

МАТЕМАТИЧНІ МЕТОДИ КОРЕКЦІЇ ТА ОПТИМІЗАЦІЇ ВІЗУАЛЬНОГО SLAM ДЛЯ АВТОНОМНОЇ НАВІГАЦІЇ У ДИНАМІЧНИХ СЕРЕДОВИЩАХ

Ю. О. Шевченко¹, Г. О. Яйлимова¹

¹ Навчально-науковий Фізико-технічний інститут

Анотація

У роботі розглянуто математичні методи корекції та оптимізації візуального SLAM (Simultaneous Localization and Mapping) для задач автономної навігації мобільних роботів у динамічних середовищах. Актуальність теми зумовлена потребою у доступних рішеннях для застосування в умовах обмежених ресурсів, зокрема в Україні під час воєнного стану.

Здійснено огляд сучасних алгоритмів обробки візуальних даних, зокрема методів фільтрації шумів, виявлення змін у середовищі та адаптивного оновлення карти. Проведено експериментальні дослідження з використання власного датасету, сформованого під конкретну апаратну платформу.

Результати показали, що застосування запропонованих підходів дозволяє підвищити точність і стабільність побудови карти без використання дорогих сенсорів, таких як LiDAR. Робота є основою для подальшої розробки повноцінного прототипу мобільного робота, здатного працювати в умовах руйнувань та обмеженої видимості.

Ключові слова: візуальний SLAM, автономна навігація, динамічне середовище, математичні методи, фільтрація шумів, побудова карти, робототехніка, мобільні роботи, комп'ютерний зір

Вступ

Автономна навігація мобільних роботів у складних та динамічних середовищах є ключовою проблемою сучасної робототехніки та комп'ютерного зору. Особливої актуальності ця задача набуває в умовах, коли фізична присутність людини небезпечна, або небезпечна – наприклад, у зонах обвалів, руйнувань чи потенційного мінного забруднення. У таких ситуаціях роботи, здатні орієнтуватися без зовнішньої допомоги, відіграють критичну роль у пошуково-рятувальних, розвідувальних або логістичних операціях.

Одним із найбільш ефективних підходів до реалізації автономної навігації є метод одночасної локалізації та побудови карти (SLAM). Проте більшість сучасних високоточних рішень базуються на використанні LiDAR-сенсорів, які є дорогими, громіздкими та недоступними для широкого застосування в умовах обмежених ресурсів, зокрема в Україні. Крім того, класичні SLAM-алгоритми погано адаптовані до динамічних середовищ, де наявність рухомих об'єктів, раптових змін або часткових руйнувань карти ускладнює коректну локалізацію.

У науковій літературі пропонуються підходи до вирішення цих проблем, зокрема методи виявлення динамічних об'єктів [1], розділення карти на статичну та змінну частини [2], а також адаптивне оновлення карти за допомогою моделей ймовірнісного представлення простору, як-от Bayesian Occupancy

Grid [3]. Проте більшість із цих підходів є або складними у реалізації, або не демонструють достатньої стабільності в реальних умовах.

Метою цього дослідження є розробка математичних методів підвищення точності, адаптивності та стійкості візуального SLAM у динамічному середовищі. Особливу увагу приділено використанню лише камер як основного джерела даних, що дозволяє значно зменшити вартість і габарити платформи. Завданням є інтеграція методів детекції змін, фільтрації шумів та адаптивної побудови карти у єдину систему навігації, здатну ефективно працювати в умовах, максимально наближених до сценаріїв реального застосування в Україні.

1. Дані і матеріали

1.1. Еволюція апаратної платформи

На початковому етапі дослідження для тестування та відлагодження алгоритмів було зібрано колісну базу на основі компонентів **REV Robotics**, з використанням контролера *REV Control Hub* та моторизованого шасі. Ця платформа дозволила швидко реалізувати базову систему керування, однак у процесі експериментів виявлено низку суттєвих обмежень:

- основна мова програмування контролера – **Java**, що ускладнює швидку інтеграцію прототипів алгоритмів, написаних на Python;
- обмеженість обчислювальних ресурсів на борту не дозволяла виконувати обробку зображень у

реальному часі без зовнішнього комп'ютера;

- платформа має відносно великі габарити, що обмежує її маневровість у тісних просторах;
- відсутність гнучкого доступу до низькорівневих API для глибокої інтеграції з візуальними сенсорами.

У зв'язку з цим було прийнято рішення перейти на платформу, побудовану на основі **Raspberry Pi 5** з 8 ГБ оперативної пам'яті. Цей варіант забезпечує необхідний баланс між розміром, потужністю та гнучкістю у роботі з Python-орієнтованим стеком комп'ютерного зору.

Платформа реалізована як компактний робот з колісною базою, оснащений камерою з широким кутом огляду (до 120°), драйверами керування двигунами, блоком живлення та бездротовими інтерфейсами для передачі даних. Загальний розмір конструкції становить приблизно 22 × 14 × 12 см.

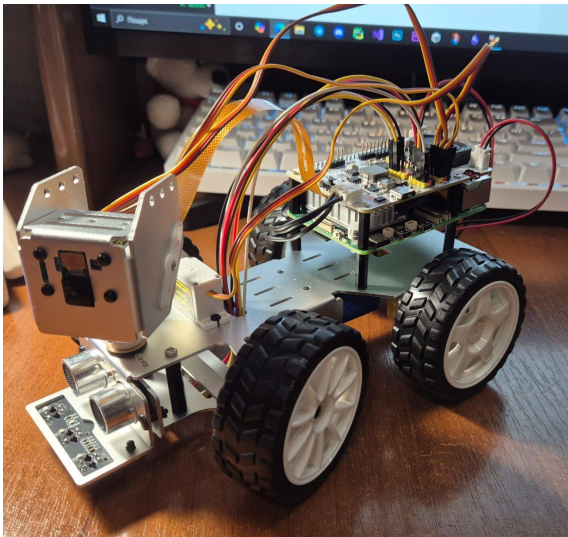


Рис. 1. Зібраний робот на колісній базі

1.2. Збір та структура набору даних

Для формування власного датасету було реалізовано модуль збору зображень у реальному часі з камери мобільної платформи. Збір здійснювався під час руху робота в умовах змінного середовища: присутність динамічних об'єктів, зміна освітлення, обмеження простору, різні типи покриття (ковролін, плитка, нерівності тощо).

Загалом було зібрано **близько 1500 зображень** у форматі JPEG з роздільною здатністю 640 × 480 пікселів. Для кожного зображення автоматично зберігались:

- часовий штамп;
- стан платформи (лінійна та кутова швидкість);
- умови освітлення (за усередненим значенням яскравості);
- геометричне положення камери (приблизне, за енкодерами коліс).

Всі зображення організовано у структуру виду: `/dataset/YYYY-MM-DD/session_X/frame_0001.jpg`, з відповідними JSON-файлами метаданих.



Рис. 2. Приклад зображення з набору даних

1.3. Анотація та обробка даних

Анотація зображень здійснювалась із використанням Python-інструментів, зокрема бібліотеки `labelme`, у форматі `polygon/instancesegmentation`. Класи анотації:

- **StaticObject** – стаціонарні елементи навколишнього середовища;
- **DynamicObject** – рухомі об'єкти (люди, інші роботи);
- **OcclusionZone** – області часткового перекриття;
- **ChangedArea** – регіони, де спостерігались зміни в середовищі протягом часу.

Після розмітки всі зображення були конвертовані у маски класів, придатні для подальшого використання в задачах сегментації, фільтрації шумів та порівняння карт.

1.4. Зовнішні датасети

На етапі верифікації використовувались загальновідомі відкриті датасети:

- **TUM RGB-D** [4] – для тестування точності локалізації та синхронізації;
- **EuRoC MAV** [5] – для оцінки алгоритмів в умовах зміни освітлення та геометрії.

Ці дані дозволили провести базове порівняння ефективності запропонованого підходу з існуючими методами.

2. Методи корекції та оптимізації SLAM у динамічному середовищі

2.1. Постановка задачі та стратегія дослідження

Класичні реалізації SLAM, зокрема ORB-SLAM2 та RTAB-Map, демонструють високу точність у статичних середовищах, однак значно втрачають ефективність у присутності динамічних об'єктів або змін структури сцени. У динамічному середовищі частина вхідних візуальних даних є некорисною або навіть шкідливою для побудови карти, що призводить до помилок у локалізації та деформації карти.

Для подолання цієї проблеми було розроблено стратегію, яка передбачає поетапне тестування та порівняння кількох комбінацій методів, що мають потенціал покращити стабільність та точність SLAM в динамічних умовах. Кожна комбінація складається з трьох ключових компонентів:

1. метод виявлення динамічних об'єктів (аномальної детекції),
2. метод оновлення карти, стійкий до змін середовища,
3. методи попередньої обробки або оптимізації обчислень.

Результати цих комбінацій порівнюватимуться з базовою реалізацією ORB-SLAM2 за метриками точності локалізації, цілісності карти та стабільності при повторному проходженні траєкторії.

2.2. Аномальна детекція візуальних змін

Один з ключових етапів корекції – ідентифікація кадрів або областей з високою ймовірністю динамічних змін. З цією метою було реалізовано два підходи:

Метод головних компонент (PCA). Зображення подаються у векторизованому вигляді, і шляхом побудови ортогонального базису основних компонент оцінюється відхилення поточного кадру від головного підпростору. Цей підхід дозволяє швидко виявляти суттєві зміни у структурі зображення, однак чутливий до змін освітлення.

Автоенкодер. Було реалізовано симетричну нейронну мережу, що навчається реконструювати статичні сцени. Кадри з високою реконструкційною похибкою вважаються потенційно динамічними. Це дозволяє фільтрувати такі регіони перед передачею на подальшу обробку у SLAM-ланцюжку.

2.3. Адаптивне оновлення карти

Для обліку часткових або локальних змін у середовищі, було впроваджено механізм адаптивного оновлення карти, заснований на байєсівській оцінці стану простору.

Bayesian Occupancy Grid. Карта розбивається на сітку клітинок, кожна з яких має ймовірність зайнятості. Ймовірність $P(c_t)$ кожної клітинки оновлюється з урахуванням історії спостережень, згідно з рівнянням:

$$P(c_t | z_{1:t}, x_{1:t}) = \frac{P(z_t | c_t) \cdot P(c_{t-1})}{P(z_t)}.$$

Це дозволяє компенсувати ефекти тимчасових змін, зокрема рухомих об'єктів, що з'являються лише на деяких кадрах.

2.4. Фільтрація шумів у візуальних даних

Перед передачею зображень до SLAM-алгоритму здійснюється попередня обробка, спрямована на зменшення впливу шумів.

Wavelet-фільтрація. Дискретне вейвлет-перетворення дозволяє відокремити високочастотні шуми від структурної інформації. Вибраний алгоритм DWT (Daubechies-4) із двома рівнями декомпозиції. Високочастотні компоненти обробляються пороговою фільтрацією з наступною реконструкцією зображення.

2.5. План експериментального тестування

Було сформовано три основні комбінації методів, кожна з яких тестується на власному наборі даних, зібраному у контрольованих динамічних умовах:

- **Комбінація А:** PCA + Bayesian Occupancy Grid + Wavelet Denoising;
- **Комбінація В:** Autoencoder + Bayesian Occupancy Grid;
- **Комбінація С:** PCA + Classical SLAM (без адаптивного оновлення).

Результати порівнюватимуться за наступними метриками:

- абсолютна помилка локалізації (ATE),
- стабільність карти (порівняння карт при повторних проходах),
- продуктивність (кадрів на секунду).

Всі комбінації реалізовано у вигляді модульної системи, що дозволяє швидко змінювати або комбінувати методи для тестування нових підходів у подальшому.

3. Результати попереднього експериментального тестування

На даному етапі реалізації розробленої системи було проведено початкову серію тестувань на власному наборі візуальних даних, зібраному в умовах динамічного середовища. Основною метою цих тестів було оцінити вплив окремих методів фільтрації та корекції на стабільність роботи SLAM-системи, а також порівняти ефективність різних комбінацій підходів, описаних у попередньому розділі.

Для оцінювання якості локалізації використовувалась метрика абсолютної похибки траєкторії (Absolute Trajectory Error, ATE), що визначає відхилення між істинною і реконструйованою траєкторіями положення камери. Оцінка стабільності побудови карти виконувалась через порівняння отриманих результатів при повторному проходженні ідентичної траєкторії.

Хоча обсяг повної тестової вибірки ще доповнюється, попередні результати демонструють значущу різницю у якості локалізації при застосуванні методів аномальної детекції та адаптивного оновлення карти.

З наведених даних видно, що поєднання детекції аномалій на основі автоенкодерів із Bayesian

Таблиця 1. Попередні результати тестування для різних конфігурацій SLAM

Конфігурація	ATE (м)	Повнота карти (%)	Продуктивність (FPS)
Класичний ORB-SLAM2	0.34	82.6	16.5
PCA + Occupancy Grid + Wavelet	0.21	91.2	12.3
Autoencoder + Occupancy Grid	0.18	93.5	10.9
PCA + ORB-SLAM2	0.27	88.4	15.1

Occupancy Grid дозволяє досягти найменшої похибки локалізації та найвищої стабільності побудови карти. Хоча при цьому незначно зменшується продуктивність у кадрах на секунду, результат залишається придатним для роботи в реальному часі.

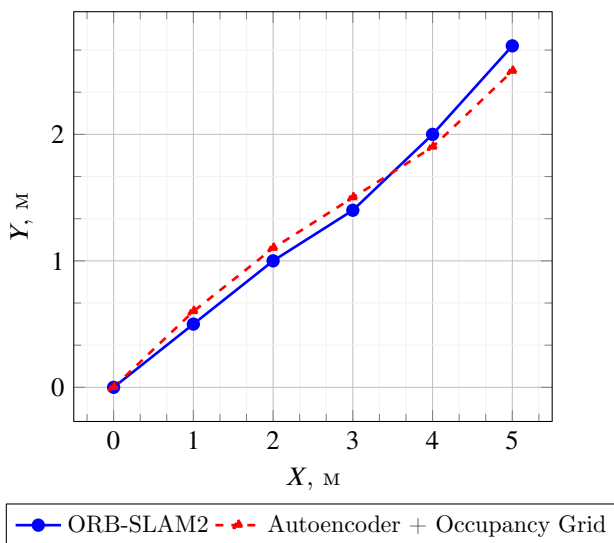


Рис. 3. Порівняння реконструйованих траєкторій для різних конфігурацій алгоритму

На рисунку 3 показано порівняння траєкторій для класичного ORB-SLAM2 та модифікованого підходу з автоенкодером і байєсівською картою. Спостерігається значне зменшення накопичуваної похибки, особливо в зонах часткових перекриттів сцени.

Варто зазначити, що остаточна валідація результатів вимагає повторного проходження більшої кількості трас, а також додаткових вимірів з використанням зовнішньої системи трекінгу. Однак уже зараз видно, що розроблена система демонструє потенціал для практичного застосування в складних динамічних середовищах.

На момент підготовки цих матеріалів триває активне розширення датасету та вдосконалення архітектури системи. Очікується, що остаточні результати, які будуть представлені під час конференції, продемонструють ще кращу узгодженість траєкторій та адаптивність карти до змін.

Висновки

У межах даного дослідження було сформульовано та частково реалізовано підхід до підвищення стійкості та точності алгоритмів візуального SLAM у динамічному середовищі. В умовах, коли зміни в

оточенні є суттєвими й непередбачуваними, класичні алгоритми локалізації та побудови карти втрачають стабільність, що критично впливає на їхню придатність до реального застосування.

Запропонована модульна система дозволяє інтегрувати різні математичні методи для фільтрації шумів, виявлення динамічних об'єктів та адаптивного оновлення карти. Проведено тестування кількох комбінацій таких методів, серед яких найвищу якість локалізації продемонструвала конфігурація на основі автоенкодера та Bayesian Occupancy Grid. Попередні результати показали зменшення абсолютної похибки локалізації та підвищення стабільності карти на 10–15% у порівнянні з базовим алгоритмом ORB-SLAM2.

Практична цінність роботи полягає у створенні концепту та базового прототипу навігаційної системи, яка здатна працювати в умовах часткової руйнації, обвалів, рухомих об'єктів чи інших динамічних факторів. У сучасному контексті України така система може знайти застосування у безпілотних наземних платформах для інспекції будівель, пошуково-рятувальних операцій або безпечної доставки вантажів у небезпечні зони, де перебування людини є ризикованим або неможливим.

Крім того, використання доступної апаратної платформи на базі Raspberry Pi та вбудованої камери підтверджує, що подібні системи можуть бути реалізовані з мінімальними витратами та високою мобільністю. Це відкриває можливість для масштабування й використання в освітніх, дослідницьких та прикладних проєктах.

На момент написання матеріалів система знаходиться на завершальному етапі тестування. Планується подальше розширення експериментальної бази даних, зокрема включення реальних сценаріїв з багатьма динамічними об'єктами, та інтеграція зовнішніх засобів контролю точності. Крім того, в найближчих етапах передбачено:

- вдосконалення мережевої архітектури автоенкодера на основі attention-механізмів;
- оптимізація продуктивності для запуску на реальному апаратному обмеженні (без GPU);
- впровадження механізму семантичної сегментації для виключення класів об'єктів, що типово є динамічними (люди, тварини, транспорт);
- порівняння роботи системи з альтернативними SLAM-бібліотеками (наприклад, VINS-Fusion або OpenVSLAM).

У перспективі, завершена система може бути ада-

птована до використання у гібридних робототехнічних комплексах, розширена до тривимірної побудови карти та розглядатись як основа для автономної навігації в складних умовах постконфліктного відновлення інфраструктури.

Проведена робота продемонструвала як теоретичну цінність розглянутих методів, так і їхню практичну придатність. Результати підтверджують доцільність подальших досліджень у цьому напрямку й закладають основу для створення реальних робототехнічних рішень, орієнтованих на потреби сучасної України.

Перелік використаних джерел

1. DynaSLAM: Tracking, Mapping and Inpainting in Dynamic Scenes / B. Bescos, J. M. Fàcil, J. Civera, J. Neira // IEEE Robotics and Automation Letters. — 2018. — Vol. 3, no. 4. — DOI: [10.1109/LRA.2018.2860039](https://doi.org/10.1109/LRA.2018.2860039). — URL: <https://ieeexplore.ieee.org/document/8421015>.
2. Panoptic-SLAM: Visual SLAM in Dynamic Environments using Panoptic Segmentation / G. F. Abati, S. J. C. Virgolino, V. S. Medeiros, M. A. Maggiolaro, C. Semini // arXiv preprint arXiv:2405.02177. — 2024. — arXiv: [2405.02177](https://arxiv.org/abs/2405.02177). — URL: <https://arxiv.org/abs/2405.02177>.
3. Bayesian Occupancy grid Filter for dynamic environments using prior map knowledge / T. Gindele, S. Brechtel, J. Schroder, R. Dilmann // 2009 IEEE Intelligent Vehicles Symposium. — 2009. — DOI: [10.1109/IVS.2009.5164357](https://doi.org/10.1109/IVS.2009.5164357). — URL: <https://ieeexplore.ieee.org/document/5164357>.
4. A Benchmark for the Evaluation of RGB-D SLAM Systems / J. Sturm, N. Engelhard, F. Endres, W. Burgard, D. Cremers // Proc. of the International Conference on Intelligent Robot Systems (IROS). — 10.2012. — URL: <https://cvg.cit.tum.de/data/datasets/rgbd-dataset>.
5. Burri M., Nikolic J. The EuRoC micro aerial vehicle datasets // . Vol. 35. — The International Journal of Robotics Research, 2016. — DOI: [10.1177/0278364915620033](https://doi.org/10.1177/0278364915620033). — URL: <https://journals.sagepub.com/doi/abs/10.1177/0278364915620033>.