УДК 004.896

МАШИННЕ НАВЧАННЯ ПРИ КЕРУВАННІ БУДІВЕЛЬНИМИ РОБОТАМИ

Мірошник В.А.

Мобільна робототехніка активно розвивається, знаходячи застосування в численних сферах: від перевезень до сільського господарства, від медицини до рятувальних робіт та застосування у цілях оборони. Використання таких роботів значно підвищує продуктивність, якість і безпеку виконуваних завдань.

Перспективною галуззю застосування мобільних роботів є будівництво. Будівельні мобільні роботи (БМР) можуть виконують земляні роботи, перевозити матеріали, здійснювати моніторинг процесу будівництва з метою перевірки їх якості та строків виконання робіт тощо.

Розробка та використання БМР потребує особливої уваги до планування їхнього руху [1], адже будівельні майданчики характеризуються складним, неструктурованим середовищем. У таких умовах ефективно працюють системи керування на основі штучного інтелекту.

У даній роботі для побудови БДР шляху пропонується використовувати машинне навчання з підкріпленням. Такий підхід до машинного навчання не вимагає попередньо зібраних даних [2], адже навчання відбувається через взаємодію з середовищем за допомогою датчиків і системи винагород за прийняті рішення. Якість навчання оцінюється за допомогою Q-функції, яка показує, наскільки цінною є дія з точки зору максимізації майбутньої винагороди:

$$Q(s,a) = r + \gamma \max_{a'} Q(s',a'), \tag{1}$$

де s – поточний стан робота БМР; a – дія БМР; s' – попередній стан БМР; a' – попередня дія БМР; γ – коефіцієнт швидкості навчання, що варіюється в діапазоні від 0 до 1.

Навчання БМР здійснювалося у Simulink із використанням додатку до MATLAB Reinforcement Learning Toolbox. Метою навчання було знайти оптимальний маршрут на будівельному майданчику, уникаючи зіткнень із перешкодами. Для моделювання використано чотириколісний БМР з бортовим поворотом, модель якого побудована у Simulink. Для навчання обрано алгоритм Deep Deterministic Policy Gradient (DDPG), що поєднує роботу двох нейронних мереж – актора і критика. Мережа актора генерує найкращу дію, тоді як критик оцінює її, обчислюючи значення Q-функції. Така взаємодія дозволяє постійно вдосконалювати політику актора. Для знаходження нових шляхів оптимізації DDPG іноді виконує випадкові дії.

Винагорода для робота розраховувалася на основі його поточних координат стану, виконаної дії, досягнення цільової точки та наявності зіткнення з перешкодою за формулою :

$$R = k_1 e^{\alpha x^2} + k_2 e^{\beta y} + k_3 \sum_{i=1}^{m} u_i^2 + k_5 g + k_6 c + k_7, \qquad (2)$$

де k_1, k_2, …, k_6 – вагові коефіцієнти, що встановлюють рівень винагороди (при доданому значенні) або штрафу (при від’ємному значенні) за відповідну дію або стан; k_7 – від'ємний коефіцієнт, що діє як базовий штраф за кожен крок та стимулює швидке досягнення мети; α, β – коефіцієнти, що регулюють чутливість винагороди до змін координат x та y; u_i – i-й керуючий впив; m – кількість керуючих впливів; g – бінарна функція, що свідчить про досягнення або недосягнення БМР мети; c – бінарна функція, що свідчить про наявність або відсутність зіткнення робота з перешкодою.

Контролер формує бажану силу тяги, що забезпечує лінійний рух БМР, змінюючи v, та момент сили обертання БМР.

Навчання здійснювалося протягом 500 епізодів, кожен із яких починався з випадкового положення. Епізод завершуався, якщо БМР досягав мети, стикався з перешкодою або перевищував задану кількість ітерацій.

У процесі навчання робот отримував винагороди за наближення до мети та штрафи за зіткнення з перешкодами. На рис. 1 зображено приклад шляху,

пройденого БМР.

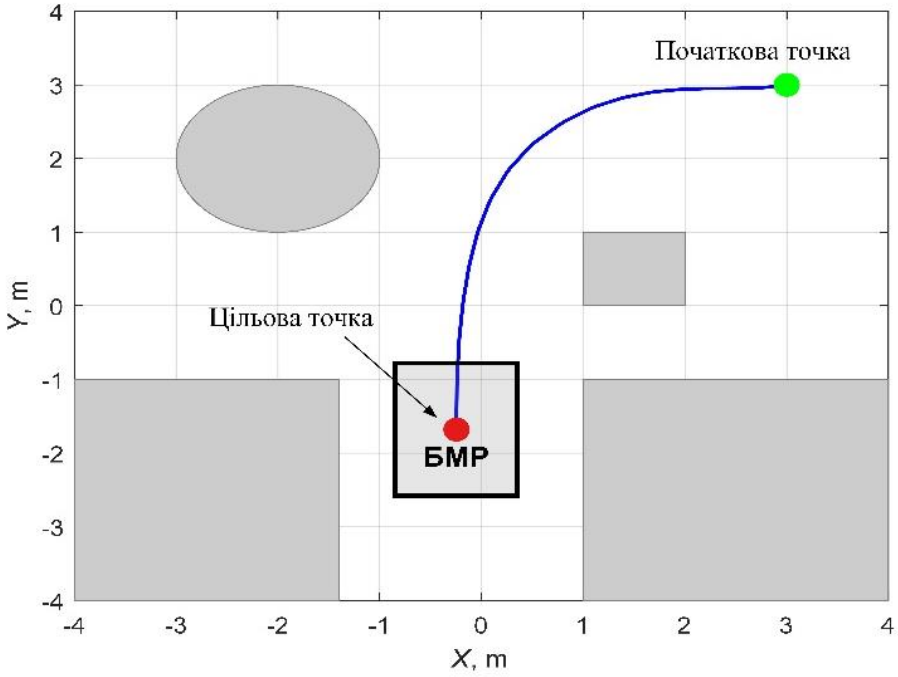

Рис. 1. Побудований БМР шлях

Як видно з рис. 1, робот успішно досяг цільової точки, уникаючи зіткнень із перешкодами.

Література:

1. Гурко О. Г., Гурко В. О., Кучеренко А. Ю. Керування рухом фронтального навантажувача за заданою траєкторією. *Вісник ХНАДУ*, 2023, вип. 101, т. 1. С. 26–34. doi: 10.30977/BUL.2219-5548.2023.101.0.26

2. Бочкарьов О. Навчання з підкріпленням в автономних інтелектуальних системах : навч. посіб. Львів : Вид. Марченко Т.В, 2024. 125 с.