

Р. М. ГАЛАГАН

кандидат технічних наук, доцент,
доцент кафедри автоматизації та систем неруйнівного контролю
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0000-0001-7470-8392

А. А. САКУТА

магістрантка кафедри автоматизації та систем неруйнівного контролю
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0009-0004-5889-0488

А. С. МОМОТ

доктор філософії, доцент,
старший викладач кафедри автоматизації та систем неруйнівного контролю
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0000-0001-9092-6699

О. В. МУРАВЬОВ

кандидат технічних наук, доцент,
доцент кафедри автоматизації та систем неруйнівного контролю
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0000-0002-7699-0245

О. А. ПОВШЕНКО

доктор філософії,
асистент кафедри автоматизації та систем неруйнівного контролю
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0000-0003-2998-5950

ПЕРСПЕКТИВИ РОЗВИТКУ ЧОТИРИНОГИХ РОБОТІВ З ФУНКЦІЄЮ ГОЛОСОВОГО УПРАВЛІННЯ

Сучасний етап розвитку робототехніки характеризується переходом від детермінованих спеціалізованих до інтелектуальних автономних систем, здатних функціонувати у складних та непередбачуваних середовищах. Особливу увагу привертають чотириногі мобільні роботи, які поєднують високу стійкість, прохідність і маневреність, забезпечуючи ефективність там, де колісні або гусеничні мобільні платформи втрачають працездатність. Разом з тим, розробка таких систем залишається складним завданням, що потребує удосконалення механічної архітектури, оптимізацію енергоспоживання, впровадження нових алгоритмів керування та систем взаємодії з людиною. Метою досліджень є комплексний аналіз сучасного стану розвитку чотириногих роботів і визначення перспектив їх удосконалення. У результаті проведеного аналізу визначено основні тенденції розвитку галузі: удосконалення біонічних конструктивних рішень, впровадження технологій штучного інтелекту (зокрема, глибокого навчання з підкріпленням) у системи керування рухом та розвиток природномовної взаємодії між людиною і роботом. На основі узагальнення даних запропоновано концептуальну архітектуру інтелектуального чотиринового робота з інтегрованим голосовим управлінням, що об'єднує модулі штучного інтелекту, розпізнавання мовлення та комп'ютерного зору. Практичне значення роботи полягає у формуванні узагальненої моделі взаємодії людина–робот, здатної забезпечити інтуїтивне керування мобільними платформами та підвищити ефективність їх використання у складних реальних умовах, в яких робот повинен не тільки орієнтуватись на голосові команди людини, але й враховувати візуальний контекст.

Ключові слова: чотириногий робот, робототехніка, голосове управління, взаємодія людина-робот, штучний інтелект, навчання з підкріпленням, розпізнавання мовлення, великі мовні моделі.

R. M. GALAGAN

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Automation
and Non-Destructive Testing Systems
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
ORCID: 0000-0001-7470-8392

A. A. SAKUTA

Master’s Student at the Department of Automation
and Non-Destructive Testing Systems
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
ORCID: 0009-0004-5889-0488

A. S. MOMOT

Doctor of Philosophy, Associate Professor,
Senior Lecturer at the Department of Automation
and Non-Destructive Testing Systems
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
ORCID: 0000-0001-9092-6699

O. V. MURAVIOV

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Automation
and Non-Destructive Testing Systems
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
ORCID: 0000-0002-7699-0245

O. A. POVSHENKO

Doctor of Philosophy, Assistant at the Department of Automation
and Non-Destructive Testing Systems
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
ORCID: 0000-0003-2998-5950

PROSPECTS FOR THE DEVELOPMENT OF QUADRUPED ROBOTS WITH VOICE CONTROL FUNCTIONALITY

The current stage of robotics development is characterized by a transition from deterministic specialized systems to intelligent autonomous systems capable of functioning in complex and unpredictable environments. Quadruped mobile robots are particularly noteworthy, combining high stability, passability, and maneuverability, ensuring efficiency where wheeled or tracked mobile platforms lose their performance. At the same time, the development of such systems remains a challenging task that requires improvements in mechanical architecture, optimization of energy consumption, and the introduction of new control algorithms and human-robot interaction systems. The aim of the research is to conduct a comprehensive analysis of the current state of development of quadruped robots and to identify prospects for their improvement. The analysis identified the main trends in the development of the industry: improvement of bionic design solutions, introduction of artificial intelligence technologies (in particular, deep learning with reinforcement) into motion control systems, and development of natural language interaction between humans and robots. Based on the generalization of the data, a conceptual architecture of an intelligent quadruped robot with integrated voice control is proposed, combining modules of artificial intelligence, speech recognition, and computer vision. The practical significance of the work lies in the formation of a generalized model of human-robot interaction capable of providing intuitive control of mobile platforms and increasing the efficiency of their use in complex real-world conditions, in which the robot must not only respond to human voice commands, but also take into account the visual context.

Key words: quadruped robot, robotics, voice control, human-robot interaction, artificial intelligence, reinforcement learning, speech recognition, large language models.

Постановка проблеми

Сучасний розвиток робототехніки демонструє значні досягнення у сфері створення мобільних роботів, проте залишається низка невіршених завдань, що стосуються їх ефективності та універсальності [1]. Колісні та гусеничні платформи, попри простоту конструкції, мають суттєві обмеження при пересуванні по пересіченій місцевості, зокрема на крутих схилах, болотистих, кам'янистих чи засніжених ділянках. Це робить їх малоефективними у тих сферах, де ключовим фактором є високий рівень прохідності.

Вирішенням цієї проблеми з конструкторської точки зору є чотириногі роботи (рис. 1), які здатні адаптуватися до складних умов середовища і водночас є простішими за інші багатонігі аналоги [2]. Конструктивні особливості чотириногих роботів розширюють можливості їхнього використання у промислових завданнях, пошуково-рятувальних операціях, інспекційних і сервісних застосуваннях, а також у повсякденному житті.



Рис. 1. Приклади чотириногих роботів: а – ANYmal ANYbotics Quadruped Robot; б – Unitree Go2; в – DEEP Robotics Lite 3; г – Boston Dynamics Spot

Хоча чотириногі роботи характеризуються високою прохідністю та стійкістю, їхнє широке практичне впровадження обмежується низкою технічних і програмних проблем: складністю механічної архітектури, значним енергоспоживанням, високою вартістю компонентів і потребою в адаптивних алгоритмах керування, здатних працювати у реальному часі [3].

У таблиці 1 наведені переваги та недоліки чотириногих роботів порівняно із дво-, шести-, восьминогими та колісними роботами, а також із біологічними аналогами.

Аналіз останніх досліджень і публікацій

Останні роки характеризуються стрімким прогресом у сфері створення чотириногих роботів завдяки розвитку технологій штучного інтелекту (ШІ), сенсорики та високопродуктивних обчислювальних систем [4]. Аналіз наукових публікацій дає змогу виокремити чотири завдання, які необхідно вирішувати під час проектування мобільних чотириногих роботів нового покоління:

Таблиця 1

Переваги та недоліки чотириногих роботів

Характеристика	Переваги	Недоліки
Мобільність та прохідність	Висока прохідність (здатність пересуватися пересіченою місцевістю, долати схили, сходи, кам'янисті та нерівні поверхні)	Адаптивність поступається біологічним аналогам (менша гнучкість, швидкість та динаміка порівняно з тваринами)
Стійкість та функціональність	Оптимальний баланс стійкості та вантажопідйомності порівняно з двоногими роботами	Складні алгоритми управління, чутливі до відмови сенсорів або приводів (може призвести до втрати мобільності)
Конструкція та управління	Простіші за конструкцією та управлінням порівняно з шестиногими чи восьминогими роботами	Складна механічна конструкція з великою кількістю приводів, що підвищує масу та вартість порівняно із колісними, гусеничними та двоногими роботами
Економіка та технічне обслуговування	Прийнятний компроміс між простою мобільністю колісних платформ і менш стійкими двоногими	Ускладнене технічне обслуговування та вищі енергетичні витрати порівняно з колісними аналогами, що обмежує їхню автономність

1) розробка та впровадження нових конструктивних рішень, спрямованих на зниження маси, енергоспоживання та підвищення автономності чотириногих платформ без втрати стабільності;

2) розробка методів керування та планування рухів із урахуванням динаміки тіла та нерівностей поверхні;

3) інтеграція інтелектуальних алгоритмів прийняття рішень, які дозволяють роботам самонавчатися, оптимізувати рух і адаптувати свою поведінку до контексту середовища;

4) забезпечення ефективної взаємодії людина–робот (Human-Robot Interaction, HRI) на інтуїтивному рівні з одночасним вирішенням проблеми точності і надійності голосової комунікації.

Зокрема, у роботі [5] проаналізовано історичну еволюцію та ключові технологічні віхи розвитку чотириногих робототехніки, приділено увагу механічній архітектурі, стратегіям управління рівновагою та навігації. Автори підкреслюють, що ефективність таких роботів значною мірою залежить від оптимального розподілу маси та розташування центрів обертання у кінцівках. Сучасні тенденції спрямовані на зменшення маси конструкції та використання легких матеріалів із високою питомою міцністю, що дозволяє підвищити енергоефективність без втрати стійкості.

На основі детального аналізу варіантів дизайну, стратегій керування та алгоритмів планування руху чотириногих роботів автори статті [6] доводять, що збалансована взаємодія між жорсткістю приводів і пасивною еластичністю з'єднань сприяє стабільності руху на нерівних поверхнях.

Дослідження [7] продемонструвало, що використання нейронних мереж дозволяє роботам навчатися нових типів рухів (наприклад, рисі чи галопу), навіть якщо вони не були явно запрограмовані у навчальних даних. Це свідчить про потенціал самоорганізованих систем у плануванні складних траєкторій. Рухові стратегії роботів-антропоїдів нового покоління, призначених для роботи у складних і змінних ландшафтах, обговорюються у роботі [8], автори якої пропонують алгоритми, що забезпечують компроміс між точністю стабілізації та швидкістю, що є критичним для застосувань у рятувальних операціях та промисловій інспекції.

Аналіз літературних джерел, що наведені вище, а також робіт [9, 10] дозволяє зробити висновок, що чотириногі роботи посідають провідне місце серед мобільних робототехнічних систем завдяки поєднанню високої прохідності, стійкості та відносної конструктивної простоти. Водночас їхнє широке впровадження обмежується високою складністю механічної архітектури, значними енергетичними витратами й потребою у складних алгоритмах управління. Одним із ключових напрямів розвитку є впровадження глибокого навчання з підкріпленням (Deep Reinforcement Learning, DRL), яке дозволяє роботам формувати адаптивні моделі поведінки без ручного налаштування параметрів керування. Інтеграція ШІ в систему прийняття рішень дозволяє поєднати дані з сенсорів, камер і лідарів у єдину когнітивну модель, що визначає оптимальні дії робота залежно від умов середовища.

Паралельно з технічним удосконаленням рухових систем зростає увага до природної взаємодії між людиною і роботом. Інтеграція голосових інтерфейсів розглядається як засіб підвищення інтуїтивності управління, особливо у випадках, коли оператор виконує паралельні завдання. Використання природної мови дозволяє зменшити когнітивне навантаження та підвищити швидкість реакції системи у стресових або динамічних умовах. Водночас залишається низка викликів, пов'язаних із контекстним розпізнаванням команд, шумозахищеністю та багатомовною підтримкою, що визначає подальші напрями розвитку HRI.

Узагальнюючи результати аналізу, можна зазначити, що сучасні дослідження спрямовані на створення інтегрованих систем, які поєднують біонічні принципи конструкції, алгоритми глибокого навчання та природномовну взаємодію. Це формує основу для розроблення чотириногих роботів нового покоління – автономних, енергоефективних і здатних до контекстно-залежної комунікації з людиною.

Формулювання мети дослідження

Метою роботи є комплексний аналіз сучасного стану, тенденцій і перспектив розвитку чотириногих роботів

з акцентом на інтеграцію систем ШІ та голосового управління у процес прийняття рішень і взаємодії людина–робот. Методи дослідження ґрунтуються на аналітичному огляді сучасних публікацій, в яких описані механічні, алгоритмічні та комунікаційні аспекти побудови чотириногих роботів, порівняльному аналізі та систематизації наукових результатів. Особлива увага в роботі приділена обговоренню можливостей голосових технологій, що створюють інтуїтивно зрозумілий канал взаємодії між людиною та роботом у контексті реальних сценаріїв HRI. Також зосереджено увагу на використанні глибокого навчання з підкріпленням для оптимізації руху чотириногих роботів та підвищення їх автономності.

Викладення основного матеріалу дослідження

Принципи реалізації руху чотириногих роботів

Реалізація руху чотириногих роботів базується на біонічних принципах і інженерних рішеннях, що забезпечують баланс між стабільністю, прохідністю та енергоефективністю. Для цього використовуються різні схеми ходи та алгоритми контролю рівноваги – від класичних методів планування рухів до сучасних підходів із використанням ШІ, які дозволяють адаптувати пересування до складних і динамічних умов [3].

Стабільність руху чотириногих роботів у неструктурованих середовищах може бути забезпечена завдяки використанню різних підходів, одним із яких є морфологічна адаптація. У роботі [11] описана реконфігурована конструкція корпусу чотиринового робота, кінцівки якого створені на основі механізму Хокена. Така архітектура дозволяє роботу змінювати положення тіла між вертикальною та низькою позами, виконуючи плавні переходи завдяки вигнутій траєкторії руху кінцівок і плечовим шарнірам, що сприяють активній і пасивній адаптації до рельєфу. Перевагами є підвищена прохідність і стійкість на складному рельєфі за рахунок пасивної/активної підлаштуваності, а недоліками – ускладнення конструкції та керування, а також збільшення маси. Тому перспективним є поєднання морфологічної адаптації з прогнозним керуванням (model predictive control, MPC) для енергоощадних переходів між позами.

Для забезпечення стабільного динамічного пересування чотириногих роботів (зокрема, рисою) може бути застосований метод, що ґрунтується на контролі центроїдального імпульсу [12]. Бажаний рух центру мас генерується на основі моделі лінійного інвертованого маятника. Для обчислення необхідних прискорень шарнірів та сил реакції опори використовується квадратичне програмування, яке мінімізує похибку відстеження бажаної швидкості зміни центроїдального імпульсу, одночасно враховуючи динамічні обмеження та вплив тертя. Висновки свідчать про те, що цей алгоритм є надійним (зокрема, до помилок маси робота) та ефективним для динамічних дій. Перевагами є передбачуваність і формальні гарантії стабілізації. Але варто врахувати і обмеження у вигляді зниження гнучкості у сильно нелінійних або невідомих умовах. Перспективним напрямом розвитку є гібридизація з DRL (наприклад, використання MPC як «захисного» шару).

Для досягнення надійних та динамічних стрибкових рухів чотиринового робота запропоновано використовувати DRL на базі алгоритму Soft-Actor Critic (SAC) [13]. Результатом є розробка єдиного, узагальненого алгоритму зворотного зв'язку, який працює в реальному часі, здатний відстежувати багато різних траєкторій стрибків і успішно переноситься з симуляції в реальність без додаткового налаштування, демонструючи надійні стрибки навіть на нерівній поверхні. Це надає системі високу адаптивність до збурень і невизначеностей, здатність виконувати узагальнення на нові траєкторії. Обмеженням є необхідність у великих обчислювальних витратах на навчання.

У роботі [14] розглядається стратегія руху чотириногих роботів, в основі якої лежать природні переходи, що забезпечує стабільне та енергоефективне пересування на різних місцевостях та швидкостях. Основним методом є створення карти пересування, яка комплексно інтегрує критерії вартості транспортування і стабільності для керування вибором типу руху, подібно до того, як тварини (наприклад, кіт) використовують рись для низької швидкості чи галоп для високої. Реалізація плавного перемикавання між моделями руху досягається за допомогою афінних перетворень для параметрів типів руху та кінцевого автомата, що визначає порядок переходу. Результати експериментів підтверджують, що ця стратегія перевершує базові методи у досягненні одночасно підвищеної ефективності локомоції та її стабільності. Перевагами є енергоощадність і робастність у широкому діапазоні швидкостей. Але існує потреба у ретельному налаштуванні критеріїв переходів. Перспективним варіантом є автоматичний вибір типу ходи на основі політик DRL з урахуванням енергобюджету та ризику ковзання.

Загалом, сучасні підходи до реалізації руху чотириногих роботів поєднують класичні методи динамічного керування з алгоритмами глибокого навчання з підкріпленням, що забезпечує як стабільність, так і адаптивність у складних умовах. Результати досліджень демонструють успішне застосування таких рішень для пересування рисою, подолання сходів і виконання стрибків навіть на нерівних поверхнях. На сьогоднішній день оптимальною є гібридна архітектура, де класичні моделі гарантують безпеку й базову стабілізацію, а ML/DRL-політики відповідають за високорівневе планування, вибір типу ходи та адаптацію до середовища

Інтеграція ШІ в систему прийняття рішень чотириногих роботів

Розвиток сучасної робототехніки критично залежить від інтеграції ШІ, що значно підвищує адаптивність, точність керування та ефективність прийняття рішень роботами у складних і неструктурованих середовищах

[15]. У фокусі – перехід від «реактивних» контролерів до багаторівневих когнітивних систем, які поєднують сприйняття, прогноз і планування.

Беззаперечно, інтеграція технологій ШІ у систему прийняття рішень чотириногих роботів є важливим напрямом їх розвитку. Використання алгоритмів машинного навчання (machine learning, ML) дозволяє підвищити адаптивність до змінних умов, оптимізувати рух по складному рельєфу та покращити інтерпретацію голосових команд у контексті конкретної ситуації. У роботі [16] оцінено поточний прогрес, ключові успіхи та відкриті виклики DRL у реальних роботизованих застосуваннях. DRL демонструє найкращі результати у задачах з високою невизначеністю, але потребує механізмів безпеки та валідації політик.

Огляд [17] зосереджується на застосуванні глибокого навчання з підкріпленням для створення високоефективних контролерів чотириногих роботів, приділяючи особливу увагу систематизації алгоритмів, подоланню розриву між симуляцією та реальними умовами і визначенню перспективних напрямів подальших досліджень. Таким чином, ключовими трендами є ієрархічні політики, рандомізація середовищ, комбінування з традиційними системами в якості проміжного прошарку.

У роботі [18] представлена ієрархічна система навчання з підкріпленням для чотириногих роботів типу ANYmal C, завдяки якій робот може виконувати завдання, пов'язані із маніпулюванням об'єктами у захищених та обмежених середовищах зі стінами. Методика використовує дворівневу систему керування, яка навчається за допомогою алгоритму Proximal Policy Optimization (PPO). Контролер високого рівня визначає бажані лінійну та кутову швидкості тіла на основі загальної мети та покращеної функції винагороди, що включає штрафи за наближення до стін і використовує ключові точки об'єкта для оцінки помилки позиціонування. Контролер низького рівня перетворює ці команди швидкості на команди позиції суглобів. Симуляційні випробування були проведені у фреймворку NVIDIA Isaac Sim і продемонстрували високу точність та енергоефективність. Таким чином, система «швидкості → суглоби» з контекстною винагородою підвищує керованість у тісних просторах.

У дослідженні [19] проаналізовано поєднання конструктивних можливостей чотириногих роботів із методами ШІ та голосового управління. Застосування глибокого навчання з підкріпленням у поєднанні з пропріоцептивним і тактильним зворотним зв'язком дозволило створити «сліпі» моделі руху, здатні підтримувати стабільність у темряві чи запиленому середовищі. Також досліджено інтеграцію хмарних сервісів розпізнавання мовлення (Amazon Alexa, Google Dialogflow, Microsoft LUIS) для інтуїтивної взаємодії з роботом у реальному часі. Експериментальні результати засвідчили високу ефективність DRL для подолання сходів та перешкод, тоді як голосове управління виявило залежність від шумового середовища. Автори визначають перспективним напрямом поєднання стійких алгоритмів руху з удосконаленими мовними моделями для застосувань у рятувальних та інспекційних завданнях. Це підтверджує, що мультимодальність (пропріоцепція + тактильні/аудіо/візуальні канали) є ключем до робастності.

Таким чином, результати останніх досліджень послідовно демонструють трансформаційний вплив ШІ та машинного навчання на робототехніку, підкреслюючи, що для майбутнього розгортання роботів у складних умовах (включаючи маніпуляції та комунікацію) необхідно продовжувати вдосконалювати багаторівневі моделі на базі DRL та інтегрувати і розвивати мовні моделі для HRI.

Принципи реалізації голосового управління чотириногими роботами

Попри значний прогрес у розвитку мобільних робототехнічних систем, зокрема чотириногих роботів, завдання забезпечення ефективної та інтуїтивно зрозумілої взаємодії між людиною і машиною залишається відкритим. Традиційні методи управління (зокрема через пульти або програмні інтерфейси) часто вимагають спеціальної підготовки, що уповільнює реакцію та знижує ефективність застосування у критичних умовах. Одним із перспективних шляхів вирішення цієї проблеми є впровадження голосового управління, яке забезпечує природний спосіб комунікації з роботом [20]. Водночас впровадження голосових технологій супроводжується серйозними викликами: необхідністю забезпечення високої точності розпізнавання мовлення в умовах шуму та розробкою алгоритмів контекстної інтерпретації команд. Окрім того, реалізація HRI через голосове управління потребує додатково вирішення деяких етичних (зокрема, довіра та впевненість у сприйнятті команд роботом) та практичних підходів до взаємодії людини з роботом [21]. Складність полягає у тому, що одна й та сама команда може мати різне значення залежно від ситуації (наприклад, контексту завдання, положення робота чи стану середовища), що потребує семантичного розуміння, а не лише акустичного розпізнавання.

Роль голосових інтерфейсів у сучасній робототехніці постійно зростає, адже роботи дедалі частіше інтегруються в «людиноцентричні» середовища, утворюючи галузь соціальної робототехніки [22]. Такі системи мають сприймати мову не лише як набір команд, а як інструмент комунікації, що враховує інтонацію, контекст, наміри та навіть емоційні ознаки. Перспективним напрямом є використання відкритих моделей типу Whisper (OpenAI) та wav2vec 2.0 (Meta), які демонструють високу універсальність, можливість донавчання під конкретні мови та стабільну роботу в умовах фонових шумів. Whisper, зокрема, показує середню достовірність понад 95% для коротких команд при належному попередньому очищенні сигналу, що робить його ефективним варіантом для локальних систем HRI [23].

Під час впровадження голосового управління важливим викликом є підвищення точності розпізнавання команд. Наприклад, деякі дослідження показують, що для мов зі складною морфологією, таких як українська, точність розпізнавання залишається нижчою, ніж для англійської [24]. Це означає, що для команд різною мовою ключовим фактором для досягнення необхідної точності є правильно обрана мовна модель.

У роботі [25] розглядається ефективність голосового керування колаборативним роботом DOBOT Magician. Автори використали Arduino-комплект і перевірили 22 оригінальні голосові команди, які були перепрограмовані для виконання рухів робота. Було проведено 2200 симуляцій (100 для кожної команди), результати яких аналізувалися з використанням програми Audacity. Жодна з 22 команд не досягла 100% точності. Результати показали, що команди, які складаються з декількох слів точніше розпізнаються, ніж короткі (однослівні) чи надто довгі. Це свідчить про необхідність оптимізації синтаксису команд і врахування когнітивного навантаження оператора.

Варто зауважити, що правильне розпізнавання команд залежить не тільки від обраної мовної моделі, але й від якості голосових даних, які передаються моделі для розпізнавання. Тому вибір мікрофону, зменшення шумів електронного тракту, вибір способів придушення шумів акустичного тракту також відіграють важливу роль. Зокрема, для попередньої обробки мовного сигналу використовуються мел-частотні кепстральні коефіцієнти (MFCC), що забезпечують виділення стабільних ознак незалежно від акустичних умов [26].

Особливістю розвитку голосової взаємодії між людиною та роботом є те, що розробники не створюють нових мовних моделей та не імплементують їх в апаратне забезпечення робота. Натомість використовуються різноманітні хмарні служби. Наприклад, в роботі [27] для синтезу мовлення використовується сервіс Amazon Polly, а для розпізнавання – Google Cloud Speech-to-Text API. Також авторами розглядається потреба у використанні під час спілкування із роботом команд-активаторів, розпізнавши які робот розумітиме, що далі він отримає вказівки від оператора. Адже оточуюче середовище може генерувати багато шумів, слів сторонніх людей, які не потрібно виконувати, що значно ускладнює процес інтерпретації потрібних вказівок, а не просто будь-чого.

Таким чином, сучасні системи HRI часто не потребують створення власних мовних моделей, оскільки активно використовують готові хмарні рішення, найпоширеніші з яких наведені у таблиці 2.

Таблиця 2

Порівняння хмарних сервісів для розпізнавання та синтезу мовлення

Сервіс	Розробник	Переваги	Обмеження
Google Cloud Speech-to-Text	Google	Висока точність для 120+ мов, підтримка потокового режиму	Погіршення результатів при нестабільному інтернеті
Amazon Polly / Lex	Amazon	Суміщення синтезу та діалогових сценаріїв	Менша точність для неанглійських мов
Microsoft Azure Speech	Microsoft	Добра інтеграція з системами керування	Не підтримує навчання користувачьких моделей
Whisper	OpenAI	Висока універсальність, можливість локальної роботи	Високе споживання обчислювальних ресурсів при великих моделях

Наступним етапом розвитку HRI є перехід від командного управління до когнітивної взаємодії, у якій робот не просто виконує інструкцію, а інтерпретує мету користувача. Це можливо завдяки використанню великих мовних моделей (Large Language Models, LLM), які дозволяють перетворювати природномовні вказівки у послідовність оптимальних дій [28]. Такі системи дають змогу оператору спілкуватися з роботом на семантичному рівні без попереднього програмування команд.

Також останні дослідження вказують на важливість у взаємодії людина-робот враховувати не тільки голосове управління, але й візуальний контекст подій та навколишню обстановку. Наприклад, після отримання команди робот може додатково завдяки обробці відеопотоку аналізувати просторову інформацію, виконувати певні дії та генерувати текстовий зворотний зв'язок оператору, забезпечуючи таким чином більш чітке уявлення про поточний стан і наміри робота, що є необхідним для ефективної та безпечної роботи [29]. Поєднання LLM та систем машинного зору (зокрема, тривимірного) дозволяє роботам краще усвідомлювати середовище, інтерпретувати мовні команди в контексті просторових моделей і підвищити автономність [30]. Завдяки візуальному сприйняттю робот паралельно з інтерпретацією голосових команд також може аналізувати жести та позу людини, що надають додаткову інформацію про наміри останньої [31]. Таким чином підвищується обізнаність робота про наміри людини, внаслідок чого покращується взаємодія людина-робот.

Отже, розвиток взаємодії людина-робот на основі голосових команд є одним із важливих напрямів сучасної робототехніки, що безпосередньо залежить від досягнень у сфері великих мовних моделей. Сучасні підходи до голосового управління чотириногими роботами демонструють перехід від окремих систем розпізнавання мовлення до інтегрованих мультимодальних HRI-комплексів, у яких мовна модель, зорове сприйняття та алгоритми ШІ утворюють єдине середовище прийняття рішень.

Для українського контексту найбільш перспективним є використання відкритих моделей типу Whisper з подальшим донавчанням на локальних наборах команд, що дозволить створити адаптивну систему взаємодії людина-робот без залежності від комерційних API.

Перспективним напрямком є реалізація гібридної архітектури «LLM (розуміння наміру) → планувальник (MPC/граф планів) → локальний контролер (DRL)», із чіткими інтерфейсами даних між рівнями та метриками безпеки. У поєднанні з системами комп'ютерного зору це дозволить роботу розпізнавати просторові відносини, аналізувати жести, позу, напрямок погляду користувача та формувати відповідну поведінку.

Концептуальна архітектура чотириноного інтелектуального робота з інтегрованим голосовим управлінням

На основі проведеного аналізу пропонується узагальнена концептуальна архітектура чотириноного інтелектуального робота (рис. 2), яка визначає основні функціональні рівні системи, взаємозв'язки між ними та напрямки подальшого розвитку. Архітектура поєднує модулі сприйняття, інтелектуального аналізу, планування руху та природномовної взаємодії, що забезпечує автономність, адаптивність і когнітивну поведінку робота.

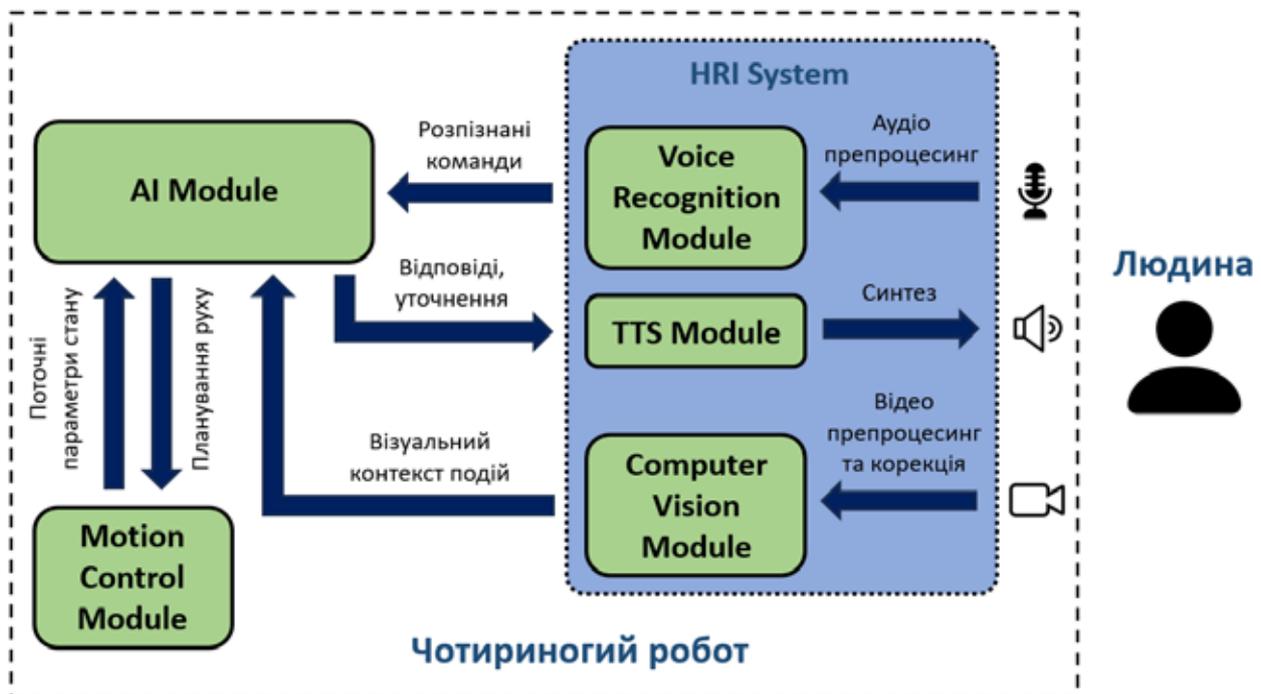


Рис. 2. Концептуальна архітектура чотириноного інтелектуального робота

Центральним компонентом архітектури є HRI System, яка забезпечує двосторонню комунікацію між користувачем і роботом. Система HRI складається із трьох модулів нижнього рівня і призначена для організації взаємодії між людиною та роботом на високому рівні абстракції. Система має забезпечувати не просто розпізнавання роботом команд оператора, але й розуміння запитань, складних вказівок, які визначають набір дій, що необхідно виконати, правильну інтерпретацію намірів оператора. Для цього необхідно впроваджувати адаптивний цикл через LLM, завдяки якому робот навчається уточнювати команди користувача, адаптуючись до стилю мовлення, наголосу чи емоційного стану. Усі модулі HRI System функціонують у єдиному комунікаційному середовищі, що спрощує обмін даними та синхронізацію між компонентами.

Модуль розпізнавання голосу (Voice Recognition Module) відповідає за захоплення, фільтрацію, попередню обробку та перетворення мовного сигналу в текстову форму. Реалізація може базуватись на моделях Whisper (OpenAI), wav2vec 2.0 (Meta), SpeechT5 (Microsoft), Kaldi з подальшим донавчанням під специфіку української мови. Для забезпечення надійності сигналу передбачено попередню обробку з використанням MFCC, шумопригнічення та компенсацію ревербераційних спотворень.

Під час проектування модуля розпізнавання голосу потрібно приділити особливу увагу вибору якісного мікрофону, який забезпечував би високу якість акустичного сигналу (голосу користувача), та впровадженню сучасних методів попередньої обробки сигналів з метою виділення корисних даних з мінімальними спотвореннями.

Повноцінна робота системи HRI ґрунтується на двосторонній комунікації. Це означає, що чотириногий робот має формувати для оператора голосові або текстові відповіді, невербальні сигнали (світлові, рухові тощо). Тому

в систему HRI включено модуль TTS (Text-To-Speech). Зокрема, завдяки цьому модулю робот може уточнити контекст команди (наприклад, уточнити наміри оператора), щоб в подальшому правильно спланувати дії. Модуль забезпечує голосовий зворотний зв'язок від робота до користувача. Використання систем типу Amazon Polly, Google Text-to-Speech або відкритих моделей SpeechT5 / Bark дає змогу роботу підтверджувати дії чи надавати короткі звіти про виконання.

Третім модулем нижнього рівня, який входить до системи HRI, є модуль комп'ютерного зору (Computer Vision Module). Цей модуль реалізує сприйняття просторової сцени за допомогою камер, лідарів та сенсорів глибини. На цьому етапі застосовується обробка зображень, детектування об'єктів (YOLO, Detectron тощо), побудова карти середовища та ідентифікація положення оператора. Отримана візуальна інформація інтегрується із голосовими даними для контекстної інтерпретації команд.

Зібрані від сенсорів та трансформовані модулями нижнього рівня системи HRI дані у заданому форматі передаються у модуль штучного інтелекту (AI Module). Цей модуль виконує роль «когнітивного ядра» системи, у якому відбувається інтерпретація отриманих команд, планування дій і генерація поведінки робота, зокрема:

- інтерпретація текстових команд у контексті поточного стану робота та середовища;
- формування семантичних представлень команд і трансформація їх у цілі;
- планування дій, прийняття рішень, оптимізація траєкторій і поведінки з урахуванням кінематичних, енергетичних та інших обмежень;
- аналіз поточних параметрів стану робота (позиція, швидкість, кутові положення, силові характеристики, контакт із поверхнею тощо), що отримані від модуля керування рухом;
- самонавчання на основі DRL (алгоритми SAC, PPO тощо) та корекція поведінки за результатами взаємодії;
- використання мультимодальних LLM (наприклад, LLaMA 2, Gemma, Mistral) для інтерпретації запитів природною мовою та генерації логічних відповідей.

Таким чином, AI Module реалізує адаптивний цикл «слухати → розуміти → планувати → виконувати → аналізувати», який забезпечує поступове покращення поведінки робота та точності реакції на мовні команди.

На основі аналізу усього масиву даних AI Module формує план руху для модуля керування рухом (Motion Control Module). Він є фізичним рівнем виконання команд. Модуль керування рухом приймає оптимізовані траєкторії або координати, що сформовані модулем ШІ, і перетворює їх у сигнали для виконавчих механізмів.

Основні завдання модуля керування рухом:

- контроль кінематики ніг, планування точок опори та стабілізація центру маси;
- обробка даних інерційних сенсорів (гіроскопи, акселерометри), енкодерів, лідарів;
- реалізація механізмів зворотного зв'язку для адаптації до нерівностей та ковзання;
- реалізація команд від AI Module на рівні виконавчих механізмів (наприклад, сервоприводів з PID) для низькорівневого стабілізування.

Запропонована архітектура інтегрує мовні, зорові та когнітивні підсистеми у єдину платформу. Її особливістю є використання гібридної схеми: нижній рівень забезпечує фізичну стабільність і керування рухом, тоді як верхній рівень реалізує інтелектуальне прийняття рішень на основі ШІ та LLM. Такий підхід формує основу для створення мобільних роботів нового покоління, здатних до природної комунікації, самонавчання та адаптації до змін середовища.

Висновки

Чотириногі роботи демонструють високий потенціал для застосування в рятувальних, промислових та сервісних завданнях, де важливими є прохідність, стійкість і швидке реагування на команди. У результаті проведеного аналізу узагальнено сучасні тенденції розвитку чотириногих роботів та встановлено, що ключовим напрямом еволюції є інтеграція біонічних конструктивних принципів, алгоритмів ШІ та природномовної взаємодії. Такий підхід дозволяє сформулювати новий клас мобільних платформ, здатних до самонавчання, контекстного прийняття рішень і природної комунікації з користувачем.

Запропоновано концептуальну архітектуру інтелектуального чотириноного робота, яка поєднує модулі розпізнавання мовлення, комп'ютерного зору, ШІ та керування рухом. На відміну від існуючих рішень, архітектура забезпечує двосторонню комунікацію людина–робот, здатність до адаптації на основі глибокого навчання з підкріпленням і семантичну інтерпретацію команд за допомогою великих мовних моделей.

Сформована узагальнена модель інтелектуальних HRI-систем, яка є придатною для реалізації в роботах промислового, сервісного та рятувального призначення. Використання відкритих мовних і візуальних моделей дозволяє адаптувати архітектуру для локальних застосувань без потреби у хмарних обчисленнях, що особливо важливо для мобільних і автономних систем.

Серед невирішених завдань залишаються підвищення точності розпізнавання мовлення для української мови, створення стійких до шуму локальних моделей автоматичного розпізнавання мовлення, а також забезпечення безпечного та етичного прийняття рішень роботами у сценаріях спільної роботи з людиною. Необхідною є також подальша розробка методів оцінки довіри до ШІ та запобігання помилкам контекстної інтерпретації команд.

Перспективи подальших досліджень полягають у вдосконаленні інтегрованих мультимодальних систем HRI, які поєднують LLM, DRL і комп'ютерний зір у єдину когнітивну платформу. Синтез цих підходів створить передумови для появи автономних роботів нового покоління, здатних не лише ефективно пересуватись у складному середовищі, а й сприймати, аналізувати та прогнозувати поведінку людини у реальному часі.

Список використаної літератури

1. Raj, R., & Kos, A. (2022). A comprehensive study of mobile robot: History, developments, applications, and future research perspectives. *Applied Sciences*, 12(14), 6951. <https://doi.org/10.3390/app12146951>
2. Biswal, P., & Mohanty, P.K. (2021). Development of quadruped walking robots: A review. *Ain Shams Engineering Journal*, 12(2), 2017–2031. <https://doi.org/10.1016/j.asej.2020.11.005>
3. Fan, Y., Pei, Z., Wang, C., Li, M., Tang, Z., & Liu, Q. (2024). A review of quadruped robots: Structure, control, and autonomous motion. *Advanced Intelligent Systems*, 6(6), 2300783. <https://doi.org/10.1002/aisy.202300783>
4. Kawaharazuka, K., Inoue, S., Suzuki, T., Yuzaki, S., Sawaguchi, S., Okada, K., & Inaba, M. (2024). MEVIUS: A quadruped robot easily constructed through e-commerce with sheet metal welding and machining. *Proceedings of the 2024 IEEE-RAS 23rd International Conference on Humanoid Robots (Humanoids)*, 631–636. <https://doi.org/10.1109/Humanoids58906.2024.10769853>
5. Li, Q., Cicirelli, F., Vinci, A., Guerrieri, A., Qi, W., & Fortino, G. (2025). Quadruped robots: Bridging mechanical design, control, and applications. *Robotics*, 14(5), 57. <https://doi.org/10.3390/robotics14050057>
6. Azeez, S. A., Mandava, R. K., & Naik, N. S. (2025). A novel review on quadruped robots design variants, gait modulation, and motion planning schemes. *Journal of Field Robotics*, 42(12), 3615–3693. <https://doi.org/10.1002/rob.22575>
7. Yamamoto, H., Kim, S., Ishii, Y., & Ikemoto, Y. (2020). Generalization of movements in quadruped robot locomotion by learning specialized motion data. *Robomech Journal*, 7(29). <https://doi.org/10.1186/s40648-020-00174-1>
8. Saha Kotha, S., Akter, N., Abhi, S. H., Das, S. K., Islam, M. R., Ali, M. F., ... & Hasan, M. M. (2024). Next generation legged robot locomotion: A review on control techniques. *Heliyon*, 10(18), e37237. <https://doi.org/10.1016/j.heliyon.2024.e37237>
9. Taheri, H., & Mozayani, N. (2023). A study on quadruped mobile robots. *Mechanism and Machine Theory*, 190, 105448. <https://doi.org/10.1016/j.mechmachtheory.2023.105448>
10. Chai, H., Li, Y., Song, R., Zhang, G., Zhang, Q., Liu, S., ... & Yang, Z. (2022). A survey of the development of quadruped robots: Joint configuration, dynamic locomotion control method and mobile manipulation approach. *Biomimetic Intelligence and Robotics*, 2(1), 100029. <https://doi.org/10.1016/j.birob.2022.100029>
11. Yuan, J., Wang, S., Wang, B., Shi, R., Wu, X., Li, L., Li, W., Wang, Z., & Dai, Z. (2024). Design of a quadruped robot with morphological adaptation through reconfigurable sprawling structure and method. *Advanced Intelligent Systems*, 6(6), 2300645. <https://doi.org/10.1002/aisy.202300645>
12. Liu, M., Qu, D., Xu, F., Zou, F., Di, P., & Tang, C. (2019). Quadrupedal robots whole-body motion control based on centroidal momentum dynamics. *Applied Sciences*, 9(7), 1335. <https://doi.org/10.3390/app9071335>
13. Bellegarda, G., Nguyen, C., & Nguyen, Q. (2024). Robust quadruped jumping via deep reinforcement learning. *Robotics and Autonomous Systems*, 182, 104799. <https://doi.org/10.1016/j.robot.2024.104799>
14. Zhang, D., Chen, X., Zhong, Z., Xu, M., Zheng, Z., & Lu, H. (2025). Novel multi-gait strategy for stable and efficient quadruped robot locomotion. *Journal of System Simulation*, 37(1), 2. <https://doi.org/10.48550/arXiv.2410.09336>
15. Hussain, M. D., Rahman, M. H., & Ali, N. M. (2024). Artificial intelligence and machine learning enhance robot decision-making adaptability and learning capabilities across various domains. *International Journal of Science and Engineering*, 1(3), 14–27. <https://doi.org/10.62304/ijse.v1i3.161>
16. Tang, C., Abbatematteo, B., Hu, J., Chandra, R., Martín-Martín, R., & Stone, P. (2024). Deep reinforcement learning for robotics: A survey of real-world successes. *arXiv Preprint arXiv:2408.03539*. <https://doi.org/10.48550/arXiv.2408.03539>
17. Zhang, H., He, L., & Wang, D. (2022). Deep reinforcement learning for real-world quadrupedal locomotion: A comprehensive review. *Intelligent Robotics*, 2(3), 275–297. <https://doi.org/10.20517/ir.2022.20>
18. Azimi, D., & Hoseinnezhad, R. (2025). Hierarchical reinforcement learning for quadrupedal robots: Efficient object manipulation in constrained environments. *Sensors*, 25(5), 1565. <https://doi.org/10.3390/s25051565>
19. Ranasinghe, U., Islam, R., Anavatti, S., & Garrat, M. (2025) A review of reinforcement learning techniques for quadruped robot control and locomotion in complex terrains. *SSRN Preprint*. <https://dx.doi.org/10.2139/ssrn.5183855>
20. Norda, M., Engel, C., Rennies, J., Appell, J.-E., Lange, S. C., & Hahn, A. (2024). Evaluating the efficiency of voice control as human-machine interface in production. *IEEE Transactions on Automation Science and Engineering*, 21(3), 4817–4828. <https://doi.org/10.1109/TASE.2023.3302951>
21. Han, J., & Conti, D. (2025). Recent advances in human-robot interactions. *Applied Sciences*, 15(12), 6850. <https://doi.org/10.3390/app15126850>

22. Badr, A. A., & Abdul-Hassan, A. K. (2020). A review on voice-based interface for human–robot interaction. *Iraqi Journal for Electrical and Electronic Engineering*, 16(2), 1–12. <https://doi.org/10.37917/ijeec.16.2.10>
23. Radford, A., Kim, J.W., Xu, T., Brockman, G., Mcleavey, C. & Sutskever, I. (2023). Robust Speech Recognition via Large-Scale Weak Supervision. *Proceedings of the 40th International Conference on Machine Learning*, PMLR 202:28492–28518. <https://proceedings.mlr.press/v202/radford23a.html>
24. Бронніков, А.І., & Онишко, В.О. (2017). Обробка інформації при голосовому керуванні у робототехніці. *Системи обробки інформації*, 3(149), 85–87.
25. Janiček, M., Ružarovský, R., Velíšek, K., & Holubek, R. (2021). Analysis of voice control of a collaborative robot. *Journal of Physics: Conference Series*, 1781, 012025. <https://doi.org/10.1088/1742-6596/1781/1/012025>
26. Kishor, I., Mamodiya, U., Saini, S., & Bossoufi, B. (2025). Voice-enabled human–robot interaction: Adaptive self-learning systems for enhanced collaboration. *Robotica*, 43(6), 2143–2171. <https://doi.org/10.1017/S0263574725000438>
27. Mehrizi, K. (2021). Quadrupedal robotic guide dog with vocal human–robot interaction. *arXiv Preprint arXiv:2111.03718*. <https://doi.org/10.48550/arXiv.2111.03718>
28. Ismail, S., Arbues, A., Cotterell, R., Zurbrügg, R., & Alonso, C. A. (2024). NARRATE: Versatile language architecture for optimal control in robotics. *Proceedings of the 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 9628–9635. <https://doi.org/10.1109/IROS58592.2024.10801425>
29. Abbas, A. N., & Beleznai, C. (2024). TalkWithMachines: Enhancing human–robot interaction through large/vision language models. *Proceedings of the 2024 Eighth IEEE International Conference on Robotic Computing (IRC)*, 253–258. <https://doi.org/10.1109/IRC63610.2024.00039>
30. Mehta, V., Sharma, C., & Thiyagarajan, K. (2025). Large language models and 3D vision for intelligent robotic perception and autonomy. *Sensors*, 25(20), 6394. <https://doi.org/10.3390/s25206394>
31. Lai, Y., Yuan, S., Nassar, Y., Fan, M., Weber, T., & Rättsch, M. (2025). NVP-HRI: Zero shot natural voice and posture-based human–robot interaction via large language model. *Expert Systems with Applications*. 268, 126360. <https://doi.org/10.1016/j.eswa.2024.126360>

References

1. Raj, R., & Kos, A. (2022). A comprehensive study of mobile robot: History, developments, applications, and future research perspectives. *Applied Sciences*, 12(14), 6951. <https://doi.org/10.3390/app12146951>
2. Biswal, P., & Mohanty, P.K. (2021). Development of quadruped walking robots: A review. *Ain Shams Engineering Journal*, 12(2), 2017–2031. <https://doi.org/10.1016/j.asej.2020.11.005>
3. Fan, Y., Pei, Z., Wang, C., Li, M., Tang, Z., & Liu, Q. (2024). A review of quadruped robots: Structure, control, and autonomous motion. *Advanced Intelligent Systems*, 6(6), 2300783. <https://doi.org/10.1002/aisy.202300783>
4. Kawaharazuka, K., Inoue, S., Suzuki, T., Yuzaki, S., Sawaguchi, S., Okada, K., & Inaba, M. (2024). MEVIUS: A quadruped robot easily constructed through e-commerce with sheet metal welding and machining. *Proceedings of the 2024 IEEE-RAS 23rd International Conference on Humanoid Robots (Humanoids)*, 631–636. <https://doi.org/10.1109/Humanoids58906.2024.10769853>
5. Li, Q., Cicirelli, F., Vinci, A., Guerrieri, A., Qi, W., & Fortino, G. (2025). Quadruped robots: Bridging mechanical design, control, and applications. *Robotics*, 14(5), 57. <https://doi.org/10.3390/robotics14050057>
6. Azeez, S. A., Mandava, R. K., & Naik, N. S. (2025). A novel review on quadruped robots design variants, gait modulation, and motion planning schemes. *Journal of Field Robotics*, 42(12), 3615–3693. <https://doi.org/10.1002/rob.22575>
7. Yamamoto, H., Kim, S., Ishii, Y., & Ikemoto, Y. (2020). Generalization of movements in quadruped robot locomotion by learning specialized motion data. *Robomech Journal*, 7(29). <https://doi.org/10.1186/s40648-020-00174-1>
8. Saha Kotha, S., Akter, N., Abhi, S. H., Das, S. K., Islam, M. R., Ali, M. F., ... & Hasan, M. M. (2024). Next generation legged robot locomotion: A review on control techniques. *Heliyon*, 10(18), e37237. <https://doi.org/10.1016/j.heliyon.2024.e37237>
9. Taheri, H., & Mozayani, N. (2023). A study on quadruped mobile robots. *Mechanism and Machine Theory*, 190, 105448. <https://doi.org/10.1016/j.mechmachtheory.2023.105448>
10. Chai, H., Li, Y., Song, R., Zhang, G., Zhang, Q., Liu, S., ... & Yang, Z. (2022). A survey of the development of quadruped robots: Joint configuration, dynamic locomotion control method and mobile manipulation approach. *Biomimetic Intelligence and Robotics*, 2(1), 100029. <https://doi.org/10.1016/j.birob.2022.100029>
11. Yuan, J., Wang, S., Wang, B., Shi, R., Wu, X., Li, L., Li, W., Wang, Z., & Dai, Z. (2024). Design of a quadruped robot with morphological adaptation through reconfigurable sprawling structure and method. *Advanced Intelligent Systems*, 6(6), 2300645. <https://doi.org/10.1002/aisy.202300645>
12. Liu, M., Qu, D., Xu, F., Zou, F., Di, P., & Tang, C. (2019). Quadrupedal robots whole-body motion control based on centroidal momentum dynamics. *Applied Sciences*, 9(7), 1335. <https://doi.org/10.3390/app9071335>
13. Bellegarda, G., Nguyen, C., & Nguyen, Q. (2024). Robust quadruped jumping via deep reinforcement learning. *Robotics and Autonomous Systems*, 182, 104799. <https://doi.org/10.1016/j.robot.2024.104799>

14. Zhang, D., Chen, X., Zhong, Z., Xu, M., Zheng, Z., & Lu, H. (2025). Novel multi-gait strategy for stable and efficient quadruped robot locomotion. *Journal of System Simulation*, 37(1), 2. <https://doi.org/10.48550/arXiv.2410.09336>
15. Hussain, M. D., Rahman, M. H., & Ali, N. M. (2024). Artificial intelligence and machine learning enhance robot decision-making adaptability and learning capabilities across various domains. *International Journal of Science and Engineering*, 1(3), 14–27. <https://doi.org/10.62304/ijse.v1i3.161>
16. Tang, C., Abbatematteo, B., Hu, J., Chandra, R., Martín-Martín, R., & Stone, P. (2024). Deep reinforcement learning for robotics: A survey of real-world successes. *arXiv Preprint arXiv:2408.03539*. <https://doi.org/10.48550/arXiv.2408.03539>
17. Zhang, H., He, L., & Wang, D. (2022). Deep reinforcement learning for real-world quadrupedal locomotion: A comprehensive review. *Intelligent Robotics*, 2(3), 275–297. <https://doi.org/10.20517/ir.2022.20>
18. Azimi, D., & Hoseinnezhad, R. (2025). Hierarchical reinforcement learning for quadrupedal robots: Efficient object manipulation in constrained environments. *Sensors*, 25(5), 1565. <https://doi.org/10.3390/s25051565>
19. Ranasinghe, U., Islam, R., Anavatti, S., & Garrat, M. (2025) A review of reinforcement learning techniques for quadruped robot control and locomotion in complex terrains. *SSRN Preprint*. <https://dx.doi.org/10.2139/ssrn.5183855>
20. Norda, M., Engel, C., Rennies, J., Appell, J.-E., Lange, S. C., & Hahn, A. (2024). Evaluating the efficiency of voice control as human–machine interface in production. *IEEE Transactions on Automation Science and Engineering*, 21(3), 4817–4828. <https://doi.org/10.1109/TASE.2023.3302951>
21. Han, J., & Conti, D. (2025). Recent advances in human–robot interactions. *Applied Sciences*, 15(12), 6850. <https://doi.org/10.3390/app15126850>
22. Badr, A. A., & Abdul-Hassan, A. K. (2020). A review on voice-based interface for human–robot interaction. *Iraqi Journal for Electrical and Electronic Engineering*, 16(2), 1–12. <https://doi.org/10.37917/ijeec.16.2.10>
23. Radford, A., Kim, J.W., Xu, T., Brockman, G., Mcleavey, C. & Sutskever, I. (2023). Robust Speech Recognition via Large-Scale Weak Supervision. *Proceedings of the 40th International Conference on Machine Learning*, PMLR 202:28492–28518. <https://proceedings.mlr.press/v202/radford23a.html>
24. Bronnikov, A., & Onyshko, V. (2017). Information processing for speech control in robotics. *Information Processing Systems*, 3(149), 85–87.
25. Janiček, M., Ružarovský, R., Velíšek, K., & Holubek, R. (2021). Analysis of voice control of a collaborative robot. *Journal of Physics: Conference Series*, 1781, 012025. <https://doi.org/10.1088/1742-6596/1781/1/012025>
26. Kishor, I., Mamodiya, U., Saini, S., & Bossoufi, B. (2025). Voice-enabled human–robot interaction: Adaptive self-learning systems for enhanced collaboration. *Robotica*, 43(6), 2143–2171. <https://doi.org/10.1017/S0263574725000438>
27. Mehrizi, K. (2021). Quadrupedal robotic guide dog with vocal human–robot interaction. *arXiv Preprint arXiv:2111.03718*. <https://doi.org/10.48550/arXiv.2111.03718>
28. Ismail, S., Arbues, A., Cotterell, R., Zurbrugg, R., & Alonso, C. A. (2024). NARRATE: Versatile language architecture for optimal control in robotics. *Proceedings of the 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 9628–9635. <https://doi.org/10.1109/IROS58592.2024.10801425>
29. Abbas, A. N., & Beleznaï, C. (2024). TalkWithMachines: Enhancing human–robot interaction through large/vision language models. *Proceedings of the 2024 Eighth IEEE International Conference on Robotic Computing (IRC)*, 253–258. <https://doi.org/10.1109/IRC63610.2024.00039>
30. Mehta, V., Sharma, C., & Thiyagarajan, K. (2025). Large language models and 3D vision for intelligent robotic perception and autonomy. *Sensors*, 25(20), 6394. <https://doi.org/10.3390/s25206394>
31. Lai, Y., Yuan, S., Nassar, Y., Fan, M., Weber, T., & Rättsch, M. (2025). NVP-HRI: Zero shot natural voice and posture-based human–robot interaction via large language model. *Expert Systems with Applications*. 268, 126360. <https://doi.org/10.1016/j.eswa.2024.126360>

Дата першого надходження рукопису до видання: 28.11.2025

Дата прийнятого до друку рукопису після рецензування: 26.12.2025

Дата публікації: 31.12.2025