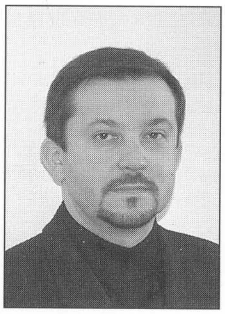
Gabriel G. KOST
POLITECHNIKA ŚLĄSKA W GLIWICACH, KATEDRA AUTOMATYZACJI PROCESÓW TECHNOLOGICZNYCH I ZINTEGROWANYCH SYSTEMÓW WYTWARZANIA

Sterowanie stochastyczne jako podstawa planowania ruchu robotów manipulacyjnych

Dr inż. Gabriel G. KOST



ur. 17. lutego 1960r., ukończył studia na Wydziale Mechanicznym Technologicznym Politechniki Śląskiej w Gliwicach w 1984r. uzyskując tytuł magistra inżyniera mechanika na kierunku studiów Technologia Maszyn. Po zakończeniu studiów rozpoczął prace w Instytucie Budowy Maszyn tego Wydziału. W 1991 roku uzyskał stopień doktora nauk technicznych w specjalności Budowa i Eksploatacja Maszyn za pracę na temat: „System komputerowego wspomagania programowania przegubowych robotów przemysłowych".
W swojej działalności naukowej dr inż. Gabriel Kost zajmuje się zagadnieniami komputerowego wspomagania prac inżynierskich z zakresu technologicznego przygotowania produkcji, a w szczególności programowania zrobotyzowanych, elastycznych systemów technologicznych i robotów przemysłowych. Obecnie jest adiunktem w Zakładzie Automatyzacji Procesów Technologicznych Katedry Automatyzacji Procesów Technologicznych i Zintegrowanych Systemów Wytwarzania Wydziału Mechanicznego Technologicznego Politechniki Śląskiej w Gliwicach.
e-mail: gabriel.kost@polsl.pl

Streszczenie

Artykuł przedstawia metodę komputerowo wspomaganego programowania off-line robotów przemysłowych wykorzystującej do analizy przestrzeni roboczej manipulacyjnych robotów przemysłowych procesy decyzyjne Markowa. Przedstawiono koncepcję oraz wymagania jakie powinien spełniać sprawny informatycznie i technologicznie komputerowy system programowania off-line robotów przemysłowych. Koncepcję systemu oparto na metodzie uczenia się ze wzmocnieniem. Zdefiniowano podstawowe parametry omawianego procesu planowania opartego na sterowaniu stochastycznym, takie jak: proces Markowa, funkcje prawdopodobieństw warunkowych przejścia i osiągnięcia typowanych przemieszczeń robota.

Abstract

The scope of the paper is the presentation of computer-aided method of programming off-line industrial robots based on Markov's decision processes applied to analyze the operation space of manipulative industrial robots. The concept of a computer system of programming off-line industrial robots is discussed as well as the requirements that it should fulfill to secure IT and technical efficiency. The concept of the system is based on a method of machine learning with reinforcement. The fundamental parameters of the stochastic process planning are defined, including: the Markov's decision process, the space of the robot state, conditional probability functions of robot passage and destinations as specified by stochastic planning process.

1. Wstęp

Każde zadanie planowania ruchu robota (zadanie planowania - ZP), związane jest z problemem analizy przestrzeni. Dla robotów stacjonarnych jest to zawsze problem trójwymiarowy. Spośród znanych metod zajmujących się tym zagadnieniem najszerzej reprezentowane są metody geometryczne [2,3,4], a w śród nich metody grafowe, oparte na technice wielopoziomowego rastrowania przestrzeni i budowania grafu - drzewa decyzyjnego będącego jej reprezentantem, którego wierzchołki określone są parametrami geometrycznymi utworzonych rastrów [2,3]. Tak prowadzona analiza powoduje, że budowany graf jest bardzo rozległy. Wymusza to konieczność zastosowania technik przeszukiwania heurystycznego do poszukiwania rozwiązań dopuszczalnych ścieżek bezkolizyjnych. Wymóg stosowania heurez jest powodem małej efektywności obliczeniowej tych metod. Ich przeciwieństwem są techniki stosowane do rozwiązywania ZP w robotyce mobilnej. Tam, konieczność skutecznego radzenia sobie robota w zmiennym otoczeniu wymusza, by planowanie ruchu było szybkie i skuteczne tak, by umożliwić sterowanie robotem w czasie rzeczywistym.

Konfrontacja możliwości metod planowania robotyki mobilnej i stacjonarnej pozwala dostrzec, że przyczyną, z powodu której metody robotyki stacjonarnej są mało efektywne jest wymiarowość ZP. Jej konsekwencją jest to, że scena robota manipulacyjnego zawsze klasyfikowana jest jako statyczna, w której rozlokowanie poszczególnych obiektów - przeszkód robota, ich kształt, wielkość, nie ulega zmianom. Dzięki takiemu warunkowi analiza przestrzeni nie musi przebiegać dynamicznie, stosownie do zmian w niej zachodzących. Jednak w warunkach przemysłowych, środowisko robota manipulacyjnego nie spełnia w pełni warunków statyczności.

Spostrzeżenia te pozwalają na podjęcie próby wykorzystania możliwości robotyki mobilnej w zadaniu planowania bezkolizyjnych ścieżek robotów manipulacyjnych, co drogą odpowiednio poprowadzonego procesu adaptacji pozwoliłoby uzyskać efektywną czasowo metodę planowania bezkolizyjnego ruchu robotów, przez co mogłaby ona być wykorzystana w systemach programowania off-line robotów przemysłowych.

2. Adaptacja metody uczenia się ze wzmocnieniem

Jedną z prostszych metod stosowanych w planowaniu ruchu robotów mobilnych jest metoda uczenia się ze wzmocnieniem. Charakteryzuje się ona tym, że [3]:
- uczy pewnego typu ogólnej reakcji, tzw. ogólnej strategii działania i nie generuje trwałych, określonych procedur postępowania oraz umożliwia wykorzystanie strategii lokalnych, wspomagających proces uczenia, charakterystycznych dla ZP,
- wykorzystuje typowy i zarazem prosty mechanizm korekcji zachowań ucznia jakim jest ocena jego postępowania,
- nie wymaga śledzenia otoczenia drogą szybkich interakcji układów sensorowych w układzie robot (uczeń) - otoczenie.

Adaptując metodę uczenia się ze wzmocnieniem do ZP, wykorzystano zasadę nagradzania i możliwość istnienia lokalnej strategii postępowania ucznia, który w omawianym ZP identyfikowany jest z pewnym blokiem funkcyjnym zwanym PLANEREM ruchu robota. Zadaniem PLANERA jest, na podstawie posiadanych informacji o konfiguracji geometrycznej otoczenia robota, poprawnie kreować ruch robota, tak, by w oparciu o przyjętą ogólną strategię działania doprowadzić go do wskazanego celu, jakim jest wyznaczone miejsce w przestrzeni otaczającej robota. Przebieg tego procesu pokazuje rys. 1. Planowanie bezkolizyjnego ruchu robota bazuje zatem na informacji definiującej aktualny stan otoczenia, wyrażonej zestawem danych geometrycznych o konfiguracji przestrzennej wszystkich obiektów sceny robota. Na tej podstawie planer wyciąga odpowiednie wnioski o możliwościach typowania przemieszczenia robota, zgodnie z przyjętą ogólną strategią Ψ działania PLANERA:

Rys. 1. Schemat działania PLANERA ścieżek robota [3]

$$\xi = \max_{i \to \min} E\left[\sum_{i=1}^{k} \omega_i \cdot r(P_{w_i})\right] \quad (1)$$

gdzie:

- $E[\bullet]$ - uśredniona (estymowana) wartość zdobywanych ocen,
- ω_i - waga uzyskiwanych ocen określona dla kroku i wg odpowiedniej funkcji wag,
- $r(P_{wi})$ - funkcja nagród uzyskiwanych za zrealizowanie zaplanowanego ruchu do P_{wi},
- i - kolejny zaplanowany krok.

3. Warunki zadania

Działanie PLANERA, ze względu na potencjalne skutki, musi się odbywać etapami. Etapy te, związane są z przemieszczeniem efektora robota do kolejnego wyznaczonego miejsca w przestrzeni sceny identyfikowanego z kolejnym punktem węzłowym P_{wi} poszukiwanej ścieżki bezkolizyjnej, nazywać będziemy krokami t_i. Planowanie powinno być zatem procesem iteracyjnym. Każda z iteracji ZP powinna prowadzić do osiągnięcia kolejnego punktu węzłowego P_{wi}. Poszukiwaną ścieżkę możemy zatem zapisać jako:

$$TR=(P_{w1},P_{w2},...,P_{wi},...,P_{wk}) = (t_1 \cup t_2 \cup ... \ t_i \cup ... t_k) \quad (2)$$

gdzie:

$$P_{wi+1}=P_{wi}+t_{i+1}$$

Oczywiście, każdy z wyznaczonych przez PLANER P_{wi}, musi spełniać warunek osiągalności, to znaczy musi być możliwy do osiągnięcia przez robot, czyli musi spełniać warunek:

$$T_Q^{-1}(q_i) = P_w \ \text{ oraz } \ T_Q(P_{wi}) = q_i \quad (3)$$

gdzie: T_Q - transformacja prosta kinematyki robota, T_Q^{-1} - transformacja odwrotna modelu kinematycznego manipulatora, q_i - wektor współrzędnych uogólnionych robota określony w przestrzeni stanów wewnętrznych (konfiguracyjnych) robota Q ($\bigcup_i q_i = Q$).

4. Sterowanie stochastyczne

Jedną z cech metod grafowych analizy przestrzeni, np. stosowanej powszechnie dla robotów manipulacyjnych metody drzew oktalnych [2,4,3], jest zjawisko błądzenia [3]. Polega ono na przypadkowym typowaniu kolejnego rastra (oktantu) spośród sąsiadujących z rastrem bieżącym. Błądzenie jest główną przyczyną znacznego rozrastania się drzewa decyzyjnego, reprezentującego analizowaną przestrzeń oktalną. Jednym ze sposobów uniknięcia tego problemu może być przyjęcie, jako podstawy poruszania się PLANERA w zrastrowanej przestrzeni, zasady sterowania stochastycznego [3].

Sterowanie stochastyczne jest procesem polegającym na losowym, opisanym pewną funkcją gęstości prawdopodobieństwa rozciągniętą na wszystkie rastry analizowanej przestrzeni [3,5], typowaniu kolejnego kroku t_i. Funkcja ta ułatwia typowanie kolejnego rastra, stosownie do szans osiągnięcia zaplanowanego celu działania i może być wykorzystana do oceny (parametr $r(P_{wi})$, (1)) PLANERA typującego kolejny krok t_i zgodnie z wymaganiami metody uczenia się i obowiązującej strategii działania Ψ.

W planowaniu ruchu robotów opartym na sterowaniu stochastycznym, wykorzystuje się procesy decyzyjne Markowa [5,4,3]. Stanowią one model matematyczny tego procesu [1,4]. Cechą charakterystyczną procesów decyzyjnych Markowa (PDM) jest to, że opisują one zjawiska, których stan przyszły s_{i+1} zależy wyłącznie od stanu bieżącego s_i [5]. Informacja o tym, że analizowany układ znajduje się w bieżącym stanie s_i, wystarcza do tego, by określić szanse (prawdopodobieństwo) jego przejścia do kolejnego stanu s_{i+1}.

O ciągu zdarzeń $X(s)$ zbudowanym na skończonym zbiorze S dyskretnych i uporządkowanych stanów ($s_i \in S$ i $i=0,1,...,k$), na którym rozpięte jest prawdopodobieństwo wystąpienia s_i o rozkładzie $p(s_i)$ i relacja przejścia do stanu kolejnego s_{i+1} o rozkładzie $p(s_i,s_{i+1})$ mówi się, że jest *łańcuchem Markowa* [5] wtedy, gdy zachodzi równość:

$$\begin{aligned} P(X(s_{k+1}) = s_{k+1} \,|\, X_k = s_k, X_{k-1} = s_{k-1},..., X_0 = s_0) &= \\ = P(X(s_{k+1}) = s_{k+1} \,|\, X_k = s_k) &= \quad (4) \\ = p(s_k, s_{k+1}) & \end{aligned}$$

przy warunku, że $k \geq 0$. Konsekwencjami tych założeń zrealizowanych w przestrzeni oktalnej jest to, że:

- zbiór S określający przestrzeń stanów analizowanego układu: przestrzeń oktalna-robot, jest skończony, niezależnie od dokonanego poziomu p podziału oktalnego,
- każdy element (oktant) $s_i \in S$ jest dyskretny i na kolejnych podziałach tworzy również dyskretnie zdefiniowane stany (oktanty kolejnych podziałów),
- warunkiem wystarczającym, by funkcje prawdopodobieństwa mogły być zdefiniowane na S, jest, by były określone dla każdego $s_i \in S$

Zdefiniowany (4) łańcuch Markowa wyrażony jako ciągła funkcja czasu nazywany jest procesem decyzyjnym Markowa (PDM) [5]. W robotyce czas, który jest podstawowym parametrem PDM oznacza, że dowolne zdarzenie wpływające na ruch robota może zajść w dowolnej chwili [4]. Tym samym czas nie musi spełniać roli argumentu PDM, by obowiązywały dla niego wszystkie zależności dotyczące (4) [5].

5. Analiza przestrzeni robota

Przestrzeń, w której możliwe jest zdefiniowanie i zrealizowanie ścieżki robota, musi być obszarem wolnym od obiektów będących dla niego przeszkodami. Przestrzeń kolizyjną robota Π_{OB} wyznaczają przestrzenie zajmowane przez wszystkie obiekty sceny Π_{OB}, czyli:

$$\Pi_{OB} = \Pi_{OB^1} + \Pi_{OB^2} + ... + \Pi_{OB^n} = \bigcup_{j=1}^{n} \Pi_{OB^j}$$

gdzie: j - liczba obiektów w scenie.

Ścieżkę robota, która spełnia warunek bezkolizyjności:

$$TR \cap \Pi_{OB} = \underset{i}{\forall} t_i \cap \underset{j}{\forall} \left(\Pi_{OB^j}\right) = 0 \quad (5)$$

nazywać będziemy ścieżką bezkolizyjną i wtedy pomiędzy nią, a obiektami sceny nie zachodzi żadna relacja incydencji. Zależność ta określa podstawowy warunek bezpiecznej ścieżki robota.

Proponowana metoda planowania zakłada, że analiza przestrzeni robota przeprowadzona jest w oparciu o jej rastrową dyskretyzację na bazie metody oktalnej. Zdyskretyzowana scena robota zdefiniowana w swoim własnym układzie U_{SC} i poddana analizie oktalnej na poziomie p charakteryzuje się skończoną liczbą oktantów. Uporządkowanie dokonywanych podziałów oktalnych pozwala wyznaczyć rekurencyjne zależności, dzięki którym możliwe jest ustalenie współrzędnych środka wskazanego oktantu (punktu podporowego P_{wi}), stosownie do przyjętego planu działania [2,3]. Przyjmuje się, że środek geometryczny każdego oktantu identyfikuje go w sposób jednoznaczny i wskazuje punkt węzłowy $P_{wi}(x_i,y_i,z_i)$ rozpoznawany przez PLANER jako stan s_i układu. Dwa sąsiadujące ze sobą oktanty (tab. 1): bieżący $s_i{:}P_{wi}$ i wskazany jako kolejny $s_{i+1}{:}P_{wi+1}$, określają krok t_{i+1} planowanej ścieżki.

Przejście robota ze stanu $s_i(P_{wi})$ do stanu $s_{i+1}(P_{wi+1})$ może nastąpić wtedy, kiedy stany te są stanami sąsiednimi (tab. 1), czyli takimi, które mają z s_i bok wspólny. Wszystkie stany sąsiednie s_{i+1} względem stanu bieżącego s_i tworzą zbiór stanów sąsiednich ${}_sS = \bigcup_i s_{i+1}$. Liczebność tego zbioru jest równa:

$$N = card({}_sS) = \sum_{p'} \left\{ \left[\left(2^3\right)^{p_{s_i}} - 2 \right] + \left(\frac{2^{3p'}}{2} - 1 \right) \right\} \quad (6)$$

gdzie: p - początkowy (startowy) poziom podziału oktalnego, $p' = |p_i - p_{s_i}|$, p_{s_i} - poziom podziału, na którym powstał stan bieżący s.

Do zbioru stanów sąsiednich, spełniających warunki określone w definicji, należy zatem sześć sąsiednich rastrów (oktantów) o parametrach punktów węzłowych określonych w tablicy 1 [3].

Tab. 1. Współrzędne punktów węzłowych w zbiorze stanów sąsiednich

	Stany sąsiednie
${}_sS$	$s_{i+1}(x_{i+1}-t_{i+1},y_{i+1},z_{i+1})$
	$s_{i+1}(x_{i+1}+t_{i+1},y_{i+1},z_{i+1})$
	$s_{i+1}(x_{i+1},y_{i+1}+t_{i+1},z_{i+1})$
	$s_{i+1}(x_{i+1},y_{i+1}-t_{i+1},z_{i+1})$
	$s_{i+1}(x_{i+1},y_{i+1},z_{i+1}-t_{i+1})$
	$s_{i+1}(x_{i+1},y_{i+1},z_{i+1}+t_{i+1})$

Dzięki uporządkowaniu stosowanych podziałów, stany s_i są stanami dyskretnymi, opisanymi parametrami geometrycznymi P_{wi}. Dlatego też przestrzeń stanów zadania planowania jest przestrzenią zdyskretyzowaną.

6. Model PLANERA ścieżek bezkolizyjnych

Niech robot znajduje się w pewnym, punkcie węzłowym P_{wi} identyfikowanym z początkiem ścieżki, wyznaczającym stan s_i powstałym na poziomie p podziału oktalnego. Stan ten nazwano stanem początkowym. Przejście PLANERA ze stanu początkowego s_i do końcowego s_k będzie się składało wg (1) z k kroków. Niech ten ciąg kroków jest realizowany zgodnie z ustaloną, ogólną strategią postępowania PLANERA Ψ, określająca podzbiór S^Ψ, gdzie:

$$\bigcup_{\Psi : i=1}^{k} s_i = S^\Psi \quad i \quad S^\Psi \in S \tag{7}$$

Kolejne kroki t_{i+1} robota będą się odbywały pomiędzy wskazanymi przez PLANER, zgodnie z Ψ, stanami s_i i s_{i+1}. Ciąg zrealizowanych kroków pomiędzy kolejnymi stanami: $s_1 \to s_2 \to \ldots \to s_{k-1} \to s_k$ wyznacza planowaną ścieżkę robota TR. Zgodnie z przyjętą zasadą sterowania stochastycznego, wytypowanie kolejnego stanu odbywa się z pewnym prawdopodobieństwem, pobytu PLANERA w stanie s_i określonym funkcją $p(s_i)$, zaś przejście t_{i+1} jest zrealizowane z prawdopodobieństwem równym $p(k,s_i,s_k)$, gdzie k - liczba zrealizowanych kroków t_k planowanej ścieżki. Ponieważ kolejne przejścia są możliwe tylko w obrębie zbioru stanów sąsiednich wtedy, że $k=1$, czyli $p(1,s_i,s_{i+1}) \equiv p(s_i,s_{i+1})$.

Żeby zatem działania PLANERA mogły być oparte na mechanizmie wzmocnienia (1), konieczne jest zdefiniowanie składników funkcji oceny, która będzie stymulować pożądane działania PLANERA. Zatem nagrodę, określoną funkcją $r(P_{wi})$, związano z prawdopodobieństwem zrealizowania wytypowanego przez PLANER przejścia ze stanu s_i do s_{i+1}. Jeżeli wybór stanu s_{i+1} zależy wyłącznie od s_i, wtedy zgodnie z PDM (4):

$$r(P_{wi+1}) = r_{i+1} = p(s_i) \cdot p(s_i, s_{i+1}) \tag{8}$$

Dla ZP oznacza to, że stan końcowy s_k planowanej ścieżki robota jest osiągalny z położenia początkowego s_1, poprzez osiągnięcie wszystkich wyznaczonych kolejno przez PLANER stanów pośrednich robota (P_{wi}) i wtedy, dla k kroków [3,5]:

$$r_k(P_{wk}) = p(s_1) \cdot p(s_1,s_2) \cdot p(s_2,s_3) \cdot \ldots \cdot p(s_i,s_{i+1}) \cdot \ldots \cdot p(s_{k-1},s_k)$$

PLANER wybierając stan niezgodny z Ψ zrealizuje go w ogólnym przypadku z prawdopodobieństwem przeciwnym, równym:

$$\overline{p}(s_i,s_{i+1}) = 1 - p(s_i,s_{i+1})$$

rozłożonym na wszystkie pozostałe (nie wybrane) stany sąsiednie.

Relacje, jakie zachodzą dla omawianych zachowań PLANERA, pozwalają zdefiniować tzw. macierz M funkcji przejść PDM dla omawianego zadania w postaci (9) [3]:

$$M(k,s_i,s_j) = \left[\begin{array}{c|cccccc} {}^{\ \ i}_{j} & 1 & 2 & 3 & 4 & \ldots & k \\ \hline 1 & p(s_1) & p(s_1,s_2) & 0 & 0 & \ldots & 0 \\ 2 & \overline{p}(s_2,s_3) & p(s_2) & p(s_2,s_3) & & \ldots & 0 \\ 3 & & \overline{p}(s_3,s_4) & p(s_3) & p(s_3,s_4) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ k & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \tag{9}$$

Przedstawiona macierz funkcji przejść M (9) jest macierzą stochastyczną [68] spełniającą warunek, że $\sum_{j=1}^{k} p(s_i,s_j) = 1$ dla każdego j, co oznacza, że może być uważana za reprezentację omawianego PDM. Stan s_k jest stanem dla którego $p(s_k,s_k)=1$ co oznacza, że PLANER realizujący swoje działania w oparciu o PDM nie może już z niego wyjść i kończy działanie.

6.1. Funkcje rozkładu prawdopodobieństwa

Dokonany, początkowy podział oktalny p przestrzeni sceny robota pozwala na wyznaczenie prawdopodobieństwa osiągnięcia wybranego stanu s_i. Jeżeli parametrem tego podziału p przestrzeni jest długość boku b^p oktantu, wtedy liczba N^p oktantów należących do przestrzeni bezkolizyjnej Π_w jest równa:

$$N_w = \mathrm{mod}\left(\frac{\Pi_w}{\left(b^p\right)^3}\right) \tag{10}$$

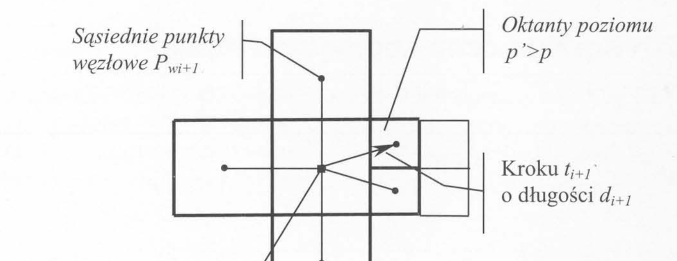
Zakładając, że lokalnie dla wybranych przez PLANER oktantów mogą zachodzić podziały na wyższych poziomach p', to ich liczba w przestrzeni bezkolizyjnej jest równa:

$$N_w^{p'} = (N_w - 1) + \sum_{p'} (2^{3p'} - 1) \tag{11}$$

co pozwala określić prawdopodobieństwo, wybrania dowolnego sąsiedniego rastra „pustego" jako:

$$p(s_w) = \frac{1}{N_w^{p'}} = \frac{1}{(N_w - 1) + \sum_{p'} (2^{3p'} - 1)} \tag{12}$$

Zależność (12) określa prawdopodobieństwo osiągnięcia stanu s_i zgodnie z ogólną strategią Ψ działania PLANERA.



Rys. 2. Układ zbioru stanów sąsiednich (rzut na płaszczyznę)

Analogicznie, ustalono wartośćfunkcji rozkładu pradopodobieństwa zrealizowania przejścia, na przestrzeni jednego kroku t_{i+1} o długości równej d_{i+1}, pomiędzy dwoma sąsiednimi stanami (rys. 2), jako:

$$p(s_i, s_{i+1}) = \frac{d_{i+1}}{D_{P_{wi+1}}} \tag{13}$$

gdzie:

d_{i+1} - długość drogi do wytypowanego stanu s_{i+1},

przy czym:

$$d_{i+1} = \begin{cases} b^p & \text{dla} \quad p \\ \dfrac{\left[\left(b^p + b^{p'}_{i+1}\right)^2 + \left(b^{p'}_{i+1}\right)^2\right]^{\frac{1}{2}}}{2} & \text{dla} \quad p' > p \end{cases} \qquad (14)$$

$Dp_{w_{i+1}}$ - suma długości wszystkich dróg prowadzących do wszystkich stanów sąsiednich s_{i+1}.

Z przeprowadzonego rozumowania wynika, że ustalona jak w (13) wartość prawdopodobieństwa przejścia robota ze stanu bieżącego s_i do kolejnego s_{i+1}, wyznaczonego przez PLANER, zależy wprost od odległości związanego z nim punktu węzłowego P_{wi+1}. Zatem stosownie do (1), PLANER powinien wybierać w pierwszej kolejności punkty węzłowe określające kolejne stany należące do zbioru stanów sąsiednich s_{i+1}, które będą możliwie najdalej oddalone od położenia bieżącego s_i, czyli te dla których $d_{i+1} \rightarrow \max$, gdyż dla nich wartość zdobywanej oceny będzie największa.

6.2. Ocena podjętych działań

Reasumując, można dzięki określonym prawdopodobieństwom (12), (13) w sposób jednoznaczny stwierdzić, że wartość oceny, jaką uzyska PLANER za wyznaczenie i osiągnięcie kolejnego stanu s_{i+1} oraz przejście do niego, a tym samym kolejnego punktu węzłowego określającego poszukiwaną ścieżkę bezkolizyjną, wyniesie wg (8), (12) i (13):

$$r_{i+1} = \frac{1}{(N_w - 1) + \sum\limits_p (2^{3p} - 1)} \cdot \frac{d_{i+1}}{D_{sP_{w_{i+1}}}} \qquad (15)$$

co zgodnie z (1) daje:

$$\xi = \max_{i \rightarrow \min} \left[E\left(\sum_{i=1}^{k} \omega_i \cdot \frac{1}{\left[(N_w - 1) + \sum\limits_p (2^{3p} - 1)\right] \cdot D_{sP_{w_{i+1}}}} \cdot d_{i+1} \right) \right] \qquad (14)$$

Zależność ta, może być również uznana za proste kryterium optymalizacyjne poszukiwanej ścieżki, gdyż $i \rightarrow min$ oznacza, że najlepsza jest ta spośród znalezionych, która umożliwia uzyskanie najwyższej oceny przy najmniejszej liczbie kroków PLANERA.

6.3. Korekta oceny podjętych działań

Jak już wspomniano oparcie omawianego ZP na rastrowej analizie przestrzeni, wymaga czasami przeprowadzenie, lokalnie dla wybranych oktantów uszczegółowienia analizy i dokonania kolejnych podziałów. Zakładając, że w przyjętym PDM wszystkie reakcje PLANERA są typowe [3], uogólniono wartość zdobywanych ocen (8) do postaci:

$$r_{i+1}(P_{wi+1}) = \gamma_{i+1} \cdot p(s_i) \cdot p(s_i, s_{i+1}) \qquad (17)$$

gdzie: γ_{i+1} - przyjęty współczynnik proporcjonalności, $\gamma = (0,1]$. Przy czym $\gamma = 1$ dla oktalnego podziału p przyjętego za podstawowy i $\gamma < 1$ proporcjonalnie dla podziałów p' na wyższych poziomach ($p'>p$). γ wyznacza zatem zależność pomiędzy prawdopodobieństwem zajścia stanu s_i na początkowym (startowym) poziomie p podziału oktalnego. Tym samym funkcję oceny r (8) dla przyjętego PDM ustalono w postaci:

$$r_{i+1} = \begin{cases} \gamma_{i+1} \cdot \left(p(s_i) \cdot p(s_i, s_{i+1})\right), & \text{gdy} \quad s_{i+1} \in \Psi \\ 0, & \text{gdy} \quad i = 0, \ p(s_i, s_i) = 0 \\ \gamma_{i+1} \cdot \left(1 - p(s_i) \cdot p(s_i, s_{i+1})\right), & \text{gdy} \quad s_{i+1} \notin \Psi \end{cases} \qquad (18)$$

Współczynnik proporcjonalności γ określony dla p' jako:

$$\gamma_{i+1} = \frac{1}{\left(2^3\right)^{p'}} \rangle 0, \quad p' = 0 \qquad (19)$$

wynika z konieczności dokonania podziałów wytypowanej oktantu dla stanu s_{i+1}.

7. Wnioski

Zaprezentowany w pracy sposób definiowania planera bezkolizyjnych ścieżek robotów przemysłowych opartego na zasadzie sterowania stochastycznego przedstawia ogólne podstawy nowej metody planowania. Jej istotnym elementem jest sposób zdefiniowania funkcji rozkładu prawdopodobieństwa przejścia do kolejnego punktu węzłowego P_{wi+1} w kroku t_{i+1} i prawdopodobieństwa $p(s_{i+1})$ pobytu PLANERA w stanie s_i. Wartość tych dwóch funkcji w istotny sposób zależy od przyjętej strategii działania PLANERA Ψ, która powinna umożliwić związanie postaci funkcji rozkładu prawdopodobieńst z przyjętym w procesie planowania kryterium optymalizacyjnym i może istotnie wzmocnić pożądaną efektywność działania zastosowanego procesu sterowania stochastycznego opartego na PDM.

Zgodnie z przyjętym mechanizmem oceniania działań PLANERA na jakość planowanych działań i ich skuteczność wpływać będzie również funkcja wag, która pozwala na modyfikację zdobywanej przez PLANER oceny zależnie od lokalnych warunków w jakich przebiegać będzie proces planowania. Ponieważ zdefiniowane funkcje $p(s_i)$ i $p(s_{i+1})$ są funkcjami o stałym rozkładzie, funkcja wag (1) pozwoli na związanie lokalnych uwarunkowań układu sceny z przyjętą strategią działania Ψ i omawianym PDM, poprzez uwzględnienie takich warunków ZP jak:

- optymalny poziom podziału oktalnego p i p' decydujący i szczegółowości dokonywanej alanilzy przestrzeni i procesu planowania, a co za tym idzie otrzymywanego drzewa decyzyjnego ZP,
- zagęszczenie przestrzeni sceny przeszkodami,
- bliskość prowadzonego procesu planowania względem omijanych przeszkód,
- wpływu dopuszczalnych w zadaniu strategii lokalnych na proces sterowania stochaastycznego działaniami PLANERA.

8. Literatura

[1] Sutton R.S., Barto A.G.: Reinforcement Learning: An Introduction. MIT Press 1998

[2] Latombe J.C.: Robot Motion Planning. Kluwer Academic Publisher, 1993.

[3] Kost G.G.: Planowanie bezkolizyjnych ścieżek stacjonarnych I manipulacyjnych robotów przemysłowych. Zeszyty Naukowe Politechniki Śląskiej, s. Mechanika z.148, Wydawnictwo Politechniki Śląskiej, Gliwice 2004.

[4] Dulęba I.: Metody i algorytmy planowania ruchu robotów mobilnych i manipulacyjnych. Akademicka Oficyna Wydawnicza EXIT, Warszawa 2001.

[5] Iosifescu M.: Skończone procesy Markowa i ich zastosowanie. PWN, Warszawa 1988.

Title: Robot's Motion Planning Based On The Stochastic Process Control

Artykuł recenzowany