

ЛАКОЗА Сергій

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

<https://orcid.org/0000-0001-6354-1611>

s.lakoz@kpi.ua

СТЕПАНОВ Євген

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

<https://orcid.org/0009-0000-2307-7121>

canstep@gmail.com

ДАНЧУКОВ Олег

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

<https://orcid.org/0009-0007-1970-244X>

oleg.donn03@gmail.com

РОЗРОБКА НЕЙРОННОЇ МЕРЕЖІ ДЛЯ ОЦІНКИ ПАРАМЕТРІВ СТОХАСТИЧНОЇ МОДЕЛІ ШУМІВ У СИГНАЛІ АКСЕЛЕРОМЕТРА

МЕМС-акселерометри широко застосовуються у багатьох прикладних задачах, проте вони характеризуються значними випадковими та систематичними похибками. Це обмежує їх точність у реальних умовах експлуатації. Сучасні роботи у сфері калібрування та оцінювання параметрів акселерометрів спрямовані на підвищення автономності, адаптивності та стійкості до зовнішніх впливів. Проте вони, зазвичай, обмежуються умовами навчання або ресурсами апаратного забезпечення. Це формує потребу розробки методів оцінювання, здатних працювати в умовах високої невизначеності, з обмеженою інформацією та з процесами принципово стохастичними по своїй природі. Для цих задач актуальними є алгоритми машинного навчання, які здатні моделювати нелінійні залежності, враховувати часові та частотні властивості шумових процесів.

У статті розроблена нейронна мережа для оцінки параметрів стохастичних процесів у сигналі акселерометра. Розглянуто питання розробки та моделювання роботи такої нейронної мережі. У нейронній мережі застосовано архітектуру, яка одночасно обробляє часову та спектральну інформацію. Розроблена модель нейронної мережі має дві паралельні гілки: часову та спектральну. Після окремої роботи результати кожної з гілок об'єднуються у спільний латентний простір і подаються на фінальний регресор, який оцінює параметри моделі шумових сигналів. У якості моделі стохастичних процесів обрано гаусівсько-марківські процеси. У роботі описані питання щодо навчання розробленої нейронної мережі, надано рекомендації щодо генерації навчальних даних. Для підвищення стабільності та ефективності навчання використовується два стандартні механізми оптимізації процесу тренування – *EarlyStopping* та *ReduceLROnPlateau*. Це дозволяє підвищити якість збіжності, зменшити ризик перенавчання та забезпечити отримання моделі з високими узагальнюючими властивостями. Отримана модель нейронної мережі продемонструвала здатність ефективно оцінювати параметри стохастичних процесів, що формують випадкові шуми МЕМС-акселерометрів.

Ключові слова: МЕМС-акселерометр, оцінювання, нейронні мережі, калібрування, шум, стохастичні моделі

LAKOZA Serhii, STEPANOV Yevhen, DANCHUKOV Oleh

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

DEVELOPMENT OF NEURAL NETWORK FOR STOCHASTIC NOISE MODEL PARAMETERS ESTIMATING IN ACCELEROMETER SIGNAL

MEMS accelerometers are widely used in many applications but they have significant random and systematic errors. It limits their accuracy under real-world conditions of work. Modern studies in the field of calibration and accelerometers parameter estimation are focused at increasing autonomy, adaptability and resilience to external influences. These studies are usually limited by training conditions or power of hardware resources. Thus it requires the development of estimation methods capable to operate under high uncertainty, with limited information, with processes that nature is fundamentally stochastic. For these tasks are highly relevant machine learning algorithms which capable model non-linear dependencies, to account time and spectral properties of noise processes.

In this work neural network for estimating parameters of stochastic processes in accelerometer signal is developed. The issues of designing and modeling operation of such neural network are considered. The neural network utilizes an architecture that simultaneously processes time and spectral information. The developed neural network model features two parallel branches: time and spectral. After separate information processing, the results from each branch are combined into shared latent space and fed into the final regressor. It that estimates the noise signal model parameters. Gauss-Markov processes are selected as the model for the stochastic processes. In this paper the aspects of training the developed neural network are described and recommendations for training data generation are provided. To increase training stability and efficiency, two standard training process optimization mechanisms are used — *EarlyStopping* and *ReduceLROnPlateau*. This allows to improve convergence quality, to reduce risk of overfitting and ensures the acquisition of the model with high generalization capabilities. The developed neural network model demonstrates effective ability to estimate the of stochastic processes parameters that form the random noise in MEMS-accelerometers signal.

Keywords: MEMS-accelerometer, estimation, neural network, calibration, noise, stochastic models



ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Мікроелектромеханічні (МЕМС) акселерометри широко застосовуються у безпілотних літальних апаратах, робототехніці, транспортних системах та мобільній електроніці. Їх популярність зумовлена низькою вартістю, компактністю та можливістю роботи в автономних навігаційних системах. Проте МЕМС-акселерометри характеризуються значними випадковими та систематичними похибками, що обмежує їх точність у реальних умовах експлуатації. Точність визначення орієнтації безпілотного літального апарата суттєво залежить від характеристик інерціальної вимірювальної системи, зокрема від шумових властивостей МЕМС-акселерометрів. Вплив температури, вібрацій, нелінійностей, випадкових процесів усередині сенсора безпосередньо впливають на алгоритми розв'язання задач орієнтації та навігації, оскільки акселерометр є джерелом інформації про вектор прискорення, який використовується для оцінювання нахилів апарату (крен та тангаж) та визначення місцеположення. У реальних умовах шумові збурення призводять до накопичення похибок, спотворення сигналу та зниження точності фільтрації.

Білий шум акселерометра є швидкозмінним, некорельованим та рівномірно розподіленим у всій смузі частот. У системах орієнтації він проявляється як випадкові коливання вимірюваного вектора прискорення. У разі використання фільтрів типу Маджвіка, Махоні або комплементарного фільтра білий шум впливає на коливання оцінених кутів нахилу, збільшує дисперсію оцінки, зменшує точність у високочастотному діапазоні [1]. Оскільки кути нахилу визначаються з нормалізованого прискорення, то наявність шуму призводить до помилки в оцінці гравітаційного вектора та, відповідно, до похибок кутів крену і тангажу. Дана похибка не накопичується у часі, але суттєво погіршує миттєву точність.

Корельовані шуми (нестабільність нульового сигналу, випадкові коливання) мають значно серйозніший вплив. Вони проявляються повільним зсувом вимірюваного значення прискорення. У задачі визначення орієнтації це означає, що фільтр отримує гравітаційний вектор з систематичним зміщенням, а це призводить до спотворень оцінки кутів, виникає псевдонахил апарата, фільтр генерує повільні помилки, що накопичуються. Це особливо критично при малих прискореннях, коли внесок власної динаміки БПЛА є мінімальним, і гравітаційний вектор має найбільшу вагу в алгоритмі.

У сучасних БПЛА часто використовуються комбіновані рішення, у яких прискорення інтегрується для оцінювання швидкості та положення. У такому разі корельовані шуми та дрейф нуля акселерометра призводять до більш серйозних наслідків: похибка швидкості зростає лінійно, похибка положення — квадратично, зміщення нуля формує псевдоприскорення, що імітує рух об'єкта з постійним прискоренням, появляються повільні систематичні тренди в оцінках стану [1, 2]. Таким чином, для інерціальної навігації навіть невеликий дрейф нуля призводить до накопичення розходження траєкторії без зовнішніх поправок.

Традиційні методи калібрування ґрунтуються на статичних або багатопозиційних процедурах та забезпечують прийнятний результат лише в лабораторних умовах. У динамічних та невизначених середовищах їх ефективність суттєво знижується. Тому актуальним є застосування алгоритмів машинного навчання, які здатні моделювати нелінійні залежності, враховувати часові та частотні властивості шумових процесів. Однак більшість робіт зосереджена або на систематичних похибках, або на використанні класичних підходів без глибокого аналізу стохастичної структури сигналу.

Метою даної роботи є розроблення нейронної мережі, яка здатна оцінювати параметри стохастичної моделі випадкової складової у сигналі акселерометра.

МОЖЛИВОСТІ НЕЙРОННИХ МЕРЕЖ ДЛЯ ОЦІНЮВАННЯ МОДЕЛЕЙ СИГНАЛІВ АКСЕЛЕРОМЕТРІВ

Штучна нейронна мережа (ШНМ) — це обчислювальна модель, яка побудована по аналогії функціонування біологічних нейронних мереж, зокрема людського мозку. Вона складається з взаємопов'язаних елементів обробки, які називаються штучними нейронами або вузлами, що працюють разом для обробки та передачі інформації. ШНМ використовуються для розв'язання широкого кола задач, що пов'язані з розпізнаванням та прийняттям рішень. Серед задач, які розв'язуються нейронними мережами виділяють кілька основних типів: класифікація, регресія, кластеризація, прогнозування та оптимізації поведінки агентів [3].

Після тривалого періоду спаду інтересу до нейронних мереж, у середині 1980-х розпочався їх новий етап розвитку завдяки роботам Девіда Румельхарта, Джеффрі Гінтона та Рональда Вільямса, які у 1986 році опублікували статтю «Learning representations by back-propagating errors». У ній було математично обґрунтовано принцип навчання багатопшарових перцептронів (Multilayer Perceptrons, MLP). Важливим теоретичним результатом стало доведення теореми універсальної апроксимації, згідно з якою багатопшарова

нейронна мережа з одним прихованим прошарком і нелінійною функцією активації здатна апроксимувати будь-яке неперервне відображення з довільною точністю [4]. Подальший розвиток обчислювальних ресурсів і поява великих наборів даних у 2000-х роках сприяли поширенню архітектур з глибинним навчання (deep learning), зокрема згорткових нейронних мереж та рекурентних мереж. Вони стали основою сучасних систем розпізнавання образів, мовлення та обробки природної мови [5].

Актуальними у використанні для різної обробки даних з первинних сенсорів є наступні архітектури нейронних мереж: багатошаровий перцептрон (MLP), згорткова нейронна мережа (CNN), рекурентна нейронна мережа (RNN), LSTM, GRU.

Повнозв'язні багатошарові перцептрони нейронні мережі (MLP) є універсальними апроксиматорами неперервних функцій, які здатні вирішувати задачі класифікації та регресії. Основним недоліком MLP є велика кількість параметрів та відсутність механізмів локальної інваріантності, що ускладнює роботу з великими зображеннями чи послідовностями [3, 4].

До переваг CNN відносять: автоматичне виділення ознак, локальність сприйняття, спільність ваг (один і той самий фільтр використовується для всіх ділянок сигналу чи зображення), ієрархічне подання (перші шари вчать розпізнавати прості ознаки, а глибші – складні патерни), стійкість до шуму та зсувів, ефективність для великих даних, можливість адаптації до нових задач з меншим обсягом даних, підтримка різних типів вхідних даних.

До недоліків CNN відносяться: висока обчислювальна складність, велика потреба у даних, «чорна скринька» (важко пояснити, які параметри виявляються і чому), чутливість до гіперпараметрів, втрати просторової інформації при зменшенні розмірності карт, складність генерації поза навчальною вибіркою, потенційна надмірність глибоких мереж.

Розглядаючи CNN у контексті аналізу вихідного сигналу з датчиків можна сказати, що даний тип нейронних мереж здатен виявляти особливості сигналу та кореляцію між даними. CNN є придатним інструментом для вирішення питань обробки сигналів датчиків, але не єдиним [5].

RNN – це тип мережі, в якій присутній зворотний зв'язок (рекурсія), що дозволяє враховувати попередні стани при обробці поточних вхідних даних. На відміну від CNN або MLP, де всі входи незалежні, RNN має пам'ять. Тобто RNN вміє пам'ятати, що було раніше, у момент часу $t-1$.

RNN добре підходять для роботи з часовими рядами, послідовностями вимірювань і сигналів, де поточне значення залежить від попередніх. Це робить їх природним вибором для задач аналізу сигналів з датчиків, прогнозування станів або компенсації дрейфу у вимірювальних системах. Проте базові RNN мають обмеження – при обробці довгих послідовностей градієнти можуть зникати або вибухати, що ускладнює навчання. Для подолання цих проблем були розроблені розширені архітектури – мережі з довгою короткочасною пам'яттю (LSTM) та мережі на основі гейтованих рекурентних блоків (GRU), які дозволяють ефективно зберігати інформацію на довгих часових інтервалах.

LSTM-мережі використовують спеціальні механізми керування потоком інформації – «вхідні», «вихідні» та «забувальні» вентиля. Вони визначають, які дані слід зберегти у внутрішньому стані, а які – відкинути. Завдяки цьому LSTM можуть моделювати складні часові залежності. Це робить їх ефективними для задач динамічного калібрування акселерометрів, де сигнал має інерційний характер і залежить від попередніх вимірювань.

Модифікація LSTM на основі GRU (Gated Recurrent Unit) спрощує архітектуру LSTM, об'єднуючи частину вентилів і зменшуючи кількість параметрів, що робить її придатною для реалізації у вбудованих системах з обмеженими ресурсами. У контексті оцінки параметрів акселерометрів такі мережі можуть навчатися безпосередньо на потоках даних у реальному часі, забезпечуючи компенсацію дрейфу, температурних і механічних похибок без необхідності зовнішніх еталонів. Отже, рекурентні нейронні мережі, особливо LSTM та GRU, відкривають можливості для створення адаптивних систем калібрування, здатних враховувати часові залежності сигналів та змінювати параметри в процесі роботи, що є критично важливим у динамічно змінюваних умовах.

АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ

У статті [6] розглянуто підхід, який використовує MLP для калібрування параметрів масивів MEMS-акселерометрів. Автори показали, що сучасні масиви інерційних датчиків все частіше застосовуються в складних навігаційних задачах, де класичні лінійні моделі калібрування не враховують міжсигнальні залежності, нелінійні впливи конструкції та динамічні збурення в реальному часі.

Запропонований метод базується на одношаровому або багатошаровому перцептроні, який приймає на вхід вихідні сигнали всіх сенсорів масиву та додаткові контекстні параметри (наприклад, швидкість руху, темп зміни температури). Мережа навчається на попередньо зібраному наборі даних, що містить вимірювання з відомими калібрувальними параметрами. Після навчання модель здатна враховувати нелінійні взаємозв'язки між вимірюваннями різних акселерометрів, коригувати зміщення, масштабні коефіцієнти та перехресну чутливість у режимі неперервної оцінки.

Основною перевагою розробленого підходу є можливість одночасної компенсації декількох видів

похибок без необхідності створення складних аналітичних моделей для кожного елементу окремо. Це знижує об'єм ручного налаштування та спрощує інтеграцію калібрування у кінцеві навігаційні алгоритми. Автори також підкреслюють, що мережа демонструє високу стабільність і можливість узагальнення на вимірювання, які раніше не зустрічалися під час тренування.

Недоліками використаного підходу автори відзначають потребу у широкому спектрі навчальних даних, що охоплюють усі режими експлуатації масиву датчиків, а також необхідність ретельної вибірки ознак для уникнення перенавчання. Також робота не розглядає вплив часових залежностей сигналу, що може обмежувати точність оцінювання в умовах сильних вібрацій або змінних нелінійних динамічних режимах. Незважаючи на це, стаття демонструє ефективність нейронних підходів для комплексної корекції похибок у масивах MEMS-акселерометрів та підтверджує перспективність MLP методів у задачах калібрування сенсорів [6].

У статті [7] пропонується підхід до калібрування недорогих інерціальних вимірювальних модулів (IBM) із використанням глибокої згорткової нейронної мережі (CNN). Авторами відповідно розроблено нейронну мережу Calib-Net. Розроблений підхід орієнтований на IBM із широко розповсюджених споживчих пристроїв, де традиційні методи калібрування не забезпечують достатньої точності через нелінійні спотворення, перехресну чутливість та шумові компоненти, що залежать від динаміки руху.

Ключова ідея полягає у використанні CNN для виділення релевантних ознак із часових сигналів акселерометрів та гіроскопів без потреби в складних попередніх моделях або ручних обчислювальних міркуваннях. Структура мережі побудована таким чином, що вона приймає на вхід вікна вимірювань IBM, а на виході повертає оцінки калібрувальних параметрів (зміщення нуля, масштабні коефіцієнти, компоненти перехресної чутливості).

Перевагами методу є здатність одночасно оцінювати кілька параметрів калібрування без розділення на окремі підзадачі, а також автоматичне виділення виразних патернів у даних, що складно формалізувати аналітично. Експериментальна оцінка показала значне зменшення похибок IBM після корекції Calib-Net порівняно з класичними підходами.

Певні обмеження у використанні вносить вимога великого обсягу тренувальних даних, зібраних у різних режимах руху, а також можливе зниження точності при екстраполяції поза навчальну доменну область. Незважаючи на це, робота демонструє практичну ефективність глибинних CNN для калібрування бюджетних IBM і підтверджує перспективність даного напрямку при формуванні високоточних навігаційних систем [8].

У статті [9] розроблена радіально базисна нейронна мережа для самокалібрування MEMS IBM. У роботі показано вбудоване навчання із застосуванням гаусових нейронних топологій: підхід орієнтований на адаптацію до змінних у часі значень зміщень та теплових ефектів, що є корисним у сценаріях, де характеристики датчиків змінюються під час експлуатації. Основна увага авторів зосереджена на реалізації онлайн-адаптації моделі. Це надає можливості оновлювати параметри калібрування під час роботи сенсора, використовуючи вбудовані обчислення, які дозволяють виконувати корекцію похибок безпосередньо на мікроконтролері або сенсорному модулі без залучення зовнішніх обчислювальних ресурсів.

Запропонована у статті RBF-NN формує локальні апроксимації похибок за допомогою гаусово-базисних функцій. Її ключовою особливістю є динамічне оновлення центрів і радіусів базисних функцій у процесі роботи сенсора. Завдяки цьому мережа здатна адаптуватися до змін середовища, температури, механічного навантаження та деградації елементів. Перевагою такого підходу є висока швидкість, що дозволяє застосовувати його у вбудованих системах БПЛА, робототехніці та переносних пристроях. Окрім того, алгоритм ефективно використовує пам'ять мікроконтролера, оскільки використовує компактну топологію, яка збільшується лише за потреби. Експерименти, приведені у статті, показують здатність моделі зменшувати зміщення нуля та похибки масштабного коефіцієнта IBM у змінних режимах.

Серед недоліків автори відзначають чутливість алгоритму до вибору початкових центрів та параметрів радіальних функцій, а також можливість накопичення помилок за тривалих періодів роботи без зовнішніх корекційних сигналів [10].

У статті [11] запропоновано алгоритм самокалібрування інерціальних вимірювальних модулів, встановлених на двоколісних транспортних засобах. Основна проблема, яку розглядають автори, полягає у тому, що телематичні модулі (e-Box) часто монтуються з відхиленнями. Відповідно осі IBM не збігаються з системою координат транспортного засобу, що спричиняє додаткові похибки у вимірюваннях.

Розроблений авторами алгоритм складається з двох послідовних кроків. Спочатку визначаються кути крену та тангажу шляхом порівняння орієнтації вимірюваного вектора прискорення з напрямком гравітації. Робиться припущення, що транспортний засіб рухається рівною поверхнею зі сталою швидкістю, а вертикальне прискорення є малим.

Другий крок – оцінювання кута курсу. Після визначення крену та тангажу автори використовують вимірювання кутових швидкостей гіроскопа та кінематичні співвідношення, притаманні руху двоколісних транспортних засобів. Приймаючи припущення, що бокові прискорення малі, курс визначають через узгодження вимірюваних і очікуваних гіроскопічних сигналів.

Запропонований метод є повністю автономним, не потребує стендових умов і може виконуватись під

час звичайної експлуатації. Основною перевагою є можливість корекції похибок установки ІВМ без додаткових датчиків або маркерів орієнтації. Недоліки пов'язані з тим, що алгоритм покладається на спрощені динамічні припущення, які можуть порушуватися під час різких маневрів або руху по нерівній поверхні. Незважаючи на це, автори демонструють успішну експериментальну перевірку та високу точність у стандартних міських умовах [12].

У статті [13] розглядається проблема оптимізації процесу калібрування інерційних МЕМС-датчиків у промислових умовах. Автори відзначають, що традиційні послідовні методи калібрування не враховують індивідуальні відмінності між сенсорами, викликані виробничими варіаціями, що призводить до надмірних витрат часу та ресурсів. Для вирішення цієї проблеми запропоновано інтелектуальну систему калібрування, яка поєднує алгоритми машинного навчання з системою моніторингу виробничого процесу.

Ключовим елементом підходу є створення «розумного» фреймворку калібрування, що використовує історичні вимірювальні дані для прогнозування оптимальної робочої точки датчика. Як базову модель автори застосували регресійний ансамблевий алгоритм XGBoost, який аналізує дані на рівні компонентів і сенсорів, визначаючи індивідуальні параметри налаштування для кожного зразка. Додатково система забезпечує моніторинг у реальному часі – відстежує стабільність процесу калібрування та автоматично виявляє аномалії або відхилення.

Результати експериментальної перевірки показали, що впровадження розробленої моделі дозволяє скоротити загальний час калібрування приблизно на 23,8 % без втрати точності вимірювань. Крім того, інтегрована система моніторингу підвищує гнучкість виробничого процесу, забезпечуючи стабільність і адаптивність калібрування при зміні умов чи партій сенсорів.

Таким чином, робота [13] демонструє ефективність інтелектуальних підходів до калібрування у промислових умовах, але водночас підкреслює необхідність подальшого розвитку у напрямку використання глибоких нейронних мереж, здатних працювати в умовах високої невизначеності.

Huang і Wu у роботі [14] запропонували підхід до динамічного калібрування акселерометрів на основі моделі «sequence-to-sequence», побудованої з використанням довготривалої короткочасної пам'яті (LSTM). Метою дослідження було відновлення високоточного сигналу акселерометра за даними з дешевших МЕМС-сенсорів, що зазнають значних спотворень під час високошвидкісних або ударних навантажень.

Розроблена архітектура LSTM-мережі навчалася на парах синхронізованих сигналів двох типів датчиків, що дозволяло моделі вивчати часові закономірності похибок і прогнозувати коригований сигнал у режимі реального часу. Отримані результати продемонстрували, що LSTM-мережі здатні ефективно апроксимувати нелінійні динамічні спотворення та відновлювати форму «еталонного» сигналу з прийнятною точністю навіть у разі короткочасних ударів.

Водночас автори відзначають, що такі рекурентні архітектури мають вищі обчислювальні вимоги, що ускладнює їх застосування на енергообмежених вбудованих платформах. Крім того, точність навчання значною мірою залежить від наявності великої кількості пар даних «низької» та «високої» якості, що обмежує універсальність методу. Незважаючи на це, робота [14] є важливим прикладом застосування рекурентних нейронних мереж для динамічного калібрування та компенсації похибок акселерометрів у реальному часі.

Проведений аналіз сучасних наукових праць показує, що калібрування та оцінювання параметрів акселерометрів і інерційних вимірювальних модулів еволюціонує від класичних аналітичних та регресійних методів до інтелектуальних систем, побудованих на основі машинного та глибокого навчання. Ранні підходи, такі як багатопарові перцептрони і методи регресії забезпечують високу точність у статичних або контрольованих умовах, але втрачають ефективність при зміні середовища чи динамічних навантажень. У таких умовах краще себе проявляють рекурентні та згорткові нейронні архітектури, які навчилися враховувати часові та просторові залежності сигналів, що дозволяє підвищити точність оцінювання у динамічних сценаріях.

Таким чином, можна зробити висновок, що сучасні дослідження у сфері калібрування та оцінювання параметрів акселерометрів спрямовані на підвищення автономності, адаптивності та стійкості до зовнішніх впливів. Проте більшість існуючих рішень залишаються обмеженими або умовами навчання (наявністю пар даних «еталон/спотворення»), або ресурсами апаратного забезпечення. Це формує науково-практичну потребу для розроблення методів оцінювання, здатних працювати в умовах високої невизначеності, з обмеженою інформацією та принципово стохастичними процесами по своїй природі.

ГАУСІВСЬКО-МАРКІВСЬКІ ПРОЦЕСИ ЯК МОДЕЛІ СТОХАСТИЧНИХ ЗБУРЕНЬ АКСЕЛЕРОМЕТРА

Навіть у стані спокою МЕМС-акселерометр має випадкові складові у вихідному сигналі. Ці збурення мають внутрішнє походження та не пов'язані із зовнішніми механічними чи температурними впливами. Основними джерелами таких збурень є: термічні флуктуації механічної конструкції датчика, 1/f-шум (фліккер-шум) електронних схем, механічні мікрорезонанси та паразитні моди коливань, електростатичні флуктуації ємнісного мосту, корельований шум аналогових підсилювачів та АЦП [15]. Ці процеси формують такі типи випадкових похибок: білий шум прискорення, дрейф зсуву нуля, низькочастотні корельовані шуми,

повільні зміни масштабного коефіцієнта.

Важливо відзначити, що масштабний коефіцієнт не дрейфує як шумовий процес. Його зміни, як правило, є детермінованими та спричинені повільними варіаціями температури, механічними напруженнями та старінням матеріалу. Тому основну увагу в рамках стохастичного моделювання зосереджують на білих та корельованих шумах, а також на дрейфі зсуву нуля, які є справді стохастичними [15].

Стохастичні процеси відіграють ключову роль у моделюванні похибок МЕМС-акселерометрів, оскільки значна частина внутрішніх збурень сенсора має випадкову природу та проявляється у вигляді шумів різного спектрального характеру. Одним із найбільш поширених класів стохастичних моделей, що застосовуються для опису таких збурень, є гаусівсько-марківські процеси (ГМП) [16].

Гаусівський процес визначається тим, що будь-яка його скінченна вибірка має багатовимірний нормальний розподіл. Марківський процес характеризується тим, що майбутній стан залежить лише від поточного, а не від усієї попередньої історії. Таким чином, гаусівсько-марківський процес – це стохастичний процес, у якому теперішній стан повністю описує ймовірнісний прогноз подальшої еволюції, а всі реалізації процесу мають нормальний розподіл [17]. Марківські властивості роблять ГМП особливо зручними в задачах прогнозування та фільтрації, зокрема в алгоритмах типу Калмана.

ГМП першого та другого порядків. ГМП першого порядку описується рівнянням стохастичної динаміки виду

$$\dot{x}(t) = -\frac{1}{\tau}x(t) + \sigma \omega(t), \quad (1)$$

де $x(t)$ – корельований шум, τ – час кореляційного процесу, σ – інтенсивність шуму, $\omega(t)$ – білий гаусівський шум.

Цей процес формує низькочастотну складову шуму сенсора: швидкі випадкові зміни згладжуються, а тренд змінюється повільно. У спектральній області ГМП першого порядку характеризується підсиленням на низьких частотах, що добре відповідає природі дрейфу нульового сигналу МЕМС-акселерометрів [16, 18].

Для практичного застосування у моделюванні та подальшому використанні в нейронних мережах (тобто в дискретному часі) неперервне рівняння (1) приводиться до дискретної форми:

$$x_{k+1} = e^{-\frac{\Delta t}{\tau}} x_k + \sigma \sqrt{1 - e^{-\frac{2\Delta t}{\tau}}} \omega_k, \quad (2)$$

де $\omega_k \sim N(0, 1)$.

Отримане рівняння (2) дозволяє моделювати корельований шум акселерометра з правильною автокореляцією й стаціонарною дисперсією.

Для моделювання шумів, що зумовлені механічними резонансами конструкції МЕМС-датчиків, використовується розширена модель – ГМП другого порядку. Вона описує випадкові коливання демпфованого осцилятора під дією білого шуму:

$$\ddot{x}(t) + 2\zeta\omega_0\dot{x}(t) + \omega_0^2x(t) = \sigma \omega(t), \quad (3)$$

де ω_0 – власна частота паразитної моди, ζ – коефіцієнт демпфування, σ – інтенсивність шумового збурення [18].

Така модель адекватно відтворює вузькосмугові компоненти спектра, що викликаються мікрорезонансами, нерівномірністю пружних властивостей або коливанням підсилювальних ланцюгів. Особливо важливо це для сенсорів низької вартості, де механічні паразитні моди проявляються сильніше [16].

$$x_{k+1} = e^{-\frac{\Delta t}{\tau}} x_k + \sigma \sqrt{1 - e^{-\frac{2\Delta t}{\tau}}} \omega_k. \quad (4)$$

При моделюванні вищеприведену модель прийнято записувати у формі простору станів. Введемо вектор зміни стану $\mathbf{s} = \begin{bmatrix} x \\ \dot{x} \end{bmatrix}$ і перейдемо до матрично-векторної форми, тоді отримуємо запи:

$$\dot{\mathbf{s}}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{B}\omega(t), \quad (5)$$

де

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -2\zeta\omega_0 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 \\ \sigma \end{bmatrix}. \quad (6)$$

Дискретна форма ГМП другого порядку має вид:

$$\mathbf{s}_{k+1} = \mathbf{F} \mathbf{s}_k + \mathbf{w}_k. \quad (7)$$

Дискретна перехідна матриця та матриця коваріації дорівнюють відповідно

$$\mathbf{F} = e^{\mathbf{A}\Delta t},$$

$$\mathbf{Q}_d = \sigma^2 \int_0^{\Delta t} e^{\mathbf{A}\tau} \mathbf{B} \mathbf{Q}_c \mathbf{B}^T e^{\mathbf{A}^T \tau} d\tau, \quad (8)$$

де $\mathbf{Q}_c$ — безперервна коваріація джерела шуму.

При побудові нейронних моделей калібрування МЕМС-акселерометрів особливо важливо забезпечити, щоб тренувальні дані відтворювали реальні статистичні властивості шумів сенсора. Використання гаусівсько-марківських процесів дає можливість: синтезувати правдоподібні послідовності

корельованих шумів, коректно моделювати довготривалий дрейф зсуву нуля, уникати перенавчання моделі на ідеалізованих або фізично нереалістичних даних, покращувати здатність нейронної мережі компенсувати низькочастотні складові похибки [16].

ГМП стають ключовим математичним інструментом при моделюванні внутрішніх шумів МЕМС-акселерометрів та забезпечують адекватність набору даних, що використовуються для навчання.

СТРУКТУРА НЕЙРОННОЇ МЕРЕЖІ ТА ЇЇ НАВЧАННЯ ДЛЯ ОЦІНКИ СТОХАСТИЧНИХ ПОХИБОК МЕМС-АКСЕЛЕРОМЕТРІВ

Розроблена нейронна мережа (НМ) для оцінки параметрів стохастичних процесів у сигналі акселерометра показана на рис.1. У представлений НМ застосовано архітектуру, яка одночасно обробляє часову та спектральну інформацію, що дозволяє підвищити точність оцінювання стохастичних характеристик сенсора.

Оскільки стохастична складова сигналу акселерометра є складною за структурою, то для ідентифікації її параметрів, а саме білого та корельованого шумів, необхідно враховувати його спектральну складову, іншими словами – розкласти сигнал у частотний спектр. Тому розроблена модель НМ має дві паралельні гілки: часову та спектральну.

Часова гілка забезпечує виділення локальних кореляційних структур, повільних змін, квазістатичного дрейфу, поведінки станів ГМП у часі та впливу зміщення нуля.

Спектральна гілка обробляє спектральні складові сигналу, отримані шляхом застосування дискретного швидкого перетворення Фур'є (ШПФ). Ця гілка дозволяє ефективно виявляти такі параметри ГМП, як: натуральна частота ω_0 , коефіцієнт демпфування ζ , інтенсивність шумового збурення σ , шум вимірювання R [19].

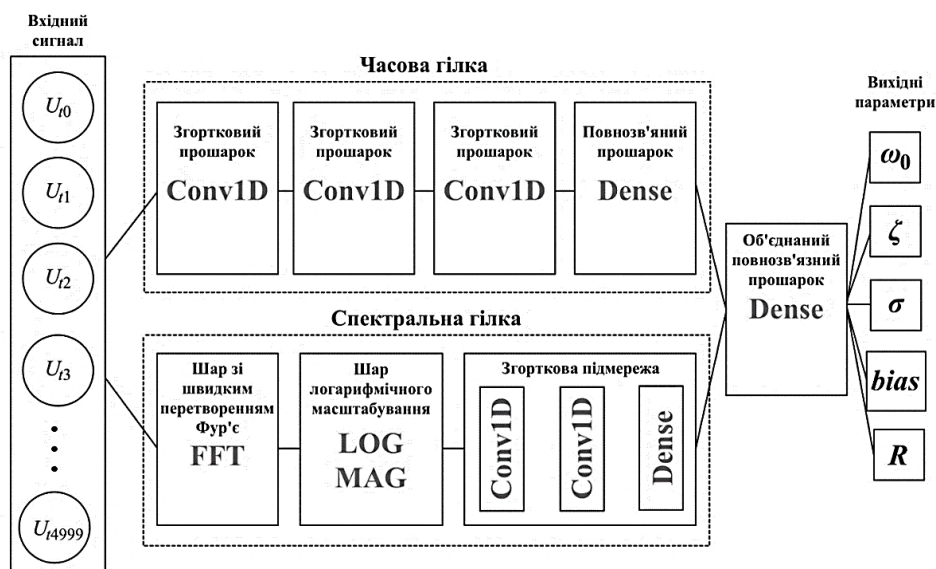


Рис. 1. Схема НМ для оцінки стохастичної складової сигналу акселерометра

Особливістю часової гілки є використання трьох одновимірних згорткових прошарків: Conv1D (32, kernel_size=9, activation function=ELU), Conv1D(64, kernel_size=9, activation function=ELU), Conv1D(128, kernel_size=9, activation function=ELU). Між прошарками застосовано відповідно функції стиснення бібліотеки Keras: MaxPool1D, MaxPool1D та GlobalAveragePooling1D. Це дозволяє витягнути складові локальної кореляції шумів, фільтрувати високочастотні компоненти, виділити поведінку шумової моделі у часі (ГМП-вклад), розпізнати зміщення та повільні дрейфи. Отже, на виході часової гілки ми отримуємо вектор ознак розмірністю 128.

В основі роботи спектральної гілки, як було написано вище, лежить швидке перетворення Фур'є:

$$X_f(k) = |\text{ШПФ}(x)|,$$

де $k = 0 \dots 4999$.

На виході ШПФ ми отримуємо амплітудний спектр в частотному діапазоні від 0 до 50 Гц, прийнявши що частота дискретизації сигналу становить 100 Гц (саме на неї налаштовується розроблена НМ).

До отриманого амплітудного спектру застосовано логарифмічне перетворення, яке стискає величезний динамічний діапазон (замість 1, 100, $1 \cdot 10^{-6}$ отримуємо 0, 4.6, -11.5). Це критично для НМ, оскільки неадекватне співвідношення між спектральними складовими просто затирає частину спектру з низькими амплітудами. Таким чином приведення до логарифмічної шкали дозволяє побачити слабкі компоненти,

нормалізувати діапазон значень, виділяє форму резонансної кривої та полегшує виявлення параметрів [20]:

$$S(k) = \log(1 + X_f). \quad (9)$$

На наступному кроці спектральне представлення подається у невеличку підмережу, яка складається з двох згорткових прошарків і одного повнозв'язного: Conv1D (32, kernel_size=9, activation function=ELU), Conv1D (64, kernel_size=9, activation function=ELU) та Dense(128, activation function=ELU).

Маючи результати роботи часової та спектральної гілки, необхідно об'єднати їх результати, формуючи єдиний вектор ознак. Далі цей вектор подається у фінальні три повнозв'язні прошарки: Dense(256, activation function=ELU), Dense(128, activation function=ELU), Dense(5, activation function=Linear). Тобто гілки після окремої обробки об'єднуються у спільний латентний простір, після чого подаються на фінальний регресор, який оцінює п'ять параметрів моделі шуму, згадані вище.

Розглянемо частину питань, що стосуються навчання розробленої НМ. Як завжди, воно починається з підготовки навчальних даних. Для симуляції скористаємося формулами (7)-(8) і доповнимо їх білим шумом $n(t)$. Корельований шум отримаємо з дискретних форм для рівнянь ГМП першого та другого порядків. НМ повинна визначити: рівень білого шуму, час кореляції дрейфу, амплітуду корельованого процесу, а також можливо резонансні ефекти. Таким чином НМ має відновлювати параметри випадкових процесів, які ми для неї генеруємо.

У модельному процесі навчання першим кроком треба задати модельні параметри для ГМП другого порядку. На базі цих параметрів розраховуються матриці F та Q_d , які використовуються у моделі простору станів для генерації шумових складових сигналу. Отриманий зашумлений сигнал необхідно зберегти. Процес повторюється декілька тисяч разів, змінюючи модельні параметри шумів та орієнтацію об'єкта у просторі. Також можна додавати різні випадкові коливальні рухи, щоб НМ не прив'язувалася до нульового положення. Описана схема зображена на рис.2. Приклади згенерованих зашумлених сигналів при нерухомому об'єкті та при наявності змінного прискорення показано на рис. 3 та рис. 4. Приклад амплітудного спектру одного із згенерованих сигналів показано на рис. 5.

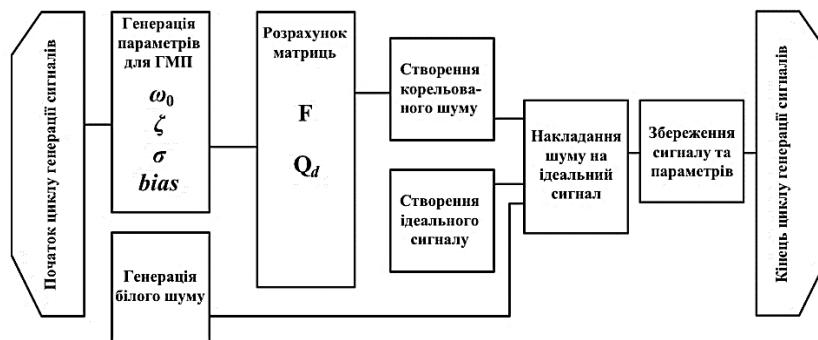


Рис. 2. Схема генерації навчальних даних для нейронної мережі

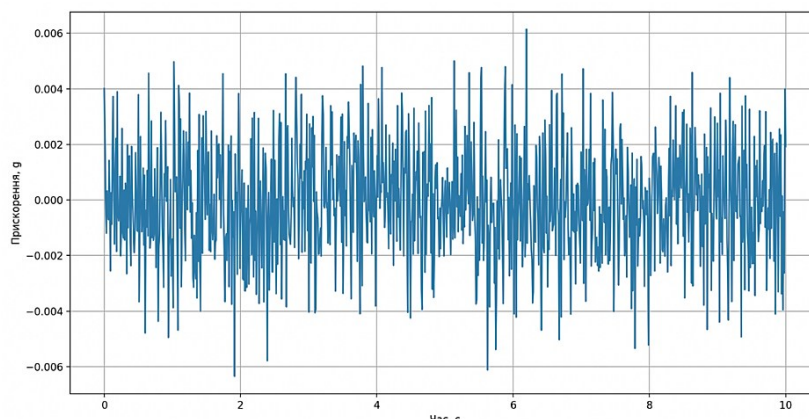


Рис. 3. Приклад згенерованого шуму при відсутності руху об'єкта (нульовий вихідний сигнал)

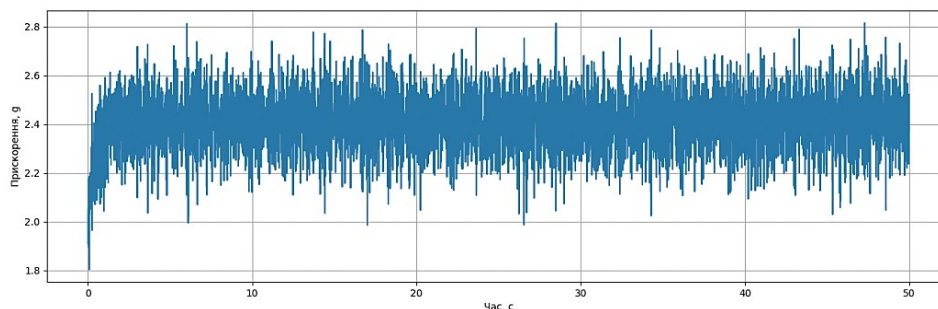


Рис. 4. Приклад згенерованого сигналу зі змінним прискоренням

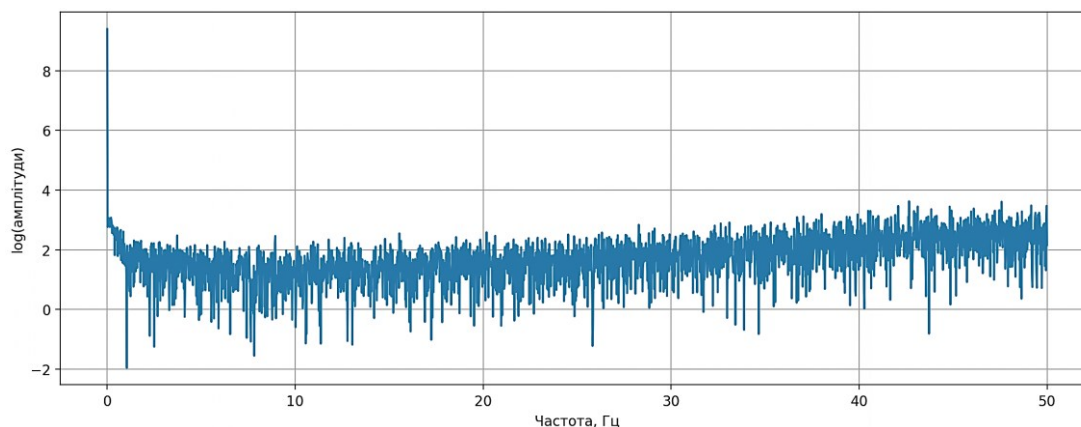


Рис. 5. Приклад амплітудно-частотного спектру згенерованого сигналу

Отже, на вхід до нейронної мережі необхідно подавати сигнал у вигляді масиву значень. Наприклад, це зручно організувати за допомогою csv-файла. Структура використовуваного csv-файла показана в таблиці 1.

Таблиця 1.

Структура csv файлу для навчання НМ

Тип	Сигнал			Параметри сигналу				
Заголовок	t0	...	t4999	omega0	zeta	sigma	bias	R

Значення сигналу у csv-файлі розмічені полями від t_0 до t_{N-1} , де N – розмір згенерованого сигналу. На виході від нейронної мережі очікується отримати параметри сигналу, що у csv-файлі розмічені полями: ω_0 , ζ , σ , $bias$ та R .

Перед подачею до НМ дані проходять робастне нормування, що підвищує стійкість до аномальних шумів та не змінює структуру спектру:

$$\tilde{x}_i = \frac{x_i - \text{медіана}(x)}{\text{MAB}(x) + \varepsilon},$$

де MAB – медіана абсолютного відхилення.

Для підвищення стабільності та ефективності навчання моделі пропонується використовувати два стандартні механізми оптимізації процесу тренування – EarlyStopping та ReduceLROnPlateau.

Механізм EarlyStopping здійснює автоматичну зупинку навчання, якщо значення функції втрат на валідаційній вибірці протягом 10 епох не демонструє покращення. Це дозволяє уникнути перенавчання моделі та скорочує тривалість тренування. Параметр `restore_best_weights = True` забезпечує повернення ваг моделі до стану найкращої епохи.

Другий механізм, ReduceLROnPlateau, контролює швидкість навчання (learning rate) та автоматично зменшує її у 2 рази після 5 епох без покращення валідаційної похибки. Це дозволяє моделі «зробити крок назад» і продовжити оптимізацію в умовах, коли великий коефіцієнт навчання заважає точному налаштуванню параметрів.

Використання цих двох інструментів значно підвищує якість збіжності, зменшує ризик перенавчання та забезпечує отримання моделі з найкращими узагальнюючими властивостями.

Приклад створення описаної вище НМ на мові програмування Python показано в лістингу нижче.

```
import tensorflow as tf
from tensorflow.keras import layers, models
import numpy as np

window_size = 5000 # довжина сигналу
# Вхідний прошарок
inputs = layers.Input(shape=(window_size, 1), dtype=tf.float32)

# Часова гілка
t = layers.Conv1D(32, 9, activation='elu', padding='same')(inputs)
t = layers.MaxPool1D(4)(t)

t = layers.Conv1D(64, 9, activation='elu', padding='same')(t)
t = layers.MaxPool1D(4)(t)

t = layers.Conv1D(128, 9, activation='elu', padding='same')(t)
t = layers.GlobalAveragePooling1D()(t)
time_features = layers.Dense(128, activation='elu')(t)

# Спектральна гілка на основі ШПФ
# 1) Видаляємо останній вимір, приводимо до форми -> shape (batch,
window_size)
squeeze = layers.Lambda(lambda x: tf.squeeze(x, axis=-1))(inputs)

# 2) Розрахунок ШПФ
rfft = layers.Lambda(lambda x: tf.signal.rfft(x), name='rfft')(squeeze)

# 3) амплітудне і логарифмічне масштабування
fft_mag = layers.Lambda(lambda x: tf.math.log1p(tf.abs(x)),
name='fft_mag')(rfft)

# 4) Розширення виміру для Conv1D: (batch, nfreq, 1)
fft_c = layers.Lambda(lambda x: tf.expand_dims(x, axis=-1))(fft_mag)

# 5) Підмережа для спектральної гілки
f = layers.Conv1D(32, 9, activation='elu', padding='same')(fft_c)
f = layers.MaxPool1D(4)(f)
f = layers.Conv1D(64, 9, activation='elu', padding='same')(f)
f = layers.GlobalAveragePooling1D()(f)
freq_features = layers.Dense(128, activation='elu')(f)

# Об'єднання
merged = layers.concatenate([time_features, freq_features])
x = layers.Dense(256, activation='elu')(merged)
x = layers.Dense(128, activation='elu')(x)
outputs = layers.Dense(5, activation='linear')(x)

model = models.Model(inputs=inputs, outputs=outputs)

model.compile(optimizer=tf.keras.optimizers.Adam(1e-4),
loss='mse', metrics=['mae'])
model.summary()
```

Було виконано чисельне моделювання розробленої НМ для оцінювання стохастичних похибок МЕМС-акселерометрів. Похибки оцінювання параметрів випадкових сигналів на валідаційній вибірці приведено у таблиці 2. У процесі моделювання власна частота згенерованих ГМП ω_0 задавалася у діапазоні 10...70 Гц. Оцінювання цього параметру відбувалося з похибкою близько 1%. Коефіцієнт демпфування ζ генерувався у діапазоні 0.01...0.2, а похибка оцінювання становила 20-30 %. Похибка оцінювання просторового шуму σ у рівнянні невелика. Зміщення *bias* у випадковому процесі генерувалося у діапазоні від -0.1 до 0.1, відповідна похибка оцінювання – 28.5%. Шум вимірювання *R* оцінювався адекватно зданій моделі у діапазоні 0.001...0.02.

Таблиця 2.

Перевірка результатів навчання на валідаційній вибірці

Тип метрики	ω_0	ζ	σ	$bias$	R
Середня абсолютна похибка	0.59307843	0.04136802	0.02695927	0.05724134	0.0246821
Середньоквадратична похибка (RMSE)	0.8712104	0.05349036	0.03765524	0.06970523	0.03653709

**ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ
І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ**

Розроблено архітектуру нейронної мережі для оцінки параметрів стохастичної моделі випадкової складової сигналу акселерометра. Розроблена НМ має дві паралельні гілки (часову та спектральну), що дає змогу одночасно враховувати короткочасні кореляційні залежності та частотні компоненти шуму. Навчання моделі здійснено з використанням оптимізаційних процедур ранньої зупинки та адаптивного зменшення швидкості навчання, що забезпечило її стійку збіжність та здатність до узагальнення. Отримана модель продемонструвала здатність ефективно оцінювати параметри стохастичних процесів, що формують шум акселерометра, та може бути використана для подальшої компенсації цих похибок у системах навігації.

Запропонований підхід не вимагає складного лабораторного обладнання і може працювати в умовах високої невизначеності, що робить його придатним для широкого спектра застосувань. Розроблена НМ має широкі перспективи інтеграції в інерціальні вимірювальні модулі безпілотних літальних апаратів, мобільної робототехніки та автономних навігаційних систем.

Запропонований підхід потребує подальших розвідок у практичному застосуванні до сигналів реальних акселерометрів різних виробників, оптимізації структури під визначені зразки мікроконтролерів та інтеграції у вбудовані системи.

References

1. Grewal, M. S., & Andrews, A. P. Kalman Filtering: Theory and Practice with MATLAB. 4th ed. Hoboken: Wiley, 2015. 640 p.
2. Mahony, R., Hamel, T., & Pflimlin, J. "Nonlinear complementary filters on SO(3)." IEEE TAC, 2008.
3. Fedorin, I. V. (2022). *Methods and Technologies of Computational Intelligence: A Textbook*. Kyiv: Igor Sikorsky Kyiv Polytechnic Institute, 314 p.
4. Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366.
5. Goodfellow, I., Bengio, Y., Courville, A. Deep Learning. MIT Press, 2016.
6. PESTI, Richard; SARCEVIC, Peter; ODRY, Akos. Artificial neural network-based MEMS accelerometer array calibration. *International Journal of Intelligent Robotics and Applications*, 2025, 9.4: 1459-1479.
7. LI, Ruihao, et al. Calib-Net: Calibrating the low-cost IMU via deep convolutional neural network. *Frontiers in Robotics and AI*, 2022, 8: 772583.
8. Gao Y., Fourati H., Georgy J. Inertial Motion Tracking: The Role of Deep Learning. *Frontiers in Robotics and AI*, 2021, vol. 8, article 772583. DOI:10.3389/frobt.2021.772583.
9. PAU, Danilo Pietro; TOGNOCCHI, Simone; MARCON, Marco. Learning online MEMS calibration with time-varying and memory-efficient Gaussian neural topologies. *Sensors*, 2025, 25.12: 3679.
10. Kim D., Lee J., Lee S. *MEMS IMU Signal Restoration Using Deep Learning-Based Denoising Architecture*. *Sensors*, 2025, vol. 25, no. 12, article 3679.
11. GELMINI, Simone, et al. Self-calibration algorithm for an imu in two-wheeled vehicles: design and experimental validation. In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2018. p. 1751-1756.
12. Mazzocato F., Marzorati A. Real-Time Self-Calibration of Low-Cost Inertial Measurement Units. Technical Report, Politecnico di Milano, 2020. Режим доступу: <https://re.public.polimi.it/>
13. Podder I., Fischl T., Bub U. Smart calibration and monitoring: leveraging artificial intelligence to improve MEMS-based inertial sensor calibration. *Complex Intell. Syst.*, 2024, 10, 7451–7474. DOI:10.1007/s40747-024-01531-y.
14. ZHEN, Huang; WU, Yuhui. Realization of shock accelerometer sequence-to-sequence calibration based on deep learning. In: *Journal of Physics: Conference Series*. IOP Publishing, 2021. p. 012089.
15. IEEE Std 952-1997. Specification Format and Test Procedure for Linear, Single-Axis, Nongyroscopic Accelerometers.
16. El-Sheimy, N., Hou, H., & Niu, X. Analysis and Modeling of Inertial Sensors Using Allan Variance. *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 1, pp. 140–149, 2008.
17. Jazwinski, A. H. Stochastic Processes and Filtering Theory. Academic Press, 1970.
18. Brown, R. G., & Hwang, P. Y. C. Introduction to Random Signals and Applied Kalman Filtering. Wiley, 2012.
19. Kiranyaz, S., Ince, T., & Gabbouj, M. "1-D convolutional neural networks for signal processing." *IEEE Signal Processing Magazine*, 2021.
20. Lyons, R. G. Understanding Digital Signal Processing. 3rd ed. Upper Saddle River: Prentice Hall, 2011. 992 p.