

Roman ZAJDEL*

STEROWANIE ROBOTEM MOBILNYM KHEPERA Z WYKORZYSTANIEM LOGIKI ROZMYTEJ I UCZENIA ZE WZMOCNIENIEM

Opisano miniaturowy robot mobilny Khepera. Zaprezentowano układ nawigacyjny realizujący zadanie poszukiwania celu. Rolę regulatora adaptacyjnego pełni regulator rozmyty Takagi - Sugeno. Uczenie reguł odbywa się za pośrednictwem prostej sieci neuronowej. Zaprezentowano wyniki sterowania robotem mobilnym Khepera.

1. WSTĘP

Teoria zbiorów rozmytych po raz pierwszy została wykorzystana w 1975 roku przez E. Mamdaniego i S. Assiliana do sterowania modelem silnika parowego. Jednakże prawdziwą „eksplozję” zastosowań logiki rozmytej zanotowano w połowie lat 80 - tych, kiedy to M. Togai i H. Watanabe zaprezentowali układ rozmyty VLSI, wykorzystujący wnioskowanie rozmyte. Od tego czasu logikę rozmytą stosuje się w wielu dziedzinach przemysłu, jak i w sprzęcie gospodarstwa domowego. Można ją odnaleźć w pralkach dostosowujących swój cykl prania do stopnia zabrudzenia ładunku, odkurzaczach, klimatyzatorach czy też w kamerach VHS. Innym ciekawym projektem jest sterowanie kolejką w Sendai (Japonia), gdzie regulator rozmyty z powodzeniem zastąpił doświadczonego maszynistę [10, 11]. Należy zauważyć, że chociaż logika rozmyta rozwija się już od 32 lat, to w większości aplikacji liczba reguł jest bardzo mała. Wynika to z trudności w doborze reguł, które najczęściej podaje „ekspert” posiadający pewną wiedzę o obiekcie sterowania [6]. Stąd pojawiają się projekty łączenia logiki rozmytej z sieciami neuronowymi. Zastosowanie sieci neuronowych pozwala na znaczne uproszczenie procesu strojenia regulatora rozmytego. W pracach [1, 2, 7, 8] zaprezentowano zastosowanie sieci neuronowych uczonych ze wzmacnieniem do sterowania robotami mobilnymi. W pracy [1] użyto sieci neuronowej uczonej ze wzmacnieniem do strojenia regulatora rozmytego Mamdaniego. Sieć składała się z dwóch warstw elementów neuronowych, zaproponowanych przez Sutton'a i Barto: ASE (*associative search element*) - służących do wyznaczenia wyjścia regulatora i ACE (*associative critic element*), których zadaniem była predykcja wewnętrznego sygnału wzmacnionego używanego do uczenia elementów ASE [3]. W niniejszym artykule zaprezentowano strojenie reguł regulatora rozmytego Takagi-Sugeno za pomocą sieci złożonej tylko z elementów ACE, których specyficzną (różną od Hebba) metodę uczenia opisano w rozdziale 5.1. Odmiernym problemem jest wytworzenie sygnału wzmacniającego, generowanego przez układ krytyka. Jego zadaniem jest ocena podjętej przez układ uczący akcji. Jeśli podjęte działanie daje wynik pozytywny, w sensie pożądanego zachowania systemu, to powinno nastąpić wzmacnienie tej tendencji, zaś w przeciwnym przypadku - jeśli wynik jest negatywny, to powinno nastąpić

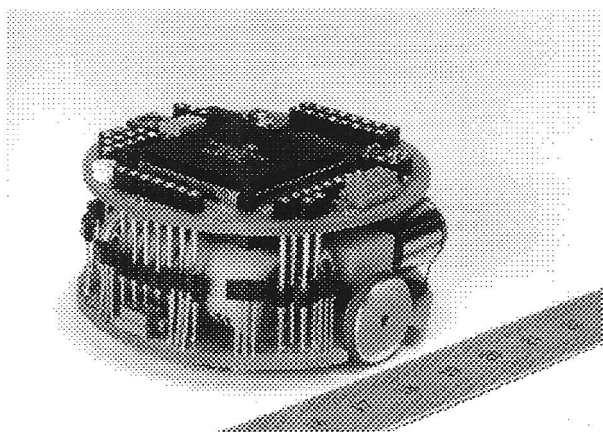
* Politechnika Rzeszowska, Katedra Automatyki i Informatyki, 35-959 Rzeszów, ul. W. Pola 2

osłabienie tej tendencji działania systemu. Lektura prac [1, 2, 3, 7, 8] może prowadzić do wniosku, że nie jest to zadanie łatwe i z reguły zależy ono od sporej wiedzy projektanta systemu sterującego o pożądanym zachowaniu kontrolowanego procesu. Przykładem niech będzie tutaj praca [8], w której sygnał wzmacniony został określony funkcją aż czterech zmiennych. W niniejszym opracowaniu etap budowy układu krytyka (rozdział 5.2), sprowadzono do sformułowania dwóch reguł.

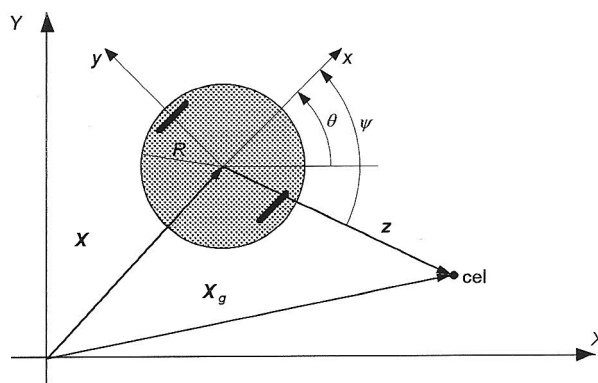
Obiektem sterowania był miniaturowy robot mobilny opisany w rozdziale 2. W rozdziale 3 przedstawiono układ sterowania realizujący zadanie poszukiwania celu, zaś rozdział 4 opisuje zastosowany regulator rozmyty. W rozdziale 6 zamieszczono wyniki przeprowadzonych eksperymentów.

2. ROBOT MOBILNY KHEPERA

Miniaturowy robot mobilny Khepera został zaprojektowany i wykonany w Laboratorium Mikroinformatyki Politechniki w Lozannie (Szwajcaria) [4, 5]. Dzięki niewielkim wymiarom (średnica 52mm, wysokość 30mm) i małej wadze (około 70g) jest dobrym obiektem do testowania algorytmów sterowania (rys. 1). Elementarna konfiguracja robota posiada dwa niezależne koła napędzane przez silniki DC, które pozwalają na uzyskanie maksymalnej prędkości 1m/s. Cztery akumulatory Ni-Cd pozwalają na 20 minutową jazdę autonomiczną a 8 czujników IR (*infra red*), umieszczonych na obwodzie robota, umożliwia pomiar światła otoczenia oraz wykrywanie przeszkód. W górnej części robota umieszczono mikrokontroler Motorola 68331 taktowany zegarem 16MHz. Łącze szeregowe o maksymalnej prędkości transmisji wynoszącej 38400 bodów umożliwia sterowanie robotem przy pomocy komputera nadrzędnego. Elementarna konfiguracja z rys.1 może być uzupełniona o takie dodatkowe moduły jak: chwytak - pozwalający na aktywne oddziaływanie robota ze środowiskiem zewnętrznym, moduł wizyjny, moduł komunikacji radiowej, czy też łącze szeregowe na podczerwień.



Rys. 1. Robot Khepera

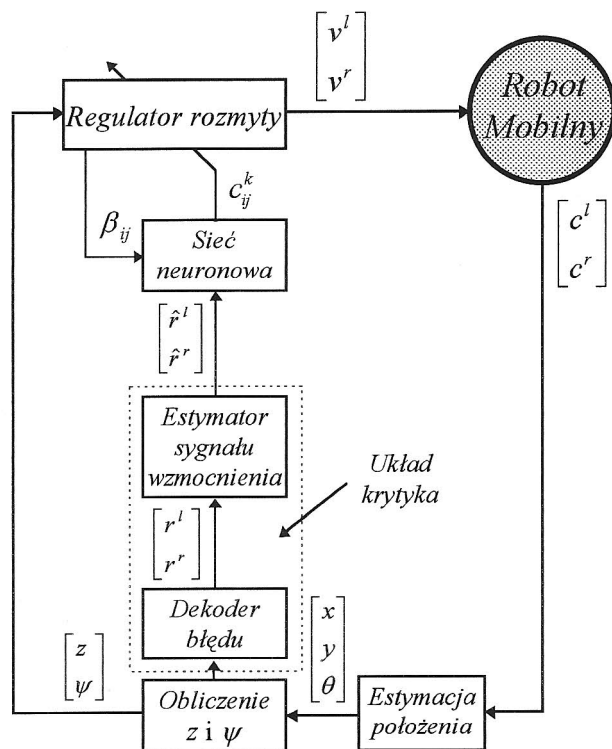


Rys. 2. Oznaczenia

3. UKŁAD STEROWANIA

Zadaniem regulatora rozmytego opisanego w następnym rozdziale będzie doprowadzenie robota do celu (rys. 2), czyli minimalizowanie odległości $z = \|X_g - X\|$ ($\|\cdot\|$ - norma Euklidesowa) robota od celu. $X = [x, y]$ oznacza pozycję robota w globalnym

układzie współrzędnych (X, Y) ; $X_g = [x_g, y_g]$ - współrzędne celu; θ - orientację robota; ψ - odchylenie robota od celu ($\psi = \theta - \arctan \frac{y_g - y}{x_g - x}$), a R - promień robota ($R = 26mm$).



Rys. 3. Układ sterowania

Na rys. 3 przedstawiono schemat adaptacyjnego układu sterowania robotem, realizującego zadanie poszukiwania celu. Na podstawie drogi pokonanej przez poszczególne koła (c^l i c^r) układ *estymacji położenia* szacuje współrzędne robota $[x, y, \theta]$. Następnie obliczane są zmienne wejściowe *regulatora rozmytego* z i ψ (rozdz. 4). Ponadto, na podstawie odchylenia robota od celu (kąta ψ) i odległości robota od celu (z) *dekodek błędów* wraz z *układem estymatora* wyznacza sygnał wzmocnienia odpowiednio dla lewego i prawego koła ($\hat{r}^l$, $\hat{r}^r$), który jest podstawą algorytmu *uczenia ze wzmocnieniem*, opisanego w rozdz. 5. Na zakończenie cyklu pracy układu regulacji *regulator rozmyty* na podstawie uaktualnionych wartości następników reguł c_{ij}^k wyznacza wartości sterowań, tj. prędkości lewego koła (v^l) i prawego koła (v^r).

4. REGULATOR ROZMYTY

Wyjście rozmytego regulatora Takagi-Sugeno w przypadku niesprzecznego i

z	ψ	c_{ij}^l	c_{ij}^r
A_1 (Near)	B_1 (Negative)	c_{11}^l	c_{11}^r
A_1 (Near)	B_2 (Zero)	c_{12}^l	c_{12}^r
A_1 (Near)	B_3 (Positive)	c_{13}^l	c_{13}^r
A_2 (Far)	B_1 (Negative)	c_{21}^l	c_{21}^r
A_2 (Far)	B_2 (Zero)	c_{22}^l	c_{22}^r
A_2 (Far)	B_3 (Positive)	c_{23}^l	c_{23}^r

Tab. 1. Tablica reguł

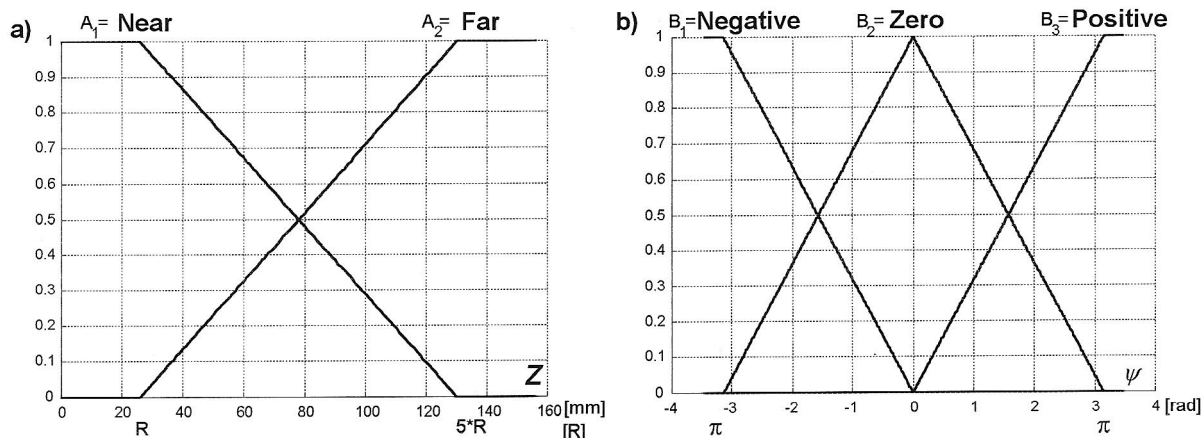
zupełnego systemu reguł sterowania obliczane jest na podstawie tablicy reguł (Tab. 1) według następującej zależności [6]:

$$v^k = \frac{\beta_{11}c_{11}^k + \beta_{12}c_{12}^k + \dots + \beta_{23}c_{23}^k}{\beta_{11} + \beta_{12} + \dots + \beta_{23}}, \quad (1)$$

gdzie współczynniki β_{ij} wyrażają się wzorem:

$$\beta_{ij} = A_i(z) \cdot B_j(\psi), \quad (2)$$

natomiast $A_i(z)$ i $B_j(\psi)$ są stopniami przynależności odpowiednio wejść z i ψ do zbiorów rozmytych A_i i B_j , a c_{ij}^k - wartościami sterowań (prędkościami sterującymi odpowiednio lewym i prawym kołem) z tablicy reguł (Tab. 1). Ponadto i jest indeksem zbioru zmiennej wejściowej z , j - zmiennej ψ , zaś $k = l, r$ - indeksem wyjścia regulatora odpowiednio dla lewego i prawego koła.



Rys. 4. Wejściowe funkcje przynależności dla: a) odległości robota od celu z , b) odchylenia od celu ψ .

Zaproponowano dwa zbiory rozmyte (*Near* i *Far*) dla zmiennej z i trzy zbiory dla zmiennej ψ (*Negative*, *Zero* i *Positive*) (rys. 4). Liczba reguł oraz wejściowe funkcje przynależności, które dobrano na zasadzie równomiernego podziału dziedziny zmiennych wejściowych z i ψ , będą niezmiennie. Uczeniu wzmocnionemu będą podlegały wartości sterowań c_{ij}^k .

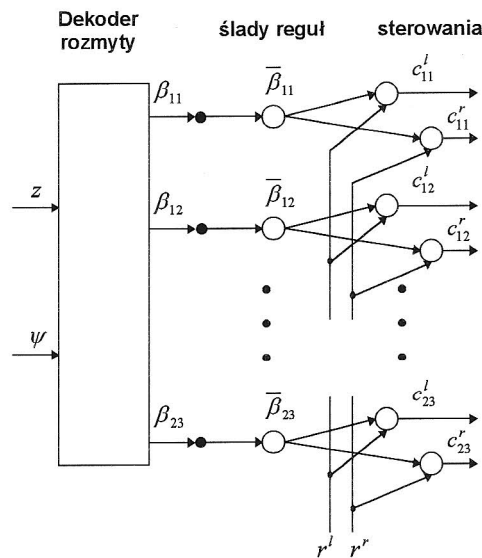
5. UCZENIE ZE WZMOCNIENIEM

5.1. ALGORYTM UCZENIA

Na rysunku 5 przedstawiono sieć neuronową umożliwiającą uczenie sterowań rozmytego regulatora zmodyfikowanego. Dekoder rozmyty, będący częścią regulatora rozmytego, oblicza poziomy zapłonu poszczególnych reguł zgodnie z zależnością (2). Następnie wyznaczane są tzw. ślady poszczególnych reguł $\bar{\beta}_{ij}$ (3) mówiące o przewidywanej aktywności reguły jako średniej ważonej jej poprzedniego śladu i aktualnego poziomu aktywacji [1,2]

$$\bar{\beta}_{ij}(t+1) = \lambda \bar{\beta}_{ij}(t) + (1 - \lambda) \beta_{ij}(t), \quad (3)$$

gdzie $0 \leq \lambda \leq 1$ jest współczynnikiem zapominania śladu reguły. Ślad reguły (opisany szerzej w pracy [2]) spełnia podobną rolę do uśrednionego parametru dopasowania (*eligibility trace*), stosowanego w typowym algorytmie uczenia ze wzmocnieniem [2,3]. Różnica polega na sposobie wyznaczenia składnika $\beta_{ij}(t)$ wzoru (3). W algorytmie typowym określany on bywa jako iloczyn dwu czynników, z których pierwszy zależy od aktualnie podjętej akcji (wartości sterowań), a drugi od wystąpienia określonego pobudzenia. U nas określony jest on zależnością (2). Jeżeli λ jest bliskie 0 to ślad reguły „zapomina” o „historii aktywności” reguły, liczy się tylko aktywność bieżąca (wyrażona przez $\beta_{ij}(t)$). Natomiast przy λ bliskim 1 ślad reguły zależy głównie od swojej „historii”, zaś w niewielkim stopniu od chwilowych zmian aktywności reguły. W eksperymentach przyjęto $\lambda = 0.5$.



Rys. 5. Sieć neuronowa realizująca uczenie ze wzmocnieniem.

Do uczenia sterowań c_{ij}^k zastosowano metodę uczenia ze wzmocnieniem. Jest to odmiana uczenia pod nadzorem, w którym nie występuje informacja o wartościach pożądanых na wyjściu systemu, a jedynie informacja, czy podjęta przez system akcja daje wyniki pozytywne w sensie pożądanego zachowania systemu, czy negatywne. Jest uczeniem bardziej uniwersalnym w zastosowaniu, gdyż nie wymaga znajomości sygnałów pożądanых na wyjściu systemu. Odbyna się ono według zależności

$$c_{ij}^k(t+1) = c_{ij}^k(t) + \eta \hat{r}^k(t) \bar{\beta}_{ij}(t), \quad (4)$$

w której $\hat{r}^k$ jest przewidywanym wewnętrznym sygnałem wzmocnionym (rozdział 5.3), zaś η współczynnikiem uczenia. Jest to klasyczna zależność określająca uczenie ze wzmocnieniem [2,3], różniącą się tylko definicją uśrednionego parametru dopasowania (u nas $\bar{\beta}_{ij}(t)$) wprowadzaną w pracy [1].

Jedyną trudnością uczenia ze wzmocnieniem może być budowa dekodera błędu (sformułowanie bazy wiedzy) oceniającego aktualnie podjętą akcję. W przypadku sterowania dwukołowym robotem mobilnym tworzenie sygnału wzmocnionego opisano poniżej.

5.2. ZEWNĘTRZNY SYGNAŁ WZMOCNIONY

Zadaniem dekodera błędu jest ocena akcji podjętej aktualnie przez system (regulator) i wytworzenie sygnału wzmacniającego tendencje pozytywne i osłabiającego tendencje

negatywne systemu. W przypadku sterowania robotem realizującym zadanie poszukiwania celu dekodery błędów powinny zapewnić zwracanie się robota do celu i podążanie w jego kierunku co można wyrazić następująco:

$$\begin{aligned} & \text{(Jeżeli } \psi > 0 \text{) to („zwiększaj_}\psi\text{” } c_{ij}^l \text{ i „zmniejszaj_}\psi\text{” } c_{ij}^r) \\ & \text{ i (Jeżeli } z > \varepsilon \text{) to („zwiększaj_}z\text{” } c_{ij}^l \text{ i } c_{ij}^r) \end{aligned}$$

w przeciwnym wypadku

$$\begin{aligned} & \text{(Jeżeli } \psi < 0 \text{) to („zmniejszaj_}\psi\text{” } c_{ij}^l \text{ i „zwiększaj_}\psi\text{” } c_{ij}^r) \\ & \text{ i (Jeżeli } z > \varepsilon \text{) to („zwiększaj_}z\text{” } c_{ij}^l \text{ i } c_{ij}^r), \end{aligned}$$

gdzie $\varepsilon = R$ jest dokładnością pozycjonowania a „zwiększaj_” „zmniejszaj_” oraz „zwiększaj_z” współczynnikami określającymi sposób zmiany sterowań.

W sposób formalny powyższą zależność można zapisać następująco:

$$\begin{aligned} r^l(z, \psi) &= \text{sgn}(\psi) \cdot \rho_\psi + \text{step}(z - \varepsilon) \cdot \rho_z \\ r^r(z, \psi) &= -\text{sgn}(\psi) \cdot \rho_\psi + \text{step}(z - \varepsilon) \cdot \rho_z \end{aligned} \quad (5)$$

gdzie: $r^k(z, \psi)$ jest zewnętrznym sygnałem wzmocnionym, step - funkcją skoku jednostkowego

$$\text{step}(x) = \begin{cases} 1 & \text{gdy } x > 0 \\ 0 & \text{gdy } x \leq 0 \end{cases}$$

a ρ_ψ i ρ_z są współczynnikami określającymi wagę poszczególnych akcji, tj. odpowiednio zwracania się do celu („zwiększaj_” „zmniejszaj_”) i zmniejszania odległości („zwiększaj_z”). Przyjęto $\rho_\psi = 0.5$ i $\rho_z = 0.05$, co oznacza że robot przede wszystkim powinien zwracać się do celu, a następnie do niego dojeżdżać.

Jeżeli założymy, że uczenie będzie prowadzone tylko dla $z > \varepsilon$, gdzie $\text{step}(z - \varepsilon) = 1$, oraz po podstawieniu współczynników liczbowych do zależność (5) otrzymamy funkcje wzmacniające uczenie

$$\begin{aligned} r^l(\psi) &= \text{sgn}(\psi) \cdot 0.5 + 0.05 = \begin{cases} 0.55 & \text{dla } \psi > 0 \\ -0.45 & \text{dla } \psi < 0 \end{cases} \\ r^r(\psi) &= -\text{sgn}(\psi) \cdot 0.5 + 0.05 = \begin{cases} -0.45 & \text{dla } \psi > 0 \\ 0.55 & \text{dla } \psi < 0 \end{cases} \end{aligned} \quad (6)$$

Zauważmy, że przyjęte ρ_z jest 10 razy mniejsze od ρ_ψ . Nasuwa się więc pytanie: Czy nie można by pominąć wpływu „akcji” związanej ze zmniejszaniem odległości, czyli przyjąć, że $\rho_z = 0$? Otóż można tak zrobić [1], lecz osiągnięcie poszukiwanego celu w trakcie uczenia jest wówczas procesem długotrwałym. Eksperymentalnie stwierdzono, że minimalna nawet zależność funkcji wzmacniających uczenie od z ($\rho_z > 0$) znacznie skraca czas uczenia.

Podsumowując można stwierdzić, że przy budowie dekodera błędów wymagane jest jasne sformułowanie wymagań odnośnie procesu uczenia oparte na znajomości systemu sterowania. Im większa jest nasza wiedza, tym lepszy możemy zbudować dekodery i tym lepsze osiągniemy rezultaty.

5.3. WEWNĘTRZNY SYGNAŁ WZMOCNIONY

Zadaniem układu estymacji wewnętrznego sygnału wzmocnienia (rys. 3) jest predykcja sygnału $\hat{r}^k$. Zgodnie z pracami [1,2,3,7] jego estymatę w chwili t określa zależność

$$\hat{r}^k(t) = r^k(t) + \gamma p^k(t) - p^k(t-1) \quad (7)$$

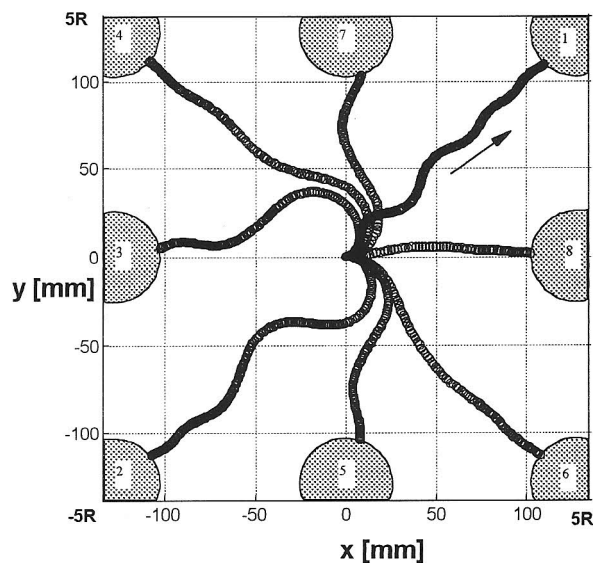
gdzie $0 \leq \gamma < 1$ jest współczynnikiem określającym wygaszanie (zapominanie) wzmocnień odległych w czasie, $r^k(t)$ - zewnętrznym sygnałem wzmocnionym (rozdział 5.2) a $p^k(t)$ - estymatą wskaźnika jakości wyrażoną przez wzór

$$p^k(t) = G \left(\sum_{i=1}^{L_z} \sum_{j=1}^{L_\psi} c_{ij}^k(t) \beta_{ij}^k(t) \right) \quad (8)$$

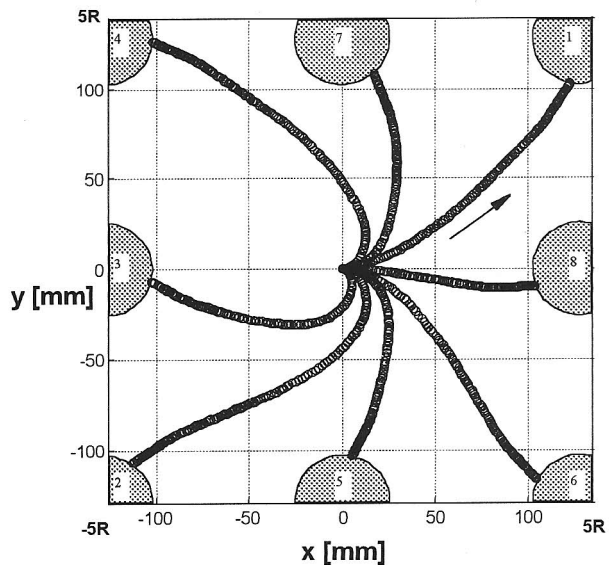
gdzie: $L_z = 2$ i $L_\psi = 3$ są współczynnikami określającymi liczbę zbiorów rozmytych przyjętych odpowiednio dla zmiennej z i ψ , zaś $G(\cdot)$ oznacza funkcję logistyczną określoną jako $G(x) = 2 / (1 + e^{-x}) - 1$.

6. WYNIKI EKSPERYMENTÓW

Program, napisany w języku C, generował sterowanie co $55ms$. Komunikacja z robotem odbywała się za pośrednictwem łącza szeregowego. Początkowe wartości sterowań c_{ij}^k były małymi liczbami losowymi z przedziału $(0 \div 0.1)$. Eksperyment obejmował poszukiwanie ośmiu celów o współrzędnych $(5R, 5R)$, $(-5R, -5R)$, $(-5R, 0)$, $(-5R, 5R)$, $(0, -5R)$, $(5R, -5R)$, $(0, 5R)$ i $(5R, 0)$. Robot rozpoczynał poszukiwanie celu zawsze z punktu o współrzędnych $(0, 0)$, a kończył, gdy odległość od celu była mniejsza od promienia robota R . Trajektorie robota w trakcie uczenia przedstawiono na rys. 6. Szarymi kołami zaznaczono położenie punktów docelowych wraz z otoczeniem równym R .



Rys. 6. Trajektorie robota w trakcie uczenia wzmocnionego



Rys. 7. Trajektorie robota realizującego zadanie poszukiwania celu z wykorzystaniem nauczonego zbioru reguł

z	ψ	c_{ij}^l	c_{ij}^r
Near	Negative	-2.02	2.21
Near	Zero	1.92	1.65
Near	Positive	2.32	-2.14
Far	Negative	-8.27	9.48
Far	Zero	4.28	3.03
Far	Positive	7.82	-6.91

Tabela 2. Tablica reguł realizująca zadanie poszukiwania celu

W wyniku uczenia otrzymano wartości sterowań, które przedstawiono w tabeli 2. Jednostką prędkości jest *impuls/10ms*, co odpowiada rzeczywistej prędkości *8mm/s*.

W celu sprawdzenia poprawności pracy nauczonego regulatora rozmytego przeprowadzono eksperymenty, których celem było osiągnięcie wspomnianych wcześniej punktów. Zadane punkty zostały osiągnięte (rys. 7).

VII. PODSUMOWANIE

Zastosowana metoda uczenia wzmocnionego pozwoliła na adaptacyjny dobór sterowań rozmytego regulatora zmodyfikowanego. Ingerencja „eksperta” sprowadziła się do podania funkcji wzmocnienia, które formułują pożądaną cechę zachowania się robota w trakcie procesu uczenia. Zalety tej metody uwidaczniają się w przypadku dużej liczby reguł, gdy intuicyjne ich sformułowanie staje się uciążliwe lub praktycznie niemożliwe.

Praca finansowana przez Komitet Badań Naukowych w ramach grantu nr 8 T11A 028 11.

LITERATURA

- [1] BEOM H., CHO H., *A Sensor Based Navigation for a Mobile Robot Using Fuzzy Logic and Reinforcement Learning*, IEEE SMC, Vol. 25, No. 3, 464-477, March 1995.
- [2] BERENJI H.R., *A Reinforcement Learning-Based Architecture for Fuzzy Logic Control*, International Journal of Approximate Reasoning, Vol. 6, No. 2, 267-292, February 1992.
- [3] BARTO A., SUTTON R.S., ANDERSON C.W., *Neuronlike Adaptive Elements That Can Solve Difficult Learning Control Problems*, IEEE SMC, Vol. 13, No. 5, 834-846, September/October 1983.
- [4] K-Team, *Khepera user manual*, Lusanne, 1995.
- [5] K-Team, *General I/O turret user manual*, Lusanne, 1995.
- [6] KLUSKA J., *Sterowanie z logiką rozmytą*, Zeszyty Naukowe Politechniki Rzeszowskiej, Seria Elektrotechnika, Zeszyt 12, Nr 104, Rzeszów 1992.
- [7] LIN C.T., *A neural fuzzy control system with structure and parameter learning*, Fuzzy Sets and Systems 70 (1995) 183-212.
- [8] MILLÁN J., *Rapid, Safe, and Incremental Learning of Navigation Strategies*, IEEE SMC, Part B: CYBERNETICS, Vol. 26, No. 3, 408-420, June 1996.
- [9] ZAJDEL R., *Synteza regulatora rozmytego i neuronowego dla obiektów nieliniowych*, Konferencja Naukowo - Techniczna, Zastosowania Komputerów w Elektrotechnice, Poznań/Kiekrz 1996, s. 161 - 164.
- [10] ZAJDEL R., *Układy logiki rozmytej i sieci neuronowych. Sprzęt i oprogramowanie - część I. Pomiary Automatyka Kontrola*, 8, 209 - 212, 1996.
- [11] ZAJDEL R., *Układy logiki rozmytej i sieci neuronowych. Sprzęt i oprogramowanie - część II. Pomiary Automatyka Kontrola*, 9, 237 - 240, 1996.