Накладаючи обмеження $c_i(K)$ на коефіцієнт фільтра K (6), можна вирішувати задачу оптимізації (5) для ковзаючого вікна з кроком n , що ітеративно корегується на основі похибки конкретного сенсора для вибраних проміжків K , що відповідає змінам у середовищі [11]:

$$c_i(K) \leq 0, \forall i \in I \quad (6)$$

Вирішувати задачу оптимізації фільтра пропонується через застосування методів нелінійного програмування, таких як метод штрафів, де штраф вводять за відхилення від моделі, що дозволяє більш точно корегувати оптимальний стан фільтра x_{k+1}^{opt} (3) порівняно із лінійною версією класичного фільтра Калмана.

Таблиця 1

Порівняльна таблиця RMSE методів

Датасет	Запропонований метод	EagerMOT	ORB-SLAM
MH_01_easy	0.07	0.09	0.16
MH_02_easy	0.04	0.07	0.15
MH_03_easy	0.05	0.06	0.14
MH_04_difficult	0.09	0.13	0.25
MH_05_difficult	0.12	0.19	0.32
MH_06_difficult	0.05	0.06	0.13

Запропонований комбінований метод порівнюється з класичними існуючими методами SLAM одометрично-інерціального позиціонування ORB-SLAM [8] та EagerMOT [7] відповідно. Критерієм порівняння виступає середньоквадратична похибка системи (RMSE), отримана в результаті моделювання. Порівняння відбувається на основі датасету EuRoC, який можна використовувати для моделювання тестування одометричних методів і одометрично-інерційних методів, і це один із найбільш часто використовуваних наборів даних для оцінки точності позиціонування методів SLAM базованих на одометрії.

Значення похибки усіх послідовностей даних є нижчими на 10-20% (табл. 1) на різних наборах даних, у порівнянні з іншими методами. Таким чином, можна сказати, що точність розробленого методу краща, а похибка моделювання є мінімальною.

Інші проблеми. Реалізація SLAM-методів у реальному часі стикається з низкою обчислювальних викликів, особливо при роботі з динамічними середовищами. Перша проблема – це високі вимоги до обчислювальних потужностей. Методи SLAM, такі як ORB-SLAM або VDO-SLAM, потребують інтенсивних обчислень для відстеження особливостей, оптимізації мап та оцінки руху. Це особливо критично для систем із семантичною сегментацією (наприклад, Mask R-CNN), які вимагають значних ресурсів GPU.

Друга проблема – необхідність розділення на фронтенд і бекенд. Фронтенд (наприклад, трекінг особливостей і попередня обробка даних) повинен працювати в реальному часі, щоб забезпечити оперативну реакцію системи. Бекенд (оптимізація мапи, глобальна узгодженість) може працювати асинхронно, але повинен ефективно масштабуватися, щоб уникнути накопичення помилок. Якщо бекенд не встигає за фронтендом, це призводить до затримок і погіршення точності.

Третя проблема – онлайн-обробка в умовах обмежених ресурсів. На відміну від офлайн-методів, які можуть використовувати всі доступні дані, онлайн-SLAM мусить працювати з обмеженим буфером кадрів, уникаючи надмірного навантаження на пам'ять. Це потребує ефективних стратегій вибіркового оновлення мапи, таких як віконна оптимізація або інкрементальні методи (наприклад, iSAM2).

Крім того, динамічні об'єкти ускладнюють обчислення, оскільки їхній рух порушує припущення про статичність сцени. Традиційні SLAM-системи або ігнорують такі об'єкти (що призводить до помилок), або вимагають додаткових обчислень для їх трекінгу (як у DyOb-SLAM). Це ще більше навантажує систему, особливо при високій щільності рухомих об'єктів.

Ці виклики залишаються активними напрямками досліджень, оскільки сучасні SLAM-системи прагнуть до балансу між точністю, швидкістю та енергоефективністю.

Висновки

В даній роботі запропоновано метод SLAM, що використовує комбінацію різних рішень для кожного етапу моделювання, що показують найкращу ефективність, таких як ORB сегментація для статичних об'єктів і Mask R-CNN для динамічних, що дозволяє досягти балансу між точністю та продуктивністю (рис. 2). Він також включає моделювання оптичного потоку PWC-Net для покращення трекінгу рухомих об'єктів (рис. 1a), та оптимізований за допомогою методів нелінійного програмування алгоритм злиття та оцінки поточного стану системи на основі фільтру Калмана (4). Порівняльний аналіз показав, що метод демонструє кращі результати на основі оцінки похибки змодельованої системи порівняно з аналогами на 10-20% для різних наборів даних (табл. 1).

Також було досліджено типи існуючих SLAM-систем для роботи в динамічних середовищах, зокрема класичні статичні системи на основі ORB, сучасні гібридні підходи з використанням семантичної сегментації.

Серед ключових проблем, які досліджувалися, є обробка динамічних об'єктів, точність трекінгу в умовах швидких змін середовища, а проблема злиття даних з різних сенсорів. Було виявлено, що традиційні SLAM-системи, які ігнорують рухомі об'єкти, демонструють значні похибки в динамічних сценах, тоді як сучасні підходи з семантичною сегментацією вимагають

надмірних обчислювальних ресурсів.

Основні проблеми, які залишаються актуальними для SLAM-систем, включають високу обчислювальну складність при інтеграції нейромереж, необхідність ефективного розподілу обчислень між фронтендом і бекендом, а також обробку сцен з високою щільністю динамічних об'єктів. Для вирішення цих проблем перспективними напрямками є використання легких архітектур нейромереж, хмарних обчислень та інкрементних методів оптимізації.

Список використаної літератури:

1. Bailey T., Durrant-Whyte H. Simultaneous localization and mapping (SLAM): part II. IEEE Robotics & Automation Magazine. 2006. Vol. 13, no. 3. P. 108–117. URL: <https://doi.org/10.1109/mra.2006.1678144>.
2. Yarovoi A., Cho Y. K. Review of simultaneous localization and mapping (SLAM) for construction robotics applications. Automation in Construction. 2024. Vol. 162. P. 105344. URL: <https://doi.org/10.1016/j.autcon.2024.105344>.
3. VDO-SLAM: A Visual Dynamic Object-aware SLAM System / J. Zhang et al. URL: <https://arxiv.org/abs/2005.11052>.
4. Visual SLAM in dynamic environments based on object detection / Y.-b. Ai et al. Defence Technology. 2020. URL: <https://doi.org/10.1016/j.dt.2020.09.012>.
5. Liu Y., Miura J. RDS-SLAM: Real-Time Dynamic SLAM Using Semantic Segmentation Methods. IEEE Access. 2021. Vol. 9. P. 23772–23785. URL: <https://doi.org/10.1109/access.2021.3050617>.
6. Wadud, R. A., & Sun, W. (2022). DyOb-SLAM: Dynamic Object Tracking SLAM System. arXiv. URL: <https://doi.org/10.48550/arXiv.2211.01941>.
7. Kim, A., Osep, A., & Leal-Taixe, L. (2021). EagerMOT: 3D Multi-Object Tracking via Sensor Fusion.
8. Mur-Artal R., Montiel J. M. M., Tardos J. D. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. IEEE Transactions on Robotics. 2015. Vol. 31, no. 5. P. 1147–1163. URL: <https://doi.org/10.1109/tro.2015.2463671>.
9. Applying SLAM Algorithm Based on Nonlinear Optimized Monocular Vision and IMU in the Positioning Method of Power Inspection Robot in Complex Environment / C. Wang et al. Mathematical Problems in Engineering. 2022. Vol. 2022. P. 1–14. URL: <https://doi.org/10.1155/2022/3378163>.
10. Choi K.-S., Lee S.-G. Enhanced SLAM for a mobile robot using extended Kalman Filter and neural networks. International Journal of Precision Engineering and Manufacturing. 2010. Vol. 11, no. 2. P. 255–264. URL: <https://doi.org/10.1007/s12541-010-0029-9>.
11. Liu H. Identifying and updating local optimization methods in extended Kalman filter SLAM. Applied and Computational Engineering. 2023. Vol. 4, no. 1. P. 569–573. URL: <https://doi.org/10.54254/2755-2721/4/2023325>.

Автор статті

Мороз Михайло – аспірант, Державний університет інформаційно-комунікаційних технологій, Київ, Україна.

ORCID: 0009-0002-9546-394X

Author of the article

Moroz Mykhailo – postgraduate, State University of Information and Communication Technologies, Kyiv, Ukraine.

ORCID: 0009-0002-9546-394X