

ЗАСТОСУВАННЯ АЛГОРИТМУ VISUAL INERTIAL SLAM ДЛЯ НАВІГАЦІЇ АВТОНОМНОГО РОБОТА

Вінницький національний технічний
університет

У даній роботі досліджується застосування алгоритму Візуально-інерційного SLAM (*Visual Inertial Simultaneous Localization and Mapping*) для навігації автономного робота. З розвитком технологій автономної навігації, постає потреба в ефективних методах, які дозволяють роботам орієнтуватися у складних середовищах. Візуально-інерційний SLAM поєднує дані з камер та інерціальних вимірювальних пристроїв, що забезпечує високу точність локалізації та побудови карти.

Протягом десятиліть було досліджено кілька підходів до SLAM на основі фільтра частинок, розширеного фільтра Калмана та нейронних мереж. Ці підходи розроблені для побудови карт у різних представленнях, включаючи представлення на основі орієнтирів, поверхонь, полігональної сітки та сітки зайнятості. Нині вони широко використовуються для таких додатків, як планування руху роботів у невідомому середовищі та використовуються в безпілотних автомобілях, безпілотних літальних апаратах та автономних підводних апаратах.

Ця робота спрямований на одночасну локалізацію та нанесення на карту робота в невідомому внутрішньому середовищі, використовуючи дані одометрії інерційного вимірювального пристрою (*Inertial Measurement Unit*) та особливі точки, виявлені стереокамерою. Для досягнення цієї мети реалізовано розширений підхід на основі фільтра Калмана. Фільтр Калмана є оптимальною лінійною оцінкою для моделей лінійної системи з адитивним незалежним білим шумом як у системах прогнозування, так і в системах спостереження. Розширений фільтр Калмана адаптував методи числення, а саме багатовимірні розкладання в ряд Тейлора, для лінеаризації нелінійної моделі.

Запропоновані рішення можуть бути корисними для вдосконалення автономних систем, таких як роботи для доставки, дронів і мобільних платформ, а також для розробки нових методів інтеграції сенсорних даних у сучасних технологіях навігації.

Ключові слова: візуально-інерційний SLAM, інерційний вимірювальний пристрій, розширений фільтр Калмана.

Вступ

Одночасна локалізація та картографування (SLAM) використовується для картографування невідомих середовищ за допомогою робота з рядом датчиків, які постійно збирають дані про оточення. Ці дані потрібно обробити, щоб створити чисту та точну карту. Робот може зіткнутися з деякими перешкодами або датчик може збирати шумні дані, які потрібно очистити. Ось де корисний візуально-інерційний SLAM. В роботі використовуються дані інерційного вимірювального пристрою (ІВП) та стереозображення автомобіля. Характеристики були отримані із зображень і забезпечені відповідними даними кутової та лінійної швидкості.

Роль робота виконує автомобіль, який їздить по дорогах загального користування та має GPS/ІВП, датчик Velodyne LiDAR, перспективні камери та налаштування стереокамери «риб'яче око». Надані дані попередньо синхронізуються.

Метою роботи є дослідження застосування візуально-інерційного SLAM при використанні розширеного фільтра Калмана (ЕФК). Використовується модель спостереження, щоб отримати відхилення між прогнозованим розташуванням об'єктів від ІВП та фактичними даними, отриманими камерою для виправлення розташування об'єктів.

Набір даних для алгоритму

У цьому розділі розглядається процес збору та підготовки набору даних для методів Visual Inertial SLAM, що реалізуються на основі даних, отриманих з автомобіля, оснащеного стереокамерами та інерціальним вимірювальним пристроєм (ІВП) [1].

Збір даних був здійснений за допомогою автомобіля, на якому були встановлені декілька стереокамер, що забезпечують отримання високоякісних зображень з різних ракурсів, та ІВП, яка вимірює прискорення та кутову швидкість. Системи спостереження, які складаються з стереокамер, дозволяють генерувати глибину сцени, тоді як ІВП забезпечує важливу інформацію про динаміку руху, що є критично важливим для точного злиття даних у Visual Inertial SLAM.

Дані були зібрані в різних умовах, включаючи міські пейзажі, дороги з різними характеристиками та зовнішні простори. Це дозволило отримати різноманітний набір даних, що підвищує загальну ефективність та надійність алгоритмів SLAM у різних сценаріях.

Після збору даних наступним етапом стало їх перетворення в зручний для обробки формат. Використовуючи бібліотеку NumPy, дані були збережені у форматі .prz, що дозволяє зберігати масиви даних разом з метаданими. Формат .prz є ефективним для зберігання багатовимірних масивів, що ідеально підходить для даних з камер та ІВП, оскільки дозволяє швидко завантажувати та обробляти великі обсяги інформації.

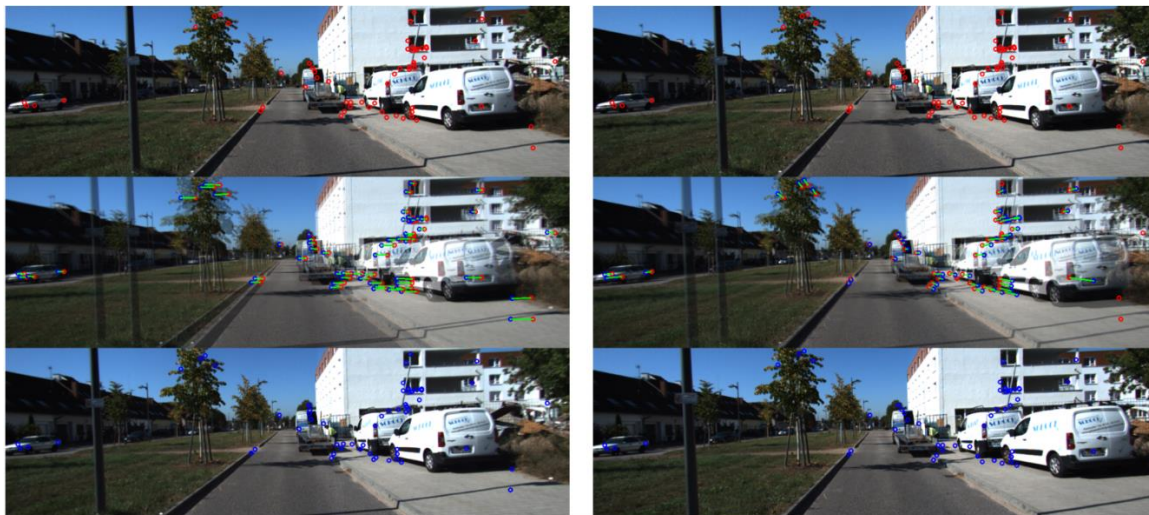


Рисунок 1 – Візуальні функції, що збігаються між лівим і правим кадрами камери (ліворуч) і часом (праворуч)

Кожен файл .prz містить наступні компоненти:

- **ІВП дані:** Вимірювання прискорення та кутової швидкості, які дозволяють моделювати рух автомобіля.
- **Часові мітки:** Важливі для синхронізації даних з камер і ІВП.

Дані, зібрані за допомогою стереокамер та ІВП, надають унікальну можливість для дослідження і тестування алгоритмів візуально-інерційного SLAM. Завдяки високій точності та різноманітності сценаріїв, він може бути використаний для розробки, навчання та валідації нових методів злиття даних, а також для покращення існуючих алгоритмів. Це особливо важливо для розробників, які прагнуть підвищити точність та надійність своїх систем в умовах, де традиційні методи можуть давати збій.

Таким чином, дані, зібрані з використанням стереокамер та ІВП, забезпечують надійну основу для подальших досліджень у галузі візуально-інерційного SLAM, відкриваючи нові можливості для розвитку автономних систем навігації.

Візуально-інерційний SLAM

Візуально-інерційний SLAM окрім візуальних даних отриманих з камери включає також інерційні дані для оцінки положення датчика та реконструкції карти. Інерційні дані отримуються за допомогою використанням інерційного вимірювального пристрою (ІВП), який складається з набору приладів, а саме: гіроскопа, акселерометра та, додатково, магнітометра. Таким чином ІВП здатний надавати інформацію щодо кутової швидкості (гіроскоп) і прискорення (акселерометр) уздовж осей x , y та z , а також, крім того, щодо магнітного поля навколо пристрою (магнітометр). Хоча додавання ІВП може збільшити інформаційну насиченість середовища та забезпечити більш високу точність, це також збільшує складність алгоритму, особливо на етапі ініціалізації. Це пов'язано з тим що, крім початкової оцінки позиції камери, алгоритм також має оцінити позицію ІВП. Алгоритми візуально-інерційного SLAM можна розділити відповідно до типу об'єднання даних камери та ІВП, які можуть бути слабо або тісно пов'язані. Методи слабого зв'язку не об'єднують дані ІВП для оцінки позиції: замість цього дані ІВП використовуються для оцінки орієнтації та змін у положенні датчика [2]. В свою чергу, тісно пов'язані методи базуються на об'єднанні даних камери та ІВП в рівняння руху, що призводить до оцінки стану, яка враховує дані обох пристроїв.

Крім того, алгоритми візуально-інерційного SLAM представляють різні реалізації відповідно до їх внутрішнього підходу, який може бути заснований на фільтрації або на основі оптимізації. Методи на основі фільтрації для візуально-інерційного SLAM покладаються на вилучення ознак, тоді як методи на основі оптимізації (також відомі як підходи на основі ключових кадрів) покладаються на глобальну оптимізацію, яка підвищує точність системи, а також обчислювальну вартість алгоритму. Результат

роботи візуально-інерційного прямого методу SLAM [3] наведено на рисунку 6. Представлені зразки зображень і реконструйовані сцени в результаті руху камери по кімнаті. Орієнтовна траєкторія руху камери показана червоним кольором [4].

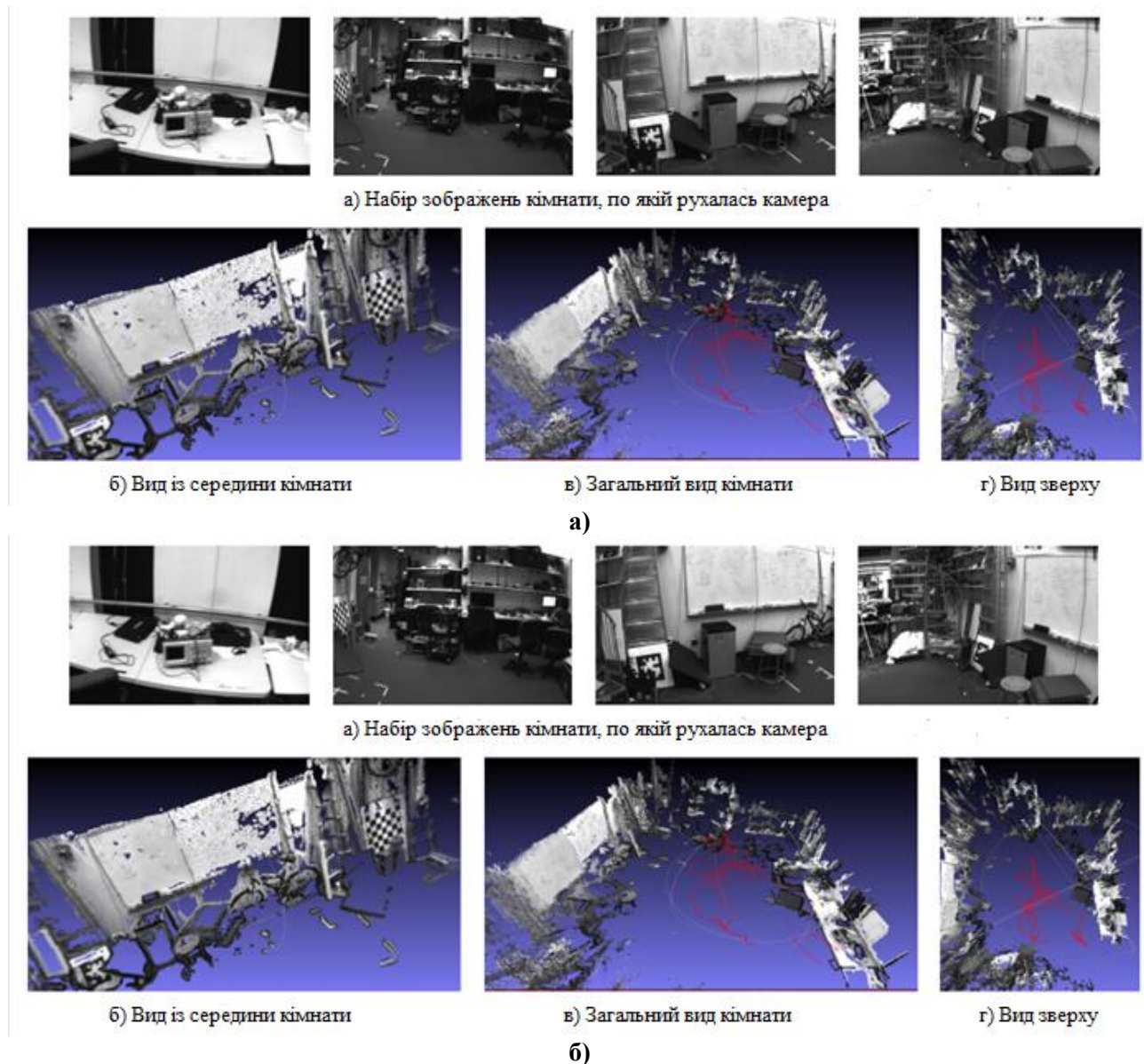


Рис. 2 – Результат роботи алгоритму візуально-інерційного прямого методу SLAM: а) зразки зображень кімнати, по якій рухається камера б) реконструйовані сцени

Технічний підхід

Розширений фільтр Калмана — це нелінійна версія фільтра Калмана, яка лінеаризує оцінку поточного середнього значення та коваріації за допомогою підходу відповідності моментів.

Використання EKF потрібне для вирішення двох завдань: локалізація ІВП через передбачення за допомогою EKF та картографування орієнтирів через оновлення EKF [5].

Щоб виконати перше завдання, потрібно знайти середнє значення $\mu \in SE$ і коваріацію $\Sigma \in R^{6 \times 6}$ прогнозованого результату. Отже, для цього потрібно буде розв'язати рівняння (1) і рівняння (2) з $w_t \sim N(0, W)$:

$$\mu_{t+1|t} = \exp(-\tau \hat{u}_t) \mu_{t|t} \quad (1)$$

$$\Sigma_{t+1|t} = \exp(-\tau \hat{u}_t) \Sigma_{t|t} \exp(-\tau \hat{u}_t)^T + W \quad (2)$$

де τ — це часова дискретизація, $\hat{u}_t \in R^{4 \times 4}$ — мапа керуючого входу \hat{u}_t , а $\hat{u}_t \in R^{6 \times 6}$ є спряженим \hat{u}_t .

Щоб виконати друге завдання, потрібно знайти середнє значення $\mu \in R^{3M}$ і коваріацію $\Sigma \in R^{3M \times 3M}$ нашого орієнтира. Отже, нам потрібно буде розв'язати рівняння (3), (4) і (5):

$$K_{t+1|t} = K_{t+1|t} H_{t+1|t}^T (H_{t+1|t} \Sigma_{t+1|t} H_{t+1|t}^T + I \otimes V) \quad (3)$$

$$\mu_{t+1|t} = \exp(K_{t+1|t} (z_{t+1} - \bar{z}_{t+1})^\wedge) \quad (4)$$

$$\Sigma_{t+1|t} = (I - K_{t+1|t} H_{t+1|t}) \Sigma_{t+1|t} \quad (5)$$

де \widetilde{z}_{t+1} – це прогнозоване спостереження, $H_{t+1|t} \in \mathbb{R}^{4N \times 6}$.

Результати

Результати роботи візуально-інерційного SLAM для 3 різних наборів даних, включаючи їхні відповідні шляхи та положення орієнтирів перед фільтрацією, наведені на рисунках 3-5. На рисунках показано розрахункову траєкторію та двовимірне положення візуальних елементів. Червона лінія на малюнку — це орієнтовна траєкторія робота, а зелені точки — положення орієнтирів. Для кожного тестового випадку порівнюються результати картографування наземних знаків на основі ІВП та візуального SLAM. Перший випадок оцінює траєкторію лише за допомогою вимірювання ІВП та прогнозування EKF і зосереджується на оцінці положення орієнтирів, тоді як другий одночасно прогнозує та оновлює положення орієнтирів і позу робота.

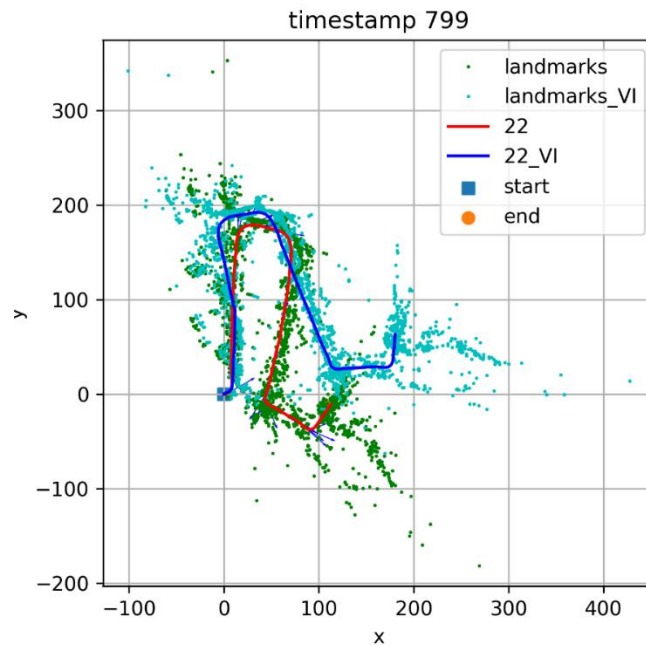


Рисунок 3 – Набір даних 22

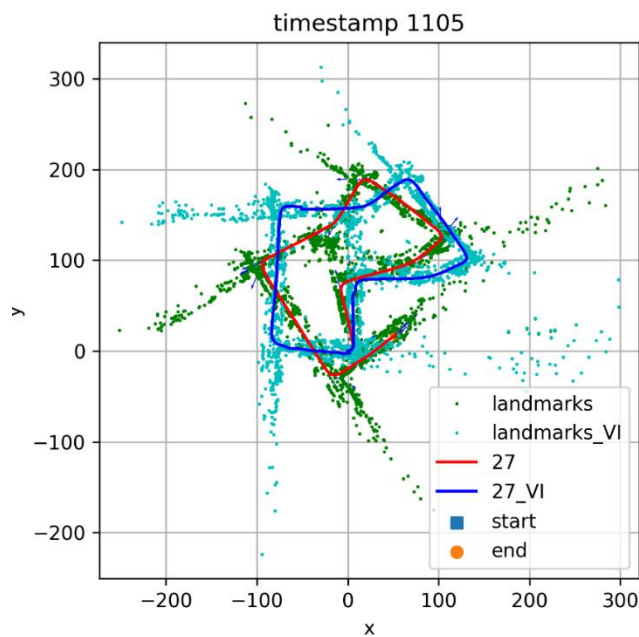


Рисунок 4 – Набір даних 27

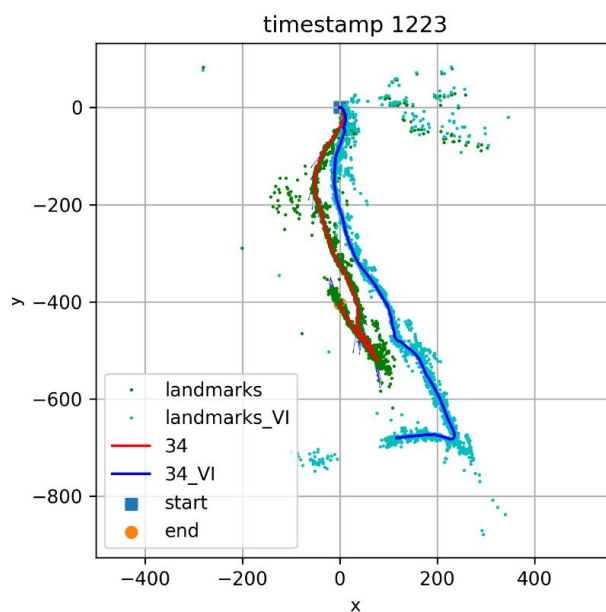


Рисунок 5 – Набір даних 33

Проаналізувавши отримані дані, можна зробити такі висновки:

- На рисунку 3 видно, що на початкових етапах шлях і орієнтири відносно не відрізнялись, проте далі чітко видно, що найбільші зміни відбуваються через поворот.
- На рисунку 4 видно, що шлях і орієнтири тут більш відхилені з самого початку, але відфільтрований шлях є набагато чистішим і точнішим порівняно з іншими прикладами.
- На рисунку 5 видно, на початку шлях і орієнтири відносно не змінилися. Однак є значне відхилення в моменті, де транспортний засіб починає поворот.

Висновки

У цій роботі досліджується використання алгоритму візуально-інерційного SLAM для навігації автономних роботів, що поєднує дані з камер та інерціальних вимірювальних пристроїв для досягнення високої точності локалізації та картографування. Основна мета — одночасна локалізація та створення карти в невідомих внутрішніх середовищах, використовуючи дані одометрії та особливі точки, виявлені стереокамерою. Для цього реалізовано розширений фільтр Калмана, що адаптує чисельні методи для лінеаризації нелінійних моделей. Запропоновані рішення можуть покращити автономні системи, такі як роботи для доставки та дрони, а також сприяти новим методам інтеграції сенсорних даних у навігаційних технологіях.

Отже, можна сказати, що алгоритм працює добре, проте ще потребує удосконалення. Для цього потрібно буде:

- Збільшити кількість орієнтирів, щоб отримати вищу точність. В роботі використовувалася лише частина доступних даних, задля економії часу побудови маршруту.
- Спробувати іншу дисперсію шуму моделі руху та спостереження.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. UCSD ECE276A: Sensing & Estimation in Robotics [Електронний ресурс]. – Режим доступу: <https://natanaso.github.io/ece276a/index.html>
2. Chen, C.; Zhu, H.; Li, M.; You, S. A Review of Visual-Inertial Simultaneous Localization and Mapping from Filtering-Based and Optimization-Based Perspectives. *Robotics* 2018, 7, 45 <https://www.mdpi.com/2218-6581/7/3/45>
3. A. Concha, G. Loianno, V. Kumar and J. Civera, "Visual-inertial direct SLAM," 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 2016, pp. 1331-1338, http://webdiis.unizar.es/~jcivera/papers/concha_etal_icra16.pdf
4. Жарков А. В., Маслій Р. В., Гармаш В. В.. Аналіз підходів VISUAL SLAM для задачі навігації автономного робота // Вісник Хмельницького національного університету. сер. Технічні науки. 2024. вип. 335. № 3(1). С. 67-77.- DOI: <https://doi.org/10.31891/2307-5732-2024-335-3-10>.
5. Simultaneous localization and mapping with the extended Kalman filter [Електронний ресурс]. – Режим доступу: https://www.iri.upc.edu/people/jsola/JoanSola/objectes/curs_SLAM/SLAM2D/SLAM%20course.pdf

Жарков Анатолій Володимирович — аспірант кафедри АІТ, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, с. Вила, Вінницького р-ну, Вінницької обл., e-mail: fkca.lakitjav@gmail.com

Маслій Роман Васильович – доцент кафедри АІТ, Вінницький національний технічний університет, Вінниця, e-mail: maslij.r.v@vntu.edu.ua

A.V. Zharkov
R.V. Maslii

APPLICATION OF VISUAL INERTIAL SLAM ALGORITHM FOR AUTONOMOUS ROBOT NAVIGATION

Vinnitsia National Technical University

This work investigates the application of the Visual-Inertial SLAM (Simultaneous Localization and Mapping) algorithm for the navigation of an autonomous robot. With the development of autonomous navigation technologies, there is a need for effective methods that allow robots to navigate in complex environments. Visual-Inertial SLAM combines data from cameras and inertial measurement devices, which ensures high accuracy of localization and map construction.

Several SLAM approaches based on particle filter, extended Kalman filter, and neural networks have been investigated over the decades. These approaches are designed to construct maps in a variety of representations, including landmark, surface, polygon grid, and occupancy grid representations. They are now widely used for applications such as robot motion planning in unknown environments and are used in self-driving cars, unmanned aerial vehicles, and autonomous underwater vehicles.

This work aims to simultaneously localize and map a robot in an unknown indoor environment using IMU odometry data and special points detected by a stereo camera. To achieve this goal, an extended approach based on the Kalman filter is implemented. The Kalman filter is an optimal linear estimator for linear system models with additive independent white noise in both forecasting and observational systems. The extended Kalman filter adapted computational methods, namely multivariate Taylor series expansions, to linearize the nonlinear model.

The proposed solutions can be useful for the improvement of autonomous systems such as delivery robots, drones and mobile platforms, as well as for the development of new methods of integrating sensor data in modern navigation technologies.

Keywords: visual inertial SLAM, inertial measurement unit, EKF.

Zharkov Anatoliy — Department of Intelligent Information Technologies and Automation, Vinnitsia National Technical University, village Vyla, Vinnitsia district, Vinnitsia region, e-mail: fkca.lakitjav@gmail.com

Maslii Roman V. – associate professor at the Department of AIIT, Vinnitsia National Technical University, Vinnitsia, email: maslij.r.v@vntu.edu.ua