

CONTROLADOR NEURAL NEBULOSO AUTÔNOMO DE DIREÇÃO E VELOCIDADE PARA ROBÔS MÓVEIS

PAULO R. CRESTANI^(a), MAURÍCIO F. FIGUEIREDO^(b), FERNANDO J. VON ZUBEN^(c)

Departamento de Engenharia de Computação e Automação Industrial,
Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas,
DCA/FEEC/Unicamp – Caixa Postal 6101, Campinas-SP, 13083-970
prcrest@dca.fee.unicamp.br ^(a), vonzuben@dca.fee.unicamp.br ^(c)

Departamento de Informática, Centro de Tecnologia, Universidade Estadual de Maringá,
DIN/UEM, Maringá-PR, 87020-900, mauricio@din.uem.br ^(b)

Resumo - O projeto de controladores autônomos para robôs móveis se apresenta como uma das tarefas mais atraentes e desafiadoras da Inteligência Artificial. Os esforços devotados à pesquisa científica sobre o tema refletem sua importância, em parte pelas aplicações práticas que proporciona (e.g. exploração de ambientes hostis ao homem, mecanismos de vigilância, transporte), em parte pela possibilidade de concepção de novas estratégias de aprendizado de propósito geral. Neste contexto, o presente trabalho propõe um modelo de controlador baseado em redes neurais nebulosas, que considera tanto a direção de deslocamento quanto a velocidade do robô como termos controláveis, enquanto que os modelos tradicionais se atêm ao controle da direção de deslocamento do robô. O controlador proposto exibe aprendizagem mediante interações com o ambiente e emprega comportamentos instintivos inatos para reflexos imediatos.

Abstract – The design of autonomous controllers for mobile robots is one of the most attractive and challenging tasks in Artificial Intelligence. The efforts devoted to scientific research in this theme reflects its importance, partially due to the practical applications involved (e.g. exploration of hostile environment, vigilance mechanisms, transport), and partially due to the possibility of designing general-purpose learning strategies. In this sense, this work proposes a fuzzy-neural-network-based controller that considers both, the direction and the velocity of navigation as controllable terms, while the traditional models deal only with the robot direction. The proposed controller exhibits learning by means of interactions with the environment and applies instinctive behaviors for instant reactions.

1. Introdução

O problema de navegação autônoma de robôs móveis a ser considerado neste trabalho pode ser definido como uma busca de alvos em um ambiente de navegação contendo uma série de obstáculos, os quais podem ser fixos ou móveis, podendo inclusive delimitar o espaço físico destinado à navegação. Os alvos podem representar uma presa a ser capturada ou um local onde alguma tarefa deva ser executada. O papel do controlador autônomo do robô é encontrar uma trajetória que o leve de um ponto inicial a um alvo, evitando colisões contra os obstáculos. Muitos trabalhos incorporam variações desta tarefa básica, proporcionando diferentes enfoques de investigação, tal como acontece nas pesquisas envolvendo múltiplos robôs, obstáculos móveis e múltiplos alvos. Dentre os aspectos mais interessantes a

serem considerados para a análise de desempenho está custo da trajetória em termos da distância percorrida. Entretanto, modelos que se preocupam excessivamente com o custo do percurso (Sebastian, 2001) geralmente acabam produzindo controladores dedicados a um único tipo de ambiente de navegação, apresentando dificuldades de adaptação a ambientes com características distintas.

Dentre os modelos de navegação autônoma propostos na literatura, é comum assumir que a velocidade seja um parâmetro constante (Sebastian, 2001; Iwakoshi et al. 1998; Jerbic et al., 1999; Jim & Giles, 2001), não sendo levada em consideração pelo mecanismo de controle. Nestes modelos, apenas a direção de deslocamento do robô é manipulada. Embora alguns modelos apresentem mecanismos de variação de velocidade (Nolfi & Floreano, 1999), frequente-

mente ela não é enfatizada como um parâmetro central do desempenho da navegação do robô, sendo utilizada apenas como um meio de controle direcional por meio da aplicação individual das rodas do robô.

Floreano & Mondada (1996) descrevem um modelo de sistema de controle para robôs móveis baseado em uma rede neural treinada por um algoritmo genético, sendo a velocidade de deslocamento do robô um dos parâmetros da função de aptidão. Embora o robô aprenda a controlar sua velocidade, o treinamento evolutivo é lento, de acordo com os resultados descritos. Além disso, nem mesmo os indivíduos mais adaptados conseguem atingir velocidades próximas da máxima permitida.

Diversas vantagens podem ser enumeradas em um modelo que considera o papel da velocidade no problema da navegação autônoma: maior flexibilidade nas manobras de desvio de obstáculos com a diminuição da velocidade, o que permite uma navegação mais segura; melhor aproveitamento de condições favoráveis à aproximação do alvo, como trechos do percurso sem obstáculos, permitindo um aumento da velocidade; e possibilidade de controlar a aproximação do alvo. Neste sentido, Kubota et al. (2001) apresentam um modelo interessante baseado em algoritmos genéticos e aprendizagem de regras nebulosas.

O modelo de controlador autônomo para robôs móveis que será descrito a seguir baseia-se em uma arquitetura modular e hierárquica de redes neurais nebulosas, concebida para agir tanto sobre a direção de deslocamento do robô quanto sobre sua velocidade de navegação. Entretanto, nenhum conhecimento inicial sobre manipulação coordenada da velocidade e da direção de deslocamento é embutido no robô. Este conhecimento será adquirido automaticamente ao longo da navegação, de forma que o robô melhore seu desempenho enquanto interage com o ambiente.

O restante do texto está organizado da seguinte maneira. Na seção 2, características essenciais do robô a ser modelado são descritas. Uma visão geral da arquitetura do controlador autônomo será apresentada na seção 3. Uma perspectiva mais focada na arquitetura da rede neural de coordenação do controlador, principal componente do sistema de navegação, é considerada na seção 4. Na seção 5, os mecanismos de aprendizagem deste módulo são apre-

sentados. Na seção 6, diversos testes realizados em um ambiente de simulação computacional são analisados e comentados. Por fim, na seção 7, algumas conclusões pertinentes ao trabalho são apresentadas.

2. O robô móvel

Para a concepção do controlador autônomo, assume-se um robô móvel dotado de 37 sensores de obstáculos, 36 sensores de alvos, um sensor de colisão, um sensor de captura de alvos e um sensor de velocidade. Os sensores de obstáculos são distribuídos ao longo da parte frontal do robô e detectam a distância ao obstáculo na sua direção. Os sensores de alvos correspondem a um conjunto de antenas posicionadas em torno do robô que determinam o ângulo de desvio entre a direção de deslocamento do robô e a direção do alvo no ambiente, o qual deve possuir um emissor de ondas eletromagnéticas anexado a si. A existência de múltiplos sensores de obstáculos e de alvos permite a formação de padrões de sensoria-mento unidimensionais, os quais fornecem informações pontuais (relativo a cada sensor) e também relacionais (relativo ao vetor ordenado de sensores).

A cada passo da navegação, após a leitura dos sensores, é realizado um ajuste na direção e velocidade do robô, o que caracteriza o controlador como sendo reativo. Tanto o ambiente de navegação quanto o robô são parte de um ambiente de simulação computacional desenvolvido para o projeto, e respeitam o maior elenco possível de restrições encontradas em uma implementação prática, embora o efeito do atrito do terreno contra as rodas não seja considerado.

3. A arquitetura do controlador autônomo

O controlador neural desenvolvido baseia-se no modelo proposto por Figueiredo (1997), em que dois grupos de redes neurais nebulosas hierarquicamente estabelecidos respondem pelo comportamento do robô. Este modelo será estendido para atuar também sobre a velocidade de deslocamento.

A Fig. 1 apresenta o controlador em forma de módulos, representando redes neurais nebulosas. A partir das leituras dos sensores de obstáculos e alvos, o módulo de desvio de obstáculos (MDO) e o módulo de busca de alvos (MBA) geram sinais independentes (U_{DO} e U_{BA}) de ajuste na direção do robô. Além disto, a partir da leitura do sensor de velocidade

de, o módulo de decremento de velocidade (MDV) e o módulo de incremento de velocidade (MIV) geram sinais independentes (U_{DV} e U_{IV}) de ajuste da velocidade atual do robô. Estes módulos, respondem por comportamentos inatos e podem ser representados por redes neurais treinadas através de algoritmos genéticos (Crestani & Von Zuben; 2001). Não obstante as vantagens de um modelo conexionista, os módulos instintivos também poderiam ser descritos a partir de um conjunto reduzido de regras de inferência do tipo: Se <antecedente> então <conseqüente>.

Em diversas situações, os sinais destes módulos instintivos são conflitantes, de forma que um Módulo de Coordenação (MC), estabelecido em um nível hierárquico superior, pondera tais sinais antes que eles estimulem os neurônios motores do controlador.

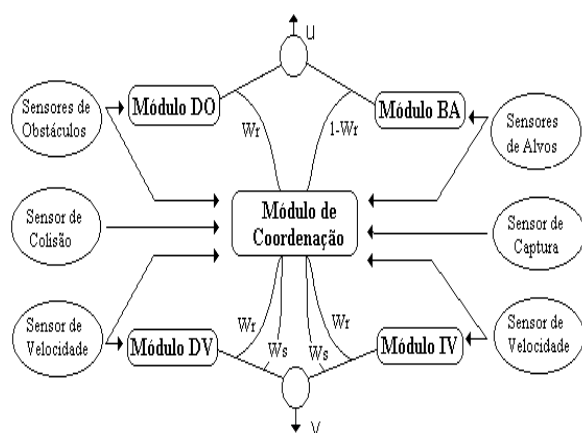


Figura 1 – Arquitetura do controlador

Os dois neurônios motores do robô recebem os sinais dos módulos instintivos, já devidamente ponderados pelo MC, e agem diretamente sobre o robô, fornecendo um sinal referente à nova direção u , conforme Eq.(1) e um sinal referente à nova velocidade v , conforme Eq.(2).

$$u = (1 - Wr) \cdot U_{BA} + (1 - Ws) \cdot Wr \cdot U_{DO} \quad (1)$$

$$v = U_{IV} \cdot (1 - \max(Wr, Ws)) + U_{DV} \cdot (\max(Wr, Ws)) \quad (2)$$

4. Arquitetura do Módulo de Coordenação

A Fig. 2 apresenta a arquitetura do MC. De acordo com o modelo, as entradas x_i (sensores) e os pesos sinápticos w_i , $i = 1, \dots, N$, são parâmetros de conjuntos nebulosos $\Phi_{x_i}(x)$ e $\Omega_{w_i}(x)$, respectivamente, e as entradas x'_l e w'_l $l = 1, \dots, N_A$ são parâmetros dos conjuntos nebulosos $\Phi'_{x_l}(x)$ e $\Omega'_l(w_l(x))$, respectivamente, todos os quais possuem perfis triangulares (isósceles) e largura da base fixa e igual a 3

nos resultados apresentados por Figueiredo (1997). Durante o processamento sináptico, estes conjuntos nebulosos são comparados pelo operador monotônico nebuloso de medida de possibilidade $\Delta(\cdot)$, e o valor encontrado na operação é ponderado pelo peso sináptico s_i através da t-norma produto $T_2(\cdot)$. Os pesos sinápticos w_i e w'_l representam memórias com relação aos padrões induzidos pelas respectivas entradas, enquanto que s_i e s'_l representam reforços à modificação das entradas pelas sinapses. Após o processamento sináptico, um operador de agregação nebuloso $A(\cdot)$, que corresponde à s-norma máximo, agrega as entradas e a função de ativação identidade $f(x)=x$ determina a saída do j -ésimo neurônio da rede, donde seguem y_j , conforme a Eq. (3) e y'_j , conforme a Eq. (4). Da agregação dos sinais y_j e y'_l por meio do operador $\max = \bigvee_{j=1}^Q (\cdot)$, que é a s-norma máximo,

procedem os sinais Wr e Ws respectivamente, onde Q é o número de neurônios da rede.

Os sensores de colisão e de captura são ocultados na Fig. 2, pois sua função é apenas disparar os processos de aprendizagem do MC, desencadeados em duas situações específicas: após uma colisão ou quando um alvo é atingido.

$$y_j = f_1 \left(\left(\bigwedge_{i=1}^N A(T_2(\Delta(\Omega_{x_i}(x), \Phi_{w_i}(x)), s_i)) + T_2(\Delta(\Omega_{x_{(i)N+1}}(x), \Phi_{w_{(N+1)}}(x)), s_{(N+1)})) \right) / 2 \right) \quad (3)$$

$$y'_j = f_2 \left(\left(\bigwedge_{l=1}^{N_A} A(T_2(\Delta(\Omega_{x'_l}(x), \Phi_{w'_l}(x)), s'_l)) + T_2(\Delta(\Omega_{x'_{(N+1)}}(x), \Phi_{w'_{(N+1)}}(x)), s_{(N+1)})) \right) / 2 \right) \quad (4)$$

5. Aprendizagem do Módulo de Coordenação

Diferentemente dos módulos instintivos, o MC não é uma estrutura inata do robô. Desta forma, enquanto aqueles estão prontos para serem utilizados, antes mesmo do início da navegação, este ainda precisa sofrer um processo de aprendizagem por reforço, baseado na interação do robô com o ambiente. Além disso, o MC corresponde a uma rede neural construtiva, de forma que a cada alvo atingido e a cada colisão um novo neurônio é agregado à sua estrutura, sendo seu papel estabelecido no sentido de incorporar o conhecimento resultante da interação.

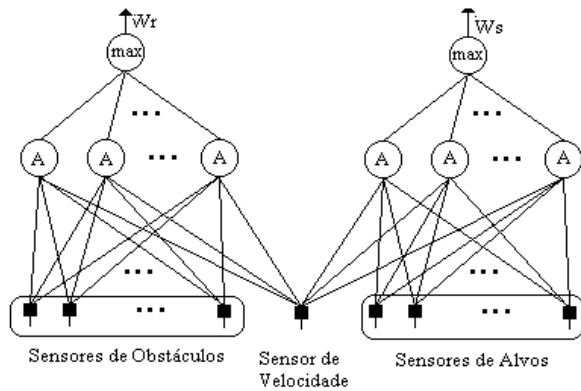


Figura 2 – Arquitetura do Módulo de Coordenação

O processo de aprendizagem do MC obedece à teoria do condicionamento operante, observada em sistemas biológicos (Baum, 1999). Quando há uma colisão, o robô instintivamente reage, mudando sua direção em 180°. Ao mesmo tempo, o padrão unidimensional de distância a obstáculos e a velocidade do robô no momento da colisão são registrados por um neurônio da rede. A correlação temporal destes padrões de aproximação (ação) com a situação de colisão (consequência), dispara um processo de punição. Caso este processo seja suficientemente e convenientemente repetido, o sistema termina por associar a ação à consequência, desenvolvendo um comportamento de aversão a obstáculos. Um processo análogo ocorre quando um alvo é atingido. Agora, há um reforço sobre a relação entre navegar na direção do alvo, aproximando-se deste e a satisfação dos objetivos do sistema (atingir o alvo).

Além da criação de um novo neurônio, o processo de aprendizagem da rede também prevê o ajuste sináptico dos neurônios já criados de forma a se conformarem à nova situação apresentada. As equações básicas de ajuste do i -ésimo peso ligado ao j -ésimo neurônio são dadas (Eq. (5) e Eq. (6)), e valem para todos os neurônios da primeira camada do MC.

$$w_{ij}(t+1) = w_{ij}(t) + (V_{COMP}(t) - w_{ij}(t)) / (1 + s_{ij}(t)) \quad (5)$$

$$s_{ij}(t+1) = s_{ij}(t) + (SIMIL - s_{ij}(t)) / (\alpha + 1) \quad (6)$$

O termo V_{COMP} corresponde à entrada sensorial mais similar ao peso w_{ij} . $SIMIL$ resulta da aplicação do operador monotônico $\Delta(\cdot)$ a esta entrada, e α é o fator de aprendizagem, cujo valor varia de acordo com o número e a ordem dos neurônios.

6. Simulações e Testes

No ambiente empregado para a realização das simulações, os obstáculos são representados por polígonos pretos, os alvos por cruzes, as regiões de colisão por pontos, os robôs por triângulos e as trajetórias por linhas pretas. Além disto, quando há mais de um alvo em um ambiente, apenas um é apresentado por vez ao robô, sendo que apenas após este ter sido atingido um novo alvo será apresentado.

Na primeira simulação (Fig. 3), o robô é colocado em um ambiente simples, contendo apenas um alvo. No início da simulação, o MC ainda não havia sido treinado e o robô possuía apenas o comportamento inato de busca pelo alvo. Seis colisões são observadas contra o obstáculo colocado entre o robô e o alvo, até que o sistema autônomo aprenda uma estratégia eficaz de coordenação, desviando-se do obstáculo e atingindo o alvo. Neste momento o MC conta com 7 neurônios (6 na estrutura de controle de desvio e 1 na estrutura de controle de busca).

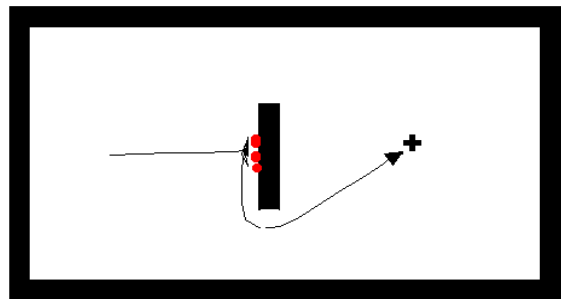


Figura 3 – Início do processo de aprendizagem

Para demonstrar como o robô se utiliza da variação de sua velocidade de deslocamento para melhorar suas manobras de desvio e aproximar-se do alvo com segurança, considerando-se um controlador cujo MC esteja treinado, a próxima simulação (Fig. 4) apresenta as velocidades desenvolvidas pelo robô ao longo da navegação. A velocidade mínima permitida é de 0,1 ud/i (unidade de deslocamento por iteração) e a velocidade máxima é de 1,0 u.d./i. Nesta simulação, é possível observar como o robô diminui sua velocidade ao aproximar-se de um obstáculo ou ao percorrer um corredor, ao passo que, havendo espaço livre, o robô chega a atingir 0,8 u.d./i., ficando próximo da velocidade máxima. Ao aproximar-se do alvo, o robô também diminui sua velocidade.

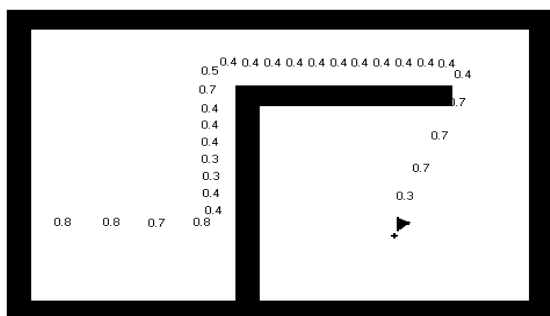


Figura 4 – Registro da velocidade desenvolvida.

A Fig. 5 apresenta uma simulação realizada em um ambiente semelhante a um conjunto de salas e corredores, sendo os alvos estrategicamente posicionados de forma que o robô deve percorrer praticamente todo o ambiente. Nenhuma colisão ocorreu e o robô conseguiu atingir todos os alvos. Como o ambiente possui um número muito grande de paredes, a velocidade média do robô foi de 0,6 u.d./i. Para ser bem sucedido em um ambiente complexo como este, o MC deve apresentar um número considerável de neurônios. No caso apresentado, este número era de 30. Assim, a aprendizagem ilustrada na Fig. 3 representa apenas uma etapa do processo. Além disso, a navegação do robô não é restrita ao ambiente utilizado durante o treinamento.

No modelo de Figueiredo (1997), onde a velocidade não era levada em consideração pelo controlador, o número de neurônios necessários para a tarefa de coordenação era menor. Entretanto, o aumento do tamanho do MC, e do período de treinamento, tem como benefícios um controle mais refinado, e melhor desempenho durante a navegação. Além disso, apesar do tempo necessário para o processo de aprendizagem no novo sistema refletir a progressiva diminuição da velocidade do robô mediante as ocorrências de colisões, o MC passa a agir sobre os atuadores já após a primeira colisão, o que confere ao sistema uma resposta rápida ao problema.

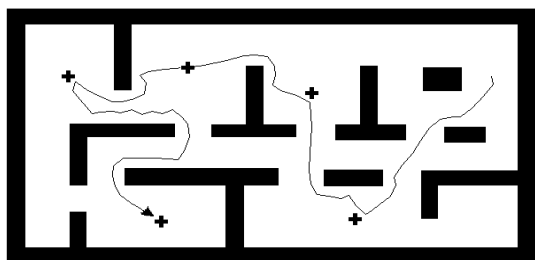


Figura 5 – Navegação em ambiente complexo.

A Fig. 6 compara os desempenhos dos modelos sem controle de velocidade (velocidade igual a 0,2 ud./i) em (a) e com controle de velocidade em (b). Em nenhum dos casos, houve colisões. Porém, além de desenvolver uma trajetória mais curta, o robô com controle de velocidade necessitou de menos tempo (1 minuto e 39 segundos) para completar o percurso que o robô sem este controle (5 minutos e 57 segundos), sob as mesmas condições de operação.

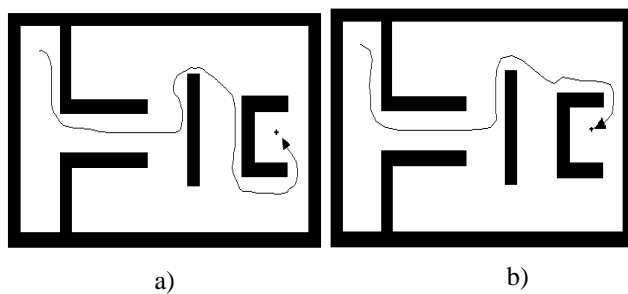


Figura 6 – Comparação entre modelos

A Fig. 7 apresenta uma simulação em que quatro robôs, todos com controle de velocidade e previamente treinados em conjunto foram colocados em um mesmo ambiente de navegação. Embora submetidos às mesmas condições de treinamento, o número de neurônios no MC de cada robô ao final do processo foi diferente, variando de 16 a 22. Esta diferença pode estar ligada a questões tão sutis quanto à posição de um robô com relação aos demais no momento em que um novo alvo é apresentado, mas também pode refletir colisões entre dois robôs, pois cada robô representa um obstáculo aos demais.

Simulações como a apresentada na Fig. 7, se tornam mais interessantes quando os robôs são capazes de controlar suas velocidades, pois caso contrário, sua habilidade de manobras é menor, acarretando um grande número de colisões entre os robôs, o que complica a navegação. Alguns resultados sobre emergência de comportamentos coletivos em agentes autônomos inteligentes podem ser encontrados em Karsai (1999) e em Asada et al. (1999). Holland & Melhuish (1999) estudam os efeitos da sinergia em agentes robóticos simples.

Em Jim & Giles (2001), um modelo de comunicação explícita entre os agentes é proposto. Os robôs apresentados na Fig. 7, ao contrário, não foram programados para apresentarem comunicação ou comportamentos de colaboração mútua, mas há indícios de que o processo de aprendizagem se torne mais

eficaz quando diversos robôs são treinados simultaneamente em um mesmo ambiente. Enquanto navegam juntos, colidem repetidas vezes entre si, acelerando o processo de treinamento. Além disso, como se movimentam na direção do alvo, acabam exercendo o papel de guias uns dos outros. Embora sejam competidores na busca pelo alvo, colaboram indiretamente para se manterem no caminho certo.

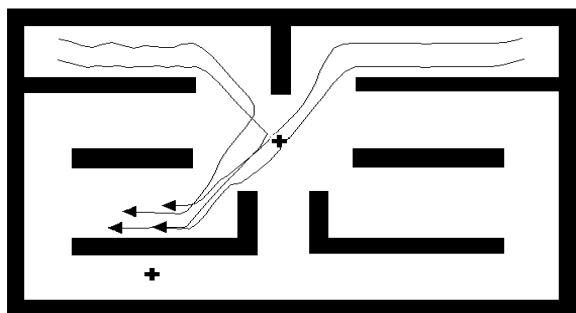


Figura 7 – Robôs navegando em formação.

7. Conclusões

No presente artigo, apresentou-se uma arquitetura de controlador autônomo para navegação de robôs móveis baseada em redes neurais nebulosas hierarquicamente estruturadas. Além da direção de deslocamento, o ajuste da velocidade do robô também passa a ser tarefa do controlador autônomo, no modelo apresentado. As vantagens desta extensão foram apresentadas ao longo do texto. Por fim, foi mostrado que a realização de simulações com múltiplos robôs em um mesmo ambiente abre espaço para novos estudos relacionados a comportamentos coletivos de agentes inteligentes artificiais. Este novo modelo pressupõe, no entanto, maior disponibilidade de recursos computacionais, tais como memória e processamento, quando comparado a modelos de controle de navegação mais simples.

Agradecimento

Os resultados apresentados neste artigo são decorrentes da execução de projetos financiados pela FAPESP (processo 00/04991-1) e pelo CNPq (processos 300910/96-7 e 300982/97-3).

Referências Bibliográficas

- Asada, M., Uchibe, E., Hosoda, K., 1999; Cooperative behavior acquisition for mobile robots in dynamically changing real worlds via vision-based reinforcement learning and development, *Artificial Intelligence*, Vol. 110, pp. 275 – 292
- Baum, W. M., 1999; *Compreendendo o behaviorismo: ciência, comportamento e cultura*; pp. 69-89, Editora Artes Médicas Sul; Porto Alegre.
- Crestani, P. R. C. & Von Zuben, F.; 2001; *Evoluindo comportamentos inatos em robôs móveis*; Relatório Técnico; RT-DCA 02/2001, DCA/FEEC/Unicamp; Campinas-SP;
- Figueiredo, M. F., 1997; *Redes neurais aplicadas em problemas de modelagem e controle autônomo*; Tese de Doutorado, FEEC/Unicamp, Campinas-SP.
- Floreano, D. & Mondada, F., 1996; Evolution of homing navigation in a real mobile robot; *IEEE Transactions on Systems, Man and Cybernetics*, Vol XX, NO. Y., pp. 1-13.
- Holland, O. & Melhuish, C.; 1999; Stigmergy, self-organization, and sorting in collective robotics; *Artificial Life*; Vol. 5, Num. 2; pp.173-202; MIT Press.
- Iwakoshi, Y., Furuhashi, T, Uchikawa, Y., 1998; A fuzzy classifier system for evolutionary learning of robot behaviors"; *Applied Mathematics and Computation*, Vol. 91, pp. 73 – 81.
- Jerbic, B, Grolinger, K, Vranjes, B, 1999; Autonomous agent based on reinforcement learning and adaptive shadowed network; *Artificial Intelligence in Engineering*; Vol. 13, pp. 141 – 157.
- Jim, Kam-Chuen & Giles, C. L.; 2001; Talking helps: evolving communicating agents for the predator-prey pursuit problem; *Artificial Life*; Vol. 6, Num. 3; pp. 237-254; MIT Press.
- Karsai, I.; 1999; Decentralized control of construction behavior in paper wasps: an overview of the stigmergy approach"; *Artificial Life*; Vol. 5, Núm. 2; pp.117-136; MIT Press.
- Kubota, N., Morioka, T, Kojima, F., Fukuda, T., 2001; Learning of mobile robots using perception-based genetic algorithm"; *Measurement*, Vol. 29, pp. 237 – 248.
- Nolfi, S. & Floreano, D.; 1999; Coevolving predator and prey robots: do 'arms races' arise in artificial evolution?; *Artificial Life*; Vol. 4, Num. 4; pp.311-335; MIT Press.
- Sebastian, T., Fox D., Burgard, W., Dellaert, F. ; 2001, Robust Monte Carlo localization for mobile robots, *Artificial Intelligence*, Vol. 138, pp. 99 – 141.