

PROJETO DE PESQUISA

Processos atencionais e aprendizado de máquina para sistemas robóticos

Aluno: Erik de Godoy Perillo

Orientadora: Profa. Dra. Esther Luna Colombini

Resumo

Entender o ambiente ao seu redor é uma tarefa fundamental para o desafio de se obter máquinas autônomas que interagem com o mundo de forma semelhante à nossa. No entanto, a alta dimensionalidade dos dados captados por sensores usados para este fim é em geral problemática, muitas vezes havendo redundância e irrelevância de informação. Nos seres humanos este filtro sensorial é realizado pela Atenção. Neste contexto, este projeto propõe a aplicação de técnicas de aprendizado de máquina sobre informações sensoriais previamente processadas por processos atencionais em uma tarefa de navegação autônoma. Adicionalmente, planeja-se implementar uma estrutura que permita o uso das técnicas propostas para projetos robóticos em geral que utilizem GPUs embarcadas.

Universidade Estadual de Campinas

24 de abril de 2016

1 Introdução

Máquinas capazes de realizar tarefas complexas, perigosas ou maçantes são objeto de interesse e desejo do ser humano há tempos. Robôs industriais já dominam os mais variados setores de produção. Não apenas parados em linhas de montagem, sistemas atuais podem locomover-se em regiões uniformes e previsíveis – como galpões de estoques – trabalhando vinte e quatro horas por dia [16]. Um desafio ainda em aberto, entretanto, é a concepção de robôs que lidam com o imprevisível, reagindo de forma apropriada às mais diversas situações e ambientes do mundo real.

Sistemas que interagem com o ambiente, objetos e pessoas com uma variedade de maneiras semelhante à nossa têm o potencial de ser benéficos para a sociedade. Entretanto, diversos são os desafios associados ao processo de concepção de máquinas gerais desse tipo. Um deles, certamente, é a compreensão do ambiente. Entender o que lhe cerca é fundamental para uma interação complexa com o resto do mundo.

Ter ciência de sua relação com outros objetos e as possíveis interações com cada um deles é essencial para a navegação de robôs por casas, ruas e até mesmo locais onde a presença de humanos seria impossível ou indesejável. Essa é uma das tarefas principais que máquinas como carros autônomos, robôs de serviço [6] e de resgate devem cumprir bem. Para tal, os robôs são dotados de sensores como Técnicas que utilizam sensores como câmeras [11] e LIDAR [9] têm sido usadas com sucesso para identificação e classificação de entidades em um ambiente com o qual o robô interage.

1.1 Motivação

Um problema de sensores usados para tarefas mais complexas de navegação em geral – como câmera, LIDAR e outros sensores – é que o volume de dados a ser processado podem ser demasiadamente grande. Para um robô que interage continuamente com o ambiente, é improvável que em todos os instantes toda informação proveniente de seus sensores seja processada ou mesmo necessária. Lidar com todo o conjunto de dados para depois determinar qual parte é relevante pode ser uma tarefa computacionalmente custosa. Isso é especialmente crítico para sistemas robóticos que precisam de respostas rápidas para interagir com o ambiente em que estão inseridos.

Além disso, o *hardware* utilizado em robôs autônomos tem sido, por muitos anos, limitado às restrições geradas pela necessidade de embarcá-lo em sistemas compactos. Entretanto, avanços recentes têm mudado isso. Além do aumento da eficiência e diminuição do tamanho de CPUs impulsionado

pelo mercado *mobile*, esforços têm sido feitos para a concepção de computadores com processadores gráficos (GPUs) embarcados, que permitem o tratamento eficiente de dados naturalmente característicos dos sensores empregados. A *NVIDIA* lançou recentemente os modelos *Jetson TK1* e *Jetson TX1* [4], que são sistemas computacionais completos com GPUs poderosas e energeticamente eficientes, desenvolvidos especialmente para serem embarcados em projetos móveis. O uso de GPUs em robótica abre a possibilidade do uso de técnicas mais sofisticadas e computacionalmente caras que até então não podiam ser exploradas. Sua natureza paralela permite um grande benefício de operações que podem ser divididas em subproblemas e realizados ao mesmo tempo. Esse é o caso de diversas tarefas realizadas no processamento de sinais e por técnicas de aprendizado de máquina [1], por exemplo.

1.2 Processos atencionais e aprendizado de máquina

O fato de sensores darem uma quantidade muito grande de informação que nem sempre é necessária motiva uma abordagem inspirada no ser humano. Não damos foco a todas as conversas que escutamos ou em tudo que vemos, mas sim apenas na parte que nos convém ou cuja natureza é extremamente atrativa. Nos seres humanos este filtro sensorial é realizado pela Atenção [3]. Técnicas que fazem uso de processos atencionais têm sido usadas para a segmentação de áreas de relevância na percepção e entendimento do ambiente com sucesso [15] [2]. Com essas abordagens, pode-se filtrar dados irrelevantes ou redundantes, dando prioridade e foco a elementos relevantes do ambiente, aumentando consideravelmente a eficiência computacional dessas técnicas.

Porém, não só o foco é necessário. Em ambientes dos mais variados é preciso, além de identificar, classificar entidades a fim de interagir com elas de forma apropriada. O aprendizado de máquina tem sido uma área de grande sucesso nessa tarefa. Diversas técnicas podem ser utilizadas, como *SVMs* [10] e redes neurais [8] a fim de se classificar entidades em geral com alta taxa de acertos.

Neste sentido, este projeto busca uma associação dessas técnicas: o uso dos processos atencionais em um estágio inicial de processamento pode atuar como um filtro na informação, passando adiante o que é relevante para o estágio de classificação, a ser suportado por técnicas de aprendizado de máquina. Espera-se que tal abordagem possibilite uma percepção eficiente para sistemas autônomos o que permitiria a realização de tarefas mais sofisticadas, trazendo benefícios objetivos para a área da percepção em robótica.

2 Objetivos

O objetivo geral deste trabalho envolve a aplicação de técnicas de aprendizado de máquina sobre informações sensoriais previamente processadas pelo modelo atencional proposto por em [2] em uma tarefa de navegação autônoma.

Mais especificamente, objetivamos:

- Obter técnicas que permitam o entendimento de ambientes diversos para a navegação de sistemas robóticos autônomos por meio do uso de processos atencionais e aprendizado de máquina;
- Avaliar a efetividade e eficiência da abordagem proposta comparado a outras técnicas utilizadas atualmente;
- Construir uma plataforma para uso em GPUs embarcadas com ferramentas que permitam o uso das técnicas desenvolvidas para aplicações variadas de robótica.

3 Materiais e Métodos

Para atendimento dos objetivos propostos, serão realizadas as seguintes etapas:

- FASE 1: Revisão bibliográfica. Nesta etapa, será realizada a revisão bibliográfica: a) do modelo atencional [2], buscando quais adequações serão necessárias para o projeto em questão, e b) das técnicas de Aprendizado de Máquina mais apropriadas para a tarefa;
- FASE 2: Preparação do ambiente computacional. Nesta fase, o ambiente para construção do sistema será preparado (configuração do ambiente, instalação de bibliotecas e softwares, instalação de pacotes gráficos, configurações, etc);
- FASE 3: Modelagem e implementação de processos atencionais. Nessa etapa será proposta a base do sistema atencional e testes preliminares serão realizados;
- FASE 4: Modelagem e implementação da técnica de aprendizado de máquina. Nessa fase a técnica escolhida na FASE 1 será implementada e testes preliminares serão realizados;

- FASE 5: Integração dos sistemas. Nessa fase será realizada a integração do sistema atencional à técnica de aprendizado de máquina;
- FASE 6: Testes comparativos. Nessa fase serão realizados testes e ensaios com o agente robótico executando o sistema integrado;
- FASE 7: Construção de plataforma para GPU. Nesta etapa o sistema proposto será adequado para ser executado na GPU;
- FASE 8: Relatório final. Redação do relatório final.

3.1 Ambiente

O desafio *Trekking* [12] é uma categoria da série de competições de robótica realizada pela *Robocore* [13]. Nesta categoria, o robô deve completar um percurso em um campo de futebol passando por diversas bases em uma ordem pré-definida. Em uma certa região do campo, há obstáculos dos quais o robô deve desviar (Figura 1). O percurso deve ser feito no menor tempo possível.



Figura 1: a) O ambiente de locomoção do robô na categoria Trekking b) Robô utilizado na competição.

A *Equipe Phoenix de Robótica da Unicamp* [5] é formada por estudantes de graduação de diversos cursos de Engenharia e Computação da Universidade Estadual de Campinas. A equipe desenvolve atualmente um robô para participar na categoria *Trekking*, apresentado na Figura 1 b). O robô consiste em um carro elétrico com tração nas quatro rodas e locomoção por motores *brushless*. O motor é controlado por meio de um conjunto da *NXP* composto da placa *FRDM-K64F* e da placa *FRDM-STBC-AGM01* [14], que contém diversos sensores inerciais como magnetômetro, giroscópio e acelerômetro. Há também um kit de desenvolvimento com GPU embarcada *NVIDIA Jetson TX1*.

A *Jetson TX1* possui um processador *ARM Quad-Core Cortex-A57*, 4GB de memória *RAM* e uma GPU *NVIDIA Maxwell* com 256 núcleos de processamento [7]. Seu sistema operacional é o Ubuntu com *kernel Linux*. Duas câmeras *Logitech C270* são os sensores principais usados pela *Jetson TX1* para a ambientação do robô no campo.

A pesquisa usará este robô como caso de uso para avaliação da abordagem a ser desenvolvida, com o uso da *Jetson TX1* como plataforma para implementação das técnicas. O robô usa a câmera como principal sensor, podendo contar com o apoio de sensores auxiliares.

Na tarefa de navegação proposta o robô deverá reconhecer o terreno no qual pode mover-se, os obstáculos do qual deve desviar e bases nas quais deve chegar.

3.2 Avaliação da abordagem proposta

Para efeitos de avaliação, serão comparadas os cenários onde:

1. Uma abordagem específica é utilizada. Um simples *threshold* de cor pode ser capaz de identificar objetos imagem e classificá-los dado que diferentes classes de objetos têm diferentes cores.
2. Aprendizado de máquina puro para reconhecimento de objetos. Nessa abordagem, será dado foco igual a toda a informação. Assim, todo o conjunto deverá ser analisado pelos algoritmos de classificação.
3. Processos atencionais são utilizados em um estágio anterior ao de classificação com aprendizado de máquina, com a intenção de selecionar subconjuntos de toda informação dada pelo sensor que tenham maior relevância para serem analisados no estágio de classificação.

Com essa comparação, espera-se avaliar cada técnica pelos seguintes critérios:

- Desempenho: o custo computacional da técnica.
- Eficiência: quão corretamente a técnica identifica e classifica as entidades de interesse.
- Generalidade: possibilidade da técnica ser utilizadas em outros cenários.

3.3 Construção de plataforma para navegação usando GPUs

Com as técnicas obtidas pelo projeto, espera-se construir uma plataforma de uso geral para robôs autônomos que usem GPUs embarcadas. *frameworks* de aprendizado de máquina serão estudados para se decidir quais usar. Toda biblioteca será escolhida de forma que se possa utilizar o processamento da GPU. O sistema será construído em *C++* e *Python*.

4 Cronograma

Tarefa/mês	Jul	Ago	Set	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun
FASE 1	x	x										
FASE 2	x	x	x									
FASE 3		x	x	x	x							
FASE 4		x	x	x	x	x						
FASE 5					x	x	x	x	x			
FASE 6							x	x	x			
FASE 7										x	x	
FASE 8											x	x

Referências

- [1] Dan C. Cireşan et al. *Flexible, High Performance Convolutional Neural Networks for Image Classification*. 2011.
- [2] E. L. Colombini, A. da Silva Simões e C. H. Costa Ribeiro. “An Attentional Model for Autonomous Mobile Robots”. Em: *IEEE Systems Journal* PP.99 (2016), pp. 1–12. ISSN: 1932-8184.
- [3] Paul E Dux e R Marois. “The attentional blink: a review of data and theory.” Em: *Attention Perception Psychophysics* 71.8 (2009), pp. 1683–1700.
- [4] *Embedded Systems*. 2016. URL: <http://www.nvidia.com/object/embedded-systems.html> (acesso em 19/04/2016).
- [5] *Equipe Phoenix de Robótica da Unicamp*. 2016. URL: <http://phoenixunicamp.com.br> (acesso em 19/04/2016).

- [6] *International Federation of Robotics (IFR)*. 2016. URL: www.worldrobotics.org (acesso em 19/04/2016).
- [7] *Jetson TX1 Development Kit*. 2016. URL: <https://developer.nvidia.com/embedded/buy/jetson-tx1-devkit> (acesso em 19/04/2016).
- [8] Alex Krizhevsky, Ilya Sutskever e Geoffrey E. Hinton. “ImageNet Classification with Deep Convolutional Neural Networks”. Em: *Advances in Neural Information Processing Systems 25*. Ed. por F. Pereira et al. Curran Associates, Inc., 2012, pp. 1097–1105. URL: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.
- [9] J. Levinson et al. “Towards fully autonomous driving: Systems and algorithms”. Em: *Intelligent Vehicles Symposium (IV), 2011 IEEE*. Jun. de 2011, pp. 163–168.
- [10] Larry M. Manevitz e Malik Yousef. “One-class Svms for Document Classification”. Em: *J. Mach. Learn. Res.* 2 (mar. de 2002), pp. 139–154. ISSN: 1532-4435.
- [11] N. P. Papanikolopoulos, P. K. Khosla e T. Kanade. “Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision”. Em: *IEEE Transactions on Robotics and Automation* 9.1 (fev. de 1993), pp. 14–35. ISSN: 1042-296X.
- [12] *Regras robô Trekking*. 2016. URL: https://www.robocore.net/upload/attachments/robocore__regras_robo_trekking_610.pdf (acesso em 19/04/2016).
- [13] *Robocore*. 2016. URL: https://www.robocore.net/modules.php?name=GR_Eventos (acesso em 19/04/2016).
- [14] *Sensor Toolbox Development Platform*. 2016. URL: <http://goo.gl/kIwBFn> (acesso em 19/04/2016).
- [15] C. Siagian e L. Itti. “Rapid Biologically-Inspired Scene Classification Using Features Shared with Visual Attention”. Em: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.2 (fev. de 2007), pp. 300–312. ISSN: 0162-8828.
- [16] Peter R. Wurman, Raffaello D’Andrea e Mick Mountz. “Coordinating Hundreds of Cooperative, Autonomous Vehicles in Warehouses”. Em: *Proceedings of the 19th National Conference on Innovative Applications of Artificial Intelligence - Volume 2*. IAAI’07. Vancouver, British Columbia, Canada: AAAI Press, 2007, pp. 1752–1759. ISBN: 9781577353232.