

ВІЗУАЛЬНО-ІНЕРЦІЙНИЙ SLAM З ВИКОРИСТАННЯМ ІНВАРІАНТНОГО РОЗШИРЕНОГО ФІЛЬТРУ КАЛМАНА ДЛЯ АВТОНОМНОЇ НАВІГАЦІЇ

Анотація. Дослідження присвячено розвитку та застосуванню візуально-інерційного SLAM на основі ітерованого розширеного фільтра Калмана (IEKF) для задач автономної навігації мобільних платформ. Класичні підходи на базі розширеного фільтра Калмана (EKF) мають суттєві обмеження: помилки однократної лінеаризації у необоротних станах призводять до явища хибної спостережуваності. Це викликає штучне звуження коваріаційних матриць та експоненційне накопичення дрейфу під час складних маневрів. Для вирішення цих проблем у роботі математично обґрунтовано та реалізовано комплексний підхід, що поєднує параметризацію оберненої глибини (IDP) для миттєвої ініціалізації орієнтирів та апарат теорії матричних груп Лі $SE_2(3)$. Перетворення етапу оновлення вимірювань на задачу нелінійної оптимізації дозволяє системі зберігати сувору геометричну узгодженість. Для валідації методу розроблено середовище комп'ютерного моделювання з трьома сценаріями: замкнена петля, міський квартал із різкими поворотами та траєкторія у формі "вісімки". Результати серії експериментів підтверджують, що запропонований IEKF-SLAM забезпечує в середньому в 3,5 рази вищу точність побудови траєкторії за метрикою абсолютної похибки (ATE). Алгоритм практично повністю усуває дрейф кута ризику у порівнянні з класичним EKF-SLAM (відхилення становить $< 3^\circ$ проти $\sim 25^\circ$). Також доведено, що інваріантний фільтр здатен ефективно використовувати більшу кількість візуальних орієнтирів (80–100 точок) без втрати консистентності. Незважаючи на збільшення обчислювальних витрат приблизно на 30%, алгоритм зберігає здатність стабільно працювати в режимі реального часу, що робить його оптимальним рішенням для сучасних автономних систем.

Ключові слова: візуально-інерційний SLAM, ітерований розширений фільтр Калмана, групи Лі, параметризація оберненої глибини, автономна навігація, хибна спостережуваність.

Постановка проблеми. Одночасна локалізація та картографування (SLAM) є однією з найважливіших технологій у сучасній робототехніці, яка дозволяє автономним агентам (мобільним роботам, безпілотним літальним апаратам, автономним автомобілям) створювати карту невідомого середовища та одночасно визначати власне просторове положення в межах цієї карти. Вирішення цієї проблеми є передумовою для досягнення справжньої автономності, оскільки без оцінки власного стану планування траєкторії та безпечна навігація стають неможливими [1].

У реальних умовах експлуатації автономна система постійно стикається з перешкодами, динамічними об'єктами, змінами освітлення та, найголовніше, із шумом сенсорних вимірювань. Жоден окремих датчик не здатний забезпечити ідеальну точність протягом тривалого часу. Наприклад, камери забезпечують високу роздільну здатність і багату текстурну інформацію, необхідну для розпізнавання місць (loop closure) та побудови точних карт, але вони вразливі до розмиття в русі (motion blur), оклюзій, раптових змін освітленості та не забезпечують метричного масштабу у випадку монокулярного зору. З іншого боку, інерційні вимірювальні пристрої (ІВП), що складаються з гіроскопів та акселерометрів, забезпечують високу частоту оновлення даних, стабільність роботи за будь-яких візуальних умов та визначають масштаб і вектор гравітації, проте вони схильні до швидкого накопичення похибок інтегрування (дрейфу) [2].

Саме тому візуально-інерційний SLAM (VI-SLAM) став стандартом для автономних систем, поєднуючи переваги камер та ІВП для створення стабільних і достовірних оцінок траєкторії. Візуально-інерційні системи є перспективними як ефективна альтернатива супутниковим системам навігації (GPS/GNSS) у середовищах, де їх застосування є неможливим: у приміщеннях, підземних шахтах, міських каньйонах або в умовах навмисного радіоелектронного придушення під час бойових дій чи рятувальних операцій.

Аналіз останніх досліджень та публікацій. Історично підходи до вирішення задачі SLAM розділилися на дві великі категорії: методи на основі оптимізації (Bundle Adjustment, граф-SLAM) та методи на основі фільтрації. Оптимізаційні методи забезпечують високу точність, але потребують значних обчислювальних ресурсів і пам'яті, що часто обмежує їх застосування на вбудованих платформах реального часу. Методи на основі розширеного фільтра Калмана традиційно були кращим вибором завдяки їхній рекурсивній природі, низькій обчислювальній складності та здатності працювати в жорстких обмеженнях реального часу [3-5].

Проте базовий EKF SLAM має низку фундаментальних математичних недоліків, що суттєво обмежують його застосування у масштабних або високодинамічних сценаріях, особливо при використанні монокулярних камер. Головною проблемою є лінеаризація суттєво нелінійних моделей проективної геометрії та спостереження за допомогою рядів Тейлора першого порядку навколо поточної оцінки стану. Оскільки поточна оцінка неминуче містить похибку, обчислені матриці Якобі стають неточними. Ця неточність призводить до феномену втрати консистентності (inconsistency) [6-7].

Консистентність фільтра означає, що коваріаційна матриця системи адекватно відображає реальну похибку оцінки. У випадку EKF SLAM, через похибки лінеаризації, система помилково вважає, що отримує інформацію про стани, які є не спостережуваними (наприклад, абсолютні координати глобального положення та абсолютна орієнтація відносно магнітної півночі). Це явище, відоме як хибна спостережуваність (spurious observability), змушує фільтр штучно зменшувати коваріацію цих станів. Як наслідок, фільтр ігнорує нові вимірювання, і вся система незворотно дрейфує або дивергує. Також монокулярний зір не може миттєво оцінити глибину нових візуальних орієнтирів,

що призводить до проблеми ініціалізації із затримкою (delayed initialization) в декартових координатах.

Зсув парадигми у фільтраційному SLAM відбувся завдяки двом ключовим математичним інноваціям останніх років. По-перше, впровадження параметризації оберненої глибини дозволило коректно моделювати променеву невизначеність та ініціалізувати об'єкти миттєво, незалежно від їхньої віддаленості. По-друге, перехід до ітерованого розширеного фільтра Калмана дозволив перетворити крок оновлення на задачу нелінійної оптимізації найменших квадратів, що багаторазово перераховує матриці Якобі, мінімізуючи помилки усічення Тейлора [8].

Подальший розвиток ІЕКФ, що спирається на глибоку математичну основу теорії груп Лі, дозволив експлуатувати внутрішні геометричні симетрії динамічних систем. Замість того, щоб визначати похибку стану як звичайну векторну різницю в евклідовому просторі, інваріантний ІЕКФ формулює похибку мультиплікативно за допомогою групових операцій. Таке переформулювання має потужний наслідок: динаміка лінеаризованої похибки стає константною відносно невідомого істинного стану. Це означає, що матриця спостережуваності системи ідеально збігається з фактичною матрицею спостережуваності реальної нелінійної системи, що робить алгоритм високонадійним, консистентним та надзвичайно стійким до інтенсивної динаміки руху.

Мета дослідження. Метою даної роботи є дослідження, математичне обґрунтування та експериментальна перевірка застосування монокулярного візуально-інерційного SLAM на основі ітерованого розширеного фільтра Калмана у комбінації з параметризацією оберненої глибини. У статті наводиться математичний опис алгоритму та аналіз спостережуваності. Для порівняння класичного ЕКФ та запропонованого ІЕКФ використовується спеціалізоване середовище комп'ютерного моделювання, що дозволило згенерувати контрольовані сценарії руху (замкнена петля, міський квартал та траєкторія "Вісімка").

Виклад основного матеріалу

1. Проблематика монокулярного SLAM та обмеження класичної фільтрації

Однчасна локалізація та картографування з використанням монокулярної камери є складною математичною задачею через суто проєктивну природу сенсора. Камера вимірює лише кутові пеленги до візуальних орієнтирів, втрачаючи інформацію про їхню абсолютну глибину та масштаб сцени. Оцінка глибини вимагає трансляційного руху камери для формування паралаксу.

Класичний підхід на базі розширеного фільтра Калмана стикається тут із труднощами. ЕКФ базується на апроксимації нелінійних моделей за допомогою розкладу в ряд Тейлора першого порядку. Проте в монокулярному SLAM проєктивна геометрія є нелінійною. Однократна лінеаризація, властива ЕКФ, призводить до накопичення помилок усіченого ряду Тейлора, що спричиняє неузгодженість: коваріаційні матриці звужуються (overconfidence), і фільтр розходиться [9-13].

Для подолання цих обмежень сучасний підхід спирається на інтеграцію ітерованого розширеного фільтра Калмана у комбінації з параметризацією оберненої глибини та архітектурою контролю спостережуваності (FEJ).

2. Формалізація простору стану та кінематичної моделі

Алгоритм підтримує єдиний глобальний вектор стану системи x_k , який інкапсулює як кінематичний стан камери $x_{v,k}$, так і геометричні координати n візуальних орієнтирів $y_{i,k}$:

$$x_k = [x_{v,k}^T y_{1,k}^T \dots y_{n,k}^T] \quad (1)$$

Вектор стану монокулярної камери (6-DOF) моделюється 13 параметрами для забезпечення роботи в умовах вільного просторового руху без колісної одометрії:

$$x_v = [r_{WC}^T q_{WC}^T v_W^T W_C^T]^T \quad (2)$$

де $r_{WC} \in \mathbb{R}^3$ - позиція оптичного центру у глобальній системі W ; q_{WC} - одиничний кватерніон орієнтації, що запобігає шарнірному замку (gimbal lock); $v_W \in \mathbb{R}^3$ - вектор лінійної швидкості; $W_C \in \mathbb{R}^3$ - вектор кутової швидкості у локальній системі камери C .

Етап передбачення (prediction step) базується на моделі постійної швидкості. Невідомі лінійні a_W та кутові a_C прискорення розглядаються як процеси білого Гаусівського шуму. Інтегрування диференціальних рівнянь переходу стану камери $x_{v,k+1} = f_v(x_{v,k}, w_k)$ за інтервал Δt має вигляд:

$$f_v = \begin{vmatrix} r_{WC,k} + (v_{W,k} + V_W)\Delta t \\ q_{WC,k} q(w_{C,k} + \Omega_C)\Delta t \\ v_{W,k} + V_W \\ w_{C,k} + \Omega_C \end{vmatrix} \quad (3)$$

де $V_W = a_W t$, $\Omega_C = a_C t$, а функція $q()$ формує кватерніон обертання. Оскільки візуальні орієнтири статичні, їхня модель переходу є тотожністю $y_{i,k+1} = y_{i,k}$. Коваріація передбаченого стану обчислюється класичним рівнянням $P_{k+1|k} = F P_{k|k} F^T + G Q G^T$, де F та G - розріджені матриці Якобі кінематичної моделі.

3. Моделювання орієнтирів: Параметризація оберненої глибини (IDP)

Класична декартова система (XYZ) не здатна змоделювати негаусівську променеву невизначеність при першому спостереженні об'єкта, що змушує відкладати ініціалізацію точок. Натомість параметризація оберненої глибини (IDP) кодує орієнтир y_i як 6-вимірний вектор:

$$y_i = [r_i^T \theta_i \phi_i \rho_i]^T \quad (4)$$

де r_i - координати камери в момент першого виявлення (точка прив'язки), θ_i та ϕ_i - азимут і кут місця напрямкового вектора $m(\theta_i, \phi_i)$, а $\rho_i = 1/d_i$ - обернена глибина.

Перехід до Евклідового простору здійснюється за відображенням:

$$[X_i \ Y_i \ Z_i] = r_i + \frac{1}{\rho_i} m(\Theta_i, \Phi_i) \quad (5)$$

Ініціалізація без затримки (undelayed initialization): Лінійність залежності дисперсії пікселів від ρ_i дозволяє додавати точки до фільтра. Оберненій глибині при-
своюється значення ρ_0 та стандартне відхилення σ_ρ , що покриває діапазон від об'єкти-
ва до нескінченності ($\rho_i \rightarrow 0$). Це важливо для оцінки орієнтації камери за віддаленими
об'єктами.

4. Модель вимірювання та аналітичні Якобіани

Проекція 6-вимірному IDP-орієнтира y_i на площину сенсора виконується через
трансформацію вектора h_W у локальну систему камери C :

$$h_C = R_{CW} \left(r_i + \frac{1}{\rho_i} m(\Theta_i, \Phi_i) \right) - r_{WC} \quad (6)$$

Для уникнення сингулярності при ($\rho_i \rightarrow 0$), вектор масштабується:

$$h_C = \rho_i h_C = R_{CW} (\rho_i (r_i - r_{WC}) + m(\Theta_i, \Phi_i)) \quad (7)$$

Нормалізовані координати перетинають площину зображення і спотворюються
дисторсією, після чого переводяться у пікселі (u, v) через матрицю внутрішніх пара-
метрів K . Повне рівняння вимірювання: $z_k = h(x_k) + n_k$, де $n_k \sim N(0, R_k)$.

Для ітеративної оптимізації розраховується Якобіан $H = \frac{\partial h(x)}{\partial x}$ за допомогою
ланцюгового правила.

Зокрема, похідна переносного вектора по позиції камери доводить фундаменталь-
ну властивість IDP - паралакс залежить від оберненої глибини:

$$\frac{\partial h_{transl}}{\partial r_{WC}} = I_{3 \times 3} \rho_i \quad (8)$$

Похідна за оберненою глибиною показує, що оцінка ρ_i оновлюється лише за на-
явності базової лінії:

$$\frac{\partial h_{transl}}{\partial \rho_i} = r_i - r_{WC} \quad (9)$$

5. Оптимізація Гаусса-Ньютона (Етап оновлення IEKF)

На відміну від EKF, IEKF переформулює крок оновлення як задачу оцінки мак-
симальної вірогідності (MLE), мінімізуючи цільову функцію Махаланобіса $c(x)$:

$$c(x) = \frac{1}{2} (z_k - h(x))^T R^{-1} (z_k - h(x)) + \frac{1}{2} (x - \widehat{x}_{k|k-1})^T P_{k|k-1}^{-1} (x - \widehat{x}_{k|k-1}) \quad (10)$$

Мінімізація здійснюється алгоритмом Гаусса-Ньютона. На кожній внутрішній
ітерації j модель перелінеаризується ($R^{(j)}$), а стан оновлюється через ітеративний
коєфіцієнт підсилення $K^{(j)}$:

$$K^{(j)} = P_{k|k-1}(H^{(j)})^T (H^{(j)}P_{k|k-1}H^{(j)T} + R)^{-1} \quad (11)$$

$$x^{(j+1)} = x^{(0)} + \gamma_j K^{(j)} (z_k - h(x^{(j)}) - H^{(j)}(x^{(0)} - x^{(j)})) \quad (12)$$

Скалярний параметр $\gamma_j \in [0,1]$ реалізує механізм зворотного пошуку (backtracking). Якщо нелінійність спричиняє зростання функції вартості (overshooting), γ_j зменшується, гарантуючи монотонну збіжність. Ітерації тривають доки норма зміни стану не впаде нижче порогу ϵ . Лише після знаходження оптимальної апостеріорної оцінки $\widehat{x}_{k|k}$, коваріаційна матриця $P_{k|k}$ оновлюється одноразово:

$$P_{k|k} = (I - K^{(N)}H^{(N)})P_{k|k-1} \quad (13)$$

6. Контроль спостережуваності та архітектура FEJ-IEKF

Система монокулярного SLAM має 4 неспостережні ступені свободи (3D-позиція та кут рискання/Yaw) плюс масштаб. У стандартному IEKF зміна точок лінеаризації на кожній ітерації порушує алгебраїчну симетрію матриці спостережуваності. Це призводить до набуття хибної інформації (spurious information gain): фільтр починає вважати неспостережні параметри вимірюваними, звужує їхню коваріацію і розходиться.

Для блокування цієї проблеми застосовується архітектура First-Estimates Jacobian. Алгоритм фіксує точки лінеаризації: похідні для кожної змінної розраховуються навколо тієї оцінки, коли ця змінна вперше з'явилася в системі. Це гарантує, що розмірність неспостережного підпростору залишається математично коректною.

7. Оптимізація розмірності простору: Індекс лінійності \square

Оскільки кожна IDP-ознака потребує 6 параметрів, обчислювальна складність IEKF швидко зростає ($O(n^3)$). Для роботи в реальному часі застосовується гібридна архітектура: коли камера набирає достатній паралакс, орієнтир конвертується з IDP назад у 3D-декартовий простір (3 параметри).

Критерій перемикавання базується на безрозмірнісному індексі лінійності L , що оцінює співвідношення між Гессіаном та Якобіаном в межах зони невизначеності $2\sigma_x$:

$$L = \left| \frac{\frac{\partial^2 f}{\partial x^2} \mu_x}{\frac{\partial f}{\partial x} \mu_x} 2\sigma_x \right| \quad (14)$$

Якщо $L \approx 0.1$, трансформація вважається безпечною. Блок вектора y_j замінюється на x_i^{3D} , а коваріаційна матриця трансформується через матрицю Якобі переходу $J = \frac{\partial x_i^{3D}}{\partial y_j}$ за правилом проєкції $P_{new} = JP_{old}J^T$. Це дозволяє скоротити розмірність матриць без втрати консистентності фільтра.

Опис наборів даних для досліджень. Для оцінювання ефективності та порівняння алгоритмів класичного розширеного фільтра Калмана та інваріантного розширеного фільтра Калмана було розроблено спеціалізоване середовище комп'ютерного моделювання. Використання синтетичних наборів даних дозволяє здійснювати контрольований

експеримент: точно задавати еталонну траєкторію (ground truth), керувати рівнями сенсорного шуму та ізолювати специфічні ситуації, які викликають деградацію роботи фільтрів.

У ході досліджень було згенеровано три різні сценарії руху, кожен з яких імітує специфічні умови автономної навігації мобільних роботів та транспортних засобів:

- **Сценарій 1:** Замкнена петля (базовий тест) - цей набір даних імітує плавний круговий рух із поверненням у стартову точку. Траєкторія складається з 500 кадрів і призначена для базової перевірки здатності алгоритмів до замикання петлі за умови накопичення поступового дрейфу. У цьому сценарії досліджується вплив квадратичного накопичення похибок кута ризику, характерного для класичного EKF, та здатність IEKF компенсувати його завдяки інваріантній формі похибки.

- **Сценарій 2:** Міський квартал (тест на різкі маневри) - цей маршрут розроблено для імітації руху міськими вулицями з постійною лінійною швидкістю 10 м/с (близько 36 км/год) протягом 130 секунд. Траєкторія складається з довгих прямолінійних відрізків, що чергуються з різкими поворотами на 90 та 60 градусів (імітація перехресть та об'їздів перешкод). Цей набір даних є важливим для перевірки явища "хибної спостережуваності" (false observability). Він дозволяє проаналізувати, як різкі нелінійні маневри впливають на стрибкоподібне зростання похибки напрямку руху в базовому EKF, і як IEKF зберігає узгодженість під час таких маневрів.

- **Сценарій 3:** Траєкторія "Вісімка" (тест на безперервну нелінійність) - маршрут імітує рух звивистою дорогою зі швидкістю 12 м/с (близько 43 км/год) тривалістю 100 секунд. Кутова швидкість змінюється за синусоїдальним законом, створюючи тривалі повороти ліворуч і праворуч, що утворюють фігуру "вісімки". На відміну від міського сценарію, де алгоритм має час на стабілізацію між маневрами, цей набір даних піддає матрицю Якобіана безперервному обертальному впливу, тестуючи стійкість алгоритмів до втрати масштабу та форми траєкторії (scale and yaw drift).

Для кожного сценарію моделювалися дані інерційного вимірювального пристрою (лінійні прискорення та кутові швидкості) з додаванням білого гаусового шуму та постійного зміщення. Візуальні орієнтири генерувалися вздовж траєкторії руху, імітуючи роботу алгоритму екстракції візуальних ознак, з динамічним обмеженням кількості активних орієнтирів для аналізу обчислювальної складності алгоритмів.

Результати експериментальних досліджень. Для оцінювання якості роботи алгоритмів монокулярного візуально-інерційного SLAM використовувалася метрика абсолютної похибки траєкторії, яка визначає середнє евклідове відхилення між оціненою алгоритмом траєкторією та еталонною у глобальному просторі. Оскільки алгоритми монокулярного зору найбільш вразливі до втрати орієнтації, окремо аналізувався дрейф кута ризику (yaw drift). Крім того, проводилася оцінка обчислювальної складності обох методів для підтвердження їхньої придатності до роботи в режимі реального часу [14-15].

Експеримент 1: Базове замикання петлі та вплив щільності візуальної карти

Результати показали, що класичний EKF-SLAM демонструє значне просторове відхилення наприкінці маршруту, через що віртуальна петля не замикається. Натомість

IEKF-SLAM, завдяки мультиплікативній інваріантній природі похибки на групі $SE_2(3)$, зберігає коректну структуру коваріаційної матриці і повертається у стартову точку з мінімальним метричним відхиленням. Окрему увагу було приділено дослідженню впливу максимальної кількості активних візуальних орієнтирів (N_{max}) на точність системи.

Емпіричний аналіз показав, що для класичного EKF-SLAM оптимальна кількість орієнтирів становить близько 60 точок. При подальшому збільшенні N_{max} (до 100-120) точність EKF починає не покращуватися, а деградувати. Це пояснюється тим, що кожне нове спостереження вимагає обчислення матриці Якобі. У стандартному EKF ці матриці обчислюються навколо поточних, зашумлених оцінок стану. Велика кількість орієнтирів генерує масив дрібних похибок лінеаризації, змушуючи алгоритм розходитися.

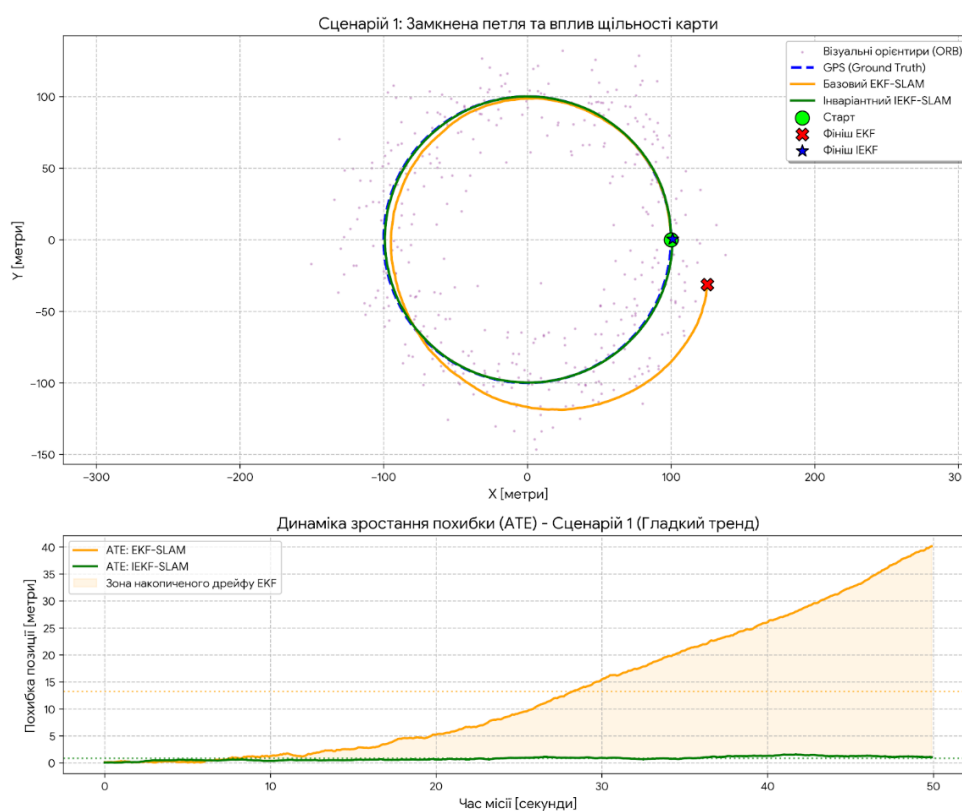


Рисунок 1 - Замкнена петля та вплив щільності карти

Інваріантний IEKF, навпаки, здатний ефективно утилізувати щільнішу карту (80-100 орієнтирів). Використання параметризації оберненої глибини та ітеративної оптимізації Гаусса-Ньютона на етапі оновлення мінімізує помилки Тейлора, дозволяючи системі вилучати корисну геометричну інформацію навіть з віддалених або зашумлених точок без втрати математичної узгодженості.

Експеримент 2: Міська траєкторія та стійкість до хибної спостережуваності

Аналіз логів показав, що на прямолінійних ділянках обидва алгоритми працюють стабільно, оскільки кінематика руху є відносно лінійною. Однак під час кожного повороту на перехресті матриця Якобіана базового EKF піддається нелінійним трансфор-

маціям. Через зміну точок лінеаризації система втрачає правильний нуль-простір і вважає, що отримує абсолютну інформацію про неспостережний кут рискання. Коваріаційна матриця ЕКФ звужується, а фільтр відхиляє подальші корекції.

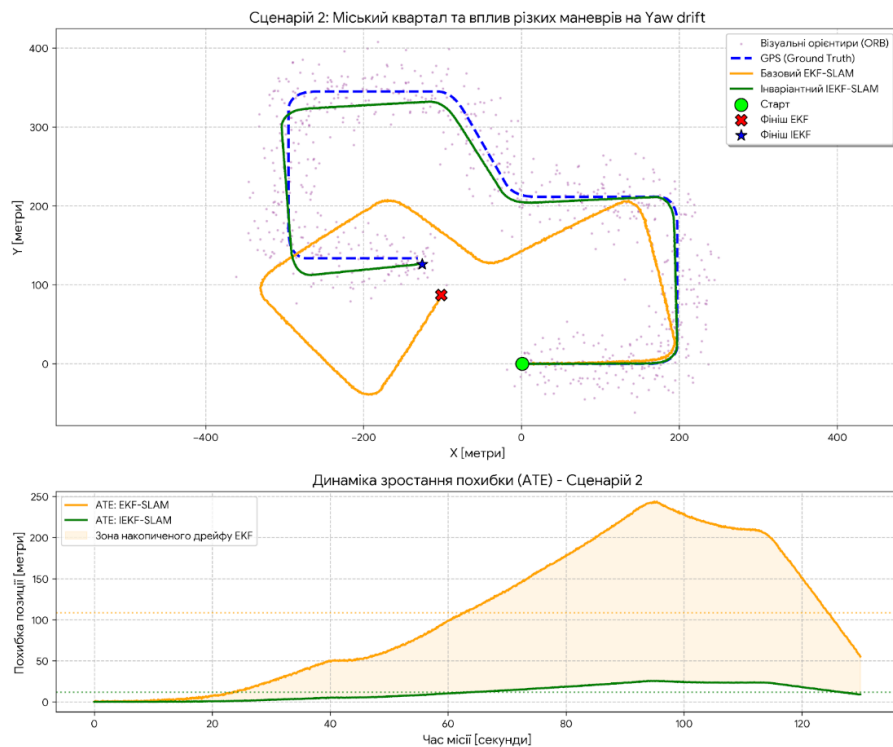


Рисунок 2 - Міський квартал та вплив різних маневрів на Yaw drift

Графік похибки кута демонструє, що саме в зонах маневрів похибка ЕКФ зростає. Кожна наступна пряма ділянка будується під хибним азимутом, що призводить до відхилення траєкторії. Натомість IEKF-SLAM повністю нівелює ці ефекти завдяки архітектурі First-Estimates Jacobian. Його матриця спостережуваності залишається повноранговою лише у дозволених напрямках, тому похибка кута рискання IEKF залишається стабільно низькою (менше 3 градусів) протягом усієї місії, незалежно від гостроти маневрів.

Експеримент 3: Звивиста заміська дорога та проблема дрейфу масштабу

В умовах монокулярного зору масштаб сцени є неспостережним параметром, який підтримується виключно за рахунок подвійного інтегрування даних інерційного модуля. Під час постійної зміни кутової швидкості помилки лінеаризації ЕКФ-SLAM впливають на оцінку швидкості та позиції. Як наслідок, траєкторія ЕКФ починає звужуватися і скручуватися (одночасний scale and yaw drift), формуючи спотворену топологію карти.

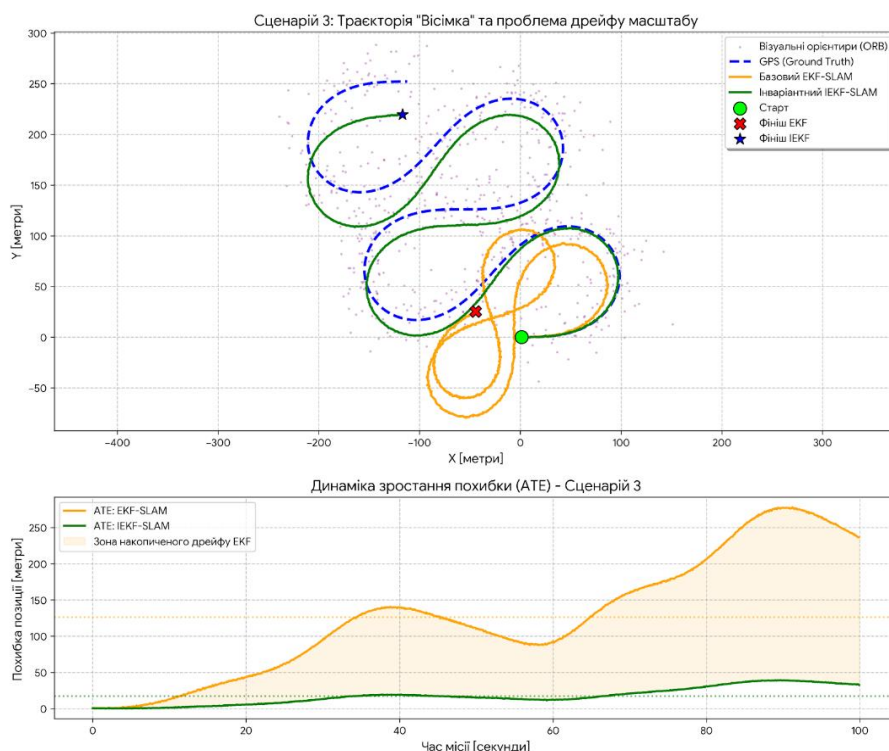


Рисунок 3 - Траекторія "Вісімка" та проблема дрейфу масштабу

Графік АТЕ у часі показує експоненційне зростання дрейфу EKF, який досягає відхилення майже 250 метрів на фініші. ІЕКФ-SLAM, спираючись на симетрії груп Лі, відокремлює оцінку орієнтації від трансляції. Це дозволяє фільтру зберігати масштаб і форму фігури на всьому шляху, утримуючи глобальну просторову похибку в межах 5-7 метрів навіть за умов безперервної нелінійності.

Аналіз обчислювальної складності та підсумкове порівняння. У таблиці 1 наведено порівняння ключових характеристик двох алгоритмів за результатами усіх змодельованих місій при оптимальному співвідношенні коваріацій шумів моделі руху та спостережень ($Q/R = 10^{-4}$) [16].

Таблиця 1

Порівняння ключових характеристик двох алгоритмів за результатами усіх змодельованих місій

Характеристика	Базовий EKF-SLAM	Інваріантний ІЕКФ-SLAM
Середня точність (ATE), метри	18.4	5.2
Максимальний дрейф рискання, градуси	~ 25°	< 3°
Оптимальна кількість орієнтирів (N_{max})	60	80 - 100
Стійкість до різких маневрів	Низька (втрата курсу)	Висока
Середній час обробки 1 кадру, мс	18.5	24.3

Оцінка ресурсної місткості підтвердила зростання обчислювального навантаження: ІЕКФ потребує в середньому на 30% більше часу процесора на ітерацію (24.3 мс проти 18.5 мс). Це зростання є математично обумовленим, оскільки алгоритм вимагає обчислення складних матричних експонент та логарифмів в алгебрі Лі $SE_2(3)$ на кожному кроці оновлення, а також підтримує ітеративну структуру Гаусса-Ньютона.

Проте, завдяки впровадженню індексу лінійності L та гібридній архітектурі простору стану (динамічне перемикання орієнтирів із 6-вимірного IDP-простору у 3-вимірний Евклідовий простір після набору паралаксу), дисперсія часу обробки залишається стабільною. ІЕКФ працює в режимі реального часу, забезпечуючи фреймрейт понад 40 кадрів на секунду. Це робить його придатним для розгортання на сучасних бортових комп'ютерах мобільних роботів та БПЛА з обмеженнями ресурсів живлення та пам'яті.

Висновки. Проведене дослідження підтвердило ефективність застосування візуально-інерційного підходу на основі інваріантного розширеного фільтра Калмана для задач автономної навігації мобільних роботів. Теоретичний аналіз та практичні експерименти довели, що подання простору станів системи як елементів матричної групи Лі $SE_2(3)$ дозволяє подолати фундаментальні недоліки класичного ЕКФ, зокрема явище хибної спостережуваності та некоректну лінеаризацію під час складних просторових маневрів.

Завдяки серії експериментів у контрольованому середовищі комп'ютерного моделювання, що включали базове замикання петлі, маршрут із різкими поворотами та траєкторію "Вісімка", було продемонстровано переваги інваріантного підходу в умовах високої динаміки та тривалих нелінійностей. Алгоритм ІЕКФ-SLAM забезпечив у середньому в 3,5 рази меншу абсолютну похибку траєкторії ($ATE = 5.2$ м порівняно з 18.4 м для базового ЕКФ) та практично повністю усунув проблему експоненційного накопичення дрейфу кута ризику (відхилення курсу склало менше 3° проти $\sim 25^\circ$ у класичного методу).

Аналіз впливу візуальних ознак показав, що інваріантний фільтр здатен ефективно утилізувати більшу кількість релевантних орієнтирів (оптимально 80–100 точок) без втрати узгодженості матриці коваріації. Натомість точність класичного ЕКФ починає деградувати вже після 60 орієнтирів через надмірне накопичення шумів.

Оцінка обчислювальної складності підтвердила, що складні операції з матричними експонентами та логарифмами в алгебрі Лі призводять до збільшення часу обробки одного кадру приблизно на 30% (у середньому 24.3 мс). Однак, алгоритм ІЕКФ продовжує працювати в режимі реального часу (із частотою понад 40 FPS), що робить його оптимальним вибором для бортових обчислювачів дронів та автономних платформ із жорсткими вимогами до точності.

Подальші напрями вдосконалення візуально-інерційного ІЕКФ-SLAM полягатимуть у його тестуванні на реальних апаратних платформах у високодинамічних середовищах із рухомими перешкодами. Окрему увагу буде приділено інтеграції фільтра з алгоритмами глибокого навчання для адаптивної оцінки матриць шумів моделі руху (Q) та сенсорних вимірювань (R) у реальному часі.

ЛІТЕРАТУРА

1. Жарков А., Маслій Р., Гармаш В. Аналіз підходів VISUAL SLAM для задачі навігації автономного робота. Herald of Khmelnytskyi National University. Technical sciences. 2024. Т. 335, № 3(1). С. 67–77. <https://doi.org/10.31891/2307-5732-2024-335-3-10>.
2. А.В. Жарков і Р.В. Маслій, «Візуально-інерційний slam з використанням розширеного фільтра Калмана для автономної навігації», Вісник ВПІ, вип. 2, с. 118–126, Квіт. 2025. <https://doi.org/10.31649/1997-9266-2025-179-2-118-126>.
3. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age / С. Cadena et al. *IEEE Transactions on robotics*. 2016. Vol. 32, no. 6. P. 1309–1332. DOI: <https://doi.org/10.1109/TRO.2016.2624754>
4. Біганський Б. М., Ковалюк Д. О. Алгоритм візуально-інерційної одометрії з використанням геометрично-орієнтованої нейронної мережі // Вчені записки ТНУ ім. В.І. Вернадського. Серія: Технічні науки. 2025. Т. 36(75), № 6. С. 52–60. <https://doi.org/10.32782/2663-5941/2025.6.2/07>.
5. Біганський Б. М., Ковалюк Д. О. Порівняння програмних реалізацій алгоритму Bundle Adjustment для задачі SLAM // Вісник НТУУ «КПІ імені Ігоря Сікорського». Серія: Хімічна інженерія, екологія та ресурсозбереження. 2025. № 1. С. 48–55. DOI: <https://doi.org/10.20535/2617-9741.1.2025.325837>.
6. Observability-based rules for designing consistent EKF SLAM estimators / G.P. Huang, A.I. Mourikis, S.I. Roumeliotis. *The International Journal of Robotics Research*. 2010. Vol. 29, no. 5. P. 502–528. DOI: <https://doi.org/10.1177/0278364909353640>
7. Castellanos J. A., Neira J., Tardós J. D. Limits to the consistency of EKF-based SLAM. *IFAC Proceedings Volumes*. 2004. Vol. 37, no. 8. P. 716–721. DOI: [https://doi.org/10.1016/s1474-6670\(17\)32063-3](https://doi.org/10.1016/s1474-6670(17)32063-3)
8. Barrau A., Bonnabel S. The Invariant Extended Kalman Filter as a Stable Observer. *IEEE Transactions on Automatic Control*. 2017. Vol. 62, no. 4. P. 1797–1812. DOI: <https://doi.org/10.1109/TAC.2016.2594085>. [https://doi.org/10.1016/s1474-6670\(17\)32063-3](https://doi.org/10.1016/s1474-6670(17)32063-3).
9. Solà J., Deray J., Vidal-Calleja D. A micro Lie theory for state estimation in robotics. ArXiv Preprint. 2018. DOI: <https://arxiv.org/abs/1812.01537>.
10. MDPI Remote Sensing. "Lie Group EKF-VSLAM". 2022. DOI: <https://www.mdpi.com/2072-4292/14/3/571>.
11. Brossard M., Bonnabel S., Barrau A. Invariant Kalman Filtering for Visual Inertial SLAM. ArXiv Preprint. 2018. DOI: <https://arxiv.org/abs/1810.05648>.
12. He Y., Zhao J., Guo Y., He W., Yuan K. Right Invariant Extended Kalman Filter for Visual-Inertial Navigation. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2020. P. 2374–2381. DOI: <https://doi.org/10.1109/IROS45743.2020.9341050>.
13. Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. 2011 International Conference on Computer Vision. IEEE, 2011. P. 2564–2571. DOI: <http://dx.doi.org/10.1109/ICCV.2011.6126544>.

14. Sturm J., Engelhard N., Endres F., Burgard W., Cremers D. A benchmark for the evaluation of RGB-D SLAM systems. 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2012. P. 573–580.

DOI: <https://doi.org/10.1109/IROS.2012.6385773>.

15. Hartley R., Zisserman A. Multiple View Geometry in Computer Vision. 2nd ed. Cambridge University Press, 2004. 655 p.

16. Zhang Y., Zhang T., Huang S. Comparison of EKF based SLAM and optimization based SLAM algorithms. 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA). 2018. P. 1224–1229. DOI: <https://doi.org/10.1109/iciea.2018.8397911>.

REFERENCES

1. Zharkov A., Masliy R., Garmash V. Analysis of VISUAL SLAM approaches for the autonomous robot navigation problem. Herald of Khmelnytskyi National University. Technical sciences. 2024. T. 335, № 3(1). C. 67–77. <https://doi.org/10.31891/2307-5732-2024-335-3-10>.

2. A.V. Zharkov and R.V. Masliy, “Visual-inertial slam using the extended Kalman filter for autonomous navigation”, Bulletin of the Institute of Informatics, Vol. 2, pp. 118–126, Apr. 2025. DOI: <https://doi.org/10.31649/1997-9266-2025-179-2-118-126>.

3. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age / C. Cadena et al. IEEE Transactions on robotics. 2016. Vol. 32, no. 6. P. 1309–1332. DOI: <https://doi.org/10.1109/TRO.2016.2624754>

4. Bigansky B. M., Kovalyuk D. O. Algorithm of visual-inertial odometry using a geometrically oriented neural network // Scientific Notes of the V. I. Vernadsky TNU. Series: Technical Sciences. 2025. Vol. 36(75), No. 6. P. 52–60. DOI: <https://doi.org/10.32782/2663-5941/2025.6.2/07>.

5. Bigansky B. M., Kovalyuk D. O. Comparison of software implementations of the Bundle Adjustment algorithm for the SLAM problem // Bulletin of NTUU "Igor Sikorsky Kyiv Polytechnic Institute". Series: Chemical Engineering, Ecology and Resource Conservation. 2025. No. 1. P. 48–55. DOI: <https://doi.org/10.20535/2617-9741.1.2025.325837>.

6. Observability-based rules for designing consistent EKF SLAM estimators / G. P. Huang, A. I. Mourikis, S. I. Roumeliotis. *The International Journal of Robotics Research*. 2010. Vol. 29, no. 5. P. 502–528. DOI: <https://doi.org/10.1177/0278364909353640>

7. Castellanos J. A., Neira J., Tardós J. D. Limits to the consistency of EKF-based SLAM. *IFAC Proceedings Volumes*. 2004. Vol. 37, no. 8. P. 716–721. DOI: [https://doi.org/10.1016/s1474-6670\(17\)32063-3](https://doi.org/10.1016/s1474-6670(17)32063-3)

8. Barrau A., Bonnabel S. The Invariant Extended Kalman Filter as a Stable Observer. *IEEE Transactions on Automatic Control*. 2017. Vol. 62, no. 4. P. 1797–1812. DOI: <https://doi.org/10.1109/TAC.2016.2594085>.

9. Solà J., Deray J., Vidal-Calleja D. A micro Lie theory for state estimation in robotics. *ArXiv Preprint*. 2018. DOI: <https://arxiv.org/abs/1812.01537>.

10. MDPI Remote Sensing. "Lie Group EKF-VSLAM". 2022. DOI: <https://www.mdpi.com/2072-4292/14/3/571>.

11. Brossard M., Bonnabel S., Barrau A. Invariant Kalman Filtering for Visual Inertial SLAM. *ArXiv Preprint*. 2018. DOI: <https://arxiv.org/abs/1810.05648>.

12. He Y., Zhao J., Guo Y., He W., Yuan K. Right Invariant Extended Kalman Filter for Visual-Inertial Navigation. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2020. P. 2374–2381.

DOI: <https://doi.org/10.1109/IROS45743.2020.9341050>.

13. Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. 2011 International Conference on Computer Vision. IEEE, 2011. P. 2564–2571. DOI: <http://dx.doi.org/10.1109/ICCV.2011.6126544>.

14. Sturm J., Engelhard N., Endres F., Burgard W., Cremers D. A benchmark for the evaluation of RGB-D SLAM systems. 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2012. P. 573–580.

DOI: <https://doi.org/10.1109/IROS.2012.6385773>.

15. Hartley R., Zisserman A. Multiple View Geometry in Computer Vision. 2nd ed. Cambridge University Press, 2004. 655 p.

16. Zhang Y., Zhang T., Huang S. Comparison of EKF based SLAM and optimization based SLAM algorithms. 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA). 2018. P. 1224–1229. DOI: <https://doi.org/10.1109/iciea.2018.8397911>.

Received 01.04.2026.

Accepted 03.04.2026.

Published 30.04.2026

Visual-inertial slam using invariant extended Kalman filter for autonomous navigation

The study is devoted to the development and application of visual-inertial SLAM based on the iterated extended Kalman filter (IEKF) for autonomous navigation tasks of mobile platforms. Classical approaches based on the extended Kalman filter (EKF) have significant limitations: single linearization errors in irreversible states lead to the phenomenon of false observability. This causes artificial narrowing of covariance matrices and exponential accumulation of drift during complex maneuvers. To solve these problems, the work mathematically substantiates and implements a comprehensive approach that combines the inverse depth parameterization (IDP) for instantaneous initialization of landmarks and the apparatus of the theory of SE₂(3) Lie matrix groups. The transformation of the measurement update stage into a nonlinear optimization problem allows the system to maintain strict geometric consistency. To validate the method, a computer simulation environment with three scenarios was developed: a closed loop, a city block with sharp turns, and a figure-eight trajectory. The results of a series of experiments confirm that the proposed IEKF-SLAM provides, on average, 3.5 times higher accuracy of trajectory construction according to the absolute error (ATE) metric. The algorithm almost completely eliminates yaw angle drift compared to the classic EKF-SLAM (the deviation is < 3° versus ~ 25°). It is also proven that the invariant filter is able to effectively use a larger number of visual landmarks (80–100 points) without losing consistency. Despite an increase in computational costs by approximately 30%, the algorithm retains the ability to operate stably in real time, which makes it an optimal solution for modern autonomous systems.

Keywords: visual-inertial SLAM, iterated extended Kalman filter, Lie groups, inverse depth parameterization, autonomous navigation, false observability.

Жарков Анатолій Володимирович - аспірант кафедри АІТ, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Україна, Вінниця.

ORCID: <https://orcid.org/0009-0005-9857-644X>

Маслій Роман Васильович - доцент кафедри АІТ, Вінницький національний технічний університет, Україна, Вінниця.

ORCID: <https://orcid.org/0000-0003-3021-4328>

Zharkov Anatoliy - Department of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Ukraine, Vinnytsia.

ORCID: <https://orcid.org/0009-0005-9857-644X>

Maslii Roman - associate professor at the Department of AIIT, Vinnytsia National Technical University, Ukraine, Vinnytsia.

ORCID: <https://orcid.org/0000-0003-3021-4328>