

Викладено метод аналізу теплообмінних процесів в акумуляторі та вказано шляхи їх вирішення. Наведені основи оптимізації досліджуваних процесів.

Математичне моделювання, крайові умови, перетворення Ханкеля і Лапласа, критерій Прандтля, функціональний оператор.

The article a method for the analysis of heat transfer processes in the accumulator and the indicated solutions. The foundations of the optimization of the processes studied.

Mathematical modeling, boundary conditions, Hankel transform and Laplace, Prandle criterion , functional operator.

УДК 007.52

ОРІЄНТАЦІЯ МОБІЛЬНОГО РОБОТА В ПРОСТОРІ ТЕПЛИЦІ ЗА ДОПОМОГОЮ ЙМОВІРНІСНИХ АВТОМАТІВ ТА СТИМУЛЮЮЧОГО НАВЧАННЯ

***В.П. Лисенко, І.М. Болбот, кандидати технічних наук
І.І. Чернов, аспірант****

Розглянуто метод вирішення задачі орієнтування мобільного робота за допомогою ймовірнісних автоматів. Опрацьовано теоретичну складову ймовірнісних автоматів теплиці як поля для переміщення робота. На основі теоретичних досліджень створено програмне забезпечення, що дозволяє розраховувати шлях пересування робота в просторі теплиці із урахуванням імовірних перешкод.

Імовірність, орієнтування мобільного робота, тепличне господарство, стимулююче навчання.

Дотепер часу успішно вирішені завдання глобальної навігації, але її використання стає проблематичним, коли потрібно, щоб робот орієнтувався всередині приміщень. Взагалі робота всередині закритих приміщень характеризується безліччю найрізноманітніших перешкод, починаючи від нерівномірності освітлення і закінчуючи проблемами відображення радіосигналів. У такому разі середовище вважається погано обумовленим, з ненадійними каналами зв'язку, з принциповою неточністю і невизначеністю. Тоді очевидною стає необхідність дослідження способів орієнтування, здатних працювати в цьому середовищі.

Мета досліджень – створення алгоритму орієнтування мобільного робота в просторі блочної теплиці, що має деяку стохастичну природу.

Науковий керівник – кандидат технічних наук, професор В.П. Лисенко

© В.П. Лисенко, І.М. Болбот, І.І. Чернов, 2014

Матеріали та методика досліджень. За результатами аналізу сучасних систем орієнтації у просторі роботів та засобів машинного навчання порівнювались різні алгоритми пошуку оптимального шляху із урахуванням слабкої формалізації умов при постановці задачі.

Результати досліджень. Питання управління роботами розглядається в рамках так званої інтелектуальної робототехніки. Саме поняття є нині не цілком визначеним. Більшість визначень інтелектуального робота зводяться до перерахування інтелектуальних властивостей, якими він повинен володіти («зір», «слух», «планування» і т.д., аж до вказівки конкретних механізмів штучного інтелекту). Менш формальним визначенням є здатність системи вирішувати завдання, сформульовані в загальному вигляді [2].

Оскільки теплиця є зоною роботи не тільки робота, але й людей, що обслуговують її та доглядають за рослинами, роботу недостатньо пересуватись визначеним маршрутом, що задає оператор. Він повинен шляхом накопичення досвіду вчитись оминати перешкоди, що стають на його шляху та брати до уваги ймовірність тієї чи іншої події, та відповідно вибудовувати свій маршрут згідно з цим прогнозам.

Для побудови керуючих імовірнісних автоматів використовуються методи, які оцінюють реакцію автомата і певним чином перерозподіляють ймовірності, пов'язані з цією реакцією. У загальному випадку це зводиться до перевірки певної умови, на основі якого робиться висновок про те, була реакція хорошою чи поганою. При цьому результуючий автомат буде більшою мірою слідувати визначеній стратегії, оскільки алгоритм навчання не враховує, що «хороша» реакція на цьому кроці може призвести до поганого результату на наступних кроках. Пропонується підхід, заснований на використанні алгоритмів стимулюючого навчання.

Нехай заданий мобільний робот, який взаємодіє з навколишнім середовищем, робить якісь дії. Середовище реагує на дії робота і певним чином рангує, або карає його. Для того, щоб робот зміг вибрати правильну тактику поведінки, йому слід стежити за рангами (покараннями) від навколишнього середовища. Такий вид навчання називається стимулюючим (рис. 1). Поведінка робота щодо означеної задачі характеризується такими параметрами: S – множина станів, в яких може перебувати мобільний робот; A – множина доступних робота дій; t – дискретний крок за часом; на кожному кроці робот розпізнає поточний стан $s_t \in S$ і вибирає дію $a_t \in A$; після цього навколишнє середовище повертає мобільному роботу ранг $r_t = (s_t, a_t)$ і переносить його в новий стан $s_{t+1} = \delta(s_t, a_t)$.

Для розглянутого випадку функції δ і r є частиною навколишнього середовища і роботу невідомі. Також розглядаються лише кінцеві множини S і A .

Завдання робота полягає в тому, щоб знайти оптимальну стратегію вибору наступної дії щодо поточного стану $\pi: S \rightarrow A$. Для оцінки оптимальності знайдених стратегій використовуються різні моделі, але на практиці перевага віддається моделі нескінченного горизонту. Відповідно

до цієї моделі оптимальною називається стратегія, що максимізує ранг робота відповідно до такої формули [3]:

$$V^\pi(s_t) \equiv r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots \equiv \sum_{i=0}^{\infty} \gamma^i r_{t+i}, \quad (1)$$

де $0 \leq \gamma \leq 1$ – деяка константа, яка дозволяє більшою мірою враховувати нагороду, отриману від навколишнього середовища тут і зараз, ніж у найближчому майбутньому. Це справедливо для більшості реальних завдань, оскільки, чим раніше робот отримає нагороду тим краще. Нагорода, яку він теоретично повинен отримати через двадцять кроків, може і не дістатися роботу, якщо у нього в запасі їх залишилося всього десять.



Рис.1. Цикл навчання

Оптимальну стратегію поведінки робота можна виразити таким співвідношенням [3]:

$$\pi^* \equiv \arg_{\pi} \max V^\pi(s), \quad (\forall s). \quad (2)$$

Для зручності позначимо $V^*(s)$ функцію прибутку, відповідну оптимальній стратегії. Тоді попередню тотожність (2) можна переписати у вигляді [3]:

$$\pi^*(s) \equiv \arg_{\pi} \max [r(s, a) + \gamma V^*(\vartheta(s, a))]. \quad (3)$$

Головне завдання робота – знайти оптимальну стратегію π^* . Однак безпосередньо обчислити функцію $\pi: S \rightarrow A$ досить важко, оскільки вся інформація, доступна роботу, – це послідовність заохочень від навколишнього середовища $r(s_i, a_i)$, де $i = 0, 1, 2, \dots$

Цю проблему можна вирішити введенням так званої Q-функції [3]:

$$Q(s, a) \equiv r(s, a) + \gamma V^*(\vartheta(s, a)). \quad (4)$$

Вираз (5) для кожної пари стан – дія (s, a) повертає максимальний ранг, яку отримує робот, якщо, перебуваючи в стані s , вибере дію a . Цю тотожність для оптимальної стратегії можна переписати у вигляді [3]:

$$\pi^*(s) \equiv \arg_{\pi} \max Q(s, a). \quad (5)$$

Ця формула показує, що для знаходження оптимальної стратегії роботу замість функції V^* достатньо обчислити функцію Q , що дозволяє

вибирати оптимальне дію для поточного стану без будь-яких знань про функції δ і r .

Таким чином, для одержання оптимальної стратегії потрібно знайти значення функції Q , яка обчислюється безпосередньо через взаємодію робота з навколишнім середовищем (через послідовність $r(s_i, a_i)$, де $i = 0, 1, 2, \dots$).

Внутрішньою моделлю для цього методу є Q -функція, яка показує потенційний ранг робота для всіх пар стан–дія. На основі цієї моделі для кожного стану робота логічно зіставити ймовірності переходів вихідного автомата кожному з можливих дій робота пропорційно очікуваного прибутку для цієї дії. Таким чином, перебуваючи в стані s_i , робот повинен вибрати дію a_j ($0 \leq j \leq n_j$) з імовірністю [4]:

$$p_j = \frac{R(s_i, a_j)}{\sum_{k=0}^{n_i} R(s_i, a_k)}, \quad (6)$$

$$R(s_i, a_k) = \begin{cases} Q(s_i, a_k) & \text{при } Q(s_i, a_k) \geq 0 \\ 0 & \text{при } Q(s_i, a_k) < 0 \end{cases}, \quad (7)$$

де n_i – число доступних робота дій у стані s_i (при цьому a_0 означає, що робот не робить ніяких дій і залишається в поточному стані).

Вираз (7) задає нульові ймовірності діям з від'ємним значенням Q -функції. Таким чином, виключається можливість того, що робот здійснить заздалегідь не вигідну для нього дію. Прикладом заздалегідь не вигідної дії може бути спроба мобільного робота вчинити дію «йти вперед», коли перед ним знаходиться стіна. У цьому випадку робот залишається на місці і отримує негативне число при «неправильній» дії. При традиційному підході, коли робот слідує оптимальній стратегії π^* , така дія виключається в силу того, що потенційний ранг для нього буде в будь-якому випадку менше, ніж для будь-якої осмисленої дії.

Для реалізації алгоритму необхідно задати деякі умови для його успішного виконання:

- мобільний робот рухається в просторі теплиці та повинен пройти під час свого руху певні контрольні точки, що задаються оператором перед запуском робота;
- оскільки робот під час руху між рядами рослин знаходиться на напрямних, кількість його дій у цей момент обмежена рухом вперед чи назад;
- заїзд на напрямні здійснюється за допомогою кольорової розмітки, що нанесено на підлогу теплиці (рис.2);
- за наявності перешкоди на маршруті робот визначає її ультразвуковим датчиком; якщо є можливість її обминути, то здійснюється маневрування; за неможливості дістатися до певної цілі робот подає звуковий сигнал та відправляє повідомлення оператору і переходить до наступної цілі;

- для спрощення орієнтації робота умовно простір теплиці розбито на сектори; зміну сектора робот відстежує за допомогою кольорових міток на горщиках рослин.

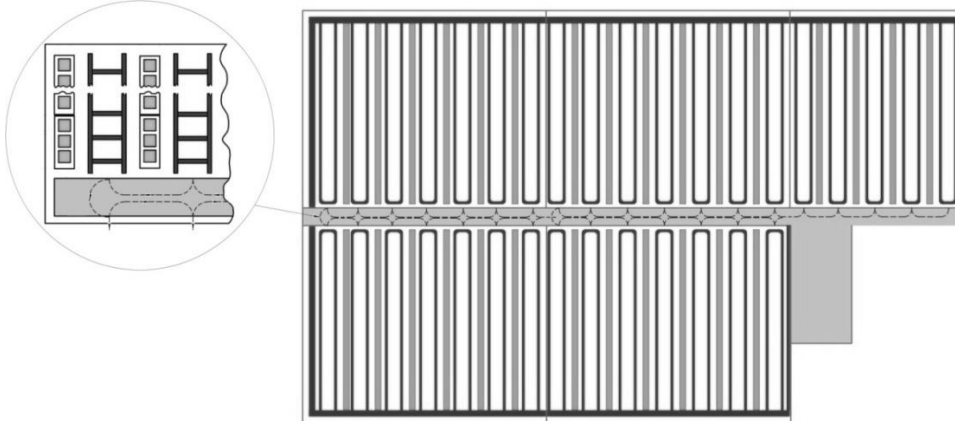


Рис.2. Візуальний план теплиці та розмітка для спрощення орієнтування робота

Для демонстрації роботи алгоритму розраховується рух робота по лівому блоку теплиці (рис.3). На рисунку темнішим відтінком показано клітинки, на які робот не може заїхати, світлішим – клітинки цілі, які визначені оператором.

Клітинки-стани нумеруються рядками, починаючи від 1 для верхньої лівої клітинки і закінчуючи 400 для нижньої правої. Дії робота теж нумеруються, починаючи з 1: 1 – рух вгору, 2 – рух вправо, 3 – рух вниз, 4 – рух вліво.

Виконуючи навчання мобільного робота згідно з описаними алгоритмами отримуємо автоматну модель поставленої задачі, що являє собою стохастичну матрицю переходів. Рядки цієї матриці відповідають клітинкам-станам мобільного робота, стовпці – можливим діям. Таким чином, стохастична матриця переходів зіставляє з кожною парою стан-дія значення імовірності цієї дії. Значення відповідних ймовірностей стохастичної матриці переходів розраховуються за виразом (6).

2	4	6	8	10	12	14	16	18	
22	24	26	28	30	32	34	36	38	
42	44	46	48	50	52	54	56	58	
62	64	66	68	70	72	74	76	78	
82	84	86	88	90	92	94	96	98	
102	104	106	108	110	112	114	116	118	
122	124	126	128	130	132	134	136	138	
142	144	146	148	150	152	154	156	158	
162	163	164	165	166	167	168	169	170	171
182	183	184	185	186	187	188	189	190	191
202	203	204	205	206	207	208	209	210	211
222	223	224	225	226	227	228	229	230	231
242	244	246	248	250	252	254	256	258	
262	264	266	268	270	272	274	276	278	
282	284	286	288	290	292	294	296	298	
302	304	306	308	310	312	314	316	318	
322	324	326	328	330	332	334	336	338	
342	344	346	348	350	352	354	356	358	
362	364	366	368	370	372	374	376	378	
382	384	386	388	390	392	394	396	398	

Рис. 3. Поле переміщень робота по лівому блоку теплиці

Нульові ймовірності вибору дії в таблиці (рис. 3) показують спробу робота вийти за межі поля, що призводить до негативної оцінки сумарного прибутку. Із таблиці виключені клітинки темного кольору, тому що робот не може на них знаходитися.

Для візуальної демонстрації на рис.4 показано інтерфейс користувача. Розмір поля задається користувачем. Також користувач визначає: цілі – лівий клік мишею (зелений колір) та перешкоди – правий клік (червоний колір).

Значення кількості ітерації визначає кількість циклів навчання, параметр рівня навчання позначає те, як добре робот вивчає успішні маршрути (не рекомендується виставляти цей параметр занадто високим). Значення штрафів та нагород визначають, яку вигоду робот отримує при знаходженні цілі або при спробі виходу з поля, та встановлюються експериментально.

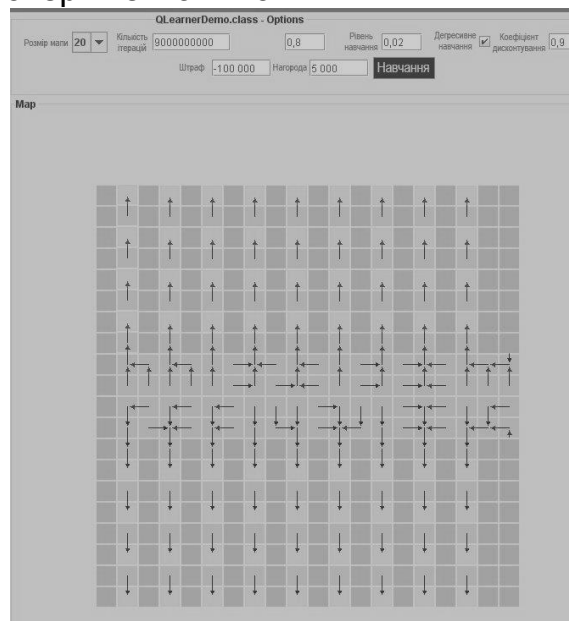


Рис.4. Інтерфейс програмної реалізації алгоритму

Для успішного проходження маршруту алгоритму потрібно пройти період навчання на накопичення досвіду спроб та помилок. Із накопиченням досвіду кількість невдалих спроб і час проходження маршруту зменшується, а кількість отриманих рангів навпаки зростає (рис.5).

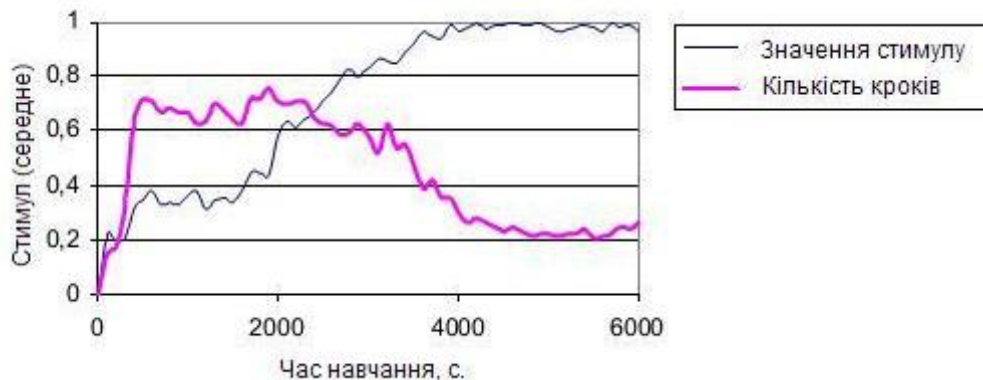


Рис.5. Графік залежності часу навчання та отриманих рангів

Висновки

Враховуючи стохастичну природу можливих перешкод у теплиці, доцільно для орієнтування мобільного робота в просторі блочної теплиці, використовувати ймовірнісні алгоритми машинного навчання. Після періоду навчання ефективність роботи алгоритму стимулюючого навчання значно збільшується і дозволяє мобільному роботу самостійно пересуватися в просторі блочної теплиці, оминаючи перешкоди.

Список літератури

1. Добрынин Д. А. Моделирование некоторых форм адаптивного поведения интеллектуальных роботов / Д.А. Добрынин, В.Э. Карпов // Информационные технологии и вычислительные системы. – 2006. – №2. – С.45–54.
2. Карпов В.Э. К вопросу об управлении мобильным роботом в условиях общей постановки задачи / В.Э. Карпов // Вестник компьютерных и информационных технологий.– 2008. – № 1. – С. 2–9.
3. Ghnemat R. Agent based Modeling for Spatial self organization /Complex Systems – Computer Sciences Implementation Applications to Engineering. 2007.
4. Mitchel T. Machine Learning. McGraw Hill Education (ISE Editions). 1997.

Рассмотрен метод решения задачи по ориентированию мобильного робота на основе использования вероятностных автоматов. Проработана теоретическая составляющая вероятностных автоматов теплицы как поля для перемещения робота. На основе теоретических исследований создано программное обеспечение, позволяющее рассчитывать путь передвижения робота в пространстве теплицы с учетом возможных препятствий.

Вероятность, ориентирование мобильного робота, тепличное хозяйство, стимулирующее обучение.

Considered the method of solving the problem orienteering mobile robot based on the use of probabilistic automata. Elaborated theoretical component of probabilistic automata, greenhouses as fields to move the robot. On the basis of theoretical studies created software that allows you to calculate the path of movement of the robot in space greenhouses considering possible obstacles.

Probability, orientation of mobile robot, greenhouses, stimulating learning.