

Um Sistema de Visão Computacional Baseado em Cores Aplicado ao Controle de um Robô Móvel

M. G. Quiles, Mestrando, *ICMC-USP*, e R. A. F. Romero, Professora Associada, *ICMC-USP*

Abstract—Many works in last few years have developed vision system to control robots, but a lot of them don't deal with color processing, which, in some cases, can bring more information than gray level images. Some artificial intelligent techniques like neural networks have commonly been applied to develop systems based on visual information. This paper presents a vision color system based on neural networks applied to a robot control. This system enables the robot to follow objects based on their colors.

Palavras chaves—Redes Neurais Artificiais, RGB, Robôs Móveis, Segmentação de Imagens Coloridas, Visão Computacional.

I. INTRODUÇÃO

Há várias décadas, os pesquisadores tentam construir máquinas (robôs) para auxiliar ou substituir os seres humanos em diversas tarefas que envolvam riscos e em ambientes hostis ou onde a rapidez e a precisão de determinados movimentos são essenciais. Somente nas últimas décadas, com o grande crescimento dos recursos computacionais disponíveis, os pesquisadores têm conseguido desenvolver robôs capazes de executar tarefas mais complexas que as executadas pelos primeiros robôs estritamente mecânicos, comumente instaladas em chão de fábricas.

Porém, em muitas aplicações reais, os robôs fixos não se mostram adequados. A simples tarefa de entregar documentos entre diversos locais de um prédio é um exemplo disto. Assim, o desenvolvimento de robôs móveis é de extrema importância tanto para essas pequenas tarefas como para objetivos maiores, como é o caso da exploração espacial ou exploração marítima, que são ambientes desconhecidos e que podem trazer riscos à vida do ser humano.

Diversos mecanismos para o controle de robôs móveis têm sido utilizados, como sonares, laser e visão. Este trabalho focalizará o uso de redes neurais artificiais (RNAs) aplicadas no desenvolvimento de um sistema de visão para o controle de um robô.

Vários trabalhos utilizando visão no controle de robôs já foram desenvolvidos. Em Pomerleau [6], um sistema de controle baseado em visão utilizando redes neurais foi construído com o objetivo de guiar um veículo autonomamente sobre uma pista. Outros sistemas, também utilizando redes neurais, foram desenvolvidos em [1], [4] e [10]. No trabalho proposto por Blanck [1], foi desenvolvido um sistema para capacitar o robô a seguir uma bola sobre uma superfície. Em [4], o robô foi treinado para caminhar em um corredor desviando dos possíveis obstáculos. Tyler [10]

apresenta um sistema de visão que tem como objetivo indicar qual a parte da bola e do gol que são visíveis na imagem, de tal forma que um sistema de controle possa utilizar essas informações na tomada de decisão do robô. Um outro trabalho utilizando visão e RNAs foi desenvolvido em [11] em que uma interface para o controle de um robô baseado em gestos foi proposta. Neste trabalho, uma câmera é responsável pela captura de gestos produzidos por um humano posicionado em frente do robô. Com base nestes gestos o robô obedece a comandos como: siga-me, pare, etc. Medeiros [5], utilizou uma rede neural para capacitar um robô do tipo Pioneer I a percorrer sobre um caminho delimitado por duas faixas laterais brancas.

No entanto, em todos os trabalhos citados não é realizado um tratamento de cores no sistema de visão. O tratamento de cores pode ser de grande importância em diversos domínios. Por exemplo, para o desenvolvimento de controladores relacionados a futebol de robôs, a análise de cores pode ser um fator decisivo [7], [8].

No desenvolvimento de um sistema de visão com tratamento de cores, a utilização de um módulo de segmentação pode ser essencial a fim de destacar o objeto de interesse no quadro que está sendo analisado, facilitando ao sistema a identificação do mesmo. A segmentação é um dos primeiros passos em análise de imagens e de reconhecimento de padrões, sendo considerada uma componente essencial nos sistemas de reconhecimento de padrões, determinando a qualidade final dos resultados do processo de reconhecimento [2].

Em Simões [7], [8] foi desenvolvido um método de segmentação de imagens coloridas baseado em classificação de cores. Para o processo de classificação de cores, uma rede neural do tipo MLP (Multi-Layer Perceptron [3]) foi treinada para reconhecer as cores que representam o domínio de futebol de robôs. Desta forma, o processo de segmentação consiste na apresentação de cada pixel presente na imagem à rede neural que informa a qual classe de cor esse pixel pertence. A rede neural foi configurada com 3 elementos na camada de entrada (entrada RGB) e $n+1$ elementos na camada de saída, em que n representa o número de cores significativas e um elemento na camada de saída é responsável por classificar as cores que não são significativas para o domínio da aplicação. Simões [9] apresentou um estudo da metodologia desenvolvida em [7], [8] para três espaços de cores: HSV, YUV e RGB, sendo que o espaço de cor RGB se mostrou mais adequado ao domínio do futebol de robôs.

O presente artigo objetiva a construção de um sistema de controle baseado em visão aplicado a um robô perseguidor de objetos determinados pela cor.

Para o desenvolvimento deste sistema dois módulos principais foram implementados: Um módulo de visão e um módulo de controle.

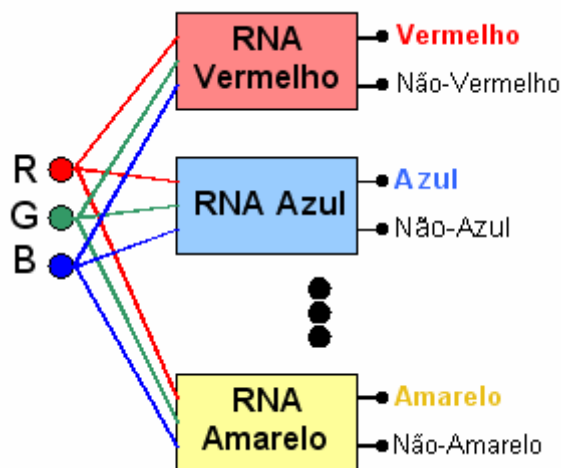


Figura 1. Sistema de classificação de cores composto por um conjunto de redes neurais do tipo MLP.

O módulo de visão é responsável pelo pré-processamento, segmentação e identificação do posicionamento do objeto no campo visual. Para a tarefa de segmentação, um sistema de classificação de cores semelhante ao apresentado em [7], [8] é utilizado. No entanto, foram utilizadas diversas redes neurais ao invés de apenas uma no processo de classificação de cores (Figura 1) e, além disso, tanto as imagens coletadas para a realização do treinamento quanto as imagens coletadas para análise em tempo real são obtidas através de uma câmera acoplada no próprio robô, ao contrário do cenário alvo de [7], [8] em que a câmera estava posicionada no teto do campo de futebol de robôs.

Este artigo está organizado do seguinte modo. Na seção 2 é apresentado o sistema desenvolvido o qual é constituído pelos módulos de visão e de controle. Na seção 3, são descritos os experimentos realizados bem como os resultados obtidos. Finalmente, na seção 4, são apresentados a conclusão e os trabalhos futuros possíveis.

II. DESCRIÇÃO DO SISTEMA

O sistema proposto por este artigo é composto por dois módulos: o primeiro responsável pelo pré-processamento, segmentação e reconhecimento das imagens e o segundo módulo implementa o controlador responsável por capacitar o robô a perseguir a bola pelo chão de uma sala.

Esses dois módulos são descritos a seguir.

A. Módulo de Visão

Como relatado anteriormente, este módulo é dividido em algumas fases: pré-processamento (redução) da imagem, segmentação e o reconhecimento das imagens. A primeira fase é responsável pela captura da imagem a partir da câmera a

uma resolução de 320x240 pixels e pela redução da mesma à resolução de 80x60 pixels. Para a redução da imagem foi realizada a média dos pixels.

O processo de segmentação de imagens é realizado por um sistema de classificação de cores. O sistema classificador é composto por diversas redes neurais (Figura 1), no qual cada rede tem por objetivo classificar uma determinada cor, separando-a das outras cores presentes na imagem. Por exemplo, quando a cor vermelha for solicitada, cada pixel da imagem é classificado em vermelho e não vermelho (demais cores). A classificação é feita com base nos componentes RGB de cada pixel da imagem. Neste caso, a rede neural solicitada recebe os três componentes (RGB) e apresenta na saída da rede o resultado de classificação que pode assumir dois valores: classificado como a cor selecionada ou classificado como demais cores (Ex: cor vermelha e cores não vermelhas).



Figura 2. Resultado do processo de segmentação por classificação de cores utilizando a representação de cores RGB.

Como o objetivo do processo de segmentação implementado neste trabalho é de isolar a cor de interesse de um determinado objeto, a imagem segmentada resultante é composta por apenas duas cores: cor de interesse (representada pela cor branca) e demais cores (representada pela cor preta), de forma semelhante ao apresentado em [7], [8] considerando a cor azul.

Na Figura 2 é mostrada a imagem resultante obtida através do processo de segmentação, na qual observa-se que apenas os pixels que representam o objeto de interesse (cor de interesse), aparecem como branco na imagem segmentada, e

todos os demais pixels presentes na imagem são convertidos em preto.

Por exemplo, quando o robô for configurado para perseguir um objeto vermelho, cada pixel da imagem é submetido à rede neural responsável pela classificação da cor vermelha, sendo as demais redes neurais desconsideradas.

Após o processo de segmentação, uma outra rede neural MLP é utilizada no reconhecimento da posição do objeto na imagem. Duas abordagens foram utilizadas no sistema de reconhecimento.

Na primeira abordagem foi utilizada uma rede neural com 14 saídas onde as primeiras 13 saídas representam respectivamente cada um dos 13 blocos presentes na imagem como é mostrado na Figura 3 e a última saída indica se o objeto está ou não presente na imagem que está sendo analisada. Os 13 blocos que correspondem à imagem foram obtidos empiricamente com base em análises realizadas em imagens contendo os objetos.

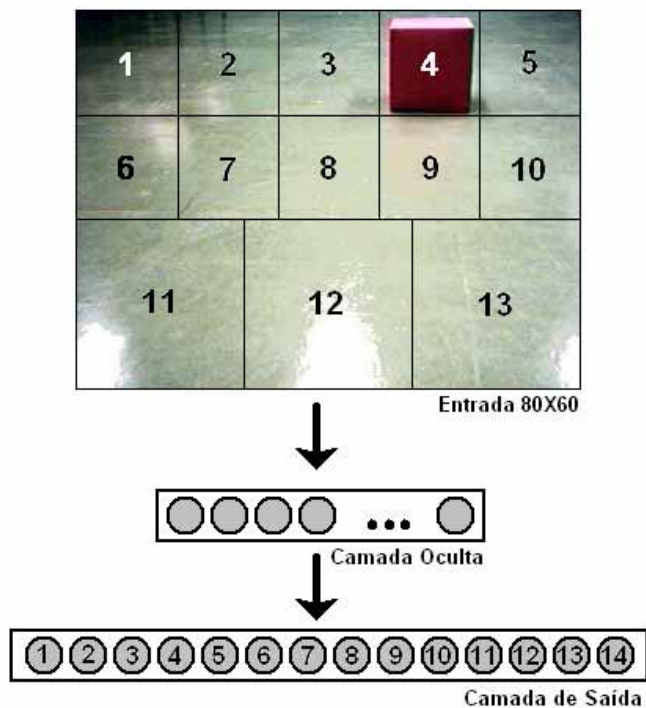


Figura 3. Rede neural MLP utilizada no processo de reconhecimento da posição do objeto implementada na primeira abordagem.

A segunda abordagem utilizada consiste na implementação de uma rede neural com apenas 3 saídas, em que as duas primeiras representam as respectivas posições X e Y do objeto na imagem e a terceira saída, P, indica se o objeto está ou não presente na imagem (Figura 4). Vale ressaltar que as posições X e Y não são referentes ao objeto no espaço real (3D), mas sim as posições X e Y do centro do objeto no quadro capturado.

B. Módulo de Controle

Com base nos dados obtidos a partir do módulo de visão, dois controladores baseados em comportamento (Behaviors) reativo [13]-[14] foram implementados a fim de capacitar o robô a perseguir o objeto pelo chão de uma sala.

O primeiro controlador, desenvolvido com base no sistema de visão implementado na primeira abordagem, é composto por dois comportamentos:

SE objeto está presente no quadrante Q

ENTÃO

Vel. de Translação = V

Vel. de Rotação = R

onde Q representa o quadrante onde o objeto está presente na imagem (saídas da rede neural: 1-13) (Figura 3), V é a velocidade de translação do Robô e R é a sua velocidade de rotação. A velocidade de translação no robô é configurada por alta, média e baixa. A velocidade de rotação é definida como “vire fortemente para a esquerda ou para a direita” ou “vire moderadamente para a esquerda ou para a direita” ou igual a zero, de acordo com o posicionamento do objeto:

- Translação Alta: quadrantes 1-5;
- Translação Média: quadrantes 6-10;
- Translação Baixa: quadrantes 11-13;
- Rotação Esq. Forte: quadrantes: 1 e 6;
- Rotação Esq. Moderada: quadrantes: 2, 7 e 11;
- Rotação Dir. Moderada: quadrantes: 4, 9 e 13;
- Rotação Dir. Forte: quadrantes: 5 e 10;
- Rotação Zero: quadrantes 3, 8 e 12.

E o segundo comportamento, toma a decisão de parar o robô quando o objeto não está presente no campo visual do robô:

SE objeto não está visível (Saída 14=0)

ENTÃO Pare

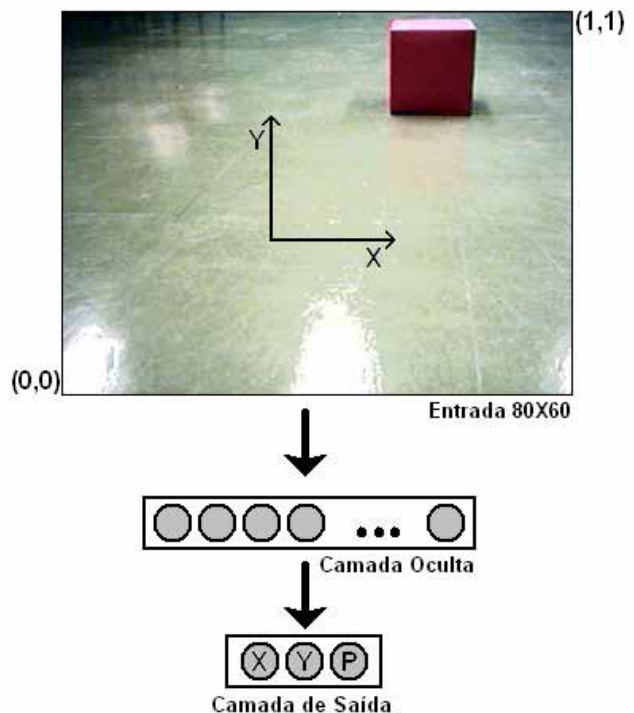


Figura 4. Rede neural MLP utilizada no processo de reconhecimento da posição do objeto implementada na segunda abordagem.

O segundo controlador foi implementado com base nos dados recebidos do sistema de visão desenvolvido utilizando a

segunda abordagem, na qual apenas 3 saídas na rede neural compõe a saída do sistema de visão (Figura 4). Diferenciando do controlador apresentado anteriormente, este recebe parâmetros não binários (X e Y) permitindo que diversos níveis de velocidade sejam configurados no robô. Dois comportamentos foram implementados neste controlador:

SE objeto está presente na imagem ($P=1$)

ENTÃO

Vel. de Translação = $C1*Y$

Vel. de Rotação = $C2*(X-0,5)$

SE objeto não está visível ($P=0$)

ENTÃO Pare

onde C1 e C2 são constantes empíricas que definem a velocidade de translação e rotação máximas, respectivamente, e X, Y e P são as saídas do módulo de visão.

III. EXPERIMENTOS REALIZADOS

Foram realizados experimentos em tempo-real com um robô móvel Pioneer I da ActivMedia acoplado com uma câmera WebCam da Creative. A comunicação entre o robô e o sistema desenvolvido foi realizada utilizando-se o simulador Saphira [12].

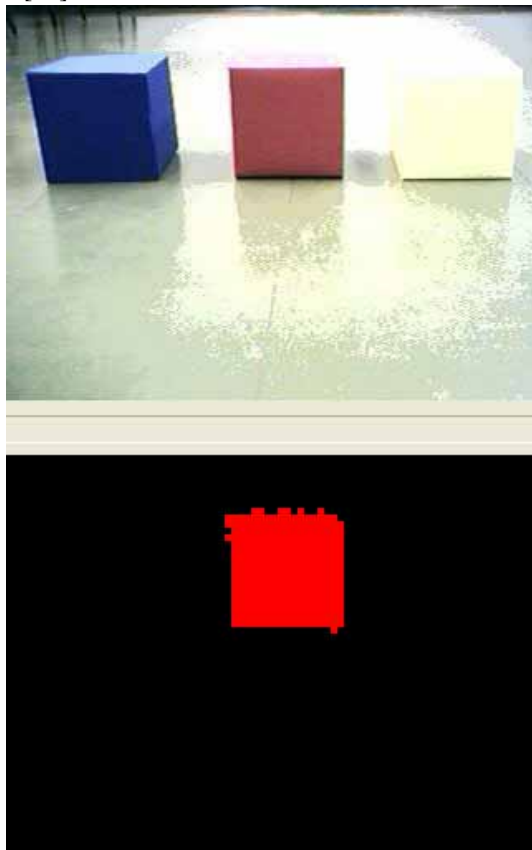


Figura 5. Resultado do processo de segmentação para a cor vermelha.

Diversos experimentos foram realizados. Os primeiros experimentos, sem a utilização do robô, foram feitos com o objetivo de testar a qualidade do processo de segmentação para as imagens capturadas pela câmera. A princípio foram

utilizadas apenas duas cores: azul e vermelho. Nas Figuras 5 e 6 são apresentados os resultados do processo de segmentação para as duas cores, Vermelha e Azul, respectivamente.

Após os ajustes do sistema de segmentação, como topologia da rede, número de neurônios nas camadas intermediárias, foram realizados alguns experimentos com o objetivo de encontrar uma topologia adequada para o sistema de reconhecimento da posição dos objetos nas imagens segmentadas.

Assim que o módulo de visão foi totalmente implementado e configurado foram realizados testes em tempo real com o robô utilizando as duas abordagens propostas.

Para realização dos testes, um humano conduzia o objeto de cor determinada pela sala, observando o comportamento do robô que, em todos os casos, foi capaz de perseguir o objeto pelo chão de uma sala atuando em tempo real.

Com a utilização da segunda abordagem os movimentos do robô foram mais suaves do que os apresentados pela primeira abordagem. Isto se deve ao fato de que na segunda abordagem os valores apresentados ao controlador do robô são não binários (Figura 4), diferenciando da primeira abordagem, no qual as saídas do sistema de visão apresentam valores binários (Figura 3).

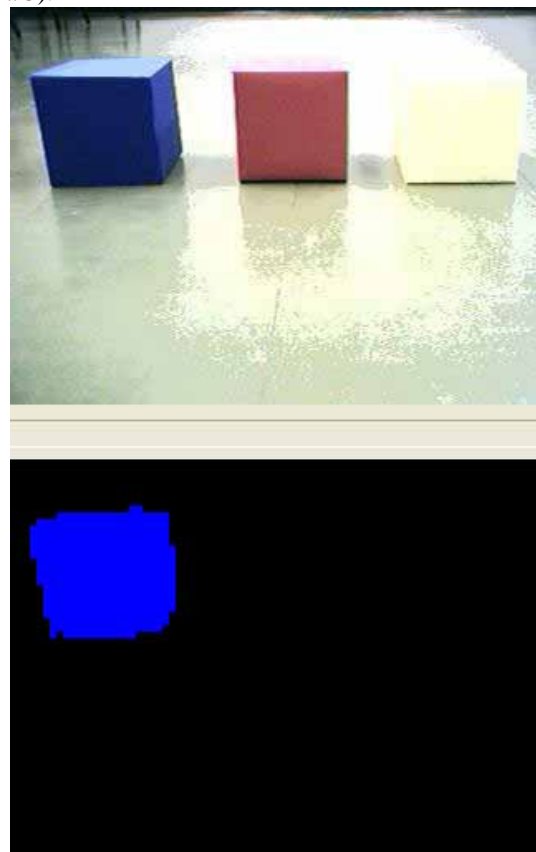


Figura 6. Resultado do processo de segmentação para a cor azul.

IV. CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho foi apresentado um sistema de visão computacional baseado em cores composto por dois módulos: módulo de visão e módulo de controle, no qual o módulo de

visão foi responsável por capturar e realizar todo o processamento dos quadros até a identificação do posicionamento do objeto no campo visual do robô e o módulo de controle foi responsável por coordenar o robô a perseguir a bola pelo chão de uma sala.

Nos testes realizados ficou demonstrado que o sistema desenvolvido comportou-se de forma muito satisfatória, mesmo utilizando uma câmera simples no processo de captura de imagens. Graças ao sistema proposto, o robô Pioneer foi capaz de perseguir objetos de cores determinadas e atuando em tempo real.

O sistema de segmentação utilizando uma rede neural para cada cor se mostrou mais adequado do que o sistema utilizando apenas uma rede neural para todas as cores, pois é possível uma maior generalização da rede neural para o reconhecimento de cada cor em diversas condições de luminosidades e sombras presentes no ambiente.

Apesar do tempo de treinamento da rede neural para o sistema de identificação da posição do objeto ser elevado, depois de treinada, a rede responde de uma forma muito rápida, o que é uma grande vantagem em se tratando de um sistema em tempo real.

Como trabalho futuro, pretende-se acoplar ao sistema um método para reconhecimento de objetos específicos como cubos, bolas, pirâmides, e não apenas realizar o reconhecimento pela cor como foi proposto neste artigo. Além disso, pretende-se modificar o comportamento que trata o caso em que o objeto não se encontra no campo visual do robô, fazendo com que este além de perseguir o objeto, também seja capaz de encontrá-lo. Isto permitirá que o robô responda a comandos do tipo: Encontre e persiga um cubo azul.

V. REFERÊNCIAS

- [1] D. S. Blank e J. O. Ross, "Incorporating a Connectionist Vision Module into a Fuzzy, Behavior-Based Robot Controller," In *Proc. 1997 Midwest Artificial Intelligence and Cognitive Science Society Conference*.
- [2] H. D. Cheng, X. H. Jiang, Y. Sun & J. Wang, "Color Image Segmentation: Advances and Prospects," *Pattern Recognition*, vol. 34, pp. 2259-2281, 2001.
- [3] S. Haykin, *Redes Neurais: Princípios e Prática*, 2ª edição, Porto Alegre: Bookman, 2001, p. 900.
- [4] M. Jonsson, P. A. Wilberg e N. Wickström, "Vision-Based Low-Level Navigation Using a Feed-Forward Neural Network," In *Proc. 1997 International Workshop on Mechatronic Computer Systems for Perception and Action*, Pisa-Italy.
- [5] D. M. R. Medeiros, e R. A. F. Romero, "Sistema de Controle Autônomo de Robôs Móveis com Utilização de Redes Neurais Artificiais," In *Proc. 2002 WORKCOMP – V Workshop de Computação*.
- [6] D. Pomerleau, "Neural Network Vision for Robot Driving". In *The Handbook of Brain Theory and Neural Networks*, 1995.
- [7] A. S. Simões e A. H. R. Costa, "Segmentação de Imagens por Classificação de Cores: Uma Abordagem Neural para a Representação RGB," In *Proc. 2000 Workshop de Computação WORKCOMP*, ITA – São José dos Campos, pp. 25-31.
- [8] A. S. Simões, e A. H. R. Costa, "Using Neural Color Classification in Robotic Soccer Domain," In *Proc. 2000 International Joint Conference IBERAMIA*.
- [9] A. S. Simões, e A. H. R. Costa, "Classificação de Cores por Redes Neurais Artificiais: Um Estudo do Uso de Diferentes Sistemas de Representação de Cores no Futebol de Robôs Móveis Autônomos," In *Proc. 2001 ENIA*.

- [10] L. G. Tyler e C. A. Czarnecki, "A Neural Vision Based Controller for a Robot Footballer," *Image Processing and its Application* vol. 465, pp. 77-81, 1999.
- [11] R. A. F. Romero, S. Waldherr e S. Thrun, "A Gesture-Based Interface for Human-Robot," *Autonomous Robots* vol. 2, pp. 151-173, 2000.
- [12] K. Konolige, *Saphira Manual Version 6.1*, ActivMedia Inc, 1997.
- [13] C. Ribeiro, A. H. R. Costa e R. A. F. Romero, "Robôs Móveis Inteligentes: Princípios e Técnicas," *Capítulo de Livro da I Jornada de Atualização em Inteligência Artificial*, 2001, XXI Congresso da SBC.
- [14] R. C. Arkin, *Behavior Based Robotics*, Cambridge: The MIT Press, 1998, p. 491.

VI. BIOGRAFIAS



computacional.

Marcos Gonçalves Quiles nasceu em Juaram- MT em 24 de março de 1980. Recebeu o título de Bacharel em Ciência da Computação pela Universidade Estadual de Londrina em 2003. Atualmente é aluno do programa de Mestrado em Ciências de Computação e Matemática Computacional junto ao Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo, orientado pela Profa. Dra. Roseli Ap. Francelin Romero. Seus interesses incluem computação bioinspirada (rede neurais artificiais e algoritmos evolutivos), robótica e visão



Roseli Aparecida Francelin Romero é Professora Livre-Docente do Departamento de Ciências de Computação e Estatística do ICMC, Universidade de São Paulo. Doutora pela Faculdade de Engenharia Elétrica da UNICAMP, Campinas-SP, 1993. Realizou um programa de Pós-Doutorado, junto a CMU - Carnegie Mellon University, Pittsburgh-PA, de 1996 a 1998. Ela tem trabalhado com redes neurais, lógica fuzzy, reconhecimento de imagens e algoritmos de aprendizado com aplicações em robôs móveis. É consultora ad hoc-CNPq, desde 1992 e também da FAPESP. Orientou 10 dissertações de mestrado e orientou/co-orientou 4 teses de doutorado. Atualmente, orienta 3 alunos de doutorado e 4 de mestrado. Participou do comitê de programa de diversas conferências nacionais e internacionais. Ela também já participou de 8 projetos de Pesquisa, dentre os quais ela foi coordenadora de 3 projetos.